

# Analysis III für Studierende der Ingenieurwissenschaften

Ingenuin Gasser

Fachbereich Mathematik  
Universität Hamburg

Technische Universität Hamburg–Harburg  
Wintersemester 2021/22

basierend auf dem Foliensatz von Jens Struckmeier Wintersemester 2020/21

# Inhalte der Vorlesung Analysis III.

- 1 Partielle Ableitungen, Differentialoperatoren.
- 2 Vektorfelder, vollständiges Differential, Richtungsableitungen.
- 3 Mittelwertsätze, Satz von Taylor.
- 4 Extrema, Satz über implizite Funktionen.
- 5 Implizite Darstellung von Kurven und Flächen.
- 6 Extrema bei Gleichungsnebenbedingungen.
- 7 Newton–Verfahren, nichtlineare Gleichungen und Ausgleichsrechnung.
- 8 Bereichsintegrale, Satz von Fubini, Transformationssatz.
- 9 Potentiale, Integralsatz von Green, Integralsatz von Gauß.
- 10 Greensche Formeln, Integralsatz von Stokes.

# Kapitel 1. Differentialrechnung mehrerer Variablen

## 1.1 Partielle Ableitungen

Im Folgenden sei

$f(x_1, \dots, x_n)$  eine skalare Funktion, die von  $n$  Variablen abhängt

**Beispiel:** Die Zustandsgleichung eines idealen Gases lautet  $pV = RT$ .

Jede der drei Größen  $p$  (Druck),  $V$  (Volumen) und  $T$  (Temperatur) läßt sich als Funktion der anderen darstellen, wobei  $R$  die universelle Gaskonstante ist.

$$p = p(V, T) = \frac{RT}{V}$$

$$V = V(p, T) = \frac{RT}{p}$$

$$T = T(p, V) = \frac{pV}{R}$$

# 1.1. Partielle Ableitungen

**Definition:** Sei  $D \subset \mathbb{R}^n$  offen,  $f : D \rightarrow \mathbb{R}$ ,  $x^0 \in D$ .

- $f(x)$  heißt in  $x^0$  nach  $x_i$  **partiell differenzierbar**, falls der Grenzwert

$$\begin{aligned}\frac{\partial f}{\partial x_i}(x^0) &:= \lim_{t \rightarrow 0} \frac{f(x^0 + te_i) - f(x^0)}{t} \\ &= \lim_{t \rightarrow 0} \frac{f(x_1^0, \dots, x_i^0 + t, \dots, x_n^0) - f(x_1^0, \dots, x_i^0, \dots, x_n^0)}{t}\end{aligned}$$

existiert, wobei  $e_i$  den  $i$ -ten Einheitsvektor bezeichnet. Den Grenzwert nennt man die **partielle Ableitung** von  $f(x)$  nach  $x_i$  im Punkt  $x^0$ .

- Existieren für jeden Punkt  $x^0$  die partiellen Ableitungen nach jeder Variablen  $x_i$ ,  $i = 1, \dots, n$  und sind diese **stetige Funktionen**, so nennt man  $f(x)$  **stetig partiell differenzierbar** oder eine  **$C^1$ -Funktion**.

# Beispiele.

- Betrachte die Funktion

$$f(x_1, x_2) = x_1^2 + x_2^2$$

Für einen Punkt  $x^0 \in \mathbb{R}^2$  existieren beide partiellen Ableitungen und diese sind auch stetig:

$$\frac{\partial f}{\partial x_1}(x^0) = 2x_1, \quad \frac{\partial f}{\partial x_2}(x^0) = 2x_2$$

Die Funktion ist also eine  $\mathcal{C}^1$ -Funktion.

- Die Funktion

$$f(x_1, x_2) = x_1 + |x_2|$$

ist im Punkt  $x^0 = (0, 0)^T$  partiell differenzierbar nach der Koordinate  $x_1$ , aber die partielle Ableitung nach  $x_2$  existiert im Ursprung **nicht!**

# Konkretes technisches Beispiel.

Der Schalldruck einer eindimensionalen Schallwelle ist gegeben durch

$$p(x, t) = A \sin(\alpha x - \omega t)$$

Die partielle Ableitung

$$\frac{\partial p}{\partial x} = \alpha A \cos(\alpha x - \omega t)$$

beschreibt zu einer festen Zeit  $t$  die **örtliche** Änderungsrate des Schalldrucks.

Die partielle Ableitung

$$\frac{\partial p}{\partial t} = -\omega A \cos(\alpha x - \omega t)$$

beschreibt für einen festen Ort  $x$  die **zeitliche** Änderung des Schalldruckes.

# Differentiationsregeln

- Sind  $f, g$  partiell nach  $x_i$  differenzierbar,  $\alpha, \beta \in \mathbb{R}$ , so gelten die Regeln

$$\frac{\partial}{\partial x_i} (\alpha f(x) + \beta g(x)) = \alpha \frac{\partial f}{\partial x_i}(x) + \beta \frac{\partial g}{\partial x_i}(x)$$

$$\frac{\partial}{\partial x_i} (f(x) \cdot g(x)) = \frac{\partial f}{\partial x_i}(x) \cdot g(x) + f(x) \cdot \frac{\partial g}{\partial x_i}(x)$$

$$\frac{\partial}{\partial x_i} \left( \frac{f(x)}{g(x)} \right) = \frac{\frac{\partial f}{\partial x_i}(x) \cdot g(x) - f(x) \cdot \frac{\partial g}{\partial x_i}(x)}{g(x)^2} \quad \text{für } g(x) \neq 0$$

- Man verwendet alternativ die Bezeichnungen:

$$D_i f(x^0) \quad \text{oder} \quad f_{x_i}(x^0)$$

für die partielle Ableitung von  $f(x)$  nach  $x_i$  in  $x^0$ .

# Gradient und Nabla-Operator.

**Definition:** Sei  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$  offen, und partiell differenzierbar.

- Man bezeichnet den **Zeilenvektor**

$$\text{grad } f(x^0) := \left( \frac{\partial f}{\partial x_1}(x^0), \dots, \frac{\partial f}{\partial x_n}(x^0) \right)$$

als **Gradient** von  $f(x)$  in  $x^0$ .

- Weiterhin bezeichnet man den symbolischen Vektor

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

als **Nabla-Operator**.

- So bekommt man den **Spaltenvektor**

$$\nabla f(x^0) := \left( \frac{\partial f}{\partial x_1}(x^0), \dots, \frac{\partial f}{\partial x_n}(x^0) \right)^T$$



# Weitere Differentiationsregeln.

Seien  $f(x)$  und  $g(x)$  partiell differenzierbar. Dann gelten die folgenden **Differentiationsregeln**:

$$\text{grad}(\alpha f + \beta g) = \alpha \cdot \text{grad} f + \beta \cdot \text{grad} g$$

$$\text{grad}(f \cdot g) = g \cdot \text{grad} f + f \cdot \text{grad} g$$

$$\text{grad}\left(\frac{f}{g}\right) = \frac{1}{g^2}(g \cdot \text{grad} f - f \cdot \text{grad} g), \quad g \neq 0$$

## Beispiele:

- Sei  $f(x, y) = e^x \cdot \sin y$ . Dann gilt:

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

- Für  $r(x) := \|x\|_2 = \sqrt{x_1^2 + \dots + x_n^2}$  gilt

$$\text{grad} r(x) = \frac{x}{r(x)} = \frac{x}{\|x\|_2} \quad \text{für } x \neq 0,$$

wobei  $x = (x_1, \dots, x_n)$  Zeilenvektor.

# Partiell differenzierbar impliziert nicht Stetigkeit.

**Beobachtung:** Eine (nach allen Koordinaten) partiell differenzierbare Funktion ist nicht notwendigerweise eine **stetige** Funktion.

**Beispiel:** Betrachte die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definiert durch

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

Die Funktion ist auf **ganz**  $\mathbb{R}^2$  partiell differenzierbar, und es gilt

$$f_x(0, 0) = f_y(0, 0) = 0$$

$$\frac{\partial f}{\partial x}(x, y) = \frac{y}{(x^2 + y^2)^2} - 4 \frac{x^2 y}{(x^2 + y^2)^3}, \quad (x, y) \neq (0, 0)$$

$$\frac{\partial f}{\partial y}(x, y) = \frac{x}{(x^2 + y^2)^2} - 4 \frac{xy^2}{(x^2 + y^2)^3}, \quad (x, y) \neq (0, 0)$$

## Beispiel (Fortsetzung).

Berechnung der partiellen Ableitungen im Ursprung  $(0, 0)$ :

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{t \rightarrow 0} \frac{f(t, 0) - f(0, 0)}{t} = \frac{t \cdot 0}{(t^2 + 0^2)^2} - 0 = 0$$

$$\frac{\partial f}{\partial y}(0, 0) = \lim_{t \rightarrow 0} \frac{f(0, t) - f(0, 0)}{t} = \frac{0 \cdot t}{(0^2 + t^2)^2} - 0 = 0$$

**Aber:** Im Nullpunkt  $(0, 0)$  ist die Funktion **nicht** stetig, denn

$$\lim_{n \rightarrow \infty} f\left(\frac{1}{n}, \frac{1}{n}\right) = \frac{\frac{1}{n} \cdot \frac{1}{n}}{\left(\frac{1}{n} \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{n}\right)^2} = \frac{\frac{1}{n^2}}{\frac{4}{n^4}} = \frac{n^2}{4} \rightarrow \infty$$

und somit gilt

$$\lim_{(x,y) \rightarrow (0,0)} f(x, y) \neq f(0, 0) = 0$$

# Beschränktheit der Ableitungen impliziert Stetigkeit.

Um die Stetigkeit einer partiell differenzierbaren Funktion zu garantieren, benötigt man zusätzliche Voraussetzungen an  $f$ .

**Satz:** Ist  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$  offen, in einer Umgebung von  $x^0 \in D$  partiell differenzierbar, und sind die partiellen Ableitungen  $\frac{\partial f}{\partial x_i}$ ,  $i = 1, \dots, n$ , dort **beschränkt**, so ist  $f(x)$  **stetig** in  $x^0$ .

**Beachte:** In unserem vorigem Beispiel sind die partiellen Ableitungen in einer Umgebung der Null  $(0, 0)$  **nicht** beschränkt, denn es gilt

$$\frac{\partial f}{\partial x}(x, y) = \frac{y}{(x^2 + y^2)^2} - 4 \frac{x^2 y}{(x^2 + y^2)^3} \quad \text{für } (x, y) \neq (0, 0)$$

# Beweis des Satzes.

Für  $\|x - x^0\|_\infty < \varepsilon$ ,  $\varepsilon > 0$  hinreichend klein, schreiben wir:

$$\begin{aligned} f(x) - f(x^0) &= (f(x_1, \dots, x_{n-1}, x_n) - f(x_1, \dots, x_{n-1}, x_n^0)) \\ &+ (f(x_1, \dots, x_{n-1}, x_n^0) - f(x_1, \dots, x_{n-2}, x_{n-1}^0, x_n^0)) \\ &\vdots \\ &+ (f(x_1, x_2^0, \dots, x_n^0) - f(x_1^0, \dots, x_n^0)) \end{aligned}$$

Für jede Differenz auf der linken Seite, betrachten wir  $f$  als univariate Funktion:

$$g(x_n) - g(x_n^0) := f(x_1, \dots, x_{n-1}, x_n) - f(x_1, \dots, x_{n-1}, x_n^0)$$

Da  $f$  partiell differenzierbar, ist  $g$  differenzierbar und es gilt der Mittelwertsatz:

$$g(x_n) - g(x_n^0) = g'(\xi_n)(x_n - x_n^0)$$

für ein geeignetes  $\xi_n$  zwischen  $x_n$  und  $x_n^0$ .

# Beweis des Satzes (Fortsetzung).

Anwendung des **Mittelwertsatzes** auf jeden Term der rechten Seite ergibt somit

$$\begin{aligned} f(x) - f(x^0) &= \frac{\partial f}{\partial x_n}(x_1, \dots, x_{n-1}, \xi_n) \cdot (x_n - x_n^0) \\ &+ \frac{\partial f}{\partial x_{n-1}}(x_1, \dots, x_{n-2}, \xi_{n-1}, x_n^0) \cdot (x_{n-1} - x_{n-1}^0) \\ &\vdots \\ &+ \frac{\partial f}{\partial x_1}(\xi_1, x_2^0, \dots, x_n^0) \cdot (x_1 - x_1^0) \end{aligned}$$

Mit der Beschränktheit der partiellen Ableitungen gilt

$$|f(x) - f(x^0)| \leq C_1|x_1 - x_1^0| + \dots + C_n|x_n - x_n^0|$$

für  $\|x - x^0\|_\infty < \varepsilon$ , und damit ist  $f(x)$  **stetig** in  $x^0$ , denn es gilt

$$f(x) \rightarrow f(x^0) \quad \text{für } \|x - x^0\|_\infty \rightarrow 0$$

# Höhere Ableitungen.

**Definition:** Eine skalare Funktion  $f(x)$  sei auf einer offenen Menge  $D \subset \mathbb{R}^n$  partiell differenzierbar. Sind die partiellen Ableitungen erneut partiell differenzierbar, so erhält man sämtliche **partiellen Ableitungen zweiter Ordnung** von  $f$  mit

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

**Beispiel:** Partielle Ableitungen zweiter Ordnung einer Funktion  $f(x, y)$ :

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial x} \right), \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right), \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y^2}$$

Seien nun  $i_1, \dots, i_k \in \{1, \dots, n\}$ . Dann definiert man rekursiv

$$\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \dots \partial x_{i_1}} := \frac{\partial}{\partial x_{i_k}} \left( \frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \partial x_{i_{k-2}} \dots \partial x_{i_1}} \right)$$

# Ableitungen höherer Ordnung.

**Definition:** Die Funktion  $f(x)$  heißt  $k$ -fach partiell differenzierbar, falls alle Ableitungen der Ordnung  $k$ ,

$$\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \dots \partial x_{i_1}} \quad \text{für alle } i_1, \dots, i_k \in \{1, \dots, n\},$$

auf  $D$  existieren.

Alternative Notationen:

$$\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \dots \partial x_{i_1}} = D_{i_k} D_{i_{k-1}} \dots D_{i_1} f = f_{x_{i_1} \dots x_{i_k}}$$

Sind alle Ableitungen  $k$ -ter Ordnung stetig, so heißt die Funktion  $f(x)$   $k$ -fach stetig partiell differenzierbar oder auch  $C^k$ -Funktion auf  $D$ . Stetige Funktionen  $f(x)$  nennt man auch  $C^0$ -Funktionen.

**Beispiel:** Für die Funktion  $f(x_1, \dots, x_n) = \prod_{i=1}^n x_i^i$  gilt  $\frac{\partial^n f}{\partial x_n \dots \partial x_1} = ?$



# Partielle Ableitungen sind nicht beliebig vertauschbar.

**ACHTUNG:** Die Reihenfolge, in der die partiellen Ableitungen durchzuführen sind, ist im Allgemeinen **nicht** beliebig vertauschbar!

**Beispiel:** Für die Funktion

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

berechnet man direkt

$$f_{xy}(0, 0) = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x}(0, 0) \right) = -1$$

$$f_{yx}(0, 0) = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y}(0, 0) \right) = +1$$

d.h.  $f_{xy}(0, 0) \neq f_{yx}(0, 0)$ .

# Vertauschbarkeitssatz von Schwarz.

**Satz:** Ist  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$  offen, eine  $C^2$ -Funktion, so gilt

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x_1, \dots, x_n) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_1, \dots, x_n)$$

für alle  $i, j \in \{1, \dots, n\}$ .

**Beweisidee:**

Zweifache Anwendung des Mittelwertsatzes.

**Folgerung:**

Ist  $f(x)$  eine  $C^k$ -Funktion, so kann man die Reihenfolge der Differentiationen zur Berechnung der partiellen Ableitungen bis zur  $k$ -ten Ordnung **beliebig** vertauschen!

# Beispiel zur Vertauschbarkeit partieller Ableitungen.

Berechne für die Funktion

$$f(x, y, z) = y^2 z \sin(x^3) + (\cosh y + 17e^{x^2})z^2$$

die partielle Ableitung dritter Ordnung  $f_{xyz}$ .

Die Reihenfolge der partiellen Ableitungen ist vertauschbar, da  $f \in \mathcal{C}^3$ .

- Differenziere zunächst nach  $z$ :

$$\frac{\partial f}{\partial z} = y^2 \sin(x^3) + 2z(\cosh y + 17e^{x^2})$$

- Differenziere dann  $f_z$  nach  $x$  (damit fällt  $\cosh y$  raus):

$$\begin{aligned} f_{zx} &= \frac{\partial}{\partial x} \left( y^2 \sin(x^3) + 2z(\cosh y + 17e^{x^2}) \right) \\ &= 3x^2 y^2 \cos(x^3) + 68xze^{x^2} \end{aligned}$$

- Für die partielle Ableitung von  $f_{zx}$  nach  $y$  erhalten wir schließlich

$$f_{xyz} = 6x^2 y \cos(x^3)$$

# Der Laplace-Operator.

Der Laplace-Operator ist definiert durch

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

Für eine skalare Funktion  $u(x) = u(x_1, \dots, x_n)$  gilt somit

$$\Delta u = \sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2} = u_{x_1 x_1} + \dots + u_{x_n x_n}$$

Beispiele für wichtige partielle Differentialgleichungen zweiter Ordnung:

$$\Delta u - \frac{1}{c^2} u_{tt} = 0 \quad (\text{Wellengleichung})$$

$$\Delta u - \frac{1}{k} u_t = 0 \quad (\text{Wärmeleitungsgleichung})$$

$$\Delta u = 0 \quad (\text{Laplace-Gleichung oder Potentialgleichung})$$

# Vektorwertige Funktionen.

**Definition:** Sei  $f : D \rightarrow \mathbb{R}^m$ ,  $D \subset \mathbb{R}^n$  offen,  $f = (f_1, \dots, f_m)^T$ , eine vektorwertige Funktion.

Die Funktion  $f$  heißt **partiell differenzierbar** in  $x^0 \in D$ , falls für alle  $i = 1, \dots, n$  die Grenzwerte

$$\frac{\partial f}{\partial x_i}(x^0) = \lim_{t \rightarrow 0} \frac{f(x^0 + te_i) - f(x^0)}{t}$$

existieren. Die Berechnung erfolgt komponentenweise

$$\frac{\partial f}{\partial x_i}(x^0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_i} \\ \vdots \\ \frac{\partial f_m}{\partial x_i} \end{pmatrix} \quad \text{für } i = 1, \dots, n$$

# Vektorfelder.

**Definition:** Für  $m = n$  nennt man die Funktion  $f : D \rightarrow \mathbb{R}^n$  ein **Vektorfeld** auf  $D$ . Ist jede Koordinatenfunktion  $f_i(x)$  von  $f = (f_1, \dots, f_n)^T$  eine  $C^k$ -Funktion, so nennt man  $f$  ein  **$C^k$ -Vektorfeld**.

## Beispiele für Vektorfelder:

- Geschwindigkeitsfelder von strömenden Flüssigkeiten oder Gasen;
- elektromagnetische Felder;
- Temperaturgradienten in Festkörpern.

**Definition:** Für ein partiell differenzierbares Vektorfeld  $f : D \rightarrow \mathbb{R}^n$  definiert man die **Divergenz** in  $x \in D$  durch

$$\operatorname{div} f(x^0) := \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(x^0)$$

oder

$$\operatorname{div} f(x) = \nabla^T f(x) = (\nabla, f(x))$$

# Rechenregeln und Rotation.

Es gelten die folgenden Rechenregeln:

$$\operatorname{div}(\alpha f + \beta g) = \alpha \operatorname{div} f + \beta \operatorname{div} g \quad \text{für } f, g : D \rightarrow \mathbb{R}^n$$

$$\operatorname{div}(\varphi \cdot f) = (\nabla \varphi, f) + \varphi \operatorname{div} f \quad \text{für } \varphi : D \rightarrow \mathbb{R}, f : D \rightarrow \mathbb{R}^n$$

**Bemerkung:** Ist  $f : D \rightarrow \mathbb{R}$  eine  $\mathcal{C}^2$ -Funktion, so gilt für den Laplace-Operator

$$\Delta f = \operatorname{div}(\nabla f)$$

**Definition:** Für ein partiell differenzierbares Vektorfeld im  $\mathbb{R}^3$ ,  $f : D \rightarrow \mathbb{R}^3$ ,  $D \subset \mathbb{R}^3$  offen, definiert man die **Rotation** durch

$$\operatorname{rot} f(x^0) := \left( \frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right)^T \Bigg|_{x^0}$$

# Alternativ Notationen und weitere Rechenregeln.

$$\operatorname{rot} f(\mathbf{x}) = \nabla \times f(\mathbf{x}) = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ f_1 & f_2 & f_3 \end{vmatrix}$$

**Bemerkung:** Es gelten die folgenden Rechenregeln:

$$\operatorname{rot}(\alpha f + \beta g) = \alpha \operatorname{rot} f + \beta \operatorname{rot} g$$

$$\operatorname{rot}(\varphi \cdot f) = (\nabla \varphi) \times f + \varphi \operatorname{rot} f$$

**Bemerkung:** Ist  $\varphi : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^3$ , eine  $\mathcal{C}^2$ -Funktion, so folgt

$$\operatorname{rot}(\nabla \varphi) = 0,$$

mit dem Vertauschbarkeitssatz von Schwarz, d.h. Gradientenfelder sind stets **rotationsfrei**.



## 1.2 Das vollständige Differential

**Definition:** Sei  $D \subset \mathbb{R}^n$  offen,  $x^0 \in D$  und  $f : D \rightarrow \mathbb{R}^m$ . Die Funktion  $f(x)$  heißt **differenzierbar** in  $x^0$  (oder **vollständig differenzierbar** bzw. **total differenzierbar** in  $x_0$ ), falls es eine lineare Abbildung

$$l(x, x^0) := A \cdot (x - x^0)$$

mit einer Matrix  $A \in \mathbb{R}^{m \times n}$  gibt, für die die Approximationseigenschaft

$$f(x) = f(x^0) + A \cdot (x - x^0) + o(\|x - x^0\|)$$

gilt, d.h.

$$\lim_{x \rightarrow x^0} \frac{f(x) - f(x^0) - A \cdot (x - x^0)}{\|x - x^0\|} = 0.$$

# Das vollständige Differential und die Jacobi-Matrix.

**Bezeichnungen:** Man nennt die lineare Abbildung  $l$  das **vollständige Differential** oder das **totale Differential** von  $f(x)$  im Punkt  $x^0$ , und man bezeichnet  $l$  mit  $df(x^0)$ .

Die zugehörige Matrix  $A$  heißt **Jacobi-Matrix** oder **Funktionalmatrix** von  $f(x)$  im Punkt  $x^0$  und wird mit  $Jf(x^0)$  (manchmal auch mit  $Df(x^0)$  oder  $f'(x^0)$ ) bezeichnet.

**Bemerkung:** Für  $m = n = 1$  erhalten wir die bekannte Beziehung

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(|x - x_0|)$$

für die Ableitung  $f'(x_0)$  im Punkt  $x_0$ .

**Bemerkung:** Im Fall einer skalaren Funktion ( $m = 1$ ) ist  $A = a$  ein Zeilenvektor und  $a(x - x^0)$  ein Skalarprodukt  $\langle a^T, x - x^0 \rangle$ .

# Vollständige und partielle Differenzierbarkeit.

**Satz:** Sei  $f : D \rightarrow \mathbb{R}^m$ ,  $x^0 \in D \subset \mathbb{R}^n$ ,  $D$  offen.

- a) Ist  $f(x)$  in  $x^0$  differenzierbar, so ist  $f(x)$  auch stetig in  $x^0$ .
- b) Ist  $f(x)$  in  $x^0$  differenzierbar, so ist das (vollständige) Differential und damit auch die Jacobi-Matrix eindeutig bestimmt und es gilt

$$Jf(x^0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^0) & \dots & \frac{\partial f_1}{\partial x_n}(x^0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x^0) & \dots & \frac{\partial f_m}{\partial x_n}(x^0) \end{pmatrix} = \begin{pmatrix} Df_1(x^0) \\ \vdots \\ Df_m(x^0) \end{pmatrix}$$

- c) Ist  $f(x)$  eine  $C^1$ -Funktion auf  $D$ , so ist  $f(x)$  auf  $D$  differenzierbar.

## Beweis von a).

Ist  $f$  in  $x^0$  differenzierbar, so gilt nach Definition

$$\lim_{x \rightarrow x^0} \frac{f(x) - f(x^0) - A \cdot (x - x^0)}{\|x - x^0\|} = 0$$

Daraus folgt aber

$$\lim_{x \rightarrow x^0} \|f(x) - f(x^0) - A \cdot (x - x^0)\| = 0$$

und wir erhalten

$$\begin{aligned} \|f(x) - f(x^0)\| &\leq \|f(x) - f(x^0) - A \cdot (x - x^0)\| + \|A \cdot (x - x^0)\| \\ &\rightarrow 0 \quad \text{für } x \rightarrow x^0 \end{aligned}$$

Damit ist die Funktion  $f$  stetig im Punkt  $x^0$ .

## Beweis von b).

Sei  $x = x^0 + te_i$ ,  $|t| < \varepsilon$ ,  $i \in \{1, \dots, n\}$ . Da  $f$  im Punkt  $x^0$  differenzierbar ist, folgt

$$\lim_{x \rightarrow x^0} \frac{f(x) - f(x^0) - A \cdot (x - x^0)}{\|x - x^0\|_\infty} = 0$$

Wir schreiben nun

$$\begin{aligned} \frac{f(x) - f(x^0) - A \cdot (x - x^0)}{\|x - x^0\|_\infty} &= \frac{f(x^0 + te_i) - f(x^0)}{|t|} - \frac{tAe_i}{|t|} \\ &= \frac{t}{|t|} \cdot \left( \frac{f(x^0 + te_i) - f(x^0)}{t} - Ae_i \right) \\ &\rightarrow 0 \quad \text{für } t \rightarrow 0 \end{aligned}$$

Daraus folgt

$$\lim_{t \rightarrow 0} \frac{f(x^0 + te_i) - f(x^0)}{t} = Ae_i \quad i = 1, \dots, n$$

# Beispiele.

- Betrachte die skalare Funktion  $f(x_1, x_2) = x_1 e^{2x_2}$ . Dann lautet die Jacobi-Matrix:

$$Jf(x_1, x_2) = Df(x_1, x_2) = e^{2x_2} (1, 2x_1)$$

- Betrachte die Funktion  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  definiert durch

$$f(x_1, x_2, x_3) = \begin{pmatrix} x_1 x_2 x_3 \\ \sin(x_1 + 2x_2 + 3x_3) \end{pmatrix}$$

Die Jacobi-Matrix ergibt sich in der Form

$$Jf(x_1, x_2, x_3) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \end{pmatrix} = \begin{pmatrix} x_2 x_3 & x_1 x_3 & x_1 x_2 \\ \cos(s) & 2 \cos(s) & 3 \cos(s) \end{pmatrix}$$

wobei  $s = x_1 + 2x_2 + 3x_3$ .

## Weitere Beispiele.

- Sei  $f(x) = Ax$ ,  $A \in \mathbb{R}^{m \times n}$  und  $x \in \mathbb{R}^n$ . Dann gilt

$$Jf(x) = A \quad \text{für alle } x \in \mathbb{R}^n$$

- Sei  $f(x) = x^T Ax = \langle x, Ax \rangle$ ,  $A \in \mathbb{R}^{n \times n}$  und  $x \in \mathbb{R}^n$ .  
Dann gilt

$$\begin{aligned} \frac{\partial f}{\partial x_j} &= \langle e_j, Ax \rangle + \langle x, Ae_j \rangle \\ &= e_j^T Ax + x^T Ae_j \\ &= x^T (A^T + A) e_j \end{aligned}$$

Daraus folgt

$$Jf(x) = \text{grad}f(x) = x^T (A^T + A)$$

# Differentiationsregeln.

## Satz:

- a) **Linearität:** Sind  $f, g : D \rightarrow \mathbb{R}^m$  differenzierbar in  $x^0 \in D$ ,  $D$  offen, so ist auch  $\alpha f(x^0) + \beta g(x^0)$ ,  $\alpha, \beta \in \mathbb{R}$ , differenzierbar in  $x^0$  und es gilt

$$d(\alpha f + \beta g)(x^0) = \alpha df(x^0) + \beta dg(x^0)$$

$$J(\alpha f + \beta g)(x^0) = \alpha Jf(x^0) + \beta Jg(x^0)$$

- b) **Kettenregel:** Ist  $f : D \rightarrow \mathbb{R}^m$  differenzierbar in  $x^0 \in D$ ,  $D$  offen, und ist  $g : E \rightarrow \mathbb{R}^k$  differenzierbar in  $y^0 = f(x^0) \in E \subset \mathbb{R}^m$ ,  $E$  offen, so ist  $g \circ f$  ebenfalls in  $x^0$  differenzierbar.

Für die Differentiale gilt

$$d(g \circ f)(x^0) = dg(y^0) \circ df(x^0)$$

und analog für die Jacobi-Matrizen

$$J(g \circ f)(x^0) = Jg(y^0) \cdot Jf(x^0)$$



## Beispiel zur Kettenregel.

Sei  $h : I \rightarrow \mathbb{R}^n$ ,  $I \subset \mathbb{R}$  Intervall, eine in  $t_0 \in I$  differenzierbare Kurve mit Werten in  $D \subset \mathbb{R}^n$ ,  $D$  offen, und  $f : D \rightarrow \mathbb{R}$  eine in  $x^0 = h(t_0)$  differenzierbare skalare Funktion.

Dann ist auch die Hintereinanderausführung

$$(f \circ h)(t) = f(h_1(t), \dots, h_n(t))$$

in  $t_0$  differenzierbar, und für die Ableitung gilt:

$$\begin{aligned}(f \circ h)'(t_0) &= Jf(h(t_0)) \cdot Jh(t_0) \\ &= \operatorname{grad} f(h(t_0)) \cdot h'(t_0) \\ &= \sum_{k=1}^n \frac{\partial f}{\partial x_k}(h(t_0)) \cdot h'_k(t_0)\end{aligned}$$

# Richtungsableitungen.

**Definition:** Sei  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$  offen,  $x^0 \in D$ , und  $v \in \mathbb{R} \setminus \{0\}$  ein Vektor. Dann heißt

$$D_v f(x^0) := \lim_{t \rightarrow 0} \frac{f(x^0 + tv) - f(x^0)}{t}$$

die **Richtungsableitung (Gateaux–Ableitung)** von  $f(x)$  in Richtung  $v$ .

**Beispiel:** Sei  $f(x, y) = x^2 + y^2$  und  $v = (1, 1)^T$ . Dann gilt für die Richtungsableitung in Richtung  $v$ :

$$\begin{aligned} D_v f(x, y) &= \lim_{t \rightarrow 0} \frac{(x+t)^2 + (y+t)^2 - x^2 - y^2}{t} \\ &= \lim_{t \rightarrow 0} \frac{2xt + t^2 + 2yt + t^2}{t} \\ &= 2(x + y) \end{aligned}$$

# Bemerkungen.

- Für  $v = e_j$  ist die Richtungsableitung in Richtung  $v$  gegeben durch die partielle Ableitung nach der Koordinatenrichtung  $x_j$ :

$$D_v f(x^0) = \frac{\partial f}{\partial x_j}(x^0)$$

- Ist  $v$  ein Einheitsvektor, also  $\|v\| = 1$ , so beschreibt die Richtungsableitung  $D_v f(x^0)$  den **Anstieg** (bzw. die **Steigung**) von  $f(x)$  in Richtung  $v$ .
- Ist  $f(x)$  in  $x^0$  differenzierbar, so existieren sämtliche Richtungsableitungen von  $f(x)$  in  $x^0$  und mit  $h(t) = x^0 + tv$  gilt

$$D_v f(x^0) = \frac{d}{dt}(f \circ h)|_{t=0} = \text{grad } f(x^0) \cdot v$$

Dies folgt unmittelbar aus der Anwendung der Kettenregel.

# Eigenschaften des Gradienten.

**Satz:** Sei  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$  offen, in  $x^0 \in D$  differenzierbar. Dann gilt

- a) Der Gradientenvektor  $\text{grad } f(x^0) \in \mathbb{R}^n$  steht senkrecht auf der **Niveaumenge**

$$N_{x^0} := \{x \in D \mid f(x) = f(x^0)\}$$

Im Fall  $n = 2$  nennt man die Niveaumengen auch **Höhenlinien**, im Fall  $n = 3$  heißen die Niveaumengen auch **Äquipotentialflächen**.

- 2) Der Gradient  $\text{grad } f(x^0)$  gibt die Richtung des steilsten Anstiegs von  $f(x)$  in  $x^0$  an.

**Beweisidee:**

- a) Anwendung der Kettenregel.  
b) Für beliebige Richtung  $v$  gilt mit der Cauchy-Schwarzschen Ungleichung

$$|D_v f(x^0)| = |(\text{grad } f(x^0), v)| \leq \|\text{grad } f(x^0)\|_2$$

Gleichheit wird für  $v = \text{grad } f(x^0) / \|\text{grad } f(x^0)\|_2$  angenommen.

# Krummlinige Koordinaten.

**Definition:** Sei  $\Phi : U \rightarrow V$ ,  $U, V \subset \mathbb{R}^n$  offen, eine  $\mathcal{C}^1$ -Abbildung, für die die Jacobimatrix  $J\Phi(u^0)$  an jeder Stelle  $u^0 \in U$  regulär ist.

Weiterhin existiere die Umkehrabbildung  $\Phi^{-1} : V \rightarrow U$  und diese sei ebenfalls eine  $\mathcal{C}^1$ -Abbildung.

Dann definiert  $x = \Phi(u)$  eine **Koordinatentransformation** von den Koordinaten  $u$  auf  $x$ .

**Beispiel:** Betrachte für  $n = 2$  die **Polarkoordinaten**  $u = (r, \varphi)$  mit  $r > 0$  und  $-\pi < \varphi < \pi$  und setze

$$x = r \cos \varphi$$

$$y = r \sin \varphi$$

mit den **kartesischen Koordinaten**  $x = (x, y)$ .

# Umrechnung der partiellen Ableitungen.

Für alle  $u \in U$  mit  $x = \Phi(u)$  gelten die Relationen

$$\Phi^{-1}(\Phi(u)) = u$$

$$J\Phi^{-1}(x) \cdot J\Phi(u) = I_n \quad (\text{Kettenregel})$$

$$J\Phi^{-1}(x) = (J\Phi(u))^{-1}$$

Sei nun  $\tilde{f} : V \rightarrow \mathbb{R}$  eine gegebene Funktion und setze

$$f(u) := \tilde{f}(\Phi(u))$$

Dann folgt aus der Kettenregel:

$$\frac{\partial f}{\partial u_i} = \sum_{j=1}^n \frac{\partial \tilde{f}}{\partial x_j} \frac{\partial \Phi_j}{\partial u_i} =: \sum_{j=1}^n g^{ij} \frac{\partial \tilde{f}}{\partial x_j}$$

mit

$$g^{ij} := \frac{\partial \Phi_j}{\partial u_i}, \quad G(u) := (g^{ij}) = (J\Phi(u))^T$$

# Notationen.

Wir verwenden die abkürzende Schreibweise

$$\frac{\partial}{\partial u_i} = \sum_{j=1}^n g^{ij} \frac{\partial}{\partial x_j}$$

Analog lassen sich die partiellen Ableitungen nach  $x_i$  durch die partiellen Ableitungen nach  $u_j$  ausdrücken mit

$$\frac{\partial}{\partial x_i} = \sum_{j=1}^n g_{ij} \frac{\partial}{\partial u_j}$$

wobei

$$(g_{ij}) := (g^{ij})^{-1} = (J\Phi)^{-T} = (J\Phi^{-1})^T$$

Man erhält diese Beziehungen durch Anwendung der Kettenregel auf  $\Phi^{-1}$ .

# Beispiel: Polarkoordinaten.

Wir betrachten die Polarkoordinaten

$$x = \Phi(u) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$$

Dann berechnet man

$$J\Phi(u) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

und damit

$$(g^{ij}) = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -r \sin \varphi & r \cos \varphi \end{pmatrix} \quad (g_{ij}) = \begin{pmatrix} \cos \varphi & -\frac{1}{r} \sin \varphi \\ \sin \varphi & \frac{1}{r} \cos \varphi \end{pmatrix}$$



# Partielle Ableitungen für die Polarkoordinaten.

Für die Umrechnung der partiellen Ableitungen bekommt man nun

$$\frac{\partial}{\partial x} = \cos \varphi \frac{\partial}{\partial r} - \frac{1}{r} \sin \varphi \frac{\partial}{\partial \varphi}$$

$$\frac{\partial}{\partial y} = \sin \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \varphi \frac{\partial}{\partial \varphi}$$

**Beispiel:** Umrechnung des **Laplace-Operator** auf Polarkoordinaten

$$\frac{\partial^2}{\partial x^2} = \cos^2 \varphi \frac{\partial^2}{\partial r^2} - \frac{\sin(2\varphi)}{r} \frac{\partial^2}{\partial r \partial \varphi} + \frac{\sin^2 \varphi}{r^2} \frac{\partial^2}{\partial \varphi^2} + \frac{\sin(2\varphi)}{r^2} \frac{\partial}{\partial \varphi} + \frac{\sin^2 \varphi}{r} \frac{\partial}{\partial r}$$

$$\frac{\partial^2}{\partial y^2} = \sin^2 \varphi \frac{\partial^2}{\partial r^2} + \frac{\sin(2\varphi)}{r} \frac{\partial^2}{\partial r \partial \varphi} + \frac{\cos^2 \varphi}{r^2} \frac{\partial^2}{\partial \varphi^2} - \frac{\sin(2\varphi)}{r^2} \frac{\partial}{\partial \varphi} + \frac{\cos^2 \varphi}{r} \frac{\partial}{\partial r}$$

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} + \frac{1}{r} \frac{\partial}{\partial r}$$

# Beispiel: Kugelkoordinaten.

Wir betrachten die Kugelkoordinaten

$$x = \Phi(u) = \begin{pmatrix} r \cos \varphi \cos \theta \\ r \sin \varphi \cos \theta \\ r \sin \theta \end{pmatrix}$$

Die Jacobi-Matrix ist dann gegeben durch:

$$J\Phi(u) = \begin{pmatrix} \cos \varphi \cos \theta & -r \sin \varphi \cos \theta & -r \cos \varphi \sin \theta \\ \sin \varphi \cos \theta & r \cos \varphi \cos \theta & -r \sin \varphi \sin \theta \\ \sin \theta & 0 & r \cos \theta \end{pmatrix}$$

# Partielle Ableitungen für die Kugelkoordinaten.

Für die Umrechnung der partiellen Ableitungen bekommt man nun

$$\frac{\partial}{\partial x} = \cos \varphi \cos \theta \frac{\partial}{\partial r} - \frac{\sin \varphi}{r \cos \theta} \frac{\partial}{\partial \varphi} - \frac{1}{r} \cos \varphi \sin \theta \frac{\partial}{\partial \theta}$$

$$\frac{\partial}{\partial y} = \sin \varphi \cos \theta \frac{\partial}{\partial r} + \frac{\cos \varphi}{r \cos \theta} \frac{\partial}{\partial \varphi} - \frac{1}{r} \sin \varphi \sin \theta \frac{\partial}{\partial \theta}$$

$$\frac{\partial}{\partial z} = \sin \theta \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \frac{\partial}{\partial \theta}$$

**Beispiel:** Umrechnung des **Laplace-Operators** auf Kugelkoordinaten

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r^2 \cos^2 \theta} \frac{\partial^2}{\partial \varphi^2} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{\tan \theta}{r^2} \frac{\partial}{\partial \theta}$$

# Kapitel 1. Differentialrechnung mehrerer Variablen

## 1.3 Mittelwertsätze und Taylor-Entwicklungen

**Satz (Mittelwertsatz):** Sei  $f : D \rightarrow \mathbb{R}$  eine auf einer offenen Menge  $D \subset \mathbb{R}^n$  differenzierbare, skalare Funktion. Weiterhin seien  $a, b \in D$  Punkte in  $D$ , so dass die Verbindungsstrecke

$$[a, b] := \{a + t(b - a) \mid t \in [0, 1]\}$$

ganz in  $D$  liegt. Dann gibt es eine Zahl  $\theta \in (0, 1)$  mit

$$f(b) - f(a) = \text{grad } f(a + \theta(b - a)) \cdot (b - a)$$

**Beweis:** Wir setzen

$$h(t) := f(a + t(b - a))$$

Aus dem Mittelwertsatz für **eine** Veränderliche folgt dann mit der Kettenregel

$$\begin{aligned} f(b) - f(a) &= h(1) - h(0) = h'(\theta) \cdot (1 - 0) \\ &= \text{grad } f(a + \theta(b - a)) \cdot (b - a) \end{aligned}$$

# Definition und Beispiel.

**Definition:** Gilt die Bedingung  $[a, b] \subset D$  für **alle** Punkte  $a, b \in D$ , so heißt die Menge  $D$  **konvex**.

**Beispiel zum Mittelwertsatz:** Gegeben sei die skalare Funktion

$$f(x, y) := \cos x + \sin y$$

Offensichtlich gilt

$$f(0, 0) = f(\pi/2, \pi/2) = 1 \quad \Rightarrow \quad f(\pi/2, \pi/2) - f(0, 0) = 0$$

Nach dem Mittelwertsatz existiert ein  $\theta \in (0, 1)$  mit

$$\text{grad } f \left( \theta \begin{pmatrix} \pi/2 \\ \pi/2 \end{pmatrix} \right) \cdot \begin{pmatrix} \pi/2 \\ \pi/2 \end{pmatrix} = 0$$

In der Tat gilt diese Beziehung für  $\theta = \frac{1}{2}$ .

# Mittelwertsatz gilt nur für **skalare** Funktionen.

**Beachte:** Der Mittelwertsatz für mehrere Variablen gilt nur für **skalare** Funktionen, aber i.A. nicht für **vektorwertige** Funktionen!

**Beispiel:** Betrachte die **vektorwertige** Funktion

$$f(t) := \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \quad t \in [0, \pi/2]$$

Nun gilt

$$f(\pi/2) - f(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

und

$$f' \left( \theta \frac{\pi}{2} \right) \cdot \left( \frac{\pi}{2} - 0 \right) = \frac{\pi}{2} \begin{pmatrix} -\sin(\theta\pi/2) \\ \cos(\theta\pi/2) \end{pmatrix}$$

**ABER:** Die Vektoren auf der rechten Seite haben die Längen  $\sqrt{2}$  bzw.  $\pi/2$  !

# Der Mittelwert–Abschätzungssatz für vektorwertige Funktionen.

**Satz:** Die Funktion  $f : D \rightarrow \mathbb{R}^m$  sei differenzierbar auf der offenen Menge  $D \subset \mathbb{R}^n$ . Weiterhin seien  $a, b$  Punkte in  $D$  mit  $[a, b] \subset D$ . Dann gibt es ein  $\theta \in (0, 1)$  mit

$$\|f(b) - f(a)\|_2 \leq \|Jf(a + \theta(b - a)) \cdot (b - a)\|_2$$

**Beweisidee:** Anwendung des Mittelwertsatzes auf die **skalare** Funktion  $g(x)$  definiert durch

$$g(x) := (f(b) - f(a))^T f(x) \quad (\text{Skalarprodukt!})$$

**Bemerkung:** Eine andere (abgeschwächte) Form der Mittelwert–Abschätzung ist

$$\|f(b) - f(a)\| \leq \sup_{\xi \in [a, b]} \|Jf(\xi)\| \cdot \|(b - a)\|$$

wobei  $\|\cdot\|$  eine beliebige Vektor– bzw. zugehörige Matrixnorm ist.

# Taylor-Entwicklungen: Notationen.

Zunächst definieren wir einen **Multiindex**  $\alpha \in \mathbb{N}_0^n$  als

$$\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$$

Weiterhin sei

$$|\alpha| := \alpha_1 + \dots + \alpha_n \quad \alpha! := \alpha_1! \cdots \alpha_n!$$

Ist  $f : D \rightarrow \mathbb{R}$   $|\alpha|$ -mal stetig differenzierbar, so setzen wir

$$D^\alpha = D_1^{\alpha_1} D_2^{\alpha_2} \dots D_n^{\alpha_n} = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}},$$

wobei  $D_i^{\alpha_i} = \underbrace{D_i \dots D_i}_{\alpha_i\text{-mal}}$  und wir schreiben

$$x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n} \quad \text{für } x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$



# Der Satz von Taylor.

## Satz: (Satz von Taylor)

Sei  $D \subset \mathbb{R}^n$  offen und konvex,  $f : D \rightarrow \mathbb{R}$  eine  $C^{m+1}$ -Funktion und sei  $x_0 \in D$ . Dann gilt für  $x \in D$  die **Taylor-Entwicklung**

$$f(x) = T_m(x; x_0) + R_m(x; x_0)$$
$$T_m(x; x_0) = \sum_{|\alpha| \leq m} \frac{D^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha = \underbrace{f(x_0)}_{T_0} + \underbrace{\text{gerade } f(x_0) \cdot (x - x_0)}_{T_1} + \dots$$
$$R_m(x; x_0) = \sum_{|\alpha|=m+1} \frac{D^\alpha f(x_0 + \theta(x - x_0))}{\alpha!} (x - x_0)^\alpha$$

*Handwritten notes:*  
- Above the first equation:  $\frac{\partial}{\partial x_1} f(x_0) (x_1 - x_1^0) + \frac{\partial}{\partial x_2} f(x_0) (x_2 - x_2^0) + \dots$   
- In the second equation:  $T_0$  and  $T_1$  are written below the terms, with a bracket under the second term.

mit einem geeigneten  $\theta \in (0, 1)$ .

**Bezeichnung:** In der obigen Taylor-Entwicklung heißt  $T_m(x; x_0)$  **Taylor-Polynom  $m$ -ten Grades** und  $R_m(x; x_0)$  wird als **Lagrange-Restglied** bezeichnet.

# Herleitung der Taylorsche Formel.

Wir definieren eine skalare Funktion **einer** Variablen  $t \in [0, 1]$  als

$$g(t) := f(x_0 + t(x - x_0))$$

$g(0) = f(x_0)$   
 $g(1) = f(x)$

und berechnen die Taylor-Entwicklung **um**  $t = 0$ . Es gilt: *1-dimensional*

$$g(1) = \underbrace{g(0)}_{T_1(1;0)} + g'(0) \cdot (1 - 0) + \underbrace{\frac{1}{2}g''(\xi) \cdot (1 - 0)^2}_{R_1(1;0)} \quad \text{für ein } \xi \in (0, 1).$$

Die Berechnung von  $g'(0)$  liefert mit der Kettenregel

$$\begin{aligned} g'(0) &= \left. \frac{d}{dt} f(x_1^0 + t(x_1 - x_1^0), x_2^0 + t(x_2 - x_2^0), \dots, x_n^0 + t(x_n - x_n^0)) \right|_{t=0} \\ &= D_1 f(x_0) \cdot (x_1 - x_1^0) + \dots + D_n f(x_0) \cdot (x_n - x_n^0) = \text{grad } f(x_0) \cdot (x - x^0) \\ &= \sum_{|\alpha|=1} \frac{D^\alpha f(x_0)}{\alpha!} \cdot (x - x_0)^\alpha \end{aligned}$$

# Fortsetzung der Herleitung.

Berechnung von  $g''(0)$  liefert

$$g''(0) = \frac{d^2}{dt^2} f(x_0 + t(x - x_0)) \Big|_{t=0} = \frac{d}{dt} \sum_{k=1}^n D_k f(x_0 + t(x - x_0)) (x_k - x_k^0) \Big|_{t=0}$$

*Handwritten notes:*  $f(x)$  above the first term;  $g(x) = f(x_0 + t(x - x_0)) \cdot (x - x_0)$  above the sum.

$$= \sum_{i,j=1}^h (x_j - x_j^0) \frac{\partial^2}{\partial x_i \partial x_j} f(x_0 + t(x - x_0)) (x_i - x_i^0) \Big|_{t=0} = (*)$$
$$= D_{11} f(x_0) (x_1 - x_1^0)^2 + D_{21} f(x_0) (x_1 - x_1^0) (x_2 - x_2^0)$$

$$+ \dots + D_{ij} f(x_0) (x_i - x_i^0) (x_j - x_j^0) + \dots +$$

$$+ D_{n-1,n} f(x_0) (x_{n-1} - x_{n-1}^0) (x_n - x_n^0) + D_{nn} f(x_0) (x_n - x_n^0)^2$$

$$(*) = \sum_{i,j=1}^h (x_j - x_j^0) \frac{\partial^2}{\partial x_i \partial x_j} f(x_0) (x_i - x_i^0) = (x - x^0)^T A (x - x^0)$$
$$= \sum_{|\alpha|=2} \frac{D^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha \quad (\text{Vertauschungssatz von Schwarz!})$$

*Handwritten note:*  $Q_{ij}$  below the sum.

**Nun:** Beweis der Taylor-Formel mittels vollständiger Induktion!

# Beweis des Satzes von Taylor.

Die Funktion

$$g(t) := f(x^0 + t(x - x^0))$$

ist  $(m + 1)$ -mal stetig differenzierbar, und es gilt

$$g(1) = \sum_{k=0}^m \frac{g^{(k)}(0)}{k!} + \frac{g^{(m+1)}(\theta)}{(m+1)!} \quad \text{für ein } \theta \in [0, 1].$$

Weiterhin gilt (per Induktion über  $k$ )

$$\frac{g^{(k)}(0)}{k!} = \sum_{|\alpha|=k} \frac{D^\alpha f(x^0)}{\alpha!} (x - x^0)^\alpha$$

und

$$\frac{g^{(m+1)}(\theta)}{(m+1)!} = \sum_{|\alpha|=m+1} \frac{D^\alpha f(x^0 + \theta(x - x^0))}{\alpha!} (x - x^0)^\alpha$$

## Beispiel zur Taylor-Entwicklung.

- 1 Berechne das Taylor-Polynom  $T_2(x; x_0)$  zweiten Grades der Funktion

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

zum Entwicklungspunkt  $(x, y, z) = (1, 2, 0)^T$ .

- 2 Die Berechnung von  $T_2(x; x_0)$  benötigt die partiellen Ableitungen bis zur zweiten Ordnung.
- 3 Diese Ableitungen müssen am Punkt  $(x, y, z) = (1, 2, 0)^T$  ausgewertet werden.
- 4 Als Ergebnis erhält man  $T_2(x; x_0)$  in der Form

$$T_2(x; x_0) = 4z(x + y - 2)$$

- 5 Berechnung auf Folie.

$$f(x, y, z) = xy^2 \sin z \quad x^0 = (1, 2, 0)$$

$$f(1, 2, 0) = 0$$

$$T_0 = 0$$

$$\text{grad } f(x, y, z) = (y^2 \sin z, 2xy \sin z, xy^2 \cos z)$$

$$\text{grad } f(1, 2, 0) \cdot \begin{pmatrix} x-1 \\ y-2 \\ z-0 \end{pmatrix} = (0, 0, 4) \cdot \begin{pmatrix} x-1 \\ y-2 \\ z \end{pmatrix} = 4z$$

$$T_1 = 0 + 4z$$

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

$$\begin{pmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\frac{1}{1!1!0!} \frac{\partial^2 f}{\partial x \partial z} \Big|_{x^0} (x-1)z$$

$$\alpha = (1, 0, 1)$$

$$\frac{1}{0!1!1!} \frac{\partial^2 f}{\partial y \partial z} \Big|_{x^0} (y-2)z$$

$$\alpha = (0, 1, 1)$$

## Bemerkung zum Restglied eines Taylor-Polynoms.

**Bemerkung:** Das Restglied eines Taylor-Polynoms enthält **alle** partiellen Ableitungen der Ordnung  $(m + 1)$ :

$f: \mathbb{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}$

$$R_m(x; x_0) = \sum_{|\alpha|=m+1} \frac{D^\alpha f(x_0 + \theta(x - x_0))}{\alpha!} (x - x_0)^\alpha$$

Sind all diese Ableitungen in der Nähe von  $x_0$  durch eine Konstante  $C$  beschränkt, so gilt die **Restgliedabschätzung**

$$\boxed{|R_m(x; x_0)|} \leq \frac{n^{m+1}}{(m+1)!} C \boxed{\|x - x_0\|_\infty^{m+1}}$$

Für die Approximationsgüte des Taylor-Polynoms einer  $\mathcal{C}^{m+1}$ -Funktion folgt daher

$$f(x) = T_m(x; x_0) + O(\|x - x_0\|^{m+1})$$

**Spezialfall**  $m = 1$ : Für eine  $\mathcal{C}^2$ -Funktion  $f(x)$  bekommt man

$$f(x) = \underbrace{f(x^0) + \text{grad } f(x^0) \cdot (x - x^0)}_{T_1(x; x_0)} + \underbrace{O(\|x - x^0\|^2)}_{\text{Restglied}}$$

# Die Hesse-Matrix.

Man nennt die Matrix

$$Hf(x_0) := \begin{pmatrix} f_{x_1 x_1}(x_0) & \dots & f_{x_1 x_n}(x_0) \\ \vdots & & \vdots \\ f_{x_n x_1}(x_0) & \dots & f_{x_n x_n}(x_0) \end{pmatrix}$$

die **Hesse-Matrix** von  $f(x)$  im Punkt  $x_0$ .

Hesse-Matrix = Jacobi-Matrix des Gradienten  $\nabla f$

Die Taylor-Entwicklung einer  $\mathcal{C}^3$ -Funktion lautet daher

$$f(x) = \underbrace{f(x_0) + \text{grad } f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^T Hf(x_0)(x - x_0)}_{T_2(x, x_0)} + O(\|x - x_0\|^3)$$

Die Hesse-Matrix einer  $\mathcal{C}^2$ -Funktion ist symmetrisch.



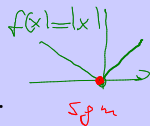
# Kapitel 2. Anwendungen der Differentialrechnung mehrerer Variablen

## 2.1 Extrema von Funktionen mehrerer Veränderlichen

**Definition:** Sei  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$ , und  $x^0 \in D$ . Dann hat  $f(x)$  in  $x^0$

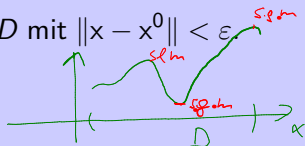
- ein **globales Maximum**, falls  $f(x) \leq f(x^0)$  für alle  $x \in D$ .
- ein **strenges globales Maximum**, falls  $f(x) < f(x^0)$  für alle  $x \in D$ .
- ein **lokales Maximum**, falls es ein  $\varepsilon > 0$  gibt mit

$$f(x) \leq f(x^0) \quad \text{für alle } x \in D \text{ mit } \|x - x^0\| < \varepsilon.$$



- ein **strenges lokales Maximum**, falls es ein  $\varepsilon > 0$  gibt mit

$$f(x) < f(x^0) \quad \text{für alle } x \in D \text{ mit } \|x - x^0\| < \varepsilon.$$



Analoge Definitionen für Minima.

# Notwendige Bedingung für lokale Extrema.

**Satz:** Besitzt eine  $C^1$ -Funktion  $f(x)$  in einem Punkt  $x^0 \in D^0$  ein **lokales Extremum** (Minimum oder Maximum), so gilt

$$\text{grad } f(x^0) = 0 \in \mathbb{R}^n$$

$f(x) = |x|$   
 $f \notin C^1$   $f'(0) \nexists$   
 $x=0$  Extremum

**Beweis:** Für ein beliebiges  $v \in \mathbb{R}^n$ ,  $v \neq 0$  ist die Funktion

$$\varphi(t) := f(x^0 + tv)$$

in einer Umgebung von  $t^0 = 0$  stetig differenzierbar.

Weiterhin hat  $\varphi(t)$  bei  $t^0 = 0$  ein lokales Extremum. Damit folgt:

$\varphi \in C^1$   $\Rightarrow 0 = \varphi'(0) = \text{grad } f(x^0) \cdot v$   $\Rightarrow \text{grad } f(x^0) \cdot v = 0$   
 $\forall v \neq 0$

*Handwritten notes:*  $f$  in  $x_0$  Extremum  
Kettenregel

Da dies für alle  $v \neq 0$  gilt, folgt die Bedingung:

$$\text{grad } f(x^0) = (0, \dots, 0)^T$$

*Handwritten notes:* z.B.  $v = \text{grad } f(x_0)$   
 $\Rightarrow |\text{grad } f(x_0)|^2 = 0$

# Bemerkungen zu lokalen Extremwerten.

## Bemerkungen:

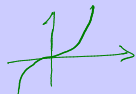
$grad f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  Vektorfeld

- Die Bedingung  $grad f(x^0) = 0$  liefert gewöhnlich ein **nichtlineares** Gleichungssystem zur Berechnung von  $x = x^0$  mit  $n$  Gleichungen und  $n$  Unbekannten.
- Die Punkte  $x^0 \in D^0$  mit  $grad f(x^0) = 0$  nennt man **stationäre Punkte** von  $f(x)$ . Stationäre Punkte sind **nicht** notwendigerweise lokale Extremwerte. Zum Beispiel besitzt die Funktion

$$f(x, y) := x^2 - y^2$$

$$f(x) = x^3 \quad 0 \text{ stat. P.}$$
$$f'(0) = 0$$

$$grad f(x, y) = 2(x, -y)$$



den Gradienten

und hat daher nur einen stationären Punkt  $x^0 = (0, 0)^T$ . Der Punkt  $x^0$  ist jedoch ein **Sattelpunkt** von  $f$ , d.h. in jeder Umgebung von  $x^0$  gibt es zwei Punkte  $x^1$  und  $x^2$  mit

$$f(x^1) < f(x^0) < f(x^2).$$

# Klassifikation stationärer Punkte.

**Satz:** Sei  $f(x)$  eine  $\mathcal{C}^2$ -Funktion auf  $D^0$  und  $x^0 \in D^0$  ein stationärer Punkt von  $f(x)$ , d.h.  $\text{grad } f(x^0) = 0$ .

## a) Notwendige Bedingung

Ist  $x^0$  ein lokales Extremum von  $f(x)$ , so gilt:

$x^0$  lokales Minimum  $\Rightarrow H f(x^0)$  positiv semidefinit

$x^0$  lokales Maximum  $\Rightarrow H f(x^0)$  negativ semidefinit

## b) Hinreichende Bedingung

Ist  $H f(x^0)$  positiv definit (bzw. negativ definit), so ist  $x^0$  ein strenges lokales Minimum (bzw. Maximum) von  $f(x)$ .

Ist  $H f(x^0)$  indefinit, so ist  $x^0$  ein Sattelpunkt, d.h. es gibt in jeder Umgebung von  $x^0$  Punkte  $x^1$  und  $x^2$  mit  $f(x^1) < f(x^0) < f(x^2)$ .

$x_0$  stat. P.,  $f \in C^2$

$x_0$  strong local minimum  $\not\Rightarrow Hf(x_0)$  positiv definit

$$f(x) = x^4$$

$x_0 = 0$  stat. pt.

$f''(x_0) = 0$   $x_0$  str. loc. min.



# Beweis des Satzes, Teil a).

Sei  $x^0$  ein lokales Minimum. Für  $v \neq 0$  und  $\varepsilon > 0$  hinreichend klein folgt aus der Taylor-Formel

$$f(x^0 + \varepsilon v) - f(x^0) = \underbrace{\frac{1}{2}(\varepsilon v)^T H f(x^0 + \theta \varepsilon v)(\varepsilon v)}_{R_1(x; x_0)} \geq 0 \quad (1)$$

*Handwritten notes:*  
-  $x = x_0 + \varepsilon v$   
-  $x - x_0 = \varepsilon v$   
-  $\text{grad } f(x_0) = 0$   
-  $\text{grad } f(x_0) \cdot \varepsilon v = 0$   
-  $\text{weil lokales Min.}$

mit  $\theta = \theta(\varepsilon, v) \in (0, 1)$ .

Der Gradient in der Taylorentwicklung verschwindet,  $\text{grad } f(x^0) = 0$ , denn  $x^0$  ist stationär.

Aus (1) folgt

$$v^T H f(x^0 + \theta \varepsilon v) v \geq 0 \quad (2)$$

Da  $f(x)$  eine  $\mathcal{C}^2$ -Funktion ist, ist die Hesse-Matrix eine stetige Abbildung. Im Grenzwert  $\varepsilon \rightarrow 0$  folgt daher aus (2),

$$v^T H f(x^0) v \geq 0 \quad \forall v \in \mathbb{R}^n$$

d.h.  $H f(x^0)$  ist positiv semidefinit.

## Beweis des Satzes, Teil b).

Ist  $Hf(x^0)$  positiv definit, so ist  $Hf(x)$  ebenfalls in einer hinreichend kleinen Umgebung  $x \in K_\varepsilon(x^0) \subset D$  um  $x^0$  positiv definit. Dies folgt aus der Stetigkeit der zweiten partiellen Ableitungen.

Für  $x \in K_\varepsilon(x^0)$ ,  $x \neq x^0$  gilt damit

$$\begin{aligned} f(x) - f(x^0) &= \frac{1}{2}(x - x^0)^T Hf(x^0 + \theta(x - x^0))(x - x^0) \\ &> 0 \end{aligned}$$

mit  $\theta \in (0, 1)$ , d.h.  $f(x)$  hat in  $x^0$  ein strenges lokales Minimum.

Ist  $Hf(x^0)$  indefinit, so existieren zu Eigenwerten von  $Hf(x^0)$  mit verschiedenen Vorzeichen gewisse Eigenvektoren  $v$   $w$  mit

$$v^T Hf(x^0)v > 0$$

$$w^T Hf(x^0)w < 0$$

und somit ist  $x^0$  ein Sattelpunkt.

# Bemerkungen.

- Ein stationärer Punkt  $x^0$  mit  $\det Hf(x^0) = 0$  heißt **ausgeartet**. Die Hesse-Matrix besitzt dann den **Eigenwert  $\lambda = 0$** .
- Ist  $x^0$  **nicht** ausgeartet, so gibt es 3 Fälle für die Eigenwerte von  $Hf(x^0)$ :
  - alle EW sind strikt positiv  $\Rightarrow x^0$  ist strenges lokales Minimum
  - alle EW sind strikt negativ  $\Rightarrow x^0$  ist strenges lokales Maximum
  - es gibt strikt pos. und neg. EW  $\Rightarrow x^0$  Sattelpunkt
- Die folgenden Implikationen gelten (**aber für keine die Umkehrung**)

$$\begin{array}{ccc} x^0 \text{ lokales Minimum} & \Leftarrow & x^0 \text{ strenges lokales Minimum} \\ \downarrow & & \uparrow \text{ i.A.} \\ Hf(x^0) \text{ positiv semidefinit} & \Leftarrow & Hf(x^0) \text{ positiv definit} \end{array}$$



## Weitere Bemerkung.

- Ist  $f(x)$  eine  $C^3$ -Funktion,  $x^0$  ein stationärer Punkt von  $f(x)$  und  $Hf(x^0)$  positiv definit, so gilt die Abschätzung:

$$(x - x^0)^T Hf(x^0) (x - x^0) \geq \lambda_{\min} \cdot \|x - x^0\|^2$$

wobei  $\lambda_{\min}$  den kleinsten Eigenwert der Hesse-Matrix bezeichnet.

Nach dem Satz von Taylor gilt dann:

$$\begin{aligned} f(x) - f(x^0) &\geq \frac{1}{2} \lambda_{\min} \|x - x^0\|^2 + R_3(x; x^0) \\ &\geq \|x - x^0\|^2 \left( \frac{\lambda_{\min}}{2} - C \|x - x^0\| \right) \end{aligned}$$

mit einer geeigneten Konstanten  $C > 0$ .

Um  $x^0$  wächst  $f(x)$  somit mindestens quadratisch mit dem Abstand von  $x^0$ .

# Beispiel.

Wir betrachten die Funktion

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

$$f(x, y) := y^2(x - 1) + x^2(x + 1)$$

und suchen die stationären Punkte:

$$\text{grad } f(x, y) = (y^2 + x(3x + 2), 2y(x - 1))^T$$

*z. ge.  
z. h. d.*

Die Bedingung  $\text{grad } f(x, y) = 0$  liefert die beiden stationären Punkte

$$x^0 = (0, 0)^T \quad \text{und} \quad x^1 = (-2/3, 0)^T.$$

Die jeweiligen Hesse-Matrizen von  $f$  an den Stellen  $x^0$  und  $x^1$  lauten

$$Hf(x^0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \quad \text{und} \quad Hf(x^1) = \begin{pmatrix} -2 & 0 \\ 0 & -10/3 \end{pmatrix}$$

Die Matrix  $Hf(x^0)$  ist indefinit, also ist  $x^0$  ein Sattelpunkt,  $Hf(x^1)$  ist negativ definit, somit ist  $x^1$  ein strenges lokales Maximum von  $f(x)$ .

# Kapitel 2. Anwendungen der Differentialrechnung mehrerer Variablen

## 2.2 Implizit definierte Funktionen

**Ziel:** Untersuche die Lösungsmengen von *nichtlinearen* Gleichungssystemen der Form

$$g(x) = 0$$

mit  $g : D \rightarrow \mathbb{R}^m$ ,  $D \subset \mathbb{R}^n$ , d.h. wir betrachten  $m$  Gleichungen für  $n$  Unbekannte mit

$$m < n.$$

**Also:** Es gibt *weniger* Gleichungen als Unbekannte.

Man nennt dann das Gleichungssystem *unterbestimmt* und die Lösungsmenge  $G \subset \mathbb{R}^n$  enthält gewöhnlich *unendlich* viele Punkte.

# Auflösbarkeit von (nichtlinearen) Gleichungen.

**Frage:** Kann man das System  $g(x) = 0$  nach bestimmten Unbekannten, zum Beispiel den letzten  $m$  Variablen  $x_{n-m+1}, \dots, x_n$  **auflösen**?

**Mit anderen Worten:** Existiert eine Funktion  $f(x_1, \dots, x_{n-m})$  mit

$$g(x) = 0 \iff (x_{n-m+1}, \dots, x_n)^T = f(x_1, \dots, x_{n-m})$$

**Terminologie:** "Auflösen" bedeutet also die letzten  $m$  Variablen durch die ersten  $n - m$  Variablen zu beschreiben.

**Weitere Frage:** Nach welchen  $m$  Variablen lässt sich das Gleichungssystem auflösen? Ist die Auflösung *global* auf dem Definitionsbereich  $D$  möglich oder nur *lokal* auf einer Teilmenge  $\tilde{D} \subset D$ ?

**Geometrische Interpretation:** Die Lösungsmenge  $G$  von  $g(x) = 0$  lässt sich (zumindest lokal) als Graph einer Funktion  $f : \mathbb{R}^{n-m} \rightarrow \mathbb{R}^m$  darstellen.

# Beispiel.

Die Kreisgleichung

$$g(x, y) = x^2 + y^2 - r^2 = 0 \quad \text{mit } r > 0$$

definiert ein **unterbestimmtes** nichtlineares Gleichungssystem, denn wir haben **zwei** Unbekannte  $(x, y)$ , aber nur **eine** Gleichung.

Die Kreisgleichung lässt sich **lokal** auflösen und definiert dabei die folgenden vier Funktionen:

$$y = \sqrt{r^2 - x^2}, \quad -r \leq x \leq r$$

$$y = -\sqrt{r^2 - x^2}, \quad -r \leq x \leq r$$

$$x = \sqrt{r^2 - y^2}, \quad -r \leq y \leq r$$

$$x = -\sqrt{r^2 - y^2}, \quad -r \leq y \leq r$$

## Beispiel.

Sei  $g(x)$  eine affin-lineare Funktion, d.h.  $g$  hat die Form

$$g(x) = Cx + b \quad \text{für } C \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$$

Wir spalten die Variablen  $x$  in zwei Vektoren auf

$$x^{(1)} = (x_1, \dots, x_{n-m})^T \in \mathbb{R}^{n-m} \quad \text{und} \quad x^{(2)} = (x_{n-m+1}, \dots, x_n)^T \in \mathbb{R}^m$$

Aufspaltung der Matrix  $C = [B, A]$  ergibt die Darstellung

$$g(x) = Bx^{(1)} + Ax^{(2)} + b$$

mit  $B \in \mathbb{R}^{m \times (n-m)}$ ,  $A \in \mathbb{R}^{m \times m}$ .

Das Gleichungssystem  $g(x) = 0$  ist genau dann nach den Variablen  $x^{(2)}$  (eindeutig) auflösbar, falls  $A$  regulär ist:

$$g(x) = 0 \quad \iff \quad x^{(2)} = -A^{-1}(Bx^{(1)} + b) = f(x^{(1)})$$

## Fortsetzung des Beispiels.

**Frage:** Wie kann man die Matrix  $A$  in Abhängigkeit von  $g$  schreiben?

Aus der Darstellung

$$g(x) = Bx^{(1)} + Ax^{(2)} + b$$

erkennt man direkt, dass

$$A = \frac{\partial g}{\partial x^{(2)}}(x^{(1)}, x^{(2)})$$

gilt, d.h.  $A$  ist die Jacobi-Matrix der Abbildung

$$x^{(2)} \rightarrow g(x^{(1)}, x^{(2)})$$

für festes  $x^{(1)}$ !

**Fazit:** Auflösbarkeit ist somit gegeben, falls die Jacobi-Matrix regulär ist.

# Satz über implizite Funktionen.

**Satz:** Sei  $g : D \rightarrow \mathbb{R}^m$  eine  $\mathcal{C}^1$ -Funktion,  $D \subset \mathbb{R}^n$  offen. Die Variablen in  $D$  seien  $(x, y)$  mit  $x \in \mathbb{R}^{n-m}$  und  $y \in \mathbb{R}^m$ . Der Punkt  $(x^0, y^0) \in D$  sei eine Lösung von  $g(x^0, y^0) = 0$ .

Falls die Jacobi-Matrix

$$\frac{\partial g}{\partial y}(x^0, y^0) := \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(x^0, y^0) & \dots & \frac{\partial g_1}{\partial y_m}(x^0, y^0) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial y_1}(x^0, y^0) & \dots & \frac{\partial g_m}{\partial y_m}(x^0, y^0) \end{pmatrix}$$

regulär ist, so gibt es Umgebungen  $U$  von  $x^0$  und  $V$  von  $y^0$ ,  $U \times V \subset D$  und eine eindeutig bestimmte stetig differenzierbare Funktion  $f : U \rightarrow V$  mit

$$f(x^0) = y^0 \quad \text{und} \quad g(x, f(x)) = 0 \quad \text{für alle } x \in U$$

und

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



## Beispiel.

Für die Kreisgleichung  $g(x, y) = x^2 + y^2 - r^2 = 0$ ,  $r > 0$  findet man im Punkt  $(x^0, y^0) = (0, r)$

$$\frac{\partial g}{\partial x}(0, r) = 0, \quad \frac{\partial g}{\partial y}(0, r) = 2r \neq 0$$

Man kann also in einer Umgebung von  $(0, r)$  die Kreisgleichung nach  $y$  auflösen:

$$f(x) = \sqrt{r^2 - x^2}$$

Die Ableitung  $f'(x)$  kann man durch **implizite Differentiation** berechnen:

$$g(x, y(x)) = 0 \quad \implies \quad g_x(x, y(x)) + g_y(x, y(x))y'(x) = 0$$

Also

$$2x + 2y(x)y'(x) = 0 \quad \implies \quad y'(x) = f'(x) = -\frac{x}{y(x)}$$

## Ein weiteres Beispiel.

Betrachte die Gleichung  $g(x, y) = e^{y-x} + 3y + x^2 - 1 = 0$ .

Es gilt

$$\frac{\partial g}{\partial y}(x, y) = e^{y-x} + 3 > 0 \quad \text{für alle } x \in \mathbb{R}.$$

Die Gleichung ist also für jedes  $x \in \mathbb{R}$  nach  $y =: f(x)$  auflösbar und  $f(x)$  ist eine stetig differenzierbare Funktion. Implizite Differentiation liefert

$$e^{y-x}(y' - 1) + 3y' + 2x = 0 \quad \implies \quad y' = \frac{e^{y-x} - 2x}{e^{y-x} + 3}$$

Erneute Differentiation liefert

$$e^{y-x}y'' + e^{y-x}(y' - 1)^2 + 3y'' + 2 = 0 \quad \implies \quad y' = -\frac{2 + e^{y-x}(y' - 1)^2}{e^{y-x} + 3}$$

**Aber:** Explizites Auflösen nach  $y$  (mit Hilfe elementarer Funktionen) ist in diesem Fall nicht möglich!

# Allgemeine Bemerkung.

Implizites Differenzieren einer durch

$$g(x, y) = 0, \quad \frac{\partial g}{\partial y} \neq 0$$

implizit definierten Funktion  $y = f(x)$ , mit  $x, y \in \mathbb{R}$ , ergibt

$$f'(x) = -\frac{g_x}{g_y}$$

$$f''(x) = -\frac{g_{xx}g_y^2 - 2g_{xy}g_xg_y + g_{yy}g_x^2}{g_y^3}$$

Daher ist der Punkt  $x^0$  ein **stationärer** Punkt von  $f(x)$ , falls gilt

$$g(x^0, y^0) = g_x(x^0, y^0) = 0 \quad \text{und} \quad g_y(x^0, y^0) \neq 0$$

Weiter ist  $x^0$  ein **lokales Maximum** (bzw. **Minimum**), falls

$$\frac{g_{xx}(x^0, y^0)}{g_y(x^0, y^0)} > 0 \quad \left( \text{bzw.} \quad \frac{g_{xx}(x^0, y^0)}{g_y(x^0, y^0)} < 0 \right)$$

# Implizite Darstellung ebener Kurven.

Betrachte die Lösungsmenge einer skalaren Gleichungen

$$g(x, y) = 0$$

Falls gilt

$$\text{grad } g = (g_x, g_y) \neq 0$$

so definiert  $g(x, y)$  lokal eine Funktion  $y = f(x)$  oder  $x = \bar{f}(y)$ .

**Definition:** Ein Lösungspunkt  $(x^0, y^0)$  der Gleichung  $g(x, y) = 0$  mit

- $\text{grad } g(x^0, y^0) \neq 0$  heißt **regulärer** Punkt,
- $\text{grad } g(x^0, y^0) = 0$  heißt **singulärer** Punkt.

**Beispiel:** Betrachte wieder die Kreisgleichung

$$g(x, y) = x^2 + y^2 - r = 0 \quad \text{mit } r > 0.$$

Auf der Kreislinie liegen **keine** singulären Punkte!

# Horizontale und vertikale Tangenten.

## Bemerkung:

a) Gilt für einen regulären Punkt  $(x^0, y^0)$

$$g_x(x^0) = 0 \quad \text{und} \quad g_y(x^0) \neq 0$$

so besitzt die Lösungskurve eine **horizontale Tangente** in  $x^0$ .

b) Gilt für einen regulären Punkt  $(x^0, y^0)$

$$g_x(x^0) \neq 0 \quad \text{und} \quad g_y(x^0) = 0$$

so besitzt die Lösungskurve eine **vertikale Tangente** in  $x^0$ .

c) Ist  $x^0$  ein **singulärer Punkt**, so wird die Lösungsmenge bei  $x^0$  "in zweiter Näherung" durch folgende **quadratische Gleichung** approximiert.

$$g_{xx}(x^0)(x - x^0)^2 + 2g_{xy}(x^0)(x - x^0)(y - y^0) + g_{yy}(x^0)(y - y^0)^2 = 0$$

# Bemerkungen.

Wegen c) erhält man für  $g_{xx}, g_{xy}, g_{yy} \neq 0$ :

$\det Hg(x^0) > 0$  :  $x^0$  ist ein **isolierter Punkt** der Lösungsmenge

$\det Hg(x^0) < 0$  :  $x^0$  ist ein **Doppelpunkt**

$\det Hg(x^0) = 0$  :  $x^0$  ist ein **Rückkehrpunkt** bzw. eine **Spitze**

## Geometrische Interpretation:

- Gilt  $\det Hg(x^0) > 0$ , so sind beide Eigenwerte von  $Hg(x^0)$  entweder strikt positiv oder strikt negativ, d.h.  $x^0$  ist ein strenges lokales **Minimum** oder **Maximum** von  $g(x)$ .
- Gilt  $\det Hg(x^0) < 0$ , so haben die beiden Eigenwerte von  $Hg(x^0)$  ein unterschiedliches Vorzeichen, d.h.  $x^0$  ist ein **Sattelpunkt** von  $g(x)$ .
- Gilt  $\det Hg(x^0) = 0$ , so ist der stationäre Punkt  $x^0$  von  $g(x)$  **ausgeartet**.

# Beispiel 1.

Betrachte den singulären Punkt  $x^0 = 0$  der impliziten Gleichung

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

Berechnung der partiellen Ableitungen bis zur Ordnung 2:

$$g_x = y^2 + 3x^2 - 4x$$

$$g_y = 2y(x - 1)$$

$$g_{xx} = 6x - 4$$

$$g_{xy} = 2y$$

$$g_{yy} = 2(x - 1)$$

$$Hg(0) = \begin{pmatrix} -4 & 0 \\ 0 & -2 \end{pmatrix}$$

Also ist  $x^0 = 0$  ein **isolierter Punkt**.

## Beispiel 2.

Betrachte den singulären Punkt  $x^0 = 0$  der impliziten Gleichung

$$g(x, y) = y^2(x - 1) + x^2(x + q^2) = 0$$

Berechnung der partiellen Ableitungen bis zur Ordnung 2:

$$g_x = y^2 + 3x^2 + 2xq^2$$

$$g_y = 2y(x - 1)$$

$$g_{xx} = 6x + 2q^2$$

$$g_{xy} = 2y$$

$$g_{yy} = 2(x - 1)$$

$$Hg(0) = \begin{pmatrix} 2q^2 & 0 \\ 0 & -2 \end{pmatrix}$$

Also ist  $x^0 = 0$  für  $q \neq 0$  ein **Doppelpunkt**.



## Beispiel 3.

Betrachte den singulären Punkt  $x^0 = 0$  der impliziten Gleichung

$$g(x, y) = y^2(x - 1) + x^3 = 0$$

Berechnung der partiellen Ableitungen bis zur Ordnung 2:

$$g_x = y^2 + 3x^2$$

$$g_y = 2y(x - 1)$$

$$g_{xx} = 6x$$

$$g_{xy} = 2y$$

$$g_{yy} = 2(x - 1)$$

$$Hg(0) = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}$$

Also ist  $x^0 = 0$  eine **Spitze** (bzw. ein **Rückkehrpunkt**).

# Implizite Darstellung von Flächen.

- Die Lösungsmenge einer skalaren Gleichung  $g(x, y, z) = 0$  ist für  $\text{grad } g \neq 0$  lokal eine **Fläche** im  $\mathbb{R}^3$ .
- Für die **Tangentialebene** in  $x^0 = (x^0, y^0, z^0)^T$  mit  $g(x^0) = 0$  und  $\text{grad } g(x^0) \neq 0^T$  bekommen wir für  $\Delta x^0 = x - x^0$  mit Taylor-Entwicklung

$$\text{grad } g \cdot \Delta x^0 = g_x(x^0)(x - x^0) + g_y(x^0)(y - y^0) + g_z(x^0)(z - z_0) = 0$$

d.h. der Gradient steht senkrecht auf der Fläche  $g(x, y, z) = 0$ .

- Ist zum Beispiel  $g_z(x^0) \neq 0$ , so gibt es lokal bei  $x^0$  eine Darstellung der Form

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

und für die **partielle Ableitungen** von  $f(x, y)$  bekommt man

$$\text{grad } f(x, y) = (f_x, f_y) = -\frac{1}{g_z}(g_x, g_y) = \left( -\frac{g_x}{g_z}, \frac{g_y}{g_z} \right)$$

mit dem Satz über implizite Funktionen.

# Das Umkehrproblem.

**Frage:** Lässt sich ein vorgegebenes Gleichungssystem

$$y = f(x)$$

mit  $f : D \rightarrow \mathbb{R}^n$ ,  $D \subset \mathbb{R}^n$  offen, nach  $x$  auflösen, also **invertieren**?

**Satz:** (**Umkehrsatz**)

Sei  $f : D \rightarrow \mathbb{R}^n$ ,  $D \subset \mathbb{R}^n$  offen, eine  $\mathcal{C}^1$ -Funktion. Ist für ein  $x^0 \in D$  die Jacobi-Matrix  $Jf(x^0)$  regulär, so gibt es Umgebungen  $U$  und  $V$  von  $x^0$  und  $y^0 = f(x^0)$ , so dass  $f$  den Bereich  $U$  **bijektiv** auf  $V$  abbildet.

Die Umkehrfunktion  $f^{-1} : V \rightarrow U$  ist ebenfalls eine  $\mathcal{C}^1$ -Funktion und es gilt für alle  $x \in U$ :

$$Jf^{-1}(y) = (Jf(x))^{-1}, \quad y = f(x)$$

**Bemerkung:** Man nennt dann  $f$  lokal einen  $\mathcal{C}^1$ -Diffeomorphismus.

# Kapitel 2. Anwendungen der Differentialrechnung mehrerer Variablen

## 2.3 Extremalprobleme unter Nebenbedingungen

**Frage:** Welche Abmessungen sollte eine Metalldose haben, damit bei vorgegebenem Volumen der Materialverbrauch am geringsten ist?

**Lösungsansatz:** Sei  $r > 0$  der Radius und  $h > 0$  die Höhe der Dose. Dann gilt

$$V = \pi r^2 h$$

$$O = 2\pi r^2 + 2\pi rh$$

Setze bei vorgegebenem Volumen  $c \in \mathbb{R}_+$ , und mit  $x := r, y := h$ ,

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

$$g(x, y) = \pi x^2 y - c = 0$$

Bestimme das Minimum der Funktion  $f(x, y)$  auf der Menge

$$G := \{(x, y) \in \mathbb{R}_+^2 \mid g(x, y) = 0\}$$

# Lösung des restringierten Minimierungsproblems.

Aus  $g(x, y) = \pi x^2 y - c = 0$  folgt

$$y = \frac{c}{\pi x^2}$$

Einsetzen in  $f(x, y)$  ergibt

$$h(x) := 2\pi x^2 + 2\pi x \frac{c}{\pi x^2} = 2\pi x^2 + \frac{2c}{x}$$

Bestimme das Minimum der Funktion  $h(x)$ :

$$h'(x) = 4\pi x - \frac{2c}{x^2} = 0 \quad \Rightarrow \quad 4\pi x = \frac{2c}{x^2} \quad \Rightarrow \quad x = \left(\frac{c}{2\pi}\right)^{1/3}$$

Hinreichende Bedingung

$$h''(x) = 4\pi + \frac{4c}{x^3} \quad \Rightarrow \quad h''\left(\left(\frac{c}{\pi}\right)^{1/3}\right) = 12\pi > 0$$

# Allgemeine Formulierung des Problems.

Bestimme die Extremwerte der Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  unter den Nebenbedingungen

$$g(x) = 0$$

wobei  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

Die Nebenbedingungen lauten also

$$g_1(x_1, \dots, x_n) = 0$$

$$\vdots$$

$$g_m(x_1, \dots, x_n) = 0$$

**Alternativ:** Bestimme die Extremwerte der Funktion  $f(x)$  auf der Menge

$$G := \{x \in \mathbb{R}^n \mid g(x) = 0\}$$

# Die Lagrange-Funktion und das Lagrange-Lemma.

Wir definieren die **Lagrange-Funktion**

$$F(x) := f(x) + \sum_{i=1}^m \lambda_i g_i(x)$$

und suchen die Extremwerte von  $F(x)$  für festes  $\lambda = (\lambda_1, \dots, \lambda_m)^T$ .

Die Zahlen  $\lambda_i$ ,  $i = 1, \dots, m$  nennt man **Lagrange-Multiplikatoren**.

**Satz:** (**Lagrange-Lemma**) Minimiert (bzw. maximiert)  $x^0$  die Lagrange-Funktion  $F(x)$  (für ein festes  $\lambda$ ) über  $D$  und gilt  $g(x^0) = 0$ , so liefert  $x^0$  das Minimum (bzw. Maximum) von  $f(x)$  über  $G := \{x \in D \mid g(x) = 0\}$ .

**Beweis:** Für ein beliebiges  $x \in D$  gilt nach Voraussetzung

$$f(x^0) + \lambda^T g(x^0) \leq f(x) + \lambda^T g(x)$$

Wählt man speziell  $x \in G$ , so ist  $g(x) = g(x^0) = 0$ , also auch  $f(x^0) \leq f(x)$ .

# Eine notwendige Bedingung für lokale Extrema.

Sind  $f$  und  $g_i$ ,  $i = 1, \dots, m$ ,  $C^1$ -Funktionen, so ist eine notwendige Bedingung für eine Extremstelle  $x^0$  von  $F(x)$  gegeben durch

$$\text{grad } F(x) = \text{grad } f(x) + \sum_{i=1}^m \lambda_i \text{grad } g_i(x) = 0$$

Zusammen mit den Nebenbedingungen  $g(x) = 0$  ergibt sich ein (nichtlineares) Gleichungssystem mit  $(n + m)$  Gleichungen und  $(n + m)$  Unbekannten  $x$  und  $\lambda$ . Die Lösungen  $(x^0, \lambda^0)$  sind die Kandidaten für die gesuchten Extremstellen, denn diese erfüllen die o.g. notwendige Bedingung.

**Alternativ:** Definiere eine Lagrange-Funktion

$$G(x, \lambda) := f(x) + \sum_{i=1}^m \lambda_i g_i(x)$$

und suche die Extremstellen von  $G(x, \lambda)$  bezüglich  $x$  **und**  $\lambda$ .



# Einige Bemerkungen zu hinreichenden Bedingungen.

- 1 Man kann auch eine **hinreichende** Bedingung aufstellen:  
Sind die Funktionen  $f$  und  $g$   $\mathcal{C}^2$ -Funktionen und ist die Hesse-Matrix  $HF(x^0)$  der Lagrange-Funktion positiv (bzw. negativ) definit, so ist  $x^0$  tatsächlich ein strenges lokales Minimum (bzw. Maximum) von  $f(x)$  auf  $G$ .
- 2 In den meisten Anwendungen ist die hinreichende Bedingung allerdings **nicht** erfüllt, obwohl  $x^0$  ein strenges lokales Extremum ist.
- 3 Insbesondere kann man aus der Indefinitheit der Hesse-Matrix  $HF(x^0)$  **nicht** schließen, dass  $x^0$  kein Extremwert ist.
- 4 Ähnlich problematisch ist die hinreichende Bedingung, die man aus der Hesse-Matrix für die Lagrange-Funktion  $G(x, \lambda)$  bezüglich  $x$  **und**  $\lambda$  erhält.

# Ein Beispiel zu restringierten Minimierungsproblems.

Gesucht seien die Extrema von  $f(x, y) := xy$  auf der Kreisscheibe

$$K := \{(x, y)^T \mid x^2 + y^2 \leq 1\}$$

Da die betrachte Funktion  $f$  stetig und  $K \subset \mathbb{R}^2$  kompakt ist, folgt aus der Min–Max–Eigenschaft die Existenz von globalen Maxima und Minima auf  $K$ .

Wir betrachten zunächst das Innere  $K^0$  von  $K$ , also die **offene** Menge

$$K^0 := \{(x, y)^T \mid x^2 + y^2 < 1\}$$

Die notwendige Bedingung für einen Extremwert lautet nun

$$\text{grad } f = (y, x) = 0$$

Somit ist der Ursprung  $x^0 = 0$  Kandidat für ein (lokales) Extremum.

## Fortsetzung des Beispiels.

Die Hesse-Matrix im Ursprung, gegeben durch

$$Hf(0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

und ist **indefinit**. Daher ist  $x^0$  ein **Sattelpunkt**.

Die Extrema der Funktion müssen also auf dem Rand liegen, der eine **Gleichungsnebenbedingung** darstellt:

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

Wir suchen also die Extremwerte von  $f(x, y) = xy$  unter der Nebenbedingung  $g(x, y) = 0$ .

Die Lagrange-Funktion lautet

$$F(x, y) = xy + \lambda(x^2 + y^2 - 1)$$

# Komplettierung des Beispiels.

Damit ergibt sich das (nichtlineare) Gleichungssystem

$$y + 2\lambda x = 0$$

$$x + 2\lambda y = 0$$

$$x^2 + y^2 = 1$$

mit den vier Lösungen

$$\lambda = \frac{1}{2} \quad : \quad x^{(1)} = (\sqrt{1/2}, -\sqrt{1/2})^T \quad x^{(2)} = (-\sqrt{1/2}, \sqrt{1/2})^T$$

$$\lambda = -\frac{1}{2} \quad : \quad x^{(3)} = (\sqrt{1/2}, \sqrt{1/2})^T \quad x^{(4)} = (-\sqrt{1/2}, -\sqrt{1/2})^T$$

**Minima** und **Maxima** lassen sich nun einfach aus den **Funktionswerten** ablesen

$$f(x^{(1)}) = f(x^{(2)}) = -1/2 \quad f(x^{(3)}) = f(x^{(4)}) = 1/2$$

d.h. Minima sind  $x^{(1)}$  und  $x^{(2)}$ , Maxima sind  $x^{(3)}$  und  $x^{(4)}$ .

# Lagrange–Multiplikatoren–Regel.

**Satz:** Seien  $f, g_1, \dots, g_m : D \rightarrow \mathbb{R}$  jeweils  $\mathcal{C}^1$ -Funktionen, und sei  $x^0 \in D$  ein lokales Extremum von  $f(x)$  unter der Nebenbedingung  $g(x) = 0$ . Weiterhin gelte die **Regularitätsbedingung**

$$\text{rang} \left( Jg(x^0) \right) = m$$

Dann existieren **Lagrange–Multiplikatoren**  $\lambda_1, \dots, \lambda_m$ , so dass für die **Lagrange Funktion**

$$F(x) := f(x) + \sum_{i=1}^m \lambda_i g_i(x)$$

die folgende **notwendige Bedingung erster Ordnung** gilt:

$$\text{grad } F(x^0) = 0$$

# Notwendige Bedingung zweiter Ordnung und hinreichende Bedingung.

**Satz:** 1) Ist  $x^0 \in D$  ein **lokales Minimum** von  $f(x)$  unter der Nebenbedingung  $g(x) = 0$ , ist die Regularitätsbedingung erfüllt und sind  $\lambda_1, \dots, \lambda_m$  zugehörige Lagrange-Multiplikatoren, so ist die Hesse-Matrix  $HF(x^0)$  der Lagrange-Funktion **positiv semidefinit** auf dem Tangentialraum

$$TG(x^0) := \{y \in \mathbb{R}^n \mid \text{grad } g_i(x^0) \cdot y = 0 \text{ für } i = 1, \dots, m\}$$

d.h., es gilt  $y^T HF(x^0) y \geq 0$  für alle  $y \in TG(x^0)$ .

2) Ist für einen Punkt  $x^0 \in G$  die Regularitätsbedingung erfüllt, existieren Lagrange-Multiplikatoren  $\lambda_1, \dots, \lambda_m$ , so dass  $x^0$  ein stationärer Punkt der zugehörigen Lagrange-Funktion ist, und ist die Hesse-Matrix  $HF(x^0)$  **positiv definit** auf dem Tangentialraum  $TG(x^0)$ , d.h., gilt

$$y^T HF(x^0) y > 0 \quad \forall y \in TG(x^0) \setminus \{0\},$$

so ist  $x^0$  ein **strenges lokales Minimum** von  $f(x)$  unter der Nebenbedingung  $g(x) = 0$ .

## Beispiel.

Man bestimme das globale Maximum der Funktion

$$f(x, y) = -x^2 + 8x - y^2 + 9$$

unter der Nebenbedingung

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

Die Lagrange-Funktion ist

$$F(x, y) = -x^2 + 8x - y^2 + 9 + \lambda(x^2 + y^2 - 1)$$

Aus der notwendigen Bedingung ergibt sich das nichtlineare System

$$-2x + 8 = -2\lambda x$$

$$-2y = -2\lambda y$$

$$x^2 + y^2 = 1$$

## Fortsetzung des Beispiels.

Aus der notwendigen Bedingung ergibt sich das nichtlineare System

$$-2x + 8 = -2\lambda x$$

$$-2y = -2\lambda y$$

$$x^2 + y^2 = 1$$

Aus der ersten Gleichung folgt  $\lambda \neq 1$ . Verwendet man dies in der zweiten Gleichung, so gilt  $y = 0$ . Aus der dritten Gleichung erkennt man sofort  $x = \pm 1$ . Demnach sind die beiden Punkte  $(x, y) = (1, 0)$  und  $(x, y) = (-1, 0)$  Kandidaten für das globale Maximum. Wegen

$$f(1, 0) = 16 \quad f(-1, 0) = 0$$

wird das globale Maximum von  $f(x, y)$  unter der Nebenbedingung  $g(x, y) = 0$  im Punkt  $(x, y) = (1, 0)$  angenommen.



## Ein weiteres Beispiel.

Man bestimme die lokalen Extremwerte der Funktion

$$f(x, y, z) = 2x + 3y + 2z$$

auf dem Durchschnitt des Zylindersmantels

$$M_Z := \{(x, y, z)^T \in \mathbb{R}^3 \mid x^2 + y^2 = 2\}$$

mit der Ebene

$$E := \{(x, y, z)^T \in \mathbb{R}^3 \mid x + z = 1\}$$

**Umformulierung:** Bestimme die Extremwerte der Funktion  $f(x, y, z)$  unter den Nebenbedingungen

$$g_1(x, y, z) := x^2 + y^2 - 2 = 0$$

$$g_2(x, y, z) := x + z - 1 = 0$$

## Fortsetzung des Beispiel.

Die Jacobi-Matrix

$$Jg(x) = \begin{pmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

hat den Rang 2, d.h. wir können über die Lagrange-Funktion Extremwerte bestimmen:

$$F(x, y, z) = 2x + 3y + 2z + \lambda_1(x^2 + y^2 - 2) + \lambda_2(x + z - 1)$$

Die notwendige Bedingung ergibt das nichtlineare Gleichungssystem

$$2 + 2\lambda_1 x + \lambda_2 = 0$$

$$3 + 2\lambda_1 y = 0$$

$$2 + \lambda_2 = 0$$

$$x^2 + y^2 = 2$$

$$x + z = 1$$

## Fortsetzung des Beispiel.

Die notwendige Bedingung ergibt das nichtlineare Gleichungssystem

$$2 + 2\lambda_1 x + \lambda_2 = 0$$

$$3 + 2\lambda_1 y = 0$$

$$2 + \lambda_2 = 0$$

$$x^2 + y^2 = 2$$

$$x + z = 1$$

Aus der ersten und dritten Gleichung folgt

$$2\lambda_1 x = 0$$

Aus der zweiten Gleichung folgt  $\lambda_1 \neq 0$ , also  $x = 0$ .  
Damit ergeben sich die möglichen Extremwerte als

$$(x, y, z) = (0, \sqrt{2}, 1) \quad (x, y, z) = (0, -\sqrt{2}, 1)$$

# Komplettierung des Beispiel.

Die möglichen Extremwerte sind also

$$(x, y, z) = (0, \sqrt{2}, 1) \quad (x, y, z) = (0, -\sqrt{2}, 1)$$

und liegen offensichtlich auf der Mantelfläche  $M_Z$  des Zylinders  $Z$  mit

$$Z = \{(x, y, z)^T \in \mathbb{R}^3 \mid x^2 + y^2 \leq 2\}$$

$$M_Z = \{(x, y, z)^T \in \mathbb{R}^3 \mid x^2 + y^2 = 2\}$$

Man berechnet nun die zugehörigen Funktionswerte

$$f(0, \sqrt{2}, 1) = 3\sqrt{2} + 2$$

$$f(0, -\sqrt{2}, 1) = -3\sqrt{2} + 2$$

Daher liegt im Punkt  $(x, y, z) = (0, \sqrt{2}, 1)$  ein Maximum und im Punkt  $(x, y, z) = (0, -\sqrt{2}, 1)$  ein Minimum vor.