

Vorlesung

Mathematik für Studierende der Physik II

Vorlesungsskript SoSe 2020

Dr. Nils Ackermann*

Johann Wolfgang Goethe-Universität, Frankfurt

Stand: 15. Februar 2022

*Basierend auf Skripten von Prof. Dr. Tobias Weth und Dr. Sven Jarohs

Inhaltsverzeichnis

1	Das Riemann-Integral	4
2	Konvergenz von Funktionenfolgen und Funktionenreihen	19
2.1	Fourierreihen	22
3	Lineare Algebra II	30
3.1	Die Determinante einer Matrix	30
3.2	Eigenwerte und Eigenvektoren	40
3.2.1	Diagonalisierung von Matrizen	42
3.2.2	Normalform von Matrizen	44
3.3	Spezielle Matrizen	50
3.3.1	Orthogonale und unitäre Matrizen	52
3.3.2	Selbstadjungierte Matrizen	54
3.3.3	Positiv definite Matrizen	56
3.3.4	Die reelle Jordansche Normalform	58
4	Die Topologie des \mathbb{R}^N	60
4.1	Metrische Räume	60
4.2	Äquivalente Normen	66
4.3	Stetige Abbildungen	67
4.4	Lineare stetige Abbildungen	70
4.5	Kompaktheit	73
5	Differenzierbarkeit	76
5.1	Richtungsableitungen und partielle Ableitungen	76
5.2	Totale Differenzierbarkeit	78
5.3	Höhere Ableitungen	85
5.4	Einblick in die Vektoranalysis	90
5.5	Lokale Extrema	93
6	Homogene Lineare Gewöhnliche Differentialgleichungen mit konstanten Koeffizienten	97
6.1	Die Exponentialfunktion einer Matrix	98
6.2	Fundamentalsysteme	100
6.3	Skalare Differentialgleichungen höherer Ordnung	104
6.3.1	Allgemeine Form	104
6.3.2	Gleichungen zweiter Ordnung	106

7	Implizite Funktionen	107
7.1	Vollständigkeit, Fixpunkte und Nullstellen	107
7.2	Auflösung von Gleichungen durch implizite Funktionen	110
7.3	Reguläre Punkte und reguläre Werte	116
7.4	Lokale Diffeomorphie	119
7.5	Extremwertbestimmung unter Nebenbedingungen und Untermannigfaltigkeiten	122

1 Das Riemann-Integral

Im Folgenden seien $a, b \in \mathbb{R}$, $a < b$ fest gewählt und $I := [a, b] \subseteq \mathbb{R}$.

Bemerkung 1.1. Eine exemplarische Ausgangsfrage der Integrationstheorie: Für welche Funktionen $f: I \rightarrow \mathbb{R}$ kann man die Fläche zwischen dem Graphen von f und der x -Achse sinnvoll definieren und berechnen? Dabei soll die Fläche über der x -Achse mit positiven Vorzeichen und die Fläche unter der x -Achse mit negativen Vorzeichen eingehen. Naive Idee hierzu: Man zerlege I in kleine Teilintervalle und approximiere die gesuchte Fläche durch Rechtecke über diesen Teilintervallen. Die Formalisierung dieser Idee, welche auf Darboux zurückgeht, führt auf das *Riemann-Integral*.

Wir benötigen zunächst einige Definitionen. Ferner vereinbaren wir folgende praktische Schreibweise: Ist M eine Menge und $f: M \rightarrow \mathbb{R}$ eine Funktion, so schreiben wir $\sup_{x \in M} f(x)$ oder kurz $\sup_M f$ anstelle von

$$\sup f(M) = \sup\{f(x) \mid x \in M\}$$

und analog für \inf , \max und \min .

Definition 1.2. Seien $D \subseteq \mathbb{C}$ und $f: D \rightarrow \mathbb{C}$ eine Funktion.

- (a) f heißt *beschränkt*, falls $f(D)$ eine beschränkte Teilmenge von \mathbb{C} ist.
- (b) Ist f beschränkt, so setzt man

$$\|f\|_\infty := \sup_D |f|$$

und nennt diese Zahl die *Supremumsnorm* oder die ∞ -Norm von f .

Definition 1.3. (a) Eine endliche Menge $Z \subseteq I$ mit $a, b \in Z$ heißt *Zerlegung von I* . Man schreibt $Z = \{x_0, \dots, x_n\}$ mit $a = x_0 < x_1 < x_2 < \dots < x_n = b$. Die Punkte x_i , $i = 0, \dots, n$, heißen *Teilpunkte von Z* und die Zahl

$$\delta(Z) := \max_{k=1, \dots, n} (x_k - x_{k-1})$$

heißt *Feinheit von Z* .

- (b) Sind Z und Z' Zerlegungen von I mit $Z \subseteq Z'$, so heißt Z' *Verfeinerung von Z* . Offensichtlich gilt dann $\delta(Z') \leq \delta(Z)$.

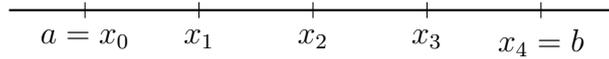


Abbildung 1.1: Äquidistante Zerlegung

Beispiel 1.4 (Äquidistante Zerlegung). Seien $n \in \mathbb{N}$ und

$$Z_n := \left\{ a + \frac{k}{n}(b-a) \mid k = 0, \dots, n \right\}.$$

Dann zerlegt Z_n das Intervall $I = [a, b]$ in n gleichgroße Intervalle, siehe Abb. 1.1. Es gilt:

$$Z_m \supseteq Z_n \quad \Leftrightarrow \quad n \text{ teilt } m.$$

Ferner gilt

$$\delta(Z_n) = \frac{b-a}{n} \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Insbesondere kann ein Intervall I mit beliebig kleiner Feinheit zerlegt werden.

Definition 1.5 (Riemann-Integral). Seien $I = [a, b]$ und $f: I \rightarrow \mathbb{R}$ beschränkt.

- (a) Sei $Z = \{x_0, x_1, \dots, x_n\}$ eine Zerlegung von I . Ferner seien *Stützstellen* $S := \{\xi_1, \dots, \xi_n\}$ mit $\xi_k \in [x_{k-1}, x_k]$, $k = 1, 2, \dots, n$, gegeben. Die Zahl

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

heißt dann *Riemann-Summe von f bezüglich Z und S* , siehe Abb. 1.2.

- (b) Die Funktion f heißt *Riemann-integrierbar (auf I)*, falls es $c \in \mathbb{R}$ mit der folgenden Eigenschaft gibt: Für jedes $\varepsilon > 0$ existiert ein $\delta > 0$, so dass für jede Zerlegung $Z := \{x_0, x_1, \dots, x_n\}$ von I mit $\delta(Z) < \delta$ und jede Wahl von Stützstellen $S := \{\xi_1, \dots, \xi_n\}$ zu Z gilt:

$$(1.1) \quad \left| c - \sum_{k=1}^n f(\xi_k)(x_k - x_{k-1}) \right| \leq \varepsilon.$$

Salopp gesagt gilt also $\lim_{\delta(Z) \rightarrow 0} \sum_{k=1}^n f(\xi_k)(x_k - x_{k-1}) = c$, unabhängig von der Wahl der Stützstellen. Die Menge der Riemann-integrierbaren Funktionen auf I bezeichnen wir mit $\mathcal{R}(I)$. Die Zahl c heißt *Riemann-Integral von f über I* und wird wie folgt bezeichnet:

$$\int_a^b f \quad \text{oder} \quad \int_a^b f(x) \, dx \quad \text{oder} \quad \int_I f \quad \text{oder} \quad \int_I f(x) \, dx.$$

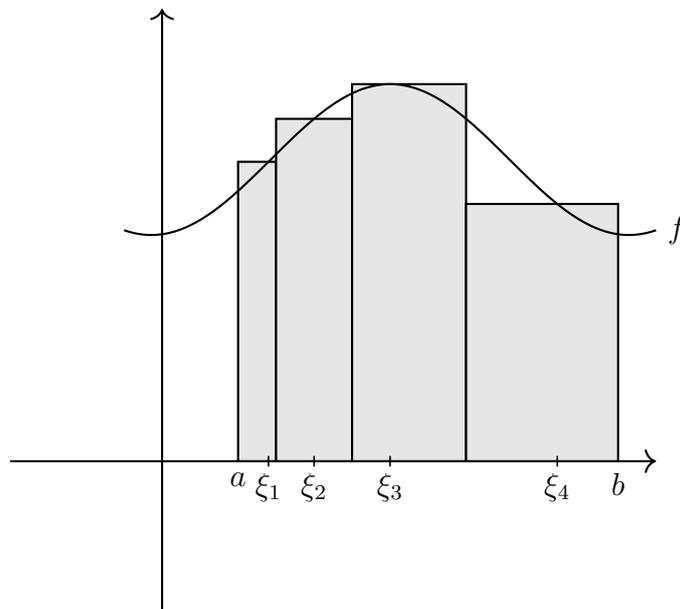


Abbildung 1.2: Riemann-Summe

Bemerkung 1.6. Gilt (1.1) für gewisse c, ε, Z und mit beliebigen Stützstellen zu Z , dann folgt sofort, dass f beschränkt ist. Darum ist die Definition von Riemann-Integrierbarkeit nur für beschränkte Funktionen sinnvoll.

Beispiel 1.7. (a) Seien $\lambda \in \mathbb{R}$ und $f: I \rightarrow \mathbb{R}$ definiert durch $f(x) := \lambda$ für $x \in I$, d.h. f ist konstant. Dann gilt

$$\sum_{k=1}^n f(\xi_k)(x_k - x_{k-1}) = \sum_{k=1}^n \lambda(x_k - x_{k-1}) = \lambda(b - a)$$

für jede Zerlegung $Z = \{x_0, \dots, x_n\}$ von I und jede Wahl von Stützstellen $\{\xi_1, \dots, \xi_n\}$ zu Z . Demnach ist $f \in \mathcal{R}(I)$ und $\int_a^b f = \lambda(b - a)$.

(b) Seien speziell $I := [0, 1]$ und

$$f: [0, 1] \rightarrow \mathbb{R}, \quad f(x) := \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases} \quad (\text{Dirichletfunktion}).$$

Seien ferner $Z = \{x_0, \dots, x_n\}$ eine beliebige Zerlegung von I sowie

$$S_1 := \{\xi_1, \dots, \xi_n\} \subseteq \mathbb{Q} \quad \text{und} \quad S_2 := \{\zeta_1, \dots, \zeta_n\} \subseteq \mathbb{R} \setminus \mathbb{Q}$$

zwei Stützstellenmengen zu Z (diese existieren immer!). Dann gilt

$$\sum_{k=1}^n f(\xi_k)(x_k - x_{k-1}) = 1 \quad \text{und} \quad \sum_{k=1}^n f(\zeta_k)(x_k - x_{k-1}) = 0.$$

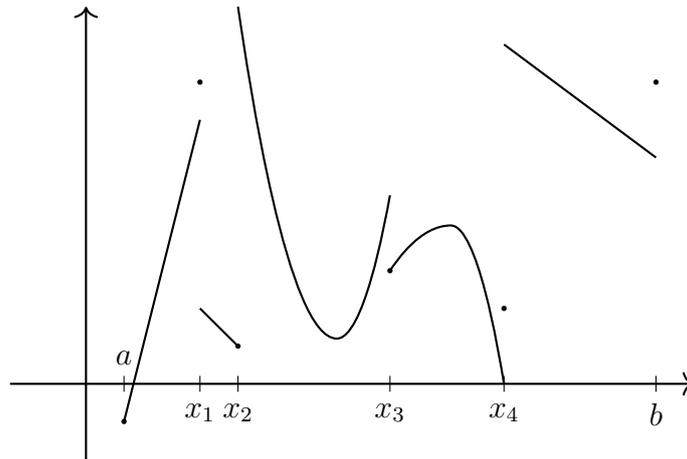


Abbildung 1.3: Stückweise stetige Funktion

Somit ist f nicht Riemann-integrierbar.

Satz 1.8. Seien $f, g \in \mathcal{R}(I)$ und $\alpha \in \mathbb{R}$. Dann gilt:

- (a) $\alpha f \in \mathcal{R}(I)$ und $\int(\alpha f) = \alpha \int f$.
- (b) $f + g \in \mathcal{R}(I)$ und $\int(f + g) = \int f + \int g$.
- (c) Ist $f \leq g$ in I , d.h. $f(x) \leq g(x)$ für alle $x \in I$, so folgt $\int f \leq \int g$.
- (d) $|f| \in \mathcal{R}(I)$ und $|\int f| \leq \int |f|$ (Standardabschätzung).
- (e) $f \cdot g \in \mathcal{R}(I)$.

Beweis. Siehe z.B. [19, Abschnitt 79.3, S. 454]. □

Bemerkung 1.9. Aus den Eigenschaften (a)–(c) von Satz 1.8 folgt, dass die Menge $\mathcal{R}(I)$ ein \mathbb{R} -Vektorraum und das Riemann-Integral ein monotonen lineares *Funktional* auf $\mathcal{R}(I)$ ist (also eine Abbildung $\mathcal{R}(I) \rightarrow \mathbb{R}$).

Korollar 1.10. Ist $f \in \mathcal{R}(I)$, so gilt $|\int f| \leq \|f\|_\infty(b - a)$.

Beweis. Sei $g: I \rightarrow \mathbb{R}$ definiert durch $g(x) := \|f\|_\infty$ für $x \in I$. Dann ist $g \in \mathcal{R}(I)$ (nach Beispiel 1.7(a)) und $|f| \leq g$, also

$$\left| \int f \right| \stackrel{\text{Satz 1.8(d)}}{\leq} \int |f| \stackrel{\text{Satz 1.8(c)}}{\leq} \int g \stackrel{\text{Beispiel 1.7(a)}}{=} \|f\|_\infty(b - a). \quad \square$$

Definition 1.11. Seien $a, b \in \mathbb{R}$, $a < b$ und $I := [a, b]$. Eine Funktion $f: I \rightarrow \mathbb{R}$ heißt *stückweise stetig*, wenn sie höchstens in endlich vielen Punkten unstetig ist (siehe Abb. 1.3).

Satz 1.12. Seien $f, g: I \rightarrow \mathbb{R}$ Funktionen.

- (a) Ist f beschränkt und stückweise stetig, so gilt $f \in \mathcal{R}(I)$.
- (b) Ist $f \in \mathcal{R}(I)$ und unterscheiden sich f und g nur in endlich vielen Punkten, dann gilt auch $g \in \mathcal{R}(I)$ und $\int_a^b f = \int_a^b g$

Beweis. Siehe z.B. [19, Sätze 84.3 und 84.4]. □

Beispiel 1.13. Seien speziell $I := [0, 1]$ und $f: I \rightarrow \mathbb{R}$ definiert durch $f(x) := x^3$. Nach Satz 1.12 ist $f \in \mathcal{R}(I)$. Seien

$$Z_n := \left\{ x_k^n := \frac{k}{n} \mid k = 0, \dots, n \right\}$$

Zerlegungen mit Stützstellenmengen

$$S_n := \left\{ \xi_k^n := \frac{k}{n} \mid k = 1, \dots, n \right\},$$

für $n \in \mathbb{N}$. Es folgt:

$$\sum_{k=1}^n f(\xi_k^n)(x_k^n - x_{k-1}^n) = \sum_{k=1}^n \frac{k^3}{n^3} \cdot \frac{1}{n} = \frac{1}{n^4} \sum_{k=1}^n k^3 \stackrel{\text{Übung}}{=} \frac{1}{n^4} \cdot \frac{n^2(n+1)^2}{4} = \frac{(1 + \frac{1}{n})^2}{4}$$

und daher

$$\int_0^1 f = \lim_{n \rightarrow \infty} \sum_{k=1}^n f(\xi_k^n)(x_k^n - x_{k-1}^n) = \frac{1}{4}.$$

Im Folgenden schreiben wir für $a, b \in \mathbb{R}$, $a < b$ stets $\mathcal{R}(a, b)$ statt $\mathcal{R}([a, b])$ und für $f \in \mathcal{R}(a, b)$ setzen wir

$$\int_b^a f := - \int_a^b f \quad \text{und} \quad \int_c^c f := 0 \quad \text{für } c \in [a, b].$$

Satz 1.14 (Intervalladditivität). Seien $a, b, c \in \mathbb{R}$ mit $a < b < c$ und $f: [a, c] \rightarrow \mathbb{R}$ beschränkt. Dann gilt:

$$f \in \mathcal{R}(a, c) \quad \Leftrightarrow \quad f|_{[a, b]} \in \mathcal{R}(a, b) \quad \text{und} \quad f|_{[b, c]} \in \mathcal{R}(b, c),$$

und in diesem Fall ist $\int_a^c f = \int_a^b f + \int_b^c f$.

Beweis. Siehe z.B. [19, Sätze 84.5 und 84.6]. □

Satz 1.15 (Mittelwertsatz der Integralrechnung). Seien $a, b \in \mathbb{R}$, $a < b$, $I := [a, b]$ sowie $f: I \rightarrow \mathbb{R}$ stetig und $g \in \mathcal{R}(I)$ mit $g \geq 0$ in I . Dann existiert $\xi \in [a, b]$ mit

$$\int_a^b fg = f(\xi) \int_a^b g.$$

Insbesondere folgt für $g \equiv 1$

$$\int_a^b f = f(\xi)(b - a).$$

Beweis. Da f stetig ist, existiert $m := \min_I f$ und $M := \max_I f$. Da $g \geq 0$ in I gilt, ist

$$mg(t) \leq f(t)g(t) \leq Mg(t) \text{ für } t \in I,$$

also nach Satz 1.8(c) und (e):

$$m \int_a^b g \leq \int_a^b fg \leq M \int_a^b g.$$

Somit existiert $c \in [m, M]$ mit $c \int_a^b g = \int_a^b fg$. Da f stetig ist, gibt es nach dem Zwischenwertsatz ein $\xi \in I$ mit $c = f(\xi)$. Die Behauptung folgt. \square

_____ Ende der Sitzung vom 20.4. _____

Im Folgenden sei stets I ein beliebiges nicht entartetes Intervall in \mathbb{R} , d.h. I besteht aus mehr als einem Punkt.

Definition 1.16. Sei $f: I \rightarrow \mathbb{R}$ eine Funktion. Eine differenzierbare Funktion $F: I \rightarrow \mathbb{R}$ heißt *Stammfunktion* von f , wenn $F' = f$ gilt.

Bemerkung 1.17. Sind $F, G: I \rightarrow \mathbb{R}$ Stammfunktionen einer Funktion $f: I \rightarrow \mathbb{R}$, so ist $F - G$ konstant auf I : Es gilt nämlich $(F - G)' = F' - G' = f - f = 0$. Demnach ist die Funktion $F - G$ sowohl monoton wachsend als auch fallend, also konstant. Alle Stammfunktionen von f haben also die Form $F + C$ mit einer Konstante $C \in \mathbb{R}$.

Satz 1.18 (Hauptsatz der Differential- und Integralrechnung). *Seien $f: I \rightarrow \mathbb{R}$ stetig, $a \in I$ und $F: I \rightarrow \mathbb{R}$ definiert durch*

$$F(x) := \int_a^x f \quad \text{für } x \in I.$$

Dann gilt:

- (a) F ist eine Stammfunktion von f .
- (b) Für $b, c \in I$ und jede Stammfunktion $G := I \rightarrow \mathbb{R}$ von f ist

$$\int_b^c f := G(c) - G(b).$$

Man schreibt statt $G(c) - G(b)$ häufig auch

$$G(x) \Big|_{x=b}^{x=c} \quad \text{oder} \quad G(x) \Big|_b^c \quad \text{oder} \quad G \Big|_b^c.$$

Beweis. Zu (a): Seien $x \in I$ und $(x_n)_n$ eine Folge in $I \setminus \{x\}$ mit $\lim_{n \rightarrow \infty} x_n = x$. Nach Satz 1.15 existiert für jedes $n \in \mathbb{N}$ ein ξ_n zwischen x und x_n derart, dass

$$F(x_n) - F(x) = \int_a^{x_n} f - \int_a^x f \stackrel{\text{Satz 1.14}}{=} \int_x^{x_n} f = f(\xi_n)(x_n - x) \quad \text{für } n \in \mathbb{N}.$$

Da $\lim_{n \rightarrow \infty} \xi_n = x$ gilt und f in x stetig ist, folgt

$$\lim_{n \rightarrow \infty} \frac{F(x_n) - F(x)}{x_n - x} = \lim_{n \rightarrow \infty} f(\xi_n) = f(x).$$

Somit ist F in x differenzierbar mit $F'(x) = f(x)$.

Zu (b): Es gilt

$$\int_b^c f \stackrel{\text{Satz 1.14}}{=} \int_a^c f - \int_a^b f = F(c) - F(b) \stackrel{\text{(a) und Bemerkung 1.17}}{=} G(c) - G(b). \quad \square$$

Definition 1.19. Sei $f: I \rightarrow \mathbb{R}$ stetig. Mit $\int f$ bezeichnen wir eine unspezifische Stammfunktion von f , das *unbestimmte Integral* von f . Wenn F eine Stammfunktion von f ist, dann existieren folgende Schreibweisen:

$$F = \int f, \quad F(u) = \int^u f = \int f(x) dx \quad \text{oder} \quad F(x) = \int f(x) dx.$$

Mischformen davon sind **nicht** erlaubt.

Bemerkung 1.20. Nach Satz 1.18(b) dienen die Unbestimmten Integrale der Berechnung von *bestimmten Integralen*, d.h. Integralen mit festgelegten Grenzen. Bei der Verwendung unbestimmter Integrale kann man aus $\int f = F$ und $\int f = G$ nicht schließen, dass $F = G$ gilt. Man kann lediglich $F - G \equiv \text{konst.}$ folgern. Die Schreibweise $F = \int f$ ist äquivalent zu $F' = f$. Das unbestimmte Integral einer Funktion zu bilden ist also in diesem Sinne die Umkehrung der Differentiation.

Bemerkung und Beispiel 1.21. Seien $a, b \in \mathbb{R}$, $a < b$.

(a) Es gilt

$$\int_a^b x^s dx = \frac{x^{s+1}}{s+1} \Big|_a^b, \quad \text{falls} \quad \begin{cases} s \in \mathbb{N}_0, & \text{oder} \\ s \in \mathbb{Z}, s \leq -2 & \text{und } 0 \notin [a, b], \text{ oder} \\ s \in \mathbb{R} \setminus \{-1\} & \text{und } a > 0. \end{cases}$$

Ferner gilt

$$\int_a^b \frac{1}{x} dx = \log|x| \Big|_a^b \quad \text{falls } 0 \notin [a, b],$$

denn $\frac{d}{dx} \log|x| = \frac{1}{x}$ für $x \neq 0$.

(b) Da $\arctan'(x) = \frac{1}{1+x^2}$ für $x \in \mathbb{R}$ ist, gilt

$$\int_a^b \frac{1}{1+x^2} dx = \arctan(x) \Big|_a^b.$$

(c) Es ist $\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$ für $x \in (-1, 1)$, also

$$\int_a^b \frac{1}{\sqrt{1-x^2}} dx = \arcsin x \Big|_a^b, \quad \text{falls } a, b \in (-1, 1).$$

(d) Ist $f \in \mathcal{R}([a, b]) \setminus C([a, b])$, so ist im Allgemeinen $F(x) = \int_a^x f$ für $f \in \mathcal{R}(I)$ keine Stammfunktion von f . Es lässt sich jedoch zeigen, dass F immerhin noch Lipschitzstetig auf $[a, b]$ ist (Übung!).

Satz 1.22 (Partielle Integration). *Seien $a < b$ und $f, g \in C^1([a, b])$. Dann gilt:*

$$\int_a^b f'g = fg \Big|_a^b - \int_a^b fg' \quad \text{bzw.} \quad \int f'g = fg - \int fg'.$$

Beweis. Nach Voraussetzung ist $(fg)' = f'g + fg': I \rightarrow \mathbb{R}$ stetig mit Stammfunktion $fg: I \rightarrow \mathbb{R}$. Die Behauptung folgt also aus Satz 1.18. \square

Beispiel 1.23. Es gilt

$$\int \log x dx = \int \underbrace{1}_{f'} \cdot \underbrace{\log(x)}_g dx = x \log x - \int x \cdot \frac{1}{x} dx = x \log x - x.$$

Satz 1.24 (Substitutionsregel). *Seien $a, b \in \mathbb{R}$, $a < b$, $\varphi: [a, b] \rightarrow I$ stetig differenzierbar sowie $f: I \rightarrow \mathbb{R}$ stetig. Dann gilt*

$$\int_{\varphi(a)}^{\varphi(b)} f = \int_a^b (f \circ \varphi) \varphi'.$$

Beweis. Sei $F: I \rightarrow \mathbb{R}$ eine Stammfunktion zu f (diese existiert nach Satz 1.18(a)). Dann ist $F \circ \varphi: [a, b] \rightarrow \mathbb{R}$ nach Kettenregel eine Stammfunktion der stetigen Funktion $(f \circ \varphi) \varphi': [a, b] \rightarrow \mathbb{R}$. Also liefert Satz 1.18(b)

$$\int_a^b (f \circ \varphi) \varphi' = F(\varphi(b)) - F(\varphi(a)) = \int_{\varphi(a)}^{\varphi(b)} f. \quad \square$$

Bemerkung 1.25 (Merkregeln zur Substitution). (a) Für $c = \varphi(a)$ und $d = \varphi(b)$,

also $a \in \varphi^{-1}(c)$ und $b \in \varphi^{-1}(d)$, schreiben wir

$$\int_c^d f(x) dx = \left| \begin{array}{l} x = \varphi(y) \\ \frac{dx}{dy} = \varphi'(y) \\ dx = \varphi'(y) dy \end{array} \right| = \int_a^b f(\varphi(y)) \varphi'(y) dy.$$

Diese Substitution kann von links nach rechts oder von rechts nach links gelesen werden, je nach Erfordernis. Falls φ injektiv ist, dann gilt einfach $a = \varphi^{-1}(c)$ und $b = \varphi^{-1}(d)$.

- (b) Falls φ' keine Nullstellen besitzt, dann ist φ streng monoton, also injektiv, und wir haben

$$\int_a^b f(\varphi(x)) dx = \left| \begin{array}{l} y = \varphi(x) \\ x = \varphi^{-1}(y) \\ dx = \frac{1}{\varphi'(\varphi^{-1}(y))} dy \end{array} \right| = \int_{\varphi(a)}^{\varphi(b)} \frac{f(y)}{\varphi'(\varphi^{-1}(y))} dy.$$

- (c) Auf unbestimmte Integrale angewendet, muss man nach der Integration die Substitution rückgängig machen: Kennt man die Stammfunktion F von f , dann gilt:

$$\int f(\varphi(y)) \varphi'(y) dy = \left| \begin{array}{l} x = \varphi(y) \\ dx = \varphi'(y) dy \end{array} \right| = \int f(x) dx = F(x) = F(\varphi(y)).$$

Kennt man die Stammfunktion G von $(f \circ \varphi)\varphi'$ und besitzt φ' keine Nullstelle, dann gilt

$$\int f(x) dx = \left| \begin{array}{l} x = \varphi(y) \\ y = \varphi^{-1}(x) \\ dx = \varphi'(y) dy \end{array} \right| = \int f(\varphi(y)) \varphi'(y) dy = G(y) = G(\varphi^{-1}(x)).$$

Beispiel 1.26. Seien $a, b \in \mathbb{R}$, $a < b$.

- (a) Seien $c, d \in \mathbb{R}$, $c \neq 0$.

$$\int_a^b f(ct + d) dt = \left| \begin{array}{l} x = ct + d \\ t = \frac{x}{c} - \frac{d}{c} \\ dt = \frac{1}{c} dx \end{array} \right| = \frac{1}{c} \int_{ca+d}^{cb+d} f(x) dx$$

für alle stetigen Funktionen $f: [\min\{ca + d, cb + d\}, \max\{ca + d, cb + d\}] \rightarrow \mathbb{R}$. Hier haben wir also mit einer *affin linearen Transformation* substituiert.

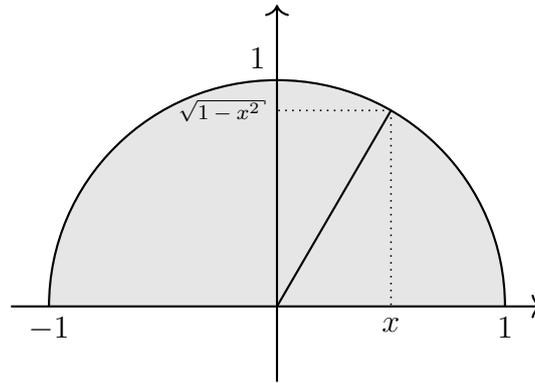


Abbildung 1.4: Halbkreisfläche

(b) Sei $\varphi: [a, b] \rightarrow (0, \infty)$ stetig differenzierbar. Dann gilt

$$\int \frac{\varphi'(x)}{\varphi(x)} dx = \left| \begin{array}{l} y = \varphi(x) \\ dy = \varphi'(x) dx \end{array} \right| = \int \frac{1}{y} dy = \log y = \log(\varphi(x)).$$

Konkretes Beispiel: Mit $\varphi(x) = x^2 + 1$ ist

$$\int_a^b \frac{2x}{x^2 + 1} dx = \int_a^b \frac{\varphi'(x)}{\varphi(x)} dx = \log(\varphi(x)) \Big|_a^b = \log \frac{\varphi(b)}{\varphi(a)} = \log \frac{b^2 + 1}{a^2 + 1}.$$

(c) Mit (b) und partieller Integration erhalten wir

$$\begin{aligned} \int \arctan x dx &= \int \underbrace{1}_{f'} \cdot \underbrace{\arctan x}_g dx = x \arctan x - \int \frac{x}{x^2 + 1} dx \\ &= x \arctan x - \frac{1}{2} \log(x^2 + 1). \end{aligned}$$

(d) Auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$ ist der Sinus bijektiv und stetig differenzierbar und der Cosinus ≥ 0 .
Es folgt

$$\begin{aligned} \int_{-1}^1 \sqrt{1-x^2} dx &= \left| \begin{array}{l} x = \sin t \\ t = \arcsin x \\ dx = \cos t dt \end{array} \right| = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{1-\sin^2 t} \cos t dt = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 t dt \\ &= \frac{1}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (1 + \cos 2t) dt = \left| \begin{array}{l} y = 2t \\ dy = 2 dt \end{array} \right| = \frac{\pi}{2} + \frac{1}{4} \int_{-\pi}^{\pi} \cos y dy = \frac{\pi}{2} + \frac{1}{4} \sin y \Big|_{-\pi}^{\pi} = \frac{\pi}{2}. \end{aligned}$$

Damit haben wir den Flächeninhalt des Halbkreises mit Radius eins, also des Graphen der Funktion $x \mapsto \sqrt{1-x^2}$ auf $[-1, 1]$, ausgerechnet, siehe Abb. 1.4.

Für die Bestimmung von Stammfunktionen rationaler Funktionen überführen wir diese in eine einfacher integrierbare Form:

Bemerkung 1.27 (Partialbruchzerlegung). Seien P, Q reelle, teilerfremde Polynome, sei $N_Q = \{x_1, x_2, \dots, x_n\}$, $n \in \mathbb{N}_0$, die Menge aller reellen Nullstellen von Q und sei $R: \mathbb{R} \setminus N_Q \rightarrow \mathbb{R}$ die durch $R := P/Q$ gegebene rationale Funktion. Ohne Einschränkung nehmen wir an, dass $\text{grad } P < \text{grad } Q$ gilt (sonst Zählergrad durch Polynomdivision reduzieren).

Es existieren $m \in \mathbb{N}$ und Zahlen $k_\nu, \ell_\mu \in \mathbb{N}$ sowie $b_\mu, c_\mu \in \mathbb{R}$ mit

$$Q(x) = (x - x_1)^{k_1} \cdots (x - x_n)^{k_n} \cdot (x^2 + b_1x + c_1)^{\ell_1} \cdots (x^2 + b_mx + c_m)^{\ell_m}.$$

Hier seien die quadratischen Polynome paarweise verschieden und ohne reelle Nullstellen. Ferner existieren Zahlen $A_{\nu,k}, B_{\mu,\ell}, C_{\mu,\ell} \in \mathbb{R}$ und eine Darstellung

$$(1.2) \quad R(x) = \sum_{\nu=1}^n \left(\frac{A_{\nu,1}}{(x - x_\nu)} + \cdots + \frac{A_{\nu,k_\nu}}{(x - x_\nu)^{k_\nu}} \right) + \sum_{\mu=1}^m \left(\frac{B_{\mu,1}x + C_{\mu,1}}{(x^2 + b_\mu x + c_\mu)} + \cdots + \frac{B_{\mu,\ell_\mu}x + C_{\mu,\ell_\mu}}{(x^2 + b_\mu x + c_\mu)^{\ell_\mu}} \right),$$

die *Partialbruchzerlegung* von R .

Die Stammfunktionen der rationalen Funktionen der Form $\frac{A}{(x-a)^k}$ erhält man dann mit Hilfe von Bemerkung und Beispiel 1.21(a) und Beispiel 1.26(a). Die rationalen Funktionen der Form $\frac{Bx+C}{(x^2+bx+c)^\ell}$ integriert man nach einer quadratischen Ergänzung und einer Substitution im Falle $\ell = 1$ direkt (siehe Beispiel 1.26(b)). Im Fall $\ell \geq 2$ gilt eine Rekursionsformel, siehe Merkblatt zur Integration.

Beispiel 1.28. Wir betrachten

$$R(x) := \frac{3x^3 - 4x^2 + 3x - 26}{x^4 + 2x^2 - 8x + 5}.$$

Für den Nenner $Q(x) := x^4 + 2x^2 - 8x + 5$ finden wir durch Ausprobieren und mit Abspalten von Lineartermen mittels Polynomdivision $Q(x) = (x - 1)^2(x^2 + 2x + 5)$. Wegen $x^2 + 2x + 5 = (x + 1)^2 + 4 > 0$ hat Q keine weiteren reellen Nullstellen. Wir machen also den Ansatz

$$(1.3) \quad R(x) = \frac{A}{x - 1} + \frac{B}{(x - 1)^2} + \frac{Cx + D}{x^2 + 2x + 5}.$$

Multiplikation mit $(x - 1)^2$ liefert

$$\frac{3x^3 - 4x^2 + 3x - 26}{x^2 + 2x + 5} = R(x)(x - 1)^2 = A(x - 1) + B + \frac{Cx + D}{x^2 + 2x + 5}(x - 1)^2.$$

Einsetzen von 1 für x liefert $B = -3$. Schließlich setzen wir in (1.3) nacheinander 0, -1

und 2 für x ein und erhalten aus dem resultierenden linearen Gleichungssystem $A = 2$, $C = 1$ und $D = -1$, also

$$(1.4) \quad R(x) = \frac{2}{x-1} - \frac{3}{(x-1)^2} + \frac{x-1}{x^2+2x+5}.$$

Wir rechnen

$$\begin{aligned} \int \frac{x-1}{x^2+2x+5} dx &= \frac{1}{4} \int \frac{x-1}{1 + \left(\frac{x}{2} + \frac{1}{2}\right)^2} dx = \left. \begin{array}{l} y = \frac{x}{2} + \frac{1}{2} \\ x = 2y - 1 \\ dx = 2 dy \end{array} \right| \\ &= \frac{1}{2} \int \frac{2y}{1+y^2} dy - \int \frac{1}{1+y^2} dy = \frac{1}{2} \log(1+y^2) - \arctan y \\ &= \frac{1}{2} \log \left(1 + \left(\frac{x}{2} + \frac{1}{2} \right)^2 \right) - \arctan \left(\frac{x}{2} + \frac{1}{2} \right). \end{aligned}$$

Also folgt aus (1.4)

$$\int R(x) dx = 2 \log|x-1| + \frac{3}{x-1} + \frac{1}{2} \log \left(1 + \left(\frac{x}{2} + \frac{1}{2} \right)^2 \right) - \arctan \left(\frac{x}{2} + \frac{1}{2} \right).$$

Probe!

Beachten Sie bitte das Merkblatt zu den Techniken zur Berechnung unbestimmter Integrale!

_____ Ende der Sitzung vom 23.4. _____

Definition 1.29 (uneigentliche Riemann-Integrale). Sind $a, b \in \mathbb{R} \cup \{\pm\infty\}$, $a < b$ die Randpunkte des Intervalls I , und sei $f: I \rightarrow \mathbb{R}$ eine Funktion.

(a) Ist $a \in I$, $b \notin I$ und $f|_{[a,y]} \in \mathcal{R}(a,y)$ für alle $y \in I$, $y > a$, so setzt man

$$\int_a^b f := \lim_{\substack{y \rightarrow b \\ y \in I}} \int_a^y f,$$

falls dieser Grenzwert in \mathbb{R} existiert.

(b) Sind $b \in I$, $a \notin I$ und $f|_{[y,b]} \in \mathcal{R}(y,b)$ für alle $y \in I$, $y < b$, so setzt man

$$\int_a^b f := \lim_{\substack{y \rightarrow a \\ y \in I}} \int_y^b f,$$

falls dieser Grenzwert in \mathbb{R} existiert.

(c) Es gelte $a, b \notin I$, und es sei $f|_{[y,z]} \in \mathcal{R}(y, z)$ für alle $y, z \in I, y < z$. Dann setzt man

$$\int_a^b f := \int_a^c f + \int_c^b f,$$

falls die Integrale auf der rechten Seite für ein $c \in I$ im Sinne von (a) und (b) in \mathbb{R} existieren. Man sieht leicht mit Satz 1.14 (Intervalladditivität), dass diese Integrale dann auch für jede Wahl von $c \in I$ existieren, und dass die Summe unabhängig von c ist.

Beispiel 1.30. (a) Für $s > -1$ existiert

$$\int_0^1 t^s dt = \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^1 t^s dt = \lim_{\varepsilon \rightarrow 0} \frac{1}{s+1} (1 - \varepsilon^{s+1}) = \frac{1}{s+1}.$$

Für $s \leq -1$ existiert $\int_0^1 t^s dt$ nicht.

(b) Für $s < -1$ existiert

$$\int_1^{\infty} t^s dt = \lim_{y \rightarrow \infty} \int_1^y t^s dt = \lim_{y \rightarrow \infty} \frac{1}{s+1} (y^{s+1} - 1) = -\frac{1}{s+1}.$$

Für $s \geq -1$ existiert $\int_1^{\infty} t^s dt$ nicht.

(c) Das uneigentliche Integral $\int_0^{\infty} t^s dt$ existiert für kein $s \in \mathbb{R}$ wegen (a) und (b).

Bemerkung 1.31. (a) Der Fall $s \in (-1, 0)$ in Beispiel 1.30(a) zeigt: Uneigentliche Riemann-Integrale können existieren, selbst wenn der Integrand unbeschränkt ist (also nicht Riemann-integrierbar).

(b) Sind $a, b \in \mathbb{R}, a < b$ und $f \in \mathcal{R}(a, b)$, dann existiert das uneigentliche Integral von f auf (a, b) und stimmt mit dem Riemann-Integral überein (Übung).

Satz 1.32 (Majorantenkriterium). *Seien $a, b \in \mathbb{R} \cup \{\pm\infty\}, a < b$ die Randpunkte des Intervalls I , und seien $f, g: I \rightarrow \mathbb{R}$ Funktionen mit $f, g \in \mathcal{R}(y, z)$ für alle $y, z \in I, y < z$. Ferner gelte $|f| \leq g$ auf I . Dann gilt: Existiert das uneigentliche Integral $\int_a^b g$, so auch das uneigentliche Integral $\int_a^b f$.*

Beweis. Wir betrachten nur den Fall $I = [a, \infty)$ mit $a \in \mathbb{R}$. Die anderen Fälle lassen sich analog behandeln.

Sei $(x_n)_n$ eine Folge in $[a, \infty)$ mit $\lim_{n \rightarrow \infty} x_n = \infty$, und seien

$$c_n := \int_a^{x_n} f, \quad d_n := \int_a^{x_n} g.$$

Nach Voraussetzung existiert dann

$$\lim_{n \rightarrow \infty} d_n = \int_a^{\infty} g.$$

Für $m, n \in \mathbb{N}$ gilt

$$|c_n - c_m| = \left| \int_{x_m}^{x_n} f \right| \leq \left| \int_{x_m}^{x_n} g \right| = |d_n - d_m|.$$

Somit ist $(c_n)_n$ eine Cauchyfolge, d.h.

$$(1.5) \quad \lim_{n \rightarrow \infty} \int_a^{x_n} f = \lim_{n \rightarrow \infty} c_n \quad \text{existiert.}$$

Ist $(y_n)_n$ eine weitere Folge in $[a, \infty)$ mit $\lim_{n \rightarrow \infty} y_n = \infty$, so kann man das obige Argument auf die „gemischte Folge“ $(x_1, y_1, x_2, y_2, \dots)$ anwenden. Dies liefert die Unabhängigkeit des Grenzwertes in (1.5) von der gewählten Folge $(x_n)_n$, und somit existiert

$$\int_a^\infty f = \lim_{x \rightarrow \infty} \int_a^x f. \quad \square$$

Satz 1.33 (Integralkriterium zur Reihenkonvergenz). *Ist $f: [c, \infty) \rightarrow [0, \infty)$ eine monoton fallende und stetige Funktion für ein $c \in \mathbb{N}$, so konvergiert die Reihe $\sum_{k=c}^\infty f(k)$ genau dann, wenn das uneigentliche Integral $\int_c^\infty f$ existiert.*

Beweis. Nach Voraussetzung gilt

$$f(k) \leq f(t) \leq f(k-1) \quad \text{für } k \geq c+1 \text{ und } t \in [k-1, k],$$

also folgt

$$f(k) \leq \int_{k-1}^k f \leq f(k-1) \quad \text{für } k \geq c+1.$$

Summation ergibt

$$(1.6) \quad \sum_{k=c+1}^N f(k) \leq \int_c^N f \leq \sum_{k=c}^{N-1} f(k) \quad \text{für } N \in \mathbb{N}, N \geq c+1.$$

Der Grenzwert $\mu := \lim_{x \rightarrow \infty} f(x) \in \mathbb{R} \cup \{-\infty\}$ existiert im erweiterten Sinne, da f monoton fällt. Konvergiert die Reihe $\sum_{k=c}^\infty f(k)$, dann ist $(f(k))_k$ eine Nullfolge, also $\mu = 0$. Existiert das uneigentliche Integral $\int_c^\infty f$, dann folgt auch $\mu = 0$. In beiden Fällen liefert dies $f \geq 0$, was wir im Folgenden ohne Einschränkung der Allgemeinheit annehmen werden.

Da f nichtnegativ ist, ist die Funktion $y \mapsto \int_c^y f$ monoton wachsend und es folgt:

$$\begin{aligned} \sum_{k=c}^\infty f(k) \quad \text{konvergiert} & \Leftrightarrow \sup_{\substack{N \in \mathbb{N} \\ N \geq c}} \sum_{k=c}^N f(k) < \infty & \stackrel{(1.6)}{\Leftrightarrow} \\ \sup_{y \geq c} \int_c^y f < \infty & \Leftrightarrow \int_c^\infty f = \lim_{y \rightarrow \infty} \int_c^y f \quad \text{existiert (in } \mathbb{R}). & \square \end{aligned}$$

Beispiel 1.34. Sei $s \in \mathbb{R}$.

- (a) Nach Beispiel 1.30(b) und Satz 1.33 konvergiert die Reihe $\sum_{k=1}^{\infty} k^s$ genau dann, wenn $s < -1$ ist.
- (b) Man kann mit Hilfe von Satz 1.33 zeigen, dass die Reihe $\sum_{k=2}^{\infty} \frac{1}{k(\log k)^s}$ genau dann konvergiert, wenn $s > 1$ ist.

2 Konvergenz von Funktionenfolgen und Funktionenreihen

Wir lassen hier einige Beweise aus, da sie schon im letzten Semester behandelt wurden.

Definition 2.1. Seien $D \subseteq \mathbb{C}$ und $f: D \rightarrow \mathbb{C}$ eine Funktion. Eine Folge $(f_n)_n$ von Funktionen $f_n: D \rightarrow \mathbb{C}$, $n \in \mathbb{N}_0$ heißt

- *punktweise konvergent gegen f* , wenn $\lim_{n \rightarrow \infty} f_n(z) = f(z)$ für alle $z \in D$ gilt.
- *gleichmäßig konvergent gegen f* , wenn f_n und f für fast alle n beschränkt sind und zusätzlich $\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0$ gilt. Man sagt dann auch: $\lim_{n \rightarrow \infty} f_n = f$ in der ∞ -Norm.

Bemerkung 2.2. Gleichmäßige Konvergenz erzwingt punktweise Konvergenz, denn für alle $z \in D$ gilt:

$$\lim_{n \rightarrow \infty} |f_n(z) - f(z)| \leq \lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0.$$

Umgekehrt ist eine punktweise konvergente Folge nicht notwendigerweise gleichmäßig konvergent.

Satz 2.3. Sei $D \subseteq \mathbb{C}$ und sei $(f_n)_n$ eine Folge stetiger Funktionen $f_n: D \rightarrow \mathbb{C}$ welche gleichmäßig gegen eine Funktion $f: D \rightarrow \mathbb{C}$ konvergiert. Dann ist auch f stetig.

Satz 2.4. Seien $a, b \in \mathbb{R}$, $a < b$ und sei $(f_n)_n$ eine Folge stetiger Funktionen $f_n: [a, b] \rightarrow \mathbb{R}$, welche **gleichmäßig** gegen eine Funktion $f: [a, b] \rightarrow \mathbb{R}$ konvergiert. Dann ist f stetig und es gilt

$$\int_a^b f = \lim_{n \rightarrow \infty} \int_a^b f_n$$

Definition 2.5. Sei $D \subseteq \mathbb{C}$, seien $f, f_n: D \rightarrow \mathbb{C}$, $n \in \mathbb{N}_0$ Funktionen, und sei $g_n := \sum_{k=0}^n f_k: D \rightarrow \mathbb{C}$ für $n \in \mathbb{N}_0$. Die Funktionenreihe $\sum_{k=0}^{\infty} f_k$ heißt

- *punktweise konvergent gegen f* , wenn die Folge $(g_n)_n$ punktweise gegen f konvergiert.
- *gleichmäßig konvergent gegen f* , wenn die Folge $(g_n)_n$ gleichmäßig gegen f konvergiert.

Satz 2.6 (Konvergenzkriterium von Weierstraß). Seien $D \subseteq \mathbb{C}$ und $f_n: D \rightarrow \mathbb{C}$, $n \in \mathbb{N}_0$ beschränkte Funktionen mit $\sum_{n=0}^{\infty} \|f_n\|_\infty < \infty$. Dann konvergiert die Reihe $\sum_{n=0}^{\infty} f_n$ gleichmäßig gegen eine Funktion $f: D \rightarrow \mathbb{C}$. Ist ferner f_n stetig für alle $n \in \mathbb{N}_0$, so ist auch f stetig.

Definition 2.7. In einem normierten Raum V verwenden wir für $x \in V$ und $r > 0$ die folgenden Schreibweisen:

$$B_r(x) := \{y \in V \mid \|y - x\| < r\} \quad \text{offene Kugel mit Zentrum } x \text{ und Radius } r$$

$$\overline{B}_r(x) := \{y \in V \mid \|y - x\| \leq r\} \quad \text{abgeschlossene Kugel mit Zentrum } x \text{ und Radius } r.$$

Für $V = \mathbb{R}$ gilt $B_r(x) = (x - r, x + r)$ und $\overline{B}_r(x) = [x - r, x + r]$. Für $V = \mathbb{C}$ sagen wir meist *Kreisscheibe* statt Kugel. In diesen Fällen ist die Norm durch den Betrag gegeben.

Satz 2.8 (Konvergenz von Potenzreihen). Seien $\zeta \in \mathbb{C}$ und $(a_n)_{n \geq 0} \subseteq \mathbb{C}$. Sei ferner r der Konvergenzradius von $\sum_n a_n(z - \zeta)^n$ (siehe Definition Mathe I). Dann gilt:

- (a) Ist $z \in B_r(\zeta)$, so konvergiert die Reihe $\sum_{n=0}^{\infty} a_n(z - \zeta)^n$ absolut.
- (b) Sei $s \in [0, r)$. Dann ist die Reihe der Funktionen $f_n: \overline{B}_s(\zeta) \rightarrow \mathbb{C}$, $f_n(z) = a_n(z - \zeta)^n$, gleichmäßig konvergent. Man sagt kurz: Die Potenzreihe $\sum_{n=0}^{\infty} a_n(z - \zeta)^n$ konvergiert gleichmäßig auf $\overline{B}_s(\zeta)$.
- (c) Die Funktion $B_r(\zeta) \rightarrow \mathbb{C}$, $f(z) = \sum_{n=0}^{\infty} a_n(z - \zeta)^n$, ist stetig.
- (d) Ist $z \in \mathbb{C} \setminus \overline{B}_r(\zeta)$, so divergiert die Reihe $\sum_{n=0}^{\infty} a_n(z - \zeta)^n$.

Definition 2.9. Seien I ein offenes Intervall, $\xi \in I$ und $f \in C^\infty(I)$. Die Potenzreihe

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(\xi)}{n!} (x - \xi)^n$$

heißt *Taylorreihe von f in ξ* .

Beispiel 2.10. (a) Seien $f = \log: (0, \infty) \rightarrow \mathbb{R}$, $\xi = 1$. Dann ist

$$\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (x - 1)^k$$

die zugehörige Taylorreihe in 1.

(b) Seien $\alpha \in \mathbb{R}$ und $g \in C^\infty((-1, \infty))$ gegeben durch

$$g(x) = (1 + x)^\alpha.$$

Dann ist die *binomische Reihe* $\sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$ die Taylorreihe von g in 0.

Hier verwenden wir die Definition der *verallgemeinerten Binomialkoeffizienten* für $\alpha \in \mathbb{R}$:

$$\binom{\alpha}{0} := 1 \quad \text{und} \quad \binom{\alpha}{k} := \frac{\alpha(\alpha - 1)(\alpha - 2) \dots (\alpha - k + 1)}{k!}, \quad k \in \mathbb{N}.$$

- (c) Die Taylorreihen von \exp , \sin , \cos etc. in 0 sind durch die Reihendarstellungen aus ihrer Definition gegeben.

Bemerkung 2.11. Die Taylorreihe einer C^∞ -Funktion muss nicht mit der eigentlichen Funktion übereinstimmen. Zum Beispiel ist

$$f: \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) := \begin{cases} e^{-\frac{1}{x^2}}, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

eine C^∞ -Funktion mit $f^{(n)}(0) = 0$ für alle $n \in \mathbb{N}$. Die Taylorreihe von f im Entwicklungspunkt 0 ist also identisch 0, obwohl $f(x) > 0$ für alle $x \in \mathbb{R} \setminus \{0\}$ gilt.

Satz 2.12 (Satz von Bernstein). *Seien $\xi \in \mathbb{R}$, $r > 0$ und $f \in C^\infty(\xi - r, \xi + r)$ eine Funktion, welche **eine** der folgenden Eigenschaften hat:*

- (i) *Es existiert $N \in \mathbb{N}$ mit $f^{(n)}(x) \geq 0$ für alle $n \geq N$, $x \in (\xi - r, \xi + r)$ oder*
- (ii) *Es existiert $N \in \mathbb{N}$ mit $(-1)^n f^{(n)}(x) \geq 0$ für alle $n \geq N$, $x \in (\xi - r, \xi + r)$.*

Dann konvergiert die Taylorreihe von f in allen Punkten aus $(\xi - r, \xi + r)$ gegen f .

Beweis. Siehe z.B. [6, Abschnitt 10.3.5, S. 374-378]. □

Bemerkung 2.13. Für einige der Beispiele aus Beispiel 2.10 können wir Satz 2.12 anwenden: In (a) erfüllt $-\log$ die Bedingung (ii), und in (b) erfüllt entweder g oder $-g$ die Bedingung (ii).

Satz 2.14. *Jede Potenzreihe $f(x) = \sum_{k=0}^{\infty} a_k(x-\xi)^k$ mit Konvergenzradius r ist innerhalb ihres Konvergenzintervalls beliebig oft differenzierbar und es gilt*

- (a) $f^{(n)}(x) = \sum_{k=n}^{\infty} \frac{k!}{(k-n)!} a_k(x-\xi)^{k-n}$ für $x \in (\xi - r, \xi + r)$, $n \in \mathbb{N}_0$.
- (b) $a_n = \frac{1}{n!} f^{(n)}(\xi)$ für $n \in \mathbb{N}_0$.

Insbesondere ist die Potenzreihenentwicklung einer C^∞ -Funktion innerhalb des Konvergenzradius eindeutig und entspricht demnach der Taylorreihe von f .

Beispiel 2.15. Wir wollen nun eine Taylorreihe von $\arctan: \mathbb{R} \rightarrow \mathbb{R}$ im Entwicklungspunkt 0 bestimmen. Dafür ist zu bemerken, dass für $x \in (-1, 1)$

$$\arctan'(x) = \frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n}$$

gilt. Hierbei ist $x \in (-1, 1)$ zu wählen da die Reihe auf der rechten Seite Konvergenzradius 1 besitzt. Setze nun

$$g: (-1, 1) \rightarrow \mathbb{R}, \quad g(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1},$$

dann ist g stetig differenzierbar mit $g'(x) = \arctan'(x)$ für $x \in (-1, 1)$ nach Satz 2.14. Schließlich ist $g(0) = 0 = \arctan(0)$ und damit folgt

$$\arctan(x) = g(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \quad \text{für } x \in (-1, 1).$$

Mit Hilfe des Leibniz-Kriteriums folgt, dass die Reihe auch für ± 1 konvergiert. Es lässt sich zeigen, dass dabei dann die Gleichheit auch für $x \in [-1, 1]$ gilt. Dies ist jedoch nicht direkt offensichtlich. Die Reihe konvergiert nicht für $|x| > 1$, obwohl $\arctan(x)$ dort definiert ist.

Satz 2.16 (Weierstraß'scher Approximationssatz). *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f: [a, b] \rightarrow \mathbb{R}$ stetig. Dann kann man f auf dem kompakten Intervall $[a, b]$ gleichmäßig durch reelle Polynome approximieren: Für jedes $\varepsilon > 0$ existiert ein reelles Polynom p mit $\|f - p\|_{\infty} < \varepsilon$.*

Beweisskizze. Es lässt sich zeigen, dass die Funktionen

$$B_n^{(f)}(x) := \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}, \quad x \in [0, 1], n \in \mathbb{N}$$

im Fall $a = 0, b = 1$ für $n \rightarrow \infty$ gleichmäßig gegen f konvergieren. Den allgemeinen Fall beweist man, indem man die Komposition von f mit einer affinen Abbildung von $[0, 1]$ auf $[a, b]$ betrachtet. $B_n^{(f)}$ bezeichnet man auch als *Bernstein-Polynom*. \square

Ende der Sitzung vom 27.4.

2.1 Fourierreihen

Definition 2.17. Sei $q > 0$. Eine Funktion $f: \mathbb{R} \rightarrow \mathbb{C}$ heißt *q-periodisch*, wenn gilt

$$f(x+q) = f(x) \quad \text{für alle } x \in \mathbb{R}.$$

Offenbar gilt dann auch $f(x+kq) = f(x)$ für alle $k \in \mathbb{Z}$. Falls

$$p := \min\{q > 0 \mid f \text{ ist } q\text{-periodisch}\}$$

existiert, dann heißt p die (*minimale*) *Periode* von f .

Bemerkung 2.18. (a) Ist f periodisch, stetig und nicht konstant, dann kann man zeigen, dass die minimale Periode existiert.

(b) Es gibt periodische Funktionen, die keine minimale Periode besitzen, z.B. die konstanten Funktionen. Man kann aber auch nicht konstante (unstetige) periodische Funktionen ohne minimale Periode konstruieren.

(c) Ist f periodisch mit Periode p und $L > 0$, so ist

$$\tilde{f}: \mathbb{R} \rightarrow \mathbb{C}, \quad \tilde{f}(x) = f\left(\frac{p}{L}x\right)$$

periodisch mit Periode L . Ohne Einschränkung der Allgemeinheit werden wir daher im Folgenden meist 2π -periodische Funktionen betrachten.

Es seien $(a_n)_{n \in \mathbb{N}_0}, (b_n)_{n \in \mathbb{N}} \subseteq \mathbb{C}$ und $f: \mathbb{R} \rightarrow \mathbb{C}$ gegeben durch

$$(2.1) \quad f(x) := \frac{a_0}{2} + \sum_{k=1}^n [a_k \cos(kx) + b_k \sin(kx)].$$

Dann ist f eine 2π -periodische Funktion. Um diese Darstellung von f zu vereinfachen, setzen wir $c_0 := \frac{a_0}{2}$ sowie

$$(2.2) \quad c_k := \begin{cases} \frac{1}{2}(a_k - ib_k), & k > 0, \\ \frac{1}{2}(a_{-k} + ib_{-k}), & k < 0. \end{cases}$$

Also gilt $a_k = c_k + c_{-k}$ und $b_k = i(c_k - c_{-k})$ für $k > 0$. Es gilt dann nach Wahl der c_k

$$\begin{aligned} \sum_{k=-n}^n c_k e^{ikx} &= c_0 + \sum_{k=1}^n c_k e^{ikx} + \sum_{k=-n}^{-1} c_k e^{ikx} \\ &= \frac{a_0}{2} + \sum_{k=1}^n c_k e^{ikx} + \sum_{k=1}^n c_{-k} e^{-ikx} \\ &= \frac{a_0}{2} + \sum_{k=1}^n \left[\frac{1}{2}(a_k - ib_k) e^{ikx} + \frac{1}{2}(a_k + ib_k) e^{-ikx} \right] \\ &= \frac{a_0}{2} + \sum_{k=1}^n \left[\frac{a_k}{2} (e^{ikx} + e^{-ikx}) - \frac{ib_k}{2} (e^{ikx} - e^{-ikx}) \right] \\ &= \frac{a_0}{2} + \sum_{k=1}^n [a_k \cos(kx) + b_k \sin(kx)] = f(x). \end{aligned}$$

Solch eine Summe $f(x) = \sum_{k=-n}^n c_k e^{ikx} = \sum_{k=-n}^n c_k (e^{ix})^k$ heißt *trigonometrisches Polynom*. Beobachtung: Ist k_0 der größte gemeinsame Teiler derjenigen k mit $c_k \neq 0$, dann ist $\frac{2\pi}{k_0}$ die Periode von f .

Seien $g, h: I \rightarrow \mathbb{C}$, $I \subseteq \mathbb{R}$ ein kompaktes Intervall, komplexwertige Funktionen. Die Funktion g heißt *Riemann-integrierbar*, $g \in \mathcal{R}(I)$, wenn $\operatorname{Re} g$ und $\operatorname{Im} g$ Riemann-integrierbar sind. In diesem Fall definieren wir das Integral von g durch

$$\int_I g := \int_I \operatorname{Re} g + i \int_I \operatorname{Im} g.$$

Sind $\operatorname{Re} h$ und $\operatorname{Im} h$ differenzierbar, dann definieren wir analog die Ableitung von h durch $h' = (\operatorname{Re} h)' + i(\operatorname{Im} h)'$. Ist g stetig, dann gilt das Analogon des Hauptsatzes (Satz 1.18) für g . Wir bezeichnen wieder mit $\int g$ eine unbestimmte Stammfunktion und berechnen Integrale genauso wie im reellen mit Hilfe von Stammfunktionen.

Satz 2.19. *Ist f ein trigonometrisches Polynom, so sind die Koeffizienten a_k und b_k bzw. c_k eindeutig bestimmt durch*

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) \, dx & k = 0, 1, \dots, n \\ b_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) \, dx & k = 1, 2, \dots, n \\ c_k &= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} \, dx & k = -n, -n+1, \dots, n-1, n. \end{aligned}$$

Beweis. Ist $\alpha \in \mathbb{R} \setminus \{0\}$, dann hat $e^{i\alpha x}$ die Stammfunktion $\frac{1}{i\alpha} e^{i\alpha x}$ (einfache Rechnung). Für $k, \ell \in \mathbb{Z}$ folgt daraus und aus der $2\pi i$ -Periodizität der Exponentialfunktion

$$(2.3) \quad \frac{1}{2\pi} \int_0^{2\pi} e^{ikx} e^{-i\ell x} \, dx = \frac{1}{2\pi} \int_0^{2\pi} e^{i(k-\ell)x} \, dx = \delta_{k\ell}.$$

Dies liefert

$$\frac{1}{2\pi} \int_0^{2\pi} \left(\sum_{k=-n}^n c_k e^{ikx} \right) e^{-i\ell x} \, dx = \sum_{k=-n}^n c_k \frac{1}{2\pi} \int_0^{2\pi} e^{ikx} e^{-i\ell x} \, dx = \sum_{k=-n}^n c_k \delta_{k\ell} = c_\ell.$$

Die restlichen Identitäten folgen dann aus (2.2). □

Bemerkung 2.20. Satz 2.19 gilt auch, wenn f die folgende Form mit gleichmäßig konvergenten Reihen hat:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)] = \sum_{k=-\infty}^{\infty} c_k e^{ikx}.$$

Definition 2.21. Sei $f \in \mathcal{R}(0, 2\pi)$ 2π -periodisch. Für $k \in \mathbb{Z}$ bezeichnen wir mit

$$c_k := c_k(f) := \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x) \, dx$$

den k -ten *Fourierkoeffizienten*. Ferner nennen wir die formale Reihe

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

Fourierreihe von f (ohne eine Aussage über ihre Konvergenz zu machen).

Im Folgenden wollen wir die Frage klären, ob und in welchem Sinn Fourierreihen konvergieren.

Definition 2.22. Wir setzen

$$P_{2\pi} := \{f \in \mathcal{R}(0, 2\pi) \mid f \text{ ist } 2\pi\text{-periodisch}\}$$

sowie

$$\langle f, g \rangle := \frac{1}{2\pi} \int_0^{2\pi} \overline{f}g \quad \text{und} \quad \|f\| := \langle f, f \rangle^{\frac{1}{2}}.$$

Eine Folge $(f_n)_n \subseteq P_{2\pi}$ konvergiert im *Quadratmittel* (bzw. im L^2 -Sinn) gegen $f \in P_{2\pi}$, wenn gilt

$$\|f_n - f\|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f_n - f|^2 \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Es lässt sich zeigen, dass $\langle \cdot, \cdot \rangle$ ein Semiskalarprodukt ist (nur positiv **semidefinit**). Identifiziert man alle Funktionen $f, g \in P_{2\pi}$, die $\|f - g\| = 0$ erfüllen, so ist $(P_{2\pi}, \langle \cdot, \cdot \rangle)$ nach dieser Identifikation ein Prähilbertraum.

Konvergiert (f_n) gleichmäßig gegen f , so konvergiert $(f_n)_n$ auch im Quadratmittel gegen f . Die Umkehrung ist im allgemeinen falsch (Übung).

Wir definieren für $k \in \mathbb{Z}$ die Funktionen $e_k: \mathbb{R} \rightarrow \mathbb{C}$ durch $e_k(x) := e^{ikx}$. Dann gilt $\langle e_k, e_\ell \rangle = \delta_{k\ell}$, wie schon im Beweis von Satz 2.19 gezeigt, d.h. (e_k) ist eine *Orthonormalfolge*. Für $f \in P_{2\pi}$ haben wir $c_k = c_k(f) = \langle e_k, f \rangle$.

Satz 2.23 (Entwicklungssatz). *Seien $f \in P_{2\pi}$ und $c_k = \langle e_k, f \rangle$ wie oben. Dann haben wir für*

$$s_n := \sum_{k=-n}^n c_k e_k, \quad n \in \mathbb{N}_0,$$

folgende Aussagen:

(a) $\|f - s_n\|^2 = \|f\|^2 - \sum_{k=-n}^n |c_k|^2$ für alle $n \in \mathbb{N}_0$.

(b) Es gilt die Besselsche Ungleichung

$$(2.4) \quad \sum_{k=-\infty}^{\infty} |c_k|^2 \leq \|f\|^2.$$

Insbesondere folgt $c_k \rightarrow 0$ für $k \rightarrow \pm\infty$.

(c) Die Folge $(s_n)_n$ konvergiert im Quadratmittel gegen f , d.h. $\|s_n - f\| \rightarrow 0$ für $n \rightarrow \infty$, genau dann, wenn die Parsevalsche Gleichung gilt:

$$(2.5) \quad \sum_{k=-\infty}^{\infty} |c_k|^2 = \|f\|^2$$

Beweis. Wir haben

$$\|s_n\|^2 = \sum_{k,\ell=-n}^n \overline{c_k} c_\ell \langle e_k, e_\ell \rangle = \sum_{k,\ell=-n}^n \overline{c_k} c_\ell \delta_{k\ell} = \sum_{k=-n}^n |c_k|^2$$

und

$$\langle f, s_n \rangle = \sum_{k=-n}^n c_k \langle f, e_k \rangle = \sum_{k=-n}^n c_k \overline{c_k} = \sum_{k=-n}^n |c_k|^2.$$

Daraus folgt

$$\|f - s_n\|^2 = \|f\|^2 - \langle f, s_n \rangle - \langle s_n, f \rangle + \|s_n\|^2 = \|f\|^2 - \sum_{k=-n}^n |c_k|^2,$$

also (a). Die anderen beiden Aussagen folgen hieraus, da stets $\|f - s_n\|^2 \geq 0$ gilt. \square

Satz 2.24 (Vollständigkeitsatz). *Sei $f \in P_{2\pi}$. Dann konvergiert die Fourierreihe von f im quadratischen Mittel gegen f , d.h. die Parsevalsche Gleichung gilt.*

Beweis. Siehe z.B. [28, Satz 8.16]. Der Beweis beruht auf einer Approximation von f zunächst durch eine 2π -periodische stetige Funktion und dann durch ein trigonometrisches Polynom mit Hilfe des Approximationssatzes von Stone-Weierstrass. \square

Bemerkung 2.25. Man kann zeigen, dass Satz 2.24 sogar für nicht periodische Funktionen in $\mathcal{R}(0, 2\pi)$ gilt, siehe z.B. [17, Satz 141.1].

Im Allgemeinen gilt in Satz 2.24 keine gleichmäßige Konvergenz der Fourierreihe und auch keine punktweise Konvergenz. Unter Zusätzlichen Bedingungen kann man dies aber erreichen:

Satz 2.26. *Sei $f \in P_{2\pi}$ stetig und stückweise stetig differenzierbar. Dann konvergiert die Fourierreihe von f gleichmäßig gegen f .*

Beweis. Sei $0 = t_0 < t_1 < \dots < t_m = 2\pi$ eine Zerlegung von $[0, 2\pi]$ so, dass $f|_{[t_{j-1}, t_j]}$ für $j = 1, \dots, m$ stetig differenzierbar ist und setze

$$\varphi: \mathbb{R} \rightarrow \mathbb{C}, \quad \varphi(x) := \begin{cases} f'(x), & x \neq t_j + 2k\pi, \\ 0, & x = t_j + 2k\pi. \end{cases}$$

Dann ist $\varphi \in P_{2\pi}$ nach Satz 1.12(a). Seien γ_k die Fourierkoeffizienten von φ . Dann liefert die Besselsche Ungleichung (Satz 2.23(b))

$$\sum_{k=-\infty}^{\infty} |\gamma_k|^2 \leq \|\varphi\|^2 < \infty.$$

Ferner gilt für die Fourierkoeffizienten c_k von f

$$\begin{aligned} c_k &= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx = \frac{1}{2\pi} \sum_{j=1}^m \int_{t_{j-1}}^{t_j} f(x) e^{-ikx} dx \\ &= \frac{i}{2k\pi} \sum_{j=1}^m f(x) e^{-ikx} \Big|_{t_{j-1}}^{t_j} - \frac{i}{2k\pi} \sum_{j=1}^m \int_{t_{j-1}}^{t_j} \varphi(x) e^{-ikx} dx \\ &= 0 - \frac{i}{2k\pi} \int_0^{2\pi} \varphi(x) e^{-ikx} dx = -\frac{i}{k} \gamma_k. \end{aligned}$$

Die Ungleichung zwischen dem geometrischen und arithmetischen Mittel (GAM) liefert $|c_k| \leq \frac{1}{2}(k^{-2} + |\gamma_k|^2)$. Folglich ist $\sum_{k=-\infty}^{\infty} |c_k|$ konvergent. Wegen $|e_k| \equiv 1$ ist $\sum_{k=-\infty}^{\infty} c_k e^{ikx}$ also gleichmäßig konvergent und konvergiert gegen eine stetige Funktion g (Satz 2.6). Schließlich liefert der Vollständigkeitsatz (Satz 2.24)

$$\|f - g\| \leq \|f - s_n\| + \|s_n - g\| \rightarrow 0 \quad \text{für } n \rightarrow \infty,$$

wobei $s_n := \sum_{k=-n}^n c_k e_k$ sei. Also ist $\|f - g\| = 0$. Angenommen, es würde $f \neq g$ gelten. Dann gäbe es x_0 mit $f(x_0) \neq g(x_0)$, und da g und f stetig sind, existierten $\varepsilon > 0$ und ein nicht ausgeartetes Intervall $I \subseteq [0, 2\pi]$ mit $|f - g| \geq \varepsilon$ auf I . Es folgte

$$2\pi \|f - g\|^2 = \int_0^{2\pi} |f - g|^2 \geq \int_I |f - g|^2 \geq \varepsilon^2 |I| > 0.$$

Dies wäre ein Widerspruch zu $\|f - g\| = 0$, also muss $f = g$ sein. \square

Beispiel 2.27. (a) Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ die 2π -periodische Funktion mit $f(x) = |x|$ für $|x| \leq \pi$. Dann konvergiert die Fourierreihe gleichmäßig nach Satz 2.26. Zur Berechnung der Fourierkoeffizienten beobachten wir, dass f gerade ist. Die Formel für die Koeffizienten b_k in Satz 2.19 liefert dann $b_k = 0$ für alle $k \in \mathbb{N}$. Somit muss f aus einer reinen Kosinusreihe bestehen, d.h. es gilt

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(kx).$$

Wir rechnen für $k \in \mathbb{N}_0$:

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx = \frac{1}{\pi} \int_0^{\pi} x \cos(kx) dx + \frac{1}{\pi} \int_{\pi}^{2\pi} (-x + 2\pi) \cos(kx) dx \\ &= \frac{1}{\pi} \int_0^{\pi} x \cos(kx) dx + \frac{1}{\pi} \int_{\pi}^{2\pi} (-x + 2\pi) \cos(k(-x + 2\pi)) dx \\ &= \frac{1}{\pi} \int_0^{\pi} x \cos(kx) dx - \frac{1}{\pi} \int_{\pi}^0 x \cos(kx) dx = \frac{2}{\pi} \int_0^{\pi} x \cos(kx) dx. \end{aligned}$$

Insbesondere gilt also $a_0 = \frac{2}{\pi} \int_0^\pi x \, dx = \pi$ sowie für $k \in \mathbb{N}$

$$\begin{aligned} a_k &= \frac{2}{\pi} \int_0^\pi x \cos(kx) \, dx = \frac{2}{\pi} \frac{1}{k} x \sin(kx) \Big|_0^\pi - \frac{2}{k\pi} \int_0^\pi \sin(kx) \, dx \\ &= \frac{2}{k^2\pi} \cos(kx) \Big|_0^\pi = \frac{2}{k^2\pi} ((-1)^k - 1) = \begin{cases} 0, & k \text{ gerade, } k \neq 0, \\ -\frac{4}{k^2\pi}, & k \text{ ungerade.} \end{cases} \end{aligned}$$

Es folgt also für $x \in [-\pi, \pi]$

$$|x| = \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\cos((2k-1)x)}{(2k-1)^2}.$$

(b) (Sägefunktion) Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ die 2π -periodische Funktion mit $f(x) = x$ in $(-\pi, \pi]$. Es ist dann

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} \, dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} x e^{-ikx} \, dx.$$

Es folgt $c_0 = 0$ und für $k \neq 0$

$$\begin{aligned} c_k &= \frac{i}{2k\pi} x e^{-ikx} \Big|_{-\pi}^{\pi} - \frac{i}{2k\pi} \int_{-\pi}^{\pi} e^{-ikx} \, dx \\ &= \frac{i}{2k\pi} (\pi e^{-ik\pi} + \pi e^{ik\pi}) + \frac{1}{2k^2\pi} e^{-ikx} \Big|_{-\pi}^{\pi} \\ &= \frac{i}{k} \cos(k\pi) + 0 = \frac{i}{k} (-1)^k. \end{aligned}$$

Also ist die Fourierreihe gegeben durch

$$\sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} (-1)^k \frac{i}{k} e^{ikx} = \sum_{k=1}^{\infty} (-1)^k \frac{i}{k} (e^{ikx} - e^{-ikx}) = 2 \sum_{k=1}^{\infty} (-1)^{k+1} \frac{\sin(kx)}{k}.$$

Wiederum war zu erwarten, dass f eine reine Sinusreihe als Fourierreihe hat, da f ungerade ist. Es lässt sich zeigen, dass die Reihe für alle $x \in \mathbb{R} \setminus \{(2k+1)\pi \mid k \in \mathbb{Z}\}$ gegen $f(x)$ und in den Punkten $(2k+1)\pi$ gegen 0 konvergiert.

Die letzte Tatsache aus Beispiel 2.27(b) lässt sich verallgemeinern:

Satz 2.28. *Ist $f \in P_{2\pi}$, $x_0 \in [0, 2\pi)$ und existieren*

$$\begin{aligned} f_{\pm}(x_0) &:= \lim_{h \rightarrow 0_{\pm}} f(x_0 + h), \\ f'_{\pm}(x_0) &:= \lim_{h \rightarrow 0_{\pm}} \frac{1}{h} (f(x_0 + h) - f_{\pm}(x_0)), \end{aligned}$$

dann gilt

$$\lim_{n \rightarrow \infty} \sum_{k=-n}^n c_k e^{ikx_0} = \frac{1}{2}(f_+(x_0) + f_-(x_0)).$$

Beweis. Siehe z.B. [17, 136.1 Dirichletsche Regel, S. 138].

□

Ende der Sitzung vom 30.4.

3 Lineare Algebra II

3.1 Die Determinante einer Matrix

In Zusammenhang mit Matrizen und der Untersuchung von linearen Gleichungssystemen ist die *Determinantenfunktion* (oder kurz: *Determinante*) nützlich. Insbesondere wird sie uns ein weiteres Hilfsmittel geben, um lineare Unabhängigkeit von Vektoren zu testen und quadratische Gleichungssysteme zu lösen.

Im Folgenden sei stets $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Ferner seien $E_n \in \mathbb{K}^{n \times n}$ die Einheitsmatrix in n Dimensionen und $\{e_1, \dots, e_n\}$ die kanonische Basis von \mathbb{K}^n , d.h. $(e_i)_j = \delta_{ij}$.

Definition 3.1. Eine Funktion $\det: \underbrace{\mathbb{K}^n \times \dots \times \mathbb{K}^n}_{n \text{ Argumente}} \rightarrow \mathbb{K}$ heißt *Determinantenfunktion*

(bzw. kurz *Determinante*), wenn sie folgende Eigenschaften besitzt:

- (a) $\det(\cdot)$ ist *multilinear*, d.h. in jedem ihrer n Argumente linear: Seien $a_1, \dots, a_n, b \in \mathbb{K}^n$ und $r \in \mathbb{K}$ dann gilt für jedes $j \in \{1, \dots, n\}$

$$\begin{aligned} \det(a_1, a_2, \dots, a_{j-1}, ra_j + b, a_{j+1}, \dots, a_n) \\ = r \cdot \det(a_1, a_2, \dots, a_n) + \det(a_1, a_2, \dots, a_{j-1}, b, a_{j+1}, \dots, a_n). \end{aligned}$$

- (b) $\det(\cdot)$ ist *alternierend*, d.h. bei Vertauschung zweier Argumente ändert die Determinante nur ihr Vorzeichen: Seien $a_1, \dots, a_n \in \mathbb{K}^n$, dann gilt für $1 \leq i < j \leq n$

$$\begin{aligned} \det(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_{j-1}, a_j, a_{j+1}, \dots, a_n) \\ = - \det(a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_{j-1}, a_i, a_{j+1}, \dots, a_n). \end{aligned}$$

- (c) $\det(\cdot)$ ist *normiert*, d.h. $\det(e_1, \dots, e_n) = 1$.

Bemerkung 3.2. Eine Determinantenfunktion wird auch *normierte alternierende Multilinearform auf \mathbb{K}^n* genannt. Für $A = (a_1 \mid a_2 \mid \dots \mid a_n) \in \mathbb{K}^{n \times n}$ setzen wir $\det(A) := \det(a_1, \dots, a_n)$ und interpretieren eine Determinantenfunktion als eine Funktion $\mathbb{K}^{n \times n} \rightarrow \mathbb{K}$, welche bzgl. der Spalten multilinear und alternierend ist und ferner $\det(E_n) = 1$ erfüllt. Weiter unten werden wir zeigen, dass zu jedem $n \in \mathbb{N}$ genau eine Determinantenfunktion existiert.

Beispiel 3.3. Für die folgenden Matrizen, die auch *Elementarmatrizen* genannt werden, können wir nun bereits „die“ Determinante ausrechnen:

- (a) Für $i, j \in \{1, \dots, n\}$, $i \neq j$ betrachte die Matrix E_{ij} , welche aus der Einheitsmatrix $E_n = (e_1 \mid \dots \mid e_n)$ durch Vertauschung der i -ten und j -ten Zeile entsteht, d.h.

$$E_{ij} := \begin{pmatrix} 1 & & \dots & & 0 \\ & \ddots & & & \\ & & 0 & \dots & 1 \\ \vdots & & \vdots & \ddots & \vdots \\ & & 1 & \dots & 0 \\ & & & & \ddots & \\ 0 & & \dots & & & 1 \end{pmatrix} \in \mathbb{K}^{n \times n}.$$

Dann gilt $\det(E_{ij}) = -1$. Ist $A \in \mathbb{K}^{n \times n}$, dann vertauscht E_{ij} durch Multiplikation von links die i -te Zeile mit der j -ten Zeile von A und durch Multiplikation von rechts die i -te Spalten mit der j -ten Spalte von A .

- (b) Für $i, j \in \{1, \dots, n\}$, $i \neq j$ und $r \in \mathbb{K}$ betrachte die Matrix $E_{ij}(r)$, welche die Einheitsmatrix beschreibt und an der i, j -ten Stelle den Wert r hat, d.h.

$$E_{ij}(r) := \begin{pmatrix} 1 & & \dots & & 0 \\ & \ddots & & & \\ & & 1 & 0 & r \\ \vdots & & \vdots & \ddots & \vdots \\ & & & & 1 \\ & & & & \ddots & \\ 0 & & \dots & & & 1 \end{pmatrix} \in \mathbb{K}^{n \times n}.$$

Wir haben nach Vertauschen des i -ten Arguments mit dem j -ten Argument (beachte $i \neq j$)

$$\det(e_1, \dots, e_{j-1}, e_i, e_{j+1}, \dots, e_n) = -\det(e_1, \dots, e_{j-1}, e_i, e_{j+1}, \dots, e_n) = 0.$$

Daraus folgt

$$\begin{aligned} \det(E_{ij}(r)) &= \det(e_1, \dots, e_{j-1}, e_j + r e_i, e_{j+1}, \dots, e_n) \\ &= \det(E_n) + r \det(e_1, \dots, e_{j-1}, e_i, e_{j+1}, \dots, e_n) = 1. \end{aligned}$$

Ist $A \in \mathbb{K}^{n \times n}$, dann beschreibt $E_{ij}(r)$ durch Multiplikation von links die Addition des r -fachen der j -ten Zeile von A zur i -ten Zeile von A und durch Multiplikation von rechts die Addition des r -fachen der i -ten Spalte von A zur j -ten Spalte von A .

Satz 3.4. „Die“ Determinantenfunktion hat folgende Eigenschaften: Sei $A = (a_1 \mid \dots \mid a_n) \in \mathbb{K}^{n \times n}$.

(a) Ist eine Spalte von A gleich 0 oder stimmen zwei Spalten überein, dann gilt $\det(A) = 0$.

(b) Addition eines Vielfachen einer Spalte von A zu einer anderen Spalte ändert die Determinante von A nicht, d.h. für $r \in \mathbb{K}$ und $k \neq j$ ist

$$\det(A) = \det(a_1, a_2, \dots, a_{j-1}, a_j + ra_k, a_{j+1}, \dots, a_n) (= \det(A \cdot E_{kj}(r))).$$

(c) $\det(A) = 0$ gilt genau dann, wenn die Spalten von A linear abhängig sind.

Beweis. Zu (a): Sei zunächst $A = (a_1 \mid \dots \mid a_n) \in \mathbb{K}^{n \times n}$ derart, dass $a_j = 0$ für ein $j \in \{1, \dots, n\}$ gilt. Dann folgt mit der Linearität der Determinante im j -ten Argument

$$\begin{aligned} \det(A) &= \det(a_1, \dots, a_{j-1}, 0, a_{j+1}, \dots, a_n) \\ &= \det(a_1, \dots, a_{j-1}, 0 \cdot a_j, a_{j+1}, \dots, a_n) \\ &= 0 \cdot \det(a_1, \dots, a_{j-1}, a_j, a_{j+1}, \dots, a_n) = 0. \end{aligned}$$

Sei nun $A = (a_1 \mid \dots \mid a_n) \in \mathbb{K}^{n \times n}$ derart, dass $a_i = a_j$ für ein Paar $i, j \in \{1, \dots, n\}$, $i \neq j$, gilt. Dann ist

$$\begin{aligned} \det(A) &= \det(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_{j-1}, a_i, a_{j+1}, \dots, a_n) \\ &= -\det(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_{j-1}, a_i, a_{j+1}, \dots, a_n) = -\det(A). \end{aligned}$$

Demnach ist $\det(A) = 0$.

Aussage (b) folgt nun direkt aus der Multilinearität und (a), denn für $r \in \mathbb{K}$ ist

$$\begin{aligned} \det(a_1, \dots, a_{j-1}, a_j + ra_k, a_{j+1}, \dots, a_n) &= \det(A) + r \det(a_1, \dots, a_{j-1}, a_k, a_{j+1}, \dots, a_n) \\ &= \det(A) + r \cdot 0 = \det(A). \end{aligned}$$

Zu (c): Betrachte zunächst „ \Leftarrow “: Da die Spalten von $A = (a_1 \mid \dots \mid a_n) \in \mathbb{K}^{n \times n}$ linear abhängig sind, existiert $k \in \{1, \dots, n\}$ derart, dass $a_k = \sum_{j \neq k} r_j a_j$ für $r_j \in \mathbb{K}$, $1 \leq j \leq n$, $j \neq k$, gilt. Es folgt mit der Multilinearität der Determinantenfunktion

$$\begin{aligned} \det(A) &= \det \left(a_1, \dots, a_{k-1}, \sum_{j \neq k} r_j a_j, a_{k+1}, \dots, a_n \right) \\ &= \sum_{j \neq k} r_j \det(a_1, \dots, a_{k-1}, a_j, a_{k+1}, \dots, a_n) \stackrel{\text{nach (a)}}{=} 0. \end{aligned}$$

Für „ \Rightarrow “ wählen wir den Beweis per Kontraposition, d.h. wir zeigen: Sind die Spalten von A linear unabhängig, dann ist $\det(A) \neq 0$. Sei also $A = (a_1 \mid \dots \mid a_n) \in \mathbb{K}^{n \times n}$ mit linear unabhängigen Spalten gegeben. Dann ist A invertierbar und wir setzen $x_k := A^{-1}e_k$ für $k = 1, \dots, n$. Es folgt

$$(3.1) \quad Ax_k = e_k \quad \text{für } 1 \leq k \leq n,$$

wobei die x_k mit dieser Eigenschaft eindeutig bestimmt sind. Dies liefert

$$\begin{aligned}
 1 &= \det(e_1, \dots, e_n) = \det(Ax_1, \dots, Ax_n) \\
 &= \det\left(\sum_{k=1}^n (x_1)_k a_k, \sum_{k=1}^n (x_2)_k a_k, \dots, \sum_{k=1}^n (x_n)_k a_k\right) \\
 &= \sum_{k_1=1}^n (x_1)_{k_1} \det\left(a_{k_1}, \sum_{k=1}^n (x_2)_k a_k, \dots, \sum_{k=1}^n (x_n)_k a_k\right) \\
 &= \sum_{k_1=1}^n \cdots \sum_{k_n=1}^n (x_1)_{k_1} \cdots (x_n)_{k_n} \det(a_{k_1}, \dots, a_{k_n}).
 \end{aligned}$$

Den letzten Faktor können wir nun wie folgt vereinfachen: Zunächst können wir die n -fache Summe reduzieren auf die Summanden, in denen jeweils gilt $k_i \neq k_j$ für $i \neq j$, da sonst $\det(a_{k_1}, \dots, a_{k_n}) = 0$ ist. Schließlich lässt sich der letzte Faktor durch endlich viele Vertauschungen von Spalten in $\det(A)$ überführen. Für k_1, \dots, k_n mit $k_i \neq k_j$ für $i \neq j$ gilt dann

$$\det(a_{k_1}, \dots, a_{k_n}) = \pm \det A.$$

Das liefert

$$(3.2) \quad 1 = \left(\sum_{\substack{k_1, \dots, k_n=1 \\ k_i \neq k_j (i \neq j)}}^n (x_1)_{k_1} \cdots (x_n)_{k_n} \cdot (\text{„}\pm \text{ abhängig von } k_1, \dots, k_n \text{“}) \right) \det(A).$$

Es folgt also $\det(A) \neq 0$. □

Korollar 3.5. Sei $A \in \mathbb{K}^{n \times n}$, dann sind folgende Aussagen äquivalent:

- (a) $\det(A) \neq 0$.
- (b) Die Spalten von A sind linear unabhängig.
- (c) Die Spalten von A bilden eine Basis von \mathbb{K}^n .
- (d) $Ax = b$ hat eine eindeutige Lösung für alle $b \in \mathbb{K}^n$.
- (e) A ist invertierbar.

Beweis. (a) \Leftrightarrow (b) nach Satz 3.4(b) und (b)–(e) sind äquivalent nach Mathe 1. □

Aus der Eindeutigkeit der Darstellung (3.2) im Beweis von Satz 3.4 folgt:

Satz 3.6. Es gibt eine eindeutig bestimmte Determinantenfunktion $\det: \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$. Hat $A \in \mathbb{K}^{n \times n}$ linear abhängige Spalten, dann gilt $\det(A) = 0$. Andernfalls ist $\det(A)$ durch (3.2) gegeben.

Bemerkung 3.7. Die sogenannten Leibnizformel gibt den Wert der Determinante explizit unter Zuhilfenahme von Permutationen der Indizes an. Da die Formel aber für große n kein effizientes Verfahren zur Berechnung von Determinanten darstellt, verzichten wir darauf und verweisen auf die Literatur, z.B. [22, Abschnitt 3.3.3].

Mit Hilfe der Eindeutigkeit der Determinante haben wir folgende Eigenschaft:

Satz 3.8 (Determinantenmultiplikationssatz). *Es seien $A, B \in \mathbb{K}^{n \times n}$. Dann gilt*

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

Beweis. 1. Fall: $\det(B) = 0$. Dann sind die Spalten von B linear abhängig nach Korollar 3.5. Die Darstellung

$$\begin{aligned} B \cdot A &= (b_1 \mid \dots \mid b_n) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \\ &= \left(\sum_{k=1}^n a_{k1} b_k \mid \dots \mid \sum_{k=1}^n a_{kn} b_k \right) \end{aligned}$$

zeigt, dass auch die Spalten von $B \cdot A$ linear abhängig sind. Es folgt also

$$\det(B \cdot A) = 0 = 0 \cdot \det(A) = \det(B) \det(A).$$

2. Fall: $\det(B) \neq 0$. Wir haben

$$\det(BA) = \frac{\det(BA)}{\det(B)} \det(B),$$

müssen also zeigen, dass $\frac{\det(BA)}{\det(B)} = \det(A)$ gilt. Wir werden hierfür die Eindeutigkeit der Determinantenfunktion ausnutzen: Definiere $F(A) := \frac{\det(BA)}{\det(B)}$. Beachte, dass $\det(BA)$ linear von den Spalten von A abhängt. Also ist $F(\cdot)$ multilinear. Vertauschen zweier Spalten in A liefert auch das vertauschen zweier Spalten von $B \cdot A$. Es folgt also, dass $F(\cdot)$ alternierend ist. Schließlich gilt für $A = E_n$: $B \cdot E_n = B$ und damit $F(E_n) = 1$. Satz 3.6 liefert $F(A) = \det(A)$. \square

Korollar 3.9. $A \in \mathbb{K}^{n \times n}$ ist invertierbar genau dann, wenn $\det(A) \neq 0$ ist. Ist dies der Fall, dann gilt

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

Beweis. Die erste Behauptung wurde schon in Korollar 3.5 gezeigt. Ist A invertierbar, dann folgt aus $E_n = A \cdot A^{-1}$ mit Satz 3.8

$$1 = \det(E_n) = \det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1}). \quad \square$$

Korollar 3.10. Die Determinantenfunktion erfüllt die folgenden Eigenschaften für Zeilenoperationen:

- (a) $\det(\cdot)$ ist alternierend in den Zeilen, d.h. bei Vertauschung zweier Zeilen ändert die Determinante nur ihr Vorzeichen.
- (b) Addition eines Vielfachen einer Zeile zu einer anderen Zeile ändert die Determinante nicht.

Beweis. Mit E_{ij} und $E_{ij}(r)$ lassen sich durch Multiplikation von links Zeilenoperationen durchführen. Der Determinantenmultiplikationssatz und $\det(E_{ij}) = -1$ bzw. $\det(E_{ij}(r)) = 1$ liefert die Aussagen (a) und (b). \square

Ende der Sitzung vom 4.5.

Wir wollen nun eine Methode entwickeln, um die Determinante einer Matrix explizit auszurechnen.

Beispiel 3.11. Es seien $a, b, c, d \in \mathbb{K}$ und wir betrachten die Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Dann gilt wegen der Multilinearität in den Spalten und wegen Korollar 3.10:

$$\begin{aligned} \det(A) &= \det \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} + \det \begin{pmatrix} 0 & b \\ c & d \end{pmatrix} \\ &= a \det \begin{pmatrix} 1 & b \\ 0 & d \end{pmatrix} + c \det \begin{pmatrix} 0 & b \\ 1 & d \end{pmatrix} \\ &= a \left(\underbrace{\det \begin{pmatrix} 1 & b \\ 0 & 0 \end{pmatrix}}_{=0} + \det \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix} \right) - c \det \begin{pmatrix} 1 & d \\ 0 & b \end{pmatrix} \\ &= ad - c \left(\underbrace{\det \begin{pmatrix} 1 & d \\ 0 & 0 \end{pmatrix}}_{=0} + \det \begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix} \right) \\ &= ad - bc. \end{aligned}$$

In dieser Rechnung haben wir gesehen, dass wir die Berechnung der Determinante auf einfache Fälle reduzieren können. Um die Methode zu verallgemeinern, benötigen wir noch einige Hilfsmittel.

Lemma 3.12. Es seien $B \in \mathbb{K}^{n \times n}$ und $A \in \mathbb{K}^{(n+1) \times (n+1)}$ durch

$$A = \begin{pmatrix} 1 & * & \cdots & * \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$$

gegeben. Dann gilt $\det(A) = \det(B)$.

Beweis. Wir schreiben $A = (e_1 \mid a_2 \mid \dots \mid a_{n+1})$ und $B = (b_1 \mid \dots \mid b_n)$. Zunächst können wir durch Addition von Vielfachen der ersten Spalte e_1 von A zu den anderen Spalten erreichen, dass die erste Zeile die Form $(1, 0, \dots, 0)$ hat, bei gleich bleibender Determinante. Wir werden nun wieder die Eindeutigkeit der Determinante ausnutzen und definieren dafür eine Abbildung $F: \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$ durch

$$F(B) := \det \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}.$$

Offensichtlich ist F multilinear und alternierend für Spalten aus B . Ferner gilt $F(E_n) = \det(E_{n+1}) = 1$. Damit ist F eine Determinantenfunktion. Satz 3.6 liefert $\det(A) = F(B) = \det(B)$, was zu zeigen war. \square

Definition 3.13. Für $A \in \mathbb{K}^{n \times n}$ und $i, j \in \{1, \dots, n\}$ sei $A_{ij} \in \mathbb{K}^{(n-1) \times (n-1)}$ die Matrix, die aus A durch Weglassen der i -ten Zeile und der j -ten Spalte entsteht. A_{ij} wird i, j -te Minorante von A genannt.

Beispiel 3.14. Ist

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

dann ist zum Beispiel $A_{12} = c$.

Ist

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

dann ist zum Beispiel

$$A_{12} = \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}$$

Erinnerung 3.15. Für $A \in \mathbb{K}^{m \times n}$ ist

$$A^t = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{pmatrix}^t := \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{1n} & \cdots & \cdots & a_{mn} \end{pmatrix} \in \mathbb{K}^{n \times m}$$

die zu A transponierte Matrix. Sie entsteht durch Vertauschung jeder Spalte mit der jeweiligen Zeile. Für $n = m$ entspricht dies gerade der Spiegelung der Einträge entlang der Diagonalen. Hat A die Einträge (a_{ij}) und A^t die Einträge (a_{ij}^t) , dann gilt also $a_{ij}^t = a_{ji}$. Ist $B \in \mathbb{K}^{n,p}$, dann folgt aus der Definition der Matrixmultiplikation sofort $(AB)^t = B^t A^t$.

Proposition 3.16. Für $A \in \mathbb{K}^{n \times n}$ gilt $\det(A^t) = \det(A)$.

Beweis. Übung! □

Bemerkung 3.17. Als Folgerung aus Proposition 3.16 gelten alle Resultate für Spaltenoperationen in Bezug auf die Determinante auch für Zeilenoperationen. Insbesondere ist die Determinantenfunktion bezüglich der Zeilen alternierend multilinear.

Satz 3.18. Für jedes $j \in \{1, \dots, n\}$ und $A \in \mathbb{K}^{n \times n}$ gilt die Laplace'sche-Entwicklung nach der j -ten Spalte:

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

Für jedes $i \in \{1, \dots, n\}$ und $A \in \mathbb{K}^{n \times n}$ gilt die Laplace'sche-Entwicklung nach der i -ten Zeile:

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

Die Terme $(-1)^{k+j} \det(A_{kj})$ werden auch Kofaktoren genannt.

Beweis. Wir schreiben $A = (a_1 | \dots | a_n)$. Es folgt mit den bisher bewiesenen Eigenschaften der Determinantenfunktion und mit Lemma 3.12:

$$\begin{aligned} \det(A) &= \det(a_1, \dots, a_{j-1}, a_j, a_{j+1}, \dots, a_n) \\ &= \det(a_1, \dots, a_{j-1}, \sum_{k=1}^n a_{kj} e_k, a_{j+1}, \dots, a_n) \\ &= \sum_{k=1}^n a_{kj} \det(a_1, \dots, a_{j-1}, e_k, a_{j+1}, \dots, a_n) \\ &= \sum_{k=1}^n a_{kj} (-1)^{j-1} \det(e_k, a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n) \\ &= \sum_{k=1}^n a_{kj} (-1)^{j-1+k-1} \det \begin{pmatrix} 1 & a_{k1} & \dots & a_{k(j-1)} & a_{k(j+1)} & \dots & a_{kn} \\ 0 & & & & & & \\ \vdots & & & A_{kj} & & & \\ 0 & & & & & & \end{pmatrix} \\ &= \sum_{k=1}^n (-1)^{k+j} a_{kj} \det(A_{kj}). \end{aligned}$$

Die Aussage für die Zeilenentwicklung folgt aus der Spaltenentwicklung und aus Proposition 3.16. □

Speziell haben wir:

Korollar 3.19. (a) Für $a, b, c, d \in \mathbb{K}$ gilt: $\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$.

(b) Für $A \in \mathbb{K}^{3 \times 3}$ gilt die sogenannte Regel von Sarrus:

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

Beispiel 3.20. Wir können nun mit den obigen Hilfsmitteln die Determinante der folgenden Matrizen berechnen:

(a)

$$\det \begin{pmatrix} 1 & 2 & 0 & 3 \\ 3 & 2 & 0 & 1 \\ 1 & 2 & 0 & 4 \\ 0 & 0 & 2 & 0 \end{pmatrix} = -2 \cdot \det \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 2 & 4 \end{pmatrix} \\ = -2 \cdot (1 \cdot 2 \cdot 4 + 2 \cdot 1 \cdot 1 + 3 \cdot 3 \cdot 2 - 3 \cdot 2 \cdot 1 \\ - 2 \cdot 3 \cdot 4 - 1 \cdot 1 \cdot 2) \\ = -2 \cdot (28 - 32) = 8.$$

(b)

$$\det \begin{pmatrix} 1 & 0 & 3 & 4 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 4 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{pmatrix} - 3 \cdot \det \begin{pmatrix} 1 & 0 & 3 \\ 1 & 2 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ = (-2 + 4) - 3 \cdot 3 = -7.$$

Satz 3.21 (Cramersche Regel für lineare Gleichungssysteme). Es sei $A \in \mathbb{K}^{n \times n}$ mit $\det(A) \neq 0$. Dann ist die eindeutig bestimmte Lösung von $Ax = b$, $b = (b_1, \dots, b_n)^t \in \mathbb{K}^n$ gegeben durch $x = (x_1, \dots, x_n)^t \in \mathbb{K}^n$ mit

$$x_i = \frac{1}{\det(A)} \det \begin{pmatrix} a_{11} & \dots & a_{1(i-1)} & b_1 & a_{1(i+1)} & \dots & a_{1n} \\ \vdots & & & \vdots & & & \vdots \\ a_{n1} & \dots & a_{n(i-1)} & b_n & a_{n(i+1)} & \dots & a_{nn} \end{pmatrix}, \quad i = 1, \dots, n.$$

Beweis. Nach Korollar 3.5 hat $Ax = b$ eine eindeutige Lösung. Ferner gilt Folgendes: Ist

x eine Lösung von $Ax = b$, so gilt:

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Also gilt für jedes $i \in \{1, \dots, n\}$:

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \dots + \begin{pmatrix} x_i a_{1i} - b_1 \\ \vdots \\ x_i a_{ni} - b_n \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = 0.$$

Es folgt also, dass die Spalten der Matrix

$$M := \begin{pmatrix} a_{11} & \dots & a_{1(i-1)} & (x_i a_{1i} - b_1) & a_{1(i+1)} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n(i-1)} & (x_i a_{ni} - b_n) & a_{n(i+1)} & \dots & a_{nn} \end{pmatrix}$$

linear abhängig sind. Also gilt nach Satz 3.4

$$0 = \det(M) = x_i \det(A) - \det \begin{pmatrix} a_{11} & \dots & a_{1(i-1)} & b_1 & a_{1(i+1)} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n(i-1)} & b_n & a_{n(i+1)} & \dots & a_{nn} \end{pmatrix}.$$

Auflösen nach x_i liefert die Behauptung. □

Satz 3.22 (Matrixinversion). *Sei $A \in \mathbb{K}^{n \times n}$ invertierbar, dann ist*

$$A^{-1} = \frac{1}{\det(A)} B, \text{ wobei}$$

$$B = ((-1)^{i+j} \det(A_{ji})) = \begin{pmatrix} \det(A_{11}) & -\det(A_{21}) & \dots & (-1)^{n+1} \det(A_{n1}) \\ -\det(A_{12}) & \det(A_{22}) & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^{1+n} \det(A_{1n}) & \dots & \dots & \det(A_{nn}) \end{pmatrix}.$$

Beweis. Für eine invertierbare Matrix A schreiben wir $A^{-1} = (x_1 \mid \dots \mid x_n)$. Dann gilt $Ax_j = e_j$ für $j = 1, 2, \dots, n$. Setzen wir jetzt in der Cramerschen Regel $b = e_j$ für die i -te Spalte von A ein und entwickeln die Determinante der so entstandenen Matrix nach der i -ten Spalte, dann folgt die Behauptung. □

Beispiel 3.23. Betrachte $n = 2$ und $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{K}^{2 \times 2}$ mit $\det(A) = ad - bc \neq 0$.

Dann folgt mit Satz 3.22:

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Ende der Sitzung vom 7.5.

3.2 Eigenwerte und Eigenvektoren

Mit Hilfe der Determinante wollen wir nun eines der wichtigsten Werkzeuge zur Charakterisierung von Matrizen entwickeln. Ziel ist es Vektoren zu finden, auf die eine Matrix nur wie eine skalare Multiplikation wirkt. Solche Vektoren werden wir *Eigenvektoren* nennen.

Definition 3.24. Sei $A \in \mathbb{K}^{n \times n}$.

- (a) Wir nennen $\lambda \in \mathbb{K}$ einen *Eigenwert* von A , falls es $v \in \mathbb{K}^n \setminus \{0\}$ gibt mit

$$(3.3) \quad Av = \lambda v.$$

Wir nennen v in diesem Fall *Eigenvektor* von A zum Eigenwert λ . Gleichung (3.3) wird auch als *Eigenwertproblem* bezeichnet.

- (b) Für $\lambda \in \mathbb{K}$ setzen wir $\mathcal{E}_A(\lambda) := \{v \in \mathbb{K}^n \mid Av = \lambda v\}$. Ist λ ein Eigenwert von A , so wird $\mathcal{E}_A(\lambda)$ *Eigenraum von A (zum Wert λ)* genannt.

Bemerkung 3.25. (a) Es kann zu einem Eigenwert mehrere linear unabhängige Eigenvektoren geben. Für alle $\lambda \in \mathbb{K}$ und $A \in \mathbb{K}^{n \times n}$ ist $\mathcal{E}_A(\lambda)$ der Untervektorraum der Lösungen von

$$(3.4) \quad (A - \lambda E_n)v = 0,$$

d.h. es gilt $\mathcal{E}_A(\lambda) = \text{Kern}(A - \lambda E_n)$. Ist λ kein Eigenwert, so gilt $\mathcal{E}_A(\lambda) = \{0\}$.

- (b) Eigenwerte und Eigenvektoren kann man genauso für allgemeine lineare Abbildungen definieren. Ist L eine lineare Abbildung auf einem Funktionenraum (wir nennen L dann auch einen linearen Operator), dann werden Eigenvektoren auch *Eigenfunktionen* von A genannt. Eigenfunktionen sind ein Grundelement um komplizierte Gleichungen zu charakterisieren. In einem Saiteninstrument bestimmt der Eigenwert, welcher Ton gerade zu hören ist, und die zugehörige Eigenfunktion beschreibt die Schwingungsform, die die Saite dann annimmt. In der Quantenmechanik beschreiben Eigenwerte mögliche Messergebnisse von physikalischen Größen (z.B. Drehimpuls, Geschwindigkeit, etc.).

Bei der Untersuchung von Eigenwerten und Eigenvektoren sind wir nur an **nichttrivialen** Lösungen von (3.4) interessiert, im Sinne von $v \neq 0$. Eine solche Lösung existiert

für gegebenes λ genau dann, wenn $\det(A - \lambda E_n) = 0$ gilt. Die Multilinearität der Determinante zeigt, dass $p_A(\lambda) := \det(A - \lambda E_n)$ ein Polynom vom Grad n ist. Die Nullstellen von p_A in \mathbb{K} sind dabei genau die Eigenwerte von A . Diese hängen insbesondere davon ab, ob wir $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ betrachten.

Definition 3.26. Es sei $A \in \mathbb{K}^{n \times n}$ eine Matrix. Die Abbildung $p_A: \mathbb{K} \rightarrow \mathbb{K}$, $p_A(\lambda) := \det(A - \lambda E_n)$, heißt *charakteristisches Polynom* von A .

Beispiel 3.27.

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \in \mathbb{C}^{3 \times 3}, \quad A - \lambda E_3 = \begin{pmatrix} 2 - \lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & -1 & -\lambda \end{pmatrix}.$$

Dann ist das charakteristische Polynom von A gegeben durch $p_A(\lambda) = \det(A - \lambda E_3) = (2 - \lambda)(\lambda^2 + 1)$. Die Nullstellen in \mathbb{C} sind $\lambda_1 = 2$, $\lambda_2 = i$ und $\lambda_3 = -i$. Lösen der entsprechenden Gleichungen (3.4) liefert als Eigenvektoren z.B. $v_1 = e_1$, $v_2 = (0, 1, i)^t$ und $v_3 = (0, i, 1)^t$ sowie $\mathcal{E}(\lambda_j) = \text{Spann}(v_j)$ für $j = 1, 2, 3$. Für $\mathbb{K} = \mathbb{R}$ erhalten wir nur den Eigenwert 2 und nur einen Eigenraum $\mathcal{E}(2)$.

Satz 3.28. Seien $A \in \mathbb{K}^{n \times n}$ und $m \in \mathbb{N}$. Ferner seien $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ (paarweise) verschiedenen Eigenwerte von A mit zugehörigen Eigenvektoren $v_1, \dots, v_m \in \mathbb{K}^n$. Dann sind die v_1, \dots, v_m linear unabhängig.

Beweis. Wir beweisen diese Aussage per Induktion.

Induktionsanfang $m = 1$: Nach Definition ist ein Eigenvektor nicht der Nullvektor. Es folgt die lineare Unabhängigkeit.

Induktionsbehauptung: Für ein $m \in \mathbb{N}$ gilt die Aussage.

Induktionsschritt $m \mapsto m + 1$: Es seien $\lambda_1, \dots, \lambda_m, \lambda_{m+1}$ paarweise verschiedene Eigenwerte von A mit zugehörigen Eigenvektoren v_1, \dots, v_m, v_{m+1} . Nach der Induktionsbehauptung folgt, dass die v_1, \dots, v_m linear unabhängig sind. Seien nun $c_1, \dots, c_m, c_{m+1} \in \mathbb{K}$ mit

$$(3.5) \quad \sum_{i=1}^{m+1} c_i v_i = 0$$

gegeben. Anwenden der Matrix A auf beiden Seiten liefert:

$$0 = A \cdot 0 = A \left(\sum_{i=1}^{m+1} c_i v_i \right) = \sum_{i=1}^{m+1} A(c_i v_i) = \sum_{i=1}^{m+1} c_i \lambda_i v_i.$$

Da auch

$$\lambda_{m+1} \sum_{i=1}^{m+1} c_i v_i = 0$$

gilt, folgt

$$\begin{aligned} 0 &= \lambda_{m+1} \sum_{i=1}^{m+1} c_i v_i - \sum_{i=1}^{m+1} c_i \lambda_i v_i = \sum_{i=1}^{m+1} (\lambda_{m+1} - \lambda_i) c_i v_i \\ &= \sum_{i=1}^m (\lambda_{m+1} - \lambda_i) c_i v_i. \end{aligned}$$

Da v_1, \dots, v_m nach Induktionsbehauptung linear unabhängig sind und $\lambda_{m+1} - \lambda_i \neq 0$ ist, folgt $c_1 = c_2 = \dots = c_m = 0$. Aus Gleichung (3.5) folgt nun $c_{m+1} v_{m+1} = 0$, also $c_{m+1} = 0$. Da die Koeffizienten c_1, \dots, c_{m+1} beliebig waren, sind die v_1, \dots, v_{m+1} linear unabhängig. \square

Korollar 3.29. *Es sei $A \in \mathbb{K}^{n \times n}$ eine Matrix mit n (paarweise) verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ und zugehörigen Eigenvektoren $v_1, \dots, v_n \in \mathbb{K}^n$. Dann bildet $\{v_1, \dots, v_n\}$ eine Basis von \mathbb{K}^n und es gilt demnach*

$$(3.6) \quad \mathbb{K}^n = \mathcal{E}_A(\lambda_1) \oplus \mathcal{E}_A(\lambda_2) \oplus \dots \oplus \mathcal{E}_A(\lambda_n).$$

3.2.1 Diagonalisierung von Matrizen

Eine Matrix $A = (a_{ij}) \in \mathbb{K}^{n \times n}$ mit $a_{ij} = 0$ für $i \neq j$ heißt *Diagonalmatrix*. Für diese verwenden wir die Notation $A = \text{Diag}(d_1, d_2, \dots, d_n)$ mit $d_k := a_{kk} \in \mathbb{K}$. Die Struktur von Eigenwerten und Eigenvektoren ist für eine solche Diagonalmatrix besonders einfach: Das charakteristische Polynom ist $p_A(\lambda) = (d_1 - \lambda) \cdot \dots \cdot (d_n - \lambda)$, d.h. die d_k sind genau die Eigenwerte, und es existieren n linear unabhängige zugehörige Eigenvektoren, nämlich die Standardbasis $\{e_1, \dots, e_n\}$. Hier können Eigenwerte mehrfach auftreten, also mehrfache Nullstellen von p_A sein.

Erinnerung 3.30. Wir identifizieren Matrizen $A \in \mathbb{K}^{n \times n}$ mit linearen Abbildungen $L_A \in \text{Hom}(\mathbb{K}^n)$ (meist schreiben wir einfach A statt L_A). Allgemeiner wissen wir aus Mathe 1, dass für einen n -dimensionalen \mathbb{K} -Vektorraum V mit einer Basis $B = \{v_1, \dots, v_n\}$ eine Korrespondenz zwischen \mathbb{K} -linearen Abbildungen $L \in \text{Hom}(V)$ und Matrizen $A \in \mathbb{K}^{n \times n}$ besteht: Sei $\Psi_B \in \text{Hom}(\mathbb{K}^n, V)$ der durch $\Psi_B(e_k) = v_k$, $k = 1, \dots, n$, gegebene Isomorphismus. Dann ist $A = (a_{ij})$ die Darstellungsmatrix von L bezüglich B genau dann, wenn $a_{ij} = e_i^t A e_j = e_i^t \Psi_B^{-1} L \Psi_B e_j = (\Psi_B^{-1} L \Psi_B)_{ij}$ gilt. Dies liest man aus folgendem kommutativen Diagramm ab:

$$\begin{array}{ccc} V & \xrightarrow{L} & V \\ \Psi_B \uparrow & & \uparrow \Psi_B \\ \mathbb{K}^n & \xrightarrow{A} & \mathbb{K}^n \end{array}$$

Die Darstellungsmatrix A von L hängt also von der Wahl der Basis B ab.

Im Falle von $V = \mathbb{K}^n$ identifizieren wir Vektoren $v \in V$ mit den Spalten ihrer Koordinaten bezüglich der kanonischen Basis. Zu einer beliebigen **geordneten Basis**

$B = \{v_1, \dots, v_n\}$ (die nicht die kanonische Basis $\{e_1, \dots, e_n\}$ zu sein braucht) formen wir die **Matrix** $B := (v_1 \mid \dots \mid v_n)$. Mit den Bezeichnungen von oben ist dann $\Psi_B = B$ und A hat bezüglich der Basis B die Darstellungsmatrix $B^{-1}AB$.

Definition 3.31. Zwei Matrizen $A, \tilde{A} \in \mathbb{K}^{n \times n}$ heißen *ähnlich* (kurz $A \sim \tilde{A}$), wenn eine invertierbare Matrix $B \in \mathbb{K}^{n \times n}$ existiert, so dass

$$\tilde{A} = B^{-1}AB$$

gilt. Ähnlichkeit ist eine Äquivalenzrelation auf $\mathbb{K}^{n \times n}$.

Bemerkung 3.32. Offensichtlich stimmen die charakteristischen Polynome von ähnlichen Matrizen überein, denn

$$\begin{aligned} \det((B^{-1}AB) - \lambda E_n) &= \det(B^{-1}(A - \lambda E_n)B) \\ &= \det(B^{-1}) \det(A - \lambda E_n) \det(B) = \det(A - \lambda E_n). \end{aligned}$$

Somit haben ähnliche Matrizen auch dieselben Eigenwerte.

Ein Basiswechsel liefert zu einer gegebenen Matrix $A \in \mathbb{K}^{n \times n}$ eine ähnliche Matrix als Darstellungsmatrix von L_A bezüglich der neuen Basis. Wir wollen Basiswechsel untersuchen, die möglichst einfache neue Darstellungsmatrizen liefern. Als ersten Fall betrachten wir:

Definition 3.33. Eine Matrix $A \in \mathbb{K}^{n \times n}$ heißt *diagonalisierbar*, wenn sie einer Diagonalmatrix ähnlich ist.

Satz 3.34. Eine Matrix $A \in \mathbb{K}^{n \times n}$ ist genau dann diagonalisierbar, wenn \mathbb{K}^n eine Basis $B = (v_1 \mid \dots \mid v_n)$ aus Eigenvektoren zu den Eigenwerten $\lambda_1, \dots, \lambda_n$ (mit Wiederholungen aufgezählt) von A besitzt. In diesem Fall ist $B^{-1}AB = \text{Diag}(\lambda_1, \dots, \lambda_n)$.

Beweis. Wenn A diagonalisierbar ist, also $B^{-1}AB = \text{Diag}(\lambda_1, \dots, \lambda_n)$ mit einer invertierbaren Matrix B und $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ gilt, dann erfüllt für jedes $k = 1, \dots, n$ der Vektor $v_k := Be_k$ (die k -te Spalte von B):

$$Av_k = BB^{-1}ABe_k = B\text{Diag}(\lambda_1, \dots, \lambda_n)e_k = B\lambda_k e_k = \lambda_k v_k.$$

Somit ist (λ_k, v_k) ein Eigenwert-Eigenvektor-Paar von A . Da B invertierbar ist, ist $\{v_1, \dots, v_n\}$ linear unabhängig, also eine Basis von \mathbb{K}^n .

Hat A umgekehrt n Eigenwert-Eigenvektor-Paare (λ_k, v_k) , so dass $B = \{v_1, \dots, v_n\}$ eine Basis von \mathbb{K}^n ist, dann ist $B := (v_1 \mid \dots \mid v_n)$ invertierbar und es gilt:

$$(B^{-1}AB)_{ij} = e_i^t B^{-1}ABe_j = e_i^t B^{-1}Av_j = \lambda_j e_i^t B^{-1}v_j = \lambda_j e_i^t e_j = \lambda_j \delta_{ij},$$

also $B^{-1}AB = \text{Diag}(\lambda_1, \dots, \lambda_n)$. □

Beispiel 3.35. Betrachte $A := \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$. Dann ist

$$p_A(\lambda) = \det \begin{pmatrix} 3-\lambda & 2 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 - 3\lambda + 2 = (\lambda - 1)(\lambda - 2).$$

Also sind $\lambda_1 = 1$ und $\lambda_2 = 2$ die Eigenwerte von A . Wir suchen nun Eigenvektoren v zum Eigenwert 1. Diese lösen

$$\begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow v_1 = -v_2.$$

Es folgt

$$\mathcal{E}_A(1) = \left\{ r \begin{pmatrix} 1 \\ -1 \end{pmatrix} \mid r \in \mathbb{R} \right\}.$$

Wir suchen nun Eigenvektoren v zum Eigenwert 2. Diese lösen

$$\begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow v_1 = -2v_2.$$

Es folgt

$$\mathcal{E}_A(2) = \left\{ r \begin{pmatrix} -2 \\ 1 \end{pmatrix} \mid r \in \mathbb{R} \right\}.$$

Wir wählen zu 1 den Eigenvektor $(1, -1)^t$ und zu 2 den Eigenvektor $(-2, 1)^t$, dann ist

$$\begin{aligned} \begin{pmatrix} 1 & -2 \\ -1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ -1 & 1 \end{pmatrix} &= - \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -4 \\ -1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}. \end{aligned}$$

Korollar 3.36. Sei $A \in \mathbb{K}^{n \times n}$ eine Matrix mit n (paarweise) verschiedenen Eigenwerten. Dann ist A diagonalisierbar.

Beweis. Dies folgt aus Korollar 3.29 und Satz 3.34. □

Bemerkung 3.37. Die Existenz von n verschiedenen Eigenwerten ist für die Diagonalisierbarkeit nicht notwendig: Z.B. ist E_n für jedes n von diagonalärer Gestalt (also diagonalisierbar) und hat den n -fachen Eigenwert 1 mit den Eigenvektoren e_1, \dots, e_n .

3.2.2 Normalform von Matrizen

Nicht alle Matrizen sind diagonalisierbar:

Beispiel 3.38. Die Matrix $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ hat das charakteristische Polynom $p_A(\lambda) = \lambda^2$ mit der einzigen (doppelten) Nullstelle 0, dem einzigen Eigenwert. Ein Eigenvektor zu diesem

Eigenwert ist $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Für jeden Eigenvektor $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{K}^2$ gilt

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ 0 \end{pmatrix},$$

also $x_2 = 0$. Damit ist $\mathcal{E}_A(0) = \text{Spann}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$ und es existiert keine Basis von \mathbb{K}^2 aus Eigenvektoren von A . Nach Satz 3.34 ist A also nicht diagonalisierbar.

Wir wollen auch für eine allgemeine Matrix $A \in \mathbb{K}^{n \times n}$ eine möglichst einfache Darstellungsmatrix bezüglich einer geeigneten Basis finden.

Definition 3.39. Es sei $A \in \mathbb{C}^{n \times n}$ und λ ein Eigenwert von A . Wir bezeichnen mit der *algebraischen Vielfachheit von λ* die Vielfachheit von λ als Nullstelle des charakteristischen Polynoms p_A . Wir bezeichnen mit der *geometrischen Vielfachheit von λ* die Anzahl der linear unabhängigen Eigenvektoren zum Eigenwert λ , also die Zahl $\dim \mathcal{E}_A(\lambda)$.

Bemerkung 3.40. (a) Für einen Eigenwert λ ist stets $\dim \mathcal{E}_A(\lambda) \geq 1$.

(b) Man kann zeigen, dass die geometrische Vielfachheit eines Eigenwertes immer kleiner oder gleich der algebraischen Vielfachheit ist.

(c) Für diagonalisierbare Matrizen stimmt die geometrische Vielfachheit eines Eigenwertes immer mit der algebraischen Vielfachheit überein.

Beispiel 3.41. Für $A = E_2$ ist 1 ein Eigenwert mit algebraischer Vielfachheit 2 und geometrischer Vielfachheit 2. Für $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ist 0 ein Eigenwert mit algebraischer Vielfachheit 2 und geometrischer Vielfachheit 1.

Ende der Sitzung vom 11.5.

Wir betrachten zunächst nur Matrizen über \mathbb{C} , da in diesem Fall das charakteristische Polynom nach dem Hauptsatz der Algebra immer in Linearfaktoren zerfällt.

Ist λ ein Eigenwert von $A \in \mathbb{C}^{n \times n}$ der algebraischen Vielfachheit α , dann haben wir für $G := (A - \lambda E_n)$ eine aufsteigende Folge von Untervektorräumen

$$(3.7) \quad \text{Kern } G \subseteq \text{Kern } G^2 \subseteq \text{Kern } G^3 \subseteq \dots$$

Aus dem Satz von Cayley-Hamilton, den wir hier nicht beweisen wollen, folgt $\text{Kern } G^k = \text{Kern } G^{k+1}$ für alle $k \geq \alpha$ und $\dim \text{Kern } G^\alpha = \alpha$. Ferner gilt für $k \in \mathbb{N}$:

$$(3.8) \quad G(\text{Kern } G^{k+1} \setminus \text{Kern } G^k) \subseteq \text{Kern } G^k \setminus \text{Kern } G^{k-1}.$$

Definition und Satz 3.42. In der obigen Situation ist die Zahl

$$\eta := \min\{k \in \mathbb{N} \mid \text{Kern}(A - \lambda E_n)^k = \text{Kern}(A - \lambda E_n)^{k+1}\}$$

wohldefiniert und erfüllt $\eta \leq \alpha$. Der Untervektorraum

$$\mathcal{H}_A(\lambda) := \text{Kern}(A - \lambda E_n)^\alpha = \text{Kern}(A - \lambda E_n)^\eta$$

heißt *Hauptraum* (verallgemeinerter *Eigenraum*) von λ , und η heißt *Index des Hauptraumes* von λ . Es gilt $\dim \mathcal{H}_A(\lambda) = \alpha$. Ein Vektor $v \in \text{Kern}(A - \lambda E_n)^k \setminus \text{Kern}(A - \lambda E_n)^{k-1}$ heißt *Nebeneigenvektor k -ter Stufe* von λ . Die Nebeneigenvektoren 1-ter Stufe sind also die Eigenvektoren.

Beweis. [12, Abschnitte 4.5 und 4.6]. □

Satz und Definition 3.43 (Jordanblock). *Die Matrix $A \in \mathbb{C}^{n \times n}$ besitze nur einen Eigenwert λ , und dieser habe die geometrische Vielfachheit 1. Dann ist A ähnlich zu der Matrix*

$$J_n(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} \in \mathbb{C}^{n \times n}.$$

Diese nennen wir Jordanblock der Länge n zum Eigenwert λ .

Beweisskizze. Für die algebraische Vielfachheit von λ gilt $\alpha = n$, d.h. es muss $\mathcal{H}_A(\lambda) = \mathbb{C}^n$ gelten. Sei $v = v^{(1)}$ ein Eigenvektor zu λ , und sei $v^{(2)}$ eine Lösung von $(A - \lambda E_n)v^{(2)} = v$. Dann ist $v^{(2)}$ ein Nebeneigenvektor 2-ter Stufe zu λ . Induktiv sei $v^{(k)}$ für $k = 2, \dots, n$ jeweils eine Lösung von $(A - \lambda E_n)v^{(k)} = v^{(k-1)}$, also ein Nebeneigenvektor k -ter Stufe zu λ . Man kann zeigen, dass diese Lösungen jeweils existieren. Die geordnete Basis

$$B = (v \mid v^{(2)} \mid \dots \mid v^{(n)})$$

liefert dann $B^{-1}AB = J_n(\lambda)$. □

Satz und Definition 3.44 (Jordanmatrix). *Die Matrix $A \in \mathbb{C}^{n \times n}$ besitze nur einen Eigenwert λ , und dieser habe die geometrische Vielfachheit γ . Ferner sei η der Index von $\mathcal{H}(\lambda) = \mathbb{C}^n$. Dann existieren Zahlen $s_1 \leq s_2 \leq \dots \leq s_\gamma \in \mathbb{N}$, $s_\gamma = \eta$, mit $\sum_{k=1}^\gamma s_k = n$, so dass A zu einer Blockmatrix*

$$J_n^*(\lambda) = \begin{pmatrix} J_{s_\gamma}(\lambda) & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & J_{s_{\gamma-1}}(\lambda) & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & J_{s_1}(\lambda) \end{pmatrix}$$

ähnlich ist. Eine solche Matrix nennen wir (geordnete) Jordanmatrix der Länge n zum Eigenwert λ .

Beweis. Wir setzen $G := A - \lambda E_N$. Seien γ die geometrische und α die algebraische Vielfachheit von λ . Ferner sei η der Index von $\mathcal{H}_A(\lambda) = \text{Kern } G^\eta$. Wir setzen weiter

$U_k := \text{Kern } G^k$, $k \in \mathbb{N}_0$. Dann haben wir wegen (3.7) und (3.8) die aufsteigende Folge

$$\{0\} = U_0 \subsetneq U_1 \subsetneq U_2 \subsetneq \cdots \subsetneq U_\eta = \mathcal{H}_A(\lambda)$$

und die Inklusionen

$$G(U_k \setminus U_{k-1}) \subseteq U_{k-1} \setminus U_{k-2}$$

für $k \geq 2$. Wir werden „Ergänzungsräume“ V_k mit Dimension $q_k := \dim U_k - \dim U_{k-1}$ mit den folgenden Eigenschaften konstruieren:

$$\begin{aligned} U_\eta &= U_{\eta-1} \oplus V_\eta \\ U_{\eta-1} &= U_{\eta-2} \oplus V_{\eta-1} \\ &\vdots \\ U_2 &= U_1 \oplus V_2 \\ U_1 &= V_1 = \mathcal{E}_A(\lambda) \end{aligned}$$

und

$$G(V_k) \subseteq V_{k-1} \quad \text{für } k = 2, \dots, \eta.$$

Nach Definition des Index η gilt hierbei $q_k > 0$ für $k = 2, \dots, \eta$. Die entsprechenden Elemente von $V_k \setminus \{0\}$ sind dabei Nebeneigenvektoren k -ter Stufe (auch Hauptvektoren genannt).

Rezept:

- Bestimme Dimensionen und Basen der Räume U_k , $k = 1, 2, \dots, \eta$ (durch Gauß-Eliminierung);
- wähle (beliebige) linear unabhängige Nebeneigenvektoren η -ter Stufe $v_1^{(\eta)}, v_2^{(\eta)}, \dots, v_{q_\eta}^{(\eta)} \in U_\eta \setminus U_{\eta-1}$, so dass für $V_\eta := \text{Spann}(v_1^{(\eta)}, v_2^{(\eta)}, \dots, v_{q_\eta}^{(\eta)})$ gilt: $V_\eta \cap U_{\eta-1} = \{0\}$;
- definiere $v_j^{(\eta-1)} := Gv_j^{(\eta)}$ für $j = 1, 2, \dots, q_\eta$; dann ist die Menge $\{v_1^{(\eta-1)}, v_2^{(\eta-1)}, \dots, v_{q_\eta}^{(\eta-1)}\} \subseteq U_{\eta-1} \setminus U_{\eta-2}$ linear unabhängig (Übung);
- ergänze diese zu einer linear unabhängigen Menge $\{v_1^{(\eta-1)}, v_2^{(\eta-1)}, \dots, v_{q_{\eta-1}}^{(\eta-1)}\} \subseteq U_{\eta-1} \setminus U_{\eta-2}$, so dass für $V_{\eta-1} := \text{Spann}(v_1^{(\eta-1)}, v_2^{(\eta-1)}, \dots, v_{q_{\eta-1}}^{(\eta-1)})$ gilt: $V_{\eta-1} \cap U_{\eta-2} = \{0\}$;
- wiederhole diesen Prozess, bis im letzten Schritt die Eigenvektoren $v_j^{(1)} = Gv_j^{(2)}$, $j = 1, 2, \dots, q_2$ festgelegt und zu einer Basis $\{v_1^{(1)}, v_2^{(1)}, \dots, v_\gamma^{(1)}\}$ von $\mathcal{E}_A(\lambda)$ ergänzt wurden;
- ordne die Basis von $\mathcal{H}_A(\lambda)$ wie folgt an:

$$v_1^{(1)}, v_1^{(2)}, \dots, v_1^{(\eta)}, v_2^{(1)}, v_2^{(2)}, \dots, v_3^{(1)}, v_3^{(2)}, \dots, v_\gamma^{(1)}, v_\gamma^{(2)}, \dots$$

Jedem Eigenvektor folgen die entsprechenden Nebeneigenvektoren. Jede solche Gruppe entspricht einem Jordanblock.

- Die Anzahl der Jordanblöcke der Länge r ist

$$(3.9) \quad q_r - q_{r+1} = \dim U_r - \dim U_{r-1} - (\dim U_{r+1} - \dim U_r) \\ = 2 \dim U_r - \dim U_{r-1} - \dim U_{r+1}.$$

- Wenn A reell ist und nur reelle Eigenwerte hat, dann können wir oben \mathbb{C} durch \mathbb{R} ersetzen, da alle Nebeneigenvektoren mit reellen Koordinaten gewählt werden können.

Erklärung:

- Die zu einem Jordanblock $J_r(\lambda)$ gehörende Kette von Basisvektoren endet mit einem Vektor $v_j^{(r)}$, welcher nicht als Bild $Gv_j^{(r+1)}$ erhalten wurde, sondern im Basisergänzungsschritt für V_r . Das erklärt, warum die Anzahl dieser Blöcke $q_r - q_{r+1}$ ist.
- Die Länge der Jordanblöcke ist in obigem Schema automatisch abnehmend.
- Die Jordansche Normalform kann man also auch bestimmen, ohne die Nebeneigenvektoren festzulegen, sobald man die Dimensionen der U_k kennt. \square

Bemerkung 3.45. (a) Unter der Maßgabe, die Jordanblöcke in absteigender Länge zu ordnen (*geordnete Jordanmatrix*), ist die Jordanmatrix zu einem Eigenwert eindeutig bestimmt, aber die zugehörige Basis aus Nebeneigenvektoren ist es nicht.

- (b) Eine geordnete Jordanmatrix der Länge n zum Eigenwert λ hat in der Diagonalen nur λ , in der oberen Nebendiagonalen nur die Werte 0 und 1, und an allen anderen Stellen nur 0 als Einträge. Die Länge der 1-Ketten in der oberen Nebendiagonalen nimmt dabei nach rechts unten ab. Achtung: Dort, wo zwei Jordanblöcke aneinanderstoßen hat die obere Nebendiagonale eine 0 als Eintrag! Es ist eine Frage der Kombinatorik, wie viele verschiedene geordnete Jordanmatrizen prinzipiell in gegebener Dimension zu einem Eigenwert existieren können.

Satz 3.46 (Jordansche Normalform). *Die Matrix $A \in \mathbb{C}^{n \times n}$ habe m Eigenwerte λ_k mit algebraischen Vielfachheiten α_k , $k = 1, 2, \dots, m$, d.h. $\sum_{k=1}^m \alpha_k = n$. Dann ist A zu einer Blockmatrix*

$$\begin{pmatrix} J_{\alpha_1}^*(\lambda_1) & & & 0 \\ & J_{\alpha_2}^*(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{\alpha_m}^*(\lambda_m) \end{pmatrix}$$

ähnlich, wobei die $J_{\alpha_k}^(\lambda_k)$ Jordanmatrizen sind.*

Beweisskizze. Für jedes k ist der Hauptraum $\mathcal{H}_A(\lambda_k)$ für A *invariant*, d.h. $A(\mathcal{H}_A(\lambda_k)) \subseteq \mathcal{H}_A(\lambda_k)$. Dies folgt daraus, dass A mit jeder Potenz von $(A - \lambda E_n)$ kommutiert, wie man leicht induktiv zeigen kann. Man beweist nun $\mathcal{H}_{\alpha_1}(\lambda_1) \oplus \cdots \oplus \mathcal{H}_{\alpha_m}(\lambda_m) = \mathbb{C}^n$, die *Hauptraumzerlegung*, und macht zunächst einen Basiswechsel, so dass eine zu A ähnliche diagonale Blockmatrix entsteht, deren m Blöcke jeweils nur einen der Eigenwerte λ_k besitzen. Danach wendet man Satz und Definition 3.44 auf diese Blöcke in den Haupträumen gesondert an. \square

Beispiel 3.47. Wir bestimmen die Jordansche Normalform J und eine Basismatrix B von

$$A := \begin{pmatrix} 2 & 1 & 0 & -2 \\ -\frac{1}{2} & 1 & \frac{1}{2} & 2 \\ -1 & -1 & 3 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

A hat nur den Eigenwert $\lambda = 2$ mit algebraischer Vielfachheit 4, wir können also rein reell rechnen. Wir setzen

$$G := A - 2E_4 = \begin{pmatrix} 0 & 1 & 0 & -2 \\ -\frac{1}{2} & -1 & \frac{1}{2} & 2 \\ -1 & -1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Es folgt

$$G^2 = \begin{pmatrix} -\frac{1}{2} & -1 & \frac{1}{2} & 2 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{2} & -1 & \frac{1}{2} & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad G^3 = 0.$$

Demnach haben wir $\eta = 3$, also $U_3 = \mathbb{R}^4$. Wir bestimmen eine Basis für U_2 . Nach einem Schritt der Gauß-Eliminierung erhalten wir aus G^2 die Matrix

$$E_{31}(-1)G^2 = \begin{pmatrix} -\frac{1}{2} & -1 & \frac{1}{2} & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Demnach hat G^2 Rang 1 und es gilt $\dim U_2 = 3$. Ferner ist

$$U_2 = \text{Kern } G^2 = \text{Kern } E_{31}(-1)G^2 = \text{Spann} \left(\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right).$$

Bestimmung der Eigenvektoren liefert $\dim U_1 = 2$ und

$$U_1 = \mathcal{E}_A(2) = \text{Spann} \left(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} \right).$$

An dieser Stelle können wir schon sagen: J hat $2 \dim U_3 - \dim U_4 - \dim U_2 = 8 - 4 - 3 = 1$ Jordanblöcke $J_3(2)$ der Länge 3, $2 \dim U_2 - \dim U_3 - \dim U_1 = 6 - 4 - 2 = 0$ Jordanblöcke $J_2(2)$ der Länge 2 und $2 \dim U_1 - \dim U_2 - \dim U_0 = 4 - 3 - 0 = 1$ Jordanblöcke $J_1(2)$ der Länge 1.

Wir wählen $v_1^{(3)} = (0, 0, 0, 1)^t \in U_3 \setminus U_2$ und setzen $V_3 := \text{Spann}(v_1^{(3)})$. Es ergibt sich $v_1^{(2)} = Gv_1^{(3)} = (-2, 2, 2, 0)^t \in U_2 \setminus U_1$ und $V_2 := \text{Spann}(v_1^{(2)})$. Es ergibt sich $v_1^{(1)} = Gv_1^{(2)} = (2, 0, 2, 0)^t$. Wir ergänzen diesen noch durch die Wahl von $v_2^{(1)} := (0, 2, 0, 1)^t$ zu einer Basis von $\mathcal{E}_A(2)$ und erhalten die Basismatrix

$$B := (v_1^{(1)}, v_1^{(2)}, v_1^{(3)}, v_2^{(1)}) = \begin{pmatrix} 2 & -2 & 0 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Es ergibt sich die Normalform

$$J = B^{-1}AB = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Bemerkung 3.48. (a) Die Jordansche Normalform einer diagonalisierbaren Matrix ist eine Diagonalmatrix.

(b) Aus der Jordanschen Normalform folgt, dass für $A \in \mathbb{C}^{n \times n}$ stets $\det(A) = \prod_{k=1}^n \lambda_k$ gilt, wobei die Eigenwerte λ_i von A entsprechend ihrer algebraischen Vielfachheit mehrfach aufgezählt werden.

3.3 Spezielle Matrizen

In den folgenden Abschnitten wollen wir Matrizen untersuchen, die bestimmte geometrische Eigenschaften haben. Um diese zu beschreiben, ist es nützlich, ein bestimmtes Skalarprodukt in \mathbb{K}^n zu verwenden:

Definition 3.49. Auf \mathbb{R}^n bzw. \mathbb{C}^n verwenden wir das sogenannte *Standardskalarprodukt*:

$$\langle u, v \rangle := \bar{u}^t v = \begin{cases} \sum_{i=1}^n u_i v_i & u = (u_1, \dots, u_n) \in \mathbb{R}^n, v = (v_1, \dots, v_n) \in \mathbb{R}^n; \\ \sum_{i=1}^n \bar{u}_i v_i & u = (u_1, \dots, u_n) \in \mathbb{C}^n, v = (v_1, \dots, v_n) \in \mathbb{C}^n. \end{cases}$$

Das Standardskalarprodukt induziert die *euklidische Norm*

$$|u| := \sqrt{\langle u, u \rangle} = \sqrt{\sum_{i=1}^n |u_i|^2}$$

auf \mathbb{R}^n bzw. \mathbb{C}^n . Mit $u \perp v$ drücken wir aus, dass u und v orthogonal sind, also $\langle u, v \rangle = 0$ erfüllen.

Lemma 3.50. Seien $(V, \langle \cdot, \cdot \rangle)$ ein \mathbb{K} -Skalarproduktraum und $b_1, \dots, b_m \in V \setminus \{0\}$ paarweise orthogonale Vektoren. Dann sind die b_1, \dots, b_m linear unabhängig.

Beweis. Seien $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ mit

$$\sum_{k=1}^m \lambda_k b_k = 0$$

gegeben. Dann folgt für jedes $j \in \{1, \dots, m\}$:

$$0 = \langle b_j, 0 \rangle = \left\langle b_j, \sum_{k=1}^m \lambda_k b_k \right\rangle = \sum_{k=1}^m \lambda_k \langle b_j, b_k \rangle = \lambda_j \langle b_j, b_j \rangle.$$

Da $b_j \neq 0$ ist, folgt $\lambda_j = 0$ für alle $j \in \{1, \dots, m\}$. Also sind die b_1, \dots, b_m linear unabhängig. \square

Bemerkung 3.51. Ist $\dim(V) = n$ und sind die Vektoren b_1, \dots, b_n paarweise orthogonal, so folgt, dass b_1, \dots, b_n eine Orthogonalbasis von V ist.

Bemerkung 3.52. Sei $v \in \mathbb{K}^n \setminus \{0\}$ ein Vektor mit $|v| = 1$. Dann beschreibt die lineare Abbildung

$$(3.10) \quad P_v: \mathbb{K}^n \rightarrow \mathbb{K}^n, \quad u \mapsto \langle v, u \rangle v$$

die Orthogonalprojektion eines Vektors u auf den Unterraum $\text{Spann}(v)$: Es gilt $u = (u - P_v u) + P_v u$, $(u - P_v u) \perp P_v u$ und $P_v u \in \text{Spann}(v)$ (Übung).

Mit dem folgenden Verfahren lässt sich aus einer beliebigen Basis von \mathbb{K}^n eine Orthonormalbasis (ONB) gewinnen:

Satz 3.53 (Gram-Schmidtsches Orthogonalisierungsverfahren). Sei V ein endlich dimensionaler \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Dann hat V eine ONB.

Beweis. Sei $\{b_1, \dots, b_n\}$ eine Basis von V . Wir setzen

$$\begin{aligned} c_1 &:= \frac{1}{|b_1|} b_1 \\ c_2 &:= \frac{1}{|b_2 - P_{c_1} b_2|} (b_2 - P_{c_1} b_2), \\ &\vdots \\ c_k &:= \frac{1}{|b_k - \sum_{j=1}^{k-1} P_{c_j} b_k|} \left(b_k - \sum_{j=1}^{k-1} P_{c_j} b_k \right), \quad k = 2, \dots, n. \end{aligned}$$

Nach Konstruktion sind die c_i normiert und paarweise orthogonal (Übung). □

3.3.1 Orthogonale und unitäre Matrizen

Definition 3.54. (a) Eine Matrix $U \in \mathbb{R}^{n \times n}$ heißt *orthogonal*, falls für alle $v, w \in \mathbb{R}^n$ gilt: $\langle Uv, Uw \rangle = \langle v, w \rangle$.

(b) Eine Matrix $U \in \mathbb{C}^{n \times n}$ heißt *unitär*, falls für alle $v, w \in \mathbb{C}^n$ gilt: $\langle Uv, Uw \rangle = \langle v, w \rangle$.

Bemerkung 3.55. Orthogonale bzw. unitäre Matrizen werden auch *winkelerhaltend* oder *normerhaltend* genannt. Sie sind sogenannte *Isometrien* im \mathbb{K}^n bezüglich des Standardskalarprodukts.

Satz 3.56. Es sei $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} . Dann ist $A \in \mathbb{K}^{n \times n}$ genau dann orthogonal bzw. unitär, wenn die Spalten von A eine ONB in \mathbb{K}^n bilden.

Beweis. Seien a_1, \dots, a_n die Spalten von A .

„ \Leftarrow “: Sei $\{a_1, \dots, a_n\}$ eine ONB von \mathbb{K}^n und seien $v = (v_1, \dots, v_n)^t$ und $w = (w_1, \dots, w_n)^t$ zwei Vektoren in \mathbb{K}^n . Es folgt

$$\begin{aligned} \langle Av, Aw \rangle &= \left\langle A \left(\sum_{i=1}^n v_i e_i \right), A \left(\sum_{j=1}^n w_j e_j \right) \right\rangle = \left\langle \sum_{i=1}^n v_i A e_i, \sum_{j=1}^n w_j A e_j \right\rangle \\ &= \left\langle \sum_{i=1}^n v_i a_i, \sum_{j=1}^n w_j a_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \bar{v}_i w_j \langle a_i, a_j \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n \bar{v}_i w_j \delta_{ij} = \sum_{i=1}^n \bar{v}_i w_i = \langle v, w \rangle. \end{aligned}$$

„ \Rightarrow “: Sei A orthogonal bzw. unitär. Es folgt

$$\langle a_i, a_j \rangle = \langle A e_i, A e_j \rangle = \langle e_i, e_j \rangle = \delta_{ij},$$

d.h. die a_1, \dots, a_n bilden ein Orthonormalsystem. Aus Lemma 3.50 folgt die Behauptung. \square

Ende der Sitzung vom 14.5.

Definition 3.57 (Gruppen). Eine Menge G mit einer Verknüpfung $G \times G \rightarrow G$, $(a, b) \mapsto a \cdot b$, heißt *Gruppe*, wenn die Verknüpfung folgende Eigenschaften hat:

- (i) *Assoziativität*: Für alle $a, b, c \in G$ gilt $(a \cdot b) \cdot c = a \cdot (b \cdot c)$;
- (ii) Es gibt ein *Neutrales Element* $e \in G$, so dass $a \cdot e = e \cdot a = a$ für alle $a \in G$ gilt;
- (iii) Zu jedem $a \in G$ existiert ein *Inverses Element* $a^{-1} \in G$ mit $a \cdot a^{-1} = a^{-1} \cdot a = e$.

Bemerkung 3.58. Das neutrale Element und die Inversen Elemente sind in einer Gruppe eindeutig bestimmt.

Satz und Definition 3.59. *Die Menge*

$$\mathrm{GL}(n, \mathbb{K}) := \{A \in \mathbb{K}^{n \times n} \mid A \text{ invertierbar}\} = \{A \in \mathbb{K}^{n \times n} \mid \det(A) \neq 0\}$$

ist bezüglich der Matrixmultiplikation eine Gruppe, die allgemeine lineare Gruppe (general linear group) vom Grad n über \mathbb{K} .

Beweis. Sind $A, B \in \mathrm{GL}(n, \mathbb{K})$, so gilt $\det(AB) = \det(A) \det(B) \neq 0$, d.h. $AB \in \mathrm{GL}(n, \mathbb{K})$. Damit ist die Matrixmultiplikation eine Verknüpfung in $\mathrm{GL}(n, \mathbb{K})$. Die Assoziativität der Matrixmultiplikation wurde in Mathe 1 gezeigt. Das neutrale Element ist E_n , und es erfüllt (ii). Ist $A \in \mathrm{GL}(n, \mathbb{K})$, so betrachten wir die assoziierten Abbildungen L_A und $L_{A^{-1}} = L_A^{-1}$ in $\mathrm{Hom}(\mathbb{K}^n)$. Aus $L_A \circ L_A^{-1} = L_A^{-1} \circ L_A = \mathrm{id}$ folgt dann (iii). \square

Definition und Satz 3.60 (Untergruppe). Eine Teilmenge H einer Gruppe G mit Verknüpfung „ \cdot “ heißt *Untergruppe von G* , falls die Einschränkung der Verknüpfung von G auf H wohldefiniert ist und H bezüglich dieser Verknüpfung eine Gruppe ist. Dies ist erfüllt, wenn für alle $a, b \in H$ auch $a \cdot b$ und a^{-1} in H liegen. In diesem Fall stimmen die neutralen Elemente von H und G sowie, für $u \in H$, die Inversen von u in H und G überein.

Beweis. Übung. \square

Satz und Definition 3.61. (a) $A \in \mathbb{R}^{n \times n}$ ist orthogonal genau dann, wenn $A \in \mathrm{GL}(n, \mathbb{R})$ und $A^{-1} = A^t$ gelten. Wir setzen

$$\mathrm{O}(n) := \{A \in \mathrm{GL}(n, \mathbb{R}) \mid A \text{ is orthogonal}\}.$$

Es gilt $|\det(A)| = 1$ für alle $A \in \mathrm{O}(n)$. Wir setzen

$$\mathrm{SO}(n) := \{A \in \mathrm{O}(n) \mid \det(A) = 1\}.$$

Die Mengen $\mathrm{O}(n)$ und $\mathrm{SO}(n)$ sind Untergruppen von $\mathrm{GL}(n, \mathbb{R})$ und heißen orthogonale Gruppe bzw. spezielle orthogonale Gruppe.

(b) $A \in \mathbb{C}^{n \times n}$ ist unitär genau dann, wenn $A \in \text{GL}(n, \mathbb{C})$ und $A^{-1} = \overline{A}^t$ gelten. Wir setzen

$$U(n) := \{A \in \text{GL}(n, \mathbb{C}) \mid A \text{ is unitär}\}.$$

Es gilt $|\det(A)| = 1$ für alle $A \in U(n)$. Die Menge $U(n)$ ist eine Untergruppe von $\text{GL}(n, \mathbb{C})$ und heißt unitäre Gruppe.

(c) Sei $A \in O(n)$ oder $A \in U(n)$. Dann haben alle Eigenwerte von A den Betrag 1.

Beweis. Wir zeigen nur: $A \in \mathbb{C}^{n \times n}$ ist unitär genau dann, wenn $A \in \text{GL}(n, \mathbb{C})$ und $A^{-1} = \overline{A}^t$ gelten. Der Beweis der entsprechenden Aussage für $\mathbb{K} = \mathbb{R}$ geht analog.

„ \Rightarrow “: Seien $x, y \in \mathbb{C}^n$ beliebig. Dann gilt

$$(3.11) \quad \langle x, y \rangle = \langle Ax, Ay \rangle = (\overline{Ax})^t(Ay) = \overline{x}^t \overline{A}^t Ay = \langle x, \overline{A}^t Ay \rangle.$$

Also ist nach umstellen

$$0 = \langle x, (E_n - \overline{A}^t A)y \rangle.$$

Für die Wahl $x := (E_n - \overline{A}^t A)y$ ist $\langle x, x \rangle = 0$. Es folgt $x = 0$, d.h. $(E_n - \overline{A}^t A)y = 0$ für alle $y \in \mathbb{C}^n$. Daraus folgt $E_n = \overline{A}^t A$, d.h. A ist injektiv, also bijektiv. Da A^{-1} eindeutig bestimmt ist, folgt $A^{-1} = \overline{A}^t$.

„ \Leftarrow “: Sei $A \in \text{GL}(n, \mathbb{C})$ mit $A^{-1} = \overline{A}^t$ gegeben. Mit Verwendung der Rechnung in (3.11) folgt für alle $x, y \in \mathbb{C}^n$:

$$\langle x, y \rangle = \langle x, \overline{A}^t Ay \rangle = \langle Ax, Ay \rangle.$$

Somit ist A unitär.

Beweise der restlichen Behauptungen: Übung! □

Reelle orthogonale Matrizen beschreiben geometrische Operationen im \mathbb{R}^n . Im einfachsten interessanten Fall $n = 2$ haben wir:

Lemma 3.62. Sei $A \in O(2)$.

(a) Für $\det(A) = 1$ gilt $A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$ mit $\varphi \in (-\pi, \pi]$ (Drehung in \mathbb{R}^2).

(b) Für $\det(A) = -1$ gilt $A = \begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}$ mit $\varphi \in (-\pi, \pi]$ (Spiegelung in \mathbb{R}^2 an einer Geraden).

Beweis. Übung. □

3.3.2 Selbstadjungierte Matrizen

Für $v, w \in \mathbb{K}^n$ und $A \in \mathbb{K}^{n \times n}$ gilt

$$\langle v, Aw \rangle = \overline{v}^t Aw = \overline{v}^t \overline{\overline{A}^t} w = \overline{(\overline{A}^t v)}^t w = \langle \overline{A}^t v, w \rangle.$$

Das motiviert folgende

Definition 3.63. Für $A \in \mathbb{K}^{n \times n}$ heißt $A^* := \overline{A}^t$ die zu A *adjungierte Matrix* oder die *Adjungierte von A* . Es gilt dann für alle $v, w \in \mathbb{K}^n$: $\langle v, Aw \rangle = \langle A^*v, w \rangle$. Ist $A^* = A$, dann heißt A *selbstadjungiert* und es gilt für alle $v, w \in \mathbb{K}^n$: $\langle Av, w \rangle = \langle v, Aw \rangle$. Ist $\mathbb{K} = \mathbb{C}$, dann nennt man A in diesem Fall auch *hermitesch*. Ist $\mathbb{K} = \mathbb{R}$, dann ist selbstadjungiert gleichbedeutend mit $A^t = A$ und man nennt eine solche Matrix A auch *symmetrisch*.

Lemma 3.64. Sei $A \in \mathbb{K}^{n \times n}$ selbstadjungiert. Dann gilt:

- (a) Alle Eigenwerte von A sind reell.
- (b) Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal (bzgl. des Standardskalarproduktes).

Beweis. Es genügt $A \in \mathbb{C}^{n \times n}$ zu betrachten. Zu (a): Sei $\lambda \in \mathbb{C}$ ein Eigenwert von A mit Eigenvektor $v \in \mathbb{C}^n$. Dann gilt

$$\overline{\lambda}|v|_2^2 = \overline{\lambda}\langle v, v \rangle = \langle \lambda v, v \rangle = \langle Av, v \rangle = \langle v, Av \rangle = \langle v, \lambda v \rangle = \lambda\langle v, v \rangle = \lambda|v|^2.$$

Es folgt $\overline{\lambda} = \lambda$, also $\lambda \in \mathbb{R}$.

Zu (b): Seien $\lambda_1 \neq \lambda_2$ zwei Eigenwerte von A mit zugehörigen Eigenvektoren $v_1, v_2 \in \mathbb{K}^n \setminus \{0\}$. Dann ist wegen (a)

$$\begin{aligned} (\lambda_1 - \lambda_2)\langle v_1, v_2 \rangle &= \langle \lambda_1 v_1, v_2 \rangle - \langle v_1, \lambda_2 v_2 \rangle = \langle Av_1, v_2 \rangle - \langle v_1, Av_2 \rangle \\ &= \langle v_1, Av_2 \rangle - \langle v_1, Av_2 \rangle = 0. \end{aligned}$$

Es folgt $\langle v_1, v_2 \rangle = 0$ und damit die Behauptung. □

Satz 3.65. Sei $A \in \mathbb{K}^{n \times n}$ selbstadjungiert. Dann ist A diagonalisierbar. Genauer: Es gibt eine orthogonale (falls $\mathbb{K} = \mathbb{R}$) bzw. eine unitäre (falls $\mathbb{K} = \mathbb{C}$) Matrix U mit

$$U^{-1}AU = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wo $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ die Eigenwerte von A sind. Hierbei sind die Spalten von U normierte Eigenvektoren von A . Im Falle $\mathbb{K} = \mathbb{R}$ nennt man diesen Basiswechsel auch Hauptachsentransformation.

Beweis. Für $\mathbb{K} = \mathbb{C}$ zerfällt das charakteristische Polynom p_A in Linearfaktoren. Für $\mathbb{K} = \mathbb{R}$ zerfällt p_A auch über \mathbb{C} , und da nach Lemma 3.64(a) alle Nullstellen von p_A reell sind, zerfällt p_A auch über \mathbb{R} . In beiden Fällen ist die Summe der algebraischen Vielfachheiten der Eigenwerte gleich n . Es reicht noch zu zeigen, dass für jeden Eigenwert seine geometrische Vielfachheit gleich seiner algebraischen Vielfachheit ist. Anschließend kann nämlich jeder Eigenraum nach Satz 3.53 mit einer ONB ausgestattet werden, und die Gesamtheit dieser ONBen liefert eine ONB für \mathbb{C}^n aus Eigenvektoren von A , wegen Lemma 3.64(b). Die Diagonalisierbarkeit folgt dann aus Satz 3.34.

Sei also λ ein Eigenwert. Angenommen, die geometrische Vielfachheit von λ ist kleiner als die algebraische Vielfachheit. Dann existiert ein Nebeneigenvektor 2. Stufe

$$v \in \text{Kern}(A - \lambda E_n)^2 \setminus \text{Kern}(A - \lambda E_n).$$

Nach (3.8) erhalten wir

$$w := (A - \lambda E_n)v \subseteq \text{Kern}(A - \lambda E_n) \setminus \{0\} = \mathcal{E}_A(\lambda) \setminus \{0\},$$

d.h. w ist ein Eigenvektor zum Eigenwert λ . Dies liefert

$$\begin{aligned} 0 \neq \langle w, w \rangle &= \langle w, (A - \lambda E_n)v \rangle = \langle w, Av \rangle - \lambda \langle w, v \rangle = \langle Aw, v \rangle - \lambda \langle w, v \rangle \\ &= \langle \lambda w, v \rangle - \lambda \langle w, v \rangle = \lambda \langle w, v \rangle - \lambda \langle w, v \rangle = 0, \end{aligned}$$

Widerspruch! Also muss die geometrische Vielfachheit gleich der algebraischen sein. \square

Bemerkung 3.66. Eine Matrix $A \in \mathbb{K}^{n \times n}$ heißt *normal*, wenn $A^*A = AA^*$ gilt. Zum Beispiel sind orthogonale, unitäre und selbstadjungierte Matrizen normal. Für $\mathbb{K} = \mathbb{C}$ kann man zeigen, dass auch normale Matrizen mittels einer ONB aus Eigenvektoren diagonalisierbar sind. Die Eigenwerte sind dann aber nicht notwendigerweise reell. Reelle normale Matrizen besitzen immerhin nach dem Wechsel zu einer geeigneten ONB eine Normalform als Blockdiagonalmatrix aus 1×1 - und 2×2 -Matrizen, wobei letztere positive Vielfache von Matrizen aus $\text{SO}(2)$ sind.

3.3.3 Positiv definite Matrizen

Definition 3.67. Eine selbstadjungierte Matrix $A \in \mathbb{K}^{n \times n}$ heißt

- (a) *positiv semidefinit* (nichtnegativ, $A \geq 0$), wenn $\langle v, Av \rangle \geq 0$ für alle $v \in \mathbb{K}^n$ gilt;
- (b) *positiv definit* (positiv, $A > 0$), wenn $\langle v, Av \rangle > 0$ für alle $v \in \mathbb{K}^n \setminus \{0\}$ gilt.
- (c) *negativ (semi)definit*, wenn $-A$ positiv (semi)definit ist.
- (d) *indefinit*, wenn A weder positiv noch negativ semidefinit ist.

Bemerkung 3.68. Auf \mathbb{K}^n lässt sich für eine gegebene Matrix $A \in \mathbb{K}^{n \times n}$ eine Abbildung wie folgt definieren:

$$\langle \cdot, \cdot \rangle_A: \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}, \quad (v, w) \mapsto \langle v, Aw \rangle.$$

Wir beobachten: $\langle \cdot, \cdot \rangle_A$ ist bilinear (falls $\mathbb{K} = \mathbb{R}$) bzw. sesquilinear (falls $\mathbb{K} = \mathbb{C}$).

- (a) Es gilt $\langle v, w \rangle_A = \overline{\langle w, v \rangle_A}$ genau dann, wenn A eine selbstadjungierte Abbildung ist.
- (b) $\langle v, w \rangle_A$ beschreibt ein Skalarprodukt genau dann, wenn A selbstadjungiert **und** positiv definit ist.

- (c) Für **jedes** Skalarprodukt $\langle \cdot, \cdot \rangle_*$ auf \mathbb{K}^n gibt es eine selbstadjungierte Abbildung A mit $\langle \cdot, \cdot \rangle_* = \langle \cdot, \cdot \rangle_A$.

Wir schreiben im letzten Fall auch

$$\langle v, w \rangle_* = \langle v|A|w \rangle = \langle v, Aw \rangle = \langle Av, w \rangle = \langle v, w \rangle_A.$$

Satz 3.69. Seien $A \in \mathbb{C}^{n \times n}$ selbstadjungiert und $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A (mit Mehrfachnennung). Dann gilt:

- (a) $A \geq 0$ genau dann, wenn $\lambda_i \geq 0$ für alle $i \in \{1, \dots, n\}$.
 (b) $A > 0$ genau dann, wenn $\lambda_i > 0$ für alle $i \in \{1, \dots, n\}$.

Beweis. Nach Satz 3.65 gibt es eine ONB von \mathbb{K}^n aus Eigenvektoren v_1, \dots, v_n von A zu zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Sei $v \in \mathbb{K}^n$ gegeben durch $v = \sum_{i=1}^n c_i v_i$. Dann folgt

$$\begin{aligned} \langle v, Av \rangle &= \left\langle \sum_{i=1}^n c_i v_i, A \left(\sum_{i=1}^n c_i v_i \right) \right\rangle = \left\langle \sum_{i=1}^n c_i v_i, \sum_{i=1}^n c_i A v_i \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n \bar{c}_i c_j \langle v_i, A v_j \rangle = \sum_{i=1}^n \sum_{j=1}^n \bar{c}_i c_j \langle v_i, \lambda_j v_j \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_j \bar{c}_i c_j \langle v_i, v_j \rangle = \sum_{i=1}^n \sum_{j=1}^n \lambda_j \bar{c}_i c_j \delta_{ij} \\ &= \sum_{i=1}^n \lambda_i |c_i|^2. \end{aligned}$$

Zu „ \Leftarrow “: Ist $\lambda_i \geq 0$ bzw. > 0 für alle $i = 1, \dots, n$, so gilt für $v = \sum_{i=1}^n c_i v_i \neq 0$:

$$\langle v, Av \rangle = \sum_{i=1}^n \lambda_i |c_i|^2 \geq 0 \quad \text{bzw.} \quad > 0.$$

Zu „ \Rightarrow “: Ist A nichtnegativ bzw. positiv, so gilt für die Basisvektoren v_j , $j = 1, \dots, n$:

$$\lambda_j = \langle v_j, A v_j \rangle \geq 0 \quad \text{bzw.} \quad \lambda_j = \langle v_j, A v_j \rangle > 0.$$

Es folgt die Behauptung. □

Beispiel 3.70. Die Matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ ist eine selbstadjungierte Matrix in \mathbb{K}^2 . Das charakteristische Polynom ist $p_A(\lambda) = (1 - \lambda)^2 - 1 = \lambda^2 - 2\lambda = \lambda(\lambda - 2)$. Also hat A die Eigenwerte 0 und 2. Beide sind nichtnegativ und daher ist A eine positiv semidefinite Matrix.

Definition 3.71. Es sei $A \in \mathbb{K}^{n \times n}$. Wir bezeichnen die Summe der Diagonalelemente,

$$\text{Spur}(A) := a_{11} + a_{22} + \dots + a_{nn} = \sum_{i=1}^n a_{ii}$$

als *Spur* von A .

Lemma 3.72. Sei $A \in \mathbb{R}^{2 \times 2}$ symmetrisch. Dann gilt

- (a) A ist positiv (negativ) definit genau dann, wenn $\det(A) > 0$ und $\text{Spur}(A) := a_{11} + a_{22} > (<) 0$.
- (b) A ist positiv (negativ) semidefinit genau dann, wenn $\det(A) \geq 0$ und $\text{Spur}(A) := a_{11} + a_{22} \geq (\leq) 0$.
- (c) A ist indefinit genau dann, wenn $\det(A) < 0$.

Beweis. Übung. □

Ende der Sitzung vom 18.5.

3.3.4 Die reelle Jordansche Normalform

Das charakteristische Polynom einer reellen Matrix $A \in \mathbb{R}^{n \times n}$ zerfällt im Allgemeinen nicht über \mathbb{R} . Dennoch kann man in diesem Fall einen Basiswechsel in \mathbb{R}^n angeben, so dass die Matrix in eine einfache Form überführt wird. Diese gewinnt man aus der komplexen Normalform, indem man jeweils Real- und Imaginärteil von Eigenvektoren, Nebeneigenvektoren und Eigenwerten zur Konstruktion der Basis und der Normalform nimmt. Dies funktioniert, weil die echt komplexen Nullstellen des reellen charakteristischen Polynoms jeweils in Paaren zueinander komplex konjugierter Zahlen auftreten. Wir stellen hier nur die sich ergebende Normalform der Matrix dar, ohne näher auf die Konstruktion der Basis einzugehen.

Bemerkung 3.73 (Reelle Jordansche Normalform). Die reelle Matrix $A \in \mathbb{R}^{n \times n}$ habe die reellen Eigenwerte $\lambda_1, \dots, \lambda_\ell$ mit den algebraischen Vielfachheiten $\alpha_1, \dots, \alpha_\ell$ und die echt komplexen Eigenwertpaare $(\mu_1, \bar{\mu}_1), \dots, (\mu_m, \bar{\mu}_m)$ mit den algebraischen Vielfachheiten $\alpha'_1, \dots, \alpha'_m$. Seien $J_{\alpha_j}^*(\lambda_j)$ die Jordanmatrizen zu den reellen Eigenwerten wie im Abschnitt 3.2.2. Für ein j und einen komplexen Eigenvektor zum Eigenwertpaar $x_j \pm iy_j$ mit $x_j, y_j \in \mathbb{R}$, $y_j \neq 0$, hat der zugehörige reelle Jordanblock der Länge $2k$ die Gestalt (in Blockmatrixform)

$$J'_k(\mu_j) := \begin{pmatrix} M_j & E_2 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & E_2 \\ 0 & \cdots & \cdots & 0 & M_j \end{pmatrix}$$

mit k 2×2 -Matrizen $M_j := \begin{pmatrix} x_j & -y_j \\ y_j & x_j \end{pmatrix}$. Offensichtlich ist M_j das positive Vielfache einer Drehmatrix, siehe Lemma 3.62(a). Die reellen Jordanmatrizen zum Eigenwertpaar $x_j \pm iy_j$ sind dann

$$J''_{\alpha'_j}(\mu_j) := \begin{pmatrix} J'_{s_\gamma}(\mu_j) & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & J'_{s_{\gamma-1}}(\mu_j) & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & J'_{s_1}(\mu_j) \end{pmatrix}$$

und haben die Länge $2\alpha'_j$. Die Matrix A ist ähnlich zu einer reellen Matrix der Gestalt

$$\begin{pmatrix} J^*_{\alpha_1}(\lambda_1) & & & & & & & 0 \\ & \ddots & & & & & & \\ & & J^*_{\alpha_\ell}(\lambda_\ell) & & & & & \\ & & & J''_{\alpha'_1}(\mu_1) & & & & \\ & & & & \ddots & & & \\ 0 & & & & & \ddots & & J''_{\alpha'_m}(\mu_m) \end{pmatrix}.$$

4 Die Topologie des \mathbb{R}^N

4.1 Metrische Räume

Definition 4.1 (Metrik). Sei X eine Menge. Eine Metrik auf X ist eine Abbildung $d: X \times X \rightarrow [0, \infty)$ derart, dass für $x, y, z \in X$ gilt:

$$(M1) \quad d(x, y) = 0 \Leftrightarrow x = y \quad (\text{Definitheit})$$

$$(M2) \quad d(x, y) = d(y, x) \quad (\text{Symmetrie})$$

$$(M3) \quad d(x, z) \leq d(x, y) + d(y, z) \quad (\text{Dreiecksungleichung})$$

Das Paar (X, d) heißt *metrischer Raum*.

Beispiel 4.2. (a) Ist $(V, \|\cdot\|)$ ein normierter Raum, so ist (V, d) ein metrischer Raum mit der durch die Norm induzierten Metrik $d: V \times V \rightarrow [0, \infty)$, $d(x, y) := \|x - y\|$. Es gilt also

$$\text{Skalarproduktraum} \quad \longrightarrow \quad \text{normierter Raum} \quad \longrightarrow \quad \text{metrischer Raum}$$

mit induzierten Strukturen. Alle Definition und Sätze über metrische Räume übertragen sich daher direkt auf normierte Räume und Skalarprodukträume.

Spezielle induzierte Metriken sind zum Beispiel $d_p(x, y) = |x - y|_p$ für $x, y \in \mathbb{R}^N$ und $p \in [1, \infty]$.

(b) Sei X eine beliebige Menge. Die *diskrete* Metrik d auf X ist definiert durch

$$d(x, y) := \begin{cases} 0, & x = y; \\ 1, & x \neq y. \end{cases}$$

Im Folgenden sei (X, d) stets ein metrischer Raum.

Definition 4.3 (Grenzwert). Sei $(x_n)_n$ eine Folge in X und $a \in X$. Gilt $\lim_{n \rightarrow \infty} d(x_n, a) = 0$, so nennt man a *Grenzwert* der Folge $(x_n)_n$ und schreibt

$$a = \lim_{n \rightarrow \infty} x_n \quad \text{oder} \quad x_n \rightarrow a \text{ für } n \rightarrow \infty. \quad ((x_n)_n \text{ konvergiert gegen } a)$$

Satz 4.4. *Konvergiert eine Folge $(x_n)_n$ in X , so ist der Grenzwert eindeutig bestimmt.*

Beweis. Sind $a, b \in X$ mit $\lim_{n \rightarrow \infty} d(x_n, a) = 0 = \lim_{n \rightarrow \infty} d(x_n, b)$, so gilt

$$0 \leq d(a, b) \stackrel{(M3)}{\leq} d(a, x_n) + d(x_n, b) \rightarrow 0 \quad \text{für } n \rightarrow \infty,$$

also $d(a, b) = 0$ und somit $a = b$ wegen (M1). \square

Beispiel 4.5. Betrachte $x_n = (1, \frac{1}{n}) \in \mathbb{R}^2$ für $n \in \mathbb{N}$ und $a = (1, 0) \in \mathbb{R}^2$. Für $1 \leq p \leq \infty$ gilt $\lim_{n \rightarrow \infty} x_n = a$ im metrischen Raum (\mathbb{R}^2, d_p) , da

$$d_p(x_n, a) = |x_n - a|_p = \left| \left(0, \frac{1}{n} \right) \right|_p = \frac{1}{n} \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Definition 4.6. (a) Für $a \in X$, $r > 0$ sei

$$\begin{aligned} B_r(a) &:= \{y \in X \mid d(y, a) < r\} && \text{(offene } r\text{-Kugel um } a) \\ \overline{B}_r(a) &:= \{y \in X \mid d(y, a) \leq r\} && \text{(abgeschlossene } r\text{-Kugel um } a) \end{aligned}$$

(b) In einem normierten Raum $(V, \|\cdot\|)$ heißt $\overline{B}_1(0) = \{v \in V \mid \|v\| \leq 1\}$ *Einheitskugel* von V .

Definition 4.7. (a) Sei $A \subseteq X$. Der *Durchmesser* von A ist definiert durch

$$\text{diam}(A) := \sup\{d(x, y) \mid x, y \in A\} \in \{-\infty\} \cup [0, \infty].$$

(b) A heißt *beschränkt*, falls $\text{diam}(A) < \infty$.

(c) Eine Folge $(x_n)_n$ in X heißt *beschränkt*, falls die Menge $\{x_n \mid n \in \mathbb{N}\}$ beschränkt ist.

Bemerkung 4.8. (a) Es gilt $\text{diam}(\overline{B}_r(a)) \leq 2r$ für $a \in X, r > 0$.

(b) Ist $A \subseteq B \subseteq X$, so gilt $\text{diam}(A) \leq \text{diam}(B)$.

(c) Eine nichtleere Teilmenge $A \subseteq X$ ist beschränkt genau dann, wenn $a \in X$ und $r > 0$ existiert mit $A \subseteq B_r(a)$.

Beweis. **(a):** Für $x, y \in \overline{B}_r(a)$ gilt $d(x, y) \leq d(x, a) + d(a, y) \leq 2r$, und somit folgt $\text{diam}(\overline{B}_r(a)) \leq 2r$.

(b): Dies folgt direkt aus der Definition von diam .

(c): „ \Leftarrow “: Gilt $A \subseteq B_r(a)$ für ein $a \in X, r > 0$, so folgt

$$\text{diam}(A) \stackrel{(b)}{\leq} \text{diam}(B_r(a)) \stackrel{(b)}{\leq} \text{diam}(\overline{B}_r(a)) \stackrel{(a)}{\leq} 2r < \infty,$$

und somit ist A beschränkt.

„ \Rightarrow “: Nach Voraussetzung ist $r' := \text{diam}(A) \in [0, \infty)$. Mit $r := r' + 1$ und beliebig gewähltem $a \in A$ gilt dann

$$d(x, a) \leq \text{diam}(A) < r \quad \text{für alle } x \in A,$$

also $A \subseteq B_r(a)$. □

Definition 4.9. Sei $A \subseteq X$ und $a \in X$. A heißt

- (a) *Umgebung von a* , wenn es $\varepsilon > 0$ gibt mit $B_\varepsilon(a) \subseteq A$.
- (b) *offen* (in X), wenn es zu jedem $x \in A$ ein $\varepsilon > 0$ gibt mit $B_\varepsilon(x) \subseteq A$, d.h. wenn A Umgebung von allen Punkten in A ist.
- (c) *abgeschlossen* (in X), wenn $X \setminus A$ offen ist.

Satz 4.10. In jedem metrischen Raum (X, d) gilt:

- (a) Für $a \in X$, $r > 0$ ist $B_r(a)$ offen und $\overline{B_r(a)}$ abgeschlossen.
- (b) X , \emptyset sind jeweils offen **und** abgeschlossen.

Beweis. **(a):** Sei $x \in B_r(a)$, d.h. $d(x, a) < r$. Mit $\varepsilon := r - d(x, a) > 0$ gilt für $y \in B_\varepsilon(x)$:

$$d(y, a) \leq d(y, x) + d(x, a) < \varepsilon + d(x, a) = r, \quad \text{also } y \in B_r(a).$$

Somit ist $B_\varepsilon(x) \subseteq B_r(a)$. Es folgt: $B_r(a)$ ist offen.

Zum Beweis der Abgeschlossenheit von $\overline{B_r(a)}$ zeigen wir:

$$(4.1) \quad X \setminus \overline{B_r(a)} \text{ ist offen.}$$

Sei $x \in X \setminus \overline{B_r(a)}$, d.h. $d(x, a) > r$. Mit $\varepsilon := d(x, a) - r > 0$ gilt für $y \in B_\varepsilon(x)$:

$$d(y, a) \geq d(x, a) - d(x, y) > d(x, a) - \varepsilon = r, \quad \text{also } y \in X \setminus \overline{B_r(a)}.$$

Somit ist $B_\varepsilon(x) \subseteq X \setminus \overline{B_r(a)}$. Es folgt (4.1).

(b): ist trivial. □

Beispiel 4.11. Sei $A := \{(t, 0) \mid 1 < t < 3\} \subseteq \mathbb{R}^2$. Dann ist A weder offen noch abgeschlossen im metrischen Raum (\mathbb{R}^2, d_2) .

Man kann offene Mengen auch als Ausgangspunkt der Untersuchungen wählen:

Definition 4.12. Seien Y eine Menge und $\mathcal{O} \subseteq 2^Y$ (d.h. \mathcal{O} ist eine Familie von Teilmengen von Y) mit den folgenden Eigenschaften:

- (i) $Y, \emptyset \in \mathcal{O}$;
- (ii) wenn $U_i \in \mathcal{O}$ für alle $i \in I$ gilt, wo I eine beliebige Indexmenge ist, dann folgt $\bigcup_{i \in I} U_i \in \mathcal{O}$;

(iii) wenn $U_i \in \mathcal{O}$ für alle $i \in I$ gilt, wo I eine endliche Indexmenge ist, dann folgt $\bigcap_{i \in I} U_i \in \mathcal{O}$.

Dann heißt \mathcal{O} eine *Topologie auf Y* . Die Elemente von \mathcal{O} nennt man auch *offene Mengen*, und das Paar (Y, \mathcal{O}) heißt *topologischer Raum*. Eine Teilmenge $A \subseteq Y$ heißt *abgeschlossen* wenn $Y \setminus A$ offen ist.

Bemerkung 4.13. Sei Y ein topologischer Raum. Ist $A_i \subseteq Y$ abgeschlossen für $i \in I$, so ist $\bigcap_{i \in I} A_i$ abgeschlossen. Ist ferner I endlich, so ist auch $\bigcup_{i \in I} A_i$ abgeschlossen

Beweis. Für $i \in I$ ist A_i abgeschlossen, also $X \setminus A_i$ offen. Nach Definition 4.12(ii) ist

$$X \setminus \bigcap_{i \in I} A_i = \bigcup_{i \in I} (X \setminus A_i)$$

offen, also $\bigcap_{i \in I} A_i$ abgeschlossen. Ist $I = \{1, \dots, n\}$, so folgt aus Definition 4.12(iii), dass

$$X \setminus \bigcup_{i \in I} A_i = \bigcap_{i \in I} (X \setminus A_i)$$

offen, also $\bigcup_{i \in I} A_i$ abgeschlossen ist. □

Satz 4.14. Die Familie der (metrisch) offenen Mengen in X bilden eine Topologie auf X , die metrische Topologie oder die von der Metrik induzierte Topologie.

Beweis. Wegen Satz 4.10(b) reicht es, die Eigenschaften in Definition 4.12(ii) und (iii) für die metrisch offenen Mengen zu zeigen. Seien dazu I eine Indexmenge und $A_i \subseteq X$ offen für $i \in I$. Sei $x \in A := \bigcup_{i \in I} A_i$, d.h. $x \in A_{i_0}$ für ein $i_0 \in I$. Da A_{i_0} offen ist, existiert $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq A_{i_0} \subseteq A$. Es folgt, dass A offen ist.

Sei ferner I endlich, d.h. $I := \{1, \dots, n\}$ und $x \in \bigcap_{i=1}^n A_i$, d.h. $x \in A_i$ für $i = 1, \dots, n$. Nach Voraussetzung existieren $\varepsilon_i > 0$ mit $B_{\varepsilon_i}(x) \subseteq A_i$ für $i = 1, \dots, n$. Mit $\varepsilon := \min_{1 \leq i \leq n} \varepsilon_i > 0$ gilt dann

$$B_\varepsilon(x) \subseteq \bigcap_{i=1}^n B_{\varepsilon_i}(x) \subseteq \bigcap_{i=1}^n A_i.$$

Es folgt die Offenheit von $\bigcap_{i=1}^n A_i$. □

Bemerkung 4.15. Wir betrachten \mathbb{R} mit der von der Betragsnorm induzierten Metrik, um zu zeigen, dass in Satz 4.14 und in Bemerkung 4.13 die Endlichkeitsvoraussetzung nicht unwesentlich ist. Unendliche Schnitte offener Mengen sind im Allgemeinen nicht offen: Für alle $n \in \mathbb{N}$ ist $(-\frac{1}{n}, \frac{1}{n})$ offen in \mathbb{R} , aber $\{0\} = \bigcap_{n \in \mathbb{N}} (-\frac{1}{n}, \frac{1}{n})$ ist nicht offen. Unendliche Vereinigungen abgeschlossener Mengen sind dementsprechend im Allgemeinen auch nicht abgeschlossen: Für alle $n \in \mathbb{N}$ ist z.B. $[\frac{1}{n}, 1]$ abgeschlossen in \mathbb{R} , aber $(0, 1] = \bigcup_{n \in \mathbb{N}} [\frac{1}{n}, 1]$ ist nicht abgeschlossen.

Definition 4.16. Sei $A \subseteq X$ und $a \in X$.

(a) Wir setzen

$$\begin{aligned}\mathring{A} &:= \bigcup_{B \subseteq A, B \text{ offen}} B && (\text{Offener Kern oder Inneres von } A) \\ \bar{A} &:= \bigcap_{B \supseteq A, B \text{ abgeschl.}} B && (\text{Abschluss von } A) \\ \partial A &:= \bar{A} \setminus \mathring{A} && (\text{Rand von } A).\end{aligned}$$

Wir verwenden außerdem die Schreibweise $\text{int } A$ (*interior of A*) statt \mathring{A} für kompliziertere Ausdrücke A .

(b) a heißt

- *innerer Punkt* von A , falls $a \in \mathring{A}$.
- *Berührungspunkt* von A , falls $a \in \bar{A}$.
- *Randpunkt* von A , falls $a \in \partial A$.
- *Häufungspunkt* von A , falls $a \in \overline{A \setminus \{a\}}$.

Satz 4.17. *Seien $A, C \subseteq X$. Dann gilt:*

- (a) \mathring{A} ist offen, und $\bar{A}, \partial A$ sind abgeschlossen.
 (b) A offen $\Leftrightarrow A = \mathring{A}$.
 (c) A abgeschlossen $\Leftrightarrow A = \bar{A}$.
 (d) $\overline{\bar{A}} = \bar{A}$ und $\text{int}(\text{int } A) = \text{int } A$.
 (e) $\overline{X \setminus A} = X \setminus \mathring{A}$, $X \setminus \bar{A} = \text{int}(X \setminus A)$ und $\partial A = \bar{A} \cap \overline{X \setminus A}$.
 (f) Ist $A \subseteq C$, so folgt $\mathring{A} \subseteq \mathring{C}$ und $\bar{A} \subseteq \bar{C}$.

Beweis. (e):

$$(4.2) \quad \overline{X \setminus A} = \bigcap_{\substack{B \supseteq X \setminus A \\ B \text{ abgeschl.}}} B = \bigcap_{\substack{B' \subseteq A \\ B' \text{ offen}}} X \setminus B' = X \setminus \bigcup_{\substack{B' \subseteq A \\ B' \text{ offen}}} B' = X \setminus \mathring{A}.$$

Sei ferner $D := X \setminus A$. Dann gilt $X \setminus \mathring{D} \stackrel{(4.2)}{=} \overline{X \setminus D} = \bar{A}$, also $\mathring{D} = X \setminus \bar{A}$. Schließlich ist

$$\partial A = \bar{A} \setminus \mathring{A} = \bar{A} \cap (X \setminus \mathring{A}) \stackrel{(4.2)}{=} \bar{A} \cap \overline{X \setminus A}.$$

(a): Wegen Satz 4.14 ist \mathring{A} offen und \bar{A} abgeschlossen. Somit ist auch

$$\partial A \stackrel{(e)}{=} \bar{A} \cap \overline{X \setminus A} \quad \text{abgeschlossen.}$$

(b) und (c) folgen aus (a) und Definition 4.16. (d) folgt aus (a)–(c). (f) folgt direkt aus Definition 4.16. \square

Definition 4.18. Für $a \in X$ und $A \subseteq X$ sei

$$\text{dist}(a, A) := \inf\{d(a, y) \mid y \in A\} \quad (\text{Abstand von } a \text{ zu } A)$$

Satz 4.19. Seien $A \subseteq X$ und $a \in X$. Dann gilt:

- (a) Genau dann ist a innerer Punkt von A , wenn $\varepsilon > 0$ existiert mit $B_\varepsilon(a) \subseteq A$, d.h. wenn A eine Umgebung von a ist.
- (b) Äquivalent sind:
 - (i) $a \in \overset{\circ}{A}$ (d.h. a ist Berührungspunkt von A).
 - (ii) $\text{dist}(a, A) = 0$
 - (iii) Es existiert eine Folge $(x_n)_n$ in A mit $\lim_{n \rightarrow \infty} x_n = a$.
- (c) a ist Häufungspunkt von A genau dann, wenn eine Folge $(x_n)_n$ in $A \setminus \{a\}$ existiert mit $\lim_{n \rightarrow \infty} x_n = a$.

Beweis. (a): „ \Rightarrow “: Sei $a \in \overset{\circ}{A}$. Da $\overset{\circ}{A}$ offen ist, existiert $\varepsilon > 0$ mit $B_\varepsilon(a) \subseteq \overset{\circ}{A} \subseteq A$.

„ \Leftarrow “: Gilt umgekehrt $B_\varepsilon(a) \subseteq A$ für ein $\varepsilon > 0$, so folgt $B_\varepsilon(a) \subseteq \overset{\circ}{A}$, da $B_\varepsilon(a)$ offen ist. Insbesondere folgt $a \in \overset{\circ}{A}$.

(b): Es gilt:

$$(4.3) \quad \begin{aligned} a \in \overline{A} &\Leftrightarrow a \notin X \setminus \overline{A} = \text{int}(X \setminus A) \\ &\stackrel{(a)}{\Leftrightarrow} \text{für alle } \varepsilon > 0 \text{ ist } B_\varepsilon(a) \not\subseteq X \setminus A \\ &\Leftrightarrow \text{für alle } \varepsilon > 0 \text{ ist } B_\varepsilon(a) \cap A \neq \emptyset. \end{aligned}$$

„(i) \Rightarrow (iii)“: Es folgt aus (4.3), dass für alle $n \in \mathbb{N}$ ein $x_n \in B_{\frac{1}{n}}(a) \cap A$ existiert, d.h. $d(x_n, a) < \frac{1}{n} \rightarrow 0$ für $n \rightarrow \infty$.

„(iii) \Rightarrow (ii)“: Es gilt $\text{dist}(a, A) \leq d(a, x_n)$ für alle $n \in \mathbb{N}$, also $\text{dist}(a, A) = 0$.

„(ii) \Rightarrow (i)“: Für alle $\varepsilon > 0$ existiert ein $x \in A$ mit $d(x, a) < \varepsilon$, d.h. für alle $\varepsilon > 0$ ist $B_\varepsilon(a) \cap A \neq \emptyset$, und (4.3) liefert daher (i).

(c): folgt aus (b) angewandt auf $A \setminus \{a\}$. □

Korollar 4.20. $A \subseteq X$ ist abgeschlossen genau dann, wenn alle Folgen in A , die in X konvergieren, ihren Grenzwert in A haben.

Beweis. Dies folgt direkt aus Satz 4.17(c) und Satz 4.19(b). □

Satz und Definition 4.21. Ist $M \subseteq X$, so erfüllt die Einschränkung von d auf $M \times M$ ebenfalls die metrischen Axiome, d.h. (M, d) ist ebenfalls ein metrischer Raum. Diese Metrik auf M heißt die von X induzierte Metrik und die zugehörige Topologie auf M heißt die induzierte (oder auch relative) Topologie von M . Die entsprechenden offenen und abgeschlossenen Mengen heißen relativ offen bzw. relativ abgeschlossen. Dabei gilt:

- (a) A ist relativ offen in M genau dann, wenn eine offene Teilmenge $A' \subseteq X$ existiert mit $A = M \cap A'$.
- (b) A ist relativ abgeschlossen in M genau dann, wenn $\overline{A}^X \cap M \subseteq A$ gilt. Hierbei bezeichne \overline{A}^X den Abschluss von A in X . In diesem Fall gilt $A = \overline{A}^X \cap M$.

Der Beweis dieser Aussagen verbleibt als Übung.

Beispiel 4.22. Sei $X = \mathbb{R}^2$ und $M = \mathbb{Z}^2 \subseteq \mathbb{R}^2$, jeweils betrachtet mit der Metrik d_2 . Dann ist jede Teilmenge von M relativ offen und abgeschlossen in M , aber keine nichtleere Teilmenge von M ist offen (in X).

4.2 Äquivalente Normen

Satz 4.23 (Bolzano-Weierstraß). Sei $(x_k)_k$ ein Folge in \mathbb{R}^N derart, dass $C > 0$ existiert mit $|x_k|_\infty \leq C$ für alle $k \in \mathbb{N}$. Dann gibt es eine Teilfolge $(x_{k_j})_j$ und $a \in \mathbb{R}^N$ mit $\lim_{j \rightarrow \infty} x_{k_j} = a$ bezüglich $|\cdot|_\infty$.

Beweis. Induktion nach der Raumdimension N .

- (a) „ $N = 1$, d.h. $\mathbb{R}^N = \mathbb{R}$, $|\cdot|_\infty = |\cdot|$ “: In diesem Fall folgt die Behauptung aus dem Satz von Bolzano-Weierstraß aus Mathe 1.
- (b) „ $N \Rightarrow N + 1$ “: Schreibe $x_k = (y_k, t_k) \in \mathbb{R}^{N+1} = \mathbb{R}^N \times \mathbb{R}$ mit $y_k \in \mathbb{R}^N$, $t_k \in \mathbb{R}$. Mit $(x_k)_k$ sind dann auch die Folgen $(y_k)_k$ und $(t_k)_k$ bzgl. $|\cdot|_\infty$ bzw. $|\cdot|$ beschränkt. Nach Induktionsannahme existiert $y \in \mathbb{R}^N$ und eine Teilfolge $(y_{k_j})_j$ von $(y_k)_k$ mit $\lim_{j \rightarrow \infty} y_{k_j} = y$ in $(\mathbb{R}^N, |\cdot|_\infty)$. Zudem existiert nach Satz von Bolzano-Weierstraß eine Teilfolge $(t_{k_{j_\ell}})_\ell$ von $(t_{k_j})_j$ und $t \in \mathbb{R}$ mit $\lim_{\ell \rightarrow \infty} t_{k_{j_\ell}} = t$. Mit $a = (y, t) \in \mathbb{R}^{N+1}$ folgt nun

$$|x_{k_{j_\ell}} - a|_\infty = \max\{|y_{k_{j_\ell}} - y|_\infty, |t_{k_{j_\ell}} - t|_\infty\} \rightarrow 0 \quad \text{für } \ell \rightarrow \infty.$$

Es folgt $\lim_{\ell \rightarrow \infty} x_{k_{j_\ell}} = a$ bzgl. $|\cdot|_\infty$ und somit die Behauptung in \mathbb{R}^{N+1} . \square

Definition 4.24. Sei V ein \mathbb{R} -Vektorraum. Zwei Normen $\|\cdot\|$ und $\|\cdot\|_*$ auf V heißen äquivalent, falls $c_1, c_2 > 0$ existieren mit

$$c_1 \|v\| \leq \|v\|_* \leq c_2 \|v\| \quad \text{für } v \in V.$$

Wir schreiben dann $\|\cdot\| \sim \|\cdot\|_*$.

Es ist eine einfache Übung zu zeigen, dass die Äquivalenz von Normen eine Äquivalenzrelation auf der Menge aller Normen in einem Vektorraum ist.

Satz 4.25. Alle Normen auf \mathbb{R}^N sind äquivalent.

Beweis. Sei $\|\cdot\|$ eine beliebige Norm auf \mathbb{R}^N . Wir zeigen: $\|\cdot\| \sim |\cdot|_\infty$. Für $x = (x_1, \dots, x_N) = \sum_{i=1}^N x_i e_i \in \mathbb{R}^N$ gilt

$$(4.4) \quad \|x\| \leq \sum_{i=1}^N |x_i| \|e_i\| \leq |x|_\infty \underbrace{\sum_{i=1}^N \|e_i\|}_{=: c_2} = c_2 |x|_\infty.$$

Noch zu zeigen:

$$(4.5) \quad \text{Es gibt } c_1 > 0 \text{ mit } c_1 |x|_\infty \leq \|x\| \text{ für alle } x \in \mathbb{R}^N \setminus \{0\}.$$

Angenommen, dies gilt nicht. Dann existieren $y_k \in \mathbb{R}^N \setminus \{0\}$, $k \in \mathbb{N}$ mit $\frac{1}{k} |y_k|_\infty \geq \|y_k\|$. Setze $z_k := \frac{y_k}{|y_k|_\infty}$. Dann ist

$$|z_k|_\infty = 1 \quad \text{und} \quad \|z_k\| = \frac{\|y_k\|}{|y_k|_\infty} \leq \frac{1}{k} \quad \text{für alle } k \in \mathbb{N}.$$

Somit folgt $\lim_{k \rightarrow \infty} z_k = 0$ bzgl. $\|\cdot\|$. Nach Satz 4.23 existiert $a \in \mathbb{R}^N$ und eine Teilfolge $(z_{k_j})_j$ mit $\lim_{j \rightarrow \infty} |z_{k_j} - a|_\infty = 0$ und somit $\lim_{j \rightarrow \infty} \|z_{k_j} - a\| = 0$ wegen (4.4). Es folgt $a = 0$. Andererseits hat man

$$|1 - |a|_\infty| = \left| |z_{k_j}|_\infty - |a|_\infty \right| \leq |z_{k_j} - a|_\infty \rightarrow 0 \quad \text{für } j \rightarrow \infty.$$

Also ist $|a|_\infty = 1$ im Widerspruch zu $a = 0$. Somit folgt (4.5), und aus (4.4) und (4.5) folgt: $\|\cdot\| \sim |\cdot|_\infty$. Mit der Transitivität von \sim folgt die Äquivalenz von je zwei Normen auf \mathbb{R}^N . □

Bemerkung 4.26. Es ist einfach, Satz 4.25 mittels eines linearen Isomorphismus auf allgemeine endlichdimensionale \mathbb{R} -Vektorräume auszuweiten.

Bemerkung 4.27 (Topologie des \mathbb{R}^N). Die Eigenschaften „Offenheit“ und „Abgeschlossenheit“ sowie die Folgenkonvergenz ändern sich bei Übergang zu einer äquivalenten Norm **nicht**. Insbesondere ist die Topologie des \mathbb{R}^N nach Satz 4.25 unabhängig von der gewählten Norm $\|\cdot\|$. Mit der $|\cdot|_\infty$ -Norm sieht man, dass eine Folge (x_k) in \mathbb{R}^N genau dann konvergiert, wenn sie komponentenweise konvergiert, d.h. wenn alle Koordinatenfolgen $(x_k^i)_k$ in \mathbb{R} konvergieren.

Falls nicht anders angegeben, betrachten wir \mathbb{R}^N ab jetzt immer mit der euklidischen Norm $|\cdot|_2$, schreiben sie aber ohne index: $|\cdot| := |\cdot|_2$.

4.3 Stetige Abbildungen

Im Folgenden seien (X, d_X) , (Y, d_Y) und (Z, d_Z) metrische Räume.

Definition 4.28 (Abbildungsgrenzwert). Sei $A \subseteq X$, $f: A \rightarrow Y$ eine Abbildung und $a \in X$ ein Häufungspunkt von A . Ein Punkt $b \in Y$ heißt Grenzwert von f in a , wenn für alle Folgen $(x_n)_n$ in $A \setminus \{a\}$ mit $\lim_{n \rightarrow \infty} x_n = a$ gilt: $\lim_{n \rightarrow \infty} f(x_n) = b$. Man schreibt dann $b = \lim_{x \rightarrow a} f(x)$.

Definition 4.29 (Stetigkeit). (a) Sei $f: X \rightarrow Y$ eine Abbildung. f heißt *stetig in einem Punkt* $a \in X$, falls $\lim_{x \rightarrow a} f(x) = f(a)$ gilt. f heißt *stetig* (oder genauer: *stetig in X*), falls f in allen Punkten aus X stetig ist.

(b) Ist $M \subseteq X$, $f: M \rightarrow Y$ eine Abbildung und $a \in M$, so heißt f *stetig in a* , wenn $f|_M$ in a stetig ist. Man betrachtet in diesem Fall in der Stetigkeitsdefinition also den metrischen Teilraum (M, d_X) anstelle von (X, d_X) .

Bemerkung 4.30. Für eine Abbildung $f: X \rightarrow Y$ und $a \in X$ sind äquivalent:

- (i) f ist stetig in a
- (ii) Für alle $\varepsilon > 0$ existiert $\delta > 0$ mit $f(B_\delta(a)) \subseteq B_\varepsilon(f(a))$.
- (iii) Für jede Umgebung $V \subseteq Y$ von $f(a)$ ist $f^{-1}(V) \subseteq X$ eine Umgebung von a .

(ii) kann offensichtlich auch folgendermaßen ausgedrückt werden: Für alle $\varepsilon > 0$ existiert $\delta > 0$ derart, dass für alle $x \in X$ mit $d_X(x, a) < \delta$ gilt: $d_Y(f(x), f(a)) < \varepsilon$.

Satz 4.31. Sei $a \in X$, $\lambda, \mu \in \mathbb{R}$. Sind $f, g: X \rightarrow \mathbb{R}$ in a stetig, so auch

- (a) $\lambda f + \mu g: X \rightarrow \mathbb{R}$.
- (b) $f \cdot g: X \rightarrow \mathbb{R}$.
- (c) $\frac{f}{g}: D_g \rightarrow \mathbb{R}$, falls $a \in D_g := \{x \in X \mid g(x) \neq 0\}$.

Beweis. Wie in Mathe 1. □

Beispiel 4.32. (a) Aus Bemerkung 4.27 folgt direkt die Stetigkeit der Projektionsabbildungen

$$P_i: \mathbb{R}^N \rightarrow \mathbb{R}, \quad P_i(x) = x_i \quad (i = 1, \dots, N).$$

(b) Jede Polynomfunktion

$$P: \mathbb{R}^N \rightarrow \mathbb{R}, \quad P(x) = \sum_{i=1}^k \lambda_i x_1^{\alpha_{i1}} \cdots x_n^{\alpha_{in}}$$

mit $k \in \mathbb{N}$, $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ und $\alpha_{ij} \in \mathbb{N}_0$ für $1 \leq i \leq k$, $1 \leq j \leq N$ ist stetig. Dies folgt aus (a) und Satz 4.31.

Bemerkung 4.33. Eine Abbildung $f: X \rightarrow \mathbb{R}^N$ ist genau dann in $a \in X$ stetig, wenn ihre Komponentenfunktionen $f_i: X \rightarrow \mathbb{R}$, $i = 1, \dots, n$, in a stetig sind.

Beweis. Sei $(x_k)_k$ eine Folge in X mit $\lim_{k \rightarrow \infty} x_k = a$. Dann:

$$\lim_{k \rightarrow \infty} f(x_k) = f(a) \stackrel{\text{Bemerkung 4.27}}{\Leftrightarrow} \lim_{k \rightarrow \infty} f_i(x_k) = f_i(a) \quad \text{für } i = 1, \dots, n.$$

Dies liefert die Behauptung. □

Satz 4.34. Seien $f: X \rightarrow Y$, $g: Y \rightarrow Z$ Abbildungen und $a \in X$. Ist f in a stetig und g in $f(a)$ stetig, so ist $g \circ f: X \rightarrow Z$ in a stetig.

Beweis. Wie in Mathe 1. □

Satz 4.35. Für $f: X \rightarrow Y$ sind folgende Eigenschaften äquivalent.

- (i) f stetig.
- (ii) Für jede offene Menge $V \subseteq Y$ ist $f^{-1}(V) \subseteq X$ offen.
- (iii) Für jede abgeschlossene Menge $V \subseteq Y$ ist $f^{-1}(V) \subseteq X$ abgeschlossen.

Beweis. **(i) \Rightarrow (iii):** Sei $V \subseteq Y$ abgeschlossen, und sei $(x_n)_n$ eine Folge in $f^{-1}(V)$ mit $x_n \rightarrow a \in X$ für $n \rightarrow \infty$. Gemäß Korollar 4.20 ist zu zeigen: $a \in f^{-1}(V)$. Da f in a stetig ist, folgt $\lim_{n \rightarrow \infty} f(x_n) = f(a)$, wobei $f(x_n) \in V$ für alle n . Mit Korollar 4.20 folgt: $f(a) \in V$, also $a \in f^{-1}(V)$.

(iii) \Rightarrow (ii): Ist $V \subseteq Y$ offen, so ist $Y \setminus V \subseteq Y$ abgeschlossen, also nach (iii) auch

$$f^{-1}(Y \setminus V) = X \setminus f^{-1}(V)$$

als Teilmenge von X . Also ist $f^{-1}(V) \subseteq X$ offen.

(ii) \Rightarrow (i): Sei $a \in X$ und $\varepsilon > 0$. Da $B_\varepsilon(f(a)) \subseteq Y$ offen ist, ist mit (ii) auch

$$f^{-1}\left(B_\varepsilon(f(a))\right) \subseteq X \quad \text{offen.}$$

Somit existiert $\delta > 0$ mit $B_\delta(a) \subseteq f^{-1}\left(B_\varepsilon(f(a))\right)$, also $f(B_\delta(a)) \subseteq B_\varepsilon(f(a))$. Mit Bemerkung 4.30 folgt: f ist stetig in a . □

Bemerkung 4.36. Satz 4.35 ist sehr nützlich, um Offenheit oder Abgeschlossenheit von Teilmengen des \mathbb{R}^N zu beweisen. So ist z.B. die Menge

$$A := \{(x, y) \in \mathbb{R}^2 \mid 14x^2 + 3y^4 - 2xy < 5\} \subseteq \mathbb{R}^2$$

offen, da sie sich als Urbild $A = f^{-1}((-\infty, 0))$ der offenen Teilmenge $(-\infty, 0) \subseteq \mathbb{R}$ unter der stetigen Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = 14x^2 + 3y^4 - 2xy - 5$$

schreiben lässt (vgl. Beispiel 4.32(b)).

Definition 4.37. Eine Abbildung $f: X \rightarrow Y$ heißt

- *gleichmäßig stetig*, falls für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert derart, dass für alle $x, y \in X$ die Implikation gilt:

$$d_X(x, y) < \delta \quad \Rightarrow \quad d_Y(f(x), f(y)) < \varepsilon.$$

- *Lipschitz-stetig*, falls ein $L \geq 0$ existiert derart, dass

$$(4.6) \quad d_Y(f(x), f(y)) \leq L d_X(x, y) \quad \text{für alle } x, y \in X.$$

Eine Zahl L , für die (4.6) gilt, nennt man auch *Lipschitz-Konstante von f* .

Bemerkung 4.38. Für Abbildungen $f: X \rightarrow Y$ gelten folgende Implikationen:

$$f \text{ Lipschitz-stetig} \quad \Rightarrow \quad f \text{ gleichmäßig stetig} \quad \Rightarrow \quad f \text{ stetig.}$$

4.4 Lineare stetige Abbildungen

Im Folgenden seien V, W stets normierte \mathbb{R} -Vektorräume mit Normen $\|\cdot\|_V$ und $\|\cdot\|_W$. Ferner sei $\text{Hom}(V, W) := \{T: V \rightarrow W \mid T \text{ } \mathbb{R}\text{-linear.}\}$.

Satz 4.39. Sei $T \in \text{Hom}(V, W)$ und

$$\|T\| := \sup\{\|Tx\|_W \mid x \in V, \|x\|_V \leq 1\} \quad \in [0, \infty].$$

Dann sind äquivalent:

- (i) T ist Lipschitz-stetig.
- (ii) T ist in 0 stetig.
- (iii) $\|T\| < \infty$.

Beachte: Wie man leicht sieht, ist $\|T\|$ die kleinste Zahl in $[0, \infty]$ derart, dass die Ungleichung

$$(4.7) \quad \|Tx\|_W \leq \|T\| \|x\|_V \quad \text{für alle } x \in V$$

gilt (unter der Vereinbarung $\infty \cdot 0 = 0$ und $\infty \cdot a = \infty$ für $a > 0$). Ist $V \neq \{0\}$, so ist

$$\|T\| = \sup_{x \in V \setminus \{0\}} \frac{\|Tx\|_W}{\|x\|_V} = \sup_{x \in V \setminus \{0\}} \left\| T \left(\frac{x}{\|x\|_V} \right) \right\|_W.$$

Beweis. „(i) \Rightarrow (ii)“: trivial.

„(ii) \Rightarrow (iii)“: Zu $\varepsilon = 1$ existiert $\delta > 0$ mit $T(B_\delta(0)) \subseteq B_1(T(0)) = B_1(0)$, d.h.

$$\|Tx\|_W < 1 \quad \text{für alle } x \in V \text{ mit } \|x\|_V < \delta.$$

Sei nun $x \in V$ mit $\|x\|_V \leq 1$. Dann ist $\|\frac{\delta}{2}x\|_V \leq \frac{\delta}{2} < \delta$, und somit

$$\|Tx\|_W = \frac{2}{\delta} \|T(\frac{\delta}{2}x)\|_W < \frac{2}{\delta}.$$

Es folgt: $\|T\| \leq \frac{2}{\delta} < \infty$.

„(iii) \Rightarrow (i)“: Für $x, y \in V$ gilt wegen (4.7):

$$\|Tx - Ty\|_W = \|T(x - y)\|_W \leq \|T\| \|x - y\|_V.$$

Also ist T Lipschitz-stetig mit Lipschitz-Konstante $\|T\|$. □

Ende der Sitzung vom 28.5.

Satz und Definition 4.40. Die Menge $\mathcal{L}(V, W) := \{T \in \text{Hom}(V, W) \mid T \text{ stetig}\}$ ist ein normierter Raum bzgl. der in Satz 4.39 definierten Norm

$$\|\cdot\|: \mathcal{L}(V, W) \rightarrow [0, \infty) \quad (\text{Operatornorm}).$$

Wir betrachten $\mathcal{L}(V, W)$ stets mit dieser Norm. Ist $V = W$, so schreiben wir $\mathcal{L}(V)$ anstelle von $\mathcal{L}(V, V)$.

Beweis (Normeigenschaften). Seien $S, T \in \mathcal{L}(V, W)$. Dann gilt:

$$\|T\| = 0 \Leftrightarrow \|Tx\|_W = 0 \quad \text{für alle } x \in V \Leftrightarrow T = 0.$$

Für $\lambda \in \mathbb{R}$ ist ferner

$$\|\lambda T\| = \sup\{\underbrace{\|\lambda Tx\|_W}_{|\lambda| \|Tx\|_W} \mid \|x\|_V \leq 1\} = |\lambda| \sup\{\|Tx\|_W \mid \|x\|_V \leq 1\} = |\lambda| \|T\| < \infty,$$

insbesondere ist also $\lambda T \in \mathcal{L}(V, W)$. Schließlich ist für alle $x \in V$

$$\|(T + S)x\|_W = \|Tx + Sx\|_W \leq \|Tx\|_W + \|Sx\|_W \leq (\|T\| + \|S\|) \|x\|_V$$

also $\|T + S\| \leq \|T\| + \|S\| < \infty$, insbesondere ist also $S + T \in \mathcal{L}(V, W)$. □

Satz 4.41 (Submultiplikativität der Operatornorm). Sei Z ein weiterer normierter Raum mit Norm $\|\cdot\|_Z$, und seien $T \in \mathcal{L}(V, W)$ und $S \in \mathcal{L}(W, Z)$. Dann ist

$$ST \in \mathcal{L}(V, Z) \quad \text{und} \quad \|ST\| \leq \|S\| \|T\|$$

Beweis. Für $x \in V$ ist

$$\|STx\|_Z \leq \|S\| \|Tx\|_W \leq \|S\| \|T\| \|x\|_V,$$

also folgt $\|ST\| \leq \|S\| \|T\|$. □

Satz 4.42. Ist $\dim V < \infty$, so ist jede lineare Abbildung $T: V \rightarrow W$ stetig.

Beweis. Sei $T \in \text{Hom}(V, W)$. Wir definieren eine weitere Norm $\|\cdot\|_T$ auf V durch

$$\|x\|_T = \|x\|_V + \|Tx\|_W \quad \text{für } x \in V \quad (\text{sog. Graphennorm}).$$

Die Normeigenschaften sind klar. Aus Bemerkung 4.26 folgt: $\|\cdot\|_T \sim \|\cdot\|_V$. Insbesondere existiert also $C > 0$ mit

$$\|x\|_T \leq C\|x\|_V \quad \text{für } x \in V, \text{ also } \|Tx\|_W \leq (C-1)\|x\|_V \quad \text{für } x \in V.$$

Es folgt $T \in \mathcal{L}(V, W)$ mit $\|T\| \leq C-1$. □

Beispiel 4.43. Wir betrachten $L \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^m)$ und die Operatornorm

$$\|L\| := \sup_{x \in \mathbb{R}^N \setminus \{0\}} \frac{|Lx|}{|x|}$$

bzgl. der zugrundeliegenden Norm $|\cdot|$ sowohl auf \mathbb{R}^N als auch auf \mathbb{R}^m . Wir setzen $T := L^*L \in \mathcal{L}(\mathbb{R}^N)$. Hier ist $L^* = L^t$ der adjungierte Operator, wie in Definition 3.63. Wir lassen aber unterschiedliche Dimensionen im Urbild- und Bildraum zu. Dann ist T symmetrisch, denn

$$\langle L^*Lx, y \rangle = \langle Lx, Ly \rangle = \langle x, L^*Ly \rangle \quad \text{für alle } x, y \in \mathbb{R}^N.$$

Wegen Satz 3.65 besitzt \mathbb{R}^N eine Orthonormalbasis von Eigenvektoren v_1, \dots, v_N von T zu reellen Eigenwerten $\lambda_1 \leq \dots \leq \lambda_N$. Diese Eigenwerte sind allesamt nichtnegativ, denn es gilt

$$\lambda_1|v_1|^2 = \langle \lambda_1 v_1, v_1 \rangle = \langle Tv_1, v_1 \rangle = \langle Lv_1, Lv_1 \rangle = |Lv_1|^2.$$

Ist nun $x = \sum_{i=1}^N \mu_i v_i \in \mathbb{R}^N$, so folgt aus der Orthonormalität

$$|Lx|^2 = \langle Tx, x \rangle = \left\langle \sum_{i=1}^N \mu_i \lambda_i v_i, \sum_{i=1}^N \mu_i v_i \right\rangle = \sum_{i=1}^N \lambda_i \mu_i^2 \leq \lambda_N \sum_{i=1}^N \mu_i^2 = \lambda_N |x|^2,$$

wobei für $x = v_N$ Gleichheit gilt. Es folgt

$$\|L\| = \sqrt{\lambda_N}.$$

Mit anderen Worten: Die Operatornorm von L bzgl. $|\cdot|$ ist die Wurzel des größten Eigenwertes von $L^*L \in \mathcal{L}(\mathbb{R}^N)$.

Im Spezialfall $N = m$ und L symmetrisch erhält man

$$\|L\| = \max\{|\lambda| \mid \lambda \text{ Eigenwert von } L\},$$

denn L^*L hat als Eigenwerte die Quadrate der Eigenwerte von L .

4.5 Kompaktheit

Im Folgenden sei (X, d) stets ein metrischer Raum. Zur Motivation dieses Kapitels betrachten wir folgende Frage: Unter welchen Bedingungen an X nimmt jede stetige Funktion $f: X \rightarrow \mathbb{R}$ ein Maximum und ein Minimum an? Wir wissen bereits aus Mathe 1, dass dies für „kompakte“ Intervalle $X = [a, b] \subseteq \mathbb{R}$ gilt. Eine allgemeinere Antwort auf diese Frage erhält man vermöge einer allgemeinen Betrachtung des Kompaktheitsbegriffs.

Definition 4.44. (a) (X, d) heißt *kompakt*, wenn jede Folge $(x_k)_k$ in X eine in X konvergente Teilfolge besitzt.

(b) Eine Teilmenge $A \subseteq X$ heißt *kompakt*, wenn der metrische Raum (A, d) kompakt ist.

Beispiel 4.45. (a) $A := \{\frac{1}{n} \mid n \in \mathbb{N}\} \subseteq \mathbb{R}$ ist nicht kompakt: Betrachte die Folge $x_n := 1/n$ für $n \in \mathbb{N}$. Dann gilt $(x_n) \subseteq A$, aber $\lim_{n \rightarrow \infty} x_n = 0 \notin A$. Somit besitzt (x_n) keine in A konvergente Teilfolge, d.h. mit Grenzwert in A .

(b) $A_0 := A \cup \{0\} \subseteq \mathbb{R}$ ist kompakt: Sei (x_n) eine Folge in A_0 . Falls die Menge der Folgenglieder endlich ist, dann besitzt (x_n) eine konstante Teilfolge. Diese konvergiert also in A_0 . Andernfalls besitzt (x_n) eine streng monoton fallende Teilfolge, welche gegen 0, also in A_0 , konvergiert.

Satz 4.46. Sei $A \subseteq X$ kompakt. Dann gilt:

(a) A ist beschränkt und abgeschlossen in X .

(b) Ist $B \subseteq A$ abgeschlossen, so ist auch B kompakt.

Beweis. (a): **A ist beschränkt:** Angenommen, A wäre unbeschränkt. Für ein beliebiges festes $z \in A$ existiert dann eine Folge $(x_n) \subseteq A$ mit $d(x_n, z) \rightarrow \infty$ für $n \rightarrow \infty$ (andernfalls wäre $A \subseteq B_R(z)$ mit einem geeigneten $R > 0$, also A beschränkt). Da A kompakt ist, hätte (x_n) eine konvergente Teilfolge (x_{n_k}) mit Grenzwert $x^* \in A$. Es würde also gelten

$$\infty = \lim_{k \rightarrow \infty} d(x_{n_k}, z) = d(x^*, z) < \infty,$$

ein Widerspruch. Hier haben wir verwendet, dass die Abbildung $x \mapsto d(x, z)$ Lipschitzstetig ist (siehe Übungen). Also muss A beschränkt sein.

A ist abgeschlossen: Sei $(x_k)_k$ eine Folge in A mit $x_k \rightarrow x^* \in X$ für $k \rightarrow \infty$. Wegen der Kompaktheit von A existiert eine Teilfolge $(x_{k_j})_j$ mit $x_{k_j} \rightarrow y^* \in A$ für $j \rightarrow \infty$. Dann ist aber $x^* = y^* \in A$, und mit Korollar 4.20 folgt die Abgeschlossenheit von A .

(b): Sei $(x_k)_k$ eine Folge in B . Da A kompakt ist, existiert eine Teilfolge $(x_{k_j})_j$ mit $x_{k_j} \rightarrow x^* \in A$ für $j \rightarrow \infty$. Da $B \subseteq A$ abgeschlossen ist, folgt mit Korollar 4.20: $x^* \in B$. Also hat jede Folge in B eine in B konvergente Teilfolge, d.h. B ist kompakt. \square

Satz 4.47 (Heine-Borel). Sei $A \subseteq \mathbb{R}^N$. Dann ist A kompakt genau dann, wenn A beschränkt und abgeschlossen ist.

Beweis. „ \Rightarrow “: Satz 4.46(a).

„ \Leftarrow “: Sei $(x_k)_k$ eine Folge in A . Da A beschränkt ist, existiert $C > 0$ mit $|x_k|_\infty \leq C$ für alle $k \in \mathbb{N}$. Nach Satz 4.23 existiert also eine Teilfolge $(x_{k_j})_j$ mit $x_{k_j} \rightarrow x^* \in \mathbb{R}^N$ für $j \rightarrow \infty$. Da ferner $A \subseteq \mathbb{R}^N$ abgeschlossen ist, folgt $x^* \in A$. Also hat jede Folge in A eine in A konvergente Teilfolge, und dies war zu zeigen. \square

Satz 4.48. *Sei (Y, d) ein weiterer metrischer Raum und $f: X \rightarrow Y$ stetig. Ist (X, d) kompakt, so gilt:*

- (a) $f(X)$ ist auch kompakt.
- (b) f ist gleichmäßig stetig.
- (c) Ist $Y = \mathbb{R}$, so nimmt f auf X ein Maximum und ein Minimum an.

Beweis. **(a):** Sei $(y_n) \subseteq f(X)$. Dann existiert $(x_n) \subseteq X$ mit $y_n = f(x_n)$. Weil X kompakt ist, hat (x_n) eine konvergente Teilfolge $(x_{n_k})_k$ mit $x^* := \lim_{k \rightarrow \infty} x_{n_k}$. Aus der Stetigkeit von f folgt $\lim_{k \rightarrow \infty} y_{n_k} = \lim_{k \rightarrow \infty} f(x_{n_k}) = f(x^*) \in f(X)$, d.h. (y_n) hat eine in $f(X)$ konvergente Teilfolge. Da $(y_n) \subseteq f(X)$ beliebig war, ist $f(X)$ kompakt.

(b): Angenommen, dies gilt nicht. Dann existieren $\varepsilon > 0$ und Folgen $(x_n)_n, (y_n)_n$ in X mit

$$(4.8) \quad \lim_{n \rightarrow \infty} d(x_n, y_n) = 0 \quad \text{und} \quad d(f(x_n), f(y_n)) \geq \varepsilon \quad \text{für alle } n \in \mathbb{N}.$$

Da X kompakt ist, existiert eine Teilfolge $(x_{n_j})_j$ von $(x_n)_n$ und $x^* \in X$ mit $\lim_{j \rightarrow \infty} x_{n_j} = x^*$. Wegen (4.8) folgt dann $\lim_{j \rightarrow \infty} y_{n_j} = x^*$ aus $d(y_{n_j}, x^*) \leq d(y_{n_j}, x_{n_j}) + d(x_{n_j}, x^*)$. Da f in x^* stetig ist, folgt schließlich

$$\lim_{j \rightarrow \infty} f(x_{n_j}) = f(x^*) = \lim_{j \rightarrow \infty} f(y_{n_j}).$$

Dies liefert

$$\lim_{j \rightarrow \infty} d(f(x_{n_j}), f(y_{n_j})) \leq \lim_{j \rightarrow \infty} d(f(x_{n_j}), f(x^*)) + \lim_{j \rightarrow \infty} d(f(x^*), f(y_{n_j})) = 0,$$

im Widerspruch zu $d(f(x_{n_j}), f(y_{n_j})) \geq \varepsilon$ für alle $n \in \mathbb{N}$. Die Behauptung folgt.

(c): Nach (a) ist $f(X) \subseteq \mathbb{R}$ kompakt und somit beschränkt und abgeschlossen nach Satz 4.46. Insbesondere ist

$$-\infty < \inf f(X) \leq \sup f(X) < \infty.$$

Zudem existieren Folgen $(x_k)_k, (y_k)_k$ in $f(X)$ mit $\lim_{k \rightarrow \infty} x_k = \inf f(X)$ und $\lim_{k \rightarrow \infty} y_k = \sup f(X)$. Mit Korollar 4.20 folgt nun

$$\inf f(X), \sup f(X) \in f(X),$$

und dies ist die Behauptung. \square

Bemerkung 4.49. Sei (Y, d) ein weiterer metrischer Raum. Eine bijektive Abbildung $f: X \rightarrow Y$ heißt *Homöomorphismus*, wenn f und f^{-1} stetig sind. Nach Satz 4.35 und Satz 4.48(a) gilt für ein solches f und beliebige Teilmengen $A \subseteq X$:

- (a) A ist offen in X genau dann, wenn $f(A)$ offen in Y ist.
- (b) A ist abgeschlossen in X genau dann, wenn $f(A)$ abgeschlossen in Y ist.
- (c) A ist kompakt genau dann, wenn $f(A)$ kompakt ist.

Bemerkung 4.50. Die Einheitskugel $\overline{B}_1(0)$ des normierten Raums $(C([0, 1]), \|\cdot\|_\infty)$ ist beschränkt und abgeschlossen, aber nicht kompakt: Die Folge der Funktionen $f_n \in \overline{B}_1(0)$, $f_n(x) := x^n$, hat keine bzgl. $\|\cdot\|_\infty$ konvergente (d.h. gleichmäßig konvergente) Teilfolge.

5 Differenzierbarkeit

5.1 Richtungsableitungen und partielle Ableitungen

Ist $I \subseteq \mathbb{R}$ offen, $a \in I$ und $f: I \rightarrow \mathbb{R}$ differenzierbar in a , so beschreibt die Ableitung

$$f'(a) = \lim_{t \rightarrow 0} \frac{f(a+t) - f(a)}{t}$$

die lokale Änderungsrate bzw. geometrisch die Steigung von f bei a . Ist f auf einer offenen Teilmenge $U \subseteq \mathbb{R}^N$ definiert und $a \in U$, so kann man die Änderungsrate/Steigung von f in verschiedene Richtungen betrachten.

Definition 5.1. Sei $U \subseteq \mathbb{R}^N$ offen, $f: U \rightarrow \mathbb{R}^m$ eine Abbildung und $a \in U$.

(a) Sei $v \in \mathbb{R}^N$. Existiert

$$\partial_v f(a) := \lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t} \in \mathbb{R}^m,$$

so heißt $\partial_v f(a)$ die *Richtungsableitung* von f bei a in Richtung v .

(b) Sei $e_j \in \mathbb{R}^N$ der j -te Einheitsvektor, $j = 1 \dots N$. Existiert

$$\partial_{e_j} f(a) = \lim_{t \rightarrow 0} \frac{f(a+te_j) - f(a)}{t},$$

so heißt $\partial_{e_j} f(a)$ die j -te *partielle Ableitung* von f in a .

Wir schreiben kurz $\partial_j f(a)$ anstelle von $\partial_{e_j} f(a)$. Andere übliche Schreibweisen: $D_j f(a)$, $\frac{\partial f}{\partial x_j}(a)$, $\frac{\partial}{\partial x_j} \Big|_{x=a} f(x)$.

Bemerkung 5.2. Sei $U \subseteq \mathbb{R}^N$ offen, $f: U \rightarrow \mathbb{R}$ eine Funktion und $x \in U$ fest. Dann existiert $\varepsilon > 0$ so, dass die Funktionen

$$g_j: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}, \quad g_j(t) = f(\underbrace{x_1, \dots, x_{j-1}, x_j + t, x_{j+1}, \dots, x_N}_{\in U}), \quad j = 1 \dots N$$

wohldefiniert sind, und im Falle der Existenz der Grenzwerte gilt:

$$\partial_j f(x) = \lim_{t \rightarrow 0} \frac{f(x+te_j) - f(x)}{t} = \lim_{t \rightarrow 0} \frac{g_j(t) - g_j(0)}{t} = g_j'(0).$$

Man betrachtet für die j -te partielle Ableitung also alle x_i , $i \neq j$ als konstant und differenziert bzgl. x_j . Ist ferner $v \in \mathbb{R}^N$ fest gewählt, so existiert $\varepsilon > 0$ (abhängig von v) so, dass auch

$$g_v: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}, \quad g_v(t) = f(\underbrace{x + tv}_{\in U})$$

wohldefiniert ist, und im Falle der Existenz der Grenzwerte gilt $\partial_v f(x) = g'_v(0)$.

Konkretes Beispiel hierzu: Betrachte

$$f: \mathbb{R}^N \setminus \{0\} \rightarrow \mathbb{R}, \quad f(x) = |x| = \sqrt{x_1^2 + \dots + x_N^2}.$$

Für festes $x \in \mathbb{R}^N \setminus \{0\}$ ist dann $\partial_j f(x) = \frac{1}{2\sqrt{x_1^2 + \dots + x_N^2}} 2x_j = \frac{x_j}{|x|}$ für $j = 1, \dots, N$. Für $v \in \mathbb{R}^N$ ist

$$g_v(t) = f(x + tv) = \sqrt{(x_1 + tv_1)^2 + (x_2 + tv_2)^2 + \dots + (x_N + tv_N)^2}$$

und somit $\partial_v f(x) = g'_v(0) = \frac{v_1 x_1 + \dots + v_N x_N}{\sqrt{x_1^2 + \dots + x_N^2}} = \sum_{j=1}^N \partial_j f(x) v_j$.

Ende der Sitzung vom 4.6.

Bemerkung 5.3. Sei $U \subseteq \mathbb{R}^N$ offen, $a \in U$, $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}: U \rightarrow \mathbb{R}^m$ eine Abbildung und $v \in \mathbb{R}^N$. Dann gilt wegen Bemerkung 4.27:

$$\partial_v f(a) \text{ existiert} \Leftrightarrow \partial_v f_i(a) \text{ existiert für } i = 1 \dots m.$$

Im Falle der Existenz gilt

$$\partial_v f(a) = \begin{pmatrix} \partial_v f_1(a) \\ \vdots \\ \partial_v f_m(a) \end{pmatrix} \quad \text{und} \quad \partial_j f(a) = \begin{pmatrix} \partial_j f_1(a) \\ \vdots \\ \partial_j f_m(a) \end{pmatrix} \in \mathbb{R}^m \quad \text{für } j = 1, \dots, N$$

Definition 5.4. Seien $U \subseteq \mathbb{R}^N$ offen, $a \in U$ und $f: U \rightarrow \mathbb{R}^m$ eine Abbildung. Falls alle partiellen Ableitungen $\partial_j f(a) \in \mathbb{R}^m$, $j = 1 \dots N$, existieren, so heißt die Matrix

$$J_f(a) = (\partial_1 f(a) \mid \dots \mid \partial_N f(a)) = \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_N f_1(a) \\ \vdots & & \vdots \\ \partial_1 f_m(a) & \dots & \partial_N f_m(a) \end{pmatrix}$$

Jacobimatrix von f in a . Im Spezialfall $m = 1$ ist der *Gradient von f in a* definiert durch

$$\nabla f(a) := \underbrace{J_f(a)^t}_{\in \mathbb{R}^{1 \times N}} = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_N f(a) \end{pmatrix} \in \mathbb{R}^N.$$

Beispiel 5.5. Betrachte $U := \mathbb{R} \times (0, \infty) \subseteq \mathbb{R}^2$ und

$$f: U \rightarrow \mathbb{R}^2, \quad f(x_1, x_2) = \begin{pmatrix} x_1^3 \log x_2 \\ \sin x_1 \cos x_2 \end{pmatrix}.$$

Dann gilt für $x \in U$:

$$\begin{aligned} \partial_1 f_1(x) &= 3x_1^2 \log x_2, & \partial_2 f_1(x) &= \frac{x_1^3}{x_2} \\ \partial_1 f_2(x) &= \cos x_1 \cos x_2, & \text{und} & \quad \partial_2 f_2(x) = -\sin x_1 \sin x_2, \end{aligned}$$

und somit

$$J_f(x) = \begin{pmatrix} 3x_1^2 \log x_2 & \frac{x_1^3}{x_2} \\ \cos x_1 \cos x_2 & -\sin x_1 \sin x_2 \end{pmatrix}.$$

Bemerkung 5.6. Seien $U \subseteq \mathbb{R}^N$ offen, $a \in U$ und $f: U \rightarrow \mathbb{R}^m$ eine Abbildung. Selbst wenn alle Richtungsableitungen $\partial_v f(a)$, $v \in \mathbb{R}^N$ existieren, wird das lokale Änderungsverhalten von f um a durch diese Ableitungen im Allgemeinen nicht vollständig wiedergegeben. Wir betrachten hierzu als Beispiel die Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} \frac{x_1^2 x_2}{x_1^4 + x_2^2}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Dann existieren alle Richtungsableitungen von f im Punkt $a = 0 \in \mathbb{R}^2$, denn für $v \in \mathbb{R}^2 \setminus \{0\}$, $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, gilt:

$$\partial_v f(0) = \lim_{t \rightarrow 0} \frac{f(tv) - f(0)}{t} = \lim_{t \rightarrow 0} \frac{t^3 v_1^2 v_2}{t(t^4 v_1^4 + t^2 v_2^2)} = \lim_{t \rightarrow 0} \frac{v_1^2 v_2}{t^2 v_1^4 + v_2^2} = \begin{cases} 0, & v_2 = 0, \\ \frac{v_1^2}{v_2}, & v_2 \neq 0. \end{cases}$$

Allerdings ist f in 0 nicht einmal stetig, da

$$\lim_{n \rightarrow \infty} f\left(\frac{1}{n}, \frac{1}{n^2}\right) = \frac{1}{2} \neq f(0) = 0$$

gilt.

5.2 Totale Differenzierbarkeit

Ist $I \subseteq \mathbb{R}$ offen, $a \in I$ und $f: I \rightarrow \mathbb{R}$ in a differenzierbar mit $c := f'(a)$, so ist $L_c: \mathbb{R} \rightarrow \mathbb{R}$, $L_c(t) := ct$, die einzige lineare Abbildung mit

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - L_c(h)}{h} = 0.$$

Diese Eigenschaft kennzeichnet $t \mapsto f(a) + L_c(t - a)$ als diejenige Gerade, welche f in der Nähe von a am besten approximiert (anschaulich: die Tangente). Wir wollen das Konzept der „besten linearen Approximation“ nutzen, um die Differenzierbarkeit und die Ableitung einer Funktion in mehreren Variablen zu definieren.

Satz und Definition 5.7. Seien $U \subseteq \mathbb{R}^N$ offen und $f: U \rightarrow \mathbb{R}^m$ eine Abbildung.

- (a) f heißt differenzierbar (auch: total differenzierbar) in $a \in U$, falls es eine lineare Abbildung $L \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^m)$ gibt mit

$$(5.1) \quad \lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - Lh}{|h|} = 0 \quad \text{in } \mathbb{R}^m.$$

- (b) Gilt (5.1), so existieren alle Richtungsableitungen von f in a , und es gilt

$$(5.2) \quad \partial_v f(a) = Lv \quad \text{für alle } v \in \mathbb{R}^N.$$

Also ist L durch (5.1) eindeutig bestimmt.

- (c) Wir setzen

$$df(a) := L \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^m)$$

und nennen $df(a)$ die Ableitung von f in a (auch totales Differenzial von f in a).

- (d) f heißt differenzierbar, wenn f in allen $a \in U$ differenzierbar ist. In diesem Fall nennt man die Abbildung $df: U \rightarrow \mathcal{L}(\mathbb{R}^N, \mathbb{R}^m)$, $a \mapsto df(a)$ die Ableitung von f .

Beweis von (b). Sei $v \in \mathbb{R}^N \setminus \{0\}$. Dann folgt:

$$\lim_{\substack{t \rightarrow 0 \\ t \in \mathbb{R}}} \left| \frac{f(a+tv) - f(a)}{t} - Lv \right| = |v| \lim_{t \rightarrow 0} \frac{|f(a+tv) - f(a) - L(tv)|}{|tv|} \stackrel{(5.1)}{=} 0,$$

da $tv \rightarrow 0$ in \mathbb{R}^N für $t \rightarrow 0$. Es folgt

$$\partial_v f(a) = \lim_{\substack{t \rightarrow 0 \\ t \in \mathbb{R}}} \frac{f(a+tv) - f(a)}{t} = Lv,$$

wie behauptet. □

Bemerkung 5.8. Sei $U \subseteq \mathbb{R}^N$ offen, $f: U \rightarrow \mathbb{R}^m$ eine Abbildung und $a \in U$.

- (a) In (5.1) kann man $|\cdot|$ durch eine beliebige Norm auf \mathbb{R}^N ersetzen, ohne dass sich die Definition ändert. Dies folgt aus der Äquivalenz aller Normen auf \mathbb{R}^N .

- (b) Man sieht leicht:

f ist differenzierbar in $a \Leftrightarrow f_i: U \rightarrow \mathbb{R}$ ist differenzierbar in a für $i = 1 \dots m$.

Gilt dies, so folgt

$$df(a)v = \begin{pmatrix} df_1(a)v \\ \vdots \\ df_m(a)v \end{pmatrix} \in \mathbb{R}^m \text{ für } v \in \mathbb{R}^N \text{ und } J_f(a) = \begin{pmatrix} J_{f_1}(a) \\ \vdots \\ J_{f_m}(a) \end{pmatrix} \in \mathbb{R}^{m \times N}.$$

Man kann die Komponenten von f also getrennt auf Differenzierbarkeit untersuchen.

- (c) Ist f in a differenzierbar, so gilt $\partial_v f(a) = df(a)v$ für $v \in \mathbb{R}^N$ wegen (5.2), insbesondere also $\partial_i f(a) = df(a)e_i$, für $i = 1 \dots N$. Schreibt man $v = \sum_{j=1}^N v_j e_j$, so folgt

$$df(a)v = \sum_{j=1}^N v_j df(a)e_j = \sum_{j=1}^N v_j \partial_j f(a) = J_f(a) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix}.$$

Also ist $J_f(a)$ die *Darstellungsmatrix* von $df(a) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^m)$ bzgl. der kanonischen Basen von \mathbb{R}^N und \mathbb{R}^m . Um im Matrixkalkül konsistent rechnen zu können, schreiben wir Vektoren im \mathbb{R}^N , welche als „Ableitungsrichtungen“ auftauchen, ab jetzt immer als Spalten. Wir erhalten dann

$$df(a)v = J_f(a)v = \partial_v f(a) \quad \text{für alle } v \in \mathbb{R}^N.$$

- (d) Sei speziell $m = 1$ und $f: U \rightarrow \mathbb{R}$ in a differenzierbar. Dann folgt aus (c)

$$\partial_v f(a) = \langle \nabla f(a), v \rangle \leq |\nabla f(a)| |v| \quad \text{für alle } v \in \mathbb{R}^N.$$

Gleichheit gilt hier speziell für $v = \mu \nabla f(a)$ mit $\mu \geq 0$. Also gibt der Vektor $\nabla f(a)$ die Richtung des steilsten Anstiegs von f an (falls $\nabla f(a) \neq 0$ ist).

Bemerkung 5.9. Jede lineare Abbildung $\in \mathcal{L}(\mathbb{R}, \mathbb{R})$ ist von der Form $L_c: \mathbb{R} \rightarrow \mathbb{R}$, $L_c t = ct$ für $t \in \mathbb{R}$. Ist also $I \subseteq \mathbb{R}$ offen, $a \in I$ und $f: I \rightarrow \mathbb{R}$ eine Funktion, so gilt:

f ist differenzierbar in a im Sinne von Satz und Definition 5.7 mit $df(a) = L_c$

$$\Leftrightarrow \lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - L_c h}{|h|} = 0$$

$$\Leftrightarrow \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = \lim_{h \rightarrow 0} \frac{L_c h}{h} = c$$

$$\Leftrightarrow f \text{ ist in } a \text{ im Sinne von Mathe 1 differenzierbar mit } f'(a) = c.$$

In diesem Fall gilt also $df(a)v = f'(a)v$ für $v \in \mathbb{R}$.

Ist allgemeiner $\gamma: I \rightarrow \mathbb{R}^m$ eine Abbildung, so ist γ in a genau dann differenzierbar, wenn alle Komponenten $\gamma_i: I \rightarrow \mathbb{R}$ in a im Sinne von Mathe 1 differenzierbar sind, und

dann gilt:

$$d\gamma(a)t = t\gamma'(a) \quad \text{für } t \in \mathbb{R} \text{ mit } \gamma'(a) := \begin{pmatrix} \gamma'_1(a) \\ \vdots \\ \gamma'_m(a) \end{pmatrix} \in \mathbb{R}^m.$$

Beispiel 5.10. Sei $L \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^m)$ und $c \in \mathbb{R}^m$, und sei $f: \mathbb{R}^N \rightarrow \mathbb{R}^m$ definiert durch $f(x) = Lx + c$. Diese nennt man *Affin lineare Abbildungen*. Dann ist f differenzierbar mit $df(a) = L$ für alle $a \in \mathbb{R}^N$.

Beweis.

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - Lh}{|h|} = \lim_{h \rightarrow 0} \frac{0}{|h|} = 0 \quad \text{für alle } a \in \mathbb{R}^N. \quad \square$$

Satz 5.11. Sei $U \subseteq \mathbb{R}^N$ offen, $a \in U$ und $f: U \rightarrow \mathbb{R}^m$ in a differenzierbar. Dann ist f in a stetig.

Beweis. Nach Voraussetzung ist $\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - df(a)h}{|h|} = 0$, also auch

$$0 = \lim_{h \rightarrow 0} (f(a+h) - f(a) - df(a)h) = \lim_{h \rightarrow 0} (f(a+h) - f(a)) = \lim_{x \rightarrow a} f(x) - f(a).$$

Also ist f nach Definition 4.28 in a stetig. □

Das wichtigste Kriterium um Differenzierbarkeit zu prüfen liefert der folgende

Satz und Definition 5.12. Seien $U \subseteq \mathbb{R}^N$ offen und $f: U \rightarrow \mathbb{R}^m$ eine Abbildung. Dann sind äquivalent:

- (i) Die partiellen Ableitungen $\partial_j f(a)$, $j = 1, \dots, N$, existieren in jedem Punkt $a \in U$, und die Abbildungen $\partial_j f: U \rightarrow \mathbb{R}^m$, $a \mapsto \partial_j f(a)$, sind stetig.
- (ii) f ist differenzierbar, und die Ableitung $df: U \rightarrow \mathcal{L}(\mathbb{R}^N, \mathbb{R}^m)$, $a \mapsto df(a)$, ist stetig.

Gilt dies, so nennt man f stetig differenzierbar. Wir setzen ferner

$$C^1(U, \mathbb{R}^m) := \{f: U \rightarrow \mathbb{R}^m \mid f \text{ stetig differenzierbar}\} \quad \text{und} \quad C^1(U) := C^1(U, \mathbb{R}).$$

Beweis. [28, Satz 9.21]. □

Bemerkung 5.13. Wegen Satz und Definition 5.12 und Bemerkung 5.8(c) haben wir die Implikationen

$$\begin{aligned} \text{stetig partiell differenzierbar} &\Leftrightarrow \text{stetig differenzierbar} \\ &\Rightarrow \text{differenzierbar} \Rightarrow \text{partiell differenzierbar.} \end{aligned}$$

Bemerkung 5.14. Ist $U \subseteq \mathbb{R}^2$ offen, $a \in U$ und $f: U \rightarrow \mathbb{R}$ in a differenzierbar, so nennt man den Graphen der „Linearisierung“ (d.h. der besten affin linearen Approximation)

$$g_a: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad g_a(v) = f(a) + df(a)(v - a),$$

von f bei a die *Tangentialebene von Graph f im Punkt $(a, f(a)) \in \mathbb{R}^3$* .

Als Beispiel seien $U := \mathbb{R}^2$ und $f: U \rightarrow \mathbb{R}$ definiert durch $f(x) = |x|^2$. Dann ist $\partial_1 f(x) = 2x_1$ und $\partial_2 f(x) = 2x_2$ für $x \in \mathbb{R}^2$. Somit sind $\partial_1 f, \partial_2 f: \mathbb{R}^2 \rightarrow \mathbb{R}$ stetig, also $f \in C^1(U)$. Für $a, v \in \mathbb{R}^2$ gilt ferner

$$df(a)v = J_f(a)v = (2a_1, 2a_2) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 2a_1 v_1 + 2a_2 v_2,$$

also $g_a(v) = |a|^2 + 2a_1(v_1 - a_1) + 2a_2(v_2 - a_2)$. Für $a = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ erhält man z.B. $g_a(v) = 1 + 2(v_1 - 1)$. (Übung: Skizziere f und Graph g_a .)

Satz 5.15 (Kettenregel). *Seien $U \subseteq \mathbb{R}^N$, $V \subseteq \mathbb{R}^m$ offen, $a \in U$ und $f: U \rightarrow \mathbb{R}^m$, $g: V \rightarrow \mathbb{R}^\ell$ Abbildungen, für die gilt:*

- (a) $f(U) \subseteq V$
- (b) f ist in a differenzierbar
- (c) g ist in $f(a)$ differenzierbar

Dann ist auch $g \circ f: U \rightarrow \mathbb{R}^\ell$ in a differenzierbar, und es gilt

$$(5.3) \quad d(g \circ f)(a) = \underbrace{dg(f(a))}_{\in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^\ell)} \cdot \underbrace{df(a)}_{\in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^m)} \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^\ell)$$

sowie

$$(5.4) \quad J_{g \circ f}(a) = \underbrace{J_g(f(a))}_{\in \mathbb{R}^{\ell \times m}} \cdot \underbrace{J_f(a)}_{\in \mathbb{R}^{m \times N}} \in \mathbb{R}^{\ell \times N}.$$

Beweis. [28, Satz 9.15]. □

Beispiel 5.16. Seien

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad f(x) = \begin{pmatrix} \sin(x_1 x_2) \\ x_1 - x_2 \\ x_2^2 \end{pmatrix} \quad \text{und} \quad g: \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad g(y) = \begin{pmatrix} y_1 y_2 \\ y_3 \end{pmatrix}.$$

Dann sind f und g stetig differenzierbar mit

$$J_f(x) = \begin{pmatrix} \cos(x_1 x_2) x_2 & \cos(x_1 x_2) x_1 \\ 1 & -1 \\ 0 & 2x_2 \end{pmatrix} \quad \text{und} \quad J_g(y) = \begin{pmatrix} y_2 & y_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Mit Satz 5.15 folgt zum Beispiel für $a = (0, 1) \in \mathbb{R}^2$, also $f(a) = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$:

$$J_{g \circ f}(a) = J_g(0, -1, 1) J_f(0, 1) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Satz 5.17. Sei $U \subseteq \mathbb{R}^N$ offen, $a \in U$ und $\lambda \in \mathbb{R}$.

- (a) Sind $f, g: U \rightarrow \mathbb{R}^m$ differenzierbar in a , so auch
- $f + g: U \rightarrow \mathbb{R}^m$ mit $d(f + g)(a) = df(a) + dg(a)$.
 - $\lambda f: U \rightarrow \mathbb{R}^m$ mit $d(\lambda f)(a) = \lambda df(a)$.
- (b) Sind $f, g: U \rightarrow \mathbb{R}$ differenzierbar in a , so auch fg mit $d(fg)(a) = g(a)df(a) + f(a)dg(a)$.
- (c) Ist $f: U \rightarrow \mathbb{R} \setminus \{0\}$ differenzierbar in a , so ist auch $\frac{1}{f}: U \rightarrow \mathbb{R}$ differenzierbar in a , mit $d\left(\frac{1}{f}\right)(a) = -\frac{1}{f(a)^2}df(a)$.

Beweis. Übung. □

Ende der Sitzung vom 8.6.

Definition 5.18. Sei $I = [a, b] \subseteq \mathbb{R}$, $a < b$ und $U \subseteq \mathbb{R}^N$ offen.

- (a) Eine stetige Abbildung $\gamma: I \rightarrow \mathbb{R}^N$ heißt *Kurve*. γ heißt C^1 -Kurve, falls γ stetig differenzierbar ist. In diesem Fall definiert man ferner
- für $t \in [a, b]$ den *Geschwindigkeitsvektor* $\gamma'(t) = (\gamma'_1(t), \dots, \gamma'_N(t))^t \in \mathbb{R}^N$ (ebenfalls gebräuchlich ist die Bezeichnung $\dot{\gamma}(t)$).
 - die *Länge* $\ell(\gamma) = \int_a^b |\gamma'|$ von γ .
- (b) Eine stetige Abbildung $f: U \rightarrow \mathbb{R}^N$ heißt *Vektorfeld*.
- (c) Ist $\gamma: [a, b] \rightarrow U$ eine C^1 -Kurve und $f: U \rightarrow \mathbb{R}^N$ ein Vektorfeld, so heißt

$$\int_{\gamma} f := \int_a^b \langle f \circ \gamma, \gamma' \rangle \quad \text{Kurvenintegral von } f \text{ längs } \gamma.$$

Beispiel 5.19. (a) Seien $x, v \in \mathbb{R}^N$. Dann ist $\gamma: [a, b] \rightarrow \mathbb{R}^N, \gamma(t) = x + tv$ eine C^1 -Kurve mit $\gamma'(t) = v$ für alle t und $\ell(\gamma) = |v|(b - a)$.

- (b) $\gamma: [0, 2\pi] \rightarrow \mathbb{R}^2, \gamma(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$ ist eine C^1 -Kurve mit $\gamma'(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$ für $t \in [0, 2\pi]$.
Es folgt: $\ell(\gamma) = \int_0^{2\pi} |\gamma'| dt = \int_0^{2\pi} \underbrace{\sqrt{\sin^2 t + \cos^2 t}}_{=1} dt = 2\pi$.

(c) Ist $U \subseteq \mathbb{R}^N$ offen und $\Phi \in C^1(U)$, so ist $\nabla\Phi: U \rightarrow \mathbb{R}^N, x \mapsto \nabla\Phi(x)$ ein Vektorfeld.

Beispiel: Ist $\Phi: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$ gegeben durch $\Phi(x) = \frac{1}{|x|}$, so ist

$$f(x) := \nabla\Phi(x) = -\frac{x}{|x|^3} \quad \text{für } x \in \mathbb{R}^2 \setminus \{0\}.$$

Sei γ wie in (b). Dann gilt wegen $|\gamma| \equiv 1$:

$$\begin{aligned} \int_{\gamma} f &= \int_0^{2\pi} \langle f \circ \gamma, \gamma' \rangle = \int_0^{2\pi} \langle -\begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \rangle dt \\ &= \int_0^{2\pi} (\cos t \sin t - \sin t \cos t) dt = 0. \end{aligned}$$

Dies ist kein Zufall, wie wir in Satz 5.20 sehen werden.

(d) Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gegeben durch $f(x) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$ und die Kurve γ wie in (b) definiert.

Dann ist

$$\int_{\gamma} f = \int_0^{2\pi} \langle \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \rangle dt = \int_0^{2\pi} (\sin^2 t + \cos^2 t) dt = 2\pi.$$

Eine wichtige Anwendung der Kettenregel ist das Berechnen von Kurvenintegralen für Gradientenvektorfelder:

Satz 5.20. Sei $U \subseteq \mathbb{R}^N$ offen, $\Phi \in C^1(U)$ und $\gamma: [a, b] \rightarrow U$ eine C^1 -Kurve. Dann gilt:

$$\int_{\gamma} \nabla\Phi = \Phi(\gamma(b)) - \Phi(\gamma(a)).$$

Ist γ geschlossen, d.h. gilt $\gamma(b) = \gamma(a)$, so folgt $\int_{\gamma} \nabla\Phi = 0$.

Beweis. Nach dem Hauptsatz der Differential- und Integralrechnung ist $\Phi(\gamma(b)) - \Phi(\gamma(a)) = \int_a^b (\Phi \circ \gamma)' dt$, wobei nach Kettenregel gilt:

$$(\Phi \circ \gamma)'(t) = J_{\Phi \circ \gamma}(t) = J_{\Phi}(\gamma(t))J_{\gamma}(t) = \langle \nabla\Phi(\gamma(t)), \gamma'(t) \rangle \quad \text{für } t \in (a, b).$$

Es folgt

$$\Phi(\gamma(b)) - \Phi(\gamma(a)) = \int_a^b \langle \nabla\Phi \circ \gamma, \gamma' \rangle = \int_{\gamma} \nabla\Phi,$$

wie behauptet. □

Definition 5.21. Für $x, y \in \mathbb{R}^N$ heißt die Punktmenge $[x, y] := \{(1-t)x + ty \mid t \in [0, 1]\}$ die (*Verbindungs-*) *Strecke zwischen x und y* . Eine Teilmenge $U \subseteq \mathbb{R}^N$ heißt *konvex*, wenn für alle $x, y \in U$ gilt: $[x, y] \subseteq U$.

Satz 5.22 (Schrankensatz). Seien $U \subseteq \mathbb{R}^N$ offen und $f \in C^1(U, \mathbb{R}^m)$. Seien ferner $x, y \in U$ mit $[x, y] \subseteq U$. Dann gilt: $|f(y) - f(x)| \leq M|y - x|$, mit $M := \max_{0 \leq t \leq 1} \|df((1-t)x + ty)\|$.

Beweis. Wir setzen $g(t) := f((1-t)x + ty)$ für $t \in [0, 1]$. Dann folgt aus der Kettenregel $g'(t) = df((1-t)x + ty)(y - x)$ und daher mit dem Hauptsatz der Differential- und Integralrechnung und $v := y - x$

$$\begin{aligned} |f(y) - f(x)| &= |g(1) - g(0)| = \left| \int_0^1 g'(t) dt \right| = \left| \int_0^1 df((1-t)x + ty)v dt \right| \\ &\leq \int_0^1 |df((1-t)x + ty)v| dt \leq \int_0^1 M|v| dt \leq M|v|. \quad \square \end{aligned}$$

Bemerkung 5.23. (a) Der Schrankensatz erfüllt eine ähnliche Funktion wie der Mittelwertsatz für reellwertige Funktionen auf Intervallen. Allerdings gibt es für \mathbb{R}^m -wertige Abbildungen, $m \geq 2$, kein zum Mittelwertsatz analoges Resultat. Ist z.B. $\gamma: [0, 2\pi] \rightarrow \mathbb{R}^2$ definiert durch $\gamma(t) = (\cos t, \sin t)$, so ist

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \frac{\gamma(2\pi) - \gamma(0)}{2\pi} \neq \gamma'(\tau) = \begin{pmatrix} -\sin \tau \\ \cos \tau \end{pmatrix} \quad \text{für } \tau \in [0, 2\pi].$$

(b) Falls U konvex ist, dann ist der Schrankensatz auf alle Punkte $x, y \in U$ anwendbar.

5.3 Höhere Ableitungen

Definition 5.24. Sei $U \subseteq \mathbb{R}^N$ offen. Induktiv definieren wir für $k \in \mathbb{N}$, $k \geq 2$: Eine Abb. $f: U \rightarrow \mathbb{R}^m$ heißt *k-mal stetig differenzierbar*, wenn alle partiellen Ableitungen $\partial_j f(x)$, $x \in U$, $j = 1 \dots N$ existieren und die Abbildungen $\partial_j f: U \rightarrow \mathbb{R}^m$ ($k-1$)-mal stetig differenzierbar sind. Wir setzen

$$C^k(U, \mathbb{R}^m) := \{f: U \rightarrow \mathbb{R}^m \mid f \text{ k-mal stetig differenzierbar}\}.$$

Ferner sei $C^0(U, \mathbb{R}^m) := C(U, \mathbb{R}^m) = \{f: U \rightarrow \mathbb{R}^m \mid f \text{ stetig}\}$. Offensichtlich gilt $C^0(U, \mathbb{R}^m) \supseteq C^1(U, \mathbb{R}^m) \supseteq C^2(U, \mathbb{R}^m) \supseteq \dots$. Wir setzen auch

$$C^\infty(U, \mathbb{R}^m) := \bigcap_{k \in \mathbb{N}} C^k(U, \mathbb{R}^m) \quad (\text{beliebig oft stetig differenzierbare Abbildungen}).$$

Schließlich verwenden wir die abkürzende Schreibweise $C^k(U) := C^k(U, \mathbb{R})$ für $k \in \mathbb{N}_0 \cup \{\infty\}$.

Definition 5.25. Sei $U \subseteq \mathbb{R}^N$ offen, $f \in C^2(U)$ und $x \in U$.

(a) $H_f(x) := (\partial_i \partial_j f(x))_{ij} \in \mathbb{R}^{N \times N}$ heißt *Hessematrix* von f in x .

(b) Man setzt $\partial_i^2 := \partial_i \partial_i$ und $\Delta f(x) := \text{Spur } H_f(x) = \sum_{i=1}^N \partial_i^2 f(x)$. Das Symbol $\Delta = \sum_{i=1}^N \partial_i^2$ heißt *Laplace-Operator*.

Beispiel 5.26. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch $f(x) = e^{x_1} x_1 (1 + x_2)$. Dann ist

$$\begin{aligned} \partial_1 f(x) &= e^{x_1} (1 + x_1)(1 + x_2), & \partial_2 f(x) &= x_1 e^{x_1}, \\ \partial_1 \partial_2 f(x) &= e^{x_1} (1 + x_1), & \partial_2 \partial_1 f(x) &= e^{x_1} (1 + x_1), \\ \partial_1^2 f(x) &= e^{x_1} (2 + x_1)(1 + x_2) & \text{und} & \partial_2^2 f(x) = 0. \end{aligned}$$

Es folgt

$$\nabla f(x) = e^{x_1} \begin{pmatrix} (1+x_1)(1+x_2) \\ x_1 \end{pmatrix}, \quad H_f(x) = e^{x_1} \begin{pmatrix} (2+x_1)(1+x_2) & 1+x_1 \\ 1+x_1 & 0 \end{pmatrix},$$

und $\Delta f(x) = e^{x_1} (2 + x_1)(1 + x_2)$ für $x \in \mathbb{R}^2$.

Die Symmetrie von $H_f(x)$ ist allgemein gültig:

Satz 5.27 (von Schwarz). *Seien $U \subseteq \mathbb{R}^N$ offen und $f \in C^2(U, \mathbb{R}^m)$. Dann gilt*

$$\partial_i \partial_j f = \partial_j \partial_i f \quad \text{für } i, j = 1 \dots N.$$

Ist insbesondere $m = 1$, d.h. $f \in C^2(U)$, so ist $H_f(x)$ eine symmetrische Matrix.

Beweis. [28, Satz 9.41]. □

Korollar 5.28. *Ist $f \in C^k(U, \mathbb{R}^m)$, $U \subseteq \mathbb{R}^N$ offen und sind $i_1, \dots, i_k \in \{1, \dots, N\}$, so gilt:*

$$\partial_{i_1} \cdots \partial_{i_k} f = \partial_{i_{\pi(1)}} \cdots \partial_{i_{\pi(k)}} f$$

für jede Permutation (Bijektion) $\pi: \{1, \dots, k\} \rightarrow \{1, \dots, k\}$.

Beweis. Dies folgt aus Satz 5.27 durch vollständige Induktion und daraus, dass sich jede Permutation als Hintereinanderausführung von Transpositionen (d.h. Vertauschungen) schreiben lässt. □

Definition und Bemerkung 5.29. Sei $U \subseteq \mathbb{R}^N$ offen und $f \in C^1(U, \mathbb{R}^N)$ ein Vektorfeld. Eine Funktion $\Phi \in C^2(U)$ heißt *Potential* für das Vektorfeld f , wenn $f = \nabla \Phi$ gilt.

Es ist wichtig zu wissen, ob f ein Potential besitzt, denn in diesem Fall gilt nach Satz 5.20:

$$\int_{\gamma} f = \Phi(\gamma(b)) - \Phi(\gamma(a)) \quad \text{für alle } C^1\text{-Kurven } \gamma: [a, b] \rightarrow U.$$

Notwendig für die Existenz eines Potentials Φ ist nach Satz 5.27 die Gültigkeit folgender *Integrabilitätsbedingungen*:

$$(5.5) \quad \partial_i f_j = \partial_i \partial_j \Phi = \partial_j \partial_i \Phi = \partial_j f_i \quad \text{für alle } i, j = 1, \dots, N.$$

Ist U sternförmig, d.h. gibt es ein $x_0 \in U$ mit $[x_0, y] \subseteq U$ für alle $y \in U$, dann ist (5.5) auch hinreichend für die Existenz eines Potentials. Für einen Beweis letzterer Aussage siehe z.B. [17, Satz 182.2].

Beispiel 5.30. Sei $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definiert durch $F(x) = \begin{pmatrix} x_2^2 \\ x_1 \end{pmatrix}$. Hat F ein Potential? Nein, da $\partial_1 F_2(x) = 1 \neq 2x_2 = \partial_2 F_1(x)$ für $x \in \mathbb{R}^2$ mit $x_2 \neq \frac{1}{2}$.

Definition 5.31. Ein Element $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{N}_0^N$ heißt *Multiindex*. Wir definieren die *Länge von α* als $|\alpha| := \alpha_1 + \dots + \alpha_N$ und die *Fakultät von α* als $\alpha! := \alpha_1! \cdots \alpha_N!$. Für $v \in \mathbb{R}^N$ sei zudem $v^\alpha := v_1^{\alpha_1} \cdots v_N^{\alpha_N}$, und im Falle $f \in C^k(U)$ und $|\alpha| \leq k$ sei $\partial^\alpha f := \partial_1^{\alpha_1} \cdots \partial_N^{\alpha_N} f$ (andere Schreibweise: $D^\alpha f$).

Satz 5.32 (von Taylor). *Seien $U \subseteq \mathbb{R}^N$ offen, $a \in U$ und $f \in C^n(U)$ für ein $n \in \mathbb{N}_0$. Sei ferner*

$$V := \{v \in \mathbb{R}^N \mid a + tv \in U \text{ für } t \in [0, 1]\}.$$

Dann gilt

$$(5.6) \quad f(a+v) = \sum_{|\alpha| \leq n} \frac{\partial^\alpha f(a)}{\alpha!} v^\alpha + R_n(v) \quad \text{für } v \in V,$$

wobei

$$(5.7) \quad \lim_{v \rightarrow 0} \frac{R_n(v)}{|v|^n} = 0$$

gilt. Ist zudem $f \in C^{n+1}(U)$, so existiert $\tau = \tau(v) \in [0, 1]$ mit

$$(5.8) \quad R_n(v) = \sum_{|\alpha|=n+1} \frac{\partial^\alpha f(a + \tau v)}{\alpha!} v^\alpha$$

Beweisskizze. Ohne Einschränkung sei $v \neq 0$. Wir setzen $w := \frac{v}{|v|}$ und $g(t) := f(a + tw)$ für $t \in [0, 1]$. Dann gelten

$$g'(t) = df(a + tw) w = \frac{1}{|v|} \sum_{i=1}^N \partial_i f(a + tw) v_i$$

$$g''(t) = \frac{1}{|v|^2} \sum_{i,j=1}^N \partial_j \partial_i f(a + tw) v_i v_j$$

$$g'''(t) = \frac{1}{|v|^3} \sum_{i,j,k=1}^N \partial_k \partial_j \partial_i f(a + tw) v_i v_j v_k$$

u.s.w.

Aus dem Satz von Taylor (Mathe 1) für Funktionen auf Intervallen folgt nun

$$f(a+v) = g(|v|) = \sum_{\ell=0}^n \frac{g^{(\ell)}(0)}{\ell!} |v|^\ell + R_n(v),$$

mit $R_n(v)/|v|^n \rightarrow 0$ für $v \rightarrow 0$. Wir setzen die obigen Formeln für die Ableitungen von g ein. Nach Korollar 5.28 ist die Reihenfolge der partiellen Ableitungen irrelevant, so dass wir sie für übereinstimmende Gruppen von Indizes zusammenfassen können. Genaues Abzählen liefert dann die erste Behauptung. Die zweite folgt genauso. \square

Bemerkung 5.33. Seien $U \subseteq \mathbb{R}^N$ offen und $f \in C^2(U)$. Dann gilt für $v, w \in \mathbb{R}^N$:

$$\partial_v f = dfv = \sum_{j=1}^N \partial_j f v_j = \langle \nabla f, v \rangle$$

und

$$\partial_w \partial_v f = \sum_{j=1}^N d(\partial_j f v_j) w = \sum_{i,j=1}^N \partial_i \partial_j f v_j w_i = \langle H_f v, w \rangle = \langle v, H_f w \rangle.$$

Definition und Bemerkung 5.34. Seien $U \subseteq \mathbb{R}^N$ offen, $a \in U$ und $f \in C^k(U)$. Dann heißt

$$T_a^k f: U \rightarrow \mathbb{R}, \quad T_a^k f(x) := \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(a)}{\alpha!} (x-a)^\alpha$$

das k -te Taylorpolynom von f in a . Gemäß Satz 5.32 gilt

$$\lim_{x \rightarrow a} \frac{f(x) - T_a^k f(x)}{|x-a|^k} = 0,$$

und man kann zeigen, dass $T_a^k f$ das einzige Polynom vom Grad kleiner gleich k mit dieser Eigenschaft ist. Speziell erhält man (vgl. Bemerkung 5.33):

- $T_a^1 f(x) = f(a) + \langle \nabla f(a), x-a \rangle$
- $T_a^2 f(x) = f(a) + \langle \nabla f(a), x-a \rangle + \frac{1}{2} \langle H_f(a)(x-a), (x-a) \rangle$

Beispiel 5.35. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch $f(x) = e^{x_1} x_1 (1+x_2)$. Nach Beispiel 5.26 ist

$$\nabla f(x) = e^{x_1} \begin{pmatrix} (1+x_1)(1+x_2) \\ x_1 \end{pmatrix} \quad \text{und} \quad H_f(x) = e^{x_1} \begin{pmatrix} (2+x_1)(1+x_2) & 1+x_1 \\ 1+x_1 & 0 \end{pmatrix},$$

In $0 \in \mathbb{R}^2$ erhält man

$$f(0) = 0, \quad \nabla f(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad H_f(0) = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}.$$

Es folgt

$$\begin{aligned} T_0^2 f(x) &= f(0) + \langle \nabla f(0), x \rangle + \frac{1}{2} \langle H_f(0)x, x \rangle \\ &= 0 + x_1 + \frac{1}{2}(2x_1^2 + x_1x_2 + x_1x_2 + 0) = x_1 + x_1^2 + x_1x_2. \end{aligned}$$

Nun zu $T_0^3 f$: Sei $\alpha \in \mathbb{N}_0^2$ mit $|\alpha| = 3$. Fallunterscheidung:

$$\alpha = (3, 0) \Rightarrow \alpha! = 6 \text{ und } \partial^\alpha f(x) = \partial_1^3 f(x) = \frac{\partial}{\partial x_1} e^{x_1} (2+x_1)(1+x_2) = e^{x_1} (3+x_1)(1+x_2)$$

$$\alpha = (2, 1) \Rightarrow \alpha! = 2 \text{ und } \partial^\alpha f(x) = \partial_1^2 \partial_2 f(x) = \frac{\partial}{\partial x_1} e^{x_1} (1+x_1) = e^{x_1} (2+x_1)$$

$$\alpha = (1, 2) \Rightarrow \partial^\alpha f(x) = \partial_1 \partial_2^2 f(x) = \frac{\partial}{\partial x_1} 0 = 0$$

$$\alpha = (0, 3) \Rightarrow \partial^\alpha f(x) = \partial_2^3 f(x) = 0.$$

Es folgt

$$\sum_{|\alpha|=3} \frac{\partial^\alpha f(0)}{\alpha!} x^\alpha = \frac{1}{2} x_1^3 + x_1^2 x_2,$$

also

$$T_0^3 f(x) = x_1 + x_1^2 + x_1x_2 + \frac{x_1^3}{2} + x_1^2 x_2 \quad \text{für } x \in \mathbb{R}^2.$$

Wir mussten nur 4 und nicht 8 dritte partielle Ableitungen ausrechnen.

Bemerkung 5.36. Andere gebräuchliche Schreibweise für höhere Ableitungen:

$$\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(x) \quad \text{anstelle von} \quad \partial_{i_1} \cdots \partial_{i_k} f(x).$$

Ende der Sitzung vom 15.6.

Definition und Bemerkung 5.37 (Landau-Symbole). Seien $U \subseteq \mathbb{R}^N$, $x^* \in \mathbb{R}^N$ ein Häufungspunkt von U und $f, g: U \rightarrow \mathbb{R}^m$ Funktionen.

(a) $f(x) = o(g(x))$ (gesprochen: $f(x)$ ist klein-O von $g(x)$) für $x \rightarrow x^*$, falls gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in B_\delta(x^*) \cap U: |f(x)| \leq \varepsilon |g(x)|.$$

(b) $f(x) = O(g(x))$ (gesprochen: $f(x)$ ist groß-O von $g(x)$) für $x \rightarrow x^*$, falls gilt:

$$\exists C > 0 \exists \delta > 0 \forall x \in B_\delta(x^*) \cap U: |f(x)| \leq C |g(x)|.$$

(c) Falls $g(x) \neq 0$ für x nahe bei x^* gilt, so ist $f(x) = o(g(x))$ für $x \rightarrow x^*$ äquivalent zu $\lim_{x \rightarrow x^*} \frac{|f(x)|}{|g(x)|} = 0$, und $f(x) = O(g(x))$ für $x \rightarrow x^*$ äquivalent zu $\limsup_{x \rightarrow x^*} \frac{|f(x)|}{|g(x)|} < \infty$.

Bemerkung 5.38. Die Darstellung (5.6) aus dem Satz von Taylor kann man mit Landau-Symbolen jetzt auch wie folgt schreiben:

$$f(a+v) = \sum_{|\alpha| \leq n} \frac{\partial^\alpha f(a)}{\alpha!} v^\alpha + o(|v|^n) \quad \text{für } v \rightarrow 0.$$

5.4 Einblick in die Vektoranalysis

Als *Vektoranalysis* bezeichnet man eine spezielle symbolische Art, die Differentialrechnung mehrerer Veränderlicher zu schreiben, insbesondere in \mathbb{R}^3 . Diese wurde historisch vor allem in der Physik und verwandten Gebieten entwickelt. Die Begriffe können im Rahmen der Differentialgeometrie für alle Raumdimensionen verallgemeinert werden. Ein zentrales Element ist der sogenannte *Nabla-Operator*:

$$\nabla := \begin{pmatrix} \partial_1 \\ \vdots \\ \partial_N \end{pmatrix}.$$

Sei $U \subseteq \mathbb{R}^N$ eine offene Teilmenge. Häufig bezeichnet man in diesem Zusammenhang eine stetige Funktion $f: U \rightarrow \mathbb{R}$ als *Skalarfeld*, um den Gegensatz zu Vektorfeldern zu betonen. Der Nabla-Operator induziert dann eine lineare Abbildung wie folgt:

$$\nabla: C^1(U) \rightarrow C(U, \mathbb{R}^N), \quad f \mapsto \nabla f := \begin{pmatrix} \partial_1 f \\ \vdots \\ \partial_N f \end{pmatrix}.$$

Der Nabla-Operator ordnet einem stetig differenzierbaren Skalarfeld also das entsprechende Gradientenfeld zu.

Andererseits induziert der Nabla-Operator eine lineare Abbildung auf stetig differenzierbaren Vektorfeldern:

$$\nabla: C^1(U, \mathbb{R}^N) \rightarrow C(U), \quad v \mapsto \nabla \cdot v := \operatorname{div} v := \partial_1 v_1 + \partial_2 v_2 + \dots + \partial_N v_N.$$

Hier heißt $\operatorname{div} v$ die *Divergenz von v*. Diese suggestive Schreibweise betont die Analogie zum „Skalarprodukt von ∇ und v “.

Für ein Skalarfeld $f \in C^2(U)$ kann man ∇ auf das Gradientenfeld ∇f anwenden und erhält

$$\nabla \cdot (\nabla f) = \operatorname{div} \nabla f = \Delta f = \sum_{i=1}^N \partial_i^2 f.$$

Also entspricht $\nabla \cdot \nabla$ formal dem linearen Laplace-Operator:

$$\nabla \cdot \nabla = \Delta: C^2(U) \rightarrow C(U), \quad v \mapsto \Delta v.$$

In \mathbb{R}^3 beschreiben wir eine zusätzliche multiplikative Struktur:

Definition 5.39. Seien $v = (v_1, v_2, v_3)$ und $w = (w_1, w_2, w_3)$ Vektoren in \mathbb{R}^3 . Dann wird das *Kreuzprodukt* zwischen v und w definiert als

$$v \times w := \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}.$$

Für stetig differenzierbare Vektorfelder definiert man nun die *Rotation* als linearen Differentialoperator wie folgt:

$$\text{rot}: C^1(U, \mathbb{R}^3) \rightarrow C(U, \mathbb{R}^3), \quad \text{rot } v := \nabla \times v := \begin{pmatrix} \partial_2 v_3 - \partial_3 v_2 \\ \partial_3 v_1 - \partial_1 v_3 \\ \partial_1 v_2 - \partial_2 v_1 \end{pmatrix}.$$

Bemerkung 5.40. (a) Für $v, w \in \mathbb{R}^3$ gilt stets $\langle v, v \times w \rangle = \langle w, v \times w \rangle = 0$, d.h. das Kreuzprodukt zwischen v und w erzeugt einen Vektor, der Senkrecht auf v und w steht.

(b) Ein Vektorfeld $v \in C^1(U, \mathbb{R}^3)$ heißt *rotationsfrei* oder *wirbelfrei*, wenn $\text{rot } v = 0$ gilt, und *divergenzfrei* oder *quellenfrei*, wenn $\text{div } v = 0$ gilt.

(c) Sind $U \subseteq \mathbb{R}^3$ offen, $f \in C^2(U)$ und $v \in C^2(U, \mathbb{R}^3)$, so gilt $\nabla \times (\nabla f) = 0$ und $\nabla \cdot (\nabla \times v) = 0$, wie man leicht mit Hilfe des Satzes von Schwarz nachrechnet. Gradientenfelder sind also rotationsfrei, und die Rotation eines Vektorfeldes ist divergenzfrei.

Die Merkgeln sind also:

- (a) ∇ auf Skalarfeld gibt den Gradient;
- (b) $\nabla \cdot$ auf Vektorfeld gibt die Divergenz;
- (c) $\nabla \times$ auf Vektorfeld gibt die Rotation (in \mathbb{R}^3);
- (d) formal gilt $\nabla \cdot \nabla = \Delta$.

Die Helmholtzzerlegung eines C^2 -Vektorfeldes erlaubt es, dieses als Summe eines divergenzfreien Vektorfeldes und eines rotationsfreien Vektorfeldes zu schreiben. Wir notieren einen Spezialfall genauer:

Satz 5.41 (Helmholtz). Seien $r > 0$ und $v \in C^2(B_{r+1}(0), \mathbb{R}^3)$ ein Vektorfeld. Dann gibt es zwei Vektorfelder $A, B \in C^1(B_r(0), \mathbb{R}^3)$, so dass gilt

$$v = A + B, \quad \nabla \cdot A = 0, \quad \text{und} \quad \nabla \times B = 0.$$

Beweis. [1]. □

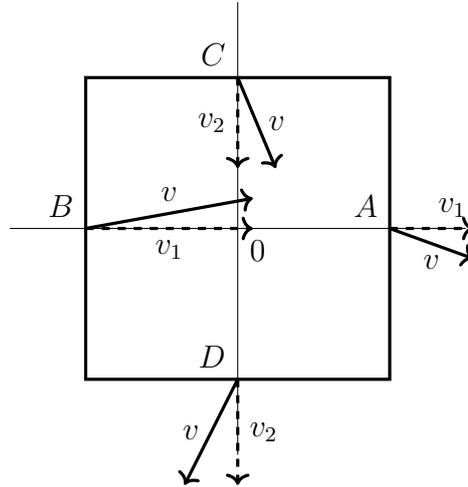


Abbildung 5.1: Fluss durch den Rand eines Quadrats

Wir haben in Bemerkung 5.8(d) schon die anschauliche Interpretation gesehen, dass der Gradient eines Skalarfeldes in die Richtung des Stärksten Anstiegs zeigt.

Wir beschreiben die physikalische Bedeutung der Divergenz eines Vektorfeldes. In der Strömungslehre modelliert man die statische Strömungsdichte einer Flüssigkeit oder eines Gases in einer Teilmenge \mathbb{R}^3 mit Hilfe eines stetig differenzierbaren Vektorfeldes v . In jedem Punkt x beschreibt der Vektor $v(x)$ dabei die Richtung und Flussdichte der Strömung. Letztere bedeutet, wie viel Masse in einer Zeiteinheit durch eine Einheitsquerschnittsfläche strömt. Um das Strömungsverhalten zu verstehen, möchte man berechnen, wie viel Masse in einer Zeiteinheit durch den Rand (die Oberfläche) ∂U eines Volumens $U \subseteq \mathbb{R}^3$ **nach außen** strömt. Eine positive Zahl für diesen Strom bedeutet dabei, dass aus U insgesamt Masse ausströmt, eine negative Zahl bedeutet, dass insgesamt Masse einströmt. Es hat sich als nützlich erwiesen, für diese Untersuchung das Verhältnis aus diesem Oberflächenstrom und dem Volumen $\text{vol}(U)$ von U im Grenzwert $\text{vol}(U) \rightarrow 0$ zu bestimmen.

Wir geben eine heuristische Erklärung in \mathbb{R}^2 dafür, dass dieser Grenzwert die Divergenz des Vektorfeldes ist. Dazu beschränken wir uns auf achsenparallele Quadrate mit Mittelpunkt 0 und Seitenlänge 2ε : In Abb. 5.1 sind einige Werte von v und die jeweiligen auf dem Rand senkrecht stehenden Anteile eingezeichnet, welche in diesem speziellen Fall mit den Koordinaten von v übereinstimmen. Die Punkte der Seite A sind durch $\{\varepsilon e_1 + s e_2 \mid s \in [-\varepsilon, \varepsilon]\}$ gegeben. Der auf der Seite A senkrecht stehende, nach außen zeigende Anteil ist in diesen Punkten durch $v_1(\varepsilon e_1 + s e_2)$ gegeben. Der Gesamtfluss $\mathcal{F}_\varepsilon(A)$ durch die Seite A hindurch nach draußen ist dann nach Approximation durch das erste Taylorpolynom:

$$\begin{aligned} \mathcal{F}_\varepsilon(A) &= \int_{-\varepsilon}^{\varepsilon} v_1(\varepsilon e_1 + s e_2) \, ds = \int_{-\varepsilon}^{\varepsilon} (v_1(0) + \partial_1 v_1(0)\varepsilon + \partial_2 v_1(0)s + o(\varepsilon)) \, ds \\ &= 2\varepsilon v_1(0) + 2\varepsilon^2 \partial_1 v_1(0) + o(\varepsilon^2) \end{aligned}$$

für $\varepsilon \rightarrow 0$. Die Punkte der Seite B sind $\{-\varepsilon e_1 + s e_2 \mid s \in [-\varepsilon, \varepsilon]\}$, und der auf der Seite B senkrecht stehende, nach außen zeigende Anteil ist in diesen Punkten durch $-v_1(-\varepsilon e_1 + s e_2)$ gegeben. Es folgt wie oben

$$\mathcal{F}_\varepsilon(B) = -2\varepsilon v_1(0) + 2\varepsilon^2 \partial_1 v_1(0) + o(\varepsilon^2).$$

Analog erhalten wir

$$\begin{aligned}\mathcal{F}_\varepsilon(C) &= 2\varepsilon v_2(0) + 2\varepsilon^2 \partial_2 v_2(0) + o(\varepsilon^2), \\ \mathcal{F}_\varepsilon(D) &= -2\varepsilon v_2(0) + 2\varepsilon^2 \partial_2 v_2(0) + o(\varepsilon^2).\end{aligned}$$

Der Grenzwert des Verhältnisses von Gesamtfluss zu Quadratfläche ist dann

$$\begin{aligned}\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{F}_\varepsilon(A) + \mathcal{F}_\varepsilon(B) + \mathcal{F}_\varepsilon(C) + \mathcal{F}_\varepsilon(D)}{4\varepsilon^2} \\ = \lim_{\varepsilon \rightarrow 0} \frac{4\varepsilon^2(\partial_1 v_1(0) + \partial_2 v_2(0)) + o(\varepsilon^2)}{4\varepsilon^2} = \partial_1 v_1(0) + \partial_2 v_2(0) = \operatorname{div} v(0).\end{aligned}$$

Dieses Argument lässt sich analog in allen Raumdimensionen durchführen.

Die Bedeutung des Laplace-Operators liegt unter anderem in der Modellierung von Diffusionsprozessen und von höherdimensionalen Wellengleichungen. Im ersten Beispiel folgt für die Diffusion einer Konzentration $\psi: [0, \infty) \times U \rightarrow \mathbb{R}$ in einem Raumgebiet U aus den Fickschen Gesetzen die Gleichung für die zeitliche Entwicklung:

$$\partial_t \psi = \Delta_x \psi.$$

Hier bedeutet Δ_x , dass der Δ -Operator nur auf die x -Variablen wirkt, wenn man $\psi(t, x)$ mit $t \in [0, \infty)$ und $x \in U$ schreibt. Die Gleichung kann auch die zeitliche Änderung einer Temperaturverteilung beschreiben (*Wärmeleitungsgleichung*).

Und schließlich beschreibt die Rotation eines Vektorfeldes v das Verhältnis des Kurvenintegrals über die Randkurve γ einer Fläche $A \subseteq \mathbb{R}^3$ und dem Flächeninhalt $\operatorname{vol}_2(A)$ im Grenzwert $\operatorname{vol}_2(A) \rightarrow 0$. Diese wird analog zur Divergenz verwendet, um Wegintegrale von Vektorfeldern über Randkurven von Flächen zu berechnen.

Die Beziehung zwischen Divergenz und Oberflächenstrom sowie Rotation und Kurvenintegralen werden wir erst mit den Sätzen von Gauß und Stokes in Mathe 3 vollständig klären können.

5.5 Lokale Extrema

Definition 5.42. Seien (X, d) ein metrischer Raum und $f: X \rightarrow \mathbb{R}$, $a \in X$.

- (a) a heißt *lokales Minimum* von f , wenn $r > 0$ existiert mit $f(x) \geq f(a)$ für alle $x \in B_r(a)$.

- (b) a heißt *striktes lokales Minimum* von f , wenn $r > 0$ existiert mit $f(x) > f(a)$ für alle $x \in B_r(a) \setminus \{a\}$.
- (c) Gilt $f(x) \geq f(a)$ für alle $x \in X$, so heißt a *globales Minimum* von f .
- (d) Analog zu (a)–(c) definiert man: a ist (striktes) lokales bzw. globales *Maximum* von f (Ungleichungen umkehren!).
- (e) a heißt lokales (bzw. globales) *Extremum* von f , wenn a ein lokales (bzw. globales) Minimum oder Maximum von f ist.

Bemerkung 5.43. Häufig spricht man auch von einer *Minimalstelle* statt von einem Minimum oder nennt den Funktionswert an einer Minimalstelle ein Minimum. Analoges gilt für Maxima und für Extrema.

Beispiel 5.44. (a) Seien $X = \mathbb{R}$ und $f: X \rightarrow \mathbb{R}, f(x) = x^4 - 2x^2$. Dann ist 0 ein lokales Maximum von f , aber kein globales Maximum. 1 und -1 sind globale Minima von f .

(b) Sei $X = [0, 1]$ und $f: X \rightarrow \mathbb{R}, f(x) = x$. Dann ist 0 ein globales Minimum von f und 1 ein globales Maximum.

(c) Sei $X = (0, 1)$ und $f: X \rightarrow \mathbb{R}, f(x) = x$. Dann hat f keine lokalen Extrema.

(d) Sei $X = \mathbb{R}^2$ und $f: X \rightarrow \mathbb{R}, f(x) = x_1^2$. Dann ist $(0, t)$ für alle $t \in \mathbb{R}$ ein globales Minimum von f , aber *kein* striktes Minimum.

Satz 5.45. Seien $U \subseteq \mathbb{R}^N$ offen, $f \in C^1(U)$ und $a \in U$ ein lokales Extremum von f . Dann gilt $\nabla f(a) = 0$.

Beweis. Ohne Einschränkung sei a ein lokales Maximum von f und $r > 0$ derart, dass $f(x) \leq f(a)$ für alle $x \in B_r(a)$. Dann gilt für $i = 1, \dots, N$:

$$\lim_{t \rightarrow 0^-} \frac{f(a + te_i) - f(a)}{t} \geq 0 \quad \text{und} \quad \lim_{t \rightarrow 0^+} \frac{f(a + te_i) - f(a)}{t} \leq 0, \quad \text{also } \partial_i f(a) = 0.$$

Es folgt $\nabla f(a) = 0$. □

Definition und Bemerkung 5.46. Sei $U \subseteq \mathbb{R}^N$ offen und $f \in C^1(U)$. Ein Punkt $a \in U$ heißt *kritischer Punkt* von f , falls $\nabla f(a) = 0$, anderenfalls *regulärer Punkt*. Nach Satz 5.45 ist jedes lokale Extremum ein kritischer Punkt von f . Die Umkehrung gilt nicht: Die Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}, f(x) = x_1^2 - x_2^2$ hat in $(0, 0)$ einen kritischen Punkt, aber kein lokales Extremum, da

$$f(0, 0) = 0 \quad \text{und} \quad f(t, 0) > 0, f(0, t) < 0 \quad \text{für } t \in \mathbb{R} \setminus \{0\}.$$

Im Folgenden werden wir anhand der Hessematrix untersuchen, unter welchen Bedingungen kritische Punkte lokale Extrema sind.

Satz 5.47. Sei $U \subseteq \mathbb{R}^N$ offen, $f \in C^2(U)$, $a \in U$ ein kritischer Punkt von f . Dann gilt:

- (a) $H_f(a)$ positiv definit $\Rightarrow a$ striktes lokales Minimum
- (b) $H_f(a)$ negativ definit $\Rightarrow a$ striktes lokales Maximum
- (c) $H_f(a)$ indefinit $\Rightarrow a$ kein lokales Extremum.
- (d) a lokales Minimum $\Rightarrow H_f(a)$ positiv semidefinit
- (e) a lokales Maximum $\Rightarrow H_f(a)$ negativ semidefinit

Gilt (c), so nennt man a auch einen Sattelpunkt von f .

Beweis. Da a ein kritischer Punkt von f ist, gilt nach Bemerkung 5.33 und dem Satz von Taylor

$$(5.9) \quad f(a+v) = f(a) + \frac{1}{2} \langle H_f(a)v, v \rangle + R_2(v) \quad \text{für } v \in \mathbb{R}^N \text{ mit } a+tv \in U \text{ für } t \in [0, 1],$$

wobei $\lim_{v \rightarrow 0} \frac{R_2(v)}{|v|^2} = 0$ gilt. Insbesondere gilt für $v \in \mathbb{R}^N \setminus \{0\}$:

$$(5.10) \quad \begin{aligned} \lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t^2} &= \lim_{t \rightarrow 0} \frac{1}{t^2} \left[\frac{1}{2} \langle H_f(a)tv, tv \rangle + R_2(tv) \right] \\ &= \frac{1}{2} \langle H_f(a)v, v \rangle + |v|^2 \lim_{t \rightarrow 0} \frac{R_2(tv)}{|tv|^2} = \frac{1}{2} \langle H_f(a)v, v \rangle. \end{aligned}$$

Die Hesse-Matrix dominiert also das Wachstumsverhalten von f in der Nähe des kritischen Punktes a , und die Eigenschaften der Hesse-Matrix stehen im Zusammenhang damit, ob a ein Extremum ist und wenn ja, welchen Typs.

Nun zunächst zu (d): Da a lokales Minimum von f ist, gilt für $v \in \mathbb{R}^N \setminus \{0\}$

$$0 \leq \lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t^2} \stackrel{(5.10)}{=} \frac{1}{2} \langle H_f(a)v, v \rangle.$$

Also ist $H_f(a)$ positiv semidefinit.

(e) analog.

Zu (c): Da $H_f(a)$ indefinit ist, existieren $v, w \in \mathbb{R}^N \setminus \{0\}$ mit $\frac{1}{2} \langle H_f(a)v, v \rangle < 0$ und $\frac{1}{2} \langle H_f(a)w, w \rangle > 0$. Aus (5.10) folgt dann

$$\lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t^2} < 0 \quad \text{und} \quad \lim_{t \rightarrow 0} \frac{f(a+tw) - f(a)}{t^2} > 0.$$

Somit existiert $\varepsilon > 0$ mit

$$f(a+tv) < f(a) < f(a+tw) \quad \text{für } 0 < |t| < \varepsilon.$$

Also ist a kein lokales Extremum von f .

Zu (a): Wir setzen $m := \min_{|x|=1} \langle H_f(a)x, x \rangle$. Dann gilt für alle $x \in \mathbb{R}^N$: $m|x|^2 \leq \langle H_f(a)x, x \rangle$. Da $m > 0$ vorausgesetzt ist, existiert $\varepsilon > 0$ mit

$$B_\varepsilon(a) \subseteq U \quad \text{und} \quad |R_2(v)| < \frac{m}{4}|v|^2 \quad \text{für } v \in B_\varepsilon(0) \setminus \{0\},$$

also

$$f(a+v) - f(a) \stackrel{(5.9)}{=} \frac{1}{2} \langle H_f(a)v, v \rangle + R_2(v) > \frac{m}{2}|v|^2 - \frac{m}{4}|v|^2 \geq \frac{m}{4}|v|^2 > 0$$

für $v \in B_\varepsilon(0) \setminus \{0\}$. Es folgt $f(x) > f(a)$ für $x \in B_\varepsilon(a) \setminus \{a\}$, und somit ist a ein striktes lokales Minimum von f .

(b) analog. □

Beispiel 5.48. Sei $f \in C^2(\mathbb{R}^2)$ definiert durch $f(x, y) = x^3 + y^3 - 3xy$. Dann gilt:

$$\nabla f(x, y) = \begin{pmatrix} 3x^2 - 3y \\ 3y^2 - 3x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow x^2 = y \text{ und } y^2 = x \Leftrightarrow \begin{pmatrix} x \\ y \end{pmatrix} \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}.$$

Also sind $p = (0, 0)$ und $q = (1, 1)$ die kritischen Punkte von f . Für $(x, y) \in \mathbb{R}^2$ ist zudem

$$H_f(x, y) = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}, \text{ also } H_f(p) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix} \quad \text{und} \quad H_f(q) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$$

Da $\det H_f(p) = -9$, ist $H_f(p)$ indefinit und somit p kein lokales Extremum von f (Sattelpunkt). Da ferner $\det H_f(q) = 27 > 0$ und $\text{Spur } H_f(q) = 12 > 0$, ist $H_f(q)$ positiv definit und somit q ein striktes lokales Minimum. f hat keine globalen Extrema, da

$$f(x, x) = 2x^3 - 3x^2 \rightarrow \begin{cases} \infty & \text{für } x \rightarrow \infty \\ -\infty & \text{für } x \rightarrow -\infty \end{cases}$$

Ende der Sitzung vom 18.6.

6 Homogene Lineare Gewöhnliche Differentialgleichungen mit konstanten Koeffizienten

Wir betrachten *gewöhnliche Differentialgleichungen*, also Gleichungen für Funktionen $u: I \rightarrow \mathbb{K}^N$, I ein Intervall, die auch Ableitungen von u enthalten. In diesem Kapitel schreiben wir die Ableitungen von u bis zur dritten Ordnung mit Punkt, also \dot{u} , \ddot{u} und \dddot{u} . Da wir $\mathbb{K} = \mathbb{C}$ zulassen, müssen wir noch die Ableitung komplex vektorwertiger Funktionen definieren. Im Folgenden bilden wir Real- und Imaginärteil eines Vektors in \mathbb{C}^N Koordinatenweise:

Definition und Bemerkung 6.1. Sei I ein Intervall. Eine Funktion $u: I \rightarrow \mathbb{C}^N$, $u = v + iw$ mit Real- und Imaginärteilen $v, w: I \rightarrow \mathbb{R}^N$, heißt *differenzierbar in $t_0 \in I$* , falls v und w in t_0 differenzierbar sind. In diesem Fall ist

$$\dot{u}(t_0) := \dot{v}(t_0) + i\dot{w}(t_0)$$

die *Ableitung von u in t_0* . Ist u in jedem Punkt von I differenzierbar, dann heißt u *differenzierbar*. Es gilt

$$\dot{u}(t_0) = \lim_{t \rightarrow t_0} \frac{u(t) - u(t_0)}{t - t_0} \quad \text{für alle } t_0 \in I.$$

Die gewohnten Rechenregeln für Ableitungen gelten auch für diese komplex vektorwertigen Funktionen.

Speziell konzentrieren wir uns auf *homogene lineare Gleichungen*; das sind solche, für die jede Linearkombination von Lösungen wieder eine Lösung ist. Ferner hängen hier die in der Gleichung auftretenden Koeffizienten nicht von $t \in I$ ab, sind also konstant. Wie wir später sehen werden, können wir dann immer $I = (-\infty, \infty)$ annehmen.

Wir behandeln hier zwei Gestalten dieser Gleichungen: *Systeme erster Ordnung* haben die Form

$$(6.1) \quad \dot{u} = Au,$$

mit der *Koeffizientenmatrix* $A \in \mathbb{K}^{N \times N}$. Erster Ordnung bedeutet hier, dass außer u nur die erste Ableitung von u auftaucht. Die zweite Form sind *skalare Gleichungen n -ter*

Ordnung für eine Funktion $u: I \rightarrow \mathbb{K}$:

$$(6.2) \quad u^{(n)} + a_{n-1}u^{(n-1)} + \dots + a_1\dot{u} + a_0u = 0,$$

mit Koeffizienten $a_0, a_1, \dots, a_{n-1} \in \mathbb{K}$. Man vergewissert sich leicht, dass diese Gleichungen linear sind. Wir beginnen mit Systemen erster Ordnung und führen die skalaren Gleichungen höherer Ordnung dann auf erstere zurück.

Die Anwendung in den Wissenschaften hat meist die folgende Gestalt: Ein Modell für das betrachtete Phänomen liefert eine Differentialgleichung (DGL), deren Lösungen das beobachtete Verhalten wiedergeben sollen. Zu einer gegebenen DGL und einem Anfangszustand oder *Anfangswert* möchte man die Lösungen u bestimmen, welche den Anfangswert annehmen. Im Falle von (6.1) haben Anfangswerte die Form $u(0) = u_0$, d.h. der Wert $u_0 \in \mathbb{K}^N$ soll von der Lösung u zum Zeitpunkt 0 angenommen werden. Im Falle von (6.2) hat der Anfangswert die Form

$$(6.3) \quad \begin{pmatrix} u(0) \\ \dot{u}(0) \\ \ddot{u}(0) \\ \vdots \\ u^{(n-1)}(0) \end{pmatrix} = \begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_{n-1} \end{pmatrix} \in \mathbb{K}^N.$$

Die Kombination einer gewöhnlichen DGL mit einem Anfangswert nennt man auch *Anfangswertproblem* (AWP).

6.1 Die Exponentialfunktion einer Matrix

Die DGL $\dot{u} = au$ für Funktionen $u: \mathbb{R} \rightarrow \mathbb{K}$, mit $a \in \mathbb{K}$ und Anfangswert $u(0) = u_0$, hat die Lösung $u(t) = u_0 e^{at}$. Dies liefert die Motivation, Systeme erster Ordnung mit Hilfe der Exponentialfunktion für Matrizen zu lösen. Wir betrachten diese zunächst für komplexe Matrizen.

Wir bezeichnen hier die Operatornorm (bezüglich $|\cdot|$ in \mathbb{C}^N , analog zur reellen Operatornorm) auf $\mathbb{C}^{N \times N} = \mathcal{L}(\mathbb{C}^N)$ durch $\|\cdot\|$. Für $A \in \mathbb{C}^{N \times N}$ definieren wir induktiv die Potenzen $A^k := A \cdot A^{k-1}$ für $k \in \mathbb{N}$, wobei $A^0 := E_N$ sei. Aus Satz 4.41 folgt induktiv $\|A^k\| \leq \|A\|^k$ für alle $k \in \mathbb{N}_0$.

Satz und Definition 6.2. Sei $A \in \mathbb{C}^{N \times N}$. Dann konvergiert die Reihe der Exponentialfunktion von A ,

$$(6.4) \quad e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k,$$

in $\mathbb{C}^{N \times N}$ bezüglich der Norm $\|\cdot\|$ (also bezüglich jeder Norm, und Koordinatenweise).

Beweisskizze. Sei $(s_n) \subseteq \mathbb{C}^{N \times N}$ die Teilsummenfolge, mit $s_n := \sum_{k=0}^n \frac{1}{k!} A^k$. Für $m, n \in \mathbb{N}$,

$m < n$, gilt

$$\|s_n - s_m\| = \left\| \sum_{k=m+1}^n \frac{1}{k!} A^k \right\| \leq \sum_{k=m+1}^n \frac{1}{k!} \|A^k\| \leq \sum_{k=m+1}^n \frac{1}{k!} \|A\|^k.$$

Da $e^{\|A\|} = \sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k$ konvergiert, ist (s_n) also eine Cauchyfolge. Insbesondere ist jede Koordinatenfolge von (s_n) eine Cauchyfolge in \mathbb{C} , also konvergent. Damit konvergiert (s_n) in $\mathbb{C}^{N \times N}$ Koordinatenweise, also auch bezüglich jeder Norm. \square

Satz 6.3. Für $A \in \mathbb{C}^{N \times N}$ ist die Funktion $\mathbb{R} \rightarrow \mathbb{C}^{N \times N}$, $t \mapsto e^{tA}$, stetig differenzierbar und es gilt

$$(6.5) \quad \frac{d}{dt} e^{tA} = A e^{tA}.$$

Beweisskizze. Für jede Konstante $C > 0$ und $t \in [-C, C]$ gilt

$$\sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k |t|^k \leq \sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k C^k = e^{\|A\|C} < \infty.$$

Dies ist auf $[-C, C]$ also eine konvergente Majorante für die matrixwertige Funktionenreihe $\sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k$. Mit einer leichten Verallgemeinerung des Konvergenzkriteriums von Weierstraß, Satz 2.6, auf matrixwertige Funktionen folgt, dass e^{tA} und somit $A e^{tA}$ stetig von t abhängt. Genauso wie bei Potenzreihen mit reellen Koeffizienten zeigt man, dass gliedweise differenziert werden darf. Es folgt

$$\frac{d}{dt} e^{tA} = \frac{d}{dt} \sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} A^k t^{k-1} = A \sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k = A e^{tA}. \quad \square$$

Korollar 6.4. Eine Lösung des AWP

$$(6.6) \quad \dot{u} = Au, \quad u(0) = u_0$$

ist $u(t) = e^{tA} u_0$.

Im nächsten Abschnitt benötigen wir noch einige Eigenschaften der Matrix-Exponentialfunktion:

Proposition 6.5. Seien $A, B \in \mathbb{C}^{N \times N}$.

- (a) Falls $AB = BA$ gilt (A und B kommutieren), dann folgt $e^{A+B} = e^A e^B$.
- (b) Ist B invertierbar, dann folgt $e^{B^{-1}AB} = B^{-1} e^A B$.
- (c) Für eine Diagonalmatrix gilt $e^{\text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_N)} = \text{Diag}(e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_N})$.

Beweis. **(a):** Genauso wie für die skalare Exponentialfunktion. Dass A und B kommutieren ist dabei entscheidend.

(b):

Wegen $(B^{-1}AB)^2 = B^{-1}ABB^{-1}AB = B^{-1}A^2B$ folgt induktiv $(B^{-1}AB)^k = B^{-1}A^k B$ und somit

$$e^{B^{-1}AB} = \sum_{k=0}^{\infty} \frac{1}{k!} (B^{-1}AB)^k = \sum_{k=0}^{\infty} \frac{1}{k!} B^{-1}A^k B = B^{-1} \left(\sum_{k=0}^{\infty} \frac{1}{k!} A^k \right) B = B^{-1} e^A B.$$

(c): Übung. □

Bemerkung 6.6. Man überlegt sich leicht, dass für eine Blockmatrix A , mit quadratischen Matrizen A_1, A_2, \dots, A_k auf der Diagonalen, gilt:

$$(6.7) \quad A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_k \end{pmatrix} \Rightarrow e^A = \begin{pmatrix} e^{A_1} & 0 & \cdots & 0 \\ 0 & e^{A_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{A_k} \end{pmatrix}.$$

6.2 Fundamentalsysteme

Es fehlt noch, die zeitabhängigen Koordinatenfunktionen der Matrix e^{tA} zu berechnen. Damit werden wir Lösungen von (6.6) explizit in elementaren Funktionen angeben können. Dies ist eine Besonderheit der linearen DGL mit konstanten Koeffizienten. Für allgemeinere DGL kann man die Lösungen nicht immer durch Elementarfunktionen ausdrücken.

Zunächst habe $A \in \mathbb{C}^{k \times k}$ nur einen Eigenwert $\lambda \in \mathbb{C}$ und nur einen zugehörigen Eigenvektor v . Dann formen wir die Basismatrix $B := (v \mid v^{(2)} \mid \cdots \mid v^{(k)})$ mit den zugehörigen Nebeneigenvektoren. Damit gilt $A = B J_k(\lambda) B^{-1}$, d.h. der Jordanblock $J_k(\lambda) = \lambda E_k + C_k$ der Länge k ist die Jordansche Normalform von A . Hier setzen wir

$$C_k := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} \in \mathbb{C}^{k \times k}.$$

Es gilt $C_k e_1 = 0$ und $C_k e_j = e_{j-1}$ für $j = 2, 3, \dots, k$. Induktiv erhält man für $j = 1, 2, \dots, k$:

$$(6.8) \quad C_k^\ell e_j = \begin{cases} 0, & j = 1, 2, \dots, \ell, \\ e_{j-\ell}, & j = \ell + 1, \dots, k, \end{cases}$$

also insbesondere $C_k^k = 0$ (man nennt ganz allgemein eine Matrix, für die eine Potenz verschwindet, *nilpotent*). Es ergibt sich

$$(6.9) \quad e^{tC_k} = \sum_{j=0}^{k-1} \frac{1}{j!} C_k^j t^j = \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{k-1}}{(k-1)!} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{t^2}{2} \\ \vdots & & \ddots & \ddots & t \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Aus Proposition 6.5(a) folgt außerdem

$$(6.10) \quad e^{tJ_k(\lambda)} = e^{\lambda t E_k} e^{tC_k} = e^{\lambda t} e^{tC_k},$$

da die Matrizen E_k und C_k kommutieren. Zusammenfassend erhalten wir aus Proposition 6.5(b)

$$(6.11) \quad e^{tA} = e^{\lambda t} B e^{tC_k} B^{-1},$$

wo e^{tC_k} durch (6.9) gegeben ist.

Möchte man das AWP (6.6) lösen, dann braucht man B^{-1} nicht zu berechnen. Man löst nur das lineare Gleichungssystem $Bx = u_0$ nach x auf und erhält die Lösung $u(t) = e^{\lambda t} B e^{tC_k} x$, eine Linearkombination der Spalten von $e^{\lambda t} B e^{tC_k}$, mit den Koordinaten von x als Koeffizienten. Aus (6.9) lesen wir direkt ab:

$$(6.12) \quad e^{\lambda t} B e^{tC_k} = e^{\lambda t} \left(v \left| v^{(2)} + tv \right| v^{(3)} + tv^{(2)} + \frac{t^2}{2}v \left| \cdots \right| v^{(k)} + tv^{(k-1)} + \cdots + \frac{t^{k-1}}{(k-1)!}v \right)$$

Jede Spalte von $e^{\lambda t} B e^{tC_k}$ ist eine Lösung von $\dot{u} = Au$, denn es gilt:

$$\frac{d}{dt} e^{\lambda t} B e^{tC_k} = \frac{d}{dt} e^{tA} B = A e^{tA} B = A e^{\lambda t} B e^{tC_k}.$$

Im allgemein Fall einer Matrix $A \in \mathbb{C}^{N \times N}$ ist die Jordansche Normalform J von A eine diagonale Blockmatrix aus Jordanblöcken:

$$B^{-1}AB = J = \begin{pmatrix} J_{s_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{s_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{s_k}(\lambda_k) \end{pmatrix}.$$

Wir haben hier eine andere Darstellung gewählt als in Satz 3.46. Dabei ist k die Anzahl linear unabhängiger Eigenvektoren von A . In der Reihe der Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_k$ können Wiederholungen auftreten. Wegen Bemerkung 6.6 kann man die Exponentialfunktion

e^{tJ} blockweise berechnen:

$$e^{tJ} = \begin{pmatrix} e^{tJ_{s_1}(\lambda_1)} & 0 & \cdots & 0 \\ 0 & e^{tJ_{s_2}(\lambda_2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{tJ_{s_k}(\lambda_k)} \end{pmatrix}.$$

Zusammen mit den obigen Ergebnissen für einen Jordanblock erhalten wir:

Bemerkung 6.7. Rezept, um explizit eine (komplexe) Lösung des AWP (6.6) angeben zu können:

- konstruiere die geordnete Basismatrix B aus Eigenvektoren und Nebeneigenvektoren, wie in Abschnitt 3.2.2;
- für jeden Jordanblock $J_k(\lambda)$ und die zugehörige Gruppe aus einem Eigenvektor und Nebeneigenvektoren bilde die k Lösungen (Spalten) von $e^{\lambda t} B e^{t\tilde{C}_k}$ wie in (6.12); hier ergänzen wir Nullen in geeigneter Position, um aus C_k eine $\mathbb{C}^{N \times N}$ -Matrix \tilde{C}_k zu machen;
- die Matrixfunktion $F: \mathbb{R} \rightarrow \mathbb{C}^{N \times N}$, $F(t) := B e^{tJ}$ aller dieser Lösungen, als Spalten aufgefasst, ist ein (*komplexes*) *Fundamentalsystem*, d.h. eine zeitabhängige Matrix aus linear unabhängigen Spalten, die Lösungen von $\dot{u} = Au$ sind. Eine Lösung des AWP (6.6) erhält man als Linearkombination $F(t)x$, wobei $x \in \mathbb{C}^N$ das lineare Gleichungssystem $Bx = u_0$ löst. Auch $t \mapsto e^{tA}$ ist ein Fundamentalsystem, welches die Zusatzeigenschaft hat, dass es in $t = 0$ den Wert E_N annimmt. Für ein beliebiges Fundamentalsystem $G(t)$ gilt $e^{tA} = G(t)G(0)^{-1}$.

Ende der Sitzung vom 22.6.

Beispiel 6.8. Wir berechnen für

$$A := \begin{pmatrix} 2 & 1 & 0 & -2 \\ -\frac{1}{2} & 1 & \frac{1}{2} & 2 \\ -1 & -1 & 3 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

aus Beispiel 3.47 die Lösung des AWP

$$(6.13) \quad \dot{u} = Au, \quad u(0) = u_0 := \begin{pmatrix} -4 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

sowie e^{tA} . Wie in Gleichung (6.12) erhalten wir unter Berücksichtigung der zwei in Beispiel 3.47 erhaltenen Jordanblöcke

$$\begin{aligned} Be^{tJ} &= e^{2t} \left(v_1^{(1)} \mid v_1^{(2)} + tv_1^{(1)} \mid v_1^{(3)} + tv_1^{(2)} + \frac{t^2}{2}v_1^{(1)} \mid v_2^{(1)} \right) \\ &= e^{2t} \begin{pmatrix} 2 & -2 + 2t & -2t + t^2 & 0 \\ 0 & 2 & 2t & 2 \\ 2 & 2 + 2t & 2t + t^2 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}. \end{aligned}$$

Auflösung des LGS $Bx = u_0$ liefert $x = (-1, 1, 1, 0)^t$. Damit ist die Lösung von (6.13)

$$Be^{tJ}x = e^{2t} \begin{pmatrix} t^2 - 4 \\ 2t + 2 \\ t^2 + 4t \\ 1 \end{pmatrix}.$$

Nach Berechnung der Inversen von B erhalten wir

$$\begin{aligned} e^{tA} &= Be^{tJ}B^{-1} = \frac{1}{4}e^{2t} \begin{pmatrix} 2 & -2 + 2t & -2t + t^2 & 0 \\ 0 & 2 & 2t & 2 \\ 2 & 2 + 2t & 2t + t^2 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & -2 & 1 & 4 \\ 1 & 2 & -1 & 0 \end{pmatrix} \\ &= e^{2t} \begin{pmatrix} 1 - \frac{t^2}{4} & t - \frac{t^2}{2} & \frac{t^2}{4} & -2t + t^2 \\ -\frac{t}{2} & 1 - t & \frac{t}{2} & 2t \\ -t - \frac{t^2}{4} & -t - \frac{t^2}{2} & 1 + t + \frac{t^2}{4} & 2t + t^2 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Bemerkung 6.9. Falls $A \in \mathbb{R}^{N \times N}$ komplexe Eigenwerte hat, dann fasst man A als Element von $\mathbb{C}^{N \times N}$ auf und löst das AWP (6.6) wie in Bemerkung 6.7, unter Verwendung der Basismatrix B und der Jordanschen Normalform J . Ist nämlich der Anfangswert u_0 reell, dann erhält man auf diese Weise automatisch eine reelle Lösung. Das liegt an der Eindeutigkeit, welche wir aber erst in Mathe 3 beweisen werden. Ebenso berechnet man e^{tA} und erhält automatisch ein reelles Fundamentalsystem. Möchte man ein reelles Fundamentalsystem (FS) bestimmen, ohne B^{-1} zu berechnen, so geht man ausgehend von einem komplexen FS (z.B. Be^{tJ}) wie folgt vor: Ist $u(t)$ eine Spalte, die zu einem reellen Eigenwert gehört, dann existiert $\tau \in \mathbb{C}$, so dass τu reell ist; man ersetzt u durch τu . Bei einem Paar $\lambda, \bar{\lambda}$ von Eigenwerten ersetzt man die zugehörigen Lösungen durch Real- und Imaginärteile aller Lösungen, die zu λ gehören.

Beispiel 6.10. Wir betrachten $A := \begin{pmatrix} 1 & -1 \\ 4 & 1 \end{pmatrix}$ und finden den Eigenwert $1 + 2i$ mit Eigenvektor $v := \begin{pmatrix} 1 \\ -2i \end{pmatrix}$ und den Eigenwert $1 - 2i$ mit Eigenvektor $\bar{v} = \begin{pmatrix} 1 \\ 2i \end{pmatrix}$. Mit $B := \begin{pmatrix} 1 & 1 \\ -2i & 2i \end{pmatrix}$ ist die komplexe Jordansche Normalform dann $J = \begin{pmatrix} 1+2i & 0 \\ 0 & 1-2i \end{pmatrix}$. Dies liefert wegen

$$e^{t(1\pm 2i)} = e^t e^{\pm 2it}$$

$$\begin{aligned} e^{tA} &= B e^{tJ} B^{-1} = \frac{1}{4} e^t \begin{pmatrix} 1 & 1 \\ -2i & 2i \end{pmatrix} \cdot \begin{pmatrix} e^{2it} & 0 \\ 0 & e^{-2it} \end{pmatrix} \cdot \begin{pmatrix} 2 & i \\ 2 & -i \end{pmatrix} \\ &= \frac{1}{4} e^t \begin{pmatrix} 1 & 1 \\ -2i & 2i \end{pmatrix} \cdot \begin{pmatrix} \cos(2t) + i \sin(2t) & 0 \\ 0 & \cos(2t) - i \sin(2t) \end{pmatrix} \cdot \begin{pmatrix} 2 & i \\ 2 & -i \end{pmatrix} \\ &= e^t \begin{pmatrix} \cos(2t) & -\frac{1}{2} \sin(2t) \\ 2 \sin(2t) & \cos(2t) \end{pmatrix}. \end{aligned}$$

Ende der Sitzung vom 25.6.

6.3 Skalare Differentialgleichungen höherer Ordnung

6.3.1 Allgemeine Form

Wir betrachten für $n \in \mathbb{N}$ und $a_0, a_1, \dots, a_{n-1} \in \mathbb{C}$ die homogene lineare (*skalare*) *Differentialgleichung n-ter Ordnung*

$$(6.14) \quad Lu := u^{(n)} + a_{n-1}u^{(n-1)} + \dots + a_1\dot{u} + a_0u = 0$$

mit dem Anfangswert

$$(6.15) \quad (u, \dot{u}, \dots, u^{(n-1)})(0) = U_0 \in \mathbb{C}^n.$$

Hier steht L also für einen Differentialoperator. Ferner setzen wir

$$A := \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \\ -a_0 & -a_1 & \dots & \dots & -a_{n-1} \end{pmatrix},$$

und betrachten das System erster Ordnung

$$(6.16) \quad \dot{v} = Av.$$

Satz 6.11 (Reduktionsprinzip). (a) *Ist u eine Lösung von (6.14), dann ist*

$$(6.17) \quad v := \begin{pmatrix} u \\ \dot{u} \\ \ddot{u} \\ \vdots \\ u^{(n-1)} \end{pmatrix}$$

eine Lösung von (6.16).

- (b) Ist $v = (v_1, v_2, \dots, v_n)^t$ eine Lösung von (6.16), dann ist $u := v_1$ eine Lösung von (6.14) und es gilt (6.17).

Beweis. Dies folgt durch Koordinatenweise Betrachtung von (6.16). □

Die Behandlung skalarer Gleichungen haben wir also auf spezielle Systeme erster Ordnung zurückgeführt. Es reicht, für Lösungen v von (6.16) nur die erste Koordinatenfunktion anzugeben, weil die restlichen Koordinaten, wie in (6.17) angegeben, dann die Ableitungen sind. Ein Fundamentalsystem von (6.14) hat also die Form

$$(6.18) \quad \begin{pmatrix} u_1 & u_2 & \cdots & u_n \\ \dot{u}_1 & \dot{u}_2 & \cdots & \dot{u}_n \\ \vdots & \vdots & & \vdots \\ u_1^{(n-1)} & u_2^{(n-1)} & \cdots & u_{n-1}^{(n-1)} \end{pmatrix}$$

mit Lösungen u_1, u_2, \dots, u_n von (6.14).

Wir definieren mit

$$P_L(\lambda) := \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

das *charakteristische Polynom von L*. Wenn man $\det(A - \lambda E_n)$ nach der letzten Zeile entwickelt, dann sieht man, dass $p_A = (-1)^n P_L$ gilt. Demnach stimmen die Eigenwerte von A und die Nullstellen von P_L samt ihrer algebraischen Vielfachheiten überein. Sei nun λ eine (komplexe) Nullstelle von P_L der algebraischen Vielfachheit α . Die Teilmatrix der Länge $n - 1$ in der rechten oberen Ecke von $A - \lambda E_n$ ist

$$\begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ -\lambda & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\lambda & 1 \end{pmatrix},$$

mit Determinante 1. Die oberen $n - 1$ Zeilen von $A - \lambda E_n$ sind also linear unabhängig, d.h. $A - \lambda E_n$ hat Rang $n - 1$. Daraus folgt $\dim \mathcal{E}_A(\lambda) = 1$. Man prüft sofort, dass $w := (1, \lambda, \lambda^2, \dots, \lambda^{n-1})^t$ ein Eigenvektor von A zum Eigenwert λ ist. Dies liefert $\mathcal{E}_A(\lambda) = \text{Spann}(w)$. Zu λ existiert demnach nur ein Jordanblock, welcher die Länge α haben muss. Aus den ersten Koordinaten der Lösungen aus (6.12) erhalten wir durch geschicktes lineares Kombinieren die linear unabhängigen Lösungen

$$\{e^{\lambda t}, te^{\lambda t}, \dots, t^{\alpha-1}e^{\lambda t}\}.$$

Wir fassen zusammen: Aus den m komplexen Nullstellen λ_k von P_L der Vielfachheit α_k

ergibt sich als Fundamentalsystem von (6.14):

$$(6.19) \quad \{t^j e^{\lambda_k t} \mid k = 1, 2, \dots, m, j = 0, 1, \dots, \alpha_k - 1\}.$$

Um im Falle $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$ ein reelles Fundamentalsystem zu erhalten, ersetzt man wieder die Lösungen, die zu einem Eigenwertpaar $\lambda, \bar{\lambda}$ gehören durch Real- und Imaginärteil der Lösungen, die zu λ gehören.

Ist das Fundamentalsystem $F: \mathbb{R} \rightarrow \mathbb{C}^{n \times n}$ durch (6.18) gegeben, so ist die Lösung des AWP $F(t)x$, wobei $x \in \mathbb{C}^n$ durch $F(0)x = U_0$ gegeben ist.

6.3.2 Gleichungen zweiter Ordnung

Wir behandeln noch den speziellen Fall der reellen linearen skalaren DGL zweiter Ordnung

$$(6.20) \quad \ddot{u} + b\dot{u} + c = 0$$

mit $b, c \in \mathbb{R}$ und dem charakteristischen Polynom $P(\lambda) = \lambda^2 + b\lambda + c$.

Fall $b^2/4 - c > 0$: Wir erhalten die unterschiedlichen reellen Eigenwerte $\lambda_{1,2} := -\frac{b}{2} \pm \sqrt{b^2/4 - c}$ und das FS $\{e^{\lambda_1 t}, e^{\lambda_2 t}\}$. Ist $b > 0$, so folgt $\lambda_2 < \lambda_1 < 0$, und die Lösungen fallen im Betrag exponentiell ab. Ist $b < 0$, so folgt $\lambda_1 > \lambda_2 > 0$, und die Lösungen wachsen im Betrag exponentiell an. Ist $b = 0$, dann folgt $\lambda_2 < 0 < \lambda_1$, es gibt betragsmäßig exponentiell wachsende und fallende Lösungen.

Fall $b^2/4 - c = 0$: Es gibt nur einen reellen Eigenwert $\lambda = -\frac{b}{2}$. Als FS erhalten wir $\{e^{\lambda t}, te^{\lambda t}\}$. Im Fall $b = 0$ gilt $\lambda = 0$ und ein FS ist $\{1, t\}$.

Fall $b^2/4 - c < 0$: Mit $\delta := -\frac{b}{2}$ und $\omega := \sqrt{c - b^2/4} > 0$ erhalten wir die zueinander komplex konjugierten Eigenwerte $\lambda := \delta + i\omega$ und $\bar{\lambda}$. Ein komplexes FS ist

$$\{e^{\lambda t}, e^{\bar{\lambda} t}\} = \{e^{\delta t} e^{i\omega t}, e^{\delta t} e^{-i\omega t}\},$$

und als reelles FS erhalten wir hieraus

$$\{e^{\delta t} \cos(\omega t), e^{\delta t} \sin(\omega t)\}.$$

Je nachdem, ob b positiv oder negativ ist, haben diese Lösungen beide exponentiell fallende oder wachsende Amplituden. Und für $b = 0$ sind alle Lösungen periodisch mit Periode $2\pi/\omega$. Im Unterfall $c > 0$ und $b = 0$ ($b > 0$) nennt man (6.20) den (*gedämpften*) *harmonischen Oszillator*.

7 Implizite Funktionen

7.1 Vollständigkeit, Fixpunkte und Nullstellen

Auch in diesem Kapitel haben alle Vektorräume einen der Skalarkörper \mathbb{R} oder \mathbb{C} .

Definition 7.1. Sei (X, d) ein metrischer Raum.

- (a) Eine Folge $(x_n)_n$ in X heißt *Cauchyfolge*, wenn für alle $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert mit $d(x_n, x_m) < \varepsilon$ für $m, n \geq N$.
- (b) (X, d) heißt *vollständig*, wenn jede Cauchyfolge in X konvergiert.
- (c) Ein vollständiger normierter Raum $(V, \|\cdot\|)$ heißt *Banachraum*.
- (d) Ein vollständiger Skalarproduktraum $(V, \langle \cdot, \cdot \rangle)$ heißt *Hilbertraum*.

Bemerkung 7.2. Ist (X, d) ein metrischer Raum und $(x_n)_n$ eine Folge in X , so gilt:

- (a) Ist $(x_n)_n$ eine Cauchyfolge, so ist $(x_n)_n$ beschränkt.
- (b) Ist $(x_n)_n$ konvergent, so ist $(x_n)_n$ eine Cauchyfolge.
- (c) Ist $(x_n)_n$ eine Cauchyfolge und besitzt $(x_n)_n$ eine konvergente Teilfolge in X , so ist $(x_n)_n$ selbst konvergent.

Beweis. Übung. □

Satz 7.3. Jeder endlich dimensionale normierte Raum $(V, \|\cdot\|)$ ist vollständig, d.h. ein Banachraum.

Beweis. Sei $(x_n)_n$ eine Cauchyfolge in V . Dann ist $(x_n)_n$ beschränkt. Also existiert $R > 0$ mit $x_n \in \overline{B}_R(0)$ für alle $n \in \mathbb{N}$. Als beschränkte und abgeschlossene Teilmenge von V ist $\overline{B}_R(0)$ nach Satz 4.47 kompakt. Daher besitzt $(x_n)_n$ eine konvergente Teilfolge. Nach Bemerkung 7.2(c) ist $(x_n)_n$ dann selbst auch konvergent. □

Satz 7.4. Sei $I := [a, b] \subseteq \mathbb{R}$ ein kompaktes Intervall mit $a < b$. Dann ist $(C(I), \|\cdot\|_\infty)$ ein Banachraum.

Beweis. Sei $(f_n)_n$ eine Cauchyfolge in $(C(I), \|\cdot\|_\infty)$, und sei $c_n := \sup_{\substack{m \in \mathbb{N} \\ m > n}} \|f_m - f_n\|_\infty$ für $n \in \mathbb{N}$. Die Cauchybedingung liefert $\lim_{n \rightarrow \infty} c_n = 0$. Ist $t \in I$, so gilt zudem für $m, n \in \mathbb{N}$

$$|f_m(t) - f_n(t)| \leq \|f_m - f_n\|_\infty.$$

Also ist $(f_n(t))_n$ eine Cauchyfolge in \mathbb{R} und somit konvergent. Definiere nun $f: I \rightarrow \mathbb{R}$ durch

$$f(t) := \lim_{n \rightarrow \infty} f_n(t) \quad \text{für } t \in I.$$

Dann gilt

$$|f(t) - f_n(t)| = \lim_{m \rightarrow \infty} |f_m(t) - f_n(t)| \leq \sup_{\substack{m \in \mathbb{N} \\ m \geq n}} |f_m(t) - f_n(t)| \leq c_n \quad \text{für } t \in I \text{ und } n \in \mathbb{N}.$$

Also folgt

$$0 \leq \|f - f_n\|_\infty \leq c_n \rightarrow 0 \quad \text{für } n \rightarrow \infty,$$

d.h. die Folge $(f_n)_n$ konvergiert gleichmäßig gegen f . Mit Satz 2.3 folgt die Stetigkeit von f . Die Folge (f_n) hat im normierten Raum $(C(I), \|\cdot\|_\infty)$ also den Grenzwert f . \square

Bemerkung 7.5. (a) Ist (K, d) ein *kompakter* metrischer Raum und $(V, \|\cdot\|_V)$ ein normierter Raum, so setzt man

$$C(K, V) := \{f: K \rightarrow V \mid f \text{ stetig}\}.$$

Ähnlich wie für $C(I)$ sieht man Folgendes:

- $C(K, V)$ ist ein normierter \mathbb{R} -Vektorraum mit Norm $\|\cdot\|_\infty$ definiert durch

$$\|f\|_\infty := \sup_{x \in K} \|f(x)\|_V \stackrel{\text{Satz 4.48(c)}}{=} \max_{x \in K} \|f(x)\|.$$

- Ist $(V, \|\cdot\|_V)$ ein Banachraum, so auch $(C(K, V), \|\cdot\|_\infty)$.

- (b) Seien V ein normierter Raum und W ein Banachraum. Dann ist $\mathcal{L}(V, W)$, versehen mit der Operatornorm, ebenfalls ein Banachraum.

Beweis. Sei $(T_n)_n$ eine Cauchyfolge in $\mathcal{L}(V, W)$, und sei $c_n := \sup_{m \geq n} \|T_n - T_m\|$ für $n \in \mathbb{N}$. Die Cauchybedingung liefert dann $\lim_{n \rightarrow \infty} c_n = 0$. Für alle $x \in V$ gilt

$$\sup_{m \geq n} \|T_m x - T_n x\|_W \leq \sup_{m \geq n} \|T_m - T_n\| \|x\|_V = c_n \|x\|_V \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Somit ist $(T_n x)_n$ eine Cauchyfolge in W , und nach Voraussetzung existiert

$$Tx := \lim_{n \rightarrow \infty} T_n x \in W \quad \text{für alle } x \in V.$$

Aus der Linearität der Abbildungen T_n folgt direkt, dass auch die Zuordnung $x \mapsto Tx$ linear ist. Ferner ist

$$\|Tx\|_W = \lim_{n \rightarrow \infty} \|T_n x\|_W \leq \left(\sup_{n \in \mathbb{N}} \|T_n\| \right) \|x\|_V \quad \text{für alle } x \in V,$$

d.h. T ist stetig. Schließlich ist

$$\|(T - T_n)x\|_W = \lim_{m \rightarrow \infty} \|(T_m - T_n)x\|_W \leq \left(\sup_{m \geq n} \|T_m - T_n\| \right) \|x\|_V = c_n \|x\|_V$$

für alle $x \in V$, d.h. $\|T - T_n\| \leq c_n \rightarrow 0$ für $n \rightarrow \infty$. Dies zeigt die Vollständigkeit von $\mathcal{L}(V, W)$. \square

Satz 7.6. *Seien (X, d) ein vollständiger metrischer Raum und $A \subseteq X$. Der metrische Raum (A, d) ist genau dann vollständig, wenn A als Teilmenge von X abgeschlossen ist. Insbesondere ist jede abgeschlossene Teilmenge eines Banachraums vollständig.*

Beweis. „ \Rightarrow “: Sei $x \in \overline{A}$. Dann gilt $x = \lim_{n \rightarrow \infty} x_n$ für eine Folge $(x_n)_n$ in A . Nach Bemerkung 7.2(a) ist $(x_n)_n$ eine Cauchyfolge in X und somit auch in A , und damit konvergiert $(x_n)_n$ nach Voraussetzung in A . Es folgt $x \in A$. Insgesamt folgt $A = \overline{A}$.

„ \Leftarrow “: Sei $(x_n)_n$ eine Cauchyfolge in A . Da X vollständig ist, existiert $x = \lim_{n \rightarrow \infty} x_n \in X$, wobei $x \in \overline{A} = A$ nach Voraussetzung. Also konvergiert $(x_n)_n$ in A . \square

Ende der Sitzung vom 29.6.

Definition 7.7. Seien (X, d) ein metrischer Raum und $f: X \rightarrow X$ eine Abbildung. Existiert ein $q \in (0, 1)$ mit

$$d(f(x), f(y)) \leq qd(x, y) \quad \text{für alle } x, y \in X,$$

so nennt man f eine *Kontraktion* (oder genauer: *q-Kontraktion*). Mit anderen Worten: $f: X \rightarrow X$ ist Lipschitzstetig und besitzt eine Lipschitzkonstante $q < 1$.

Satz 7.8 (Banachscher Fixpunktsatz (BFPS)). *Seien (X, d) ein vollständiger metrischer Raum mit $X \neq \emptyset$ und $\varphi: X \rightarrow X$ eine Kontraktion. Dann hat φ genau einen Fixpunkt, d.h. es gibt genau ein $x \in X$ mit $\varphi(x) = x$. Ferner gilt*

$$\lim_{n \rightarrow \infty} \varphi^n(y) = x \quad \text{für alle } y \in X \quad (\text{wobei } \varphi^n := \underbrace{\varphi \circ \dots \circ \varphi}_{n\text{-mal}} \text{ sei.})$$

Beweis. Nach Voraussetzung existiert $q \in (0, 1)$ mit

$$(7.1) \quad d(\varphi(y), \varphi(z)) \leq qd(y, z) \quad \text{für alle } y, z \in X.$$

Um die Eindeutigkeit zu zeigen, seien $x_1, x_2 \in X$ Fixpunkte von φ . Mit (7.1) folgt dann

$$d(x_1, x_2) = d(\varphi(x_1), \varphi(x_2)) \leq qd(x_1, x_2), \quad \text{also } d(x_1, x_2) = 0 \text{ und damit } x_1 = x_2.$$

Um die Existenz und Konvergenz zu zeigen, seien $y \in X$ beliebig und $y_n := \varphi^n(y)$ für $n \in \mathbb{N}_0$ (also $y_0 = y$). Wir zeigen per Induktion:

$$(7.2) \quad d(y_{n+1}, y_n) \leq q^n d(y_1, y) \quad \text{für alle } n \in \mathbb{N}_0.$$

„ $n = 0$ “: klar.

„ $n \Rightarrow n + 1$ “:

$$d(y_{n+2}, y_{n+1}) = d(\varphi(y_{n+1}), \varphi(y_n)) \stackrel{(7.1)}{\leq} qd(y_{n+1}, y_n) \stackrel{\text{IA}}{\leq} qq^n d(y_1, y) = q^{n+1} d(y_1, y).$$

Für $m, n \in \mathbb{N}$, $m > n$, folgt nun mit der Dreiecksungleichung

$$d(y_m, y_n) \leq \sum_{j=n}^{m-1} d(y_{j+1}, y_j) \stackrel{(7.2)}{\leq} d(y_1, y) \sum_{j=n}^{m-1} q^j \leq d(y_1, y) q^n \sum_{j=0}^{\infty} q^j = \frac{q^n}{1-q} d(y_1, y) =: c_n.$$

Da $c_n \rightarrow 0$ für $n \rightarrow \infty$ gilt, ist $(y_n)_n$ eine Cauchyfolge. Da X vollständig ist, existiert $x := \lim_{n \rightarrow \infty} y_n$. Da ferner φ stetig ist, folgt

$$\varphi(x) = \varphi\left(\lim_{n \rightarrow \infty} y_n\right) = \lim_{n \rightarrow \infty} \varphi(y_n) = \lim_{n \rightarrow \infty} y_{n+1} = x.$$

Also ist x der eindeutige Fixpunkt von φ , und es folgt $\lim_{n \rightarrow \infty} \varphi^n(y) = x$ für alle $y \in X$, wie behauptet. \square

Korollar 7.9 (Parameterabhängiger BFPS). *Seien X, Y metrische Räume, Y vollständig und $\varphi: X \times Y \rightarrow Y$ eine Abbildung. Ferner existiere $q \in (0, 1)$, so dass $\varphi(x, \cdot)$ für alle $x \in X$ eine q -Kontraktion ist. Sei $g: X \rightarrow Y$ die eindeutige Abbildung mit $\varphi(x, g(x)) = g(x)$ für alle $x \in X$, welche nach dem BFPS existiert.*

- (a) *Ist $\varphi(\cdot, y)$ für alle $y \in Y$ stetig, dann ist auch g stetig.*
- (b) *Existiert $L \geq 0$, so dass $\varphi(\cdot, y)$ für alle $y \in Y$ Lipschitzstetig mit Lipschitzkonstante L ist, dann ist auch g Lipschitzstetig.*

Beweis. Für $x_1, x_2 \in X$ setzen wir $y_i := g(x_i)$. Dann gilt:

$$\begin{aligned} d(y_1, y_2) &= d(\varphi(x_1, y_1), \varphi(x_2, y_2)) \leq d(\varphi(x_1, y_1), \varphi(x_1, y_2)) + d(\varphi(x_1, y_2), \varphi(x_2, y_2)) \\ &\leq qd(y_1, y_2) + d(\varphi(x_1, y_2), \varphi(x_2, y_2)), \end{aligned}$$

also

$$d(y_1, y_2) \leq \frac{1}{1-q} d(\varphi(x_1, y_2), \varphi(x_2, y_2)).$$

Daraus folgen beide Behauptungen. \square

7.2 Auflösung von Gleichungen durch implizite Funktionen

Bemerkung und Beispiel 7.10. (a) Viele Probleme und Gleichungen der Analysis hängen von Parametern ab, und man möchte die Abhängigkeit der Lösungen von diesen Parametern untersuchen. Zum Beispiel seien $\lambda > 0$ und $P_\lambda: \mathbb{R} \rightarrow \mathbb{R}$ durch

$P_\lambda(x) := \lambda x^3 + 2x - 3$ gegeben. Dann ist $P'_\lambda(x) = 3\lambda x^2 + 2 > 0$ für alle $x \in \mathbb{R}$, also P_λ streng monoton wachsend in x . Da ferner gilt

$$P_\lambda(x) \rightarrow \begin{cases} \infty & \text{für } x \rightarrow \infty, \\ -\infty & \text{für } x \rightarrow -\infty, \end{cases}$$

folgt, dass P_λ für jedes $\lambda > 0$ genau eine Nullstelle $g(\lambda)$ hat. Frage: Ist die Abbildung $g: (0, \infty) \rightarrow \mathbb{R}$, $\lambda \mapsto g(\lambda)$, differenzierbar? Falls ja, so kann man ihre Ableitung g' dort *implizit* berechnen, wo g bekannt ist. Seien dazu $U := (0, \infty) \times \mathbb{R}$ und $F \in C^1(U)$ definiert durch $F(\lambda, x) = P_\lambda(x)$. Dann ist

$$\partial_1 F(\lambda, x) = x^3 \quad \text{und} \quad \partial_2 F(\lambda, x) = P'_\lambda(x) = 3\lambda x^2 + 2 \quad \text{für } (\lambda, x) \in U.$$

Da ferner $F(\lambda, g(\lambda)) = P_\lambda(g(\lambda)) = 0$ für alle $\lambda > 0$ gilt, folgt mit der Kettenregel:

$$0 = \frac{d}{d\lambda} F(\lambda, g(\lambda)) = J_F(\lambda, g(\lambda)) \begin{pmatrix} 1 \\ g'(\lambda) \end{pmatrix} = \partial_1 F(\lambda, g(\lambda)) + \partial_2 F(\lambda, g(\lambda))g'(\lambda),$$

also

$$g'(\lambda) = -\frac{\partial_1 F(\lambda, g(\lambda))}{\partial_2 F(\lambda, g(\lambda))} = -\frac{g(\lambda)^3}{3\lambda g(\lambda)^2 + 2} \quad \text{für } \lambda > 0.$$

Z.B. ist $g(1) = 1$ und damit $g'(1) = -\frac{1}{5}$.

Eine notwendige Bedingung, die Ableitung von g (deren Existenz wir noch nicht gezeigt haben!) zu berechnen, ist also $\partial_2 F(\lambda, g(\lambda)) \neq 0$. In Satz 7.12 werden wir sehen, dass diese Bedingung sogar hinreichend für die Differenzierbarkeit von g ist.

(b) Sei allgemeiner $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ eine C^1 -Funktion mit Nullstellenmenge

$$M := \{(x, y) \in \mathbb{R}^2 \mid F(x, y) = 0\}.$$

Frage: Können wir die Gleichung $F(x, y) = 0$ nach x oder y auflösen? Mit anderen Worten: Können wir M als Graph einer Funktion $y = y(x)$ oder $x = x(y)$ schreiben (evtl. mit C^1 Funktionen?).

Ist F linear und nicht identisch Null, dann gilt $F(x, y) = ax + by$, mit $a, b \in \mathbb{R}$ und $a^2 + b^2 \neq 0$. Ferner ist $M = \text{Kern } F$ ein eindimensionaler Untervektorraum. Für $b \neq 0$ ist $\partial_y F \equiv b \neq 0$ und $M = \{(x, y) \in \mathbb{R}^2 \mid y = -\frac{a}{b}x\}$ der Graph einer Funktion $y = y(x)$. Für $a \neq 0$ ist $\partial_x F \equiv a \neq 0$ und $M = \{(x, y) \in \mathbb{R}^2 \mid x = -\frac{b}{a}y\}$ der Graph einer Funktion $x = x(y)$.

Ist F nicht linear, dann erhalten wir als Nullstellenmengen nicht mehr Untervektorräume, sondern im besten Fall „eindimensionale gekrümmte Mengen“. Diese kann man im Allgemeinen nicht insgesamt als Graphen einer Funktion darstellen. Sei z.B. $F(x, y) = x^2 + y^2 - 1$. Dann ist $M := \{(x, y) \mid |(x, y)| = 1\}$ der Einheitskreis in \mathbb{R}^2 , und diese Menge ist offensichtlich kein Graph einer Funktion. Allerdings existieren zu jedem $(x, y) \in M$ Umgebungen $U \subseteq \mathbb{R}^2$ von x bzw. $V \subseteq \mathbb{R}^2$ von y und

- eine C^1 -Funktion $g: U \rightarrow V$ mit

$$M \cap (U \times V) = \text{Graph } g = \{(\xi, g(\xi)) \mid \xi \in U\}$$

- **oder** eine C^1 -Funktion $g: V \rightarrow U$ mit

$$M \cap (U \times V) = \{(g(\xi), \xi) \mid \xi \in V\}.$$

Z.B. kann man für $(x, y) = (1, 0)$ die Mengen $U = (0, \infty)$, $V = (-1, 1)$ wählen. Damit ist

$$M \cap (U \times V) = \{(\sqrt{1 - \xi^2}, \xi) \mid \xi \in (-1, 1)\} \quad (\text{rechter Halbkreis}).$$

Der untenstehende Satz 7.12 liefert ein allgemeines Kriterium für die **lokale** Auflösbarkeit von Gleichungen bzw. die **lokale** Darstellbarkeit von Nullstellenmengen als Graphen.

Satz 7.11. Die Menge $\text{GL}(N, \mathbb{R})$ ist offen in $\mathbb{R}^{N \times N}$, und die Abbildung $\text{GL}(N, \mathbb{R}) \rightarrow \text{GL}(N, \mathbb{R})$, $A \mapsto A^{-1}$, ist beliebig oft differenzierbar.

Beweis. Die Determinante $\det: \mathbb{R}^{N \times N} \rightarrow \mathbb{R}$ ist eine Polynomfunktion, also beliebig oft differenzierbar. Somit ist $\text{GL}(N, \mathbb{R})$ nach Satz 4.35 offen in $\mathbb{R}^{N \times N}$ als Urbild von $\mathbb{R} \setminus \{0\}$ unter \det . Weiterhin gilt nach Satz 3.22:

$$A^{-1} = \frac{1}{\det A} \tilde{A}^t \quad \text{mit } \tilde{A} = \left((-1)^{i+j} A_{ij} \right)_{ij},$$

wobei A_{ij} die i, j -te Minorante von A ist. Als Determinanten stellen die A_{ij} beliebig oft differenzierbare Funktionen von A dar. Dasselbe gilt also auch für die Abbildung

$$\text{GL}(N, \mathbb{R}) \rightarrow \text{GL}(N, \mathbb{R}), \quad A \mapsto A^{-1}. \quad \square$$

Satz 7.12 (über implizite Funktionen). Seien $k, m \in \mathbb{N}$ und $N = k + m$, $U \subseteq \mathbb{R}^N$ offen, $F \in C^1(U, \mathbb{R}^m)$. Wir schreiben $z = (x, y) \in \mathbb{R}^N$ mit $x = (x_1, \dots, x_k)^t \in \mathbb{R}^k$, $y = (y_1, \dots, y_m)^t \in \mathbb{R}^m$ sowie

$$J_F(z) = \left(\partial_x F(z) \quad \partial_y F(z) \right) \in \mathbb{R}^{m \times N}$$

mit

$$\partial_x F(z) := \left(\frac{\partial F}{\partial x_1}(z) \mid \frac{\partial F}{\partial x_2}(z) \mid \cdots \mid \frac{\partial F}{\partial x_k}(z) \right) \in \mathbb{R}^{m \times k}$$

und

$$\partial_y F(z) := \left(\frac{\partial F}{\partial y_1}(z) \mid \frac{\partial F}{\partial y_2}(z) \mid \cdots \mid \frac{\partial F}{\partial y_m}(z) \right) \in \mathbb{R}^{m \times m}$$

für $z \in U$.

Sei ferner $z_0 = (x_0, y_0) \in U$ eine Nullstelle von F , welche $\partial_y F(z_0) \in \text{GL}(m, \mathbb{R})$ erfülle. Dann existieren $\varepsilon, \delta > 0$ und eine Abbildung $g: B_\varepsilon^k(x_0) \rightarrow B_\delta^m(y_0)$ mit

$$(7.3) \quad \{(x, y) \in B_\varepsilon^k(x_0) \times B_\delta^m(y_0) \mid F(x, y) = 0\} = \text{Graph } g.$$

Insbesondere ist $g(x_0) = y_0$. Ferner ist g stetig differenzierbar, es gilt

$$(7.4) \quad J_g(x_0) = -(\partial_y F(z_0))^{-1} \partial_x F(z_0) \in \mathbb{R}^{m \times k},$$

und g ist so oft stetig differenzierbar wie F (falls F von höherer Differenzierbarkeitsordnung ist).

Beweis. Wir setzen $A := \partial_y F(z_0)$. Aufgrund der Stetigkeit von F und dF und Satz 7.11 können wir dann $\varepsilon > 0, \delta > 0$ so klein wählen, dass gilt:

$$(7.5) \quad \overline{B}_\varepsilon^k(x_0) \times \overline{B}_\delta^m(y_0) \subseteq U,$$

$$(7.6) \quad \|\partial_y F(z) - A\| \leq \frac{1}{2\|A^{-1}\|} \quad \text{für } z \in \overline{B}_\varepsilon^k(x_0) \times \overline{B}_\delta^m(y_0),$$

$$(7.7) \quad \partial_y F(z) \in \text{GL}(m, \mathbb{R}) \quad \text{für } z \in \overline{B}_\varepsilon^k(x_0) \times \overline{B}_\delta^m(y_0),$$

$$(7.8) \quad |F(x, y_0)| < \frac{\delta}{2\|A^{-1}\|} \quad \text{für } x \in \overline{B}_\varepsilon^k(x_0).$$

Wir setzen $X := \overline{B}_\varepsilon^k(x_0)$, $Y := \overline{B}_\delta^m(y_0)$ und definieren $\varphi: X \times Y \rightarrow \mathbb{R}^m$ durch

$$\varphi(x, y) := y - A^{-1}F(x, y).$$

Es gilt: $(x, y) \in X \times Y$ ist genau dann eine Nullstelle von F , wenn y ein Fixpunkt von $\varphi(x, \cdot)$ ist. Weiter folgt

$$(7.9) \quad \partial_x \varphi(z) = -A^{-1} \partial_x F(z), \quad \partial_y \varphi(z) = E_m - A^{-1} \partial_y F(z), \quad \partial_y \varphi(z_0) = 0.$$

Somit liefert (7.8)

$$(7.10) \quad |\varphi(x, y_0) - y_0| \leq \|A^{-1}\| |F(x, y_0)| \leq \frac{\delta}{2} \quad \text{für alle } x \in X,$$

und (7.6) liefert

$$\|\partial_y \varphi(z)\| = \|A^{-1}(A - \partial_y F(z))\| \leq \frac{1}{2} \quad \text{für alle } z \in X \times Y.$$

Zusammen mit dem Schrankensatz, Satz 5.22, folgt hieraus

$$(7.11) \quad |\varphi(x, y_1) - \varphi(x, y_2)| \leq \frac{1}{2} |y_1 - y_2| \quad \text{für alle } x \in X, y_1, y_2 \in Y.$$

Für $(x, y) \in X \times Y$ erhalten wir aus (7.10), (7.11) und wegen $|y - y_0| \leq \delta$

$$|\varphi(x, y) - y_0| \leq |\varphi(x, y) - \varphi(x, y_0)| + |\varphi(x, y_0) - y_0| \leq \frac{1}{2}|y - y_0| + \frac{\delta}{2} \leq \delta.$$

Demnach gilt $\varphi: X \times Y \rightarrow Y$. Wegen (7.11) ist $\varphi(x, \cdot)$ für jedes $x \in X$ eine $\frac{1}{2}$ -Kontraktion. Da ferner $\|\partial_x \varphi\|$ auf der kompakten Menge $X \times Y$ stetig, also beschränkt ist, existiert nach dem Schrankensatz $L > 0$ mit

$$(7.12) \quad |\varphi(x_1, y) - \varphi(x_2, y)| \leq L|x_1 - x_2| \quad \text{für alle } x_1, x_2 \in X, y \in Y.$$

Der BFPS mit Parameterabhängigkeit, Korollar 7.9, liefert nun, dass eine Lipschitzstetige Funktion $g: X \rightarrow Y$ existiert, so dass (7.3) gilt.

Wir zeigen, dass g in x_0 differenzierbar ist und dass (7.4) gilt. Für $h \in B_\varepsilon^k(0)$ folgt aus der Differenzierbarkeit von φ , der Fixpunkteigenschaft und (7.9)

$$\begin{aligned} & |g(x_0 + h) - g(x_0) + A^{-1}\partial_x F(z_0)h| \\ &= |\varphi(x_0 + h, g(x_0 + h)) - \varphi(x_0, g(x_0)) - \partial_x \varphi(z_0)h| \\ &= |\varphi(x_0 + h, g(x_0 + h)) - \varphi(x_0, g(x_0)) - d\varphi(z_0)(h, g(x_0 + h) - g(x_0))| \\ &= o(|(h, g(x_0 + h) - g(x_0))|) \\ &= o(|h|) \quad \text{für } h \rightarrow 0. \end{aligned}$$

Die letzte asymptotische Aussage ist richtig, weil g Lipschitzstetig ist: Es gilt dann nämlich $g(x_0 + h) - g(x_0) = O(|h|)$ für $h \rightarrow 0$. Damit ist die Differenzierbarkeit in x_0 gezeigt.

Für jedes $x \in X$ sind wegen (7.7) die Voraussetzungen des Satzes im Punkt $(x, g(x))$ anstelle von z_0 wieder erfüllt, und aus dem bisher Bewiesenen folgt dann — unter Berücksichtigung der punktwisen Eindeutigkeit der Lösungsfunktion g — die Differenzierbarkeit von g in x mit

$$(7.13) \quad J_g(x) = -(\partial_y F(x, g(x)))^{-1} \partial_x F(x, g(x)).$$

Da dF stetig ist und die Inverse $GL(m, \mathbb{R}) \rightarrow GL(m, \mathbb{R})$ gemäß Satz 7.11 sogar unendlich oft differenzierbar, folgt die Stetigkeit von dg , d.h. g ist stetig differenzierbar. Ist F sogar zweimal stetig differenzierbar, dann ist dF stetig differenzierbar. Weil g stetig differenzierbar ist, liefert (7.13), dass dg stetig differenzierbar ist, also dass g auch zweimal stetig differenzierbar ist. Induktiv zeigt man so, dass g so oft stetig differenzierbar ist wie F . \square

Ende der Sitzung vom 2.7.

Bemerkungen und Beispiele 7.13. (a) Fortsetzung von Bemerkung und Beispiel 7.10(b): Sei $F: \mathbb{R}^2 \rightarrow \mathbb{R}$, $F(x, y) = x^2 + y^2 - 1$ und $M = \{(x, y) \mid |(x, y)| = 1\} \subseteq \mathbb{R}^2$. Sei ferner $z_0 = (x_0, y_0) \in M$ mit $y_0 \neq 0$. Dann ist $\partial_2 F(z_0) = 2y_0 \neq 0$. Al-

so existieren nach Satz 7.12 Umgebungen $B_\varepsilon(x_0), B_\delta(y_0)$ und eine C^1 -Abbildung $g: B_\varepsilon(x_0) \rightarrow B_\delta(y_0)$ derart, dass gilt:

$$M \cap (B_\varepsilon(x_0) \times B_\delta(y_0)) = \text{Graph } g = \left\{ (x, g(x)) \mid x \in B_\varepsilon(x_0) \right\}.$$

In diesem speziellen Fall ist $g(x) = \text{sign}(y_0)\sqrt{1-x^2}$ für $x \in B_\varepsilon(x_0)$.

(b) Wir betrachten das (unterbestimmte) nichtlineare Gleichungssystem

$$(7.14) \quad \left\{ \begin{array}{l} 3x^2 - y^2 + u^3 + v^2 = 4 \\ 2x^2 - y^3 + u^2 - v = 1 \end{array} \right\} \quad (x, y, u, v) \in \mathbb{R}^4.$$

Offensichtlich ist $z_0 = (1, 1, 1, 1) \in \mathbb{R}^4$ eine Lösung. Kann man (7.14) lokal bei z_0 durch eine C^1 -Abbildung nach (u, v) auflösen? Um dies zu klären, schreiben wir (7.14) in der Form $F(x, y, u, v) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ mit

$$F \in C^1(\mathbb{R}^4, \mathbb{R}^2), \quad F(x, y, u, v) = \begin{pmatrix} 3x^2 - y^2 + u^3 + v^2 - 4 \\ 2x^2 - y^3 + u^2 - v - 1 \end{pmatrix}.$$

Dann ist $F(z_0) = 0$ und

$$\partial_{(u,v)} F(z) = \begin{pmatrix} \frac{\partial F_1}{\partial u}(z) & \frac{\partial F_1}{\partial v}(z) \\ \frac{\partial F_2}{\partial u}(z) & \frac{\partial F_2}{\partial v}(z) \end{pmatrix} = \begin{pmatrix} 3u^2 & 2v \\ 2u & -1 \end{pmatrix} \quad \text{für } z = (x, y, u, v) \in \mathbb{R}^4,$$

also $\partial_{(u,v)} F(z_0) = \begin{pmatrix} 3 & 2 \\ 2 & -1 \end{pmatrix} \in \text{GL}(2, \mathbb{R})$, mit Determinante -7 . Nach Satz 7.12 existieren also $\varepsilon, \delta > 0$ und eine C^1 -Abbildung $g: B_\varepsilon^2((1, 1)) \rightarrow B_\delta^2((1, 1))$ derart, dass für $z = (x, y, u, v) \in B_\varepsilon^2((1, 1)) \times B_\delta^2((1, 1))$ gilt:

$$F(z) = 0, \text{ (d.h. } z \text{ löst (7.14))} \Leftrightarrow (u, v) = g(x, y).$$

Wir bestimmen nun $J_g(1, 1)$: Es ist $(\partial_{(u,v)} F(z_0))^{-1} = -\frac{1}{7} \begin{pmatrix} -1 & -2 \\ -2 & 3 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 1 & 2 \\ 2 & -3 \end{pmatrix}$ und

$$\partial_{(x,y)} F(z_0) := \begin{pmatrix} \frac{\partial F_1}{\partial x}(z_0) & \frac{\partial F_1}{\partial y}(z_0) \\ \frac{\partial F_2}{\partial x}(z_0) & \frac{\partial F_2}{\partial y}(z_0) \end{pmatrix} = \begin{pmatrix} 6x & -2y \\ 4x & -3y^2 \end{pmatrix} \Big|_{z_0} = \begin{pmatrix} 6 & -2 \\ 4 & -3 \end{pmatrix}.$$

Also folgt mit Satz 7.12:

$$J_g(1, 1) = -\frac{1}{7} \begin{pmatrix} 1 & 2 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} 6 & -2 \\ 4 & -3 \end{pmatrix} = -\frac{1}{7} \begin{pmatrix} 14 & -8 \\ 0 & 5 \end{pmatrix}.$$

Im nächsten Abschnitt werden wir die Frage behandeln, ob man (7.14) lokal bei z_0 auch nach anderen Variablen, z.B. nach (x, y) , auflösen kann.

7.3 Reguläre Punkte und reguläre Werte

Definition 7.14. Sei $U \subseteq \mathbb{R}^N$ offen, $f \in C^1(U, \mathbb{R}^m)$.

- (a) $z \in U$ heißt *regulärer Punkt* von f , falls $df(z) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^m)$ surjektiv ist.
- (b) $w \in \mathbb{R}^m$ heißt *regulärer Wert* von f , falls die Menge $f^{-1}(w) \subseteq U$ nur aus regulären Punkten von f besteht.
- (c) Nichtreguläre Punkte bzw. Werte von f nennt man *singulär* oder *kritisch*.

Bemerkung 7.15. In der Situation von Definition 7.14 gilt für $z \in U$, $w \in \mathbb{R}^m$:

- (a) z ist ein regulärer Punkt von f .
 - \Leftrightarrow Rang $J_f(z) = m$.
 - \Leftrightarrow die Zeilen von $J_f(z)$ sind linear unabhängig im \mathbb{R}^N .
 - \Leftrightarrow $J_f(z)$ hat m linear unabhängige Spalten.
 Ist z ein regulärer Punkt von f , dann muss also insbesondere gelten $m \leq N$.
- (b) Ist $w \notin f(U)$ so ist w ein regulärer Wert von f .
- (c) Ist $m = 1$, so ist z regulärer Punkt von f genau dann, wenn $\nabla f(z) \neq 0$ ist.

Beispiel 7.16. (a) Sei $f \in C^\infty(\mathbb{R}^2)$ definiert durch $f(x, y) = x^2 - y^2 + 1$. Dann gilt:

$$\nabla f(x, y) = 2 \begin{pmatrix} x \\ -y \end{pmatrix} = 0 \Leftrightarrow (x, y) = (0, 0).$$

Also ist $(0, 0)$ der einzige kritische Punkt von f und somit $1 = f(0, 0)$ der einzige kritische Wert von f . Man beachte aber: $(1, 1)$ ist ein regulärer Punkt von f , aber $f(1, 1) = 1$ ist **kein** regulärer Wert von f ! Das Urbild $f^{-1}(1)$ besteht aus den Diagonalengeraden, welche sich im Ursprung senkrecht schneiden. In einer Umgebung von $(0, 0)$ lässt sich diese Menge daher nicht als Graph einer Funktion darstellen (Übung).

Zur Veranschaulichung ist es hilfreich, die Mengen $f^{-1}(c) \subseteq \mathbb{R}^2$ (sogenannte Niveaumengen von f) für $c > 1$, $c = 1$ und $c < 1$ in der Nähe des Nullpunkts im \mathbb{R}^2 zu skizzieren. Man kann generell zeigen, dass sich die topologische Gestalt der Niveaumengen von f höchstens bei Überschreitung von kritischen Werten ändern kann.

- (b) Sei $f \in C^1(\mathbb{R}^3, \mathbb{R}^2)$ definiert durch $f(x, y, z) := \begin{pmatrix} x^2 + y^2 + z^2 \\ (x-1)^2 + y^2 + z^2 \end{pmatrix}$. Dann gilt:

$$J_f(x, y, z) = 2 \begin{pmatrix} x & y & z \\ x-1 & y & z \end{pmatrix} \text{ hat Rang } 2 \Leftrightarrow y \neq 0 \text{ oder } z \neq 0.$$

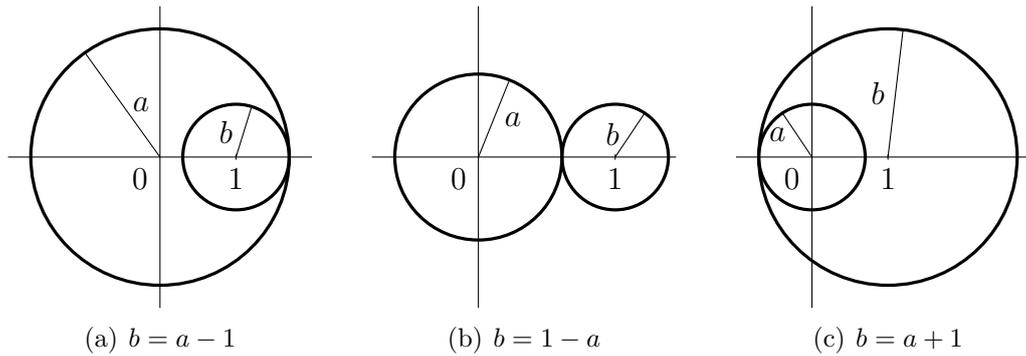


Abbildung 7.1: $f^{-1}(a^2, b^2)$ für kritische Werte (a^2, b^2) .

Also ist $\{(x, 0, 0) \mid x \in \mathbb{R}\} \subseteq \mathbb{R}^3$ die Menge der kritischen Punkte und

$$\{(x^2, (x-1)^2) \mid x \in \mathbb{R}\}$$

die Menge der kritischen Werte von f .

Seien nun $a, b > 0$. Geometrisch ist $f^{-1}(a^2, b^2)$ der Schnitt einer Kugeloberfläche (der sogenannten Sphäre) mit Radius a und Mittelpunkt $(0, 0, 0)$ mit einer Sphäre mit Radius b und Mittelpunkt $(1, 0, 0)$, siehe Abb. 7.1. Dabei gilt:

$$\begin{aligned} (a^2, b^2) \text{ ist ein kritischer Wert von } f & \\ \Leftrightarrow \text{es gibt } x \in \mathbb{R} \text{ mit } a^2 = x^2 \text{ und } b^2 = (x-1)^2 & \\ \Leftrightarrow b = a - 1 \text{ oder } b = 1 - a \text{ oder } b = 1 + a & \\ \Leftrightarrow \text{die beiden Sphären berühren sich genau in einem Punkt.} & \end{aligned}$$

Anhand der folgenden Variante des Satzes von der impliziten Funktion sehen wir, dass sich das Urbild eines regulären Wertes nach einer Permutation der Koordinaten von \mathbb{R}^N lokal als Graph schreiben lässt. Mit einer Koordination permutation in \mathbb{R}^N sei dabei ein $T \in \mathcal{L}(\mathbb{R}^N)$ bezeichnet, welches für eine geeignete Permutation (bijektive Abbildung) $\pi: \{1, 2, \dots, N\} \rightarrow \{1, 2, \dots, N\}$ durch $Te_j = e_{\pi(j)}$, $j = 1, \dots, N$ gegeben ist.

Satz 7.17. Seien $U \subseteq \mathbb{R}^N$ offen, $f \in C^1(U, \mathbb{R}^m)$ mit $m < N$ und $z \in U$ ein regulärer Punkt von f . Seien ferner $k := N - m$ und $w := f(z)$. Dann existieren offene Teilmengen $\tilde{U} \subseteq \mathbb{R}^k$ und $V \subseteq U$ mit $z \in V$, eine Abbildung $g \in C^1(\tilde{U}, \mathbb{R}^m)$ und eine Koordinatenpermutation $T \in \mathcal{L}(\mathbb{R}^N)$ derart, dass gilt:

$$f^{-1}(w) \cap V = \left\{ T(x, g(x)) \mid x \in \tilde{U} \right\}.$$

Insbesondere lässt sich das Urbild $f^{-1}(w)$ eines regulären Wertes $w \in f(U)$ von f lokal als Graph einer eindeutigen C^1 -Abbildung schreiben, wenn man die Koordinaten geeignet permutiert.

Beweis. Da z ein regulärer Punkt von f ist, sind für eine geeignete Permutation π der Koordinaten in \mathbb{R}^N die Vektoren $\partial_{\pi(k+1)}f(z), \dots, \partial_{\pi(N)}f(z) \in \mathbb{R}^m$ linear unabhängig. Sei $T \in \mathcal{L}(\mathbb{R}^N)$ die zu π gehörige Koordinatenpermutation, und seien $U_0 := T^{-1}(U)$ sowie $z_0 := T^{-1}z \in U_0$. Sei ferner $F \in C^1(U_0, \mathbb{R}^m)$ definiert durch $F(u) = f(Tu) - w$. Dann gilt $F(z_0) = 0$ und (mit Kettenregel)

$$\partial_j F(z_0) = dF(z_0)e_j = df(Tz_0)Te_j = df(z)e_{\pi(j)} = \partial_{\pi(j)}f(z), \quad j = 1, \dots, N.$$

Nach Voraussetzung sind also die Vektoren

$$\partial_{k+1}F(z_0), \dots, \partial_N F(z_0) \in \mathbb{R}^m$$

linear unabhängig. Wir schreiben nun wie in Satz 7.12 die Elemente von U_0 als Tupel (x, y) mit $x = (x_1, \dots, x_k) \in \mathbb{R}^k$ und $y = (y_1, \dots, y_m) \in \mathbb{R}^m$. Insbesondere sei $z_0 = (x_0, y_0)$. Gemäß der Notation von Satz 7.12 ist dann $\partial_y F(z_0)$ die von den Spalten $\partial_{k+1}F(z_0), \dots, \partial_N F(z_0)$ gebildete Matrix, und diese ist invertierbar. Nach Satz 7.12 existieren also $\varepsilon, \delta > 0$ und eine C^1 -Abbildung $g: B_\varepsilon^k(x_0) \rightarrow B_\delta^m(y_0)$ mit

$$(7.15) \quad F^{-1}(0) \cap (B_\varepsilon^k(x_0) \times B_\delta^m(y_0)) = \{(x, g(x)) \mid x \in B_\varepsilon^k(x_0)\}.$$

Wir setzen $V := T(B_\varepsilon(x_0) \times B_\delta(y_0))$ und $\tilde{U} := B_\varepsilon^k(x_0)$. Dann ist V offen, da $T \in \text{GL}(N, \mathbb{R})$ ist. Beachtet man nun, dass $T(F^{-1}(0)) = f^{-1}(w)$ gilt, so folgt aus (7.15)

$$f^{-1}(w) \cap V = T(F^{-1}(0) \cap (B_\varepsilon^k(x_0) \times B_\delta^m(y_0))) = \{T(x, g(x)) \mid x \in \tilde{U}\},$$

wie behauptet. □

Bemerkung 7.18. Um höhere Ableitungen einer implizit gegebenen Funktion zu berechnen, ist es am einfachsten, wie folgt vorzugehen: Lösen wir zum Beispiel $f(x, y, z) = 0$ nach y durch eine Funktion g auf, dann leiten wir die Gleichung

$$f(x, g(x, z), z) = 0$$

nach x ab und erhalten (mit $f_x := \partial_x f$ usw.):

$$(7.16) \quad f_x(x, g(x, z), z) + f_y(x, g(x, z), z)g_x(x, z) = 0,$$

kurz $f_x + f_y g_x = 0$, also $g_x = -f_x/f_y$. Analog erhält man $g_z = -f_z/f_y$. Um höhere Ableitungen zu erhalten, können wir entweder diese Ausdrücke direkt ableiten oder (7.16) weiter differenzieren und dann nach der gewünschten Ableitung von g auflösen.

Beispiel 7.19. Die Gleichung

$$(7.17) \quad y^5 - 2y^3 \cos(x - z) + 1 = 0$$

hat die Lösung $w_0 = (0, 1, 0)$. Da es hoffnungslos ist, explizite Formeln für weitere Lösungen zu finden, zeigen wir, dass (7.17) in der Nähe von w_0 nach y durch eine

Funktion $g(x, z)$ auflösbar ist. Sei also

$$F(x, y, z) := y^5 - 2y^3 \cos(x - z) + 1.$$

Es gilt $\partial_y F(x, y, z) = 5y^4 - 6y^2 \cos(x - z)$ und daher $\partial_y F(0, 1, 0) = -1 \neq 0$. Nach Satz 7.17 existiert eine beliebig oft differenzierbare Funktion $g = g(x, z)$ in einer Umgebung von $(0, 0)$ in \mathbb{R}^2 mit $g(0, 0) = 1$ und

$$(7.18) \quad F(x, g(x, z), z) = g^5 - 2g^3 \cos(x - z) + 1 \equiv 0.$$

Hier und im Folgenden lassen wir der Übersicht halber die variablen Funktionsargumente weg. Ableiten von (7.18) nach x ergibt

$$(7.19) \quad 5g^4 g_x - 6g^2 g_x \cos(x - z) + 2g^3 \sin(x - z) \equiv 0.$$

In $(0, 0)$ ausgewertet liefert dies

$$5g_x(0, 0) - 6g_x(0, 0) = 0, \quad \text{also } g_x(0, 0) = 0.$$

Ableiten von (7.19) nach z und Auswertung in $(0, 0)$ liefert, zusammen mit den schon bekannten Werten:

$$5g_{xz}(0, 0) - 6g_{xz}(0, 0) - 2 = 0, \quad \text{also } g_{xz}(0, 0) = -2.$$

Weitere partielle Ableitungen von g in $(0, 0)$ bestimmt man analog.

Ende der Sitzung vom 6.7.

7.4 Lokale Diffeomorphie

Bemerkung 7.20. (a) Seien $U \subseteq \mathbb{R}^N$, $V \subseteq \mathbb{R}^m$ offen, und sei $f: U \rightarrow V$ eine C^1 -Abbildung, die bijektiv ist und eine stetig differenzierbare Inverse $f^{-1}: V \rightarrow U$ besitzt. Differenzieren der Gleichungen $\text{id}_U = f^{-1} \circ f$ und $\text{id}_V = f \circ f^{-1}$ liefert nach Kettenregel:

$$\begin{aligned} E_N &= J_{\text{id}_U}(x) = J_{f^{-1}}(f(x)) J_f(x) && \text{für alle } x \in U \text{ und} \\ E_m &= J_{\text{id}_V}(y) = J_f(f^{-1}(y)) J_{f^{-1}}(y) && \text{für alle } y \in V, \text{ d.h.} \\ E_m &= J_f(x) J_{f^{-1}}(f(x)) && \text{für alle } x \in U. \end{aligned}$$

Für ein $x \in U$ ist demnach $df(x)$ injektiv, also $N \leq m$, und $df^{-1}(f(x))$ ist injektiv, also $m \leq N$. Es folgt $m = N$ und $J_{f^{-1}}(f(x)) = (J_f(x))^{-1}$ für alle $x \in U$.

(b) Im eindimensionalen Fall gilt speziell (vgl. Mathe 1): Ist $U \subseteq \mathbb{R}$ ein offenes Intervall und $f \in C^1(U)$ mit $f'(x) \neq 0$ für alle $x \in U$, so ist $V := f(U)$ offen und $f: U \rightarrow V$ hat eine C^1 -Umkehrfunktion $f^{-1}: V \rightarrow U$ mit $(f^{-1})'(f(x)) = \frac{1}{f'(x)}$.

- (c) Die obigen Überlegungen motivieren folgende Frage: Ist im mehrdimensionalen Fall $m = N > 1$ die Invertierbarkeit der Matrix $J_f(x_0)$ in einem Punkt $x_0 \in U$ hinreichend für die „lokale Invertierbarkeit“ von f in der Nähe des Punktes x_0 ?

Definition 7.21. Seien $U, V \subseteq \mathbb{R}^N$ offen. Eine C^1 -Abbildung $f: U \rightarrow V$ heißt *Diffeomorphismus*, falls f bijektiv ist und $f^{-1}: V \rightarrow U$ ebenfalls eine C^1 -Abbildung ist. Wir schreiben: $\text{Diff}(U, V) = \{f: U \rightarrow V \mid f \text{ Diffeomorphismus}\}$.

Beispiel 7.22. (a) Sei $f: \mathbb{R}^N \rightarrow \mathbb{R}^N$ affin linear, also gegeben durch $f(x) = Tx + c$, wobei $T \in \mathcal{L}(\mathbb{R}^N)$ invertierbar und $c \in \mathbb{R}^N$ sei. Dann ist $f \in \text{Diff}(\mathbb{R}^N, \mathbb{R}^N)$ mit $f^{-1}(y) = T^{-1}(y - c)$ für $y \in \mathbb{R}^N$.

- (b) $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^3$ ist eine bijektive C^1 -Abbildung mit $f^{-1}: \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$f^{-1}(y) = \begin{cases} \sqrt[3]{y}, & y \geq 0, \\ -\sqrt[3]{|y|} & y < 0. \end{cases}$$

f^{-1} ist stetig, aber in 0 nicht differenzierbar. Also ist f ein Homöomorphismus, aber kein Diffeomorphismus.

Bemerkung 7.23. Seien $U, V, W \subseteq \mathbb{R}^N$ offen und $f \in \text{Diff}(U, V), g \in \text{Diff}(V, W)$. Dann gilt:

- (a) $f^{-1} \in \text{Diff}(V, U)$. Dabei ist $df^{-1}(f(x)) = (df(x))^{-1}$ und $J_{f^{-1}}(f(x)) = (J_f(x))^{-1}$ für alle $x \in U$ gemäß Bemerkung 7.20(a).
- (b) $g \circ f \in \text{Diff}(U, W)$.

Satz 7.24 (lokale Diffeomorphie bzw. von der inversen Funktion). *Seien $U \subseteq \mathbb{R}^N$ offen, $f \in C^1(U, \mathbb{R}^N)$ und $a \in U$ mit $J_f(a) \in \text{GL}(N, \mathbb{R})$. Dann existieren offene Umgebungen $\tilde{U} \subseteq U$ von a und $V \subseteq \mathbb{R}^N$ von $f(a)$ derart, dass $f|_{\tilde{U}}: \tilde{U} \rightarrow V$ ein Diffeomorphismus ist.*

Beweis. Wir betrachten die Abbildung

$$F: \mathbb{R}^N \times U \rightarrow \mathbb{R}^N, \quad F(x, y) = f(y) - x,$$

und wenden den Satz über implizite Funktion Satz 7.12 auf F im Punkt $z_0 = (f(a), a)$ an. Dies ist möglich, da

$$\frac{\partial F}{\partial y}(z_0) = J_f(a) \in \mathbb{R}^{N \times N}$$

nach Voraussetzung invertierbar ist. Somit existieren $\varepsilon, \delta > 0$ sowie eine Abbildung $g: B_\varepsilon(f(a)) \rightarrow B_\delta(a)$ derart, dass für $(x, y) \in B_\varepsilon(f(a)) \times B_\delta(a)$ gilt:

$$(7.20) \quad x = f(y) \quad \Leftrightarrow \quad F(x, y) = 0 \quad \Leftrightarrow \quad y = g(x).$$

Ferner ist g stetig differenzierbar. Da f stetig ist, ist $\tilde{U} := f^{-1}(B_\varepsilon(f(a))) \cap B_\delta(a)$ eine offene Teilmenge von U , und wegen (7.20) ist auch

$$V := f(\tilde{U}) = g^{-1}(\tilde{U}) \subseteq B_\varepsilon(f(a)) \subseteq \mathbb{R}^N$$

offen. Aus (7.20) folgt nun, dass $f|_{\tilde{U}}: \tilde{U} \rightarrow V$ bijektiv ist mit der Inversen $g|_V: V \rightarrow \tilde{U}$. Da f und g stetig differenzierbar sind, folgt $f|_{\tilde{U}} \in \text{Diff}(\tilde{U}, V)$, wie behauptet. \square

Korollar 7.25. *Ist $U \subseteq \mathbb{R}^N$ offen und $f \in C^1(U, \mathbb{R}^N)$ eine Abbildung mit $\det J_f(x) \neq 0$ für alle $x \in U$, so gilt:*

- (a) $f(U) \subseteq \mathbb{R}^N$ ist offen.
- (b) Ist f injektiv, so ist $f \in \text{Diff}(U, f(U))$.

Beweis. Sei $y \in f(U)$ beliebig und $x \in U$ mit $f(x) = y$. Nach Satz 7.24 existieren offene Umgebungen $\tilde{U} \subseteq U$ von x und $V \subseteq \mathbb{R}^N$ von y derart, dass $f|_{\tilde{U}}: \tilde{U} \rightarrow V$ ein Diffeomorphismus ist. Insbesondere ist $V = f(\tilde{U}) \subseteq f(U)$. Es folgt, dass $f(U)$ in \mathbb{R}^N offen ist.

Ist zudem f injektiv mit der Inversen $f^{-1}: f(U) \rightarrow U$, so ist $f^{-1}|_V = \left(f|_{\tilde{U}}\right)^{-1}$ stetig differenzierbar. Da $y \in f(U)$ beliebig gewählt war, folgt die stetige Differenzierbarkeit von f^{-1} , und somit ist $f \in \text{Diff}(U, f(U))$. \square

Beispiel 7.26 (Polarkoordinaten). Seien $U := (0, \infty) \times \mathbb{R} \subseteq \mathbb{R}^2$ und $f \in C^1(U, \mathbb{R}^2)$ definiert durch $f(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$. Dann ist

$$J_f(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} \quad \text{für } (r, \varphi) \in U.$$

Also ist $\det J_f(r, \varphi) = r(\cos^2 \varphi + \sin^2 \varphi) = r > 0$ und damit $J_f(r, \varphi)$ invertierbar für alle $(r, \varphi) \in U$. Somit sind die Sätze Satz 7.24 und Korollar 7.25(a) auf f anwendbar. Allerdings ist f nicht injektiv, da

$$f(r, \varphi + k2\pi) = f(r, \varphi) \quad \text{für alle } (r, \varphi) \in U, k \in \mathbb{Z}.$$

Wir betrachten nun die offenen Teilmengen

$$U' := (0, \infty) \times (-\pi, \pi) \subseteq U, \quad V := \mathbb{R}^2 \setminus \{(x, 0) \mid x \leq 0\}$$

und zeigen:

$$(7.21) \quad f|_{U'} \in \text{Diff}(U', V).$$

Dazu beachten wir, dass man nach Mathe 1 jede Zahl $z \in \mathbb{C} \setminus \{0\}$ in eindeutiger Weise als $z = re^{i\varphi}$ mit $r > 0$ und $\varphi \in (-\pi, \pi]$ schreiben kann. Durch Übergang zu Real- und Imaginärteil folgt, dass man jeden Vektor $(x, y) \in \mathbb{R}^2 \setminus \{0\}$ in eindeutiger Weise als $(x, y) = (r \cos \varphi, r \sin \varphi) = f(r, \varphi)$ schreiben kann mit $(r, \varphi) \in (0, \infty) \times (-\pi, \pi]$. Da ferner $f((0, \infty) \times \{\pi\}) = \{(x, 0) \mid x < 0\}$ ist, bildet f also U' bijektiv auf V ab. Mit Korollar 7.25(b) folgt (7.21). Die Abbildung $P := f|_{U'}$ heißt *Polarkoordinatenabbildung*.

Wie berechnet man nun die Jacobimatrix von P^{-1} z.B. in $(1, 1) \in V$? Es ist $(1, 1) = P(\sqrt{2}, \frac{\pi}{4})$, wobei $J_P(\sqrt{2}, \frac{\pi}{4}) = \begin{pmatrix} \frac{1}{\sqrt{2}} & -1 \\ \frac{1}{\sqrt{2}} & 1 \end{pmatrix}$. Also gilt nach Bemerkung 7.20 bzw.

Bemerkung 7.23:

$$J_{P^{-1}}(1, 1) = \left(J_P \left(\sqrt{2}, \frac{\pi}{4} \right) \right)^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

7.5 Extremwertbestimmung unter Nebenbedingungen und Untermannigfaltigkeiten

Bemerkung und Beispiel 7.27. Seien $D \subseteq \mathbb{R}^N$ offen und $f \in C^1(D)$. Oft interessiert man sich für die Extremwerte von $f|_M$, wobei M eine gegebene Teilmenge von D ist. Man sucht also Extrema x von f „unter der Nebenbedingung $x \in M$ “.

Konkretes Beispiel: Sei $D = \mathbb{R}^3$,

$$M = \left\{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + \frac{y^2}{4} + \frac{z^2}{9} = 1 \right\} \quad (\text{Ellipsoid})$$

und $q = (0, 0, 1)$. Da M abgeschlossen und beschränkt, also kompakt ist, existiert ein Punkt $p \in M$, der den Abstand zu q minimiert. Es gilt also

$$|p - q| = \text{dist}(q, M).$$

Wie berechnet man einen solchen Punkt p ?

Ist $f \in C^1(\mathbb{R}^3)$ gegeben durch $f(v) = |v - q|^2$, so ist p ein globales Minimum von $f|_M$. Offensichtlich ist p aber kein kritischer Punkt von f , denn der einzige kritische Punkt von f ist q selbst. Welche Bedingung erfüllt $\nabla f(p)$ dann?

Definition 7.28. Seien $M \subseteq \mathbb{R}^N$ eine Menge und $p \in M$. Ein Vektor $v \in \mathbb{R}^N$ heißt *Tangententialvektor* von M in p , falls es $\varepsilon > 0$ und eine C^1 -Kurve $\gamma: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^N$ gibt mit $\gamma((-\varepsilon, \varepsilon)) \subseteq M$, $\gamma(0) = p$ und $\gamma'(0) = v$.

Satz 7.29. Seien $D \subseteq \mathbb{R}^N$ offen, $M \subseteq D$, $f \in C^1(D)$ und $p \in M$ ein lokales Extremum von $f|_M$. Dann steht $\nabla f(p)$ senkrecht auf allen Tangentialvektoren von M in p , d.h. für jeden Tangentialvektor v von M in p gilt

$$\langle \nabla f(p), v \rangle = 0.$$

Beweis. Seien $v \in \mathbb{R}^N$ ein Tangentialvektor von M in p und $\gamma: (-\varepsilon, \varepsilon) \rightarrow M$ eine C^1 -Kurve mit $\gamma(0) = p$, $\gamma'(0) = v$. Nach Voraussetzung ist 0 ein lokales Extremum der Funktion $f \circ \gamma: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$. Also ist

$$0 = (f \circ \gamma)'(0) = \langle \nabla f(\gamma(0)), \gamma'(0) \rangle = \langle \nabla f(p), v \rangle. \quad \square$$

Definition 7.30. Seien $D \subseteq \mathbb{R}^N$ offen und $f \in C^1(D)$. Seien ferner $M \subseteq D$ und $p \in M$. Wir nennen $p \in M$ einen *kritischen Punkt* von $f|_M$, falls $\langle \nabla f(p), v \rangle = 0$ für alle Tangentialvektoren v von M in p ist.

Bemerkung 7.31. Nach Satz 7.29 ist jedes lokale Extremum p von $f|_M$ ein kritischer Punkt von $f|_M$. Dies passt also zur Definition „ohne die Nebenbedingung“. Für allgemeine Mengen M und Punkte $p \in M$ ist die Menge aller Tangentialvektoren von M in p aber ein kompliziertes Objekt. Daher betrachten wir im Folgenden spezielle Mengen $M \subseteq \mathbb{R}^N$, nämlich solche, die lokal wie ein (gekrümmter) Ausschnitt eines Untervektorraums kleinerer Dimension aussehen.

Definition 7.32. Seien $0 \leq k \leq N$ und $M \subseteq \mathbb{R}^N$. M heißt *k-dimensionale Untermannigfaltigkeit* des \mathbb{R}^N , falls es zu jedem Punkt $p \in M$ eine offene Umgebung $V \subseteq \mathbb{R}^N$ von p und einen Diffeomorphismus $\tau: V \rightarrow \mathbb{R}^N$ gibt, so dass $M \cap V = \tau^{-1}(\mathbb{R}^k \times \{0_{N-k}\})$ gilt. Wir nennen das Paar (V, τ) eine *Einbettungskarte für M bei p* , und das Paar $(M \cap V, \tau|_{M \cap V})$ eine *Karte für M bei p* .

Bemerkung 7.33. (a) Wenn $D \subseteq \mathbb{R}^N$ offen ist und $M \subseteq D$ eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^N , dann ist für jeden Diffeomorphismus $\Phi: D \rightarrow \mathbb{R}^N$ das Bild $\Phi(M)$ auch eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^N .

(b) Sei wie in Definition 7.32 (V, τ) eine Einbettungskarte für M bei p , und sei V_1 eine offene Umgebung von p mit $V_1 \subseteq V$. Dann ist auch $(V_1, \tau|_{V_1})$ eine Einbettungskarte für M bei p .

(c) Aus (b) folgt: Eine Menge $M \subseteq \mathbb{R}^N$ ist eine k -dimensionale Untermannigfaltigkeit genau dann, wenn M lokal jeweils eine k -dimensionale Untermannigfaltigkeit ist, d.h., wenn es zu jedem Punkt $p \in M$ eine offene Umgebung V von p in \mathbb{R}^N gibt, so dass $M \cap V$ eine k -dimensionale Untermannigfaltigkeit ist. Um dies zu sehen, bildet man für eine Einbettungskarte (V_1, τ) für M jeweils die Einbettungskarte $(V \cap V_1, \tau|_{V \cap V_1})$ für $M \cap V$.

Beispiel 7.34. Seien $0 \leq k \leq N$, $U \subseteq \mathbb{R}^k$ offen und $h \in C^1(U, \mathbb{R}^{N-k})$. Dann ist

$$M := \text{Graph } h = \{(x, h(x)) \mid x \in U\}$$

eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^N .

Beweis. Seien $m := N - k$, $V := U \times \mathbb{R}^m$ und $\sigma \in C^1(V, V)$ definiert durch $\sigma(x, y) := (x, h(x) + y)$. Dann gilt für $(x, y) \in V$

$$J_\sigma(x, y) = \begin{pmatrix} E_k & 0 \\ J_h(x) & E_m \end{pmatrix}.$$

Für $v := (v_1, v_2) \in \mathbb{R}^k \times \mathbb{R}^m$ folgt aus $J_\sigma(x, y)v = 0$ also $v_1 = 0$ und $J_h(x)v_1 + v_2 = 0$, also auch $v_2 = 0$. Daher ist $d\sigma(x, y)$ injektiv, also ein linearer Isomorphismus, und σ ist nach dem Satz über die Umkehrfunktion lokal diffeomorph in allen Punkten. Wegen der offensichtlichen Bijektivität ist σ also ein Diffeomorphismus mit $\text{Bild}(\sigma) = V$, und es gilt, mit $\tau := \sigma^{-1}$, $M \cap V = M = \sigma(\mathbb{R}^k \times \{0\}) = \tau^{-1}(\mathbb{R}^k \times \{0\})$. Demnach ist (V, τ) eine Einbettungskarte für jeden Punkt in M . \square

Satz 7.35 (Satz vom regulären Wert). *Seien $D \subseteq \mathbb{R}^N$ offen, $g \in C^1(D, \mathbb{R}^m)$, mit $m \leq N$, und $w \in \mathbb{R}^m$ ein regulärer Wert von g . Dann ist $M := g^{-1}(w)$ eine $(N - m)$ -dimensionale Untermannigfaltigkeit des \mathbb{R}^N .*

Beweis. Setze $k := N - m$ und sei $p \in M$. Nach dem Satz über implizite Funktionen existiert eine offene Teilmenge $U \subseteq \mathbb{R}^k$, eine offene Umgebung $V \subseteq D$ von p sowie Abbildungen $h \in C^1(U, \mathbb{R}^m)$ und $T \in \text{GL}(N, \mathbb{R})$, mit

$$M \cap V = T(\text{Graph } h).$$

Nach Beispiel 7.34 ist $\text{Graph } h$ eine k -dimensionale Untermannigfaltigkeit von \mathbb{R}^N . Weil T insbesondere ein Diffeomorphismus ist, folgt aus Bemerkung 7.33(a), dass dann auch $M \cap V$ eine k -dimensionale Untermannigfaltigkeit von \mathbb{R}^N ist. Da $p \in M$ beliebig war, folgt die Behauptung aus Bemerkung 7.33(c). \square

Beispiel 7.36. Wir betrachten $N \geq 2$, $g: \mathbb{R}^N \rightarrow \mathbb{R}$, $g(x) := |x|^2$ und die sogenannte *Einheitssphäre*

$$S^{N-1} := g^{-1}(1) = \{x \in \mathbb{R}^N \mid |x| = 1\} = \partial \bar{B}_1(0) \subseteq \mathbb{R}^N.$$

Da $\nabla g(x) = 2x \neq 0$ für alle $x \in S^{N-1}$ gilt, ist 1 ein regulärer Wert von g . Somit ist S^{N-1} nach Satz 7.35 eine $(N - 1)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^N .

Bemerkung 7.37. Seien M eine k -dimensionale Untermannigfaltigkeit von \mathbb{R}^N , $p \in M$, (V, τ) eine Einbettungskarte für M bei p und $P \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^{N-k})$ die Projektion auf die letzten $N - k$ Koordinaten (hier weichen wir etwas von der bisherigen Definition von Projektionen ab). Dann ist $g := P \circ \tau$ eine C^1 -Funktion, und es gilt $M \cap V = g^{-1}(0)$. Ferner ist 0 regulär für g , denn es gilt $dg(z) = Pd\tau(z)$, P ist surjektiv, und $d\tau(z)$ ist für jedes $z \in V$ ein linearer Isomorphismus. M lässt sich also lokal als Urbild eines regulären Wertes darstellen. Wie im Beweis von Satz 7.35 liefert dies, zusammen mit dem Satz über implizite Funktionen, dass M sich lokal als das linear isomorphe Bild eines Graphen darstellen lässt.

Für einen Vektorraum V und eine Teilmenge $A \subseteq V$ definieren wir den *von A aufgespannten Untervektorraum* oder auch die *lineare Hülle von A* durch

$$\text{Spann}(A) := \bigcap_{\substack{W \text{ Untervektorraum von } V \\ A \subseteq W}} W.$$

Offensichtlich ist $\text{Spann}(A)$ ein Untervektorraum von V .

Definition und Satz 7.38. Sei $V \subseteq \mathbb{R}^N$ ein Untervektorraum. Wir setzen: $V^\perp := \{w \in \mathbb{R}^N \mid \langle w, v \rangle = 0 \text{ für alle } v \in V\}$ (*orthogonales Komplement von V*). Dann gilt: V^\perp

ist ein Untervektorraum von \mathbb{R}^N mit

$$(7.22) \quad \mathbb{R}^N = V \oplus V^\perp \quad \text{und} \quad (V^\perp)^\perp = V.$$

Ist $A \subseteq \mathbb{R}^N$ eine beliebige Teilmenge, dann setzen wir $A^\perp := \text{Spann}(A)^\perp$.

Beweis. Sei P die Orthogonalprojektion auf V gemäß Aufgabe 17 von Blatt 5. Es gilt offensichtlich $V^\perp = \text{Kern } P$ und daher $\mathbb{R}^N = V \oplus V^\perp$. Außerdem ist $Q := I - P$ die Orthogonalprojektion auf V^\perp mit $V = \text{Kern } Q = (V^\perp)^\perp$. \square

Definition 7.39. Seien $M \subseteq \mathbb{R}^N$ eine k -dimensionale Untermannigfaltigkeit und $p \in M$. Dann setzen wir:

- (a) $T_p M := \{v \in \mathbb{R}^N \mid v \text{ Tangentialvektor von } M \text{ in } p\}$ (*Tangentenraum von } M \text{ in } p*),
- (b) $N_p M := (T_p M)^\perp$ (*Normalraum von } M \text{ in } p*).

Bemerkung 7.40. Die Menge $T_p M$ ist wegen Definition 7.28 lokal definiert. Um sie zu untersuchen, können wir uns also auf $M \cap V$ für eine beliebige offene Umgebung von p in \mathbb{R}^N beschränken.

Satz 7.41. Sei M eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^N und $p \in M$. Seien ferner $m := N - k$, $V \subseteq \mathbb{R}^N$ eine offene Umgebung von p , $g \in C^1(V, \mathbb{R}^m)$ und $w \in \mathbb{R}^m$ ein regulärer Wert von g , so dass $M \cap V = g^{-1}(w)$ gilt. Solch V , g und w existieren nach Bemerkung 7.37. Dann gelten:

- (a) $T_p M = \text{Kern } dg(p) := \{v \in \mathbb{R}^N \mid dg(p)v = 0\}$,
- (b) $N_p M = \text{Spann}\{\nabla g_1(p), \dots, \nabla g_m(p)\}$.

Es folgt, dass $T_p M$ ein k -dimensionaler und $N_p M$ ein m -dimensionaler Untervektorraum von \mathbb{R}^N ist.

Beweis. (a): Weil p ein regulärer Punkt von g ist, folgt aus Dimensionsgründen $\dim \text{Kern } dg(p) = k$. Wir können nach Anwendung affiner Diffeomorphismen (also Abbildungen $x \mapsto Tx + c$ mit $T \in \text{GL}(N, \mathbb{R})$ und $c \in \mathbb{R}^N$) annehmen, dass $p = 0$, $w = 0$ und $\text{Kern } dg(0) = \mathbb{R}^k \times \{0\}$ gelten (Übung!).

Sei $v \in T_0 M$. Dann existieren $\varepsilon > 0$ und eine C^1 -Kurve $\gamma: (-\varepsilon, \varepsilon) \rightarrow M \cap V$ mit $\gamma(0) = 0$ und $\gamma'(0) = v$. Es folgt $g \circ \gamma \equiv 0$ und somit $0 = (g \circ \gamma)'(0) = dg(0)v$, also $v \in \text{Kern } dg(0)$.

Sei umgekehrt $v \in \text{Kern } dg(0)$. Wir schreiben wieder (x, y) für die Punkte in $\mathbb{R}^k \times \mathbb{R}^m$ und entsprechend $v = (v_1, 0)$ mit $v_1 \in \mathbb{R}^k$. Die Abbildung

$$\partial_y g(0) = dg(0)|_{\{0\} \times \mathbb{R}^m} \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$$

ist ein linearer Isomorphismus, weil $dg(0)$ surjektiv ist und

$$\partial_x g(0) = dg(0)|_{\mathbb{R}^k \times \{0\}} = dg(0)|_{\text{Kern } dg(0)} = 0$$

gilt. Nach dem Satz über implizite Funktionen existieren dann offene Nullumgebungen $U_1 \subseteq \mathbb{R}^k$, $U_2 \subseteq \mathbb{R}^m$ und eine C^1 -Funktion $h: U_1 \rightarrow U_2$, so dass für $U := U_1 \times U_2$ gilt: $U \subseteq V$, $M \cap U = \text{Graph } h$, $h(0) = 0$ und

$$J_h(0) = -(\partial_y g(0))^{-1} \partial_x g(0) = 0.$$

Nun sei $\varepsilon > 0$ so klein gewählt, dass $tv_1 \in U_1$ für alle $t \in (-\varepsilon, \varepsilon)$ gilt. Wir definieren eine C^1 -Kurve $\gamma: (-\varepsilon, \varepsilon) \rightarrow M$ durch $\gamma(t) := (tv_1, h(tv_1))$. Es folgt $\gamma(0) = 0$ und $\gamma'(0) = (v_1, dh(0)v_1) = (v_1, 0) = v$. Demnach ist $v \in T_0M$.

(b): Für $v \in \mathbb{R}^N$ gilt:

$$v \in \text{Kern } dg(p) \Leftrightarrow J_g(p)v = 0 \Leftrightarrow \langle \nabla g_i(p), v \rangle = 0 \text{ für } i = 1, \dots, m.$$

Also ist

$$\text{Kern } dg(p) = \{\nabla g_1(p), \dots, \nabla g_m(p)\}^\perp = (\text{Spann}\{\nabla g_1(p), \dots, \nabla g_m(p)\})^\perp$$

und somit

$$N_pM = (T_pM)^\perp \stackrel{(a)}{=} (\text{Kern } dg(p))^\perp \stackrel{(7.22)}{=} \text{Spann}\{\nabla g_1(p), \dots, \nabla g_m(p)\}. \quad \square$$

Wir führen eine weitere Möglichkeit ein, Untermannigfaltigkeiten darzustellen:

Definition 7.42. Seien $0 \leq k \leq N$, $M \subseteq \mathbb{R}^N$ und $p \in M$. Seien weiterhin (U, V, ψ) ein Tripel aus offenen Mengen $U \subseteq \mathbb{R}^k$ und $V \subseteq \mathbb{R}^N$ und einer Funktion $\psi \in C^1(U, \mathbb{R}^N)$ mit folgenden Eigenschaften:

- (i) $p \in V$,
- (ii) ψ bildet U bijektiv auf $M \cap V$ ab,
- (iii) $\psi^{-1}: M \cap V \rightarrow U$ ist stetig und
- (iv) $d\psi(x) \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^N)$ ist injektiv für alle $x \in U$.

Dann heißt (U, V, ψ) *lokale k -dimensionale Parametrisierung von M bei p* und M heißt *in p lokal k -dimensional parametrisierbar*. Ist M in allen Punkten lokal k -dimensional parametrisierbar, dann nennen wir M *lokal k -dimensional parametrisierbar*. Falls eine lokale k -dimensionale Parametrisierung mit $M \cap V = M$ existiert, dann heißt diese *globale k -dimensionale Parametrisierung von M* und M heißt *global k -dimensional parametrisierbar*.

Bemerkung 7.43. (a) Mit einer k -dimensionalen lokalen Parametrisierung von M bei p übertragen wir gewissermaßen einen Ausschnitt des Koordinatensystems von \mathbb{R}^k auf eine Umgebung von p in M .

- (b) Eine Einbettungskarte (V, τ) für eine k -dimensionale Untermannigfaltigkeit M des \mathbb{R}^N liefert mit

$$U := \tau(V) \cap (\mathbb{R}^k \times \{0_{N-k}\})$$

und $\psi := \tau^{-1}|_U$ die lokale Parametrisierung (U, V, ψ) . Jede k -dimensionale Untermannigfaltigkeit des \mathbb{R}^N ist also lokal k -dimensional parametrisierbar.

- (c) In Beispiel 7.34 ist (U, \mathbb{R}^N, ψ) mit $\psi(x) := (x, h(x))$ eine globale Parametrisierung von $M = \text{Graph } h$: Sei dazu $V := \mathbb{R}^N$ und sei $P \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^k)$ die Projektion auf die ersten k Koordinaten. Dann gilt $P \circ \psi = \text{id}_U$ und $\psi \circ (P|_M) = \text{id}_M$, d.h. ψ bildet U bijektiv auf M ab. Ferner ist $\psi \in C^1(U, \mathbb{R}^N)$ und $\psi^{-1} = P|_M$ stetig. Aus $d\psi(x)v = (v, dh(x)v)$ für $x \in U$ und $v \in \mathbb{R}^k$ folgt $\text{Kern } d\psi(x) = \{0\}$, d.h. $d\psi(x)$ ist injektiv.
- (d) Eine nichtleere **kompakte** k -dimensionale Untermannigfaltigkeit M des \mathbb{R}^N besitzt keine globale Parametrisierung. Andernfalls gäbe es nämlich eine nichtleere offene Teilmenge $U \subseteq \mathbb{R}^k$ und einen Homöomorphismus $\psi: U \rightarrow M$. Es würde folgen, dass U kompakt ist, ein Widerspruch.

Beispiel 7.44. Es gibt viele Möglichkeiten, S^{N-1} lokal zu parametrisieren, z.B. Graphen über offenen Einheitskugeln in $(N-1)$ -dimensionalen Unterräumen. Dafür bräuchte man allerdings mindestens $N+1$ lokale Parametrisierungen. Die folgende *stereographische Projektion* liefert eine Überdeckung von S^{N-1} durch lediglich zwei lokale Parametrisierungen: Seien dazu $U := \mathbb{R}^{N-1}$ und $V_{\pm} = \mathbb{R}^N \setminus \{\pm e_N\}$, wobei $e_N := (0, \dots, 0, 1)$ der N -te Einheitsvektor sei. Dann ist $S^{N-1} = (S^{N-1} \cap V_+) \cup (S^{N-1} \cap V_-)$. Wir definieren

$$\psi, \varphi: U \rightarrow \mathbb{R}^N, \quad \begin{cases} \psi(x) = \frac{1}{1+|x|^2} \begin{pmatrix} 2x \\ |x|^2 - 1 \end{pmatrix}, \\ \varphi(x) = \frac{1}{1+|x|^2} \begin{pmatrix} 2x \\ 1 - |x|^2 \end{pmatrix}. \end{cases}$$

Dann gilt:

- Ist $p \in S^{N-1} \setminus \{e_N\}$, so ist $\psi: U \rightarrow S^{N-1} \cap V_+$ eine lokale Parametrisierung bei p
- Ist $p \in S^{N-1} \setminus \{-e_N\}$, so ist $\varphi: U \rightarrow S^{N-1} \cap V_-$ eine lokale Parametrisierung bei p

Beweis: Übung.

Satz 7.45. Seien $0 \leq k \leq N$ und $U \subseteq \mathbb{R}^k$ offen. Ferner sei $\psi \in C^1(U, \mathbb{R}^N)$ injektiv, $d\psi(x)$ sei injektiv für jedes $x \in U$, und für $M := \text{Bild } \psi$ sei $\psi^{-1}: M \rightarrow U$ stetig. Dann ist M eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^N , ψ ist eine globale k -dimensionale Parametrisierung von M und es gilt $T_p M = \text{Bild } d\psi(x)$ für jedes $x \in U$ und $p = \psi(x)$. Ist ferner $(\tilde{U}, \tilde{V}, \varphi)$ eine weitere lokale k -dimensionale Parametrisierung von M bei einem Punkt $p \in M$, dann ist für $W := \tilde{V} \cap M$ die Parameterwechselabbildung

$$\psi^{-1} \circ \varphi: \varphi^{-1}(W) \rightarrow \psi^{-1}(W)$$

ein Diffeomorphismus.

Beweis. Sei $m := N - k$. Wie im Beweis von Satz 7.41 nehmen wir an, dass $p = 0$, $\psi(0) = 0$ und $\text{Bild } d\psi(0) = \mathbb{R}^k \times \{0_m\}$ gelten. Dies können wir tun, weil $\dim(\text{Bild } d\psi(0)) = k$ wegen der Injektivität von $d\psi(0)$ gilt.

Seien $P: \mathbb{R}^N \rightarrow \mathbb{R}^k$ die Projektion auf die ersten k Koordinaten und $Q: \mathbb{R}^N \rightarrow \mathbb{R}^m$ die Projektion auf die letzten m Koordinaten. Betrachte die Funktion $\sigma: U \rightarrow \mathbb{R}^k$ mit $\sigma := P\psi$. Dann ist $d\sigma(0) = Pd\psi(0) \in \mathcal{L}(\mathbb{R}^k)$ surjektiv, also ein linearer Isomorphismus. Demnach existiert nach dem Satz über die Umkehrfunktion eine offene Teilmenge $U_1 \subseteq U$, so dass $\sigma|_{U_1}$ ein Diffeomorphismus ist. Wir setzen $\mu := Q\psi \circ \sigma^{-1} \in C^1(\sigma(U_1), \mathbb{R}^m)$. Es folgt für alle $x \in U_1$:

$$\psi(x) = (P\psi(x), Q\psi(x)) = (\sigma(x), \mu(\sigma(x))),$$

d.h., $\psi(U_1) = \text{Graph } \mu$. Die Stetigkeit von ψ^{-1} und Satz 4.35 liefern weiter, dass $\psi(U_1) = (\psi^{-1})^{-1}(U_1)$ in M relativ offen ist. Nach Satz und Definition 4.21(a) gibt es eine offene Menge $V \subseteq \mathbb{R}^N$, so dass $M \cap V = \psi(U_1) = \text{Graph } \mu$ gilt. Genauso wie im Beweis von Satz 7.41(a) erhält man hieraus $T_0M = \mathbb{R}^k \times \{0\} = \text{Bild } d\psi(0)$.

Da $p \in M$ beliebig war, haben wir also gezeigt, dass M lokal jeweils ein affin diffeomorphes Bild des Graphen einer C^1 -Funktion auf einer offenen Teilmenge von \mathbb{R}^k mit Bild in \mathbb{R}^m ist. Mit Beispiel 7.34 und Bemerkung 7.33(c) folgt, dass M eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^N ist.

Wir setzen $\rho := \psi^{-1} \circ \varphi$. Sei $q \in W$. Wir können nach Anwendung affin diffeomorpher Transformationen annehmen, dass $0 \in U \cap \tilde{U}$, $q = \varphi(0) = \psi(0)$ und $\text{Bild}(d\varphi(0)) = \text{Bild}(d\psi(0)) = \mathbb{R}^k \times \{0_m\}$ gelten. Wie oben finden wir eine offene Nullumgebung $U' \subseteq U \cap \tilde{U}$, Diffeomorphismen $\sigma_i: U' \rightarrow \mathbb{R}^k$ und Abbildungen $\mu_i \in C^1(\sigma_i(U'), \mathbb{R}^m)$ mit $\varphi(x) = (\sigma_1(x), \mu_1(\sigma_1(x)))$ und $\psi(x) = (\sigma_2(x), \mu_2(\sigma_2(x)))$ für alle $x \in U'$. Es folgt $\psi^{-1} = \sigma_2^{-1} \circ P$ und daher $\rho = \sigma_2^{-1} \circ \sigma_1$ in einer offenen Nullumgebung. Da $q \in W$ beliebig war, ist also ρ lokal ein Diffeomorphismus. Die Bijektivität von $\rho: \varphi^{-1}(W) \rightarrow \psi^{-1}(W)$ liefert nun die Behauptung. \square

Beispiel 7.46. Die Stetigkeit von ψ^{-1} in Satz 7.45 kann nicht weggelassen werden: Betrachte $U := (-2 - \pi/2, \pi + 1)$ und $\psi: U \rightarrow \mathbb{R}^2$ mit

$$\psi(t) := \begin{cases} \left(t + 1 + \frac{\pi}{2}, 0\right), & t \in \left(-2 - \frac{\pi}{2}, -\frac{\pi}{2}\right), \\ (1 + \cos t, 1 + \sin t), & t \in \left[-\frac{\pi}{2}, \pi\right] \\ (0, \pi + 1 - t), & t \in (\pi, \pi + 1). \end{cases}$$

Dann ist ψ eine injektive C^1 -Kurve mit $\psi'(t) \neq (0, 0)$ für alle $t \in U$. Die Umkehrung ψ^{-1} ist auf dem Bild von ψ jedoch nicht stetig, weil $(0, 0) = \psi(-1 - \pi/2) = \lim_{t \rightarrow (\pi+1)^-} \psi(t)$ gilt, siehe Abb. 7.2. Tatsächlich kann man zeigen, dass bei $(0, 0)$ keine Einbettungskarte für $\text{Bild } \psi$ existiert (anschaulich ist das klar).

Korollar 7.47. Seien $0 \leq k \leq N$ und $M \subseteq \mathbb{R}^N$ lokal k -dimensional parametrisierbar. Dann ist M eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^N .

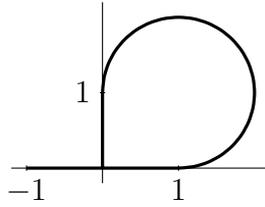


Abbildung 7.2: Bild ψ ist keine Untermannigfaltigkeit.

Beweis. Folgt sofort aus Satz 7.45 und Bemerkung 7.33(c). □

Wir fassen noch einmal alle Möglichkeiten zusammen, k -dimensionale Untermannigfaltigkeiten zu charakterisieren:

Satz 7.48. Sei $M \subseteq \mathbb{R}^N$. Dann sind die folgenden Aussagen äquivalent:

- (i) M ist eine k -dimensionale Untermannigfaltigkeit von \mathbb{R}^N ;
- (ii) M lässt sich lokal als Urbild eines regulären Wertes einer C^1 -Funktion $U \rightarrow \mathbb{R}^{N-k}$ darstellen, wo $U \subseteq \mathbb{R}^N$ offen ist;
- (iii) M lässt sich lokal als affin diffeomorphes Bild des Graphen einer C^1 -Funktion $U \rightarrow \mathbb{R}^{N-k}$ darstellen, wo $U \subseteq \mathbb{R}^k$ offen ist;
- (iv) M ist lokal k -dimensional parametrisierbar.

Beweis. „(i) \Rightarrow (ii)“ und „(ii) \Rightarrow (iii)“: Bemerkung 7.37.

„(iii) \Rightarrow (iv)“: Bemerkung 7.43(c).

„(iv) \Rightarrow (i)“: Korollar 7.47. □

Ende der Sitzung vom 13.7.

Satz 7.49 (Lagrange-Multiplikatoren). Seien $D \subseteq \mathbb{R}^N$ offen, $f \in C^1(D)$, $g \in C^1(D, \mathbb{R}^m)$ mit $m \leq N$ und $M = g^{-1}(c)$ für einen regulären Wert $c \in g(D)$ von g . Dann gilt für $p \in M$: p ist kritischer Punkt von $f|_M$ genau dann, wenn $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ existieren mit

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \dots + \lambda_m \nabla g_m(p).$$

Dies trifft insbesondere dann zu, wenn p ein lokales Extremum von $f|_M$ ist. Die Zahlen $\lambda_1, \dots, \lambda_m$ heißen Lagrange-Multiplikatoren.

Beweis. Nach Definition 7.39 und Satz 7.41 gilt:

$$p \text{ ist kritischer Punkt von } f|_M \Leftrightarrow \nabla f(p) \in N_p M = \text{Spann}\{\nabla g_1(p), \dots, \nabla g_m(p)\}.$$

Dies liefert die Behauptung. □

Beispiel 7.50 (Fortsetzung von Bemerkung und Beispiel 7.27). Seien $D = \mathbb{R}^3$, $q = (0, 0, 1) \in D$ und seien $f, g \in C^1(\mathbb{R}^3)$ definiert durch

$$f(x, y, z) = \left\| \begin{pmatrix} x \\ y \\ z \end{pmatrix} - q \right\|_2^2 = x^2 + y^2 + (z - 1)^2 \quad \text{und} \quad g(x, y, z) = x^2 + \frac{y^2}{4} + \frac{z^2}{9}.$$

Dann ist $M := g^{-1}(1) \subseteq \mathbb{R}^3$ der Ellipsoid wie in Bemerkung und Beispiel 7.27 definiert. Es ist

$$\nabla g(x, y, z) = \begin{pmatrix} 2x \\ \frac{1}{2}y \\ \frac{2}{9}z \end{pmatrix} \neq 0 \quad \text{für } (x, y, z) \neq 0,$$

also insbesondere für $(x, y, z) \in M$. Also folgt, dass 1 ein regulärer Wert von g ist. Wie schon in Bemerkung und Beispiel 7.27 bemerkt, existiert $p = (x, y, z) \in M$ mit $|p - q| = \text{dist}(q, M)$, d.h. p ist ein globales Minimum von $f|_M$. Im Folgenden bestimmen wir p und somit auch $\text{dist}(q, M)$ für den Punkt $q = (0, 0, 1)$.

Nach Satz 7.49 existiert ein Lagrange-Multiplikator $\lambda \in \mathbb{R}$ mit $\nabla f(p) = \lambda \nabla g(p)$, d.h.

$$\begin{pmatrix} 2x \\ 2y \\ 2(z - 1) \end{pmatrix} = \lambda \begin{pmatrix} 2x \\ \frac{1}{2}y \\ \frac{2}{9}z \end{pmatrix} \Leftrightarrow \begin{array}{l} (1 - \lambda)x = 0 \quad \text{(I)} \\ (4 - \lambda)y = 0 \quad \text{(II)} \\ (9 - \lambda)z = 9 \quad \text{(III)} \end{array}$$

1. Fall: Ist $x \neq 0$ dann folgt mit Gleichung (I) $\lambda = 1$ und dann mit (II), (III) $y = 0$, $z = \frac{9}{8}$. Da

$$1 = g(x, y, z) = g\left(x, 0, \frac{9}{8}\right) = x^2 + \frac{9}{64}$$

ist, folgt $x = \pm \sqrt{1 - \frac{9}{64}} = \pm \sqrt{\frac{55}{64}}$ und damit ist

$$f(x, y, z) = \frac{55}{64} + \frac{1}{64} = \frac{7}{8}.$$

2. Fall: Ist $y \neq 0$ dann folgt mit Gleichung (II) $\lambda = 4$ und dann mit (I), (III) $x = 0$ und $z = \frac{9}{5}$. Da

$$1 = g(x, y, z) = g\left(0, y, \frac{9}{5}\right) = \frac{y^2}{4} + \frac{9}{25}$$

ist, folgt $y = \pm 2\sqrt{1 - \frac{9}{25}} = \pm \frac{8}{5}$, und damit ist

$$f(x, y, z) = 4\left(1 - \frac{9}{25}\right) + \frac{16}{25} = 4 + \frac{-36 + 16}{25} = \frac{16}{5}.$$

3. Fall: Ist $x = 0 = y$ dann folgt mit der Bedingung $g(x, y, z) = 1$, dass $1 = \frac{z^2}{9}$, also

$z = \pm 3$. Es folgt

$$f(0, 0, z) = (z - 1)^2 = \begin{cases} 4, & \text{falls } z = 3 \\ 16, & \text{falls } z = -3. \end{cases}$$

Folglich sind alle kritischen Punkte von $f|_M$ gegeben durch $(\pm\sqrt{\frac{55}{64}}, 0, \frac{9}{8})$, $(0, \pm\frac{8}{5}, \frac{9}{5})$ und $(0, 0, \pm 3)$. Es folgt, dass das gesuchte globale Minimum von $f|_M$ angenommen wird bei $p = (\sqrt{\frac{55}{64}}, 0, \frac{9}{8})$ und $\tilde{p} = (-\sqrt{\frac{55}{64}}, 0, \frac{9}{8})$. Es folgt

$$\text{dist}(q, M) = \sqrt{|q - p|^2} = \sqrt{f(p)} = \sqrt{\frac{7}{8}}.$$

Proposition 7.51. *Seien $N \geq 2$, $A \in \mathbb{R}^{N \times N}$ symmetrisch und $f \in C^2(\mathbb{R}^N)$ definiert durch $f(x) = \langle Ax, x \rangle$. Dann gilt für $x \in S^{N-1}$: x ist ein kritischer Punkt von $f|_{S^{N-1}}$ genau dann, wenn x ein Eigenvektor von A ist. Ist dies erfüllt, so ist $f(x)$ der zugehörige Eigenwert.*

Beweis. Zunächst gilt für $x, h \in \mathbb{R}^N$ wegen der Symmetrie von A :

$$\begin{aligned} f(x+h) &= \langle A(x+h), x+h \rangle = \langle Ax, x \rangle + \langle Ah, x \rangle + \langle Ax, h \rangle + \langle Ah, h \rangle \\ &= f(x) + 2\langle Ax, h \rangle + O(|h|^2) \end{aligned}$$

für $h \rightarrow 0$. Daraus lesen wir ab, dass f differenzierbar ist mit $\nabla f(x) = 2Ax$.

Wie in Beispiel 7.36 ist $S^{N-1} = g^{-1}(1)$, für $g \in C^1(\mathbb{R}^N)$ definiert durch $g(x) = |x|^2$ und den regulären Wert 1 von g . Nach Satz 7.49 gilt für $x \in S^{N-1}$: x ist ein kritischer Punkt von $f|_{S^{N-1}}$ genau dann, wenn es $\lambda \in \mathbb{R}$ mit $\nabla f(x) = \lambda \nabla g(x)$ gibt, also wenn $Ax = \lambda x$ gilt. In diesem Fall ist $f(x) = \langle Ax, x \rangle = \langle \lambda x, x \rangle = \lambda$. Es folgt die Behauptung. \square

Satz 7.52. *Seien $D \subseteq \mathbb{R}^N$ offen, $f \in C^1(D)$, $M \subseteq D$ eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^N und $p \in M$. Sei ferner (U, V, ψ) eine lokale k -dimensionale Parametrisierung von M bei p und sei $q := \psi^{-1}(p)$. Es folgt, dass p genau dann ein kritischer Punkt von $f|_M$ ist, wenn q ein kritischer Punkt von $f \circ \psi$ ist.*

Beweis. Für $v \in \mathbb{R}^k$ gilt

$$(7.23) \quad \langle \nabla(f \circ \psi)(q), v \rangle = d(f \circ \psi)(q)v = df(p)d\psi(q)v = \langle \nabla f(p), d\psi(q)v \rangle.$$

Sei q ein kritischer Punkt von $f \circ \psi$, d.h. $\nabla(f \circ \psi)(q) = 0$. Da $d\psi(q): \mathbb{R}^k \rightarrow T_p M$ ein linearer Isomorphismus ist, hat jeder Tangentialvektor in p eine Darstellung $d\psi(q)v$ mit $v \in \mathbb{R}^k$. Wegen (7.23) ist dann $\nabla f(p) \in N_p M$, also p ein kritischer Punkt von $f|_M$. Sei nun umgekehrt p ein kritischer Punkt von $f|_M$, also $\nabla f(p) \in N_p M$. Dann gilt $\langle \nabla f(p), d\psi(q)v \rangle = 0$ für jedes $v \in \mathbb{R}^k$. Insbesondere folgt für $v := \nabla(f \circ \psi)(q)$ aus (7.23)

$$|\nabla(f \circ \psi)(q)|^2 = \langle \nabla(f \circ \psi)(q), \nabla(f \circ \psi)(q) \rangle = \langle \nabla f(p), d\psi(q)\nabla(f \circ \psi)(q) \rangle = 0,$$

also $\nabla(f \circ \psi)(q) = 0$. Mit anderen Worten ist q ein kritischer Punkt von $f \circ \psi$. \square

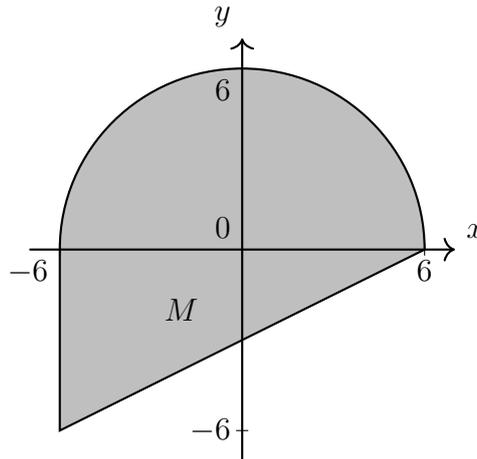


Abbildung 7.3: Nebenbedingung für Extrema

Dieser Satz liefert ein Kriterium, kritische Punkte der Einschränkung einer Funktion auf eine Untermannigfaltigkeit in dem Fall zu finden, wo explizite Parametrisierungen bekannt sind.

Im Folgenden bedeute $\cup, \dot{\cup}$ die Vereinigung von disjunkten Mengen.

Bemerkung 7.53. Häufig möchte man die Extremwerte einer Funktion f auf einer Menge $M = \dot{\cup}_{k=1}^n M_k$ finden, wo die Mengen M_k Untermannigfaltigkeiten mit u.U. verschiedenen Dimensionen sind. Jede lokale Extremalstelle $x \in M$ von $f|_M$ ist aber auch eine lokale Extremalstelle von $f|_{M_k}$, wo $x \in M_k$ ist. Also sucht man kritische Punkte von $f|_{M_k}$ für alle $k = 1, \dots, n$. Man beachte dabei: Die N -dimensionalen Untermannigfaltigkeiten von \mathbb{R}^N sind genau die offenen Teilmengen von \mathbb{R}^N , und die 0-dimensionalen Untermannigfaltigkeiten von \mathbb{R}^N sind genau die Teilmengen, die nur aus isolierten Punkten bestehen.

Beispiel 7.54. Wir bestimmen Minimum und Maximum der Funktion $f(x, y) := x^2 - xy + y^2$ auf der kompakten Menge M , die in Abb. 7.3 als graue Region dargestellt ist. Dazu betrachten wir die Zerlegung $M = \dot{\cup}_{k=0}^4 M_k$ mit

$$\begin{aligned}
 M_0 &:= \{(6, 0), (-6, 0), (-6, -6)\}, \\
 M_1 &:= \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 36, y > 0\}, \\
 M_2 &:= \{(-6, t) \mid t \in (-6, 0)\}, \\
 M_3 &:= \{(t, t/2 - 3) \mid t \in (-6, 6)\}, \\
 M_4 &:= M \setminus \left(\bigcup_{k=0}^3 M_k \right).
 \end{aligned}$$

Auf M_0 nimmt f nur den Wert 36 an.

M_1 : Sei $D := \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$. Wir haben

$$f(z) = \langle Az, z \rangle \quad \text{für } z = (x, y)$$

mit $A := \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix}$. Wegen $f(z) = |z|^2 f(z/|z|) = 36f(z/6)$ für $z \in M_1$ reicht es, die kritischen Punkte von $f|_{S^1 \cap D}$ zu bestimmen und die zugehörigen Werte mit 36 zu multiplizieren. Kritische Punkte von f_{S^1} entsprechen nach Proposition 7.51 den Eigenvektoren von A , und Werte von f sind die zugehörigen Eigenwerte von A . Berechnung der Eigenwerte von A liefert $\lambda = 1/2$ und $\lambda = 3/2$, mit entsprechenden Eigenvektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ in $S^1 \cap D$. Dies sind alle möglichen kritischen Punkte von $f|_{S^1 \cap D}$, und die entsprechen Werte von $f|_{M_1}$ sind $36 \cdot (1/2) = 18$ und $36 \cdot (3/2) = 54$.

M_2 : Wir verwenden die globale Parametrisierung $\psi: (-6, 0) \rightarrow \mathbb{R}^2$, $\psi(t) := \begin{pmatrix} -6 \\ t \end{pmatrix}$ und suchen die kritischen Punkte von $g := f \circ \psi$ gemäß Satz 7.52. Aus $g(t) = 36 + 6t + t^2$ folgt $g'(t) = 6 + 2t$. g hat also nur den kritischen Punkt $t = -3 \in (-6, 0)$, dies entspricht $\psi(-3) = \begin{pmatrix} -6 \\ -3 \end{pmatrix}$ als einzigem kritischen Punkt von $f|_{M_2}$ mit dem Wert $f(-6, -3) = 27$.

M_3 : Ähnlich wie eben benutzen wir die Parametrisierung $t \mapsto \begin{pmatrix} t \\ t/2 - 3 \end{pmatrix}$ auf $(-6, 6)$ und betrachten $g(t) := f(t, t/2 - 3) = 3t^2/4 + 9$. Mit $g'(t) = 3t/2$ erhalten wir $t = 0 \in (-6, 6)$ als einzigen kritischen Punkt von g , bzw. $(0, -3)$ als einzigen kritischen Punkt von $f|_{M_3}$ mit dem Wert 9.

M_4 : Dies ist eine offene Teilmenge von \mathbb{R}^2 . Wir bestimmen die kritischen Punkte von $f|_{M_4}$. Wegen $\nabla f(z) = 2Az$ und da Kern $A = \{0\}$ gilt, ist $(x, y) = (0, 0)$ der einzige kritische Punkt von $f|_{M_4}$, mit Wert 0.

Zusammengefasst gilt also $\min_M f = 0$ und $\max_M f = 54$.

Zum Abschluss stellen wir noch die Matrixgruppe

$$\text{SL}(N) := \{A \in \mathbb{R}^{N \times N} \mid \det(A) = 1\}$$

(*Special Linear Group*) als Untermannigfaltigkeit dar. Diese und andere *Lie-Gruppen* spielen eine wichtige Rolle in der Untersuchung von Symmetrien in der Physik. Man ist daher an den analytischen Eigenschaften interessiert.

Satz 7.55. $\text{SL}(N)$ ist für jedes $N \geq 2$ eine $(N^2 - 1)$ -dimensionale Untermannigfaltigkeit des $\mathbb{R}^{N \times N}$. Der Tangentialraum von $\text{SL}(N)$ im Punkt $E_N \in \text{SL}(N)$ wird gerade durch die $N \times N$ -Matrizen A mit $\text{Spur}(A) = 0$ beschrieben:

$$(7.24) \quad T_{E_N} \text{SL}(N) = \{A \in \mathbb{R}^{N \times N} \mid \text{Spur}(A) = 0\}.$$

Beweis. Wegen $\text{SL}(N) = \det^{-1}(1)$ wollen wir den Satz vom regulären Wert, Satz 7.35, verwenden. Zunächst gilt $\det \in C^1(\mathbb{R}^{N \times N})$, da es sich um eine Polynomfunktion $\mathbb{R}^{N^2} \rightarrow \mathbb{R}$ handelt. Sei nun $(i, j) \in \{1, \dots, N\}^2$. Dann folgt aus der Entwicklung nach der i -ten Zeile,

$$\det(A) = \sum_{k=1}^N (-1)^{i+k} a_{ik} \det(A_{ik}),$$

unmittelbar

$$(7.25) \quad \partial_{(i,j)}(\det)(A) = \frac{\partial}{\partial a_{ij}} \det(A) = (-1)^{i+j} \det(A_{ij}).$$

Angenommen, $\nabla(\det)(A) = 0$ würde für mindestens eine Matrix $A \in \mathrm{SL}(N)$ gelten. Aus (7.25) folgte dann $\det(A_{ij}) = 0$ für alle $(i, j) \in \{1, \dots, N\}^2$. Dies würde $\det(A) = 0$ nach sich ziehen, im Widerspruch zu $\det(A) = 1$. Demnach ist 1 also ein regulärer Wert von $\det(\cdot)$, und $\mathrm{SL}(N)$ eine $N^2 - 1$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^{N^2} .

Um (7.24) zu zeigen, setzen wir $X := \{A \in \mathbb{R}^{N \times N} \mid \mathrm{Spur}(A) = 0\}$.

„ \supseteq “: Sei $A \in X$. Für $\gamma(t) := \exp(tA)$ folgt mit Aufgabe 37 (Blatt 10)

$$\det(\gamma(t)) = e^{\mathrm{Spur}(tA)} = e^{t \mathrm{Spur}(A)} = e^0 = 1.$$

Also ist $\gamma(t) \in \mathrm{SL}(N)$ für alle $t \in \mathbb{R}$. Offensichtlich gilt $\gamma(0) = E_N$ und $\gamma'(0) = A \exp(0 \cdot A) = A$ und damit $A \in T_{E_N} \mathrm{SL}(N)$.

„ \subseteq “: In E_N gilt für die partiellen Ableitungen wegen Gleichung (7.25):

$$\partial_{(i,j)}(\det)(E_N) = (-1)^{i+j} \det(E_{ij}) = \delta_{ij}.$$

Sei $A = (a_{ij})_{1 \leq i, j \leq N} \in T_{E_N} \mathrm{SL}(N)$. Nach Satz 7.41 gilt: $T_{E_N} \mathrm{SL}(N) = \mathrm{Kern} \, d(\det)(E_N)$. Es folgt

$$0 = d(\det)(E_N)A = \sum_{i,j=1}^N \delta_{ij} a_{ij} = \sum_{i=1}^N a_{ii} = \mathrm{Spur}(A),$$

also $A \in X$. □

Ende der Sitzung vom 16.7.

Literaturverzeichnis

- [1] *Helmholtz decomposition*, Wikipedia.
- [2] H. Amann, *Ordinary differential equations*, De Gruyter Studies in Mathematics, vol. 13, Walter de Gruyter & Co., Berlin, 1990, An introduction to nonlinear analysis, Translated from the German by Gerhard Metzen. MR 1071170
- [3] H. Amann and J. Escher, *Analysis. II*, Birkhäuser Verlag, Basel, 2008, Translated from the 1999 German original by Silvio Levy and Matthew Cargo. MR 2419362
- [4] M. Artin, *Algebra*, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Verlag, Basel, 1993, Translated from the 1991 English original by Annette A'Campo. MR 1266624
- [5] M. Barner and F. Flohr, *Analysis. II*, second ed., de Gruyter Lehrbuch. [de Gruyter Textbook], Walter de Gruyter & Co., Berlin, 1989. MR 1091712
- [6] M. Barner and F. Flohr, *Analysis. I*, fourth ed., de Gruyter Lehrbuch. [de Gruyter Textbook], Walter de Gruyter & Co., Berlin, 1991. MR 1137515
- [7] C. Blatter, *Analysis. II*, second ed., Heidelberger Taschenbücher [Heidelberg Paperbacks], vol. 152, Springer-Verlag, Berlin-New York, 1979. MR 592590
- [8] T. Bröcker, *Analysis. II*, Bibliographisches Institut, Mannheim, 1992. MR 1204000
- [9] T. Bröcker, *Lineare Algebra und analytische Geometrie*, Grundstudium Mathematik. [Basic Study of Mathematics], Birkhäuser Verlag, Basel, 2003, Ein Lehrbuch für Physiker und Mathematiker. [A textbook for physicists and mathematicians]. MR 1992380
- [10] V. Drumm and W. Weil, *Lineare Algebra und Analytische Geometrie*, Universität Karlsruhe, 1989, Vorlesungsskript.
- [11] K. Endl and W. Luh, *Analysis. II*, fourth ed., Akademische Verlagsgesellschaft, Wiesbaden, 1978, Eine integrierte Darstellung, Studien-Texte: Mathematik. MR 537036
- [12] G. Fischer, *Lineare Algebra*, 14 ed., Vieweg + Teubner Verlag, 2003.
- [13] G.B. Folland, *Real analysis*, second ed., Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1999, Modern techniques and their applications, A Wiley-Interscience Publication. MR 1681462

- [14] O. Forster, *Analysis 2. Differentialrechnung im \mathbf{R}^n* , Verlag Vieweg, Braunschweig, 1977, Gewöhnliche Differentialgleichungen, Rororo Vieweg: Mathematik Grundkurs, No. 31. MR 0492102
- [15] O. Forster, *Analysis 1*, expanded ed., Grundkurs Mathematik. [Foundational Course in Mathematics], Vieweg + Teubner, Wiesbaden, 2011, Differential- und Integralrechnung einer Veränderlichen. [Differential and integral calculus of one variable]. MR 2840783
- [16] H. Grauert and H.C. Grunau, *Lineare Algebra und analytische Geometrie*, R. Oldenbourg Verlag, Munich, 1999. MR 1725707
- [17] H. Heuser, *Lehrbuch der Analysis. Teil 2*, sixth ed., Mathematische Leitfäden. [Mathematical Textbooks], B. G. Teubner, Stuttgart, 1991. MR 1147182
- [18] H. Heuser, *Gewöhnliche Differentialgleichungen*, fifth ed., Mathematische Leitfäden. [Mathematical Textbooks], B. G. Teubner, Stuttgart, 2006, Einführung in Lehre und Gebrauch. [Introduction to theory and application]. MR 2382060
- [19] H. Heuser, *Lehrbuch der Analysis. Teil 1*, revised ed., Vieweg + Teubner, Wiesbaden, 2009. MR 2502662
- [20] R.A. Horn and C.R. Johnson, *Matrix analysis*, second ed., Cambridge University Press, Cambridge, 2013. MR 2978290
- [21] E. Kamke, *Differentialgleichungen*, B. G. Teubner, Stuttgart, 1977, Lösungsmethoden und Lösungen. I: Gewöhnliche Differentialgleichungen, Neunte Auflage, Mit einem Vorwort von Detlef Kamke. MR 0466672
- [22] M. Köcher, *Lineare Algebra und analytische Geometrie*, Springer-Lehrbuch, Springer-Verlag Berlin Heidelberg, 1997.
- [23] K. Königsberger, *Analysis. 2*, Springer-Lehrbuch. [Springer Textbook], Springer-Verlag, Berlin, 1993, Grundwissen Mathematik. [Basic Knowledge in Mathematics]. MR 1251736
- [24] K. Königsberger, *Analysis. 1*, sixth ed., Springer-Lehrbuch. [Springer Textbook], Springer-Verlag, Berlin, 2004. MR 2374633
- [25] P. Lancaster and M. Tismenetsky, *The theory of matrices*, second ed., Computer Science and Applied Mathematics, Academic Press, Inc., Orlando, FL, 1985. MR 792300
- [26] P. Plaschko and K. Brod, *Höhere mathematische Methoden für Ingenieure und Physiker*, Hochschultext. [University Textbooks], Springer-Verlag, Berlin, 1989. MR 1010041

- [27] W. Rudin, *Principles of mathematical analysis*, third ed., McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976, International Series in Pure and Applied Mathematics. MR 0385023
- [28] W. Rudin, *Analysis*, R. Oldenbourg Verlag, 1998.
- [29] W. Walter, *Gewöhnliche Differentialgleichungen*, fifth ed., Springer-Lehrbuch. [Springer Textbook], Springer-Verlag, Berlin, 1993, Eine Einführung. [An introduction]. MR 1231977
- [30] W. Walter, *Analysis 2*, Springer Berlin Heidelberg, 1995.
- [31] W. Walter, *Ordinary differential equations*, Graduate Texts in Mathematics, vol. 182, Springer-Verlag, New York, 1998, Translated from the sixth German (1996) edition by Russell Thompson, Readings in Mathematics. MR 1629775
- [32] D. Werner, *Einführung in die höhere Analysis*, corrected ed., Springer-Lehrbuch. [Springer Textbook], Springer-Verlag, Berlin, 2009, Topologische Räume, Funktionentheorie, gewöhnliche Differentialgleichungen, Maß- und Integrationstheorie, Funktionalanalysis. [Topological spaces, function theory, ordinary differential equations, measure and integration theory, functional analysis]. MR 2730632