

Skript zur Vorlesung

Analysis II

Sommersemester 2016

Kombinations-Bachelorstudiengang Mathematik mit
Lehramtsbezug/-option

Dozent: Prof. Dr. Andreas Filler

Dieses Skript wurde von Prof. Dr. Helga Baum erarbeitet und im Sommersemester 2015 eingesetzt. Es wurde für das Sommersemester 2016 nur geringfügig modifiziert. Der Aufbau der Vorlesung kann an einigen Stellen von diesem Skript abweichen.

Inhaltsverzeichnis

6	Integralrechnung für Funktionen einer reellen Variablen	1
6.1	Das Riemann-Integral	1
6.2	Integration und Differentiation	10
6.3	Uneigentliche Riemann-Integrale	27
6.4	Geometrische Anwendungen des Riemann-Integrals	30
6.4.1	Parametrisierte Kurven und ihre Länge	31
6.4.2	Flächeninhalt ebener Gebiete	37
6.5	Analytische Anwendung: Differentialgleichungen mit getrennten Variablen .	42
7	Metrische Räume und stetige Abbildungen	51
7.1	Definition und Beispiele metrischer Räume	51
7.2	Konvergente Folgen und Cauchy-Folgen in metrischen Räumen	55
7.3	Spezielle Teilmengen in metrischen Räumen	59
7.4	Stetige Abbildungen und ihre Eigenschaften	64
8	Differentialrechnung für Funktionen mehrerer reeller Variablen	73
8.1	Differenzierbare Abbildungen, das Differential, die Richtungsableitungen und die partiellen Ableitungen	74
8.2	Kettenregeln und Mittelwertsätze	82
8.3	Die Jacobi-Matrix, der Gradient und die Hesse-Matrix	87
8.4	Die Taylorformel für Funktionen mehrerer reeller Variablen	92
8.5	Lokale Extrema für Funktionen mehrerer reeller Variablen	95
8.6	Koordinatentransformationen und der Satz über den lokalen Diffeomorphismus	98
8.7	Auflösen von Gleichungen. Der Satz über implizite Funktionen	105
8.8	Gleichungsdefinierte Flächen, ihre Tangentialebenen und Normalen	109
9	Integralrechnung für Funktionen mehrerer reeller Variablen (Ein Ausblick)	113
9.1	Das Problem der Volumendefinition für Teilmengen des \mathbb{R}^n	113
9.2	Das Jordan-Volumen	115

9.3	Das Riemann-Integral für Funktionen mehrerer reeller Variablen	117
9.4	Der Satz von Fubini und das Prinzip von Cavalieri	121
9.5	Die Transformationsformel für Riemann-Integrale	127
	Sachverzeichnis	131

Integralrechnung für Funktionen einer reellen Variablen

In diesem Kapitel beginnen wir mit der Integralrechnung. Die Integralrechnung wird durch zwei verschiedene Problemstellungen motiviert:

1. *Bestimmung von Längen, Flächeninhalten und Volumnen.*

Die Integralrechnung stellt Methoden bereit, mit denen man geometrische Größen wie Längen, Flächeninhalte und Volumen berechnen kann.

2. *Lösung von Differentialgleichungen.*

Viele Prozesse in der Natur werden durch Differentialgleichungen modelliert. Die Aufgabe besteht dann immer darin, aus gegebenen Ableitungen einer Funktion die Funktion selbst zu bestimmen. Dazu muß man den Prozeß des Differenzierens "umkehren". Auch dabei hilft die Integralrechnung.

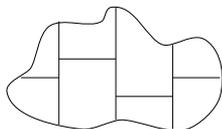
6.1 Das Riemann-Integral

Die erste Motivation für die Entwicklung der Integralrechnung entstand aus dem Versuch, Flächeninhalte und Volumen zu berechnen. Wir werden hier zunächst das Riemann-Integral für reelle Funktionen einer Variablen behandeln, mit dessen Hilfe man Flächeninhalte für Teilmengen des \mathbb{R}^2 berechnen kann. In Kapitel ?? werden wir dann die Integralrechnung für Funktionen mehrerer reeller Variablen behandeln, mit deren Hilfe man Volumen für Teilmengen des \mathbb{R}^3 berechnen kann. Das heute so genannte Riemann-Integral wurde in rigoroser Form von Bernard Riemann in seiner Habilitationsschrift (1854) definiert. Für einen kurzen historischen Überblick zur Entwicklung der Integralrechnung verweisen wir auf das 9. Kapitel des Buches *Analysis I* von W. Walter.

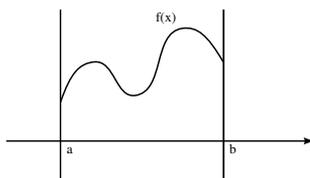
Zunächst formulieren wir einige Forderungen, die ein vernünftig definierter geometrischer Flächeninhalt $\mu(A)$ für Teilmengen $A \subset \mathbb{R}^2$ erfüllen sollte:

1. $\mu(A) \geq 0$.
2. $A \subset B \implies \mu(A) \leq \mu(B)$.
3. $\mu([a, b] \times [c, d]) = (b - a) \cdot (d - c)$.
4. Haben zwei Teilmengen $A, B \subset \mathbb{R}^2$ höchstens Randpunkte gemeinsam, so gilt $\mu(A \cup B) = \mu(A) + \mu(B)$.
5. $\mu(A)$ ist invariant gegen Euklidische Bewegungen (z.B. Drehungen, Verschiebungen, Spiegelungen).

Legt man diese Eigenschaften zugrunde, so kann man den Flächeninhalt einer Teilmenge $A \subset \mathbb{R}^2$ bestimmen, in dem man sie in Mengen "einfacher" Form zerlegt. Z.B. kann man Mengen der Form



in Teile der Form



zerlegen, wobei die obere Begrenzungskurve durch einen Funktionsgraphen gegeben wird. Es genügt deshalb, den Flächeninhalt von Mengen der Form

$$A = \{(x, y) \in \mathbb{R}^2 \mid x \in [a, b], 0 \leq y \leq f(x)\},$$

zu kennen, wobei f eine nicht-negative Funktion auf dem Intervall $[a, b]$ ist. Dies motiviert die jetzt folgende Definition des so-genannten Riemann-Integrals.

Definition 6.1. Sei $I = [a, b] \subset \mathbb{R}$ ein abgeschlossenes Intervall mit $a < b$.

Eine Menge von Punkten $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ mit $a = x_0 < \dots < x_n = b$ heißt Zerlegung von I . $I_k := [x_{k-1}, x_k]$ bezeichnet das k -te Teilintervall von \mathcal{P} und $L(I_k) := x_k - x_{k-1}$ seine Länge. $\|\mathcal{P}\| := \max\{L(I_k) \mid k = 1, \dots, n\}$ heißt die Feinheit der Zerlegung \mathcal{P} .

Eine Zerlegung $\tilde{\mathcal{P}}$ heißt Verfeinerung der Zerlegung \mathcal{P} (symbolisch: $\tilde{\mathcal{P}} \geq \mathcal{P}$), falls $\tilde{\mathcal{P}} = \mathcal{P}$ oder $\tilde{\mathcal{P}}$ aus \mathcal{P} durch Hinzunahme weiterer Teilungspunkte entsteht.

Definition 6.2. Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion, $\mathcal{P} = \{x_0, \dots, x_n\}$ eine Zerlegung von $[a, b]$ und $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ ein Tupel von Zahlen mit $\xi_k \in I_k$ für $k = 1, \dots, n$. Des Weiteren bezeichne

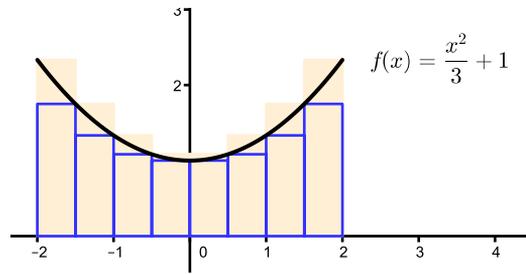
$$m_k := \inf\{f(x) \mid x \in I_k\} \quad \text{und} \quad M_k := \sup\{f(x) \mid x \in I_k\}.$$

Dann heißt

$$\underline{S}(f, \mathcal{P}) := \sum_{k=1}^n m_k \cdot L(I_k) \quad \text{Untersumme von } f \text{ bzgl. } \mathcal{P},$$

$$\overline{S}(f, \mathcal{P}) := \sum_{k=1}^n M_k \cdot L(I_k) \quad \text{Obersumme von } f \text{ bzgl. } \mathcal{P},$$

$$S(f, \mathcal{P}, \xi) := \sum_{k=1}^n f(\xi_k) \cdot L(I_k) \quad \text{Riemannsche Summe für } f \text{ bzgl. der Zerlegung } \mathcal{P} \text{ und der Stützstellen } \xi = (\xi_1, \dots, \xi_n).$$



Satz 6.1 Mit den obigen Bezeichnungen gilt:

1. Sind ξ beliebige Stützstellen von \mathcal{P} , so gilt $\underline{S}(f, \mathcal{P}) \leq S(f, \mathcal{P}, \xi) \leq \overline{S}(f, \mathcal{P})$.
2. Für $\tilde{\mathcal{P}} \geq \mathcal{P}$ gilt

$$\inf_{x \in [a, b]} f(x) \cdot (b-a) \leq \underline{S}(f, \mathcal{P}) \leq \underline{S}(f, \tilde{\mathcal{P}}) \leq \overline{S}(f, \tilde{\mathcal{P}}) \leq \overline{S}(f, \mathcal{P}) \leq \sup_{x \in [a, b]} f(x) \cdot (b-a).$$

3. Für beliebige Zerlegungen \mathcal{P} und \mathcal{P}^* gilt $\underline{S}(f, \mathcal{P}) \leq \overline{S}(f, \mathcal{P}^*)$.

Beweis. Die ersten beiden Aussagen folgen unmittelbar aus der Definition. Für die 3. Aussage betrachten wir die gemeinsame Unterteilung $\tilde{\mathcal{P}} := \mathcal{P} \cup \mathcal{P}^*$, die eine Verfeinerung von \mathcal{P} sowie von \mathcal{P}^* ist, und wenden 2. darauf an. \square

Definition 6.3. Wir nennen

$$\int_a^b f(x) dx := \sup \{ \underline{S}(f, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } [a, b] \} \quad \text{unteres Riemann Integral von } f \text{ über } [a, b],$$

$$\int_a^b f(x) dx := \inf \{ \overline{S}(f, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } [a, b] \} \quad \text{oberes Riemann-Integral von } f \text{ über } [a, b].$$

Aus Satz 6.1 folgt, dass das obere und das untere Riemann-Integral existieren und dass gilt:

$$\int_a^b f(x) dx \leq \int_a^b f(x) dx.$$

Definition 6.4. Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt über $[a, b]$ Riemann-integrierbar, wenn sie beschränkt ist und das obere und das untere Riemann-Integral von f über $[a, b]$ übereinstimmen. In diesem Fall nennt man

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

das Riemann-Integral von f über $[a, b]$.

$\mathcal{R}([a, b], \mathbb{R})$ bezeichnet die Menge der Riemann-integrierbaren Funktionen von $[a, b]$ in \mathbb{R} .

Wir setzen außerdem:

$$\int_a^a f(x) dx := 0 \quad \text{und}$$

$$\int_b^a f(x) dx := - \int_a^b f(x) dx \quad \text{falls } f \in \mathcal{R}([a, b], \mathbb{R}).$$

Satz 6.2 (Riemannsches Integrierbarkeitskriterium)

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion. Dann gilt:

f ist genau dann über $[a, b]$ Riemann-integrierbar, wenn zu jedem $\varepsilon > 0$ eine Zerlegung $\mathcal{P}(\varepsilon)$ von $[a, b]$ existiert, so dass

$$\overline{S}(f, \mathcal{P}(\varepsilon)) - \underline{S}(f, \mathcal{P}(\varepsilon)) < \varepsilon.$$

Beweis. (\Leftarrow): Sei $n \in \mathbb{N}$ und \mathcal{P}_n eine Zerlegung von $[a, b]$ mit $\overline{S}(f, \mathcal{P}_n) - \underline{S}(f, \mathcal{P}_n) < \frac{1}{n}$. Dann gilt

$$0 \leq \int_a^b f(x) dx - \int_a^b f(x) dx \leq \overline{S}(f, \mathcal{P}_n) - \underline{S}(f, \mathcal{P}_n) \leq \frac{1}{n}.$$

Der Grenzübergang $n \rightarrow \infty$ liefert die Behauptung (Sandwich-Lemma).

(\Rightarrow): Sei $\varepsilon > 0$ gegeben. Nach Definition des Supremums und des Infimums gibt es Zerlegungen \mathcal{P} und $\tilde{\mathcal{P}}$ von $[a, b]$, so dass

$$\int_a^b f(x) dx - \underline{S}(f, \mathcal{P}) < \frac{\varepsilon}{2} \quad \text{und} \quad \overline{S}(f, \tilde{\mathcal{P}}) - \int_a^b f(x) dx < \frac{\varepsilon}{2}. \quad (6.1)$$

Für die Zerlegung $\mathcal{P}^* := \mathcal{P} \cup \tilde{\mathcal{P}}$ gilt dann wegen der Riemann-Integrierbarkeit von f

$$\begin{aligned} \overline{S}(f, \mathcal{P}^*) - \underline{S}(f, \mathcal{P}^*) &\leq \overline{S}(f, \tilde{\mathcal{P}}) - \underline{S}(f, \mathcal{P}) \\ &\leq \left(\overline{S}(f, \tilde{\mathcal{P}}) - \int_a^b f(x) dx \right) + \left(\int_a^b f(x) dx - \underline{S}(f, \mathcal{P}) \right) \\ &\stackrel{(6.1)}{<} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned} \quad \square$$

Satz 6.3 Jede monotone Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist Riemann-integrierbar.

Beweis. Wir beweisen die Behauptung für monoton wachsende Funktionen. Den Beweis für monoton fallende Funktionen führt man analog.

Da f monoton wachsend ist, gilt $f(a) \leq f(x) \leq f(b)$ für alle $x \in [a, b]$, d.h. f ist beschränkt. Wir prüfen nun das Riemannsche Integrierbarkeitskriterium. Sei $n \in \mathbb{N}$. Wir wählen die äquidistante Zerlegung $\mathcal{P}_n := \{x_0, \dots, x_n\}$ von $[a, b]$ mit $x_k := a + k \cdot \frac{b-a}{n}$. Dann gilt $L(I_k) = x_k - x_{k-1} = \frac{b-a}{n} = \|\mathcal{P}_n\|$ für alle $k = 1, \dots, n$. Da f monoton wachsend ist, erhalten wir

$$\begin{aligned}
\overline{S}(f, \mathcal{P}_n) - \underline{S}(f, \mathcal{P}_n) &= \sum_{k=1}^n f(x_k) L(I_k) - \sum_{k=1}^n f(x_{k-1}) L(I_k) \\
&= \left(\sum_{k=1}^n f(x_k) - f(x_{k-1}) \right) \frac{b-a}{n} \\
&= (f(b) - f(a)) \cdot \frac{b-a}{n}.
\end{aligned}$$

Für $\varepsilon > 0$ existiert ein $n \in \mathbb{N}$ so dass $(f(b) - f(a)) \frac{b-a}{n} < \varepsilon$. Somit folgt die Behauptung aus dem Riemannschem Integrierbarkeitskriterium. \square

Satz 6.4 *Jede stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist Riemann-integrierbar.*

Beweis. Da f stetig und $[a, b]$ kompakt ist, ist f beschränkt. Für $n \in \mathbb{N}$ sei \mathcal{P}_n wieder die äquidistante Zerlegung von $[a, b]$ in n Teilintervalle. Da f stetig ist, nimmt f auf jedem Intervall $I_k = [x_{k-1}, x_k]$ ein Maximum und ein Minimum an. Seien $c_k, d_k \in I_k$ mit $f(c_k) = \min\{f(x) \mid x \in I_k\}$ und $f(d_k) = \max\{f(x) \mid x \in I_k\}$. Dann gilt

$$\overline{S}(f, \mathcal{P}_n) - \underline{S}(f, \mathcal{P}_n) = \sum_{k=1}^n (f(d_k) - f(c_k)) L(I_k) = \left(\sum_{k=1}^n f(d_k) - f(c_k) \right) \frac{b-a}{n}. \quad (6.2)$$

Sei nun $\varepsilon > 0$. Die stetige Abbildung f ist auf dem kompakten Intervall $[a, b]$ gleichmäßig stetig (Satz 4.14). Folglich existiert ein $\delta > 0$, so dass für alle $v, w \in [a, b]$ mit $|v - w| < \delta$ gilt:

$$|f(v) - f(w)| < \frac{\varepsilon}{b-a}.$$

Wir wählen jetzt $n \in \mathbb{N}$ so groß, dass $\frac{b-a}{n} < \delta$. Dann gilt $|d_k - c_k| < \delta$ für $k = 1, \dots, n$ und aus (6.2) folgt

$$\overline{S}(f, \mathcal{P}_n) - \underline{S}(f, \mathcal{P}_n) < \left(\sum_{k=1}^n \frac{\varepsilon}{b-a} \right) \frac{b-a}{n} = \varepsilon.$$

Das Riemannsche Integrierbarkeitskriterium liefert dann wieder die Behauptung. \square

Der nächste Satz zeigt, dass man das Riemann-Integral für eine Riemann-integrierbare Funktion als Grenzwert einer *beliebigen Folge* von Riemannschen Summen ausrechnen kann.

Satz 6.5 *Sei $f : [a, b] \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion. Dann gilt für jede Folge (\mathcal{P}_m) von Zerlegungen von $[a, b]$ mit $\lim_{m \rightarrow \infty} \|\mathcal{P}_m\| = 0$ und für beliebige Stützstellen ξ_m von \mathcal{P}_m :*

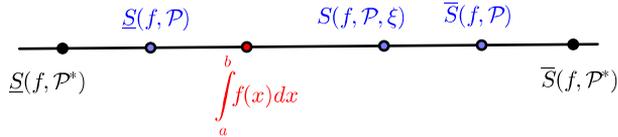
$$\int_a^b f(x) dx = \lim_{m \rightarrow \infty} S(f, \mathcal{P}_m, \xi_m).$$

Beweis. Sei $\varepsilon > 0$ gegeben und \mathcal{P}^* eine Zerlegung von $[a, b]$ mit

$$\bar{S}(f, \mathcal{P}^*) - \underline{S}(f, \mathcal{P}^*) < \frac{\varepsilon}{2},$$

die nach dem Riemannschem Integrierbarkeitskriterium (Satz 6.2) existiert.

Aus Satz 6.1 erhält man dann für alle Verfeinerungen $\mathcal{P} \geq \mathcal{P}^*$ und alle Stützstellen ξ von \mathcal{P}



$$\left| \int_a^b f(x) dx - S(f, \mathcal{P}, \xi) \right| < \frac{\varepsilon}{2}. \quad (6.3)$$

Da f beschränkt ist, existiert eine Konstante $C > 0$ so dass $|f(x)| \leq C$ für alle $x \in [a, b]$. Die Zerlegung \mathcal{P}^* habe r Teilintervalle. Da $\|\mathcal{P}_m\| \rightarrow 0$, existiert ein $m_0 \in \mathbb{N}$, so dass

$$2Cr\|\mathcal{P}_m\| < \frac{\varepsilon}{2} \quad \forall m \geq m_0.$$

Wir betrachten nun \mathcal{P}_m mit $m \geq m_0$. Seien J_1, J_2, \dots, J_s diejenigen Teilintervalle von \mathcal{P}_m , die durch Punkte von \mathcal{P}^* echt zerlegt werden. Dann ist $s \leq r$. Wir fügen nun zu den Stützstellen ξ_m von \mathcal{P}_m neue Stützstellen hinzu, so dass Stützstellen $\tilde{\xi}_m$ von $\mathcal{P}^* \cup \mathcal{P}_m$ entstehen. Damit erhalten wir mit Hilfe der Dreiecksungleichung

$$\begin{aligned} |S(f, \mathcal{P}_m, \xi_m) - S(f, \mathcal{P}_m \cup \mathcal{P}^*, \tilde{\xi}_m)| &\leq 2C \sum_{\ell=1}^s L(J_\ell) \\ &\leq 2Cs\|\mathcal{P}_m\| \leq 2Cr\|\mathcal{P}_m\| < \frac{\varepsilon}{2}. \end{aligned} \quad (6.4)$$

Da $\mathcal{P}_m \cup \mathcal{P}^* \geq \mathcal{P}^*$, folgt aus (6.3) und (6.4)

$$\begin{aligned} \left| S(f, \mathcal{P}_m, \xi_m) - \int_a^b f(x) dx \right| &\leq |S(f, \mathcal{P}_m, \xi_m) - S(f, \mathcal{P}_m \cup \mathcal{P}^*, \tilde{\xi}_m)| + \\ &\quad + \left| S(f, \mathcal{P}_m \cup \mathcal{P}^*, \tilde{\xi}_m) - \int_a^b f(x) dx \right| \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Dies gilt für alle $m \geq m_0$. Somit folgt $\lim_{m \rightarrow \infty} S(f, \mathcal{P}_m, \xi_m) = \int_a^b f(x) dx$. \square

Satz 6.5 ist gut geeignet, um Rechenregeln für Riemann-Integrale zu beweisen und Abschätzungen für Summen auf Abschätzungen für Integrale zu übertragen.

Satz 6.6 1. Sei $f : [a, c] \rightarrow \mathbb{R}$ und $a < b < c$. Ist $f|_{[a,b]} \in \mathcal{R}([a,b], \mathbb{R})$ und $f|_{[b,c]} \in \mathcal{R}([b,c], \mathbb{R})$, so ist $f \in \mathcal{R}([a,c], \mathbb{R})$ und es gilt

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx.$$

2. Ist $f \in \mathcal{R}([a,b], \mathbb{R})$ und $[\alpha, \beta] \subset [a,b]$, so ist $f|_{[\alpha,\beta]} \in \mathcal{R}([\alpha,\beta], \mathbb{R})$.

Beweis. Übungsaufgabe. □

Satz 6.7 (Rechenregeln für das Riemann-Integral) Es seien $f, g : [a,b] \rightarrow \mathbb{R}$ zwei Riemann-integrierbare Funktionen. Dann gilt:

1. Für alle $\mu, \lambda \in \mathbb{R}$ ist die Funktion $\mu f + \lambda g$ Riemann-integrierbar und

$$\int_a^b (\mu f(x) + \lambda g(x)) dx = \mu \int_a^b f(x) dx + \lambda \int_a^b g(x) dx.$$

2. Sei $f \leq g$. Dann gilt

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

3. $|f|$ ist Riemann-integrierbar und

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

4. Für jedes $p \in [1, \infty)$ ist $|f|^p$ Riemann-integrierbar.

5. Das Produkt $f \cdot g$ ist Riemann-integrierbar.

Beweis. Den Beweis lassen wir als Übungsaufgabe. Man kann z.B. das Riemannsche Integrierbarkeitskriterium, Satz 6.5 und die entsprechenden Eigenschaften für Summen benutzen. □

Die Definition des Riemann-Integrals rechtfertigt folgende Definition:

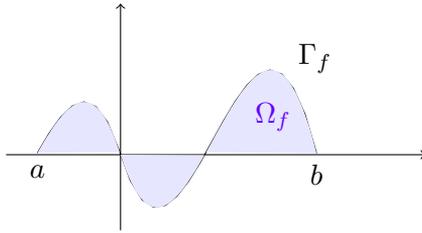
Definition 6.5. Sei $f : [a,b] \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion und $\Omega_f \subset \mathbb{R}^2$ das Gebiet zwischen dem Graphen Γ_f von f und der x -Achse, d.h.

$$\Omega_f := \{(x, y) \in \mathbb{R}^2 \mid 0 \leq y \leq f(x) \text{ falls } f(x) \geq 0, \quad f(x) \leq y \leq 0 \text{ falls } f(x) \leq 0\}.$$

Dann heißt

$$\text{Area}(\Omega_f) := \int_a^b |f(x)| dx$$

Flächeninhalt des Gebietes Ω_f .



Satz 6.8 (Mittelwertsatz der Integralrechnung)

Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und $g : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar mit $g \geq 0$. Dann existiert ein $\eta \in [a, b]$ so dass

$$\int_a^b f(x)g(x) dx = f(\eta) \cdot \int_a^b g(x) dx.$$

Beweis. Da f stetig und $[a, b]$ kompakt ist, existieren $m := \min\{f(x) \mid x \in [a, b]\}$ und $M := \max\{f(x) \mid x \in [a, b]\}$. Da $g \geq 0$ folgt

$$mg(x) \leq f(x)g(x) \leq Mg(x) \quad \forall x \in [a, b].$$

und mit den Rechenregeln aus Satz 6.7 somit

$$m \int_a^b g(x) dx \leq \int_a^b f(x)g(x) dx \leq M \int_a^b g(x) dx.$$

Sei $\int_a^b g(x) dx \neq 0$. Dann ist dieses Integral positiv und wir erhalten

$$m \leq \frac{\int_a^b f(x)g(x) dx}{\int_a^b g(x) dx} \leq M.$$

Wir setzen $\mu := \frac{\int_a^b f(x)g(x) dx}{\int_a^b g(x) dx}$. Dann gilt $\mu \in [m, M]$ und

$$\int_a^b f(x)g(x) dx = \mu \cdot \int_a^b g(x) dx.$$

Im Fall $\int_a^b g(x) dx = 0$ gilt die letzte Gleichung für ein beliebiges $\mu \in [m, M]$. Da f stetig ist, existiert nach dem Zwischenwertsatz für stetige Funktionen ein $\eta \in [a, b]$ mit $f(\eta) = \mu$. Dies zeigt die Behauptung des Satzes. \square

Folgerung 6.1 Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, dann existiert ein $\eta \in [a, b]$, so dass gilt

$$\int_a^b f(x) dx = f(\eta) \cdot (b - a).$$

Beweis. Dies folgt aus Satz 6.8 mit $g \equiv 1$. □

Für stetige Funktionen $f \geq 0$ bedeutet das, dass der Flächeninhalt zwischen der x -Achse und dem Graphen von f gleich der Fläche des Rechtecks mit den Kantenlängen $[a, b]$ und $f(\eta)$ ist.

Folgerung 6.2 Sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion mit $f \geq 0$ und $\int_a^b f(x) dx = 0$. Dann gilt $f = 0$.

Beweis. Sei $c \in (a, b)$. Wir betrachten Folgen (x_n) und (y_n) mit $a \leq x_n < c < y_n \leq b$ und $x_n \rightarrow c, y_n \rightarrow c$. Da $f \geq 0$, gilt

$$0 \leq \int_{x_n}^{y_n} f(x) dx \leq \int_a^b f(x) dx = 0.$$

Nach dem Mittelwertsatz existiert ein $\eta_n \in [x_n, y_n]$ mit

$$0 = \int_{x_n}^{y_n} f(x) dx = f(\eta_n)(y_n - x_n).$$

Folglich ist $f(\eta_n) = 0$. Nach Konstruktion konvergiert die Folge (η_n) gegen c . Da f stetig ist, konvergiert dann auch $f(\eta_n)$ gegen $f(c)$. Folglich ist $f(c) = 0$ für jedes $c \in (a, b)$. Für die Randpunkte folgt dies dann wegen der Stetigkeit von f . □

Man kann mit Hilfe von Satz 6.5 auch Riemann-Integrale berechnen:

Einfaches Beispiel: Berechne das Riemann-Integral $\int_a^b x^2 dx$:

Die Funktion $f : [a, b] \rightarrow \mathbb{R}$, definiert durch $f(x) := x^2$, ist stetig, also Riemann-integrierbar. Sei $\mathcal{P}_n = \{x_0, \dots, x_n\}$ die äquidistante Unterteilung von $[a, b]$ mit $x_k = a + k \cdot \frac{b-a}{n}$. Wir wählen die Stützstellen $\xi = (x_1, \dots, x_n)$. Dann gilt

$$\begin{aligned} S(f, \mathcal{P}_n, \xi) &= \sum_{k=1}^n f(x_k) L(I_k) \\ &= \sum_{k=1}^n \left(a + k \cdot \frac{b-a}{n} \right)^2 \cdot \frac{b-a}{n} \\ &= \frac{b-a}{n} \cdot \sum_{k=1}^n \left(a^2 + 2ka \cdot \frac{b-a}{n} + k^2 \cdot \frac{(b-a)^2}{n^2} \right) \\ &= a^2(b-a) + 2a \frac{(b-a)^2}{n^2} \cdot \sum_{k=1}^n k + \frac{(b-a)^3}{n^3} \cdot \sum_{k=1}^n k^2 \\ &= a^2(b-a) + 2a \frac{(b-a)^2}{n^2} \cdot \frac{n(n+1)}{2} + \frac{(b-a)^3}{n^3} \cdot \frac{n(n+1)(2n+1)}{6}. \end{aligned}$$

Es folgt

$$\begin{aligned} \int_a^b x^2 dx &= \lim_{n \rightarrow \infty} S(f, \mathcal{P}_n, \xi) \\ &= a^2(b-a) + a(b-a)^2 + \frac{1}{3}(b-a)^3 \\ &= a^2b - a^3 + ab^2 + a^3 - 2a^2b + \frac{1}{3}(b^3 - 3b^2a + 3ba^2 - a^3) \\ &= \frac{1}{3}(b^3 - a^3). \end{aligned}$$

Dies ist eine etwas umständliche Methode zur Berechnung des Riemann-Integrals. Im nächsten Abschnitt lernen wir einfachere Verfahren zur Berechnung des Riemann-Integrals kennen.

6.2 Integration und Differentiation

Wir wollen jetzt nachweisen, dass die Integration für gewisse Klassen von Funktionen die Umkehrung der Differentiation ist.

Im Folgenden bezeichne $I \subset \mathbb{R}$ immer ein *beliebiges* beschränktes oder unbeschränktes Intervall, das aus mehr als einem Punkt besteht.

Definition 6.6. Eine differenzierbare Funktion $F : I \rightarrow \mathbb{R}$ heißt *Stammfunktion* einer Funktion $f : I \rightarrow \mathbb{R}$, wenn $F' = f$.

Satz 6.9 Sind $F : I \rightarrow \mathbb{R}$ und $G : I \rightarrow \mathbb{R}$ zwei Stammfunktionen von $f : I \rightarrow \mathbb{R}$, dann gilt $G = F + c$, wobei c eine reelle Konstante ist.

Beweis. Wir betrachten die Funktion $G - F : I \rightarrow \mathbb{R}$. Nach Voraussetzung gilt $(G - F)' = G' - F' = f - f = 0$. Folglich ist $G - F$ eine konstante Funktion, also $G = F + c$ für ein $c \in \mathbb{R}$. \square

Wenn eine Stammfunktion von f existiert, so ist sie also bis auf eine additive Konstante eindeutig bestimmt. Wir interessieren uns nun für die *Existenz* von Stammfunktionen.

Satz 6.10 (Hauptsatz der Differential- und Integralrechnung - Teil 1)

Sei $f : I \rightarrow \mathbb{R}$ eine stetige Funktion und $a \in I$. Wir definieren die Funktion $F : I \rightarrow \mathbb{R}$ durch das Riemann-Integral

$$F(x) := \int_a^x f(t) dt.$$

Dann ist F eine Stammfunktion von f .

Beweis. Wir betrachten den Differenzenquotienten von F in $x \in I$:

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \left(\int_a^{x+h} f(t) dt - \int_a^x f(t) dt \right) = \frac{1}{h} \int_x^{x+h} f(t) dt.$$

Nach dem Mittelwertsatz der Integralrechnung (Satz 6.8) existiert ein $\eta(h) \in [x, x+h]$ (bzw. $\eta(h) \in [x+h, x]$, falls $h < 0$) mit

$$\int_x^{x+h} f(t) dt = f(\eta(h)) \cdot h.$$

Wegen $\lim_{h \rightarrow 0} \eta(h) = x$ und der Stetigkeit von f existiert der Grenzwert

$$F'(x) = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} f(\eta(h)) = f(x). \quad \square$$

Bemerkung: *Stetige* Funktionen besitzen also immer eine Stammfunktionen, die durch ihr Riemann-Integral definiert ist. Man beachte aber:

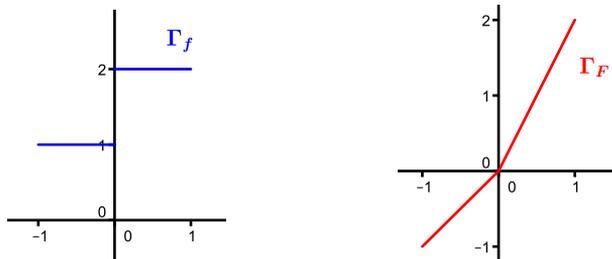
- Nicht jede Riemann-integrierbare Funktion besitzt eine Stammfunktion. Die Funktion $f : [-1, 1] \rightarrow \mathbb{R}$ mit

$$f(x) := \begin{cases} 1 & \text{falls } x \in [-1, 0], \\ 2 & \text{falls } x \in (0, 1]. \end{cases}$$

ist Riemann-integrierbar, besitzt aber keine Stammfunktion. Eine stetige Funktion $F : [-1, 1] \rightarrow \mathbb{R}$ mit $F'(x) = f(x)$ für alle $x \neq 0$ hätte nämlich die Form

$$F(x) = \begin{cases} x + c & \text{falls } x \in [-1, 0], \\ 2x + c & \text{falls } x \in (0, 1], \end{cases}$$

für eine Konstante $c \in \mathbb{R}$. Eine solche Funktion F ist aber in $x = 0$ nicht differenzierbar.



- Es gibt auch Funktionen, die Stammfunktionen besitzen, aber nicht Riemann-integrierbar sind. Wir betrachten dazu die Funktion $F : [0, 1] \subset \mathbb{R} \rightarrow \mathbb{R}$

$$F(x) := \begin{cases} x^2 \sin \frac{1}{x^2} & x \in (0, 1], \\ 0 & x = 0. \end{cases}$$

F ist auf $[0, 1]$ differenzierbar, aber $f := F'$ ist auf $[0, 1]$ nicht beschränkt (Übungsaufgabe).

Satz 6.11 (Hauptsatz der Differential- und Integralrechnung - Teil 2)

Sei $f : I \rightarrow \mathbb{R}$ stetig und $F : I \rightarrow \mathbb{R}$ eine Stammfunktion von f . Dann gilt für alle $a, b \in I$:

$$\int_a^b f(x) dx = F(b) - F(a).$$

Beweis. Wir betrachten die Funktion $F_0 : I \rightarrow \mathbb{R}$ mit $F_0(t) := \int_a^t f(x) dx$. F_0 ist eine Stammfunktion von f mit $F_0(a) = 0$ und $F_0(b) = \int_a^b f(x) dx$. Für die Stammfunktion F gilt dann $F = F_0 + c$ für eine Konstante $c \in \mathbb{R}$ und folglich

$$F(b) - F(a) = F_0(b) - F_0(a) = F_0(b) = \int_a^b f(x) dx. \quad \square$$

Bezeichnung: Man setzt: $F(x) \Big|_a^b := F(b) - F(a)$. Die Formel aus Satz 6.11 schreibt sich dann in der Form: $\int_a^b f(x) dx = F(x) \Big|_a^b$. Deshalb wird häufig die folgende Symbolik für die Stammfunktion F von f benutzt¹:

$$\int f(x) dx = F(x) \quad \text{auf } I.$$

Man nennt

$$\int f(x) dx \quad \text{unbestimmtes Integral von } f,$$

$$\int_a^b f(x) dx \quad \text{bestimmtes Integral von } f.$$

f heißt *Integrand*, x *Integrationsvariable*, $[a, b]$ das *Integrationsintervall*, a und b die *untere* bzw. *obere Integrationsgrenze*.

In der folgenden Liste stellen wir einige wichtige Grundintegrale zusammen. Der Beweis erfolgt durch Ableiten der Stammfunktion.

¹ Diese Bezeichnung ist etwas problematisch, da die Stammfunktion nur bis auf eine Konstante eindeutig bestimmt ist. Man merke sich deshalb, was mit dieser Symbolik gemeint ist.

Wichtige Grundintegrale:

$$(1) \int x^\alpha dx = \frac{1}{\alpha+1} x^{\alpha+1} \quad \text{für } \alpha \in \mathbb{R} \text{ mit } \alpha \neq -1.$$

Dies gilt für nicht ganzzahlige $\alpha \in \mathbb{R}$ auf $I = (0, \infty)$, für $\alpha \in \mathbb{N}_0$ auf \mathbb{R} und für $\alpha \in \mathbb{Z}$ mit $\alpha \leq -1$ auf $I = (-\infty, 0) \cup (0, \infty)$.

$$(2) \int \frac{1}{x} dx = \ln|x| \quad \text{auf } (0, \infty) \text{ und } (-\infty, 0).$$

$$(3) \int e^{ax} dx = \frac{1}{a} \cdot e^{ax} \quad \text{für } a \in \mathbb{R} \text{ mit } a \neq 0, \quad \text{auf } \mathbb{R}.$$

$$(4) \int \cos x dx = \sin x, \quad \int \sin x dx = -\cos x \quad \text{auf } \mathbb{R}.$$

$$(5) \int \cosh x dx = \sinh x, \quad \int \sinh x dx = \cosh x \quad \text{auf } \mathbb{R}.$$

$$(6) \int \frac{1}{1+x^2} dx = \arctan x \quad \text{auf } \mathbb{R}.$$

$$(7) \int \frac{1}{1-x^2} dx = \frac{1}{2} \ln \left| \frac{1+x}{1-x} \right| = \begin{cases} \operatorname{artanh}(x) & \text{auf } (-1, 1) \\ \operatorname{arcoth}(x) & \text{auf } (-\infty, -1) \cup (1, \infty) \end{cases}.$$

$$(8) \int \frac{1}{\sqrt{1+x^2}} dx = \ln(x + \sqrt{x^2+1}) = \operatorname{arsinh}(x) \quad \text{auf } \mathbb{R}.$$

$$(9) \int \frac{1}{\sqrt{1-x^2}} dx = \arcsin(x) \quad \text{auf } (-1, 1).$$

$$(10) \int \frac{1}{\sqrt{x^2-1}} dx = \ln|x + \sqrt{x^2-1}| = \begin{cases} \operatorname{arcosh}(x) & \text{auf } (1, \infty) \\ -\operatorname{arcosh}(-x) & \text{auf } (-\infty, -1) \end{cases}.$$

Wir beweisen jetzt die beiden wichtigsten Rechenregeln für Integrale, die *partielle Integration* und die *Substitutionsregel*.

Satz 6.12 (Partielle Integration)

Seien $g, f : I \rightarrow \mathbb{R}$ stetig differenzierbare Funktionen. Dann gilt für alle $a, b \in I$:

$$\int_a^b f'(x)g(x) dx = - \int_a^b f(x)g'(x) dx + f(x)g(x) \Big|_a^b,$$

bzw. für die unbestimmten Integrale

$$\int f'(x)g(x) dx = - \int f(x)g'(x) dx + f(x)g(x) \quad \text{auf } I.$$

Beweis. Für $F := f \cdot g$ gilt nach der Produktregel für Ableitungen $F' = f' \cdot g + f \cdot g'$. Die Linearität des Integrals und der Hauptsatz der Differential- und Integralrechnung liefert dann für alle $a, b \in I$

$$\begin{aligned} \int_a^b f'(x)g(x) dx + \int_a^b f(x)g'(x) dx &= \int_a^b (f'(x)g(x) + f(x)g'(x)) dx \\ &= \int_a^b (f \cdot g)'(x) dx = f(x) \cdot g(x) \Big|_a^b. \end{aligned} \quad \square$$

Satz 6.13 (Substitutionsregel)

Sei $\phi : I \rightarrow \mathbb{R}$ stetig differenzierbar und $f : J := \phi(I) \rightarrow \mathbb{R}$ stetig. Dann gilt für alle $c, d \in I$

$$\int_c^d f(\phi(t))\phi'(t) dt = \int_{\phi(c)}^{\phi(d)} f(x) dx.$$

Ist ϕ zusätzlich streng monoton, so existiert die Umkehrfunktion $\phi^{-1} : J \rightarrow I$ und es gilt für alle $a, b \in J$

$$\int_a^b f(x) dx = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f(\phi(t))\phi'(t) dt.$$

(Formale Regel: Man substituiert $x = \phi(t)$, setzt formal $dx = \phi'(t)dt$ und ändert die Integrationsgrenzen entsprechend).

Beweis. Wir bemerken zunächst, dass wegen der Stetigkeit von ϕ mit I auch $J = \phi(I) \subset \mathbb{R}$ ein Intervall ist (Zwischenwertsatz). Sei nun $F : J \rightarrow \mathbb{R}$ eine Stammfunktion von f . Dann gilt für $F \circ \phi : I \rightarrow \mathbb{R}$ nach der Kettenregel für Ableitungen

$$(F \circ \phi)'(t) = F'(\phi(t)) \cdot \phi'(t) = f(\phi(t)) \cdot \phi'(t) \quad \forall t \in I.$$

Nach Satz 6.11 folgt für $c, d \in I$:

$$\int_c^d f(\phi(t))\phi'(t) dt = (F \circ \phi)(d) - (F \circ \phi)(c) = F(\phi(d)) - F(\phi(c)) = \int_{\phi(c)}^{\phi(d)} f(x) dx.$$

Wenn ϕ streng monoton ist, so ist ϕ injektiv und $\phi : I \rightarrow J$ somit bijektiv. Folglich existiert die Umkehrfunktion $\phi^{-1} : J \rightarrow I$. Für gegebene $a, b \in J$ kann man also eindeutige bestimmte $c, d \in I$ finden mit $\phi(c) = a$ und $\phi(d) = b$. Die 2. Behauptung folgt dann aus der ersten. \square

Bemerkung: In der unbestimmten Form schreibt sich die Substitutionsregel als

$$\int f(\phi(t))\phi'(t) dt = \int f(x) dx \Big|_{x=\phi(t)} \quad \text{auf } I$$

bzw. im Fall, dass die Umkehrfunktion ϕ^{-1} existiert als

$$\int f(x) dx = \int f(\phi(t))\phi'(t) dt \Big|_{t=\phi^{-1}(x)} \quad \text{auf } J.$$

Beispiel 1: Sei $f : I \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion ohne Nullstellen auf I . Dann gilt

$$\int \frac{f'(x)}{f(x)} dx = \ln |f(x)| \quad \text{auf } I.$$

Dies sieht man durch Ableiten der Funktion auf der rechten Seite oder durch die Substitution $y = f(x)$, $dy = f'(x) dx$:

$$\int \frac{f'(x)}{f(x)} dx = \int \frac{dy}{y} \Big|_{y=f(x)} = \ln |y| \Big|_{y=f(x)} = \ln |f(x)| \quad \text{auf } I.$$

So ist zum Beispiel

$$\int \tan x dx = \int \frac{\sin x}{\cos x} dx = - \int \frac{\cos' x}{\cos x} dx = - \ln |\cos x| \quad \text{auf } \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Beispiel 2: Berechnung von $I_n(x) := \int \sin^n(x)$, $n \in \mathbb{N}_0$.

Für $n = 0, 1$ gilt:

$$I_0(x) = x \quad \text{und} \quad I_1(x) = -\cos x.$$

Für $n \geq 2$ leiten wir mittels partieller Integration eine Rekursionsformel her:

$$\begin{aligned} I_n(x) &= - \int \sin^{n-1}(x) \cdot \cos'(x) dx \\ &\stackrel{\text{part. Int.}}{=} - \sin^{n-1}(x) \cdot \cos(x) + \int \frac{d}{dx} \sin^{n-1}(x) \cdot \cos(x) dx \\ &= - \sin^{n-1}(x) \cdot \cos(x) + (n-1) \int \sin^{n-2}(x) (1 - \sin^2(x)) dx \\ &= - \sin^{n-1}(x) \cdot \cos(x) + (n-1)I_{n-2}(x) - (n-1)I_n(x). \end{aligned}$$

Daraus erhält man die Rekursionsformel:

$$I_n(x) = -\frac{1}{n} \cdot \sin^{n-1}(x) \cdot \cos(x) + \frac{n-1}{n} I_{n-2}(x) \quad \text{für } n \geq 2.$$

Für $J_n(x) := \int \cos^n(x) dx$ berechnet man analog

$$J_n(x) = \frac{1}{n} \cos^{n-1}(x) \cdot \sin(x) + \frac{n-1}{n} J_{n-2}(x) \quad \text{für } n \geq 2$$

mit $J_0(x) = x$ und $J_1(x) = \sin(x)$.

Beispiel 3: Integrale rationaler Funktionen $\int \frac{a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0}{b_n x^n + b_{n-1} x^{n-1} + \dots + b_1 x + b_0} dx$:

Solche Integrale werden mit Hilfe der *Partialbruchzerlegung*² berechnet. Dabei stellt man die rationale Funktion $R(x) := \frac{a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0}{b_n x^n + b_{n-1} x^{n-1} + \dots + b_1 x + b_0}$ als Summe *einfacher* Funktionen, die sogenannten *Partialbrüche*, dar. Im folgenden beschreiben wir dieses Verfahren zunächst für komplexe rationale Funktionen und spezialisieren es anschließend auf den reellen Fall.

1. Komplexe Partialbruchzerlegung

Sei $Q \in \mathbb{C}[z]$ ein komplexes Polynom vom Grad $n \geq 1$, d.h.

$$Q(z) := b_n z^n + b_{n-1} z^{n-1} + \dots + b_1 z + b_0$$

mit $b_0, b_1, \dots, b_n \in \mathbb{C}$ und $b_n \neq 0$. Nach dem Fundamentalsatz der Algebra hat Q n komplexe Nullstellen und lässt sich in komplexe Linearfaktoren zerlegen:

Sind ξ_1, \dots, ξ_k die verschiedenen komplexen Nullstellen von Q und ν_1, \dots, ν_k ihre jeweilige Vielfachheit, dann gilt

$$Q(z) = b_n (z - \xi_1)^{\nu_1} \cdot (z - \xi_2)^{\nu_2} \cdot \dots \cdot (z - \xi_k)^{\nu_k}.$$

Satz 6.14 (Partialbruchzerlegung komplexer rationaler Funktionen)

Seien $P, Q \in \mathbb{C}[z]$ komplexe Polynome mit $\deg(P) < \deg(Q)$. Q habe die verschiedenen Nullstellen ξ_1, \dots, ξ_k mit der jeweiligen Vielfachheit ν_1, \dots, ν_k . Dann existieren eindeutig bestimmte Konstanten $A_{j1}, A_{j2}, \dots, A_{j\nu_j} \in \mathbb{C}$ für $j \in \{1, \dots, k\}$, so dass

$$\frac{P(z)}{Q(z)} = \sum_{j=1}^k \left(\frac{A_{j1}}{(z - \xi_j)} + \frac{A_{j2}}{(z - \xi_j)^2} + \dots + \frac{A_{j\nu_j}}{(z - \xi_j)^{\nu_j}} \right) \quad \forall z \notin \{\xi_1, \dots, \xi_k\}. \quad (6.5)$$

Beweis. Wir führen den Beweis durch Induktion über $\deg(Q)$.

Ind.-Anfang: Sei $\deg(Q) = 1$. Da nach Voraussetzung $\deg(P) < \deg(Q)$, hat P den Grad Null. Es gilt also $P(z) = a_0$ und $Q(z) = b_1 z + b_0 = b_1(z - \xi_1)$ mit $\xi_1 = -\frac{b_0}{b_1}$. Daraus folgt

$$\frac{P(z)}{Q(z)} = \frac{a_0}{b_1(z - \xi_1)} = \frac{A_{11}}{z - \xi_1}$$

für die eindeutig bestimmte Zahl $A_{11} := \frac{a_0}{b_1}$.

Ind.-Schritt: Wir setzen voraus, dass die Behauptung für alle rationalen Funktionen $\tilde{R} = \frac{\tilde{P}}{\tilde{Q}}$ mit $\deg(\tilde{Q}) \leq n - 1$ gilt und zeigen sie dann für rationale Funktionen $R = \frac{P}{Q}$ mit $\deg(Q) = n$:

Sei ξ_1 eine Nullstelle von Q mit Vielfachheit ν_1 . Dann kann man die Linearfaktoren $(z - \xi_1)^{\nu_1}$ abspalten und erhält $Q(z) = (z - \xi_1)^{\nu_1} \cdot S(z)$, wobei S ein komplexes Polynom vom Grad $n - \nu_1$ mit $S(\xi_1) \neq 0$ ist. Sei $a := \frac{P(\xi_1)}{S(\xi_1)}$. Dann gilt

² Den Beweis der Sätze 6.14 und 6.15 haben wir in der Vorlesung nicht vorgeführt.

$$\frac{P(z)}{Q(z)} - \frac{a}{(z - \xi_1)^{\nu_1}} = \frac{P(z) - a \cdot S(z)}{(z - \xi_1)^{\nu_1} \cdot S(z)}. \quad (*)$$

Nach Definition von a ist gilt $P(\xi_1) - aS(\xi_1) = 0$, d.h. ξ_1 ist eine Nullstelle des Polynoms $P - aS$. Ist $P - aS = 0$, so folgt aus $(*)$ sofort die Induktionsbehauptung. Ist $P - aS \neq 0$, so kann man den Linearfaktor $(z - \xi_1)$ abspalten, d.h. es gilt $P(z) - aS(z) = (z - \xi_1)\tilde{S}(z)$ für ein Polynom \tilde{S} . Es folgt

$$\frac{P(z)}{Q(z)} \stackrel{(*)}{=} \frac{a}{(z - \xi_1)^{\nu_1}} + \frac{\tilde{S}(z)}{(z - \xi_1)^{\nu_1-1} \cdot S(z)}.$$

Das Nennerpolynom im zweiten Summanden hat den Grad $n - 1$ und als seine Nullstellen kommen alle Nullstellen von Q vor, wobei ξ_1 mit Vielfachheit $\nu_1 - 1$ auftritt. Das Zählerpolynom \tilde{S} im zweiten Summanden hat den Grad $< n - 1$. Nach Ind.-Voraussetzung kann man die rationale Funktion im zweiten Summanden entsprechend der Behauptung des Satzes zerlegen mit eindeutig bestimmten Koeffizienten $A_{j\ell}$. $A_{1\nu_1} := a$ ist ebenfalls eindeutig bestimmt. \square

2. Reelle Partialbruchzerlegung

Sei $Q \in \mathbb{R}[x]$ nun ein reelles Polynom vom Grad $n \geq 1$, d.h.

$$Q(x) = b_n x^n + b_{n-1} x^{n-1} + \dots + b_1 x^1 + b_0$$

mit $b_0, b_1, \dots, b_n \in \mathbb{R}$ und $b_n \neq 0$. Dann tritt mit jeder echt komplexen Nullstelle ξ von Q auch die konjugiert-komplexe Zahl $\bar{\xi}$ als Nullstelle mit der gleichen Vielfachheit auf. Aus

$$0 = Q(\xi) = b_n \xi^n + b_{n-1} \xi^{n-1} + \dots + b_1 \xi^1 + b_0$$

folgt nämlich durch Konjugieren

$$0 = \overline{Q(\xi)} = \overline{b_n \xi^n + b_{n-1} \xi^{n-1} + \dots + b_1 \xi^1 + b_0} = b_n \bar{\xi}^n + b_{n-1} \bar{\xi}^{n-1} + \dots + b_1 \bar{\xi}^1 + b_0 = Q(\bar{\xi}).$$

Das quadratische Polynom $(x - \xi)(x - \bar{\xi})$ ist dann reell und reell-irreduzibel, da

$$\begin{aligned} (x - \xi)(x - \bar{\xi}) &= x^2 - (\xi + \bar{\xi})x + \xi\bar{\xi} = x^2 - 2 \operatorname{Re}(\xi)x + |\xi|^2 \\ &=: x^2 + px + q, \end{aligned}$$

wobei für die Koeffizienten p und q gilt $\frac{p^2}{4} - q = \operatorname{Re}(\xi)^2 - |\xi|^2 = -\operatorname{Im}(\xi)^2 < 0$.

Sind $\lambda_1, \dots, \lambda_s$ die reellen Nullstellen von Q mit der jeweiligen Vielfachheit μ_1, \dots, μ_s und $\xi_1, \bar{\xi}_1, \dots, \xi_r, \bar{\xi}_r$ die echt komplexen Nullstellen von Q mit der Vielfachheit $\nu_1, \nu_1, \dots, \nu_r, \nu_r$, so hat die (komplexe) Linearfaktorzerlegung von $Q(x)$ die Form

$$Q(x) = b_n (x - \lambda_1)^{\mu_1} \cdot \dots \cdot (x - \lambda_s)^{\mu_s} \cdot (x - \xi_1)^{\nu_1} (x - \bar{\xi}_1)^{\nu_1} \cdot \dots \cdot (x - \xi_r)^{\nu_r} (x - \bar{\xi}_r)^{\nu_r}.$$

Wir fassen die Produkte $(x - \xi_j)(x - \bar{\xi}_j)$ zu einem reellen Polynom zusammen und erhalten die *reelle Faktorzerlegung* von $Q(x)$:

$$Q(x) = b_n (x - \lambda_1)^{\mu_1} \cdot \dots \cdot (x - \lambda_s)^{\mu_s} \cdot \underbrace{(x^2 + p_1 x + q_1)^{\nu_1}}_{\text{reell-irreduzibel}} \cdot \dots \cdot \underbrace{(x^2 + p_r x + q_r)^{\nu_r}}_{\text{reell-irreduzibel}}.$$

Satz 6.15 (Partialbruchzerlegung reeller rationaler Funktionen)

Seien $P, Q \in \mathbb{R}[x]$ zwei reelle Polynome mit $\deg(P) < \deg(Q)$ und sei

$$Q(x) = b_n(x - \lambda_1)^{\mu_1} \cdot \dots \cdot (x - \lambda_s)^{\mu_s} \cdot \underbrace{(x^2 + p_1x + q_1)^{\nu_1}}_{\text{reell-irreduzibel}} \cdot \dots \cdot \underbrace{(x^2 + p_r x + q_r)^{\nu_r}}_{\text{reell-irreduzibel}}$$

die reelle Faktorzerlegung von $Q(x)$. Dann existieren eindeutig bestimmte reelle Zahlen $A_{j1}, A_{j2}, \dots, A_{j\mu_j}$ für $j = 1, \dots, s$, und $B_{\ell 1}, B_{\ell 2}, \dots, B_{\ell\nu_\ell}, C_{\ell 1}, C_{\ell 2}, \dots, C_{\ell\nu_\ell}$ für $\ell = 1, \dots, r$, so dass für alle $x \in \mathbb{R}$ mit $x \notin \{\lambda_1, \dots, \lambda_s\}$ gilt:

$$\begin{aligned} \frac{P(x)}{Q(x)} &= \sum_{j=1}^s \left(\frac{A_{j1}}{(x - \lambda_j)} + \frac{A_{j2}}{(x - \lambda_j)^2} + \dots + \frac{A_{j\mu_j}}{(x - \lambda_j)^{\mu_j}} \right) \\ &+ \sum_{\ell=1}^r \left(\frac{B_{\ell 1}x + C_{\ell 1}}{(x^2 + p_\ell x + q_\ell)} + \frac{B_{\ell 2}x + C_{\ell 2}}{(x^2 + p_\ell x + q_\ell)^2} + \dots + \frac{B_{\ell\nu_\ell}x + C_{\ell\nu_\ell}}{(x^2 + p_\ell x + q_\ell)^{\nu_\ell}} \right). \end{aligned} \quad (6.6)$$

Beweis. Wir machen zunächst die komplexe Partialbruchzerlegung wie in Satz 6.14 und stellen $\frac{P(x)}{Q(x)}$ als Summe komplexer Partialbrüche nach Formel (6.5) dar. Es existieren also eindeutig bestimmte komplexe Zahlen $A_{j1}, \dots, A_{j\mu_j}$ und $a_{\ell 1}, \dots, a_{\ell\nu_\ell}, b_{\ell 1}, \dots, b_{\ell\nu_\ell}$, so dass für alle $x \in \mathbb{R}$ mit $x \notin \{\lambda_1, \dots, \lambda_s\}$ gilt

$$\begin{aligned} \frac{P(x)}{Q(x)} &= \sum_{j=1}^s \left(\frac{A_{j1}}{(x - \lambda_j)} + \frac{A_{j2}}{(x - \lambda_j)^2} + \dots + \frac{A_{j\mu_j}}{(x - \lambda_j)^{\mu_j}} \right) \\ &+ \sum_{\ell=1}^r \left(\frac{a_{\ell 1}}{(x - \xi_\ell)} + \frac{a_{\ell 2}}{(x - \xi_\ell)^2} + \dots + \frac{a_{\ell\nu_\ell}}{(x - \xi_\ell)^{\nu_\ell}} \right) \\ &+ \sum_{\ell=1}^r \left(\frac{b_{\ell 1}}{(x - \bar{\xi}_\ell)} + \frac{b_{\ell 2}}{(x - \bar{\xi}_\ell)^2} + \dots + \frac{b_{\ell\nu_\ell}}{(x - \bar{\xi}_\ell)^{\nu_\ell}} \right). \end{aligned} \quad (6.7)$$

Da $P(x)$ und $Q(x)$ reelle Polynome sind, gilt $\overline{\frac{P(x)}{Q(x)}} = \frac{P(x)}{Q(x)}$ für alle $x \in \mathbb{R}$ mit $Q(x) \neq 0$. Dies liefert

$$0 = \frac{P(x)}{Q(x)} - \overline{\frac{P(x)}{Q(x)}} = \sum_{j=1}^s \sum_{k=1}^{\mu_j} \frac{A_{jk} - \overline{A_{jk}}}{(x - \lambda_j)^k} + \sum_{\ell=1}^r \sum_{k=1}^{\nu_\ell} \left(\frac{a_{\ell k} - \overline{b_{\ell k}}}{(x - \xi_\ell)^k} + \frac{b_{\ell k} - \overline{a_{\ell k}}}{(x - \bar{\xi}_\ell)^k} \right)$$

für alle $x \in \mathbb{R}$ mit $Q(x) \neq 0$. Wegen der Eindeutigkeit der Konstanten in der komplexen Partialbruchzerlegung (der linken Seite $0 = \frac{0}{Q(x)}$) folgt $A_{jk} = \overline{A_{jk}}$, also $A_{jk} \in \mathbb{R}$, und $b_{\ell k} = \overline{a_{\ell k}}$.

Wir fassen in (6.7) nun noch die komplexen Partialbrüche vom Typ $\frac{a}{(x - \xi)^k} + \frac{\bar{a}}{(x - \bar{\xi})^k}$ zu reellen Partialbrüchen zusammen:

$$\frac{a}{(x - \xi)^k} + \frac{\bar{a}}{(x - \bar{\xi})^k} = \frac{a(x - \bar{\xi})^k + \bar{a}(x - \xi)^k}{(x^2 + px + q)^k} =: \frac{S(x)}{(x^2 + px + q)^k},$$

wobei $S(x)$ ein reelles Polynom von Grad $\leq k$ ist. Ist $k = 1$, so hat dieser Bruch bereits die gewünschte Form. Ist $k > 1$, so teilen wir $S(x)$ durch $(x^2 + px + q)$ mit Rest:

$$S(x) = S_1(x)(x^2 + px + q) + (D_1x + E_1)$$

und wiederholen dies ggf., d.h.,

$$\begin{aligned} \frac{a}{(x-\xi)^k} + \frac{\bar{a}}{(x-\bar{\xi})^k} &= \frac{S_1(x)}{(x^2+px+q)^{k-1}} + \frac{D_1x+E_1}{(x^2+px+q)^k} \\ &= \frac{S_2(x)}{(x^2+px+q)^{k-2}} + \frac{D_2x+E_2}{(x^2+px+q)^{k-1}} + \frac{D_1x+E_1}{(x^2+px+q)^k} \\ &\vdots \\ &= \frac{D_mx+E_m}{(x^2+px+q)} + \dots + \frac{D_2x+E_2}{(x^2+px+q)^{k-1}} + \frac{D_1x+E_1}{(x^2+px+q)^k}. \end{aligned}$$

Damit erhalten wir die Behauptung des Satzes. \square

Methode zur Berechnung der Koeffizienten in der Partialbruchzerlegung:

Sei $R(x) := \frac{P(x)}{Q(x)}$ eine reelle rationale Funktion mit der Partialbruchzerlegung (6.6) aus Satz 6.15. Wir multiplizieren Gleichung (6.6) mit dem Nennerpolynom $Q(x)$ und erhalten daraus eine Polynom-Gleichung:

$$P(x) = \text{Polynom} \dots$$

Das Polynom auf der rechten Seite enthält die gesuchten reellen Zahlen A_{jk} , B_{ℓ} und C_{ℓ} in seinen Koeffizienten. Der Koeffizientenvergleich für beide Polynome liefert ein Gleichungssystem für diese Zahlen, das man eindeutig lösen kann.

Beispiel: Partialbruchzerlegung von $R(x) := \frac{3x^3-9x^2+14x+10}{x^4-6x^3+10x^2}$

Es gilt $Q(x) := x^4 - 6x^3 + 10x^2 = x^2(x^2 - 6x + 10)$. Q hat $\lambda_1 = 0$ als doppelte reelle Nullstelle. Der letzte Faktor ist ein reell-irreduzibles quadratisches Polynom. Der Ansatz für die Partialbruchzerlegung von $R(x)$ lautet somit:

$$R(x) = \frac{A}{x} + \frac{B}{x^2} + \frac{Cx+D}{x^2-6x+10}$$

für eindeutig bestimmte reelle Zahlen A, B, C, D . Multiplizieren wir diese Gleichung mit $Q(x)$, so erhalten wir die Polynom-Gleichung

$$\begin{aligned} 3x^3 - 9x^2 + 14x + 10 &= Ax(x^2 - 6x + 10) + B(x^2 - 6x + 10) + Cx^3 + Dx^2 \\ &= (A+C)x^3 + (-6A+B+D)x^2 + (10A-6B)x + 10B. \end{aligned}$$

Der Koeffizientenvergleich liefert das Gleichungssystem

$$3 = A + C, \quad -9 = -6A + B + D, \quad 14 = 10A - 6B, \quad 10 = 10B$$

mit der Lösung $A = 2$, $B = 1$, $C = 1$, $D = 2$. Die Partialbruchzerlegung lautet somit:

$$R(x) = \frac{2}{x} + \frac{1}{x^2} + \frac{x+2}{x^2-6x+10}.$$

Anwendung auf die Berechnung der Integrale rationaler Funktionen

Seien $P, Q \in \mathbb{R}[x]$ reelle Polynome und $I \subset \mathbb{R}$ ein Intervall, das keine Nullstellen von Q enthält. Wir wollen das Integral

$$\int \frac{P(x)}{Q(x)} dx \quad \text{auf } I$$

berechnen. Dazu gehen wir folgendermaßen vor:

a) Ist $\deg(P) \geq \deg(Q)$, so dividieren wir P durch Q und erhalten

$$\frac{P(x)}{Q(x)} = P_1(x) + \frac{P_2(x)}{Q(x)},$$

wobei $P_1, P_2 \in \mathbb{R}[x]$ reelle Polynome sind und $\deg(P_2) < \deg(Q)$ gilt.

b) Wir zerlegen $\frac{P_2(x)}{Q(x)}$ in Partialbrüche nach Satz 6.15.

c) Es bleibt nun, die Integrale der auftretenden Partialbrüche auf I zu berechnen.

Wir geben die relevanten Stammfunktionen an, die man aus den Grundintegralen nach geeigneter Substitution erhält. In den folgenden Formeln sei $k \in \mathbb{N}$ mit $k > 1$, $\lambda, p, q \in \mathbb{R}$ und $q - \frac{p^2}{4} > 0$.

$$\int \frac{1}{(x - \lambda)} dx = \ln|x - \lambda| \quad (\text{Subst. } y = x - \lambda).$$

$$\int \frac{1}{(x - \lambda)^k} dx = -\frac{1}{k-1} \cdot \frac{1}{(x - \lambda)^{k-1}} \quad (\text{Subst. } y = x - \lambda).$$

$$\int \frac{1}{x^2 + px + q} dx = \frac{1}{\sqrt{q - p^2/4}} \cdot \arctan \frac{x + p/2}{\sqrt{q - p^2/4}} \quad (\text{Subst. } y = \frac{x + p/2}{\sqrt{q - p^2/4}}).$$

$$\int \frac{2x + p}{(x^2 + px + q)} dx = \ln|x^2 + px + q| \quad (\text{nach Beispiel 1}).$$

$$\int \frac{2x + p}{(x^2 + px + q)^k} dx = -\frac{1}{k-1} \cdot \frac{1}{(x^2 + px + q)^{k-1}} \quad (\text{Subst. } y = x^2 + px + q).$$

Das noch fehlende Integral

$$\int \frac{1}{(x^2 + px + q)^k} dx$$

führt man mit Hilfe der Rekursionsformel

$$\int \frac{1}{(y^2 + 1)^k} dy = \frac{1}{(2k-2)} \cdot \frac{y}{(y^2 + 1)^{k-1}} + \frac{2k-3}{2k-2} \cdot \int \frac{1}{(y^2 + 1)^{k-1}} dy$$

schrittweise auf den arctan zurück.

Für das obige Beispiel erhalten wir mit diesen Formeln:

$$\begin{aligned} \int \frac{3x^3 - 9x^2 + 14x + 10}{x^4 - 6x^3 + 10x^2} dx &= \int \frac{2}{x} dx + \int \frac{1}{x^2} dx + \int \frac{x+2}{x^2 - 6x + 10} dx \\ &= 2 \ln|x| - \frac{1}{x} + \frac{1}{2} \int \frac{2x-6}{x^2 - 6x + 10} dx + 5 \int \frac{1}{x^2 - 6x + 10} dx \\ &= 2 \ln|x| - \frac{1}{x} + \frac{1}{2} \ln|x^2 - 6x + 10| + 5 \arctan(x-3). \end{aligned}$$

Wir betrachten nun noch zwei weitere Anwendungen des Hauptsatzes der Differential- und Integralrechnung, die man zur Approximation von Funktionen bzw. von Integralen benutzen kann. Zunächst erhalten wir eine weitere Formel für das Restglied in der Taylorentwicklung einer $(n + 1)$ -mal stetig differenzierbaren Funktion, die wir in Abschnitt 5.4. bereits angekündigt hatten.

Satz 6.16 (Integralform des Restgliedes in der Taylorentwicklung)

Sei $f : I \rightarrow \mathbb{R}$ eine $(n + 1)$ -mal stetig differenzierbare Funktion, $x_0 \in I$ und

$$f(x) = \underbrace{f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n}_{=: T_n(f, x_0)(x) \text{ } n\text{-tes Taylorpolynom von } f \text{ in } x_0} + R_n(f, x_0)(x)$$

die Taylorformel n -ter Ordnung von f in x_0 auf I . Dann gilt für das Restglied $R_n(f, x_0)(x)$

$$R_n(f, x_0)(x) = \frac{1}{n!} \int_{x_0}^x (x - t)^n f^{(n+1)}(t) dt.$$

Beweis. Wir zeigen die Behauptung durch vollständige Induktion über $n \in \mathbb{N}_0$.

Induktionsanfang: Für $n = 0$ ist die zu beweisende Formel

$$R_0(f, x_0)(x) = f(x) - f(x_0) = \int_{x_0}^x f'(t) dt$$

gerade der Hauptsatz der Differential- und Integralrechnung.

Induktionsschritt: Wir setzen voraus, dass die Behauptung für ein beliebiges aber festes $n \in \mathbb{N}_0$ gilt und beweisen sie für $(n + 1)$:

$$\begin{aligned} R_{n+1}(f, x_0)(x) &= f(x) - T_{n+1}(f, x_0)(x) \\ &= f(x) - T_n(f, x_0)(x) - \frac{f^{(n+1)}(x_0)}{(n+1)!}(x - x_0)^{n+1} \\ &= R_n(f, x_0)(x) - \frac{f^{(n+1)}(x_0)}{(n+1)!}(x - x_0)^{n+1}. \end{aligned} \quad (6.8)$$

Für das n -te Restglied $R_n(f, x_0)(x)$ erhalten wir aus der Induktionsvoraussetzung

$$\begin{aligned} R_n(f, x_0)(x) &= \frac{1}{n!} \int_{x_0}^x (x - t)^n f^{(n+1)}(t) dt \\ &= - \int_{x_0}^x \frac{d}{dt} \left(\frac{(x - t)^{n+1}}{(n+1)!} \right) f^{(n+1)}(t) dt \\ &\stackrel{\text{part. Int.}}{=} - \frac{(x - t)^{n+1}}{(n+1)!} f^{(n+1)}(t) \Big|_{x_0}^x + \frac{1}{(n+1)!} \int_{x_0}^x (x - t)^{n+1} f^{(n+2)}(t) dt \\ &= \frac{1}{(n+1)!} \int_{x_0}^x (x - t)^{n+1} f^{(n+2)}(t) dt + \frac{f^{(n+1)}(x_0)}{(n+1)!} (x - x_0)^{n+1}. \end{aligned}$$

Mit (6.8) folgt dann die Induktionsbehauptung. \square

In den meisten Fällen lässt sich das Integral einer Funktion nicht durch elementare Funktionen ausdrücken. Das tritt bereits auf, wenn man z.B. die Länge einer Ellipse berechnen will. In diesen Fällen hilft Satz 6.5, das Integral zu approximieren. Der folgende Satz erlaubt es, den Fehler bei Approximation des Riemann-Integrals durch Trapezsummen abzuschätzen.

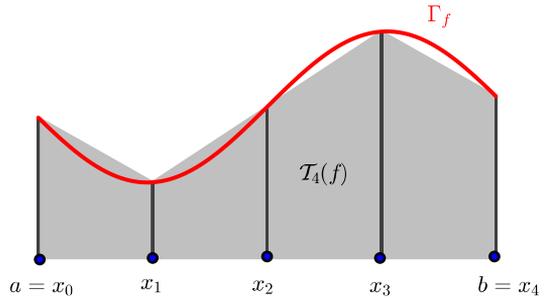
Satz 6.17 (Trapez-Regel)

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine 2-mal stetig differenzierbare Funktion, $K := \max\{f''(\xi) \mid \xi \in [a, b]\}$ das Maximum von f'' auf $[a, b]$, $\mathcal{P}_n = \{x_0, x_1, \dots, x_n\}$ die äquidistante Zerlegung von $[a, b]$ in n Teilintervalle und

$$\mathcal{T}_n(f) = \frac{b-a}{n} \cdot \sum_{k=1}^n \frac{(f(x_k) + f(x_{k-1})))}{2}$$

die Trapezsumme von f zur Schrittweite $\frac{b-a}{n}$.
Dann gilt

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \mathcal{T}_n(f)$$



mit der Fehlerabschätzung

$$\left| \int_a^b f(x) dx - \mathcal{T}_n(f) \right| \leq \frac{(b-a)^3 \cdot K}{12} \cdot \frac{1}{n^2}.$$

Beweis. 1. Die Darstellung des Riemann-Integrals als Grenzwert der Trapez-Summen $\mathcal{T}_n(f)$ wurde in Übungsaufgabe 6 (Analysis 2) bewiesen.

2. Für die Fehlerabschätzung betrachten wir zunächst einen Spezialfall und zeigen:

Ist $g : [0, 1] \rightarrow \mathbb{R}$ eine 2-mal stetig differenzierbare Funktion, so existiert ein $\xi \in [0, 1]$ mit

$$\int_0^1 g(x) dx = \frac{1}{2}(g(0) + g(1)) - \frac{1}{12}g''(\xi) = \mathcal{T}_1(g) - \frac{1}{12}g''(\xi). \tag{6.9}$$

Sei dazu $\psi : [0, 1] \rightarrow \mathbb{R}$ die Funktion $\psi(x) := \frac{1}{2}(x - x^2)$. Dann gilt $\psi \geq 0$, $\psi(0) = \psi(1) = 0$, $\psi'(x) = \frac{1}{2} - x$ und $\psi''(x) = -1$. Mit partieller Integration folgt

$$\begin{aligned} \int_0^1 \psi(x)g''(x) dx &\stackrel{p.Int.}{=} - \int_0^1 \psi'(x)g'(x) dx + \underbrace{\psi(x)g'(x)}_{=0} \Big|_0^1 \\ &\stackrel{p.Int.}{=} \int_0^1 \psi''(x)g(x) dx - \psi'(x)g(x) \Big|_0^1 \\ &= - \int_0^1 g(x) dx + \frac{1}{2}(g(1) + g(0)). \end{aligned} \tag{6.10}$$

Aus dem Mittelwertsatz der Integralrechnung (anwendbar wegen $\psi \geq 0$) folgt andererseits die Existenz eines $\xi \in [0, 1]$ mit

$$\int_0^1 \psi(x)g''(x) dx = g''(\xi) \int_0^1 \psi(x) dx = \frac{1}{2} \cdot g''(\xi) \int_0^1 (x-x^2) dx = \frac{1}{12}g''(\xi). \quad (6.11)$$

Aus (6.10) und (6.11) folgt die Behauptung (6.9).

3. Wir beweisen nun die Fehlerabschätzung des Satzes. Für die Teilungspunkte der äquidistanten Zerlegung $\mathcal{P}_n = \{x_0, x_1, \dots, x_n\}$ gilt $x_k = a + k \cdot \frac{b-a}{n}$. Sei $h := \frac{b-a}{n}$. Dann folgt mit der Substitution $\phi_k : [0, 1] \rightarrow [x_{k-1}, x_k]$,

$$\phi_k(t) := x_{k-1} + t \cdot h,$$

unter Benutzung von Schritt 2.:

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{k=1}^n \int_{x_{k-1}}^{x_k} f(x) dx \stackrel{\text{Subst.}}{=} \sum_{k=1}^n \int_0^1 f(\phi_k(t)) \cdot h dt = h \cdot \sum_{k=1}^n \int_0^1 \underbrace{(f \circ \phi_k)}_{=g}(t) dt \\ &\stackrel{(6.9)}{=} h \cdot \sum_{k=1}^n \left(\frac{f(x_k) + f(x_{k-1})}{2} - \frac{1}{12}(f \circ \phi_k)''(\xi_k) \right) \\ &= \mathcal{T}_n(f) - \frac{h}{12} \cdot \sum_{k=1}^n (f \circ \phi_k)''(\xi_k) \end{aligned}$$

für gewisse $\xi_k \in [0, 1]$. Aus der Kettenregel für die Ableitungen folgt

$$\begin{aligned} (f \circ \phi_k)'(t) &= f'(\phi_k(t)) \cdot \phi_k'(t) = f'(\phi_k(t)) \cdot h, \\ (f \circ \phi_k)''(t) &= f''(\phi_k(t)) \cdot h^2. \end{aligned}$$

Folglich gilt:

$$\begin{aligned} \left| \int_a^b f(x) dx - \mathcal{T}_n(f) \right| &= \frac{h^3}{12} \cdot \left| \sum_{k=1}^n f''(\phi_k(\xi_k)) \right| \leq \frac{h^3}{12} \cdot \sum_{k=1}^n |f''(\phi_k(\xi_k))| \\ &\leq \frac{h^3}{12} \cdot nK = \frac{(b-a)^3 \cdot K}{12} \cdot \frac{1}{n^2}. \end{aligned}$$

□

Eine andere Möglichkeit, Integrale näherungsweise zu berechnen, besteht darin, den Integranden zu approximieren.

Definition 6.7. Sei (f_n) eine Folge von Funktionen $f_n : I \rightarrow \mathbb{R}$ und $f : I \rightarrow \mathbb{R}$ eine weitere Funktion. Man sagt:

i) (f_n) konvergiert punktweise gegen f (Bez: $f_n \rightarrow f$), wenn $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ für alle $x \in I$,

ii) (f_n) konvergiert gleichmäßig gegen f (Bez: $f_n \rightrightarrows f$), wenn für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass

$$|f(x) - f_n(x)| < \varepsilon \quad \forall n \geq n_0 \quad \underline{\text{und}} \quad \forall x \in I.$$

Die gleichmäßige Konvergenz von (f_n) ist stärker als die punktweise Konvergenz. Sie bedeutet, dass die Zahlenfolge $(f_n(x))$ für jedes $x \in I$ mit der "gleichen Geschwindigkeit" gegen $f(x)$ konvergiert. Dies hat wichtige Eigenschaften der Grenzfunktion f zur Folge.

Satz 6.18 Sei (f_n) eine Folge stetiger Funktionen $f_n : I \rightarrow \mathbb{R}$, die gleichmäßig gegen $f : I \rightarrow \mathbb{R}$ konvergent. Dann ist die Grenzfunktion f ebenfalls stetig.

Beweis. Sei $x_0 \in I$ beliebig fixiert. Wir zeigen, dass f in x_0 stetig ist. Sei $\varepsilon > 0$. Da (f_n) gleichmäßig gegen f konvergiert, existiert ein $n_0 \in \mathbb{N}$ mit

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3} \quad \forall n \geq n_0, \forall x \in I.$$

Die Abbildung f_{n_0} ist in x_0 stetig, folglich gibt es ein $\delta > 0$, so dass

$$|f_{n_0}(x_0) - f_{n_0}(x)| < \frac{\varepsilon}{3} \quad \forall x \in I \text{ mit } |x_0 - x| < \delta.$$

Für alle $x \in I$ mit $|x_0 - x| < \delta$ erhalten wir damit

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_{n_0}(x)| + |f_{n_0}(x) - f_{n_0}(x_0)| + |f_{n_0}(x_0) - f(x_0)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Somit ist f in $x_0 \in I$ stetig. □

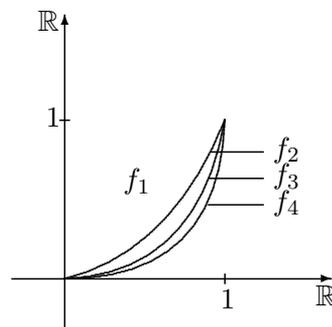
Beispiel 1: Eine nicht gleichmäßig konvergente Funktionenfolge.

Seien $f_n : [0, 1] \rightarrow \mathbb{R}$ die stetigen Funktionen $f_n(x) := x^n$.

(f_n) konvergiert punktweise gegen die Funktion

$$f(x) := \lim_{n \rightarrow \infty} f_n(x) = \begin{cases} 0 & \text{falls } 0 \leq x < 1 \\ 1 & \text{falls } x = 1 \end{cases}$$

f ist in 1 nicht stetig (insbesondere ist (f_n) nicht gleichmäßig gegen f konvergent).



Beispiel 2: Gleichmäßige Konvergenz von Potenzreihen.

Sei $f(x) := \sum_{n=1}^{\infty} a_n(x-x_0)^n$ eine reelle Potenzreihe mit positivem Konvergenzradius $R > 0$

und $p_N(x) := \sum_{n=0}^N a_n(x-x_0)^n$ ihre N -te Partialsumme. Dann konvergiert die Folge (p_N) auf jedem abgeschlossenen Intervall $[x_0 - r, x_0 + r]$ mit $0 < r < R$ gleichmäßig gegen f . (Man sagt dann, dass die Potenzreihe $f(x)$ auf $[x_0 - r, x_0 + r]$ gleichmäßig konvergiert.)

Beweis. Sei r_1 eine reelle Zahl mit $0 < r < r_1 < R$. Dann ist $f(x_0 + r_1) = \sum_{n=0}^{\infty} a_n r_1^n$ konvergent, die Folge $(a_n r_1^n)$ somit eine Nullfolge und insbesondere beschränkt. Sei $C \in \mathbb{R}^+$ mit $|a_n r_1^n| \leq C$ für alle $n \in \mathbb{N}_0$. Dann gilt für alle $x \in [x_0 - r, x_0 + r]$:

$$|a_n(x - x_0)^n| = |a_n r_1^n| \cdot \frac{|x - x_0|^n}{r_1^n} \leq C \left(\frac{r}{r_1}\right)^n = C\theta^n \quad \text{mit } \theta := \frac{r}{r_1} \in (0, 1).$$

Daraus folgt für alle $x \in [x_0 - r, x_0 + r]$:

$$\begin{aligned} |f(x) - p_N(x)| &= \left| \sum_{n=N+1}^{\infty} a_n(x - x_0)^n \right| \leq \sum_{n=N+1}^{\infty} |a_n(x - x_0)^n| \\ &\leq C \cdot \sum_{n=N+1}^{\infty} \theta^n = C \cdot \theta^{N+1} \cdot \frac{1}{1 - \theta} \quad (\text{GW geometr. Reihe}). \end{aligned}$$

Da (θ^{N+1}) eine Nullfolge ist, existiert für jedes $\varepsilon > 0$ ein $N_0 \in \mathbb{N}_0$, so dass

$$|f(x) - p_N(x)| < \varepsilon \quad \forall N \geq N_0 \quad \text{und } x \in [x_0 - r, x_0 + r],$$

(p_N) konvergiert somit auf $[x_0 - r, x_0 + r]$ gleichmäßig gegen f . \square

Satz 6.19 (Vertauschbarkeit von Limes und Integral)

Sei (f_n) eine Folge stetiger³ Funktionen $f_n : [a, b] \rightarrow \mathbb{R}$, die gleichmäßig gegen eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ konvergiert. Dann ist f Riemann-integrierbar und für das Riemann-Integral gilt:

$$\int_a^b f(x) dx = \int_a^b \lim_{n \rightarrow \infty} f_n(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

Beweis. Da (f_n) gleichmäßig gegen f konvergiert, ist f stetig und somit Riemann-integrierbar. Nach Definition existiert zu beliebig gegebenem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass

$$|f(x) - f_n(x)| < \frac{\varepsilon}{3(b-a)} \quad \forall n \geq n_0 \quad \text{und } \forall x \in [a, b]. \quad (6.12)$$

Wir approximieren die Riemann-Integrale durch Riemannsche Summen.

Sei $\mathcal{P}_m = \{x_0, x_1, \dots, x_m\}$ die äquidistante Zerlegung von $[a, b]$ in m Teilintervalle und ξ_m das Stützstellentupel $\xi_m = (x_1, x_2, \dots, x_m)$. Dann gilt für die Riemannschen Summen

$$\begin{aligned} |S(f, \mathcal{P}_m, \xi_m) - S(f_n, \mathcal{P}_m, \xi_m)| &= \left| \sum_{k=1}^m (f(x_k) - f_n(x_k)) \cdot \frac{b-a}{m} \right| \\ &\leq \sum_{k=1}^m |f(x_k) - f_n(x_k)| \cdot \frac{b-a}{m} \\ &\stackrel{(6.12)}{<} \frac{\varepsilon}{3(b-a)} \cdot (b-a) = \frac{\varepsilon}{3} \end{aligned} \quad (6.13)$$

³ Dieser Satz gilt auch allgemeiner für eine Folge gleichmäßig konvergenter Riemann-integrierbarer Funktionen. Die Grenzfunktion ist dann auch Riemann-integrierbar.

für alle $n \geq n_0$ und $m \in \mathbb{N}$. Da die Riemannschen Summen von f bzw. f_n gegen die jeweiligen Riemann-Integrale konvergieren (Satz 6.5), existiert ein $m_0 \in \mathbb{N}$ und ein $m_1(n) \in \mathbb{N}$, so dass

$$\left| \int_a^b f(x) dx - S(f, \mathcal{P}_m, \xi_m) \right| < \frac{\varepsilon}{3} \quad \forall m \geq m_0 \quad (6.14)$$

und

$$\left| \int_a^b f_n(x) dx - S(f_n, \mathcal{P}_m, \xi_m) \right| < \frac{\varepsilon}{3} \quad \forall m \geq m_1(n). \quad (6.15)$$

Für alle $n \geq n_0$ folgt aus (6.13), (6.14) und (6.15) mit einem beliebig gewählten $m \geq \max(m_0, m_1(n))$:

$$\begin{aligned} \left| \int_a^b f(x) dx - \int_a^b f_n(x) dx \right| &\leq \left| \int_a^b f(x) dx - S(f, \mathcal{P}_m, \xi_m) \right| + \left| S(f, \mathcal{P}_m, \xi_m) - S(f_n, \mathcal{P}_m, \xi_m) \right| + \\ &\quad + \left| S(f_n, \mathcal{P}_m, \xi_m) - \int_a^b f_n(x) dx \right| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Dies zeigt die Behauptung des Satzes. \square

Als Spezialfall erhalten wir

Satz 6.20 Ist $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine reelle Potenzreihe mit positivem Konvergenzradius $R > 0$ und $\psi : [c, d] \rightarrow \mathbb{R}$ eine stetige Funktion mit Bild im Konvergenzintervall $(x_0 - R, x_0 + R)$ von f . Dann gilt:

$$\int_c^d f(\psi(t)) dt = \int_c^d \sum_{n=0}^{\infty} a_n(\psi(t) - x_0)^n dt = \sum_{n=0}^{\infty} a_n \int_c^d (\psi(t) - x_0)^n dt.$$

Beweis. Da ψ stetig ist, ist $\psi([c, d])$ ein abgeschlossenes Intervall in Konvergenzintervall von f , d.h. es existiert ein $0 < r < R$ so dass $\psi([c, d]) \subset [x_0 - r, x_0 + r]$. Bezeichne (p_N) die Partialsummenfolge von f . Da (p_N) auf $[x_0 - r, x_0 + r]$ gleichmäßig gegen f konvergiert, konvergiert auch $p_n \circ \psi$ auf $[c, d]$ gleichmäßig gegen $f \circ \psi$. Dann folgt die Behauptung aus Satz 6.19. \square

Beispiel: Berechne das Integral $I = \int_0^{2\pi} \sqrt{1 - k^2 \cos^2(t)} dt$ für $0 < k < 1$.

Dieses Integral können wir nicht durch elementare Funktionen ausdrücken (es gehört zu den sogenannten *elliptischen Integralen*). Wir werden es bei der Berechnung der Länge einer Ellipse wiedertreffen.

Wir geben I durch Reihenentwicklung des Integranden an. Wir betrachten die Funktion $f(x) = \sqrt{1+x}$. Dann ist der Integrand von I gegeben durch $f \circ \psi$ mit $\psi(t) = -k^2 \cos^2(t)$. In Kapitel ?? haben wir die Taylorentwicklung von f um $x_0 = 0$ bestimmt (Beispiel 5, Abschnitt 5.5) und dabei die Binomialreihe $B_{\frac{1}{2}}(x)$ erhalten: Es gilt

$$f(x) = \sqrt{1+x} = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} x^n = B_{\frac{1}{2}}(x) \quad \forall x \in (-1, 1).$$

Nach Definition ist $\binom{\frac{1}{2}}{0} = 1$ und $\binom{\frac{1}{2}}{1} = \frac{1}{2}$. Für $n \geq 2$ berechnet man

$$\begin{aligned} \binom{\frac{1}{2}}{n} &= \frac{\frac{1}{2} \cdot (\frac{1}{2} - 1) \cdot (\frac{1}{2} - 2) \cdot \dots \cdot (\frac{1}{2} - (n-1))}{n!} = \frac{1}{2^n} \cdot \frac{1 \cdot (1-2) \cdot (1-4) \cdot \dots \cdot (1-2(n-1))}{1 \cdot 2 \cdot \dots \cdot n} \\ &= (-1)^{n-1} \cdot \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-3)}{2 \cdot 4 \cdot 6 \cdot \dots \cdot (2n)}. \end{aligned}$$

Aus Satz 6.20 folgt damit:

$$\begin{aligned} I &= \int_0^{2\pi} \sqrt{1 - k^2 \cos^2(t)} dt = \int_0^{2\pi} \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-1)^n k^{2n} \cos^{2n}(t) dt \\ &= \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-1)^n \cdot \int_0^{2\pi} \cos^{2n}(t) dt \cdot k^{2n} \\ &\stackrel{\text{Aufg. 9}}{=} 2\pi \left(1 - \sum_{n=1}^{\infty} \left[\left(\frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)}{2 \cdot 4 \cdot 6 \cdot \dots \cdot (2n)} \right)^2 \cdot \frac{1}{2n-1} \cdot k^{2n} \right] \right) \\ &= 2\pi \left(1 - \frac{1}{4} k^2 - \frac{3}{64} k^4 - \frac{5}{256} k^6 - \dots \right). \end{aligned}$$

6.3 Uneigentliche Riemann-Integrale

Bisher haben wir das Riemann-Integral für Funktionen über *abgeschlossenen* Intervallen definiert. Wir wollen die Integraldefinition jetzt auf offene, halb-offene bzw. unendliche Intervalle ausweiten.

Definition 6.8. Eine Funktion $f : [a, b) \subset \mathbb{R} \rightarrow \mathbb{R}$ mit $a < b \leq +\infty$ heißt *uneigentlich Riemann-integrierbar*, wenn $f|_{[a, s]}$ für alle abgeschlossenen Intervalle $[a, s] \subset [a, b)$ Riemann-integrierbar ist und der Grenzwert

$$\int_a^b f(x) dx := \lim_{s \rightarrow b^-} \int_a^s f(x) dx$$

existiert.

Eine Funktion $f : (a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ mit $-\infty \leq a < b$ heißt *uneigentlich Riemann-integrierbar*, wenn $f|_{[s, b]}$ für alle abgeschlossenen Intervalle $[s, b] \subset (a, b]$ Riemann-integrierbar ist und der Grenzwert

$$\int_a^b f(x) dx := \lim_{s \rightarrow a^+} \int_s^b f(x) dx$$

existiert.

Eine Funktion $f : (a, b) \subset \mathbb{R} \rightarrow \mathbb{R}$ mit $-\infty \leq a < b \leq +\infty$ heißt *uneigentlich Riemann-integrierbar*, wenn $f|_{[s,t]}$ für jedes abgeschlossene Intervall $[s, t] \subset (a, b)$

Riemann-integrierbar ist und die uneigentlichen Riemann-Integrale $\int_{x_0}^b f(x) dx$ und $\int_a^{x_0} f(x) dx$

für ein $x_0 \in (a, b)$ existieren. Wir definieren dann

$$\int_a^b f(x) dx := \int_a^{x_0} f(x) dx + \int_{x_0}^b f(x) dx.$$

(Dies ist unabhängig von der Wahl von x_0 .)

Die Rechenregeln für Riemann-Integrale (Linearität, Monotonie, Aufspaltungseigenschaft, ...) übertragen sich mit Hilfe der Grenzwertsätze für Folgen auf die uneigentlichen Integrale.

Beispiel 1: Die Gamma-Funktion

Die Gamma-Funktion $\Gamma : \mathbb{R}^+ \rightarrow \mathbb{R}$ ist definiert durch

$$\Gamma(x) := \int_0^{\infty} t^{x-1} e^{-t} dt.$$

Wir zeigen, dass das uneigentliche Integral $\Gamma(x)$ für jedes $x \in \mathbb{R}^+$ existiert:

(1) Für $t > 0$ gilt $0 < t^{x-1} e^{-t} < t^{x-1}$. Für $0 < \alpha < 1$ erhalten wir damit

$$0 < \int_{\alpha}^1 t^{x-1} e^{-t} dt < \int_{\alpha}^1 t^{x-1} dt = \frac{1}{x} t^x \Big|_{\alpha}^1 = \frac{1}{x} - \frac{\alpha^x}{x} < \frac{1}{x}.$$

Da die Funktion $g(\alpha) = \int_{\alpha}^1 t^{x-1} e^{-t} dt$ für $\alpha \rightarrow 0^+$ monoton wachsend und durch $\frac{1}{x}$ beschränkt ist, existiert der Grenzwert

$$\lim_{\alpha \rightarrow 0^+} \int_{\alpha}^1 t^{x-1} e^{-t} dt = \int_0^1 t^{x-1} e^{-t} dt.$$

(2) Es gilt $t^{x-1} e^{-t} = \frac{t^{x+1}}{e^t} \cdot \frac{1}{t^2}$. Da $\lim_{t \rightarrow \infty} \frac{t^{x+1}}{e^t} = 0$, existiert ein $M > 0$, so dass

$$t^{x-1} e^{-t} \leq \frac{1}{t^2} \quad \forall t \geq M.$$

Daraus folgt für alle $\beta > M$

$$\int_M^{\beta} t^{x-1} e^{-t} dt \leq \int_M^{\beta} \frac{1}{t^2} dt = -\frac{1}{t} \Big|_M^{\beta} = \frac{1}{M} - \frac{1}{\beta} < \frac{1}{M}.$$

Da $h(\beta) = \int_M^\beta t^{x-1} e^{-t} dt$ für $\beta \rightarrow \infty$ monoton wachsend und durch $\frac{1}{M}$ beschränkt ist, existiert der Grenzwert

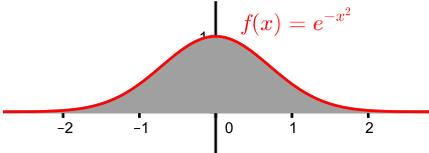
$$\lim_{\beta \rightarrow \infty} \int_M^\beta t^{x-1} e^{-t} dt = \int_M^\infty t^{x-1} e^{-t} dt.$$

Folglich existiert

$$\int_0^\infty t^{x-1} e^{-t} dt = \lim_{\alpha \rightarrow 0^+} \int_\alpha^1 t^{x-1} e^{-t} dt + \int_1^M t^{x-1} e^{-t} dt + \lim_{\beta \rightarrow \infty} \int_M^\beta t^{x-1} e^{-t} dt = \Gamma(x).$$

Die Gamma-Funktion spielt eine wichtige Rolle in der Zahlentheorie und der komplexen Analysis⁴. Es gilt z.B. $\Gamma(n) = (n-1)!$ für alle natürlichen Zahlen (Übungsaufgabe), d.h. die Gamma-Funktion setzt die Fakultäten auf \mathbb{R}^+ fort. Man findet auch π in der Gamma-Funktion wieder. Es gilt: $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Wir benutzen dies, um das in der Stochastik wichtige Integral über die Gaußsche Glockenfunktion $f(x) := e^{-x^2}$ zu berechnen.

Beispiel 2: Das Gauß-Integral $\int_{-\infty}^\infty e^{-x^2} dx = \sqrt{\pi}$:



Mit Hilfe der Gamma-Funktion berechnen wir mit der Substitution $t = x^2$, $dt = 2x dx$:

$$\int_0^\infty e^{-x^2} dx = \int_0^\infty e^{-t} \frac{1}{2\sqrt{t}} dt = \frac{1}{2} \int_0^\infty t^{-\frac{1}{2}} e^{-t} dt = \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}.$$

Da die Gaußsche Glockenfunktion eine gerade Funktion ist, folgt für das Gauß-Integral

$$\int_{-\infty}^\infty e^{-x^2} dx = 2 \cdot \int_0^\infty e^{-x^2} dx = \sqrt{\pi}.$$

Als Anwendung des uneigentlichen Riemann-Integrals beweisen noch wir ein Kriterium für die Konvergenz von Reihen:

Satz 6.21 (Integral-Vergleichskriterium für Reihen)

Sei $f : [n_0, \infty) \rightarrow \mathbb{R}$ eine monoton-fallende, positive Funktion, wobei $n_0 \in \mathbb{N}_0$. Dann gilt:

$$\sum_{n=n_0}^\infty f(n) \text{ konvergiert} \iff \int_{n_0}^\infty f(x) dx \text{ existiert.}$$

Im Fall der Konvergenz gilt

$$0 \leq \sum_{n=n_0}^\infty f(n) - \int_{n_0}^\infty f(x) dx \leq f(n_0).$$

⁴ Wichtige Eigenschaften der Gamma-Funktion werden z.B. im Buch von O. Forster: *Analysis 1*, 11. Auflage, § 20, bewiesen.

Beweis. Sei $n \in \mathbb{N}$ beliebig mit $n \geq n_0$. Da f monoton fallend ist, gilt $f(n) \geq f(x) \geq f(n+1)$ für alle $x \in [n, n+1]$. Aus der Monotonie des Integrals folgt

$$f(n) = \int_n^{n+1} f(n) dx \geq \int_n^{n+1} f(x) dx \geq \int_n^{n+1} f(n+1) dx = f(n+1).$$

Daraus folgt für die N -te Partialsumme der Reihe

$$\sum_{n=n_0}^N f(n) \geq \int_{n_0}^{N+1} f(x) dx \geq \sum_{n=n_0}^N f(n+1) = \sum_{n=n_0+1}^{N+1} f(n).$$

Die erste Ungleichung zeigt Behauptung \implies , die 2. Ungleichung zeigt Behauptung \impliedby . Außerdem gilt im Fall der Konvergenz

$$\sum_{n=n_0}^{\infty} f(n) - f(n_0) \leq \int_{n_0}^{\infty} f(x) dx \leq \sum_{n=n_0}^{\infty} f(n)$$

und damit die Abschätzung. □

6.4 Geometrische Anwendungen des Riemann-Integrals

Wir werden die Integralrechnung jetzt benutzen, um die Länge von Kurven sowie Flächeninhalte "komplizierterer" ebener Gebiete zu bestimmen.

Zunächst erinnern wir an einige Kenntnisse aus der Vorlesung *Lineare Algebra und analytische Geometrie I*. \mathbb{R}^n bezeichnet wie üblich die Menge der n -Tupel reeller Zahlen

$$\mathbb{R}^n := \{(x_1, x_2, \dots, x_n) \mid x_k \in \mathbb{R} \text{ für } k = 1, \dots, n\}.$$

Auf dem Vektorraum \mathbb{R}^n ist das Euklidische Skalarprodukt $\langle \cdot, \cdot \rangle$ und die Euklidische Norm $\|\cdot\|$ definiert. Für Vektoren $x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ ist

$$\begin{aligned} \langle x, y \rangle &:= x_1 y_1 + x_2 y_2 + \dots + x_n y_n, \\ \|x\| &:= \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}. \end{aligned}$$

Es gilt die *Dreiecksungleichung* (Δ)

$$\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in \mathbb{R}^n,$$

sowie die *Cauchy-Schwarzsche Ungleichung* (CSU)

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| \quad \forall x, y \in \mathbb{R}^n.$$

Der Winkel $\angle(x, y)$ zwischen zwei Vektoren $x, y \in \mathbb{R}^n \setminus \{0\}$ ist die eindeutig bestimmte Zahl $\alpha \in [0, \pi]$ mit

$$\cos \alpha = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}.$$

Legt man in der Ebene, dem 3-dimensionalen (oder höher-dimensionalen Raum) ein kartesisches Koordinatensystem fest, so kann man sie ebenfalls durch \mathbb{R}^2 , \mathbb{R}^3 (bzw. \mathbb{R}^n) beschreiben, indem man jedem Punkt seine kartesischen Koordinaten zuordnet.

In gesamten Abschnitt 6.4 bezeichnet I wieder ein Intervall, das aus mehr als einem Punkt besteht.

6.4.1 Parametrisierte Kurven und ihre Länge

Wir verwenden einen Kurvenbegriff, der in der Kinematik wurzelt. Er ist die mathematische Abstraktion der Bewegung eines Punktes in der Ebene \mathbb{R}^2 oder im Raum \mathbb{R}^3 , die durch die Angabe des Ortes $\gamma(t)$ zum Zeitpunkt t beschrieben wird.

Definition 6.9. Eine parametrisierte Kurve im \mathbb{R}^n ist eine Abbildung $\gamma : I \rightarrow \mathbb{R}^n$ mit

$$t \in I \mapsto \gamma(t) =: (\gamma_1(t), \gamma_2(t), \dots, \gamma_n(t)) \in \mathbb{R}^n,$$

deren Komponentenfunktionen $\gamma_k : I \rightarrow \mathbb{R}$, $k = 1, \dots, n$, stetig sind.

$\gamma : I \rightarrow \mathbb{R}^n$ heißt differenzierbar (stetig differenzierbar, ...), wenn jede der Komponentenfunktionen γ_k die entsprechende Eigenschaft hat.

Die Elemente von I heißen die Parameter von γ . Die Bildmenge $\Gamma := \gamma(I) \subset \mathbb{R}^n$ nennt man auch die Spur von γ . γ heißt dann eine Parametrisierung der Menge $\Gamma \subset \mathbb{R}^n$.

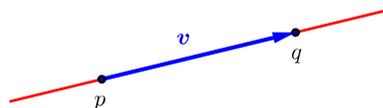
Statt parametrisierter Kurve sagen wir auch kurz nur Kurve.

Bei unserem Kurvenbegriff interessieren wir uns also nicht nur für die Spur der Kurve im \mathbb{R}^n , sondern auch dafür, wie die Kurve durchlaufen wird (z.B. mit welchem Zeitplan).

Beispiel 1 Geraden im \mathbb{R}^n

Sei $L \subset \mathbb{R}^n$ eine Gerade, $p, q \in L$ zwei verschiedene Punkte und $v := q - p$ der Verbindungsvektor von p nach q . Eine Parametrisierung von L ist gegeben durch $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ mit

$$\gamma(t) := p + t \cdot v.$$



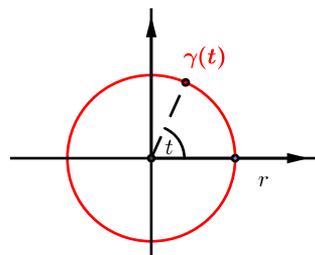
Beispiel 2 Der Kreis vom Radius r

Sei $K_r := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2\}$ der Kreis vom Radius r . Man kann ihn auf verschiedene Weise parametrisieren:

$$\gamma(t) := (r \cos t, r \sin t), \quad t \in [0, 2\pi],$$

$$\delta(t) := (r \cos t, -r \sin t), \quad t \in [0, 2\pi],$$

$$\eta(t) := (r \cos 2t, r \sin 2t), \quad t \in [0, 2\pi],$$



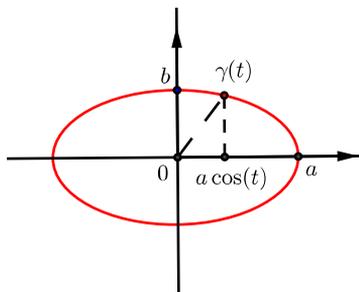
sind drei verschiedene Parametrisierungen des Kreises, die verschiedene Durchlaufeschaften beschreiben.

Beispiel 3 Die Ellipse mit den Halbachsen a und b

Sei $E := \{(x, y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\}$ die Ellipse mit den Halbachsen $a, b > 0$. Die Kurve

$$\gamma(t) := (a \cos t, b \sin t), \quad t \in [0, 2\pi],$$

ist eine Parametrisierung von E .

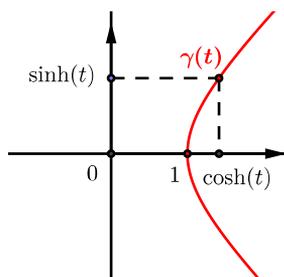


Beispiel 4 Die Hyperbel

Sei $H := \{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 = 1, x > 0\}$ die Hyperbel. Die Kurve

$$\gamma(t) := (\cosh t, \sinh t), \quad t \in \mathbb{R},$$

ist eine Parametrisierung von H .

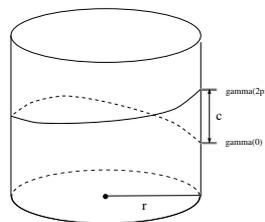


Beispiel 5: Die Schraubenlinie

Wir betrachten die Schraubenlinie $\gamma : \mathbb{R} \rightarrow \mathbb{R}^3$, gegeben durch

$$\gamma(t) := (r \cos t, r \sin t, \frac{c}{2\pi}t), \quad t \in \mathbb{R},$$

wobei $r \in \mathbb{R}^+$ und $c \in \mathbb{R}$ Konstanten sind. Die Spur von γ schraubt sich mit konstanter Ganghöhe c um den Zylinder vom Radius r .

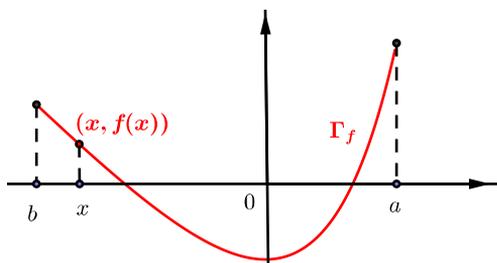


Beispiel 6: Ein Funktionsgraph

Sei $f : I \rightarrow \mathbb{R}$ eine stetige Funktion und $\Gamma_f := \{(x, f(x)) \in \mathbb{R}^2 \mid x \in I\}$ ihr Funktionsgraph. Γ_f kann man offensichtlich durch die Kurve

$$\gamma(x) := (x, f(x)), \quad x \in I,$$

parametrisieren.



Das Bild einer Kurve $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ kann ein ganzes Quadrat ausfüllen (*Peano-Kurven*). Beispiele dafür findet man in K. Königsberger: *Analysis I*, Kapitel 12, Aufgabe 14, oder in H. Sagan: *Space-Filling Curves*. Solche (pathologischen) Fälle wollen wir ausschließen. Wir setzen deshalb im Folgenden voraus, dass unsere Kurven $\gamma : I \rightarrow \mathbb{R}^n$ nicht nur stetig, sondern differenzierbar sind. In diesem Fall können wir auch Tangenten an Kurven und Winkel zwischen Kurven analytisch beschreiben.

Definition 6.10. Sei $\gamma : I \rightarrow \mathbb{R}^n$ eine differenzierbare Kurve. Dann heißt

$$\gamma'(t) := (\gamma'_1(t), \gamma'_2(t), \dots, \gamma'_n(t))$$

Tangentialvektor von γ im Parameter $t \in I$. Die Kurve γ heißt im Parameter $t \in I$ regulär, wenn $\gamma'(t) \neq 0$. γ heißt regulär, wenn sie in jedem Parameter regulär ist.

Wir definieren für Folgen im \mathbb{R}^n den folgenden Konvergenzbegriff:

Definition 6.11. Eine Folge $(x_m = (x_{m1}, x_{m2}, \dots, x_{mn}))_{m=1}^{\infty}$ konvergiert gegen $z = (z_1, z_2, \dots, z_n)$, wenn die Folge der Komponenten $(x_{mk})_{m=1}^{\infty}$ gegen z_k konvergiert, d.h.

$$\lim_{m \rightarrow \infty} x_m = z \quad :\iff \quad \lim_{m \rightarrow \infty} x_{mk} = z_k \quad \forall k \in \{1, \dots, n\}.$$

Dann gilt für den Tangentialvektor:

$$\gamma'(t) = \lim_{h \rightarrow 0} \frac{\gamma(t+h) - \gamma(t)}{h}.$$

Ist $\gamma'(t) \neq 0$, so beschreibt die Gerade

$$\text{Tan}_t \gamma := \gamma(t) + \mathbb{R} \cdot \gamma'(t)$$

die Tangente von γ im Parameter t .

Sie entsteht als Grenzgerade der Sekanten $\text{Sek}_t(h)$ durch $\gamma(t)$ und $\gamma(t+h)$ bei $h \rightarrow 0$:

$$\text{Sek}_t(h) := \gamma(t) + \mathbb{R} \cdot \frac{\gamma(t+h) - \gamma(t)}{h}.$$

Sind γ und δ zwei reguläre differenzierbare Kurven, die sich im Punkt $p \in \mathbb{R}^n$ schneiden, d.h. $p = \gamma(t_0) = \delta(s_0)$, dann versteht man unter dem Schnittwinkel von γ und δ im Punkt p den Winkel zwischen den Tangentialvektoren $\gamma'(t_0)$ und $\delta'(s_0)$:

$$\angle_p(\gamma, \delta) := \angle(\gamma'(t_0), \delta'(s_0)) = \arccos \frac{\langle \gamma'(t_0), \delta'(s_0) \rangle}{\|\gamma'(t_0)\| \cdot \|\delta'(s_0)\|}.$$

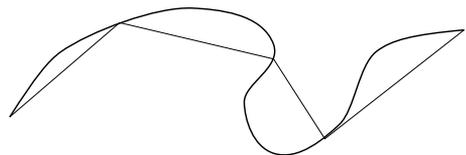
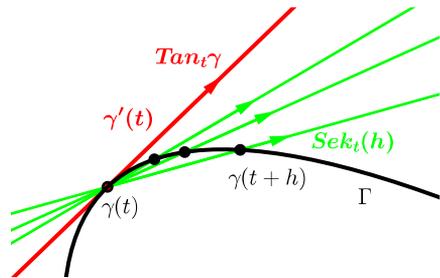
Um die Länge einer Kurve $\gamma : I \rightarrow \mathbb{R}^n$ zu definieren, benutzen wir die geometrische Intuition. Bereits im Altertum haben Mathematiker den Kreisumfang berechnet, in dem sie ihn durch einbeschriebene reguläre n -Ecke approximiert haben. Für allgemeine Kurven knüpft man an dieses Verfahren an.

Sei $\mathcal{P} = \{t_0, t_1, \dots, t_m\}$ eine endliche Menge von Teilungspunkten des Intervalls I mit $t_0 < t_1 < \dots < t_m$. Ist $I = [a, b]$, so setzt man $t_0 = a$ und $t_m = b$. Dann beschreibt

$$L(\gamma, \mathcal{P}) := \sum_{k=1}^m \|\gamma(t_k) - \gamma(t_{k-1})\|$$

die Länge des durch die Zerlegung \mathcal{P} definierten Sehnenpolygons durch die Punkte $\gamma(t_0), \gamma(t_1), \dots, \gamma(t_m)$.

Ist $\tilde{\mathcal{P}} \supseteq \mathcal{P}$, so folgt aus der Dreiecksungleichung für die Norm $\|\cdot\|$, dass $L(\gamma, \tilde{\mathcal{P}}) \geq L(\gamma, \mathcal{P})$.



Definition 6.12. Eine Kurve $\gamma : I \rightarrow \mathbb{R}^n$ heißt rektifizierbar, wenn die Menge der Längen aller einbeschriebenen Sehnenpolygone beschränkt ist. In diesem Fall nennt man

$$L(\gamma) := \sup\{L(\gamma, \mathcal{P}) \mid \mathcal{P} \text{ endliche Zerlegung von } I\}$$

die Länge von γ .

Ist γ rektifizierbar und $I = I_1 \cup I_2$ eine Zerlegung von I in zwei Teilintervalle, so sind $\gamma|_{I_1}$ und $\gamma|_{I_2}$ ebenfalls rektifizierbar und es gilt

$$L(\gamma) = L(\gamma|_{I_1}) + L(\gamma|_{I_2}).$$

Nicht jede differenzierbare Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^n$ ist rektifizierbar (ein Beispiel werden wir in den Übungen sehen), aber stetig differenzierbare Kurven sind dies immer und ihre Länge kann man mit Hilfe des Riemann-Integrals berechnen. Um dies zu beweisen, machen wir zunächst eine Vorbetrachtung.

Sei $f : [a, b] \rightarrow \mathbb{R}^n$ eine stetige⁵ Abbildung mit den Komponenten $f = (f_1, \dots, f_n)$. Dann ist $\|f\| : [a, b] \rightarrow \mathbb{R}$ ebenfalls stetig. Wir definieren das Riemann-Integral von f komponentenweise durch

$$\int_a^b f(x) dx := \left(\int_a^b f_1(x) dx, \dots, \int_a^b f_n(x) dx \right).$$

Satz 6.22 Sei $f : [a, b] \rightarrow \mathbb{R}^n$ stetig. Dann gilt

$$\left\| \int_a^b f(x) dx \right\| \leq \int_a^b \|f(x)\| dx.$$

Beweis. Für den Beweis benutzen wir die Cauchy-Schwarzsche Ungleichung.

Sei $v := \int_a^b f(x) dx =: (v_1, \dots, v_n) \in \mathbb{R}^n$. Dann gilt:

$$\begin{aligned} \|v\|^2 &= \langle v, v \rangle = \sum_{k=1}^n v_k \cdot \int_a^b f_k(x) dx = \int_a^b \sum_{k=1}^n v_k \cdot f_k(x) dx = \int_a^b \langle v, f(x) \rangle dx \\ &\stackrel{CSU}{\leq} \int_a^b \|v\| \cdot \|f(x)\| dx = \|v\| \cdot \int_a^b \|f(x)\| dx. \end{aligned}$$

Ist $v = 0$, so ist die Behauptung des Satzes offensichtlich erfüllt. Ist $v \neq 0$, so teilen wir durch $\|v\| \neq 0$ und erhalten die Behauptung. \square

Satz 6.23 Sei $\gamma : [a, b] \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Kurve. Dann ist γ rektifizierbar und es gilt

$$L(\gamma) = \int_a^b \|\gamma'(t)\| dt.$$

⁵ d.h. alle Komponentenfunktionen seien stetig

Beweis. (1) Wir zeigen zuerst, dass $\int_a^b \|\gamma'(t)\| dt$ eine obere Schranke für die Längen aller einbeschriebenen Sehnenpolygone ist. Insbesondere ist γ dann rektifizierbar.

Sei $\mathcal{P} = \{t_0, t_1, \dots, t_m\}$ eine beliebige Zerlegung von $[a, b]$. Dann gilt

$$\begin{aligned} L(\gamma, \mathcal{P}) &= \sum_{k=1}^m \|\gamma(t_k) - \gamma(t_{k-1})\| = \sum_{k=1}^m \left\| \int_{t_{k-1}}^{t_k} \gamma'(t) dt \right\| \\ &\stackrel{S.6.22}{\leq} \sum_{k=1}^m \int_{t_{k-1}}^{t_k} \|\gamma'(t)\| dt = \int_a^b \|\gamma'(t)\| dt. \end{aligned}$$

Insbesondere gilt

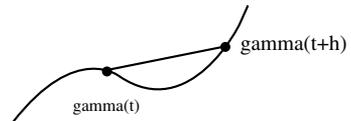
$$L(\gamma) := \sup\{L(\gamma, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } [a, b]\} \leq \int_a^b \|\gamma'(t)\| dt. \quad (6.16)$$

(2) Wir zeigen nun, dass in (6.16) Gleichheit gilt. Wir betrachten dazu die Hilfsfunktion $\ell : [a, b] \rightarrow [0, L(\gamma)]$,

$$\ell(t) := L(\gamma|_{[a,t]}) = \sup\{L(\gamma|_{[a,t]}, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } [a, t]\}.$$

Sei $t \in [a, b]$, $h > 0$ und $t+h \in [a, b]$. Wir wenden (6.16) auf $\gamma|_{[t, t+h]}$ an und erhalten

$$\|\gamma(t+h) - \gamma(t)\| \leq L(\gamma|_{[t, t+h]}) \leq \int_t^{t+h} \|\gamma'(s)\| ds.$$



Wegen

$$L(\gamma|_{[t, t+h]}) = L(\gamma|_{[a, t+h]}) - L(\gamma|_{[a, t]}) = \ell(t+h) - \ell(t)$$

folgt dann

$$\left\| \frac{\gamma(t+h) - \gamma(t)}{h} \right\| \leq \frac{\ell(t+h) - \ell(t)}{h} \leq \frac{1}{h} \int_t^{t+h} \|\gamma'(s)\| ds = \frac{1}{h} (F(t+h) - F(t)),$$

wobei F die Stammfunktion von $\|\gamma'\|$ bezeichnet. Durch Limesbildung $\lim_{h \rightarrow 0^+}$ ergibt sich

$$\|\gamma'(t)\| \leq \lim_{h \rightarrow 0^+} \frac{\ell(t+h) - \ell(t)}{h} \leq F'(t) = \|\gamma'(t)\|.$$

Also gilt

$$\lim_{h \rightarrow 0^+} \frac{\ell(t+h) - \ell(t)}{h} = \|\gamma'(t)\|.$$

Auf analoge Weise zeigt man

$$\lim_{h \rightarrow 0^-} \frac{\ell(t+h) - \ell(t)}{h} = \|\gamma'(t)\|.$$

Also ist ℓ differenzierbar und es gilt $\ell'(t) = \|\gamma'(t)\|$. Damit erhalten wir

$$L(\gamma) = \ell(b) = \ell(b) - \ell(a) = \int_a^b \ell'(t) dt = \int_a^b \|\gamma'(t)\| dt.$$

□

Bemerkung:

1. Sei $I \subset \mathbb{R}$ ein offenes, halboffenes oder unbeschränktes Intervall mit den Grenzen a und b , $-\infty \leq a < b \leq +\infty$, und $\gamma : I \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Kurve. Dann ist γ genau dann rektifizierbar, wenn das uneigentliche Riemann-Integral $\int_a^b \|\gamma'(t)\| dt$ existiert. In diesem Fall gilt wie für abgeschlossene Intervalle $L(\gamma) = \int_a^b \|\gamma'(t)\| dt$. Ist der Grenzwert des uneigentlichen Riemann-Integrals $+\infty$, so sagen wir γ hat unendliche Länge.
2. Man nennt $\gamma : I \rightarrow \mathbb{R}^n$ stückweise stetig differenzierbar, wenn γ stetig ist und es endlich viele Teilungspunkte $a = t_0 < t_1 < t_2 < \dots < t_m = b$ des Definitionsbereiches I gibt, so dass $\gamma|_{(t_{k-1}, t_k)}$ stetig differenzierbar ist. Dann gilt

$$L(\gamma) = \sum_{k=1}^m L(\gamma|_{(t_{k-1}, t_k)}) = \sum_{k=1}^m \int_{t_{k-1}}^{t_k} \|\gamma'(t)\| dt.$$

Beispiel 1: *Kreisumfang, Länge eines Kreisbogens*

Wir betrachten den Kreis vom Radius r , parametrisiert durch $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$ mit $\gamma(t) := (r \cos t, r \sin t)$. Dann ist

$$\gamma'(t) = (-r \sin t, r \cos t) \quad \text{und} \quad \|\gamma'(t)\| = r.$$

Folglich erhalten wir für die Länge

$$L(\gamma) = \int_0^{2\pi} r dt = 2\pi r \quad \text{sowie} \quad L(\gamma|_{[0, \varphi]}) = \varphi \cdot r.$$

Beispiel 2: *Länge der Schraubenlinie*

Wir betrachten die Schraubenlinie $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^3$, gegeben durch $\gamma(t) = (r \cos t, r \sin t, ct)$. Dann ist

$$\gamma'(t) = (-r \sin t, r \cos t, c) \quad \text{und} \quad \|\gamma'(t)\| = \sqrt{r^2 + c^2}.$$

Folglich erhalten wir für die Länge der Schraubenlinie bei einer Umdrehung

$$L(\gamma) = 2\pi \sqrt{r^2 + c^2}.$$

Beispiel 3: *Umfang der Ellipse*

Wir betrachten die Ellipse $E := \{(x, y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\}$, wobei $a, b \in \mathbb{R}^+$. Es gelte $a > b$. Sei $k^2 := 1 - \frac{b^2}{a^2} < 1$. Wir parametrisieren die Ellipse E durch die Kurve $\gamma : [0, 2\pi) \rightarrow \mathbb{R}^2$ mit $\gamma(t) = (a \cos t, b \sin t)$. Dann gilt

$$\gamma'(t) = (-a \sin t, b \cos t) \quad \text{und} \quad \|\gamma'(t)\| = \sqrt{a^2 \sin^2 t + b^2 \cos^2 t}$$

und wir erhalten

$$L(\gamma) = a \int_0^{2\pi} \sqrt{\sin^2 t + \frac{b^2}{a^2} \cos^2 t} dt = a \int_0^{2\pi} \sqrt{1 - k^2 \cos^2 t} dt.$$

Das Integral hatten wir in Abschnitt 6.2. mittels Reihenentwicklung berechnet. Es folgt

$$L(\gamma) = 2\pi a \cdot \left(1 - \sum_{n=1}^{\infty} \left[\left(\frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)}{2 \cdot 4 \cdot 6 \cdot \dots \cdot (2n)} \right)^2 \cdot \frac{1}{2n-1} \cdot k^{2n} \right] \right)$$

Beispiel 4: Länge eines Graphen

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion und $\Gamma_f := \{(x, f(x)) \mid x \in [a, b]\}$ ihr Graph. Wir parametrisieren Γ_f durch $\gamma : [a, b] \rightarrow \mathbb{R}^2$ mit $\gamma(x) := (x, f(x))$. Dann gilt

$$\gamma'(x) = (1, f'(x)) \quad \text{und} \quad \|\gamma'(x)\| = \sqrt{1 + f'(x)^2}.$$

Folglich ist die Länge

$$L(\gamma) = \int_a^b \sqrt{1 + f'(x)^2} dx.$$

Wir haben die geometrischen Eigenschaften Tangente, Schnittwinkel, Länge einer Kurve mit Hilfe ihrer Parametrisierung erklärt. Man kann eine Spurkurve Γ natürlich auf verschiedene Weisen (z.B. mit unterschiedlicher Geschwindigkeit) durchlaufen. Sinnvollerweise sollten sich ihre geometrischen Eigenschaften bei Umparametrisierung nicht ändern. Wir werden uns in einer Übungsaufgabe davon überzeugen, dass dies tatsächlich so ist.

6.4.2 Flächeninhalt ebener Gebiete

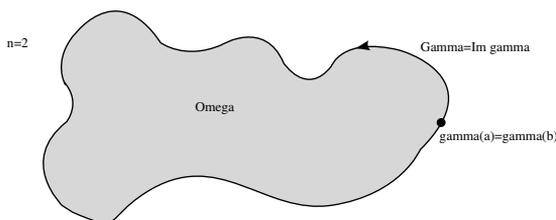
In diesem Abschnitt wollen wir Methoden herleiten, mit den wir den Flächeninhalt "komplizierterer" ebener Gebiete berechnen können.

Definition 6.13. Eine Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^2$ heißt

- geschlossen, falls $\gamma(a) = \gamma(b)$.
- einfach, falls $\gamma : (a, b) \rightarrow \mathbb{R}^2$ injektiv ist.

Eine einfache, geschlossene Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^2$ heißt positiv-orientiert, falls sie entgegen dem Uhrzeigersinn durchlaufen wird.

Sei nun $\gamma : [a, b] \rightarrow \mathbb{R}^2$ eine einfache, geschlossene und positiv-orientierte Kurve:

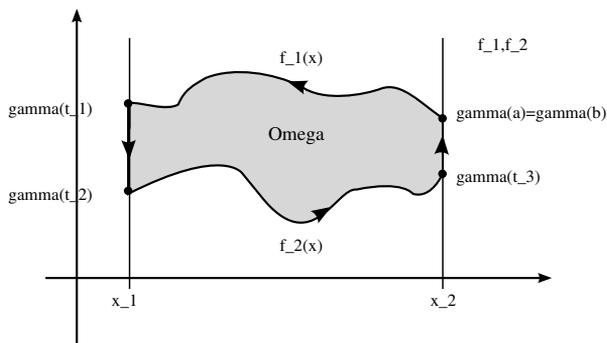


Ω bezeichne das von $\Gamma := \text{Im}(\gamma)$ umschlossene Gebiet. Wir geben nun eine Formel an, mit der man den Flächeninhalt von Ω mit Hilfe der Randkurve berechnen kann.

Satz 6.24 Sei $\gamma : [a, b] \rightarrow \mathbb{R}^2$ eine einfache, geschlossene, positiv-orientierte, (stückweise) stetig differenzierbare Kurve und sei $\Omega \subset \mathbb{R}^2$ das von $\Gamma := \text{Im}(\gamma)$ umschlossene Gebiet. Weiterhin seien $x, y : [a, b] \rightarrow \mathbb{R}$ die Koordinatenfunktionen von γ , d.h., $\gamma(t) =: (x(t), y(t))$. Dann gilt für den Flächeninhalt von Ω :

$$\begin{aligned} \text{Area}(\Omega) &= - \int_a^b y(t)x'(t) dt \stackrel{\text{p.I.}}{=} \int_a^b y'(t)x(t) dt \\ &= \frac{1}{2} \int_a^b (x(t)y'(t) - y(t)x'(t)) dt. \end{aligned}$$

Beweis. 1) Wir betrachten zunächst Kurven γ , deren Spur aus zwei zur y -Achse parallelen Strecken und zwei Bögen, die Graphen von Funktionen f_1 und f_2 mit $0 < f_1 < f_2$ sind, besteht.



Nach Definition des Riemann-Integrals ist

$$\text{Area}(\Omega) = \int_{x_1}^{x_2} f_1(x) dx - \int_{x_1}^{x_2} f_2(x) dx. \tag{*}$$

Für die obige Kurve $\gamma(t) = (x(t), y(t))$ gilt

$$\gamma(t) = \begin{cases} (x(t), f_1(x(t))), & t \in [a, t_1] \\ (x_1, y(t)), & t \in [t_1, t_2] \\ (x(t), f_2(x(t))), & t \in [t_2, t_3] \\ (x_2, y(t)), & t \in [t_3, b]. \end{cases}$$

Die Substitution $x = x(t)$, $dx = x'(t)dt$ mit $x_1 = x(t_1)$ und $x_2 = x(a)$ liefert für das erste Integral in (*)

$$\int_{x_1}^{x_2} f_1(x) dx = \int_{t_1}^a f_1(x(t)) \cdot x'(t) dt = - \int_a^{t_1} y(t)x'(t) dt.$$

Analog ergibt sich mit $x = x(t)$, $dx = x'(t)dt$ sowie $x_1 = x(t_2)$ und $x_2 = x(t_3)$ für das zweite Integral in (*)

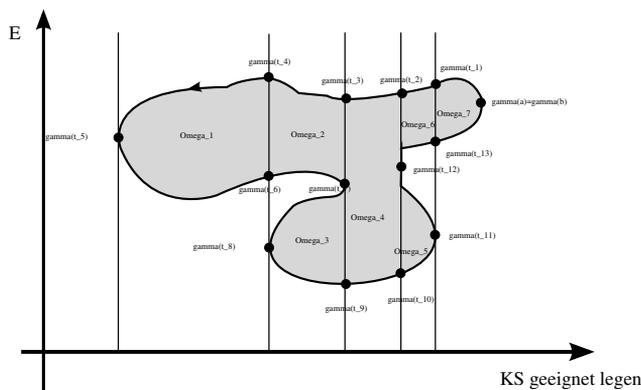
$$\int_{x_1}^{x_2} f_2(x) dx = \int_{t_2}^{t_3} y(t)x'(t) dt.$$

Damit ergibt sich insgesamt

$$Area(\Omega) = - \left(\int_a^{t_1} y(t)x'(t) dt + \int_{t_2}^{t_3} y(t)x'(t) dt \right) = - \int_a^b y(t)x'(t) dt,$$

da $x(t) = const$ für $t \in [t_1, t_2]$ und $t \in [t_3, b]$.

2) Wir betrachten nun den allgemeinen Fall. Zunächst sei die Kurve γ stetig differenzierbar.



Beh.: Man kann Ω in eine endliche Zahl von Gebieten zerlegen, die die Form aus 1) haben. Um das einzusehen, legen wir das Koordinatensystem so, dass Γ im positiven Quadranten liegt. Sei E die y -Achse. Der Abstand des Punktes $\gamma(t)$ zur y -Achse E ist dann gegeben durch

$$x(t) = dist(E, \gamma(t)).$$

Da das Intervall $[a, b]$ kompakt ist, hat $x : [a, b] \rightarrow \mathbb{R}^+$ nur endlich viele kritische Werte. Wir zeichnen in diesen Punkten, wie im Bild dargestellt, die zu E parallelen Geraden. Dann zerlegt sich Ω in Gebiete der Form 1). Ist nämlich $x'(t) \neq 0$, so ist $x' > 0$ oder $x' < 0$. Daher ist $x(t)$ zwischen den kritischen Punkten von $x : [a, b] \rightarrow \mathbb{R}^+$ streng monoton wachsend oder streng monoton fallend. Deshalb existiert dort eine Umkehrfunktion $t = t(x)$ und das entsprechende Teilstück von γ kann man als Graph der Funktion $f(x) := y(t(x))$ darstellen:

$$\gamma(t) = (x(t), y(t)) = (x(t), f(x(t))).$$

Nun können wir die Fläche von Ω leicht berechnen:

$$\begin{aligned}
 \text{Area}(\Omega) &= \sum_{i=1}^m \text{Area}(\Omega_i) \\
 &= -\left(\int_a^{t_1} y(t)x'(t) dt + \int_{t_3}^b y(t)x'(t) dt + \int_{t_1}^{t_2} y(t)x'(t) dt + \int_{t_2}^{t_3} y(t)x'(t) dt + \dots \right) \\
 &= -\int_a^b y(t)x'(t) dt,
 \end{aligned}$$

wobei wir im letzten Schritt benutzt haben, dass alle Parameterabschnitte $[t_i, t_{i+1}]$ genau einmal auftreten. Ist γ nur stückweise stetig differenzierbar, so zeichnen wir durch die "Ecken" von γ zusätzliche parallele Geraden zu E und schließen dann analog. \square

Folgerung 6.3 Ist die Randkurve γ in Polarkoordinaten gegeben, d.h., gilt

$$\gamma(t) = r(t)e^{i\varphi(t)} = (r(t) \cos \varphi(t), r(t) \sin \varphi(t)),$$

so ist

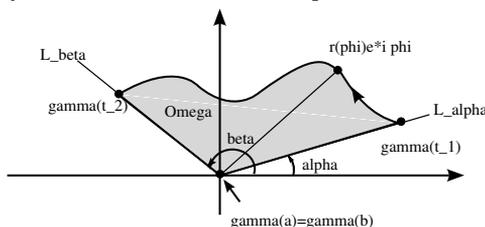
$$\text{Area}(\Omega) = \frac{1}{2} \int_a^b r^2(t)\varphi'(t) dt.$$

Folgerung 6.4 (Leibnizsche Sektorformel)

Das Gebiet Ω sei begrenzt durch die Strahlen L_α und L_β mit den Winkeln α bzw. β zur x -Achse, $0 \leq \alpha < \beta \leq 2\pi$, und eine durch Polarkoordinaten beschriebene Kurve $r = r(\varphi)$, wobei der Winkel $\varphi \in [\alpha, \beta]$ beim Durchlauf durch die Kurve streng monoton wächst.

Dann gilt

$$\text{Area}(\Omega) = \frac{1}{2} \int_\alpha^\beta r^2(\varphi) d\varphi.$$



Beweis. Wir parametrisieren die Randkurve von Ω durch $\gamma(t) = r(t)e^{i\varphi(t)}$, $t \in [a, b]$,

$$\varphi(t) = \begin{cases} \alpha, & t \in [a, t_1] \\ \varphi(t), & t \in [t_1, t_2] \\ \beta, & t \in [t_2, b] \end{cases}$$

Aus Folgerung 6.3 erhalten wir

$$\text{Area}(\Omega) = \frac{1}{2} \int_{t_1}^{t_2} r^2(t)\varphi'(t) dt.$$

Wir substituieren $\varphi = \varphi(t)$, $d\varphi = \varphi'(t)dt$. Da $\varphi(t)$ auf $[t_1, t_2]$ streng monoton wachsend ist, existiert eine Umkehrfunktion $t = t(\varphi)$. Somit können wir r in Abhängigkeit von φ darstellen, $r = r(\varphi)$, und erhalten insgesamt

$$\text{Area}(\Omega) = \frac{1}{2} \int_{\alpha}^{\beta} r^2(\varphi) d\varphi.$$

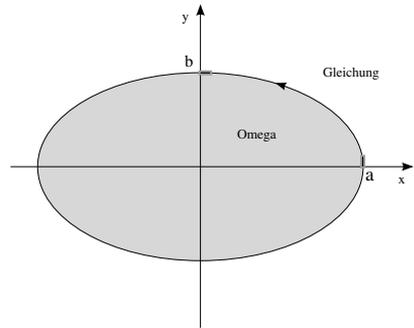
□

Beispiel 1: *Flächeninhalt einer Ellipse*

Sei Ω die von einer Ellipse mit den Halbachsen a und b eingeschlossene Fläche.

Wir parametrisieren die Ellipse durch die Kurve $\gamma(t) = (a \cos t, b \sin t) =: (x(t), y(t))$ mit $t \in [0, 2\pi]$.

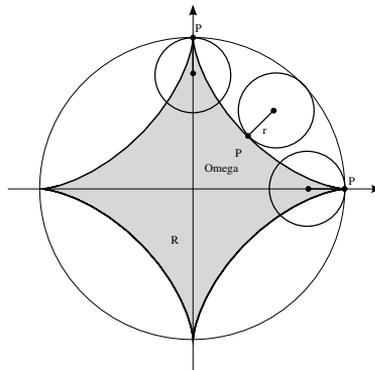
Dann ist nach Satz 6.24



$$\text{Area}(\Omega) = \frac{1}{2} \int_0^{2\pi} (x(t)y'(t) - y(t)x'(t)) dt = \frac{1}{2} \int_0^{2\pi} ab \cdot (\cos^2 t + \sin^2 t) dt = \pi ab.$$

Beispiel 2: *Flächeninhalt der Sternkurve (Astroide)*

Ein Kreis vom Radius $r = \frac{R}{4}$ rolle auf der Innenseite eines Kreises vom Radius R entlang.



Der Weg, den der feste Punkt P während des Rollens zurücklegt, wird Astroide genannt und ist bestimmt durch die Gleichung $x^{\frac{2}{3}} + y^{\frac{2}{3}} = R^{\frac{2}{3}}$. Wir können die Astroide parametrisieren durch $\gamma : [0, 2\pi] \rightarrow \mathbb{R}$ mit $\gamma(t) := R (\cos^3 t, \sin^3 t)$. Dann folgt für den Flächeninhalt des Gebietes Ω , das durch die Astroide begrenzt wird

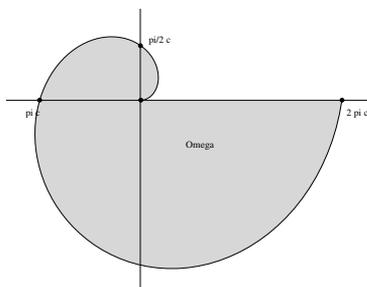
$$\begin{aligned}
 \text{Area}(\Omega) &= \frac{1}{2}R^2 \int_0^{2\pi} (\cos^3 t \cdot 3 \sin^2 t \cdot \cos t + 3 \cos^2 t \cdot \sin t \cdot \sin^3 t) dt \\
 &= \frac{3}{2}R^2 \int_0^{2\pi} \sin^2 t \cos^2 t (\cos^2 t + \sin^2 t) dt \\
 &= \frac{3}{2}R^2 \int_0^{2\pi} (\sin t \cos t)^2 dt = \frac{3}{2}R^2 \int_0^{2\pi} \left(\frac{1}{2} \sin 2t\right)^2 dt \\
 &= \frac{3}{8}R^2 \int_0^{2\pi} \sin^2 2t dt, \quad x := 2t, dx = 2dt \\
 &= \frac{3}{16}R^2 \int_0^{4\pi} \sin^2 x dx = \frac{3}{16}R^2 \cdot 8 \underbrace{\int_0^{\frac{\pi}{2}} \sin^2 x dx}_{=\frac{\pi}{4}} = \frac{3}{8}\pi R^2.
 \end{aligned}$$

Beispiel 3: Die Archimedische Spirale

Die archimedische Spirale ist die in Polarkoordinaten durch $r(\varphi) = c\varphi$, mit einer Konstanten $c \in \mathbb{R}$, beschriebene Kurve. In Euklidischen Koordinaten ist sie also durch

$$\gamma(\varphi) = r(\varphi)e^{i\varphi} = (c\varphi \cos \varphi, c\varphi \sin \varphi), \quad \varphi \in [0, 2\pi],$$

gegeben.



Nach der Leibnizschen Sektorformel ist der Flächeninhalt des Gebietes Ω

$$\text{Area}(\Omega) = \frac{1}{2} \int_0^{2\pi} r^2(\varphi) d\varphi = \frac{1}{2}c^2 \int_0^{2\pi} \varphi^2 d\varphi = \frac{1}{2}c^2 \cdot \frac{1}{3}8\pi^3 = \frac{4}{3}\pi^3 c^2.$$

6.5 Analytische Anwendung: Differentialgleichungen mit getrennten Variablen

Eine Differentialgleichung ist eine Gleichung, die eine Beziehung zwischen einer Funktion und ihren Ableitungen beschreibt. Auf solche Gleichungen trifft man immer, wenn man dynamische Prozesse in der Natur mathematisch modellieren möchte.

Definition 6.14. Sei $F : U \subset \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$ eine Funktion. Dann heißt die Gleichung

$$y^{(n)} = F(x, y, y', \dots, y^{(n-1)}) \quad (6.17)$$

(gewöhnliche) Differentialgleichung n -ter Ordnung.

Eine Lösung von (6.17) ist eine n -mal differenzierbare Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f^{(n)}(x) = F(x, f(x), f'(x), \dots, f^{(n-1)}(x)) \quad \forall x \in I.$$

Ist zusätzlich $(x_0, y_0, y_1, \dots, y_{n-1}) \in U$, dann heißt die Gleichung (6.17) zusammen mit den Bedingungen

$$y(x_0) = y_0, y'(x_0) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1} \quad (6.18)$$

Anfangswertproblem (AWP) n -ter Ordnung.

Eine Lösung des Anfangswertproblems ist eine n -mal differenzierbare Funktion $f : I \rightarrow \mathbb{R}$ mit $x_0 \in I$, die die Differentialgleichung (6.17) erfüllt und zusätzlich den Anfangsbedingungen

$$f(x_0) = y_0, f'(x_0) = y_1, \dots, f^{(n-1)}(x_0) = y_{n-1}$$

genügt.

Man möchte nun folgende Fragen beantworten:

- Unter welchen Bedingungen an F gibt es Lösungen der Differentialgleichung (6.17) bzw. des Anfangswertproblems (6.17) und (6.18)?
- Unter welchen Bedingungen an F ist die Lösung eindeutig bestimmt?
- Wie lange "lebt" die Lösung (d.h. wie groß ist der Definitionsbereich I)?
- Wie kann man Lösungen bestimmen?

Zur Antwort auf diese Fragen gibt es eine weit entwickelte Theorie der *gewöhnlichen Differentialgleichungen*, die man während der Masterphase des Lehramtsstudiums in einer Wahlpflicht-Vorlesung kennenlernen kann.

Beispiel 1: Die Differentialgleichung $y^{(n)} = 0$:

In diesem Fall ist $F = 0$. Wir kennen die Lösungen bereits aus Kapitel 5. Jede Lösung $f : \mathbb{R} \rightarrow \mathbb{R}$ ist ein Polynom höchstens $(n - 1)$ -ten Grades, d.h.

$$f(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots + a_1x + a_0, \quad x \in \mathbb{R}.$$

Aus den Anfangsbedingungen $f(x_0) = y_0, f'(x_0) = y_1, \dots, f^{(n-1)}(x_0) = y_{n-1}$ kann man die Koeffizienten a_0, a_1, \dots, a_{n-1} eindeutig bestimmen.

Beispiel 2: Die Differentialgleichung $y' = c \cdot y$, mit $c \in \mathbb{R}$ konstant:

In diesem Fall ist $F(x, y) = c \cdot y$. Die Funktionen der Gestalt

$$f(x) = k \cdot e^{cx}, \quad x \in \mathbb{R},$$

mit einer Konstante $k \in \mathbb{R}$, lösen offensichtlich die Differentialgleichung. Es gibt keine anderen Lösungen: Ist nämlich $g : \mathbb{R} \rightarrow \mathbb{R}$ eine weitere Lösung, dann gilt für die Funktion $h : \mathbb{R} \rightarrow \mathbb{R}$ mit $h(x) := g(x) \cdot e^{-cx}$:

$$h'(x) = g'(x)e^{-cx} - cg(x)e^{-cx} = cg(x)e^{-cx} - cg(x)e^{-cx} = 0 \quad \forall x \in \mathbb{R}.$$

Folglich ist die Funktion h konstant, d.h. es existiert eine Konstante $k \in \mathbb{R}$ mit $g(x) = ke^{cx}$ für alle $x \in \mathbb{R}$. Aus der Anfangsbedingung $f(x_0) = y_0$ kann man die Konstante k eindeutig bestimmen.

Wir wollen uns hier mit einem einfachen Typ einer Differentialgleichung 1. Ordnung befassen, der *Differentialgleichung mit getrennten Variablen*.

Definition 6.15. Eine Differentialgleichung 1. Ordnung $y' = F(x, y)$ heißt *Differentialgleichung mit getrennten Variablen*, wenn die Funktion F die Gestalt

$$F(x, y) = p(x) \cdot q(y) \tag{6.19}$$

für stetige Funktionen $p : J_1 \rightarrow \mathbb{R}$ und $q : J_2 \rightarrow \mathbb{R}$ hat, wobei J_1 und J_2 offene Intervalle sind und q keine Nullstelle auf J_2 besitzt.

Satz 6.25 (Differentialgleichung mit getrennten Variablen)

Wir betrachten ein Anfangswertproblem mit getrennten Variablen

$$y' = p(x) \cdot q(y), \quad y(x_0) = y_0, \tag{6.20}$$

mit Bedingungen wie in (6.19). Sei G eine Stammfunktion von $\frac{1}{q}$ auf J_2 . Dann ist G auf J_2 injektiv und das Anfangswertproblem (6.20) besitzt auf einem hinreichend kleinen Intervall $I \subset J_1$ um x_0 eine eindeutig bestimmte Lösung. Diese ist gegeben durch

$$f(t) = G^{-1} \left(G(y_0) + \int_{x_0}^t p(x) dx \right). \tag{6.21}$$

Beweis. 1) Wir nehmen zuerst an, dass die Differentialgleichung (6.20) eine Lösung f besitzt und bestimmen ihre Form. Sei also $f : I \rightarrow \mathbb{R}$ eine differenzierbare Funktion mit $f'(x) = p(x) \cdot q(f(x))$ für alle $x \in I$ und $f(x_0) = y_0$. Dann ist insbesondere $f(I) \subset J_2$. Wir integrieren die Gleichung $\frac{f'(x)}{q(f(x))} = p(x)$ von x_0 bis $t \in I$:

$$\int_{x_0}^t \frac{f'(x)}{q(f(x))} dx = \int_{x_0}^t p(x) dx.$$

Im ersten Integral führen wir die Substitution $y = f(x)$ und $dy = f'(x) dx$ durch und erhalten mit der Stammfunktion G von $\frac{1}{q}$

$$G(f(t)) - G(y_0) = \int_{y_0}^{f(t)} \frac{1}{q(y)} dy = \int_{x_0}^t p(x) dx.$$

Dies liefert die folgende *implizite* Beschreibung von f :

$$G(f(t)) = G(y_0) + \int_{x_0}^t p(x) dx \quad \forall t \in I. \quad (6.22)$$

Insbesondere liegt die Zahl auf der rechten Seite von (6.22) für alle $t \in I$ im Bild von G . Da $\frac{1}{q}$ stetig ist und auf J_2 keine Nullstellen besitzt, gilt $G' = \frac{1}{q} > 0$ auf J_2 oder $G' = \frac{1}{q} < 0$ auf J_2 . D.h. G ist streng monoton und somit injektiv. Folglich existiert die Umkehrfunktion $G^{-1} : \text{Im}(G) \rightarrow J_2$. Somit kann man die Gleichung (6.22) nach $f(t)$ auflösen und erhält eine *explizite* Formel für f :

$$f(t) = G^{-1}\left(G(y_0) + \int_{x_0}^t p(x) dx\right) \quad \forall t \in I. \quad (6.23)$$

Wenn es also eine Lösung $f : I \rightarrow \mathbb{R}$ des AWP (6.20) gibt, so ist sie durch die Gleichung (6.23) eindeutig bestimmt.

2) Wir zeigen nun die Existenz einer Lösung. Dazu *definieren* wir uns eine Funktion durch die Gleichung (6.23): Wir betrachten ein Intervall $I \subset J_1$ um x_0 , das so klein gewählt ist, dass $G(y_0) + \int_{x_0}^t p(x) dx$ für alle $t \in I$ im Bild von G liegt und definieren $f : I \rightarrow \mathbb{R}$ durch

$$f(t) := G^{-1}\left(G(y_0) + \int_{x_0}^t p(x) dx\right) \quad \forall t \in I.$$

f löst tatsächlich das Anfangswertproblem (6.20), denn nach Kettenregel gilt für alle $t \in I$:

$$f'(t) = \underbrace{(G^{-1})'\left(G(y_0) + \int_{x_0}^t p(x) dx\right)}_{= G(f(t))} \cdot p(t) = \frac{1}{G'(f(t))} \cdot p(t) = q(f(t)) \cdot p(t)$$

und $f(x_0) = G^{-1}(G(y_0)) = y_0$. □

Satz 6.25 rechtfertigt das folgende **formale Vorgehen** zur Lösung einer Differentialgleichung mit getrennten Variablen $y' = p(x) \cdot q(y)$:

Wir setzen $y' = \frac{dy}{dx}$ und bringen alle Ausdrücke, die y enthalten auf die linke Seite und alle Ausdrücke, die x enthalten auf die rechte Seite ("Trennung der Variablen"). Danach integrieren wir links über y und rechts über x :

$$y' = \frac{dy}{dx} \implies \frac{dy}{dx} = p(x) \cdot q(y) \implies \frac{1}{q(y)} dy = p(x) dx. \quad (6.24)$$

a) Bei gegebener Anfangsbedingung $y(x_0) = y_0$ integriert man links über $[y_0, y(t)]$ und rechts über $[x_0, t]$:

$$\stackrel{(6.24)}{\implies} \int_{y_0}^{y(t)} \frac{1}{q(y)} dy = \int_{x_0}^t p(x) dx.$$

Ist G eine Stammfunktion von $\frac{1}{q}$, so erhält man daraus die impliziten Form für die Lösung $y(t)$

$$G(y(t)) = G(y_0) + \int_{x_0}^t p(x) dx,$$

und durch Anwenden der Umkehrfunktion G^{-1} (falls man sie berechnen kann) die explizite Form der Lösung

$$y(t) = G^{-1}\left(G(y_0) + \int_{x_0}^t p(x) dx\right).$$

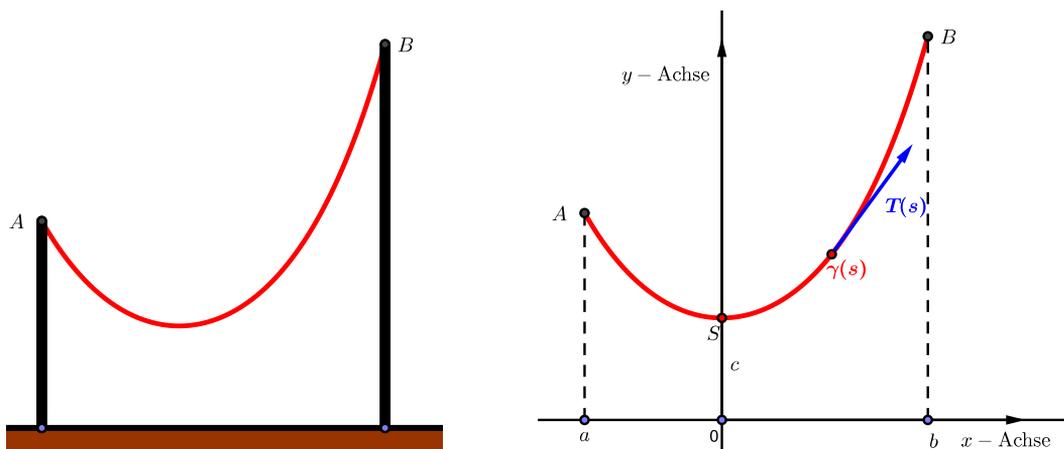
b) Ist keine Anfangsbedingung vorgegeben, so benutzt man bei der Integration die Stammfunktionen (unbestimmte Integration):

$$\begin{aligned} \stackrel{(6.24)}{\implies} \int \frac{1}{q(y)} dy \Big|_{y=y(x)} &= \int p(x) dx \\ \implies G(y(x)) &= \int p(x) dx + C, \end{aligned}$$

wobei G eine Stammfunktion von $\frac{1}{q}$ und C eine reelle Konstante ist. Auch diese Gleichung kann man ggf. nach $y(x)$ auflösen. Die Konstante C ist durch Festlegung eines Anfangswertes $y(x_0) = y_0$ zu bestimmen.

Beispiel: Die Kettenlinie (Seilkurve)

Wir wollen dieses Verfahren jetzt benutzen, um die Gleichung für die Kettenlinie herzuleiten. Wir hängen ein Seil (oder eine Kette) zwischen zwei Pfosten so hoch auf, dass es nicht auf der Erde schleift. Wie können wir die entstehende Form des Seils mathematisch beschreiben? Wir leiten eine Gleichung unter idealisierten Annahmen her (Seil ist vollkommen biegsam, das Gewicht des Seils pro Längeneinheit ist überall gleich).



Zur mathematischen Beschreibung legen wir ein Koordinatensystem so fest, dass das Seil in der xy -Ebene hängt, die x -Achse auf dem Erdboden liegt und die y -Achse durch den tiefsten Punkt S des Seils läuft. c sei die Höhe von S über dem Boden und a und b die x -Koordinate der Aufhängungspunkte A bzw. B . Wir wollen das Seil als Graph einer Funktion $f : [a, b] \rightarrow \mathbb{R}$ beschreiben.

Wir betrachten die rechte Hälfte des Seils. Das Seil habe zwischen S und B die Länge L . Jedem Punkt P auf der rechten Seilhälfte ordnen wir als Parameter die Länge s des Seils zwischen S und P zu, d.h. die rechte Seilhälfte ist beschrieben als Spur der Kurve $\gamma : [0, L] \rightarrow \mathbb{R}^2$, wobei das Seil vom Seilpunkt S bis zum Seilpunkt $\gamma(s)$ die Länge s hat. Auf den Seilpunkt $\gamma(s)$ wirkt eine Spannungskraft $T(s)$ tangential an die Kurve γ (durch die Aufhängung an den Pfosten). Gleichzeitig wirkt auf jeden Seilabschnitt die Gewichtskraft durch das Gewicht des Seils senkrecht nach unten. Wir nehmen (idealisiert) an, dass keine weiteren Kräfte auf das Seil wirken.

Wir leiten zunächst eine Formel für die Spannungskraft her. Dazu benutzen wir das folgende physikalische Gesetz: Das hängende Seil befindet sich im Gleichgewichtszustand, wenn sich in jedem beliebig kleinen Seilabschnitt die wirkenden Kräfte aufheben. D.h. für zwei beliebig nahe Seilpunkte $\gamma(s)$ und $\gamma(s+h)$ mit $h > 0$ gilt

$$T(s+h) - T(s) + G_s(h) = 0,$$

wobei $T(s+h)$ die Spannungskraft im Punkt $\gamma(s+h)$, $-T(s)$ die Gegenspannung im Punkt $\gamma(s)$ und $G_s(h)$ die Gewichtskraft ist, die durch das Gewicht des Seils zwischen $\gamma(s)$ und $\gamma(s+h)$ entsteht. Sei $g > 0$ das konstante Gewicht des Seils pro Längeneinheit.

Da h die Länge des Seils zwischen $\gamma(s)$ und $\gamma(s+h)$ ist, folgt für den Vektor der Gewichtskraft:

$$G_s(h) = (0, -h \cdot g).$$

Wir erhalten somit für den Spannungsvektor

$$\frac{T(s+h) - T(s)}{h} = (0, g)$$

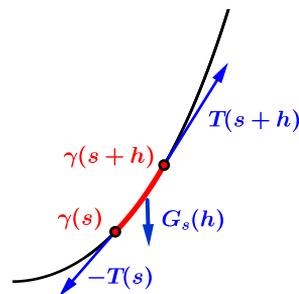
und beim Grenzübergang $h \rightarrow 0$ für seine Ableitung

$$T'(s) = (0, g).$$

Bezeichnet $T(s) =: (T_1(s), T_2(s))$, so folgt für die Komponenten des Spannungsvektors:

$$T_1(s) = c_1 \quad \text{und} \quad T_2(s) = gs + c_2$$

wobei c_1 und c_2 reelle Konstanten sind. c_1 ist positiv (da wir auf der rechten Seilhälfte sind) und $c_2 = 0$, da im Punkt $S = \gamma(0)$ der Spannungsvektor parallel zur x -Achse ist, d.h. $T_2(0) = 0$ gilt. Wir erhalten für die Spannungskraft im Seilpunkt $\gamma(s)$ somit



$$T(s) = (c_1, sg) \quad \text{mit } c_1 \in \mathbb{R}^+.$$

Dies genügt nun, um die gesuchte Funktionsgleichung für das Seil herzuleiten. Wir suchen eine Funktion $f : [0, b] \rightarrow \mathbb{R}$, so dass das Seil zwischen S und B der Graph von f ist. Mit unserer Parametrisierung γ gilt für einen Seilpunkt $(x, f(x))$:

$$\gamma(s) = (x, f(x)), \quad \text{wobei } s \text{ die Länge des Seil von } S \text{ bis } (x, f(x)) \text{ ist.}$$

Der Tangentialvektor an den Graphen der Funktion f im Punkt $(x, f(x))$ ist gegeben durch $(1, f'(x))$, d.h. die Vektoren $T(s)$ und $(1, f'(x))$ sind parallel und zeigen in die gleiche Richtung. Folglich gilt für ihren Anstieg

$$\frac{f'(x)}{1} = \frac{T_2(s)}{T_1(s)} = \underbrace{\frac{g}{c_1}}_{=: k} \cdot s.$$

Die Länge s des Seils zwischen $S = (0, f(0))$ und $(x, f(x))$ berechnet sich in der Parametrisierung als Graph durch

$$s = \int_0^x \sqrt{1 + f'(t)^2} dt.$$

Daraus erhalten wir

$$f'(x) = k \cdot \int_0^x \sqrt{1 + f'(t)^2} dt.$$

und durch Ableiten nach x

$$f''(x) = k \cdot \sqrt{1 + f'(x)^2}.$$

Betrachtet man analog die linke Seilhälfte, so erhält man die gleiche Differentialgleichung für f auf dem Intervall $[a, 0]$.

Unsere physikalischen Annahmen führen also auf die folgende Differentialgleichung 2. Ordnung für die gesuchte Funktion $f : [a, b] \rightarrow \mathbb{R}$

$$\boxed{f''(x) = k \cdot \sqrt{1 + f'(x)^2} \quad \text{Differentialgleichung der Kettenlinie (Seilkurve)}}$$

mit den Anfangsbedingungen $f(0) = c$ und $f'(0) = 0$, wobei $k > 0$ eine Materialkonstante des Seils und c die Höhe des tiefsten Punktes des Seils über dem Boden ist.

Lösung der Differentialgleichung der Kettenlinie

Wir setzen $z(x) := f'(x)$. Dann hat die Differentialgleichung der Kettenlinie, betrachtet als Differentialgleichung für z , die Form einer Differentialgleichung mit getrennten Variablen:

$$z' = k \cdot \sqrt{1 + z^2} \quad \text{mit der Anfangsbedingung } z(0) = 0.$$

Wir setzen $z' = \frac{dz}{dx}$ und trennen die Variablen:

$$\frac{1}{\sqrt{1 + z^2}} dz = k \cdot dx.$$

Nach Integration folgt

$$\int_0^{z(t)} \frac{1}{\sqrt{1+z^2}} dz = \int_0^t k dx$$

und daraus mit der Stammfunktion arsinh von $\frac{1}{\sqrt{1+z^2}}$

$$\operatorname{arsinh}(z(t)) - \operatorname{arsinh}(0) = kt.$$

Durch Anwenden von \sinh folgt

$$f'(t) = z(t) = \sinh(kt) \quad \forall t \in \mathbb{R}.$$

f erhalten wir dann durch nochmaliges Integrieren:

$$f(x) = f(0) + \int_0^x f'(t) dt = c + \frac{1}{k} \cdot \cosh(kt) \Big|_{t=0}^{t=x},$$

d.h.

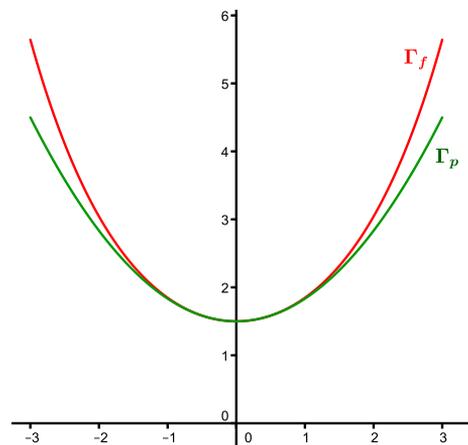
$$f(x) = c - \frac{1}{k} + \frac{1}{k} \cosh(kx) \quad \forall x \in \mathbb{R}$$

Legt man das Koordinatensystem jetzt noch so, dass $c = \frac{1}{k}$ gilt, so erhält man *in diesem* Koordinatensystem

$$f(x) = \frac{1}{k} \cosh(kx) \quad \forall x \in \mathbb{R}.$$

Die Reihenentwicklung von \cosh zeigt, inwieweit diese Funktion von einer Parabel, die ein Schüler vielleicht als Lösung vermuten würde, abweicht:

$$\begin{aligned} f(x) &= \frac{1}{k} \cosh(kx) = \frac{1}{k} \cdot \frac{e^{kx} + e^{-kx}}{2} \\ &= \frac{1}{k} \cdot \sum_{n=0}^{\infty} \frac{(kx)^{2n}}{(2n)!} \\ &= \underbrace{\frac{1}{k} + \frac{k}{2}x^2}_{=: p(x) \text{ Parabel}} + \frac{k^3}{24}x^4 + \dots \end{aligned}$$



Metrische Räume und stetige Abbildungen

In der bisherigen Vorlesung haben wir reelle Funktionen $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ und ihre Eigenschaften (Stetigkeit, Differenzierbarkeit, Integrierbarkeit) behandelt. Diese drei Eigenschaften haben wir mit Hilfe des Grenzwertbegriffes für reelle Zahlenfolgen definiert. Das zentrale Hilfsmittel zur Definition des Grenzwertes war der Abstand $|x - y|$ zwischen zwei reellen Zahlen x und y . Die bisher betrachtete Funktionenklasse reicht für die Beschreibung vieler Prozesse in der Natur nicht aus. Bereits wenn wir z.B. den Luftdruck oder die Temperatur in den Punkten der Erdoberfläche E beschreiben wollen, benötigen wir Funktionen vom Typ $f : E \rightarrow \mathbb{R}$, die von mehr als einer reellen Variablen abhängen. Wir werden deshalb in diesem Kapitel eine allgemeinere Klasse von Abbildungen $f : X \rightarrow Y$ betrachten und zunächst nur voraussetzen, dass im Definitionsbereich X und im Wertebereich Y Abstandsfunktionen gegeben sind, die die Definition von Grenzwerten ermöglichen. Dies genügt, um den Begriff der Stetigkeit für solche Abbildungen zu definieren. Wir werden sehen, dass sich zentrale Sätze über stetige reelle Funktionen $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$, die wir in Kapitel 4 behandelt haben, auf diesen allgemeineren Fall übertragen lassen.

7.1 Definition und Beispiele metrischer Räume

Definition 7.1. Sei X eine beliebige nichtleere Menge. Eine Abstandsfunktion (oder Metrik) auf X ist eine Abbildung $d : X \times X \rightarrow \mathbb{R}$ mit folgenden Eigenschaften

1. $d(p, q) \geq 0 \quad \forall p, q \in X \quad \text{und} \quad d(p, q) = 0 \iff p = q$ (Positivität),
2. $d(p, q) = d(q, p) \quad \forall p, q \in X$ (Symmetrie),
3. $d(p, q) \leq d(p, r) + d(r, q) \quad \forall p, q, r \in X$ (Dreiecksungleichung).

Das Paar (X, d) heißt metrischer Raum. Die Elemente von X heißen Punkte des metrischen Raumes, die Zahl $d(p, q)$ nennt man den Abstand zwischen p und q .

Beispiel 1: Die Standardmetriken auf \mathbb{R} und \mathbb{C} .

Sei $|x|$ der Betrag einer reellen Zahl x . Dann ist $d : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ mit

$$d(x, y) := |x - y| \quad \text{für } x, y \in \mathbb{R},$$

eine Abstandsfunktion auf \mathbb{R} .

Analog erhält man aus dem Betrag von komplexen Zahlen eine Abstandsfunktion auf \mathbb{C} :

$$d(z, w) := |z - w| \quad \text{für } z, w \in \mathbb{C}.$$

Diese durch den Betrag definierte Abstandsfunktion d nennen wir die *Standardmetrik* auf \mathbb{R} bzw. \mathbb{C} . Ist nichts anderes vereinbart, so seien \mathbb{R} und \mathbb{C} immer mit dieser Abstandsfunktion versehen.

Beispiel 2: Auf einer Menge X können verschiedene Abstandsfunktionen existieren.

Die folgenden drei Abbildungen d_1, d_2, d_3 sind z.B. Abstandsfunktionen auf \mathbb{R}^2 :

- Der "Luftlinienabstand" (Standardmetrik oder Euklidischer Abstand):

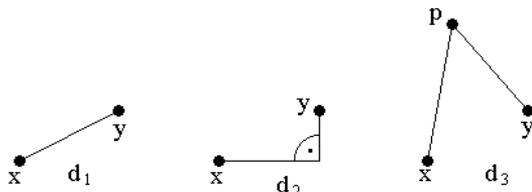
$$d_1(x, y) := \|x - y\| := \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}, \quad \text{wobei } x = (x_1, x_2), y = (y_1, y_2).$$

- Die "Taxifahrer-Metrik" (oder "Mannheimer-Metrik" – in Mannheim muß man rechtwinklige Straßen langlaufen, um von einem Punkt zum anderen zu kommen) :

$$d_2(x, y) := |x_1 - y_1| + |x_2 - y_2|.$$

- Die "Metrik der französischen Eisenbahn" (um von einer Stadt zur anderen zu kommen, muß man über Paris fahren): Sei p ein fixierter Punkt.

$$d_3(x, y) := \begin{cases} d_1(x, p) + d_1(p, y) & \text{falls } x \neq y, \\ 0 & \text{falls } x = y. \end{cases}$$



Beispiel 3: Auf jeder nichtleeren Menge X existiert eine Abstandsfunktion.

Wir definieren $d: X \times X \rightarrow \mathbb{R}$ durch

$$d(x, y) := \begin{cases} 0 & \text{falls } x = y, \\ 1 & \text{falls } x \neq y. \end{cases}$$

Dann ist (X, d) ein metrischer Raum. (X, d) heißt *diskreter metrischer Raum* und d die *diskrete Metrik*.

Beispiel 4: Das kartesische Produkt metrischer Räume.

Seien $(X_1, d_1), \dots, (X_n, d_n)$ metrische Räume. Wir betrachten die Produktmenge

$$X := X_1 \times X_2 \times \dots \times X_n := \{(p_1, \dots, p_n) \mid p_i \in X_i, i \in \{1, \dots, n\}\}$$

und definieren die *Produktmetrik* auf X durch

$$d((p_1, \dots, p_n), (q_1, \dots, q_n)) := \sqrt{d_1(p_1, q_1)^2 + \dots + d_n(p_n, q_n)^2} = \sqrt{\sum_{j=1}^n d_j(p_j, q_j)^2}.$$

Dann ist (X, d) ein metrischer Raum. Er heißt das *kartesische Produkt* der metrischen Räume $(X_1, d_1), \dots, (X_n, d_n)$.

Beweis. Die Positivität und Symmetrie sind aus der Definition sofort ersichtlich. Wir zeigen die Dreiecksungleichung. Seien dazu $p = (p_1, \dots, p_n)$, $q = (q_1, \dots, q_n)$ und $r = (r_1, \dots, r_n)$ drei beliebige Punkte in X . Unter Benutzung der Dreiecksungleichung für die Metriken d_j und der Cauchy-Schwarzschen Ungleichung (CSU) für das Euklidische Skalarprodukt (siehe Kapitel 6.4 oder Vorlesung *Lineare Algebra I*) erhalten wir

$$\begin{aligned}
 d(p, q)^2 &= \sum_{j=1}^n d_j(p_j, q_j)^2 \\
 &\stackrel{\Delta}{\leq} \sum_{j=1}^n \left(d_j(p_j, r_j) + d_j(r_j, q_j) \right)^2 \\
 &= \sum_{j=1}^n d_j(p_j, r_j)^2 + 2d_j(p_j, r_j) \cdot d_j(r_j, q_j) + d_j(r_j, q_j)^2 \\
 &= d(p, r)^2 + d(r, q)^2 + 2 \sum_{j=1}^n d_j(p_j, r_j) \cdot d_j(r_j, q_j) \\
 &\stackrel{\text{CSU}}{\leq} d(p, r)^2 + d(r, q)^2 + 2 \sqrt{\sum_{j=1}^n d_j(p_j, r_j)^2} \cdot \sqrt{\sum_{j=1}^n d_j(r_j, q_j)^2} \\
 &= d(p, r)^2 + d(r, q)^2 + 2 d(p, r) \cdot d(r, q) \\
 &= (d(p, r) + d(r, q))^2
 \end{aligned}$$

Da der Abstand nicht-negativ ist, können wir in der Ungleichung auf beiden Seiten die Wurzel ziehen und erhalten die Dreiecksungleichung für d . \square

Beispiel 5: Zwei Abstandsfunktionen auf der Menge der stetigen Funktionen $C^0([a, b], \mathbb{R})$. Sei $C^0([a, b], \mathbb{R})$ der Vektorraum der stetigen Funktionen von $[a, b]$ in \mathbb{R} . Dann definieren

$$\begin{aligned}
 d_\infty(f, g) &:= \max\{|f(x) - g(x)| \mid x \in [a, b]\}, & f, g \in C^0([a, b], \mathbb{R}), \\
 d_1(f, g) &:= \int_a^b |f(x) - g(x)| dx,
 \end{aligned}$$

Abstandsfunktionen auf $C^0([a, b], \mathbb{R})$ (Übungsaufgabe).

Beispiel 6: Metriken auf normierten Vektorräumen.

Sei V ein reeller Vektorraum. Eine Norm auf V ist eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$ mit

1. $\|v\| \geq 0$ für alle $v \in V$ und $\|v\| = 0 \iff v = 0$ (Positivität).
2. $\|\lambda \cdot v\| = |\lambda| \cdot \|v\|$ für alle $v \in V$ und $\lambda \in \mathbb{R}$ (Homogenität).
3. $\|v + w\| \leq \|v\| + \|w\|$ für alle $v, w \in V$ (Dreiecksungleichung).

Das Paar $(V, \|\cdot\|)$ heißt *normierter Vektorraum*. Dann ist $d : V \times V \rightarrow \mathbb{R}$, definiert durch

$$d(v, w) := \|v - w\| \quad v, w \in V,$$

eine Abstandsfunktion auf V . Sie heißt die *durch $\|\cdot\|$ induzierte Metrik*. Ein normierter Vektorraum $(V, \|\cdot\|)$ ist immer mit dieser Metrik versehen (falls nicht ausdrücklich anders gesagt).

Beweis. Die Positivität und Symmetrie von d folgt aus der 1. und 2. Eigenschaft für die Norm. Die Dreiecksungleichung für d erhält man aus der Dreiecksungleichung für die Norm. Für $v, w, u \in V$ gilt nämlich

$$d(v, u) = \|v - u\| = \|(v - w) + (w - u)\| \leq \|v - w\| + \|w - u\| = d(v, w) + d(w, u). \quad \square$$

Dies verallgemeinert die Situation in Beispiel 1. Dort ist $V = \mathbb{R}$ und $\|\cdot\| = |\cdot|$. In der analytischen Geometrie wird die *Euklidische Norm* auf dem \mathbb{R}^n benutzt

$$\|v\| := \sqrt{v_1^2 + v_2^2 + \dots + v_n^2} \quad \text{für } v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n,$$

und der davon induzierte *Euklidische Abstand* zwischen zwei Punkten $p = (p_1, \dots, p_n)$ und $q = (q_1, \dots, q_n)$:

$$d(p, q) := \|p - q\| = \sqrt{(p_1 - q_1)^2 + \dots + (p_n - q_n)^2}.$$

(siehe Vorlesung *Lineare Algebra I* und Kapitel 6.4. dieser Vorlesung). Ist der \mathbb{R}^n mit diesem Abstand versehen, so nennt man ihn den *Euklidischen Raum*. Er ist das n -fache kartesische Produkt von \mathbb{R} mit der Standardmetrik. In der Vorlesung *Elementargeometrie* des Lehramtsstudiums werden Sie andere Geometrien kennenlernen (hyperbolische Geometrie oder sphärische Geometrie), in denen die Abstände zweier Punkte mit anderen Abstandsfunktionen gemessen werden.

Definition 7.2. Zwei metrische Räume (X, d_X) und (Y, d_Y) heißen *isometrisch*, wenn eine bijektive Abbildung $\phi : X \rightarrow Y$ existiert, so dass

$$d_X(p, q) = d_Y(\phi(p), \phi(q)) \quad \text{für alle } p, q \in X.$$

Die Abbildung ϕ heißt dann *Isometrie* zwischen (X, d_X) und (Y, d_Y) .

In isometrischen metrischen Räumen stimmen alle Eigenschaften, die mittels der Abstandsfunktion charakterisiert werden, überein.

Für spätere Zwecke definieren wir das Analogon von offenen Intervallen im \mathbb{R} , offenen Kreisscheiben im \mathbb{R}^2 (bzw. in \mathbb{C}) und offenen Kugeln im Euklidischen Raum \mathbb{R}^3 für beliebige metrische Räume:

Definition 7.3. Sei (X, d) ein metrischer Raum, $p \in X$ und $\varepsilon \in \mathbb{R}^+$. Die Menge

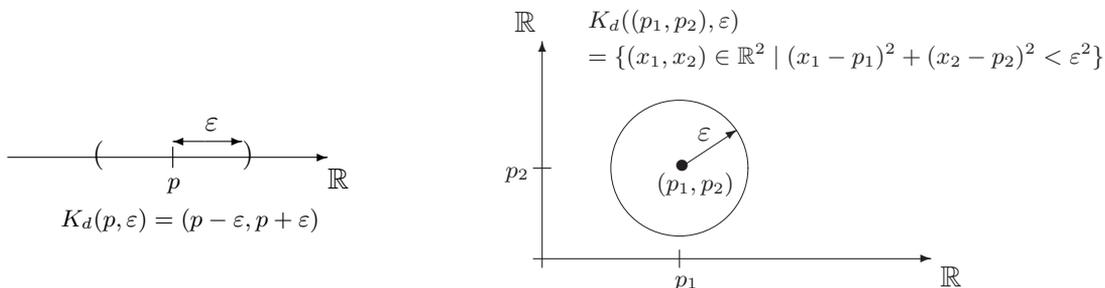
$$K(p, \varepsilon) := \{x \in X \mid d(p, x) < \varepsilon\}$$

heißt ε -Kugel in (X, d) um p .

Beispiel 1: Sei \mathbb{R}^n der Euklidische Raum. Dann ist

$$K(p, \varepsilon) = \{x \in \mathbb{R}^n \mid \|x - p\| < \varepsilon\} = \{x \in \mathbb{R}^n \mid (x_1 - p_1)^2 + \dots + (x_n - p_n)^2 < \varepsilon^2\}$$

die übliche n -dimensionale Kugel vom Radius ε um den Punkt $p = (p_1, \dots, p_n) \in \mathbb{R}^n$ ohne ihren Rand. Für $n = 1$ sind die ε -Kugeln offene Intervalle, für $n = 2$ Kreisscheiben ohne ihren Rand.

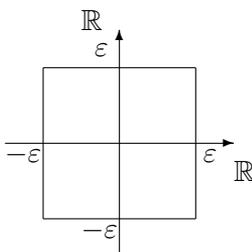


Beispiel 2: Sei \mathbb{R}^2 mit folgender Metrik versehen:

$$d_\infty((x_1, x_2), (y_1, y_2)) = \max(|x_1 - y_1|, |x_2 - y_2|).$$

(\mathbb{R}^2, d_∞) ist tatsächlich ein metrischer Raum (Übungsaufgabe). Die ε -Kugel um den Nullpunkt ist für diese Metrik ein Quadrat mit der Seitenlänge 2ε ohne seinen Rand:

$$K(0, \varepsilon) = \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| < \varepsilon, |x_2| < \varepsilon\}.$$



Beispiel 3: Sei (X, d) ein diskreter metrischer Raum. Dann gilt für die ε -Kugeln

$$K(x, \varepsilon) = \begin{cases} \{x\} & \text{falls } \varepsilon \leq 1 \\ X & \text{falls } \varepsilon > 1. \end{cases}$$

7.2 Konvergente Folgen und Cauchy-Folgen in metrischen Räumen

Definition 7.4. Sei (X, d) ein metrischer Raum und (x_n) eine Folge in X .

Wir sagen, dass (x_n) gegen $p \in X$ konvergiert, falls zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass $d(p, x_n) < \varepsilon$ für alle $n \geq n_0$, d.h. falls $x_n \in K(p, \varepsilon)$ für alle $n \geq n_0$.

Der Punkt p heißt Grenzwert (GW) der Folge (x_n) . Besitzt eine Folge (x_n) einen Grenzwert, so heißt sie konvergent. Besitzt die Folge (x_n) keinen Grenzwert, so heißt sie divergent.

Für eine gegen p konvergente Folge (x_n) schreiben wir:

$$\lim_{n \rightarrow \infty} x_n = p \quad \text{oder} \quad x_n \xrightarrow[n \rightarrow \infty]{} p \quad \text{oder kurz} \quad x_n \longrightarrow p.$$

Aus der Definition folgt unmittelbar:

Satz 7.1 Eine Folge (x_n) in einem metrischen Raum (X, d) konvergiert genau dann gegen $p \in X$, wenn die Folge der Abstände $(d(p, x_n))$ in \mathbb{R} gegen Null konvergiert:

$$x_n \longrightarrow p \text{ in } (X, d) \iff d(p, x_n) \longrightarrow 0 \text{ in } \mathbb{R}.$$

Satz 7.2 Der Grenzwert einer konvergenten Folge eines metrischen Raumes ist eindeutig bestimmt.

Beweis. Der Beweis wird analog zum Beweis des entsprechenden Satzes für konvergente Folgen in \mathbb{R} geführt. Wir überlassen das Übertragen als Übungsaufgabe. \square

Definition 7.5. Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt beschränkt, falls es eine Kugel $K(p, M)$ des metrischen Raumes gibt, die A enthält. Eine Folge (x_n) in (X, d) heißt beschränkt, wenn die Menge der Folgenglieder $\{x_1, x_2, x_3, \dots\}$ beschränkt ist.

Satz 7.3 Jede konvergente Folge (x_n) eines metrischen Raumes (X, d) ist beschränkt.

Beweis. Sei (x_n) eine konvergente Folge in (X, d) und $p \in X$ ihr Grenzwert. Wir betrachten die Kugel $K(p, 1)$ um p vom Radius 1. Dann existiert ein $n_0 \in \mathbb{N}$, so dass $x_n \in K(p, 1)$ für alle $n \geq n_0$. Wir setzen nun

$$M := \max(d(p, x_1), d(p, x_2), \dots, d(p, x_{n_0-1})).$$

Dann gilt $x_n \in K(p, M + 1)$ für alle $n \in \mathbb{N}$. Folglich ist (x_n) beschränkt. \square

Als nächstes charakterisieren wir die Konvergenz von Folgen im kartesischen Produkt metrischer Räume.

Satz 7.4 Seien $(X_1, d_1), \dots, (X_k, d_k)$ metrische Räume und (X, d) das kartesische Produkt dieser Räume. Eine Folge (x_n) in X mit $x_n = (x_{n1}, x_{n2}, \dots, x_{nk})$ konvergiert genau dann gegen $p = (p_1, p_2, \dots, p_k) \in X$, wenn für jedes $j \in \{1, \dots, k\}$ die Komponentenfolge $(x_{nj})_{n=1}^\infty$ in (X_j, d_j) gegen p_j konvergiert.

Beweis. Nach Definition der Produktmetrik ist

$$d(\underbrace{(p_1, \dots, p_k)}_p, \underbrace{(x_{n1}, \dots, x_{nk})}_{x_n}) = \sqrt{\sum_{i=1}^k d_i(p_i, x_{ni})^2} \geq d_j(p_j, x_{nj}) \quad (*)$$

für jedes $j \in \{1, \dots, k\}$. Sei nun (x_n) gegen p in (X, d) konvergent. Dann existiert für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit $d(p, x_n) < \varepsilon$ für alle $n \geq n_0$. Nach Abschätzung (*) folgt daraus $d_j(p_j, x_{nj}) < \varepsilon$ für alle $n \geq n_0$ und jedes $j \in \{1, \dots, k\}$. Somit konvergiert (x_{nj}) gegen p_j in (X_j, d_j) für jedes $j \in \{1, \dots, k\}$.

Sei umgekehrt $\lim_{n \rightarrow \infty} x_{nj} = p_j$ für jedes $j \in \{1, \dots, k\}$ und $\varepsilon > 0$. Dann existieren $n_{0j} \in \mathbb{N}$ so dass $d_j(p_j, x_{nj}) < \frac{\varepsilon}{\sqrt{k}}$ für alle $n \geq n_{0j}$. Es folgt

$$d(p, x_n) = \sqrt{\sum_{j=1}^k d_j(p_j, x_{nj})^2} < \sqrt{\sum_{j=1}^k \frac{\varepsilon^2}{k}} = \varepsilon$$

für alle $n \geq m_0 = \max(n_{01}, \dots, n_{0k})$. Also konvergiert x_n gegen p im Produktraum (X, d) . \square

Wir betrachten als Beispiel den Euklidischen Raum \mathbb{R}^k . Er ist das k -fache kartesische Produkt von \mathbb{R} mit der Standardmetrik. Die komponentenweise Konvergenz aus Satz 7.4 entspricht der Konvergenz von Folgen im Euklidischen Raum, die wir in Abschnitt 6.4 bereits benutzt hatten. Als Konsequenz erhält man aus den Grenzwertsätzen für Zahlenfolgen, dass auch die algebraischen Strukturen auf dem Euklidischen Vektorraum \mathbb{R}^k mit der Grenzwertbildung verträglich sind:

Satz 7.5 *Seien (x_n) und (y_n) zwei Folgen im Euklidischen Raum \mathbb{R}^k , die gegen p bzw. q konvergieren, $\langle \cdot, \cdot \rangle$ das Euklidische Skalarprodukt und $\| \cdot \|$ die Euklidische Norm. Dann gilt:*

1. $\lim_{n \rightarrow \infty} (\lambda \cdot x_n + \mu \cdot y_n) = \lambda \cdot \lim_{n \rightarrow \infty} x_n + \mu \cdot \lim_{n \rightarrow \infty} y_n = \lambda \cdot p + \mu \cdot q$ für alle $\lambda, \mu \in \mathbb{R}$.
2. $\lim_{n \rightarrow \infty} \langle x_n, y_n \rangle = \langle \lim_{n \rightarrow \infty} x_n, \lim_{n \rightarrow \infty} y_n \rangle = \langle p, q \rangle$.
3. $\lim_{n \rightarrow \infty} \|x_n\| = \left\| \lim_{n \rightarrow \infty} x_n \right\| = \|p\|$.

Wie bei Zahlenfolgen möchte man auch für Folgen in metrischen Räumen die Konvergenz untersuchen, ohne den Grenzwert bereits zu kennen. Dies geht in den *vollständigen* metrischen Räumen, die wir jetzt definieren werden.

Definition 7.6. *Sei (X, d) ein metrischer Raum. Eine Folge (x_n) in (X, d) heißt Cauchy-Folge, wenn zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass $d(x_n, x_m) < \varepsilon$ für alle $n, m \geq n_0$.*

Satz 7.6 *Jede konvergente Folge in einem metrischen Raum ist eine Cauchy-Folge.*

Beweis. Sei (x_n) eine gegen p konvergente Folge und $\varepsilon > 0$. Dann existiert ein $n_0 \in \mathbb{N}$ mit $d(p, x_n) < \frac{\varepsilon}{2}$ für alle $n \geq n_0$. Für $n, m \geq n_0$ folgt aus der Dreiecksungleichung

$$d(x_n, x_m) \leq d(x_n, p) + d(p, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Somit ist (x_n) eine Cauchy-Folge. \square

Wir wissen bereits aus Kapitel 3.1.3, dass die Umkehrung von Satz 7.6 nicht gilt. Es gibt metrische Räume mit Cauchy-Folgen, die nicht konvergieren.

Definition 7.7. *Ein metrischer Raum (X, d) heißt vollständig, wenn jede Cauchy-Folge in (X, d) konvergiert.*

In einem *vollständigen* metrischen Raum kann man also die Konvergenz einer Folge untersuchen, ohne ihren Grenzwert zu kennen. Man prüft dafür die *Cauchy-Bedingung*

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \text{ mit } d(x_n, x_m) < \varepsilon \quad \forall n, m \geq n_0.$$

Beispiel 1: Die reellen Zahlen \mathbb{R} mit der Standardmetrik $d(x, y) = |x - y|$, sowie die komplexen Zahlen \mathbb{C} mit der Standardmetrik $d(z, w) := |z - w|$ sind vollständige metrische Räume.

Beispiel 2: Der metrische Raum der rationalen Zahlen \mathbb{Q} mit der Standardmetrik $d(x, y) = |x - y|$ ist *nicht* vollständig (siehe Kapitel 3.1.3).

Beispiel 3: Wir betrachten $X = (-\frac{\pi}{2}, \frac{\pi}{2}) \subset \mathbb{R}$ mit den Metriken d und \widehat{d} :

$$\begin{aligned} d(x, y) &= |x - y|, \\ \widehat{d}(x, y) &= |\tan(x) - \tan(y)|. \end{aligned}$$

Dann ist (X, d) nicht vollständig, aber (X, \widehat{d}) vollständig (Übungsaufgabe).

Beispiel 4: Sei $X = C^0([a, b], \mathbb{R})$ der Vektorraum der stetigen Funktionen von $[a, b]$ in \mathbb{R} mit den Metriken d_∞ und d_1 :

$$\begin{aligned} d_\infty(f, g) &= \max\{|f(x) - g(x)| \mid x \in [a, b]\}, \\ d_1(f, g) &= \int_a^b |f(x) - g(x)| dx. \end{aligned}$$

Dann ist (X, d_∞) vollständig, aber (X, d_1) nicht vollständig (Übungsaufgabe).

Die Vollständigkeit eines metrischen Raumes (X, d) ist also von der auf X gewählten Metrik abhängig.

Satz 7.7 Sind $(X_1, d_1), \dots, (X_k, d_k)$ vollständige metrische Räume, so ist auch das kartesische Produkt (X, d) dieser metrischen Räume vollständig.

Beweis. Sei (x_n) mit $x_n = (x_{n1}, \dots, x_{nk})$, eine Cauchy-Folge im kartesischen Produkt $X = X_1 \times \dots \times X_k$ und $\varepsilon > 0$. Dann existiert ein $n_0 \in \mathbb{N}$, so dass

$$d(x_n, x_m) = \sqrt{\sum_{i=1}^k d_i(x_{ni}, x_{mi})^2} < \varepsilon \quad \forall n, m \geq n_0.$$

Daraus erhalten wir $d_j(x_{nj}, x_{mj}) < \varepsilon$ für alle $n, m \geq n_0$ und jedes $j \in \{1, \dots, k\}$. Folglich sind die Folgen $(x_{nj})_{n=1}^\infty$ Cauchy-Folgen in (X_j, d_j) für jedes $j \in \{1, \dots, k\}$. Da (X_j, d_j) vollständige metrische Räume sind, konvergieren die Folgen $(x_{nj})_{n=1}^\infty$ gegen ein $p_j \in X_j$. Somit gilt nach Satz 7.4 $\lim_{n \rightarrow \infty} x_n = (p_1, \dots, p_k) \in X_1 \times \dots \times X_k$. \square

Folgerung 7.1 *Der Euklidische Raum \mathbb{R}^k ist vollständig.*

Bemerkung: Am Ende von Kapitel 7.4 werden wir zeigen, dass *jede* Norm auf dem Vektorraum \mathbb{R}^k eine vollständige Metrik auf \mathbb{R}^k induziert. Insbesondere sind die normierten Vektorräume $(\mathbb{R}^k, \|\cdot\|_p)$ für $p \in [1, \infty)$ und $(\mathbb{R}^k, \|\cdot\|_\infty)$ mit

$$\|x\|_p := \left(\sum_{j=1}^k |x_j|^p \right)^{\frac{1}{p}},$$

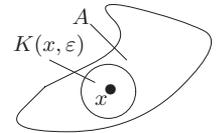
$$\|x\|_\infty := \max\{|x_1|, \dots, |x_k|\},$$

vollständig (siehe auch Übungsaufgabe 18). $\|\cdot\|_2$ ist die Euklidische Norm.

7.3 Spezielle Teilmengen in metrischen Räumen

In diesem Abschnitt werden wir spezielle Sorten von Teilmengen metrischer Räume definieren, die die von uns bisher im "naiven" Sinne benutzten Bezeichnungen *offene* und *abgeschlossene* Intervalle, Randpunkte von Intervallen bzw. Rand von Gebieten in der Ebene verallgemeinern.

Definition 7.8. *Eine Teilmenge A eines metrischen Raumes (X, d) heißt offen, wenn zu jedem $p \in A$ eine ε -Kugel $K(p, \varepsilon)$ existiert mit $K(p, \varepsilon) \subset A$.*



Beispiel 1: Die ε -Kugeln $K(p, \varepsilon)$ eines metrischen Raumes (X, d) sind offen.

Um dies zu zeigen, betrachten wir einen beliebigen Punkt $x \in K(p, \varepsilon)$ und zeigen, dass $K(x, \varepsilon - d(p, x)) \subset K(p, \varepsilon)$: Sei $z \in K(x, \varepsilon - d(p, x))$. Dann gilt $d(x, z) < \varepsilon - d(p, x)$ und folglich $d(z, x) + d(x, p) < \varepsilon$. Aus der Dreiecksungleichung folgt dann $d(p, z) \leq d(p, x) + d(x, z) < \varepsilon$. Folglich liegt z in $K(p, \varepsilon)$.

Insbesondere sind die Intervalle (a, b) im metrischen Raum \mathbb{R} (mit der Standardmetrik) in diesem Sinne offen. Die Intervalle der Form $(-\infty, a)$ und (a, ∞) sind ebenfalls offen.

Definition 7.9. *Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt abgeschlossen, wenn ihr Komplement $X \setminus A$ offen ist.*

Beispiel 2: Es gilt $[a, b] = \mathbb{R} \setminus ((-\infty, a) \cup (b, \infty))$, $[a, \infty) = \mathbb{R} \setminus (-\infty, a)$ und $(-\infty, a] = \mathbb{R} \setminus (a, \infty)$. Folglich sind die Intervalle $[a, b]$, $[a, \infty)$, $(-\infty, a]$ abgeschlossen.

Beispiel 3: Sei (X, d) der diskrete metrische Raum.

Dann ist *jede* Teilmenge $A \subset X$ offen, denn für jedes $p \in A$ gilt $K(p, \frac{1}{2}) = \{p\} \subset A$. Folglich ist auch *jede* Teilmenge von X abgeschlossen.

Die Eigenschaft einer Menge, abgeschlossen zu sein, kann man auch durch konvergente Folgen charakterisieren. In Kapitel 4.3. hatten wir abgeschlossene Teilmengen in \mathbb{R} und \mathbb{C} auf diese Weise definiert.

Satz 7.8 Sei (X, d) ein metrischer Raum. Dann gilt:

Eine Teilmenge $A \subset X$ ist genau dann abgeschlossen, wenn der Grenzwert jeder konvergenten Folge von Punkten aus A selbst in A liegt.

Beweis. (\implies) Sei $A \subset X$ abgeschlossen und (a_n) eine Folge in A , die gegen $x \in X$ konvergiert. Angenommen $x \notin A$. Dann ist $x \in X \setminus A$. Da $X \setminus A$ offen ist, existiert eine ε -Kugel $K(x, \varepsilon) \subset X \setminus A$. Da (a_n) gegen x konvergiert, existiert ein $n_0 \in \mathbb{N}$, so dass $a_n \in K(x, \varepsilon)$ für alle $n \geq n_0$. Dann können diese a_n nicht in A liegen (Widerspruch).

(\impliedby) Angenommen A ist nicht abgeschlossen, d.h. $X \setminus A$ nicht offen. Dann existiert ein $x \in X \setminus A$, so dass $K(x, \varepsilon) \cap A \neq \emptyset$ für alle $\varepsilon > 0$. Wir wählen nun $a_n \in K(x, \frac{1}{n}) \cap A$. Dann ist (a_n) eine Folge in A mit $d(x, a_n) \leq \frac{1}{n} \rightarrow 0$. Folglich konvergiert (a_n) gegen x . Nach Voraussetzung liegt aber x nicht in A (Widerspruch). \square

Man kann jede Teilmenge $A \subset X$ in eine offene und eine abgeschlossene Menge einschließen. Dazu definieren wir das Innere $\text{Int}(A)$ und den Abschluß $\text{cl}(A)$ von A , so dass

$$\text{Int}(A) \subset A \subset \text{cl}(A).$$

Definition 7.10. Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge von X .

1. Das Innere von A ist die Menge

$$\text{Int}(A) := \{x \in A \mid \exists \varepsilon > 0 \text{ mit } K(x, \varepsilon) \subset A\}.$$

Ein Punkt $x \in \text{Int}(A)$ heißt innerer Punkt von A .

2. Der Abschluss von A ist die Menge

$$\text{cl}(A) := \{x \in X \mid \forall \varepsilon > 0 \text{ gilt } K(x, \varepsilon) \cap A \neq \emptyset\}.$$

3. Der Rand von A ist die Menge

$$\partial A := \text{cl}(A) \setminus \text{Int}(A).$$

Ein Punkt $x \in \partial A$ heißt Randpunkt von A .

Satz 7.9 Für jede Teilmenge A eines metrischen Raumes gilt:

1. $\text{Int}(A)$ ist offen.
2. $\text{cl}(A) = X \setminus \text{Int}(X \setminus A)$. Insbesondere ist $\text{cl}(A)$ abgeschlossen.
3. $\partial A = X \setminus (\text{Int}(A) \cup \text{Int}(X \setminus A))$. Insbesondere ist ∂A abgeschlossen.
4. Ein Punkt $x \in X$ ist genau dann ein Randpunkt von A , wenn für alle $\varepsilon > 0$ gilt

$$K(x, \varepsilon) \cap A \neq \emptyset \text{ und } K(x, \varepsilon) \cap (X \setminus A) \neq \emptyset.$$

Beweis. Zu 1: Sei $x \in \text{Int}(A)$. Dann existiert eine ε -Kugel $K(x, \varepsilon) \subset A$. Für jeden Punkt $z \in K(x, \varepsilon)$ gilt $K(z, \varepsilon - d(x, z)) \subset K(x, \varepsilon) \subset A$. Folglich ist jeder Punkt von $K(x, \varepsilon)$ ein innerer Punkt von A , d.h. $K(x, \varepsilon) \subset \text{Int}(A)$. Somit ist $\text{Int}(A)$ offen.

Zu 2: Es gilt:

$$\begin{aligned} x \in X \setminus \text{Int}(X \setminus A) &\iff x \in X \text{ und } x \notin \text{Int}(X \setminus A) \\ &\iff \forall \varepsilon > 0 \text{ gilt } K(x, \varepsilon) \not\subset X \setminus A \\ &\iff \forall \varepsilon > 0 \text{ gilt } K(x, \varepsilon) \cap A \neq \emptyset \\ &\iff x \in \text{cl}(A). \end{aligned}$$

Zu 3: Aus $\text{cl}(A) = X \setminus \text{Int}(X \setminus A)$ folgt

$$\partial A = \text{cl}(A) \setminus \text{Int}(A) = (X \setminus \text{Int}(X \setminus A)) \setminus \text{Int}(A) = X \setminus \underbrace{(\text{Int}(X \setminus A) \cup \text{Int}(A))}_{\text{offen}}.$$

Zu 4: Daraus erhalten wir

$$\begin{aligned} x \in \partial A &\iff x \notin \text{Int}(A) \text{ und } x \notin \text{Int}(X \setminus A) \\ &\iff \forall \varepsilon > 0 \text{ gilt } K(x, \varepsilon) \cap (X \setminus A) \neq \emptyset \text{ und } K(x, \varepsilon) \cap A \neq \emptyset. \end{aligned}$$

□

Beispiel 4: Wir betrachten \mathbb{R} mit der Standardmetrik $d(x, y) = |x - y|$.

Für $A = [a, b)$ erhalten wir $\text{Int}(A) = (a, b)$, $\text{cl}(A) = [a, b]$ und $\partial A = \{a, b\}$.

Beispiel 5: Sei \mathbb{R}^2 der Euklidische Raum mit der Euklidischen Norm $\|\cdot\|$.

Für $A = K(x, \varepsilon)$ gilt:

- $\text{Int}(A) = K(x, \varepsilon)$,
- $\partial A = \{y \in \mathbb{R}^2 \mid \|x - y\| = \varepsilon\}$,
- $\text{cl}(A) = \{y \in \mathbb{R}^2 \mid \|x - y\| \leq \varepsilon\}$.

Als nächstes definieren wir die kompakten Teilmengen eines metrischen Raumes, die wir für die metrischen Räume \mathbb{R} und \mathbb{C} bereits aus Kapitel 4.3 kennen.

Ist (x_n) eine Folge in (X, d) . Eine Folge der Form (x_{n_k}) mit $n_1 < n_2 < n_3 < \dots$ heißt *Teilfolge* von (x_n) .

Definition 7.11. Eine Teilmenge A eines metrischen Raumes (X, d) heißt *kompakt*, falls jede Folge in A eine konvergente Teilfolge besitzt, deren Grenzwert in A liegt¹.

Beispiel 6: Eine Teilmenge $A \subset \mathbb{R}$ bzw. $B \subset \mathbb{C}$ ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.

Satz 7.10 Jede kompakte Teilmenge eines metrischen Raumes ist abgeschlossen und beschränkt.

¹ Man kann kompakte Mengen äquivalent auch durch Eigenschaften offener Überdeckungen definieren. Wir wollen dies hier aus Zeitgründen nicht tun, sondern verweisen dazu auf O. Deiser. Analysis 2.

Beweis. Sei A eine kompakte Teilmenge des metrischen Raumes (X, d) .

1. Wir zeigen mit dem Folgenkriterium aus Satz 7.8, dass A abgeschlossen ist:

Sei (a_n) eine Folge in A , die gegen $x \in X$ konvergiert. Wir müssen zeigen, dass x in A liegt. Da A kompakt ist, besitzt (a_n) eine Teilfolge (a_{n_k}) , die gegen einen Punkt $a \in A$ konvergiert. Es gilt aber $a = \lim_{k \rightarrow \infty} a_{n_k} = \lim_{n \rightarrow \infty} a_n = x$. Folglich ist $x = a \in A$.

2. Wir zeigen, dass A beschränkt ist:

Sei $p \in X$. Angenommen, A wäre nicht beschränkt. Dann gilt $A \not\subset K(p, n)$ für jedes $n \in \mathbb{N}$. Wir wählen jeweils einen Punkt $a_n \in A \setminus K(p, n)$. Dann ist (a_n) eine Folge in A mit $d(a_n, p) \geq n$. Da A kompakt ist, besitzt (a_n) eine Teilfolge (a_{n_k}) , die gegen ein $a \in A$ konvergiert. Dann existiert ein $k_0 \in \mathbb{N}$, so dass $d(a, a_{n_k}) < 1$ für alle $k \geq k_0$. Daraus folgt für $k \geq k_0$:

$$d(p, a_{n_k}) \leq d(p, a) + d(a, a_{n_k}) < d(p, a) + 1.$$

Dies ist ein Widerspruch, da $d(p, a_{n_k}) \geq n_k \rightarrow \infty$. □

Satz 7.11 *Sei (X, d) ein metrischer Raum und $B \subset X$ kompakt. Dann ist jede abgeschlossene Teilmenge $A \subset B$ ebenfalls kompakt.*

Beweis. Sei (a_n) eine beliebige Folge in A . Dann ist (a_n) auch Folge in B und besitzt, da B kompakt ist, eine in B konvergente Teilfolge (a_{n_j}) . Sei $b = \lim_{j \rightarrow \infty} a_{n_j}$. Da A abgeschlossen ist, liegt der Grenzwert b in A . Also enthält (a_n) eine konvergente Teilfolge mit Grenzwert in A . Damit ist A kompakt. □

Satz 7.12 *Sei (X, d) das kartesische Produkt der metrischen Räume $(X_1, d_1), \dots, (X_k, d_k)$ und seien $A_j \subset X_j$ kompakte Mengen in (X_j, d_j) , $j = 1, \dots, k$. Dann ist die Menge $A := A_1 \times A_2 \times \dots \times A_k$ ebenfalls kompakt in (X, d) .*

Beweis. Den Beweis wird analog zum Beweis von Satz 7.7 geführt. Wir überlassen ihn deshalb als Übungsaufgabe. □

Beispiel 7: Die Quader $W := [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_k, b_k] \subset \mathbb{R}^k$ sind kompakte Teilmengen des Euklidischen Raumes \mathbb{R}^k .

Wir erhalten daraus den folgenden Satz:

Satz 7.13 *Eine Teilmenge des Euklidischen Raumes \mathbb{R}^k ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.*

Beweis. Nach Satz 7.10 ist jede kompakte Menge abgeschlossen und beschränkt. Wir müssen die Umkehrung zeigen. Sei $A \subset \mathbb{R}^k$ beschränkt und abgeschlossen. Da A beschränkt ist, gibt es eine Kugel, die A enthält. Jede Kugel des Euklidischen Raumes ist in einem Quader W enthalten. Folglich gilt $A \subset W$. Da W kompakt und A abgeschlossen ist, ist A nach Satz 7.11 kompakt. □

Beispiel 8: Die Sphäre $S_r^{n-1} := \{x \in \mathbb{R}^n \mid \|x\| = r\} \subset \mathbb{R}^n$ und die abgeschlossene Kugel $D_r^n := \{x \in \mathbb{R}^n \mid \|x\| \leq r\} \subset \mathbb{R}^n$ sind kompakt.

Beispiel 9: Es gibt metrische Räume mit abgeschlossenen und beschränkten Teilmengen, die *nicht* kompakt sind.

Sei z.B. X eine unendliche Menge mit der diskreten Metrik

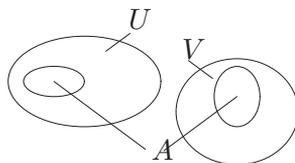
$$d(x, y) = \begin{cases} 0 & x = y \\ 1 & x \neq y. \end{cases}$$

In diesem metrischen Raum ist *jede* Teilmenge abgeschlossen und beschränkt. Eine abzählbare Teilmenge $A := \{a_1, a_2, \dots\} \subset X$ ist aber nicht kompakt: Eine Folge in (X, d) ist genau dann konvergent, wenn sie ab einem bestimmten Index konstant ist. Folglich besitzt die Folge (a_n) keine konvergente Teilfolge.

Abschließend verallgemeinern wir eine weitere Eigenschaft, die Intervalle in \mathbb{R} haben: Sie zerfallen nicht in zwei disjunkte offene Mengen.

Definition 7.12. Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt *zusammenhängend*, wenn es keine offenen und zueinander disjunkten Teilmengen $U, V \subset X$ gibt, so dass $A \subset U \cup V$, $A \cap U \neq \emptyset$ und $A \cap V \neq \emptyset$.

Beispiel 10: Sei $X = \mathbb{R}^2$, $U, V \subset X$ offene Teilmengen von X und $A \subset X$ wie im Bild. Dann ist A nicht zusammenhängend.



Wir beschreiben nun die zusammenhängenden Mengen im metrischen Raum \mathbb{R} mit der Standardmetrik $d(x, y) = |x - y|$.

Satz 7.14 Eine Teilmenge $A \subset \mathbb{R}$ ist genau dann zusammenhängend, wenn sie ein Intervall ist.

Beweis. (\implies) Sei $A \subset \mathbb{R}$ zusammenhängend. Angenommen A wäre kein Intervall. Dann existieren $a, b \in A$ mit $[a, b] \not\subset A$, d.h. es gibt ein $x \in (a, b)$ mit $x \notin A$. Wir betrachten die Mengen $U := (-\infty, x)$ und $V := (x, \infty)$. U und V sind disjunkte offene Teilmengen in \mathbb{R} mit $A \subset U \cup V = \mathbb{R} \setminus \{x\}$. Außerdem ist $A \cap U \neq \emptyset$, da $a \in A \cap U$ und $A \cap V \neq \emptyset$, da $b \in A \cap V$. Dies widerspricht aber der Voraussetzung, dass A zusammenhängend ist.

(\impliedby) Sei A ein Intervall in \mathbb{R} . Angenommen, A wäre nicht zusammenhängend. Dann existieren offene, disjunkte Mengen $U, V \subset \mathbb{R}$, so dass $A \subset U \cup V$ und $a \in A \cap U$, $b \in A \cap V$. Da A ein Intervall ist, gilt $[a, b] \subset A \subset U \cup V$. Wir betrachten die Funktion $f : U \cup V \rightarrow \mathbb{R}$, definiert durch

$$f(x) := \begin{cases} 1 & \text{falls } x \in U, \\ -1 & \text{falls } x \in V. \end{cases}$$

f ist stetig: Sei $x \in U$. Da U offen ist, gibt es eine Kugel $K(x, \varepsilon) \subset U$. Folglich gilt für jede Folge (x_n) in $U \cup V$, die gegen x konvergiert, dass $x_n \in K(x, \varepsilon) \subset U$ für alle $n \geq n_0$. Somit ist $f(x_n) = 1$ für alle $n \geq n_0$, d.h. $f(x_n)$ konvergiert gegen $f(x) = 1$. Nach dem Folgenkriterium für stetige reelle Funktionen ist f in x stetig. Analog zeigt man, dass f in jedem Punkt $x \in V$ stetig ist. Folglich ist $f|_{[a,b]}$ stetig. Es gilt $f(a) = 1$ und $f(b) = -1$, f nimmt aber nur die beiden Werte -1 und 1 an. Dies widerspricht dem Zwischenwertsatz für stetige reelle Funktionen. \square

7.4 Stetige Abbildungen und ihre Eigenschaften

Wir definieren die Stetigkeit von Abbildungen zwischen beliebigen metrischen Räumen völlig analog zum Fall reeller Funktionen. Wir werden sehen, dass sich wichtige Eigenschaften, die wir für stetige reelle Funktionen aus Kapitel 4 kennen, auf den allgemeinen Fall übertragen.

In diesem Abschnitt bezeichnen (X, d_X) und (Y, d_Y) stets metrische Räume.

Definition 7.13. Sei $f : A \subset X \rightarrow Y$ eine Abbildung.

1. f heißt im Punkt $p \in A$ stetig, falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass gilt:

$$\forall x \in A \text{ mit } d_X(p, x) < \delta \implies d_Y(f(p), f(x)) < \varepsilon,$$

bzw., so dass für die Kugeln gilt:

$$f(K_X(p, \delta) \cap A) \subset K_Y(f(p), \varepsilon).$$

2. Die Abbildung f heißt stetig, wenn sie in jedem Punkt $p \in A$ stetig ist.

Satz 7.15 (Folgenkriterium für Stetigkeit)

Eine Abbildung $f : A \subset X \rightarrow Y$ ist genau dann in $p \in A$ stetig, wenn für jede Folge (x_n) in A , die gegen p konvergiert, die Bildfolge $(f(x_n))$ gegen $f(p)$ konvergiert.

Beweis. Der Beweis ist völlig analog zum Beweis des entsprechenden Satz für reellwertige Funktionen. \square

Wie wir bereits für reelle Funktionen gesehen hatten, ist dieses Kriterium ein äußerst nützliches Hilfsmittel, um weitere Eigenschaften stetiger Abbildungen mit Hilfe von Grenzwertsätzen für konvergente Folgen zu beweisen. Zunächst betrachten wir einige Beispiele, die sofort aus dem Folgenkriterium folgen.

Beispiel 1:

- Die identische Abbildung $Id_X : X \rightarrow X$, $Id_X(x) := x$, ist stetig.
- Die konstante Abbildung $c_p : X \rightarrow X$, $c_p(x) := p$, ist stetig.

Beispiel 2: Die Projektionen auf die Faktoren eines Produktraumes sind stetig:

Seien $(X_1, d_1), \dots, (X_k, d_k)$ metrische Räume und (X, d) ihr kartesisches Produkt. Dann ist jede Projektion

$$\begin{aligned} \pi_j : X = X_1 \times \dots \times X_k &\longrightarrow X_j \\ (x_1, \dots, x_k) &\longmapsto x_j \end{aligned}$$

stetig.

Beispiel 3: Alle algebraischen Operationen auf dem Euklidischen Vektorraum \mathbb{R}^k sind stetig, d.h.

die Summe	$+$	$:\mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}^k,$
die skalare Multiplikation	\cdot	$:\mathbb{R} \times \mathbb{R}^k \rightarrow \mathbb{R}^k,$
das Euklidische Skalarprodukt	$\langle \cdot, \cdot \rangle$	$:\mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R},$
die Euklidische Norm	$\ \cdot\ $	$:\mathbb{R}^k \rightarrow \mathbb{R}$

sind stetige Abbildungen. Dies folgt mit dem Folgenkriterium aus den Grenzwertsätzen (Satz 7.5) für Folgen im \mathbb{R}^k .

Beispiel 4: Jede lineare Abbildung $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ zwischen den Euklidischen Vektorräumen \mathbb{R}^k und \mathbb{R}^m ist stetig.

Beweis: Sei (v_n) eine Folge von Vektoren im \mathbb{R}^k mit $v_n = (v_{n1}, \dots, v_{nk})$, die gegen einen Vektor $w = (w_1, \dots, w_k) \in \mathbb{R}^k$ konvergiert. Dann konvergiert nach Satz 7.4 für jedes $j = 1, \dots, k$ die Komponentenfolge (v_{nj}) gegen w_j . Sei $e_j := (0, \dots, 0, \underbrace{1}_{j.\text{Stelle}}, 0, \dots, 0)$ der

j -te kanonische Basisvektor des \mathbb{R}^k . Dann folgt aus der Linearität von L und der Stetigkeit der linearen Operationen auf \mathbb{R}^m :

$$L(v_n) = L\left(\sum_{j=1}^k v_{nj}e_j\right) = \sum_{j=1}^k v_{nj}L(e_j) \longrightarrow \sum_{j=1}^k w_jL(e_j) = L\left(\sum_{j=1}^k w_je_j\right) = L(w).$$

Nach dem Folgenkriterium ist L somit stetig.

Satz 7.16 Seien $f : A \subset X \rightarrow Y$ und $g : B \subset Y \rightarrow Z$ zwei Abbildungen zwischen metrischen Räumen, $f(A) \subset B$ und $g \circ f : A \subset X \rightarrow Z$ die Verknüpfung von f und g . Dann gilt: Ist f in $p \in A$ und g in $f(p) \in B$ stetig, so ist $g \circ f$ in $p \in A$ stetig.

Insbesondere ist die Verknüpfung stetiger Abbildungen ebenfalls stetig.

Beweis. Analog zum Beweis des analogen Satzes über stetige reellwertige Funktionen. \square

Satz 7.17 Sei $f : A \subset X \rightarrow Y_1 \times \dots \times Y_k$ eine Abbildung in das kartesische Produkt metrischer Räume und $f = (f_1, \dots, f_k)$ die Komponentendarstellung von f :

$$f(x) =: (f_1(x), \dots, f_k(x)) \quad \forall x \in X.$$

Dann gilt: Die Abbildung f ist genau dann in $p \in A$ stetig, wenn jede Komponentenfunktion $f_j : A \subset X \rightarrow Y_j$, $j = 1, \dots, k$, in p stetig ist.

Beweis. (\implies) Da $f_j = \pi_j \circ f$, folgt dies aus Beispiel 2 und Satz 7.16.

(\impliedby) Sei (x_n) eine Folge in A , die gegen $p \in A$ konvergiert. Da alle f_j in p stetig sind, konvergiert $(f_j(x_n))$ gegen $f_j(p) \in Y_j$ für jedes $j = 1, \dots, k$. Somit konvergiert $(f(x_n)) = ((f_1(x_n), \dots, f_k(x_n)))$ gegen $(f_1(p), \dots, f_k(p)) = f(p)$, d.h. f ist in p stetig. \square

Dies entspricht dem von uns in Kapitel 6.4. für parametrisierte Kurven $\gamma : I \subset \mathbb{R} \rightarrow \mathbb{R}^k$ benutzten Stetigkeitsbegriff.

Neben dem Folgenkriterium ist auch das folgende topologische Kriterium für die Stetigkeit von Abbildungen nützlich:

Satz 7.18 (Topologisches Kriterium für Stetigkeit)

Eine Abbildung $f : X \rightarrow Y$ zwischen metrischen Räumen ist genau dann stetig, wenn das Urbild $f^{-1}(U) \subset X$ jeder offenen Menge $U \subset Y$ offen ist.

Beweis. (\implies) Sei $f : X \rightarrow Y$ stetig und $U \subset Y$ offen. Wir zeigen, dass das Urbild

$$f^{-1}(U) := \{x \in X \mid f(x) \in U\} \subset X$$

dann ebenfalls offen ist.

Ist $f^{-1}(U) = \emptyset$, so ist die Behauptung erfüllt, da \emptyset offen ist. Sei also $f^{-1}(U) \neq \emptyset$ und $p \in f^{-1}(U)$ ein beliebig gewählter Punkt. Dann gilt $f(p) \in U$, und da $U \subset Y$ offen ist, existiert ein $\varepsilon > 0$ mit $K_Y(f(p), \varepsilon) \subset U$. Da f in p stetig ist, gibt es ein $\delta > 0$, so dass

$$f(K_X(p, \delta)) \subset K_Y(f(p), \varepsilon) \subset U.$$

Aus der Definition des Urbildes folgt dann:

$$K_X(p, \delta) \subset f^{-1}(U).$$

Somit ist $f^{-1}(U)$ offen in X .

(\impliedby) Sei $f^{-1}(U) \subset X$ offen für jede offene Menge $U \subset Y$. Wir zeigen, dass f stetig ist. Wir wählen einen beliebigen Punkt $p \in X$ und ein beliebiges $\varepsilon > 0$. Da die Kugel $K_Y(f(p), \varepsilon)$ in Y offen ist, ist nach Voraussetzung auch das Urbild $f^{-1}(K_Y(f(p), \varepsilon))$ in X offen. Es enthält den Punkt p . Folglich existiert ein $\delta > 0$, so dass $K_X(p, \delta) \subset f^{-1}(K_Y(f(p), \varepsilon))$. Nach Definition des Urbildes folgt daraus $f(K_X(p, \delta)) \subset K_Y(f(p), \varepsilon)$. Das bedeutet aber, dass f in p stetig ist. Da $p \in X$ beliebig gewählt war, ist f stetig. \square

Genauso wie für reelle Funktionen hat man für Abbildungen zwischen metrischen Räumen zwei stärkere Stetigkeitsbegriffe.

Definition 7.14. Sei $f : A \subset X \rightarrow Y$ eine Abbildung zwischen metrischen Räumen.

1. f heißt *gleichmäßig stetig*, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass gilt:

$$\forall p, q \in A \text{ mit } d_X(p, q) < \delta \implies d_Y(f(p), f(q)) < \varepsilon.$$

(Im Unterschied zur Definition der Stetigkeit hängt hier die Größe von δ nur von ε , aber nicht von p oder q ab.)

2. f heißt *lipschitzstetig*, wenn es eine positive Konstante $L \in \mathbb{R}^+$ gibt, so dass gilt:

$$d_Y(f(p), f(q)) \leq L \cdot d_X(p, q) \quad \forall p, q \in A.$$

L heißt *Lipschitz-Konstante* von f .

Satz 7.19 Sei $f : A \subset X \rightarrow Y$ eine Abbildung zwischen metrischen Räumen. Dann gilt:

1. f ist lipschitzstetig $\implies f$ ist gleichmäßig stetig.
2. f ist gleichmäßig stetig $\implies f$ ist stetig.

Beweis. Analog zum Beweis des analogen Satzes in Teil 1. □

Beispiel 5: Jede lineare Abbildung $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ zwischen den Euklidischen Vektorräumen \mathbb{R}^k und \mathbb{R}^m ist sogar lipschitzstetig (Übungsaufgabe).

Wir betrachten nun das Verhalten von stetigen Abbildungen auf zusammenhängenden Mengen und verallgemeinern den Zwischenwertsatz für reelle Funktionen.

Satz 7.20 Sei $f : X \rightarrow Y$ eine stetige Abbildung zwischen metrischen Räumen. Dann ist das Bild jeder zusammenhängenden Menge ebenfalls zusammenhängend.

Beweis. Sei $A \subset X$ eine zusammenhängende Menge. Wir wollen zeigen, dass dann auch das Bild $f(A) \subset Y$ zusammenhängend ist. Angenommen $f(A) \subset Y$ wäre nicht zusammenhängend. Dann existieren offene Mengen $U, V \subset Y$ mit

- $U \cap V = \emptyset$,
- $f(A) \subset U \cup V$,
- $f(A) \cap U \neq \emptyset$ und $f(A) \cap V \neq \emptyset$.

Da f stetig ist, sind nach Satz 7.18 die Urbilder $f^{-1}(U)$ und $f^{-1}(V)$ ebenfalls offen. Desweiteren gilt für diese Urbilder

- $f^{-1}(U) \cap f^{-1}(V) = f^{-1}(U \cap V) = f^{-1}(\emptyset) = \emptyset$.
- $A \subset f^{-1}(f(A)) \subset f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V)$.
- Nach Voraussetzung existiert ein $a \in A$ mit $f(a) \in U$ und ein $b \in A$ mit $f(b) \in V$.
Dann gilt: $a \in A \cap f^{-1}(U) \neq \emptyset$ und $b \in A \cap f^{-1}(V) \neq \emptyset$.

Folglich ist A nicht zusammenhängend, was ein Widerspruch zur Voraussetzung ist. □

Wir definieren noch einen weiteren Zusammenhangsbegriff für Teilmengen metrischer Räume (X, d) .

Definition 7.15. Eine Teilmenge $A \subset X$ heißt *wegzusammenhängend*, falls zu je zwei Punkten $a, b \in A$ eine stetige Abbildung $\omega : [0, 1] \subset \mathbb{R} \rightarrow X$ existiert mit $\omega([0, 1]) \subset A$, $\omega(0) = a$ und $\omega(1) = b$.

Die Abbildung ω heißt *Weg* in A von a nach b .

Beispiel 6: *Beziehung zwischen den beiden Zusammenhangsbegriffen*

a) Für Teilmengen $A \subset \mathbb{R}$ sind die Zusammenhangsbegriffe äquivalent:

$$A \text{ zusammenhängend} \stackrel{S, 7.14}{\iff} A \text{ Intervall} \iff A \text{ wegzusammenhängend.}$$

b) In metrischen Räumen ist jede wegzusammenhängende Menge auch zusammenhängend, aber die Umkehrung gilt i.a. nicht (für ein Beispiel siehe O. Deiser: Analysis 2, Seite 198).

c) Konvexe Mengen im Euklidischen Raum \mathbb{R}^k :

Eine Teilmenge $A \subset \mathbb{R}^k$ heißt *konvex*, wenn mit je zwei Punkten $x, y \in A$ auch die Verbindungsstrecke $\overline{xy} := \{x + t(y - x) \mid t \in [0, 1]\}$ vollständig in A liegt.

Jede konvexe Menge ist wegzusammenhängend, da die Abbildung $\omega : [0, 1] \rightarrow A \subset \mathbb{R}^k$,

$$\omega(t) := x + t(y - x),$$

stetig ist und die Punkte x und y verbindet.

Zum Beispiel ist jede Kugel $K(x_0, r)$ und ihr Abschluß konvex.

Der nächste Satz ist eine weitere Verallgemeinerung des Zwischenwertsatzes für reelle Funktionen.

Satz 7.21 *Sei $f : X \rightarrow Y$ eine stetige Abbildung. Dann ist das Bild jeder wegzusammenhängenden Teilmenge ebenfalls wegzusammenhängend.*

Beweis. Sei $A \subset X$ wegzusammenhängend. Wir wollen zeigen, dass das Bild $f(A)$ ebenfalls wegzusammenhängend ist. Seien $x, y \in f(A)$ zwei Punkte in $f(A)$. Dann existieren $a, b \in A$ mit $x = f(a)$, $y = f(b)$. Da A wegzusammenhängend ist, gibt es einen (stetigen) Weg $\omega : [0, 1] \rightarrow A \subset X$ von a nach b . Wir betrachten die Abbildung $f \circ \omega : [0, 1] \rightarrow f(A) \subset Y$. Da f und ω stetig sind, ist $f \circ \omega$ ebenfalls stetig. Weiterhin gilt $(f \circ \omega)(0) = f(\omega(0)) = f(a) = x$ und $(f \circ \omega)(1) = f(\omega(1)) = f(b) = y$, d.h. $f \circ \omega$ ist ein Weg in $f(A)$ von x nach y . \square

Als nächstes beschäftigen wir uns mit dem Verhalten von stetigen Abbildungen auf kompakten Mengen. Die für stetige reelle Funktionen bekannten Aussagen kann man durch Ersetzen der Metrik direkt auf den Fall von Abbildungen zwischen metrischen Räumen übertragen:

Satz 7.22 *Sei $f : X \rightarrow Y$ eine stetige Abbildung. Dann gilt:*

1. *Das Bild jeder kompakten Menge ist kompakt.*
2. *f ist auf jeder kompakten Menge $K \subset X$ gleichmäßig stetig.*
3. *Ist f insbesondere reell-wertig, so nimmt f auf jeder kompakten Menge $K \subset X$ ein Maximum und ein Minimum an, d.h. es existieren Punkte $\xi_1, \xi_2 \in K$ mit*

$$f(\xi_1) \leq f(x) \leq f(\xi_2) \quad \forall x \in K.$$

Beweis. Beweis von 1. ist analog zum Beweis von Satz 4.12, Beweis von 2. analog zum Beweis von Satz 4.14, Beweis von 3. analog zum Beweis von Satz 4.13. \square

Definition 7.16. Eine Abbildung $f : X \rightarrow Y$ zwischen zwei metrischen Räumen heißt *Homöomorphismus*, wenn f bijektiv ist und f und f^{-1} stetig sind.

Zwei metrische Räume X und Y heißen *homöomorph*, falls es einen Homöomorphismus $f : X \rightarrow Y$ gibt.

Bei homöomorphen metrischen Räumen bleibt zwar die Abstandsmessung nicht erhalten, aber alle topologischen Eigenschaften von Teilmengen (offen, abgeschlossen, kompakt, zusammenhängend, wegzusammenhängend, ...) übertragen sich auf die Bildmenge. Die Konvergenz einer Folge überträgt sich ebenfalls auf die Konvergenz der Bildfolge.

Beispiel 7: Die inverse Abbildung einer bijektiven, stetigen Abbildung ist im Allgemeinen nicht stetig, wie das folgende Beispiel zeigt:

Seien $X = (0, 1) \cup \{2\} \subset \mathbb{R}$ und $Y = (0, 1] \subset \mathbb{R}$, versehen mit der Standardmetrik $d(x, y) = |x - y|$. Wir betrachten die Abbildung $f : X \rightarrow Y$ gegeben durch

$$f(t) := \begin{cases} t & \text{falls } t \in (0, 1) \\ 1 & \text{falls } t = 2. \end{cases}$$

f ist stetig und bijektiv. Aber $f^{-1} : (0, 1] \rightarrow (0, 1) \cup \{2\}$ ist nach Zwischenwertsatz nicht stetig.

Für kompakte metrische Räume gilt aber folgender nützliche Satz:

Satz 7.23 (Satz über die Stetigkeit der inversen Abbildung)

Sei $f : X \rightarrow Y$ eine bijektive stetige Abbildung zwischen metrischen Räumen und sei X kompakt. Dann ist die inverse Abbildung $f^{-1} : Y \rightarrow X$ stetig.

Beweis. Wir benutzen das topologische Kriterium für Stetigkeit (Satz 7.18). Sei $U \subset X$ offen. Wir zeigen, dass das Urbild $(f^{-1})^{-1}(U) = f(U) \subset Y$ offen ist.

Da $U \subset X$ offen ist, ist $X \setminus U \subset X$ abgeschlossen. Da X kompakt ist, ist $X \setminus U \subset X$ ebenfalls kompakt (Satz 7.11). Dann ist nach Satz 7.22 auch das Bild $f(X \setminus U)$ in Y kompakt und insbesondere abgeschlossen (Satz 7.10). Da f bijektiv ist, gilt $f(X \setminus U) = Y \setminus f(U)$. Also ist die Teilmenge $f(U) \subset Y$ offen. \square

Definition 7.17. Sei (X, d) ein metrischer Raum. Eine Abbildung $f : X \rightarrow X$ heißt *kontrahierend* (oder *Kontraktion*), wenn sie Lipschitzstetig mit einer Lipschitzkonstanten $0 < L < 1$ ist, das heißt wenn eine Konstante $0 < L < 1$ existiert, so dass

$$d(f(x), f(y)) \leq L \cdot d(x, y) \quad \text{für alle } x, y \in X.$$

Für *kontrahierende* Abbildungen in *vollständigen* metrischen Räumen gilt der folgende sehr nützliche Fixpunktsatz, den wir für reelle Funktionen bereits kennen (Übungsaufgabe 35, Analysis 1).

Satz 7.24 (Banachscher Fixpunktsatz)

Sei (X, d) ein vollständiger metrischer Raum und $f : X \rightarrow X$ eine kontrahierende Abbildung. Dann hat f genau einen Fixpunkt. Diesen Fixpunkt erhält man konstruktiv:

Sei $x_0 \in X$ ein beliebiger Punkt und $x_n := \underbrace{f \circ \dots \circ f}_{n\text{-mal}}(x_0) =: f^n(x_0)$. Dann konvergiert die Folge (x_n) gegen den Fixpunkt von f .

Beweis. (1) Eindeutigkeit des Fixpunktes: Angenommen es existieren zwei Fixpunkte ξ und η von f , dh. es gelte $f(\xi) = \xi$ und $f(\eta) = \eta$. Aus der Kontraktivität von f erhält man

$$d(\xi, \eta) = d(f(\xi), f(\eta)) \leq L \cdot d(\xi, \eta).$$

Da aber $0 < L < 1$ gilt, folgt $d(\xi, \eta) = 0$, also $\xi = \eta$.

(2) Existenz des Fixpunktes: Sei $x_0 \in X$ ein beliebig gewählter Punkt. Wir definieren eine Folge von Punkten in X durch

$$\begin{aligned} x_1 &:= f(x_0), \\ x_2 &:= f(x_1) = f^2(x_0), \\ &\vdots \\ x_n &:= f(x_{n-1}) = f^2(x_{n-2}) = \dots = f^n(x_0). \end{aligned}$$

Dann gilt

$$\begin{aligned} d(x_{n+1}, x_n) &= d(f(x_n), f(x_{n-1})) \\ &\leq L \cdot d(x_n, x_{n-1}) = L \cdot d(f(x_{n-1}), f(x_{n-2})) \\ &\leq L^2 \cdot d(x_{n-1}, x_{n-2}) = L^2 \cdot d(f(x_{n-2}), f(x_{n-3})) \\ &\vdots \\ &\leq L^n \cdot d(x_1, x_0). \end{aligned}$$

Für $n > m$ folgt mit der Dreiecksungleichung und der Formel für die geometrische Summe

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n-1}) + d(x_{n-1}, x_{n-2}) + \dots + d(x_{m+1}, x_m) \\ &\leq (L^{n-1} + \dots + L^m) \cdot d(x_0, x_1) \\ &= L^m (L^0 + L^1 + \dots + L^{n-1-m}) \cdot d(x_0, x_1) \\ &= L^m \cdot \frac{1 - L^{n-m}}{1 - L} \cdot d(x_0, x_1) \\ &\leq L^m \cdot \frac{1}{1 - L} \cdot d(x_0, x_1). \end{aligned}$$

Da $0 < L < 1$, ist (L^m) eine Nullfolge. Somit ist (x_n) eine Cauchyfolge in X . Da der metrische Raum X vollständig ist, konvergiert diese Cauchyfolge gegen einen Punkt ξ

aus X . Wir zeigen nun, dass dieser Grenzwert ξ der gesuchte Fixpunkt von f ist. Dazu betrachten wir

$$\begin{aligned} d(f(\xi), \xi) &\leq d(f(\xi), x_n) + d(x_n, \xi) \\ &= d(f(\xi), f(x_{n-1})) + d(x_n, \xi) \\ &\leq L \underbrace{d(\xi, x_{n-1})}_{\rightarrow 0} + \underbrace{d(x_n, \xi)}_{\rightarrow 0}. \end{aligned}$$

Daraus folgt $d(f(\xi), \xi) = 0$, also $f(\xi) = \xi$. □

Abschließend beweisen wir einen Satz, der zeigt, dass alle Eigenschaften, die wir in diesem Kapitel für den Euklidischen Raum \mathbb{R}^k bewiesen haben, auch gelten, wenn wir den \mathbb{R}^k mit einer *beliebigen* Norm versehen.

Satz 7.25 Sei \mathbb{R}^k der Euklidische Vektorraum mit der Euklidischen Norm $\|\cdot\|$ und sei $N : \mathbb{R}^k \rightarrow \mathbb{R}$ eine weitere Norm auf \mathbb{R}^k . Dann gilt:

1. $N : \mathbb{R}^k \rightarrow \mathbb{R}$ ist Lipschitzstetig.
2. Es existieren Konstanten $a, b \in \mathbb{R}^+$, so dass

$$a \cdot \|x\| \leq N(x) \leq b \cdot \|x\| \quad \forall x \in \mathbb{R}^k.$$

Beweis. 1) Wir zeigen mittels der Cauchy-Schwarzschen Ungleichung und der Normeigenschaften die Existenz von $b \in \mathbb{R}^+$:

Sei $x \in \mathbb{R}^k$ und $x = \sum_{j=1}^k x_j e_j$ die Basisdarstellung von x bezüglich der kanonischen Basis (e_1, \dots, e_k) im \mathbb{R}^k . Dann gilt:

$$\begin{aligned} N(x) &= N\left(\sum_{j=1}^k x_j e_j\right) \stackrel{\Delta}{\leq} \sum_{j=1}^k N(x_j e_j) = \sum_{j=1}^k |x_j| N(e_j) \\ &\stackrel{CSU}{\leq} \sqrt{\sum_{j=1}^k |x_j|^2} \cdot \sqrt{\sum_{j=1}^k N(e_j)^2} \\ &= \|x\| \cdot \underbrace{\sqrt{\sum_{j=1}^k N(e_j)^2}}_{=: b > 0}. \end{aligned}$$

2) Für eine Norm N gilt die umgekehrte Dreiecksungleichung

$$|N(x) - N(y)| \leq N(x - y) \quad \forall x, y \in \mathbb{R}^k. \quad (7.1)$$

Dies folgt aus zweifacher Anwendung der Dreiecksungleichung für N :

$$\begin{aligned} N(x) &= N(y + (x - y)) \stackrel{\Delta}{\leq} N(y) + N(x - y), \\ N(y) &= N(x + (y - x)) \stackrel{\Delta}{\leq} N(x) + N(y - x) = N(x) + N(x - y). \end{aligned}$$

Durch Umstellen erhalten wir $\pm(N(x) - N(y)) \leq N(x - y)$. Also gilt (7.1).

3) N ist Lipschitzstetig mit der Lipschitz-Konstanten $b \in \mathbb{R}^+$, da

$$|N(x) - N(y)| \stackrel{2)}{\leq} N(x - y) \stackrel{1)}{\leq} b \cdot \|x - y\| \quad \forall x, y \in \mathbb{R}^k.$$

4) Wir zeigen die Existenz von $a \in \mathbb{R}^+$:

Für $x = 0$ gilt die im Satz behauptete Ungleichung immer. Sei also $x \neq 0$. Dann gilt

$$N(x) = N\left(\frac{x}{\|x\|}\right) = \|x\| \cdot N\left(\frac{x}{\|x\|}\right). \quad (7.2)$$

Da $\left\|\frac{x}{\|x\|}\right\| = \frac{\|x\|}{\|x\|} = 1$, liegt $\frac{x}{\|x\|}$ in der Euklidischen Sphäre $S^{k-1} := \{y \in \mathbb{R}^k \mid \|y\| = 1\}$. Die Abbildung $N : \mathbb{R}^k \rightarrow \mathbb{R}$ ist stetig und die Menge $S^{k-1} \subset \mathbb{R}^k$ ist kompakt. Folglich existiert das Minimum

$$a := \min\{N(y) \mid y \in S^{k-1}\}$$

(Satz 7.22). Dann ist $a = N(p)$ für ein $p \in S^{k-1}$, also $a > 0$. Aus (7.2) folgt dann

$$N(x) \geq \|x\| \cdot a \quad \forall x \in \mathbb{R}^k. \quad \square$$

Als Konsequenz erhalten wir

Satz 7.26 Sei $\|\cdot\|$ die Euklidische Norm und N eine weitere Norm auf dem Vektorraum \mathbb{R}^k . Dann stimmen alle topologischen Eigenschaften von $(\mathbb{R}^k, \|\cdot\|)$ und (\mathbb{R}^k, N) überein. D.h. für Folgen (x_n) im \mathbb{R}^k und für Teilmengen $A \subset \mathbb{R}^k$ gilt:

1. (x_n) konvergiert gegen p in $(\mathbb{R}^k, \|\cdot\|)$ \iff (x_n) konvergiert gegen p in (\mathbb{R}^k, N) .
2. (x_n) ist Cauchy-Folge in $(\mathbb{R}^k, \|\cdot\|)$ \iff (x_n) ist Cauchy-Folge in (\mathbb{R}^k, N) .
3. A ist offen (abgeschlossen, beschränkt, kompakt, zusammenhängend bzw. wegzusammenhängend) bzgl. $\|\cdot\|$ genau dann, wenn A offen (abgeschlossen, beschränkt, kompakt, zusammenhängend bzw. wegzusammenhängend) bzgl. N ist.

Insbesondere ist der normierte Vektorraum (\mathbb{R}^k, N) für jede Norm N vollständig.

Beweis. Aus der Abschätzung in Satz 7.25 folgt für die Kugeln $K_{eu}(x, \varepsilon)$ bzgl. der Norm $\|\cdot\|$ und $K_N(x, \varepsilon)$ bzgl. der Norm N :

$$K_{eu}(x, \varepsilon) \subset K_N(x, b \cdot \varepsilon) \quad \text{und} \quad K_N(x, \varepsilon) \subset K_{eu}(x, \frac{\varepsilon}{a}) \quad \forall \varepsilon > 0.$$

Da Konvergenz von Folgen, Cauchy-Folgen, offene und beschränkte Mengen mittels Abständen bzw. ε -Kugeln definiert sind, folgt die Äquivalenz dieser Eigenschaften für $\|\cdot\|$ und N . Abgeschlossene und kompakte Mengen sind durch Konvergenzeigenschaften von Folgen charakterisiert. Folglich stimmen diese Eigenschaften für beide Normen überein. Der Zusammenhang ist durch offene Mengen, der Wegzusammenhang durch stetige Abbildungen, also ebenfalls durch Konvergenzeigenschaften charakterisiert, folglich stimmen auch diese Eigenschaften für beide Normen überein. \square

Differentialrechnung für Funktionen mehrerer reeller Variablen

Nachdem wir in Kapitel 5 die Differentialrechnung für Funktionen *einer* reellen Variablen behandelt haben, wollen wir uns jetzt mit Funktionen bzw. Abbildungen befassen, die von *mehreren* reellen Variablen abhängen, d.h. mit Abbildungen der Form $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$. Auch für solche Abbildungen werden wir einen Differenzierbarkeitsbegriff einführen, der der Abbildung f in jedem Punkt $p \in U$ eine Ableitung zuordnet. Aus den Eigenschaften der Ableitung möchte man wieder Schlüsse für die Eigenschaften von f selbst ziehen. Für eine Abbildung $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ ist die Ableitung im Punkt $p \in U$ eine *lineare Abbildung* vom \mathbb{R}^k in den \mathbb{R}^m . *Lineare* Abbildungen sind im Gegensatz zu beliebigen Abbildungen sehr "einfach". Durch die lineare Struktur kann man mit ihnen sehr gut umgehen (siehe die Vorlesungen über lineare Algebra). Wir können also Methoden der linearen Algebra benutzen, um mit Hilfe der Ableitungen analytische Eigenschaften von beliebigen Abbildungen $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ zu untersuchen. Uns interessieren dabei z.B. folgende Fragen:

- Wie kann man lokale Extrema von Funktionen $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ finden?
- Wie kann man Tangentialebenen an Flächen im \mathbb{R}^3 berechnen.
- Wann gibt es Lösungen der Gleichung $f(x_1, \dots, x_k) = 0$. Wie sieht Lösungsmenge aus?
- Wie kann man einer Abbildung $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^k$ ansehen, ob sie (wenigstens lokal) umkehrbar ist? Wann ist sie eine Koordinatentransformation?

Für die Differentialrechnung spielt neben der linearen Struktur der benutzen Räume wieder die Konvergenz von Folgen eine entscheidende Rolle. Wir versehen deshalb die Räume \mathbb{R}^k bzw. \mathbb{R}^m mit der Euklidischen Norm $\|\cdot\|$ und betrachten den dadurch induzierten Konvergenzbegriff. Wie wir aus Kapitel 7.4 wissen, ist es für Konvergenzfragen allerdings egal, welche Norm auf dem \mathbb{R}^k bzw. \mathbb{R}^m wir benutzen.

In den Vektorräumen \mathbb{R}^k benutzen wir oft die kanonische Basis (e_1, \dots, e_k) , wobei

$$e_i := (0, \dots, 0, \underbrace{1}_{i.\text{Stelle}}, 0, \dots, 0) \in \mathbb{R}^k.$$

Im folgenden seien die Definitionsbereiche $U \subset \mathbb{R}^k$ der Abbildungen immer *offene* Teilmengen, d.h. jeder Punkt $p \in U$ ist ein innerer Punkt.

8.1 Differenzierbare Abbildungen, das Differential, die Richtungsableitungen und die partiellen Ableitungen

Bevor wir uns mit der Differenzierbarkeit von Abbildungen mehrerer reeller Variablen befassen, erinnern wir uns nochmal daran, wie wir die Differenzierbarkeit für Funktionen einer Variablen in Kapitel 5 definiert hatten:

Eine Funktion $f : U \subset \mathbb{R} \rightarrow \mathbb{R}$ heißt in $p \in U$ differenzierbar, wenn der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(p+h) - f(p)}{h} =: f'(p) \in \mathbb{R} \quad (8.1)$$

existiert. Den Begriff des Grenzwertes kann man für Abbildungen mehrerer reeller Variablen genauso definieren wie für reelle Funktionen einer Variablen.

Definition 8.1. Sei $\phi : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine Abbildung.

Man sagt: ϕ besitzt in $q \in U$ den Grenzwert $y \in \mathbb{R}^m$, wenn für jede Folge (x_n) in $U \setminus \{q\}$, die gegen q konvergiert, die Bildfolge $(\phi(x_n))$ gegen y konvergiert.

Bezeichnung: $\lim_{x \rightarrow q} \phi(x) = y$.

Die Bedingung (8.1) lässt sich auch schreiben als

$$\lim_{h \rightarrow 0} \frac{f(p+h) - f(p) - f'(p) \cdot h}{h} = 0$$

oder in der Form

$$f(p+h) = f(p) + f'(p) \cdot h + r_p(h),$$

wobei $r_p(h) \in \mathbb{R}$ ein Fehlerterm ist, für den

$$\lim_{h \rightarrow 0} \frac{r_p(h)}{h} = 0$$

gilt. Die Zuordnung $L : h \in \mathbb{R} \mapsto f'(p) \cdot h \in \mathbb{R}$ ist dabei eine lineare Abbildung.

Das motiviert die folgende Definition für Abbildungen, die von *mehreren* reellen Variablen abhängen:

Definition 8.2. Eine Abbildung $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ heißt in $p \in U$ differenzierbar, wenn eine lineare Abbildung $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ existiert, so dass gilt

$$\lim_{h \rightarrow 0} \frac{f(p+h) - f(p) - L(h)}{\|h\|} = 0. \quad (8.2)$$

Die lineare Abbildung $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ heißt Ableitung von f in p oder auch Differential von f in p . Wir bezeichnen die Ableitung bzw. das Differential mit $L =: df_p$.

$f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ heißt differenzierbar, wenn f in jedem $p \in U$ differenzierbar ist.¹

¹ In den Ingenieurwissenschaften sagt man auch f ist in p total differenzierbar und nennt df_p das totale Differential von f in p .

Die Bedingung (8.2) ist äquivalent zu

$$\lim_{h \rightarrow 0} \frac{\|f(p+h) - f(p) - L(h)\|}{\|h\|} = 0. \quad (8.3)$$

Wir können (8.2) auch in der folgenden Form schreiben:

$$f(p+h) = f(p) + L(h) + r_p(h) \quad \text{mit} \quad \lim_{h \rightarrow 0} \frac{r_p(h)}{\|h\|} = 0. \quad (8.4)$$

Satz 8.1 Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ in $p \in U$ differenzierbar. Dann gilt:

1. Das Differential $df_p : \mathbb{R}^k \rightarrow \mathbb{R}^m$ ist eindeutig bestimmt.

2. f ist in p stetig.

3. Sei $f =: (f_1, \dots, f_m) : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ die Komponentendarstellung von f .

Die Abbildung f ist genau dann in $p \in U$ differenzierbar, wenn jede Komponente f_1, \dots, f_m in $p \in U$ differenzierbar ist. In diesem Fall gilt

$$df_p = ((df_1)_p, \dots, (df_m)_p).$$

Beweis. Zu 1) Seien $L, \tilde{L} : \mathbb{R}^k \rightarrow \mathbb{R}^m$ zwei lineare Abbildungen mit

$$f(p+h) - f(p) = L(h) + r_p(h) = \tilde{L}(h) + \tilde{r}_p(h)$$

und $\lim_{h \rightarrow 0} \frac{r_p(h)}{\|h\|} = \lim_{h \rightarrow 0} \frac{\tilde{r}_p(h)}{\|h\|} = 0$. Dann folgt

$$\lim_{h \rightarrow 0} \frac{L(h) - \tilde{L}(h)}{\|h\|} = 0. \quad (8.5)$$

Da beide Abbildungen linear sind, gilt $L(0) = \tilde{L}(0) = 0$. Sei nun $x \in \mathbb{R}^k$, $x \neq 0$ und $t \in \mathbb{R}^+$. Wir setzen in (8.5) $h := tx$ mit $t \rightarrow 0^+$ ein. Wegen der Linearität von L und \tilde{L} erhalten wir

$$0 = \lim_{t \rightarrow 0^+} \frac{L(tx) - \tilde{L}(tx)}{\|tx\|} = \lim_{t \rightarrow 0^+} \frac{L(x) - \tilde{L}(x)}{\|x\|} = \frac{L(x) - \tilde{L}(x)}{\|x\|}.$$

Also ist $L(x) = \tilde{L}(x)$ für alle $x \in \mathbb{R}^k$.

Zu 2) Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ in $p \in U$ differenzierbar. Wir benutzen das Folgenkriterium, um die Stetigkeit von f in $p \in U$ zu zeigen. Sei (x_n) eine gegen p konvergente Folge in U und $x_n \neq p$. Dann gilt nach Dreiecksungleichung

$$\begin{aligned} \|f(x_n) - f(p)\| &\leq \|f(x_n) - f(p) - L(x_n - p)\| + \|L(x_n - p)\| \\ &= \underbrace{\frac{\|f(x_n) - f(p) - L(x_n - p)\|}{\|x_n - p\|}}_{\rightarrow 0} \cdot \underbrace{\|x_n - p\|}_{\rightarrow 0} + \underbrace{\|L(x_n - p)\|}_{\rightarrow 0}. \end{aligned}$$

Die ersten beiden Konvergenzen folgen aus der Definition der Differenzierbarkeit und der Stetigkeit der Norm. Für die dritte Konvergenz benutzen wir außerdem die Stetigkeit der

linearen Abbildung L . Folglich konvergiert die Folge $(f(x_n))$ gegen $f(p)$, d.h. f ist in $p \in U$ stetig.

Zu 3) Die Bedingung

$$\lim_{h \rightarrow 0} \frac{f(p+h) - f(p) - L(h)}{\|h\|} = 0$$

für eine Abbildung $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ ist äquivalent zur Bedingung an die Komponenten

$$\lim_{h \rightarrow 0} \frac{f_j(p+h) - f_j(p) - L_j(h)}{\|h\|} = 0 \quad \forall j = 1, \dots, m,$$

wobei $L =: (L_1, \dots, L_m)$ die Komponenten von L bezeichnen. $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ ist genau dann linear, wenn alle Komponenten $L_j : \mathbb{R}^k \rightarrow \mathbb{R}$ linear sind. Somit ist f genau dann in p differenzierbar, wenn alle f_1, \dots, f_m in p differenzierbar sind und es gilt $df_p = L = (L_1, \dots, L_m) = ((df_1)_p, \dots, (df_m)_p)$. \square

Beispiel 1: Für Abbildungen $f : U \subset \mathbb{R} \rightarrow \mathbb{R}^m$, die nur von einer reellen Variablen abhängen, stimmen die bisherigen Ableitungsbegriffe überein, wenn man den Vektorraum \mathbb{R}^m mit den linearen Abbildungen $L(\mathbb{R}, \mathbb{R}^m)$ identifiziert. Es gilt:

$$f'(p) = f'(p) \cdot 1 = df_p(1) \in \mathbb{R}^m.$$

Beispiel 2: *Das Differential einer linearen Abbildung.*

Eine lineare Abbildung $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ ist in jedem Punkt $p \in \mathbb{R}^k$ differenzierbar und es gilt $dL_p = L$ für alle $p \in \mathbb{R}^k$:

Beweis: Wegen der Linearität von L gilt

$$\lim_{h \rightarrow 0} \frac{L(p+h) - L(p) - L(h)}{\|h\|} = \lim_{h \rightarrow 0} \frac{L(p) + L(h) - L(p) - L(h)}{\|h\|} = \lim_{h \rightarrow 0} \underbrace{\frac{0}{\|h\|}}_{=0} = 0.$$

L erfüllt also die Bedingungen für das Differential, folglich ist $L = dL_p$.

Definition 8.3. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine Abbildung und $\mathbf{a} \in \mathbb{R}^k$ ein fixierter Vektor. Man sagt: f besitzt in $p \in U$ eine Ableitung in Richtung \mathbf{a} , falls der Grenzwert

$$\nabla_{\mathbf{a}} f(p) := \lim_{t \rightarrow 0} \frac{f(p + t\mathbf{a}) - f(p)}{t} \in \mathbb{R}^m \quad (\text{hier ist } t \in \mathbb{R})$$

existiert. Dieser Grenzwert heißt die Richtungsableitung von f in Richtung \mathbf{a} an der Stelle p .

Bei der Richtungsableitung wird die Abbildung f also nur entlang des Geradenstückes $\{p + t\mathbf{a} \mid t \in (-\varepsilon, \varepsilon)\} \subset U$ betrachtet. Bezeichnet $\varphi_{\mathbf{a}} : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^m$ die Abbildung

$$\varphi_{\mathbf{a}}(t) := f(p + t\mathbf{a}),$$

so gilt

$$\nabla_{\mathbf{a}} f(p) = \lim_{t \rightarrow 0} \frac{\varphi_{\mathbf{a}}(t) - \varphi_{\mathbf{a}}(0)}{t} = \varphi'_{\mathbf{a}}(0).$$

Satz 8.2 Ist $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ in $p \in U$ differenzierbar, so besitzt f in p in jeder Richtung $\mathbf{a} \in \mathbb{R}^k$ die Richtungsableitung und es gilt

$$df_p(\mathbf{a}) = \nabla_{\mathbf{a}}f(p).$$

Beweis. Sei f in p differenzierbar, $df_p : \mathbb{R}^k \rightarrow \mathbb{R}^m$ das Differential und $\mathbf{a} \in \mathbb{R}^k$. Für $\mathbf{a} = 0$ folgt die Behauptung aus der Linearität von df_p . Für $\mathbf{a} \neq 0$ betrachten wir $h = t\mathbf{a}$. Dann gilt nach Definition der Differenzierbarkeit und der Linearität von df_p :

$$0 = \lim_{t \rightarrow 0} \frac{\|f(p + t\mathbf{a}) - f(p) - df_p(t\mathbf{a})\|}{\|t\mathbf{a}\|} = \frac{1}{\|\mathbf{a}\|} \cdot \lim_{t \rightarrow 0} \left\| \frac{f(p + t\mathbf{a}) - f(p)}{t} - df_p(\mathbf{a}) \right\|.$$

Damit existiert

$$\nabla_{\mathbf{a}}f(p) = \lim_{t \rightarrow 0} \frac{f(p + t\mathbf{a}) - f(p)}{t} = df_p(\mathbf{a}). \quad \square$$

Beispiel 3: Eine Abbildung kann in einem Punkt p alle Richtungsableitungen besitzen, ohne in diesem Punkt differenzierbar zu sein.

Wir betrachten dazu die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, definiert durch

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2+y^4} & \text{falls } (x, y) \neq (0, 0), \\ 0 & \text{falls } (x, y) = (0, 0). \end{cases}$$

Die Funktion f ist im Punkt $p := (0, 0)$ nicht stetig, also auch nicht differenzierbar. Um dies einzusehen, betrachten wir die Folge (a_n) mit $a_n := (\frac{1}{n^2}, \frac{1}{n})$. Dann gilt

$$a_n = (\frac{1}{n^2}, \frac{1}{n}) \rightarrow (0, 0), \quad f(a_n) = \frac{\frac{1}{n^4}}{\frac{1}{n^4} + \frac{1}{n^4}} = \frac{1}{2} \rightarrow \frac{1}{2} \neq f((0, 0)).$$

Folglich ist f in $p = (0, 0)$ nicht stetig.

Es existieren aber alle Richtungsableitungen $\nabla_{\mathbf{a}}f(p)$. Ist $\mathbf{a} = (0, 0)$, so folgt aus der Definition sofort $\nabla_{\mathbf{a}}f(x_0) = 0$. Für einen beliebigen Vektor $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ mit $\mathbf{a} \neq (0, 0)$ gilt:

$$\frac{f(p + t\mathbf{a}) - f(p)}{t} = \frac{f(t\mathbf{a})}{t} = \frac{t^3 a_1 a_2^2}{t(t^2 a_1^2 + t^4 a_2^4)} = \frac{a_1 a_2^2}{a_1^2 + t^2 a_2^4} \xrightarrow{t \rightarrow 0} \begin{cases} \frac{a_2^2}{a_1} & a_1 \neq 0, \\ 0 & a_1 = 0. \end{cases}$$

Somit existiert $\nabla_{\mathbf{a}}f(p)$ für alle $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$.

Wichtige Richtungsableitungen sind die Richtungsableitungen in Richtung der Koordinatenachsen, d.h. in Richtung der kanonischen Basisvektoren e_1, \dots, e_k des \mathbb{R}^k .

Definition 8.4. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine Abbildung und $p \in U$.

Man sagt: f besitzt in $p \in U$ die partielle Ableitung nach der i -ten Koordinate, wenn die Richtungsableitung $\nabla_{e_i}f(p)$ von f in p in Richtung e_i existiert.

Bezeichnung:

$$\frac{\partial f}{\partial x_i}(p) := \nabla_{e_i}f(p) = \lim_{t \rightarrow 0} \frac{f(p + te_i) - f(p)}{t} \in \mathbb{R}^m.$$

f heißt partiell differenzierbar, wenn alle partiellen Ableitungen

$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_k} : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ existieren.

Sei $p = (p_1, \dots, p_k)$ und $\varphi_i : (p_i - \varepsilon, p_i + \varepsilon) \subset \mathbb{R} \rightarrow \mathbb{R}^m$ die Funktion

$$\varphi_i(x) := f(p_1, \dots, p_{i-1}, x, p_{i+1}, \dots, p_k).$$

Dann gilt nach Definition

$$\frac{\partial f}{\partial x_i}(p) = \varphi_i'(p_i).$$

Dies motiviert den Namen *partielle Ableitung nach der i-ten Koordinate*. Man hält alle Koordinaten außer der i -ten fest und leitet im üblichen Sinne nach der i -ten Koordinate ab.

Beispiel 4: Wir betrachten die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, definiert durch $f(x, y) = \sin(xy^2)$. Die partiellen Ableitungen sind dann

$$\frac{\partial f}{\partial x}(x, y) = y^2 \cos(xy^2) \quad \text{und} \quad \frac{\partial f}{\partial y}(x, y) = 2xy \cos(xy^2).$$

Satz 8.3 Ist $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ in $p \in U$ differenzierbar, so existieren die partiellen Ableitungen $\frac{\partial f}{\partial x_1}(p), \dots, \frac{\partial f}{\partial x_k}(p)$ und es gilt

$$df_p(h) = \sum_{i=1}^k h_i \cdot \frac{\partial f}{\partial x_i}(p), \quad \text{wobei } h = (h_1, \dots, h_k) \in \mathbb{R}^k. \quad (8.6)$$

Beweis. Nach Satz 8.2 wissen wir, dass die Richtungsableitungen $\nabla_{e_i} f(p)$, und somit die partiellen Ableitungen von f in $p \in U$ existieren. Außerdem folgt aus der Definition der partiellen Ableitungen für $h = (h_1, \dots, h_k) = \sum_{i=1}^k h_i e_i$:

$$df_p(h) = df_p\left(\sum_{i=1}^k h_i e_i\right) = \sum_{i=1}^k h_i \cdot df_p(e_i) = \sum_{i=1}^k h_i \cdot \frac{\partial f}{\partial x_i}(p). \quad \square$$

Mit der Formel (8.6) kann man die Differentiale einer *differenzierbaren* Abbildung mit Hilfe der partiellen Ableitungen ausrechnen. Es gilt aber noch mehr! Die partiellen Ableitungen sind ein wichtiges Hilfsmittel, um festzustellen, ob eine Abbildung $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ überhaupt differenzierbar ist. Wir wissen aus Beispiel 3, dass die partiellen Ableitungen existieren können, ohne dass die Abbildung differenzierbar ist. Es gilt aber folgender stärkerer Satz.

Satz 8.4 (Hauptkriterium für Differenzierbarkeit)

Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine Abbildung und $p \in U$.

Wenn die partiellen Ableitungen $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ (auf der gesamten Menge U !) existieren und im Punkt $p \in U$ stetig sind, dann ist f in $p \in U$ differenzierbar.

Beweis. Auf Grund von Satz 8.1, Punkt 3) genügt es, die Behauptung für reellwertige Abbildungen zu beweisen. Sei also nun $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ reellwertig, $p = (p_1, \dots, p_k)$ und $h = (h_1, \dots, h_k) \in \mathbb{R}^k$ so klein, dass $h \in K(p, \varepsilon) \subset U$. Dann gilt

$$\begin{aligned}
f(p+h) - f(p) &= f(p_1 + h_1, p_2, \dots, p_k) - f(p_1, \dots, p_k) \\
&+ f(p_1 + h_1, p_2 + h_2, p_3, \dots, p_k) - f(p_1 + h_1, p_2, \dots, p_k) \\
&+ f(p_1 + h_1, p_2 + h_2, p_3 + h_3, p_4, \dots, p_k) - f(p_1 + h_1, p_2 + h_2, p_3, \dots, p_k) \\
&\vdots \\
&+ f(p_1 + h_1, \dots, p_k + h_k) - f(p_1 + h_1, \dots, p_{k-1} + h_{k-1}, p_k).
\end{aligned}$$

Wir betrachten die Funktion $\varphi_1(x) := f(x, p_2, \dots, p_k)$. Nach Voraussetzung existiert die Ableitung $\varphi_1'(x) = \frac{\partial f}{\partial x_1}(x, p_2, \dots, p_k)$ für alle x zwischen p_1 und $p_1 + h_1$. Wir wenden den Mittelwertsatz der Differentialrechnung auf φ_1 an, und erhalten eine Zahl ξ_1 zwischen p_1 und $p_1 + h_1$, so dass gilt

$$f(p_1 + h_1, p_2, \dots, p_k) - f(p_1, \dots, p_k) = h_1 \cdot \frac{\partial f}{\partial x_1}(\xi_1, p_2, \dots, p_k).$$

Analog existiert für alle $j \in \{2, \dots, k\}$ ein ξ_j zwischen p_j und $p_j + h_j$, so dass gilt

$$\begin{aligned}
&f(p_1 + h_1, \dots, p_j + h_j, p_{j+1}, \dots, p_k) - f(p_1 + h_1, \dots, p_{j-1} + h_{j-1}, p_j, p_{j+1}, \dots, p_k) \\
&= h_j \cdot \frac{\partial f}{\partial x_j}(\underbrace{p_1 + h_1, \dots, p_{j-1} + h_{j-1}, \xi_j, p_{j+1}, \dots, p_k}_{=: c_j = c_j(h)}).
\end{aligned}$$

Folglich gilt

$$\begin{aligned}
f(p+h) - f(p) &= \sum_{j=1}^k h_j \cdot \frac{\partial f}{\partial x_j}(c_j) = \underbrace{\sum_{j=1}^k h_j \cdot \frac{\partial f}{\partial x_j}(p)}_{=: L(h), L \text{ linear}} + \underbrace{\sum_{j=1}^k h_j \cdot \left(\frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(p) \right)}_{=: r_p(h)}.
\end{aligned}$$

Es bleibt zu zeigen, dass $\lim_{h \rightarrow 0} \frac{r_p(h)}{\|h\|} = 0$. Mit der Cauchy-Schwarzschen Ungleichung erhalten wir:

$$|r_p(h)| = \left| \sum_{j=1}^k h_j \cdot \underbrace{\left(\frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(p) \right)}_{=: b_j = b_j(h)} \right| \stackrel{CSU}{\leq} \|h\| \cdot \|b\|,$$

wobei $b := (b_1, \dots, b_k)$. Folglich gilt

$$0 \leq \frac{|r_p(h)|}{\|h\|} \leq \|b(h)\|.$$

Für $h \rightarrow 0$ gilt $c_j(h) \rightarrow p$ und somit wegen der Stetigkeit der partiellen Ableitungen in p auch $b_j(h) \rightarrow 0$, also $\|b(h)\| \rightarrow 0$. Daraus folgt

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

Folglich ist f in $p \in U$ differenzierbar und es gilt

$$df_p(h) = \sum_{j=1}^k h_j \cdot \frac{\partial f}{\partial x_j}(p).$$

□

Geometrische Bedeutung der Ableitung

1. Zur Erinnerung: Für eine reelle differenzierbare Funktion, die von einer Variablen abhängt, beschreibt man mit Hilfe der Ableitung die Tangente an den Graphen von f .

Sei $f : U \subset \mathbb{R} \rightarrow \mathbb{R}$ eine differenzierbare Funktion. Dann ist ihr Graph eine Kurve im \mathbb{R}^2 :

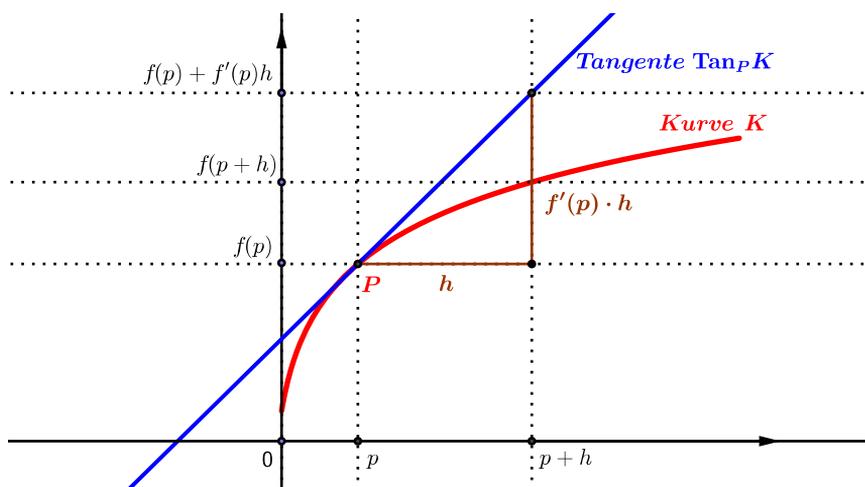
$$K := \{(x, f(x)) \mid x \in U\} \subset \mathbb{R}^2.$$

Sei $p \in U$. Die Tangente an K im Punkt $P := (p, f(p))$ ist der Graph der Funktion $T_1(f, p) : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$T_1(f, p)(x) := f(p) + f'(p)(x - p) \quad (1. \text{ Taylorpolynom} = \text{Tangentenfunktion}),$$

d.h.

$$\begin{aligned} \text{Tan}_P K &= \{(x, f(p) + f'(p)(x - p)) \mid x \in \mathbb{R}\} \\ &\stackrel{h=x-p}{=} \{(p, f(p)) + h \cdot (1, f'(p)) \mid h \in \mathbb{R}\} \\ &= P + \mathbb{R} \cdot (1, f'(p)) \subset \mathbb{R}^2. \end{aligned}$$



2. Für eine reelle differenzierbare Funktion, die von zwei Variablen abhängt, beschreibt man mit Hilfe der Ableitung die Tangentialebene an den Graphen von f .

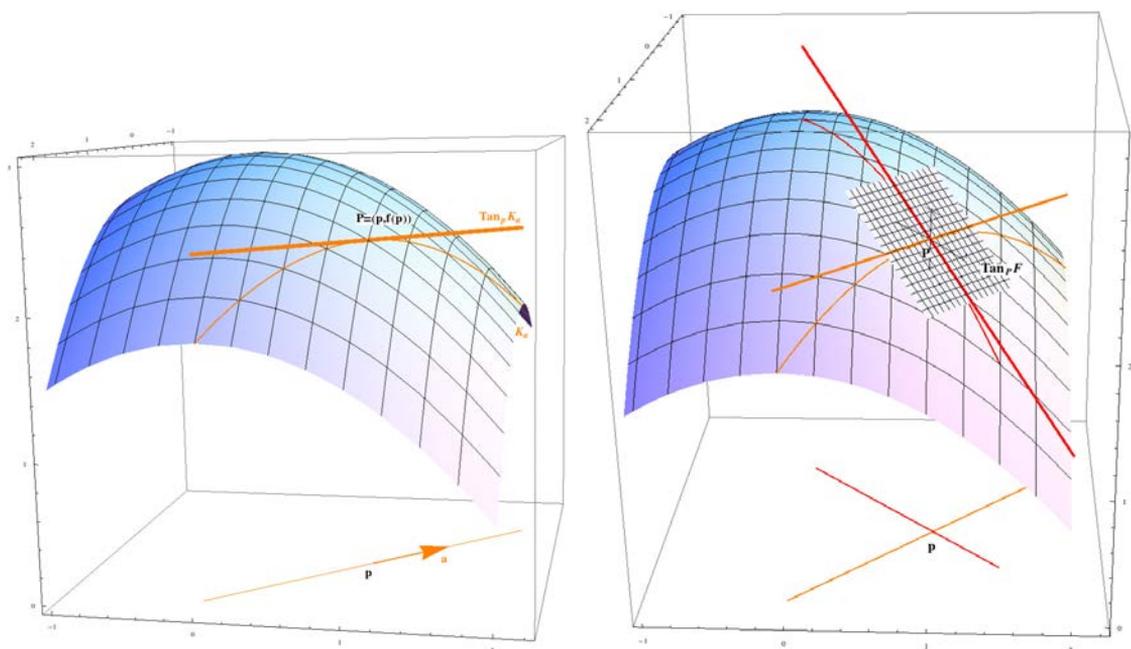
Sei $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ eine differenzierbare Funktion. Der Graph von f ist in diesem Fall eine Fläche im \mathbb{R}^3 :

$$F := \{(x, y, f(x, y)) \mid (x, y) \in U\} \subset \mathbb{R}^3.$$

Sei $p \in U$ und $df_p : \mathbb{R}^2 \rightarrow \mathbb{R}$ das Differential von f in p . Für jeden Vektor $\mathbf{a} \in \mathbb{R}^2$ mit $\mathbf{a} \neq 0$ betrachten wir die Funktion

$$\varphi_{\mathbf{a}} : t \in (-\varepsilon, \varepsilon) \mapsto f(p + t\mathbf{a}),$$

die das Verhalten von f entlang des Geradenstückes $\{p + t\mathbf{a} \mid t \in (-\varepsilon, \varepsilon)\} \subset U$ beschreibt.



Vielen Dank an Christoph Stadtmüller für die beiden Flächenbilder.

Der Graph $K_{\mathbf{a}}$ von $\varphi_{\mathbf{a}}$ ist eine Kurve auf der Fläche F , die durch den Punkt $P = (p, f(p)) = (p, \varphi_{\mathbf{a}}(0))$ läuft. Wir wenden 1) auf die Funktion $\varphi_{\mathbf{a}}$ an und erhalten für die Tangente an $K_{\mathbf{a}}$ im Punkt $P = (p, f(p))$

$$\begin{aligned} \text{Tan}_P K_{\mathbf{a}} &= \{(p + t\mathbf{a}, \varphi_{\mathbf{a}}(0) + \varphi'_{\mathbf{a}}(0) \cdot t \mid t \in \mathbb{R}\} \\ &= \{(p + t\mathbf{a}, f(p) + \nabla_{\mathbf{a}} f(p) \cdot t \mid t \in \mathbb{R}\} \\ &= \{(p, f(p)) + t \cdot (\mathbf{a}, df_p(\mathbf{a})) \mid t \in \mathbb{R}\} \\ &= P + \mathbb{R} \cdot (\mathbf{a}, df_p(\mathbf{a})). \end{aligned}$$

Die Ebene, die von allen diesen Tangenten $\text{Tan}_P K_{\mathbf{a}}$, für $\mathbf{a} \in \mathbb{R}^2$, gebildet wird, nennt man die *Tangentialebene an die Fläche F im Punkt P* und bezeichnet sie mit $\text{Tan}_P F$. Für die Tangentialebene gilt also:

$$\begin{aligned} \text{Tan}_P F &= P + \{(\mathbf{a}, df_p(\mathbf{a})) \mid \mathbf{a} \in \mathbb{R}^2\} \\ &\stackrel{\text{Basis } e_1, e_2}{=} P + \mathbb{R} \cdot (1, 0, \frac{\partial f}{\partial x}(p)) + \mathbb{R} \cdot (0, 1, \frac{\partial f}{\partial y}(p)). \end{aligned} \quad (8.7)$$

Die Tangentialebene an die Fläche F im Punkt P kann man ebenfalls als Graph des 1. Taylorpolynoms $T_1(f, p) : \mathbb{R}^2 \rightarrow \mathbb{R}$ beschreiben:

Das 1. Taylorpolynom ist in diesem Fall ein Polynom vom Grad 1 in den zwei Variablen x und y :

$$T_1(f, p)(x, y) := f(p) + \frac{\partial f}{\partial x}(p)(x - p_1) + \frac{\partial f}{\partial y}(p)(y - p_2) \quad \begin{aligned} &(1. \text{ Taylorpolynom} \\ &= \text{Tangentialebenenfunktion}). \end{aligned}$$

Aus (8.7) folgt für die Tangentialebene durch Einsetzen der Parameter $x - p_1$ und $y - p_2$

$$\text{Tan}_p F = \{(x, y, f(p) + \frac{\partial f}{\partial x}(p)(x - p_1) + \frac{\partial f}{\partial y}(p)(y - p_2) \mid (x, y) \in \mathbb{R}^2\}.$$

Wir werden auf die Taylorentwicklung für Abbildungen, die von mehreren Variablen abhängen, in Abschnitt 8.4 zurückkommen.

8.2 Kettenregeln und Mittelwertsätze

Für die Differentiale gelten die gleichen Rechenregeln wie für die Ableitungen von Funktionen einer reellen Variablen.

Satz 8.5 (Rechenregeln für Differentiale)

Seien $f, g : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$, $\varphi, \psi : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ und $p \in U$.

1. Sind f und g in p differenzierbar, so ist $f + g$ in p differenzierbar und es gilt

$$d(f + g)_p = df_p + dg_p.$$

2. Sind f und φ in p differenzierbar, so ist $\varphi \cdot f$ in p differenzierbar und es gilt

$$d(\varphi \cdot f)_p = d\varphi_p \cdot f(p) + \varphi(p) \cdot df_p.$$

3. Sind f und ψ in p differenzierbar und $\psi(p) \neq 0$, so ist $\frac{f}{\psi}$ in einer Umgebung von p definiert, in p differenzierbar und es gilt

$$d\left(\frac{f}{\psi}\right)_p = \frac{df_p \cdot \psi(p) - f(p) \cdot d\psi_p}{\psi(p)^2}.$$

Beweis. Die Aussagen 1) und 2) beweist man durch Nachprüfen der Bedingung (8.4). Die Aussage 3) folgt aus der Produktregel 2) und der bereits bekannten Quotientenregel für Funktionen einer Variablen:

$$d\left(\frac{1}{\psi}\right)_p(h) = \left(\frac{1}{\psi}\right)'(p) \cdot h = -\frac{\psi'(p)}{\psi(p)^2} \cdot h = -\frac{d\psi_p(h)}{\psi(p)^2}.$$

Wir lassen die Ausführung als Übungsaufgabe. □

Satz 8.6 (Kettenregel für Funktionen mehrerer Variablen)

Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ in p differenzierbar, $g : V \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ in $q := f(p)$ differenzierbar und $f(U) \subset V$. Dann ist die Verknüpfung $g \circ f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^n$ in p differenzierbar und es gilt

$$d(g \circ f)_p = dg_{f(p)} \circ df_p.$$

Für $k = 1$ gilt insbesondere: $(g \circ f)'(p) = dg_{f(p)}(f'(p))$.

Beweis. Wir setzen $L := df_p$ und $\tilde{L} := dg_q$ und betrachten die Abbildungen

$$\begin{aligned}\varphi(x) &:= f(x) - f(p) - L(x - p), \\ \psi(y) &:= g(y) - g(q) - \tilde{L}(y - q), \\ \varrho(x) &:= (g \circ f)(x) - (g \circ f)(p) - (\tilde{L} \circ L)(x - p).\end{aligned}$$

Nach Voraussetzung gilt

$$\lim_{x \rightarrow p} \frac{\varphi(x)}{\|x - p\|} = 0 \quad \text{und} \quad \lim_{y \rightarrow q} \frac{\psi(y)}{\|y - q\|} = 0. \quad (8.8)$$

Es ist zu zeigen, dass $\lim_{x \rightarrow p} \frac{\varrho(x)}{\|x - p\|} = 0$ gilt. Da

$$\begin{aligned}\varrho(x) &= g(f(x)) - g(f(p)) - \tilde{L}(L(x - p)) \\ &= g(f(x)) - g(q) - \tilde{L}(f(x) - f(p)) - \varphi(x) \\ &= g(f(x)) - g(q) - \tilde{L}(f(x) - q) + \tilde{L}(\varphi(x)) \\ &= \psi(f(x)) + \tilde{L}(\varphi(x)),\end{aligned}$$

genügt es dazu $\lim_{x \rightarrow p} \frac{\|\tilde{L}(\varphi(x))\|}{\|x - p\|} = 0$ und $\lim_{x \rightarrow p} \frac{\|\psi(f(x))\|}{\|x - p\|} = 0$ zu zeigen. Wir benutzen dazu die Abschätzung von $\|L(x)\|$ und $\|\tilde{L}(\varphi(x))\|$ durch die Operatornormen. Da $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ und $\tilde{L} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ linear sind, gilt für alle $x \in U$

$$\begin{aligned}\|L(x)\| &\leq \|L\| \cdot \|x\| \quad \text{und} \\ \|\tilde{L}(\varphi(x))\| &\leq \|\tilde{L}\| \cdot \|\varphi(x)\|,\end{aligned}$$

wobei $\|L\|$ und $\|\tilde{L}\|$ die Operatornormen von L bzw. \tilde{L} sind (siehe Übungsaufgabe 23, Analysis 2).

1) $\lim_{x \rightarrow p} \frac{\|\tilde{L}(\varphi(x))\|}{\|x - p\|} = 0$, denn

$$0 \leq \lim_{x \rightarrow p} \frac{\|\tilde{L}(\varphi(x))\|}{\|x - p\|} \leq \|\tilde{L}\| \cdot \underbrace{\lim_{x \rightarrow p} \frac{\|\varphi(x)\|}{\|x - p\|}}_{=0 \text{ nach (8.8)}} = 0.$$

2) $\lim_{x \rightarrow p} \frac{\|\psi(f(x))\|}{\|x - p\|} = 0$:

Nach (8.8) gilt $\lim_{y \rightarrow q} \frac{\|\psi(y)\|}{\|y - q\|} = 0$. Folglich existiert für jedes $\varepsilon > 0$ ein $\delta > 0$, so dass

$$\|\psi(f(x))\| < \varepsilon \cdot \|f(x) - q\| \quad \text{für alle } x \in U \text{ mit } \|f(x) - q\| < \delta.$$

Da f in p stetig ist, existiert ein $\delta_1 > 0$, so dass

$$\|f(x) - q\| < \delta \quad \text{für alle } x \in U \text{ mit } \|x - p\| < \delta_1.$$

Für $x \in U$ mit $\|x - p\| < \delta_1$ gilt also

$$\begin{aligned} \|\psi(f(x))\| &< \varepsilon \cdot \|f(x) - q\| = \varepsilon \cdot \|\varphi(x) + L(x - p)\| \\ &< \varepsilon \cdot \|\varphi(x)\| + \varepsilon \cdot \|L\| \cdot \|x - p\| \end{aligned}$$

und damit

$$\frac{\|\psi(f(x))\|}{\|x - p\|} < \varepsilon \cdot \underbrace{\frac{\|\varphi(x)\|}{\|x - p\|}}_{\rightarrow 0 \text{ für } x \rightarrow p} + \varepsilon \cdot \|L\| \quad \forall |x - p| < \delta_1.$$

Wir erhalten daraus

$$0 \leq \limsup_{x \rightarrow p} \frac{\|\psi(f(x))\|}{\|x - p\|} \leq \varepsilon \cdot \|L\| \quad \forall \varepsilon > 0$$

und somit

$$\lim_{x \rightarrow p} \frac{\|\psi(f(x))\|}{\|x - p\|} = 0.$$

Damit ist die Kettenregel bewiesen.

Ist $k = 1$, so ist $g \circ f : U \subset \mathbb{R} \rightarrow \mathbb{R}^n$ eine Funktion, die nur von *einer* reellen Variablen abhängt. Dann erhalten wir die folgende spezielle Form der Kettenregel für die Ableitung in p :

$$(g \circ f)'(p) = (g \circ f)'(p) \cdot 1 = d(g \circ f)_p(1) = dg_{f(p)}(df_p(1)) = dg_{f(p)}(f'(p)). \quad \square$$

Satz 8.7 (Kettenregel für partielle Ableitungen)

Seien $g_1, \dots, g_n : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ reellwertige Funktionen, die in $p \in U$ differenzierbar sind. Sei $V \subset \mathbb{R}^n$ eine offene Menge mit $(g_1, \dots, g_n)(U) \subset V$ und $f : V \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine in $(g_1(p), \dots, g_n(p))$ differenzierbare Abbildung. Dann ist die Abbildung $F : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$,

$$F(x) := f(g_1(x), \dots, g_n(x)),$$

in $p \in U$ differenzierbar und für ihre partiellen Ableitungen gilt

$$\frac{\partial F}{\partial x_i}(p) = \sum_{j=1}^n \frac{\partial f}{\partial y_j}(g_1(p), \dots, g_n(p)) \cdot \frac{\partial g_j}{\partial x_i}(p), \quad i = 1, \dots, k.$$

Ist $k = 1$, so gilt insbesondere:

$$F'(p) = \sum_{j=1}^n \frac{\partial f}{\partial y_j}(g_1(p), \dots, g_n(p)) \cdot g'_j(p).$$

Beweis. Wir betrachten $g : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^n$, definiert durch $g(x) := (g_1(x), \dots, g_n(x))$. Dann gilt $F = f \circ g$ und nach Kettenregel für die Differentiale $dF_p = df_{g(p)} \circ dg_p$. Für die partiellen Ableitungen folgt daraus

$$\begin{aligned} \frac{\partial F}{\partial x_i}(p) &= dF_p(e_i) = df_{g(p)}(dg_p(e_i)) = df_{g(p)}\left(\sum_{j=1}^n (dg_j)_p(e_i) \cdot e_j\right) \\ &= df_{g(p)}\left(\sum_{j=1}^n \frac{\partial g_j}{\partial x_i}(p) \cdot e_j\right) = \sum_{j=1}^n \frac{\partial g_j}{\partial x_i}(p) \cdot df_{g(p)}(e_j) \\ &= \sum_{j=1}^n \frac{\partial g_j}{\partial x_i}(p) \cdot \frac{\partial f}{\partial y_j}(g(p)). \end{aligned} \quad \square$$

Mit Hilfe der Kettenregeln kann man den Mittelwertsatz für reelle Funktionen verallgemeinern.

Satz 8.8 (Mittelwertsätze für Funktionen mehrerer reeller Variablen)

1. Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine differenzierbare reellwertige Funktion und $x, y \in U$ zwei Punkte in U , für die die gesamte Strecke \overline{xy} zwischen x und y ebenfalls in U liegt. Dann existiert ein $\xi \in \overline{xy}$ mit $\xi \neq x$, $\xi \neq y$, so dass

$$f(y) - f(x) = df_{\xi}(y - x).$$

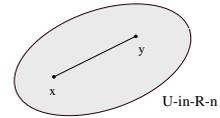
2. Sei $f : [a, b] \rightarrow \mathbb{R}^n$ stetig und auf (a, b) differenzierbar, dann existiert ein $\eta \in (a, b)$, so dass

$$\|f(b) - f(a)\| \leq \|f'(\eta)\| \cdot (b - a).$$

Beweis. Zu 1) Wir betrachten die reelle Funktion $\varphi : [0, 1] \rightarrow \mathbb{R}$ definiert durch $\varphi(t) := f(x + t(y - x))$.

Nach Voraussetzung ist φ auf $[0, 1]$ differenzierbar. Daher existiert nach dem Mittelwertsatz von Lagrange ein $\theta \in (0, 1)$ mit

$$\varphi(1) - \varphi(0) = \varphi'(\theta) \cdot (1 - 0) = \varphi'(\theta).$$



Nach Definition von h und aufgrund der Kettenregel ist dies äquivalent zu

$$f(y) - f(x) = h(1) - h(0) = h'(\theta) = \underbrace{df_{x + \theta(y - x)}}_{=: \xi}(y - x) = df(\xi)(y - x).$$

Zu 2) Wir betrachten die reelle Funktion $\phi : [a, b] \rightarrow \mathbb{R}$, definiert durch das Skalarprodukt

$$\phi(t) := \langle f(b) - f(a), f(t) \rangle.$$

ϕ ist auf $[a, b]$ stetig und auf (a, b) differenzierbar. Nach dem Mittelwertsatz von Lagrange existiert ein $\eta \in (a, b)$ mit

$$\phi(b) - \phi(a) = (b - a) \cdot \phi'(\eta) = (b - a) \cdot \langle f(b) - f(a), f'(\eta) \rangle.$$

Aus der Cauchy-Schwarzschen Ungleichung folgt dann

$$|\phi(b) - \phi(a)| \leq (b - a) \cdot |\langle f(b) - f(a), f'(\eta) \rangle| \leq (b - a) \cdot \|f(b) - f(a)\| \cdot \|f'(\eta)\|.$$

Andererseits ist

$$\phi(b) - \phi(a) = \langle f(b) - f(a), f(b) - f(a) \rangle = \|f(b) - f(a)\|^2.$$

Daraus folgt die Behauptung

$$\|f(b) - f(a)\| \leq (b - a) \|f'(\eta)\|.$$

□

Wir betrachten zwei Anwendungen der Mittelwertsätze.

Satz 8.9 Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ differenzierbar und U offen und wegzusammenhängend. Dann ist die Abbildung f genau dann konstant, wenn $df_p = 0$ für alle $p \in U$.

Beweis. Wegen Satz 8.1 3) genügt es, die Behauptung für jede Komponente von f , d.h. für reellwertige Funktionen zu beweisen. Sei also oBdA $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$.

(\implies) Wenn die Abbildung f konstant ist, so gilt für die Richtungsableitung in jedem Punkt $p \in U$ und in jeder Richtung $\mathbf{a} \in \mathbb{R}^k$

$$df_p(\mathbf{a}) = \nabla_{\mathbf{a}} f(p) = \lim_{t \rightarrow 0} \frac{f(p + t\mathbf{a}) - f(p)}{t} = \lim_{t \rightarrow 0} \frac{0}{t} = 0.$$

Somit ist $df_p = 0$ für alle $p \in U$.

(\impliedby) Sei $df_p = 0$ für alle $p \in U$ und sei $x_0 \in U$ ein fixierter Punkt. Da U offen ist, existiert ein $\varepsilon(x_0) > 0$, so dass $K(x_0, \varepsilon(x_0)) \subset U$. Die Kugel ist konvex, folglich liegt für jedes $z \in K(x_0, \varepsilon(x_0))$ mit $z \neq x_0$ die Strecke $\overline{x_0 z}$ ebenfalls in $K(x_0, \varepsilon(x_0))$. Nach Satz 8.8 existiert ein $\xi \in \overline{x_0 z}$ mit

$$f(z) - f(x_0) = df_{\xi}(z - x_0) \stackrel{Vor.}{=} 0.$$

Folglich ist f auf der Kugel $K(x_0, \varepsilon(x_0))$ konstant mit dem Wert $f(x_0)$.

Seien nun $p, q \in U$ beliebige, aber fixierte Punkte. Da U wegzusammenhängend ist, existiert eine stetige Abbildung $\sigma : [0, 1] \rightarrow U \subset \mathbb{R}^k$ mit $\sigma(0) = p$ und $\sigma(1) = q$. Die Menge der Kugeln $\{K(\sigma(t), \varepsilon(\sigma(t)))\}_{t \in [0,1]}$ überdeckt $\sigma([0, 1])$. Da $[0, 1]$ kompakt und σ stetig ist, ist auch $\sigma([0, 1])$ kompakt.

Folglich findet man endlich viele dieser Kugeln $K(\sigma(t_1), \varepsilon_1), \dots, K(\sigma(t_r), \varepsilon_r)$, die die Menge $\sigma([0, 1])$ bereits überdecken (Übungsaufgabe). Auf jeder dieser Kugeln gilt aber

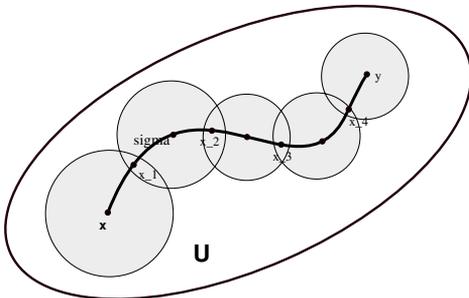
$$f|_{K(\sigma(t_j), \varepsilon_j)} = \text{konst.} = f(\sigma(t_j)).$$

Daraus folgt

$$f(p) = f(\sigma(t_1)) = f(p_1) = f(\sigma(t_2)) = f(p_2) = \dots = f(\sigma(t_r)) = f(q),$$

mit $p_i \in K(\sigma(t_i), \varepsilon_i) \cap K(\sigma(t_{i+1}), \varepsilon_{i+1})$. Folglich ist f auf U konstant. □

Man kann zeigen, dass offene und zusammenhängende Teilmengen $U \subset \mathbb{R}^k$ auch wegzusammenhängend sind. Satz 8.9 gilt deshalb auch für offene, zusammenhängende Mengen $U \subset \mathbb{R}^k$.



Satz 8.10 Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine differenzierbare Abbildung und U eine offene und konvexe Menge. Es existiere eine Konstante $C \in \mathbb{R}^+$, so dass

$$\|df_p\| \leq C \quad \forall p \in U,$$

wobei $\|df_p\|$ die Operatornorm von df_p bezeichnet. Dann gilt

$$\|f(y) - f(x)\| \leq C \cdot \|y - x\| \quad \forall x, y \in U.$$

Insbesondere ist f auf U Lipschitzstetig.

Beweis. Seien $x, y \in U$ fixiert. Wir parametrisieren die Strecke von x nach y durch $\gamma : [0, 1] \rightarrow U$ mit $\gamma(t) := x + t(y - x)$ und betrachten die Abbildung $g : [0, 1] \rightarrow \mathbb{R}^m$ mit $g(t) := f(\gamma(t))$. Dann gilt nach Kettenregel

$$g'(t) = df_{x+t(y-x)}(y - x)$$

und daher

$$\|g'(t)\| \leq \|df_{x+t(y-x)}\| \cdot \|y - x\| \leq C \cdot \|y - x\| \quad \forall t \in [0, 1].$$

Aus dem Mittelwertsatz (Satz 8.8, 2) folgt die Existenz einer Zahl $\eta \in (0, 1)$ mit

$$\|f(y) - f(x)\| = \|g(1) - g(0)\| \leq \|g'(\eta)\| \leq C \cdot \|y - x\|.$$

□

8.3 Die Jacobi-Matrix, der Gradient und die Hesse-Matrix

Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine in $p \in U$ differenzierbare Abbildung und $df_p : \mathbb{R}^k \rightarrow \mathbb{R}^m$ ihr Differential in p . Aus der Vorlesung *Lineare Algebra* ist bekannt, dass man lineare Abbildungen durch Matrizen beschreiben kann, wenn man im Urbild- und im Bildvektorraum jeweils eine Basis fixiert. Wir können das Differential df_p also auch durch eine Matrix angeben.

Wir fixieren dazu in den Vektorräumen \mathbb{R}^k und \mathbb{R}^m jeweils die kanonischen Basen (e_1, \dots, e_k) bzw. $(\hat{e}_1, \dots, \hat{e}_m)$. Dann entspricht jeder linearen Abbildung $A : \mathbb{R}^k \rightarrow \mathbb{R}^m$ die $(m \times k)$ -Matrix $[A] := (A_{ij})$, definiert durch

$$A(e_j) =: \sum_{i=1}^m A_{ij} \hat{e}_i \quad j = 1, \dots, k.$$

In der j -ten Spalte der Matrix $[A]$ stehen die Komponenten des Vektors $A(e_j) \in \mathbb{R}^m$ bzgl. der kanonischen Basis $(\hat{e}_1, \dots, \hat{e}_m)$ im \mathbb{R}^m . Der Anwendung der linearen Abbildung A auf den Vektor $x = \sum_{j=1}^k x_j e_j = (x_1, \dots, x_k) \in \mathbb{R}^k$ entspricht dabei die Anwendung der Matrix $[A] = (A_{ij})$ auf den Spaltenvektor $(x_1, \dots, x_k)^t$:

$$\begin{array}{ccc} \mathbb{R}^k & \xrightarrow{A} & \mathbb{R}^m \\ \text{transponieren} = t \downarrow & & \downarrow t = \text{transponieren} \\ \mathbb{R}^k & \xrightarrow{[A]} & \mathbb{R}^m \end{array} .$$

Sei $f = (f_1, \dots, f_m)$ die Komponentendarstellung von f . Dann gilt für das Differential $df_p : \mathbb{R}^k \rightarrow \mathbb{R}^m$:

$$df_p(e_j) = \frac{\partial f}{\partial x_j}(p) = \sum_{i=1}^m \frac{\partial f_i}{\partial x_j}(p) \cdot \hat{e}_i \quad j = 1, \dots, k.$$

Die zu df_p gehörende Matrix $[df_p]$ hat also die Form

$$[df_p] = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \frac{\partial f_1}{\partial x_2}(p) & \cdots & \frac{\partial f_1}{\partial x_k}(p) \\ \frac{\partial f_2}{\partial x_1}(p) & \frac{\partial f_2}{\partial x_2}(p) & \cdots & \frac{\partial f_2}{\partial x_k}(p) \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \frac{\partial f_m}{\partial x_2}(p) & \cdots & \frac{\partial f_m}{\partial x_k}(p) \end{pmatrix} = \begin{pmatrix} \frac{\partial f^t}{\partial x_1}(p) & \frac{\partial f^t}{\partial x_2}(p) & \cdots & \frac{\partial f^t}{\partial x_k}(p) \end{pmatrix}.$$

Definition 8.5. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine Abbildung, die in $p \in U$ partiell differenzierbar ist und seien $f = (f_1, \dots, f_m)$ ihre Komponenten. Dann nennt man die Matrix

$$Jf_p := \left(\frac{\partial f_i}{\partial x_j}(p) \right)_{\substack{i=1, \dots, m \text{ Zeilen} \\ j=1, \dots, k \text{ Spalten}}}$$

Jacobi-Matrix von f im Punkt p .

Ist $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ in $p \in U$ differenzierbar, so gilt also $[df_p] = Jf_p$.

Eine besondere Bedeutung haben die Vektoren, die in den Zeilen der Jacobi-Matrix auftreten.

Definition 8.6. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ eine reellwertige Funktion, die im Punkt $p \in U$ partiell differenzierbar ist. Dann heißt der Vektor

$$\text{grad} f(p) := \left(\frac{\partial f}{\partial x_1}(p), \frac{\partial f}{\partial x_2}(p), \dots, \frac{\partial f}{\partial x_k}(p) \right) \in \mathbb{R}^k$$

Gradient von f im Punkt $p \in U$.

Für reellwertige partiell differenzierbare Funktionen $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ stimmt die Jacobi-Matrix also mit dem Gradienten überein:

$$Jf_p = \text{grad} f(p) = \left(\frac{\partial f}{\partial x_1}(p), \dots, \frac{\partial f}{\partial x_k}(p) \right).$$

Die Jacobi-Matrix einer partiell differenzierbaren Abbildung $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ mit den Komponenten $f = (f_1, \dots, f_m)$ schreibt sich durch die Gradienten der Komponentenfunktionen in der Form:

$$Jf_p = \begin{pmatrix} \text{grad} f_1(p) \\ \text{grad} f_2(p) \\ \vdots \\ \text{grad} f_m(p) \end{pmatrix}.$$

Der folgende Satz erklärt die geometrische Bedeutung des Gradienten einer Funktion. Ist $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ in p differenzierbar, so gilt für alle $\mathbf{a} = (a_1, \dots, a_k) \in \mathbb{R}^k$ mit dem Euklidischen Skalarprodukt $\langle \cdot, \cdot \rangle$

$$df_p(\mathbf{a}) = \sum_{j=1}^k \frac{\partial f}{\partial x_j}(p) \cdot a_j = \langle \text{grad} f(p), \mathbf{a} \rangle. \quad (8.9)$$

Wir erhalten daraus folgende geometrische Interpretation von $\text{grad} f(p)$:

Satz 8.11 Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ in $p \in U$ differenzierbar.

1. $\text{grad} f(p) = 0$ genau dann, wenn $df_p = 0$.
2. Ist $\text{grad} f(p) \neq 0$, so gibt $\text{grad} f(p) \in \mathbb{R}^k$ diejenige Richtung an, in der die Funktion f im Punkt p am schnellsten wächst, d.h. diejenige Richtung mit dem größten Funktionsanstieg.
3. Sei $c \in \mathbb{R}$ im Bild von f und bezeichne $M_c := \{x \in U \mid f(x) = c\}$ die Niveaufläche von f zum Niveau c . Ist $p \in M_c$ ein Punkt auf dieser Niveaufläche und $\gamma : I \subset \mathbb{R} \rightarrow M_c$ eine differenzierbare Kurve mit $\gamma(t_0) = p$ und Spur in M_c , dann gilt:

$$\gamma'(t_0) \perp \text{grad} f(p),$$

d.h. der Vektor $\text{grad} f(p)$ steht senkrecht auf allen Tangenten an die Niveaufläche M_c im Punkt p .

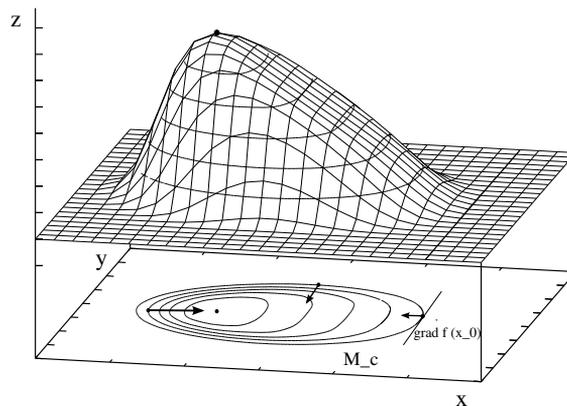
Beispiel: Höhenlinien

Wir betrachten zur Illustration von Satz 8.11 die Abbildung, die die Höhe eines Ortes über dem Meeresspiegel angibt:

$$f : U \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$(x, y) \longmapsto$ Höhe des Ortes über dem Meeresspiegel.

Dann sieht man die Niveauflächen von f als Höhenlinien auf der Landkarte.



Beweis von Satz 8.11 Zu 1) Das folgt unmittelbar aus (8.9), da das Euklidische Skalarprodukt nicht ausgeartet ist.

Zu 2) Sei $\mathbf{a} \in \mathbb{R}^n$ mit $\|\mathbf{a}\| = 1$. Dann folgt aus der Cauchy-Schwarzschen Ungleichung

$$\nabla_{\mathbf{a}} f(p) = \langle \text{grad} f(p), \mathbf{a} \rangle \stackrel{CSU}{\leq} \|\text{grad} f(p)\| \cdot \|\mathbf{a}\| = \|\text{grad} f(p)\|. \quad (*)$$

Der maximal mögliche Wert von $\nabla_{\mathbf{a}} f(p)$, wenn man die Vektoren \mathbf{a} mit $\|\mathbf{a}\| = 1$ durchläuft, kann also höchstens gleich $\|\text{grad} f(p)\|$ sein. In der Cauchy-Schwarzschen Ungleichung (*) wird aber die Gleichheit angenommen, und zwar genau dann, wenn die auftretenden Vektoren linear abhängig sind und in die gleiche Richtung zeigen, d.h. genau dann, wenn $\mathbf{a} = \frac{\text{grad} f(p)}{\|\text{grad} f(p)\|}$. Die Richtung mit dem größten Funktionsanstieg ist also diejenige, in die der Vektor $\text{grad} f(p)$ zeigt.

Zu 3) Sei $\gamma : I \rightarrow M_c$ eine differenzierbare Kurve mit $\gamma(t_0) = p$ und Spur in M_c . Dann gilt $f(\gamma(t)) = c$ für alle $t \in I$. Nach der Kettenregel ist dann

$$0 = (f \circ \gamma)'(t_0) = df_p(\gamma'(t_0)) = \langle \text{grad} f(p), \gamma'(t_0) \rangle.$$

Also steht der Gradient $\text{grad} f(p)$ senkrecht auf dem Tangentialvektor $\gamma'(t_0)$. □

Wir definieren als nächstes die partiellen Ableitungen höherer Ordnung.

Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine Funktion, für die die partielle Ableitung $\frac{\partial f}{\partial x_i} : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ existiert. Wenn die partielle Ableitung der Funktion $\frac{\partial f}{\partial x_i} : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ nach der j -ten Koordinate existiert, so bezeichnet man sie mit

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

Auf diese Weise entstehen partielle Ableitungen höherer Ordnung. Die Funktion

$$\frac{\partial^n f}{\partial x_{i_n} \cdots \partial x_{i_1}} := \frac{\partial}{\partial x_{i_n}} \left(\frac{\partial^{n-1} f}{\partial x_{i_{n-1}} \cdots \partial x_{i_1}} \right) : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$$

heißt (falls sie existiert) n -te partielle Ableitung von f nach den Variablen x_{i_1}, \dots, x_{i_n} . Leitet man n -mal nach der gleichen Variablen ab, so schreibt man zur Abkürzung

$$\frac{\partial^n f}{\partial x_i \cdots \partial x_i} =: \frac{\partial^n f}{\partial x_i^n}.$$

Definition 8.7. Eine Funktion $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ heißt n -mal stetig differenzierbar (C^n -Funktion oder von der Klasse C^n) mit $n \in \mathbb{N}$, wenn alle partiellen Ableitungen von f der Ordnung $\leq n$ existieren und stetig sind.

Bezeichnung: $C^n(U, \mathbb{R}^m) := \{f : U \rightarrow \mathbb{R}^m \mid f \text{ } C^n\text{-Funktion}\}$.

Offensichtlich gilt für $n \in \mathbb{N}$

$$C^n(U, \mathbb{R}^m) \subset C^{n-1}(U, \mathbb{R}^m) \subset C^{n-2}(U, \mathbb{R}^m) \subset \dots \subset C^0(U, \mathbb{R}^m).$$

Beispiel: Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ die Funktion $f(x, y) := x^3y^2$.

Dann erhalten wir für die partiellen Ableitungen erster und zweiter Ordnung

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= 3x^2y^2, & \frac{\partial f}{\partial y}(x, y) &= 2x^3y \\ \frac{\partial^2 f}{\partial x^2}(x, y) &= 6xy^2, & \frac{\partial^2 f}{\partial y^2}(x, y) &= 2x^3, & \frac{\partial^2 f}{\partial y \partial x}(x, y) &= 6x^2y = \frac{\partial f}{\partial x \partial y}(x, y). \end{aligned}$$

f ist also 2-fach stetig differenzierbar. In diesem Beispiel sind die gemischten 2. partiellen Ableitungen unabhängig von der Reihenfolge der Differentiation.

Im Allgemeinen hängen die partiellen Ableitungen von der Reihenfolge der Variablen ab, nach denen abgeleitet wird (siehe Übungsaufgabe 29). Der folgende Satz gibt eine hinreichende Bedingung dafür an, dass man die Reihenfolge der partiellen Ableitungen vertauschen kann.

Satz 8.12 (Lemma von Schwarz)

Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine n -mal stetig differenzierbare Abbildung. Dann sind alle partiellen Ableitungen der Ordnung $\leq n$ unabhängig von der Reihenfolge des Differenzierens. Insbesondere gilt für $f \in C^2(U, \mathbb{R}^m)$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Beweis. Es genügt wieder, die Behauptung für reellwertige Funktionen zu zeigen. Außerdem genügt es, die Behauptung für die 2. partiellen Ableitungen zu zeigen. Die Behauptung für die partiellen Ableitungen höherer Ordnung beweist man dann durch Induktion. Bei den 2. partiellen Ableitungen läßt man bis auf zwei Variablen alle anderen fest. Es genügt also für jede C^2 -Funktion $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}$$

zu zeigen. Das werden wir jetzt tun. Sei $p = (p_1, p_2) \in U$ und seien $h = (h_1, h_2) \in \mathbb{R}^2$ so gewählt, dass $h_1, h_2 \neq 0$ und $h \in K(p, \varepsilon) \subset U$.

Wir betrachten die Funktion $\varphi(x) := f(x, p_2 + h_2) - f(x, p_2)$. Dann ist φ auf dem Intervall $[p_1 - |h_1|, p_1 + |h_1|]$ differenzierbar. Nach dem Mittelwertsatz existiert ein ξ_1 zwischen p_1 und $p_1 + h_1$, so dass gilt

$$\varphi(p_1 + h_1) - \varphi(p_1) = h_1 \cdot \varphi'(\xi_1).$$

Sei nun

$$F(h_1, h_2) := f(p_1 + h_1, p_2 + h_2) - f(p_1 + h_1, p_2) - f(p_1, p_2 + h_2) + f(p_1, p_2).$$

Dann gilt

$$\begin{aligned} F(h_1, h_2) &= \varphi(p_1 + h_1) - \varphi(p_1) = h_1 \cdot \varphi'(\xi_1) \\ &= h_1 \left(\frac{\partial f}{\partial x_1}(\xi_1, p_2 + h_2) - \frac{\partial f}{\partial x_1}(\xi_1, p_2) \right). \end{aligned}$$

Da die Funktion $\frac{\partial f}{\partial x_1}$ auf U nach der 2. Variable differenzierbar ist, können wir den Mittelwertsatz auch auf die Funktion $\frac{\partial f}{\partial x_1}(\xi_1, \cdot)$ anwenden. Es existiert also ein ξ_2 zwischen p_2 und $p_2 + h_2$, so dass

$$F(h_1, h_2) = h_1 \cdot h_2 \cdot \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_1, \xi_2). \quad (\star)$$

Verfährt man analog mit der Funktion $\psi(x) := f(p_1 + h_1, x) - f(p_1, x)$, so existieren nach dem Mittelwertsatz ein $\tilde{\xi}_1$ zwischen p_1 und $p_1 + h_1$ und ein $\tilde{\xi}_2$ zwischen p_2 und $p_2 + h_2$, so dass

$$F(h_1, h_2) = h_1 \cdot h_2 \cdot \frac{\partial^2 f}{\partial x_1 \partial x_2}(\tilde{\xi}_1, \tilde{\xi}_2). \quad (\star\star)$$

Da $h_1, h_2 \neq 0$, folgt aus (\star) und $(\star\star)$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(\tilde{\xi}_1, \tilde{\xi}_2) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_1, \xi_2).$$

Bei $h = (h_1, h_2) \rightarrow (0, 0)$ konvergieren sowohl (ξ_1, ξ_2) als auch $(\tilde{\xi}_1, \tilde{\xi}_2)$ gegen $p = (p_1, p_2)$. Da die partiellen Ableitungen $\frac{\partial^2 f}{\partial x_1 \partial x_2}$ und $\frac{\partial^2 f}{\partial x_2 \partial x_1}$ in (p_1, p_2) nach Voraussetzung stetig sind, folgt mit $h = (h_1, h_2) \rightarrow (0, 0)$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(p) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(p).$$

□

Definition 8.8. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ eine 2-mal stetig differenzierbare reellwertige Funktion. Die symmetrische $(k \times k)$ -Matrix

$$\text{Hess}f(p) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(p) \right)_{i,j=1,\dots,k}$$

heißt Hesse-Matrix von f im Punkt $p \in U$.

8.4 Die Taylorformel für Funktionen mehrerer reeller Variablen

Funktionen einer reellen Variablen $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$, die auf einem Intervall I definiert sind, haben wir durch ihre Taylorpolynome approximiert. Ist f $(n+1)$ -mal stetig differenzierbar und $p \in I$, so gilt auf I die Taylorformel

$$f(x) = \underbrace{f(p) + f'(p)(x-p) + \frac{f''(p)}{2!}(x-p)^2 + \dots + \frac{f^{(n)}(p)}{n!}(x-p)^n}_{=: T_n(f,p)(x) \text{ } n\text{-tes Taylorpolynom}} + \underbrace{R_n(f,p)(x)}_{n\text{-tes Restglied}$$

wobei das n -te Restglied $R_n(f,p)(x)$ z.B. in der Lagrange-Form

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

für ein geeignetes ξ zwischen p und x zu beschreiben ist.

Man kann auch Funktionen mehrerer reeller Variablen durch Taylorpolynome approximieren. Wie im Fall einer reellen Variablen wird sich dies als wichtiges Hilfsmittel bei der Untersuchung von Extremwerten erweisen. Um die Taylorformel für Funktionen mehrerer Variablen beschreiben zu können, führen wir zunächst einige Bezeichnungen ein.

Definition 8.9. Ein Multiindex der Länge k ist ein k -Tupel von natürlichen Zahlen

$$\alpha = (\alpha_1, \dots, \alpha_k), \quad \alpha_j \in \mathbb{N}_0.$$

Die Zahl $|\alpha| := \alpha_1 + \dots + \alpha_k$ heißt die Ordnung des Multiindex α und die Zahl

$\alpha! := (\alpha_1!) \cdot \dots \cdot (\alpha_k!)$ die Fakultät von α . Für $y = (y_1, \dots, y_k) \in \mathbb{R}^k$ sei

$$y^\alpha := y_1^{\alpha_1} \cdot \dots \cdot y_k^{\alpha_k}.$$

Für eine reellwertige Funktion $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ bezeichne (sofern existent)

$$\frac{\partial^{|\alpha|} f}{\partial x^\alpha} := \frac{\partial^{|\alpha|} f}{(\partial x_1)^{\alpha_1} \dots (\partial x_k)^{\alpha_k}}.$$

Definition 8.10. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ eine n -mal stetig differenzierbare Funktion. Das Taylorpolynom n -ter Ordnung von f im Entwicklungspunkt $p \in U$ ist das Polynom n -ten Grades in den Unbestimmten x_1, \dots, x_k gegeben durch

$$\begin{aligned} T_n(f,p)(x) &:= \sum_{|\alpha| \leq n} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(p) \cdot (x-p)^\alpha \quad (\alpha \text{ sind hier Multiindizes der Länge } k) \\ &= f(p) + \sum_{|\alpha|=1} \frac{1}{\alpha!} \frac{\partial f}{\partial x^\alpha}(p) \cdot (x-p)^\alpha + \sum_{|\alpha|=2} \frac{1}{\alpha!} \frac{\partial^2 f}{\partial x^\alpha}(p) \cdot (x-p)^\alpha + \dots + \\ &\quad + \sum_{|\alpha|=n} \frac{1}{\alpha!} \frac{\partial^n f}{\partial x^\alpha}(p) \cdot (x-p)^\alpha. \end{aligned}$$

Es gilt:

$$\begin{aligned} |\alpha| = 1 &\iff \alpha = (0, \dots, 0, \overset{i}{1}, 0, \dots, 0) \quad \text{für } i = 1, \dots, k \\ |\alpha| = 2 &\iff \alpha = (0, \dots, 0, \overset{i}{2}, 0, \dots, 0) \quad \text{für } 1 \leq i \leq k \text{ oder} \\ &\quad \alpha = (0, \dots, 0, \overset{i}{1}, 0, \dots, 0, \overset{j}{1}, 0, \dots, 0) \quad \text{für } 1 \leq i < j \leq k \end{aligned}$$

Daraus folgt für das erste und zweite Taylorpolynom:

$$\begin{aligned} T_1(f, p)(x) &= f(p) + \sum_{i=1}^k \frac{\partial f}{\partial x_i}(p) \cdot (x_i - p_i) \\ &= f(p) + \langle \text{grad} f(p), x - p \rangle. \end{aligned} \tag{8.10}$$

$$\begin{aligned} T_2(f, p)(x) &= T_1(f, p)(x) + \frac{1}{2} \sum_{i,j=1}^k \frac{\partial^2 f}{\partial x_i \partial x_j}(p) \cdot (x_i - p_i) \cdot (x_j - p_j) \\ &= f(p) + \langle \text{grad} f(p), x - p \rangle + \frac{1}{2} (x - p) \cdot \text{Hess} f(p) \cdot (x - p)^t. \end{aligned} \tag{8.11}$$

Der Graph des ersten Taylorpolynoms beschreibt die Tangentialebene an die Hyperfläche $F := \{(x, f(x) \mid x \in U\} \subset \mathbb{R}^{k+1}$ im Punkt $P = (p, f(p))$ (siehe Abschnitt 8.1). Der Graph des zweiten Taylorpolynoms ist die sogenannte *Schmieghquadratik an F im Punkt P*. Die Normalformen solcher Quadriken und ihre geometrische Gestalt lernen Sie in der Vorlesung *Lineare Algebra 2* kennen.

Satz 8.13 (Taylorformel n-ter Ordnung)

Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ eine $(n + 1)$ -mal stetig differenzierbare reellwertige Funktion und $p, x \in U$ zwei Punkte, für die die Strecke \overline{px} in U liegt. Dann existiert ein $\xi \in \overline{px}$ mit $\xi \neq p, \xi \neq x$, so dass

$$f(x) = T_n(f, p)(x) + \underbrace{\sum_{|\alpha|=n+1} \frac{1}{\alpha!} \frac{\partial^{n+1} f}{\partial x^\alpha}(\xi) \cdot (x - p)^\alpha}_{=: R_n(f, p)(x) \text{ n-tes Restglied}}$$

Beweis. Wir parametrisieren die Strecke \overline{px} durch $\sigma : [0, 1] \rightarrow U$ mit $\sigma(t) := p + t(x - p)$ und betrachten die Funktion $g : [0, 1] \rightarrow \mathbb{R}$ mit $g(t) := f(\sigma(t)) = f(p + t(x - p))$. Dann ist $g \in C^{n+1}([0, 1], \mathbb{R})$. Auf g wenden wir die Taylorformel für Funktionen einer reellen Variablen an und erhalten: Es existiert ein $\theta \in (0, 1)$, so dass

$$g(1) = \sum_{j=0}^n \frac{1}{j!} g^{(j)}(0) + \underbrace{\frac{1}{(n + 1)!} g^{(n+1)}(\theta)}_{\text{Lagrange-Restglied}}$$

Mit der Kettenregel für partielle Ableitungen folgt:

$$\begin{aligned} g^{(0)}(t) &= g(t) = f(\sigma(t)), \\ g^{(1)}(t) &= \sum_{i=1}^k \frac{\partial f}{\partial x_i}(\sigma(t)) \cdot (x_i - p_i) = \sum_{|\alpha|=1} \frac{\partial^\alpha f}{\partial x^\alpha}(\sigma(t)) \cdot (x - p)^\alpha, \\ g^{(2)}(t) &= \sum_{i,j=1}^k \frac{\partial^2 f}{\partial x_i \partial x_j}(\sigma(t)) \cdot (x_i - p_i) \cdot (x_j - p_j) \tag{*} \\ &= \sum_{|\alpha|=2} \frac{2!}{\alpha!} \frac{\partial^\alpha f}{\partial x^\alpha}(\sigma(t)) \cdot (x - p)^\alpha. \tag{**} \end{aligned}$$

Der Faktor $\frac{2!}{\alpha!}$ in (**) ist notwendig, da in (*) $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ zweimal auftritt, während $\frac{\partial^2 f}{\partial x_i^2}$ in (**) nur einmal vorhanden ist.

Analog beweist man mit der Kettenregel und etwas Kombinatorik mittels Induktion für die j -te Ableitung

$$g^{(j)}(t) = \sum_{|\alpha|=j} \frac{j!}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(\sigma(t)) \cdot (x - p)^\alpha.$$

Daraus folgt mit $\xi := p + \theta(x - p) = \sigma(\theta)$ die Taylorformel

$$f(x) = \sum_{|\alpha| \leq n} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(p) \cdot (x - p)^\alpha + \sum_{|\alpha|=n+1} \frac{1}{\alpha!} \frac{\partial^{n+1} f}{\partial x^\alpha}(\xi) \cdot (x - p)^\alpha.$$

□

8.5 Lokale Extrema für Funktionen mehrerer reeller Variablen

Für 2-mal stetig differenzierbare Funktionen $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ kennen wir Kriterien für das Vorliegen lokaler Extremwerte. Wir wollen dies jetzt auf den Fall von Funktionen mehrerer Variablen verallgemeinern.

Dazu erinnern wir zunächst an einige Kenntnisse aus der Vorlesung *Lineare Algebra*.

Definition 8.11. Sei $A = (A_{ij})$ eine symmetrische reelle $(n \times n)$ -Matrix. A heißt
 positiv definit (symbolisch $A > 0$), falls $xAx^t > 0 \quad \forall x = (x_1, \dots, x_n) \in \mathbb{R}^n, x \neq 0$,
 negativ definit (symbolisch $A < 0$), falls $xAx^t < 0 \quad \forall x = (x_1, \dots, x_n) \in \mathbb{R}^n, x \neq 0$,
 positiv semidefinit (symbolisch $A \geq 0$), falls $xAx^t \geq 0 \quad \forall x = (x_1, \dots, x_n) \in \mathbb{R}^n$,
 negativ semidefinit (symbolisch $A \leq 0$), falls $xAx^t \leq 0 \quad \forall x = (x_1, \dots, x_n) \in \mathbb{R}^n$,
 indefinit, falls A weder positiv semidefinit noch negativ semidefinit ist.

Algebraische Fakten:

1. Jede symmetrische reelle $(n \times n)$ -Matrix ist diagonalisierbar, d.h. sie hat n reelle Eigenwerte.
2. *Eigenwert-Kriterium:* Für eine symmetrische reelle $(n \times n)$ -Matrix A gilt:

$$A > 0 \iff \text{alle Eigenwerte von } A \text{ sind positiv.}$$

$$A < 0 \iff \text{alle Eigenwerte von } A \text{ sind negativ.}$$

$$A \geq 0 \iff \text{alle Eigenwerte von } A \text{ sind } \geq 0.$$

$$A \leq 0 \iff \text{alle Eigenwerte von } A \text{ sind } \leq 0.$$

3. *Determinanten-Kriterium:* Für eine symmetrische reelle $(n \times n)$ -Matrix (A_{ij}) bezeichne

$$A[k] := \begin{pmatrix} A_{11} & \dots & A_{1k} \\ \vdots & \dots & \vdots \\ A_{k1} & \dots & A_{kk} \end{pmatrix}$$

ihren k -ten Hauptminor. Dann gilt:

$$\begin{aligned} A > 0 &\iff \det A[k] > 0 \quad \forall k = 1, \dots, n. \\ A < 0 &\iff (-1)^k \cdot \det A[k] > 0 \quad \forall k = 1, \dots, n. \end{aligned}$$

Definition 8.12. Sei $B \subset \mathbb{R}^n$ eine beliebige Teilmenge und $f : B \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine reellwertige Funktion. Man sagt:

1. f nimmt in $p \in B$ ein lokales Minimum an, falls ein $\varepsilon > 0$ existiert, so dass $f(x) \geq f(p)$ für alle $x \in B$ mit $\|x - p\| < \varepsilon$.
2. f nimmt in $p \in B$ ein isoliertes (oder striktes) lokales Minimum an, falls ein $\varepsilon > 0$ existiert, so dass $f(x) > f(p)$ für alle $x \in B$ mit $0 < \|x - p\| < \varepsilon$.
3. f nimmt in $p \in B$ ein lokales Maximum an, falls ein $\varepsilon > 0$ existiert, so dass $f(x) \leq f(p)$ für alle $x \in B$ mit $\|x - p\| < \varepsilon$.
4. f nimmt in $p \in B$ ein isoliertes (oder striktes) lokales Maximum an, falls ein $\varepsilon > 0$ existiert, so dass $f(x) < f(p)$ für alle $x \in B$ mit $0 < \|x - p\| < \varepsilon$.

Definition 8.13. Sei $U \subset \mathbb{R}^n$ offen und $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar. $p \in U$ heißt kritischer Punkt von f , wenn $df_p = 0$ (bzw. dazu äquivalent, wenn $\text{grad} f(p) = 0$). Anderenfalls heißt $p \in U$ regulärer Punkt von f .

Satz 8.14 Sei $U \subset \mathbb{R}^n$ eine offene Menge und $f : U \rightarrow \mathbb{R}$ 2-mal stetig differenzierbar.

1. Hat f in $p \in U$ ein lokales Maximum (lokales Minimum), so ist p ein kritischer Punkt von f und die Hesse-Matrix $\text{Hess} f(p)$ ist negativ semidefinit (positiv semidefinit).
2. Ist $p \in U$ ein kritischer Punkt von f und die Hesse-Matrix $\text{Hess} f(p)$ negativ definit (positiv definit), so nimmt f in p ein isoliertes lokales Maximum (isoliertes lokales Minimum) an.

Beweis. Wir zeigen die Behauptungen nur für den Fall eines lokalen Maximums. Die Aussagen für das lokale Minimum folgt dann durch Übergang von f zu $-f$.

(1) f habe in $p \in U$ ein lokales Maximum. Sei $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$. Da U offen ist, existiert ein $\varepsilon > 0$, so dass $K(p, \varepsilon) \subset U$. Folglich ist die Funktion $g_{\mathbf{a}} : \left(-\frac{\varepsilon}{\|\mathbf{a}\|}, \frac{\varepsilon}{\|\mathbf{a}\|}\right) \rightarrow \mathbb{R}$ mit $g_{\mathbf{a}}(t) := f(p + t\mathbf{a})$ korrekt definiert. $g_{\mathbf{a}}$ ist 2-mal stetig differenzierbar hat in $t = 0$ ein lokales Maximum. Somit gilt $g'_{\mathbf{a}}(0) = 0$ und $g''_{\mathbf{a}}(0) \leq 0$. Aus der Kettenregel für Differentiale folgt:

$$0 = g'_{\mathbf{a}}(0) = df_p(\mathbf{a}).$$

Dies gilt für alle $\mathbf{a} \in \mathbb{R}^n$. Folglich ist $df_p = 0$, also $p \in U$ ein kritischer Punkt von f .

Aus der Kettenregel für partielle Ableitungen folgt mit $\gamma(t) := p + t\mathbf{a}$:

$$\begin{aligned} g'_{\mathbf{a}}(t) &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\gamma(t)) \cdot a_i, \\ g''_{\mathbf{a}}(t) &= \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\gamma(t)) \cdot a_i a_j. \end{aligned}$$

Folglich ist

$$0 \geq g_{\mathbf{a}}''(0) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(p) \cdot a_i a_j = \mathbf{a} \cdot \text{Hess}f(p) \cdot \mathbf{a}^t.$$

Dies gilt für alle $\mathbf{a} \in \mathbb{R}^n$, somit ist die Hesse-Matrix von f in p negativ semidefinit.

(2) Sei $p \in U$ ein kritischer Punkt von f und $\text{Hess}f(p) < 0$. Wir zeigen zunächst, dass es eine Kugel $K(p, r) \subset U$ um p gibt, so dass $\text{Hess}f(\xi) < 0$ für alle $\xi \in K(p, r)$. Dazu benutzen wir das Determinanten-Kriterium für negative Definitheit. Für $k = 1, \dots, n$ bezeichne $H(\xi)[k]$ den k -ten Hauptminor der Hesse-Matrix $\text{Hess}f(\xi)$ und $\phi_k : U \rightarrow \mathbb{R}$ die Funktion

$$\phi_k(\xi) := (-1)^k \cdot \det H(\xi)[k].$$

Da f 2-mal stetig differenzierbar und die Determinanten-Funktion stetig ist, sind die Funktionen ϕ_k stetig. Da nach Voraussetzung $\text{Hess}f(p) < 0$, gilt $\phi_k(p) > 0$ für alle $k = 1, \dots, n$. Wegen der Stetigkeit der Funktionen ϕ_k existiert dann eine Kugel $K(p, r) \subset U$, so dass $\phi_k(\xi) > 0$ für alle $\xi \in K(p, r)$ und alle $k = 1, \dots, n$. Folglich ist $\text{Hess}f(\xi) < 0$ für alle $\xi \in K(p, r)$.

Wir approximieren die Funktion f nun bei p durch das Taylorpolynom 1. Grades und erhalten für alle $x \in K(p, r)$ ein $\xi \in \overline{px} \subset K(p, r)$, so dass

$$f(x) = f(p) + \langle \text{grad}f(p), x - p \rangle + \frac{1}{2}(x - p) \cdot \text{Hess}f(\xi) \cdot (x - p)^t$$

(siehe Satz 8.13). Da p ein kritischer Punkt von f ist, gilt $\text{grad}f(p) = 0$. Außerdem wissen wir aus dem oben Bewiesenen, dass die Hesse-Matrix $\text{Hess}f(\xi)$ negativ definit ist. Daraus folgt

$$f(x) < f(p)$$

für alle $x \in K(p, r)$ mit $x \neq p$. Dies zeigt, dass f in p ein isoliertes lokales Maximum annimmt. \square

Beispiel: Wir betrachten die Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ mit

$$f(x, y, z) := 35 - 6x + 2z + x^2 - 2xy + 2y^2 + 2yz + 3z^2.$$

Wir wollen untersuchen, in welchen Punkten f lokale Extremwerte annimmt. Dazu gehen wir folgendermaßen vor:

1. *Bestimme die kritischen Punkte von f :*

Sei $p = (x, y, z)$. Es gilt

$$\text{grad}f(p) = (-6 + 2x - 2y, -2x + 4y + 2z, 2 + 2y + 6z).$$

Folglich ist $p = (x, y, z)$ genau dann ein kritischer Punkt von f , wenn das Gleichungssystem

$$\begin{pmatrix} 2 & -2 & 0 \\ -2 & 4 & 2 \\ 0 & 2 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ -2 \end{pmatrix}$$

erfüllt ist. Dieses hat genau eine Lösung, nämlich $p_0 = (8, 5, -2)$.

2. Bestimme die Hesse-Matrix von f in den kritischen Punkten, also hier im Punkt p_0 :

$$\text{Hess}f(p_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x \partial x} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y \partial y} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z \partial z} \end{pmatrix} (p_0) = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 4 & 2 \\ 0 & 2 & 6 \end{pmatrix} =: H.$$

3. Untersuche $\text{Hess}f(p_0)$ auf Definitheit:

Dazu kann man die Eigenwerte von H bestimmen oder die Determinanten der Hauptminoren von H ausrechnen. Wir betrachten letzteres:

$$\det H[1] = H_{11} = 2, \quad \det H[2] = \det \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix} = 4, \quad \det H[3] = \det H = 16.$$

Alle diese Determinanten sind positiv, folglich ist $\text{Hess}f(p_0)$ positiv definit. f hat in p_0 somit ein striktes lokales Minimum. Weitere lokale Extrema existieren nicht.

Achtung: Die Kriterien von Satz 8.14 gelten nur auf *offenen* Teilmengen. Sucht man lokale bzw. globale Extremwerte einer Funktion f auf Mengen $B \subset \mathbb{R}^n$, die nicht offen sind, so wendet man die Kriterien zunächst auf die Punkte im Inneren von B an (die Menge $\text{Int}(B)$ ist offen). Anschließend muß man das Verhalten von f in den restlichen Punkten von B mit anderen Methoden untersuchen.

8.6 Koordinatentransformationen und der Satz über den lokalen Diffeomorphismus

Definition 8.14. Es seien $U, V \subset \mathbb{R}^n$ offene Teilmengen des \mathbb{R}^n .

Eine Abbildung $f : U \rightarrow V$ heißt *Diffeomorphismus* (oder *Koordinatentransformation*), wenn f bijektiv ist und sowohl f als auch f^{-1} stetig differenzierbar sind.

Offensichtlich ist jeder Diffeomorphismus auch ein Homöomorphismus. Die Verknüpfung zweier Diffeomorphismen ist wieder ein Diffeomorphismus.

Mit Hilfe eines Diffeomorphismus $f : U \rightarrow V$ ordnet man jedem Punkt p aus V neue Koordinaten zu, nämlich diejenigen $(x_1, \dots, x_n) \in U$ mit $f(x_1, \dots, x_n) = p$. Wir werden sehen, wie nützlich solche Koordinatentransformationen sind. Wie im eindimensionalen Fall vereinfachen sie z.B. oft die Berechnung von Integralen. In der *linearen Algebra* werden sie benutzt, um die geometrische Gestalt von Quadriken zu bestimmen (Hauptachsentransformation). Deshalb werden wir uns in diesem Abschnitt damit beschäftigen, wie man feststellen kann, ob eine Abbildung $f : U \rightarrow V$ ein Diffeomorphismus (bzw. eine Koordinatentransformation) ist. Zunächst sehen wir uns einige Beispiele an.

Beispiel 1: Umsortierung der Koordinaten des \mathbb{R}^n

Sei ρ eine Permutation der Zahlen $\{1, \dots, n\}$. Die Abbildung $f_\rho : \mathbb{R}^n \rightarrow \mathbb{R}^n$, die die Reihenfolge der Koordinaten des \mathbb{R}^n entsprechend der Permutation ρ vertauscht, d.h.

$$f_\rho(x_1, x_2, \dots, x_n) := (x_{\rho(1)}, x_{\rho(2)}, \dots, x_{\rho(n)}),$$

ist eine Koordinatentransformation.

Beispiel 2: Affine Koordinatentransformation

Sei $A \in L(\mathbb{R}^n, \mathbb{R}^n)$ ein (linearer) Isomorphismus und $\mathbf{b} \in \mathbb{R}^n$ ein Vektor. Dann heißt die Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit

$$f(x) := A(x) + \mathbf{b}$$

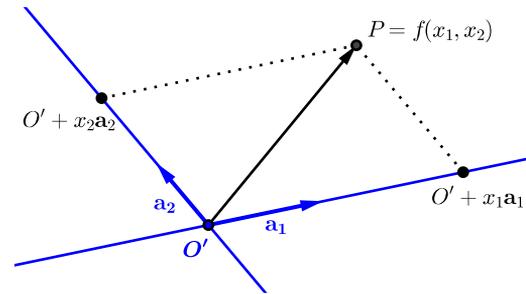
affine Koordinatentransformation. f ist bijektiv, die Umkehrabbildung ist gegeben durch $f^{-1}(y) = A^{-1}(y) - A^{-1}(\mathbf{b})$. f ist stetig differenzierbar, da

$$\frac{\partial f}{\partial x_i}(x) = df_x(e_i) = dA_x(e_i) = A(e_i) = \text{konstant}.$$

Analog zeigt man, dass f^{-1} stetig differenzierbar ist.

Die Koordinaten (x_1, \dots, x_n) des Punktes $P = f(x_1, \dots, x_n)$ sind die Koordinaten von P im Koordinatensystem $(O'; \mathbf{a}_1, \dots, \mathbf{a}_n)$ mit dem Ursprung $O' := O + \mathbf{b}$ und den Basisvektoren $\mathbf{a}_i := Ae_i$, denn

$$P = f(x_1, \dots, x_n) = \mathbf{b} + A\left(\sum_{i=1}^n x_i e_i\right) = O' + \sum_{i=1}^n x_i \mathbf{a}_i.$$



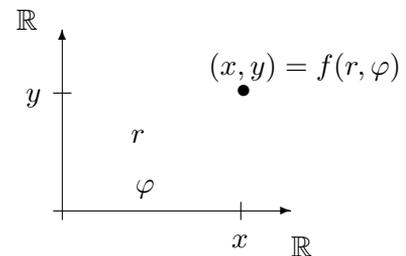
Beispiel 3: Polarkoordinaten in der Ebene

Wir betrachten die Abbildung

$$f : (0, \infty) \times (0, 2\pi) \longrightarrow \mathbb{R}^2 \setminus ([0, \infty) \times \{0\})$$

$$f(r, \varphi) := (r \cos(\varphi), r \sin(\varphi)).$$

(r, φ) heißen *Polarkoordinaten des Punktes* $(x, y) = f(r, \varphi)$.



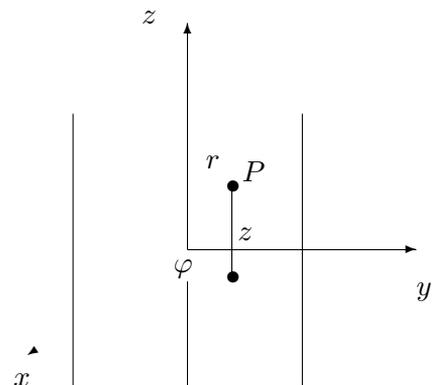
Beispiel 4: Zylinderkoordinaten im \mathbb{R}^3

Wir betrachten die Abbildung

$$f : \underbrace{(0, \infty) \times (0, 2\pi)}_{\text{Polarkoordinaten}} \times \mathbb{R} \longrightarrow (\mathbb{R}^2 \setminus ([0, \infty) \times \{0\})) \times \mathbb{R}$$

$$f(r, \varphi, z) := (r \cos(\varphi), r \sin(\varphi), z).$$

(r, φ, z) heißen *Zylinderkoordinaten des Punktes* $P = f(r, \varphi, z)$.



Beispiel 5: *Kugelkoordinaten im \mathbb{R}^3*

Wir betrachten die Abbildung

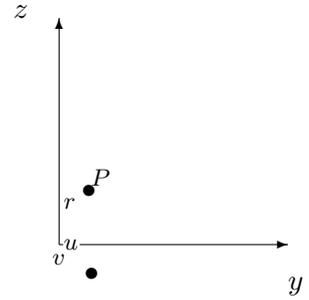
$$f : (0, \infty) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times (0, 2\pi) \longrightarrow \mathbb{R}^3 \setminus ([0, \infty) \times \{0\} \times \mathbb{R})$$

$$f(r, u, v) := (r \cos(u) \cos(v), r \cos(u) \sin(v), r \sin(u)).$$

(r, u, v) heißen *Kugelkoordinaten des Punktes*

$$P = f(r, u, v).$$

$(r \cos(u), v)$ sind dabei die Polarkoordinaten der Projektion (P_1, P_2) von P auf die xy -Ebene.



Die Bijektivität der Abbildungen in den letzten drei Beispielen folgt aus ihrer geometrischen Konstruktion. Die Abbildungen sind offensichtlich stetig differenzierbar. Um zu überprüfen, ob die inverse Abbildung ebenfalls stetig differenzierbar ist, müßte man sie ausrechnen. Wir werden im folgenden eine Möglichkeit angeben, mit denen man sich die explizite Berechnung der inversen Abbildung ersparen kann.

Zunächst hat man die folgende *notwendige* Bedingung dafür, dass $f : U \rightarrow V$ ein Diffeomorphismus ist.

Satz 8.15 *Sei $f : U \rightarrow V$ ein Diffeomorphismus zwischen offenen Mengen des \mathbb{R}^n . Dann ist das Differential $df_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ für alle $p \in U$ ein Isomorphismus und seine inverse Abbildung ist gegeben durch*

$$(df_p)^{-1} = (df^{-1})_{f(p)}.$$

Beweis. Da f stetig differenzierbar ist, ist f auch differenzierbar.

Wir wenden die Kettenregel für Differentiale auf die Abbildungen $f \circ f^{-1} = \text{Id}_V$ und $f^{-1} \circ f = \text{Id}_U$ an. Für alle $q \in V$ und $p \in U$ gilt dann

$$\begin{aligned} d(f \circ f^{-1})_q &= df_{f^{-1}(q)} \circ (df^{-1})_q = (d \text{Id}_{\mathbb{R}^n})_q = \text{Id}_{\mathbb{R}^n} \quad \text{und} \\ d(f^{-1} \circ f)_p &= (df^{-1})_{f(p)} \circ df_p = (d \text{Id}_{\mathbb{R}^n})_p = \text{Id}_{\mathbb{R}^n}. \end{aligned}$$

Betrachten wir $q = f(p)$, so folgt $(df^{-1})_{f(p)} = (df_p)^{-1}$. □

Bemerkung 1: Die Jacobi-Matrix Jf_p ist die Darstellungsmatrix des Differentials df_p . Folglich gilt:

$$df_p : \mathbb{R}^n \rightarrow \mathbb{R}^n \text{ ist ein Isomorphismus} \iff Jf_p \text{ ist invertierbar} \iff \det Jf_p \neq 0.$$

Bemerkung 2: Im 1-dimensionalen gilt auch die Umkehrung von Satz 8.15.

Sei $U \subset \mathbb{R}$ ein offenes Intervall und $f : U \subset \mathbb{R} \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion mit $f'(x) \neq 0$ für alle $x \in U$. Dann ist f streng monoton. Insbesondere folgt dann bereits, dass $f : U \rightarrow f(U)$ bijektiv ist. Außerdem ist die Umkehrfunktion stetig differenzierbar. $f : U \rightarrow f(U)$ ist also ein Diffeomorphismus.

Im \mathbb{R}^n mit $n > 1$ gilt diese Aussage nicht mehr. Eine stetig differenzierbare Abbildung

$f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, deren Jacobi-Matrix in jedem Punkt invertierbar ist, muß nicht mal injektiv sein. Als Beispiel betrachten wir die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, definiert durch

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

Diese Abbildung ist stetig differenzierbar und für die Determinante der Jacobi-Matrix von f gilt:

$$\det Jf_{(x,y)} = \det \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} = e^{2x} > 0.$$

f ist aber offensichtlich nicht injektiv. Es gibt also keine globale Umkehrabbildung von $f : U \rightarrow f(U)$.

Wir wollen nun untersuchen, unter welchen Bedingungen eine stetig differenzierbare Abbildung $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ wenigstens *lokal* eine (stetig differenzierbare) Umkehrfunktion besitzt.

Definition 8.15. Eine Abbildung $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt *lokaler Diffeomorphismus* um $p \in U$, wenn es offene Umgebungen² $\tilde{U} \subset U$ von p und $\tilde{V} \subset \mathbb{R}^n$ von $f(p)$ gibt, so dass $f|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein Diffeomorphismus ist.

Satz 8.16 (Satz über den lokalen Diffeomorphismus)

Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar, $p \in U$ und df_p ein Isomorphismus. Dann ist f ein lokaler Diffeomorphismus um p .

Beweis. 1. Schritt: Reduktion auf eine einfache Situation:

a) Wir können annehmen, dass $df_p = \text{Id}_{\mathbb{R}^n}$ gilt.

Anderenfalls betrachten wir die Abbildung $\tilde{f} := (df_p)^{-1} \circ f$. Da $(df_p)^{-1}$ linear ist, folgt aus der Kettenregel für Differentiale $d\tilde{f}_p = (df_p)^{-1} \circ df_p = \text{Id}_{\mathbb{R}^n}$. Ist $\tilde{f}|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein Diffeomorphismus, so ist $f|_{\tilde{U}} : \tilde{U} \rightarrow df_p(\tilde{V})$ ein Diffeomorphismus.

b) Wir können außerdem annehmen, dass $p = 0$ und $f(p) = 0$ gilt.

Anderenfalls betrachten wir die Abbildung \hat{f} , definiert durch $\hat{f}(x) := f(x + p) - f(p)$. Dann gilt $\hat{f}(0) = 0$ und $d\hat{f}_0 = df_p$. Ist $\hat{f}|_{U_0} : U_0 \rightarrow V_0$ ein Diffeomorphismus um 0 und $\tilde{U} := U_0 + p$, $\tilde{V} := V_0 + f(p)$, dann ist $f|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein Diffeomorphismus um p .

Es genügt also zu zeigen: Ist $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar, $0 \in U$, $f(0) = 0$ und $df_0 = \text{Id}_{\mathbb{R}^n}$, so ist f ein lokaler Diffeomorphismus um 0.

2. Schritt: Wir konstruieren offene Mengen U_0 und V_0 um 0, so dass $f|_{U_0} : U_0 \rightarrow V_0$ bijektiv ist: Da f stetig differenzierbar ist, ist die Abbildung

$$x \in U \mapsto Jf_x \in M(n, n, \mathbb{R})$$

stetig (bzgl. der Euklidischen Norm auf den Matrizen). Da die gewählte Norm für Konvergenzfragen keine Rolle spielt (siehe Satz 7.26) ist auch die Abbildung

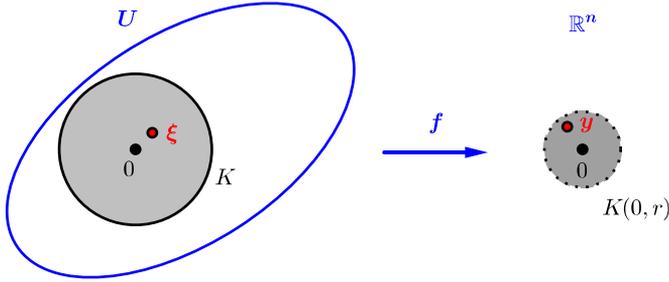
² Mit *Umgebung eines Punktes* p meinen wir eine Menge, die den Punkt p enthält.

$$x \in U \longmapsto df_x \in L(\mathbb{R}^n, \mathbb{R}^n)$$

stetig bzgl. der Operatornorm auf $L(\mathbb{R}^n, \mathbb{R}^n)$. Folglich existiert ein $r > 0$, so dass die abgeschlossene Kugel $K := \{x \in \mathbb{R}^n \mid \|x\| \leq 2r\}$ in U liegt und für alle $x \in K$ gilt

$$\|df_x - df_0\| = \|df_x - \text{Id}_{\mathbb{R}^n}\| < \frac{1}{2}. \quad (8.12)$$

Sei $y \in K(0, r)$ fixiert. Wir zeigen nun mit Hilfe des Banachschen Fixpunktsatzes, dass es genau ein $\xi \in K(0, 2r) = \text{Int}(K)$ gibt mit $f(\xi) = y$.



Dazu betrachten wir die Abbildung $\phi : U \rightarrow \mathbb{R}^n$ mit

$$\phi(x) := y + x - f(x).$$

Dann gilt $\phi(\xi) = \xi$ genau dann, wenn $f(\xi) = y$. Die Urbilder von y bei f sind also die Fixpunkte von ϕ . Außerdem gilt $\phi(0) = y$ und $d\phi_x = \text{Id}_{\mathbb{R}^n} - df_x$ für alle $x \in U$. Aus (8.12) folgt dann

$$\|d\phi_x\| = \|\text{Id}_{\mathbb{R}^n} - df_x\| < \frac{1}{2} \quad \forall x \in K.$$

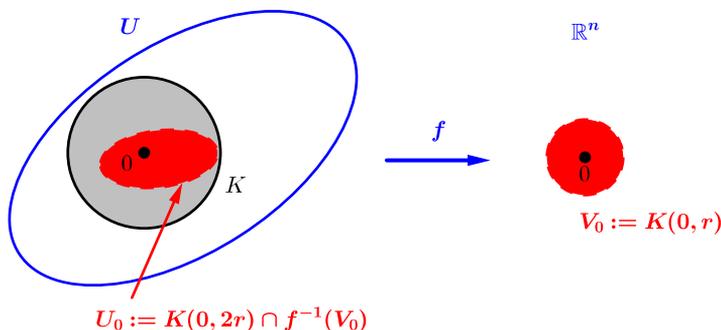
Somit ist ϕ auf K Lipschitzstetig mit der Lipschitzkonstanten $\frac{1}{2}$ (siehe Satz 8.10), d.h. es gilt

$$\|\phi(b) - \phi(a)\| \leq \frac{1}{2}\|b - a\| \quad \forall a, b \in K. \quad (8.13)$$

Insbesondere ist

$$\begin{aligned} \|\phi(x)\| &\leq \|\phi(x) - \phi(0)\| + \|\phi(0)\| \\ &\stackrel{(8.13)}{\leq} \frac{1}{2}\|x\| + \|y\| < r + r = 2r \quad \forall x \in K, \end{aligned}$$

d.h., ϕ bildet die abgeschlossene Kugel K in die offene Kugel $K(0, 2r) \subset K$ ab. Da $K \subset \mathbb{R}^n$ abgeschlossen und \mathbb{R}^n ein vollständiger metrischer Raum ist, ist K ebenfalls vollständig. $\phi : K \rightarrow K$ ist also eine kontrahierende Abbildung auf einem vollständigen metrischen Raum. Der Banachsche Fixpunktsatz (Satz 7.24) liefert dann die Existenz eines Fixpunktes ξ von ϕ auf K und seine Eindeutigkeit. Da ϕ in $K(0, 2r)$ abbildet, gilt $\xi \in K(0, 2r)$. Es gibt also genau ein $\xi \in K(0, 2r)$ mit $f(\xi) = y$. Die Kugel $V_0 := K(0, r)$ ist eine offene Umgebung von $f(0) = 0$, die Menge $U_0 := K(0, 2r) \cap f^{-1}(V_0)$ ist eine offene Umgebung von 0 (topologisches Stetigkeitskriterium) und nach Konstruktion ist die Abbildung $f|_{U_0} : U_0 \rightarrow V_0$ bijektiv.



3. *Schritt:* Wir zeigen nun, dass die Umkehrabbildung $g := (f|_{U_0})^{-1} : V_0 \rightarrow U_0$ stetig ist. Dazu betrachten wir die Abbildung $\varphi : U \rightarrow \mathbb{R}^n$ mit $\varphi(x) = x - f(x)$ für alle $x \in U$. φ stimmt mit der Abbildung ϕ (für $y = 0$) aus dem 2. Schritt überein. Für $a, b \in U_0$ folgt damit

$$b - a = \varphi(b) + f(b) - (\varphi(a) + f(a))$$

und somit

$$\|b - a\| \leq \|\varphi(b) - \varphi(a)\| + \|f(b) - f(a)\| \stackrel{(8.13)}{\leq} \frac{1}{2}\|b - a\| + \|f(b) - f(a)\|.$$

Also gilt

$$\|b - a\| \leq 2 \cdot \|f(b) - f(a)\| \quad \forall b, a \in U_0. \quad (8.14)$$

Für $u, w \in V_0$ setzen wir $a := g(u)$, $b := g(w)$. Dann folgt aus (8.14)

$$\|g(w) - g(u)\| \leq 2 \cdot \|w - u\| \quad \forall w, u \in V_0.$$

g ist also lipschitzstetig und somit stetig.

4. *Schritt:* Das Differential $df_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist für jedes $p \in U_0$ ein Isomorphismus. Angenommen, es gäbe ein $v \in \mathbb{R}^n$ mit $v \neq 0$ und $df_p(v) = 0$. Dann ergibt

$$\|v\| = \|v - df_p(v)\| = \|(\text{Id}_{\mathbb{R}^n} - df_p)(v)\| \leq \|\text{Id}_{\mathbb{R}^n} - df_p\| \cdot \|v\| \stackrel{(8.12)}{\leq} \frac{1}{2}\|v\|$$

einen Widerspruch.

5. *Schritt:* Die Umkehrabbildung $g := (f|_{U_0})^{-1} : V_0 \rightarrow U_0$ ist differenzierbar.

Sei $q \in V_0$ ein beliebiger Punkt und $p = g(q) \in U_0$. Wir zeigen, dass g in q differenzierbar ist. Da f in p differenzierbar ist, gilt für $p + h \in U_0$

$$f(p + h) - f(p) = df_p(h) + r_p(h) \quad \text{mit} \quad \lim_{h \rightarrow 0} \frac{r_p(h)}{\|h\|} = 0. \quad (8.15)$$

Sei $y := f(p + h)$. Dann gilt $g(y) = p + h = g(q) + h$. Aus der Gleichung (8.15) erhalten wir h durch Anwendung von $(df_p)^{-1}$:

$$h = g(y) - g(q) = \underbrace{(df_p)^{-1}(y - q)}_{\text{linear}} - \underbrace{(df_p)^{-1}(r_p(h))}_{\text{Restglied}}.$$

Wir müssen nun für das Restglied zeigen, dass

$$\lim_{y \rightarrow q} \frac{\|(df_p)^{-1}(r_p(h))\|}{\|y - q\|} = 0. \quad (8.16)$$

Es gilt

$$\frac{\|(df_p)^{-1}(r_p(h))\|}{\|y - q\|} \leq \|(df_p)^{-1}\| \cdot \frac{\|r_p(h)\|}{\|y - q\|} = \|(df_p)^{-1}\| \cdot \frac{\|r_p(h)\|}{\|h\|} \cdot \frac{\|h\|}{\|f(p+h) - f(p)\|}.$$

Aus (8.14) folgt

$$\|h\| = \|(p+h) - p\| \leq 2 \cdot \|f(p+h) - f(p)\|.$$

Wir erhalten also die Abschätzung

$$\frac{\|(df_p)^{-1}(r_p(h))\|}{\|y - q\|} \leq 2 \cdot \|(df_p)^{-1}\| \cdot \frac{\|r_p(h)\|}{\|h\|}.$$

Da g stetig ist, folgt aus $y \rightarrow q$, dass $h = g(y) - g(q) \rightarrow 0$. Dies liefert mit (8.15) die Restgliedeigenschaft (8.16). Somit ist g in $q \in V_0$ differenzierbar und es gilt $dg_q = (df_p)^{-1}$.

6. Schritt: Als letztes zeigen wir, dass $g := (f|_{U_0})^{-1} : V_0 \rightarrow U_0$ stetig differenzierbar ist.

Die Jacobi-Matrix von g in $q \in V_0$ ist die inverse Matrix der Jacobi-Matrix von f in $p = g(q)$. Die Vorschrift zur Berechnung inverser Matrizen zeigt, dass die partiellen Ableitungen von g in q rationale Funktionen der partiellen Ableitungen von f in p sind, d.h. es existieren rationale Funktionen R_{kl} in n^2 Variablen, so dass

$$\frac{\partial g_k}{\partial y_l}(q) = R_{kl} \left(\frac{\partial f_i}{\partial x_j}(g(q)), \quad i, j = 1, \dots, n \right).$$

Da f stetig differenzierbar ist, sind die partiellen Ableitungen von f stetig. g ist stetig (Schritt 3). Rationale Funktionen sind auch stetig. Somit sind die partiellen Ableitungen von g als Verknüpfung stetiger Funktionen ebenfalls stetig. \square

Abschließend beweisen wir ein Kriterium für das Vorliegen eines globalen Diffeomorphismus (bzw. einer Koordinatentransformation).

Satz 8.17 (Globaler Umkehrsatz)

Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar und sei das Differential $df_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ für alle $p \in U$ ein Isomorphismus. Dann gilt:

1. $V := f(U) \subset \mathbb{R}^n$ ist offen.
2. Ist f zusätzlich injektiv, so ist $f : U \rightarrow V$ ein Diffeomorphismus.

Beweis. (1) Nach Satz 8.16 existiert für alle $p \in U$ eine offene Umgebung $U_p \subset U$ von p und V_p von $f(p)$, so dass $f|_{U_p} : U_p \rightarrow V_p$ ein Diffeomorphismus ist. Dann gilt $U = \bigcup_{p \in U} U_p$ und

$$f(U) = \bigcup_{p \in U} f|_{U_p}(U_p) = \bigcup_{p \in U} V_p =: V.$$

Die Menge V ist insbesondere offen (Siehe Übungsaufgabe 20).

(2) Ist f injektiv, so ist $f : U \rightarrow V = f(U)$ bijektiv. Da $f|_{U_p} : U_p \rightarrow V_p$ für alle $p \in U$ ein Diffeomorphismus ist und Differenzierbarkeit eine lokale Eigenschaft ist, ist $f : U \rightarrow V = f(U)$ ein (globaler) Diffeomorphismus. \square

Mit diesem Satz kann man nun leicht zeigen, dass Abbildungen aus den Beispielen 3, 4 und 5, die die Polarkoordinaten, die Zylinderkoordinaten und die Kugelkoordinaten definieren, tatsächlich Koordinatentransformationen sind. Man berechnet dazu die Jacobi-Matrizen und zeigt, dass ihre Determinate in jedem Punkt ungleich Null ist. (Übungsaufgabe).

8.7 Auflösen von Gleichungen. Der Satz über implizite Funktionen

Eine Gerade in der Euklidischen Ebene \mathbb{R}^2 kann man auf verschiedene Weise beschreiben:

a) *Implizite Beschreibung* als Lösungsmenge der Gleichung

$$f(x, y) := \alpha x + \beta y + \gamma = 0, \quad (*)$$

wobei α, β, γ reelle Konstanten mit $(\alpha, \beta) \neq (0, 0)$ sind.

b) *Explizite Beschreibung* als Graph einer Funktion $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ über der x -Achse oder einer Funktion $\psi : \mathbb{R} \rightarrow \mathbb{R}$ über der y -Achse mit

$$\varphi(x) := mx + d \quad \text{bzw.} \quad \psi(y) = \tilde{m}y + \tilde{d}.$$

Ist $\beta \neq 0$, so erhält man φ durch Auflösen der Gleichung $(*)$ nach der y -Variablen:

$$y = \varphi(x) := -\frac{\alpha}{\beta}x - \frac{\gamma}{\beta}.$$

Ist $\alpha \neq 0$ so kann man $(*)$ nach x auflösen und erhält

$$x = \psi(y) := -\frac{\beta}{\alpha}y - \frac{\gamma}{\alpha}.$$

In der expliziten Form erkennt man die Gestalt der Geraden sofort.

In der *Lineare Algebra* wird dieses Verfahren auf Systeme von m linearen Gleichungen mit $N \geq m$ Unbestimmten verallgemeinert. Auch hier löst man die Gleichungen nach einzelnen Unbestimmten auf (Gauß-Algorithmus).

Wir wollen jetzt die analoge Frage für *nicht-lineare* Gleichungen stellen. Wir betrachten m Funktionen f_1, \dots, f_m , die von $N \geq m$ Variablen (x_1, \dots, x_N) abhängen und das System der Gleichungen

$$\begin{aligned} f_1(x_1, \dots, x_N) &= 0, \\ f_2(x_1, \dots, x_N) &= 0, \\ &\vdots \\ f_m(x_1, \dots, x_N) &= 0, \end{aligned}$$

oder in Kurzfassung: $f(x) = 0$ für $f = (f_1, \dots, f_m)$ und $x = (x_1, \dots, x_N)$. Dann stellen sich folgende Fragen:

- Hat dieses Gleichungssystem eine Lösung?
- Von wievielen Parametern hängt die Lösungsmenge ab?
(D.h. wie "groß" ist die Lösungsmenge?)
- Wann kann man die Lösungsmenge *explizit* als Graph einer Funktion φ beschreiben?

Um diese Fragen zu beantworten, versucht man wieder, die Gleichungen nach einer oder mehreren Variablen aufzulösen.

Wir sehen uns zunächst zwei Beispiele in der Ebene an, die die typische Effekte zeigen, die auftreten, wenn man eine *nicht-lineare* Gleichung nach einer Variablen aufzulösen versucht.

Beispiel 1: Der Kreis S^1

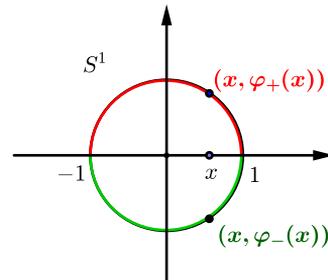
Wir betrachten in der Ebene den Kreis S^1 um $(0,0)$ vom Radius 1. S^1 ist die Lösungsmenge der Gleichung

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

Man kann die Gleichung

$x^2 + y^2 - 1 = 0$ nach y auflösen und erhält

$$y = \pm \sqrt{1 - x^2} =: \varphi_{\pm}(x).$$



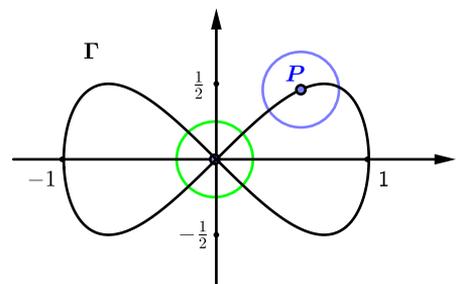
Es gibt hier also zwei Auflösungen φ_{\pm} der Gleichung. Beide sind nur auf dem Intervall $[-1, 1]$ definiert und beschreiben jeweils nur einen Teil der Lösungsmenge S^1 . Eine *globale* Beschreibung von S^1 als Graph einer Funktion gibt es nicht.

Beispiel 2: Die Lemniskate von Gerono

Sei $\Gamma \subset \mathbb{R}^2$ die Lösungsmenge der Gleichung

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

Betrachtet man einen beliebigen Punkt $P \in \Gamma$ mit $P \neq (0,0)$, so gibt es eine kleine Umgebung U von P , für die man $\Gamma \cap U$ als Graph einer Funktion über der x -Achse oder der y -Achse schreiben kann. Im Punkt $(0,0)$ geht das allerdings nicht, da Γ mit zwei Zweigen durch $(0,0)$ läuft. In keiner noch so kleinen Umgebung von $(0,0)$ kann man Γ nach x oder y auflösen.



Lemniskate von Gerono

Im nächsten Satz geben wir Bedingungen dafür an, wann man die Lösungsmenge einer Gleichung $f(x) = 0$ wenigstens *lokal* nach einer bestimmten Zahl von Variablen auflösen kann, d.h. wann man sie *lokal* als Graph einer Funktion φ beschreiben kann. Dabei kann

man in der Regel zwar φ nicht konkret angeben, man erhält aber nützliche konkrete Formeln für ihre partiellen Ableitungen.

Satz 8.18 (Satz über implizite Funktionen)

Sei $U \subset \mathbb{R}^n \times \mathbb{R}^m$ offen und $f : U \rightarrow \mathbb{R}^m$ eine stetig differenzierbare Abbildung.

(x_1, \dots, x_n) (y_1, \dots, y_m)

Wir betrachten einen Punkt $(a, b) \in U$ mit

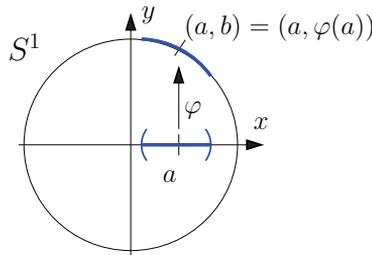
1. $f(a, b) = 0$, d.h. (a, b) löst die Gleichung $f(x, y) = 0$.
2. Die letzten m Spalten der Jacobi-Matrix $Jf_{(a,b)}$ sind linear unabhängig, d.h.,

$$\det \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a, b) & \frac{\partial f_1}{\partial y_2}(a, b) & \dots & \frac{\partial f_1}{\partial y_m}(a, b) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(a, b) & \frac{\partial f_m}{\partial y_2}(a, b) & \dots & \frac{\partial f_m}{\partial y_m}(a, b) \end{pmatrix} \neq 0.$$

Dann existieren offene Umgebungen $U_0 \subset U$ von (a, b) und $A(a) \subset \mathbb{R}^n$ von a , sowie eine eindeutig bestimmte Funktion $\varphi : A(a) \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit

1. $\varphi(a) = b$,
2. $f(x, \varphi(x)) = 0$ für alle $x \in A(a)$.
3. $f^{-1}(0) \cap U_0 = \text{graph}(\varphi)$,

d.h., in einer Umgebung des Punktes (a, b) ist die Lösungsmenge der Gleichung $f(x, y) = 0$ durch den Graphen der Funktion φ beschrieben.



4. $\varphi : A(a) \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist stetig differenzierbar und die partiellen Ableitungen sind gegeben durch

$$\underbrace{\left(\frac{\partial \varphi_l}{\partial x_i}(x) \right)}_{\in M(m,n,\mathbb{R})} = - \underbrace{\left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x)) \right)^{-1}}_{\in M(m,m,\mathbb{R})} \cdot \underbrace{\left(\frac{\partial f_k}{\partial x_i}(x, \varphi(x)) \right)}_{\in M(m,n,\mathbb{R})} \quad \forall x \in A(a).$$

Beweis. 1) Konstruktion von φ (d.h. Auflösung von $f(x, y) = 0$ nach $y = (y_1, \dots, y_m)$): Wir wenden dazu den Satz über den lokalen Diffeomorphismus an. Wir betrachten die Abbildung $F : U \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^m$, definiert durch

$$F(x, y) := (x, f(x, y)).$$

F ist stetig differenzierbar und nach Voraussetzung gilt:

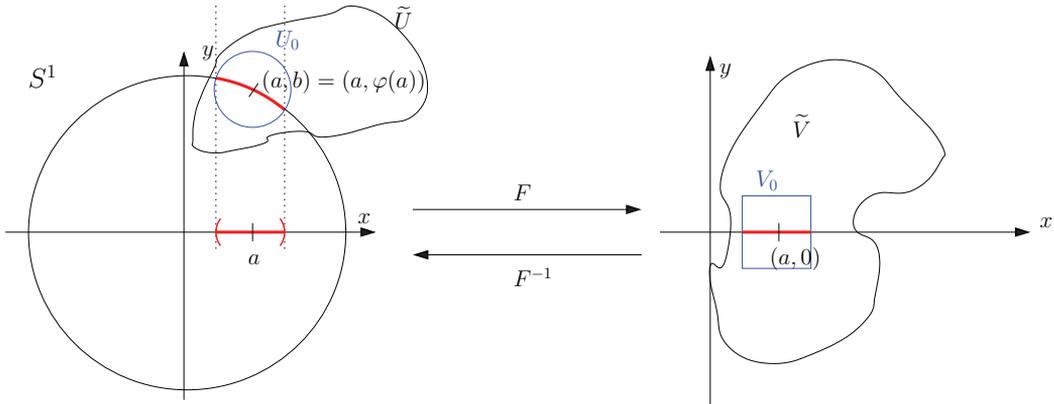
$$\det JF_{(a,b)} = \det \begin{pmatrix} E_n & & 0 \\ & \frac{\partial f_1}{\partial y_1}(a, b) & \dots & \frac{\partial f_1}{\partial y_m}(a, b) \\ * & \vdots & & \vdots \\ & \frac{\partial f_m}{\partial y_1}(a, b) & \dots & \frac{\partial f_m}{\partial y_m}(a, b) \end{pmatrix} \neq 0.$$

Nach dem Satz über den lokalen Diffeomorphismus (Satz 8.16) existieren offene Umgebungen $\tilde{U} \subset U$ um (a, b) und $\tilde{V} \subset \mathbb{R}^n \times \mathbb{R}^m$ um $F(a, b) = (a, 0)$, so dass $F|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein Diffeomorphismus ist. Da \tilde{V} offen ist, können wir eine offene Umgebung $A(a) \subset \mathbb{R}^n$ von a

und einen offenen Würfel $(-\varepsilon, \varepsilon)^m \subset \mathbb{R}^m$ wählen, so dass $V_0 := A(a) \times (-\varepsilon, \varepsilon)^m \subset \tilde{V}$. Sei nun $U_0 := F^{-1}(V_0) \subset \tilde{U}$. Dann ist $F|_{U_0} : U_0 \rightarrow V_0$ ein Diffeomorphismus. Wir betrachten die inverse Abbildung $(F|_{U_0})^{-1} : V_0 \rightarrow U_0$ mit den Komponenten

$$(F|_{U_0})^{-1} =: (F_1^-, F_2^-).$$

$\begin{matrix} \uparrow & \uparrow \\ \mathbb{R}^n & \mathbb{R}^m \end{matrix}$



Dann erhalten wir

$$(x, y) = F(F|_{U_0})^{-1}(x, y) = F(F_1^-(x, y), F_2^-(x, y)) \stackrel{\text{Def.}}{=} (F_1^-(x, y), f(F_1^-(x, y), F_2^-(x, y))).$$

Folglich gilt

$$\begin{aligned} F_1^-(x, y) &= x, \\ f(x, F_2^-(x, y)) &= y \quad \forall (x, y) \in V_0. \end{aligned}$$

Dies zeigt insbesondere, dass

$$f^{-1}(0) \cap U_0 = \{(x, F_2^-(x, 0)) \mid x \in A(a)\}.$$

Wir definieren nun $\varphi : A(a) \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ durch

$$\varphi(x) := F_2^-(x, 0).$$

Dann gilt:

1. $\varphi(a) = F_2^-(a, 0) = b$.
2. $f(x, \varphi(x)) = f(x, F_2^-(x, 0)) = 0$ für alle $x \in A(a)$.
3. $f^{-1}(0) \cap U_0 = \{(x, \varphi(x)) \mid x \in A(a)\} = \text{graph}(\varphi)$.

Die Eindeutigkeit der Abbildung φ mit den Eigenschaften 1) - 3) folgt aus der Konstruktion.

2) *Differenzierbarkeit von φ :*

$F|_{U_0} : U_0 \rightarrow V_0$ ist ein Diffeomorphismus, folglich ist die inverse Abbildung $(F|_{U_0})^{-1}$ stetig differenzierbar. Das gleiche gilt dann auch für ihre Komponente F_2^- . Da $\varphi(x) = F_2^-(x, 0)$, ist auch φ stetig differenzierbar.

3) Die Formel für die partiellen Ableitungen von φ :

Wir wenden die Kettenregel für partielle Ableitungen auf die Funktion

$$g : A(a) \subset \mathbb{R}^n \rightarrow \mathbb{R}^m,$$

$$g(x) := f(x, \varphi(x)) = f(x_1, \dots, x_n, \varphi_1(x_1, \dots, x_n), \dots, \varphi_m(x_1, \dots, x_n)) = 0,$$

an und erhalten für $x \in A(a)$:

$$\begin{aligned} 0 &= \frac{\partial g}{\partial x_i}(x) \\ &= \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x, \varphi(x)) \cdot \underbrace{\frac{\partial x_j}{\partial x_i}}_{=\delta_{ij}} + \sum_{l=1}^m \frac{\partial f}{\partial y_l}(x, \varphi(x)) \cdot \frac{\partial \varphi_l}{\partial x_i}(x) \\ &= \frac{\partial f}{\partial x_i}(x, \varphi(x)) + \sum_{l=1}^m \frac{\partial f}{\partial y_l}(x, \varphi(x)) \cdot \frac{\partial \varphi_l}{\partial x_i}(x). \end{aligned}$$

Für die Komponenten (f_1, \dots, f_m) der Funktion f bedeutet dies in Matrixschreibweise

$$-\left(\frac{\partial f_k}{\partial x_i}(x, \varphi(x))\right) = \left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x))\right) \cdot \left(\frac{\partial \varphi_l}{\partial x_i}(x)\right). \quad (*)$$

Da die Abbildung $F|_{U_0} : U_0 \rightarrow V_0$ ein Diffeomorphismus ist, ist die Matrix $\left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x))\right)$ für jedes $x \in A(a)$ invertierbar. Nach Multiplikation der Gleichung (*) mit der Inversen der Matrix $\left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x))\right)$ erhalten wir für die Ableitungen von φ

$$\left(\frac{\partial \varphi_l}{\partial x_i}(x)\right) = -\left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x))\right)^{-1} \cdot \left(\frac{\partial f_k}{\partial x_i}(x, \varphi(x))\right).$$

□

8.8 Gleichungsdefinierte Flächen, ihre Tangentialebenen und Normalen

Am Ende von Abschnitt 8.1 hatten wir eine Formel für die Tangentialebenen an Flächen angegeben, die als Graph einer reellen Funktion in 2 Variablen gegeben sind:

Sei $\varphi : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion und F der Graph von φ , d.h.

$$F := \{(x, \varphi(x)) \mid x \in U\}.$$

Dann gilt für die Tangentialebene an F im Punkt $p = (x, y, \varphi(x, y)) \in F$

$$\text{Tan}_p F = p + \mathbb{R} \cdot \left(1, 0, \frac{\partial \varphi}{\partial x}(x, y)\right) + \mathbb{R} \cdot \left(0, 1, \frac{\partial \varphi}{\partial y}(x, y)\right).$$

Wir wollen nun mit Hilfe des Satzes über implizite Funktionen eine Formel für die Tangentialebenen an gleichungsdefinierte Flächen in Termen der definierenden Gleichung herleiten.

Definition 8.16. Sei $f : U \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion und

$$M := \{(x, y, z) \in U \mid f(x, y, z) = 0\} = f^{-1}(0) \subset \mathbb{R}^3$$

die Lösungsmenge der Gleichung $f(x, y, z) = 0$.

M heißt reguläre Fläche, wenn jeder Punkt $p \in M$ ein regulärer Punkt von f ist, d.h. wenn $\text{grad}f(p) \neq 0$ für alle $p \in M$.

Der Tangentialraum an M im Punkt $p \in M$ ist die Menge der Vektoren

$$T_p M := \{v \in \mathbb{R}^3 \mid \exists \text{ diffb. Kurve } \gamma : (-\varepsilon, \varepsilon) \rightarrow M \text{ mit } \gamma(0) = p \text{ und } \gamma'(0) = v\}.$$

Beispiel: Der Graph $F = \text{graph}(\varphi)$ ist eine reguläre Fläche.

Wir betrachten die Funktion $f : U \times \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f(x, y, z) := z - \varphi(x, y).$$

Dann gilt $F = f^{-1}(0) = \{(x, y, z) \in U \times \mathbb{R} \mid z = \varphi(x, y)\}$. Für den Gradienten von f erhalten wir

$$\text{grad}f(x, y, z) = \left(-\frac{\partial\varphi}{\partial x}(x, y), -\frac{\partial\varphi}{\partial y}(x, y), 1 \right) \neq 0.$$

Folglich ist F eine reguläre Fläche.

Satz 8.19 Sei $f : U \subset \mathbb{R}^3$ stetig differenzierbar, $M = f^{-1}(0) \subset \mathbb{R}^3$ eine reguläre Fläche und $p \in M$. Dann ist der Tangentialraum $T_p M$ ein 2-dimensionaler Vektorraum und es gilt

$$T_p M = \{v \in \mathbb{R}^3 \mid v \perp \text{grad}f(p)\} = \{v \in \mathbb{R}^3 \mid \langle v, \text{grad}f(p) \rangle = 0\}.$$

Beweis. Sei $p = (p_1, p_2, p_3) \in M$. Da p ein regulärer Punkt von f ist, gilt $\text{grad}f(p) \neq 0$, d.h. mindestens eine der partiellen Ableitungen $\frac{\partial f}{\partial x}(p)$, $\frac{\partial f}{\partial y}(p)$, $\frac{\partial f}{\partial z}(p)$ ist von Null verschieden. Wir nehmen oBdA. den 3. Fall an (sonst sortieren wir die Koordinaten um). Nach dem Satz über implizite Funktionen kann man die Gleichung $f(x, y, z) = 0$ in der Umgebung von p nach z auflösen, d.h. existiert eine offene Umgebung $A \subset \mathbb{R}^2$ von (p_1, p_2) und eine stetig differenzierbare Funktion $\varphi : A \rightarrow \mathbb{R}$, so dass $\varphi(p_1, p_2) = p_3$ und $f(x, y, \varphi(x, y)) = 0$ für alle $(x, y) \in A$.³ Es folgt mit der Kettenregel durch partielles Ableiten nach x und y :

$$0 = \frac{\partial f}{\partial x}(p) + \frac{\partial f}{\partial z}(p) \cdot \frac{\partial \varphi}{\partial x}(p_1, p_2), \quad (8.17)$$

$$0 = \frac{\partial f}{\partial y}(p) + \frac{\partial f}{\partial z}(p) \cdot \frac{\partial \varphi}{\partial y}(p_1, p_2). \quad (8.18)$$

Wir zeigen nun die Formel für den Tangentialraum.

1) $T_p M \subset \{v \in \mathbb{R}^3 \mid v \perp \text{grad}f(p)\}$:

Sei $v \in T_p M$. Dann existiert eine differenzierbare Kurve $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ mit $\gamma(0) = p$ und $\gamma'(0) = v$. Da $f(\gamma(t)) = 0$ für alle $t \in (-\varepsilon, \varepsilon)$, gilt nach Kettenregel für Differentiale

³ Man könnte jetzt die Formeln für die Tangentialebenen eines Graphen aus Abschnitt 8.1 benutzen. Wir werden die Behauptung aber nochmal direkt beweisen.

$$0 = df_{\gamma(t)}(\gamma'(t)) = \langle \text{grad}f(\gamma(t)), \gamma'(t) \rangle = 0,$$

und somit für $t = 0$:

$$0 = \langle \text{grad}f(p), v \rangle = 0.$$

Also ist v senkrecht zum Vektor $\text{grad}f(p)$.

2) $\{v \in \mathbb{R}^3 \mid v \perp \text{grad}f(p)\} \subset T_pM$:

Sei $v = (v_1, v_2, v_3) \in \mathbb{R}^3$ ein Vektor, der senkrecht zu $\text{grad}f(p)$ steht. Dann gilt

$$0 = \langle \text{grad}f(p), v \rangle = \frac{\partial f}{\partial x}(p) \cdot v_1 + \frac{\partial f}{\partial y}(p) \cdot v_2 + \frac{\partial f}{\partial z}(p) \cdot v_3.$$

Setzen wir in diese Gleichung (8.17) und (8.18) ein, so erhalten wir

$$0 = -\frac{\partial f}{\partial z}(p) \cdot \frac{\partial \varphi}{\partial x}(p_1, p_2) \cdot v_1 - \frac{\partial f}{\partial z}(p) \cdot \frac{\partial \varphi}{\partial y}(p_1, p_2) \cdot v_2 + \frac{\partial f}{\partial z}(p) \cdot v_3,$$

und nach Division durch $\frac{\partial f}{\partial z}(p) \neq 0$:

$$v_3 = \frac{\partial \varphi}{\partial x}(p_1, p_2) \cdot v_1 + \frac{\partial \varphi}{\partial y}(p_1, p_2) \cdot v_2. \quad (8.19)$$

Wir betrachten nun das parametrisierte Geradenstück $\sigma : (-\varepsilon, \varepsilon) \rightarrow A$ mit $\sigma(t) = (p_1, p_2) + t(v_1, v_2)$ und die differenzierbare Kurve $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ mit $\gamma(t) = (\sigma(t), \varphi(\sigma(t)))$. Dann gilt $\gamma(0) = (p_1, p_2, \varphi(p_1, p_2)) = p$ und nach Kettenregel

$$\begin{aligned} \gamma'(0) &= \left(v_1, v_2, d\varphi_{(p_1, p_2)}(v_1, v_2) \right) = \left(v_1, v_2, \frac{\partial \varphi}{\partial x}(p_1, p_2) \cdot v_1 + \frac{\partial \varphi}{\partial y}(p_1, p_2) \cdot v_2 \right) \\ &\stackrel{(8.19)}{=} (v_1, v_2, v_3) = v. \end{aligned}$$

Somit gilt $v \in T_pM$. □

Definition 8.17. Sei $f : U \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion, $M = f^{-1}(0) \subset \mathbb{R}^3$ eine reguläre Fläche und $p \in M$. Dann heißt die Ebene

$$\text{Tan}_pM := p + T_pM \subset \mathbb{R}^3$$

Tangentialebene an M im Punkt p und die Gerade

$$\text{Nor}_pM := p + \mathbb{R} \cdot \text{grad}f(p) \subset \mathbb{R}^3$$

die Normale an M im Punkt p .

Beispiel 1: *Das einschalige Hyperboloid*

Wir betrachten das einschalige Hyperboloid⁴

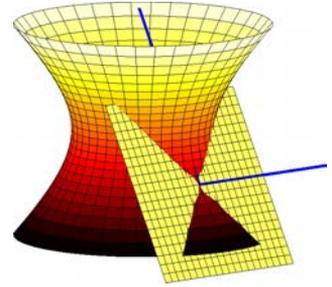
$$M := \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = z^2 + 1\}.$$

M ist die Nullstellenmenge der Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ mit

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

Da $\text{grad}f(x, y, z) = (2x, 2y, -2z) \neq 0$ für alle Punkte $p = (x, y, z) \in M$, ist M eine reguläre Fläche im \mathbb{R}^3 . Für den Tangentialraum in einem Punkt $p = (p_1, p_2, p_3) \in M$ erhalten wir:

$$T_p M = \{v \in \mathbb{R}^3 \mid \langle \text{grad}f(p), v \rangle = 0\} = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid p_1 v_1 + p_2 v_2 = p_3 v_3\}.$$

**Beispiel 2:** *Das Paraboloid*

Wir betrachten das Paraboloid

$$M := \{(x, y, z) \in \mathbb{R}^3 \mid z = x^2 + y^2 + 1\}.$$

M ist der Graph der Funktion $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$, gegeben durch

$$\varphi(x, y) := x^2 + y^2 + 1.$$

Wir können M als Nullstellenmenge der Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ mit

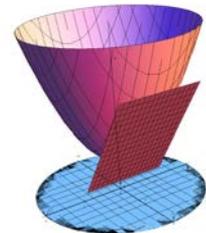
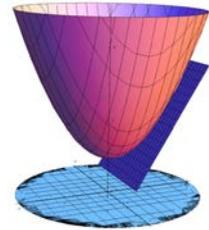
$$f(x, y, z) = x^2 + y^2 - z + 1$$

beschreiben. Für den Gradienten von f gilt

$$\text{grad}f(x, y, z) = (2x, 2y, -1) \neq 0.$$

M ist also eine reguläre Fläche im \mathbb{R}^3 und für den Tangentialraum im Punkt $p = (x, y, x^2 + y^2 + 1) \in M$ erhalten wir:

$$\begin{aligned} T_p M &= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid 2xv_1 + 2yv_2 - v_3 = 0\} \\ &= \{(v_1, v_2, 2xv_1 + 2yv_2) \mid v_1, v_2 \in \mathbb{R}\} \\ &= \mathbb{R} \cdot (1, 0, 2x) + \mathbb{R} \cdot (0, 1, 2y). \end{aligned}$$



⁴ Vielen Dank an Christoph Stadtmüller und Thomas Neukirchner für die Bilder

Integralrechnung für Funktionen mehrerer reeller Variablen (Ein Ausblick)

Wir wollen nun auch die Integralrechnung auf Funktionen mehrerer reeller Variablen verallgemeinern. Die Motivation zur Einführung des Integrals ist auch hier Frage, wie man Volumen von Teilmengen im \mathbb{R}^3 (oder im \mathbb{R}^n) vernünftig definieren kann und wie man dieses Volumen auch möglichst effektiv berechnen kann.

9.1 Das Problem der Volumendefinition für Teilmengen des \mathbb{R}^n

In Kapitel 6 hatten wir das Riemann-Integral von Funktionen einer Variablen benutzt, um den Flächeninhalt *gewisser* Teilmengen des \mathbb{R}^2 zu definieren und zu berechnen. Wir wollen uns nun zunächst mit der Frage beschäftigen, ob man für *jede* Teilmenge A des \mathbb{R}^n ein Volumen $\mu_n(A)$ definieren kann, so dass eine Reihe geometrisch vernünftiger Eigenschaften erfüllt sind? Für $n = 2$ soll μ_2 alle Flächeninhaltsformeln aus Kapitel 6 erfüllen. Insbesondere soll das Volumen $\mu_2(A)$ für eine klassische ebene Figur $A \subset \mathbb{R}^2$ (Dreieck, Trapez, Rechteck, ...) der aus der Schule bekannte elementargeometrische Flächeninhalt sein. Das Volumen $\mu_3(K)$ eines Körpers $K \subset \mathbb{R}^3$ (Kegel, Zylinder, Kugel, ...) soll mit dem aus der Schule bekannten elementargeometrischen Volumen übereinstimmen. Wir stellen an eine Volumenfunktion μ_n außerdem die folgenden vernünftigen Forderungen:

μ_n sei eine Funktion

$$\mu_n : A \subset \mathbb{R}^n \mapsto \mu_n(A) =: \text{''Volumen von } A\text{''} \in [0, \infty) \cup \{\infty\} =: [0, \infty]$$

mit den folgenden Eigenschaften:

- 1) Ist $A \subset B$, so gilt $\mu_n(A) \leq \mu_n(B)$ (*Monotonie*).
- 2) μ_n ist *translationsinvariant*, das heißt für $p \in \mathbb{R}^n$ gilt $\mu_n(A + p) = \mu_n(A)$.
- 3) Ist $W = [a_1, b_1] \times \dots \times [a_n, b_n]$ ein Quader im \mathbb{R}^n , so gilt $\mu_n(W) = \prod_{i=1}^n (b_i - a_i)$.
- 4) Sind A_1, A_2, \dots paarweise disjunkte Teilmengen, so gilt $\mu_n\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu_n(A_n)$ (*σ -Additivität*).

Satz 9.1 *Es existiert keine Funktion $\mu_n : \mathcal{P}(\mathbb{R}^n) \rightarrow [0, \infty]$ mit den Eigenschaften 1)–4).*

Beweis. Angenommen $\mu_n : \mathcal{P}(\mathbb{R}^n) \rightarrow [0, \infty]$ ist eine Funktion mit den Eigenschaften 1)–4). Wir konstruieren uns eine (etwas merkwürdige) Menge $B \subset \mathbb{R}^n$, die dies zum Widerspruch führt. Dazu betrachten wir auf dem Würfel $[0, 1]^n \subset \mathbb{R}^n$ folgende Äquivalenzrelation:

$$x \sim y \iff x - y \in \mathbb{Q}^n.$$

Sei $A \subset [0, 1]^n$ eine Teilmenge des Würfels, die aus jeder dieser Äquivalenzklassen genau ein Element enthält. Wir definieren die Menge B nun durch

$$B := \bigcup_{r \in [-1, 1]^n \cap \mathbb{Q}^n} A + r.$$

Dann gilt:

- a) B ist die Vereinigung abzählbar vieler Mengen, da \mathbb{Q}^n abzählbar ist.
 b) B ist eine disjunkte Vereinigung:

Seien $r, s \in [-1, 1]^n \cap \mathbb{Q}^n$ mit $r \neq s$. Angenommen es gäbe ein $b \in (A + r) \cap (A + s)$. Dann ist $b = a + r = a' + s$ für $a, a' \in A$. Daraus folgt $a - a' = s - r \in \mathbb{Q}^n$, folglich sind a und a' äquivalent und somit gleich (denn A enthält aus jeder Äquivalenzklasse genau ein Element). Dann wäre aber $r = s$ (Widerspruch). Für verschiedene r, s gilt also $(A + r) \cap (A + s) = \emptyset$.

- c) Nach Konstruktion gilt $[0, 1]^n \subset B \subset [-1, 2]^n$.

Wegen der Monotonie von μ_n und der Festlegung des Würfelvolumens gilt dann

$$1 = \mu_n([0, 1]^n) \leq \mu_n(B) \leq \mu_n([-1, 2]^n) = 3^n.$$

Wegen der σ -Additivität und der Translationsinvarianz erhält man

$$\mu_n(B) = \sum_{r \in \mathbb{Q}^n \cap [-1, 1]^n} \mu_n(A + r) = \sum_{r \in \mathbb{Q}^n \cap [-1, 1]^n} \mu_n(A).$$

Es folgt also

$$1 \leq \sum_{r \in \mathbb{Q}^n \cap [-1, 1]^n} \mu_n(A) \leq 3^n.$$

$\mu_n(A)$ wird dabei unendlich oft summiert. Dies geht aber nicht für eine Menge A mit $0 \leq \mu_n(A) \leq \infty$. Dies liefert den Widerspruch. \square

Ausweg: Man definiert ein n -dimensionales Volumen nicht für alle Teilmengen, sondern nur für eine bestimmte Auswahl von Teilmengen.

- 1) Das Jordan-Volumen: $\mu_n(A)$ wird für bestimmte *beschränkte* Mengen $A \subset \mathbb{R}^n$ definiert. Dies ist die historisch erste Variante zur Präzisierung des Volumenbegriffes und geometrisch leicht zugänglich. Alle uns in dieser Vorlesung interessierenden geometrisch 'vernünftigen' beschränkten Mengen gehören dazu.
- 2) Das Lebesgue-Maß: $\mu_n(A)$ wird für eine viel größere Klasse von Mengen definiert, z.B. auch für unbeschränkte Mengen. Das ist die modernere Variante. Die Definition ist technisch viel aufwändiger. Die zugehörige Integrationstheorie (Lebesgue-Integral) hat aber wesentlich bessere Eigenschaften als das Riemann-Integral).

9.2 Das Jordan–Volumen

Ähnlich wie beim Riemann-Integral in Kapitel 6, wollen wir die 'Größe' einer beschränkten Menge $A \subset \mathbb{R}^n$ dadurch messen, dass wir sie von Innen und von Außen durch kleine Quader approximieren und die 'Größe' von A durch das geometrische Volumen dieser approximierenden Quader angeben.

Das Volumen des Quaders $W = [a_1, b_1] \times \dots \times [a_n, b_n] \subset \mathbb{R}^n$ ist die Zahl

$$\text{vol}(W) := \prod_{i=1}^n (b_i - a_i).$$

Definition 9.1. Eine Zerlegung des Quaders W ist ein n -Tupel $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$ von Zerlegungen \mathcal{P}_k der Kanten $[a_k, b_k]$. Die Zerlegung $\mathcal{P}' = (\mathcal{P}'_1, \dots, \mathcal{P}'_n)$ von W heißt Verfeinerung der Zerlegung $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$, wenn \mathcal{P}'_k für jedes $k = 1, \dots, n$ eine Verfeinerung von \mathcal{P}_k ist (Bezeichnung: $\mathcal{P}' \geq \mathcal{P}$).

Sei $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$ eine Zerlegung von W mit $\mathcal{P}_k = \{x_{k0}, x_{k1}, \dots, x_{kr_k}\}$ wobei

$$a_k = x_{k0} < x_{k1} < \dots < x_{kr_k} = b_k.$$

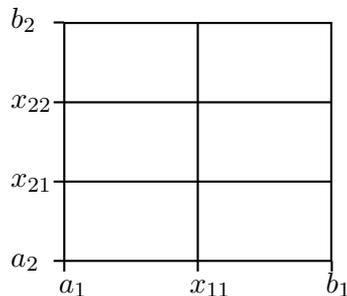
Durch

$$W_{\alpha_1 \dots \alpha_n} := [x_{1\alpha_1-1}, x_{1\alpha_1}] \times \dots \times [x_{n\alpha_n-1}, x_{n\alpha_n}]$$

für $1 \leq \alpha_i \leq r_i$ werden kleinere Quader definiert, die W zerlegen. Somit ergibt sich für den Quader W

$$W = \bigcup_{\alpha_1=1}^{r_1} \dots \bigcup_{\alpha_n=1}^{r_n} W_{\alpha_1 \dots \alpha_n},$$

$$\text{vol}(W) = \sum_{\alpha_1 \dots \alpha_n} \text{vol}(W_{\alpha_1 \dots \alpha_n}).$$



Definition 9.2. Sei $A \subset \mathbb{R}^n$ eine beschränkte Teilmenge, $\text{Int}(A)$ ihr Inneres und $\text{cl}(A)$ ihr Abschluß. Wir wählen einen Quader $W \subset \mathbb{R}^n$ mit $A \subset W$. Sei \mathcal{P} eine Zerlegung von W . Dann heißt

$$\overline{S}(A, W, \mathcal{P}) := \sum_{W_{\alpha_1 \dots \alpha_n} \cap \text{cl}(A) \neq \emptyset} \text{vol}(W_{\alpha_1 \dots \alpha_n}) \quad \text{Obersumme von } A \text{ bzgl. } \mathcal{P},$$

$$\underline{S}(A, W, \mathcal{P}) := \sum_{W_{\alpha_1 \dots \alpha_n} \subset \text{Int}(A)} \text{vol}(W_{\alpha_1 \dots \alpha_n}) \quad \text{Untersumme von } A \text{ bzgl. } \mathcal{P}.$$

Für zwei beliebige Zerlegungen \mathcal{P} und \mathcal{P}^* von W gilt

$$\underline{S}(A, W, \mathcal{P}) \leq \overline{S}(A, W, \mathcal{P}^*).$$

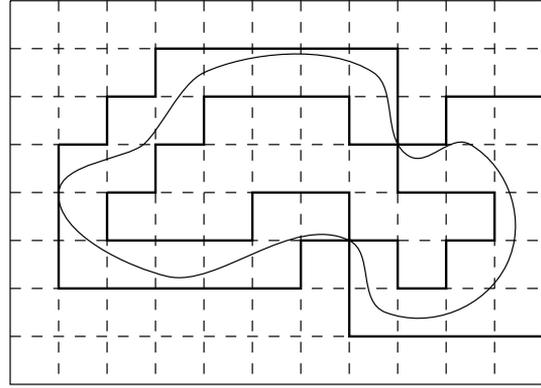
Definition 9.3. Die Zahl

$$\underline{\text{vol}}(A) := \sup \{ \underline{S}(A, W, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } W \}$$

heißt *inneres Volumen* von A , die Zahl

$$\overline{\text{vol}}(A) := \inf \{ \overline{S}(A, W, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } W \}$$

heißt *äußeres Volumen* von A .



Diese Definition ist unabhängig von der Wahl des Würfels $W \supset A$ und es gilt

$$0 \leq \underline{\text{vol}}(A) \leq \overline{\text{vol}}(A) < +\infty.$$

Definition 9.4. Eine Teilmenge $A \subset \mathbb{R}^n$ heißt *Jordan-messbar*, wenn sie beschränkt ist und $\underline{\text{vol}}(A) = \overline{\text{vol}}(A)$ gilt. In diesem Fall heißt

$$\text{vol}(A) := \underline{\text{vol}}(A) = \overline{\text{vol}}(A)$$

das n -dimensionale *Jordan-Volumen* (oder der n -dimensionale *Jordan-Inhalt*) von A . Eine Teilmenge $A \subset \mathbb{R}^n$ heißt *Jordansche Nullmenge*, wenn sie beschränkt ist und wenn $\overline{\text{vol}}(A) = 0$ gilt.

$\mathcal{J}(\mathbb{R}^n)$ bezeichnet die Menge der Jordan-messbaren Teilmengen des \mathbb{R}^n .

Im Fall $n = 2$ nennt man das Jordan-Volumen auch *Flächeninhalt* von A und bezeichnet es mit $\text{Area}(A)$. Im Fall $n = 3$ nennt man das Jordan-Volumen kurz *Volumen* von A .

Jordansche Nullmengen sind Jordan-messbar und haben das Jordan-Volumen 0. Jede Teilmenge einer Jordanschen Nullmenge ist selbst eine Jordansche Nullmenge. Der folgende Satz bietet ein gut handhabbares Kriterium, um Jordan-Messbarkeit nachzuweisen:

Satz 9.2 (Kriterium für Jordan-Messbarkeit)

Sei $A \subset \mathbb{R}^n$ eine beschränkte Menge und ∂A ihr Rand. Dann gilt

$$\overline{\text{vol}}(A) - \underline{\text{vol}}(A) = \overline{\text{vol}}(\partial A).$$

Insbesondere ist eine beschränkte Menge $A \subset \mathbb{R}^n$ genau dann Jordan-messbar, wenn ihr Rand ∂A eine Jordansche Nullmenge ist.

Beispiele für Jordan-messbare Mengen

- Jeder Quader $W = [a_1, b_1] \times \dots \times [a_n, b_n] \subset \mathbb{R}^n$ ist Jordan-messbar mit den Jordan-Volumen $\text{vol}(W) = (b_1 - a_1) \cdot \dots \cdot (b_n - a_n)$.
- Die Menge $[0, 1]^2 \cap \mathbb{Q}^2 \subset \mathbb{R}^2$ ist nicht Jordan-messbar.
- Sei $\varphi : K \subset \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ eine stetige Funktion auf einer kompakten Menge $K \subset \mathbb{R}^{n-1}$. Dann ist der Graph $\Gamma_\varphi := \{(x, \varphi(x)) \mid x \in K\} \subset \mathbb{R}^n$ eine Jordansche Nullmenge.

- d) Jedes beschränkte Gebiet $A \subset \mathbb{R}^2$, dessen Rand ∂A aus Stücken besteht, die man als Graphen stetiger Funktionen $\varphi : [a, b] \rightarrow \mathbb{R}$ darstellen kann, ist Jordan-meßbar. Z.B. ist jedes Gebiet Ω_f unter dem Graph einer stetigen nicht-negativen Funktion $f : [a, b] \rightarrow \mathbb{R}$ Jordan-meßbar (solche Gebiete haben wir in Kapitel 6.1 betrachtet).
- e) Jedes beschränkte Gebiet $\Omega \subset \mathbb{R}^2$, dessen Rand $\Gamma = \partial\Omega$ man durch eine einfache stückweise stetig differenzierbare Kurve parametrisieren kann, ist Jordan-meßbar (solche Gebiete haben wir in Kapitel 6.4.2 betrachtet).
- f) Jede beschränkte Teilmenge $B \subset \mathbb{R}^3$, deren Rand ∂B aus Graphen von stetigen Funktionen $\varphi : K \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ auf kompakten Mengen K besteht, ist Jordan-meßbar. Solche Ränder liegen z.B. oft als gleichungsdefinierte Flächen vor (siehe Kapitel 8.7.).

Mit dem Kriterium aus Satz 9.2 kann man auch die folgenden Eigenschaften Jordan-meßbarer Mengen beweisen:

Satz 9.3 1. Ist $A \subset \mathbb{R}^n$ Jordan-meßbar, so sind auch das Innere $\text{Int}(A)$, der Abschluß $\text{cl}(A)$ und der Rand ∂A Jordan-meßbar und es gilt

$$\text{vol}(\text{Int}(A)) = \text{vol}(A) = \text{vol}(\text{cl}(A)) \quad \text{und} \quad \text{vol}(\partial A) = 0.$$

2. Sind A und B Jordan-meßbar, so sind auch $A \cup B$, $A \cap B$ und $A \setminus B$ Jordan-meßbar.

3. Das Jordan-Volumen definiert eine Volumenfunktion

$$\mu_n : A \in \mathcal{J}(\mathbb{R}^n) \mapsto \text{vol}(A) \in [0, \infty)$$

mit den Eigenschaften 1)–4) aus Abschnitt 9.1.

Im nächsten Abschnitt werden wir das Riemann-Integral für Funktionen mehrerer reeller Variablen definieren. Es wird uns Methoden liefern, um das Jordan-Volumen zu berechnen.

9.3 Das Riemann-Integral für Funktionen mehrerer reeller Variablen

Die Definition des Riemann-Integrals im Mehrdimensionalen verläuft völlig analog zum 1-dimensionalen Fall (siehe Kapitel 6).

Sei $W \subset \mathbb{R}^n$ ein Quader und $f : W \rightarrow \mathbb{R}$ eine beschränkte Funktion. Wir betrachten eine Zerlegung \mathcal{P} von W mit den Teilquadern $W_{\alpha_1 \dots \alpha_n}$.

Als *Obersumme von f bezüglich \mathcal{P}* bezeichnet man die Zahl

$$\bar{S}(f, \mathcal{P}) := \sum_{W_{\alpha_1 \dots \alpha_n}} \sup(f|_{W_{\alpha_1 \dots \alpha_n}}) \cdot \text{vol}(W_{\alpha_1 \dots \alpha_n}).$$

Als *Untersumme von f bezüglich \mathcal{P}* bezeichnet man

$$\underline{S}(f, \mathcal{P}) := \sum_{W_{\alpha_1 \dots \alpha_n}} \inf(f|_{W_{\alpha_1 \dots \alpha_n}}) \cdot \text{vol}(W_{\alpha_1 \dots \alpha_n}).$$

Für zwei beliebige Zerlegungen \mathcal{P} und \mathcal{P}^* von W gilt wieder

$$\underline{S}(f, \mathcal{P}) \leq \overline{S}(f, \mathcal{P}^*).$$

Wir erhalten also

$$\begin{aligned} \sup_{\mathcal{P}} \underline{S}(f, \mathcal{P}) &:= \sup \{ \underline{S}(f, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } W \} \\ &\leq \inf \{ \overline{S}(f, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } W \} \\ &=: \inf_{\mathcal{P}} \overline{S}(f, \mathcal{P}). \end{aligned}$$

Definition 9.5. Eine Funktion $f : W \rightarrow \mathbb{R}$ heißt Riemann-integrierbar, wenn sie beschränkt ist und $\sup_{\mathcal{P}} \underline{S}(f, \mathcal{P}) = \inf_{\mathcal{P}} \overline{S}(f, \mathcal{P})$ gilt.

In diesem Fall heißt die Zahl

$$\int_W f(x) dx := \sup_{\mathcal{P}} \underline{S}(f, \mathcal{P}) = \inf_{\mathcal{P}} \overline{S}(f, \mathcal{P})$$

das Riemann-Integral von f über W . (Dabei ist $x = (x_1, \dots, x_n) \in \mathbb{R}^n$).

Auch im Mehrdimensionalen gilt:

Satz 9.4 (Riemannsches Integrierbarkeitskriterium)

Sei W ein Quader im \mathbb{R}^n . Eine Funktion $f : W \rightarrow \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn sie beschränkt ist und für jedes $\varepsilon > 0$ eine Zerlegung \mathcal{P} von W existiert, so daß

$$\overline{S}(f, \mathcal{P}) - \underline{S}(f, \mathcal{P}) < \varepsilon.$$

Wie im 1-dimensionalen Fall beweist man folgende Eigenschaften des Riemann-Integrals:

Satz 9.5 Seien $f, g : W \subset \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integrierbare Funktionen auf einem Quader W . Dann gilt

1. Für reelle Zahlen λ und μ ist die Funktion $\lambda f + \mu g$ ebenfalls Riemann-integrierbar und

$$\int_W (\lambda \cdot f(x) + \mu \cdot g(x)) dx = \lambda \cdot \int_W f(x) dx + \mu \cdot \int_W g(x) dx.$$

2. Das Produkt $f \cdot g : W \rightarrow \mathbb{R}$ ist Riemann-integrierbar.
3. Der Betrag $|f| : W \rightarrow [0, \infty)$ ist Riemann-integrierbar und

$$\left| \int_W f(x) dx \right| \leq \int_W |f(x)| dx.$$

4. Ist $f \leq g$, so gilt $\int_W f(x) dx \leq \int_W g(x) dx$.

Im folgenden Satz sind die drei wichtigsten Beispiele für Riemann-integrierbare Funktionen zusammengestellt, die wir im Folgenden benutzen werden.

Satz 9.6 Sei $W \subset \mathbb{R}^n$ ein Quader.

1. Jede stetige Funktion $f : W \rightarrow \mathbb{R}$ ist Riemann-integrierbar.
2. Sei $f : W \rightarrow \mathbb{R}$ beschränkt und die Menge der Unstetigkeitsstellen von f eine Jordansche Nullmenge. Dann ist f Riemann-integrierbar.
3. Sei $A \subset \mathbb{R}^n$ eine Jordan-meßbare Menge und W ein Quader im \mathbb{R}^n , der A enthält. Mit $\chi_A : W \rightarrow \mathbb{R}$ bezeichnen wir die charakteristische Funktion von A , das heißt die Funktion

$$\chi_A(x) := \begin{cases} 1 & \text{falls } x \in A, \\ 0 & \text{falls } x \in W \setminus A. \end{cases}$$

Dann ist χ_A Riemann-integrierbar und für das Jordan-Volumen von A gilt

$$\text{vol}(A) = \int_W \chi_A(x) dx.$$

Beweis. Wir beweisen 3). Sei \mathcal{P} eine Zerlegung von W . Dann gilt

$$\begin{aligned} \overline{S}(\chi_A, \mathcal{P}) &= \sum_{W^* \in \mathcal{P}} \sup(\chi_A|_{W^*}) \cdot \text{vol}(W^*) = \sum_{W^* \cap A \neq \emptyset} \text{vol}(W^*) \\ &\leq \sum_{W^* \cap \text{cl}(A) \neq \emptyset} \text{vol}(W^*) = \overline{S}(A, W, \mathcal{P}), \\ \underline{S}(\chi_A, \mathcal{P}) &= \sum_{W^* \in \mathcal{P}} \inf(\chi_A|_{W^*}) \cdot \text{vol}(W^*) = \sum_{W^* \subset A} \text{vol}(W^*) \\ &\geq \sum_{W^* \subset \text{Int}(A)} \text{vol}(W^*) = \underline{S}(A, W, \mathcal{P}). \end{aligned}$$

Somit gilt

$$\underline{S}(A, W, \mathcal{P}) \leq \underline{S}(\chi_A, \mathcal{P}) \leq \overline{S}(\chi_A, \mathcal{P}) \leq \overline{S}(A, W, \mathcal{P})$$

und folglich

$$\underline{\text{vol}}(A) = \sup_{\mathcal{P}} \underline{S}(A, W, \mathcal{P}) \leq \sup_{\mathcal{P}} \underline{S}(\chi_A, \mathcal{P}) \leq \inf_{\mathcal{P}} \overline{S}(\chi_A, \mathcal{P}) \leq \inf_{\mathcal{P}} \overline{S}(A, W, \mathcal{P}) = \overline{\text{vol}}(A). \quad (9.1)$$

Da A Jordan-meßbar ist, gilt $\underline{\text{vol}}(A) = \overline{\text{vol}}(A) = \text{vol}(A)$. Folglich muß in (9.1) überall Gleichheit stehen. D.h. χ_A ist Riemann-integrierbar und es gilt

$$\text{vol}(A) = \int_W \chi_A(x) dx. \quad \square$$

Wir wollen jetzt Funktionen über Jordan-meßbaren Mengen $A \subset \mathbb{R}^n$ integrieren.

Ist W ein Quader, der A enthält, und $f : W \rightarrow \mathbb{R}$ Riemann-integrierbar, so ist auch das Produkt $f \cdot \chi_A : W \rightarrow \mathbb{R}$ Riemann-integrierbar (Satz 9.5 und Satz 9.6). Wir schneiden

f hierbei außerhalb von A ab und ersetzen die Funktionswerte dort durch Null. Auf A stimmen f und $f \cdot \chi_A$ überein.

Haben wir eine nur auf A definierte Funktion $f : A \rightarrow \mathbb{R}$ gegeben, so setzen wir f zunächst auf W durch Null fort, d.h. wir definieren zusätzlich $f(x) := 0$ für alle $x \in W \setminus A$. Ist f z.B. auf A stetig, so liegen die Unstetigkeitsstellen der Fortsetzung auf dem Rand ∂A . Da A Jordan-meßbar ist, ist der Rand ∂A eine Jordansche Nullmenge und deshalb 'klein' genug, um die Riemann-Integrierbarkeit der Fortsetzung zu sichern (siehe Satz 9.6).

Definition 9.6. Sei $A \subset W$ Jordan-meßbar und $f : W \rightarrow \mathbb{R}$ Riemann-integrierbar. Unter dem Riemann-Integral von f über A verstehen wir die Zahl

$$\int_A f(x) dx := \int_W (f \cdot \chi_A)(x) dx.$$

Insbesondere gilt für das Jordan-Volumen einer Jordan-meßbaren Menge $A \subset \mathbb{R}^n$

$$\text{vol}(A) = \int_A 1 dx.$$

Die Rechenregeln von Satz 9.5 gelten auch für das Integral über eine Jordan-meßbare Menge A . Außerdem erhält man durch Einsetzen der Definition:

Satz 9.7 Sei $f : W \rightarrow \mathbb{R}$ Riemann-integrierbar, $A, B \subset W$ Jordan-meßbar und disjunkt und $N \subset W$ eine Jordansche Nullmenge. Dann gilt

$$\begin{aligned} \int_{A \cup B} f(x) dx &= \int_A f(x) dx + \int_B f(x) dx, \\ \int_A f(x) dx &= \int_{A \setminus N} f(x) dx, \\ \int_N f(x) dx &= 0. \end{aligned}$$

Völlig analog man zum 1-dimensionalen Fall kann man den Mittelwertsatz beweisen.

Satz 9.8 (Mittelwertsatz)

Sei $f : W \rightarrow \mathbb{R}$ stetig, $g : W \rightarrow \mathbb{R}$ Riemann-integrierbar und $g \geq 0$. Dann existiert ein $\xi \in W$ mit

$$\int_W f(x)g(x) dx = f(\xi) \cdot \int_W g(x) dx.$$

Insbesondere existiert ein $\xi_0 \in W$ mit

$$\int_W f(x) dx = f(\xi_0) \cdot \text{vol}(W).$$

In den folgenden beiden Abschnitten werden wir Methoden kennenlernen, mit denen man die Berechnung des mehrdimensionalen Riemann-Integrals auf die Berechnung des Integrals im 1-dimensionalen Fall zurückführen kann. Die Methoden für Letzteres haben wir in Kapitel 6 behandelt.

9.4 Der Satz von Fubini und das Prinzip von Cavalieri

Der zentrale Satz der mehr-dimensionalen Integralrechnung ist der Satz von Fubini. Mit diesem Satz führt man die Integration über einem Quader $W_1 \times W_2 \subset \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$ auf die Integration über die Faktoren W_1 und W_2 zurück. Dies erlaubt es dann, die Integration auf die 1-dimensionalen Riemann-Integrale zurückzuführen, die wir aus Kapitel 6 kennen.

Satz 9.9 (Satz von Fubini)

Seien $W_1 \subset \mathbb{R}^n$ und $W_2 \subset \mathbb{R}^m$ zwei Quader, $W := W_1 \times W_2$ ihr Produkt im \mathbb{R}^{n+m} und $f : W_1 \times W_2 \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion.

1. Sei $x \in W_1$ fixiert und $g_x : W_2 \rightarrow \mathbb{R}$ definiert durch $g_x(y) := f(x, y)$. Wir setzen voraus, dass die Funktion g_x für alle $x \in W_1$ Riemann-integrierbar ist. Dann ist die Funktion $G : W_1 \rightarrow \mathbb{R}$, definiert durch die partielle Integration über W_2

$$G(x) := \int_{W_2} g_x(y) dy = \int_{W_2} f(x, y) dy,$$

Riemann-integrierbar und es gilt

$$\int_{W_1 \times W_2} f(x, y) d(x, y) = \int_{W_1} G(x) dx = \int_{W_1} \left(\int_{W_2} f(x, y) dy \right) dx.$$

2. Sei $y \in W_2$ fixiert und $h_y : W_1 \rightarrow \mathbb{R}$ definiert durch $h_y(x) := f(x, y)$. Wir setzen voraus, dass die Funktion h_y für alle $y \in W_2$ Riemann-integrierbar ist. Dann ist die Funktion $H : W_2 \rightarrow \mathbb{R}$, definiert durch die partielle Integration über W_1

$$H(y) := \int_{W_1} h_x(y) dx = \int_{W_1} f(x, y) dx,$$

Riemann-integrierbar und es gilt

$$\int_{W_1 \times W_2} f(x, y) d(x, y) = \int_{W_2} H(y) dy = \int_{W_2} \left(\int_{W_1} f(x, y) dx \right) dy.$$

Bemerkung: Ist $f : W_1 \times W_2 \rightarrow \mathbb{R}$ Riemann-integrierbar, so müssen die Funktionen $g_x : W_2 \rightarrow \mathbb{R}$ und $h_y : W_1 \rightarrow \mathbb{R}$ nicht unbedingt Riemann-integrierbar sein. Dies muß man deshalb zusätzlich voraussetzen.

Ist $f : W_1 \times W_2 \rightarrow \mathbb{R}$ stetig, so sind $g_x : W_2 \rightarrow \mathbb{R}$ und $h_y : W_1 \rightarrow \mathbb{R}$ ebenfalls stetig und folglich auch Riemann-integrierbar. In diesem Fall können wir den Satz von Fubini ohne zusätzliche Voraussetzungen über g_x und h_y anwenden und erhalten die Formel für die iterierte Integration

$$\int_{W_1 \times W_2} f(x, y) d(x, y) = \int_{W_1} \left(\int_{W_2} f(x, y) dy \right) dx = \int_{W_2} \left(\int_{W_1} f(x, y) dx \right) dy.$$

Beweis des Satzes von Fubini. Sei $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$ eine Zerlegung von W_1 mit den Teilwürfeln W^* , $\tilde{\mathcal{P}} = (\tilde{\mathcal{P}}_1, \dots, \tilde{\mathcal{P}}_m)$ eine Zerlegung von W_2 mit den Teilwürfeln W^{**} . Dann ist $\mathcal{Q} := (\mathcal{P}, \tilde{\mathcal{P}})$ eine Zerlegung von $W_1 \times W_2$ mit den Teilwürfeln $W^* \times W^{**}$ und es gilt

$$\underline{S}(f, W_1 \times W_2, \mathcal{Q}) = \sum_{W^* \in \mathcal{P}, W^{**} \in \tilde{\mathcal{P}}} \inf(f|_{W^* \times W^{**}}) \cdot \text{vol}(W^*) \cdot \text{vol}(W^{**}).$$

Ist $x \in W^*$, so gilt $\inf(f|_{W^* \times W^{**}}) \leq \inf(g_x|_{W^{**}})$. Für festes W^* und $x \in W^*$ erhalten wir also

$$\begin{aligned} \sum_{W^{**} \in \tilde{\mathcal{P}}} \inf(f|_{W^* \times W^{**}}) \cdot \text{vol}(W^{**}) &\leq \sum_{W^{**} \in \tilde{\mathcal{P}}} \inf(g_x|_{W^{**}}) \cdot \text{vol}(W^{**}) = \underline{S}(g_x, W_2, \tilde{\mathcal{P}}) \\ &\leq \int_{W_2} g_x(y) dy = G(x). \end{aligned}$$

Daraus folgt

$$\sum_{W^{**} \in \tilde{\mathcal{P}}} \inf(f|_{W^* \times W^{**}}) \cdot \text{vol}(W^{**}) \leq \inf\{G(x) \mid x \in W^*\} = \inf(G|_{W^*}).$$

Somit ist

$$\underline{S}(f, W_1 \times W_2, \mathcal{Q}) \leq \sum_{W^* \in \mathcal{P}} \inf(G|_{W^*}) \cdot \text{vol}(W^*) = \underline{S}(G, W_1, \mathcal{P}). \quad (*)$$

Analog zeigt man,

$$\overline{S}(f, W_1 \times W_2, \mathcal{Q}) \geq \overline{S}(G, W_1, \mathcal{P}). \quad (**)$$

Daraus erhalten wir

$$\underline{S}(f, W_1 \times W_2, \mathcal{Q}) \stackrel{(*)}{\leq} \underline{S}(G, W_1, \mathcal{P}) \leq \overline{S}(G, W_1, \mathcal{P}) \stackrel{(**)}{\leq} \overline{S}(f, W_1 \times W_2, \mathcal{Q}).$$

Da f auf $W_1 \times W_2$ integrierbar ist, folgt

$$\int_{W_1 \times W_2} f(x, y) d(x, y) = \sup_{\mathcal{Q}} \underline{S}(f, W_1 \times W_2, \mathcal{Q}) = \inf_{\mathcal{Q}} \overline{S}(f, W_1 \times W_2, \mathcal{Q}).$$

Folglich gilt

$$\sup \underline{S}(G, W_1, \mathcal{P}) = \inf \overline{S}(G, W_1, \mathcal{P}) = \int_{W_1 \times W_2} f(x, y) d(x, y).$$

$G : W_1 \rightarrow \mathbb{R}$ ist also Riemann-integrierbar und es gilt

$$\int_{W_1} \left(\int_{W_2} f(x, y) dy \right) dx = \int_{W_1 \times W_2} f(x, y) d(x, y).$$

Offensichtlich gilt der gleiche Beweis, wenn man die Rollen von W_1 und W_2 vertauscht. Dies beweist dann den 2. Teil von Satz 9.9. \square

Folgerung 9.1 (Iterierte Integration stetiger Funktionen auf Quadern)

Sei $f : W = [a_1, b_1] \times \dots \times [a_n, b_n] \rightarrow \mathbb{R}$ eine stetige Funktion. Dann kann man das Riemann-Integral von f über W durch schrittweise Integration nach jeweils einer Variablen berechnen und das Ergebnis hängt nicht von der gewählten Reihenfolge der Integration ab. D.h. für jede Permutation σ von $\{1, 2, \dots, n\}$ gilt:

$$\begin{aligned} \int_W f(x) dx &= \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \dots \int_{a_{n-1}}^{b_{n-1}} \left(\int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \right) dx_{n-1} \dots \right) dx_1 \\ &= \int_{a_{\sigma(1)}}^{b_{\sigma(1)}} \left(\int_{a_{\sigma(2)}}^{b_{\sigma(2)}} \dots \int_{a_{\sigma(n-1)}}^{b_{\sigma(n-1)}} \left(\int_{a_{\sigma(n)}}^{b_{\sigma(n)}} f(x_1, \dots, x_n) dx_{\sigma(n)} \right) dx_{\sigma(n-1)} \dots \right) dx_{\sigma(1)}. \end{aligned}$$

Diese Rechenregel kann man auf die Integration über meßbaren Mengen $A \subset \mathbb{R}^n$ übertragen. Dazu schneidet man A mit Hyperebenen, die parallel zu den Koordinatenebenen sind.

Definition 9.7. Sei $A \subset \mathbb{R}^n$ und $c \in \mathbb{R}$ eine fixierte Zahl. Die Menge

$$A_c := \{(x_1, \dots, x_{n-1}) \mid (x_1, \dots, x_{n-1}, c) \in A\} = \text{proj}_{\mathbb{R}^{n-1}} (A \cap \{x_n = c\}) \subset \mathbb{R}^{n-1}.$$

heißt c -Schnitt von A entlang der n -ten Koordinate. Analog definiert man den c -Schnitt entlang der anderen Koordinaten.

Folgerung 9.2 (Iterierte Integration stetiger Funktionen auf Jordan-Mengen)

Sei $A \subset \mathbb{R}^n$ eine Jordan-meßbare Menge, $c \in \mathbb{R}$ und sei der c -Schnitt

$$A_c := \{(x_1, \dots, x_{n-1}) \mid (x_1, \dots, x_{n-1}, c) \in A\} \subset \mathbb{R}^{n-1}$$

Jordan-meßbar für alle $c \in \mathbb{R}$. Wir wählen ein Intervall $[c_1, c_2]$ mit $A_c = \emptyset$ für alle $c \notin [c_1, c_2]$. Sei $f : A \rightarrow \mathbb{R}$ eine stetige Funktion. Dann gilt

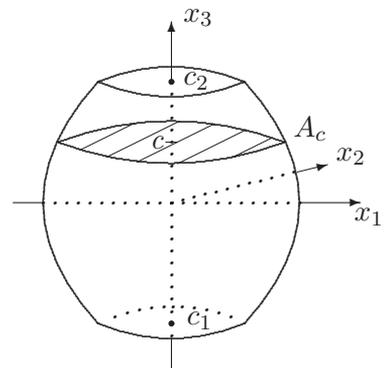
$$\int_A f(x) dx = \int_{c_1}^{c_2} \left(\int_{A_c} f(x_1, \dots, x_{n-1}, c) d(x_1, \dots, x_{n-1}) \right) dc.$$

Die analoge Formel gilt für die Schnitte entlang der anderen Koordinaten.

Beweis. Da A beschränkt ist, gibt es einen Quader $W := W_1 \times [c_1, c_2]$, der A enthält. Nach Definition des Riemann-Integrals von f über A gilt

$$\int_A f(x) dx = \int_W \chi_A(x) f(x) dx.$$

Mit Satz dem Satz von Fubini folgt dann



$$\int_A f(x) dx = \int_{W_1 \times [c_1, c_2]} \chi_A(x) f(x) dx = \int_{c_1}^{c_2} \left(\int_{W_1} (\chi_A(y, c) \cdot f(y, c)) dy \right) dc,$$

wobei $y = (x_1, \dots, x_{n-1})$. Es gilt $y \in A_c$ genau dann, wenn $(y, c) \in A$. Folglich ist $\chi_A(y, c) = \chi_{A_c}(y)$ und wir erhalten

$$\int_A f(x) dx = \int_{c_1}^{c_2} \left(\int_{W_1} \chi_{A_c}(y) \cdot f(y, c) dy \right) dc = \int_{c_1}^{c_2} \left(\int_{A_c} f(y, c) dy \right) dc$$

□

Setzt man der letzten Folgerung $f = 1$ ein, so erhält man eine Methode zur Berechnung des n -dimensionalen Jordan-Volumens $vol_n(A)$ von $A \in \mathcal{J}(\mathbb{R}^n)$ durch die $(n-1)$ -dimensionalen Jordan-Volumen $vol_{n-1}(A_c)$ der Schnittmengen $A_c \subset \mathbb{R}^{n-1}$.

Folgerung 9.3 (Das Prinzip von Cavalieri zur Volumenberechnung)

Sei $A \subset \mathbb{R}^n$ eine Jordan-messbare Menge und seien alle c -Schnitte $A_c \subset \mathbb{R}^{n-1}$ nach einer fixierten Variablen ebenfalls Jordan-messbar. Dann gilt

$$vol_n(A) = \int_{c_1}^{c_2} vol_{n-1}(A_c) dc,$$

wobei $[c_1, c_2]$ so gewählt ist, dass $A_c = \emptyset$ für alle $c \notin [c_1, c_2]$.

Beispiel 1: Sei $f : W = [a_1, b_1] \times [a_2, b_2] \rightarrow \mathbb{R}$ eine stetige Funktion der Form

$$f(x, y) = h_1(x) \cdot h_2(y).$$

Dann gilt

$$\int_W f(x, y) d(x, y) = \int_{a_1}^{b_1} h_1(x) dx \cdot \int_{a_2}^{b_2} h_2(y) dy.$$

Wir benutzen zum Beweis die Iterationsformel und erhalten

$$\begin{aligned} \int_W f(x, y) d(x, y) &= \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} h_1(x) \cdot h_2(y) dy \right) dx = \int_{a_1}^{b_1} h_1(x) \cdot \left(\int_{a_2}^{b_2} h_2(y) dy \right) dx \\ &= \int_{a_1}^{b_1} h_1(x) dx \cdot \int_{a_2}^{b_2} h_2(y) dy. \end{aligned}$$

Eine analoge Formel gilt für drei oder mehr Variablen.

Beispiel 2: Wir betrachten die Funktion $f : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$,

$$f(x, y) := x^2 + y^4,$$

und berechnen das Integral von f über $[0, 1] \times [0, 1]$:

$$\begin{aligned} \int_{[0,1] \times [0,1]} f(x, y) d(x, y) &= \int_0^1 \left(\int_0^1 (x^2 + y^4) dx \right) dy = \int_0^1 \left(\left. \left(\frac{1}{3}x^3 + y^4x \right) \right|_0^1 \right) dy \\ &= \int_0^1 \left(\frac{1}{3} + y^4 \right) dy = \frac{1}{3} + \frac{1}{5} = \frac{8}{15}. \end{aligned}$$

Beispiel 3: Wir betrachten die Kreisscheibe $D := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ und berechnen das Integral der Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, mit $f(x, y) := x + y$, über D .

Der Rand von D , d.h. die Menge $\partial D = \{(x, y) \mid f(x, y) = x^2 + y^2 = 1\} = S^1$, ist eine Jordansche Nullmenge (man kann sie als Graphen von Funktionen darstellen). Also ist die Kreisscheibe D Jordan-meßbar und wir können rechnen

$$\int_D f(x, y) d(x, y) = \int_{-1}^1 \left(\int_{D_c} f(x, c) dx \right) dc.$$

Da

$$D_c = \{x \in \mathbb{R} \mid x^2 + c^2 \leq 1\} = [-\sqrt{1-c^2}, \sqrt{1-c^2}]$$

folgt

$$\begin{aligned} \int_D f(x, y) d(x, y) &= \int_{-1}^1 \left(\int_{-\sqrt{1-c^2}}^{\sqrt{1-c^2}} (x + c) dx \right) dc = \int_{-1}^1 \left(\frac{1}{2}x^2 + cx \right) \Big|_{-\sqrt{1-c^2}}^{\sqrt{1-c^2}} dc \\ &= \int_{-1}^1 2c\sqrt{1-c^2} dc = 0. \quad (\text{ungerade Funktion}) \end{aligned}$$

Beispiel 4: Das Volumen von Mengen unter Funktionsgraphen.

Sei $A \subset \mathbb{R}^n$ Jordan-meßbar, $f : A \rightarrow [0, \infty)$ stetig und

$$\Omega_f := \{(x, y) \in A \times \mathbb{R} \mid 0 \leq y \leq f(x)\}$$

die Menge unter dem Funktionsgraphen von f . Die Menge Ω_f ist Jordan-meßbar, da der Rand von Ω_f aus Graphen stetiger Funktionen besteht. Für das Jordan-Volumen von Ω_f gilt:

$$\text{vol}(\Omega_f) = \int_A f(x) dx.$$

Ist $A = [a, b] \subset \mathbb{R}$, so entspricht dies dem Flächeninhalt von Ω_f , den wir in Kapitel 6 definiert haben.

Zur Berechnung des Integrals wählen wir einen Quader $W_1 \subset \mathbb{R}^n$, der A enthält. Nach Definition ist dann

$$\text{vol}(\Omega_f) = \int_{W_1 \times [0, \sup f]} \chi_{\Omega_f}(x, y) d(x, y) = \int_{W_1} \left(\int_0^{\sup f} \chi_{\Omega_f}(x, y) dy \right) dx.$$

Es gilt

$$\chi_{\Omega_f}(x, y) = \begin{cases} \chi_A(x) & \text{falls } y \in [0, f(x)], \\ 0 & \text{sonst.} \end{cases}$$

Daraus folgt

$$\text{vol}(\Omega_f) = \int_{W_1} \left(\int_0^{f(x)} \underbrace{\chi_A(x)}_{\text{konst.}} dy \right) dx = \int_{W_1} \chi_A(x) \cdot f(x) dx = \int_A f(x) dx.$$

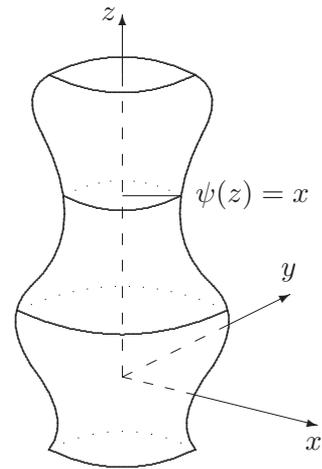
Beispiel 5: Das Volumen von Rotationskörpern im \mathbb{R}^3 .

Sei $\psi : [a, b] \rightarrow \mathbb{R}$ eine nicht-negative stetige Funktion. Wir betrachten in der zx -Ebene den Graphen der Funktion $x = \psi(z)$ und rotieren diese Kurve um die z -Achse. Die Menge V_ψ , die von der dabei entstehenden Fläche eingeschlossen wird, heißt der von ψ erzeugte *Rotationskörper*. Sie ist beschrieben durch

$$V_\psi := \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 \leq \psi(z)^2\}.$$

Der Rotationskörper V_ψ ist Jordan-meßbar, da sein Rand eine reguläre Fläche, d.h. eine Jordansche Nullmenge ist. Dann gilt für das Volumen von V_ψ

$$\text{vol}(V_\psi) = \pi \cdot \int_a^b \psi^2(z) dz.$$



Zum Beweis betrachten wir die Schnitte von V_ψ nach der 3. Koordinate. Der z -Schnitt von V_ψ ist die Kreisscheibe $D_{\psi(z)}^2$ vom Radius $\psi(z)$. Eine Kreisscheibe D_r^2 vom Radius r hat das Jordan-Volumen (d.h. in diesem Fall den Flächeninhalt) $\text{Area}(D_r^2) = \pi r^2$. Aus dem Prinzip von Cavalieri folgt damit

$$\text{vol}(V_\psi) = \int_a^b \text{Area}(D_{\psi(z)}^2) dz = \pi \cdot \int_a^b \psi^2(z) dz.$$

Beispiel 6: Das Volumen der Kugel im \mathbb{R}^3 vom Radius R .

Wir bezeichnen die abgeschlossene Kugel im \mathbb{R}^3 vom Radius R mit D_R^3 , d.h.

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

Für das Volumen von D_R^3 gilt

$$\text{vol}(D_R^3) = \frac{4}{3}\pi R^3.$$

Zum Beweis betrachten wir die Kugel als Rotationskörper. Die rotierende Funktion in der zx -Ebene ist hierbei $\psi : [-R, R] \rightarrow \mathbb{R}$ mit

$$\psi(z) = \sqrt{R^2 - z^2}.$$

Folglich gilt nach Beispiel 5:

$$\text{vol}(D_R^3) = \pi \int_{-R}^R (R^2 - z^2) dz = \pi \left(R^2 z \Big|_{-R}^R - \frac{1}{3} z^3 \Big|_{-R}^R \right) = \pi \left(2R^3 - \frac{2}{3}R^3 \right) = \frac{4}{3}\pi R^3.$$

9.5 Die Transformationsformel für Riemann-Integrale

Die zweite wichtige Formel zur Integralberechnung ist die Transformationsformel. Mit dieser Formel kann man das Riemann-Integral über einer Jordan-meßbaren Menge B , die "kompliziert" aussieht, auf ein Riemann-Integral über einer Menge A zurückführen, die "einfacher" aussieht (die z.B. ein Quader ist). Dazu benutzt man eine Koordinatentransformation, die die Menge B durch Koordinaten aus der einfacheren Menge A beschreibt.

Satz 9.10 Sei $\phi : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^n$ ein Diffeomorphismus (d.h. eine Koordinatentransformation) zwischen den offenen Mengen U und V , $f : V \rightarrow \mathbb{R}$ eine stetige Funktion und $A \subset U$ Jordan-meßbar. Dann ist $\phi(A)$ ebenfalls Jordan-meßbar und es gilt

$$\int_{\phi(A)} f(y) dy = \int_A f(\phi(x)) \cdot |\det J\phi_x| dx.$$

Diese Formel verallgemeinert die Substitutionsregel für Riemann-Integrale von Funktionen einer Variablen: Sei $\phi : [a, b] \rightarrow [c, d]$ ein Diffeomorphismus. In diesem Fall ist die Jacobi-Matrix $J\phi_x = \phi'(x)$. Wir führen die Substitution $y = \phi(x)$, $dy = \phi'(x) dx$ durch. Dann gilt

$$\int_c^d f(y) dy = \int_{\phi^{-1}(c)}^{\phi^{-1}(d)} f(\phi(x)) \phi'(x) dx.$$

Es ist entweder $\phi'(x) > 0$ für alle $x \in [a, b]$ oder $\phi'(x) < 0$ für alle $x \in [a, b]$, denn ϕ' ist stetig und $\phi'(x) \neq 0$, da ϕ ein Diffeomorphismus ist. Ist $\phi'(x) > 0$ für alle $x \in [a, b]$, so ist ϕ streng monoton wachsend und es gilt $\phi^{-1}(c) = a$ und $\phi^{-1}(d) = b$. Ist $\phi'(x) < 0$ für alle $x \in [a, b]$, so ist ϕ streng monoton fallend und es gilt $\phi^{-1}(c) = b$ und $\phi^{-1}(d) = a$. Folglich erhalten wir

$$\int_c^d f(y) dy = \int_a^b f(\phi(x)) \cdot |\phi'(x)| dx = \int_a^b f(\phi(x)) \cdot |\det J\phi_x| dx.$$

In den folgenden Beispielen demonstrieren wir die Anwendung der Transformationsformel.

Beispiel 7: Wir berechnen das Integral $\int_{(x-a)^2+y^2 \leq a^2} (x^2 + y^2) d(x, y)$.

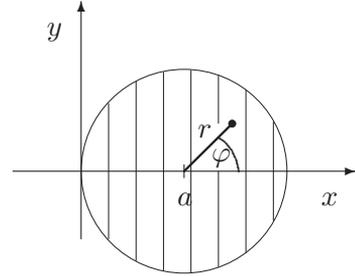
Der Integrationsbereich ist die Kreisscheibe um $(a, 0)$ vom Radius a , d.h.

$$B := \{(x, y) \in \mathbb{R}^2 \mid (x - a)^2 + y^2 \leq a^2\}.$$

Um das Integral zu berechnen, beschreiben wir B (bis auf eine Nullmenge) durch Polarkoordinaten um $(a, 0)$:

Jeden Punkt (x, y) von B schreiben wir als $x = a + r \cos \varphi$ und $y = r \sin \varphi$. Die Abbildung

$$\begin{aligned} \phi : (0, \infty) \times (0, 2\pi) &\longrightarrow \mathbb{R}^2 \setminus \underbrace{([0, \infty) \times \{0\}}_{:=N \text{ Nullmenge}} \\ (r, \varphi) &\longmapsto (a + r \cos \varphi, r \sin \varphi). \end{aligned}$$



ein Diffeomorphismus. Es gilt $B \setminus N = \phi((0, a) \times (0, 2\pi))$. Die Determinante der Jacobi-Matrix von ϕ ist

$$\det J\phi_{(r, \varphi)} = \det \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} = r > 0.$$

Wir erinnern uns daran, dass man zum Integrationsbereich Nullmengen beliebig hinzunehmen oder herausnehmen kann, ohne das Integral zu verändern (siehe Satz 9.7). Aus der Transformationsformel folgt dann

$$\begin{aligned} \int_B (x^2 + y^2) d(x, y) &= \int_{B \setminus N} (x^2 + y^2) d(x, y) \\ &\stackrel{\text{Trafo}}{=} \int_{(0, a) \times (0, 2\pi)} (a + r \cos \varphi)^2 + r^2 \sin^2 \varphi \cdot |\det J\phi_{(r, \varphi)}| d(r, \varphi) \\ &= \int_{[0, a] \times [0, 2\pi]} (a^2 r + 2r^2 a \cos \varphi + r^3 \cos^2 \varphi + r^3 \sin^2 \varphi) d(r, \varphi) \\ &\stackrel{\text{Fubini}}{=} \int_0^a \left(\int_0^{2\pi} (a^2 r + 2r^2 a \cos \varphi + r^3) d\varphi \right) dr \\ &= 2\pi \int_0^a (a^2 r + r^3) dr = 2\pi \left(a^2 \frac{r^2}{2} \Big|_0^a + \frac{r^4}{4} \Big|_0^a \right) \\ &= \frac{3}{2} \pi a^4. \end{aligned}$$

Beispiel 8: Wir berechnen das Integral $\int_{x^2+y^2+z^2 \leq 1} |xy| d(x, y, z)$.

Der Integrationsbereich ist die Kugel D_1^3 vom Radius 1. Wir betrachten die Kugelkoordinaten im \mathbb{R}^3

$$\phi : (0, \infty) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times (0, 2\pi) \longrightarrow \mathbb{R}^3 \setminus \underbrace{([0, \infty) \times \{0\} \times \mathbb{R})}_{:=L \text{ Nullmenge}}$$

mit

$$\phi(r, u, v) := (r \cos u \cos v, r \cos u \sin v, r \sin u).$$

Die Menge $D_1^3 \setminus L$ wird durch die Kugelkoordinaten aus der Menge

$$A := (0, 1] \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times (0, 2\pi)$$

beschrieben, d.h. es gilt $\phi(A) = D_1^3 \setminus L$. ϕ ist ein Diffeomorphismus mit der Jacobi-Matrix

$$J\phi_{(r,u,v)} = \begin{pmatrix} \cos u \cdot \cos v & -r \cdot \sin u \cdot \cos v & -r \cdot \cos u \cdot \sin v \\ \cos u \cdot \sin v & -r \cdot \sin u \cdot \sin v & r \cdot \cos u \cdot \cos v \\ \sin u & r \cos u & 0 \end{pmatrix}.$$

Für die Determinante der Jacobi-Matrix berechnet man

$$|\det J\phi_{(r,u,v)}| = r^2 \cdot \cos u.$$

Wir benutzen wieder, dass man zu Integrationsgebieten Nullmenge beliebig hinzufügen oder wegnehmen kann, ohne das Integral zu verändern und erhalten aus der Transformationsformel

$$\begin{aligned} \int_{D_1^3} |xy| d(x, y, z) &= \int_{D_1^3 \setminus L} |xy| d(x, y, z) \\ &\stackrel{\text{Trafo}}{=} \int_A |r \cos u \cos v| \cdot |r \cos u \sin v| \cdot |\det J\phi_{(r,u,v)}| d(r, u, v) \\ &= \int_{[0,1] \times [-\frac{\pi}{2}, \frac{\pi}{2}] \times [0,2\pi]} r^4 \cos^3 u \cdot |\sin v \cos v| d(r, u, v) \\ &\stackrel{\text{Bsp.1}}{=} \int_0^1 r^4 dr \cdot \int_{-\pi/2}^{\pi/2} \cos^3 u du \cdot \int_0^{2\pi} \frac{1}{2} |\sin 2v| dv. \end{aligned}$$

Da

$$\frac{1}{2} \int_0^{2\pi} |\sin 2v| dv = \frac{1}{4} \int_0^{4\pi} |\sin w| dw = \int_0^\pi \sin w dw = 2,$$

und

$$\int_{-\pi/2}^{\pi/2} \cos^3 u \, du \stackrel{P.Int}{=} \frac{\cos^2 u \sin u}{3} \Big|_{\pi/2}^{\pi/2} + \frac{2}{3} \int_{-\pi/2}^{\pi/2} \cos u \, du = \frac{4}{3}$$

folgt

$$\int_{D_1^3} |xy| \, d(x, y, z) = \frac{8}{15}.$$

Beispiel 9: Geometrische Interpretation der Determinante

Sei $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein linearer Isomorphismus. Dann gilt für jede Jordan-messbare Teilmenge $A \subset \mathbb{R}^n$

$$\text{vol}(L(A)) = |\det(L)| \cdot \text{vol}(A).$$

Zum Beweis benutzen wir, dass $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein Diffeomorphismus ist und wenden die Transformationsformel an:

$$\begin{aligned} \text{vol}(L(A)) &= \int_{L(A)} 1 \, dx = \int_A |\det dL_x| \, dx = \int_A |\det L| \, dx \\ &= |\det L| \int_A 1 \, dx = |\det L| \cdot \text{vol}(A). \end{aligned}$$

Daraus erhalten wir die aus der Vorlesung *Lineare Algebra* bekannte geometrische Interpretation der Determinante einer regulären Matrix:

Sei $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ eine $(n \times n)$ -Matrix, deren Spaltenvektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ linear unabhängig sind. Wir betrachten den linearen Isomorphismus $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$, der die kanonischen Basisvektoren e_i auf \mathbf{a}_i abbildet ($i = 1, \dots, n$). Dann ist $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ gerade die Darstellungsmatrix $[L]$ von L bzgl. der kanonischen Basis des \mathbb{R}^n . Die Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ spannen das Parallelotop

$$\mathcal{P}(\mathbf{a}_1, \dots, \mathbf{a}_n) := \left\{ \sum_{i=1}^n \lambda_i \mathbf{a}_i \mid 0 \leq \lambda_i \leq 1 \right\} \subset \mathbb{R}^n$$

auf. Dann gilt

$$\mathcal{P}(\mathbf{a}_1, \dots, \mathbf{a}_n) = L([0, 1]^n)$$

und wir erhalten

$$\text{vol}(\mathcal{P}(\mathbf{a}_1, \dots, \mathbf{a}_n)) = \text{vol}(L([0, 1]^n)) = |\det L| \cdot \text{vol}([0, 1]^n) = |\det[L]| = |\det(\mathbf{a}_1 \dots \mathbf{a}_n)|.$$

Dies liefert

$$|\det(\mathbf{a}_1 \dots \mathbf{a}_n)| = \text{vol}(\mathcal{P}(\mathbf{a}_1, \dots, \mathbf{a}_n)).$$

Sachverzeichnis

- C^n , 90
- $K_d(p, \varepsilon)$, 54
- $L(\gamma)$, 34
- d , 51
- $d(p, q)$, 51
- df_p , 74
- $x_n \rightarrow p$, 55
- y^α , 93
- $\alpha!$, 93
- ∂A , 60
- $C^n(U, \mathbb{R})$, 90
- $\text{Int}(A)$, 60
- $\text{cl}(A)$, 60
- $\lim_{n \rightarrow \infty} x_n$, 55
- $x_n \xrightarrow{n \rightarrow \infty} p$, 55
- $\frac{\partial f}{\partial x_i}(p)$, 77
- $\frac{\partial^n f}{\partial x_{i_1} \dots \partial x_{i_n}}$, 90
- \int_a^∞ , 27
- $\nabla_{\mathbf{a}} f(p)$, 76

- abgeschlossen, 59
- Ableitung, 74
 - partielle, 77
- Abschluss einer Menge, 60
- Abstand, 51
- Abstandsfunktion, 51
- äußeres Volumen von A , 116
- archimedische Spirale, 42
- Astroide, 41

- Banachscher Fixpunktsatz, 70
- beschränkte Folge, 56
- beschränkte Teilmenge, 56

- Cauchy-Bedingung, 58
- Cauchy-Folge, 57
- C^n -Funktion, 90

- Diffeomorphismus, 98
- Differential, 74
- differenzierbar, 74
- Differenzierbarkeit
 - hinreichende Bedingung für, 78
- diskrete Metrik, 52
- diskreter metrischer Raum, 52
- divergente Folge, 55
- Dreiecksungleichung, 51

- einfache Kurve, 37
- Ellipse, 32
 - Flächeninhalt, 41
 - Umfang, 36
- ε -Kugel, 54

- Feinheit der Zerlegung eines Intervalls, 2
- Fixpunktsatz, Banachscher, 70
- Folge
 - divergente, 55
 - konvergente, 55

- geschlossene Kurve, 37
- gleichmäßig konvergente
 - Folge von Funktionen, 24
- gleichmäßig stetig, 66
- Grenzwert
 - Eindeutigkeit, 56
 - einer Folge, 55

- homöomorph, 69
- Homöomorphismus, 69
- Hyperbel, 32

- implizite Funktionen
 - Satz über, 107
- innerer Punkt, 60
- Inneres einer Menge, 60
- inneres Volumen von A , 116

- Integration
 - partielle, 13
- Integrierbarkeitskriterium
 - Riemannsches, 4
- isoliertes lokales Maximum, 96
- isoliertes lokales Minimum, 96
- Isometrie, 54
- isometrische metrische Räume, 54
- Iterierte Integration, 123

- Jordan-meßbare Menge, 116
- Jordan-Volumen, 116
- Jordansche Nullmenge, 116

- kartesisches Produkt metrischer Räume, 52
- Kettenregel, 82
 - für partielle Ableitungen, 84
- kompakt, 61
- kontrahierende Abbildung, 69
- Kontraktion, 69
- konvergente Folge, 55
- konvexe Menge, 68
- Koordinatentransformation, 98
- Kugelkoordinaten, 100
- Kurve
 - einfache, 37
 - geschlossene, 37
 - parametrisierte, 31
 - positiv-orientierte, 37

- Länge
 - einer rektifizierbaren Kurve, 34
 - einer stückweisen C^1 -Kurve, 36
 - eines Intervalls, 2
- Leibnizsche Sektorformel, 40
- Lipschitz-Konstante, 67
- lipschitzstetig, 67
- lokales Maximum, 96
- lokales Minimum, 96
- Luftlinienabstand, 52

- Mannheimer-Metrik, 52
- Maximum
 - lokales, 96
 - striktes lokales, 96
- Metrik, 51
 - der französischen Eisenbahn, 52
 - diskrete, 52
- metrische Räume
 - isometrische, 54
 - kartesisches Produkt, 52
- metrischer Raum, 51
 - abgeschlossene Teilmenge, 59
 - diskreter, 52
 - innerer Punkt, 60
 - offene Teilmenge, 59
 - Rand einer Teilmenge, 60
 - vollständiger, 57
- Minimum
 - lokales, 96
 - striktes lokales, 96
- Mittelwertsatz
 - für Funktionen mehrerer Veränderlicher, 85
- Mittelwertsatz für Riemann-Integrale, 120
- Multiindex, 93
 - Ordnung eines, 93

- n -te partielle Ableitung, 90
- Norm, 53
- Normale an eine Fläche, 111
- normierter Vektorraum, 53

- oberes Riemann-Integral, 3
- Obersumme von A bzgl. \mathcal{P} , 115
- Obersumme von f bzgl. \mathcal{P} , 117
- offen, 59

- parametrisierte Kurve, 31
- Parametrisierung einer Kurve, 31
- Partialbruchzerlegung
 - komplexer rationaler Funktionen, 16
 - reeller rationaler Funktionen, 18
- partielle Ableitung, 77
 - der Ordnung n , 90
 - höherer Ordnung, 90
- partielle Ableitungen
 - Kettenregel, 84
 - Vertauschbarkeit, 91
- partielle Integration, 13
- positiv-orientierte Kurve, 37
- Positivität, 51
- Prinzip von Cavalieri, 124
- Produktmetrik, 52
- punktweise konvergente
 - Folge von Funktionen, 23

- Rand einer Menge, 60
- Randpunkt einer Menge, 60
- reguläre Fläche, 110
- rektifizierbare Kurve, 34
- Richtungsableitung, 76
- Riemann-Integral für Funktionen mehrerer
 - Variabler, 118
- Riemann-integrierbar

- uneigentlich, 27
- Riemann-Integral über eine Jordan-meßbare Menge, 120
- Riemann-integrierbare Funktion mehrerer Variablen, 118
- Riemannsches Integrierbarkeitskriterium, 4
- Rotationskörper, 126
- Satz über
 - die Stetigkeit der inversen Abbildung, 69
 - die Umkehrabbildung, 101
 - implizite Funktionen, 107
- Satz von
 - Schwarz, 91
- Satz von Fubini, 121
- Schnitte Jordan-meßbarer Mengen, 123
- Schraubenlinie, 32, 36
- Sektorformel, Leibnitzsche, 40
- Spirale, archimedische, 42
- Standardmetrik, 52
- Sternkurve, 41
- stetige Abbildung, 64
- stetige Abbildung in einem Punkt, 64
- striktes lokales Maximum, 96
- striktes lokales Minimum, 96
- Substitutionsregel, 14
- Symmetrie, 51
- Tangentialebene an eine Fläche, 111
- Tangentialraum an eine Fläche, 110
- Taylorformel
 - mehrdimensionale, 94
- Teilintervall, 2
- Transformationsformel für Riemann-Integrale, 127
- Umkehrfunktion
 - Ableitung der, 100
- Umkehrsatz, 101
- uneigentlich Riemann-integrierbar, 27
- unteres Riemann Integral, 3
- Untersumme von A bzgl. \mathcal{P} , 115
- Untersumme von f bzgl. \mathcal{P} , 117
- Verfeinerung der Zerlegung eines Intervalls, 2
- vollständiger metrischer Raum, 57
- Weg, 67
- wegzusammenhängend, 67
- Zerlegung eines Intervalls, 2
- Zerlegung eines Quaders, 115
- zusammenhängende Teilmenge, 63
- Zylinderkoordinaten, 99