

Analysis 2

Manfred Dobrowolski*

9. April 2014

Inhaltsverzeichnis

1	Der euklidische Raum \mathbb{R}^n	1
1.1	Grundlegende Begriffe	1
1.2	Topologische Räume	2
1.3	Die Topologie des \mathbb{R}^n	3
1.4	Folgen im \mathbb{R}^n	4
1.5	Kompakte Teilmengen des \mathbb{R}^n	5
1.6	Die Geometrie des \mathbb{R}^n	6
2	Stetige und differenzierbare Funktionen	9
2.1	Funktionen und ihre Darstellung	9
2.2	Lineare und polynomiale Transformationen	9
2.3	Stetige Funktionen	11
2.4	Stetige Funktionen auf kompakten Mengen, gleichmäßige Stetigkeit	12
2.5	Richtungsableitung und partielle Ableitung	13
2.6	Differenzierbare Funktionen	14
2.7	Das Gradientenverfahren	15
2.8	Höhere Ableitungen	16
2.9	Vektorwertige Funktionen und Kettenregel	17
2.10	Mittelwertsatz und Satz von Taylor	18
2.11	Symmetrische Matrizen	19
2.12	Relative Extrema	21
3	Metrische und normierte Räume	26
3.1	Metrische Räume	26
3.2	Der Banachsche Fixpunktsatz	27
3.3	Der Banachsche Fixpunktsatz im \mathbb{R}^n	28
3.4	Banach Räume	28
3.5	Das Newton-Verfahren	30

*Institut für Mathematik, Universität Würzburg, Emil-Fischer-Straße 30, 97047 Würzburg

4	Der Satz über implizite Funktionen	32
4.1	Implizite Funktionen – Der Fall $n = m = 1$	32
4.2	Implizite Funktionen – Der allgemeine Fall	32
4.3	Der Umkehrsatz	35
4.4	Extremwertaufgaben mit Nebenbedingungen	36
4.5	Konvexität	39
5	Kurven	42
5.1	Kurven im \mathbb{R}^n	42
5.2	Kurvenlänge	44
5.3	Geometrie der Kurven	46
5.4	Kurvenintegrale	48
5.5	Wegintegrale	50
6	Gewöhnliche Differentialgleichungen	54
6.1	Gleichungen erster Ordnung	54
6.2	Lösungsmethoden für Gleichungen erster Ordnung	55
6.3	Systeme gewöhnlicher Differentialgleichungen	56
6.4	Existenz von Lösungen	57
6.5	Lineare Systeme	58
6.6	Dynamische Systeme	63
6.7	Stabile Punkte dynamischer Systeme	64
6.8	Ljapunovfunktionen	66
6.9	Die Exponentialfunktion für Matrizen	68
6.10	Der Stabilitätssatz	71

1 Der euklidische Raum \mathbb{R}^n

1.1 Grundlegende Begriffe Der \mathbb{R}^n besteht aus *Punkten* $x = (x_1, \dots, x_n)$, die gelegentlich auch als *Vektoren* bezeichnet werden. x_1, \dots, x_n heißen die *Koordinaten* des Punktes x . Für Vektoren x, y und reelle Zahlen α sind *Addition* und *Skalarmultiplikation* erklärt durch

$$x + y = (x_1 + y_1, \dots, x_n + y_n), \quad \alpha x = (\alpha x_1, \dots, \alpha x_n).$$

Die geometrische Bedeutung dieser Operationen ist die gleiche wie wir sie bei den ebenen Vektoren im Kapitel über komplexe Analysis kennengelernt haben. Der Vektor e_i , der an der i -ten Stelle eine Eins besitzt und ansonsten aus lauter Nullen besteht, heißt i -ter *Einheitsvektor*. Aus der Definition von Addition und Skalarmultiplikation folgt dann

$$x = \sum_{i=1}^n x_i e_i.$$

Das *Skalarprodukt* zweier Vektoren x und y ist definiert durch

$$x \cdot y = x_1 y_1 + \dots + x_n y_n = \sum_{i=1}^n x_i y_i$$

und wird manchmal auch mit xy oder (x, y) bezeichnet. Der *Betrag* des Vektors x ist dann

$$|x| = \left(\sum_{i=1}^n x_i^2 \right)^{1/2} = (x, x)^{1/2}$$

und entspricht der Länge der Strecke $\overline{0x}$ oder anders ausgedrückt, dem Abstand zwischen dem Nullpunkt und x . Der Abstand zwischen zwei Punkten x und y errechnet sich damit zu $|x - y|$.

Lemma Für das Skalarprodukt gilt die *Cauchy-Ungleichung*

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

Beweis: Wir beweisen die Ungleichung mit einem *Homogenitätsargument*, das in dieser Form sehr häufig vorkommt. Zunächst ist die Ungleichung richtig, wenn einer der beiden Vektoren verschwindet. Für $\tilde{x}, \tilde{y} \neq 0$ kann man die Cauchy-Ungleichung durch die Setzung $x = |\tilde{x}|^{-1} \tilde{x}$, $y = |\tilde{y}|^{-1} \tilde{y}$ auf den Fall $|x| = |y| = 1$ zurückführen und dadurch die Homogenität der Cauchy-Ungleichung ausnutzen. Für solche x, y erhalten wir aus der Youngschen Ungleichung

$$|(x, y)| = \left| \sum_{i=1}^n x_i y_i \right| \leq \sum_{i=1}^n |x_i| |y_i| \leq \frac{1}{2} \sum_{i=1}^n |x_i|^2 + \frac{1}{2} \sum_{i=1}^n |y_i|^2 = 1.$$

□

Sei X ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Eine Abbildung $\|\cdot\| : X \rightarrow \mathbb{R}$ heißt *Norm* auf X , wenn die folgenden Axiome erfüllt sind.

(N1) $\|x\| \geq 0$ und $x = 0 \Leftrightarrow \|x\| = 0$.

(N2) $\|\alpha x\| = |\alpha| \|x\| \quad \forall \alpha \in \mathbb{K}$ (Homogenität).

(N3) $\|x + y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung).

Die Normaxiome lassen sich für $\|\cdot\| = |\cdot|$ leicht nachweisen: Die beiden ersten Eigenschaften folgen direkt aus der Definition. Für die Dreiecksungleichung verwendet man die Cauchy-Ungleichung

$$\begin{aligned} |x + y|^2 &= (x + y) \cdot (x + y) = |x|^2 + 2x \cdot y + |y|^2 \\ &\leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2. \end{aligned}$$

Aus der Dreiecksungleichung erhält man mit Hilfe von

$$|x| = |x - y + y| \leq |x - y| + |y|$$

die *umgekehrte Dreiecksungleichung*

$$||x| - |y|| \leq |x - y|.$$

Ferner gilt die eigentliche Dreiecksungleichung

$$|x - z| \leq |x - y + y - z| \leq |x - y| + |y - z|,$$

die besagt, dass in einem Dreieck jede Seitenlänge kleiner oder gleich der Summe der beiden anderen Seitenlängen ist.

1.2 Topologische Räume Sei X eine beliebige Menge. Eine *Topologie* τ auf X ist ein System von Teilmengen von X , die *offene Mengen* genannt werden, mit:

(T1) \emptyset und X sind offen.

(T2) Die Vereinigung beliebig vieler offener Mengen ist offen.

(T3) Der Durchschnitt endlich vieler offener Mengen ist offen.

Das Paar (X, τ) heißt dann *topologischer Raum*. Ein topologischer Raum heißt *Hausdorff-Raum*, wenn zusätzlich das folgende *Trennungsaxiom* erfüllt ist:

(T4) Zu allen $x, y \in X$ mit $x \neq y$ gibt es offene Mengen A, B mit $x \in A$, $y \in B$ und $A \cap B = \emptyset$.

Für eine beliebige Menge ist die Potenzmenge von X eine Topologie auf X , die *diskrete Topologie* genannt wird und (X, τ) zu einem Hausdorff-Raum macht. Eine weitere Möglichkeit, eine beliebige Menge zu topologisieren, besteht in der Wahl $\tau = \{X, \emptyset\}$. Wenn X aus mehr als einem Element besteht, so ist für diese Topologie das Trennungsaxiom nicht erfüllt.

Sei (X, τ) ein topologischer Raum und $A \subset X$ eine beliebige Teilmenge. Dann ist A zusammen mit den Mengen $\{M \cap A : M \in \tau\}$ ein topologischer Raum. Diese Topologie heißt *Relativtopologie auf A* . Wenn A nicht selber offen in X ist, so sind die offenen Mengen der Relativtopologie nicht notwendig offen in X .

$A \subset X$ heißt *abgeschlossen*, wenn das Komplement $A^c = X \setminus A$ offen ist. Da die abgeschlossenen Mengen durch Komplementbildung definiert sind, gilt: Der beliebige Durchschnitt und die endliche Vereinigung abgeschlossener Mengen sind abgeschlossen.

Ferner definieren wir:

$\text{int } A$ = Vereinigung aller in A enthaltenen offenen Mengen,

\overline{A} = Durchschnitt aller abgeschlossenen Mengen, die A enthalten.

$\text{int } A$ heißt das *Innere* von A , \overline{A} der *Abschluß* von A . Nach Definition ist $\text{int } A$ offen und \overline{A} abgeschlossen.

In der diskreten Topologie sind alle Mengen A offen und abgeschlossen und es gilt $A = \text{int } A = \bar{A}$. In der Topologie $\tau = \{X, \emptyset\}$ sind nur die leere Menge und der ganze Raum offen und abgeschlossen.

Sei (X, τ) ein topologischer Raum. Eine Menge $U \subset X$ heißt *Umgebung* eines $x \in X$, wenn U offen ist mit $x \in U$. x heißt *innerer Punkt* einer Menge $A \subset X$, wenn eine Umgebung von x in A enthalten ist. $x \in X$ heißt *Berührungspunkt* von $A \subset X$, wenn in jeder Umgebung von x mindestens ein Punkt von A liegt. $x \in X$ heißt *Randpunkt* von $A \subset X$, wenn in jeder Umgebung von x mindestens ein Punkt von A und mindestens ein Punkt von A^c liegt. Mit ∂A wird die Menge der Randpunkte von A bezeichnet.

Mit diesen Definitionen lassen sich offene und abgeschlossene Mengen auch anders charakterisieren:

Lemma Sei (X, τ) ein topologischer Raum und $A \subset X$. Dann gilt:

- (a) A ist offen \Leftrightarrow Jedes $x \in A$ ist innerer Punkt von A ,
- (b) A ist abgeschlossen \Leftrightarrow Jeder Berührungspunkt von A gehört zu A ,
- (c) $\text{int } A =$ Menge der inneren Punkte von A ,
- (d) $\bar{A} =$ Menge der Berührungspunkte von A .

Beweis: (a): Die Richtung „ \Leftarrow “ gilt wegen $A = \cup_{x \in A} U(x)$, wobei $U(x)$ eine ganz in A liegende Umgebung von x ist. Wenn umgekehrt A offen ist, so ist A auch eine ganz in A enthaltene Umgebung eines jeden $x \in A$.

(b) folgt aus (a) durch Betrachtung des Komplements. (c) und (d) folgen direkt aus (a) bzw. (b).

□

Eine Folge (x_k) im topologischen Raum X *konvergiert* gegen $x \in X$, wenn in jeder Umgebung von x fast alle (also alle bis auf endlich viele) Folgenglieder liegen. Bei allgemeinen topologischen Räumen kann es zu Pathologien kommen, die unserer anschaulichen Vorstellung widersprechen. Ist beispielsweise $\tau = \{X, \emptyset\}$, so ist X die einzige Umgebung eines jeden $x \in X$. Damit ist in dieser Topologie jede Folge gegen jedes $x \in X$ konvergent. Ist dagegen X ein Hausdorff-Raum, so lassen sich zwei verschiedene Punkte durch offene Mengen trennen. In diesen offenen Mengen können aber nicht jeweils fast alle Folgenglieder liegen. Damit ist gezeigt, dass in einem Hausdorff-Raum der Grenzwert einer konvergenten Folge eindeutig bestimmt ist.

1.3 Die Topologie des \mathbb{R}^n Für $x \in \mathbb{R}^n$ und $r \geq 0$ ist die *offene beziehungsweise abgeschlossene Kugel* mit Mittelpunkt x und Radius r definiert durch

$$B_r(x) = \{y : |x - y| < r\}, \quad \tilde{B}_r(x) = \{y : |x - y| \leq r\}$$

Bei der offenen Kugel gehört der Rand $\{x : |x - y| = r\}$ nicht dazu, sie ist also zum Rande hin „offen“.

Sei $A \subset \mathbb{R}^n$. $x \in A$ heißt *innerer Punkt* von A , wenn es ein $r > 0$ gibt mit $B_r(x) \subset A$. A heißt *offen* oder *offene Menge*, wenn jeder Punkt von A innerer Punkt ist. A heißt *abgeschlossen*, wenn das *Komplement* A^c von A , $A^c = \mathbb{R}^n \setminus A$ offen ist. Ferner wird auch die leere Menge \emptyset als offen definiert.

Schauen wir uns diese Definitionen an einigen Beispielen genauer an. Da die leere Menge als offen definiert wurde, ist ihr Komplement, der \mathbb{R}^n , abgeschlossen. Andererseits gilt für jedes $x \in \mathbb{R}^n$, dass $B_1(x) \subset \mathbb{R}^n$. Damit ist der \mathbb{R}^n auch offen. Somit sind der \mathbb{R}^n und die leere Menge sowohl offen als auch abgeschlossen. Sind der \mathbb{R}^n und die leere Menge eher pathologische Fälle, deren Eigenschaften rein formal aus den Definitionen erschlossen werden, so besitzen die Begriffe „offen“ und „abgeschlossen“ ansonsten eine anschauliche Bedeutung: Bei einer abgeschlossenen Menge gehört der „Rand“ zur Menge dazu, bei einer offenen Menge dagegen nicht. Als Beispiel betrachten wir die r -Kugel $B_r(x)$ für $r > 0$. Für $y \in B_r(x)$ gilt $a = |x - y| < r$. Für $b = r - a$ ist dann $B_b(y) \subset B_r(x)$. Damit ist jeder Punkt von $B_r(x)$ innerer Punkt und die offene Kugel trägt

ihren Namen zu recht: Sie ist auch eine offene Menge im Sinne der Definition. Eine abgeschlossene Kugel ist dagegen nicht offen: In jeder offenen Kugel mit einem Randpunkt als Mittelpunkt befinden sich auch immer Punkte, die nicht zur abgeschlossenen Kugel gehören. Um zu entscheiden, ob die abgeschlossene Kugel $\tilde{B}_r(x)$ auch eine abgeschlossene Menge ist, betrachten wir das Komplement

$$\tilde{B}_r(x)^c = \{y : |x - y| > r\}.$$

Ist $z \in \tilde{B}_r(x)^c$, so gilt $|x - z| = a > r$. Die offene Kugel um z mit Radius $a - r$ liegt demnach ganz in $\tilde{B}_r(x)^c$. Damit ist $\tilde{B}_r(x)$ eine abgeschlossene Menge.

Das eindimensionale halboffene Intervall $I = (a, b] \subset \mathbb{R}$ ist ein Beispiel für eine Menge, die weder offen noch abgeschlossen ist.

Satz Der \mathbb{R}^n mit den oben definierten offenen Mengen ist ein Hausdorff-Raum.

Beweis: Dass die leere Menge und der \mathbb{R}^n offen sind, hatten wir bereits gesehen. Seien $A_i, i \in I$ offen. Ist $x \in \cup A_i$, so ist $x \in A_{i_0}$ für ein i_0 . Da A_{i_0} offen ist, gilt $B_r(x) \subset A_{i_0}$ für ein $r > 0$. Diese Kugel liegt auch in der Vereinigung der A_i , womit gezeigt ist, dass x innerer Punkt der Vereinigung ist. Sind A_1, \dots, A_k offene Mengen und $x \in \cap A_i$, so ist $B_{r_i}(x) \subset A_i$. Mit $r = \min\{r_1, \dots, r_k\}$ gilt dann $B_r(x) \subset \cap A_i$ und x ist innerer Punkt des Durchschnitts.

Kommen wir nun zum Nachweis von (T4). Seien $x, y \in \mathbb{R}^n$ mit $x \neq y$. Mit $r = |x - y|/2$ gilt $B_r(x) \cap B_r(y) = \emptyset$. \square

Das eindimensionale Beispiel $I_k = (-1/k, 1/k)$ mit $\cap I_k = \{0\}$ zeigt, dass beliebige Durchschnitte offener Mengen nicht offen zu sein brauchen.

Wir hatten jede offene Menge A mit $x \in A$ als Umgebung des Punktes x bezeichnet. Die Kugel $B_\varepsilon(x)$ wird auch ε -Umgebung genannt. Nach unserer Definition hieß ein Punkt *Randpunkt* einer Menge A , wenn sich in jeder Umgebung sowohl Punkte von A als auch Punkte von A^c befinden. Auf den \mathbb{R}^n angewandt stimmt diese Definition mit unserer anschaulichen Vorstellung des Randes überein. Abgesehen von der leeren Menge und dem \mathbb{R}^n besitzt jede Menge einen nichtleeren Rand.

Ist $A \subset \mathbb{R}^n$, so können wir die Menge A wie im vorigen Abschnitt topologisieren, indem wir $B \subset A$ als offen definieren, wenn es eine (im \mathbb{R}^n) offene Menge $B' \subset \mathbb{R}^n$ gibt mit $B = A \cap B'$. Auf diese Weise vererben sich die im \mathbb{R}^n gültigen Axiome auf diese *Relativtopologie*. Ist beispielsweise $\{B_i\}_{i \in I}$ ein System von offenen Mengen in dieser Relativtopologie, so gibt es offene Teilmengen B'_i des \mathbb{R}^n mit $B_i = A \cap B'_i$. Da $\cup B'_i$ offen im \mathbb{R}^n ist, gehört auch $A \cap (\cup B'_i)$ zur Relativtopologie.

1.4 Folgen im \mathbb{R}^n Bei Folgen $(a_k)_{k \in \mathbb{N}}$ im \mathbb{R}^n verwenden wir die Notation $a_k = (a_{k,1}, \dots, a_{k,n})$. Die Definitionen von Häufungspunkt und Grenzwert einer Folge können wir wörtlich aus dem eindimensionalen Fall übernehmen: $a \in \mathbb{R}^n$ ist *Häufungspunkt* der Folge, wenn in jeder Umgebung von a unendlich viele Folgenglieder liegen; a ist *Grenzwert* der Folge, wenn in jeder Umgebung von a fast alle Folgenglieder liegen. Damit konvergiert a_k gegen a , $a_k \rightarrow a$, genau dann, wenn es zu jedem $\varepsilon > 0$ ein $K \in \mathbb{N}$ gibt mit $|a_k - a| < \varepsilon$ für alle $k \geq K$. In die elementaren Ungleichungen

$$|x_i| \leq |x| \leq \sqrt{n} \max_i \{|x_i|\} \quad \forall x \in \mathbb{R}^n$$

können wir für x die Differenz $a - a_k$ einsetzen und erhalten unmittelbar: Die Folge (a_k) konvergiert genau dann gegen a , wenn alle zugehörigen eindimensionalen Komponentenfolgen $(a_{k,i})$ gegen a_i konvergieren. Denn die obige Ungleichung zeigt, dass, wenn $|a - a_k|$ klein wird, auch jede Komponente $|a_i - a_{k,i}|$ klein wird und umgekehrt. Damit kann mit den Methoden der eindimensionalen Analysis Folgen auf Konvergenz untersucht werden. Insbesondere folgt: Ist $a_k \rightarrow a$ und $b_k \rightarrow b$ sowie $\alpha, \beta \in \mathbb{R}$, so gilt

$$\alpha a_k + \beta b_k \rightarrow \alpha a + \beta b.$$

Eine Folge (a_k) heißt *beschränkt*, wenn es eine Konstante M gibt mit $|a_k| \leq M$ für alle k .

Jede streng monoton steigende Folge (k_l) mit $k_l \in \mathbb{N}$ erzeugt eine Teilfolge $(a_{k_l})_{l \in \mathbb{N}}$ der Folge (a_k) . Genau wie im eindimensionalen Fall gilt der

Satz [Bolzano-Weierstraß] Jede beschränkte Folge besitzt eine konvergente Teilfolge.

Beweis: Jede beschränkte Folge ist in einem Würfel $\times_{i=1}^n [-M, M]$ enthalten. Dieser Würfel wird in 2^n Teilwürfel der Kantenlänge M unterteilt. Von diesen Würfeln enthält einer unendlich viele Folgenglieder, aus dem wir ein Folgenglied auswählen. Durch fortgesetzte Teilung der Würfel und anschließende Wahl eines Folgengliedes wird die Teilfolge ganz analog zum eindimensionalen Fall konstruiert. \square

Das folgende Lemma ist zwar recht einfach gestrickt, sollte aber immer präsent sein:

Lemma Sei $A \subset \mathbb{R}^n$. $a \in \mathbb{R}^n$ ist genau dann Berührungspunkt von A , wenn es eine Folge (a_k) in A gibt mit $a_k \rightarrow a$.

Beweis: Nach der Definition des Berührungspunkts gibt es zu jeder Kugel $B_{1/k}(a)$ ein Element $a_k \in A \cap B_{1/k}(a)$. Offenbar gilt dann $a_k \rightarrow a$. Ist a kein Berührungspunkt von A , so gibt es eine ε -Umgebung von a , die keine Elemente von A besitzt. Demnach kann a auch nicht als Grenzwert einer Folge in A dargestellt werden. \square

Eine Folge (a_k) des \mathbb{R}^n heißt *Cauchy-Folge*, wenn es zu jedem $\varepsilon > 0$ ein $K \in \mathbb{N}$ gibt mit $|a_k - a_l| < \varepsilon$ für alle $k, l \geq K$. Genau wie bei der Konvergenz zeigt man, dass eine Folge (a_k) genau dann eine Cauchy-Folge ist, wenn die Folgen der Komponenten $(a_{k,i})_{k \in \mathbb{N}}$ für $i = 1, \dots, n$ Cauchy-Folgen des \mathbb{R}^1 sind. Insbesondere besitzt jede Cauchy-Folge einen Grenzwert.

1.5 Kompakte Teilmengen des \mathbb{R}^n Sei (X, τ) ein topologischer Raum. Eine Menge $A \subset X$ heißt *folgenkompakt*, wenn jede Folge in A eine in A konvergente Teilfolge besitzt. A heißt *(überdeckungs)kompakt*, wenn jede Überdeckung von A durch offene Mengen, offene Überdeckung genannt, eine endliche Teilüberdeckung besitzt.

Beispiel Sei $A = (0, 1) \subset \mathbb{R}$. Für

$$A_i = \left\{ x \in \mathbb{R} : x > \frac{1}{i} \right\}, \quad i \in \mathbb{N}$$

gilt $\cup A_i = (0, \infty) \supset (0, 1)$, aber endlich viele der A_i reichen nicht aus, um $(0, 1)$ zu überdecken. Kurz: Die A_i enthalten keine endliche Teilüberdeckung und das Intervall $(0, 1)$ ist nicht kompakt.

Satz Sei $A \subset \mathbb{R}^n$. Die folgenden Aussagen sind äquivalent:

- (a) A ist folgenkompakt.
- (b) A ist kompakt.
- (c) A ist abgeschlossen und beschränkt.

Bemerkung Für später: (a) und (b) sind auch im vollständigen metrischen Raum äquivalent, aber nicht in allgemeinen topologischen Räumen. Die Äquivalenz von (a) und (b) zu (c) beruht auf der Tatsache, dass der \mathbb{R}^n ein endlich dimensionaler Raum ist.

Beweis: „(a) \Rightarrow (c)“: Sei A folgenkompakt. Dann ist A auch beschränkt, weil man ansonsten eine Folge (a_k) in A mit $|a_k| \rightarrow \infty$ wählen kann. Nach Lemma 1.4 bestehen die Grenzwerte von Folgen in A aus den Berührungspunkten von A . Daher gehören alle Berührungspunkte von A zu A , demnach ist A abgeschlossen.

„(c) \Rightarrow (a)“: Ist A beschränkt, so besitzt nach dem Satz von Bolzano-Weierstraß jede Folge in A eine konvergente Teilfolge. Die Abgeschlossenheit garantiert nach Lemma 1.4, dass der Grenzwert zu A gehört.

„(b) \Rightarrow (c)“: Sei A kompakt. Wäre A unbeschränkt, so könnte man zur offenen Überdeckung $\{B_k(0)\}_{k \in \mathbb{N}}$ keine endliche Teilüberdeckung finden, weil eine solche immer nur beschränkte Mengen überdecken kann. Angenommen, A ist nicht abgeschlossen. Dann gibt es einen Berührungspunkt a

von A , der nicht zu A gehört. Die offene Überdeckung bestehend aus den Mengen $U_k = \{x \in \mathbb{R}^n : |x-a| > 1/k\}$, $k \in \mathbb{N}$, besitzt wieder keine endliche Teilüberdeckung, weil bei dieser eine Umgebung von a nicht mit überdeckt wird.

„(c) \Rightarrow (b)“: Angenommen, $\{A_i\}_{i \in I}$ wäre eine offene Überdeckung der beschränkten und abgeschlossenen Menge A , die keine endliche Teilüberdeckung enthält. A läßt sich durch abgeschlossene Kugeln $\tilde{B}^1, \dots, \tilde{B}^i$ vom Radius 1 überdecken. Unter diesen muß es eine Kugel $\tilde{B}_1(x_1)$ geben, so daß die Menge $U_1 = \tilde{B}_1(x_1) \cap A$ sich ebenfalls nicht durch endlich viele dieser A_i überdecken läßt. U_1 wird durch endlich viele Kugeln vom Radius $1/2$ überdeckt und wieder wird eine Kugel $\tilde{B}_{1/2}(x_2)$ gefunden, so dass sich $U_2 = \tilde{B}_{1/2}(x_2) \cap A$ nicht durch endlich viele A_i überdecken läßt. Dieses Argument wird mit $1/k$ iteriert. Wir erhalten eine Folge von abgeschlossenen Mengen U_k , die sich nicht durch endlich viele A_i überdecken lassen. Nach Konstruktion gilt $U_{k+1} \subset U_k \subset A$. Für beliebige Punkte $y_k \in U_k$ ist (y_k) eine Cauchy-Folge, die aufgrund der Abgeschlossenheit von A einen Grenzwert $y \in A$ besitzt. Für ein i ist $y \in A_i$ und mit A_i offen ist $B_r(y) \subset A_i$ mit $r > 0$. Daher $U_k \subset A_i$ für genügend großes k . Widerspruch! \square

$x \in \mathbb{R}^n$ heißt *Häufungspunkt* der Menge A , wenn in jeder Umgebung von x unendlich viele Punkte von A liegen. Wählt man in dieser Definition als Umgebungen speziell $B_{1/k}(x)$ und nimmt für jedes $k \in \mathbb{N}$ ein $a_k \in B_{1/k} \cap A \setminus \{x\}$, so erhält man: x ist genau dann Häufungspunkt der Menge A , wenn es eine Folge (a_k) in A gibt mit $a_k \neq x$ und $a_k \rightarrow x$.

Man unterscheide Häufungspunkte von Folgen und Mengen. Beispielsweise hat eine konstante Folge (a_k) einen Häufungspunkt, die zugehörige Menge $\{a_k\}$ dagegen nicht.

1.6 Die Geometrie des \mathbb{R}^n Eine offene und zusammenhängende Menge heißt *Gebiet*. Dabei heißt eine offene Menge genau dann zusammenhängend, wenn sie sich nicht als Vereinigung zweier nichtleerer disjunkter offener Mengen darstellen lässt. Beispiele für Gebiete sind offene Intervalle des \mathbb{R}^1 sowie offene Kugeln. Dagegen bilden zwei sich berührende offene Kugeln wie etwa $B_1(0) \cup B_1(2e_1)$ kein Gebiet.

Ist $x_0 \in \mathbb{R}^n$ und $a \in \mathbb{R}^n \setminus \{0\}$, so ist die Gerade durch x_0 entlang a gegeben durch

$$\{y = x_0 + \lambda a : \lambda \in \mathbb{R}\}.$$

Die Strecke \overline{xy} zwischen den Punkten x und y lässt sich analytisch beschreiben durch

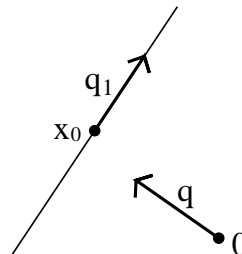
$$\{z = \lambda x + (1 - \lambda)y : \lambda \in [0, 1]\}.$$

Sind x_0, x_1, \dots, x_k Punkte des \mathbb{R}^n , so ist der zugehörige *Polygonzug* definiert durch

$$P(x_0, \dots, x_k) = \overline{x_0, x_1} \cup \dots \cup \overline{x_{k-1}, x_k}.$$

Eine *Hyperebene* des \mathbb{R}^n ist eine $(n-1)$ -dimensionale Ebene, also eine Menge der Form $H = x_0 + U$ mit einem $(n-1)$ -dimensionalen Teilraum U . U wird durch $n-1$ linear unabhängige Vektoren q_1, \dots, q_{n-1} aufgespannt,

$$U = \left\{ x = \sum_{i=1}^{n-1} \lambda_i q_i, \lambda_i \in \mathbb{R} \text{ für } i = 1, \dots, n-1 \right\}.$$



Sei $q \in \mathbb{R}^n$ mit $|q| = 1$, der auf allen q_i senkrecht steht. Ein Punkt $x = \sum_{i=1}^n \alpha_i q_i$ gehört genau dann zu U , wenn $\alpha = (x, q) = 0$ ist. Er gehört genau dann zu H , wenn $x - x_0 \in U$, wenn also $(x, q) = (x_0, q) = \alpha$ erfüllt ist. Diese Darstellung der Hyperebene heißt *Hessesche Normalenform*,

$$H = \{x : (x, q) = \alpha\}, \quad |q| = 1,$$

wobei der Vektor q senkrecht zur Ebene H steht und $\alpha = (x_0, q)$ ist. Es ist anschaulich klar, dass das Lot des Nullpunktes auf die Ebene die Gestalt βq besitzt. Aus der Gleichung für die Hyperebene erhalten wir $\beta = \alpha$, womit gezeigt ist, dass $|\alpha|$ den Abstand der Ebene zum Nullpunkt angibt.

Hat man umgekehrt eine Ebene in der Darstellung $(x, \tilde{q}) = \tilde{\alpha}$ mit $\tilde{q} \neq 0$ gegeben, so bringt man sie zuerst auf die Form $(x, q) = \alpha$ mit $|q| = 1$. Anschließend ergänzt man q mit q_1, \dots, q_{n-1} zu einer Orthogonalbasis des \mathbb{R}^n und setzt $U = \text{span}\{q_1, \dots, q_{n-1}\}$. Mit $x_0 = \alpha q$ ist dann $H = x_0 + U$ die Hyperebene in expliziter Darstellung.

Beispiel Sei $n = 3$. Wir bestimmen die Ebene H durch die Punkte e_1, e_2, e_3 in expliziter und in Hessescher Normalenform. Die Vektoren $e_2 - e_1, e_3 - e_1$ sind linear unabhängig, also

$$H = \{x = e_1 + \lambda(e_2 - e_1) + \mu(e_3 - e_1) : \lambda, \mu \in \mathbb{R}\}.$$

Zur Bestimmung der Hesseschen Normalenform müssen wir ein $q \in \mathbb{R}^3$ finden mit

$$q \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = 0, \quad q \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = 0, \quad |q| = 1$$

also $q = (1, 1, 1)/\sqrt{3}$. Daher

$$H = \left\{x : \frac{x_1 + x_2 + x_3}{\sqrt{3}} = \frac{1}{\sqrt{3}}\right\}.$$

Aufgaben

1.1 Zeigen Sie, dass

$$\sum_{i=1}^n |x_i| \leq \sqrt{n}|x| \quad \forall x \in \mathbb{R}^n.$$

Hinweis: Verwenden Sie die Cauchy-Ungleichung.

1.2 Zeigen Sie für $x, y \in \mathbb{R}^n$ die *Parallelogrammgleichung*

$$2|x|^2 + 2|y|^2 = |x + y|^2 + |x - y|^2.$$

Interpretieren Sie die Gleichung für $n = 2$ geometrisch.

1.3 Beweisen Sie: Ist a ein Randpunkt der Menge $A \subset \mathbb{R}^n$, so gibt es eine Folge (a_k) , $a_k \in A$, mit $a_k \rightarrow a$.

1.4 Welche der folgenden ebenen Mengen sind offen bzw. abgeschlossen?

a) $\bigcup_{n=1}^{\infty} \left(\frac{1}{n+1}, \frac{1}{n}\right) \times (0, n)$

b) $\left\{x = \left(\frac{1}{m}, \frac{1}{n}\right) : m, n \in \mathbb{Z} \setminus \{0\}\right\}$

c) $\bigcup_{n=1}^{\infty} B_{1/n}\left(\left(\frac{1}{n}, n\right)\right)$

1.5 Für Teilmengen $A, B \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ erklären wir Summe, Differenz und Skalarmultiplikation durch

$$A + B = \{x = y + z : y \in A, z \in B\}$$

$$A - B = \{x = y - z : y \in A, z \in B\}$$

$$\lambda A = \{x = \lambda y : y \in A\}.$$

Beweisen Sie oder widerlegen Sie durch Gegenbeispiel:

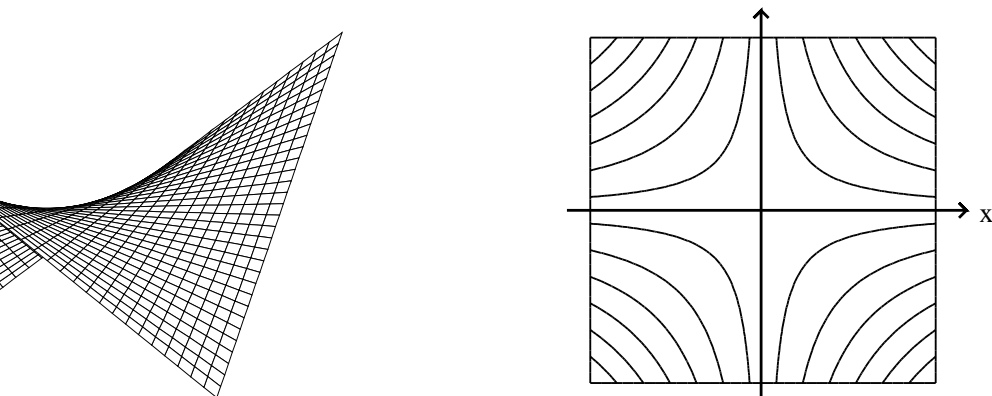
- a) Ist A offen, B beliebig, so ist $A + B$ offen.
- b) Sind A, B kompakt, so ist $A + B$ kompakt.
- c) Sind A, B abgeschlossen, so ist $A + B$ abgeschlossen.
- d) $2A = A + A$.
- e) Besitzt A einen inneren Punkt, so gibt es ein $r > 0$ mit $B_r(0) \subset A - A$.

1.6 Sei $n = 3$. Bestimmen Sie die Hessesche Normalenform der Ebene durch die drei Punkte e_1, e_2 und $e_3 - 3e_1$. Bestimmen Sie den Abstand der Ebene zum Nullpunkt.

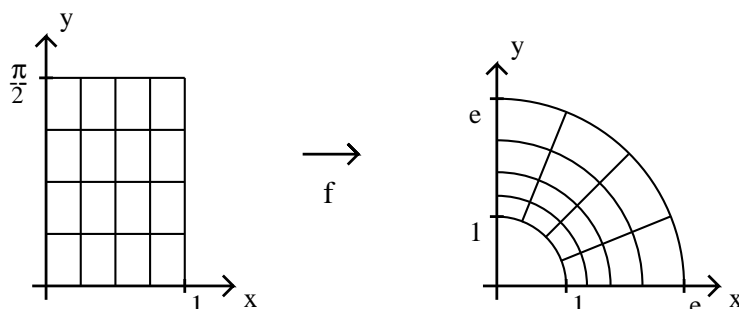
2 Stetige und differenzierbare Funktionen

2.1 Funktionen und ihre Darstellung Sei $D \subset \mathbb{R}^n$. Eine Abbildung $f : D \rightarrow \mathbb{R}^m$ heißt auch *Funktion mehrerer Veränderlicher*. Ist $m > 1$, so betonen wir häufig den geometrischen Aspekt einer solchen Funktion und nennen sie *Transformation*. Für $m = 1$ wird die Funktion auch als *reellwertig*, für $m > 1$ auch als *vektorwertig* bezeichnet.

Im Fall $n = 2$ bezeichnen wir die einzelnen Punkte auch mit (x, y) statt (x_1, x_2) . In diesem Fall lassen sich die reellwertigen Funktionen visualisieren, indem der zugehörige Graph $G_f = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D \text{ und } z = f(x, y)\}$ als Fläche abgebildet wird oder indem die Höhenlinien $f(x, y) = c$ eingezeichnet werden. In dieser Form begegnen uns Funktionen zweier Veränderlicher durch die Höhenlinien auf Wanderkarten.



Auf dem linken Bild sehen wir den Graphen der Funktion $f(x, y) = xy$, auf dem rechten die Höhenlinien dieser Funktion im Bereich $[-1, 1] \times [-1, 1]$.



Transformationen $f : D \rightarrow \mathbb{R}^2$ lassen sich visualisieren, indem man den Definitionsbereich D mit einem Gitternetz überzieht und das Bild des Netzes unter der Abbildung zeichnet. Das obige Beispiel zeigt die Transformation $f(x, y) = (\exp x \cos y, \exp x \sin y)$, was der komplexen Exponentialfunktion entspricht.

2.2 Lineare und polynomiale Transformationen Die Funktion $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt *linear*, wenn für alle $x, y \in \mathbb{R}^n$ und alle $\alpha, \beta \in \mathbb{R}$ gilt

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y).$$

Mit $x = \sum_{j=1}^n x_j e_j$ erhalten wir

$$(2.1) \quad f(x) = \sum_{j=1}^n x_j f(e_j).$$

Eine lineare Transformation ist daher schon durch ihre Werte auf den kanonischen Einheitsvektoren $\{e_j\}_{j=1, \dots, n}$ eindeutig bestimmt. Wir fassen die $f(e_j)$ als Spaltenvektoren auf und stellen sie in einem qua-

dratischen Schema, das *Matrix* genannt wird, zusammen,

$$A = (f(e_1) | \dots | f(e_n)) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

mit $a_{ij} = f(e_j)_i$, $i = 1, \dots, m$, $j = 1, \dots, n$. Wir definieren das Produkt einer Matrix mit einem Spaltenvektor durch

$$Ax = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix},$$

was man sich am besten mit dem Stichwort „Zeile mal Spalte“ merken kann. Aufgrund von (2.1) gilt dann für die lineare Abbildung f und ihrer zugehörigen Matrix A , dass $f(x) = Ax$. A kann daher als lineare Abbildung des \mathbb{R}^n in den \mathbb{R}^m aufgefasst werden, die wir in der Notation nicht von der Matrix A unterscheiden. Die Menge der $(m \times n)$ -Matrizen bezeichnen wir mit $\mathbb{R}^{m \times n}$.

In der mehrdimensionalen Analysis schreiben wir aus typographischen Gründen die Koordinaten der Punkte meist zeilenförmig, im Zusammenhang mit Matrizen wird ein Punkt jedoch immer als Spaltenvektor aufgefasst, für den das Matrix-Vektor-Produkt definiert ist.

Sind $g: \mathbb{R}^m \rightarrow \mathbb{R}^p$ und $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ lineare Transformationen mit Matrixdarstellungen $A \in \mathbb{R}^{p \times m}$ und $B \in \mathbb{R}^{m \times n}$, so ist die Komposition $h(x) = g(f(x))$ wiederum linear und es gilt

$$\begin{aligned} h(x) = g(f(x)) &= \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{p1} & \dots & a_{pm} \end{pmatrix} \begin{pmatrix} b_{11}x_1 + \dots + b_{1n}x_n \\ \vdots \\ b_{m1}x_1 + \dots + b_{mn}x_n \end{pmatrix} \\ &= \begin{pmatrix} \sum_l a_{1l}(b_{l1}x_1 + \dots + b_{ln}x_n) \\ \vdots \\ \sum_l a_{pl}(b_{l1}x_1 + \dots + b_{ln}x_n) \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{p1} & \dots & c_{pn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

mit

$$C = (c_{ij})_{i=1, \dots, p, j=1, \dots, n}, \quad c_{ij} = \sum_{l=1}^m a_{il}b_{lj}.$$

Demnach wird die Komposition der Transformationen g und f dargestellt durch die Matrix $C \in \mathbb{R}^{p \times n}$. Wir nennen C das *Matrizenprodukt* von A und B und schreiben $C = AB$. Auch hier gilt die Regel „Zeile mal Spalte“. Da die Komposition von Abbildungen assoziativ ist, gilt auch für das Matrizenprodukt $(AB)C = A(BC)$.

Beispiel Seien $g, f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ linear mit $g(e_1) = e_1 + e_2$, $g(e_2) = 2e_1$, $f(e_1) = 2e_2$, $f(e_2) = e_1$. Dann sind die Matrixdarstellungen

$$g(x) = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} x, \quad f(x) = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} x,$$

und

$$g(f(x)) = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} x = \begin{pmatrix} 4 & 1 \\ 0 & 1 \end{pmatrix} x.$$

Die *Einheitsmatrix* $E = E_n = (e_{ij})_{i,j=1, \dots, n}$ ist definiert durch $e_{ij} = \delta_{ij}$. Für jede Matrix $A \in \mathbb{R}^{n \times n}$ gilt $AE = EA = A$.

Die $(m \times n)$ -Matrizen bilden mit der komponentenweisen Addition und der durch $\lambda A = (\lambda a_{ij})$ definierten Skalarmultiplikation einen linearen Vektorraum der Dimension mn . Wir werden zeigen, dass man diesen Raum normieren kann mit

$$|A| = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2} \quad (= \text{euklidische Norm})$$

oder alternativ mit

$$\|A\|_2 = \sup_{x \neq 0} \frac{|Ax|}{|x|} \quad (= \text{induzierte Norm}).$$

Für die euklidische Matrixnorm folgen die Normaxiome unmittelbar, da sie mit dem Betrag für Vektoren des \mathbb{R}^{mn} übereinstimmt. Auch für die induzierte Matrixnorm lassen sich die Normaxiome leicht nachweisen, die Dreiecksungleichung folgt aus

$$\begin{aligned} \|A + B\|_2 &= \sup_{x \neq 0} \frac{|(A + B)x|}{|x|} \leq \sup_{x \neq 0} \frac{|Ax| + |Bx|}{|x|} \\ &\leq \sup_{x \neq 0} \frac{|Ax|}{|x|} + \sup_{x \neq 0} \frac{|Bx|}{|x|} = \|A\|_2 + \|B\|_2. \end{aligned}$$

Eine Matrixnorm $\|\cdot\|$ heißt *verträglich* mit der euklidischen Vektornorm $|\cdot|$, wenn

$$|Ax| \leq \|A\| |x|$$

für alle $x \in \mathbb{R}^n$ erfüllt ist. Für die induzierte Norm ist das richtig, denn $\|A\|_2$ ist die kleinste Zahl, für die die obige Abschätzung richtig ist. Aber auch die euklidische Matrixnorm ist verträglich mit der euklidischen Vektornorm, weil mit der Cauchy-Ungleichung gilt

$$\left| \sum_j a_{ij} x_j \right|^2 \leq \sum_j |a_{ij}|^2 |x|^2 \Rightarrow |Ax|^2 \leq \sum_{i,j} |a_{ij}|^2 |x|^2 = |A|^2 |x|^2.$$

Mit der Matrixmultiplikation ist im Gegensatz zum \mathbb{R}^n noch eine weitere Operation auf dem Matrizenraum definiert. Wir nennen eine Matrixnorm $\|\cdot\|$ *submultiplikativ*, wenn

$$\|AB\| \leq \|A\| \|B\| \quad \text{für alle } A \in \mathbb{R}^{p \times m} \text{ und } B \in \mathbb{R}^{m \times n}.$$

Beide oben eingeführten Matrixnormen genügen dieser Bedingung. Für die euklidische Norm erhalten wir mit Hilfe der Cauchy-Ungleichung ähnlich wie zuvor

$$\begin{aligned} |AB|^2 &\leq \sum_{i,j} \left| \sum_l a_{il} b_{lj} \right|^2 \leq \sum_{i,j} \left(\sum_l |a_{il}|^2 \sum_l |b_{lj}|^2 \right) \\ &= \sum_i \sum_l |a_{il}|^2 \sum_j \sum_l |b_{lj}|^2 = |A|^2 |B|^2. \end{aligned}$$

Für die induzierte Norm folgt aus der Verträglichkeit mit der euklidischen Vektornorm

$$|ABx| \leq \|A\|_2 |Bx| \leq \|A\|_2 \|B\|_2 |x|.$$

Nach Division durch $|x|$ folgt die behauptete Ungleichung.

Einen Vektor $\alpha \in \mathbb{N}_0^n$ nennen wir *Multiindex*. Wir setzen $|\alpha| = \alpha_1 + \dots + \alpha_n$. Mit Hilfe der Multiindezes lassen sich Polynome des \mathbb{R}^n sehr einfach notieren. Wir schreiben

$$x^\alpha = x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \dots \cdot x_n^{\alpha_n}$$

für ein *Monom*. Das allgemeine *Polynom vom Grad $\leq k$* besitzt dann die Gestalt

$$p(x) = \sum_{|\alpha| \leq k} a_\alpha x^\alpha, \quad a_\alpha \in \mathbb{R}.$$

2.3 Stetige Funktionen Sei $f : D \rightarrow \mathbb{R}^m$ und $\xi \in \mathbb{R}^n$ sei ein Häufungspunkt von $D \subset \mathbb{R}^n$. Wir sagen, f *konvergiert gegen* $a \in \mathbb{R}^m$ für $x \rightarrow \xi$,

$$\lim_{x \rightarrow \xi} f(x) = a \quad \text{oder} \quad f(x) \rightarrow a \text{ für } x \rightarrow \xi,$$

wenn für alle Folgen (x_k) mit $x_k \rightarrow \xi$ und $x_k \neq \xi$ gilt $f(x_k) \rightarrow a$. Die Funktion f heißt *stetig* in $\xi \in D$, wenn für jede Folge (x_k) mit $x_k \rightarrow \xi$ gilt $f(x_k) \rightarrow f(\xi)$. f heißt *stetig in* D , wenn f in jedem Punkt von D stetig ist.

Da eine Folge von Vektoren genau dann konvergent ist, wenn jede Komponente konvergiert, gilt: f ist genau dann stetig, wenn jede Komponente f_i von f stetig ist.

Diese Begriffe unterscheiden sich nicht von denen, die wir bei der Analysis einer Veränderlichen kennengelernt haben. Mit Beweis wie im eindimensionalen Fall zeigt man: Die Funktion f ist genau dann stetig in $\xi \in D$, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x \in D$ mit $|x - \xi| < \delta$ folgt

$$|f(x) - f(\xi)| < \varepsilon.$$

Beispiel Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) = \begin{cases} 0 & \text{für } (x, y) = 0 \\ \frac{xy}{x^2 + y^2} & \text{sonst} \end{cases}.$$

f ist in beiden Koordinatenrichtungen (als eindimensionale Funktion) im Nullpunkt stetig. Aber für $x \neq 0$ gilt $f(\pm x, x) = 1/2$, womit f im Nullpunkt unstetig ist.

Wir können die Stetigkeit von f auch topologisch charakterisieren:

Satz Sei $D \subset \mathbb{R}^n$ und $f : D \rightarrow \mathbb{R}^m$. Dann sind äquivalent:

- (a) Die Funktion f ist in D stetig.
- (b) Die Urbilder offener Mengen sind offen (bezüglich der Relativtopologie von D).

Beweis: f sei stetig in D . Sei $G \subset \mathbb{R}^m$ offen. Ist $f^{-1}(G) = \emptyset$, so ist $f^{-1}(G)$ jedenfalls offen. Ist andernfalls $\xi \in f^{-1}(G)$, so wird eine δ -Umgebung von ξ auf eine ε -Umgebung von $f(\xi) \in G$ abgebildet. Da G offen ist, kann das ε so klein gewählt werden, dass

$$f(B_\delta(\xi)) \subset B_\varepsilon(f(\xi)) \subset G.$$

Damit ist ξ innerer Punkt von $f^{-1}(G)$ in der Relativtopologie von D .

f erfülle die Bedingung (b). Sei $\xi \in D$. Das Urbild von $B_\varepsilon(f(\xi))$ ist relativ offen, ξ also innerer Punkt von D bezüglich der Relativtopologie,

$$f(B_\delta(\xi) \cap D) \subset B_\varepsilon(f(\xi)).$$

Damit ist f in ξ stetig. \square

Korollar f stetig \Leftrightarrow Die Urbilder abgeschlossener Mengen sind abgeschlossen.

Beweis: $f^{-1}(G)$ ist für offenes G relativ offen. Damit ist $f^{-1}(G^c)$ relativ abgeschlossen. \square

Wie im eindimensionalen Fall beweist man: Summe, Produkt und Komposition stetiger Funktionen sind stetig. Sind f, g stetig, g rellwertig und nichtverschwindend, so ist auch der Quotient f/g stetig.

Beispiele (i) Das Skalarprodukt $(\cdot, \cdot) : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ ist stetig, weil es sich aus Summen und Produkten zusammensetzt. Auch der Betrag $|\cdot| : \mathbb{R}^n \rightarrow \mathbb{R}$ ist stetig, weil er aus der Komposition des Skalarprodukt und der Wurzelfunktion besteht.

(ii) Jede lineare Funktion $f(x) = Ax$ ist stetig, weil sie sich zusammensetzt aus Summen und Produkten.

(iii) Jedes Polynom ist stetig, weil ein Monom aus Produkten besteht und ein Polynom die Summe von Monomen ist.

2.4 Stetige Funktionen auf kompakten Mengen, gleichmäßige Stetigkeit Den folgenden Sätzen sind wir im Prinzip schon begegnet, mit Hilfe des Kompaktheitsbegriffs können sie eleganter formuliert werden.

Satz Das stetige Bild einer kompakten Menge ist kompakt. Insbesondere nimmt eine stetige rellwertige Funktion auf einer kompakten Menge Maximum und Minimum an.

Beweis: Sei D kompakt und $f : D \rightarrow \mathbb{R}^m$ stetig. Sei (y_k) eine Folge in $f(D)$ mit zugehöriger Urbildfolge (x_k) . Da D kompakt ist, gibt es eine Teilfolge (x_{k_l}) mit $x_{k_l} \rightarrow x$. Da f stetig ist, gilt $y_{k_l} = f(x_{k_l}) \rightarrow f(x) = y \in f(D)$. \square

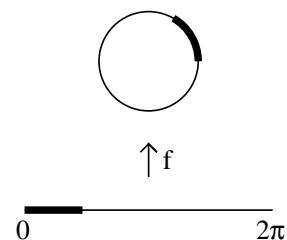
Satz Sei $D \subset \mathbb{R}^n$ kompakt und $f : D \rightarrow \mathbb{R}^m$ stetig und injektiv. Dann ist auch die Umkehrfunktion $f^{-1} : f(D) \rightarrow D$ stetig.

Beweis: Sei $y_k \rightarrow y$ eine konvergente Folge in $f(D)$ mit zugehöriger Urbildfolge $x_k = f^{-1}(y_k)$. Da D kompakt ist, gibt es eine Teilfolge mit $x_{k_l} \rightarrow x$. Wegen der Stetigkeit von f folgt $y_{k_l} = f(x_{k_l}) \rightarrow f(x) = y$ und damit $x = f^{-1}(y)$. Analog gilt: Jede Teilfolge von (x_k) besitzt eine Teilfolge, die gegen x mit $f(x) = y$ konvergiert. Damit konvergiert die ganze Folge (x_k) gegen x , was die Stetigkeit von f^{-1} impliziert. \square

Beispiel Sei $D = [0, 2\pi)$. Die Abbildung $f = (f_1, f_2)^T$ mit

$$f_1(x) = \cos x, \quad f_2(x) = \sin x,$$

bildet D stetig und bijektiv auf den Einheitskreis ab. Die Inverse ist aber unstetig im Punkt $(1, 0)$, denn die relativ offene Menge $[0, a)$ wird nicht auf eine offene Menge (in der Relativtopologie des Einheitskreises) abgebildet. Die Voraussetzung, dass D kompakt ist, ist im letzten Satz ganz wesentlich.



$f : D \rightarrow \mathbb{R}^m$ heißt gleichmäßig stetig, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt mit

$$\forall x, y \in D \quad |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Mit gleichem Beweis wie im Eindimensionalen zeigt man: Eine auf einer kompakten Menge D stetige Funktion ist dort gleichmäßig stetig.

Auch bezüglich der Konvergenz von Funktionenfolgen haben wir die gleichen Begriffsbildungen wie im eindimensionalen Fall. $f_k \rightarrow f$ gilt *punktweise*, wenn $f_k(x) \rightarrow f(x)$ für alle $x \in D$. $f_k \rightarrow f$ gilt *gleichmäßig*, wenn zu jedem $\varepsilon > 0$ fast alle Folgenglieder der Bedingung $|f_k(x) - f(x)| < \varepsilon$ für alle $x \in D$ genügen. Wie im eindimensionalen Fall beweist man: Der gleichmäßige Limes einer Folge stetiger Funktionen ist stetig.

2.5 Richtungsableitung und partielle Ableitung

In diesem und den folgenden Abschnitten ist $D \subset \mathbb{R}^n$ immer ein Gebiet.

Einen Vektor $e \in \mathbb{R}^n$ mit $|e| = 1$ bezeichnen wir als *Richtung*. Sei $f : D \rightarrow \mathbb{R}$ und $\xi \in D$. f heißt in ξ in *Richtung* e *differenzierbar*, wenn der Grenzwert

$$\frac{\partial}{\partial e} f(\xi) = \frac{\partial f}{\partial e}(\xi) = \lim_{t \rightarrow 0} \frac{f(\xi + te) - f(\xi)}{t}$$

existiert. Dies ist gleichbedeutend mit der Differenzierbarkeit der eindimensionalen Funktion

$$(2.2) \quad \phi(t) = f(\xi + te)$$

im Punkt $t = 0$. Anschaulich bedeutet die Richtungsableitung, dass f auf den Geradenabschnitt $\{\xi + te\}$ eingeschränkt und die so erhaltene Funktion „eindimensional“ differenziert wird.

Die Richtungsableitungen in den Koordinatenrichtungen heißen *partielle Ableitungen* und werden mit $\frac{\partial f}{\partial x_i}$ oder kürzer mit f_i bezeichnet. Im Falle $n = 2$ verwenden wir auch die Schreibweise f_x, f_y .

Für eine partielle Ableitung nimmt ϕ in (2.2) eine besonders einfache Gestalt an. Man bestimmt die partielle Ableitung nach der Variablen x_i , indem man alle anderen x_j festhält und eine gewöhnliche eindimensionale Ableitung nach der Variablen x_i durchführt.

Beispiel Sei $n = 2$ und

$$f(x, y) = x^2 \exp(x + y^2).$$

Zur Bestimmung von f_x differenzieren wir nach x , wobei y als eine Konstante angesehen wird,

$$f_x(x, y) = 2x \exp(x + y^2) + x^2 \exp(x + y^2) = (x^2 + 2x) \exp(x + y^2),$$

und analog

$$f_y(x, y) = 2yx^2 \exp(x + y^2).$$

Zur Bestimmung der Ableitung im Punkt (x_0, y_0) in Richtung $e = (a, b)$ stellen wir die zugehörige Funktion ϕ aus (2.2) explizit auf,

$$\phi(t) = (x_0 + at)^2 \psi(t) \quad \text{mit} \quad \psi(t) = \exp((x_0 + at) + (y_0 + bt)^2),$$

und differenzieren sie mit Hilfe von Produkt- und Kettenregel nach t ,

$$\phi'(t) = 2(x_0 + at)a\psi(t) + (x_0 + at)^2\psi'(t)(a + 2(y_0 + bt)b),$$

daher

$$\frac{\partial}{\partial e} f(x_0, y_0) = \phi'(0) = (2ax_0 + ax_0^2 + 2x_0y_0b) \exp(x_0 + y_0^2).$$

2.6 Differenzierbare Funktionen Die Funktion $f : D \rightarrow \mathbb{R}$ heißt in $\xi \in D$ *differenzierbar*, wenn es eine lineare Funktion $L : \mathbb{R}^n \rightarrow \mathbb{R}$ gibt mit

$$f(\xi + h) - f(\xi) = L(h) + o(|h|) \quad \text{für } h \rightarrow 0.$$

Äquivalent können wir hier $h = x - \xi$ setzen,

$$f(x) = f(\xi) + L(x - \xi) + o(|x - \xi|) \quad \text{für } x \rightarrow \xi,$$

was bedeutet, dass die *Tangentialebene*

$$T(x) = f(\xi) + L(x - \xi)$$

eine $o(|x - \xi|)$ -Approximation von f in Umgebung von ξ ist.

Satz Ist f in ξ differenzierbar, so existieren alle partiellen Ableitungen und es gilt

$$L(h) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\xi) h_i.$$

Beweis: Wir setzen $h = te_i$ in die Definition der Differenzierbarkeit ein,

$$f(\xi + te_i) - f(\xi) = tL(e_i) + o(|t|).$$

Damit erfüllt $L(e_i)$ die Definition der partiellen Ableitung. Da eine lineare Funktion durch ihre Werte auf den Einheitsvektoren eindeutig bestimmt ist, haben wir den Satz bewiesen. \square

Wir schreiben die partiellen Ableitungen als Zeilenvektor,

$$\text{grad } f(\xi) = \left(\frac{\partial f}{\partial x_1}(\xi), \dots, \frac{\partial f}{\partial x_n}(\xi) \right)$$

mit dem *Gradienten* $\text{grad } f$. Mit dieser Darstellung gilt dann für differenzierbares f

$$(2.3) \quad f(\xi + h) - f(\xi) = \text{grad } f(\xi) \cdot h + o(|h|) \quad \text{für } h \rightarrow 0.$$

Für die Tangentialebene haben wir

$$z = Tx = f(\xi) + \text{grad } f(\xi)(x - \xi),$$

oder, als Hyperebene des \mathbb{R}^{n+1} geschrieben,

$$\frac{\partial}{\partial x_1} f(\xi)x_1 + \dots + \frac{\partial}{\partial x_n} f(\xi)x_n - z = \beta.$$

Aus der Hesseschen Normalenform folgt, dass

$$\nu = (\text{grad } f(\xi), -1) \in \mathbb{R}^{n+1}$$

eine Normale der Tangentialebene im Punkt $(\xi, f(\xi)) \in \mathbb{R}^{n+1}$ ist.

Für den Gradienten ist vor allem in der ingenieurwissenschaftlichen Literatur die Notation ∇ (gesprochen: *nabla*) in Gebrauch.

Satz Ist f im Punkt ξ differenzierbar, so ist f in ξ nach allen Richtungen differenzierbar und es gilt

$$\frac{\partial f}{\partial e}(\xi) = \text{grad } f(\xi) \cdot e.$$

Für $n \geq 2$ nimmt $\frac{\partial f}{\partial e}(\xi)$ alle Werte zwischen $-|\text{grad } f(\xi)|$ und $|\text{grad } f(\xi)|$ an. Falls $\text{grad } f(\xi) \neq 0$, wird der größte Wert der Richtungsableitung für die Richtung

$$e^* = \frac{\text{grad } f(\xi)}{|\text{grad } f(\xi)|}$$

angenommen.

Bemerkung Dies ist fast schon ein geflügeltes Wort: Der Gradient zeigt in Richtung des stärksten Anstiegs der Funktion.

Beweis: In (2.3) setzen wir $h = te$, teilen durch t und erhalten

$$\frac{1}{t}(f(\xi + te) - f(\xi)) = \text{grad } f(\xi) \cdot e + o(1),$$

was die behauptete Darstellung der Richtungsableitung ergibt.

Die möglichen Werte der Richtungsableitungen sind also gegeben durch

$$g(e) = \text{grad } f(\xi) \cdot e.$$

Wegen $|e| = 1$ folgt $|g(e)| \leq |\text{grad } f(\xi)|$, der maximale Wert wird damit in der Tat von $e = e^*$ angenommen und der minimale von $e = -e^*$. Da die Funktion $g(e)$ stetig von e abhängt, wird auch jeder Wert dazwischen angenommen. Dazu müssen wir allerdings den Vektor e^* unter der Nebenbedingung $|e| = 1$ stetig in den Vektor $-e^*$ überführen können, was nur im Falle $n \geq 2$ möglich ist. \square

Leider gibt es für die Differenzierbarkeit einer Funktion kein einfaches notwendiges und hinreichendes Kriterium. Wie die Übungsaufgabe 2.9 zeigt, reicht die Existenz der partiellen Ableitungen für die Differenzierbarkeit nicht aus. Definiere den Raum

$$C^1(D) = \{f : D \rightarrow \mathbb{R} : \text{Alle partiellen Ableitungen von } f \text{ existieren und sind in } D \text{ stetig}\}.$$

Satz Ist $f \in C^1(D)$, so ist f in D differenzierbar.

Beweis: Wir betrachten nur den Fall $n = 2$, für höhere Dimensionen verläuft der Beweis völlig analog. Es gilt

$$\begin{aligned} f(\xi + h) - f(\xi) &= f(\xi_1 + h_1, \xi_2 + h_2) - f(\xi_1, \xi_2) \\ &= (f(\xi_1 + h_1, \xi_2 + h_2) - f(\xi_1, \xi_2 + h_2)) + (f(\xi_1, \xi_2 + h_2) - f(\xi_1, \xi_2)) \end{aligned}$$

Wir wenden auf beide Ausdrücken den eindimensionalen Mittelwertsatz an

$$\begin{aligned} f(\xi_1 + h_1, \xi_2 + h_2) - f(\xi_1, \xi_2 + h_2) &= f_1(\eta_1, \xi_2 + h_2)h_1, \\ f(\xi_1, \xi_2 + h_2) - f(\xi_1, \xi_2) &= f_2(\xi_1, \eta_2)h_2, \end{aligned}$$

wobei η_1 auch von h_2 abhängt. Es gibt daher Punkte ζ_i mit $\zeta_i \rightarrow \xi$ für $h \rightarrow 0$ und

$$f(\xi + h) - f(\xi) = \sum_{i=1}^2 \frac{\partial f}{\partial x_i}(\zeta_i)h_i.$$

Da die partiellen Ableitungen als stetig vorausgesetzt wurden, gilt $\frac{\partial f}{\partial x_i}(\zeta_i) \rightarrow \frac{\partial f}{\partial x_i}(\xi_i)$, was (2.3) beweist. \square

2.7 Das Gradientenverfahren ist ein Algorithmus zur Bestimmung von $f(x) \rightarrow \text{Min}$ für eine stetig differenzierbare Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Ausgehend von einem Startvektor $x_0 \in \mathbb{R}^n$ bestimmt man sukzessive x_1, x_2, \dots . Sei $x_i \in \mathbb{R}^n$ bereits konstruiert. Sei $r_i = \text{grad } f(x_i)$. Ist $r_i = 0$, so endet der Algorithmus. Ist $r_i \neq 0$, so bestimmt man ein t_0 aus

$$f(x_i - t_0 r_i) = \min_{t > 0} f(x_i - t r_i).$$

Existiert solch ein t_0 nicht, so besitzt $\phi(t) = f(x_i - t r_i)$ für $t > 0$ kein Minimum und der Algorithmus muss abgebrochen werden. Andernfalls setze

$$x_{i+1} = x_i - t_0 r_i$$

und verfähre mit x_{i+1} entsprechend weiter.

Das Gradientenverfahren sucht also die nächste Iterierte immer in Richtung des stärksten Abstiegs der Funktion f .

2.8 Höhere Ableitungen Sind die partiellen Ableitungen einer Funktion ebenfalls differenzierbar, so können Ableitungen zweiter Ordnung gebildet werden,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right).$$

Für partielle Ableitungen höherer Ordnung schreiben wir

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} \quad \text{oder} \quad f_{x_{i_1} \dots x_{i_k}}.$$

Beispiel Für $n = 2$ betrachten wir wieder die Funktion $f(x, y) = x^2 \exp(x + y^2)$ mit partiellen Ableitungen $f_x(x, y) = (x^2 + 2x) \exp(x + y^2)$ und $f_y(x, y) = 2yx^2 \exp(x + y^2)$. Für die gemischten Ableitungen erhalten wir

$$f_{xy} = 2y(x^2 + 2x) \exp(x + y^2)$$

$$f_{yx} = 2yx^2 \exp(x + y^2) + 2x \exp(x + y^2) = f_{xy}$$

Wie wir später sehen werden, ist es kein Zufall, dass diese Ableitungen übereinstimmen.

$f : D \rightarrow \mathbb{R}$ heißt *von der Klasse* $C^k = C^k(D)$, wenn alle partiellen Ableitungen der Ordnung $\leq k$ existieren und in D stetig sind.

Satz [Schwarz] Für $f \in C^2(D)$ gilt

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} \quad \text{in } D \quad \text{für alle } i, j = 1, \dots, n.$$

Beweis: Es ist völlig ausreichend, nur den ebenen Fall $n = 2$ zu betrachten, denn die Behauptung ist rein zweidimensional. Sei also $n = 2$ und $\xi = 0$. Wir werden zeigen, dass der Ausdruck

$$D_{x,y,h} f(0) = \frac{1}{h^2} (f(h, h) + f(0, 0) - f(h, 0) - f(0, h))$$

für $h \rightarrow 0$ sowohl gegen $f_{xy}(0)$ als auch gegen $f_{yx}(0)$ konvergiert. Mit

$$\phi(x) = \frac{1}{h} (f(x, h) - f(x, 0)), \quad \psi(y) = \frac{1}{h} (f(h, y) - f(0, y))$$

folgt

$$D_{x,y,h} f(0) = \frac{1}{h} (\phi(h) - \phi(0)) = \frac{1}{h} (\psi(h) - \psi(0)).$$

Mit dem eindimensionalen Mittelwertsatz

$$D_{x,y,h} f(0) = \phi'(x_1) = \psi'(y_2)$$

gilt

$$D_{x,y,h} f(0) = \frac{1}{h} (f_x(x_1, h) - f_x(x_1, 0)) = \frac{1}{h} (f_y(h, y_2) - f_y(0, y_2))$$

Auf die linke Differenz wenden wir den eindimensionalen Mittelwertsatz bezüglich der zweiten Variable und auf die rechte den Mittelwertsatz bezüglich der ersten Variable an,

$$D_{x,y,h} f(0) = f_{xy}(x_1, y_1) = f_{yx}(x_2, y_2).$$

Grenzübergang $h \rightarrow 0$ liefert aufgrund der Stetigkeit der zweiten Ableitungen $f_{xy}(0) = f_{yx}(0)$. \square

Korollar Für $f \in C^m(D)$ hängen alle partiellen Ableitungen der Ordnung $\leq m$ nicht von der Reihenfolge der einzelnen Differentiationen ab.

Beispiel Sei

$$f(x, y) = \frac{xy^3}{x^2y^2} \quad \text{für } (x, y) \neq (0, 0)$$

und $f(0, 0) = 0$. Es gilt

$$\text{grad } f(x, y) = \left(\frac{y^3(y^2 - x^2)}{(x^2 + y^2)^2}, \frac{xy^2(3x^2 + y^2)}{(x^2 + y^2)^2} \right)$$

sowie

$$f_{xy}(0, y) = 1, \quad f_{y,x}(x, 0) = 0,$$

also $f_{xy}(0, 0) \neq f_{yx}(0, 0)$. Dies ist nur möglich, weil $f \notin C^2$.

2.9 Vektorwertige Funktionen und Kettenregel Eine vektorwertige Funktion $f : D \rightarrow \mathbb{R}^m$ mit $f = (f_1, \dots, f_m)$ heißt in $\xi \in D$ differenzierbar, wenn es eine lineare Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ gibt mit

$$f(\xi + h) - f(\xi) = L(h) + o(|h|).$$

Da die j -te Komponente dieser Gleichung,

$$f_j(\xi + h) - f_j(\xi) = L_j(h) + o(|h|),$$

genau dem reellwertigen Fall entspricht, ist eine Transformation genau dann differenzierbar, wenn jede ihrer Komponente differenzierbar ist. Aus dem reellwertigen Fall kennen wir auch die Gestalt von L ,

$$f(\xi + h) - f(\xi) = \begin{pmatrix} \text{grad } f_1(\xi) \\ \vdots \\ \text{grad } f_m(\xi) \end{pmatrix} \cdot h + o(|h|),$$

wobei $\text{grad } f_i$ als Zeilenvektor zu interpretieren ist. Wir nennen die $(m \times n)$ -Matrix

$$L = f' = \frac{\partial f}{\partial x} = \frac{\partial (f_1 \dots f_m)}{\partial (x_1 \dots x_n)} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Funktionalmatrix von f . Falls f differenzierbar ist, gilt

$$f(\xi + h) - f(\xi) = f'(\xi)h + o(|h|).$$

Wir schreiben $f \in C^k(D)^m$, wenn jede Komponente von f von der Klasse C^k ist.

Beispiel Sei $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ und $f(x) = Ax + b$ sei die zugehörige affin lineare Funktion. Mit $f_i(x) = \sum_j a_{ij}x_j + b_j$ folgt $f_{i,x_j}(x) = a_{ij}$ und damit $f'(x) = A$.

Satz [Kettenregel] Sei D ein Gebiet des \mathbb{R}^n und $f : D \rightarrow \mathbb{R}^m$ sei in $\xi \in D$ differenzierbar. Ferner sei E ein Gebiet des \mathbb{R}^m und $g : E \rightarrow \mathbb{R}^p$ sei in $f(\xi) \in E$ differenzierbar. Dann ist auch $h = g \circ f$ in ξ differenzierbar mit

$$\frac{\partial h}{\partial x}(\xi) = \frac{\partial g}{\partial y}(f(\xi)) \cdot \frac{\partial f}{\partial x}(\xi).$$

Dabei bezeichnet „ \cdot “ das Matrizenprodukt der $(p \times m)$ -Matrix $\frac{\partial g}{\partial y}$ mit der $(m \times n)$ -Matrix $\frac{\partial f}{\partial x}$.

Bemerkung Im Spezialfall eines reellwertigen g lautet die Kettenregel

$$\begin{aligned} \frac{\partial}{\partial x_i} h(x) &= \frac{\partial}{\partial x_i} g(f_1(x), \dots, f_m(x)) \\ &= \sum_{j=1}^m \frac{\partial}{\partial y_j} g(f_1(x), \dots, f_m(x)) \frac{\partial}{\partial x_i} f_j(x). \end{aligned}$$

Wie im eindimensionalen lautet die Regel auch hier „äußere Ableitung mal innerer Ableitung“, allerdings wird über die einzelnen partiellen Ableitungen summiert.

Beweis: Wir brauchen nur den obigen Spezialfall zu beweisen, der allgemeine Fall folgt dann aus der Definition des Matrizenprodukts. Da nur die partielle Ableitung betrachtet wird, können wir $n = 1$ setzen und erhalten

$$\begin{aligned} \frac{d}{dx} h(x) &= \lim_{h \rightarrow 0} \frac{1}{h} (g(f_1(x+h), \dots, f_m(x+h)) - g(f_1(x), \dots, f_m(x))) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} (g(f_1(x+h), \dots, f_m(x+h)) - g(f_1(x), f_2(x+h), \dots, f_m(x+h))) \\ &\quad + \dots + g(f_1(x), \dots, f_{m-1}(x), f_m(x+h)) - g(f_1(x), \dots, f_m(x))) \end{aligned}$$

Auf jede dieser Differenzen wenden wir den eindimensionalen Mittelwertsatz an und erhalten für $\eta_j \in [f_j(x), f_j(x+h)]$

$$\begin{aligned} & g(\dots, f_{j-1}(x), f_j(x+h), f_{j+1}(x+h), \dots) - g(\dots, f_{j-1}(x), f_j(x), f_{j+1}(x+h), \dots) \\ &= \frac{\partial}{\partial y_j} g(\dots, f_{j-1}(x), \eta_j, f_{j+1}(x+h), \dots) \cdot (f_j(x+h) - f_j(x)) \\ &= \frac{\partial}{\partial y_j} g(\dots, \eta_j, \dots) \cdot \frac{d}{dx} f_j(\xi_j) h, \quad \xi \in (x, x+h). \end{aligned}$$

Die Behauptung folgt wegen der Stetigkeit von g und f durch Grenzübergang $h \rightarrow 0$. \square

2.10 Mittelwertsatz und Satz von Taylor **Satz** [Mittelwertsatz] Sei $f : D \rightarrow \mathbb{R}$ und $a, b \in D$, so dass die Verbindungsstrecke \overline{ab} in D liegt. Wenn f auf \overline{ab} differenzierbar ist, so gilt

$$(2.4) \quad f(b) - f(a) = \text{grad } f(\xi) \cdot (b - a) \quad \text{für ein } \xi \in \overline{ab}.$$

Beweis: Die Funktion

$$\phi(t) = f(a + t(b - a))$$

ist aufgrund der Kettenregel im Intervall $[0, 1]$ differenzierbar mit

$$\phi'(t) = \sum_{i=1}^n \frac{\partial}{\partial x_i} f(a + t(b - a)) (b_i - a_i).$$

Aus dem eindimensionalen Mittelwertsatz $\phi(1) - \phi(0) = \phi'(\tau)$ folgt die Behauptung mit $\xi = a + \tau(b - a)$. \square
Der Mittelwertsatz in Integralform lässt sich mit gleichem Beweis zeigen,

$$f(b) - f(a) = \int_0^1 \text{grad } f(tb + (1-t)a) dt \cdot (b - a).$$

Man beachte, dass der Mittelwertsatz in der Form (2.4) nicht für vektorwertige Funktionen $f : D \rightarrow \mathbb{R}^m$ gilt, weil das angegebene ξ in jeder Komponente ein anderes sein kann. Dagegen bleibt der Mittelwertsatz in Integralform richtig,

$$(2.5) \quad f(b) - f(a) = \int_0^1 f'(tb + (1-t)a) dt \cdot (b - a),$$

wobei „ \cdot “ als Matrix-Vektor Produkt zu interpretieren ist.

Erinnert sei an die eindimensionale Taylorformel für eine Funktion $\phi : [0, 1] \rightarrow \mathbb{R}$,

$$\phi(t) = \phi(0) + \phi'(0)t + \frac{1}{2!}\phi''(0)t^2 + \dots + \frac{1}{m!}\phi^{(m)}(0)t^m + R_m$$

mit Restglied

$$R_m = \frac{1}{(m+1)!}\phi^{(m)}(\theta t)t^{m+1}, \quad 0 < \theta < 1.$$

Sei $f : D \rightarrow \mathbb{R}$ von der Klasse C^{m+1} . Liegt die Strecke $\overline{\xi \xi + h}$ innerhalb von D , so können wir für $\phi(t) = f(\xi + th)$ die Ableitungen mit der Kettenregel bestimmen,

$$\phi'(t) = \text{grad } f(\xi + th) \cdot h$$

$$\phi''(t) = \text{grad} (\text{grad } f(\xi + th) \cdot h) \cdot h = \sum_{i,j=1}^n \frac{\partial^2 f(\xi + th)}{\partial x_i \partial x_j} h_i h_j.$$

Für die k -te Ableitung von ϕ wird also über alle k -ten partiellen Ableitungen von f , die mit der entsprechenden Potenz von h multipliziert werden, summiert. Damit gilt:

Satz [Satz von Taylor] Sei $f \in C^{m+1}(D)$ und die Strecke $\overline{\xi, \xi + h}$ in D enthalten. Dann gilt

$$f(\xi + h) = f(\xi) + \frac{1}{1!} \sum_{i=1}^n f_i(\xi) h_i + \dots + \frac{1}{m!} \sum_{i_1, \dots, i_m=1}^n f_{i_1, \dots, i_m}(\xi) h_{i_1} \dots h_{i_m} + R_m(h; \xi)$$

mit dem Restglied

$$R_m(h; \xi) = \frac{1}{(m+1)!} \sum_{i_1, \dots, i_{m+1}=1}^n f_{i_1, \dots, i_{m+1}}(\xi + \theta h) h_{i_1} \dots h_{i_{m+1}}, \quad \theta \in (0, 1).$$

Setzen wir im Satz von Taylor $h = x - \xi$, so gilt

$$f(x) = T_m(x; \xi) + R_m(x; \xi)$$

wobei $T_m(x; \xi)$ ein Polynom in x vom Grade $\leq m$ ist und das Restglied die Form $R_m(x; \xi) = O(|x - \xi|^{m+1})$ besitzt.

Beispiel Im Falle $n = 2$ schreiben wir (h, k) für das Inkrement und erhalten für das Taylorpolynom vom Grade 2 wegen $f_{xy} = f_{yx}$

$$\begin{aligned} f(x+h, y+k) &= f(x, y) + f_x(x, y)h + f_y(x, y)k \\ &\quad + \frac{1}{2} f_{xx}(x, y)h^2 + f_{xy}(x, y)hk + \frac{1}{2} f_{yy}(x, y)k^2 + O(|h|^3 + |k|^3). \end{aligned}$$

Als Beispiel betrachten wir die Taylorentwicklung für die Potenzfunktion $f(x, y) = y^x = \exp(x \ln y)$. Es gilt

$$\begin{aligned} f_x &= \exp(x \ln y) \ln y = y^x \ln y, & f_y &= xy^{x-1} \\ f_{xx} &= y^x (\ln y)^2, & f_{xy} &= xy^{x-1} \ln y + y^{x-1}, & f_{yy} &= x(x-1)y^{x-2} \end{aligned}$$

Das Taylorpolynom mit Entwicklungspunkt $(1, 1)$ ist daher

$$T_2(h, k) = 1 + k + hk.$$

Wir testen die Güte dieser Formel, indem wir $0.9^{1.1}$ angenähert bestimmen, $T_2(0.1, -0.1) = 1 - 0.1 - 0.01 = 0.89$, der exakte Wert ist $0.8905\dots$

2.11 Symmetrische Matrizen Zu $A \in \mathbb{R}^{m \times n}$ definieren wir die *transponierte Matrix* $A^T \in \mathbb{R}^{n \times m}$ durch $A^T = (a_{ij}^T)_{i=1, \dots, n, j=1, \dots, m}$ mit $a_{ij}^T = a_{ji}$. $A \in \mathbb{R}^{n \times n}$ heißt *symmetrisch*, wenn $A = A^T$.

Einer symmetrischen Matrix können wir die *quadratische Form*

$$q_A(x) = \sum_{i,j=1}^n a_{ij} x_i x_j = (Ax, x)$$

zuordnen. Für quadratische Formen und ihre zugehörigen symmetrischen Matrizen definieren wir:

q_A		Bezeichnung für q_A	Bezeichnung für A
> 0	für alle $x \neq 0$	positiv	positiv definit
≥ 0	für alle x	nicht negativ	positiv semidefinit
< 0	für alle $x \neq 0$	negativ	negativ definit
≤ 0	für alle x	nicht positiv	negativ semidefinit

Beispielsweise gilt für die Einheitsmatrix $q_E = (Ex, x) = |x|^2$. q_E ist damit positiv und die Einheitsmatrix positiv definit.

Ferner bezeichnen wir eine quadratische Form und die zugehörige Matrix als *indefinit*, wenn es x, y gibt mit $q_A(x) > 0$ und $q_A(y) < 0$.

Seien $A_m = (a_{ij})_{i,j=1,\dots,m}$ für $m = 1, \dots, n$ die *Hauptuntermatrizen* von A . Die Determinanten $\Delta_m = \det A_m$ heißen dann *Hauptminoren*, also

$$\Delta_1 = a_{11}, \quad \Delta_2 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \Delta_3 = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \dots$$

Für den Beweis des folgenden Lemmas sei auf die Literatur verwiesen.

Lemma Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Dann gilt die folgende Tabelle, wobei ein Eintrag in den hinteren Spalten so zu verstehen ist, dass er äquivalent zum Eintrag in der ersten Spalte ist.

Bezeichnung für A	Eigenwerte	Hauptminoren	Bemerkung
positiv definit	$\lambda > 0$	$\Delta_m > 0$	1
positiv semidefinit	$\lambda \geq 0$	–	2
negativ definit	$\lambda < 0$	$(-1)^m \Delta_m > 0$	
negativ semidefinit	$\lambda \leq 0$	–	2
indefinit	$\exists \lambda_1 > 0$ und $\exists \lambda_2 < 0$	–	3

1 Eine Matrix ist genau dann positiv definit, wenn eine *Cholesky-Zerlegung* von A existiert, wenn also $A = LL^T$ mit einer linken unteren Dreiecksmatrix L . Dies stellt zugleich die effektivste Möglichkeit dar, um zu entscheiden, ob A positiv definit ist oder nicht.

2 Das Hauptminoren-Kriterium versagt bei semidefiniten Matrizen. Einfachstes Beispiel sind die Matrizen

$$A_{\pm} = \begin{pmatrix} 0 & 0 \\ 0 & \pm 1 \end{pmatrix}.$$

Es gilt $\Delta_1 = \Delta_2 = 0$, aber A_+ ist positiv und A_- ist negativ semidefinit.

3 Um nachzuweisen, dass A indefinit ist, kann man auch einfach Vektoren x und y angeben mit $q_A(x) > 0$ und $q_A(y) < 0$.

Beispiel Für

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

gilt $\Delta_1 = -1$ und $\Delta_2 = 1$, also ist $(-1)^m \Delta_m > 0$ erfüllt und die Matrix negativ definit. Wenn das Hauptminoren-Kriterium überhaupt zum Ziel führt, ist es natürlich einfacher als das Ausrechnen der Eigenwerte. Wir erhalten das charakteristische Polynom

$$\det(A - \lambda E) = (-2 - \lambda)(-1 - \lambda) - 1 \Rightarrow \lambda_{\pm} = \frac{\pm\sqrt{5} - 3}{2} < 0.$$

Satz Sei A symmetrisch. Ist A positiv definit (indefinit), so gibt es ein $\varepsilon > 0$, so dass alle Matrizen B mit $|B - A| < \varepsilon$ ebenfalls positiv definit (indefinit) sind.

Beweis: Die Einheitssphäre $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ ist abgeschlossen und beschränkt, also kompakt. Ist A positiv definit, so nimmt die zugehörige quadratische Form q_A auf ihr das Minimum an. Es gibt daher ein $\alpha > 0$ mit

$$q_A(x) \geq \alpha \quad \text{für alle } x \text{ mit } |x| = 1.$$

Für eine beliebige symmetrische Matrix B gilt dann

$$q_B(x) = (Bx, x) = (Ax, x) + ((B - A)x, x).$$

Aus der Cauchy-Ungleichung und der Verträglichkeit der euklidischen Matrix-Norm mit der euklidischen Vektornorm erhalten wir für $|x| = 1$

$$|((B - A)x, x)| \leq |(B - A)x| \leq |B - A|,$$

zusammen mit der letzten Gleichung für $|B - A| \leq \frac{\alpha}{2}$ daher

$$q_B(x) \geq \alpha - \frac{\alpha}{2} = \frac{\alpha}{2} \quad \text{für alle } x \text{ mit } |x| = 1.$$

Wegen der Homogenität der quadratischen Form q_B folgt hieraus $q_B(x) > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$. Damit ist B positiv definit.

Ist A indefinit, so gibt es Vektoren x, y mit $|x| = |y| = 1$ und ein $\alpha > 0$ mit $q_A(x) \geq \alpha$ und $q_A(y) \leq -\alpha$. Für $|B - A| \leq \frac{\alpha}{2}$ folgt genau wie im positiv definiten Fall, dass $q_B(x) \geq \frac{\alpha}{2}$ und $q_B(y) \leq -\frac{\alpha}{2}$. Damit ist auch B indefinit. \square

2.12 Relative Extrema Die Definitionen für ein relatives Extremum, ein relatives Minimum oder Maximum können wir wörtlich aus dem eindimensionalen Fall übernehmen. $\xi \in D$ heißt *relatives Minimum* von f , wenn es eine Umgebung U von ξ gibt mit $f(\xi) \leq f(x)$ für alle $x \in U$. Gilt sogar $f(\xi) < f(x)$ für alle $x \in U \setminus \{\xi\}$, so heißt das relative Minimum strikt. Für relative Maxima verwenden wir die gleichen Begriffsbildungen.

Auch die Herleitung der notwendigen und hinreichenden Bedingungen verläuft ganz ähnlich zum eindimensionalen Fall. Ist ξ ein relatives Extremum, so können wir f auf die Koordinatenrichtungen $\xi + te_i$ einschränken. Die eindimensionale Funktion $\phi(\xi + te_i)$ besitzt dann ein relatives Extremum an der Stelle $t = 0$. Für differenzierbares f gilt daher $\phi'(0) = 0$ und daher $f_i(\xi) = 0$.

Zur Herleitung der Bedingungen zweiter Ordnung schreiben wir zunächst das Restglied R_1 der Taylorentwicklung in kompakterer Form. Wir stellen die zweiten partiellen Ableitungen in Form einer $(n \times n)$ -Matrix zusammen,

$$H_{f(x)} = \begin{pmatrix} f_{11}(x) & \dots & f_{1n}(x) \\ \vdots & & \vdots \\ f_{1n}(x) & \dots & f_{nn}(x) \end{pmatrix},$$

die *Hesse-Matrix* genannt wird. Für $f \in C^2$ ist sie wegen des Satzes von Schwarz symmetrisch. Erinnern wir uns daran, dass die Vektoren im Zusammenhang mit Matrizen als Spaltenvektoren interpretiert werden, so schreibt sich der Satz von Taylor für $m = 1$ in der Form,

$$(2.6) \quad f(\xi + h) - f(\xi) = \text{grad } f(\xi)h + \frac{1}{2}h^T H_{f(\xi + \theta h)}h, \quad \text{mit } 0 < \theta < 1.$$

Da $\text{grad } f(\xi)$ im lokalen Extremum verschwindet, entscheidet das Verhalten der von der Hesse-Matrix induzierten quadratischen Form über die Art des extremalen Punktes ξ .

Satz [Notwendige Bedingungen für einen Extremwert] Die Funktion $f : D \rightarrow \mathbb{R}$ besitze in $\xi \in D$ ein relatives Minimum oder Maximum.

- (a) Ist $f \in C^1$, so gilt $\text{grad } f(\xi) = 0$.
- (b) Ist $f \in C^2$, so gilt zusätzlich

$$\begin{aligned} H_{f(\xi)} &\text{ ist positiv semidefinit, falls } \xi \text{ Minimum,} \\ H_{f(\xi)} &\text{ ist negativ semidefinit, falls } \xi \text{ Maximum.} \end{aligned}$$

Beweis: (b) Sei ξ ein relatives Minimum. Angenommen, $H_{f(\xi)}$ wäre nicht positiv semidefinit. Dann gibt es ein $v \in \mathbb{R}^n$ mit $|v| = 1$ und $v^T H_{f(\xi)}v < 0$. Da $H_{f(x)}$ stetig von x abhängt, gilt

$$v^T H_{f(x)}v < 0 \quad \text{für alle } x \text{ mit } |x - \xi| < \varepsilon.$$

Dies widerspricht aber für $h = tv$, t genügend klein, (2.6), weil die linke Seite für kleine h als nichtnegativ vorausgesetzt wurde. \square

Satz [Hinreichende Bedingung für einen Extremwert] Für die Funktion $f \in C^2$ sei in $\xi \in D$ die Bedingung $\text{grad } f(\xi) = 0$ erfüllt. Ferner sei $H_{f(\xi)}$ positiv (negativ) definit. Dann besitzt f in ξ ein striktes relatives Minimum (Maximum).

Beweis: Dies folgt direkt aus der Stetigkeit von $H_{f(x)}$ sowie (2.6) und Satz 2.11. \square

Satz [indefiniten Fall] Für die Funktion $f \in C^2$ sei in $\xi \in D$ die Bedingung $\text{grad } f(\xi) = 0$ erfüllt. Ferner sei $H_{f(\xi)}$ indefinit. Dann besitzt f in ξ einen Sattelpunkt, es gibt also in jeder genügend kleinen Umgebung $U(\xi)$ Punkte $x, y \in U(\xi)$ mit $f(x) > f(\xi)$ und $f(y) < f(\xi)$.

Beweis: Dies folgt ebenfalls aus der Stetigkeit von $H_{f(x)}$ sowie (2.6) und Satz 2.11. \square

Beispiel Zu Punkten $\xi_1, \dots, \xi_k \in \mathbb{R}^n$ wollen wir das Minimum von $f(x) = \sum_{i=1}^k |x - \xi_i|^2$ bestimmen. Mit

$$\text{grad } f(x) = 2 \sum_{i=1}^k (x - \xi_i) = 0$$

folgt $x_0 = \frac{1}{k} \sum \xi_i$, was physikalisch klar ist: x_0 ist der Schwerpunkt der Massepunkte ξ_1, \dots, ξ_k . Die Hesse-Matrix ist $H = 2kE$, x_0 damit striktes lokales Minimum.

Beispiel Sei $n = 2$ und $f(x, y) = e^{xy} + x^2 + \lambda y^2$ mit einem Parameter $\lambda > 0$. Wir bestimmen die ersten und zweiten Ableitungen von f ,

$$f_x(x, y) = ye^{xy} + 2x, \quad f_y(x, y) = xe^{xy} + 2\lambda y,$$

$$f_{xx}(x, y) = y^2 e^{xy} + 2, \quad f_{xy}(x, y) = (1 + xy)e^{xy}, \quad f_{yy}(x, y) = x^2 e^{xy} + 2\lambda.$$

$(x, y) = (0, 0)$ ist offenbar stationärer Punkt, d.h. $\text{grad } f(0, 0) = (0, 0)$. In diesem Fall erhalten wir die Hesse-Matrix

$$H_{f(0,0)} = \begin{pmatrix} 2 & 1 \\ 1 & 2\lambda \end{pmatrix}$$

mit den Hauptminoren $\Delta_1 = 2$ und $\Delta_2 = 4\lambda - 1$. Damit liegt für $\lambda > 1/4$ ein Minimum vor, weil die Hesse-Matrix positiv definit ist. Für $0 < \lambda < 1/4$ haben wir entsprechend einen Sattelpunkt vor uns. Für $\lambda = 1/4$ versagen dagegen die hinreichenden Kriterien und wir müssen gut hinschauen,

$$f(x, y) = 1 + \left(x + \frac{1}{2}y\right)^2 + (e^{xy} - 1 - xy).$$

Aus dem Satz von Taylor bekommen wir

$$e^s = 1 + s + \frac{1}{2}e^\xi s^2, \quad 0 < \xi < s.$$

Für $s \neq 0$ ist das Restglied strikt positiv, woraus die Ungleichung $e^s > 1 + s$ für $s \neq 0$ folgt. Aus der vorigen Darstellung von f folgt daher

$$f(x, y) > 1 + \left(x + \frac{1}{2}y\right)^2 \quad \text{für } xy \neq 0.$$

Der Punkt $(0, 0)$ ist daher für $\lambda = 1/4$ ein globales striktes Minimum.

Für $0 < \lambda < 1/4$ erhalten wir weitere stationäre Punkte, die nicht zu unserem Thema gehören und hier nur der Vollständigkeit halber angegeben werden sollen: Mit

$$\mu = \sqrt{\lambda}, \quad \eta = \sqrt{-\frac{\ln(2\mu)}{\mu}} \quad \text{für } 0 < \mu < \frac{1}{2},$$

sind

$$\pm(\xi, \eta) \quad \text{mit } \xi = -\mu\eta.$$

lokale Minima von f .

Aufgaben

2.1 Zeigen Sie: Ist $a \in \mathbb{R}^n$ mit $|a| = 1$ und $a \cdot x = 0$ die Gleichung für die zugehörige Hyperebene, so ist

$$P = E_n - aa^T$$

die Matrix der orthogonalen Projektion auf die Hyperebene.

Hinweis: Im Zusammenhang mit Matrizen fassen wir Vektoren immer als Spaltenmatrizen auf. Es gilt daher

$$a \cdot a^T = (a_i a_j)_{i,j=1,\dots,n} \in \mathbb{R}^{n \times n}$$

2.2 Bestimmen Sie den Grenzwert der Funktion

$$f(x, y) = xy/(x^2 + y^2)$$

im Punkt $x_0 = (1, 1)$. Der Definitionsbereich ist dabei die Menge $\{(x, y) \in \mathbb{R}^2 : x \neq 0 \text{ und } y \neq 0\}$.

2.3 Geben Sie ein Beispiel für eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, für die gilt $\lim_{t \rightarrow 0} f(ta) = f(0)$ für alle $a \in \mathbb{R}^2 \setminus \{0\}$, die aber im Nullpunkt unstetig ist.

2.4 Zeigen Sie, dass

$$S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$$

kompakt ist.

2.5 Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ stetig mit $f(x) > 0$ für $x \neq 0$ und $f(\lambda x) = \lambda f(x)$ für alle $x \in \mathbb{R}^n$ und $\lambda \geq 0$. Zeigen Sie: Es gibt Konstanten $m, M > 0$ mit

$$m|x| \leq f(x) \leq M|x|.$$

Hinweis: Betrachten Sie zunächst den Fall $|x| = 1$ und verwenden Sie die vorige Aufgabe.

2.6 Zu nichtleerem $A \subset \mathbb{R}^n$ und $x \in \mathbb{R}^n$ definieren wir den *Abstand*

$$\text{dist}(x, A) = \inf\{|x - y| : y \in A\}.$$

Zeigen Sie: Ist A kompakt, so gibt es ein $y_0 \in A$ mit $\text{dist}(x, A) = |x - y_0|$.

Hinweis: Verwenden Sie eine *Minimalfolge*, also eine Folge (y_k) mit $y_k \in A$ und $|x - y_k| \rightarrow \text{dist}(x, A)$.

2.7 $f : D \rightarrow \mathbb{R}^m$, $D \subset \mathbb{R}^n$, heißt *gleichmäßig stetig*, wenn es zu jedem $\epsilon > 0$ ein $\lambda > 0$ gibt mit

$$|x - y| < \lambda \Rightarrow |f(x) - f(y)| < \epsilon.$$

Zeigen sie: Ist f stetig auf D und D kompakt, so ist f gleichmäßig stetig auf D .

Bemerkung und Hinweis: Im Gegensatz zur Stetigkeit muss das zu findende $\delta > 0$ unabhängig von x, y sein. Verwenden Sie einen indirekten Beweis mit der Widerspruchannahme: Es gibt ein $\epsilon > 0$, sodass es für alle $\frac{1}{k}$, $k \in \mathbb{N}$, Punkte $x_k, y_k \in D$ gibt mit $|x_k - y_k| \leq \frac{1}{k}$ und $|f(y_k) - f(x_k)| \geq \epsilon$. Wenden Sie auf die Folgen (x_k) und (y_k) die Kompaktheit an.

2.8 Bestimmen Sie die partiellen Ableitungen von f .

a) $f(x, y) = x \ln(xy)$, $D = \{(x, y) : xy > 0\}$,

b) $f(x) = x \cdot x$, $D = \mathbb{R}^n$.

2.9 Sei $f(x, y) = |xy|^{1/3}$.

a) Zeigen Sie, dass $\pm e_1$ und $\pm e_2$ die einzigen Richtungen sind, in denen die Richtungsableitung von f im Punkt $(x, y) = 0$ existiert.

b) Zeigen Sie, dass f im Nullpunkt stetig ist.

2.10 Zeigen Sie: Wenn $\frac{\partial}{\partial a} f(x)$ existiert, so existiert auch $\frac{\partial}{\partial(-a)} f(x)$ mit

$$\frac{\partial}{\partial a} f(x) = -\frac{\partial}{\partial(-a)} f(x).$$

2.11 Bestimmen Sie die Tangentialebene der Funktion $f(x, y) = 3x^2y + 2xy^2$ im Punkt $(1, -2, 2)$.

2.12 Bestimmen Sie $\frac{1,02}{0,98}$ angenähert, indem Sie die Tangentengleichung von $f(x, y) = \frac{x}{y}$ im Punkt $(x_0, y_0) = (1, 1)$ im entsprechenden Punkt auswerten. Wie groß ist der Fehler zur exakten Division?

2.13 Bestimmen Sie die Richtungsableitung mit Hilfe von $\frac{\partial}{\partial a} f(x_0) = \text{grad } f(x) \cdot a$.

a) $f(x, y) = x \exp(xy)$, $(x_0, y_0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $a = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$,

b) $f(x, y, z) = rx^2 + sy^2 + tz^2$, $(x_0, y_0, z_0) = e_1$, $a = e_3$.

2.14 (Eulers Formel) Eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt *homogen* von Grad $p \in \mathbb{R}$, wenn

$$f(tx) = t^p f(x) \quad \forall t > 0 \quad \forall x \neq 0.$$

Zeigen Sie, dass für eine solche Funktion gilt

$$\text{grad } f(x) \cdot x = p f(x) \quad \forall x \neq 0.$$

2.15 Sei A eine symmetrische $(n \times n)$ -Matrix mit Elementen a_{ij} und $q(x) = \sum_{i,j=1}^n a_{ij} x_i x_j$ die zugehörige quadratische Form. Ferner sei $q(x) > 0$ für $x \neq 0$. Bestimmen Sie $\text{grad } f(x)$ für $f(x) = q(x)^{p/2}$, $p \in \mathbb{R}$, und rechnen Sie Eulers Formel für diesen Spezialfall nach.

2.16 Seien $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ differenzierbar. Bestimmen Sie die partiellen Ableitungen der Ordnung ≤ 2 für die Funktion

$$F(x, y) = f(x, y, g(x, y))$$

2.17 Mit

$$\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$$

bezeichnen wir den *Laplace-Operator*. Zeigen Sie für $n \geq 3$

$$-\Delta |x|^{2-n} = 0 \quad \text{in } \mathbb{R}^n \setminus \{0\}$$

2.18 Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

$$f(x, y) = \frac{xy^3}{x^2 + y^2} \quad \text{für } (x, y) \neq (0, 0), \quad f(0, 0) = 0.$$

Bestimmen Sie $f_x(x, y)$, $f_y(x, y)$, $f_{xy}(0, 0)$, $f_{yx}(x, 0)$ und zeigen Sie damit, dass $f_{xy}(0, 0) \neq f_{yx}(0, 0)$.

2.19 Sei

$$g(x, y) = \begin{pmatrix} xy \\ x^2 y \end{pmatrix}, \quad f(s, t) = \begin{pmatrix} s + t \\ s^2 - t^2 \end{pmatrix}$$

a) Bestimmen Sie $h(s, t) = g(f(s, t))$ und $h'(s, t)$.

b) Bestimmen Sie $g'(x, y)$ und $f'(s, t)$ und anschließend $h'(s, t)$ nach der Kettenregel.

2.20 Sei $g : \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar und $f(x) = Ax$ mit $A \in \mathbb{R}^{n \times n}$. Bestimmen Sie die Hesse-Matrix von $h(x) = g(f(x))$.

2.21 Zeigen Sie: Ist p ein Polynom vom Grade $\leq k$, so stimmt das Taylorpolynom $T_m(x; \xi)$ für jedes $\xi \in \mathbb{R}^n$ und $m \geq k$ mit p überein.

2.22 Für $n = 2$ und $h = 0$ definieren wir den Differenzenquotienten

$$D_{xy,h}u(x, y) = \frac{1}{4h^2} (u(x+h, y+h) - u(x+h, y-h) - u(x-h, y+h) + u(x-h, y-h))$$

Zeigen sie: Für $u \in C^4$ gilt $|D_{xy,h}u(x, y) - u_{xy}(x, y)| \leq ch^2$, wobei c nur von den vierten partiellen Ableitungen von u abhängt.

Hinweis: Verwenden Sie Taylorentwicklung im Punkt $(x, y) = 0$. Stellen Sie das Restglied nicht explizit auf, sondern verwenden Sie das Landausche Symbol.

2.23 Sei $f(x, y) = x \exp(x + y)$. Bestimmen Sie das lineare Taylorpolynom $T_1((x, y); (0, 0))$. Bestimmen Sie eine möglichst kleine Konstante M mit

$$|f(x, y) - T_1(x, y)| \leq M(x^2 + y^2), \quad |x|, |y| \leq 1.$$

2.24 Man untersuche $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^3 + y^3 + 3xy$, auf lokale Extrema.

2.25 Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) = (y - x^2)(y - 2x^2)$. Zeigen Sie:

a) Für jedes $e = (a, b) \neq 0$ hat die Funktion $\phi : \mathbb{R} \rightarrow \mathbb{R}$, $\phi(t) = f(at, bt)$ in 0 ein striktes lokales Minimum.

b) f hat in $(0, 0)$ kein lokales Minimum.

Hinweis: Zu a): Man stelle $\phi(t)$ auf und weise die Eigenschaft direkt, ohne Rückgriff auf die Analysis nach.

Zu b): Auch hier vermeide man die analytischen Hilfsmittel.

3 Metrische und normierte Räume

3.1 Metrische Räume Mit der metrischen Struktur wird der aus dem \mathbb{R}^n bekannte Abstandsbegriff abstrahiert. Wir können uns einen metrischen Raum als eine Punktmenge vorstellen, in der Entfernungen zwischen den Punkten definiert sind, die den folgenden plausiblen Bedingungen genügen müssen.

Sei X eine Menge. Eine Abbildung $d : X \times X \rightarrow [0, \infty)$ heißt *Metrik* auf X , wenn die folgenden Bedingungen erfüllt sind.

- (i) $d(x, y) = 0 \Leftrightarrow x = y$,
- (ii) $d(x, y) = d(y, x)$ (Symmetrie),
- (iii) $d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungleichung).

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

Aus der Definition der Metrik folgt die *Vierecksungleichung*

$$(3.1) \quad |d(x, y) - d(x', y')| \leq d(x, x') + d(y, y'),$$

denn die Dreiecksungleichung liefert

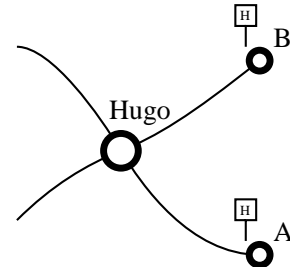
$$d(x, y) \leq d(x, x') + d(x', y') + d(y, y'), \quad d(x', y') \leq d(x, x') + d(x, y) + d(y, y'),$$

womit (3.1) gezeigt ist.

Beispiele (i) Der \mathbb{K}^n mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ mit der Metrik $d(x, y) = |x - y|$ ist metrischer Raum.

(ii) Jede Teilmenge eines metrischen Raumes ist mit der gleichen Abstandsfunktion selber ein metrischer Raum.

(iii) (ERLANGER METRIK) Um in Erlangen mit dem Bus von einem Ortsteil in den benachbarten zu kommen (Fußweg 5'), muss man zuerst zum zentralen Busbahnhof fahren, dort umsteigen und dann im wesentlichen die gleiche Strecke wieder zurückfahren. Diese Metrik, die nicht nur bei den Mathematikern, sondern auch bei den Benutzern des öffentlichen Nahverkehrs immer wieder neu auf große Begeisterung stößt, kann folgendermaßen abstrakt definiert werden: Grundraum ist der \mathbb{R}^2 mit dem Ursprung als ausgezeichneten Punkt, die Metrik ist



$$d(x, y) = \begin{cases} |x - y| & \text{wenn } x = \lambda y \text{ für ein } \lambda \in \mathbb{R}, \\ |x| + |y| & \text{sonst.} \end{cases}$$

Der Beweis der Dreiecksungleichung macht einige Fallunterscheidungen notwendig, ist ansonsten trivial. Ich habe übrigens diese Metrik unter dem Namen „Französische Eisenbahn Metrik“ kennengelernt (fährt immer über Paris), in einem amerikanischen Buch wird sie „Washington D.C. Metrik“ genannt. Die Probleme sind also überall die gleichen.

Eine Folge (x_k) heißt *konvergent*, wenn es zu jedem $\varepsilon > 0$ ein $K \in \mathbb{N}$ gibt mit

$$d(x_k, x) < \varepsilon \quad \text{für alle } k \geq K.$$

Im Falle $X = \mathbb{R}^n$ bekommen wir unseren wohlbekanntem Konvergenzbegriff zurück.

Eine Folge (x_k) im metrischen Raum X heißt *Cauchy-Folge*, wenn es zu jedem $\varepsilon > 0$ ein $K \in \mathbb{N}$ gibt mit

$$d(x_k, x_l) < \varepsilon \quad \text{für alle } k, l \geq K.$$

X heißt *vollständig*, wenn jede Cauchy-Folge gegen ein $x \in X$ konvergiert.

Jede konvergente Folge ist Cauchy-Folge, denn aus $d(x_k, x) < \varepsilon$, $d(x_l, x) < \varepsilon$ für alle $k, l \geq K$ folgt mit der Dreiecksungleichung $d(x_k, x_l) < 2\varepsilon$.

Beispiele (i) Der metrische Raum \mathbb{Q} mit der Standardmetrik $d(x, y) = |x - y|$ ist unvollständig. Als (x_k) wählen wir eine beliebige Folge in \mathbb{Q} , die gegen ein Element $x \in \mathbb{R} \setminus \mathbb{Q}$ konvergiert. Diese Folge ist Cauchy-Folge in \mathbb{R} , weil sie dort konvergiert, somit auch eine Cauchy-Folge in \mathbb{Q} , die aber in \mathbb{Q} keinen Grenzwert besitzt.

(ii) \mathbb{R} mit der Standardmetrik $d(x, y) = |x - y|$ ist ein vollständiger metrischer Raum. Denn wenn (x_k) eine Cauchy-Folge in \mathbb{R} ist, so ist diese beschränkt und enthält eine konvergente Teilfolge $x_{k_m} \rightarrow x$. Da der Betrag stetig ist, können wir in $|x_{k_m} - x_l| < \varepsilon$ zum Grenzwert übergehen und erhalten $|x - x_l| \leq \varepsilon$ für alle $l \geq K$. Damit ist die ganze Folge gegen x konvergent.

(iii) Der \mathbb{R}^n ist ebenfalls vollständig unter der Standardmetrik. Denn die Komponenten einer Cauchy-Folge des \mathbb{R}^n bilden selber eine Cauchy-Folge in \mathbb{R} . Mit \mathbb{R} ist daher auch der \mathbb{R}^n vollständig.

3.2 Der Banachsche Fixpunktsatz Für eine Abbildung $T : X \rightarrow X$ in einem metrischen Raum X möchten wir die *Fixpunktgleichung*

$$(3.2) \quad T\bar{x} = \bar{x}$$

mit Hilfe des einfachsten Verfahrens, der *sukzessiven Approximation*

$$(3.3) \quad x_{k+1} = Tx_k, \quad x_0 \in X \text{ vorgegeben,}$$

lösen. Da schon einfachste Beispiele im \mathbb{R}^1 zeigen, dass dieses Verfahren auch bei Existenz eines Fixpunktes nicht konvergieren muss, benötigen wir einschränkende Voraussetzungen an die Abbildung T .

Seien (X, d_x) und (Y, d_y) metrische Räume. $T : X \rightarrow Y$ heißt *lipschitzstetig*, wenn es ein $L \in \mathbb{R}_+$ gibt mit

$$d_y(Tx_1, Tx_2) \leq Ld_x(x_1, x_2) \quad \text{für alle } x_1, x_2 \in X.$$

L heißt dann *Lipschitzkonstante*. Wenn $X = Y$ und $L < 1$ in dieser Abschätzung gewählt werden kann, so heißt T *Kontraktion*.

Satz [Banachscher Fixpunktsatz] Sei (X, d) ein vollständiger metrischer Raum und $T : X \rightarrow X$ eine Kontraktion. Dann besitzt T genau einen Fixpunkt \bar{x} und die Folge der sukzessiven Approximation (3.3) konvergiert für alle Startwerte $x_0 \in X$ gegen \bar{x} . Weiter gilt die Fehlerabschätzung

$$d(x_k, \bar{x}) \leq \frac{L^k}{1-L} d(x_0, Tx_0).$$

Beweis: Es gibt höchstens einen Fixpunkt, denn für Fixpunkte $\bar{x}, \bar{y} \in X$ folgt

$$d(\bar{x}, \bar{y}) = d(T\bar{x}, T\bar{y}) \leq Ld(\bar{x}, \bar{y})$$

und damit $d(\bar{x}, \bar{y}) = 0$ und $\bar{x} = \bar{y}$.

Aus (3.3) erhalten wir

$$d(x_k, x_{k+1}) = d(Tx_{k-1}, Tx_k) \leq Ld(x_{k-1}, x_k) \leq L^k d(x_0, Tx_0)$$

und aus der Dreiecksungleichung

$$d(x_k, x_{k+l}) \leq \sum_{i=0}^{l-1} d(x_{k+i}, x_{k+i+1}) \leq \sum_{i=0}^{l-1} L^{k+i} d(x_0, Tx_0) = \frac{L^k - L^{k+l}}{1-L} d(x_0, Tx_0).$$

Damit ist (x_k) Cauchy-Folge und konvergiert wegen der Vollständigkeit von X gegen ein $x \in X$. Aufgrund der Stetigkeit von T kann man in der Gleichung (3.3) zum Grenzwert $k \rightarrow \infty$ gehen und erhält $x = Tx$. Die Fehlerabschätzung ergibt sich aus der letzten Ungleichung für $l \rightarrow \infty$. \square

3.3 Der Banachsche Fixpunktsatz im \mathbb{R}^n In diesem Abschnitt betrachten wir $X = D \subset \mathbb{R}^n$ versehen mit der Standardmetrik $d(x, y) = |x - y|$. Damit ist (X, d) ein metrischer Raum, die Frage ist aber, ob (\mathbb{R}^n, d) seine Vollständigkeit auf (D, d) vererbt.

Satz (D, d) ist genau dann vollständig, wenn D abgeschlossen ist.

Beweis: Sei (x_k) eine Cauchy-Folge in D . Diese ist auch eine Cauchy-Folge im \mathbb{R}^n und hat daher einen Grenzwert $x \in \mathbb{R}^n$. Ist D abgeschlossen, so gilt $x \in D$, denn andernfalls wäre x innerer Punkt von D^c und $x_k \rightarrow x$ nicht möglich. Für die andere Richtung nimmt man eine Cauchy-Folge, die gegen einen Berührungspunkt $x \notin D$ konvergiert. \square

Wir wollen uns nun der Frage zuwenden, wann im Falle $X = D \subset \mathbb{R}^n$ eine Funktion einer Lipschitzbedingung genügt. Ein Gebiet D des \mathbb{R}^n heißt *konvex*, wenn zu je zwei Punkten $x, y \in D$ auch die Verbindungsstrecke \overline{xy} zu D gehört.

Lemma Sei $D \subset \mathbb{R}^n$ ein konvexes Gebiet. Dann ist jede in D stetig differenzierbare Funktion $f : D \rightarrow \mathbb{R}^m$ mit beschränkter Ableitung $\sup_{x \in \Omega} |f'(x)| = M$ Lipschitzstetig in D mit Lipschitzkonstante M ,

$$(3.4) \quad |f(x) - f(y)| \leq M|x - y| \quad \text{für alle } x, y \in D.$$

Beweis: Für $x, y \in D$ erhalten wir aus dem Mittelwertsatz

$$f(x) - f(y) = \int_0^1 f'(tx + (1-t)y)(x - y) dt.$$

In dieser Gleichung setzen wir Beträge, verwenden $|Ax| \leq |A||x|$ und schätzen $|f'|$ durch M ab. \square

Eine differenzierbare Funktion mit unbeschränkter Ableitung ist nicht Lipschitzstetig, wie z.B. die eindimensionale Funktion $f(x) = x^2$ mit $f(x) - f(y) = (x + y)(x - y)$ zeigt.

Beispiel Sei

$$f(x) = \frac{1}{2}x + \frac{1}{8}(x - 1)^2 + \frac{1}{2}.$$

Wir zeigen, dass f den vollständigen metrischen Raum $[0, 2]$ auf sich abbildet und dort eine Kontraktion ist. Es gilt $f(0) = \frac{5}{8}$ und $f(2) = 1 + \frac{1}{8} + \frac{1}{2} < 2$. Ferner ist $f'(x) = \frac{1}{2} + \frac{1}{4}(x - 1)$ mit Nullstelle $x = 3$. Damit besitzt f im Intervall $[0, 2]$ keine Extremwerte und bildet das Intervall $[0, 2]$ auf sich ab. Für die Ableitung gilt $f'(0) = \frac{1}{2}$, $f'(2) = \frac{3}{4}$. Ferner ist $f'' = \frac{1}{4}$. Damit besitzt auch f' keine Extremwerte im Intervall $[0, 2]$ und es gilt $|f'(x)| \leq \frac{3}{4}$. Damit ist f eine Kontraktion. Der vom Banachschen Fixpunktsatz garantierte Fixpunkt ist $\bar{x} = 1$.

3.4 Banach Räume Sei X ein nicht notwendig endlich dimensionaler linearer Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Eine Abbildung $\|\cdot\| : X \rightarrow \mathbb{R}$ heißt *Norm*, wenn sie den folgenden Bedingungen genügt:

- (i) $\|x\| \geq 0$ und $\|x\| = 0 \Leftrightarrow x = 0$
- (ii) $\|\alpha x\| = |\alpha| \|x\|$ für alle $\alpha \in \mathbb{K}$ und $x \in X$.
- (iii) $\|x + y\| \leq \|x\| + \|y\|$.

Da wir diese Axiome für den \mathbb{R}^n mit $\|\cdot\| = |\cdot|$ nachgewiesen haben, ist er ein normierter Raum. Auch die Schlußfolgerungen, die wir aus diesen Axiomen gezogen haben, bleiben mit gleichem Beweis für allgemeine normierte Räume gültig, nämlich die umgekehrte Dreiecksungleichung

$$\left| \|x\| - \|y\| \right| \leq \|x - y\|.$$

sowie die eigentliche Dreiecksungleichung

$$\|x - z\| \leq \|x - y\| + \|y - z\|.$$

Aus den Normaxiomen folgt sofort, dass

$$d(x, y) = \|x - y\|_X$$

eine Metrik auf X ist. Jeder normierte Raum ist damit auch ein metrischer Raum, sodass alle Begriffsbildungen aus den letzten Abschnitten verwendet werden können. Insbesondere heißt eine Folge (x_k) in X *Cauchy-Folge*, wenn es zu jedem $\varepsilon > 0$ ein $K \in \mathbb{N}$ gibt mit

$$\|x_k - x_l\| < \varepsilon \quad \text{für alle } k, l \geq K.$$

Der normierte Raum heißt *vollständig*, wenn jede Cauchy-Folge einen Grenzwert besitzt, wenn also ein $x \in X$ existiert mit $\|x_k - x\| < \varepsilon$ für alle $k \geq K(\varepsilon)$. Ein vollständiger normierter Raum heißt *Banach Raum*. Die Räume \mathbb{R}^n und \mathbb{C}^n mit der Norm $\|x\|_X = |x|$ sind demnach nicht nur vollständige metrische Räume, sondern auch Banach Räume.

Interessanter als die genannten endlich dimensionalen Räume sind die Funktionenräume, von denen einer vorgestellt werden soll.

Satz Sei $D \subset \mathbb{R}^n$ eine kompakte Menge. Dann ist der Raum

$$C(D)^m = \{f : D \rightarrow \mathbb{R}^m : f \text{ stetig auf } D\}$$

mit der Norm

$$\|f\|_\infty = \max_{x \in D} |f(x)|$$

ein Banach Raum.

Beweis: Ist f stetig, so ist auch die reellwertige Funktion $|f|$ stetig und nimmt nach Satz 2.3 das Maximum an. Die Norm ist daher auf $C(D)^m$ wohldefiniert. Die Normaxiome lassen sich leicht überprüfen, die Dreiecksungleichung folgt aus

$$\begin{aligned} \|f + g\|_\infty &= \max_{x \in D} |f(x) + g(x)| \leq \max_{x \in D} (|f(x)| + |g(x)|) \\ &\leq \max_{x \in D} |f(x)| + \max_{x \in D} |g(x)| = \|f\|_\infty + \|g\|_\infty. \end{aligned}$$

Nun zeigen wir die Vollständigkeit des Raumes. Sei (f_k) eine Cauchy-Folge in $C(D)^m$, also

$$(3.5) \quad |f_k(x) - f_l(x)| < \varepsilon \quad \text{für alle } k, l \geq K \text{ und für alle } x \in D.$$

Insbesondere gilt $|f_k(x) - f_l(x)| < \varepsilon$ für jedes $x \in D$. Damit sind auch die Folgen $(f_k(x))$ Cauchy-Folgen des \mathbb{R}^m und besitzen einen Grenzwert, den wir $f(x)$ nennen. Da der Absolutbetrag stetig ist, können wir in (3.5) zum Grenzwert $l \rightarrow \infty$ übergehen und erhalten

$$|f_k(x) - f(x)| \leq \varepsilon \quad \text{für alle } k \geq K \text{ und für alle } x \in D.$$

Damit konvergiert (f_k) gegen f in der Norm von $C(D)^m$, insbesondere ist f als *gleichmäßiger* Grenzwert der Folge (f_k) stetig. Damit ist $f \in C(D)^m$ gezeigt. \square

Zum Schluss noch eine Variante des Banachschen Fixpunktsatzes, die sich später noch als nützlich erweisen wird:

Lemma Sei X ein Banach Raum mit Identität $E : X \rightarrow X$. Für ein $r > 0$ sei $T : B_r(0) \rightarrow X$ eine Kontraktion,

$$\|Tx - Ty\| \leq L\|x - y\| \quad \text{für alle } x, y \in B_r(0) \text{ mit einem } L < 1.$$

Gilt $\|T(0)\| < r(1 - L)$, so existiert genau ein $x \in B_r(0)$ mit $(E + T)x = 0$.

Beweis: Wir haben zu zeigen, dass $S = -T$ in $B_r(0)$ genau einen Fixpunkt besitzt. Für $x \in B_r(0)$ ist

$$\|Sx\| \leq \|Sx - S(0)\| + \|S(0)\| \leq L\|x\| + \|S(0)\| < r.$$

Damit verbleibt die Folge der sukzessiven Approximation $x_{k+1} = Sx_k$, $x_0 = 0$, in der Kugel $B_r(0)$. Wie im Beweis des Banachschen Fixpunktsatzes erhalten wir

$$(3.6) \quad \|x_{k+l} - x_k\| \leq \frac{L^k - L^{k+l}}{1 - L} \|x_1\|.$$

Damit ist (x_k) Cauchy-Folge und besitzt einen Grenzwert x . Aus (3.6) für $k = 1$ erhalten wir mit Hilfe der Dreiecksungleichung

$$\|x_{l+1}\| \leq \|x_1\| + \frac{L}{1 - L} \|x_1\| = \frac{1}{1 - L} \|x_1\| = \frac{1}{1 - L} \|S(0)\|.$$

und damit

$$\|x\| \leq \|x - x_{l+1}\| + \|x_{l+1}\| \leq \|x - x_{l+1}\| + \frac{1}{1 - L} \|S(0)\|.$$

Der erste Term auf der rechten Seite kann beliebig klein gemacht werden, der zweite ist wegen der Voraussetzung an $\|S(0)\|$ kleiner als r . Damit ist auch $\|x\| < r$ gezeigt.

Da S stetig ist, kann in $x_{k+1} = Sx_k$ zum Grenzwert übergegangen werden. Also ist $Sx = x$. Die Eindeutigkeit folgt wie im Beweis des Banachschen Fixpunktsatzes. \square

3.5 Das Newton-Verfahren Sei $D \subset \mathbb{R}^m$ ein Gebiet und $f \in C^2(D)^m$. Wir wollen eine Nullstelle von f bestimmen, was der Auflösung des nichtlinearen $m \times m$ -Gleichungssystems $f(x) = 0$ oder

$$\begin{aligned} f_1(x_1, \dots, x_m) &= 0 \\ &\vdots \\ f_m(x_1, \dots, x_m) &= 0. \end{aligned}$$

entspricht. Im Allgemeinen wird dies analytisch nicht möglich sein, weshalb man zu numerischen Verfahren greifen muss, von denen das Newton-Verfahren das wichtigste ist. Zu einem Startwert $x^0 \in D$, der möglichst in der Nähe einer Nullstelle ξ liegen sollte, sind die Iterierten des Newton-Verfahrens definiert durch

$$x^{k+1} = x^k - f_x(x^k)^{-1} f(x^k), \quad k \geq 0.$$

f_x ist hier die Kurzschreibweise für die Funktionalmatrix $\frac{\partial f}{\partial x}$. In jedem Schritt des Verfahrens muss die Funktionalmatrix bestimmt und ein lineares Gleichungssystem mit rechter Seite $f(x^k)$ gelöst werden. Natürlich muss in irgendeiner Form garantiert werden, dass die Funktionalmatrix im Punkt x^k auch regulär ist.

Man leitet das Newton-Verfahren aus dem Satz von Taylor ab. Zu x^k möchte man eine Korrektur η finden mit $f(x^k + \eta) = 0$, also

$$0 = f(x^k + \eta) = f(x^k) + f_x(x^k)\eta + O(|\eta|^2).$$

Vernachlässigt man den quadratischen Term und löst nach η auf, so $\eta = -f_x(x^k)^{-1} f(x^k)$ und mit $x^{k+1} = x^k + \eta$ erhält man das Newton-Verfahren.

Satz Sei $f \in C^2(D)^m$ mit $f(\xi) = 0$ für ein $\xi \in D$. Ferner sei $f_x(\xi)$ regulär. Dann gibt es eine Umgebung U von ξ , so dass die Iterierten des Newton-Verfahrens für alle Startwerte $x^0 \in U$ definiert sind und gegen ξ konvergieren. Für diese Startwerte gilt die Fehlerabschätzung

$$|x^{k+1} - \xi| \leq c|x^k - \xi|^2$$

mit einer Konstanten c , die von f , aber nicht von k abhängt.

Beweis: Taylorentwicklung liefert

$$0 = f(\xi) = f(x^k) + f_x(x^k)(\xi - x^k) + O(|\xi - x^k|^2).$$

Wir lösen dies nach $x^k - \xi$ auf

$$x^k - \xi = f_x(x^k)^{-1}f(x^k) + f_x(x^k)^{-1}O(|\xi - x^k|^2)$$

und setzen für x^k die Newton-Gleichung $x^k = x^{k+1} + f_x(x^k)^{-1}f(x^k)$ ein

$$x^{k+1} - \xi = f_x(x^k)^{-1}O(|\xi - x^k|^2).$$

Ist $|f_x^{-1}(\xi)| \leq M$, so gibt es ein $B_r(\xi)$, $r > 0$, mit $|f(x)^{-1}| \leq 2M$ für alle $x \in B_r(\xi)$. In $\overline{B_r(\xi)}$ sind die zweiten Ableitungen beschränkt, $|f_{xx}(x)| \leq K$. Mit einer Konstanten $c(n)$, die nur von der Raumdimension abhängt, folgt dann

$$|x^{k+1} - \xi| \leq 2Mc(n)K|x^k - \xi|^2 = c|x^k - \xi|^2.$$

Die gesuchte Umgebung U ist nun der Schnitt von $B_r(\xi)$ mit der Kugel $2Mc(n)K|x - \xi| \leq \frac{1}{2}$. Ist $x^0 \in U$, so zeigt die letzte Gleichung, dass sich in jedem Schritt des Newton Verfahrens der Fehler mindestens halbiert. Alle Iterierten verbleiben daher in U und konvergieren gegen ξ . \square

Dieser Beweis gibt das Verhalten des Newton-Verfahrens sehr gut wieder. Der Startwert muss in der Regel sehr gut sein. In diesem Fall beobachtet man häufig, dass die Konvergenz zunächst nicht besonders schnell ist, aber wenn der Fehler genügend klein geworden ist, kommt man in den Genuss der quadratischen Konvergenz: Bei jedem Schritt verdoppelt sich die Anzahl der gültigen Stellen.

Von den zahlreichen Varianten des Newton-Verfahrens sei hier das *vereinfachte Newton-Verfahren* erwähnt, das später für theoretische Zwecke verwendet wird. Man geht dabei von der Überlegung aus, dass, wenn die Iterierten nahe bei der Nullstelle liegen, sie sich nur noch wenig ändern. Dann ändert sich auch $f_x(x^k)$ nur wenig und man kann das einmal berechnete f_x für die nächsten Iterationen behalten. Man erspart sich so die Neuberechnung der partiellen Ableitungen und kann, sofern das lineare Gleichungssystem mit Gauss-Elimination gelöst wird, die *LR-Zerlegung* abspeichern und wiederverwenden.

Aufgaben

3.1 Geben sie auf \mathbb{R} eine Metrik an, sodass die Folge $(k)_{k \in \mathbb{N}}$ eine Cauchy-Folge ist, die keinen Grenzwert besitzt. Machen Sie den Raum vollständig, indem Sie weitere Elemente hinzufügen.

3.2 Sei (X, d) ein vollständiger metrischer Raum und $T : X \rightarrow X$ eine Abbildung mit der Eigenschaft, dass T^m eine Kontraktion ist für ein $m \in \mathbb{N}$. Zeigen Sie, dass T einen eindeutigen Fixpunkt besitzt.

4 Der Satz über implizite Funktionen

4.1 Implizite Funktionen – Der Fall $n = m = 1$ Beispiel Wir betrachten im \mathbb{R}^2 die Funktion $f(x, y) = x^2 + y^2 - r^2$ und fragen uns, wann wir das Nullstellengebilde $f(x, y) = 0$ lokal nach der Variablen y auflösen können, wann es also eine eindimensionale Funktion g gibt mit $f(x, g(x)) = 0$. Für $|x| < r$ lassen sich zwei Funktionen $g(x) = \pm\sqrt{r^2 - x^2}$ angeben. In diesem Beispiel können wir die Problematik der Fragestellung genau studieren: Es muss nicht immer eine solche implizite Funktion geben und selbst wenn es sie gibt, ist sie oft nur lokal in einer y -Umgebung eindeutig bestimmt.

Satz [Implizite Funktionen im Fall $n = m = 1$] Sei $B = B_r(\xi, \eta)$ die offene Kugel des \mathbb{R}^2 mit Radius $r > 0$ und sei $f \in C(B)$ mit $f(\xi, \eta) = 0$. Ferner sei f streng monoton wachsend (oder fallend) bezüglich y in B . Dann existiert ein Rechteck $R = J_x \times J_y$,

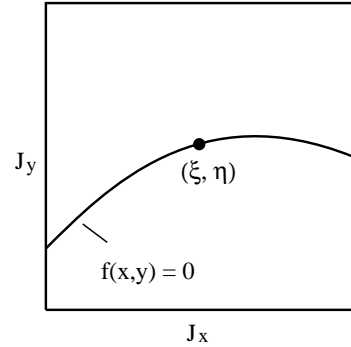
$$J_x = [\xi - \alpha, \xi + \alpha], \quad J_y = [\eta - \beta, \eta + \beta], \quad \alpha, \beta > 0,$$

und eine stetige Funktion $g : J_x \rightarrow J_y$ mit

$$f(x, g(x)) = 0 \quad \forall x \in J_x,$$

und

$$f(x, y) \neq 0 \quad \text{für } y \neq g(x).$$



Beweis: Aufgrund der strengen Monotonie von f in y gilt für $\beta = r/2$

$$f(\xi, \eta + \beta) > 0 \quad \text{und} \quad f(\xi, \eta - \beta) < 0.$$

Da f stetig ist, gibt es ein $\alpha > 0$ mit

$$f(x, \eta + \beta) > 0 \quad \text{und} \quad f(x, \eta - \beta) < 0 \quad \forall x \in J_x = [\xi - \alpha, \xi + \alpha].$$

Da f für festes x streng monoton in y ist, gibt es nach dem Zwischenwertsatz zu jedem $x \in J_x$ ein eindeutig bestimmtes $g(x)$ mit $f(x, g(x)) = 0$. Da $g(x)$ eindeutig bestimmt ist, folgt $f(x, y) \neq 0$ für $y \neq g(x)$.

Wir müssen nun noch zeigen, dass die Funktion g stetig ist. Sei $x' \in J_x$ und $\varepsilon > 0$ vorgegeben. Aus

$$f(x', g(x') + \varepsilon) > 0, \quad f(x', g(x') - \varepsilon) < 0,$$

folgt wegen der Stetigkeit von f für alle x in einer Umgebung von x'

$$f(x, g(x') + \varepsilon) > 0, \quad f(x, g(x') - \varepsilon) < 0.$$

Wegen $f(x, g(x)) = 0$ ist dies nur möglich, wenn $|g(x) - g(x')| < \varepsilon$ erfüllt ist. Damit ist die Funktion g stetig und der Satz bewiesen. \square

Die Monotoniebedingung an die Funktion f können wir garantieren, indem wir f als stetig differenzierbar voraussetzen und $f_y(\xi, \eta) \neq 0$ fordern. In diesem Fall ist auch $f_y(x, y) \neq 0$ in einer Umgebung von (ξ, η) und die Funktion f in dieser Umgebung streng monoton in y .

4.2 Implizite Funktionen – Der allgemeine Fall Hier betrachten wir Funktionen $f(x, y)$ in den Variablen $x \in \mathbb{R}^n$ und $y \in \mathbb{R}^m$ mit Werten im \mathbb{R}^m , also $f : D \rightarrow \mathbb{R}^m$ mit $D \subset \mathbb{R}^{n+m}$. Für $f(\xi, \eta) = 0$ stellen wir die Frage, ob wir das vom Parameter x abhängende nichtlineare $(m \times m)$ -

Gleichungssystem

$$\begin{aligned} f_1(x_1, \dots, x_n, y_1, \dots, y_m) &= 0 \\ &\vdots \\ f_m(x_1, \dots, x_n, y_1, \dots, y_m) &= 0 \end{aligned}$$

für (x, y) in einer Umgebung von (ξ, η) nach y_1, \dots, y_m auflösen können. Ist dies für jedes x in einer Umgebung von ξ möglich, so können wir die Lösung als Funktion $g(x)$ schreiben, für die dann gilt $f(x, g(x)) = 0$.

Beispiel Ist f affin linear, also von der Gestalt

$$f(x, y) = Ax + By + c,$$

mit $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{m \times m}$ und $c \in \mathbb{R}^m$, so lässt sich $f(x, y) = 0$ jedenfalls dann nach y auflösen, wenn B regulär ist. In diesem Fall gilt

$$y = g(x) = -(B^{-1}Ax + B^{-1}c).$$

Satz [Satz über implizite Funktionen] Sei $D \subset \mathbb{R}^{n+m}$ ein Gebiet und $f \in C^1(D)^m$. Sei ferner

$$f(\xi, \eta) = 0 \quad \text{und} \quad \frac{\partial f}{\partial y}(\xi, \eta) \text{ regulär.}$$

Dann gibt es Umgebungen $U(\xi) \subset D$ und $V(\eta) \subset \mathbb{R}^m$ und eine Funktion $g : U \rightarrow V$, $g \in C^1(U)^m$, mit

$$f(x, g(x)) = 0 \quad \text{für alle } x \in U, \quad f(x, y) \neq 0 \quad \text{für alle } y \in V \text{ mit } y \neq g(x).$$

Im Definitionsbereich von g ist die Funktionalmatrix f_y regulär und für die Funktionalmatrix von g gilt

$$(4.1) \quad \frac{\partial g}{\partial x}(x) = -\left(\frac{\partial f}{\partial y}(x, g(x))\right)^{-1} \frac{\partial f}{\partial x}(x, g(x)),$$

wobei die rechte Seite als Matrizenprodukt der $(m \times m)$ -Matrix $\left(\frac{\partial f}{\partial y}\right)^{-1}$ mit der $(m \times n)$ -Matrix $\frac{\partial f}{\partial x}$ zu verstehen ist. Ist zusätzlich $f \in C^k$, so ist auch $g \in C^k$ und die höheren Ableitungen von g können aus der Formel (4.1) durch Differenzieren bestimmt werden.

Beweisidee Wir wollen $f(x, y) = 0$ für x in Umgebung von ξ und y in Umgebung von η lösen. Dies ist ein $m \times m$ -Gleichungssystem in y , x ist nur ein Parameter. Das Newton-Verfahren zur Lösung dieses Problems ist dann

$$y_{k+1} = y_k - \frac{\partial f}{\partial y}(x, y_k)^{-1} f(x, y_k).$$

Da x nur wenig von ξ abweicht, verwenden wir das vereinfachte Newton-Verfahren und ersetzen die exakte Funktionalmatrix $f_y(x, y)$ durch $B = f_y(\xi, \eta)$.

Beweis: Wir können $\xi = 0$ und $\eta = 0$ wählen, also $f(0, 0) = 0$ voraussetzen. Für die Funktionalmatrizen verwenden wir hier immer die Kurzform, beispielsweise f_y statt $\frac{\partial f}{\partial y}$. Es sei $B = f_y(0, 0) \in \mathbb{R}^{m \times m}$, E_m die Einheitsmatrix des \mathbb{R}^m und $S(x, y) = B^{-1}f(x, y) - y$. Es gilt

$$(4.2) \quad f(x, y) = 0 \Leftrightarrow y + S(x, y) = 0.$$

Um an eine Nullstelle von f zu kommen, muss demnach die Fixpunktgleichung für $-S(x, \cdot)$ gelöst werden. S ist stetig differenzierbar mit $S_y = B^{-1}f_y - E_m$. Wir zeigen, dass $T = S(x, \cdot)$ die Voraussetzungen von Lemma 3.4 erfüllt, wenn x nahe dem Nullpunkt liegt.

Aufgrund der Definition von B gilt $S_y(0,0) = 0$. Da S_y stetig ist, existiert zu $L \in (0,1)$ ein $\delta > 0$ mit $|S_y(x,y)| < L$ für alle $(x,y) \in B_\delta(0) \times B_\delta(0)$. Für $x \in B_\delta(0)$ und $y, y' \in B_\delta(0)$ liefert der Mittelwertsatz in der Form (2.5)

$$|S(x,y) - S(x,y')| = \left| \int_0^1 S_y(x, y' + t(y - y'))(y - y') dt \right| \leq L|y - y'|.$$

Da $S(0,0) = 0$ und $S(\cdot, 0)$ stetig ist, existiert ein $s \leq \delta$ mit $|S(x,0)| < \delta(1-L)$ für alle $x \in B_s(0)$. Nach Lemma 3.4 existiert zu $x \in B_s(0)$ genau ein $y \in B_\delta(0)$ mit $y + S(x,y) = 0$. Dieses y nennen wir $g(x)$. Es gilt dann $g : B_s(0) \rightarrow B_\delta(0)$ und $f(x, g(x)) = 0$ für alle $x \in B_s(0)$.

Nun weisen wir die Stetigkeit von g nach. Wegen (4.2) gilt für $x, x' \in B_s(0)$

$$0 = g(x) + S(x, g(x)) = g(x') + S(x', g(x'))$$

und daher

$$\begin{aligned} |g(x) - g(x')| &\leq |S(x', g(x')) - S(x, g(x'))| + |S(x, g(x')) - S(x, g(x))| \\ &\leq |S(x', g(x')) - S(x, g(x'))| + L|g(x) - g(x')|, \end{aligned}$$

also $|g(x) - g(x')| \leq |S(x', g(x')) - S(x, g(x'))| / (1-L)$. Die Stetigkeit von g folgt daher aus der Stetigkeit von S bezüglich der ersten Variable. Man beachte, dass wir die Differenzierbarkeit von f nach x bisher nicht verwendet haben.

Als letztes zeigen wir die Differenzierbarkeit von g . Für $x \in B_s(0)$ ist für genügend kleine h auch $x+h \in B_s(0)$. Dann folgt aus dem Mittelwertsatz 2.10

$$0 = f(x+h, g(x+h)) - f(x, g(x)) = f_x(\xi, \eta)h + f_y(\xi, \eta)(g(x+h) - g(x))$$

mit (ξ, η) auf der Verbindungsstrecke der Punkte $(x, g(x))$ und $(x+h, g(x+h))$. Durch eventuelles Verkleinern von $B_s(0)$ können wir erreichen, dass die Funktionalmatrix $f_y(\xi, \eta)$ regulär ist, denn die Determinante ist eine stetige Funktion der Koeffizienten einer Matrix. Für diese x folgt dann

$$g(x+h) - g(x) = -f_y(\xi, \eta)^{-1} f_x(\xi, \eta)h = -f_y(x, y)^{-1} f_x(x, y)h + o(h)$$

Hier haben wir verwendet, dass f und g stetige Funktionen sind. Mit $\lim g(x+h) = g(x)$ folgt $\lim f(\xi, \eta) = f(x, g(x))$. Damit ist g differenzierbar mit $g_x = -f_y^{-1} f_x$. Aus dieser Darstellung folgt auch, dass g_x stetig ist, denn die rechte Seite ist es. Die Stetigkeit von f_y^{-1} folgt aus der Cramerschen Regel: f_y^{-1} ist eine rationale Funktion der Elemente von f_y und stetig, solange der Nenner $\det f_y$ nicht verschwindet. \square

Wir wollen die Formel (4.1) durch *implizite Differentiation* nachvollziehen. Da wir wegen des Satzes über implizite Funktionen wissen, dass die implizit gegebene Funktion g existiert und differenzierbar ist, können wir die Gleichung $h(x) = f(x, g(x)) = 0$ nach x mit der Kettenregel differenzieren,

$$0 = \underbrace{\frac{\partial}{\partial x} [f(x, g(x))]}_{\text{totale Ableitung}} = \frac{\partial f}{\partial x}(x, g(x)) + \frac{\partial f}{\partial y}(x, g(x)) \frac{\partial g}{\partial x}(x)$$

Da f_y regulär ist, können wir diese Formel nach g_x auflösen und erhalten (4.1).

Auch die nach dem Satz über implizite Funktionen existierenden höheren Ableitungen berechnet man am besten durch implizite Differentiation. Im Spezialfall $m = n = 1$ erhalten wir aus $f(x, g(x)) = 0$

$$f_x + f_y g' = 0$$

und indem wir diese Gleichung weiter differenzieren,

$$f_{xx} + f_{xy} g' + f_{yx} g' + f_{yy} (g')^2 + f_y g'' = 0.$$

Dies lösen wir nach g'' auf und erhalten mit $g' = -f_x/f_y$

$$g'' = -\frac{1}{f_y} \left(f_{xx} - \frac{2f_{xy}f_x}{f_y} + \frac{f_{yy}f_x^2}{f_y^2} \right) = -\frac{1}{f_y^3} (f_{xx}f_y^2 - 2f_{xy}f_xf_y + f_{yy}f_x^2).$$

Wir bleiben beim Spezialfall $m = n = 1$ und wollen die Höhenlinien einer Funktion $f : D \rightarrow \mathbb{R}$ untersuchen. Wir nennen die Menge der Punkte (x, y) mit $f(x, y) = c$ die c -Höhenlinie von f . Aber ist dies tatsächlich eine Linie? Wenn $\text{grad } f(x, y) \neq 0$ in einem Punkt der Höhenlinie erfüllt ist, so muss wenigstens eine der beiden partiellen Ableitungen nicht verschwinden. Nach dem Satz über implizite Funktionen kann die Gleichung $f(x, y) = c$ durch $y(x)$ oder $x(y)$ lokal implizit aufgelöst werden. In jedem Fall ist die Höhenlinie in der Tat eine Linie.

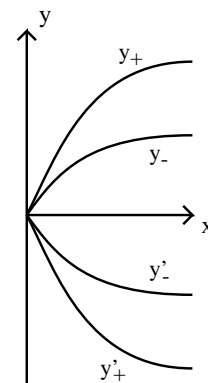
Nun betrachten wir den Spezialfall $n = 1$, was dem Auflösen eines nichtlinearen $(m \times m)$ -Gleichungssystems, das von einem Parameter x abhängt, entspricht,

$$\begin{aligned} f_1(x, y_1, \dots, y_m) &= 0 \\ &\vdots \\ f_m(x, y_1, \dots, y_m) &= 0. \end{aligned}$$

Beispiel Sei $f_1(x, y_1, y_2) = y_1y_2$, $f_2(x, y_1, y_2) = y_1^2 + y_2^2 - x$. Für die Funktionalmatrix von f bezüglich y erhalten wir

$$\frac{\partial f}{\partial y}(x, y) = \begin{pmatrix} y_2 & y_1 \\ 2y_1 & 2y_2 \end{pmatrix}$$

mit $\det f_y(y_1, y_2) = 2y_2^2 - 2y_1^2$. Die Funktionaldeterminante verschwindet genau dann, wenn $y_1 = \pm y_2$. Die Lösungen von $f(x, y) = 0$ sind $y_{\pm} = (0, \pm\sqrt{x})$ und $y'_{\pm} = (\pm\sqrt{x}, 0)$. Für $x > 0$ gibt es daher vier Lösungen, die für $x = 0$ zusammenfallen. Dort ist entsprechend die Funktionalmatrix f_y singular. Wir können die vom Parameter x abhängenden Lösungen in ein *Verzweigungsdiagramm* einzeichnen, das in symbolischer Form Auskunft gibt, wie und für welche x -Werte Lösungen zusammenfallen bzw. sich verzweigen.



Ganz allgemein gilt: Solange die Funktionalmatrix f_y regulär ist, können im Verzweigungsdiagramm weder Verzweigungspunkte noch senkrechte Tangenten auftreten. Beides schließt der Satz über implizite Funktionen aus.

4.3 Der Umkehrsatz Wir betrachten einen Spezialfall des Satzes über implizite Funktionen, nämlich $n = m$ der Form

$$\begin{aligned} f_1(x, \dots, x_n) &= y_1 \\ &\vdots \\ f_n(x_1, \dots, x_n) &= y_n \end{aligned}$$

und fragen uns, wann die Transformation f lokal umkehrbar ist, oder anders ausgedrückt, ob es eine lokal definierte Transformation g gibt mit $f(g(y)) = y$.

Satz Sei $f \in C^1(D)^n$ in einem Gebiet $D \subset \mathbb{R}^n$. Für $\xi \in D$ gelte $f(\xi) = \eta$ und die $n \times n$ -Matrix $f_x(\xi)$ sei regulär. Dann gibt es eine Umgebung $U(\xi)$ mit

- (a) $V = f(U)$ ist Umgebung von η
- (b) $f : U \rightarrow V$ ist bijektiv.

(c) Die Umkehrabbildung $g : V \rightarrow U$ ist stetig differenzierbar mit

$$\frac{\partial}{\partial y} g(y) = \left(\frac{\partial}{\partial x} f(x) \right)^{-1} \quad \text{mit } y = f(x) \text{ und } x \in U(\xi).$$

Ist $f \in C^k$, so auch $g \in C^k$.

Beweis: Wir können den Satz über implizite Funktionen anwenden, wobei zu beachten ist, dass hier die Rollen von x und y vertauscht sind. Wir schreiben $F(x, y) = f(x) - y$. Ist nun $f(\xi) = \eta$ und die Matrix $f_x(\eta)$ regulär, so entspricht das $F(\xi, \eta) = 0$ und $F_x(\xi, \eta)$ regulär. Nach dem Satz über implizite Funktionen gibt es daher eine Umgebung U von ξ und eine Umgebung V von η sowie eine Funktion $g : V \rightarrow U$ stetig differenzierbar mit $F(g(y), y) = 0$ oder $f(g(y)) = y$. f ist damit bijektiv von $U' = g(V) \subset U$ nach V . Für die Funktionalmatrix von g erhalten wir

$$\frac{\partial g}{\partial y} = - \left(\frac{\partial F}{\partial x} \right)^{-1} \frac{\partial F}{\partial y} = \left(\frac{\partial f}{\partial x} \right)^{-1}.$$

Dies stellt eine Verallgemeinerung der wohlbekannteren Formel für die eindimensionale Umkehrfunktion dar.

f ist auf U definiert und stetig, U' ist die Urbildmenge der offenen Menge V und damit selber offen. \square

Als Anwendung betrachten wir Polarkoordinaten im \mathbb{R}^2 ,

$$x = r \cos \phi, \quad y = r \sin \phi,$$

mit Funktionalmatrix

$$\frac{\partial(x, y)}{\partial(r, \phi)} = \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix}.$$

Die Funktionaldeterminante ist r . Damit ist die Transformation lokal umkehrbar beispielsweise auf der offenen Menge $r > 0$ und $0 < \phi < 2\pi$. Offenbar ist die Transformation auf diesem Bereich auch global umkehrbar mit Umkehrfunktion $r = (x^2 + y^2)^{1/2}$ und $\phi = \arg(x, y)$ oder

$$\phi = \arctan \frac{y}{x} \quad \text{falls } x \neq 0, \quad \phi = \operatorname{arccot} \frac{x}{y} \quad \text{falls } y \neq 0.$$

Seien $U, V \subset \mathbb{R}^n$ offen und $f : U \rightarrow V$ sei bijektiv. Sind f und die Umkehrabbildung $g = f^{-1}$ offen, so heißt f *Homöomorphismus*. Ist f zusätzlich von der Klasse C^1 mit $\frac{\partial f}{\partial x}$ regulär, so heißt f *Diffeomorphismus*.

4.4 Extremwertaufgaben mit Nebenbedingungen Sei $D \subset \mathbb{R}^2$ ein Gebiet und $f, g \in C^1(D)$. Gesucht ist $(x, y) \in D$, das die Aufgabe

$$f(x, y) \rightarrow \text{Min} \quad \text{unter allen } (x, y) \in D \text{ mit } g(x, y) = 0$$

löst. Nach dem Satz über implizite Funktionen wird der Bereich

$$Z = \{(x, y) \in D : g(x, y) = 0\}$$

sich i.a. aus Linien zusammensetzen. Sei (ξ, η) eine Lösung dieses Optimierungsproblems mit $\operatorname{grad} g(\xi, \eta) \neq 0$. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $g_y(\xi, \eta) \neq 0$. Nach dem Satz über implizite Funktionen gibt es Umgebungen U von ξ und V von η sowie eine Funktion $h : U \rightarrow V$ mit

$$g(x, h(x)) = 0 \quad \forall x \in U.$$

Die Funktion

$$F(x) = f(x, h(x))$$

besitzt daher ein lokales Minimum für $x = \xi$, also

$$F'(\xi) = f_x(\xi, h(\xi)) + f_y(\xi, h(\xi))h'(\xi) = 0.$$

Nach dem Satz über implizite Funktionen ist $h' = -g_x/g_y$, also

$$f_x(\xi, h(\xi)) + \frac{-f_y(\xi, h(\xi))}{g_y(\xi, h(\xi))}g_x(\xi, h(\xi)) = 0.$$

Mit $\lambda_0 = -\frac{f_y}{g_y}(\xi, h(\xi))$ folgt $f_y + \lambda_0 g_y = 0$ und nach der letzten Formel $f_x + \lambda_0 g_x = 0$, zusammen also

$$(4.3) \quad \text{grad } f(\xi, \eta) + \lambda_0 \text{grad } g(\xi, \eta) = 0.$$

Dies ist die notwendige Bedingung für einen Extremwert der restringierten Extremwertaufgabe. Zusammen mit der Restriktion $g(x, y) = 0$ bildet (4.3) ein nichtlineares Gleichungssystem in den drei Variablen x, y, λ , mit dessen Lösung man sich alle Kandidaten für einen Extremwert verschaffen kann.

Nun betrachten wir den allgemeinen Fall von m Restriktionen in einem Gebiet $D \subset \mathbb{R}^n$. f sei eine auf D definierte reellwertige Funktion und $g : D \rightarrow \mathbb{R}^m$. Die zulässige Menge

$$Z = \{x \in D : g(x) = 0\}$$

oder ausgeschrieben

$$\begin{aligned} g_1(x_1, \dots, x_n) &= 0, \\ &\vdots \\ g_m(x_1, \dots, x_n) &= 0, \end{aligned}$$

wird im Allgemeinen eine $(n - m)$ -dimensionale Fläche sein.

Sei $m \leq n$. $\xi \in Z$ heißt *regulärer Punkt* von $g(x) = 0$, wenn die Funktionalmatrix $g_x(\xi)$ Rang m hat. Dieser Bedingung entspricht im oben betrachteten Fall $n = 2$ und $m = 1$, dass $\text{grad } g(\xi, \eta) \neq 0$.

Satz Seien $f \in C^1(D)$, $g \in C^1(D)^m$ und $m \leq n$. $\xi \in D$ sei lokales Extremum von f unter der Bedingung $g(x) = 0$. Ist ferner ξ ein regulärer Punkt von $g(x) = 0$, so gibt es ein $\lambda_0 \in \mathbb{R}^m$, so dass das *Lagrange-Funktional*

$$L(x, \lambda) = f(x) + \lambda \cdot g(x) = f(x) + \sum_{j=1}^m \lambda_j g_j(x)$$

stationär in $(\xi, \lambda_0) \in \mathbb{R}^{m+n}$ wird, d.h.

$$\frac{\partial L}{\partial x}(\xi, \lambda_0) = \text{grad } f(\xi) + \sum_{j=1}^m \lambda_{0,j} \text{grad } g_j(\xi) = 0$$

$$\frac{\partial L}{\partial \lambda}(\xi, \lambda_0) = g(\xi) = 0.$$

Beweis: Der Beweis verläuft völlig analog zum oben betrachteten Fall. Da die Funktionalmatrix $g_x(\xi)$ den Rang $m \leq n$ besitzt, gibt es m linear unabhängige Spalten. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass dies die letzten m sind. Wir schreiben $x = (r, y)$ mit $r \in$

\mathbb{R}^{m-n} und $y \in \mathbb{R}^m$ sowie $\xi = (\xi_r, \xi_y)$ und $f(r, y), g(r, y)$. Nach Voraussetzung ist die $(m \times m)$ -Matrix $g_y(\xi_r, \xi_y)$ regulär. Nach dem Satz über implizite Funktionen gibt es eine Funktion $h : U(\xi_r) \rightarrow V(\xi_y)$ mit

$$g(r, h(r)) = 0 \quad \forall r \in U(\xi_r).$$

$f(r, h(r))$ besitzt ein relatives Extremum in ξ_r , also

$$0 = \frac{\partial}{\partial r}(f(\xi_r, h(\xi_r))) = \frac{\partial f}{\partial r}(\dots) + \frac{\partial f}{\partial y}(\dots) \frac{\partial h}{\partial r}(\xi_r).$$

Nach dem Satz über implizite Funktionen gilt $h_r = -g_y^{-1}g_r$ und damit

$$\underbrace{\frac{\partial f}{\partial r}(\dots)}_{\in \mathbb{R}^{1 \times (n-m)}} - \underbrace{\frac{\partial f}{\partial y}(\dots)}_{\in \mathbb{R}^{1 \times m}} \underbrace{\frac{\partial g}{\partial y}(\dots)^{-1}}_{\in \mathbb{R}^{m \times m}} \underbrace{\frac{\partial g}{\partial r}(\dots)}_{\in \mathbb{R}^{m \times (n-m)}} = 0.$$

Mit $\lambda_0 = -f_y(\dots)g_y(\dots)^{-1} \in \mathbb{R}^m$ folgt

$$\frac{\partial f}{\partial y}(\xi) + \lambda_0 \frac{\partial g}{\partial y}(\xi) = 0, \quad \frac{\partial f}{\partial r}(\xi) + \lambda_0 \frac{\partial g}{\partial r}(\xi) = 0,$$

zusammen also

$$\frac{\partial f}{\partial x}(\xi) + \lambda_0 \frac{\partial g}{\partial x}(\xi) = 0.$$

□

Beispiel Im Falle $n = 2$ wollen wir die Funktion $f(x, y) = (x-1)^2 + y^2$ unter der Nebenbedingung $g(x, y) = x^2 = 0$ minimieren. Der Bereich, in dem die Extremwerte gesucht werden, ist damit

$$Z = \{(0, y) : y \in \mathbb{R}\}.$$

Dieses Problem hat offenbar die Lösung $(\xi, \eta) = (0, 0)$, die notwendige Bedingung

$$\text{grad } f(0, 0) + \lambda \text{grad } g(0, 0) = \left(\begin{array}{c} 2(x-1) + 2\lambda x \\ 2y + \lambda \cdot 0 \end{array} \right) \Big|_{(x,y)=(0,0)} = \left(\begin{array}{c} -2 \\ 0 \end{array} \right) \stackrel{!}{=} \left(\begin{array}{c} 0 \\ 0 \end{array} \right).$$

wird jedoch von keinem $\lambda \in \mathbb{R}$ erfüllt. Grund dafür ist $\text{grad } g = 0$ auf Z , es gibt also keine regulären Punkte in Z . Verwenden wir dagegen die äquivalente Parametrisierung $g(x, y) = x$, so ist die notwendige Bedingung

$$\left(\begin{array}{c} 2(x-1) + \lambda \\ 2y + \lambda \cdot 0 \end{array} \right) = 0, \quad x = 0,$$

eindeutig lösbar. Wir sagen in einem solchen Fall: Z wird durch g regulär parametrisiert.

Bemerkung Für $m = n$ und $\text{rang } g_x(\xi) = n$ ist die notwendige Bedingung leer, weil nach dem Satz über implizite Funktionen ξ ein isolierter Punkt ist.

Beispiele (i) Sei A eine symmetrische $(n \times n)$ -Matrix. Wir betrachten das Problem

$$q_A(x) = x^T A x \rightarrow \text{Min} \quad \text{unter } |x|^2 = 1.$$

Mit dem zugehörigen Lagrange-Funktional

$$L(x, \lambda) = x^T A x + \lambda(1 - |x|^2)$$

folgt

$$\frac{\partial}{\partial x} L(x, \lambda) = 2Ax - 2\lambda x = 0.$$

Damit sind die Eigenwerte und die zugehörigen Eigenvektoren die stationären Punkte dieses Optimierungsproblems. Es gilt

$$\lambda_{\min} = \min_{|x|=1} q_A(x) = \min_x R(x),$$

$$\lambda_{\max} = \max_{|x|=1} q_A(x) = \max_x R(x),$$

mit dem *Rayleigh-Quotienten*

$$R(x) = \frac{x^T A x}{x^T x} = \frac{q_A(x)}{|x|^2}$$

(ii) Sei $A \in \mathbb{R}^{m \times n}$ mit $m < n$ und $\text{rang } A = m$. Da das unterbestimmte Gleichungssystem $Ax = b$ mehrdeutig ist, sucht man diejenige Lösung, die eine vorgegebene Funktion minimiert. Im einfachsten Fall löst man das Problem

$$|x|^2 \rightarrow \text{Min} \quad \text{unter } Ax = b.$$

Für $\lambda \in \mathbb{R}^m$ lautet das zugehörige Lagrange-Funktional

$$L(x, \lambda) = |x|^2 + \lambda(Ax - b)$$

mit der notwendigen Bedingung

$$(4.4) \quad 2x + A^T \lambda = 0.$$

Wir multiplizieren diese Gleichung von links mit A . Da A Rang m hat, ist AA^T invertierbar und wir erhalten

$$2Ax + AA^T \lambda = 0 \quad \Rightarrow \quad \lambda = -2(AA^T)^{-1}b.$$

Dies mit (4.4) kombiniert ergibt die Lösung

$$x = A^T(AA^T)^{-1}b.$$

4.5 Konvexität

Sei X ein \mathbb{K} -Vektorraum ($\mathbb{K} = \mathbb{R}$ oder \mathbb{C}).

Eine Teilmenge A von X heißt *konvex*, wenn mit $x, y \in A$ auch die Verbindungsstrecke in A liegt, wenn also $tx + (1-t)y \in A$ für alle $t \in [0, 1]$.

Eine auf einer konvexen Teilmenge A des Vektorraums X definierte reellwertige Funktion f heißt *konvex*, wenn für alle $x, y \in A$ und $t \in [0, 1]$ gilt

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

f heißt *konkav*, wenn $-f$ konvex ist.



Eine konvexe und eine konkave Funktion

Sind $x_1, \dots, x_k \in X$ und $t_1, \dots, t_k \in \mathbb{R}$ mit $0 \leq t_i \leq 1$ und $\sum_i t_i = 1$, so heißt $\sum_i t_i x_i$ *Konvexkombination* der x_i .

Die *konvexe Hülle* einer Menge $A \subset X$ besteht aus allen Konvexkombinationen von Elementen von A .

Satz Sei X ein \mathbb{K} -Vektorraum.

(a) Eine Menge $A \subset X$ ist genau dann konvex, wenn alle Konvexkombinationen von Elementen von A in A liegen.

(b) Sei $A \subset X$ konvex. Eine Funktion $f : A \rightarrow \mathbb{R}$ ist genau dann konvex, wenn für alle Konvexkombinationen $\sum_i t_i x_i$, $x_i \in A$, gilt

$$f\left(\sum_{i=1}^k t_i x_i\right) \leq \sum_{i=1}^k t_i f(x_i).$$

(c) Die konvexe Hülle einer Menge $A \subset X$ ist konvex.

Beweis: (a) Für konvexes A ist die Behauptung für $k = 2$ erfüllt. Für $k > 2$ verwenden wir Induktion über k . Für eine Konvexkombination $\sum_{i=1}^k t_i x_i$ mit $0 < t_k < 1$ folgt aus

$$y = \sum_{i=1}^{k-1} \frac{1}{1-t_k} t_i x_i \in A,$$

daß

$$\sum_{i=1}^k t_i x_i = (1-t_k)y + t_k x_k \in A.$$

(b) ist wie (a).

(c) folgt aus der Tatsache, daß eine Konvexkombination von Konvexkombinationen wiederum eine Konvexkombination ergibt. \square

Satz Sei $\Omega \subset \mathbb{R}^n$ ein konvexes Gebiet. $f \in C^2(\Omega)$ ist genau dann konvex (konkav), wenn die Matrix $(D^2 f(x))$ für alle $x \in \Omega$ positiv (negativ) semidefinit ist.

Beweis: Sei $x_0 \in \Omega$ und $y \in \mathbb{R}^n \setminus \{0\}$. Die Funktion $\phi(t) = f(x_0 + ty)$ ist in einer Umgebung von 0 definiert. Wir schreiben den Satz von Taylor in der Form

$$\phi(t) - \phi(0) - \phi'(0)t = \frac{1}{2}\phi''(\tau)t^2, \quad \tau \in (0, t).$$

Ist f und damit ϕ konvex, so ist die linke Seite nichtnegativ, weil die Tangente einer konvexen Funktion unterhalb ihres Graphen liegt. Division durch t^2 und Grenzübergang $t \rightarrow 0$ liefern $\phi''(0) \geq 0$ und damit nach der Kettenregel $y^T D^2 f(x_0) y \geq 0$. Die umgekehrte Richtung zeigt man analog. \square

Beispiel Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische positiv definite Matrix. Für die Aufgabe

$$F(x) = (Ax, x) - 2(b, x) \rightarrow \text{Min}$$

erhalten wir $F_x(x) = Ax - b$ und $H_F = A$. Damit ist F strikt konvex und die Aufgabe $F(x) \rightarrow \text{Min}$ besitzt ein eindeutiges Minimum, nämlich die Lösung des linearen Gleichungssystems $Ax = b$. Also: Bei symmetrischen positiv definiten Matrizen lässt sich die Lösung des linearen Gleichungssystems durch ein Minimierungsproblem charakterisieren.

Aufgaben

4.1 Die Gleichung $z^3 + z + xy = 1$ hat für jedes $(x, y) \in \mathbb{R}^2$ genau eine Lösung $g(x, y)$. Man zeige, dass $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ differenzierbar ist und berechne $g'(1, 1)$. Man untersuche g auf Extrema.

4.2 Man zeige, dass das Gleichungssystem

$$x^2 + uy + e^v = 0,$$

$$2x + u^2 - uv = 5,$$

in einer Umgebung von $(x, y) = (2, 5)$ durch eine C^1 -Transformation

$$(x, y) \mapsto (u(x, y), v(x, y))$$

mit $u(2, 5) = -1$ und $v(2, 5) = 0$ aufgelöst werden kann und berechne die Funktionalmatrix dieser Transformation.

4.3 Man zeige, dass die Gleichung $y^2 + xz + z^2 - e^{xz} = 1$ in einer Umgebung des Punktes $(0, -1, 1)$ in der Form $z = g(x, y)$ eindeutig auflösbar ist und berechne die Taylorentwicklung von g um den Punkt $(0, -1)$ bis zu den Gliedern der Ordnung 2.

4.4 Man bestimme den achsenparallelen Quader größten Volumens, der dem Ellipsoid $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$ eingeschrieben ist.

4.5 Es seien a_1, \dots, a_n beliebige positive Zahlen und p_1, \dots, p_n positive Zahlen mit $p_1 + \dots + p_n = 1$. Man zeige

$$a_1^{p_1} \cdot \dots \cdot a_n^{p_n} \leq p_1 a_1 + \dots + p_n a_n$$

4.6 Bestimmen Sie das Maximum der Funktion $f(x) = x_1 + \dots + x_n$ unter der Nebenbedingung $x_1^2 + \dots + x_n^2 = 1$.

4.7 Bestimmen Sie das Maximum der Funktion $f(x) = 5x_1 + x_2 - 3x_3$ unter den Nebenbedingungen $x_1 + x_2 + x_3 = 1$ und $x_1^2 + x_2^2 + x_3^2 = 1$.

5 Kurven

5.1 Kurven im \mathbb{R}^n Wir haben eine anschauliche Vorstellung davon, was eine „Kurve“ im \mathbb{R}^n ist, etwa in der Ebene der Graph einer Funktion $f : [a, b] \rightarrow \mathbb{R}$. Auch der Rand eines Gebiets des \mathbb{R}^2 , z.B. die Einheitskreislinie, ist eine Kurve. Diese ist aber nicht als Graph darstellbar, sondern als Zusammensetzung von Graphen. Bei der Definition, was eine allgemeine Kurve sein soll, hilft die kinematische Vorstellung eines physikalischen Partikels, das in einem Zeitintervall $[a, b]$ einen Weg $\{x(t), t \in [a, b]\}$ im \mathbb{R}^n durchläuft. Dabei kann es passieren, dass die Bahnkurve gewisse Punkte mehrfach durchläuft, z.B. die ebenen Bahnkurven

$$x(t) = (t^2 - 1, t^3 - t), \quad t \in \mathbb{R}, \quad x(t) = (\cos(2t), \sin(2t)), \quad t \in [0, 2\pi].$$

Die erste hat einen sog. „Doppelpunkt“ $x(1) = x(-1)$ und bei der zweiten durchläuft der Punkt den Einheitskreis zweimal. Wir müssen also unterscheiden zwischen dem „Weg“, den ein Partikel in dem Zeitintervall durchläuft, und der entstehenden „Kurve“ als Punktmenge des \mathbb{R}^n . Ein solcher Weg hat eine natürliche Orientierung, gegeben durch die zeitliche Bewegung des Partikels.

Unter einem *Weg* verstehen wir eine stetige Abbildung $\phi : [a, b] \rightarrow \mathbb{R}^n$ auf einem abgeschlossenen (nicht degenerierten) Intervall $I = [a, b] \subset \mathbb{R}$. (Die Fälle $I = (-\infty, b]$, $I = [a, \infty)$ und $I = (-\infty, \infty)$ sind gelegentlich zugelassen.) Die Bildmenge

$$\Gamma_\phi = \{\phi(t) \in \mathbb{R}^n, t \in [a, b]\} \subset \mathbb{R}^n$$

eines Weges $\phi : [a, b] \rightarrow \mathbb{R}^n$ wird als *Kurve* im \mathbb{R}^n bezeichnet, mit *Parameterdarstellung* ϕ . Im Falle $I = [a, b]$ nennen wir $\phi(a)$ und $\phi(b)$ den *Anfangspunkt* sowie den *Endpunkt* der Kurve. Im Falle $\phi(a) = \phi(b)$ heißt diese *geschlossen*. Ist $\phi(t_1) = \phi(t_2)$ für zwei Parameterwerte $t_1 \neq t_2$, so hat die Kurve dort einen *Doppelpunkt*. Wir werden im Folgenden von einer Kurve als einem geometrischen Objekt bzw. Punktmenge in \mathbb{R}^n sprechen, die durch einen Weg $\phi : [a, b] \rightarrow \mathbb{R}^n$ parametrisiert ist. Dabei können unterschiedliche Parametrisierungen durchaus dieselbe Kurve erzeugen; z.B. gehört zu den Wegen

$$\phi(t) = (\cos(t), \sin(t)), \quad t \in [0, 2\pi],$$

$$\phi(t) = (\cos(t), -\sin(t)), \quad t \in [0, 2\pi],$$

$$\phi(t) = (\cos(2t), \sin(2t)), \quad t \in [0, \pi],$$

als Kurve jeweils die Einheitskreislinie, nur auf unterschiedliche Weise durchlaufen. Häufig hat man für eine geometrische Kurve keine vollständige Parameterdarstellung gegeben, sondern nur für Teilstücke, aus denen die Kurve zusammengesetzt ist. Durch Aneinanderfügung von endlich vielen Kurvenstücken $\Gamma_1, \dots, \Gamma_r$ erhält man eine Kurve; diese heißt geschlossen, wenn der Endpunkt von Γ_r gleich dem Anfangspunkt von Γ_1 ist.

Beispiele Wir geben weitere einfache Beispiele von Wegen und Kurven:

(i) Für eine stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ kann die Abbildung $\phi(t) := (t, f(t))$, $t \in [a, b]$, als Weg im \mathbb{R}^2 aufgefasst werden; die zugehörige Kurve ist der Graph der Funktion f . Wir sprechen in diesem Fall auch von einer *expliziten Parametrisierung* der Kurve durch die Funktion f .

(ii) Der Abbildung $\phi : \mathbb{R} \rightarrow \mathbb{R}^2$ mit $\phi(t) = (a_1 + v_1 t, a_2 + v_2 t) = a + vt$ beschreibt eine Gerade in der Ebene durch den Punkt $a = (a_1, a_2)$ in Richtung des Vektors $v = (v_1, v_2)$.

(iii) Die Abbildung $\phi : \mathbb{R} \rightarrow \mathbb{R}^3$ mit $\phi(t) = (r \cos t, r \sin t, ct)$ beschreibt eine Schraubenlinie im \mathbb{R}^3 .

Sei $\Gamma \subset \mathbb{R}^n$ eine Kurve mit (wie immer als stetig vorausgesetzter) Parametrisierung $\phi : [a, b] \rightarrow \mathbb{R}^n$. Ist ϕ injektiv, so nennen wir die Kurve eine *Jordan-Kurve* und die Parametrisierung einen *Jordan-Weg*. Ist $\phi : [a, b] \rightarrow \mathbb{R}^n$ injektiv mit $\phi(a) = \phi(b)$ als einzigem Doppelpunkt, so sprechen wir von einer *geschlossenen Jordan-Kurve*.

Es scheint anschaulich klar, dass zu einer geschlossenen ebenen Jordan-Kurve immer ein wohl-definiertes, von dieser umschlossenes Gebiet existiert. Dies ist die Aussage des Jordanschen Kurvensatzes. Der Beweis dieser Aussage ist sehr schwierig und kann hier nicht gegeben werden.

Satz [Jordanscher Kurvensatz] Jede geschlossene ebene Jordan-Kurve $\Gamma \subset \mathbb{R}^2$ zerlegt \mathbb{R}^2 in zwei Gebiete, die von ihr berandet werden

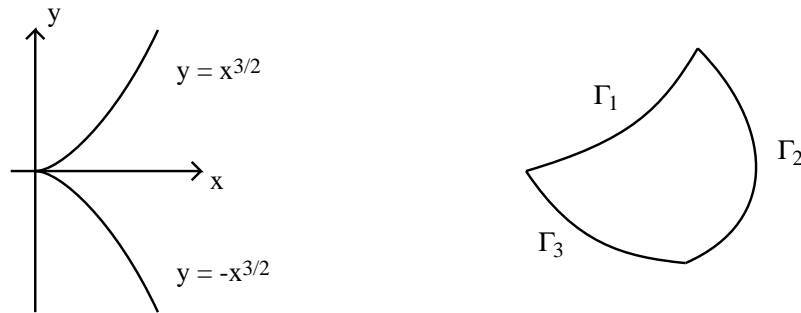
$$\mathbb{R}^2 \setminus \Gamma = G_1 \cup G_2, \quad G_1 \cap G_2 = \emptyset, \quad \partial G_1 = \partial G_2 = \Gamma.$$

Genau eines dieser beiden Gebiete ist beschränkt. Dieses wird als das Innengebiet von Γ bezeichnet.

Eine Kurve heißt *stetig differenzierbar*, wenn für sie eine stetig differenzierbare Parametrisierung $\phi : I \rightarrow \mathbb{R}^n$ (kurz eine C^1 -Parametrisierung) existiert. Eine C^1 -Parametrisierung mit der Eigenschaft $\phi'(t) \neq 0$ für alle $t \in [a, b]$ wird als *regulär* oder *glatt* bezeichnet. Singuläre Parameterwerte $t \in [a, b]$ bzw. die zugehörigen Punkte $\phi(t) \in \mathbb{R}^n$ sind solche mit $\phi'(t) = 0$. Eine Kurve heißt *stückweise differenzierbar*, wenn sie aus endlich vielen differenzierbaren Kurvenstücken besteht (siehe Bild rechts).

Beispiel Die *Neilsche Parabel* hat die Parametrisierung $\phi(t) = (t^2, t^3)$, $t \in \mathbb{R}$. Die zugehörige Kurve ist (siehe Bild links)

$$\Gamma = \{(x, y) \in \mathbb{R}^2 : x \geq 0, y = \pm x^{3/2}\}.$$



Wegen $\phi'(t) = (2t, 3t^2)$ liegt für $t = 0$ in der Spitze der Kurve ein singulärer Punkt. Allgemein sprechen wir von einer *impliziten Darstellung* einer Kurve des \mathbb{R}^2 , wenn $\Gamma = \{(x, y) : f(x, y) = 0\}$ für eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Eine implizite gegebene Kurve genügt nicht genau unserer Definition, weil sie nicht orientiert ist, was aber hier ohne Belang ist. Die Neilsche Parabel lässt sich als Nullstellengebilde der Funktion $f(x, y) = y^2 - x^3$, darstellen. Dies ist gleichzeitig ein Beispiel für die Tatsache, dass eine glatte Funktion ein nichtglattes Nullstellengebilde besitzen kann. Dies erklärt sich aus $\text{grad } f(0, 0) = 0$, so dass der Satz über implizite Funktionen weder für die lokale Auflösung des Nullstellengebildes nach x noch nach y herangezogen werden kann.

Eine injektive Parametrisierung $\phi : [a, b] \rightarrow \mathbb{R}^n$ der Kurve Γ prägt dieser eine Orientierung auf, d.h. eine Reihenfolge, in der ihre Punkte für wachsenden Parameter t erreicht werden. Das folgende Lemma zeigt, dass alle regulären, injektiven Parametrisierungen einer Kurve isomorph sind.

Lemma Seien $\phi : I = [a, b] \rightarrow \mathbb{R}^n$ und $\phi^* : I^* = [a^*, b^*] \rightarrow \mathbb{R}^n$ zwei reguläre, injektive Parametrisierungen derselben Kurve $\Gamma \subset \mathbb{R}^n$. Dann gibt es genau eine bijektive Abbildung $h : I^* \rightarrow I$ derart, dass gilt

$$\phi^* = \phi \circ h.$$

Sind die Parametrisierungen ϕ und ϕ^* (stückweise) stetig differenzierbar, so auch h , und es ist $h' \neq 0$, wo die Ableitung existiert. Also ist h monoton steigend oder monoton fallend; im letzteren Fall durchlaufen die Punkte $\phi^*(t)$, $t \in I^*$, und die Punkte $\phi(t)$, $t \in I$, die Kurve zueinander im entgegengesetzten Sinn, d.h. die zugehörigen Wege haben entgegengesetzte Orientierungen.

Beweis: Da ϕ injektiv und stetig ist, existiert nach Satz 2.4 die Umkehrabbildung ϕ^{-1} und ist stetig. Folglich ist auch die Abbildung $h := \phi^{-1} \circ \phi^* : I^* \rightarrow I$ stetig und bijektiv. Ferner ist

$\phi \circ h = \phi \circ \phi^{-1} \circ \phi^* = \phi^*$. Gilt nun $\phi^* = \phi \circ h^*$ mit einer zweiten bijektiven Abbildung $h^* : I^* \rightarrow I$, so ist $h = \phi^{-1} \circ \phi^* = \phi^{-1} \circ \phi \circ h^*$, d.h. h ist eindeutig bestimmt. Die stetige und bijektive Abbildung $h : I^* \rightarrow I$ ist notwendig streng monoton.

Es bleibt die Differenzierbarkeit von h zu zeigen. Sei $t_0^* \in I^*$ und $t_0 = h(t_0^*) \in I$. Wegen $\phi'(t_0) \neq 0$ muss für eine Komponente $\phi'_j(t_0) \neq 0$ sein. Sei etwa $\phi'_j(t_0) > 0$. Dann ist auch $\phi'_j(t) > 0$ in einer ganzen Intervallumgebung $I_0 \subset I$ von t_0 . Dazu gehört ein Intervall $I_0^* = h^{-1}(I_0) \subset I^*$. In I_0 ist ϕ_j streng monoton wachsend, und die Umkehrfunktion ϕ_j^{-1} ist im Intervall $\phi_j(I_0)$ ebenfalls (stückweise) stetig differenzierbar. Wegen $\phi_j^* = \phi_j \circ h$ ist $h = \phi_j^{-1} \circ \phi_j^*$ in I_0^* . Also ist h in I_0^* (stückweise) stetig differenzierbar. Aus $\phi^* = \phi \circ h$ folgt dann $\phi^{*'} = \phi'(h)h'$ und, wegen $\phi^{*'}, \phi' \neq 0$, schließlich $h' \neq 0$. \square

5.2 Kurvenlänge Zur Definition der Länge einer Kurve als Punktmenge des \mathbb{R}^n beschränken wir uns auf die Teilklasse der Jordan-Kurven mit Parametrisierungen $\phi : I \rightarrow \mathbb{R}^n$ auf kompakten Intervallen $I = [a, b]$. Der Zusatz „Jordan“ wird daher meist weglassen. Wir wollen die Länge einer solchen Kurve Γ des \mathbb{R}^n bestimmen. Sei Z eine Zerlegung des Parameterintervalls $[a, b]$ in Teilpunkte $a = t_0 < t_1 < \dots < t_m = b$, mit Feinheit $|Z| = \max_{k=1, \dots, m} (t_k - t_{k-1})$. Die Menge solcher Zerlegungen wird mit $\mathcal{Z}(a, b)$ bezeichnet. Wir approximieren die Kurve durch einen Polygonzug $p_Z(\Gamma)$, der aus den Strecken durch die Punkte $\phi(t_k)$ besteht. Die Länge $|p_Z(\Gamma)|$ des Polygonzugs ist die Summe der Längen seiner einzelnen linearen Teilstücke, welche wiederum als die euklidischen Abstände ihrer Endpunkte definiert sind, d.h.

$$|p_Z(\Gamma)| = \sum_{k=1}^m |\phi(t_k) - \phi(t_{k-1})|.$$

Die Länge des Polygonzugs nimmt bei Hinzunahme eines weiteren Punktes $t \in (t_{k-1}, t_k)$ wegen $|t_k - t_{k-1}| \leq |t_k - t| + |t - t_{k-1}|$ nicht ab. Hieraus folgt, dass im Falle

$$L = \sup_{Z \in \mathcal{Z}(a, b)} |p_Z(\Gamma)| < \infty$$

für jede Zerlegungsfolge $(Z_k) \subset \mathcal{Z}(a, b)$, mit $|Z_k| \rightarrow 0$ ($k \rightarrow \infty$) die zugehörigen Polygonzuglängen gegen L konvergieren.

Eine Kurve Γ heißt *rektifizierbar* mit der Länge $|\Gamma|$, wenn die Längen aller Polygonzüge $p_Z(\Gamma)$ gleichmäßig beschränkt sind mit

$$|\Gamma| = \sup_{Z \in \mathcal{Z}(a, b)} |p_Z(\Gamma)| = \lim_{Z \in \mathcal{Z}(a, b), |Z| \rightarrow 0} |p_Z(\Gamma)|.$$

Satz Ist die Parameterdarstellung $\phi : [a, b] \rightarrow \mathbb{R}^n$ der Kurve Γ stetig differenzierbar, so ist sie rektifizierbar, und ihre Länge ist gegeben durch

$$|\Gamma| = \int_a^b |\phi'(t)| dt,$$

wobei der Wert des Integrals nicht von der (stetig differenzierbaren) Parametrisierung ϕ der Kurve Γ abhängt.

Beweis: Mit ϕ' ist auch $|\phi'|$ stetig und damit integrierbar. Sei $Z \in \mathcal{Z}(a, b)$ eine Zerlegung. Mit

dem Mittelwertsatz erhalten wir

$$\begin{aligned}
p_Z(\Gamma) &= \sum_{k=1}^m |\phi(t_k) - \phi(t_{k-1})| \\
&= \sum_{k=1}^m (t_k - t_{k-1}) \sqrt{\left(\frac{\phi_1(t_k) - \phi_1(t_{k-1})}{t_k - t_{k-1}}\right)^2 + \dots + \left(\frac{\phi_n(t_k) - \phi_n(t_{k-1})}{t_k - t_{k-1}}\right)^2} \\
&= \sum_{k=1}^m (t_k - t_{k-1}) \sqrt{\phi_1'(\tau_{k,1})^2 + \dots + \phi_n'(\tau_{k,n})^2}
\end{aligned}$$

mit $\tau_{k,i} \in [t_{k-1}, t_k]$. Hier entsteht das kleine Problem, dass die $\tau_{k,i}$ in der Regel verschieden sind. Wegen der gleichmäßigen Stetigkeit von ϕ' auf dem Intervall $[a, b]$ existiert zu beliebig gewähltem $\varepsilon > 0$ ein $h_\varepsilon > 0$, so dass für jede Zerlegung Z mit Feinheit $h \leq h_\varepsilon$ gilt

$$\max_{k=1, \dots, m} |\phi_i'(\tau_{k,1}) - \phi_i'(\tau_{k,i})| < \varepsilon, \quad i = 1, \dots, n.$$

Ferner können wir wegen der gleichmäßigen Stetigkeit von $|\phi'(x)|$ (eventuell nach Verkleinerung von h_ε) auch annehmen, dass

$$\left| \int_a^b |\phi'(t)| dt - \sum_{k=1}^m (t_k - t_{k-1}) |\phi'(\tau_{k,1})| \right| < \varepsilon.$$

Aus den beiden letzten Abschätzungen folgt dann

$$\begin{aligned}
&\left| \int_a^b |\phi'(t)| dt - \sum_{k=1}^m (t_k - t_{k-1}) \sqrt{\phi_1'(\tau_{k,1})^2 + \dots + \phi_n'(\tau_{k,n})^2} \right| \\
&\leq \left| \int_a^b |\phi'(t)| dt - \sum_{k=1}^m (t_k - t_{k-1}) |\phi'(\tau_{k,1})| \right| \\
&\quad + \sum_{k=1}^m (t_k - t_{k-1}) \left(|\phi'(\tau_{k,1})| - \sqrt{\phi_1'(\tau_{k,1})^2 + \dots + \phi_n'(\tau_{k,n})^2} \right) \\
&< \varepsilon + \sum_{k=1}^m (t_k - t_{k-1}) \sum_{i=1}^n |\phi'(\tau_{k,1}) - \phi'(\tau_{k,i})| \leq \varepsilon + |b - a|n\varepsilon.
\end{aligned}$$

Damit ist die Konvergenz

$$|p_Z(\Gamma)| \rightarrow \int_a^b |\phi'(t)| dt$$

bewiesen.

Wir zeigen nun die Unabhängigkeit der Kurvenlänge von der Wahl der Parametrisierung. Sei $\psi : [c, d] \rightarrow \mathbb{R}^n$ eine zweite injektive, stetig differenzierbare Parametrisierung der Kurve Γ . Dann ist nach Lemma 5.1 die Abbildung

$$h = \psi^{-1} \circ \phi : [a, b] \rightarrow [c, d], \quad h(a) = c, \quad h(b) = d,$$

bijektiv und stetig differenzierbar. Mit der Substitutionsregel gilt dann wegen $h' \neq 0$

$$\begin{aligned}
\int_a^b |\phi'(t)| dt &= \int_a^b |(\psi \circ h)'| dt = \int_a^b |\psi'(h(t))h'(t)| dt \\
&= \int_a^b |\psi(h(t))| |h'(t)| dt = \int_c^d |\psi'(s)| ds.
\end{aligned}$$

□

Ist die Kurve aus endlich vielen Stücken Γ_j , $j = 1, \dots, N$ zusammengesetzt, wobei die Γ_j regulär parametrisiert sind, so gilt natürlich

$$|\Gamma| = \sum_{j=1}^N |\Gamma_j| = \sum_{j=1}^N \int_{t_{j-1}}^{t_j} |\phi'(t)| dt,$$

wobei die Γ_j an den Punkten t_1, \dots, t_{n-1} miteinander verknüpft sind.

5.3 Geometrie der Kurven Im Folgenden wollen wir einige geometrische Eigenschaften von Kurven studieren. Da wir uns dabei der Parametrisierung $\phi : [a, b] \rightarrow \mathbb{R}^n$ der Kurve bedienen, müssen wir darauf achten, inwieweit die gefundenen Eigenschaften eventuell von der gewählten Parametrisierung abhängen.

Sei Γ eine stetig differenzierbare Jordan-Kurve mit Parametrisierung $\phi : [a, b] \rightarrow \mathbb{R}^n$. Für $t_0 \in [a, b]$ wird der Vektor $\phi'(t_0)$ *Tangentenvektor* an die Kurve Γ im Punkt $\phi(t_0)$ und die Gerade durch $\phi(t_0)$ in Richtung $\phi'(t_0)$ *Tangente* genannt. Falls $\phi'(t_0) \neq 0$ ist, ist der Tangenten-Einheitsvektor gegeben durch

$$\tau(t_0) = |\phi(t_0)|^{-1} \phi'(t_0).$$

Der Tangentenvektor an eine stetig differenzierbare Jordan-Kurve lässt sich als Limes von Sekantenvektoren auffassen

$$\phi'(t_0) = \lim_{h \rightarrow 0} \frac{\phi(t_0 + h) - \phi(t_0)}{h}.$$

Ist $\phi^* : [a, b] \rightarrow \mathbb{R}^n$ eine zweite Parametrisierung der Kurve Γ , so gibt es nach Lemma 5.1 einen Diffeomorphismus $h : [a^*, b^*] \rightarrow [a, b]$ derart, dass $\phi^*(t^*) = \phi(h(t^*))$, $t^* \in [a^*, b^*]$. Nach der Kettenregel gilt dann, $\phi^{*\prime}(t_0^*) = \phi'(h(t_0^*))h'(t_0^*)$, d.h. die durch die beiden Parametrisierungen ϕ und ϕ^* der Kurve Γ im Punkt $\phi(t_0) = \phi^*(t_0^*)$ erzeugten Tangenten stimmen überein.

Parametrisierung mit der Bogenlänge Sei Γ eine rektifizierbare Kurve mit regulärer Parametrisierung $\phi : [a, b] \rightarrow \mathbb{R}^n$. Für $t \in [a, b]$ bezeichnen wir mit $\Gamma_{[a,t]}$ die Teilkurve mit der Parametrisierung $\phi|_{[a,t]}$. Die Bogenlängenfunktion

$$s(t) = |\Gamma_{[a,t]}| = \int_a^t |\phi'(\tau)| d\tau, \quad t \in [a, b],$$

ist stetig und wegen $|\phi'(t)| > 0$ streng monoton wachsend. Die Integraldarstellung gilt auch, wenn die Parametrisierung ϕ nur stückweise stetig differenzierbar ist. In diesem Sinne ist dann auch $s(t)$ stückweise stetig differenzierbar mit der (stückweisen) Ableitung

$$s'(t) = |\phi'(t)| > 0, \quad t \in [a, b].$$

Also existiert die Umkehrung $t(s)$ mit denselben Eigenschaften. Die Parametrisierung

$$\psi(s) = \phi(t(s)), \quad s \in [0, |\Gamma|]$$

beschreibt ebenfalls die Kurve Γ ; sie wird als die Parametrisierung von Γ mit der Bogenlänge bezeichnet.

Lemma Für die Parametrisierung $\psi[0, |\Gamma|] \rightarrow \mathbb{R}^n$ einer Kurve mit der Bogenlänge gilt $|\psi'(s)| = 1$.

Beweis:

$$\psi'(s) = \phi'(t) \frac{d}{ds} t(s) = \phi'(t) \left(\frac{d}{dt} s(t) \right)^{-1} = \phi'(t) |\phi'(t)|^{-1}$$

und damit $|\psi'(s)| = 1$. \square

Die zur Parametrisierung mit der Bogenlänge gehörende Bogenlängenfunktion ist

$$|\Gamma_{[0,s]}| = \int_0^s |\psi'(\tau)| d\tau = s,$$

was nach Konstruktion zu erwarten war. Die Parametrisierung einer Kurve mit der Bogenlänge ist die natürlichste unter vielen anderen Parametrisierungsmöglichkeiten; sie bietet manchmal theoretische Vorteile, ist aber praktisch kaum nutzbar. Der Parameter s hat eine nur von der Kurve und der ihr aufgeprägten Orientierung abhängige Bedeutung, und die Darstellung $\psi(s)$ ist abgesehen von der Orientierung wegen des Ausschlusses von Doppelpunkten eindeutig bestimmt (bei geschlossenen Kurven bis auf den Anfangspunkt).

Beispiel In der von zwei zueinander orthonormalen Einheitsvektoren $\eta, \xi \in \mathbb{R}^n$ aufgespannten Ebene wird der Kreis $K_r(0) \subset \mathbb{R}^n$ mit Mittelpunkt im Ursprung und Radius $r > 0$ betrachtet. Dieser besitzt die Parametrisierung

$$\phi(t) = r(\eta \cos t + \xi \sin t), \quad t \in [0, 2\pi].$$

Durch Übergang zur Bogenlänge ergibt sich die Darstellung

$$(5.1) \quad \psi(s) = r\left(\eta \cos \frac{s}{r} + \xi \sin \frac{s}{r}\right), \quad s \in [0, 2\pi r].$$

Krümmung einer Kurve Verläuft eine Kurve in der Umgebung eines ihrer Punkte nicht geradlinig, so spricht man von einer gekrümmten Kurve. Um ein Maß für die Krümmung einer Kurve in jedem ihrer Punkte zu gewinnen, gehen wir von ihrer natürlichen Parametrisierung mit der Bogenlänge aus. Die Definition soll mit der Vorstellung verträglich sein, dass eine Gerade die Krümmung Null und ein Kreis eine konstante Krümmung hat, und dass bei Kreisen mit zunehmendem Radius die Krümmung abnimmt.

Für eine Kurve $\Gamma \subset \mathbb{R}^n$ mit zweimal stetig differenzierbarer Parametrisierung $\psi : [0, |\Gamma|] \rightarrow \mathbb{R}^n$ mit der Bogenlänge ist die *Krümmung* definiert durch

$$\kappa(s) = |\psi''(s)|, \quad s \in [0, |\Gamma|].$$

Beispiel Die Kreiskurve in der (x, y) -Ebene um den Nullpunkt mit Radius r hat die Länge $2\pi r$. Ihre Parametrisierung mit der Bogenlänge lautet also

$$x(s) = r \cos(s/r), \quad y(s) = r \sin(s/r), \quad s \in [0, 2\pi r].$$

Die zugehörige Formel für die Krümmung ist

$$\kappa(s) = \sqrt{x''(s)^2 + y''(s)^2} = \frac{1}{r}.$$

Die Krümmung der Kreiskurve ist also konstant $\kappa = 1/r$ und wächst mit abnehmendem Radius.

Differentiation der Beziehung $1 = \psi'(s) \cdot \psi'(s)$ ergibt $\psi''(s) \cdot \psi'(s) = 0$, d.h.: Der *Krümmungsvektor* $\psi''(s)$ steht senkrecht zum Tangentenvektor $\psi'(s)$. Im Fall $\psi''(s) \neq 0$ heißt der Einheitsvektor

$$\nu(s) = \frac{\psi''(s)}{|\psi''(s)|}$$

Hauptnormalvektor zur Kurve Γ im Punkt $\psi(t)$.

Beispiel Als kleine Übung wollen wir den Krümmungsvektor für eine reguläre C^2 -Parametrisierung

$\phi : [a, b] \rightarrow \mathbb{R}^n$ bestimmen. Wir schreiben $\phi' = \frac{d}{dt}\phi$. Dann gilt aufgrund der Definition der Bogenlänge

$$s'(t) = |\phi'(t)| \quad \Rightarrow \quad \frac{d}{ds}t(s) = |\phi'(t(s))|^{-1}.$$

Wegen $|\frac{d}{ds}\psi| = 1$ muss die Ableitung von ψ nach der Bogenlänge mit dem Tangenteneinheitsvektor übereinstimmen

$$\tau(t) = \frac{\phi'(t)}{|\phi'(t)|} = \frac{d}{ds}\psi(s(t)).$$

Ferner benötigen wir

$$\begin{aligned} \frac{d}{ds}|\phi'(t(s))| &= \frac{d}{ds}\sqrt{|\phi'_1(t(s))|^2 + \dots + |\phi'_n(t(s))|^2} \\ &= \frac{1}{2}|\phi'|^{-1} \sum_{j=1}^n 2\phi'_j\phi''_j \frac{d}{ds}t(s) = |\phi'|^{-2}(\phi', \phi'') \end{aligned}$$

Mit diesen Beziehungen folgt dann mit Produkt- und Kettenregel

$$\frac{d}{ds}\tau(t(s)) = \frac{d}{ds} \frac{\phi'(t(s))}{|\phi'(t(s))|} = \frac{\phi''|\phi'|^{-1}}{|\phi'|} - \frac{\phi'|\phi'|^{-2}(\phi', \phi'')}{|\phi'|^2},$$

daher also

$$\frac{d^2}{ds^2}\psi(s) = \frac{d}{ds}\tau(t(s)) = \frac{\phi''|\phi'|^2 - \phi'(\phi', \phi'')}{|\phi'|^4}.$$

Die Größen

$$r(s) = \frac{1}{|\psi''(s)|}, \quad m(s) = \psi(s) + \frac{\psi''(s)}{|\psi''(s)|^2}$$

werden *Krümmungsradius* sowie *Krümmungsmittelpunkt* genannt. Diese Bezeichnung ist dadurch motiviert, dass der Krümmungskreis $K(s)$, d.h. der Kreis in der von Tangenten- und Hauptnormalenvektor $\tau(s)$ bzw. $\nu(s)$ aufgespannten Ebene mit Mittelpunkt $m(s)$ und Radius $r(s)$, die Kurve Γ im Punkt $\psi(s)$ berührt und dort dieselbe Tangente und Krümmung besitzt. Um dies zu sehen, sei ein Wert s_0 fixiert und $r_0 = r(s_0)$, $m_0 = m(s_0)$, $\tau_0 = \tau(s_0)$, $\nu_0 = \nu(s_0)$ gesetzt. Die Parameterdarstellung des zugehörigen Krümmungskreises mit der Bogenlänge lautet nach (5.1)

$$\omega(s) = m_0 + r_0 \left(-\nu_0 \cos \frac{s-s_0}{r_0} + \tau_0 \sin \frac{s-s_0}{r_0} \right), \quad s \in [0, 2\pi r_0].$$

Für $s = s_0$ ist $\omega(s_0) = m_0 - r_0\nu_0 = \psi(s_0)$, d.h. der Krümmungskreis berührt die Kurve im Punkt $\psi(s_0)$. Weiter gilt

$$\begin{aligned} \omega'(s) &= \nu_0 \sin \frac{s-s_0}{r_0} + \tau_0 \cos \frac{s-s_0}{r_0} \Rightarrow \omega'(s_0) = \tau_0 = \psi'(s_0), \\ \omega''(s) &= \frac{\nu_0}{r_0} \cos \frac{s-s_0}{r_0} - \frac{\tau_0}{r_0} \sin \frac{s-s_0}{r_0} \Rightarrow \omega''(s_0) = \frac{\nu_0}{r_0} = \psi''(s_0). \end{aligned}$$

Damit haben Krümmungskreis und Kurve Γ in $\psi(s_0)$ die gleiche Tangente und Krümmung.

5.4 Kurvenintegrale Sei Γ eine rektifizierbare Jordan-Kurve mit einer C^1 -Parametrisierung $\phi : [a, b] \rightarrow \mathbb{R}^n$ mit Bogenlängenfunktion $s(t)$. Für eine Funktion $f : \Gamma \rightarrow \mathbb{R}$ wird das Integral

$$\int_{\Gamma} f(x(s)) ds := \int_a^b f(\phi(t)) |\phi'(t)| dt,$$

sofern es existiert, das *Kurvenintegral* von f über Γ genannt. Man bezeichnet $ds = |\phi'(t)| dt$ als *Streckenelement* entlang der Kurve Γ .

Satz Ist f entlang Γ stückweise stetig, so existiert das Kurvenintegral und hat für alle C^1 -Parametrisierungen denselben Wert.

Beweis: Der Integrand ist stückweise stetig und damit regelintegrierbar.

Sei $\psi : \Gamma \rightarrow \mathbb{R}^n$ die Parametrisierung der Kurve mit der Bogenlänge. Nach Lemma 5.1 gibt es eine Bijektion $h : [0, |\Gamma|] \rightarrow [a, b]$ derart, dass $\phi(t) = \psi(h(t))$. Ferner gilt $\phi' = \psi'(h(t))h'(t)$ und daher wegen $|\psi'(s)| = 1$

$$|\phi'(t)| = |\psi'(h(t))| |h'(t)| = |h'(t)|.$$

Mit Integration durch Substitution folgt dann

$$\int_0^{|\Gamma|} f(x(s)) ds = \int_a^b f(\phi(t)) |h'(t)| dt = \int_a^b f(\phi(t)) |\phi'(t)| dt.$$

□

Für das Kurvenintegral gelten dieselben Regeln wie für das Regelintegral, wobei die Beweise genauso geführt werden können. Die drei wichtigsten Eigenschaften sind:

- Linearität: Für integrierbare Funktionen $f, g : \Gamma \rightarrow \mathbb{R}$ und $\alpha, \beta \in \mathbb{R}$ gilt

$$\int_{\Gamma} (\alpha f(x) + \beta g(x)) ds = \alpha \int_{\Gamma} f(x) ds + \beta \int_{\Gamma} g(x) ds.$$

- Additivität: Für eine disjunkte Zerlegung $\Gamma = \Gamma_1 \cup \Gamma_2$ gilt

$$\int_{\Gamma} f(x) ds = \int_{\Gamma_1} f(x) ds + \int_{\Gamma_2} f(x) ds.$$

f ist genau dann über Γ integrierbar, wenn es über Γ_1 und Γ_2 integrierbar ist.

- Beschränktheit: Ist f über Γ integrierbar, so gilt

$$\left| \int_{\Gamma} f(x) ds \right| \leq |\Gamma| \sup_{x \in \Gamma} |f(x)|.$$

Beispiele (i) Mit $\phi : [a, b] \rightarrow \mathbb{R}^n$ mit $\phi(t) = (t, 0, \dots, 0)$ gilt

$$\int_{\Gamma} f(x) ds = \int_a^b f(t) dt \quad \text{wegen } |\phi'(t)| = 1.$$

Das Kurvenintegral über einem Intervall stimmt mit dem gewöhnlichen Integral überein.

(ii) Ist die Kurve mit einer Masse der Dichte $\rho(x)$ belegt, so ergibt die Gesamtmasse

$$\mu(|\Gamma|) = \int_{\Gamma} \rho(x) ds = \int_a^b \rho(\phi(t)) \phi'(t) dt.$$

(iii) Das Parabelstück

$$\Gamma = \{(x, y) \in \mathbb{R}^2 : y = x^2, 0 \leq x \leq 1\}$$

hat die Parametrisierung $\phi(t) = (t, t^2)$, $0 \leq t \leq 1$. Das Kurvenintegral einer Funktion $f : \Gamma \rightarrow \mathbb{R}$ ist gegeben durch

$$\int_{\Gamma} f(x) ds = \int_0^1 f(x(t)) \sqrt{1 + 4t^2} dt.$$

Für $f = 1$ ergibt sich

$$\begin{aligned} |\Gamma| &= \int_0^1 \sqrt{1 + 4t^2} dt = \frac{1}{2} \int_0^2 \sqrt{1 + u^2} du \\ &= \frac{1}{4} (u\sqrt{1 + u^2} + \ln(u + \sqrt{1 + u^2})) \Big|_0^2 = \frac{1}{4} (2\sqrt{5} + \ln(2 + \sqrt{5})) = 1,46894\dots \end{aligned}$$

5.5 Wegintegrale Hier betrachten wir Wegintegrale von Vektorfeldern, wie sie etwa in der Physik auftreten, z.B. Kraftfelder, elektrische Felder, Geschwindigkeitsfelder und so fort.

Sei $\phi : [a, b] \rightarrow \mathbb{R}^n$ ein differenzierbarer Jordan-Weg und $\Gamma \subset \mathbb{R}^n$ die durch diesen parametrisierte (und orientierte) Jordan-Kurve. Für ein stetiges Vektorfeld $v = (v_1, \dots, v_n) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt

$$\int_{\Gamma} v(x) \cdot ds = \int_a^b v(\phi(t)) \cdot \phi'(t) dt = \sum_{i=1}^n \int_a^b v_i(\phi(t)) \phi'_i(t) dt$$

das *Wegintegral* (oder auch *Zirkulationsintegral* oder *Arbeitsintegral*) von v über Γ . Das Vorzeichen des Wegintegrals hängt dabei von der Orientierung des Weges ab.

Zwischen dem Wegintegral eines Vektorfelds und dem Kurvenintegral einer skalaren Funktion besteht eine enge Beziehung. Der Tangenteneinheitsvektor zur Kurve im Punkt $x \in \Gamma$ ist gegeben durch

$$\tau = \frac{\phi'(t)}{|\phi'(t)|}, \quad x = \phi(t), \quad t \in [a, b].$$

Mit der Tangentialkomponente $v_{\tau} = v \cdot \tau$ eines Vektorfelds $v : \mathbb{R}^n \rightarrow \mathbb{R}^n$ gilt daher

$$\int_{\Gamma} v(x) \cdot ds = \int_a^b v(\phi(t)) \cdot \phi'(t) dt = \int_a^b v_{\tau}(\phi(t)) |\phi'(t)| dt = \int_{\Gamma} v_{\tau} ds,$$

für den Wert des Wegintegrals „zählt“ daher nur die Tangentialkomponente des Vektorfelds. Gleichzeitig haben wir damit gezeigt, dass bei (stückweise) regulärer Parametrisierung der Wert des Wegintegrals nicht von der Parametrisierung, sondern nur von der Orientierung abhängt.

Das Wegintegral hat wie das Kurvenintegral die folgenden Eigenschaften:

- Linearität: Für integrierbare Vektorfelder $v, w : \Gamma \rightarrow \mathbb{R}^n$ und $\alpha, \beta \in \mathbb{R}$ gilt

$$\int_{\Gamma} (\alpha v(x) + \beta w(x)) \cdot ds = \alpha \int_{\Gamma} v(x) \cdot ds + \beta \int_{\Gamma} w(x) \cdot ds.$$

- Additivität: Für eine disjunkte Zerlegung $\Gamma = \Gamma_1 \cup \Gamma_2$ gilt

$$\int_{\Gamma} v(x) \cdot ds = \int_{\Gamma_1} v(x) \cdot ds + \int_{\Gamma_2} v(x) \cdot ds.$$

v ist genau dann über Γ integrierbar, wenn es über Γ_1 und Γ_2 integrierbar ist.

- Beschränktheit: Für beschränkte Vektorfelder gilt

$$\left| \int_{\Gamma} v(x) \cdot ds \right| \leq |\Gamma| \sup_{x \in \Gamma} |v(x)|.$$

Ist v ein Kraftfeld, so bedeutet das Wegintegral $\int_{\Gamma} v(x) \cdot ds$ über eine Kurve die Arbeit, welche von dem Kraftfeld an einem Massepunkt entlang der Kurve Γ geleistet wird. Wie wir bereits festgestellt haben, ist das Wegintegral unabhängig von der Parametrisierung der Kurve gleicher Orientierung, insbesondere unabhängig von der Geschwindigkeit, mit der das Kraftfeld durchlaufen wird. Für ein konstantes Feld $v = \hat{v}$ gilt

$$\int_{\Gamma} v(x) \cdot ds = \sum_{i=1}^n \hat{v}_i \int_a^b \phi'_i(t) dt = \hat{v}(\phi(b) - \phi(a)).$$

In diesem Fall hängt das Wegintegral nur von den Endpunkten der Kurve ab. Als Beispiel betrachten wir das Schwerfeld an der Erdoberfläche, das wir bei kurzen Wegen als konstant $v = (0, 0, -g)$ ansehen können,

$$\int_{\Gamma} v(x) \cdot ds = - \int_a^b g \phi'_3(t) dt = g(\phi_3(a) - \phi_3(b)).$$

Die durch das Schwerfeld geleistete Arbeit hängt nur von der Höhendifferenz ab.

Im Folgenden untersuchen wir, für welche Felder das Wegintegral nur von den Endpunkten der Kurve abhängt.

Ein Vektorfeld $v : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt *Gradientenfeld*, wenn es Gradient einer Funktion $f : D \rightarrow \mathbb{R}$ ist, also

$$v = \text{grad } f = \nabla f.$$

In diesem Fall wird f *Stammfunktion* von v genannt. Im physikalischen Kontext wird v auch als *Potentialfeld* und $U = f$ als *Potential* bezeichnet.

Hat ein Vektorfeld v eine Stammfunktion $f \in C^2(D)$, also $v_i = \partial_i f$, so folgt aus dem Satz von Schwarz, dass $\partial_j v_i = \partial_i v_j$ für alle $1 \leq i, j \leq n$. Dies ist natürlich ein spezieller Fall und im Allgemeinen gar nicht erfüllt. Im \mathbb{R}^2 gilt beispielsweise für $v = (0, x)$, dass $\partial_x v_2 = 1$, aber $\partial_y v_1 = 0$.

Wir nennen das Wegintegral $\int_{\Gamma} v(x) \cdot ds$ *wegunabhängig*, wenn sein Wert nur von den Endpunkten der Kurve Γ abhängt.

Satz (a) Sei $D \subset \mathbb{R}^n$ ein Gebiet und v ein stetiges Vektorfeld. Für eine beliebige orientierte Jordan-Kurve $\Gamma \subset D$ ist das Wegintegral

$$I_{\Gamma}(v) = \int_{\Gamma} v(x) \cdot ds$$

genau dann wegunabhängig, wenn v ein Gradientenfeld ist.

(b) Ist f eine Stammfunktion von v , so hat das Wegintegral über eine beliebige Kurve Γ mit Anfangspunkt x_a und Endpunkt x_b den Wert

$$I_{\Gamma}(v) = f(x_b) - f(x_a).$$

insbesondere ist für jede geschlossene Kurve $I_{\Gamma}(v) = 0$.

(c) Eine Stammfunktion von v existiere. Dann erhält man ausgehend von einem festen Punkt $a \in D$ durch

$$f(x) = \int_{\Gamma(a,x)} v(y) \cdot ds, \quad x \in D,$$

eine Stammfunktion von v . Dabei ist $\Gamma(a,x)$ eine beliebige Jordan-Kurve in D , die a mit x verbindet.

(d) Die Stammfunktion eines Vektorfelds ist bis auf eine Konstante eindeutig bestimmt.

Beweis: (a) Sei v ein Gradientenfeld mit Stammfunktion f . Für eine beliebige C^1 -Parametrisierung $\phi : [a, b] \rightarrow \mathbb{R}^n$ von Γ gilt nach der Kettenregel

$$\frac{d}{dt} f(\phi(t)) = \sum_{i=1}^n \partial_i f(\phi(t)) \phi'_i(t) = \sum_{i=1}^n v_i(\phi(t)) \phi'_i(t) = v(\phi(t)) \cdot \phi'(t),$$

daher

$$(5.2) \quad \int_{\Gamma} v(x) \cdot ds = \int_a^b v(\phi(t)) \cdot \phi'(t) dt = \int_a^b \frac{d}{dt} f(\phi(t)) dt = f(x_b) - f(x_a).$$

Sei nun umgekehrt das Wegintegral wegunabhängig. Dann ist die Funktion

$$f(x) = \int_{\Gamma(x_a,x)} v(y) \cdot ds, \quad x \in D,$$

wohldefiniert. Für $\xi \in D$ und hinreichend kleinem $h \in \mathbb{R}^n$ gilt

$$(5.3) \quad \int_{\Gamma(x_a,\xi+h)} v(y) \cdot ds = \int_{\Gamma(x_a,\xi)} v(y) \cdot ds + \int_{\Gamma(\xi,\xi+h)} v(y) \cdot ds.$$

Da das Integral wegunabhängig ist, können wir in $\int_{\Gamma_{(\xi, \xi+h)}} v(y) \cdot ds$ einfach die Verbindungsstrecke von ξ und $\xi + h$ nehmen. Diese wird parametrisiert durch $\phi : [0, 1] \rightarrow \mathbb{R}^n$, $\phi(t) = \xi + th$,

$$\int_{\Gamma_{(\xi, \xi+h)}} v(y) \cdot ds = \int_0^1 v(\phi(t)) \cdot h dt = v(\xi)h + o(|h|).$$

Aus (5.3) folgt dann

$$f(\xi + h) = f(\xi) + v(\xi) \cdot h + o(|h|).$$

Damit ist f differenzierbar mit $\nabla f = v$.

(b) folgt aus (5.2).

(c) folgt aus dem Beweis von (a).

(d) Für die Differenz $h = f - g$ zweier Stammfunktionen von v gilt $\nabla h = 0$ in D . Damit ist h konstant. \square

Wir hatten bereits gesehen, dass für die Existenz einer Stammfunktion eines Vektorfeldes $v \in C^1$ notwendigerweise

$$(5.4) \quad \partial_i v_j = \partial_i \partial_j f = \partial_j \partial_i f = \partial_j v_i, \quad i, j = 1, \dots, n$$

gelten muss. Damit kann man im konkreten Fall leicht nachprüfen, ob eine Stammfunktion existieren kann.

Beispiele (i) Für das Vektorfeld $v = (ye^{xy}, xe^{xy} + 2y)$ gilt

$$\partial_x v_y = e^{xy} + xye^{xy} = \partial_y v_x.$$

In der Tat ist $f(x, y) = e^{xy} + y^2$ eine Stammfunktion von v .

(ii) Das durch

$$v(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$$

definierte Vektorfeld $v : \mathbb{R}^2 \setminus \{0\}$ erfüllt zwar die Verträglichkeitsbedingungen (5.4),

$$\partial_y v_x = \frac{-(x^2 + y^2) + 2y^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{(x^2 + y^2) - 2x^2}{(x^2 + y^2)^2} = \partial_x v_y.$$

Das Wegintegral entlang des geschlossenen Einheitskreises Γ mit Parametrisierung $\phi(t) = (\cos t, \sin t)$ ist jedoch nicht Null wegen

$$\int_{\Gamma} v(x, y) \cdot ds = \int_0^{2\pi} v(x, y) \cdot \phi'(t) dt = \int_0^{2\pi} \left(\frac{y \sin t}{x^2 + y^2} + \frac{x \cos t}{x^2 + y^2} \right) dt = \int_0^{2\pi} dt = 2\pi.$$

Die Bedingungen (5.4) reichen daher nicht aus, um auf eine Stammfunktion schließen zu können. Wir sagen, ein Gebiet $D \subset \mathbb{R}^n$ ist einfach zusammenhängend, wenn sich jede geschlossene Jordan-Kurve in stetiger Weise innerhalb von D zu einem Punkt zusammenziehen lässt. Im \mathbb{R}^2 darf ein einfach zusammenhängendes Gebiet keine Löcher haben und im \mathbb{R}^3 ist zum Beispiel ein Torus ausgeschlossen.

Satz Gelten für ein Vektorfeld $v \in C^1(D)^n$ im einfach zusammenhängenden Gebiet D die Bedingungen (5.4), so existiert eine Stammfunktion $f \in C^2(D)$ von v .

Beweis: Sei zunächst die Einheitskugel $B_1(0)$ in D enthalten. Wir setzen für $x \in B_1(0)$ und $\phi(t) = tx$

$$f(x) = \int_0^1 v(\phi(t)) \cdot \phi'(t) dt = \int_0^1 v(tx) \cdot x dt.$$

Wegen $\partial_i v_j = \partial_j v_i$ gilt

$$\frac{\partial}{\partial x_i} (v(tx) \cdot x) = \sum_{j=1}^n \partial_i v_j(tx) x_j t + v_i(tx) = \nabla v_i(tx) \cdot (tx) + v_i(tx)$$

sowie

$$\frac{\partial}{\partial t} (v_i(tx)t) = \sum_{j=1}^n \partial_j v_i(tx) t x_j + v_i(tx) = \nabla v_i(tx) \cdot (tx) + v_i(tx).$$

Damit gilt

$$\begin{aligned} \frac{\partial}{\partial x_i} f(x) &= \int_0^1 \frac{\partial}{\partial x_i} (v(tx) \cdot x) dt = \int_0^1 (\nabla v_i(tx) \cdot (tx) + v_i(tx)) dt \\ &= \int_0^1 \frac{\partial}{\partial t} (v_i(tx)t) dt = t v_i(tx) \Big|_0^1 = v_i(x) \end{aligned}$$

Bei einem allgemeinen Gebiet können wir mit diesem Verfahren auf einer kleinen Kugel eine Stammfunktion konstruieren und diese Konstruktion am Rande dieser Kugel mit den dort bekannten Werten von f fortsetzen. Auf diese Weise erschließen wir immer größere Bereiche von D . Treffen wir dann auf einen bereits erschlossenen Bereich, so müssen die vorhandenen Werte mit den neu zu definierenden Werte übereinstimmen. Andernfalls hätten wir ein nichtverschwindendes Wegintegral über eine geschlossene Kurve, was einen Widerspruch auslösen würde, weil wir die Kurve zu einem Punkt zusammenziehen könnten. \square

Beispiel Ein Kraftfeld $F : D \subset \mathbb{R}^n$ heißt *konservativ*, wenn die von ihm entlang einer Kurve geleistete Arbeit nur von den Endpunkten abhängt. Anders ausgedrückt: Wenn die die entlang einer geschlossenen Kurve geleistete Arbeit Null ist. Nach Satz 5.5 ist ein Kraftfeld konservativ genau dann, wenn es ein Gradientenfeld ist. Als Beispiel betrachten wir die Schwerkraft, welche von einer im Ursprung befindlichen Masse μ erzeugt wird. Nach dem Newtonschen Gesetz hat dieses die Gestalt

$$F(x) = -\gamma \frac{\mu x}{|x|^3}, \quad x \in \mathbb{R}^3 \setminus \{0\},$$

mit der Gravitationskonstante $\gamma > 0$. Es gilt

$$\partial_i \frac{x_j}{|x|^3} = \frac{|x|^3 \delta_{ij} - 3|x|^2 x_j x_i |x|^{-1}}{|x|^6} = \frac{|x|^2 \delta_{ij} - 3x_j x_i}{|x|^5}, \quad i, j = 1, 2, 3.$$

Es gilt daher $\partial_i F_j(x) = \partial_j F_i(x)$ im einfach zusammenhängenden Gebiet $\mathbb{R}^3 \setminus \{0\}$. Das Potential lässt sich leicht errechnen,

$$U(x) = -\gamma \frac{\mu}{|x|}, \quad x \in \mathbb{R}^3 \setminus \{0\}.$$

Damit ist das Schwerefeld wie bekannt konservativ. Da sich mehrere Kraftfelder linear überlagern, erzeugen mehrere Massen μ_i , $i = 1, \dots, N$, in Punkten x_1, \dots, x_N folgendes Kraftfeld mit zugehörigem Potential

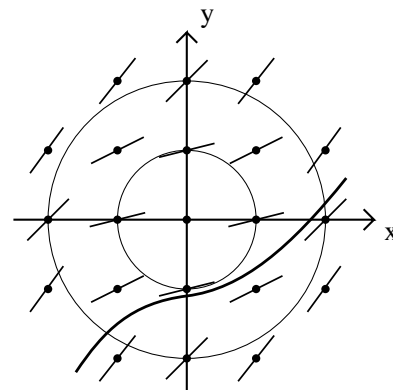
$$F(x) = -\gamma \sum_{i=1}^N \frac{\mu_i (x - x_i)}{|x - x_i|^3}, \quad U(x) = -\gamma \sum_{i=1}^N \frac{\mu_i}{|x - x_i|}$$

6 Gewöhnliche Differentialgleichungen

6.1 Gleichungen erster Ordnung Sei $f : [a, b] \times \mathbb{R} \rightarrow \mathbb{R}$. Gesucht ist eine stetig differenzierbare Funktion $y : [a, b] \rightarrow \mathbb{R}$ mit

$$(6.1) \quad y'(x) = f(x, y(x)) \quad \text{für } x \in [a, b].$$

Wir bezeichnen (6.1) als *gewöhnliche Differentialgleichung*. Die Gleichung ist so zu verstehen, dass für die Lösung im Punkt $y(x)$ die Steigung $y'(x)$ durch $f(x, y(x))$ vorgegeben ist. Man kann eine solche Differentialgleichung graphisch lösen, indem man sich das *Richtungsfeld* in einem Gitter einzeichnet. Die nebenstehende Zeichnung zeigt $f(x, y) = x^2 + y^2$ und eine zugehörige Lösungskurve, die so eingezeichnet wird, dass die Steigung der Tangente der Kurve in jedem Punkt mit der durch f vorgegebenen Steigung übereinstimmt. Aus dieser Zeichnung wird auch klar, dass die Lösungen nicht eindeutig sind, sondern in der Regel eine einparametrische Schar $y(x; c)$ bilden. Man stellt deshalb eine *Anfangswertaufgabe*



$$(6.2) \quad y'(x) = f(x, y(x)) \quad \text{für } x \in [a, b], \quad y(a) = y_0.$$

Verläuft das Richtungsfeld ähnlich wie in der obigen Zeichnung, so wird die Lösung der Anfangswertaufgabe eindeutig bestimmt sein. Allerdings lernen wir später ein Beispiel kennen, in dem auch die Anfangswertaufgabe mehrere Lösungen besitzt.

Beispiele (i) $y(t)$ sei die Populationsdichte einer Art zum Zeitpunkt t , beispielsweise die Anzahl einer Insektenpopulation pro Flächeneinheit. Sind die Nahrungsquellen der Population unbeschränkt, so hängt das Wachstum der Population nur von der Wachstumsrate a ab. Es gilt daher für die Entwicklung der Population die Differentialgleichung $y'(t) = ay(t)$ mit der Anfangswertvorgabe $y(t_0) = y_0$. Durch Einsetzen bestätigt man, dass $y(t) = y_0 \exp(a(t - t_0))$ die Lösung dieses Anfangswertproblems ist. Wir sprechen daher auch vom exponentiellen Wachstum, das allerdings, da ja unsere Welt endlich ist, nur eine gewisse Weile beibehalten werden kann.

(ii) Seien $y_1(t), y_2(t)$ die Konzentrationen zweier Stoffe, die miteinander chemisch reagieren. Die Reaktion $y_1 + y_2 \rightarrow 2y_1$ wird beschrieben durch

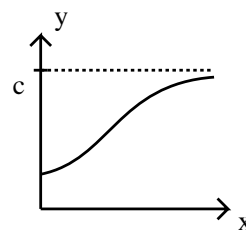
$$y_1 + y_2 = \text{const},$$

$$y_1' = ay_1y_2, \quad y_2' = -ay_1y_2,$$

denn die Wahrscheinlichkeit, dass ein Molekül des einen Stoffes auf ein Molekül des anderen Stoffes trifft und es damit zu einer Umwandlung von y_2 zu y_1 kommt, ist proportional zu y_1y_2 . Addieren wir die beiden Differentialgleichungen, so ergibt sich $y_1' + y_2' = 0$, also $y_1 + y_2 = c = \text{const}$. Wir setzen $y_2 = c - y_1$ in die erste Differentialgleichung ein und erhalten

$$(6.3) \quad y_1' = ay_1(c - y_1), \quad y_1(0) = y_0 \text{ mit } 0 < y_0 < c.$$

In dieser Form kann die Differentialgleichung auch als Populationsmodell interpretiert werden. c ist dann die Kapazität des Lebensraums, also die maximale Populationsdichte, die der Lebensraum gerade noch ernähren kann. Ist $y_1 \ll c$, so verläuft das Wachstum nahezu exponentiell mit Wachstumsrate ac , für $y_1 = c$ ist $y_1' = 0$, y_1 kann daher c nicht



überwinden. Tatsächlich besitzt die Lösungskurve $y_1 = c$ als Asymptote. Ihr charakteristischer S-förmiger Verlauf kann in der Natur häufig beobachtet werden, wenn eine Art einen neuen Lebensraum erobert. Man nennt die Lösungskurve auch *logistische Kurve*, entsprechend heißt (6.3) *logistische Differentialgleichung*.

6.2 Lösungsmethoden für Gleichungen erster Ordnung In vielen Fällen lassen sich Differentialgleichungen mit *getrennten Veränderlichen*, also von der Form $y'(x) = f(x)g(y)$ analytisch lösen. Wir schreiben mit $y' = dy/dx$

$$\frac{dy}{g(y)} = f(x) dx.$$

und erhalten die allgemeine Lösung durch unbestimmte Integration auf beiden Seiten,

$$\int^y \frac{d\xi}{g(\xi)} = \int^x f(\xi) d\xi.$$

Für die Lösung des Anfangswertproblems $y(x_0) = y_0$ folgt aus bestimmter Integration

$$(6.4) \quad \int_{y_0}^y \frac{d\xi}{g(\xi)} = \int_{x_0}^x f(\xi) d\xi.$$

Diese Formeln gelten nur, solange $g(y) \neq 0$ und die Funktion $\phi(y) = \int^y d\xi/g(\xi)$ eindeutig umkehrbar ist. In diesem Fall können wir dieses eher formale Vorgehen auch streng beweisen. Bezeichnen wir eine Stammfunktion von f mit F , so erhalten wir aus (6.4) die Funktion

$$y(x) = \phi^{-1}(F(x) - F(x_0) + \phi(y_0)).$$

Es gilt $y(x_0) = \phi^{-1}(\phi(y_0)) = y_0$, die Anfangsbedingung ist also erfüllt. Mit Kettenregel und Differentiation der Umkehrfunktion gilt

$$\begin{aligned} y'(x) &= \frac{d}{dz} \phi^{-1}(z) \frac{d}{dx} \{ (F(x) - F(x_0) + \phi(y_0)) \} \\ &= \frac{1}{\phi'(y)} f(x) = g(y) f(x). \end{aligned}$$

In der Praxis wendet man dieses Lösungsverfahren besser nur formal an und macht anschließend die Probe.

Beispiele (i) Für das Anfangswertproblem $y' = y^2$, $y(0) = 1$ folgt

$$\int_1^y \frac{d\xi}{\xi^2} = \int_0^x 1 d\xi,$$

daher $-y^{-1} + 1 = x$, also $y(x) = 1/(1-x)$. Da die Probe gelingt, ist dies tatsächlich die Lösung, die allerdings nur für $x < 1$ existiert.

(ii) Für $y' = y^{1/2}$, $y(0) = 0$, folgt $2y^{1/2}(x) = x$ und damit $y(x) = \frac{1}{4}x^2$. Auch hier gelingt die Probe, aber

$$y_s(x) = \begin{cases} 0 & \text{für } 0 \leq x \leq s \\ \frac{1}{4}(x-s)^2 & \text{für } x > s \end{cases}$$

sind für alle $s \geq 0$ ebenfalls Lösungen. Wie wir später sehen werden, kommt eine solche Mehrdeutigkeit nur bei nichtglatten rechten Seiten vor.

(iii) Das Richtungsfeld der Differentialgleichung $y' = e^y \sin x$ ist 2π -periodisch bezüglich x . Mit $y(x)$ sind daher auch $y(x + 2k\pi)$ Lösungen. Durch Trennung der Veränderlichen erhält man

$$\int e^{-y} dy = -e^{-y} = \int \sin x dx = -\cos x - c,$$

daher

$$y(x; c) = -\ln(\cos x + c) \quad \text{falls } \cos x + c > 0.$$

Die Lösung des Anfangswertproblems $y(0) = \eta$ ist daher $y(x) = -\ln(\cos x + e^{-\eta} - 1)$. Ist nun $\eta = -\ln a$, so ist $y(x) = -\ln(\cos x + a - 1)$. Die Lösungen um den Punkt $a = 2$ verhalten sich völlig unterschiedlich: Für $a > 2$ existiert die Lösung für alle $x \in \mathbb{R}$, für $a = 2$ dagegen nur im Bereich $|x| < \pi$.

Für Gleichungen der Form $y' = f(ax + by + c)$, $a, b, c \in \mathbb{R}$, verwendet man die Substitution

$$u(x) = ax + by(x) + c, \quad y'(x) = \frac{1}{b}(u'(x) - a),$$

die auf die Differentialgleichung

$$u' = bf(u) + a$$

mit getrennten Veränderlichen führt.

Beispiel Für $y' = (x + y)^2$ folgt mit $u(x) = y(x) + x$,

$$u' = u^2 + 1, \quad \text{damit } u = \tan(x + c).$$

Die Rücksubstitution liefert damit die allgemeine Lösung $y(x; c) = \tan(x + c) - x$.

Für die *homogene Differentialgleichung* $y' = f(y/x)$ verwendet man die Substitution

$$u(x) = \frac{y(x)}{x} \quad \text{also } y' = u + xu' = f(u),$$

daher ist die Differentialgleichung mit getrennten Veränderlichen

$$u' = \frac{f(u) - u}{x}$$

zu lösen.

Beispiel Für das Anfangswertproblem

$$y' = \frac{y}{x} - \frac{x^2}{y^2}, \quad y(1) = 1,$$

erhalten wir mit $u = y/x$ das Problem für u ,

$$u' = -\frac{1}{xu^2}, \quad u(1) = 1,$$

mit Lösung

$$\int_1^u \xi^2 d\xi = -\int_1^x \frac{d\xi}{\xi}, \quad \text{also } \frac{u^3 - 1}{3} = -\ln x.$$

Die Lösung lautet daher

$$y = x \sqrt[3]{1 - 3 \ln x} \quad \text{für } 0 < x < \sqrt[3]{e}.$$

6.3 Systeme gewöhnlicher Differentialgleichungen Wir gehen nun zu Systemen von Differentialgleichungen über. Zu $f : [a, b] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ gehört das *System 1. Ordnung*

$$y'(x) = f(x, y(x)), \quad a \leq x \leq b,$$

oder ausgeschrieben

$$\begin{aligned} y_1'(x) &= f_1(x, y_1(x), \dots, y_n(x)), \\ &\vdots \\ y_n'(x) &= f_n(x, y_1(x), \dots, y_n(x)), \end{aligned}$$

es sind damit n unbekannte Funktionen gesucht, für die auch n Differentialgleichungen gestellt werden. Damit eine eindeutige Lösung erwartet werden kann, benötigen wir in diesem Fall n Bedingungen an die Lösung. Wir betrachten daher das Anfangswertproblem

$$(6.5) \quad y'(x) = f(x, y(x)) \quad \text{in } [a, b], \quad y(a) = y_0 \in \mathbb{R}^n.$$

Jedes System höherer Ordnung lässt sich als System erster Ordnung schreiben. Beispielsweise führen wir in der Differentialgleichung n -ter Ordnung

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

neue Variablen ein:

$$v_1 = y, \quad v_2 = y', \quad \dots, \quad v_n = y^{(n-1)},$$

und erhalten das äquivalente System

$$\begin{aligned} v_1' &= v_2 \\ &\vdots \\ v_{n-1}' &= v_n \\ v_n' &= f(x, v_1, \dots, v_n) \end{aligned}$$

Unter einer Lösung des Anfangswertproblems (6.5) verstehen wir analog zum Fall einer skalaren Gleichung eine Funktion $y \in C^1([a, b])^n$, für die die Differentialgleichung in jedem Punkt des Intervalls erfüllt ist.

6.4 Existenz von Lösungen Wir ordnen dem Anfangswertproblem (6.5) die Integralgleichung

$$(6.6) \quad y(x) = y_0 + \int_a^x f(\xi, y(\xi)) \, d\xi$$

zu.

Lemma Sei f stetig bezüglich x und y . Dann ist das Anfangswertproblem (6.5) zur Integralgleichung (6.6) äquivalent: Jede Lösung $y \in C^1([a, b])^n$ des Anfangswertproblems (6.5) ist auch eine Lösung $y \in C([a, b])^n$ der Integralgleichung (6.5) und umgekehrt.

Beweis: Die Integralgleichung folgt aus der Differentialgleichung, indem man letztere von a nach b integriert und den Hauptsatz der Differentialrechnung anwendet. Dies ist möglich, weil die auf dem kompakten Intervall $[a, b]$ stetige Funktion y auch beschränkt ist. Da f ebenfalls stetig ist, ist auch die Funktion $f(x, y(x))$ auf $[a, b]$ stetig und beschränkt.

Ist umgekehrt $y \in C([a, b])^n$ eine Lösung der Integralgleichung, so folgt $y(a) = y_0$. Mit y stetig ist auch $f(x, y(x))$ stetig und ihre Stammfunktion stetig differenzierbar. Die Integralgleichung darf also differenziert werden und man erhält die Differentialgleichung (6.5). \square

Wir sagen, f ist *lipschitzstetig bezüglich y* , wenn

$$(6.7) \quad |f(x, y) - f(x, z)| \leq L|y - z| \quad \text{für alle } x \in [a, b], \quad y, z \in \mathbb{R}^n.$$

Nach Lemma 3.3 ist eine in y differenzierbare Funktion lipschitzstetig bezüglich y , wenn $\sup_{(x,y)} |f_y(x, y)|$ endlich ist.

Satz f sei stetig und erfülle eine Lipschitzbedingung bezüglich y . Dann existiert genau eine Lösung des Anfangswertproblems (6.5).

Beweis: Wegen des vorigen Lemmas brauchen wir nur die Existenz einer Lösung der Integralgleichung (6.6) zu zeigen. Nach Satz 3.4 ist $C([a, b])^n$ mit der Norm $\|y\|_\infty = \max_x |y(x)|$ ein Banach Raum, daher ist $C([a, b])^n$ mit der Metrik

$$d(y, z) = \max_{x \in [a, b]} |y(x) - z(x)|$$

ein vollständiger metrischer Raum. Mit Hilfe des Banachschen Fixpunktsatzes zeigen wir nun, dass der Operator

$$Ty(x) = y_0 + \int_a^x f(\xi, y(\xi)) d\xi$$

einen Fixpunkt in $C([a, b])^n$ besitzt, sofern $b - a$ genügend klein ist. Da die Komposition stetiger Funktionen stetig ist, bildet T den Raum $C([a, b])^n$ auf sich selber ab. Die Kontraktionseigenschaft folgt aus

$$\begin{aligned} d(y, z) &= \max_{x \in [a, b]} |Ty(x) - Tz(x)| = \max_{x \in [a, b]} \left| \int_a^x (f(\xi, y(\xi)) - f(\xi, z(\xi))) d\xi \right| \\ &\leq L \max_{x \in [a, b]} \int_a^x |y(\xi) - z(\xi)| d\xi \leq L(b - a) d(y, z), \end{aligned}$$

indem $b - a$ so klein gewählt wird, dass $L(b - a) < 1$.

Mit dem beschriebenen Verfahren können wir das Anfangswertproblem (6.6) auf einem möglicherweise kleinen Intervall $[a, b']$ mit $y(b') = y_1$ lösen. Wir setzen diese Lösung fort, indem wir zum Anfangswertproblem $y' = f(x, y)$ in $[b', b'']$, $y(b') = y_1$ übergehen. Für dieses können wir Existenz auf einem genauso großen Intervall wie zuvor beweisen. Auf diese Weise erhalten wir eine stetige Funktion, die der Integralgleichung auf dem ganzen Intervall $[a, b]$ genügt und daher auch eine Lösung des Anfangswertproblems ist. \square

Satz [Existenz lokaler Lösungen] In einer Umgebung des Punktes $(a, y_0) \in \mathbb{R}^{n+1}$ sei f stetig und stetig differenzierbar bezüglich y . Dann existiert ein $b > a$, so dass das Anfangswertproblem (6.6) im Intervall $[a, b]$ eindeutig lösbar ist.

Beweis: Die Umgebung von $(a, y_0) \in \mathbb{R}^{n+1}$ enthält einen kompakten Zylinder der Form

$$Q = \{(x, y) \in \mathbb{R}^{n+1} : a \leq x \leq b, |y - y_0| \leq K\}.$$

Da f_y stetig auf dieser kompakten Menge ist, ist f lipschitzstetig bezüglich y mit Konstante $L = \max_Q |f_y(x, y)| < \infty$. Wir untersuchen den gleichen Operator wie im Beweis des letzten Satzes auf dem metrischen Raum

$$V = \{y \in C([a, b]) : \|y - y_0\|_\infty \leq K\}.$$

Für $y \in V$ gilt

$$\begin{aligned} \|Ty - y_0\|_\infty &= \max_x \left| \int_a^x (f(\xi, y(\xi)) - f(\xi, y_0)) d\xi + \int_a^x f(\xi, y_0) d\xi \right| \\ &\leq L \max_x \int_a^x |y(\xi) - y_0| d\xi + \max_x \int_a^x |f(\xi, y_0)| d\xi \\ &\leq L(b - a)K + (b - a)\|f(\cdot, y_0)\|_{\infty; [a, b]}. \end{aligned}$$

Wir wählen $b - a > 0$ so klein, dass $\|Ty - y_0\|_\infty \leq K$ erfüllt ist. Dann gilt $T : V \rightarrow V$. Der Rest des Beweises verläuft genauso wie der Beweis des letzten Satzes. \square

6.5 Lineare Systeme Sei $A : [a, b] \rightarrow \mathbb{R}^{n \times n}$ eine stetige matrixwertige Funktion und $d : [a, b] \rightarrow \mathbb{R}^n$ sei ebenfalls stetig. Dann heißt

$$y'(x) = A(x)y(x) + d(x)$$

lineares System. Ist $d = 0$, so heißt das lineare System *homogen*. Ist A konstant, so heißt das System *linear mit konstanten Koeffizienten*.

Satz Das Anfangswertproblem für ein lineares System ist in ganz $[a, b]$ eindeutig lösbar. Die Lösungsmenge eines linearen Systems ist ein n -dimensionaler affin linearer Raum.

Beweis: $f(x, y) = A(x)y + d(x)$ ist offenbar stetig und lipschitz bezüglich y ,

$$|f(x, y) - f(x, z)| = |A(x)(y - z)| \leq \max_{x \in [a, b]} |A(x)| |y - z|.$$

Damit sind alle Voraussetzungen des Existenzsatzes 6.4 erfüllt.

Sei y_{ih} eine beliebige Lösung des inhomogenen Systems. Seien $y_1(x), \dots, y_n(x)$ die Lösungen der Anfangswertprobleme

$$y_i'(x) = A(x)y_i(x), \quad y_i(a) = e_i, \quad i = 1, \dots, n,$$

wobei e_i den i -ten Einheitsvektor bezeichnet. Die Funktionen y_i sind linear unabhängig, weil sie schon im Punkt a linear unabhängig sind. Sei $y(x) = \sum_i c_i y_i(x)$. Dann gilt

$$\begin{aligned} (y(x) + y_{ih}(x))' &= \sum_{i=1}^n c_i y_i'(x) + y_{ih}'(x) = \sum_{i=1}^n c_i A(x)y_i(x) + A(x)y_{ih}(x) + d(x) \\ &= A(x)(y(x) + y_{ih}(x)) + d(x). \end{aligned}$$

Also: Jede Linearkombination der y_i plus y_{ih} ist eine Lösung des Systems. Insbesondere wird das Anfangswertproblem $y(a) = y_0 = \sum_i y_{0,i} e_i$ durch $y = \sum_i y_{0,i} y_i + y_{ih}$ gelöst, wenn wir an y_{ih} die zusätzliche Bedingung $y_{ih}(a) = 0$ stellen. Da das Anfangswertproblem eindeutig lösbar ist, kann jede Lösung mit einer Lösung des Anfangswertproblems identifiziert werden. Der Lösungsraum ist daher

$$y_{ih} + \text{span} \{y_1, \dots, y_n\}.$$

□

Für ein homogenes System liefert dieser Beweis, dass der Lösungsraum ein n -dimensionaler Vektorraum ist. Insbesondere gilt: Sind y_1, y_2 Lösungen des homogenen Systems, so ist für $\alpha, \beta \in \mathbb{R}$ auch $\alpha y_1 + \beta y_2$ eine Lösung des homogenen Systems. Dies nennt man das *Superpositionsprinzip* für die Lösungen des linearen homogenen Systems. Eine Basis des Lösungsraums erhält man wie im obigen Beweis durch Lösen der Anfangswertprobleme $y_i(a) = e_i$.

Nun betrachten wir den Spezialfall eines homogenen Systems mit konstanter Matrix $A \in \mathbb{R}^{n \times n}$. Ist v ein Eigenvektor von A , also $Av = \lambda v$, so ist offenbar $y(x) = e^{\lambda x} v$ eine Lösung des Systems $y' = Ay$ wegen $y'(x) = \lambda e^{\lambda x} v = A(e^{\lambda x} v)$. Ist der Eigenwert λ komplex, so ist auch die zugehörige Lösung komplex. Wir können aus den beiden konjugiert komplexen Eigenwerten λ und $\bar{\lambda}$ auch zwei reellwertige Lösungen definieren, nämlich

$$y_1 = \text{Re}(e^{\lambda x} v) = \frac{1}{2}(e^{\lambda x} v + e^{\bar{\lambda} x} \bar{v}), \quad y_2 = \text{Im}(e^{\lambda x} v) = \frac{1}{2i}(e^{\lambda x} v - e^{\bar{\lambda} x} \bar{v}).$$

Ist die Matrix A diagonalisierbar, so kann der gesamte Lösungsraum mit Hilfe der Eigenwerte und Eigenvektoren konstruiert werden.

Ist A nicht diagonalisierbar, so besagt die Jordansche Normalform, dass die Eigen- und Hauptvektoren eine Basis des \mathbb{C}^n bilden. Wir nennen $v_1, \dots, v_k \in \mathbb{C}^n$ eine *Jordan-Kette* zum Eigenwert λ , wenn

$$Av_1 = \lambda v_1 \quad \text{und} \quad Av_j = \lambda v_j + v_{j-1} \quad \text{für } j = 2, \dots, k.$$

v_1 ist dann der Eigenvektor der Kette, v_2, \dots, v_k sind die Hauptvektoren. Die Jordan-Kette liefert

k linear unabhängige Lösungen des homogenen Systems, nämlich

$$(6.8) \quad \begin{aligned} y_1(x) &= e^{\lambda x} v_1 \\ y_2(x) &= e^{\lambda x} (xv_1 + v_2) \\ &\vdots \\ y_k(x) &= e^{\lambda x} \left(\frac{x^k}{k!} v_1 + \frac{x^{k-1}}{(k-1)!} v_2 + \dots + v_k \right) \end{aligned}$$

Die Lösungen sind schon deshalb linear unabhängig, weil sie ein unterschiedliches Wachstumsverhalten haben. Durch Bestimmung von y'_j und Ay_j weist man unmittelbar nach, dass die y_j Lösungen des homogenen Systems sind.

Ist λ komplex, so sind die Eigen- und Hauptvektoren komplex konjugiert. Bezeichnen wir die komplexen Lösungen mit $y_j^\lambda, y_j^{\bar{\lambda}}$, so bilden die zugehörigen reellen Lösungen

$$(6.9) \quad y_{j,1} = \frac{1}{2}(y_j^\lambda + y_j^{\bar{\lambda}}), \quad y_{j,2} = \frac{1}{2i}(y_j^\lambda - y_j^{\bar{\lambda}})$$

einen Anteil des Lösungsraums der Dimension $2k$. Fassen wir die Ergebnisse zusammen:

Satz Ist A konstant, so bilden die n linear unabhängigen Lösungen in (6.8) bzw. (6.9) eine Basis des Lösungsraums des Systems $y' = Ay$.

Beweis: Der Satz über die Jordansche Normalform besagt, dass jede Matrix genau n linear unabhängige Eigen- und Hauptvektoren besitzt. \square

Wir gehen zu inhomogenen Systemen $y' = A(x)y + d(x)$ über. Seien y_1, \dots, y_n die n linear unabhängigen Lösungen des homogenen Systems $y' = A(x)y$, die wir nach dem letzten Satz zumindest im Fall konstanter Koeffizienten berechnen können. Diese n Lösungen stellen wir zu einer matrixwertigen Funktion

$$Y(x) = [y_1(x) | y_2(x) | \dots | y_n(x)] \in \mathbb{R}^{n \times n}$$

zusammen, die *Fundamentalmatrix* genannt wird. Für jeden Vektor $c \in \mathbb{R}^n$ ist dann das Matrix-Vektor-Produkt $Y(x)c$ eine Lösung des homogenen Systems. Da der Lösungsraum des inhomogenen Systems affin linear ist, benötigen wir nur *eine* Lösung des inhomogenen System und kennen dann den gesamten Lösungsraum.

Satz [Variation der Konstanten] Die Funktion

$$y_{ih}(x) = Y(x) \int_a^x Y(\xi)^{-1} d(\xi) d\xi$$

ist die Lösung des inhomogenen Systems mit $y_{ih}(a) = 0$.

Beweis: Mit der vektorwertigen Funktion

$$c(x) = \int_a^x Y(\xi)^{-1} d(\xi) d\xi, \quad c(x) = (c_1(x), \dots, c_n(x)),$$

gilt $y_{ih}(x) = \sum_j c_j(x) y_j(x)$. Aus dieser Darstellung leitet sich der Name „Variation der Konstanten“ ab, weil in der allgemeinen Lösung des homogenen Problems $y(x) = \sum_i c_i y_i(x)$ die Konstanten c_i

„variiert“, also durch Funktionen $c_i(x)$ ersetzt werden. Es gilt

$$\begin{aligned} y'_{ih}(x) &= \sum_{j=1}^n c'_j(x)y_j(x) + \sum_{j=1}^n c_j(x)y'_j(x) \\ &= \sum_{j=1}^n (Y(x)^{-1}d(x))_j y_j(x) + \sum_{j=1}^n c_j(x)A(x)y_j \\ &= d(x) + \sum_{j=1}^n c_j(x)A(x)y_j(x), \end{aligned}$$

andererseits ist

$$A(x)y_{ih}(x) = \sum_{j=1}^n c_j(x)A(x)y_j(x).$$

□

Beispiele (i) Wir wollen den Lösungsraum der Differentialgleichung

$$y' = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} y + \begin{pmatrix} 0 \\ e^{2x} \end{pmatrix}$$

bestimmen. Das charakteristische Polynom ist

$$\det \begin{pmatrix} 1 - \lambda & 1 \\ 1 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2 - 1 = 0$$

mit den Lösungen $\lambda_1 = 0$ und $\lambda_2 = 2$. Die zugehörigen Eigenvektoren sind $v_1 = (1, -1)$ und $v_2 = (1, 1)$. Die allgemeine Lösung des homogenen Systems ist daher

$$y(x) = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2 e^{2x} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Die Fundamentalmatrix

$$Y = \begin{pmatrix} 1 & e^{2x} \\ -1 & e^{2x} \end{pmatrix}$$

besitzt die Inverse

$$Y^{-1} = \frac{1}{2e^{2x}} \begin{pmatrix} e^{2x} & -e^{2x} \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ e^{-2x} & e^{-2x} \end{pmatrix}.$$

Eine Lösung y_{ih} des inhomogenen Systems erhalten wir mit Hilfe von

$$\int_0^x Y^{-1} \begin{pmatrix} 0 \\ e^{2\xi} \end{pmatrix} d\xi = \frac{1}{2} \int_0^x \begin{pmatrix} -e^{2\xi} \\ 1 \end{pmatrix} d\xi = \frac{1}{4} \begin{pmatrix} -e^{2x} + 1 \\ 2x \end{pmatrix},$$

also

$$y_{ih}(x) = \frac{1}{4}(-e^{2x} + 1) \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \frac{1}{2} x e^{2x} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

(ii) Hier betrachten wir das System

$$y' = \begin{pmatrix} 1 & 2 & 3 \\ -3 & -2 & -1 \\ 1 & 1 & 1 \end{pmatrix} y$$

mit Eigenwerten

$$\det(A - \lambda E) = \det \begin{pmatrix} 1 - \lambda & 2 & 3 \\ -3 & -2 - \lambda & -1 \\ 1 & 1 & 1 - \lambda \end{pmatrix} = -\lambda(\lambda^2 + 1) = 0,$$

also $\lambda_1 = 0$, $\lambda_2 = i$, $\lambda_3 = -i$. Wir bestimmen die zugehörigen Eigenvektoren

$$\lambda_1 = 0: \quad Ax = \begin{pmatrix} x_1 + 2x_2 + 3x_3 \\ -3x_1 - 2x_2 - x_3 \\ x_1 + x_2 + x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow v_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

$$\lambda_2 = i: \quad (A - iE)x = \begin{pmatrix} (1 - i)x_1 + 2x_2 + 3x_3 \\ -3x_1 - (2 + i)x_2 - x_3 \\ x_1 + x_2 + (1 - i)x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow v_2 = \begin{pmatrix} 5 \\ -7 + 4i \\ 3 - i \end{pmatrix}$$

Für $\lambda_3 = -i$ ist $v_3 = \bar{v}_2 = (5, -7 - 4i, 3 + i)$. Damit erhalten wir das komplexe Fundamentalsystem

$$y_1(x) = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad y_2(x) = e^{ix} \begin{pmatrix} 5 \\ -7 + 4i \\ 3 - i \end{pmatrix}, \quad y_3(x) = e^{-ix} \begin{pmatrix} 5 \\ -7 - 4i \\ 3 + i \end{pmatrix},$$

das auch reell geschrieben werden kann,

$$y_{r,1}(x) = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad y_{r,2}(x) = \frac{1}{2}(y_1(x) + y_2(x)), \quad y_{r,3}(x) = \frac{1}{2i}(y_1(x) - y_2(x)),$$

daher

$$y_{r,2}(x) = \cos x \begin{pmatrix} 5 \\ -7 \\ 3 \end{pmatrix} - \sin x \begin{pmatrix} 0 \\ 4 \\ -1 \end{pmatrix}, \quad y_{r,3}(x) = -\sin x \begin{pmatrix} 5 \\ -7 \\ 3 \end{pmatrix} + \cos x \begin{pmatrix} 0 \\ 4 \\ -1 \end{pmatrix}.$$

(iii) Das System $y'(x) = Ay(x)$ mit

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

besitzt wegen $\det(A - \lambda E) = -(\lambda + 2)^3 = 0$ den dreifachen Eigenwert $\lambda = -2$, wegen

$$(A + 2E)v_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ 0 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

aber nur einen Eigenvektor. Wir suchen daher nichttriviale Vektoren v_2, v_3 in der zugehörigen Jordan-Kette

$$(A + 2E)v_2 = v_1 \Rightarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \Rightarrow v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

$$(A + 2E)v_3 = v_2 \Rightarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Das Fundamentalsystem ist daher

$$y_1(x) = e^{-2x}v_1 = e^{-2x} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

$$y_2(x) = e^{-2x}(v_2 + xv_1) = e^{-2x} \begin{pmatrix} x \\ 1 \\ 0 \end{pmatrix},$$

$$y_3(x) = e^{-2x}(v_3 + xv_2 + \frac{1}{2}x^2v_1) = e^{-2x} \begin{pmatrix} x^2/2 \\ x \\ 1 \end{pmatrix},$$

und die Fundamentalmatrix

$$Y(x) = e^{-2x} \begin{pmatrix} 1 & x & x^2/2 \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix}.$$

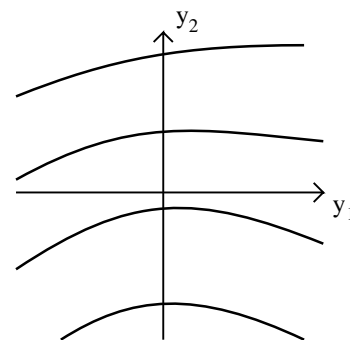
6.6 Dynamische Systeme Sei $D \subset \mathbb{R}^n$ ein Gebiet und $f : D \rightarrow \mathbb{R}^n$ sei von der Klasse C^1 . Wir studieren die Lösungen des Systems gewöhnlicher Differentialgleichungen

$$y'(x) = f(y(x)).$$

Da f nicht von x abhängt, heißt $y' = f(y)$ (*autonomes dynamisches System*). Nach dem lokalen Existenzsatz 6.4 existieren eindeutige Lösungen solange, wie sie in D verbleiben. Aus dieser Tatsache folgt, dass Lösungskurven sich nicht kreuzen oder verzweigen können. Ist nämlich \bar{y} ein Kreuzungs- oder Verzweigungspunkt, so stellen wir die Anfangsbedingung $y(0) = \bar{y}$ und erhalten einen Widerspruch zur eindeutigen Lösbarkeit. Ist $y(x)$ eine Lösung, so ist auch $y(x - x_0)$ eine Lösung. Eine Lösung hängt daher hauptsächlich von ihrem Startpunkt y_0 ab, wir schreiben daher auch $y(x) = y(x, y_0)$.

Dynamische Systeme im \mathbb{R}^2 kann man graphisch darstellen, indem man die Kurve $(y_1(x), y_2(x))$ in ein ebenes Koordinatensystem einzeichnet. Der Verlauf der Lösung ist dann sichtbar, nicht aber die Zeit, die die Lösung braucht.

Die Grundfrage der Theorie dynamischer Systeme lautet: Wie verhalten sich die Lösungen $y(x, y_0)$ für $x \rightarrow \infty$ in Abhängigkeit von y_0 ? Um ein Gefühl für die Möglichkeiten des Langfristverhaltens zu bekommen, untersuchen wir zunächst lineare Systeme im \mathbb{R}^2 .



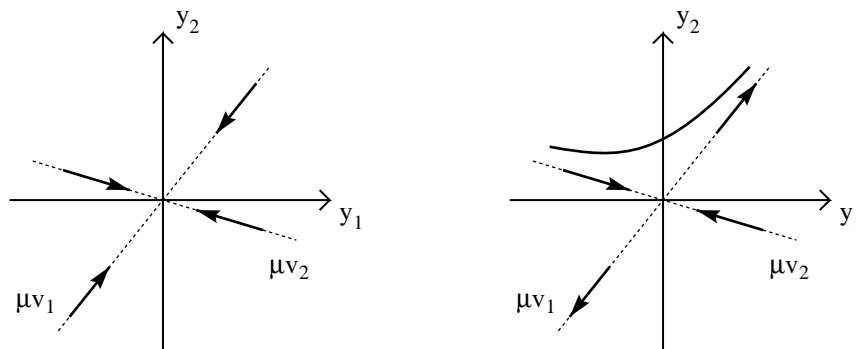
Beispiel Wir betrachten $y' = Ay$ mit einer konstanten 2×2 -Matrix A , die als diagonalisierbar vorausgesetzt wird. Es gibt dann Eigenwerte $\lambda_1, \lambda_2 \in \mathbb{C}$ mit zugehörigen Eigenvektoren $v_1, v_2 \in \mathbb{C}^2$. Nach dem letzten Abschnitt schreibt sich die allgemeine Lösung des Systems in der Form

$$y(x) = c_1 e^{\lambda_1 x} v_1 + c_2 e^{\lambda_2 x} v_2.$$

Wir erhalten die folgenden Fälle:

1. Fall: Beide Eigenwerte sind reell. Sind beide Eigenwerte negativ, entsteht das Bild links. Alle Pfeile des Richtungsfelds zeigen zum Nullpunkt, daher $y(x, y_0) \rightarrow 0$ für alle $y_0 \in \mathbb{R}$. Ist hingegen ein Eigenwert positiv, entsteht das Bild rechts. Wenn y_0 den positiven Eigenvektor enthält, gilt

$y(x, y_0) \rightarrow \infty$. Da jedes Modell die Wirklichkeit nicht exakt abbildet, wird es immer Nebeneffekte geben, die nicht berücksichtigt werden und die die Lösung stören. Auch wenn man die Lösung numerisch berechnet, sind die Rundungsfehler ein dauernder Störfaktor. In einer Situation wie dieser, wo die Lösung nur auf einer dünnen Menge von Anfangswerten gegen den Nullpunkt konvergiert, wird man sowohl in der Natur als auch in der Numerik ein abstoßendes Verhalten des Nullpunkts vorfinden wie die im Bild rechts eingezeichnete Lösung veranschaulicht.

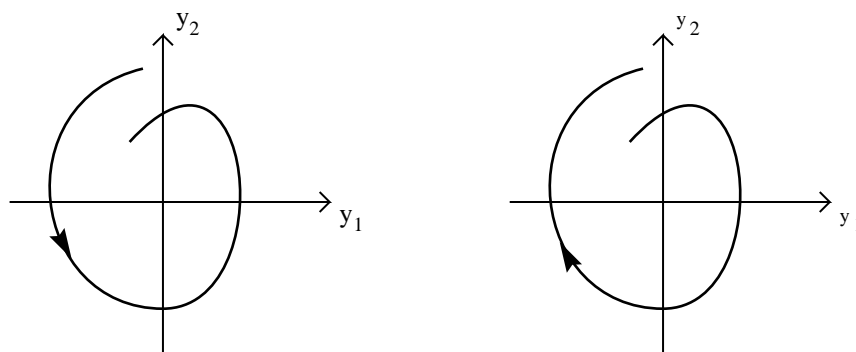


2. Fall: Die Eigenwerte sind komplex konjugiert. Dann bilden die komplexwertigen Funktionen $e^{\lambda x} v$ und $e^{\bar{\lambda} x} \bar{v}$ eine Basis des Lösungsraums. Mit $\lambda = a + ib$, $v = w + iz$ erhalten wir die reellen Lösungen

$$y_1(x) = \frac{1}{2}(e^{\lambda x} v + e^{\bar{\lambda} x} \bar{v}) = \frac{1}{2} e^{ax} (e^{ibx} (w + iz) + e^{-ibx} (w - iz)) = e^{ax} (\cos(bx) w - \sin(bx) z)$$

$$y_2(x) = \frac{1}{2i}(e^{\lambda x} v - e^{\bar{\lambda} x} \bar{v}) = \frac{1}{2i} e^{ax} (e^{ibx} (w + iz) - e^{-ibx} (w - iz)) = e^{ax} (\sin(bx) w + \cos(bx) z)$$

Über das Konvergenzverhalten entscheidet hier der Realteil a von λ . Ist er negativ, so münden die Lösungen spiralförmig in den Nullpunkt ein, was man auch als „Strudelpunkt“ bezeichnet (siehe Bild links). Für $a > 0$ ist der Strudelpunkt abstoßend und für $a = 0$ sind alle Lösungen periodisch.



Fassen wir zusammen: Gilt $\operatorname{Re} \lambda_1, \operatorname{Re} \lambda_2 < 0$, so konvergiert die Lösung für alle Startwerte gegen den Nullpunkt. Besitzt aber einer der Eigenwerte einen positiven Realteil, so konvergiert die Lösung für „fast alle“ Startwerte gegen unendlich.

6.7 Stabile Punkte dynamischer Systeme Wir kehren zum allgemeinen Fall $y' = f(y)$ zurück und nehmen an, dass $\lim_{x \rightarrow \infty} y(x) = \bar{y}$ für einen Punkt $\bar{y} \in D$. Wegen der Stetigkeit von f folgt hieraus $\lim_{x \rightarrow \infty} f(y(x)) = f(\bar{y})$.

Es gibt ein $\varepsilon_0 > 0$, so dass $\tilde{B}_{\varepsilon_0}(\bar{y}) \subset D$. Auf dieser kompakten Menge ist die Funktionalmatrix f_y durch eine Konstante M beschränkt, woraus nach Lemma 3.3 die Lipschitzstetigkeit von f mit Lipschitzkonstante M folgt.

Sei $0 < \varepsilon \leq \varepsilon_0$ beliebig vorgegeben. Dann gibt es ein x_0 mit $|y(x) - \bar{y}| \leq \varepsilon$ für alle $x \geq x_0$. Insbesondere folgt aus der Dreiecksungleichung, dass $|y(x_0 + 1) - y(x_0)| \leq 2\varepsilon$. Wir integrieren

$y' = f(y)$ in jeder Komponente und wenden den Mittelwertsatz der Integralrechnung an

$$y(x_0 + 1) - y(x_0) = \int_{x_0}^{x_0+1} f(y(\xi)) d\xi = \begin{pmatrix} f_1(y(x_1)) \\ \vdots \\ f_n(y(x_n)) \end{pmatrix}$$

mit Punkten $y(x_i)$ mit $x_0 < x_i < x_0 + 1$ für $i = 1, \dots, n$. Hieraus folgt $|f_i(y(x_i))| \leq 2\varepsilon$. Nach Voraussetzung gilt $|y(x_i) - \bar{y}| \leq \varepsilon$. Mit der Dreiecksungleichung und der Lipschitzstetigkeit der f_i folgt dann

$$|f_i(\bar{y})| \leq |f_i(\bar{y}) - f_i(y(x_i))| + |f_i(y(x_i))| \leq M|\bar{y} - y(x_i)| + 2\varepsilon \leq (M + 2)\varepsilon.$$

Da ε beliebig gewählt war, folgt $f(\bar{y}) = 0$. Wenn eine Lösung eines dynamischen Systems konvergiert, so ist der Grenzwert eine Nullstelle von f . Aus der Eindeutigkeit der Lösungen dynamischer Systeme lässt sich noch eine interessante Folgerung ziehen: Ist $y_0 \neq \bar{y}$ und gilt $y(x, y_0) \rightarrow \bar{y}$, so ist $y(x, y_0) \neq \bar{y}$ für alle x , der Grenzwert wird also nie erreicht. Zum Beweis nimmt man das Gegenteil an, dass also $y(x_0, y_0) = \bar{y}$. Man stellt nun die Anfangswertaufgabe mit umgekehrter Zeitrichtung $-y'(x) = f(y(x))$, $y(x_0) = \bar{y}$. Wegen $f(\bar{y}) = 0$ ist diese Lösung konstant, was einen Widerspruch zur Eindeutigkeit bedeutet.

Jedes $\bar{y} \in D$ mit $f(\bar{y}) = 0$ heißt *stationärer Punkt* des dynamischen Systems. Ein stationärer Punkt heißt *asymptotisch stabil*, wenn es eine Umgebung $U(\bar{y})$ gibt mit

$$\lim_{x \rightarrow \infty} y(x, y_0) \rightarrow \bar{y} \quad \text{für alle } y_0 \in U(\bar{y}).$$

\bar{y} heißt *stabil*, wenn es zu jeder Umgebung $U(\bar{y})$ eine Umgebung $V(\bar{y}) \subset U(\bar{y})$ gibt mit $y(x, y_0) \in U(\bar{y})$ für alle $y_0 \in V(\bar{y})$.

Beispiel [Lotka-Volterra-Gleichungen] $x(t)$ und $y(t)$ seien die Populationsdichten zweier Arten, die sich in der Zeit t entwickeln. x ist die Räuberpopulation, die sich ausschließlich von der Beutepopulation y ernährt. Die Lotka-Volterra-Gleichungen sind das einfachste Modell zur Beschreibung der entstehenden Dynamik,

$$\begin{aligned} \dot{x} &= -\alpha_1 x + \beta_1 xy, \\ \dot{y} &= \alpha_2 y - \beta_2 xy, \end{aligned}$$

wobei alle Konstanten positiv sind. Der Räuber verliert Biomasse durch den ersten Term, α_1 ist daher die Todesrate für den Räuber. Bei der Beute wird umgekehrt vorausgesetzt, dass für sie genügend Nahrung vorhanden ist, α_2 ist daher die Geburtsrate für die Beute. Die Wahrscheinlichkeit, dass ein Räuber auf eine Beute stößt ist proportional zu xy , damit ist der Gewinn an Biomasse für den Räuber und der Verlust für die Beute ebenfalls proportional zu xy .

Neben dem trivialen Gleichgewichtspunkt $x = y = 0$ gibt es nur einen weiteren

$$(6.10) \quad \xi = \frac{\alpha_2}{\beta_2}, \quad \eta = \frac{\alpha_1}{\beta_1}$$

Mit $y(t) = y(x(t))$ folgt aus der Kettenregel $\dot{y} = \frac{dy}{dx} \dot{x}$. Für $\dot{x} \neq 0$ erhalten wir daher

$$\frac{dy}{dx} = \frac{\dot{y}}{\dot{x}} = \frac{\alpha_2 y - \beta_2 xy}{-\alpha_1 x + \beta_1 xy} = \frac{\alpha_2 - \beta_2 x}{x} \frac{y}{-\alpha_1 + \beta_1 y}.$$

Damit liegt eine Differentialgleichung mit getrennten Veränderlichen vor. Wir können durch den y -Term teilen und auf beiden Seiten unbestimmt integrieren. Mit einer Integrationskonstanten c folgt dann

$$(6.11) \quad \alpha_1 \ln y - \beta_1 y + \alpha_2 \ln x - \beta_2 x = c.$$

Ist dagegen $\dot{y} \neq 0$, so kann man mit $\frac{dx}{dy}$ genauso verfahren und kommt ebenfalls auf (6.11). (6.11) lässt sich daher als Erhaltungsgleichung interpretieren, was die Beliebtheit der Lotka-Volterra-Gleichungen in der mathematischen Biologie erklärt. Wir wenden die Exponentialfunktion auf (6.11) an und erhalten mit einer neuen Konstanten $c > 0$

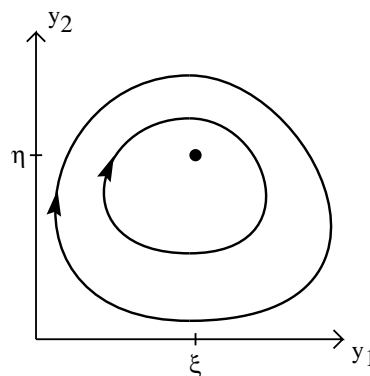
$$(6.12) \quad f(x)g(y) = c \quad \text{mit} \quad f(x) = \frac{x^{\alpha_2}}{e^{\beta_2 x}} \quad \text{und} \quad g(y) = \frac{y^{\alpha_1}}{e^{\beta_1 y}}.$$

Es gilt $f(0) = 0$ und $\lim_{x \rightarrow \infty} f(x) = 0$. Wegen $f \geq 0$ muss der einzige kritische Punkt

$$\left(\frac{x^{\alpha_2}}{e^{\beta_2 x}} \right)' = \frac{\alpha_2 x^{\alpha_2-1}}{e^{\beta_2 x}} - \frac{x^{\alpha_2} \beta_2}{e^{\beta_2 x}} \stackrel{!}{=} 0 \quad \Rightarrow \quad x = \xi = \frac{\alpha_2}{\beta_2}$$

das globale Maximum sein. Entsprechend nimmt die Funktion g ihr globales Maximum im Punkt $\eta = \frac{\alpha_1}{\beta_1}$ an. Aus obiger Darstellung der Ableitung folgt auch, dass beide Funktionen streng monoton steigend bis zum Maximum und streng monoton fallend hinter dem Maximum sind.

Das maximale c in (6.12) wird für die konstante Lösung $x = \xi$, $y = \eta$ angenommen. Für $0 < c < f(\xi)g(\eta)$ gibt (6.12) die Gleichung für die Höhenlinien der Lösung an. Aufgrund der Gestalt von f und g sind alle diese Lösungen periodisch und umrunden den Gleichgewichtspunkt. Damit ist (ξ, η) stabil.



Ist T die Periode einer Lösung, so sind

$$\bar{x} = \frac{1}{T} \int_0^T x(t) dt, \quad \bar{y} = \frac{1}{T} \int_0^T y(t) dt$$

die Durchschnittswerte. Aus den Lotka-Volterra-Gleichungen folgt

$$\frac{d}{dt} \ln y = \alpha_2 - \beta_2 x,$$

daher

$$0 = \frac{1}{T} \int_0^T \frac{d}{dt} \ln y dt = \alpha_2 - \beta_2 \bar{x} \quad \Rightarrow \quad \bar{x} = \frac{\alpha_2}{\beta_2} = \xi.$$

Analog folgt $\bar{y} = \alpha_1/\beta_1 = \eta$. Damit sind die Populationen im Durchschnitt immer gleich groß.

6.8 Ljapunovfunktionen Die Stabilitätseigenschaften stationärer Punkte des autonomen Systems $y' = f(y)$, $f \in C^1(\mathbb{R}^n)^n$, kann in vielen Fällen mit der Methode der Ljapunovfunktionen untersucht werden. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $\bar{x} = 0$ stationärer Punkt des Systems ist, also $f(0) = 0$ erfüllt ist. Ferner nehmen wir an, dass der stationäre Punkt 0 isoliert ist, dass also $f(y) \neq 0$ in einer Umgebung $U(0) \setminus \{0\}$ erfüllt ist.

Eine Funktion $E : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt *Ljapunovfunktion* für das System $y' = f(y)$, wenn sie die folgenden Eigenschaften besitzt:

- (a) E ist stetig differenzierbar.
- (b) E verschwindet im Nullpunkt und ist sonst positiv.
- (c) $\sum_{i=1}^n \partial_i E(y) f_i(y)$ verschwindet im Nullpunkt und ist sonst ≤ 0 .

Gilt in (c) sogar $<$ statt \leq , so heißt E *strikte Ljapunovfunktion*.

In vielen Fällen kann man sich $E(y)$ als Energie des Systems im Zustand y vorstellen. Der Gleichgewichtspunkt ist dann der Zustand minimaler Energie. Für eine Trajektorie $y(x)$ gilt bei Existenz einer Ljapunov-Funktion

$$(6.13) \quad \frac{d}{dx}E(y(x)) = \sum_{i=1}^n \partial_i E(y(x)) y'_i(x) = \sum_{i=1}^n \partial_i E(y(x)) f_i(y(x)) \leq 0,$$

die Energie ist also entlang einer Lösung monoton fallend.

Satz Das System $y' = f(y)$ habe den isolierten stationären Punkt 0 und besitze eine Ljapunovfunktion in Umgebung von 0. Dann ist 0 stabil. Ist die Ljapunovfunktion strikt, so ist 0 asymptotisch stabil.

Beweis: Der Definitionsbereich der Ljapunovfunktion enthalte die Sphäre $S_{\varepsilon_0}^{n-1}$, also die Punkte y mit $|y| = \varepsilon_0$. Für $0 < \varepsilon \leq \varepsilon_0$ sei

$$m = \min_{y \in S_{\varepsilon}^{n-1}} E(y).$$

Dann ist $m > 0$, und wegen der Stetigkeit von E finden wir ein $0 < \delta < \varepsilon$ mit $E(y) < m$ für alle $|y| \leq \delta$. Starten wir nun mit $y(0) = y_0$, $|y_0| \leq \delta$, unser System, so kann die zugehörige Lösung $y(x)$ wegen (6.13) die Kugel $B_{\varepsilon}(0)$ nicht verlassen. Damit ist 0 stabiler Gleichgewichtspunkt.

Sei nun E strenge Ljapunovfunktion. Für Startpunkte $|y_0| \leq \delta$ mit einem $\delta > 0$ aus dem vorigen Beweisschritt verbleibt die zugehörige Lösung $y(x)$ innerhalb des Definitionsbereichs von E . Da (6.13) nun mit der strikten Ungleichung gilt, ist $E(y(x))$ streng monoton fallend und nach unten durch 0 beschränkt. Damit konvergiert $E(y(x))$ gegen ein $\lambda \geq 0$. Angenommen, es gilt $\lambda > 0$. Dann gibt es ein $0 < \rho < \delta$ mit $E(y) < \lambda$ für $y \in B_{\rho}(0)$. Auf der kompakten Menge $K = \{y : \rho \leq |y| \leq \varepsilon_0\}$ gilt

$$\sum_{i=1}^n \partial_i E(y) f_i(y) \leq M < 0.$$

Da die Lösung wegen der Annahme $\lambda > 0$ die Menge $B_{\rho}(0)$ nie erreicht, verbleibt sie für alle $x \geq x_0$ in K . Daher gilt für diese $x \geq x_0$

$$E(y(x)) = E(y(x_0)) + \int_{x_0}^x \frac{d}{dt} E(y(t)) dt \leq E(y(x_0)) + M(x - x_0).$$

Für genügend große x folgt hieraus $E(y(x)) < 0$ und damit ein Widerspruch zu $E(y) \geq 0$. Daher ist $\lambda = 0$ und somit $y(x) \rightarrow 0$. \square

Beispiele (i) Die Bewegung eines Massepunktes x in der Zeit t genüge der Differentialgleichung

$$mx''(t) + rx'(t) + kx(t) = 0$$

mit Konstanten $m, r, k > 0$. Zu dieser Differentialgleichung gehört das System erster Ordnung

$$\begin{aligned} x' &= F(x, v) = v, \\ v' &= G(x, v) = -\frac{k}{m}x - \frac{r}{m}v. \end{aligned}$$

Der einzige stationäre Punkt dieses Systems ist $(0, 0)$. Die Gesamtenergie dieses Systems setzt sich zusammen aus potentieller und kinetischer Energie, also

$$E(x, v) = \frac{1}{2}kx^2 + \frac{1}{2}mv^2.$$

Es gilt $E(x, v) \geq 0$ und $E(x, v) = 0$ nur für $(x, v) = (0, 0)$. Ferner ist

$$\begin{aligned} \frac{d}{dt}E(x(t), v(t)) &= \frac{\partial E}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial E}{\partial v} \frac{\partial v}{\partial t} \\ &= kxv + mv \left(-\frac{k}{m}x - \frac{r}{m}v \right) = -rv^2. \end{aligned}$$

Damit ist E eine Ljapunovfunktion und $(0, 0)$ stabil.

Natürlich bekommt man bei diesem linearen System genauere Auskünfte, wenn man einfach die Eigenwerte der Systemmatrix

$$\frac{1}{m} \begin{pmatrix} 0 & m \\ -k & -r \end{pmatrix}$$

ausrechnet, nämlich $m\lambda_{\pm} = -\frac{r}{2} \pm \sqrt{\frac{r^2}{4} - km}$, also $\operatorname{Re} \lambda_{\pm} < 0$. Damit ist $(0, 0)$ sogar asymptotisch stabil.

(ii) Das autonome System

$$x' = -x^3 + y, \quad y' = -x - y^5$$

hat $(0, 0)$ als einzigen stationären Punkt. Mit dem Ansatz $E(x, y) = ax^2 + by^2$ mit noch zu bestimmenden $a, b > 0$ erhalten wir

$$\begin{aligned} \frac{d}{dt}E(x(t), y(t)) &= 2ax(-x^3 + y) + 2by(-x - y^5) \\ &= -2ax^4 + (2a - 2b)xy - 2by^6. \end{aligned}$$

Demnach gelingt schon mittels $a = b = 1$ der Nachweis, dass $(0, 0)$ asymptotisch stabil und global anziehend ist.

6.9 Die Exponentialfunktion für Matrizen Mit \mathbb{K} bezeichnen wir immer einen der Körper \mathbb{R} oder \mathbb{C} . Für $A \in \mathbb{K}^{n \times n}$ ist die Matrix-Exponentialfunktion definiert durch

$$(6.14) \quad \exp A = e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

Wegen

$$(6.15) \quad \left\| \sum_{k=0}^n \frac{A^k}{k!} \right\| \leq \sum_{k=0}^n \frac{\|A^k\|}{k!} \leq \sum_{k=0}^n \frac{\|A\|^k}{k!} \leq \exp(\|A\|)$$

konvergiert die Reihe absolut und $\exp : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}^{n \times n}$.

Satz (a) Die Reihe (6.14) konvergiert gleichmäßig auf beschränkten Teilmengen von $\mathbb{K}^{n \times n}$.

(b) Die Funktion $\exp : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}^{n \times n}$ ist stetig.

(c) Es gilt $e^0 = I$ und $e^{A+B} = e^A e^B$, sofern $AB = BA$.

(d) Für alle regulären Matrizen $C \in \mathbb{K}^{n \times n}$ gilt $C^{-1}e^A C = e^{C^{-1}AC}$.

(e) Für alle $A \in \mathbb{K}^{n \times n}$ ist e^A invertierbar mit $(e^A)^{-1} = e^{-A}$.

Beweis: (a) Für $\|A\| \leq M$ gilt

$$\left\| \sum_{k=n}^{\infty} \frac{A^k}{k!} \right\| \leq \sum_{k=n}^{\infty} \frac{M^k}{k!} \rightarrow 0 \text{ für } n \rightarrow \infty.$$

(b) Da auf beschränkten Teilmengen gleichmäßige Konvergenz der Reihe vorliegt und A^n stetig ist, ist auch die Grenzfunktion $\exp A$ stetig in A . Damit ist \exp auf ganz $\mathbb{K}^{n \times n}$ stetig.

(c) $e^0 = I$ ist klar. Für Matrizen A, B mit $AB = BA$ können wir die binomische Formel anwenden und erhalten

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

also

$$\begin{aligned} e^{A+B} &= \sum_{n=0}^{\infty} \frac{1}{n!} (A+B)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} A^k B^{n-k} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{A^k}{k!} \frac{B^{n-k}}{(n-k)!} = \sum_{i=0}^{\infty} \frac{A^i}{i!} \sum_{j=0}^{\infty} \frac{B^j}{j!} = e^A e^B. \end{aligned}$$

(d) Für $A \in \mathbb{K}^{n \times n}$ und reguläres C gilt $C^{-1}A^n C = (C^{-1}AC)^n$.

(e) Nach (c) ist $e^A e^{-A} = I$. Also ist e^A invertierbar mit $(e^A)^{-1} = e^{-A}$. \square

Beispiele (i) Für die Diagonalmatrix $D = \text{diag}(d_1, \dots, d_n)$ gilt $D^k = \text{diag}(d_1^k, \dots, d_n^k)$ und daher $e^D = \text{diag}(e^{d_1}, \dots, e^{d_n})$.

(ii) Für die $n \times n$ Matrix

$$N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \quad \text{ist} \quad N^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Daher schiebt die Multiplikation von N mit N^k immer die Einser eine Diagonale nach rechts oben bis man schließlich $N^n = 0$ erhält. Hieraus folgt für $x \in \mathbb{R}$

$$e^{Nx} = \sum_{k=0}^{\infty} \frac{1}{k!} (Nx)^k = \begin{pmatrix} 1 & x & \frac{x^2}{2} & \dots & \frac{x^{n-1}}{(n-1)!} \\ & 1 & x & \ddots & \vdots \\ & & 1 & \ddots & \frac{x^2}{2} \\ & & & \ddots & x \\ 0 & & & & 1 \end{pmatrix}$$

(iii) Für ein Jordan-Kästchen

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \ddots & \ddots \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix} \in \mathbb{K}^{n \times n}$$

gilt $J_n(\lambda) = \lambda I + N$ mit N wie in Beispiel (ii). Da λI und N kommutieren, folgt aus dem obigen Satz (c)

$$e^{J_n(\lambda)} = e^{\lambda} e^N.$$

Die Berechnung von e^A ist also immer auf einfache Weise möglich, wenn man die Jordansche Normalform von A kennt. Ist nämlich $A = T^{-1}JT$ mit $J = \text{diag}(J_{n_1}(\lambda_1), \dots, J_{n_k}(\lambda_k))$, so gilt

$$(6.16) \quad e^{Ax} = T^{-1}e^{Jx}T, \quad e^{Jx} = \text{diag}(e^{J_{n_1}(\lambda_1)x}, \dots, e^{J_{n_k}(\lambda_k)x})$$

mit

$$(6.17) \quad e^{J_n(\lambda)x} = e^{\lambda x} \begin{pmatrix} 1 & x & \frac{x^2}{2} & \cdots & \frac{x^{n-1}}{(n-1)!} \\ & 1 & x & \ddots & \vdots \\ & & 1 & \ddots & \frac{x^2}{2} \\ & & & \ddots & x \\ 0 & & & & 1 \end{pmatrix}$$

Hieraus erhalten wir:

Lemma Ist $\operatorname{Re} \lambda < \alpha$ für alle Eigenwerte λ von A , so gilt

$$\|e^{Ax}\| \leq ce^{\alpha x} \quad \text{für alle } x \geq 0.$$

Beweis: Aus (6.16) folgt

$$\|e^{Ax}\| \leq \|T^{-1}\| \max_{1 \leq i \leq k} \|e^{J_{n_i}(\lambda_i)x}\| \|T\|.$$

Wir brauchen daher nur $\|e^{J_n(\lambda)x}\|$ für einen Eigenwert λ abzuschätzen. Für $0 \leq x < 1$ bleibt e^{Nx} beschränkt und für $x \geq 1$ erhalten wir aus (6.17)

$$\|e^{J_n(\lambda)x}\| \leq c_1 e^{\operatorname{Re} \lambda x} x^{n-1} \leq ce^{\alpha x},$$

weil die e -Funktion jede Potenz erschlägt. \square

Satz Sei $A \in \mathbb{K}^{n \times n}$. Die Funktion $\phi : \mathbb{R} \rightarrow \mathbb{K}^{n \times n}$ mit $\phi(x) = e^{Ax}$ ist auf ganz \mathbb{R} unendlich oft differenzierbar mit

$$\frac{d}{dx} \phi(x) = Ae^{Ax} = e^{Ax} A.$$

Beweis: Es gilt

$$\frac{d}{dx} (Ax)^n = nx^{n-1} A^n = nA(Ax)^{n-1}.$$

Damit konvergiert die gliedweise differenzierte Reihe auf endlichen Abschnitten von \mathbb{R} gleichmäßig gegen $Ae^{Ax} = e^{Ax} A$. Nach einem Satz der Analysis I stellt daher die Exponentialreihe eine differenzierbare Funktion dar, die gliedweise differenziert werden kann. Die Ableitung darf man offenbar weiter differenzieren und man erhält für die n -te Ableitung $\phi^{(n)}(x) = A^n e^{Ax}$. \square

Damit wird das homogene Anfangswertproblem mit konstanten Koeffizienten,

$$y' = Ay, \quad y(x_0) = y_0 \in \mathbb{K}^n$$

gelöst durch

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

Analog lässt sich die Lösung der inhomogenen Differentialgleichung

$$y' = Ay + d(x), \quad y(x_0) = y_0 \in \mathbb{K}^n,$$

für stetiges d eleganter schreiben als wir das im vorigen Abschnitt getan haben. Für x im Definitionsbereich von d gilt nämlich

$$y(x) = e^{A(x-x_0)} y_0 + \int_{x_0}^x e^{A(x-t)} d(t) dt$$

wegen $y(x_0) = y_0$ und

$$\begin{aligned} y'(x) &= \frac{d}{dx} \left(e^{A(x-x_0)} y_0 + e^{Ax} \int_{x_0}^x e^{-At} b(t) dt \right) \\ &= A e^{A(x-x_0)} y_0 + A e^{Ax} \int_{x_0}^x e^{-At} d(t) dt + e^{Ax} e^{-Ax} d(x) \\ &= Ay(x) + d(x). \end{aligned}$$

Beispiel Gesucht ist die Lösung von

$$y_1' = y_2 + 1, \quad y_2' = -y_1 \quad \text{mit } y_1(0) = y_2(0) = 1$$

oder

$$y' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} y + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{mit } y(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Die Eigenwerte von $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ sind i und $-i$ mit Eigenvektoren $(1, -i)^T$ und $(1, i)^T$, also

$$e^{Ax} = \frac{1}{2i} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} e^{ix} & 0 \\ 0 & e^{-ix} \end{pmatrix} \begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix} = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}.$$

Aus obiger Lösungsformel erhalten wir

$$\begin{aligned} y(x) &= \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \int_0^x \begin{pmatrix} \cos(x-t) & \sin(x-t) \\ -\sin(x-t) & \cos(x-t) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} dt \\ &= \begin{pmatrix} \cos x + \sin x \\ \cos x - \sin x \end{pmatrix} + \int_0^x \begin{pmatrix} \cos(x-t) \\ -\sin(x-t) \end{pmatrix} dt \\ &= \begin{pmatrix} \cos x + \sin x \\ \cos x - \sin x \end{pmatrix} + \begin{pmatrix} -\sin(x-t) \\ -\cos(x-t) \end{pmatrix} \Big|_0^x \\ &= \begin{pmatrix} \cos x + \sin x \\ \cos x - \sin x \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} -\sin x \\ -\cos x \end{pmatrix} = \begin{pmatrix} \cos x + 2 \sin x \\ 2 \cos x - \sin x - 1 \end{pmatrix} \end{aligned}$$

6.10 Der Stabilitätssatz **Satz** Sei \bar{y} stationärer Punkt von $y' = f(y)$ und $A = f_y(\bar{y}) \in \mathbb{R}^{n \times n}$ besitze die Eigenwerte $\lambda_1, \dots, \lambda_k$, $k \leq n$. Es sei $\gamma = \max\{\operatorname{Re} \lambda_1, \dots, \operatorname{Re} \lambda_k\}$. Dann gilt:

- (a) Wenn $\gamma < 0$, so ist \bar{y} asymptotisch stabil.
- (b) Wenn $\gamma > 0$, so ist \bar{y} nicht stabil.
- (c) Wenn $\gamma = 0$, so lässt sich keine Aussage über Stabilität oder Instabilität machen.

Beweis: Wir zeigen nur die wichtige Aussage (a).

Wir können $\bar{y} = 0$ annehmen. Für beliebiges y folgt aus der Definition der Differenzierbarkeit im Nullpunkt

$$f(y) = f(0) + Ay + g(y), \quad A = f_y(0), \quad \frac{|g(y)|}{|y|} \rightarrow 0 \quad \text{für } y \rightarrow 0.$$

Es genügt also, das System

$$y' = Ay + g(y), \quad \operatorname{Re} \lambda(A) < 0,$$

in Umgebung des Nullpunkts zu betrachten.

Wir wissen, dass das Systems mit Anfangswert $y(0) = y_0$ eine Lösung auf dem Intervall $[0, M]$ mit einem möglicherweise kleinen $M > 0$ besitzt. Für $0 \leq x \leq M$ gilt dann

$$y(x) = e^{Ax}y_0 + \int_0^x e^{A(x-t)}g(y(t)) dt.$$

Es gibt ein positives α , so dass jeder Eigenwert λ der Matrix A die Bedingung $\operatorname{Re} \lambda < -\alpha$ erfüllt. Nach Lemma 6.9 existiert eine Konstante c_0 , die wir > 1 wählen können, mit

$$\|e^{Ax}\| \leq c_0e^{-\alpha x}, \quad x \geq 0.$$

Hiermit folgt die Abschätzung

$$|y(x)| \leq c_0e^{-\alpha x}|y_0| + c_0 \int_0^x e^{-\alpha(x-t)}|g(y(t))| dt, \quad 0 \leq x \leq M.$$

Wir wählen ein ε_0 mit $0 < \varepsilon_0 < \alpha$. Dann wählen wir ein $\delta > 0$ mit

$$|y| \leq \delta \Rightarrow |g(y)| \leq \frac{\varepsilon_0}{c_0}|y|,$$

was wegen $|g(y)| = o(|y|)$ möglich ist. Als letztes wählen wir ein $\varepsilon > 0$ mit $c_0\varepsilon \leq \delta$.

Wir zeigen nun für Startwerte y_0 mit $|y_0| \leq \varepsilon$, dass $|y(x)| \rightarrow 0$. Wegen $c_0 > 1$ und $|y_0| \leq \varepsilon$ ist $|y(0)| < \delta$ und damit auch $|y(x)| \leq \delta$ für genügend kleine x . Für solche x haben wir dann

$$\begin{aligned} |y(x)| &\leq c_0e^{-\alpha x}|y_0| + c_0 \int_0^x e^{-\alpha(x-t)}|g(y(t))| dt \\ &\leq c_0\varepsilon e^{-\alpha x} + c_0 \int_0^x e^{-\alpha(x-t)} \frac{\varepsilon_0}{c_0}|y(t)| dt \\ &\leq \delta e^{-\alpha x} + \varepsilon_0 \int_0^x e^{-\alpha(x-t)}|y(t)| dt \end{aligned}$$

und daher

$$|e^{\alpha x}y(x)| \leq \delta + \varepsilon_0 \int_0^x |e^{\alpha t}y(t)| dt.$$

Wir setzen $u(x) = |e^{\alpha x}y(x)|$. Dann gilt $u(x) \geq 0$ und

$$u(x) \leq \delta + \varepsilon_0 \int_0^x u(t) dt.$$

Aus dem Gronwallschen Lemma, das gleich bewiesen wird, folgt dann

$$u(x) \leq \delta e^{\varepsilon_0 x},$$

und damit

$$|y(x)| \leq \delta e^{-(\alpha - \varepsilon_0)x}.$$

Wegen $\varepsilon_0 < \alpha$ verbleibt $y(x)$ für alle $x \geq 0$ in der δ -Umgebung des Nullpunkts und $y(x) \rightarrow 0$. \square

Lemma [Gronwallsches Lemma] Seien $\varepsilon, \delta \geq 0$. Gilt für eine stetige Funktion $u \geq 0$ auf dem Intervall $I = [0, a]$

$$u(x) \leq \delta + \varepsilon \int_0^x u(t) dt \quad \text{für alle } x \in I,$$

so folgt die Abschätzung

$$u(x) \leq \delta e^{\varepsilon x} \quad \text{für alle } x \in I.$$

Beweis: Mit $v(x) = \int_0^x u(t) dt$ erhalten wir unter Verwendung von $v \geq 0$

$$\begin{aligned}
 v'(x) &\leq \delta + \varepsilon v(x) \\
 (v(x)e^{-\varepsilon x})' &\leq \delta e^{-\varepsilon x} \\
 \int_0^x (v(t)e^{-\varepsilon t})' dt &\leq \delta \int_0^x e^{-\varepsilon t} dt \\
 (6.18) \quad v(x)e^{-\varepsilon x} &\leq \frac{\delta}{\varepsilon}(1 - e^{-\varepsilon x}).
 \end{aligned}$$

Aus $u(x) \leq \delta + \varepsilon v(x)$ und (6.18) folgt die Behauptung. \square

Beispiele (i) Das autonome System

$$x' = -x^3 + y, \quad y' = -x - y^5$$

hatten wir in Abschnitt 6.8 mit der Methode der Ljapunovfunktionen behandelt und dabei herausgefunden, dass der stationäre Punkt $(0, 0)$ asymptotisch stabil und global anziehend ist. Die rechte Seite lässt sich schreiben in der Form

$$f(x, y) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + g(x, y)$$

mit $g(x, y) = o(|x| + |y|)$. Die Differentialgleichung hat also genau die Form, wie wir sie im Beweis des letzten Satz verwendet haben. Die Eigenwerte der Matrix sind $\pm i$, so dass wir mittels des letzten Satzes keine Aussage über das Stabilitätsverhalten machen können.

(ii) Für das analoge System

$$x' = x^3 + y, \quad y' = -x + y^5$$

erhalten wir mit der gleichen Ljapunovfunktion wie zuvor, $E(x, y) = x^2 + y^2$,

$$\begin{aligned}
 \frac{d}{dt}E(x(t), y(t)) &= 2x(x^3 + y) + 2y(-x + y^5) \\
 &= 2x^4 + (2 - 2)xy + 2y^6 > 0 \quad \text{für } (x, y) \neq 0.
 \end{aligned}$$

Das Prinzip der Ljapunovfunktionen schlägt natürlich auch hier zu: Der Punkt $(0, 0)$ ist noch nicht einmal stabil.

Diese Beispiele zeigen, dass man bei nichtlinearen rechten Seiten $f(y)$ mit $\operatorname{Re} \lambda(f_y(0)) = 0$ keine Aussage über Stabilität oder Instabilität machen kann.

Aufgaben

6.1 Sei $k : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ eine stetige Funktion. Wir betrachten den *volterraschen Integraloperator*

$$Tu(t) = \int_0^1 k(s, t)u(s) ds + u_0$$

für ein $u_0 \in \mathbb{R}$.

a) Zeigen Sie $T : C([0, 1]) \rightarrow C([0, 1])$ ist lipschitzstetig.

b) Formulieren Sie eine Bedingung an k , so dass das Problem $Tu = u$ in $C([0, 1])$ eine eindeutige Lösung besitzt.

6.2 Bestimmen Sie die Lösung der folgenden Anfangswertprobleme:

a) $y' = \frac{3y^3}{x^7}, y(1) = 1,$

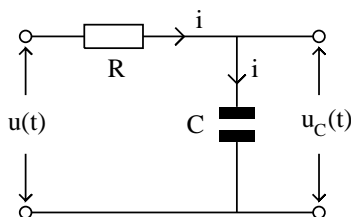
b) $y' = 3x^2y, y(0) = 1.$

6.3 Man bestimme die allgemeine Lösung der homogenen linearen Differentialgleichung

$$y'(x) = a(x)y(x)$$

mit Hilfe von $A(x) = \int a(\xi) d\xi.$

6.4 Ein RC-Glied besteht aus einem Widerstand R und einem in Serie geschalteten Kondensator



C . Gesucht ist die Spannung am Kondensator $u_C(t)$ in Abhängigkeit der Eingangsspannung $u(t)$. Der durch R fließende Strom $i = (u - u_C)/R$ ist der Ladestrom des Kondensators,

$$i = C \frac{du_C}{dt},$$

was auf die lineare Differentialgleichung

$$(*) \quad \frac{du_C}{dt} + \frac{1}{T} u_C = \frac{1}{T} u(t)$$

führt, wobei $T = RC$ die *Zeitkonstante* des RC-Glieds bezeichnet.

a) Man eliminiere T aus der Gleichung $(*)$ durch eine lineare Transformation. In den übrigen Aufgabenteilen kann daher mit $T = 1$ gerechnet werden.

b) Man bestimme die Lösung des Anfangswertproblems $u_C(0) = u_0$ und konstantem u (=Ladevorgang des Kondensators).

c) Man bestimme die allgemeine Lösung von $(*)$ für $u(t) = \sin t$.

6.5 Man bestimme die allgemeine Lösung des Systems

$$y'(x) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} y(x) - \begin{pmatrix} e^{-x} \\ 2e^{-x} \end{pmatrix}.$$

6.6 Sei

$$Ly = y^n + a_{n-1}y^{n-1} + \dots + a_0y$$

der allgemeine lineare Differentialoperator n -ter Ordnung mit konstanten Koeffizienten. Wir können diesen als System erster Ordnung schreiben und wie oben angegeben ein Fundamentalsystem konstruieren. Hier soll ein einfacherer Lösungsweg besprochen werden. Wir ordnen L das *charakteristische Polynom*

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0$$

zu, das nach dem Fundamentalsatz der Algebra l Nullstellen besitzt, deren Vielfache sich zu n summieren. Man zeige: Ist λ eine k -fache Nullstelle, so sind

$$y_1(x) = e^{\lambda x}$$

$$y_2(x) = e^{\lambda x}(x + 1)$$

\vdots

$$y_k(x) = e^{\lambda x} \left(\frac{x^k}{k!} + \frac{x^{k-1}}{(k-1)!} + \dots + 1 \right)$$

linear unabhängige Lösungen der homogenen Gleichung $Ly = 0$.

Bemerkung: Bei einer Gleichung n -ter Ordnung entspricht eine mehrfache Nullstelle immer dem Fall einer nichttrivialen Jordan-Kette des zugehörigen linearen Systems.

6.7 Man zeige, dass für alle Lösungen des Systems

$$y_1' = -y_1^3 + y_2, \quad y_2' = -y_1 - y_2$$

$\lim_{x \rightarrow \infty} y(x) = 0$ erfüllt ist.

Hinweis: Man weise nach, dass das Funktional $E(y) = |y|^2$ für $y \neq 0$ streng monoton fallend auf den Lösungskurven ist.

Literatur

- [1] Königsberger, K.: Analysis II, Springer Verlag, Berlin 1992
- [2] Stoer, J.: Einführung in die Numerische Mathematik I, Springer-Verlag, Berlin 1983
- [3] Stoer, J. – Bulirsch, R.: Einführung in die Numerische Mathematik II, Springer-Verlag, Berlin 1973
- [4] Walter, W.: Gewöhnliche Differentialgleichungen, Heidelberger Taschenbücher 110, Springer-Verlag 1972
- [5] Walter, W.: Analysis II, Springer-Verlag, Berlin 1997