

Analysis III für Studierende der Ingenieurwissenschaften

Jens Struckmeier

Fachbereich Mathematik
Universität Hamburg

Technische Universität Hamburg–Harburg
Wintersemester 2022/23

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

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

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

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

- 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 $(0,0)$:

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

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

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:

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

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 ξ_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 $\mathcal{C}^2$ -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):

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

- 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

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

Alternativ Notationen und weitere Rechenregeln.

$$\operatorname{rot} f(x) = \nabla \times f(x) = \begin{vmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ f_1 & f_2 & f_3 \end{vmatrix}$$

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 $\mathcal{C}^1$ -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

$$\begin{aligned} \frac{\partial f}{\partial x_i} &= \langle e_i, Ax \rangle + \langle x, Ae_i \rangle \\ &= e_i^T Ax + x^T Ae_i \\ &= x^T (A^T + A) e_i \end{aligned}$$

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 heit

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

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 fr 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 $v = e_i$ ist die Richtungsableitung in Richtung v gegeben durch die partielle Ableitung nach der Koordinatenrichtung x_i :

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

- 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 Φ^{-1} .

Beispiel: Polarkoordinaten.

Wir betrachten die Polarkoordinaten

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

Dann berechnet man

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

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

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

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 $\mathcal{C}^{m+1}$ -Funktion und sei $x_0 \in D$. Dann gilt für $x \in D$ die **Taylor-Entwicklung**

$$f(x) = T_m(x; x_0) + R_m(x; x_0)$$

$$T_m(x; x_0) = \sum_{|\alpha| \leq m} \frac{D^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha$$

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

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 Taylorschen Formel.

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

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

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

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

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

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

Fortsetzung der Herleitung.

Berechnung von $g''(0)$ liefert

$$\begin{aligned} g''(0) &= \left. \frac{d^2}{dt^2} f(x_0 + t(x - x_0)) \right|_{t=0} = \frac{d}{dt} \sum_{k=1}^n D_k f(x_0 + t(x - x_0)) (x_k - x_k^0) \Big|_{t=0} \\ &= D_{11} f(x_0) (x_1 - x_1^0)^2 + D_{21} f(x_0) (x_1 - x_1^0) (x_2 - x_2^0) \\ &\quad + \dots + D_{ij} f(x_0) (x_i - x_i^0) (x_j - x_j^0) + \dots + \\ &\quad + D_{n-1,n} f(x_0) (x_{n-1} - x_{n-1}^0) (x_n - x_n^0) + D_{nn} f(x_0) (x_n - x_n^0)^2 \\ &= \sum_{|\alpha|=2} \frac{D^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha \quad (\text{Vertauschungssatz von Schwarz!}) \end{aligned}$$

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.

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

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

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

- ❷ Die Berechnung von $T_2(x; x_0)$ benötigt die partiellen Ableitungen bis zur zweiten Ordnung.

- ❸ Diese Ableitungen müssen am Punkt $(x, y, z) = (1, 2, 0)^T$ ausgewertet werden.

- ❹ Als Ergebnis erhält man $T_2(x; x_0)$ in der Form

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

- ❺ Berechnung auf Folie.

Bemerkung zum Restglied eines Taylor-Polynoms.

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

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

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

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

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

$$f(x) = f(x^0) + \text{grad } f(x^0) \cdot (x - x^0) + O(\|x - x^0\|^2).$$

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 ∇f

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

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

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 $\mathcal{C}^1$ -Funktion $f(x)$ in einem Punkt $x^0 \in D^0$ ein lokales Extremum (Minimum oder Maximum), so gilt

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

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:

$$\varphi'(0) = \text{grad } f(x^0) v = 0$$

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

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

Bemerkungen zu lokalen Extremwerten.

Bemerkungen:

- Die Bedingung $\text{grad } f(x^0) = 0$ 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 $\text{grad } f(x^0) = 0$ nennt man **stationäre Punkte** von $f(x)$. Stationäre Punkte sind **nicht** notwendigerweise lokale Extremwerte. Zum Beispiel besitzt die Funktion

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

den Gradienten

$$\text{grad } f(x, y) = 2(x, -y)$$

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

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

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

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),

$$v^T H f(x^0) v \geq 0$$

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

$$v^T Hf(x^0) v > 0 \quad w^T Hf(x^0) w < 0$$

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**)

x^0 lokales Minimum $\Leftarrow x^0$ strenges lokales Minimum



$Hf(x^0)$ positiv semidefinit $\Leftarrow Hf(x^0)$ positiv definit

Weitere Bemerkung.

- Ist $f(x)$ eine $\mathcal{C}^3$ -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

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

und suchen die stationären Punkte:

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

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:

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

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 Diffentiation** berechnen:

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

Also

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

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

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

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:

- a) 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 $g(x)$.
- b) 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 $g(x)$.
- c) 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 λ_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 λ) 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 λ . Die Lösungen (x^0, λ^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** λ .

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** λ 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 Beispiels.

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 Beispiels.

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.

Kapitel 2. Anwendungen der Differentialrechnung mehrerer Variablen

2.4 Das Newton–Verfahren

Ziel: Wir suchen die Nullstellen einer Funktion $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R}^n$:

$$f(x) = 0$$

- Wir kennen bereits die **Fixpunktiteration**

$$x^{k+1} := \Phi(x^k)$$

mit Startwert x^0 und Iterationsvorschrift $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

- Konvergenzaussagen liefert der **Banachsche Fixpunktsatz**.

Vorteil: Dieses Verfahren ist **ableitungsfrei**.

Nachteile:

- das numerische Verfahren konvergiert zu langsam (nur linear),
- es gibt keine eindeutige Iterationsvorschrift.

Zur Konstruktion des Newton–Verfahrens.

Ausgangspunkt: Gegeben sei eine $\mathcal{C}^1$ –Funktion $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R}^n$ offen.

Wir suchen eine Nullstelle von f , d.h ein $x^* \in D$ mit

$$f(x^*) = 0$$

Konstruktion des Newton–Verfahrens:

Die Taylor–Entwicklung von $f(x)$ um einen Startwert x^0 lautet

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

Setzen wir $x = x^*$, so folgt

$$Jf(x^0)(x^* - x^0) \approx -f(x^0)$$

Eine Näherungslösung für x^* ist dann x^1 , $x^1 \approx x^*$, die Lösung des linearen Gleichungssystems

$$Jf(x^0)(x^1 - x^0) = -f(x^0)$$

Das Newton–Verfahrens als Algorithmus.

Das **Newton–Verfahren** kann man somit wie folgt als Algorithmus formulieren.

Algorithmus (**Newton–Verfahren**):

(1) FOR $k = 0, 1, 2, \dots$

(2a) Löse $Jf(x^k) \cdot \Delta x^k = -f(x^k);$

(2b) Setze $x^{k+1} = x^k + \Delta x^k;$

- Man löst in *jedem* Newton–Schritt ein lineares Gleichungssystem.
- Die Lösung Δx^k heißt **Newton–Korrektur**.
- Das Newton–Verfahren ist **skalierungsinvariant**.

Skalierungsinvarianz des Newton–Verfahrens.

Satz: Das Newton–Verfahren ist invariant unter linearen Transformationen der Form

$$f(x) \rightarrow g(x) = Af(x) \quad \text{für } A \in \mathbb{R}^{n \times n} \text{ regulär,}$$

d.h. die Iterierten für f und g sind in diesem Fall identisch.

Beweis: Bildet man das Newton–Verfahren für $g(x)$, so lautet die Newton–Korrektur

$$\begin{aligned} \Delta x^k &= -(Jg(x^k))^{-1} \cdot g(x^k) \\ &= -(AJf(x^k))^{-1} \cdot Af(x^k) \\ &= -(Jf(x^k))^{-1} \cdot A^{-1}A \cdot f(x^k) \\ &= -(Jf(x^k))^{-1} \cdot f(x^k) \end{aligned}$$

womit die Newton–Korrektur von f und g übereinstimmen.

Bei gleichem Startwert x^0 stimmen somit auch alle Iterierten x^k überein.

Zur lokalen Konvergenz des Newton–Verfahrens.

Satz: Sei $f : D \rightarrow \mathbb{R}^n$ eine $\mathcal{C}^1$ –Funktion, $D \subset \mathbb{R}^n$ offen und konvex. Sei $x^* \in D$ eine Nullstelle von f , d.h. $f(x^*) = 0$.

Weiterhin sei die Jacobi–Matrix $Jf(x)$ regulär für $x \in D$, und es gelte eine **Lipschitz–Bedingung**

$$\|(Jf(x))^{-1}(Jf(y) - Jf(x))\| \leq L\|y - x\| \quad \text{für alle } x, y \in D,$$

mit einem $L > 0$. Dann ist das Newton–Verfahren für alle Startwerte $x^0 \in D$ mit

$$\|x^0 - x^*\| < \frac{2}{L} =: r \quad \text{und} \quad K_r(x^*) \subset D$$

wohldefiniert mit $x^k \in K_r(x^*)$, $k = 0, 1, 2, \dots$, und die Newton–Iterierten x^k konvergieren **quadratisch** gegen x^* , d.h.

$$\|x^{k+1} - x^*\| \leq \frac{L}{2} \|x^k - x^*\|^2$$

Weiterhin ist x^* die eindeutige Nullstelle von $f(x)$ innerhalb der Kugel $K_r(x^*)$.

Das gedämpfte Newton–Verfahren.

Weitere Beobachtungen:

- Das Newton–Verfahren konvergiert zwar quadratisch, aber nur **lokal**.
- **Globale** Konvergenz kann ggf. durch einen Dämpfungsterm erreicht werden:

Algorithmus (**Gedämpftes Newton–Verfahren**):

(1) FOR $k = 0, 1, 2, \dots$

(2a) Löse $Jf(x^k) \cdot \Delta x^k = -f(x^k);$

(2b) Setze $x^{k+1} = x^k + \lambda_k \Delta x^k;$

Frage: Wie wählt man die **Dämpfungsfaktoren** λ_k ?

Wahl des Dämpfungsparameters.

Strategie: Verwende eine **Testfunktion** $T(x) = \|f(x)\|$, womit gilt

$$T(x) \geq 0, \quad \forall x \in D$$

$$T(x) = 0 \Leftrightarrow f(x) = 0$$

Wähle nun $\lambda_k \in (0, 1)$ so, dass die Folge $T(x^k)$ streng monoton fällt, d.h.

$$\|f(x^{k+1})\| < \|f(x^k)\| \quad \text{für } k \geq 0.$$

In der Nähe der gesuchten Lösung x^* sollte $\lambda_k = 1$ gewählt werden, um (lokale) quadratische Konvergenz zu sichern.

Der folgende Satz garantiert die Existenz eines Dämpfungsparameters.

Satz: Sei f eine $\mathcal{C}^1$ -Funktion auf der offenen und konvexen Menge $D \subset \mathbb{R}^n$. Für $x^k \in D$ mit $f(x^k) \neq 0$ gibt es dann ein $\mu_k > 0$, sodass

$$\|f(x^k + \lambda \Delta x^k)\|_2^2 < \|f(x^k)\|_2^2 \quad \text{für alle } \lambda \in (0, \mu_k).$$

Dämpfungsstrategie.

Für die **Startiteration** $k = 0$: Wähle $\lambda_0 \in \{1, \frac{1}{2}, \frac{1}{4}, \dots, \lambda_{min}\}$ **möglichst groß**, sodass gilt

$$\|f(x^0)\|_2 > \|f(x^0 + \lambda_0 \Delta x^0)\|_2$$

Für **nachfolgende Iterationen** $k > 0$: Setze $\lambda_k = \lambda_{k-1}$.

IF $\|f(x^k)\|_2 > \|f(x^k + \lambda_k \Delta x^k)\|_2$ **THEN**

- $x^{k+1} := x^k + \lambda_k \Delta x^k$
- $\lambda_k := 2\lambda_k$, falls $\lambda_k < 1$.

ELSE

- Bestimme $\mu = \max\{\lambda_k/2, \lambda_k/4, \dots, \lambda_{min}\}$ mit

$$\|f(x^k)\|_2 > \|f(x^k + \lambda_k \Delta x^k)\|_2$$

- $\lambda_k := \mu$

END

Kapitel 3. Integralrechnung mehrerer Variabler

3.1 Bereichsintegrale

Gegeben sei eine Funktion $f : D \rightarrow \mathbb{R}$ mit Definitionsbereich $D \subset \mathbb{R}^n$.

Ziel: Berechnung des Volumens unterhalb des Graphen von $f(x)$:

$$V = \int_D f(x) dx$$

Erinnerung Analysis II: Bestimmtes Riemann-Integral einer Funktion $f(x)$ über dem Intervall $[a, b]$:

$$I = \int_a^b f(x) dx$$

Das Integral I war als Grenzwert von Riemannscher Ober- und Untersumme definiert, falls diese Grenzwerte jeweils existierten und übereinstimmten.

Konstruktionsprinzip für Bereichsintegrale.

Vorgehensweise: Analog dem eindimensionalen Fall.

Aber: der Definitionsbereich D ist komplizierter.

Startpunkt: Betrachten zunächst den Fall zweier Variablen, $n = 2$, und einen Definitionsbereich $D \subset \mathbb{R}^2$ der Form

$$D = [a_1, b_1] \times [a_2, b_2] \subset \mathbb{R}^2$$

d.h. D ist ein kompakter Quader (Rechteck).

Weiterhin sei $f : D \rightarrow \mathbb{R}$ eine beschränkte Funktion.

Definition: Man nennt $Z = \{(x_0, x_1, \dots, x_n), (y_0, y_1, \dots, y_m)\}$ eine **Zerlegung** des Quaders $D = [a_1, b_1] \times [a_2, b_2]$, falls gilt

$$a_1 = x_0 < x_1 < \dots < x_n = b_1$$

$$a_2 = y_0 < y_1 < \dots < y_m = b_2$$

Mit $Z(D)$ wird die **Menge der Zerlegungen** von D bezeichnet.

Zerlegungen und Riemannsche Summen.

Definition:

- Die **Feinheit** einer Zerlegung $Z \in \mathcal{Z}(D)$ ist gegeben durch

$$\|Z\| := \max_{i,j} \{|x_{i+1} - x_i|, |y_{j+1} - y_j|\}$$

- Für eine vorgegebene Zerlegung Z nennt man die Mengen

$$Q_{ij} := [x_i, x_{i+1}] \times [y_j, y_{j+1}]$$

die **Teilquader** der Zerlegung Z . Das **Volumen** des Teilquaders Q_{ij} ist

$$\text{vol}(Q_{ij}) := (x_{i+1} - x_i) \cdot (y_{j+1} - y_j)$$

- Für beliebige Punkte $x_{ij} \in Q_{ij}$ der jeweiligen Teilquader nennt man

$$R_f(Z) := \sum_{i,j} f(x_{ij}) \cdot \text{vol}(Q_{ij})$$

eine **Riemannsche Summe** zur Zerlegung Z .

Riemannsche Ober– und Untersummen.

Definition:

Analog zum Integral einer Variablen heißen für eine Zerlegung Z

$$U_f(Z) := \sum_{i,j} \inf_{x \in Q_{ij}} f(x) \cdot \text{vol}(Q_{ij})$$

$$O_f(Z) := \sum_{i,j} \sup_{x \in Q_{ij}} f(x) \cdot \text{vol}(Q_{ij})$$

die **Riemannsche Untersumme** bzw. **Riemannsche Obersumme** von $f(x)$.

Bemerkung:

Eine Riemannsche Summe zur Zerlegung Z liegt stets zwischen der Unter– und Obersumme dieser Zerlegung, d.h. es gilt

$$U_f(Z) \leq R_f(Z) \leq O_f(Z)$$

Bemerkung.

Entsteht eine Zerlegung Z_2 aus der Zerlegung Z_1 durch Hinzunahme weiterer Zwischenpunkte x_i und/oder y_j , so gilt

$$U_f(Z_2) \geq U_f(Z_1) \quad \text{und} \quad O_f(Z_2) \leq O_f(Z_1)$$

Für zwei beliebige Zerlegungen Z_1 und Z_2 gilt stets:

$$U_f(Z_1) \leq O_f(Z_2)$$

Frage: Was passiert mit den Unter- und Obersummen im Grenzwert $\|Z\| \rightarrow 0$:

$$U_f := \sup\{U_f(Z) : Z \in \mathcal{Z}(D)\}$$

$$O_f := \inf\{O_f(Z) : Z \in \mathcal{Z}(D)\}$$

Beobachtung: Die beiden Werte U_f und O_f existieren, da Unter- und Obersumme monoton und beschränkt sind.

Riemannsche Ober– und Unterintegrale.

Definition:

- ① Das **Riemannsches Unter– bzw. Oberintegral** der Funktion $f(x)$ über D ist gegeben durch

$$\int_{\underline{D}} f(x) dx := \sup \{ U_f(Z) : Z \in Z(D) \}$$

$$\int_{\overline{D}} f(x) dx := \inf \{ O_f(Z) : Z \in Z(D) \}$$

- ② Die Funktion $f(x)$ nennt man **Riemann–integrierbar** über D , falls Unter– und Oberintegral übereinstimmen. Das **Riemann–Integral** von $f(x)$ über D ist dann gegeben durch

$$\int_D f(x) dx := \int_{\underline{D}} f(x) dx = \int_{\overline{D}} f(x) dx$$

Bemerkung.

Wir haben bis jetzt “nur” den Fall von **zwei** Variablen betrachtet:

$$f : D \rightarrow \mathbb{R}, \quad D \in \mathbb{R}^2$$

In höheren Dimensionen, $n > 2$, ist die Vorgehensweise analog.

Schreibweise: für $n = 2$ und $n = 3$

$$\int_D f(x, y) dx dy \quad \text{bzw.} \quad \int_D f(x, y, z) dx dy dz$$

oder auch

$$\iint_D f(x, y) dx dy \quad \text{bzw.} \quad \iiint_D f(x, y, z) dx dy dz$$

Elementare Eigenschaften des Integrals.

Satz:

a) Linearität

$$\int_D (\alpha f(x) + \beta g(x)) dx = \alpha \int_D f(x) dx + \beta \int_D g(x) dx$$

b) Monotonie

Gilt $f(x) \leq g(x)$ für alle $x \in D$, so folgt:

$$\int_D f(x) dx \leq \int_D g(x) dx$$

c) Positivität

Gilt für alle $x \in D$ die Beziehung $f(x) \geq 0$, d.h. $f(x)$ ist **nicht-negativ**, so folgt:

$$\int_D f(x) dx \geq 0$$

Weitere Eigenschaften des Integrals.

Satz:

- a) Sind D_1 , D_2 und D Quader, $D = D_1 \cup D_2$ und $\text{vol}(D_1 \cap D_2) = 0$, so ist $f(x)$ genau dann über D integrierbar, falls $f(x)$ über D_1 und D_2 integrierbar ist, und es gilt

$$\int_D f(x) dx = \int_{D_1} f(x) dx + \int_{D_2} f(x) dx$$

- b) Für das Integral gilt die folgende **Abschätzung**

$$\left| \int_D f(x) dx \right| \leq \sup_{x \in D} |f(x)| \cdot \text{vol}(D)$$

- c) **Riemannsches Kriterium**

$f(x)$ ist genau dann über D integrierbar, falls gilt:

$$\forall \varepsilon > 0 \quad \exists Z \in \mathcal{Z}(D) \quad : \quad O_f(Z) - U_f(Z) < \varepsilon$$

Der Satz von Fubini.

Satz: (Satz von Fubini) Ist $f : D \rightarrow \mathbb{R}$ integrierbar, $D = [a_1, b_1] \times [a_2, b_2]$ ein Quader, und existieren die Integrale

$$F(x) = \int_{a_2}^{b_2} f(x, y) dy \quad \text{und} \quad G(y) = \int_{a_1}^{b_1} f(x, y) dx$$

für alle $x \in [a_1, b_1]$ bzw. $y \in [a_2, b_2]$, so gelten

$$\int_D f(x) dx = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y) dy dx$$

$$\int_D f(x) dx = \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x, y) dx dy$$

Bedeutung:

Der Satz von Fubini erlaubt Reduktion auf eindimensionale Integrale.

Beispiel.

Gegeben sei der Quader $D = [0, 1] \times [0, 2]$ sowie die Funktion

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

Stetige Funktionen sind – wie wir gleich zeigen werden – über Quadern integrierbar. Daher können wir den Satz von Fubini anwenden:

$$\begin{aligned} \int_D f(x, y) dx &= \int_0^2 \int_0^1 f(x, y) dx dy = \int_0^2 \left[2x - \frac{x^2 y}{2} \right]_{x=0}^{x=1} dy \\ &= \int_0^2 \left(2 - \frac{y}{2} \right) dy = \left[2y - \frac{y^2}{4} \right]_{y=0}^{y=2} = 3 \end{aligned}$$

Bemerkung: Der Satz von Fubini verlangt als Voraussetzung die Integrierbarkeit von $f(x, y)$. Die Existenz der beiden Integrale $F(x)$ und $G(y)$ alleine garantiert die Integrierbarkeit von $f(x, y)$ **nicht**!

Die charakteristische Funktion.

Definition: Für $D \subset \mathbb{R}^n$ kompakt und $f : D \rightarrow \mathbb{R}$ beschränkt setzen wir

$$f^*(x) := \begin{cases} f(x) & : \text{ falls } x \in D \\ 0 & : \text{ falls } x \in \mathbb{R}^n \setminus D \end{cases}$$

Speziell für $f(x) = 1$ heißt $f^*(x)$ die **charakteristische Funktion** von D . Die charakteristische Funktion von D wird mit $\chi_D(x)$ bezeichnet.

Sei nun Q der kleinste Quader mit $D \subset Q$. Dann heißt die Funktion $f(x)$ **integrierbar** über D , falls $f^*(x)$ über Q integrierbar ist, und wir setzen

$$\int_D f(x) dx := \int_Q f^*(x) dx$$

Messbarkeit und Nullmengen.

Definition: Die kompakte Menge $D \subset \mathbb{R}^n$ heißt **messbar**, falls das Integral

$$\text{vol}(D) := \int_D 1 dx = \int_Q \chi_D(x) dx$$

existiert. Man nennt dann $\text{vol}(D)$ das **Volumen** von D im $\mathbb{R}^n$.

Die kompakte Menge D heißt **Nullmenge**, falls D messbar ist und $\text{vol}(D) = 0$ gilt.

Bemerkung:

- Ist die Menge D selbst ein Quader, so folgt $Q = D$, und somit gilt

$$\int_D f(x) dx = \int_Q f^*(x) dx = \int_Q f(x) dx$$

d.h. die eingeführten Integrationsbegriffe stimmen überein.

- Quader sind messbare Mengen.
- $\text{vol}(D)$ in diesem Fall das *tatsächliche* Volumen des Quaders im $\mathbb{R}^n$.

Drei wichtige Eigenschaften der Integration.

Bei der mehrdimensionalen Integration gelten die folgenden Aussagen.

Satz: Sei $D \subset \mathbb{R}^n$ kompakt. Dann ist D genau dann messbar, falls der Rand ∂D von D eine Nullmenge ist.

Satz: Sei $D \subset \mathbb{R}^n$ kompakt und messbar, und sei $f : D \rightarrow \mathbb{R}$ stetig. Dann ist $f(x)$ integrierbar über D .

Satz: (Mittelwertsatz) Ist $D \subset \mathbb{R}^n$ kompakt, zusammenhängend und messbar, und ist $f : D \rightarrow \mathbb{R}$ stetig, so gibt es einen Punkt $\xi \in D$ mit

$$\int_D f(x) dx = f(\xi) \cdot \text{vol}(D)$$

Normalbereiche.

Definition:

- Eine Teilmenge $D \subset \mathbb{R}^2$ heißt **Normalbereich**, falls es stetige Funktionen g, h bzw. $\tilde{g}, \tilde{h}$ gibt mit

$$D = \{(x, y) \mid a \leq x \leq b \text{ und } g(x) \leq y \leq h(x)\}$$

bzw.

$$D = \{(x, y) \mid \tilde{a} \leq y \leq \tilde{b} \text{ und } \tilde{g}(y) \leq x \leq \tilde{h}(y)\}$$

- Eine Teilmenge $D \subset \mathbb{R}^3$ heißt **Normalbereich**, falls es eine Darstellung

$$D = \{ (x_1, x_2, x_3) \mid a \leq x_i \leq b, \ g(x_i) \leq x_j \leq h(x_i) \\ \text{und } \varphi(x_i, x_j) \leq x_k \leq \psi(x_i, x_j) \}$$

gibt mit einer Permutation (i, j, k) von $(1, 2, 3)$ und stetigen Funktionen g, h, φ und ψ .

Projizierbare Mengen.

Definition: Eine Teilmenge $D \subset \mathbb{R}^n$ heißt **projizierbar** in Richtung x_i , $i \in \{1, \dots, n\}$, falls es eine messbare Menge $B \subset \mathbb{R}^{n-1}$ und stetige Funktionen φ, ψ gibt, so dass

$$D = \{ x \in \mathbb{R}^n \mid \tilde{x} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)^T \in B \\ \text{und } \varphi(\tilde{x}) \leq x_i \leq \psi(\tilde{x}) \}$$

Bemerkung:

- Projizierbare Mengen sind stets messbar. Damit sind auch alle Normalbereiche messbar, denn Normalbereiche sind projizierbar.
- Häufig läßt sich der Integrationsbereich D als Vereinigung endlich vieler Normalbereiche darstellen. Solche Bereich sind dann ebenfalls messbar.

Integration über Normalbereiche und projizierbare Mengen.

Satz: Ist $f(x)$ auf einem Normalbereich

$$D = \{ (x, y) \in \mathbb{R}^2 : a \leq x \leq b \text{ und } g(x) \leq y \leq h(x) \}$$

eine **stetige** Funktion, so gilt

$$\int_D f(x) dx = \int_a^b \int_{g(x)}^{h(x)} f(x, y) dy dx$$

Analoge Beziehungen gelten für höhere Dimensionen: Ist $D \subset \mathbb{R}^n$ eine **projizierbare Menge** in Richtung x_i , d.h. D besitzt eine Darstellung der Form

$$D = \{ x \in \mathbb{R}^n \mid \tilde{x} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)^T \in B \\ \text{und } \varphi(\tilde{x}) \leq x_i \leq \psi(\tilde{x}) \}$$

so gilt

$$\int_D f(x) dx = \int_B \left(\int_{\varphi(\tilde{x})}^{\psi(\tilde{x})} f(x) dx_i \right) d\tilde{x}$$

Beispiel.

Gegeben sei die Funktion

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

Berechne das Integral über der durch zwei Parabeln begrenzten Fläche

$$D := \{(x, y) \mid -1 \leq x \leq 1 \text{ und } x^2 \leq y \leq 2 - x^2\}$$

Die Menge D ist ein Normalbereich und $f(x, y)$ ist stetig, somit gilt

$$\begin{aligned} \int_D f(x, y) dx &= \int_{-1}^1 \left(\int_{x^2}^{2-x^2} (x + 2y) dy \right) dx = \int_{-1}^1 [xy + y^2]_{x^2}^{2-x^2} dx \\ &= \int_{-1}^1 (x(2 - x^2) + (2 - x^2)^2 - x^3 - x^4) dx \\ &= \int_{-1}^1 (-2x^3 - 4x^2 + 2x + 4) dx = \frac{16}{3} \end{aligned}$$

Beispiel.

Zu berechnen ist das Volumen des **Rotationsparaboloids**

$$V := \{(x, y, z)^T \mid x^2 + y^2 \leq 1 \text{ und } x^2 + y^2 \leq z \leq 1\}$$

Darstellung von V als **Normalbereich**

$$V = \{(x, y, z)^T \mid -1 \leq x \leq 1, -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2} \text{ und } x^2 + y^2 \leq z \leq 1\}$$

Damit gilt

$$\begin{aligned} \text{vol}(V) &= \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{x^2+y^2}^1 dz dy dx = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (1 - x^2 - y^2) dy dx \\ &= \int_{-1}^1 \left[(1 - x^2)y - \frac{y^3}{3} \right]_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} dx = \frac{4}{3} \int_{-1}^1 (1 - x^2)^{3/2} dx \\ &= \frac{1}{3} \left[x(\sqrt{1-x^2})^3 + \frac{3}{2}x\sqrt{1-x^2} + \frac{3}{2}\arcsin(x) \right]_{-1}^1 = \frac{\pi}{2} \end{aligned}$$

Integration über allgemeine Integrationsbereiche.

Definition: Sei $D \subset \mathbb{R}^n$ eine kompakte und messbare Menge. Man nennt $Z = \{D_1, \dots, D_m\}$ eine **allgemeine Zerlegung** von D , falls die Mengen D_k kompakt, messbar und zusammenhängend sind und falls gilt

$$\bigcup_{j=1}^m D_j = D \quad \text{und} \quad \forall i \neq j : D_i^0 \cap D_j^0 = \emptyset.$$

Weiterhin heißt

$$\text{diam}(D_j) := \sup \{ \|x - y\| \mid x, y \in D_j \}$$

der **Durchmesser** der Menge D_j und

$$\|Z\| := \max \{ \text{diam}(D_j) \mid j = 1, \dots, m \}$$

die **Feinheit** der allgemeinen Zerlegung Z .

Riemannsche Summe für allgemeine Zerlegungen.

Für eine stetige Funktion $f : D \rightarrow \mathbb{R}$ definiert man die **Riemannschen Summen**

$$R_f(Z) = \sum_{j=1}^m f(x^j) \operatorname{vol}(D_j)$$

mit beliebigen $x^j \in D_j$, $j = 1, \dots, m$.

Satz: Für jede Folge $(Z_k)_{k \in \mathbb{N}}$ allgemeiner Zerlegungen von D mit $\|Z_k\| \rightarrow 0$ (für $k \rightarrow \infty$) und für jede Folge zugehöriger Riemannscher Summen $R_f(Z_k)$ gilt

$$\lim_{k \rightarrow \infty} R_f(Z_k) = \int_D f(x) dx$$

Schwerpunkte von Flächen und Körpern.

Eine wichtige **Anwendung** der Bereichsintegrale ist die Berechnung der **Schwerpunkte** von Flächen und Körpern.

Definition: Sei $D \subset \mathbb{R}^2$ (bzw. $\mathbb{R}^3$) eine messbare Menge und $\rho(x)$, $x \in D$, eine vorgegebene Massendichte. Dann ist der **Schwerpunkt** der Fläche (bzw. des Körpers) D gegeben durch

$$x_s := \frac{\int_D \rho(x)x dx}{\int_D \rho(x) dx}$$

Das Zählerintegral (über eine vektorwertige Funktion) ist hierbei koordinatenweise zu berechnen.

Beispiel.

Zu berechnen ist der Schwerpunkt der Pyramide P

$$P := \left\{ (x, y, z)^T \mid \max(|y|, |z|) \leq \frac{ax}{2h}, \quad 0 \leq x \leq h \right\}$$

Berechne das Volumen von P unter Annahme konstanter Dichte wie folgt.

$$\begin{aligned} \text{vol}(P) &= \int_0^h \int_{-\frac{ax}{2h}}^{\frac{ax}{2h}} \int_{-\frac{ax}{2h}}^{\frac{ax}{2h}} dz \, dy \, dx \\ &= \int_0^h \int_{-\frac{ax}{2h}}^{\frac{ax}{2h}} \frac{ax}{h} dy \, dx \\ &= \int_0^h \left(\frac{ax}{h} \right)^2 dx = \frac{1}{3} a^2 h \end{aligned}$$

Fortsetzung des Beispiels.

Weiterhin gilt

$$\begin{aligned} \int_0^h \int_{-\frac{ax}{2h}}^{\frac{ax}{2h}} \int_{-\frac{ax}{2h}}^{\frac{ax}{2h}} \begin{pmatrix} x \\ y \\ z \end{pmatrix} dz dy dx &= \int_0^h \int_{-\frac{ax}{2h}}^{\frac{ax}{2h}} \begin{pmatrix} \frac{ax^2}{h} \\ \frac{axy}{h} \\ 0 \end{pmatrix} dy dx \\ &= \int_0^h \begin{pmatrix} \frac{a^2 x^3}{h^2} \\ 0 \\ 0 \end{pmatrix} dx \\ &= \begin{pmatrix} \frac{1}{4} a^2 h^2 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

Der Schwerpunkt von P liegt daher im Punkt $x_s = (\frac{3}{4}h, 0, 0)^T$.

Trägheitsmomente von Flächen und Körpern.

Eine weitere wichtige **Anwendung** der Bereichsintegrale ist die Berechnung der **Trägheitsmomente** von Flächen und Körpern.

Definition: (**Trägheitsmoment bezüglich einer Achse**)

Sei $D \subset \mathbb{R}^2$ (bzw. $\mathbb{R}^3$) eine messbare Menge, $\rho(x)$ bezeichne für $x \in D$ eine Massendichte und $r(x)$ den Abstand des Punktes $x \in D$ von einer vorgegebenen Drehachse.

Dann besitzt D bezüglich dieser Achse das Trägheitsmoment

$$\Theta := \int_D \rho(x) r^2(x) dx$$

Beispiel.

Wir berechnen das Trägheitsmoment des **homogenen Zylinders**

$$Z := \{ (x, y, z)^T : x^2 + y^2 \leq r^2, -l/2 \leq z \leq l/2 \}$$

bezüglich der x -Achse bei konstanter Dichte ρ wie folgt.

$$\begin{aligned}\Theta &= \int_Z \rho(y^2 + z^2) d(x, y, z) = \rho \int_Z (y^2 + z^2) d(x, y, z) \\&= \rho \int_{-r}^r \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \int_{-l/2}^{l/2} (y^2 + z^2) dz dy dx \\&= \rho \int_{-r}^r \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \left(ly^2 + \frac{l^3}{12} \right) dy dx \\&= \rho \frac{\pi l r^2}{12} (3r^2 + l^2)\end{aligned}$$

Der Transformationssatz.

Ziel: Verallgemeinerung der (eindimensionalen) [Substitutionsregel](#)

$$\int_{\varphi(a)}^{\varphi(b)} f(x) dx = \int_a^b f(\varphi(t)) \varphi'(t) dt$$

Satz: ([Transformationssatz](#)) Sei $\Phi : U \rightarrow \mathbb{R}^n$, $U \subset \mathbb{R}^n$ offen, eine $\mathcal{C}^1$ -Abbildung. $D \subset U$ sei eine kompakte, messbare Menge, so dass Φ auf D^0 einen $\mathcal{C}^1$ -Diffeomorphismus bildet. Dann ist auch $\Phi(D)$ kompakt und messbar, und für jede stetige Funktion $f : \Phi(D) \rightarrow \mathbb{R}$ gilt die [Transformationsformel](#)

$$\int_{\Phi(D)} f(x) dx = \int_D f(\Phi(u)) |\det J\Phi(u)| du$$

Bemerkung: Man beachte, dass im Transformationssatz die Bijektivität von Φ nur auf dem inneren Bereich D^0 von D gefordert wird – nicht jedoch auf dem Rand ∂D !

Beispiel.

Berechne den Schwerpunkt eines homogenen **Kugeloktanten**

$$V = \{(x, y, z)^T \mid x^2 + y^2 + z^2 \leq 1 \text{ und } x, y, z \geq 0\}$$

Es ist einfacher, den Schwerpunkt in **Kugelkoordinaten** zu berechnen:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \cos \varphi \cos \psi \\ r \sin \varphi \cos \psi \\ r \sin \psi \end{pmatrix} = \Phi(r, \varphi, \psi)$$

Die Transformation ist auf ganz $\mathbb{R}^3$ definiert und mit

$$D = [0, 1] \times \left[0, \frac{\pi}{2}\right] \times \left[0, \frac{\pi}{2}\right]$$

gilt $\Phi(D) = V$. Weiterhin ist Φ auf D^0 ein $\mathcal{C}^1$ -Diffeomorphismus mit

$$\det J\Phi(r, \varphi, \psi) = r^2 \cos \psi$$

Fortsetzung des Beispiels.

Nach dem Transformationssatz folgt

$$\text{vol}(V) = \int_V dx = \int_0^1 \int_0^{\pi/2} \int_0^{\pi/2} r^2 \cos \psi \, d\psi \, d\varphi \, dr = \frac{\pi}{6}$$

und

$$\begin{aligned} \text{vol}(V) \cdot x_s &= \int_V x \, dx = \int_0^1 \int_0^{\pi/2} \int_0^{\pi/2} (r \cos \varphi \cos \psi) r^2 \cos \psi \, d\psi \, d\varphi \, dr \\ &= \int_0^1 r^3 \, dr \cdot \int_0^{\pi/2} \cos \varphi \, d\varphi \cdot \int_0^{\pi/2} \cos^2 \psi \, d\psi = \frac{\pi}{16} \end{aligned}$$

Daraus folgt $x_s = \frac{3}{8}$.

Analog berechnet man $y_s = z_s = \frac{3}{8}$.

Der Steinersche Satz.

Satz: (Steinerscher Satz) Für das Trägheitsmoment eines homogenen Körpers K mit Gesamtmasse m gilt bezüglich einer vorgegebenen Drehachse A

$$\Theta_A = md^2 + \Theta_S$$

Hierbei ist S die zu A parallele Achse durch den Schwerpunkt x_s des Körpers K und d der Abstand des Schwerpunktes x_s von der Achse A .

Beweisidee: Setze $x := \Phi(u) = x_s + u$. Dann gilt mit dem Einheitsvektor a in Richtung der Achse A

$$\begin{aligned}\Theta_A &= \rho \int_K (\langle x, x \rangle - \langle x, a \rangle^2) dx \\ &= \rho \int_D (\langle x_s + u, x_s + u \rangle - \langle x_s + u, a \rangle^2) dx\end{aligned}$$

wobei

$$D := \{x - x_s \mid x \in K\}$$

Kapitel 3. Integralrechnung mehrerer Variabler

3.2 Kurvenintegrale

Für eine stückweise $\mathcal{C}^1$ -Kurve $c : [a, b] \rightarrow D$, $D \subset \mathbb{R}^n$, und eine stetige **skalare** Funktion $f : D \rightarrow \mathbb{R}$ hatten wir das **Kurvenintegral erster Art** definiert durch

$$\int_c f(x) ds := \int_a^b f(c(t)) \|\dot{c}(t)\| dt$$

wobei $\|\cdot\|$ die euklidische Norm bezeichnet.

Erweiterung: Kurvenintegrale über **vektorwertige** Funktionen, d.h.

$$\int_c f(x) dx := ?$$

Anwendung: Ein Massenpunkt bewegt sich entlang $c(t)$ in einem Kraftfeld $f(x)$.

Frage: Welche **physikalische** Arbeit muss entlang der Kurve geleistet werden?

Kurvenintegrale zweiter Art.

Definition: Für ein stetiges Vektorfeld $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R}^n$ offen, und eine stückweise $\mathcal{C}^1$ -Kurve $c : [a, b] \rightarrow D$ definieren wir das **Kurvenintegral zweiter Art** durch

$$\int_c f(x) dx := \int_a^b \langle f(c(t)), \dot{c}(t) \rangle dt$$

Herleitung: Approximiere die Kurve durch einen Streckenzug mit Ecken $c(t_i)$, wobei

$$Z = \{a = t_0 < t_1 < \cdots < t_m = b\}$$

eine Zerlegung des Intervalls $[a, b]$ ist.

Dann gilt für die in einem Kraftfeld $f(x)$ entlang der Kurve $c(t)$ geleistete Arbeit die Näherungsformel:

$$A \approx \sum_{i=0}^{m-1} \langle f(c(t_i)), c(t_{i+1}) - c(t_i) \rangle$$

Fortsetzung der Herleitung.

Daraus folgt:

$$\begin{aligned} A &\approx \sum_{j=1}^n \sum_{i=0}^{m-1} f_j(c(t_i))(c_j(t_{i+1}) - c_j(t_i)) \\ &= \sum_{j=1}^n \sum_{i=0}^{m-1} f_j(c(t_i)) \dot{c}_j(\tau_{ij})(t_{i+1} - t_i) \end{aligned}$$

Für eine Folge von Zerlegungen Z mit $\|Z\| \rightarrow 0$ konvergiert die linke Seite gegen das oben definierte **Kurvenintegral zweiter Art**.

Bemerkung: Für eine geschlossene Kurve $c(t)$, d.h. $c(a) = c(b)$, schreibt man das Kurvenintegral auch als

$$\oint_c f(x) dx$$

Eigenschaften des Kurvenintegrals zweiter Art.

- Linearität:

$$\int_c (\alpha f(x) + \beta g(x)) dx = \alpha \int_c f(x) dx + \beta \int_c g(x) dx$$

- Es gilt:

$$\int_{-c} f(x) dx = - \int_c f(x) dx,$$

wobei $(-c)(t) := c(b + a - t)$, $a \leq t \leq b$, den inversen Weg bezeichnet.

- Es gilt

$$\int_{c_1+c_2} f(x) dx = \int_{c_1} f(x) dx + \int_{c_2} f(x) dx$$

wobei $c_1 + c_2$ den aus c_1 und c_2 zusammengesetzten Weg bezeichnet, sodass der Endpunkt von c_1 der Anfangspunkt von c_2 ist.

Weitere Eigenschaften des Kurvenintegrals zweiter Art.

- Das Kurvenintegral zweiter Art ist **parametrisierungsinvariant**.
- Es gilt

$$\int_c f(x) dx = \int_a^b \langle f(c(t)), T(t) \rangle \|\dot{c}(t)\| dt = \int_c \langle f, T \rangle ds$$

mit dem **Tangenten-Einheitsvektor** $T(t) := \frac{\dot{c}(t)}{\|\dot{c}(t)\|}$.

- Formale Schreibweise:

$$\int_c f(x) dx = \int_c \sum_{i=1}^n f_i(x) dx_i = \sum_{i=1}^n \int_c f_i(x) dx_i$$

mit

$$\int_c f_i(x) dx_i := \int_a^b f_i(c(t)) \dot{c}_i(t) dt$$

Beispiel.

Für $x \in \mathbb{R}^3$ sei

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

$$c(t) := (\cos t, \sin t, at)^T \quad \text{mit } 0 \leq t \leq 2\pi$$

Dann berechnet man

$$\begin{aligned} \int_c f(x) dx &= \int_c (-y dx + x dy + z^2 dz) \\ &= \int_0^{2\pi} (-\sin t)(-\sin t) + \cos t \cos t + a^2 t^2 a) dt \\ &= \int_0^{2\pi} (1 + a^3 t^2) dt \\ &= 2\pi + \frac{a^3}{3}(2\pi)^3 \end{aligned}$$

Die Zirkulation eines Feldes längs einer Kurve.

Definition: Ist $u(x)$ ein Geschwindigkeitsfeld eines strömenden Mediums, so nennt man das Kurvenintegral $\oint_C u(x) dx$ entlang einer geschlossenen Kurve auch die **Zirkulation** des Feldes $u(x)$.

Beispiel: Für das Feld $u(x, y) = (y, 0)^T \in \mathbb{R}^2$ erhält man längs der Kurve $c(t) = (r \cos t, 1 + r \sin t)^T$, $0 \leq t \leq 2\pi$ die Zirkulation

$$\begin{aligned}\oint_C u(x) dx &= \int_0^{2\pi} (1 + r \sin t)(-r \sin t) dt \\&= \int_0^{2\pi} (-r \sin t - r^2 \sin^2 t) dt \\&= \left[r \cos t - \frac{r^2}{2}(t - \sin t \cos t) \right]_0^{2\pi} = -\pi r^2\end{aligned}$$

Wirbelfreie Vektorfelder.

Definition: Ein stetiges Vektorfeld $f(x)$, $x \in D \subset \mathbb{R}^n$, heißt **wirbelfrei**, falls dessen Kurvenintegral längs **aller** geschlossenen stückweise $\mathcal{C}^1$ -Kurven $c(t)$ in D verschwindet, d.h.

$$\oint_c f(x) dx = 0 \quad \text{für alle geschlossenen } c.$$

Bemerkung: Ein Vektorfeld ist genau dann wirbelfrei, wenn der Wert des Kurvenintegrals $\int_c f(x) dx$ nur vom Anfangs- und Endpunkt des Weges, jedoch nicht vom konkreten Verlauf der Kurve c abhängt. In diesem Fall nennt man das Kurvenintegral **wegunabhängig**.

Frage: Welche Kriterien für das Vektorfeld $f(x)$ **garantieren** die Wegunabhängigkeit des Kurvenintegrals?

Zusammenhängende Mengen.

Definition: Eine Teilmenge $D \subset \mathbb{R}^n$ heißt **zusammenhängend**, falls je zwei Punkte in D durch eine stückweise $\mathcal{C}^1$ -Kurve verbunden werden können:

$$\forall x^0, y^0 \in D : \exists c : [a, b] \rightarrow D \quad : \quad c(a) = x^0 \wedge c(b) = y^0$$

Eine offene und zusammenhängende Menge $D \subset \mathbb{R}^n$ nennt man auch ein **Gebiet** in $\mathbb{R}^n$.

Bemerkung: Eine **offene** Menge $D \subset \mathbb{R}^n$ ist genau dann **nicht** zusammenhängend, wenn es **disjunkte**, offene Mengen $U_1, U_2 \subset \mathbb{R}^n$ gibt mit

$$U_1 \cap D \neq \emptyset, \quad U_2 \cap D \neq \emptyset, \quad D \subset U_1 \cup U_2$$

Nicht zusammenhängende offene Mengen sind also – im Gegensatz zu zusammenhängenden Mengen – in (zumindest) zwei disjunkte offene Mengen trennbar.

Gradientenfelder, Stammfunktionen, Potentiale.

Definition: Sei $f : D \rightarrow \mathbb{R}^n$ ein Vektorfeld auf einem Gebiet $D \subset \mathbb{R}^n$. Das Vektorfeld nennt man ein **Gradientenfeld**, falls es eine skalare $\mathcal{C}^1$ -Funktion $\varphi : D \rightarrow \mathbb{R}$ gibt mit

$$f(x) = \nabla \varphi(x)$$

Die Funktion $\varphi(x)$ heit dann **Stammfunktion** oder **Potential** von $f(x)$, und das Vektorfeld $f(x)$ nennt man **konservativ**.

Bemerkung: Ein Massenpunkt bewege sich in einem **konservativen** Kraftfeld $K(x)$, d.h. K besitzt ein Potential $\varphi(x)$, sodass $K(x) = \nabla \varphi(x)$. Dann liefert die Funktion $U(x) = -\varphi(x)$ gerade die **potentielle Energie**:

$$K(x) = m\ddot{x} = -\nabla U(x)$$

Multipliziert man diese Beziehung mit $\dot{x}$, so folgt

$$m\langle \ddot{x}, \dot{x} \rangle + \langle \nabla U(x), \dot{x} \rangle = \frac{d}{dt} \left(\frac{1}{2} m \|\dot{x}\|^2 + U(x) \right) = 0$$

Hauptsatz für Kurvenintegrale.

Satz: (Hauptsatz für Kurvenintegrale)

Sei $D \subset \mathbb{R}^n$ ein Gebiet und $f(x)$ ein stetiges Vektorfeld auf D .

- 1) Besitzt $f(x)$ ein Potential $\varphi(x)$, so gilt für alle stückweisen $\mathcal{C}^1$ -Kurven $c : [a, b] \rightarrow D$:

$$\int_c f(x) dx = \varphi(c(b)) - \varphi(c(a))$$

Insbesondere ist das Kurvenintegral wegunabhängig und $f(x)$ ist wirbelfrei.

- 2) Umgekehrt gilt: Ist $f(x)$ wirbelfrei, so besitzt $f(x)$ ein Potential $\varphi(x)$.
Ist $x^0 \in D$ ein fester Punkt, und bezeichnet c_x (für $x \in D$) eine beliebige, die Punkte x^0 und x verbindende stückweise $\mathcal{C}^1$ -Kurve in D , so ist $\varphi(x)$ gegeben durch:

$$\varphi(x) = \int_{c_x} f(x) dx + \text{const.}$$

Beispiel I.

Das zentrale Kraftfeld

$$K(x) := \frac{x}{\|x\|^3}$$

besitzt das Potential

$$U(x) = -\frac{1}{\|x\|} = -(x_1^2 + x_2^2 + x_3^2)^{-1/2}$$

denn es gilt

$$\nabla U(x) = (x_1^2 + x_2^2 + x_3^2)^{-3/2} (x, y, z)^T = \frac{x}{\|x\|^3}$$

Für die längs einer stückweisen $\mathcal{C}^1$ -Kurve $c : [a, b] \rightarrow \mathbb{R}^3 \setminus \{0\}$ geleistete Arbeit gilt dann

$$A = \int_c K(x) dx = \left(\frac{1}{\|c(a)\|} - \frac{1}{\|c(b)\|} \right)$$

Beispiel II.

Das Vektorfeld

$$f(x) := \begin{pmatrix} 2xy + z^3 \\ x^2 + 3z \\ 3xz^2 + 3y \end{pmatrix}$$

besitzt das Potential

$$\varphi(x) = x^2y + xz^3 + 3yz$$

Für eine beliebige $\mathcal{C}^1$ -Kurve $c(t)$ von $P = (1, 1, 2)$ nach $Q = (3, 5, -2)$ gilt

$$\int_c f(x) dx = \varphi(Q) - \varphi(P) = -9 - 15 = -24$$

Interpretiert man $f(x)$ als elektrisches Feld, so gibt das Kurvenintegral zweiter Art die **Spannung** zwischen den beiden Punkten P und Q an.

Beispiel III.

Wir betrachten das Vektorfeld

$$f(x, y) = \frac{1}{x^2 + y^2} \begin{pmatrix} -y \\ x \end{pmatrix} \quad \text{mit } (x, y)^T \in D = \mathbb{R}^2 \setminus \{0\}$$

Für den Einheitskreis $c(t) := (\cos t, \sin t)^T$, $0 \leq t \leq 2\pi$, bekommt man

$$\begin{aligned} \int_c f(x) dx &= \int_0^{2\pi} \langle f(c(t), \dot{c}(t)) \rangle dt \\ &= \int_0^{2\pi} \left\langle \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \right\rangle dt \\ &= \int_0^{2\pi} 1 dt = 2\pi \end{aligned}$$

$f(x, y)$ ist somit nicht wirbelfrei und besitzt auf D kein Potential.

Bedingungen für Potentiale.

Bemerkung: Ist $f(x)$, $x \in D \subset \mathbb{R}^3$, ein $\mathcal{C}^1$ -Vektorfeld mit Potential $\varphi(x)$, so folgt

$$\operatorname{rot} f(x) = \operatorname{rot} (\nabla \varphi(x)) = 0 \quad \text{für alle } x \in D$$

Somit ist $\operatorname{rot} f(x) = 0$ eine **notwendige Bedingung** für die Existenz eines Potentials ist.

Definiert man für ein Vektorfeld $f : D \rightarrow \mathbb{R}^2$, $D \subset \mathbb{R}^2$, die **skalare** Rotation

$$\operatorname{rot} f(x, y) := \frac{\partial f_2}{\partial x}(x, y) - \frac{\partial f_1}{\partial y}(x, y)$$

so ist $\operatorname{rot} f(x, y) = 0$ auch in zwei Dimensionen eine **notwendige Bedingung**.

Die Bedingung

$$\operatorname{rot} f(x) = 0$$

ist eine **hinreichende Bedingung**, falls das Gebiet D **einfach zusammenhängend** ist, d.h. keine "Löcher" enthält.

Beispiel.

Wir betrachten erneut das Vektorfeld

$$f(x, y) = \frac{1}{x^2 + y^2} \begin{pmatrix} -y \\ x \end{pmatrix} \quad \text{mit } (x, y)^T \in D = \mathbb{R}^2 \setminus \{0\}$$

Berechnet man die Rotation, so ergibt sich

$$\begin{aligned} \text{rot} \left[\frac{1}{r^2} \begin{pmatrix} -y \\ x \end{pmatrix} \right] &= \frac{\partial}{\partial x} \left(\frac{x}{x^2 + y^2} \right) + \frac{\partial}{\partial x} \left(\frac{y}{x^2 + y^2} \right) \\ &= \frac{1}{x^2 + y^2} - \frac{2x^2}{(x^2 + y^2)^2} + \frac{1}{x^2 + y^2} - \frac{2y^2}{(x^2 + y^2)^2} \\ &= 0 \end{aligned}$$

Die Rotation von $f(x, y)$ verschwindet.

Allerdings besitzt $f(x, y)$ auf der Menge $D = \mathbb{R}^2 \setminus \{0\}$ kein Potential.

Das Gebiet ist nämlich **nicht** einfach zusammenhängend.

Der Integralsatz von Green für Vektorfelder im $\mathbb{R}^2$.

Satz: (Integralsatz von Green)

Sei $f(x)$ ein $\mathcal{C}^1$ -Vektorfeld auf dem Gebiet $D \subset \mathbb{R}^2$. Weiterhin sei $K \subset D$ kompakt und bezüglich beider Koordinatenrichtungen projizierbar, sodass K von einer geschlossenen, stückweisen $\mathcal{C}^1$ -Kurve $c(t)$ berandet wird.

Die Parametrisierung von $c(t)$ sei so gewählt, dass K stets links zur Durchlaufrichtung liegt (positiver Umlauf). Dann gilt:

$$\oint_c f(x) \, dx = \int_K \operatorname{rot} f(x) \, dx$$

Bemerkung:

Der Greensche Integralsatz gilt auch für kompakte Bereiche, die sich in *endlich* viele, bezüglich beider Koordinatenrichtungen projizierbarer Bereiche zerlegen lassen, in so genannte **Greensche Bereiche**.

Alternative Formulierung des Greenschen Satzes I.

Wir hatten gesehen, dass die Beziehung

$$\oint_C f(x) dx = \oint_C \langle f, T \rangle ds$$

gilt, wobei $T(t) = \frac{\dot{c}(t)}{\|\dot{c}(t)\|}$ den Tangenteneinheitsvektor bezeichnet.

Daraus folgt mit dem Integralsatz von Green

$$\int_K \operatorname{rot} f(x) dx = \oint_{\partial K} \langle f, T \rangle ds$$

Ist $f(x)$ ein Geschwindigkeitsfeld, so ist die durch f beschriebene Strömung unter der Bedingung $\operatorname{rot} f(x) = 0$ wirbelfrei, denn

$$\oint_C f(x) dx$$

ist gerade die Zirkulation von $f(x)$.

Alternative Formulierung des Greenschen Satzes II.

Ersetzt man in der obigen Gleichungen den Vektor T durch den äußeren Normaleneinheitsvektor $n = (T_2, -T_1)^T$, so folgt

$$\begin{aligned}\oint_{\partial K} \langle f, n \rangle ds &= \oint_{\partial K} (f_1 T_2 - f_2 T_1) ds = \oint_{\partial K} \left\langle \begin{pmatrix} -f_2 \\ f_1 \end{pmatrix}, T \right\rangle ds \\ &= \int_K \operatorname{rot} \begin{pmatrix} -f_2 \\ f_1 \end{pmatrix} dx = \int_K \operatorname{div} f dx\end{aligned}$$

und damit die Beziehung

$$\int_K \operatorname{div} f(x) dx = \oint_{\partial K} \langle f, n \rangle ds$$

Ist $f(x)$ das Geschwindigkeitsfeld einer Strömung, so beschreibt die rechte Seite den **Gesamtfluss** der Strömung durch den Rand von K . Gilt also $\operatorname{div} f(x) = 0$, so ist die Strömung **quellen- und senkenfrei** (oder **divergenzfrei**).

Nochmal zurück zur Existenz von Potentialen.

Folgerung: Ist $\operatorname{rot} f(x) = 0$ für alle $x \in D$, $D \subset \mathbb{R}^2$ ein Gebiet, so folgt

$$\oint_c f(x) dx = 0$$

für jede geschlossene stückweise C^1 -Kurve, die einen Greenschen Bereich $B \subset D$ vollständig umrandet.

Definition: Ein Gebiet $D \subset \mathbb{R}^n$ heißt **einfach zusammenhängend**, falls sich jede geschlossene Kurve $c : [a, b] \rightarrow D$ stetig innerhalb von D auf einen Punkt in D zusammenziehen lässt. Genauer: es gibt für $x^0 \in D$ eine stetige Abbildung

$$\Phi : [a, b] \times [0, 1] \rightarrow D$$

mit $\Phi(t, 0) = c(t)$, für alle $t \in [a, b]$ und $\Phi(t, 1) = x^0 \in D$, für alle $t \in [a, b]$. Die Abbildung $\Phi(t, s)$ nennt man eine **Homotopie**.

Integrabilitätsbedingung für Potentiale.

Satz: Sei $D \subset \mathbb{R}^n$ ein einfach zusammenhängendes Gebiet. Ein C^1 -Vektorfeld $f : D \rightarrow \mathbb{R}^n$ besitzt genau dann ein Potential auf D , falls die Integrabilitätsbedingung

$$Jf(x) = (Jf(x))^T \quad \text{für alle } x \in D$$

erfüllt ist, d.h. falls gilt

$$\frac{\partial f_k}{\partial x_j} = \frac{\partial f_j}{\partial x_k} \quad \forall j, k$$

Bemerkung: Für $n = 2, 3$ stimmt die Integrabilitätsbedingung mit

$$\operatorname{rot} f(x) = 0$$

überein.

Beispiel.

Für $x \in \mathbb{R}^3 \setminus \{0\}$ sei das Vektorfeld

$$f(x) = \begin{pmatrix} \frac{2xy}{r^2} + \sin z \\ \ln r^2 + \frac{2y^2}{r^2} + ze^y \\ \frac{2yz}{r^2} + e^y + x \cos z \end{pmatrix} \quad \text{mit } r^2 = x^2 + y^2 + z^2$$

gegeben.

Wir wollen untersuchen, ob $f(x)$ ein Potential besitzt.

Die Menge $D = \mathbb{R}^3 \setminus \{0\}$ ist offensichtlich **einfach zusammenhängend**.
Weiterhin gilt

$$\operatorname{rot} f(x) = 0$$

Also besitzt $f(x)$ ein Potential.

Berechnung des Potentials.

Es muss gelten: $f(x) = \nabla\varphi(x)$. Demnach folgt:

$$\frac{\partial\varphi}{\partial x} = f_1(x, y, z) = \frac{2xy}{r^2} + \sin z$$

Durch Integration bezüglich der Variablen x ergibt sich:

$$\varphi(x) = y \ln r^2 + x \sin z + c(y, z)$$

mit einer unbekannten Funktion $c(y, z)$.

Einsetzen in die Gleichung

$$\frac{\partial\varphi}{\partial y} = f_2(x, y, z) = \ln r^2 + \frac{2y^2}{r^2} + ze^y$$

liefert

$$\ln r^2 + \frac{2y^2}{r^2} + \frac{\partial c}{\partial y} = \ln r^2 + \frac{2y^2}{r^2} + ze^y$$

Berechnung des Potentials (Fortsetzung).

Daraus folgt die Bedingung

$$\frac{\partial c}{\partial y} = ze^y$$

und somit gilt

$$c(y, z) = ze^y + d(z)$$

für eine unbekannte Funktion $d(z)$. Wir haben damit:

$$\varphi(x) = y \ln r^2 + x \sin z + ze^y + d(z)$$

Die letzte Bedingung lautet

$$\frac{\partial \varphi}{\partial z} = f_3(x, y, z) = \frac{2yz}{r^2} + e^y + x \cos z$$

Daraus folgt $d'(z) = 0$ und das Potential ist gegeben durch

$$\varphi(x) = y \ln r^2 + x \sin z + ze^y + c \quad \text{für } c \in \mathbb{R}$$

Kapitel 3. Integralrechnung mehrerer Variabler

3.3 Oberflächenintegrale

Definition: Sei $D \subset \mathbb{R}^2$ ein Gebiet und $p : D \rightarrow \mathbb{R}^3$ eine $\mathcal{C}^1$ -Abbildung

$$x = p(u) \quad \text{mit } x \in \mathbb{R}^3 \text{ und } u = (u_1, u_2)^T \in D \subset \mathbb{R}^2$$

Sind für alle $u \in D$ die beiden Vektoren

$$\frac{\partial p}{\partial u_1} \quad \text{und} \quad \frac{\partial p}{\partial u_2}$$

linear unabhängig, so heißt

$$F := \{p(u) \mid u \in D\}$$

eine **Fläche** bzw. ein **Flächenstück**. Die Abbildung $x = p(u)$ nennt man dann eine **Parametrisierung** oder **Parameterdarstellung** der Fläche F .

Beispiel I.

Wir betrachten für gegebenes $r > 0$ die Abbildung

$$p(\varphi, z) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{pmatrix} \quad \text{für } (\varphi, z) \in \mathbb{R}^2.$$

Die dadurch parametrisierte Fläche ist ein **unbeschränkter Zylinder** im $\mathbb{R}^3$.

Schränken wir den Definitionsbereich ein, etwa

$$(\varphi, z) \in K := [0, 2\pi] \times [0, H] \subset \mathbb{R}^2$$

so erhalten wir einen **beschränkten Zylinder** der Höhe H .

Die partiellen Ableitungen

$$\frac{\partial p}{\partial \varphi} = \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \\ 0 \end{pmatrix}, \quad \frac{\partial p}{\partial z} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

von $p(\varphi, z)$ sind linear unabhängig auf ganz $\mathbb{R}^2$.

Beispiel II.

Der Graph einer skalaren $\mathcal{C}^1$ -Funktion $\varphi : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^2$, ist eine **Fläche**.

Eine **Parametrisierung** ist gegeben durch

$$p(u_1, u_2) := \begin{pmatrix} u_1 \\ u_2 \\ \varphi(u_1, u_2) \end{pmatrix} \quad \text{für } u \in D$$

Die partiellen Ableitungen

$$\frac{\partial p}{\partial u_1} = \begin{pmatrix} 1 \\ 0 \\ \varphi_{u_1} \end{pmatrix}, \quad \frac{\partial p}{\partial u_2} = \begin{pmatrix} 0 \\ 1 \\ \varphi_{u_2} \end{pmatrix}$$

sind **linear unabhängig**.

Die Tangentialebene einer Fläche.

Die beiden linear unabhängigen Vektoren

$$\frac{\partial \mathbf{p}}{\partial u_1}(\mathbf{u}^0) \quad \text{und} \quad \frac{\partial \mathbf{p}}{\partial u_2}(\mathbf{u}^0)$$

liegen **tangential** an die Fläche F .

Sie spannen die **Tangentialebene** $T_{\mathbf{x}^0}F$ der Fläche F im Punkt $\mathbf{x}^0 = \mathbf{p}(\mathbf{u})$ auf.

Die Tangentialebene hat die Parameterdarstellung

$$T_{\mathbf{x}^0}F : \mathbf{x} = \mathbf{x}^0 + \lambda \frac{\partial \mathbf{p}}{\partial u_1}(\mathbf{u}^0) + \mu \frac{\partial \mathbf{p}}{\partial u_2}(\mathbf{u}^0) \quad \text{für } \lambda, \mu \in \mathbb{R}.$$

Frage: Wie kann man den Flächeninhalt einer gegebenen Fläche F berechnen?

Das Oberflächenintegral eines Flächenstückes.

Definition: Sei $p : D \rightarrow \mathbb{R}^3$ die Parameterdarstellung einer Fläche, und sei $K \subset D$ kompakt, messbar und zusammenhängend. Dann wird der Flächeninhalt von $p(K)$ definiert durch das **Oberflächenintegral**

$$\int_{p(K)} do := \int_K \left\| \frac{\partial p}{\partial u_1}(u) \times \frac{\partial p}{\partial u_2}(u) \right\| du$$

Dabei nennt man den Term

$$do := \left\| \frac{\partial p}{\partial u_1}(u) \times \frac{\partial p}{\partial u_2}(u) \right\| du$$

das **Oberflächenelement** der Fläche $x = p(u)$.

Bemerkung: Das Oberflächenintegral ist insbesondere **unabhängig** von der speziellen Parametrisierung der Fläche. Dies folgt aus dem Transformationssatz.

Beispiel.

Für die Mantelfläche des Zylinders $Z = p(K)$ mit

$$K := [0, 2\pi] \times [0, H] \subset \mathbb{R}^2$$

und

$$x = p(\varphi, z) := \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{pmatrix} \quad \text{für } (\varphi, z) \in \mathbb{R}^2$$

erhält man mit

$$\left\| \frac{\partial p}{\partial \varphi} \times \frac{\partial p}{\partial z} \right\| = r$$

den Wert

$$O(Z) = \int_Z do = \int_K rd(\varphi, z) = \int_0^{2\pi} \int_0^H rdz d\varphi = 2\pi rH$$

Beispiel.

Ist die Fläche der Graph einer skalaren Funktion, d.h. $x_3 = \varphi(x_1, x_2)$, so gilt für die zugehörigen Tangentialvektoren

$$\frac{\partial \mathbf{p}}{\partial x_1} \times \frac{\partial \mathbf{p}}{\partial x_2} = \begin{pmatrix} 1 \\ 0 \\ \varphi_{x_1} \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ \varphi_{x_2} \end{pmatrix} = \begin{pmatrix} -\varphi_{x_1} \\ -\varphi_{x_2} \\ 1 \end{pmatrix}$$

Damit ergibt sich

$$\left\| \frac{\partial \mathbf{p}}{\partial x_1} \times \frac{\partial \mathbf{p}}{\partial x_2} \right\| = \sqrt{1 + \varphi_{x_1}^2 + \varphi_{x_2}^2}$$

und

$$\begin{aligned} O(p(K)) &= \int_{p(K)} do \\ &= \int_K \sqrt{1 + \varphi_{x_1}^2 + \varphi_{x_2}^2} d(x_1, x_2) \end{aligned}$$

Beispiel.

Für die Oberfläche des Paraboloids P , gegeben durch

$$P := \{(x_1, x_2, x_3)^T \in \mathbb{R}^3 \mid x_3 = 2 - x_1^2 - x_2^2, x_1^2 + x_2^2 \leq 2\},$$

gilt

$$\begin{aligned} O(P) &= \int_{x_1^2 + x_2^2 \leq 2} \sqrt{1 + 4x_1^2 + x_2^2} \, d(x_1, x_2) \\ &= \int_0^{\sqrt{2}} \int_0^{2\pi} \sqrt{1 + 4r^2} \, r \, d\varphi \, dr = \pi \int_0^2 \sqrt{1 + 4s} \, ds \\ &= \pi \left[\frac{1}{6} (1 + 4s)^{3/2} \right]_0^2 = \pi \left(\frac{1}{6} (27 - 1) \right) = \frac{13}{3} \pi \end{aligned}$$

Bemerkung.

Für das Kreuzprodukt zweier Vektoren $a, b \in \mathbb{R}^3$ gilt

$$\|a \times b\|^2 = \|a\|^2 \|b\|^2 - \langle a, b \rangle^2$$

Daraus folgt

$$\left\| \frac{\partial p}{\partial x_1} \times \frac{\partial p}{\partial x_2} \right\|^2 = \left\| \frac{\partial p}{\partial x_1} \right\|^2 \left\| \frac{\partial p}{\partial x_2} \right\|^2 - \left\langle \frac{\partial p}{\partial x_1}, \frac{\partial p}{\partial x_2} \right\rangle^2$$

Definiert man

$$E := \left\| \frac{\partial p}{\partial x_1} \right\|^2, \quad F := \left\langle \frac{\partial p}{\partial x_1}, \frac{\partial p}{\partial x_2} \right\rangle^2, \quad G := \left\| \frac{\partial p}{\partial x_2} \right\|^2,$$

so ergibt sich die Beziehung

$$do = \sqrt{EG - F^2} d(u_1, u_2)$$

Beispiel.

Für das Oberflächenelement der **Sphäre**

$$S_r^2 = \{(x_1, x_2, x_3)^T \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = r^2\}$$

ergeben sich mit der Parametrisierung über Kugelkoordinaten

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = r \begin{pmatrix} \cos \varphi \cos \theta \\ \sin \varphi \cos \theta \\ \sin \theta \end{pmatrix} \quad \text{für } (\varphi, \theta) \in [0, 2\pi] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

die Beziehungen

$$\frac{\partial \mathbf{p}}{\partial \varphi} = r \begin{pmatrix} -\sin \varphi \cos \theta \\ \cos \varphi \cos \theta \\ 0 \end{pmatrix} \quad \text{und} \quad \frac{\partial \mathbf{p}}{\partial \theta} = r \begin{pmatrix} -\cos \varphi \sin \theta \\ -\sin \varphi \sin \theta \\ \cos \theta \end{pmatrix}$$

Daraus folgt

$$E = r^2 \cos^2 \theta, \quad F = 0, \quad G = r^2$$

Fortsetzung des Beispiels.

Mit

$$E = r^2 \cos^2 \theta, \quad F = 0, \quad G = r^2$$

folgt aus der Beziehung

$$do = \sqrt{EG - F^2} d(u_1, u_2)$$

daher

$$do = r^2 \cos \theta d(\varphi, \theta) \quad \text{für } (\varphi, \theta) \in [0, 2\pi] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

Wir können nun die Oberfläche der Kugel wie folgt berechnen.

$$\begin{aligned} O &= \int_{S_r^2} do = \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} r^2 \cos \theta d\varphi d\theta \\ &= 2\pi r^2 \sin \theta \Big|_{-\pi/2}^{\pi/2} = 4\pi r^2 \end{aligned}$$

Oberflächenintegrale erster und zweiter Art.

Definition: Sei $x = p(u)$ eine $\mathcal{C}^1$ -Parametrisierung einer Fläche $F = p(K)$, wobei $K \subset D$ kompakt, messbar und zusammenhängend ist.

- Für eine stetige Funktion $f : F \rightarrow \mathbb{R}$ ist das **Oberflächenintegral 1. Art** definiert durch

$$\int_F f(x) \, do := \int_K f(p(u)) \left\| \frac{\partial p}{\partial u_1} \times \frac{\partial p}{\partial u_2} \right\| du$$

- Für ein stetiges Vektorfeld $f : F \rightarrow \mathbb{R}^3$ ist das **Oberflächenintegral 2. Art** definiert durch

$$\int_F f(x) \, do := \int_K \left\langle f(p(u)), \frac{\partial p}{\partial u_1} \times \frac{\partial p}{\partial u_2} \right\rangle du$$

Alternative Darstellung für Oberflächenintegrale.

Andere Darstellungen des Oberflächenintegrals 2. Art

Der Einheitsnormalenvektor $n(x)$ auf der Fläche F ist gegeben durch

$$n(x) = n(p(u)) = \frac{\frac{\partial p}{\partial u_1} \times \frac{\partial p}{\partial u_2}}{\left\| \frac{\partial p}{\partial u_1} \times \frac{\partial p}{\partial u_2} \right\|}$$

Wir schreiben daher auch

$$\begin{aligned} \int_F f(x) \, do &= \int_K \left\langle f(p(u)), \frac{\partial p}{\partial u_1} \times \frac{\partial p}{\partial u_2} \right\rangle du \\ &= \int_K \langle f(p(u)), n(p(u)) \rangle \left\| \frac{\partial p}{\partial u_1} \times \frac{\partial p}{\partial u_2} \right\| du \\ &= \int_F \langle f(x), n(x) \rangle \, do \end{aligned}$$

Interpretation der Oberflächenintegrale.

Bemerkung:

- Ist $f(x)$ die Dichte einer massenbelegten Fläche, so liefert das Integral 1. Art gerade die Gesamtmasse der Fläche.
- Ist $f(x)$ ein Geschwindigkeitsfeld einer stationären Strömung, so liefert das Integral 2. Art die Flüssigkeitsmenge, die pro Zeiteinheit durch die Fläche F strömt, d.h. den **Fluss** von $f(x)$ durch die Fläche F .
- Ist F eine geschlossene Fläche, d.h. die Oberflächen eines kompakten und einfach zusammenhängenden Körpers im $\mathbb{R}^3$, so schreiben wir

$$\oint_F f(x) \, d\sigma \quad \text{bzw.} \quad \oint_F f(x) \, d\sigma$$

Die Parametrisierung ist dabei so gewählt, dass der Einheitsnormalenvektor $n(x)$ nach außen weist.

Der Integralsatz von Gauß.

Satz: (**Integralsatz von Gauß**) Sei $G \subset \mathbb{R}^3$ ein kompakter und messbarer Standardbereich, d.h. G sei bezüglich jeder Koordinate projizierbar. Der Rand ∂G bestehe aus endlich vielen glatten Flächenstücken mit äußerer Normale $n(x)$.

Ist $f : D \rightarrow \mathbb{R}^3$ ein $\mathcal{C}^1$ -Vektorfeld mit $G \subset D$, so gilt

$$\int_G \operatorname{div} f(x) \, dx = \oint_{\partial G} f(x) \, do$$

Interpretation des Gaußschen Integralsatzes: Die linke Seite ist ein Bereichsintegral über die skalare Funktion $g(x) := \operatorname{div} f(x)$. Die rechte Seite ist ein Oberflächenintegral 2. Art bezüglich des Vektorfeldes $f(x)$. Ist $f(x)$ das Geschwindigkeitsfeld einer **inkompressiblen** Strömung, so gilt $\operatorname{div} f(x) = 0$ und daher

$$\oint_{\partial G} f(x) \, do = 0$$

Beispiel.

Wir betrachten das Vektorfeld

$$f(x) = x = (x_1, x_2, x_3)^T$$

und die Kugel K :

$$K := \{(x_1, x_2, x_3)^T \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 \leq 1\}$$

Dann gilt offensichtlich

$$\operatorname{div} f(x) = 3$$

und damit

$$\int_K \operatorname{div} f(x) \, dx = 3 \cdot \operatorname{vol}(K) = 4\pi$$

Das entsprechende Oberflächenintegral lässt sich am besten durch Übergang auf Kugelkoordinaten, d.h. durch die Parametrisierung der Kugel mit Kugelkoordinaten, berechnen.

Die Formeln von Green.

Satz: (**Formeln von Green**) Die Menge $G \subset \mathbb{R}^3$ erfülle die Voraussetzungen des Gaußschen Integralsatzes. Für $\mathcal{C}^2$ -Funktionen $f, g : D \rightarrow \mathbb{R}$, $G \subset D$, gelten dann die Relationen:

$$\int_G (f \Delta g + \langle \nabla f, \nabla g \rangle) dx = \oint_{\partial G} f \frac{\partial g}{\partial n} do$$

$$\int_G (f \Delta g - g \Delta f) dx = \oint_{\partial G} \left(f \frac{\partial g}{\partial n} - g \frac{\partial f}{\partial n} \right) do$$

Hierbei bezeichnet

$$\frac{\partial f}{\partial n}(x) = D_n f(x) \quad \text{für } x \in \partial G$$

die Richtungsableitung von $f(x)$ in Richtung des äußeren Einheitsnormalenvektors $n(x)$.

Beweis der Greenschen Formeln.

Wir setzen

$$F(x) = f(x) \cdot \nabla g(x)$$

Dann gilt

$$\begin{aligned}\operatorname{div} F(x) &= \frac{\partial}{\partial x_1} \left(f \cdot \frac{\partial g}{\partial x_1} \right) + \frac{\partial}{\partial x_2} \left(f \cdot \frac{\partial g}{\partial x_2} \right) + \frac{\partial}{\partial x_3} \left(f \cdot \frac{\partial g}{\partial x_3} \right) \\ &= f \cdot \Delta g + \langle \nabla f, \nabla g \rangle\end{aligned}$$

Wir wenden nun den **Gaußschen Integralsatz** an:

$$\begin{aligned}\int_G (f \Delta g + \langle \nabla f, \nabla g \rangle) dx &= \int_G \operatorname{div} F(x) dx = \oint_{\partial G} \langle F, n \rangle do \\ &= \oint_{\partial G} f \langle \nabla g, n \rangle do = \oint_{\partial G} f \frac{\partial g}{\partial n} do\end{aligned}$$

Die zweite Greensche Formel folgt direkt durch Vertauschen von f und g .

Der Integralsatz von Stokes.

Satz: (Integralsatz von Stokes)

Sei $f : D \rightarrow \mathbb{R}^3$ ein $\mathcal{C}^1$ -Vektorfeld auf einem Gebiet $D \subset \mathbb{R}^3$.

Weiter sei $F = p(K)$ eine Fläche in D , $F \subset D$, mit der Parametrisierung $x = p(u)$, $u \in \mathbb{R}^2$, und $K \subset \mathbb{R}^2$ sei ein Greenscher Bereich.

Der Rand ∂K werde durch eine stückweise glatte $\mathcal{C}^1$ -Kurve c parametrisiert, deren Bild $\tilde{c}(t) := p(c(t))$ dann den Rand ∂F der Fläche F parametrisiert.

Die Orientierung der Randkurve $\tilde{c}(t)$ sei hierbei so gewählt, dass $n(\tilde{c}(t)) \times \dot{\tilde{c}}(t)$ in Richtung der Fläche weist.

Dann gilt

$$\int_F \operatorname{rot} f(x) \, d\sigma = \oint_{\partial F} f(x) \, dx$$

Beispiel.

Gegeben sei das Vektorfeld

$$f(x, y, z) = (-y, x, -z)^T$$

und die geschlossene Kurve $c : [0, 2\pi] \rightarrow \mathbb{R}^3$ sei parametrisiert durch

$$c(t) = (\cos t, \sin t, 0)^T \quad \text{für } 0 \leq t \leq 2\pi$$

Dann gilt:

$$\begin{aligned} \oint_c f(x) \, dx &= \int_0^{2\pi} \langle f(c(t)), \dot{c}(t) \rangle \, dt \\ &= \int_0^{2\pi} \left\langle \begin{pmatrix} -\sin t \\ \cos t \\ 0 \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \\ 0 \end{pmatrix} \right\rangle \, dt \\ &= \int_0^{2\pi} (\sin^2 t + \cos^2 t) \, dt = 2\pi \end{aligned}$$

Fortsetzung des Beispiels.

Wir definieren nun eine Fläche $F \subset \mathbb{R}^3$, die durch die Kurve $c(t)$ berandet wird:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos \varphi \cos \psi \\ \sin \varphi \cos \psi \\ \sin \psi \end{pmatrix} =: p(\varphi, \psi)$$

mit $(\varphi, \psi) \in K = [0, 2\pi] \times [0, \pi/2]$, d.h. die Fläche F ist gerade die obere Kugelhälfte.

Der Integralsatz von Stokes besagt nun:

$$\int_F \operatorname{rot} f(x) \, d\sigma = \oint_{c=\partial F} f(x) \, dx$$

Wir haben bereits die rechte Seite, ein **Kurvenintegral 2. Art**, berechnet:

$$\oint_{c=\partial F} f(x) \, dx = 2\pi$$

Komplettierung des Beispiels.

Es bleibt also das **Oberflächenintegral 2. Art**:

$$\int_F \operatorname{rot} f(x) \, d\sigma = \int_K \left\langle \operatorname{rot} f(p(\varphi, \psi)), \frac{\partial p}{\partial \varphi} \times \frac{\partial p}{\partial \psi} \right\rangle d\varphi d\psi$$

Beachte: Die rechte Seite ist ein **Bereichsintegral**.

Man rechnet $\operatorname{rot} f(x) = (0, 0, 2)^T$ direkt nach, sowie

$$\frac{\partial p}{\partial \varphi} \times \frac{\partial p}{\partial \psi} = \begin{pmatrix} \cos \varphi \cos^2 \psi \\ \sin \varphi \cos^2 \psi \\ \sin \psi \cos \psi \end{pmatrix}$$

Daraus folgt:

$$\int_F \operatorname{rot} f(x) \, d\sigma = \int_0^{\pi/2} \int_0^{2\pi} 2 \sin \psi \cos \psi \, d\varphi d\psi = 2\pi \int_0^{\pi/2} \sin(2\psi) \, d\psi = 2\pi$$