

# Elemente der Analysis III

Wolf P. Barth

Nürnberg, SS 07, WS 07/08

Version vom 14. März 2007

Mathematisches Institut der Universität  
Bismarckstr. 1 1/2, D-91054 Erlangen

## Inhaltsverzeichnis

<b>5</b>	<b>Der Zahlenraum <math>\mathbb{R}^n</math></b>	<b>2</b>
5.1	Funktionen und Abbildungen . . . . .	2
5.2	Abstand und Topologie . . . . .	6
5.3	Konvergenz, kompakte Mengen . . . . .	12
5.4	Stetigkeit . . . . .	16
<b>6</b>	<b>Differentialrechnung mit mehreren Veränderlichen</b>	<b>22</b>
6.1	Partielle Ableitungen . . . . .	22
6.2	Höhere partielle Ableitungen, Taylor-Formel . . . . .	31
6.3	Lokale Extrema . . . . .	36
6.4	Die Kettenregel . . . . .	44

## 5 Der Zahlenraum $\mathbb{R}^n$

Der Zahlenraum  $\mathbb{R}^n$  besteht aus allen  $n$ -tupeln reeller Zahlen. So ein  $n$ -tupel nennt man auch Vektor. Wir wollen Vektoren immer als Spaltenvektoren

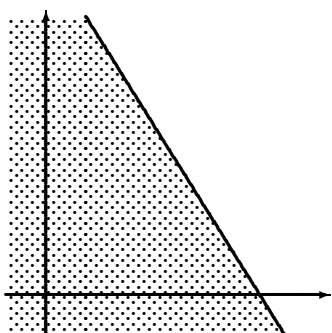
$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

schreiben, d.h., immer, wenn nichts dagegenspricht.

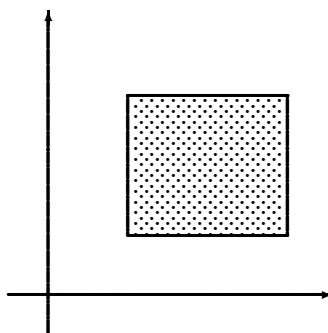
Die algebraischen Eigenschaften dieses Zahlenraums untersucht man in der Linearen Algebra. Wir beschäftigen uns mit seinen analytischen Eigenschaften. Die Analysis im  $\mathbb{R}^n$  ist vor allem deswegen schwieriger, als die eindimensionale Analysis aus den ersten beiden Semestern, weil es im  $\mathbb{R}^n$  sehr viel mehr verschiedenartige Teilmengen gibt, als in der reellen Geraden. Die einfachsten davon, die immer wieder vorkommen werden, sind - außer den linearen und affinen Unterräumen - z.B.

Halbräume	$\{\mathbf{x} \in \mathbb{R}^n : a_1 x_1 + a_2 x_2 + \dots + a_n x_n \leq c\}$	$a_1, a_2, \dots, a_n, c \in \mathbb{R}$
Quader	$\{\mathbf{x} \in \mathbb{R}^n : a_1 \leq x_1 \leq b_1, \dots, a_n \leq x_n \leq b_n\}$	$a_1 < b_1, \dots, a_n < b_n \in \mathbb{R}$
Kugeln	$\{\mathbf{x} \in \mathbb{R}^n : (x_1 - a_1)^2 + \dots + (x_n - a_n)^2 \leq r^2\}$	$a_1, a_2, \dots, a_n, r \in \mathbb{R}$

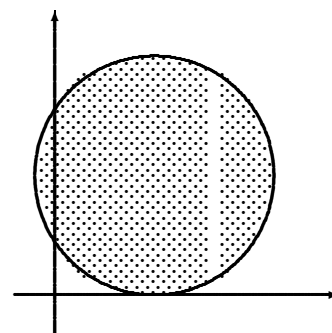
Ihre zweidimensionalen Versionen sehen so aus:



Halbebene



Rechteck



Kreisscheibe

### 5.1 Funktionen und Abbildungen

Wenn man mehr als eine Dimension zur Verfügung hat, kann man Abbildungen betrachten, deren Definitionsmenge etwa im  $\mathbb{R}^m$  liegt, und deren Bildmenge zum  $\mathbb{R}^n$  gehört:

$$F : U \rightarrow V, \quad U \subset \mathbb{R}^m, V \subset \mathbb{R}^n.$$

Solche Abbildungen  $F$  kann man sich natürlich jetzt nur viel schwerer vorstellen, als die Funktionen der eindimensionalen Analysis. Wir wollen uns ihnen deswegen schrittweise nähern.

**Der Fall  $m = 1$ : Kurven.** Eine Abbildung eines Intervalls  $[a, b] \subset \mathbb{R}$  in den  $\mathbb{R}^n$  heißt *Kurve*. Man schreibt so eine Kurve wie folgt:

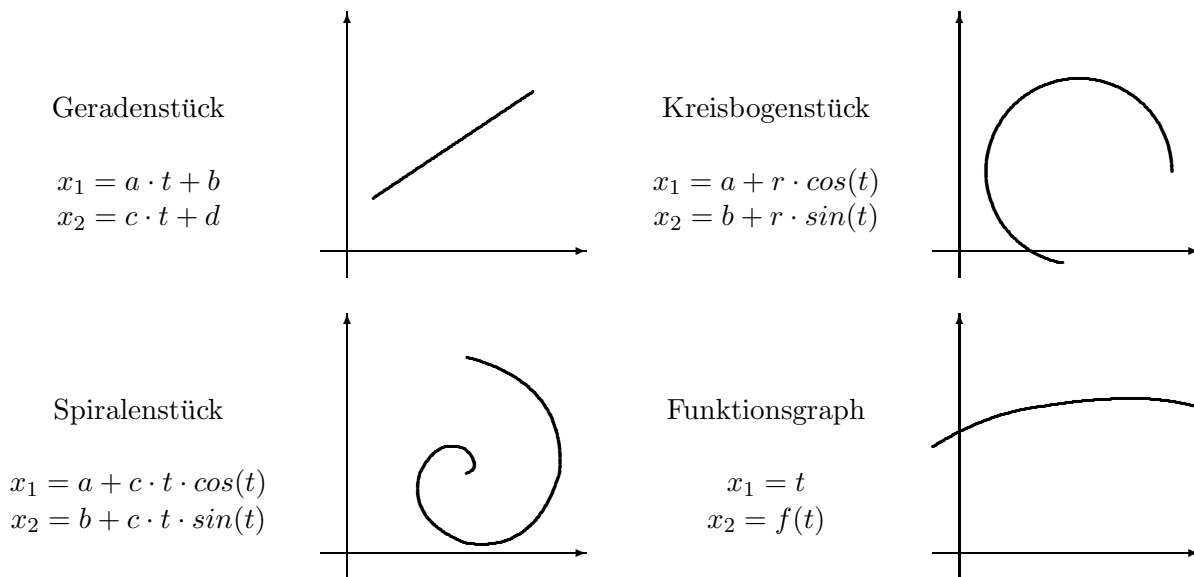
$$\phi : [a, b] \ni t \mapsto \phi(t) = \mathbf{x}(t) \in \mathbb{R}^n$$

Am besten stellt man sich  $t$  als einen Parameter (etwa die Zeit) vor, von dem der Bildpunkt abhängt. Dieser Bildvektor

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

ist also ein  $n$ -tupel von ganz normalen Funktionen  $x_\nu(t)$ ,  $\nu = 1, \dots, n$ .

Für  $n = 2$  oder  $n = 3$  kann man sich solche Kurven noch recht gut vorstellen. Zeichnen kann man sie eigentlich vernünftig nur für  $n = 2$  (ebene Kurven). Hier einige Beispiele:



Hier ordnen sich also auch die Graphen aller in der eindimensionalen Analysis betrachteten Funktionen ein.

Raumkurven möchte ich hier nicht mehr zeichnen, aber eine besonders schöne ist die Helix (Wendeltreppe)

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} r \cdot \cos(t) \\ r \cdot \sin(t) \\ c \cdot t \end{pmatrix}.$$

Praktisch kommen Raumkurven vor als Bahnkurven von Teilchen, solange man sich diese Teilchen punktförmig denkt: Asteroiden im Weltraum, Kanonenkugeln in der Erdatmosphäre, Elementarteilchen im Teilchenbeschleuniger, Schmetterlinge über einer Wiese, ... .

**Der Fall  $n = 1$ : Funktionen.** Eine Abbildung

$$f : \mathbb{R}^n \ni \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mapsto f(x_1, \dots, x_n) \in \mathbb{R}$$

bezeichnet man üblicherweise als *Funktion*, auch wenn sie von  $m > 1$  Variablen  $x_\mu$  abhängt. Für  $m = 1$  erhalten wir nichts neues.

Für  $m = 2$  bekommen wir Funktionen  $f(x_1, x_2)$  von zwei Variablen. Deren Graphen

$$y = f(x_1, x_2)$$

kann man sich immer noch vorstellen, als eine Fläche, die über der  $(x_1, x_2)$ -Ebene im Raum liegt. Zeichnen möchte ich die jetzt nicht mehr, das ist mir zu aufwendig. Aber hier einige Beispiele:

affine Ebene	$f(x_1, x_2) = a \cdot x_1 + b \cdot x_2 + c$
Halb-Sphäre	$f(x_1, x_2) = \sqrt{r^2 - (x_1^2 + x_2^2)}$
Paraboloid	$f(x_1, x_2) = x_1^2 + x_2^2$
Sattelfläche	$f(x_1, x_2) = x_1^2 - x_2^2$

Hierher gehören auch Verknüpfungen, wie sie aus der Algebra bekannt sind

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1 + x_2, \quad x_1 - x_2, \quad x_1 \cdot x_2, \quad x_1/x_2,$$

oder aus der Analysis

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^{x_2}, \quad \ln^{x_1}(x_2).$$

**Der Fall  $m = n$ :** Dieser Fall kommt sehr häufig vor, unter dem Namen *Transformation*. Eine Transformation ist eine bijektive Abbildung

$$F : U \rightarrow V, \quad U, V \subset \mathbb{R}^n.$$

Hierher gehören die Vektorraum-Isomorphismen

$$\mathbf{x} \mapsto A \cdot \mathbf{x}, \quad A \in GL(n, \mathbb{R}),$$

aus der linearen Algebra, oder die Affinitäten

$$\mathbf{x} \mapsto \mathbf{a} + A \cdot \mathbf{x}.$$

Außer diesen Transformationen ist für uns die wichtigste die Transformation von euklidischen Koordinaten in Polarkoordinaten  $r, \varphi$

$$\begin{pmatrix} r \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} r \cdot \cos(\varphi) \\ r \cdot \sin(\varphi) \end{pmatrix}.$$

Diese Transformation ist meistens günstig bei rotations-symmetrischen Situationen in der Ebene. Es gibt auch eine ähnliche Transformation für rotations-symmetrische Situationen im Raum, die Transformation in Kugelkoordinaten  $r, \varphi, \theta$

$$\begin{pmatrix} r \\ \varphi \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} r \cdot \cos(\varphi) \cdot \cos(\theta) \\ r \cdot \sin(\varphi) \cdot \cos(\theta) \\ r \cdot \sin(\theta) \end{pmatrix}.$$

Natürlich gibt es auch Abbildungen  $U \rightarrow V$ , mit  $U, V \in \mathbb{R}^n$ , die nicht bijektiv sind. Ziemlich häufig kommt zum Beispiel die Projektion des Raums in die  $(x_1, x_2)$ -Ebene vor

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}.$$

Abbildungen  $F : U \rightarrow V$  mit  $U, V \subset \mathbb{R}^2$  oder  $\mathbb{R}^3$  kommen auch als *Vektorfelder* vor: In jedem Punkt  $\mathbf{x} \in U$  stellt man sich den Bildpunkt (= Bildvektor)  $F(\mathbf{x}) \in V$  als Vektor angeheftet vor. Solche Vektorfelder braucht man zur Beschreibung des Feldes von Geschwindigkeitsvektoren einer Flüssigkeitsströmung, oder für elektrische oder magnetische Felder.

**Der Allgemeinfall:** Sind  $m$  und  $n > 0$ , so kann man sich eine Abbildung  $U \rightarrow V$ ,  $U \subset \mathbb{R}^m, V \subset \mathbb{R}^n$  meist nicht mehr vorstellen. Vorkommen tun sie trotzdem. So deutet man etwa den Fall  $m = 2, n = 3$  als Parametrisierung einer Fläche (in Analogie zu  $m = 1, n = 3$ , der Parametrisierung einer Kurve,) im Raum. Für uns wichtig sind aber nur die oben aufgeführten Fälle.

Noch etwas zur Schreibweise: Eine Abbildung  $F : U \rightarrow V$  mit  $U \subset \mathbb{R}^m$  und  $V \subset \mathbb{R}^n$  schreibe ich immer als

$$F : U \rightarrow V, \quad F : \mathbf{x} \mapsto F(\mathbf{x}) = \begin{pmatrix} F_1(\mathbf{x}) \\ \vdots \\ F_n(\mathbf{x}) \end{pmatrix}.$$

Dabei ist  $F_\nu(\mathbf{x})$  die  $\nu$ -te Komponente des Bildvektors  $F(\mathbf{x}) \in V \subset \mathbb{R}^n$ . Diese  $\nu$ -te Komponente ist eine Funktion  $F_\nu : U \rightarrow \mathbb{R}$ , also eine Funktion  $F_\nu(\mathbf{x}) = F_\nu(x_1, \dots, x_m)$  von  $m$  Variablen. Die Abbildung  $F$  ist genau dasselbe wie die Kollektion ihrer  $n$  Komponentenfunktionen  $F_\nu$ ,

$$F = \begin{pmatrix} F_1 \\ \vdots \\ F_n \end{pmatrix}.$$

**Beispiel 5.1** Die Komponentenfunktionen der Polarkoordinatentransformation sind

$$F_1(r, \varphi) = r \cdot \cos(\varphi), \quad F_2(r, \varphi) = r \cdot \sin(\varphi).$$

Auf der Halbebene  $x_1 > 0$  z.B. ist sie umkehrbar durch

$$r = \sqrt{x_1^2 + x_2^2}, \quad \varphi = \operatorname{arctg}\left(\frac{x_2}{x_1}\right).$$

Die Komponentenfunktionen dieser Umkehrabbildung  $G$  sind

$$G_1(x_1, x_2) = \sqrt{x_1^2 + x_2^2}, \quad G_2(x_1, x_2) = \operatorname{arctg}\left(\frac{x_2}{x_1}\right).$$

## 5.2 Abstand und Topologie

Ausgangspunkt zur Abstandsberechnung im  $\mathbb{R}^n$  ist das euklidische Skalarprodukt

$$(\mathbf{x} \cdot \mathbf{y}) = \sum_{\nu=1}^n x_{\nu} y_{\nu}$$

aus der Linearen Algebra. Seine *drei wichtigen Eigenschaften* sind wohlbekannt. Stellen wir sie hier noch einmal zusammen:

- *Bilinearität*: Das Skalarprodukt  $(\mathbf{x} \cdot \mathbf{y})$  ist linear in  $\mathbf{x}$  und  $\mathbf{y}$ . (Das brauchen wir nicht noch explizit in Formeln hinzuschreiben.)
- *Symmetrie*:  $(\mathbf{x} \cdot \mathbf{y}) = (\mathbf{y} \cdot \mathbf{x})$ .
- *Positiv-Definitheit*: Es ist stets  $(\mathbf{x} \cdot \mathbf{x}) \geq 0$  und  $(\mathbf{x} \cdot \mathbf{x}) = 0$  nur wenn  $\mathbf{x} = \mathbf{0}$ .

Die Zahl

$$\|\mathbf{x}\| := \sqrt{(\mathbf{x} \cdot \mathbf{x})} = \sqrt{\sum_{\nu=1}^n (x_{\nu})^2}$$

heißt *Länge* oder *Norm* des Vektors  $\mathbf{x} \in \mathbb{R}^n$ . Auch diese Norm-Funktion hat *drei wichtige Eigenschaften*:

- Aus der Positiv-Definitheit folgt  $\|\mathbf{x}\| \geq 0$  für alle  $\mathbf{x} \in \mathbb{R}^n$ , und  $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ .
- Aus der Bilinearität folgt

$$\|c \cdot \mathbf{x}\| = |c| \cdot \|\mathbf{x}\|$$

für alle  $c \in \mathbb{R}$  und  $\mathbf{x} \in \mathbb{R}^n$ .

- Es gilt die Cauchy-Schwarz-Ungleichung

$$|(\mathbf{x} \cdot \mathbf{y})| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\| \quad \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

Beweis der Cauchy-Schwarz-Ungleichung: Wenn  $\mathbf{y} = \mathbf{0}$  ist, steht auf beiden Seiten der Ungleichung 0, dann gilt sie also trivialerweise, wie man sagt. Wenn  $\mathbf{y} \neq \mathbf{0}$  ist, dann projizieren wir den Vektor  $\mathbf{x}$  orthogonal in die von  $\mathbf{y}$  aufgespannte Gerade  $\mathbb{R} \cdot \mathbf{y}$  und erhalten

$$\mathbf{x}' = \frac{(\mathbf{x} \cdot \mathbf{y})}{(\mathbf{y} \cdot \mathbf{y})} \mathbf{y}.$$

(Dies ist tatsächlich die Orthogonalprojektion des Vektors  $\mathbf{x}$  in die Gerade  $\mathbb{R} \cdot \mathbf{y}$ , denn man sieht sofort, dass der Differenzvektor  $\mathbf{x}' - \mathbf{x}$  auf  $\mathbf{y}$  senkrecht steht:

$$(\mathbf{x}' - \mathbf{x} \cdot \mathbf{y}) = \frac{(\mathbf{x} \cdot \mathbf{y})}{(\mathbf{y} \cdot \mathbf{y})} (\mathbf{y} \cdot \mathbf{y}) - (\mathbf{x} \cdot \mathbf{y}) = 0.$$

Aber das brauchen wir für den Beweis gar nicht.) Wir brauchen nur die Positiv-Definitheit, ausgewertet für den Vektor  $\mathbf{x}' - \mathbf{x}$ :

$$\begin{aligned}
 \|\mathbf{x}' - \mathbf{x}\|^2 &= \|\mathbf{x}'\|^2 - 2(\mathbf{x}' \cdot \mathbf{x}) + \|\mathbf{x}\|^2 \\
 &= \left(\frac{(\mathbf{x} \cdot \mathbf{y})}{(\mathbf{y} \cdot \mathbf{y})}\right)^2 \|\mathbf{y}\|^2 - 2 \cdot \frac{(\mathbf{x} \cdot \mathbf{y})}{(\mathbf{y} \cdot \mathbf{y})} \cdot (\mathbf{y} \cdot \mathbf{x}) + \|\mathbf{x}\|^2 \\
 &= -\frac{(\mathbf{x} \cdot \mathbf{y})^2}{\|\mathbf{y}\|^2} + \|\mathbf{x}\|^2 \\
 &\geq 0, \\
 \frac{(\mathbf{x} \cdot \mathbf{y})^2}{\|\mathbf{y}\|^2} &\leq \|\mathbf{x}\|^2, \\
 (\mathbf{x} \cdot \mathbf{y})^2 &\leq \|\mathbf{x}\|^2 \cdot \|\mathbf{y}\|^2.
 \end{aligned}$$

Die Cauchy-Schwarz-Ungleichung ergibt sich, wenn man aus dieser letzten Ungleichung die Wurzel zieht.  $\square$

Mit dieser Norm definiert man den Abstand (die Distanz) zweier Vektoren  $\mathbf{x}$  und  $\mathbf{y} \in \mathbb{R}^n$  als

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

Die *drei wichtigen Eigenschaften* des Abstands sind:

- Es ist stets  $d(\mathbf{x}, \mathbf{y}) \geq 0$  und  $= 0$  nur, wenn  $\mathbf{x} = \mathbf{y}$ . (Dies sieht man sofort mit der ersten wichtigen Eigenschaft der Norm.)
- Die Abstandsfunktion ist symmetrisch:  $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ . (Dies folgt aus der zweiten wichtigen Eigenschaft der Norm, wenn man dort  $c = -1$  setzt.)
- Es gilt die *Dreiecksungleichung*

$$d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) \text{ für alle } \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n.$$

Beweis der Dreiecksungleichung: Wir folgern ihre quadrierte Form aus der Cauchy-Schwarz-Ungleichung:

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

$\square$

Diese Abstandsfunktion  $d(\mathbf{x}, \mathbf{y})$  nennt man auch *Metrik*. Allgemein heißt eine Funktion  $d(\mathbf{x}, \mathbf{y})$  von zwei Elementen aus einer beliebigen Menge  $M$  eine Metrik auf  $M$ , wenn sie die drei zuletzt angeführten wichtigen Eigenschaften besitzt. Die Menge  $M$  zusammen mit einer Metrik  $d$  heißt dann ein *metrischer Raum*. Aber wir brauchen keine metrischen Räume.

Mit dieser Abstandsfunktion  $d(\mathbf{x}, \mathbf{y})$  kann man Kugeln definieren:

**Definition 5.1** Die Menge

$$\begin{aligned}\bar{K} &:= \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x} - \mathbf{a}) \leq r\} && \text{heißt abgeschlossene,} \\ K &:= \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x} - \mathbf{a}) < r\} && \text{heißt offene}\end{aligned}$$

Kugel mit Mittelpunkt  $\mathbf{a}$  und Radius  $r$ .

Den Unterschied zwischen einer abgeschlossenen und einer offenen Kugel sieht man an deren Rand: Der *Rand* der oben definierten Kugeln ist die Menge der Punkte, welche vom Mittelpunkt genau den Abstand  $r$  haben:

$$\partial(\bar{K}) = \partial(K) = \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{a}) = r\}.$$

Bei der abgeschlossenen Kugel  $\bar{K}$  gehört der Rand dazu, bei der offenen Kugel nicht. (Das ist genau derselbe Unterschied, wie bei den Intervallen auf der reellen Achse  $\mathbb{R}$ : Zum abgeschlossenen Intervall gehören die Randpunkte dazu, zum offenen Intervall nicht.) Der Radius  $r$  einer Kugel soll bei uns immer  $> 0$  sein. Für  $r = 0$  bestünde die abgeschlossene Kugel vom Radius  $r = 0$  nur aus ihrem Mittelpunkt, die offene Kugel vom Radius  $r = 0$  wäre leer.

Mit offenen Kugeln kann man offene Mengen definieren:

**Definition 5.2** Eine Teilmenge  $U \subset \mathbb{R}^n$  heißt offen, wenn zu jedem Punkt  $\mathbf{u} \in U$  eine Kugel mit Mittelpunkt  $\mathbf{u}$  und einem Radius  $r > 0$  existiert, die noch ganz zu  $U$  gehört.

**Beispiel 5.2** Sei  $c \in \mathbb{R}$  beliebig. Die Halbebene

$$H := \{\mathbf{x} \in \mathbb{R}^2 : x_1 < c\}$$

ist offen.

*Beweis.* Sei  $\mathbf{u} \in H$ , also  $u_1 < c$ . Dann gibt es einen Radius  $r > 0$  mit  $r < c - u_1$ , etwa  $r := (c - u_1)/2$ . Wir zeigen: Die Kugel  $K$  um  $\mathbf{u}$  mit diesem Radius liegt ganz in  $H$ .

In der Tat, sei  $\mathbf{x} \in K$ , also  $d(\mathbf{x}, \mathbf{u}) < r$ . Dann ist

$$\begin{aligned}x_1 - u_1 &\leq |x_1 - u_1| \\ &= \sqrt{(x_1 - u_1)^2} \\ &\leq \sqrt{(x_1 - u_1)^2 + (x_2 - u_2)^2} \\ &= d(\mathbf{x}, \mathbf{u}) \\ &< r.\end{aligned}$$

Daraus folgt

$$x_1 < u_1 + r < c$$

und  $\mathbf{x} \in H$ . □



**Beispiel 5.3** Dieselbe Halbebene, definiert mit dem  $\leq$ -Zeichen an Stelle des  $<$ -Zeichens

$$\bar{H} := \{\mathbf{x} \in \mathbb{R}^2 : x_1 \leq c\}$$

ist nicht offen.

*Beweis.* Zu  $\bar{H}$  gehören auch alle Punkte  $\mathbf{u} \in \mathbb{R}^2$  mit  $u_1 = c$ . Wir halten einen solchen Punkt  $\mathbf{u}$  fest und zeigen: Es gibt keinen Radius  $r > 0$  derart, dass die Kugel um  $\mathbf{u}$  mit diesem Radius noch ganz in  $\bar{H}$  liegt.

In der Tat, sei  $r > 0$ . Für den Punkt

$$\mathbf{x} := \begin{pmatrix} u_1 + r/2 \\ u_2 \end{pmatrix}$$

ist der Abstand zu  $\mathbf{u}$

$$d(\mathbf{x}, \mathbf{u}) = \sqrt{(x_1 - u_1)^2 + (x_2 - u_2)^2} = \frac{r}{2} < r,$$

Also gehört  $\mathbf{x}$  zur Kugel  $K$  um  $\mathbf{u}$  mit Radius  $r$ . Wegen

$$x_1 = u_1 + \frac{r}{2} = c + \frac{r}{2} > c$$

ist  $\mathbf{x} \notin \bar{H}$ . Die Kugel  $K$  liegt also nicht ganz in  $\bar{H}$ . □

**Beispiel 5.4** Offene Kugeln (im Sinn der Definition offener Kugeln) sind offen (im Sinn der Definition offener Mengen).

*Beweis.* Sei

$$K := \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{a}) < R\}$$

eine offene Kugel mit Mittelpunkt  $\mathbf{a}$  und Radius  $R > 0$ . Sei  $\mathbf{u} \in K$  ein Punkt. Wir müssen zeigen, es gibt einen Radius  $r > 0$  derart, dass die Kugel  $K_{\mathbf{u}}$  um  $\mathbf{u}$  mit diesem Radius  $r$  ganz in der Kugel  $K$  liegt.

In der Tat, weil  $\mathbf{u}$  zu  $K$  gehört ist  $d(\mathbf{u}, \mathbf{a}) < R$  und

$$r := \frac{R - d(\mathbf{u}, \mathbf{a})}{2} > 0.$$

Sei  $K_{\mathbf{u}}$  die Kugel um  $\mathbf{u}$  mit diesem Radius  $r$ . Nach der Dreiecksungleichung gilt für jeden Punkt  $\mathbf{x}$  dieser Kugel  $K_{\mathbf{u}}$

$$d(\mathbf{x}, \mathbf{a}) \leq d(\mathbf{x}, \mathbf{u}) + d(\mathbf{u}, \mathbf{a}) < \frac{R - d(\mathbf{u}, \mathbf{a})}{2} + d(\mathbf{u}, \mathbf{a}) < (R - d(\mathbf{u}, \mathbf{a})) + d(\mathbf{u}, \mathbf{a}) = R.$$

Also gehört jeder Punkt  $\mathbf{x} \in K_{\mathbf{u}}$  zu  $K$ , es gilt  $K_{\mathbf{u}} \subset K$ . □

**Beispiel 5.5** Die leere Menge  $\emptyset \subset \mathbb{R}^n$  ist offen: Mit jedem Punkt  $\mathbf{u} \in \emptyset$  gehört eine ganze Kugel um  $\mathbf{u}$  zu  $\emptyset$ . Das ist richtig, denn es gibt keinen einzigen Punkt  $\mathbf{u} \in \emptyset$ , für den wir das beweisen müssen.

**Beispiel 5.6** Sind  $U, V \subset \mathbb{R}^n$  offen, so ist auch ihre Vereinigung  $U \cup V \subset \mathbb{R}^n$  offen.

*Beweis.* Sei  $\mathbf{u} \in U \cup V$ , etwa  $\mathbf{u} \in U$ . Weil  $U$  offen ist, gibt es ein Kugel  $K$  mit Zentrum  $\mathbf{u}$ , die ganz in  $U$  liegt. Diese Kugel  $K$  gehört dann auch zu  $U \cup V$ .  $\square$

Genauso sieht man, dass die Vereinigung beliebig vieler (auch unendlich vieler) offener Mengen wieder offen ist.

**Beispiel 5.7** Sind  $U, V \subset \mathbb{R}^n$  offen, so ist auch ihr Durchschnitt  $U \cap V \subset \mathbb{R}^n$  wieder offen.

*Beweis.* Sei  $\mathbf{u} \in U \cap V$ . Weil  $U$  offen ist, gibt es eine Kugel  $K_U$  mit Zentrum  $\mathbf{u}$ , die ganz in  $U$  liegt. Und weil  $V$  offen ist, gibt es auch eine Kugel  $K_V$  um  $\mathbf{u}$ , die ganz in  $V$  liegt. Sei  $K$  diejenige der beiden Kugeln  $K_U$  und  $K_V$ , welche den kleineren Radius besitzt. Dann ist also  $K \subset K_U \subset U$  und  $K \subset K_V \subset V$ , also  $K \subset U \cap V$ .  $\square$

Genauso sieht man, dass der Durchschnitt beliebig (aber endlich) vieler offener Mengen wieder offen ist.

**Definition 5.3** Eine Teilmenge  $A \subset \mathbb{R}^n$  heißt abgeschlossen, wenn ihr Komplement  $U := \mathbb{R}^n \setminus A$  offen ist.

**Beispiel 5.8** Sei  $c \in \mathbb{R}$  beliebig. Die Halbebene

$$\bar{H} := \{\mathbf{x} \in \mathbb{R}^2 : x_1 \leq c\}$$

ist abgeschlossen.

*Beweis.* Ihr Komplement

$$\mathbb{R}^2 \setminus \bar{H} = \{\mathbf{x} \in \mathbb{R}^2 : x_1 > c\}$$

ist eine offene Halbebene (vgl. Beispiel 5.2) bei den offenen Mengen).  $\square$

**Beispiel 5.9** Abgeschlossene Kugeln (im Sinn der Definition abgeschlossener Kugeln) sind abgeschlossen (im Sinn der Definition abgeschlossener Mengen).

*Beweis.* Sei etwa

$$\bar{K} := \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{a}) \leq R\}$$

eine abgeschlossene Kugel. Wir müssen zeigen, ihr Komplement

$$U = \mathbb{R}^n \setminus \bar{K} = \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{a}) > R\}$$

ist offen. Sei dazu  $\mathbf{u} \in U$  herausgegriffen. Dann ist also  $d(\mathbf{u}, \mathbf{a}) > R$ . Wir wählen

$$r := \frac{1}{2}(d(\mathbf{u}, \mathbf{a}) - R)$$

und nehmen als  $K_{\mathbf{u}}$  die Kugel um  $\mathbf{u}$  mit diesem Radius  $r$ . Wir zeigen  $K_{\mathbf{u}} \subset U$ .

Sei dazu  $\mathbf{x} \in K_{\mathbf{u}}$ , also

$$d(\mathbf{x}, \mathbf{u}) < r = \frac{1}{2}(d(\mathbf{u}, \mathbf{a}) - R).$$

Wenn  $\mathbf{x}$  nicht in  $U$ , sondern in  $\bar{K}$  liegen würde, dann würde  $d(\mathbf{x}, \mathbf{a}) \leq R$  gelten, und aus der Dreiecksungleichung würde folgen

$$d(\mathbf{u}, \mathbf{a}) \leq d(\mathbf{u}, \mathbf{x}) + d(\mathbf{x}, \mathbf{a}) < r + R = \frac{1}{2}(d(\mathbf{u}, \mathbf{a}) - R) + R < (d(\mathbf{u}, \mathbf{a}) - R) + R = d(\mathbf{u}, \mathbf{a}),$$

Widerspruch! □

Schließlich noch eine

**Definition 5.4** Sei  $M \subset \mathbb{R}^n$  eine Teilmenge. Ein Punkt  $\mathbf{x} \in \mathbb{R}^n$  heißt Randpunkt von  $M$ , wenn jede offene Kugel  $K \subset \mathbb{R}^n$  mit Mittelpunkt  $\mathbf{x}$  sowohl Punkte aus  $M$  als auch Punkte aus  $\mathbb{R}^n \setminus M$  enthält. Die Menge aller Randpunkte von  $M$  nennt man den Rand  $\partial M$  von  $M$ .

Ist beispielsweise  $M$  offen, so ist kein Punkt  $\mathbf{x} \in M$  ein Randpunkt, denn eine ganze Kugel  $K$  um  $\mathbf{x}$  gehört ja zu  $M$ , enthält deswegen keinen Punkt aus  $\mathbb{R}^n \setminus M$ .

Ist beispielsweise  $M$  abgeschlossen, so gilt  $\partial M \subset M$ . Denn  $\mathbb{R}^n \setminus M$  ist offen, und nach Definition ist  $\partial M = \partial(\mathbb{R}^n \setminus M)$ , und deswegen ist  $\partial M \cap (\mathbb{R}^n \setminus M) = \emptyset$ .

Noch ein ganz konkretes Beispiel:

**Beispiel 5.10** Sei  $B$  die offene Einheitskugel  $\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| < 1\}$ . Ihr Rand ist die Einheitssphäre  $\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}$ .

**Aufgabe 5.1** Zeigen Sie:

- a)  $M \subset \mathbb{R}^n$  ist offen  $\Leftrightarrow M \cap \partial M = \emptyset$ .
- b)  $M \subset \mathbb{R}^n$  ist abgeschlossen  $\Leftrightarrow \partial M \subset M$ .
- c)  $\partial M$  ist abgeschlossen.

**Aufgabe 5.2** Sei  $M$  eine offene Teilmenge des  $\mathbb{R}^2$ . Beweisen Sie, dass der Rand  $\partial M$  von  $M$  keine nichtleere offene Teilmenge des  $\mathbb{R}^2$  enthält.

**Aufgabe 5.3** Sei  $A$  eine abgeschlossene Teilmenge des  $\mathbb{R}^n$  mit  $\mathbf{0} \notin A$ . Zeigen Sie, dass es eine positive reelle Zahl  $\epsilon$  gibt mit

$$\|\mathbf{a}\| \geq \epsilon \text{ für jedes } \mathbf{a} \in A.$$

**Aufgabe 5.4** a) Skizzieren Sie  $G := G_1 \cup G_2 \cup G_3 \setminus G_4$ , wobei

$$\begin{aligned} G_1 &= \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1 \text{ und } x \leq 0\}, \\ G_2 &= \{(x, y) \in \mathbb{R}^2 : x^2 + \left(y - \frac{3}{4}\right)^2 \leq \left(\frac{1}{4}\right)^2 \text{ und } x \geq 0\}, \\ G_3 &= \{(x, y) \in \mathbb{R}^2 : x^2 + \left(y + \frac{3}{4}\right)^2 \leq \left(\frac{1}{4}\right)^2 \text{ und } x \geq 0\}, \\ G_4 &= \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < \frac{1}{4} \text{ und } x \leq 0\}. \end{aligned}$$

b) Ist  $G$  abgeschlossen, offen?

**Aufgabe 5.5** Zeigen Sie durch ein Beispiel, dass die Vereinigung unendlich vieler abgeschlossener Teilmengen von  $\mathbb{R}$  nicht abgeschlossen sein muss.

**Aufgabe 5.6** Es sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  eine beschränkte Funktion, deren Graph in  $\mathbb{R}^2$  abgeschlossen ist. Beweisen Sie, dass für jede gegen  $0 \in \mathbb{R}$  konvergente Folge  $(a_n)_{n \geq 1}$  reeller Zahlen die Folge  $(f(a_n))_{n \geq 1}$  eine gegen  $f(0)$  konvergente Teilfolge besitzt.

**Aufgabe 5.7** Es sei  $p$  ein Punkt und  $M$  eine offene Teilmenge des  $\mathbb{R}^2$ . Beweisen Sie, dass auch

$$\{p + x \mid x \in M\}$$

eine offene Teilmenge des  $\mathbb{R}^2$  ist.

**Aufgabe 5.8** Beweisen Sie, dass  $(0, 0)$  ein Randpunkt der Teilmenge

$$\left\{ \left( x, \sin\left(\frac{1}{x}\right) \right) \mid x \in \mathbb{R}, x > 0 \right\}$$

des  $\mathbb{R}^2$  ist.

### 5.3 Konvergenz, kompakte Mengen

In diesem Abschnitt behandeln wir konvergente Folgen von Punkten im  $\mathbb{R}^n$ . Als erstes habe ich dabei ein Problem mit der Notation. In der eindimensionalen Analysis haben wir Folgen von Punkten immer  $(x_\nu)_{\nu \in \mathbb{N}}$  geschrieben, oder so ähnlich. Hier ist aber  $x_\nu$  die Bezeichnung für die  $\nu$ -te Komponente des Vektors  $\mathbf{x}$ . Deswegen werden wir Folgen von Vektoren als  $(\mathbf{x}^{(k)})_{k \in \mathbb{N}}$  schreiben. Der  $k$ -te Vektor dieser Folge ist also

$$\mathbf{x}^{(k)} = \begin{pmatrix} x_1^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix} \in \mathbb{R}^n.$$

**Definition 5.5** Eine Folge  $(\mathbf{x}^{(k)})_{k \in \mathbb{N}}$  von Vektoren  $\mathbf{x}^{(k)} \in \mathbb{R}^n$  konvergiert gegen den Vektor  $\mathbf{a} \in \mathbb{R}^n$ , wenn die Folge der Abstände  $d(\mathbf{x}^{(k)}, \mathbf{a})$  eine Nullfolge ist, d.h., wenn

$$\lim_{k \rightarrow \infty} d(\mathbf{x}^{(k)}, \mathbf{a}) = 0.$$

Konvergiert die Folge  $(\mathbf{x}^{(k)})$  gegen  $\mathbf{a} \in \mathbb{R}^n$ , so ist dieser Vektor  $\mathbf{a}$  durch die Folge eindeutig bestimmt:

Beweis: Sei  $\mathbf{b} \in \mathbb{R}^n$  noch ein Vektor, gegen den dieselbe Folge  $\mathbf{x}^{(k)}$  konvergiert. Für alle  $\epsilon > 0$  gibt es also ein  $N \in \mathbb{N}$  mit

$$d(\mathbf{x}^{(k)}, \mathbf{a}) < \epsilon \text{ und } d(\mathbf{x}^{(k)}, \mathbf{b}) < \epsilon \text{ falls } k > N.$$

Mit der Dreiecksungleichung folgt daraus

$$d(\mathbf{a}, \mathbf{b}) \leq d(\mathbf{a}, \mathbf{x}^{(k)}) + d(\mathbf{x}^{(k)}, \mathbf{b}) < 2\epsilon$$

für alle  $\epsilon > 0$ . Das geht nur, wenn  $d(\mathbf{a}, \mathbf{b}) = 0$  ist, also wenn  $\mathbf{a} = \mathbf{b}$  ist. □

Diese Bemerkung rechtfertigt (wie im ersten Semester) die Schreibweise

$$\mathbf{a} = \lim_{k \rightarrow \infty} \mathbf{x}^{(k)}$$

dafür, dass die Folge  $\mathbf{x}^{(k)}$  gegen  $\mathbf{a}$  konvergiert.

Man kann die Theorie der konvergenten Folgen von Vektoren im  $\mathbb{R}^n$  genauso aufbauen, wie wir das im ersten Semester für Folgen reeller Zahlen taten. Man muss nur überall den Absolutbetrag  $|x_\nu - a|$  durch den Abstand  $d(\mathbf{x}^{(k)}, \mathbf{a})$  ersetzen. Und weil alles ganz genauso geht, wie in der Dimension eins, hat sich so um 1960 herum die Gewohnheit eingebürgert, in der Analysis-Vorlesung gar nicht erst die eindimensionale Analysis zu behandeln, sondern gleich alles im  $\mathbb{R}^n$  zu machen. Das war damals der Zug der Zeit. Ich finde soetwas eine didaktische Katastrophe.

Wir wollen den umgekehrten Weg gehen, und alles soweit wie nur irgend möglich auf die uns nunmehr (hoffentlich) vertraute eindimensionale Analysis zurückführen. Dazu dient der folgende Satz.

**Satz 5.1** Eine Folge  $(\mathbf{x}^{(k)})$  von Vektoren im  $\mathbb{R}^n$  konvergiert genau dann gegen einen Vektor  $\mathbf{a} \in \mathbb{R}^n$ , wenn für  $\nu = 1, \dots, n$  die Folge der  $\nu$ -ten Komponenten der Vektoren  $\mathbf{x}^{(k)}$  gegen  $a_\nu$  konvergiert, also

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{a} \iff \lim_{k \rightarrow \infty} x_\nu^{(k)} = a_\nu \text{ für } \nu = 1, \dots, n.$$

Dem Beweis schicken wir ein Abschätzungs-Lemma voraus.

**Satz 5.2 (Abschätzungs-Lemma)** Es seien  $\mathbf{a}$  und  $\mathbf{b} \in \mathbb{R}^n$  zwei Vektoren. Dann ist

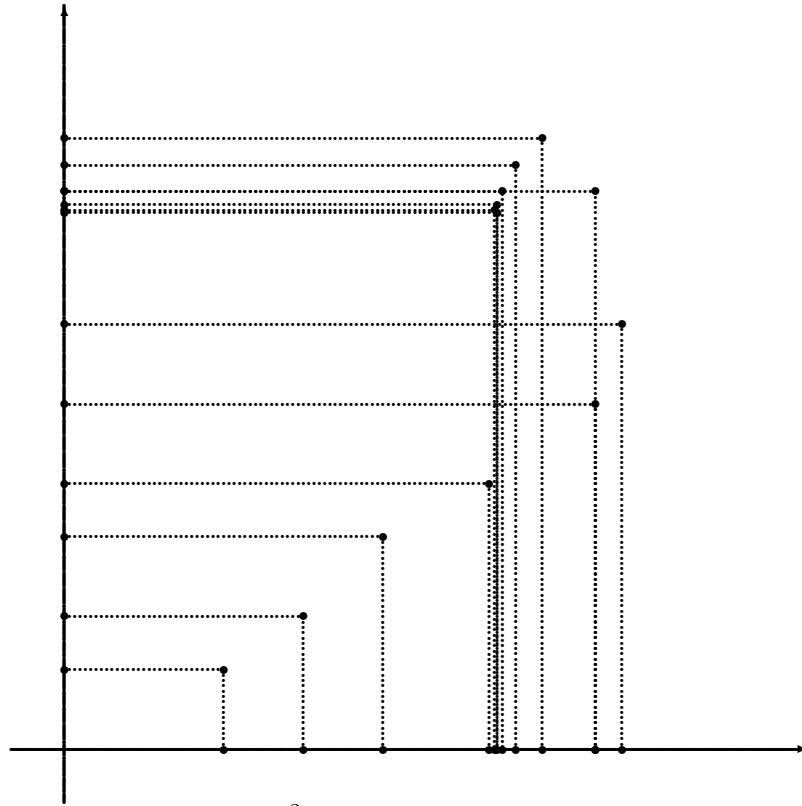
$$\max_{\nu=1}^n |a_\nu - b_\nu| \leq d(\mathbf{a}, \mathbf{b}) \leq \sqrt{n} \cdot \max_{\nu=1}^n |a_\nu - b_\nu|.$$

Beweis. Für  $\nu = 1, \dots, n$  ist

$$|a_\nu - b_\nu| = \sqrt{(a_\nu - b_\nu)^2} \leq \sqrt{\sum_{\mu=1}^n (a_\mu - b_\mu)^2} = d(\mathbf{a}, \mathbf{b}).$$

Das war die erste Ungleichung, und jetzt kommt die zweite:

$$d(\mathbf{a}, \mathbf{b}) = \sqrt{\sum_{\nu=1}^n (a_\nu - b_\nu)^2} \leq \sqrt{n \cdot \max_{\nu=1}^n (a_\nu - b_\nu)^2} = \sqrt{n} \cdot \max_{\nu=1}^n |a_\nu - b_\nu|. \quad \square$$



Eine Folge im  $\mathbb{R}^2$  und ihre beiden Komponentenfolgen

Beweis von Satz 1). a) Es sei  $\mathbf{a} = \lim \mathbf{x}^{(k)}$  vorausgesetzt. Nach dem Abschätzungs-Lemma ist für  $\nu = 1, \dots, n$  stets  $|x_\nu^{(k)} - a_\nu| \leq d(\mathbf{x}^{(k)}, \mathbf{a})$ , also

$$0 \leq \lim_{k \rightarrow \infty} |x_\nu^{(k)} - a_\nu| \leq \lim_{k \rightarrow \infty} d(\mathbf{x}^{(k)}, \mathbf{a}) = 0.$$

Also konvergiert die  $\nu$ -te Komponentenfolge  $(x_\nu^{(k)})_{k \in \mathbb{N}}$  gegen  $a_\nu$ .

b) Jetzt nehmen wir umgekehrt an, dass für  $\nu = 1, \dots, n$  gilt

$$\lim_{k \rightarrow \infty} x_\nu^{(k)} = a_\nu.$$

Zu jedem  $\epsilon > 0$  gibt es also Schranken  $N_1, \dots, N_n$  mit

$$|x_\nu^{(k)} - a_\nu| < \epsilon \text{ für } k > N_\nu.$$

Für  $k \geq \max_{\nu=1}^n N_\nu$  ist also

$$d(\mathbf{x}^{(k)}, \mathbf{a}) \leq \sqrt{n} \cdot \max_{\nu=1}^n |x_\nu^{(k)} - a_\nu| < \sqrt{n} \cdot \epsilon.$$

Das bedeutet

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{a}.$$

□

**Satz 5.3 (Konvergente Folgen und offene Mengen)** Eine Folge von Vektoren  $\mathbf{x}^{(k)} \in \mathbb{R}^n$  konvergiere gegen den Vektor  $\mathbf{a} \in \mathbb{R}^n$ . Ist  $U \subset \mathbb{R}^n$  eine offene Menge, die den Vektor  $\mathbf{a}$  enthält, so enthält  $U$  auch alle Folgenvektoren  $\mathbf{x}^{(k)}$  bis auf endlich viele davon.

Beweis. Nach Definition der offenen Mengen gibt es eine Kugel  $K$  um  $\mathbf{a}$  mit  $K \subset U$ . Es sei  $r$  deren Radius. Nach Definition der Konvergenz gibt es ein  $N(r) \in \mathbb{N}$  mit

$$d(\mathbf{x}^{(k)}, \mathbf{a}) < r \text{ für } k > N(r).$$

Für  $k > N(r)$  gehören also alle Folgen-Vektoren  $\mathbf{x}^{(k)}$  zur Kugel  $K$ , und damit zu  $U$ . □

**Satz 5.4 (Konvergente Folgen und abgeschlossene Mengen)** Eine Folge von Vektoren  $\mathbf{x}^{(k)} \in \mathbb{R}^n$  konvergiere gegen den Vektor  $\mathbf{a} \in \mathbb{R}^n$ . Ist  $A \subset \mathbb{R}^n$  eine abgeschlossene Menge, die alle Folgenvektoren  $\mathbf{x}^{(k)}$  enthält, so enthält sie auch den Grenzwert  $\mathbf{a}$ .

Beweis. Nach Definition der abgeschlossenen Mengen ist das Komplement  $U := \mathbb{R}^n \setminus A$  offen. Würde  $\mathbf{a}$  zu diesem Komplement  $U$  gehören, so würden nach Satz 5.3 auch alle Folgenvektoren  $\mathbf{x}^{(k)}$ , bis auf endlich viele, zu  $U$  gehören. Nach Voraussetzung ist dies aber nicht der Fall. Also gehört  $\mathbf{a}$  nicht zu  $U$ , sondern zu  $A$ . □

**Definition 5.6** Eine Teilmenge  $B \subset \mathbb{R}^n$  heißt beschränkt, wenn es eine Kugel  $K \subset \mathbb{R}^n$  gibt, die  $B$  enthält.

Nach dieser Definition sind z.B. alle Kugeln beschränkt, aber keine Gerade ist es. Eine Umformulierung der Definition ist: Die Menge  $B$  ist beschränkt, wenn es eine Schranke  $R$  gibt mit  $\|\mathbf{b}\| < R$  für alle  $\mathbf{b} \in B$ .

**Satz 5.5 (Konvergente und beschränkte Folgen)** a) Jede konvergente Folge im  $\mathbb{R}^n$  ist beschränkt. b) (Bolzano-Weierstraß) Jede beschränkte Folge im  $\mathbb{R}^n$  besitzt eine konvergente Teilfolge.

Beweis. a) Die Folge  $(\mathbf{x}^{(k)})$  konvergiere gegen  $\mathbf{a}$ . Dann gibt es also eine Kugel  $K_{\mathbf{a}}$  um  $\mathbf{a}$  und ein  $N \in \mathbb{N}$  mit  $\mathbf{x}^{(k)} \in K_{\mathbf{a}}$  für  $k > N$ . Außerdem gibt es eine Kugel  $K'$ , welche die endlich vielen Folgenvektoren  $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}$  enthält. Wir können eine Kugel  $K$  wählen, welche die beiden Kugeln  $K_{\mathbf{a}}$  und  $K'$  enthält. In dieser Kugel  $K$  liegen dann alle Vektoren  $\mathbf{x}^{(k)}$  der Folge.

b) Die Folge  $(\mathbf{x}^{(k)})$  sei beschränkt, es gelte etwa  $\|\mathbf{x}^{(k)}\| < R$  für alle  $k$ . Nach dem Abschätzungs-Lemma ist dann auch jede Komponentenfolge  $(x_{\nu}^{(k)})_{k \in \mathbb{N}}$  eine beschränkte Folge reeller Zahlen:

$$|x_{\nu}^{(k)}| = |x_{\nu}^{(k)} - 0| \leq d(\mathbf{x}^{(k)}, \mathbf{0}) = \|\mathbf{x}^{(k)}\|$$

Nach dem Satz von Bolzano-Weierstraß für Folgen reeller Zahlen aus dem ersten Semester besitzt die erste Komponentenfolge  $(x_1^{(k)})_{k \in \mathbb{N}}$  eine konvergente Teilfolge. Nachdem wir zur entsprechenden Teilfolge der Folge  $\mathbf{x}^{(k)}$  übergehen, können wir also annehmen, dass die erste Komponentenfolge  $(x_1^{(k)})$  konvergiert.

Diesen Schritt wiederholen wir jetzt für die Folge  $(x_2^{(k)})$  der zweiten Komponenten. Nachdem wir nochmal zu einer Teilfolge übergehen, können wir also annehmen, dass die beiden Komponentenfolgen  $(x_1^{(k)})$  und  $(x_2^{(k)})$  konvergieren. Und wenn wir diesen Schritt  $n$  mal durchgeführt haben, haben wir also eine Teilfolge der ursprünglichen Folge  $\mathbf{x}^{(k)}$  gefunden, deren  $n$  Komponentenfolgen  $(x_{\nu}^{(k)})$ ,  $\nu = 1, \dots, n$ , alle konvergieren. Nach Satz 5.1 konvergiert also diese Teilfolge auch selbst. □

Die folgende Definition brauchen wir vor allem für den Satz vom Maximum im  $\mathbb{R}^n$ :

**Definition 5.7** Eine Teilmenge  $M \subset \mathbb{R}^n$  heißt kompakt, wenn sie beschränkt und abgeschlossen ist.

Nach dieser Definition sind beispielsweise alle abgeschlossenen Kugeln kompakt. Offene Kugeln sind nicht kompakt, weil sie nicht abgeschlossen sind. Geraden sind nicht kompakt, weil sie nicht beschränkt sind.

**Satz 5.6 (Teilfolgen kompakter Mengen)** Die Menge  $M \subset \mathbb{R}^n$  sei kompakt. Dann besitzt jede Folge  $(\mathbf{x}^{(k)})_{k \in \mathbb{N}}$  von Vektoren  $\mathbf{x}^{(k)} \in M$  eine konvergente Teilfolge mit Grenzwert in  $M$ .

Beweis. Sei also  $(\mathbf{x}^{(k)})$  eine Folge von Vektoren aus  $M$ . Weil  $M$  beschränkt ist, ist auch die Folge beschränkt, und besitzt nach Satz 5.5b) eine konvergente Teilfolge. Diese Teilfolge liegt auch wieder in  $M$ . Weil  $M$  auch abgeschlossen ist, gehört der Grenzwert dieser Folge nach Satz 5.4 selbst auch zu  $M$ .  $\square$

**Aufgabe 5.9** Zeigen Sie, dass die Vereinigung endlich vieler kompakter Mengen  $K_i \subset \mathbb{R}^n$  wieder kompakt ist.

**Aufgabe 5.10** Seien  $A, B \subset \mathbb{R}$  kompakte Mengen. Zeigen Sie dass auch  $A \times B \subset \mathbb{R}^2$  kompakt ist.

**Aufgabe 5.11** Beweisen Sie die Umkehrung von Satz 5.6: Ist  $M \subset \mathbb{R}^n$  eine Menge mit der Eigenschaft, dass jede Folge  $\mathbf{x}^{(k)}$  von Punkten  $\mathbf{x}^{(k)} \in M$  eine konvergente Teilfolge mit Grenzwert in  $M$  besitzt, dann ist  $M$  kompakt.

**Aufgabe 5.12** Es sei  $C$  eine kompakte und  $U$  eine offene Teilmenge des  $\mathbb{R}^2$ . Beweisen Sie, dass auch  $C \setminus U$  eine kompakte Teilmenge des  $\mathbb{R}^2$  ist.

## 5.4 Stetigkeit

Die Stetigkeit einer Abbildung  $F : U \rightarrow V$ ,  $U \subset \mathbb{R}^m$ ,  $V \subset \mathbb{R}^n$  kann man ganz genauso definieren, wie die Stetigkeit einer Funktion einer reellen Variablen im ersten Semester, sogar auf zwei Weisen:

Sei dazu  $\mathbf{x} \in U$  und  $\mathbf{y} = F(\mathbf{x}) \in V$ .

**Definition 5.8 ( $\epsilon$ - $\delta$ -Stetigkeit))** Die Abbildung  $F$  heißt stetig in  $\mathbf{x}$ , wenn zu jedem  $\epsilon > 0$  ein  $\delta > 0$  existiert mit

$$\mathbf{x}' \in U, d(\mathbf{x}', \mathbf{x}) < \delta \quad \Rightarrow \quad d(F(\mathbf{x}'), F(\mathbf{x})) < \epsilon.$$



**Definition 5.9 (Folgen-Stetigkeit)** Die Abbildung  $F$  heißt stetig in  $\mathbf{x}$ , wenn für jede Folge  $(\mathbf{x}^{(k)})_{k \in \mathbb{N}}$  von Punkten  $\mathbf{x}^{(k)} \in U$  mit

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}$$

gilt

$$\lim_{k \rightarrow \infty} F(\mathbf{x}^{(k)}) = \mathbf{y}.$$

Beide Definitionen sind - wie im ersten Semester auch - äquivalent:

**Satz 5.7** Die Abbildung  $F$  ist in  $\mathbf{x}$  genau dann folgenstetig, wenn sie  $\epsilon$ - $\delta$ -stetig ist.

Beweis. „ $\epsilon$ - $\delta$ -Stetigkeit  $\Rightarrow$  Folgenstetigkeit“: Sei  $F$   $\epsilon$ - $\delta$ -stetig in  $\mathbf{x}$ . Sei  $\mathbf{x}^{(k)}$  eine Folge in  $U$ , die gegen  $\mathbf{x}$  konvergiert. Wir zeigen, dass die Bildfolge  $F(\mathbf{x}^{(k)})$  gegen  $F(\mathbf{x}) = \mathbf{y}$  konvergiert. Dazu müssen wir uns ein  $\epsilon > 0$  vorgeben. Nach Voraussetzung gibt es ein  $\delta > 0$  mit  $d(\mathbf{x}', \mathbf{x}) < \delta \Rightarrow d(F(\mathbf{x}'), \mathbf{y}) < \epsilon$ . Weil die Folge  $\mathbf{x}^{(k)}$  gegen  $\mathbf{x}$  konvergiert, gibt es zu diesem  $\delta$  ein  $N$  mit

$$k > N \quad \Rightarrow \quad d(\mathbf{x}^{(k)}, \mathbf{x}) < \delta.$$

Für  $k > N$  ist dann auch

$$d(F(\mathbf{x}^{(k)}), \mathbf{y}) < \epsilon.$$

Damit ist die Folgen-Stetigkeit nachgewiesen.

„Folgenstetigkeit  $\Rightarrow$   $\epsilon$ - $\delta$ -Stetigkeit“: Diese Richtung geht mit Widerspruch. Nehmen wir also an, die Abbildung  $F$  sei folgenstetig in  $\mathbf{x}$ , aber nicht  $\epsilon$ - $\delta$ -stetig. Dass die  $\epsilon$ - $\delta$ -Eigenschaft nicht erfüllt ist, heißt:

Es gibt ein  $\epsilon > 0$  derart, dass  
zu jedem  $\delta > 0$  ein

$$\mathbf{x}' \in U \text{ existiert mit } d(\mathbf{x}', \mathbf{x}) < \delta \\ \text{und } d(F(\mathbf{x}'), \mathbf{y}) > \epsilon.$$

Spielen wir das einmal durch für  $\delta := 1/k$ . Das  $\mathbf{x}' \in U$  mit  $d(\mathbf{x}', \mathbf{x}) < \delta = 1/k$ , das dazu existiert, nennen wir  $\mathbf{x}^{(k)}$ . Dann ist  $d(\mathbf{x}^{(k)}, \mathbf{x}) < 1/k$  eine Nullfolge, und die Folge  $\mathbf{x}^{(k)}$  konvergiert gegen  $\mathbf{x}$ . Für die Bildvektoren  $F(\mathbf{x}^{(k)})$  der Vektoren dieser Folge gilt immer  $d(F(\mathbf{x}^{(k)}), \mathbf{y}) > \epsilon$ . Dabei ist das  $\epsilon$  das vom Anfang des Beweises, also im Beweis fest. Deswegen kann die Folge  $F(\mathbf{x}^{(k)})$  unmöglich gegen  $\mathbf{y}$  konvergieren, im Widerspruch zur vorausgesetzten Folgenstetigkeit.  $\square$

**Satz 5.8 (Hintereinanderschalten stetiger Abbildungen)** Es seien  $U \subset \mathbb{R}^m$ ,  $V \subset \mathbb{R}^n$ ,  $W \subset \mathbb{R}^p$  Mengen und

$$F : U \rightarrow V, G : V \rightarrow W$$

stetige Abbildungen. Dann ist die zusammengesetzte Abbildung

$$G \circ F : U \rightarrow W$$

auch wieder stetig.

Beweis. Wir benützen die Folgenstetigkeit. Sei also  $\mathbf{x} \in U$  und  $\mathbf{x}^{(k)}$  eine Folge von Punkten in  $U$  mit Grenzwert  $\mathbf{x}$ . Wegen der Stetigkeit von  $F$  konvergiert die Bildfolge  $\mathbf{y}^{(k)} := F(\mathbf{x}^{(k)})$  gegen  $\mathbf{y} = F(\mathbf{x})$ .

Wegen der Stetigkeit von  $G$  konvergiert deren Bildfolge  $G(\mathbf{y}^{(k)})$  gegen  $G(\mathbf{y})$ . Insgesamt konvergiert also

$$(G \circ F)(\mathbf{x}^{(k)}) = G(F(\mathbf{x}^{(k)})) = G(\mathbf{y}^{(k)})$$

gegen

$$G(\mathbf{y}) = G(F(\mathbf{x})) = (G \circ F)(\mathbf{x}).$$

Damit ist die Folgenstetigkeit der Abbildung  $G \circ F$  nachgewiesen.  $\square$

Mit dem nächsten Satz spielen wir die Stetigkeit von Abbildungen auf die Stetigkeit ihrer Komponentenfunktionen zurück.

**Satz 5.9 (Stetigkeit und Komponentenfunktionen)** *Eine Abbildung  $F : U \rightarrow V$ ,  $U \subset \mathbb{R}^m$ ,  $V \subset \mathbb{R}^n$ , ist genau dann stetig im Punkt  $\mathbf{x} \in U$ , wenn alle ihre  $n$  Komponentenfunktionen  $F_\nu : U \rightarrow \mathbb{R}$  in diesem Punkt  $\mathbf{x}$  stetig sind.*

*Beweis.* Wir betrachten eine Folge  $(\mathbf{x}^{(k)})$  von Vektoren  $\mathbf{x}^{(k)} \in U$  und die Bildfolge  $F(\mathbf{x}^{(k)}) \in V$ . Nach Satz 5.1 konvergiert diese Bildfolge genau dann gegen  $F(\mathbf{x})$  wenn die  $n$  Komponentenfolgen  $F(\mathbf{x}^{(k)})_\nu = F_\nu(\mathbf{x}^{(k)})$  gegen  $F_\nu(\mathbf{x})$  konvergieren. Also ist die Abbildung  $F$  folgenstetig in  $\mathbf{x}$ , genau dann, wenn ihre  $n$  Komponentenfolgen in  $\mathbf{x}$  folgenstetig sind.  $\square$

**Beispiel 5.11** *Beispiele für stetige Funktionen: Die Koordinatenfunktionen  $\mathbf{x} \rightarrow x_\nu$  sind stetig, denn wenn eine Folge  $\mathbf{x}^{(k)}$  gegen  $\mathbf{x}$  konvergiert, so konvergiert nach Satz 5.1 jede ihrer Komponentenfunktionen  $x_\nu^{(k)}$  gegen  $x_\nu$ .*

*Sind  $f, g : U \rightarrow \mathbb{R}$  stetige Funktionen, so zeigt man mit der Folgenstetigkeit, und den Grenzwert-Rechenregeln aus dem ersten Semester, dass auch die Funktionen*

$$f \pm g, \quad f \cdot g, \quad f/g \text{ wo } g \neq 0$$

*wieder stetig sind.*

*Daraus folgt, dass alle Polynome*

$$\sum a_{\nu_1, \dots, \nu_n} x_1^{\nu_1} \cdot \dots \cdot x_n^{\nu_n}$$

*in  $n$  Veränderlichen stetig sind. Polynome in  $n > 1$  Veränderlichen sind sehr gefährliche Tiere. Es ist sogar schon kompliziert, solche Polynome überhaupt nur richtig hinzuschreiben. Ein besonders einfaches Polynom aber ist z.B.*

$$f(x_1, \dots, x_n) = x_1^2 + \dots + x_n^2 = \|\mathbf{x}\|^2.$$

*Mit den elementaren Funktionen aus der eindimensionalen Analysis und den Koordinatenfunktionen kann man durch Hintereinanderschaltung alle Funktionen zusammensetzen, die man so braucht. Als Beispiel betrachten wir die Norm*

$$\|\mathbf{x}\| = (g \circ f)(\mathbf{x})$$

*mit*

$$f : \mathbb{R}^n \rightarrow [0, \infty) \subset \mathbb{R}, \quad f(\mathbf{x}) = x_1^2 + \dots + x_n^2$$

*und*

$$g : [0, \infty) \rightarrow \mathbb{R}, \quad g(y) = \sqrt{y}.$$

Damit ist gezeigt, dass die Funktion

$$\mathbb{R}^n \rightarrow \mathbb{R}, \quad \mathbf{x} \rightarrow \|\mathbf{x}\|$$

stetig ist.

Für den nächsten Satz erinnern wir uns an die Definition von Bild- und Urbild-Mengen. Sei dazu  $F : U \rightarrow \mathbb{R}^n$  eine Abbildung. Ist  $M \subset U$  eine Teilmenge, so heißt

$$F(M) = \{F(\mathbf{x}) : \mathbf{x} \in M\}$$

die Bildmenge von  $M$ . Ist  $A \subset \mathbb{R}^n$  eine Teilmenge, so heißt

$$F^{-1}(A) = \{\mathbf{x} \in U : F(\mathbf{x}) \in A\}$$

die Urbildmenge von  $A$ .

**Satz 5.10 (Fundamental-Eigenschaften stetiger Abbildungen)** *Es sei  $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$  stetig.*

- a) *Ist  $A \subset \mathbb{R}^n$  abgeschlossen, so ist das Urbild  $F^{-1}(A) \subset \mathbb{R}^m$  auch wieder abgeschlossen.*
- b) *Ist  $B \subset \mathbb{R}^n$  offen, so ist das Urbild  $F^{-1}(B) \subset \mathbb{R}^m$  auch wieder offen.*
- c) *Ist  $K \subset \mathbb{R}^m$  kompakt, so ist das Bild  $F(K) \subset \mathbb{R}^n$  auch wieder kompakt.*

Beweis. a) Ist  $A \subset \mathbb{R}^n$  abgeschlossen, so ist  $B := \mathbb{R}^n \setminus A$  in  $\mathbb{R}^n$  offen. Und es ist

$$\begin{aligned} F^{-1}(A) &= \{\mathbf{x} \in \mathbb{R}^m : F(\mathbf{x}) \in A\} \\ &= \mathbb{R}^m \setminus \{\mathbf{x} \in \mathbb{R}^m : F(\mathbf{x}) \notin A\} \\ &= \mathbb{R}^m \setminus \{\mathbf{x} \in \mathbb{R}^m : F(\mathbf{x}) \in B\} \\ &= \mathbb{R}^m \setminus F^{-1}(B). \end{aligned}$$

Wir brauchen also nur die Aussage b) zu beweisen, denn dann folgt, dass  $F^{-1}(B) \subset \mathbb{R}^m$  offen und  $F^{-1}(A) = \mathbb{R}^m \setminus F^{-1}(B)$  offen in  $\mathbb{R}^m$  ist.

b) Um zu zeigen, dass  $F^{-1}(B) \subset \mathbb{R}^m$  offen ist, müssen wir beweisen: Um jeden Punkt  $\mathbf{x} \in F^{-1}(B)$  gibt es eine Kugel  $K_{\mathbf{x}} \subset \mathbb{R}^m$  mit  $\mathbf{x}$  als Mittelpunkt, die wieder ganz in  $F^{-1}(B)$  liegt. Nach Definition der Urbildmenge  $F^{-1}(B)$  ist  $\mathbf{y} := F(\mathbf{x}) \in B$ . Weil  $B$  offen ist, gibt es eine Kugel  $K_{\mathbf{y}}$  mit  $\mathbf{y}$  als Mittelpunkt, die ganz in  $B$  liegt. Sei  $\epsilon$  der Radius dieser Kugel. Weil  $F$  in  $\mathbf{x}$  stetig ist, gibt es ein  $\delta > 0$  mit  $d(\mathbf{x}', \mathbf{x}) < \delta \Rightarrow d(F(\mathbf{x}'), F(\mathbf{x})) < \epsilon$ . Mit anderen Worten: Ist  $K_{\mathbf{x}} \subset \mathbb{R}^m$  die Kugel um  $\mathbf{x}$  mit diesem Radius  $\delta$ , so liegt für alle Punkte  $\mathbf{x}' \in K_{\mathbf{x}}$  der Bildpunkt  $F(\mathbf{x}')$  in  $K_{\mathbf{y}} \subset B$ . Die ganze Kugel  $K_{\mathbf{x}}$  gehört also in der Tat zur Urbildmenge  $F^{-1}(B)$ .

c) Wir müssen zeigen, die Bildmenge  $F(K) \subset \mathbb{R}^n$  ist beschränkt und abgeschlossen.

Wenn  $F(K)$  nicht beschränkt ist, dann gibt es zu jeder natürlichen Zahl  $k$  einen Punkt  $\mathbf{y}^{(k)} \in F(K)$  mit  $\|\mathbf{y}^{(k)}\| > k$ . Die Folge  $\mathbf{y}^{(k)}$  konvergiert nicht, im Gegenteil, sie divergiert ins Unendliche. Auch keine Teilfolge dieser Folge kann konvergieren.

Wenn  $F(K) \subset \mathbb{R}^n$  nicht abgeschlossen ist, ist  $B := \mathbb{R}^n \setminus F(K)$  nicht offen. Es gibt also einen Punkt  $\mathbf{y} \in \mathbb{R}^n \setminus F(K)$ , so dass keine Kugel um  $\mathbf{y}$  noch ganz zu dieser Menge  $\mathbb{R}^n \setminus F(K)$  gehört. Es gibt also zu jedem Radius  $r = 1/k$  einen Punkt  $\mathbf{y}^{(k)}$  mit  $d(\mathbf{y}^{(k)}, \mathbf{y}) < 1/k$  aber  $\mathbf{y}^{(k)} \notin \mathbb{R}^n \setminus F(K)$ . Die Punkte  $\mathbf{y}^{(k)}$  bilden also eine Folge in  $F(K)$ , die gegen  $\mathbf{y} \notin F(K)$  konvergiert. Auch jede Teilfolge dieser Folge konvergiert gegen  $\mathbf{y}$ .

Wir sehen: Ist  $F(K)$  nicht kompakt, so gibt es eine Folge von Punkten  $\mathbf{y}^{(k)} \in F(K)$ , die nicht in  $F(K)$  konvergiert. Auch keine ihrer Teilfolgen tut das. Das führt zu einem Widerspruch: Denn jeder Punkt  $\mathbf{y}^{(k)} \in F(K)$  ist Bildpunkt  $F(\mathbf{x}^{(k)})$  eines Punktes  $\mathbf{x}^{(k)} \in K$ . Weil  $K$  kompakt ist, gibt es eine Teilfolge der Folge  $\mathbf{x}^{(k)}$ , welche einen Grenzwert  $\mathbf{x} \in K$  hat. Die Bildfolge dieser Teilfolge ist eine Teilfolge der ursprünglichen Folge  $\mathbf{y}^{(k)}$ , die wegen der Stetigkeit von  $F$  gegen  $F(\mathbf{x}) \in F(K)$  konvergiert. Widerspruch!  $\square$

Dieser Satz 5.10 ist furchtbar abstrakt. Aber man kann viel mit ihm machen. Schauen wir uns einige Anwendungen an.

**Beispiel 5.12** Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  eine stetige Funktion. Wenden wir Satz 5.10a) auf die abgeschlossene Menge  $A = \{0\} \subset \mathbb{R}$  an, so sehen wir

$$f^{-1}(A) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = 0\}$$

ist eine abgeschlossene Menge. Die Nullstellenmenge einer stetigen Funktion ist also immer abgeschlossen! Nehmen wir etwa eine Koordinatenfunktion  $x_\nu$ , so sehen wir: Jede Koordinaten-Hyperebene  $\{x_\nu = 0\}$  ist abgeschlossen. Oder nehmen wir  $f = \|\mathbf{x}\| - r$ , so finden wir: Die Sphäre

$$f^{-1}(\{0\}) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = r\}$$

ist abgeschlossen.

**Beispiel 5.13** Wenden wir jetzt Satz 5.10b) auf ein offenes Intervall  $(-\infty, r) \subset \mathbb{R}$  an, so sehen wir: Jede Menge

$$\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) < c\},$$

also die Menge der Punkte, wo  $f$  Werte  $< c$  hat, ist offen. Ein Spezialfall sind die offenen Kugeln ( $f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\|$ ,  $c = r > 0$ ).

Ein Spezialfall von Satz 5.10c) ist der Satz vom Maximum:

**Satz 5.11 (vom Maximum)** Eine stetige Funktion  $f : K \rightarrow \mathbb{R}$  auf einer kompakten Menge  $K$  nimmt auf  $K$  ihr Maximum an.

Beweis. Die Bildmenge  $f(K) \subset \mathbb{R}$  ist kompakt.

Dann ist sie auf jeden Fall beschränkt. Deswegen existiert  $m := \sup(f(K))$ , das Supremum dieser Bildmenge.

Außerdem ist  $f(K) \subset \mathbb{R}$  abgeschlossen. Nach Definition des Supremums aus dem ersten Semester gibt es eine Folge  $(y^{(k)})$  von Punkten  $y^{(k)} \in f(K)$ , die gegen das Supremum konvergiert. Deswegen gehört  $m$  zur Menge  $f(K)$  und ist deren Maximum.

Wegen  $m \in f(K)$  gibt es also ein  $\mathbf{x} \in K$  mit  $f(\mathbf{x}) = m$ . Wegen  $m \geq y$  für alle  $y \in f(K)$  ist also  $f(\mathbf{x}) \geq f(\mathbf{x}') = y$  für alle  $\mathbf{x}' \in K$ . Die Funktion  $f$  nimmt also in  $\mathbf{x} \in K$  ihren größten Wert an.  $\square$

**Aufgabe 5.13** Zeigen Sie, dass jede lineare Abbildung  $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$  stetig ist.

**Aufgabe 5.14** Gegeben sei die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  durch

$$f(x, y) = \begin{cases} \exp\left(-\frac{y^2}{x}\right) & \text{falls } x \neq 0 \\ 1 & \text{falls } x = 0 \end{cases}.$$

a) Zeigen Sie, dass für jedes  $s \in \mathbb{R}$  die Funktion

$$\mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto f(x, sx)$$

stetig an der Stelle 0 ist.

b) Ist  $f$  stetig an der Stelle  $(0, 0)$ ?

**Aufgabe 5.15** Gegeben sei die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(0, 0) = 0$  und  $f(x, y) = \frac{x^2 y}{x^4 + y^2}$  für jedes  $(x, y) \in \mathbb{R}^2$  mit  $(x, y) \neq (0, 0)$ . Zeigen Sie:

a) Für jedes  $(a, b) \in \mathbb{R}$  mit  $(a, b) \neq (0, 0)$  ist die Funktion  $g : \mathbb{R} \rightarrow \mathbb{R}$ ,  $t \mapsto f(ta, tb)$ , stetig an der Stelle  $t = 0$ .

b)  $f$  ist nicht stetig an der Stelle  $(0, 0)$ .

**Aufgabe 5.16** Die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  sei definiert durch

$$f(x, y) := \begin{cases} 0 & \text{für } (x, y) = (0, 0) \\ \frac{x \cdot |y|}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0) \end{cases}$$

Ist die Funktion  $f$  stetig in  $(0, 0)$ ?

**Aufgabe 5.17** Gegeben sei die Funktion

$$f(x, y) = e^{\frac{x}{x^2 + y^2}} \quad (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}.$$

a) Zeigen Sie unter Verwendung von Polarkoordinaten, dass  $f$  in jeder Kreisscheibe um den Nullpunkt alle positiven Werte annimmt.

b) Untersuchen Sie, ob  $f$  in  $(0, 0)$  stetig fortsetzbar ist.

**Aufgabe 5.18** Es sei  $C$  eine nichtleere, kompakte Teilmenge des  $\mathbb{R}^2$  und  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  eine stetige Funktion mit  $f(x) > 0$  für jedes  $x \in C$ . Beweisen Sie, dass es ein  $\epsilon \in \mathbb{R}$  gibt mit  $\epsilon > 0$  und

$$f(C) \cap ]-\epsilon, \epsilon[ = \emptyset.$$

## 6 Differentialrechnung mit mehreren Veränderlichen

Hier wollen wir uns also um die Ableitung von Funktionen  $f(x_1, \dots, x_n)$  in mehr als einer Variablen kümmern. Mit  $n$  Variablen  $x_1, \dots, x_n$  geht alles ganz genauso wie mit zwei Variablen. Auch in den Staatsexamensaufgaben zur Differenzierbarkeit in mehreren Veränderlichen kommen immer nur Funktionen von zwei Variablen vor. Wir wollen uns deswegen vor allem auf Funktionen  $f(x_1, x_2)$  von zwei Variablen konzentrieren. Weil es übersichtlicher ist, schreiben wir diese Funktionen als  $f(x, y)$  und nicht mehr  $f(x_1, x_2)$ .

Etwas anderes werde ich von hier an manchmal auch machen, das mir (wie jedem Mathematik-Dozenten) sehr widerstrebt: Ich werde nicht alle Sätze beweisen. Und zwar dann, wenn der Beweis ziemlich aufwendig ist, und eigentlich nichts zum Verständnis der Aussagen beiträgt. Das ändert nichts daran, dass die Sätze, bei denen ich mir das erlaube, wichtig sind. Ich gebe aber in jedem Fall die Stelle im Forster II an, wo ein interessierter Student einen vollständigen Beweis finden kann. Mit Forster II meine ich das Buch:

Forster, O.: Analysis 2, rororo vieweg, erste Auflage, 1977.

### 6.1 Partielle Ableitungen

Eine Funktion von zwei Variablen, etwa

$$f(x, y) = x \cdot y, \quad \text{oder} \quad = x^2 + y^2, \quad \text{oder} \quad = x^y$$

kann man nach jeder der beiden Variablen differenzieren. Dabei muss man dann die andere Variable als Konstante auffassen. Diese Art von Ableitung bezeichnet man als partielle Ableitung und schreibt sie mit einem gerundeten  $d$  als  $\partial/\partial x$  oder  $\partial/\partial y$ .

**Beispiel 6.1** *So ist z.B.*

$$\begin{aligned} \frac{\partial(xy)}{\partial x} &= y, & \frac{\partial(x^2 + y^2)}{\partial x} &= 2x, & \frac{\partial(x^y)}{\partial x} &= y \cdot x^{y-1} \\ \frac{\partial(xy)}{\partial y} &= x, & \frac{\partial(x^2 + y^2)}{\partial y} &= 2y, & \frac{\partial(x^y)}{\partial y} &= \ln(x) \cdot x^y. \end{aligned}$$

**Definition 6.1** *Gegeben sei eine Funktion  $f : U \rightarrow \mathbb{R}$  mit  $U \subset \mathbb{R}^2$  und ein Punkt  $(x^{(0)}, y^{(0)}) \in U$ . Diese Funktion  $f(x, y)$  heißt partiell differenzierbar in dem gegebenen Punkt, wenn die beiden Funktionen einer Variablen*

$$g_1(x) := f(x, y^{(0)}) \text{ in } x^{(0)} \quad \text{und} \quad g_2(y) = f(x^{(0)}, y) \text{ in } y^{(0)}$$

*differenzierbar sind. Die Ableitungen dieser beiden Funktionen einer Variablen heißen dann die partiellen Ableitungen nach  $x$  und  $y$  von  $f$  in dem festen Punkt:*

$$\frac{\partial f}{\partial x}(x^{(0)}, y^{(0)}) = \frac{d}{dx} f(x, y^{(0)})|_{x=x^{(0)}}, \quad \frac{\partial f}{\partial y}(x^{(0)}, y^{(0)}) = \frac{d}{dy} f(x^{(0)}, y)|_{y=y^{(0)}}.$$

**Beispiel 6.2** Noch ein Beispiel:  $f(x, y) := \sqrt{x^2 + y^2}$ , die Normfunktion. Für festes  $y^{(0)}$  ist

$$f(x, y^{(0)}) = \sqrt{x^2 + (y^{(0)})^2}$$

differenzierbar nach  $x$ , falls  $(x, y^{(0)}) \neq (0, 0)$ , mit Ableitung

$$\frac{\partial f}{\partial x} = \frac{2x}{2\sqrt{x^2 + (y^{(0)})^2}} = \frac{x}{\sqrt{x^2 + (y^{(0)})^2}}.$$

**Definition 6.2** Die Funktion  $f(x, y)$  sei partiell differenzierbar nach  $x$  und nach  $y$ . Dann heißt das Paar

$$\text{grad } f = \nabla f = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$$

der Gradient der Funktion  $f$ . Der Gradient  $\nabla f(x^{(0)}, y^{(0)})$  in einem Punkt ist also ein Vektor. Ist  $f$  partiell differenzierbar auf einer Menge  $U \subset \mathbb{R}^2$ , so ist  $\nabla f$  eine Abbildung  $U \rightarrow \mathbb{R}^2$ , ein Vektorfeld.

(Diese Definitionen gehen natürlich genauso für Funktionen in  $n$  Variablen. Der Gradient

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

ist auch hier ein Vektorfeld.)

Manches geht für die partiellen Ableitungen genauso, wie für die Ableitung in einer Variablen. Dazu gehört die notwendige Bedingung für lokale Extrema. Ist  $f : U \rightarrow \mathbb{R}$  eine Funktion, so sagt man, sie nimmt in einem Punkt  $(x^{(0)}, y^{(0)}) \in U$  ein Maximum (Minimum) an, wenn

$$f(x, y) \leq f(x^{(0)}, y^{(0)}) \text{ für alle } (x, y) \in U,$$

(bzw.  $\geq$ ). Sie nimmt ein lokales Maximum (Minimum) an, wenn es eine Kugel  $K$  um  $(x^{(0)}, y^{(0)})$  gibt, so dass die Ungleichung für alle  $(x, y) \in K$  gilt. Ein (lokales) Extremum ist ein (lokales) Maximum oder Minimum.

**Satz 6.1 (Notwendige Bedingung für lokale Extrema)** Die Funktion  $f : U \rightarrow \mathbb{R}$  sei partiell differenzierbar auf der offenen Menge  $U \subset \mathbb{R}^2$ . Ist  $\mathbf{x}^{(0)} = (x^{(0)}, y^{(0)})$  ein lokales Extremum für  $f$ , so verschwinden dort die beiden partiellen Ableitungen:

$$\frac{\partial f}{\partial x}(\mathbf{x}^{(0)}) = \frac{\partial f}{\partial y}(\mathbf{x}^{(0)}) = 0.$$

Beweis. Weil  $U$  offen ist, gibt es eine ganze Kugel mit einem Radius  $r > 0$ , die noch ganz in  $U$  liegt. Die beiden Funktionen einer Variablen

$$g_1(x) = f(x, y^{(0)}) \quad \text{und} \quad g_2(y) = f(x^{(0)}, y)$$

sind dann differenzierbar auf den Intervallen  $(x^{(0)} - r, x^{(0)} + r)$ , bzw.  $(y^{(0)} - r, y^{(0)} + r)$  und nehmen in  $x^{(0)}$ , bzw.  $y^{(0)}$  beide ein lokales Extremum an. Wegen des notwendigen Kriteriums für lokale Extrema aus der Analysis einer Variablen verschwinden ihre Ableitungen in den betreffenden Punkten. Das sind aber die beiden partiellen Ableitungen von  $f$  im Punkt  $\mathbf{x}^{(0)}$ .  $\square$

**Beispiel 6.3** Betrachten wir die wohlbekannte Funktion  $f(x, y) := x^2 + y^2$ . Wenn wir ihr (globales) Minimum im Nullpunkt noch nicht kennen würden, könnten wir es mit dem notwendigen Kriterium für Extrema suchen. Dazu müssten wir die beiden partiellen Ableitungen

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y$$

= 0 setzen. Das ergibt die Bedingung  $x = y = 0$ , den Nullpunkt.

**Beispiel 6.4** Ändern wir die Funktion etwas ab:  $f(x, y) = x^2 - y^2$ . Die hat jetzt im Nullpunkt kein Extremum mehr nicht einmal ein lokales, denn

$$f(x, 0) = x^2 > 0 \text{ für } x \neq 0, \quad f(0, y) = -y^2 < 0 \text{ für } y \neq 0.$$

Wenn wir lokale Extrema mit dem notwendigen Kriterium suchen, müssen wir genauso

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = -2y$$

= 0 setzen, und erhalten wieder den Nullpunkt. Das ist aber kein lokales Extremum. Immerhin können wir noch schließen: Außerhalb des Nullpunkts ist das notwendige Kriterium sicher nicht erfüllt, da gibt es keinerlei lokales Extremum.

Aus Beispiel 6.4 lernen wir, dass das notwendige Kriterium für lokale Extrema keineswegs hinreichend ist. Das war ja auch nicht zu erwarten, weil es in einer Variablen auch nicht hinreichend war. Nur: in einer Variablen war das eine subtile Sache, und hing mit den höheren Ableitungen zusammen. In mehreren Variablen geht das Kriterium auf grandiose Weise schief: Es kann für  $x$  erfüllt sein, weil da ein Minimum vorliegt, und auch für  $y$ , weil da ein Maximum vorliegt. Und das Minimum in  $x$ -Richtung, zusammen mit dem Maximum in  $y$ -Richtung ergibt nie und nimmer mehr ein Extremum (zumindest, wenn das echte Maxima/Minima waren). Man nennt das dann einen *Sattelpunkt*.

**Definition 6.3** Ein Punkt  $(x^{(0)}, y^{(0)})$  wo

$$\frac{\partial f}{\partial x}(x^{(0)}, y^{(0)}) = \frac{\partial f}{\partial y}(x^{(0)}, y^{(0)}) = 0$$

ist, heißt kritischer oder stationärer Punkt von  $f$ , egal ob dort ein Extremum vorliegt oder nicht.

**Beispiel 6.5** Gegeben sei die Funktion

$$f(x, y) := 6 \cdot (\sin(y) + \cos(y)) - 12 \cdot x^2 - 11.$$

Gesucht seien das absolute Maximum und das absolute Minimum von  $f$  auf der Menge

$$\{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq \pi - x^2\}.$$

Diese Menge ist unten beschränkt von der  $x$ -Achse, oben von der nach unten geöffneten Parabel  $y = \pi - x^2$ . Diese Menge ist begrenzt und auch abgeschlossen, wegen der  $\leq$ -Zeichen in ihrer Definition.



Damit ist die Menge kompakt. Die Funktion  $f$  ist stetig, nimmt hier also ihr Maximum und ihr Minimum an, und diese gilt es zu finden.

Wer in der letzten Zeit aufgepasst hat, weiß, was man tun muss: die partiellen Ableitungen bilden

$$f_x = -24 \cdot x, \quad f_y = 6 \cdot (\cos(y) - \sin(y))$$

und  $= 0$  setzen:

$$f_x = 0 \Leftrightarrow x = 0, \quad f_y = 0 \Leftrightarrow \cos(y) = \sin(y).$$

Wegen  $0 \leq y \leq \pi$  ist  $\cos(y) = \sin(y)$  nur für  $y = \pi/4$ , wo  $\sin(y) = \cos(y) = \sqrt{2}/2$ . Damit haben wir einen kritischen Punkt gefunden:

$$(x, y) = \left(0, \frac{\pi}{4}\right) \quad \text{mit} \quad f(x, y) = 6\sqrt{2} - 11.$$

Einerseits ist dies ja ein Erfolg. Andererseits: Handelt es sich hier um das Minimum oder das Maximum, oder beides? oder keines von beiden? Und wenn unser Punkt einer von beiden ist, wo ist dann der andere? Es scheint, die Probleme gehen jetzt erst los!

Eines ist klar: Wir haben den einzigen stationären Punkt der Funktion  $f$  in der gegebenen Menge identifiziert. Die Lösung findet man, wenn man Satz 6.1 nochmal liest, und auch auf die Unterstreichung achtet: In einem lokalen Extremum verschwinden die partiellen Ableitungen nur dann, wenn man eine ganze Kreisscheibe um den kritischen Punkt herum betrachtet. Auf dem Rand ist das Kriterium nicht anwendbar! Wir müssen noch den Rand untersuchen. Der besteht aus zwei Teilen:

a) Das Geradenstück

$$\{(x, 0) : -\sqrt{\pi} \leq x \leq \sqrt{\pi}\}.$$

Hier ist die Funktion

$$f(x, 0) = 6 \cdot \cos(0) - 12x^2 - 11 = -12x^2 - 5,$$

eine nach unten geöffnete Parabel symmetrisch zum Scheitel bei  $x = 0$ . Auf unserem Geradenstück ist das Maximum bei  $x = 0$ , das Minimum in den Randpunkten bei  $x = \pm\sqrt{\pi}$ . Die Werte sind:

	$(x, y)$	$f(x, y)$
Maximum	$(0, 0)$	$-5$
Minimum	$(\pm\sqrt{\pi}, 0)$	$-12\pi - 5$

b) Der Parabelbogen

$$\{(x, y) = (x, \pi - x^2) : -\sqrt{\pi} \leq x \leq \sqrt{\pi}\}.$$

Dort ist die Funktion

$$f(x, \pi - x^2) = 6(\sin(\pi - x^2) + \cos(\pi - x^2)) - 12x^2 - 11 = 6(\sin(x^2) - \cos(x^2)) - 12x^2 - 11.$$

Die Ableitung dieser Funktion ist

$$6 \cdot (\cos(x^2) + \sin(x^2)) \cdot 2x - 24x = 12x \cdot (\cos(x^2) + \sin(x^2) - 2).$$

Hier ist die Klammer  $\neq 0$ , weil nie

$$\cos(x^2) = \sin(x^2) = 1$$

sein kann. Die Ableitung verschwindet nur für  $x = 0$ . Damit erhalten wir den Punkt

$(x, y)$	$f(x, y)$
$(0, \pi)$	$-6 - 11 = -17$

Weil die Endpunkte des Parabelbogens auch die Endpunkte des Geradenstücks aus a) sind, brauchen wir sie hier nicht noch einmal zu betrachten.

Insgesamt haben wir die folgenden fünf Kandidaten für globale Extrema:

$(x, y)$	$f(x, y)$
$(0, \frac{\pi}{4})$	$6\sqrt{2} - 11$
$(0, 0)$	$-5$
$(\pm\sqrt{\pi}, 0)$	$-12\pi - 5$
$(0, \pi)$	$-17$

Wir müssen nur noch die Funktionswerte vergleichen: Offensichtlich ist

$$-12\pi - 5 < -17 < -5.$$

Und

$$6\sqrt{2} - 11 > -5 \quad \Leftrightarrow \quad 6\sqrt{2} > 6$$

ist auch klar. Damit lautet die Antwort in Tabellenform

	$(x, y)$	$f(x, y)$
globales Maximum	$(0, \pi/4)$	$6\sqrt{2} - 11$
globales Minimum	$(0, \pi)$	$-17$

Manches ist also ganz anders als in einer Variablen. Dazu gehört auch die Sache mit der linearen Approximation: Eine differenzierbare Funktion  $f(x)$  in einer Variablen kann man linear approximieren:

$$f(x + h) = f(x) + f'(x) \cdot h + \varphi(h)$$

mit

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

Die Gerade

$$h \mapsto f(x) + f'(x) \cdot h$$

heißt die Tangente an den Graphen der Funktion  $f$  im Punkt  $(x, f(x))$ .

Für eine partiell differenzierbare Funktion  $f(x, y)$  kann man ganz analog eine *Tangential-Ebene* an deren Graphen im Punkt  $(x, y, f(x, y))$  definieren. Das ist die Ebene, die durch die affin-lineare Abbildung

$$(h, k) \mapsto f(x, y) + \frac{\partial f}{\partial x}(x, y) \cdot h + \frac{\partial f}{\partial y}(x, y) \cdot k$$

parametrisiert wird. Die approximiert den Funktionsgraphen schön linear in  $x$ -Richtung und in  $y$ -Richtung, aber es gibt zwischen diesen beiden Richtungen noch viele andere Richtungen, gegeben durch Geraden  $\mathbb{R} \cdot (a, b)$ ,

$$t \mapsto t \cdot (a, b), \quad a \neq 0 \neq b.$$

Und in diesen Richtungen kann die Approximation leider sehr böse schiefgehen.

**Beispiel 6.6** *Wie dramatisch das aussehen kann, zeigt die Funktion*

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

*Außerhalb des Nullpunkts ist der Nenner, wie sich das gehört, von 0 verschieden, und die Funktion ist dort überall partiell differenzierbar. Auf beiden Koordinatenachsen ist  $f(x, y) \equiv 0$ , identisch = 0. Deswegen ist  $f$  auch im Nullpunkt partiell differenzierbar mit den partiellen Ableitungen*

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0.$$

*Aber wenn wir uns dem Nullpunkt in Richtung der Geraden  $\mathbb{R} \cdot (a, b)$  nähern, wird die Funktion*

$$f(ta, tb) = \frac{t^2 \cdot ab}{t^4 \cdot (a^2 + b^2)^2} = \frac{1}{t^2} \cdot \frac{ab}{(a^2 + b^2)^2}.$$

*Für  $t \rightarrow 0$  wächst  $f(ta, tb)$  über alle Berge, und ist nie und nimmer linear zu approximieren, schon gar nicht durch die harmlose Tangentialebene*

$$(h, k) \mapsto \frac{\partial f}{\partial x}(0, 0) \cdot h + \frac{\partial f}{\partial y}(0, 0) \cdot k \equiv 0.$$

Natürlich ist der Grund für dieses absonderliche Verhalten, der dass die Funktion  $f$  ziemlich künstlich ist. Bei einigermaßen vernünftigen Funktionen passiert soetwas nicht:

**Satz 6.2 (Stetig partiell differenzierbare Funktionen)** *Die Funktion  $f : U \rightarrow \mathbb{R}$  sei partiell differenzierbar auf der offenen Menge  $U \subset \mathbb{R}^2$  und die beiden partiellen Ableitungen  $\partial f/\partial x$  sowie  $\partial f/\partial y$  seien auf  $U$  stetig. Dann kann man  $f$  in jedem Punkt  $(x, y) \in U$  linear approximieren,*

$$f(x + h, y + k) = f(x, y) + \frac{\partial f}{\partial x}(x, y) \cdot h + \frac{\partial f}{\partial y}(x, y) \cdot k + \varphi(h, k)$$

mit

$$\lim_{(h, k) \rightarrow (0, 0)} \frac{\varphi(h, k)}{\|(h, k)\|} = 0.$$

Das ist genau die analoge Formel zu der linearen Approximation in einer Variablen. Ich möchte diesen Satz hier nicht beweisen. Einen Beweis finden Sie z.B. im Forster II auf p. 48.

**Definition 6.4** *Eine Funktion  $f : U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^2$  offen, heißt im Punkt  $(x, y) \in U$  total differenzierbar, wenn sie linear approximierbar ist, wenn also die Formel in Satz 6.2 gilt.*

Dieses Wort „total differenzierbar“ hat sich eingebürgert, wahrscheinlich vor allem deswegen, weil es sprachlich so ein schöner Gegensatz zu „partiell differenzierbar“ ist, aber wohl auch deswegen, weil man aus der Eigenschaft „total differenzierbar“ schließen kann, dass die Funktion  $f$  in jeder Richtung - nicht nur in den beiden Koordinatenrichtungen - differenzierbar ist:

**Satz 6.3 (Stetigkeit, Richtungsableitung)** Ist  $f : U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^2$  offen, in  $(x, y) \in U$  total differenzierbar, so gilt:

a)  $f$  ist in  $(x, y)$  stetig.

b) Für jeden Richtungsvektor  $(a, b) \neq (0, 0)$  ist die Funktion einer Variablen

$$g(t) = f(x + t \cdot a, y + t \cdot b)$$

differenzierbar im Punkt  $t = 0$  mit Ableitung

$$\frac{dg}{dt} \Big|_{t=0} = \frac{\partial f}{\partial x}(x, y) \cdot a + \frac{\partial f}{\partial y}(x, y) \cdot b.$$

Beweis. a) Es ist

$$\lim_{(h,k) \rightarrow (0,0)} [f(x+h, y+k) - f(x, y)] = \lim_{h \rightarrow 0} \frac{\partial f}{\partial x}(x, y) \cdot h + \lim_{k \rightarrow 0} \frac{\partial f}{\partial y}(x, y) \cdot k + \lim_{(h,k) \rightarrow (0,0)} \varphi(h, k)$$

mit

$$\lim_{(h,k) \rightarrow (0,0)} \varphi(h, k) = \lim_{(h,k) \rightarrow (0,0)} \frac{\varphi(h, k)}{\|(h, k)\|} \cdot \lim_{(h,k) \rightarrow (0,0)} \|(h, k)\| = 0.$$

Daraus folgt

$$\lim_{(h,k) \rightarrow (0,0)} f(x+h, y+k) = f(x, y),$$

also die Stetigkeit von  $f$  in  $(x, y)$ .

b) Wir setzen ganz einfach  $(h, k) = (t \cdot a, t \cdot b)$  in die Approximationsformel ein:

$$\begin{aligned} g(t) &= f(x, y) + \frac{\partial f}{\partial x}(x, y) \cdot t \cdot a + \frac{\partial f}{\partial y}(x, y) \cdot t \cdot b + \varphi(ta, tb) \\ &= f(x, y) + \left( \frac{\partial f}{\partial x}(x, y) \cdot a + \frac{\partial f}{\partial y}(x, y) \cdot b \right) \cdot t + \varphi(ta, tb) \end{aligned}$$

Dabei ist

$$\lim_{t \rightarrow 0} \frac{\varphi(ta, tb)}{|t|} = \lim_{(ta, tb) \rightarrow 0} \frac{\varphi(ta, tb)}{\sqrt{t^2 a^2 + t^2 b^2}} \cdot \sqrt{a^2 + b^2} = 0 \cdot \sqrt{a^2 + b^2} = 0.$$

Also ist  $g(t)$  im Punkt  $t = 0$  differenzierbar, und für seine Ableitung gilt die angegebene Formel.  $\square$

Fassen wir zusammen: Es gelten die Implikationen

$\text{stetig partiell differenzierbar} \Rightarrow \text{total differenzierbar} \Rightarrow \text{partiell differenzierbar.}$
--

**Aufgabe 6.1** Für  $x > 0$  und  $y > 0$  sei

$$G(x, y) = x^y + (\ln x) \cdot [(\arctan(\sin(\cos(xy))))^3 - \ln(x + y)].$$

Man berechne die partielle Ableitung 1. Ordnung von  $G$  nach der zweiten Variablen an der Stelle  $(1, 2)$ .

**Aufgabe 6.2** Für welche  $(x, y) \in \mathbb{R}^2$  ist

$$\begin{aligned} f : \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x, y) &\mapsto |x \cdot y| \end{aligned}$$

partiell nach  $x$  differenzierbar?

**Aufgabe 6.3** Gegeben sei die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  durch

$$f(x, y) := \begin{cases} \sin(x) \cdot \arctan\left(\ln\left(\frac{1}{\sqrt{x^2+y^2}}\right)\right) & \text{falls } (x, y) \neq (0, 0) \\ 0 & \text{falls } (x, y) = (0, 0) \end{cases}$$

Zeigen Sie, dass  $f$  im Punkt  $(0, 0)$  partiell nach  $x$  differenzierbar ist und bestimmen Sie dort diese erste partielle Ableitung nach  $x$ .

**Aufgabe 6.4** Bestimmen Sie die Tangentialebene im Punkt  $(1, -2, 2)$  der Fläche im  $\mathbb{R}^3$ , die durch die Gleichung

$$z = 3x^2y + 2xy^2$$

gegeben ist.

**Aufgabe 6.5** Bestimmen Sie die Tangentialebene an den Graphen der Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) = x^3 + y^3 - 3x - 12y$  im Punkt  $(2/0/2)$ .

**Aufgabe 6.6** Sei  $f(x, y) = 1 + x^4 + ay^4$  mit  $a \in \mathbb{R}$ . Man diskutiere in Abhängigkeit von  $a$ , ob  $f$  in  $(0, 0)$  ein lokales Extremum hat.

**Aufgabe 6.7** Gegeben sei die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f(x, y) := x^2 + 2xy - 4x - 4y$  und das Quadrat  $Q = [0, 2] \times [0, 2]$ .

a) Man begründe, dass es  $a, b \in Q$  gibt mit  $f(a) = \inf(f(I))$  und  $f(b) = \sup(f(I))$ .

b) Man zeige  $f(Q) = [-8, 0]$ .

**Aufgabe 6.8** Gegeben sei die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad f(x, y) := \sin(x) \cdot \cos(y).$$

Bestimmen Sie alle globalen Extrema von  $f$  auf dem Quadrat

$$\{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq \pi, 0 \leq y \leq \pi\}.$$

**Aufgabe 6.9** Gegeben sei die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad f(x, y) := e^{-(x^2+y^2)} \cdot (x^2 - 2y^2).$$

- a) Bestimmen Sie  $\inf f(\mathbb{R}^2)$  und  $\sup f(\mathbb{R}^2)$ .  
b) Untersuchen Sie  $f$  auf lokale Extrema.

**Aufgabe 6.10** Zeigen Sie, dass die Funktion

$$f : \mathbb{R} \times ]0, \infty[ \rightarrow \mathbb{R}, \quad f(x, y) = y^x,$$

kein lokales Extremum besitzt.

**Aufgabe 6.11** Es seien  $a = (a_1, a_2), b = (b_1, b_2)$  und  $c = (c_1, c_2)$  drei verschiedene Punkte der Ebene  $\mathbb{R}^2$ . Bestimmen Sie alle Punkte  $p \in \mathbb{R}^2$  mit der Eigenschaft, dass

$$\|p - a\|^2 + 2\|p - b\|^2 + 3\|p - c\|^2$$

minimal ist.

**Aufgabe 6.12** Beweisen Sie, dass die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = e^x \cdot \cos(y) + e^y \cdot \cos(x)$$

keine lokalen Extremstellen besitzt. Versuchen Sie aber nicht, die kritischen Stellen von  $f$  explizit zu bestimmen!

**Aufgabe 6.13** Bestimmen Sie auf der abgeschlossenen Kreisscheibe

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$$

alle lokalen und globalen Extrema der Funktion

- a)  $f_1(x, y) := 3x^2 - 2xy + 3y^2$ .  
b)  $f_2(x, y) := e^{x^2+y^2} \cdot (x - 2y)^2$ .  
c)  $f_3(x, y) := x^3 - 3xy^2$ .

**Aufgabe 6.14** a) Bestimmen Sie drei positive reelle Zahlen, deren Summe gleich 1 und deren Produkt maximal ist.

b) Bei der Post dürfen nur Pakete verschickt werden, bei denen die Summe aus Länge und Umfang (Umfang =  $2 \times$  Breite +  $2 \times$  Höhe) nicht größer als  $d$  cm ist. Welches Maß für Länge, Breite und Höhe muss man wählen, damit das Paket größtmögliches Volumen hat und verschickt werden kann?

## 6.2 Höhere partielle Ableitungen, Taylor-Formel

Wenn alle partiellen Ableitungen  $\partial f/\partial x_\nu$  einer Funktion  $f$  wieder stetig sind, dann nennt man die Funktion *stetig partiell differenzierbar*. Und dann können diese partiellen Ableitungen auch wieder partiell differenzierbar sein. Die partiellen Ableitungen der partiellen Ableitungen heißen dann zweite partielle Ableitungen. Für

$$\frac{\partial}{\partial x_\mu} \left( \frac{\partial f}{\partial x_\nu} \right)$$

schreibt man etwas kürzer

$$\frac{\partial^2 f}{\partial x_\mu \partial x_\nu}.$$

Und für

$$\frac{\partial}{\partial x_\nu} \left( \frac{\partial f}{\partial x_\nu} \right)$$

schreibt man

$$\frac{\partial^2 f}{\partial x_\nu^2}.$$

Manchmal kürzt man noch weiter ab:

$$\frac{\partial^2 f}{\partial x_\mu \partial x_\nu} = \partial_\mu \partial_\nu f = f_{\mu,\nu}, \quad \frac{\partial^2 f}{\partial x_\nu^2} = \partial_\nu^2 f = f_{\nu,\nu}.$$

Eine Funktion  $f(x, y)$  kann also die folgenden vier zweiten partiellen Ableitungen haben:

$$\frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y \partial x}, \quad \frac{\partial^2 f}{\partial y^2}.$$

Diese zweiten partiellen Ableitungen kann man zu einer Matrix

$$H_f := \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

anordnen. Diese Matrix heißt *Hesse-Matrix* der Funktion  $f$ . (Natürlich geht das auch bei Funktionen von  $n > 2$  Veränderlichen. Dann ist die Hesse-Matrix keine  $2 \times 2$ -Matrix, sondern eine  $n \times n$ -Matrix.)

Auch diese zweiten partiellen Ableitungen können wieder alle stetig sein. Dann heißt die Funktion *zweimal stetig differenzierbar*.

**Satz 6.4 (Symmetrie der Ableitungen)** *Die Funktion  $f$  sei zweimal stetig partiell differenzierbar auf der offenen Menge  $U \subset \mathbb{R}^2$ . Dann ist*

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Auch hier möchte ich den Beweis nicht ausführen, und stattdessen auf Forster II, p.40, verweisen.

Satz 6.4 kann man so ausdrücken: Es kommt bei der Bildung der zweiten partiellen Ableitungen nicht auf die Reihenfolge an, in der man nach  $x$  oder  $y$  differenziert.

**Beispiel 6.7** Wir betrachten die Funktion

$$f(x, y) = x^y.$$

Ihre ersten partiellen Ableitungen haben wir schon einmal ausgerechnet:

$$\frac{\partial f}{\partial x} = y \cdot x^{y-1}, \quad \frac{\partial f}{\partial y} = x^y \cdot \ln(x).$$

Daraus erhalten wir dann die folgenden zweiten partiellen Ableitungen

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= y(y-1) \cdot x^{y-2} & \frac{\partial^2 f}{\partial x \partial y} &= y \cdot x^{y-1} \cdot \ln(x) + x^y \cdot \frac{1}{x} \\ \frac{\partial^2 f}{\partial y \partial x} &= x^{y-1} + y \cdot x^{y-1} \cdot \ln(x) & \frac{\partial^2 f}{\partial y^2} &= x^y \cdot \ln(x)^2. \end{aligned}$$

Die beiden gemischten zweiten partiellen Ableitungen sind untereinander gleich, und zwar

$$= x^{y-1} \cdot (y \cdot \ln(x) + 1),$$

ganz, wie Satz 6.4 das behauptet.

Wenn die zweiten partiellen Ableitungen wieder partiell differenzierbar sind, kann man nochmal partiell differenzieren, und erhält so die dritten partiellen Ableitungen. Und so kann man weitermachen. Das führt dann zu den höheren partiellen Ableitungen.

Wenn die  $k$ -ten partiellen Ableitungen wieder alle stetig sind, kann man Satz 6.4 anwenden, wo man statt der Funktion  $f$  ihre  $(k-2)$ -ten partiellen Ableitungen einsetzt. Man sieht: Auch bei den höheren partiellen Ableitungen kommt es nicht auf die Reihenfolge an, in der man nach  $x$  oder  $y$  differenziert. Dann gibt es also die folgenden  $k+1$  partiellen Ableitungen  $k$ -ter Ordnung:

$$\frac{\partial^k f}{\partial x^k}, \quad \frac{\partial^k f}{\partial x^{k-1} \partial y}, \quad \frac{\partial^k f}{\partial x^{k-2} \partial y^2}, \quad \dots, \quad \frac{\partial^k f}{\partial y^k}.$$

Und wozu braucht man diese höheren partiellen Ableitungen? Antwort: für die Taylor-Reihe der Funktion  $f$ . Das kann ja lustig werden, bei so viel partiellen Ableitungen! Nehmen wir mal an, die Funktion  $f(x, y)$  habe partielle Ableitungen beliebig großer Ordnung  $k$  (sei unendlich oft partiell differenzierbar). Wir fixieren also einen Entwicklungspunkt  $(x, y)$  und wollen die Funktionswerte  $f(x+h, y+k)$  in den ausgelenkten Punkten  $(x+h, y+k)$  durch eine Reihe ausdrücken. Was kann dabei nur herauskommen?

Gehen wir systematisch vor, und entwickeln die Funktion  $f(x+h, y+k)$  erst mal nach  $x$ , d.h. nach der Auslenkung  $h$ . Die Variable  $y+k$  lassen wir dabei erst mal fest, ein Parameter, von dem die Entwicklung abhängt. Wir erhalten dann eine ganz normale Taylor-Reihe

$$\sum_{\mu=0}^{\infty} \frac{1}{\mu!} \frac{\partial^{\mu} f}{\partial x^{\mu}}(x, y+k) \cdot h^{\mu}.$$

Die in den Koeffizienten auftretenden Funktionen können wir jetzt im zweiten Schritt nach  $y$ , d.h., nach  $k$  in Taylor-Reihen entwickeln:

$$\frac{\partial^{\mu} f}{\partial x^{\mu}}(x, y+k) = \sum_{\nu=1}^n \frac{1}{\nu!} \frac{\partial^{\nu}}{\partial y^{\nu}} \left( \frac{\partial^{\mu} f}{\partial x^{\mu}}(x, y) \right) \cdot k^{\nu}.$$



Setzen wir diese Entwicklungen in die erste Taylor-Reihe ein, so entsteht die Doppelreihe

$$\sum_{\mu, \nu=1}^{\infty} \frac{1}{\mu! \nu!} \frac{\partial^{\mu+\nu} f}{\partial y^{\nu} \partial x^{\mu}}(x, y) \cdot h^{\mu} k^{\nu}.$$

Um etwas Ordnung in diese Doppelreihe zu bringen sammeln wir die partiellen Ableitungen gleicher Ordnung  $m = \mu + \nu$ . Wir vertauschen - der Schönheit wegen - die Reihenfolge der partiellen Ableitungen, und benutzen außerdem

$$\frac{1}{\mu! \nu!} = \frac{1}{m!} \binom{m}{\mu} \quad (m = \mu + \nu).$$

So entsteht

$$\sum_{m=0}^{\infty} \frac{1}{m!} \sum_{\mu=0}^m \binom{m}{\mu} \cdot \frac{\partial^m f}{\partial x^{\mu} \partial y^{m-\mu}}(x, y) \cdot h^{\mu} k^{m-\mu}.$$

Das ist die *Taylor-Reihe* der Funktion  $f$  in zwei Variablen - ein Ungetüm.

Und dabei haben wir uns noch gar nicht um die üblichen Komplikationen gekümmert, nämlich:

- Die Taylor-Reihe braucht nicht zu konvergieren,
- und wenn sie konvergiert, dann braucht sie nicht gegen die Funktion  $f(x+h, y+k)$  zu konvergieren.

Ob diese Taylor-Reihe gegen die Funktion konvergiert, hängt - wie bei Funktionen einer Variablen - vom Restglied  $R_N$  ab. Das Restglied ist jetzt eine Funktion von zwei Variablen:  $R_N(h, k)$ . Mit diesem Restglied entsteht die *Taylor-Formel*

$$f(x+h, y+k) = \sum_{m=0}^N \frac{1}{m!} \sum_{\mu=0}^m \binom{m}{\mu} \cdot \frac{\partial^m f}{\partial x^{\mu} \partial y^{m-\mu}}(x, y) \cdot h^{\mu} k^{m-\mu} + R_N(h, k).$$

Schauen wir uns die ersten Fälle einmal an:

$N=1$ : Die Taylorformel wird

$$f(x+h, y+k) = f(x, y) + \frac{\partial f}{\partial x}(x, y)h + \frac{\partial f}{\partial y}(x, y)k + R_1(h, k).$$

Das ist die lineare Approximation der Funktion  $f$  durch die Tangentialebene an ihren Graphen.

$N=2$ : Wir bekommen

$$\begin{aligned} f(x+h, y+k) &= f(x, y) + \frac{\partial f}{\partial x}(x, y)h + \frac{\partial f}{\partial y}(x, y)k \\ &+ \frac{1}{2} \cdot \left( \frac{\partial^2 f}{\partial x^2}(x, y)h^2 + 2 \cdot \frac{\partial^2 f}{\partial x \partial y}(x, y)hk + \frac{\partial^2 f}{\partial y^2}(x, y)k^2 \right) \\ &+ R_2(h, k). \end{aligned}$$

Natürlich könnten wir auch noch die Formel für  $n = 3$  hinschreiben, aber die ist mir in Staatsexamens-Aufgaben noch nicht begegnet. Es scheint überhaupt so, als ob Taylor-Reihen in mehreren Variablen jetzt in Staatsexamens-Aufgaben nicht mehr vorkommen (ohne Gewähr!).

Jede Taylor-Formel ist so gut wie ihr Restglied. Eine Formel dafür steht im Forster II auf p. 56. In zwei Variablen sieht die so aus:

**Satz 6.5 (Restglied)** *Ist  $U \subset \mathbb{R}^2$  offen, die Funktion  $f : U \rightarrow \mathbb{R}$  mindestens  $N + 1$ -mal stetig partiell differenzierbar, ist  $(x, y) \in U$  ein Entwicklungspunkt, und  $K \subset U$  eine Kugel um diesen Entwicklungspunkt, so gibt es für jeden Punkt  $(x + h, y + k) \in K$  einen Zwischenpunkt  $(x + t \cdot h, y + t \cdot k)$ ,  $0 \leq t \leq 1$  so, dass*

$$R_N(h, k) = \frac{1}{(N+1)!} \sum_{\mu=0}^{N+1} \binom{N+1}{\mu} \frac{\partial^{N+1} f}{\partial x^\mu \partial y^{N+1-\mu}}(x + th, y + tk) h^\mu k^{N+1-\mu}.$$

Gelegentlich braucht man diese Formel (für  $N = 2$ ) tatsächlich, aber meistens reicht die folgende Information aus:

**Satz 6.6 (Abschätzung für das Restglied)** *Es seien die Voraussetzungen von Satz 6.5 erfüllt, allerdings braucht  $f$  jetzt nur  $N$ -mal stetig partiell differenzierbar zu sein. Dann gilt*

$$\lim_{(h,k) \rightarrow (0,0)} \frac{R_N(h, k)}{\|(h, k)\|^N} = 0.$$

Beweis. Wir benützen die Taylorformel bis zur Ordnung  $N - 1$

$$f(x + h, y + k) = \sum_{m=0}^{N-1} \frac{1}{m!} \sum_{\mu=0}^m \binom{m}{\mu} \cdot \frac{\partial^m f}{\partial x^\mu \partial y^{m-\mu}}(x, y) \cdot h^\mu h^{m-\mu} + R_{N-1}(h, k).$$

Die verändern wir folgendermaßen:

$$\begin{aligned} f(x + h, y + k) &= \sum_{m=0}^N \frac{1}{m!} \sum_{\mu=0}^m \binom{m}{\mu} \cdot \frac{\partial^m f}{\partial x^\mu \partial y^{m-\mu}} \cdot h^\mu h^{m-\mu} \\ &+ \left[ R_{N-1}(h, k) - \frac{1}{N!} \sum_{\mu=0}^N \binom{N}{\mu} \cdot \frac{\partial^N f}{\partial x^\mu \partial y^{N-\mu}}(x, y) \cdot h^\mu k^{N-\mu} \right]. \end{aligned}$$

Der Ausdruck in der eckigen Klammer ist dann das Restglied  $R_N$ . Mit der Formel für  $R_{N-1}$  aus Satz 6.5 wird dieser Ausdruck

$$R_N(k, h) = \frac{1}{N!} \sum_{\mu=0}^N \binom{N}{\mu} \cdot \left[ \frac{\partial^N f}{\partial x^\mu \partial y^{N-\mu}}(x + t \cdot h, y + t \cdot k) - \frac{\partial^N f}{\partial x^\mu \partial y^{N-\mu}}(x, y) \right] \cdot h^\mu k^{N-\mu}.$$

Weil die partiellen Ableitungen  $N$ -ter Ordnung von  $F$  stetig vorausgesetzt sind, geht für  $(h, k) \rightarrow (0, 0)$  der Ausdruck in der eckigen Klammer gegen 0.

Wir brauchen also nur noch für festes  $\mu$  zu zeigen: Für  $(h, k) \rightarrow (0, 0)$  bleibt

$$q(h, k) := \frac{h^\mu k^{N-\mu}}{\|(h, k)\|^N}$$

beschränkt. Nun gilt für alle  $t \in \mathbb{R}$ ,  $t > 0$ ,

$$q(th, tk) = \frac{t^N \cdot h^\mu k^{N-\mu}}{|t|^N \|(h, k)\|^N} = q(h, k).$$

Deswegen nimmt  $q(h, k)$  alle Werte, die es überhaupt annimmt, auch auf einer kleinen Kugel  $\|(h, k)\| = \epsilon > 0$  an. (Setzen Sie  $t := \epsilon / \|(h, k)\|$ .) Diese Kugel ist kompakt, die Funktion  $h^\mu k^{N-\mu}$  ist darauf stetig, und nach dem Satz vom Maximum beschränkt, etwa  $|h^\mu k^{N-\mu}| \leq C'$ . Daraus folgt:

$$|q(h, k)| \leq \frac{C'}{\epsilon}$$

ist tatsächlich beschränkt. □

**Aufgabe 6.15**  $h : \mathbb{R} \rightarrow \mathbb{R}$  und  $k : \mathbb{R} \rightarrow \mathbb{R}$  seien zweimal stetig differenzierbare Funktionen;  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  sei durch  $F(x, y) := h(x + k(y))$  definiert. Man berechne die partiellen Ableitungen zweiter Ordnung von  $F$ .

**Aufgabe 6.16** Warum gibt es keine partiell differenzierbare Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit

$$(\text{grad } f)(x, y) = (\arctan(xy), e^x \sin y)$$

für jedes  $(x, y) \in \mathbb{R}^2$ ?

**Aufgabe 6.17** Mit Hilfe der Taylor-Formel (Entwicklung bei  $(x_0, y_0) = (1, 0)$ ) beweise man

$$\frac{1}{x} e^y \geq 2 - x + y$$

für  $0 < x \leq 1$  und  $y \geq 0$ .

**Aufgabe 6.18** Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  eine zweimal stetig partiell differenzierbare Funktion mit  $f(0, 0) = 0$ , die an der Stelle  $(0, 0)$  ein lokales Extremum hat. Zeigen Sie mit Hilfe der Taylor-Formel, dass es eine Konstante  $C$  gibt, so dass für jedes  $(x, y) \in [-1, 1] \times [-1, 1]$  gilt:

$$|f(x, y)| \leq C(|x| + |y|)^2.$$

**Aufgabe 6.19** Zeigen Sie, dass die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \text{ mit } f(0, 0) = 0 \text{ und } f(x, y) = \frac{x^3 y - x y^3}{x^2 + y^2} \text{ für } (x, y) \neq (0, 0)$$

im Punkt  $(0, 0)$  zweimal partiell differenzierbar ist, und die Hesse-Matrix von  $f$  im Punkt  $(0, 0)$  nicht symmetrisch ist. Ist  $f$  zweimal stetig partiell differenzierbar?

**Aufgabe 6.20** Gegeben sei die Funktion  $f : ]0, \infty[ \times ]0, \infty[ \rightarrow \mathbb{R}$  mit  $f(x, y) = xy + \frac{a}{x} + \frac{a}{y}$ ,  $a > 0$ . Bestimmen Sie das Taylor-Polynom zweiten Grades von  $f$  im Entwicklungspunkt  $(a, a)$ .

**Aufgabe 6.21** Bestimmen Sie mit Hilfe der Taylor-Formel alle zweimal stetig partiell differenzierbaren Funktionen  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit konstanten zweiten partiellen Ableitungen.

**Aufgabe 6.22** Es seien  $F$  und  $G$  zweimal stetig partiell differenzierbare Funktionen von  $\mathbb{R}$  in  $\mathbb{R}$ , und  $u : \mathbb{R}^2 \rightarrow \mathbb{R}$  sei durch  $u(t, x) = F(x + t) + G(x - t)$  definiert. Man beweise

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}.$$

**Aufgabe 6.23** Es sei  $f : ]0, \infty[ \times ]0, \infty[ \rightarrow \mathbb{R}$  die durch  $f(x, y) = x^y$  für  $x > 0$  und  $y > 0$  gegebene reellwertige Funktion. Man berechne das Taylor-Polynom 2. Ordnung von  $f$  im Entwicklungspunkt  $(1, 1)$ .

### 6.3 Lokale Extrema

Hier wollen wir das hinreichende Kriterium für lokale Extrema differenzierbarer Funktionen von mehreren Variablen beweisen. Wir müssen dabei immer annehmen, dass wir einen inneren Punkt des Definitionsbereiches der Funktion  $f$  haben. Also ist die Voraussetzung für diesen ganzen Abschnitt:

$U \subset \mathbb{R}^2$  ist offen,  $f : U \rightarrow \mathbb{R}$  ist zweimal stetig partiell differenzierbar.

Wir können dann auch gleich annehmen, dass  $U = K$  eine offene Kugel um den Punkt  $(x, y)$  ist, in dem wir die Extremumseigenschaft für  $f$  nachweisen wollen.

Natürlich muss in diesem Punkt das notwendige Kriterium für lokale Extrema erfüllt sein: Die beiden ersten partiellen Ableitungen müssen dort verschwinden,

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial y}(x, y) = 0.$$

Information über die Funktionswerte von  $f$  beziehen wir aus der Taylor-Formel (Satz 6.5) für  $N = 2$  um den Entwicklungspunkt  $(x, y)$ . Wenn die ersten partiellen Ableitungen  $= 0$  sind, sieht die folgendermaßen aus:

$$f(x + h, y + k) = f(x, y) + \frac{1}{2} \left( \frac{\partial^2 f}{\partial x^2}(x, y) \cdot h^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(x, y) \cdot hk + \frac{\partial^2 f}{\partial y^2}(x, y) \cdot k^2 \right) + R_2(h, k)$$

mit

$$\lim_{(h,k) \rightarrow (0,0)} \frac{R_2(h, k)}{h^2 + k^2} = 0.$$

Die partiellen Ableitungen zweiter Ordnung kürzen wir ab:

$$a := \frac{\partial^2 f}{\partial x^2}(x, y), \quad b := \frac{\partial^2 f}{\partial x \partial y}(x, y), \quad c := \frac{\partial^2 f}{\partial y^2}(x, y).$$

Damit wird die Taylor-Formel

$$f(x+h, y+k) = f(x, y) + \frac{1}{2}(a \cdot h^2 + 2b \cdot hk + c \cdot k^2) + R_2(h, k).$$

Der homogene Anteil zweiter Ordnung (bis auf den Faktor 1/2)

$$a \cdot h^2 + 2b \cdot hk + c \cdot k^2,$$

ist das, was man in der linearen Algebra eine *quadratische Form* nennt. Mit der Hesse-Matrix

$$H = H_f = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

und dem Vektor

$$\mathbf{h} := (h, k)$$

kann man diesen Ausdruck schreiben als

$$\mathbf{h} \cdot H \cdot \mathbf{h}^t.$$

Und es wird darauf ankommen, ob dieser Anteil für alle  $\mathbf{h} \neq \mathbf{0}$  größer als 0 oder kleiner als 0 ist. Und genau das untersucht man in der Linearen Algebra. Dort hat man sogar eine eigenen Definition.

**Definition 6.5** Eine reelle symmetrische  $n \times n$ -Matrix  $H$  heißt positiv definit, wenn für alle Vektoren  $\mathbf{h} \in \mathbb{R}^n$ ,  $\mathbf{h} \neq \mathbf{0}$  gilt

$$\mathbf{h} \cdot H \cdot \mathbf{h}^t > 0.$$

Sie heißt negativ definit, wenn dieser Ausdruck immer  $< 0$  ist, und sie heißt indefinit, wenn es Vektoren  $\mathbf{h}$  gibt, für die das Resultat  $> 0$  ist, und andere, für die es  $< 0$  ist.

**Satz 6.7** a) Die  $n \times n$ -Matrix  $H$  ist genau dann positiv definit, wenn es ein  $m > 0$  gibt so, dass für alle Einheitsvektoren  $\mathbf{h}_1 \in \mathbb{R}^n$  (mit  $\|\mathbf{h}_1\| = 1$ ) gilt

$$\mathbf{h}_1 \cdot H \cdot \mathbf{h}_1^t \geq m.$$

b) Die  $n \times n$ -Matrix  $H$  ist genau dann negativ definit, wenn es ein  $m < 0$  gibt so, dass für alle Einheitsvektoren  $\mathbf{h}_1 \in \mathbb{R}^n$  gilt

$$\mathbf{h}_1 \cdot H \cdot \mathbf{h}_1^t \leq m.$$

c) Die  $n \times n$ -Matrix  $H$  ist genau dann indefinit, wenn es Einheitsvektoren  $\mathbf{h}_+$  und  $\mathbf{h}_-$  gibt mit

$$\mathbf{h}_+ \cdot H \cdot \mathbf{h}_+^t > 0, \quad \mathbf{h}_- \cdot H \cdot \mathbf{h}_-^t < 0.$$

Beweis. a) Sei  $H$  positiv definit vorausgesetzt. Dann ist für alle Einheitsvektoren  $\mathbf{h}_1 \in \mathbb{R}^n$

$$\mathbf{h}_1 \cdot H \cdot \mathbf{h}_1^t > 0,$$

denn Einheitsvektoren sind  $\neq \mathbf{0}$ . Nun ist die Funktion

$$\mathbf{h} \mapsto \mathbf{h} \cdot H \cdot \mathbf{h}^t = \sum_{\mu, \nu=1}^n H_{\mu, \nu} h_\mu h_\nu$$

auf  $\mathbb{R}^n$  ein Polynom und damit stetig. Weiter ist die Einheitskugel

$$\{\mathbf{h}_1 \in \mathbb{R}^n : \|\mathbf{h}_1\| = 1\}$$

beschränkt und abgeschlossen, also kompakt. Nach dem Satz vom Maximum gibt es einen Einheitsvektor, in dem sie ihr Minimum  $m$  annimmt. Auch dieses Minimum  $m$  ist  $> 0$ . Mit diesem Minimum  $m$  gilt somit

$$\mathbf{h}_1 \cdot H \cdot \mathbf{h}_1^t \geq m$$

für alle Einheitsvektoren  $\mathbf{h}_1$ .

Sei nun umgekehrt diese Ungleichung vorausgesetzt. Jeden Vektor  $\mathbf{h} \in \mathbb{R}^n$ ,  $\mathbf{h} \neq \mathbf{0}$ , kann man schreiben

$$\mathbf{h} = \|\mathbf{h}\| \cdot \mathbf{h}_1 \text{ mit dem Einheitsvektor } \mathbf{h}_1 := \frac{1}{\|\mathbf{h}\|} \cdot \mathbf{h}.$$

Damit wird

$$\mathbf{h} \cdot H \cdot \mathbf{h}^t = \|\mathbf{h}\|^2 \cdot \mathbf{h}_1 \cdot H \cdot \mathbf{h}_1^t \geq \|\mathbf{h}\|^2 \cdot m > 0.$$

b) Im negativ definiten Fall verläuft der Beweis genau so, man muss nur das  $>$ -Zeichen ersetzen durch  $<$ .

c) ist offensichtlich. □

**Satz 6.8 (Lokale Extrema, hinreichende Bedingung)** Die Funktion  $f : U \rightarrow \mathbb{R}$  sei zweimal stetig partiell differenzierbar auf der offenen Menge  $U \subset \mathbb{R}^2$ . In einem Punkt  $\mathbf{x}^{(0)} \in U$  gelte

$$\frac{\partial f}{\partial x}(\mathbf{x}^{(0)}) = \frac{\partial f}{\partial y}(\mathbf{x}^{(0)}) = 0.$$

Ist die Hesse-Matrix  $H_f$  von  $f$  im Punkt  $\mathbf{x}^{(0)}$  definit, so liegt ein lokales Extremum vor, und zwar ein

$$\begin{aligned} \text{Maximum} &\Leftrightarrow H_f \text{ negativ definit} \\ \text{Minimum} &\Leftrightarrow H_f \text{ positiv definit.} \end{aligned}$$

Ist die Hesse-Matrix indefinit, so liegt kein lokales Extremum vor.

Beweis. Wir benutzen die Taylor-Formel für  $N = 2$

$$f(\mathbf{x}^{(0)} + \mathbf{h}) = f(\mathbf{x}^{(0)}) + \frac{1}{2} \cdot \mathbf{h} \cdot H \cdot \mathbf{h}^t + R_2(\mathbf{h}) \quad \text{mit} \quad \lim_{\mathbf{h} \rightarrow 0} \frac{R_2(\mathbf{h})}{\|\mathbf{h}\|^2} = 0.$$

a) Sei  $H = H_f$  positiv definit und  $m > 0$  wie in Satz 6.7 a) gewählt. Dann gilt also

$$\mathbf{h} \cdot H \cdot \mathbf{h}^t \geq m \cdot \|\mathbf{h}\|^2.$$

Andererseits gibt es ein  $r > 0$  mit

$$\frac{|R_2(\mathbf{h})|}{\|\mathbf{h}\|^2} < \frac{m}{4}, \quad \text{d.h.} \quad |R_2(\mathbf{h})| < \frac{m}{4} \cdot \|\mathbf{h}\|^2$$

für  $\|\mathbf{h}\| < r$ . Für diese  $\mathbf{h}$  ist dann

$$\frac{1}{2} \cdot \mathbf{h} \cdot H \cdot \mathbf{h}^t + R_2(\mathbf{h}) \geq \left(\frac{m}{2} - \frac{m}{4}\right) \|\mathbf{h}\|^2 > 0$$

und

$$f(\mathbf{x}^{(0)} + \mathbf{h}) > f(\mathbf{x}^{(0)}).$$

In  $\mathbf{x}^{(0)}$  hat die Funktion  $f$  ein lokales Minimum.

b) Ist  $H$  negativ definit, so sieht man mit Satz 6.7 b) ganz analog, dass  $\mathbf{x}^{(0)}$  ein lokales Maximum für  $f$  ist.

c) Sei nun  $H$  indefinit. Nach Satz 6.7 c) gibt es ein  $m > 0$  und Einheitsvektoren  $\mathbf{h}_+, \mathbf{h}_-$  mit

$$\mathbf{h}_+ \cdot H \cdot \mathbf{h}_+^t \geq m, \quad \mathbf{h}_- \cdot H \cdot \mathbf{h}_-^t < -m.$$

Wie im Beweis von a) findet man ein  $r_+ > 0$  so, dass für  $|t| < r_+$  gilt

$$\frac{1}{2} \cdot (t\mathbf{h}_+) \cdot H \cdot (t\mathbf{h}_+)^t + R_2(t\mathbf{h}_+) > 0.$$

Ebenso gibt es ein  $r_- > 0$  mit

$$\frac{1}{2} \cdot (t\mathbf{h}_-) \cdot H \cdot (t\mathbf{h}_-)^t + R_2(t\mathbf{h}_-) < 0$$

für  $|t| < r_-$ . In jeder Kugel mit einem Radius

$$r < \min\{r_+, r_-\}$$

gibt es deswegen Punkte

$$\mathbf{x}^{(0)} + t \cdot \mathbf{h}_+ \text{ mit } f(\mathbf{x}^{(0)} + t \cdot \mathbf{h}_+) > f(\mathbf{x}^{(0)})$$

und Punkte

$$\mathbf{x}^{(0)} + t \cdot \mathbf{h}_- \text{ mit } f(\mathbf{x}^{(0)} + t \cdot \mathbf{h}_-) < f(\mathbf{x}^{(0)}).$$

In  $\mathbf{x}^{(0)}$  hat die Funktion  $f$  weder ein lokales Maximum, noch ein lokales Minimum. □

Natürlich bleibt jetzt noch das Problem, einer symmetrischen Matrix  $H$  anzusehen, ob sie positiv definit, negativ definit oder indefinit ist. (Dummerweise gibt es leider auch noch symmetrische Matrizen, die nichts von alledem sind, wo unsere wundervolle Theorie nicht greift.)

Einer Diagonalmatrix  $D$  sieht man sehr leicht an, ob sie eine dieser drei Eigenschaften besitzt: Sie ist

positiv definit	⇔	alle Diagonal-Einträge sind $> 0$
negativ definit	⇔	alle Diagonal-Einträge sind $< 0$
indefinit	⇔	es gibt Diagonal-Einträge $> 0$ und $< 0$ .

Nicht jede symmetrische Matrix  $H$  ist eine Diagonalmatrix. Aber in der Linearen Algebra beweist man auch den Satz von der Hauptachsen-Transformation: Jede symmetrische Matrix  $H$  lässt sich mit einer orthogonalen Matrix  $T$  in eine Diagonalmatrix  $D = T^t \cdot H \cdot T$  transformieren. Die Diagonaleinträge von  $D$  sind dabei die Eigenwerte von  $H$ . Daraus folgt: Die symmetrische Matrix  $H$  ist

positiv definit	⇔	alle ihre Eigenwerte sind $> 0$
negativ definit	⇔	alle ihre Eigenwerte sind $< 0$
indefinit	⇔	sie hat Eigenwerte $> 0$ und $< 0$ .

Das ist sehr schön, vor allem, wenn man die Eigenwerte kennt. Bei  $2 \times 2$ -Matrizen kann man die immer ziemlich leicht ausrechnen. Bei größeren Matrizen ist es i.A. schwierig. Dafür gibt es dann das *Hurwitz-Kriterium*. Da muss man alle linken oberen Unterdeterminanten (Haupt-Minoren) der Matrix ausrechnen. Sind alle diese Haupt-Minoren  $> 0$ , so ist die Matrix positiv definit. Wir brauchen nur den Fall  $n = 2$ . Da sieht dieses Kriterium so aus:

**Satz 6.9 (Hurwitz-Kriterium)** Die symmetrische  $2 \times 2$ -Matrix

$$H = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

ist

$$\begin{aligned} \text{positiv definit} &\Leftrightarrow \det(H) = ac - b^2 > 0 \text{ und } a > 0, \\ \text{negativ definit} &\Leftrightarrow \det(H) = ac - b^2 > 0 \text{ und } a < 0, \\ \text{indefinit} &\Leftrightarrow \det(H) = ac - b^2 < 0. \end{aligned}$$

Beweis. Betrachten wir zuerst den Fall

$$\det(H) = ac - b^2 > 0.$$

Hier kann nicht  $a = 0$  gelten. Deswegen können wir mit quadratischer Ergänzung schreiben

$$\begin{aligned} (x, y) \cdot H \cdot (x, y)^t &= a \cdot x^2 + 2b \cdot xy + c \cdot y^2 \\ &= a \cdot \left( x^2 + \frac{2b}{a} \cdot xy + \frac{c}{a} \cdot y^2 \right) \\ &= a \cdot \left( \left( x + \frac{b}{a}y \right)^2 + \left( \frac{c}{a} - \frac{b^2}{a^2} \right) y^2 \right) \\ &= a \cdot \left( \left( x + \frac{b}{a}y \right)^2 + \frac{ac - b^2}{a^2} y^2 \right) \end{aligned}$$

Hier ist der erste Summand in der Klammer als Quadrat eine Zahl  $\geq 0$ . Und für  $y \neq 0$  ist wegen  $ac > b^2$  der rechte Summand echt  $> 0$ . Für  $y \neq 0$  ist also die ganze Klammer echt  $> 0$ . Ist aber  $y = 0$  und  $x \neq 0$ , so ist der erste Summand in der Klammer  $= x^2 > 0$ . Damit ist für alle Vektoren  $(x, y) \in \mathbb{R}^2, (x, y) \neq (0, 0)$  die Klammer echt  $> 0$ . Somit ist, abhängig vom Vorzeichen von  $a$ , die Matrix  $H$  entweder positiv oder negativ definit.

Sei nun

$$\det(H) = ac - b^2 < 0.$$

Wenn  $a \neq 0$  ist, berechnen wir

$$\begin{aligned} (b, -a) \cdot \begin{pmatrix} a & b \\ b & c \end{pmatrix} \cdot \begin{pmatrix} b \\ -a \end{pmatrix} &= a \cdot (ac - b^2), \\ (1, 0) \cdot \begin{pmatrix} a & b \\ b & c \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= a. \end{aligned}$$

Wenn  $a > 0$  ist, dann fällt das erste Ergebnis negativ, das zweite positiv aus. Wenn  $a < 0$  ist, dann ist es umgekehrt. Wir sehen: Die Matrix  $H$  ist indefinit, falls  $a \neq 0$ .

Es bleibt der Fall  $a = 0$ , wo dann natürlich  $b \neq 0$  ist. Jetzt berechnen wir

$$(x, 1) \cdot \begin{pmatrix} 0 & b \\ b & c \end{pmatrix} \cdot \begin{pmatrix} x \\ 1 \end{pmatrix} = 2bx + c.$$

Durchläuft hier  $x$  alle reellen Zahlen so beschreibt das Ergebnis eine Gerade mit Steigung  $2b \neq 0$ . Für geeignete Wahl von  $x$  ergeben sich positive oder negative Werte. Auch jetzt ist  $H$  indefinit.  $\square$



**Beispiel 6.8** Es seien für die Funktion

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

alle lokalen Extrema gesucht.

Wir differenzieren

$$\begin{aligned} f_x &= 2x - e^{xy} \cdot y \\ f_y &= 2y - e^{xy} \cdot x \end{aligned}$$

und sicherheitshalber nochmal

$$\begin{aligned} f_{xx} &= 2 - e^{xy} \cdot y^2 \\ f_{xy} &= -e^{xy} \cdot xy - e^{xy} = -e^{xy} \cdot (xy + 1) \\ f_{yy} &= 2 - e^{xy} \cdot x^2 \end{aligned}$$

Die notwendigen Bedingungen  $f_x = f_y = 0$  schreiben wir als homogenes lineares Gleichungssystem

$$\begin{pmatrix} 2x & y \\ 2y & x \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -e^{xy} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

mit dem nicht-trivialen Lösungsvektor  $(1, -e^{xy})$ . Wegen der nicht-trivialen Lösung muss die Determinante der Koeffizientenmatrix

$$\det \begin{pmatrix} 2x & y \\ 2y & x \end{pmatrix} = 2x^2 - 2y^2 = 0$$

sein. Wir finden also

$$x = \pm y.$$

Leider sind wir noch nicht fertig. Wenn  $x = y$  ist müssen wir für  $x = y \neq 0$  die Bedingung

$$2 - e^{xy} = 0, \quad 2 - e^{x^2} = 0, \quad x^2 = \ln(2)$$

auswerten und finden

$$x = y = \sqrt{\ln(2)}.$$

Natürlich ist auch noch

$$x = y = 0$$

eine Möglichkeit. Wenn  $x = -y \neq 0$  sein sollte, dann haben wir

$$2 + e^{xy} = 0$$

auszuwerten. Aber wegen  $e^{xy} > 0$  gibt es da keine Lösung. Es gibt also nur die beiden folgenden stationären Punkte

$$(x, y) = (\sqrt{\ln(2)}, \sqrt{\ln(2)}), \quad (x, y) = (0, 0).$$

Jetzt müssen wir die Hesse-Matrix

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}$$

untersuchen. Mit den oben ausgerechneten zweiten Ableitungen finden wir

$(x, y)$	$H$	$\det(H)$
$(0, 0)$	$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$	3
$(\sqrt{\ln(2)}, \sqrt{\ln(2)})$	$\begin{pmatrix} 2(1 - \ln(2)) & -2(\ln(2) + 1) \\ -2(\ln(2) + 1) & 2(1 - \ln(2)) \end{pmatrix}$	$4((1 - \ln(2))^2 - (\ln(2) + 1)^2)$ $= -16\ln(2)^2$

Im ersten Fall ist die Determinante positiv und das Hurwitz-Kriterium zeigt:  $H$  ist positiv definit, der kritische Punkt ist ein lokales Minimum. Im zweiten Fall ist die Determinante negativ und das Hurwitz-Kriterium zeigt:  $H$  ist indefinit. Der kritische Punkt ist ein Sattelpunkt, weder lokales Maximum, noch lokales Minimum.

**Aufgabe 6.24** Sei

$$f(x, y) := \frac{1}{x} - \frac{1}{y} + 4x - 9y \quad (x \neq 0, y \neq 0).$$

Man bestimme alle lokalen Extrema von  $f$ .

**Aufgabe 6.25** Bestimmen Sie das absolute Minimum und Maximum von

$$f(x, y) := 2xy - x + y$$

auf  $M := [-2, 2] \times [-2, 2]$ .

**Aufgabe 6.26** Man bestimme Minimum und Maximum von

$$f(x, y) = e^{-(x^2+y^2)}(2x^2 + 3y^2)$$

auf

$$M = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 4\}.$$

**Aufgabe 6.27** Gegeben sei die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto 2x^2 + 2y^2 - 4xy - x^4 - y^4.$$

a) Untersuchen Sie  $f$  auf relative Maxima und Minima.

b) Bestimmen Sie die Tangentialebene an die Fläche

$$G = \{(x, y, z) \in \mathbb{R}^3 \mid z = f(x, y)\} \text{ im Punkt } P(1, 1, -2).$$

**Aufgabe 6.28** Gegeben sei die Funktion  $f : ]0, \infty[ \times ]0, \infty[ \rightarrow \mathbb{R}$ , mit

$$f(x, y) = xy + \frac{a}{x} + \frac{a}{y}, \quad a > 0.$$

Untersuchen Sie  $f$  auf relative Maxima und Minima.

**Aufgabe 6.29** Für  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , definiert durch  $f(x, y) = y^2 + \cos x$ , bestimme man die lokalen Extremstellen in  $\mathbb{R}^2$  und die absoluten Extremstellen in

$$M := \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq \pi/2, -1 \leq y \leq 1\}.$$

**Aufgabe 6.30** Gegeben sei die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) = x^3 + y^3 - 3x - 12y$ . Untersuchen Sie  $f$  auf relative Extrema; geben Sie gegebenenfalls die Art der Extrema an.

**Aufgabe 6.31** Gegeben sei die Funktion

$$f(x, y) = e^{-x^2-y^2}(x^2 + y^2 + xy + 3).$$

Man bestimme Infimum und Supremum der Bildmenge  $f(\mathbb{R}^2)$ .

**Aufgabe 6.32** Für  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f(x, y) = 2x^3 - xy^2 + 5x^2 + y^2$  bestimme man die relativen (= lokalen) Extrema.

**Aufgabe 6.33** Gegeben sei die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  durch

$$f(x, y) = (x^2 + 2y^2 - 1) \cdot (x^2 + y^2 - 1).$$

- Skizzieren Sie die Menge  $\{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$  und markieren Sie die Bereiche, wo  $f(x, y)$  positiv, bzw. negativ ist.
- Berechnen Sie den Gradienten und die Hesse-Matrix von  $f$ .
- Bestimmen Sie alle lokalen Extrema.

## 6.4 Die Kettenregel

Genau wie Funktionen einer Veränderlichen kann man auch Funktionen von mehreren Veränderlichen ineinander schachteln. Da gibt es allerdings eine Vielzahl von Möglichkeiten. Die einfachste davon ist folgende:

- Es ist  $f(x, y)$  eine Funktion von zwei Veränderlichen.
- Jede der beiden Veränderlichen ist eine Funktion  $x = x(t)$ , bzw.  $y = y(t)$  einer dritten Veränderlichen  $t$ .

Und die geschachtelte Funktion

$$f(x(t), y(t))$$

wollen wir nach  $t$  differenzieren. Die Physiker haben dafür eine sehr intuitive Notation:

$$df = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} \cdot dt + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt} \cdot dt.$$

Dabei stellen sie sich  $df$  und  $dt$  (in der Tradition von Leibniz) als infinitesimal kleine Größen vor. Ein normaler Mensch kann sich sowas nicht vorstellen. Mathematiker sind normale Menschen. Machen wir also präzise, was zu beweisen ist:

**Satz 6.10 (Kettenregel)** Die Funktion  $f(x, y)$  sei stetig differenzierbar auf einer offenen Menge  $U \subset \mathbb{R}^2$  und  $(x_0, y_0) \in U$  ein Punkt. Weiter seien

$$x(t) \text{ und } y(t) \text{ für } |t - t_0| < r$$

Funktionen mit

- $(x(t), y(t)) \in U$  für  $|t - t_0| < r$ ,
- $x(t_0) = x_0, y(t_0) = y_0$ ,
- $x(t)$  und  $y(t)$  sind differenzierbar in  $t = t_0$ .

Dann ist auch die geschachtelte Funktion

$$f(x(t), y(t)), \quad |t - t_0| < r$$

differenzierbar in  $t = t_0$  mit

$$\left. \frac{df(x(t), y(t))}{dt} \right|_{t=t_0} = \frac{\partial f}{\partial x}(x_0, y_0) \cdot \frac{dx}{dt}(t_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot \frac{dy}{dt}(t_0).$$

Bevor wir uns an den Beweis dieses Satzes wagen, erst zwei Beispiele:

**Beispiel 6.9 (Richtungsableitung)** Wir betrachten die Funktionen

$$x(t) = x_0 + t \cdot a, \quad y(t) = y_0 + t \cdot b, \quad t_0 = 0.$$

Dann lautet also unsere Formel

$$\left. \frac{df(x_0 + t \cdot a, y_0 + t \cdot b)}{dt} \right|_{t=0} = \frac{df}{dx}(x_0, y_0) \cdot a + \frac{df}{dy}(x_0, y_0) \cdot b.$$

Das ist genau die Formel aus Satz 6.3 b).

**Beispiel 6.10** Die Funktion  $f(x, y)$  sei stetig differenzierbar auf einer offenen Menge, die den Einheitskreis

$$x = \cos(t), y = \sin(t), t \in \mathbb{R}$$

enthält. Dann ist in einem Punkt

$$x_0 = \cos(t_0), y_0 = \sin(t_0)$$

des Einheitskreises

$$\left. \frac{df(\cos(t), \sin(t))}{dt} \right|_{t=t_0} = \frac{\partial f}{\partial x}(x_0, y_0) \cdot (-\sin(t_0)) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot \cos(t_0).$$

Beweis von Satz 6.10: Weil  $f$  total differenzierbar ist, können wir approximieren (Satz 6.2)

$$f(x(t), y(t)) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \cdot (x(t) - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot (y(t) - y_0) + \varphi(x(t), y(t))$$

mit

$$\lim_{\substack{(x, y) \rightarrow (x_0, y_0) \\ (x, y) \neq (x_0, y_0)}} \frac{\varphi(x, y)}{\|(x, y) - (x_0, y_0)\|} = 0 \quad (**).$$

Für  $t \neq t_0$  schreiben wir das etwas um:

$$\frac{f(x(t), y(t)) - f(x_0, y_0)}{t - t_0} = \frac{\partial f}{\partial x}(x_0, y_0) \cdot \frac{x(t) - x_0}{t - t_0} + \frac{\partial f}{\partial y}(x_0, y_0) \cdot \frac{y(t) - y_0}{t - t_0} + \frac{\varphi(x(t), y(t))}{t - t_0}.$$

Weil  $x(t)$  und  $y(t)$  in  $t = t_0$  differenzierbar sind, gilt

$$\lim_{t \rightarrow t_0} \frac{x(t) - x_0}{t - t_0} = \frac{dx}{dt}(t_0), \quad \lim_{t \rightarrow t_0} \frac{y(t) - y_0}{t - t_0} = \frac{dy}{dt}(t_0).$$

Die Behauptung steht da, wenn wir zeigen können

$$\lim_{t \rightarrow t_0} \frac{\varphi(x(t), y(t))}{t - t_0} = 0.$$

Das ist natürlich genau das Problem. Wir müssen etwas ausholen:

1) Wegen

$$\lim_{\substack{t \rightarrow t_0 \\ t \neq t_0}} \frac{x(t) - x_0}{t - t_0} = \frac{dx}{dt}(t_0), \quad \lim_{\substack{t \rightarrow t_0 \\ t \neq t_0}} \frac{y(t) - y_0}{t - t_0} = \frac{dy}{dt}(t_0),$$

gibt es ein  $\tau_1$  mit  $0 < \tau_1 < r$  so, dass für

$$|t - t_0| < \tau_1, \quad t \neq t_0$$

gilt

$$\left( \frac{x(t) - x_0}{t - t_0} \right)^2 \leq \left( \frac{dx}{dt}(t_0) \right)^2 + 1, \quad \left( \frac{y(t) - y_0}{t - t_0} \right)^2 \leq \left( \frac{dy}{dt}(t_0) \right)^2 + 1,$$

$$\frac{\| (x(t), y(t)) - (x_0, y_0) \|}{|t - t_0|} \leq \sqrt{\frac{dx}{dt}(t_0)^2 + \frac{dy}{dt}(t_0)^2 + 2} =: C.$$

2) Wir nehmen Zuflucht zu einem beliebigen  $\epsilon > 0$ . Wegen (\*\*) gibt es ein  $\delta(\epsilon) > 0$  so, dass aus

$$\| (x, y) - (x_0, y_0) \| < \delta$$

folgt  $(x, y) \in U$  und, falls  $(x, y) \neq (x_0, y_0)$ ,

$$\frac{|\varphi(x, y)|}{\| (x, y) - (x_0, y_0) \|} < \epsilon.$$

Weil  $x(t)$  und  $y(t)$  stetig in  $t_0$  sind, gibt es ein  $\tau_2 > 0$  mit

$$|t - t_0| < \tau_2 \quad \Rightarrow \quad \| (x(t), y(t)) - (x_0, y_0) \| < \delta \left( \frac{\epsilon}{C} \right).$$

Für

$$|t - t_0| < \min\{\tau_1, \tau_2\}$$

ist dann

- falls zufällig  $(x(t), y(t)) = (x_0, y_0)$  sein sollte

$$\frac{\varphi(x(t), y(t))}{|t - t_0|} = 0 < \epsilon,$$

- und falls  $(x(t), y(t)) \neq (x_0, y_0)$  ist

$$\frac{|\varphi(x(t), y(t))|}{|t - t_0|} = \underbrace{\frac{|\varphi(x(t), y(t))|}{\| (x(t), y(t)) - (x_0, y_0) \|}}_{< \epsilon/C} \cdot \underbrace{\frac{\| (x(t), y(t)) - (x_0, y_0) \|}{|t - t_0|}}_{< C} < \epsilon.$$

Der Beweis ist geschafft. □

Der wichtigste Spezialfall der Situation in Satz 6.10 sind Niveaulinien. Ist  $f(x, y)$  eine Funktion, so heißen die Teilmengen

$$\{(x, y) : f(x, y) = \text{const}\}$$

ihres Definitionsbereichs *Niveaulinien*. Ist etwa  $f(x, y) = x^2 - y^2$ , so sind dies die Kreise  $x^2 + y^2 = c$ . Man ist daran interessiert, diese Niveaulinien zu parametrisieren, etwa durch

$$x \mapsto (x, g(x)) \quad \text{mit} \quad f(x, g(x)) = c.$$

Im Fall der obigen Kreise sind solche Parametrisierungen z.B.

$$y = \pm \sqrt{c - x^2}.$$

Hier muss man allerdings  $x \neq \pm\sqrt{c}$ , d.h.,  $y \neq 0$  voraussetzen, weil sonst die Wurzel keine richtige Funktion ist. Ist  $f(x, g(x)) = \text{const}$ , so folgt mit der Kettenregel

$$0 = \frac{d}{dx} f(x, g(x)) = f_x + f_y \cdot g'(x).$$

Falls  $f_y \neq 0$  ist erhält man daraus

$$g'(x) = -\frac{f_x}{f_y}.$$

Im Fall der Kreise lautet diese Gleichung

$$0 = 2x + 2y \cdot g'(x), \quad g'(x) = -\frac{x}{y}.$$

Hier kennt man die Funktion  $y(x) = g(x) = \sqrt{c - x^2}$  und ihre Ableitung

$$g'(x) = \frac{-2x}{2\sqrt{c - x^2}} = -\frac{x}{y}.$$

Aber man kann die Kettenregel auch anwenden, wenn man die Funktion  $y = g(x)$  nicht explizit kennt.

**Beispiel 6.11** *Es sei*

$$f(x, y) := x^3 + y^3 - 3xy.$$

Die Gleichung  $f(x, y) = 0$  hat in  $y$  den Grad 3. Solche Gleichungen kann man im Prinzip auflösen, aber das ist sehr aufwendig und unübersichtlich. Wir betrachten etwa den Punkt

$$(x_0, y_0) = (\sqrt[3]{2}, \sqrt[3]{4}).$$

Hier ist

$$f(x_0, y_0) = 2 + 4 - 3 \cdot \sqrt[3]{8} = 6 - 6 = 0.$$

Die Niveaulinie  $f(x, y) = 0$  zur Konstanten  $c = 0$  geht durch diesen Punkt. Wenn  $y = g(x)$  eine Parametrisierung dieser Niveaulinie ist, können wir ableiten

$$\begin{aligned} 0 &= \frac{d}{dx} f(x, g(x)) \\ &= f_x + f_y \cdot g'(x) \\ &= 3x^2 - 3y + (3y^2 - 3x) \cdot g'(x) \\ g'(x) &= -\frac{x^2 - y}{y^2 - x} \text{ falls } y^2 \neq x. \end{aligned}$$

Im Punkt  $(x_0, y_0)$  ist

$$y^2 - x = \sqrt[3]{16} - \sqrt[3]{2} > 0, \quad x^2 - y = \sqrt[3]{4} - \sqrt[3]{4} = 0.$$

Ohne die Funktion  $g(x)$  zu kennen, sehen wir: In  $x_0$  hat sie einen kritischen Punkt. Aber wir können die Gleichung

$$x^2 - y + (y^2 - x) \cdot g'(x) = x^2 - g(x) + (g^2(x) - x) \cdot g'(x) = 0$$

noch weiter ableiten

$$2x - g'(x) + (2g(x)g'(x) - 1) \cdot g'(x) + (g^2(x) - x) \cdot g''(x) = 0.$$

Für  $x = x_0$  ist  $g'(x) = 0$  und dort finden wir

$$2x_0 + (y_0^2 - x_0) \cdot g''(x_0) = 0.$$

Wir wissen schon  $y_0^2 - x_0 > 0$ . Damit folgt

$$g''(x_0) = -\frac{2x_0}{y_0^2 - x_0} < 0.$$

Wir sehen, in  $x_0$  hat die Funktion  $g(x)$  ein Maximum. Und das sehen wir, ohne die Funktion  $g(x)$  überhaupt zu kennen!

Die Kettenregel kann man nun in verschiedene Richtungen verallgemeinern. Zuerst eine Verallgemeinerung, die keine neuen Begriffe, sondern nur etwas Schreibaufwand erfordert:

Wir betrachten eine Funktion  $f(x_1, \dots, x_n)$  von  $n > 2$  Veränderlichen, wo die Variablen

$$x_1 = x_1(t), \dots, x_n = x_n(t),$$

wieder differenzierbare Funktionen einer neuen Variablen  $t$  sind. Der Beweis von Satz 6.10 verallgemeinert sich wörtlich und liefert die Kettenregel in der Form

$$\frac{df}{dt}(x_1(t), \dots, x_n(t)) = \frac{\partial f}{\partial x_1} \cdot \frac{dx_1}{dt} + \dots + \frac{\partial f}{\partial x_n} \cdot \frac{dx_n}{dt}.$$

**Beispiel 6.12** Sei etwa

$$f(x_1, \dots, x_n) = r(x_1, \dots, x_n) = \sqrt{x_1^2 + \dots + x_n^2}$$

der Abstand vom Nullpunkt. Die partiellen Ableitungen sind

$$\frac{\partial r}{\partial x_\nu} = \frac{1}{2\sqrt{x_1^2 + \dots + x_n^2}} \cdot 2x_\nu = \frac{x_\nu}{r}.$$

Bewegt sich ein Vektor

$$\mathbf{x}(t) = (x_1(t), \dots, x_n(t)) \in \mathbb{R}^n$$

in Abhängigkeit von  $t$ , so ist die Ableitung seines Abstands vom Nullpunkt

$$\frac{d}{dt} \|\mathbf{x}(t)\| = \sum_{\nu=1}^n \frac{\partial r}{\partial x_\nu} \cdot \frac{dx_\nu}{dt} = \sum_{\nu=1}^n \frac{x_\nu}{r} \cdot \frac{dx_\nu}{dt}.$$

Es gibt aber auch eine sehr weitreichende begriffliche Verallgemeinerung: Die Ableitung von Abbildungen  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^k$ . Sind  $\Phi_1(\mathbf{x}), \dots, \Phi_p(\mathbf{x})$  die Koordinaten von  $\Phi(\mathbf{x})$  im  $\mathbb{R}^p$ , so ist

$$\Phi(\mathbf{x}) = \begin{pmatrix} \Phi_1(x_1, \dots, x_n) \\ \vdots \\ \Phi_p(x_1, \dots, x_n) \end{pmatrix}.$$

**Definition 6.6** Die Abbildung  $\Phi$  heißt stetig differenzierbar, wenn alle Komponentenfunktionen  $\Phi_1, \dots, \Phi_p$  stetig differenzierbar bezüglich  $x_1, \dots, x_n$  sind. Die  $p \times n$ -Matrix der partiellen Ableitungen

$$J_\Phi := \begin{pmatrix} \frac{\partial \Phi_1}{\partial x_1} & \dots & \frac{\partial \Phi_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \Phi_p}{\partial x_1} & \dots & \frac{\partial \Phi_p}{\partial x_n} \end{pmatrix}$$



heißt dann Jacobi-Matrix oder Funktionalmatrix von  $\Phi$ . Die lineare Abbildung

$$\Phi'(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^k, \quad \mathbf{h} \mapsto J_\Phi(\mathbf{x}) \cdot \mathbf{h}$$

heißt die Ableitung der Abbildung  $\Phi$  im Punkt  $\mathbf{x}$ .

**Beispiel 6.13** Die Transformation in Polarkoordinaten

$$\mathbb{R}^2 \ni \begin{pmatrix} r \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cdot \cos(\varphi) \\ r \cdot \sin(\varphi) \end{pmatrix} \in \mathbb{R}^2$$

ist eine solche Abbildung  $\Phi$ . Ihre Jacobi-Matrix ist

$$\begin{pmatrix} \cos(\varphi) & -r \cdot \sin(\varphi) \\ \sin(\varphi) & r \cdot \cos(\varphi) \end{pmatrix}.$$

**Beispiel 6.14 (Trivial)** Die Identität  $\Phi = id : \mathbb{R}^n \rightarrow \mathbb{R}^n$  hat die Komponentenfunktionen

$$\Phi_1(\mathbf{x}) = x_1, \dots, \Phi_n(\mathbf{x}) = x_n.$$

Ihre Funktionalmatrix ist deswegen

$$J_{id} = \left( \frac{\partial x_\mu}{\partial x_\nu} \right) = (\delta_{\mu,\nu}) = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix},$$

die Einheitsmatrix  $\mathbb{1}_n$ .

So, und jetzt werden wir Abbildungen

$$\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^p, \quad \Psi : \mathbb{R}^p \rightarrow \mathbb{R}^q$$

ineinanderschachteln. Da müssen wir ziemlich genau mit der Notation für die Koordinaten buchhalten:

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{\Phi} & \mathbb{R}^p & \xrightarrow{\Psi} & \mathbb{R}^q \\ x_1 & & y_1(x_1, \dots, x_n) & & z_1(y_1, \dots, y_p) \\ \vdots & & \vdots & & \vdots \\ x_n & & y_p(x_1, \dots, x_n) & & z_q(y_1, \dots, y_p) \end{array}$$

Die Komponentenfunktionen der zusammengesetzten Abbildung

$$\Psi \circ \Phi : \mathbb{R}^n \rightarrow \mathbb{R}^q$$

sind

$$z_k(y_1(x_1, \dots, x_n), \dots, y_p(x_1, \dots, x_n)), \quad k = 1, \dots, q.$$

Wenn  $\Phi$  und  $\Psi$  stetig differenzierbar sind, dann folgt aus Satz 6.10, dass diese Komponentenfunktionen stetig nach  $x_1, \dots, x_n$  differenzierbar sind mit

$$\frac{\partial z_k}{\partial x_i} = \sum_{j=1}^p \frac{\partial z_k}{\partial y_j} \cdot \frac{\partial y_j}{\partial x_i}.$$

Für die Funktionalmatrizen

$$J_{\Phi} = \left( \frac{\partial y_j}{\partial x_i} \right), \quad J_{\Psi} = \left( \frac{\partial z_k}{\partial y_j} \right), \quad J_{\Psi \circ \Phi} = \left( \frac{\partial z_k}{\partial x_i} \right)$$

Bedeutet das:

$$J_{\Psi \circ \Phi} = \left( \frac{\partial z_k}{\partial y_j} \right)_{k,j} \cdot \left( \frac{\partial y_j}{\partial x_i} \right)_{j,i} = J_{\Psi} \cdot J_{\Phi}.$$

Damit haben wir bewiesen:

**Satz 6.11 (Kettenregel für Abbildungen)** Die Abbildungen  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^p$  und  $\Psi : \mathbb{R}^p \rightarrow \mathbb{R}^q$  wie oben seien stetig differenzierbar. Dann ist auch die Abbildung  $\Psi \circ \Phi : \mathbb{R}^n \rightarrow \mathbb{R}^q$  stetig differenzierbar und ihre Funktionalmatrix im Punkt  $\mathbf{x} \in \mathbb{R}^n$  ist das Matrizenprodukt

$$J_{\Psi \circ \Phi}(\mathbf{x}) = J_{\Psi}(\Phi(\mathbf{x})) \cdot J_{\Phi}(\mathbf{x}).$$

**Beispiel 6.15 (Ableitung der Umkehrabbildung)** Die stetig differenzierbare Abbildung  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  sei bijektiv mit stetig differenzierbarer Umkehrabbildung  $\Psi = \Phi^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Dann ist also

$$\Psi \circ \Phi = id_{\mathbb{R}^n}, \quad \Phi \circ \Psi = id_{\mathbb{R}^n}.$$

Mit Beispiel 6.14 sehen wir

$$J_{\Psi}(\Phi(\mathbf{x})) \cdot J_{\Phi}(\mathbf{x}) = \mathbb{1}_n, \quad J_{\Phi}(\mathbf{x}) \cdot J_{\Psi}(\Phi(\mathbf{x})) = \mathbb{1}_n.$$

Die Ableitung der Umkehrabbildung  $\Psi = \Phi^{-1}$  im Punkt  $\Phi(\mathbf{x}) = \mathbf{y}$  ist die inverse Matrix

$$\left( \Phi^{-1} \right)'(\mathbf{y}) = \left( \Phi'(\mathbf{x}) \right)^{-1}.$$

Das ist die Verallgemeinerung auf mehrere Veränderliche der Formel für die Ableitung der Umkehrfunktion

$$\left( f^{-1} \right)'(y) = \frac{1}{f'(x)}, \quad y = f(x).$$

**Beispiel 6.16** Der einfachste nicht ganz triviale Spezialfall für Beispiel 6.15 ist die Transformation

$$\Phi : \begin{pmatrix} r \\ \varphi \end{pmatrix} \rightarrow \begin{pmatrix} x = r \cdot \cos(\varphi) \\ y = r \cdot \sin(\varphi) \end{pmatrix}$$

von euklidischen Koordinaten  $(x, y)$  in Polarkoordinaten  $(r, \varphi)$ . Die zugehörige Funktionalmatrix

$$J_{\Phi} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos(\varphi) & -r \cdot \sin(\varphi) \\ \sin(\varphi) & r \cdot \cos(\varphi) \end{pmatrix}$$

haben wir schon einmal ausgerechnet (Beispiel 6.13). Ihre inverse Matrix ist

$$J_{\Phi}^{-1} = \frac{1}{r} \begin{pmatrix} r \cdot \cos(\varphi) & r \cdot \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

Die Umkehrabbildung  $\Phi^{-1} : (x, y) \rightarrow (r, \varphi)$  ist nicht ganz einfach anzugeben. Einfach ist

$$r = \sqrt{x^2 + y^2}.$$

Aber - vor allem wegen der Mehrdeutigkeit des Winkels  $\varphi$  - ist dieser als Funktion von  $x$  und  $y$  nicht so einfach anzugeben. Wir beschränken uns auf die rechte Halbebene  $x > 0$ . Da ist

$$\frac{y}{x} = \frac{r \cdot \cos(\varphi)}{r \cdot \sin(\varphi)} = \tan(\varphi), \quad \varphi = \arctan\left(\frac{y}{x}\right).$$

Damit wird für  $x > 0$

$$\Phi^{-1} : \begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} r = \sqrt{x^2 + y^2} \\ \varphi = \arctan\left(\frac{y}{x}\right) \end{pmatrix}.$$

Die partiellen Ableitungen

$$\frac{\partial r}{\partial x} = \frac{x}{r}, \quad \frac{\partial r}{\partial y} = \frac{y}{r}$$

haben wir in Beispiel 6.12 schon mal ausgerechnet. Die partiellen Ableitungen von  $\varphi$  nach  $x$  und  $y$  erhalten wir als

$$\frac{\partial \varphi}{\partial x} = \frac{\partial}{\partial x} \arctan\left(\frac{y}{x}\right) = \frac{1}{1 + (y/x)^2} \cdot -\frac{y}{x^2} = -\frac{y}{x^2 + y^2} = -\frac{y}{r^2},$$

$$\frac{\partial \varphi}{\partial y} = \frac{\partial}{\partial y} \arctan\left(\frac{y}{x}\right) = \frac{1}{1 + (y/x)^2} \cdot \frac{1}{x} = \frac{x}{x^2 + y^2} = \frac{x}{r^2}.$$

Und die Funktionalmatrix der Umkehrabbildung wird

$$J_{\Phi^{-1}} = \begin{pmatrix} \frac{x}{r} & \frac{y}{r} \\ -\frac{y}{r^2} & \frac{x}{r^2} \end{pmatrix} = \frac{1}{r} \begin{pmatrix} x & y \\ -\frac{y}{r} & \frac{x}{r} \end{pmatrix}.$$

Das sieht etwas anders aus, als das oben erhaltene Ergebnis für die Umkehrmatrix  $(J_{\Phi})^{-1}$ . Das liegt daran, dass die Matrix oben in  $r$  und  $\varphi$  ausgedrückt ist. Wenn wir in unser letztes Ergebnis

$$x = r \cdot \cos(\varphi), \quad y = r \cdot \sin(\varphi)$$

einsetzen, stellen sich beide Matrizen als identisch heraus.

**Aufgabe 6.34** Auf der Halbebene

$$\{(x, y) \in \mathbb{R}^2 : x + y > 0\}$$

sei  $y(x)$  eine Niveau-Linie der Funktion

$$f(x, y) := e^{x-y} - (x+y)^2.$$

Bestimmen Sie  $y'(x)$  als Funktion von  $x$  und  $y$ .

**Aufgabe 6.35** Bestimmen Sie die Funktionalmatrix  $J_\Phi$  der Abbildung

$$\Phi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} u = x^2 - y^2 \\ v = 2xy \end{pmatrix}$$

sowie deren Inverse  $J_\Phi^{-1}$ .

**Aufgabe 6.36 (Elliptische Koordinaten)** Betrachtet werde die Abbildung

$$\Phi : \begin{pmatrix} t \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} x = \cos(\varphi) \cdot \cosh(t) \\ y = \sin(\varphi) \cdot \sinh(t) \end{pmatrix}.$$

- a) Bestimmen Sie die Funktionalmatrix dieser Abbildung, sowie für  $(t, \varphi) \neq (0, 0)$  deren Inverse.  
b) Zeigen Sie, dass die Kurven  $t = \text{const}$  Ellipsen und die Kurven  $\varphi = \text{const}$  Hyperbeln sind.

## Index

- abgeschlossene Menge, 10
- Ableitung
  - partielle, 23
- Ableitung einer Abbildung, 49
- Abstand, 7
  
- beschränkte Menge, 15
- Bildmenge, 19
  
- Cauchy-Schwarz-Ungleichung, 6
  
- Differenzierbarkeit
  - totale, 27
- Dreiecksungleichung, 7
  
- Extremum
  - lokales, 23
  
- Funktionalmatrix, 49
  
- Gradient, 23
  
- Hesse-Matrix, 31
  
- indefinit, 37
  
- Jacobi-Matrix, 49
  
- Kettenregel, 44
  - für Abbildungen, 50
- kompakte Menge, 16
- konvergente Folge von Vektoren, 13
- kritischer Punkt, 24
- Kugel, 8
  
- Maximum
  - lokales, 23
- Metrik, 7
- metrischer Raum, 7
- Minimum
  - lokales, 23
  
- negativ definit, 37
- Niveaulinie, 46
- Norm, 6
  
- offene Menge, 8
  
- partielle Ableitungen, 23
- positiv definit, 37
- Punkt
  - kritischer, 24
  - stationärer, 24
  
- Rand, 11
  - einer Kugel, 8
- Richtungsableitung, 28
  
- Sattelpunkt, 24
- Skalarprodukt
  - euklidisches, 6
- stationärer Punkt, 24
- Stetigkeit, 16
  
- Tangential-Ebene, 26
- Taylor-Formel, 33
- Taylor-Reihe, 33
- totale Differenzierbarkeit, 27
  
- Urbildmenge, 19
  
- Vektorfeld, 23
  
- zweite partielle Ableitungen, 31