

**Analysis III für
Studierende der Ingenieurwissenschaften**

Ingenuin Gasser
Fachbereich Mathematik
Universität Hamburg

Technische Universität Hamburg–Harburg
Wintersemester 2009/10

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 Größe p , V und T läßt sich als Funktion der anderen darstellen:

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

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

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

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

- 1) $f(\mathbf{x})$ heißt in \mathbf{x}^0 nach x_i **partiell differenzierbar**, falls der Grenzwert

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

existiert. Dabei bezeichnet e_i den i -ten Einheitsvektor.

Den Grenzwert nennt man die **partielle Ableitung** von $f(\mathbf{x})$ nach x_i im Punkt \mathbf{x}^0 .

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

Beispiele:

1) Betrachte die Funktion

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

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

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

Die Funktion ist also eine $C^1(\mathbb{R}^2)$ -Funktion.

2) Die Funktion

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

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

Konkretes Beispiel:

Der Schalldruck einer 1–d 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.

Bemerkungen:

1) Sind f, g partiell nach x_i differenzierbar, $\alpha, \beta \in \mathbb{R}$, so gelten:

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

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

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

2) Man verwendet auch **andere Bezeichnungen**:

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

Definition: Sei $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^n$ offen, in einem Punkt $\mathbf{x}^0 \in \mathbb{R}^n$ (nach allen Koordinaten) partiell differenzierbar.

1) Der **Zeilenvektor**

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

heißt die **Ableitung von $f(\mathbf{x})$ in \mathbf{x}^0** .

2) Man schreibt auch als **Spaltenvektor**:

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

und bezeichnet dies als **Gradient**, den symbolischen Vektor

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

als den **Nabla-Operator**.

Bemerkung: Seien $f(\mathbf{x})$ und $g(\mathbf{x})$ partiell differenzierbar auf D .

Dann gelten die folgenden **Differentiationsregeln**:

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

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

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

Beispiel:

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

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

2) Für $r(\mathbf{x}) := \|\mathbf{x}\|_2$ gilt:

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

Wichtige Beobachtung:

Eine nach allen Koordinaten partiell differenzierbare Funktion ist

NICHT unbedingt stetig!!

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

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

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

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

Beispiel: (Fortsetzung)

Berechnung der partiellen Ableitungen im Punkt $(x^0, y^0) = (0, 0)$:

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

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

Aber: Im Punkt $(x^0, y^0) = (0, 0)$ ist die Funktion **nicht** stetig:

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

Also gilt:

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

Damit eine partiell differenzierbare Funktion **auch** stetig ist, benötigt man zusätzliche Eigenschaften, z.B.

Alle partiellen Ableitungen sind beschränkt.

Satz:

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

Bemerkung:

In unserem **Beispiel** sind die partiellen Ableitungen in einer Umgebung vom Punkt $(x^0, y^0) = (0, 0)$ **nicht** beschränkt:

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

Beweis des Satzes:

Für $\|\mathbf{x} - \mathbf{x}^0\|_\infty < \varepsilon$, ε hinreichend klein, schreiben wir:

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

Bei jeder Differenz auf der linken Seite, betrachten wir f als eine Funktion einer Variablen, zum Beispiel

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

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 .

Anwendung des **MWS** für Funktionen einer Variablen auf jeden Term der rechten Seite ergibt:

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

Sind die partiellen Ableitungen in der Umgebung $\|\mathbf{x} - \mathbf{x}^0\|_\infty < \varepsilon$ **beschränkt**, so gilt:

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

und damit ist $f(\mathbf{x})$ **stetig** in \mathbf{x}^0 , denn

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

Höhere Ableitungen

Definition:

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

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

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

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

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

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

Definition: (Fortsetzung)

Die Funktion $f(\mathbf{x})$ heißt **k -fach partiell differenzierbar**, falls alle Ableitungen der Ordnung k auf D existieren:

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

Sind all diese Ableitungen zudem stetig, so heißt die Funktion $f(\mathbf{x})$ **k -fach stetig partiell differenzierbar** oder auch **C^k -Funktion** auf D , $k = 1, 2, 3, \dots$

Stetige Funktionen $f(\mathbf{x})$ nennt man auch **C^0 -Funktionen**.

Beispiel: Gegeben sei die Funktion $f(x_1, \dots, x_n) = \prod_{i=1}^n x_i^i$.

Dann gilt:

$$\frac{\partial^n f}{\partial x_n \dots \partial x_1} = ?$$

ACHTUNG:

Die Reihenfolge, in der die partiellen Ableitungen durchzuführen sind, ist

i. Allg. **nicht beliebig vertauschbar!**

Beispiel: Gegeben sei die Funktion

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

Man berechnet direkt

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

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

Satz: (Vertauschbarkeitssatz von Schwarz)

Ist $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^n$ offen, eine C^2 -Funktion, so gilt für alle $i, j \in \{1, \dots, n\}$:

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

Beweisidee:

Zweifache Anwendung des Mittelwertsatzes.

Folgerung:

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

Beispiel: Gegeben sei die Funktion

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

Zu berechnen ist die partielle Ableitung dritter Ordnung f_{xyz} .

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

1) Leite zunächst nach z ab:

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

2) Jetzt leiten wir f_z nach x (damit fällt $\cosh y$ raus):

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

3) Für die partielle Ableitung von f_{zx} nach y erhalten wir:

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

Der **Laplace–Operator** ist definiert durch

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

Für eine skalare Funktion $u(\mathbf{x}) = u(x_1, \dots, x_n)$ ist also

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

Bedeutung: Partielle Differentialgleichungen zweiter Ordnung

$$u_{tt} = \Delta u \quad (\text{Wellengleichung})$$

$$u_t = \Delta u \quad (\text{Wärmeleitungsgleichung})$$

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

Vektorwertige Funktionen

Definition: Sei $f : D \rightarrow \mathbb{R}^m$, $D \subset \mathbb{R}^n$ offen, eine vektorwertige Funktion.

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

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

existieren.

Die Berechnung erfolgt komponentenweise

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_i} \\ \vdots \\ \frac{\partial f_m}{\partial x_i} \end{pmatrix}$$

Definition:

Für $m = n$ nennt man die Funktion $f : D \rightarrow \mathbb{R}^n$ ein **Vektorfeld** auf D . Ist jede Koordinatenfunktion $f_i(\mathbf{x})$ 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 oder
- Temperaturgradienten in Festkörpern.

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

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

Andere Notationen:

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

Bemerkung:

Es gelten die folgenden Rechenregeln:

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

$$\operatorname{div}(\varphi \cdot \mathbf{f}) = (\nabla \varphi, \mathbf{f}) + \varphi \operatorname{div} \mathbf{f}$$

Bemerkung:

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

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

Definition:

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

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

Andere Notationen:

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

Bemerkung:

Es gelten die folgenden Rechenregeln:

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

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

Bemerkung:

Ist $\varphi : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^3$, eine C^2 -Funktion, so folgt aufgrund des Vertauschbarkeitssatzes von Schwarz

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

d.h. Gradientenfelder sind stets rotationsfrei.

1.2 Das vollständige Differential

Definition:

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

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

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

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

d.h.

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

erfüllt ist.

Bezeichnungen:

Man nennt die lineare Abbildung l das (**vollständige** oder **totale**)

Differential von $f(x)$ im Punkt x^0 und 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 aus Analysis I 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 $\langle a^T, x - x^0 \rangle$.

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

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

$$\mathbf{J}f(\mathbf{x}^0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}^0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}^0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}^0) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}^0) \end{pmatrix} = \begin{pmatrix} Df_1(\mathbf{x}^0) \\ \vdots \\ Df_m(\mathbf{x}^0) \end{pmatrix}$$

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

Bemerkung:

Man beachte, dass **differenzierbar** hier **vollständig/total** differenzierbar bedeutet.

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 $\mathbf{x} = \mathbf{x}^0 + t\mathbf{e}_i$, $|t| < \varepsilon$, $i \in \{1, \dots, n\}$.

Da f im Punkt \mathbf{x}^0 differenzierbar ist, folgt

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

Wir schreiben nun

$$\begin{aligned} \frac{f(\mathbf{x}) - f(\mathbf{x}^0) - \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}^0)}{\|\mathbf{x} - \mathbf{x}^0\|_\infty} &= \frac{f(\mathbf{x}^0 + t\mathbf{e}_i) - f(\mathbf{x}^0) - \frac{t\mathbf{A}\mathbf{e}_i}{|t|}}{|t|} \\ &= \frac{t}{|t|} \cdot \left(\frac{f(\mathbf{x}^0 + t\mathbf{e}_i) - f(\mathbf{x}^0)}{t} - \mathbf{A}\mathbf{e}_i \right) \\ &\rightarrow 0 \quad (t \rightarrow 0) \end{aligned}$$

Daraus folgt

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x}^0 + t\mathbf{e}_i) - f(\mathbf{x}^0)}{t} = A\mathbf{e}_i \quad i = 1, \dots, n$$

Beispiel:

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

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

- b) 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 $s = x_1 + 2x_2 + 3x_3$.

Beispiele:

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

$$\mathbf{J}f(\mathbf{x}) = \mathbf{A} \quad \forall \mathbf{x} \in \mathbb{R}^n$$

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

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

Daraus folgt

$$\mathbf{J}f(\mathbf{x}) = Df(\mathbf{x}) = \mathbf{x}^T (\mathbf{A}^T + \mathbf{A})$$

Satz: (Differentiationsregeln)

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

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

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

2) **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 $\mathbf{x}^0 = h(t_0)$ 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:

$$\begin{aligned} (f \circ h)'(t_0) &= \mathbf{J}f(h(t_0)) \cdot \mathbf{J}h(t_0) \\ &= Df(h(t_0)) \cdot \mathbf{h}'(t_0) \\ &= \sum_{k=1}^n \frac{\partial f}{\partial x_k}(h(t_0)) \cdot h_k(t_0) \end{aligned}$$

Richtungsableitungen

Definition: Sei $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^n$ offen, $\mathbf{x}^0 \in D$. Für einen Vektor $\mathbf{v} \in \mathbb{R} \setminus \{0\}$ heißt

$$D_{\mathbf{v}} f(\mathbf{x}^0) := \lim_{t \rightarrow 0} \frac{f(\mathbf{x}^0 + t\mathbf{v}) - f(\mathbf{x}^0)}{t}$$

die **Richtungsableitung (Gateaux-Ableitung)** von $f(\mathbf{x})$ in Richtung \mathbf{v} .

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

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

Bemerkung:

- 1) Für $v = e_j$ ist die Richtungsableitung gerade die partielle Ableitung nach x_j :

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

- 2) Ist v ein Einheitsvektor, also $\|v\| = 1$, so beschreibt die Richtungsableitung $D_v f(\mathbf{x}^0)$ die Steigung von $f(\mathbf{x})$ in Richtung v .
- 3) Ist $f(\mathbf{x})$ in \mathbf{x}^0 differenzierbar, so existieren sämtliche Richtungsableitungen von $f(\mathbf{x})$ in \mathbf{x}^0 und es gilt:

$$D_v f(\mathbf{x}^0) = \text{grad } f(\mathbf{x}^0) \cdot v$$

Dies folgt aus Anwendung der Kettenregel für $h(t) := \mathbf{x}^0 + tv$

$$D_v f(\mathbf{x}^0) = \frac{d}{dt}(f \circ h)|_{t=0} = \text{grad } f(\mathbf{x}^0) \cdot h'(0)$$

Satz: (Eigenschaften des Gradienten)

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

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

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

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

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

Beweisidee:

- 1) Anwendung der Kettenregel

2) Cauchy–Schwarzsche Ungleichung

$$|D_{\mathbf{v}} f(\mathbf{x}^0)| \leq \|\text{grad } f(\mathbf{x}^0)\|_2$$

Krummlinige Koordinaten

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

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

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

Beispiel: Polarkoordinaten

Betrachte für $n = 2$ die (Polar)Koordinaten $\mathbf{u} = (r, \varphi)$ mit $r > 0$ und $-\pi < \varphi < \pi$ und setze

$$x = r \cos \varphi$$

$$y = r \sin \varphi$$

mit den kartesischen Koordinaten $\mathbf{x} = (x, y)$.

Umrechnung von partiellen Ableitungen von u auf x

Zunächst gelten für alle $u \in U$ mit $x = \Phi(u)$ die folgenden Relationen:

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

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

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

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

$$g^{ij} := \frac{\partial \Phi_j}{\partial u_i},$$

$$\mathbf{G}(\mathbf{u}) := (g^{ij}) = (\mathbf{J} \Phi(\mathbf{u}))^T$$

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:

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

mit

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

Man erhält diese Beziehungen durch Anwendung der Kettenregel auf die Funktion 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

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

und damit

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

Die Umrechnung der partiellen Ableitungen lautet:

$$\begin{aligned}\frac{\partial}{\partial x} &= \cos \varphi \frac{\partial}{\partial r} - \frac{1}{r} \sin \varphi \frac{\partial}{\partial \varphi} \\ \frac{\partial}{\partial y} &= \sin \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \varphi \frac{\partial}{\partial \varphi}\end{aligned}$$

Damit lässt sich etwa der **Laplace-Operator** auf Polarkoordinaten umrechnen:

$$\begin{aligned}\frac{\partial^2}{\partial x^2} &= \cos^2 \varphi \frac{\partial^2}{\partial r^2} - \frac{\sin(2\varphi)}{r} \frac{\partial^2}{\partial r \partial \varphi} + \frac{\sin^2 \varphi}{r^2} \frac{\partial^2}{\partial \varphi^2} + \frac{\sin(2\varphi)}{r^2} \frac{\partial}{\partial \varphi} + \frac{\sin^2 \varphi}{r} \frac{\partial}{\partial r} \\ \frac{\partial^2}{\partial y^2} &= \sin^2 \varphi \frac{\partial^2}{\partial r^2} + \frac{\sin(2\varphi)}{r} \frac{\partial^2}{\partial r \partial \varphi} + \frac{\cos^2 \varphi}{r^2} \frac{\partial^2}{\partial \varphi^2} - \frac{\sin(2\varphi)}{r^2} \frac{\partial}{\partial \varphi} + \frac{\cos^2 \varphi}{r} \frac{\partial}{\partial r} \\ \Delta &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} + \frac{1}{r} \frac{\partial}{\partial r}\end{aligned}$$

Beispiel: (Kugelkoordinaten)

Wir betrachten die Kugelkoordinaten

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

Die Jacobi-Matrix ist dann gegeben durch:

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

Partielle Ableitungen:

$$\begin{aligned} \frac{\partial}{\partial x} &= \cos \varphi \cos \theta \frac{\partial}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial}{\partial \varphi} - \frac{1}{r} \cos \varphi \sin \theta \frac{\partial}{\partial \theta} \\ \frac{\partial}{\partial y} &= \sin \varphi \cos \theta \frac{\partial}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial}{\partial \varphi} - \frac{1}{r} \sin \varphi \sin \theta \frac{\partial}{\partial \theta} \\ \frac{\partial}{\partial z} &= \sin \theta \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \frac{\partial}{\partial \theta} \end{aligned}$$

Laplace-Operator:

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

1.3 Mittelwertsätze und Taylor-Entwicklungen

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

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

ganz in D liegt.

Dann gibt es eine Zahl $\theta \in (0, 1)$ mit

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

Beweis: Wir setzen

$$h(t) := f(\mathbf{a} + t(\mathbf{b} - \mathbf{a}))$$

Aus dem MWS für eine Veränderliche folgt dann mit der Kettenregel

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

Bemerkung: 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 $\theta = \frac{1}{2}$.

ACHTUNG:

Der Mittelwertsatz für mehrere Variablen gilt nur für **skalare** Funktionen!!!

Beispiel: Betrachte die **vektorwertige** Funktion

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

Nun gilt

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

und

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

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

Satz: (Mittelwert–Abschätzungssatz)

Die Funktion $f : D \rightarrow \mathbb{R}^m$ sei differenzierbar auf der offenen Menge $D \subset \mathbb{R}^n$. Ferner seien \mathbf{a}, \mathbf{b} Punkte in D mit $[\mathbf{a}, \mathbf{b}] \subset D$.

Dann existiert ein $\theta \in (0, 1)$ mit

$$\|f(\mathbf{b}) - f(\mathbf{a})\|_2 \leq \|\mathbf{J}f(\mathbf{a} + \theta(\mathbf{b} - \mathbf{a}))\|_2 \cdot \|\mathbf{b} - \mathbf{a}\|_2$$

Beweis:

Anwendung des Mittelwertsatzes auf die **skalare** Funktion $g(\mathbf{x})$ definiert durch

$$g(\mathbf{x}) := (f(\mathbf{b}) - f(\mathbf{a}))^T f(\mathbf{x})$$

Bemerkung:

Eine andere (abgeschwächte) Form der Mittelwert–Abschätzung ist

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

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

Taylor–Entwicklungen für skalare Funktionen mehrerer Variablen

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

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

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

$$\mathbf{x}^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$$

Satz von Taylor:

Sei $D \subset \mathbb{R}^n$ offen und konvex, $f : D \rightarrow \mathbb{R}$ eine C^{m+1} -Funktion und sei $\mathbf{x}_0 \in D$.

Dann gilt für alle $\mathbf{x} \in D$ die folgenden Entwicklung nach Taylor:

$$f(\mathbf{x}) = T_m(\mathbf{x}; \mathbf{x}_0) + R_m(\mathbf{x}; \mathbf{x}_0)$$

$$T_m(\mathbf{x}; \mathbf{x}_0) = \sum_{|\alpha| \leq m} \frac{D^\alpha f(\mathbf{x}_0)}{\alpha!} (\mathbf{x} - \mathbf{x}_0)^\alpha$$

$$R_m(\mathbf{x}; \mathbf{x}_0) = \sum_{|\alpha| = m+1} \frac{D^\alpha f(\mathbf{x}_0 + \theta(\mathbf{x} - \mathbf{x}_0))}{\alpha!} (\mathbf{x} - \mathbf{x}_0)^\alpha$$

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

Beachte: Summation über $|\alpha| \leq m$ und $|\alpha| = m + 1$.

Herleitung der Taylorformel:

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

$$g(t) := f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))$$

und berechnen die Taylor-Entwicklung um $t = 0$.

Es gilt:

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

Berechnung von $g'(0)$:

$$\begin{aligned} g'(0) &= \frac{d}{dt} f(x_1^0 + t(x_1 - x_1^0), x_2^0 + t(x_2 - x_2^0), \dots, x_n^0 + t(x_n - x_n^0))|_{t=0} \\ &= D_1 f(\mathbf{x}_0)(x_1 - x_1^0) + \dots + D_n f(\mathbf{x}_0)(x_n - x_n^0) \\ &= \sum_{|\alpha|=1} \frac{D^\alpha f(\mathbf{x}_0)}{\alpha!} (\mathbf{x} - \mathbf{x}_0)^\alpha \end{aligned}$$

Taylor-Entwicklung um $t = 0$

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(0) + \frac{1}{6}g^{(3)}(\xi), \quad \xi \in (0, 1)$$

Berechnung von $g''(0)$:

$$\begin{aligned} \frac{1}{2}g''(0) &= \frac{1}{2} \frac{d^2}{dt^2} f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) \Big|_{t=0} \\ &= \frac{1}{2} (D_{11}f(\mathbf{x}_0)(x_1 - x_1^0)^2 + D_{21}f(\mathbf{x}_0)(x_1 - x_1^0)(x_2 - x_2^0) \\ &\quad + \dots + D_{ij}f(\mathbf{x}_0)(x_i - x_i^0)(x_j - x_j^0) + \dots + \\ &\quad + D_{n-1,n}f(\mathbf{x}_0)(x_{n-1} - x_{n-1}^0)(x_n - x_n^0) + D_{nn}f(\mathbf{x}_0)(x_n - x_n^0)^2 \\ &= \sum_{|\alpha|=2} \frac{D^\alpha f(\mathbf{x}_0)}{\alpha!} (\mathbf{x} - \mathbf{x}_0)^\alpha \quad (\text{Vertauschungssatz von Schwarz!}) \end{aligned}$$

Beweis der Taylor-Formel mittels vollständiger Induktion!

Beispiel:

Berechne das Taylor–Polynom $T_2(\mathbf{x}; \mathbf{x}_0)$ zweiten Grades der Funktion

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

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

Zur Berechnung von $T_2(\mathbf{x}; \mathbf{x}_0)$ benötigen wir die partiellen Ableitungen bis zur zweiten Ordnung.

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

Als Ergebnis erhält man $T_2(\mathbf{x}; \mathbf{x}_0)$ in der Form

$$T_2(\mathbf{x}; \mathbf{x}_0) = 4z(x + y - 2)$$

Berechnung auf Folie

Beispiel:

Man berechne das Taylor–Polynom $T_3(\mathbf{x}; \mathbf{x}_0)$ dritten Grades der Funktion

$$f(x, y) = e^y \cos x$$

zum Entwicklungspunkt $\mathbf{x}_0 = (0, 0)^T$:

- 1) unter Verwendung des Taylorschen Satzes,
- 2) unter Verwendung bekannter Reihenentwicklungen.

Berechnung auf Folie

Bemerkung:

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

$$R_m(\mathbf{x}; \mathbf{x}_0) = \sum_{|\alpha|=m+1} \frac{D^\alpha f(\mathbf{x}_0 + \theta(\mathbf{x} - \mathbf{x}_0))}{\alpha!} (\mathbf{x} - \mathbf{x}_0)^\alpha$$

Sind all diese Ableitungen in der Nähe von \mathbf{x}_0 beschränkt durch C , so gilt die **Restgliedabschätzung**

$$|R_m(\mathbf{x}; \mathbf{x}_0)| \leq \frac{n^{m+1}}{(m+1)!} C \|\mathbf{x} - \mathbf{x}_0\|_\infty^{m+1}$$

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

$$f(\mathbf{x}) = T_m(\mathbf{x}; \mathbf{x}_0) + O\left(\|\mathbf{x} - \mathbf{x}_0\|^{m+1}\right)$$

Bemerkung: (Fortsetzung)

2) Man nennt die Matrix

$$\mathbf{H}f(\mathbf{x}_0) := \begin{pmatrix} f_{x_1 x_1}(\mathbf{x}_0) & \cdots & f_{x_1 x_n}(\mathbf{x}_0) \\ \vdots & & \vdots \\ f_{x_n x_1}(\mathbf{x}_0) & \cdots & f_{x_n x_n}(\mathbf{x}_0) \end{pmatrix}$$

die **Hesse-Matrix** von $f(\mathbf{x})$ im Punkt \mathbf{x}_0 .

Hesse-Matrix = Jacobi-Matrix des Gradienten ∇f

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

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \text{grad } f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + O(\|\mathbf{x} - \mathbf{x}_0\|^3)$$

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

Kapitel 2: Anwendung der Differentialrechnung mehrerer Variablen

2.1 Extrema von Funktionen mehrerer Variablen

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

1) $f(\mathbf{x})$ hat in \mathbf{x}^0 ein **globales Maximum**, falls gilt:

$$\forall \mathbf{x} \in D : f(\mathbf{x}) \leq f(\mathbf{x}^0)$$

2) $f(\mathbf{x})$ hat in \mathbf{x}^0 ein **strenges globales Maximum**, falls gilt:

$$\forall \mathbf{x} \in D \setminus \{\mathbf{x}^0\} : f(\mathbf{x}) < f(\mathbf{x}^0)$$

3) $f(\mathbf{x})$ hat in \mathbf{x}^0 ein **lokales Maximum**, falls es ein ε gibt mit:

$$\forall \mathbf{x} \in D : \|\mathbf{x} - \mathbf{x}^0\| < \varepsilon \Rightarrow f(\mathbf{x}) \leq f(\mathbf{x}^0)$$

4) $f(\mathbf{x})$ hat in \mathbf{x}^0 ein **strenges lokales Maximum**, falls es ein ε gibt mit:

$$\forall \mathbf{x} \in D : 0 < \|\mathbf{x} - \mathbf{x}^0\| < \varepsilon \Rightarrow f(\mathbf{x}) < f(\mathbf{x}^0)$$

Analoge Definitionen für minimale Werte (siehe **Analysis I**)

Satz: (Notwendige Bedingung I)

Besitzt $f(\mathbf{x})$ mit $f \in \mathcal{C}^2$ in einem Punkt $\mathbf{x}^0 \in D^0$ ein lokales Extremum (Minimum oder Maximum), so gilt $\text{grad } f(\mathbf{x}^0) = (0, \dots, 0)^T$.

Beweis:

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

$$\varphi(t) := f(\mathbf{x}^0 + t\mathbf{v})$$

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

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

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

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

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

Bemerkung:

1) Die Bedingung $\text{grad } f(\mathbf{x}^0) = (0, \dots, 0)^T$ definiert gewöhnlich ein nichtlineares Gleichungssystem zur Berechnung von $\mathbf{x} = \mathbf{x}^0$, wobei n Gleichungen für n Unbekannte gegeben sind.

2) Die Punkte $\mathbf{x}^0 \in D^0$ mit $\text{grad } f(\mathbf{x}^0) = (0, \dots, 0)^T$ nennt man **stationäre Punkte** von $f(x)$. Stationäre Punkte sind **nicht** immer 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 $\mathbf{x}^0 = (0, 0)^T$, dieser ist jedoch ein **Sattelpunkt** von $f(x)$.

Satz: (Klassifikation stationärer Punkte)

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

1) Notwendige Bedingung II

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

2) 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: (zu Teil 1))

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

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

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

Der Gradient in der Taylorentwicklung entfällt, da $\text{grad } f(\mathbf{x}^0) = (0, \dots, 0)^T$ gilt.

Wegen (1) gilt auch

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

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

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

d.h. $\mathbf{H} f(\mathbf{x}^0)$ ist positiv semidefinit.

Beweis: (zu Teil 2))

Ist $\mathbf{H} f(\mathbf{x}^0)$ positiv definit, so ist $\mathbf{H} f(\mathbf{x})$ ebenfalls in einer hinreichend kleinen Umgebung $\mathbf{x} \in K_\varepsilon(\mathbf{x}^0) \subset D$ positiv definit.

Dies folgt aus der Stetigkeit der zweiten partiellen Ableitungen.

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

$$f(\mathbf{x}) - f(\mathbf{x}^0) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^0)^T \mathbf{H} f(\mathbf{x}^0 + \theta(\mathbf{x} - \mathbf{x}^0))(\mathbf{x} - \mathbf{x}^0) > 0$$

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

Ist $\mathbf{H} f(\mathbf{x}^0)$ indefinit, so existieren zu Eigenwerten von $\mathbf{H} f(\mathbf{x}^0)$ mit verschiedenen Vorzeichen gewisse Eigenvektoren \mathbf{v} , \mathbf{w} mit (zum Beispiel)

$$\mathbf{v}^T \mathbf{H} f(\mathbf{x}^0) \mathbf{v} > 0 \quad \mathbf{w}^T \mathbf{H} f(\mathbf{x}^0) \mathbf{w} < 0$$

Analog zu Teil 1) sieht man dann, dass $\varepsilon_1, \varepsilon_2 > 0$ existieren mit

$$f(\mathbf{x}^0 + \varepsilon \mathbf{v}) - f(\mathbf{x}^0) = \frac{1}{2}(\varepsilon \mathbf{v})^T \mathbf{H} f(\mathbf{x}^0 + \theta_1 \varepsilon \mathbf{v}) (\varepsilon \mathbf{v}) > 0$$

für alle $\varepsilon \in (0, \varepsilon_1)$ und

$$f(\mathbf{x}^0 + \varepsilon \mathbf{w}) - f(\mathbf{x}^0) = \frac{1}{2}(\varepsilon \mathbf{w})^T \mathbf{H} f(\mathbf{x}^0 + \theta_1 \varepsilon \mathbf{w}) (\varepsilon \mathbf{w}) < 0$$

für alle $\varepsilon \in (0, \varepsilon_2)$.

Damit ist gezeigt, dass \mathbf{x}^0 ein **Sattelpunkt** von $f(\mathbf{x})$ ist.

Bemerkung: (Geometrische Interpretation)

Die Hesse-Matrix kann positive und negative Eigenwerte besitzen.

Die zugehörigen Eigenvektoren geben dabei diejenigen Richtungen an, in denen die Funktion wächst beziehungsweise fällt.

Bemerkung:

1) Ein stationärer Punkt x^0 mit $\det Hf(x^0) = 0$ heißt **ausgeartet**.

Die Hesse-Matrix besitzt dann einen Eigenwert $\lambda = 0$.

Ist x^0 **nicht** ausgeartet, so existieren genau drei Fälle:

Alle Eigenwerte $> 0 \Rightarrow x^0$ ist strenges lokales Minimum

Alle Eigenwerte $< 0 \Rightarrow x^0$ ist strenges lokales Maximum

\exists Eigenwerte $\lambda_1 \cdot \lambda_2 < 0 \Rightarrow x^0$ Sattelpunkt

2) Es gelten die folgenden Implikationen:

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

\Downarrow

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

\Uparrow

Bemerkung: (Fortsetzung)

- 3) Ist $f(\mathbf{x})$ eine \mathcal{C}^3 -Funktion, \mathbf{x}^0 ein stationärer Punkt von $f(\mathbf{x})$ und $\mathbf{H}f(\mathbf{x}^0)$ positiv definit, so gilt die Abschätzung:

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

wobei λ_{min} den **kleinsten** Eigenwert der Hesse-Matrix bezeichnet.

Nach dem Satz von Taylor gilt dann:

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

mit einer geeigneten Konstanten $C > 0$.

In der Nähe von \mathbf{x}^0 wächst $f(\mathbf{x})$ also mindestens quadratisch mit dem Abstand von \mathbf{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 $\text{grad } f(x, y) = (0, 0)^T$ liefert die beiden stationären Punkte

$$\mathbf{x}^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mathbf{x}^1 = \begin{pmatrix} -2/3 \\ 0 \end{pmatrix}$$

Die Hesse-Matrizen an den Stellen \mathbf{x}^0 und \mathbf{x}^1 lauten

$$\mathbf{H}f(\mathbf{x}^0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \quad \mathbf{H}f(\mathbf{x}^1) = \begin{pmatrix} -2 & 0 \\ 0 & -10/3 \end{pmatrix}$$

Die Matrix $\mathbf{H}f(\mathbf{x}^0)$ ist indefinit, also ist \mathbf{x}^0 ein Sattelpunkt, $\mathbf{H}f(\mathbf{x}^1)$

ist negativ definit, also ist x^1 ein strenges lokales Maximum von $f(x)$.

2.2 Implizit definierte Funktionen

Untersuche die Lösungsmengen von nichtlinearen Gleichungssystemen der Form

$$\mathbf{g}(\mathbf{x}) = 0$$

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

Insbesondere gelte:

$$m < n$$

d.h. wir haben **weniger** Gleichungen als Unbekannte.

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

Frage: Kann man das System $g(\mathbf{x}) = 0$ nach bestimmten Unbekannten, zum Beispiel den letzten m Variablen x_{n-m+1}, \dots, x_n , **auflösen**? Existiert eine Funktion $f(x_1, \dots, x_{n-m})$ mit

$$g(\mathbf{x}) = 0 \Leftrightarrow \begin{pmatrix} x_{n-m+1} \\ \vdots \\ x_n \end{pmatrix} = f(x_1, \dots, x_{n-m})$$

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 $g(\mathbf{x}) = 0$ lässt sich (zumindest lokal) als Graph einer Funktion f darstellen.

Beispiel:

Die Kreisgleichung

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

definiert ein **unterbestimmtes** nichtlineares Gleichungssystem.

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 $g(\mathbf{x})$ eine affin-lineare Funktion, d.h.

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

Wir spalten die Variablen \mathbf{x} in zwei Vektoren auf

$$\begin{aligned} \mathbf{x}^{(1)} &= (x_1, \dots, x_{n-m})^T \in \mathbb{R}^{n-m} \\ \mathbf{x}^{(2)} &= (x_{n-m+1}, \dots, x_n)^T \in \mathbb{R}^m \end{aligned}$$

Aufspaltung der Matrix \mathbf{C} ergibt die Darstellung

$$g(\mathbf{x}) = \mathbf{B}\mathbf{x}^{(1)} + \mathbf{A}\mathbf{x}^{(2)} + \mathbf{b}$$

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

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

$$g(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x}^{(2)} = -\mathbf{A}^{-1}(\mathbf{B}\mathbf{x}^{(1)} + \mathbf{b})$$

$$= \mathbf{f}(\mathbf{x}^{(1)})$$

Beispiel: (Fortsetzung)

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 $x^{(1)}$!

Auflösbarkeit ist also möglich, falls die Jacobi-Matrix regulär ist.

Dies führt auf den

Satz über implizite Funktionen

Satz über implizite Funktionen:

Sei $g : D \rightarrow \mathbb{R}^m$ eine 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

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

regulär ist, 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

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

und

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

Beispiele:

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

man im Punkt $(x^0, y^0) = (0, r)$

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

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

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

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

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

Also

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

Beispiele: (Fortsetzung)

2) Betrachte die Gleichung

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

Es gilt:

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

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:

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

Eine **explizite** Auflösung nach y ist in diesem Fall nicht möglich!

Bemerkung: Implizites Differenzieren einer durch

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

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

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

Gilt

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

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

Definition:

- 1) Ein Lösungspunkt (x^0, y^0) der Gleichung $g(x, y) = 0$ mit $\text{grad } g(x^0, y^0) \neq (0, 0)^T$ heißt **regulärer Punkt**.
- 1) Ein Lösungspunkt (x^0, y^0) der Gleichung $g(x, y) = 0$ mit $\text{grad } g(x^0, y^0) = (0, 0)^T$ heißt **singulärer Punkt**.

Lemma:

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

$$g_x(\mathbf{x}^0) = 0, \quad g_y(\mathbf{x}^0) \neq 0$$

so besitzt die Lösungskurve eine **horizontale Tangente** in \mathbf{x}^0 .

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

$$g_x(\mathbf{x}^0) \neq 0, \quad g_y(\mathbf{x}^0) = 0$$

so besitzt die Lösungskurve eine **vertikale Tangente** in \mathbf{x}^0 .

3) Ist \mathbf{x}^0 ein **singulärer Punkt** so wird die Lösungsmenge bei \mathbf{x}^0 durch folgende **quadratische Gleichung** approximiert:

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

Wegen 3) erhält man für $g_{xx}, g_{xy}, g_{yy} \neq 0^T$:

$\det Hg(x^0) > 0$: 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** oder auch **Spitze**

Interpretation:

1) 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 $g(x)$.

2) 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 $g(x)$.

3) Gilt $\det Hg(x^0) = 0$, so ist der stationäre Punkt x^0 von $g(x)$ **ausgeartet**.

Beispiele: Wir betrachten jeweils den singulären Punkt \mathbf{x}^0 :

1) Gegeben sei die implizite Gleichung

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

Berechnung der partiellen Ableitungen bis zur Ordnung 2:

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

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

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

$$g_{xy} = 2y$$

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

$$\mathbf{H}g(\mathbf{0}) = \begin{pmatrix} -4 & 0 \\ 0 & -2 \end{pmatrix}$$

Also ist $\mathbf{x}^0 = \mathbf{0}$ ein isolierter Punkt.

Beispiele: (Fortsetzung)

2) Gegeben sei die implizite Gleichung

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

Berechnung der partiellen Ableitungen bis zur Ordnung 2:

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

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

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

$$g_{xy} = 2y$$

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

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

Also ist $\mathbf{x}^0 = \mathbf{0}$ für $q \neq 0$ ein Doppelpunkt.

Beispiele: (Fortsetzung)

3) Gegeben sei die implizite Gleichung

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

Berechnung der partiellen Ableitungen bis zur Ordnung 2:

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

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

$$g_{xx} = 6x$$

$$g_{xy} = 2y$$

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

$$\mathbf{H}g(\mathbf{0}) = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}$$

Also ist $\mathbf{x}^0 = \mathbf{0}$ ein Rückkehrpunkt.

Implizite Darstellung von Flächen

Lösungsmenge einer skalaren Gleichung $g(x, y, z) = 0$ ist für $\text{grad } g \neq 0^T$ lokal eine Fläche im \mathbb{R}^3 .

Tangentialebene in \mathbf{x}^0 mit $g(\mathbf{x}^0) = 0$ und $\text{grad } g(\mathbf{x}^0) \neq 0^T$:

$$g_x(\mathbf{x}^0)(x - x^0) + g_y(\mathbf{x}^0)(y - y^0) + g_z(\mathbf{x}^0)(z - z_0) = 0$$

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

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

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

Partielle Ableitungen von $f(x, y)$:

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

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: Sei $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R}^n$ offen, eine 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 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 C^1 -Diffeomorphismus.

2.3 Extremalprobleme unter Nebenbedingungen

Frage:

Welche Abmessungen sollte eine Metalldose haben, damit bei gegebenem Volumen der Materialverbrauch am geringsten ist?

Sei r der Radius und h die Höhe. Dann gilt

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

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

Setze bei vorgebenem $c \in \mathbb{R}_+$

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

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

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

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

Lösung:

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

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

Allgemein:

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

$$\mathbf{g}(\mathbf{x}) = 0$$

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

Die Nebenbedingungen lauten also

$$\begin{aligned} g_1(x_1, \dots, x_n) &= 0 \\ &\vdots \\ g_m(x_1, \dots, x_n) &= 0 \end{aligned}$$

Alternativ: Bestimme die Extremwerte der Funktion $f(\mathbf{x})$ auf der Menge

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

Die Lagrange–Funktion:

Wir definieren folgende erweiterte Funktion $F(\mathbf{x})$:

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

und suchen die Extremwerte von $F(\mathbf{x})$ 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) \mathbf{x}^0 die Lagrange–Funktion $F(\mathbf{x})$ (für ein festes λ) über D und gilt $\mathbf{g}(\mathbf{x}^0) = 0$, so liefert \mathbf{x}^0 zugleich das Minimum (bzw. Maximum) von $f(\mathbf{x})$ über $G := \{\mathbf{x} \in D : \mathbf{g}(\mathbf{x}) = 0\}$.

Beweis: Für ein beliebiges $\mathbf{x} \in D$ gilt nach Voraussetzung

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

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

Bemerkung:

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

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

Zusammen mit den Nebenbedingungen $\mathbf{g}(\mathbf{x}) = 0$ ergibt sich ein (nichtlineares) Gleichungssystem mit $(n + m)$ Gleichungen und $(n + m)$ Unbekannten \mathbf{x} und λ .

Die Lösungen $(\mathbf{x}^0, \lambda^0)$ sind die Kandidaten für die gesuchten Extremstellen (**Notwendige Bedingung**).

Alternativ: Definiere eine Lagrange-Funktion

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

und suche die Extremstellen von $G(\mathbf{x}, \lambda)$ bezüglich \mathbf{x} und λ .

Bemerkung:

Man kann auch eine **hinreichende** Bedingung aufstellen:

Sind die Funktionen f und g sogar C^2 -Funktionen und ist die Hesse-Matrix $H_F(x^0)$ der Lagrange-Funktion positiv (bzw. negativ) definit, so ist x^0 tatsächlich ein strenges lokales Minimum (bzw. Maximum) von $f(x)$ auf G .

In den meisten Anwendungen ist die hinreichende Bedingung allerdings **nicht** erfüllt, obwohl x^0 ein strenges lokales Extremum **ist**.

Insbesondere kann man aus der Indefinitheit der Hesse-Matrix $H_F(x^0)$ **nicht** schließen, dass x^0 kein Extremwert ist.

Ähnlich problematisch ist die hinreichende Bedingung, die man aus der Hesse-Matrix für die Lagrange-Funktion $G(x, \lambda)$ bezüglich x **und** λ erhält.

Beispiel:

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

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

Da die betrachte Funktion stetig ist 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

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

Dies ist eine offene Menge und die notwendige Bedingung für einen Extremwert lautet

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

Also erhalten wir als Kandidaten den Ursprung $\mathbf{x}^0 = 0$.

Beispiel: (Fortsetzung)

Die Hesse-Matrix im Punkt $x^0 = 0$ lautet

$$H^f(\mathbf{0}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

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

Die Extrema der Funktion müssen also auf dem Rand liegen, der ein 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

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

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

$$\lambda = \frac{1}{2} \quad : \quad \mathbf{x}^{(1)} = (\sqrt{0.5}, -\sqrt{0.5})^T \quad \mathbf{x}^{(2)} = (-\sqrt{0.5}, \sqrt{0.5})^T$$

$$\lambda = -\frac{1}{2} \quad : \quad \mathbf{x}^{(3)} = (\sqrt{0.5}, \sqrt{0.5})^T \quad \mathbf{x}^{(4)} = (-\sqrt{0.5}, -\sqrt{0.5})^T$$

Minima und Maxima lassen sich nun einfach aus den Funktionswerten ablesen

$$f(\mathbf{x}^{(1)}) = f(\mathbf{x}^{(2)}) = -0.5 \quad f(\mathbf{x}^{(3)}) = f(\mathbf{x}^{(4)}) = 0.5$$

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

Satz: (Lagrange–Multiplikatoren–Regel)

Seien $f, g_1, \dots, g_m : D \rightarrow \mathbb{R}$ jeweils C^1 –Funktionen, und sei $\mathbf{x}^0 \in D$ ein lokales Extremum von $f(\mathbf{x})$ unter der Nebenbedingung $\mathbf{g}(\mathbf{x}) = 0$.

Ferner gelte die **Regularitätsbedingung**

$$\text{Rang}(\mathbf{J} \mathbf{g}(\mathbf{x}^0)) = m$$

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

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

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

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

Bemerkung:

1) Notwendige Bedingung zweiter Ordnung

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

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

d.h., es gilt

$$y^T \mathbf{H}F(\mathbf{x}^0) y \geq 0 \quad \forall y \in TG(\mathbf{x}^0)$$

Bemerkung: (Fortsetzung)

2) Hinreichende Bedingung

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

$$\mathbf{y}^T \mathbf{H}F(\mathbf{x}^0) \mathbf{y} > 0 \quad \forall \mathbf{y} \in TG(\mathbf{x}^0)$$

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

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

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

Die notwendige Bedingung ergibt das nichtlineare System

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

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

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

Beispiel: (Fortsetzung)

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

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

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

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

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

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

Beispiel: Man bestimme die lokalen Extremwerte der Funktion

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

auf dem Durchschnitt des Zylinders

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

mit der Ebene

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

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

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

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

Beispiel: (Fortsetzung)

Die Jacobi-Matrix

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

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

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

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

$$2 + \lambda_2 = 0$$

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

$$x + z = 1$$

Beispiel: (Fortsetzung)

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

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

$$2 + \lambda_2 = 0$$

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

$$x + z = 1$$

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

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

Die möglichen Extremwerte sind also

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

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

2.4 Das Newton–Verfahren für nichtlineare Gleichungssysteme

Problem: Wir suchen die Nullstellen einer Funktion $f : D \rightarrow \mathbb{R}^n$ mit

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

$$f(x) = 0$$

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: Verfahren ist **ableitungsfrei**

Nachteile:

- Numerisches Verfahren konvergiert langsam (linear),
- Keine eindeutige Iterationsvorschrift.

Alternative: Newton–Verfahren für vektorwertige Funktionen

Gegeben sei eine C^1 –Funktion $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R}^n$ offen. Wir suchen eine Nullstelle, d.h. ein $\mathbf{x}^* \in D$ mit

$$f(\mathbf{x}^*) = \mathbf{0}$$

Taylor–Entwicklung von $f(\mathbf{x})$ um einen Startwert \mathbf{x}^0 lautet

$$f(\mathbf{x}) = f(\mathbf{x}^0) + \mathbf{J}f(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0) + o(\|\mathbf{x} - \mathbf{x}^0\|)$$

Setzen wir $\mathbf{x} = \mathbf{x}^*$, so folgt

$$\mathbf{J}f(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0) \approx -f(\mathbf{x}^0)$$

Eine Näherungslösung für \mathbf{x}^* ist dann \mathbf{x}^1 , die Lösung des LGS

$$\mathbf{J}f(\mathbf{x}^0)(\mathbf{x}^1 - \mathbf{x}^0) = -f(\mathbf{x}^0)$$

Das **Newton–Verfahren** lautet somit

für $k = 0, 1, 2, \dots$

$$\mathbf{J}f(\mathbf{x}^k) \cdot \Delta \mathbf{x}^k = -f(\mathbf{x}^k)$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \Delta \mathbf{x}^k$$

Man muss also im Newton–Verfahren in jedem Iterationsschritt ein LGS lösen. Die Lösung $\Delta \mathbf{x}^k$ heißt **Newton–Korrektur**.

Das Newton–Verfahren ist **skalierungsinvariant**:

Satz: Das Newton–Verfahren ist invariant unter allen Transformationen (Skalierungen) der Form

$$f(\mathbf{x}) \rightarrow \mathbf{g}(\mathbf{x}) = A\mathbf{f}(\mathbf{x})$$

mit einer regulären Matrix $A \in \mathbb{R}^{(n,n)}$, d.h. die Iterierten für f und \mathbf{g} sind identisch.

Satz: (Konvergenzsatz)

Sei $f : D \rightarrow \mathbb{R}^n$ eine C^1 -Funktion, $D \subset \mathbb{R}^n$ offen und konvex.

Sei $\mathbf{x}^* \in D$ mit $f(\mathbf{x}^*) = 0$. Die Jacobi-Matrix $\mathbf{J}f(\mathbf{x})$ sei regulär für $\mathbf{x} \in D$, und es gelte eine Lipschitz-Bedingung

$$\|(\mathbf{J}f(\mathbf{x}))^{-1}(\mathbf{J}f(\mathbf{y}) - \mathbf{J}f(\mathbf{x}))\| \leq L\|\mathbf{y} - \mathbf{x}\|$$

für alle $\mathbf{x}, \mathbf{y} \in D$ mit $L > 0$.

Für alle Startwerte $\mathbf{x}^0 \in D$ mit

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

ist dann das Newton-Verfahren wohldefiniert mit $\mathbf{x}^k \in K_r(\mathbf{x}^*)$, $k = 0, 1, 2, \dots$, und die Newton-Iterierten \mathbf{x}^k konvergieren **quadratisch** gegen \mathbf{x}^* :

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

Ferner ist \mathbf{x}^* die einzige Nullstelle von $f(\mathbf{x})$ innerhalb $K_r(\mathbf{x}^*)$.

Das gedämpfte Newton–Verfahren

Das Newton–Verfahren konvergiert zwar quadratisch, aber nur **lokal**.

Globale Konvergenz kann durch einen Dämpfungsterm verbessert werden:

für $k = 0, 1, 2, \dots$

$$\mathbf{J}f(\mathbf{x}^k) \cdot \Delta \mathbf{x}^k = -f(\mathbf{x}^k)$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda_k \Delta \mathbf{x}^k$$

Gedämpfte Newton–Verfahren mit **Dämpfungsparameter** $\lambda_k \in (0, 1]$.

Frage: Wie wählt man die Dämpfungsfaktoren λ_k ?

Verwende eine **Testfunktion** $T(\mathbf{x})$ mit

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

$$T(\mathbf{x}) = 0 \Leftrightarrow f(\mathbf{x}) = 0$$

sodass die Folge $T(\mathbf{x}^k)$ monoton fällt.

Bemerkung:

Befindet man sich beim gedämpften Newton–Verfahren *in der Nähe* der gesuchten Lösung \mathbf{x}^* , so sollte $\lambda_k = 1$ gewählt werden. Nur das sichert die (lokale) quadratische Konvergenz.

Beispiel:

Bei praktischen Anwendungen verwendet man häufig die Testfunktion

$$T(\mathbf{x}) = \|\mathbf{f}(\mathbf{x})\|_2^2$$

Für diesen Fall gilt der folgende Satz.

Satz:

Sei \mathbf{f} eine C^1 –Funktion auf der offenen und konvexen Menge $D \subset \mathbb{R}^n$.

Für $\mathbf{x}^k \in D$ mit $\mathbf{f}(\mathbf{x}^k) \neq 0$ gilt dann:

$$\exists \mu_k : \forall \lambda \in (0, \mu_k) : \|\mathbf{f}(\mathbf{x}^k + \lambda \Delta x^k)\|_2^2 < \|\mathbf{f}(\mathbf{x}^k)\|_2^2$$

Dämpfungsstrategie:

$k = 0$: Man wähle $\lambda_0 \in \{1, \frac{1}{2}, \frac{1}{4}, \dots, \lambda_{min}\}$ möglichst groß,

so dass gilt:

$$T(\mathbf{x}^0 + \lambda_0 \Delta \mathbf{x}^0) < T(\mathbf{x}^0)$$

$k > 0$: $\lambda_k := \lambda_{k-1}$

falls $T(\mathbf{x}^k + \lambda_k \Delta \mathbf{x}^k) < T(\mathbf{x}^k)$:

$$\mathbf{x}^{k+1} := \mathbf{x}^k + \lambda_k \Delta \mathbf{x}^k$$

$$\lambda_k := \begin{cases} \lambda_k & : \text{falls } \lambda_k = 1 \\ 2\lambda_k & : \text{sonst} \end{cases}$$

Dämpfungsstrategie: (Fortsetzung)

sonst:

$$\mu = \max \left\{ \frac{\lambda_k}{2}, \frac{\lambda_k}{4}, \dots, \lambda_{min} \right\}$$

so dass $T(\mathbf{x}^k + \mu \Delta \mathbf{x}^k) < T(\mathbf{x}^k)$

$$\lambda_k := \mu$$

Bemerkung: Die Testfunktion ist **nicht** skalierungsinvariant.

Daher ist der sogenannte **natürliche Monotonietest** sinnvoll:

$$\| \mathbf{Jf}(\mathbf{x}^k)^{-1} \mathbf{f}(\mathbf{x}^{k+1}) \| < \| \mathbf{Jf}(\mathbf{x}^k)^{-1} \mathbf{f}(\mathbf{x}^k) \|$$

Es gilt dabei:

$$\| \Delta \mathbf{x}^k \| = \| \mathbf{Jf}(\mathbf{x}^k)^{-1} \mathbf{f}(\mathbf{x}^k) \|$$

Kapitel 3: Integralrechnung mehrerer Variablen

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 $f(\mathbf{x})$:

$$V = \int_D f(\mathbf{x}) dx$$

Erinnerung Analysis II:

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

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

Das Integral war als Grenzwert von Ober- und Untersumme, für den Fall, dass die Feinheit der Zerlegung gegen Null strebt und ein gemeinsamer Grenzwert existiert, definiert.

Konstruktionsprinzip des Integrals mehrerer Veränderlichen analog,
aber der Definitionsbereich D ist komplizierter.

Zunächst betrachten wir $n = 2$ und einen Definitionsbereich D der Form

$$D = [a_1, b - 1] \times [a_2, b_2] \subset \mathbb{R}^2$$

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

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

Definition:

- 1) 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 gelten

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

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

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

Definition: (Fortsetzung)

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

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

3) 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

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

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

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

eine **Riemannsche Summe** zur Zerlegung Z .

Definition: (Fortsetzung)

5) Analog zum Integral einer Variablen heißen

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

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

die **Riemannsche Unter- bzw. Obersumme** der Funktion $f(\mathbf{x})$ zur Zerlegung Z .

Bemerkung:

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

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

Bemerkung:

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

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

Für zwei beliebige Zerlegungen Z_1 und Z_2 gilt stets:

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

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

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

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

Die beiden Werte U_f und O_f existieren, da Unter- und Obersumme geeignete Monotonieeigenschaften besitzen.

Definition:

- 1) **Riemannsches Unter- bzw. Oberintegral** der Funktion $f(x)$ über D :

$$\int_D^* f(x) dx := \sup\{U_f(Z) : Z \in \mathbf{Z}(D)\}$$

$$\int_D^* f(x) dx := \inf\{O_f(Z) : Z \in \mathbf{Z}(D)\}$$

- 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

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

Bemerkung:

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

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

betrachtet.

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

Schreibweise für $n = 2$ und $n = 3$:

$$\int_D f(x, y) dx dy \qquad \int_D f(x, y, z) dx dy dz$$

oder auch

$$\iint_D f(x, y) dx dy \qquad \iiint_D f(x, y, z) dx dy dz$$

Satz:

1) Linearität

$$\int_D (\alpha f(\mathbf{x}) + \beta g(\mathbf{x})) d\mathbf{x} = \alpha \int_D f(\mathbf{x}) d\mathbf{x} + \beta \int_D g(\mathbf{x}) d\mathbf{x}$$

2) Monotonie

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

$$\int_D f(\mathbf{x}) d\mathbf{x} \leq \int_D g(\mathbf{x}) d\mathbf{x}$$

3) Positivität

Gilt für alle $\mathbf{x} \in D$ die Beziehung $f(\mathbf{x}) \geq 0$, so folgt:

$$\int_D f(\mathbf{x}) d\mathbf{x} \geq 0$$

Satz: (Fortsetzung)

- 4) 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 $f(\mathbf{x})$ genau dann über D integrierbar, falls $f(\mathbf{x})$ über D_1 und D_2 integrierbar ist, und es gilt:

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

- 5) Es gilt folgende **Abschätzung** für das Integral

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

- 6) **Riemannsches Kriterium:**

$f(\mathbf{x})$ ist genau dann über D integrierbar, falls gilt:

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

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

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

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

$$\begin{aligned} \int_D f(\mathbf{x}) d\mathbf{x} &= \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y) dy dx && \text{bzw.} \\ \int_D f(\mathbf{x}) d\mathbf{x} &= \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x, y) dx dy \end{aligned}$$

Bedeutung: Rückführung auf eindimensionale Integration möglich

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

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

Stetige Funktionen sind über Quadern integrierbar (kommt gleich), daher können wir den Satz von Fubini anwenden:

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

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

Definition:

Sei $D \subset \mathbb{R}^n$ eine kompakte Menge und $f : D \rightarrow \mathbb{R}$ beschränkt.
Wir setzen

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

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

Sei nun Q der kleinste Quader mit $D \subset Q$. Dann definieren wir:

- 1) Die Funktion $f(\mathbf{x})$ heißt **integrierbar** über D , falls $f^*(\mathbf{x})$ über Q integrierbar ist. In diesem Fall setzen wir

$$\int_D f(\mathbf{x}) d\mathbf{x} := \int_Q f^*(\mathbf{x}) d\mathbf{x}$$

Definition: (Fortsetzung)

2) Die kompakte Menge heißt **messbar**, falls das Integral

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

existiert.

Man nennt dann $\text{vol}(D)$ das **Volumen** von D .

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 der Integrationsbegriff von oben stimmt mit dem vorhergehend diskutierten überein.

Das durch $\text{vol}(D)$ bezeichnete Volumen ist dann auch das tatsächliche Volumen des Quaders (im \mathbb{R}^n).

Wir fassen drei wichtige Eigenschaften der mehrdimensionalen Integration zusammen:

1) **Satz:**

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

2) **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 .

3) **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

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

Definition:

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

$$D = \{(x, y) : a \leq x \leq b \wedge g(x) \leq y \leq h(x)\}$$

bzw.

$$D = \{(x, y) : \bar{a} \leq y \leq \bar{b} \wedge \bar{g}(y) \leq x \leq \bar{h}(y)\}$$

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

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

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

Definition: (Fortsetzung)

- 3) 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

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

Bemerkung:

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

Satz:

Ist $f(\mathbf{x})$ stetig auf einem Normalbereich

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

so gilt

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

Bemerkung:

Analoge Beziehungen gelten für höhere Dimensionen. Ist $D \subset \mathbb{R}^n$ eine projizierbare Menge, so gilt

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

Beispiel: Gegeben sei die Funktion

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

Berechne das Integral über der durch zwei Parabeln begrenzten Fläche

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

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

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

$$= \int_{-1}^1 (-2x^3 - 4x^2 + 2x + 4) dx = \frac{16}{3}$$

Beispiel: Zu berechnen ist das Volumen des Rotationsparaboloids:
ids:

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

Darstellung von V als **Normalbereich**

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

Damit gilt:

$$\begin{aligned} \text{vol}(V) &= \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{x^2+y^2}^1 dz dy dx = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (1-x^2-y^2) dy dx \end{aligned}$$

$$= \dots = \frac{4}{3} \int_{-1}^1 (1 - x^2)^{3/2} dx = \frac{\pi}{2}$$

Integration über allgemeine Integrationsbereiche

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

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

gelten.

Ferner heißt

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

der **Durchmesser** der Menge D_j und

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

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

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

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

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:

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

Anwendungen der Bereichsintegrale auf die Berechnung von Schwerpunkten oder Trägheitsmomenten von Flächen und Körpern.

A) Schwerpunkt einer Fläche oder eines Körpers:

Definition:

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

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

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

Beispiel: Zu berechnen ist der Schwerpunkt der Pyramide P :

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

Unter der Annahme konstanter Dichte berechnen wir das Volumen von P :

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

Beispiel: (Fortsetzung)

Weiter 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 $\mathbf{x}_s = (\frac{3}{4}h, 0, 0)^T$.

B) Trägheitsmomente von Flächen oder Körpern:

Definition: (Trägheitsmoment bezüglich einer Achse)

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

Dann besitzt D bezüglich dieser Achse das Trägheitsmoment

$$\ominus \text{Achse} := \int_D \rho(\mathbf{x}) r^2(\mathbf{x}) d\mathbf{x}$$

Beispiel: Gegeben sei der homogene Zylinder Z :

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

Wir berechnen das Trägheitsmoment bezüglich der x -Achse:

$$\ominus_{x\text{-Achse}} = \int_Z \rho(y^2 + z^2) d(x, y, z)$$

unter der Annahme konstanter Dichte ρ .

Beispiel: (Fortsetzung)

Es gilt:

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

Transformationssatz:

Dies verallgemeinert die Substitutionsregel aus Analysis II

Satz:

Sei $\Phi : U \rightarrow \mathbb{R}^n$, $U \subset \mathbb{R}^n$ offen, eine C^1 -Abbildung. $D \subset U$ sei eine kompakte, messbare Menge, so dass Φ auf D^0 einen C^1 -Diffeomorphismus bildet.

Dann ist auch $\Phi(D)$ kompakt und messbar, und für jede stetige Funktion $f : \Phi(D) \rightarrow \mathbb{R}$ gilt:

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

Bemerkung:

Man beachte, dass im Transformationssatz die Bijektivität von Φ nur

auf dem inneren Bereich D^0 gefordert wird – nicht jedoch auf dem Rand!

Beispiel: Berechne den Schwerpunkt eines homogenen Kugeloktanten:

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

Hier ist es 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 $\Phi(D) = V$.

Weiter ist Φ auf der offenen Menge D^0 ein C^1 -Diffeomorphismus

und

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

Beispiel: (Fortsetzung)

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 $y_s = z_s = \frac{3}{8}$.

Beispiel: Der Steinersche 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 \mathbf{x}_s des Körpers K und d der Abstand des Schwerpunktes \mathbf{x}_s von der Achse A .

Idee zur Herleitung: Setze $\mathbf{x} := \Phi(\mathbf{u}) = \mathbf{x}_s + \mathbf{u}$. Dann gilt:

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

wobei

$$D := \{\mathbf{x} - \mathbf{x}_s : \mathbf{x} \in K\}$$

3.2 Kurvenintegrale

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

1. Art definiert durch

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

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

Erweiterung auf Kurvenintegrale über vektorwertige Funktionen, d.h.

$$\int_c \mathbf{f}(\mathbf{x}) dx := ?$$

Anwendung und Interpretation:

Ein Massenpunkt bewegt sich entlang $c(t)$ in einem Kraftfeld $\mathbf{f}(\mathbf{x})$.

Welche **Arbeit** muss entlang der Kurve geleistet werden?

Definition: Für ein stetiges Vektorfeld $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R}^n$ offen, und eine stückweise C^1 -Kurve $c : [a, b] \rightarrow D$ definieren wir das **Kurvenintegral 2. Art** durch

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

Herleitung:

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

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

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

Dann gilt für die in einem Kraftfeld $f(\mathbf{x})$ entlang der Kurve $c(t)$

geleistete Arbeit die Näherungsformel:

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

Daraus folgt:

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

Für eine Folge von Zerlegungen Z mit $\|Z\| \rightarrow 0$ konvergiert die linke Seite gegen das oben definierte **Kurvenintegral 2. Art**.

Bemerkung:

Für eine geschlossene Kurve $\mathbf{c}(t)$, d.h. $\mathbf{c}(a) = \mathbf{c}(b)$, schreibt man das Kurvenintegral auch als

$$\oint_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \, d\mathbf{x}$$

Eigenschaften des Kurvenintegrals 2. Art:

1) Linearität:

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

2) Es gilt:

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

wobei $(-c)(t) := c(b + a - t), a \leq t \leq b$

3) Es gilt:

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

wobei der Endpunkt von c_1 der Anfangspunkt von c_2 ist.

4) Das Kurvenintegral 2. Art ist **parametrisierungsinvariant**.

5) Es gilt:

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

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

6) Formale Schreibweise:

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

mit

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

Beispiel: Für $\mathbf{x} \in \mathbb{R}^3$ sei

$$\mathbf{f}(\mathbf{x}) := (-y, x, z^2)^T$$

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

Dann berechnet man

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

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

Beispiel:

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

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

Definition:

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

$$\forall c : \oint_c f(\mathbf{x}) dx = 0$$

Bemerkung: Ein Vektorfeld ist genau dann wirbelfrei, wenn der Wert des Kurvenintegrals $\int_c f(\mathbf{x}) dx$ nur vom Anfangs- und Endpunkt des Weges, jedoch nicht vom konkreten Verlauf der Kurve c abhängt.

Man sagt: das Kurvenintegral ist **wegunabhängig**.

Frage:

Welche Kriterien für das Vektorfeld $f(\mathbf{x})$ **garantieren** die Wegunabhängigkeit des Kurvenintegrals?

Definition:

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

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

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 zusammenhängende Mengen – in (zumindest) zwei disjunkte offene Bereiche trennbar.

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 C^1 -Funktion $\varphi : D \rightarrow \mathbb{R}$ gibt mit

$$f(\mathbf{x}) = \nabla \varphi(\mathbf{x})$$

Die Funktion $\varphi(\mathbf{x})$ heißt dann **Stammfunktion** oder **Potential** von $f(\mathbf{x})$.

Bemerkung:

Ein Massenpunkt bewege sich in einem **konservativen** Kraftfeld $\mathbf{K}(\mathbf{x})$,
d.h. \mathbf{K} besitzt ein Potential $\varphi(\mathbf{x})$.

Dann ist $U(\mathbf{x}) = -\varphi(\mathbf{x})$ gerade die **potentielle Energie**:

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

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

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

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 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

C^1 -Kurve in D , so ist $\varphi(\mathbf{x})$ gegeben durch:

$$\varphi(\mathbf{x}) = \int_{c_x} \mathbf{f}(\mathbf{x}) d\mathbf{x} + c, \quad c = \text{const.}$$

Beispiel:

Das zentrale Kraftfeld

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

besitzt das Potential

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

Denn es gilt:

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

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

$$A = \int_{\mathbf{c}} \mathbf{K}(\mathbf{x}) d\mathbf{x} = \left(\frac{1}{\|\mathbf{c}(a)\|} - \frac{1}{\|\mathbf{c}(b)\|} \right)$$

Beispiel:

Das Vektorfeld

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

besitzt das Potential

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

Für eine beliebige, die Punkte $P = (1, 1, 2)$ und $Q = (3, 5, -2)$ verbindende C^1 -Kurve $c(t)$ gilt also:

$$\int_c \mathbf{f}(\mathbf{x}) \, d\mathbf{x} = \varphi(Q) - \varphi(P) = -9 - 15 = -24$$

Interpretiert man $\mathbf{f}(\mathbf{x})$ als elektrisches Feld, so gibt das Kurvenintegral 2. Art die Spannung zwischen den beiden Punkten P und Q an.

Beispiel: Wir betrachten das Vektorfeld

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

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

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

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

Bemerkung:

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

$$\operatorname{rot} f(\mathbf{x}) = \operatorname{rot} (\nabla \varphi(\mathbf{x})) = 0 \quad \forall \mathbf{x} \in D$$

Daraus folgt, dass $\operatorname{rot} f(\mathbf{x}) = 0$ 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

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

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

Beispiel:

Wir betrachten wieder das Vektorfeld

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

Berechnet man die Rotation, so ergibt sich

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

Die Rotation von $f(x, y)$ verschwindet zwar, aber $f(x, y)$ besitzt auf der

Menge $D = \mathbb{R}^2 \setminus \{0\}$ kein Potential.

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

Satz: (Integralsatz von Green)

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

Die Parametrisierung 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**.

Andere Darstellungen des Integralsatzes von Green:

Wir hatten gesehen, dass die Beziehung

$$\oint_C \mathbf{f}(\mathbf{x}) d\mathbf{x} = \oint_C \langle \mathbf{f}, \mathbf{T} \rangle ds$$

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

Daraus folgt aber mit dem Integralsatz von Green

$$\int_{\bar{K}} \operatorname{rot} \mathbf{f}(\mathbf{x}) d\mathbf{x} = \oint_{\partial K} \langle \mathbf{f}, \mathbf{T} \rangle ds$$

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

$$\oint_C \mathbf{f}(\mathbf{x}) d\mathbf{x}$$

ist gerade die Zirkulation von $\mathbf{f}(\mathbf{x})$.

Andere Darstellungen des Integralsatzes von Green II:

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

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

und damit die Beziehung

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

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

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

$$\oint_c \mathbf{f}(\mathbf{x}) \, d\mathbf{x} = 0$$

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 D auf einen Punkt in D zusammenziehen lässt, d.h. genauer:

Es gibt eine stetige Abbildung

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

mit $\Phi(t, 0) = c(t)$, $\forall t \in [a, b]$ und $\Phi(t, 1) = \mathbf{x}^0 \in D$, $\forall t \in [a, b]$.

Die Abbildung $\Phi(t, s)$ heißt auch **Homotopie**.

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**

$$\forall \mathbf{x} \in D : \mathbf{J}f(\mathbf{x}) = (\mathbf{J}f(\mathbf{x}))^T$$

erfüllt ist.

Ausgeschrieben bedeutet dies

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

Bemerkung:

In den Fällen $n = 2, 3$ stimmt die Integrabilitätsbedingung mit der Bedingung

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

überein.

Beispiel:

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

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

gegeben.

Wir wollen untersuchen, ob $\mathbf{f}(\mathbf{x})$ ein Potential besitzt und dieses gegebenenfalls bestimmen.

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

Weiter gilt:

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

Also besitzt $f(x)$ ein Potential.

Berechnung des Potentials:

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

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

Durch Integration bezüglich der Variablen x ergibt sich:

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

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

Einsetzen in die Gleichung

$$\frac{\partial\varphi}{\partial y} = f_2 = \ln r^2 + \frac{2y^2}{r^2} + ze^y$$

liefert

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

Daraus folgt die Bedingung

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

und es gilt

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

Wir haben damit:

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

mit der noch unbekanntem Funktion $d(z)$.

Die letzte Bedingung lautet

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

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

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

3.3 Oberflächenintegrale

Definition:

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

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

Sind für alle $\mathbf{u} \in D$ die beiden Vektoren

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

linear unabhängig, so heißt

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

eine **Fläche** bzw. ein **Flächenstück**.

Die Abbildung $\mathbf{x} = p(\mathbf{u})$ nennt man eine **Parametrisierung** oder

Parameterdarstellung der Fläche F .

Beispiel:

Wir betrachten für gegebenes $r > 0$ die Abbildung

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

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 sind

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

und damit linear unabhängig auf ganz \mathbb{R}^2 .

Beispiel:

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

Eine Parametrisierung ist etwa gegeben durch

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

Die partiellen Ableitungen

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

sind linear unabhängig.

Tangentialebene:

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

Die Tangentialebene hat die Parameterdarstellung

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

Frage:

Wie kann ich den Flächeninhalt einer gegebenen Fläche F berechnen?

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

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

Dabei nennt man den Term

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

auch das **Oberflächenelement** der Fläche $x = p(\mathbf{u})$.

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

Beispiel:

Für die Mantelfläche des Zylinders $Z = \mathbf{p}(K)$ mit

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

und

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

erhält man wegen

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

den Wert

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

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

$$O(\mathbf{p}(K)) = \int_{\mathbf{p}(K)} do$$

$$= \int_K \sqrt{1 + \varphi_{x_1}^2 + \varphi_{x_2}^2} \, d(x_1, x_2)$$

Beispiel:

Wir berechnen die Oberfläche des Paraboloids P gegeben durch

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

Dann gilt:

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

Bemerkung:

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

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

Daraus folgt

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

Definiert man

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

so ergibt sich die Beziehung

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

Beispiel:

Für das Oberflächenelement der Sphäre

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

ergeben sich mit der Parametrisierung über Kugelkoordinaten

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

die Beziehungen

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

Daraus folgt

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

Beispiel: (Fortsetzung)

Aus der Beziehung

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

folgt daher

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

Wir können nun die Oberfläche der Kugel berechnen:

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

Definition:

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

- 1) Für eine stetige Funktion $f : F \rightarrow \mathbb{R}$ definiert man das

Oberflächenintegral 1. Art durch

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

- 2) Für ein stetiges Vektorfeld $\mathbf{f} : F \rightarrow \mathbb{R}^3$ definiert man das

Oberflächenintegral 2. Art durch

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

Anderer Darstellungen des Oberflächenintegrals 2. Art

Der Einheitsnormalenvektor $\mathbf{n}(\mathbf{x})$ auf der Fläche F ist gegeben durch

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

Wir schreiben daher auch

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

$$= \int_F \langle \mathbf{f}(\mathbf{x}), \mathbf{n}(\mathbf{x}) \rangle d\sigma$$

Bemerkung:

1) Physikalische Interpretation der Oberflächenintegrale:

Ist $\rho(\mathbf{x})$ die Dichte einer massenbelegten Fläche, so liefert das Integral 1. Art gerade die Gesamtmasse der Fläche.

Ist $f(\mathbf{x})$ 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 $f(\mathbf{x})$ durch die Fläche F .

2) Ist F eine geschlossene Fläche, d.h. die Oberflächen eines kompakten und einfach zusammenhängenden Körpers im \mathbb{R}^3 , so schreiben wir wiederum

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

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

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 $\mathbf{n}(\mathbf{x})$.

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

$$\int_G \operatorname{div} \mathbf{f}(\mathbf{x}) \, d\mathbf{x} = \int_{\partial G} \mathbf{f}(\mathbf{x}) \, d\mathbf{o}$$

Interpretation:

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

$$\int_{\partial G} \mathbf{f}(\mathbf{x}) \, d\mathbf{o} = 0$$

Beispiel:

Wir betrachten das Vektorfeld $f(\mathbf{x}) = \mathbf{x}$ und die Kugel K :

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

Dann gilt offensichtlich

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

und damit

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

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

Satz: (Formeln von Green)

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

$$\begin{aligned} \int_G (f \Delta g + \langle \nabla f, \nabla g \rangle) dx &= \int_{\partial G} f \frac{\partial g}{\partial \mathbf{n}} do \\ \int_G (f \Delta g - g \Delta f) dx &= \int_{\partial G} \left(f \frac{\partial g}{\partial \mathbf{n}} - g \frac{\partial f}{\partial \mathbf{n}} \right) do \end{aligned}$$

Hierbei bezeichnet

$$\frac{\partial f}{\partial \mathbf{n}}(\mathbf{x}) = D_{\mathbf{n}} f(\mathbf{x}), \quad \mathbf{x} \in \partial G$$

die Richtungsableitung von $f(\mathbf{x})$ in Richtung des äußeren Einheitsnormalenvektors $\mathbf{n}(\mathbf{x})$.

Beweis: Wir setzen

$$\mathbf{F}(\mathbf{x}) = f(\mathbf{x}) \cdot \nabla g(\mathbf{x})$$

Dann gilt

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

Wir wenden nun den Gaußschen Integralsatz an:

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

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

Satz: (Integralsatz von Stokes)

Sei $f : D \rightarrow \mathbb{R}^3$ ein 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 $\mathbf{x} = p(\mathbf{u})$, $\mathbf{u} \in \mathbb{R}^2$.

$K \subset \mathbb{R}^2$ sei ein Greenscher Bereich. Der Rand ∂K werde durch eine stückweise glatte 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 $\mathbf{n}(\tilde{c}(t)) \times \dot{\tilde{c}}(t)$ in Richtung der Fläche weist.

Dann gilt

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

Beispiel:

Gegeben sei das Vektorfeld

$$\mathbf{f}(x, y, z) = (-y, x, -z)^T$$

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

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

Dann gilt:

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

$$= \int_0^{2\pi} (\sin^2 t + \cos^2 t) dt = 2\pi$$

Beispiel: (Fortsetzung)

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

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

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:

$$\int_F \operatorname{rot} \mathbf{f}(\mathbf{x}) \, d\mathbf{o} = \int_{c=\partial F} \mathbf{f}(\mathbf{x}) \, d\mathbf{x}$$

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

$$\int_{c=\partial F} \mathbf{f}(\mathbf{x}) \, d\mathbf{x} = 2\pi$$

Beispiel: (Fortsetzung)

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

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

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

Man berechnet direkt, dass $\operatorname{rot} f(\mathbf{x}) = (0, 0, 2)^T$ gilt und

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

Daraus folgt:

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