

# Mathematik II für Physiker\*

Sommersemester 2013

Christian Hainzl mit Vorlage von Stefan Teufel, Simon Mayer und Mario Laux  
Mathematisches Institut  
Uni Tübingen

18. Juli 2013

\*Diese vorläufige Version des Skriptums ist nur zum Gebrauch parallel zum Besuch der Vorlesung gedacht. Das Studium des Skripts kann den Besuch der Vorlesung **nicht** ersetzen! Falls Sie Fehler finden, teilen Sie mir diese (auch die offensichtlichen) bitte mit!



# Inhaltsverzeichnis

<b>1</b>	<b>Eindimensionale Differentialrechnung</b>	<b>1</b>
1.1	Differenzierbarkeit und Ableitungsregeln	1
1.2	Der Mittelwertsatz	6
1.3	Der Satz von Taylor und Taylorreihen	7
<b>2</b>	<b>Mehrdimensionale Differenzierbarkeit</b>	<b>13</b>
2.1	Taylor Formel und lokale Extrema	26
<b>3</b>	<b>Einschub: Lineare Algebra</b>	<b>31</b>
3.1	Determinanten	31
3.2	Eigenwerte und Eigenvektoren	39
3.3	Symmetrische Matrizen und lokale Minima	45
<b>4</b>	<b>Implizite Funktionen</b>	<b>49</b>
<b>5</b>	<b>Integralrechnung</b>	<b>57</b>
5.1	Unbestimmtes Integral	57
5.2	Riemann-Integral	62
5.3	Hauptsatz der Differential- und Integralrechnung	66
5.4	Satz von Fubini, Substitutionsregel	71
<b>6</b>	<b>Kurvenintegrale</b>	<b>81</b>
<b>7</b>	<b>Oberflächenintegrale</b>	<b>99</b>



# 1 Eindimensionale Differentialrechnung

Wir beginnen mit reelwertigen Funktionen einer Variable,  $f : \mathbb{R} \rightarrow \mathbb{R}$ .

## 1.1 Differenzierbarkeit und Ableitungsregeln

### 1.1 Erinnerung. Das Tangentenproblem

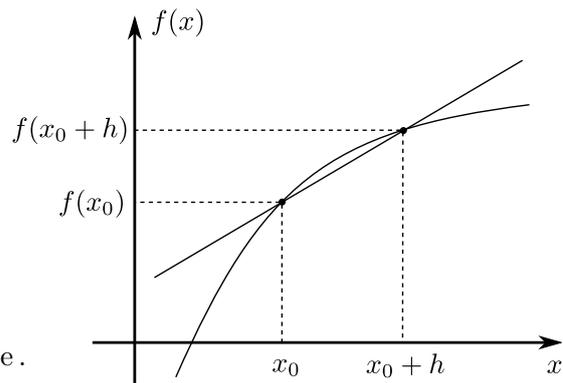
Wann hat eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  eine eindeutig bestimmte Tangente im Punkt  $x_0$  und welche Steigung hat diese?

Approximation durch Differenzenquotienten:

$$\text{Steigung der Sekante} = \frac{\Delta f}{\Delta x} := \frac{f(x_0 + h) - f(x_0)}{h}.$$

Für  $h \rightarrow 0$  erhält man den Differentialquotienten:

$$\frac{df}{dx} := \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \text{Steigung der Tangente.}$$



### 1.2 Definition. Differenzierbarkeit in einem Punkt

Sei  $I \subset \mathbb{R}$  ein Intervall. Eine Funktion  $f : I \rightarrow \mathbb{R}$  heißt im Punkt  $x_0 \in I$  **differenzierbar**, wenn der Grenzwert

$$f'(x_0) := \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \quad (*)$$

existiert. Dann heißt  $f'(x_0)$  die **Ableitung** von  $f$  an der Stelle  $x_0$ . Man schreibt dafür auch

$$\frac{d}{dx} f(x_0) \quad \text{oder} \quad \frac{df}{dx}(x_0) \quad \text{oder} \quad \left. \frac{df}{dx}(x) \right|_{x=x_0}.$$

**1.3 Bemerkung.** In (\*) ist gemeint, dass die Funktion  $x \mapsto \frac{f(x) - f(x_0)}{x - x_0}$  von  $I \setminus \{x_0\}$  nach  $\mathbb{R}$  einen Grenzwert bei  $x_0$  im Sinne von der Definition aus dem vorigen Semester hat.

Es sei erinnert, dass dies genau dann der Fall ist, wenn der Limes

$$\lim_{n \rightarrow \infty} \frac{f(x_n) - f(x_0)}{x_n - x_0}$$

für jede Folge  $(x_n)$  in  $I \setminus \{x_0\}$  mit  $\lim_{n \rightarrow \infty} x_n = x_0$  existiert.

### 1.4 Satz. Lineare Approximierbarkeit

Eine Funktion  $f : I \rightarrow \mathbb{R}$  ist in  $x_0 \in I$  genau dann differenzierbar, wenn gilt

$$(*) \quad f(x) = \underbrace{f(x_0)}_{\text{konstant}} + a \cdot \underbrace{(x - x_0)}_{\text{linear}} + \underbrace{\varphi_{x_0}(x - x_0)}_{\text{Rest}}$$

für ein  $a \in \mathbb{R}$  und eine Funktion  $\varphi_{x_0} : \{h \in \mathbb{R} \mid h + x_0 \in I\} \rightarrow \mathbb{R}$  mit

$$\lim_{h \rightarrow 0} \frac{\varphi_{x_0}(h)}{h} = 0.$$

Der „Restterm“  $\varphi_{x_0}$  geht für  $x \rightarrow x_0$  also schneller als linear gegen Null. Es gilt dann  $a = f'(x_0)$ .

## 1 Eindimensionale Differentialrechnung

*Beweis.* Ist  $f$  in  $x_0$  differenzierbar, so gilt

$$\frac{f(x) - f(x_0)}{x - x_0} \rightarrow f'(x_0) \quad \text{für } x \rightarrow x_0.$$

Für  $\varphi_{x_0}(h) := f(x_0 + h) - f(x_0) - h \cdot f'(x_0)$  gilt somit

$$\lim_{h \rightarrow 0} \frac{\varphi_{x_0}(h)}{h} = \lim_{x \rightarrow x_0} \frac{\varphi_{x_0}(x - x_0)}{x - x_0} = \left( \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \right) - f'(x_0) = 0.$$

Also gilt (\*) mit  $a = f'(x_0)$ .

Ist umgekehrt  $f$  in  $x_0$  linear approximierbar, so gilt

$$\frac{f(x) - f(x_0)}{x - x_0} = a + \frac{\varphi_{x_0}(x - x_0)}{x - x_0} \rightarrow a \quad \text{für } x \rightarrow x_0,$$

d.h.  $f$  ist differenzierbar bei  $x_0$  mit  $f'(x_0) = a$ . □

### 1.5 Definition. Differenzierbare Funktionen

Eine Funktion  $f : I \rightarrow \mathbb{R}$  heißt differenzierbar, falls sie in jedem Punkt  $x_0 \in I$  differenzierbar ist.

### 1.6 Satz. Differenzierbare Funktionen sind stetig

Ist  $f : I \rightarrow \mathbb{R}$  differenzierbar in  $x_0$ , so ist  $f$  stetig in  $x_0$ .

*Beweis.* Aus der linearen Approximierbarkeit folgt

$$\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} (f(x_0) + a(x - x_0) + \varphi_{x_0}(x - x_0)) = f(x_0).$$

Also ist  $f$  stetig in  $x_0$ . □

**1.7 Beispiele.** (a) Die konstante Funktion  $f(x) \equiv c$  hat überall die Ableitung 0, da

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{0}{x - x_0} = 0.$$

(b) Die Identität  $f(x) = x$  hat überall die Ableitung 1, da

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{x - x_0}{x - x_0} = \lim_{x \rightarrow x_0} 1 = 1.$$

(c) Die Funktion  $f(x) = x^2$  hat die Ableitung

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{x^2 - x_0^2}{x - x_0} = \lim_{x \rightarrow x_0} (x + x_0) = 2x_0.$$

### 1.8 Satz. Ableitung der Exponentialfunktion

Die Exponentialfunktion  $\exp : \mathbb{R} \rightarrow \mathbb{R}$  hat in jedem Punkt  $x \in \mathbb{R}$  die Ableitung  $\exp(x)$ .

*Beweis.* Wir zeigen zunächst  $(e^x)'|_{x=0} = \lim_{x \rightarrow 0} \frac{e^x - e^0}{x - 0} = \lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1 = e^0$  und folgern danach  $(e^x)' = e^x$  aus dem Exponentialgesetz.

Sei  $x_n$  eine Nullfolge. Dann gilt

$$\frac{e^{x_n} - 1}{x_n} = \frac{y_n}{\ln(y_n + 1)} = \frac{1}{\frac{1}{y_n} \ln(y_n + 1)} = \frac{1}{\ln((y_n + 1)^{1/y_n})},$$

mit  $y_n = e^{x_n} - 1$ , sodass  $y_n$  auch eine Nullfolge ist. Nehmen wir zuerst an, dass  $y_n \geq 0$  ist. Dann finden wir natürliche Zahlen  $k_n$ , sodass  $k_n \leq 1/y_n \leq k_n + 1$ . und

$$\left(1 + \frac{1}{k_n + 1}\right)^{k_n} \leq (y_n + 1)^{1/y_n} \leq \left(1 + \frac{1}{k_n}\right)^{k_n + 1}.$$

Da  $\left(1 + \frac{1}{k_n}\right)^{k_n} \rightarrow e$  folgt

$$\lim \frac{e^{x_n} - 1}{x_n} = \lim \frac{1}{\ln((y_n + 1)^{1/y_n})} = \frac{1}{\ln e} = 1,$$

wobei wir die Stetigkeit des Logarithmus verwendet haben. Bei nicht positiven Folgen zerlegen wir diese in positiven und negativen Anteil und machen obiges Argument fuer die separaten Folgen.

Also ist

$$\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1$$

und für  $x_0 \in \mathbb{R}$  gilt damit

$$\lim_{x \rightarrow x_0} \frac{e^x - e^{x_0}}{x - x_0} = e^{x_0} \lim_{x \rightarrow x_0} \frac{e^{x-x_0} - 1}{x - x_0} = e^{x_0}.$$

□

### 1.9 Beispiel. Die Betragsfunktion

Die Betragsfunktion  $\text{abs}: x \mapsto |x|$  ist für  $x \neq 0$  differenzierbar

$$\text{abs}'(x) = \begin{cases} -1 & \text{für } x < 0 \\ 1 & \text{für } x > 0. \end{cases}$$

Im Punkt  $x = 0$  ist sie nicht differenzierbar, da

$$\lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{|x| - |0|}{x - 0} = 1 \quad \text{und} \quad \lim_{\substack{x \rightarrow 0 \\ x < 0}} \frac{|x| - |0|}{x - 0} = -1.$$

Um zusammengesetzte Funktionen bequem differenzieren zu können, benötigen wir noch die bekannten Ableitungsregeln.

### 1.10 Satz. Ableitungsregeln

Seien  $f, g: I \rightarrow \mathbb{R}$  differenzierbar (in  $x_0$ ). Dann gilt:

(a) Für  $a, b \in \mathbb{R}$  ist  $af + bg: I \rightarrow \mathbb{R}$  differenzierbar (in  $x_0$ ) und es gilt

$$(af + bg)' = af' + bg'. \quad \text{Linearität der Ableitung}$$

(b)  $f \cdot g$  ist differenzierbar (in  $x_0$ ) und es gilt

$$(f \cdot g)' = f'g + f \cdot g'. \quad \text{Produktregel}$$

(c) Dort wo  $g(x) \neq 0$  ist, gilt

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}. \quad \text{Quotientenregel}$$

*Beweis.* (a) Das folgt aus den Regeln für Grenzwerte.

(b) Da  $f$  insbesondere auch stetig ist, folgt die Behauptung, indem man in

$$\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} = f(x) \frac{g(x) - g(x_0)}{x - x_0} + g(x_0) \frac{f(x) - f(x_0)}{x - x_0}$$

den Limes  $x \rightarrow x_0$  bildet.

(c) Zunächst ist  $\left(\frac{1}{g}\right)' = \frac{-g'}{g^2}$ , da analog zu (b)

$$\frac{\frac{1}{g(x)} - \frac{1}{g(x_0)}}{x - x_0} = \frac{1}{g(x_0)g(x)} \frac{g(x_0) - g(x)}{(x - x_0)} \xrightarrow{x \rightarrow x_0} \frac{-g'(x_0)}{g^2(x_0)}$$

gilt. Mit der Produktregel folgt dann sofort

$$\left(\frac{f}{g}\right)' = f' \cdot \frac{1}{g} + f \left(\frac{1}{g}\right)' = \frac{f'g - fg'}{g^2}.$$

□

**1.11 Satz. Kettenregel**

Seien  $f : I \rightarrow \mathbb{R}$  und  $g : J \rightarrow \mathbb{R}$  differenzierbar mit  $g(J) \subset I$ . Dann ist auch  $f \circ g : J \rightarrow \mathbb{R}$  differenzierbar und es gilt

$$(f \circ g)'(x) = (f' \circ g)(x) \cdot g'(x) \quad \left( := f'(g(x)) g'(x) \right)$$

*Beweis.* Da  $f$  differenzierbar ist gilt

$$f(g(x+h)) - f(g(x)) = f'(g(x))(g(x+h) - g(x)) + \varphi_x^f(g(x+h) - g(x)),$$

sodass  $\lim_{s \rightarrow 0} \frac{\varphi_x^f(s)}{s} = 0$ . Auf Grund der Differenzierbarkeit von  $g$  gilt weiters

$$g(x+h) - g(x) = g'(x)h + \varphi_x^g(h).$$

Damit sehen

$$f(g(x+h)) - f(g(x)) = f'(g(x))g'(x)h + \text{Rest}_h,$$

wobei  $\lim_{h \rightarrow 0} \text{Rest}_h/h = 0$ , was die Behauptung zeigt. □

**1.12 Satz. Die Ableitung der Umkehrfunktion**

Es sei  $f : I \rightarrow J$  stetig und bijektiv,  $I, J \subset \mathbb{R}$  Intervalle. Es sei  $f$  differenzierbar in  $x_0 \in I$  mit  $f'(x_0) \neq 0$ . Dann ist die Inverse  $f^{-1} : J \rightarrow I$  differenzierbar in  $y_0 := f(x_0) \in J$  mit

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.$$

*Beweis.* (a) **Graphisch:** Eine Spiegelung an der Diagonalen invertiert die Steigung.

(b) **Analytisch:** Es gilt mit den Voraussetzungen, dass  $f^{-1} : J \rightarrow I$  ebenfalls stetig ist. Für jede Folge  $(y_n)$  in  $J$  mit  $y_n \rightarrow y_0, y_n \neq y_0$ , gilt daher  $x_n := f^{-1}(y_n) \in I, x_n \neq x_0$  und  $x_n \rightarrow f^{-1}(y_0) = x_0$ . Da  $f$  in  $x_0$  differenzierbar ist, folgt

$$\lim_{n \rightarrow \infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} = \lim_{n \rightarrow \infty} \frac{x_n - x_0}{f(x_n) - f(x_0)} = \frac{1}{f'(x_0)}$$

(c) Wenn wir einmal wissen dass  $f^{-1}$  differenzierbar ist, dann kann man den Wert der Ableitung leicht folgendermaßen ausrechnen. Einerseits gilt

$$\frac{d}{dx} f^{-1} \circ f(x) = \frac{d}{dx} x = 1,$$

andererseits mittels Kettenregel

$$1 = \frac{d}{dx} f^{-1} \circ f(x) = f^{-1}'(f(x)) f'(x),$$

also, mit  $y = f(x)$  und  $x = f^{-1}(y)$ ,

$$f^{-1}'(y) = \frac{1}{f'(f^{-1}(y))}.$$

□

Nun können wir die uns bisher bekannten Funktionen differenzieren:

**1.13 Beispiele.** (a) Für  $f_n(x) := x^n, n \in \mathbb{N}_0$ , gilt

$$f_0'(x) = 0 \quad \text{und} \quad f_n'(x) = nx^{n-1} \quad \text{für } n \in \mathbb{N}.$$

**Beweis: Übung:** Für Polynome  $p(x) = \sum_{j=0}^n a_j x^j$  gilt also

$$p'(x) = \sum_{j=1}^n a_j j x^{j-1} = \sum_{j=0}^{n-1} a_{j+1} (j+1) x^j.$$

(b) Die **Logarithmusfunktion**  $\ln : (0, \infty) \rightarrow \mathbb{R}$  hat mit Satz 1.12 die Ableitung

$$\ln'(y) = \frac{1}{\exp'(\ln(y))} = \frac{1}{\exp(\ln(y))} = \frac{1}{y}.$$

(c) Die Ableitung von **Sinus und Kosinus**.

Wir wissen bereits, dass  $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$ , und aus

$$0 \leq 1 - \cos x = 2 \sin^2 \frac{x}{2} \leq 2 \left(\frac{x}{2}\right)^2 = \frac{x^2}{2}$$

folgt

$$\lim_{x \rightarrow 0} \frac{\cos x - 1}{x} = 0.$$

Mit Hilfe der Additionstheoreme folgt dann

$$\begin{aligned} \sin'(x) &= \lim_{h \rightarrow 0} \frac{\sin(x+h) - \sin x}{h} = \lim_{h \rightarrow 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \lim_{h \rightarrow 0} \left( \cos x \frac{\sin h}{h} + \sin x \frac{\cos h - 1}{h} \right) = \cos x \end{aligned}$$

und

$$\begin{aligned} \cos'(x) &= \lim_{h \rightarrow 0} \frac{\cos(x+h) - \cos x}{h} = \lim_{h \rightarrow 0} \frac{\cos x \cos h - \sin x \sin h - \cos x}{h} \\ &= \lim_{h \rightarrow 0} \left( -\sin x \frac{\sin h}{h} + \cos x \frac{\cos h - 1}{h} \right) = -\sin x. \end{aligned}$$

(d) Für  $\arcsin : (-1, 1) \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$  ergibt sich

$$\arcsin'(y) = \frac{1}{\sin'(\arcsin y)} = \frac{1}{\cos(\arcsin y)} = \frac{1}{\sqrt{1 - \sin^2(\arcsin y)}} = \frac{1}{\sqrt{1 - y^2}},$$

da  $\cos x > 0$  für  $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ .

(e) Für  $\arccos : (-1, 1) \rightarrow (0, \pi)$  zeigen Sie in den Übungen, dass  $\arccos'(y) = -\frac{1}{\sqrt{1-y^2}}$ .

(f) Ebenfalls in den Übungen wird  $\tan' x = \frac{1}{\cos^2 x} = 1 + \tan^2 x$  mit Hilfe der Quotientenregel gezeigt.

(g) Schließlich gilt für  $\arctan : \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$

$$\arctan' y = \frac{1}{\tan'(\arctan y)} = (1 + \tan^2(\arctan y))^{-1} = (1 + y^2)^{-1} = \frac{1}{1 + y^2}.$$

Wie Sie aus der Schule wissen, liefern die Nullstellen der Ableitung Kandidaten für lokale Extrema.

#### 1.14 Definition. Lokale Extrema

Sei  $I \subset \mathbb{R}$  ein offenes Intervall. Wir sagen, dass  $f : I \rightarrow \mathbb{R}$  bei  $x_0 \in I$  ein

- **lokales Maximum** hat, falls

$$\exists \delta > 0 \forall x \in B_\delta(x_0) \cap I : f(x) \leq f(x_0)$$

- **lokales Minimum** hat, falls

$$\exists \delta > 0 \forall x \in B_\delta(x_0) \cap I : f(x) \geq f(x_0)$$

#### 1.15 Satz. Notwendiges Kriterium für lokales Extremum

Sei  $I$  ein offenes Intervall. Hat  $f : I \rightarrow \mathbb{R}$  bei  $x_0 \in I$  ein lokales Extremum (Maximum oder Minimum) und ist  $f$  bei  $x_0$  differenzierbar, so ist  $f'(x_0) = 0$ .

*Beweis.* Sei  $x_0$  lokales Maximum, dann gilt

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{\substack{x \rightarrow x_0 \\ x > x_0}} \frac{f(x) - f(x_0)}{x - x_0} \leq 0$$

und

$$f'(x_0) = \lim_{\substack{x \rightarrow x_0 \\ x < x_0}} \frac{f(x) - f(x_0)}{x - x_0} \geq 0$$

also  $f'(x_0) = 0$ . Für ein Minimum verfähre man analog. □

## 1.2 Der Mittelwertsatz

### 1.16 Satz. Satz von Rolle

Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig und in  $(a, b)$  differenzierbar und  $f(a) = f(b)$ . Dann existiert ein  $x_0 \in (a, b)$  mit  $f'(x_0) = 0$ .

*Beweis.* Ist  $f(x) = f(a)$  für alle  $x \in (a, b)$ , so ist  $f'(x) = 0$  für alle  $x \in (a, b)$ . Andernfalls ist  $\sup f > f(a)$  oder  $\inf f < f(a)$ . Deshalb nimmt  $f$  sein Maximum (oder Minimum) in einem Punkt  $x_0 \in (a, b)$  an, wo nach Satz 1.15  $f'(x_0) = 0$  gilt. □

### 1.17 Satz. Mittelwertsatz

Ist  $f : [a, b] \rightarrow \mathbb{R}$  stetig und in  $(a, b)$  differenzierbar, so gibt es (mindestens) ein  $x_0 \in (a, b)$  mit

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}$$

Steigung der Tangente bei  $x_0$  = Steigung der Sekante durch die Punkte  $(a, f(a))$  und  $(b, f(b))$ .

*Beweis.* Satz von Rolle "gekippt": Betrachte die Funktion

$$h : [a, b] \rightarrow \mathbb{R}, \quad h(x) := f(x) - (x - a) \frac{f(b) - f(a)}{b - a}.$$

Es ist  $g$  in  $(a, b)$  differenzierbar, in  $[a, b]$  stetig und es gilt  $h(a) = f(a) = h(b)$ . Also liefert der Satz von Rolle die Existenz eines  $x_0 \in (a, b)$  mit

$$0 = h'(x_0) = f'(x_0) - \frac{f(b) - f(a)}{b - a}.$$

□

**1.18 Korollar.** Sei  $f : (a, b) \rightarrow \mathbb{R}$ . Dann gilt

$$f \text{ ist differenzierbar mit } f'(x) = 0 \text{ für alle } x \in (a, b) \iff f \text{ ist konstant.}$$

*Beweis.* Übung. □

**1.19 Bemerkung.** Dieses Korollar scheint offensichtlich, spielt aber eine wichtige Rolle, da es die Eindeutigkeit von Lösungen von Differentialgleichungen liefert: haben zwei Funktionen die gleiche Ableitung, also  $f'(x) = g'(x)$  und stimmen sie für ein  $x_0$  überein, also  $f(x_0) = g(x_0)$ , so muss schon  $f(x) = g(x)$  für alle  $x$  gelten.

*Beweis.* Nach dem Korollar impliziert  $f'(x) - g'(x) = 0$ , dass  $f(x) - g(x)$  konstant ist. Die Konstante kann aber wegen  $f(x_0) = g(x_0)$  nur Null sein. □

**1.20 Korollar.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig und differenzierbar auf  $(a, b)$ . Dann gilt:

- (a) Es ist  $f$  genau dann monoton wachsend (bzw. fallend), wenn für alle  $x \in (a, b)$  gilt, dass  $f'(x) \geq 0$  (bzw.  $f'(x) \leq 0$ ).

(b) Ist  $f'(x) > 0$  (bzw.  $f'(x) < 0$ ) für alle  $x \in (a, b)$ , so ist  $f$  streng monoton wachsend (bzw. fallend).

*Beweis.* (a) „ $\Rightarrow$ “ liest man aus dem Differenzquotienten ab.

„ $\Leftarrow$ “ Für  $a \leq x < y \leq b$  existiert nach dem Mittelwertsatz ein  $\xi \in (x, y)$  mit  $f(y) - f(x) = f'(\xi)(y - x)$ .

Ist  $f'(\xi) \geq 0$  so folgt  $f(y) \geq f(x)$ , ist  $f'(\xi) \leq 0$  so folgt  $f(y) \leq f(x)$ .

(b) Wie (a) „ $\Leftarrow$ “, nur mit strikter Ungleichung. □

**1.21 Bemerkung.** Es gibt aber streng monotone Funktionen deren Ableitung in einem Punkt verschwindet, z.B.  $x \mapsto x^3$ .

### 1.22 Satz. Verallgemeinerter Mittelwertsatz

Seien  $f, g : [a, b] \rightarrow \mathbb{R}$  stetig und in  $(a, b)$  differenzierbar. Sei weiterhin  $g'(x) \neq 0$  für alle  $x \in (a, b)$ . Dann gibt es ein  $x_0 \in (a, b)$  mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x_0)}{g'(x_0)}.$$

(Für  $g(x) = x$  ist das der obige Mittelwertsatz.)

*Beweis.* Zunächst ein Hinweis: Man kann nicht einfach die Aussagen des Mittelwertsatzes

$$f(b) - f(a) = (b - a)f'(x_0), \quad g(b) - g(a) = (b - a)f'(\tilde{x}_0)$$

durcheinander teilen, da ja im Allgemeinen  $x_0 \neq \tilde{x}_0$  gilt. Allerdings kann man den Beweis des Mittelwertsatzes eins zu eins auf die allgemeinere Situation übertragen: Wegen  $g'(x) \neq 0$  ist mit dem Mittelwertsatz

$$g(b) - g(a) = (b - a)g'(x_0) \neq 0.$$

Wir können also

$$h(x) := f(x) - (g(x) - g(a)) \frac{f(b) - f(a)}{g(b) - g(a)}$$

betrachten. Wieder ist  $h$  in  $(a, b)$  differenzierbar und  $h(a) = f(a) = h(b)$ . Also gibt es nach Rolle ein  $x_0 \in (a, b)$  mit

$$0 = h'(x_0) = f'(x_0) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(x_0).$$

□

## 1.3 Der Satz von Taylor und Taylorreihen

### 1.23 Definition. Ableitungen höherer Ordnung

Ist  $f : I \rightarrow \mathbb{R}$  differenzierbar, so ist  $f' : I \rightarrow \mathbb{R}$  wieder eine Funktion. Falls  $f'$  ebenfalls differenzierbar ist, so nennt man  $f'' := (f')' =: f^{(2)}$  die zweite Ableitung von  $f$ . Entsprechend definiert man die  $n$ -te Ableitung  $f^{(n)}$ ,

$$f^{(n+1)} := (f^{(n)})' \quad \text{für } n \geq 1.$$

Man schreibt auch

$$f^{(n)} = \frac{d^n f}{dx^n} = \frac{d^n}{dx^n} f = \left( \frac{d}{dx} \right)^n f.$$

### 1.24 Satz. Leibniz-Regel

Sind  $f : I \rightarrow \mathbb{R}$  und  $g : I \rightarrow \mathbb{R}$  je  $n$ -mal differenzierbar, so ist auch das Produkt  $fg$  der Funktionen  $n$ -mal differenzierbar und es gilt

$$(f \cdot g)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)},$$

wobei  $f^{(0)} := f$  ist.

*Beweis. durch Induktion:*  $n = 0 \quad (f \cdot g)^{(0)} = f \cdot g = \sum_{k=0}^0 \binom{0}{k} f^{(k)} g^{(0-k)}$

$n \Rightarrow n + 1$

$$\begin{aligned} (f \cdot g)^{(n+1)} &= \frac{d}{dx} (f \cdot g)^{(n)} = \frac{d}{dx} \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)} = \sum_{k=0}^n \binom{n}{k} (f^{(k+1)} g^{(n-k)} + f^{(k)} g^{(n-k+1)}) \\ &= \binom{n}{0} f^{(0)} g^{(n+1)} + \left[ \binom{n}{0} + \binom{n}{1} \right] f^{(1)} g^{(n)} \\ &\quad + \dots + \left[ \binom{n}{n-1} + \binom{n}{n} \right] f^{(n)} g^{(1)} + \binom{n}{n} f^{(n+1)} g^{(0)} \\ &= \binom{n+1}{0} f^{(0)} g^{(n+1)} + \binom{n+1}{1} f^{(1)} g^{(n)} + \dots + \binom{n+1}{n} f^{(n)} g^{(1)} \\ &\quad + \binom{n+1}{n+1} f^{(n+1)} g^{(0)} \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} f^{(k)} g^{(n+1-k)}. \end{aligned}$$

□

### 1.25 Definition. Stetig differenzierbare Funktionen

Eine Funktion  $f : I \rightarrow \mathbb{R}$  heißt **stetig differenzierbar**, falls  $f$  differenzierbar ist und  $f' : I \rightarrow \mathbb{R}$  stetig ist. Eine Funktion  $f : I \rightarrow \mathbb{R}$  heißt  **$n$ -mal stetig differenzierbar**, falls  $f$   $n$ -mal differenzierbar ist und  $f^{(n)} : I \rightarrow \mathbb{R}$  stetig ist. Die Funktionen  $f^{(k)}$ ,  $k = 0, \dots, n - 1$  sind dann offenbar auch stetig, da sie differenzierbar sind.

Mit  $C^n(I)$  bezeichnet man den Raum der  $n$ -mal stetig differenzierbaren Funktionen auf  $I$ .

Wir kommen nun zu einem der wichtigsten Sätze der Analysis, dem Satz von Taylor. Wir haben in Satz 1.4 gesehen, dass Differenzierbarkeit äquivalent zu linearer Approximierbarkeit ist. Der Satz von Taylor sagt nun, dass sich Funktionen in  $C^{n+1}(I)$  durch Polynome vom Grade  $n$  lokal bis auf Fehler der Ordnung  $(x - x_0)^{n+1}$  approximieren lassen.

### 1.26 Satz. Satz von Taylor

Sei  $f \in C^{n+1}(I)$  und  $x_0 \in I$ . Dann läßt sich  $f(x)$  für  $x \in I$  wie folgt näherungsweise durch ein Polynom (das Taylorpolynom vom Grad  $n$  bei  $x_0$ ) in  $(x - x_0)$  darstellen:

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + R_n(x, x_0) \\ &= \underbrace{\sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k}_{\text{Taylorpolynom vom Grad } n \text{ bei } x_0} + R_n(x, x_0). \end{aligned}$$

Das Restglied hat die Form

$$R_n(x, x_0) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1},$$

wobei  $\xi = \xi(x)$  zwischen  $x_0$  und  $x$  liegt, also  $\xi = x_0 + \theta(x - x_0)$  mit  $0 < \theta < 1$ .

**1.27 Bemerkung.** (a) Für  $n = 1$  ist sowohl die Annahme,  $f \in C^2$ , als auch die Aussage,  $R_2 = C(x - x_0)^2$ , stärker als in Satz 1.4. Dort haben wir lediglich  $f$  differenzierbar angenommen und vom Restglied wissen wir dann nur, dass

$$\lim_{x \rightarrow x_0} \frac{R_2(x, x_0)}{x - x_0} = 0.$$

(b) Die Form des sog. Lagrangeschen Restglieds merkt man sich wie folgt:  $R_n(x)$  sieht so aus, wie das  $(n + 1)$ -te Glied aussehen würde, nur wird  $f^{(n+1)}$  an der Stelle  $\xi$  statt bei  $x_0$  ausgewertet. Dabei hängt  $\xi$  nun von  $x$  ab und wir kennen  $\xi(x)$  im Allgemeinen nicht.

Sobald wir integrieren können, werden wir noch eine andere Form des Restglieds kennenlernen.

*Beweis. des Satzes von Taylor:* Für festes  $x \in I$  seien

$$F(t) := f(x) - \sum_{j=0}^n \frac{f^{(j)}(t)}{j!} (x-t)^j \quad \text{und} \quad G(t) := \frac{(x-t)^{n+1}}{(n+1)!}.$$

Nun ist

$$F'(t) := \frac{d}{dt} F(t) = - \sum_{j=0}^n \left( \frac{f^{(j+1)}(t)}{j!} (x-t)^j - \frac{f^{(j)}(t)}{(j-1)!} (x-t)^{j-1} \right) = - \frac{f^{(n+1)}(t)}{n!} (x-t)^n$$

und

$$G'(t) = - \frac{(x-t)^n}{n!}.$$

Setzen wir nun  $t = x_0$  bzw.  $x$ , so liefert der verallgemeinerte Mittelwertsatz, Satz 1.22, ein  $\xi$  zwischen  $x_0$  und  $x$  mit

$$\frac{F(x) - F(x_0)}{G(x) - G(x_0)} = \frac{F'(\xi)}{G'(\xi)} = f^{(n+1)}(\xi).$$

Da

$$F(x) = G(x) = 0, \quad \text{folgt} \quad F(x_0) = G(x_0) f^{(n+1)}(\xi),$$

also die Behauptung. □

Ist eine Funktion  $f$  beliebig oft differenzierbar, man schreibt dann  $f \in C^\infty(I)$ , so heißt die Reihe

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x-x_0)^n$$

die Taylorreihe von  $f$  an der Stelle  $x_0$ .

Es stellt sich die Frage, ob die Taylorreihe konvergiert, und wenn ja, ob ihre Summe dann gleich  $f(x)$  ist. Letzteres gilt offenbar genau dann, wenn das Restglied  $R_n(x)$  für  $n \rightarrow \infty$  gegen Null geht.

### 1.28 Satz. Hinreichendes Kriterium für lokales Extremum

Sei  $I$  ein offenes Intervall und  $f \in C^2(I)$ .

(a) Hat  $f$  bei  $x_0 \in I$  ein lokales Maximum (bzw. Minimum), so gilt notwendigerweise

$$f'(x_0) = 0 \quad \text{und} \quad f''(x_0) \leq 0 \quad (\text{bzw. } f''(x_0) \geq 0).$$

(b) Gilt bei  $x_0 \in I$

$$f'(x_0) = 0 \quad \text{und} \quad f''(x_0) < 0 \quad (\text{bzw. } f''(x_0) > 0),$$

so hat  $f$  bei  $x_0$  ein lokales Maximum (bzw. Minimum).

*Beweis.* Das folgt sofort aus der Taylorentwicklung um den Punkt  $x_0$  (Übungsaufgabe). □

### 1.29 Satz. Regel von de l'Hospital

Es seien  $f, g : (a, b) \rightarrow \mathbb{R}$  differenzierbar, wobei  $-\infty \leq a < b \leq \infty$  und  $g'(x) \neq 0$  für alle  $x \in (a, b)$ . Weiter gelte

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0$$

oder

$$\lim_{x \rightarrow a} g(x) = \pm \infty.$$

Dann gilt

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$$

falls der rechtsstehende Grenzwert existiert. Die analogen Aussagen gelten für  $x \rightarrow b$ .

## 1 Eindimensionale Differentialrechnung

*Beweis.* Wir betrachten den Fall  $a \in \mathbb{R}$  und  $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0$ , können also mittels der Setzung  $f(a) = g(a) = 0$  die Funktionen  $f$  und  $g$  stetig auf das Intervall  $[a, b]$  fortsetzen. Der verallgemeinerte Mittelwertsatz liefert dann für  $x \in (a, b)$

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(x_0)}{g'(x_0)} \quad \text{für ein } a < x_0 < x.$$

Daraus folgt die Behauptung dann für  $x \rightarrow a$ . Für die anderen Fälle siehe W. Walter, Analysis I, Satz 10.11.  $\square$

Eine der wichtigsten Eigenschaften der Analysis ist Konvexität (Konkavität).

**1.30 Definition.** Eine Funktion  $f : A \rightarrow \mathbb{R}$  heißt **konvex**,  $A \subset \mathbb{R}^n$  falls  $\forall x, y \in A$  und  $t \in [0, 1]$  gilt, dass

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

Gilt stets  $<$  falls  $x \neq y$  und  $t \in (0, 1)$ , so heißt  $f$  **strikt konvex**. Es gilt weiters  $f$  (strikt) konkav  $\Leftrightarrow -f$  strikt konvex.

**1.31 Satz.** Seien  $I \subseteq \mathbb{R}$  ein offenes Intervall und  $f : I \rightarrow \mathbb{R}$  zweimal differenzierbar. Dann gilt:  $f$  ist konvex genau dann wenn  $f''(x) \geq 0 \forall x \in I$ .

*Beweis.* Es genügt zu zeigen:  $f$  konvex  $\Leftrightarrow f'$  ist monoton wachsend.

$\Rightarrow$ : Wir zeigen dass für beliebiges  $y \in I$ , die Funktion  $x \mapsto \frac{f(x) - f(y)}{x - y}$  eine monoton wachsende Funktion ist in  $I \setminus \{y\}$ . Wir zeigen dies zuerst für  $x > y$ . Also wir nehmen  $x_2 > x_1$ , und beide  $> y$ . Dann gibt es  $t \in [0, 1]$  sodass  $x_1 = (1-t)y + tx_2$  und aufgrund der Konvexität

$$f(x_1) = f((1-t)y + tx_2) \leq (1-t)f(y) + tf(x_2),$$

und somit

$$\frac{f(x_1) - f(y)}{x_1 - y} \leq \frac{t(f(x_2) - f(y))}{t(x_2 - y)} = \frac{f(x_2) - f(y)}{x_2 - y}.$$

Gilt nun  $x_1 < x_2 < y$ , dann haben wir mit  $x_2 = (1-t)x_1 + ty$ , dass

$$f(x_2) - f(y) \leq (1-t)(f(x_1) - f(y)).$$

Da  $x_2 - y = (1-t)(x_1 - y) < 0$ , gilt

$$\frac{f(x_2) - f(y)}{x_2 - y} \geq \frac{(1-t)(f(x_1) - f(y))}{(1-t)(x_1 - y)} = \frac{f(x_1) - f(y)}{x_1 - y}.$$

Sei nun  $f$  differenzierbar, sodass der Differenzenquotient existiert, dann gilt für  $x_1 < x_2$ , dass

$$f'(x_1) = \lim_{x \rightarrow x_1^+} \frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{f(x_1) - f(x_2)}{x_1 - x_2} \quad (1.1)$$

$$\leq \lim_{x \rightarrow x_2^+} \frac{f(x) - f(x_2)}{x - x_2} = f'(x_2) \quad (1.2)$$

$\Leftarrow$ : Sei nun umgekehrt  $f'(x)$  monoton wachsend, und  $x < y$ . Wähle  $z = (1-t)x + ty$ . Dann gilt mit dem Mittelwertsatz und der Monotonie der Ableitung, dass

$$\frac{f((1-t)x + ty) - f(x)}{t(y-x)} = \frac{f(z) - f(x)}{z-x} \leq \frac{f(y) - f(z)}{y-z} = \frac{f(y) - f((1-t)x + ty)}{(1-t)(y-x)},$$

was

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$$

impliziert.  $\square$

Damit können wir auch folgern, dass jedes lokale Minimum einer konvexen Funktion automatisch auch ein globales Minimum ist. Also, falls eine konvexe Funktion differenzierbar ist, so ist jeder kritische Punkt automatisch ein globales Minimum.

**1.32 Satz.** Sei  $f : I \rightarrow \mathbb{R}$  konvex, dann gilt

- (a) Jedes lokale Minimum von  $f$  ist ein globales Minimum.
- (b) Ist  $f$  strikt konvex so gibt es höchstens eine Stelle eines Globalen Minimums.
- (c) Ist  $f$  differenzierbar, und  $f'(x_0) = 0$ , so ist  $x_0$  Stelle eines globalen Minimums.

*Beweis.* (a) Sei  $x_0$  Stelle eines lokalen Minimums, als es gibt  $\varepsilon$ , sodass  $f(x) \geq f(x_0)$  für alle  $x \in U_\varepsilon(x_0)$ . Sei  $y \in I$  beliebig. Dann gibt es  $t \in [0, 1]$ , sodass  $z = (1 - t)x_0 + ty \in U_\varepsilon(x_0)$ , und  $f(x_0) \leq f(z) \leq (1 - t)f(x_0) + tf(y)$ , und somit auch  $f(x_0) \leq f(y)$ .

- (b) Sei  $x \neq y$  zwei Stellen globaler Minima, dann gilt

$$f\left(\frac{1}{2}x + \frac{1}{2}y\right) < \frac{1}{2}f(x) + \frac{1}{2}f(y) = f(x),$$

somit haetten wir Widerspruch.

- (c) Übung.

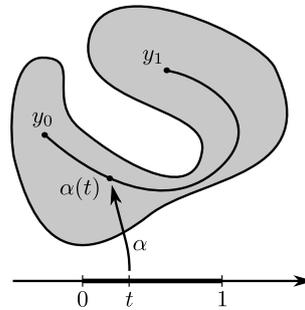
□



## 2 Mehrdimensionale Differenzierbarkeit

### 2.1 Definition. Wegzusammenhang

Sei  $X = \mathbb{R}^n$ . Eine Teilmenge  $Y \subset X$  heißt **wegzusammenhängend**, wenn es zu je zwei Punkten  $y_0, y_1 \in Y$  einen **Weg** von  $y_0$  nach  $y_1$  in  $Y$  gibt, d.h. eine stetige Abbildung  $\alpha : [0, 1] \rightarrow Y$  mit  $\alpha(0) = y_0$  und  $\alpha(1) = y_1$ .



**2.2 Satz.** Seien  $X$  und  $Y$  Teilmengen von  $\mathbb{R}^n$ , bzw  $\mathbb{R}^m$ , und sei  $f : X \rightarrow Y$  stetig. Dann gilt: Ist  $A \subset X$  wegzusammenhängend, so ist auch  $f(A) \subset Y$  wegzusammenhängend.

*Beweis.* Seien  $y_0, y_1 \in f(A)$  beliebig und  $x_0, x_1 \in A$  so, dass  $f(x_0) = y_0$  und  $f(x_1) = y_1$ . Dann gibt es nach Annahme einen stetigen Weg  $\alpha : [0, 1] \rightarrow A$  mit  $\alpha(0) = x_0$  und  $\alpha(1) = x_1$ . Dann ist auch  $\beta = f \circ \alpha$  stetig und damit ein Weg von  $(f \circ \alpha)(0) = f(x_0) = y_0$  nach  $(f \circ \alpha)(1) = f(x_1) = y_1$ . Also ist  $f(A)$  wegzusammenhängend.  $\square$

### 2.3 Definition. Gebiet

Sei  $X = \mathbb{R}^n$ . Eine nichtleere Teilmenge  $U \subset X$  heißt **Gebiet in  $X$** , falls  $U$  offen und wegzusammenhängend ist.

Im Folgenden wird meist aus rein praktischen Gründen vorausgesetzt, dass Funktionen auf Gebieten definiert sind.

**2.4 Bemerkung.** Die Gebiete in  $\mathbb{R}$  sind genau die offenen Intervalle  $(a, b) \subset \mathbb{R}$  mit  $a < b$ ,  $a \in \{-\infty\} \cup \mathbb{R}$  und  $b \in \mathbb{R} \cup \{\infty\}$ . *Beweis.* Übungsaufgabe.

### 2.5 Definition. Partielle Ableitung

Sei  $n \geq 1$  und  $G \subset \mathbb{R}^n$  ein Gebiet. Für  $x \in G$  und  $j \in \{1, \dots, n\}$  heißt eine Funktion  $f : G \rightarrow \mathbb{R}$  im **Punkt  $x$  in die  $j$ -te Koordinatenrichtung partiell differenzierbar**, wenn der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(x + h e_j) - f(x)}{h}$$

existiert, wobei  $e_j$  den  $j$ -ten kanonischen Basisvektor des  $\mathbb{R}^n$  bezeichnet. Wir schreiben dann

$$\frac{\partial f}{\partial x_j}(x) = D_j f(x) = \partial_j f(x) := \lim_{h \rightarrow 0} \frac{f(x + h e_j) - f(x)}{h}$$

und nennen diese Zahl die  **$j$ -te partielle Ableitung von  $f$  in  $x$** . Die drei angegebenen Schreibweisen sind alle gebräuchlich und wir werden sie auch alle verwenden.

**2.6 Bemerkung.** Zur Erinnerung: Dass der  $\lim_{h \rightarrow 0}$  existiert bedeutet, dass für alle Nullfolgen  $(h_k)_{k \in \mathbb{N}}$  in  $\mathbb{R} \setminus \{0\}$  mit  $h_k$  derart, dass  $x + h_k e_j \in G$  ist, der Grenzwert

$$\lim_{k \rightarrow \infty} \frac{f(x + h_k e_j) - f(x)}{h_k}$$

existiert und damit auch für all diese Folgen gleich ist.

**2.7 Bemerkung.** Man führt also in dieser Definition die Differenzierbarkeit einer Funktion mehrerer Veränderlicher auf die Situation einer reellen Veränderlichen zurück. Setzt man  $g(x_j) := f(x_1, x_2, \dots, x_j, \dots, x_n)$  für festes  $x_i$  falls  $i \neq j$ , so ist  $\frac{\partial f}{\partial x_j}(x) = g'(x_j)$ .

**2.8 Beispiel. Euklidische Norm**

Betrachte die Radiusfunktion

$$r : \mathbb{R}^n \rightarrow [0, \infty), \quad x \mapsto r(x) = \|x\|_2 = \sqrt{x_1^2 + \dots + x_n^2}.$$

Dann ist  $r$  für alle  $x \in \mathbb{R}^n \setminus \{0\}$  und  $j \in \{1, \dots, n\}$  partiell in die  $j$ -te Richtung differenzierbar und es gilt

$$\partial_j r(x) = \frac{1}{2} \frac{2x_j}{\sqrt{x_1^2 + \dots + x_n^2}} = \frac{x_j}{\|x\|}.$$

**2.9 Definition. Partielle und stetige partielle Differenzierbarkeit**

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f$  eine reelle Funktion auf  $G$ , also  $f : G \rightarrow \mathbb{R}$ .

- (a) Für  $x \in G$  heißt  $f$  in  $x$  **partiell differenzierbar**, wenn  $f$  in alle  $n$  Richtungen partiell differenzierbar ist. Man nennt dann den Vektor

$$\text{grad } f(x) = \nabla f(x) = (\partial_1 f(x), \dots, \partial_n f(x)) \in \mathbb{R}^n$$

den **Gradienten von  $f$**  in  $x$ .

- (b) Es heißt  $f$  **partiell differenzierbar in  $G$** , falls  $f$  in allen Punkten in  $G$  partiell differenzierbar ist.
- (c) Es heißt  $f$  **stetig partiell differenzierbar in  $G$** , falls  $f$  partiell differenzierbar in  $G$  ist und die partiellen Ableitungen  $\partial_j f : G \rightarrow \mathbb{R}, j = 1, \dots, n$ , stetig sind.

**2.10 Bemerkung.** Man beachte, dass die partielle Differenzierbarkeit einer Funktion  $f : G \rightarrow \mathbb{R}$  in einem Punkt  $x \in G$  im Allgemeinen nicht, wie im Fall  $n = 1$ , die Stetigkeit von  $f$  in  $x$  nach sich zieht. Betrachte z.B. die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,

$$f(x, y) = \begin{cases} 0 & \text{wenn } xy = 0 \\ 1 & \text{wenn } xy \neq 0. \end{cases}$$

Dann ist  $f$  nämlich in  $(x, y) = (0, 0)$  partiell differenzierbar mit  $\text{grad } f(0, 0) = (0, 0)$ , aber offenbar nicht stetig. Wir werden jedoch sehen, dass jede Funktion  $f$  die in einem Gebiet stetig partiell differenzierbar ist, dort auch stetig ist.

**2.11 Bemerkung. Rechenregeln für partielle Ableitungen**

Da der Begriff der partiellen Differenzierbarkeit einer Funktion  $f : G \rightarrow \mathbb{R}, G \subset \mathbb{R}^n$  ein Gebiet, auf den der Differenzierbarkeit einer Funktion in einer Veränderlichen zurückgeführt werden kann, gelten auch die bekannten Rechenregeln:

- (a) **Linearität:** Sind  $f, g : G \rightarrow \mathbb{R}$  partiell differenzierbar und  $\lambda \in \mathbb{R}$ , dann sind auch  $f + g : G \rightarrow \mathbb{R}$  und  $\lambda f : G \rightarrow \mathbb{R}$  partiell differenzierbar und es gilt für  $j \in \{1, \dots, n\}$

$$\begin{aligned} \partial_j(f + g) &= \partial_j f + \partial_j g \\ \partial_j(\lambda f) &= \lambda \partial_j f. \end{aligned}$$

Insbesondere ist also

$$C^1(G) := \{f : G \rightarrow \mathbb{R} \mid f \text{ stetig partiell differenzierbar}\}$$

ein reeller Vektorraum.

- (b) **Produktregel:** Sind  $f, g : G \rightarrow \mathbb{R}$  partiell differenzierbar, so auch ihr Produkt  $fg : G \rightarrow \mathbb{R}$  und ihr Quotient  $f/g : G \rightarrow \mathbb{R}$  (falls  $g \neq 0$  auf  $G$ ).

Für  $j \in \{1, \dots, n\}$  gilt dann

$$\begin{aligned}\partial_j(fg) &= (\partial_j f)g + f(\partial_j g) \\ \partial_j\left(\frac{f}{g}\right) &= \frac{(\partial_j f)g - f(\partial_j g)}{g^2}.\end{aligned}$$

- (c) **Kettenregel:** Ist  $f : G \rightarrow \mathbb{R}$  partiell differenzierbar und  $h : \mathbb{R} \rightarrow \mathbb{R}$  differenzierbar, so ist auch  $h \circ f : G \rightarrow \mathbb{R}$  partiell differenzierbar und es gilt

$$\partial_j(h \circ f) = (h' \circ f) \partial_j f.$$

**2.12 Beispiel.** Ist  $f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$  eine **rotationssymmetrische Funktion**, d.h.  $f(x) = f(y)$  falls  $\|x\| = \|y\|$ , so gibt es eine Funktion  $h : (0, \infty) \rightarrow \mathbb{R}$ , so dass  $f(x) = h(\|x\|)$  für alle  $x \in \mathbb{R}^n \setminus \{0\}$ . Wähle z.B.  $h(r) = f(r, 0, 0, \dots, 0)$ . Es ist also  $f = h \circ r$  mit  $r(x) = \|x\|$  wie zuvor. Ist nun  $f$  partiell differenzierbar, so ist auch  $h$  differenzierbar ( $h'(r) = \partial_1 f(r, 0, \dots, 0)$ ) und es gilt für  $j \in \{1, \dots, n\}$

$$\begin{aligned}\partial_j f(x) &= \partial_j(h \circ r)(x) = (h' \circ r)(x) \partial_j r(x) \\ &= h'(\|x\|) \frac{x_j}{\|x\|},\end{aligned}$$

also  $\text{grad} f(x) = h'(\|x\|) \frac{x}{\|x\|}$ .

### 2.13 Definition. Vektorwertige Funktionen

Seien  $m, n \geq 1$ ,  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}^m$  eine vektorwertige Funktion, also  $f = (f_1, \dots, f_m)$  mit  $f_i : G \rightarrow \mathbb{R}$  für  $i = 1, \dots, m$ . Wir sagen, dass,  $f$  (**stetig**) **partiell differenzierbar** ist, falls jede Komponente  $f_i$  (stetig) partiell differenzierbar ist.

### 2.14 Definition. Vektorfeld

Sei  $G \subset \mathbb{R}^n$ . Eine Abbildung  $f : G \rightarrow \mathbb{R}^n$  heißt ein **Vektorfeld** auf  $G$ .

### 2.15 Beispiel. Gradient als Vektorfeld

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}$  partiell differenzierbar. Dann ist  $\text{grad} f : G \rightarrow \mathbb{R}^n$  ein Vektorfeld.

### 2.16 Definition. Divergenz und Laplace

Sei  $G \subset \mathbb{R}^n$  ein Gebiet.

- (a) Für ein partiell differenzierbares Vektorfeld  $f : G \rightarrow \mathbb{R}^n$  heißt die Funktion

$$\begin{aligned}\text{div} f : G &\rightarrow \mathbb{R} \\ x &\mapsto \text{div} f(x) = \sum_{j=1}^n \frac{\partial f_j}{\partial x_j}(x)\end{aligned}$$

die **Divergenz** von  $f$ .

- (b) Sei  $f : G \rightarrow \mathbb{R}$  partiell differenzierbar und sei auch  $\text{grad} f : G \rightarrow \mathbb{R}^n$  partiell differenzierbar, so heißt die Funktion

$$\begin{aligned}\Delta f : G &\rightarrow \mathbb{R} \\ x &\mapsto \Delta f(x) = \text{div}(\text{grad} f)(x) = \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j^2}\end{aligned}$$

**Laplace von  $f$ .**

**2.17 Beispiel.** (a) Die Identität  $\text{id} : G \rightarrow \mathbb{R}^n, x \mapsto x$ , ist ein partiell differenzierbares Vektorfeld auf  $G$ . Ihre Divergenz ist

$$\text{div}(\text{id})(x) = \sum_{j=1}^n \frac{\partial \text{id}_j}{\partial x_j}(x) = \sum_{j=1}^n \underbrace{\frac{\partial x_j}{\partial x_j}}_{=1}(x) = n.$$

(b) Die Abbildung  $f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}^n, x \mapsto \frac{x}{\|x\|} = \text{grad } r$  ist ein partiell differenzierbares Vektorfeld auf  $\mathbb{R}^n \setminus \{0\}$ . Die Divergenz ist

$$\begin{aligned} \text{div} f(x) &= \sum_{j=1}^n \frac{\partial}{\partial x_j} \left( \frac{x_j}{\|x\|} \right) = \sum_{j=1}^n \left( \frac{1}{\|x\|} - \frac{x_j^2}{\|x\|^3} \right) \\ &= n \frac{1}{\|x\|} - \frac{\|x\|^2}{\|x\|^3} = \frac{n-1}{\|x\|}. \end{aligned}$$

Also ist  $\Delta r = \text{div}(\text{grad } r) = \frac{n-1}{r}$ .

**2.18 Definition. Richtungsableitung**

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}^m$  eine Funktion. Unter der Richtungsableitung von  $f$  im Punkt  $x \in G$  in Richtung  $v \in \mathbb{R}^n$ , mit  $\|v\| = 1$ , versteht man den Differentialquotienten

$$\partial_v f(x) = D_v f(x) = \frac{d}{dh} f(x + hv)|_{h=0} = \lim_{h \rightarrow 0} \frac{f(x + hv) - f(x)}{h},$$

falls dieser existiert. Für  $v = e_j$  ist  $\partial_v f$  also gleich der  $j$ -ten partiellen Ableitung  $\partial_j f$ .

**2.19 Satz. Richtungsableitung und Gradient**

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}$  stetig differenzierbar. Dann gilt für jedes  $x \in G$  und jeden Vektor  $v \in \mathbb{R}^n$ , mit  $\|v\| = 1$ ,

$$\partial_v f(x) = \langle \nabla f(x), v \rangle$$

wobei  $\nabla f = \text{grad } f$ .

*Proof.* Der Beweis wird einfach aus der verallgemeinerten Kettenregel folgen. Wir skizzieren hier einen Beweis für  $n = 2$ : Mit  $v = (v_1, v_2)$  und  $x = (x_1, x_2)$  gilt

$$\begin{aligned} f(x + hv) - f(x) &= f(x_1 + hv_1, x_2 + hv_2) - f(x_1, x_2 + hv_2) + f(x_1, x_2 + hv_2) - f(x_1, x_2) \\ &= \partial_1 f(x_1, x_2 + hv_2) hv_1 + \varphi_1(hv_1) + \partial_2 f(x_1, x_2) hv_2 + \varphi_2(hv_2). \end{aligned} \quad (2.1)$$

Mit der Stetigkeit der partiellen Ableitung und der Definition von Differenzierbarkeit von Funktionen von einer Variablen folgt, dass

$$\lim_{h \rightarrow 0} \frac{f(x + hv) - f(x)}{h} = \nabla f(x) \cdot v.$$

□

**2.20 Bemerkung.** (a) Ist  $\nabla f(x) \neq 0$ , so ist der Winkel  $\theta$  zwischen den Vektoren  $v$  und  $\nabla f(x)$  definiert und es gilt

$$\partial_v f(x) = \langle \nabla f(x), v \rangle = \|\nabla f(x)\| \|v\| \cos \theta.$$

Die Richtungsableitung ist also maximal, falls  $v$  und  $\nabla f(x)$  die gleiche Richtung haben. Der Vektor  $\nabla f(x)/\|\nabla f(x)\|$  gibt also die Richtung des stärksten Anstiegs von  $f$  an.

**2.21 Bemerkung.** Betrachten wir nun den Spezialfall,  $A \subseteq \mathbb{R}^2$ , und  $f : A \rightarrow \mathbb{R}$  stetig differenzierbar. Dann ist der Graph von  $f$  die Fläche

$$\{(x, y, f(x, y)) : (x, y) \in A\}.$$

Die Tangentialebene an den Graphen im Punkt  $(\bar{x}, \bar{y}, f(\bar{x}, \bar{y}))$  ist gegeben durch:

$$z = f(\bar{x}, \bar{y}) + f_x(\bar{x}, \bar{y})(x - \bar{x}) + f_y(\bar{x}, \bar{y})(y - \bar{y}),$$

und die Tangente an den Graphen in Richtung  $v = (v_1, v_2)$ ,  $\|v\| = 1$ , ist gegeben durch:

$$t_v : \phi(t) = \begin{pmatrix} \bar{x} \\ \bar{y} \\ f(\bar{x}, \bar{y}) \end{pmatrix} + t \begin{pmatrix} v_1 \\ v_2 \\ \nabla f(\bar{x}, \bar{y}) \cdot v \end{pmatrix}.$$

Die Steigung der Tangente in Richtung oder äquivalent dazu die Steigung des Graphen in Richtung  $v$  im Punkt  $(\bar{x}, \bar{y}, f(\bar{x}, \bar{y}))$  lautet

$$\partial f_v(\bar{x}, \bar{y}) = \nabla f(\bar{x}, \bar{y}) \cdot v.$$

Sucht man eine Tangente an den Graphen mit einer vorgegebenen Steigung  $k$ , so müssen  $v_1$  und  $v_2$  so gesucht werden, dass

$$v_1^2 + v_2^2 = 1 \quad \nabla f(\bar{x}, \bar{y}) \cdot v = k$$

gilt. Zwei interessante Spezialfälle sind die Tangenten in Richtung des steilsten Anstieges und die waagrechte Tangente.

## 2.22 Definition. Die $r$ -mal stetig partiell differenzierbaren Funktionen $C^r(G)$

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $r \in \mathbb{N}$ . Eine Funktion  $f : G \rightarrow \mathbb{R}$  heißt  $r$ -mal stetig partiell differenzierbar, wenn für alle  $j = (j_1, \dots, j_r)$  mit  $j_1, \dots, j_r \in \{1, \dots, n\}$  gilt:

- $f$  ist stetig partiell differenzierbar
- $\partial_{j_1} f$  ist stetig partiell differenzierbar
- $\partial_{j_2}(\partial_{j_1} f)$  ist stetig partiell differenzierbar.
- ⋮
- $\partial_{j_{r-1}} \cdots \partial_{j_1} f$  ist stetig partiell differenzierbar, also  $\partial_{j_r} \cdots \partial_{j_1} f$  ist stetig.

Wiederholte Anwendung von 2.11 (a) liefert, dass

$$C^r(G) := \{f : G \rightarrow \mathbb{R} \mid f \text{ ist } r\text{-mal stetig partiell differenzierbar}\}$$

ein reeller Vektorraum ist.

**2.23 Notation.** Ist  $f$   $r$ -mal stetig partiell differenzierbar, so schreibt man auch

$$\frac{\partial^r f}{\partial x_{j_r} \cdots \partial x_{j_1}}(x) = \partial_{j_r} \cdots \partial_{j_1} f(x)$$

für jedes  $j = (j_1, \dots, j_r) \in \{1, \dots, n\}^r$ .

## 2.24 Satz. von Schwarz

Sei  $G \subset \mathbb{R}^n$  ein Gebiet,  $f : G \rightarrow \mathbb{R}$  zweimal stetig partiell differenzierbar und  $1 \leq i, j \leq n$ . Dann vertauschen die partiellen Ableitungen, d.h. für alle  $x \in G$  gilt

$$\partial_i \partial_j f(x) = \partial_j \partial_i f(x).$$

## 2 Mehrdimensionale Differenzierbarkeit

*Beweis.* Wir beschränken uns im Beweis auf den Fall  $\mathbb{R}^2$ . Für  $(h, k) \in \mathbb{R}^2$  sei

$$A(h, k) = f(x + h, y + k) - f(x + h, y) - f(x, y + k) + f(x, y).$$

Wir werden zeigen dass

$$\lim_{(h,k) \rightarrow (0,0)} \frac{A(h, k)}{hk} = f_{xy}(x, y)$$

und

$$\lim_{(h,k) \rightarrow (0,0)} \frac{A(h, k)}{hk} = f_{yx}(x, y),$$

womit dann die Behauptung folgt. Seien

$$\begin{aligned} \psi(h, k) &= f(x + h, y + k) - f(x, y + k) \\ \phi(h, k) &= f(x + h, y + k) - f(x + h, y) \end{aligned}$$

Mittels zweimaliger Anwendung des Mittelwertsatzes bekommen wir für geeignete  $\theta, \theta' \in [0, 1]$

$$\begin{aligned} A(h, k) &= \psi(h, k) - \psi(h, 0) = \psi_y(h, \theta k)k = [f_y(x + h, y + \theta k) - f_y(x, y + \theta k)]k \\ &= f_{xy}(x + \theta' h, y + \theta k)hk, \end{aligned}$$

als auch für geeignete  $\nu, \nu'$ , dass

$$\begin{aligned} A(h, k) &= \phi(h, k) - \phi(0, k) = \phi_x(\nu h, k)h = [f_x(x + \nu h, y + k) - f_x(x + \nu h, y)]h \\ &= f_{xy}(x + \nu h, y + \nu' k)hk. \end{aligned}$$

Obige Aussage folgt durch die Annahme dass  $f$  zweimal stetig differenzierbar ist in einer Umgebung von  $(x, y)$ . Denn unter dieser Voraussetzung gilt, dass  $f_{xy}(x + \theta' h, y + \theta k)$  als auch  $f_{xy}(x + \nu h, y + \nu' k)$  konvergieren für  $(h, k) \rightarrow (0, 0)$ .  $\square$

### 2.25 Definition. Hesse-Matrix

Für  $f \in C^2(G)$  und  $x \in G$  nennt man die  $n \times n$ -Matrix

$$\text{Hess}f(x) = \begin{pmatrix} \partial_1 \partial_1 f(x) & \partial_1 \partial_2 f(x) & \cdots & \partial_1 \partial_n f(x) \\ \partial_2 \partial_1 f(x) & \partial_2 \partial_2 f(x) & \cdots & \partial_2 \partial_n f(x) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_n \partial_1 f(x) & \partial_n \partial_2 f(x) & \cdots & \partial_n \partial_n f(x) \end{pmatrix}$$

die **Hesse-Matrix von  $f$  in  $x$** . Wegen Satz 2.24 ist  $\text{Hess}f(x)$  symmetrisch.

**2.26 Bemerkung.** Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Dann ist die zweite Ableitung von  $f$  in Richtung  $v$  gegeben durch

$$\left. \frac{d^2}{dt^2} \right|_{t=0} f(x + tv) = \langle v, \text{Hess}f(x)v \rangle.$$

In dem Sinne ist die Matrix  $\text{Hess}f(x)$  die zweite Ableitung von  $f$  an  $x$ , und die Eigenschaften dieser Matrix werden entscheiden sein, ob an dem Punkt  $x$  ein lokales Minimum oder Maximum sein kann. Dies wird von den Eigenwerten der Matrix  $\text{Hess}f(x)$  abhängen. Dazu kommen wir später.

**2.27 Beispiel.** Für  $f \in C^2(G)$  gilt

$$\begin{aligned} \Delta f(x) &= \text{div}(\text{grad}f)(x) = \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j^2}(x) \\ &= \text{Spur}(\text{Hess}f)(x). \end{aligned}$$

### 2.28 Definition. Rotation eines Vektorfeldes

Sei  $G \subset \mathbb{R}^3$  ein Gebiet und  $v : G \rightarrow \mathbb{R}^3$  ein partiell differenzierbares Vektorfeld,  $v = (v_1, v_2, v_3)$ . Man definiert die **Rotation von  $v$**  durch  $\operatorname{rot} v : G \rightarrow \mathbb{R}^3$ ,

$$\operatorname{rot} v := \left( \frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right).$$

### 2.29 Notation. Nabla-Operator

Führt man  $\nabla$  als "vektoriellen Operator" ein, den sogenannten **Nabla-Operator**,

$$\nabla = \left( \frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right) = (\partial_1, \dots, \partial_n),$$

so schreiben sich Gradient, Divergenz, Rotation und Laplace folgendermaßen:

$$\begin{aligned} \operatorname{grad} f &= (\partial_1 f, \dots, \partial_n f) = \nabla f \\ \operatorname{div} v &= \sum_{j=1}^n \frac{\partial v_j}{\partial x_j} = \langle \nabla, v \rangle = \nabla \cdot v \\ \Delta f &= \operatorname{div}(\operatorname{grad} f) = \nabla \cdot \nabla f \\ \operatorname{rot} v &= \nabla \times v. \end{aligned}$$

**2.30 Korollar. zu Satz 2.24.** Sei  $G \subset \mathbb{R}^3$  ein Gebiet.

- (a) Für  $f \in C^2(G)$  gilt  $\operatorname{rot}(\operatorname{grad} f) = 0$ .
- (b) Für  $v \in C^2(G, \mathbb{R}^3)$  gilt  $\operatorname{div}(\operatorname{rot} v) = 0$ .

*Beweis.* Übungsaufgabe: Man rechne nach, dass man  $\nabla$  formal wie einen Vektor behandeln kann:

$$\begin{aligned} \operatorname{rot}(\operatorname{grad} f) &= \nabla \times \nabla f = 0, \\ \operatorname{div}(\operatorname{rot} v) &= \nabla \cdot (\nabla \times v) = 0. \end{aligned}$$

□

**2.31 Beispiel.** Sei  $F : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$  zweimal stetig partiell differenzierbar und rotationssymmetrisch, also  $F(x) = f(r(x))$  mit  $r(x) = \|x\|$  und  $f \in C^2((0, \infty))$ . Dann ist auch  $\Delta F : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$  rotationssymmetrisch, also  $\Delta F = g \circ r$  für eine stetige Funktion  $g : (0, \infty) \rightarrow \mathbb{R}$  und es gilt

$$g(r) = f''(r) + \frac{n-1}{r} f'(r).$$

*Beweis.* Übungsaufgabe

□

Wir führen nun einen etwas anderen Differenzierbarkeitsbegriff ein, der im Gegensatz zur partiellen Differenzierbarkeit geometrisch motiviert ist und die geometrische Bedeutung der Ableitung als **lineare Approximation** in den Vordergrund stellt.

### 2.32 Definition. Totale Differenzierbarkeit

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}^m$  eine Abbildung. Es heißt  $f$  in einem Punkt  $x \in G$  **total differenzierbar** (oder einfach nur **differenzierbar**), wenn es eine lineare Abbildung

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

gibt und ein  $\delta > 0$  mit  $B_\delta(x) \subset G$ , so dass für  $h \in B_\delta(0)$  und

$$\varphi(h) := f(x+h) - f(x) - Ah$$

gilt

$$\lim_{h \rightarrow 0} \frac{\varphi(h)}{\|h\|} = 0.$$

**2.33 Bemerkung.** (a) Ist  $\varphi : \mathbb{R}^n \supset B_\delta(0) \rightarrow \mathbb{R}^m$  eine Funktion und  $k \in \mathbb{N}_0$ , so sagt man, dass  $\varphi$  von höherer als  $k$ -ter Ordnung in 0 verschwindet und schreibt  $\varphi = o(\|h\|^k)$ , falls

$$\lim_{h \rightarrow 0} \frac{\varphi(h)}{\|h\|^k} = 0.$$

Mit dieser Schreibweise ist also dann  $f : G \rightarrow \mathbb{R}^m$  differenzierbar in  $x$  wenn sich  $f$  in  $x$  bis auf einen Fehler der Ordnung  $o(\|h\|)$  linear approximieren läßt, also wenn

$$f(x+h) = f(x) + Ah + o(\|h\|).$$

(b) **Zur Erinnerung:** Eine lineare Abbildung  $A : V \rightarrow W$  zwischen Vektorräumen  $V$  und  $W$  wird nach Wahl von Basen in  $V$  und  $W$  durch eine  $(m \times n)$ -Matrix  $(a_{ij})$  beschrieben.

Wir werden im Folgenden die Wahl der Basis nur dann explizit machen, wenn wir **nicht** die kanonische Basis  $(e_1, \dots, e_n)$  von  $\mathbb{R}^n$  bzw.  $(e_1, \dots, e_m)$  von  $\mathbb{R}^m$  zugrunde legen.

(c) Natürlich sagen wir, dass  $f : G \rightarrow \mathbb{R}^m$  **total differenzierbar** ist, wenn  $f$  in allen Punkten  $x \in G$  total differenzierbar ist.

**2.34 Beispiel. Quadratische Formen**

Sei  $C = (c_{ij}) \in M(n \times n, \mathbb{R})$  eine symmetrische  $n \times n$ -Matrix und

$$\begin{aligned} f : \mathbb{R}^n &\rightarrow \mathbb{R} \\ x &\mapsto f(x) = \langle x, Cx \rangle = \sum_{i,j=1}^n c_{ij}x_i x_j \end{aligned}$$

die zugehörige quadratische Form. Für  $x, h \in \mathbb{R}^n$  gilt

$$\begin{aligned} f(x+h) &= \langle x+h, Cx+Ch \rangle = \\ &= \langle x, Cx \rangle + \langle h, Cx \rangle + \langle x, Ch \rangle + \langle h, Ch \rangle \\ &= \langle x, Cx \rangle + 2\langle Cx, h \rangle + \langle h, Ch \rangle \\ &= f(x) + Ah + \varphi(h) \end{aligned}$$

mit  $A := 2Cx \in M(1 \times n)$  und  $\varphi(h) = \langle h, Ch \rangle$ .

Da  $|\varphi(h)| \leq \|h\| \cdot \|Ch\| \leq \|C\| \cdot \|h\|^2$ , gilt  $\lim_{h \rightarrow 0} \frac{\varphi(h)}{\|h\|} = 0$ . Also ist  $f$  in  $x$  differenzierbar.

**2.35 Definition. Die Norm einer linearen Abbildung**

Die **Norm** einer linearen Abbildung  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  definiert man durch

$$\|T\| := \sup \{ \|Tx\|_{\mathbb{R}^m} \mid x \in \mathbb{R}^n \text{ mit } \|x\|_{\mathbb{R}^n} \leq 1 \}.$$

**2.36 Bemerkung.** (a) Es gilt  $\|T\| < \infty$ , da  $f : x \mapsto \|Tx\|_{\mathbb{R}^m}$  stetig ist ( $f = \|\cdot\| \circ T$ ) und  $B_1(0) \subset \mathbb{R}^n$  kompakt, also  $f$  als stetige Abbildung auf einem Kompaktum beschränkt ist.

(b) Es gilt für beliebig  $x \in \mathbb{R}^n$ :

$$\|Tx\|_{\mathbb{R}^m} = \|x\|_{\mathbb{R}^n} \|T \frac{x}{\|x\|}\|_{\mathbb{R}^m} \leq \|T\| \|x\|_{\mathbb{R}^n}.$$

**2.37 Satz. und Definition. Stetigkeit diffbarer Funktionen und die Jacobi-Matrix**

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}^m$  eine Abbildung, die im Punkt  $x \in G$  differenzierbar sei, also

$$f(x+h) = f(x) + Ah + o(\|h\|)$$

mit der Matrix  $A = (a_{ij}) \in M(m \times n, \mathbb{R})$ .

Dann gilt:

(a)  $f$  ist im Punkt  $x$  stetig.

(b) Alle Komponenten  $f_i : G \rightarrow \mathbb{R}$ ,  $1 \leq i \leq m$ , von  $f$  sind in  $x$  partiell differenzierbar mit

$$\frac{\partial f_i}{\partial x_j}(x) = a_{ij}.$$

Aus (b) folgt insbesondere, dass die Matrix  $A$  durch die differenzierbare Abbildung  $f$  eindeutig bestimmt ist. Man nennt  $A$  das **Differential**, die **Jacobi-Matrix** oder die **Funktionalmatrix** von  $f$  im Punkte  $x$  und schreibt:

$$(Df)(x) := J_f(x) := \left( \frac{\partial f_i}{\partial x_j}(x) \right)_{ij}.$$

Das Differential  $Df(x)$  ist also die lineare Approximation an  $f$  im Punkt  $x$ :

$$f(x+h) = f(x) + Df(x)h + o(\|h\|).$$

*Beweis.* (a)  $\lim_{h \rightarrow 0} f(x+h) = f(x) + \lim_{h \rightarrow 0} (Ah + o(\|h\|)) = f(x)$ .

(b) Für  $i = 1, \dots, m$  und  $h \in \mathbb{R}^n$  ist

$$f_i(x+h) = f_i(x) + \sum_{j=1}^n a_{ij}h_j + o(\|h\|),$$

also für  $k \in \mathbb{R}$  und  $h = ke_j$

$$f_i(x+ke_j) = f_i(x) + ka_{ij} + o(\|ke_j\|).$$

Damit folgt für die partielle Ableitung von  $f_i$  in Richtung  $e_j$

$$\frac{\partial f_i}{\partial x_j}(x) := \lim_{k \rightarrow 0} \frac{f_i(x+ke_j) - f_i(x)}{k} = a_{ij} + \lim_{k \rightarrow 0} \frac{o(k)}{k} = a_{ij}.$$

□

### 2.38 Satz. Stetig partiell differenzierbar $\Rightarrow$ total differenzierbar

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}$  eine partiell differenzierbare Funktion. Falls alle partiellen Ableitungen  $\partial_j f$  im Punkt  $x \in G$  stetig sind, so ist  $f$  in  $x$  total differenzierbar.

*Beweis.* Für  $h \in \mathbb{R}^n$  hinreichend klein sei

$$z^{(i)} := x + \sum_{j=1}^i h_j e_j, \quad i = 0, \dots, n.$$

Es gilt  $z^{(0)} = x$  und  $z^{(n)} = x+h$ . Die Punkte  $z^{(i-1)}$  und  $z^{(i)}$  unterscheiden sich nur in der  $i$ -ten Koordinate. Nach dem Mittelwertsatz für differenzierbare Funktionen einer Veränderlichen gibt es deshalb ein  $\theta_i \in [0, 1]$ , so dass

$$f(z^{(i)}) - f(z^{(i-1)}) = \partial_i f(y^{(i)})h_i$$

wobei

$$y^{(i)} = z^{(i-1)} + \theta_i h_i e_i.$$

Daraus folgt

$$f(x+h) - f(x) = \sum_{i=1}^n \partial_i f(y^{(i)})h_i$$

Setzt man  $a_i = \partial_i f(x)$ , so gilt

$$f(x+h) = f(x) + \sum_{i=1}^n a_i h_i + \varphi(h)$$

mit

$$\varphi(h) = \sum_{i=1}^n (\partial_i f(y^{(i)}) - a_i) h_i.$$

Wegen der Stetigkeit von  $\partial_i f$  in  $x$  gilt

$$\lim_{h \rightarrow 0} (\partial_i f(y^{(i)}) - a_i) = \partial_i f(\lim_{h \rightarrow 0} y^{(i)}) - a_i = \partial_i f(x) - a_i = 0,$$

also

$$\lim_{h \rightarrow 0} \frac{\varphi(h)}{\|h\|} = 0.$$

□

**2.39 Korollar.** Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}^m$  stetig partiell differenzierbar. Dann ist  $f$  total differenzierbar und somit stetig.

**2.40 Merkregel.** Es gelten also die Implikationen:

$$\begin{array}{c} \text{stetig partiell differenzierbar} \Rightarrow \text{total differenzierbar} \Rightarrow \text{partiell differenzierbar} \\ \Downarrow \\ \text{stetig} \end{array}$$

Die Umkehrungen gelten im Allgemeinen nicht!

Im Folgenden werden wir oft “stetig partiell differenzierbar” durch “stetig differenzierbar” abkürzen, da nach obigem eine total differenzierbare Funktion mit stetiger Ableitung ja insbesondere stetige partielle Ableitungen hat.

**2.41 Satz. Kettenregel**

Seien  $G \subset \mathbb{R}^n$  und  $H \subset \mathbb{R}^m$  Gebiete und  $g : G \rightarrow \mathbb{R}^m$  und  $f : H \rightarrow \mathbb{R}^k$  Abbildungen mit  $g(G) \subset H$ . Die Abbildung  $g$  sei im Punkt  $x \in G$  differenzierbar und die Abbildung  $f$  sei im Punkt  $y := g(x)$  differenzierbar. Dann ist die Komposition

$$f \circ g : G \rightarrow \mathbb{R}^k$$

im Punkt  $x$  differenzierbar und für ihr Differential gilt:

$$D(f \circ g)(x) = \underbrace{Df(g(x))}_{k \times m\text{-Matrix}} \cdot \underbrace{Dg(x)}_{m \times n\text{-Matrix}}$$

*Beweis.* Sei  $A := Dg(x)$  und  $B = Df(y)$ . Es ist zu zeigen, dass  $D(f \circ g)(x) = BA$ . Nach Voraussetzung gelten

$$\begin{aligned} g(x+h) &= g(x) + Ah + \varphi(h) \\ f(y+\eta) &= f(y) + B\eta + \psi(\eta) \end{aligned}$$

mit  $\varphi(h) = o(\|h\|)$  und  $\psi(\eta) = o(\|\eta\|)$ . Wählt man

$$\eta := g(x+h) - g(x) = Ah + \varphi(h)$$

so ergibt sich

$$\begin{aligned} (f \circ g)(x+h) &= f(g(x) + \eta) = f(g(x) + Ah + \varphi(h)) \\ &= f(g(x)) + BAh + B\varphi(h) + \psi(Ah + \varphi(h)) \\ &= (f \circ g)(x) + BAh + \chi(h) \end{aligned}$$

mit  $\chi(h) = B\varphi(h) + \psi(Ah + \varphi(h))$ . Es bleibt also zu zeigen, dass  $\chi(h) = o(\|h\|)$ . Mit  $\varphi(h) = o(\|h\|)$  ist auch  $B\varphi(h) = o(\|h\|)$ . Außerdem gibt es eine Konstante  $K > 0$ , so dass  $\|\varphi(h)\| \leq K\|h\|$  für alle hinreichend kleinen  $h$ . Wegen  $\psi(\eta) = o(\|\eta\|)$  gilt  $\psi(\eta) = \|\eta\|\psi_1(\eta)$  mit  $\lim_{\eta \rightarrow 0} \psi_1(\eta) = 0$ . Damit ergibt sich

$$\|\psi(Ah + \varphi(h))\| \leq (\|A\| + K)\|h\| \cdot \|\psi_1(Ah + \varphi(h))\|,$$

also

$$\lim_{h \rightarrow 0} \frac{\psi(Ah + \varphi(h))}{\|h\|} = 0.$$

□

*Beweis.* [Beweis von Satz 2.19] Sei  $g : (-\delta, \delta) \rightarrow G$  definiert durch  $g(t) := x + tv$  und  $h := f \circ g : (-\delta, \delta) \rightarrow \mathbb{R}$ . Nach Definition der Richtungsableitung ist

$$\partial_v f(x) = \frac{d}{dt} f(x + tv)|_{t=0} = \frac{d}{dt} h(0).$$

Aus der Kettenregel folgt

$$\begin{aligned} \frac{d}{dt} h(t) &= \underbrace{Df(g(t))}_{1 \times n} \cdot \underbrace{Dg(t)}_{n \times 1} = \sum_{j=1}^n \underbrace{\frac{\partial f}{\partial x_j}(g(t))}_{(\text{grad } f)_j} \underbrace{\frac{dg_j}{dt}(t)}_{=v_j} \\ &= \langle v, \text{grad } f(g(t)) \rangle, \end{aligned}$$

also  $\frac{d}{dt} h(0) = \langle v, \text{grad } f(x) \rangle$ .

□

**2.42 Bemerkung.** Für stetig differenzierbares  $f : G \rightarrow \mathbb{R}^m$  ist die Richtungsableitung somit gegeben durch

$$\partial_v f(x) = \begin{pmatrix} \partial_v f_1(x) \\ \vdots \\ \partial_v f_m(x) \end{pmatrix} = \begin{pmatrix} \nabla f_1(x) \\ \vdots \\ \nabla f_m(x) \end{pmatrix} \cdot v = Df(x) \cdot v,$$

wobei das Produkt das Matrix-Vektor Produkt ist.

**2.43 Korollar.** Sei  $G$  ein Gebiet,  $f : G \rightarrow \mathbb{R}^m$  stetig partiell differenzierbar und  $Df(x) = 0$  für alle  $x \in G$ . Dann ist  $f$  auf  $G$  konstant.

*Beweis.* Sei  $x_0 \in G$  und  $h \in \mathbb{R}^n$  so dass  $x_0 + th \in G$  für alle  $t \in [0, 1]$ . Dann gilt für  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  nach dem Mittelwertsatz, angewandt auf die Funktion  $g(t) = f(x_0 + th)$  ein  $\theta$  gibt sodass

$$f(x_0 + th) = f(x_0) + \nabla f(x_0 + \theta h) \cdot h = f(x_0).$$

Somit folgt die Aussage für konvexe Bereiche. Da Gebiete wegzusammenhängend sind und alle Wege durch Polygonzüge approximiert werden können, lässt sich die Aussage unmittelbar auf Gebiete verallgemeinern. Die Verallgemeinerung auf Vektorwertige Funktionen  $f$  ist offensichtlich. □

## 2.44 Beispiel. Der Laplaceoperator in Polarkoordinaten

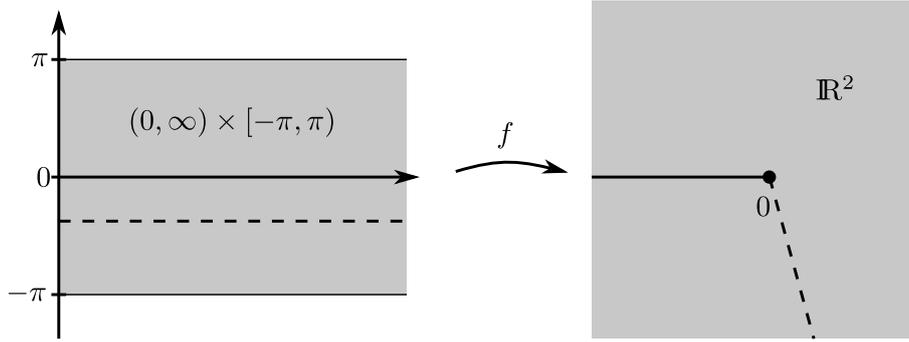
Die Abbildung

$$f : (0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}^2 \setminus \{0\}, \quad (r, \varphi) \mapsto f(r, \varphi) = (r \cos \varphi, r \sin \varphi) =: (x(r, \varphi), y(r, \varphi))$$

versieht den  $\mathbb{R}^2 \setminus \{0\}$  mit Polarkoordinaten. Eingeschränkt z.B. auf  $(0, \infty) \times [-\pi, \pi)$  ist  $f$  sogar bijektiv. Sei  $G \subset \mathbb{R}^2$  ein Gebiet und  $u \in C^2(G, \mathbb{R})$ , dann drückt man  $u$  in Polarkoordinaten aus, indem man  $u \circ f$  betrachtet, also  $u$  als Funktion von  $(r, \varphi)$  schreibt. Genauso ist  $\Delta u$  in Polarkoordinaten durch  $(\Delta u) \circ f$  gegeben. Ziel ist es nun,  $(\Delta u) \circ f$  durch Differentiation an  $u \circ f$  ausdrücken. Und tatsächlich gilt

$$(\Delta u) \circ f = \frac{\partial^2(u \circ f)}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2(u \circ f)}{\partial \varphi^2} + \frac{1}{r} \frac{\partial(u \circ f)}{\partial r}. \quad (2.2)$$

*Beweis.* Nachrechnen mit Kettenregel (Übungsaufgabe). □



Aber wie kommt man auf den Ausdruck (2.2)? Dazu betrachten wir zunächst das Differential von  $u$  in Polarkoordinaten. Mit der Kettenregel gilt

$$D(u \circ f) = Du \circ f \cdot Df, \quad \text{also} \quad Du \circ f = D(u \circ f) \cdot (Df)^{-1}, \quad (2.3)$$

wobei  $(Df)^{-1}$  in unserem Beispiel leicht berechnet werden kann:

$$\begin{aligned} (Df)^{-1} &= \begin{pmatrix} \frac{\partial x}{\partial r}(r, \varphi) & \frac{\partial x}{\partial \varphi}(r, \varphi) \\ \frac{\partial y}{\partial r}(r, \varphi) & \frac{\partial y}{\partial \varphi}(r, \varphi) \end{pmatrix}^{-1} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}^{-1} \\ &= \frac{1}{r} \begin{pmatrix} r \cos \varphi & r \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{\sin \varphi}{r} & \frac{\cos \varphi}{r} \end{pmatrix}. \end{aligned}$$

Wir haben also in (2.3) eine allgemeine Formel für das Differential  $Du$  von  $u$  ausgedrückt in Koordinaten geben durch  $f$ . Dabei benötigen wir nur die Invertierbarkeit von  $Df$ . In unserem Beispiel ergibt sich

$$\begin{aligned} Du \circ f &= \left( \frac{\partial(u \circ f)}{\partial r}, \frac{\partial(u \circ f)}{\partial \varphi} \right) \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{\sin \varphi}{r} & \frac{\cos \varphi}{r} \end{pmatrix} \\ &= \left( \cos \varphi \frac{\partial(u \circ f)}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial(u \circ f)}{\partial \varphi}, \sin \varphi \frac{\partial(u \circ f)}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial(u \circ f)}{\partial \varphi} \right) \\ &=: \underbrace{\left( \cos \varphi \frac{\partial u}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial u}{\partial \varphi} \right)}_{\frac{\partial u}{\partial x} \circ f}, \underbrace{\left( \sin \varphi \frac{\partial u}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial u}{\partial \varphi} \right)}_{\frac{\partial u}{\partial y} \circ f} = \text{grad } u \circ f, \end{aligned} \quad (2.4)$$

wobei hier die Komponenten bezüglich der kanonischen Basis des  $\mathbb{R}^2$  stehen und  $\text{grad } u$  als Zeilenvektor aufgefasst wird. Im letzten Schritt und im Folgenden unterdrücken wir manchmal  $f$ , d.h. an Stelle von  $(u \circ f)(r, \varphi)$  schreiben wir einfach  $u(r, \varphi)$  und entsprechend  $\partial_r u(r, \varphi)$  statt  $\partial_r(u \circ f)(r, \varphi)$ . In dieser verkürzten Notation können wir nun auch  $\Delta u \circ f$  berechnen:

$$\begin{aligned} \frac{\partial^2}{\partial x^2} u &= \left( \cos \varphi \frac{\partial}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial}{\partial \varphi} \right)^2 u = \left( \cos \varphi \frac{\partial}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial}{\partial \varphi} \right) \left( \cos \varphi \frac{\partial u}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial u}{\partial \varphi} \right) \\ &= \cos^2 \varphi \frac{\partial^2 u}{\partial r^2} + \frac{\cos \varphi \sin \varphi}{r^2} \frac{\partial u}{\partial \varphi} - \frac{\cos \varphi \sin \varphi}{r} \frac{\partial^2 u}{\partial r \partial \varphi} \\ &\quad - \frac{\cos \varphi \sin \varphi}{r} \frac{\partial^2 u}{\partial r \partial \varphi} + \frac{\sin^2 \varphi}{r} \frac{\partial u}{\partial r} + \frac{\cos \varphi \sin \varphi}{r^2} \frac{\partial u}{\partial \varphi} + \frac{\sin^2 \varphi}{r^2} \frac{\partial^2 u}{\partial \varphi^2} \end{aligned}$$

und analog

$$\begin{aligned} \frac{\partial^2}{\partial y^2} u &= \left( \sin \varphi \frac{\partial}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial}{\partial \varphi} \right)^2 u = \left( \sin \varphi \frac{\partial}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial}{\partial \varphi} \right) \left( \sin \varphi \frac{\partial u}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial u}{\partial \varphi} \right) \\ &= \sin^2 \varphi \frac{\partial^2 u}{\partial r^2} - \frac{\cos \varphi \sin \varphi}{r^2} \frac{\partial u}{\partial \varphi} + \frac{\cos \varphi \sin \varphi}{r} \frac{\partial^2 u}{\partial r \partial \varphi} \\ &\quad + \frac{\cos \varphi \sin \varphi}{r} \frac{\partial^2 u}{\partial r \partial \varphi} + \frac{\cos^2 \varphi}{r} \frac{\partial u}{\partial r} - \frac{\cos \varphi \sin \varphi}{r^2} \frac{\partial u}{\partial \varphi} + \frac{\cos^2 \varphi}{r^2} \frac{\partial^2 u}{\partial \varphi^2}. \end{aligned}$$

Addiert man die beiden Ausdrücke, so ergibt sich (2.2).

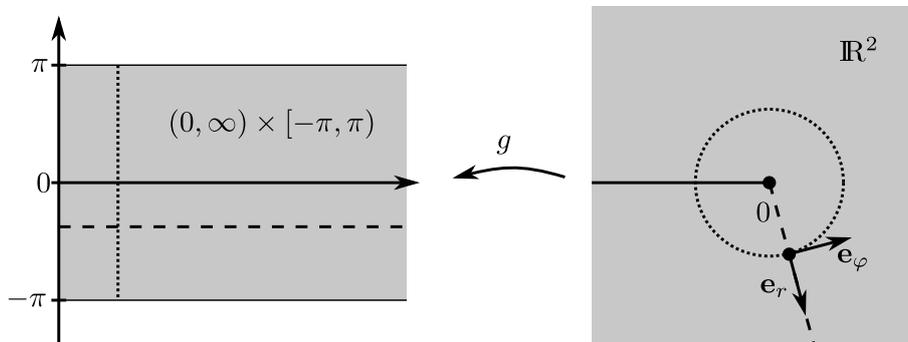
Nochmals zurück zu (2.4): Die Matrix  $Du \circ f$  ist die Linearisierung der Abbildung  $u : \mathbb{R}^2 \rightarrow \mathbb{R}$  im Punkte  $f(r, \varphi)$  bezüglich der kanonischen Basis des  $\mathbb{R}^2$ . Man kann nun aber auch noch versuchen, die lineare Abbildung  $Du \circ f$  bezüglich einer den Koordinaten angepassten Basis darzustellen, nämlich bezüglich der aus der Physik bekannten Basisvektoren  $\mathbf{e}_r$  und  $\mathbf{e}_\varphi$ . Um diese vernünftig zu definieren, schränken wir zunächst  $f$  auf  $(0, \infty) \times (-\pi, \pi)$  ein und erhalten so einen **Diffeomorphismus**

$$f : G := (0, \infty) \times (-\pi, \pi) \rightarrow \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\} =: D$$

Seine Umkehrung ist gegeben durch  $g : D \rightarrow G$ ,  $g(x, y) = (r(x, y), \varphi(x, y))$  mit  $r(x, y) = \sqrt{x^2 + y^2}$  und

$$\varphi(x, y) = \begin{cases} \arctan(x/y) & \text{für } x > 0, y > 0 \\ \pi/2 & \text{für } x = 0, y > 0 \\ \vdots & \end{cases}$$

Wir werden später zeigen, dass die Invertierbarkeit von  $Df$  ganz allgemein die lokale Invertierbarkeit von  $f$  impliziert.



Die Zeilenvektoren von  $Dg$  bilden nun eine Basis des  $\mathbb{R}^2$  und stehen senkrecht auf den jeweiligen Koordinatenlinien, da sie ja gerade durch den Gradienten der Koordinatenfunktionen gegeben sind,

$$Dg = \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \varphi}{\partial x} & \frac{\partial \varphi}{\partial y} \end{pmatrix} = \begin{pmatrix} \text{grad } r \\ \text{grad } \varphi \end{pmatrix}.$$

Um nun  $Dg$  auszurechnen, differenziert man nicht etwa  $g$  (was bei komplizierten  $f$ 's oft gar nicht explizit geht), sondern verwendet nochmals die Kettenregel. Weil  $g \circ f = \text{id}$  ist, ist nach der Kettenregel

$$Dg(f(r, \varphi)) \cdot Df(r, \varphi) = D(\text{id})(r, \varphi) = E,$$

wobei  $E$  wie immer die Einheitsmatrix bezeichnet. Also ist

$$Dg \circ f = (Df)^{-1},$$

was wir für unser Beispiel oben schon berechnet haben. Der Basiswechsel von der neuen Basis in die kanonische ist durch  $(Dg)^T = ((Df)^{-1})^T$  gegeben. Multiplikation von (2.4) durch  $((Df)^{-1})^T$  von rechts (nur im Urbildraum wird die Basis gewechselt) liefert das Differential  $Du$  als Matrix bezüglich der durch  $f$  beschriebenen Basisvektoren

$$(Du \circ f)(Dg \circ f)^T = D(u \circ f) \cdot (Df)^{-1} \cdot ((Df)^{-1})^T.$$

Im Beispiel der Polarkoordinaten ergibt sich

$$\text{grad } u = \left( \cos \varphi \frac{\partial u}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial u}{\partial \varphi}, \sin \varphi \frac{\partial u}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial u}{\partial \varphi} \right) \begin{pmatrix} \cos \varphi & -\frac{\sin \varphi}{r} \\ \sin \varphi & \frac{\cos \varphi}{r} \end{pmatrix} = \left( \frac{\partial u}{\partial r}, \frac{1}{r^2} \frac{\partial u}{\partial \varphi} \right),$$

also  $\text{grad } u = \partial_r u \text{ grad } r + r^{-2} \partial_\varphi u \text{ grad } \varphi$ , bzw. nach Normierung mit  $\mathbf{e}_r = \text{grad } r$  und  $\mathbf{e}_\varphi = r^{-1} \text{grad } \varphi$

$$\text{grad } u = \frac{\partial u}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial u}{\partial \varphi} \mathbf{e}_\varphi.$$

## 2.1 Taylor Formel und lokale Extrema

Wir haben gezeigt, dass für eine zweimal stetig differenzierbare Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  die Taylor Entwicklung

$$\begin{array}{cccc}
 f(x+h) & = & f(x) & + & f'(x) \cdot h & + & \frac{1}{2} f''(x) \cdot h^2 & + & o(\|h\|^2) \\
 & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\
 & & 0. \text{ Ordnung} & & 1. \text{ Ordnung} & & 2. \text{ Ordnung} & & \text{Fehler höherer} \\
 & & = \text{konst.} & & = \text{linear} & & = \text{quadratisch} & & \text{Ordnung}
 \end{array}$$

eine lokale quadratische Approximation an die Funktion  $f$  liefert. Für Funktionen  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  ist die erste Ableitung  $Df(x) = \text{grad}f(x)^T$  ein Zeilenvektor und die zweite Ableitung  $\text{Hess}f(x)$  eine Matrix. Die naheliegende Verallgemeinerung der Taylorsche Formel für solche  $f$  ist also

$$\begin{aligned}
 f(x+h) &= f(x) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) h_j + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i h_j + o(\|h\|^2) \\
 &= f(x) + \langle \text{grad}f(x), h \rangle + \frac{1}{2} \langle h, \text{Hess}f(x)h \rangle + o(\|h\|^2).
 \end{aligned}$$

### 2.45 Notation. Multiindices und iterierte Richtungsableitung

Um die höheren Terme der Taylor-Entwicklung günstig zu notieren, führt man folgende Schreibweisen ein:

(a) **Multiindices:** Für  $\alpha \in \mathbb{N}_0^n$ ,  $\alpha = (\alpha_1, \dots, \alpha_n)$ , sei

$$\begin{aligned}
 |\alpha| &= \alpha_1 + \dots + \alpha_n = \sum_{j=1}^n \alpha_j \\
 \alpha! &= \alpha_1! \cdots \alpha_n! = \prod_{j=1}^n \alpha_j! ,
 \end{aligned}$$

für eine  $|\alpha|$ -mal stetig partiell differenzierbare Funktion  $f : G \rightarrow \mathbb{R}$  sei

$$\partial^\alpha f := \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} ,$$

und schließlich für  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n} = \prod_{j=1}^n x_j^{\alpha_j} .$$

(b) **Iterierte Richtungsableitung:** Für  $f \in C^k(G)$  und  $h \in \mathbb{R}^n$  sei

$$(h \cdot \nabla)^k f(x) := \partial_h^k f(x) := \sum_{j_1=1}^n \cdots \sum_{j_k=1}^n h_{j_1} \cdots h_{j_k} \partial_{j_1} \cdots \partial_{j_k} f(x) .$$

**2.46 Lemma.** Sei  $G \subset \mathbb{R}^n$  ein Gebiet,  $f : G \rightarrow \mathbb{R}$  eine  $k$ -mal stetig partiell differenzierbare Funktion,  $x \in G$  und  $h \in \mathbb{R}^n$  derart, dass die geradlinige Verbindung von  $x$  nach  $x+h$  ganz in  $G$  verläuft, also  $\{x+ht \mid t \in [0,1]\} \subset G$ . Dann ist die Funktion  $\varphi : [0,1] \rightarrow \mathbb{R}$ ,  $\varphi(t) = f(x+th)$  auch  $k$ -mal stetig differenzierbar und es gilt

$$\frac{d^k \varphi}{dt^k}(t) = ((h \cdot \nabla)^k f)(x+th) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(x+th) h^\alpha .$$

**2.47 Bemerkung.** (a) Die Notation  $\sum_{|\alpha|=k}$  bedeutet, dass sich die Summe über alle  $n$ -Tupel  $\alpha \in \mathbb{N}_0^n$  erstreckt, für die  $|\alpha| = k$  ist. Davon gibt es  $\binom{n-1+k}{k}$  Stück.

(b) Ist  $\alpha = (\alpha_1, \dots, \alpha_n)$  mit  $k = |\alpha| = \alpha_1 + \dots + \alpha_n$ , so gibt es  $\frac{k!}{\alpha!}$  Möglichkeiten, eine  $k$ -elementige Menge  $M$  in  $n$  disjunkte Teilmengen  $S_1, \dots, S_n$  zu zerlegen,  $M = S_1 \dot{\cup} \dots \dot{\cup} S_n$ , so dass  $S_i$  gerade  $\alpha_i$  Elemente hat ( $i = 1, \dots, n$ ). Oder anders formuliert: es gibt  $\frac{k!}{\alpha!}$  Möglichkeiten  $k$  verschiedene Kugeln auf  $n$  Urnen  $S_1, \dots, S_n$  zu verteilen, sodass in der  $j$ -ten Urne genau  $\alpha_j$  Kugeln liegen. Für  $n = 2$  ist beispielsweise

$$\frac{k!}{\alpha_1! \alpha_2!} = \frac{k!}{\alpha_1(k-\alpha_1)!} = \binom{k}{\alpha_1}.$$

*Beweis.* von Lemma 2.46. Nach der Kettenregel ist

$$\frac{d}{dt} \varphi(t) = Df(x+th)h = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x+th) h_j.$$

Nochmals die Kettenregel liefert

$$\frac{d^2}{dt^2} \varphi(t) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(x+th) h_j h_i = \sum_{i_1, i_2=1}^n \partial_{i_1} \partial_{i_2} f(x+th) h_{i_1} h_{i_2},$$

und  $k$ -malige Anwendung schließlich

$$\frac{d^k}{dt^k} \varphi(t) = \sum_{i_1, \dots, i_k=1}^n \partial_{i_1} \dots \partial_{i_k} f(x+th) h_{i_1} \dots h_{i_k} = ((h \cdot \nabla)^k f)(x+th).$$

Da die Reihenfolge der Differentiationen aber gemäß Satz 2.24 keine Rolle spielt, fassen wir Terme, in denen  $\alpha_1$ -mal nach der ersten Koordinate,  $\alpha_2$ -mal nach der zweiten Koordinate etc. abgeleitet wird, zusammen. Es gibt nun nach Bemerkung 2.47 gerade  $\frac{k!}{\alpha_1! \dots \alpha_n!}$  solche  $k$ -Tupel  $(i_1, \dots, i_k)$  in denen  $\alpha_1$ -mal der Wert 1,  $\alpha_2$ -mal der Wert 2,  $\dots$ , und  $\alpha_n$ -mal der Wert  $n$  vorkommt. Durch Zusammenfassen der Summanden ergibt sich also

$$\frac{d^k}{dt^k} \varphi(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(x+th) h^\alpha.$$

□

### 2.48 Satz. Satz von Taylor

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}$  eine  $(k+1)$ -mal stetig partiell differenzierbare Funktion. Sei  $x \in G$  und  $h \in \mathbb{R}^n$  derart, dass die Strecke  $[x, x+h] := \{x+th \mid t \in [0, 1]\}$  ganz in  $G$  liegt. Dann gibt es ein  $\theta \in [0, 1]$  so, dass

$$\begin{aligned} f(x+h) &= \sum_{m=0}^k \frac{((h \cdot \nabla)^m f)(x)}{m!} + \frac{((h \cdot \nabla)^{k+1} f)(x+\theta h)}{(k+1)!} \\ &= \sum_{m=0}^k \sum_{|\alpha|=m} \frac{\partial^\alpha f}{\alpha!}(x) h^\alpha + \sum_{|\alpha|=k+1} \frac{\partial^\alpha f}{\alpha!}(x+\theta h) h^\alpha. \end{aligned}$$

**2.49 Bemerkung.** Man nennt

$$P_{f,x}^{(k)}(h) := \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha$$

das **Taylorpolynom vom Grad  $k$  von  $f$  in  $x$** . Schreiben wir

$$P_{f,x}^{(k)}(h) := P_0(h) + P_1(h) + \dots + P_k(h),$$

so ist

$$\begin{aligned} P_0(h) &= f(x) \\ P_1(h) &= \partial_1 f(x) h_1 + \cdots + \partial_n f(x) h_n = \langle \text{grad} f(x), h \rangle \\ P_2(h) &= \sum_{\alpha_1 + \alpha_2 = 2} \frac{1}{\alpha_1! \alpha_2!} \partial_1^{\alpha_1} \partial_2^{\alpha_2} f(x) h_1^{\alpha_1} h_2^{\alpha_2} \\ &= \frac{1}{2} \left( \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i h_j \right) = \frac{1}{2} \langle h, \text{Hess} f(x) h \rangle. \end{aligned}$$

Für Funktionen  $f : G \rightarrow \mathbb{R}^m$  erhält man so eine Taylorentwicklung für jede Komponente  $f_j, j = 1 \dots, m$ .

*Beweis. von Satz 2.48.* Betrachte die Kurve  $\gamma : [0, 1] \rightarrow G, \gamma(t) = x + th$ , und setze  $\varphi : [0, 1] \rightarrow \mathbb{R}, \varphi(t) = f \circ \gamma(t)$ . Dann ist  $\varphi$  eine  $(k + 1)$ -mal stetig differenzierbare Funktion, also existiert nach der Taylor-Formel für Funktionen in einer Veränderlichen ein  $\theta \in [0, 1]$ , so dass gilt

$$\varphi(1) = \sum_{m=0}^k \frac{\varphi^{(m)}(0)}{m!} + \frac{\varphi^{(k+1)}(\theta)}{(k+1)!}.$$

Nach Lemma 2.46 erhält man

$$\begin{aligned} f(x+h) = \varphi(1) &= \sum_{m=0}^k \frac{((h \cdot \nabla)^m f)(x)}{m!} + \frac{((h \cdot \nabla)^{k+1} f)(x + \theta h)}{(k+1)!} \\ &= \underbrace{\sum_{m=0}^k \sum_{|\alpha|=m} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha}_{\sum_{|\alpha| \leq k}} + \sum_{|\alpha|=k+1} \frac{\partial^\alpha f(x + \theta h)}{\alpha!} h^\alpha. \end{aligned}$$

□

**2.50 Korollar.** Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f \in C^k(G, \mathbb{R})$ . Dann gilt für jedes  $x \in G$  und  $\delta > 0$  mit  $B_\delta(x) \subset G$ , dass für  $h \in B_\delta(0)$

$$f(x+h) = P_{f,x}^{(k)}(h) + o(\|h\|^k).$$

*Beweis.* Nach Taylors Satz gibt es für jedes  $h \in B_\delta(0)$  ein  $\theta = \theta(h) \in [0, 1]$ , so dass

$$f(x+h) = \sum_{|\alpha| \leq k-1} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha + \sum_{|\alpha|=k} \frac{\partial^\alpha f(x + \theta h)}{\alpha!} h^\alpha.$$

Wir setzen  $\varphi : B_\delta(0) \rightarrow \mathbb{R}$  als

$$\varphi(h) := \sum_{|\alpha|=k} \left( \frac{\partial^\alpha f(x + \theta h)}{\alpha!} - \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha \right)$$

Wegen der Stetigkeit von  $\partial^\alpha f$  ist

$$\lim_{h \rightarrow 0} \frac{1}{\alpha!} (\partial^\alpha f(x + \theta h) - \partial^\alpha f(x)) = 0$$

und deshalb gilt

$$\frac{\| \frac{1}{\alpha!} (\partial^\alpha f(x + \theta h) - \partial^\alpha f(x)) \cdot h^\alpha \|}{\|h\|^k} \leq \frac{1}{\alpha!} (\partial^\alpha f(x + \theta h) - \partial^\alpha f(x)) \cdot \frac{\|h^\alpha\|}{\|h\|^k} \rightarrow 0$$

für  $h \rightarrow 0$ , denn  $\frac{\|h^\alpha\|}{\|h\|^k} = \frac{|h_1|^{\alpha_1} \cdots |h_n|^{\alpha_n}}{\|h\|^{\alpha_1 + \cdots + \alpha_n}} \leq 1$ . Also ist

$$f(x+h) = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha + \varphi(h)$$

mit  $\varphi(h) = o(\|h\|^k)$ .

□

**2.51 Bemerkung.** (a) Für  $k = 0$  erhält man die Aussage, dass eine stetige Funktion stetig in  $x$  ist, denn

$$f(x+h) = \sum_{|\alpha| \leq 0} \frac{\partial^\alpha f}{\alpha!}(x) + o(\|h\|) = f(x) + o(1),$$

also  $\lim_{h \rightarrow 0} f(x+h) = f(x)$ .

(b) Für  $k = 1$  erhält man die Aussage, dass eine stetig differenzierbare Funktion in  $x$  stetig differenzierbar ist, denn

$$f(x+h) = \sum_{|\alpha| \leq 1} \frac{\partial^\alpha f}{\alpha!}(x) + o(\|h\|) = f(x) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) h_i + o(\|h\|) = f(x) + Df(x)h + o(\|h\|).$$

(c) Aber für  $k = 2$  erhält man nun die sehr nützliche Aussage, dass eine 2-mal stetig differenzierbare Funktion die folgende Darstellung erlaubt:

$$f(x+h) = \sum_{|\alpha| \leq 2} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha + o(\|h\|^2) = f(x) + \langle \text{grad} f(x), h \rangle + \frac{1}{2} \langle h, \text{Hess} f(x) h \rangle + o(\|h\|^2).$$

### 2.52 Definition. Lokale Extrema

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $x \in G$ . Man sagt, dass eine Funktion  $f : G \rightarrow \mathbb{R}$  in  $x$  ein **lokales Maximum** hat, wenn es ein  $\delta > 0$  gibt, so dass für alle  $y \in B_\delta(x)$  gilt

$$f(y) \leq f(x).$$

Ist  $f(y) \geq f(x)$  für alle  $y \in B_\delta(x)$ , so spricht man von einem **lokalen Minimum**. Hat  $f$  in  $x$  ein lokales Maximum oder Minimum, so spricht man von einem **lokalen Extremum**.

**2.53 Proposition.** Sei  $f : G \rightarrow \mathbb{R}$ ,  $x \in G$  und  $f$  habe in  $x$  ein lokales Extremum. Falls  $f$  in  $x$  partiell differenzierbar ist, so gilt

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

*Beweis.* Definiert man für  $\delta > 0$  klein genug die Funktionen  $h_i : (-\delta, \delta) \rightarrow \mathbb{R}$  durch

$$h_i(t) = f(x + te_i),$$

so ist

$$\frac{d}{dt} h_i(0) = \frac{\partial f}{\partial x_i}(x).$$

Mit  $f$  hat aber auch  $h_i$  in 0 ein lokales Extremum, also gilt  $\frac{d}{dt} h_i(0) = 0$  für  $i = 1, \dots, n$ . Somit ist auch  $\text{grad} f(x) = 0$ .  $\square$

Wie im eindimensionalen Fall ist  $\nabla f(x) = 0$  nur ein notwendiges, aber kein hinreichendes Kriterium für das Vorliegen eines lokalen Extremums. Die Lösungen der Gleichung  $\nabla f(x) = 0$  liefern somit die Kandidaten für die lokalen Extrema. Um ein hinreichendes Kriterium zu finden, betrachtet man wie im Fall von  $f : \mathbb{R} \rightarrow \mathbb{R}$  die zweite Ableitung.

Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Dann ist, wie oben schon erwähnt, die zweite Ableitung von  $f$  in Richtung  $v$  gegeben durch

$$\left. \frac{d^2}{dt^2} \right|_{t=0} f(x + tv) = \langle v, \text{Hess} f(x) v \rangle.$$

Falls nun  $\nabla f(x) = 0$ , dann kann somit garantiert werden, dass  $f(x)$  ein lokales Minimum ist, falls für alle  $v \in \mathbb{R}^n$  gilt dass

$$\langle v, \text{Hess} f(x) v \rangle > \gamma_0 \|v\|^2,$$

für ein  $\gamma_0 > 0$ . Dies ist garantiert falls alle Eigenwerte von  $\text{Hess} f(x)$  positiv sind. Wir werden sehen, dass symmetrische Matrizen immer rein reelle Eigenwerte besitzt und diese für Diagonalmatrizen direkt von Einträgen abgelesen werden können. Deshalb ist offensichtlich, dass  $f(x, y) = x^2 + y^2$  den Ursprung als lokales Minimum hat, wogegen dies bei  $g(x, y) = x^2 - y^2$  nicht der Fall ist. Um allgemeinere Probleme behandeln zu können werden wir zwei Kapitel der linearen Algebra einschieben.



# 3 Einschub: Lineare Algebra

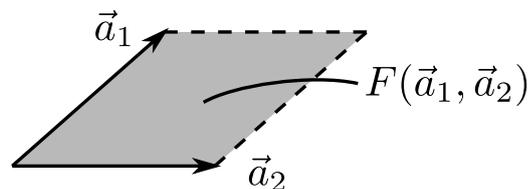
## 3.1 Determinanten

### Die $2 \times 2$ -Determinante

Das lineare Gleichungssystem  $Ax = b$  mit  $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = (\vec{a}_1, \vec{a}_2)$  und  $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$  ist genau dann universell und eindeutig lösbar, wenn  $A$  bijektiv ist, also wenn  $\text{Span}\{\vec{a}_1, \vec{a}_2\} = \mathbb{R}^2$  ist, d.h. wenn  $\vec{a}_1$  und  $\vec{a}_2$  linear unabhängig sind.

Wann sind zwei Vektoren in  $\mathbb{R}^2$  linear unabhängig?

Genau dann, wenn das von ihnen aufgespannte Parallelogramm nicht entartet ist, also wenn die (orientierte) Fläche



$$F(\vec{a}_1, \vec{a}_2) = |\vec{a}_1||\vec{a}_2| \sin \varphi = |\vec{a}_1||\vec{a}_2| \cos(\varphi + \frac{\pi}{2}) = \vec{a}_1 \cdot \vec{a}_2^\perp = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \cdot \begin{pmatrix} a_{22} \\ a_{12} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

ungleich Null ist.

Für zwei Vektoren  $x, y \in \mathbb{R}^2$  nennt man

$$\det(x, y) = x_1y_2 - x_2y_1$$

die Determinante bzw. die Abbildung

$$\det : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto \det(x, y)$$

die Determinantenform auf  $\mathbb{R}^2$ .

Die Determinante der Matrix  $A = (\vec{a}_1, \vec{a}_2)$  ist  $\det A := \det(\vec{a}_1, \vec{a}_2)$ .

Die Determinantenform ist

- (a) linear in jedem Argument, d.h.  $x \mapsto \det(x, y)$  und  $y \mapsto \det(x, y)$  sind Linearformen.
- (b) antisymmetrisch bzw. alternierend, d.h.  $\det(x, y) = -\det(y, x)$ .
- (c) normiert, d.h.  $\det(e_1, e_2) = 1$  für die kanonischen Einheitsvektoren.

### Die $3 \times 3$ Determinante

Auch in drei Dimensionen sind Vektoren genau dann linear unabhängig, wenn das von ihnen aufgespannte Volumen

$$\text{Vol}(x, y, z) = x \cdot (y \times z)$$

von Null verschieden ist. Hier ist

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3, (x, y) \mapsto x \times y := \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}$$

das Vektorprodukt im  $\mathbb{R}^3$ . Der Vektor  $x \times y$  steht senkrecht auf der von  $x$  und  $y$  aufgespannten Ebene und sein Betrag  $|x \times y|$  liefert den Flächeninhalt des von  $x$  und  $y$  aufgespannten Parallelogramms.

Wieder ist die Abbildung

$$\det : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}, \quad (x, y, z) \mapsto x \cdot (y \times z)$$

eine alternierende, normierte Multilinearform, d.h.

- (a)  $\det$  ist linear in jeder Variablen,
- (b) Vertauschung zweier Argumente verändert das Vorzeichen,  $\det(x, y, z) = -\det(y, x, z) = \det(y, z, x)$  usw.,
- (c)  $\det(e_1, e_2, e_3) = 1$ .

Wieder kann man auch die Determinante einer  $3 \times 3$ -Matrix  $A = (\vec{a}_1, \vec{a}_2, \vec{a}_3)$ ,  $\vec{a}_j \in \mathbb{R}^3$ , durch die Determinante der Spaltenvektoren definieren:  $\det A := \det(\vec{a}_1, \vec{a}_2, \vec{a}_3)$ .

Man könnte nun die Determinante in  $\mathbb{R}^n$  mittels des von  $n$  Vektoren aufgespannten Volumens definieren. Das ist aber unpraktisch und im Moment ist ja auch noch gar nicht klar, wie dieses Volumen genau definiert ist. Stattdessen nehmen wir deshalb die Eigenschaften (a), (b) und (c) als Grundlage der Definition.

Es gibt einen guten Grund, allgemein Determinanten auch für komplexwertige Funktionen zu definieren. Deshalb bezeichnet  $\mathbb{K}^n$  im weiteren entweder  $\mathbb{R}$  oder  $\mathbb{C}$ .

### 3.1 Definition. Determinantenform

Eine Abbildung

$$F : \underbrace{\mathbb{K}^n \times \cdots \times \mathbb{K}^n}_{m\text{-mal}} \rightarrow \mathbb{K} \quad (n \geq 2)$$

heißt

- (a) **Multilinearform auf  $\mathbb{K}^n$** , kurz  **$m$ -Form**, wenn  $F$  in jedem der  $m$  Argumente linear ist:

$$F(x_1, \dots, x_j + \alpha y, \dots, x_m) = F(x_1, \dots, x_j, \dots, x_m) + \alpha F(x_1, \dots, y, \dots, x_m)$$

für alle  $x_1, \dots, x_m, y \in \mathbb{K}^n$ ,  $\alpha \in \mathbb{K}$ ,  $j = 1, \dots, m$ .

- (b) **Alternierende  $m$ -Form auf  $\mathbb{K}^n$** , wenn (a) gilt und wenn  $F$  beim Vertauschen zweier Argumente das Vorzeichen ändert:

$$F(x_1, \dots, x_i, \dots, x_j, \dots, x_m) = -F(x_1, \dots, x_j, \dots, x_i, \dots, x_m) \quad \forall i \neq j \in \{1, \dots, m\}.$$

- (c) **Determinantenform auf  $\mathbb{K}^n$** , wenn  $m = n$  gilt, (a) und (b) erfüllt sind, und  $F(e_1, \dots, e_n) = 1$  für die kanonische Basis  $(e_1, \dots, e_n)$  des  $\mathbb{K}^n$  gilt.

**3.2 Satz.** Für eine alternierende  $m$ -Form  $F$  auf  $\mathbb{K}^n$  gilt: Sind  $a_1, \dots, a_m$  linear abhängig, so folgt  $F(a_1, \dots, a_m) = 0$ . Insbesondere ist  $F(a_1, \dots, a_m) = 0$ , falls zwei Einträge gleich sind.

*Beweis.* Für zwei gleiche Einträge ergibt das Vertauschen  $F(\dots, a, \dots, a, \dots) = -F(\dots, a, \dots, a, \dots)$  also  $F(\dots, a, \dots, a, \dots) = 0$ .

Seien  $a_1, \dots, a_m \in \mathbb{K}^n$  linear abhängig, etwa  $a_1 = \sum_{k=2}^m \alpha_k a_k$ . Dann gilt wegen der Linearität im ersten Argument

$$F(a_1, \dots, a_m) = \sum_{k=2}^m \alpha_k F(a_k, a_2, \dots, a_k, \dots, a_m) = 0.$$

□

Es folgt, dass jede alternierende  $m$ -Form auf  $\mathbb{K}^n$  mit  $m > n$  identisch Null sein muss.

### 3.3 Satz. Kriterium für das Alternieren einer $m$ -Form

Verschwindet eine  $m$ -Form  $F(a_1, \dots, a_m)$  falls zwei benachbarte Argumente gleich sind, also

$$a_k = a_{k+1} \text{ für ein } k \in \{1, \dots, m-1\} \Rightarrow F(a_1, \dots, a_m) = 0,$$

so ist sie alternierend.

*Beweis.* Wegen

$$\begin{aligned} 0 &= F(\dots, a + b, a + b, \dots) \\ &= F(\dots, a, a, \dots) + F(\dots, a, b, \dots) + F(\dots, b, a, \dots) + F(\dots, b, b, \dots) \\ &= F(\dots, a, b, \dots) + F(\dots, b, a, \dots) \end{aligned}$$

ändert  $F$  sein Vorzeichen bei Vertauschungen benachbarter Argumente. Sind  $a_j$  und  $a_k$  für  $j < k$  nicht benachbart, so erreicht man durch sukzessives Vertauschen benachbarter Argumente, dass  $a_j$  und  $a_k$  die Plätze tauschen, wobei man das mit  $2(k - j) - 1$  Vertauschungen erreicht,

$$\dots, a_j, \underbrace{a_{j+1}, \dots, a_{k-1}, a_k}_{k-j} \dots$$

Man benötigt beispielsweise  $k - j$  Vertauschungen, um  $a_j$  an die  $k$ -te Stelle zu bringen. Dann steht  $a_k$  an der  $k - 1$ -ten Stelle und man benötigt  $k - j - 1$  Vertauschungen um  $a_k$  and die  $j$ -te Stelle zu bringen.

Bei jeder Vertauschung ändert  $F$  das Vorzeichen, also

$$\begin{aligned} F(\dots, a_j, \dots, a_k, \dots) &= (-1)^{2(k-j)-1} F(\dots, a_k, \dots, a_j, \dots) \\ &= -F(\dots, a_k, \dots, a_j, \dots). \end{aligned}$$

□

**3.4 Satz. Existenz und Eindeutigkeit der Determinantenform auf  $\mathbb{K}^n$**

- (a) Für jedes  $n \geq 2$  gibt es genau eine Determinantenform auf  $\mathbb{K}^n$ . Sie wird mit  $\det_n$  oder einfach nur  $\det$  bezeichnet.
- (b) Hat die  $n \times n$ -Matrix  $A = (a_{jk})$  die Spalten  $a_1, \dots, a_n \in \mathbb{K}^n$  so heißt  $\det(a_1, \dots, a_n)$  die Determinante von  $A$ , bezeichnet mit

$$|A| = \det A = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{vmatrix}.$$

- (c) Die Determinante einer  $n \times n$ -Matrix  $A = (a_{jk})$  läßt sich wie folgt durch  $(n - 1) \times (n - 1)$ -Determinanten ausdrücken:

$$\begin{aligned} |A| &= \sum_{j=1}^n (-1)^{i+j} a_{ij} |A_{ij}| && \text{(Entwicklung nach der } i\text{-ten Zeile)} \\ &= \sum_{j=1}^n (-1)^{j+k} a_{jk} |A_{jk}| && \text{(Entwicklung nach der } k\text{-ten Spalte).} \end{aligned}$$

Dabei sei  $A_{ij}$  diejenige  $(n - 1) \times (n - 1)$ -Matrix, welche aus  $A$  durch Streichen der  $i$ -ten Zeile und der  $j$ -ten Spalte hervorgeht.

**3.5 Beispiele.** (a) Durch Entwicklung nach der letzten Spalte erhalten wir

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} - a_{23} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} + a_{33} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \\ &= a_{13}(a_{21}a_{32} - a_{31}a_{22}) - a_{23}(a_{11}a_{32} - a_{31}a_{12}) + a_{33}(a_{11}a_{22} - a_{21}a_{12}) \\ &= \vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3). \end{aligned}$$

**Merkschema für  $3 \times 3$ -Determinanten:**

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= \underline{a_{11}a_{22}a_{33}} + \underline{a_{12}a_{23}a_{31}} + \underline{a_{13}a_{21}a_{32}} \\ &\quad - \underline{a_{31}a_{22}a_{13}} - \underline{a_{32}a_{23}a_{11}} - \underline{a_{33}a_{21}a_{12}} \end{aligned}$$

(b) In konkreten Beispielen entwickelt man zweckmäßigerweise nach denjenigen Spalten bzw. Zeilen, die möglichst viele Nullen enthalten, z.B. hier nach der 2-ten Spalte

$$\begin{vmatrix} 2 & 1 & 2 & 2 \\ 1 & 0 & 1 & 1 \\ -2 & 0 & 1 & -2 \\ 3 & 4 & 1 & 0 \end{vmatrix} = -1 \begin{vmatrix} 1 & 1 & 1 \\ -2 & 1 & -2 \\ 3 & 1 & 0 \end{vmatrix} + 4 \begin{vmatrix} 2 & 2 & 2 \\ 1 & 1 & 1 \\ -2 & 1 & -2 \end{vmatrix} = -(-6 - 2 - 3 + 2) = 9.$$

Wir beweisen Satz 3.4 in mehreren Schritten.

**3.6 Satz.** Stimmen zwei alternierende  $n$ -Formen  $F$  und  $G$  auf dem  $\mathbb{K}^n$  auf der kanonischen Basis überein, so sind sie gleich. Insbesondere gibt es höchstens eine Determinantenform auf dem  $\mathbb{K}^n$ .

*Beweis.* Wir zeigen, dass für jede alternierende  $n$ -Form  $H$  auf  $\mathbb{K}^n$  gilt:

$$H(e_1, \dots, e_n) = 0 \Rightarrow H = 0.$$

Daraus folgt dann

$$F(e_1, \dots, e_n) = G(e_1, \dots, e_n) \Rightarrow F = G,$$

weil  $H = F - G$  selbst wieder eine alternierende  $n$ -Form ist.

Sei also  $H(e_1, \dots, e_n) = 0$ . Dann ist auch  $H(e_{\pi(1)}, \dots, e_{\pi(n)}) = 0$  für jede Permutation

$$\pi : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\},$$

da jede Permutation Komposition von Vertauschungen ist.

Wegen der Multilinearität ist aber dann für beliebige  $a_1, \dots, a_n \in \mathbb{K}^n$

$$\begin{aligned} H(a_1, \dots, a_n) &= H\left(\sum_{j=1}^n a_{j1} e_j, a_2, \dots, a_n\right) = \sum_{j=1}^n a_{j1} H(e_j, a_2, \dots, a_n) = \dots = \\ &= \sum_{j=1}^n \dots \sum_{k=1}^n a_{j1} \dots a_{kn} H(e_j, \dots, e_k) = 0. \end{aligned}$$

□

Es bleibt also die Existenz einer Determinantenform  $\det_n$  auf  $\mathbb{K}^n$  zu zeigen.

Dazu verwenden wir die Entwicklung nach Zeilen und zeigen per Induktion, dass diese tatsächlich eine Determinantenform liefert.

**Induktionsanfang:** Für  $n = 2$  liefert  $\det_2(A) = a_{11}a_{22} - a_{12}a_{21}$  eine Determinantenform in den Spalten von  $A$ .

**Induktionsschritt:** Sei  $\det_{n-1}$  eine Determinantenform auf  $\mathbb{R}^{n-1}$  und sei  $i \in \{1, \dots, n\}$  beliebig aber fest. Für  $A \in M(n \times n, \mathbb{K})$  sei nun

$$\det_n(A) := \sum_{j=1}^n (-1)^{i+j} a_{ij} \det_{n-1}(A_{ij}). \tag{3.1}$$

Es ist zu zeigen, dass  $\det_n$  eine Determinantenform in den Spalten von  $A$  ist:

(a) **det<sub>n</sub> ist multilinear**, denn  $a_{ij} \det_{n-1}(A_{ij})$  ist linear in jeder Spalte von  $A$ :

in der  $j$ -ten Spalte, da  $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} \mapsto a_{ij}$  linear ist,

in den anderen Spalten, da  $\det_{n-1} A_{ij}$  nach Induktionsannahme eine Multilinearform ist.

- (b) **det<sub>n</sub> ist alternierend:** nach Satz 3.3 müssen wir nur zeigen, dass  $\det_n(A) = 0$  wenn zwei benachbarte Spalten gleich sind.

Sei etwa  $\vec{a}_k = \vec{a}_{k+1}$ . Für  $j \neq k$  und  $j \neq k+1$  hat  $A_{ij}$  dann zwei gleiche Spalten, weswegen  $\det_{n-1}A_{ij} = 0$  gilt. Damit bleibt in (3.1) nur noch

$$\det_n(A) = (-1)^{i+k} a_{ik} \det_{n-1}(A_{ik}) + (-1)^{i+k+1} a_{i,k+1} \det_{n-1}(A_{i,k+1}).$$

Wegen  $\vec{a}_k = \vec{a}_{k+1}$  ist aber  $a_{ik} = a_{i,k+1}$  und  $A_{ik} = A_{i,k+1}$  und somit  $\det_n|A| = 0$ .

- (c) **det<sub>n</sub> ist normiert,** da

$$\begin{aligned} \det_n(E_n) &= \sum_{j=1}^n (-1)^{i+j} \delta_{ij} \det_{n-1}(E_{ij}) \\ &= (-1)^{2i} \det_{n-1}(E_{ii}) = \det_{n-1}(E_{n-1}) = 1. \end{aligned}$$

Also gibt es für jedes  $n \geq 2$  eine Determinantenform auf  $\mathbb{K}^n$ , welche nach Satz 3.6 eindeutig ist.

Insbesondere liefert die Entwicklung nach der  $i$ -ten Zeile für jedes  $i \in \{1, \dots, n\}$  das gleiche Ergebnis.

Die Behauptung zur Entwicklung nach Spalten folgt aus  $\det A = \det A^T$ , der Aussage des nächsten Satzes.

### 3.7 Satz. Determinante der transponierten Matrix

Sei  $A^T = (a_{ki})$  die zu  $A = (a_{ik}) \in M(n, \mathbb{K})$  transponierte Matrix, also

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{1n} & \cdots & \cdots & a_{nn} \end{pmatrix} \quad \text{für} \quad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{pmatrix}.$$

Dann gilt  $|A^T| = |A|$ .

*Beweis.* Für  $n = 2$  gilt  $\begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ .

Sei also  $n \geq 3$ . Es gilt  $|A^T| = \det(z_1, \dots, z_n)$ , wobei  $z_1, \dots, z_n$  die Spalten von  $A^T$  bzw. die Zeilenvektoren von  $A$  sind.

Sei andererseits  $D(z_1, \dots, z_n) := |A|$  aufgefasst als Funktion der Zeilen von  $A$ .

Wir werden zeigen, dass  $D$  eine Determinantenform ist. Wegen der Eindeutigkeit gilt dann aber  $D = \det$  und somit  $|A| = |A^T|$ .

- (a)  **$D$  ist linear  $i$ -ten Argument:** Entwicklung nach  $i$ -ter Zeile

$$D(z_1, \dots, z_n) = |A| = \sum_{j=1}^n (-1)^{i+j} a_{ij} |A_{ij}|$$

$z_i$  kommt in  $|A_{ij}|$  nicht vor und  $z_i = \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix} \mapsto a_{ij}$  ist linear.

- (b)  **$D$  ist alternierend:** Sei  $z_k = z_{k+1}$ . In der Entwicklung nach der  $i$ -ten Zeile für  $i \neq k$  und  $i \neq k+1$  ( $n \geq 3$ ) ist jeder Term gleich Null, da  $A_{ij}$  für jedes  $j$  zwei gleiche Zeilen hat.

Damit ist  $|A_{ij}| = 0$ , da  $A_{ij}$  nicht vollen Rang hat.

- (c)  $D(e_1, \dots, e_n) = |E_n| = 1$ .

□

Nun können wir noch den Teil von Satz 3.4 über die Entwicklung nach Spalten beweisen:

$$|A| = |A^T| = \sum_{j=1}^n (-1)^{k+j} a_{kj}^T |(A^T)_{kj}| = \sum_{j=1}^n (-1)^{k+j} a_{jk} |A_{jk}|.$$

↑  
Entwicklung nach  $k$ -ter Zeile

**3.8 Satz. Multiplikationssatz für Determinanten**

Für  $A, B \in M(n, \mathbb{K})$  gilt

$$|AB| = |A| \cdot |B|.$$

*Beweis.* Sind  $b_1, \dots, b_n$  die Spalten von  $B$ , so hat  $AB$  die Spalten  $Ab_1, \dots, Ab_n$ . Für festes  $A$  und beliebiges  $B$  betrachten wir

$$F(b_1, \dots, b_n) = \det(Ab_1, \dots, Ab_n) = |AB|.$$

Als Komposition der linearen Abbildungen  $b_i \mapsto Ab_i$  und  $\det$  ist  $F$  multilinear.  $F$  ist alternierend, da  $\det$  alternierend ist und es gilt

$$F(e_1, \dots, e_n) = \det(a_1, \dots, a_n) = |A|$$

Auch durch  $G(b_1, \dots, b_n) = |A| \det(b_1, \dots, b_n)$  ist eine alternierende  $n$ -Form mit  $G(e_1, \dots, e_n) = |A|$  gegeben. Falls  $|A| = 0$  ist, so gilt  $G = F$ . Falls  $|A| \neq 0$  ist, so sind  $\frac{G}{|A|}$  und  $\frac{F}{|A|}$  Determinantenformen und wegen der Eindeutigkeit gleich. Also  $G = F$ , d.h.  $|A| \cdot |B| = |AB|$ . □

**3.9 Korollar.** Eine Matrix  $A \in M(n, \mathbb{K})$  ist genau dann invertierbar, wenn  $|A| \neq 0$  ist. Dann gilt

$$|A^{-1}| = \frac{1}{|A|}.$$

*Beweis.* Falls  $A$  invertierbar ist, so folgt aus dem Multiplikationssatz

$$1 = |E| = |A \cdot A^{-1}| = |A| |A^{-1}|,$$

also

$$|A| \neq 0 \quad \text{und} \quad |A^{-1}| = \frac{1}{|A|}.$$

Ist  $A$  nicht invertierbar, so ist  $A$  nicht surjektiv. Also sind die Spalten von  $A$  linear abhängig und es gilt nach Satz 3.2, dass  $|A| = 0$  ist. □

**3.10 Zusammenfassung.** Für eine  $n \times n$  Matrix  $A$  sind also äquivalent

- (a)  $|A| \neq 0$
- (b)  $A$  ist invertierbar.
- (c) Das Gleichungssystem  $Ax = b$  ist universell und eindeutig lösbar.
- (d) Die Spalten von  $A$  sind linear unabhängig.
- (e) Die Zeilen von  $A$  sind linear unabhängig.
- (f)  $\text{Rang} A = n$
- (g)  $A : \mathbb{K}^n \rightarrow \mathbb{K}^n$  ist injektiv.
- (h)  $A : \mathbb{K}^n \rightarrow \mathbb{K}^n$  ist surjektiv.

**3.11 Satz. und Definition. Die Determinante eines Endomorphismus**

Sei  $V$  ein Vektorraum der Dimension  $n < \infty$ ,  $L \in \mathcal{L}(V)$  und  $\mathcal{B}$  eine Basis von  $V$ . Die Zahl  $|M_{\mathcal{B}}(L)|$  hängt nicht von der Basis  $\mathcal{B}$  ab und wird mit  $\det L$  bezeichnet.

*Beweis.* Ähnliche Matrizen haben gemäß des Multiplikationssatzes gleiche Determinanten,

$$|S^{-1}AS| = |S^{-1}| |A| |S| = |A|.$$

Da für zwei Basen  $\mathcal{A}$  und  $\mathcal{B}$  die Matrizen  $M_{\mathcal{A}}(L)$  und  $M_{\mathcal{B}}(L)$  ähnlich sind, folgt die Aussage. □

**3.12 Satz. Die Cramersche Regel**

Für eine invertierbare  $n \times n$ -Matrix  $A = (a_1, \dots, a_n)$  über  $\mathbb{K}$  und  $b \in \mathbb{K}^n$  ist die eindeutige Lösung der Gleichung  $Ax = b$  gegeben durch

$$x_i = \frac{1}{|A|} \det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n) \quad \text{für } i = 1, \dots, n.$$

*Beweis.* Wir schreiben die Gleichung  $Ax = b$  als

$$x_1 a_1 + \dots + x_n a_n = b$$

bzw.

$$x_1 a_1 + \dots + (x_i a_i - b) + \dots + x_n a_n = 0.$$

Also sind die Spalten der Matrix

$$\begin{pmatrix} a_{11} & \dots & (x_i a_{1i} - b_1) & \dots & a_{1n} \\ \vdots & & & & \vdots \\ a_{n1} & \dots & (x_i a_{ni} - b_n) & \dots & a_{nn} \end{pmatrix}$$

linear abhängig und ihre Determinante somit Null. Linearität in der  $i$ -ten Spalte liefert schließlich

$$x_i \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} - \begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ \vdots & & & & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{vmatrix} = 0.$$

□

**3.13 Bemerkung.** Die Cramersche Regel ist nicht zur expliziten Berechnung von Lösungen geeignet, zeigt aber z.B. die stetige Abhängigkeit der Lösungen von  $A$  und  $b$ .

**Determinante und Volumen im  $\mathbb{R}^n$** 

Seien  $\vec{a}_1, \dots, \vec{a}_n$  Vektoren im  $\mathbb{R}^n$ , dann ist das von ihnen aufgespannte **Parallelotop** die Menge

$$P(\vec{a}_1, \dots, \vec{a}_n) := \left\{ \sum_{j=1}^n \alpha_j \vec{a}_j \mid 0 \leq \alpha_j \leq 1 \text{ für } j = 1, \dots, n \right\}.$$

Für  $n = 1$  ist  $P(\vec{a}_1)$  eine Strecke, für  $n = 2$  ist  $P(\vec{a}_1, \vec{a}_2)$  ein Parallelogramm und für  $n = 3$  ist  $P(\vec{a}_1, \vec{a}_2, \vec{a}_3)$  ein Spat. Der  $n$ -dimensionale Einheitswürfel ist  $P(\vec{e}_1, \dots, \vec{e}_n)$ .

Wir haben uns bereits überlegt, dass

$$\text{Fläche}(P(\vec{a}_1, \vec{a}_2)) = |\det(\vec{a}_1, \vec{a}_2)|$$

und

$$\text{Volumen}(P(\vec{a}_1, \vec{a}_2, \vec{a}_3)) = |\det(\vec{a}_1, \vec{a}_2, \vec{a}_3)|$$

gilt. Dieser Zusammenhang besteht auch für  $n \geq 4$ . Da wir das "Volumen" von Teilmengen des  $\mathbb{R}^n$  noch nicht definiert haben, erheben wir diese Setzung zur Definition. Das müssen wir natürlich begründen. Dazu stellen wir fest, dass wir folgende Eigenschaften a priori von jeder vernünftigen Definition des  $n$ -dimensionalen Volumens

$$\text{Vol}(P(a_1, \dots, a_n)) =: V(a_1, \dots, a_n)$$

fordern müssen:

(a) **Homogenität in jeder Richtung**

$$V(\dots, a_{i-1}, \alpha a_i, a_{i+1}, \dots) = |\alpha| V(a_1, \dots, a_n)$$

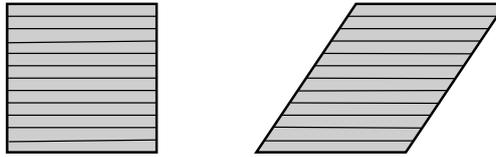
für jedes  $i = 1, \dots, n$  und  $\alpha \in \mathbb{R}$ .

(b) **Cavalierisches Prinzip**

$$V(\dots, a_{i-1}, a_i + \alpha a_k, a_{i+1}, \dots) = V(a_1, \dots, a_n)$$

für  $i = 1, \dots, n$  und  $k \neq i$ .

“Ein Papierstapel ändert sein Volumen nicht, wenn man die Blätter gegeneinander verschiebt”:



(c) **Der Einheitswürfel hat Volumen 1**

$$V(e_1, \dots, e_n) = 1.$$

**3.14 Satz. Eindeutigkeit des “Volumens”**

Es gibt genau eine Funktion  $V : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$  welche die drei Eigenschaften (a), (b) und (c) hat. Sie ist gegeben durch

$$V(a_1, \dots, a_n) = |\det(a_1, \dots, a_n)|$$

und heißt Volumen des Parallelotops  $P(a_1, \dots, a_n)$ .

*Beweis.*  $|\det(a_1, \dots, a_n)|$  hat offensichtlich die Eigenschaften (a), (b) und (c). Erfülle nun auch  $V$  (a), (b) und (c) und sei

$$F(a_1, \dots, a_n) := \frac{V(a_1, \dots, a_n)}{|\det(a_1, \dots, a_n)|} \det(a_1, \dots, a_n)$$

falls  $a_1, \dots, a_n$  linear unabhängig und  $F(a_1, \dots, a_n) = 0$  sonst. Wir zeigen, dass  $F$  eine Determinantenform ist. Dann folgt aus der Eindeutigkeit  $F = \det$  und somit die Aussage des Satzes.

Wir zeigen zunächst, dass  $F$  multilinear ist. Dazu setzen wir  $a_1 = \alpha x + \beta y$  und weisen o.B.d.A. die Linearität im ersten Argument nach.

(i) Falls  $a_2, \dots, a_n$  linear abhängig sind, gilt

$$F(\alpha x + \beta y, a_2, \dots, a_n) = 0 = 0 + 0 = \alpha F(x, a_2, \dots, a_n) + \beta F(y, a_2, \dots, a_n).$$

(ii) Seien also  $a_2, \dots, a_n$  linear unabhängig. Falls  $\beta y = \sum_{j=2}^n t_j a_j$ , dann ist

$$\begin{aligned} F(\alpha x + \beta y, \dots) &= \frac{V(\alpha x + \beta y, \dots)}{|\det(\alpha x + \beta y, \dots)|} \det(\alpha x + \beta y, \dots) \\ &= \frac{V(\alpha x, \dots)}{|\det(\alpha x, \dots)|} \det(\alpha x, \dots) \\ &= \frac{|\alpha| V(x, \dots)}{|\alpha| |\det(x, \dots)|} \alpha \det(x, \dots) \\ &= \alpha F(x, \dots) = \alpha F(x, \dots) + \beta \underbrace{F(y, \dots)} = 0. \end{aligned}$$

Dabei folgt  $V(\alpha x + \beta y, \dots) = V(\alpha x, \dots)$  aus dem Cavalierischen Prinzip.

(iii) Seien schließlich  $\beta y, a_2, \dots, a_n$  linear unabhängig und  $\alpha x = t_1 \beta y + \sum_{j=2}^n t_j a_j$ , dann ist

$$\begin{aligned} F(\alpha x + \beta y, \dots) &= F((1 + t_1)\beta y, \dots) = (1 + t_1)\beta F(y, \dots) = \beta F(y, \dots) + F(t_1 \beta y, \dots) \\ &= \beta F(y, \dots) + F(\alpha x, \dots) = \beta F(y, \dots) + \alpha F(x, \dots), \end{aligned}$$

wobei wir im ersten und im vorletzten Schritt wieder das Cavalierische Prinzip verwendet haben.

Also ist  $F$  multilinear. Da  $F(\dots, a, a, \dots) = 0$  per Definition gilt, ist  $F$  alternierend und wegen  $V(e_1, \dots, e_n) = 1$  ist  $F$  eine Determinantenform. □

## 3.2 Eigenwerte und Eigenvektoren

### 3.15 Definition. Eigenwerte und Eigenvektoren

Sei  $V$  ein Vektorraum über  $\mathbb{K}$  und  $L : V \rightarrow V$  ein Endomorphismus. Eine Zahl  $\lambda \in \mathbb{K}$  heißt **Eigenwert** von  $L$ , falls es einen Vektor  $v \neq 0$  aus  $V$  mit

$$Lv = \lambda v$$

gibt. Ein solcher Vektor heißt dann **Eigenvektor** von  $L$  zum Eigenwert  $\lambda$ .

**3.16 Lemma.** Ist  $A = M_{\mathcal{B}}(L)$  die Matrix von  $L \in \mathcal{L}(V)$  bezüglich einer Basis  $\mathcal{B} = (v_1, \dots, v_n)$  von  $V$ , so hat  $A$  genau dann **Diagonalgestalt**,

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wenn für alle  $j = 1, \dots, n$  der Basisvektor  $v_j$  Eigenvektor von  $L$  zum Eigenwert  $\lambda_j$  ist.

*Beweis.* Es gilt  $Lv_j = \lambda_j v_j \Leftrightarrow M_{\mathcal{B}}(L)e_j = \lambda_j e_j \Leftrightarrow$  die  $j$ -te Spalte von  $M_{\mathcal{B}}(L)$  ist gleich  $\lambda_j e_j$ . □

### 3.17 Definition. Diagonalisierbarkeit

Ein Endomorphismus für den eine Basis aus Eigenvektoren existiert heißt diagonalisierbar.

**3.18 Beispiel.** Sei  $V = \mathbb{R}^2$ , also  $\mathbb{K} = \mathbb{R}$ .

- (a) **Spiegelung  $L$  an der “ $y$ -Achse”:** Es ist  $Le_1 = -e_1$  und  $Le_2 = e_2$ . Somit hat  $L$  die Eigenwerte  $\lambda_1 = -1$  und  $\lambda_2 = 1$  zu den Eigenvektoren  $e_1$  und  $e_2$ . In der Basis aus Eigenvektoren, also der kanonischen Basis, hat  $L$  die Diagonalgestalt

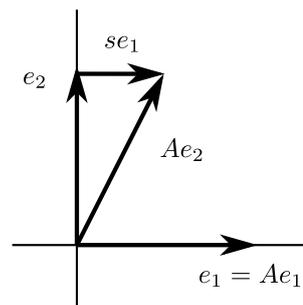
$$M_{\mathcal{K}}(L) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

- (b) **Drehung  $D_\varphi$  um den Winkel  $\varphi$  um den Ursprung:** Für  $\varphi \notin \pi\mathbb{Z}$  hat  $D_\varphi$  keinen Eigenvektor, da jeder Vektor unter  $D_\varphi$  seine Richtung ändert. Also hat  $D_\varphi$  dann auch keinen (reellen) Eigenwert und ist nicht diagonalisierbar.
- (c) **Scherung:**

Eine Scherung des  $\mathbb{R}^2$  ist gegeben durch

$$A = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \quad \text{mit } s \neq 0.$$

$A$  hat nur den Eigenvektor  $e_1$  zum Eigenwert  $\lambda = 1$ , also  $Ae_1 = e_1$ . Auch  $A$  ist somit nicht diagonalisierbar.



**3.19 Bemerkung.** Ein Vektor  $v \in V \setminus \{0\}$  ist genau dann Eigenvektor von  $L \in \mathcal{L}(V)$  zum Eigenwert  $\lambda \in \mathbb{K}$ , wenn

$$v \in \text{Kern}(L - \lambda \mathbb{1}_V)$$

gilt. Denn  $v \in \text{Kern}(L - \lambda \mathbb{1}_V) \Leftrightarrow (L - \lambda \mathbb{1}_V)v = 0 \Leftrightarrow Lv = \lambda v$ .

Also ist  $\lambda \in \mathbb{K}$  genau dann Eigenwert von  $L \in \mathcal{L}(V)$ , wenn  $\text{Kern}(L - \lambda \mathbb{1}_V) \neq \{0\}$ .

### 3.20 Definition. Eigenraum und geometrische Vielfachheit

Sei  $\lambda$  ein Eigenwert von  $L \in \mathcal{L}(V)$ , dann heißt der Unterraum

$$E_\lambda := \text{Kern}(L - \lambda \mathbb{1}_V)$$

von  $V$  der **Eigenraum** von  $L$  zum Eigenwert  $\lambda$ . Die Dimension  $\dim E_\lambda$  heißt **geometrische Vielfachheit** des Eigenwertes  $\lambda$ .

**3.21 Bemerkung.** Jeder Vektor  $v \in E_\lambda$  mit  $v \neq 0$  ist Eigenvektor zum Eigenwert  $\lambda$ . Bei der Wahl von Eigenvektoren besteht also immer eine Freiheit. Das geometrisch relevante Objekt ist der Eigenraum. Eigenräume zu verschiedenen Eigenwerten haben nur die Null gemeinsam, denn es kann ja nicht  $Lv = \lambda v = \mu v$  für  $\lambda \neq \mu$  gelten. Es gilt aber mehr:

**3.22 Lemma. Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig**

Sei  $L \in \mathcal{L}(V)$  und seien  $v_1, \dots, v_\ell$  Eigenvektoren zu verschiedenen Eigenwerten  $\lambda_1, \dots, \lambda_\ell$ , also  $\lambda_i \neq \lambda_j$  für  $i \neq j$ , dann sind  $v_1, \dots, v_\ell$  linear unabhängig.

*Beweis. Induktionsanfang:* Für  $\ell = 1$  ist  $v_1$  linear unabhängig, da nach Definition  $v_1 \neq 0$  gilt.

*Induktionsschritt:* Die Aussage sei richtig für  $\ell$ . Seien nun  $v_1, \dots, v_{\ell+1}$  Eigenvektoren zu verschiedenen Eigenwerten  $\lambda_1, \dots, \lambda_{\ell+1}$  und sei

$$\alpha_1 v_1 + \dots + \alpha_{\ell+1} v_{\ell+1} = 0.$$

Daraus erhält man durch Anwendung von  $L$  bzw.  $\lambda_{\ell+1} \mathbb{1}_V$  die Gleichungen

$$\begin{aligned} \alpha_1 \lambda_1 v_1 + \dots + \alpha_{\ell+1} \lambda_{\ell+1} v_{\ell+1} &= 0 \\ \alpha_1 \lambda_{\ell+1} v_1 + \dots + \alpha_{\ell+1} \lambda_{\ell+1} v_{\ell+1} &= 0 \end{aligned}$$

und durch Subtraktion

$$\alpha_1 (\lambda_1 - \lambda_{\ell+1}) v_1 + \dots + \alpha_\ell (\lambda_\ell - \lambda_{\ell+1}) v_\ell = 0.$$

Nach Induktionsannahme sind  $v_1, \dots, v_\ell$  aber linear unabhängig, somit ist

$$\alpha_1 (\lambda_1 - \lambda_{\ell+1}) = \dots = \alpha_\ell (\lambda_\ell - \lambda_{\ell+1}) = 0.$$

Wegen  $\lambda_i \neq \lambda_j$  für  $i \neq j$  folgt  $\alpha_1 = \dots = \alpha_\ell = 0$ , damit  $\alpha_{\ell+1} v_{\ell+1} = 0$ , also  $\alpha_{\ell+1} = 0$  und schließlich die lineare Unabhängigkeit von  $v_1, \dots, v_{\ell+1}$ . □

**3.23 Bemerkung.** Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig. Damit kann ein Endomorphismus  $L$  auf einem  $n$ -dimensionalen Raum höchstens  $n$  verschiedene Eigenwerte haben. In diesem Fall bilden die Eigenvektoren dann eine Basis und  $L$  ist diagonalisierbar. Das Vorliegen von  $n$  verschiedenen Eigenwerten ist allerdings keine notwendige Bedingung für Diagonalisierbarkeit, beispielsweise hat die Einheitsmatrix nur den einen Eigenwert  $\lambda = 1$  und die Nullmatrix nur den Eigenwert  $\lambda = 0$ . Beide sind aber diagonal.

Wie bereits bemerkt, ist  $\lambda \in \mathbb{K}$  genau dann Eigenwert von  $L \in \mathcal{L}(V)$ , wenn  $\text{Kern}(L - \lambda \mathbb{1}_V) \neq \{0\}$ , also wenn  $L - \lambda \mathbb{1}_V$  nicht injektiv ist. Daraus ergibt sich das

**3.24 Lemma.** Sei  $\dim V = n < \infty$  und  $L \in \mathcal{L}(V)$ . Dann ist  $\lambda \in \mathbb{K}$  genau dann ein Eigenwert von  $L$ , wenn  $\det(L - \lambda \mathbb{1}_V) = 0$  gilt.

Wird  $L$  bezüglich einer Basis von  $V$  durch die Matrix  $A = (a_{ij})$  beschrieben, so ist

$$\det(L - \lambda \mathbb{1}_V) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix}.$$

Diese von  $\lambda$  abhängige Determinante ist ein Polynom vom Grade  $n$  in  $\lambda$  und heißt das charakteristische Polynom von  $L$ .

**3.25 Lemma.** Sei  $\dim V = n < \infty$  und  $L \in \mathcal{L}(V)$ . Dann gibt es Zahlen  $a_0, \dots, a_n \in \mathbb{K}$  mit

$$P_L(\lambda) := \det(L - \lambda \mathbb{1}_V) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + a_0$$

für alle  $\lambda \in \mathbb{K}$ . Es gilt

$$a_n = (-1)^n, \quad a_{n-1} = (-1)^{n-1} \text{Spur} L := (-1)^{n-1} \sum_{j=1}^n a_{jj} \quad \text{und} \quad a_0 = \det L.$$

Man nennt  $P_L$  das **charakteristische Polynom** von  $L$ .

*Beweis.* Sei  $A = M_{\mathcal{B}}(L)$  bezüglich einer Basis  $\mathcal{B}$  von  $V$ . Wir bestimmen

$$\det(L - \lambda \mathbf{1}_V) = \det(A - \lambda E_n).$$

Aus

$$\det(A - \lambda B) = \sum_{j=1}^n (-1)^{i+j} (a_{ij} - \lambda b_{ij}) \det((A - \lambda B)_{ij})$$

sieht man sofort durch Induktion, dass für beliebige  $n \times n$ -Matrizen  $A$  und  $B$   $\det(A - \lambda B) = c_n \lambda^n + \dots + c_1 \lambda + c_0$  für geeignete  $c_0, \dots, c_n \in \mathbb{K}$  ist.

Für  $B = E_n$  ergibt Induktion und Entwicklung nach der ersten Zeile

$$\det(A - \lambda E_n) = (a_{11} - \lambda) \det((A - \lambda E_n)_{11}) + R(\lambda).$$

Es ist  $R(\lambda)$  ein Polynom vom Grad  $n - 2$ , da  $\det(A - \lambda E_n)_{1j}$  ein Polynom vom Grad  $n - 2$  ist.

Nun ist nach Induktionsannahme

$$\det((A - \lambda E_n)_{11}) = (-1)^{n-1} \lambda^{n-1} + (-1)^{n-2} \sum_{j=2}^n a_{jj} \lambda^{n-2} + \dots$$

und somit

$$\det(A - \lambda E_n) = (-1)^n \lambda^n + (-1)^{n-1} \sum_{j=1}^n a_{jj} \lambda^{n-1} + \dots$$

Der konstante Term ergibt sich aus

$$a_0 = P_L(0) = \det(A - 0 \cdot E_n) = \det A.$$

□

**3.26 Korollar.** Die Eigenwerte sind die Nullstellen des charakteristischen Polynoms.

**3.27 Korollar.** Sei  $\dim V = n$  und  $L \in \mathcal{L}(V)$ . Seien  $\lambda_1, \dots, \lambda_\ell$  verschiedene Eigenwerte von  $L$  mit geometrischen Vielfachheiten  $n_1, \dots, n_\ell$  und  $(v_1^{(i)}, \dots, v_{n_i}^{(i)})$  jeweils eine Basis des Eigenraums  $E_{\lambda_i}$  zum Eigenwert  $\lambda_i$ .

Dann sind auch

$$v_1^{(1)}, \dots, v_{n_1}^{(1)}, v_1^{(2)}, \dots, v_{n_2}^{(2)}, \dots, v_1^{(\ell)}, \dots, v_{n_\ell}^{(\ell)}$$

linear unabhängig.

Insbesondere ist die Summe  $n_1 + \dots + n_\ell$  der geometrischen Vielfachheiten höchstens  $n$  und  $L$  ist genau dann diagonalisierbar, wenn  $n_1 + \dots + n_\ell = n$  ist.

*Beweis.* Ist  $\sum_{i=1}^{\ell} \sum_{k=1}^{n_i} \alpha_k^{(i)} v_k^{(i)} = 0$ , so sind nach Lemma 3.22 die Vektoren  $\sum_{k=1}^{n_i} \alpha_k^{(i)} v_k^{(i)} \in E_{\lambda_i}$  gleich Null und wegen der linearen Unabhängigkeit der  $v_1^{(i)}, \dots, v_{n_i}^{(i)}$  verschwinden alle Koeffizienten  $\alpha_k^{(i)}$ .

Also ist  $(v_1^{(1)}, \dots, v_{n_1}^{(1)}, \dots, v_1^{(\ell)}, \dots, v_{n_\ell}^{(\ell)})$  ein linear unabhängiges  $n_1 + \dots + n_\ell$ -Tupel von Eigenvektoren.

Im Fall  $n_1 + \dots + n_\ell = n$  ist es eine Basis aus Eigenvektoren und  $L$  somit diagonalisierbar.

Ist nun andererseits  $L$  als diagonalisierbar vorausgesetzt und sei  $m_i$  die Anzahl der Eigenvektoren zum Eigenwert  $\lambda_i$  in einer Basis aus Eigenvektoren, so ist  $m_i \leq \dim E_{\lambda_i} = n_i$ , also

$$n = m_1 + \dots + m_\ell \leq n_1 + \dots + n_\ell \leq n,$$

woraus  $n_1 + \dots + n_\ell = n$  und  $m_j = n_j$  für alle  $j = 1, \dots, \ell$  folgt.

□

### Der Fundamentalsatz der Algebra

Jedes komplexe Polynom vom Grade  $n \geq 1$ , d.h. jede Abbildung  $P : \mathbb{C} \rightarrow \mathbb{C}$  von der Form

$$P(z) = c_n z^n + \dots + c_1 z + c_0,$$

wobei  $c_0, \dots, c_n \in \mathbb{C}$  und  $c_n \neq 0$ , hat mindestens eine Nullstelle. Der Satz wird im 3. Semester bewiesen mittels komplexer Analysis.

**3.28 Korollar.** Für  $n \geq 1$  hat jeder Endomorphismus eines  $n$ -dimensionalen komplexen Vektorraumes mindestens einen Eigenwert.

Nun kann man einfach aus dem Fundamentalsatz folgern:

**3.29 Lemma.** Jedes komplexe Polynom  $P$  zerfällt in Linearfaktoren: ist  $P(z) = c_n z^n + \dots + c_1 z + c_0$  mit  $c_n \neq 0$ , und sind  $z_1, \dots, z_r \in \mathbb{C}$  die paarweise verschiedenen Nullstellen von  $P$ , so gilt

$$P(z) = c_n \prod_{i=1}^r (z - z_i)^{m_i}$$

mit wohlbestimmten Exponenten  $m_i \geq 1$ , welche man die Vielfachheit der Nullstelle nennt.

*Beweis.* Wir wissen es gibt eine Nullstelle  $z_1$ . Mittels binomischen Lehrsatz kann man

$$P(z) = P(z - z_1 + z_1) = a_n (z - z_1)^n + \dots + a_1 (z - z_1) + b(z_1)$$

schreiben, mit geeigneten Koeffizienten  $a_i$ . Da  $P(z_1) = 0$ , folgt  $b(z_1) = 0$ , und damit

$$P(z) = (z - z_1)Q(z),$$

wobei  $Q(z)$  ein Polynom vom Grad  $n - 1$  ist. Dieses hat nun wieder mindestens eine Nullstelle  $z_2$ , usw..  $\square$

**3.30 Bemerkung.** Im Allgemeinen sind zwei Matrizen  $A, B \in M(n, \mathbb{C})$  nicht vertauschbar, also  $[A, B] = AB - BA \neq 0$ . Sind aber  $A, B$  diagonalisierbar, dann gilt  $[A, B] = 0$  genau dann wenn sich eine gemeinsame Basis aus Eigenvektoren für  $A, B$  finden läßt.

*Beweis.* Sei  $v$  Eigenvektor von  $A$  zum Eigenwert  $\lambda$ , dann gilt

$$ABv = BAv = \lambda Bv,$$

also  $Bv$  liegt im Eigenraum  $E_\lambda^A$ , also läßt den Eigenraum invariant. Damit wird in den Übungen die Aussage des Satzes bewiesen.  $\square$

### 3.31 Definition. Algebraische Vielfachheit

Sei  $\dim V = n < \infty$ ,  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$  und  $L \in \mathcal{L}(V)$ . Falls  $\lambda_0$  eine  $k$ -fache Nullstelle des charakteristischen Polynoms ist, d.h.

$$P_L(\lambda) = (\lambda - \lambda_0)^k Q(\lambda) \quad \text{mit} \quad Q(\lambda_0) \neq 0,$$

so heißt  $k$  die **algebraische Vielfachheit** des Eigenwertes  $\lambda_0$ .

### 3.32 Beispiel. (a) Spiegelung an der $y$ -Achse:

$$L = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow P_L(\lambda) = \det(L - \lambda E_2) = \det \begin{pmatrix} -1 - \lambda & 0 \\ 0 & 1 - \lambda \end{pmatrix} = (\lambda + 1)(\lambda - 1).$$

Die Nullstellen und somit die Eigenwerte sind  $\lambda_1 = 1$ ,  $\lambda_2 = -1$ .

(b) **Drehung um Winkel  $\varphi$  in der Ebene:**

$$D_\varphi = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

Also ist

$$P_{D_\varphi}(\lambda) = \det \begin{pmatrix} \cos \varphi - \lambda & -\sin \varphi \\ \sin \varphi & \cos \varphi - \lambda \end{pmatrix} = (\cos \varphi - \lambda)^2 + \sin^2 \varphi > 0 \quad \text{falls } \varphi \notin \pi\mathbb{Z}.$$

Für  $\varphi = 0$  ist  $D_\varphi = E_2$  und  $P_{E_2}(\lambda) = (1 - \lambda)^2$  hat die doppelte Nullstelle  $\lambda_1 = 1$ . Also hat  $D_0$  den Eigenwert  $\lambda_1 = 1$ . Seine geometrische Vielfachheit ist 2, da der Eigenraum  $E_{\lambda_1} = \mathbb{R}^2$  ist. Für  $\varphi = \pi$  ist  $D_\varphi = -E_2$ , der einzige Eigenwert also  $\lambda = -1$ . Für alle anderen  $\varphi \in (0, 2\pi)$  hat  $P_{D_\varphi}(\lambda)$  keine Nullstellen, und somit hat  $D_\varphi$  dann auch keine Eigenwerte.

(c) **Scherung:**

$$A = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \Rightarrow P_A(\lambda) = \det \begin{pmatrix} 1 - \lambda & s \\ 0 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2.$$

$P_A(\lambda)$  hat also die doppelte Nullstelle  $\lambda_1 = 1$ , der Eigenraum zum Eigenwert  $\lambda_1$  ist aber nur eindimensional,  $E_{\lambda_1} = \text{Span}(e_1)$ . Die geometrische Vielfachheit ist hier also echt kleiner als die algebraische Vielfachheit.

**3.33 Lemma.** Die geometrische Vielfachheit eines Eigenwertes  $\lambda$  ist immer kleiner als die algebraische Vielfachheit.

*Beweis.* Ergänzt man eine Basis des  $n_i$ -dimensionalen Eigenraumes  $E_{\lambda_i}$  zu einer Basis von  $V$ , so hat die Matrix von  $L$  bezüglich dieser Basis die Form

$$\left( \begin{array}{cc|c} \overbrace{\lambda_i & 0}^{n_i} & & \\ & \ddots & \\ 0 & & \lambda_i & * \\ \hline & & 0 & * \end{array} \right)$$

Daraus liest man ab, dass in  $P_L(\lambda) = \det(M(L) - \lambda E_n)$  der Faktor  $(\lambda_i - \lambda)$  mindestens  $n_i$ -mal vorkommt.  $\square$

**3.34 Korollar.** Sei  $V$  ein komplexer  $n$ -dimensionaler Vektorraum und  $L \in \mathcal{L}(V)$ . Sind  $\lambda_1, \dots, \lambda_\ell$  die Eigenwerte von  $L$  mit algebraischen Vielfachheiten  $m_1, \dots, m_\ell$  so gilt  $m_1 + \dots + m_\ell = n$ . Insbesondere ist  $L$  genau dann diagonalisierbar, wenn für jeden Eigenwert gilt  $n_i = \dim E_{\lambda_i} = m_i$ , also geometrische Vielfachheit = algebraische Vielfachheit.

**3.35 Beispiele.** (a) Sei  $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ , dann ist

$$\begin{aligned} P_A(\lambda) &= \det(A - \lambda E_3) = \begin{vmatrix} 1 - \lambda & 2 & 3 \\ 4 & 5 - \lambda & 6 \\ 7 & 8 & 9 - \lambda \end{vmatrix} \\ &= (1 - \lambda)(5 - \lambda)(9 - \lambda) + 2 \cdot 6 \cdot 7 + 3 \cdot 4 \cdot 8 - 7 \cdot (5 - \lambda) \cdot 3 - 8 \cdot 6 \cdot (1 - \lambda) - (9 - \lambda) \cdot 4 \cdot 2 \\ &= -\lambda^3 + 15\lambda^2 - 18\lambda + 0 = -\lambda(\lambda^2 - 15\lambda + 18) \end{aligned}$$

$$\Rightarrow \lambda_1 = 0, \lambda_2 = \frac{15 + 3\sqrt{17}}{2}, \lambda_3 = \frac{15 - 3\sqrt{17}}{2}$$

Um die Eigenräume zu bestimmen, müsste man nun die Gleichungssysteme

$$(A - \lambda_1 E_3)v_1 = 0, (A - \lambda_2 E_3)v_2 = 0, (A - \lambda_3 E_3)v_3 = 0$$

lösen. Da alle Eigenwerte einfach sind (d.h. einfache Nullstellen des charakteristischen Polynoms), bilden die zugehörigen Eigenvektoren eine Basis von  $\mathbb{R}^3$  und  $A$  ist diagonalisierbar.

(b) Sei  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ , dann ist

$$P_A(\lambda) = \det(A - \lambda E_2) = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = (\lambda + 1)(\lambda - 1).$$

Die Eigenwerte sind also  $\lambda_1 = 1$  und  $\lambda_2 = -1$  und die zugehörigen Eigenräume liest man in diesem Fall einfach ab:

$$E_{\lambda_1} = \left\{ \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid \alpha \in \mathbb{R} \right\} \quad \text{und} \quad E_{\lambda_2} = \left\{ \alpha \begin{pmatrix} 1 \\ -1 \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}.$$

### Anwendung: System linearer Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten.

Die Differentialgleichungen

$$\begin{aligned} \dot{x}_1(t) &= a_{11}x_1(t) + \dots + a_{1n}x_n(t) \\ &\vdots \\ \dot{x}_n(t) &= a_{n1}x_1(t) + \dots + a_{nn}x_n(t) \end{aligned}$$

sind gekoppelt und in Kurzform schreiben wir

$$\dot{x}(t) = Ax(t).$$

Es sei die Anfangsbedingung  $x(0) = a \in \mathbb{K}^n$  gegeben und wir suchen die Lösung für alle anderen Zeiten, also eine Funktion  $x : \mathbb{R} \rightarrow \mathbb{K}^n$  welche  $\frac{d}{dt}x(t) = Ax(t)$  und  $x(0) = a$  erfüllt.

Falls  $A$  diagonalisierbar ist, also

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = D$$

für eine reguläre Matrix  $S$  gilt, dann ergibt sich für den Vektor  $y(t) = S^{-1}x(t)$

$$\dot{y}(t) = S^{-1}\dot{x}(t) = S^{-1}Ax(t) = S^{-1}ASS^{-1}x(t) = S^{-1}ASy(t) = Dy(t).$$

Wir erhalten also die entkoppelten Differentialgleichungen

$$\dot{y}_1(t) = \lambda_1 y_1(t), \dots, \dot{y}_n(t) = \lambda_n y_n(t).$$

Die Anfangsdaten für  $y$  ergeben sich aus denen für  $x$  durch

$$y(0) = S^{-1}x(0) = S^{-1}a =: b.$$

Die Lösungen für das entkoppelte System kann man direkt ablesen,

$$y_1(t) = b_1 e^{\lambda_1 t}, \dots, y_n(t) = b_n e^{\lambda_n t}$$

oder kurz

$$y(t) = e^{Dt}b := \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Durch Rücktransformation bekommt man auch wieder

$$x(t) = Sy(t) = Se^{Dt}b = Se^{Dt}S^{-1}a =: e^{At}a.$$

### 3.3 Symmetrische Matrizen und lokale Minima

Wir betrachten im folgenden symmetrische Matrizen über  $\mathbb{K} = \mathbb{R}$ . Also  $A \in M(n, \mathbb{R})$ . Bisher spielte Orthogonalität keine Rolle. Wichtig war nur die lineare Unabhängigkeit von Basisvektoren. Es stellt sich heraus, dass symmetrische Matrizen vor allem dadurch charakterisiert sind, dass ihre Eigenvektoren orthogonal aufeinander stehen. Zu diesem Zwecke verwenden wir hier das innere Produkt von Vektoren. Da Vektoren nun auch komplex sein können, müssen wir das normale innere Produkt, das man aus der Schule kennt auf komplexe Vektoren anpassen.

Sei nun das innere Produkt von nicht notwendigerweise reellen Vektoren  $u, v \in \mathbb{C}^n$  definiert als

$$\langle u, v \rangle = \sum_{i=1}^n \bar{u}_i v_i.$$

Dies ist notwendig, weil Eigenwerte und Vektoren a-priori komplex sein können.

#### 3.36 Definition. Symmetrie und Selbstadjungiertheit

- (a) Eine  $n \times n$ -Matrix  $A$  heißt **symmetrisch** oder **selbstadjungiert**, wenn  $A = A^*$  gilt. Also wenn  $a_{ij} = \bar{a}_{ji}$ .
- (b) Eine lineare Abbildung  $T \in \mathcal{L}(V)$  auf einem Skalarproduktraum  $V$  heißt **symmetrisch** oder **selbstadjungiert**, wenn

$$\langle Tu, v \rangle = \langle u, Tv \rangle \text{ für alle } u, v \in V.$$

Wir werden zeigen, dass selbstadjungierte Matrizen auf  $\mathbb{R}^n$  immer diagonalisierbar sind. Das folgt im Wesentlichen schon aus den folgenden sehr simplen Beobachtungen:

**3.37 Satz.** Sei  $A \in M(n, \mathbb{R})$  symmetrisch. Dann gilt:

- (a) Alle Eigenwerte von  $A$  sind reell.
- (b) Eigenvektoren von  $A$  zu verschiedenen Eigenwerten sind orthogonal, und können reell gewählt werden.
- (c) Ist  $u$  ein Eigenvektor von  $A$ , so ist der zu  $u$  orthogonale Teilraum

$$W = \{u\}^\perp = \{v \in \mathbb{R}^n \mid \langle u, v \rangle = 0\}$$

$A$ -invariant, d.h.  $A(W) \subset W$  (also  $Aw \in W \forall w \in W$ ).

*Beweis.* (a) Sei  $Au = \lambda u$  mit  $u \neq 0$ . Sei nun das innere Produkt von nicht notwendigerweise reellen Vektoren  $u, v$  definiert als

$$\langle u, v \rangle = \sum_{i=1}^n \bar{u}_i v_i.$$

Dies ist notwendig, weil Eigenwerte und Vektoren a-priori komplex sein können. Da  $A$  nun selbstadjungiert ist, und sei  $Au = \lambda u$  mit  $u \neq 0$ , dann gilt

$$\lambda \langle u, u \rangle = \langle u, \lambda u \rangle = \langle u, Au \rangle = \langle Au, u \rangle = \langle \lambda u, u \rangle = \bar{\lambda} \langle u, u \rangle,$$

also folgt mit  $\langle u, u \rangle \neq 0$ , dass  $\lambda = \bar{\lambda}$ .

- (b) Sei  $Au = \lambda u$  und  $Av = \mu v$  für  $\lambda \neq \mu$  und  $u, v \neq 0$ , dann ist

$$\lambda \langle u, v \rangle = \langle \lambda u, v \rangle = \langle Au, v \rangle = \langle u, Av \rangle = \langle u, \mu v \rangle = \mu \langle u, v \rangle,$$

also folgt mit  $\lambda \neq \mu$ , dass  $\langle u, v \rangle = 0$ . Weiters gilt dass  $u + \bar{u}$  ebenfalls Eigenvektor zu Eigenwert  $\lambda$  ist, und per Definitionem reell.

- (c) Sei  $Au = \lambda u$  und  $w \in W$ , also  $\langle w, u \rangle = 0$ . Dann ist auch

$$\langle Aw, u \rangle = \langle w, Au \rangle = \langle w, \lambda u \rangle = \lambda \langle w, u \rangle = 0,$$

also  $Aw \in W$ .

□

**3.38 Bemerkung.** Es sei bemerkt, dass ganz allgemein gilt, dass lineare Abbildungen  $L$  orthogonale Komplemente von Eigenvektoren  $\{u\}^\perp$  nur dann invariant lassen können, falls die anderen orthogonal auf  $u$  stehen.

**3.39 Satz.** Jeder symmetrische Matrix  $A \in M(n, \mathbb{R})$  besitzt wenigstens einen Eigenwert (welcher dann gemäß Satz 3.37 reell ist.)

*Beweis.* Das charakteristische Polynom  $P_A$  von  $A$  sicherlich eine Nullstelle  $\lambda \in \mathbb{C}$ . Diese ist Eigenwert der Matrix  $A$ . Nach Satz 3.37 ist aber  $\lambda$  reell. □

**3.40 Satz. Diagonalisierbarkeit symmetrischer Matrizen**

Jede symmetrische Matrix  $A \in M(n, \mathbb{R})$  ist diagonalisierbar, besitzt also eine Basis aus zueinander orthogonalen Eigenvektoren.

*Beweis.* Nach Satz 3.39 hat  $A$  einen reellen Eigenwert  $\lambda_1$ , und zugehörigen Eigenvektor  $u_1$ . Der orthogonale Raum  $V_1 : \{u_1\}^\perp$  ist  $n - 1$  dimensional und invariant unter  $A$ . Das heisst die lineare Abbildung  $A : V_1 \rightarrow V_1$  lässt sich durch eine  $(n - 1) \times (n - 1)$  Matrix  $A_2$  darstellen. Diese ist wieder symmetrisch, und hat somit einen reellen Eigenwert  $\lambda_2$ , mit Eigenvektor  $u_2$ . usw. Nach  $n$  Schritten. gelangen wir zu  $n$  orthogonalen Eigenvektoren  $\{u_1, \dots, u_n\}$ . Sei  $S$  die Matrix  $S = (u_1, u_2, \dots, u_n)$ . Dann ist leicht zu sehen dass

$$S^{-1}AS = \text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Also die Diagonalmatrix mit Einträgen der Eigenwerte auf der Diagonalen. □

**3.41 Definition.** Eine reelle Matrix  $A \in M(n, \mathbb{R})$

- (a) ist **positiv definit** definit, wenn alle Eigenwerte positiv sind.
- (b) **negativ definit** falls alle Eigenwerte negativ sind.
- (c) **indefinit** wenn es sowohl positive als auch negative Eigenwerte gibt.

Das heisst falls eine Matrix 0-Eigenwerte hat, dann fällt sie in keine der obigen Kategorien.

**3.42 Bemerkung.** Man nennt die Nullstellen von  $\text{grad} f$  auch die **kritischen Punkte** von  $f$ .

**3.43 Korollar.** Eine reelle symmetrische  $n \times n$ -Matrix  $A = (a_{ij})$  ist

- (i) **positiv definit**, wenn für alle  $x \in \mathbb{R}^n \setminus \{0\}$  gilt, dass  $\langle x, Ax \rangle_{\mathbb{R}^n} = \sum_{i,j=1}^n a_{ij} x_i x_j > 0$ .
- (ii) **negativ definit**, wenn für alle  $x \in \mathbb{R}^n \setminus \{0\}$  gilt, dass  $\langle x, Ax \rangle_{\mathbb{R}^n} < 0$ .
- (iii) **indefinit**, wenn es  $x, y \in \mathbb{R}^n$  gibt mit  $\langle x, Ax \rangle > 0$  und  $\langle y, Ay \rangle < 0$ .

*Beweis.* Beweis: Übung. □

**3.44 Satz.** Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f \in C^2(G, \mathbb{R})$ . Sei  $x \in G$  eine Nullstelle von  $\text{grad} f$ , also  $\text{grad} f(x) = 0$ .

- (a) Ist  $\text{Hess} f(x)$  positiv definit, so hat  $f$  in  $x$  ein lokales Minimum.
- (b) Ist  $\text{Hess} f(x)$  negativ definit, so hat  $f$  in  $x$  ein lokales Maximum.
- (c) Ist  $\text{Hess} f(x)$  indefinit, so hat  $f$  in  $x$  kein lokales Extremum, also  $x$  ist ein Sattelpunkt.

**3.45 Lemma.** Sei  $A \in \text{Sym}_n$  positiv definit und  $\lambda_0 > 0$  der kleinste Eigenwert von  $A$ . Dann gilt für alle  $x \in \mathbb{R}^n$ , dass

$$\langle x, Ax \rangle \geq \lambda_0 \|x\|^2.$$

*Beweis.* Stelle  $x$  in einer Orthonormalbasis aus Eigenvektoren dar. □

*Beweis.* von Satz 3.44. (a) Sei  $\delta > 0$  so klein, dass  $B_\delta(x) \subset G$  ist. Nach Korollar 2.50 gilt dann für  $h \in B_\delta(0)$

$$f(x+h) = f(x) + \frac{1}{2} \langle h, \text{Hess}f(x)h \rangle + \varphi(h)$$

mit  $\varphi(h) = o(\|h\|^2)$ . Wähle nun  $0 < \delta' < \delta$  so klein, dass  $|\varphi(h)| \leq \frac{\lambda_0}{4} \|h\|^2$  für alle  $h \in B_{\delta'}(0)$  gilt, wobei  $\lambda_0 > 0$  der kleinste Eigenwert von  $\text{Hess}f(x)$  sei. Dann ist für alle  $h \in B_{\delta'}(0) \setminus \{0\}$

$$\begin{aligned} f(x+h) &= f(x) + \frac{1}{2} \langle h, \text{Hess}f(x)h \rangle + \varphi(h) \\ &\geq f(x) + \frac{\lambda_0}{2} \|h\|^2 - \frac{\lambda_0}{4} \|h\|^2 \\ &= f(x) + \frac{\lambda_0}{4} \|h\|^2 > f(x). \end{aligned}$$

Also ist  $f(x)$  ein striktes lokales Minimum.

(b) Ist  $\text{Hess}f(x)$  negativ definit, so ist  $\text{Hess}(-f)(x) = -\text{Hess}f(x)$  positiv definit. Also hat  $-f$  ein lokales Minimum und somit  $f$  ein lokales Maximum.

(c) Laut Voraussetzung gibt es mindestens einen normierten Eigenvektor  $h_1$  von  $\text{Hess}f(x)$  zu positiven Eigenwert  $\lambda_1 > 0$  und einen normierten Vektor  $h_2$  zu negativen Eigenwert  $\lambda_2 < 0$ . Dann sieht man zum Beispiel, dass

$$\begin{aligned} f(x+th_2) &= f(x) + \frac{1}{2} \langle th_2, \text{Hess}f(x)th_2 \rangle + \varphi(th) \\ &= f(x) + \lambda_2/2t^2 + o(t^2) \end{aligned}$$

gilt und somit  $f(x+th_2) < f(x)$  in kleiner Umgebung. Umgekehrt folgt, dass in Richtung  $h_1$  die Funktion  $f(x+th_1)$  ein lokales Minimum besitzt als Funktion von  $t$ . Damit ist  $x$  ein Sattelpunkt. also

$$f(x+th) \geq f(x) + \frac{\alpha}{4} t^2 > f(x)$$

für  $0 < t < \delta$  und  $\delta$  klein genug. Also gibt es in jeder Umgebung von  $x$  einen Punkt  $y = x+th$ , sodass  $f(y) > f(x)$  ist. Das gleiche Argument für  $\tilde{h}$  mit  $\tilde{\alpha} = \langle \tilde{h}, \text{Hess}f(x)\tilde{h} \rangle < 0$  zeigt, dass in jeder Umgebung von  $x$  auch ein Punkt  $\tilde{y}$  mit  $f(\tilde{y}) < f(x)$  liegt. Also hat  $f$  in  $x$  kein lokales Extremum.  $\square$

### 3.46 Beispiel. Minimum und Sattel

- (a) Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 + y^2 + c$ . Dann ist  $\text{grad}f(x, y) = (2x, 2y)$ , also  $\text{grad}f(x, y) = 0 \Leftrightarrow (x, y) = (0, 0)$ . Weiterhin ist

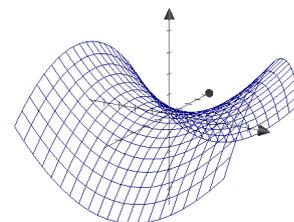
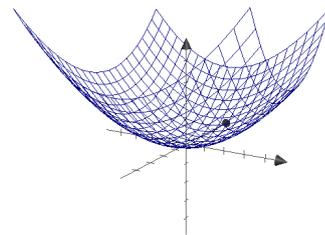
$$\text{Hess}f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \text{Hess}f(0, 0)$$

positiv definit. Somit hat  $f$  bei  $(0, 0)$  ein striktes lokales Minimum.

- (b) Die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 - y^2 + c$  erfüllt  $\text{grad}f(0, 0) = (0, 0)$  und

$$\text{Hess}f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Also hat  $f$  in  $(0, 0)$  kein lokales Extremum.



**3.47 Bemerkung.** Ist  $\text{Hess}f(x)$  nur positiv semidefinit, gilt also nur  $\langle h, \text{Hess}f(x)h \rangle \geq 0$  für alle  $h \in \mathbb{R}^n \setminus \{0\}$  und  $\text{grad}f(x) = 0$ , so kann man noch nicht entscheiden, ob in  $x$  ein lokales Minimum vorliegt. Beispiele:

### 3 Einschub: Lineare Algebra

- (i)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 + y^4$  hat lokales Minimum bei  $(0, 0)$
- (ii)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 + y^3$  hat kein lokales Minimum bei  $(0, 0)$
- (iii)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2$  hat lokales entartetes Minimum bei  $(0, 0)$ .

Wegen Satz 3.44 (c) ist aber die Semidefinitheit ein notwendiges Kriterium für das Vorliegen eines Extremums.

# 4 Implizite Funktionen

Eine Funktion  $g : \mathbb{R} \rightarrow \mathbb{R}$  ist nicht immer „explizit“ in der Form  $y = g(x)$  gegeben, sondern häufig nur „implizit“ durch eine Gleichung der Form

$$F(x, g(x)) = 0.$$

Hierbei wäre  $F : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto F(x, y)$  eine „explizit“ gegebene Funktion. Man möchte dann die implizite Gleichung  $F(x, y) = 0$  explizit machen, d.h. „nach  $y$  auflösen“ und in der Form  $y = g(x)$  schreiben.

**4.1 Beispiel.** Betrachte  $F : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 + y^2 - 1$ . Dann ist

$$C = \{(x, y) \in \mathbb{R}^2 \mid F(x, y) = 0\}$$

die Einheitskreislinie in  $\mathbb{R}^2$ .  $C$  ist aber nicht der Graph einer Funktion  $g : \mathbb{R} \rightarrow \mathbb{R}$ , denn

- (i) Für  $x \in \mathbb{R}$  mit  $|x| > 1$  gibt es kein  $y \in \mathbb{R}$  mit  $(x, y) \in C$ .
- (ii) Für  $x \in \mathbb{R}$  mit  $|x| < 1$  gibt es gleich zwei  $y \in \mathbb{R}$  mit  $(x, y) \in C$ , nämlich  $y = \pm\sqrt{1 - x^2}$ .

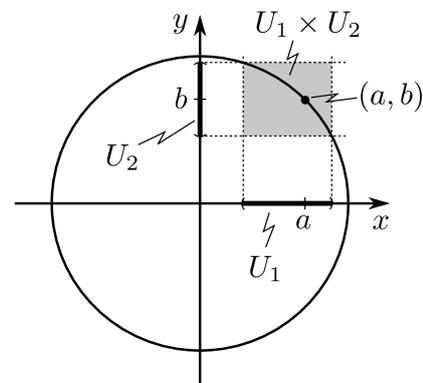
Gibt man allerdings  $(a, b) \in C$  mit  $b \neq 0$  vor, so kann man  $F(x, y)$  lokal um  $(a, b)$  nach  $y$  auflösen, d.h. es gibt eine Umgebung  $U_1 \subset \mathbb{R}$  von  $a$  und eine Umgebung  $U_2$  von  $b$  und eine Funktion  $g : U_1 \rightarrow U_2$  so, dass für alle  $(x, y) \in U_1 \times U_2$  gilt

$$F(x, y) = 0 \iff y = g(x).$$

Ist etwa  $b > 0$ , so wähle  $\varepsilon > 0$  klein genug,  $U_1 = (a - \varepsilon, a + \varepsilon)$ ,  $U_2 = (b - \varepsilon, b + \varepsilon)$  und

$$g(x) = +\sqrt{1 - x^2}.$$

**Aber:** Für  $b = 0$  kann man in keiner Umgebung von  $(1, 0)$  oder  $(-1, 0)$  nach  $y$  auflösen, wohl aber nach  $x$ , durch  $x = \pm\sqrt{1 - y^2}$ . Die Punkte  $(\pm 1, 0) \in C$  sind aber genau die Punkte, wo  $\frac{\partial F}{\partial y} = 2y$  verschwindet, d.h. die Tangente an  $C$  vertikal ist.



Der folgende Satz über implizite Funktionen gibt eine hinreichende Bedingung dafür an, dass man eine implizite Gleichung  $F(x, y) = 0$  lokal um einen Punkt  $(a, b)$  mit  $F(a, b) = 0$  in der Form  $y = g(x)$  explizit machen kann.

## 4.2 Satz. Satz über implizite Funktionen

Sei  $G \subset \mathbb{R}^{n+m} = \mathbb{R}_x^n \times \mathbb{R}_y^m$  ein Gebiet und  $F : G \rightarrow \mathbb{R}^m$  eine stetig partiell differenzierbare Funktion. Sei  $(a, b) \in G$  derart, dass  $F(a, b) = 0$  ist und

$$\frac{\partial F}{\partial y}(a, b) := \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \dots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \dots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}(a, b)$$

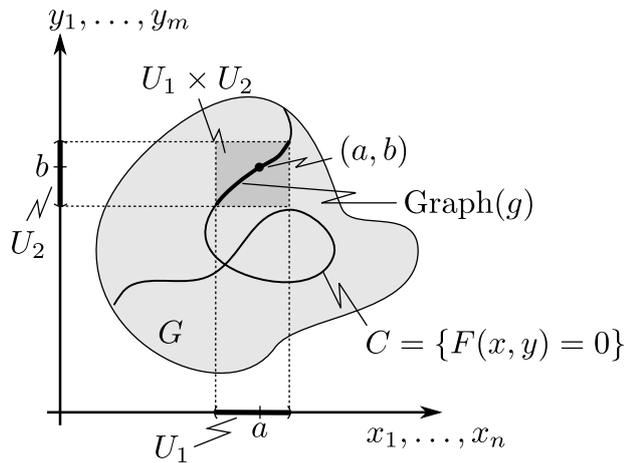
invertierbar ist.

Dann existieren offene Umgebungen  $U_1 \subset \mathbb{R}^n$  von  $a$  und  $U_2 \subset \mathbb{R}^m$  von  $b$  mit  $U_1 \times U_2 \subset G$  und eine stetig partiell differenzierbare Funktion  $g : U_1 \rightarrow U_2$ , sodass für alle  $(x, y) \in U_1 \times U_2$  gilt:

$$F(x, y) = 0 \iff y = g(x).$$

Das Differential von  $g$  ist

$$Dg(x) = -[\partial_y F(x, g(x))]^{-1} \partial_x F(x, g(x)).$$



**4.3 Bemerkung.** Die Bedingung  $\det\left(\frac{\partial F}{\partial y}(a, b)\right) \neq 0$  ist gleichbedeutend damit, dass keine Richtungsableitung von  $F$  in „ $y$ -Richtung“ verschwindet, also  $\vec{v} \cdot \frac{\partial F}{\partial y}(a, b) \neq 0$  für alle  $\vec{v} \in \mathbb{R}^m \setminus \{0\}$ . Dadurch wird sichergestellt, dass es eine Umgebung  $U_2$  von  $b$  gibt, so dass für  $y \in U_2$

$$F(a, y) = 0 \iff y = b.$$

**4.4 Bemerkung.** Die Methode, um den Hauptsatz zu beweisen, ist das Newton-Verfahren zum Auffinden von Nullstellen. Die Idee ist, dass man von einem Startpunkt  $x_0$  startet und falls  $f'(x_0) \neq 0$  ist, dann definiert man  $x_1$  durch den Schnittpunkt der Tangente an  $x_0$  mit der  $x$ -Achse, also  $x_1 = x_0 - f(x_0)/f'(x_0)$ , und allgemeiner

$$x_{n+1} = x_n - f(x_n)/f'(x_n).$$

Besitzt  $f$  nun auf dem Intervall  $B_r(x_0)$  eine stetige Ableitung dritter Ordnung und die erste Ableitung verschwinde in keinem Punkt von  $B_r$ . Gebe es weiters eine Konstante  $q < 1$  sodass

$$\sup_{x \in \bar{B}_r} \frac{|f(x)f''(x)|}{|f'(x)|^2} \leq q, \quad \frac{|f(x_0)|}{|f'(x_0)|} \leq (1 - q)r,$$

Besitzt die Gleichung  $f(x) = 0$  genau eine Lösung in  $B_r(x_0)$ .

*Beweis.* Man wendet den Banach'schen Fixpunktsatz auf  $g(x) = x - f(x)/f'(x)$  an. Dies ist möglich, weil

$$\|g(x) - g(y)\| \leq \sup_{\xi \in B_r} \|g'(\xi)\| \|x - y\|$$

und  $|g'(x)| = |f(x)f''(x)|/|f'(x)|^2 \leq q$  ist, also eine Kontraktionskonstante  $q$  besitzt und mit

$$\|g(x) - x_0\| \leq \|g(x) - g(x_0)\| + \|g(x_0) - x_0\| \leq q\|x - x_0\| + \frac{|f(x_0)|}{|f'(x_0)|} \leq qr + (1 - q)r \leq r$$

Also  $g : \bar{B}_r(x_0) \rightarrow \bar{B}_r(x_0)$  eine Selbstabbildung und Kontraktion, somit gibt es genau einen Fixpunkt  $g(x) = x \iff f(x) = 0$  in  $B_r$ . □

**4.5 Bemerkung.** Mit etwas stärkeren Voraussetzungen kann man auch etwas anders vorgehen. Sei beispielsweise  $|1 - f(x)/f'(x_0)| \leq 1/2$  auf  $B_\delta(x_0)$ . Sei weiters  $\|f(x_0)/f'(x_0)\| \leq \delta/2$ , dann kann man eine Nullstelle in  $B_\delta(x_0)$  auch mittels der etwas vereinfachten Funktion  $\varphi(x) = x - f(x)/f'(x_0)$  finden. Diese Vorgehensweise ist für den Hauptsatz über implizite Funktionen ausreichend, weil wir nämlich bei der Gleichung  $F(x, y) = 0$ , einen zweiten Parameter  $x$  haben, dessen Intervall wir so anpassen können, dass die Voraussetzungen erfüllt werden können.

*Beweis. von Satz 4.2.* Wir zerlegen den Beweis in die folgenden drei Schritte:

- (a) Wir werden zunächst für  $U_1, U_2$  klein genug mit Hilfe des Banachschen Fixpunktsatzes zeigen, dass  $g : U_1 \rightarrow U_2$  mit  $F(x, g(x)) = 0$  existiert und eindeutig bestimmt ist.

(b) Wiederum mit dem Banachschem Fixpunktsatz zeigen wir dann die Stetigkeit von  $g$ .

(c) Die stetige Differenzierbarkeit von  $g$  folgt dann aus der stetigen Differenzierbarkeit von  $F$ .

Zu (a): Betrachte die Funktion  $\Phi : G \rightarrow \mathbb{R}^m$  gegeben durch

$$\Phi(x, y) = y - B^{-1}F(x, y)$$

wobei  $B = \frac{\partial F}{\partial y}(a, b)$ . Dann gilt

$$\begin{aligned} F(x, y) = 0 &\Leftrightarrow B^{-1}F(x, y) = 0 \\ &\Leftrightarrow y - B^{-1}F(x, y) = y \\ &\Leftrightarrow \Phi(x, y) = y. \end{aligned}$$

Also ist  $F(x, y) = 0$ , genau dann wenn  $y$  ein Fixpunkt der Abbildung  $y \rightarrow \Phi(x, y)$  ist.

Weil  $\partial_y \Phi(a, b) = E_m - B^{-1}\partial_y F(a, b) = E_m - B^{-1}B = 0$  ist, ist  $\|\partial_y \Phi(x, y)\| \leq \frac{1}{2}$  in einer hinreichend kleinen Umgebung von  $(a, b)$ , die wiederum das Produkt zweier abgeschlossener Kugeln  $\overline{B}_{r_1}(a) \times \overline{B}_{r_2}(b)$  enthält. Wir präsentieren der Einfachheit halber den weiteren Beweis für  $n = 2$ , und  $m = 1$ . Also  $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ .

Wir betrachten nun die Abbildung

$$y \rightarrow \phi(x, y) = \varphi_x(y) = y - B^{-1}F(x, y). \quad (4.1)$$

Der Mittelwertssatz impliziert dann, dass für  $(x, y)$  und  $(x, y')$  in dieser Umgebung

$$\|\varphi_x(y) - \varphi_x(y')\| = \|\Phi(x, y) - \Phi(x, y')\| \leq \sup_{(x, y) \in \overline{B}_{r_1} \times \overline{B}_{r_2}} \|\partial_y \Phi(x, y)\| \|y - y'\| \leq \frac{1}{2} \|y - y'\| \quad (4.2)$$

gilt. Daraus folgt nun wieder, dass

$$\Phi(x, \cdot) = \varphi_x(\cdot) : \overline{B}_{r_2}(b) \rightarrow \overline{B}_{r_2}(b)$$

für jedes feste  $x$  in einer geeigneten Kugel  $\overline{B}_{r_3}(a) \subset \overline{B}_{r_1}(a)$  eine Kontraktion ist: für  $y \in \overline{B}_{r_2}(b)$  ist

$$\|\varphi_x(y) - b\| \leq \|\varphi_x(y) - \varphi_x(b)\| + \|\varphi_x(b) - b\| \leq \frac{1}{2} \|y - b\| + \frac{1}{2} r_2 \leq r_2,$$

wenn man  $\overline{B}_{r_3}(a)$  so klein wählt, dass  $\|\varphi_x(b) - b\| \leq r_2/2$  ist für  $x \in \overline{B}_{r_3}(a)$ . Letzteres ist wegen  $\Phi(a, b) = 0$  und der Stetigkeit von  $\Phi$  immer möglich. Nach dem Banachschem Fixpunktsatz hat  $\Phi(x, \cdot)$  genau einen Fixpunkt, den wir  $g(x)$  nennen. Also gibt es zu jedem  $x \in \overline{B}_{r_3}$  genau ein  $y = g(x) \in \overline{B}_{r_2}$ , sodass  $F(x, y) = 0$  ist. Insbesondere ist  $g(a) = 0$ .

Wir führen die Anwendung des Banach'schen Fixpunktsatzes explizit aus: Sei

$$g_n(x) = \varphi_x \circ \dots \circ \varphi_x(b) = \varphi_x^n(b),$$

oder  $g_n(x) = \varphi_x(g_{n-1}(x))$ , dann ist jetzt mittels (4.2) leicht zu sehen, dass

$$\|g_n(x) - g_m(x)\| \leq \|g_n(x) - g_{n-1}(x)\| + \dots + \|g_{m+1}(x) - g_m(x)\| \leq \frac{1}{2^{m-1}} \|\varphi_x(b) - b\| \leq \frac{1}{2^{m-1}} r_2, \quad (4.3)$$

und somit  $g_n(x)$  Cauchyfolge. Damit existiert ein Limes  $g(x)$  mit  $g_n(x) \rightarrow_{n \rightarrow \infty} g(x)$ , und  $g(x) = \lim g_n(x) = \lim \varphi_x(g_n(x)) = \varphi_x(g(x))$ , aufgrund der Stetigkeit von  $\varphi_x(y)$ , was eine Folgerung aus der Stetigkeit von  $F(x, y)$  ist. Damit gilt somit  $F(x, g(x)) = 0$ .

Zu (b):

Wir haben in (4.3) sogar gezeigt, dass  $g_n$  gleichmässig gegen  $g(x)$  geht, da der Fehlerterm nicht von  $x$  abhängt, damit folgt aus dem letzten Semester dass  $g(x)$  stetig ist.

Dies sieht man explizit aufgrund von

$$\|g(x) - g(y)\| \leq \|g(x) - g_n(x)\| + \|g_n(x) - g_n(y)\| + \|g_n(y) - g(y)\|,$$

#### 4 Implizite Funktionen

und der Tatsache, dass  $g_n(x)$  stetig ist und damit gleichmäßig stetig auf  $\overline{B_{r_3}(a)}$ , damit gibt es für beliebiges  $\varepsilon$ , ein  $n$  sodass  $\|g(z) - g_n(z)\| \leq \frac{\varepsilon}{3}$ , und ein  $\delta$ , so dass  $\|g_n(x) - g_n(y)\| \leq \frac{\varepsilon}{3}$ , wenn  $\|x - y\| \leq \delta$ .

Zu (c): Da  $F$  stetig differenzierbar ist, gilt mit dem Mittelwertsatz

$$\begin{aligned} 0 &= F(x, g(x)) - F(x_0, g(x_0)) = F(x, g(x)) - F(x_0, g(x)) + F(x_0, g(x)) - F(x_0, g(x_0)) \\ &= \partial_x F(\tilde{x}_j, g(x)) \cdot (x - x_0) + \partial_y F_j(x_0, \tilde{y}_j) \cdot (g(x) - g(x_0)) \end{aligned}$$

für geeignete Zwischenstellen  $\tilde{x}_j, \tilde{y}_j$ . Also ist

$$\frac{g(x) - g(x_0)}{x - x_0} = \partial_y F_j(x_0, g(\bar{x}))^{-1} \partial_x F(\tilde{x}_j, g(x)),$$

und die rechte Seite ist stetig für  $x \rightarrow x_0$  was schließlich die Existenz und Stetigkeit der partiellen Ableitungen von  $g$  zeigt.  $\square$

**4.6 Beispiel.** Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}$  stetig differenzierbar. Wir betrachten die Niveaufläche

$$N_c = \{x \in \mathbb{R}^n \mid f(x) = c\}$$

zu  $c \in \mathbb{R}$  von  $f$ . Ist nun  $\text{grad}(f)(a) \neq 0$  für  $a \in N_c$ , so behaupten wir, dass  $N_c$  bei  $a$  lokal aussieht wie der Graph einer stetig differenzierbaren Funktion  $g$  in  $(n - 1)$  Veränderlichen. Sei z.B.  $\frac{\partial f}{\partial x_n}(a) \neq 0$  und  $\bar{x} = (x_1, \dots, x_{n-1})$  und  $y = x_n$ . Dann ist

$$x \in N_c \Leftrightarrow \underbrace{f(\bar{x}, y) - c}_{=: F(\bar{x}, y)} = 0$$

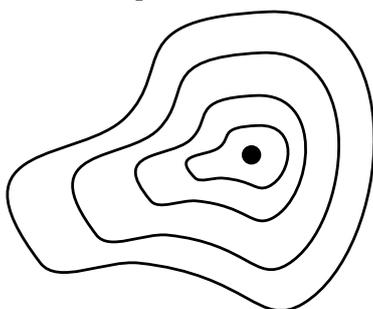
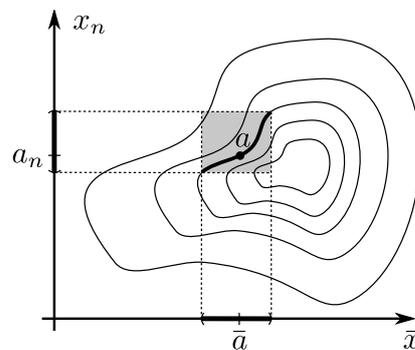
und

$$\frac{\partial F}{\partial y}(\bar{a}, a_n) = \frac{\partial f}{\partial x_n}(\bar{a}, a_n) \neq 0.$$

Nach dem Satz über implizite Funktionen existiert eine Umgebung  $\bar{U} \times U_n$  von  $a$  und eine  $C^1$  Funktion  $g : \bar{U} \rightarrow U_n$  mit  $F(\bar{x}, y) = 0 \Leftrightarrow g(\bar{x}) = y$ . Somit ist

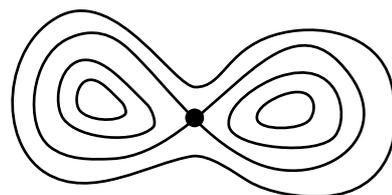
$$N_c \cap (\bar{U} \times U_n) = \text{Graph}(g) \cap (\bar{U} \times U_n).$$

An den kritischen Punkten von  $f$ , also dort wo  $\text{grad} f = 0$  ist, gilt das nicht: z.B. ist bei strikten lokalen Extrema  $N_c$  ein einzelner Punkt und bei einem Sattelpunkt hat  $N_c$  eine Selbstdurchschneidung.



← striktes Extremum

Sattelpunkt →



Wir kommen nun zu der Frage nach der lokalen Umkehrbarkeit einer stetig differenzierbaren Funktion  $f : \mathbb{R}^n \supset G \rightarrow D \subset \mathbb{R}^n$ . Insbesondere ist es oft wichtig, z.B. bei Koordinationstransformationen, dass auch  $g = f^{-1} : D \rightarrow G$  wieder stetig differenzierbar ist.

#### 4.7 Definition. Diffeomorphismus

Seien  $G, D \subset \mathbb{R}^n$  Gebiete. Eine stetig differenzierbare Abbildung  $f : G \rightarrow D$  heißt **Diffeomorphismus**, wenn  $f$  bijektiv und  $f^{-1} : D \rightarrow G$  stetig differenzierbar ist.

**4.8 Bemerkung.** Ist  $f : G \rightarrow D$  ein Diffeomorphismus, so ist für jedes  $x \in G$  die Ableitung  $Df(x)$  invertierbar und es gilt mit  $g = f^{-1}$ , dass  $(Df(x))^{-1} = Dg(f(x))$ .

*Beweis.* Wegen  $g \circ f = \text{id}$  liefert die Kettenregel

$$E = D(\text{id}) = D(g \circ f) = Dg \circ f \cdot Df.$$

□

**4.9 Bemerkung.** Nicht jede stetig differenzierbare Bijektion hat auch eine stetig differenzierbare Umkehrung:  $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^3$ , ist bijektiv, stetig differenzierbar, die Umkehrabbildung  $g : \mathbb{R} \rightarrow \mathbb{R}, y \mapsto \sqrt[3]{y}$  ist aber im Nullpunkt nicht differenzierbar. Beachte, dass die nach Bemerkung 4.8 notwendige Bedingung  $f' \neq 0$  für  $x = 0$  nicht erfüllt ist, und somit  $f$  kein Diffeomorphismus sein kann.

#### 4.10 Satz. Satz über die Umkehrabbildung

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}^n$  eine stetig differenzierbare Funktion. Ist nun  $a \in G$  so, dass  $Df(a) \in M(n \times n, \mathbb{R})$  invertierbar ist, so existiert eine offene Umgebung  $U \subset G$  von  $a$  so, dass  $f|_U : U \rightarrow V$  mit  $V = f(U)$  ein Diffeomorphismus ist.

*Beweis.* Sei  $b := f(a)$ . Man möchte die Gleichung  $y = f(x)$  in einer Umgebung von  $(x, y) = (a, b)$  nach  $x$  auflösen, also  $x = g(y)$  schreiben. Dazu betrachtet man die Abbildung

$$F : G \times \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$F(x, y) = y - f(x).$$

Es ist

$$\frac{\partial F}{\partial x}(a, b) = -Df(a)$$

nach Voraussetzung invertierbar. Deshalb liefert der Satz über implizite Funktionen Umgebungen  $\tilde{U} \subset G$  von  $a$  und  $V \subset \mathbb{R}^n$  von  $b$  und ein stetig differenzierbares  $g : V \rightarrow \tilde{U}$ , so dass für alle  $(x, y) \in \tilde{U} \times V$  gilt

$$F(x, y) = 0 \Leftrightarrow x = g(y).$$

Setzt man noch  $U := f^{-1}(V) \cap \tilde{U}$ , so gilt für alle  $(x, y) \in U \times V$

$$f(x) = y \Leftrightarrow F(x, y) = 0 \Leftrightarrow x = g(y).$$

Also ist  $f|_U : U \rightarrow V$  bijektiv,  $g = f^{-1}$  und  $g$  stetig differenzierbar und somit  $f|_U$  ein Diffeomorphismus. □

#### 4.11 Beispiel. Kugelkoordinaten

Kugelkoordinaten des  $\mathbb{R}^3$  sind gegeben durch

$$f : (0, \infty) \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^3, \quad (r, \vartheta, \varphi) \mapsto (r \sin \vartheta \cos \varphi, r \sin \vartheta \sin \varphi, r \cos \vartheta)$$

Eingeschränkt auf

$$G = (0, \infty) \times (0, \pi) \times (-\pi, \pi)$$

ist  $f$  injektiv und für die Funktionaldeterminante  $\det(Df) : G \rightarrow \mathbb{R}$  gilt

$$\det(Df) = \begin{vmatrix} \sin \vartheta \cos \varphi & r \cos \vartheta \cos \varphi & -r \sin \vartheta \sin \varphi \\ \sin \vartheta \sin \varphi & r \cos \vartheta \sin \varphi & r \sin \vartheta \cos \varphi \\ \cos \vartheta & -r \sin \vartheta & 0 \end{vmatrix} = r^2 \sin \vartheta$$

für alle  $(r, \vartheta, \varphi) \in G$ . Deshalb ist  $f$  ein lokaler Diffeomorphismus. Da  $f$  bijektiv ist, ist es aber auch global ein Diffeomorphismus auf sein Bild  $D := f(G) \subset \mathbb{R}^3$ .

Häufig sucht man Extrema von Funktionen  $f : \mathbb{R}^n \supset G \rightarrow \mathbb{R}$  unter einer Nebenbedingung, die man durch  $h(x) = 0$  für eine geeignete Funktion  $h : G \rightarrow \mathbb{R}$  beschreiben kann. Beispielsweise sucht man das Maximum einer Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  auf der Einheitskreislinie in  $\mathbb{R}^2$ , also unter der Nebenbedingung  $h(x) = \|x\| - 1 = 0$ .

**4.12 Definition. Extrema unter Nebenbedingungen**

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und seien  $f, h : G \rightarrow \mathbb{R}$  stetig differenzierbar. Sei  $M = \{x \in G \mid h(x) = 0\}$  und  $a \in M$ . Man sagt,  $f$  habe bei  $a$  ein **lokales Maximum (bzw. Minimum) unter Nebenbedingung  $h = 0$** , wenn es eine offene Umgebung  $U \subset G$  von  $a$  gibt, sodass für alle  $x \in U \cap M$  gilt

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

Der Satz über implizite Funktionen liefert ein notwendiges Kriterium für das Vorliegen lokaler Extrema unter Nebenbedingungen.

**4.13 Satz. Satz über Extrema unter Nebenbedingungen**

Seien  $f, h : G \rightarrow \mathbb{R}$  stetig differenzierbar,  $G \subset \mathbb{R}^n$  ein Gebiet und  $a \in M = \{x \in G \mid h(x) = 0\}$ . Es habe  $f$  ein lokales Extremum unter der Nebenbedingung  $h = 0$  in  $a$  **und** es sei  $\text{grad}h(a) \neq 0$ . Dann gibt es ein  $\lambda \in \mathbb{R}$  mit

$$\text{grad}f(a) = \lambda \text{grad}h(a),$$

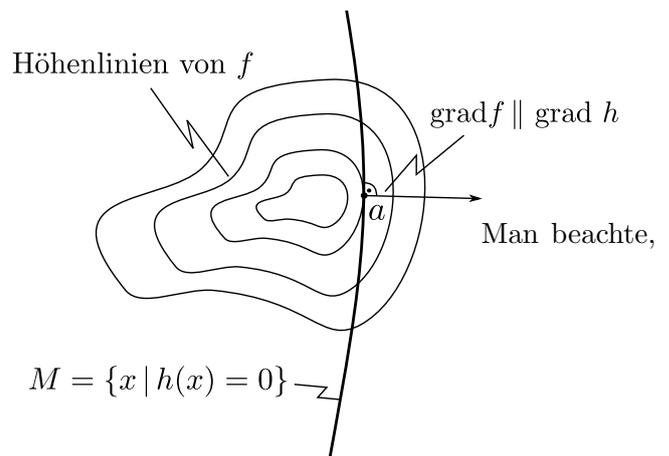
**4.14 Bemerkung. Geometrische Bedeutung**

Der Gradient von  $h$  in  $a$  steht senkrecht auf der Niveaulfläche  $M$  von  $h$ . Damit  $f$  auf  $M$  ein lokales Extremum hat, müssen nur die Richtungsableitungen  $\langle v, \nabla f(a) \rangle$  von  $f$  tangential an  $M$  verschwinden, also

$$\langle v, \nabla f(a) \rangle = 0 \quad \text{falls} \quad \langle v, \nabla h(a) \rangle = 0,$$

d.h. der Gradient von  $f$  muss auf  $M$  senkrecht stehen. Da  $M$  Kodimension 1 hat, folgt daraus

$$\nabla f(a) \parallel \nabla h(a).$$



dass Satz 4.13 das Analogon zu Proposition 2.53 für den Fall ohne Nebenbedingungen ist. Es ist  $\text{grad}f(a) = \lambda \text{grad}h(a)$  eine notwendige, aber keine hinreichende Bedingung für das Vorliegen eines lokalen Extremums unter der Nebenbedingung  $h = 0$ . Eine hinreichende Bedingung ist beispielsweise wieder, dass zusätzlich zu  $\text{grad}f(a) = \lambda \text{grad}h(a)$  die Hessematrix  $\text{Hess}f(a)$  eingeschränkt auf das orthogonale Komplement des Aufspans von  $\nabla h$  positiv bzw. negativ definit ist.

*Beweis.* von Satz 4.13. Da  $\text{grad}h(a) \neq 0$  ist, ist wenigstens eine partielle Ableitung von  $h$  in  $a$  von Null verschieden, sagen wir  $\frac{\partial h}{\partial x_n}(a) \neq 0$ . Sei  $\bar{a} := (a_1, \dots, a_{n-1})$ . Der Satz über implizite Funktionen liefert dann Umgebungen  $V \subset \mathbb{R}^{n-1}$  von  $\bar{a}$  und  $I \subset \mathbb{R}$  von  $a_n$  mit  $V \times I \subset G$  und ein stetig differenzierbares  $g : V \rightarrow I$ , so dass für alle  $(\bar{x}, x_n) \in V \times I$  gilt

$$h(\bar{x}, x_n) = 0 \quad \Leftrightarrow \quad x_n = g(\bar{x}).$$

Hat nun  $f : G \rightarrow \mathbb{R}$  ein lokales Extremum unter der Nebenbedingung  $h = 0$  in  $a \in M$ , so hat  $f(\bar{x}, g(\bar{x})) =: f \circ \varphi(\bar{x})$  ein lokales Extremum in  $\bar{x}$  (ohne Nebenbedingung). Hierbei ist  $\varphi : V \rightarrow M \subset \mathbb{R}^n$ ,  $\varphi(\bar{x}) = (\bar{x}, g(\bar{x}))$ . Es ist also

$$\text{grad}(f \circ \varphi)(\bar{a}) = 0.$$

Für  $i = 1, \dots, n - 1$  ergibt die Kettenregel

$$0 = \frac{\partial (f \circ \varphi)}{\partial x_i}(\bar{a}) = \sum_{j=1}^n \partial_j f(\varphi(\bar{a})) \frac{\partial \varphi_j}{\partial x_i}(\bar{a}) = \partial_i f(a) + \partial_n f(a) \frac{\partial g}{\partial x_i}(\bar{a}). \quad (*)$$

Andererseits gilt wegen  $(h \circ \varphi)(\bar{x}) = 0 \forall \bar{x} \in V$ , dass

$$0 = \frac{\partial(h \circ \varphi)}{\partial x_i}(\bar{a}) = \partial_i h(a) + \partial_n h(a) \frac{\partial g}{\partial x_i}(\bar{a}),$$

also wegen  $\frac{\partial h}{\partial x_n}(a) \neq 0$

$$\frac{\partial g}{\partial x_i}(\bar{a}) = - \left( \frac{\partial h}{\partial x_n}(a) \right)^{-1} \frac{\partial h}{\partial x_i}(a). \quad (**)$$

Setzen wir nun

$$\lambda := \frac{\frac{\partial f}{\partial x_n}(a)}{\frac{\partial h}{\partial x_n}(a)} \in \mathbb{R},$$

so folgt zunächst für  $i = n$

$$\frac{\partial f}{\partial x_n}(a) = \lambda \frac{\partial h}{\partial x_n}(a)$$

aber dann auch für  $i = 1, \dots, n-1$ , wenn man  $(**)$  in  $(*)$  einsetzt, dass

$$\frac{\partial f}{\partial x_i}(a) = \lambda \frac{\partial h}{\partial x_i}(a),$$

also  $\text{grad} f(a) = \lambda \text{grad} h(a)$ . □

**4.15 Beispiel.** Sei  $f : K \rightarrow \mathbb{R}$ ,  $(x, y) \mapsto y^2 - x^2$ , auf der Kreisscheibe  $K := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$  definiert. Da  $f$  stetig ist und  $K$  kompakt, nimmt  $f$  auf  $K$  sein Supremum  $c := \sup f(x, y)$  an. Wir wollen  $c$  und die Stellen  $(x, y) \in K$  wo  $f$  den Wert  $c$  annimmt, berechnen. Dazu suchen wir zunächst lokale Extrema im Inneren und auf dem Rand.

(i) **Innere Punkte:** Da  $\text{grad} f(x, y) = (-2x, 2y) \neq (0, 0)$  für  $(x, y) \neq (0, 0)$  und

$$\text{Hess} f(0, 0) = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$$

indefinit ist, hat  $f$  in  $\overset{\circ}{K} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$  kein lokales Extremum. Die Extrema von  $f$  müssen also auf dem Rand

$$M = \{(x, y) \mid x^2 + y^2 = 1\}$$

liegen.

(ii) **Randpunkte:** Betrachte die Nebenbedingung  $h = 0$  für

$$h : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto x^2 + y^2 - 1.$$

Ein Maximum von  $f$  auf einem Randpunkt  $a \in M$  ist sicherlich auch ein Maximum von  $f$  unter der Nebenbedingung  $h = 0$ . Es muss also gelten, dass

$$\text{grad} f(a) = \lambda \text{grad} h(a)$$

für ein  $\lambda \in \mathbb{R}$ . Beachte, dass  $\text{grad} h(x, y) = (2x, 2y) \neq 0$  für  $(x, y) \in M$ . Wir erhalten somit die Gleichungen

$$\begin{aligned} \text{(I.)} \quad & -2x &= & \lambda(2x) \\ \text{(II.)} \quad & 2y &= & \lambda(2y) \\ \text{(III.)} \quad & x^2 + y^2 &= & 1 \end{aligned}$$

Ist nun  $x \neq 0$ , so folgt aus I., dass  $\lambda = -1$  und dann aus II. und III., dass  $y = 0$  und  $x = \pm 1$ . Ist dagegen  $y \neq 0$ , so folgt  $x = 0$  und  $y = \pm 1$ . Man erhält also die vier Kandidaten

$$a \in \{(1, 0), (0, 1), (-1, 0), (0, -1)\}.$$

Da  $f(\pm 1, 0) = -1$  und  $f(0, \pm 1) = +1$  ist, gilt  $c = +1$  und das Supremum wird genau in den Punkten  $(0, 1)$  und  $(0, -1)$  angenommen.

**4.16 Bemerkung.** Man leitet die Gleichungen I. - III. aus Beispiel 4.15 oft folgendermaßen ab: statt  $\nabla f = 0$  fordert man bei NB  $h = 0$ , dass für ein  $\lambda \in \mathbb{R}$

$$\nabla(f + \lambda h) = 0 \quad (= \text{I.} + \text{II.})$$

und eben

$$h = 0 \quad (= \text{III.})$$

gelten. Man nennt  $\lambda$  auch den Lagrangeschen Multiplikator.

**4.17 Beispiel.** Wir geben nun einen alternativen Beweis für die Tatsache, dass jede symmetrische Matrix mindestens einen reellen Eigenwert hat. Sei dazu  $A \in M(n \times n, \mathbb{R})$  symmetrisch und

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto \langle x, Ax \rangle, \quad \text{sowie} \quad h : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto \langle x, x \rangle - 1.$$

Die Fläche  $\{h = 0\}$  ist die Einheitssphäre  $S^{n-1}$  im  $\mathbb{R}^n$  und ist kompakt. Damit nimmt die stetige Funktion  $f$  auf  $S^{n-1}$  ihr Maximum und ihr Minimum an. Sei  $x_0$  ein Punkt an dem  $f$  auf  $S^{n-1}$  maximal wird. Dann gilt nach Satz 4.13, dass

$$\nabla f(x_0) = 2Ax_0 = \lambda \nabla h(x_0) = \lambda 2x_0$$

für ein  $\lambda \in \mathbb{R}$ . Also ist  $\lambda$  Eigenwert zum Eigenvektor  $x_0$ .

**4.18 Bemerkung. Mehrere Nebenbedingungen**

Liegen mehrere Nebenbedingungen  $h_1, \dots, h_k$  vor, sucht man also ein Extremum von  $f$  auf der  $(n - k)$ -dimensionalen "Fläche"

$$\{h_1 = 0\} \cap \dots \cap \{h_k = 0\},$$

so ergibt sich die notwendige Bedingung, dass Richtungsableitungen  $\langle v, \nabla f(a) \rangle$  von  $f$  die tangential an alle Hyperflächen  $M_j := \{h_j = 0\}$  liegen, verschwinden. Also

$$\langle v, \nabla f(a) \rangle = 0 \quad \text{falls} \quad \langle v, \nabla h_j(a) \rangle = 0 \quad \forall j = 1, \dots, k.$$

Damit ist eine notwendige Bedingung für das Vorliegen eines Extremums von  $f|_{\cap_{j=1}^k M_j}$ , dass

$$\nabla f(a) \in \text{span}\{\nabla h_1(a), \dots, \nabla h_k(a)\}. \quad (*)$$

Fasst man  $(h_1, \dots, h_k)$  als Vektorwertige Funktion  $h : \mathbb{R}^n \rightarrow \mathbb{R}^k$  auf, so kann man (\*) wie gehabt aus

$$\nabla(f + \lambda \cdot h) = 0$$

herleiten, wobei der Lagrangemultiplikator  $\lambda$  nun ein Vektor im  $\mathbb{R}^k$  ist.

# 5 Integralrechnung

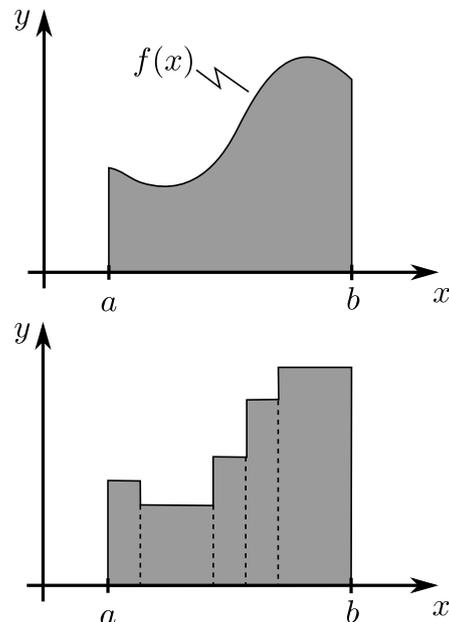
## 5.1 Motivation. Das Flächenproblem

Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine Funktion. Kann man dem Gebiet

$$\{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b \text{ und } 0 \leq y \leq f(x)\}$$

unter dem Graphen von  $f$  einen Flächeninhalt zuordnen?

**Idee:** Approximiere allgemeine Funktionen durch sogenannte Riemannsummen wobei die Zerlegung immer feiner wird. Wir werden sehen, dass für Funktionen, die nett genug sind, diese Summen konvergieren. Den Limes bezeichnen wir dann als  $\int_a^b f(x)dx$ .



A-priori unterscheiden wir über zwei Verschiedene Integraltypen, dem **bestimmten Integral**, das welches zum Berechnen von Flächen und Volumina verwendet wird. Und das **unbestimmte Integral**. Dabei handelt es sich um die algebraische Aufgabe eine Funktion  $F(x)$  zu finden, dessen Ableitung  $f$  ergibt.

## 5.2 Definition. Stammfunktion

Eine Funktion  $F : [a, b] \rightarrow \mathbb{C}$  heißt **Stammfunktion** von  $f : [a, b] \rightarrow \mathbb{C}$ , wenn  $F$  differenzierbar ist und

$$F' = f$$

gilt. Wir schreiben dann auch  $F(x) = \int f(x)dx$ .

Der Inhalt des Hauptsatzes der Differential und Integralrechnung wird dann besagen, dass zwischen, den beiden Definitionen es tatsächlich einen Zusammenhang gibt, nämlich dass der Flächeninhalt in der Tat durch die Stammfunktion berechnet werden kann, via

$$\int_a^b f(x)dx = F(b) - F(a).$$

## 5.1 Unbestimmtes Integral

**5.3 Bemerkung.** Man sieht sofort, dass sich zwei verschiedene Stammfunktionen von  $f$  um eine Konstante unterscheiden. (*Integrationskonstante*). Im Folgenden wollen wir eine Gleichung der Gestalt  $\int f(x)dx = F(x)$  verstehen als  $F$  ist *eine* Stammfunktion von  $f$ . Nach Definition der Stammfunktion kann man solche Behauptungen stets durch Differentiation verifizieren, so etwa an folgenden Beispielen:

**5.4 Beispiele.** Die folgenden durch Differentiation zu beweisenden Aussagen gelten auf ganz  $\mathbb{R}$ , falls nicht andere Gültigkeitsintervalle angegeben sind:

(a) Für  $\alpha \in \mathbb{R} \setminus \{-1\}$

$$\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1}$$

auf  $\mathbb{R}$  für  $\alpha \in \mathbb{N}_0$ , auf  $\mathbb{R}_- \cup \mathbb{R}_+$  für  $\alpha \in \{-2, -3, -4, \dots\}$  und auf  $\mathbb{R}_+$  sonst.

(b)

$$\int c dx = cx, \quad \int \frac{1}{x} dx = \ln x, \quad \int e^x dx = e^x,$$

(c)

$$\int \cos x dx = \sin x, \quad \int \sin x dx = -\cos x, \quad \int \cosh x dx = \sinh x, \quad \int \sinh x dx = \cosh x,$$

(d)

$$\int \tanh x dx = \ln \cosh x, \quad \int \frac{1}{1+x^2} dx = \arctan x,$$

(e)

$$\int \frac{1}{1-x^2} dx = \frac{1}{2} \ln \left| \frac{1+x}{1-x} \right| \quad \text{auf } (-\infty, -1), (-1, 1), \text{ und } (1, \infty)$$

(f)

$$\int \frac{1}{\sqrt{1+x^2}} dx = \operatorname{arcsinh} x, \quad \int \frac{1}{\sqrt{1-x^2}} dx = \arcsin x \quad \text{auf } (-1, 1),$$

(g)

$$\int \frac{1}{\sqrt{x^2-1}} dx = \begin{cases} \operatorname{arcosh} x & \text{auf } (1, \infty) \\ -\operatorname{arcosh}(-x) & \text{auf } (-\infty, -1) \end{cases} .$$

(h)

$$\int \tan x dx = -\ln |\cos x| \quad \text{auf } \left] (2k-1)\frac{\pi}{2}, (2k+1)\frac{\pi}{2} \right[ \quad (k \in \mathbb{Z})$$

**5.5 Satz.** Sei  $I \subset \mathbb{R}$  ein Intervall. Dann gilt:

(a)  $f_1, f_2, \dots, f_n : I \rightarrow \mathbb{R}$  möge die Stammfunktion  $F_1, \dots, F_n$  besitzen, dann ist  $\alpha_1 F_1 + \dots + \alpha_n F_n$  eine Stammfunktion für  $\alpha f_1 + \dots + \alpha_n f_n$ .

(b)  $f, g : I \rightarrow \mathbb{R}$ ,  $f$  habe die Stammfunktion  $F$ ,  $g$  sei differenzierbar und  $Fg'$  besitze eine Stammfunktion. Dann hat  $fg$  die Stammfunktion

$$\int fg dx = Fg - \int Fg' dx$$

*Beweis.* (b)  $(FG - \int Fg' dx)' = (Fg)' - Fg' = fg.$  □

**5.6 Beispiele.** (a)

$$\int \ln(x) dx = \int 1 \cdot \ln(x) dx = x \ln(x) - \int x \cdot \frac{1}{x} dx = x \ln(x) - x.$$

(b)

$$\int x e^x dx = x e^x - \int e^x dx = x e^x - e^x$$

und analog

$$\int x^2 e^x dx = x^2 e^x - \int 2x e^x dx = (x^2 - 2x + 2) e^x.$$

Auf diese Weise kann für jedes Polynom  $p$  das Integral

$$\int p(x) e^x dx$$

berechnet werden.

(c) Manchmal reproduziert sich der Integrand nach mehrfacher partieller Integration,

$$\int f(x)g(x)dx = h(x) + c \int f(x)g(x)dx.$$

Falls  $c \neq 1$ , so ist man fertig, denn dann ist

$$\int fg = \frac{1}{1-c}h.$$

Beispielsweise ist

$$\begin{aligned} \int \cos^2 x \, dx &= \int \cos x \cdot \cos x \, dx = \sin x \cdot \cos x + \int \sin x \cdot \sin x \, dx \\ &= \sin x \cdot \cos x + \int (1 - \cos^2 x) \, dx \\ &= \sin x \cdot \cos x + x - \int \cos^2 x \, dx. \end{aligned}$$

Also gilt

$$\int^x \cos^2 x \, dx = \frac{1}{2}(\sin x \cos x + x).$$

### 5.7 Satz. Substitutionsregel

Es sei  $f : [a, b] \rightarrow \mathbb{C}$  stetig und  $g : [c, d] \rightarrow [a, b]$  stetig differenzierbar, so dass  $g' \neq 0$  auf  $[c, d]$ . Sei  $\Phi(t)$  die Stammfunktion von  $f \circ g(t)g'(t)$ . Dann ist  $\phi \circ g^{-1}(x)$  die Stammfunktion von  $f(x)$ .

*Beweis.* Sei  $\phi(t)$  die Stammfunktion von  $f \circ g(t)g'(t)$ . Dann gilt nach der Kettenregel

$$\phi(g^{-1}(x))' = \phi'(g^{-1}(x))g^{-1}(x)' = f(x)g'(g^{-1}(x))g^{-1}(x)' = f(x)g(g^{-1}(x))' = f(x).$$

□

**5.8 Bemerkung.** Am einfachsten merkt man sich die Substitutionsregel folgendermassen.

$$\int f(x)dx = \int f(g(t))g'(t)dt \Big|_{t=g^{-1}(x)}.$$

Setzt man  $x = g(t)$ , dann bekommt man *formal* die Beziehung  $dx = g'(t)dt$  zwischen den Differenzialen, also *formal*

$$\int f(x)dx = \int f(g(t))g'(t)dt.$$

**5.9 Beispiel.** Ein häufiger Spezialfall ist:

$$\int \frac{g'(t)}{g(t)} dt = \int \frac{1}{x} dx = \ln |x| = \ln |g(t)|.$$

Oder

$$\int f(at + b)dt = \frac{1}{a} \int f(at + b)g'(t)dt = \frac{1}{a} \int f(x)dx = \frac{1}{a}F(at + b),$$

wobei  $F(x)$  die Stammfunktion von  $f(x)$  ist.

### 5.10 Bemerkung. Merkhilfe zur Substitutionsregel

Eine gute Merkhilfe ist auch folgendes,

$$\int f(g(x))g'(x)dx = \int_c^d f(g(x))d(g(x)) = F(g(x))$$

wobei  $F(x)$  eine Stammfunktion von  $f(x)$  ist.

Formal hilft sie einem aber schon jetzt bei der Anwendung der Substitutionsregel. Beispielsweise ist

$$\int \frac{f(\ln x)}{x} dx = \int f(\ln x) d(\ln x) = F(\ln x)$$

oder

$$\int_f (\sin x) \cos x dx = \int f(\sin x) d(\sin x) = F(\sin x)$$

oder

$$\int \frac{x^3}{(x^4 + 1)^2} dx = \int \frac{1}{(x^4 + 1)^2} \frac{d(x^4)}{4} = \frac{1}{4} \int \frac{1}{(y + 1)^2} dy \Big|_{y=x^4} = -\frac{1}{4} \frac{1}{x^4 + 1}$$

Sind  $p, q$  reelle Polynome mit  $p(x) \neq 0$  in  $[a, b]$ , so lässt sich die Stammfunktion

$$\int \frac{q(x)}{p(x)} dx$$

explizit angeben. Das Verfahren beruht auf der sogenannten Partialbruchzerlegung, d.h. der Zerlegung der rationalen Funktion in eine Summe einfacher Bausteine, die explizit integriert werden können.

**5.11 Satz. Komplexe Partialbruchzerlegung**

Es seien  $q, p$  Polynome und es habe  $p$  den Grad  $n$  und die Nullstellen  $\{z_1, \dots, z_k\}$  mit Vielfachheiten  $\{l_1, \dots, l_k\}$ , also  $p(z) = a_n \prod_{j=1}^k (z - z_j)^{l_j}$  (Fundamentalsatz der Algebra).

Dann gibt es eindeutig bestimmte Zahlen  $A_{ij} \in \mathbb{C}$  und ein Polynom  $h(z)$  mit

$$\begin{aligned} \frac{q(z)}{p(z)} &= h(z) + \sum_{j=1}^k \sum_{m=1}^{l_j} \frac{A_{jm}}{(z - z_j)^m} \\ &= h(z) + \frac{A_{11}}{(z - z_1)} + \frac{A_{12}}{(z - z_1)^2} + \dots + \frac{A_{1l_1}}{(z - z_1)^{l_1}} \\ &\quad + \dots \\ &\quad + \frac{A_{k1}}{(z - z_k)} + \dots + \frac{A_{kl_k}}{(z - z_k)^{l_k}}. \end{aligned}$$

*Beweis.* Das Polynom  $h$  ergibt sich eindeutig aus Polynomdivision mit Rest,  $q = h \cdot p + r$ , also

$$\frac{q(z)}{p(z)} = h(z) + \frac{r(z)}{p(z)} \quad \text{mit Grad } r < \text{Grad } p.$$

Es reicht also  $\frac{q}{p}$  mit Grad  $q < \text{Grad } p$  zu betrachten. Induktion nach  $n = \text{Grad } p$  liefert für  $n = 1$

$$\frac{q(z)}{p(z)} = \frac{c}{z - z_1}.$$

$n - 1 \Rightarrow n$ : Sei  $z_0$  Nullstelle von  $p$  der Ordnung  $l \geq 1$ , also

$$p(z) = (z - z_0)^l \tilde{p}(z)$$

mit Grad  $\tilde{p} \leq n - 1$ ,  $\tilde{p}(z_0) \neq 0$ . Daraus folgt für alle  $z$  mit  $p(z) \neq 0$ , dass

$$\frac{q(z)}{\tilde{p}(z)} - \frac{q(z_0)}{\tilde{p}(z_0)} = \frac{1}{\tilde{p}(z)} \frac{q(z)\tilde{p}(z_0) - q(z_0)\tilde{p}(z)}{\tilde{p}(z_0)} = \frac{(z - z_0)\tilde{q}(z)}{\tilde{p}(z)}$$

mit einem Polynom  $\tilde{q}$  vom Grad  $\tilde{q} \leq n - 2$ . Also ist

$$\frac{q(z)}{p(z)} = \frac{q(z)}{(z - z_0)^l \tilde{p}(z)} = \frac{q(z_0)}{\tilde{p}(z_0)} \frac{1}{(z - z_0)^l} + \frac{\tilde{q}(z)}{\tilde{p}(z)} \frac{1}{(z - z_0)^{l-1}}$$

Auf den letzten Term kann man die Induktionsvoraussetzung anwenden, da der Grad des Nenners  $n - 1$  ist. Die Eindeutigkeit folgt, wenn wir

$$\sum_{j=1}^k \sum_{m=1}^{l_j} \frac{A_{jm}}{(z - z_j)^m} = 0 \quad \Rightarrow \quad A_{jm} = 0 \quad \text{für alle } j, m$$

zeigen können. Das sieht man aber leicht: Multiplikation der linken Seite mit  $(z - z_j)^{l_j}$  und Auswertung bei  $z = z_j$  liefert  $A_{jl_j} = 0$ . Multiplikation der linken Seite mit  $(z - z_j)^{l_j - 1}$  und Auswertung bei  $z = z_j$  liefert  $A_{j,l_j - 1} = 0$  usw.  $\square$

### 5.12 Satz. Reelle Partialbruchzerlegung

Seien  $q, p$  reelle Polynome,  $p$  vom Grad  $n$  und

$$p(x) = a_n \prod_{j=1}^k (x - x_j)^{l_j} \prod_{j=1}^r (x^2 + 2b_j x + c_j)^{m_j}$$

mit  $b_j^2 < c_j$ . Dann gibt es eindeutig bestimmte Zahlen  $A_{jl}, B_{jm}, C_{jm}$  so, dass

$$\frac{q(x)}{p(x)} = h(x) + \sum_{j=1}^k \sum_{l=1}^{l_j} \frac{A_{jl}}{(x - x_j)^l} + \sum_{j=1}^r \sum_{m=1}^{m_j} \frac{B_{jm}x + C_{jm}}{(x^2 + 2b_j x + c_j)^m}.$$

*Beweis.* Man fasse die Terme mit komplex konjugierten Nullstellen aus Satz 5.11 zusammen.  $\square$

### 5.13 Beispiel. Betrachte

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

wobei die Nullstellen  $0, 1, -1$  des Nenners alle einfach sind. Polynomdivision mit Rest liefert zunächst

$$\left( \begin{array}{r} x^4 \\ -x^4 + x^2 \\ \hline x^2 \end{array} + 2 \right) : (x^3 - x) = x + \frac{x^2 + 2}{x^3 - x}$$

Es ist also

$$\frac{x^2 + 2}{x^3 - x} = \frac{A_1}{x} + \frac{A_2}{x - 1} + \frac{A_3}{x + 1}$$

und Multiplikation mit  $x^3 - x$  ergibt

$$x^2 + 2 = (x^2 - 1)A_1 + x(x + 1)A_2 + x(x - 1)A_3.$$

Einsetzen von

$$\begin{array}{lll} x = 1 & \text{liefert} & 3 = 2A_2 \Rightarrow A_2 = \frac{3}{2} \\ x = 0 & \text{liefert} & 2 = -A_1 \Rightarrow A_1 = -2 \\ x = -1 & \text{liefert} & 3 = 2A_3 \Rightarrow A_3 = \frac{3}{2}. \end{array}$$

Insgesamt ist also

$$\frac{x^4 + 2}{x^3 - x} = x - \frac{2}{x} + \frac{3}{2} \frac{1}{x - 1} + \frac{3}{2} \frac{1}{x + 1}.$$

## 5.2 Riemann-Integral

Das Riemann-Integral ist definiert durch

$$\int_a^b f(x)dx \approx \sum_{i=1}^n f(\xi_i) \underbrace{(x_i - x_{i-1})}_{=|I_i|}$$

$$= S(f, z_n, \sigma_n)$$

mit  $\sigma_n = \{\xi_1, \dots, \xi_n\}$  und  $z_n = \{x_1, \dots, x_n\}$ . Für  $n \rightarrow \infty$  gilt Gleichheit:

$$\int_a^b f(x)dx = \lim_{n \rightarrow \infty} S(f, z_n, \sigma_n)$$

Vergleiche auch Abbildung 5.1. Diese Definition des Integrals bringt einige Nachteile mit sich:

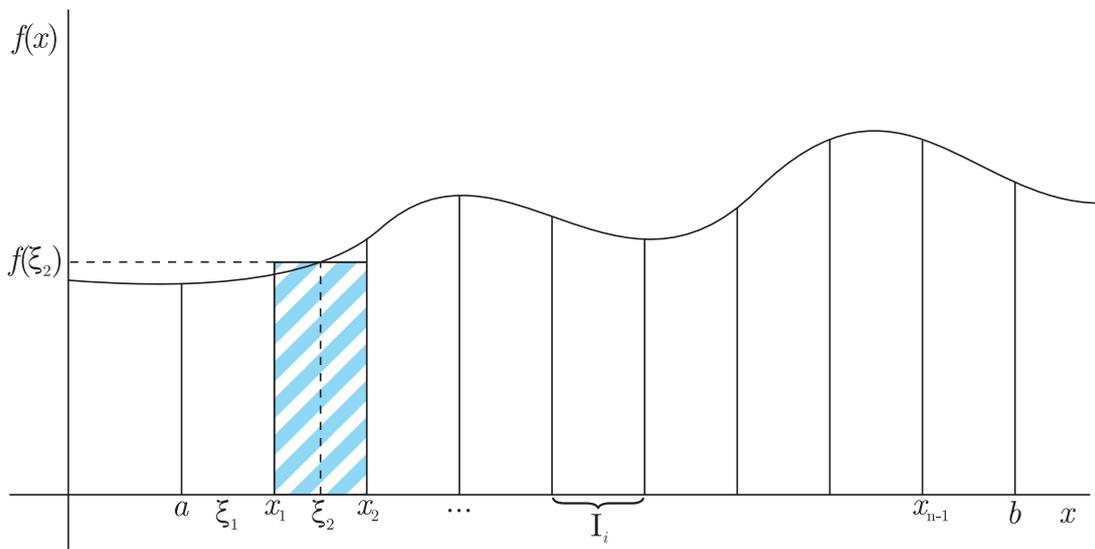


Abbildung 5.1: Riemann-Integral

- Es gibt eine große Klasse von einfachen Funktionen, die NICHT Riemann-integrierbar ist, zum Beispiel

$$f : [0, 1] \rightarrow \mathbb{R} \quad f(x) = \begin{cases} 0 & x \text{ irrational} \\ 1 & x \text{ rational} \end{cases}$$

- Sei  $f_n$  stetig und monoton. Dann existiert der Limes

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x)dx$$

allerdings ist

$$\lim_{n \rightarrow \infty} f_n(x) = f(x)$$

nicht Riemann-integrierbar.

- Die Menge

$$\left\{ f \mid \int_a^b |f|dx < \infty \right\}$$

ist nicht abgeschlossen wenn  $f_n \in L$ .

- $\|f\|_1 = \int_a^b |f|(x)dx$ ,  $\|\cdot\|$  keine Norm wenn  $\int f dx$  durch die Riemann-Summe definiert ist.

Alle diese Nachteile behebt das

## Lebesgue-Integral

Das Lebesgue-Integral ist definiert durch

$$\int_a^b f := \lim_{n \rightarrow \infty} \sum_k \frac{k}{2^n} \mu A_k^n$$

mit

$$A_k^n = f^{-1} \left( \left[ \frac{k}{2^n}, \frac{k+1}{2^n} \right] \right)$$

Das heißt das Bild von  $f$  wird zerlegt. Wenn  $f$  fast überall stetig ist, so sind das Riemann- und das Lebesgue-Integral identisch. Behandle der Einfachheit halber zuerst Riemann-Integrale, da diese einfacher explizit zu berechnen sind.

### 5.2.1 Definitionen.

Sei  $[a, b] \subset \mathbb{R}$  abgeschlossenes Intervall. Es heißt

$$z := \{x_0, \dots, x_n\} \quad \text{mit } a = x_0 \text{ und } b = x_n$$

Zerlegung von  $[a, b]$  mit Feinheit

$$|z| = \max\{|x_{i-1} - x_i| \mid i \in \{1, \dots, n\}\}$$

Das Intervall  $I$  setzt sich dann zusammen aus

$$I = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$$

Dann ist

$$z_1 \times z_2 \times \dots \times z_n$$

Zerlegung von  $I$  mit Feinheit

$$|z| = \max\{|z_i| \mid i \in \{1, \dots, n\}\}$$

$I_m$  sind Teilrechtecke,  $l \in \{1, \dots, m\}$ . Mit dem System von Zwischenpunkten

$$\sigma := \{\xi_1, \dots, \xi_m\}$$

ist die zugehörige Riemann-Summe

$$f : I \rightarrow \mathbb{R} \quad S(f, z, \sigma) := \sum_{i=1}^n f(\xi_i) |I_i|$$

Man sagt eine Funktion  $f$  heißt Riemann-integrierbar genau dann, wenn

$$\exists A \in \mathbb{R} : \forall \epsilon > 0 \exists \delta > 0 \forall z : (|z| < \delta \Rightarrow |S(f, z, \sigma) - A| < \epsilon \forall \sigma)$$

Es ist dann

$$\int_I f(x) dx = A$$

### 5.2.2 Bemerkung.

$f$  ist Riemann-integrierbar auf  $I$  ( $f \in \mathcal{R}(I)$ ) genau dann, wenn für alle Folgen mit Zerlegungen  $z_n$  und  $|z_n| \rightarrow 0$  gilt, dass die Riemann-Summe konvergiert für alle Systeme von Zwischenpunkten:

$$S(f, z_n, \sigma_n) \rightarrow \int_I f dx$$

*Proof.* Dazu gilt nachzuweisen, dass alle Folgen  $S(f, z_n, \sigma_n)$  konvergieren. Ihre Grenzwerte stimmen dann notwendigerweise überein. Die Zerlegung ist eindeutig: Nehme zwei Zerlegungen  $z_n, |z_n| \rightarrow 0$  und  $z'_n, |z'_n| \rightarrow 0$  dann kann man

$$z''_n = z_n \cup z'_n$$

bilden.  $z''_n$  ist so, dass sowohl  $S(f, z_n, \sigma_n)$  als auch  $S(f, z'_n, \sigma'_n)$  Teilfolgen von  $S(f, z''_n, \sigma''_n)$  sind. Also gilt:

$$\Rightarrow \lim_{n \rightarrow \infty} S(f, z_n, \sigma_n) = \lim_{n \rightarrow \infty} S(f, z'_n, \sigma'_n) = \lim_{n \rightarrow \infty} S(f, z''_n, \sigma''_n)$$

□

### 5.2.3 Satz. Cauchysches Integritätskriterium

Sei  $I \subset \mathbb{R}^n$  abgeschlossenes Intervall und  $f : I \rightarrow \mathbb{R}$ . Es gilt dann:

$$f \in \mathcal{R}(I) \Leftrightarrow \forall \epsilon > 0 \exists \delta > 0 \forall z_1, z_2 \text{ Zerlegung von } I \text{ mit} \\ |z_1| < \delta \text{ und } |z_2| < \delta : |S(f, z_1, \sigma_1) - S(f, z_2, \sigma_2)| < \epsilon \forall \sigma_1, \sigma_2$$

### 5.2.4 Satz.

Sei  $I \subset \mathbb{R}^n$  abgeschlossenes Intervall und  $f : I \rightarrow \mathbb{R}$  stetig.  $\Rightarrow f$  Riemann-integrierbar.

*Proof.* Da  $f$  stetig auf  $I$  ist  $f$  gleichmäßig stetig. Das heißt

$$\forall \epsilon > 0 \exists \delta > 0 \forall J : |f(x) - f(y)| < \frac{\epsilon}{2|I|} \text{ für } x, y \in J \text{ mit } |J| < \delta$$

Wähle zwei Zerlegungen  $z_1$  und  $z_2$  von  $I$ . Es bezeichne  $z$  die Vereinigung der beiden Zerlegungen:  $z = z_1 \cup z_2$ .  $z$  hat dann die Zerlegung  $J_{ij}$  mit Zwischenpunkten  $\eta_{ij}$ . Es ist dann

$$\sum_{j=1}^{k_i} |J_{ij}| = |I_i|$$

Die Riemann-Summen von  $z$  und  $z_1$  sind dann

$$S(f, z_1, \sigma_1) = \sum_{i=1}^m f(\xi_i) |I_i| \\ S(f, z, \sigma) = \sum_{i=1}^m \sum_{j=1}^{k_i} f(\eta_{ij}) |J_{ij}|$$

Bildet man deren Differenz erhält man

$$|S(f, z_1, \sigma_1) - S(f, z, \sigma)| = \left| \sum_{i=1}^m f(\xi_i) |I_i| - \sum_{i=1}^m \sum_{j=1}^{k_i} f(\eta_{ij}) |J_{ij}| \right| \\ = \left| \sum_{i=1}^m \sum_{j=1}^{k_i} [f(\xi_i) |J_{ij}| - f(\eta_{ij}) |J_{ij}|] \right| \\ \leq \sum_{i=1}^m \sum_{j=1}^{k_i} |f(\xi_i) - f(\eta_{ij})| |J_{ij}| \\ < \frac{\epsilon}{2|I|} \underbrace{\sum_{i,j} |J_{ij}|}_{=|I|} = \frac{\epsilon}{2}$$

Analog bildet man die Differenz von  $S(f, z, \sigma)$  und  $S(f, z_2, \sigma_2)$ . Es gilt dann

$$|S(f, z, \sigma) - S(f, z_2, \sigma_2)| < \frac{\epsilon}{2}$$

Also gilt

$$\begin{aligned} &\Rightarrow |S(f, z_1, \sigma_1) - S(f, z_2, \sigma_2)| < \epsilon \\ &\Rightarrow f \text{ Riemann-integrierbar} \end{aligned}$$

□

### 5.2.5 Bemerkung.

$A \subset \mathbb{R}^n$  heißt Nullmenge falls gilt:

$$\forall \epsilon > 0 \exists I_1, I_2, \dots, I_n, \dots \subset \mathbb{R}^n \text{ mit } A \subseteq \bigcup_i I_i \text{ und } \sum_{i=1}^{\infty} |I_i| < \epsilon$$

Beispiel:  $\mathbb{Q}^n$  ist Nullmenge.

### 5.2.6 Satz. Lebesgue

Es sei  $I \subset \mathbb{R}^n$  abgeschlossen und  $f : I \rightarrow \mathbb{R}$ . Dann gilt

$$f \text{ ist Riemann-integrierbar} \Leftrightarrow f \text{ beschränkt und } \{x \in I \mid f \text{ unstetig in } x\} \text{ ist Nullmenge}$$

*Proof.* ohne

□

### 5.2.7 Satz.

Es sei  $f : [a, b] \rightarrow \mathbb{R}$  monoton und beschränkt.  $\Rightarrow f$  Riemann-integrierbar

*Proof.* Übung

□

### 5.2.8 Satz.

Es seien  $I \subset \mathbb{R}^n$ ,  $f, g : I \rightarrow \mathbb{R}$ ,  $f, g \in \mathcal{R}(I)$ . Dann gilt

(a)  $f + g \in \mathcal{R}(I)$ , das heißt:

$$\int_I (f(x) + g(x)) dx = \int_I f(x) dx + \int_I g(x) dx$$

$cf \in \mathcal{R}(I)$ ,  $c \in \mathbb{R}$ , das heißt:

$$\int_I cf(x) dx = c \int_I f(x) dx$$

(b) Wenn  $f(x) \geq g(x) \quad \forall x \in I$

$$\Rightarrow \int_I f(x) dx \geq \int_I g(x) dx$$

(c)  $|f| \in \mathcal{R}(I)$  und

$$\int_I f(x) dx \leq \int_I |f(x)| dx \leq \sup_{x \in I} |f(x)| |I|$$

(d)  $f^2, g^2, fg \in \mathcal{R}(I)$  und

$$\int_I fg(x) dx \leq \sqrt{\int_I f^2(x) dx} \sqrt{\int_I g^2(x) dx} \quad (\text{Cauchy-Schwarz-Ungleichung für Integrale})$$

*Proof.* Übung

□

**5.2.9 Definition.**

Sei  $f : A \rightarrow \mathbb{R}, A \subset I \subset \mathbb{R}^n$ .

$$f_I(x) := \begin{cases} f(x) & x \in A \\ 0 & x \notin A \end{cases}$$

Es gilt dann

$$f \text{ ist Riemann-integrierbar} \Leftrightarrow f_I \in \mathcal{R}(I)$$

Man schreibt

$$\int_A f(x)dx := \int_I f_I(x)dx$$

Wir wollen auch Volumina durch Integrale berechnen. Dies funktioniert ja einfach dadurch, dass man die Funktion konstante  $f = 1$  über den gewünschten Bereich integriert. Sollte dies durch Riemann-integral wohldefiniert sein, so nennen wir eine solche Menge Jordan-messbar. Falls der Rand eines Gebietes nett, also glatt, genug ist, ist dies immer möglich.

**5.2.10 Definition. Jordan-messbar**

$A \subset \mathbb{R}^n$  heißt Jordan-messbar genau dann, wenn die charakteristische Funktion

$$i_A = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

Riemann-integrierbar ist. Der Jordan-Inhalt  $|A|$  ist dann

$$|A| := \int_A i_A(x)dx = \int_A 1dx$$

Für  $n = 2$  ist der Jordan-Inhalt der Flächeninhalt und für  $n = 3$  das Volumen.

**5.2.11 Bemerkung.**

Aus Satz 5.2.6 (Lebesgue) ergeben sich sofort folgende Aussagen:

- (a)  $A \subset \mathbb{R}^n$  beschränkt.  $A$  ist Jordan-messbar  $\Leftrightarrow \partial A$  (Rand von  $A$ ) ist Nullmenge
- (b)  $A \subset \mathbb{R}^n$  Jordan-messbar und  $f : A \rightarrow \mathbb{R}$ . Dann ist

$$f \in \mathcal{R}(A) \Leftrightarrow f \text{ beschränkt und } \{x \in A \mid f \text{ unstetig in } A\} \text{ ist Nullmenge}$$

- (c) Die Aussagen von Satz 5.2.8 gelten auch für  $A$  anstelle von  $I$ .
- (d)  $m := \inf\{f(x) \mid x \in A\}$  und  $M := \sup\{f(x) \mid x \in A\}$

$$m|A| \leq \int_A f(x)dx \leq M|A|$$

**5.3 Hauptsatz der Differential- und Integralrechnung**

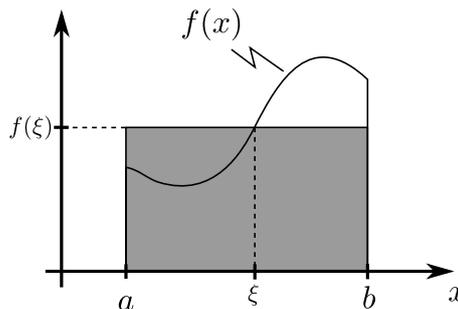
Bis hierhin sind bestimmte und unbestimmte Integrale zwei unterschiedliche Dinge. In diesem Kapitel werden wir sehen, dass die bestimmten integrale sich durch Randwerte von unbestimmten Integralen ausrechnen lassen. Analoges werden wir dann für höhere Dimensionen sehen. Auch dort wird ein Volumenintegral durch iteriertes Aufleiten ausgerechnet.

Wir beginnen zurest mit dem Mittelwertsatz der Integralrechnung.

**5.14 Satz. Der Mittelwertsatz der Integralrechnung**

- (a) Ist  $f : [a, b] \rightarrow \mathbb{R}$  stetig, so existiert ein  $\xi \in (a, b)$  mit

$$\int_a^b f(x) dx = f(\xi) \cdot (b - a).$$



- (b) Ist  $f : [a, b] \rightarrow \mathbb{R}$  stetig und  $\varphi : [a, b] \rightarrow \mathbb{R}$  eine positive Riemann-integrierbare Funktion, so existiert ein  $\xi \in (a, b)$  mit

$$\int_a^b f(x)\varphi(x) dx = f(\xi) \int_a^b \varphi(x) dx.$$

*Beweis.* (a) folgt aus (b) für  $\varphi(x) = 1$ .

- (b) Ist  $\min f = \max f$ , so ist  $f$  konstant und die Behauptung gilt für jedes  $\xi \in [a, b]$ . Ist  $\min f < \max f$ , so gilt

$$\min f \cdot \int_a^b \varphi(x) dx < \int_a^b f(x)\varphi(x) dx < \max f \cdot \int_a^b \varphi(x) dx$$

also

$$\min f < \frac{\int_a^b f(x)\varphi(x) dx}{\int_a^b \varphi(x) dx} < \max f.$$

Mit dem Zwischenwertsatz für die stetige Funktion  $f$  folgt die Behauptung. □

### 5.15 Satz. Zweiter Hauptsatz der Differential- und Integralrechnung

Für eine stetige Funktion  $f : [a, b] \rightarrow \mathbb{R}$  und jedes  $x \in [a, b]$  sei

$$F(x) := \int_a^x f(t) dt.$$

Dann ist  $F : [a, b] \rightarrow \mathbb{R}$  stetig differenzierbar und

$$\frac{d}{dx} F(x) = f(x) \quad \text{für alle } x \in [a, b].$$

*Beweis.* Seien  $x, x+h \in [a, b]$  mit  $h > 0$ . Dann gilt aufgrund der Stetigkeit von  $f$ , dass

$$\begin{aligned} \frac{F(x+h) - F(x)}{h} &= \frac{1}{h} \left( \int_a^{x+h} f(t) dt - \int_a^x f(t) dt \right) \\ &= \frac{1}{h} \int_x^{x+h} f(t) dt \\ &= \frac{1}{h} (x+h-x) f(\xi_h), \end{aligned}$$

mit  $\xi_h \in [x, x+h]$  nach dem Mittelwertsatz der Integralrechnung. Im letzten Schritt wurde die Stetigkeit verwendet. Damit folgt mit der Stetigkeit von  $f$  dass

$$F'(x) := \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} f(\xi_h) = f(x).$$

□

**5.16 Bemerkung.** Der zweite Hauptsatz besagt also, dass jedes stetige  $f$  eine Stammfunktion hat, nämlich

$$F(x) = \int_a^x f(t) dt.$$

Offenbar ist mit  $F(x)$  auch  $F(x) + c$ ,  $c \in \mathbb{R}$  eine Stammfunktion von  $f$ . Das sind dann aber auch schon alle Stammfunktionen. Denn seien  $F$  und  $\tilde{F}$  Stammfunktionen von  $f$ , so gilt  $(F - \tilde{F})' = f - f = 0$ , also  $F - \tilde{F} = \text{konstant}$ .

**5.17 Korollar. Erster Hauptsatz der Differential- und Integralrechnung**

Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig und  $F$  eine Stammfunktion von  $f$ . Dann gilt

$$\int_a^b f(x) dx = F(b) - F(a) =: [F(x)]_a^b.$$

*Beweis.* Für  $F(x) := \int_a^x f(t) dt$  gilt offenbar

$$\int_a^b f(x) dx = F(b) = F(b) - F(a).$$

Mit der vorangegangenen Bemerkung ist jede Stammfunktion  $\tilde{F}$  von  $f$  von der Form  $\tilde{F} = F + c$ , also

$$\tilde{F}(b) - \tilde{F}(a) = F(b) - F(a) = \int_a^b f(x) dx.$$

□

Damit haben wir für eindimensionale Funktionen den Zusammenhang zwischen Aufleitung und bestimmter Integration hergestellt.

**5.18 Bemerkung.** Äquivalent zu Korollar 5.17 formuliert man den ersten Hauptsatz auch oft so: Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig differenzierbar, so ist  $f(b) - f(a) = \int_a^b f'(x) dx$ .

Dies kann man gut dazu verwenden, um auf gegebenem Intervall, den Abstand  $\|f(b) - f(a)\|$  durch das Supremum der Ableitung abzuschätzen, analog zum Mittelwertsatz der Differentialrechnung.

Im Falle von Funktionen  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  lässt sich allerdings eine allgemeine Abschätzung auf konvexen Gebieten  $A$  herleiten, mittels  $S = \sup_{x \in A} \|Df(x)\|$ . Es gilt nämlich

$$\|f(x) - f(y)\| \leq S \|x - y\|,$$

was man leicht durch Anwendung des Mittelwertsatzes auf  $g(t) = f(x + t(y - x))$  sieht.

**5.19 Korollar. Integralrestglied der Taylorformel**

Es sei  $f \in C^{n+1}([a, b])$  für ein  $n \in \mathbb{N}_0$ . Dann gilt für  $x, x_0 \in [a, b]$

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \int_{x_0}^x \frac{f^{(n+1)}(t)}{n!} (x - t)^n dt.$$

*Beweis.* Wir setzen für  $t \in [a, b]$

$$F(t) := f(x) - \sum_{k=0}^n \frac{f^{(k)}(t)}{k!} (x - t)^k.$$

Es ist  $F(t)$  stetig differenzierbar und  $F(x) = 0$ , also gilt mit dem Hauptsatz, Satz 5.17,

$$F(x_0) = F(x) - \int_{x_0}^x F'(t) dt = - \int_{x_0}^x F'(t) dt.$$

Mit

$$F'(t) = - \sum_{k=0}^n \frac{f^{(k+1)}(t)}{k!} (x - t)^k + \sum_{k=1}^n \frac{f^{(k)}(t)}{(k-1)!} (x - t)^{k-1} = - \frac{f^{(n+1)}(t)}{n!} (x - t)^n$$

folgt die Behauptung. □

**5.20 Bemerkung.** Mit Hilfe des Mittelwertsatzes erhält man aus dem Integralrestglied das Lagrangsche Restglied der Taylorformel: Es gilt

$$\begin{aligned} R_n(x, x_0) &= \int_{x_0}^x \frac{f^{(n+1)}(t)}{n!} (x-t)^n dt = \frac{f^{(n+1)}(\xi)}{n!} \int_{x_0}^x (x-t)^n dt \\ &= \frac{f^{(n+1)}(\xi)}{n!} \frac{(x-x_0)^{n+1}}{(n+1)} = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-x_0)^{n+1}, \end{aligned}$$

da  $(x-t)^n$  im Integranden immer ein festes Vorzeichen hat.

### 5.21 Satz. Partielle Integration

Es seien  $f, g : [a, b] \rightarrow \mathbb{R}$ . Es sei  $f$  stetig,  $g$  stetig differenzierbar und  $F$  eine Stammfunktion von  $f$ . Dann gilt

$$\int_a^b fg = [Fg]_a^b - \int_a^b Fg',$$

wobei  $[Fg]_a^b := (Fg)(b) - (Fg)(a)$ .

*Beweis.* Das ist einfach der erste Hauptsatz, denn

$$\int_a^b fg = \int_a^b [(Fg)' - Fg'] = [Fg]_a^b - \int_a^b Fg'.$$

□

### 5.22 Beispiele. (a)

$$\int_a^b xe^x dx = xe^x \Big|_a^b - \int_a^b e^x dx = [xe^x - e^x]_a^b$$

und analog

$$\int_a^b x^2 e^x dx = x^2 e^x \Big|_a^b - \int_a^b 2xe^x dx = [(x^2 - 2x + 2)e^x]_a^b.$$

Auf diese Weise kann für jedes Polynom  $p$  das Integral

$$\int_a^b p(x)e^x dx$$

berechnet werden.

(b) Manchmal reproduziert sich der Integrand nach mehrfacher partieller Integration,

$$\int_a^b f(x)g(x)dx = [h(x)]_a^b + c \int_a^b f(x)g(x)dx.$$

Falls  $c \neq 1$ , so ist man fertig, denn dann ist

$$\int_a^b fg = \frac{1}{1-c} [h]_a^b.$$

### 5.23 Satz. Substitutionsregel

Es sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig und  $g : [c, d] \rightarrow [a, b]$  stetig differenzierbar. Dann gilt

$$\int_{g(c)}^{g(d)} f = \int_a^b (f \circ g) g'.$$

Falls entweder  $g' > 0$ , oder  $g' < 0$ , dann kann man es auch folgendermassen ausdrücken

$$\int_{g([c,d])} f = \int_{[c,d]} (f \circ g) |g'|.$$

*Beweis.* Sei  $F$  Stammfunktion von  $f$ . Und wir nehmen an  $a = g(c)$  und  $b = g(d)$ . Dann folgt aus der Kettenregel und dem Hauptsatz

$$\begin{aligned} \int_{g(c)}^{g(d)} f(x) dx &= F(g(d)) - F(g(c)) = (F \circ g)(d) - (F \circ g)(c) \\ &= \int_c^d (F \circ g)'(y) dy = \int_c^d (F' \circ g)(y) g'(y) dy \\ &= \int_c^d (f \circ g)(y) g'(y) dy. \end{aligned}$$

□

Das ganze lässt sich auch in Form von Riemmansummen beschreiben, mittels

$$\begin{aligned} \int_a^b f(x) dx &\sim \sum_i f(x_i)(x_i - x_{i-1}) = \sum_i f(g(t_i))(g(t_i) - g(t_{i-1})) \\ &\sim \sum_j f(g(t_i))g'(t_i)(t_i - t_{i-1}) \sim \int_c^d f(g(t))g'(t) dt, \quad (5.1) \end{aligned}$$

wobei  $x_i = g(t_i)$  ist. Dabei sieht man, dass der Term  $g'(t)$  die Vergrößerung oder Verkleinerung der Teilintervalle kompensiert.

In höheren Dimensionen wird diese Kompensation von der Determinante von  $Dg$  beschrieben.

#### 5.24 Bemerkung. Merkhilfe zur Substitutionsregel

Eine gute Merkhilfe ist folgende Notation,

$$\int_c^d f(g(x)) g'(x) dx = \int_c^d f(g(x)) d(g(x)) = \int_{g(c)}^{g(d)} f(y) dy$$

da  $\frac{dg}{dx} = g'(x)$ , also  $g'(x)dx = dg$ . Wenn wir Differentialformen einführen, wird diese Schreibweise einen präzisen Sinn bekommen.

Formal hilft sie einem aber schon jetzt bei der Anwendung der Substitutionsregel. Beispielsweise ist

$$\int_c^d \frac{f(\ln x)}{x} dx = \int_c^d f(\ln x) d(\ln x) = \int_{\ln c}^{\ln d} f(y) dy,$$

oder

$$\int_c^d f(\sin x) \cos x dx = \int_c^d f(\sin x) d(\sin x) = \int_{\sin c}^{\sin d} f(y) dy$$

oder

$$\int_a^b \frac{x^3}{(x^4 + 1)^2} dx = \int_a^b \frac{1}{(x^4 + 1)^2} \frac{d(x^4)}{4} = \frac{1}{4} \int_{a^4}^{b^4} \frac{1}{(y + 1)^2} dy = \frac{1}{4} \left( \frac{1}{a^4 + 1} - \frac{1}{b^4 + 1} \right).$$

**5.25 Beispiel.** Mit  $f(y) = \frac{1}{y}$  gilt, dass

$$\int_c^d \frac{g'(x)}{g(x)} dx = \int_c^d f(g(x)) g'(x) dx = \int_{g(c)}^{g(d)} f(y) dy = \ln \left( \frac{g(d)}{g(c)} \right)$$

Beispielsweise folgt so für  $\tan x = \frac{\sin x}{\cos x} = -\frac{\cos' x}{\cos x}$ , dass

$$\int_c^d \tan x dx = \ln \left( \frac{\cos c}{\cos d} \right).$$

### 5.26 Beispiel. Die Kreisfläche

Um die Fläche  $F$  des Kreises mit Radius  $r$  zu bestimmen, berechnen wir die Fläche des halben Kreises als Integral über die Funktion  $f(x) = \sqrt{r^2 - x^2}$  von  $-r$  bis  $r$ ,

$$\frac{F}{2} = \int_{-r}^r f(x) dx = \int_{-r}^r \sqrt{r^2 - x^2} dx.$$

Geometrisch liegt es nahe, das Integral durch den Winkel statt durch den Achsenabschnitt zu parametrisieren, also  $g : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [-r, r]$ ,  $x = g(\varphi) = r \sin \varphi$ , zu substituieren. Das liefert dann

$$\begin{aligned} \frac{F}{2} &= \int_{g(-\frac{\pi}{2})}^{g(\frac{\pi}{2})} f(x) dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(g(\varphi)) dg(\varphi) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{r^2 - g(\varphi)^2} g'(\varphi) d\varphi \\ &= r^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{1 - \sin^2 \varphi} \cos \varphi d\varphi = r^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 \varphi d\varphi = \frac{\pi}{2} r^2. \end{aligned}$$

## 5.4 Satz von Fubini, Substitutionsregel

Im folgenden leiten wir Rechenregeln für mehrdimensionale Integrale her. Solche werden dann iterativ mittels Aufleiten berechnet. Grundlage dazu ist der Satz von Fubini.

### 5.4.1 Satz. Fubini

Seien  $I_x \subset \mathbb{R}^n, I_y \subset \mathbb{R}^m$ . Das Intervall  $I \subset \mathbb{R}^{n+m}$  ist dann  $I = I_x \times I_y$ . Sei ferner  $f \in \mathcal{R}(I)$  und  $\forall y \in I_y$  existiere  $g(y) = \int_{I_x} f(x, y) dx$ . Dann ist  $g(y) \in \mathcal{R}(I)$  und

$$\int_I f(z) dz = \int_{I_y} \left( \int_{I_x} f(x, y) dx \right) dy$$

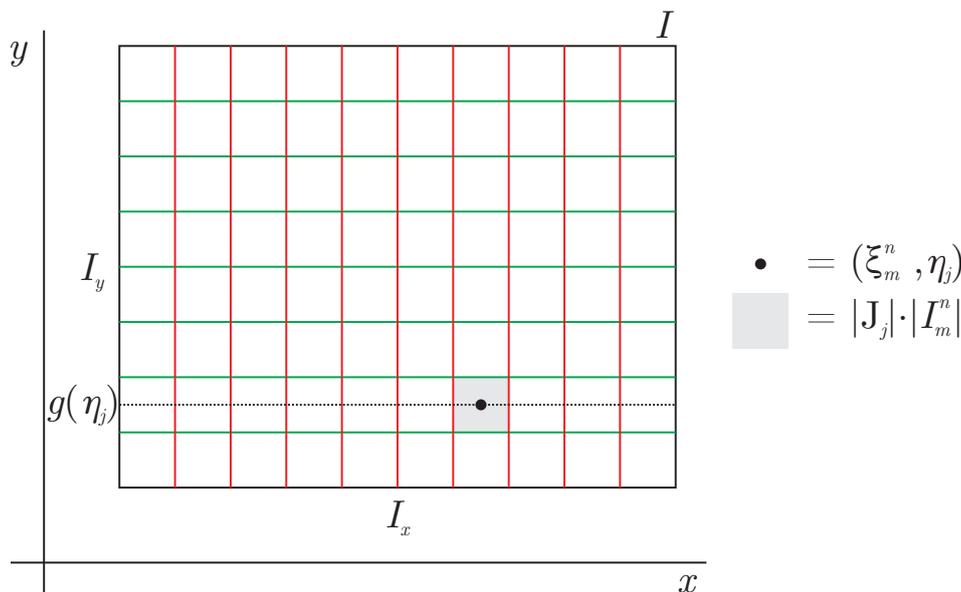


Abbildung 5.2: Zerlegung des Rechtecks  $I$

*Proof.* Da  $f$  Riemann-integrierbar ist, kann man schreiben

$$\forall \epsilon > 0 \exists \delta > 0 \forall z : |z| < \delta : \left| S(f, z, \sigma) - \int_I f(z) dz \right| < \epsilon \forall \sigma$$

## 5 Integralrechnung

Sei nun  $z_n^x = \{\xi_1^n, \dots, \xi_r^n\}$  eine Folge von Zerlegungen (vergleiche Abbildung 5.2) von  $I_x$  und  $z^y = \{\eta_1, \dots, \eta_s\}$  die Zerlegung von  $I_y$  mit  $|z_n^x|, |z^y| < \delta \forall n \in \mathbb{N}$ . Es ist  $\lim_{n \rightarrow \infty} |z_n^x| = 0$ . Setzt man dieses nun oben ein erhält man, dass gelten muss

$$\left| \sum_{m,j} f(\xi_m^n, \eta_j) |I_m^n| |J_j| - \int_I f(z) dz \right| < \epsilon$$

Es bleibt zu zeigen, dass

$$\left| \sum_m g(\eta_m) |J_m| - \int_I f(z) dz \right| < \epsilon$$

Verwende, dass

$$g(\eta_j) = \int_{I_x} f(x, \eta_j) dx = \lim_{n \rightarrow \infty} \sum_m f(\xi_m^n, \eta_j) |I_m^n|$$

und

$$\sum_j g(\eta_j) |J_j| = \lim_{n \rightarrow \infty} \sum_j \sum_m f(\xi_m^n, \eta_j) |I_m^n| |J_j|$$

Schreibe also um das obige zu zeigen

$$\begin{aligned} \left| \sum_j g(\eta_j) |J_j| - \int_I f(z) dz \right| &= \left| \lim_{n \rightarrow \infty} \sum_j \sum_m f(\xi_m^n, \eta_j) |I_m^n| |J_j| - \int_I f(z) dz \right| \\ &\leq \lim_{n \rightarrow \infty} \left| \sum_j \sum_m f(\xi_m^n, \eta_j) |I_m^n| |J_j| - \int_I f(z) dz \right| < \epsilon \\ &\Rightarrow g \in \mathcal{R}(I), \int_{I_y} g(y) dy = \int_I f(z) dz \end{aligned}$$

□

### 5.4.2 Korollar.

Sei  $I = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$  und  $f : I \rightarrow \mathbb{R}$  stetig. Dann gilt:

$$\int_I f(x) dx = \int_{a_1}^{b_1} \left( \int_{a_2}^{b_2} \dots \left( \int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \right) \dots dx_2 \right) dx_1$$

### 5.4.3 Satz.

Sei  $A \subset \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$  beschränkt und Jordan-messbar und  $f \in \mathcal{R}(A)$ . Es sei weiterhin

$$P_n(A) = \{x \mid \exists y \in \mathbb{R}^m : (x, y) \in A\}$$

und

$$A_x = \{y \in \mathbb{R}^m \mid (x, y) \in A\}$$

Ferne existiere

$$g(x) = \int_{A_x} f(x, y) dy$$

dann existiert

$$\int_{P_n(A)} g(x) dx$$

und es gilt

$$\int_A f(x, y) d(x, y) = \int_{P_n(A)} \left( \int_{A_x} f(x, y) dy \right) dx$$

*Proof.* folgt aus Fubini. □

#### 5.4.4 Korollar.

Sei  $A \subset \mathbb{R}^n$  beschränkt und Jordan-messbar sowie  $a < b \in \mathbb{R}$  so, dass

$$\forall x = (x_1, \dots, x_n) \in A : a \leq x_1 \leq b$$

und

$$\forall \xi \in [a, b] : Q(\xi) = A \cap \{(\xi, x_2, \dots, x_n)\}$$

in  $\mathbb{R}^{n-1}$  Jordan-messbar mit  $(n-1)$ -dimensionalem Inhalt  $g(\xi)$ . Dann ist

$$|A| = \int_a^b g(\xi) d\xi$$

#### 5.4.5 Beispiel. Kreisscheibe mit Radius 1

Im rechten oberen Quadranten ist die Kreisscheibe definiert durch die Funktion  $y = \sqrt{1-x^2}$ , im rechten unteren Quadranten durch  $y = -\sqrt{1-x^2}$ . Die Fläche ist dann

$$\int_A 1 dx = \int_{-1}^1 \left( \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1 dy \right) dx = 2 \int_{-1}^1 \sqrt{1-x^2} dx$$

Substituiere nun  $x = \sin t$ ,  $dx = \cos t dt$ . Es ergibt sich

$$2 \int_{-\pi/2}^{\pi/2} \sqrt{1-\sin^2 t} \cos t dt = 2 \int_{-\pi/2}^{\pi/2} \cos^2 t dt = \pi$$

### Integrieren über Normalbereiche

Betrachte die Mengen (Abbildung 5.3)

$$B = \{(x, y) \in \mathbb{R}^2, a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\}$$

$$C = \{(x, y) \in \mathbb{R}^2, c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\}$$

$$A = \{(x, y, z) \in \mathbb{R}^3, a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x), \psi_1(x, y) \leq z \leq \psi_2(x, y)\}$$

#### 5.4.6 Satz.

Sei  $f : B, C \rightarrow \mathbb{R}$  stetig. Es gilt für das Integral von  $f$ :

$$\int_B f(x, y) d(x, y) = \int_a^b \left( \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy \right) dx$$

und

$$\int_C f(x, y) d(x, y) = \int_c^d \left( \int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx \right) dy$$

Für  $A$  gilt

$$\int_A f(x, y, z) d(x, y, z) = \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} \int_{\psi_1(x, y)}^{\psi_2(x, y)} f(x, y, z) dz dy dx$$

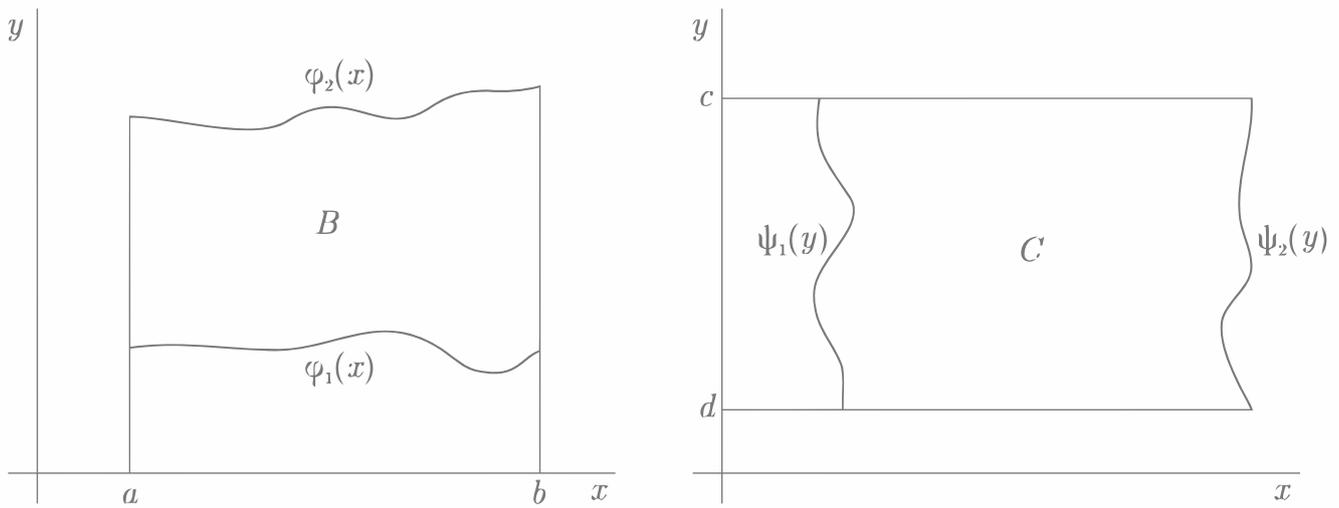


Abbildung 5.3: Normalbereiche

**5.4.7 Beispiel.**

Sei  $B$  ein Gebiet in  $\mathbb{R}^2$  das begrenzt ist durch  $y = x^2$ ,  $y = x^3$ , das heißt:

$$B = \{(x, y) \mid 0 \leq x \leq 1, x^3 \leq y \leq x^2\}$$

Sei außerdem  $f(x, y) = x$ . Gesucht ist jetzt  $\int_B f(z) dz$ . Es ist

$$\begin{aligned} \int_B f(z) dz &= \int_0^1 \int_{x^3}^{x^2} x dy dx \\ &= \int_0^1 x(x^2 - x^3) dx \\ &= \frac{1}{4} - \frac{1}{5} \end{aligned}$$

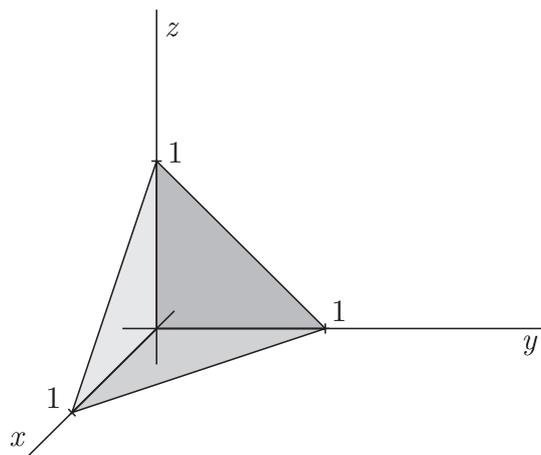


Abbildung 5.4: Die Flächen umschließen einen Tetraeder

**5.4.8 Beispiel.**

$$\int_1^4 \left( \int_{\sqrt{x}}^2 f(x, y) dy \right) dx = \int_1^2 \left( \int_1^{y^2} f(x, y) dx \right) dy$$

**5.4.9 Beispiel.**

Berechne das Integral

$$\int_0^1 \int_x^1 e^{y^2} dy dx.$$

Dies kann auf diese Art nicht explizit berechnet werden, allerdings durch ändern der Integrationsreihenfolge. Denn

$$\int_0^1 \int_x^1 e^{y^2} dy dx = \int_0^1 \int_0^y e^{y^2} dx dy = \int_0^1 y e^{y^2} dy = \frac{1}{2} \int_0^1 e^{y^2} d(y^2) = \frac{1}{2}(e - 1).$$

**5.4.10 Beispiel.**

Das Integral  $\int_D xy d(x, y)$  wobei  $D$  das Gebiet ist, das von der Geraden  $y = x - 1$  und der Parabel  $y^2 = 2x + 6$  eingeschlossen wird. Dieses kann als Typ  $C$ -Integral der Form

$$\int_{-2}^4 \left( \int_{\frac{1}{2}y^2 - 3}^{y+1} xy dx \right) dy,$$

berechnet werden als auch als Typ- $B$  Integral, allerdings muss man dort den Bereich in zwei Teile zerlegen,

$$\int_{-3}^{-1} \int_{-\sqrt{2x+6}}^{\sqrt{2x+6}} xy dy dx + \int_{-1}^5 \int_{x-1}^{\sqrt{2x+6}} xy dy dx.$$

In beiden Fällen ist das Resultat 36.

**5.4.11 Beispiel.**

Berechne das Volumen des Bereichs, der durch die Flächen  $x = 0$ ,  $y = 0$ ,  $z = 0$ ,  $x + y + z = 1$  begrenzt wird (Abbildung 5.4). Die Menge der Punkte innerhalb des eingeschlossenen Bereichs ist

$$A = \{(x, y, z), 0 \leq x \leq 1, 0 \leq y \leq 1 - x, 0 \leq z \leq 1 - x - y\}$$

Das Volumenintegral lautet dann

$$\begin{aligned} \int_0^1 \int_0^{1-x} \int_0^{1-x-y} dz dy dx &= \int_0^1 \int_0^{1-x} (1-x-y) dy dx \\ &= \int_0^1 \left[ (y - xy - \frac{y^2}{2}) \right]_{y=0}^{1-x} dx \\ &= \int_0^1 1 - x - x(1-x) - \frac{(1-x)^2}{2} dx \\ &= \frac{1}{2} \int_0^1 (x-1)^2 dx \\ &= \frac{1}{6} \end{aligned}$$

**5.4.12 Satz. Substitutionsregel**

Die Integration über krummlinige Gebiete lässt sich oft auf ein Integral über ein ( $n$ -dimensionales) Rechteck reduzieren. Die Information darüber, wie sehr jeder Punkt durch die Abbildung gestaucht oder gestreckt wird, ist natürlich wichtig für die Gewichtung.

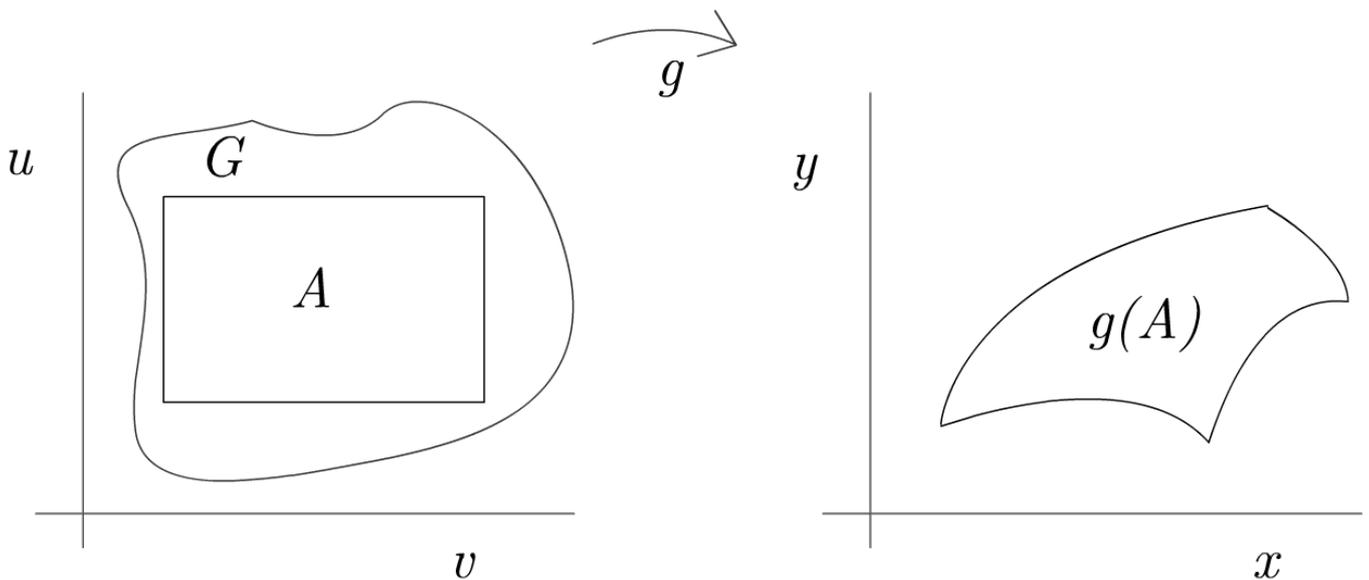


Abbildung 5.5: Darstellung der Substitution

Sei  $A \subseteq \mathbb{R}^n$  kompakt und Jordan-messbar.  $A \subseteq G, G$  offen, sodass  $g \in C^1(G, \mathbb{R}^n)$ , mit

$$\det Dg(z) > 0 \text{ oder} \\ \det Dg(z) < 0 \forall z \in G$$

und  $g$  injektiv. Sei  $f : g(A) \rightarrow \mathbb{R}$  stetig, dann gilt:

$$\int_{g(A)} f(x) dx = \int_A f(g(z)) |\det Dg(z)| dz$$

Dass die Determinante diese Eigenschaft besitzt sei hier für zwei Dimensionen verdeutlicht.

$$x = g_1(u, v) \\ y = g_2(u, v)$$

Ein Flächenstück in  $A$  sei ein Rechteck mit den Seitenlängen  $\Delta u \Delta v$ . Unter der Abbildung  $g$  verformt es sich. Unter der Annahme, dass es genügend klein ist, betrachten wir es als Parallelogramm. Die Richtung der aufspannenden Vektoren ist gegeben durch die Ableitungen von  $g$  längs der alten Koordinaten  $u$  und  $v$ :

$$\frac{\partial g}{\partial u}, \frac{\partial g}{\partial v} \in \mathbb{R}^2$$

Die Fläche des Parallelogramms ist gegeben durch

$$\left| \frac{\partial g}{\partial u} \times \frac{\partial g}{\partial v} \right| \Delta u \Delta v$$

Da wir das Kreuzprodukt nur im  $\mathbb{R}^3$  definiert haben, führen wir eine Transformation durch, welche jedem Vektor die z-Komponente 0 anfügt.

$$\frac{\partial g}{\partial u} = \begin{pmatrix} \frac{\partial g_1}{\partial u} \\ \frac{\partial g_2}{\partial u} \\ 0 \end{pmatrix}, \quad \frac{\partial g}{\partial v} = \begin{pmatrix} \frac{\partial g_1}{\partial v} \\ \frac{\partial g_2}{\partial v} \\ 0 \end{pmatrix}$$

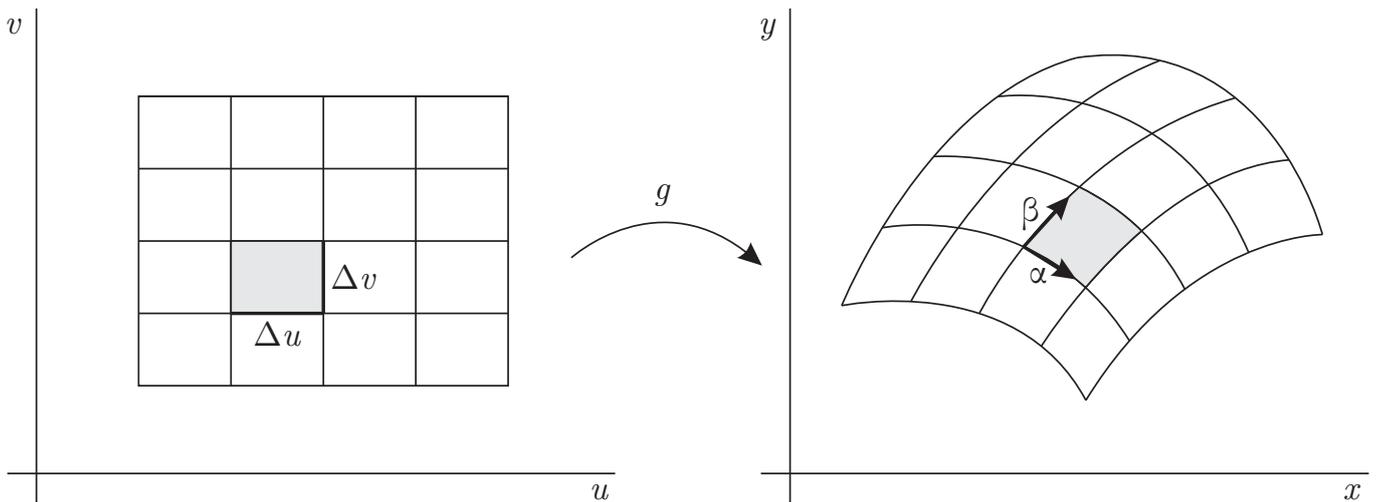


Abbildung 5.6: Transformation von Parallelogramm auf Rechteck. Im Bild ist  $\alpha = \frac{\partial g}{\partial u}$  und  $\beta = \frac{\partial g}{\partial v}$

So sieht man den Zusammenhang zwischen Kreuzprodukt der Richtungsableitungen und Determinante der Funktionalmatrix von  $g'$ :

$$\left| \frac{\partial g}{\partial u} \times \frac{\partial g}{\partial v} \right| = \left| \frac{\partial g_1}{\partial u} \frac{\partial g_2}{\partial v} - \frac{\partial g_2}{\partial u} \frac{\partial g_1}{\partial v} \right| = \left| \det \begin{pmatrix} \frac{\partial g_1}{\partial u} & \frac{\partial g_1}{\partial v} \\ \frac{\partial g_2}{\partial u} & \frac{\partial g_2}{\partial v} \end{pmatrix} \right| = |\det Dg(z)|$$

Die Substitutionsregel analog im n-dimensionalen Fall mit

$$z = (x_1, \dots, x_n), \quad g = \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix} \quad \text{und} \quad Dg(z) = \frac{\partial g}{\partial (x_1, \dots, x_n)} \quad (n \times n \text{ Matrix})$$

(a) POLARKOORDINATEN

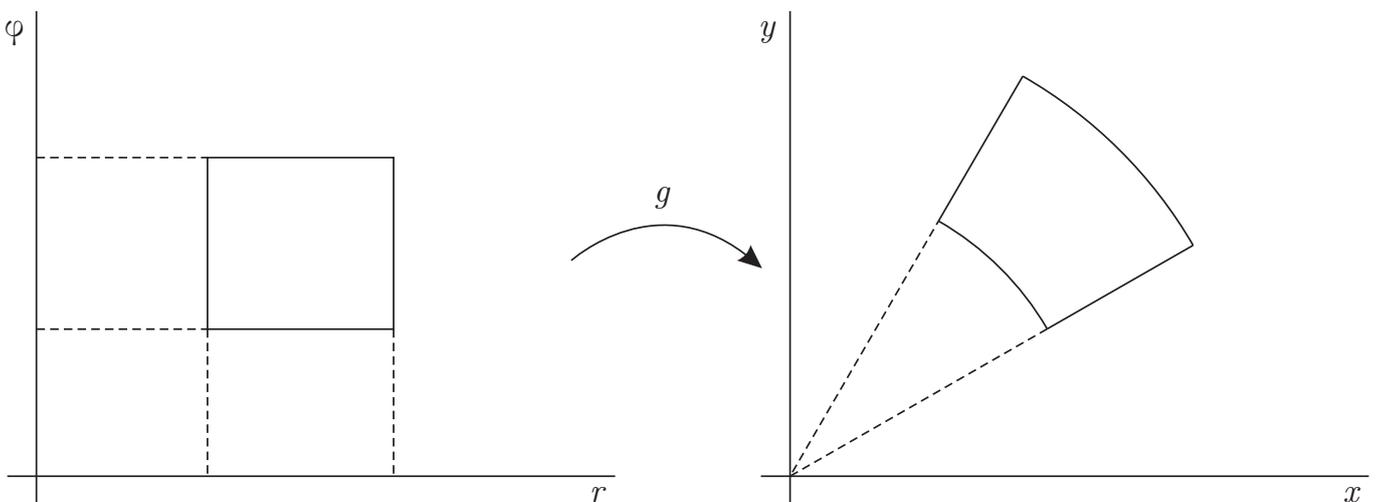


Abbildung 5.7: Transformation von kartesischen zu Polarkoordinaten

$$\begin{pmatrix} x \\ y \end{pmatrix} = g \begin{pmatrix} r \\ \varphi \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$$

$$\Rightarrow gD(z) = \begin{pmatrix} \frac{\partial g_1}{\partial r} & \frac{\partial g_1}{\partial \varphi} \\ \frac{\partial g_2}{\partial r} & \frac{\partial g_2}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

$$\Rightarrow |\det Dg(z)| = r$$

Für das Integral bedeutet dies:

$$\int_{g(A)} f(x, y) \, dx dy = \int_{\varphi_1}^{\varphi_2} \int_{r_1}^{r_2} f(r \cos \varphi, r \sin \varphi) r \, dr d\varphi$$

(b) ZYLINDERKOORDINATEN

Für Zylinderkoordinaten ist das Volumenelement das gleiche:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = g \begin{pmatrix} r \\ \varphi \\ z \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{pmatrix}$$

$$\Rightarrow Dg = \begin{pmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\Rightarrow |\det Dg| = r$$

mit dem Volumenelement  $dV = dx dy dz$  lautet das Integral:

$$\int_{g(A)} f(x, y, z) \, dV = \int_{z_1}^{z_2} \int_{\varphi_1}^{\varphi_2} \int_{r_1}^{r_2} f(r \cos \varphi, r \sin \varphi, z) r \, dr d\varphi dz$$

(c) KUGELKOORDINATEN Die Berechnung des Volumenelements für Kugelkoordinaten:

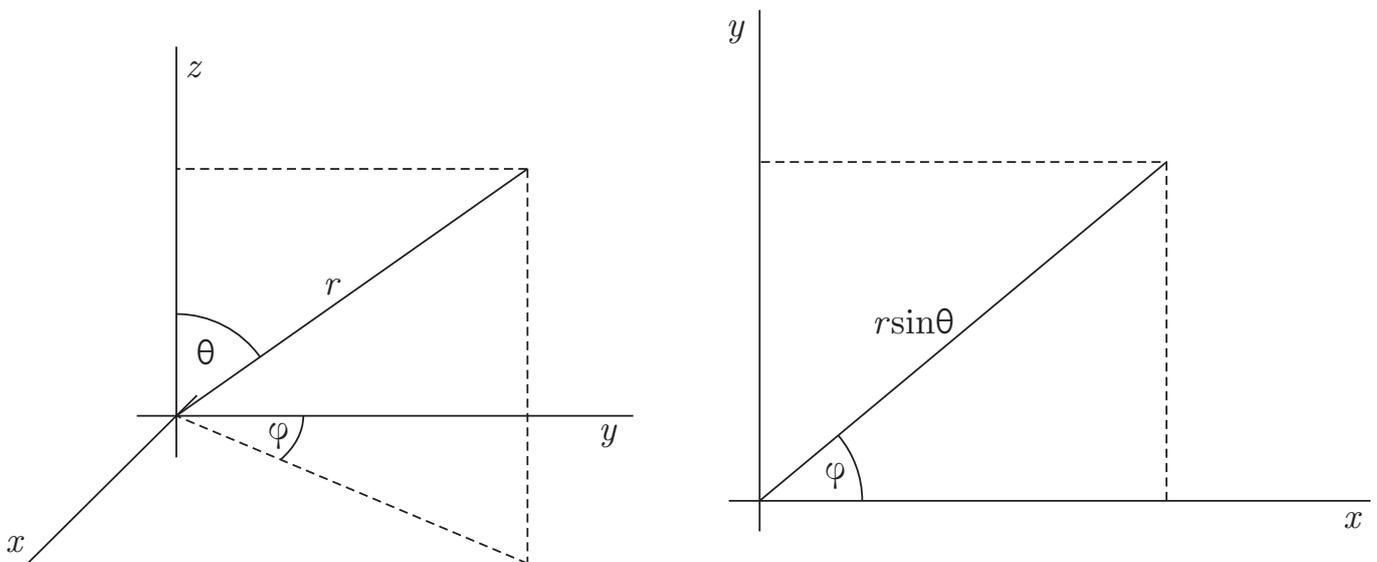


Abbildung 5.8: Die Transformation in Kugelkoordinaten

$$\begin{aligned}
\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= g \begin{pmatrix} r \\ \theta \\ \varphi \end{pmatrix} = \begin{pmatrix} r \sin \theta \cos \varphi \\ r \sin \theta \sin \varphi \\ r \cos \theta \end{pmatrix} \\
\Rightarrow Dg &= \begin{pmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix} \\
\Rightarrow \det Dg &= \cos \theta \det \begin{pmatrix} r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ r \cos \theta \sin \varphi & \sin \theta \cos \varphi \end{pmatrix} \\
&\quad - (-r \sin \theta) \det \begin{pmatrix} \sin \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & \sin \theta \cos \varphi \end{pmatrix} \\
&= (r^2 \cos^2 \theta \sin \theta \cos^2 \varphi + r^2 \cos^2 \theta \sin \theta \sin^2 \varphi) \\
&\quad + r \sin \theta (r \sin^2 \theta \cos^2 \varphi + r \sin^2 \theta \sin^2 \varphi) \\
&= r^2 \cos^2 \theta \sin \theta + r^2 \sin \theta \sin^2 \varphi \\
&= r^2 \sin \theta
\end{aligned}$$

#### 5.4.13 Beispiel.

Mittels Koordinatentransformation läßt sich leicht der Flächeninhalt des Ellipsoids, das von den Punkten  $(\frac{x}{a})^2 + (\frac{y}{b})^2 \leq r^2$  beschrieben wird. Durch die Abbildung  $x = au$ ,  $y = bv$  lässt sich das Ellipsoid auf einen Kreis in  $u, v$  transformieren. Da die Determinante von  $Dg(u, v)$  gleich  $ab$  ist, ist der Flächeninhalt somit  $r^2 \pi ab$ .

#### 5.4.14 Beispiel.

Es sei ein Zylinder gegeben, aus welchem am unteren Ende eine Halbkugel mit dem selben Radius herausgeschnitten ist. Mit anderen Worten, der Boden wird ersetzt durch die Oberfläche einer Halbkugel. Die Höhe sei 4, ( $0 \leq z \leq 4$ ), der Radius sei 1 ( $r \leq 1$ ), und der Boden  $z = \sqrt{1 - r^2}$ . Das Volumen sei  $V$ :

$$\begin{aligned}
V &= \int_0^{2\pi} \int_0^1 \int_{\sqrt{1-r^2}}^4 r \, dz \, dr \, d\varphi \\
&= \int_0^{2\pi} \int_0^1 (4 - \sqrt{1-r^2}) r \, dr \, d\varphi \\
&= 2\pi \left[ 2r^2 + \frac{1}{3}(1-r^2)^{\frac{3}{2}} \right]_0^1
\end{aligned}$$

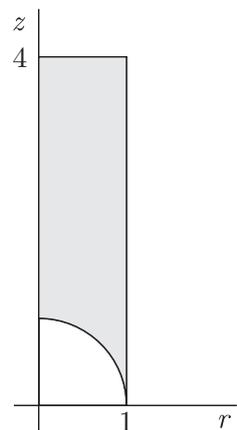


Abbildung 9

#### 5.4.15 Beispiel.

Gegeben ist eine Kugel mit Radius 1, in welche ein umgekehrter Kegel mit Öffnungswinkel  $\frac{\pi}{2}$  reingelegt ist. Gesucht ist das Volumen zwischen dem Kegel und der Kugel. Die Kugel ist in Zylinderkoordinaten gegeben durch  $z = \sqrt{1 - r^2}$ , der Kegel durch  $z = \sqrt{x^2 + y^2} = r$ . Das Volumenintegral lautet dann

$$V = \int_0^{2\pi} \int_0^{1/\sqrt{2}} \int_r^{\sqrt{1-r^2}} r dz dr d\varphi$$

Das selbe Volumen lässt sich einfacher in Kugelkoordinaten ausdrücken:

$$V = \int_0^{2\pi} \int_0^{\pi/4} \int_0^1 r^2 \sin \theta dr d\theta d\varphi$$

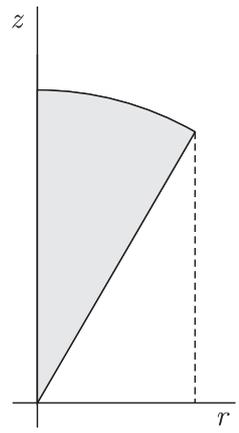


Abbildung 10

#### 5.4.16 Beispiel.

Besonders interessant ist die Berechnung des Volumens des Objektes, das oberhalb des Kegels  $z = \sqrt{x^2 + y^2}$  liegt und unterhalb der Kugel mit Radius  $1/2$  und Mittelpunkt  $(0, 0, 1/2)$ , die die Gleichung  $x^2 + y^2 + z^2 = z$  erfüllt. Da in Kugelkoordinaten gilt  $z = r \cos \theta$ , lässt sich die Gleichung für die Kugel umschreiben als  $r^2 = r \cos \theta$ , und damit ist der Bereich durch

$$0 \leq r \leq 1, 0 \leq \theta \leq \pi/4, 0 \leq \phi \leq 2\pi$$

gegeben.

# 6 Kurvenintegrale

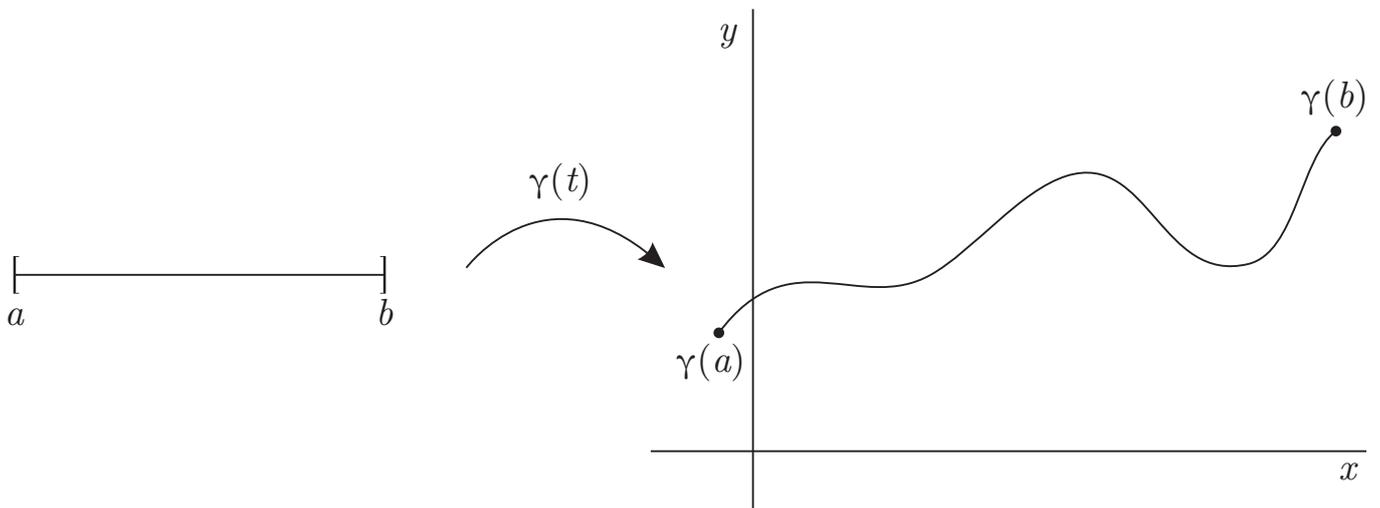


Abbildung 6.1: Kurve mit Parameterdarstellung  $\gamma$

**6.0.17 Definition. Parameterdarstellung** (a) Sei  $a < b$ ,  $\gamma : [a, b] \rightarrow \mathbb{R}^n$  heißt PARAMETERDARSTELLUNG einer Kurve mit Parameterintervall  $[a, b]$ . Der Anfangspunkt der Kurve ist  $\gamma(a)$ , der Endpunkt ist  $\gamma(b)$ .

(b) Zwei Parameterdarstellungen

$$\gamma_1 : [a, b] \rightarrow \mathbb{R}^n$$

$$\gamma_2 : [c, d] \rightarrow \mathbb{R}^n$$

heißen ÄQUIVALENT genau dann, wenn es eine stetige, monoton wachsende Abbildung von einem Intervall in das andere gibt:

$$\gamma_1 \sim \gamma_2 \Leftrightarrow \exists \varphi : [a, b] \rightarrow [c, d] \text{ stetig und monoton wachsend}$$

Es ist dann  $\gamma_2(t) = \gamma_1 \circ \varphi(t)$ .

**6.1 Beispiele.** (a) Die Gerade zwischen zwei Punkten  $P, Q \in \mathbb{R}^n$  kann durch

$$\gamma(t) = P + t(Q - P) = tQ + (1 - t)P \quad t \in [0, 1]$$

parametrisiert werden.

(b) Der Kreis mit Radius 1 hat die Gleichung  $x^2 + y^2 = 1$ , und kann durch

$$\gamma(t) = (\cos t, \sin t), \quad t \in [0, 2\pi)$$

parametrisiert werden. Der Kreis mit Radius  $r$  und Mittelpunkt  $(a, b)$  hat dann die Gleichung  $(x - a)^2 + (y - b)^2 = r^2$ , und die Parametrisierung

$$\gamma(t) = (r \cos t - a, r \sin t - b), \quad t \in [0, 2\pi).$$

Mittels  $t = 5s$  zum Beispiel ändert sich die Kurve nicht, wenn man sie als Menge von Punkten sieht, aber die Durchlaufgeschwindigkeit ändert sich.

- (c) Die Schnittkurve des Zylinders  $x^2 + y^2 = 1$  und der Ebene  $y + z = 2$  lässt sich parametrisieren indem man vom Zylinder erkennt, dass  $x = \cos t, y = \sin t$  die Gleichung erfüllt. Die  $z$ -Komponente erhält man von der Ebene. Nämlich  $z = 2 - y = 2 - \sin t$ , womit

$$\gamma(t) = (\cos t, \sin t, 2 - \sin t). \quad t \in [0, 2\pi).$$

**6.0.18 Definition. Länge einer Kurve**

Die Länge einer Kurve lässt sich annähern durch eine Aneinanderreihung von geraden Stücken (vgl. Abbildung 6.2). Das Supremum für beliebig viele solcher Stücke ist dann die Länge der Kurve:

$$L := \sup \left\{ \sum_{i=1}^m \|\gamma(t_i) - \gamma(t_{i-1})\| \mid m \in \mathbb{N}, \{t_0 = a, t_1, \dots, t_m = b\} \text{ ist Zerlegung von } [a, b] \right\}$$

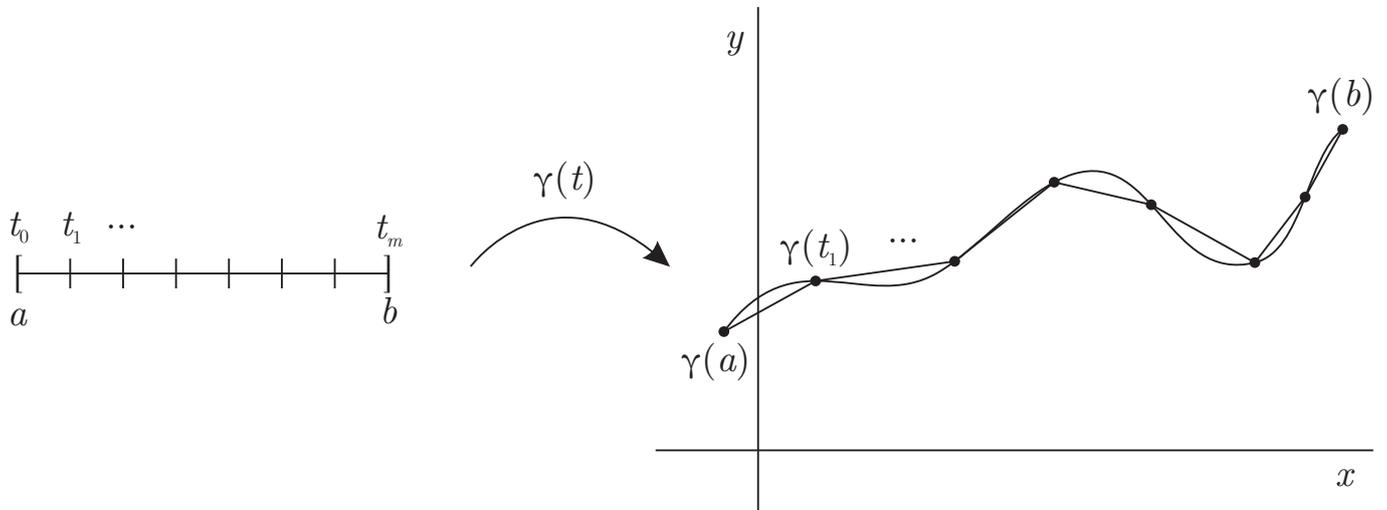


Abbildung 6.2: Annäherung einer Kurve durch Polygone

**6.0.19 Definition.**

Sei  $C$  eine Kurve mit Parameterdarstellung  $\gamma : [a, b] \rightarrow \mathbb{R}^n$ .  $C$  heißt rektifizierbar genau dann, wenn die Länge der Kurve endlich ist:

$$C \text{ heißt rektifizierbar} \Leftrightarrow L(C) < \infty$$

**6.0.20 Satz.**

Die Begriffe „rektifizierbar“ und Länge einer Kurve hängen nicht von der speziellen Parameterdarstellung ab.

*Proof.* Seien  $\gamma_1$  und  $\gamma_2$  Parameterdarstellungen einer Kurve mit  $\gamma_1 : [a, b] \rightarrow \mathbb{R}^n, \gamma_2 : [c, d] \rightarrow \mathbb{R}^n$ . Sei  $\varphi : [a, b] \rightarrow [c, d]$  und dadurch  $\gamma_1(t) = \gamma_2 \circ \varphi(t), \{t_0, \dots, t_m\}$  sei Zerlegung von  $[a, b]$ . Dann ist  $\{\varphi(t_0), \varphi(t_1), \dots, \varphi(t_m)\}$  Zerlegung von  $[c, d]$ . Es gilt dann

$$\sum_{i=1}^m \|\gamma_1(t_i) - \gamma_1(t_{i-1})\| = \sum_{i=1}^m \|\gamma_2(\varphi(t_i)) - \gamma_2(\varphi(t_{i-1}))\|$$

Die Länge der ersten Kurve

$$L_1 = \sup \left\{ \sum_{i=1}^m \|\gamma_1(t_i) - \gamma_1(t_{i-1})\| \mid \{t_0, \dots, t_m\} \text{ Zerlegung von } [a, b] \right\}$$

ist dann definitionsgemäß kleiner als die der zweiten Kurve:

$$\Rightarrow L_1 \leq L_2$$

Verwendet man andersherum die Umkehrfunktion von  $\varphi$ , so ergibt sich

$$\begin{aligned} \gamma_2(t) &= \gamma_1 \circ \varphi^{-1}(t) \\ \Rightarrow L_2 &\leq L_1 \end{aligned}$$

Also müssen die beiden Längen gleich sein

$$\Rightarrow L_1 = L_2$$

□

### 6.0.21 Definitionen.

Zu Kurven:

- (a) Eine Kurve  $C$  heißt stetig differenzierbar, wenn es eine stetig differenzierbare Parameterdarstellung gibt.
- (b) Eine Kurve heißt Jordan-Kurve, wenn eine Parameterdarstellung  $\gamma : [a, b] \rightarrow \mathbb{R}^n$  existiert und diese injektiv ist.
- (c) Seien  $C_1, C_2$  Kurven in  $\mathbb{R}^n$  mit Parameterdarstellungen  $\gamma_1 : [a, b] \rightarrow \mathbb{R}^n$  und  $\gamma_2 : [b, c] \rightarrow \mathbb{R}^n$ . Der Endpunkt von  $C_1$  sei der Anfangspunkt von  $C_2$ . Dann heißt die durch die Parameterdarstellung

$$\gamma : [a, c] \rightarrow \mathbb{R}^n \quad t \mapsto \begin{cases} \gamma_1(t) & t \in [a, b] \\ \gamma_2(t) & t \in [b, c] \end{cases}$$

dargestellte Kurve Summe von  $C_1$  und  $C_2$ , man schreibt:  $C_1 + C_2$ . Unter  $-C_1$  versteht man die Kurve mit Parameterdarstellung

$$\gamma : [0, 1] \rightarrow \mathbb{R}^n \quad t \mapsto \gamma_1(b + t(a - b))$$

mit  $\gamma_1 : [a, b] \rightarrow \mathbb{R}^n$ .

- (d) Eine Kurve  $C$  heißt stückweise stetig differenzierbar, wenn stetig differenzierbare Kurven  $C_1, \dots, C_n$  mit  $C = C_1 + \dots + C_n$  existieren.

### 6.0.22 Satz.

Sei  $C$  stetig differenzierbare Kurve in  $\mathbb{R}^n$  mit Parameterdarstellung  $\gamma \in C^1([a, b], \mathbb{R}^n)$ . Dann gilt

$$L(C) = \int_a^b \|\gamma'(t)\| dt < \infty$$

*Proof.* Es sei  $z = \{t_0 = a, t_1, \dots, t_m = b\}$  eine Zerlegung des Intervalls  $[a, b]$  mit Feinheit  $|z|$ . Nähere die Länge der Kurve durch eine Aneinanderreihung von geraden Stücken an:

$$V_z(\gamma) = \sum_{i=1}^m \|\gamma(t_i) - \gamma(t_{i-1})\|$$

Es ist zu zeigen, dass

$$\forall \epsilon > 0 \exists \delta > 0 \quad |z| < \delta \quad \left| V_z(\gamma) - \int_a^b \|\gamma'(t)\| dt \right| < \epsilon$$

Bemerke, dass der Differenzvektor  $\gamma(t_i) - \gamma(t_{i-1})$  sich aus  $n$  Komponenten zusammensetzt:

$$\gamma(t_i) - \gamma(t_{i-1}) = (\gamma_1(t_i) - \gamma_1(t_{i-1}), \dots, \gamma_n(t_i) - \gamma_n(t_{i-1}))$$

Nach dem Mittelwertsatz der Differentialrechnung gibt es nun einen Zwischenpunkt  $\tau_i = (\tau_i^1, \dots, \tau_i^n) \in \mathbb{R}^n$  sodass gilt

$$(\gamma_1(t_i) - \gamma_1(t_{i-1}), \dots, \gamma_n(t_i) - \gamma_n(t_{i-1})) = (\gamma_1'(\tau_i^1)(t_i - t_{i-1}), \dots, \gamma_n'(\tau_i^n)(t_i - t_{i-1})) = \gamma'(\tau_i)(t_i - t_{i-1})$$

mit  $\gamma'(\tau_i) = (\gamma_1'(\tau_i^1), \dots, \gamma_n'(\tau_i^n))$ . Es folgt damit für  $V_z$ :

$$V_z(\gamma) = \sum_{i=1}^m \int_{t_{i-1}}^{t_i} \|\gamma'(\tau_i)\| dt$$

Damit gilt

$$\begin{aligned} \left| V_z(\gamma) - \int_a^b \|\gamma'(t)\| dt \right| &\leq \sum_{i=1}^m \left| \int_{t_{i-1}}^{t_i} \|\gamma'(\tau_i)\| dt - \int_{t_{i-1}}^{t_i} \|\gamma'(t)\| dt \right| \\ &\leq \sum_{i=1}^m \int_{t_{i-1}}^{t_i} \left| \|\gamma'(\tau_i)\| - \|\gamma'(t)\| \right| dt = (*) \end{aligned}$$

Es gilt für den Integranden:

$$\left| \|\gamma'(\tau_i)\| - \|\gamma'(t)\| \right| \leq \|\gamma'(\tau_i) - \gamma'(t)\| = \sqrt{\sum_{k=1}^n |\gamma_k'(\tau_i^k) - \gamma_k'(t)|^2}$$

Wähle nun  $\delta$  so, dass für  $|t - s| < \delta$ ,  $t, s \in [a, b]$

$$|\gamma_k'(s) - \gamma_k'(t)| < \frac{\epsilon}{\sqrt{n(b-a)}} \quad \forall k \in \{1, \dots, n\}$$

$\gamma'$  ist stetig und insbesondere gleichmäßig stetig auf dem definierten Intervall, also gilt

$$\begin{aligned} (*) &\leq \sum_{i=1}^m \int_{t_{i-1}}^{t_i} \sqrt{\sum_{k=1}^n |\gamma_k'(\tau_i) - \gamma_k'(t)|^2} dt \\ &< \sum_{i=1}^m \sqrt{\frac{n\epsilon^2}{n(b-a)^2}} (t_i - t_{i-1}) = \epsilon \end{aligned}$$

□

### 6.0.23 Beispiel.

Zur Länge einer Kurve

- (a) Möchte man die Länge einer eindimensionalen Funktion auf einem Intervall bestimmen, so wählt man die Parametrisierung

$$\gamma(t) = (t, f(t))$$

Die Ableitung  $\gamma'$  ist dann

$$\gamma'(t) = (1, f'(t))$$

Damit ist die Länge

$$L = \int_a^b \|\gamma'(t)\| dt = \int_a^b \sqrt{1 + (f'(t))^2} dt$$

(b) Gegeben sei eine Ellipse durch die Parametrisierung

$$\gamma(t) = (a \cos t, b \sin t) \quad 0 \leq t \leq 2\pi$$

Deren Ableitung nach  $t$  lautet

$$\gamma'(t) = (-a \sin t, b \cos t)$$

Die Ellipse wird im mathematisch positiven Sinne (gegen den Uhrzeigersinn) durchlaufen, der Startpunkt liegt bei  $(a, 0)$ . Die Länge ist

$$\begin{aligned} L &= \int_0^{2\pi} \sqrt{a^2 \sin^2 t + b^2 \cos^2 t} dt \\ &= \int_0^{2\pi} \sqrt{a^2 + (b^2 - a^2) \cos^2 t} dt \\ &= \int_0^{2\pi} a \sqrt{1 - k^2 \cos^2 t} dt && \text{mit } k = \frac{\sqrt{a^2 - b^2}}{a} \\ &= 4 \int_0^{\pi/2} a \sqrt{1 - k^2 \cos^2 t} dt \end{aligned}$$

#### 6.0.24 Bemerkung.

Es bezeichne die  $s(t)$  die Bogenlänge einer Kurve

$$s(t) = \int_a^t \|\gamma'(\tau)\| d\tau \quad \frac{ds}{dt} = \|\gamma'(t)\|$$

Parametrisiert man die Kurve nun mit  $s$ , sprich  $t \mapsto s(t)$  so kann man umgekehrt, falls  $\|\dot{\gamma}(t)\| \neq 0$  für alle  $t \in [a, b]$ , auch  $s \mapsto t(s)$  verwenden, ist dann

$$\tilde{\gamma}(s) := \gamma \circ t(s) = \gamma(t(s))$$

Der Geschwindigkeitsvektor dieser Bogenlängenparametrisierung ist

$$\tilde{\gamma}'(s) = \frac{d}{ds} \tilde{\gamma} = \frac{d}{ds} (\gamma(t(s))) = \gamma' \frac{dt}{ds} = \frac{\gamma'}{\frac{ds}{dt}} = \frac{\gamma'(t(s))}{\|\gamma'(t(s))\|}$$

Der neue Geschwindigkeitsvektor ist also normiert:

$$\Rightarrow \|\tilde{\gamma}'\| = 1$$

#### 6.0.25 Definitionen.

Sei  $\gamma : [a, b] \rightarrow \mathbb{R}^n$  eine stetig differenzierbare Parameterdarstellung der Kurve  $C$ ,  $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  ein stetiges Vektorfeld.

(a) Das Kurvenintegral von  $\mathbf{F}$  längs der Kurve  $C$  ist

$$\int_a^b \mathbf{F}(\gamma(t)) \cdot \gamma'(t) dt$$

Symbolisch schreibt man

$$\int_C \mathbf{F} \cdot d\mathbf{x}$$

(b)  $\mathbf{F} = (P, Q)$  für  $n = 2$ ,  $\mathbf{F} = (P, Q, R)$  für  $n = 3$ . Man schreibt

$$\int_C \mathbf{F} \cdot d\mathbf{x} = \int_C P dx + Q dy + R dz$$

### 6.0.26 Satz.

Seien  $C_1, C_2$  stückweise stetig differenzierbare Kurven und  $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  ein stetiges Vektorfeld.

(a)

$$\int_{-C_1} \mathbf{F} \cdot d\mathbf{x} = - \int_{C_1} \mathbf{F} \cdot d\mathbf{x}$$

(b)

$$\int_{C_1+C_2} \mathbf{F} \cdot d\mathbf{x} = \int_{C_1} \mathbf{F} \cdot d\mathbf{x} + \int_{C_2} \mathbf{F} \cdot d\mathbf{x}$$

(c)

$$\left| \int_{C_1} \mathbf{F} \cdot d\mathbf{x} \right| \leq \sup_{x \in C_1} \|\mathbf{F}(x)\| L(C_1)$$

*Proof.* (c) Sei  $\gamma$  Parametrisierung von  $C_1$

$$\begin{aligned} \left| \int_{C_1} \mathbf{F} \cdot d\mathbf{x} \right| &= \left| \int_a^b \mathbf{F}(\gamma(t)) \cdot \gamma'(t) dt \right| \\ &\leq \int_a^b \|\mathbf{F}(\gamma(t))\| \|\gamma'(t)\| dt \\ &\leq \sup_{x \in C_1} \|\mathbf{F}(x)\| \int_a^b \|\gamma'(t)\| dt \\ &= \sup_{x \in C_1} \|\mathbf{F}(x)\| L(C_1) \end{aligned}$$

□

### 6.0.27 Beispiel.

In der  $x - y$ -Ebene sei das Vektorfeld  $\mathbf{F}$  gegeben durch

$$\mathbf{F} = (y, y - x)$$

oder in der Form

$$\mathbf{F} = (P, Q) \text{ mit } P = y \text{ und } Q = y - x.$$

Wir betrachten zwei Kurven vom Ursprung nach  $(1, 1)$ , gegeben durch  $C_1$  und  $C_2$ .  $C_1$  ist zusammengesetzt aus einem horizontalen Stück nach  $(1, 0)$ ,  $C_{11}$  und einem vertikalen Stück von dort bis zum Endpunkt,  $C_{12}$ . Wir verwenden folgende Parametrisierung für  $C_1$ :

$$\begin{aligned} C_{11} : \gamma_{11}(t) &= (t, 0) & 0 \leq t \leq 1 \\ C_{12} : \gamma_{12}(t) &= (1, t) & 0 \leq t \leq 1 \end{aligned}$$

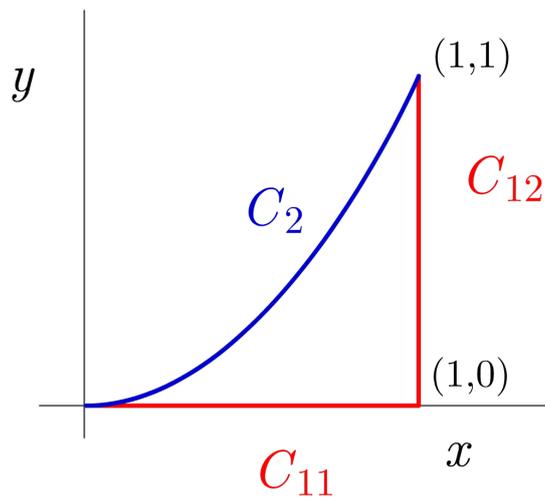


Abbildung 6.3: Zwei Kurven

So folgt für die Integrale:

$$\begin{aligned}
 \int_{C_{11}} \mathbf{F} \cdot d\mathbf{x} &= \int_0^1 \mathbf{F}(\gamma_{11}(t)) \cdot \gamma'_{11}(t) dt \\
 &= \int_0^1 (0, 0 - t) \cdot (1, 0) dt = 0 \\
 \int_{C_{12}} \mathbf{F} \cdot d\mathbf{x} &= \int_0^1 \mathbf{F}(\gamma_{12}(t)) \cdot \gamma'_{12}(t) dt \\
 &= \int_0^1 (t, t - 1) \cdot (0, 1) dt \\
 &= \int_0^1 (t - 1) dt = -\frac{1}{2}
 \end{aligned}$$

Die Kurve  $C_2$  sei gegeben durch ein Stück einer Parabel  $y = x^2$ .

$$C_2 : \gamma_2(t) = (t, t^2) \quad 0 \leq t \leq 1$$

Oder als einzelne Parameter  $x$  und  $y$  betrachtet:

$$x(t) = t, \quad y(t) = t^2$$

so hat das Integral die Form

$$\begin{aligned}
 \int_{C_2} \mathbf{F} \cdot d\mathbf{x} &= \int_{C_2} P dx + Q dy \\
 &= \int_0^1 y \frac{dx}{dt} dt + \int_0^1 (y-x) \frac{dy}{dt} dt \\
 &= \int_0^1 t^2 dt + \int_0^1 (t^2-t) 2t dt \\
 &= \frac{1}{3} + 2 \left[ \frac{t^4}{4} - \frac{t^3}{3} \right]_0^1 = \frac{1}{3} - \frac{1}{6} = \frac{1}{6}
 \end{aligned}$$

### 6.0.28 Bemerkung.

Es bezeichne  $\mathbf{T}$  den normierten Tangentialvektor zu einer Kurve, die durch  $\gamma(t)$  parametrisiert ist:

$$\mathbf{T}(\gamma(t)) = \frac{\gamma'(t)}{\|\gamma'(t)\|}$$

Damit kann man das Wegintegral über ein Vektorfeld  $\mathbf{F}$  umschreiben:

$$\begin{aligned}
 \int_C \mathbf{F} \cdot d\mathbf{x} &= \int_a^b \mathbf{F}(\gamma(t)) \cdot \gamma'(t) dt \\
 &= \int_a^b \mathbf{F}(\gamma(t)) \cdot \gamma'(t) \frac{\|\gamma'(t)\|}{\|\gamma'(t)\|} dt \\
 &= \int_a^b \mathbf{F}(\gamma(t)) \cdot \mathbf{T}(\gamma(t)) \|\gamma'(t)\| dt \\
 &= \int_C \mathbf{F} \cdot \mathbf{T} ds
 \end{aligned}$$

mit dem Linienelement  $ds = \|\gamma'(t)\| dt$ . Wenn nun  $f = \mathbf{F} \cdot \mathbf{T} = 1$  ist, gilt

$$\int_C f ds = \int_C 1 ds = L(C)$$

**6.2 Bemerkung.** Auf diese Art und Weise lässt sich so manch Wegintegral mit rein geometrischen Überlegungen berechnen. Sei nun  $F_1(x, y) = (x, y)$ , und  $C_r$  ein Kreis mit Radius  $r$  und Mittelpunkt  $(0, 0)$ , dann ist klarerweise

$$\int_{C_r} F_1 \cdot d\mathbf{x} = 0,$$

da  $F_1$  an jedem Punkt des Kreises orthogonal auf den Tangentialvektor  $\mathbf{T}$  steht.

Sei nun  $F_2 = (y, -x)$ , dann lässt sich das Wegintegral über den Kreis ebenfalls sofort hinschreiben, da  $F_2$  parallel zum Tangentialvektor verläuft und an jedem Punkt die Länge  $r$  hat. Damit ist

$$\int_{C_r} F_1 \cdot d\mathbf{x} = \int_{C_r} \mathbf{F} \cdot \mathbf{T} ds = \int_{C_r} r ds = rL(C_r) = 2\pi r^2.$$

Somit kann man auch Wegintegrale für Linearkombinationen von  $F_1$  und  $F_2$  leicht berechnen.

### 6.0.29 Definition. Gradientenfeld

Sei  $A \subseteq \mathbb{R}^n$  Gebiet.  $\mathbf{F}$  heißt GRADIENTENFELD  $\Leftrightarrow \exists \varphi \in C^1(A, \mathbb{R}^n)$  ist Skalarfeld mit  $\mathbf{F} = \nabla \varphi$ .  $-\varphi$  heißt Potential,  $\varphi$  Stammfunktion.

### 6.0.30 Satz.

$A \subseteq \mathbb{R}^n$  Gebiet,  $\mathbf{F} : A \rightarrow \mathbb{R}^n$  stetiges Gradientenfeld mit Stammfunktion  $\varphi : A \rightarrow \mathbb{R}$ . Sei  $C$  Kurve in  $\mathbb{R}^n$  mit Anfangspunkt  $a \in \mathbb{R}^n$  und Endpunkt  $b \in \mathbb{R}^n$ . Das Kurvenintegral hängt dann lediglich von diesen Punkten ab:

$$\int_C \mathbf{F} \cdot dx = \varphi(a) - \varphi(b)$$

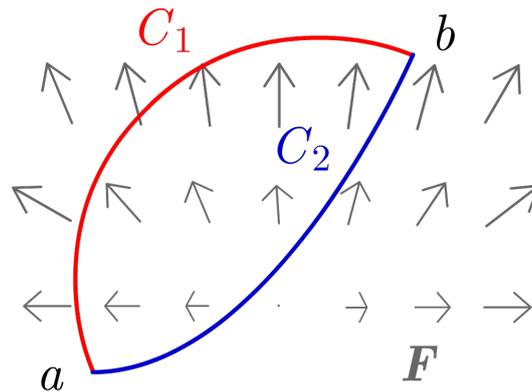


Abbildung 6.4: Ein konservatives Kraftfeld, das von 2 verschiedenen Kurven durchlaufen wird.

*Proof.*

$$\begin{aligned} \int_C \mathbf{F} \cdot dx &= \int_0^1 \mathbf{F}(\gamma(t)) \cdot \gamma'(t) dt \\ &= \int_0^1 \nabla \varphi(\gamma(t)) \cdot \gamma'(t) dt \\ &= \int_0^1 \frac{d}{dt} \varphi(\gamma(t)) dt \\ &= \varphi(\gamma(1)) - \varphi(\gamma(0)) \\ &= \varphi(b) - \varphi(a) \end{aligned}$$

□

### 6.0.31 Definition.

$A \subseteq \mathbb{R}^n$  Gebiet,  $\mathbf{F}$  stetiges Vektorfeld  $\mathbf{F} : A \rightarrow \mathbb{R}^n$ .  $\mathbf{F}$  heißt konservativ genau dann, wenn  $\forall a, b \in \mathbb{R}^n, \forall C_1, C_2$  (stückweise stetig differenzierbar) mit Anfangspunkt  $a$  und Endpunkt  $b$  gilt:

$$\int_{C_1} \mathbf{F} \cdot dx = \int_{C_2} \mathbf{F} \cdot dx$$

### 6.0.32 Satz.

Sei  $A \subseteq \mathbb{R}^n$  Gebiet,  $\mathbf{F} : A \rightarrow \mathbb{R}^n$  stetig. Dann:

$$\mathbf{F} \text{ ist konservativ} \Leftrightarrow \mathbf{F} \text{ ist Gradientenfeld}$$

*Proof.* Eine Richtung haben wir schon gezeigt.

⇐: Siehe oben

⇒:

$$\varphi(x) = \int_a^x \mathbf{F}(z) \cdot dz$$

Es bist zu zeigen, dass die so definierte Funktion  $\varphi$  folgendes erfüllt

$$\begin{aligned} \nabla\varphi &= \mathbf{F} \\ \Rightarrow \varphi(x+h) - \varphi(x) &= \mathbf{F}(x) \cdot h + \psi(h), \end{aligned}$$

wobei  $\psi(h)$  schneller gegen 0 geht als linear.

$$\varphi(x+h) - \varphi(x) = \int_a^{x+h} \mathbf{F}(z) \cdot dz - \int_a^x \mathbf{F}(z) \cdot dz = \int_x^{x+h} \mathbf{F}(z) \cdot dz$$

$$\gamma : t \mapsto x + th \quad t \in \mathbb{R}, h \in \mathbb{R}^n$$

$$\begin{aligned} & \int_0^1 \mathbf{F}(x+th) \cdot \gamma'(t) dt \\ &= \int_0^1 \mathbf{F}(x+th) \cdot h dt \\ &= \mathbf{F}(x) \cdot h + \left[ \int_0^1 \mathbf{F}(x+th) \cdot h dt - \mathbf{F}(x) \cdot h \right] \\ &= \mathbf{F}(x) \cdot h + \psi(h) \end{aligned}$$

$$\begin{aligned} \psi(h) &= \int_0^1 \mathbf{F}(x+th) \cdot h - \mathbf{F}(x) \cdot h \\ &= \int_0^1 [\mathbf{F}(x+th) - \mathbf{F}(x)] \cdot h dt \\ &\Rightarrow \|\psi(h)\| \leq \sup_{t \in [0,1]} \|\mathbf{F}(x+th) - \mathbf{F}(x)\| \cdot \|h\| \\ &\Leftrightarrow \frac{\|\psi(h)\|}{\|h\|} \leq \sup_{t \in [0,1]} \|\mathbf{F}(x+th) - \mathbf{F}(x)\| \rightarrow 0 \text{ f\u00fcr } h \rightarrow 0 \text{ da } \mathbf{F} \text{ stetig.} \end{aligned}$$

□

Sei  $\mathbf{F}$  ein Gradientenfeld und  $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$  und  $\nabla\varphi = \mathbf{F}$ . Die Partiellen Ableitungen von  $\varphi$  sind die Komponenten von  $\mathbf{F}$ :

$$\begin{aligned} \partial_x \varphi &= F_1 \\ \partial_y \varphi &= F_2 \\ \Rightarrow \frac{\partial^2 \varphi}{\partial y \partial x} &= \frac{\partial}{\partial y} F_1 \text{ und} \\ \frac{\partial^2 \varphi}{\partial x \partial y} &= \frac{\partial}{\partial x} F_2 \end{aligned}$$

Wir nehmen an, dass  $\varphi$  zweimal stetig partiell differenzierbar ist, sodass die Ableitungen vertauschen:

$$\begin{aligned}\frac{\partial^2 \varphi}{\partial y \partial x} &= \frac{\partial^2 \varphi}{\partial x \partial y} \\ \Rightarrow \frac{\partial}{\partial y} F_1 &= \frac{\partial}{\partial x} F_2\end{aligned}$$

### 6.0.33 Beispiel.

$\mathbf{F} = (P, Q)$  Gradientenfeld

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

Nun ist das Skalarfeld  $\varphi$  zu finden mit  $\nabla \varphi = \mathbf{F}$

$$\begin{aligned}P &= 2xy & \Rightarrow \frac{\partial P}{\partial y} &= 2x \\ Q &= x^2 + 4y & \Rightarrow \frac{\partial Q}{\partial x} &= 2x\end{aligned}$$

$$\begin{aligned}P &= \varphi_x = 2xy \\ \Rightarrow \varphi &= x^2 y + c(y) \\ Q &= \varphi_y = x^2 + c'(y) = x^2 + 4y \\ \Rightarrow c'(y) &= 4y \\ \Rightarrow c(y) &= 2y^2 \\ \Rightarrow \varphi &= x^2 y + 2y^2 \\ \Rightarrow \int_C \mathbf{F} \cdot d\mathbf{x} &= \varphi(b) - \varphi(a)\end{aligned}$$

Mit Start und Endpunkten  $a, b \in \mathbb{R}^2$ .

### 6.0.34 Definition.

Eine Menge  $A \subseteq \mathbb{R}^n$  heißt STERNFÖRMIG, wenn es einen Punkt  $a$  gibt, von dem jeder Punkt der Menge mit einer Geraden durch  $a$  verbunden werden kann.

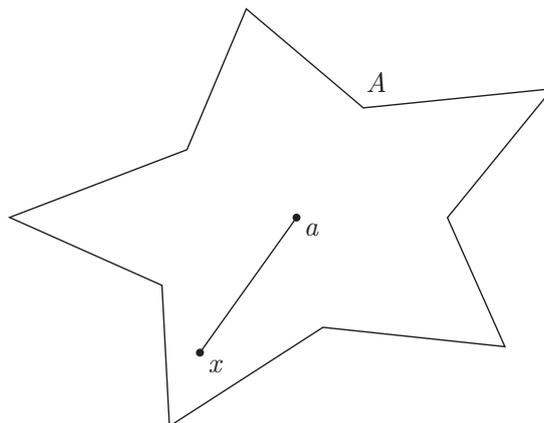


Abbildung 6.5: Eine sternförmige Menge

### 6.0.35 Definition.

Eine Menge  $A \subseteq \mathbb{C}^n$  heißt KONVEX, wenn zwei beliebige Punkte der Menge durch eine Gerade verbunden werden können, sodass die Gerade noch in  $A$  ist. Insbesondere ist jede konvexe Menge sternförmig.

**6.0.36 Satz.**

Sei  $A \subseteq \mathbb{R}^n$ , ein Gebiet,  $\mathbf{F} \in C^1(A, \mathbb{R}^n)$ . Ist  $\mathbf{F} = (F_1, \dots, F_n)$  Gradientenfeld, so gilt:

$$\frac{\partial}{\partial x_i} F_j = \frac{\partial}{\partial x_j} F_i \quad \forall i, j \in \{1, \dots, n\} \quad (6.1)$$

Ist  $A$  sternförmig, so gilt auch die Umkehrung, also aus (6.1) folgt  $\mathbf{F}$  ist Gradientenfeld.

*Proof.* Die Gerade sei parametrisiert durch  $\gamma(t) = a + t(x - a)$ .

$$\begin{aligned} \varphi(x) &= \int_a^x \mathbf{F}(z) \cdot dz \\ &= \int_0^1 \mathbf{F}(a + t(x - a)) \cdot (x - a) dt \\ &= \sum_{k=1}^n \int_0^1 F_k(a + t(x - a)) \cdot (x_k - a_k) dt \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial x_j} \varphi(x) &= \int_0^1 F_j(a + t(x - a)) dt + \sum_{k=1}^n \int_0^1 \left( \frac{\partial}{\partial x_j} F_k \right) (a + t(x - a)) t \cdot (x_k - a_k) dt \\ &= \int_0^1 F_j(a + t(x - a)) dt + \sum_{k=1}^n \int_0^1 \left( \frac{\partial}{\partial x_k} F_j \right) (a + t(x - a)) \cdot (x_k - a_k) t dt \\ &= \int_0^1 F_j(a + t(x - a)) dt + \int_0^1 \left( \frac{d}{dt} F_j(a + t(x - a)) \right) t dt \\ &= \int_0^1 F_j(a + t(x - a)) dt + F_j(a + t(x - a)) t \Big|_{t=0}^1 - \int_0^1 F_j(a + t(x - a)) dt \\ &= F_j(x) \end{aligned}$$

□

**6.0.37 Korollar.**

Es sei  $A \subseteq \mathbb{R}^3$  konvex und  $\mathbf{F} \in C^1(A, \mathbb{R})$ . Dann gelten folgende Äquivalenzen

$$\begin{aligned} \mathbf{F} \text{ ist konservativ} &\Leftrightarrow \mathbf{F} \text{ ist Gradientenfeld} \\ \mathbf{F} = \nabla \varphi, \varphi : A \rightarrow \mathbb{R} &\Leftrightarrow \mathbf{F} \text{ ist rotationsfrei (rot } \mathbf{F} = 0) \\ \mathbf{F} \text{ konservativ} &\Leftrightarrow \int_{C_1} \mathbf{F} \cdot d\mathbf{x} = \int_{C_2} \mathbf{F} \cdot d\mathbf{x} \text{ für beliebiges } C_1 \text{ und } C_2 \\ &\text{mit gleichem Anfangs- und Endpunkt} \\ &\Leftrightarrow \varphi(x) = \int_a^x \mathbf{F}(z) dz \end{aligned}$$

Es vertauschen die ersten Ableitungen der Komponenten von  $\mathbf{F} = (F^1, F^2, F^3)$ :

$$\frac{\partial F^i}{\partial x_j} = \frac{\partial F^j}{\partial x_i}$$

### 6.0.38 Bemerkung. Masse

Mit einer Dichtefunktion  $\rho : \mathbb{R}^n \rightarrow \mathbb{R}$ , der Bogenlänge  $S(t) = \int_a^t \|\gamma'(\tau)\| d\tau$  und dem Zwischenpunkt  $t'_i$ , der zwischen  $t_i$  und  $t_{i-1}$  liegt, ist die Masse gegeben durch

$$\sum_{i=1}^n \rho(\gamma(t'_i)) \cdot \underbrace{(S(t_i) - S(t_{i-1}))}_{\approx S'(t_i)(t_i - t_{i-1})}$$

Es ergibt sich als Grenzwert

$$\begin{aligned} M &= \int_a^b \rho(\gamma(t)) S'(t) dt \\ &= \int_a^b \rho(\gamma(t)) \|\gamma'(t)\| dt \end{aligned}$$

### 6.0.39 Definition.

Es sei  $f$  eine skalare Funktion, das Linienintegral über die Kurve  $C$  ist

$$\int_C f(x) ds = \int_a^b f(\gamma(t)) \|\gamma'(t)\| dt$$

### 6.0.40 Definition.

Sei  $D \subseteq \mathbb{R}^2$  ein beschränktes Gebiet. Der Rand von  $D$  ( $\partial D$ ) besteht aus endlicher Vereinigung stückweiser stetig differenzierbarer Jordan-Kurven  $C_i, i \in \{1, \dots, n\}$ .  $C_i$  heißt POSITIV ORIENTIERT falls  $\gamma_i$ , Parametrisierung von  $C_i$ , so ist, dass der Vektor  $(-\gamma'_2, \gamma'_1)$  in  $D$  hineinzeigt. Anders formuliert:  $D$  liegt links von  $\gamma_i$ .

### 6.0.41 Satz. Green'scher Integralsatz

Sei  $D \subseteq \mathbb{R}^2$ ,  $\mathbf{F} = (F^1, F^2) = (P, Q) \in C^1(D, \mathbb{R}^2)$ .  $C = \partial D$  sei positiv orientierte Randkurve. Dann gilt

$$\begin{aligned} \int_C \mathbf{F} \cdot d\mathbf{x} &= \oint_{\partial D} \mathbf{F} \cdot d\mathbf{x} = \iint_D \left( \frac{\partial F^2}{\partial x} - \frac{\partial F^1}{\partial y} \right) d(x, y) \\ &= \iint_D (Q_x - P_y) d(x, y) \end{aligned}$$

*Proof.* Annahme:  $D$  ist Normalbereich, das heißt

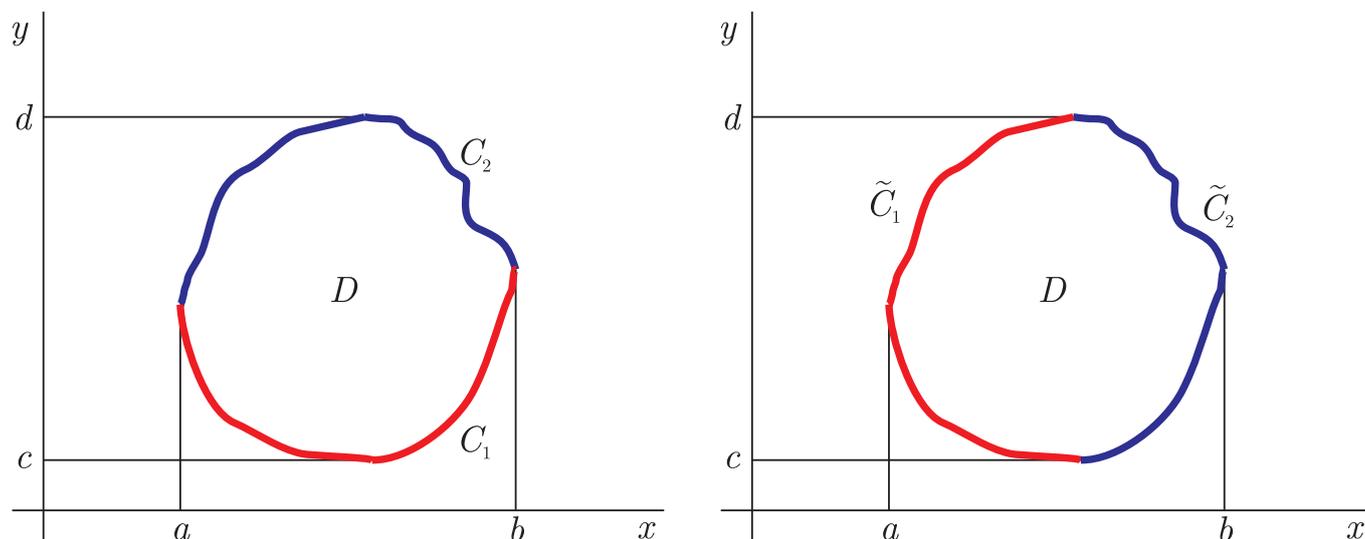
$$\begin{aligned} D &= \{(x, y) \mid a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\} \\ &\stackrel{!}{=} \{(x, y) \mid c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\} \end{aligned}$$

Siehe auch Abbildung 6.6. Wir zeigen zuerst, dass

$$\oint_{\partial D} P dx = - \iint_D P_y d(x, y)$$

Parametrisiere dazu

$$\begin{aligned} C_1 : \gamma_1(t) &= (t, \varphi_1(t)) \quad a \leq t \leq b \\ -C_2 : \gamma_2(t) &= (t, \varphi_2(t)) \quad a \leq t \leq b \end{aligned}$$

Abbildung 6.6: Der selbe Bereich  $D$  wird mit verschiedenen Kurven umrundet

Es ist dann

$$\begin{aligned}
 \oint_{\partial D} P dx &= \int_a^b P(\gamma_1(t))(\gamma_1'(t))_1 dt - \int_a^b P(\gamma_2(t))(\gamma_2'(t))_2 dt \\
 &= \int_a^b P(\gamma_1(t)) dt - \int_a^b P(\gamma_2(t)) dt \\
 &= \int_a^b P(t, \varphi_1(t)) dt - \int_a^b P(t, \varphi_2(t)) dt \\
 &= - \int_a^b (P(t, \varphi_2(t)) - P(t, \varphi_1(t))) dt \\
 &= - \int_a^b \int_{\varphi_1(t)}^{\varphi_2(t)} \frac{\partial P}{\partial y}(t, y) dy dt \\
 &= - \int_a^b \int_{\varphi_1(t)}^{\varphi_2(t)} \frac{\partial P}{\partial y}(x, y) dy dx
 \end{aligned}$$

wobei im vorletzten Schritt der Hauptsatz der Integralrechnung angewendet wurde:

$$f(b) - f(a) = \int_a^b f'(t) dt$$

Nun zeigen wir weiterhin, dass

$$\int_C Q dy = \iint_D Q_x d(x, y)$$

Seien dazu die Kurven  $\tilde{C}_1 = (\psi_1(y), y)$ ,  $\tilde{C}_2 = (\psi_2(y), y)$ , spalte das Integral dementsprechend auf:

$$\begin{aligned} \int_C Q dy &= \int_{\tilde{C}_1} Q dy + \int_{\tilde{C}_2} Q dy \\ &= \int_c^d Q(\psi_2(y), y) dy - \int_c^d Q(\psi_1(y), y) dy \\ &= \int_c^d \int_{\psi_1(y)}^{\psi_2(y)} Q_x(x, y) dx dy \\ &= \iint_D Q_x d(x, y) \end{aligned}$$

2. Version: Wir betrachten nun einen etwas anderen und allgemeineren Zugang, der auch auf beliebige Gebiete verallgemeinert werden kann. Wie nehmen zuerst ein Quadrat beliebiger Länge  $a$ , und zeigen den Satz dort. Allgemein wird die Aussage dadurch erweitert, dass schrittweise eine beliebige Fläche durch kleine Quadrate approximiert wird.

Zuerst betrachten wir ein kleines Quadrat am Punkt  $(x, y)$  mit Seitenlänge  $\varepsilon$ . Dort gilt

$$\begin{aligned} \int_{C_\varepsilon} F \cdot d\mathbf{x} &= \int_0^\varepsilon (P(x+t, y) - P(x+t, y+\varepsilon)) dt \\ &\quad + \int_0^\varepsilon (Q(x+\varepsilon, y+t) - Q(x, y+t)) dt \\ &= \varepsilon^2 (Q_x(x, y) - P_y(x, y)) + \varepsilon^3 K(\theta, \eta), \end{aligned}$$

wobei  $(\theta, \eta) \in [x, x+\varepsilon] \times [y, y+\varepsilon]$ , und der Satz von Taylor verwendet wurde. Das Wegintegral von  $F = (P, Q)$  entlang eines Quadrates  $C_a$  mit Seitenlänge  $a$ , das in kleine Quadrate der Länge  $\varepsilon$  geteilt wird gilt

$$\begin{aligned} \int_{C_a} F \cdot d\mathbf{x} &= \sum_{i=1}^{a^2/\varepsilon^2} \int_{C_\varepsilon} F \cdot d\mathbf{x} \\ &= \sum_{i=1}^{a^2/\varepsilon^2} [\varepsilon^2 (Q_x - P_y)(x_i, y_i) + \varepsilon^3 K(\theta_i, \eta_i)]. \end{aligned}$$

Der erste Teil der rechten Seite ist das Riemannintegral von

$$\iint_{Q_a} (Q_x - P_y) d(x, y),$$

und der zweite Term geht gegen 0 für  $\varepsilon \rightarrow 0$ , falls  $K$  auf  $Q_a$  uniform beschränkt. □

### 6.0.42 Beispiel.

Berechne das Wegintegral von  $F = (3y - e^{\sin x}, 7x + \sqrt{y^4 + 1})$  entlang des Kreises im Ursprung mit Radius 3. Den Einheitskreis bezeichne wir mit  $D$ . Dann gilt

$$\int_{\partial D} F \cdot d\mathbf{x} = \int \int_D (7 - 3) = 36\pi.$$

### 6.0.43 Beispiel.

Sei  $F = \frac{1}{x^2+y^2}(-y, x)$ , dann kann man sich überzeugen, dass für jede Kurve  $C$ , die den Ursprung beinhaltet das Wegintegral gleich  $2\pi$  ist. Jedes Wegintegral das den Ursprung nicht beinhaltet verschwindet. Diese Beispiel bildet in der komplexen Analysis den Core der gesamten Theorie.

Zuerst zeigt man mittels dem Satz von Green, dass je zwei Kurven das gleiche Resultat bringen, indem man das Gebiet so konstruiert, dass der Innenteil der kleiner Kurve nicht Teil des Gebietes ist. Dann ist die Funktion  $F$  glatt genug um Green anzuwenden, und man sieht, dass das Integral verschwindet, da ausserhalb des Ursprungs  $Q_x - P_y = 0$  gilt. Und man rechnet dann das Wegintegral für einen Kreis mit Radius eins explizit aus.

#### 6.0.44 Bemerkung.

Möchte man den Fluss eines Vektorfeldes  $\mathbf{F} = (P, Q)$  durch eine Kurve  $C = \partial D$  berechnen, welche durch die Parametrisierung

$$\gamma : [a, b] \rightarrow \mathbb{R}^2 \quad t \mapsto (\gamma_1(t), \gamma_2(t))$$

gegeben ist, so benötigt man dazu den Normalenvektor

$$\mathbf{N} = (\gamma_2'(t), -\gamma_1'(t)) \frac{1}{\|\gamma'(t)\|}$$

Der Fluss ist dann

$$\begin{aligned} \int_C \mathbf{F} \cdot \mathbf{N} ds &= \int_a^b \mathbf{F}(\gamma(t)) \cdot (\gamma_2'(t), -\gamma_1'(t)) \frac{\|\gamma'(t)\|}{\|\gamma'(t)\|} dt \\ &= \int_a^b (P(\gamma(t))\gamma_2'(t) - Q(\gamma(t))\gamma_1'(t)) dt \\ &= \int_a^b \mathbf{G} \cdot \gamma'(t) dt & \mathbf{G} &= (-Q, P) = (G^1, G^2) \\ &= \iint_D \left( \frac{\partial}{\partial x} G^2 - \frac{\partial}{\partial y} G^1 \right) d(x, y) \\ &= \iint_D (\partial_x P + \partial_y Q) d(x, y) \\ &= \iint_D \operatorname{div} \mathbf{F} d(x, y) \end{aligned}$$

wobei der Satz von Green angewendet wurde. Es ergibt sich also

$$\int_C \mathbf{F} \cdot \mathbf{N} ds = \iint_D \operatorname{div} \mathbf{F} d(x, y)$$

#### 6.0.45 Korollar.

Sei  $\mathbf{F} = \frac{1}{2}(-y, x)$  und  $C$  positiv orientierte Randkurve des Gebiets  $D$ . Dann gilt:

$$\int_C \mathbf{F} \cdot d\mathbf{x} = \int_C P dx + Q dy = |D| = \text{Flächeninhalt von } D$$

*Proof.* Nach Green gilt

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{x} &= \int_C P dx + Q dy \\ &= \iint_D (Q_x - P_y) d(x, y) \\ &= \iint_D \left( \frac{1}{2} - \left( -\frac{1}{2} \right) \right) d(x, y) \\ &= \iint_D 1 d(x, y) \\ &= |D|\end{aligned}$$

□

Mithilfe dieser Tatsache kann man durch Umfahren einer Fläche deren Flächeninhalt bestimmen, zum Beispiel mit einem Planimeter.

#### **6.0.46 Bemerkung.**

Ist das Vektorfeld  $\mathbf{F}$  von der Form  $\mathbf{F} = (P, Q, 0)$ , so ist  $\text{rot } \mathbf{F} = (0, 0, Q_x - P_y)$ . Wendet man den Satz von Green an, erhält man:

$$\begin{aligned}\int_{\partial D} \mathbf{F} \cdot d\mathbf{x} &= \iint_D (Q_x - P_y) d(x, y) \\ &= \iint_D \text{rot } \mathbf{F} \cdot \mathbf{N} d(x, y) \\ &= \iint_D \text{rot } \mathbf{F} \cdot d\mathbf{S}\end{aligned}$$



# 7 Oberflächenintegrale

## 7.0.47 Definition.

Sei  $K \subset \mathbb{R}^2$  kompakt. Ist  $\varphi : K \rightarrow \mathbb{R}^3$  stetig differenzierbar, so heißt  $\varphi|_K$  PARAMETERDARSTELLUNG einer Fläche in  $\mathbb{R}^3$ . Sei weiterhin  $T \subset \mathbb{R}^2$  kompakt,  $\psi$  Parameterdarstellung mit Parameterbereich  $T$ . Die beiden Abbildungen  $\varphi$  und  $\psi$  sind ÄQUIVALENT, wenn  $\psi = \varphi \circ g$ ,  $g : T \rightarrow K$  injektiv und stetig differenzierbar und entweder  $\det g' > 0$  oder  $\det g' < 0$  auf  $T$  ist.

## 7.0.48 Beispiel.

Betrachte konkrete Beispiele zur Anschauung

- (a) Um die Oberfläche, die die obere Hälfte der Einheitskugel beschreibt zu parametrisieren orientiert man sich an den Kugelkoordinaten. Man erhält dann als Parametrisierung:

$$\varphi(u, v) = (\cos u \sin v, \sin u \sin v, \cos v) \quad 0 \leq u \leq 2\pi \quad 0 \leq v \leq \frac{\pi}{2}$$

- (b) Die Parameterdarstellung der Oberfläche eines Zylinders mit Radius 1 und Höhe 2 lautet

$$\varphi(u, v) = (\cos u, \sin u, v) \quad 0 \leq u \leq 2\pi \quad 0 \leq v \leq 2$$

- (c) Eine ebene Fläche wird parametrisiert durch Linearkombination der aufspannenden Vektoren (mit Startpunkt  $A$ ):

$$\varphi(u, v) = A + \vec{a}u + \vec{b}v \quad 0 \leq u \leq 1\pi \quad 0 \leq v \leq 1$$

- (d) Der Graph der Funktion  $z = f(x, y)$  ist

$$\varphi(x, y) = (x, y, f(x, y))$$

Die Fläche in (c) ist trivialerweise durch  $\|\vec{a} \times \vec{b}\| = \left\| \frac{\partial \varphi}{\partial u} \times \frac{\partial \varphi}{\partial v} \right\|$  gegeben. Diese Anschauung erlaube folgende Definition der Oberfläche.

## 7.0.49 Definition.

Sei  $\varphi : K \rightarrow \mathbb{R}^3$  eine Parameterdarstellung der Fläche  $A$ . Der Flächeninhalt von  $A$  ist

$$I(A) = \iint_K \left\| \frac{\partial \varphi}{\partial u} \times \frac{\partial \varphi}{\partial v} \right\| d(u, v) = \iint_A 1 dS$$

## 7.0.50 Definition.

Sei  $\varphi : K \rightarrow \mathbb{R}^3$  eine Parameterdarstellung einer Fläche  $A \subset \mathbb{R}^3$ .

- (a) Sei  $f$  stetiges Skalarfeld auf  $A$ ,  $f \in C(\varphi(K), \mathbb{R})$ . Dann ist das Oberflächenintegral von  $f$  über  $A$  symbolisch definiert durch

$$\iint_A f dS = \iint_K f(\varphi(u, v)) \left\| \frac{\partial \varphi}{\partial u} \times \frac{\partial \varphi}{\partial v} \right\| d(u, v)$$

- (b) Sei  $\mathbf{F}$  stetiges Vektorfeld auf  $A$ ,  $\mathbf{F} \in C(\varphi(A), \mathbb{R}^3)$ . Mit dem Normalenvektor

$$\mathbf{N}(u, v) = \frac{\varphi_u \times \varphi_v}{\|\varphi_u \times \varphi_v\|}$$

ist dann das Oberflächenintegral von  $\mathbf{F}$  auf  $A$  gegeben durch

$$\begin{aligned} \iint_K \mathbf{F}(\varphi(u, v)) \cdot \mathbf{N} \|\varphi_u \times \varphi_v\| d(u, v) &= \iint_K \mathbf{F}(\varphi(u, v)) \cdot (\varphi_u \times \varphi_v) d(u, v) \\ &= \iint_A \mathbf{F} \cdot d\mathbf{S} \end{aligned} \quad (\text{symbolisch})$$

### 7.0.51 Beispiel.

Die Fläche von (d), also der Fläche eines Graphen  $z = f(x, y)$ , mit  $(x, y) \in K$  ist einfach zu berechnen mittels  $\varphi_x = (1, 0, f_x(x, y))$ ,  $\varphi_y = (0, 1, f_y(x, y))$  was für die Fläche die Formel

$$\iint_K \sqrt{1 + f_x^2 + f_y^2} d(x, y)$$

liefert.

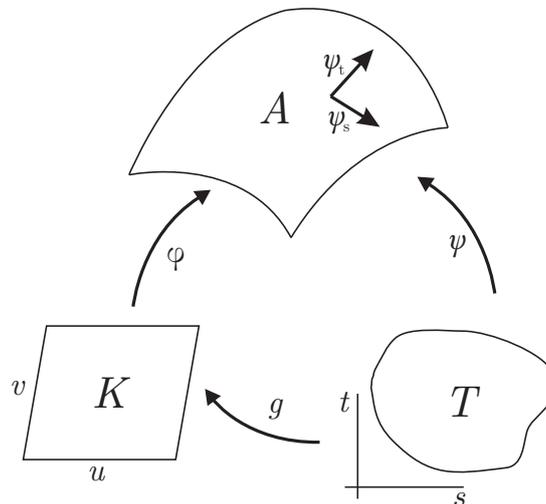


Abbildung 7.1: Zwei Parametrisierungen derselben Fläche  $A$

### 7.0.52 Satz.

Flächenintegrale  $\int_A \mathbf{F} \cdot d\mathbf{S}$ ,  $\int_A f dS$  sind unabhängig von der Parametrisierung.

*Proof.* Es sei  $\psi = \varphi \circ g$  (vgl. Abbildung 7.1). Es ist

$$\frac{\partial(u, v)}{\partial(s, t)} = \begin{pmatrix} \frac{\partial g_1}{\partial s} & \frac{\partial g_1}{\partial t} \\ \frac{\partial g_2}{\partial s} & \frac{\partial g_2}{\partial t} \end{pmatrix}.$$

Für die partiellen Ableitungen von  $\psi = \varphi \circ g$  gilt

$$\begin{aligned} \frac{\partial}{\partial s} \varphi \circ g &= \varphi_u \frac{\partial g_1}{\partial s} + \varphi_v \frac{\partial g_2}{\partial s} \\ \frac{\partial}{\partial t} \varphi \circ g &= \varphi_u \frac{\partial g_1}{\partial t} + \varphi_v \frac{\partial g_2}{\partial t}. \end{aligned}$$

Damit kann man schreiben

$$\begin{aligned}
 \int_T \mathbf{F}(\psi(s, t)) \cdot (\psi_s \times \psi_t) d(s, t) &= \int_T \mathbf{F}(\varphi \circ g) \cdot \left( \frac{\partial}{\partial s} \varphi \circ g \times \frac{\partial}{\partial t} \varphi \circ g \right) d(s, t) \\
 &= \int_T \mathbf{F}(\varphi \circ g) \cdot \left( \left( \varphi_u \frac{\partial g_1}{\partial s} + \varphi_v \frac{\partial g_2}{\partial s} \right) \times \left( \varphi_u \frac{\partial g_1}{\partial t} + \varphi_v \frac{\partial g_2}{\partial t} \right) \right) d(s, t) \\
 &= \int_T \mathbf{F}(\varphi \circ g) \cdot (\varphi_u \times \varphi_v) \left( \frac{\partial g_1}{\partial s} \frac{\partial g_2}{\partial t} - \frac{\partial g_2}{\partial s} \frac{\partial g_1}{\partial t} \right) d(s, t) \\
 &= \int_T \mathbf{F}(\varphi \circ g) \cdot (\varphi_u \times \varphi_v) \det \left( \frac{\partial(u, v)}{\partial(s, t)} \right) d(s, t) \\
 &= \int_{g(T)} \mathbf{F}(\varphi) \cdot (\varphi_u \times \varphi_v) d(u, v)
 \end{aligned}$$

□

## Integralsätze von Stokes und Gauß

### 7.0.53 Satz. Stokes

Sei  $A \subset \mathbb{R}^3$  eine Fläche mit Parametrisierung  $\varphi : K \rightarrow \mathbb{R}^3$ ,  $\varphi \in C^2(K, \mathbb{R}^3)$ .  $C$  sei die positiv orientierte Randkurve von  $K$  mit Parametrisierung  $\gamma : [a, b] \rightarrow \mathbb{R}^2$ ,  $\gamma$  (stückweise) stetig differenzierbar. Mit  $\partial A$  bezeichnen wir die Randkurve in  $\mathbb{R}^3$  mit Parametrisierung  $\varphi \circ \gamma$ . Sei  $F$  stetig differenzierbares Vektorfeld,  $F : A \rightarrow \mathbb{R}^3$ . Dann gilt:

$$\begin{aligned}
 \int_C F \cdot d\mathbf{x} &= \int_a^b \mathbf{F}(\gamma(t)) \cdot \gamma'(t) dt = \iint_A \operatorname{rot} \mathbf{F} \cdot d\mathbf{S} \\
 &= \iint_K \operatorname{rot} \mathbf{F}(\varphi(u, v)) \cdot (\varphi_u \times \varphi_v) d(u, v)
 \end{aligned}$$

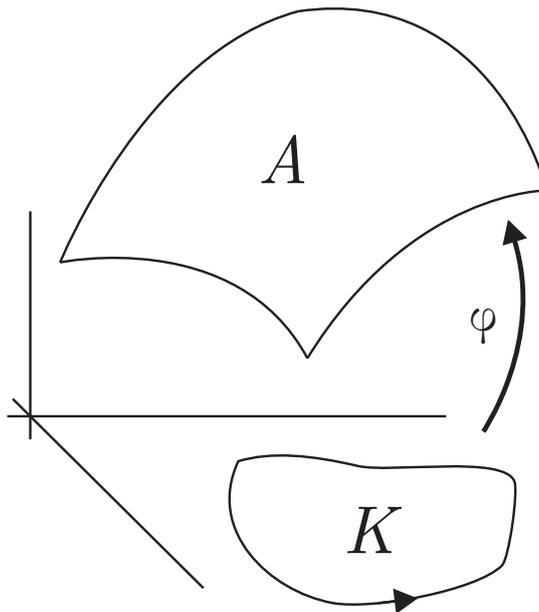


Abbildung 7.2: Zur Visualisierung des Satzes von Stokes.

*Proof.* Sei  $A$  eine Fläche,  $\partial A$  ihr Rand (der stückweise stetig differenzierbar ist). Wir werden den Integralsatz von Stokes durch Rückführung auf den Satz von Green beweisen. Sei  $\gamma(t)$  eine Parametrisierung von  $\partial K$ :  $\gamma[a, b] \rightarrow \mathbb{R}^2$ . Die zugehörige Parametrisierung von  $\partial A$  ist dann  $[a, b] \rightarrow \varphi \circ \gamma$ . Dabei gilt für die Ableitung von  $\varphi \circ \gamma$

$$\frac{d}{dt} \varphi \circ \gamma = \varphi_u \dot{\gamma}_1 + \varphi_v \dot{\gamma}_2.$$

Schreibe darum für das Kurvenintegral

$$\begin{aligned} \int_{\partial A} \mathbf{F} \cdot d\mathbf{x} &= \int_a^b \mathbf{F}(\varphi \circ \gamma) \cdot \frac{d}{dt}(\varphi \circ \gamma) dt \\ &= \int_a^b \begin{pmatrix} \mathbf{F} \cdot \varphi_u \\ \mathbf{F} \cdot \varphi_v \end{pmatrix} \cdot \frac{d}{dt} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} dt \\ &= \int_{\partial K} P dx + Q dy && \text{mit } P = \mathbf{F} \cdot \varphi_u \text{ und } Q = \mathbf{F} \cdot \varphi_v \\ &= \iint_K \underbrace{\left( \frac{\partial}{\partial u} \mathbf{F} \cdot \varphi_v - \frac{\partial}{\partial v} \mathbf{F} \cdot \varphi_u \right)}_{=\Gamma} d(u, v) \\ &\stackrel{*}{=} \iint_K \text{rot } \mathbf{F} \cdot (\varphi_u \times \varphi_v) d(u, v) \\ &= \iint_A \text{rot } \mathbf{F} \cdot d\mathbf{S}. \end{aligned}$$

Wobei \* erst noch gezeigt werden muss. Wir werden den Ausdruck  $\Gamma$  auf diese Form bringen. Betrachte zuerst die Ableitung des Skalarproduktes von  $\mathbf{F}$  und der Ableitungen von  $\varphi$ :

$$\begin{aligned} \frac{\partial}{\partial u} \sum_{i=1}^3 F^i \varphi_v^i &= \sum_{i=1}^3 \left( \frac{\partial}{\partial x_j} F^i \varphi_u^j \varphi_v^i + F^i \varphi_{vu}^i \right) \\ \frac{\partial}{\partial v} \sum_{i=1}^3 F^i \varphi_u^i &= \sum_{i=1}^3 \left( \frac{\partial}{\partial x_j} F^i \varphi_v^j \varphi_u^i + F^i \varphi_{uv}^i \right) \\ \Rightarrow \Gamma &= \sum_{i,j=1}^3 \left( \frac{\partial}{\partial x_j} F^i \varphi_u^j \varphi_v^i - \frac{\partial}{\partial x_j} F^i \varphi_v^j \varphi_u^i \right). \end{aligned}$$

Dies bezeichnet die Rotation des Vektorfeldes  $\mathbf{F}$  skalar multipliziert mit  $(\varphi_u \times \varphi_v)$ , denn es ist

$$\begin{aligned} \text{rot } \mathbf{F} \cdot (\varphi_u \times \varphi_v) &= \begin{pmatrix} \partial_2 F^3 - \partial_3 F^2 \\ \partial_3 F^1 - \partial_1 F^3 \\ \partial_1 F^2 - \partial_2 F^1 \end{pmatrix} \cdot (\varphi_u^2 \varphi_v^3 - \varphi_u^3 \varphi_v^2) \\ &= \partial_2 F^3 \varphi_u^2 \varphi_v^3 - \partial_2 F^3 \varphi_u^3 \varphi_v^2 + \dots \\ &= \sum_{i,j=1}^3 \left( \frac{\partial}{\partial x_j} F^i \varphi_u^j \varphi_v^i - \frac{\partial}{\partial x_j} F^i \varphi_v^j \varphi_u^i \right). \end{aligned}$$

□

### 7.0.54 Beispiel.

Um den Fluss der Funktion  $F = (0, 0, 2)$  durch die Oberfläche der oberen Halbkugel mit Radius 1 zu berechnen, kann man die Tatsache dass  $\nabla \times f = F$  ist, mit  $f = (y, -x, 0)$ , aber das Wegintegral von  $f$  über den Rand ist offensichtlich gleich  $2\pi$ .

### 7.0.55 Satz. Gauß

Sei  $A \subset \mathbb{R}^3$  eine Fläche mit Parametrisierung  $\varphi : K \rightarrow \mathbb{R}^3$ . Sei  $A$  geschlossene Fläche,  $V$  der von  $A$  eingeschlossene Bereich,  $\partial V = A$ . Sei  $A$  so parametrisiert, durch  $\varphi$ , dass  $\mathbf{N}$ , die Einheitsnormale, nach außen zeigt. Dann gilt

$$\begin{aligned} \iint_A \mathbf{F} \cdot d\mathbf{S} &= \iint_K \mathbf{F}(\varphi(u, v)) \cdot \mathbf{N} \|\varphi_u \times \varphi_v\| d(u, v) \\ &= \iiint_V \operatorname{div} \mathbf{F} d(x, y, z) \end{aligned}$$

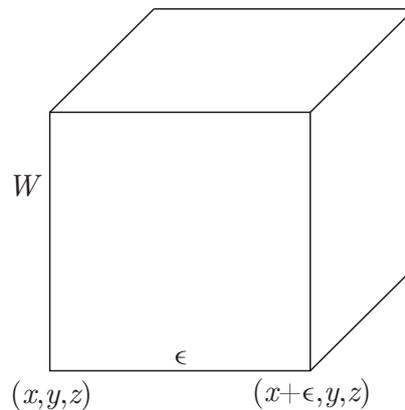


Abbildung 7.3: Ein infinitesimal kleiner Würfel mit Kantenlänge  $\epsilon$

*Proof.* Beweise den Satz von Gauß zunächst für einen infinitesimal kleinen Würfel  $W$  mit Kantenlänge  $\epsilon$  (vgl. Abbildung 7.3). Zeige, dass

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^3} \iiint_W \mathbf{F} \cdot d\mathbf{S} = \operatorname{div} \mathbf{F}$$

Betrachte dazu beispielhaft den Anteil der Oberfläche des Würfels, der senkrecht zur  $z$ -Achse steht. Das Skalarprodukt mit der Flächennormalen liefert dann die dritte Komponente des Vektorfelds  $\mathbf{F} = (F^1, F^2, F^3)$  wobei der Deckel ein positives, der Boden ein negatives Vorzeichen liefert:

$$\int_y^{y+\epsilon} \int_x^{x+\epsilon} F^3(x, y, z + \epsilon) - F^3(x, y, z) d(x, y) = \epsilon^3 \frac{\partial F^3}{\partial z} + \mathcal{O}(\epsilon^4)$$

Für die beiden Flächen senkrecht zur  $y$ - sowie zur  $x$ -Achse ergibt sich analog  $\epsilon^3 \frac{\partial F^2}{\partial y}$  und  $\epsilon^3 \frac{\partial F^1}{\partial x}$  und damit insgesamt

$$\iiint_W \mathbf{F} \cdot d\mathbf{S} = \epsilon^3 \left( \frac{\partial F^1}{\partial x} + \frac{\partial F^2}{\partial y} + \frac{\partial F^3}{\partial z} \right) + \mathcal{O}(\epsilon^4)$$

Denkt man sich nun einen größeren Würfel mit Kantenlänge 1, so kann man sich diesen aus vielen kleinen

Würfeln der Kantenlänge  $\epsilon$  zusammengesetzt vorstellen. Es ergibt sich damit

$$\begin{aligned}\iint_{W_1} \mathbf{F} \cdot d\mathbf{S} &= \lim_{\epsilon \rightarrow 0} \sum_{i=1}^{1/\epsilon^3} \iint_W \mathbf{F} \cdot d\mathbf{S} \\ &= \lim_{\epsilon \rightarrow 0} \sum_{i=1}^{1/\epsilon^3} \epsilon^3 \operatorname{div} \mathbf{F}(x_i) + \mathcal{O}(\epsilon^4) \\ &= \iiint_V \operatorname{div} \mathbf{F} d(x, y, z)\end{aligned}$$

Für Normalbereiche lässt sich der Satz von Gauß auch anders beweisen, sei dazu  $V$  Normalbereich, das heißt:

$$\begin{aligned}V &= \{(x, y, z) \mid (x, y) \in D_1 \varphi_1 \leq z \leq \varphi_2\} \\ &= \{(x, y, z) \mid (y, z) \in D_2 \psi_1 \leq x \leq \psi_2\} \\ &= \{(x, y, z) \mid (x, z) \in D_3 \Gamma_1 \leq y \leq \Gamma_2\}\end{aligned}$$

Das Vektorfeld  $\mathbf{F}$  sei wie üblich  $\mathbf{F} = (P, Q, R)$ . Man kann unter Verwendung, dass  $V$  Normalbereich ist schreiben

$$\begin{aligned}\iint_D \int_{\varphi_1(x,y)}^{\varphi_2(x,y)} R_z dz d(x, y) &= \iint_D R(x, y, \varphi_2(x, y)) - R(x, y, \varphi_1(x, y)) d(x, y) \\ &= \iint_A (0, 0, R) \cdot d\mathbf{S}\end{aligned}$$

Analog schreibt man für die anderen beiden Komponenten

$$\begin{aligned}\iint_V Q_y d(x, y, z) &= \iint_A (0, Q, 0) \cdot d\mathbf{S} \\ \iint_V P_x d(x, y, z) &= \iint_A (P, 0, 0) \cdot d\mathbf{S}\end{aligned}$$

und damit insgesamt

$$\iint_V \operatorname{div} \mathbf{F} d(x, y, z) = \iiint_V (P_x + Q_y + R_z) d(x, y, z) = \iint_A \mathbf{F} \cdot d\mathbf{S}$$

□

### 7.0.56 Beispiel.

Nehmen wir die Kraft, die von einer Punktladung herkommt,  $F(x) = q \frac{\vec{x}}{|\vec{x}|^3}$ . Dann gilt, dass der Fluss durch jede Fläche, die den Ursprung umschließt, den gleichen Wert ergibt. Das sieht man mittels Satz von Gauss und der Tatsache, dass die Divergenz von  $F$  verschwindet ausserhalb vom Ursprung, da

$$\nabla \cdot \frac{\vec{x}}{|\vec{x}|^3} = 3 \frac{1}{|\vec{x}|^3} - 3 \frac{\vec{x} \cdot \vec{x}}{|\vec{x}|^5} = 0.$$

Und der Fluss durch die Einheitssphäre ist einfach abzulesen und gleich  $q$  mal der Fläche der Sphäre, also  $q4\pi$ .

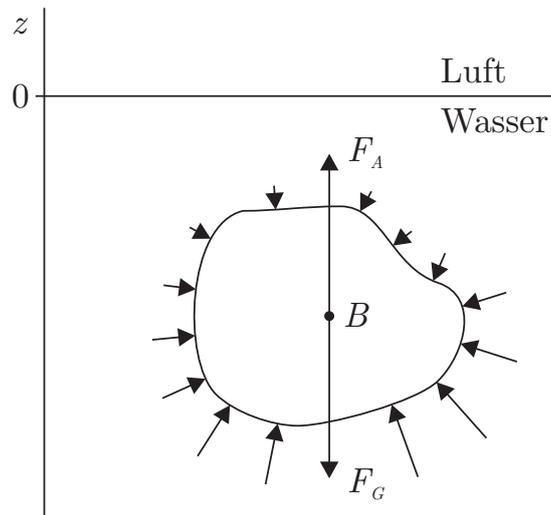


Abbildung 7.4: Der Druck bewirkt eine Auftriebskraft, da er mit der Tiefe zunimmt

## Anwendungen des Satzes von Gauß

### Archimedisches Prinzip

Das Archimedische Prinzip besagt, dass ein Körper im Wasser eine Auftriebskraft erfährt, die bestimmt wird durch die Gewichtskraft des vom Körper verdrängten Wassers.

*Proof.* Um das mit Hilfe des Satzes von Gauß zu zeigen gehe folgendermaßen vor: Aus der Physik weiß man, dass der Druck mit der Tiefe des Wassers zunimmt:

$$p = -\rho g z \quad z = 0 : \text{Wasseroberfläche.}$$

Mit dem Satz von Gauß kann man zeigen, dass für eine skalare Funktion  $f$  gilt

$$\int_{\partial B} f d\mathbf{S} = \int_B \nabla f dV,$$

denn schreibt man den Normalenvektor  $\mathbf{N}$  auf die folgende Weise

$$\mathbf{N} = \sum_{i=1}^3 N_i \mathbf{e}_i, \quad N_i = \langle \mathbf{N}, \mathbf{e}_i \rangle$$

mit der Standardbasis  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ , so kann man für das Oberflächenintegral schreiben

$$\begin{aligned} \int_{\partial B} f d\mathbf{S} &= \sum_{i=1}^3 \mathbf{e}_i \int_{\partial B} \langle \mathbf{N}, f \mathbf{e}_i \rangle dS \\ &= \sum_{i=1}^3 \mathbf{e}_i \int_B \operatorname{div}(f \mathbf{e}_i) dV \\ &= \int_B \operatorname{grad} f dV \\ &= \int_B \nabla f dV. \end{aligned}$$

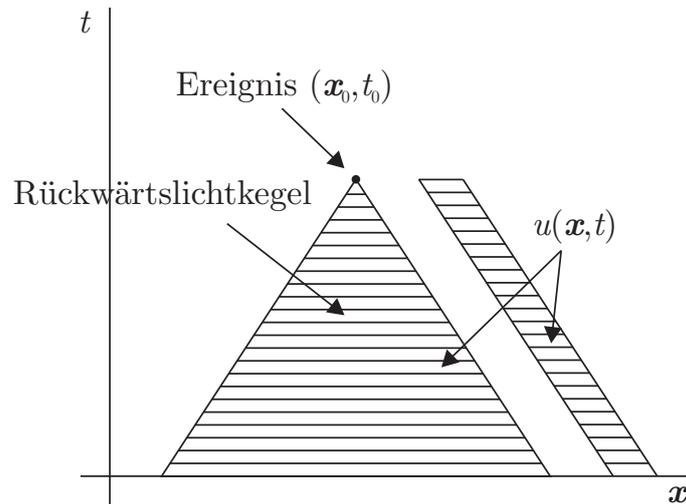


Abbildung 7.5: Der Teil außerhalb des Rückwärtslichtkegels kann das Ereignis  $(\mathbf{x}_0, t_0)$  nicht beeinflussen, da Energie und somit Information nicht schneller als Licht übertragen werden kann.

Wendet man dies an auf  $f = -p$  so erhält man für die Auftriebskraft

$$\begin{aligned}
 \mathbf{F}_A &= - \int_{\partial B} p d\mathbf{S} \\
 &= - \int_B \nabla p dV \\
 &= \rho g \mathbf{e}_3 \int_B dV \\
 &= \rho g \mathbf{e}_3 \text{vol}(B).
 \end{aligned}$$

□

### Kausalitätsprinzip für Lösungen der Wellengleichung

Die allgemeine Wellengleichung lautet

$$\frac{\partial^2}{\partial t^2} u = c^2 \Delta_x u$$

mit  $u(\mathbf{x}, t)$  2-mal stetig differenzierbar. In einer Dimension haben allgemeine Lösungen die Form

$$u(x, t) = f(x - ct) + g(x + ct) \quad f, g \in C^2(\mathbb{R}, \mathbb{R}).$$

Was wir zeigen möchten ist, dass nur der Teil einer Welle  $u(\mathbf{x}, t)$ , der im Lichtkegel liegt, das Ereignis  $(\mathbf{x}_0, t_0)$  beeinflussen kann (vgl. Abbildung 7.5). Die mathematische Formulierung dafür lautet: Sind  $u_1, u_2$  Lösungen der Wellengleichung und ist für ein  $t < t_0$

$$\left. \begin{aligned}
 u_1(\mathbf{x}, t) &= u_2(\mathbf{x}, t) \\
 \dot{u}_1(\mathbf{x}, t) &= \dot{u}_2(\mathbf{x}, t)
 \end{aligned} \right\} \forall \mathbf{x} \in K_t$$

mit dem Boden des Rückwärtslichtkegels

$$K_t := \{\mathbf{x} \in \mathbb{R}^3 \mid \|\mathbf{x} - \mathbf{x}_0\| \leq c(t_0 - t)\},$$

dann gilt

$$u_1(\mathbf{x}_0, t_0) = u_2(\mathbf{x}_0, t_0).$$

*Proof.* Die Idee ist, zu zeigen, dass aus dem Rückwärtslichtkegel nur Energie entweichen kann (siehe auch Abbildung 7.5) d.h., dass innerhalb des Kegels die Energie mit der Zeit höchstens konstant bleiben kann und somit die Energiedichte bei der Spitze ( $t = t_0$ ) verschwinden muss, wenn sie schon am Boden ( $t < t_0$ ) verschwindet.

**Schritt 1 (Kontinuitätsgleichung)** Sei die Energiedichte definiert durch

$$\epsilon_t(u) := \dot{u}^2 + c^2 \|\nabla_x u\|^2$$

mit dem räumlichen Gradienten

$$\nabla_x u = \left( \frac{\partial}{\partial x_1} u, \frac{\partial}{\partial x_2} u, \frac{\partial}{\partial x_3} u \right).$$

Es sei

$$u_t(\mathbf{x}) := u(\mathbf{x}, t).$$

Der Träger einer skalaren Funktion  $f$  ist

$$\text{supp}(f) := \overline{\{\mathbf{x} \in \mathbb{R}^3 \mid f(\mathbf{x}) \neq 0\}}.$$

Existiert zu einer Dichtefunktion  $\rho$  und einer Stromdichte  $\mathbf{j}$  eine Kontinuitätsgleichung

$$\dot{\rho} + \text{div } \mathbf{j} = 0,$$

so resultiert die Änderung der in einem Volumen  $D \subset \mathbb{R}^3$  enthaltene Größe  $Q_D = \int_D \rho dV$  aus dem Strom, der die Oberfläche durchfließt:

$$\frac{d}{dt} Q_D = \int_D \dot{\rho} dV = - \int_{\partial D} \mathbf{j} \cdot d\mathbf{S}.$$

Dies ist eine direkte Konsequenz des Satzes von Gauß und man spricht von der (globalen) Erhaltung der Größe  $Q$ . Solange die Stromdichte  $\mathbf{j}$  auf ein kompaktes Gebiet  $\bar{D} \subset \mathbb{R}^3$  beschränkt bleibt, d.h.  $\text{supp}(\mathbf{j}) \subset \bar{D}$ , ist die Größe  $Q$  nicht nur global, sondern auch lokal innerhalb des Volumens  $\bar{D}$  erhalten:

$$\frac{d}{dt} Q_{\bar{D}} = - \int_{\partial \bar{D}} \mathbf{j} \cdot d\mathbf{S} = 0.$$

Dabei wurde verwendet, dass  $\mathbf{j}$  da stetig, auf dem Rand von  $\bar{D}$  identisch 0 sein muss (wenn  $\mathbf{j}$  es außerhalb von  $D$  auch ist).

In unserem Fall findet man die Kontinuitätsgleichung

$$\dot{\epsilon}_t + \text{div}_x \mathbf{j}_E = 0$$

für die Energiedichte  $\epsilon_t$  und die Energiestromdichte  $\mathbf{j}_E = -2c^2 \dot{u} \nabla_x u$ , denn

$$\begin{aligned} 2c^2 \text{div}_x (\dot{u} \nabla_x u) &= 2c^2 (\langle \nabla_x \dot{u}, \nabla_x u \rangle + \dot{u} \Delta_x u) \\ &= 2c^2 \langle \nabla_x \dot{u}, \nabla_x u \rangle + 2\dot{u} \ddot{u} \\ &= \dot{\epsilon}_t(u). \end{aligned}$$

Dabei wurde die Wellengleichung verwendet, sowie

$$\|\nabla_x u\|^2 = \langle \nabla_x u, \nabla_x u \rangle.$$

Das obige Argument lässt sich hier also anwenden und es folgt die Erhaltung der Energie  $E_t[u] = \int \epsilon_t(u) dV$ .

**Schritt 2 (Energieerhaltung)** Sei  $\Omega \subset \mathbb{R}^3 \times \mathbb{R}$  offen mit Rand  $\partial\Omega$ . Dann ist

$$\int_{\partial\Omega} \begin{pmatrix} -2c^2 \dot{u} \nabla_x u \\ \epsilon_t(u) \end{pmatrix} \cdot d\mathbf{S} = \int_{\Omega} \text{div} \begin{pmatrix} -2c^2 \dot{u} \nabla_x u \\ \epsilon_t(u) \end{pmatrix} dV = 0,$$

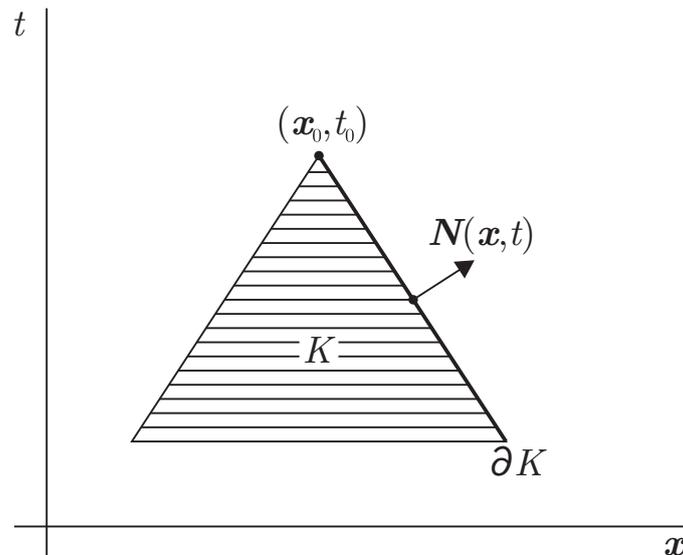


Abbildung 7.6: Der Normalenvektor steht senkrecht auf dem Rand des Rückwärtslichtkegels

denn es ist nach Schritt 1

$$\operatorname{div} \begin{pmatrix} -2c^2 \dot{u} \nabla_x u \\ \epsilon_t(u) \end{pmatrix} = \dot{\epsilon}_t(u) - 2c^2 \operatorname{div}_x(\dot{u} \nabla_x u) = 0.$$

**Schritt 3 (Anwendung auf Kegel Teil 1)** Sei  $(\mathbf{x}_0, t_0) \in \mathbb{R}^3 \times \mathbb{R}$  Dann gilt für die Einheitsnormale

$$\mathbf{N} = \begin{pmatrix} N_1 \\ N_0 \end{pmatrix}$$

des Rückwärtslichtkegels

$$K := \{(\mathbf{x}, t) \in \mathbb{R}^3 \times \mathbb{R} \mid \|\mathbf{x}_0 - \mathbf{x}\| \leq c|t - t_0|\}$$

folgende Gleichung

$$N_0^2 = c^2 \|\mathbf{N}_1\|^2.$$

Argumentation: Der Rand von  $K$  ist gegeben durch  $\|\mathbf{x}_0 - \mathbf{x}\| = c|t - t_0|$  und somit Teil der Niveaumenge  $h(\mathbf{x}, t) = \|\mathbf{x}_0 - \mathbf{x}\|^2 - c^2|t - t_0|^2 = 0$ , vgl. Abbildung 7.6. Also ist die Einheitsnormale parallel zum Gradienten dieser Niveaumenge:

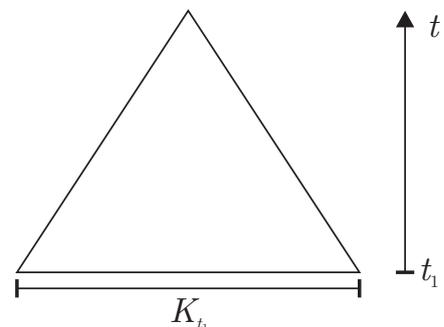
$$\mathbf{N} \parallel \nabla h \quad \text{oder} \quad \mathbf{N} = \lambda \begin{pmatrix} \mathbf{x}_0 - \mathbf{x} \\ c^2|t - t_0| \end{pmatrix}.$$

Also kann man schreiben

$$\begin{aligned} c^2 \|\mathbf{N}_1\|^2 &= \lambda^2 c^2 \|\mathbf{x}_0 - \mathbf{x}\|^2 \\ &= \lambda^2 c^4 |t - t_0|^2 \\ &= N_0^2. \end{aligned}$$

**Schritt 4 (Anwendung auf Kegel Teil 2)** Sei  $t < t_0$  und der Rückwärtslichtkegel wie vorhin

$$K_t := \{\mathbf{x} \in \mathbb{R}^3 \mid \|\mathbf{x} - \mathbf{x}_0\| \leq c|t - t_0|\}.$$



Für  $t_1 < t_2 < t$  gilt dann

$$\int_{K_{t_2} \times \{t_2\}} \epsilon_t(u) dV - \int_{K_{t_1} \times \{t_1\}} \epsilon_t(u) dV \leq 0.$$

Denn: Wendet man Schritt 2 an auf

$$\Omega = \{(\mathbf{x}, t) \in K \mid t_1 \leq t \leq t_2\}$$

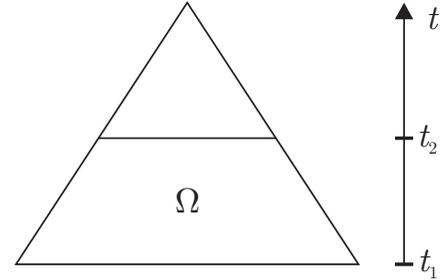


Abbildung 8

erhält man

$$\begin{aligned} 0 &= \int_{\partial\Omega} \begin{pmatrix} -2c^2\dot{u}\nabla_x u \\ \epsilon_t(u) \end{pmatrix} \cdot d\mathbf{S} \\ &= \int_{K_{t_2} \times \{t_2\}} \epsilon_t(u) dV - \int_{K_{t_1} \times \{t_1\}} \epsilon_t(u) dV + \int_M \langle \begin{pmatrix} -2c^2\dot{u}\nabla_x u \\ \epsilon_t(u) \end{pmatrix}, \mathbf{N} \rangle dS \end{aligned}$$

mit der Mantelfläche  $M$  von  $\Omega$  sowie der Tatsache, dass auf dem Deckel von  $\Omega$  die Einheitsnormale gegeben ist durch  $\mathbf{N} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$  beziehungsweise auf dem Boden von  $\Omega$  durch  $\mathbf{N} = \begin{pmatrix} \mathbf{0} \\ -1 \end{pmatrix}$ . Man muss nun noch zeigen, dass das hintere Integral größer oder gleich Null ist:

$$\begin{aligned} \int_M \langle \begin{pmatrix} -2c^2\dot{u}\nabla_x u \\ \epsilon_t(u) \end{pmatrix}, \mathbf{N} \rangle dS &= \int_M -\langle 2c^2\dot{u}\nabla_x u, \mathbf{N}_1 \rangle + N_0 \epsilon_t(u) dV \\ &\geq \int_M N_0 (-2c\dot{u}\|\nabla_x u\| + \dot{u}^2 + c^2\|\nabla_x u\|^2) dS \\ &= \int_M N_0 (c\|\nabla_x u\| - \dot{u})^2 dS \\ &\geq 0. \end{aligned}$$

Denn es gilt nach Schritt 3:

$$-\langle 2c^2\dot{u}\nabla_x u, \mathbf{N}_1 \rangle \geq -\|\nabla_x u\| \|\mathbf{N}_1\| 2c^2\dot{u} = -\|\nabla_x u\| N_0 2c\dot{u}.$$

**Schritt 5 (Energieabnahme im Rückwärtslichtkegel)** Falls für ein  $t < t_0$  gilt, dass

$$\left. \begin{aligned} u(\mathbf{x}, t) &= 0 \\ \dot{u}(\mathbf{x}, t) &= 0 \end{aligned} \right\} \forall \mathbf{x} \in K_t,$$

so gilt  $u(\mathbf{x}_0, t_0) = 0$ . Denn nach Schritt 4 ist

$$0 = \int_{K_t \times \{t\}} \epsilon_t(u) dV \geq \int_{K_{t_1} \times \{t_1\}} \epsilon_t(u) dV \geq 0$$

da die Energiedichte  $\epsilon_t(u)$  positiv ist. Es folgt

$$\Rightarrow \int_{K_{t_1} \times \{t_1\}} \epsilon_t(u) dV = 0.$$

Damit ist  $\epsilon_t(u) \equiv 0$  auf  $K_{t_1} \times \{t_1\}$  und auch  $\epsilon_t(u) \equiv 0$  auf  $\overset{\circ}{K}$  (Inneres von  $K$ ).

$$\Rightarrow u \equiv 0 \text{ auf } \overset{\circ}{K}.$$

Da  $u$  stetig ist

$$\Rightarrow u \equiv 0 \text{ auf } K \text{ insbesondere auch für } (\mathbf{x}_0, t_0).$$

**Schritt 6 (Superposition)** Wenn gilt, dass

$$\begin{aligned} u_1(\mathbf{x}, t) &\equiv u_2(\mathbf{x}, t) \\ \dot{u}_1(\mathbf{x}, t) &\equiv \dot{u}_2(\mathbf{x}, t) \end{aligned} \text{ auf } K_t \times \{t\},$$

dann gilt

$$u_1(\mathbf{x}_0, t_0) = u_2(\mathbf{x}_0, t_0).$$

Argumentation: Eine Superposition  $u := u_1 - u_2$  der beiden Lösungen ist wieder eine Lösung der Wellengleichung und erfüllt die Voraussetzungen von Schritt 5 und es gilt

$$\begin{aligned} u(\mathbf{x}_0, t_0) &= u_1(\mathbf{x}_0, t_0) - u_2(\mathbf{x}_0, t_0) = 0 \\ \Rightarrow u_1(\mathbf{x}_0, t_0) &= u_2(\mathbf{x}_0, t_0). \end{aligned}$$

□

### Transporttheorem (Fluiddynamik)

Die Stromlinien sind durch  $\Phi^t(\mathbf{x})$  gegeben wobei  $\Phi^t : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  ein Diffeomorphismus ist. Das zugehörige Geschwindigkeitsvektorfeld ist  $\mathbf{v}(\Phi^t(\mathbf{x})) = \frac{\partial}{\partial t} \Phi^t(\mathbf{x})$ . Das Transporttheorem ist

$$\frac{d}{dt} \int_{\Phi^t(B)} f(\mathbf{y}, t) dV = \int_{\Phi^t(B)} \frac{\partial}{\partial t} f(\mathbf{y}, t) dV + \int_{\partial \Phi^t(B)} f(\mathbf{y}, t) \mathbf{v}(\mathbf{y}) \cdot d\mathbf{S}.$$

Die Funktion  $f$  kann dabei beispielsweise eine Massendichte oder Energiedichte oder einfach  $f \equiv 1$  sein.

*Proof.* Lemma: Aus der Linearen Algebra weiß man, dass gilt

$$\frac{d}{dt} \det(A(t)) = \text{tr}(\dot{A}(t)A^{-1}(t)) \det(A(t)).$$

Auf den Beweis dafür wird hier verzichtet. Man kann also schreiben

$$\begin{aligned} \partial_t |\det(D\Phi^t)| &= \text{sgn}(\det(D\Phi^t)) \text{tr}(\partial_t D\Phi^t (D\Phi^t)^{-1}) \det(D\Phi^t) \\ &= |\det(D\Phi^t)| \text{tr}(D\dot{\Phi}^t D(\Phi^{-t}) \circ \Phi^t) \\ &= |\det(D\Phi^t)| \text{tr}(D(\dot{\Phi}^t \circ \Phi^{-t}) \circ \Phi^t) \\ &= |\det(D\Phi^t)| \text{div}(\dot{\Phi}^t \circ \Phi^{-t}) \circ \Phi^t. \end{aligned}$$

Schreibe also

$$\begin{aligned}
 \frac{d}{dt} \int_{\Phi^t(B)} f(\mathbf{x}, t) dV &= \frac{d}{dt} \int_B f(\Phi^t(\mathbf{x}), t) |\det(D\Phi^t)| dV \\
 &= \int_B \left[ \frac{\partial f}{\partial t}(\Phi^t(\mathbf{x}), t) + \langle \nabla_{\mathbf{x}} f \circ \Phi^t, \dot{\Phi}^t \rangle + f(\Phi^t(\mathbf{x}), t) \operatorname{div}(\dot{\Phi}^t \circ \Phi^{-t}) \circ \Phi^t \right] |\det(D\Phi^t)| dV \\
 &= \int_{\Phi^t(B)} \left[ \frac{\partial f}{\partial t}(\mathbf{y}, t) + \langle \nabla_{\mathbf{x}} f, \dot{\Phi}^t \circ \Phi^{-t} \rangle + f(\mathbf{y}, t) \operatorname{div}(\dot{\Phi}^t \circ \Phi^{-t}) \right] dV \\
 &= \int_{\Phi^t(B)} \left[ \frac{\partial f}{\partial t}(\mathbf{y}, t) + \operatorname{div}(f(\mathbf{y}, t) \dot{\Phi}^t \circ \Phi^{-t}) \right] dV \\
 &= \int_{\Phi^t(B)} \frac{\partial f}{\partial t}(\mathbf{y}, t) dV + \int_{\partial\Phi^t(B)} f(\mathbf{y}, t) \dot{\Phi}^t \circ \Phi^{-t} \cdot d\mathbf{S} \\
 &= \int_{\Phi^t(B)} \frac{\partial f}{\partial t}(\mathbf{y}, t) dV + \int_{\partial\Phi^t(B)} f(\mathbf{y}, t) \mathbf{v}(\mathbf{y}) \cdot d\mathbf{S}.
 \end{aligned}$$

□

Zur Anwendung: Man beobachtet bei verschiedenen Körpern den selben Zusammenhang zwischen Volumen und Oberfläche

	Volumen		Oberfläche
Kugel	$\frac{4}{3}\pi r^3$	$\frac{\partial r}{\partial t}$	$4\pi r^2$
Torus	$2\pi^2 r^2 R$	$\frac{\partial r}{\partial t}$	$4\pi r^2 R$
Kreisfläche	$\pi r^2$	$\frac{\partial r}{\partial t}$	$2\pi r$

Die Frage ist: wann funktioniert das? Die Antwort liefert das Transporttheorem mit  $f \equiv 1$ , es funktioniert wenn ein Referenzkörper  $B \subset \mathbb{R}^n$  und wie oben ein Diffeomorphismus  $\Phi^r$  existiert, so dass der zu untersuchende Körper gerade  $\Phi^r(B)$  ist und das zugehörige Skalarprodukt  $\langle \mathbf{v}(\mathbf{y}), \mathbf{N}(\mathbf{y}) \rangle \equiv 1$  auf  $\partial\Phi^r(B)$ . Denn dann ist

$$\int_{\partial\Phi^r(B)} f(\mathbf{y}, t) \mathbf{v}(\mathbf{y}) \cdot d\mathbf{S} = \int_{\partial\Phi^r(B)} dS = \operatorname{area}(\partial\Phi^r(B)).$$

Zum Beispiel ist dies gegeben für eine Kugel mit Radius 1 um den Ursprung:  $B = B_1(0)$ . Der Fluss ist  $\Phi^r(\mathbf{x}) = r\mathbf{x} = \mathbf{y}$  und damit das Vektorfeld  $\mathbf{v}(\mathbf{y}) = \mathbf{x}$ . Es zeigt immer in die selbe Richtung wie der Normalenvektor auf der Kugeloberfläche, also ist das Skalarprodukt  $\langle \mathbf{v}(\mathbf{y}), \mathbf{N}(\mathbf{y}) \rangle \equiv 1$  und damit

$$\frac{d}{dr} \int_{\Phi^r(B)} f(\mathbf{y}, t) dV = \frac{d}{dr} \operatorname{vol}(\Phi^r(B)) = \int_{\partial\Phi^r(B)} \mathbf{x} \cdot d\mathbf{S} = \int_{\partial\Phi^r(B)} dS = \operatorname{area}(\partial\Phi^r(B)).$$