Analysis II

Sommersemester 2003

W. Ebeling

Institut für Algebraische Geometrie
Leibniz Universität Hannover
Postfach 6009
30060 Hannover
E-mail: ebeling@math.uni-hannover.de

1 Der $\mathbb{R}^n$ als normierter Vektorraum

Den $\mathbb{R}^n$ kennen Sie aus der linearen Algebra.

$$\mathbb{R}^n := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;\middle|\; x_1, \ldots, x_n \in \mathbb{R} \right\}.$$

Der $\mathbb{R}^n$ ist ein reeller Vektorraum der Dimension n mit der Addition und Skalarmultiplikation wie folgt:

$$x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \lambda \in \mathbb{R},$$

$$x + y := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix},$$

$$\lambda x := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

Die Vektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

bilden eine Basis des $\mathbb{R}^n$; sie wird als *Standardbasis* bezeichnet.

Wir wollen nun Normen auf dem $\mathbb{R}^n$ betrachten. Wir hatten bereits die Norm einer Funktion betrachtet, dies war eine Norm auf dem Vektorraum der beschränkten Funktionen.

Definition Es sei V ein reeller Vektorraum. Unter einer *Norm* auf V versteht man eine Funktion

$$\begin{array}{rccc} ||\ ||: & V & \to & \mathbb{R} \\ & x & \mapsto & ||x|| \end{array}$$

mit folgenden Eigenschaften:

(i) $||x|| \geq 0, ||x|| = 0 \Leftrightarrow x = 0$,

(ii) $||\lambda x|| = |\lambda| \cdot ||x||$ für alle $\lambda \in \mathbb{R}, x \in V$.

(iii) $||x + y|| \leq ||x|| + ||y||$ für alle $x, y \in V$.

Ein *normierter Vektorraum* ist ein Paar $(V, ||\ ||)$, das aus einem Vektorraum V und einer Norm $||\ ||$ auf V besteht. Ist klar, um welche Norm es sich handelt, schreibt man meist nur V statt $(V, ||\ ||)$.

Definition Es sei X eine Menge. Unter einer *Metrik* auf X versteht man eine Abbildung

$$\begin{array}{llll} d\colon & X \times X & \to & \mathbb{R} \\ & (x,y) & \mapsto & d(x,y) \end{array}$$

mit folgenden Eigenschaften:

(i) $d(x,y) = 0 \Leftrightarrow x = y$

(ii) $d(x,y) = d(y,x)$ für alle $x, y \in X$ (*Symmetrie*)

(iii) $d(x,z) \leq d(x,y) + d(y,z)$ für alle $x, y, z \in X$ (*Dreiecksungleichung*).

Ein *metrischer Raum* ist ein Paar (X, d), X Menge, d Metrik. Man nennt $d(x,y)$ den *Abstand* oder die *Distanz* der Punkte x und y bzgl. d.

Bemerkung 1.1 Aus den Axiomen folgt, dass $d(x,y) \geq 0$ für alle $x, y \in X$.

Beweis. Wende Dreiecksungleichung auf x, y, x an:

$$0 \overset{\text{(i)}}{=} d(x,x) \leq d(x,y) + d(y,x) \overset{\text{(ii)}}{=} 2d(x,y).$$

□

Satz 1.1 *Es sei* $(V, ||\ ||)$ *ein normierter Vektorraum. Dann wird durch*

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

eine Metrik d *auf* V *definiert.*

Beweis.

(i) $d(x,y) = 0 \Leftrightarrow ||x-y|| = 0 \Leftrightarrow x - y = 0 \Leftrightarrow x = y$.

(ii) $d(x,y) = ||x-y|| = |-1|\,||x-y|| = ||y-x|| = d(y,x)$.

(iii) $d(x,z) = ||x-z|| = ||x-y+y-z|| \leq ||x-y|| + ||y-z|| = d(x,y) + d(y,z)$.

□

Beispiel 1.1 Die *Maximumsnorm* auf dem $\mathbb{R}^n$

$$||x||_{\max} := \max\{|x_1|, \ldots, |x_n|\} \text{ für } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

Aufgabe 1.1 Man zeige, dass die Maximumsnorm eine Norm auf dem $\mathbb{R}^n$ ist.

Beispiel 1.2 Im $\mathbb{R}^2$ berechnet sich die Länge eines Vektors $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ durch $||x|| = \sqrt{x_1^2 + x_2^2}$. Es liegt daher nahe, die Norm

$$||x|| := \sqrt{x_1^2 + \cdots + x_n^2}$$

zu betrachten. Dies ist die sogenannte *euklidische Norm.* Sie kann mit Hilfe des *Standardskalarprodukts* (oder *euklidischen Skalarprodukts*)

$$\langle x, y\rangle := x_1 y_1 + \cdots + x_n y_n$$

(für $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$) eingeführt werden durch

$$||x|| = \sqrt{\langle x, x\rangle}.$$

Wieder sind die Eigenschaften (i) und (ii) unmittelbar einzusehen. Beweis von (iii): Um die Dreiecksungleichung

$$\sqrt{\langle x+y, x+y\rangle} \leq \sqrt{\langle x, x\rangle} + \sqrt{\langle y, y\rangle}$$

zu beweisen, geht man durch Quadrieren zu der äquivalenten Ungleichung

$$\langle x+y, x+y\rangle \leq \langle x, x\rangle + 2\sqrt{\langle x, x\rangle\langle y, y\rangle} + \langle y, y\rangle$$

über, die gleichbedeutend ist mit

$$\langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle \leq \langle x, x\rangle + 2\sqrt{\langle x, x\rangle\langle y, y\rangle} + \langle y, y\rangle.$$

Diese Ungleichung ist äquivalent mit der Cauchy-Schwarzschen Ungleichung

$$\langle x, y\rangle^2 \leq \langle x, x\rangle\langle y, y\rangle$$

die Gegenstand des nächstens Lemmas ist.

Lemma 1.1 (Cauchy-Schwarzsche Ungleichung) *Für alle $x, y \in \mathbb{R}^n$ gilt*

$$\langle x, y\rangle^2 \leq \langle x, x\rangle\langle y, y\rangle$$

Beweis. Wir betrachten das quadratische Polynom

$$\langle tx + y, tx + y\rangle = \langle x, x\rangle t^2 + 2\langle x, y\rangle t + \langle y, y\rangle.$$

Für $x \neq 0$ kann dieses Polynom keine zwei verschiedenen reellen Nullstellen besitzen. Daher muß für die Diskriminante

$$\langle x, y\rangle^2 - \langle x, x\rangle\langle y, y\rangle \leq 0$$

gelten. □

Beispiel 1.3 Es sei $p \geq 1$. Die l_p-Norm auf $\mathbb{R}^n$ ist erklärt durch

$$||x||_p := (|x_1|^p + \cdots |x_n|^p)^{\frac{1}{p}}.$$

Beweis, dass dies eine Norm ist, als Aufgabe. Die l_2-Norm ist die euklidische Norm; die Maximumsnorm kann als Grenzfall „$p = \infty$“ betrachtet werden.

Wir wollen nun den Begriff der Konvergenz von Folgen vom $\mathbb{R}^1$ auf den $\mathbb{R}^n$ übertragen. Da es für $n \geq 2$ jeweils beliebig viele Normen gibt, könnte die Definition von der Wahl der Norm abhängen. Tatsächlich ist dies aber nicht der Fall, wie wir zeigen werden. Zunächst verwenden wir zur Definition der Konvergenz die Maximumsnorm.

Definition Die Folge (x_k), $x_k \in \mathbb{R}^n$ heißt *konvergent gegen* $a \in \mathbb{R}^n$, wenn

$$\lim_{k\to\infty} ||x_k - a||_{\max} = 0$$

gilt. Man schreibt dann auch

$$\lim_{k\to\infty} x_k = a.$$

Bemerkung 1.2 Wie im eindimensionalen Fall zeigt man, dass eine Folge (x_k) nicht gegen zwei verschiedene Elemente a und b konvergieren kann. Wenn es ein a gibt, so dass (x_k) gegen a konvergiert, dann sagt man kurz, die Folge (x_k) sei konvergent.

Satz 1.2 *Die Folge (x_k) ist genau dann konvergent, wenn jede der n Komponentenfolgen (x_{ki}) konvergiert.*

Beweis. Dies ergibt sich aus der Abschätzung

$$|x_{ki} - a_i| \leq ||x_k - a||_{\max} \leq |x_{k1} - a_1| + \cdots + |x_{kk} - a_n|$$

für $i = 1, \ldots, n$. □

Auch den Satz von Bolzano-Weierstraß kann man auf den $\mathbb{R}^n$ übertragen.

Definition Ein Folge (x_k) heißt *beschränkt*, falls es ein $r \in \mathbb{R}$ gibt mit

$$||x_k||_{\max} < r \text{ für alle } k \in \mathbb{N}.$$

Satz 1.3 (Bolzano-Weierstraß) *Jede beschränkte Folge (x_k) besitzt eine konvergente Teilfolge.*

Beweis. Betrachte die n Komponentenfolgen der Folge (x_k), diese sind jeweils beschränkte Zahlenfolgen. Nach dem Satz von Bolzano-Weierstraß für $\mathbb{R}$ besitzt $x_{k,1}$ eine konvergente Teilfolge $(x_{k_l,1})$. Ebenso besitzt $(x_{k_l,2})$ eine konvergente Teilfolge $(x_{k_{l_m},2})$ usw. Nach n Schritten erhalten wir so eine Folge von Vektoren mit der Eigenschaft, dass jede Komponentenfolge konvergiert. □

Um zu zeigen, dass wir die Konvergenzdefinition statt mit der Maximumsnorm $||\ ||_{\max}$ ebensogut mit einer beliebigen Norm $||\ ||$ hätten fassen können, beweisen wir

Satz 1.4 *Zu jeder Norm $||\ ||\colon \mathbb{R}^n \to \mathbb{R}$ gibt es positive Zahlen α und β so, dass für alle $x \in \mathbb{R}^n$ gilt*

$$||x|| \leq \alpha ||x||_{\max} \text{ und } ||x||_{\max} \leq \beta ||x||.$$

Bemerkung 1.3 Aus diesen beiden Abschätzungen ergibt sich unmittelbar, dass gilt:

$$\lim_{k\to\infty} ||x_k - a||_{\max} = 0 \Leftrightarrow \lim_{k\to\infty} ||x_k - a|| = 0.$$

Beweis des Satzes.
(a)

$$\begin{aligned} ||x|| &= ||x_1 e_1 + \cdots + x_n e_n|| \\ &\leq |x_1|\,||e_1|| + \cdots + |x_n|\,||e_n|| \qquad \text{((ii), (iii))} \\ &\leq ||x||_{\max}||e_1|| + \cdots + ||x_n||_{\max}||e_n|| \\ &\leq (||e_1|| + \cdots + ||e_n||)||x||_{\max}. \end{aligned}$$

Also können wir $\alpha = ||e_1|| + \cdots + ||e_n||$ wählen.

(b) Die Abschätzung rechts beweisen wir indirekt. Angenommen, es gäbe kein derartiges $\beta > 0$. Dann könnte man zu $\beta = 1, 2, \ldots, k, \ldots$ jeweils einen Vektor x_k finden, so dass

$$||x_k||_{\max} > k||x_k||.$$

Nach (ii) würde dann für $k = 1, 2, 3, \ldots$ gelten

$$\left|\left|\frac{x_k}{||x_k||_{\max}}\right|\right| < \frac{1}{k}.$$

Für $y_k = \frac{x_k}{||x_k||_{\max}}$ hätten wir also

$$||y_k|| < \frac{1}{k}. \qquad (*)$$

Wegen $||y_k||_{\max} = 1$ ist die Folge (y_k) beschränkt; nach Satz 1.3 gibt es also eine konvergente Teilfolge (y_{k_l}), die gegen ein $z \in \mathbb{R}^n$ konvergiert, d. h. es gilt

$$\lim_{l \to \infty} ||y_{k_l} - z||_{\max} = 0.$$

Aus der bereits bewiesenen Abschätzung $||z - y_{k_l}|| \leq a||z - y_{k_l}||_{\max}$ ergibt sich

$$||z|| = ||z - y_{k_l} + y_{k_l}|| \leq ||z - y_{k_l}|| + ||y_{k_l}|| \leq a||z - y_{k_l}||_{\max} + ||y_{k_l}||.$$

Wegen $(*)$ folgt

$$\lim_{l \to \infty} ||y_{k_l}|| = 0,$$

also $||z|| = 0$, also $z = 0$ nach (i).

Andererseits gilt

$$1 = ||y_{k_l}||_{\max} = ||y_{k_l} - z + z||_{\max} \leq ||y_{k_l} - z||_{\max} + ||z||_{\max}.$$

Hieraus folgt $||z||_{\max} \geq 1$, also $z \neq 0$, ein Widerspruch. □

Definition Die Norm $||\ ||' \colon V \to \mathbb{R}$ heißt *äquivalent* zur Norm $||\ || \colon V \to \mathbb{R}$, wenn es positive Zahlen α, β gibt, so dass für alle $x \in V$ gilt:

$$\alpha||x|| \leq ||x||' \leq \beta||x||.$$

Satz 1.5 *Die Äquivalenz von Normen ist eine Äquivalenzrelation.*

Beweis. (i) Jede Norm ist zu sich selbst äquivalent: Setze $\alpha = \beta = 1$.

(ii) Symmetrie: Aus
$$\alpha||x|| \leq ||x||' \leq \beta||x||$$
folgt die Ungleichung
$$\frac{1}{\beta}||x||' \leq ||x|| \leq \frac{1}{\alpha}||x||'.$$

(iii) (Transitivität) Aus
$$\alpha||x|| \leq ||x||' \leq \beta||x||$$
und
$$\alpha'||x||' \leq ||x||'' \leq \beta'||x||'$$
folgt
$$\alpha\alpha'||x|| \leq ||x||'' \leq \beta\beta'||x||.$$
□

Korollar 1.1 *Je zwei Normen auf $\mathbb{R}^n$ sind äquivalent.*

Beweis. Satz 1.4 besagt, dass eine beliebige Norm zur Maximumsnorm äquivalent ist. Damit folgt die Behauptung aus Satz 1.5. □

Beispiel 1.4 Für die Maximumsnorm und die euklidische Norm gilt:
$$||x||_{\max} \leq |x| \leq \sqrt{n}||x||_{\max}.$$
Beweis als Aufgabe.

2 Topologie des $\mathbb{R}^n$

Im folgenden sei $|| \, ||$ eine Norm auf $\mathbb{R}^n$.

Definition Es sei $r > 0$ und $a \in \mathbb{R}^n$. Die Menge
$$B(a, r) := \{x \in \mathbb{R}^n \mid ||x - a|| < r\}$$
heißt *offene Kugel um a vom Radius r* bezüglich der Norm $|| \, ||$. Allgemeiner für metrischen Raum (X, d):
$$B(a, r) := \{x \mid d(x, a) < r\}.$$

Beispiel 2.1 1) Man skizziere offene Kugeln bezüglich der Maximumsnorm in $\mathbb{R}^2$ und $\mathbb{R}^3$!

2) In der euklidischen Norm sind die offenen Kugeln gewöhnliche Kugeln.

3) In der l_1-Norm?

Definition Eine Teilmenge $U \subset \mathbb{R}^n$ heißt *Umgebung* des Punktes $x \in \mathbb{R}^n$, falls ein $\varepsilon > 0$ existiert, so dass
$$B(x, \varepsilon) \subset U.$$
Insbesondere ist $B(x, \varepsilon)$ selbst eine Umgebung von x. Man nennt $B(x, \varepsilon)$ auch die *ε-Umgebung* von x.

Bemerkung 2.1 Der Umgebungsbegriff ist unabhängig von der gewählten Norm: Es sei $|| \, ||'$ eine weitere Norm mit

$$\alpha||x|| \leq ||x||' \leq \beta||x||$$

für geeignete positive Zahlen α und β und alle $x \in \mathbb{R}^n$. Wir bezeichnen mit $B'(a,\varepsilon)$ die Kugel um a bzgl. $|| \, ||'$. Dann gilt

$$B'(a,\varepsilon) \subset B\left(a, \frac{\varepsilon}{\alpha}\right), \qquad B\left(a, \frac{\varepsilon}{\beta}\right) \subset B'(a,\varepsilon).$$

Denn:

$$||x-a||' < \varepsilon \Rightarrow \alpha||x-a|| < \varepsilon \Rightarrow ||x-a|| < \frac{\varepsilon}{\alpha}$$

$$||x-a|| < \frac{\varepsilon}{\beta} \Rightarrow \beta||x-a|| < \varepsilon \Rightarrow ||x-a||' < \varepsilon.$$

Die Definition der Konvergenz einer Folge (bzw. Reihe) kann jetzt auch mit Hilfe des Umgebungsbegriffes formuliert werden.

Satz 2.1 *Eine Folge (x_k) aus $\mathbb{R}^n$ konvergiert genau dann gegen ein $a \in \mathbb{R}^n$, wenn gilt: Zu jeder Umgebung U von a existiert ein $k_0 \in \mathbb{N}$, so dass für alle $k \geq k_0$ gilt*

$$x_k \in U.$$

Bemerkung 2.2 Dass eine Folge nur einen Grenzwert haben kann, ergibt sich daraus, dass zwei verschiedene Punkte disjunkte Umgebungen besitzen ($\mathbb{R}^n$ ist ein sogenannter *Hausdorff-Raum*).

Definition Eine Teilmenge $U \subset \mathbb{R}^n$ heißt *offen*, wenn sie Umgebung jedes ihrer Punkte ist, d. h. wenn zu jedem $x \in U$ ein $\varepsilon > 0$ existiert, so dass

$$B(x,\varepsilon) \subset U.$$

Beispiel 2.2 Es seien $a, b \in \mathbb{R}$, $a < b$. Das Intervall (a,b) ist offen in $\mathbb{R}$, denn es gilt für $x \in (a,b)$:

$$B(x,\varepsilon) \subset (a,b) \text{ für } \varepsilon = \min(|a-x|, |b-x|).$$

Ebenso sind die uneigentlichen Intervalle $(a,+\infty)$, $(-\infty,a)$ offen. Sind die Intervalle $[a,b]$ und $[a,b)$ offen?

Aufgabe 2.1 Es sei $a \in \mathbb{R}^n$ und $r > 0$. Man zeige, dass $B(a,r)$ offen ist.

Satz 2.2 *Für die offenen Mengen des $\mathbb{R}^n$ gilt:*

(a) $\emptyset$ und $\mathbb{R}^n$ sind offen,

(b) Sind U und V offen, so ist auch der Durchschnitt $U \cap V$ offen,

(c) Sei $U_i, i \in I$, eine Familie offener Teilmengen von $\mathbb{R}^n$. Dann ist auch die Vereinigung $\bigcup_{i \in I} U_i$ offen.

Beweis. a) $\mathbb{R}^n$ ist offen, da für jeden Punkt x aus $\mathbb{R}^n$ auch eine Kugel um x in $\mathbb{R}^n$ enthalten ist. Die leere Menge ist offen, da es keinen Punkt $x \in \emptyset$ gibt, zu dem es eine ε-Umgebung $B(x, \varepsilon) \subset \emptyset$ geben müsste.

b) Sei $x \in U \cap V$. Da U, V offen, gibt es $\varepsilon_1 > 0, \varepsilon_2 > 0$ mit

$$B(x, \varepsilon_1) \subset U \text{ und } B(x, \varepsilon_2) \subset V.$$

Für $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$ gilt dann

$$B(x, \varepsilon) \subset U \cap V.$$

Also ist $U \cap V$ offen.

c) Ist $x \in \bigcup_{i \in I} U_i$, so gibt es ein $i_0 \in I$ mit $x \in U_{i_0}$. Da U_{i_0} offen ist, existiert ein $\varepsilon > 0$ mit

$$B(x, \varepsilon) \subset U_{i_0} \subset \bigcup_{i \in I} U_i.$$

□

Bemerkung 2.3 Aus Satz 2.2(b) folgt durch wiederholte Anwendung, dass ein Durchschnitt von endlichen vielen offenen Menge wieder offen ist. Dies gilt nicht mehr für Durchschnitte von unendlich vielen Mengen:

$$\{0\} = \bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n} \right).$$

Definition Eine Teilmenge $A \subset \mathbb{R}^n$ heißt *abgeschlossen*, wenn ihr Komplement $\mathbb{R}^n \setminus A$ offen ist.

Beispiel 2.3 1) Für $a, b \in \mathbb{R}, a \leq b$, ist das Intervall $[a, b]$ abgeschlossen, denn sein Komplement

$$\mathbb{R} \setminus [a, b] = (-\infty, a) \cup (b, \infty)$$

ist nach Satz 2.2(c) offen. Ebenso sind die Intervalle $[a, \infty)$ und $(-\infty, a]$ abgeschlossen.

2) Sind $A_1 \subset \mathbb{R}^k$ und $A_2 \subset \mathbb{R}^m$ abgeschlossen, so ist auch $A_1 \times A_2 \subset \mathbb{R}^{k+m}$ abgeschlossen. Denn sei $(x, y) \in \mathbb{R}^k \times \mathbb{R}^m, (x, y) \notin A_1 \times A_2$. Dann gilt $x \notin A_1$ oder $y \notin A_2$. Sei etwa $x \notin A_1$. Da A_1 abgeschlossen ist, gibt es ein $\varepsilon > 0$, so dass $B(x, \varepsilon) \subset \mathbb{R}^k \setminus A_1$. Daraus folgt

$$B((x, y), \varepsilon) \subset \mathbb{R}^{k+m} \setminus (A_1 \times A_2),$$

das Komplement von $A_1 \times A_2$ ist also offen.

Insbesondere folgt daraus, dass die Quader

$$Q := \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i, i = 1, \ldots, n\}.$$

$a_i, b_i \in \mathbb{R}$, $a_i \leq b_i$, abgeschlossen im $\mathbb{R}^n$ sind.

3) $\emptyset, \mathbb{R}^n$ sind abgeschlossen, da Komplemente von $\mathbb{R}^n, \emptyset$.

4) Für $a, b \in \mathbb{R}$, $a < b$ ist das Intervall $[a, b) \subset \mathbb{R}$ weder offen noch abgeschlossen.

Definition Es sei $M \subset \mathbb{R}^n$. Dann heißt $x \in \mathbb{R}^n$ *innerer Punkt* von M, wenn es eine Umgebung U von x gibt mit $U \subset M$.

Definition Es sei $M \subset \mathbb{R}^n$. Dann heißt $x \in \mathbb{R}^n$ *Häufungspunkt* von M, wenn in jeder Umgebung U von x ein $y \in M$ liegt mit $x \neq y$.

Definition Es sei $M \subset \mathbb{R}^n$. Dann heißt $x \in \mathbb{R}$ *Randpunkt* von M, wenn in jeder Umgebung von x mindestens ein Punkt von M und mindestens ein Punkt von $\mathbb{R}^n \setminus M$ liegt.

Die Menge aller Randpunkte von M bezeichnen wir mit ∂M.

Satz 2.3 *Eine Teilmenge $U \subset \mathbb{R}^n$ ist genau dann offen, wenn sie nur aus inneren Punkten besteht.*

Beweis. folgt unmittelbar aus der Definition. □

Bemerkung 2.4 Ein innerer Punkt von M gehört stets zu M; ein Häufungspunkt von M dagegen braucht nicht zu M gehören. Ein innerer Punkt von M ist stets auch Häufungspunkt von M; das Umgekehrte braucht offensichtlich nicht der Fall zu sein; z. B. sind a, b Häufungspunkte von (a, b), aber nicht innere Punkte. Was sind Häufungspunkte und innere Punkte von $\mathbb{Q} \subset \mathbb{R}$?

Satz 2.4 *Für eine Teilmenge $A \subset \mathbb{R}^n$ sind die folgenden Aussagen äquivalent:*

(i) A ist abgeschlossen,

(ii) Jeder Häufungspunkt von A gehört zu A,

(iii) Ist $(x_k)_{k\in\mathbb{N}}$ eine Folge von Punkten $x_k \in A$, die gegen einen Punkt $x \in \mathbb{R}^n$ konvergiert, so liegt x schon in A.

Beweis. (i) $\Rightarrow$ (ii): Es sei A abgeschlossen. Dann ist $\mathbb{R}^n \setminus A$ offen. Angenommen, es gäbe einen Häufungspunkt x von A, der nicht zu A, also zu $\mathbb{R}^n \setminus A$ gehört. Da x Häufungspunkt von A ist, liegt in jeder Umgebung von x ein Punkt von A. Also kann $x \in \mathbb{R}^n \setminus A$ nicht innerer Punkt von $\mathbb{R}^n \setminus A$ sein, Widerspruch zu $\mathbb{R}^n \setminus A$ offen.

(ii) $\Rightarrow$ (iii): Es sei $(x_k)_{k\in\mathbb{N}}$ eine Folge von Punkten $x_k \in A$, die gegen einen Punkt $x \in \mathbb{R}^n$ konvergiert. Dann ist x Häufungspunkt von A, denn nach Definition der Konvergenz liegt in jeder Umgebung von x ein $x_k \in A$. Nach Voraussetzung gehören alle Häufungspunkte von A zu A, also gilt $x \in A$.

(iii) $\Rightarrow$ (i): Das Folgenkriterium sei erfüllt; wir müssen zeigen, dass dann A abgeschlossen ist. Sei $x \in \mathbb{R}^n \setminus A$. Falls für jedes $\varepsilon > 0$ gilt $B(x, \varepsilon) \cap A \neq \emptyset$, so können wir zu jedem $k \geq 1$ ein $x_k \in A$ mit $||x - x_k|| < \frac{1}{k}$ finden. Dann gilt $\lim_{k\to\infty} x_k = x \in A$, Widerspruch zu $x \in X \setminus A$. Also gibt es doch ein $\varepsilon > 0$ mit $B(x, \varepsilon) \cap A = \emptyset$, d. h. $B(x, \varepsilon) \subset \mathbb{R}^n \setminus A$. Also ist $\mathbb{R}^n \setminus A$ offen. □

Beispiel 2.4 1) Im $\mathbb{R}^n$ ist der Rand der Einheitskugel

$$B = \{x \in \mathbb{R}^n \mid ||x|| \leq 1\}$$

die Einheitssphäre

$$\partial B = \{x \in \mathbb{R}^n \mid ||x|| = 1\}.$$

2) Was ist der Rand von $\mathbb{Q}$ in $\mathbb{R}$?

Satz 2.5 *Es sei $M \subset \mathbb{R}^n$. Dann gilt:*

(a) Die Menge $M \setminus \partial M$ ist offen,

(b) Die Menge $M \cup \partial M$ ist abgeschlossen,

(c) Der Rand ∂M ist abgeschlossen.

Beweis. (a) Sei $a \in M \setminus \partial M$ beliebig. Dann gibt es ein $\varepsilon > 0$, so dass

$$B(a, \varepsilon) \cap (\mathbb{R}^n \setminus M) = \emptyset,$$

denn sonst wäre a ein Randpunkt von M. Dann gilt auch

$$B(a, \varepsilon) \cap \partial M = \emptyset,$$

denn wäre $y \in B(a, \varepsilon) \cap \partial M$, dann $B(a, \varepsilon) \cap (\mathbb{R}^n \setminus M) \neq \emptyset$, da $B(a, \varepsilon)$ Umgebung von y. Insgesamt also

$$B(a, \varepsilon) \subset M \setminus \partial M.$$

Also ist $M \setminus \partial M$ offen.

(b) Wir setzen $M' := \mathbb{R}^n \setminus M$. Aus der Definition des Randes folgt $\partial M = \partial M'$. Nach Teil (a) ist $M' \setminus \partial M'$ offen, also ist

$$\mathbb{R}^n \setminus (M' \setminus \partial M') = (\mathbb{R}^n \setminus M') \cup \partial M' = M \cup \partial M$$

abgeschlossen.

(c) Es gilt $\partial M = (M \cup \partial M) \setminus (M \setminus \partial M)$, also

$$\mathbb{R} \setminus \partial M = (\mathbb{R}^n \setminus (M \cup \partial M)) \cup (M \setminus \partial M).$$

Nach Teil (a) und (b) ist dies offen, also ∂M abgeschlossen. □

Definition Ist M Teilmenge des $\mathbb{R}^n$, so heißt $\mathring{M} := M \setminus \partial M$ das *Innere* oder der *offene Kern* von M und $\overline{M} := M \cup \partial M$ die *abgeschlossene Hülle* von M.

Man kann die oben definierten Begriffe in einen noch abstrakteren Rahmen stellen. Alle bisherigen Definitionen benutzen nur die Existenz einer Metrik auf dem $\mathbb{R}^n$ und sind auch gültig, wenn man den $\mathbb{R}^n$ durch einen beliebigen metrischen Raum ersetzt. Ebenso gelten auch die entsprechenden Sätze für einen beliebigen metrischen Raum.

Tatsächlich kann man aber sogar auf die Metrik verzichten und die offenen Mengen als Grundbegriff nehmen.

Definition Es sei X eine Menge. Ein System $\mathcal{T}$ von Teilmengen von X heißt *Topologie* auf X, falls gilt:

(i) $\emptyset, X \in \mathcal{T}$,

(ii) $U, V \in \mathcal{T} \Rightarrow U \cap V \in \mathcal{T}$,

(iii) Ist I eine beliebige Indexmenge und $U_i \in \mathcal{T}$ für alle $i \in I$, so gilt $\bigcup_{i \in I} U_i \in \mathcal{T}$.

Ein *topologischer Raum* ist ein Paar $(X, \mathcal{T})$, bestehend aus einer Menge X und einer Topologie $\mathcal{T}$ auf X. Eine Teilmenge $U \subset X$ heißt *offen*, wenn sie zu $\mathcal{T}$ gehört. Eine Teilmenge $A \subset X$ heißt *abgeschlossen*, wenn ihr Komplement $X \setminus A$ offen ist.

Nach Satz 2.2 bildet das System der offenen Mengen des $\mathbb{R}^n$ oder allgemeiner eines metrischen Raumes eine Topologie im Sinne der obigen Definition. Ein metrischer Raum ist also in natürlicher Weise auch ein topologischer Raum.

Definition Es sei $(X, \mathcal{T})$ ein topologischer Raum und $x \in X$ ein Punkt. Eine Teilmenge $V \subset X$ heißt *Umgebung* von x, wenn es eine offene Menge $U \subset X$ mit $x \in U \subset V$ gibt.

Mit der Definition der offenen Kugel

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

im Falle eines metrischen Raumes (X, d) ist diese Definition mit der früher Gegebenen äquivalent.

Definition Ein topologischer Raum $(X, \mathcal{T})$ heißt *Hausdorff-Raum*, falls zu je zwei Punkten $x, y \in X$, $x \neq y$, Umgebungen U von x und V von y existieren mit $U \cap V = \emptyset$.

Satz 2.6 *Ein metrischer Raum (X, d) ist ein Hausdorff-Raum.*

Beweis. Es seien $x, y \in X$ mit $x \neq y$. Setze $\varepsilon := \frac{1}{2} d(x, y)$. Dann ist $\varepsilon > 0$, und

$$U := B(x, \varepsilon), \qquad V := B(y, \varepsilon)$$

sind Umgebungen von x bzw. y. Diese Umgebungen sind disjunkt, denn gäbe es $z \in U \cap V$, so würde folgen

$$2\varepsilon = d(x, y) \leq d(x, z) + d(z, y) < \varepsilon + \varepsilon,$$

also $2\varepsilon < 2\varepsilon$, Widerspruch! $\square$

Wir stellen noch einmal alle Begriffe zusammen: die Reihenfolge entspricht dem Abstraktheitsgrad:

$\mathbb{R}^n$
|
normierter Vektorraum
|
metrischer Raum
|
Hausdorff-Raum
|
topologischer Raum

Wir wollen nun noch einige spezielle Eigenschaften des $\mathbb{R}^n$ studieren.

Definition Eine Folge $(x_k)_{k\in\mathbb{N}}$ von Punkten aus dem $\mathbb{R}^n$ heißt *Cauchyfolge*, wenn gilt: Zu jedem $\varepsilon > 0$ existiert ein $k_0 \in \mathbb{N}$, so dass für alle $k, m \geq k_0$ gilt

$$||x_k - x_m|| < \varepsilon.$$

Bemerkung 2.5 Jede konvergente Folge ist eine Cauchyfolge (Beweis wie für $n = 1$, oder mit Satz 1.2).

Satz 2.7 *Im $\mathbb{R}^n$ konvergiert jede Cauchyfolge.*

Beweis. Es sei $x_k = (x_{k1}, x_{k2}, \ldots, x_{kn})$, $k \in \mathbb{N}$, eine Cauchyfolge im $\mathbb{R}^n$. Da

$$|x_{ki} - x_{mi}| \leq ||x_k - x_m||_{\max},$$

ist für jedes $i \in \{1, \ldots, n\}$ die Folge $(x_{ki})_{k\in\mathbb{N}}$ eine Cauchyfolge in $\mathbb{R}$, also wegen der Vollständigkeit von $\mathbb{R}$ konvergent. Nach Satz 1.2 konvergiert dann die Folge $(x_k)_{k\in\mathbb{N}}$ im $\mathbb{R}^n$. □

Definition Für eine Teilmenge $M \subset \mathbb{R}^n$ nennt man

$$\delta(M) := \sup_{x,y\in M} ||y - x||$$

den *Durchmesser* von M. Die Menge M heißt *beschränkt*, falls $\delta(M) < \infty$.

Bemerkung 2.6 M beschränkt $\Leftrightarrow$ Es existieren $a \in \mathbb{R}^n$ und $r \in \mathbb{R}_+^*$, so dass $M \subset B(a, r)$.

Aufgabe 2.2 Man gebe eine Abschätzung für $\delta(B(a, r))$ an.

Satz 2.8 (Schachtelungsprinzip, Satz von Cantor) *Es sei*

$$A_0 \supset A_1 \supset A_2 \supset \cdots$$

eine Folge von beschränkten nichtleeren abgeschlossenen Teilmengen des $\mathbb{R}^n$ mit

$$\lim_{k\to\infty} \delta(A_k) = 0.$$

Dann gibt es genau einen Punkt $x \in \mathbb{R}^n$, der allen Mengen A_k angehört:

$$\bigcap_{k=1}^{\infty} A_k = \{x\}.$$

Beweis. Wählt man aus jeder Menge A_k beliebig viele Punkte x_k aus, so erhält man eine Cauchyfolge: Denn für $k, m \geq k_0$ gilt wegen $A_k \subset A_{k_0}$, $A_m \subset A_{k_0}$

$$||x_k - x_m|| \leq \delta(A_{k_0}).$$

Nach Satz 2.7 konvergiert die Folge (x_k); ihr Grenzwert sei x. Da $x_m \in A_k$ für alle $m \geq k$, folgt aus Satz 2.4, dass $x \in A_k$.

Aus $\lim_{k\to\infty} \delta(A_k) = 0$ folgt, dass x der einzige Punkt des Durchschnitts aller A_k ist. □

Wir kommen nun zu dem wichtigen Begriff der kompakten Menge.

Definition Es sei M ein Teilmenge des $\mathbb{R}^n$. Unter einer *offenen Überdeckung* von M versteht man eine Familie $(U_i)_{i\in I}$ von offenen Teilmengen $U_i \subset \mathbb{R}^n$ mit

$$M \subset \bigcup_{i\in I} U_i.$$

Dabei ist I eine beliebige (endliche oder unendliche) Indexmenge.

Definition Eine Teilmenge $M \subset \mathbb{R}^n$ heißt *kompakt*, wenn es zu jeder offenen Überdeckung $(U_i)_{i\in I}$ von M eine endliche Teilüberdeckung gibt, d. h. es endlich viele Indizes $i_1, \ldots, i_k \in I$ gibt, so dass

$$M \subset U_{i_1} \cup U_{i_2} \cup \cdots \cup U_{i_k}.$$

(„M hat die *Heine-Borelsche Überdeckungseigenschaft.*")

Warnung Dies besagt nicht, dass M kompakt ist, wenn M eine endliche offene Überdeckung besitzt (so was gibt es immer!).

Satz 2.9 *Es sei $(x_k)_{k\in\mathbb{N}}$ eine Folge im $\mathbb{R}^n$, die gegen den Punkt $a \in \mathbb{R}^n$ konvergiert. Dann ist die Menge*

$$A := \{x_k \mid k \in \mathbb{N}\} \cup \{a\}$$

kompakt.

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

Da $a \in A$, gibt es ein $i^* \in I$, so dass $a \in U_{i^*}$. Da U_{i^*} offen, ist U_{i^*} Umgebung von a. Wegen $\lim_{k\to\infty} x_k = a$ gibt es also ein $k_0 \in \mathbb{N}$, so dass

$$x_k \in U_i^* \text{ für alle } k \geq k_0.$$

Außerdem liegt jedes x_k in einem gewissen U_{i_k}. Es gilt dann

$$A \subset U_{i_0} \cup U_{i_1} \cup \cdots \cup U_{i_{k_0}} \cup U_{i^*}.$$

Wir haben also eine endliche Teilüberdeckung gefunden. □

Bemerkung 2.7 Der Satz gilt i. a. nicht mehr, wenn man aus A den Grenzwert der Folge weglässt: Es sei

$$A := \left\{ \frac{1}{k} \;\middle|\; k \in \mathbb{N} \setminus \{0\} \right\} \subset \mathbb{R}.$$

Behauptung: A ist nicht kompakt.

Beweis. Wir setzen

$$U_1 := \left(\frac{1}{2}, 2\right), \quad U_k := \left(\frac{1}{k+1}, \frac{1}{k-1}\right) \text{ für } k \geq 2.$$

Jedes Intervall U_k ist offen, also ist $(U_k)_{k\geq 1}$ eine offene Überdeckung von A. Jedes U_k enthält genau einen Punkt von A, nämlich $\frac{1}{k}$. Deshalb wird A von keinem endlichen Teilsystem $(U_{k_1}, U_{k_2}, \ldots, U_{k_m})$ überdeckt. □

Satz 2.10 *Es seien $a_\nu, b_\nu \in \mathbb{R}$, $a_\nu \le b_\nu$, für $\nu = 1, \ldots, n$. Dann ist der abgeschlossene Quader*

$$Q := \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_\nu \le x_\nu \le b_\nu\}$$

kompakt.

Beweis. Es sei $(U_i)_{i \in I}$ eine offene Überdeckung von Q. Wir führen einen indirekten Beweis.

Annahme: Q kann nicht durch endlich viele U_{i_k} überdeckt werden.

Konstruieren durch vollständige Induktion eine Folge von abgeschlossenen Teilquadern

$$Q_0 \supset Q_1 \supset Q_2 \supset \cdots$$

mit

(i) Q_m kann nicht durch endlich viele U_{i_k} überdeckt werden,

(ii) $\delta(Q_m) = 2^{-m}\delta(Q)$.

Setze $Q_0 = Q$. Sei Q_m schon konstruiert,

$$Q_m = I_1 \times I_2 \times \cdots \times I_n, \quad I_\nu \in \mathbb{R} \text{ abgeschlossene Intervalle.}$$

Zerlege

$$I_\nu = I_\nu^{(1)} \cup I_\nu^{(2)}, \quad \delta(I_\nu^{(1)}) = \delta(I_\nu^{(2)}) = \frac{1}{2}\delta(I_\nu)$$

und setze

$$Q_m^{(s_1,\ldots,s_n)} := I_1^{(s_1)} \times I_2^{(s_2)} \times \cdots \times I_n^{(s_n)}.$$

$$\begin{array}{r|c|c|l}
\multicolumn{3}{c}{} & Q_m \\ \cline{2-3}
I_2^{(2)}\{ & Q_m^{(1,2)} & Q_m^{(2,2)} & \\ \cline{2-3}
I_2^{(1)}\{ & Q_m^{(1,1)} & Q_m^{(2,1)} & \\ \cline{2-3}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{\underbrace{\qquad\quad}_{I_1^{(1)}}} & \multicolumn{1}{c}{\underbrace{\qquad\quad}_{I_1^{(2)}}} &
\end{array}$$

Wir erhalten so 2^n Quader mit

$$\bigcup_{s_1,\ldots,s_n} Q_m^{(s_1,\ldots,s_n)} = Q_m.$$

Da Q_m nicht von endlich vielen U_{i_k} überdeckt werden kann, gibt es mindestens einen der Quader $Q_m^{(s_1,\ldots,s_n)}$, der nicht von endlich vielen U_{i_k} überdeckt werden kann. Dieser sei Q_{m+1}. Es gilt

$$\delta(Q_{m+1}) = \frac{1}{2}\delta(Q_m) = 2^{-m-1}\delta(Q).$$

Deshalb hat Q_{m+1} wieder die Eigenschaften (i) und (ii).

Nach Satz 2.8 gibt es $a \in Q$ mit $a \in Q_m$ für alle $m \in \mathbb{N}$. Da $(U_i)_{i \in I}$ eine Überdeckung von Q ist, gibt es $i_0 \in I$, so dass $a \in U_{i_0}$. Da U_{i_0} offen ist, gibt es ein $\varepsilon > 0$, so dass

$$B(a, \varepsilon) \subset U_{i_0}.$$

Sei m so groß, dass $\delta(Q_m) < \varepsilon$. Da $a \in Q_m$, gilt

$$Q_m \subset B(a, \varepsilon) \subset U_{i_0}, \quad \text{Widerspruch zu (i).}$$

□

Satz 2.11 *Jede kompakte Teilmenge $A \subset \mathbb{R}^n$ ist beschränkt und abgeschlossen.*

Beweis. (a) Es sei $a \in \mathbb{R}^n$ beliebig. Da

$$\mathbb{R}^n = \bigcup_{k=1}^{\infty} B(a,k),$$

ist $(B(a,k))_{k\geq 1}$ eine offene Überdeckung von A, es gibt also $k_1, \ldots, k_m$ mit

$$A \subset \bigcup_{j=1}^{m} B(a,k_j).$$

Für $k = \max\{k_1, \ldots, k_m\}$ gilt also $A \subset B(a,k)$, d.h. A ist beschränkt.

(b) Es sei $x \in \mathbb{R}^n \backslash A$ beliebig. Für $k \geq 1$ setze

$$U_k := \{y \in \mathbb{R}^n \mid ||y - x|| > \frac{1}{k}\}.$$

Die Menge U_k ist offen und es gilt

$$\bigcup_{k=1}^{\infty} U_k = \mathbb{R}^n \backslash \{x\} \supset A.$$

Da A kompakt ist, gibt es $k_1, \ldots, k_m$ mit

$$A \subset \bigcup_{j=1}^{m} U_{k_j}.$$

Für $k = \max\{k_1, \ldots, k_m\}$ gilt daher

$$B(x, \frac{1}{k}) \subset \mathbb{R}^n \backslash A.$$

Also ist $\mathbb{R}^n \backslash A$ offen, also A abgeschlossen. □

Satz 2.12 *Es sei $K \subset \mathbb{R}^n$ kompakt und $A \subset K$ abgeschlossen. Dann ist auch A kompakt.*

Beweis. Es sei $(U_i)_{i\in I}$ eine offene Überdeckung von A. $\mathbb{R}^n \backslash A$ ist offen und es gilt

$$K \subset \mathbb{R}^n = (\mathbb{R}^n \backslash A) \cup \bigcup_{i\in I} U_i.$$

Da K kompakt, gibt es $i_1, \ldots, i_k \in I$ mit

$$A \subset K \subset (\mathbb{R}^n \backslash A) \cup U_{i_1} \cup \cdots \cup U_{i_k}.$$

Daraus folgt

$$A \subset U_{i_1} \cup \cdots \cup U_{i_k}.$$

□

Satz 2.13 (Heine-Borel) *Eine Teilmenge $A \subset \mathbb{R}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Beweis. "$\Rightarrow$": Dies ist Satz 2.11.

"$\Leftarrow$": Ist A beschränkt und abgeschlossen, so ist A in einem (abgeschlossenen) Quader Q enthalten. Der Quader Q ist nach Satz 2.10 kompakt. Nach Satz 2.12 ist dann A kompakt. □

Satz 2.14 (kompakt = folgenkompakt) *Eine Teilmenge $A \subset \mathbb{R}^n$ ist genau dann kompakt, wenn gilt: jede Folge $(x_k)_{k\in\mathbb{N}}$ in A besitzt eine konvergente Teilfolge mit Grenzwert in A.*

Beweis. "$\Rightarrow$": Es sei A kompakt. Nach Satz 2.13 ist A beschränkt und abgeschlossen. Es sei $(x_k)_{k\in\mathbb{N}}$ eine Folge in A. Dann ist diese Folge beschränkt, besitzt also nach dem Satz von Bolzano-Weierstraß (Satz 1.3) eine konvergente Teilfolge mit Grenzwert $a \in \mathbb{R}^n$. Da A abgeschlossen ist, folgt aus Satz 2.4: $a \in A$.

"$\Leftarrow$": Die Menge A erfülle das Folgenkriterium. Es sei $(x_k)_{k\in\mathbb{N}}$ eine Folge in A, die gegen $x \in \mathbb{R}^n$ konvergiert. Da jede Teilfolge einer konvergenten Folge gegen denselben Grenzwert konvergiert, muss x in A liegen. Also ist A abgeschlossen.

Angenommen, A ist nicht beschränkt. Dann gäbe es zu jedem $k \in \mathbb{N}$ ein $x_k \in A$ mit

$$||x_k|| > k.$$

Die Folge $(x_k)_{k\in\mathbb{N}}$ besitzt aber keine konvergente Teilfolge, Widerspruch! □

3 Stetige Abbildungen

Wir wollen uns nun mit Abbildungen aus dem $\mathbb{R}^n$ in den $\mathbb{R}^m$ befassen.

Es sei $D \subset \mathbb{R}^n$ beliebig,

$$\begin{array}{rccc} f: & D & \longrightarrow & \mathbb{R}^m \\ & x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \longmapsto & f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix} \end{array}.$$

Ausführlich

$$f(x) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}$$

Für eine Funktion $f\colon \mathbb{R} \to \mathbb{R}$ haben wir auch geschrieben:

$$y = f(x).$$

Entsprechend wird eine Abbildung $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$ durch m Abbildungsgleichungen

$$\begin{array}{ll} y_1 & = f_1(x_1, \dots, x_n) \\ \vdots & \qquad\qquad\qquad\qquad (x \in D) \\ y_m & = f_m(x_1, \dots, x_n) \end{array}$$

beschrieben.

Beispiel 3.1 (1) Lineare Abbildungen

$$y = Ax, \quad A = (a_{ij})^{1\le i\le m}_{1\le j\le n} \quad m \times n\text{-Matrix},$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

(2) $m = 1$: Normfunktion

Für $m = 1$ spricht man von einer (reellwertigen) *Funktion* f,

$$\begin{array}{cccc} f: & D & \longrightarrow & \mathbb{R} \\ & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \longmapsto & f(x_1, \ldots, x_n) \end{array}.$$

Der *Graph* von f ist die Menge

$$\Gamma_f := \{(x, y) \in D \times \mathbb{R} \mid y = f(x)\} \subset \mathbb{R}^{n+1}.$$

Beispiel 3.2 $n = 2$: $D = \{(x_1, x_2) \mid x_1^2 + x_2^2 \le 1\}$

$$\begin{array}{cccc} f: & D & \longrightarrow & \mathbb{R} \\ & (x_1, x_2) & \longmapsto & \sqrt{1 - x_1^2 - x_2^2} \end{array}.$$

Man skizziere den Graph dieser Funktion!

Die Menge

$$N_f(c) = \{x \in D \mid f(x) = c\}$$

für $c \in \mathbb{R}$ nennt man die zu c gehörende *Niveaumenge* (*Niveaukurve*, *Niveaufläche*) von f.

Wie sehen in Beispiel 3.2 die Niveaumengen aus?

Wir definieren nun die Stetigkeit einer Abbildung $f\colon D \to \mathbb{R}^m$.

Definition Es sei $D \subset \mathbb{R}^n$, $f\colon D \to \mathbb{R}^m$ eine Abbildung. Dann heißt f *stetig* in $a \in D$, wenn gilt: Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $x \in D$ mit $||x - a|| < \delta$ gilt:

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

Bemerkung 3.1 Es kommt bei dieser Definition nach § 1 nicht darauf an, welche Normen im $\mathbb{R}^n$ und $\mathbb{R}^m$ verwendet werden; beide Normen werden mit dem gleichen Symbol bezeichnet.

Satz 3.1 *Eine Abbildung $f\colon D \to \mathbb{R}^m$ ist genau dann stetig in $a \in D$, wenn gilt: Zu jeder Umgebung V von $f(a)$ gibt es eine Umgebung U von a, so dass gilt:*

$$f(U \cap D) \subset V.$$

Beweis. "$\Leftarrow$": Es sei $\varepsilon > 0$ gegeben. Zu $V = B(f(a), \varepsilon)$ gibt es dann eine Umgebung U von a, so dass $f(U \cap D) \subset V$. Wähle $\delta > 0$ so, dass $B(a, \delta) \subset U$.

"$\Rightarrow$": Es sei V eine Umgebung von $f(a)$. Dann gibt es ein $\varepsilon > 0$ mit $B(f(a), \varepsilon) \subset V$. Da f stetig, gibt es ein $\delta > 0$ mit $f(B(a, \delta) \cap D) \subset B(f(a), \varepsilon) \subset V$. Man kann also $U = B(a, \delta)$ setzen. □

Um zu zeigen, dass man die Stetigkeit auch mit Hilfe von konvergenten Folgen definieren kann, führen wir einen Grenzwertbegriff für Abbildungen ein.

Definition Es sei $D \subset \mathbb{R}^n$, $f\colon D \to \mathbb{R}^m$ eine Abbildung, und a ein Häufungspunkt von D. Man schreibt

$$\lim_{x \to a} f(x) = b,$$

falls für jede Folge $(x_k)_{k \in \mathbb{N}}$, $x_k \in D$, $x_k \neq a$, mit $\lim_{k \to \infty} x_k = a$ gilt:

$$\lim_{k \to \infty} f(x_k) = b.$$

Satz 3.2 *Eine Abbildung $f\colon D \to \mathbb{R}^m$ ist in $a \in D$ stetig genau dann, wenn gilt:*

$$\lim_{x \to a} f(x) = f(a).$$

Beweis. Genau wie im eindimensionalen Fall, siehe I, Satz 11.2. □

Bemerkung 3.2 Wie im eindimensionalen Fall folgt: $\lim_{x \to a} f(x) = b \Leftrightarrow$ Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $x \in D$ mit $0 < ||x - a|| < \delta$ gilt: $||f(x) - b|| < \varepsilon$.

Beispiel 3.3 Betrachte die durch

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

definierte Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$. Die Funktion f ist stetig in $(x, y) \neq (0, 0)$, wie sich aus den Rechenregeln ergeben wird.

Die Funktion f ist nicht stetig in $(0, 0)$:

$$\begin{aligned} f(0, y) &= f(x, 0) = 0, \\ f(x, x) &= \frac{2x^2}{x^2 + x^2} = 1 \text{ für } x \neq 0. \end{aligned}$$

In jeder Umgebung von $(0, 0)$ gibt es Punkte, in denen f den Wert 1 annimmt (Aus „partieller Stetigkeit“, d.h. Stetigkeit der Funktionen $x \mapsto f(x, y_0)$, $y \mapsto f(x_0, y)$, folgt nicht die Stetigkeit von f!).

Bei Abbildungen kann man die Stetigkeit dadurch nachweisen, dass man prüft, ob die m Komponentenfunktionen stetig sind.

Satz 3.3 *Die Abbildung $f\colon D \to \mathbb{R}^m$, $f = (f_1, \ldots, f_m)$ ist genau dann in $a \in D$ stetig, wenn die reellen Funktionen $f_1, \ldots, f_m$ in a stetig sind.*

Beweis. Entweder mit Hilfe von Satz 3.2 und Satz 1.2, oder mit Hilfe der folgenden Abschätzung für $i = 1, \ldots, m$:

$$|f_i(x) - f_i(a)| \leq ||f(x) - f(a)||_{\max} \leq |f_1(x) - f_1(a)| + \cdots + |f_m(x) - f_m(a)|.$$

□

Definition Eine Abbildung $f \colon D \to \mathbb{R}^m$ heißt *stetig*, wenn f stetig in a ist für alle $a \in D$.

Es gibt nun für solche stetigen Abbildungen eine Kennzeichnung, die auch als Definition für stetige Abbildungen zwischen topologischen Räumen genommen werden kann:

Satz 3.4 *Es sei $D \subset \mathbb{R}^n$ offen. Eine Abbildung $f \colon D \to \mathbb{R}^m$ ist genau dann stetig, wenn gilt: Das Urbild jeder offenen Teilmenge des $\mathbb{R}^m$ ist eine offene Teilmenge des $\mathbb{R}^n$.*

Beweis. "$\Rightarrow$": Es sei $B \subset \mathbb{R}^m$ offen, $A := f^{-1}(B)$, $a \in A$, d.h. $f(a) \in B$. Da f in a stetig ist, gibt es zu jeder Umgebung V von $f(a)$ eine Umgebung U von a mit $f(U \cap D) \subset V$. Wählen wir nun $V \subset B$ (möglich, da B offen), dann gilt

$$U \cap D \subset f^{-1}(V) \subset f^{-1}(B) = A,$$

d. h. die offene Umgebung $U \cap D$ des Punktes a ist in A enthalten. Deshalb ist A offen.

"$\Leftarrow$": Es sei V offene Umgebung von $b = f(a)$. Nach Voraussetzung ist $U = f^{-1}(V)$ eine offene Teilmenge des $\mathbb{R}^n$, somit auch Umgebung von a. Es gilt $U \subset D$ und

$$f(U) \subset V,$$

d. h. f ist in a stetig. □

Definition Ist D eine beliebige Teilmenge des $\mathbb{R}^n$, so heißen die Durchschnitte von D mit offenen (bzw. abgeschlossenen) Mengen des $\mathbb{R}^n$ *relativ offen* (bzw. *relativ abgeschlossen*) bezüglich D.

Satz 3.5 *Es sei $D \subset \mathbb{R}^n$ eine beliebige Teilmenge. Eine Abbildung $f \colon D \to \mathbb{R}^n$ ist genau dann stetig, wenn gilt: Das Urbild jeder offenen Teilmenge des $\mathbb{R}^m$ ist relativ offen bezüglich D.*

Die folgenden Regeln gelten sowohl für die Stetigkeit an einer Stelle $a \in D$ wie auch für die Stetigkeit in ganz D.

Satz 3.6 *Mit $f \colon D \to \mathbb{R}^m$ und $g \colon D \to \mathbb{R}^m$ sind auch die Abbildungen $f + g$ und λf $(\lambda \in \mathbb{R})$ stetig.*

Beweis. (a) Der Beweis für $f + g$ ergibt sich aus der Abschätzung

$$\begin{aligned} ||(f+g)(x) - (f+g)(a)|| &= ||f(x) - f(a) + g(x) - g(a)|| \\ &\leq ||f(x) - f(a)|| + ||g(x) - g(a)||: \end{aligned}$$

Es sei $\varepsilon > 0$ gegeben und $\delta > 0$ so gewählt, dass für alle $x \in D$ mit $||x - a|| < \delta$ gilt

$$||f(x) - f(a)|| < \frac{\varepsilon}{2}, \quad ||g(x) - g(a)|| \leq \frac{\varepsilon}{2}.$$

Dann gilt auch $||(f + g)(x) - (f + g)(a)|| < \varepsilon$.

(b) Der Beweis für λf ergibt sich analog aus der Gleichung

$$||(\lambda f)(x) - (\lambda f)(a)|| = |\lambda| \, ||f(x) - f(a)||.$$

□

Satz 3.7 *Sind $f : D \to \mathbb{R}$, $g : D \to \mathbb{R}$ stetig, so ist $f \cdot g$ überall und $\frac{f}{g}$ an allen Stellen, in denen g nicht den Wert 0 annimmt, stetig.*

Beweis. (a) Es sei $a \in D$ und $\varepsilon > 0$ beliebig vorgegeben. Nach Voraussetzung gibt es dann ein (für f und g zugleich brauchbares) $\delta > 0$, so dass für alle $x \in D$ mit $||x - a|| < \delta$ folgt:

$$|f(x) - f(a)| < \varepsilon \text{ und } |g(x) - g(a)| < \varepsilon.$$

Damit ergibt sich für diese x

$$\begin{aligned} |f(x)g(x) - f(a)g(a)| &= |f(x)g(x) - f(a)g(x) + f(a)g(x) - f(a)g(a)| \\ &\leq |g(x)| \cdot |f(x) - f(a)| + |f(a)| \cdot |g(x) - g(a)| \\ &\leq (|g(a)| + \varepsilon) \cdot \varepsilon + |f(a)| \cdot \varepsilon \\ &= \varepsilon(|f(a)| + |g(a)| + \varepsilon). \end{aligned}$$

Daraus folgt die Stetigkeit von $f \cdot g$.

(b) Wegen (a) genügt es, die Stetigkeit von $\frac{1}{g}$ zu zeigen. Für $a \in D$ gelte $g(a) \neq 0$. Dann gilt

$$\left| \frac{1}{g(x)} - \frac{1}{g(a)} \right| = \frac{|g(x) - g(a)|}{|g(x)| \cdot |g(a)|} \leq \frac{\varepsilon}{(|g(a)| - \varepsilon) \cdot |g(a)|}.$$

Daraus folgt die Stetigkeit von $\frac{1}{g}$. □

Beispiel 3.4 Ein *Monom* auf dem $\mathbb{R}^n$ vom Grad r ist eine Funktion $\mathbb{R}^n \to \mathbb{R}$ der Gestalt

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1^{k_1} x_2^{k_2} \cdot \ldots \cdot x_n^{k_n},$$

wobei $k_1, \ldots, k_n \in \mathbb{N}$ und $k_1 + \cdots + k_n = r$ ist. Eine *Polynomfunktion* $F : \mathbb{R}^n \to \mathbb{R}$ vom Grad $\leq r$ ist eine Linearkombination von Monomen vom Grad $\leq r$,

$$F(x_1, \ldots, x_n) = \sum_{\substack{k_1, \ldots, k_n \\ k_1 + \cdots + k_n \leq r}} c_{k_1 \cdots k_n} x_1^{k_1} \cdot \ldots \cdot x_n^{k_n},$$

wobei $c_{k_1 \cdots k_n} \in \mathbb{R}$. Da die Koordinatenfunktionen $(x_1, \ldots, x_n) \mapsto x_\nu$ und die konstanten Funktionen stetig sind, folgt durch wiederholte Anwendung von Satz 3.6 und Satz 3.7, dass alle Polynomfunktionen auf dem $\mathbb{R}^n$ stetig sind.

Aufgabe 3.1 Man zeige, dass die Menge $U \subset \mathbb{R}^{n^2}$ aller $n \times n$-Matrizen A mit $\det A \neq 0$ offen ist.

Auch das Hinteranderausführen von stetigen Abbildungen ergibt wieder eine stetige Abbildung (wie für $m = n = 1$):

Satz 3.8 *Es seien $f \colon D_1 \to \mathbb{R}^p$, $g \colon D_2 \to \mathbb{R}^m$ stetige Abbildungen mit $D_2 \subset \mathbb{R}^p$, $f(D_1) \subset D_2$. Dann ist $g \circ f$ stetig.*

Beweis. Es sei $a \in D_1$ und $b := f(a) \in D_2$. Es sei $(x_k)_{k\in\mathbb{N}}$, $x_k \in D_1$, $x_k \neq a$, eine beliebige Folge mit $\lim_{k\to\infty} x_k = a$. Da f stetig ist, gilt

$$\lim_{k\to\infty} f(x_k) = f(a).$$

Aus der Stetigkeit von g folgt

$$\lim_{k\to\infty} g(f(x_k)) = g(f(a)),$$

also

$$\lim_{k\to\infty} (g \circ f)(x_k) = (g \circ f)(a).$$

Die Behauptung folgt somit aus Satz 3.2. □

Wir betrachten nun Eigenschaften von stetigen Abbildungen.

Satz 3.9 *Es sei $f \colon D \to \mathbb{R}^m$ eine stetige Abbildung. Ist $D \subset \mathbb{R}^n$ kompakt, so ist auch $f(D) \subset \mathbb{R}^m$ kompakt.*

Beweis. Es sei $(U_i)_{i\in I}$ eine offene Überdeckung von $f(D)$. Nach Satz 3.5 ist $f^{-1}(U_i)$ relativ offen in D, d. h. es gilt $f^{-1}(U_i) = D \cap V_i$ für eine offene Menge V_i und

$$D \subset \bigcup_{i\in I} V_i.$$

Da D kompakt ist, gibt es endlich viele Indizes $i_1, \ldots, i_k \in I$, so dass $D \subset \bigcup_{m=1}^k V_{i_m}$. Daraus folgt

$$f(D) \subset \bigcup_{m=1}^{k} U_{i_m}.$$

□

Satz 3.10 *Es sei $D \subset \mathbb{R}^n$ kompakt und $f \colon D \to \mathbb{R}$ stetig. Dann ist die Funktion f beschränkt und nimmt ihr (absolutes) Maximum und Minimum an, d.h. es gibt $a, b \in D$ mit*

$$f(a) = \inf\{f(x) \mid x \in D\}, \quad f(b) = \sup\{f(x) \mid x \in D\}.$$

Beweis. Nach Satz 3.9 ist $A := f(D)$ kompakt. Nach Satz 2.11 ist A beschränkt und abgeschlossen. Da A beschränkt ist, sind $\sup(A)$ und $\inf(A)$ endlich. Es existieren Folgen $x_k \in A$, $y_k \in A$ $(k \in \mathbb{N})$ mit $\lim_{k\to\infty} x_k = \sup(A)$ und $\lim_{k\to\infty} y_k = \inf(A)$. Da A abgeschlossen ist, folgt aus Satz 2.4:

$$\inf(A), \sup(A) \in A.$$

□

Aufgabe 3.2 Es sei $A \subset \mathbb{R}^n$ und $x \in \mathbb{R}^n$. Der *Abstand* von x zur Menge A wird definiert als

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

(i) Es sei f die durch $f(x) = \operatorname{dist}(x, A)$ auf $\mathbb{R}^n$ definierte Funktion. Man zeige, dass f stetig ist.

(ii) Für eine weitere Teilmenge $K \subset \mathbb{R}^n$ sei

$$\operatorname{dist}(K, A) := \inf_{x \in K} \operatorname{dist}(x, A) = \inf_{x \in K, y \in A} ||x - y||.$$

Man beweise: Ist $A \subset \mathbb{R}^n$ abgeschlossen und $K \subset \mathbb{R}^n$ kompakt und $A \cap K = \emptyset$, so gilt

$$\operatorname{dist}(K, A) > 0.$$

Gilt dies auch unter der schwächeren Voraussetzung, dass K nur abgeschlossen ist?

Aus Satz 3.10 folgt natürlich noch nicht, dass tatsächlich alle Werte zwischen $f(a)$ und $f(b)$ angenommen werden. Damit dies der Fall ist, braucht man andere Voraussetzungen über D. Für $n = 1$ reicht es, D als Intervall, d. h. als „zusammenhängend“ anzunehmen (Zwischenwertsatz). Diesen Begriff wollen wir nun einführen.

Definition Es sei $M \subset \mathbb{R}^n$. Ein *Weg in* M ist eine stetige Abbildung $\gamma \colon [0, 1] \to M$. Man sagt, $\gamma(0)$ und $\gamma(1)$ sind durch den Weg γ miteinander verbunden.

Eine Teilmenge $M \subset \mathbb{R}^n$ heißt *wegzusammenhängend*, wenn je zwei Punkte aus M durch einen Weg in M miteinander verbunden werden können.

Beispiel 3.5 (1) Intervalle in $\mathbb{R}$.

(2) Einheitskugel, Einheitssphäre, Kreisring.

Satz 3.11 *Es sei $f \colon D \to \mathbb{R}^m$ eine stetige Abbildung. Ist $D \subset \mathbb{R}^n$ wegzusammenhängend, so ist auch $f(D)$ wegzusammenhängend.*

Beweis. Es seien $a, b \in f(D)$, α, β Urbilder von a, b. Nach Voraussetzung können α und β durch einen Weg γ in D verbunden werden. Nach Satz 3.8 ist $f \circ \gamma \colon [0, 1] \to f(D)$ stetig, also ein Weg, der a und b verbindet. □

Bemerkung 3.3 Speziell sind die Bildmengen von Funktionen $f \colon D \to \mathbb{R}$, D wegzusammenhängend, immer Intervalle.

Auch I, Satz 14.4 lässt sich verallgemeinern.

Definition Eine Abbildung $f \colon D \to \mathbb{R}^m$ heißt *gleichmäßig stetig*, wenn gilt: Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$, so dass für alle $x, y \in D$ mit $||x - y|| < \delta$ gilt

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

Satz 3.12 *Ist $D \subset \mathbb{R}^n$ kompakt und $f \colon D \to \mathbb{R}^m$ stetig, so ist f gleichmäßig stetig.*

Beweis. Der Beweis von I, Satz 14.4 kann wörtlich übertragen werden (wobei die Absolutbeträge durch entsprechende Normen zu ersetzen sind).

Alternativ: Es sei $\varepsilon > 0$ vorgegeben. Dann gibt es zu jedem $\xi \in D$ ein $\delta(\xi) > 0$, so dass für alle $\eta \in B(\xi, \delta(\xi)) \cap D$ gilt:

$$||f(\eta) - f(\xi)|| < \frac{\varepsilon}{2}.$$

Da

$$\bigcup_{\xi \in D} B\left(\xi, \frac{1}{2}\delta(\xi)\right) \supset D$$

und D kompakt ist, gibt es Punkte $\xi_1, \ldots, \xi_k \in D$ mit

$$\bigcup_{j=1}^{k} B\left(\xi_j, \frac{1}{2}\delta(\xi_j)\right) \supset D.$$

Sei $\delta_0 := \frac{1}{2}\min(\delta(\xi_1), \ldots, \delta(\xi_k))$. Seien $x, x' \in D$ mit $||x - x'|| < \delta_0$. Dann gibt es ein $j \in \{1, \ldots, k\}$ mit $x \in B(\xi_j, \frac{1}{2}\delta(\xi_j))$. Dann gilt $x' \in B(\xi_j, \delta(\xi_j))$. Es folgt

$$||f(x) - f(\xi_j)|| < \frac{\varepsilon}{2} \text{ und } ||f(x') - f(\xi_j)|| < \frac{\varepsilon}{2},$$

also $||f(x) - f(x')|| < \varepsilon$. □

Die Umkehrabbildung einer stetigen Abbildung – falls sie existiert – braucht nicht wieder stetig zu sein.

Definition Eine injektive stetige Abbildung $f\colon D \to \mathbb{R}^m$, für die auch die Umkehrabbildung

$$f^{-1}\colon f(D) \to \mathbb{R}^n$$

stetig ist, heißt *Homöomorphismus* von D auf $f(D)$. Die Mengen D und $f(D)$ heißen dann zueinander *homöomorph*.

Aufgabe 3.3 Man zeige, dass die Abbildung

$$f\colon \mathbb{R}^n \to B, \quad x \mapsto \frac{1}{1 + ||x||}x,$$

ein Homöomorphismus von $\mathbb{R}^n$ auf $B = \{y \in \mathbb{R}^n \mid ||y|| < 1\}$ ist.

Satz 3.13 *Ist $f\colon D \to \mathbb{R}^m$ eine injektive stetige Abbildung mit kompakter Definitionsmenge D, dann ist auch f^{-1} stetig, d.h. $f\colon D \to f(D)$ ist ein Homöomorphismus.*

Beweis. Es sei $U \subset \mathbb{R}^n$ offen. Wir müssen zeigen, dass das Urbild von U unter f^{-1}, nämlich $f(U)$, relativ offen in $f(D)$ ist.

Setze $A := \mathbb{R}^n \backslash U$. Dann ist A abgeschlossen, nach Satz 2.12 $A \cap D$ als abgeschlossene Teilmenge von D kompakt und nach Satz 3.9 daher auch $f(A \cap D)$ kompakt, also insbesondere abgeschlossen. Daher ist $\mathbb{R}^m \backslash f(A \cap D)$ offen. Nun gilt aber

$$f(U) = (\mathbb{R}^m \backslash f(A \cap D)) \cap f(D),$$

also ist $f(U)$ relativ offen in $f(D)$. □

Die gleichmäßige Konvergenz von Abbildungsfolgen wird entsprechend wie im eindimensionalen Fall eingeführt.

Definition Es sei $f\colon D \subset \mathbb{R}^n \to \mathbb{R}^m$ eine beschränkte Abbildung. Die Zahl

$$||f|| := \sup_{x \in D} ||f(x)||$$

heißt *Norm* von f.

Dies ist eine Norm auf dem Vektorraum aller beschränkten Abbildungen von D in $\mathbb{R}^m$, dieser Raum wird damit zu einem normierten Vektorraum. Der Beweis verläuft wie im Fall $n = m = 1$ (I, Satz 14.1).

Ist D kompakt und f stetig, so gilt

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

Definition Es sei $(f_k)_{k \in \mathbb{N}}$ eine Folge von Abbildungen $f_k\colon D \subset \mathbb{R}^n \to \mathbb{R}^m$. Dann heißt (f_k) *gleichmäßig konvergent* gegen $f\colon D \to \mathbb{R}^m$, wenn gilt

$$\lim_{k \to \infty} ||f_k - f|| = 0.$$

Bemerkung 3.4 (i) Dabei ist natürlich vorausgesetzt, dass die Abbildungen $f_k - f$ beschränkt sind.

(ii) Die Eigenschaft einer Abbildungsfolge, gleichmäßig konvergent zu sein, hängt nicht von der gewählten Norm im $\mathbb{R}^m$ ab.

Satz 14.3 aus Analysis I kann samt Beweis auf den Fall von Abbildungen übertragen werden, d.h. es gilt

Satz 3.14 *Wenn die Abbildungen $f_k\colon D \to \mathbb{R}^m$ stetig sind und die Folge $(f_k)_{k \in \mathbb{N}}$ gleichmäßig gegen $f\colon D \to \mathbb{R}^m$ konvergiert, dann ist f stetig.*

Für *lineare* Abbildungen $f\colon \mathbb{R}^n \to \mathbb{R}^m$ ist die Norm zunächst nicht erklärt, da f auf $\mathbb{R}^n$ nicht beschränkt ist, falls $f \neq 0$. Man kann aber trotzdem in zweckmäßiger Weise eine Norm einführen. Für eine lineare Abbildung gilt

$$f(\lambda x) = \lambda f(x) \text{ für alle } \lambda \in \mathbb{R}, x \in \mathbb{R}^n \qquad \text{(Homogenität).}$$

Deswegen ist f durch seine Werte auf der Einheitskugel $||x|| = 1$ schon vollständig festgelegt. Da die Einschränkung von f auf diese kompakte Menge stetig ist, existiert für sie die Norm gemäß obiger Definition. Deswegen definiert man

Definition

$$||f|| := \sup_{||x||=1} ||f(x)|| = \sup_{x \neq 0} \frac{||f(x)||}{||x||}.$$

Satz 3.15 (i) *Durch obige Definition wird eine Norm auf dem Vektorraum* $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ *aller linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ erklärt.*

(ii) *Es gilt für $f \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ und alle $x \in \mathbb{R}^n$:*

$$||f(x)|| \leq ||f|| \cdot ||x||.$$

Beweis. (ii) folgt direkt aus der Definition.

(i) Zum Beweis der Dreiecksungleichung: Es seien $f, g \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$. Dann gilt für alle $x \in \mathbb{R}^n$:

$$||(f+g)(x)|| = ||f(x) + g(x)|| \leq ||f(x)|| + ||g(x)|| \leq (||f|| + ||g||)||x||,$$

also auch

$$||f+g|| = \sup_{||x||=1} ||(f+g)(x)|| \leq ||f|| + ||g||.$$

□

4 Kurven im $\mathbb{R}^n$

Es sei $I \subset \mathbb{R}$ ein beliebiges (eigentliches oder uneigentliches) Intervall, das mehr als einen Punkt enthält. Wir wollen nun zunächst stetige Abbildungen von I in den $\mathbb{R}^n$ betrachten. Man bezeichnet solche Abbildungen als Wege oder Kurven.

Definition Eine *Kurve* im $\mathbb{R}^n$ ist eine stetige Abbildung

$$\gamma\colon I \to \mathbb{R}^n.$$

Beispiel 4.1 (1) Es sei $r > 0$.

$$\begin{array}{cccc} \gamma\colon & [0, 2\pi] & \to & \mathbb{R}^2 \\ & t & \mapsto & \begin{pmatrix} r\cos t \\ r\sin t \end{pmatrix} \end{array}$$

Das Bild von γ ist ein Kreis vom Radius r.

(2) Es sei $a \in \mathbb{R}^n$, $v \in \mathbb{R}^n \backslash 0$.

$$\begin{array}{cccc} \gamma\colon & \mathbb{R} & \to & \mathbb{R}^n \\ & t & \mapsto & a + tv \end{array}$$

Das Bild von γ ist eine Gerade im $\mathbb{R}^n$ durch a mit Richtungsvektor v.

(3) Es seien $r > 0$ und $c \neq 0$, $r, c \in \mathbb{R}$. Die *Schraubenlinie* ist definiert durch

$$\begin{array}{cccc} \gamma\colon & \mathbb{R} & \to & \mathbb{R}^3 \\ & t & \mapsto & \begin{pmatrix} r\cos t \\ r\sin t \\ ct \end{pmatrix} \end{array}.$$

Man skizziere das Bild von γ!

(4) Es sei $f\colon I \to \mathbb{R}$ stetig. Der Graph dieser Funktion ist eine Kurve im $\mathbb{R}^2$:

$$\begin{array}{cccc} \gamma\colon & I & \to & \mathbb{R}^2 \\ & t & \mapsto & \begin{pmatrix} t \\ f(t) \end{pmatrix} \end{array}.$$

(5) Kurven können sehr kompliziert sein: Es gibt zum Beispiel eine Kurve, die ein Dreieck ausfüllt (Peano-Kurve). Zur Konstruktion siehe Barner/Flohr, Analysis II, p. 40f.

Kinematische Interpretation einer Kurve $\gamma\colon I \to \mathbb{R}^n$*:* Man fasst die Variable $t \in I$ als Zeit und $\gamma(t) \in \mathbb{R}^n$ als Ort auf. Die Kurve beschreibt dann die zeitliche Bewegung eines Punktes im $\mathbb{R}^n$ (manchmal sehr nützlich).

Da die Bildmengen von Kurven recht kompliziert sein können, wollen wir nun einschränkende Bedingungen an γ stellen, die garantieren, dass die möglichen Bildmengen der anschaulichen Kurvenvorstellung besser gerecht werden. Wir wollen nun definieren, wann eine Kurve differenzierbar heißt.

Definition Eine Kurve $\gamma\colon I \to \mathbb{R}^n$ heißt *differenzierbar* in $t_0 \in I$, falls der Grenzwert

$$\lim_{t\to t_0} \frac{\gamma(t)-\gamma(t_0)}{t-t_0}$$

existiert.

Wie im Falle $n = 1$ kann man zeigen, dass, falls der Grenzwert existiert, dieser eindeutig bestimmt ist.

Definition Ist $\gamma\colon I \to \mathbb{R}^n$ differenzierbar in t_0, so heißt

$$\gamma'(t_0) = \lim_{t\to t_0} \frac{\gamma(t)-\gamma(t_0)}{t-t_0}$$

der *Tangentialvektor* von γ zum Parameterwert t_0. Falls $\gamma'(t_0) \neq 0$, so heißt der auf den Betrag 1 normierte Vektor

$$\frac{\gamma'(t_0)}{||\gamma'(t_0)||},$$

wobei $||\quad||$ die euklidische Norm darstellt, der *Tangenten-Einheitsvektor* von γ in t_0.

Geometrische Interpretation: $\gamma'(t_0)$ läßt sich als Limes von Sekanten auffassen (Bild!).

Physikalische Interpretation:
$\frac{\gamma(t)-\gamma(t_0)}{t-t_0}$ Vektor der Durchschnittsgeschwindigkeit zwischen den Zeitpunkten t und t_0,
$\gamma'(t_0)$ momentaner Geschwindigkeitsvektor,
$\|\gamma'(t_0)\| = \sqrt{|\gamma_1'(t_0)|^2 + \cdots + |\gamma_n'(t_0)|^2}$ Momentangeschwindigkeit.

Um die allgemeine Definition der Differenzierbarkeit vorzubereiten, geben wir auch noch eine „quotientenfreie“ Charakterisierung der Differenzierbarkeit einer Kurve.

Satz 4.1 *Eine Kurve* $\gamma\colon I \to \mathbb{R}^n$ *ist genau dann differenzierbar in* $t_0 \in I$ *mit Tangentialvektor* $\gamma'(t_0)$*, wenn es eine in* t_0 *stetige Abbildung* $\varrho\colon I \to \mathbb{R}^n$ *gibt, so dass gilt:*

$$\gamma(t) = \gamma(t_0) + \gamma'(t_0)(t-t_0) + \varrho(t)(t-t_0)$$

und

$$\varrho(t_0) = 0.$$

Beweis. Setze

$$\varrho(t) = \begin{cases} \frac{\gamma(t)-\gamma(t_0)}{t-t_0} - \gamma'(t_0) & \text{für } t \neq t_0, \\ 0 & \text{für } t = t_0. \end{cases}$$

Die Kurve γ ist genau dann differenzierbar in t_0 mit Tangentialvektor $\gamma'(t_0)$, wenn ϱ stetig ist. □

Die Bedingung von Satz 4.1 kann so interpretiert werden: Durch

$$g(t) = \gamma(t_0) + \gamma'(t_0)(t - t_0)$$

ist die Parameterdarstellung einer Geraden gegeben, die in der Umgebung von t_0 die Kurve γ gut approximiert, d.h.

$$\lim_{t \to t_0} \frac{\gamma(t) - g(t)}{t - t_0} = 0.$$

Satz 4.2 *Eine Kurve $\gamma = (\gamma_1, \ldots, \gamma_n)\colon I \to \mathbb{R}^n$ ist genau dann differenzierbar in t_0, wenn alle Funktionen γ_i differenzierbar in t_0 sind.*

Beweis. Dies ist klar. □

Bemerkung 4.1 Eine Kurve $\gamma\colon I \to \mathbb{R}^n$ braucht nicht notwendig injektiv zu sein. Gilt $\gamma(t_1) = \gamma(t_2) =: x$ für $t_1 \neq t_2$, so heißt x *Doppelpunkt* von γ. In x hat γ i.A. zwei verschiedene Tangentialvektoren.

Aufgabe 4.1 Es sei

$$\begin{array}{rccc} \gamma\colon & \mathbb{R} & \to & \mathbb{R}^2 \\ & t & \mapsto & \begin{pmatrix} t^2 - 1 \\ t^3 - t \end{pmatrix} \end{array}.$$

Man zeige $\gamma(\mathbb{R}) = \{(x, y) \in \mathbb{R}^n \mid y^2 = x^2 + x^3\}$ und skizziere $\gamma(\mathbb{R})$. (Die Kurve heißt *kartesisches Blatt.*)

Definition Die Kurve $\gamma\colon I \to \mathbb{R}^n$ heißt *stetig differenzierbar*, wenn

$$\begin{array}{rccc} \gamma'\colon & I & \to & \mathbb{R}^n \\ & t & \mapsto & \gamma'(t) \end{array}$$

stetig ist.

Definition Es sei $\gamma\colon I \to \mathbb{R}^n$ eine stetig differenzierbare Kurve. Ein Parameterwert $t \in I$ mit $\gamma'(t) \neq 0$ heißt *regulär*, einer mit $\gamma'(t) = 0$ heißt *singulär*. Die Kurve γ heißt *regulär*, wenn alle Parameterwerte $t \in I$ regulär sind.

Aufgabe 4.2 Man zeige, dass 0 ein singulärer Parameterwert der Kurve (*Neilsche Parabel*)

$$\begin{array}{rccc} \gamma\colon & \mathbb{R} & \to & \mathbb{R}^2 \\ & t & \mapsto & \begin{pmatrix} t^2 \\ t^3 \end{pmatrix} \end{array}$$

ist. Man zeige $\gamma(\mathbb{R}) = \{(x, y) \in \mathbb{R}^2 \mid y^2 = x^3\}$.

Bogenlänge

Es sei $I = [a,b] \subset \mathbb{R}$, $a < b$, ein abgeschlossenes Intervall und $\gamma\colon [a,b] \to \mathbb{R}^n$ eine stetig differenzierbare Kurve. Wir wollen die Länge der Kurve beschreiben. Dazu approximieren wir die Kurve durch Streckenzüge.

Nehmen wir zu diesem Zweck eine Einteilung $E\colon a = t_0 < t_1 < \cdots < t_q = b$ des Intervalls $[a.b]$. Zu dieser gehört ein Streckenzug mit den Eckpunkten $\gamma(t_k), k = 0, \ldots, q$.

Definition Die *Länge* des Streckenzuges bezüglich E ist die Zahl

$$\lambda(E) := \sum_{k=1}^{q} ||\gamma(t_k) - \gamma(t_{k-1})||.$$

Definition Es sei (E_m) eine Folge von Einteilungen, deren Feinheitsgrade $\varphi(E_m)$ gegen 0 konvergieren. Dann heißt die Zahl

$$L(\gamma) := \lim_{m\to\infty} \lambda(E_m)$$

die *(Bogen)länge der Kurve* γ.

Aus dem nächsten Satz folgt, dass dieser Grenzwert immer existiert und nicht von der gewählten Folge von Einteilungen abhängt.

Satz 4.3 *Für die Länge $L(\gamma)$ einer stetig differenzierbaren Kurve $\gamma\colon [a,b] \to \mathbb{R}^n$ gilt*

$$L(\gamma) = \int_a^b ||\gamma'(t)||\, dt.$$

Zum Beweis von Satz 4.3 benötigen wir einen Hilfssatz.

Lemma 4.1 *Es sei $\gamma\colon [a,b] \to \mathbb{R}^n$ stetig differenzierbar. Dann gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass*

$$\left\| \frac{\gamma(t) - \gamma(\tau)}{t - \tau} - \gamma'(t) \right\| \leq \varepsilon$$

für alle $t, \tau \in [a,b]$ mit $0 < |t - \tau| \leq \delta$.

Beweis. (a) Zunächst $n = 1$: $\gamma'\colon [a,b] \to \mathbb{R}$ ist nach Voraussetzung stetig, also sogar gleichmäßig stetig. Zu $\varepsilon > 0$ gibt es also ein $\delta > 0$, so dass für $s, t \in [a,b]$ mit $|t - s| \leq \delta$ gilt

$$|\gamma'(t) - \gamma'(s)| \leq \varepsilon.$$

Sei nun $t, \tau \in [a,b]$ mit $0 < |t - \tau| \leq \delta$. Nach dem MWS gibt es ein s zwischen t und τ, so dass

$$\frac{\gamma(t) - \gamma(\tau)}{t - \tau} = \gamma'(s).$$

Also ist

$$\left| \frac{\gamma(t) - \gamma(\tau)}{t - \tau} - \gamma'(t) \right| = |\gamma'(s) - \gamma'(t)| \leq \varepsilon.$$

(b) Nun n beliebig und $\gamma = (\gamma_1, \ldots, \gamma_n)$. Nach Beispiel 1.4 gilt

$$\left\| \frac{\gamma(t) - \gamma(\tau)}{t - \tau} - \gamma'(t) \right\| \leq \sqrt{n} \cdot \max_{i=1,\ldots,n} \left| \frac{\gamma_i(t) - \gamma_i(\tau)}{t - \tau} - \gamma_i'(t) \right|,$$

also folgt die Behauptung aus Teil (a). $\square$

Beweis von Satz 4.3. Wir versuchen zunächst den gleichen Beweis wie in Analysis I. Es gilt

$$\lambda(E) = \sum_{k=1}^{q} \left\| \frac{\gamma(t_k) - \gamma(t_{k-1})}{t_k - t_{k-1}} \right\| \cdot (t_k - t_{k-1}).$$

Diese Summe müssen wir mit einer Riemannschen Summe der Funktion $\|\gamma'(t)\|$ bezüglich E vergleichen:

$$\sum_{k=1}^{q} \|\gamma'(\tau_k)\|(t_k - t_{k-1})$$

für Zwischenstellen $\tau_k \in (t_{k-1}, t_k)$. Für $n = 1$ können wir einfach den MWS anwenden. Für beliebiges n ist das nicht mehr möglich, da die Zwischenstellen i.A. für jede Komponentenfunktion verschieden sind.

Wir werden aber zeigen, dass bei genügend kleinem Feinheitsgrad der Einteilung E die Länge $\lambda(E)$ einer Riemannschen Summe der obigen Art beliebig nahe kommt.

Es sei $\varepsilon > 0$ vorgegeben. Nach dem Lemma existiert ein $\delta > 0$ mit folgender Eigenschaft: Hat die Einteilung

$$E\colon a = t_0 < t_1 < \cdots < t_q = b$$

einen Feinheitsgrad $\varphi(E) < \delta$, so gilt

$$\left\| \frac{\gamma(t_k) - \gamma(t_{k-1})}{t_k - t_{k-1})} - \gamma'(t_k) \right\| \leq \frac{\varepsilon}{2(b-a)}$$

für $k = 1, \ldots, q$. Daraus folgt

$$\begin{aligned}
&\left| \sum_{k=1}^{q} \|\gamma(t_k) - \gamma(t_{k-1})\| - \sum_{k=1}^{q} \|\gamma'(t_k)\|(t_k - t_{k-1}) \right| \\
&\leq \sum_{k=1}^{q} \left\| \frac{\gamma(t_k) - \gamma(t_{k-1})}{t_k - t_{k-1}} - \gamma'(t_k) \right\| \cdot (t_k - t_{k-1}) \\
&\leq \sum_{k=1}^{q} \frac{\varepsilon}{2(b-a)} \cdot (t_k - t_{k-1}) = \frac{\varepsilon}{2}.
\end{aligned}$$

Nun sei (E_m) eine Folge von Einteilungen mit $\lim_{m\to\infty} \varphi(E_m) = 0$. Nach I, Satz 15.8, gilt dann

$$\lim_{m\to\infty} \sum_{k=1}^{q_m} \|\gamma'(t_k^{(m)})\|(t_k^{(m)} - t_{k-1}^{(m)}) = \int_a^b \|\gamma'(t)\|\, dt.$$

Also gibt es ein $m_0 \in \mathbb{N}$, so dass für alle $m \geq m_0$ gilt:

$$\varphi(E_m) < \delta$$

und

$$\left| \int_a^b \|\gamma'(t)\| \, dt - \sum_{k=1}^{q_m} \|\gamma'(t_k^{(m)})\| (t_k^{(m)} - t_{k-1}^{(m)}) \right| < \frac{\varepsilon}{2}.$$

Dann folgt aber

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

also

$$\lim_{m \to \infty} \underbrace{\sum_{k=1}^{q_m} \|\gamma(t_k^{(m)}) - \gamma(t_{k-1}^{(m)})\|}_{=\lambda(E_m)} = \int_a^b \|\gamma'(t)\| \, dt.$$

□

Aufgabe 4.3 Man berechne die Länge des Kreisbogens des Einheitskreises.

Aufgabe 4.4 Läßt man einen Kreis auf einer Geraden abrollen, so beschreibt ein beliebig auf der Peripherie gewählter Punkt eine *Zykloide*.

(a) Man zeige, dass die Zykloide die Parameterdarstellung

$$\begin{array}{rccc} \gamma\colon & [0, 2\pi] & \to & \mathbb{R}^2 \\ & t & \mapsto & \begin{pmatrix} t - \sin t \\ 1 - \cos t \end{pmatrix} \end{array}$$

besitzt.

(b) Man berechne die Länge der Zykloide.

Aufgabe 4.5 Man berechne die Länge der Schraubenlinie.

Aufgabe 4.6 Es sei $f\colon [a, b] \to \mathbb{R}$ stetig und

$$\begin{array}{rccc} \gamma\colon & [a, b] & \to & \mathbb{R}^2 \\ & t & \mapsto & \begin{pmatrix} t \\ f(t) \end{pmatrix} \end{array}.$$

Man zeige

$$L(\gamma) = \int_a^b \sqrt{1 + (f'(t))^2} \, dt.$$

Parametertransformation

Es sei $f\colon [a,b] \to \mathbb{R}^n$ eine Kurve, $[\alpha,\beta] \subset \mathbb{R}$ ein weiteres Intervall und

$$\varphi\colon [\alpha,\beta] \to [a,b]$$

eine bijektive stetige Abbildung. Dann ist die zusammengesetzte Abbildung

$$g := f \circ \varphi\colon [\alpha,\beta] \to \mathbb{R}^n$$

wieder eine Kurve im $\mathbb{R}^n$. Man sagt, dass die Kurve g aus f durch die *Parametertransformation* φ hervorgeht. Sind

$$\varphi\colon [\alpha,\beta] \to [a,b]$$

und

$$\varphi^{-1}\colon [a,b] \to [\alpha,\beta]$$

stetig differenzierbar, so nennt man φ eine *C^1-Parametertransformation.*

Die Bildmengen von f und g sind dieselben, da $g(t) = f(\varphi(t))$ für alle $t \in [\alpha,\beta]$, aber sie werden i.A. verschieden durchlaufen.

Da $\varphi\colon [\alpha,\beta] \to [a,b]$ stetig und bijektiv ist, tritt genau einer der beiden folgenden Fälle auf:

1. φ ist streng monoton wachsend. Man nennt φ dann *orientierungstreu.*
2. φ ist streng monoton fallend. Dann heißt φ *orientierungsumkehrend.*

Ist φ eine C^1-Parametertransformation, so folgt aus $\varphi^{-1} \circ \varphi = \mathrm{id}$ mit der Kettenregel

$$(\varphi^{-1})'(\varphi(t)) \cdot \varphi'(t) = 1,$$

d. h. $\varphi'(t) \neq 0$ für alle $t \in [\alpha,\beta]$. Es gilt

$$\begin{aligned} \varphi \text{ ist orientierungstreu} \quad &\Leftrightarrow \quad \varphi'(t) > 0 \text{ für alle } t. \\ \varphi \text{ ist orientierungsumkehrend} \quad &\Leftrightarrow \quad \varphi'(t) < 0 \text{ für alle } t. \end{aligned}$$

Die Bogenlänge bleibt unter C^1-Parametertransformationen invariant: Es sei $\gamma\colon [a,b] \to \mathbb{R}^n$ eine stetig differenzierbare Kurve, $\varphi\colon [\alpha,\beta] \to [a,b]$ eine C^1-Parametertransformation. O.B.d.A. sei φ orientierungstreu, d.h. $\varphi'(\tau) > 0$ für alle $\tau \in [\alpha,\beta]$. Es sei

$$\widetilde{\gamma} = \gamma \circ \varphi\colon [\alpha,\beta] \to \mathbb{R}^n.$$

Dann gilt:

$$L(\widetilde{\gamma}) = L(\gamma).$$

Dies folgt aus der Substitutionsregel wie folgt

$$L(\widetilde{\gamma}) = \int_\alpha^\beta \|\widetilde{\gamma}'(\tau)\|\, d\tau = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} \|\gamma'(\varphi(\tau))\| \cdot \varphi'(\tau)\, d\tau = \int_a^b \|\gamma'(t)\|\, dt = L(\gamma).$$

Jede reguläre Kurve besitzt eine ausgezeichnete Parameterdarstellung, nämlich eine solche, für die der Geschwindigkeitsvektor überall den Betrag 1 hat: Es sei $\gamma\colon [a,b] \to \mathbb{R}^n$ eine stetig differenzierbare Kurve. Wir versuchen eine orientierungstreue C^1-Parametertransformation $\varphi\colon [\alpha,\beta] \to [a,b]$ so zu bestimmen, dass $\widetilde{\gamma} = \gamma \circ \varphi$ die gewünschte Bedingung

$$\|\widetilde{\gamma}'(s)\| = 1$$

erfüllt. Dies führt auf

$$\|\gamma'(\varphi(s))\| \cdot \varphi'(s) = 1.$$

Die Umkehrfunktion $\psi = \varphi^{-1}$ von φ muss daher der Bedingung

$$\psi'(\varphi(s)) \cdot \varphi'(s) = 1, \quad \varphi(s) = t, \quad \psi'(t) = \|\gamma'(t)\|$$

genügen, d.h.

$$s = \psi(t) = s_0 + \int_a^t \|\gamma'(u)\|\, du.$$

Die frei wählbare Konstante s_0 kann man z.B. gleich 0 setzen, so dass $t = a$ dem Punkt $s = 0$ entspricht.

Wegen der Voraussetzung $\|\gamma'(t)\| > 0$ ist ψ streng monoton wachsend, besitzt also eine Umkehrfunktion φ, die ebenso wie ψ stetig differenzierbar ist.

Da $s = \psi(t)$ die Länge der zum Intervall $[a,t]$ gehörenden Kurve ist, wird s als *Bogenlängenparameter* bezeichnet, und man sagt, die Kurve

$$\begin{array}{llcl} \widetilde{\gamma}\colon & [0, L(\gamma)] & \to & \mathbb{R}^n \\ & s & \mapsto & \gamma(s) \end{array}$$

ist *durch die Bogenlänge parametrisiert.*

5 Differenzierbare Abbildungen

Wir führen nun den Begriff der Differenzierbarkeit einer Abbildung $f\colon \mathbb{R}^n \to \mathbb{R}^m$ ein. Dabei setzen wir voraus, dass die Definitionsmenge von f offen ist. Diese Vorausetzung ist ausreichend für die Anwendungen.

Die Grundidee der Differentialrechnung besteht darin, eine beliebige Abbildung in der Umgebung eines festen Punktes durch eine lineare Abbildung zu approximieren.

Definition Es sei $U \subset \mathbb{R}^n$ offen. Eine Abbildung $f\colon U \to \mathbb{R}^m$ heißt *differenzierbar in $a \in U$*, wenn es eine lineare Abbildung $T\colon \mathbb{R}^n \to \mathbb{R}^m$ und eine in a stetige Abbildung $r\colon U \to \mathbb{R}^m$ gibt, so dass für alle $x \in U$ gilt

$$f(x) = f(a) + T(x-a) + r(x)||x-a||$$

und

$$r(a) = 0.$$

Wann ist $f\colon U \to \mathbb{R}^m$ also differenzierbar in $a \in U$? Es muss eine lineare Abbildung T geben und die Abbildung

$$r(x) = \begin{cases} \frac{f(x)-f(a)-T(x-a)}{||x-a||} & \text{für } x \neq a \\ 0 & \text{für } x = a \end{cases}$$

muss in a stetig sein, d.h. es muss gelten (diese Bedingung ist unabhängig von der Norm!)

$$\lim_{x \to a} r(x) = 0.$$

Wie man so ein T findet, werden wir gleich studieren. Zunächst zeigen wir:

Satz 5.1 *Zu der Abbildung $f\colon U \to \mathbb{R}^m$ gibt es höchstens eine lineare Abbildung $T\colon \mathbb{R}^n \to \mathbb{R}^m$ und eine in a stetige Abbildung $r\colon U \to \mathbb{R}^m$, so dass für alle $x \in U$ gilt*

$$f(x) = f(a) + T(x-a) + r(x)||x-a||$$

und $r(a) = 0$.

Beweis. Angenommen, es gäbe lineare Abbildungen T_1, T_2 mit entsprechenden Abbildung r_1, r_2, so dass

$$f(x) = f(a) + T_1(x-a) + r_1(x)||x-a||,$$

$$f(x) = f(a) + T_2(x-a) + r_2(x)||x-a||,$$

also

$$(T_1 - T_2)(x-a) = (r_2(x) - r_1(x))||x-a||$$

und

$$\lim_{x \to a} r_1(x) = \lim_{x \to a} r_2(x) = 0.$$

Wähle nun $h \in \mathbb{R}^n$ beliebig und betrachte x mit $x - a = th$, $t > 0$ genügend klein. Dann gilt

$$(T_1 - T_2)h = (r_2(a+th) - r_1(a+th))||h||.$$

Wegen $\lim_{t \to 0}(r_2(a+th) - r_1(a+th)) = 0$ folgt daraus

$$(T_1 - T_2)h = 0 \text{ für alle } h \in \mathbb{R}^n.$$

Also folgt $T_1 = T_2$. □

Definition Ist f in $a \in U$ differenzierbar, gilt also

$$f(x) = f(a) + T(x-a) + r(x)||x-a||$$

mit $\lim_{x \to a} r(x) = 0$, so heißt die lineare Abbildung T die *Ableitung* von f in a. Bezeichnung: $T = f'(a)$.

Nun betrachten wir zunächst den Spezialfall $m = 1$, d. h. $f\colon U \to \mathbb{R}$. In diesem Fall ist die Ableitung eine Linearform; sie wird durch einen Vektor, den *Gradienten*, $\operatorname{grad} f(a) \in \mathbb{R}^n$ unter Verwendung des Skalarprodukts gegeben durch

$$f'(a)h = \langle \operatorname{grad} f(a), h \rangle$$

für alle $h \in \mathbb{R}^n$. Nach Definition gilt

$$f(a+h) - f(a) = \langle \operatorname{grad} f(a), h \rangle + r(a+h)||h||$$

mit
$$\lim_{h\to 0} r(a+h) = 0.$$
Für hinreichend kleines $||h||$ ist der Zuwachs $f(a+h) - f(a)$ daher durch $\langle \operatorname{grad} f(a), h\rangle$ gut wiedergegeben. Es ergibt sich:

Der Gradient $\operatorname{grad} f(a)$ *steht senkrecht auf der Niveaufläche* $f(x) = const.$ *durch* a *und zeigt in die Richtung des größten Anstiegs der Funktion* f.

Denn diejenigen Vektoren h, die orthogonal zu $\operatorname{grad} f(a)$ sind, führen zu Nachbarpunkten von a, in denen sich die Funktionswerte nur wenig von $f(a)$ unterscheiden. Andererseits nimmt das Skalarprodukt $\langle \operatorname{grad} f(a), h\rangle$ als Funktion von h seinen größten Wert an, wenn h ein positives Vielfaches von $\operatorname{grad} f(a)$ ist.

Eine weitere Deutung des Gradienten: Betrachte Graph Γ_f der Funktion $f\colon U \to \mathbb{R}$ als Hyperfläche im $\mathbb{R}^{n+1}$:
$$\Gamma_f = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ z \end{pmatrix} \in \mathbb{R}^{n+1} \;\middle|\; z = f(x_1, \ldots, x_n) \right\}$$
In der Umgebung von $a \in U$ wird die Hyperfläche durch die Hyperebene
$$T_a\Gamma_f = \left\{ \begin{pmatrix} x \\ z \end{pmatrix} \in \mathbb{R}^{n+1} \;\middle|\; z = f(a) + \langle \operatorname{grad} f(a), x - a\rangle \right\}$$
approximiert: dies ist die *Tangentialhyperebene* an Γ_f in a.

Für den Vektor $\begin{pmatrix} -\operatorname{grad} f(a) \\ 1 \end{pmatrix}$ gilt
$$\begin{aligned} &\left\langle \begin{pmatrix} -\operatorname{grad} f(a) \\ 1 \end{pmatrix}, \begin{pmatrix} x - a \\ z - f(a) \end{pmatrix} \right\rangle \\ = \quad & -\langle \operatorname{grad} f(a), x - a\rangle + z - f(a) = 0 \text{ für alle } \begin{pmatrix} x \\ z \end{pmatrix} \in T_a\Gamma_f. \end{aligned}$$
Also ist $\begin{pmatrix} -\operatorname{grad} f(a) \\ 1 \end{pmatrix}$ der Normalenvektor an $T_a\Gamma_f$.

Anderer Speziallfall: $n = 1$, m beliebig. In diesem Fall ist die Ableitung eine lineare Abbildung von $\mathbb{R}$ nach $\mathbb{R}^m$. Sie wird durch einen Vektor des $\mathbb{R}^m$ gegeben. Damit erhalten wir die alte Definition aus §4.

Wir bestimmt man nun die lineare Abbildung $f'(a)$? Die lineare Abbildung $T = f'(a)$ ist festgelegt, wenn man ihre Bildvektoren für eine Basis $v_1, v_2, \ldots, v_n$ des $\mathbb{R}^n$ kennt. Wir wollen zunächst für einen beliebigen Vektor $v \in \mathbb{R}^n$ mit $||v|| = 1$ den Bildvektor $T(v)$ berechnen: Setze dazu $x = a + tv$ ein:
$$f(a+tv) = f(a) + T(tv) + r(a+tv)|t| \cdot ||v||.$$
Es folgt für $t \neq 0$:
$$T(v) = \frac{f(a+tv) - f(a)}{t} - r(a+tv)\frac{|t|}{t},$$
also
$$T(v) = \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t} =: D_v f(A).$$

Den rechts stehenden Grenzwert nennt man die *Richtungsableitung* $D_v f(a)$ von f in a bezüglich v.

Der Vektor $D_v f(a)$ ist der Tangentialvektor zum Zeitpunkt $t = 0$ der Kurve $f(a + tv)$ (Skizze!).

Die lineare Abbildung T ist also durch die Richtungsableitungen bezüglich der Vektoren $v_1, \ldots, v_n$ einer Basis des $\mathbb{R}^n$ festgelegt. Wählt man speziell die Standardbasis $e_1, \ldots, e_n$ des $\mathbb{R}^n$, so heißen die betreffenden Richtungsableitungen auch *partielle Ableitungen*.

Statt $D_{e_i} f(a)$ schreibt man auch

$$D_i f(a), \quad \frac{\partial f}{\partial x_i}, \quad f_{x_i}.$$

Es gilt also:

$$D_i f(a) := \lim_{t \to 0} \frac{f(a + te_i) - f(a)}{t}.$$

Die partiellen Ableitungen einer Funktion $f \colon U \to \mathbb{R}$ kann man als gewöhnliche Ableitungen von Funktionen einer Veränderlichen interpretieren: Sei

$$a = (a_1, \ldots, a_n) \in U.$$

Für $i = 1, \ldots, n$ betrachten wir die Funktionen (Schnitte durch Graphen von f)

$$\xi \mapsto f_i(\xi) = f(a_1, \ldots, a_{i-1}, \xi, a_{i+1}, \ldots, a_n)$$

Es gilt

$$D_i f(a) = \lim_{h \to 0} \frac{f_i(a_i + h) - f_i(a_i)}{h} = f_i'(a_i)$$

Die partielle Ableitung bezüglich der i-ten Koordinatenrichtung ist also nichts anderes als die gewöhnliche Ableitung bezüglich der i-ten Variablen bei Festhaltung der übrigen $n - 1$ Veränderlichen.

Hat man nun eine beliebige differenzierbare Abbildung $f \colon U \to \mathbb{R}^m$, so existieren alle partiellen Ableitungen und es gilt für $h = \sum_{i=1}^n h_i e_i$:

$$f'(a)(h) = \sum_{i=1}^n h_i f'(a)(e_i) = \sum_{i=1}^n h_i D_i f(a).$$

Die Ableitung $f'(a)$ ist also durch die partiellen Ableitungen $D_i f(a)$ festgelegt.

Stellt man weiter die Vektoren $D_i f(a)$ mittels der Standardbasis des $\mathbb{R}^n$ dar, d. h. zerlegt

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix},$$

so ergibt sich die folgende Matrixdarstellung von $f'(a)$:

$$Df(a) := \begin{pmatrix} D_1 f_1(a) & D_2 f_1(a) & \cdots & D_n f_1(a) \\ D_1 f_2(a) & D_2 f_2(a) & \cdots & D_n f_2(a) \\ \vdots & \vdots & & \vdots \\ D_1 f_m(a) & D_2 f_m(a) & \cdots & D_n f_m(a) \end{pmatrix}.$$

Diese $m \times n$-Matrix heißt *Jacobi-Matrix* der differenzierbaren Abbildung f in a (oder *Funktionalmatrix*).

Für den Gradienten von f in a gilt

$$\operatorname{grad} f(a) = (D_1 f(a), \dots, D_n f(a)) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right).$$

Demnach *Vorgehen beim Nachweis der Differenzierbarkeit:*

1) Untersuche zunächst, ob partielle Ableitungen $D_i f_j(a)$ existieren.

2) Falls diese existieren, setze $T = (D_i f_j(a))$ und prüfe, ob

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - Th}{||h||} = 0$$

gilt.

Warnung Allein aus der Existenz der partiellen Ableitungen folgt die Differenzierbarkeit noch nicht, siehe das folgende Beispiel (4).

Beispiel 5.1 (1)

$$\begin{array}{rccc} f\colon & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \mapsto & \begin{pmatrix} x_1^2 - x_2^2 \\ 2x_1 x_2 \end{pmatrix} \end{array}$$

$$T = \begin{pmatrix} 2a_1 & -2a_2 \\ 2a_2 & 2a_1 \end{pmatrix}$$

Mit diesem T erhält man für $r(a+h)$:

$$\begin{aligned} r(a+h) &= \begin{pmatrix} r_1(a+h) \\ r_2(a+h) \end{pmatrix}, \\ r_1(a+h) &= \frac{(a_1+h_1)^2 - (a_2+h_2)^2 - a_1^2 + a_2^2 - 2a_1 h_1 + 2a_2 h_2}{||h||} \\ &= \frac{h_1^2 - h_2^2}{||h||} \\ r_2(a+h) &= \frac{2(a_1+h_1)(a_2+h_2) - 2a_1 a_2 - 2a_2 h_1 - 2a_1 h_2}{||h||} \\ &= \frac{2h_1 h_2}{||h||} \end{aligned}$$

Verwendet man die Maximumsnorm, so ergibt sich

$$|r_1(a+h)| \leq 2||h||, \quad |r_2(a+h)| \leq 2||h||,$$

d.h. $\lim_{h \to 0} r_1(a+h) = \lim_{h \to 0} r_2(a+h) = 0$. Also ist f differenzierbar in a mit

$$Df(a) = \begin{pmatrix} 2a_1 & -2a_2 \\ 2a_2 & 2a_1 \end{pmatrix}$$

(2) Es sei f eine affine Abbildung vom $\mathbb{R}^n$ nach $\mathbb{R}^m$, d. h.

$$f(x) = Ax + c \quad (x \in \mathbb{R}^n)$$

mit $A \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, $c \in \mathbb{R}^m$ konstant.

Dann gilt

$$\begin{aligned} f(a+h) &= A(a+h) + c \\ &= Aa + Ah + c \\ &= Aa + c + Ah \\ &= f(a) + Ah. \end{aligned}$$

In diesem Fall ist also $r(a+h) = 0$ für alle $h \in \mathbb{R}^n$, also f differenzierbar in jedem $a \in \mathbb{R}^n$ und

$$Df(a) = A.$$

(3) Es sei $g\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ eine bilineare Abbildung (nicht notwendig symmetrisch), $f\colon \mathbb{R}^n \to \mathbb{R}^m$ definiert durch

$$f(x) = g(x,x) \text{ für alle } x \in \mathbb{R}^n.$$

Dann ist f in jedem $a \in \mathbb{R}^n$ differenzierbar und es gilt

$$f'(a)(h) = g(a,h) + g(h,a).$$

Beweis.

$$\begin{aligned} f(a+h) &= g(a+h, a+h) \\ &= g(a,a) + g(a,h) + g(h,a) + g(h,h) \\ &= f(a) + f'(a)(h) + g(h,h). \end{aligned}$$

Es bleibt zu zeigen:

$$\lim_{h\to 0} \frac{g(h,h)}{||h||} = 0.$$

Es sei

$$||g|| := \sup_{||x||=1} ||g(x,x)||$$

Es folgt

$$\left|\left| g\left(\frac{h}{||h||}, \frac{h}{||h||}\right)\right|\right| \leq ||g||,$$

also

$$||g(h,h)|| \leq ||g|| \cdot ||h||^2.$$

Daher gilt

$$\lim_{h\to 0} \frac{g(h,h)}{||h||} = 0.$$

Spezialfall: Für $f(x) = \langle x, x\rangle$ ergibt sich

$$f'(a)(h) = \langle \operatorname{grad} f(a), h\rangle = \langle 2a, h\rangle,$$

also

$$\operatorname{grad} f(a) = 2a.$$

□

(4) Es sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x_1, x_2) = \begin{cases} 0 & \text{für } (x_1, x_2) = (0, 0) \\ \frac{x_1 x_2}{\sqrt{x_1^2 + x_2^2}} & \text{für } (x_1, x_2) \neq (0, 0) \end{cases}.$$

Ist f in $a = (0, 0)$ differenzierbar?

Wegen

$$f(x_1, 0) = 0 \text{ für alle } x_1 \in \mathbb{R}$$

$$f(0, x_2) = 0 \text{ für alle } x_2 \in \mathbb{R}$$

gilt

$$D_1 f(0) = 0, \quad D_2 f(0) = 0,$$

d.h. die partiellen Ableitungen existieren. Wäre nun f an der Stelle $a = 0$ differenzierbar, so müsste T die Nullabbildung sein. Aber

$$r(a + h) = \frac{h_1 h_2}{\sqrt{h_1^2 + h_2^2}} \cdot \frac{1}{||h||} = \frac{h_1 h_2}{h_1^2 + h_2^2}$$

(mit euklidischer Norm). Setzt man $h_1 = h_2$, so sieht man, dass $\lim_{h \to 0} r(a + h) = 0$ nicht gelten kann. Deshalb ist f nicht in 0 differenzierbar.

Satz 5.2 *Ist $f\colon U \to \mathbb{R}^m$ in a differenzierbar, so ist f dort stetig.*

Beweis. Aus

$$f(x) = f(a) + f'(a)(x - a) + r(x)||x - a||$$

folgt

$$||f(x) - f(a)|| \leq ||f'(a)|| \cdot ||x - a|| + ||r(x)|| \cdot ||x - a||.$$

Wegen $\lim_{x \to a} r(x) = 0$ ergibt sich hieraus weiter mit einer geeigneten Konstanten c:

$$||f(x) - f(a)|| \leq c||x - a||.$$

Hieraus folgt die Stetigkeit von f. □

Wir haben gesehen, dass aus der Existenz der partiellen Ableitungen allein noch nicht die Differenzierbarkeit folgt. Es gilt aber:

Satz 5.3 *Wenn die Funktion $f\colon U \to \mathbb{R}$ sämtliche partiellen Ableitungen $D_1 f, \ldots, D_n f$ besitzt und diese an der Stelle $a \in U$ stetig sind, dann ist f in a differenzierbar.*

Beweis. Wir haben zu zeigen, dass

$$\lim_{h \to 0} \frac{f(a + h) - f(a) - \sum_{i=1}^{n} D_i f(a) h_i}{||h||} = 0$$

gilt. Es sei $h \in \mathbb{R}^n$ ein Vektor mit $a + h \in U$. Setze $z^{(0)} = a$, $z^{(n)} = a + h$ und führe Zwischenpunkte ein durch

$$\begin{array}{ccc} z^{(1)} - z^{(0)} & = & h_1 e_1 \\ z^{(2)} - z^{(1)} & = & h_2 e_2 \\ \vdots & \vdots & \vdots \\ z^{(n)} - z^{(n-1)} & = & h_n e_n. \end{array}$$

Dann gilt

$$f(a+h)-f(a)=(f(z^{(n)})-f(z^{(n-1)}))+\cdots+(f(z^{(1)})-f(z^{(0)})).$$

Die Punkte $z^{(j-1)}$ und $z^{(j)}$ unterscheiden sich nur in der j-ten Koordinate. Nach dem Mittelwertsatz für differenzierbare Funktionen einer Variablen gibt es deshalb ein $\theta_j \in [0,1]$, so dass für

$$\xi^{(j)} = z^{(j-1)} + \theta_j h_j e_j$$

gilt:

$$f(z^{(j)}) - f(z^{(j-1)}) = D_j f(\xi^{(j)}) h_j.$$

Daraus folgt

$$f(a+h)-f(a)=\sum_{j=1}^{n} D_j f(\xi^{(j)}) h_j.$$

Damit erhalten wir

$$f(a+h)-f(a)-\sum_{j=1}^{n} D_j f(a) h_j = \sum_{j=1}^{n} (D_j f(\xi^{(j)}) - D_j f(a)) h_j,$$

$$\begin{aligned}\left| f(a+h)-f(a)-\sum_{j=1}^{n} D_j f(a) h_j \right| &\le \sum_{j=1}^{n} |D_j f(\xi^{(j)}) - D_j f(a)|\,|h_j| \\ &\le ||h|| \cdot \sum_{j=1}^{n} |D_j f(\xi^{(j)}) - D_j f(a)|,\end{aligned}$$

also auch

$$\frac{\left| f(a+h)-f(a)-\sum_{j=1}^{n} D_j f(a) h_j \right|}{||h||} \le \sum_{j=1}^{n} |D_j f(\xi^{(j)}) - D_j f(a)|.$$

Wegen der Stetigkeit von $D_1 f, \ldots, D_n f$ folgt die Behauptung. □

Aufgabe 5.1 Man zeige, dass die Funktion $f\colon \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$, $f(x) = (x_1^2 + \cdots + x_n^2)^s$ ($s \in \mathbb{R}$ fest), überall differenzierbar ist und berechne die Ableitung.

Wir wollen Rechenregeln für differenzierbare Abbildungen herleiten.

Satz 5.4 *Die Abbildungen $f\colon U \to \mathbb{R}^m$ und $g\colon U \to \mathbb{R}^m$ seien beide in $a \in U$ differenzierbar. Dann sind auch $f+g$ und λf ($\lambda \in \mathbb{R}$) in a differenzierbar und es gilt*

$$(f+g)'(a) = f'(a) + g'(a)$$
$$(\lambda f)'(a) = \lambda f'(a)$$

Beweis. Nach Voraussetzung gilt

$$f(a+h) = f(a) + f'(a)(h) + r_1(a+h)||h|| \text{ mit } \lim_{h\to 0} r_1(a+h) = 0,$$

$$g(a+h) = g(a) + g'(a)(h) + r_2(a+h)||h|| \text{ mit } \lim_{h\to 0} r_2(a+h) = 0.$$

Addition ergibt

$$(f+g)(a+h) = (f+g)(a) + (f'(a)+g'(a))(h) + (r_1+r_2)(a+h)||h||$$

und

$$\lim_{h\to 0}(r_1+r_2)(a+h) = \lim_{h\to 0} r_1(a+h) + \lim_{h\to 0} r_2(a+h) = 0.$$

Entsprechend erhält man die zweite Aussage, indem man die erste Gleichung mit λ multipliziert. □

Die Produktregel beweisen wir für Funktionen.

Satz 5.5 *Die Funktionen $f\colon U \to \mathbb{R}$ und $g\colon U \to \mathbb{R}$ seien beide in $a \in U$ differenzierbar. Dann ist auch $f \cdot g$ in a differenzierbar und es gilt*

$$\operatorname{grad} fg(a) = f(a) \operatorname{grad} g(a) + g(a) \operatorname{grad} f(a).$$

Beweis. Nach Voraussetzung gilt

$$f(a+h) = f(a) + \langle \operatorname{grad} f(a), h\rangle + r_1(a+h)||h|| \text{ mit } \lim_{h\to 0} r_1(a+h) = 0,$$

$$g(a+h) = g(a) + \langle \operatorname{grad} g(a), h\rangle + r_2(a+h)||h|| \text{ mit } \lim_{h\to 0} r_2(a+h) = 0.$$

Ausmultiplizieren ergibt

$$\begin{aligned}(fg)(a+h) = (fg)(a) &+ f(a)\langle \operatorname{grad} g(a), h\rangle + g(a)\langle \operatorname{grad} f(a), h\rangle \\ &+ \langle \operatorname{grad} f(a), a\rangle\langle \operatorname{grad} g(a), h\rangle \\ &+ (g(a) + \langle \operatorname{grad} g(a), h\rangle) r_1(a+h)||h|| \\ &+ (f(a) + \langle \operatorname{grad} f(a), h\rangle) r_2(a+h)||h|| \\ &+ r_1(a+h) r_2(a+h)||h||^2.\end{aligned}$$

Wegen

$$\langle \operatorname{grad} f(a), h\rangle\langle \operatorname{grad} g(a), h\rangle \leq ||\operatorname{grad} f(a)||\, ||h||\, ||\operatorname{grad} g(a)||\, ||h||$$

folgt, dass das Fehlerglied in der Form $r(a+h)||h||$, $\lim_{h\to 0} r(a+h) = 0$, darstellbar ist. □

Bemerkung 5.1 Es sei $f\colon U \to \mathbb{R}$ in a differenzierbar und $f(a) \neq 0$. Dann kann man zeigen (Beweis!):

$$\operatorname{grad} \frac{1}{f}(a) = -\frac{1}{(f(a))^2} \operatorname{grad} f(a).$$

Wie sieht die entsprechende Aussage für $\operatorname{grad} \frac{g}{f}(a)$ aus?

Satz 5.6 (Kettenregel) *Es sei $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen; $f\colon U \to \mathbb{R}^m$ in $a \in U$, $g\colon V \to \mathbb{R}^p$ in $b \in V$ differenzierbar; $f(U) \subset V$ und $b = f(a)$.*

Dann ist $g \circ f$ in a differenzierbar und es gilt

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

Beweis. Nach Voraussetzung gilt

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + r_1(a+h)||h|| \quad \text{mit } \lim_{h\to 0} r_1(a+h) = 0, \\ g(b+k) &= g(b) + g'(b)k + r_2(b+k)||k|| \quad \text{mit } \lim_{k\to 0} r_2(b+k) = 0. \end{aligned}$$

Wir setzen

$$k := f'(a)h + r_1(a+h)||h||$$

und setzen dies in die zweite Gleichung ein:

$$\begin{aligned} (g\circ f)(a+h) &= g(f(a+h)) \\ &= g(f(a)+k) \\ &= g(b) + g'(b)f'(a)h + g'(b)r_1(a+h)||h|| + r_2(b+k)||k||. \end{aligned}$$

Wir setzen für $h \neq 0$

$$\begin{aligned} R(a+h) &:= \frac{1}{||h||}(g'(b)r_1(a+h)||h|| + r_2(b+k)||k||) \\ &= g'(b)r_1(a+h) + \frac{||k||}{||h||} r_2(b+k). \end{aligned}$$

Aus der Abschätzung

$$||k|| = ||f'(a)h + r_1(a+h)||h|| \, || \leq ||h||(||f'(a)|| + ||r_1(a+h)||)$$

folgt

$$\lim_{h\to 0} R(a+h) = 0.$$

□

Aus der Kettenregel ergibt sich für die Jacobi-Matrizen von g, f und $H = g\circ f$:

$$\left.\begin{pmatrix} D_1H_1 & \cdots & D_nH_1 \\ \vdots & & \vdots \\ D_1H_p & \cdots & D_nH_p \end{pmatrix}\right|_a = \left.\begin{pmatrix} D_1g_1 & \cdots & D_mg_1 \\ \vdots & & \vdots \\ D_1g_p & \cdots & D_mg_p \end{pmatrix}\right|_{f(a)} \left.\begin{pmatrix} D_1f_1 & \cdots & D_nf_1 \\ \vdots & & \vdots \\ D_1f_m & \cdots & D_nf_m \end{pmatrix}\right|_a$$

d.h. es gilt

$$\frac{\partial H_j}{\partial x_i}(a) = \sum_{k=1}^{m} \frac{\partial g_j}{\partial y_k}(f(a)) \cdot \frac{\partial f_k}{\partial x_i}(a).$$

Korollar 5.1 *Es sei* $\gamma\colon [a,b] \to U \subset \mathbb{R}^n$ *eine differenzierbare Kurve und* $f\colon U \to \mathbb{R}$ *eine differenzierbare Funktion. Dann gilt für die Funktion* $F := f\circ\gamma\colon [a,b] \to \mathbb{R}$

$$F'(t) = f'(\gamma(t))(\gamma'(t)) = \langle \operatorname{grad} f(\gamma(t)), \gamma'(t)\rangle.$$

Wir wollen nun die Aussage präzisieren, dass der Gradient einer differenzierbaren Funktion $f : U \to \mathbb{R}$ in einem Punkt $a \in U$ senkrecht auf der Niveaufläche $f(x) = f(a)$ steht.

Satz 5.7 *Es sei $U \subset \mathbb{R}^n$ eine offene Menge, $f : U \to \mathbb{R}$ eine differenzierbare Funktion. Der Punkt $a \in U$ liege auf der Niveaufläche $N_c = \{x \in \mathbb{R}^n \mid f(x) = c\}$, $c \in \mathbb{R}$. Dann gilt für jede differenzierbare Kurve $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ $(\varepsilon > 0)$ mit $\gamma((-\varepsilon, \varepsilon)) \subset N_c$ und $\gamma(0) = a$:*

$$\langle \operatorname{grad} f(a), \gamma'(0) \rangle = 0.$$

Beweis. Es sei $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ eine differenzierbare Kurve mit $\gamma((-\varepsilon, \varepsilon)) \subset N_c$. Dann gilt $f(\gamma(t)) = c$ für alle $t \in (-\varepsilon, \varepsilon)$. Also ist die Funktion $F := f \circ \gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}$ konstant. Nach dem Spezialfall der Kettenregel (Korollar 5.1) gilt

$$0 = F'(0) = f'(\gamma(0))(\gamma'(0)) = \langle \operatorname{grad} f(a), \gamma'(0) \rangle.$$

□

Aufgabe 5.2 Eine Funktion $f \colon \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$ heißt *positiv homogen von der Ordnung* α, wenn für alle $t > 0$ gilt:

$$f(tx) = t^\alpha \cdot f(x).$$

Man zeige

$$\langle \operatorname{grad} f(x), x \rangle = \alpha f(x)$$

(*Eulersche Relation für homogene Funktionen*).

Wir wollen nun als eine weitere Anwendung der Kettenregel die Ableitung der Umkehrabbildung einer bijektiven Abbildung bestimmen.

Es sei $f \colon U \to V \subset \mathbb{R}^n$, V offen, bijektiv und in $a \in U$ differenzierbar. Setzt man voraus, dass $f^{-1} \colon V \to U$ in $b = f(a)$ differenzierbar ist, so kann man die Kettenregel auf $f^{-1} \circ f = id_U$ anwenden:

$$(f^{-1})'(f(a)) \circ f'(a) = \mathrm{id}_{\mathbb{R}^n}.$$

Es folgt also, dass $f'(a)$ umkehrbar ist und für die Ableitung von f^{-1} gilt:

$$(f^{-1})'(f(a)) = (f'(a))^{-1}.$$

Die lineare Abbildung $f'(a)$ ist genau dann umkehrbar, wenn ihre Determinante $\det f'(a)$ (*Jacobi-Determinante* oder *Funktionaldeterminante*) von Null verschieden ist.

Wir zeigen nun, dass es für die Differenzierbarkeit von f^{-1} im Punkte $f(a)$ schon genügt, dass $\det f'(a) \neq 0$ ist und f^{-1} in $f(a)$ stetig ist.

Satz 5.8 *Es seien $U, V \subset \mathbb{R}^n$ offen, $f \colon U \to V$ bijektiv, in $a \in U$ differenzierbar und es gelte $\det f'(a) \neq 0$. Weiter sei $f^{-1} \colon V \to U$ in $b = f(a)$ stetig.*

Dann ist $f^{-1} \colon V \to U$ in $b = f(a)$ differenzierbar und es gilt

$$(f^{-1})'(f(a)) = (f'(a))^{-1}.$$

Beweis. Nach Voraussetzung gilt

$$f(a + h) = f(a) + f'(a)h + r(a + h)||h|| \quad \text{mit } \lim_{h \to 0} r(a + h) = 0.$$

Es sei $k \in \mathbb{R}^n$ so gewählt, dass $b + k \in V$. Mit

$$h := f^{-1}(b + k) - f^{-1}(b)$$

folgt daraus

$$b + k = b + f'(a)f^{-1}(b + k) - f'(a)f^{-1}(b) + r(f^{-1}(b + k))||h||.$$

Wegen $\det f'(a) \neq 0$ existiert $(f'(a))^{-1}$. Wenden wir $(f'(a))^{-1}$ auf diese Gleichung an, so folgt

$$f^{-1}(b + k) = f^{-1}(b) + (f'(a))^{-1}k - (f'(a))^{-1}r(f^{-1}(b + k))||h||.$$

Wir setzen für $k \neq 0$

$$r^*(b + k) = -(f'a)^{-1}r(f^{-1}(b + k))\frac{||h||}{||k||}.$$

Damit ergibt sich mit $c := ||(f'(a))^{-1}||$

$$||f^{-1}(b + k) - f^{-1}(b)|| \leq c||k|| + ||r^*(b + k)||||k||.$$

Wegen $\lim_{h\to 0} r(a + h) = 0$ gibt es ein $\eta > 0$, so dass für $||h|| \leq \eta$ gilt:

$$||r(a + h)|| \leq \frac{1}{2c}.$$

Wegen der Stetigkeit von f^{-1} in b gibt es ein $\delta > 0$, so dass für $||k|| \leq \delta$ gilt

$$||h|| = ||f^{-1}(b + k) - f^{-1}(b)|| \leq \eta.$$

Daraus folgt für $||k|| \leq \delta$

$$||r^*(b + k)||||k|| \leq \frac{1}{2}||f^{-1}(b + k) - f^{-1}(b)||.$$

Insgesamt ergibt sich für $||k|| \leq \delta$

$$||h|| = ||f^{-1}(b + k) - f^{-1}(b)|| \leq 2c||k||.$$

Damit erhält man für $0 < ||k|| \leq \delta$

$$||r^*(b + k)|| \leq 2c^2||r(f^{-1}(b + k))||.$$

Aus der Stetigkeit von f^{-1} in b folgt damit

$$\lim_{k\to 0} r^*(b + k) = 0.$$

. □

Beispiel 5.2 (Polarkoordinatenabbildung)

$$\begin{array}{rlcl} f: & \mathbb{R}_+^* \times (0, 2\pi) & \longrightarrow & \mathbb{R}^2 \backslash (\mathbb{R}_+ \times \{0\}) \\ & \begin{pmatrix} r \\ \varphi \end{pmatrix} & \longmapsto & \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix} \end{array}$$

$$
\begin{aligned}
Df\begin{pmatrix} r \\ \varphi \end{pmatrix} &= \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} \\
\det Df\begin{pmatrix} r \\ \varphi \end{pmatrix} &= \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r(\cos^2\varphi + \sin^2\varphi) = r \neq 0 \\
(f^{-1})'\begin{pmatrix} x \\ y \end{pmatrix} &= \left(Df\begin{pmatrix} r \\ \varphi \end{pmatrix}\right)^{-1} \\
&= \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\frac{1}{r}\sin\varphi & \frac{1}{r}\cos\varphi \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix}
\end{aligned}
$$

Bisher haben wir die Ableitung einer differenzierbaren Abbildung $f\colon U \to \mathbb{R}^m$ an einer festen Stelle $a \in U$ betrachtet.

Definition Eine Abbildung $f\colon U \to \mathbb{R}^m$ heißt *differenzierbar*, wenn f in allen Punkten $a \in U$ differenzierbar ist.

Die Ableitung definiert dann eine Abbildung

$$
\begin{array}{rccl} f'\colon & U & \to & \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \\ & a & \mapsto & f'(a) \end{array}
$$

Der Vektorraum $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ ist isomorph zu $\mathbb{R}^{n\cdot m}$ (man fasse lineare Abbildungen als $n \times m$-Matrizen auf). Am Ende von §3 haben wir diesen Vektorraum mit einer natürlichen Norm versehen. Also können wir f' auffassen als Abbildung

$$
\begin{array}{rccl} f'\colon & U & \to & \mathbb{R}^{nm} \\ & a & \mapsto & f'(a) \end{array}.
$$

Definition Eine Abbildung $f\colon U \to \mathbb{R}^m$ heißt *stetig differenzierbar* (in Zeichen $f \in C^1(U)$), falls f differenzierbar und die Abbildung f' stetig ist.

Wenn f eine stetig differenzierbare umkehrbare Abbildung $\mathbb{R}^n \to \mathbb{R}^m$ ist, dann braucht die Umkehrabbildung nicht differenzierbar zu sein (z. B. $n = 1; x \leftrightarrow x^3$).

Definition Es sei $f\colon U \to \mathbb{R}^n$ stetig differenzierbar, injektiv und die Bildmenge $f(U)$ sei offen. Ist dann $f^{-1}\colon f(U) \to U$ ebenfalls stetig differenzierbar, so wird f (und auch f^{-1}) als *Diffeomorphismus* bezeichnet.

Beispiel 5.3 Die Polarkoordinatenabbildung ist ein Diffeomorphismus von dem Streifen $\mathbb{R}_+^* \times (0, 2\pi)$ auf die längs der positiven x-Achse aufgeschnittene Ebene

$$
\left\{ \begin{pmatrix} x \\ y \end{pmatrix} \;\middle|\; y \neq 0 \text{ oder } y = 0,\ x < 0 \right\}.
$$

Man sagt, dass diese beiden Mengen zueinander *diffeomorph* sind.

Wir wollen nun auch den MWS für differenzierbare Funktionen einer reellen Variablen verallgemeinern.

Satz 5.9 (Mittelwertsatz) *Es sei $U \subset \mathbb{R}^n$ offen, $a, b \in U$, so dass die Verbindungsstrecke von a und b*

$$\overline{ab} = \{(1-t)a + tb \mid 0 \le t \le 1\}$$

in U enthalten ist, $f\colon U \to \mathbb{R}$ differenzierbar. Dann gibt es einen Punkt $c \in \overline{ab}$, so dass gilt:

$$f(b) - f(a) = \langle \operatorname{grad} f(c), b - a\rangle$$

Beweis. Die Verbindungsstrecke von a und b wird durch die Kurve

$$\gamma\colon \begin{array}{lcl} [0,1] & \to & U \\ t & \mapsto & (1-t)a + tb \end{array}$$

beschrieben. Die Kurve γ ist differenzierbar und hat konstante Ableitung

$$\gamma'(t) = b - a.$$

Wende MWS in einer Variablen auf die Funktion $F = f \circ \gamma\colon [0,1] \to \mathbb{R}$ an: Demnach existiert $\vartheta \in [0,1]$, so dass

$$F(1) - F(0) = F'(\vartheta),$$

d. h. nach der Kettenregel

$$f(b) - f(a) = \langle \operatorname{grad} f(\gamma(\vartheta)), \gamma'(\vartheta)\rangle.$$

Mit $c := \gamma(\vartheta)$ folgt

$$f(b) - f(a) = \langle \operatorname{grad} f(c), b - a\rangle.$$

□

Bemerkung 5.2 Durch Umbenennung ergibt sich die Formulierung

$$f(x+h) - f(x) = \langle \operatorname{grad} f(x + \vartheta h), h\rangle \text{ mit } 0 \le \vartheta \le 1$$

falls $\{x + th \mid 0 \le t \le 1\} \subset U$.

Korollar 5.2 (Schrankensatz) *Ist $f\colon U \to \mathbb{R}$ differenzierbar, $\{x + th \mid 0 \le t \le 1\} \subset U$ und*

$$S := \sup_{t \in [0,1]} || \operatorname{grad} f(x + th)|| < \infty,$$

so gilt

$$|f(x+h) - f(x)| \le S||h||.$$

Beweis. Es gilt für ein ϑ, $0 \le \vartheta \le 1$

$$\begin{aligned} |f(x+h) - f(x)| &= |\langle \operatorname{grad} f(x + \vartheta h), h\rangle| \\ &\le || \operatorname{grad} f(x + \vartheta h)|| \, ||h|| \\ &\le S||h||. \end{aligned}$$

□

Eine Anwendung des Mittelwertsatzes ist – wie in Analysis I – eine Kennzeichnung der konstanten Funktionen.

Satz 5.10 *Es sei $U \subset \mathbb{R}^n$ offen und wegzusammenhängend, $f: U \to \mathbb{R}$ differenzierbar und* $\operatorname{grad} f(x) = 0$ *für alle $x \in U$. Dann ist f konstant.*

Beweis. Es sei $a, b \in U$ und

$$a_0 = a, a_1, a_2, \ldots, a_k = b$$

ein Streckenzug zwischen a, b, der ganz in U verläuft. Wir wenden den Mittelwertsatz auf $\overline{a_j a_{j+1}}$ an: Danach existiert $c_j \in \overline{a_j a_{j+1}}$ mit

$$f(a_{j+1}) - f(a_j) = \langle \operatorname{grad} f(c_j), a_{j+1} - a_j \rangle = 0,$$

also

$$f(a_{j+1}) = f(a_j), \quad j = 0, 1, \ldots, k-1.$$

Somit gilt auch $f(b) = f(a)$. □

Der Mittelwertsatz läßt sich nicht auf Abbildungen übertragen, da die Zwischenstellen für die Komponentenfunktionen $f_1, \ldots, f_m$ im Allgemeinen nicht dieselben sind. Es gilt

$$f_j(x + h) - f_j(x) = \langle \operatorname{grad} f_j(x + \vartheta_j h), h \rangle$$

mit $0 \leq \vartheta_j \leq 1$, $j = 1, \ldots, m$. Wendet man (bei beschränkter Ableitung) den Schrankensatz an, so erhält man

$$|f_j(x + h) - f_j(x)| \leq S_j ||h||, j = 1, \ldots, m.$$

Verwendet man im Zielraum $\mathbb{R}^m$ die Maximumsnorm und setzt

$$S = \max\{S_1, \ldots, S_m\},$$

so folgt

$$||f(x + h) - f(x)|| \leq S||h||.$$

Satz 5.11 (Schrankensatz für Abbildungen) *Ist $f: U \to \mathbb{R}^m$ differenzierbar und f' beschränkt, so existiert ein $S \in \mathbb{R}$ mit*

$$||f(x + h) - f(x)|| \leq S||h||$$

für alle genügend kleinen h.

Satz 5.12 *Es sei $U \subset \mathbb{R}^n$ offen und wegzusammenhängend, $f: U \to \mathbb{R}^m$ differenzierbar und $f'(x)$ die Nullabbildung für alle $x \in U$. Dann ist f konstant.*

Beweis. Wende Satz 5.10 auf die einzelnen Komponentenfunktionen an. □

6 Höhere Ableitungen, Taylorsche Formel, lokale Extrema

Wir wollen nun höhere Ableitungen definieren.

Es sei $U \subset \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$ eine Funktion, deren sämtliche partielle Ableitungen existieren und stetig sind. Nach Satz 5.3 ist eine solche Funktion dann differenzierbar. Wir nennen f eine *stetig differenzierbare* Funktion.

Sind alle partiellen Ableitung $D_i f\colon U \to \mathbb{R}$ selbst wieder (partiell) differenzierbar, so nennt man f *zweimal partiell differenzierbar*. Dann existieren die partiellen Ableitungen 2. Ordnung

$$D_j D_i f.$$

Definition Eine Funktion $f\colon U \to \mathbb{R}$ heißt $(k+1)$*-mal partiell differenzierbar*, wenn f k-mal partiell differenzierbar ist und alle partiellen Ableitungen k-ter Ordnung

$$D_{i_k} \cdots D_{i_2} D_{i_1} f\colon U \to \mathbb{R}$$

partiell differenzierbar sind.

Eine Funktion $f\colon U \to \mathbb{R}$ heißt *k-mal stetig differenzierbar*, wenn f k-mal partiell differenzierbar ist und alle partiellen Ableitungen der Ordnung $\leq k$ stetig sind.

Satz 6.1 (H. A. Schwarz) *Es sei $U \subset \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$ zweimal stetig differenzierbar. Dann gilt für alle $a \in U$ und alle $i,j = 1,\ldots,n$*

$$D_j D_i f(a) = D_i D_j(a).$$

Beweis. O. B. d. A. $n = 2, i = 1, j = 2, a = 0$. Statt (x_1, x_2) schreiben wir wieder (x, y). Es gibt ein $\delta > 0$, so dass

$$W_\delta := \{(x,y) \in \mathbb{R}^2 \mid |x| < \delta, |y| < \delta\} \subset U.$$

Für $|x|, |y| < \delta$ betrachte die Funktionen $F_y, G_x\colon (-\delta,\delta) \to \mathbb{R}$

$$F_y(x) := f(x,y) - f(x,0),$$

$$G_x(y) := f(x,y) - f(0,y).$$

Nach dem MWS aus Analysis I gibt es $\xi, \tilde{\eta}$ mit $|\xi| \leq |x|, |\tilde{\eta}| \leq |y|$ mit

$$F_y(x) - F_y(0) = F_y'(\xi)x,$$

$$G_x(y) - G_x(0) = G_x'(\tilde{\eta})y.$$

Nun gilt

$$F_y'(\xi) = D_1 f(\xi, y) - D_1 f(\xi, 0),$$
$$G_x'(\tilde{\eta}) = D_2 f(x, \tilde{\eta}) - D_2 f(0, \tilde{\eta}).$$

Nochmalige Anwendung des MWS, diesmal auf $y \mapsto D_1 f(\xi, y)$ bzw. $x \mapsto D_2 f(x, \tilde{\eta})$, liefert $\eta, \tilde{\xi}$ mit $|\eta| \leq |y|$, $|\tilde{\xi}| \leq |x|$ und

$$D_1 f(\xi, y) - D_1 f(\xi, 0) = D_2 D_1 f(\xi, \eta) y,$$

$$D_2 f(x, \tilde{\eta}) - D_2 f(0, \tilde{\eta}) = D_1 D_2 f(\tilde{\xi}, \tilde{\eta}) x.$$

Insgesamt ergibt sich also

$$f(x, y) - f(x, 0) - f(0, y) + f(0, 0) = D_2 D_1 f(\xi, \eta) xy,$$

$$f(x, y) - f(0, y) - f(x, 0) + f(0, 0) = D_1 D_2 f(\tilde{\xi}, \tilde{\eta}) xy.$$

Daraus folgt für $xy \neq 0$

$$D_2 D_1 f(\xi, \eta) = D_1 D_2 f(\tilde{\xi}, \tilde{\eta}).$$

((ξ, η) und $(\tilde{\xi}, \tilde{\eta})$ hängen von (x, y) ab!) Mit $(x, y) \to 0$ gilt auch

$$(\xi, \eta) \to 0 \text{ und } (\tilde{\xi}, \tilde{\eta}) \to 0.$$

Aus der Stetigkeit von $D_1 D_2 f$ und $D_2 D_1 f$ folgt damit

$$D_2 D_1 f(0, 0) = D_1 D_2 f(0, 0).$$

□

Korollar 6.1 *Es sei $U \subset \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$ eine k-mal stetig differenzierbare Funktion. Dann gilt*

$$D_{i_k} \cdots D_{i_2} D_{i_1} f = D_{i_{\pi(k)}} \cdots D_{i_{\pi(2)}} D_{i_{\pi(1)}} f$$

für jede Permutation π der Zahlen $1, \ldots, k$.

Beweis. Durch Induktion über k unter Verwendung der Tatsache, dass jede Permutation ein Produkt von Transpositionen ist. □

Notation Andere Schreibweisen:

$$D_j D_i f = \frac{\partial^2 f}{\partial x_j \partial x_i} = f_{x_i x_j}, \quad D_i D_i f = \frac{\partial^2 f}{\partial x_i^2} = f_{x_i x_i},$$

$$D_{i_k} \cdots D_{i_1} f = \frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}} \text{ usw.}$$

Für ein n-Tupel natürlicher Zahlen

$$\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$$

sei

$$|\alpha| := \alpha_1 + \cdots + \alpha_n,$$

$$\alpha! := \alpha_1! \alpha_2! \cdot \ldots \cdot \alpha_n!.$$

Ist f eine $|\alpha|$-mal stetig differenzierbare Funktion, so setzt man

$$D^\alpha f := D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, \quad D_i^{\alpha_i} = \underbrace{D_i \cdots D_i}_{\alpha_i\text{-mal}}$$

Für $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ sei

$$x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}.$$

Satz 6.2 *Es sei $U \subset \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$ eine k-mal stetig differenzierbare Funktion. Es sei $x \in U$ und $h \in \mathbb{R}^n$, so dass die Strecke $\{x + th \mid 0 \le t \le 1\}$ ganz in U liegt. Dann ist die Funktion*

$$\begin{array}{llll} g\colon & [0,1] & \to & \mathbb{R} \\ & t & \mapsto & f(x+th) \end{array}$$

k-mal stetig differenzierbar und es gilt

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

($\sum_{|\alpha|=k}$ bedeutet: Die Summe erstreckt sich über alle n-Tupel

$$\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$$

mit $|\alpha| = k$.)

Beweis. (a) Wir zeigen zunächst

$$g^{(k)}(t) = \sum_{i_1,\dots,i_k=1}^{n} D_{i_k} \cdots D_{i_1} f(x+th) h_{i_1} \cdots h_{i_k}.$$

Induktion über k: $k = 1$:

$$\begin{aligned} g'(t) &= \frac{d}{dt} f(x_1 + th_1, \dots, x_n + th_n) \\ &= \sum_{i=1}^{n} D_i f(x+th) h_i \quad \text{(Kettenregel)} \end{aligned}$$

Induktionsschritt $(k-1) \to k$:

$$\begin{aligned} g^{(k)}(t) &= \frac{d}{dt}\left(\sum_{i_1,\dots,i_{k-1}}^{n} D_{i_{k-1}} \cdots D_{i_1} f(x+th) h_{i_1} \cdots h_{i_{k-1}} \right) \\ &= \sum_{j=1}^{n} D_j \left(\sum_{i_1,\dots,i_{k-1}}^{n} D_{i_{k-1}} \cdots D_{i_1} f(x+th) h_{i_1} \cdots h_{i_{k-1}} \right) h_j \\ &= \sum_{i_1,\dots,i_k=1}^{n} D_{i_k} \cdots D_{i_1} f(x+th) h_{i_1} \cdots h_{i_k}. \end{aligned}$$

(b) Nun fassen wir gleiche Indizes zusammen:

$$\{i_1, \dots, i_k\} = \{\underbrace{1, \dots, 1}_{\alpha_1}; \underbrace{2, \dots, 2}_{\alpha_2}; \dots; \underbrace{n, \dots, n}_{\alpha_n}\}$$

Nach Korollar 6.1 dürfen wir die partiellen Ableitungen beliebig vertauschen. Also gilt

$$D_{i_k} \cdots D_{i_1} f(x+th) h_{i_1} \cdots h_{i_k} = D_1^{\alpha_1} \cdots D_n^{\alpha_n} f(x+th) h_1^{\alpha_1} \cdots h_n^{\alpha_n}.$$

Nun gibt es

$$\frac{k!}{\alpha_1!\alpha_2!\cdots\alpha_n!}$$

solcher k-Tupel $(i_1,\ldots,i_k)$, bei denen i genau α_i mal vorkommt. Deshalb folgt

$$\begin{aligned} g^{(k)}(t) &= \sum_{i_1,\ldots,i_k=1}^{n} D_{i_k}\cdots D_{i_1} f(x+th)h_{i_1}\cdots h_{i_k} \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha_1!\cdots\alpha_n!} D_1^{\alpha_1}\cdots D_n^{\alpha_n} f(x+th)h_1^{\alpha_1}\cdots h_n^{\alpha_n} \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha!} D^\alpha f(x+th)h^\alpha. \end{aligned}$$

□

Satz 6.3 (Taylor-Formel) *Es sei $U \subset \mathbb{R}^n$ offen, $x \in U$ und $h \in \mathbb{R}^n$, so dass die Strecke $\{x+th \mid 0 \le t \le 1\}$ ganz in U liegt, $f\colon U \to \mathbb{R}$ eine $(k+1)$-mal stetig differenzierbare Funktion. Dann existiert ein $\vartheta \in [0,1]$, so dass*

$$\boxed{f(x+h) = \sum_{|\alpha|\le k} \frac{D^\alpha f(x)}{\alpha!} h^\alpha + \sum_{|\alpha|=k+1} \frac{D^\alpha f(x+\vartheta h)}{\alpha!} h^\alpha}$$

Beweis. Betrachte die Funktion von Satz 6.2,

$$g\colon \begin{array}{ccc} [0,1] & \to & \mathbb{R} \\ t & \mapsto & f(x+th) \end{array}.$$

Nach Satz 6.2 ist g eine $(k+1)$-mal stetig differenzierbare Funktion. Wir können daher auf g die Taylor-Formel für Funktionen einer Veränderlichen anwenden: Danach existiert ein $\vartheta \in [0,1]$, so dass

$$g(1) = \sum_{m=0}^{k} \frac{g^{(m)}(0)}{m!} + \frac{g^{(k+1)}(\vartheta)}{(k+1)!}.$$

Nach Satz 6.2 gilt

$$\begin{aligned} \frac{g^{(m)}(0)}{m!} &= \sum_{|\alpha|=m} \frac{D^\alpha f(x)}{\alpha!} h^\alpha, \\ \frac{g^{(k+1)}(\vartheta)}{(k+1)!} &= \sum_{|\alpha|=k+1} \frac{D^\alpha f(x+\vartheta h)}{\alpha!} h^\alpha. \end{aligned}$$

□

Korollar 6.2 *Es sei $U \subset \mathbb{R}^n$ offen, $x \in U$ und $\delta > 0$, so dass die Kugel $B(x,\delta)$ ganz in U liegt, $f\colon U \to \mathbb{R}$ eine k-mal stetig differenzierbare Funktion.*

Dann gilt für alle $h \in \mathbb{R}^n$ mit $||h|| < \delta$

$$f(x+h) = \sum_{|\alpha|\le k} \frac{D^\alpha f(x)}{\alpha!} h^\alpha + r(x+h)||h||^k,$$

wobei $\lim_{h\to 0} r(x+h) = 0$.

Beweis. Nach Satz 6.3 gibt es ein von h abhängiges $\vartheta \in [0,1]$ mit

$$\begin{aligned} f(x+h) &= \sum_{|\alpha|\le k-1} \frac{D^\alpha f(x)}{\alpha!}h^\alpha + \sum_{|\alpha|=k} \frac{D^\alpha f(x+\vartheta h)}{\alpha!}h^\alpha \\ &= \sum_{|\alpha|\le k} \frac{D^\alpha f(x)}{\alpha!}h^\alpha + r(x+h)||h||^k \end{aligned}$$

wobei

$$r(x+h) := \sum_{|\alpha|=k} \frac{D^\alpha f(x+\vartheta h) - D^\alpha f(x)}{\alpha!||h||^k}h^\alpha \text{ für } h \neq 0.$$

Nun gilt

$$|r(x+h)| \le \sum_{|\alpha|=k} \frac{1}{\alpha!}|D^\alpha f(x+\vartheta h) - D^\alpha f(x)|.$$

Aus der Stetigkeit von $D^\alpha f$ folgt daher

$$\lim_{h\to 0} r(x+h) = 0.$$

□

Wir wollen die Formel von Korollar 6.2 näher betrachten. Setze dazu

$$P_m(h) := \sum_{|\alpha|=m} \frac{D^\alpha f(x)}{\alpha!}h^\alpha.$$

Das Polynom P_m ist ein homogenes Polynom m-ten Grades in $h = (h_1, \dots, h_n)$ und es gilt

$$f(x+h) = \sum_{m=0}^{k} P_m(h) + r(x+h)||h||^k.$$

Wir sehen uns die Polynome P_m für $m = 0, 1, 2$ näher an:

$m = 0$: Es gibt nur ein n-Tupel $\alpha \in \mathbb{N}^n$ mit $|\alpha| = 0$, nämlich $\alpha = (0, \dots, 0)$. Daher

$$P_0(h) = \frac{D^0 f(x)}{0!}h^0 = f(x).$$

P_0 ist also eine Konstante, der Funktionswert von f in x.

$m = 1$: $\alpha \in \mathbb{N}^n$ mit $|\alpha| = 1$:

$$\alpha = e_j = (0, \dots, 0, \underbrace{1}_{\substack{j\text{-te} \\ \text{Stelle}}}, 0, \dots, 0)$$

$$D^{ej} = D_j, \quad e_j! = 1, \quad h^{e_j} = h_j$$

$$\Rightarrow P_1(h) = \sum_{j=1}^{n} D_j f(x) h_j = \langle \operatorname{grad} f(x), h\rangle$$

Für eine einmal stetig differenzierbare Funktion lautet also die Formel von Korollar 6.2

$$f(x+h) = f(x) + \langle \operatorname{grad} f(x), h\rangle + r(x+h)||h||.$$

Dies ist genau die Bedingung aus der Definition der Differenzierbarkeit.

$m = 2$: $\alpha \in \mathbb{N}^n$ mit $|\alpha| = 2$:

$$2e_i = (0, \dots, 0, \underbrace{2}_{\substack{i\text{-te}\\ \text{Stelle}}}, 0, \dots, 0)$$

$$e_i + e_j = (0, \dots, \underbrace{1}_{\substack{i\text{-te}\\ \text{Stelle}}}, \dots, \underbrace{1}_{\substack{j\text{-te}\\ \text{Stelle}}}, \cdot, 0) \qquad (i < j)$$

$$D^{2e_i} = D_i^2, \quad (2e_i)! = 2, \quad h^{2e_i} = h_i^2$$

$$D^{e_i+e_j} = D_i D_j, \quad (e_i + e_j)! = 1, \quad h^{e_i+e_j} = h_i h_j \quad (i \neq j)$$

$$\begin{aligned} \Rightarrow P_2(h) &= \frac{1}{2} \sum_{i=1}^n D_i^2 f(x) h_i^2 + \sum_{i<j} D_i D_j f(x) h_i h_j \\ &= \frac{1}{2} \sum_{i=1}^n D_i^2 f(x) h_i^2 + \frac{1}{2} \sum_{i \neq j} D_i D_j f(x) h_i h_j \\ &= \frac{1}{2} \sum_{i,j=1}^n D_i D_j f(x) h_i h_j \end{aligned}$$

(Jetzt laufen i, j unabhängig voneinander.)

Das Polynom P_2 ist also eine quadratische Form mit der Matrix $(\frac{1}{2} D_i D_j f(x))$. Wir setzen

$$Q(h, h) := 2P_2(h) = \sum_{i,j=1}^n D_i D_j f(x) h_i h_j.$$

Definition Es sei $U \subset \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$ zweimal stetig differenzierbar. Dann heißt die $n \times n$-Matrix

$$(\operatorname{Hess} f)(x) := (D_i D_j f(x))_{1 \leq j \leq n}^{1 \leq i \leq n}$$

die *Hesse-Matrix* von f in $x \in U$. Dies ist eine symmetrische Matrix, da $D_i D_j f(x) = D_j D_i f(x)$ für alle $1 \leq i, j \leq n$ nach Satz 6.1.

Die quadratische Form

$$Q(h, h) = h^t A h, \quad A := (\operatorname{Hess} f)(x)$$

heißt die *Hesse-Form* von f an der Stelle $x \in U$.

Wir erhalten den folgenden Spezialfall von Korollar 6.2:

Korollar 6.3 *Es sei $U \subset \mathbb{R}^n$ offen, $x \in U$, $f\colon U \to \mathbb{R}$ zweimal stetig differenzierbar. Dann gilt*

$$f(x + h) = f(x) + \langle \operatorname{grad} f(x), h \rangle + \frac{1}{2} Q(h, h) + r(x + h) ||h||^2,$$

wobei Q die Hesse-Form von f an der Stelle x ist.

Die Hesse-Form wird bei der Bestimmung der *lokalen Extrema* von differenzierbaren Funktionen eine Rolle spielen. Damit wollen wir uns nun befassen.

Definition Es sei $U \subset \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$ eine Funktion. Ein Punkt $a \in U$ heißt *lokales Maximum* (bzw. *lokales Minimum*) von f, falls eine Umgebung $V \subset U$ von a existiert, so dass

$$f(a) \geq f(x) \quad (\text{bzw. } f(a) \leq f(x))$$

für alle $x \in V$.

Tritt in dieser Definition der Fall $f(a) = f(x)$ nur für $x = a$ ein, so spricht man von einem *isolierten* lokalen Maximum bzw. Minimum.

Ein *lokales Extremum* ist ein lokales Maximum oder Minimum.

Satz 6.4 *Es sei $f\colon U \to \mathbb{R}$ differenzierbar und $a \in U$ ein lokales Extremum. Dann ist*

$$\operatorname{grad} f(a) = 0.$$

Beweis. Für $i = 1, \ldots, n$ betrachten wir die Funktionen

$$g_i(t) := f(a + te_i)$$

Die Funktion g_i ist auf einem gewissen Intervall $(-\varepsilon, \varepsilon) \subset \mathbb{R}$, $(\varepsilon > 0)$, definiert und dort differenzierbar. Hat f in a ein lokales Extremum, so hat g_i in 0 ein lokales Extremum. Nach I, Satz 12.7, gilt also

$$g_i'(0) = 0.$$

Wegen $g_i'(0) = D_i f(a)$ folgt

$$\operatorname{grad} f(a) = (D_1 f(a), \ldots, D_n f(a)) = 0.$$

□

Definition Es sei $f\colon U \to \mathbb{R}$ differenzierbar. Ein Punkt $a \in U$ mit $\operatorname{grad} f(a) = 0$ heißt *kritischer Punkt* von f. Der Wert $f(a)$ von f an einem kritischen Punkt a heißt *kritischer Wert* von f.

Sucht man die lokalen Extrema einer differenzierbaren Funktion f, so wird man zunächst alle kritischen Punkte von f ermitteln. Ob ein kritischer Punkt dann wirklich ein lokales Extremum ist, muß im einzelnen geprüft werden.

Aufgabe 6.1 Man bestimme die lokalen Extrema der Funktion $f\colon U \to \mathbb{R}$ mit

$$f(x) = \frac{(x_1 x_2 \cdots x_n)^{\frac{1}{n+1}}}{1 + x_1 + x_2 + \cdots + x_n},$$

wobei $U := \{x \in \mathbb{R}^n \mid x_1 > 0, \ldots, x_n > 0\}$, und beweise die *Ungleichung zwischen dem arithmetischen und geometrischen Mittel*

$$\sqrt[n+1]{y_1 \cdot \ldots \cdot y_{n+1}} \leq \frac{y_1 + \cdots y_{n+1}}{n+1}$$

für $y_1 \geq 0, \ldots, y_{n+1} \geq 0$.

Um hinreichende Bedingungen für die Existenz von lokalen Extrema einer Funktion f herleiten zu können, müssen wir die Hesse-Form von f näher betrachten. Dazu brauchen wir einige Begriffe aus der Theorie der quadratischen Formen, an die wir hier erinnern.

Definition Eine quadratische Form $Q(x,x)$ heißt

- *positiv definit*, falls $Q(x,x) > 0$ für alle $x \in \mathbb{R}^n \backslash \{0\}$.
- *positiv semidefinit*, falls $Q(x,x) \geq 0$ für alle $x \in \mathbb{R}^n$.
- *negativ definit*, falls $Q(x,x) < 0$ für alle $x \in \mathbb{R}^n \backslash \{0\}$
- *negativ semidefinit*, falls $Q(x,x) \leq 0$ für alle $x \in \mathbb{R}^n$.
- *indefinit*, falls es $x, y \in \mathbb{R}^n$ gilt mit $Q(x,x) > 0$ und $Q(y,y) < 0$.

In der Linearen Algebra II wird das folgende Kriterium von Hurwitz für die positive Definitheit hergeleitet:

Satz 6.5 (Hurwitz-Kriterium) *Es sei*

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

Die quadratische Form Q ist genau dann positiv definit, wenn für alle $k = 1, \ldots, n$ gilt

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} > 0.$$

Satz 6.6 *Es sei $f \colon U \to \mathbb{R}$ zweimal stetig differenzierbar und $a \in U$ sei ein kritischer Punkt von f. Ist dann die Hesse-Form $Q(h,h)$*

- *positiv definit, so ist a ein isoliertes lokales Minimum von f,*
- *negativ definit, so ist a ein isoliertes lokales Maximum von f,*
- *indefinit, so ist a kein lokales Extremum von f.*

Beweis. Nach Korollar 6.3 gilt in einer Umgebung von a, da $\operatorname{grad} f(a) = 0$,

$$f(a+h) = f(a) + \frac{1}{2} Q(h,h) + r(a+h)||h||^2 \tag{$*$}$$

mit $\lim_{h \to 0} r(a+h) = 0$.

(a) Wir betrachten zunächst den Fall, dass Q positiv definit ist.

Es gibt dann eine Zahl $c > 0$, so dass

$$Q(h,h) \geq c||h||^2$$

für alle $h \in \mathbb{R}^n$ gilt, da das Minimum c der quadratischen Form auf der Einheitssphäre $\{h \in \mathbb{R}^n \mid ||h|| = 1\}$ nach Voraussetzung positiv ist.

Wegen $\lim_{h\to 0} r(a+h) = 0$ gibt es ein $\delta > 0$, so dass für alle $h \in \mathbb{R}^n$ mit $||h|| < \delta$ gilt

$$|r(a+h)| < \frac{c}{4}.$$

Dann folgt aus der Gleichung $(*)$

$$f(a+h) - f(a) \geq \frac{c}{2}||h||^2 - \frac{c}{4}||h||^2,$$

d. h.

$$f(a+h) - f(a) > 0 \text{ für } h \neq 0,$$

d. h. a ist isoliertes lokales Minimum.

(b) Ist Q negativ definit, so betrachte man statt f die Funktion $-f$. Damit ist man im Fall (a).

(c) Ist Q indefinit, so gibt es einen Vektor h mit $||h|| = 1$ und

$$Q(h,h) = C_0 > 0$$

und einen Vektor k mit $||k|| = 1$ und

$$Q(k,k) = c_0 < 0.$$

Dann gilt nach $(*)$ für kleine $t \in \mathbb{R}$

$$f(a+th) - f(a) = \frac{1}{2}C_0 t^2 + r(a+th)t^2.$$

Wegen $\lim_{h\to 0} r(a+h) = 0$ gibt es ein $\delta > 0$, so dass für alle $t \in \mathbb{R}$ mit $|t| < \delta$ gilt:

$$|r(a+th)| < \frac{C_0}{4}.$$

Also folgt

$$f(a+th) > f(a) \text{ für } 0 < |t| < \delta.$$

Ebenso zeigt man, dass

$$f(a+tk) < f(a)$$

für genügend kleine $t \neq 0$ gilt. Der Punkt a kann daher in diesem Fall kein lokales Extremum sein. □

Im Fall zweier Variablen lauten die bewiesenen Ergebnisse folgendermaßen: Wir betrachten Funktionen

$$f\colon \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & f(x,y) \end{array}.$$

Ist (x_0, y_0) kritischer Punkt von f, so setzen wir

$$a := f_{xx}(x_0,y_0), \quad b := f_{xy}(x_0,y_0), \quad c := f_{yy}(x_0,y_0),$$

so dass die Hesse Matrix die Form

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

bekommt. Dann gilt nach dem Hurwitz-Kriterium

$$a > 0, ac - b^2 > 0 \Rightarrow (x_0, y_0) \text{ Minimum},$$

$$a < 0, ac - b^2 > 0 \Rightarrow (x_0, y_0) \text{ Maximum},$$

$$ac - b^2 < 0 \Rightarrow (x_0, y_0) \text{ Sattelpunkt}.$$

Aufgabe 6.2 Man bestimme die Hesse-Form im Punkt $(0,0)$ und die Art des kritischen Punktes $(0,0)$ bei den folgenden Funktionen und skizziere die Graphen:

(1) $f(x,y) := x^2 + y^2$,

(2) $f(x,y) := -x^2 - y^2$,

(3) $f(x,y) := x^2 - y^2$,

(4) $f(x,y) := x^2 + y^4$,

(5) $f(x,y) := x^2$,

(6) $f(x,y) := x^2 + y^3$.

7 Lokale Umkehrbarkeit, implizite Funktionen

Für den Beweis der wichtigen Sätze der Analysis, mit denen wir uns nun beschäftigen wollen, brauchen wir den Banachschen Fixpunktsatz. Dieser Satz gilt für beliebige vollständige metrische Räume.

Es sei (X,d) ein metrischer Raum. Die Konvergenz einer Folge und der Begriff der Cauchyfolge wird wie im $\mathbb{R}^n$ erklärt:

Eine Folge $(x_k)_{k\in\mathbb{N}}$ mit $x_k \in X$ heißt *konvergent gegen* $a \in X$, genau dann, wenn $\lim_{k\to\infty} d(x_k, a) = 0$.

Eine Folge $(x_k)_{k\in\mathbb{N}}$ aus X heißt *Cauchyfolge*, wenn gilt: Zu jedem $\varepsilon > 0$ existiert ein $k_0 \in \mathbb{N}$, so dass für alle $k, m \geq k_0$ gilt

$$d(x_k, x_m) < \varepsilon.$$

Zu einer Cauchyfolge (x_k) gibt es aber i.A. nicht notwendig ein Element $a \in X$, gegen das die Folge (x_k) konvergiert (z. B. hat nicht jede Cauchyfolge in dem metrischen Raum $\mathbb{Q} \subset \mathbb{R}$ einen Grenzwert in $\mathbb{Q}$). Man definiert:

Definition Ein metrischer Raum heißt *vollständig*, wenn in ihm jede Cauchyfolge konvergiert.

Beispiel 7.1 Der $\mathbb{R}^n$, alle abgeschlossenen Teilmengen des $\mathbb{R}^n$, und die Räume stetiger Funktionen mit kompakter Definitionsmenge sind vollständige metrische Räume.

Wie schon erwähnt, lässt sich auch der Begriff einer stetigen Abbildung auf metrischen Räume übertragen.

Definition Eine Abbildung $\Phi\colon X \to X$ heißt *kontrahierend*, wenn es eine reelle Zahl c mit $0 \leq c < 1$ gibt, so dass für alle $x, y \in X$ gilt

$$d(\Phi(x), \Phi(y)) \leq c \cdot d(x, y).$$

Satz 7.1 (Banachscher Fixpunktsatz) *Es sei (X, d) ein vollständiger metrischer Raum. Jede kontrahierende Abbildung $\Phi\colon X \to X$ besitzt genau einen Fixpunkt, d.h. es gibt genau ein $z \in X$ mit*

$$\Phi(z) = z.$$

Beweis. a) Wir beweisen zunächst, dass ein Fixpunkt z existiert. Dazu konstruieren wir rekursiv eine Cauchyfolge, deren Grenzwert z die Gleichung $\Phi(z) = z$ erfüllt.

Wähle $x_0 \in X$ beliebig und setze

$$x_{k+1} = \Phi(x_k).$$

Um zu beweisen, dass (x_k) eine Cauchyfolge ist, müssen wir $d(x_{k+p}, x_k)$ abschätzen. Nach Voraussetzung gilt für alle $j \geq 1$:

$$d(x_{j+1}, x_j) = d(\Phi(x_j), \Phi(x_{j-1})) \leq c \cdot d(x_j, x_{j-1}).$$

Durch Induktion erhalten wir daraus

$$d(x_{j+1}, x_j) \leq c^j d(x_1, x_0).$$

Nach der Dreiecksungleichung gilt bei Einschub der Punkte x_{k+p-1}, x_{k+p-2}, $\ldots$, x_{k+1}:

$$\begin{aligned} d(x_{k+p}, x_k) &\leq d(x_{k+p}, x_{k+p-1}) + \cdots + d(x_{k+1}, x_k) \\ &\leq (c^{k+p-1} + c^{k+p-2} + \cdots + c^{k+1} + c^k) d(x_1, x_0) \\ &= \frac{(1 - c^p)c^k}{1 - c} d(x_1, x_0) \end{aligned}$$

Da $0 \leq c < 1$ gilt, ergibt sich für beliebiges k und p die Abschätzung

$$d(x_{k+p}, x_k) \leq \frac{c^k}{1 - c} d(x_1, x_0).$$

Wegen $\lim_{k\to\infty} c^k = 0$ ist (x_k) also eine Cauchyfolge.

Nach Voraussetzung ist (X, d) vollständig, also konvergiert (x_k) gegen $z \in X$.

Wir behaupten, dass die Folge (x_k) auch gegen $\Phi(z)$ konvergiert. Es sei $\varepsilon > 0$ vorgegeben, $k_0 \in \mathbb{N}$, so dass für alle $k \geq k_0$ gilt

$$d(x_k, z) < \varepsilon.$$

Dann gilt

$$d(x_{k+1}, \Phi(z)) \leq c \cdot d(x_k, z) < \varepsilon.$$

Also konvergiert (x_k) auch gegen $\Phi(z)$, und wegen der Eindeutigkeit des Grenzwerts ergibt sich $\Phi(z) = z$, d.h. z ist ein Fixpunkt von Φ.

b) Die Eindeutigkeit von z ergibt sich folgendermaßen: Aus $\Phi(z) = z$ und $\Phi(z') = z'$ folgt
$$d(z, z') = d(\Phi(z), \Phi(z')) \leq c \cdot d(z, z').$$
Wegen $0 \leq c < 1$ folgt hieraus $d(z, z') = 0$, d.h. $z = z'$. □

Der Beweis des Banachschen Fixpunktsatzes liefert auch ein Verfahren, diesen Fixpunkt zu bestimmen: Man startet mit einem beliebigen $x_0 \in X$ und wendet darauf immer wieder Φ an.

Man veranschauliche sich dieses Verfahren im einfachsten Fall, nämlich der kontrahierenden Abbildung eines abgeschlossenen Intervalls in sich!

Wir wollen nun den Banachschen Fixpunktsatz anwenden. Zunächst betrachten wir folgendes Problem:

Es seien $U_1, U_2 \subset \mathbb{R}^n$ offene Mengen und $f\colon U_1 \to U_2$ eine *stetig* differenzierbare Abbildung. Wir interessieren uns dafür, ob die Abbildung f bijektiv ist und eine differenzierbare Umkehrabbildung $g\colon U_2 \to U_1$ besitzt. Wir haben bereits eine notwendige Bedingung hierfür abgeleitet, nämlich die Bedingung
$$\det f'(a) \neq 0.$$
Der nächste Satz zeigt, dass diese Bedingung für die Umkehrbarkeit auch hinreichend ist, wenn man eine Verkleinerung von U_1 und U_2 erlaubt.

Satz 7.2 (Satz von der lokalen Umkehrbarkeit) *Es sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Es sei $a \in U$ und $f'(a)$ sei umkehrbar, d.h. es gelte*
$$\det f'(a) \neq 0.$$
Dann gibt es eine offene Umgebung V von a, so dass gilt

(i) *$f|_V$ ist injektiv,*

(ii) *$f(V)$ ist offen,*

(iii) *die Umkehrabbildung g von $f|_V$ ist stetig differenzierbar,*

d.h. f bildet V diffeomorph auf $f(V)$ ab.

Beweis. Zur Vereinfachung nehmen wir an, dass $a = 0$ und $f(a) = 0$ (dies kann man erreichen, indem man statt $x \mapsto f(x)$ die Abbildung $x \mapsto f(x + a) - f(a)$ betrachtet).

Außerdem können wir noch annehmen, dass gilt
$$f'(0) = \mathrm{id}_{\mathbb{R}^n}.$$
(Denn nach Voraussetzung ist $f'(0)$ umkehrbar. Betrachtet man statt f die Abbildung $(f'(0))^{-1} \circ f$, so ist deren Ableitung die Identität.)

Gesucht ist nun eine Umgebung W von 0, so dass jedes $y \in W$ genau ein Urbild hat.

Dazu betrachte Abbildung Φ_y mit
$$\Phi_y(x) = x - f(x) + y. \qquad (*)$$
Fixpunkte von Φ_y sind Urbilder von y.

Ziel: Bestimme $r > 0$, so dass für alle $y \in B(0, r)$ gilt

(a) $\Phi_y(\overline{B(0,2r)}) \subset \overline{B(0,2r)}$ und

(b) $\Phi_y \colon \overline{B(0,2r)} \to \overline{B(0,2r)}$ kontrahierend.

Wegen $f'(0) = \mathrm{id}_{\mathbb{R}^n}$ und der Stetigkeit von f' gibt es ein $r > 0$, so dass für alle x mit $||x|| \leq 2r$ gilt

$$||\underbrace{f'(0)}_{=\mathrm{id}_{\mathbb{R}^n}} - f'(x)|| \leq \frac{1}{2}.$$

Nun gilt

$$\Phi_0(x) = x - f(x) \Rightarrow \Phi_0'(x) = \mathrm{id}_{\mathbb{R}^n} - f'(x),$$

also folgt nach dem Schrankensatz 5.11 für x mit $||x|| \leq 2r$:

$$||x - f(x)|| = ||\Phi_0(x)|| = ||\Phi_0(x) - \Phi_0(0)|| \leq ||\mathrm{id}_{\mathbb{R}^n} - f'(x)||\,||x|| \leq \frac{1}{2}||x||.$$

Falls $||y|| < r$ und $||x|| \leq 2r$ ergibt sich daher

$$||\Phi_y(x)|| = ||x - f(x) + y)|| \leq ||x - f(x)|| + ||y|| \leq \frac{1}{2}2r + r = 2r.$$

Damit ist (a) gezeigt für alle $y \in B(0,r)$.

Wir zeigen nun, dass (b) für alle $y \in B(0,r)$ gilt. Es seien $x_1, x_2 \in \overline{B(0,2r)}$. Dann gilt

$$||\Phi_y(x_2) - \Phi_y(x_1)|| \leq ||\Phi_y'||\,||x_2 - x_1|| = ||\Phi_0'||\,||x_2 - x_1|| \leq \frac{1}{2}||x_2 - x_1||.$$

Also ist Φ_y: $\overline{B(0,2r)} \to \overline{B(0,2r)}$ für jedes $y \in B(0,r)$ kontrahierend.

Nach dem Banachschen Fixpunktsatz hat Φ_y genau einen Fixpunkt x in $\overline{B(0,2r)}$, d.h. es gibt zu jedem $y \in B(0,r)$ genau ein $x \in \overline{B(0,2r)}$ mit $f(x) = y$. Also ist

$$f \colon V := f^{-1}(B(0,r)) \to B(0,r)$$

bijektiv. Die Menge V ist offen, da f stetig ist. Damit sind die Aussagen (i) und (ii) des Satzes bewiesen.

Zu (iii): Es genügt zu zeigen, dass die Umkehrabbildung g von $f|_V$ stetig ist, dann nach Satz 5.8 folgt daraus die Differenzierbarkeit von g und

$$g'(y) = (f'(g(y)))^{-1} \text{ für } y \in f(V).$$

Aus der Stetigkeit von f' ergibt sich dann die Stetigkeit von g'.

Beweis, dass g stetig ist: Es seien $x_1, x_2 \in \overline{B(0,2r)}$. Dann gilt

$$\left.\begin{aligned} \Phi_0(x_1) &= x_1 - f(x_1) \\ \Phi_0(x_2) &= x_2 - f(x_2) \end{aligned}\right\} \Rightarrow$$

$$\begin{aligned} &\Rightarrow \quad x_2 - x_1 = \Phi_0(x_2) - \Phi_0(x_1) + f(x_2) - f(x_1) \\ &\Rightarrow \quad ||x_2 - x_1|| \leq \frac{1}{2}||x_2 - x_1|| + ||f(x_2) - f(x_1)|| \\ &\Rightarrow \quad ||x_2 - x_1| \leq 2||f(x_2) - f(x_1)||. \end{aligned}$$

Für $y_1 = f(x_1)$, $y_2 = f(x_2)$ folgt daraus

$$||g(y_2) - g(y_1)|| \leq 2||y_2 - y_1||.$$

Aus dieser Abschätzung folgt die Stetigkeit von g. □

Aufgabe 7.1 Man betrachte das Gleichungssystem

$$\begin{aligned} y_1 &= x_1 + x_2 + x_3, \\ y_2 &= x_2x_3 + x_3x_1 + x_1x_3, \\ y_3 &= x_1x_2x_3. \end{aligned}$$

Man zeige: Wenn $x = a$ Lösung zu $y = b$ und $(a_1 - a_2)(a_1 - a_3)(a_2 - a_3) \neq 0$, dann lassen sich die Gleichungen für alle $y \in f(V)$, $f(V)$ geeignete Umgebung von b, nach x auflösen:

$$\begin{aligned} x_1 &= g_1(y_1, y_2, y_3) \\ x_2 &= g_2(y_1, y_2, y_3) \qquad (y \in f(V)) \\ x_3 &= g_3(y_1, y_2, y_3) \end{aligned}$$

Dabei gilt

$$(x_1 - x_2)(x_1 - x_3)(x_2 - x_3) \neq 0.$$

Für die kubische Gleichung

$$t^3 - y_1t^2 + y_2t - y_3 = 0 \tag{$*$}$$

bedeutet das: Hat $(*)$ für $y_1 = b_1, y_2 = b_2, y_3 = b_3$ drei verschiedene reelle Nullstellen, so auch für y_1, y_2, y_3 in hinreichender Nähe von b_1, b_2, b_3; die Lösungen sind (unendlich oft) differenzierbare Funktionen der Koeffizienten y_1, y_2, y_3.

Bemerkung 7.1 Im Fall $n = 1$ ist die Bedingung $f'(x) > 0$ für eine Funktion $f\colon [a, b] \to \mathbb{R}$ sogar hinreichend für die „globale“ Umkehrbarkeit. Für $n \geq 2$ gilt die entsprechende Aussage nicht, wie folgendes Beispiel zeigt:

Es sei $U = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \;\middle|\; 0 < x^2 + y^2 \right\}$,

$$f\colon \begin{array}{ccc} U & \to & \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix} \end{array}.$$

Es gilt

$$\det Df \begin{pmatrix} x \\ y \end{pmatrix} = \begin{vmatrix} 2x & -2y \\ 2y & 2x \end{vmatrix} = 4(x^2 + y^2) > 0$$

für $\begin{pmatrix} x \\ y \end{pmatrix} \in U$. Nach Satz 7.2 ist f also lokal umkehrbar; jedoch ist f nicht global umkehrbar, da $f(x, y) = f(-x, -y)$.

Ein anderes Beispiel ist die Polarkoordinatenabbildung mit größerer Definitionsmenge

$$f\colon \begin{array}{ccc} \mathbb{R}_+^* \times \mathbb{R} & \to & \mathbb{R}^2 \\ \begin{pmatrix} r \\ \varphi \end{pmatrix} & \mapsto & \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix} \end{array}.$$

Wieviele Punkte haben hier denselben Bildpunkt?

Satz über implizite Funktionen

Wir beschäftigen uns nun mit der Aufgabe, einen Überblick über die Lösungen eines Systems von m Gleichungen in n Variablen

$$\begin{array}{rcl} f_1(x_1,\dots,x_n) & = & 0 \\ \vdots & \vdots & \vdots \\ f_m(x_1,\dots,x_n) & = & 0 \end{array} \qquad (m<n)$$

zu gewinnen. Voraussetzung: $f_1,\dots,f_m$ stetig differenzierbar, Lösung $(x_1,\dots,x_n)=(c_1,\dots,c_n)$ bekannt.

Wir orientieren uns dazu an dem linearen Gleichungssystem:

$$\begin{array}{rcl} a_{11}x_1+\dots+a_{1n}x_n & = & 0 \\ \vdots & \vdots & \vdots \\ a_{m1}x_1+\dots+a_{mn}x_n & = & 0 \end{array} \qquad (m<n)$$

Angenommen, der Rang ist maximal, d.h. Rang $=m$ und die ersten m Spalten seien linear unabhängig, dann kann das Gleichungssystem mit dem Gaußschen Algorithmus auf die folgende Gestalt gebracht werden:

$$\begin{array}{llllll} x_1 & & & +a'_{1,m+1}x_{m+1}+\dots+a'_{1,n}x_n & =0 \\ & x_2 & & +a'_{2,m+1}x_{m+1}+\dots+a'_{2,n}x_n & =0 \\ & & \ddots & & \\ & & x_m & +a'_{m,m+1}x_{m+1}+\dots+a'_{m,n}x_n & =0 \end{array}$$

Das bedeutet, dass bei geeigneter Nummerierung die ersten m Variablen durch $x_{m+1},\dots,x_n$ ausgedrückt werden können. Außerdem folgt, dass eine Lösung existiert.

Im nichtlinearen Fall ist es nicht so leicht zu entscheiden, ob es überhaupt eine Lösung gibt. Wir setzen dies deshalb voraus. Wir werden dann zeigen: Dann existiert eine von $p:=n-m$ Parametern abhängige Lösungsschar, falls das Analogon der Rangbedingung erfüllt ist.

Dazu nummerieren wir die Variablen geeignet und bezeichnen sie mit $x_1,\dots,x_m,u_1,\dots,u_p$. Untersucht werden soll das Gleichungssystem

$$\begin{array}{rcl} f_1(x_1,\dots,x_m,u_1,\dots,u_p) & = & 0 \\ \vdots & \vdots & \vdots \\ f_m(x_1,\dots,x_m,u_1,\dots,u_p) & = & 0 \end{array}$$

unter der Voraussetzung

$$\begin{vmatrix} D_1f_1(c) & \cdots & D_mf_1(c) \\ \vdots & \ddots & \vdots \\ D_1f_m(c) & \cdots & D_mf_m(c) \end{vmatrix} \neq 0.$$

Die Möglichkeit, die beliebigen differenzierbaren Funktionen $f_1,\dots,f_m$ in der Nähe der Lösung c durch lineare Funktionen zu ersetzen, berechtigt zu der Vermutung, dass $x_1,\dots,x_m$ als differenzierbare Funktionen von u_1, …, u_p so darstellbar sind, dass das Gleichungssystem für alle Werte von $(u_1,\dots,u_p)$ aus einer geeigneten Umgebung des Punktes $(c_{m+1},\dots,c_n)$ erfüllt wird.

Notation Für die Formulierung des entsprechenden Satzes führen wir folgende Bezeichnungen ein:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} =: x, \qquad \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} =: a, \qquad \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} =: f,$$

$$\begin{pmatrix} u_1 \\ \vdots \\ u_p \end{pmatrix} =: u, \qquad \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} =: b, \qquad \begin{pmatrix} a \\ b \end{pmatrix} =: c,$$

$$\frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_m} \end{pmatrix}$$

$$\frac{\partial f}{\partial u} = \begin{pmatrix} \frac{\partial f_1}{\partial u_1} & \cdots & \frac{\partial f_1}{\partial u_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial u_1} & \cdots & \frac{\partial f_m}{\partial u_p} \end{pmatrix}$$

Die Rangbedingung lautet also:

$$\det \frac{\partial f}{\partial x}(a,b) \neq 0.$$

Satz 7.3 (Satz über implizite Funktionen) *Es sei $D \subset \mathbb{R}^m \times \mathbb{R}^p$ offen und $f\colon D \to \mathbb{R}^m$ sei stetig differenzierbar. Weiter gelte*

$$f(a,b) = 0 \text{ und } \det \frac{\partial f}{\partial x}(a,b) \neq 0.$$

Dann gibt es offene Umgebungen $U \subset \mathbb{R}^p$ von b und $V \subset \mathbb{R}^m$ von a mit $V \times U \subset D$ und eine eindeutig bestimmte stetig differenzierbare Abbildung

$$\varphi\colon U \to V,$$

so dass für alle $u \in U$ gilt

$$f(\varphi(u), u) = 0.$$

Ist $(x,u) \in V \times U$ mit $f(x,u) = 0$, so folgt $x = \varphi(u)$.

Für die Ableitung von φ im Punkt b gilt

$$\varphi'(b) = -\left(\frac{\partial f}{\partial x}(a,b)\right)^{-1} \circ \frac{\partial f}{\partial u}(a,b).$$

Beweis. Wir wollen den Satz über die lokale Umkehrbarkeit anwenden. Definiere

$$\begin{array}{rccc} F\colon & D \subset \mathbb{R}^m \times \mathbb{R}^p & \to & \mathbb{R}^m \times \mathbb{R}^p \\ & \begin{pmatrix} x \\ u \end{pmatrix} & \mapsto & \begin{pmatrix} f(x,u) \\ u \end{pmatrix} \end{array}$$

Es gilt

$$DF(x,u) = \begin{pmatrix} \frac{\partial f}{\partial x}(x,u) & \frac{\partial f}{\partial u}(x,u) \\ 0 & \mathrm{id}_{\mathbb{R}^p} \end{pmatrix},$$

also

$$\det DF(a,b) = \det \frac{\partial f}{\partial x}(a,b) \neq 0.$$

Nach Satz 7.2 gibt es also eine offene Umgebung W von $\begin{pmatrix} a \\ b \end{pmatrix} \in D$, so dass $F|_W$ eine stetig differenzierbare Umkehrabbildung

$$G\colon F(W) \to W$$

besitzt. Die Abbildung G hat die Form

$$G(y,v) = \begin{pmatrix} g(y,v) \\ v \end{pmatrix},$$

wobei $g\colon F(W) \subset \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}^m$ stetig differenzierbar ist.

Dieses g ergibt nun die gesuchte Abbildung φ, indem man $y = 0$ einsetzt:

$$\varphi(v) := g(0,v).$$

Die Definitionsmenge von g ist die offene Menge $F(W)$ mit $\begin{pmatrix} 0 \\ b \end{pmatrix} \in F(W)$. Es sei U eine offene Umgebung von b in $\mathbb{R}^p$, so dass für alle $v \in U$ gilt: $\begin{pmatrix} 0 \\ v \end{pmatrix} \in F(W)$. Dann ist φ auf U erklärt und stetig differenzierbar.

Wegen $F \circ G = \mathrm{id}_{F(w)}$ folgt

$$F(W) \ni \begin{pmatrix} y \\ v \end{pmatrix} \overset{G}{\mapsto} \begin{pmatrix} g(y,v) \\ v \end{pmatrix} \overset{F}{\mapsto} \begin{pmatrix} f(g(y,v),v) \\ v \end{pmatrix} = \begin{pmatrix} y \\ v \end{pmatrix},$$

also

$$f(g(y,v),v) = y \quad \text{für } \begin{pmatrix} y \\ v \end{pmatrix} \in F(W).$$

Mit $y = 0$ und $v = u$ erhält man also

$$f(\varphi(u),u) = 0 \quad \text{für alle } u \in U.$$

Dass $\varphi(u)$ die einzige Lösung ist, folgt aus der Gleichung $g(f(x,u),u) = x$:

$$f(x,u) = 0 \Rightarrow g(0,u) = x \Rightarrow x = \varphi(u).$$

Die Ableitung von φ bestimmt man nach der Kettenregel: Wegen

$$f(\varphi(u),u) = 0 \quad \text{für alle } u \in U$$

folgt (betrachte $u \mapsto f(\varphi(u),u)$)

$$\begin{pmatrix} \frac{\partial f}{\partial x}(\varphi(u),u) & \frac{\partial f}{\partial u}(\varphi(u),u) \end{pmatrix} \begin{pmatrix} \varphi'(u) \\ \mathrm{id}_{\mathbb{R}^p} \end{pmatrix} = 0,$$

also

$$\frac{\partial f}{\partial x}(\varphi(u),u) \circ \varphi'(u) + \frac{\partial f}{\partial u}(\varphi(u),u) = 0,$$

wegen $\det \frac{\partial f}{\partial x}(a,b) \neq 0$ also

$$\varphi'(b) = -\left(\frac{\partial f}{\partial x}(a,b)\right)^{-1} \circ \frac{\partial f}{\partial u}(a,b).$$

□

Bemerkung 7.2 Der Beweis liefert für alle $u \in U$ die Formel

$$\varphi'(u) = -\left(\frac{\partial f}{\partial x}(\varphi(u), u)\right)^{-1} \circ \left(\frac{\partial f}{\partial u}(\varphi(u), u)\right).$$

Beispiel 7.2 Nur eine Gleichung

$$f(x_1, \ldots, x_n) = 0. \tag{$*$}$$

Bilde $\operatorname{grad} f(x)$ und prüfe, ob eine der Komponenten von 0 verschieden ist. Gilt etwa

$$D_n f(c_1, \ldots, c_n) \neq 0,$$

so kann die Gleichung $(*)$ in einer Umgebung von $(c_1, \ldots, c_{n-1})$ nach x_n aufgelöst werden:

$$x_n = \varphi(x_1, \ldots, x_{n-1}).$$

Man kann nun $\operatorname{grad} \varphi$ gemäß der hergeleiteten Formel berechnen. Praktisch geht man so vor:

$$\begin{aligned}
f(x_1, \ldots, x_{n-1}, \varphi(x_1, \ldots, x_{n-1})) &= 0 \\
D_1 f + D_n f \cdot D_1 \varphi &= 0 \\
\Rightarrow \qquad \vdots \quad & \vdots \quad \vdots \\
D_{n-1} f + D_n f \cdot D_{n-1} \varphi &= 0 \\
\Rightarrow \quad \operatorname{grad} \varphi &= \left(-\frac{D_1 f}{D_n f}, \ldots, -\frac{D_{n-1} f}{D_n f}\right).
\end{aligned}$$

Z.B. $f(x) = x_1^2 + \cdots + x_n^2 - 1 = 0$:

$$\operatorname{grad} \varphi(x_1, \ldots, x_{n-1}) = \left(-\frac{x_1}{x_n}, \ldots, -\frac{x_{n-1}}{x_n}\right)$$

($x_n < 0$ und $x_n > 0$ unterscheiden!)

$n = 2$: $f(x, y) = 0$. Gilt etwa $f_x(x_0, y_0) \neq 0$, $f(x_0, y_0) = 0$, so kann die ebene Kurve, die durch obige Gleichung definiert ist, lokal durch

$$x = \varphi(y)$$

beschrieben werden. Es ergibt sich

$$\varphi'(y) = -\frac{f_y(x, y)}{f_x(x, y)}$$

(wobei $f(x, y) = 0$ gilt).

Die Tangenten an die Kurve können hiermit leicht bestimmt werden:

Gleichung der Tangente in (x, y) mit Koordinaten ξ, η:

$$\xi - x = -\frac{f_y(x, y)}{f_x(x, y)}(\eta - y)$$

$$\Leftrightarrow f_x(x, y)(\xi - x) + f_y(x, y)(\eta - y) = 0.$$

Ganz entsprechend für eine Hyperfläche

$$f(x_1, \ldots, x_n) = 0$$

mit $\operatorname{grad} f \neq 0$: Gleichung der Tangentialhyperebene in x:

$$\langle \operatorname{grad} f(x), \xi - x \rangle = 0, \quad \xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}.$$

Beispiel 7.3 Durch zwei Gleichungen

$$\begin{aligned} f(x, y, z) &= 0, \\ g(x, y, z) &= 0. \end{aligned}$$

wird im $\mathbb{R}^3$ im Allgemeinen eine Raumkurve als Schnitt zweier Flächen bestimmt. Die Jacobi-Matrix ist

$$\begin{pmatrix} f_x & f_y & f_z \\ g_x & g_y & g_z \end{pmatrix}.$$

Ist nun etwa $f_x g_y(c) - f_y g_x(c) \neq 0$, $f(c) = g(c) = 0$, so kann man in einer Umgebung von c die Variablen x und y als stetig differenzierbare Funktionen von z darstellen:

$$\begin{aligned} x &= \varphi(z), \\ y &= \psi(z). \end{aligned}$$

Es gilt

$$\begin{pmatrix} \varphi' \\ \psi' \end{pmatrix} = - \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}^{-1} \begin{pmatrix} f_z \\ g_z \end{pmatrix} = \begin{pmatrix} \frac{f_y g_z - f_z g_y}{f_x g_y - f_y g_x} \\ \frac{f_z g_x - f_x g_z}{f_x g_y - f_y g_x} \end{pmatrix}.$$

Dies erhält man auch aus

$$\begin{aligned} f(\varphi(z), \psi(z), z) &= 0 \\ g(\varphi(z), \psi(z), z) &= 0 \end{aligned}$$

durch Anwendung der Kettenregel:

$$\begin{aligned} f_x \cdot \varphi' + f_y \cdot \psi' + f_z &= 0 \\ g_x \cdot \varphi' + g_y \cdot \psi' + g_z &= 0. \end{aligned}$$

Extrema unter Nebenbedingungen

Wir beginnen mit einem typischen Beispiel für die Aufgabe, Maxima und Minima unter Nebenbedingungen aufzusuchen.

Beispiel 7.4 Gesucht sind diejenigen Punkte der Ellipse

$$ax^2 + 2bxy + cy^2 = 1 \qquad (a > 0, ac - b^2 > 0),$$

die von $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ maximalen bzw. minimalen Abstand haben (statt Abstand: Quadrat des Abstandes).

D. h.: Gesucht sind die Extrema der Funktion

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

unter der Nebenbedingung

$$g(x,y) = ax^2 + 2bxy + cy^2 - 1 = 0.$$

Die Funktion f hat in $\mathbb{R}^2$ genau ein Minimum und zwar in $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$; dieser Punkt interessiert aber nicht, da er der Nebenbedingung nicht genügt.

Mögliche Lösung: Bestimme Parameterdarstellung

$$x = \varphi(t), y = \psi(t)$$

für die Nebenbedingung $g(x,y) = 0$, und bestimme Extrema der Funktion

$$F(t) := f(\varphi(t), \psi(t)).$$

Verfahren, dass die mühsame explizite Bestimmung einer solchen Parameterdarstellung vermeidet:

Die kritischen Punkte von F genügen der Gleichung

$$F'(t) = D_1 f(\varphi(t), \psi(t)) \cdot \varphi'(t) + D_2 f(\varphi(t), \psi(t)) \cdot \psi'(t) = 0.$$

Wegen $g(\varphi(t), \psi(t)) = 0$ gilt für alle t:

$$D_1 g(\varphi(t), \psi(t)) \cdot \varphi'(t) + D_2 g(\varphi(t), \psi(t)) \cdot \psi'(t) = 0.$$

Die Vektoren $\begin{pmatrix} D_1 f \\ D_2 f \end{pmatrix}$ und $\begin{pmatrix} D_1 g \\ D_2 g \end{pmatrix}$ stehen also an jedem kritischen Punkt von F senkrecht auf dem Vektor $\begin{pmatrix} \varphi'(t) \\ \psi'(t) \end{pmatrix}$, sie sind also linear abhängig. Wenn $\begin{pmatrix} x \\ y \end{pmatrix}$ kritischer Punkt von F ist, gibt es also ein $\lambda \in \mathbb{R}$, so dass gilt

$$D_1 f(x,y) - \lambda D_1 g(x,y) = 2x - 2\lambda(ax + by) = 0,$$

$$D_2 f(x,y) - \lambda D_2 g(x,y) = 2y - 2\lambda(bx + cy) = 0,$$
$$g(x,y) = ax^2 + 2bxy + cy^2 - 1 = 0.$$

Die verwendete Parameterdarstellung ist verschwunden: Die kritischen Punkte von F sind die kritischen Punkte von

$$G = f - \lambda g$$

mit $g(x,y) = 0$.

Nun gehen wir in unseren Beispiel wie folgt vor: Löse zunächst das lineare Gleichungssystem

$$\begin{aligned} (1 - \lambda a)x - \lambda by &= 0, \\ -\lambda bx + (1 - \lambda c)y &= 0. \end{aligned}$$

Wegen $g(0,0) = -1$ interessieren nur die nichttrivialen Lösungen; also muss λ der Gleichung

$$\begin{vmatrix} 1-\lambda a & -\lambda b \\ -\lambda b & 1-\lambda c \end{vmatrix} = (ac-b^2)\lambda^2 - (a+c)\lambda + 1 = 0$$

genügen. Wegen $a > 0, ac - b^2 > 0$ (damit auch $c > 0$) hat diese quadratische Gleichung zwei verschiedene reelle Lösungen λ_1, λ_2, falls $(a-c)^2 + 4b^2 > 0$ gilt, d.h. falls kein Kreis vorliegt.

Die Lösungsmenge zu λ_1 (bzw. λ_2) ist eine Gerade durch den Nullpunkt; sie schneidet die Kurve $g(x,y) = 0$ in zwei Punkten (x_1, y_1) und $(-x_1, -y_1)$ (bzw. (x_2, y_2) und $(-x_2, y_2)$).

Es gibt also insgesamt 4 kritische Punkte von G mit $g(x,y) = 0$, die paarweise symmetrisch zum Nullpunkt liegen; dies sind auch wirklich Extrema von $f(x,y) = x^2 + y^2$: Scheitelpunkte der Ellipse, Verbindungsstrecken mit dem Nullpunkt sind Hauptachsen.

Das an diesem Beispiel erläuterte Verfahren läßt sich nun analog auch für Extremwertprobleme mit n Variablen und r Nebenbedingungen durchführen.

Satz 7.4 *Es sei $U \subset \mathbb{R}^n$ offen, $f, g_1, \dots, g_r \colon U \to \mathbb{R}$ stetig differenzierbare Funktionen. Die Teilmenge $M \subset U$ sei definiert durch*

$$\begin{aligned} g_1(x) &= 0 \\ \vdots\ &\ \vdots\ \ \vdots \\ g_r(x) &= 0. \end{aligned}$$

Die Ableitung der Abbildung $g = (g_1, \dots, g_r) \colon U \to \mathbb{R}^r$ habe überall den Rang r. Hat dann die eingeschränkte Funktion $f|_M$ in $a \in M$ ein lokales Extremum, so gibt es reelle Zahlen $\lambda_1, \dots, \lambda_r$ mit

$$\operatorname{grad} f(a) = \lambda_1 \operatorname{grad} g_1(a) + \cdots + \lambda_r \operatorname{grad} g_r(a).$$

(Die Zahlen $\lambda_1, \dots, \lambda_r$ heißen „Lagrangesche Multiplikatoren“*).*

Beweis. Setze

$$y := \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix}, \qquad u := \begin{pmatrix} x_{r+1} \\ \vdots \\ x_n \end{pmatrix}.$$

Wegen $\operatorname{Rang} Dg(a) = r$ können wir noch eventueller Umnummerierung der Koordinaten annehmen, dass

$$\det \frac{\partial g}{\partial y}(a) \neq 0.$$

Wir setzen

$$a' := \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix}, \qquad a'' := \begin{pmatrix} a_{r+1} \\ \vdots \\ a_n \end{pmatrix}, \qquad p := n - r.$$

Nach dem Satz über implizite Funktionen (Satz 7.3) gibt es eine offene Umgebung $V \subset \mathbb{R}^p$ von a'' und eine eindeutig bestimmte stetig differenzierbare Abbildung

$$\varphi\colon V \to \mathbb{R}^r$$

mit $\varphi(a'') = a'$, so dass für alle $u \in V$ gilt

$$g(\varphi(u), u) = 0.$$

Betrachte die Abbildung

$$\begin{array}{cccc} \Phi\colon & V & \to & \mathbb{R}^n \\ & u & \mapsto & \begin{pmatrix} \varphi(u) \\ u \end{pmatrix} \end{array},$$

wobei $\varphi\colon V \to \mathbb{R}^n$, $V \subset \mathbb{R}^p$, $p = n - r$, $\varphi(b) = a$. Aus $g_i(\Phi(u)) = 0$ und der Kettenregel ergibt sich

$$0 = D_j(g_i \circ \Phi)(a'') = \sum_{k=1}^{n} D_k g_i(a) \cdot D_j \Phi_k(a'') = \langle \operatorname{grad} g_i(a), D_j\Phi(a'')\rangle$$

für $i = 1, \ldots, r, j = 1, \ldots, p$.

Setzt man $v_i := \operatorname{grad} g_i(a), w_j = D_j\Phi(a'')$, so gilt also

$$\langle v_i, w_j\rangle = 0 \text{ für } i = 1, \ldots, r, j = 1, \ldots, p.$$

Hieraus folgt, dass $\{v_1, \ldots, v_r, w_1, \ldots, w_p\}$ eine Basis des $\mathbb{R}^n$ ist: Denn

$$\begin{aligned} & \sum_{i=1}^{r} \alpha_i v_i + \sum_{j=1}^{p} \beta_j w_j = 0 \qquad (\alpha_i, \beta_j \in \mathbb{R}) \\ \Rightarrow \quad & \left\langle \sum_{i=1}^{r} \alpha_i v_i + \sum_{j=1}^{p} \beta_j w_j, \sum_{i=1}^{r} \alpha_i v_i \right\rangle = \left\langle \sum_{i=1}^{r} \alpha_i v_i, \sum_{i=1}^{r} \alpha_i v_i \right\rangle = 0 \\ \Rightarrow \quad & \sum_{i=1}^{r} \alpha_i v_i = 0 \\ \Rightarrow \quad & \alpha_1 = \cdots = \alpha_r = 0, \end{aligned}$$

da $\{v_1, \ldots, v_r\}$ linear unabhängig wegen $\operatorname{Rang} g'(a) = r$.

Entsprechend folgt

$$\sum_{j=1}^{p} \beta_j w_j = 0$$

und daraus

$$\beta_1 = \cdots = \beta_p = 0,$$

weil wegen $\operatorname{Rang} \Phi'(a'') = p$ auch $\{w_1, \ldots, w_p\}$ linear unabhängig ist.

Also ist $\{v_1, \ldots, v_r, w_1, \ldots, w_p\}$ eine Basis des $\mathbb{R}^n$ und der Vektor $\operatorname{grad} f(a)$ erlaubt eine Darstellung

$$\operatorname{grad} f(a) = \sum_{i=1}^{r} \lambda_i v_i + \sum_{j=1}^{p} \mu_j w_j$$

mit $\lambda_i, \mu_j \in \mathbb{R}$.

Weil $F = f \circ \Phi\colon V \to \mathbb{R}$ in b einen kritischen Punkt hat, folgt

$$0 = D_j(f \circ \Phi)(a'') = \sum_{k=1}^{n} D_k f(a) \cdot D_j \Phi_k(a'') = \langle \operatorname{grad} f(a), w_j \rangle$$

für $j = 1, \ldots, p$.

Aus

$$\left\langle \sum_{i=1}^{r} \lambda_i v_i + \sum_{j=1}^{p} \mu_j w_j, \sum_{j=1}^{p} \mu_j w_j \right\rangle = 0$$

folgt aber wie oben $\mu_1 = \cdots = \mu_p = 0$, so dass wir die Behauptung

$$\operatorname{grad} f(a) = \sum_{i=1}^{r} \lambda_i \operatorname{grad} g_i(a)$$

erhalten. □

Um also ein lokales Extremum der Funktion f unter den Nebenbedingungen $g_1 = 0, \ldots, g_r = 0$ zu bestimmen, hat man also das Gleichungssystem für die $n + r$ Unbekannten $x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_r$ mit den $n + r$ Gleichungen

$$\begin{array}{rcl} D_1 f(x) & = & \lambda_1 D_1 g_1(x) + \cdots + \lambda_1 D_1 g_r(x) \\ \vdots & \vdots & \vdots \\ D_n f(x) & = & \lambda_1 D_n g_1(x) + \cdots + \lambda_r D_n g_r(x) \\ g_1(x) & = & 0 \\ \vdots & \vdots & \vdots \\ g_r(x) & = & 0 \end{array}$$

zu lösen. Ob dann eine Lösung $x_1, \ldots, x_n$ dieses Gleichungssystems wirklich ein lokales Extremum von $f|_M$ darstellt, muss auf andere Weise geprüft werden.

Definiert man $F\colon U \times \mathbb{R}^r \to \mathbb{R}$ durch

$$F(x, \lambda) := f(x) - \sum_{i=1}^{r} \lambda_i g_i(x),$$

so ist das obige Gleichungssystem äquivalent zu

$$\operatorname{grad} F(x, \lambda) = 0,$$

d.h. man sucht für F die lokalen Extrema ohne Nebenbedingungen.

8 Gewöhnliche Differentialgleichungen

Definition Es sei $G \subset \mathbb{R}^2$ und

$$\begin{array}{rccl} f\colon & G & \to & \mathbb{R} \\ & \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & f(x, y) \end{array}$$

eine stetige Funktion. Dann heißt

$$y' = f(x, y) \tag{1}$$

eine *Differentialgleichung erster Ordnung*. Unter einer *Lösung* von (1) versteht man eine differenzierbare Funktion

$$\varphi\colon I \to \mathbb{R} \qquad (I \subset \mathbb{R} \text{ Intervall})$$

mit

a) Der Graph von φ ist in G enthalten, d. h.
$$\Gamma_\varphi = \{(x, y) \in I \times \mathbb{R} \mid y = \varphi(x)\} \subset G.$$

b) $\varphi'(x) = f(x, \varphi(x))$ für alle $x \in I$.

(a) ist für b) nötig).

Bemerkung 8.1 Bei

$$y' = f(x)$$

einfach integrieren

$$\varphi(x) = c + \int_{x_0}^{x} f(t)\,dt.$$

(Im Allgemeinen $\varphi(x) = c + \int_{x_0}^{x} f(t, \underbrace{\varphi(t)}_{\text{ges. Funkt.}})dt$, $c = \varphi(x_0)$.)

Geometrische Interpretation Eine Differentialgleichung $y' = f(x, y)$ in einem Gebiet $G \subset \mathbb{R}^2$ bestimmt ein *Richtungsfeld*: In jedem Punkt $(x, y) \in G$ wird durch $y' = f(x, y)$ eine Steigung vorgegeben (Skizze!). Gesucht sind differenzierbare Funktionen, deren Graphen in jedem ihrer Punkte die vorgegebene Steigung haben.

Aufgabe 8.1 Man skizziere die Richtungsfelder und bestimme die Lösungen der folgenden Differentialgleichungen

(1) $y' = \frac{y}{x}$ in $\mathbb{R}_+^* \times \mathbb{R}$,

(2) $y' = -\frac{x}{y}$ in $\mathbb{R} \times \mathbb{R}_+^*$.

Definition Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und

$$\begin{aligned} f\colon\ G &\to \mathbb{R}^n \\ \begin{pmatrix} x \\ y \end{pmatrix} &\mapsto f(x, y) \end{aligned}$$

eine stetige Abbildung. Dann heißt

$$y' = f(x, y) \tag{2}$$

ein *System* von n Differentialgleichungen erster Ordnung. Eine *Lösung* von (2) ist eine differenzierbare Abbildung

$$\varphi\colon I \to \mathbb{R}^n \qquad (I \subset \mathbb{R} \text{ Intervall})$$

mit

a) Der Graph von φ ist in G enthalten.

b) $\varphi'(x) = f(x, \varphi(x))$ für alle $x \in I$.

Das Gleichungssystem (2) lautet ausgeschrieben

$$\begin{aligned} y_1' &= f_1(x, y_1, \dots, y_n), \\ y_2' &= f_2(x, y_1, \dots, y_n), \\ &\ \vdots \\ y_n' &= f_n(x, y_1, \dots, y_n). \end{aligned}$$

Definition Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $f\colon G \to \mathbb{R}$ eine stetige Funktion. Dann heißt

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}) \tag{3}$$

eine *Differentialgleichung n-ter Ordnung*. Eine *Lösung* von (3) ist eine n-mal differenzierbare Funktion

$$\varphi\colon I \to \mathbb{R} \qquad (I \subset \mathbb{R} \text{ Intervall})$$

mit

(a) $\{(x, y_0, \dots, y_{n-1}) \in I \times \mathbb{R}^n \mid y_\nu = \varphi^{(\nu)}(x),\ 0 \le \nu \le n-1\} \subset G$,

(b) $\varphi^{(n)}(x) = f(x, \varphi(x), \varphi'(x), \dots, \varphi^{(n-1)}(x))$ für alle $x \in I$.

Eine Differentialgleichung n-ter Ordnung schreibt also eine Bedingung für die n-te Ableitung der gesuchten Funktion φ vor. Diese Ableitung hängt sowohl von x als auch vom Wert von φ und seiner Ableitungen bis zur $(n-1)$-ten Ordnung im Punkt x ab.

Reduktion auf ein System 1. Ordnung

Man kann eine Differentialgleichung n-ter Ordnung auf ein System von Differentialgleichungen 1. Ordnung zurückführen. Dazu betrachten wir neben (3) das Differentialgleichungssystem

$$\left\{ \begin{aligned} y_0' &= y_1, \\ y_1' &= y_2, \\ &\ \vdots \\ y_{n-2}' &= y_{n-1}, \\ y_{n-1}' &= f(x, y_0, y_1, \dots, y_{n-1}). \end{aligned} \right. \tag{4}$$

Das System (4) ist äquivalent zu

$$Y' = F(x, Y), \quad Y := \begin{pmatrix} y_0 \\ \vdots \\ y_{n-2} \\ y_{n-1} \end{pmatrix}, \quad F(x, Y) := \begin{pmatrix} y_1 \\ \vdots \\ y_{n-1} \\ f(x, Y) \end{pmatrix}.$$

Es sei nun $\varphi\colon I \to \mathbb{R}$ eine Lösung von (3), d. h.

$$\varphi^{(n)}(x) = f(x, \varphi(x), \ldots, \varphi^{(n-1)}(x)).$$

Dann ist durch

$$\Phi(x) := \begin{pmatrix} \varphi(x) \\ \varphi'(x) \\ \vdots \\ \varphi^{(n-1)}(x) \end{pmatrix}, \quad x \in I$$

eine Lösung $\Phi\colon I \to \mathbb{R}^n$ von (4) definiert.

Sei umgekehrt

$$\Phi := \begin{pmatrix} \varphi_0 \\ \vdots \\ \varphi_{n-1} \end{pmatrix} : I \to \mathbb{R}^n$$

eine Lösung von (4). Dann ist $\varphi := \varphi_0\colon I \to \mathbb{R}$ eine Lösung von (3): Aus den ersten $(n-1)$ Gleichungen von (4) folgt

$$\begin{aligned} \varphi_1 &= \varphi_0' = \varphi', \\ \varphi_2 &= \varphi_1' = \varphi'', \\ &\vdots \\ \varphi_{n-1} &= \varphi_{n-2}' = \varphi^{(n-1)}. \end{aligned}$$

Da φ_{n-1} einmal differenzierbar ist, folgt daraus, dass φ n-mal differenzierbar ist. Die n-te Gleichung von (4) liefert dann

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

Aufgabe 8.2 Man bestimme zu der Differentialgleichung

$$m\ddot{x} = F(t, x, \dot{x}) \qquad (m = \text{ Masse}, F = \text{ Kraft})$$

(Newtonsches Bewegungsgesetz) ein zugehöriges System von Differentialgleichungen.

Wir werden jetzt eine Bedingung kennenlernen, die Existenz und Eindeutigkeit der Lösungen von gewöhnlichen Differentialgleichungen garantiert.

Definition Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $f\colon G \to \mathbb{R}^n$ eine Abbildung. Man sagt, f genüge in G einer *Lipschitz-Bedingung* mit der Lipschitz-Konstanten $L \geq 0$, wenn für alle $(x, y), (x, \tilde{y}) \in G$ gilt

$$||f(x, y) - f(x, \tilde{y})|| \leq L||y - \tilde{y}||.$$

Man sagt, f genüge in G *lokal einer Lipschitzbedingung*, falls jeder Punkt $(a, b) \in G$ eine Umgebung U besitzt, so dass f in $G \cap U$ einer Lipschitz-Bedingung mit einer gewissen (von U abhängigen) Konstanten $L \in \mathbb{R}_+$ genügt.

Satz 8.1 *Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und*

$$f\colon \begin{array}{ccc} G & \to & \mathbb{R}^n \\ \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & f(x,y) \end{array}$$

eine Abbildung, deren partielle Ableitungen nach den Variablen $y_1, \dots, y_n$

$$\frac{\partial f}{\partial y_1}(x,y), \dots, \frac{\partial f}{\partial y_n}(x,y)$$

existieren und stetig sind. Dann genügt f in G lokal einer Lipschitz-Bedingung.

Beweis. Es sei $(a,b) \in G$. Es gibt dann ein $r > 0$, so dass

$$V := \{(x,y) \in \mathbb{R} \times \mathbb{R}^n \mid |x-a| \le r, ||y-b|| \le r\} \subset G.$$

Die Menge V ist eine kompakte Umgebung von (a,b). Da nach Voraussetzung alle Einträge der $n \times n$-Matrix $\frac{\partial f}{\partial y}(x,y)$ stetige Funktionen sind, gilt

$$L := \sup_{(x,y)\in V} \left|\left| \frac{\partial f}{\partial y}(x,y) \right|\right| < \infty.$$

Aus dem Schrankensatz folgt nun für alle $(x,y), (x,\tilde{y}) \in V$:

$$||f(x,y) - f(x,\tilde{y})|| \le L||y - \tilde{y}||.$$

□

Satz 8.2 (Satz von Picard-Lindelöf) *Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f\colon G \to \mathbb{R}^n$ eine stetige Abbildung, die lokal einer Lipschitz-Bedingung genügt.*

Dann gibt es zu jedem $(a,c) \in G$ ein $\delta > 0$, so dass das Differentialgleichungssystem

$$y' = f(x,y) \qquad (x \in [a-\delta, a+\delta])$$

genau eine Lösung $\varphi\colon [a-\delta, a+\delta] \to \mathbb{R}^n$ auf dem Intervall $[a-\delta, a+\delta]$ mit der Anfangsbedingung $\varphi(a) = c$ hat.

Beweis. Wir wollen den Banachschen Fixpunktsatz anwenden.

a) Es gibt ein $r > 0$, so dass die Menge

$$V := \{(x,y) \in \mathbb{R} \times \mathbb{R}^n \mid |x-a| \le r, ||y-c|| \le r\}$$

ganz in G enthalten ist und f in V einer Lipschitz-Bedingung mit einer Konstanten L genügt. Wähle

$$0 < \delta < \min\left(r, \frac{1}{2L}\right)$$

und setze

$$I := [a-\delta, a+\delta].$$

b) Eine stetige Abbildung $\varphi\colon I \to \mathbb{R}^n$ genügt genau dann der Differentialgleichung

$$y' = f(x,y)$$

mit der Anfangsbedingung $\varphi(a) = c$, wenn

$$\varphi(x) = c + \int_a^x f(t, \varphi(t))\, dt \text{ für alle } x \in I.$$

(Dies folgt unmittelbar aus dem Hauptsatz der Differential- und Integralrechnung.) Die Abbildung φ ist demnach Fixelement der Abbildung

$$\Psi\colon \mathcal{C}(I) \to \mathcal{C}(I),$$

die erklärt ist durch

$$(\Psi(\varphi))(x) = c + \int_a^x f(t, \varphi(t))\, dt \text{ für alle } x \in I.$$

Da I ein abgeschlossenes Intervall ist, ist $\mathcal{C}(I)$ mit der Metrik

$$d(\varphi, \tilde{\varphi}) := \max_{x \in I} ||\varphi(x) - \tilde{\varphi}(x)||$$

ein vollständiger metrischer Raum. Wir zeigen nun, dass Φ eine kontrahierende Selbstabbildung dieses Raumes ist. Aus

$$(\Psi(\varphi))(x) - (\Psi(\tilde{\varphi}))(x) = \int_a^x (f(t, \varphi(t) - f(t, \tilde{\varphi}(t)))\, dt$$

folgt

$$\begin{aligned}
||(\Psi(\varphi))(x) - (\Psi(\tilde{\varphi}))(x)|| &\leq \operatorname{sign}(x-a) \int_a^x ||f(t, \varphi(t)) - f(t, \tilde{\varphi}(t))||\, dt \\
&\leq \left| \int_a^x L||\varphi(t) - \tilde{\varphi}(t)||\, dt \right| \\
&\leq \int_{a-\delta}^{a+\delta} L||\varphi - \tilde{\varphi}||\, dt \\
&= 2\delta L||\varphi - \tilde{\varphi}||.
\end{aligned}$$

Also folgt

$$||\Psi(\varphi) - \Psi(\tilde{\varphi})|| \leq 2\delta L||\varphi - \tilde{\varphi}||.$$

Nach Wahl von δ gilt $2\delta L < 1$, also ist $\Psi\colon \mathcal{C}(I) \to \mathcal{C}(I)$ kontrahierend.

c) Nach dem Banachschen Fixpunktsatz gibt es dann genau ein $\varphi \in \mathcal{C}(I)$ für das gilt

$$\varphi(x) = c + \int_a^x f(t, \varphi(t))\, dt \text{ für alle } x \in I.$$

Da der Integrand stetig ist, ist φ sogar differenzierbar und es gilt

$$\begin{aligned}
\varphi'(x) &= f(t, \varphi(t)) \\
\varphi(a) &= c
\end{aligned}$$

Also gibt es genau eine Lösung φ der Differentialgleichung $y' = f(x, y)$ mit der Anfangsbedingung $\varphi(a) = c$ im Intervall $[a - \delta, a + \delta]$. □

Bemerkung 8.2 Beim Beweis des Banachschen Fixpunktsatzes hatten wir gesehen, dass man das Fixelement φ durch das Iterationsverfahren

$$\varphi_{k+1} = \Psi(\varphi_k)$$

erhält. In unserem Fall bedeutet das, dass man Abbildungen $\varphi_k \colon I \to \mathbb{R}^n, k \in \mathbb{N}$, definiert durch $\varphi_0(x) = c$ für alle $x \in I$ und

$$\varphi_{k+1}(x) := c + \int_a^x f(t, \varphi_k(t))\, dt.$$

Die Folge (φ_k) konvergiert dann gleichmäßig gegen die gesuchte Lösung φ. Dieses Verfahren nennt man das *Picard-Lindelöfsche Iterationsverfahren.* In einfachen Fällen kann man mittels dieses Verfahrens die Lösungen einer Differentialgleichung explizit berechnen.

Aufgabe 8.3 Man bestimme mit dem Picard-Lindelöfschen Iterationsverfahren eine Lösung des Anfangswertproblems

$$\begin{aligned} y' &= 2xy \text{ in } \mathbb{R} \times \mathbb{R}, \\ \varphi(0) &= c. \end{aligned}$$

Wir wollen nun den Existenz- und Eindeutigkeitssatz auf Differentialgleichungen n-ter Ordnung

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

übertragen, wobei $f \colon G \to \mathbb{R}$ eine auf einer Teilmenge $G \subset \mathbb{R} \times \mathbb{R}^n$ definierte stetige Funktion ist.

Man sagt, f genüge lokal einer Lipschitz-Bedingung, wenn es zu jedem Punkt $z \in G$ eine Umgebung U und eine Konstante $L \in \mathbb{R}_+$ gibt, so dass

$$|f(x, Y) - f(x, \tilde{Y})| \le L||Y - \tilde{Y}||,$$

für alle (x, Y), $(x, \tilde{Y}) \in G \cap U$, $Y = \begin{pmatrix} y_0 \\ \vdots \\ y_{n-1} \end{pmatrix}$, $\tilde{Y} = \begin{pmatrix} \tilde{y}_0 \\ \vdots \\ \tilde{y}_{n-1} \end{pmatrix}$.

Korollar 8.1 *Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f \colon G \to \mathbb{R}$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung genügt.*

Dann gibt es zu jedem $(a, c_0, \dots, c_{n-1}) \in G$ ein $\delta > 0$, so dass die Differentialgleichung

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

genau eine Lösung $\varphi \colon [a - \delta, a + \delta] \to \mathbb{R}$ auf dem Intervall $[a - \delta, a + \delta]$ mit der Anfangsbedingung

$$\varphi(a) = c_0, \varphi'(a) = c_1, \dots, \varphi^{(n-1)}(a) = c_{n-1}$$

hat.

Beweis. Dies folgt aus Satz 8.2 durch Reduktion auf ein Differentialgleichungssystem 1. Ordnung. $\square$

Wir können aus Satz 8.2 noch eine schärfere Eindeutigkeitsaussage ableiten:

Satz 8.3 *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f\colon G \to \mathbb{R}^n$ eine stetige Abbildung, die lokal einer Lipschitz-Bedingung genügt. Seien*

$$\varphi, \psi\colon I \to \mathbb{R}^n$$

zwei Lösungen der Differentialgleichung

$$y' = f(x, y)$$

über einem Intervall $I \subset \mathbb{R}$. Gilt dann

$$\varphi(a) = \psi(a) \text{ für ein } a \in I,$$

so folgt

$$\varphi(x) = \psi(x) \text{ für alle } x \in I.$$

Beweis. a) Wir zeigen zunächst

$$\varphi(x) = \psi(x) \text{ für alle } x \in I \text{ mit } x \geq a.$$

Es sei

$$s := \sup\{x \in I \mid \varphi|_{[a,x]} = \psi|_{[a,x]}\}.$$

Falls $s = \infty$ oder $s =$ rechte Intervallgrenze, sind wir fertig. Andernfalls gibt es ein $\varepsilon > 0$, so dass $[s, s+\varepsilon] \subset I$. Da φ und ψ stetig sind, gilt $\varphi(s) = \psi(s)$. Nach Satz 8.2 gibt es ein $\delta > 0$, so dass

$$\varphi(x) = \psi(x) \text{ für alle } x \in I \text{ mit } |x - s| \leq \delta.$$

Dies steht aber im Widerspruch zur Definition von s. Daher gilt $\varphi(x) = \psi(x)$ für alle $x \in I$ mit $x \geq a$.

b) Analog zeigt man $\varphi(x) = \psi(x)$ für alle $x \in I$ mit $x \leq a$. □

Korollar 8.2 *Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f\colon G \to \mathbb{R}$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung genügt. Seien*

$$\varphi, \psi\colon I \to \mathbb{R}$$

zwei Lösungen der Differentialgleichung

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

Für einen Punkt $a \in I$ gelte

$$\varphi(a) = \psi(a), \varphi'(a) = \psi'(a), \ldots, \varphi^{(n-1)}(a) = \psi^{(n-1)}(a).$$

Dann gilt

$$\varphi(x) = \psi(x) \text{ für alle } x \in I.$$

Beweis. Dies folgt aus Satz 8.3 durch Reduktion auf eine System 1. Ordnung. □

Um eine Lösung einer Differentialgleichung n-ter Ordnung eindeutig festzulegen, muß man also nicht nur den Wert der Funktion an einer Stelle a des Definitionsintervalls vorschreiben, sondern auch alle Ableitungen der Ordnung $\leq n-1$ in a.

Beispiel 8.1 Die Differentialgleichung 2. Ordnung

$$y'' + y = 0$$

besitzt offenbar die auf $\mathbb{R}$ definierten Lösungen

$$y = \cos x \text{ und } y = \sin x$$

Allgemeine Lösung:

$$\varphi(x) = c_0 \cos x + c_1 \sin x \qquad (c_0, c_1 \in \mathbb{R})$$

(einzige Lösung mit $\varphi(0) = c_0, \varphi'(0) = c_1$, aber c_0, c_1 beliebig aus $\mathbb{R}$).

Nicht immer kann man die Lösungen einer vorgegebenen Differentialgleichung explizit angeben. Wir behanden nun einige einfache Fälle, in denen das möglich ist.

Differentialgleichungen mit getrennten Variablen

Es seien $I, J \subset \mathbb{R}$ offene Intervalle, $f\colon I \to \mathbb{R}$ und $g\colon J \to \mathbb{R}$ zwei stetige Funktionen mit $g(y) \neq 0$ für alle $y \in J$. Die Differentialgleichung

$$(5) \qquad y' = f(x)g(y) \text{ in } I \times J$$

heißt *Differentialgleichung mit getrennten Variablen.*

Satz 8.4 *Mit obigen Bezeichnungen sei $(x_0, y_0) \in I \times J$ ein Punkt. Wir definieren Funktionen $F\colon I \to \mathbb{R}, G\colon J \to \mathbb{R}$ durch*

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

Es sei $I' \subset I$ ein Intervall mit $x_0 \in I'$ und $F(I') \subset G(J)$.

Dann existiert genau eine Lösung

$$\varphi\colon I' \to \mathbb{R}$$

der Differentialgleichung (5) *mit*

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

Diese Lösung genügt der Beziehung

$$(6) \qquad G(\varphi(x)) = F(x) \textit{ für alle } x \in I'.$$

Beweis. a) Wir zeigen zunächst: Ist $\varphi\colon I' \to \mathbb{R}$ eine Lösung von (5) mit $\varphi(x_0) = y_0$, so gilt (6). Aus

$$\varphi'(x) = f(x)g(\varphi(x))$$

folgt

$$\int_{x_0}^{x} \frac{\varphi'(t)}{g(\varphi(t))}\,dt = \int_{x_0}^{x} f(t)\,dt.$$

Mit der Substitution $u = \varphi(t)$, $du = \varphi'(t)\,dt$ folgt daraus

$$\int_{y_0}^{\varphi(x)} \frac{du}{g(u)} = \int_{x_0}^{x} f(t)\,dt.$$

Dies bedeutet aber

$$G(\varphi(x)) = F(x).$$

b) Wir zeigen nun die Eindeutigkeit der Lösung mit der gegebenen Anfangsbedingung, falls sie existiert.

Da $G'(y) = \frac{1}{g(y)} \neq 0$, ist G streng monoton, besitzt also eine stetig differenzierbare Umkehrfunktion

$$H\colon G(J) \to \mathbb{R}$$

Aus (6) folgt daher

(7) $$\varphi(x) = H(F(x)) \text{ für alle } x \in I'.$$

Daraus folgt die Eindeutigkeit.

c) Wir zeigen nun die Existenz. Nehme (7) als Definition für $\varphi\colon I' \to \mathbb{R}$. Die Funktion φ ist stetig differenzierbar und es gilt

$$\varphi(x_0) = H(F(x_0)) = H(0) = y_0,$$

da $F(x_0) = 0 = G(y_0)$. Nun gilt

$$\varphi(x) = H(F(x)).$$

Daraus folgt

$$G(\varphi(x)) = F(x).$$

Differentiation ergibt

$$G'(\varphi(x))\varphi'(x) = \frac{\varphi'(x)}{g(\varphi(x))} = F'(x) = f(x),$$

also

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

Also ist φ Lösung von (5). □

Bemerkung 8.3 Man kann wie folgt formal rechnen und sich damit den eben bewiesenen Satz merken: Man schreibt $y' = f(x)g(y)$ als

$$\frac{dy}{dx} = f(x)g(y).$$

Daraus ergibt sich

$$\frac{dy}{g(y)} = f(x)\,dx,$$

also

$$\int \frac{dy}{g(y)} = \int f(x)\,dx + \text{const.}$$

Diese Gleichung hat man nach y aufzulösen, um eine Lösung $y = \varphi(x)$ der Differentialgleichung zu erhalten.

Aufgabe 8.4 Man bestimme die Lösungen des Anfangswertproblems

$$\begin{aligned} y' &= y^2 \quad \text{im } \mathbb{R}^2, \\ \varphi(0) &= c. \end{aligned}$$

(Für $c > 0$ ist $\varphi(x) = \frac{c}{1-cx}$ die eindeutig bestimmte Lösung, die auf dem Intervall $(-\infty, \frac{1}{c})$ definiert ist. Diese Lösung ist nicht auf ganz $\mathbb{R}$ fortsetzbar.)

9 Lineare Differentialgleichungen

Es sei $I \subset \mathbb{R}$ ein Intervall und $a, b\colon I \to \mathbb{R}$ seien stetige Funktionen. Dann nennt man

$$y' = a(x)y + b(x)$$

eine *lineare Differentialgleichung 1. Ordnung*, und zwar *homogen*, falls $b = 0$, sonst *inhomogen*.

Wir beschäftigen uns zunächst mit der Lösung der homogenen linearen Differentialgleichung

$$y' = a(x)y. \tag{1}$$

Satz 9.1 *Mit den obigen Bezeichnungen gilt: Sei $x_0 \in I$ und $c \in \mathbb{R}$. Dann gibt es genau eine Lösung $\varphi\colon I \to \mathbb{R}$ der Differentialgleichung $y' = a(x)y$ mit $\varphi(x_0) = c$, nämlich:*

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

Beweis. (1) ist ein Spezialfall von §8 (5), man könnte also Satz 8.4 anwenden. Hier aber folgt direkt aus der Definition von φ:

$$\varphi'(x) = a(x)\varphi(x), \qquad \varphi(x_0) = c.$$

Da die Funktion $f(x, y) = a(x)y$ in $I \times \mathbb{R}$ stetig partiell nach y differenzierbar ist, also die Lipschitz-Bedingung erfüllt, ist die Lösung eindeutig. □

Aufgabe 9.1 Man bestimme die Lösungen der Differentialgleichung $y' = ky$ ($k \in \mathbb{R}$) in $\mathbb{R}$.

Variation der Konstanten

Wir wollen nun den inhomogenen Fall

$$y' = a(x)y + b(x) \tag{2}$$

behandeln. Es sei $\varphi\colon I \to \mathbb{R}$ eine Lösung der zugehörigen homogenen Differentialgleichung, d. h.

$$\varphi'(x) = a(x)\varphi(x).$$

Wir setzen voraus, dass $\varphi(x_0) \neq 0$. Daraus folgt nach Satz 9.1, dass $\varphi(x) \neq 0$ für alle $x \in I$. Daher lässt sich eine beliebige Lösung $\psi\colon I \to \mathbb{R}$ der Gleichung (2) schreiben als

$$\psi(x) = \varphi(x)u(x)$$

mit einer stetig differenzierbaren Funktion $u\colon I \to \mathbb{R}$. Wir untersuchen jetzt, welcher Bedingung u erfüllen muss, damit ψ Lösung von (2) ist: Es gilt

$$\begin{aligned} \psi' &= \varphi' u + \varphi u', \\ a\psi + b &= a\varphi u + b. \end{aligned}$$

Wegen $\varphi' = a\varphi$ folgt

$$\psi' = a\psi + b \Leftrightarrow \varphi u' = b \Leftrightarrow u(x) = \int_{x_0}^{x} \frac{b(t)}{\varphi(t)}\, dt + c,$$

wobei c durch $\psi(x_0) = c\varphi(x_0)$ bestimmt ist.

Damit haben wir bewiesen:

Satz 9.2 *Es sei $I \subset \mathbb{R}$ ein Intervall und $a, b\colon I \to \mathbb{R}$ stetige Funktionen. Dann gibt es zu $x_0 \in I$ und $c \in \mathbb{R}$ genau eine Lösung $\psi\colon I \to \mathbb{R}$ der Differentialgleichung*

$$y' = a(x)y + b(x)$$

mit der Anfangsbedingung $\psi(x_0) = c$, nämlich

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

wobei

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

(„Variation der Konstanten“*: Lösungen der homogenen Gleichung: $\varphi(x)c$; ersetze c durch „variable“ Funktion $u(x)$)*

Aufgabe 9.2 Man bestimme die Lösungen der Differentialgleichung $y' = 2xy + x^3$.

Wir betrachten nun den Fall eines Systems von linearen Differentialgleichungen

Definition Es sei $I \subset \mathbb{R}$ ein Intervall. Die Gleichung

(3) $$y' = A(x)y,$$

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} : I \to M(n \times n, \mathbb{R})$$

stetige Abbildung, heißt *homogene(s) lineare(s) Differentialgleichung(ssystem)*. Weiter sei

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} : I \to \mathbb{R}^n$$

eine stetige Abbildung. Dann heißt

(4) $$y' = A(x)y + b(x)$$

ein *inhomogene(s) lineare(s) Differentialgleichung(ssystem)*. Das System (3) heißt das dem System (4) zugeordnete homogene Differentialgleichungssystem.

Es ist zweckmäßig, auch komplexe lineare Differentialgleichungen zu betrachten. Dabei sind A und b stetige Abbildungen $A\colon I \to M(n \times n, \mathbb{C})$, $b\colon I \to \mathbb{C}^n$, und eine Lösung ist eine differenzierbare Abbildung

$$\varphi\colon I \to \mathbb{C}^n = \mathbb{R}^{2n}$$

mit

$$\varphi'(x) = A(x)\varphi(x) + b(x) \text{ für alle } x \in I.$$

Wegen $\mathbb{C} = \mathbb{R}^2, \mathbb{C}^n = \mathbb{R}^{2n}$, ist ein System von n komplexen Differentialgleichungen äquivalent zu einem System von $2n$ reellen Differentialgleichungen.

Warnung Es sind zwar komplexe Werte zugelassen, die Variable x ist aber immer reell ($I \subset \mathbb{R}$)!

Es sei im Folgenden $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Satz 9.3 *Es sei $I \subset \mathbb{R}$ ein offenes Intervall,*

$$A\colon I \to M(n \times n, \mathbb{K}), \quad b\colon I \to \mathbb{K}^n$$

stetige Abbildungen. Dann gibt es zu jedem $x_0 \in I$ und $c \in \mathbb{K}^n$ genau eine Lösung

$$\varphi\colon I \to \mathbb{K}^n$$

der linearen Differentialgleichung

$$y' = A(x)y + b(x)$$

mit der Anfangsbedingung $\varphi(x_0) = c$.

Bemerkung 9.1 I.A. kann bei einer Differentialgleichung $y' = f(x, y)$ eine Lösung nicht über dem ganzen Intervall $I \subset \mathbb{R}$ gefunden werden, selbst wenn f auf $I \times \mathbb{R}^n$ definiert und dort lokal einer Lipschitz-Bedingung genügt, siehe $y' = y^2$.

Beweis von Satz 3. Wir setzen

$$f(x, y) := A(x)y + b(x).$$

a) Wir zeigen zunächst: Ist $J \subset I$ irgendein kompaktes Teilintervall, so genügt f in $J \times \mathbb{K}^n$ einer globalen Lipschitz-Bedingung.

Beweis. Da A stetig, folgt

$$L := \sup_{x \in J} ||A(x)|| < \infty.$$

Damit erhält man für $x \in J$, $y, \tilde{y} \in \mathbb{K}^n$.

$$||f(x, y) - f(x, \tilde{y})|| = ||A(x)(y - \tilde{y})|| \leq L||y - \tilde{y}||.$$

□

b) Nach Satz 8.3 folgt aus a) die Eindeutigkeit der Lösung.

c) Zur Existenz: Es sei $J \subset I$ ein beliebiges kompaktes Teilintervall von I mit $x_0 \in J$. Nach Satz 8.2 existiert eine Lösung $\varphi\colon [x_0-\delta, x_0+\delta] \to \mathbb{K}^n$. Dabei hängt nach dem Beweis dieses Satzes δ nur von J und der zu J gehörigen globalen Lipschitz-Konstante L ab. Indem man Satz 8.2 auf die Punkte $x_0-\delta$ und $x_0+\delta$ mit den Anfangsbedingungen $\varphi(x_0-\delta)$ bzw. $\varphi(x_0+\delta)$ anwendet, erhält man eine und nur eine Lösung im doppelt so langen Intervall $[x_0-2\delta, x_0+2\delta]$. Durch mehrmalige Anwendung dieses Fortsetzungsprozesses kann man die Lösung φ auf das ganze Intervall J fortsetzen.

Da J ein beliebiges kompaktes Teilintervall von I ist und wegen der Eindeutigkeit erhält man die Lösung sogar auf ganz I. $\square$

Satz 9.4 *Es sei $I \subset \mathbb{R}$ Intervall,*

$$A\colon I \to M(n \times n, \mathbb{K})$$

eine stetige Abbildung. Wir bezeichnen mit L_H die Menge aller Lösungen $\varphi\colon I \to \mathbb{K}^n$ der homogenen linearen Differentialgleichung

$$y' = A(x)y.$$

Dann ist L_H ein n-dimensionaler Vektorraum über $\mathbb{K}$. Für ein k-Tupel von Lösungen $\varphi_1, \dots, \varphi_k \in L_H$ sind folgende Aussagen äquivalent:

(i) *$\varphi_1, \dots, \varphi_k$ sind linear unabhängig über $\mathbb{K}$.*

(ii) *Es existiert ein $x_0 \in I$, so dass die Vektoren $\varphi_1(x_0), \dots, \varphi_k(x_0) \in \mathbb{K}^n$ linear unabhängig über $\mathbb{K}$ sind.*

(iii) *Für jedes $x_0 \in I$ sind die Vektoren $\varphi_1(x_0), \dots, \varphi_k(x_0) \in \mathbb{K}^n$ linear unabhängig über $\mathbb{K}$.*

Beweis. a) Wir zeigen: L_H ist Unterraum des Vektorraums aller Abbildungen $f\colon I \to \mathbb{K}^n$.

1) $0 \in L_H$

2) $\varphi, \psi \in L_H \Rightarrow \varphi' = A\varphi, \psi' = A\psi$
$\Rightarrow (\varphi+\psi)' = \varphi' + \psi' = A\varphi + A\psi = A(\varphi+\psi)$
$\Rightarrow \varphi + \psi \in L_H$

3) $\varphi \in L_H, \lambda \in \mathbb{K}$
$\Rightarrow (\lambda\varphi)' = \lambda\varphi' = \lambda A\varphi = A(\lambda\varphi)$
$\Rightarrow \lambda\varphi \in L_H$.

b) (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) trivial.

(i) $\Rightarrow$ (iii): Es seien $\varphi_1, \dots, \varphi_k \in L_H$ linear unabhängig und $x_0 \in I$. Es seien $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ mit

$$\lambda_1\varphi_1(x_0) + \dots + \lambda_k\varphi_k(x_0) = 0.$$

Dann folgt

$$\varphi := \lambda_1\varphi_1 + \dots + \lambda_k\varphi_k \in L_H, \quad \varphi(x_0) = 0.$$

Wegen der Eindeutigkeit der Lösung folgt daraus $\varphi = 0$. Da $\varphi_1, \ldots, \varphi_k$ linear unabhängig sind, ergibt sich

$$\lambda_1 = \cdots = \lambda_k = 0.$$

c) Wir zeigen: $\dim L_H = n$.

Es sei $\{e_1, \ldots, e_n\}$ die kanonische Basis von $\mathbb{K}^n$. Nach Satz 9.3 gibt es Lösungen $\varphi_1, \ldots, \varphi_n \in L_H$ mit $\varphi_i(x_0) = e_i$. Nach b) sind diese Lösungen dann linear unabhängig, d.h. es gilt $\dim L_H \geq n$.

Gäbe es $n+1$ linear unabhängige Lösungen $\psi_1, \ldots, \psi_{n+1}$, so müssten nach b) die Vektoren $\psi_1(x_0), \ldots, \psi_{n+1}(x_0) \in \mathbb{K}^n$ linear unabhängig sein, Widerspruch zu $\dim \mathbb{K}^n = n$. Also $\dim L_H \leq n$. □

Definition Eine Basis $(\varphi_1, \ldots, \varphi_n)$ des Vektorraums der Lösungen der Differentialgleichung $y' = A(x)y$ nennt man ein *Lösungs-Fundamentalsystem* dieser Gleichung.

Hat man eine beliebiges n-Tupel von Lösungen $(\varphi_1, \ldots, \varphi_n)$, so kann man die Lösungen φ_i als Spaltenvektoren

$$\varphi_i = \begin{pmatrix} \varphi_{1i} \\ \varphi_{2i} \\ \vdots \\ \varphi_{ni} \end{pmatrix}, \quad i = 1, \ldots, n,$$

schreiben. Dann ist $\Phi = (\varphi_1, \ldots, \varphi_n)$,

$$\Phi = \begin{pmatrix} \varphi_{11} & \cdots & \varphi_{1n} \\ \vdots & & \vdots \\ \varphi_{n1} & \cdots & \varphi_{nn} \end{pmatrix},$$

eine $n \times n$-Matrix. Nach Satz 9.4 gilt:

$$\begin{aligned} &(\varphi_1, \ldots, \varphi_n) \text{ Lösungs-Fundamentalsystem} \\ &\quad \Leftrightarrow \quad \det \Phi(x_0) \neq 0 \text{ für wenigstens ein } x_0 \in I \\ &\quad \Leftrightarrow \quad \det \Phi(x) \neq 0 \text{ für alle } x \in I. \end{aligned}$$

Ist $\Phi = (\varphi_1, \ldots, \varphi_n)$ ein Lösungs-Fundamentalsystem von $y' = A(x)y$, so erhält man eine beliebige Lösung φ durch

$$\varphi = c_1\varphi_1 + \cdots + c_n\varphi_n = \Phi c, \quad c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{K}^n.$$

Man kann die Matrix $\Phi = (\varphi_1, \ldots, \varphi_n)$ eines Lösungs-Fundamentalsystems selbst als Lösung der Differentialgleichung auffassen:

$$\begin{aligned} \Phi' &= (\varphi_1', \ldots, \varphi_n'), \\ A\Phi &= (A\varphi_1, \ldots, A\varphi_n), \\ \Phi' &= A\Phi. \end{aligned}$$

Beispiel 9.1 Man bestimme ein Lösungsfundamentalsystem des Differentialsgleichungssystems

$$\begin{cases} y_1' &= -\omega y_2, \\ y_2' &= \omega y_1, \end{cases} \Leftrightarrow \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

wobei $\omega \in \mathbb{R}$ eine Konstante ist.

Lösungen: $\varphi_i \colon \mathbb{R} \to \mathbb{R}^2$:

$$\varphi_1(x) := \begin{pmatrix} \cos \omega x \\ \sin \omega x \end{pmatrix}, \qquad \varphi_2(x) := \begin{pmatrix} -\sin \omega x \\ \cos \omega x \end{pmatrix}$$

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

Es gilt

$$\det \Phi(x) = \cos^2 \omega x + \sin^2 \omega x = 1 \text{ für alle } x.$$

Daraus folgt, dass (φ_1, φ_2) ein Lösungs-Fundamentalsystem ist.

Satz 9.5 *Es sei $I \subset \mathbb{R}$ ein Intervall,*

$$A \colon I \to M(n \times n, \mathbb{K}), \quad b \colon I \to \mathbb{K}^n$$

stetige Abbildungen. Wir bezeichnen mit L_I die Menge der Lösungen $\psi \colon I \to \mathbb{K}^n$ der inhomogenen linearen Differentialgleichung

$$y' = A(x)y + b(x)$$

und mit L_H den Vektorraum aller Lösungen der zugehörigen homogenen Differentialgleichung. Dann gilt für ein beliebiges $\psi_s \in L_I$

$$L_I = \psi_s + L_H.$$

Beweis. a) $L_I \subset \psi_s + L_H$:

Sei $\psi \in L_I$. Setze $\varphi := \psi - \psi_s$.

$\Rightarrow \varphi' = \psi' - \psi_s' = (A\psi + b) - (A\psi_s + b) = A(\psi_s - \psi) = A\varphi$

$\Rightarrow \varphi \in L_H$

$\Rightarrow \psi = \psi_s + \varphi \in \psi_s + L_H$

b) $\psi_s + L_H \subset L_I$:

Sei $\psi \in \psi_s + L_H$

$\Rightarrow \psi = \psi_s + \varphi, \quad \varphi \in L_H$

$\Rightarrow \psi' = \psi_s' + \varphi' = (A\psi_s + b) + A\varphi = A\psi + b$

$\Rightarrow \psi \in L_I$. □

Wie erhält man nun eine spezielle Lösung ψ_s der inhomogenen Gleichung?

Satz 9.6 (Variation der Konstanten) *Mit den Bezeichnungen von Satz 9.3 gilt: Sei $\Phi = (\varphi_1, \dots, \varphi_n)$ ein Lösungs-Fundamentalsystem des homogenen Systems*

$$y' = A(x)y.$$

Dann erhält man eine Lösung $\psi\colon I \to \mathbb{K}^n$ *des inhomogenen Systems*

$$y' = A(x)y + b(x)$$

durch den Ansatz

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

Dabei ist $u\colon I \to \mathbb{K}^n$ *eine differenzierbare Funktion mit*

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

Beweis.

$$\psi = \Phi u \Rightarrow \begin{cases} \psi' & = \Phi' u + \Phi u' \\ A\psi + b & = A\Phi u + b \end{cases}$$

Wegen $\Phi' = A\Phi$ gilt:

$$\psi' = A\psi + b \Leftrightarrow \Phi u' = b \Leftrightarrow u' = \int_{x_0}^{x} \Phi(t)^{-1} b(t)\, dt + \text{const.}$$

□

Aufgabe 9.3 Man bestimme die Lösungen des Differentialgleichungssystems

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

Wir übertragen jetzt die bewiesenen Resultate über lineare Differentialgleichungssystem 1. Ordnung auf lineare Differentialgleichungen n-ter Ordnung.

Es sei $I \subset \mathbb{R}$ ein Intervall, $a_k\colon I \to \mathbb{K}$, $0 \le k \le n-1$ stetige Funktionen, $b\colon I \to \mathbb{K}$ stetige Funktion. Die Gleichung

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = 0 \tag{5}$$

heißt *homogene lineare Differentialgleichung n-ter Ordnung*. Die Gleichung

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = b(x) \tag{6}$$

heißt *inhomogene lineare Differentialgleichung n-ter Ordnung*. Die Gleichung (5) heißt die (6) zugeordnete homogene Gleichung.

Satz 9.7 *Mit obigen Bezeichnungen gilt*

(a) *Sei* L_H *die Menge aller Lösungen* $\varphi\colon I \to \mathbb{K}$ *der homogenen Differentialgleichung* (5)*. Dann ist* L_H *ein n-dimensionaler Vektorraum.*

(b) *Es sei* L_I *die Menge aller Lösungen* $\psi\colon I \to \mathbb{K}$ *der inhomogenen Differentialgleichung* (6)*. Dann gilt für ein beliebiges* $\psi_s \in L_I$

$$L_I = \psi_s + L_H.$$

Beweis. Die Differentialgleichung (6) ist äquivalent mit dem inhomogenen linearen Differentialgleichungssystem 1. Ordnung:

$$(7)\qquad \begin{cases} y_0' &= y_1 \\ y_1' &= y_2 \\ &\vdots \\ y_{n-2}' &= y_{n-1} \\ y_{n-1}' &= -a_0(x)y_0 - a_1(x)y_1 - \cdots - a_{n-1}(x)y_{n-1} + b(x) \end{cases}$$

Jeder Lösung $\varphi\colon I \to \mathbb{K}$ von (6) entspricht eine Lösung

$$f = \begin{pmatrix} \varphi \\ \varphi' \\ \vdots \\ \varphi^{(n-1)} \end{pmatrix} : I \to \mathbb{K}^n$$

von (7) und umgekehrt. Entsprechend für $b = 0$. Damit folgen die Behauptungen aus den Sätzen 9.4 und 9.5. □

Definition Eine Basis $(\varphi_1, \ldots, \varphi_n)$ von L_H heißt *Lösungs-Fundamentalsystem* von (5).

Definition Es sei $(\varphi_1, \ldots, \varphi_n)$ ein n-Tupel von Lösungen der homogenen Gleichung (5). Dann heißt

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

die *Wronski-Determinante* von $(\varphi_1, \ldots, \varphi_n)$.

Satz 9.8 *Ein n-Tupel $(\varphi_1, \ldots, \varphi_n)$ von Lösungen der homogenen Gleichung* (5) *ist genau dann ein Lösungs-Fundamentalsystem, wenn für ein und damit für alle $x \in I$ die Wronski-Determinante von Null verschieden ist.*

Beweis. Dies folgt wie Satz 9.7 aus Satz 9.4. □

Aufgabe 9.4 Man bestimme die Lösungen der Differentialgleichung

$$y'' - \frac{1}{2x}y' + \frac{1}{2x^2}y = 0 \text{ auf } I = \mathbb{R}_+^*.$$

10 Lineare Differentialgleichungen mit konstanten Koeffizienten

Wir wollen nun eine Lösungstheorie von linearen Differentialgleichungen n-ter Ordnung mit konstanten Koeffizienten behandeln. Dazu betrachten wir zunächst Polynome von Differentialoperatoren.

Es sei

$$\mathbb{C}[T] = \{a_0 + a_1 T + \cdots + a_n T^n \mid a_0, \ldots, a_n \in \mathbb{C}\}$$

die Menge aller Polynome mit komplexen Koeffizienten. Betrachte ein Polynom $P(T) \in \mathbb{C}[T]$

$$P(T) = a_0 + a_1 T + \cdots + a_n T^n.$$

Wir ersetzen hierin die Unbestimmte durch $D = \frac{d}{dx}$:

$$P(D) = a_0 + a_1 D + \cdots + a_n D^n.$$

Dadurch erhält man einen *Differentialoperator*, d.h. eine Abbildung, die einer n-mal differenzierbaren Funktion

$$f \colon I \to \mathbb{C}, \qquad I \subset \mathbb{R} \text{ Intervall},$$

die neue Funktion

$$\begin{aligned} P(D)f &:= a_0 f + a_1 Df + \cdots + a_n D^n f \\ &= a_0 f + a_1 f' + \cdots + a_n f^{(n)} \end{aligned}$$

zuordnet.

Eine homogene lineare Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten

$$y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0, \qquad a_i \in \mathbb{C},$$

schreibt sich damit wie folgt:

$$P(D)y = 0,$$

wobei $P(T) \in \mathbb{C}[T]$ ein Polynom n-ten Grades mit höchstem Koeffizienten 1 ist.

Wir wollen jetzt zeigen, dass man mit Polynomen von Differentialoperatoren ganz analog rechnen kann wie mit gewöhnlichen Polynomen.

Satz 10.1 *Seien $P_1(T), P_2(T) \in \mathbb{C}[T]$ und*

$$P(T) := P_1(T) + P_2(T).$$

Dann gilt für jede genügend oft differenzierbare Funktion $f \colon I \to \mathbb{C}$

$$P(D)f = P_1(D)f + P_2(D)f.$$

Beweis. Sei

$$P_1(T) = \sum_{i=0}^{n} a_i T^i, \qquad P_2(T) = \sum_{j=0}^{m} b_j T^j.$$

O. B. d. A. $m = n$. (Falls etwa $m < n$ ergänze $b_{m+1} = \cdots = b_n = 0$.) Dann ist

$$P(T) = \sum_{i=0}^{n} (a_i + b_i) T^i.$$

Damit

$$\begin{aligned} P(D)f &= \sum_{i=0}^{n} (a_i + b_i) D^i f = \sum_{i=0}^{n} a_i D^i f + \sum_{i=0}^{n} b_i D^i f \\ &= P_1(D)f + P_2(D)f. \end{aligned}$$

□

Satz 10.2 *Seien* $P_1(T), P_2(T) \in \mathbb{C}[T]$ *und*

$$Q(T) := P_1(T)P_2(T).$$

Dann gilt für jede genügend oft differenzierbare Funktion $f\colon I \to \mathbb{C}$

$$Q(D)f = P_1(D)(P_2(D)f).$$

Beweis. Sei

$$P_1(T) = \sum_{i=0}^{n} a_i T^i, \qquad P_2(T) = \sum_{j=0}^{m} b_j T^j.$$

Dann

$$Q(T) = \sum_{k=0}^{n+m} c_k T^k \text{ mit } c_k = \sum_{i=0}^{k} a_i b_{k-i}.$$

(Dabei ist $a_i = 0$ für $i > n$ und $b_j = 0$ für $j > m$ zu setzen.) Damit

$$\begin{aligned}
Q(D)f &= \sum_{k=0}^{n+m} c_k D^k f \\
&= \sum_{k=0}^{n+m} \left(\sum_{i=0}^{k} a_i b_{k-i} \right) D^k f \\
&= \sum_{k=0}^{n+m} \left(\sum_{i+j=k} a_i b_j D^{i+j} f \right) \\
&= \sum_{k=0}^{n+m} \left(\sum_{i+j=k} a_i D^i (b_j D^j f) \right) \\
&= \sum_{i=0}^{n} a_i D^i \left(\sum_{j=0}^{m} b_j D^j f \right) \\
&= P_1(D)(P_2(D)f).
\end{aligned}$$

□

Wir beschäftigen uns nun mit der Wirkung von Differentialoperatoren $P(D)$ auf Funktionen der speziellen Gestalt $f(x) = e^{\lambda x}$.

Lemma 10.1 *Für* $\lambda \in \mathbb{C}$ *gilt*

$$De^{\lambda x} = \lambda e^{\lambda x}.$$

Beweis. Es sei

$$\lambda = \mu + i\omega, \qquad \mu, \omega \in \mathbb{R}.$$

Dann ist

$$e^{\lambda x} = e^{\mu x}(\cos \omega x + i \sin \omega x).$$

Also

$$\begin{aligned}
De^{\lambda x} &= D(e^{\lambda x}\cos\omega x) + iD(e^{\mu x}\sin\omega x)\\
&= \mu e^{\mu x}\cos\omega x - \omega e^{\mu x}\sin\omega x + i(\mu e^{\mu x} \in \omega x + \omega e^{\mu x}\cos\omega x)\\
&= (\mu + i\omega)e^{\mu x}(\cos\omega x + i\sin\omega x)\\
&= \lambda e^{\lambda x}.
\end{aligned}$$

□

Lemma 10.2 *Für jedes Polynom $P(T) \in \mathbb{C}[T]$ und jedes $\lambda \in \mathbb{C}$ gilt*

$$P(D)e^{\lambda x} = P(\lambda)e^{\lambda x}.$$

Beweis. Es sei

$$P(T) = \sum_{i=0}^{n} a_i T^i.$$

Dann gilt

$$P(D)e^{\lambda x} = \sum_{i=0}^{n} a_i D^i e^{\lambda x} = \sum_{i=0}^{n} a_i \lambda^i e^{\lambda x} = P(\lambda)e^{\lambda x}.$$

□

Bemerkung 10.1 Insbesondere folgt aus Lemma 10.2: Ist λ eine Nullstelle von $P(T)$, d. h. $P(\lambda) = 0$, so ist die Funktion $\varphi(x) = e^{\lambda x}$ eine Lösung der Differentialgleichung $P(D)y = 0$.

Satz 10.3 *Es sei*

$$P(T) = T^n + a_{n-1}T^{n-1} + \cdots + a_1 T + a_0 \in \mathbb{C}[T].$$

Das Polynom $P(T)$ habe n paarweise voneinander verschiedene Nullstellen

$$\lambda_1, \ldots, \lambda_n \in \mathbb{C}.$$

Dann bilden die Funktionen $\varphi_k \colon \mathbb{R} \to \mathbb{C}$,

$$\varphi_k(x) := e^{\lambda_k x}, \qquad k = 1, \ldots, n$$

ein Fundamentalsystem von Lösungen der Differentialgleichung

$$P(D)y = y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_0 y = 0.$$

Beweis. Daß die Funktionen φ_k Lösungen der Differentialgleichung sind, folgt aus Lemma 10.2.

Zur linearen Unabhängigkeit der φ_k: Wir berechnen die Wronski-Determinante W von $\varphi_1, \ldots, \varphi_n$. Wegen

$$\varphi_k^{(\nu)}(x) = \lambda_k^{\nu} e^{\lambda_k x}$$

gilt

$$\begin{aligned} W(0) &= \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix} \\ & \quad \text{(Vandermondesche Determinante)} \\ &= \prod_{i>j} (\lambda_i - \lambda_j) \neq 0, \end{aligned}$$

da $\lambda_1, \ldots, \lambda_n$ paarweise voneinander verschieden sind. Nach Satz 9.7 sind die Lösungen $\varphi_1, \ldots, \varphi_n$ daher linear unabhängig. □

Ein Polynom n-ten Grades

$$P(T) = a_0 + a_1 T + \cdots + a_{n-1} T^{n-1} + T^n \in \mathbb{C}[T]$$

lässt sich stets folgendermaßen in Linearfaktoren zerlegen:

$$P(T) = (T - \lambda_1)^{k_1} (T - \lambda_2)^{k_2} \cdot \ldots \cdot (T - \lambda_r)^{k_r}$$

mit paarweise verschiedenen $\lambda_j \in \mathbb{C}$ und $k_j \in \mathbb{N}, k_j \geq 1$. Wir nennen die Zahl k_j die *Vielfachheit* der Nullstelle λ_j.

Es gilt $\sum k_j = n$. Falls mindestens ein $k_j \geq 2$, erhält man mit den bisherigen Methoden weniger als n linear unabhängige Lösungen der Differentialgleichung $P(D)y = 0$. Um die noch fehlende Lösungen zu erhalten, brauchen wir noch einige Vorbereitungen.

Lemma 10.3 *Es sei $\lambda \in \mathbb{C}$ und $k \in \mathbb{N}$. Dann gilt für jede auf einem Intervall $I \subset \mathbb{R}$ k-mal differenzierbare Funktion $f\colon I \to \mathbb{C}$*

$$(D - \lambda)^k (f(x) e^{\lambda x}) = f^{(k)}(x) e^{\lambda x}.$$

Beweis. (durch vollständige Induktion nach k)

Induktionsanfang: $k = 0$ trivial.

Induktionsschritt $(k-1) \to k$*:*

$$\begin{aligned} (D - \lambda)^k (f(x) e^{\lambda x}) &= (D - \lambda)(D - \lambda)^{k-1} (f(x) e^{\lambda x}) \\ &= (D - \lambda)(f^{(k-1)}(x) e^{\lambda x}) \\ &= D(f^{(k-1)}(x) e^{\lambda x}) - \lambda f^{(k-1)}(x) e^{\lambda x} \\ &= f^{(k)}(x) e^{\lambda x} + \lambda f^{(k-1)}(x) e^{\lambda x} - \lambda f^{(k-1)}(x) e^{\lambda x} \\ &= f^{(k)}(x) e^{\lambda x}. \end{aligned}$$

□

Satz 10.4 *Das Polynom*

$$P(T) = T^n + a_{n-1} T^{n-1} + \cdots + a_1 T + a_0 \in \mathbb{C}[T]$$

habe die paarweise voneinander verschiedenen Nullstellen $\lambda_j \in \mathbb{C}$ mit den Vielfachheiten $k_j, 1 \leq j \leq r$. Dann bilden die Funktionen $\varphi_{jm} \colon \mathbb{R} \to \mathbb{C}$,

$$\varphi_{jm}(x) := x^m e^{\lambda_j x}, \quad 1 \leq j \leq r, \quad 0 \leq m \leq k_j - 1$$

ein Fundamentalsystem von Lösungen der Differentialgleichung

$$P(D)y = 0.$$

Beweis. a) Die Funktion φ_{jm} sind Lösungen der Differentialgleichung: Es gilt

$$P(T) = Q_j(T)(T - \lambda_j)^{k_j}, \quad Q_j(T) \in \mathbb{C}[T].$$

Also

$$\begin{aligned} P(D)\varphi_{jm}(x) &= Q_j(D)(D - \lambda_j)^{k_j}(x^m e^{\lambda_j x}) \\ &= Q_j(D)(D^{k_j} x^m) e^{\lambda_j x} = 0, \end{aligned}$$

da $k_j > m$.

b) Es ist noch zu zeigen, dass die Funktionen φ_{jm} linear unabhängig sind.

Eine Linearkombination der φ_{jm} hat die Gestalt

$$\sum_{j=1}^{r} g_j(x) e^{\lambda_j x},$$

wobei die g_j Polynome vom Grad $\leq k_j - 1$ sind.

Behauptung Diese Linearkombination stellt nur dann die Nullfunktion dar, wenn alle $g_j \equiv 0$ sind.

Beweis. (durch Induktion nach r)

Induktionsanfang $r = 1$*:* $g_1(x)e^{\lambda_1 x} = 0$ für alle $x \in \mathbb{R} \Rightarrow g_1 \equiv 0$.

Induktionsschritt $r \to r + 1$*:* Es sei für alle $x \in \mathbb{R}$

$$\sum_{j=1}^{r} g_j(x) e^{\lambda_j x} + g_{r+1}(x) e^{\lambda_{r+1} x} = 0.$$

Multiplikation mit $e^{-\lambda_{r+1} x}$ ergibt

$$\sum_{j=1}^{r} g_j(x) e^{\mu_j x} + g_{r+1}(x) = 0$$

für alle $x \in \mathbb{R}$, mit $\mu_j = \lambda_j - \lambda_{r+1} \neq 0$. Anwendung des Differentialoperators $D^{k_{r+1}}$ ergibt

$$\sum_{j=1}^{r} h_j(x) e^{\mu_j x} = 0 \text{ für alle } x \in \mathbb{R}.$$

Da die h_j wieder Polynome sind, folgt nach Induktionsvoraussetzung, dass die h_j identisch verschwinden.

Dies ist aber nur möglich, wenn auch die g_j identisch verschwinden, wegen

$$Dg(x)e^{\mu x} \qquad (g(x) \text{ Polynom } \not\equiv 0, \mu \neq 0)$$

$$= (g'(x) + \mu g(x))e^{\mu(x)} = h(x)e^{\mu x},$$

wobei $h(x)$ Polynom mit $\operatorname{grad} h(x) = \operatorname{grad} g(x)$, also $h(x)$ auch nicht das Nullpolynom ist. □

□

Beispiel 10.1 Die Differentialgleichung der *gedämpften Schwingung* lautet:

$$y'' + 2\mu y' + \omega_0^2 y = 0, \quad \omega_0 \in \mathbb{R}_+^*, \quad \mu \in \mathbb{R}_+,$$

$$2\mu \text{ Dämpfungsfaktor.}$$

$\mu = 0$:

$$y'' + \omega_0^2 y = 0 \Leftrightarrow P(D)y = 0, \quad P(D) = D^2 + \omega_0^2.$$

Nullstellen von $P(T)$:

$$P(\lambda) = 0 \Leftrightarrow \lambda^2 + \omega_0^2 = 0 \Leftrightarrow \lambda = \pm i\omega_0.$$

Lösungsfundamentalsystem:

$$\varphi_1(x) := e^{i\omega_0 x}, \qquad \varphi_2(x) := e^{-i\omega_0 x}.$$

Reelles Fundamentalsystem:

$$\begin{aligned} \psi_1(x) &:= \frac{1}{2}(\varphi_1(x) + \varphi_2(x)) = \cos\omega_0 x, \\ \psi_2(x) &:= \frac{1}{2i}(\varphi_1(x) - \varphi_2(x)) = \sin\omega_0 x. \end{aligned}$$

Hierbei ist ω_0 die Kreisfrequenz der ungedämpften Schwingung.

$\mu > 0$:

$$y'' + 2\mu y' + \omega_0^2 y = 0 \Leftrightarrow P(D)y = 0, \quad P(D) = D^2 + 2\mu D + \omega_0^2.$$

Nullstellen von $P(T)$:

$$P(\lambda) = \lambda^2 + 2\mu\lambda + \omega_0^2 = 0 \Leftrightarrow \lambda = -\mu \pm \sqrt{\mu^2 - \omega_0^2}.$$

1. Fall: $0 < \mu < \omega_0$:

$$\lambda_{1,2} = -\mu \pm i\sqrt{\omega_0^2 - \mu^2} = -\mu \pm i\omega \text{ mit } \omega := \sqrt{\omega_0^2 - \mu^2}.$$

Lösungs-Fundamentalsystem:

$$\varphi_1(x) = e^{-\mu x}e^{i\omega x}, \qquad \varphi_2(x) = e^{-\mu x}e^{-i\omega x}.$$

Reelles Fundamentalsystem:

$$\psi_1(x) = e^{-\mu x}\cos\omega x, \qquad \psi_2(x) = e^{-\mu x}\sin\omega x.$$

Physikalisch: Schwingungen mit abnehmender Amplitude, Kreisfrequenz $\omega < \omega_0$ (siehe Abbildung 1).

2. Fall: $\mu = \omega_0$: $P(T)$ hat Nullstelle $\lambda = -\mu$ mit Vielfachheit 2.

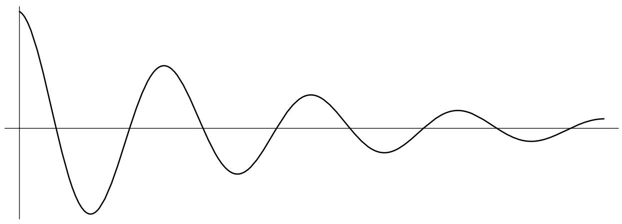

Abbildung 1: Schwache Dämpfung: $\psi_1(x) = e^{-(1/2)x} \cos 5x$

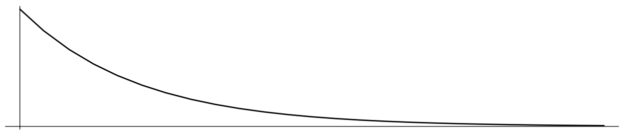

Abbildung 2: Starke Dämpfung: $\varphi_1(x) = e^{-x}$

Lösungsfundamentalsystem:

$$\varphi_1(x) = e^{-\mu x}, \qquad \varphi_2(x) = xe^{-\mu x}$$

3. Fall: $\mu > \omega_0$

$$\lambda_{1,2} = -\mu \pm \sqrt{\mu^2 - \omega_0^2} < 0$$

Lösungsfundamentalsystem:

$$\varphi_1(x) = e^{\lambda_1 x}, \qquad \varphi_2(x) = e^{\lambda_2 x}$$

Physikalisch: keine echte Schwingung mehr (siehe Abbildung 2).

Inhaltsverzeichnis