HSB
Hochschule Bremen
City University of Applied Sciences

Höhere Mathematik 2

Kapitel 7

Differentialrechnung für Funktionen mehrerer Variabler

Prof. Dr.-Ing. Dieter Kraus

INSTITUTE OF WATERACOUSTICS, SONAR ENGINEERING AND SIGNAL THEORY
iWSS
HOCHSCHULE BREMEN

Höhere Mathematik 2

Kapitel 7

Inhaltsverzeichnis

7 Differentialrechnung für Funktionen mehrerer Variabler

7.1 Einführung

Im folgenden werden Funktionen $f: D \subset \mathbb{R}^n \to \mathbb{R}$ betrachtet und deren Eigenschaften hinsichtlich Stetigkeit, Differenzierbarkeit, Extremalstellen etc. untersucht.

Geometrische Deutung:

$n=1$: $f: D \subset \mathbb{R} \to \mathbb{R}$

$$C_f = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : \ x \in D(f), \ \ y = f(x) \right\}$$

ist der Graph von f, d.h. Kurve im $\mathbb{R}^2$.

$n=2$: $f: D \subset \mathbb{R}^2 \to \mathbb{R}$

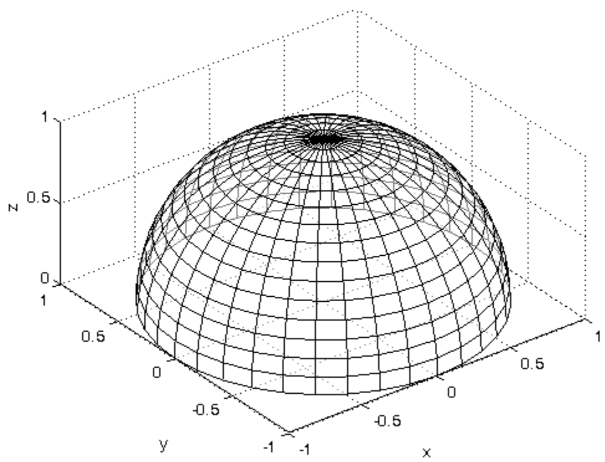
$$S_f = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : \ \begin{pmatrix} x \\ y \end{pmatrix} \in D(f), \ \ z = f(x,y) \right\}$$

ist die durch f gegebene Punktmenge, d.h. Fläche im $\mathbb{R}^3$.

Beispiel:
obere Halbkugel:

$$f(x,y) = \sqrt{1-x^2-y^2},$$

$$D(f) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : \ x^2 + y^2 \le 1 \right\}$$



Darstellungshilfen:

1) Höhenlinien

$$H_c = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : f(x,y) = c \right\} \text{ für } c = \text{konstant} \in \mathbb{R}$$

entspricht Kurve auf S_f mit konstanter Höhe c.

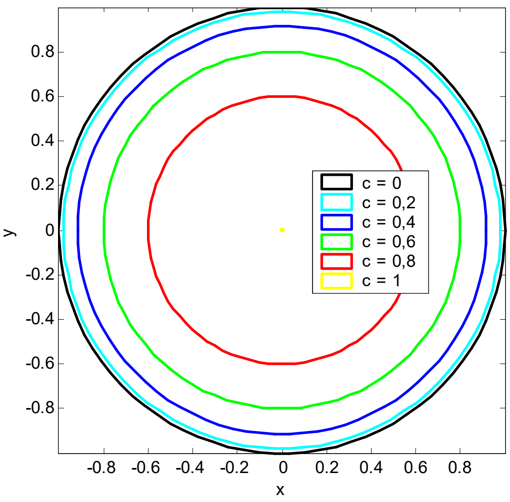
Beispiel: (*obere Halbkugel*)

$f(x,y) = \sqrt{1 - x^2 - y^2} = c$

$\Rightarrow 1 - x^2 - y^2 = c^2$

$\Rightarrow x^2 + y^2 = 1 - c^2$

$\Rightarrow$ Höhenlinien sind um $(0,0)^T$ mit dem Radius $\sqrt{1-c^2}$ gelegen.



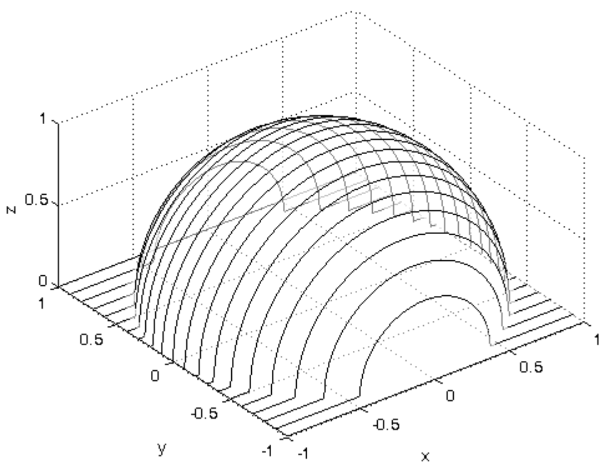
2) (x,y)-Koordinatenlinien

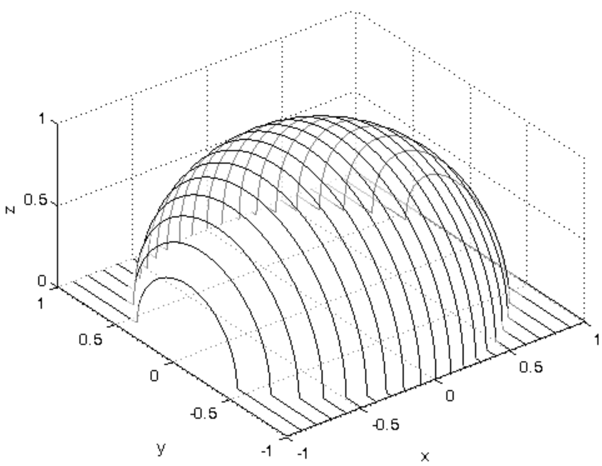
$y = y_0 = \text{konstant} \Rightarrow z = f(x, y_0)$ Kurve im $\mathbb{R}^3$

$x = x_0 = \text{konstant} \Rightarrow z = f(x_0, y)$ Kurve im $\mathbb{R}^3$

Beispiel: (*obere Halbkugel*)

$y_0: \quad z = \sqrt{1 - x^2 - y_0^2}$ $\qquad$ $x_0: \quad z = \sqrt{1 - x_0^2 - y^2}$





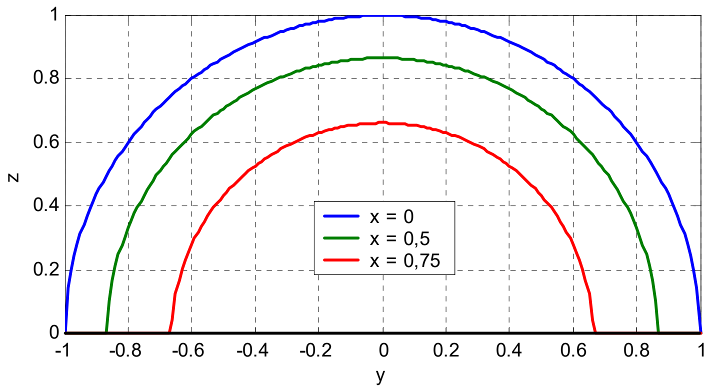
3) Schnitt mit der (x,y)- oder (y,z)- oder (x,z)-Ebene,
z.B. $x = x_0 \Rightarrow z = f(x_0, y)$ Kurve in der (y,z)-Ebene

Beispiel: (*obere Halbkugel*)

$$f(x,y) = \sqrt{1 - x^2 - y^2}$$

$$x = x_0 \Rightarrow z = \sqrt{1 - x_0^2 - y^2}$$

$$\Rightarrow \; y^2 + z^2 = 1 - x_0^2 = \text{Halbkreise in } (y,z) - \text{Ebene}$$



4) (r,φ)-Koordinatenlinien
$x = r\cos\varphi, \;\; y = r\sin\varphi$
$r = r_0 = \text{konstant} \;\Rightarrow\; z = f(r_0\cos\varphi, r_0\sin\varphi)$ Kurve im $\mathbb{R}^3$
$\varphi = \varphi_0 = \text{konstant} \;\Rightarrow\; z = f(r\cos\varphi_0, r\sin\varphi_0)$ Kurve im $\mathbb{R}^3$

Beispiel: (*obere Halbkugel*)

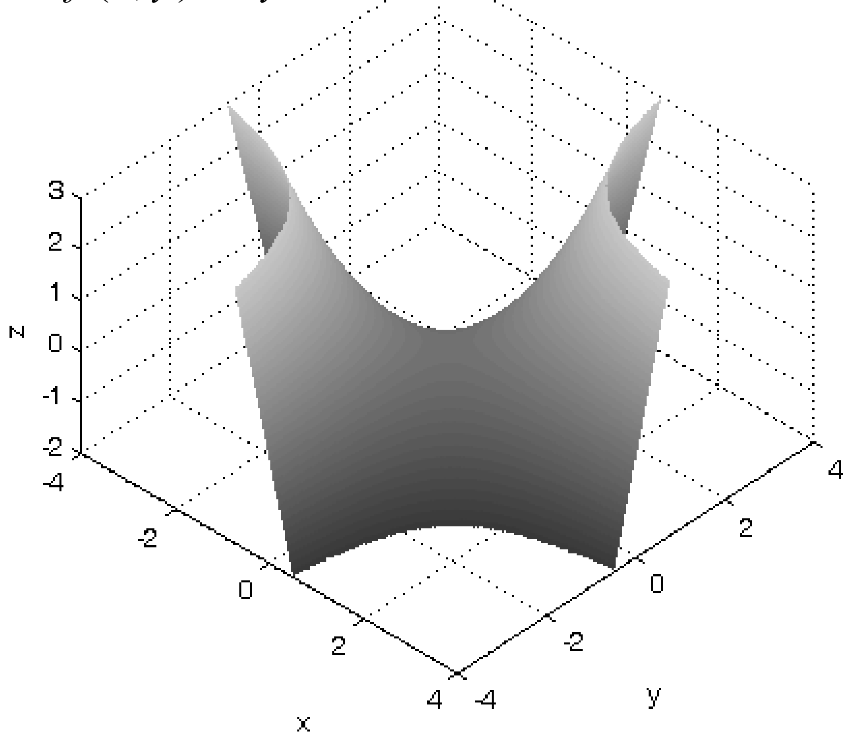
$$\begin{aligned} f(x,y) &= \sqrt{1 - x^2 - y^2} \\ &= \sqrt{1 - r^2\cos^2\varphi - r^2\sin^2\varphi} \\ &= \sqrt{1 - r^2(\cos^2\varphi + \sin^2\varphi)} = \sqrt{1 - r^2} \end{aligned}$$

$$r = r_0 \;\Rightarrow\; z = \sqrt{1 - r_0{}^2} = \text{Konstant}$$

$$\varphi = \varphi_0 \;\Rightarrow\; z = \sqrt{1 - r^2} = \text{Halbkreis in } (z,r)\text{-Ebene}$$

Beispiel:

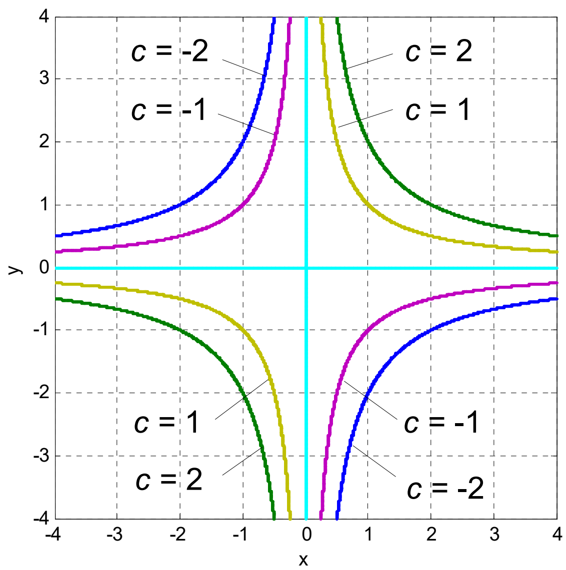
$$f(x, y) = xy$$



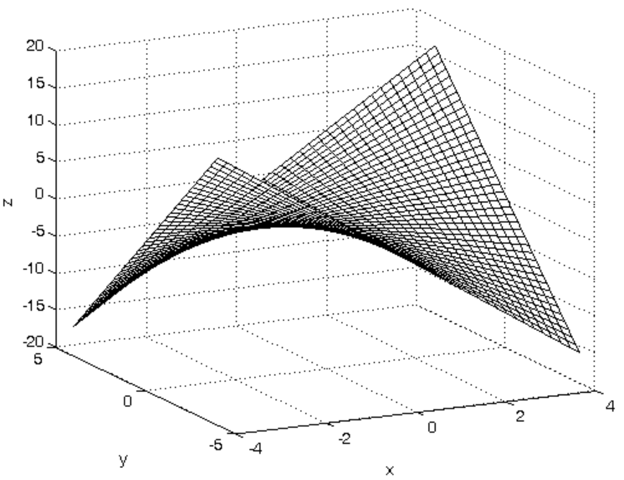
Höhenlinien

$xy = c \;\Rightarrow\; y = c/x$ falls $c \neq 0$

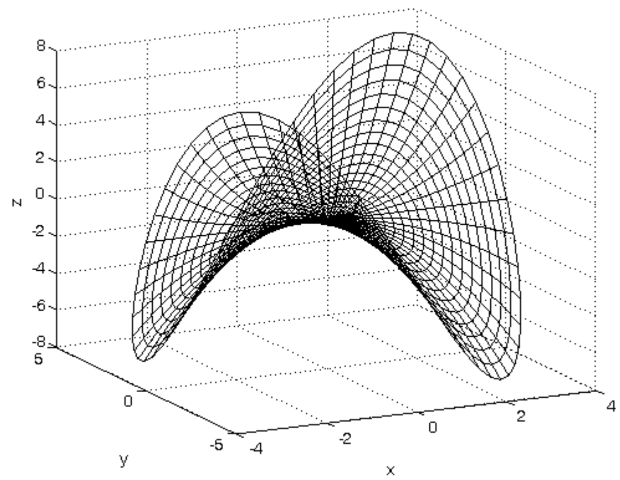
$c = 0 \;\Rightarrow\; x = 0$ (y-Achse) $\vee$ $y = 0$ (x-Achse)



(x,y)-Koordinatenlinien



(r,φ)-Koordinatenlinien



Für $n > 2$ ist die Veranschaulichung von Funktionen nicht mehr so einfach möglich.

7.2 Eigenschaften des $\mathbb{R}^n$

$$\mathbb{R}^n = \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} :\ x_i \in \mathbb{R},\ i = 1, 2, \ldots, n \right\}$$

ist mit den Operationen $+$ und λ-mal ein n-dimensionaler Vektorraum, d.h. $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gilt $\mathbf{x} + \mathbf{y},\ \lambda\mathbf{x} \in \mathbb{R}^n$.

Als Basis betrachtet man üblicherweise die Einheitsvektoren

$$\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$$

mit der 1 an der i-ten Stelle.

Die Vektoren $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ sind linear unabhängig und für $\mathbf{x} = (x_1, \ldots, x_n)^T$ gilt

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i .$$

Mit dem Skalarprodukt

$$\mathbf{x}^T\mathbf{y} = \sum_{i=1}^{n} x_i y_i\,, \qquad (\vec{x}\cdot\vec{y})$$

wird der $\mathbb{R}^n$ zum euklidschen Vektorraum.

Eigenschaften des Skalarprodukts

a) $\mathbf{x}^T\mathbf{x} \geq 0,\ \mathbf{x}^T\mathbf{x} = 0 \ \Leftrightarrow\ \mathbf{x} = \mathbf{0},$

b) $(\lambda\mathbf{x}^T)\,\mathbf{x} = \lambda(\mathbf{x}^T\mathbf{x}),$

c) $(\mathbf{x}+\mathbf{y})^T\,\mathbf{z} = \mathbf{x}^T\mathbf{z} + \mathbf{y}^T\mathbf{z},$

d) $\mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}.$

Mit der Euklidischen Norm

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^T\mathbf{x}} = \sqrt{\sum_{i=1}^{n} x_i^2} \quad \text{(entspricht der Länge des Vektors } \mathbf{x}\text{)}$$

wird der $\mathbb{R}^n$ zum normierten Vektorraum.

Eigenschaften der Norm

a) $\|\mathbf{x}\|_2 \geq 0,\ \|\mathbf{x}\|_2 = 0 \ \Leftrightarrow\ \mathbf{x} = \mathbf{0},$

b) $\|\lambda\mathbf{x}\|_2 = |\lambda|\,\|\mathbf{x}\|_2\,,$

c) $\|\mathbf{x}+\mathbf{y}\|_2 \leq \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2$ Dreiecksungleichung.

Außerdem gilt die Schwarzsche Ungleichung

$$|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\|_2\,\|\mathbf{y}\|_2\,.$$

Der Abstand zwischen zwei Vektoren $\mathbf{x}$ und $\mathbf{y} \in \mathbb{R}^n$ wird durch

$$\|\mathbf{x}-\mathbf{y}\|_2 = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

gemessen. Hierdurch wird der $\mathbb{R}^n$ zum metrischen Raum.

Eigenschaften der Metrik

a) $\|\mathbf{x}-\mathbf{y}\|_2 \geq 0$, $\|\mathbf{x}-\mathbf{y}\|_2 = 0 \;\Leftrightarrow\; \mathbf{x}=\mathbf{y}$,

b) $\|\mathbf{x}-\mathbf{y}\|_2 = \|\mathbf{y}-\mathbf{x}\|_2$,

b) $\|\mathbf{x}-\mathbf{y}\|_2 \leq \|\mathbf{x}-\mathbf{z}\|_2 + \|\mathbf{z}-\mathbf{y}\|_2$.

Die meisten Eigenschaften des $\mathbb{R}$ übertragen sich auf den $\mathbb{R}^n$ für beliebiges n, da wir auch im $\mathbb{R}^n$ den "Abstand" zwischen zwei Elementen messen können.

Unterschiede zwischen $\mathbb{R}$ und $\mathbb{R}^n$ ergeben sich dadurch, dass im $\mathbb{R}^n$

a) die Division fehlt,

b) keine Ordnungsrelation $(\leq)$ vorhanden ist,

c) die uneigentlichen Punkte $-\infty$ und ∞ nicht in Betracht gezogen werden.

Definition 7-1:

Es sei $\mathbf{a}=(a_1,...,a_n)^T, \mathbf{b}=(b_1,...,b_n)^T \in \mathbb{R}^n$ mit $-\infty < a_i < b_i < \infty$ für $i=1,2,$ ...,n. Dann heißt

a) $$[\mathbf{a},\mathbf{b}] = \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n : \; a_i \leq x_i \leq b_i \;\; i=1,2,\ldots,n \right\}$$

ein abgeschlossenes Intervall des $\mathbb{R}^n$,

b) $$(\mathbf{a},\mathbf{b}) = \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n : \; a_i < x_i < b_i \;\; i=1,2,\ldots,n \right\}$$

ein offenes Intervall des $\mathbb{R}^n$.

Man erhält für $n=1$ ein Intervall im bisherigen Sinn,
$n=2$ ein Rechteck und
$n=3$ einen Quader.

Zu einem festen $\mathbf{x}_0 \in \mathbb{R}^n$ und $\varepsilon > 0$ heißt

$$U_\varepsilon(\mathbf{x}_0) = \left\{ \mathbf{x} \in \mathbb{R}^n : \left\| \mathbf{x} - \mathbf{x}_0 \right\|_2 < \varepsilon \right\}$$

ε-Umgebung von $\mathbf{x}_0$.

Man erhält für $n = 1$ ein Intervall um $\mathbf{x}_0$, d.h. $(x_0 - \varepsilon, x_0 + \varepsilon)$,
$n = 2$ eine Kreisscheibe um $\mathbf{x}_0$ mit Radius ε und
$n = 3$ eine Kugel um $\mathbf{x}_0$ mit Radius ε.

Im $\mathbb{R}^n$ werden ε-Umgebungen allgemein Kugelumgebungen genannt.

Definition 7-2:

a) $\mathbf{x}_0 \in M$ heißt innerer Punkt von $M \subset \mathbb{R}^n$, wenn es eine ε-Umgebung $U_\varepsilon(\mathbf{x}_0)$ gibt mit $U_\varepsilon(\mathbf{x}_0) \subset M$. Die Menge der inneren Punkte von M heißt Inneres von M und wird mit $\mathring{M}$ bezeichnet.

b) $\mathbf{x}_0 \in M$ heißt Randpunkt von $M \subset \mathbb{R}^n$, wenn jede ε-Umgebung von $\mathbf{x}_0$ sowohl Punkte enthält, die zu M gehören als auch Punkte, die nicht zu

M gehören. Die Menge aller Randpunkte von M heißt Rand von M und wird mit ∂M bezeichnet.

c) Die Menge $\overline{M} = M \cup \partial M$ heißt abgeschlossene Hülle von M.

Definition 7-3:

a) Eine Menge $M \subset \mathbb{R}^n$ heißt offen, wenn alle Punkte von M innere Punkte sind.

b) Eine Menge $M \subset \mathbb{R}^n$ heißt abgeschlossen, wenn ihr Rand ∂M zu M gehört.

c) Eine Menge $M \subset \mathbb{R}^n$ heißt beschränkt, wenn es eine Konstante $K > 0$ gibt mit $\|\mathbf{x}\|_2 < K$ für alle $\mathbf{x} \in M$.

d) Eine Menge $M \subset \mathbb{R}^n$ heißt kompakt, wenn M abgeschlossen und beschränkt ist.

7.3 Folgen im $\mathbb{R}^n$

Definition 7-4:

Es sei $(\mathbf{x}_k)_{k\in\mathbb{N}}$ eine Folge im $\mathbb{R}^n$ mit

$$\mathbf{x}_k = \begin{pmatrix} x_{1,k} \\ \vdots \\ x_{n,k} \end{pmatrix} = (x_{1,k}, \ldots, x_{n,k})^T.$$

$\mathbf{a} \in \mathbb{R}^n$ heißt Grenzwert der Folge $(\mathbf{x}_k)$ genau dann, wenn

$$\lim_{k\to\infty} \left\| \mathbf{x}_k - \mathbf{a} \right\|_2 = 0.$$

In diesem Fall heißt die Folge $(\mathbf{x}_k)$ konvergent gegen $\mathbf{a}$.

Satz 7-1:

Für Folgen $(\mathbf{x}_k)_{k\in\mathbb{N}}$ im $\mathbb{R}^n$ gilt

$$\lim_{k\to\infty} \mathbf{x}_k = \mathbf{a} \Leftrightarrow \lim_{k\to\infty} x_{i,k} = a_i \quad \text{für } i = 1, 2, \ldots, n.$$

Beweis:

"$\Rightarrow$" Für $\mathbf{a} = (a_1,...,a_n)^T$ gilt $|a_i| \le \|\mathbf{a}\|_2$, also folgt $|x_{i,k} - a_i| \le \|\mathbf{x}_k - \mathbf{a}\|_2 < \varepsilon$ $\forall k > N_\varepsilon$ und $i = 1,2,...,n$ falls $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{a}$ und somit $\lim_{k\to\infty} x_{i,k}$ $= a_i$ für $i = 1,2,...,n$.

"$\Leftarrow$" $\left\|\mathbf{x}_k - \mathbf{a}\right\|_2 = \sqrt{\sum_{i=1}^n (x_{i,k} - a_i)^2} < \varepsilon$ falls $\left|x_{i,k} - a_i\right| < \varepsilon / \sqrt{n}$ für $i = 1,2,...,n$ demzufolge gilt $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{a}$ falls $\lim_{k\to\infty} x_{i,k} = a_i$ für $i = 1,2,...,n$.

Satz 7-2:

Es seien $(\mathbf{x}_k)_{k\in\mathbb{N}}$ und $(\mathbf{y}_k)_{k\in\mathbb{N}}$ zwei Folgen im $\mathbb{R}^n$ mit $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{a}$ und $\lim_{k\to\infty} \mathbf{y}_k = \mathbf{b}$ und $(\lambda_k)_{k\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ mit $\lim_{k\to\infty} \lambda_k = \lambda \in \mathbb{R}$, dann gilt

a) $\lim_{k\to\infty} (\mathbf{x}_k \pm \mathbf{y}_k) = \mathbf{a} \pm \mathbf{b}$, b) $\lim_{k\to\infty} \lambda_k \mathbf{x}_k = \lambda \mathbf{a}$,

c) $\lim_{k\to\infty} \mathbf{x}_k^T \mathbf{y}_k = \mathbf{a}^T \mathbf{b}$, d) $\lim_{k\to\infty} \left\|\mathbf{x}_k\right\|_2 = \left\|\mathbf{a}\right\|_2$.

Beweis:

a) und b) ergibt sich sofort durch Übergang auf die einzelnen Koordinaten,

c) $\mathbf{x}_k^T \mathbf{y}_k = \sum_{i=1}^{n} x_{i,k}\, y_{i,k} \xrightarrow[k\to\infty]{} \sum_{i=1}^{n} a_i b_i = \mathbf{a}^T \mathbf{b}$,

d) $\|\mathbf{x}_k\|_2 = \sqrt{\sum_{i=1}^{n} x_{i,k}^2} \xrightarrow[k\to\infty]{} \sqrt{\sum_{i=1}^{n} a_i^2} = \|\mathbf{a}\|_2$.

Analog zum Fall $n = 1$ zeigt man

Satz 7-3:

a) Jede Folge $(\mathbf{x}_k)_{k\in\mathbb{N}}$ des $\mathbb{R}^n$ besitzt höchstens einen Grenzwert.

b) Jede Teilfolge einer konvergenten Folge konvergiert gegen den gleichen Grenzwert.

c) Jede konvergente Folge $(\mathbf{x}_k)_{k\in\mathbb{N}}$ ist beschränkt, d.h. die Folge $(\|\mathbf{x}_k\|_2)_{k\in\mathbb{N}}$ ist beschränkt.

Definition 7-5: (*Cauchy-Konvergenz*)

Eine Folge $(\mathbf{x}_k)_{k\in\mathbb{N}}$ des $\mathbb{R}^n$ heißt Cauchy-konvergent genau dann, wenn für jedes beliebig kleine $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ existiert mit $\|\mathbf{x}_k - \mathbf{x}_l\|_2 < \varepsilon$ für alle $k, l > N_\varepsilon$.

Satz 7-4: (*Vollständigkeit des* $\mathbb{R}^n$)

Die Folge $(\mathbf{x}_k)_{k\in\mathbb{N}}$ des $\mathbb{R}^n$ ist genau dann konvergent, wenn die Folge $(\mathbf{x}_k)_{k\in\mathbb{N}}$ Cauchy-konvergent ist.

Anmerkung:

Ein Punkt $\mathbf{x}_0 \in \mathbb{R}^n$ heißt Häufungspunkt von $M \subset \mathbb{R}^n$, wenn in jeder ε-Umgebung von $\mathbf{x}_0$ unendlich viele Punkte aus M liegen. Dabei muss $\mathbf{x}_0$ nicht zu M gehören.

Ist $\mathbf{x}_0$ Häufungspunkt von $M \subset \mathbb{R}^n$ so gibt es eine Folge $(\mathbf{x}_k)_{k\in\mathbb{N}}$ mit $\mathbf{x}_k \in M$ und $\mathbf{x}_k \neq \mathbf{x}_0$ für $k \in \mathbb{N}$, die gegen $\mathbf{x}_0$ konvergiert.

7.4 Stetigkeit von Funktionen mehrerer Variabler

Definition 7-6:

Es sei $f: M \subset \mathbb{R}^n \to \mathbb{R}$ mit $M = D(f)$ eine reellwertige Funktion mehrerer Variabler.

a) f heißt stetig in $\mathbf{x}_0 \in M$ genau dann, wenn für jedes beliebig kleine $\varepsilon > 0$ ein $\delta_\varepsilon > 0$ existiert mit $|f(\mathbf{x}) - f(\mathbf{x}_0)| < \varepsilon$ für alle $\|\mathbf{x} - \mathbf{x}_0\|_2 < \delta_\varepsilon$,

b) f heißt stetig auf M genau dann, wenn f stetig in allen $\mathbf{x}_0 \in M$.

Beispiel:

$f: \mathbb{R}^2 \to \mathbb{R}$ mit $f(\mathbf{x}) = f(x_1, x_2) = x_1$

f ist stetig auf $\mathbb{R}^2$, denn $\mathbf{x}_0 = (x_{1,0}, x_{2,0})^T \in \mathbb{R}^2$, $\varepsilon > 0$

$$\Rightarrow \left| f(x_1, x_2) - f(x_{1,0}, x_{2,0}) \right| = \left| x_1 - x_{1,0} \right| \leq \|\mathbf{x} - \mathbf{x}_0\|_2 < \varepsilon \;\; \forall \|\mathbf{x} - \mathbf{x}_0\|_2 < \delta_\varepsilon = \varepsilon$$

Oft kann die Stetigkeit einfacher mit Hilfe von Folgen nachgewiesen werden.

Satz 7-5:

$f: M \subset \mathbb{R}^n \to \mathbb{R}$ ist stetig in $\mathbf{x} \in M$ genau dann, wenn für alle Folgen $(\mathbf{x}_k)$ aus M mit $\lim\limits_{k\to\infty} \mathbf{x}_k = \mathbf{x}$ gilt $\lim\limits_{k\to\infty} f(\mathbf{x}_k) = f(\mathbf{x})$.

Satz 7-6:

a) Summe, Differenz, Produkt und Quotient von in $\mathbf{x}$ stetigen Funktionen sind stetig in $\mathbf{x}$, wobei die Nennerfunktion des Quotienten in $\mathbf{x}$ keine Nullstelle besitzen darf.

b) Ist g stetig in $\mathbf{x}$ und f stetig in $g(\mathbf{x})$, dann ist $f\big(g(\mathbf{x})\big)$ stetig in $\mathbf{x}$.

Beispiel:

1) $f: \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = xy$

f ist stetig auf $\mathbb{R}^2$, denn $f = f_1 \cdot f_2$ mit $f_1 = f_1(x,y) = x$, $f_2 = f_2(x,y) = y$ und f_1, f_2 stetig auf $\mathbb{R}^2$ $\Rightarrow$ f ist stetig auf $\mathbb{R}^2$.

2) $f:\mathbb{R}^n \to \mathbb{R}$, f ist ein Polynom im $\mathbb{R}^n$, d.h.

$$f(x_1,...,x_n) = \sum_{i_1=0}^{k} \cdots \sum_{i_n=0}^{k} a_{i_1,i_2,\cdots,i_n} x_1^{\,i_1} x_2^{\,i_2} \cdots x_n^{\,i_n}$$

$\Rightarrow$ f ist stetig auf $\mathbb{R}^n$, da Produkt und Summe stetiger Funktionen.

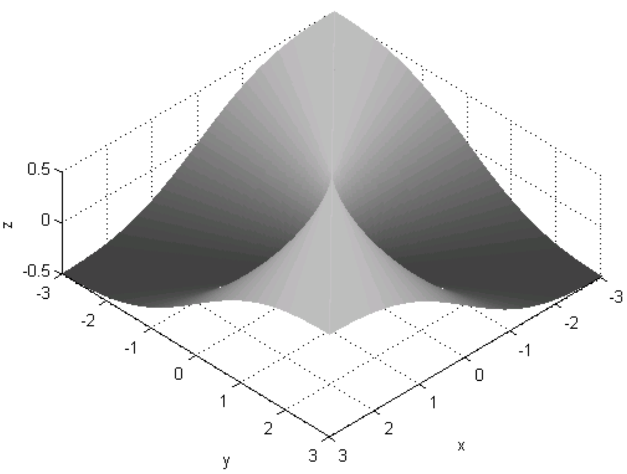
3) $f:\mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{für } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ 0 & \text{für } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{cases}$

f ist stetig auf $\mathbb{R}^2 \setminus \{\mathbf{0}\}$, da Quotient stetiger Funktionen mit Nenner $\neq 0$, aber f ist nicht stetig in 0, denn

$$\mathbf{x}_k = \begin{pmatrix} 1/k \\ 1/k \end{pmatrix} \xrightarrow[k\to\infty]{} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ aber } f(1/k, 1/k) = 1/2 \not\to f(0) = 0.$$

Anmerkung: (*partielle Stetigkeit*)
Betrachtet man die speziellen Folgen $\mathbf{x}_k = (x_{1,k}, 0)^T$ mit $x_{1,k} \to 0$ so gilt $\lim\limits_{k\to\infty} f(\mathbf{x}_k) = 0 = f(\mathbf{0})$.

Analog gilt für alle Folgen $\mathbf{x}_k = (0, x_{2,k})^T$ mit $x_{2,k} \to 0$ auch $\lim\limits_{k\to\infty} f(\mathbf{x}_k) = 0 = f(\mathbf{0})$.



Man spricht in diesem Fall von partieller Stetigkeit in $\mathbf{x} = \mathbf{0}$, d.h. Stetigkeit bzgl. der einzelnen Variablen bei Festhalten der anderen Variablen.

Das Beispiel zeigt, dass aus der partiellen Stetigkeit im allgemeinen nicht die Stetigkeit folgt.

7.5 Richtungsableitung, partielle Ableitung

Definition 7-7:

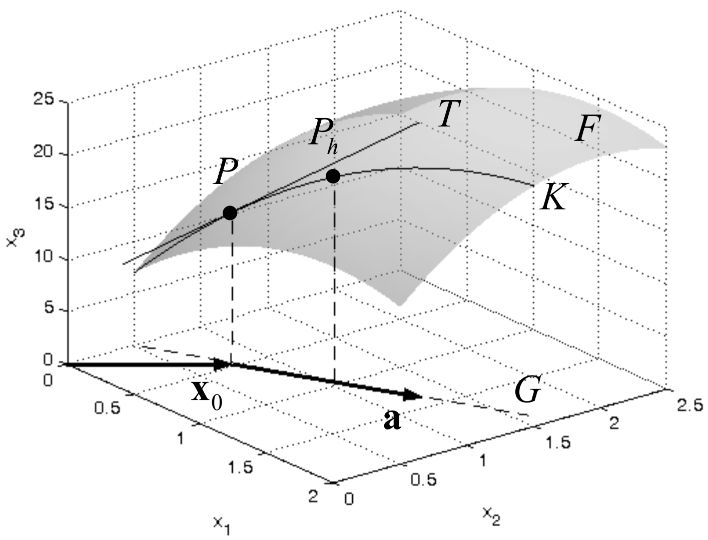
Es sei $f: M \subset \mathbb{R}^n \to \mathbb{R}$ mit $\mathbf{x}_0 \in M$ ein innerer Punkt und $\mathbf{a} \in \mathbb{R}^n$ ein fester Vektor mit $\|\mathbf{a}\|_2 = 1$. f heißt in $\mathbf{x}_0$ differenzierbar in Richtung $\mathbf{a}$, wenn der Grenzwert

$$\lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\mathbf{a}) - f(\mathbf{x}_0)}{h}$$

existiert. Diesen Grenzwert nennt man Richtungsableitung von f an der Stelle $\mathbf{x}_0$ in Richtung $\mathbf{a}$ und schreibt

$$\frac{\partial f}{\partial \mathbf{a}}(\mathbf{x}_0).$$

Geometrische Deutung für $n = 2$



Gerade durch $\mathbf{x}_0$ in Richtung $\mathbf{a}$

$$G = \{\mathbf{x} = \mathbf{x}_0 + h\mathbf{a} : h \in \mathbb{R}\},$$

Fläche im $\mathbb{R}^3$

$$F = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ f(\mathbf{x}) \end{pmatrix} : \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in M \right\},$$

Kurve auf Fläche F

$$K = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ f(\mathbf{x}) \end{pmatrix} : \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in G \right\},$$

$$P = \begin{pmatrix} x_{1,0} \\ x_{2,0} \\ f(\mathbf{x}_0) \end{pmatrix} \text{ und } P_h = \begin{pmatrix} x_{1,0} + h\,a_1 \\ x_{2,0} + h\,a_2 \\ f(\mathbf{x}_0 + h\,\mathbf{a}) \end{pmatrix} \text{ liegen auf } K.$$

Demzufolge ist

$$\frac{f(\mathbf{x}_0 + h\,\mathbf{a}) - f(\mathbf{x}_0)}{h}$$

die Steigung der Sekante durch die Punkte P und P_h. Existiert der Grenzwert

$$\frac{\partial f}{\partial \mathbf{a}}(\mathbf{x}_0),$$

so ist dieser Grenzwert die Steigung der Tangente von K im Punkt P.

Beispiel:

$f : M \subset \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) = x^2 + x\cos y$. Gesucht wird die Richtungsableitung von f in $(x_0,y_0)^T$ in Richtung $\mathbf{a} = 1/\sqrt{2} \cdot (1,1)^T$.

$$\lim_{h\to 0} \frac{f(x_0 + ha_1, y_0 + ha_2) - f(x_0, y_0)}{h}$$

$$= \lim_{h\to 0} \frac{f(x_0 + h/\sqrt{2}, y_0 + h/\sqrt{2}) - f(x_0, y_0)}{h}$$

$$= \lim_{h\to 0} \frac{1}{h}\Big((x_0 + h/\sqrt{2})^2 + (x_0 + h/\sqrt{2})\cos(y_0 + h/\sqrt{2}) - x_0^2 - x_0 \cos y_0\Big)$$

$$= \lim_{h\to 0} \frac{1}{h}\left(\sqrt{2}hx_0 + \frac{h^2}{2} + x_0\left(\cos\left(y_0 + \frac{h}{\sqrt{2}}\right) - \cos y_0\right) + \frac{h}{\sqrt{2}}\cos\left(y_0 + \frac{h}{\sqrt{2}}\right)\right)$$

$$= \lim_{h\to 0} \frac{1}{h}\left(\sqrt{2}hx_0 + \frac{h^2}{2}\right) + \lim_{h\to 0} \frac{x_0}{\sqrt{2}} \frac{\cos(y_0 + h/\sqrt{2}) - \cos y_0}{h/\sqrt{2}} + \lim_{h\to 0} \frac{\cos(y_0 + h/\sqrt{2})}{\sqrt{2}}$$

$$= \sqrt{2}x_0 - \frac{x_0}{\sqrt{2}}\sin y_0 + \frac{1}{\sqrt{2}}\cos y_0 = \frac{\partial f}{\partial \mathbf{a}}(x_0, y_0).$$

Für die Anwendung besonders wichtig sind die Richtungsableitungen in die Richtung der natürlichen Basisvektoren

$$\mathbf{e}_i = (0,\ldots,0,1,0,\ldots,0)^T \quad i = 1,2,\ldots,n.$$

Definition 7-8:

Existiert in $\mathbf{x}_0$ die Richtungsableitung in Richtung $\mathbf{e}_i$ so heißt f in $\mathbf{x}_0$ partiell differenzierbar nach der i-ten Variable x_i und man schreibt

$$f_{x_i}(\mathbf{x}_0) = \frac{\partial f}{\partial x_i}(\mathbf{x}_0) := \frac{\partial f}{\partial \mathbf{e}_i}(\mathbf{x}_0)$$

und nennt dies die partielle Ableitung von f nach x_i an der Stelle $\mathbf{x}_0$.

Anmerkung:

Da $\mathbf{x} + h\,\mathbf{e}_i = (x_1, x_2, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n)^T$ wird bei der Bildung der partiellen Ableitung nach x_i nur die i-te Variable variiert, alle anderen Variablen bleiben unverändert.

Also erhält man die partielle Ableitung nach x_i, indem man die Variablen $x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ konstant hält während man nach der i-ten Variablen x_i differenziert.

Beispiel:

1) $f(x,y) = x^2 + x\cos y$

$$f_x(x,y) = \frac{\partial f}{\partial x}(x,y) = 2x + \cos y,$$

$$f_y(x,y) = \frac{\partial f}{\partial y}(x,y) = -x\sin y.$$

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

Für $(x,y)^T \neq (0,0)^T$:

$$f_x(x,y) = \frac{y(x^2+y^2) - xy\cdot 2x}{(x^2+y^2)^2} = \frac{y}{x^2+y^2} - \frac{2x^2 y}{(x^2+y^2)^2},$$

$$f_y(x,y) = \frac{x(x^2+y^2) - xy\cdot 2y}{(x^2+y^2)^2} = \frac{x}{x^2+y^2} - \frac{2xy^2}{(x^2+y^2)^2}.$$

Für $(x,y)^T = (0,0)^T$:

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(0+h,0) - f(0,0)}{h} = 0,$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{h\to 0} \frac{f(0,0+h) - f(0,0)}{h} = 0.$$

$\Rightarrow$ Die partiellen Ableitungen existieren in jedem Punkt $(x,y)^T \in \mathbb{R}^2$.
Die Funktion ist aber in $(0,0)^T$ nicht stetig.

Anmerkung:
Aus der Existenz aller partiellen Ableitungen folgt im allgemeinen nicht die Stetigkeit einer Funktion.

Satz 7-7:
Es sei $f: M \subset \mathbb{R}^n \to \mathbb{R}$ und $\mathbf{x}_0 \in M$. Existieren in einer Umgebung $U_r(\mathbf{x}_0) \subset M$ alle partiellen Ableitungen $f_{x_1}, f_{x_2}, ..., f_{x_n}$ und sind diese dort beschränkt, dann gilt f ist stetig in $\mathbf{x}_0$.

Beweis:
Sei $\mathbf{v} = (v_1, ..., v_n)^T \in \mathbb{R}^n$ mit $\|\mathbf{v}\|_2 < r$, also $\mathbf{x}_0 + \mathbf{v} \in U_r(\mathbf{x}_0)$. Setze $\mathbf{y}_0 = \mathbf{0}$, $\mathbf{y}_i = (v_1, ..., v_i, 0, \dots, 0)^T$ $(i = 1, 2, ..., n)$ $\Rightarrow$ $\|\mathbf{y}_i\|_2 \leq \|\mathbf{v}\|_2 < r$ und $\mathbf{x}_0 + \mathbf{y}_i \in U_r(\mathbf{x}_0)$.
Dann gilt

$$f(\mathbf{x}_0 + \mathbf{v}) - f(\mathbf{x}_0) = \sum_{i=1}^{n} \big(f(\mathbf{x}_0 + \mathbf{y}_i) - f(\mathbf{x}_0 + \mathbf{y}_{i-1}) \big) = \sum_{i=1}^{n} v_i\, f_{x_i}(\boldsymbol{\xi}_i)$$

mit ξ_i auf der Strecke zwischen $\mathbf{x}_0 + \mathbf{y}_i$ und $\mathbf{x}_0 + \mathbf{y}_{i-1}$ (nach Mittelwertsatz der Differentialrechnung für eine Variable).

Aus der Beschränktheit von $f_{x_i} < k_i$ auf $U_r(\mathbf{x}_0)$ folgt

$$\left| f(\mathbf{x}_0 + \mathbf{v}) - f(\mathbf{x}_0) \right| \le \left| \sum_{i=1}^{n} v_i\, k_i \right| = \mathbf{k}^T \mathbf{v} \le \|\mathbf{k}\|_2 \, \|\mathbf{v}\|_2 = K \, \|\mathbf{v}\|_2$$

$$\Rightarrow \lim_{\mathbf{v} \to \mathbf{0}} f(\mathbf{x}_0 + \mathbf{v}) = f(\mathbf{x}_0) \;\Rightarrow\; f \text{ ist stetig in } \mathbf{x}_0.$$

Definition 7-9:

Ist $f: M \subset \mathbb{R}^n \to \mathbb{R}$ in $\mathbf{x}_0 \in M$ partiell differenzierbar nach allen Variablen x_i $(i = 1, 2, ..., n)$, so heißt der Vektor

$$\operatorname{grad} f(\mathbf{x}_0) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \frac{\partial f}{\partial x_2}(\mathbf{x}_0), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right)^T = \left(f_{x_1}(\mathbf{x}_0), \ldots, f_{x_n}(\mathbf{x}_0) \right)^T$$

der Gradient von f in $\mathbf{x}_0$.

Den Differentialoperator

$$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n} \right)^T$$

mit

$$\nabla f(\mathbf{x}_0) = \operatorname{grad} f(\mathbf{x}_0)$$

bezeichnet man als Nabla-Operator.

Beispiel:

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

$$\nabla f(x, y) = \operatorname{grad} f(x, y) = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix} = \begin{pmatrix} 2x + \cos y \\ -x \sin y \end{pmatrix}.$$

Die Verallgemeinerung der Differenzierbarkeit im $\mathbb{R}^n$ ergibt sich nicht aus der Existenz aller partiellen Ableitungen.

Definition 7-10: (*Landau-Symbol*)
Für zwei Funktionen $f, g : M \subset \mathbb{R}^n \to \mathbb{R}$ und $\mathbf{x}_0 \in M$, $k \in \mathbb{N}$ schreibt man

$$f(\mathbf{x}) = g(\mathbf{x}) + o\left(\|\mathbf{x}-\mathbf{x}_0\|_2^k\right) \text{ für } \mathbf{x} \to \mathbf{x}_0$$

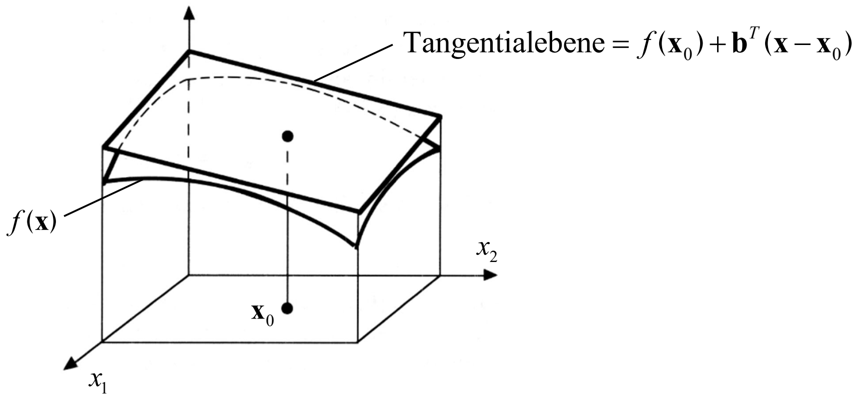
falls

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f(\mathbf{x}) - g(\mathbf{x})}{\|\mathbf{x}-\mathbf{x}_0\|_2^k} = 0.$$

Definition 7-11:
Es sei $f : M \subset \mathbb{R}^n \to \mathbb{R}$ und $\mathbf{x}_0$ ein innerer Punkt von M. Dann heißt f in $\mathbf{x}_0$ (total) differenzierbar (oder linear approximierbar), wenn ein Vektor $\mathbf{b} \in \mathbb{R}^n$ und eine Umgebung $U_r(\mathbf{x}_0) \subset M$ existiert mit

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{b}^T(\mathbf{x}-\mathbf{x}_0) + o\left(\|\mathbf{x}-\mathbf{x}_0\|_2\right)$$

für $\mathbf{x} \in U_r(\mathbf{x}_0)$.



Satz 7-8:
Ist $f : M \subset \mathbb{R}^n \to \mathbb{R}$ in $\mathbf{x}_0 \in M$ (total) differenzierbar, d.h.

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{b}^T(\mathbf{x}-\mathbf{x}_0) + o\left(\|\mathbf{x}-\mathbf{x}_0\|_2\right),$$

so existieren alle Richtungsableitungen und damit alle partiellen Ableitungen und es gilt

$$\text{a) } \mathbf{b} = \operatorname{grad} f(\mathbf{x}_0), \qquad \text{b) } \frac{\partial f}{\partial \mathbf{a}}(\mathbf{x}_0) = \mathbf{a}^T \operatorname{grad} f(\mathbf{x}_0), \ \|\mathbf{a}\|_2 = 1.$$

Beweis:

Sei $\mathbf{x} = \mathbf{x}_0 + h\mathbf{a},\ h \in \mathbb{R}\backslash\{0\} \ \Rightarrow\ \|\mathbf{x} - \mathbf{x}_0\|_2 = |h|\ \|\mathbf{a}\|_2 = |h|$. Da

$$\lim_{h\to 0} \frac{f(\mathbf{x}_0+h\mathbf{a})-f(\mathbf{x}_0)-\mathbf{b}^T(h\mathbf{a})}{h}=0 \ \Rightarrow\ \frac{\partial f}{\partial \mathbf{a}}(\mathbf{x}_0)=\lim_{h\to 0}\frac{f(\mathbf{x}_0+h\mathbf{a})-f(\mathbf{x}_0)}{h}=\mathbf{b}^T\mathbf{a}$$

mit $\mathbf{a} = \mathbf{e}_i$ folgt

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0)=\mathbf{b}^T\mathbf{e}_i=b_i \ \Rightarrow\ \operatorname{grad} f(\mathbf{x}_0)=\mathbf{b} \ \text{ und } \ \frac{\partial f}{\partial \mathbf{a}}(\mathbf{x}_0)=\mathbf{a}^T\mathbf{b}=\mathbf{a}^T\operatorname{grad} f(\mathbf{x}_0)$$

Im Falle der Differenzierbarkeit liefert b) eine einfache Methode, die Richtungsableitungen zu bestimmen.

Beispiel:

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

f ist differenzierbar auf $\mathbb{R}^2$ (siehe nächstes Beispiel).

$$\operatorname{grad} f(x, y) = \begin{pmatrix} 2x + \cos y \\ -x\sin y \end{pmatrix},\ \mathbf{a} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\frac{\partial f}{\partial \mathbf{a}}(x, y) = \mathbf{a}^T \operatorname{grad} f(x, y) = \frac{1}{\sqrt{2}}\left(2x + \cos y - x\sin y\right).$$

<u>Geometrische Deutung des Gradienten für $n = 2$</u>

$$\frac{\partial f}{\partial \mathbf{a}}(\mathbf{x}_0) = \mathbf{a}^T\operatorname{grad} f(\mathbf{x}_0) = \|\mathbf{a}\|_2\ \|\operatorname{grad} f(\mathbf{x}_0)\|_2 \cos\alpha$$

$$= \|\operatorname{grad} f(\mathbf{x}_0)\|_2 \cos\alpha,\ \text{da}\ \|\mathbf{a}\|_2 = 1$$

α bezeichnet den Winkel zwischen $\mathbf{a}$ und $\operatorname{grad} f(\mathbf{x}_0)$.

Die Richtungsableitung wird am größten, wenn $\alpha = 0$, d.h. wenn $\mathbf{a}$ in die Richtung des Gradienten zeigt. Demzufolge zeigt $\operatorname{grad} f(\mathbf{x}_0)$ an der Stelle $\mathbf{x}_0$ in die Richtung des stärksten Anstiegs der Fläche, die durch f dargestellt wird.

Satz 7-9:
Ist $f: M \subset \mathbb{R}^n \to \mathbb{R}$ in $\mathbf{x}_0$ (total) differenzierbar, dann ist f in $\mathbf{x}_0$ auch stetig.

Beweis:
Für $\mathbf{x} \neq \mathbf{x}_0$ gilt

$$\frac{f(\mathbf{x})-f(\mathbf{x}_0)}{\|\mathbf{x}-\mathbf{x}_0\|_2} = \frac{f(\mathbf{x})-f(\mathbf{x}_0)-\mathbf{b}^T(\mathbf{x}-\mathbf{x}_0)}{\|\mathbf{x}-\mathbf{x}_0\|_2} + \frac{\mathbf{b}^T(\mathbf{x}-\mathbf{x}_0)}{\|\mathbf{x}-\mathbf{x}_0\|_2}.$$

Da f in $\mathbf{x}_0$ (total) differenzierbar, existiert $U_r(\mathbf{x}_0) \subset M$ und $k > 0$ mit

$$\frac{|f(\mathbf{x})-f(\mathbf{x}_0)|}{\|\mathbf{x}-\mathbf{x}_0\|_2} \leq k + \frac{|\mathbf{b}|\|\mathbf{x}-\mathbf{x}_0\|_2}{\|\mathbf{x}-\mathbf{x}_0\|_2} = k + |\mathbf{b}| = K > 0$$

für alle $\mathbf{x} \in U_r(\mathbf{x}_0) \setminus \{\mathbf{x}_0\}$ (Schwarzsche Ungleichung)

$$\Rightarrow \; |f(\mathbf{x})-f(\mathbf{x}_0)| \leq K\|\mathbf{x}-\mathbf{x}_0\|_2 \quad \forall \mathbf{x} \in U_r(\mathbf{x}_0) \;\Rightarrow\; \lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0).$$

Anmerkung:
Aus der Existenz aller partiellen Ableitungen an der Stelle $\mathbf{x}_0$ folgt nicht die (totale) Differenzierbarkeit in $\mathbf{x}_0$, da daraus noch nicht einmal die Stetigkeit folgt.

Satz 7-10:
Existieren für $f: M \subset \mathbb{R}^n \to \mathbb{R}$ in einer Umgebung $U_r(\mathbf{x}_0) \subset M$ von $\mathbf{x}_0$ alle partiellen Ableitungen $\partial f/\partial x_1, \partial f/\partial x_2, \ldots, \partial f/\partial x_n$ und sind diese stetig in $\mathbf{x}_0$, so ist f in $\mathbf{x}_0$ (total) differenzierbar.

Beweis:
Sei $\mathbf{v} = (v_1,\ldots,v_n)^T \in \mathbb{R}^n$ mit $\|\mathbf{v}\|_2 < r$, also $\mathbf{x}_0 + \mathbf{v} \in U_r(\mathbf{x}_0)$. Setze $\mathbf{y}_0 = \mathbf{0}$, $\mathbf{y}_i = (v_1,\ldots,v_i,0,\ldots,0)^T$ $(i = 1,2,\ldots,n)$ $\Rightarrow$ $\|\mathbf{y}_i\|_2 \leq \|\mathbf{v}\|_2 < r$ und $\mathbf{x}_0 + \mathbf{y}_i \in U_r(\mathbf{x}_0)$.
Dann gilt

$$f(\mathbf{x}_0+\mathbf{v})-f(\mathbf{x}_0)=\sum_{i=1}^{n}\left(f(\mathbf{x}_0+\mathbf{y}_i)-f(\mathbf{x}_0+\mathbf{y}_{i-1})\right)=\sum_{i=1}^{n}v_i\,f_{x_i}(\boldsymbol{\xi}_i)$$

mit $\boldsymbol{\xi}_i$ auf der Strecke zwischen $\mathbf{x}_0+\mathbf{y}_i$ und $\mathbf{x}_0+\mathbf{y}_{i-1}$ und man erhält

$$f(\mathbf{x}_0+\mathbf{v})-f(\mathbf{x}_0)-\mathbf{v}^T\operatorname{grad}f(\mathbf{x}_0)=\sum_{i=1}^{n}v_i\,f_{x_i}(\boldsymbol{\xi}_i)-\sum_{i=1}^{n}v_i\,f_{x_i}(\mathbf{x}_0)$$

$$\Rightarrow\ \frac{\left|f(\mathbf{x}_0+\mathbf{v})-f(\mathbf{x}_0)-\mathbf{v}^T\operatorname{grad}f(\mathbf{x}_0)\right|}{\|\mathbf{v}\|_2}\le\sum_{i=1}^{n}\frac{|v_i|}{\|\mathbf{v}\|_2}\left|f_{x_i}(\boldsymbol{\xi}_i)-f_{x_i}(\mathbf{x}_0)\right|$$

$$\le\sum_{i=1}^{n}\left|f_{x_i}(\boldsymbol{\xi}_i)-f_{x_i}(\mathbf{x}_0)\right|.$$

Für $\mathbf{v}\to\mathbf{0}$ gilt $\boldsymbol{\xi}_i\to\mathbf{x}_0$ und wegen der Stetigkeit von f_{x_i} gilt $f_{x_i}(\boldsymbol{\xi}_i)\to f_{x_i}(\mathbf{x}_0)$ für alle $i=1,2,...,n$. Also konvergiert die rechte Seite gegen 0 und damit auch die linke Seite.

Beispiel:

1) $f(x,y)=x^2+x\cos y$

$f_x(x,y)=2x+\cos y,\ \ f_y(x,y)=-x\sin y$

$\Rightarrow\ f_x,f_y$ sind stetig auf $\mathbb{R}^2$

$\Rightarrow\ f(x,y)$ ist (total) differenzierbar auf $\mathbb{R}^2$.

2) $$f(x,y)=\begin{cases}\dfrac{x^3}{x^2+y^2} & \text{für } \begin{pmatrix}x\\y\end{pmatrix}\ne\begin{pmatrix}0\\0\end{pmatrix}\\ 0 & \text{für } \begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\end{cases}$$

f ist stetig auf $\mathbb{R}^2\backslash\{\mathbf{0}\}$, da Quotient zweier auf $\mathbb{R}^2$ stetiger Funktionen und der Nenner $\ne 0$ auf $\mathbb{R}^2\backslash\{\mathbf{0}\}$.

f ist auch stetig in 0, denn

$$\left|f(x,y)-f(0,0)\right| = |x|\left|\frac{x^2}{x^2+y^2}\right| \le |x| \to 0 \text{ für } \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$\Rightarrow$ f ist stetig auf $\mathbb{R}^2$.

$$f_x(x,y) = \frac{3x^2(x^2+y^2)-x^3 2x}{(x^2+y^2)^2} = \frac{x^4+3x^2y^2}{(x^2+y^2)^2} \text{ falls } \begin{pmatrix} x \\ y \end{pmatrix} \ne \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$f_y(x,y) = \frac{-2x^3y}{(x^2+y^2)^2} \text{ falls } \begin{pmatrix} x \\ y \end{pmatrix} \ne \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left.\begin{aligned} f_x(0,0) &= \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{h^3}{h^3} = 1 \\ f_y(0,0) &= \lim_{h\to 0}\frac{f(0,h)-f(0,0)}{h} = \lim_{h\to 0}\frac{0}{h^3} = 0 \end{aligned}\right\} \Rightarrow \operatorname{grad} f(0,0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\lim_{h\to 0} f_x(0,h) = \frac{0}{h^4} = 0 \ne \lim_{h\to 0} f_x(h,0) = \frac{h^4}{h^4} = 1 \;\Rightarrow\; f_x \text{ ist nicht stetig in } (0,0)^T.$$

f ist nicht total differenzierbar, denn

<u>Annahme:</u> f sei (total) differenzierbar in $(0,0)^T$

$$\Rightarrow \lim_{(x,y)^T\to\mathbf{0}} \frac{f(x,y)-f(0,0)-(x,y)\operatorname{grad} f(0,0)}{\left|(x,y)^T\right|} = 0$$

$$\Rightarrow \lim_{(x,y)^T\to\mathbf{0}} \frac{x^3/(x^2+y^2)-0-x}{\sqrt{x^2+y^2}} = 0,$$

z.B.: $(x,y)^T = (h,h)^T$ mit $h>0$

$$\Rightarrow \lim_{h\to 0}\frac{h^3/(2h^2)-h}{\sqrt{2h^2}} = \lim_{h\to 0}\frac{-h/2}{h\sqrt{2}} = -\frac{1}{2\sqrt{2}} \ne 0 \text{ Widerspruch.}$$

3) $$f(x,y)=\begin{cases}\dfrac{x^4}{x^2+y^2} & \text{für } \begin{pmatrix}x\\y\end{pmatrix}\neq\begin{pmatrix}0\\0\end{pmatrix}\\ 0 & \text{für } \begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\end{cases}$$

$$f_x(x,y)=\frac{4x^3(x^2+y^2)-x^4 2x}{(x^2+y^2)^2}=\frac{2x^5+4x^3y^2}{(x^2+y^2)^2} \quad \text{falls } \begin{pmatrix}x\\y\end{pmatrix}\neq\begin{pmatrix}0\\0\end{pmatrix}$$

$\Rightarrow$ f_x ist stetig auf $\mathbb{R}^2\backslash\{\mathbf{0}\}$.

$$f_x(0,0)=\lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{h^4}{h^3}=\lim_{h\to 0}h=0$$

$$\left|f_x(x,y)-f_x(0,0)\right|\leq 2|x|\left|\frac{x^4+2x^2y^2}{(x^2+y^2)^2}\right|\leq 2|x|\to 0 \quad \text{für } \begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}0\\0\end{pmatrix}$$

$\Rightarrow$ f_x ist stetig in $\mathbf{0}$ $\Rightarrow$ f_x ist stetig auf $\mathbb{R}^2$.

$$f_y(x,y)=\frac{-x^4 2y}{(x^2+y^2)^2} \quad \text{falls } \begin{pmatrix}x\\y\end{pmatrix}\neq\begin{pmatrix}0\\0\end{pmatrix}$$

$\Rightarrow$ f_y ist stetig auf $\mathbb{R}^2\backslash\{\mathbf{0}\}$.

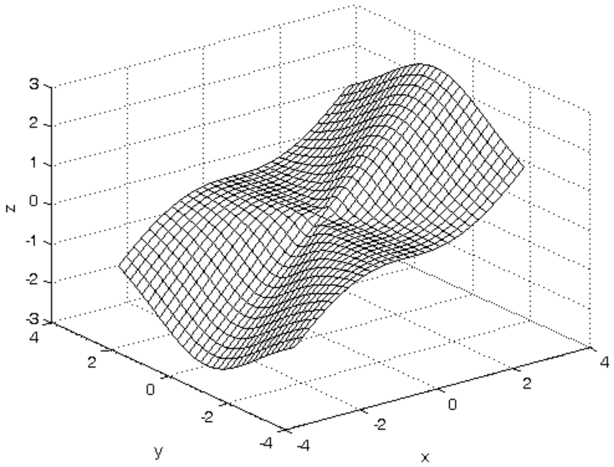
$$f_y(0,0)=\lim_{h\to 0}\frac{f(0,h)-f(0,0)}{h}=0$$

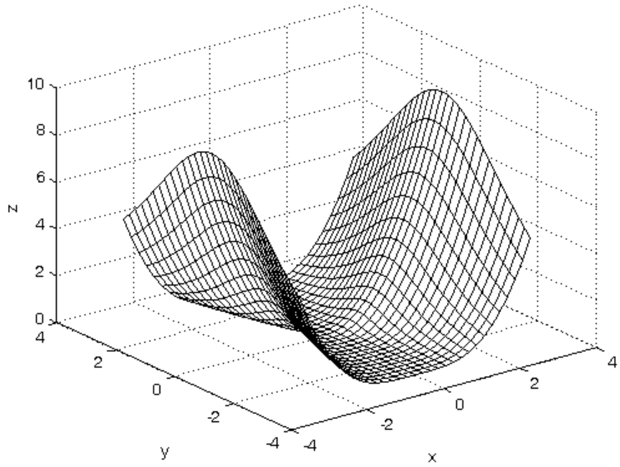
$$\left|f_y(x,y)-f_y(0,0)\right|\leq 2|y|\left|\frac{x^4}{(x^2+y^2)^2}\right|\leq 2|y|\to 0 \quad \text{für } \begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}0\\0\end{pmatrix}$$

$\Rightarrow$ f_y ist stetig in $\mathbf{0}$ $\Rightarrow$ f_y ist stetig auf $\mathbb{R}^2$

$\Rightarrow$ f ist (total) differenzierbar auf $\mathbb{R}^2$.

$$f(x,y)=\frac{x^3}{x^2+y^2} \qquad\qquad f(x,y)=\frac{x^4}{x^2+y^2}$$





7.6 Differentiation von Parameterintegralen

<u>Satz 7-11:</u>

Es sei

$$I=\left\{(x,t)^T\in\mathbb{R}^2:\ a\le x\le b,\ c\le t\le d\right\}$$

und $f: I\to\mathbb{R}$ stetig auf I. Ferner sei $F:[a,b]\to\mathbb{R}$ durch

$$F(x)=\int_c^d f(x,t)\,dt$$

definiert. Dann gilt

a) F ist stetig auf $[a,b]$

b) ist zusätzlich f auf $[a,b]$ stetig partiell nach x differenzierbar so ist F auf $[a,b]$ differenzierbar mit

$$F'(x)=\frac{d}{dx}\int_c^d f(x,t)\,dt=\int_c^d\frac{\partial f}{\partial x}(x,t)\,dt=\int_c^d f_x(x,t)\,dt.$$

Beispiel:

$$F(x)=\int_0^1 \sin(x-t)\cos t\,dt$$

Der Integrand ist stetig differenzierbar auf $\mathbb{R}^2$

$$\Rightarrow\; F'(x)=\int_0^1 \frac{\partial}{\partial x}\sin(x-t)\cos t\,dt=\int_0^1 \cos(x-t)\cos t\,dt.$$

Kriterium für gleichmäßige Konvergenz

Gilt $|f(x,t)|\le g(t)$ für alle $x\in[a,b]$ und $t\ge c$ und konvergiert $\int_c^\infty g(t)\,dt$ so ist das Integral $\int_c^\infty f(x,t)\,dt$ gleichmäßig konvergent auf $[a,b]$.

Satz 7-12:

Es sei

$$I=\left\{(x,t)^T\in\mathbb{R}^2:\; a\le x\le b,\; c\le t<\infty\right\}$$

und $f: I\to\mathbb{R}$ stetig auf I. Ferner sei $F:[a,b]\to\mathbb{R}$ durch

$$F(x)=\int_c^\infty f(x,t)\,dt$$

definiert und das Integral $\int_c^\infty f(x,t)\,dt$ gleichmäßig konvergent auf $[a,b]$. Dann gilt

a) F ist stetig auf $[a,b]$ mit

$$\lim_{x\to x_0}F(x)=\lim_{x\to x_0}\int_c^\infty f(x,t)\,dt=\int_c^\infty \lim_{x\to x_0}f(x,t)\,dt=\int_c^\infty f(x_0,t)\,dt,$$

(Grenzwert und Integration dürfen vertauscht werden)

b) ist zusätzlich f auf $[a,b]$ stetig partiell nach x differenzierbar und das Integral $\int_c^\infty f_x(x,t)\,dt$ gleichmäßig konvergent auf $[a,b]$, so ist F auf $[a,b]$ differenzierbar mit

$$F'(x)=\frac{d}{dx}\int_c^\infty f(x,t)\,dt=\int_c^\infty \frac{\partial}{\partial x}f(x,t)\,dt=\int_c^\infty f_x(x,t)\,dt.$$

(Differentiation und Integration dürfen vertauscht werden)

Beispiel:

$$F(x)=\int_0^\infty e^{-t^2}\cos(xt)\,dt$$

$$|f(x,t)|=\left|e^{-t^2}\cos(xt)\right|\le e^{-t^2}\quad \text{für alle } x\in\mathbb{R},\ t\ge 0$$

$$\int_0^\infty e^{-t^2}\,dt=\frac{\sqrt{\pi}}{2}\quad \text{konvergent,}$$

$$|f_x(x,t)|=\left|-t\,e^{-t^2}\sin(xt)\right|\le t\,e^{-t^2}\quad \text{für alle } x\in\mathbb{R},\ t\ge 0$$

$$\int_0^\infty t\,e^{-t^2}\,dt=-\frac{1}{2}e^{-t^2}\Big|_0^\infty=\frac{1}{2}\quad \text{konvergent,}$$

$$\Rightarrow\ \int_c^\infty f(x,t)\,dt\ \text{ und }\ \int_c^\infty f_x(x,t)\,dt\ \text{ sind gleichmäßig konvergent auf } \mathbb{R}.$$

$$\Rightarrow\ F'(x)=\int_0^\infty f_x(x,t)\,dt=-\int_0^\infty t\,e^{-t^2}\sin(xt)\,dt$$

$$=\frac{1}{2}e^{-t^2}\sin(xt)\Big|_0^\infty-\frac{x}{2}\int_0^\infty e^{-t^2}\cos(xt)\,dt=-\frac{x}{2}F(x),$$

$$\text{d.h. } \frac{F'(x)}{F(x)}=-\frac{x}{2}\ \Rightarrow\ \int\frac{F'(x)}{F(x)}\,dx=\int\frac{1}{F}\,dF=-\int\frac{x}{2}\,dx$$

$$\Rightarrow\ \ln|F(x)|=-\frac{x^2}{4}+\tilde{C}\ \Rightarrow\ F(x)=C\,e^{-x^2/4},$$

$$\text{da } F(0)=C=\lim_{x\to 0}\int_0^\infty e^{-t^2}\cos(xt)\,dt=\int_0^\infty e^{-t^2}\,dt=\frac{\sqrt{\pi}}{2}$$

$$\Rightarrow\ F(x)=\int_0^\infty e^{-t^2}\cos(xt)\,dt=\frac{\sqrt{\pi}}{2}\,e^{-x^2/4}.$$

7.7 Partielle Ableitungen höherer Ordnung

Es sei $f: M \subset \mathbb{R}^n \to \mathbb{R}$. Existiert in einer Teilmenge $D \subset M$ die partielle Ableitung $\partial f / \partial x_k = f_{x_k}$, so ist $f_{x_k}: D \subset \mathbb{R}^n \to \mathbb{R}$ eine Funktion auf D. Ist diese Funktion in $D_1 \subset D$ partiell differenzierbar nach x_l, so schreibt man

$$\frac{\partial f_{x_k}}{\partial x_l} = \frac{\partial}{\partial x_l}\left(\frac{\partial f}{\partial x_k}\right) = \frac{\partial^2 f}{\partial x_k \partial x_l} = f_{x_k x_l}.$$

Ist $f_{x_k x_l}$ wiederum in $D_2 \subset D_1$ partiell differenzierbar nach x_m, so schreibt man

$$\frac{\partial f_{x_k x_l}}{\partial x_m} = \frac{\partial}{\partial x_m}\left(\frac{\partial^2 f}{\partial x_k \partial x_l}\right) = \frac{\partial^3 f}{\partial x_k \partial x_l \partial x_m} = f_{x_k x_l x_m}.$$

Diesen Prozess kann man fortsetzen und man erhält die allgemeine partielle Ableitung m-ter Ordnung

$$\frac{\partial^m f}{\partial x_{l_1} \partial x_{l_2} \cdots \partial x_{l_m}} = f_{x_{l_1} x_{l_2} \ldots x_{l_m}} \quad \text{mit } l_1, l_2, \ldots, l_m \in \{1, 2, \cdots, n\}.$$

Definition 7-12:

Es sei $f: M \subset \mathbb{R}^n \to \mathbb{R}$.

a) Existiert in einem inneren Punkt $\mathbf{x}_0 \in D(f_{x_{l_1} x_{l_2} \ldots x_{l_{m-1}}})$ für ein l_m mit $m > 1$ die partielle Ableitung $\partial f_{x_{l_1} x_{l_2} \ldots x_{l_{m-1}}} / \partial x_{l_m} (\mathbf{x}_0)$, so heißt diese die partielle Ableitung m-ter Ordnung von f in $\mathbf{x}_0$ nach $x_{l_1} x_{l_2} \ldots x_{l_m}$ und man schreibt

$$f_{x_{l_1} x_{l_2} \ldots x_{l_m}}(\mathbf{x}_0) = \frac{\partial^m f}{\partial x_{l_1} \partial x_{l_2} \cdots \partial x_{l_m}}(\mathbf{x}_0).$$

b) Existieren für f in M alle partiellen Ableitungen bis zur m-ten Ordnung und sind diese auf M stetig, so heißt f auf M m-mal stetig differenzierbar und man schreibt

$$f \in C^m(M).$$

Beispiel:

1) $f(x,y,z) = 4xyz - x^2 + y^2$

$$f_x = 4yz - 2x, \quad f_y = 4xz + 2y, \quad f_z = 4xy$$

$$f_{xx} = -2, \quad f_{xy} = 4z, \quad f_{xz} = 4y$$

$$f_{yx} = 4z, \quad f_{yy} = 2 \quad f_{yz} = 4x$$

$$f_{zx} = 4y, \quad f_{zy} = 4x, \quad f_{zz} = 0$$

$\Rightarrow\ f \in C^2(\mathbb{R}^3)$, es gilt sogar $f \in C^\infty(\mathbb{R}^3)$, da die partiellen Ableitungen der Ordnung $l \geq 4$ alle gleich 0 sind.

In diesem Beispiel gilt

$f_{xy} = f_{yx},\ f_{xz} = f_{zx},\ f_{yz} = f_{zy}$

d.h. die Reihenfolge der partiellen Ableitungen ist vertauschbar.

Dass dies nicht immer gilt, zeigt das folgende Beispiel.

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

Es gilt $f_{xy}(0,0) = -1 \neq f_{yx}(0,0) = 1$, denn für $(x,y)^T = (0,0)^T$ ist

$$f_x(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = 0,$$

$$f_y(0,0) = \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} = 0$$

und für $(x,y)^T \neq (0,0)^T$ ist

$$f_x(x,y) = \frac{(3x^2y - y^3)(x^2 + y^2) - (x^3y - xy^3)\,2x}{(x^2 + y^2)^2} = \frac{x^4y - y^5 + 4x^2y^3}{(x^2 + y^2)^2},$$

$$f_y(x,y)=\frac{(x^3-3xy^2)(x^2+y^2)-(x^3y-xy^3)2y}{(x^2+y^2)^2}=\frac{x^5-xy^4-4x^3y^2}{(x^2+y^2)^2}$$

und damit schlussendlich

$$f_x(x,y)=\begin{cases} y\dfrac{x^4-y^4+4x^2y^2}{(x^2+y^2)^2} & \text{falls } \begin{pmatrix}x\\y\end{pmatrix}\neq\begin{pmatrix}0\\0\end{pmatrix} \\ 0 & \text{falls } \begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\end{cases},$$

$$f_y(x,y)=\begin{cases} x\dfrac{x^4-y^4-4x^2y^2}{(x^2+y^2)^2} & \text{falls } \begin{pmatrix}x\\y\end{pmatrix}\neq\begin{pmatrix}0\\0\end{pmatrix} \\ 0 & \text{falls } \begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\end{cases},$$

so dass

$$f_{xy}(0,0)=\lim_{h\to 0}\frac{f_x(0,h)-f_x(0,0)}{h}=\lim_{h\to 0}\frac{-h}{h}=-1,$$

$$f_{yx}(0,0)=\lim_{h\to 0}\frac{f_y(h,0)-f_y(0,0)}{h}=\lim_{h\to 0}\frac{h}{h}=1.$$

In diesem Fall sind f_{xy} und f_{yx} in $(0,0)^T$ nicht stetig.

<u>Satz 7-13:</u>

Ist $f\in C^m(M)$ mit $M\subset\mathbb{R}^n$ und $m\geq 2$, d.h. die partiellen Ableitungen m-ter Ordnung sind stetig auf M, so ist jede partielle Ableitung unabhängig von der Reihenfolge der x_l und es gilt

$$f_{x_{l_1}x_{l_2}\cdots x_{l_m}}=f_{x_{k_1}x_{k_2}\cdots x_{k_m}},$$

wenn $(k_1,k_2,...,k_m)$ eine Permutation von $(l_1,l_2,...,l_m)$ ist.

Beispiel:

$f:\mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y)=x^2+x\cos y$

$f_x = 2x+\cos y,\ f_y = -x\sin y,$

$f_{xx}=2,\ f_{yy}=-x\cos y,\ f_{xy}=-\sin y = f_{yx},$

$f_{xxx}=0,\ f_{yyy}=x\sin y,$

$f_{xxy}=f_{xyx}=f_{yxx}=0,$

$f_{xyy}=f_{yxy}=f_{yyx}=-\cos y.$

Da $\cos y, \sin y$ und x beliebig oft differenzierbar sind folgt $f\in C^\infty(\mathbb{R}^2)$.

Anmerkung:

Für $\dfrac{\partial^m f}{\partial x_l \partial x_l \cdots \partial x_l} = \dfrac{\partial}{\partial x_l}\left(\dfrac{\partial}{\partial x_l}\left(\cdots\left(\dfrac{\partial f}{\partial x_l}\right)\right)\right)$ schreibt man $\dfrac{\partial^m f}{\partial x_l^m}$.

7.8 Abbildungen von $\mathbb{R}^n \to \mathbb{R}^m$

Eine Abbildung $\mathbf{f}: M\subset\mathbb{R}^n \to \mathbb{R}^m$ mit

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

wobei $f_i: M\subset\mathbb{R}^n\to\mathbb{R}$ die i-te Koordinatenfunktion $(i=1,2,...,m)$ von $\mathbf{f}$ angibt, bezeichnet man als Vektorfeld.

Beispiel:

$$\mathbf{f}:\mathbb{R}^2\to\mathbb{R}^3 \text{ mit } \mathbf{f}(x,y)=\begin{pmatrix} 2x+y \\ 3x^2+y^2 \\ xy \end{pmatrix},$$

d.h. $f_1(x,y)=2x+y,\ f_2(x,y)=3x^2+y^2,\ f_3(x,y)=xy.$

Definition 7-13:

Es sei $\mathbf{f}: M \subset \mathbb{R}^n \to \mathbb{R}^m$ und $\mathbf{x}_0$ innerer Punkt von M, d.h. $\mathbf{x}_0 \in \mathring{M} = M \setminus \partial M$.

1) $\mathbf{f}$ heißt stetig in $\mathbf{x}_0$ genau dann, wenn für jedes $\varepsilon > 0$ ein $\delta_\varepsilon > 0$ existiert mit

$$\left\| \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0) \right\|_2 < \varepsilon \text{ für alle } \left\| \mathbf{x} - \mathbf{x}_0 \right\|_2 < \delta_\varepsilon, \; \mathbf{x} \in M.$$

2) $\mathbf{f}$ heißt (total) differenzierbar (linear approximierbar) genau dann, wenn es eine $m \times n$-Matrix $\mathbf{A} = (a_{ij})_{i=1,\ldots m;\, j=1,\ldots n}$ und eine Umgebung $U(\mathbf{x}_0) \subset M$ gibt, so dass

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \mathbf{A}(\mathbf{x} - \mathbf{x}_0) + \mathbf{o}\left(\left\| \mathbf{x} - \mathbf{x}_0 \right\|_2 \right)$$

für alle $\mathbf{x} \in U(\mathbf{x}_0)$ gilt.

Das eine Folge $(\mathbf{x}_k)_{k\in\mathbb{N}}$ aus $\mathbb{R}^n$ genau dann konvergiert, wenn jede Koordinatenfolge $(x_{i,k})_{k\in\mathbb{N}}$ $(i = 1,\ldots,n)$ konvergiert, motiviert den folgenden Satz.

Satz 7-14:

Es sei $\mathbf{f}: M \subset \mathbb{R}^n \to \mathbb{R}^m$ mit $\mathbf{f} = (f_1, f_2, \ldots, f_m)^T$ und $\mathbf{x}_0$ innerer Punkt von M. Dann gilt

1) $\mathbf{f}$ ist stetig in $\mathbf{x}_0$ genau dann, wenn jede Koordinatenfunktion $f_i(\mathbf{x})$ in $\mathbf{x}_0$ stetig ist.

2) $\mathbf{f}$ ist differenzierbar in $\mathbf{x}_0$ genau dann, wenn jede Koordinatenfunktion $f_i(\mathbf{x})$ in $\mathbf{x}_0$ differenzierbar ist.

Für die Matrix $\mathbf{A} = (a_{ij})$ gilt dann

$$a_{ij} = \frac{\partial f_i}{\partial x_j}, \quad i = 1,\ldots,m \quad j = 1,\ldots,n$$

Da diese Matrix in den Anwendungen häufig auftritt, ist die folgende Definition sinnvoll.

Definition 7-14:

Es sei $\mathbf{f}: M \subset \mathbb{R}^n \to \mathbb{R}^m$ mit $\mathbf{f} = (f_1, f_2, ..., f_m)^T$ und $\mathbf{x}_0$ innerer Punkt von M.

1) Existieren alle partiellen Ableitungen $\partial f_i / \partial x_j\,(\mathbf{x}_0)$ für $i = 1,2,...,m$ und $j = 1,2,...,n$ so heißt

$$\mathbf{A}(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}$$

Funktional- oder Jacobi-Matrix von f an der Stelle $\mathbf{x}_0$ und man schreibt

$$\mathbf{A}(\mathbf{x}_0) = \mathbf{J}(\mathbf{x}_0) = \frac{\partial(f_1, f_2, \ldots, f_m)}{\partial(x_1, x_2, \ldots, x_n)}(\mathbf{x}_0) = \frac{d\mathbf{f}}{d\mathbf{x}}(\mathbf{x}_0).$$

2) Ist $m = n$, so ist $\mathbf{A}(\mathbf{x}_0)$ eine quadratische Matrix, deren Determinante als Funktional- oder Jacobi-Determinante

$$\begin{aligned} \det\big(\mathbf{A}(\mathbf{x}_0)\big) &= \det\big(\mathbf{J}(\mathbf{x}_0)\big) \\ &= \det\left(\frac{\partial(f_1, f_2, \ldots, f_n)}{\partial(x_1, x_2, \ldots, x_n)}(\mathbf{x}_0)\right) = \det\left(\frac{d\mathbf{f}}{d\mathbf{x}}(\mathbf{x}_0)\right) \end{aligned}$$

und deren Spur als Divergenz

$$\operatorname{div}\mathbf{f}(\mathbf{x}_0) = \operatorname{Sp}\left(\frac{d\mathbf{f}}{d\mathbf{x}}(\mathbf{x}_0)\right) = \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) + \frac{\partial f_2}{\partial x_2}(\mathbf{x}_0) + \cdots + \frac{\partial f_n}{\partial x_n}(\mathbf{x}_0)$$

bezeichnet wird.

Anmerkung:

$$m=1: \qquad \mathbf{A}(\mathbf{x}_0)=\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\dots,\frac{\partial f}{\partial x_n}\right)=\left(\operatorname{grad} f(\mathbf{x}_0)\right)^T$$

$$n=1: \qquad \mathbf{A}(\mathbf{x}_0)=\left(\frac{\partial f_1}{\partial x},\frac{\partial f_2}{\partial x},\dots,\frac{\partial f_m}{\partial x}\right)^T$$

$$\text{allgemein}: \quad \mathbf{A}(\mathbf{x}_0)=\begin{pmatrix}\left(\operatorname{grad} f_1(\mathbf{x}_0)\right)^T\\ \left(\operatorname{grad} f_2(\mathbf{x}_0)\right)^T\\ \vdots\\ \left(\operatorname{grad} f_m(\mathbf{x}_0)\right)^T\end{pmatrix}=\left(\frac{\partial \mathbf{f}}{\partial x_1}(\mathbf{x}_0),\frac{\partial \mathbf{f}}{\partial x_2}(\mathbf{x}_0),\dots,\frac{\partial \mathbf{f}}{\partial x_n}(\mathbf{x}_0)\right)$$

Beispiel:

1) $\mathbf{f}:\mathbb{R}^2\to\mathbb{R}^3$ mit $\mathbf{f}(x,y)=\begin{pmatrix}f_1(x,y)\\ f_2(x,y)\\ f_3(x,y)\end{pmatrix}=\begin{pmatrix}2x+y\\ 3x^2+y^2\\ xy\end{pmatrix}$,

$$\frac{\partial(f_1,f_2,f_3)}{\partial(x,y)}(x,y)=\begin{pmatrix}2 & 1\\ 6x & 2y\\ y & x\end{pmatrix}, \quad \frac{\partial(f_1,f_2,f_3)}{\partial(x,y)}(1,2)=\begin{pmatrix}2 & 1\\ 6 & 4\\ 2 & 1\end{pmatrix}$$

2) $\mathbf{f}:\mathbb{R}^2\to\mathbb{R}^2$ mit $\mathbf{f}(r,\varphi)=\begin{pmatrix}f_1(r,\varphi)\\ f_2(r,\varphi)\end{pmatrix}=\begin{pmatrix}r\cos\varphi\\ r\sin\varphi\end{pmatrix}$,

$$\frac{\partial(f_1,f_2)}{\partial(r,\varphi)}(r,\varphi)=\begin{pmatrix}\cos\varphi & -r\sin\varphi\\ \sin\varphi & r\cos\varphi\end{pmatrix}$$

$$\det\left(\frac{\partial(f_1,f_2)}{\partial(r,\varphi)}(r,\varphi)\right)=\det\begin{pmatrix}\cos\varphi & -r\sin\varphi\\ \sin\varphi & r\cos\varphi\end{pmatrix}$$

$$=r\cos^2\varphi+r\sin^2\varphi=r\,(\cos^2\varphi+\sin^2\varphi)=r$$

$$\operatorname{div}\mathbf{f}(r,\varphi)=\operatorname{Sp}\left(\frac{\partial(f_1,f_2)}{\partial(r,\varphi)}(r,\varphi)\right)=\operatorname{Sp}\begin{pmatrix}\cos\varphi & -r\sin\varphi\\ \sin\varphi & r\cos\varphi\end{pmatrix}$$

$$=\cos\varphi+r\cos\varphi=(1+r)\cos\varphi$$

7.9 Kettenregel

Analog zum eindimensionalen Fall gilt auch für Abbildungen von $\mathbb{R}^n \to \mathbb{R}^m$ eine Kettenregel.

<u>Satz 7-15:</u>

Es sei $\mathbf{f} : M_1 \subset \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{g} : M_2 \subset \mathbb{R}^m \to \mathbb{R}^l$ und $\mathbf{h} = \mathbf{g} \circ \mathbf{f} : M_1 \to \mathbb{R}^l$ mit $\mathbf{h} = \mathbf{g}(\mathbf{f}(\mathbf{x}))$. Ist $\mathbf{x}_0$ innerer Punkt von M_1, $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ innerer Punkt von M_2, $\mathbf{f}$ differenzierbar in $\mathbf{x}_0$ und $\mathbf{g}$ differenzierbar in $\mathbf{y}_0$, dann gilt $\mathbf{h} = \mathbf{g} \circ \mathbf{f}$ ist differenzierbar in $\mathbf{x}_0$ mit

$$\underbrace{\frac{\partial(h_1, h_2, \ldots, h_l)}{\partial(x_1, x_2, \ldots, x_n)}(\mathbf{x}_0)}_{\mathbf{J}_h(\mathbf{x}_0)} = \underbrace{\frac{\partial(g_1, g_2, \ldots, g_l)}{\partial(y_1, y_2, \ldots, y_m)}\big(\mathbf{f}(\mathbf{x}_0)\big)}_{\mathbf{J}_g(\mathbf{f}(\mathbf{x}_0))} \cdot \underbrace{\frac{\partial(f_1, f_2, \ldots, f_m)}{\partial(x_1, x_2, \ldots, x_n)}(\mathbf{x}_0)}_{\mathbf{J}_f(\mathbf{x}_0)}.$$

(Produkt der Funktionalmatrizen)

<u>Anmerkung:</u>

Ist $n = l = 1$, $\mathbf{f} : \mathbb{R} \to \mathbb{R}^m$, $g : \mathbb{R}^m \to \mathbb{R}$ und $h = g \circ \mathbf{f} : \mathbb{R} \to \mathbb{R}$ mit

$$h(x) = g\big(\mathbf{f}(x)\big) = g\big(f_1(x), f_2(x), \ldots, f_m(x)\big)$$

und sind $\mathbf{f}$ und g differenzierbar, dann gilt

$$h'(x) = \frac{dh}{dx}(x) = \mathbf{J}_g\big(\mathbf{f}(x)\big) \cdot \mathbf{J}_f(x) = \big(\operatorname{grad} g\big(\mathbf{f}(x)\big)\big)^T \frac{d\mathbf{f}}{dx}(x)$$

$$= \left(\frac{\partial g}{\partial y_1}\big(\mathbf{f}(x)\big), \ldots, \frac{\partial g}{\partial y_m}\big(\mathbf{f}(x)\big)\right) \begin{pmatrix} \dfrac{df_1}{dx}(x) \\ \vdots \\ \dfrac{df_m}{dx}(x) \end{pmatrix}$$

$$= \frac{\partial g}{\partial y_1}\big(\mathbf{f}(x)\big)\frac{df_1}{dx}(x) + \cdots + \frac{\partial g}{\partial y_m}\big(\mathbf{f}(x)\big)\frac{df_m}{dx}(x) = \sum_{i=1}^{m} \frac{\partial g}{\partial y_i}\big(\mathbf{f}(x)\big)\frac{df_i}{dx}(x)$$

Beispiel:

1) $n = l = 1,\ m = 2,$

$$g:\mathbb{R}^2 \to \mathbb{R},\ g \in C^2(\mathbb{R}^2),\ \mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} : \mathbb{R} \to \mathbb{R}^2,\ f_1, f_2 \in C^2(\mathbb{R}),$$

$$h = g \circ \mathbf{f},\ h(t) = g\big(f_1(t), f_2(t)\big)$$

$$\Rightarrow h'(t) = g_x f_1' + g_y f_2'$$

$$\Rightarrow h''(t) = (g_{xx} f_1' + g_{xy} f_2') f_1' + g_x f_1'' + (g_{yx} f_1' + g_{yy} f_2') f_2' + g_y f_2''$$

$$= g_{xx} (f_1')^2 + 2 g_{xy} f_1' f_2' + g_{yy} (f_2')^2 + g_x f_1'' + g_y f_2''.$$

2) $\mathbf{g}(x,y) = \begin{pmatrix} xy \\ x-y \end{pmatrix},\ \mathbf{f}(r,\varphi) = \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix},\ \mathbf{h} = \mathbf{g} \circ \mathbf{f}$

$$\mathbf{h}(r,\varphi) = \mathbf{g}(f_1(r,\varphi), f_2(r,\varphi)) = \mathbf{g}\big(r\cos\varphi, r\sin\varphi\big)$$

$$= \begin{pmatrix} r\cos\varphi \cdot r\sin\varphi \\ r\cos\varphi - r\sin\varphi \end{pmatrix} = \begin{pmatrix} r^2 \cos\varphi \sin\varphi \\ r(\cos\varphi - \sin\varphi) \end{pmatrix} = \begin{pmatrix} h_1(r,\varphi) \\ h_2(r,\varphi) \end{pmatrix}$$

$$\Rightarrow \partial h_1 / \partial r = 2r\cos\varphi\sin\varphi,\quad \partial h_1/\partial\varphi = r^2(-\sin^2\varphi + \cos^2\varphi),$$

$$\partial h_2/\partial r = \cos\varphi - \sin\varphi,\quad \partial h_2/\partial\varphi = -r(\sin\varphi + \cos\varphi)$$

oder mit Hilfe der Kettenregel

$$\frac{\partial(h_1,h_2)}{\partial r\,\partial\varphi} = \begin{pmatrix} \partial h_1/\partial r & \partial h_1/\partial\varphi \\ \partial h_2/\partial r & \partial h_2/\partial\varphi \end{pmatrix} = \frac{\partial(g_1,g_2)}{\partial x\,\partial y} \cdot \frac{\partial(f_1,f_2)}{\partial r\,\partial\varphi}$$

$$= \begin{pmatrix} \partial g_1/\partial x & \partial g_1/\partial y \\ \partial g_2/\partial x & \partial g_2/\partial y \end{pmatrix} \cdot \begin{pmatrix} \partial f_1/\partial r & \partial f_1/\partial\varphi \\ \partial f_2/\partial r & \partial f_2/\partial\varphi \end{pmatrix}$$

$$= \begin{pmatrix} y & x \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

$$= \begin{pmatrix} r\sin\varphi & r\cos\varphi \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

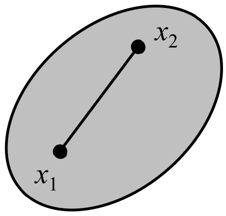
$$= \begin{pmatrix} 2r\sin\varphi\cos\varphi & -r^2(\sin^2\varphi - \cos^2\varphi) \\ \cos\varphi - \sin\varphi & -r(\sin\varphi + \cos\varphi) \end{pmatrix}.$$

7.10 Approximation höherer Ordnung, Satz von Taylor

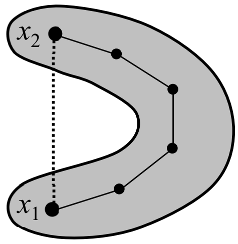
Definition 7-15:

1) $M \subset \mathbb{R}^n$ heißt konvex, wenn für zwei beliebige Punkte $\mathbf{x}_1, \mathbf{x}_2 \in M$ auch deren Verbindungslinie ganz in M liegt.

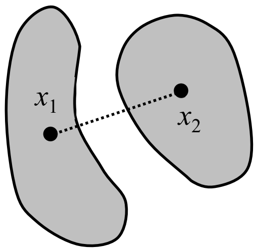
2) $G \subset \mathbb{R}^n$ heißt Gebiet, wenn G offen und zusammenhängend ist, d.h. zwei beliebige Punkte $\mathbf{x}_1, \mathbf{x}_2 \in G$ lassen sich durch einen Polygonzug verbinden, der ganz in G verläuft.



konvex,
Gebiet



nicht konvex,
zusammenhängend,
Gebiet



nicht konvex,
nicht zusammenhängend,
kein Gebiet

Der Satz von Taylor lässt sich mit Hilfe des Differentialoperators

$$\partial_{\mathbf{h}}^l = \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + \ldots + h_n \frac{\partial}{\partial x_n} \right)^l$$

übersichtlicher darstellen, wobei $\mathbf{h} \in \mathbb{R}^n$.

Beispiel:

$$l = 1: \quad \partial_{\mathbf{h}} = h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + \ldots + h_n \frac{\partial}{\partial x_n} = \sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i}$$

$$l = 2: \quad \partial_{\mathbf{h}}^2 = \left(\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i} \right)^2 = \left(\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i} \right)\left(\sum_{j=1}^{n} h_j \frac{\partial}{\partial x_j} \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \frac{\partial^2}{\partial x_i \partial x_j}$$

$$l \in \mathbb{N}: \quad \partial_{\mathbf{h}}^l = \left(\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i} \right)^l = \sum_{i_1=1}^{n} \cdots \sum_{i_l=1}^{n} h_{i_1} \cdots h_{i_l} \frac{\partial^l}{\partial x_{i_1} \cdots \partial x_{i_l}}$$

Spezialfall : $n=2,\ f: M \subset \mathbb{R}^2 \to \mathbb{R},\ f \in C^3(M)$

$$\partial_{\mathbf{h}} f = \left(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y} \right) f = h_1 \frac{\partial f}{\partial x} + h_2 \frac{\partial f}{\partial y} = h_1 f_x + h_2 f_y$$

$$\partial_{\mathbf{h}}^2 f = \left(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y} \right)^2 f = \left(h_1^2 \frac{\partial^2}{\partial x^2} + 2h_1 h_2 \frac{\partial^2}{\partial x \partial y} + h_2^2 \frac{\partial^2}{\partial y^2} \right) f$$

$$= h_1^2 f_{xx} + 2h_1 h_2 f_{xy} + h_2^2 f_{yy}$$

$$\partial_{\mathbf{h}}^3 f = \left(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y} \right)^3 f$$

$$= \left(h_1^3 \frac{\partial^3}{\partial x^3} + 3h_1^2 h_2 \frac{\partial^3}{\partial x^2 \partial y} + 3h_1 h_2^2 \frac{\partial^3}{\partial x \partial y^2} + h_2^3 \frac{\partial^3}{\partial y^3} \right) f$$

$$= h_1^3 f_{xxx} + 3h_1^2 h_2 f_{xxy} + 3h_1 h_2^2 f_{xyy} + h_2^3 f_{yyy}$$

Satz 7-16: (*Satz von Taylor*)
Ist $G \subset \mathbb{R}^n$ ein konvexes Gebiet, $f \in C^{m+1}(G)$ und $\mathbf{x}, \mathbf{x}_0 \in G$. Dann lässt sich f durch

$$f(\mathbf{x}) = T_{m,\mathbf{x}_0}(\mathbf{x}) + R_{m,\mathbf{x}_0}(\mathbf{x})$$

darstellen, wobei

$$T_{m,\mathbf{x}_0}(\mathbf{x}) = \sum_{k=0}^{m} \frac{1}{k!} \partial_{(\mathbf{x}-\mathbf{x}_0)}^k f(\mathbf{x}_0)$$

das Taylor-Polynom vom grad m und

$$R_{m,\mathbf{x}_0}(\mathbf{x}) = \frac{1}{(m+1)!} \partial_{(\mathbf{x}-\mathbf{x}_0)}^{m+1} f(\boldsymbol{\xi})$$

das Restglied mit

$$\boldsymbol{\xi} = \mathbf{x}_0 + \delta(\mathbf{x} - \mathbf{x}_0),\ 0 < \delta < 1$$

bezeichnet.

Es sei $f \in C^k(G)$ mit $k \geq 2$, dann heißt die symmetrische Matrix

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) & \cdots & f_{x_1x_n}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & \cdots & f_{x_2x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{x}) & f_{x_nx_2}(\mathbf{x}) & \cdots & f_{x_nx_n}(\mathbf{x}) \end{pmatrix}$$

Hesse-Matrix von f im Punkt $\mathbf{x}$.

Wegen

$$\partial_{\mathbf{h}} f = h_1 f_x + h_2 f_y = (h_1, h_2)\begin{pmatrix} f_x \\ f_y \end{pmatrix} = \mathbf{h}^T \operatorname{grad} f$$

und

$$\partial_{\mathbf{h}}^2 f = h_1^2 f_{xx} + 2h_1h_2 f_{xy} + h_2^2 f_{yy} = \left(h_1, h_2\right)\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \mathbf{h}^T \mathbf{H}_f \mathbf{h}$$

lauten die Taylor-Polynome für die vielbenutzten Spezialfälle $m = 1,2$

$$T_{1,\mathbf{x}_0}(\mathbf{x}) = f(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \operatorname{grad} f(\mathbf{x}_0)$$

und

$$T_{2,\mathbf{x}_0}(\mathbf{x}) = f(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \operatorname{grad} f(\mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0).$$

<u>Anmerkung:</u>

a) $f(\mathbf{x}) = f(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \operatorname{grad} f(\boldsymbol{\xi})$ (Mittelwertsatz)
mit $\boldsymbol{\xi} = \mathbf{x}_0 + \delta(\mathbf{x} - \mathbf{x}_0)$, $0 < \delta < 1$.

b) $$f(\mathbf{x}) = f(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \operatorname{grad} f(\mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + o\left(\|\mathbf{x} - \mathbf{x}_0\|_2^2\right).$$

7.11 Implizite Funktionen

Oft sind Funktionen nur in impliziter Form gegeben, z.B.

$$F(x,y) = x^2 + y^2 - 1 = 0.$$

Der gesamte Kreis ist nicht der Graph einer Funktion. Er ist aber als Vereinigung der Graphen der (impliziten) Funktionen

$$f_1(x) = \sqrt{1-x^2} \quad \text{und} \quad f_2(x) = -\sqrt{1-x^2}$$

darstellbar.

Satz 7-17: (*implizite Funktionen*)
Es sei $F: M \subset \mathbb{R}^2 \to \mathbb{R}$, U eine Umgebung von (x_0, y_0) mit $U \subset M$ und $F \in C^1(U)$ mit $F(x_0,y_0) = 0$ und $F_y(x_0,y_0) \neq 0$. Dann gibt es eine Umgebung

$$U_\delta(x_0) = \left\{ x : \left| x - x_0 \right| < \delta \right\}$$

mit $\delta > 0$ und genau eine Funktion $f: U_\delta \subset \mathbb{R} \to \mathbb{R}$ mit

$$F\big(x, f(x)\big) = 0 \ \ \forall x \in U_\delta, \ \ f \in C^1(U_\delta), \ \ f(x_0) = y_0$$

und

$$f'(x) = -\frac{F_x(x, f(x))}{F_y(x, f(x))}.$$

Merkregel:

$F\big(x, f(x)\big) = 0 \ \ \forall x \in U_\delta,$

differenzieren nach x liefert mit der Kettenregel

$$F_x\big(x, f(x)\big) \cdot 1 + F_y\big(x, f(x)\big) \cdot f'(x) = 0 \ \Rightarrow \ f'(x) = -\frac{F_x\big(x, f(x)\big)}{F_y\big(x, f(x)\big)}.$$

Beispiel:

$$F(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \ \text{ und } \ a, b > 0,$$

$$F_y(x,y)=\frac{2y}{b^2}\neq 0 \ \text{ falls } \ y\neq 0$$

$\Rightarrow$ lokal auflösbar nach y falls $y\neq 0$, d.h. es existiert $y=f(x)$ mit $F\big(x,f(x)\big)=0 \ \ \forall x\in U_\delta$.

Höhere Ableitungen

Ist $F\in C^l(U)$, dann gilt auch $f\in C^l(U_\delta)$. Man kann auch die höheren Ableitungen der Funktion implizit bestimmen.

Beispiel:

Es sei $F\big(x,f(x)\big)=0$ für alle $x\in U_\delta$ und $F_y\big(x,f(x)\big)\neq 0$.

$$\Rightarrow F_x\big(x,f(x)\big)+F_y\big(x,f(x)\big)\cdot f'(x)=0$$

$$\Rightarrow F_{xx}\big(x,f(x)\big)\cdot 1+F_{xy}\big(x,f(x)\big)\cdot f'(x)+$$
$$\Big(F_{yx}\big(x,f(x)\big)\cdot 1+F_{yy}\big(x,f(x)\big)\cdot f'(x)\Big)\cdot f'(x)+F_y\big(x,f(x)\big)\cdot f''(x)=0$$

$$\Rightarrow f''(x)=-\frac{1}{F_y\big(x,f(x)\big)}\Big(F_{xx}\big(x,f(x)\big)+2F_{xy}\big(x,f(x)\big)\cdot f'(x)$$
$$+F_{yy}\big(x,f(x)\big)\cdot (f'(x))^2\Big).$$

Anmerkung:

Analoge Aussagen sind möglich, falls bei der impliziten Gleichung

$$F(x_1,\dots,x_n,y_1,\dots,y_m)=0$$

für $F:M\subset\mathbb{R}^{n+m}\to\mathbb{R}$ eine Auflösung

$$y_1=f_1(x_1,\dots,x_n),\dots,y_m=f_m(x_1,\dots,x_n)$$

mit den Funktionen

$$f_1:U_\delta\subset\mathbb{R}^n\to\mathbb{R},\dots,f_m:U_\delta\subset\mathbb{R}^n\to\mathbb{R}$$

gesucht wird, so dass

$$F\big(x_1,\dots,x_n,f_1(x_1,\dots,x_n),\dots,f_m(x_1,\dots,x_n)\big)=0.$$

7.12 Extrema für Funktionen mehrerer Variabler

Definition 7-16: (*relative Extrema*)
Es sei $f: M \subset \mathbb{R}^n \to \mathbb{R}$. f besitzt in $\mathbf{x}_0 \in M$ ein relatives (lokales) Maximum bzw. Minimum, wenn es eine Umgebung U von $\mathbf{x}_0$ gibt mit

$$f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad \text{bzw.} \quad f(\mathbf{x}) \geq f(\mathbf{x}_0).$$

Gelten diese Ungleichungen jeweils $\forall \mathbf{x} \in M$, so besitzt f in $\mathbf{x}_0$ ein absolutes (globales) Maximum bzw. Minimum.

Satz 7-18: (*notwendige Bedingung für relative Extrema*)
Es sei $f: M \subset \mathbb{R}^n \to \mathbb{R}$, $\mathbf{x}_0 \in M$ und $U \subset M$ eine Umgebung von $\mathbf{x}_0$. Ist $f \in C^1(U)$ und hat f in $\mathbf{x}_0$ ein relatives Extremum, dann gilt

$$\operatorname{grad} f(\mathbf{x}_0) = \left(f_{x_1}(\mathbf{x}_0), f_{x_2}(\mathbf{x}_0), \ldots, f_{x_n}(\mathbf{x}_0) \right)^T = \mathbf{0}.$$

Anmerkung:
Man nennt $\mathbf{x}_0 \in \mathbb{R}^n$ einen stationären Punkt von f, wenn $\operatorname{grad} f(\mathbf{x}_0) = \mathbf{0}$ gilt. Einen stationären Punkt, der keine Extremalstelle ist, bezeichnet man als Sattelpunkt.

Satz 7-19: (*hinreichende Bedingung für relative Extrema*)
Es sei $f: M \subset \mathbb{R}^n \to \mathbb{R}$, $\mathbf{x}_0 \in M$, $U \subset M$ Umgebung von $\mathbf{x}_0$, $f \in C^2(U)$ und $\operatorname{grad} f(\mathbf{x}_0) = \mathbf{0}$ sowie

$$\mathbf{H}_f(\mathbf{x}_0) = \left(f_{x_i x_j}(\mathbf{x}_0) \right)_{i=1,\ldots,n;\, j=1,\ldots,n}$$

die Hesse-Matrix von f in $\mathbf{x}_0$, dann gilt

a) $\mathbf{H}_f(\mathbf{x}_0)$ positiv definit $\Rightarrow$ f hat in $\mathbf{x}_0$ relatives Minimum,
b) $\mathbf{H}_f(\mathbf{x}_0)$ negativ definit $\Rightarrow$ f hat in $\mathbf{x}_0$ relatives Maximum,
c) $\mathbf{H}_f(\mathbf{x}_0)$ indefinit $\Rightarrow$ f hat in $\mathbf{x}_0$ Sattelpunkt,
d) $\mathbf{H}_f(\mathbf{x}_0)$ semidefinit $\Rightarrow$ keine Aussage möglich.

Anmerkung:

a) Quadratische Form:

Eine Funktion $Q : \mathbb{R}^n \to \mathbb{R}$, der Form

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A}\mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \quad \text{mit} \quad \mathbf{A} = \mathbf{A}^T \in \mathbb{R}^{n \times n}$$

heißt quadratische Form und $\mathbf{A}$ heißt

$$\begin{array}{lll}
\text{positiv definit} & \Leftrightarrow & Q(\mathbf{x}) > 0 \text{ für alle } \mathbf{x} \in \mathbb{R}^n \backslash \{\mathbf{0}\}, \\
\text{negativ definit} & \Leftrightarrow & Q(\mathbf{x}) < 0 \text{ für alle } \mathbf{x} \in \mathbb{R}^n \backslash \{\mathbf{0}\}, \\
\text{indefinit} & \Leftrightarrow & Q(\mathbf{x}) \text{ nimmt positive} \wedge \text{negative Werte an}, \\
\text{positiv semidefinit} & \Leftrightarrow & Q(\mathbf{x}) \geq 0 \text{ für alle } \mathbf{x} \in \mathbb{R}^n \backslash \{\mathbf{0}\}, \\
\text{negativ semidefinit} & \Leftrightarrow & Q(\mathbf{x}) \leq 0 \text{ für alle } \mathbf{x} \in \mathbb{R}^n \backslash \{\mathbf{0}\}.
\end{array}$$

b) Hauptminoren von $\mathbf{A}$ ($n \times n$ Matrix):

$$D_i = \det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1i} \\ a_{21} & a_{22} & \dots & a_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ii} \end{pmatrix}$$

bezeichnet den führenden Hauptminor der Ordnung i und

$$\Delta_i = \det \begin{pmatrix} a_{l_1,l_1} & a_{l_1,l_2} & \dots & a_{l_1,l_i} \\ a_{l_2,l_1} & a_{l_2,l_2} & \dots & a_{l_2,l_i} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l_i,l_1} & a_{l_i,l_2} & \dots & a_{l_i,l_i} \end{pmatrix}, \quad 1 \leq l_1 < l_2 < \cdots < l_i \leq n$$

einen der (n über i) Hauptminoren der Ordnung i.

c) Eigenwert-/Eigenvektorzerlegung von $\mathbf{A}$ ($n \times n$ Matrix)

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{v}_i^T, \quad \mathbf{v}_i^T \mathbf{v}_j = 0 \text{ für } i \neq j$$

gibt die Eigenwert-/Eigenvektorzerlegung von $\mathbf{A}$ an, wobei λ_i die Eigenwerte und $\mathbf{v}_i$ die Eigenvektoren von $\mathbf{A}$ bezeichnet.

Es gilt

$\mathbf{A}$ ist positiv definit $\Leftrightarrow D_i > 0$ für $i = 1,2,...,n$,
$\Leftrightarrow \lambda_i > 0$ für $i = 1,2,...,n$,

$\mathbf{A}$ ist negativ definit $\Leftrightarrow (-1)^i D_i > 0$ für $i = 1,2,...,n$,
$\Leftrightarrow \lambda_i < 0$ für $i = 1,2,...,n$,
$\Leftrightarrow -\mathbf{A}$ ist positiv definit,

$\mathbf{A}$ ist positiv semidefinit $\Leftrightarrow \Delta_n = 0$ und alle $\Delta_i \geq 0$ für $i = 1,2,...,n-1$,
$\Leftrightarrow \lambda_i \geq 0$ für $i = 1,2,...,n$, wobei mindestens ein $\lambda_i = 0$ ist,

$\mathbf{A}$ ist negativ semidefinit $\Leftrightarrow \Delta_n = 0$ und alle $(-1)^i \Delta_i \geq 0$ für $i = 1,2,...,n-1$,
$\Leftrightarrow \lambda_i \leq 0$ für $i = 1,2,...,n$, wobei mindestens ein $\lambda_i = 0$ ist,
$\Leftrightarrow -\mathbf{A}$ ist positiv semidefinit,

$\mathbf{A}$ ist indefinit $\Leftrightarrow D_i$ haben andere Vorzeichen als oben,
$\Leftrightarrow$ es existieren $\lambda_i > 0$ und $\lambda_i < 0$.

Spezialfall $n = 2$:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \text{ d.h. } D_1 = a_{11} \text{ und } D_2 = \det \mathbf{A} = a_{11}a_{22} - a_{12}^2, \text{ da } a_{12} = a_{21}.$$

$\mathbf{A}$ ist positiv definit $\Leftrightarrow a_{11} > 0$ und $\det \mathbf{A} > 0$,
$\mathbf{A}$ ist negativ definit $\Leftrightarrow a_{11} < 0$ und $\det \mathbf{A} > 0$,
$\mathbf{A}$ ist positiv semidefinit $\Leftrightarrow a_{11} \geq 0$ und $\det \mathbf{A} = 0$,
$\mathbf{A}$ ist negativ semidefinit $\Leftrightarrow a_{11} \leq 0$ und $\det \mathbf{A} = 0$,
$\mathbf{A}$ ist indefinit $\Leftrightarrow \det \mathbf{A} < 0$.

Beispiel:

1) $f(x,y)=xy \;\Rightarrow\; f\in C^\infty(\mathbb{R}^2)$

<u>notwendige Bedingung</u>

$$\operatorname{grad} f(x,y)=\begin{pmatrix} f_x(x,y)\\ f_y(x,y)\end{pmatrix}=\begin{pmatrix} y\\ x\end{pmatrix}=\begin{pmatrix} 0\\ 0\end{pmatrix}$$

liefert mögliches Extremum in $(x,y)^T=(0,0)^T$.

<u>hinreichende Bedingung</u>

$$\mathbf{H}_f(x,y)=\begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y)\\ f_{yx}(x,y) & f_{yy}(x,y)\end{pmatrix}=\begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix}$$

$$\mathbf{H}_f(0,0)=\begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix},\;\det\mathbf{H}_f(0,0)=-1<0$$

$\Rightarrow\; \mathbf{H}_f(0,0)$ ist indefinit $\Rightarrow$ Sattelpunkt $(0,0)^T$.

2) $f(x,y)=(x+y)^3-12xy \;\Rightarrow\; f\in C^\infty(\mathbb{R}^2)$

<u>notwendige Bedingung</u>

$$\operatorname{grad} f(x,y)=\begin{pmatrix} f_x\\ f_y\end{pmatrix}=\begin{pmatrix} 3(x+y)^2-12y\\ 3(x+y)^2-12x\end{pmatrix}=\begin{pmatrix} 0\\ 0\end{pmatrix}\;\Rightarrow\; x=y$$

Einsetzen von $x=y$ in die zweite Gleichung liefert

$$x^2=x\;\Rightarrow\; x_1=1,\; x_2=0$$

und damit mögliche Extrema in $(x,y)^T=(0,0)^T$ und $(x,y)^T=(1,1)^T$.

<u>hinreichende Bedingung</u>

$$\mathbf{H}_f(x,y)=\begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y)\\ f_{yx}(x,y) & f_{yy}(x,y)\end{pmatrix}$$

$$=\begin{pmatrix} 6(x+y) & 6(x+y)-12\\ 6(x+y)-12 & 6(x+y)\end{pmatrix}$$

$$\mathbf{H}_f(0,0) = \begin{pmatrix} 0 & -12 \\ -12 & 0 \end{pmatrix} \Rightarrow \det \mathbf{H}_f(0,0) = -144 < 0$$

$\Rightarrow$ $\mathbf{H}_f(0,0)$ indefinit

$\Rightarrow$ Sattelpunkt in $(0,0)^T$,

$$\mathbf{H}_f(1,1) = \begin{pmatrix} 12 & 0 \\ 0 & 12 \end{pmatrix} \Rightarrow \det \mathbf{H}_f(1,1) = 144 > 0,\ h_{f,11}(1,1) = 12 > 0$$

$\Rightarrow$ $\mathbf{H}_f(1,1)$ positiv definit

$\Rightarrow$ relatives Minimum in $(1,1)^T$ mit $f(1,1) = -4$.

Aufgabe 7-1: Berechnen Sie die Richtungsableitungen der Funktionen

a) $f(x,y)=\sin(x^2y^3)$,

b) $f(x,y)=x\arctan y$

in Richtung $\mathbf{a}=(1/\sqrt{2},1/\sqrt{2})^T$ und geben Sie die Richtungen an in denen die Richtungsableitungen ihr Maximum annehmen.

Aufgabe 7-2: Berechnen Sie den Gradienten der Funktion

$$f(x,y,z)=e^{x+2y}+2x\sin z+z^2xy\,.$$

Aufgabe 7-3: Berechnen Sie alle partiellen Ableitungen 2. Ordnung der Funktionen

a) $f(x,y)=\sin(ax+by)$,

b) $f(x,y)=\ln(x^2+y)$,

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

Aufgabe 7-4: Berechnen Sie für a) die Ableitung $h'(t)$ sowie für b) die Ableitungen $h_r(r,\varphi)$ und $h_\varphi(r,\varphi)$ mit Hilfe der Kettenregel.

a) $h(t)=g\big(f_1(t),f_2(t)\big)$ mit $g(x_1,x_2)=3x_1^2+2x_1x_2+x_2^2$ und
$x_1=f_1(t)=\sin t,\quad x_2=f_2(t)=\cos t$

b) $h(r,\varphi)=g\big(f_1(r,\varphi),f_2(r,\varphi)\big)$ mit $g(x_1,x_2)=x_1^3-x_1x_2+x_2^3$ und
$x_1=f_1(r,\varphi)=r\cos\varphi,\quad x_2=f_2(r,\varphi)=r\sin\varphi$

Aufgabe 7-5: Das Potential zweier positiver Punktlandungen sei gegeben durch

$$\varphi(x,y,z)=\tilde{\varphi}\big(r_1(x,y,z),r_2(x,y,z)\big)\quad\text{mit}$$

$$\tilde{\varphi}(r_1,r_2)=\frac{1}{4\pi\varepsilon_0}\left(\frac{Q_1}{r_1}+\frac{Q_2}{r_2}\right)\quad\text{und}\quad r_1=\sqrt{(x-1)^2+y^2+z^2},\quad r_2=\sqrt{x^2+(y-1)^2+z^2}.$$

Berechnen Sie die elektrische Feldstärke im Koordinatenursprung.

Aufgabe 7-6: Approximieren Sie die folgenden Funktionen um (x_0,y_0) durch ein Taylorpolynom 2. Ordnung.

a) $f(x,y)=y\ln(y-3x),\quad (x_0,y_0)=(0,1)$

b) $f(x,y)=\cos x\cos y,\quad (x_0,y_0)=(0,0)$

Aufgabe 7-7: Überprüfen Sie, ob für die Gleichung

$$F(x,y) = x\,e^{y} - y\,e^{x} + x = 0$$

in $U(0{,}0)$ die implizite Funktion $y = f(x)$ existiert und berechnen Sie gegebenenfalls die Werte $f'(0)$ und $f''(0)$.

Aufgabe 7-8: Bestimmen Sie die Lage und Art der relativen Extrema der Funktionen

a) $f(x,y) = xy(3 - x - y)$,

b) $f(x,y) = x^2(2 - y) - y^3 + 3y^2 + 9y$

und geben Sie die zugehörigen Funktionswerte an.