

# Inhaltsverzeichnis

0.1	Stetigkeit in einer Dimension . . . . .	3
0.2	Zwei Sonderfälle . . . . .	3
<b>1</b>	<b>Differentialrechnung in höheren Dimensionen</b>	<b>5</b>
1.1	Topologie . . . . .	5
1.1.1	Korollar . . . . .	5
1.1.2	Konvention . . . . .	6
1.1.3	Definition der $\varepsilon$ -Umgebung . . . . .	6
1.1.4	Topologische Grundbegriffe . . . . .	6
1.1.5	Definition von offen und abgeschlossen . . . . .	6
1.1.6	Beispiele . . . . .	6
1.1.7	Satz . . . . .	7
1.1.8	Satz . . . . .	7
1.1.9	Satz . . . . .	8
1.1.10	Definition von beschränkt und kompakt . . . . .	8
1.2	Folgen . . . . .	8
1.2.1	Definition von Konvergenz und Beschränktheit . . . . .	8
1.2.2	Bemerkung . . . . .	8
1.2.3	Satz von Bolzano Weierstraß . . . . .	9
1.2.4	Abschließende Bemerkungen . . . . .	9
1.3	Funktionsgrenzwerte und Stetigkeit . . . . .	9
1.3.1	Definition . . . . .	9
1.3.2	Definition Grenzwert/Limes . . . . .	9
1.3.3	Korollar . . . . .	10
1.3.4	Beispiel . . . . .	10
1.3.5	Lemma Folgenkriterium . . . . .	10
1.3.6	Satz zu Grenzwerte verketteter Funktionen . . . . .	10
1.3.7	Beispiel . . . . .	10
1.3.8	Definition der Stetigkeit . . . . .	11
1.3.9	Bemerkung . . . . .	11
1.4	Partielle Ableitungen, Richtungsableitungen . . . . .	11
1.4.1	Definition der partiellen Ableitung . . . . .	11
1.4.2	Beispiel . . . . .	12
1.4.3	Definition der Richtungsableitung . . . . .	12
1.5	Total Differenzierbarkeit . . . . .	12
1.5.1	Definition der totalen Differenzierbarkeit . . . . .	12
1.5.2	Beispiele . . . . .	13
1.5.3	Satz . . . . .	13
1.5.4	Satz . . . . .	13
1.5.5	Bemerkung . . . . .	15

1.5.6	Satz zur Kettenregel . . . . .	16
1.6	Lokale Extremstellen und Mittelwertsätze . . . . .	16
1.6.1	Definition lokale/globale Extremstellen . . . . .	17
1.6.2	Satz zur notwendigen Bedingung für eine lokale Extremstelle . . . . .	17
1.6.3	Definition des kritischen Punktes . . . . .	17
1.6.4	Mittelwertsatz . . . . .	17
1.6.5	Definition eines Gebiets . . . . .	18
1.6.6	Bemerkungen zu Gebieten . . . . .	18
1.6.7	Satz . . . . .	18
1.6.8	Definition partieller Ableitungen $r$ 'ter Ordnung . . . . .	18
1.6.9	Definition der Hessematrix . . . . .	19
1.6.10	Beispiele . . . . .	19
1.6.11	Satz von Schwarz . . . . .	19
1.6.12	Satz von Taylor mit quadratischem Restglied . . . . .	19
1.6.13	Definition von Definitheit . . . . .	20
1.6.14	Beispiele . . . . .	20
1.6.15	Satz zum Hauptminorenkriterium . . . . .	21
1.6.16	Satz über die hinreichenden Bedingungen für lokale Extremstellen . . . . .	21
1.6.17	Definition der Sattelpunkte . . . . .	22
1.6.18	Satz für den Fall $n = 2$ . . . . .	22
1.6.19	Beispiel . . . . .	22
1.7	Extremstellen unter Nebenbedingungen und implizite Funktionen . . . . .	23
1.7.1	Spezialfälle . . . . .	24
1.7.2	Bemerkung . . . . .	24
1.7.3	Satz über die Umkehrfunktion . . . . .	24

# Einführung

## 0.1 Stetigkeit in einer Dimension

$f$  ist stetig in  $x_0$

$$\Leftrightarrow \lim_{x \rightarrow x_0} f(x) = f(x_0)$$

$$\Leftrightarrow \forall (x_n) \text{ mit } \lim_{n \rightarrow \infty} x_n = x_0 \text{ gilt } \lim_{n \rightarrow \infty} f(x_n) = f(x_0)$$

$$\Leftrightarrow \forall \varepsilon > 0 \quad \exists \delta \quad \text{mit} \quad |f(x) - f(x_0)| < \varepsilon \quad \forall x \in (x_0 - \delta, x_0 + \delta)$$

*Bemerkung:* Der Grenzwert von Funktionen ist über den Grenzwert von Folgen definiert und kann auch nur so überprüft werden.

## 0.2 Zwei Sonderfälle

### Skalarfeld

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

Visualisierung durch Höhenlinien:  $H_c := \{x \in \mathbb{R}^n : f(x) = c\}$

Beispiel:  $f(x, y) = x^2 + y^2$

### Vektorfeld

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

Beispiel:  $f(x, y) = \begin{pmatrix} x \\ y \end{pmatrix}$



# Kapitel 1

## Differentialrechnung in höheren Dimensionen

### 1.1 Topologie

#### Skalarprodukt

Definition:  $\langle x, y \rangle := x^\top y = \sum_{k=1}^n x_k y_k$  für  $x, y \in \mathbb{R}^n$

#### Euklidische Norm

Definition:  $\|x\|_2 := \sqrt{\langle x, x \rangle} = \sqrt{\sum_{k=1}^n x_k^2}$

#### 1.1.1 Korollar

Sei  $x \in \mathbb{R}^n$  mit  $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

1.

$$\max_{1 \leq k \leq n} |x_k| \leq \|x\| \leq \sqrt{n} \max_{1 \leq k \leq n} |x_k|$$

2. Cauchy-Schwarz-Ungleichung:

$$\forall x, y \in \mathbb{R}^n \quad : \quad |\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

Begründung (nicht Beweis!) durch alternative Definition:  $\langle x, y \rangle = \|x\| \cdot \|y\| \underbrace{\cos \alpha}_{\leq 1}$

Dabei ist  $\alpha$  der Winkel der zwischen  $x$  und  $y$  eingeschlossen wird.

Daraus folgt:

$$|\langle x, y \rangle| = \|x\| \cdot \|y\| \Leftrightarrow x, y \text{ sind lin. unabhängig : } x = \lambda y \text{ oder } y = \lambda x \text{ für } \lambda \in \mathbb{R}$$

3.  $\|\cdot\|$  ist eine Norm. Eine Norm hat folgende Eigenschaften:

- (i)  $\|x\| \geq 0$  und  $\|x\| = 0 \Leftrightarrow x = 0$
- (ii)  $\|\lambda x\| = |\lambda| \cdot \|x\|$
- (iii)  $\|x + y\| \leq \|x\| + \|y\|$  Dreiecksungleichung

### 1.1.2 Konvention

Für  $A \subset \mathbb{R}^n$  gilt für das Komplement  $A^c = \mathbb{R}^n \setminus A$

### 1.1.3 Definition der $\varepsilon$ -Umgebung

Sei  $x_0 \in \mathbb{R}^n$  und  $\varepsilon > 0$ , dann gilt für die  $\varepsilon$ -Umgebung  $U_\varepsilon(x_0)$  von  $x_0$ :

$$U_\varepsilon(x_0) := \{x \in \mathbb{R}^n : \|x - x_0\| < \varepsilon\}$$

*Bemerkung:* Die punktierte  $\varepsilon$ -Umgebung ist definiert als:  $\dot{U}_\varepsilon = U_\varepsilon(a) \setminus \{a\}$

### 1.1.4 Topologische Grundbegriffe

Sei  $A \subset \mathbb{R}^n$ , dann heißt ein Punkt  $x_0 \in \mathbb{R}^n$

- (i) ein **innerer Punkt**, wenn gilt  $\exists \varepsilon > 0$  mit  $U_\varepsilon(x_0) \subset A$   
Menge aller inneren Punkte:  $\mathring{A} = \{x \in \mathbb{R}^n : \exists \varepsilon > 0 \text{ mit } U_\varepsilon(x) \subset A\}$
- (ii) ein **Berührungspunkt**, wenn  $\forall \varepsilon > 0$  gilt  $U_\varepsilon(x_0) \cap A \neq \emptyset$   
**abgeschlossene Hülle:**  $\bar{A} = \{x \in \mathbb{R}^n : \forall \varepsilon > 0 \text{ gilt } U_\varepsilon(x_0) \neq \emptyset\}$
- (iii) ein **Häufungspunkt**, wenn  $\forall \varepsilon > 0$  gilt  $(U_\varepsilon(x_0) \setminus \{x_0\}) \cap A \neq \emptyset$   
Die Menge aller Häufungspunkte wird mit  $A'$  bezeichnet.
- (iv) ein **Randpunkt**, wenn  $\forall \varepsilon > 0$  gilt  $U_\varepsilon(x_0) \cap A \neq \emptyset$  und  $U_\varepsilon(x_0) \cap A^c \neq \emptyset$   
Menge aller Randpunkte oder auch **Rand** von  $A$  wird mit  $\partial A$  bezeichnet.

#### Korollar

- (i)  $\mathring{A} \subset A$
- (ii)  $\mathring{A} \subset \bar{A}$
- (iii)  $\partial A \subset \bar{A}$
- (iv)  $\bar{A} = \mathring{A} \cup \partial A$
- (v)  $\bar{A} = A \cup \partial A$  (schwächere Aussage als (iv))

### 1.1.5 Definition von offen und abgeschlossen

Eine Menge  $A \subset \mathbb{R}^n$  heißt

- (i) **offen**, wenn  $A = \mathring{A}$  gilt ( $A$  besteht nur aus inneren Punkten)
- (ii) **abgeschlossen**, wenn  $\partial A \subset A$  gilt (wenn der Rand in der Menge enthalten ist)

### 1.1.6 Beispiele

1. Jede  $\varepsilon$ -Umgebung  $U_\varepsilon(x_0 \in \mathbb{R}^n)$  ist offen

2. Sei  $I \subset \mathbb{R}$ , dann gilt

- (i)  $I$  ist offen, wenn  $I = (a, b)$  mit  $-\infty \leq a \leq b \leq \infty$   
für  $a = b$  gilt  $I = \emptyset$  mit  $I$  offen  
und für  $a = -\infty, b = \infty$  ist  $I$  auch offen

- (ii)  $I$  ist abgeschlossen, wenn  $I = [a, b]$  mit  $a, b \in \mathbb{R}$   
 oder  $I = (-\infty, b]$  oder  $I = [a, \infty)$  oder  $I = (-\infty, \infty) = \mathbb{R}$

(die reellen Zahlen sind offen und abgeschlossen zugleich)

### 1.1.7 Satz

für  $A \subset \mathbb{R}^n$  sind folgenden Aussagen äquivalent:

- (i)  $A$  ist abgeschlossen  $A = \overline{A}$
- (ii)  $A$  enthält alle Häufungspunkte,  $A' \subset A$
- (iii)  $A$  enthält alle Randpunkte,  $\partial A \subset A$
- (iv)  $A^c$  ist offen

### 1.1.8 Satz

- (i)  $\emptyset$  und  $\mathbb{R}^n$  sind offen.
- (ii) Die Vereinigung beliebig vieler offene Mengen ist offen:

$$\bigcup_{j \in J} (O_j \text{ offen}) = O \text{ offen}$$

- (iii) Der Durchschnitt endlich vieler offener Mengen ist offen:

$$\bigcap_{j=1}^n (O_j \text{ offen}) = O \text{ offen}$$

*Bemerkung:* Für unendlich viele offene Mengen gilt dies nicht immer:

$$\bigcap_{k=1}^{\infty} \left(-\frac{1}{k}, \frac{1}{k}\right) = (-1, 1) \cap \left(-\frac{1}{2}, \frac{1}{2}\right) \cap \left(-\frac{1}{3}, \frac{1}{3}\right) \cap \dots = \{0\} \text{ abgeschlossen}$$

### Beispiel

Seien  $A_1, A_2$  zwei abgeschlossene Mengen, dann gilt

- (i)  $A_1 \cup A_2$  ist abgeschlossen

**Beweisidee:**  $A_1$  ist abgeschlossen  $\Rightarrow A_1^c$  ist offen

$$\begin{aligned} (A_1 \cup A_2)^c &\stackrel{\text{De Morgan}}{=} \underbrace{A_1^c}_{\text{offen}} \cap \underbrace{A_2^c}_{\text{offen}} \text{ ist offen wegen Satz 1.1.8} \\ ((A_1 \cup A_2)^c)^c &= A_1 \cup A_2 \text{ ist abgeschlossen} \end{aligned}$$

**1.1.9 Satz**

- (i)  $\emptyset$  und  $\mathbb{R}^n$  sind abgeschlossen.
- (ii) Der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen:

$$\bigcap_{j \in J} (A_j \text{ abgeschlossen}) = A \text{ abgeschlossen}$$

- (iii) Die Vereinigung endlich vieler abgeschlossenen Mengen ist abgeschlossen:

$$\bigcup_{j=1}^n (A_j \text{ abgeschlossen}) = A \text{ abgeschlossen}$$

*Bemerkung:* Für unendlich viele abgeschlossene Mengen gilt dies nicht immer:

$$\bigcup_{k=1}^{\infty} \left[ -1 + \frac{1}{n}, 1 - \frac{1}{n} \right] = \{0\} \cup \left[ -\frac{1}{2}, \frac{1}{2} \right] \cup \left[ -\frac{2}{3}, \frac{2}{3} \right] \cup \dots = (-1, 1) \text{ offen}$$

**1.1.10 Definition von beschränkt und kompakt**

Eine Menge  $A \subset \mathbb{R}^n$  heißt:

- (i) **beschränkt** wenn  $\exists c > 0$  mit  $\|x\| < c \quad \forall x \in A$
- (ii) **kompakt**, wenn  $A$  abgeschlossen und beschränkt ist.

**1.2 Folgen****1.2.1 Definition von Konvergenz und Beschränktheit**

Eine Folge  $(a_k)_{k=1}^{\infty}$  heißt

- (i) **konvergent**, wenn gilt

$$\exists a \in \mathbb{R}^n \quad \text{mit} \quad \forall \varepsilon > 0 \quad \exists N(\varepsilon) : \quad \|a_k - a\| < \varepsilon \quad \forall k \geq N(\varepsilon)$$

Dann ist  $a$  der Grenzwert der Folge:

$$a = \lim_{k \rightarrow \infty} a_k \quad \text{oder} \quad a_k \xrightarrow{k \rightarrow \infty} a$$

- (ii) **beschränkt**, wenn  $\exists c > 0$  mit  $\|a_k\| < c \quad \forall k$

**1.2.2 Bemerkung**

Wenn eine Folge  $(a_k) = \begin{pmatrix} a_1^{(k)} \\ \vdots \\ a_n^{(k)} \end{pmatrix} \in \mathbb{R}^n$  konvergiert, so gilt

- (i)  $\Leftrightarrow$  jede Komponente  $a_1^{(k)}, \dots, a_n^{(k)}$  konvergiert:

$$\lim_{k \rightarrow \infty} a_k = a \quad \Leftrightarrow \quad \lim_{k \rightarrow \infty} a_i^{(k)} = a_i \quad \text{für } i = 1, \dots, n$$



(ii)  $\Leftrightarrow (a_k)$  erfüllt das **Cauchy-Kriterium**:

$$\forall \varepsilon > 0 \quad \exists N(\varepsilon) : \quad \|a_k - a_l\| < \varepsilon \quad \forall k, l \geq N(\varepsilon)$$

(iii)  $\Leftrightarrow$  jede Teilfolge von  $(a_k)$  konvergiert gegen  $a$ :  $a_{l_k} \xrightarrow{k \rightarrow \infty} a$  für  $l_1 \geq 1, l_2 \geq 2, \dots$

(iv) der Grenzwert  $a$  ist eindeutig.

### 1.2.3 Satz von Bolzano Weierstraß

Jede beschränkte Folge im  $\mathbb{R}^n$  besitzt eine konvergente Teilfolge.

#### Beispiele

(i)  $n = 1$ : Sei  $A \leq (a_k) \leq B \quad \forall k$ . Konstruiert man eine neue Schranke mit  $\frac{A+B}{2}$  so liegen wiederum  $\infty$  viele Elemente in der oberen und/oder unteren Hälfte.

(ii) Sei  $(a_k) = \begin{pmatrix} (x_k) \\ (y_k) \end{pmatrix}$  eine beschränkte Folge im  $\mathbb{R}^2$   
 $\Rightarrow (x_k), (y_k)$  sind beschränkte Folgen  
Satz von Bolzano Weierstraß  
 $\Rightarrow \exists (x_k), (y_k)$  sind konvergent

### 1.2.4 Abschließende Bemerkungen

(i) Grenzwert Rechenregeln können aus dem  $\mathbb{R}$  für  $\mathbb{R}^n$  übernommen werden.

$$\text{z.B. } a_k \xrightarrow{k \rightarrow \infty} a, \quad b_k \xrightarrow{k \rightarrow \infty} b \quad \Rightarrow \quad a_k^\top b_k \xrightarrow{k \rightarrow \infty} a^\top b$$

(ii) Es gibt viele Zusammenhänge zwischen den Eigenschaften von Folgen und den topologischen Eigenschaften von Mengen.

z.B. Sei  $A \subset \mathbb{R}^n$  und  $a \in \mathbb{R}^n$  ein Häufungspunkt

$$\Leftrightarrow \exists (a_k)_{k=1}^\infty \text{ mit } a_k \in A \setminus \{a\} \quad \forall k \quad \text{und} \quad a_k \xrightarrow{k \rightarrow \infty} a$$

## 1.3 Funktionsgrenzwerte und Stetigkeit

### 1.3.1 Definition

Eine Funktion  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  nennt man eine Funktion mit  $n$ -Veränderlichen.

$$f(x_1, \dots, x_n) = f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix} \quad \text{mit} \quad f_1, \dots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}$$

### 1.3.2 Definition Grenzwert/Limes

Sei  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  und  $a \in \overline{A}$ . Ein  $b \in \mathbb{R}^m$  heißt Grenzwert von  $f$  für  $x \rightarrow a$ , wenn gilt:

$$\forall \varepsilon > 0 \quad \exists \delta(\varepsilon) > 0 : \quad \|f(x) - b\| < \varepsilon \quad \forall x \in \dot{U}_{\delta(\varepsilon)}(a) \cap A$$

*Bemerkung:* Die Funktion  $f$  muss in  $a$  nicht stetig sein, so kann z.B. gelten:

$$\lim_{x \rightarrow a} f(x) = b \neq f(a)$$

### 1.3.3 Korollar

Sei  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m, a \in \overline{A}, b \in \mathbb{R}^m$  dann sind folgende Aussagen äquivalent:

- (i)  $f(x) \xrightarrow{x \rightarrow a} b$
- (ii)  $\|f(x) - b\| \xrightarrow{x \rightarrow a} 0 \in \mathbb{R}^1$  (Eine Norm bildet immer auf ein Skalar ab)
- (iii)  $f_1(x) \xrightarrow{x \rightarrow a} b_1, \dots, f_m(x) \xrightarrow{x \rightarrow a} b_m$

Zusätzlich gilt das **Cauchy-Kriterium**:

$$\lim_{x \rightarrow a} f(x) = b \Leftrightarrow \forall \varepsilon > 0 \exists \delta(\varepsilon) > 0 : \|f(x), f(y)\| < \varepsilon \quad \forall x, y \in \dot{U}_{\delta(\varepsilon)}(a) \cap A$$

### 1.3.4 Beispiel

Sei  $f(x, y) = \frac{xy}{x^2 + y^2}$

$$a_k = \begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} \frac{1}{k} \\ \frac{1}{k} \end{pmatrix}, \quad f(a_k) = \frac{\frac{1}{k^2}}{\frac{1}{k^2} + \frac{1}{k^2}} = \frac{1}{2} \quad \forall k$$

$$b_k = \begin{pmatrix} x_k \\ 0 \end{pmatrix} \text{ mit } x_k \xrightarrow{k \rightarrow \infty} 0, \quad f(b_k) = \frac{0}{x_k^2} \quad \forall k$$

Da  $\lim_{k \rightarrow \infty} f(a_k) = \frac{1}{2} \neq 0 = \lim_{k \rightarrow \infty} f(b_k)$  kann der Grenzwert nicht existieren.

### 1.3.5 Lemma Folgenkriterium

Sei  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m, a \in \overline{A}$

$$\underbrace{\exists b \in \mathbb{R}^m \text{ mit } \lim_{x \rightarrow a} f(x) = b}_{\text{der Grenzwert } b \text{ existiert}} \Leftrightarrow \underbrace{\begin{array}{l} \text{jede Folge } (x_k)_{k=1}^\infty \subset A \text{ mit } x_k \neq a \quad \forall k \text{ und } x_k \xrightarrow{k \rightarrow \infty} a \\ \Rightarrow f(x_k) \xrightarrow{k \rightarrow \infty} b \end{array}}_{\text{jede beliebige Folge konvergiert gegen } b}$$

### 1.3.6 Satz zu Grenzwerte verketteter Funktionen

Sei  $A \subset \mathbb{R}^n, B \subset \mathbb{R}^m, a \in \overline{A}, f : A \rightarrow B, g : \overline{B} \rightarrow \mathbb{R}^l$

$$\exists b \in \overline{B} \text{ mit } \lim_{x \rightarrow a} f(x) = b, \quad \exists c \in \mathbb{R}^l \text{ mit } \lim_{y \rightarrow b} g(y) = c \quad \Rightarrow \quad \lim_{x \rightarrow a} \underbrace{g(f(x))}_{(g \circ f)(x)} = \lim_{y \rightarrow b} g(y) = c$$

### 1.3.7 Beispiel

Sei  $f(x, y) = e^{-x^2 + y^2} = \exp(g(x, y))$  mit  $g(x, y) = x^2 + y^2$ , dabei gilt:

$$\lim_{(x,y)^\top \rightarrow (0,0)^\top} g(x, y) = \lim_{(x,y)^\top \rightarrow (0,0)^\top} x^2 + y^2 = 0 \quad \Rightarrow \quad \lim_{z \rightarrow 0} f(z) = \lim_{z \rightarrow 0} e^z = 1$$

### 1.3.8 Definition der Stetigkeit

Sei  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$

- (i)  $f$  ist **stetig** in  $a \in A$  wenn gilt:

$$\forall \varepsilon > 0 \exists \delta(\varepsilon) : \|f(x) - f(a)\| < \varepsilon \quad \forall x \in U_{\delta(\varepsilon)}(a) \cap A$$

*Bemerkung:* Es wird  $\lim_{x \rightarrow a} f(x) = f(a)$  gefordert.

Diese Definition unterscheidet sich in der nicht punktierten  $\varepsilon$ -Umgebung und es gilt  $f(a)$  anstatt b.

- (ii)  $f$  ist stetig auf  $A$ , wenn  $f$  in jedem Punkt  $a \in A$  stetig ist.

### 1.3.9 Bemerkung

- (i) Kompositionen stetiger Funktionen sind wieder stetig:  $f, g$  stetig  $\Rightarrow f + g, f - g, \dots$  stetig
- (ii) Das Folgenkriterium überträgt sich:  
Sei  $(a_k)_{k=1}^\infty$  eine Folge in  $A$  mit  $\lim_{k \rightarrow \infty} a_k = a \Leftrightarrow \lim_{k \rightarrow \infty} f(a_k) = f(a)$
- (iii) Ist  $A$  kompakt, dann nimmt eine stetige Funktion  $f : A \rightarrow \mathbb{R}$  immer ein Maximum und Minimum an:

$$\exists x_m, x_M \in A \text{ mit } f(x_m) = \min_{x \in A} f(x), f(x_M) = \max_{x \in A} f(x)$$

## 1.4 Partielle Ableitungen, Richtungsableitungen

### 1.4.1 Definition der partiellen Ableitung

Die Funktion  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  heißt **partielle differenzierbar** in  $a \in A$  nach der  $k$ -ten Variable  $x_k$  mit  $k \in \{1, \dots, n\}$  wenn der folgender Grenzwert existiert:

$$\frac{\partial}{\partial x_k} f(a) = f_{x_k}(a) = \lim_{h \rightarrow 0} \frac{f(a + h \cdot e_k) - f(a)}{h}$$

Existieren alle partielle Ableitungen  $f_{x_1}(a), \dots, f_{x_n}(a)$ , dann ist der **Gradient** von  $f$  wie folgt definiert:

$$\nabla f(a) = \begin{pmatrix} f_{x_1}(a) \\ \vdots \\ f_{x_n}(a) \end{pmatrix}$$

und die Funktion  $f$  heißt mindestens einmal partielle differenzierbar. Sind die partiellen Ableitungen  $f_{x_1}(a), \dots, f_{x_n}(a)$  zudem stetig, so heißt  $f$  einmal stetig differenzierbar:  $f \in C^1(A, \mathbb{R}^m)$  oder kurz  $f \in C^1(A)$ .

### 1.4.2 Beispiel

Sei  $f(x, y, z) = x^2 - xy + 3z$

$$\begin{aligned}
 \frac{\partial}{\partial x} f(x, y, z) &= \lim_{h \rightarrow 0} \frac{f(x+h, y, z) - f(x, y, z)}{h} \\
 &= \lim_{h \rightarrow 0} \frac{(x+h)^2 - (x+h)y + 3z - (x^2 - xy + 3z)}{h} \\
 &= \lim_{h \rightarrow 0} \frac{(x+h)^2 - x^2}{h} - \frac{(x+h)y - xy}{h} + \frac{3z - 3z}{h} \\
 &= \left( \frac{d}{dx} x^2 \right) - \left( \frac{d}{dx} x \right) y + \left( \frac{d}{dx} 0 \right) z \\
 &= 2x - y + 0 \\
 \Rightarrow \nabla f(x, y, z) &= \begin{pmatrix} 2x - y \\ -x \\ 3 \end{pmatrix}
 \end{aligned}$$

### 1.4.3 Definition der Richtungsableitung

Sei  $a, r \in \mathbb{R}^n$  mit  $\|r\| = 1$  (normiert),  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , dann heißt der folgende Grenzwert die Richtungsableitung von  $f$  bei  $a$  in Richtung  $r$ :

$$\frac{\partial}{\partial r} f(a) = f_r(a) = \lim_{h \rightarrow 0} \frac{f(a + h \cdot r) - f(a)}{h}$$

#### Bemerkung

- (i) Ist  $r = e_k$ , dann erhalten wir gerade eine partielle Ableitung.
- (ii) Es gibt Funktionen die in  $a$  in jede Richtung differenzierbar sind, aber in  $a$  nicht stetig sind!

## 1.5 Total Differenzierbarkeit

*Idee:* Differenzierbare Funktionen sind lokal im Punkt  $x_0$  linear approximierbar:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \underbrace{r(x)\|x - x_0\|}_{\tilde{r}(x)}$$

Dabei muss der Fehler  $\tilde{r}(x) = r(x)\|x - x_0\|$  *schneller gegen Null gehen als  $x$  gegen  $x_0$*  also muss  $\tilde{r}(x) = o(x - x_0)$  gelten (Landau-Notation: klein-oh).

### 1.5.1 Definition der totalen Differenzierbarkeit

Sei  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $A$  offen,  $x_0 \in A$

- (i) Die Funktion  $f$  nennt man **total differenzierbar** bei  $x_0$ , wenn eine Matrix  $A \in \mathbb{R}^{m \times n}$  existiert, mit der sich die Funktion  $f$  in einer  $\varepsilon$ -Umgebung um  $x_0$  mittels einer Hyperebene approximieren lässt:

$$f(x) = f(x_0) + A(x - x_0) + r(x)\|x - x_0\|$$

Dann nennt man die Matrix  $A = f'(x_0) = \frac{\partial}{\partial x} f(x_0)$  die total Ableitung von  $f$  in  $x_0$ .

(ii) Ist  $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$  partiell diff'bar, so nennt man die Ableitung **Jacobi-Matrix**:

$$f'(x_0) = \frac{\partial}{\partial x} f(x_0) = J_f(x_0) = \begin{pmatrix} \frac{\partial}{\partial x_1} f_1(x_0) & \cdots & \frac{\partial}{\partial x_n} f_1(x_0) \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_1} f_m(x_0) & \cdots & \frac{\partial}{\partial x_n} f_m(x_0) \end{pmatrix} \in \mathbb{R}^{m \times n}$$

*Bemerkung:* Es gilt:  $\exists f'(x_0) \Rightarrow f'(x_0) = J_f(x_0)$ , nicht aber die Gegenrichtung! Es kann also sein, dass die Jacobi-Matrix  $J_f$  existiert die Funktion aber nicht total diff'bar ist.

### 1.5.2 Beispiele

(i)

$$f(r, \varphi) = r \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \Rightarrow J_f = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

$$(ii) \quad f(x) = a + b^\top (x - x_0), \quad f: \mathbb{R}^n \rightarrow \mathbb{R}, \quad a \in \mathbb{R}, \quad b, x_0 \in \mathbb{R}^n \\ \Rightarrow f(x_0) = a, \quad f'(x_0) = b^\top$$

$$(iii) \quad f(x) = a + A(x - x_0), \quad f: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad a \in \mathbb{R}^m, \quad A \in \mathbb{R}^{m \times n}, \quad x_0 \in \mathbb{R}^n \\ \Rightarrow f(x_0) = a, \quad f'(x_0) = A$$

*Bemerkung:* Beispiel (ii) und (iii) sind lineare Funktionen.

### 1.5.3 Satz

Ist  $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  in jedem Punkt  $x_0 \in A$  total differenzierbar, so ist  $f$  stetig in  $A$ .

**Beweis:**

$$f(x) = \underbrace{f(x_0)}_{\substack{x \rightarrow x_0 \\ \rightarrow f(x_0)}} + \underbrace{A(x - x_0)}_{\substack{x \rightarrow x_0 \\ \rightarrow 0 \in \mathbb{R}^n}} + \underbrace{r(x)}_{\substack{x \rightarrow x_0 \\ \rightarrow 0 \in \mathbb{R}^m}} \underbrace{\|x - x_0\|}_{\substack{x \rightarrow x_0 \\ \rightarrow 0 \in \mathbb{R}}} \quad \text{mit } r(x) \xrightarrow{x \rightarrow x_0} 0 \\ \lim_{x \rightarrow a} f(x) = f(x_0) \quad \square$$

### 1.5.4 Satz

Sei  $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m, x_0 \in A$

a. Ist  $f$  total differenzierbar in  $x_0$ , so gilt

$$(i) \quad f'(x_0) = J_f(x_0)$$

$$(ii) \quad f \text{ ist in jede Richtung } r \text{ differenzierbar mit: } \frac{\partial}{\partial r} f(x_0) = J_f(x_0) \cdot r$$

**Beweis:** Es ist zu zeigen, dass wenn  $f$  differenzierbar in  $x_0$  die Ableitung gerade die Form  $\frac{\partial}{\partial r} f(x_0) = J_f(x_0) \cdot r$  besitzt. Für diese Ableitung muss folgendes gelten:

$$f(x) = f(x_0) + A(x - x_0) + \tilde{r}(x) \text{ mit } A = f'(x_0) \text{ und } \tilde{r} \in o(\|x - x_0\|) \Rightarrow \frac{\tilde{r}(x)}{\|x - x_0\|} \xrightarrow{x \rightarrow x_0} 0$$

$$f(x) = f(x_0 + r \cdot h) \Leftrightarrow f(x_0 + r \cdot h) - f(x_0) = A \cdot r \cdot h + \tilde{r}(x) = f'(x_0)rh + \tilde{r}(x)$$

also muss folgendes gezeigt werden:

$$\begin{aligned} & \left\| \underbrace{\frac{f(x_0 + r \cdot h) - f(x_0)}{h}}_{\text{Diff'Quotient für } \frac{\partial f}{\partial r}} - \underbrace{f'(x_0) \cdot r}_{\text{Grenzwert-kandidat}} \right\| \xrightarrow{h \rightarrow 0} 0 \\ & \left\| \frac{f(x_0 + r \cdot h) - f(x_0)}{h} - f'(x_0) \cdot r \right\| = \left\| \frac{f'(x_0)rh + \tilde{r}(x)}{h} - f'(x_0)r \right\| = \\ & \left\| \frac{f'(x_0)r}{h} \cdot h + \frac{\tilde{r}(x)}{h} - f'(x_0)r \right\| = \left\| \frac{\tilde{r}(x)}{h} \right\| = \left\| \frac{\tilde{r}(x)}{x - x_0} \right\| \xrightarrow{x \rightarrow x_0} 0 \end{aligned}$$

Ist  $r = e_k$  so erhält man gerade eine Spalte der Jacobi-Matrix.

- b. Existieren in  $x_0$  alle partiellen Ableitungen (also alle Komponenten der Jacobi-Matrix) und diese stetig sind  $\Rightarrow f$  ist in  $x_0$  total differenzierbar.

**Beweis:** Für den Fall  $n = 2, m = 1$  muss folgendes gezeigt werden:

$$\begin{aligned} & \exists \nabla f(x_0) \text{ und } \tilde{r}(x) \text{ mit } f(x) = f(x_0) + \nabla f(x_0)^\top (x - x_0) + \tilde{r}(x) \\ \text{oder } & \left\| \frac{f(x) - f(x_0)}{x - x_0} - \nabla f(x_0) \right\| = \frac{\|f(x) - f(x_0) - \nabla f(x_0)(x - x_0)\|}{\|x - x_0\|} \xrightarrow{x \rightarrow x_0} 0 \end{aligned}$$

Sei  $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$  und sei  $x_0 = a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ .

Nebenrechnung: Definition zweier Hilfsfunktionen  $g_1, g_2$ :

$$\begin{aligned} & \text{Sei } g_1(t) = f(t, x_2) \quad g_2 : \mathbb{R} \rightarrow \mathbb{R} \\ & \xrightarrow{\text{MWS}} \exists \xi_1 \in (a_1, x_1) \text{ mit } g'_1(\xi_1) = \frac{g_1(x_1) - g_1(a_1)}{x_1 - a_1} \\ & = \frac{\partial}{\partial x_1} f(\xi_1, x_2) = \frac{f(x_1, x_2) - f(a_1, x_2)}{x_1 - a_1} \\ & \Leftrightarrow f(x_1, x_2) - f(a_1, x_2) = \frac{\partial}{\partial x_1} f(\xi_1, x_2)(x_1 - a_1) \\ & \text{analog gilt für } g_2(t) = f(a_1, t) \quad g_2 : \mathbb{R} \rightarrow \mathbb{R} \\ & f(a_1, a_2) - f(a_1, x_2) = \frac{\partial}{\partial x_2} f(a_1, \xi_2)(x_2 - a_2) \end{aligned}$$

Damit gilt:

$$f(x) - f(a) = f(x_1, x_2) - f(a_1, a_2) = f(x_1, x_2) - \underbrace{f(a_1, x_2) + f(a_1, x_2) - f(a_1, a_2)}_{=0}$$

$$\begin{aligned} & \stackrel{\text{mit Resultat aus Nebenrechnung}}{=} \frac{\partial}{\partial x_1} f(\xi_1, x_2)(x_1 - a_1) + \frac{\partial}{\partial x_2} f(a_1, \xi_2)(x_2 - a_2) \\ & = \begin{pmatrix} f_{x_1}(\xi_1, x_2) \\ f_{x_2}(a_1, \xi_2) \end{pmatrix}^\top \begin{pmatrix} x_1 - a_1 \\ x_2 - a_2 \end{pmatrix} = \begin{pmatrix} f_{x_1}(\xi_1, x_2) \\ f_{x_2}(a_1, \xi_2) \end{pmatrix}^\top (x - a) \end{aligned}$$

Für  $x \rightarrow a$  gilt:

$$\begin{aligned} x_1 &\rightarrow a_1 & x_2 &\rightarrow a_2 \\ \xi_1 &\rightarrow a_1 & \xi_2 &\rightarrow a_2 \end{aligned}$$

da  $f_{x_1}, f_{x_2}$  stetig, folgt:

$$\begin{aligned} f_{x_1}(\xi_1, x_2) &\rightarrow f_{x_1}(a_1, a_2) \\ f_{x_2}(a_1, \xi_2) &\rightarrow f_{x_2}(a_1, a_2) \\ \Rightarrow \begin{pmatrix} f_{x_1}(\xi_1, x_2) \\ f_{x_2}(a_1, \xi_2) \end{pmatrix}^\top &\rightarrow \nabla f(a_1, a_2) = \nabla f(x_0) \end{aligned}$$

und es gilt:

$$\frac{\left\| f(x) - f(x_0) - \overbrace{\begin{pmatrix} f_{x_1}(\xi_1, x_2) \\ f_{x_2}(a_1, \xi_2) \end{pmatrix}^\top}^{x \rightarrow x_0 \nabla f(x_0)} (x - x_0) \right\|}{\|x - x_0\|} \xrightarrow{x \rightarrow x_0} 0$$

### 1.5.5 Bemerkung

Sei  $r$  eine Richtung mit  $\|r\| = 1$  und  $x = x_0 + r$ , dann gilt:

$$\begin{aligned} f(x) &\approx f(x_0) + \nabla f(x_0)^\top \cdot r \\ \Rightarrow 1. \text{ Fall : } & r, \nabla f(x_0) \text{ zeigen in dieselbe Richtung :} \\ & f(x) - f(x_0) \approx \|\nabla f(x_0)\| \|r\| = \|\nabla f(x_0)\| > 0 \\ \Rightarrow 2. \text{ Fall : } & r, \nabla f(x_0) \text{ zeigen in entgegengesetzte Richtungen :} \\ & f(x) - f(x_0) \approx -\|\nabla f(x_0)\| < 0 \end{aligned}$$

In allen Fällen gilt Näherungsweise:

$$-\|\nabla f(x_0)\| < \nabla f(x_0)^\top r \leq \|\nabla f(x_0)\|$$

*Fazit:* Beim Reinzoomen sind die Höhenlinien parallel. Der Gradient zeigt in Richtung des steilsten Anstieges.

### 1.5.6 Satz zur Kettenregel

Ist  $f : A \subset \mathbb{R}^n \rightarrow B \subset \mathbb{R}^m$  differenzierbar in  $a \in A$  und  $g : B \subset \mathbb{R}^m \rightarrow \mathbb{R}^l$  differenzierbar in  $b \in B$ , so gilt:

$$(g \circ f)'(a) = g' \left( \underbrace{f(a)}_{=b} \right) f'(a) = \underbrace{J_g(b)}_{\in \mathbb{R}^{l \times m}} \underbrace{J_f(a)}_{\in \mathbb{R}^{m \times n}}$$

**Beispiel aus der Strömungsmechanik:** Die Funktion  $f : \mathbb{R}^4 \rightarrow \mathbb{R}$ ,  $f = f(x, y, z, t)$  beschreibe die Eigenschaften eines Teilchens in einer Strömung. Dabei kann die Bewegung der Position im Raum  $x, y, z$  als Abhängigkeit von der Zeit beschrieben werden. Dazu definieren wir den Weg  $\gamma(t)$ :

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^4, \quad \gamma(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \\ t \end{pmatrix} \quad \text{mit} \quad \frac{d\gamma}{dt} = \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \\ \frac{dz}{dt} \\ 1 \end{pmatrix}$$

Nun leiten wir die verkettete Funktion  $\hat{f}(t) = (f \circ \gamma)(t) = f(\gamma(t))$  nach der Zeit  $t$  ab:

$$\begin{aligned} \frac{d\hat{f}}{dt} &= \left( \hat{f}(\gamma(t)) \right)' = f'(\gamma(t)) \gamma'(t) = \nabla f \cdot \frac{d\gamma}{dt} \\ &= \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix}^\top \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \\ \frac{dz}{dt} \\ 1 \end{pmatrix} \\ &= \frac{\partial f}{\partial x} \underbrace{\frac{dx}{dt}}_u + \frac{\partial f}{\partial y} \underbrace{\frac{dy}{dt}}_v + \frac{\partial f}{\partial z} \underbrace{\frac{dz}{dt}}_w + \frac{\partial f}{\partial t} \end{aligned}$$

Dabei beschreibt der Vektor  $\begin{pmatrix} u \\ v \\ w \end{pmatrix}$  die Geschwindigkeit im Raum.

## 1.6 Lokale Extremstellen und Mittelwertsätze

In einer Dimension gilt:

### 1. Mittelwertsatz

Ist  $f$  differenzierbar auf  $(a, b)$  und stetig auf  $[a, b]$ , so gilt:

$$\exists \xi \in (a, b) \text{ mit } f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

### Satz von Rolle

Ist  $f$  differenzierbar auf  $(a, b)$  und stetig auf  $[a, b]$  und gilt  $f(a) = f(b)$ , so gilt:

$$\exists \xi \in (a, b) \text{ mit } f'(\xi) = 0$$



**1.6.1 Definition lokale/globale Extremstellen**

- (i) Eine Funktion  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  (Skalarfeld) hat bei  $x_0 \in A$  ein lokales Minimum (Maximum) wenn in einer Umgebung  $U = U_\varepsilon(x_0) \cap A$  für  $\varepsilon > 0$  (offen bezüglich  $A$ ) von  $x_0$  gilt:

$$f(x_0) \stackrel{(\geq)}{\leq} f(x) \quad \forall x \in U$$

Ist bei  $x_0$  ein lokales Minimum (Maximum) dann nennt man  $x_0$  eine lokale Extremstelle.

- (ii)  $f$  besitzt in  $x_0$  ein globales Minimum (Maximum), wenn gilt:

$$f(x_0) \stackrel{(\geq)}{\leq} f(x) \quad \forall x \in A$$

**1.6.2 Satz zur notwendigen Bedingung für eine lokale Extremstelle**

Besitzt  $f : \overset{\circ}{A} \subset \mathbb{R}^n \rightarrow \mathbb{R}$  bei  $x_0 \in A$  eine lokale Extremstelle und  $f$  ist partiell differenzierbar, dann ist

$$\nabla f(x_0) = 0$$

*Bemerkung:* Der Rand ist ausgeschlossen da  $\overset{\circ}{A}$  (alle inneren Punkte) in der Definition verwendet wurde.

Auch gilt:

$$x_0 \text{ ist eine lokale Extremstelle} \not\Rightarrow f'(x_0) = 0$$

Aus  $f'(x_0)$  folgt nicht direkt die Extremstelle, denn Sattelpunkte sind keine Extremstellen.

**Beweis:** ...

**1.6.3 Definition des kritischen Punktes**

Ein  $x_0 \in \mathbb{R}^n$  mit  $\nabla f(x_0) = 0$  heißt **kritischer** oder stationärer Punkt.

**1.6.4 Mittelwertsatz**

Sei  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$  differenzierbar und sei  $G$  offen und enthalte die Menge  $\overline{a, b} = \{a, b \in G \text{ mit } a + t(b - a) : t \in [0, 1]\}$  ( $a, b$  können durch eine Gerade verbunden werden). Dann:

$$\exists \xi \in (0, 1) \quad \text{mit} \quad f(b) = f(a) + \nabla f(a + \xi(b - a))^\top (b - a)$$

*Bemerkung:*

$$\begin{aligned} h(t) &= a + t(b - a) & g(t) &= f(h(t)) \quad (\text{differenzierbar}) \\ \Rightarrow \exists \xi \in (0, 1) & \quad \text{mit} \quad g'(\xi) &= \frac{g(1) - g(0)}{1 - 0} \end{aligned}$$

**Beweis:** Definiere  $h(t) = a + t(b - a)$  und  $g : [0, 1] \rightarrow \mathbb{R}, g(t) = f(h(t))$  differenzierbar, damit gilt:

$$\begin{aligned} \exists \xi \in (0, 1) \quad \text{mit} \quad g'(\xi) &= \frac{g(1) - g(0)}{1 - 0} = g(1) - g(0) = f(a) - f(b) \\ g'(\xi) &= \frac{d}{dt} g(t)|_{t=\xi} = \frac{d}{dt} f(h(t))|_{t=\xi} \\ &\stackrel{\text{Kettenregel}}{=} f'(h(t))h'(t)|_{t=\xi} \\ &= \nabla f(a + \xi(b - a))^\top (b - a) = f(a) - f(b) \\ \Leftrightarrow f(b) &= f(a) + \nabla f(a + \xi(b - a))^\top (b - a) \end{aligned}$$

### 1.6.5 Definition eines Gebiets

(i) Eine Menge, die wie folgt konstruiert werden kann, heißt **Polygonzug**:

$$\overline{a_0, \dots, a_k} = \bigcup_{j=1}^k \overline{a_{j-1}, a_j} \quad \text{mit} \quad a_0, \dots, a_k \in \mathbb{R}^n$$

- (ii) Eine Menge  $M \subset \mathbb{R}^n$  heißt **kurvenweise zusammenhängend** wenn zu beliebigen  $a, b \in M$  eine stetige Funktion  $\gamma : [0, 1] \rightarrow M$  mit  $\gamma(0) = a, \gamma(1) = b$  existiert.
- (iii) Eine Menge  $G \subset \mathbb{R}^n$  heißt **Gebiet**, wenn  $G$  offen und kurvenweise zusammenhängend ist.

### 1.6.6 Bemerkungen zu Gebieten

- (i) Ein Gebiet  $G$  entspricht einem offenen Intervall  $(a, b) \subset \mathbb{R}$  im Eindimensionalen: Der Rand ist nicht dabei, es hat keine Inseln.
- (ii) Man kann zeigen, dass es reicht, wenn  $a, b \in G$  mit einem Polygonzug verbunden werden kann.

### 1.6.7 Satz

Sei  $G \subset \mathbb{R}^n$  ein Gebiet,  $G \neq \emptyset$ , und  $f : G \rightarrow \mathbb{R}$  differenzierbar, dann gilt:

$$f(x) = \text{konst.} \quad \Leftrightarrow \quad \nabla f(x) = 0 \quad \forall x \in G$$

**Beweis:** ...

### 1.6.8 Definition partieller Ableitungen $r$ 'ter Ordnung

Für  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  definiert man (wenn diese auch existieren) induktiv für  $x_0 \in A$  und  $k_1, \dots, k_r \in \{1, \dots, n\}$  die partiellen Ableitungen  $r$ 'ter Ordnung als:

$$\frac{\partial^n}{\partial x_{k_1} \dots \partial x_{k_r}} f(x_0) = f_{x_{k_1} \dots x_{k_r}} = \begin{cases} f(x_0) & r = 0 \\ \frac{\partial}{\partial x_{k_1}} f(x_0) & r = 1 \\ \frac{\partial}{\partial x_{k_1}} \left( \frac{\partial^{r-1}}{\partial x_{k_2} \dots \partial x_{k_r}} f(x_0) \right) & r > 1 \end{cases}$$

Existieren alle Ableitungen  $r$ 'ter Ordnung und sind diese zudem stetig, so nennt man die Funktion  $f$   $r$ -mal stetig differenzierbar:  $f \in C^r(A; \mathbb{R}^m)$ .

### 1.6.9 Definition der Hessematrix

Ist  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  2-mal stetig differenzierbar bei  $x_0 \in A$ , dann ist die **Hessematrix** wie folgt definiert:

$$H_f(x_0) = \begin{pmatrix} f_{x_1, x_1}(x_0) & \dots & f_{x_1, x_n}(x_0) \\ \vdots & & \vdots \\ f_{x_m, x_1}(x_0) & \dots & f_{x_m, x_n}(x_0) \end{pmatrix}$$

### 1.6.10 Beispiele

(i)

$$f(x, y) = 2xy^3 + y \log x$$

(ii)

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

$\Rightarrow$  Hessematrix ist nicht symmetrisch.

(iii)  $A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^n, c \in \mathbb{R}, Q : \mathbb{R}^n \rightarrow \mathbb{R}$ 

$$Q(x) = x^\top A x + b^\top x + c$$

$$\begin{aligned} \nabla Q(x) &= (A + A^\top) x + b \quad \overbrace{(\quad)}^{\text{wenn } A \text{ sym.}} (= 2Ax + b) \\ H_Q(x) &= A + A^\top \quad \underbrace{(\quad)}_{\text{wenn } A \text{ sym.}} (= 2A) \end{aligned}$$

$Q$  wird eine quadratische Funktion genannt.

### 1.6.11 Satz von Schwarz

Ist  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  in  $x_0$  2-mal stetig partielle differenzierbar ( $f \in C^2(A)$ ), dann ist  $H_f(x_0) \quad \forall x_0 \in A$  symmetrisch und es gilt:

$$f_{x_l, x_k}(x_0) = \frac{\partial}{\partial x_l \partial x_k} f(x_0) = \frac{\partial}{\partial x_k \partial x_l} f(x_0) = f_{x_k, x_l}(x_0) \quad \forall l, k \in \{1, \dots, k\}$$

### 1.6.12 Satz von Taylor mit quadratischem Restglied

Seien  $a, b \in G \subset \mathbb{R}^n, f \in C^2(G), G$  ein Gebiet, dann:

$$\exists \xi \in \overline{a, b} \quad \text{mit} \quad f(b) = f(a) + \nabla f(a)^\top (b - a) + \frac{1}{2} (b - a)^\top H_f(\xi) (b - a)$$

**Beweis:** Definiere  $g(t) = f(h(t))$  mit  $h(t) = a + t(b - a) \Rightarrow g(0) = f(a), g(1) = f(b)$ . Bei Funktionen mit einem Skalar, gilt der eindimensionale Taylor mit einer Zwi-

schenstelle  $z \in [0, 1]$ :

$$\begin{aligned} \Rightarrow \underbrace{g(1)}_{f(b)} &= \underbrace{g(0)}_{f(a)} + \underbrace{g'(0)(1-0)}_{\substack{g'(t)=\frac{d}{dt}f(h(t)) \\ =f'(h(t))h'(t) \\ =\nabla f(a+t(b-a))^\top(b-a) \\ =(b-a)^\top \nabla f(a+t(b-a))}} + \underbrace{\frac{1}{2}g''(z)(1-0)^2}_{g''(t)=\dots=(b-a)^\top H_f(a+t(b-a))(b-a)} \\ &= f(a) + \nabla f(a)^\top (b-a) + \frac{1}{2}(b-a)^\top H_f(\underbrace{\xi}_{\xi=a+t(b-a)})(b-a) \end{aligned}$$

### 1.6.13 Definition von Definitheit

Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch ( $A = A^\top$ ):

- Die durch  $Q_A(x) = x^\top A x$  definierte Funktion  $Q : \mathbb{R}^n \rightarrow \mathbb{R}$  heißt quadratische Form von  $A$ .
- Die Matrix  $A$  und ihre quadratische Form  $Q_A$  heißen:

- positiv definit**, wenn

$$Q_A(x) = x^\top A x > 0 \quad \forall x \in \mathbb{R}^n \text{ mit } x \neq 0$$

**negativ definit**, wenn

$$Q_A(x) = x^\top A x < 0 \quad \forall x \in \mathbb{R}^n \text{ mit } x \neq 0$$

oder kurz **definit** falls die Matrix  $A$  positiv oder negativ definit ist.

- semi definit** falls die Matrix  $A$  positiv semi definit oder (negativ semi definit) ist:

$$Q_A(x) = x^\top A x \stackrel{(\leq)}{\geq} 0 \quad \forall x \in \mathbb{R}^n \text{ mit } x \neq 0$$

- indefinit**, wenn ein  $x_1, x_2 \in \mathbb{R}^n$  existieren mit:

$$\underbrace{x_1^\top A x_1}_{Q_A(x_1)} < 0 \quad \text{und} \quad \underbrace{x_2^\top A x_2}_{Q_A(x_2)} > 0$$

### 1.6.14 Beispiele

- $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
- $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
- $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
- $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$

$$(v) \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$(vi) \quad A = \begin{pmatrix} a & b & 0 \\ b & c & 0 \\ 0 & 0 & d \end{pmatrix}$$

$$(vii) \quad A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

### 1.6.15 Satz zum Hauptminorenkriterium

Sei  $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \in \mathbb{R}^{n \times n}$  symmetrisch,

dann ist der  $k$ 'te Hauptminor definiert als:  $D_k = \det \underbrace{\begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}}_{k \times k \text{ Teilmatrix oben links}}$

Für die Hauptminoren  $D_1, \dots, D_n$  gilt:

- a.  $A$  ist positiv definit  $\Leftrightarrow D_1 > 0, D_2 > 0, \dots, D_n > 0$  alle Hauptminoren sind positiv
- b.  $A$  ist negativ definit  $\Leftrightarrow D_1 < 0, D_2 > 0, D_3 < 0, \dots$  oder  $D_k = \begin{cases} < 0 & k \text{ ungerade} \\ > 0 & k \text{ gerade} \end{cases}$
- c. (i)  $D_k < 0$  mit  $k$  gerade  $\Rightarrow A$  ist indefinit  
 (ii)  $D_k < 0 < D_l$  mit  $k, l$  ungerade  $\Rightarrow A$  ist indefinit

### Beispiele

$$(i) \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$(ii) \quad A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$(iii) \quad A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

### 1.6.16 Satz über die hinreichenden Bedingungen für lokale Extremstellen

Sei  $f \in C^2(U)$  in einer Umgebung  $U$  um  $x_0$  und gilt  $\nabla f(x_0) = 0$  sowie:

- (i)  $H_f(x_0)$  ist positiv definit  $\Rightarrow x_0$  ist eine lokale Minimalstelle.
- (ii)  $H_f(x_0)$  ist negativ definit  $\Rightarrow x_0$  ist eine lokale Maximalstelle.

**Beweis:** o.B.d.A nur für (i), denn (ii) folgt analog mit gedrehtem Ungleichungszeichen.

Sei  $U = U_\varepsilon(x_0)$  eine  $\varepsilon$ -Umgebung um den Punkt  $x_0 \in \mathbb{R}^n$ :

Sei  $f \in C^2(U) \Rightarrow H_f \in C^2(U, \mathbb{R}^{n \times n})$  mit allen Komponenten stetig  
 $\Rightarrow \tilde{g} : \mathbb{R}^n \rightarrow \mathbb{R}, \tilde{g}(x) = (x - x_0)^\top H_f(x)(x - x_0)$  ist stetig  
 $\Rightarrow g(x) = (x - x_0)^\top H_f(h(x))(x - x_0)$  ist stetig  
 ist  $H_f(x_0)$  pos. def.  $\Rightarrow (x - x_0)^\top H_f(x_0)(x - x_0) > 0$  für  $x \neq x_0$   
 $\Rightarrow (x - x_0)^\top H_f(h(x_0))(x - x_0) > 0$  falls  $z(x_0)$  nahe genug bei  $x_0$

Sei  $h(x_0) = x_0 + \xi(x - x_0)$  dann gilt:

$$f(x) = f(x_0) + \underbrace{\underbrace{\nabla f(x_0)^\top (x - x_0)}_{=0}}_{=0} + \underbrace{\frac{1}{2}(x - x_0)^\top \underbrace{H_f(x_0 + \xi(x - x_0))}_{\text{pos. def. in } U_\varepsilon(x_0)}(x - x_0)}_{>0 \ \forall x \in U_\varepsilon(x_0)} \geq f(x_0)$$

### 1.6.17 Definition der Sattelpunkte

Sei  $U = U_\varepsilon(x_0), f \in C^2(U), \nabla f(x_0) = 0, H_f(x_0)$  indefinit, dann besitzt die Funktion  $f$  bei  $x_0$  einen **Sattelpunkt**.

*Bemerkung:*  $\forall \varepsilon > 0$  gilt:

$$\exists x_1, x_2 \in U_\varepsilon(x_0) \text{ mit } f(x_1) > f(x_0) \text{ und } f(x_2) < f(x_0)$$

### 1.6.18 Satz für den Fall $n = 2$

Ist  $f \in C^2(U)$  für eine Umgebung  $U$  von  $x_0 = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$  gilt weiter  $f_x(a, b) = f_y(a, b) = 0$ , dann ist  $x_0$ :

- (i) eine lokale **Minimalstelle**, wenn  
 $f_{xx}(a, b) > 0$  (erster Hauptminor  $D_1$ ) und  
 $f_{xx}(a, b)f_{yy}(a, b) - 2(f_{xy}(a, b))^2 > 0$  (zweiter Hauptminor  $D_2$ )
- (ii) eine lokale **Maximalstelle**, wenn  
 $f_{xx}(a, b) < 0$  (erster Hauptminor  $D_1$ ) und  
 $f_{xx}(a, b)f_{yy}(a, b) - 2(f_{xy}(a, b))^2 > 0$  (zweiter Hauptminor  $D_2$ )
- (iii) ein **Sattelpunkt**, wenn  
 $f_{xx}(a, b) < 0$  (erster Hauptminor  $D_1$ ) oder  
 $f_{xx}(a, b)f_{yy}(a, b) - 2(f_{xy}(a, b))^2 < 0$  (zweiter Hauptminor  $D_2$ )

### 1.6.19 Beispiel

Sei  $f(x, y, z) = x^2 + xy + y^2 - \cos z \in C^2(\mathbb{R}^2)$ :

$$\nabla f(x, y, z) = \begin{pmatrix} 2x + y \\ x + 2y \\ \sin z \end{pmatrix} \stackrel{!}{=} 0$$

## 1.7 Extremstellen unter Nebenbedingungen und implizite Funktionen

*Bisher:* Optimierungsproblem ohne Nebenbedingungen:  $\begin{cases} f(x, y) \rightarrow \min \\ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \end{cases}$

Lösbar in drei Schritten:

1. lokale Minimalstellen bestimmen
2. untersuche  $f(x, y)$  für  $\left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\| \rightarrow \infty$
3. vergleiche die Resultate aus 1. und 2.

*Bemerkung:* aus  $f(x, y) \rightarrow \min$  wird  $-f(x, y) \rightarrow \max$  darum reicht es o.B.d.A. denn Fall  $\rightarrow \min$  zu betrachten.

*Jetzt:* Optimierungsproblem mit Nebenbedingungen:  $\begin{cases} f(x, y) \rightarrow \min \\ \begin{pmatrix} x \\ y \end{pmatrix} \in A \subset \mathbb{R}^2 \end{cases}$

Es ist bekannt, dass wenn  $f \in C(A)$  und  $A \neq \emptyset$  kompakt (abgeschlossen und beschränkt) dann existiert sicher eine Lösung (Satz von Weierstraß:  $f$  stetig und  $A$  kompakt  $\Rightarrow$  es wird ein Minimum (Maximum) angenommen).

Es sind zwei Fälle möglich:

- (i) globale Minimalstelle liegt in  $\overset{\circ}{A}$  (im Inneren)
- (ii) globale Minimalstelle liegt in  $\partial A$  (auf dem Rand)

Wenn zusätzlich  $f \in C^2(A)$  gilt, kann die Minimalstelle wie folgt gefunden werden:

1. bestimme lokale Minimalstellen mit  $\nabla f(x_0) = 0$  und  $H_f(x_0)$  positiv definit in  $\overset{\circ}{A}$
2. bestimme lokale Minimalstelle in  $\partial A$

*Bemerkung:* Eckpunkte müssen gesondert betrachtet werden, denn die Funktion kann an diesen Stellen nicht differenzierbar sein.

3. wähle das Minimum aus 1. und 2.

### Wiederholung im Eindimensionalen

Eine Funktion  $f : X$  (Definitionsbereich)  $\rightarrow Y$  (Bildbereich) ist:

- (i) **injektiv**, wenn

$$x_1, x_2 \in X \text{ mit } x_1 \neq x_2 \quad \Rightarrow \quad f(x_1) \neq f(x_2)$$

*Bemerkung:* Von zwei verschiedenen Punkten aus dem Definitionsbereich darf nicht auf den gleichen Punkt im Bildbereich abgebildet werden.

- (ii) **surjektiv**, wenn

$$\forall y \in Y \quad \exists x \in X \quad \text{mit} \quad f(x) = y$$

*Bemerkung:* Für alle Bildpunkte existiert ein Punkt im Definitionsbereich.

- (iii) **bijektiv**, wenn die Funktion  $f$  injektiv und surjektiv ist.

**Beispiel**

Sei  $d : C^1(\mathbb{R}) \rightarrow C(\mathbb{R})$  eine Funktion mit  $f \mapsto f'$  (Ableitungsoperator für alle einmal stetig differenzierbaren Funktionen). Für diese Funktion  $d$  gilt:

- (i)  $d$  ist nicht injektiv da:  
Seien  $f_1(x) = x, f_2(x) = x + 1$  zwei Funktionen aus  $C^1(\mathbb{R})$ . Für diese gilt:  $f_1 \neq f_2$  aber  $d(f_1) = f'_1 = 1 = f'_2 = d(f_2)$ .
- (ii)  $d$  ist surjektiv, denn nach dem Hauptsatz der Differential und Integralrechnung gilt:  
Alle stetigen Funktionen besitzen eine Stammfunktion.

**1.7.1 Spezialfälle**

- (i)  $f$  ist linear und  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m : f(x) = Ax$  mit  $A \in \mathbb{R}^{m \times n}$   
Im Fall  $m = n$  gilt:  $f$  ist bijektiv  $\Leftrightarrow A$  ist invertierbar.
- (ii) Allgemeinerer Fall z.B.  $f : [0, \infty) \times [-\pi, \pi) \rightarrow \mathbb{R}^2 : f(r, \varphi) = r \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$   
 $f$  ist surjektiv aber nicht injektiv, denn  $f(0, \varphi) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \forall \varphi \in [-\pi, \pi)$

**1.7.2 Bemerkung**

a. Folgende Aussagen sind äquivalent:

- (i)  $A$  ist invertierbar (oder regulär oder nicht singulär)
- (ii)  $Ax = b$  besitzt  $\forall b \in \mathbb{R}^n$  eine eindeutige Lösung  $x \in \mathbb{R}^n$
- (iii)  $\det A \neq 0$

b. Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  mit  $x_0 \in \mathbb{R}^n : f(x) = f(x_0) + A(x - x_0)$  mit  $\det A \neq 0$   
Die Umkehrfunktion  $f^{-1}$  kann durch Äquivalenzumformungen gebildet werden:

$$\begin{aligned} f(x_0) + A(x - x_0) &\stackrel{!}{=} y \\ \Leftrightarrow A(x - x_0) &= y - f(x_0) \\ \Leftrightarrow x - x_0 &= A^{-1}(y - f(x_0)) \\ \Leftrightarrow x &= x_0 + A^{-1}(y - f(x_0)) = f^{-1}(y) \end{aligned}$$

Für die Ableitung der Umkehrfunktion gilt:  $\frac{\partial}{\partial y} f^{-1}(y) = A^{-1}$

*Kommentar:* Differenzierbare Funktionen sind in einer hinreichend kleinen  $\varepsilon$ -Umgebung im Prinzip linear (nicht ganz korrekt, aber sehr anschaulich).

**1.7.3 Satz über die Umkehrfunktion**

Sei  $f \in C^1(G; \mathbb{R}^n)$  mit  $G \subset \mathbb{R}^n$  ein Gebiet,  $x_0 \in G$  mit  $\det f'(x_0) = \det J_f(x_0) \neq 0$ . Dann gilt in einer geeigneten offenen Umgebung  $U$  um  $x_0$ , dass

- (i)  $V = f(U)$  ist offen (das Bild  $V$  von  $U$  ist offen) und  $\det f'(x) \neq 0 \quad \forall x \in U$
- (ii)  $f : U \rightarrow V$  ist bijektiv, das heißt:  $\exists f^{-1} : V \rightarrow U$  (die Funktion ist lokal invertierbar)



**Beweisidee:** Da  $f$  stetig differenzierbar ist, gilt:

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + r(x)\|x - x_0\| \text{ mit } r(x) \xrightarrow{x \rightarrow x_0} 0 \\ &\approx f(x_0) + f'(x_0)(x - x_0) \\ &= f(x_0) + A(x - x_0) \text{ mit } \det A \neq 0 \end{aligned}$$

Wenn  $\approx$  ein  $=$  wäre so könnte die Umkehrfunktion  $f^{-1}(y)$  einfach durch Äquivalenzumformungen bestimmt werden

$\Rightarrow$  In einer hinreichend kleinen  $\varepsilon$ -Umgebung kann  $\approx$  als  $=$  angenommen werden.

$$(iii) \quad (f^{-1})'(y) = (f'(f^{-1}(y)))^{-1} \quad \Leftrightarrow \quad J_{f^{-1}}(y) = J_{f'}(f^{-1}(y))$$

**Beweisidee:** Wenn  $f^{-1} : V \rightarrow U$  existiert, dann ist  $f^{-1}$  auch stetig diff'bar.  
Beispielskizze aus dem Eindimensionalen:

$$\begin{aligned} f(x) &= e^x, \quad f'(x) = e^x, \quad \exists f^{-1}(y) = \log y \\ \Rightarrow \quad y &= f(f^{-1}(y)) \\ \xRightarrow{\text{Ableiten mit Kettenregel}} \quad 1 &= f'(f^{-1}(y)) \cdot (f^{-1})'(y) \\ \Rightarrow \quad (f^{-1})'(y) &= \log' y = \frac{1}{f'(f^{-1}(y))} = \frac{1}{e^{\log y}} = \frac{1}{y} \end{aligned}$$

Beispiel übertragen auf den allgemeinen mehrdimensionalen Fall:

$$\begin{aligned} y &= f(f^{-1}(y)) \\ \xRightarrow{\text{Ableiten mit Kettenregel}} \quad 1 &= \underbrace{f'(f^{-1}(y))}_{J_{f'}(f^{-1}(y))} \cdot \frac{d}{dy} f^{-1}(y) \\ \Rightarrow \quad \frac{d}{dy} f^{-1}(y) &= (f'(f^{-1}(y)))^{-1} \end{aligned}$$