

Analysis II

Rüdiger W. Braun

Wintersemester 2025/26

Inhaltsverzeichnis

I. Differentialrechnung mehrerer Veränderlicher	1
1. Normierte Räume	3
2. Metrische Räume	7
3. Stetige Abbildungen	15
4. Kompakte Mengen	17
5. Vollständige metrische Räume	21
6. Partielle Ableitungen	23
7. Totale Ableitungen	27
8. Die Hesse-Matrix	33
9. Mittelwertsatz und Taylor-Formel	35
10. Eigenwerte	41
11. Extremwerte und kritische Stellen	45
 II. Gewöhnliche Differentialgleichungen	 49
12. Differentialgleichungen erster Ordnung	51
13. Der Satz von Picard-Lindelöf	59
14. Lineare Differentialgleichungen	63
15. Jordansche Normalform	75
16. Lineare Differentialgleichungen mit konstanten Koeffizienten	79
17. Die Eulersche Differentialgleichung	91

Inhaltsverzeichnis

18. Der Banachsche Fixpunktsatz	93
19. Existenz- und Eindeigkeitssätze	97
20. Abhängigkeit der Lösungen von Parametern	105
21. Autonome Differentialgleichungen	109
III. Der Satz über implizite Abbildungen	113
22. Der Umkehrsatz	115
23. Der Satz über implizite Funktionen	119
24. Untermannigfaltigkeiten des $\mathbb{R}^N$	123
25. Tangentialräume und Extrema unter Nebenbedingungen	133
26. Die Weglänge	139

Teil I.

Differentialrechnung mehrerer Veränderlicher

1. Normierte Räume

1.1 Definition. Ein (reeller) *normierter Raum* besteht aus einem $\mathbb{R}$ -Vektorraum V und einer Abbildung $V \rightarrow \mathbb{R}, v \mapsto \|v\|$, welche man als *Norm* bezeichnet, so dass gilt

- (a) $\|v\| \geq 0$ für alle $v \in V$,
- (b) $\|v\| = 0$ dann und nur dann, wenn $v = 0$,
- (c) $\|\alpha v\| = |\alpha| \|v\|$ für alle $\alpha \in \mathbb{R}$ und $v \in V$,
- (d) $\|v + w\| \leq \|v\| + \|w\|$ für alle $v, w \in V$ (Dreiecksungleichung).

1.2 Lemma (Umgekehrte Dreiecksungleichung). Für $v, w \in V$ gilt $\|v - w\| \geq |\|v\| - \|w\||$.

Beweis.

$$\|v\| \leq \|v - w\| + \|w\|,$$

also $\|v - w\| \leq \|v\| - \|w\|$. Genauso sieht man $\|w - v\| \leq \|w\| - \|v\|$. $\square$

1.3 Definition. Sei $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$. Die *euklidische Norm* von x ist definiert als

$$\|x\|_2 = \sqrt{x_1^2 + \dots + x_n^2}.$$

Jetzt würde man gerne beweisen, dass die euklidische Norm ihren Namen zu Recht trägt. Dazu benötigen wir zuerst eine Schreibweise und eine Ungleichung.

1.4 Definition. Für $x, y \in \mathbb{R}^n$ definieren wir das Skalarprodukt durch

$$x \cdot y := \sum_{j=1}^n x_j y_j.$$

1.5 Satz (Cauchy-Schwarzsche Ungleichung). Sind $x = (x_1, x_2, \dots, x_n)$ und $y = (y_1, y_2, \dots, y_n)$ im $\mathbb{R}^n$, so gilt $|x \cdot y| \leq \|x\|_2 \|y\|_2$.

Beweis. Das wurde in der Linearen Algebra I bewiesen. Wir zeigen nachher die Höldersche Ungleichung, aus welcher die Cauchy-Schwarzsche Ungleichung folgt. $\square$

1.6 Satz. Die euklidische Norm ist in der Tat eine Norm.

1. Normierte Räume

Beweis. (a)–(c) sind unproblematisch. Die Dreiecksungleichung beweisen wir wie folgt

$$\begin{aligned}\|v + w\|_2^2 &= \sum_{j=1}^n (v_j + w_j)^2 = \sum_{j=1}^n v_j^2 + 2 \sum_{j=1}^n v_j w_j + \sum_{j=1}^n w_j^2 \\ &\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2. \quad \square\end{aligned}$$

1.7 Definition. Für $x \in \mathbb{R}^n$ und $p \geq 1$ definieren wir

$$\|x\|_p := \left(\sum_{j=1}^n |x_j|^p \right)^{1/p}.$$

Außerdem

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

1.8 Lemma. Für $a, b > 0$ und $p, q \geq 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$ gilt

$$a^{1/p} b^{1/q} \leq \frac{a}{p} + \frac{b}{q}.$$

Beweis. Der Logarithmus ist konkav, also

$$\ln\left(\frac{1}{p}a + \frac{1}{q}b\right) \geq \frac{1}{p} \ln a + \frac{1}{q} \ln b. \quad \square$$

1.9 Satz (Höldersche Ungleichung). Für $x, y \in \mathbb{R}^n$ und $p, q \in]1, \infty[$ mit $\frac{1}{p} + \frac{1}{q} = 1$ gilt

$$|x \cdot y| \leq \|x\|_p \|y\|_q.$$

Beweis. Wir nehmen $x \neq 0$ und $y \neq 0$ an, weil nur in diesen Fällen etwas zu beweisen ist. Wir setzen

$$\xi = \frac{1}{\|x\|_p^p} (|x_1|^p, \dots, |x_n|^p) \quad \text{und} \quad \eta = \frac{1}{\|y\|_q^q} (|y_1|^q, \dots, |y_n|^q).$$

Für $j = 1, \dots, n$ folgt aus dem Lemma, angewandt auf ξ_j und η_j , dass

$$\frac{|x_j|}{\|x\|_p} \frac{|y_j|}{\|y\|_q} = \xi_j^{1/p} \eta_j^{1/q} \leq \frac{\xi_j}{p} + \frac{\eta_j}{q} = \frac{1}{p} \frac{|x_j|^p}{\|x\|_p^p} + \frac{1}{q} \frac{|y_j|^q}{\|y\|_q^q}$$

Summiert man die linke Seite auf, so erhält man einen Wert, der mindestens so groß ist wie $\frac{|x \cdot y|}{\|x\|_p \|y\|_q}$, während auf der rechten Seite 1 herauskommt. $\square$

1.10 Satz (Minkowskische Ungleichung). Seien $x, y \in \mathbb{R}^n$ und sei $1 \leq p \leq \infty$. Dann gilt

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Beweis. Für $p = 1$ und $p = \infty$ folgt die Aussage aus der Dreiecksungleichung für Zahlen. Sei also $1 < p < \infty$. Dann gibt es $q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Definiere

$$z := (|x_1 + y_1|^{p-1}, \dots, |x_n + y_n|^{p-1}).$$

Für $j = 1, \dots, n$ erhalten wir wegen $q(p-1) = p$ die Gleichheit $z_j^q = |x_j + y_j|^p$, also $\|z\|_q = \|x + y\|_p^{p/q}$. Ferner gilt $|x_j + y_j|^p = |x_j + y_j|z_j \leq |x_j|z_j + |y_j|z_j$. Die Höldersche Ungleichung liefert daher

$$\begin{aligned} \|x + y\|_p^p &= \sum_{j=1}^n |x_j + y_j|^p \leq \sum_{j=1}^n |x_j|z_j + \sum_{j=1}^n |y_j|z_j \leq \|x\|_p \|z\|_q + \|y\|_p \|z\|_q \\ &= (\|x\|_p + \|y\|_p) \|x + y\|_p^{p/q}. \end{aligned}$$

Nun teilt man beide Seiten durch $\|x + y\|_p^{p/q}$ und berücksichtigt noch, dass $p - \frac{p}{q} = 1$. □

1.11 Korollar. Für $1 \leq p \leq \infty$ ist $\|\cdot\|_p$ eine Norm.

Übungsaufgabe: Normen ausrechnen, Einheitskreise zeichnen

Tutorium: Lineare Algebra wiederholen, Norm auf dem Raum aller Polynome

1.12 Definition. Zwei Normen $\|\cdot\|$ und $|||\cdot|||$ auf einem $\mathbb{R}$ -Vektorraum V heißen *äquivalent*, wenn es $a, b > 0$ gibt, so dass

$$a\|v\| \leq |||v||| \leq b\|v\| \quad \text{für alle } v \in V.$$

1.13 Satz. Auf dem $\mathbb{R}^n$ sind alle p -Normen mit $1 \leq p \leq \infty$ äquivalent.

Beweis.

$$\|x\|_p^p = \sum_{j=1}^n |x_j|^p \leq n \|x\|_\infty^p,$$

also $\|x\|_p \leq \sqrt[p]{n} \|x\|_\infty$. Ferner folgt aus der Hölderschen Ungleichung

$$\|x\|_1 = \sum_{j=1}^n 1|x_j| \leq \|1\|_q \|x\|_p = \sqrt[q]{n} \|x\|_p.$$

Die Ungleichung $\|x\|_\infty \leq \|x\|_1$ ist offensichtlich. □

Wir werden später den Satz beweisen, dass je zwei Normen auf dem $\mathbb{R}^n$ äquivalent sind.

2. Metrische Räume

Der Begriff der Norm verwendet die Vektorraumstruktur. Für einige Grundbegriffe wie z. B. Stetigkeit ist die Vektorraumstruktur aber nicht notwendig.

2.1 Definition. Ein *metrischer Raum* besteht aus einer Menge X und einer Abbildung $d: X \times X \rightarrow \mathbb{R}$, genannt *Metrik*, mit den folgenden Eigenschaften

- (a) $d(x, y) \geq 0$ für alle $x, y \in X$,
- (b) $d(x, y) = 0$ genau dann, wenn $x = y$,
- (c) $d(x, y) = d(y, x)$ für alle $x, y \in X$,
- (d) $d(x, z) \leq d(x, y) + d(y, z)$ für alle $x, y, z \in X$ (Dreiecksungleichung).

2.2 Beispiel. Jeder normierte Raum wird durch die Setzung $d(x, y) = \|x - y\|$ zu einem metrischen Raum.

2.3 Beispiel. Wenn X ein metrischer Raum und Y eine Teilmenge von X ist, so wird auch Y zu einem metrischen Raum, indem man die Metrik $d: X \times X \rightarrow \mathbb{R}$ auf $Y \times Y$ einschränkt. Das ist die *Teilraummetrik*.

2.4 Beispiel. Auf jeder beliebigen Menge X gibt es die folgende, ganz langweilige Metrik

$$d(x, y) = \begin{cases} 0, & \text{falls } x = y, \\ 1, & \text{sonst.} \end{cases}$$

Sie heißt *diskrete Metrik*.

Zwei Beispiele aus der Praxis:

2.5 Beispiel. (a) (Luftlinie) Die Entfernung zwischen zwei Punkten auf der Erdoberfläche ist die Länge der kürzesten Verbindung auf der Erdoberfläche. Beispielsweise beträgt die Entfernung vom Nord- zum Südpol 20 000 km, weil ein Großkreis durchlaufen werden muss.

In der Teilraummetrik des $\mathbb{R}^3$ beträgt der Abstand dagegen nur den doppelten Erdradius, also ungefähr 12 740 km.

(b) (Hamming-Abstand) Auf dem Raum $E = \{0, 1\}^N$ wird der Hamming-Abstand wie folgt erklärt: Für $x, y \in E$ ist $d(x, y)$ gleich der Anzahl der unterschiedlichen Einträge. Man sieht leicht, dass es sich um eine Metrik handelt.

2. Metrische Räume

Eine Teilmenge $C \subseteq E$ ist ein *Code*. Ein Element $x \in C$ wird durch Störung zu einem Element $y \in E$ verändert. Die Dekodierung besteht darin, dasjenige Element $z \in C$ zu finden, für welches $d(y, z)$ minimal ist. Wenn dieses Element nicht eindeutig ist, kann y nicht dekodiert werden.

2.6 Definition. Sei X ein metrischer Raum mit Metrik d . Für $a \in X$ und $r > 0$ definieren wir die *offene Kugel* mit Mittelpunkt a und Radius r als

$$B_r(a) = \{x \in X \mid d(a, x) < r\}.$$

Die *abgeschlossene Kugel* mit Mittelpunkt a und Radius r wird definiert durch

$$\bar{B}_r(a) = \{x \in X \mid d(a, x) \leq r\}.$$

Die offene und die abgeschlossene Kugel brauchen sich nicht zu unterscheiden.

2.7 Definition. Sei a ein Punkt eines metrischen Raums X . Eine Menge $U \subseteq X$ heißt *Umgebung* von a , wenn es ein $r > 0$ gibt, so dass $B_r(a) \subseteq U$.

2.8 Beispiel. Für jedes $x \in B_r(a)$ ist $B_r(a)$ eine Umgebung von x .

Beweis. Wenn $x \in B_r(a)$, dann $\epsilon := r - d(a, x) > 0$. Wir zeigen $B_\epsilon(x) \subseteq B_r(a)$. Sei dazu $y \in B_\epsilon(x)$, dann gilt $d(a, y) \leq d(a, x) + d(x, y) < d(a, x) + \epsilon = r$. $\square$

2.9 Definition. Sei X ein metrischer Raum. Eine Teilmenge $G \subseteq X$ heißt *offen* in X , wenn sie Umgebung aller ihrer Punkte ist.

2.10 Beispiel. (a) $X = \mathbb{R}$ mit der Metrik $d(x, y) = |x - y|$. Seien $c < d$ reelle Zahlen. Dann ist das Intervall $]c, d[$ offen und die Intervalle $[c, d[$, $]c, d]$ und $[c, d]$ sind nicht offen.

(b) Sei X ein metrischer Raum, sei $a \in X$ und sei $r > 0$. Dann ist $B_r(a)$ offen.

(c) Der Quadrant $Q = \{(x, y) \in \mathbb{R}^2 \mid x > 0, y > 0\}$ ist offen.

Beweis. (a) Sei zuerst $I =]c, d[$. Sei $a \in I$ beliebig. Dann sind $\epsilon_1 := a - c$ und $\epsilon_2 := d - a$ positiv. Setzt man $r := \min(\epsilon_1, \epsilon_2)$, so gilt $B_r(a) =]a - r, a + r[\subseteq I$. Also ist I Umgebung von a .

Für den zweiten Teil der Behauptung muss man ebenfalls eine Fallunterscheidung vornehmen. Ist etwa I von der Form $]c, d]$, so ist I keine Umgebung von d .

(b) Das wurde bereits in Beispiel 2.8 gezeigt. $\square$

2.11 Satz. Sei X ein metrischer Raum. Dann gelten

(a) X und $\emptyset$ sind offen in X .

(b) Ist I irgendeine Menge und ist für jedes $i \in I$ eine offene Teilmenge $G_i \subseteq X$ gegeben, dann ist $\bigcup_{i \in I} G_i$ offen in X .

(c) Sind $G_1, \dots, G_n$ endlich viele offene Mengen in X , so ist $G_1 \cap \dots \cap G_n$ offen in X .

2.12 Beispiel. $I = \mathbb{N}$ und $A_n =]-\frac{1}{n}, \frac{1}{n}[$. Dann ist jedes A_n offen, aber $\bigcap_{n \in \mathbb{N}} A_n = \{0\}$ ist nicht offen.

2.13 Definition. Sei X eine Menge. Ein System $\mathcal{T}$ von Teilmengen von X heißt *Topologie* auf X , wenn gilt

(a) $X \in \mathcal{T}$ und $\emptyset \in \mathcal{T}$.

(b) Ist I irgendeine Menge und ist für jedes $i \in I$ ein $G_i \in \mathcal{T}$ gegeben, so gilt $\bigcup_{i \in I} G_i \in \mathcal{T}$.

(c) Sind $G_1, \dots, G_n \in \mathcal{T}$ für ein $n \in \mathbb{N}$, so gilt $G_1 \cap \dots \cap G_n \in \mathcal{T}$.

Die Elemente von $\mathcal{T}$ sind die *offenen* Mengen (bzgl. $\mathcal{T}$).

Satz 2.11 besagt also, dass jede Metrik eine Topologie induziert.

2.14 Bemerkung. Eigenschaften von Umgebungen:

(a) Für $x \in X$ und $U \subseteq X$ sind äquivalent

(i) U ist Umgebung von x .

(ii) Es gibt eine offene Menge G mit $x \in G \subseteq U$.

Insbesondere hängt der Umgebungsbegriff nur von der Topologie ab.

(b) Eine Menge G ist genau dann offen, wenn es zu jedem $a \in G$ ein $r > 0$ gibt, so dass $B_r(a) \subseteq G$.

(c) Jede Obermenge einer Umgebung von x ist wieder Umgebung von x .

(d) Der Durchschnitt endlich vieler Umgebungen von x ist wieder eine Umgebung von x .

2.15 Satz. Seien $\|\cdot\|$ und $|||\cdot|||$ zwei äquivalente Normen auf einem $\mathbb{R}$ -Vektorraum V und sei $X \subseteq V$. Die Metriken d_1 und d_2 auf X seien gegeben durch $d_1(x, y) = \|x - y\|$ und $d_2(x, y) = |||x - y|||$. Dann stimmt die von d_1 induzierte Topologie mit der von d_2 induzierten Topologie überein.

Beweis. Wegen der Äquivalenz der Normen gibt es $A, B > 0$, so dass

$$A\|v\| \leq |||v||| \leq B\|v\| \quad \text{für alle } v \in V.$$

2. Metrische Räume

Sei $G \subseteq X$ offen bzgl. der Metrik d_1 und sei $a \in G$. Wir setzen $B_r^{(j)}(a) = \{x \in X \mid d_j(a, x) < r\}$. Um zu zeigen, dass G offen bzgl. d_2 ist, müssen wir $r > 0$ finden, so dass $B_r^{(2)}(a) \subseteq G$. Jedenfalls existiert $\epsilon > 0$, so dass $B_\epsilon^{(1)}(a) \subseteq G$. Setze $r = \epsilon A$. Wenn $x \in B_r^{(2)}(a)$, dann $\|x - a\| < r$, also $\|x - a\| < \frac{r}{A} = \epsilon$ und daher $x \in B_\epsilon^{(1)}(a) \subseteq G$. Vertauschung von d_1 und d_2 gibt die Rückrichtung. $\square$

2.16 Definition. Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in einem metrischen Raum X , und sei $y \in X$. Der Punkt y heißt *Grenzwert* der Folge $(x_n)_{n \in \mathbb{N}}$, wenn es zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $d(x_n, y) < \epsilon$ für alle $n \geq N$.

Man schreibt dann $\lim_{n \rightarrow \infty} x_n = y$ oder $x_n \rightarrow y$.

2.17 Satz. Sei (X, d) ein metrischer Raum. Dann sind äquivalent

- (a) $\lim_{n \rightarrow \infty} x_n = y$.
- (b) Zu jedem $\epsilon > 0$ existiert ein $N \in \mathbb{N}$, so dass $x_n \in B_\epsilon(y)$ für alle $n \geq N$.
- (c) Zu jeder Umgebung U von y existiert ein $N \in \mathbb{N}$, so dass $x_n \in U$ für alle $n \geq N$.

2.18 Korollar. Wenn eine Folge bezüglich einer Metrik konvergiert, die von einer Norm induziert wird, dann konvergiert sie auch bezüglich jeder Metrik, die von einer äquivalenten Norm induziert wird.

2.19 Satz. Keine Folge besitzt mehr als einen Grenzwert.

Beweis. Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in X . Zum Widerspruchsbeweis nehmen wir an, dass sie zwei verschiedene Grenzwerte $a \neq b$ besitzt. Zu $\epsilon = d(a, b)/2 > 0$ existieren $N_1, N_2 \in \mathbb{N}$, so dass $d(a, x_n) < \epsilon$ für alle $n \geq N_1$ und $d(b, x_n) < \epsilon$ für alle $n \geq N_2$. Für $N = \max(N_1, N_2)$ gilt

$$d(a, b) \leq d(a, x_N) + d(x_N, b) < \epsilon + \epsilon = d(a, b). \quad \square$$

2.20 Satz. Es seien (X, d) ein metrischer Raum, $Y \subseteq X$ und d_Y die Teilraummetrik auf Y . Dann ist G genau dann offen in (Y, d_Y) , wenn es eine in X offene Menge H mit $G = H \cap Y$ gibt.

Beweis. Ich bezeichne die r -Kugeln in Y mit B_r^Y , diejenigen in X mit B_r . Sei G offen in Y . Zu jedem $y \in G$ existiert $r_y > 0$, so dass $B_{r_y}^Y(y) \subseteq G$. Dann ist $H := \bigcup_{y \in G} B_{r_y}(y)$ offen in X und wegen $B_{r_y}(y) \cap Y = B_{r_y}^Y(y)$ gilt $G = H \cap Y$.

Sei nun H offen in X und sei $G = H \cap Y$. Für beliebiges $y \in G$ existiert $r > 0$ mit $B_r(y) \subseteq H$, also $B_r^Y(y) = B_r(y) \cap Y \subseteq G$. $\square$

2.21 Definition. Es sei X ein metrischer Raum, es sei $x \in X$, und es sei $A \subseteq X$.

x heißt *Berührungspunkt* von A , falls in jeder Umgebung von x ein Punkt von A liegt.

x heißt *Häufungspunkt* von A , falls in jeder Umgebung von x ein von x verschiedener Punkt von A liegt.

2.22 Beispiel. (a) Jedes Element von A ist Berührungspunkt von A .

(b) Ein Punkt $x \in \mathbb{R}$ ist genau dann ein Berührungspunkt von $A = \{\frac{1}{n} \mid n \in \mathbb{N}\}$, wenn $x \in A$ oder $x = 0$, aber nur 0 ist ein Häufungspunkt von A .

(c) Der $\mathbb{R}^n$ sei versehen mit einer Metrik, die von einer Norm $\|\cdot\|$ induziert wird. Dann ist ein Punkt $x \in \mathbb{R}^n$ genau dann Berührungspunkt von $B_r(a)$, wenn $x \in \overline{B}_r(a)$. Alle diese Punkte sind auch Häufungspunkte.

Beweis. (b) " $\subseteq$ ": Wir zeigen: Wenn $x \notin A \cup \{0\}$, dann ist x kein Berührungspunkt von A . Dazu sind verschiedene Fälle zu unterscheiden. Wenn $x < 0$, dann liegt in $B_{|x|}(x)$ kein Punkt von A . Wenn $x > 1$, dann liegt in $B_{x-1}(x)$ kein Punkt von A . Wenn $x \in]0, 1[\setminus A$, dann gibt es ein $n \in \mathbb{N}$, so dass $\frac{1}{n+1} < x < \frac{1}{n}$. Setzt man $\epsilon := \min(x - \frac{1}{n+1}, \frac{1}{n} - x)$, dann liegt in $B_\epsilon(x)$ kein Punkt von A .

" $\supseteq$ ": Alle Punkte von A sind automatisch Berührungspunkte. Um zu zeigen, dass 0 ein Berührungspunkt ist, sei $\epsilon > 0$ beliebig vorgegeben. Wenn $n > \frac{1}{\epsilon}$, dann $\frac{1}{n} \in B_\epsilon(0)$.

Wir haben sogar gezeigt, dass 0 Häufungspunkt von A ist. Wenn $x = \frac{1}{n}$ ein Element von A ist, dann setzen wir $r = \frac{1}{n} - \frac{1}{n+1} > 0$. Dann ist x das einzige Element von A in $B_r(x)$, also ist x kein Häufungspunkt.

(c): Seien $a = 0$, $x \in \overline{B}_r(0)$ und $\epsilon > 0$.

Im Fall $x = 0$ setzen wir $y = (\min(r, \frac{\epsilon}{2}), 0, \dots, 0)$, im Fall $\epsilon > |x| > 0$ setzen wir $y = 0$ und im Fall $\epsilon \leq |x| \leq r$ setzen wir $y = (1 - \frac{\epsilon}{2r})x$. In allen Fällen gilt $x \neq y \in A \cap B_\epsilon(x)$; daher ist x Häufungspunkt von $B_r(0)$.

Ist umgekehrt $x \notin \overline{B}_r(0)$, dann ist $\epsilon := \|x\| - r > 0$ und $B_\epsilon(x) \cap \overline{B}_r(0) = \emptyset$. (Analog für $a \neq 0$.) $\square$

2.23 Bemerkung. Sei X ein metrischer Raum, sei $A \subseteq X$ und $x \in X$.

(a) x ist genau dann Berührungspunkt von A , wenn es eine Folge $(x_n)_{n \in \mathbb{N}}$ in A gibt, so dass $\lim_{n \rightarrow \infty} x_n = x$.

(b) x ist genau dann Häufungspunkt von A , wenn es eine Folge $(x_n)_{n \in \mathbb{N}}$ in A gibt, so dass $\lim_{n \rightarrow \infty} x_n = x$ und $x_n \neq x$ für alle n .

2.24 Satz. Sei X ein metrischer Raum und sei $A \subseteq X$. Dann sind äquivalent:

(a) A enthält alle Berührungspunkte von A .

(b) A enthält alle Häufungspunkte von A .

(c) $X \setminus A$ ist offen in X .

Beweis. (a) $\implies$ (b): Häufungspunkte sind Berührungspunkte.

(b) $\implies$ (c): Beweis durch Kontraposition. Wir nehmen also an, $X \setminus A$ sei nicht offen. Dann gibt es ein $x \in X \setminus A$, so dass keine Umgebung von x in $X \setminus A$ enthalten

2. Metrische Räume

ist. Das bedeutet, dass jede Umgebung von x ein Element von A enthält. Daher ist x Berührungspunkt von A , und wegen $x \notin A$ sogar Häufungspunkt.

(c) $\implies$ (a): Sei $x \in X \setminus A$. Dann ist $X \setminus A$ eine Umgebung von x , die kein Element von A enthält. Folglich ist x kein Berührungspunkt von A . $\square$

2.25 Definition. Wenn eine Teilmenge A eines metrischen Raums X die äquivalenten Eigenschaften aus Satz 2.24 besitzt, dann sagt man, A sei *abgeschlossen* in X .

2.26 Bemerkung. (a) Abgeschlossene Intervalle sind abgeschlossen im Sinne dieser Definition.

(b) A ist genau dann abgeschlossen, wenn gilt: Ist $(x_n)_{n \in \mathbb{N}}$ eine Folge in A , die einen Grenzwert in X besitzt, so liegt dieser Grenzwert bereits in A .

2.27 Satz (De Morgansche Regeln).

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

2.28 Satz. Sei X ein metrischer Raum. Dann gelten

(a) $\emptyset$ und X sind abgeschlossen.

(b) Der Durchschnitt von beliebig vielen abgeschlossenen Mengen ist abgeschlossen.

(c) Die Vereinigung von endlich vielen abgeschlossenen Mengen ist abgeschlossen.

Beweis. Verwende die de Morganschen Regeln. $\square$

2.29 Beispiel. $\bigcup_{n \in \mathbb{N}} \left[-1 + \frac{1}{n}, 1 - \frac{1}{n}\right] =]-1, 1[$ ist eine Menge, die nicht abgeschlossen, aber Vereinigung von abgeschlossenen Mengen ist.

2.30 Satz. Sei X ein metrischer Raum und $A \subseteq X$ eine endliche Teilmenge. Dann ist A abgeschlossen.

Beweis. Wir brauchen das nur für einelementige Mengen $\{x\}$ zu machen. Sei also $y \in X \setminus \{x\}$. Dann $d(x, y) > 0$. Sei $0 < r < d(x, y)$. Dann $B_r(y) \subseteq X \setminus \{x\}$. Wir haben gezeigt, dass $X \setminus \{x\}$ offen ist. $\square$

2.31 Satz. Es seien (X, d) ein metrischer Raum, $Y \subseteq X$ und d_Y die Teilraummetrik auf Y . Eine Menge $A \subset Y$ ist genau dann abgeschlossen in (Y, d_Y) , wenn es eine in X abgeschlossene Menge B mit $A = B \cap Y$ gibt.

2.32 Satz. Sei $X = \mathbb{R}^n$ mit einer zu $\|\cdot\|_\infty$ äquivalenten Norm. Sei $(x^{(k)})_{k \in \mathbb{N}}$ eine Folge im $\mathbb{R}^n$ mit $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$, und sei $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$. Dann gilt $\lim_{k \rightarrow \infty} x^{(k)} = x^{(0)}$ genau dann, wenn für jedes $j \in \{1, \dots, n\}$ gilt $\lim_{k \rightarrow \infty} x_j^{(k)} = x_j^{(0)}$.

Da wir später sehen werden, dass alle Normen auf dem $\mathbb{R}^n$ äquivalent sind, ist dieser Teil der Voraussetzung überflüssig.

Man sagt: Eine Folge konvergiert genau dann im $\mathbb{R}^n$, wenn sie komponentenweise konvergiert.

Beweis. [von Satz 2.32] Sei zunächst $\lim_{k \rightarrow \infty} x^{(k)} = x^{(0)}$. Sei $j = 1, \dots, n$. Wir zeigen $\lim_{k \rightarrow \infty} x_j^{(k)} = x_j^{(0)}$. Sei dazu $\epsilon > 0$ vorgegeben. Dann existiert $N \in \mathbb{N}$, so dass $\|x^{(k)} - x^{(0)}\|_\infty < \epsilon$ für alle $k \geq N$. Wegen $|x_j^{(k)} - x_j^{(0)}| \leq \|x^{(k)} - x^{(0)}\|_\infty$ folgt $|x_j^{(k)} - x_j^{(0)}| < \epsilon$.

Nun setzen wir umgekehrt voraus, dass $\lim_{k \rightarrow \infty} x_j^{(k)} = x_j^{(0)}$ für $j = 1, \dots, n$. Sei $\epsilon > 0$ beliebig vorgegeben. Dann existiert für jedes j ein $N_j \in \mathbb{N}$, so dass $|x_j^{(k)} - x_j^{(0)}| < \epsilon$ für alle $k \geq N_j$. Sei $N = \max\{N_1, \dots, N_n\}$. Dann $\|x^{(k)} - x^{(0)}\|_\infty < \epsilon$ für alle $k \geq N$. $\square$

2.33 Definition. Sei X ein metrischer Raum, sei $A \subseteq X$. Die Menge aller Berührungspunkte von A heißt *Abschluss* von A , i. Z. $\bar{A}$.

Ein Punkt $a \in A$ ist ein *innerer Punkt* von A , wenn A Umgebung von a ist. Die Menge der inneren Punkte von A bezeichnet man mit $\mathring{A}$.

2.34 Beispiel. In $X = \mathbb{R}^n$, versehen mit einer Norm $\|\cdot\|$, gilt $\overline{B_r(a)} = \bar{B}_r(a)$. Das ist der Inhalt von Beispiel 2.22.

2.35 Bemerkung. $\bar{A}$ ist die kleinste abgeschlossene Teilmenge von X , die A umfasst.

Beweis. Wenn B eine abgeschlossene Menge ist, die A umfasst, dann ist jeder Berührungspunkt von A auch Berührungspunkt von B , also Element von B . Das zeigt $\bar{A} \subseteq B$.

Für die andere Inklusion ist zu zeigen, dass $\bar{A}$ abgeschlossen ist. Sei also x ein Berührungspunkt von $\bar{A}$. Dann existiert eine Folge $(x_n)_{n \in \mathbb{N}}$ in $\bar{A}$ mit $\lim_{n \rightarrow \infty} x_n = x$. Zu jedem n existiert ein $y_n \in A$ mit $d(x_n, y_n) < \frac{1}{n}$. Dann $d(x, y_n) \leq d(x, x_n) + \frac{1}{n} \rightarrow 0$. Also $x \in \bar{A}$. $\square$

2.36 Definition. Sei X ein metrischer Raum, $A \subseteq X$ und $x \in X$. x heißt *Randpunkt* von A , wenn x Berührungspunkt von A und von $X \setminus A$ ist. Die Menge aller Randpunkte ist der *Rand* von A , man schreibt dafür ∂A .

2.37 Bemerkung. ∂A ist abgeschlossen in X . Es gilt nämlich $\partial A = \bar{A} \cap \overline{X \setminus A}$.

2.38 Beispiel. In $\mathbb{R}^n$ mit einer Norm $\|\cdot\|$ gilt

$$\partial B_r(a) = \partial \bar{B}_r(a) = \{x \in \mathbb{R}^n \mid \|x - a\| = r\}.$$

3. Stetige Abbildungen

3.1 Definition. Seien (X, d_1) und (Y, d_2) metrische Räume, sei $f: X \rightarrow Y$ eine Abbildung, und sei $x_0 \in X$. f heißt stetig im Punkt x_0 , wenn es zu jedem $\epsilon > 0$ ein $\delta > 0$ gibt, so dass für jedes $x \in X$ mit $d_1(x, x_0) < \delta$ gilt $d_2(f(x), f(x_0)) < \epsilon$.

Eine Abbildung f heißt *stetig*, wenn sie in jedem Punkt ihres Definitionsbereichs stetig ist.

3.2 Bemerkung. f ist genau dann stetig in x_0 , wenn es zu jeder Umgebung U von $f(x_0)$ eine Umgebung V von x_0 gibt mit $f(V) \subseteq U$.

Daher hängt Stetigkeit nur von der Topologie ab.

3.3 Beispiel. (a) Alle konstanten Abbildungen sind stetig.

(b) Seien X, Y und Z metrische Räume, $f: X \rightarrow Y$ und $g: Y \rightarrow Z$ Abbildungen und sei $x_0 \in X$. Falls f stetig in x_0 und g stetig in $f(x_0)$ ist, so ist $g \circ f$ stetig in x_0 .

3.4 Satz. Seien X und Y metrische Räume, sei $f: x \rightarrow y$ eine Abbildung. Dann sind äquivalent:

(a) f ist stetig.

(b) Für jede offene Menge $G \subseteq Y$ ist $f^{-1}(G)$ offen in X .

(c) Für jede abgeschlossene Menge $A \subseteq Y$ ist $f^{-1}(A)$ abgeschlossen in X .

(d) Für jede konvergente Folge $(x_n)_{n \in \mathbb{N}}$ in X gilt

$$\lim_{n \rightarrow \infty} f(x_n) = f\left(\lim_{n \rightarrow \infty} x_n\right).$$

Beweis. (a) $\implies$ (b): Sei $x \in f^{-1}(G)$. Dann ist G eine Umgebung von $f(x)$. Da f stetig in x ist, existiert eine Umgebung V von x derart, dass $f(V) \subseteq G$. Daraus folgt $V \subseteq f^{-1}(G)$. Also ist $f^{-1}(G)$ eine Umgebung von x .

(b) $\implies$ (d): Sei $x_0 = \lim_{n \rightarrow \infty} x_n$. Sei U eine offene Umgebung von $f(x_0)$. Dann ist $V = f^{-1}(U)$ eine offene Umgebung von x_0 . Also existiert ein $N \in \mathbb{N}$, so dass $x_n \in V$ für alle $n \geq N$. Für diese n gilt $f(x_n) \in U$.

(d) $\implies$ (c): Sei A eine abgeschlossene Teilmenge von Y , und sei x_0 ein Berührungspunkt von $f^{-1}(A)$. Dann gibt es eine Folge $(x_n)_{n \in \mathbb{N}}$ in $f^{-1}(A)$ mit $x_n \rightarrow x_0$. Nach Voraussetzung gilt $f(x_n) \rightarrow f(x_0)$. Alle $f(x_n)$ liegen in A , ihr Grenzwert daher auch. Also $x_0 \in f^{-1}(A)$.

3. Stetige Abbildungen

(c) $\implies$ (a): Sei $x_0 \in X$, und sei V eine offene Umgebung von $f(x_0)$. Wegen Satz 2.24 ist $A = Y \setminus V$ abgeschlossen. Wegen (c) ist $f^{-1}(A)$ abgeschlossen in X . Also ist $X \setminus f^{-1}(A) = f^{-1}(V)$ offen. Es ist klar, dass $x_0 \in f^{-1}(V)$. $\square$

3.5 Korollar. Sei X ein metrischer Raum, sei $x_0 \in X$, und seien $f, g: X \rightarrow \mathbb{R}$ stetig in x_0 . Dann sind auch die Funktionen $f + g$, $f - g$ und $f \cdot g$ stetig in x_0 . Falls zusätzlich $g(x_0) \neq 0$, so gibt es eine Umgebung U von x_0 , so dass $g(x) \neq 0$ für alle $x \in U$, und die Funktion $f/g: U \rightarrow \mathbb{R}$ ist stetig in x_0 .

3.6 Beispiel. Der Quadrant $Q = \{(x, y) \in \mathbb{R}^2 \mid x > 0, y > 0\}$ ist offen.

Beweis. Die Abbildungen $f: \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x$, und $g: \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto y$, sind stetig und

$$Q = f^{-1}(]0, \infty[) \cap g^{-1}(]0, \infty[). \quad \square$$

3.7 Korollar. Seien X, Y metrische Räume und seien $f, g: X \rightarrow Y$ stetig. Dann ist $A = \{x \in X \mid f(x) = g(x)\}$ abgeschlossen in X .

Beweis. Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in A , welche gegen ein $x_0 \in X$ konvergiert. Dann gilt wegen der Stetigkeit von f und g

$$f(x_0) = f\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} g(x_n) = f\left(\lim_{n \rightarrow \infty} x_n\right) = g(x_0). \quad \square$$

3.8 Korollar. Sei X ein metrischer Raum und seien $f, g: X \rightarrow \mathbb{R}$ stetig. Dann ist $G = \{x \mid f(x) < g(x)\}$ offen in X und $A = \{x \mid f(x) \leq g(x)\}$ ist abgeschlossen.

Beweis. $G = h^{-1}(]-\infty, 0[)$ für $h(x) = f(x) - g(x)$ und $A = h^{-1}([0, \infty[)$. $\square$

3.9 Satz. Sei X ein metrischer Raum und sei $f: X \rightarrow \mathbb{R}^n$ eine Abbildung. Wir schreiben $f(x) = (f_1(x), \dots, f_n(x))$ mit Abbildungen $f_j: X \rightarrow \mathbb{R}, j = 1, \dots, n$. Wir versehen $\mathbb{R}^n$ mit einer zu $\|\cdot\|_\infty$ äquivalenten Norm. Dann ist f genau dann stetig, wenn alle f_j stetig sind.

Beweis. Das folgt aus Satz 2.32. $\square$

4. Kompakte Mengen

Ziel in diesem Abschnitt ist die Verallgemeinerung des Satzes von Bolzano-Weierstraß.

4.1 Definition. Eine Teilmenge A eines metrischen Raums X heißt *kompakt*, wenn jede Folge in A eine konvergente Teilfolge hat, deren Grenzwert in A liegt.

4.2 Bemerkung. (a) Alle Intervalle der Form $[a, b]$ mit $a, b \in \mathbb{R}$, $a < b$, sind kompakt. (Das ist der Satz von Bolzano-Weierstraß.)

(b) Wenn A kompakt ist, dann liegen alle Berührungspunkte von A in A . Kompakte Teilmengen sind daher abgeschlossen.

4.3 Satz. Jede abgeschlossene Teilmenge einer kompakten Menge ist kompakt.

4.4 Definition. Sei V ein normierter Raum. Eine Teilmenge A von V heißt *beschränkt*, wenn es ein $C > 0$ gibt, so dass $\|x\| \leq C$ für alle $x \in A$.

4.5 Lemma. Sei A eine kompakte Teilmenge eines normierten Raums V . Dann ist A beschränkt.

Beweis. Wenn A nicht beschränkt ist, dann gibt es zu jedem $n \in \mathbb{N}$ ein $x_n \in A$ mit $\|x_n\| \geq n$. Die Folge $(x_n)_{n \in \mathbb{N}}$ kann dann keine konvergente Teilfolge haben. $\square$

4.6 Lemma. Der $\mathbb{R}^n$ sei mit der Norm $\|\cdot\|_\infty$ versehen. Eine Teilmenge $A \subseteq \mathbb{R}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beweis. Wir müssen nur noch die Rückrichtung zeigen. Sei also $(x^{(k)})_{k \in \mathbb{N}}$ eine Folge in einer abgeschlossenen und beschränkten Menge A . Schreibe $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$. Dann ist die Folge $(x_1^{(k)})_{k \in \mathbb{N}}$ beschränkt, besitzt also eine in $\mathbb{R}$ konvergente Teilfolge $(x_1^{(k_1, m)})_{m \in \mathbb{N}}$ mit Grenzwert $x_1^{(0)}$. Die beschränkte Folge $(x_2^{(k_1, m)})_{m \in \mathbb{N}}$ besitzt wieder eine konvergente Teilfolge $(x_2^{(k_2, m)})_{m \in \mathbb{N}}$ mit Grenzwert $x_2^{(0)}$. Das machen wir für jede Komponente, bis wir eine Teilfolge $(x^{(k_n, m)})_{m \in \mathbb{N}}$ der Ausgangsfolge $(x^{(k)})_{k \in \mathbb{N}}$ erhalten, so dass für $j = 1, \dots, n$ gilt $\lim_{m \rightarrow \infty} x_j^{(k_n, m)} = x_j^{(0)}$. Aus Satz 2.32 folgt $\lim_{m \rightarrow \infty} x^{(k_n, m)} = x^{(0)}$ für $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$. Insbesondere ist $x^{(0)}$ ein Berührungspunkt von A . Da A abgeschlossen ist, liegt $x^{(0)}$ in A . $\square$

4.7 Beispiel. Abgeschlossene Quader der Form $\prod_{j=1}^n [a_j, b_j]$ sind kompakt.

4.8 Satz. Seien X, Y metrische Räume, sei $f: X \rightarrow Y$ eine stetige Abbildung, und sei $A \subseteq X$ kompakt. Dann ist $f(A)$ ebenfalls kompakt.

4. Kompakte Mengen

Prägnante Formulierung: “Das stetige Bild einer kompakten Menge ist kompakt.”

Beweis. Sei $(y_n)_{n \in \mathbb{N}}$ eine Folge in $f(A)$, dann gibt es $x_n \in A$ mit $f(x_n) = y_n$. Die Folge $(x_n)_{n \in \mathbb{N}}$ besitzt eine konvergente Teilfolge $(x_{n_m})_{m \in \mathbb{N}}$. Da f stetig ist, konvergiert auch $(f(x_{n_m}))_{m \in \mathbb{N}}$. $\square$

4.9 Satz. Sei X ein metrischer Raum, sei A eine kompakte Teilmenge von X , und sei $f: A \rightarrow \mathbb{R}$ stetig. Dann nimmt f auf A ihr Maximum und ihr Minimum an.

Beweis. $B := f(A)$ ist eine kompakte Teilmenge von $\mathbb{R}$, also beschränkt. Daher besitzt B ein Supremum. Dieses Supremum ist ein Berührungspunkt von B , liegt wegen der Abgeschlossenheit von B also in B . $\square$

4.10 Lemma. Es sei V ein normierter Raum und seien $b_1, \dots, b_n \in V$. Wenn der $\mathbb{R}^n$ mit der Supremumsnorm $\|\cdot\|_\infty$ versehen wird, dann ist die Abbildung

$$\Phi: \mathbb{R}^n \rightarrow V, \quad (w_1, \dots, w_n) \mapsto \sum_{j=1}^n w_j b_j,$$

stetig.

Beweis. Sei $L = \max_{j=1, \dots, n} \|b_j\|$. Zu vorgegebenem $\epsilon > 0$ setzen wir $\delta = \epsilon/(Ln)$. Für $(v_1, \dots, v_n)$ und $(w_1, \dots, w_n) \in \mathbb{R}^n$ mit $\|(w_1, \dots, w_n) - (v_1, \dots, v_n)\|_\infty < \delta$ gilt

$$\begin{aligned} \|\Phi(w_1, \dots, w_n) - \Phi(v_1, \dots, v_n)\| &\leq \sum_{j=1}^n |w_j - v_j| \|b_j\| \\ &\leq Ln \|(w_1, \dots, w_n) - (v_1, \dots, v_n)\|_\infty < Ln\delta = \epsilon. \quad \square \end{aligned}$$

4.11 Lemma. Es sei V ein endlich-dimensionaler, normierter Raum mit Basis $(b_1, \dots, b_n)$ und es sei $K \subseteq V$ beschränkt. Dann gibt es ein $C > 0$, so dass

$$\max_{j=1, \dots, n} |w_j| \leq C \quad \text{falls} \quad \sum_{j=1}^n w_j b_j \in K.$$

Beweis. Wir versehen den $\mathbb{R}^n$ mit der Supremumsnorm $\|\cdot\|_\infty$ und definieren $\Phi: \mathbb{R}^n \rightarrow V$ wie in Lemma 4.10. Dann ist Φ eine stetige lineare Abbildung. Wir setzen

$$A = \{(w_1, \dots, w_n) \in \mathbb{R}^n \mid \|(w_1, \dots, w_n)\|_\infty = 1\} \quad \text{und} \quad B = \Phi(A).$$

Dann ist A kompakt nach Lemma 4.6 und B ist kompakt als stetiges Bild eines Kompaktums. Daher existiert $\epsilon = \min\{\|v\| \mid v \in B\}$. Wegen $0 \notin B$ gilt $\epsilon > 0$.

Da K kompakt ist, existiert $R > 0$, so dass $K \subset B_R(0)$. Wir zeigen die Aussage für $C = R/\epsilon$. Sei dazu $v = \sum_{j=1}^n w_j b_j \in K$. Für $v = 0$ ist nichts zu zeigen. Wir können daher annehmen, dass $m = \max_{j=1, \dots, n} |w_j| > 0$. Dann $\frac{1}{m}(w_1, \dots, w_n) \in A$ und $\frac{1}{m}v \in B$. Wegen der Wahl von ϵ bedeutet die letzte Aussage, dass $\frac{1}{m}\|v\| \geq \epsilon$, also

$$m \leq \frac{\|v\|}{\epsilon} \leq \frac{R}{\epsilon} = C. \quad \square$$

4.12 Theorem (Heine-Borel). *Es sei V ein endlich dimensionaler, normierter Raum. Eine Menge $K \subseteq V$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Beweis. Nur die Rückrichtung ist zu zeigen. Dazu sei $\Phi: \mathbb{R}^n \rightarrow V$ wie in Lemma 4.10 und C sei wie in Lemma 4.11. In Lemma 4.11 haben wir gesehen, dass $K \subset \Phi([-C, C]^n)$. Die Menge auf der rechten Seite ist kompakt wegen Beispiel 4.7 und Satz 4.8. Da K eine abgeschlossene Teilmenge eines Kompaktums ist, ist sie kompakt nach Satz 4.3. $\square$

4.13 Satz. *Es seien $(V, \|\cdot\|)$ und $(W, |||\cdot|||)$ normierte Räume, und es sei $\dim V < \infty$. Für eine lineare Abbildung $A: V \rightarrow W$ setzen wir*

$$L := \sup\{|||Av||| \mid \|v\| \leq 1\}.$$

Dann gelten

- (a) $L < \infty$.
- (b) $|||Av||| \leq L\|v\|$ für alle $v \in V$.
- (c) A ist stetig.

Beweis. (a) Wegen Heine-Borel ist $\bar{B}_1(0) = \{v \in V \mid \|v\| \leq 1\}$ kompakt. Sei $(b_1, \dots, b_n)$ eine Basis von V und sei C wie in Lemma 4.11. Dann gilt für $v = \sum_{j=1}^n x_j b_j \in \bar{B}_1(0)$

$$|||Av||| \leq \sum_{j=1}^n |x_j| |||Ab_j||| \leq C \sum_{j=1}^n |||Ab_j|||.$$

- (b) Für beliebiges $v \in V$ mit $v \neq 0$ gilt

$$|||Av||| = \left| \left| \left| \|v\| A\left(\frac{v}{\|v\|}\right) \right| \right| \right| \leq \|v\| L.$$

- (c) Für gegebenes $\epsilon > 0$ sei $\delta = \frac{\epsilon}{1+L}$. Für $u, v \in V$ mit $\|u - v\| < \delta$ gilt

$$|||Au - Av||| = |||A(u - v)||| \leq L\|u - v\| < L\delta \leq \epsilon. \quad \square$$

4.14 Korollar. *Je zwei Normen auf dem $\mathbb{R}^n$ sind äquivalent.*

Beweis. Man wendet Satz 4.13 an auf die identische Abbildung $(V, \|\cdot\|) \rightarrow (V, |||\cdot|||)$. $\square$

4.15 Definition. Das L aus Satz 4.13 ist die *Operatornorm* von A . Man schreibt $\|A\|$ für die Operatornorm.

4.16 Satz. *Die Operatornorm ist tatsächlich eine Norm. Sie ist außerdem submultiplikativ, es gilt also $\|AB\| \leq \|A\| \|B\|$.*

4. Kompakte Mengen

4.17 Beispiel. Für $n, m \in \mathbb{N}$ und eine Matrix $A = (a_{k,j})_{k=1,\dots,n,j=1,\dots,m} \in \mathbb{R}^{n \times m}$ sei eine lineare Abbildung $f: (\mathbb{R}^m, \|\cdot\|_\infty) \rightarrow (\mathbb{R}^n, \|\cdot\|_\infty)$ gegeben durch $f(x) := Ax$. Für $x \in \mathbb{R}^m$ sei $Ax = (y_1, \dots, y_n)$. Dann gilt für $k = 1, \dots, n$

$$|y_k| = \left| \sum_{j=1}^m a_{k,j} x_j \right| \leq \|x\|_\infty \sum_{j=1}^m |a_{k,j}|.$$

Also

$$\|A\| \leq \max_{k=1,\dots,n} \sum_{j=1}^m |a_{k,j}|. \quad (4.1)$$

Wir behaupten die Gleichheit. Sei k ein Index, für den das Maximum in (4.1) angenommen wird. Wir definieren $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ durch

$$x_j = \begin{cases} \frac{|a_{k,j}|}{a_{k,j}}, & a_{k,j} \neq 0, \\ 1, & \text{sonst.} \end{cases}$$

Dann $\|x\|_\infty = 1$ und die k -te Komponente von Ax ist gleich

$$\sum_{j=1}^m a_{k,j} \frac{|a_{k,j}|}{a_{k,j}} = \sum_{j=1}^m |a_{k,j}|.$$

Die Operatornorm in (4.1) ist die *Zeilensummennorm*.

4.18 Definition. Sei $[a, b] \subseteq \mathbb{R}$ ein abgeschlossenes Intervall, sei X ein metrischer Raum, und sei $G \subseteq X$. Eine stetige Funktion $\gamma: [a, b] \rightarrow G$ ist ein *Weg* von $\gamma(a)$ nach $\gamma(b)$. Eine Menge G heißt *wegzusammenhängend*, wenn es zu je zwei Punkten $u, v \in G$ einen Weg von u nach v gibt.

4.19 Satz. Sei X ein metrischer Raum, sei A eine kompakte, wegzusammenhängende Teilmenge von X , und sei $f: A \rightarrow \mathbb{R}$ stetig. Dann ist $f(A)$ ein kompaktes Intervall.

Beweis. Seien $c := \min f(A)$ und $d := \max f(A)$. Dann $f(A) \subseteq [c, d]$. Wir zeigen $[c, d] \subseteq f(A)$. Seien dazu $u, v \in A$ mit $c = f(u)$ und $d = f(v)$, und sei $\gamma: [a, b] \rightarrow A$ ein Weg von u nach v . Dann ist die Funktion $g: [a, b] \rightarrow \mathbb{R}$, $g(t) = f(\gamma(t))$, stetig. Wegen Korollar 9.16 der Analysis I ist $J := g([a, b])$ ein Intervall. Wegen $c, d \in J$ folgt die Behauptung nun aus $[c, d] \subseteq J = f(\gamma([a, b])) \subseteq f(A)$. $\square$

5. Vollständige metrische Räume

5.1 Definition. Sei X ein metrischer Raum.

- (a) Eine Folge $(x_n)_{n \in \mathbb{N}}$ heißt *Cauchyfolge*, wenn es zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle $n, m \geq N$ gilt $d(x_n, x_m) < \epsilon$.
- (b) X heißt *vollständig*, wenn jede Cauchyfolge in X konvergiert.
- (c) Ein normierter Raum heißt *Banachraum*, wenn er vollständig ist.

5.2 Lemma. *Jede konvergente Folge in einem metrischen Raum ist eine Cauchyfolge.*

5.3 Satz. *Jeder endlich-dimensionale normierte Raum ist ein Banachraum.*

Beweis. Der Raum heie V , seine Norm werde mit $\|\cdot\|$ bezeichnet. Sei $(b_1, \dots, b_n)$ eine Basis von V . Dann ist

$$\Phi: (\mathbb{R}^n, \|\cdot\|_\infty) \rightarrow V, \quad (v_1, \dots, v_n) \mapsto \sum_{j=1}^n v_j b_j,$$

eine stetige lineare Bijektion, deren Inverse nach Satz 4.13 ebenfalls stetig ist. Wenn $(x_k)_{k \in \mathbb{N}}$ eine Cauchyfolge in V ist, dann ist $(\Phi^{-1}(x_k))_{k \in \mathbb{N}}$ eine Cauchyfolge im $\mathbb{R}^n$. Also sind alle Komponenten Cauchyfolgen in $\mathbb{R}$, konvergieren also. Nun wendet man wieder Φ an. $\square$

5.4 Satz. *Sei V ein endlich-dimensionaler Banachraum. Wir bezeichnen sowohl die Norm auf V als auch die zugehörige Operatornorm mit $\|\cdot\|$. Es sei $A: V \rightarrow V$ eine lineare Abbildung, für welche gilt*

$$\|\text{id} - A\| < 1.$$

Dann ist A invertierbar.

Bevor ich den Satz mit den Mitteln der Analysis beweise, zeige ich ihn zunächst mit den Mitteln der Linearen Algebra.

Wir nehmen zum Widerspruch an, dass A nicht invertierbar ist. Dann ist A nicht injektiv — hier geht die Endlichkeit der Dimension von V ein. Es gibt folglich ein $v \neq 0$ mit $Av = 0$. Für dieses v gilt $\|(\text{id} - A)v\| = \|v\|$, also $\|\text{id} - A\| \geq 1$ im Widerspruch zur Voraussetzung.

5. Vollständige metrische Räume

Beweis von Satz 5.4. Wir zeigen zunächst die Konvergenz von

$$B = \sum_{n=0}^{\infty} (\text{id} - A)^n$$

im Raum der $\text{End}_{\mathbb{R}}(V)$ der $\mathbb{R}$ -linearen Abbildungen von V in sich, den wir mit der Operatornorm versehen. Da auch der $\text{End}_{\mathbb{R}}(V)$ endliche Dimension hat, ist er vollständig.

Sei dazu $q = \|\text{id} - A\| < 1$. Wir zeigen, dass die Folge der Partialsummen von B eine Cauchyfolge ist. Sei dazu $\epsilon > 0$ vorgegeben. Wähle N mit $q^{N+1} < \epsilon(1 - q)$. Für $k, m \geq N$ mit $k < m$ gilt

$$\begin{aligned} \left\| \sum_{n=0}^m (\text{id} - A)^n - \sum_{n=0}^k (\text{id} - A)^n \right\| &\leq \sum_{n=k+1}^m \|(\text{id} - A)^n\| \leq \sum_{n=k+1}^m q^n \leq \sum_{n=k+1}^{\infty} q^n \\ &= \frac{q^{k+1}}{1 - q} \leq \frac{q^{N+1}}{1 - q} < \epsilon. \end{aligned}$$

Aus der Zwischenbehauptung folgt nun

$$(\text{id} - A)B = \lim_{k \rightarrow \infty} (\text{id} - A) \sum_{n=0}^k (\text{id} - A)^n = \lim_{k \rightarrow \infty} \sum_{n=0}^k (\text{id} - A)^{n+1} = B - \text{id}.$$

Subtrahiert man B auf beiden Seiten, so erhält man $-AB = -\text{id}$. Indem man die beiden Faktoren vertauscht, bekommt man schließlich auch noch $BA = \text{id}$. $\square$

6. Partielle Ableitungen

6.1 Definition. Es sei $U \subseteq \mathbb{R}^n$ offen, es sei $f: U \rightarrow \mathbb{R}$, und es sei $x = (x_1, x_2, \dots, x_n) \in U$. Für $i = 1, \dots, n$ setzen wir $U_i = \{t \in \mathbb{R} \mid (x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n) \in U\}$. Dann ist U_i eine offene Umgebung von x_i in $\mathbb{R}$. Man definiert $F_i: U_i \rightarrow \mathbb{R}$ durch

$$F_i(t) = f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n).$$

Wenn für jedes $i = 1, \dots, n$ die Funktion F_i in x_i differenzierbar ist, dann heißt f *partiell differenzierbar* in x . Man schreibt dann

$$\frac{\partial f}{\partial x_i}(x) = F'_i(x_i).$$

$\frac{\partial f}{\partial x_i}(x)$ heißt dann *i-te partielle Ableitung* von f in x . Die Funktion f heißt *partiell differenzierbar*, wenn sie in jedem Punkt von U partiell differenzierbar ist.

6.2 Bemerkung. (a) Man berechnet die i -te partielle Ableitung, indem man f als Funktion von x_i allein auffasst und die anderen Variablen als konstant ansieht.

(b) Für $n = 3$ schreibt man oft (x, y, z) anstelle von (x_1, x_2, x_3) . Dann schreibt man entsprechend $\frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$ usw.

6.3 Beispiel. (a) $f(x, y) = e^{xy^2}$. Dann $\frac{\partial f}{\partial x} = y^2 e^{xy^2}$ und $\frac{\partial f}{\partial y} = 2xy e^{xy^2}$.

(b) Betrachte

$$g: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{xy}{x^2 + y^2}, & \text{für } (x, y) \neq (0, 0), \\ 0, & \text{sonst.} \end{cases}$$

Die Funktion g ist überall partiell differenzierbar, obwohl sie in $(0, 0)$ unstetig ist. Den Graphen von g zeigt Abbildung 6.1.

6.4 Definition. Sei $U \subseteq \mathbb{R}^n$ offen und sei $f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix}: U \rightarrow \mathbb{R}^m$. Dann heißt f *partiell differenzierbar*, wenn alle f_j partiell differenzierbar sind. Für $x \in U$ ist

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \dots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}$$

die *Funktionalmatrix* oder *Jacobi-Matrix* von f an der Stelle x .

6. Partielle Ableitungen

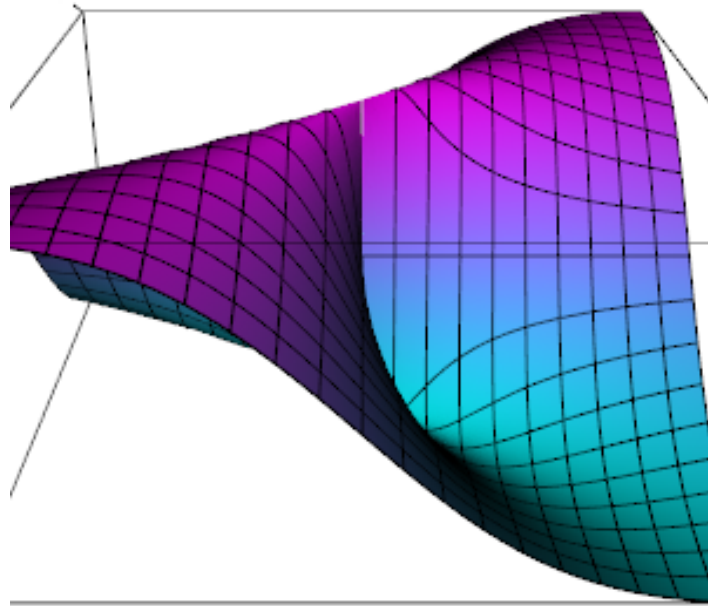


Abbildung 6.1.: Graph der Funktion g aus Beispiel 6.3

6.5 *Bemerkung.* (a) Wenn Funktionalmatrizen vorkommen, schreiben wir Vektoren als Spalten.

(b) Für skalarwertiges f ist

$$Df(x) = \left(\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \cdots \quad \frac{\partial f}{\partial x_n} \right)$$

ein Zeilenvektor. Sein Transponierter wird als *Gradient* von f bezeichnet, geschrieben

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}.$$

Das Zeichen ∇ wird als “Nabla” ausgesprochen. Für Forster ist der Gradient die Funktionalmatrix. Die hier verwendete Definition findet sich bei Kaballo, Königsberger und Jarre.

6.6 *Beispiel.* (a) Sei $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ gegeben durch

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \cos(x - y).$$

Dann

$$\nabla f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\sin(x-y) \\ \sin(x-y) \\ 0 \end{pmatrix}.$$

(b) Sei $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ gegeben durch

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + 3z \\ \sin(xz^2) \end{pmatrix}.$$

Dann

$$Df \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ z^2 \cos(xz^2) & 0 & 2xz \cos(xz^2) \end{pmatrix}.$$

6.7 Definition. Sei U offen in $\mathbb{R}^n$ und $f: U \rightarrow \mathbb{R}$ partiell differenzierbar. Wenn alle partiellen Ableitungen $\frac{\partial f}{\partial x_i}$, $i = 1, \dots, n$, wieder partiell differenzierbar sind, dann heißt f *zweimal partiell differenzierbar*. Man schreibt dann

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j} \frac{\partial f}{\partial x_i}.$$

Induktiv definiert man so für jedes $k \in \mathbb{N}$ die k -fachen partiellen Ableitungen. Die 0-te Ableitung von f ist f selbst.

6.8 Beispiel. Für $f(x, y) = e^{xy^2}$ hatten wir bereits gesehen, dass $\frac{\partial f}{\partial x} = y^2 e^{xy^2}$. Also

$$\frac{\partial^2 f}{\partial y \partial x} = 2ye^{xy^2} + 2xy^3 e^{xy^2}.$$

Durch partielle Ableitung von $\frac{\partial f}{\partial y}$ sieht man $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$. Wir werden diese Gleichheit später allgemein unter sehr schwachen Voraussetzungen an f zeigen.

6.9 Bezeichnung. Statt $x \mapsto \frac{\partial f}{\partial x_i}$ wird auch $D_i f$ geschrieben. Also beispielsweise

$$\frac{\partial^3 f}{\partial x_1 \partial x_2^2} = D_1 D_2^2 f.$$

7. Totale Ableitungen

7.1 Definition. Seien V und W endlich dimensionale, normierte Räume und sei $U \subset V$ offen. Eine Abbildung $f: U \rightarrow W$ heißt *(total) differenzierbar* in $x \in U$, wenn es eine lineare Abbildung $A: V \rightarrow W$ gibt, so dass für die durch

$$f(x + \xi) - f(x) = A(\xi) + \varphi(\xi) \quad (7.1)$$

bestimmte Abbildung φ gilt

$$\lim_{\xi \rightarrow 0} \frac{\|\varphi(\xi)\|}{\|\xi\|} = 0.$$

Die lineare Abbildung A ist die *Ableitung* von f in x . Wir schreiben $f'(x)$.

Eine Abbildung heißt *(total) differenzierbar*, wenn sie in jedem Punkt total differenzierbar ist.

7.2 Bemerkung. Im Fall $V = W = \mathbb{R}$ ist totale Differenzierbarkeit dasselbe wie Differenzierbarkeit im Sinne der Analysis I. In diesem Fall ist $A: \mathbb{R} \rightarrow \mathbb{R}$ die lineare Abbildung $\xi \mapsto f'(x)\xi$.

Beweis. Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ total differenzierbar und sei A wie in (7.1). Da A eine lineare Abbildung von $\mathbb{R} \rightarrow \mathbb{R}$ ist, ist A von der Form $\xi \mapsto a\xi$ für ein $a \in \mathbb{R}$. Es gilt

$$\frac{f(x+h) - f(x)}{h} = \frac{ah + \varphi(h)}{h} \rightarrow a, \quad h \rightarrow 0.$$

Also ist f in x differenzierbar im Sinne der Analysis I mit $f'(x) = a$.

Wenn f differenzierbar im Sinne der Analysis I ist und wir $A(\xi) = f'(x)\xi$ setzen, dann

$$\lim_{\xi \rightarrow 0} \frac{\|\varphi(\xi)\|}{\|\xi\|} = \lim_{\xi \rightarrow 0} \frac{\|f(x+\xi) - f(x) - f'(x)\xi\|}{\|\xi\|} = \lim_{\xi \rightarrow 0} \left\| \frac{f(x+\xi) - f(x)}{\xi} - f'(x) \right\| = 0.$$

□

7.3 Beispiel. Sei $A: V \rightarrow W$ linear. Dann ist A differenzierbar, und für jedes $x \in V$ gilt $A'(x) = A$. Insbesondere ist die Ableitung einer linearen Funktion konstant.

7.4 Satz. Seien V und W endlich dimensionale, normierte Räume, sei $U \subseteq V$ offen und sei $f: U \rightarrow W$ differenzierbar in $x \in U$. Dann ist f stetig in x .

7.5 Beispiel. Wir hatten in Beispiel 6.3 eine Funktion gesehen, die überall partiell differenzierbar, aber unstetig im Ursprung ist. Wegen Satz 7.4 ist diese Funktion nicht total differenzierbar.

7. Totale Ableitungen

7.6 Lemma. Sei V ein endlich dimensionaler, normierter Raum und sei $U \subseteq V$ offen. Für eine Abbildung $f: U \rightarrow \mathbb{R}^m$ seien $f_j: U \rightarrow \mathbb{R}$, $j = 1, \dots, m$, ihre Komponenten. Dann ist f genau dann total differenzierbar, wenn alle f_j total differenzierbar sind.

7.7 Definition. Sei $U \subseteq \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$. Wenn f k -mal partiell differenzierbar ist und alle partiellen Ableitungen der Ordnung $\leq k$ stetig sind, dann sagt man, f sei von der Klasse C^k . Wenn f für jedes $k \in \mathbb{N}$ von der Klasse C^k ist, so heißt f *glatt* oder von der Klasse C^∞ . Schließlich heißt f von der Klasse C^0 , wenn es stetig ist.

Eine Abbildung $g: U \rightarrow \mathbb{R}^m$ heißt von der Klasse C^k , wenn ihre Komponenten von der Klasse C^k sind.

7.8 Satz. Sei U offen in $\mathbb{R}^n$ und $f = (f_1, f_2, \dots, f_m): U \rightarrow \mathbb{R}^m$ von der Klasse C^1 . Dann ist f total differenzierbar und die Ableitung $f'(x)$ wird durch die Jacobi-Matrix $Df(x)$ gegeben.

Beweis. Wir brauchen den Beweis nur im Fall $m = 1$ zu führen. x und ξ seien gegeben. Zuerst definieren wir Zwischenpunkte

$$\begin{aligned} z^0 &= x \\ z^1 &= (x_1 + \xi_1, x_2, x_3, \dots, x_n) \\ z^2 &= (x_1 + \xi_1, x_2 + \xi_2, x_3, \dots, x_n) \\ &\dots \\ z^n &= x + \xi. \end{aligned}$$

Dann unterscheiden sich z^{i-1} und z^i nur in der i -ten Komponente. Man kann daher den Mittelwertsatz auf die i -te Partialfunktion anwenden und erhält Zwischenpunkte $\eta_i \in]0, \xi_i[$, so dass

$$f(z^i) - f(z^{i-1}) = \xi_i \frac{\partial f}{\partial x_i}(x_1 + \xi_1, \dots, x_{i-1} + \xi_{i-1}, x_i + \eta_i, x_{i+1}, \dots, x_n).$$

Also

$$\begin{aligned} f(x + \xi) - f(x) &= \sum_{i=1}^n \xi_i \frac{\partial f}{\partial x_i}(x_1 + \xi_1, \dots, x_{i-1} + \xi_{i-1}, x_i + \eta_i, x_{i+1}, \dots, x_n) \\ &= \nabla f(x) \cdot \xi + \sum_{i=1}^n \xi_i \left(\frac{\partial f}{\partial x_i}(x_1 + \xi_1, \dots, x_{i-1} + \xi_{i-1}, x_i + \eta_i, x_{i+1}, \dots, x_n) - \frac{\partial f}{\partial x_i}(x) \right). \end{aligned}$$

Also ist $\varphi(x) = \sum_{i=1}^n \xi_i \left(\frac{\partial f}{\partial x_i}(x_1 + \xi_1, \dots, x_{i-1} + \xi_{i-1}, x_i + \eta_i, x_{i+1}, \dots, x_n) - \frac{\partial f}{\partial x_i}(x) \right)$. Es folgt für $\xi \rightarrow 0$

$$\frac{\|\varphi(\xi)\|_\infty}{\|\xi\|_\infty} \leq \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(x_1 + \xi_1, \dots, x_{i-1} + \xi_{i-1}, x_i + \eta_i, x_{i+1}, \dots, x_n) - \frac{\partial f}{\partial x_i}(x) \right| \rightarrow 0.$$

Bei dem Grenzübergang wurde die Stetigkeitsvoraussetzung eingesetzt. □

Bemerkung. Mit diesem Satz und dem Satz aus der Analysis I, dass Summen, Produkte und Verkettungen von C^1 -Funktionen ebenfalls von der Klasse C^1 sind, sieht man bei vielen Funktionen sofort, dass sie total differenzierbar sind.

7.9 Satz (Kettenregel). *Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offen. Es seien $f: U \rightarrow V$ und $g: V \rightarrow \mathbb{R}^p$ Abbildungen, wobei f in $x \in U$ und g in $f(x)$ total differenzierbar ist. Dann ist $g \circ f$ in x total differenzierbar und es gilt*

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x).$$

Ausgedrückt in Jacobi-Matrizen hat diese Gleichung die Form

$$D(g \circ f)(x) = Dg(f(x)) \cdot Df(x).$$

Beweis. Wir führen in x bzw. $f(x)$ die Abbildungen φ bzw. ψ ein

$$\begin{aligned} f(x + \xi) - f(x) - f'(x)(\xi) &= \varphi(\xi) \\ g(f(x) + \eta) - g(f(x)) - g'(f(x))(\eta) &= \psi(\eta). \end{aligned}$$

Dann

$$\begin{aligned} g \circ f(x + \xi) - g \circ f(x) &= g(f(x) + f'(x)(\xi) + \varphi(\xi)) - g(f(x)) \\ &= g(f(x)) + g'(f(x))(f'(x)(\xi)) + g'(f(x))(\varphi(\xi)) \\ &\quad + \psi(f'(x)(\xi) + \varphi(\xi)) - g(f(x)) \\ &= g'(f(x))(f'(x)(\xi)) + \theta(\xi), \end{aligned}$$

wobei

$$\theta(\xi) = g'(f(x))(\varphi(\xi)) + \psi(f'(x)(\xi) + \varphi(\xi)).$$

Wir müssen $\lim_{\xi \rightarrow 0} \frac{\|\theta(\xi)\|}{\|\xi\|} = 0$ zeigen. Für den ersten Summanden von θ ist das klar. Zur Untersuchung des zweiten schreiben wir $\psi(\eta) = \|\eta\|\rho(\eta)$. Dann $\lim_{\eta \rightarrow 0} \rho(\eta) = 0$. Ferner gilt $\|\varphi(\xi)\| \leq \|\xi\|$ für alle hinreichend kleinen ξ . Also

$$\begin{aligned} \frac{\|\theta(\xi)\|}{\|\xi\|} &\leq \frac{\|f'(x)(\xi) + \varphi(\xi)\|}{\|\xi\|} \rho(f'(x)(\xi) + \varphi(\xi)) \\ &\leq (\|f'(x)\| + 1) \rho(f'(x)(\xi) + \varphi(\xi)) \rightarrow 0. \quad \square \end{aligned}$$

7.10 Korollar. *Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offen und seien $f: U \rightarrow V$ und $g: V \rightarrow \mathbb{R}^p$ von der Klasse C^k . Dann ist auch $g \circ f$ von der Klasse C^k .*

Beweis. Folgt mit vollständiger Induktion aus der Kettenregel. $\square$

7.11 Beispiel. Für $p = 1$, also $g \circ f: U \rightarrow \mathbb{R}$, erhält man

$$\frac{\partial(g \circ f)}{\partial x_i}(x) = \sum_{j=1}^m \frac{\partial g}{\partial y_j}(f(x)) \frac{\partial f_j}{\partial x_i}(x).$$

7. Totale Ableitungen

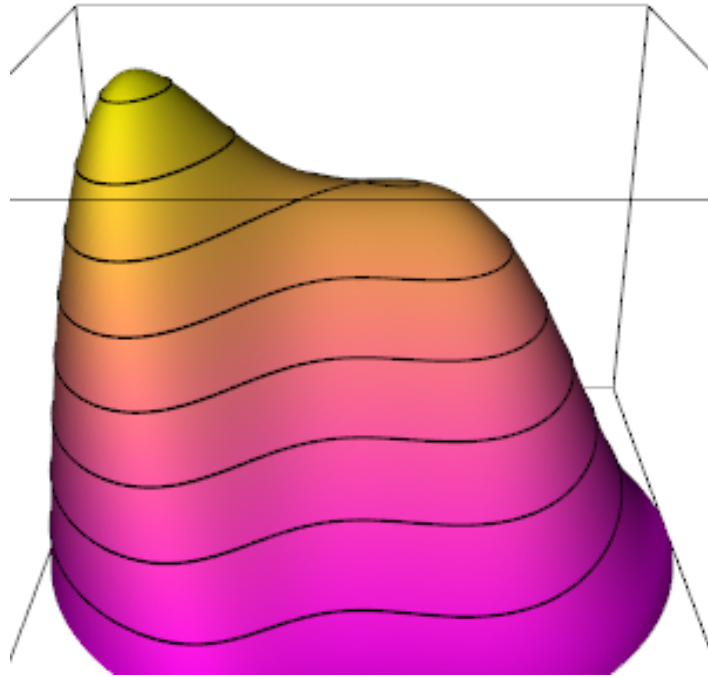


Abbildung 7.1.: Graph der Funktion

$$(x, y) \mapsto \left(4 - \frac{x^2 y^2}{20} + 3x^3 + 4x^2 - 4y + 3\right) (4 - x^2 - y^2)$$

7.12 Beispiel. Sei $V \subseteq \mathbb{R}^m$ offen, sei $g: V \rightarrow \mathbb{R}$ von der Klasse C^1 , sei $I \subseteq \mathbb{R}$ ein offenes Intervall, und sei $\gamma: I \rightarrow V$ ein C^1 -Weg, so dass $g \circ \gamma$ konstant ist. Dann gilt für jedes $t \in I$

$$\gamma'(t) \cdot \nabla g(\gamma(t)) = 0.$$

Der Weg γ verläuft entlang einer Höhenlinie von g . Im Fall $m = 2$ zeigt das Beispiel, dass die Höhenlinien senkrecht zum Gradienten verlaufen. Abbildung 7.1 zeigt ein Beispiel.

7.13 Definition. Sei $U \subseteq \mathbb{R}^n$ offen, sei $f: U \rightarrow \mathbb{R}$ eine Funktion, sei $x \in U$ und sei $v \in \mathbb{R}^n$. Dann ist $U_v = \{t \in \mathbb{R} \mid x + tv \in U\}$ eine offene Umgebung der 0 in $\mathbb{R}$. Definiere $F_v: U_v \rightarrow \mathbb{R}$ durch $F_v(t) = f(x + tv)$. Wenn F_v in 0 differenzierbar ist, dann bezeichnet man $F_v'(0)$ als *Richtungsableitung* von f an der Stelle x in Richtung v . Wir schreiben $D_v f(x)$ oder $\frac{\partial f}{\partial v}(x)$ dafür.

7.14 Satz. Sei $U \subseteq \mathbb{R}^n$ offen, sei $f: U \rightarrow \mathbb{R}$ von der Klasse C^1 , sei $x \in U$ und sei $v \in \mathbb{R}^n$. Dann existiert die Richtungsableitung $D_v f(x)$ und es gilt

$$D_v f(x) = \nabla f(x) \cdot v.$$

Beweis. Ergibt sich aus der Kettenregel, wenn man $g(t) = x + tv$ setzt. □

Wenn der Gradient nicht verschwindet, dann gibt er die Richtung des stärksten Anstiegs an. In diesem Fall ist die Richtung des stärksten Anstiegs senkrecht zu den Höhenlinien.

Zum Schluss noch ein Beispiel, bei dem die totale Differenzierbarkeit direkt mit der Definition gezeigt wird.

7.15 Beispiel. Es sei V der normierte Raum der $n \times n$ -Matrizen. Wegen der Stetigkeit der Determinanten ist die Menge $G = \{M \in V | \det(M) \neq 0\}$ der invertierbaren Matrizen offen. Wir zeigen die totale Differenzierbarkeit der Abbildung

$$f: G \rightarrow V, \quad f(M) = M^{-1}.$$

Es seien $M \in G$ und $H \in V$ derart, dass $M + H \in G$ und es sei $E \in G$ die Einheitsmatrix. Dann gilt

$$(M + H)^{-1} - M^{-1} = (M(E + M^{-1}H))^{-1} - M^{-1} = (E + M^{-1}H)^{-1}M^{-1} - M^{-1}. \quad (7.2)$$

Es sei nun $\|H\| < \frac{1}{\|M^{-1}\|}$. Wir wenden Satz 5.4 an auf $B = E + M^{-1}H$. Dann $\|E - B\| = \|M^{-1}H\| < 1$ und daher

$$(E + M^{-1}H)^{-1} = \sum_{j=0}^{\infty} (E - (E + M^{-1}H))^j = \sum_{j=0}^{\infty} (-1)^j (M^{-1}H)^j.$$

Setzt man dies in Gleichung (7.2) ein, so erhält man

$$(M + H)^{-1} - M^{-1} = -M^{-1}HM^{-1} + \sum_{j=2}^{\infty} (-1)^j (M^{-1}H)^j M^{-1}.$$

Der erste Term auf der rechten Seite ist $f'(M)(H)$, der zweite $\varphi(H)$. Mit der nächsten Abschätzung ist der Beweis beendet.

$$\frac{\|\varphi(H)\|}{\|H\|} \leq \frac{1}{\|H\|} \|M^{-1}\| \sum_{j=2}^{\infty} \|M^{-1}H\|^j = \frac{1}{\|H\|} \|M^{-1}\| \frac{\|M^{-1}H\|^2}{1 - \|M^{-1}H\|} \leq \frac{\|M^{-1}\|^3 \|H\|}{1 - \|M^{-1}H\|} \rightarrow 0.$$

8. Die Hesse-Matrix

8.1 Definition. Sei $U \subseteq \mathbb{R}^n$ offen und sei $f: U \rightarrow \mathbb{R}$ zweimal partiell differenzierbar. Die Matrix

$$H_f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

heißt *Hesse-Matrix* von f an der Stelle x .

8.2 Satz (H. A. Schwarz). Wenn f von der Klasse C^2 ist, dann ist $H_f(x)$ für jedes $x \in U$ symmetrisch, d. h.

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$$

für alle $i, j \in \{1, \dots, n\}$.

Beweis. $i = 1$ und $j = 2$ genügt. Es genügt auch, den Beweis nur im Nullpunkt zu führen. Für festes x_2 definiere

$$F_{x_2}(x_1) = f(x_1, x_2) - f(x_1, 0).$$

Aus dem eindimensionalen Mittelwertsatz erhalten wir die Existenz eines ξ zwischen 0 und x_1 mit

$$F_{x_2}(x_1) - F_{x_2}(0) = x_1 F'_{x_2}(\xi) = x_1 \left(\frac{\partial f}{\partial x_1}(\xi, x_2) - \frac{\partial f}{\partial x_1}(\xi, 0) \right).$$

In der Klammer können wir nochmal den Mittelwertsatz anwenden und erhalten ein η zwischen 0 und x_2 , so dass

$$F_{x_2}(x_1) - F_{x_2}(0) = x_1 x_2 \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi, \eta).$$

Es gilt

$$F_{x_2}(x_1) - F_{x_2}(0) = f(x_1, x_2) - f(x_1, 0) - f(0, x_2) + f(0, 0) = G_{x_1}(x_2) - G_{x_1}(0),$$

8. Die Hesse-Matrix

wobei $G_{x_1}(x_2) = f(x_1, x_2) - f(0, x_2)$. Dasselbe Argument wie oben liefert z zwischen 0 und x_1 und w zwischen 0 und x_2 mit

$$G_{x_1}(x_2) - G_{x_1}(0) = x_1 x_2 \frac{\partial^2 f}{\partial x_1 \partial x_2}(z, w).$$

Für $x_1, x_2 \neq 0$ folgt

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi, \eta) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(z, w).$$

Aus $(x_1, x_2) \rightarrow (0, 0)$ folgen $(\xi, \eta) \rightarrow (0, 0)$ und $(z, w) \rightarrow (0, 0)$. Wegen der Stetigkeit der zweiten Ableitungen folgt schließlich die Behauptung. $\square$

8.3 Beispiel. Für $f(x, y, z) = xz + e^{x+2y}$ gilt

$$H_f(x, y, z) = \begin{pmatrix} e^{x+2y} & 2e^{x+2y} & 1 \\ 2e^{x+2y} & 4e^{x+2y} & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

8.4 Korollar. Wenn f von der Klasse C^k ist, dann kommt es bei den Ableitungen von der Ordnung $\leq k$ nicht auf die Reihenfolge an.

Wir werden in den Übungen eine Funktion kennenlernen, die zweimal partiell differenzierbar ist und deren Hesse-Matrix nicht symmetrisch ist.

9. Mittelwertsatz und Taylor-Formel

Der Mittelwertsatz der Analysis I lautet: Sei $f: [a, b] \rightarrow \mathbb{R}$ stetig und differenzierbar auf $]a, b[$, dann gibt es ein $\xi \in]a, b[$ mit

$$f(b) - f(a) = f'(\xi)(b - a).$$

In dieser Form kann der Mittelwertsatz nicht auf vektorwertige Abbildungen verallgemeinert werden.

9.1 Beispiel. $f: \mathbb{R} \rightarrow \mathbb{R}^2$ sei definiert durch $f(x) = \begin{pmatrix} \cos x \\ \sin x \end{pmatrix}$. Dann $f(0) = f(2\pi)$, aber es gibt kein ξ mit $Df(\xi) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

9.2 Lemma. Sei $U \subseteq \mathbb{R}^n$ offen und sei $f: U \rightarrow \mathbb{R}^m$ von der Klasse C^1 . Dann hängt die Operatornorm $\|Df(x)\|$ stetig von x ab.

Beweis. Wir versehen den Raum $\mathbb{R}^{m \times n}$ mit der Supremumsnorm des $\mathbb{R}^{mn}$. In dieser Norm ist die Aussage klar. Da auf dem $\mathbb{R}^{mn}$ je zwei Normen äquivalent sind, folgt die Behauptung. $\square$

9.3 Notation. Die *Strecke* zwischen x und y ist die Menge $[x, y] = \{x + t(y - x) \mid 0 \leq t \leq 1\}$. Sie ist kompakt.

9.4 Theorem (Mittelwertsatz für skalarwertige Funktionen). *Es sei $U \subseteq \mathbb{R}^n$ offen und es sei $f: U \rightarrow \mathbb{R}$ total differenzierbar. Ferner seien $x, y \in U$ derart, dass $[x, y] \subset U$. Dann gibt es $\xi \in [x, y] \setminus \{x, y\}$, so dass*

$$f(y) - f(x) = f'(\xi)(y - x).$$

Beweis. Wir wenden den Mittelwertsatz der Analysis I an auf $\varphi(t) = f(x + t(y - x))$ und erhalten $\tau \in]0, 1[$ mit

$$f(y) - f(x) = \varphi(1) - \varphi(0) = \varphi'(\tau) = f'(x + \tau(y - x))(y - x).$$

Nun setzen wir $\xi = x + \tau(y - x)$. $\square$

9.5 Theorem (Mittelwertsatz für vektorwertige Abbildungen). *Es sei $U \subseteq \mathbb{R}^n$ offen, es seien $x, y \in U$, so dass $[x, y] \subseteq U$. Für jede C^1 -Abbildung $f: U \rightarrow \mathbb{R}^m$ gilt*

$$\|f(y) - f(x)\|_\infty \leq \|x - y\| \max_{\xi \in [x, y]} \|f'(\xi)\|.$$

Dabei ist $\|x - y\|$ eine beliebige Norm auf dem $\mathbb{R}^n$ und $\|f'(\xi)\|$ die Operatornorm für lineare Abbildungen von $(\mathbb{R}^n, \|\cdot\|)$ nach $(\mathbb{R}^m, \|\cdot\|_\infty)$.

9. Mittelwertsatz und Taylor-Formel

Beweis. Es sei j ein Index mit $\|f(y) - f(x)\|_\infty = |f_j(y) - f_j(x)|$. Dann gibt es ein $\xi \in [x, y] \setminus \{x, y\}$

$$\|f(y) - f(x)\|_\infty = |f_j(y) - f_j(x)| = \|f'_j(\xi)(x - y)\| \leq \|f'(\xi)\| \|x - y\|,$$

wobei $\|f'_j(\xi)\|$ die Operatornorm für lineare Abbildungen von $(\mathbb{R}^n, \|\cdot\|)$ nach $\mathbb{R}$ ist. $\square$

9.6 Theorem. Es sei $U \subseteq \mathbb{R}^n$ offen, es seien $x, y \in U$ und es gebe einen Weg $\gamma: [a, b] \rightarrow U$ mit $\gamma(a) = x$ und $\gamma(b) = y$, dessen Einschränkung auf $]a, b[$ von der Klasse C^1 ist. Für jede C^1 -Abbildung

$$f: U \rightarrow \mathbb{R}^m, \quad f(w) = \begin{pmatrix} f_1(w) \\ \vdots \\ f_m(w) \end{pmatrix},$$

gilt dann für $j = 1, \dots, m$

$$f_j(y) - f_j(x) = \int_a^b f'_j(\gamma(t)) \gamma'(t) dt.$$

Beweis. Wir wenden den Hauptsatz der Differential- und Integralrechnung aus Analysis I an auf $\varphi = f_j \circ \gamma$ und erhalten

$$f_j(y) - f_j(x) = \varphi(b) - \varphi(a) = \int_a^b \varphi'(t) dt = \int_a^b f'_j(\gamma(t)) \gamma'(t) dt. \quad \square$$

9.7 Beispiel. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gegeben als

$$f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos(x_1) + \sin(x_2) \\ \sin(x_1) + \cos(x_2) \end{pmatrix}.$$

Für $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ wollen wir $\|f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) - f\left(\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right)\|_\infty$ durch $\left\|\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right\|_\infty$ abschätzen. Wir berechnen die Funktionalmatrix

$$D(f) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -\sin x_1 & \cos x_2 \\ \cos x_1 & -\sin x_2 \end{pmatrix}.$$

Ihre Operatornorm ist die Zeilsummennorm

$$\|D(f) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\| = \max\{|\sin x_1| + |\cos x_2|, |\cos x_1| + |\sin x_2|\} \leq 2.$$

Wir haben gezeigt

$$\left\| f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - f \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\|_\infty \leq 2 \left\| \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix} \right\|_\infty.$$

9.8 Satz. $U \subseteq \mathbb{R}^n$ sei offen und wegzusammenhängend. Sei $f: U \rightarrow \mathbb{R}^m$ stetig und partiell differenzierbar mit $D(f)(x) = 0$ für alle $x \in U$. Dann ist f konstant.

Beweis. Unter diesen Voraussetzungen ist f von der Klasse C^1 . Angenommen, es gebe $u, v \in U$ mit $f(u) \neq f(v)$. Da U wegzusammenhängend ist, gibt es einen Weg $\gamma: [a, b] \rightarrow U$ von u nach v . Sei $t_0 := \inf\{t \in [a, b] \mid f(\gamma(t)) = f(v)\}$. Wegen der Stetigkeit von f und γ gelten $f(\gamma(t_0)) = f(v)$ und $t_0 > a$. Da U offen ist, existiert $\epsilon > 0$ mit $B_\epsilon(\gamma(t_0)) \subseteq U$. (Welche Norm auf dem $\mathbb{R}^n$ gewählt wurde, ist dabei unerheblich.) Sei $t \in [a, t_0[$ so gewählt, dass $\gamma(t) \in B_\epsilon(\gamma(t_0))$. Dann liegt auch die Strecke von $\gamma(t)$ nach $\gamma(t_0)$ in $B_\epsilon(\gamma(t_0))$. Der Mittelwertsatz impliziert nun den Widerspruch $f(\gamma(t)) = f(\gamma(t_0))$. $\square$

Da es wegen des Satzes von Schwarz bei C^k -Abbildungen auf die Differentiationsreihenfolge nicht ankommt, ist die folgende Schreibweise sinnvoll:

9.9 Notation. Sei $U \subseteq \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ eine Abbildung.

(a) Für $\alpha \in \mathbb{N}_0^n$ setzen wir

$$|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n, \\ \alpha! = \alpha_1! \cdot \alpha_2! \cdot \dots \cdot \alpha_n!.$$

Ein n -Tupel mit dieser Interpretation heißt *Multiindex*.

(b) Für einen Multiindex α und f von der Klasse $C^{|\alpha|}$ setzen wir

$$D^\alpha f = D_1^{\alpha_1} D_2^{\alpha_2} \dots D_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}.$$

Dabei gilt die Vereinbarung $D_i^0 f = f$.

(c) Für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ setzen wir $x^\alpha = x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \dots \cdot x_n^{\alpha_n}$.

9.10 Theorem (Taylor-Formel). Sei $U \subseteq \mathbb{R}^n$ offen und sei $f: U \rightarrow \mathbb{R}$ von der Klasse C^{k+1} . Seien $x, \xi \in \mathbb{R}^n$, so dass die Strecke $[x, x + \xi]$ in U liegt. Dann

$$f(x + \xi) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha + (k+1) \sum_{|\alpha|=k+1} \int_0^1 (1-t)^k \frac{D^\alpha f(x + t\xi)}{\alpha!} \xi^\alpha dt.$$

Beweis. Weil U offen ist, existiert ein offenes Intervall I mit $[0, 1] \subseteq I$, so dass $\{x + t\xi \mid t \in I\} \subseteq U$. Definiere $g: I \rightarrow \mathbb{R}$ durch $g(t) = f(x + t\xi)$. Dann ist g von der Klasse C^{k+1} . Die Taylor-Formel aus der Analysis I besagt

$$f(x + \xi) = g(1) = \sum_{j=0}^k \frac{g^{(j)}(0)}{j!} + \frac{1}{k!} \int_0^1 (1-t)^k g^{(k+1)}(t) dt.$$

9. Mittelwertsatz und Taylor-Formel

Wir müssen die benötigten Ableitungen von g bestimmen. Ausgehend von Satz 7.14 liefert eine einfache vollständige Induktion

$$g^{(j)}(t) = \sum_{i_1=1}^n \cdots \sum_{i_j=1}^n D_{i_1} \dots D_{i_j} f(x + t\xi) \xi_{i_1} \cdots \xi_{i_j}.$$

Definiert man den Multiindex α durch $\alpha_1 = |\{l \mid i_l = 1\}|$ usw., dann $D_{i_1} \dots D_{i_j} f(x + t\xi) \xi_{i_1} \cdots \xi_{i_j} = D^\alpha f(x + t\xi) \xi^\alpha$.

Wie viele Tupel $(i_1, \dots, i_j)$ führen zu diesem α ? Es gibt $\binom{j}{\alpha_1}$ Möglichkeiten, die Indices v mit $i_v = 1$ auszuwählen. Danach gibt es $\binom{j-\alpha_1}{\alpha_2}$ Möglichkeiten, unter den verbleibenden Indices v diejenigen mit $i_v = 2$ auszuwählen. Insgesamt ist die Anzahl der Möglichkeiten also

$$\begin{aligned} & \binom{j}{\alpha_1} \binom{j-\alpha_1}{\alpha_2} \cdots \binom{j-\alpha_1-\cdots-\alpha_{n-1}}{\alpha_n} \\ &= \frac{j!}{\alpha_1!(j-\alpha_1)!} \cdot \frac{(j-\alpha_1)!}{\alpha_2!(j-\alpha_1-\alpha_2)!} \cdots \frac{(j-\alpha_1-\cdots-\alpha_{n-1})!}{\alpha_n!0!} = \frac{j!}{\alpha!}. \end{aligned}$$

Also

$$g^{(j)}(t) = \sum_{|\alpha|=j} \frac{j!}{\alpha!} D^\alpha f(x + t\xi) \xi^\alpha.$$

Setzt man dies in die Formel aus der Analysis I ein, so erhält man die Behauptung. $\square$

Beispiel. Wir wollen die Funktion $f: x \mapsto e^{x_1} \cos(x_1 + x_2)$ im Ursprung in ihr Taylor-Polynom der Ordnung 2 entwickeln. Es gelten

$$\nabla f(x) = \begin{pmatrix} (-\sin(x_1 + x_2) + \cos(x_1 + x_2)) e^{x_1} \\ -e^{x_1} \sin(x_1 + x_2) \end{pmatrix}$$

und

$$H_f(x) = \begin{pmatrix} -2e^{x_1} \sin(x_1 + x_2) & -e^{x_1} \sin(x_1 + x_2) - e^{x_1} \cos(x_1 + x_2) \\ -e^{x_1} \sin(x_1 + x_2) - e^{x_1} \cos(x_1 + x_2) & -e^{x_1} \cos(x_1 + x_2) \end{pmatrix}.$$

Für die Taylor-Formel benötigen wir die Werte im Ursprung

$$\nabla f(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad H_f(0) = \begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix}.$$

Also

$$f(\xi) = 1 + \xi_1 - \xi_1 \xi_2 - \frac{1}{2} \xi_2^2 + R_3(0).$$

9.11 Notation. Es seien $f, g: B_r(a) \rightarrow \mathbb{R}$ zwei Funktionen. Wir schreiben $f = O(g)$, $x \rightarrow a$, wenn es ein $\epsilon \in]0, r[$ und ein $C > 0$ gibt, so dass $|f(x)| \leq Cg(x)$ für alle $x \in B_\epsilon(a)$. Wir schreiben $f = o(g)$, $x \rightarrow a$, wenn $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0$.

9.12 Korollar. Sei $U \subseteq \mathbb{R}^n$ offen und sei $f: U \rightarrow \mathbb{R}$ von der Klasse C^{k+1} und sei $x \in U$. Für

$$P_k(\xi) := \sum_{|\alpha| \leq k} \frac{D^\alpha f(x)}{\alpha!} (\xi - x)^\alpha \quad (9.1)$$

gilt

$$f(\xi) - P_k(\xi) = O(\|x - \xi\|^{k+1}), \quad \xi \rightarrow x. \quad (9.2)$$

9.13 Definition. Das Polynom P_k aus (9.1) ist das k -te *Taylor-Polynom* von f in x .

9.14 Satz. Das *Taylor-Polynom* ist das einzige Polynom, welches (9.2) erfüllt.

Beweis. Es seien Polynome P_k und Q_k vom Grad höchstens k gegeben mit $P_k(\xi) - Q_k(\xi) = O(\xi^{k+1})$. Wir müssen zeigen, dass $R := P_k - Q_k$ verschwindet. Wir nehmen zum Widerspruch $R \neq 0$ und schreiben

$$R(\xi) = \sum_{j=k_0}^k \sum_{|\alpha|=j} b_\alpha \xi^\alpha,$$

wobei wir annehmen, dass $\eta_0 := \sum_{|\alpha|=k_0} b_\alpha \xi_0^\alpha \neq 0$ für ein ξ_0 mit $\|\xi_0\| = 1$. Es sei $C_1 = \sum_{j=k_0+1}^k |b_\alpha|$. Dann gilt für t mit $0 < t < \min\left(1, \frac{|\eta_0|}{2C_1}\right)$

$$|R(t\xi_0)| = \left| \sum_{j=k_0}^k t^j \sum_{|\alpha|=j} b_\alpha \xi_0^\alpha \right| \geq t^{k_0} |\eta_0| - \sum_{j=k_0+1}^k t^j \sum |b_\alpha| \geq t^{k_0} (|\eta_0| - C_1 t) \geq t^{k_0} \frac{|\eta_0|}{2}.$$

Aus der Voraussetzung ergibt sich jetzt der Widerspruch $k_0 \geq k + 1$. $\square$

9.15 Definition. Sei $A = (a_{j,k})_{\substack{j=1,\dots,n \\ k=1,\dots,m}} \in \mathbb{R}^{n \times m}$. Die *Transponierte* von A ist die Matrix $A^T := (a_{j,k})_{\substack{k=1,\dots,m \\ j=1,\dots,n}} \in \mathbb{R}^{m \times n}$.

9.16 Korollar. Sei $U \subseteq \mathbb{R}^n$ offen, sei $f: U \rightarrow \mathbb{R}$ von der Klasse C^3 , sei $x \in U$ und sei $\bar{B}_r(x) \subseteq U$, wobei die Kugel $\bar{B}_r(x)$ bezüglich einer beliebigen Norm $\|\cdot\|$ gebildet wird. Für ξ mit $\|\xi\| < r$ ist $R_3(\xi)$ bestimmt durch

$$f(x + \xi) = f(x) + \langle \nabla f(x), \xi \rangle + \frac{1}{2} \xi^T H_f(x) \xi + R_3(\xi).$$

Dann gibt es $M \geq 0$, so dass für alle ξ mit $\|\xi\| \leq r$ gilt $|R_3(\xi)| \leq M \|\xi\|^3$.

9.17 Beispiel. Wir entwickeln $f(x, y) = e^{-x^2-y^2}$ im Ursprung bis zur zweiten Ordnung. Dann $f(0) = 1$ und

$$\nabla f(x, y) = \begin{pmatrix} -2xe^{-x^2-y^2} \\ -2ye^{-x^2-y^2} \end{pmatrix},$$

9. Mittelwertsatz und Taylor-Formel

also $\nabla f(0,0) = 0$. Ferner

$$H_f(x, y) = \begin{pmatrix} (-2 + 4x^2)e^{-x^2-y^2} & 4xye^{-x^2-y^2} \\ 4xye^{-x^2-y^2} & (-2 + 4y^2)e^{-x^2-y^2} \end{pmatrix},$$

also

$$H_f(0,0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Wir erhalten

$$\begin{aligned} f(x, y) &= 1 + \frac{1}{2} \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + O(\|(x, y)\|^3) \\ &= 1 - x^2 - y^2 + O(\|(x, y)\|^3). \end{aligned}$$

10. Eigenwerte

Ich wiederhole zentrale Begriffe der Linearen Algebra. Ich schreiben das Skalarprodukt $x \cdot y$ gelegentlich auch als $\langle x, y \rangle$.

10.1 Definition. Sei $A \in \mathbb{R}^{n \times n}$. Ein Vektor $v \in \mathbb{R}^n \setminus \{0\}$ ist ein *Eigenvektor* von A , wenn es ein $\lambda \in \mathbb{R}$ gibt, so dass $Av = \lambda v$. In diesem Fall ist λ ein *Eigenwert* von A .

10.2 Bemerkung. Für $v, w \in \mathbb{R}^n$ gilt $\langle v, w \rangle = v^T w$. Daraus ergibt sich sofort die *Wälzformel*

$$\langle Av, w \rangle = \langle v, A^T w \rangle.$$

10.3 Definition. Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt *symmetrisch*, wenn $A = A^T$.

10.4 Theorem (Hauptachsentransformation). *Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Dann sind alle Eigenwerte reell und es gibt eine Orthonormalbasis des $\mathbb{R}^n$, die aus Eigenvektoren von A besteht.*

10.5 Definition. Sei A eine reelle, symmetrische $n \times n$ -Matrix.

- (a) A heißt *positiv definit*, wenn $\langle x, Ax \rangle > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$.
- (b) A heißt *negativ definit*, wenn $\langle x, Ax \rangle < 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$.
- (c) A heißt *positiv semidefinit*, wenn $\langle x, Ax \rangle \geq 0$ für alle $x \in \mathbb{R}^n$.
- (d) A heißt *negativ semidefinit*, wenn $\langle x, Ax \rangle \leq 0$ für alle $x \in \mathbb{R}^n$.
- (e) A heißt *indefinit*, wenn es ein $x \in \mathbb{R}^n$ mit $\langle x, Ax \rangle > 0$ und ein $y \in \mathbb{R}^n$ mit $\langle y, Ay \rangle < 0$ gibt.

Man beachte, dass $\langle x, Ax \rangle = x^T Ax$. Abbildung 10.1 zeigt die Funktion $(x, y) \mapsto \langle x, Ax \rangle$ für symmetrische Matrizen A mit unterschiedlichem Definitheitsverhalten.

10.6 Lemma. *Sei A eine reelle, symmetrische $n \times n$ -Matrix.*

- (a) *A ist genau dann positiv definit, wenn alle Eigenwerte von A positiv sind (also echt größer als 0).*
- (b) *A ist genau dann positiv semidefinit, wenn alle Eigenwerte von A positiv oder 0 sind.*

10. Eigenwerte

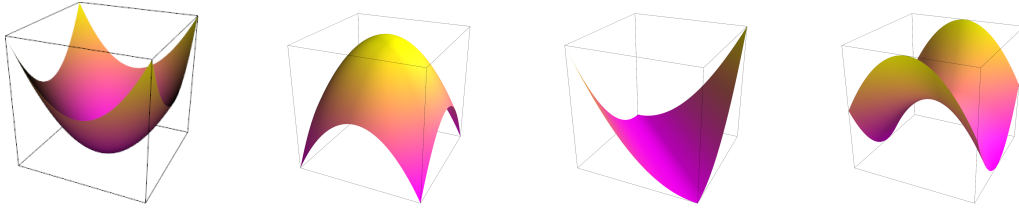


Abbildung 10.1.: Definitheitsverhalten: positiv, negativ, positiv semidefinit, indefinit

- (c) *A ist genau dann negativ definit, wenn alle Eigenwerte von A negativ sind.*
- (d) *A ist genau dann negativ semidefinit, wenn alle Eigenwerte von A negativ oder 0 sind.*
- (e) *A ist genau dann indefinit, wenn A sowohl positive als auch negative Eigenwerte besitzt.*

10.7 Satz (Hurwitz-Kriterium). Sei $A = (a_{i,j})_{1 \leq i,j \leq n}$ eine reelle, symmetrische $n \times n$ -Matrix. Sei

$$\Delta_k = \det \begin{pmatrix} a_{1,1} & \cdots & a_{1,k} \\ \vdots & & \vdots \\ a_{k,1} & \cdots & a_{k,k} \end{pmatrix}$$

die k -te Unterdeterminante.

- (a) *A ist genau dann positiv definit, wenn $\Delta_k > 0$ für alle k .*
- (b) *A ist genau dann negativ definit, wenn $(-1)^k \Delta_k > 0$ für alle k .*
- (c) *Wenn es ein k mit $\Delta_{2k} < 0$ gibt, dann ist A indefinit.*

Beweis. Wir zeigen (a). Die $k \times k$ -Matrix aus der Behauptung nennen wir A_k .

Sei zuerst A positiv definit. Wir nehmen zum Widerspruch die Existenz eines k mit $\Delta_k \leq 0$ an. Dann besitzt A_k einen Eigenwert $\lambda \leq 0$. Sei $0 \neq w \in \mathbb{R}^k$ ein Eigenvektor von A_k zum Eigenwert λ . Wir setzen

$$v = \begin{pmatrix} w_1 \\ \vdots \\ w_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n.$$

Dann $\langle v, Av \rangle = \langle w, A_k w \rangle = \lambda \langle w, w \rangle \leq 0$ im Widerspruch zur positiven Definitheit von A.

Wir zeigen nun die Rückrichtung mit vollständiger Induktion. Für $n = 1$ ist nichts zu zeigen. Die Behauptung sei für $n - 1$ richtig. Wir nehmen zum Widerspruch an, dass es eine symmetrische $n \times n$ -Matrix A gibt, die nicht positiv definit ist, für die aber $\Delta_k > 0$ für $k = 1, \dots, n$. Da $\det(A) > 0$, besitzt A dann zwei negative Eigenwerte λ_1 und λ_2 (welche nicht unbedingt verschieden sein müssen). Seien b_1 und b_2 orthogonale Eigenvektoren der Norm 1 zu λ_1 bzw. λ_2 . Mit LH bezeichnen wir die lineare Hülle. Der Raum

$$E = \text{LH}(\{b_1, b_2\}) \cap (\mathbb{R}^{n-1} \times \{(0)\})$$

besitzt mindestens die Dimension 1. Es sei $0 \neq x = \beta_1 b_1 + \beta_2 b_2 \in E$. Dann ist x von der Form

$$x = \begin{pmatrix} x_1, \\ \vdots \\ x_{n-1} \\ 0 \end{pmatrix}$$

und wir setzen

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}.$$

Damit erhalten wir

$$\langle X, A_{n-1}X \rangle = \langle x, Ax \rangle = \langle \beta_1 b_1 + \beta_2 b_2, \lambda_1 \beta_1 b_1 + \lambda_2 \beta_2 b_2 \rangle = \lambda_1 \beta_1^2 + \lambda_2 \beta_2^2 < 0.$$

Dies widerspricht der Induktionsvoraussetzung, denn die Unterdeterminanten von A_{n-1} sind alle positiv.

(b) folgt durch Anwendung von (a) auf $-A$.

(c) folgt ähnlich wie die Rückrichtung von (a). □

10.8 Beispiel. Für $a \in \mathbb{R}$ betrachte

$$A = \begin{pmatrix} -1 & -1 & 1 \\ -1 & -2 & 0 \\ 1 & 0 & a \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 2 & 0 \\ 1 & 0 & -2 \end{pmatrix}.$$

Für A gelten $\Delta_1 = -1$, $\Delta_2 = 1$ und $\Delta_3 = \det(A) = a + 2$. Also ist A negativ definit für $a < -2$. Wenn man die Unterdeterminanten von unten anfängt, sieht man, dass A indefinit ist, wenn $a > 0$.

Für B gilt $\Delta_2 = -2$. Daher ist B indefinit.

11. Extremwerte und kritische Stellen

11.1 Definition. Sei X ein metrischer Raum, $f: X \rightarrow \mathbb{R}$ eine Funktion und $x_0 \in X$. f besitzt in x_0 ein *lokales Maximum*, wenn es eine Umgebung U von x_0 gibt mit

$$f(x_0) \geq f(x) \quad \text{für alle } x \in U.$$

f besitzt in x_0 ein *striktes lokales Maximum*, wenn es eine Umgebung U von x_0 gibt mit

$$f(x_0) > f(x) \quad \text{für alle } x \in U \setminus \{x_0\}.$$

Entsprechend definiert man (strikte) lokale Minima. f besitzt in x_0 ein (striktes) lokales Extremum, wenn f in x_0 ein (striktes) lokales Maximum oder ein (striktes) lokales Minimum besitzt.

11.2 Satz. Sei $U \subseteq \mathbb{R}^n$ offen und sei $f: U \rightarrow \mathbb{R}$ partiell differenzierbar. Wenn f in x ein lokales Extremum besitzt, dann

$$\nabla f(x) = 0.$$

Beweis. O.E. besitze f in x ein lokales Maximum. Wir schätzen die erste partielle Ableitung ab. Für $h > 0$ gilt

$$\frac{f(x_1 + h, x_2, \dots, x_n) - f(x)}{h} \leq 0,$$

also $\frac{\partial f}{\partial x_1}(x) \leq 0$. Für $h < 0$ gilt aber

$$\frac{f(x_1 + h, x_2, \dots, x_n) - f(x)}{h} \geq 0,$$

also $\frac{\partial f}{\partial x_1}(x) \geq 0$. Insgesamt folgt $\frac{\partial f}{\partial x_1}(x) = 0$. Genauso können wir für die anderen partiellen Ableitungen argumentieren. $\square$

11.3 Definition. Sei $U \subseteq \mathbb{R}^n$ offen und sei $f: U \rightarrow \mathbb{R}$ partiell differenzierbar. Wenn $\nabla f(x) = 0$, so heißt x *kritische Stelle* von f .

11.4 Satz. Sei $U \subseteq \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ von der Klasse C^3 . Sei x eine kritische Stelle von f .

(a) Ist $H_f(x)$ positiv definit, so besitzt f in x ein striktes lokales Minimum.

11. Extremwerte und kritische Stellen

(b) Ist $H_f(x)$ negativ definit, so besitzt f in x ein striktes lokales Maximum.

(c) Ist $H_f(x)$ indefinit, so besitzt f in x kein lokales Extremum.

Beweis. (a) Wir benutzen das Korollar zur Taylor-Formel. Wähle also $r, M > 0$ so, dass

$$f(x + \xi) = f(x) + \langle \nabla f(x), \xi \rangle + \frac{1}{2} \langle H_f(x) \cdot \xi, \xi \rangle + R_3(\xi),$$

und $|R_3(\xi)| \leq M \|\xi\|_2^3$ für alle $\xi \in \overline{B}_r(0)$. Da $H_f(x)$ symmetrisch ist, besitzt $\mathbb{R}^n$ eine Orthonormalbasis aus Eigenvektoren von $H_f(x)$, d.h. es gibt $b_1, \dots, b_n \in \mathbb{R}^n$ und $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, so dass $H_f(x)b_i = \lambda_i b_i$ und $\langle b_i, b_j \rangle = \delta_{i,j}$ für $1 \leq i, j \leq n$. Jedes $\xi \in \mathbb{R}^n$ lässt sich schreiben als $\xi = \sum_{i=1}^n \xi_i b_i$. Dann

$$\|\xi\|_2^2 = \left\langle \sum_{i=1}^n \xi_i b_i, \sum_{j=1}^n \xi_j b_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \xi_i \xi_j \langle b_i, b_j \rangle = \sum_{i=1}^n \xi_i^2.$$

Ferner

$$\langle H_f(x) \cdot \xi, \xi \rangle = \left\langle \sum_{i=1}^n \lambda_i \xi_i b_i, \sum_{j=1}^n \xi_j b_j \right\rangle = \sum_{i=1}^n \lambda_i \xi_i^2.$$

Da $H_f(x)$ positiv definit ist, sind alle λ_i positiv. Setze $m = \min\{\lambda_i \mid 1 \leq i \leq n\} > 0$. Wähle

$$\epsilon = \min\left\{\frac{m}{4M}, r\right\}.$$

Dann gilt für alle $\xi \in B_\epsilon(0) \setminus \{0\}$

$$\begin{aligned} f(x + \xi) &= f(x) + \frac{1}{2} \sum_{i=1}^n \lambda_i \xi_i^2 + R_3(\xi) \geq f(x) + \frac{m}{2} \sum_{i=1}^n \xi_i^2 - M \|\xi\|_2^3 \\ &= f(x) + \left(\frac{m}{2} - M \|\xi\|_2\right) \|\xi\|_2^2 \geq f(x) + \frac{m}{4} \|\xi\|_2^2 > f(x). \end{aligned}$$

(b) beweist man durch Anwendung des Arguments auf $-A$.

(c) beweist man durch Anwendung auf skalare Vielfache je eines Eigenvektors zu einem positiven und einem negativen Eigenwert. $\square$

11.5 Bemerkung. (a) Wenn in einem kritischen Punkt kein Extremum vorliegt, so spricht man von einem *Sattelpunkt* von f .

(b) Satz 11.4 gilt auch dann noch, wenn f nur von der Klasse C^2 ist. Details findet man in den angegebenen Büchern.

11.6 Beispiel. (a) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^2 - y^2$. Dann ist $(0, 0)$ der einzige kritische Punkt. Dort ist H_f indefinit. Also besitzt f in $(0, 0)$ einen Sattel. Den zugehörigen Graphen zeigt Abbildung 11.1

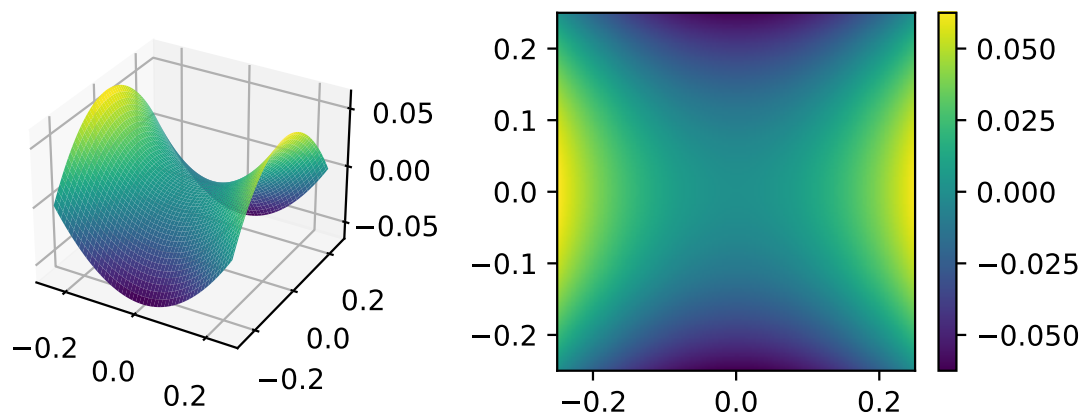


Abbildung 11.1.: Graphen zu Beispiel 11.6 (a)

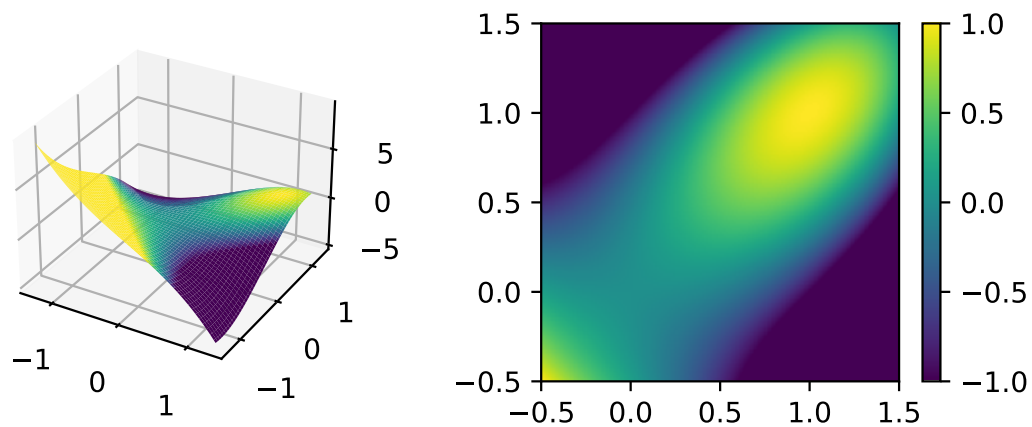


Abbildung 11.2.: Graphen zu Beispiel 11.6 (b)

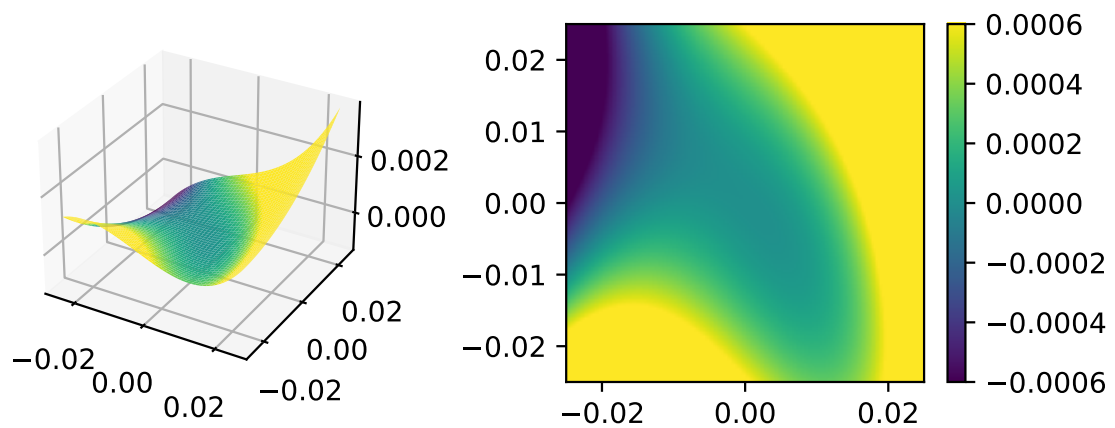


Abbildung 11.3.: Graphen zu Beispiel 11.6 (c)

11. Extremwerte und kritische Stellen

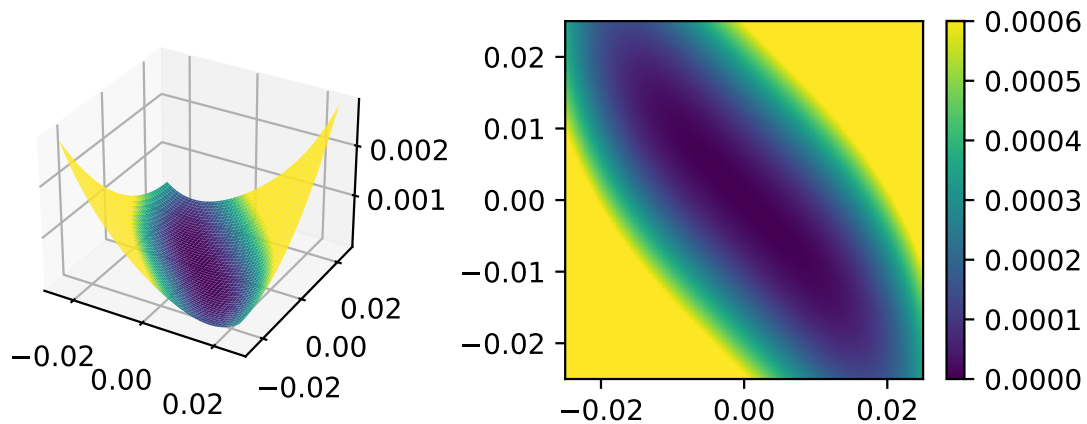


Abbildung 11.4.: Graphen zu Beispiel 11.6 (d)

- (b) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = 3xy - x^3 - y^3$. Dann gibt es zwei kritische Punkte, nämlich $(x, y) = (0, 0)$ und $(x, y) = (1, 1)$. In $(0, 0)$ ist die Hessesche indefinit, es liegt also ein Sattel vor. In $(1, 1)$ ist die Hessesche negativ definit, es liegt also ein striktes lokales Maximum vor. Den zugehörigen Graphen zeigt Abbildung 11.2
- (c) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = (x + y)^2 + 80x^3$. Der einzige kritische Punkt ist der Ursprung. Wegen $f(t, -t) = 80t^3$ liegt dort ein Sattel vor. Den zugehörigen Graphen zeigt Abbildung 11.3
- (d) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = (x + y)^2 + 900x^4$. Der einzige kritische Punkt ist der Ursprung. Für $(x, y) \neq 0$ ist $f(x, y) > 0$; dies ist für $x \neq 0$ sofort klar, und für $x = 0$ gilt $f(0, y) = y^2$. Also besitzt f ein striktes globales Minimum im Ursprung. Den zugehörigen Graphen zeigt Abbildung 11.4
- (e) Es gibt kein Polynom der Form $p(x, y) = \sum_{|\alpha| \geq 3} a_\alpha x^{\alpha_1} y^{\alpha_2}$, so dass $f(x, y) = (x + y)^2 + p(x, y)$ im Ursprung ein lokales Maximum besitzt.

Beweis. Für die Funktion $g(t) = f(0, t) = t^2 + p(0, t)$ gelten $g'(0) = 0$ und $g''(0) = 2$. Also besitzt g in $t = 0$ ein striktes lokales Minimum. $\square$

Teil II.

Gewöhnliche Differentialgleichungen

12. Differentialgleichungen erster Ordnung

12.1 Definition. Sei $U \subseteq \mathbb{R}^2$ offen und $f: U \rightarrow \mathbb{R}$ stetig. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $\varphi: I \rightarrow \mathbb{R}$ differenzierbar. Wenn für alle $x \in I$ gilt $(x, \varphi(x)) \in U$ und $\varphi'(x) = f(x, \varphi(x))$, so heißt φ Lösung der Differentialgleichung

$$y' = f(x, y). \quad (12.1)$$

Man bezeichnet (12.1) als *explizite Differentialgleichung erster Ordnung*.

Seien zusätzlich noch $x_0 \in I$ und $y_0 \in \mathbb{R}$ mit $(x_0, y_0) \in U$ gegeben. Falls $\varphi(x_0) = y_0$, so sagt man, dass φ die *Anfangsbedingung*

$$y(x_0) = y_0$$

erfüllt. Man sagt dann auch, φ sei Lösung der *Anfangswertaufgabe* $y' = f(x, y)$, $y(x_0) = y_0$.

12.2 Beispiel. (a) Sei $J \subseteq \mathbb{R}$ ein offenes Intervall und $U = J \times \mathbb{R}$. Für eine stetige Funktion $g: J \rightarrow \mathbb{R}$ definiere $f(x, y) = g(x)$. Das bedeutet, dass die Differentialgleichung $y' = g(x)$ betrachtet wird. Ihre Lösungen sind die Stammfunktionen von g .

(b) Sei $U = \mathbb{R}^2$ und sei $f: U \rightarrow \mathbb{R}$ definiert durch $f(x, y) = y$. Das bedeutet, dass die Differentialgleichung $y' = y$ betrachtet wird. Ihre Lösungen kennen wir aus der Analysis I.

Für jedes $c \in \mathbb{R}$ ist $\varphi_c(x) = ce^x$ eine Lösung, und andere Lösungen gibt es nicht.

Die Graphen von $\varphi_{1/2}$ und $\varphi_{-1/5}$ zusammen mit dem Richtungsfeld der Differentialgleichung zeigt Abbildung 12.1

12.3 Bezeichnung. Das Richtungsfeld einer expliziten Differentialgleichung erster Ordnung $y' = f(x, y)$ mit $f: U \rightarrow \mathbb{R}$ ist die Abbildung $F: U \rightarrow \mathbb{R}^2$, $(x, y) \mapsto (1, f(x, y))$. Man verdeutlicht das Richtungsfeld, indem man an den Punkt (x, y) ein (kleines) positives Vielfaches des Vektors $(1, f(x, y))$ als Pfeil anzeichnet.

Wenn $\varphi: I \rightarrow \mathbb{R}$ eine Lösung der Differentialgleichung ist, dann sind die Pfeile des Richtungsfelds tangential an die Lösungskurve $\{(x, \varphi(x)) \mid x \in I\}$.

12. Differentialgleichungen erster Ordnung

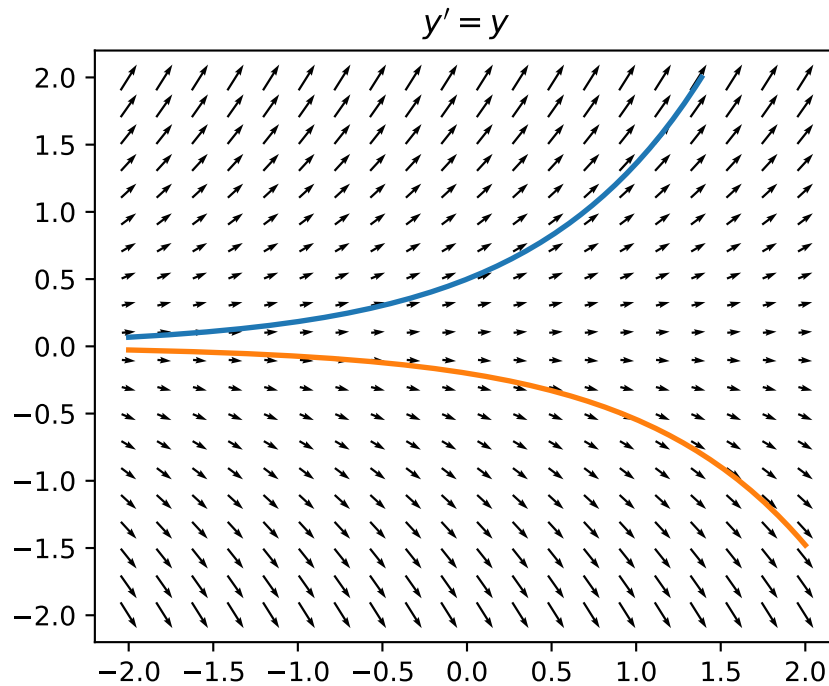


Abbildung 12.1.: Graph und Richtungsfeld zu Beispiel 12.2 (b)

12.4 Satz (Getrennte Variablen). *Es seien $I \subseteq \mathbb{R}$ und $J \subseteq \mathbb{R}$ offene Intervalle, $f: I \rightarrow \mathbb{R}$ und $g: J \rightarrow \mathbb{R}$ stetig mit $g(y) \neq 0$ für alle $y \in J$. Seien $x_0 \in I$ und $y_0 \in J$ beliebig. Dann existiert ein offenes Intervall $I_0 \subseteq I$ mit $x_0 \in I_0$, so dass die Anfangswertaufgabe*

$$y' = f(x)g(y), \quad y(x_0) = y_0,$$

genau eine Lösung $\varphi: I_0 \rightarrow \mathbb{R}$ besitzt. Diese Lösung ist von der Form $\varphi(x) = G^{-1}(F(x))$, wobei

$$F(x) = \int_{x_0}^x f(t) dt, \quad G(y) = \int_{y_0}^y \frac{ds}{g(s)}.$$

Für I_0 kann jedes offene Intervall gewählt werden, welches den Bedingungen $x_0 \in I_0$ und $F(I_0) \subseteq G(J)$ genügt.

Beweis. Da $g(y) \neq 0$ für alle $y \in J$, ist g entweder immer positiv oder immer negativ. In jedem Fall ist G streng monoton und daher invertierbar mit differenzierbarer Inverser. Speziell ist $G(J)$ offen in $\mathbb{R}$. Da F stetig ist, ist $F^{-1}(G(J))$ offen. Wegen $F(x_0) = 0 = G(y_0)$ umfasst $F^{-1}(G(J))$ also ein Intervall I_0 mit $x_0 \in I_0$. Sei jetzt φ wie in der Behauptung. Dann

$$\varphi'(x) = \frac{1}{G'(G^{-1}(F(x)))} F'(x) = g(\varphi(x))f(x).$$

Sei nun $\psi: I_0 \rightarrow \mathbb{R}$ eine weitere Lösung der Anfangswertaufgabe. Dann

$$(G \circ \psi)'(x) = G'(\psi(x))\psi'(x) = \frac{1}{g(\psi(x))}g(\psi(x))f(x) = f(x).$$

Daraus erhält man durch Integration unter Verwendung von $G(y_0) = 0$

$$(G \circ \psi)(x) = \int_{x_0}^x f(x) dx = F(x). \quad \square$$

12.5 Beispiel. In $U = \mathbb{R} \times \mathbb{R}$ sei die Anfangswertaufgabe

$$y' = xe^{-y}, \quad y(0) = 1,$$

gegeben. Das bedeutet in der Notation von Satz 12.4

$$I = \mathbb{R}, \quad J = \mathbb{R}, \quad f(x) = x, \quad g(y) = e^{-y}, \quad x_0 = 0, \quad y_0 = 1.$$

Damit berechnen wir

$$F(x) = \int_{x_0}^x f(t) dt = \int_0^x t dt = \frac{x^2}{2}$$

und

$$G(y) = \int_{y_0}^y \frac{ds}{g(s)} = \int_1^y e^s ds = e^y - e.$$

$G^{-1}(F(x))$ ist die Lösung der Gleichung $G(y) = F(x)$, also von $e^y - e = \frac{x^2}{2}$. Daher

$$\varphi(x) = G^{-1}(F(x)) = \log\left(\frac{x^2}{2} + e\right).$$

Jetzt muss noch der Definitionsbereich von φ angegeben werden. Es gelten $G(J) =]-e, \infty[$ und $F(\mathbb{R}) = [0, \infty[$. Für $I_0 = \mathbb{R}$ gilt also $F(I_0) \subseteq G(J)$. Daher ist die eindeutige Lösung der Differentialgleichung gegeben durch

$$\varphi: \mathbb{R} \rightarrow \mathbb{R}, \quad \varphi(x) = \log\left(\frac{x^2}{2} + e\right).$$

Graph und Richtungsfeld dieser Lösung zeigt Abbildung 12.2.

12.6 Beispiel. In $U = \mathbb{R}^2$ betrachten wir die Differentialgleichung

$$y' = y^2.$$

In der Situation des Satzes über getrennte Variable hat man also $f(x) = 1$ und $g(y) = y^2$. Setzt man $I = \mathbb{R}$ und $J =]0, \infty[$, so ist g nullstellenfrei. Für gegebene Anfangsdaten $(x_0, y_0) \in I \times J$ erhalten wir

$$F(x) = \int_{x_0}^x dt = x - x_0 \quad \text{und} \quad G(y) = \int_{y_0}^y \frac{ds}{s^2} = \frac{1}{y_0} - \frac{1}{y}.$$

12. Differentialgleichungen erster Ordnung

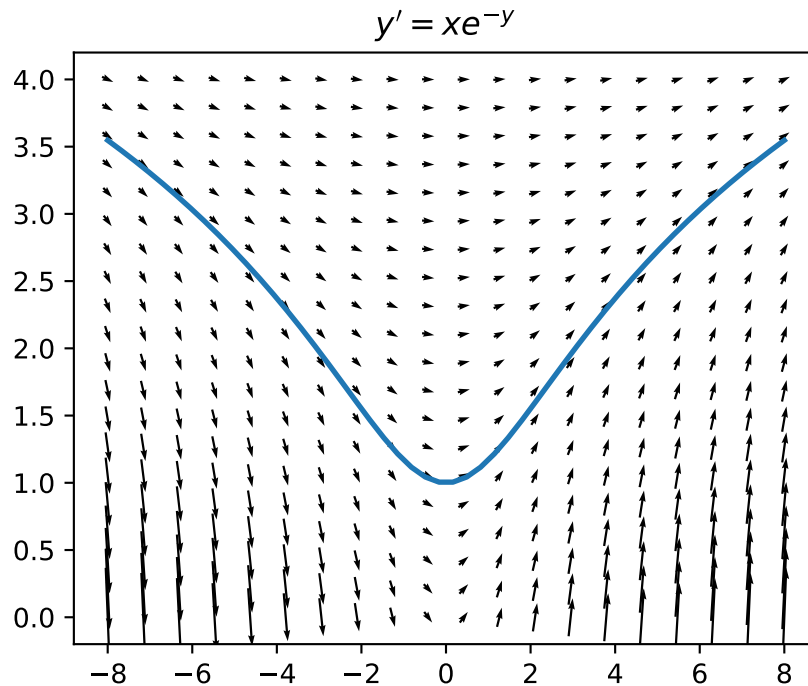


Abbildung 12.2.: Graph und Richtungsfeld zu Beispiel 12.5

Um $G^{-1}(F(x))$ zu bestimmen, müssen wir die Gleichung

$$\frac{1}{y_0} - \frac{1}{y} = x - x_0$$

nach y auflösen. Ihre Lösung ist

$$\varphi(x) = \frac{y_0}{1 - y_0(x - x_0)}.$$

Der Nenner hat eine Nullstelle bei $x_0 + \frac{1}{y_0}$. Wegen $y_0 > 0$ liegt diese Stelle rechts von x_0 . Wir erhalten also folgende Lösung der Anfangswertaufgabe

$$\varphi: \left] -\infty, x_0 + \frac{1}{y_0} \right[\rightarrow \mathbb{R}, \varphi(x) = \frac{y_0}{1 - y_0(x - x_0)}.$$

Falls dagegen $y_0 < 0$ gilt, so muss $J =]-\infty, 0[$ gewählt werden. Die Rechnung ist dieselbe, aber die Nullstelle des Nenners liegt nun links von x_0 und man erhält die Lösung

$$\varphi: \left] x_0 + \frac{1}{y_0}, \infty \right[\rightarrow \mathbb{R}, \varphi(x) = \frac{y_0}{1 - y_0(x - x_0)}.$$

Im Falle $y_0 = 0$ ist der Satz über die getrennten Variablen nicht anwendbar. Man sieht die Lösung $\varphi \equiv 0$.

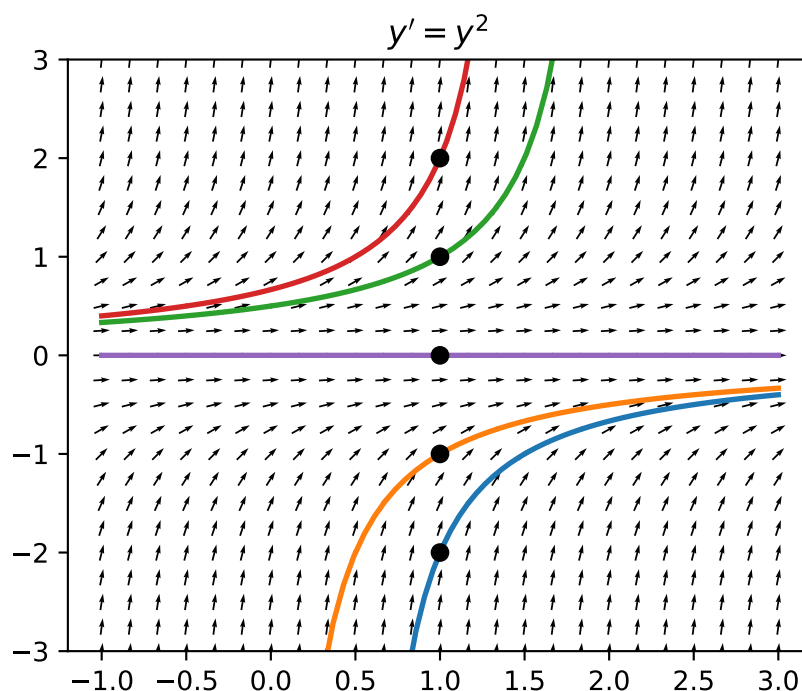


Abbildung 12.3.: Graph und Richtungsfeld zu Beispiel 12.6

Die Graphen der Lösungen der Anfangswertaufgaben zu $x_0 = 1$ und $y_0 = -2, -1, 0, 1, 2$ zeigt Abbildung 12.3.

Die Lösung ist eindeutig im folgenden Sinn: Wenn es ein Intervall I_0 mit $x_0 \in I_0$ und Lösungen $\varphi_j: I_0 \rightarrow \mathbb{R}$, $j = 1, 2$, der Anfangswertaufgabe gibt, so gilt $\varphi_1 = \varphi_2$. Wir nehmen zum Widerspruch an, dass es ein ξ mit $\varphi_1(\xi) \neq \varphi_2(\xi)$ gibt. Wir behandeln den Fall, dass $\xi > x_0$. Dann sei $x_1 = \sup\{x > x_0 | \varphi_1(x) = \varphi_2(x)\}$. Wegen der Stetigkeit der Lösungen gilt $\varphi_1(x_1) = \varphi_2(x_1) =: y_1$. Daher lösen φ_1 und φ_2 die Anfangswertaufgabe mit den Anfangsdaten (x_1, y_1) . Falls $y_1 \neq 0$, so haben wir einen Widerspruch zum Satz über getrennte Variable. Also $y_1 = 0$. Da φ_1 und φ_2 nicht beide die Nullfunktion sein können, gibt es $x_2 > x_1$ in I_0 mit, sagen wir, $\varphi_2(x_2) \neq 0$. Wir haben aber nachgerechnet, dass es keine Lösung gibt, die sowohl den Wert 0 als auch von Null verschiedene Werte annimmt.

Bemerkung. Es gibt zwei typische Fragestellungen:

- Alle Lösungen einer Anfangswertaufgabe werden gesucht.
- Die *Lösungsgesamtheit* einer Differentialgleichung wird gesucht. Das sind alle Lösungen von allen möglichen Anfangswertaufgaben.

12.7 Satz (Verheftungslemma). *Es sei $U \subseteq \mathbb{R}^2$ offen und $f: U \rightarrow \mathbb{R}$ stetig. Es seien I und J zwei Intervalle und $\varphi: I \rightarrow \mathbb{R}$ und $\psi: J \rightarrow \mathbb{R}$ zwei Lösungen der*

12. Differentialgleichungen erster Ordnung

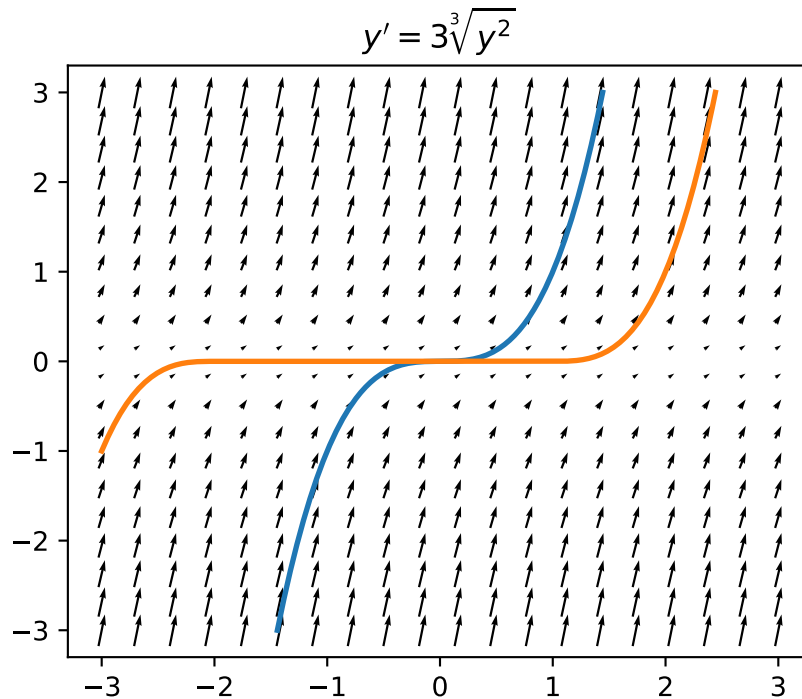


Abbildung 12.4.: Graph und Richtungsfeld zu Beispiel 12.8

Differentialgleichung $y' = f(x, y)$. Es gelte außerdem $\sup I = \inf J =: x_0$ und $\lim_{x \nearrow x_0} \varphi(x) = \lim_{x \searrow x_0} \psi(x) =: y_0$. Dann wird durch

$$\rho(x) = \begin{cases} \varphi(x), & x \in I, \\ y_0, & x = x_0, \\ \psi(x), & x \in J, \end{cases}$$

eine Lösung von $y' = f(x, y)$ gegeben.

Beweis. Sei $x < x_0$. Wegen des Mittelwertsatzes der Analysis I gibt es $\xi \in]x, x_0[$ mit

$$\frac{\rho(x) - \rho(x_0)}{x - x_0} = \rho'(\xi) = \varphi'(\xi) = f(\xi, \varphi(\xi)) \rightarrow f(x_0, y_0) = f(x_0, \rho(x_0)), \quad x \nearrow x_0.$$

Für $x > x_0$ argumentiert man analog. □

12.8 Beispiel. Sei $U = \mathbb{R}^2$ und $f(x, y) = 3y^{\frac{2}{3}}$. Gesucht ist die Lösungsgesamtheit der Differentialgleichung

$$y' = 3\sqrt[3]{y^2}.$$

Wir verwenden nun die Leibniz-Notation

$$\frac{dy}{dx} = 3\sqrt[3]{y^2}.$$

Wir bringe alle Terme in y auf die linke Seite und gehen zur Stammfunktion über

$$\int y^{-2/3} dy = \int 3 dx,$$

also

$$3y^{1/3} = 3x + C.$$

Löse nach y auf

$$y = (x - c)^3,$$

wobei $c = -C/3$ (aus Bequemlichkeit).

Für $c \in \mathbb{R}$ sei $\varphi_c: \mathbb{R} \rightarrow \mathbb{R}$ definiert durch $\varphi_c(x) = (x - c)^3$. Für $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$ mit $a \leq b$ setze

$$\varphi_{a,b}(x) = \begin{cases} \varphi_a(x), & \text{für } x \leq a, \\ 0, & \text{für } a < x < b, \\ \varphi_b(x), & \text{für } x \geq b. \end{cases}$$

Wegen des Verheftungslemmas [12.7](#) ist $\varphi_{a,b}$ differenzierbar und eine Lösung der Differentialgleichung. Jede Anfangswertaufgabe besitzt also unendlich viele Lösungen. Da alle Lösungen monoton wachsen, gibt es auch nicht noch mehr.

Die Graphen von $\varphi_{0,0}$ und $\varphi_{-2,1}$ zusammen mit dem Richtungsfeld der Differentialgleichung zeigt Abbildung [12.4](#)

13. Der Satz von Picard-Lindelöf

In diesem Abschnitt wird der Satz von Picard-Lindelöf formuliert. Der Beweis folgt in Paragraph 19 der Vorlesung.

13.1 Definition. Sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ offen und seien $f_1, \dots, f_n: U \rightarrow \mathbb{R}$ stetig. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und seien $\varphi_1, \dots, \varphi_n: I \rightarrow \mathbb{R}$ differenzierbar. Wenn für alle $x \in I$ sowohl $(x, \varphi_1(x), \dots, \varphi_n(x)) \in U$ als auch $\varphi'_i(x) = f_i(x, \varphi_1(x), \dots, \varphi_n(x))$ für $i = 1, \dots, n$ gelten, dann ist das Tupel $(\varphi_1, \dots, \varphi_n)$ eine Lösung des *expliziten Differentialgleichungssystems erster Ordnung*.

$$\begin{aligned} y'_1 &= f_1(x, y_1, \dots, y_n) \\ &\vdots \\ y'_n &= f_n(x, y_1, \dots, y_n). \end{aligned} \tag{13.1}$$

Dafür schreibt man meist einfacher $y' = f(x, y)$ mit der Interpretation $f = (f_1, \dots, f_n)$ und $\varphi = (\varphi_1, \dots, \varphi_n)$.

13.2 Definition. Seien (X, d) und (Y, d) metrische Räume und $f: X \rightarrow Y$ eine Abbildung.

- (a) f heißt *Lipschitz-stetig*, wenn es ein $L \geq 0$ gibt mit $d(f(a), f(b)) \leq Ld(a, b)$ für alle $a, b \in X$.
- (b) f heißt *lokal Lipschitz-stetig* in $x_0 \in X$, wenn es eine Umgebung U von x_0 gibt, so dass die Einschränkung $f|_U$ Lipschitz-stetig ist.

13.3 Bemerkung. (a) Lipschitz-stetige Abbildungen sind lokal Lipschitz-stetig und lokal Lipschitz-stetige Abbildungen sind stetig.

- (b) C^1 -Abbildungen sind lokal Lipschitz-stetig. Das ergibt sich sofort aus dem Mittelwertsatz.
- (c) Für $\alpha > 0$ ist die Funktion $x \mapsto |x|^\alpha$ genau dann lokal Lipschitz-stetig, wenn $\alpha \geq 1$.

Beweis. Für $0 < \alpha < 1$ gilt $|f(x) - f(0)| = |x|^\alpha = |x|^{\alpha-1}|x - 0|$. Für $\alpha < 1$ ist der Faktor vor $|x - 0|$ unbeschränkt.

Für $\alpha = 1$ verwenden wir die umgekehrte Dreiecksungleichung $||x| - |y|| \leq |x - y|$. Für $\alpha > 1$ ist f von der Klasse C^1 . $\square$

13. Der Satz von Picard-Lindelöf

13.4 Definition. Seien X , Y und Z metrische Räume, sei $U \subseteq X \times Y$, und sei $f: U \rightarrow Z$ eine Abbildung.

(a) f heißt *Lipschitz-stetig im zweiten Argument*, wenn es ein $L \geq 0$ gibt, so dass

$$d(f(x, y_1), f(x, y_2)) \leq Ld(y_1, y_2) \quad \text{für alle } (x, y_1), (x, y_2) \in U.$$

(b) f heißt *lokal Lipschitz-stetig im zweiten Argument*, wenn es zu jedem Paar $(x_0, y_0) \in U$ eine Umgebung V gibt, so dass $f|_V$ Lipschitz-stetig im zweiten Argument ist.

13.5 Bemerkung. Sei $U \subseteq \mathbb{R}^m \times \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}^k$. Wir schreiben die Elemente von U als $(x, y) = (x_1, \dots, x_m, y_1, \dots, y_n)$. Wenn die partiellen Ableitungen $\frac{\partial f}{\partial y_1}, \dots, \frac{\partial f}{\partial y_n}$ existieren und stetige Funktionen von (x, y) sind, dann ist f lokal Lipschitz-stetig im zweiten Argument.

13.6 Theorem (Existenz- und Eindeigkeitssatz von Picard und Lindelöf). *Sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ offen, sei $f: U \rightarrow \mathbb{R}^n$ stetig (in beiden Argumenten) und lokal Lipschitz-stetig im zweiten Argument, und sei $(x_0, y_0) \in U$. Dann existieren ein offenes Intervall I mit $x_0 \in I$ und eine Lösung $\varphi: I \rightarrow \mathbb{R}^n$ der Differentialgleichung $y' = f(x, y)$ mit folgenden Eigenschaften*

(a) $\varphi(x_0) = y_0$.

(b) Ist $\psi: J \rightarrow \mathbb{R}^n$ ebenfalls eine Lösung der Differentialgleichung $y' = f(x, y)$ mit $\psi(x_0) = y_0$, so gelten $J \subseteq I$ und $\psi = \varphi|_J$.

Beweis. Dieser Beweis wird auf das Ende der Vorlesung verschoben. □

13.7 Beispiel. Die Funktion $f(x, y) = 3\sqrt[3]{y^2}$ aus Beispiel 12.8 erfüllt die Voraussetzungen des Satzes von Picard-Lindelöf nicht. In der Tat haben wir Anfangswertaufgaben gefunden, deren Lösung nicht eindeutig ist.

Die Funktion $f(x, y) = y^2$ aus Beispiel 12.6 erfüllt dagegen die Voraussetzungen des Satzes. Wir hatten die Eindeutigkeit bereits aus dem Satz über getrennte Variablen erhalten.

13.8 Satz (Existenzsatz von Peano). *Sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ offen, sei $f: U \rightarrow \mathbb{R}^n$ stetig, und sei $(x_0, y_0) \in U$. Dann existiert eine Lösung $\varphi: I \rightarrow \mathbb{R}^n$ der Differentialgleichung $y' = f(x, y)$ mit $y(x_0) = y_0$.*

Den Existenzsatz von Peano werden wir in dieser Vorlesung nicht beweisen. Einen Beweis findet man beispielsweise in dem Bucha "Gewöhnliche Differentialgleichungen" von Walter.

13.9 Definition. Sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ offen, sei $f: U \rightarrow \mathbb{R}$ stetig. Ferner sei I ein offenes Intervall und sei $\varphi: I \rightarrow \mathbb{R}$ n -mal differenzierbar. Wenn für alle $x \in I$ sowohl $(x, \varphi(x), \varphi'(x), \dots, \varphi^{(n-1)}(x)) \in U$ als auch $\varphi^{(n)}(x) = f(x, \varphi(x), \varphi'(x), \dots, \varphi^{(n-1)}(x))$ gelten, dann ist φ eine Lösung der expliziten Differentialgleichung n -ter Ordnung

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}).$$

Gelten für φ außerdem noch $\varphi(x_0) = y_0, \varphi'(x_0) = y_1, \dots, \varphi^{(n-1)}(x_0) = y_{n-1}$ für gegebene Werte $(x_0, y_0, y_1, \dots, y_{n-1}) \in U$, so löst φ die Anfangswertaufgabe

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}), \quad y(x_0) = y_0, \dots, y^{(n-1)}(x_0) = y_{n-1}.$$

13.10 Satz. Für f wie oben definieren wir $F: U \rightarrow \mathbb{R}^n$ durch

$$F(x, y_0, y_1, \dots, y_{n-1}) = (y_1, \dots, y_{n-1}, f(x, y_0, y_1, \dots, y_{n-1})).$$

Mit $Y = (y_0, y_1, \dots, y_{n-1})$ lautet das Differentialgleichungssystem $Y' = F(x, Y)$ ausgeschrieben

$$\begin{aligned} y_0' &= y_1 \\ y_1' &= y_2 \\ &\vdots \\ y_{n-2}' &= y_{n-1} \\ y_{n-1}' &= f(x, y_0, \dots, y_{n-1}). \end{aligned}$$

Daher gelten

- (a) Ist φ eine Lösung der Differentialgleichung $y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$, so ist $\Phi = (\varphi, \varphi', \dots, \varphi^{(n-1)})$ eine Lösung des Differentialgleichungssystems $Y' = F(x, Y)$.
- (b) Ist $\Phi = (\varphi_0, \varphi_1, \dots, \varphi_{n-1})$ eine Lösung von $Y' = F(x, Y)$, so ist φ_0 eine Lösung von $y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$.

13.11 Beispiel (Harmonischer Oszillator). : Für $\omega > 0$ betrachten wir die Differentialgleichung zweiter Ordnung $y'' = -\omega^2 y$ mit der Anfangsbedingung $y(0) = 1, y'(0) = 0$. Wir haben also $U = \mathbb{R}^3$ und $f(x, y, y') = -\omega^2 y$. Solche Differentialgleichungen werden wir später genauer betrachten. Die Lösung dieser Anfangswertaufgabe lässt sich sofort erraten: $\varphi(x) = \cos(\omega x)$.

Wir übersetzen die Differentialgleichung wie folgt in ein System erster Ordnung:

Der Vektor im $\mathbb{R}^2$ heißt $Y = \begin{pmatrix} Y_0 \\ Y_1 \end{pmatrix}$. Die zugehörige Differentialgleichung ist $Y' = F(x, Y)$ mit

$$F(x, Y) = \begin{pmatrix} Y_1 \\ -\omega^2 Y_0 \end{pmatrix}.$$

13. Der Satz von Picard-Lindelöf

Die Anfangsbedingung ist $Y(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Die oben angegebene Lösung übersetzt sich zu

$$\Phi(x) = \begin{pmatrix} \cos(\omega x) \\ -\omega \sin(\omega x) \end{pmatrix}.$$

Diese Technik wird hauptsächlich benutzt, um Existenz- und Eindeigkeitssätze für Differentialgleichungen höherer Ordnung auf die entsprechenden Sätze für Differentialgleichungssysteme erster Ordnung zurückzuführen.

13.12 Korollar (Existenz- und Eindeigkeitssatz für Differentialgleichungen höherer Ordnung). *Sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ offen, sei $f: U \rightarrow \mathbb{R}$ stetig und lokal Lipschitz-stetig im zweiten Argument. Sei $(x_0, y_0, \dots, y_{n-1}) \in U$. Dann existiert eine Lösung $\varphi: I \rightarrow \mathbb{R}$ der Differentialgleichung*

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$$

mit

(a) $\varphi(x_0) = y_0, \varphi'(x_0) = y_1, \dots, \varphi^{(n-1)}(x_0) = y_{n-1}.$

(b) *Ist $\psi: J \rightarrow \mathbb{R}$ eine Lösung mit $\psi^{(k)}(x_0) = y_k$ für $k = 0, \dots, n-1$, so ist $J \subseteq I$ und $\psi = \varphi|_J$.*

13.13 Beispiel. (a) Die Lösung $\varphi(x) = \cos(\omega x)$ ist die einzige Lösung der Anfangswertaufgabe $y'' = -\omega^2 y$, $y(0) = 1$, $y'(0) = 0$.

(b) Die Gleichung des mathematischen Pendels ist

$$y'' = -\omega^2 \sin(y).$$

Aus dem Korollar 13.12 folgt für jede Anfangswertaufgabe die Existenz einer eindeutigen Lösung der Pendelgleichung. Diese lässt sich aber nicht geschlossen ausdrücken.

14. Lineare Differentialgleichungen

Mit $\mathbb{K}$ ist nachfolgend entweder $\mathbb{R}$ oder $\mathbb{C}$ gemeint.

14.1 Definition. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $A: I \rightarrow \mathbb{K}^{n \times n}$ und $b: I \rightarrow \mathbb{K}^n$ seien stetige Abbildungen. Definiere $f: I \times \mathbb{K}^n \rightarrow \mathbb{K}^n$ durch $f(x, y) = A(x)y + b(x)$. Dann bezeichnet man $y' = f(x, y)$ als *lineares Differentialgleichungssystem erster Ordnung*. Ist dabei $b(x) = 0$ für alle $x \in I$, so ist das System *homogen*.

14.2 Satz. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $A: I \rightarrow \mathbb{K}^{n \times n}$ und $b: I \rightarrow \mathbb{K}^n$ seien stetige Abbildungen. Sei $(x_0, y_0) \in I \times \mathbb{K}^n$. Dann besitzt die Anfangswertaufgabe $y' = A(x)y + b(x)$, $y(x_0) = y_0$, eine eindeutig bestimmte Lösung $\varphi: I \rightarrow \mathbb{K}^n$.

Beweis. Das ist eine Verschärfung des lokalen Existenz- und Eindeutigkeitssatzes. Genau wie diesen werden wir sie gegen Ende des Semesters beweisen. Die Verschärfung gegenüber dem Satz von Picard-Lindelöf besteht darin, dass der Definitionsbereich der Lösung gleich I ist. $\square$

14.3 Satz. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und sei $A: I \rightarrow \mathbb{K}^{n \times n}$ eine stetige Abbildung. Die Lösungsgesamtheit des homogenen Differentialgleichungssystems $y' = A(x)y$ bildet einen Untervektorraum des $\mathbb{K}$ -Vektorraums aller stetigen Abbildungen $I \rightarrow \mathbb{K}^n$.

14.4 Satz (Homogene lineare Differentialgleichung). Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, sei $a: I \rightarrow \mathbb{R}$ stetig. Sei $(x_0, y_0) \in I \times \mathbb{R}$. Dann existiert genau eine Lösung $\varphi: I \rightarrow \mathbb{R}$ von $y' = a(x)y$ mit $\varphi(x_0) = y_0$. Diese Lösung hat die Form

$$\varphi(x) = y_0 \exp\left(\int_{x_0}^x a(t) dt\right).$$

Beweis. Entweder durch Ansatz $\varphi(x) = \exp(g(x))$ oder mit der Methode der getrennten Variablen. $\square$

14.5 Beispiel. Bestimme die Lösungsgesamtheit der Differentialgleichung

$$y' = \frac{y}{2\sqrt{x}}, \quad x \in I =]0, \infty[, y \in \mathbb{R}.$$

Dann $a(x) = \frac{1}{2\sqrt{x}}$ und

$$\int a(x) dx = \int \frac{1}{2\sqrt{x}} dx = \sqrt{x}.$$

14. Lineare Differentialgleichungen

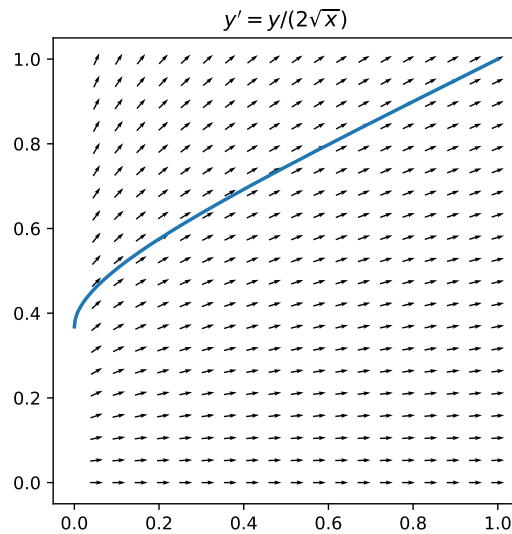


Abbildung 14.1.: Richtungsfeld und Lösungskurve zu Beispiel 14.5

Alle Lösungen der Differentialgleichung sind von der Form

$$\varphi(x) = C \exp(\sqrt{x}), \quad C \in \mathbb{R}.$$

Nun sei die Lösung gesucht, welche der Anfangsbedingung $y(1) = 1$ genügt. Das bedeutet $1 = C \exp(1)$, also $C = \frac{1}{e}$. Das Richtungsfeld und diese Lösung zeigt Abbildung 14.1.

14.6 Satz (Methode der Variation der Konstanten). *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, seien $a, b: I \rightarrow \mathbb{R}$ stetig. Definiere $f: I \times \mathbb{R} \rightarrow \mathbb{R}$ durch $f(x, y) = a(x)y + b(x)$. Sei $(x_0, y_0) \in I \times \mathbb{R}$. Dann existiert genau eine Lösung $\psi: I \rightarrow \mathbb{R}$ von $y' = a(x)y + b(x)$ mit $\psi(x_0) = y_0$. Diese Lösung hat die Form*

$$\psi(x) = \varphi(x) \left(y_0 + \int_{x_0}^x \frac{b(t)}{\varphi(t)} dt \right),$$

wobei φ die Lösung der homogenen Differentialgleichung $y' = a(x)y$ mit der Anfangsbedingung $\varphi(x_0) = 1$ ist.

Beweis. Setze $\psi(x) = C(x)\varphi(x)$, dann $\psi'(x) = C'(x)\varphi(x) + a(x)\psi(x)$. Also $\psi'(x) = a(x)\psi(x) + b(x)$ genau dann, wenn $C'(x)\varphi(x) = b(x)$. $\square$

Bemerkung. Die allgemeine Lösung einer inhomogenen linearen Differentialgleichung besteht aus der allgemeinen Lösung der zugehörigen homogenen Differentialgleichung plus einer speziellen Lösung der inhomogenen Differentialgleichung.

14.7 Beispiel. Wir suchen die Lösungsgesamtheit der linearen, inhomogenen Differentialgleichung

$$y' = \frac{y-1}{2\sqrt{x}}, \quad x \in I :=]0, \infty[, y \in \mathbb{R}.$$

Die Lösungsgesamtheit der zugehörigen homogenen Differentialgleichung kennen wir schon aus Beispiel 14.5. Es handelt sich um alle Vielfachen von $\varphi(x) = \exp(\sqrt{x})$. Mit Satz 14.6 bestimmen wir eine spezielle Lösung der inhomogenen Gleichung, nämlich

$$\begin{aligned} \varphi_i(x) &= \varphi(x) \int \frac{-1}{2\sqrt{x} \exp(\sqrt{x})} dx = e^{\sqrt{x}} \int \frac{-1}{\exp(u)} du = e^{\sqrt{x}} \exp(-u) \\ &= e^{\sqrt{x}} \exp(-\sqrt{x}) = 1. \end{aligned}$$

Die allgemeine Lösung der inhomogenen Differentialgleichung ist also von der Form

$$\psi(x) = C \exp(\sqrt{x}) + 1, \quad C \in \mathbb{R}.$$

Will man nun eine Anfangswertaufgabe wie etwa $y(1) = \frac{1}{2}$ lösen, so muss man C entsprechend anpassen. In diesem Fall ergibt sich $C = -\frac{1}{2e}$. Abbildung 14.2 zeigt die Lösungen der Anfangswertaufgaben $y(1) = \frac{1}{2}$, $y(1) = 1$ und $y(1) = \frac{3}{2}$.

Ich wiederhole den Begriff der Basis aus der Linearen Algebra.

14.8 Definition. Ein Tupel $(b_1, \dots, b_n)$ in einem $\mathbb{K}$ -Vektorraum V ist eine *Basis* von V , wenn die folgenden beiden Bedingungen erfüllt sind

- (a) $(b_1, \dots, b_n)$ ist *linear unabhängig*, das heißt, dass aus $\sum_{j=1}^n \lambda_j b_j = 0$ bereits $\lambda_1 = \dots = \lambda_n = 0$ folgt.
- (b) $(b_1, \dots, b_n)$ ist ein *Erzeugendensystem* von V , das heißt, dass es zu jedem $v \in V$ Skalare $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ gibt, so dass $v = \sum_{j=1}^n \lambda_j b_j$.

In der Linearen Algebra I wird gezeigt, dass alle Basen eines Vektorraums die gleiche Anzahl an Elementen haben. Diese Anzahl ist die *Dimension* des Vektorraums. Ferner wird dort folgendes gezeigt:

14.9 Theorem. Es sei V ein $\mathbb{K}$ -Vektorraum der Dimension n und es seien $b_1, \dots, b_n \in V$. Dann sind gleichwertig

- (a) $(b_1, \dots, b_n)$ ist eine Basis von V .
- (b) $(b_1, \dots, b_n)$ ist linear unabhängig.
- (c) $(b_1, \dots, b_n)$ ist ein Erzeugendensystem von V .

Den Zusammenhang zwischen Basen im $\mathbb{K}^n$ und Determinanten liefert der folgende Satz aus der Linearen Algebra I.

14. Lineare Differentialgleichungen

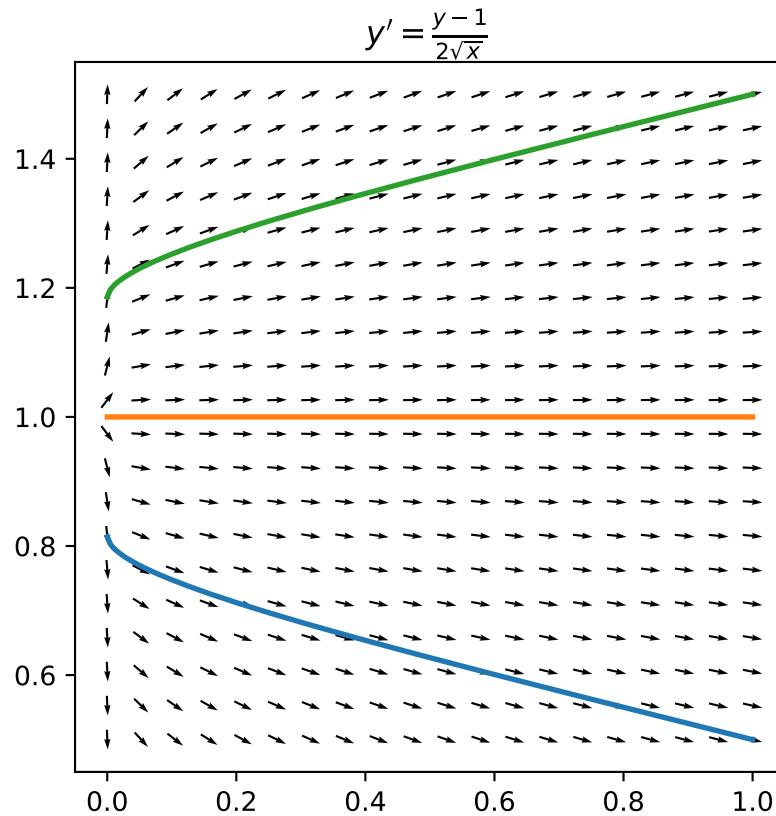


Abbildung 14.2.: Richtungsfeld und Lösungskurve zu Beispiel 14.7

14.10 Satz. Für eine Matrix $A \in \mathbb{K}^{n \times n}$ sind äquivalent

- (a) $\det A \neq 0$.
- (b) Die Spalten von A bilden eine Basis des $\mathbb{K}^n$.
- (c) Die Zeilen von A bilden eine Basis des $\mathbb{K}^n$.

14.11 Satz. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und sei $A: I \rightarrow \mathbb{K}^{n \times n}$ eine stetige Abbildung. Sei

$$L = \{\varphi: I \rightarrow \mathbb{K}^n \mid \varphi'(x) = A(x)\varphi(x)\}$$

die Lösungsgesamtheit des homogenen Differentialgleichungssystems.

Für $\varphi^1, \varphi^2, \dots, \varphi^n \in L$ sind äquivalent

- (a) $(\varphi^1, \dots, \varphi^n)$ ist eine Basis von L .
- (b) Für jedes $x_0 \in I$ ist $(\varphi^1(x_0), \dots, \varphi^n(x_0))$ eine Basis des $\mathbb{K}^n$.
- (c) Es gibt ein $x_0 \in I$, so dass $(\varphi^1(x_0), \dots, \varphi^n(x_0))$ eine Basis des $\mathbb{K}^n$ ist.

Insbesondere hat die Lösungsgesamtheit des homogenen Differentialgleichungssystems die Dimension n .

Beweis. (a) $\implies$ (b): Weil der $\mathbb{K}^n$ die Dimension n hat, genügt es zu zeigen, dass $(\varphi^1(x_0), \dots, \varphi^n(x_0))$ linear unabhängig ist. Sei dazu

$$\sum_{j=1}^n \lambda_j \varphi^j(x_0) = 0.$$

Setze $\psi(x) = \sum_{j=1}^n \lambda_j \varphi^j(x)$. Dann $\psi \in L$. In der Tat löst ψ die Anfangswertaufgabe $y' = A(x)y$, $y(x_0) = 0$. Da die Lösung der Anfangswertaufgabe eindeutig ist, folgt $\psi = 0$. Aus Voraussetzung (a) folgt $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$.

(b) $\implies$ (c): klar

(c) $\implies$ (a): Es ist nur noch zu zeigen, dass $(\varphi_1, \dots, \varphi_n)$ ein Erzeugendensystem von L ist. Sei dazu $\psi \in L$ beliebig. Setze $y_0 = \psi(x_0)$. Dann ist ψ die eindeutige Lösung der Anfangswertaufgabe $y' = Ay$, $y(x_0) = y_0$.

Andererseits ist aber $(\varphi^1(x_0), \dots, \varphi^n(x_0))$ eine Basis des $\mathbb{K}^n$. Es existieren daher $\lambda_1, \dots, \lambda_n \in \mathbb{K}^n$, so dass $y_0 = \sum_{j=1}^n \lambda_j \varphi^j(x_0)$. Für diese Wahl der λ_j löst $\varphi(x) = \sum_{j=1}^n \lambda_j \varphi_j$ die obige Anfangswertaufgabe. Wegen der Eindeutigkeit der Lösung gilt $\varphi = \psi$. $\square$

14.12 Definition. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und sei $A: I \rightarrow \mathbb{K}^{n \times n}$ eine stetige Abbildung. Sei

$$L = \{\varphi: I \rightarrow \mathbb{K}^n \mid \varphi'(x) = A(x)\varphi(x)\}$$

die Lösungsgesamtheit des homogenen Differentialgleichungssystems. Jede Basis $(\varphi^1, \dots, \varphi^n)$ von L bezeichnet man als *Fundamentalsystem* des Differentialgleichungssystems $y' = A(x)y$.

Für $\varphi^1, \dots, \varphi^n \in L$ bezeichnet man $\det(\varphi^1, \dots, \varphi^n)$ als *Wronski-Determinante*.

14.13 Bemerkung. Die Lösungen $\varphi^1, \dots, \varphi^n$ bilden genau dann ein Fundamentalsystem, wenn ihre Wronski-Determinante $W(x)$ für mindestens ein $x \in I$ nicht verschwindet. In diesem Fall verschwindet $W(x)$ für kein $x \in I$.

14.14 Beispiel.

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= -y_1 \end{aligned}$$

Zwei Lösungen sind gegeben durch

$$\varphi^1(x) = \begin{pmatrix} \sin x \\ \cos x \end{pmatrix} \quad \text{und} \quad \varphi^2(x) = \begin{pmatrix} \cos x \\ -\sin x \end{pmatrix}.$$

Ihre Wronski-Determinante ist gleich -1 . Daher ist (φ^1, φ^2) ein Fundamentalsystem.

14. Lineare Differentialgleichungen

14.15 Beispiel. Wir bestimmen ein Fundamentalsystem für das Differentialgleichungssystem

$$\begin{aligned}y_1' &= xy_2 - y_1 \\ y_2' &= 2y_2.\end{aligned}\tag{14.1}$$

In Matrixschreibweise hat die Differentialgleichung die Gestalt

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} -1 & x \\ 0 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Die allgemeine Lösung schreiben wir als Vektor $\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}$. Dann $\varphi_2(x) = C_2 e^{2x}$, und φ_1 ist eine Lösung der inhomogenen linearen Differentialgleichung

$$y_1' = -y_1 + C_2 x e^{2x}.\tag{14.2}$$

Eine Lösung der zu (14.2) gehörigen homogenen Gleichung ist $\varphi_{1,h}(x) = e^{-x}$. Nach der Formel über die Variation der Konstanten erhält man dann die folgende spezielle Lösung der inhomogenen Differentialgleichung

$$\varphi_{1,i}(x) = e^{-x} \int \frac{C_2 x e^{2x}}{e^{-x}} dx = C_2 e^{-x} \int x e^{3x} dx = C_2 e^{-x} \left(\frac{x}{3} e^{3x} - \frac{1}{9} e^{3x} \right).$$

Die Lösungsgesamtheit von (14.1) ist also

$$\left\{ \begin{pmatrix} C_1 e^{-x} + C_2 \left(\frac{x}{3} - \frac{1}{9} \right) e^{2x} \\ C_2 e^{2x} \end{pmatrix} \middle| C_1, C_2 \in \mathbb{R} \right\}.$$

Ein Fundamentalsystem ist

$$\Phi(x) = \begin{pmatrix} e^{-x} & \left(\frac{x}{3} - \frac{1}{9} \right) e^{2x} \\ 0 & e^{2x} \end{pmatrix}.$$

Seine Wronski-Determinante beträgt e^x .

14.16 Definition. Ein *komplexer normierter Raum* besteht aus einem $\mathbb{C}$ -Vektorraum V und einer Abbildung $V \rightarrow \mathbb{R}$, $v \mapsto \|v\|$, mit

- (a) $\|v\| \geq 0$ für alle $v \in V$.
- (b) $\|v\| = 0$ dann und nur dann, wenn $v = 0$.
- (c) $\|\alpha v\| = |\alpha| \|v\|$ für alle $\alpha \in \mathbb{C}$, $v \in V$.
- (d) $\|v + w\| \leq \|v\| + \|w\|$ für alle $v, w \in V$.

14.17 Beispiel. Auf $\mathbb{C}^n$ gebräuchliche Normen sind

$$\begin{aligned}\|z\|_1 &= |z_1| + \cdots + |z_n|, \\ \|z\|_2 &= \sqrt{|z_1|^2 + \cdots + |z_n|^2}, \\ \|z\|_\infty &= \max\{|z_1|, \dots, |z_n|\}.\end{aligned}$$

Man kann jeden komplexen normierten Raum auch als reellen normierten Raum auffassen. Daher sind je zwei Normen $\|\cdot\|$ und $\|\cdot\|$ auf dem $\mathbb{C}^n$ äquivalent.

14.18 Beispiel. Es sei X ein metrischer Raum. Der Raum $C_b(X, \mathbb{C})$ aller beschränkten, stetigen Funktionen $f: X \rightarrow \mathbb{C}$ ist ein komplexer normierter Raum, wenn man ihn mit der Norm

$$\|f\|_\infty = \sup_{x \in X} |f(x)|$$

versieht.

Aus der Linearen Algebra ist der folgende Satz bekannt:

14.19 Satz (Cramersche Regel). Sei $\mathbb{K}$ ein Körper. Für $A \in \mathbb{K}^{n \times n}$ mit $\det(A) \neq 0$ gilt

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

wobei $\tilde{A}$ eine $n \times n$ -Matrix ist, die aus geeigneten Unterdeterminanten von A besteht.

14.20 Bemerkung. Im Falle $n = 2$ lautet die Cramersche Regel

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

14.21 Bezeichnung. Mit $GL_n(\mathbb{K}) \subseteq \mathbb{K}^{n \times n}$ werde die Gruppe aller invertierbarer $n \times n$ -Matrizen bezeichnet.

14.22 Satz. Der $\mathbb{K}^{n \times n}$ sei mit einer beliebigen Matrixnorm versehen. Dann ist $GL_n(\mathbb{K})$ offen und die Abbildung

$$T: GL_n(\mathbb{K}) \rightarrow GL_n(\mathbb{K}), \quad A \mapsto A^{-1},$$

ist stetig.

Beweis. Verwende $\det A = \sum_{\sigma \in S_n} \text{signum}(\sigma) \prod_{j=1}^n a_{j, \sigma(j)}$ und die Cramersche Regel. □

Bemerkung. Wir hatten in Beispiel 7.15 sogar schon die Differenzierbarkeit von T gezeigt. Daraus folgt die Stetigkeit natürlich auch.

14.23 Lemma. Es sei $I \subseteq \mathbb{R}$ ein offenes Intervall und es seien $A: I \rightarrow \mathbb{K}^{m \times n}$ und $B: I \rightarrow \mathbb{K}^{k \times m}$ differenzierbare Abbildungen. Dann ist $B \circ A: I \rightarrow \mathbb{K}^{k \times n}$, $x \mapsto B(x) \circ A(x)$, differenzierbar mit $(B \circ A)'(x) = B'(x) \circ A(x) + B(x) \circ A'(x)$.

14. Lineare Differentialgleichungen

Beweis. Sei $C(x) = B(x) \circ A(x)$. Wir betrachten aus Bequemlichkeit nur $c_{1,1}$.

$$c_{1,1}(x) = \sum_{j=1}^m b_{1,j}(x) a_{j,1}(x).$$

Also

$$c'_{1,1}(x) = \sum_{j=1}^m (b'_{1,j}(x) a_{j,1}(x) + b_{1,j}(x) a'_{j,1}(x)).$$

Das ist aber das $(1, 1)$ -Element von $B'(x) \circ A(x) + B(x) \circ A'(x)$. $\square$

14.24 Definition. Sei $f: [a, b] \rightarrow \mathbb{R}^n$ stetig mit $f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix}$. Dann definieren wir das Integral komponentenweise, also durch

$$\int_a^b f(x) dx = \begin{pmatrix} \int_a^b f_1(x) dx \\ \vdots \\ \int_a^b f_n(x) dx \end{pmatrix}.$$

14.25 Bemerkung. Aus dem Hauptsatz der Differential- und Integralgleichung folgt sofort

$$\frac{d}{dx} \int_a^x f(t) dt = f(x).$$

14.26 Satz. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $A: I \rightarrow \mathbb{K}^{n \times n}$ und $b: I \rightarrow \mathbb{K}^n$ seien stetige Abbildungen. Sei L der Lösungsraum des homogenen Differentialgleichungssystems $y' = A(x)y$ und sei ψ_0 eine Lösung des inhomogenen Differentialgleichungssystems $y' = A(x)y + b(x)$. Dann ist die Lösungsgesamtheit des inhomogenen Differentialgleichungssystems gleich $\psi_0 + L$.

14.27 Satz (Variation der Konstanten). Sei L der Lösungsraum des homogenen linearen Differentialgleichungssystems $y' = A(x)y$ und sei $\varphi^1, \dots, \varphi^n$ eine Basis von L . Definiert man die Matrix

$$\Phi(x) = \begin{pmatrix} \varphi^1(x) & \varphi^2(x) & \dots & \varphi^n(x) \end{pmatrix},$$

so wird eine Lösung des inhomogenen Differentialgleichungssystems $y' = A(x)y + b(x)$ gegeben durch

$$\psi(x) = \Phi(x)u(x),$$

wobei

$$u(x) = \int_{x_0}^x \Phi^{-1}(t)b(t)dt.$$

Beweis. Wegen der Cramerschen Regel ist $\Phi^{-1}: I \rightarrow \mathbb{K}^{n \times n}$ eine stetige Abbildung. Daher existiert das Integral. Nun beweist man die Formel durch Ableiten. $\square$

14.28 Beispiel. Betrachte

$$y_1' = y_2 + \sin x$$

$$y_2' = -y_1 + 2$$

Wir hatten bereits ein Fundamentalsystem für das homogene Differentialgleichungssystem ausgerechnet: Dieses führt zu

$$\Phi(x) = \begin{pmatrix} \sin x & \cos x \\ \cos x & -\sin x \end{pmatrix}.$$

Die Cramersche Regel ergibt $\Phi^{-1}(x) = \Phi(x)$ für alle x . Damit erhalten wir

$$\begin{aligned} u(x) &= \int \begin{pmatrix} \sin x & \cos x \\ \cos x & -\sin x \end{pmatrix} \cdot \begin{pmatrix} \sin x \\ 2 \end{pmatrix} dx = \int \begin{pmatrix} \sin^2(x) + 2 \cos(x) \\ \sin(x) \cos(x) - 2 \sin(x) \end{pmatrix} dx \\ &= \begin{pmatrix} \frac{x}{2} - \frac{1}{2} \sin(x) \cos(x) + 2 \sin(x) \\ \frac{1}{2} \sin^2(x) + 2 \cos(x) \end{pmatrix}. \end{aligned}$$

Daher ist

$$\psi(x) = \begin{pmatrix} \frac{x}{2} \sin(x) + 2 \\ \frac{x}{2} \cos(x) - \frac{1}{2} \sin(x) \end{pmatrix}$$

eine spezielle Lösung des inhomogenen Systems.

Bemerkung. Wenn man ein Fundamentalsystem kennt, dann kann man alle zugehörigen inhomogenen Differentialgleichungssysteme mit der Methode der Variation der Konstanten lösen. Es gibt aber kein allgemeines Verfahren zur Bestimmung eines Fundamentalsystems.

14.29 Definition. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, seien $a_0, \dots, a_{n-1}, b: I \rightarrow \mathbb{K}$ stetig. Dann bezeichnet man

$$y^{(n)} = a_0(x)y + a_1(x)y' + \dots + a_{n-1}(x)y^{(n-1)} + b(x) \quad (14.3)$$

als *lineare Differentialgleichung n-ter Ordnung*. Ist $b(x) = 0$ für alle x , so ist sie *homogen*.

14.30 Satz. (a) Die Lösungsgesamtheit L von (14.3) mit $b \equiv 0$ ist ein n -dimensionaler $\mathbb{K}$ -Vektorraum.

(b) $\varphi_1, \dots, \varphi_n \in L$ bilden genau dann eine Basis von L , wenn für ein und damit für alle $x \in I$ gilt $W(x) \neq 0$, wobei

$$W(x) = \det \begin{pmatrix} \varphi_1(x) & \varphi_2(x) & \dots & \varphi_n(x) \\ \varphi_1'(x) & \varphi_2'(x) & \dots & \varphi_n'(x) \\ \vdots & \vdots & & \vdots \\ \varphi_1^{(n-1)}(x) & \varphi_2^{(n-1)}(x) & \dots & \varphi_n^{(n-1)}(x) \end{pmatrix}$$

die Wronski-Determinante der Differentialgleichung ist.

14. Lineare Differentialgleichungen

(c) Die Lösungsgesamtheit der inhomogenen Differentialgleichung (14.3) ist gleich $\psi + L$, wobei ψ eine Lösung der inhomogenen Differentialgleichung ist.

(d) Jede Anfangsbedingung $y(x_0) = y_0, y'(x_0) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1}$ besitzt eine eindeutig bestimmte Lösung $\varphi: I \rightarrow \mathbb{K}$.

14.31 Bemerkung (Variation der Konstanten). Der Übersetzungsmechanismus aus Satz 13.10 ermöglicht die Bestimmung einer speziellen Lösung der inhomogenen Differentialgleichung, wenn man ein Fundamentalsystem für die homogene Differentialgleichung kennt.

Wir führen das für $n = 2$ im Detail aus. Gegeben sei also die Differentialgleichung

$$y'' = a_0(x)y + a_1(x)y' + b(x). \quad (14.4)$$

Das zugehörige System erster Ordnung ist

$$Y' = A(x)Y + B(x) \quad (14.5)$$

mit

$$A(x) = \begin{pmatrix} 0 & 1 \\ a_0(x) & a_1(x) \end{pmatrix} \quad \text{und} \quad B(x) = \begin{pmatrix} 0 \\ b(x) \end{pmatrix}.$$

Es sei (φ_1, φ_2) ein Fundamentalsystem für die homogene Differentialgleichung (14.4). Dann ist

$$\Phi(x) = \begin{pmatrix} \varphi_1(x) & \varphi_2(x) \\ \varphi_1'(x) & \varphi_2'(x) \end{pmatrix}$$

ein Fundamentalsystem für (14.5). Sei $W(x)$ die zugehörige Wronski-Determinante. Dann sagt die Cramersche Regel

$$\Phi^{-1}(x) = \frac{1}{W(x)} \begin{pmatrix} \varphi_2'(x) & -\varphi_2(x) \\ -\varphi_1'(x) & \varphi_1(x) \end{pmatrix}.$$

Eine Lösung des inhomogenen Differentialgleichungssystems (14.5) ist daher

$$\Psi(x) = \Phi(x) \int \Phi^{-1}(x)B(x)dx = \Phi(x) \int \frac{1}{W(x)} \begin{pmatrix} -\varphi_2(x)b(x) \\ \varphi_1(x)b(x) \end{pmatrix} dx.$$

Die erste Komponente von $\Psi(x)$ ist dann eine Lösung $\psi(x)$ der inhomogenen Differentialgleichung (14.4). Sie hat die Gestalt

$$\psi(x) = -\varphi_1(x) \int \frac{\varphi_2(x)b(x)}{W(x)} dx + \varphi_2(x) \int \frac{\varphi_1(x)b(x)}{W(x)} dx.$$

14.32 *Beispiel.* Wir betrachten die inhomogene, lineare Differentialgleichung

$$y'' = \frac{6}{x^2}y + 1.$$

Wir sehen zwei Lösungen der homogenen Gleichung, nämlich $\varphi_1(x) = x^3$ und $\varphi_2(x) = \frac{1}{x^2}$. Die Wronski-Determinante ist also

$$W(x) = \det \begin{pmatrix} x^3 & \frac{1}{x^2} \\ 3x^2 & -\frac{2}{x^3} \end{pmatrix} = -5.$$

Eine spezielle Lösung der inhomogenen Differentialgleichung wird folglich gegeben durch

$$\psi(x) = -x^3 \int \frac{-1}{5x^2} dx + \frac{1}{x^2} \int \frac{-x^3}{5} dx = -\frac{x^3}{5x} - \frac{x^4}{20x^2} = -\frac{1}{4}x^2.$$

14.33 *Bemerkung* (D'Alembertsches Reduktionsverfahren). Wir betrachten die homogene lineare Differentialgleichung zweiter Ordnung

$$y'' = a_0(x)y + a_1(x)y'.$$

Dann besteht jedes Fundamentalsystem aus zwei linear unabhängigen Lösungen. Eine Lösung $\varphi_1 \neq 0$ sei bekannt. Dann erlaubt das Reduktionsverfahren von d'Alembert, auf einem Intervall, auf welchem φ_1 nur positive Werte annimmt, eine linear unabhängige Lösung φ_2 zu bestimmen. Dazu macht man den Ansatz $\varphi_2 = u\varphi_1$ mit einer zweimal differenzierbaren Funktion u . Man findet dann nach Einsetzen der Differentialgleichung für φ_1

$$\varphi_2''(x) = u''(x)\varphi_1(x) + 2u'(x)\varphi_1'(x) + u(x)a_0(x)\varphi_1(x) + u(x)a_1(x)\varphi_1'(x)$$

und

$$a_0(x)\varphi_2(x) + a_1(x)\varphi_2'(x) = a_0(x)u(x)\varphi_1(x) + a_1(x)u'(x)\varphi_1(x) + a_1(x)u(x)\varphi_1'(x).$$

Der Vergleich beider Ausdrücke führt durch Auslöschung der gefärbten Terme auf

$$u''(x) = u'(x) \left(a_1(x) - \frac{2\varphi_1'(x)}{\varphi_1(x)} \right).$$

Das ist eine lineare Differentialgleichung erster Ordnung für u' . Ihre Lösungen sind Vielfache von

$$u'(x) = \frac{1}{\varphi_1^2(x)} \exp \left(\int a_1(x) dx \right).$$

14.34 *Beispiel.*

$$y'' = -4y' - 4y.$$

Exponentialansatz liefert eine Lösung, nämlich $\varphi_1(x) = e^{-2x}$. Das d'Alembertsche Reduktionsverfahren führt auf die Differentialgleichung $u'(x) = 1$. Also $\varphi_2(x) = xe^{-2x}$.

15. Jordansche Normalform

15.1 Definition. Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt *diagonalisierbar*, wenn $\mathbb{C}^n$ eine Basis aus Eigenvektoren von A besitzt. Den $\mathbb{C}$ -Vektorraum der Eigenvektoren zum Eigenwert λ bezeichne ich mit $E(\lambda)$.

15.2 Bemerkung. (a) Eine Matrix $A \in \mathbb{C}^{n \times n}$ mit Eigenwerten $\lambda_1, \dots, \lambda_m$ ist also genau dann diagonalisierbar, wenn $\sum_{j=1}^m \dim E(\lambda_j) = n$.

(b) Die Eigenwerte sind die Nullstellen des charakteristischen Polynoms

$$\chi_A(z) = \det(z \operatorname{id} - A)$$

und die Eigenvektoren zum Eigenwert λ sind die Lösungen des linearen Gleichungssystems $\lambda x = Ax$.

(c) Der Satz 10.4 über die Hauptachsentransformation sagt unter anderem aus, dass jede reelle, symmetrische Matrix diagonalisierbar ist.

(d) Über einem Körper, der nicht algebraisch abgeschlossen ist, hat das charakteristische Polynom möglicherweise nicht genug Nullstellen. Deswegen arbeiten wir über $\mathbb{C}$.

(e) Die Matrix $J = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ besitzt das charakteristische Polynom $\chi_J(z) = z^2$ und ist nicht diagonalisierbar.

15.3 Beispiel. Das zur Differentialgleichung $y'' = -4y' - 4y$ gehörende System ist $Y' = AY$ mit

$$A = \begin{pmatrix} 0 & 1 \\ -4 & -4 \end{pmatrix}.$$

Das charakteristische Polynom ist $\chi_A(z) = (z + 2)^2$. Wegen $-2 \operatorname{id} - A = \begin{pmatrix} -2 & -1 \\ 4 & 2 \end{pmatrix}$ gilt $E(-2) = \operatorname{LH}\left(\left\{\begin{pmatrix} 1 \\ -2 \end{pmatrix}\right\}\right)$. Daher ist A nicht diagonalisierbar.

Es gilt aber $(-2 \operatorname{id} - A)^2 = 0$. Wir wählen ein Element aus $\ker(-2 \operatorname{id} - A)^2 \setminus E(-2)$, sagen wir $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, so gilt $(-2 \operatorname{id} - A) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 6 \end{pmatrix} \in E(-2)$. Der Satz über die Jordansche Normalform sagt aus, dass dies kein Zufall ist.

15. Jordansche Normalform

15.4 Definition. Für $\lambda \in \mathbb{C}$ und $m \in \mathbb{N}$ ist der *Jordan-Block* definiert durch

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots \\ \vdots & & & & \vdots \\ 0 & & \dots & 0 & \lambda & 1 \\ 0 & & \dots & 0 & \lambda \end{pmatrix} \in \mathbb{C}^{m \times m}.$$

15.5 Theorem (Jordan-Zerlegung). Es sei $A \in \mathbb{C}^{n \times n}$ mit $\chi_A(z) = \prod_{j=1}^{\ell} (z - \lambda_j)^{M_j}$ mit paarweise verschiedenen λ_j . Für $j = 1, \dots, \ell$ gibt es Zahlen k_j und $m_{j,1}, \dots, m_{j,k_j} \in \mathbb{N}$, sowie $T \in GL_n(\mathbb{C})$, so dass $A = T J T^{-1}$, wobei J die folgende Blockdiagonalmatrix ist

$$J = \begin{pmatrix} J_{m_{1,1}}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{m_{1,2}}(\lambda_1) & 0 & \dots \\ \vdots & & \ddots & \vdots \\ 0 & J_{m_{1,k_1}}(\lambda_1) & 0 & \vdots \\ & 0 & J_{m_{2,1}}(\lambda_2) & 0 \\ \vdots & & & \vdots \\ 0 & \dots & 0 & J_{m_{\ell,k_\ell}}(\lambda_\ell) \end{pmatrix}.$$

Bemerkung. (a) A ist genau dann diagonalisierbar, wenn alle $m_{j,i}$ gleich 1 sind. In diesem Fall bilden die Spalten von T eine Basis des $\mathbb{C}^n$ aus Eigenvektoren.

(b) In jedem Fall gilt $\sum_{i=1}^{k_j} m_{j,i} = M_j$.

(c) Der Rang von $A - \lambda_j \text{id}$ ist $n - M_j + k_j$. Quadriert man $A - \lambda_j \text{id}$, so sinkt der Rang pro Jordan-Block der Größe ≥ 2 um eins, also ist $\text{Rang}(A - \lambda_j \text{id}) - \text{Rang}(A - \lambda_j \text{id})^2$ gleich der Anzahl der Jordan-Blöcke zum Eigenwert λ_j , die mindestens die Größe 2 haben. Allgemein ist $\text{Rang}(A - \lambda_j \text{id})^{m-1} - \text{Rang}(A - \lambda_j \text{id})^m$ gleich der Anzahl der Jordan-Blöcke zum Eigenwert λ_j , die mindestens die Größe m haben. Wir können also für jedes m die Anzahl der Jordan-Blöcke dieser Größe feststellen.

15.6 Bemerkung. Wir bestimmen T (unter der Annahme, dass wir Eigenwerte exakt bestimmen können). Das geschieht für die jeweiligen Eigenwerte unabhängig voneinander. Wir machen das also nur für den ersten.

Aus der Bemerkung kennen wir die Größen aller Jordan-Blöcke zum Eigenwert λ_1 . Der größte davon habe die Größe m . Zur Abkürzung setzen wir $B = A - \lambda_1 \text{id}$.

Wir wählen ein $v_m \in \ker B^m$ mit $B^{m-1}v_m \neq 0$ und setzen $v_{m-i} = B^i v_m$ für $i = 0, \dots, m-1$. Dann bilden $v_1, \dots, v_m$ die ersten m Spaltenvektoren von T und das erste Blockdiagonalelement von J ist $J_m(\lambda_1)$.

Es sei nun p die Größe des zweitgrößten Jordan-Blocks zum Eigenwert λ_1 . Wir wählen $w_p \in \ker B^p$ mit $B^{p-1}w_p \neq 0$, welches außerdem noch linear unabhängig

zu den bereits bestimmten Spaltenvektoren von T ist, und setzen $w_{p-i} = B^i w_p$ für $i = 0, \dots, p-1$. Dann bilden $w_1, \dots, w_p$ die nächsten p Spaltenvektoren von T und das zweite Blockdiagonalelement von J ist $J_p(\lambda_1)$. So fahren wir rekursiv fort, bis alle Jordan-Blöcke konstruiert sind.

15.7 Beispiel. (a) Für $A = \begin{pmatrix} 0 & 1 \\ -4 & -4 \end{pmatrix}$ ist $B = A + 2 \text{id} = \begin{pmatrix} 2 & 1 \\ -4 & -2 \end{pmatrix}$. Es gelten $\text{Rang}(B) = 1$ und $\text{Rang}(B^2) = 0$. Es gibt also einen Jordan-Block der Größe 2 und sonst nichts. Wir wählen $v_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Dann $v_1 = Bv_2 = \begin{pmatrix} 2 \\ -4 \end{pmatrix}$. Also

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

(b) Wir betrachten ein aufwändigeres Beispiel

$$A = \begin{pmatrix} 5 & 1 & -4 & -3 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 6 & 1 & -7 & -4 \end{pmatrix}.$$

Das charakteristische Polynom ist $\chi_A(z) = (z-2)^3(z+1)$. Wir beginnen mit dem Eigenwert 2. Es gelten

$$B = A - 2 \text{id} = \begin{pmatrix} 3 & 1 & -4 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 6 & 1 & -7 & -6 \end{pmatrix} \quad \text{und} \quad B^2 = \begin{pmatrix} -9 & 0 & 9 & 9 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -18 & 0 & 18 & 18 \end{pmatrix}.$$

Also $\text{Rang}(B) = 2$ und $\text{Rang}(B^2) = 1$. Es gibt also einen Jordan-Block der Größe 2 und einen der Größe 1. Wir wählen $v_2 \in \ker B^2$ mit $Bv_2 \neq 0$, z.B. $v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$. Dann $v_1 = Bv_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$. Der erste Jordan-Block ist abgearbeitet. Der zweitgrößte hat die Größe 1. Wir benötigen ein Element aus $\ker B$, welches linear unabhängig zu v_1 ist. (Linear unabhängig zu v_2 ist es automatisch.) Wir wählen $v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$.

Schließlich benötigen wir noch einen Eigenvektor zum Eigenwert -1 . Wir wählen $v_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix}$. Dann

$$\begin{pmatrix} 5 & 1 & -4 & -3 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 6 & 1 & -7 & -4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix}^{-1}.$$

16. Lineare Differentialgleichungen mit konstanten Koeffizienten

16.1 Bemerkung. (a) Wie im vorigen Kapitel auch sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

(b) Wenn die rechte Seite einer linearen Differentialgleichung nicht von x abhängt, dann spricht man von einem System mit konstanten Koeffizienten.

(c) Die Lösungen eines linearen Differentialgleichungssystems mit konstanten Koeffizienten sind auf ganz $\mathbb{R}$ erklärt.

16.2 Definition. Sei V ein normierter Raum und sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in V .

(a) Die Reihe $\sum_{n=1}^{\infty} a_n$ heißt *konvergent*, wenn die Folge $\left(\sum_{n=1}^N a_n\right)_{N \in \mathbb{N}}$ in V konvergiert.

(b) Die Reihe $\sum_{n=1}^{\infty} a_n$ heißt *absolut konvergent*, wenn die Reihe $\sum_{n=1}^{\infty} \|a_n\|$ in $\mathbb{R}$ konvergiert.

16.3 Satz. In einem Banachraum ist jede absolut konvergente Reihe konvergent.

Beweis. Der Beweis ist wörtlich derselbe wie der von Satz 6.12 der Analysis I. $\square$

16.4 Definition. In $\mathbb{K}^{n \times n}$ bezeichne E_n die Einheitsmatrix. Wir setzen $A^0 = E_n$ für jede Matrix $A \in \mathbb{K}^{n \times n}$ und definieren

$$\exp A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k. \quad (16.1)$$

Man bezeichnet $\exp A$ als *Matrixexponential* von A . Man schreibt auch $\exp A = e^A$.

16.5 Satz. Die Exponentialreihe (16.1) konvergiert absolut.

16.6 Beispiel. (a) Für eine Diagonalmatrix

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

16. Lineare Differentialgleichungen mit konstanten Koeffizienten

gilt

$$\exp D = \begin{pmatrix} e^{d_1} & 0 & \dots & 0 \\ 0 & e^{d_2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & e^{d_n} \end{pmatrix}.$$

(b)

$$\exp \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

16.7 Lemma. Für $A, B \in \mathbb{K}^{n \times n}$ mit $AB = BA$ gilt $e^{A+B} = e^A e^B$.

Beweis. Man zeigt zunächst wie beim Beweis des Binomialsatzes aus der Analysis I für Matrizen A und B mit $AB = BA$

$$(A + B)^n = \sum_{k=0}^n \binom{n}{k} A^k B^{n-k}.$$

Man benötigt ebenfalls den Satz 6.23 der Analysis I über das Cauchy-Produkt. Auch dieser Beweis kann wörtlich übernommen werden. Damit ist der Beweis derselbe wie im skalaren Fall:

$$\exp(A) \exp(B) = \left(\sum_{n=0}^{\infty} \frac{1}{n!} A^n \right) \left(\sum_{n=0}^{\infty} \frac{1}{n!} B^n \right) = \sum_{n=0}^{\infty} c_n,$$

wobei

$$c_n = \sum_{k=0}^n \frac{1}{k!(n-k)!} A^k B^{n-k} = \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} A^k B^{n-k} = \frac{1}{n!} (A + B)^n. \quad \square$$

16.8 Satz. Sei $f: \mathbb{R} \rightarrow \mathbb{K}^{n \times n}$ gegeben durch $f(x) = \exp(xA)$. Dann ist f differenzierbar mit $f'(x) = A \exp(xA)$ für alle $x \in \mathbb{R}$.

Beweis.

$$\begin{aligned} & \frac{1}{h} (\exp((x+h)A) - \exp(xA)) - A \exp(xA) \\ &= \frac{1}{h} (\exp(hA) \exp(xA) - \exp(xA)) - A \exp(xA) \\ &= \left(\frac{1}{h} (\exp(hA) - E_n) - A \right) \exp(xA). \end{aligned}$$

Wir zeigen $\lim_{h \rightarrow 0} \frac{1}{h} (\exp(hA) - E_n) - A = 0$. Sei dazu $\epsilon > 0$ beliebig vorgegeben. Da für die skalare Funktion $f(t) = e^{t\|A\|}$ gilt $f'(0) = \|A\|$, existiert $\delta > 0$, so dass $|\frac{1}{t}(e^{t\|A\|} - 1) - \|A\|| < \epsilon$ falls $|t| < \delta$. Dann gilt für alle h mit $|h| < \delta$

$$\begin{aligned} \left\| \frac{1}{h} (\exp(hA) - E_n) - A \right\| &= \left\| \sum_{n=2}^{\infty} \frac{h^{n-1}}{n!} A^n \right\| \leq \lim_{N \rightarrow \infty} \sum_{n=2}^N \frac{|h|^{n-1}}{n!} \|A\|^n \\ &= \frac{1}{|h|} (e^{|h|\|A\|} - 1) - \|A\| < \epsilon. \quad \square \end{aligned}$$

16.9 Korollar. Ist $A \in \mathbb{K}^{n \times n}$, $x_0 \in \mathbb{R}$ und $y_0 \in \mathbb{K}^n$, so ist die einzige Lösung der Anfangswertaufgabe $y' = Ay$, $y(x_0) = y_0$, gegeben durch

$$\varphi(x) = e^{(x-x_0)A} y_0.$$

Beweis. Eindeutigkeit holen wir aus dem Existenz- und Eindeigkeitssatz für lineare Differentialgleichungen. $\square$

16.10 Beispiel. Gegeben sei das Differentialgleichungssystem

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= 0, \end{aligned}$$

zusammen mit der Anfangsbedingung $y(1) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$. In Matrixschreibweise lautet die Differentialgleichung $y' = Ay$ für $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Das Matrixexponential von A hatten wir in Beispiel 16.6 bestimmt. Daher löst die folgende Funktion die Anfangswertaufgabe

$$\varphi(x) = e^{(x-1)A} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 & x-1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 3-x \\ -1 \end{pmatrix}.$$

16.11 Lemma. Für jede Matrix $M \in \mathbb{K}^{n \times n}$ und jede invertierbare Matrix T gilt

$$\exp(TMT^{-1}) = T \exp(M) T^{-1}.$$

16.12 Satz. Wenn $A = TDT^{-1}$, dann bilden sowohl die Spalten von $\exp(xA)$ als auch die Spalten von $T \exp(xD)$ jeweils ein Fundamentalsystem für die homogene Differentialgleichung $y' = Ay$.

Beweis. Der erste Teil der Behauptung ist der Inhalt von Korollar 16.9. Der zweite Teil folgt nun aus $\exp(xA) = \exp(xTDT^{-1}) = T \exp(xD) T^{-1}$, denn die Multiplikation von rechts mit einer invertierbaren Matrix verändert lediglich das Fundamentalsystem. $\square$

16.13 Bemerkung. Insbesondere gilt: Wenn v ein Eigenvektor zum Eigenwert a ist, dann ist $e^{ax}v$ eine Lösung der homogenen Differentialgleichung.

16.14 Beispiel. Berechne $\exp(xA)$ für

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

Das charakteristische Polynom ist

$$\det(xE_2 - A) = x^2 + 1 = (x - i)(x + i).$$

16. Lineare Differentialgleichungen mit konstanten Koeffizienten

Also hat A die Eigenwerte $\pm i$. Lösung der entsprechenden linearen Gleichungssysteme liefert uns Eigenvektoren

$$v_1 = \begin{pmatrix} -i \\ 1 \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} i \\ 1 \end{pmatrix}.$$

Daraus erhalten wir mit der Cramerschen Regel

$$T = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad T^{-1} = \frac{1}{2} \begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}.$$

Mit

$$D = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

gilt dann $A = TDT^{-1}$. Also

$$\exp xA = T \exp(xD) T^{-1} = \frac{1}{2} \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e^{ix} & 0 \\ 0 & e^{-ix} \end{pmatrix} \begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix} = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}.$$

Das ist Fundamentalsystem aus Beispiel 14.14.

16.15 Beispiel.

$$\exp \left(x \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \right) = \begin{pmatrix} 1 & x & \frac{x^2}{2} \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix}.$$

16.16 Satz. Für $a \in \mathbb{C}$ und $n \in \mathbb{N}$ sei

$$J_n(a) = \begin{pmatrix} a & 1 & 0 & \cdots & 0 & 0 \\ 0 & a & 1 & \cdots & 0 & 0 \\ 0 & 0 & a & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & 1 \\ 0 & 0 & 0 & \cdots & 0 & a \end{pmatrix}$$

der Jordan-Block der Größe n zum Eigenwert a . Dann gilt

$$\begin{aligned} \exp(xJ_n(a)) &= e^{ax} \begin{pmatrix} 1 & x & \frac{x^2}{2} & \frac{x^3}{3!} & \cdots & \frac{x^{n-1}}{(n-1)!} \\ 0 & 1 & x & \frac{x^2}{2} & \cdots & \frac{x^{n-2}}{(n-2)!} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & x \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \\ &= \begin{pmatrix} e^{ax} & xe^{ax} & \frac{x^2}{2}e^{ax} & \frac{x^3}{3!}e^{ax} & \cdots & \frac{x^{n-1}}{(n-1)!}e^{ax} \\ 0 & e^{ax} & xe^{ax} & \frac{x^2}{2}e^{ax} & \cdots & \frac{x^{n-2}}{(n-2)!}e^{ax} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & xe^{ax} \\ 0 & 0 & 0 & 0 & \cdots & e^{ax} \end{pmatrix}. \end{aligned}$$

16.17 Lemma. *Das Matrixexponential einer Blockdiagonalmatrix A ist eine Blockdiagonalmatrix, deren Blöcke die Matrixexponentiale der Blöcke von A sind.*

16.18 Beispiel. Gegeben sei das Differentialgleichungssystem

$$\begin{aligned}y_1' &= 4y_4, \\y_2' &= -y_2, \\y_3' &= 6y_2 + 2y_3 + y_4, \\y_4' &= -y_1 + 4y_4.\end{aligned}$$

In Matrixschreibweise also $y' = Ay$ mit

$$A = \begin{pmatrix} 0 & 0 & 0 & 4 \\ 0 & -1 & 0 & 0 \\ 0 & 6 & 2 & 1 \\ -1 & 0 & 0 & 4 \end{pmatrix}.$$

Das charakteristische Polynom von A ist $\chi_A(\lambda) = (\lambda - 2)^3(\lambda + 1)$. Die Eigenräume von 2 und -1 sind jeweils eindimensional. Daher ist die Jordansche Normalform von A gleich

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

Um $\exp(xJ)$ zu bestimmen, schreiben wir J als Summe aus einer Diagonalmatrix und einer kommutierenden nilpotenten Matrix, also $J = D + N$ mit

$$D = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \text{und} \quad N = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Dann

$$\exp(xD) = \begin{pmatrix} e^{2x} & 0 & 0 & 0 \\ 0 & e^{2x} & 0 & 0 \\ 0 & 0 & e^{2x} & 0 \\ 0 & 0 & 0 & e^{-x} \end{pmatrix} \quad \text{und} \quad \exp(xN) = \begin{pmatrix} 1 & x & \frac{x^2}{2} & 0 \\ 0 & 1 & x & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Also

$$\exp(xJ) = \exp(xD) \circ \exp(xN) = \begin{pmatrix} e^{2x} & xe^{2x} & \frac{1}{2}x^2e^{2x} & 0 \\ 0 & e^{2x} & xe^{2x} & 0 \\ 0 & 0 & e^{2x} & 0 \\ 0 & 0 & 0 & e^{-x} \end{pmatrix}.$$

16. Lineare Differentialgleichungen mit konstanten Koeffizienten

Daraus kann man bereits ablesen: Jede Komponente einer beliebigen Lösung von $y' = Ay$ ist eine $\mathbb{C}$ -lineare Kombination der vier Funktionen e^{2x} , xe^{2x} , x^2e^{2x} und e^{-x} .

Die Transformationsmatrix T , also eine Matrix mit $A = T J T^{-1}$ bestimmt man zu

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

Wegen Korollar 16.9 braucht man T^{-1} nicht zu kennen. Ein Fundamentalsystem wird gegeben durch die Spalten von

$$T \exp(xJ) = \begin{pmatrix} 0 & -2e^{2x} & -2xe^{2x} + e^{2x} & 0 \\ 0 & 0 & 0 & -e^{-x} \\ -e^{2x} & -xe^{2x} & -\frac{x^2}{2}e^{2x} & 2e^{-x} \\ 0 & -e^{2x} & -xe^{2x} & 0 \end{pmatrix}.$$

Man kann also zu jeder Matrix das Matrixexponential ausrechnen, indem man die Jordan-Zerlegung bestimmt. Alternativ ist es aber auch möglich, einen Ansatz unter Berücksichtigung des folgenden Satzes zu machen.

16.19 Satz. Sei $A \in \mathbb{C}^{n \times n}$ mit paarweise verschiedenen, komplexen Eigenwerten $\lambda_1, \dots, \lambda_k$. Die algebraische Vielfachheit von λ_j sei α_j , die geometrische sei γ_j . Schließlich sei $\varphi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{pmatrix} : \mathbb{R} \rightarrow \mathbb{C}^n$ eine Lösung von $y' = Ay$. Dann ist jedes φ_j eine komplexe Linearkombination der Funktionen

$$f_{j,v} : x \mapsto x^v e^{\lambda_j x}, \quad 1 \leq j \leq k, \quad 0 \leq v \leq \alpha_j - \gamma_j.$$

Beweis. Das größte Jordan-Kästchen J zum Eigenwert λ_j kann nicht mehr als $\alpha_j - \gamma_j$ Hauptvektoren enthalten, die keine Eigenvektoren sind. Also hat es höchstens $\alpha_j - \gamma_j + 1$ Spalten. Die größte auftretende Potenz von x in $\exp(xJ)$ ist also $\alpha_j - \gamma_j$. $\square$

16.20 Beispiel.

$$\begin{aligned} y_1' &= y_1 + 2y_2 \\ y_2' &= -2y_1 + 5y_2. \end{aligned}$$

Dann

$$A = \begin{pmatrix} 1 & 2 \\ -2 & 5 \end{pmatrix}$$

und $\chi_A(\lambda) = (\lambda - 3)^2$. Der Ansatz lautet nun

$$\begin{aligned} y_1(x) &= \alpha_1 e^{3x} + \alpha_2 x e^{3x} \\ y_2(x) &= \alpha_3 e^{3x} + \alpha_4 x e^{3x}. \end{aligned}$$

Wir erhalten

$$\alpha_2 = -2\alpha_1 + 2\alpha_3 \quad \text{und} \quad \alpha_4 = \alpha_2.$$

Das führt zu dem Fundamentalsystem

$$\varphi_1(x) = \begin{pmatrix} e^{3x} - 2xe^{3x} \\ -2xe^{3x} \end{pmatrix} \quad \varphi_2(x) = \begin{pmatrix} 2xe^{3x} \\ e^{3x} + 2xe^{3x} \end{pmatrix}.$$

16.21 Bemerkung. Wenn A reell ist, folgt aus $Av = \lambda v$ auch $A\bar{v} = \bar{\lambda}\bar{v}$. Besitzt A also echt komplexe Eigenwerte und Eigenvektoren, so sind die jeweiligen komplex konjugierten ebenfalls Eigenwerte und Eigenvektoren. Man kombiniert dann die beiden komplexen Lösungen $e^{\lambda x}v$ und $e^{\bar{\lambda}x}\bar{v}$ zu den reellen Lösungen

$$\begin{aligned} \frac{1}{2} \left(e^{\lambda x}v + e^{\bar{\lambda}x}\bar{v} \right) &= \operatorname{Re} e^{\lambda x}v, \\ \frac{1}{2i} \left(e^{\lambda x}v - e^{\bar{\lambda}x}\bar{v} \right) &= \operatorname{Im} e^{\lambda x}v. \end{aligned}$$

Ähnliches gilt auch für Jordanblöcke.

16.22 Beispiel.

$$\begin{aligned} y_1' &= 2y_1 - y_2 \\ y_2' &= y_1 + 2y_2 \end{aligned}$$

Dann

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

Das charakteristische Polynom ist $\chi_A(x) = (x-2-i)(x-2+i)$. Also sind die Lösungen von der Form

$$\begin{aligned} y_1(x) &= \alpha_1 e^{2x} \cos(x) + \alpha_2 e^{2x} \sin(x) \\ y_2(x) &= \alpha_3 e^{2x} \cos(x) + \alpha_4 e^{2x} \sin(x). \end{aligned}$$

Die Unbekannten $\alpha_1, \dots, \alpha_4$ bestimmt man, indem man den Ansatz in die Differentialgleichung einsetzt. Man findet

$$\alpha_3 = -\alpha_2 \quad \text{und} \quad \alpha_4 = \alpha_1.$$

Damit erhält man das folgende Fundamentalsystem

$$\varphi_1(x) = \begin{pmatrix} e^{2x} \cos(x) \\ e^{2x} \sin(x) \end{pmatrix} \quad \varphi_2(x) = \begin{pmatrix} e^{2x} \sin(x) \\ -e^{2x} \cos(x) \end{pmatrix}.$$

Satz 13.10 erlaubt die Übertragung dieser Theorie auf Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten.

16. Lineare Differentialgleichungen mit konstanten Koeffizienten

16.23 Bemerkung. Zur Differentialgleichung n-ter Ordnung

$$y^{(n)} = a_0 y + a_1 y' + \dots + a_{n-1} y^{(n-1)} \quad (16.2)$$

gehört das Differentialgleichungssystem erster Ordnung $Y' = AY$ mit

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ a_0 & a_1 & a_2 & a_3 & \dots & a_{n-1} \end{pmatrix} \in \mathbb{K}^{n \times n}.$$

16.24 Lemma. $\chi_A(\lambda) = \lambda^n - a_{n-1}\lambda^{n-1} - \dots - \lambda a_1 - a_0$.

Beweis. Mit vollständiger Induktion nach n . Für $n = 1$ ist die Behauptung klar. Um von n nach $n + 1$ zu gelangen, entwickeln wir nach der ersten Spalte

$$\begin{aligned} \det \begin{pmatrix} \lambda & -1 & 0 & 0 & \dots & 0 \\ 0 & \lambda & -1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \dots & -1 \\ -a_0 & -a_1 & -a_2 & -a_3 & \dots & \lambda - a_n \end{pmatrix} \\ = \lambda \det \begin{pmatrix} \lambda & -1 & 0 & \dots & 0 \\ 0 & \lambda & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & -1 \\ -a_1 & -a_2 & -a_3 & \dots & \lambda - a_n \end{pmatrix} - (-1)^n a_0 \det \begin{pmatrix} -1 & 0 & 0 & \dots & 0 \\ \lambda & -1 & 0 & \dots & 0 \\ 0 & \lambda & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & -1 \end{pmatrix} \\ = \lambda(\lambda^n - a_{n-1}\lambda^{n-1} - \dots - a_2\lambda - a_1) - a_0. \quad \square \end{aligned}$$

16.25 Satz. Das zu der Differentialgleichung (16.2) gehörige charakteristische Polynom sei wie folgt in Linearfaktoren zerlegt

$$\lambda^n - a_{n-1}\lambda^{n-1} - \dots - a_1\lambda - a_0 = (\lambda - \lambda_1)^{v_1}(\lambda - \lambda_2)^{v_2} \dots (\lambda - \lambda_k)^{v_k},$$

wobei die λ_j paarweise verschieden sind. Dann bilden die Funktionen

$$x \mapsto x^m e^{\lambda_j x} \quad \text{mit } j = 1, \dots, k, \quad m = 0, \dots, v_j - 1 \quad (16.3)$$

ein Fundamentalsystem.

Beweis. Wegen Satz 16.19 ist der Lösungsraum ein n -dimensionaler Unterraum des von den $f_{j,v}$ aufgespannten Raums. Da die Anzahl der $f_{j,v}$ genau n beträgt, ist der Lösungsraum genau gleich dem von den $f_{j,v}$ aufgespannten Raum. $\square$

16.26 *Beispiel.* Gegeben sei die Differentialgleichung

$$y''' = y'' + y' - y.$$

Ihr charakteristisches Polynom ist $\chi(z) = (z - 1)^2(z + 1)$. Ein Fundamentalsystem wird also gegeben durch (e^x, xe^x, e^{-x}) .

16.27 *Bemerkung.* Es sei $y'' = a_1y' + a_0y$ eine lineare Differentialgleichung zweiter Ordnung, wobei $a_0, a_1 \in \mathbb{R}$. Dann sind äquivalent:

(a) Alle Lösungen der Differentialgleichung sind beschränkt auf $]0, \infty[$.

(b) $a_1 \leq 0$ und $a_0 \leq 0$, aber nicht $a_0 = a_1 = 0$.

Beweis. Das charakteristische Polynom ist $z^2 - a_1z - a_0$. Bezeichnen wir seine beiden Nullstellen mit λ_1 und λ_2 , so besagt der Satz von Vieta, dass $\lambda_1 + \lambda_2 = a_1$ und $\lambda_1\lambda_2 = -a_0$.

“ $\Rightarrow$ ”: Die Negation von (b) ist “ $a_1 > 0$ oder $a_0 > 0$ oder $a_1 = a_0 = 0$ ”. In diesen Fällen zeigen wir die Existenz einer auf $]0, \infty[$ unbeschränkten Lösung.

Sei zuerst $a_1 > 0$. Wegen der Formel von Vieta ist dann $\operatorname{Re} \lambda_1 + \operatorname{Re} \lambda_2 > 0$ und mindestens eine der beiden Lösungen $e^{\lambda_j x}$, $j = 1, 2$, ist unbeschränkt.

Sei nun $a_0 > 0$. Wenn die beiden Nullstellen nicht reell sind, dann ist $a_0 = -\lambda_1\lambda_2 = -|\lambda_1|^2 < 0$ um Widerspruch zu diesem Fall. Also sind die Nullstellen von Null verschiedene, reelle Zahlen mit unterschiedlichem Vorzeichen. Daher ist eine der beiden Lösungen $e^{\lambda_j x}$ unbeschränkt.

Im Fall $a_0 = a_1 = 0$ ist x eine Lösung.

“ $\Leftarrow$ ”: Wenn $\lambda_1, \lambda_2 \in \mathbb{R}$, dann sind wegen $a_0 \leq 0$ entweder beide ≤ 0 oder beide ≥ 0 . Wegen $a_1 \leq 0$ ist das zweite ausgeschlossen. Wenn λ_1 und λ_2 verschieden sind, dann ist $(e^{\lambda_1 x}, e^{\lambda_2 x})$ ein Fundamentalsystem aus beschränkten Funktionen. Im anderen Fall ist $(e^{\lambda_1 x}, xe^{\lambda_1 x})$ ein Fundamentalsystem. In diesem Fall ist $\lambda_1 < 0$, denn sonst $a_0 = a_1 = 0$. Daher besteht auch dieses Fundamentalsystem aus beschränkten Funktionen.

Falls λ_1 und λ_2 nicht reell sind, gilt $\operatorname{Re} \lambda_1 = \frac{1}{2}a_1 \leq 0$. Daher sind die Elemente des Fundamentalsystems $(\cos(x)e^{\lambda_1 x}, \sin(x)e^{\lambda_1 x})$ beschränkt. $\square$

16.28 *Satz.* Das zu der Differentialgleichung (16.2) gehörige charakteristische Polynom sei wie folgt in Linearfaktoren zerlegt

$$\lambda^n - a_{n-1}\lambda^{n-1} - \dots - a_1\lambda - a_0 = (\lambda - \lambda_1)^{\nu_1}(\lambda - \lambda_2)^{\nu_2} \dots (\lambda - \lambda_k)^{\nu_k},$$

wobei die λ_j paarweise verschieden sind. Weil das charakteristische Polynom reell ist, treten nicht-reelle λ_j immer zusammen mit ihrem komplex konjugierten

16. Lineare Differentialgleichungen mit konstanten Koeffizienten

auf. Daher kann man die λ_j so anordnen, dass $\lambda_{2r+1}, \dots, \lambda_k \in \mathbb{R}$ und

$$\begin{aligned}\lambda_1 &= u_1 + iv_1, \\ \lambda_2 &= u_1 - iv_1, \\ \lambda_3 &= u_2 + iv_2, \\ &\vdots \\ \lambda_{2r-1} &= u_r + iv_r, \\ \lambda_{2r} &= u_r - iv_r,\end{aligned}$$

wobei $u_1, \dots, u_r \in \mathbb{R}$ und $v_1, \dots, v_r \in \mathbb{R} \setminus \{0\}$. Dann bilden die Funktionen

$$\begin{array}{lll} x^m e^{u_j x} \cos(v_j x) & j = 1, \dots, r, & m = 0, \dots, v_j - 1 \\ x^m e^{u_j x} \sin(v_j x) & j = 1, \dots, r, & m = 0, \dots, v_j - 1 \\ x^m e^{\lambda_j x} & j = 2r + 1, \dots, k, & m = 0, \dots, v_j - 1 \end{array}$$

ein Fundamentalsystem.

16.29 Bemerkung (Ansatz für die homogene Differentialgleichung). Man kann versuchen, die Differentialgleichung (16.2) durch den Ansatz $\varphi(x) = e^{\lambda x}$ mit unbekanntem λ zu lösen. Es stellt sich heraus, dass φ genau dann die Differentialgleichung löst, wenn λ eine Nullstelle des charakteristischen Polynoms ist. In den Fällen, in denen das charakteristische Polynom mehrfache Nullstellen aufweist, kommt man freilich mit diesem Ansatz nicht weiter.

16.30 Satz (Ansatz für die inhomogene Differentialgleichung). Gegeben sei die inhomogene Differentialgleichung n -ter Ordnung

$$y^{(n)} = a_0 y + a_1 y' + \dots + a_{n-1} y^{(n-1)} + P(x) e^{\lambda x},$$

wobei P ein Polynom vom Grad d ist. Es sei k die Vielfachheit von λ als Nullstelle des charakteristischen Polynoms χ der Differentialgleichung, wobei $k = 0$, wenn $\chi(\lambda) \neq 0$. Dann gibt es eine spezielle Lösung der Form

$$\varphi(x) = Q(x) e^{\lambda x},$$

wobei Q ein Polynom vom Grad höchstens $d + k$ ist.

Beweis. Das wird im Forster, Analysis 2, §13, Satz 4 bewiesen. □

Bemerkung. Der Satz gilt auch für komplexe λ . Auf diese Weise kann man auch Inhomogenitäten mit trigonometrischen Funktionen behandeln.

Wenn man nicht komplex rechnen will, muss man für Inhomogenitäten der Form $\cos(ax) e^{\lambda x}$ einen Ansatz der Form $(Q_1(x) \cos(ax) + Q_2(x) \sin(ax)) e^{\lambda x}$ machen und entsprechend für den Sinus.

16.31 *Beispiel.* (a) $y'' = 2y' - y + 3 \cos(x)$. Das charakteristische Polynom ist $\chi(\lambda) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$. Daher bilden $\varphi_1(x) = e^x$ und $\varphi_2(x) = xe^x$ ein Fundamentalsystem für die Lösungen der homogenen Differentialgleichung.

Wir bestimmen eine spezielle Lösung der inhomogenen Differentialgleichung durch Ansatz $\psi(x) = Q_1(x) \cos(x) + Q_2(x) \sin(x)$. Wir dürfen Q_1 und Q_2 vom Grad 0 wählen, da 1 keine Nullstelle von χ ist. Der Ansatz lautet also $\psi(x) = a \cos(x) + b \sin(x)$. Dann $\psi'(x) = -a \sin(x) + b \cos(x)$ und $\psi'' = -\psi$. Eingesetzt in die Differentialgleichung erhalten wir

$$-a \cos(x) - b \sin(x) = (2b - a + 3) \cos(x) + (-2a - b) \sin(x).$$

Also ist $\psi(x) = -\frac{3}{2} \sin(x)$ eine spezielle Lösung der inhomogenen Differentialgleichung und die allgemeine Lösung der inhomogenen Differentialgleichung hat die Gestalt

$$\varphi(x) = C_1 e^x + C_2 x e^x - \frac{3}{2} \sin(x), \quad C_1, C_2 \in \mathbb{R}.$$

(b) $y'' = 2y' - y + e^x$. In diesem Fall macht man den Ansatz

$$\psi(x) = ax^2 e^x.$$

Koeffizientenvergleich liefert schließlich $a = \frac{1}{2}$.

In den Fällen, wo der Ansatz nach Satz 16.30 möglich ist, ist er auch das empfohlene Verfahren.

17. Die Eulersche Differentialgleichung

17.1 Definition. Differentialgleichungen der Form

$$y^{(n)} = a_{n-1}x^{-1}y^{(n-1)} + a_{n-2}x^{-2}y^{(n-2)} + \dots + a_1x^{1-n}y' + a_0x^{-n}y, \quad x > 0, \quad (17.1)$$

heißen *Eulersche Differentialgleichungen*.

Wir setzen

$$L(y)(x) = x^n y^{(n)}(x) - \sum_{j=0}^{n-1} a_j x^j y^{(j)}(x).$$

Dann ist (17.1) äquivalent zu $L(y) = 0$.

17.2 Satz. Die Substitution $x = e^t$ verwandelt die Eulersche Differentialgleichung (17.1) in eine lineare Differentialgleichung mit konstanten Koeffizienten.

Beweis. Wir setzen $u(t) = y(e^t)$ und zeigen mit vollständiger Induktion für $k = 0, \dots, n$ die Existenz von $b_{k,0}, \dots, b_{k,k-1}$, so dass

$$u^{(k)}(t) = e^{kt} y^{(k)}(e^t) + \sum_{j=0}^{k-1} b_{k,j} u^{(j)}(t). \quad (17.2)$$

Für $k = 0$ ist das jedenfalls wahr. Wir schließen von k nach $k + 1$.

$$\begin{aligned} u^{(k+1)}(t) &= k e^{kt} y^{(k)}(e^t) + e^{(k+1)t} y^{(k+1)}(e^t) + \sum_{j=0}^{k-1} b_{k,j} u^{(j+1)}(t) \\ &= e^{(k+1)t} y^{(k+1)}(e^t) + k u^{(k)}(t) - \sum_{j=0}^{k-1} k b_{k,j} u^{(j)}(t) + \sum_{j=0}^{k-1} b_{k,j} u^{(j+1)}(t). \end{aligned}$$

Da diese Formel die geforderte Gestalt ist, ist die Zwischenbehauptung gezeigt.

Aus (17.2) und (17.1) folgt

$$u^{(n)}(t) = e^{nt} y^{(n)}(e^t) + \sum_{j=0}^{n-1} b_{n,j} u^{(j)}(t) = \sum_{j=0}^{n-1} a_j e^{jt} y^{(j)}(e^t) + \sum_{j=0}^{n-1} b_{n,j} u^{(j)}(t).$$

Nochmalige Anwendung von (17.2) zeigt, dass diese Differentialgleichung linear mit konstanten Koeffizienten ist. $\square$

17. Die Eulersche Differentialgleichung

17.3 *Beispiel.* Gegeben sei die Eulersche Differentialgleichung

$$x^2 y'' = xy' + y.$$

Für $u(t) = y(e^t)$ bekommen wir

$$\begin{aligned} u'(t) &= e^t y'(e^t) \quad \text{und} \\ u''(t) &= e^t y'(e^t) + e^{2t} y''(e^t) = 2e^t y'(e^t) + y(e^t) = 2u'(t) + u(t). \end{aligned}$$

Ein Fundamentalsystem für die lineare Differentialgleichung $u'' = 2u' + u$ ist gegeben durch $(e^{(1+\sqrt{2})t}, e^{(1-\sqrt{2})t})$. Also ist $(x^{1+\sqrt{2}}, x^{1-\sqrt{2}})$ ein Fundamentalsystem der Eulerschen Differentialgleichung.

18. Der Banachsche Fixpunktsatz

18.1 Definition. Sei X eine Menge und $G: X \rightarrow X$ eine Abbildung. Ein Element $x \in X$ heißt *Fixpunkt* von G , wenn $G(x) = x$.

Der folgende Satz wird gerne in der Analysis I als Übungsaufgabe gestellt.

18.2 Satz. *Es sei $I \subset \mathbb{R}$ ein kompaktes Intervall und $f: I \rightarrow \mathbb{R}$ stetig mit $f(I) \subseteq I$. Dann besitzt f einen Fixpunkt.*

Beweis. Setze $g(x) = f(x) - x$. Dann $g(a) \geq 0$ und $g(b) \leq 0$ und die Behauptung ergibt sich sofort aus dem Zwischenwertsatz. $\square$

18.3 Definition. Sei X ein metrischer Raum. Eine Abbildung $G: X \rightarrow X$ heißt *kontrahierend*, wenn es ein $q < 1$ gibt, so dass

$$d(G(x), G(y)) \leq qd(x, y) \quad \text{für alle } x, y \in X.$$

18.4 Bemerkung. Eine kontrahierende Abbildung ist Lipschitz-stetig, speziell also stetig.

18.5 Satz (Banachscher Fixpunktsatz). *Sei X ein vollständiger metrischer Raum und $G: X \rightarrow X$ eine kontrahierende Abbildung. Dann besitzt G genau einen Fixpunkt.*

Idee. Wähle $x_0 \in X$ beliebig und definiere rekursiv $x_{n+1} = G(x_n)$. Man zeigt, dass $(x_n)_{n \in \mathbb{N}}$ eine Cauchyfolge ist. $\square$

Beweis. Eindeutigkeit: Seien $x, y \in X$ zwei Fixpunkte. Dann

$$d(x, y) = d(G(x), G(y)) \leq qd(x, y).$$

Das kann nur sein, wenn $d(x, y) = 0$.

Existenz: Sei $x_0 \in X$ beliebig gewählt. Definiere rekursiv für jedes $n \in \mathbb{N}_0$

$$x_{n+1} = G(x_n).$$

Dann zeigt man induktiv

$$d(x_j, x_{j+1}) \leq qd(x_{j-1}, x_j) \leq q^2d(x_{j-2}, x_{j-1}) \leq \dots \leq q^jd(x_0, x_1).$$

18. Der Banachsche Fixpunktsatz

Falls $x_1 = x_0$, so ist der Fixpunkt gefunden. Andernfalls zeigen wir, dass die Folge $(x_n)_{n \in \mathbb{N}}$ eine Cauchyfolge ist. Sei dazu $\epsilon > 0$ beliebig vorgegeben. Wegen $q < 1$ existiert dazu ein $N \in \mathbb{N}$ mit

$$q^N < \frac{\epsilon(1-q)}{d(x_0, x_1)}.$$

Seien nun $n, m \geq N$. Wir nehmen an, dass $n > m$.

$$d(x_m, x_n) \leq \sum_{j=m}^{n-1} d(x_j, x_{j+1}) \leq \sum_{j=m}^{n-1} q^j d(x_0, x_1) \leq d(x_0, x_1) \sum_{j=m}^{\infty} q^j = \frac{q^m d(x_0, x_1)}{1-q} < \epsilon.$$

Da der Raum X vollständig ist, existiert $\lim_{n \rightarrow \infty} x_n = x$. Daraus folgt

$$x = \lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} G(x_n) = G(x). \quad \square$$

Häufig ist die folgende Variante praktischer.

18.6 Satz. Sei X ein vollständiger metrischer Raum, seien $x_0 \in X$ und $R > 0$. Sei $G: \overline{B}_R(x_0) \rightarrow X$ eine Abbildung. Es gebe ein $q < 1$, so dass

$$(a) \quad d(G(x), G(y)) \leq qd(x, y) \text{ für alle } x, y \in \overline{B}_R(x_0),$$

$$(b) \quad d(G(x_0), x_0) \leq R(1-q).$$

Dann gibt es genau ein $x \in \overline{B}_R(x_0)$ mit $G(x) = x$.

Beweis. Wir zeigen $G(\overline{B}_R(x_0)) \subset \overline{B}_R(x_0)$ und wenden den vorhergehenden Satz an. Für $y \in \overline{B}_R(x_0)$ gilt

$$d(G(y), x_0) \leq d(G(y), G(x_0)) + d(G(x_0), x_0) \leq qd(y, x_0) + R(1-q) \leq qR + R - qR = R. \quad \square$$

18.7 Beispiel. Sei $G: [1, 2] \rightarrow \mathbb{R}$, $G(x) := x + \cos(x)$. Wegen $G'(x) = 1 - \sin(x)$ ist G streng monoton wachsend. Wegen $0 \leq \cos(1) \leq 1$ und $-1 \leq \cos(2) \leq 0$ folgt aus Satz 18.2 die Existenz eines Fixpunkts, der wegen der strengen Monotonie eindeutig ist. Es ist klar, dass es sich bei diesem Fixpunkt um den Wert $\frac{\pi}{2}$ handelt.

Für $x, y \in I := [1, 2]$ folgt aus dem Mittelwertsatz die Existenz eines $\xi \in]1, 2[$ mit

$$|G(x) - G(y)| = |x - y|(1 - \sin \xi).$$

Der Sinus wächst monoton auf $[1, \frac{\pi}{2}]$ und fällt monoton auf $[\frac{\pi}{2}, 2]$. Daher ist $\sin(\xi) \geq \min(\sin(1), \sin(2))$ auf $[1, 2]$ und G ist eine Kontraktion. Der Banachsche Fixpunktsatz zeigt, dass die Folge

$$x_1 = 1, \quad x_{n+1} = x_n + \cos x_n$$

gegen $\frac{\pi}{2}$ konvergiert.

Eine nicht-konstruktive Verallgemeinerung von Satz 18.6 ist der Brouwersche Fixpunktsatz, den ich voraussichtlich in der Analysis III beweisen werde (siehe Zeidler, Nonlinear Functional Analysis and Applications I, Proposition 2.5):

18.8 Theorem. *Es sei $M \subset \mathbb{R}^n$ konvex, kompakt und nicht leer, und es sei $f: M \rightarrow M$ stetig. Dann besitzt f einen Fixpunkt.*

19. Existenz- und Eindeutigkeitsätze

Wir formulieren die Anfangswertaufgabe $y' = f(x, y)$, $y(x_0) = y_0$ um in eine Fixpunktgleichung in einem Banachraum von Funktionen.

19.1 Lemma. *Sei $J \subseteq \mathbb{R}$ ein offenes Intervall, sei $H \subseteq \mathbb{R}^n$ offen und sei $f: J \times H \rightarrow \mathbb{R}^n$ stetig. Sei $(x_0, y_0) \in J \times H$. Für eine stetige Abbildung $\varphi: J \rightarrow \mathbb{R}^n$ mit $\varphi(J) \subseteq H$ definieren wir eine stetige Abbildung $G(\varphi): J \rightarrow \mathbb{R}^n$ durch*

$$G(\varphi)(x) = y_0 + \int_{x_0}^x f(t, \varphi(t)) dt.$$

Dann ist φ genau dann eine Lösung der Anfangswertaufgabe $y' = f(x, y)$, $y(x_0) = y_0$, wenn $G(\varphi) = \varphi$.

Beweis. Wenn φ eine Lösung der AWA ist, dann muss gezeigt werden, dass

$$\varphi(x) = y_0 + \int_{x_0}^x f(t, \varphi(t)) dt \quad \text{für alle } x \in J.$$

Dazu differenzieren wir beide Seiten:

$$\varphi'(x) = f(x, \varphi(x)).$$

Da beide Seiten außerdem an der Stelle $x = x_0$ übereinstimmen, ist diese Richtung gezeigt.

Wenn $G(\varphi) = \varphi$, dann müssen wir zeigen, dass φ die AWA löst. Die Differentialgleichungen ergibt sich wieder durch Differentiation und die AB durch Einsetzen. $\square$

19.2 Definition. Sei $I \subset \mathbb{R}$ ein kompaktes Intervall. Definiere

$$C(I, \mathbb{K}^n) = \{f: I \rightarrow \mathbb{K}^n \mid f \text{ stetig}\}$$

und versehe ihn mit der Norm

$$\|f\|_\infty = \max\{\|f(x)\|_\infty \mid x \in I\}.$$

Man schreibt $C(I)$ anstelle von $C(I, \mathbb{R})$.

Bemerkung. Eine Folge $(f_n)_{n \in \mathbb{N}}$ in $C(I)$ konvergiert genau dann in $C(I)$ gegen f , wenn sie gleichmäßig gegen f konvergiert.

19. Existenz- und Eindeutigkeitssätze

19.3 Satz. $C(I, \mathbb{K}^n)$ ist ein Banachraum.

Beweis. Wir brauchen das nur für $\mathbb{K} = \mathbb{R}$ zu machen. Es sei $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in $C(I, \mathbb{R}^m)$. Sei $f_{n,j}$ die j -te Komponentenfunktion von f_n , also

$$f_n(x) = (f_{n,1}(x), \dots, f_{n,m}(x)).$$

Für jedes j ist $(f_{n,j})_{n \in \mathbb{N}}$ eine Cauchyfolge in $C(I)$, und für jedes $x \in I$ ist $(f_{n,j}(x))_{n \in \mathbb{N}}$ eine Cauchyfolge in $\mathbb{R}$. Daher existiert $\lim_{n \rightarrow \infty} f_{n,j}(x) =: F_j(x)$.

Wir zeigen, dass die Folge $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen F_j konvergiert. Sei dazu $\epsilon > 0$ beliebig vorgegeben. Da $(f_{n,j})_{n \in \mathbb{N}}$ eine Cauchyfolge ist, gibt es $N \in \mathbb{N}$, so dass für alle $n, m \geq N$ gilt $\|f_{j,n} - f_{j,m}\|_\infty < \frac{\epsilon}{2}$. Sei $x \in I$ beliebig. Dann existiert $m \geq N$, so dass $|f_{j,m}(x) - F_j(x)| < \frac{\epsilon}{2}$. Daher gilt für $n \geq N$

$$|f_{j,n}(x) - F_j(x)| \leq |f_{j,n}(x) - f_{j,m}(x)| + |f_{j,m}(x) - F_j(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Da diese Abschätzung für alle $x \in I$ gilt, haben wir $\|f_{j,n} - F_j\|_\infty < \epsilon$ gezeigt.

Wegen der gleichmäßigen Konvergenz von $(f_{j,n})_{n \in \mathbb{N}}$ gegen F_j folgt aus Analysis I, Satz 17.4, die Stetigkeit von F_j . Daher ist $F = (F_1, \dots, F_m)$ in $C(I, \mathbb{R}^m)$. Da für jedes j die Folge $(f_{j,n})_{n \in \mathbb{N}}$ in $C(I)$ gegen F_j konvergiert, konvergiert $(f_n)_{n \in \mathbb{N}}$ in $C(I, \mathbb{R}^m)$ gegen F . $\square$

19.4 Satz (Lokaler Existenz- und Eindeutigkeitssatz). *Sei $U \subseteq \mathbb{R}^{n+1}$ offen, sei $f: U \rightarrow \mathbb{R}^n$ stetig (in beiden Argumenten) und lokal Lipschitz-stetig im zweiten Argument und sei $(x_0, y_0) \in U$. Dann existieren ein kompaktes Intervall $I \subset \mathbb{R}$ und ein Kompaktum $H \subset \mathbb{R}^n$, so dass $I \times H \subset U$ eine Umgebung von (x_0, y_0) ist und so dass es für jedes kompakte Intervall J mit $x_0 \in J \subseteq I$ genau eine stetige Abbildung $\varphi: J \rightarrow H$ gibt, welche die Fixpunktgleichung*

$$\varphi(x) = y_0 + \int_{x_0}^x f(t, \varphi(t)) dt \quad \text{für alle } x \in J$$

erfüllt.

Idee. Man wendet den Banachschen Fixpunktsatz auf die Integralgleichung an. Um die Kontraktionseigenschaft sicherzustellen, muss das Intervall gegebenenfalls verkleinert werden. $\square$

Beweis. Es gibt eine kompakte Umgebung $I_1 \times H_1 \subset U$ von (x_0, y_0) , so dass f auf $I_1 \times H_1$ Lipschitz-stetig in der zweiten Variablen mit Lipschitz-Konstante C ist. Sei $M = 1 + \max\{\|f(t, y_0)\|_\infty \mid t \in I_1\}$. Wähle $\delta_1 > 0$ so klein, dass $H := \overline{B}_{\delta_1}(y_0) \subseteq H_1$. Wähle nun $\epsilon > 0$ so klein, dass

$$(a) \quad q := \epsilon C < 1,$$

$$(b) \quad I := [x_0 - \epsilon, x_0 + \epsilon] \subseteq I_1,$$

$$(c) \quad \epsilon < \frac{\delta_1(1-q)}{M}.$$

Setze

$$R = \frac{\epsilon M}{1-q}.$$

Dann gilt $0 < R < \delta_1$. Ist nun ein kompaktes Intervall J mit $x_0 \in J \subseteq I$ gegeben, so erklären wir $\varphi_0 \in C(J, \mathbb{R}^n)$ durch $\varphi_0(x) = y_0$ für alle x und definieren

$$G: \overline{B}_R(\varphi_0) \rightarrow C(J, \mathbb{R}^n), \quad G(\varphi)(x) = y_0 + \int_{x_0}^x f(t, \varphi(t)) dt.$$

Wir zeigen nun, dass G kontrahierend mit Kontraktionskonstante q ist. Dazu seien $\varphi, \psi \in \overline{B}_R(\varphi_0)$ vorgeben. Dann gilt für jedes $x \in J$ wegen $R < \delta_1$ und der Lipschitz-Stetigkeit im zweiten Argument

$$\begin{aligned} \|G(\varphi)(x) - G(\psi)(x)\|_\infty &\leq \int_{x_0}^x \|f(t, \varphi(t)) - f(t, \psi(t))\|_\infty dt \\ &\leq C \int_{x_0}^x \|\varphi(t) - \psi(t)\|_\infty dt = C|x - x_0| \|\varphi - \psi\|_\infty \leq C\epsilon \|\varphi - \psi\|_\infty, \end{aligned}$$

also $\|G(\varphi) - G(\psi)\|_\infty \leq q \|\varphi - \psi\|_\infty$. Ferner gilt für jedes $x \in J$

$$\|G(\varphi_0)(x) - \varphi_0(x)\|_\infty \leq \int_{x_0}^x \|f(t, y_0)\|_\infty dt \leq \epsilon M = R(1-q).$$

Damit sind die Voraussetzungen (a) und (b) von Satz 18.6 nachgewiesen, und wir erhalten einen eindeutig bestimmten Fixpunkt φ von G . $\square$

Damit kann man nun den Satz von Picard-Lindelöf beweisen, indem man lokale Lösungen verheftet.

Theorem (Picard-Lindelöf). *Sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ offen, sei $f: U \rightarrow \mathbb{R}^n$ stetig (in beiden Argumenten) und lokal Lipschitz-stetig im zweiten Argument, und sei $(x_0, y_0) \in U$. Dann existieren ein offenes Intervall I mit $x_0 \in I$ und eine Lösung $\varphi: I \rightarrow \mathbb{R}^n$ der Differentialgleichung $y' = f(x, y)$ mit folgenden Eigenschaften*

- (a) $\varphi(x_0) = y_0$.
- (b) Ist $\psi: J \rightarrow \mathbb{R}^n$ ebenfalls eine Lösung der Differentialgleichung $y' = f(x, y)$ mit $\psi(x_0) = y_0$, so gelten $J \subseteq I$ und $\psi = \varphi|_J$.

Beweis von Satz 13.6. Wir setzen

$$L = \bigcup \{J \text{ offenes Intervall} \mid x_0 \in J\}$$

und es gibt Lösung $\varphi: J \rightarrow \mathbb{R}^n$ der AWA $y' = f(x, y)$, $y(x_0) = y_0$.

Dann ist L ein offenes Intervall mit $x_0 \in L$. Um auf L eine Lösung φ der Anfangswertaufgabe erklären zu können, muss man sich folgendes überlegen:

19. Existenz- und Eindeutigkeitssätze

Sind $\varphi_1: J_1 \rightarrow \mathbb{R}^n$ und $\varphi_2: J_2 \rightarrow \mathbb{R}^n$ Lösungen der Anfangswertaufgabe, so stimmen φ_1 und φ_2 auf $J_1 \cap J_2$ überein.

Wir nehmen zum Widerspruch an, dass φ_1 und φ_2 sich auf $\{x \in J_1 \cap J_2 \mid x > x_0\}$ unterscheiden (das “>” ist dabei o.B.d.A.). Sei $t_0 := \inf\{x > x_0 \mid \varphi_1(x) \neq \varphi_2(x)\}$. Wegen der Stetigkeit von φ_1 und φ_2 gilt $\varphi_1(t_0) = \varphi_2(t_0)$. Also lösen beide Funktionen die Anfangswertaufgabe $y(t_0) = \varphi_1(t_0)$. Wegen Satz 19.4 stimmen sie daher auf einer Umgebung von t_0 überein. Das ist ein Widerspruch zur Wahl von x_0 .

Damit ist die Zwischenbehauptung gezeigt. Beide Teile der Behauptung sind damit klar. $\square$

Die Aussage des folgenden Satzes findet man in der Literatur in der Formulierung: “Es gibt eine Lösung der Anfangswertaufgabe, die dem Rand von U beliebig nahe kommt.”

19.5 Satz. Sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ offen und sei $f: U \rightarrow \mathbb{R}^n$ stetig und lokal Lipschitz-stetig im zweiten Argument. Sei $(x_0, y_0) \in U$, und sei $\varphi: I \rightarrow \mathbb{R}^n$ eine Lösung der Anfangswertaufgabe $y' = f(x, y)$, $y(x_0) = y_0$, mit maximalem Definitionsbereich I . Sei schließlich

$$\mathcal{G} = \{(x, \varphi(x)) \mid x \in I, x \geq x_0\}$$

die rechte Seite des Graphen von φ . Dann ist $\overline{\mathcal{G}}$ keine kompakte Teilmenge von U .

Die analoge Aussage gilt für die linke Seite des Graphen.

Beweis. Sei $I =]a, b[$. Für $b = \infty$ ist die Behauptung klar. Sei also $b < \infty$. Wir nehmen als Widerspruchssannahme an, dass $\overline{\mathcal{G}}$ eine kompakte Teilmenge von U ist. Dann ist

$$M = \max\{|f(x, y)| \mid (x, y) \in \overline{\mathcal{G}}\}$$

endlich. Wir behaupten, dass $\lim_{x \nearrow b} \varphi(x)$ existiert. Sei dazu $(x_m)_{m \in \mathbb{N}}$ eine beliebige Folge in I mit $\lim_{m \rightarrow \infty} x_m = b$. Wir zeigen, dass $(\varphi(x_m))_{m \in \mathbb{N}}$ eine Cauchyfolge ist. Dazu seien $k, m \in \mathbb{N}$ mit $x_k < x_m$ gegeben. Dann folgt aus dem Mittelwertsatz

$$\begin{aligned} \|\varphi(x_m) - \varphi(x_k)\| &\leq |x_m - x_k| \max_{x_k \leq x \leq x_m} \|\varphi'(x)\| \\ &= |x_m - x_k| \max_{x_k \leq x \leq x_m} \|f(x, \varphi(x))\| \leq M|x_m - x_k|. \end{aligned}$$

Da $(x_m)_{m \in \mathbb{N}}$ eine Cauchyfolge ist, zeigt diese Abschätzung, dass auch $(\varphi(x_m))_{m \in \mathbb{N}}$ eine ist. Da dies für alle Cauchyfolgen gilt, existiert $\eta = \lim_{x \nearrow b} \varphi(x)$. Aus dem lokalen Existenz- und Eindeutigkeitssatz folgt, dass es ein offenes Intervall J und eine differenzierbare Abbildung $\psi: J \rightarrow \mathbb{R}^n$ gibt, so dass $\psi'(x) = f(x, \psi(x))$ für alle $x \in J$ und $\psi(b) = \eta$. Aus der Definition von η folgt sofort die Stetigkeit von

$$\rho: I \cup J \rightarrow V, \quad \rho(x) = \begin{cases} \varphi(x), & x < b \\ \psi(x), & x \geq b. \end{cases}$$

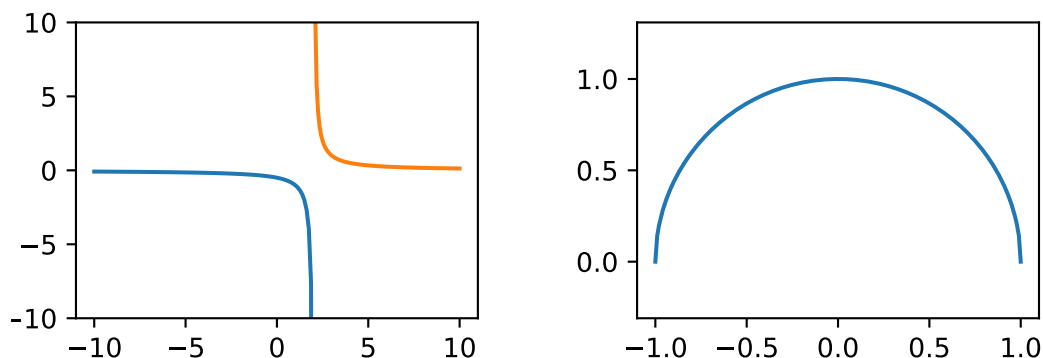


Abbildung 19.1.: Graphen zu Beispiel 19.6

Aus dem Verheftungslemma 12.7 folgt, dass ρ die Anfangswertaufgabe löst. Dies steht im Widerspruch zur Maximalität von I . $\square$

19.6 Beispiel. (a) Betrachte die Differentialgleichung

$$y' = -y^2.$$

Ihre Lösungsgesamtheit besteht aus den folgenden drei Funktionsklassen

$$\begin{aligned} \varphi:]-\infty, C[&\rightarrow \mathbb{R}, & x &\mapsto \frac{1}{x - C}, & C &\in \mathbb{R}, \\ \varphi:]C, \infty[&\rightarrow \mathbb{R}, & x &\mapsto \frac{1}{x - C}, & C &\in \mathbb{R}, \\ \varphi: \mathbb{R} &\rightarrow \mathbb{R}, & x &\mapsto 0. \end{aligned}$$

Man sieht sofort, dass in allen Fällen weder der Abschluss der rechten noch der linken Seite des Graphen kompakt in $U = \mathbb{R}^2$ ist.

(b) Betrachte die Anfangswertaufgabe

$$y' = -\frac{x}{y}, \quad y(0) = 1.$$

Die Lösung wurde in der Übung bestimmt. Es handelt sich um

$$\varphi:]-1, 1[\rightarrow \mathbb{R}, \quad \varphi(x) = \sqrt{1 - x^2}.$$

Der Abschluss des Graphen ist zwar kompakt, liegt aber nicht in U .
Die Graphen zeigt Abbildung 19.1.

19.7 Lemma. Es seien $a, M > 0$ und $u: [0, a] \rightarrow \mathbb{R}$ stetig mit

$$u(0) = 0 \quad \text{und} \quad u(x) \leq M \int_0^x u(t) dt \quad \text{für } 0 \leq x \leq a.$$

Dann gilt $u(x) \leq 0$ für alle $x \in [0, a]$.

19. Existenz- und Eindeigkeitssätze

Beweis. Angenommen, die Behauptung sei falsch. Wegen der Stetigkeit von u gilt dann $\xi := \inf\{x \leq a \mid u(x) > 0\} < a$ und es gibt ein $x_0 > \xi$ mit $x_0 - \xi < \frac{1}{M}$, so dass $u(x_0) > 0$. Wähle y mit $\xi < y \leq x_0$ und $u(y) = \max u([\xi, x_0])$. Dann

$$u(y) \leq M \int_0^y u(t) dt \leq M \int_\xi^y u(t) dt \leq M(y - \xi)u(y).$$

Wegen $u(y) > 0$ ist das ein Widerspruch zu $(y - \xi)M < 1$. $\square$

19.8 Theorem (Lemma von Gronwall). *Sei $c > 0$ und sei $I = [0, c]$. Ferner seien $w, \varphi, \psi: I \rightarrow [0, \infty[$ stetige Funktionen mit*

$$w(x) \leq \varphi(x) + \int_0^x \psi(t)w(t) dt \quad \text{für alle } x \in I.$$

Dann gilt für alle $x \in I$

$$w(x) \leq \varphi(x) + \int_0^x \psi(t)\varphi(t) \exp\left(\int_t^x \psi(s) ds\right) dt.$$

Beweis. Im Spezialfall $\varphi \equiv 0$ handelt es sich um das vorhergehende Lemma. Um im allgemeinen Fall den Term φ loszuwerden, bestimmen wir zuerst eine stetige Lösung u der Integralgleichung

$$u(x) = \varphi(x) + \int_0^x \psi(t)u(t) dt, \quad x \in I. \quad (19.1)$$

Wenn φ differenzierbar wäre, könnten wir ableiten und die Differentialgleichung lösen. Stattdessen setzen wir $v = u - \varphi$ und suchen eine stetige Lösung v der Integralgleichung

$$v(x) = \int_0^x \psi(t)\varphi(t) dt + \int_0^x \psi(t)v(t) dt.$$

Dieses v ist dann automatisch differenzierbar. Die Integralgleichung für v ist daher äquivalent zu

$$v'(x) = \psi(x)\varphi(x) + \psi(x)v(x) \quad \text{für } 0 < x < c \quad \text{und} \quad v(0) = 0.$$

Die Differentialgleichung für v ist linear inhomogen. Die Lösung der zugehörigen homogenen Gleichung ist

$$\exp\left(\int_0^x \psi(s) ds\right).$$

Daher ist eine spezielle Lösung der inhomogenen Gleichung

$$\exp\left(\int_0^x \psi(s) ds\right) \int_0^x \psi(t)\varphi(t) \exp\left(-\int_0^t \psi(s) ds\right) dt = \int_0^x \psi(t)\varphi(t) \exp\left(\int_t^x \psi(s) ds\right) dt.$$

Da diese Lösung auch noch die Anfangsbedingung erfüllt, ist gezeigt, dass

$$u(x) = \varphi(x) + \int_0^x \psi(t) \varphi(t) \exp\left(\int_t^x \psi(s) ds\right) dt$$

die Integralgleichung (19.1) erfüllt.

Setzt man für $x \in I$

$$\rho(x) := \varphi(x) + \int_0^x \psi(t) w(t) dt.$$

Nach Voraussetzung gilt $w \leq \rho$ und daher

$$\rho(x) \leq \varphi(x) + \int_0^x \psi(t) \rho(t) dt.$$

Für $h = \rho - u$ gilt wegen (19.1)

$$h(x) \leq \varphi(x) + \int_0^x \psi(t) \rho(t) dt - \varphi(x) - \int_0^x \psi(t) u(t) dt = \int_0^x \psi(t) h(t) dt, \quad h(0) = 0.$$

Aus dem vorigen Lemma folgt $h \leq 0$, wobei man $M = \sup \psi(I)$ setzt. $\square$

19.9 Theorem (Globaler Existenz- und Eindeigkeitssatz). *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und sei $f: I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig. Für jedes kompakte Teilintervall $K \subset I$ sei $f|_{K \times \mathbb{R}^n}$ Lipschitz-stetig im zweiten Argument. Sei ferner $(x_0, y_0) \in I \times \mathbb{R}^n$. Dann gibt es eine eindeutig bestimmte Lösung $\varphi: I \rightarrow \mathbb{R}^n$ der Anfangswertaufgabe $y' = f(x, y)$, $y(x_0) = y_0$.*

Beweis. Wegen des lokalen Existenz- und Eindeigkeitssatzes 13.6 existiert eine Lösung $\varphi: J \rightarrow \mathbb{R}^n$ der Anfangswertaufgabe mit maximalem Definitionsbereich. Wir müssen $J = I$ zeigen. Wir nehmen zum Widerspruch an, dass $b := \sup J < \sup I$. Wegen Satz 19.5 ist der Abschluss von

$$\mathcal{G} := \{\varphi(x) \mid x_0 < x < b\}$$

nicht kompakt in $I \times \mathbb{R}^n$. Das ist nur möglich, wenn $\mathcal{G}$ unbeschränkt ist. Wir verwenden nun das Gronwall-Lemma, um eine Schranke für $\mathcal{G}$ zu finden.

Mit C werde eine Lipschitz-Konstante für $f|_{[x_0, b] \times \mathbb{R}^n}$ bezeichnet, und es sei $M := \sup_{x_0 \leq t \leq b} \|f(t, 0)\|$. Wegen $\varphi(x) = y_0 + \int_{x_0}^x f(t, \varphi(t)) dt$ folgt

$$\begin{aligned} \|\varphi(x)\| &\leq \|y_0\| + \int_{x_0}^x \|f(t, \varphi(t))\| dt \leq \|y_0\| + \int_{x_0}^x M + C \|\varphi(t)\| dt \\ &= \|y_0\| + M(x - x_0) + \int_{x_0}^x C \|\varphi(t)\| dt. \end{aligned}$$

Mit dem Gronwall-Lemma folgt

$$\begin{aligned} \|\varphi(x)\| &\leq \|y_0\| + M(x - x_0) + \int_{x_0}^x C(\|y_0\| + M(t - x_0)) \exp\left(\int_t^x C ds\right) dt \\ &\leq \|y_0\| + M(b - x_0) + (C(\|y_0\|(b - x_0) + M(b - x_0)^2)) \exp(C(b - x_0)). \quad \square \end{aligned}$$

19. Existenz- und Eindeutigkeitssätze

Satz 14.2 folgt nun sofort.

Satz. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $A: I \rightarrow \mathbb{K}^{n \times n}$ und $b: I \rightarrow \mathbb{K}^n$ seien stetige Abbildungen. Sei $(x_0, y_0) \in I \times \mathbb{K}^n$. Dann besitzt die Anfangswertaufgabe $y' = A(x)y + b(x)$, $y(x_0) = y_0$, eine eindeutig bestimmte Lösung $\varphi: I \rightarrow \mathbb{K}^n$.

20. Abhängigkeit der Lösungen von Parametern

20.1 Definition. Sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ und $f: U \rightarrow \mathbb{R}^n$ sei stetig. Sei $\epsilon \geq 0$. Ist I ein offenes Intervall und $\varphi: I \rightarrow \mathbb{R}^n$ differenzierbar, so heißt φ *Näherungslösung* der Differentialgleichung $y' = f(x, y)$ zur Näherung ϵ , falls für alle $x \in I$ gilt

$$(a) \quad (x, \varphi(x)) \in U,$$

$$(b) \quad \|\varphi'(x) - f(x, \varphi(x))\| \leq \epsilon.$$

20.2 Satz. Sei $J \subseteq \mathbb{R}$ ein offenes Intervall, sei $V \subseteq \mathbb{R}^n$ offen, und sei $f: J \times V \rightarrow \mathbb{R}^n$ stetig und Lipschitz-stetig im zweiten Argument. Es gibt also ein $L > 0$ mit

$$\|f(x, y_1) - f(x, y_2)\| \leq L\|y_1 - y_2\| \quad \text{für alle } x \in J, y_1, y_2 \in V.$$

Sei $I \subseteq J$ ein offenes Intervall, sei $x_0 \in I$, und seien $\varphi_1, \varphi_2: I \rightarrow \mathbb{R}^n$ Näherungslösungen von $y' = f(x, y)$ zur Näherung ϵ_1 bzw. ϵ_2 . Dann gilt für jedes $x \in I$

$$\|\varphi_1(x) - \varphi_2(x)\| \leq \|\varphi_1(x_0) - \varphi_2(x_0)\| e^{L|x-x_0|} + \frac{\epsilon_1 + \epsilon_2}{L} (e^{L|x-x_0|} - 1).$$

Beweis. Es gilt für $x \in I$

$$\varphi_1(x) = \varphi_1(x_0) + \int_{x_0}^x \varphi_1'(t) dt.$$

Dasselbe gilt für φ_2 . Wendet man ferner auf $\|\varphi_1'(t) - \varphi_2'(t)\| = \|(\varphi_1'(t) - f(t, \varphi_1(t)) + (f(t, \varphi_1(t)) - f(t, \varphi_2(t)) - (\varphi_2'(t) - f(t, \varphi_2(t))))\|$ die Dreiecksungleichung an, so erhält man

$$\begin{aligned} \|\varphi_1(x) - \varphi_2(x)\| &\leq \|\varphi_1(x_0) - \varphi_2(x_0)\| + \int_{x_0}^x \|\varphi_1'(t) - \varphi_2'(t)\| dt \\ &\leq \|\varphi_1(x_0) - \varphi_2(x_0)\| + \int_{x_0}^x (\epsilon_1 + \|f(t, \varphi_1(t)) - f(t, \varphi_2(t))\| + \epsilon_2) dt \\ &\leq \|\varphi_1(x_0) - \varphi_2(x_0)\| + (\epsilon_1 + \epsilon_2)|x - x_0| + \int_{x_0}^x L\|\varphi_1(t) - \varphi_2(t)\| dt. \end{aligned}$$

Für $x \geq x_0$ können wir das Lemma von Gronwall anwenden, wobei $w(x) = \|\varphi_1(x) - \varphi_2(x)\|$, $\varphi(x) = A + B(x - x_0)$ für $A = \|\varphi_1(x_0) - \varphi_2(x_0)\|$ und $B = \epsilon_1 + \epsilon_2$ und

20. Abhängigkeit der Lösungen von Parametern

$\psi(x) = L$. Wir verwenden $\int y e^{-y} dy = -(y+1)e^{-y}$ und erhalten

$$\begin{aligned}
 \|\varphi_1(x) - \varphi_2(x)\| &\leq A + B(x - x_0) + \int_{x_0}^x L(A + B(t - x_0)) e^{L(x-t)} dt \\
 &= A e^{L(x-x_0)} + B(x - x_0) + B e^{L(x-x_0)} \int_{x_0}^x L(t - x_0) e^{L(x_0-t)} dt \\
 &= A e^{L(x-x_0)} + B(x - x_0) + \frac{B}{L} e^{L(x-x_0)} \int_0^{L(x-x_0)} y e^{-y} dy \\
 &= A e^{L(x-x_0)} + B(x - x_0) + \frac{B}{L} e^{L(x-x_0)} (1 - (L(x - x_0) + 1) e^{-L(x-x_0)}) \\
 &= A e^{L(x-x_0)} + \frac{B}{L} (e^{L(x-x_0)} - 1).
 \end{aligned}$$

Für $x < x_0$ argumentiert man analog. □

Aus diesem Satz folgt speziell noch einmal die Eindeutigkeit der Lösung der Anfangswertaufgabe.

20.3 Beispiel. Betrachte die Anfangswertaufgabe

$$\begin{aligned}
 y'''' &= \frac{1}{25} y'''' - \frac{4101}{2500} y'' + \frac{1}{25} y' - \frac{1601}{2500} y \\
 y(0) &= 1, \quad y'(0) = 0, \quad y''(0) = -1, \quad y'''(0) = 0.
 \end{aligned}$$

Die Nullstellen des charakteristischen Polynoms sind $\pm i$ und $\frac{1}{50} \pm \frac{4}{5}i$. Die eindeutig bestimmte Lösung der AWA ist $\varphi(x) = \cos(x)$. Wenn wir die Anfangswertaufgabe etwas stören

$$y(0) = 1, \quad y'(0) = \frac{1}{2500}, \quad y''(0) = -\frac{31249}{31250}, \quad y'''(0) = -\frac{1597}{3125000}$$

dann ist die eindeutige Lösung der Anfangswertaufgabe gleich

$$\psi(x) = \cos(x) + \frac{1}{1000} \sin\left(\frac{4}{5}x\right) \exp\left(\frac{x}{50}\right).$$

Graphen beider Funktionen (blau: φ , orange: ψ) zeigen die Abbildungen 20.1 bis 20.3.

Das zugehörige System ist $y' = Ay$ mit

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{1601}{2500} & \frac{1}{25} & -\frac{4101}{2500} & \frac{1}{25} \end{pmatrix}.$$

Wenn wir den $\mathbb{R}^n$ mit der Maximumsnorm $\|\cdot\|_\infty$ versehen, dann ist die Matrixnorm von A gleich der Zeilensummennorm, also

$$\|A\| = \frac{1601}{2500} + \frac{1}{25} + \frac{4101}{2500} + \frac{1}{25} = \frac{2951}{1250}.$$

Die Maximumsnorm der Differenz der Anfangswerte ist $\frac{1597}{3\,125\,000}$. Sowohl φ als auch ψ lösen die Differentialgleichung, also $\epsilon_1 = \epsilon_2 = 0$. Wir erhalten

$$|\varphi(x) - \psi(x)| \leq \frac{1597}{3\,125\,000} \exp\left(\frac{2951}{1250}|x|\right).$$

Das bedeutet beispielsweise, dass für $0 \leq x \leq 2.23$ ein Fehler von höchstens 0.1 garantiert ist. Abbildung [20.1](#) zeigt, dass der Unterschied in Wirklichkeit viel geringer ist.

20. Abhängigkeit der Lösungen von Parametern

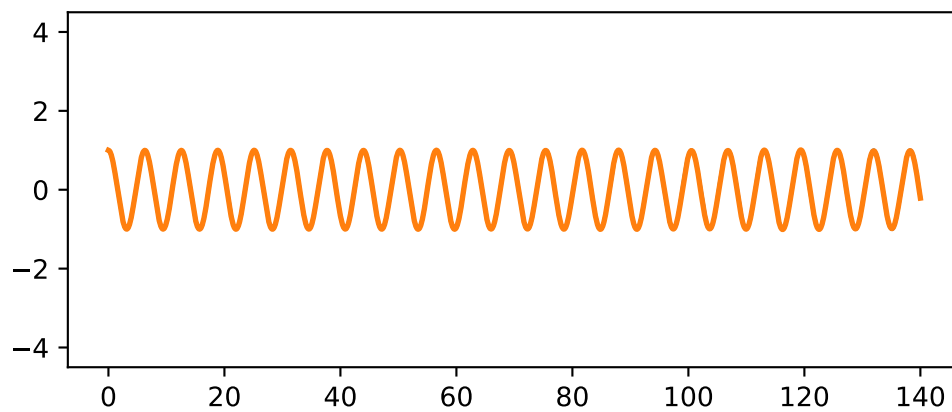


Abbildung 20.1.: Parameterabhängigkeit, blau: $\cos$, orange: gestörte Lösung

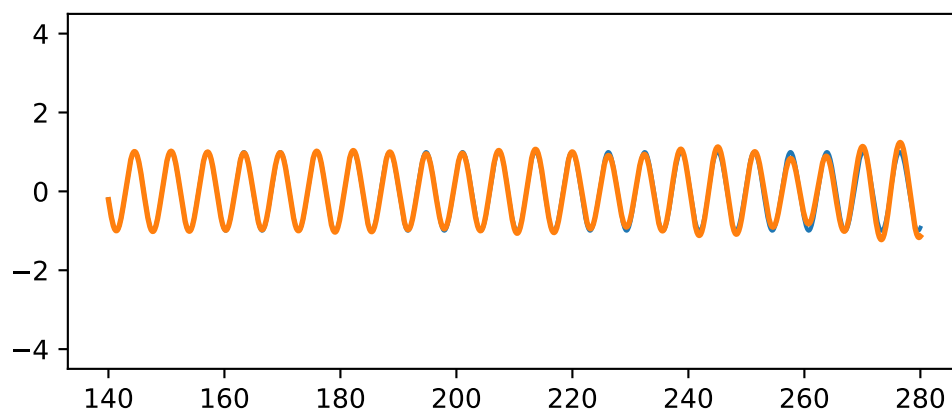


Abbildung 20.2.: Parameterabhängigkeit, blau: $\cos$, orange: gestörte Lösung

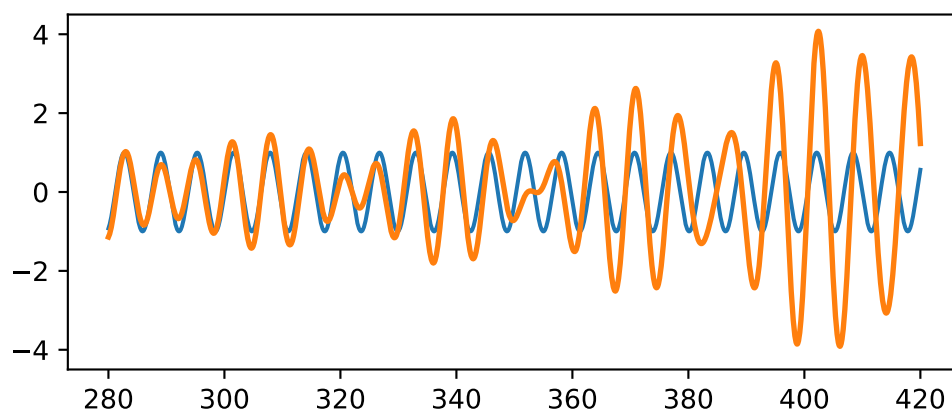


Abbildung 20.3.: Parameterabhängigkeit, blau: $\cos$, orange: gestörte Lösung

21. Autonome Differentialgleichungen

21.1 Definition. Sei $V \subseteq \mathbb{R}^n$ offen und sei $F: V \rightarrow \mathbb{R}^n$ stetig. Man nennt F auch *stetiges Vektorfeld* auf V . Dann bezeichnet man $y' = F(y)$ als *autonome* Differentialgleichung. Ihre Lösungen heißen *Integralkurven* von F .

21.2 Bemerkung. Es sei $\varphi: I \rightarrow V$ eine Lösung der autonomen Differentialgleichung $y' = F(y)$. Dann ist für jedes $a \in \mathbb{R}$ auch $\psi: I + a \rightarrow V$, $\psi(x) = \varphi(x - a)$, eine Lösung.

21.3 Definition. Es sei $V \subseteq \mathbb{R}^2$ offen und es sei $F: V \rightarrow \mathbb{R}^2$ stetig. Für eine beliebige Lösung $\varphi = (\varphi_1, \varphi_2): J \rightarrow \mathbb{R}^2$ von $y' = F(y)$ definiert man die zugehörige Lösungskurve im Phasenraum $\mathbb{R}^2$ als

$$K_\varphi = \{(\varphi_1(x), \varphi_2(x)) \mid x \in J\}.$$

Das *Phasenportrait* besteht aus sämtlichen Kurven K_φ zu Lösungen φ von $y' = F(y)$.

21.4 Beispiel. Abbildung 21.1 zeigt Phasenportraits für einige lineare Differentialgleichungssysteme.

21.5 Beispiel (Räuber-Beute-Gleichung). In einem gemeinsamen Lebensraum leben Füchse und Hasen. Die Größe der Hasenpopulation zur Zeit t sei $x(t)$, die der Fuchspopulation sei $y(t)$. Wir bezeichnen die Zeitableitungen mit $\dot{x}$ und $\dot{y}$. Ein auf den Mathematiker Volterra zurückgehendes Modell besagt, dass sich x und y entsprechend des folgenden Differentialgleichungssystems verhalten

$$\begin{aligned}\dot{x} &= ax - bxy \\ \dot{y} &= -cy + dxy\end{aligned}$$

mit positiven Konstanten a, b, c und d . Es handelt sich um ein nichtlineares Differentialgleichungssystem. Die Abbildungen 21.3 und 21.2 zeigen einige Lösungskurven.

21.6 Notation. Für $K > 0$ definiert man

$$M_K = \{(\xi, \eta) \in]0, \infty[^2 \mid \xi^c e^{-d\xi} \eta^a e^{-b\eta} = K\}.$$

21.7 Satz. Sei $\varphi: I \rightarrow \mathbb{R}^2$, $t \mapsto (x(t), y(t))$, eine Lösung der Räuber-Beute-Gleichung, so dass es ein $t_0 \in I$ mit $x(t_0) > 0$ und $y(t_0) > 0$ gibt. Dann gibt es ein K mit $(x(t), y(t)) \in M_K$ für alle $t \in I$.

21. Autonome Differentialgleichungen

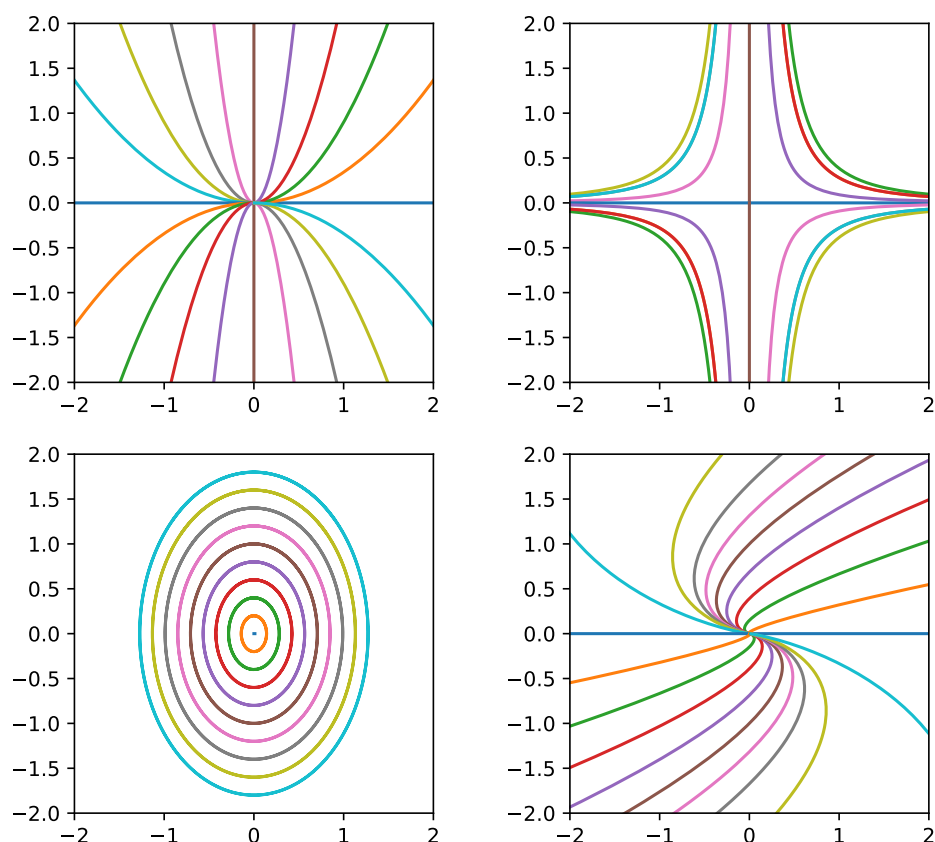


Abbildung 21.1.: Phasenportraits für $y' = Ay$ für $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ und $A = \begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix}$ im Uhrzeigersinn beginnend oben links

Beweis. Im Phasenraum können sich keine zwei Kurven überschneiden. Die Randkurven des ersten Quadranten sind selber Lösungen, also bleiben alle Lösungskurven im ersten Quadranten. Wir berechnen durch Einsetzen in der Differentialgleichung (ich schreibe das Argument t nicht hin)

$$\begin{aligned}
 \frac{d}{dt} x^c e^{-dx} y^a e^{-by} &= c \dot{x} x^{c-1} e^{-dx} y^a e^{-by} - d \dot{x} x^c e^{-dx} y^a e^{-by} \\
 &\quad + a \dot{y} x^c e^{-dx} y^{a-1} e^{-by} - b \dot{y} x^c e^{-dx} y^a e^{-by} \\
 &= acx^c e^{-dx} y^a e^{-by} - bcx^c e^{-dx} y^{a+1} e^{-by} \\
 &\quad - adx^{c+1} e^{-dx} y^a e^{-by} + bdx^{c+1} e^{-dx} y^{a+1} e^{-by} \\
 &\quad - acx^c e^{-dx} y^a e^{-by} + adx^{c+1} e^{-dx} y^{a+1} e^{-by} \\
 &\quad + bcx^c e^{-dx} y^{a+1} e^{-by} - bdx^{c+1} e^{-dx} y^{a+1} e^{-by}
 \end{aligned}$$

$$\begin{aligned}
&= (ac - ac)x^c e^{-dx} y^a e^{-by} + (-bc + bc)x^c e^{-dx} y^{a+1} e^{-by} \\
&\quad (-ad + ad)x^{c+1} e^{-dx} y^a e^{-by} + (bd - bd)x^{c+1} e^{-dx} y^{a+1} e^{-by} \\
&= 0.
\end{aligned}$$

21.8 Satz. Sei $V \subseteq \mathbb{R}^n$ offen und sei $F: V \rightarrow \mathbb{R}^n$ ein lokal Lipschitz-stetiges Vektorfeld. Sei $x_0 \in \mathbb{R}$, sei $y_0 \in V$ und sei $\varphi: I \rightarrow V$ eine Lösung der Anfangswertaufgabe $y' = F(y)$, $y(x_0) = y_0$, mit maximalem Definitionsbereich $I =]a, b[$. Falls $b < \infty$, dann gibt es für jede kompakte Menge $M \subset V$ eine Folge $(x_m)_{m \in \mathbb{N}}$ in I mit $\lim_{m \rightarrow \infty} x_m = b$ und $\varphi(x_m) \notin M$ für alle m . Die analoge Aussage gilt, falls $a > -\infty$.

Beweis. Sei $b < \infty$, und sei M eine beliebige kompakte Teilmenge von V . Aus Satz 19.5 folgt, dass $\overline{\mathcal{G}}$ keine kompakte Teilmenge von $\mathbb{R} \times V$ ist. Es gibt also eine Folge $(x_k, y_k)_{k \in \mathbb{N}}$ in $\mathcal{G}$, die keine in $\mathbb{R} \times V$ konvergente Teilfolge hat. Da $(x_k)_{k \in \mathbb{N}}$ beschränkt ist, besitzt sie eine konvergente Teilfolge $(x_{k_m})_{m \in \mathbb{N}}$ mit $\lim_{m \rightarrow \infty} x_{k_m} = \xi$. Es gilt $\xi < b$, weil sonst die Folge $(x_{k_m}, \varphi(x_{k_m}))_{m \in \mathbb{N}}$ gegen $(\xi, \varphi(\xi))$ konvergieren würde. Da M kompakt ist und $(y_{k_m})_{m \in \mathbb{N}}$ keine konvergente Teilfolge enthält, können nur endlich viele ihrer Elemente in M liegen. Daher besitzt $(x_{k_m})_{m \in \mathbb{N}}$ eine Teilfolge mit den in der Behauptung geforderten Eigenschaften. $\square$

21.9 Korollar. Es sei $\varphi: I \rightarrow \mathbb{R}^2$, $t \mapsto (x(t), y(t))$ eine maximale Lösung der Räuber-Beute Gleichung, für die es ein t_0 mit $x(t_0) > 0$ und $y(t_0) > 0$ gibt. Dann $I = \mathbb{R}$.

Beweis. Wir müssen zeigen, dass M_K beschränkt ist. Wir nehmen zum Widerspruch an, dass es eine Folge $((\xi_n, \eta_n))_{n \in \mathbb{N}}$ mit $\lim_{n \rightarrow \infty} \xi_n = \infty$ gibt. Dann $\lim_{n \rightarrow \infty} \xi_n^c e^{-d\xi_n} = 0$ und daher $\lim_{n \rightarrow \infty} \eta_n^a e^{-b\eta_n} = \infty$. Letzteres kann nicht sein, da die Funktion $]0, \infty[\rightarrow \mathbb{R}$, $\eta \mapsto \eta^a e^{-b\eta}$ beschränkt ist. Genauso argumentiert man im Fall $\lim_{n \rightarrow \infty} \eta_n = \infty$. $\square$

21.10 Satz. Seien $a, b, c, d > 0$, und sei $\varphi: \mathbb{R} \rightarrow]0, \infty[^2$ eine Lösung der Räuber-Beute Gleichung. Dann ist φ periodisch, d. h. es gibt $T > 0$ mit $\varphi(t + T) = \varphi(t)$ für alle $t \in \mathbb{R}$.

Beweis. Diesen Satz werde ich im Kapitel über Untermannigfaltigkeiten zeigen. $\square$

21. Autonome Differentialgleichungen

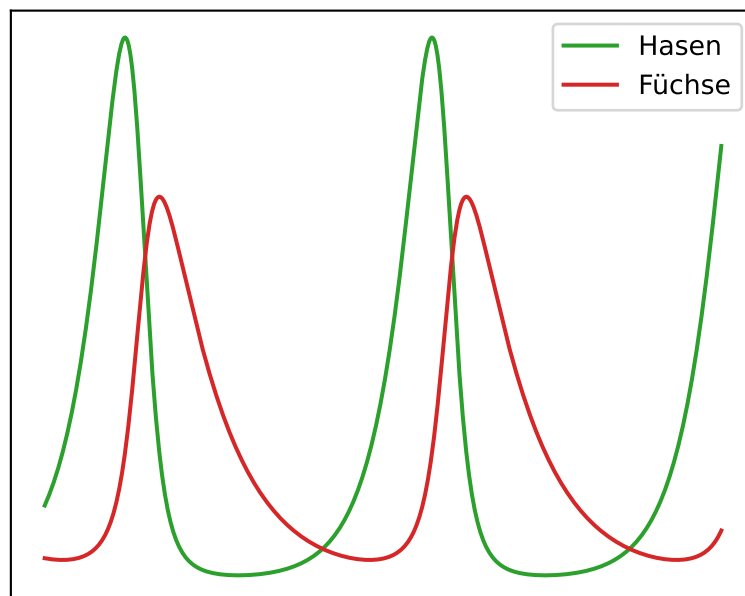


Abbildung 21.2.: Lösungen einer Räuber-Beute-Gleichung

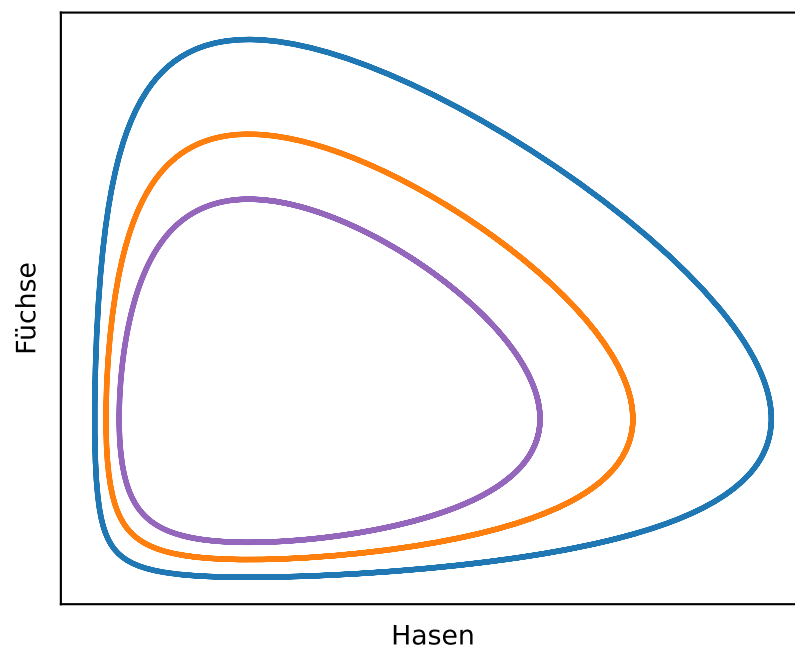


Abbildung 21.3.: Phasenportrait einer Räuber-Beute-Gleichung

Teil III.

Der Satz über implizite Abbildungen

22. Der Umkehrsatz

22.1 Definition. Es seien $W, U \subseteq \mathbb{R}^n$ offen und $k \in \mathbb{N} \cup \{\infty\}$. Eine bijektive Abbildung $\psi: U \rightarrow W$ von der Klasse C^k heißt C^k -Diffeomorphismus, wenn die Umkehrabbildung ψ^{-1} ebenfalls von der Klasse C^k ist.

22.2 Bemerkungen. (a) Die Abbildung $f:]-1, 1[\rightarrow]-1, 1[, x \mapsto x^3$, ist bijektiv und von der Klasse C^∞ , aber noch nicht mal ein C^1 -Diffeomorphismus.

(b) $\psi: U \rightarrow W$ sei ein C^1 -Diffeomorphismus. Dann ist für jedes $u \in U$ die Matrix $D(\psi)(u) \in \mathbb{R}^{n \times n}$ invertierbar. Das folgt sofort aus der Kettenregel.

(c) Die Umkehrung gilt nicht. Betrachte

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad f(x, y) = \begin{pmatrix} e^x \cos(y) \\ e^x \sin(y) \end{pmatrix}.$$

Dann ist $Df(x, y)$ für jedes $(x, y) \in \mathbb{R}^2$ invertierbar. Trotzdem ist f nicht invertierbar.

22.3 Definition. Es sei $W \subseteq \mathbb{R}^n$ offen und $k \in \mathbb{N} \cup \{\infty\}$. Eine Abbildung $f: W \rightarrow \mathbb{R}^n$ heißt *lokaler C^k -Diffeomorphismus*, falls es zu jedem Punkt $x \in W$ eine offene Umgebung U von x gibt, so dass $f(U)$ offen und $f: U \rightarrow f(U)$ ein C^k -Diffeomorphismus ist.

22.4 Theorem (Umkehrsatz). Es sei $W \subseteq \mathbb{R}^n$ offen und $k \in \mathbb{N} \cup \{\infty\}$. Eine Abbildung $f: W \rightarrow \mathbb{R}^n$ von der Klasse C^k ist genau dann ein lokaler C^k -Diffeomorphismus, wenn für alle $x \in W$ die Matrizen $Df(x)$ invertierbar sind.

22.5 Beispiel (Kugelkoordinaten). Für $q = (x, y, z) \in \mathbb{R}^3$ sei $r = \sqrt{x^2 + y^2 + z^2}$. Im Fall $q \neq 0$ gibt es genau ein $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ mit $z = r \sin \theta$. Der Winkel θ ist der Breitengrad von q . Es folgt dann $x^2 + y^2 = r^2 \cos^2 \theta$. Schreibt man jetzt (x, y) in ebenen Polarkoordinaten, so ergibt sich

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \Psi(r, \varphi, \theta) = \begin{pmatrix} r \cos(\varphi) \cos(\theta) \\ r \sin(\varphi) \cos(\theta) \\ r \sin(\theta) \end{pmatrix}.$$

φ ist dann der Längengrad von q . Wir haben gerade gesehen, dass $\Psi: [0, \infty[\times]-\pi, \pi] \times [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow \mathbb{R}^3$ surjektiv ist. Es gilt

$$D\Psi(r, \varphi, \theta) = \begin{pmatrix} \cos(\varphi) \cos(\theta) & -r \sin(\varphi) \cos(\theta) & -r \cos(\varphi) \sin(\theta) \\ \sin(\varphi) \cos(\theta) & r \cos(\varphi) \cos(\theta) & -r \sin(\varphi) \sin(\theta) \\ \sin(\theta) & 0 & r \cos(\theta) \end{pmatrix}.$$

22. Der Umkehrsatz

Dann $\det(D\Psi(r, \varphi, \theta)) = r^2 \cos \theta$. Also ist Ψ ein lokaler C^∞ -Diffeomorphismus auf $\{(r, \varphi, \theta) \in \mathbb{R}^3 \mid r \neq 0, \cos \theta \neq 0\}$. Aus der Periodizität der trigonometrischen Funktionen folgt, dass $\Psi: U \rightarrow \mathbb{R}^3 \setminus H$ ein C^∞ -Diffeomorphismus ist, wenn $U =]0, \infty[\times]-\pi, \pi[\times]-\frac{\pi}{2}, \frac{\pi}{2}[$ und $H = \{(x, 0, z) \mid x \leq 0, z \in \mathbb{R}\}$.

Die folgende Formulierung des Umkehrsatzes impliziert Theorem 22.4.

22.6 Satz. *Es sei $G \subseteq \mathbb{R}^n$ offen und $k \in \mathbb{N} \cup \{\infty\}$, und es sei $f: G \rightarrow \mathbb{R}^n$ von der Klasse C^k . Es sei $a \in G$ ein Punkt, für den $Df(a)$ invertierbar ist. Dann gibt es Umgebungen V von a und U von $f(a)$, so dass $f: V \rightarrow U$ bijektiv ist. Bezeichnet man mit $g: U \rightarrow V$ ihre Inverse, so ist g von der Klasse C^k , und es gilt*

$$Dg(f(a)) = (Df(a))^{-1}.$$

Beim Beweis des Umkehrsatzes orientiere ich mich an dem Buch von Kaballo. Er erfolgt in mehreren Schritten. Der erste ist der folgende Satz, der aus der Cramerschen Regel folgt.

22.7 Satz. *Die Abbildung $A \mapsto A^{-1}$ ist ein C^∞ -Diffeomorphismus von $GL_n(\mathbb{R})$ in sich.*

Beweis. Die Cramersche Regel besagt, dass die Einträge von A^{-1} rationale Funktionen in den Einträgen von A sind. $\square$

Die folgenden Lemmata sind Beweisschritte des Umkehrsatzes. Sie beziehen sich auf dessen Notation.

22.8 Lemma. *Man kann ohne Einschränkung $a = 0$, $f(a) = 0$ und $Df(0) = E_n$ annehmen.*

Beweis. Durch Translation des Koordinatensystems kann man $a = 0$ erreichen. Sei dann $A = Df(0)$. Setze $h(x) = A^{-1}(f(x) - f(0))$. Dann $h(0) = 0$ und $Dh(0) = E_n$ und h ist von der Klasse C^k . Wenn für Nullumgebungen U und V gilt, dass $h: V \rightarrow U$ bijektiv ist mit C^k -Inverser, dann ist $f: V \rightarrow A(U) + f(0)$ bijektiv mit C^k -Inverser $f^{-1}: y \mapsto h^{-1}(A^{-1}(y - f(0)))$. $\square$

22.9 Lemma. *Unter der Annahme von Lemma 22.8 gibt es $\delta > 0$ mit $Df(y) \in GL_n(\mathbb{R})$ für alle $y \in \overline{B}_{7\delta}(0)$ und für $\|\cdot\| = \|\cdot\|_\infty$ gilt*

$$\|f(x+h) - f(x) - h\| \leq \frac{1}{2}\|h\| \quad \text{für alle } x, x+h \in \overline{B}_{2\delta}(0). \quad (22.1)$$

Insbesondere gilt $\|f(x+h) - f(x)\| \geq \frac{1}{2}\|h\|$ für alle $x, x+h \in \overline{B}_{2\delta}(0)$, und f ist auf $\overline{B}_{2\delta}(0)$ injektiv.

Beweis. Die Abbildung $y \mapsto Df(y)$ ist stetig und $GL_n(\mathbb{R})$ ist offen. Daher existiert $\delta > 0$, so dass $\overline{B}_{7\delta}(0) \subseteq G$ und

$$Df(y) \in GL_n(\mathbb{R}) \quad \text{und} \quad \|Df(y) - E_n\| \leq \frac{1}{2} \quad \text{für alle } y \in \overline{B}_{7\delta}(0).$$

Es sei nun $x \in \overline{B}_{2\delta}(0)$ fest. Für $h \in B_{5\delta}(0)$ definiert man $F(h) = f(x+h) - h$. Dann gilt $\|DF(h)\| \leq \frac{1}{2}$. Aus dem Mittelwertsatz folgt $\|F(h) - F(0)\| \leq \frac{1}{2}\|h\|$, also

$$\|f(x+h) - f(x) - h\| \leq \frac{1}{2}\|h\|. \quad \square$$

22.10 Lemma. Für $\delta > 0$ ist der Raum $X = \{\varphi: \overline{B}_\delta(0) \rightarrow \overline{B}_{2\delta}(0) \mid \varphi \text{ stetig}\}$ vollständig, wenn man ihn mit der Metrik $d(\varphi, \psi) = \sup\{\|\varphi(y) - \psi(y)\| \mid \|y\| \leq \delta\}$ versieht.

Beweis. Nach Satz 19.3 ist $C(I, \mathbb{R}^n)$ für kompakte Intervalle I bezüglich $\|\cdot\|_\infty$ vollständig. Der Beweis von Satz 19.3 überträgt sich auf den Raum X . $\square$

22.11 Lemma. Unter der Annahme von Lemma 22.8 gilt $B_\delta(0) \subseteq f(B_{2\delta}(0)) \subseteq B_{3\delta}(0)$. Außerdem gibt es eine stetige Abbildung $g: \overline{B}_\delta(0) \rightarrow \overline{B}_{2\delta}(0)$ mit $f(g(y)) = y$ für alle $y \in \overline{B}_\delta(0)$.

Beweis. Setzt man $x = 0$ im vorigen Lemma, so erhält man $\|f(h) - h\| \leq \frac{1}{2}\|h\|$ für $h \in \overline{B}_{2\delta}(0)$. Daraus folgt $f(B_{2\delta}(0)) \subseteq B_{3\delta}(0)$.

Für X wie im vorigen Lemma betrachten wir

$$G: X \rightarrow X, \quad G(\varphi)(y) = \varphi(y) - f(\varphi(y)) + y.$$

Es ist nicht à priori klar, dass $G(\varphi) \in X$ für $\varphi \in X$. Sei also $y \in \overline{B}_\delta(0)$ beliebig. Dann $\|\varphi(y)\| \leq 2\delta$, weil $\varphi \in X$, und aus (22.1), angewandt auf $x = 0$ und $h = \varphi(y)$, folgt

$$\|G(\varphi)(y)\| \leq \|\varphi(y) - f(\varphi(y))\| + \|y\| \leq \frac{1}{2}\|\varphi(y)\| + \|y\| \leq 2\delta.$$

Beachte, dass im Fall $\|y\| < \delta$ sogar $\|G(\varphi)(y)\| < 2\delta$ folgt. Um zu zeigen, dass G eine Kontraktion ist, benutzen wir ebenfalls (22.1). Seien $\varphi, \psi \in X$ und sei $y \in \overline{B}_\delta(0)$ fest. Dann gibt es ein $h \in \mathbb{R}^n$ mit $\psi(y) = \varphi(y) + h$. Damit gilt

$$\begin{aligned} \|G(\psi)(y) - G(\varphi)(y)\| &= \|\varphi(y) + h - f(\varphi(y) + h) - \varphi(y) + f(\varphi(y))\| \\ &= \|f(\varphi(y)) - f(\varphi(y) + h) + h\| \leq \frac{1}{2}\|h\|. \end{aligned}$$

Damit ist gezeigt, dass G kontrahierend mit Kontraktionskonstante $\frac{1}{2}$ ist. Der Banachsche Fixpunktsatz impliziert die Existenz der behaupteten Abbildung g .

Sei schließlich $y \in B_\delta(0)$ gegeben. Dann folgt aus der Zusatzüberlegung, dass $\|g(y)\| = \|G(g)(y)\| < 2\delta$. Daher ist $y = f(g(y))$ in $f(B_{2\delta}(0))$. $\square$

22. Der Umkehrsatz

22.12 Lemma. *Unter der Annahme von Lemma 22.8 sei $U = B_\delta(0)$ und $V = g(U)$. Dann sind V und U offene Umgebungen von $a = 0$ bzw. $f(a) = 0$, und $f: V \rightarrow U$ ist ein C^k -Diffeomorphismus. Es gilt $D(g)(y) = (D(f)(f^{-1}(y)))^{-1}$.*

Beweis. Mit U ist nach Lemma 22.11 auch $V = f^{-1}(U)$ offen und wegen $f \circ g = \text{id}$ ist $g: U \rightarrow V$ in der Tat invertierbar mit Inverser f . Wir zeigen zunächst, dass g differenzierbar ist. Dazu seien $y, y + k \in U$ gegeben. Wir setzen

$$x := g(y) \quad \text{und} \quad h := g(y + k) - g(y).$$

Das bedeutet $g(y + k) = x + h$ und daher

$$k = f(x + h) - f(x) = D(f)(x)h + \rho(h) \tag{22.2}$$

für eine Abbildung ρ mit $\lim_{h \rightarrow 0} \frac{\|\rho(h)\|}{\|h\|} = 0$. Für $B := (Df(x))^{-1}$ erhalten wir

$$Bk = h + B\rho(h),$$

also

$$g(y + k) - g(y) = h = Bk - B\rho(h).$$

Wegen der ersten Gleichung in (22.2) folgt aus (22.1), dass $\|k\| \geq \frac{1}{2}\|h\|$. Daher impliziert $k \rightarrow 0$ auch $h \rightarrow 0$ und die Differenzierbarkeit von g in y ist gezeigt.

Genauer haben wir $D(g)(y) = (D(f)(f^{-1}(y)))^{-1}$ gezeigt. Daher folgen die k -fache Differenzierbarkeit und die Stetigkeit der Ableitungen mit vollständiger Induktion. Beim Induktionsschritt ist g von der Klasse C^{j-1} . Diese Abbildung wird verknüpft mit welchen, die mindestens von der Klasse C^k sind. $\square$

23. Der Satz über implizite Funktionen

23.1 Beispiel. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^3 + y^3 - 3xy$. Die Menge $\{(x, y) \in \mathbb{R}^2 \mid f(x, y) = 0\}$ ist eine Kurve im $\mathbb{R}^2$. Sie wird in Abbildung 23.1 gezeigt. Fast überall ist sie wenigstens lokal der Graph einer Funktion.

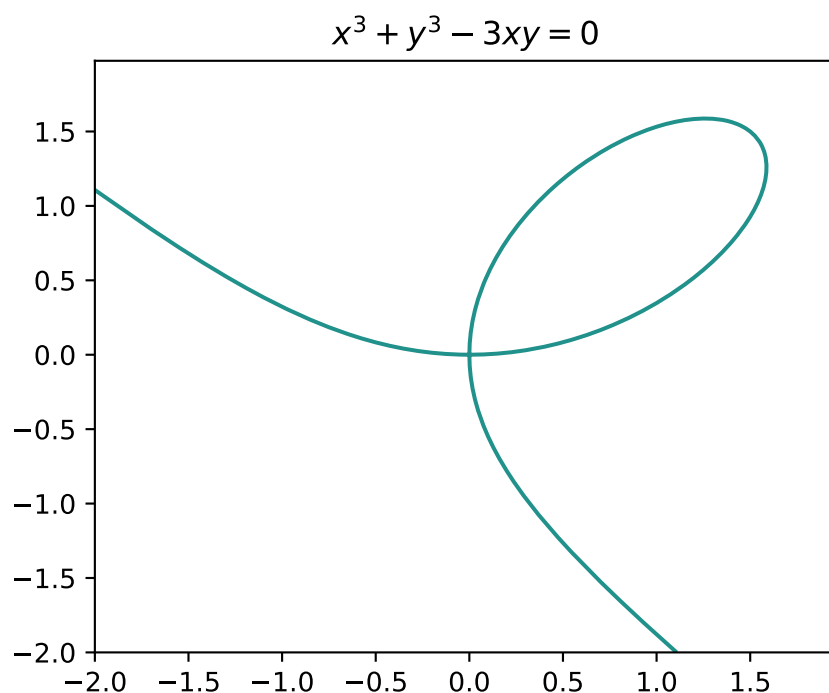


Abbildung 23.1.: Das kartesische Blatt

23.2 Definition. Wenn die Elemente von $\mathbb{R}^k \times \mathbb{R}^m$ als (x, y) mit $x \in \mathbb{R}^k$ und $y \in \mathbb{R}^m$ geschrieben werden, dann definiert man für offene Mengen $U \subseteq \mathbb{R}^k \times \mathbb{R}^m$ und Abbildungen $F: U \rightarrow \mathbb{R}^n$ mit Komponenten $F_1, \dots, F_n$

$$\frac{\partial F}{\partial y} = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial y_1} & \cdots & \frac{\partial F_n}{\partial y_m} \end{pmatrix}.$$

23.3 Theorem (Satz über implizite Funktionen). Seien $U_1 \subseteq \mathbb{R}^k$ und $U_2 \subseteq \mathbb{R}^m$ offen und seien $x_0 \in U_1$, $y_0 \in U_2$. Sei $F: U_1 \times U_2 \rightarrow \mathbb{R}^n$ von der Klasse C^p , $p \in \mathbb{N} \cup \{\infty\}$,

23. Der Satz über implizite Funktionen

mit $F(x_0, y_0) = 0$. Ferner sei die Matrix $\frac{\partial F}{\partial y}(x_0, y_0)$ invertierbar. Dann gibt es eine Umgebung $U \subseteq U_1$ von x_0 , eine Umgebung $V \subseteq U_2$ von y_0 und eine Abbildung $g: U \rightarrow V$ von der Klasse C^p mit den folgenden Eigenschaften

(a) $F(x, g(x)) = 0$ für alle $x \in U$.

(b) Wenn $F(x, y) = 0$ für $(x, y) \in U \times V$, dann $y = g(x)$.

Es gilt

$$Dg(x_0) = - \left(\frac{\partial F}{\partial y}(x_0, y_0) \right)^{-1} \circ \left(\frac{\partial F}{\partial x}(x_0, y_0) \right).$$

23.4 Beispiel. Sei $f(x, y) = x^3 + y^3 - 3xy$ und sei K die Kurve $K = \{(x, y) \mid f(x, y) = 0\}$. Wir suchen diejenigen Punkte $(x_0, y_0) \in K$, in denen die Voraussetzungen des Satzes über implizite Funktionen verletzt sind

$$0 = \frac{\partial f}{\partial y}(x_0, y_0) = 3y_0^2 - 3x_0.$$

Also $x_0 = y_0^2$. Welche davon liegen in K ?

$$0 = f(y_0^2, y_0) = y_0^6 - 2y_0^3.$$

Also $y_0 = 0$ oder $y_0 = \sqrt[3]{2}$ und daher $(x_0, y_0) = (0, 0)$ oder $(x_0, y_0) = (\sqrt[3]{4}, \sqrt[3]{2})$. Für alle anderen Punkte existieren also ein Intervall $]x_0 - \epsilon, x_0 + \epsilon[$, eine offene Umgebung U von (x_0, y_0) und eine C^∞ -Abbildung $g:]x_0 - \epsilon, x_0 + \epsilon[\rightarrow \mathbb{R}$, so dass $K \cap U = \{(x, g(x)) \mid x_0 - \epsilon < x < x_0 + \epsilon\}$.

In $(\sqrt[3]{4}, \sqrt[3]{2})$ gilt

$$\frac{\partial f}{\partial x}(\sqrt[3]{4}, \sqrt[3]{2}) = 3(\sqrt[3]{4})^2 - 3\sqrt[3]{2} = 3\sqrt[3]{2} \neq 0.$$

Daher kann dort die Kurve als Graph $x = h(y)$ geschrieben werden. Dagegen gilt $\frac{\partial f}{\partial x}(0, 0) = 0$. Dort lässt sich der Satz über implizite Funktionen nicht anwenden, und in der Tat ist K dort lokal kein Graph. (Der algebraische Geometer sagt: “ K besitzt im Ursprung eine Singularität.”)

In jedem Punkt $(x_0, y_0) \in K$ mit $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ kann man ferner die Tangente an den Graphen von g bestimmen. Es gilt nämlich

$$g'(x_0) = \frac{y_0 - x_0^2}{y_0^2 - x_0}.$$

Daher ist die Tangente an K im Punkt (x_0, y_0) gleich

$$T = (x_0, y_0) + \mathbb{R} \begin{pmatrix} y_0^2 - x_0 \\ y_0 - x_0^2 \end{pmatrix}.$$

Das gilt auch in $(\sqrt[3]{4}, \sqrt[3]{2})$, aber nicht im Ursprung. Descartes hat von Fermat verlangt, die Tangente auch im Ursprung zu bestimmen. Fermat konnte das. Wir werden das auch können.

Beweis des Satzes über implizite Abbildungen. Definiere

$$\Phi: \mathcal{U}_1 \times \mathcal{U}_2 \rightarrow \mathbb{R}^k \times \mathbb{R}^m, \quad \Phi \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ F(x, y) \end{pmatrix}.$$

Es gelten $\Phi(x_0, y_0) = (x_0, 0)$ und

$$D\Phi = \begin{pmatrix} E_k & 0 \\ \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \end{pmatrix}.$$

Daher ist $D\Phi(x_0, y_0)$ invertierbar. Aus dem Umkehrsatz folgt, dass es Umgebungen $W_1 \times V \subseteq \mathbb{R}^k \times \mathbb{R}^m$ von (x_0, y_0) und $W_2 \subseteq \mathbb{R}^{k+m}$ von $(x_0, 0)$ gibt, so dass $\Phi: W_1 \times V \rightarrow W_2$ ein C^p -Diffeomorphismus ist. Setze

$$\mathcal{U} = \{x \in W_1 \mid \begin{pmatrix} x \\ 0 \end{pmatrix} \in W_2\}$$

und definiere $g: \mathcal{U} \rightarrow V$ durch

$$\Phi^{-1} \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ g(x) \end{pmatrix}.$$

Dann ist g von der Klasse C^p , und es gilt für alle $x \in \mathcal{U}$

$$\begin{pmatrix} x \\ F(x, g(x)) \end{pmatrix} = \Phi \begin{pmatrix} x \\ g(x) \end{pmatrix} = \Phi \circ \Phi^{-1} \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}.$$

Gilt umgekehrt $F(x, y) = 0$ für $x \in \mathcal{U}$ und $y \in V$, so folgt

$$\Phi \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix},$$

also

$$\begin{pmatrix} x \\ y \end{pmatrix} = \Phi^{-1} \begin{pmatrix} x \\ 0 \end{pmatrix}$$

und damit $y = g(x)$. Die Formel für $Dg(x_0)$ erhalten wir schließlich aus der Kettenregel. $\square$

23.5 Beispiel. Der Nullpunkt ist eine Lösung des Gleichungssystems

$$\begin{aligned} (x-1)^3 + (y-1)^4 - (z-1)^5 - 1 &= 0 \\ 4(x-1)^2 + 3(y-1)^3 - 2(z-1)^4 + 1 &= 0. \end{aligned}$$

Gibt es beliebig kleine Lösungen (x, y, z) mit $x > 0$, $y < 0$ und $z > 0$?

23. Der Satz über implizite Funktionen

Wir fassen die Variablen y und z zu $Y = \begin{pmatrix} y \\ z \end{pmatrix}$ zusammen. Dann

$$\frac{\partial F}{\partial Y}(x, y, z) = \begin{pmatrix} 4(y-1)^3 & -5(z-1)^4 \\ 9(y-1)^2 & -8(z-1)^3 \end{pmatrix},$$

also ist

$$\frac{\partial F}{\partial Y}(0, 0, 0) = \begin{pmatrix} -4 & -5 \\ 9 & 8 \end{pmatrix},$$

invertierbar mit Inverser

$$\left(\frac{\partial F}{\partial Y}(0, 0, 0) \right)^{-1} = \frac{1}{13} \begin{pmatrix} 8 & 5 \\ -9 & -4 \end{pmatrix}.$$

Aus dem Satz über implizite Abbildungen folgt die Existenz einer Nullumgebung $U \subseteq \mathbb{R}$ und einer C^∞ -Abbildung $\varphi: U \rightarrow \mathbb{R}^2$, so dass für alle $x \in U$ das Tripel $(x, \varphi(x))$ das Gleichungssystem löst. Um die Ableitung von φ zu bestimmen, berechnen wir

$$\frac{\partial F}{\partial x}(x, y, z) = \begin{pmatrix} 3(x-1)^2 \\ 8(x-1) \end{pmatrix},$$

also

$$\frac{\partial F}{\partial x}(0, 0, 0) = \begin{pmatrix} 3 \\ -8 \end{pmatrix},$$

Damit

$$\varphi'(0) = -\frac{1}{13} \begin{pmatrix} 8 & 5 \\ -9 & -4 \end{pmatrix} \begin{pmatrix} 3 \\ -8 \end{pmatrix} = -\frac{1}{13} \begin{pmatrix} -16 \\ 5 \end{pmatrix}.$$

Wenn also $x > 0$ klein ist, dann $y = \frac{16}{13}x + o(x)$ und $z = -\frac{5}{13}x + o(x)$. Es gibt also keine Lösung mit $x \in U$ und der gewünschten Vorzeichenverteilung.

23.6 Beispiel. Es sei $n \in \mathbb{N}$. Wir wollen zeigen, dass es zu jeder Matrix $A \in \mathbb{R}^{n \times n}$ in der Nähe der Einheitsmatrix eine Matrix B mit $B^2 = A$ gibt.

Sei dazu $F: \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}$ definiert durch $F(A, B) = A - B^2$. Dann ist das Paar (E_n, E_n) eine Nullstelle von F . Für $H \in \mathbb{R}^{n \times n}$ gilt

$$\frac{\partial F}{\partial B}(A, B)H = BH + HB,$$

also

$$\frac{\partial F}{\partial B}(E_n, E_n)H = 2H.$$

Also sind die Voraussetzungen des Satzes über implizite Abbildungen erfüllt und es gibt eine Umgebung U der Einheitsmatrix, so dass zu jedem $A \in U$ ein $B \in \mathbb{R}^{n \times n}$ mit $B^2 = A$ existiert.

24. Untermannigfaltigkeiten des $\mathbb{R}^N$

24.1 Definition. Sei $p \in \mathbb{N} \cup \{\infty\}$. Eine Teilmenge $M \subseteq \mathbb{R}^N$ heißt *n-dimensionale Untermannigfaltigkeit* des $\mathbb{R}^N$ der Klasse C^p , wenn es für jedes $a \in M$ eine offene Umgebung $U \subseteq \mathbb{R}^N$ von a und eine Abbildung $f: U \rightarrow \mathbb{R}^{N-n}$ von der Klasse C^p gibt, so dass

$$(a) \quad M \cap U = \{x \in U \mid f(x) = 0\},$$

$$(b) \quad \text{Rang}(Df(a)) = N - n.$$

24.2 Bemerkung. $\text{Rang}(Df(a)) = N - n$ bedeutet, dass die lineare Abbildung

$$Df(a): \mathbb{R}^N \rightarrow \mathbb{R}^{N-n}$$

surjektiv ist. Das ist äquivalent dazu, dass es eine $(N-n) \times (N-n)$ -Unterdeterminante von $Df(a)$ gibt, die nicht verschwindet.

24.3 Beispiel. (a) Der Einheitskreis $S^1 = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$ ist eine C^∞ -Untermannigfaltigkeit des $\mathbb{R}^2$. Man kommt sogar mit einem von a unabhängigen f aus und kann setzen

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^2 + x_2^2 - 1.$$

(b) Genauso zeigt man, dass die Einheitssphäre S^{n-1} eine C^∞ -Untermannigfaltigkeit des $\mathbb{R}^n$ ist.

(c) Sei $U \subseteq \mathbb{R}^n$ offen, sei $g: U \rightarrow \mathbb{R}^m$ von der Klasse C^p . Dann ist der Graph $\mathcal{G}(g)$ eine n -dimensionale Untermannigfaltigkeit des $\mathbb{R}^{n+m}$ von der Klasse C^p .

Beweis. Setzt man $f: U \times \mathbb{R}^m, f(x, y) = y - g(x)$, so hat

$$Df(x, y) = (-Dg(x) \quad E_m)$$

den Rang $m = (m + n) - n$. □

(d) Für die Abbildung $h: \mathbb{R}^2 \rightarrow \mathbb{R}, h(x, y) = x^3 + y^3 - 3xy$, aus Beispiel 23.1 sind alle Höhenlinien

$$\{(x, y) \in \mathbb{R}^2 \mid h(x, y) = c\}$$

mit Ausnahme der Höhenlinien zu $c = 0$ und $c = -1$ eindimensionale Mannigfaltigkeiten der Klasse C^∞ . Einige Höhenlinien zeigt Abbildung 24.1, den 3D-Graphen zeigt Abbildung 24.2.

24. Untermannigfaltigkeiten des $\mathbb{R}^N$

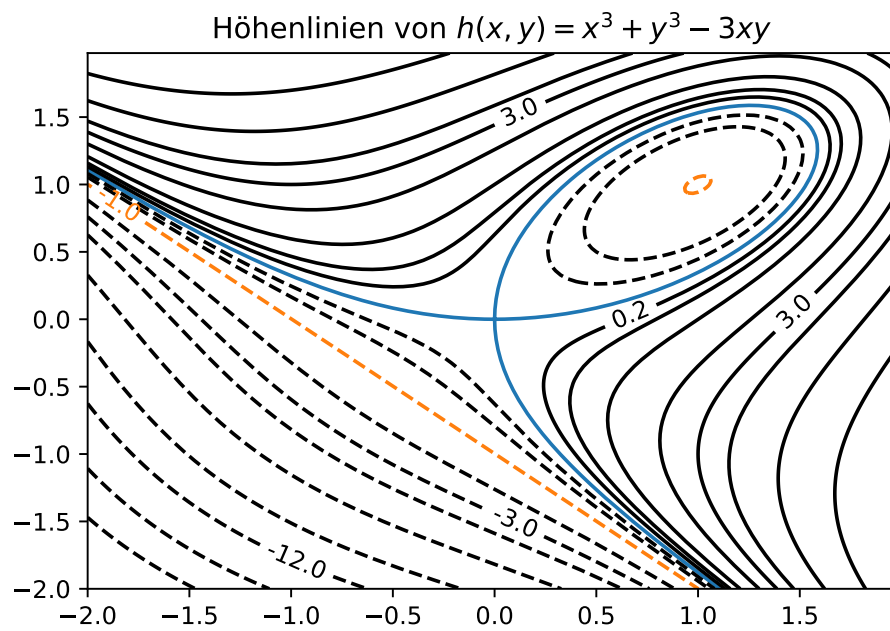


Abbildung 24.1.: Höhenlinien zu Beispiel 24.3(c). Die blaue Kurve ist das kartesische Blatt, die orange die Höhenlinie zu $c = -0.99$.

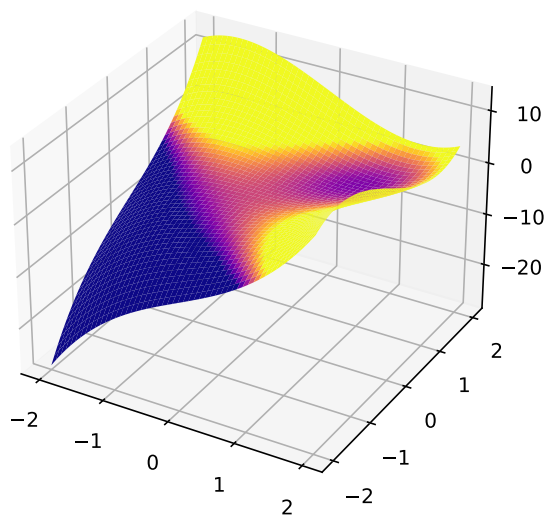


Abbildung 24.2.: Graph zu Beispiel 24.3(c)

Beweis. Für $f(x, y) = h(x, y) - c$ gilt

$$Df(x, y) = \begin{pmatrix} 3x^2 - 3y & 3y^2 - 3x \end{pmatrix}.$$

Diese Matrix hat vollen Rang, außer wenn $x^2 = y$ und $y^2 = x$, was äquivalent ist zu $(x, y) = (0, 0)$ oder $(x, y) = (1, 1)$. Im ersten Fall gilt $c = 0$, im zweiten $c = -1$. $\square$

24.4 Satz. Sei $M \subseteq \mathbb{R}^N$. Dann sind äquivalent:

- (a) M ist eine n -dimensionale C^p -Untermannigfaltigkeit von $\mathbb{R}^N$.
- (b) Für jedes $a \in M$ existieren eine offene Umgebung $U \subseteq \mathbb{R}^N$ von a , eine offene Teilmenge $V \subseteq \mathbb{R}^N$ und ein C^p -Diffeomorphismus $h: U \rightarrow V$ mit

$$h(M \cap U) = V \cap (\mathbb{R}^n \times \{0\}).$$

Beweis. (a) $\implies$ (b): Es sei $a \in M$ beliebig gewählt. Wegen $\text{Rang } Df(a) = N - n$ können wir die Variablen so umnummerieren, dass die letzten letzten $N - n$ Spalten der Matrix linear unabhängig sind. Wenn wir das gemacht haben, schreiben wir die Elemente des $\mathbb{R}^N$ in der Form $(x_1, \dots, x_n, y_1, \dots, y_{N-n})$. Entsprechend schreiben wir $a = (\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_{N-n})$. Dann ist für das f aus der Definition der Untermannigfaltigkeit die Matrix $\frac{\partial f}{\partial y}(\alpha, \beta)$ invertierbar. Es gibt also Umgebungen $W_1 \subseteq \mathbb{R}^n$ von α und $W_2 \subseteq \mathbb{R}^{N-n}$ von β und eine Abbildung $g: W_1 \rightarrow W_2$ von der Klasse C^p , so dass

$$M \cap (W_1 \times W_2) = \{(x, g(x)) \in \mathbb{R}^n \times \mathbb{R}^{N-n} \mid x \in W_1\}.$$

Setze $U = W_1 \times W_2$ und $V = \{(\xi, \eta) \in W_1 \times \mathbb{R}^{N-n} \mid \eta - g(\xi) \in W_2\}$ und definiere

$$h: U \rightarrow V, \quad h(x, y) = (x, y - g(x)).$$

Dann ist h ein C^p -Diffeomorphismus; wir können seine Inverse sogar angeben: Es handelt sich um die Abbildung $k(\xi, \eta) = (\xi, \eta + g(\xi))$. Wir zeigen nun die Mengengleichheit aus der Behauptung, indem wir beide Inklusionen nachweisen.

Sei zuerst $(x, y) \in M \cap U$. Dann $h(x, y) \in V$ nach Definition von V und $h(x, y) = (x, y - g(x)) = (x, 0)$ nach Wahl von g .

Sei nun $(\xi, 0) \in V$. Dann $k(\xi, 0) = (\xi, g(\xi)) \in M \cap U$.

(b) $\implies$ (a): Seien $h_j, j = 1, \dots, N$, die Komponenten von h . Setze

$$f: U \rightarrow \mathbb{R}^{N-n}, \quad f(x) = \begin{pmatrix} h_{n+1}(x) \\ \vdots \\ h_N(x) \end{pmatrix}.$$

Da $Dh(a)$ invertierbar ist, sind die Zeilen von $Df(a)$ linear unabhängig. Die Gleichheit $M = f^{-1}(0)$ folgt sofort aus der Definition von h . $\square$

24. Untermannigfaltigkeiten des $\mathbb{R}^N$

24.5 Definition. Seien X, Y zwei metrische Räume. Eine bijektive Abbildung $f: X \rightarrow Y$ heißt *Homöomorphismus*, wenn sowohl f als auch f^{-1} stetig sind. Wenn es einen Homöomorphismus zwischen X und Y gibt, so sind die beiden Räume *homöomorph*.

24.6 Beispiel. $\mathbb{R}$ ist homöomorph zum Einheitsintervall $]0, 1[$.

Beweis. $f(x) = \frac{1}{\pi} (\arctan(x) + \frac{\pi}{2})$ ist ein Homöomorphismus. $\square$

24.7 Definition. Sei $W \subseteq \mathbb{R}^n$ offen und sei $\varphi: W \rightarrow \mathbb{R}^N$ von der Klasse C^1 . Falls $\text{Rang } D\varphi(x) = n$ für alle $x \in W$, so ist φ eine *Immersion*.

24.8 Beispiel. (a) Die Abbildung $f: \mathbb{R} \rightarrow \mathbb{R}^2, t \mapsto \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$, ist eine C^∞ -Immersion.

(b) Die Abbildung

$$g: \mathbb{R} \rightarrow \mathbb{R}^2, \quad g(t) = \begin{cases} \begin{pmatrix} 0 \\ t^2 \end{pmatrix}, & t < 0, \\ \begin{pmatrix} t^2 \\ 0 \end{pmatrix}, & t \geq 0 \end{cases}$$

ist zwar von der Klasse C^1 , aber keine Immersion.

24.9 Bemerkung. Wenn es eine Immersion $\varphi: U \rightarrow \mathbb{R}^N$ für eine offene Menge $U \subseteq \mathbb{R}^n$ gibt, so ist $n \leq N$.

24.10 Satz. Sei $M \subseteq \mathbb{R}^N$. Dann sind äquivalent:

(a) M ist eine n -dimensionale C^p -Untermannigfaltigkeit des $\mathbb{R}^N$.

(b) Für jedes $a \in M$ existieren eine offene Umgebung $U \subseteq M$ von a in M , eine offene Teilmenge $V \subseteq \mathbb{R}^n$ und eine homöomorphe Abbildung $\varphi: U \rightarrow V$, so dass φ^{-1} als Abbildung in den $\mathbb{R}^N$ eine Immersion der Klasse C^p ist.

Beweis. (a) $\implies$ (b): Es gibt einen C^p -Diffeomorphismus $h: U_1 \rightarrow V_1$ mit $h(M \cap U_1) = V_1 \cap (\mathbb{R}^n \times \{0\})$. Setze $V = \{\xi \in \mathbb{R}^n \mid (\xi, 0) \in V_1\}$ und $U = M \cap U_1$ und

$$\psi: V \rightarrow \mathbb{R}^N, \quad \psi(\xi) = h^{-1}(\xi, 0).$$

Dann ist ψ eine Immersion. Ihre Inverse ist gegeben durch $\varphi = \pi \circ h$, wobei $\pi: \mathbb{R}^N \rightarrow \mathbb{R}^n$ die Projektion auf die ersten n Komponenten ist.

(b) $\implies$ (a): Ohne Einschränkung ist $\varphi(a) = 0$. Die Komponenten von $\psi = \varphi^{-1}$ bezeichnen wir mit $\psi_1, \dots, \psi_N$. Da ψ eine Immersion ist, sind $\frac{\partial \psi}{\partial \xi_1}(0), \dots, \frac{\partial \psi}{\partial \xi_n}(0)$ linear unabhängig. Man kann sie also zu einer Basis

$$\left(\frac{\partial \psi}{\partial \xi_1}(0), \dots, \frac{\partial \psi}{\partial \xi_n}(0), v^{(1)}, \dots, v^{(N-n)} \right)$$

des $\mathbb{R}^N$ ergänzen. Wir setzen

$$g: V \times \mathbb{R}^{N-n} \rightarrow \mathbb{R}^N, \quad g(\xi, \eta) = \psi(\xi) + \sum_{j=n+1}^N \eta_{j-n} v^{(j)}.$$

Dann ist

$$\begin{aligned} Dg(0,0) &= \begin{pmatrix} \frac{\partial \psi_1}{\partial \xi_1}(0) & \dots & \frac{\partial \psi_1}{\partial \xi_n}(0) & v_1^{(1)} & \dots & v_{(N-n)}^{(1)} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial \psi_N}{\partial \xi_1}(0) & \dots & \frac{\partial \psi_N}{\partial \xi_n}(0) & v_1^{(N)} & \dots & v_{(N-n)}^{(N)} \end{pmatrix} \\ &= (D\psi(0) \quad v^{(1)} \quad \dots \quad v^{(N-n)}) \end{aligned}$$

invertierbar. Aus dem Umkehrsatz folgt die Existenz von offenen Umgebungen $\tilde{V}$ von 0 und $\tilde{U}$ von α im $\mathbb{R}^N$, so dass $g: \tilde{V} \rightarrow \tilde{U}$ ein C^p -Diffeomorphismus ist. Wir dürfen annehmen, dass $\tilde{V} \cap \mathbb{R}^n \times \{0\} \subseteq V \times \{0\}$ und $\tilde{U} \cap M \subseteq U$. Wenn $\pi_2: \mathbb{R}^N = \mathbb{R}^n \times \mathbb{R}^{N-n} \rightarrow \mathbb{R}^{N-n}$ die Projektion auf die letzten $(N-n)$ Komponenten ist, dann setzen wir

$$f: \tilde{U} \rightarrow \mathbb{R}^{N-n}, \quad f(x) = \pi_2(g^{-1}(x)).$$

Die Kettenregel zeigt, dass $Df(\alpha)$ das Produkt einer invertierbaren Matrix mit einer vom Rang $(N-n)$ ist, also vollen Rang hat. Wir zeigen zum Schluss $f^{-1}(0) = M \cap \tilde{U}$, indem wir beide Inklusionen nachweisen.

Sei zuerst $x \in \tilde{U}$ mit $f(x) = 0$. Dann $\pi_2(g^{-1}(x)) = 0$, also $g^{-1}(x) = (y, 0)$ für ein $y \in V$. Daraus folgt $x = g(g^{-1}(x)) = g(y, 0) = \psi(y) \in \tilde{U} \cap M$.

Ist andererseits $x \in M \cap \tilde{U}$, dann $x = \psi(y)$ für $y = \varphi(x)$ und daher $g(y, 0) = \psi(y) = x$. Das bedeutet $g^{-1}(x) = (y, 0)$ und $f(x) = \pi_2(g^{-1}(x)) = 0$. $\square$

In der Vorlesung am 20.1.2026 wurde ergänzend die Menge

$$M := \left\{ \begin{pmatrix} \cos(t) \\ \sin(2t) \end{pmatrix} \mid 0 < t < 2\pi \right\}$$

betrachtet, der Rand eines Schmetterlings mit punktförmigem Rumpf. Dabei ist M durch eine Immersion definiert. Trotzdem ist M an der Stelle $t = \pi/2$ bzw. $t = 3\pi/2$ lokal keine Mannigfaltigkeit. An der Stelle $t = \pi/2$ bzw. $t = 3\pi/2$ gibt es auch keine homöomorphe Abbildung wie in Teil (b) gefordert. Hier ist die Homöomorphie-Eigenschaft also eine wesentliche Voraussetzung des obigen Satzes. Umgekehrt gibt es auch Beispiele von Abbildungen, die keine Immersion sind und deren Bild trotzdem eine Mannigfaltigkeit ist, nur wird dies durch die gewählte Abbildung nicht bewiesen.

24.11 Definition. Jedes solche φ ist eine *Karte* von M . Eine Menge $\mathcal{A}$ von Karten ist ein *Atlas* der Untermannigfaltigkeit M , wenn jeder Punkt aus M im Definitionsbereich wenigstens einer Karte aus $\mathcal{A}$ liegt.

Wenn $\varphi: U \rightarrow V$ mit $U \subseteq M$ und $V \subseteq \mathbb{R}^n$ eine Karte ist, dann bezeichnet man φ^{-1} als *lokale Parametrisierung* von M .

24. Untermannigfaltigkeiten des $\mathbb{R}^N$

24.12 Beispiel. Die Kugelkoordinaten, eingeschränkt auf $r = 1$, definieren eine Karte der S^2 , deren Definitionsbereich die Menge $\{(x, y, z) \in S^2 \mid y \neq 0 \text{ oder } x > 0\}$ ist. Analog kann man Kugelkoordinaten schaffen, deren Definitionsbereich die Menge $\{(x, y, z) \in S^2 \mid z \neq 0 \text{ oder } x < 0\}$ ist. Diese beiden Karten bilden einen Atlas der S^2 . Abbildung 24.3 zeigt die jeweils ausgelassenen Kurven.

In Kaballo, Aufgabe 21.2, werden Polarkoordinaten für beliebige Dimension erklärt. Daraus erhält man sofort einen Atlas für die S^{n-1} .

Für die Sphäre S^{n-1} gibt es aber auch andere Atlanten. Wir setzen dazu $S_{j,\pm}^{n-1} = \{x \in S^{n-1} \mid \pm x_j > 0\}$ und

$$\varphi_{j,\pm}: S_{j,\pm}^{n-1} \rightarrow B_1(0) \subset \mathbb{R}^{n-1}, \quad x \mapsto (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n).$$

Die Inverse ist dann

$$\begin{aligned} \psi_{j,\pm}(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \\ = \left(x_1, \dots, x_{j-1}, \pm \sqrt{1 - x_1^2 - \dots - x_{j-1}^2 - x_{j+1}^2 - \dots - x_n^2}, x_{j+1}, \dots, x_n \right). \end{aligned}$$

Abbildung 24.4 zeigt die sechs Karten, aus denen dieser Atlas der S^2 besteht.

Bemerkung. Wir zeigen in der Analysis III: Es sei M eine C^p -Untermannigfaltigkeit und es seien $\varphi_1: U_1 \rightarrow V_1$ und $\varphi_2: U_2 \rightarrow V_2$ zwei Karten, so dass $U_1 \cap U_2 \neq \emptyset$. Setzt man $W_1 = \varphi_1(U_1 \cap U_2)$ und $W_2 = \varphi_2(U_1 \cap U_2)$, so ist $\varphi_1 \circ \varphi_2^{-1}: W_2 \rightarrow W_1$ ein C^p -Diffeomorphismus. Die Abbildung der Form $\varphi_1 \circ \varphi_2^{-1}$ heißen *Kartenwechsel*.

Definiert man abstrakte C^p -Mannigfaltigkeiten, so kann man die Forderung, dass das Inverse einer Karte eine Immersion der Klasse C^p ist, nicht mehr stellen. Statt dessen fordert man, dass alle Kartenwechsel von der Klasse C^p sind.

24.13 Satz. Für $a, b, c, d > 0$ und $K > 0$ sei $M_K = \{(\xi, \eta) \in]0, \infty[^2 \mid f(\xi, \eta) = K\}$, wobei $f(\xi, \eta) = \xi^c e^{-d\xi} \eta^a e^{-b\eta}$. Ferner sei $K_0 = \left(\frac{c}{de}\right)^c \left(\frac{a}{be}\right)^a$. Für $0 < K < K_0$ ist M_K eine eindimensionale C^∞ -Untermannigfaltigkeit des $\mathbb{R}^2$. Sie besitzt einen Atlas aus Lösungen der Räuber-Beute-Gleichung. Für $K = K_0$ besteht M_K aus einem einzigen Punkt und für $K > K_0$ ist M_K leer.

Beweis. Für (ξ, η) mit $f(\xi, \eta) = K$ gilt $Df(\xi, \eta) = \left(\frac{c}{\xi} - d, \frac{a}{\eta} - b\right) K$. Also ist $\left(\frac{c}{\xi}, \frac{a}{\eta}\right)$ die einzige kritische Stelle von f in $]0, \infty[$. Der Funktionswert an dieser Stelle ist K_0 . Für $\|(\xi, \eta)\|_1 = R$ gilt $f(\xi, \eta) \leq R^{a+c} \exp(-\min(b, d)R)$, wird also für große R kleiner als K_0 . Da ferner $f(0, \eta) = f(\xi, 0) = 0$, besitzt f in $[0, R]^2$ genau eine globale Maximalstelle, nämlich $\left(\frac{c}{d}, \frac{a}{b}\right)$. Für alle (ξ, η) gilt daher $f(\xi, \eta) \leq K_0$, woraus die letzte Aussage der Behauptung folgt. Die vorletzte folgt daraus, dass die globale Maximalstelle strikt ist, wie die Hesse-Matrix zeigt. Es gilt nämlich $H_f\left(\frac{c}{d}, \frac{a}{b}\right) = \begin{pmatrix} -\frac{d^2}{c} & 0 \\ 0 & -\frac{b^2}{a} \end{pmatrix}$.

Sei also nun $0 < K < K_0$ und sei $(\xi_0, \eta_0) \in M_K$ beliebig. Ferner sei $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} : \mathbb{R} \rightarrow \mathbb{R}^2$ die Lösung der Anfangswertaufgabe

$$\begin{aligned} \dot{x} &= ax - bxy, & x(0) &= \xi_0 \\ \dot{y} &= -cy + dxy, & y(0) &= \eta_0. \end{aligned}$$

Im Fall $(\dot{\psi}_1(0), \dot{\psi}_2(0)) = (0, 0)$ würden $\psi_2(0) = \frac{a}{b}$ und $\psi_1(0) = \frac{c}{d}$ und daher $K = K_0$ gelten. Da wir das bereits ausgeschlossen haben, folgt $(\dot{\psi}_1(0), \dot{\psi}_2(0)) \neq (0, 0)$. Ohne Einschränkung behandeln wir nur den Fall $\dot{\psi}_1(0) \neq 0$, also $a - b\eta \neq 0$. Wegen $\frac{\partial f}{\partial \eta}(\xi, \eta) = \left(\frac{a}{\eta} - b\right) K$ gilt daher $\frac{\partial f}{\partial \eta}(\xi_0, \eta_0) \neq 0$ und wir können den Satz über implizite Abbildungen anwenden. Es gibt also eine Umgebung U von ξ_0 , eine Umgebung V von η_0 und eine C^∞ -Abbildung $g: U \rightarrow V$, so dass

$$M_K \cap (U \times V) = \left\{ \begin{pmatrix} x \\ g(x) \end{pmatrix} \middle| x \in U \right\}.$$

Wegen $\dot{\psi}_1(0) \neq 0$ gibt es Umgebungen $W \subseteq U$ von ξ_0 und T von 0 , so dass $\psi_1: T \rightarrow W$ eine C^∞ -Inverse h besitzt. Wir setzen

$$\varphi: M_K \cap (W \times V) \rightarrow T, \quad \begin{pmatrix} \xi \\ \eta \end{pmatrix} \mapsto h(\xi).$$

Für $\begin{pmatrix} \xi \\ \eta \end{pmatrix} \in M_K \cap (W \times V)$ gilt $\eta = g(\xi)$ und daher

$$\psi \left(\varphi \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right) = \varphi(h(\xi)) = \begin{pmatrix} \psi_1(h(\xi)) \\ \psi_2(h(\xi)) \end{pmatrix} = \begin{pmatrix} \xi \\ \psi_2(h(\xi)) \end{pmatrix} = \begin{pmatrix} \xi \\ g(\xi) \end{pmatrix} = \begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

und

$$\varphi(\psi(t)) = \varphi \begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix} = h(\psi_1(t)) = t$$

. Also sind ψ und φ invers zueinander. Speziell ist φ ein Homöomorphismus. Dass ψ eine Immersion ist, hatten wir schon bei der Untersuchung der kritischen Punkte von f gesehen. $\square$

24.14 Satz. Seien $a, b, c, d > 0$, und sei $\rho: \mathbb{R} \rightarrow]0, \infty[^2$ eine Lösung der Räuber-Beute-Gleichungen

$$\begin{aligned} \dot{x} &= ax - bxy, \\ \dot{y} &= -cy + dxy. \end{aligned}$$

Dann ist ρ periodisch, d. h. es gibt $T > 0$ mit $\rho(t + T) = \rho(t)$ für alle $t \in \mathbb{R}$.

24. Untermannigfaltigkeiten des $\mathbb{R}^N$

Beweis. Nach Satz 21.7 existiert $K > 0$, so dass $\rho(t) \in M_K$ für alle $t \in \mathbb{R}$. Falls $K = K_0 := \left(\frac{c}{de}\right)^c \left(\frac{a}{be}\right)^a$, so ist M_K einelementig und ρ daher konstant.

Wir brauchen also nur den Fall $K < K_0$ zu behandeln. Da M_K kompakt ist, besitzt die Folge $(\rho(n))_{n \in \mathbb{N}}$ eine konvergente Teilfolge. Wir bezeichnen den Grenzwert mit (ξ, η) . Wegen Satz 24.13 besitzt M_K in einer Umgebung U von (ξ, η) eine Karte $\varphi: M_K \cap U \rightarrow]\alpha, \beta[$, deren Inverse ψ die Räuber-Beute-Gleichung mit der Anfangsbedingung $\psi(0) = (\xi, \eta)$ löst. Nach Wahl von (ξ, η) gibt es $n < m$, so dass $\rho(n), \rho(m) \in U$ und $m - n > \beta - \alpha$. Es seien $t_1 = \varphi(\rho(n))$ und $t_2 = \varphi(\rho(m))$. Aus der Eindeutigkeit der Lösung der Anfangswertaufgabe folgt $\psi(t) = \rho(n + t - t_1)$. Das impliziert $\rho(m) = \psi(t_2) = \rho(n + t_2 - t_1) = \rho(m - T)$ für $T = m - n - t_2 + t_1 > 0$. Wegen der Eindeutigkeit der Lösung der Anfangsbedingung folgt $\rho(t) = \rho(t + T)$ für alle t . $\square$

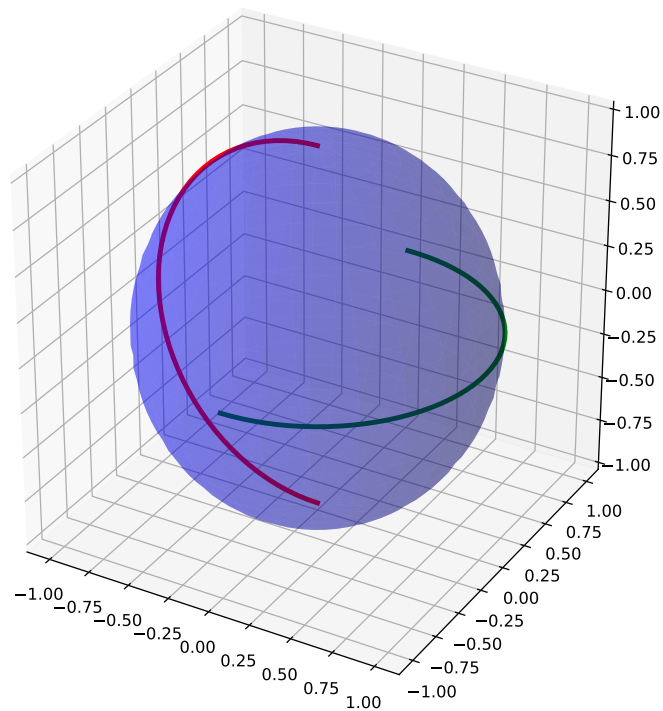


Abbildung 24.3.: Atlas der S^2 in Kugelkoordinaten

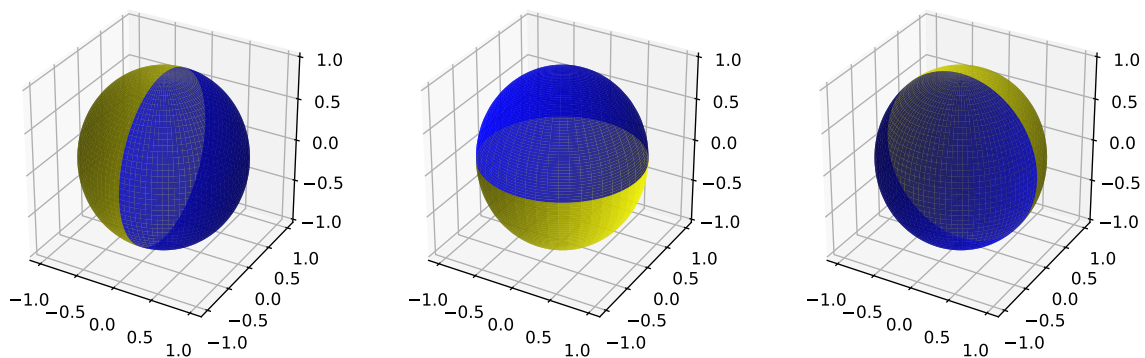


Abbildung 24.4.: Atlas der S^2 aus Graphen

25. Tangentialräume und Extrema unter Nebenbedingungen

25.1 Definition. Sei $M \subseteq \mathbb{R}^N$ und sei $a \in M$. Ein Vektor $v \in \mathbb{R}^N$ heißt *Tangentialvektor* an M in a , falls es ein offenes Intervall $I \subseteq \mathbb{R}$ mit $0 \in I$ und eine C^1 -Abbildung $\gamma: I \rightarrow M$ gibt mit $\gamma(0) = a$ und $\gamma'(0) = v$.

Die Menge $T_a(M)$ aller Tangentialvektoren an M in a bezeichnet man als den *Tangentialraum* an M in a .

25.2 Beispiel. Sei $N = (0, 0, 1)$ der Nordpol der S^2 und sei v ein Vektor der Form $v = (x, y, 0)$. Dann $v \in T_N(S^2)$.

Beweis. Sei $(x, y) \neq (0, 0)$ und $\epsilon := \|(x, y)\|_2^{-2}$. Definiere $\gamma:]-\epsilon, \epsilon[\rightarrow S^2$ durch $\gamma(t) = (tx, ty, \sqrt{1 - t^2x^2 - t^2y^2})$. Dann $\gamma'(0) = v$. $\square$

25.3 Satz. Sei M eine n -dimensionale C^p -Untermannigfaltigkeit des $\mathbb{R}^N$, sei $a \in M$. Laut Definition der Untermannigfaltigkeit gibt es eine Umgebung U von a und eine Abbildung $g: U \rightarrow \mathbb{R}^{N-n}$ von der Klasse C^p mit $M \cap U = \{x \in U \mid g(x) = 0\}$ und $\text{Rang}(Dg(a)) = N - n$.

Nach Satz 24.10 gibt es eine lokale Parametrisierung $\psi: W \rightarrow M$ und $b \in W$ mit $\psi(b) = a$.

Mit diesen Abbildungen gilt

$$\{D\psi(b)y \mid y \in \mathbb{R}^n\} = T_a(M) = \{v \in \mathbb{R}^N \mid Dg(a)v = 0\}.$$

Beweis. Setze $E = \{D\psi(b)y \mid y \in \mathbb{R}^n\}$ und $F = \{v \in \mathbb{R}^N \mid Dg(a)v = 0\}$. Beide Mengen sind Untervektorräume des $\mathbb{R}^N$. Es gelten $\dim E = \text{Rang}(D\psi(b)) = n$ und $\dim F = N - \text{Rang}(Dg(a)) = n$. Es genügt daher der Nachweis von

$$E \subseteq T_a(M) \subseteq F.$$

Sei zuerst $w = D\psi(b)y \in E$. Für ein hinreichend kleines $\epsilon > 0$ setze $I =]-\epsilon, \epsilon[$ und definiere

$$\gamma: I \rightarrow M, \quad \gamma(t) = \psi(b + ty).$$

Dann $\gamma'(0) = w$ nach der Kettenregel, und daher $w \in T_a(M)$.

Sei nun $v = \gamma'(0) \in T_a(M)$, wobei $\gamma: I \rightarrow M$ eine Abbildung wie in Definition 25.1 ist. Dann $g \circ \gamma = 0$, also nach der Kettenregel $Dg(a)\gamma'(0) = 0$. Dies zeigt $v \in F$. $\square$

25. Tangentialräume und Extrema unter Nebenbedingungen

25.4 Korollar. Der Tangentialraum $T_a(M)$ an eine n -dimensionale Untermannigfaltigkeit ist ein n -dimensionaler Vektorraum.

25.5 Beispiel. (a) $T_N(S^2) = \mathbb{R}^2 \times \{0\}$.

(b) Allgemein gilt $T_N(S^{n-1}) = \mathbb{R}^{n-1} \times \{0\}$, wenn $N = (0, \dots, 0, 1)$ den Nordpol der S^{n-1} bezeichnet. Den Fall $n = 3$ zeigt Abbildung 25.1.

Beweis. Für $f(x) = x_1^2 + \dots + x_n^2 - 1$ gilt $S^{n-1} = \{x \in \mathbb{R}^n | f(x) = 0\}$. Wegen Satz 25.3 gilt

$$T_N S^{n-1} = \{v \in \mathbb{R}^n | Df(N)v = 0\}.$$

Die Behauptung folgt nun aus

$$Df(N) = (0, \dots, 0, 2). \quad \square$$

(c) Das kartesische Blatt ist keine Untermannigfaltigkeit des $\mathbb{R}^2$.

Beweis. Sei $K = \{(x, y) \in \mathbb{R}^2 | x^3 + y^3 - 3xy\}$ das kartesische Blatt. Wir zeigen, dass $(1, 0) \in T_0 K$ und $(0, 1) \in T_0 K$. Dazu konstruieren wir ein offenes Intervall I mit $0 \in I$ und eine C^∞ -Funktion $g: I \rightarrow \mathbb{R}$, so dass $\gamma(t) := (t, t^2 g(t)) \in K$ für alle $t \in I$. Um g zu bestimmen, setzen wir $a(t, s) := 1 - 3s + t^3 s^3$, so ist a von der Klasse C^∞ und erfüllt

$$\frac{\partial a}{\partial s}(t, s) = -3 + 3t^3 s^2,$$

insbesondere also $\frac{\partial a}{\partial s}(0, 0) = -3$. Aus dem Satz über implizite Funktionen folgt daher die Existenz einer Nullumgebung $I \subseteq \mathbb{R}$ und einer C^∞ -Abbildung $g: I \rightarrow \mathbb{R}$, so dass $a(t, g(t)) = 0$ für alle $t \in I$. Nun gilt

$$f(\gamma(t)) = t^3 + t^6 g(t)^3 - 3t^3 g(t) = t^3(1 + t^3 g(t)^3 - 3g(t)) = a(t, g(t)) = 0.$$

Also ist $\gamma'(0) = (1, 0)$ ein Tangentialvektor. Vertauschung von x und y zeigt, dass auch $(0, 1)$ ein Tangentialvektor ist.

Wir haben zwei linear unabhängige Vektoren in $T_0 K$ gefunden. Also besitzt $T_0 K$ die Dimension 2. Das ist ein Widerspruch zu Korollar 25.4. $\square$

25.6 Definition. Sei $U \subseteq \mathbb{R}^N$ offen und sei $g: U \rightarrow \mathbb{R}^m$ eine Abbildung. Sei $f: U \rightarrow \mathbb{R}$ eine Funktion. Ein Punkt $a \in \mathbb{R}^N$ mit $g(a) = 0$ heißt lokale Maximalstelle von f unter der Nebenbedingung $g(x) = 0$, wenn es eine Umgebung V von a gibt, so dass $f(a) \geq f(x)$ für alle $x \in V$ mit $g(x) = 0$.

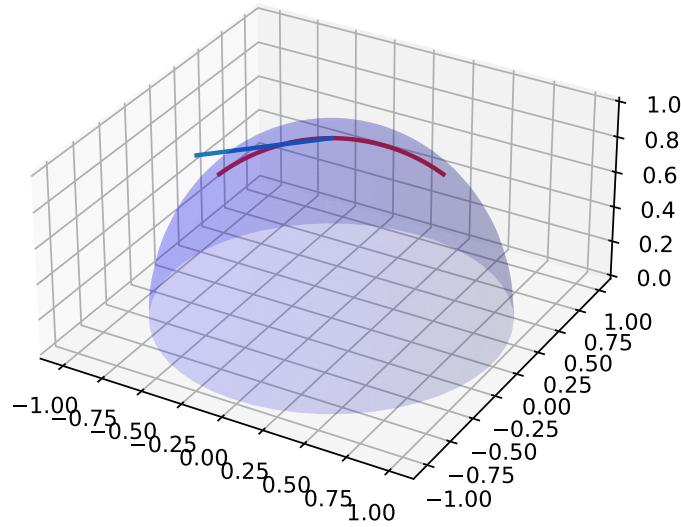


Abbildung 25.1.: Eine Tangente an die S^2 im Norpol

25.7 Satz (Methode der Lagrange-Multiplikatoren). Sei $U \subseteq \mathbb{R}^N$ offen und sei $g: U \rightarrow \mathbb{R}^m$ eine Abbildung von der Klasse C^1 mit $\text{Rang } D(g(x)) = m$ für alle $x \in U$. Sei $f: U \rightarrow \mathbb{R}$ eine Funktion von der Klasse C^1 . Ist a eine lokale Maximalstelle von f unter der Nebenbedingung $g(x) = 0$, so gibt es $\mu_1, \dots, \mu_m \in \mathbb{R}$ mit

$$\nabla f(a) + \mu_1 \nabla g_1(a) + \dots + \mu_m \nabla g_m(a) = 0. \quad (25.1)$$

Für den Beweis benutzen wir eine Definition und zwei Lemmata.

25.8 Definition. Es sei $E \subset \mathbb{R}^N$ ein $\mathbb{R}$ -Vektorraum. Sein *Orthogonalraum* ist definiert als

$$E^\perp = \{w \in \mathbb{R}^N \mid w \cdot v = 0 \text{ für alle } v \in E\}.$$

25.9 Lemma. Es sei $M \subset \mathbb{R}^N$ eine n -dimensionale Untermannigfaltigkeit, es sei $a \in M$ und es sei $g: U \rightarrow \mathbb{R}^{N-n}$ eine C^1 -Abbildung mit $M \cap U = \{x \mid g(x) = 0\}$ und $a \in U$ und $\text{Rang } Dg(a) = N-n$. Dann spannen die Vektoren $\nabla g_1(a), \dots, \nabla g_{N-n}(a)$ den Orthogonalraum $T_a(M)^\perp$ auf.

Beweis. Wegen Satz 25.3 liegen die $\nabla g_j(a)$ jedenfalls in $(T_a M)^\perp$. Wegen der Voraussetzung über den Rang von $Dg(a)$ ist das Tupel $(\nabla g_1(a), \dots, \nabla g_{N-n}(a))$ linear unabhängig. Wegen $\dim T_a M = n$ besitzt sein Orthogonalraum $(T_a M)^\perp$ die Dimension $N - n$. Daher spannt $(\nabla g_1(a), \dots, \nabla g_{N-n}(a))$ ihn auf. $\square$

25.10 Lemma. Für f und g wie in Satz 25.7 ist

$$M = \{x \in U \mid g(x) = 0\}$$

25. Tangentialräume und Extrema unter Nebenbedingungen

eine $(N - n)$ -dimensionale C^1 -Untermannigfaltigkeit des $\mathbb{R}^N$ und es gilt

$$\nabla f(a) \in (T_a M)^\perp.$$

Beweis. Der erste Teil der Behauptung folgt sofort aus den Eigenschaften von g .

Wir wählen eine offene Umgebung $V_1 \subseteq M$ von a in M , eine offene Teilmenge $W \subseteq \mathbb{R}^n$ und eine lokale Parametrisierung $\psi: W \rightarrow \mathbb{R}^N$ mit $\psi(W) = V_1$. Sei b so, dass $\psi(b) = a$. Dann besitzt $f \circ \psi$ in b ein lokales Maximum. Also

$$0 = \nabla(f \circ \psi)(b) = \nabla f(a) \cdot D\psi(b).$$

Daher $0 = \nabla f(a) \cdot (D\psi(b) \cdot w)$ für alle $w \in \mathbb{R}^n$. Wegen Satz 25.3 folgt $\nabla f(a) \in (T_a M)^\perp$. $\square$

25.11 Bemerkung. (a) Dieser Satz gilt genauso für lokale Minima. Die Funktion f bezeichnet man als *Zielfunktion*, die Gleichung $g(x) = 0$ als *Nebenbedingung*. Die Zahlen $\mu_1, \dots, \mu_m$ sind die *Lagrangeschen Multiplikatoren*.

(b) Setzt man

$$h(x, \lambda) = f(x) + \sum_{j=1}^m \lambda_j g_j(x),$$

so ist $\nabla h(a, \mu) = 0$ äquivalent (25.1) und $g(x) = 0$.

25.12 Beispiel. Der Rauminhalt einer Konservendose mit Radius r und Höhe h beträgt $\pi r h^2$, ihre Oberfläche $2\pi r^2 + 2\pi r h$. Für welche Werte von r und h ist die Oberfläche einer 1l-Konservendose minimal?

$$f(r, h) = 2\pi r^2 + 2\pi r h,$$

$$g(r, h) = \pi r^2 h - 1.$$

Es gibt nur einen Lagrangeschen Multiplikator λ . Wir erhalten drei Gleichungen

$$4\pi r + 2\pi h + \lambda 2\pi r h = 0$$

$$2\pi r + \lambda \pi r^2 = 0,$$

$$\pi r^2 h = 1.$$

Aus der zweiten folgt $\lambda = -2/r$. Einsetzen in die erste gibt $h = 2r$ und schließlich $r = 1/\sqrt[3]{2\pi}$.

25.13 Beispiel. Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Die quadratische Form $f(x) = x^T A x$ nimmt auf der S^n ihr Maximum und ihr Minimum an. Sei $v \in S^n$ ein Punkt, in dem f ihr Maximum annimmt. Wir wollen zeigen, dass v ein Eigenvektor von A ist. Für $g(x) = 1 - \sum_{j=1}^n x_j^2$ gilt $S^n = g^{-1}(0)$. Es gilt $f(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$, also $Df(x) = 2Ax$. Die Lagrange-Gleichung hat also die Gestalt

$$2Ax - 2\lambda x = 0.$$

Daher ist v ein Eigenvektor und λ der zugehörige Eigenwert.

25.14 Satz. Sei $\nabla h(a, \mu) = 0$, sei $M = \{x | g(x) = 0\}$ und sei H die Hessematrix von $x \mapsto h(x, \mu)$ in a . Falls die Einschränkung von H auf $T_a(M)$ positiv definit ist, so besitzt f in a ein striktes, lokales Minimum unter der Nebenbedingung $g = 0$. Falls sie negativ definit ist, so besitzt f dort ein striktes lokales Maximum, und falls sie indefinit ist, so besitzt sie kein lokales Extremum in a .

Beweis. Kabblo, Analysis 2, Satz 24.6. □

25.15 Bemerkung. Für die Anwendung dieses Satzes benötigen wir eine Darstellung des Tangentialraums $T_a(M)$. Dabei ist die Darstellung von $T_a(M)$ wie in Satz 25.3 nur unter der Voraussetzung möglich, dass $Dg(a)$ maximalen Rang besitzt.

25.16 Beispiel. Wir suchen das Maximum von $f(x, y, z) := xy + xz + yz$ unter der Nebenbedingung $0 = g(x, y, z) := xyz - 8$. Wir haben

$$h(x, y, z, \lambda) = xy + yz + xz - \lambda(xyz - 8).$$

Die Lagrange-Gleichungen sind

$$0 = y + z - \lambda yz$$

$$0 = x + z - \lambda xz$$

$$0 = x + y - \lambda xy$$

$$8 = xyz.$$

Wir multiplizieren die erste Gleichung mit x , die zweite mit y und die dritte mit z und kommen auf die Lösung $a = (2, 2, 2)$ und $\mu = 1$. Für beliebiges λ ist die Hessesche von $(x, y, z) \mapsto h(x, y, z, \lambda)$ gleich

$$\begin{pmatrix} 0 & 1 - \lambda z & 1 - \lambda y \\ 1 - \lambda z & 0 & 1 - \lambda x \\ 1 - \lambda y & 1 - \lambda x & 0 \end{pmatrix}.$$

Setzt man dort $x = y = z = 2$ und $\lambda = 1$ ein, so erhält man

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

Nach dem Hurwitz-Kriterium ist H indefinit. Der Tangentialraum $T_a(M)$ ist orthogonal zu $\nabla g(a)$, wird also aufgespannt von $(1, -1, 0)$ und $(1, 0, -1)$. Beide sind Eigenvektoren von H zum Eigenwert 1. Daher ist die Einschränkung von H auf $T_a M$ gleich der Einheitsmatrix. Aus dem Satz folgt, dass f in a ein striktes lokales Minimum unter der Nebenbedingung $g(x) = 0$ hat.

26. Die Weglänge

Den Begriff des Weges hatten wir in 4.18 definiert.

26.1 Definition. Sei $\alpha: [a, b] \rightarrow \mathbb{R}^n$ ein Weg. Seine Bildmenge $\{\alpha(t) \mid a \leq t \leq b\}$ wird als *Spur* bezeichnet.

26.2 Definition. Eine *Zerlegung* von $[a, b]$ ist ein $(m+1)$ -Tupel $(t_0, t_1, \dots, t_m)$ mit

$$a = t_0 < t_1 < \dots < t_m = b.$$

26.3 Definition. $\alpha: [a, b] \rightarrow \mathbb{R}^n$ heißt *stetig differenzierbar*, wenn es ein offenes Intervall I mit $[a, b] \subset I$ gibt, so dass sich α zu einer Abbildung der Klasse C^1 von I nach $\mathbb{R}^n$ fortsetzen lässt.

α heißt *stückweise stetig differenzierbar*, wenn es eine Zerlegung $(t_0, t_1, \dots, t_m)$ von $[a, b]$ gibt, so dass die Einschränkungen $\alpha|_{[t_{i-1}, t_i]}$ für $i = 1, \dots, m$ stetig differenzierbar sind.

26.4 Beispiel. (a) Wenn I ein offenes Intervall und $f: I \rightarrow \mathbb{R}$ eine C^1 -Funktion ist, dann ist für jedes kompakte Intervall $[a, b]$ der Graph der Einschränkung von f auf $[a, b]$ ein stetig differenzierbarer Weg.

(b) Der Weg $\alpha: [0, 1] \rightarrow \mathbb{R}^2$,

$$\alpha(x) = \begin{cases} \begin{pmatrix} x \\ x \sin(\frac{1}{x}) \end{pmatrix}, & x > 0, \\ 0, & x = 0, \end{cases}$$

ist nicht stückweise stetig differenzierbar.

Beweis. Für $x > 0$ gilt

$$\alpha'(x) = \begin{pmatrix} 1 \\ \sin(\frac{1}{x}) - \frac{1}{x} \cos(\frac{1}{x}) \end{pmatrix}.$$

Für $x_n = \frac{1}{2\pi n}$, $n \in \mathbb{N}$, gilt folglich

$$\alpha'(x_n) = \begin{pmatrix} 1 \\ -2\pi n \end{pmatrix}.$$

Daher existiert $\lim_{x \searrow 0} \alpha'(x)$ nicht. □

26. Die Weglänge

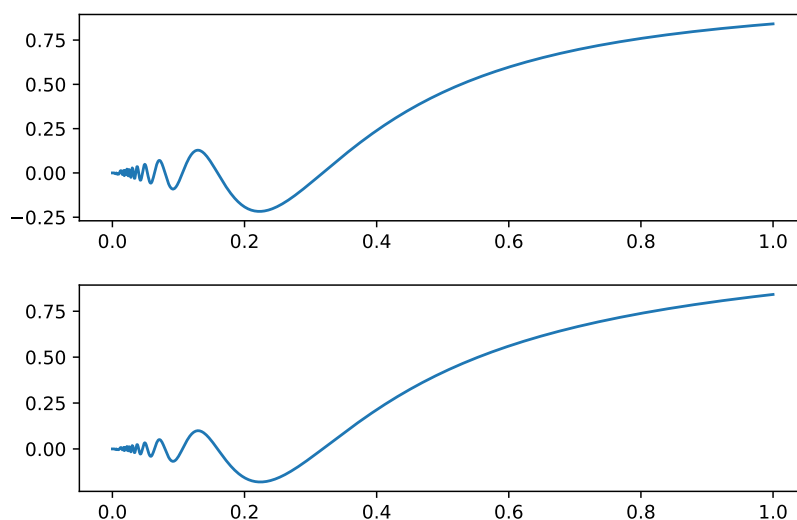


Abbildung 26.1.: Graph von $x \sin(\frac{1}{x})$ und von $x^{9/8} \sin(\frac{1}{x})$. Nur der untere der beiden Wege ist rektifizierbar.

26.5 Definition. Sei $\alpha: [a, b] \rightarrow \mathbb{R}^n$ ein stetiger Weg. Mit $\mathcal{Z}$ bezeichnen wir die Menge aller Zerlegungen von $[a, b]$. Ist $Z = (t_0, t_1, \dots, t_m) \in \mathcal{Z}$, so sei

$$L(\alpha, Z) = \sum_{i=1}^m \|\alpha(t_i) - \alpha(t_{i-1})\|_2.$$

α heißt *rektifizierbar*, wenn die Menge $\{L(\alpha, Z) \mid Z \in \mathcal{Z}\}$ beschränkt ist. In diesem Fall ist

$$L(\alpha) = \sup\{L(\alpha, Z) \mid Z \in \mathcal{Z}\}$$

die *Länge* von α .

Die Länge hängt nur von der Spur des Weges ab.

26.6 Beispiel. Die Kurve α aus Beispiel 26.4 ist nicht rektifizierbar.

Beweis. Für ein hinreichend großes N sei die Zerlegung Z definiert als die monotone Anordnung von

$$\{0, 1\} \cup \left\{ \frac{2}{\pi n} \mid n \in \mathbb{N}, n \leq N \right\}.$$

Für ungerades n gelten

$$\alpha\left(\frac{2}{\pi n}\right) = \left(\frac{2}{\pi n}, \pm \frac{2}{\pi n}\right) \quad \text{und} \quad \alpha\left(\frac{2}{\pi(n+1)}\right) = \left(\frac{2}{\pi(n+1)}, 0\right).$$

Also

$$\left\| \alpha\left(\frac{2}{\pi(n+1)}\right) - \alpha\left(\frac{2}{\pi n}\right) \right\|_2 \geq \frac{2}{\pi n}.$$

Das bedeutet, dass die Summe $L(\alpha, Z)$ nach unten durch ein positives Vielfaches eines Anfangsstücks der harmonischen Reihe abgeschätzt werden kann. Da man dieses Anfangsstücks beliebig lang wählen kann, ist α nicht rektifizierbar. $\square$

26.7 Bemerkung. Sei $Z = (t_0, t_1, \dots, t_n)$ eine Zerlegung von $[a, b]$. Eine *Verfeinerung* von Z ist eine Zerlegung $Z_1 = (s_0, \dots, s_m)$, so dass jedes t_j auch zu Z_1 gehört. In diesem Fall gilt $L(\alpha, Z) \leq L(\alpha, Z_1)$.

26.8 Definition. Sei $\alpha: [a, b] \rightarrow \mathbb{R}^n$ ein rektifizierbarer Weg. Die *Weglängenfunktion* ist definiert als

$$s(t) = L(\alpha|_{[a, t]}).$$

26.9 Bemerkung. Sei $Z = (t_0, t_1, \dots, t_m) \in \mathcal{Z}$. Ein Weg $\alpha: [t_0, t_m] \rightarrow \mathbb{R}^n$ ist genau dann rektifizierbar, wenn für alle $i \in \{1, \dots, m\}$ die Einschränkung $\alpha|_{[t_{i-1}, t_i]}$ rektifizierbar ist. In diesem Fall gilt

$$L(\alpha) = \sum_{i=1}^m L(\alpha|_{[t_{i-1}, t_i]}).$$

Insbesondere gilt für $a \leq t_1 < t \leq b$

$$s(t) = s(t_1) + L(\alpha|_{[t_1, t]}).$$

26.10 Satz. Ist α stetig differenzierbar, so ist α rektifizierbar, die Weglängenfunktion ist stetig differenzierbar mit $s'(t) = \|\alpha'(t)\|_2$ und

$$L(\alpha) = \int_a^b \|\alpha'(t)\|_2 dt.$$

Beweis. Sei $Z \in \mathcal{Z}$. Wir zeigen zuerst, dass $L(\alpha, Z) \leq \int_a^b \|\alpha'(t)\|_2 dt$. Nach einem Satz aus der Analysis I sind die Komponenten von α' gleichmäßig stetig auf dem kompakten Intervall $[a, b]$. Zu beliebig gewähltem $\epsilon > 0$ existiert daher ein $\delta > 0$, so dass für beliebige $s_1, s_2 \in [a, b]$ mit $|s_1 - s_2| < \delta$ bereits gilt $\|\alpha'(s_1) - \alpha'(s_2)\|_2 < \epsilon$. Sei $Z_1 = (t_0, t_1, \dots, t_k)$ eine Verfeinerung von Z mit $|t_i - t_{i-1}| < \delta$ für alle $i \in \{1, \dots, k\}$. Setze $M_i = \max\{\|\alpha'(t)\|_2 \mid t_{i-1} \leq t \leq t_i\}$ und sei $m_i = \min\{\|\alpha'(t)\|_2 \mid t_{i-1} \leq t \leq t_i\}$. Wegen der Wahl von δ gilt $0 \leq M_i - m_i < \epsilon$. Aus dem Mittelwertsatz folgt

$$\begin{aligned} L(\alpha, Z) &\leq L(\alpha, Z_1) = \sum_{i=1}^k \|\alpha(t_i) - \alpha(t_{i-1})\|_2 \leq \sum_{i=1}^k M_i |t_i - t_{i-1}| \\ &\leq \epsilon(b-a) + \sum_{i=1}^k m_i |t_i - t_{i-1}| \leq \epsilon(b-a) + \int_a^b \|\alpha'(t)\|_2 dt. \end{aligned} \quad (26.1)$$

Da dies für alle $\epsilon > 0$ gilt, ist die Zwischenbehauptung bewiesen. Insbesondere ist gezeigt, dass α rektifizierbar ist.

26. Die Weglänge

Wir zeigen nun die Differenzierbarkeit der Weglängenfunktion s . Sei dazu $h > 0$. Wir benutzen die Abschätzung (26.1) auf dem Intervall $[t, t+h]$

$$\left\| \frac{\alpha(t+h) - \alpha(t)}{h} \right\|_2 \leq \frac{1}{h}(s(t+h) - s(t)) \leq \frac{1}{h} \int_t^{t+h} \|\alpha'(\tau)\|_2 d\tau.$$

Der erste und der dritte Term konvergieren für $h \searrow 0$ gegen $\|\alpha'(t)\|_2$. Für $h < 0$ argumentiert man analog. Damit ist $s'(t) = \|\alpha'(t)\|_2$ gezeigt. Die letzte Aussage folgt daraus sofort. $\square$

26.11 Korollar. *Jeder stückweise stetig differenzierbare Weg ist rektifizierbar.*

26.12 Beispiel. In der Antike wurde π erklärt als halbe Länge des Einheitskreises. Die folgende Rechnung zeigt, dass diese Definition mit der modernen übereinstimmt.

Der Weg $\alpha: [-1, 1] \rightarrow \mathbb{R}^2$, $t \mapsto (t, \sqrt{1-t^2})$, beschreibt einen Halbkreis vom Radius 1. Dann

$$\alpha'(t) = \left(1, \frac{-t}{\sqrt{1-t^2}} \right).$$

Also

$$\|\alpha'(t)\|_2 = \sqrt{1 + \frac{t^2}{1-t^2}}$$

und daher

$$L(\alpha) = \int_{-1}^1 \|\alpha'(t)\|_2 dt = \int_{-1}^1 \frac{1}{\sqrt{1-t^2}} dt = \int_{-\pi/2}^{\pi/2} \frac{\cos(\varphi)}{\sqrt{1-\sin^2(\varphi)}} d\varphi = \pi.$$

26.13 Beispiel. Für $0 < \epsilon < 1$ sei der Weg $\alpha: [0, 1]$ definiert durch

$$\alpha(x) = \begin{cases} \begin{pmatrix} x \\ x^{1+\epsilon} \sin\left(\frac{1}{x}\right) \end{pmatrix}, & x > 0, \\ 0, & x = 0, \end{cases}$$

Dann ist α rektifizierbar, aber nicht stückweise stetig differenzierbar.

Beweis. Es gilt

$$\alpha'(t) = \begin{pmatrix} 1 \\ (1+\epsilon)t^\epsilon \sin\left(\frac{1}{t}\right) - t^{\epsilon-1} \cos\left(\frac{1}{t}\right) \end{pmatrix}$$

Setzt man wieder $t = \frac{1}{2\pi n}$ ein, so sieht man wie in Beispiel 26.4, dass α nicht stückweise stetig differenzierbar ist.

Ferner gilt

$$\begin{aligned} \|\alpha'(t)\|_2^2 &= 1 + (1+\epsilon)^2 t^{2\epsilon} \sin^2\left(\frac{1}{t}\right) - 2(1+\epsilon)t^{2\epsilon-1} \sin\left(\frac{1}{t}\right) \cos\left(\frac{1}{t}\right) + t^{2\epsilon-2} \cos^2\left(\frac{1}{t}\right) \\ &\leq 10t^{2\epsilon-2}. \end{aligned}$$

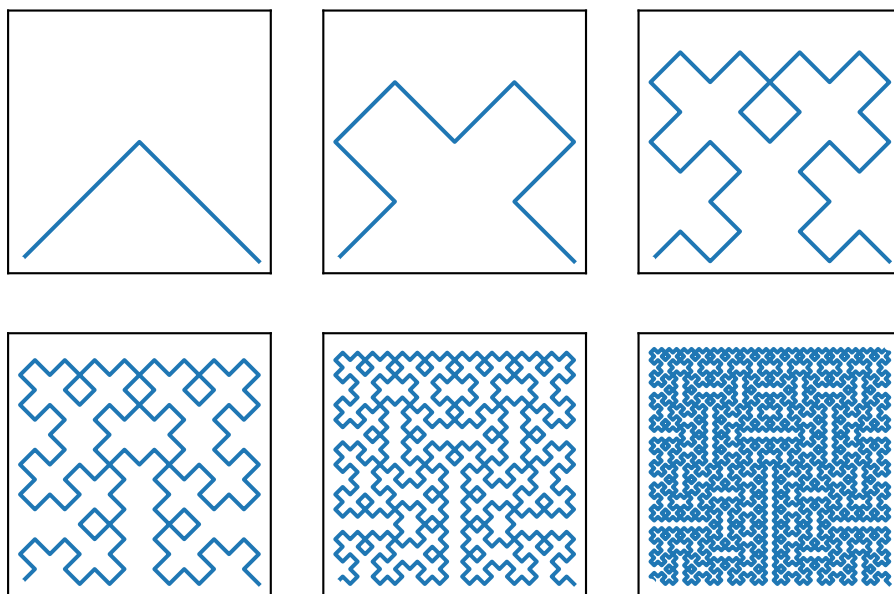


Abbildung 26.2.: Spuren von Approximationen an die Peano-Kurve aus Beispiel 26.14

Es sei $Z = (t_0, t_1, \dots, t_n)$ eine beliebige Zerlegung von $[0, 1]$. Dann ist die Länge der Einschränkung von α auf $[t_1, 1]$ gleich

$$\int_{t_1}^1 \|\alpha'(t)\|_2 \, dt \leq \sqrt{10} \int_{t_1}^1 t^{\epsilon-1} \, dt \leq \frac{\sqrt{10}}{\epsilon}. \quad \square$$

26.14 Beispiel (Peano-Kurve). Wir definieren rekursiv Funktionen $f_n = (g_n, h_n): [0, 1] \rightarrow Q := [0, 1]^2$ durch

$$g_0(t) := t, \quad h_0(t) := \frac{1}{2} - \left| t - \frac{1}{2} \right|,$$

$$g_{n+1}(t) := \begin{cases} \frac{1}{2}h_n(4t), & 0 \leq t < \frac{1}{4}, \\ \frac{1}{2}g_n(4t-1), & \frac{1}{4} \leq t < \frac{1}{2}, \\ \frac{1}{2} + \frac{1}{2}g_n(4t-2), & \frac{1}{2} \leq t < \frac{3}{4}, \\ 1 - \frac{1}{2}h_n(4t-3), & \frac{3}{4} \leq t \leq 1, \end{cases}, \quad h_{n+1}(t) := \begin{cases} \frac{1}{2}g_n(4t), & 0 \leq t < \frac{1}{4}, \\ \frac{1}{2} + \frac{1}{2}h_n(4t-1), & \frac{1}{4} \leq t < \frac{1}{2}, \\ \frac{1}{2} + \frac{1}{2}h_n(4t-2), & \frac{1}{2} \leq t < \frac{3}{4}, \\ \frac{1}{2} - \frac{1}{2}g_n(4t-3), & \frac{3}{4} \leq t \leq 1. \end{cases}$$

Die Spuren der ersten sechs Wege zeigt Abbildung 26.2.

Aus der Konstruktion sieht man, dass $\|f_{n+1}(t) - f_n(t)\| \leq 2^{1-n}$ für jedes t und jedes n gilt. Also konvergiert die Funktionenfolge $(f_n)_{n \rightarrow \infty}$ gleichmäßig gegen eine Funktion f . Wegen der Gleichmäßigkeit der Konvergenz ist f stetig.

Wir zeigen, dass f surjektiv ist. Sei dazu $x \in Q$ gegeben. Teilt man Q in 4^n Quadrate, so berührt f_n alle diese Quadrate. Daher existiert zu jedem n ein t_n mit

$$\|x - f_n(t_n)\|_\infty \leq 2^{-n}. \quad (26.2)$$

26. Die Weglänge

Die Folge $(t_n)_{n \in \mathbb{N}}$ besitzt eine konvergente Teilfolge $(t_{n_k})_{k \in \mathbb{N}}$. Ihr Grenzwert sei t_0 . Dann konvergiert $(f(t_{n_k}))_{k \in \mathbb{N}}$ einerseits gegen $f(t_0)$, andererseits aber auch gegen χ . Um letzteres zu sehen, sei $\epsilon > 0$ gegeben. Dann existiert $k_0 \in \mathbb{N}$, so dass $2^{-n_{k_0}} < \frac{\epsilon}{2}$ und $\|f_{n_k} - f\|_\infty < \frac{\epsilon}{2}$ für alle $k \geq k_0$. Für diese k gilt wegen (26.2)

$$\|\chi - f(t_0)\|_\infty \leq \|\chi - f(t_{n_k})\|_\infty + \|f(t_{n_k}) - f_{n_k}(t_{n_k})\|_\infty < 2^{-n_{k_0}} + \frac{\epsilon}{2} < \epsilon.$$

26.15 Satz. *Es sei $\alpha: [0, 1] \rightarrow Q$ surjektiv und stetig. Dann ist α nicht rektifizierbar.*

Beweis. Der Weg muss die Punkte $(\frac{j}{n}, \frac{k}{n})$, $j = 0, \dots, n$, und $k = 0, \dots, n$, miteinander verbinden. Wir wissen nicht, in welcher Reihenfolge das geschieht, aber jedenfalls beträgt die Entfernung von einem Punkt zum nächsten in der Supremumsnorm mindestens $\frac{1}{n}$. Da es mindestens $n^2 + 2n$ Segmente gibt, wird die Länge von α nach unten durch $n + 2$ abgeschätzt. $\square$