# Analysis I – III

Vorlesungsskriptum WS 2005/06 – WS 2006/07

R. Verfürth

Fakultät für Mathematik, Ruhr-Universität Bochum

## Inhaltsverzeichnis

KAPITEL I

# Aufbau des Zahlsystems

In diesem Kapitel setzen wir die natürlichen, ganzen, rationalen und reellen Zahlen als im Prinzip bekannt voraus. Unser Ziel ist es, die wesentlichen Punkte des axiomatischen Aufbaus herauszuarbeiten, ohne aber dabei bekannte Tatsachen wie z.B. die Assoziativität der Addition formal aus den Axiomen herzuleiten.
Für einen streng axiomatischen Aufbau verweisen wir auf E. Landau: Grundlagen der Analysis.
Die wesentlichen Kernpunkte unserer Darstellung sind das Induktionsprinzip und die Vollständigkeit der reellen Zahlen.
Als nicht bekannt setzen wir die komplexen Zahlen voraus. Ihre Eigenschaften stellen wir etwas ausführlicher dar.

## I.1. Die natürlichen Zahlen

Die intuitiven Vorstellungen von den natürlichen Zahlen $\mathbb{N}$, die vom Abzählen geprägt sind, lassen sich wie folgt zusammenfassen:

(1) Es gibt die kleinste Einheit 1.
(2) Jede Zahl kann man durch geeignetes „Zusammenzählen" der Einheit darstellen.
(3) Zu jeder Zahl gibt es eine noch größere.
(4) Es gibt eine „natürliche" Anordnung, die festlegt, welche von zwei Zahlen die „größere" ist.
(5) Aus praktischen Gründen ist es sinnvoll, eine Zahl 0, die „nichts" ist, hinzuzunehmen.

Diese intuitiven Vorstellungen werden in folgendem, auf G. Peano (1858-1932) zurückgehenden Axiomensystem präzisiert.

Definition I.1.1 (Peano Axiome). Die natürlichen Zahlen $\mathbb{N}$ erfüllen die folgenden Axiome:

(N1) (Unendlichkeitsaxiom) Es gibt eine Abbildung $\nu : \mathbb{N} \to \mathbb{N}^* = \mathbb{N}\backslash\{0\}$, die jeder natürlichen Zahl $n$ eine natürliche Zahl $\nu(n) \neq 0$ zuordnet, mit der Eigenschaft (injektiv)

$$n \neq m \Longrightarrow \nu(n) \neq \nu(m).$$

(N2) (Induktionsaxiom) Ist $M$ eine Teilmenge von $\mathbb{N}$ mit den Eigenschaften
(1) $0 \in M$
(2) $n \in M \Longrightarrow \nu(n) \in M$,

dann ist $M = \mathbb{N}$.

Wir nennen $\nu(0) = 1$ und $\nu(n) = n + 1$ den Nachfolger von $n$.

Der folgende Satz, den wir nicht beweisen, zeigt, dass die durch die Peano Axiome definierte Menge die Eigenschaften hat, die wir intuitiv von den natürlichen Zahlen erwarten.

Satz I.1.2. *Auf der Menge $\mathbb{N}$ können wir eindeutig eine Addition $+$, eine Multiplikation $\cdot$ und eine Anordnung $\leq$ definieren, so dass die folgenden Eigenschaften gelten:*

(1) $n + m = m + n$ (Kommutativität)
$(n + m) + k = n + (m + k)$ (Assoziativität)
$n + 0 = n$ (neutrales Element 0)
(2) $n \cdot m = m \cdot n$ (Kommutativität)
$(n \cdot m) \cdot k = n \cdot (m \cdot k)$ (Assoziativität)
$n \cdot 1 = n$ (neutrales Element 1)
(3) $(n + m) \cdot k = n \cdot k + m \cdot k$ (Distributivgesetz)
(4) $m \leq n \iff \exists\, l \in \mathbb{N} : m + l = n$
$m < n \iff \exists\, l \in \mathbb{N}^* : m + l = n \iff m \leq n$ *und* $m \neq n$
*$l$ ist durch $m$ und $n$ eindeutig festgelegt und heißt Differenz zwischen $m$ und $n$, $l = n - m$.*
(5) $0 \cdot n = 0$
$\nu(n) = n + 1$
(6) $n \leq n$ (Reflexivität)
$n \leq m$ *und* $m \leq k \Longrightarrow n \leq k$ (Transitivität)
$n \leq m$ *und* $m \leq n \Longrightarrow n = m$ (Antisymmetrie)
$0 \leq n\ \forall n \in \mathbb{N}$
(7) *Für alle $n, m \in \mathbb{N}$ gilt genau eine der Beziehungen*
$n < m$ *oder* $n = m$ *oder* $m < n$ (Totalordnung)
(8) $\forall n \in \mathbb{N}\ \nexists\, k \in \mathbb{N}$ *mit* $n < k < n + 1$.
(9) $m \leq n \iff m + l \leq n + l\ \forall l \in \mathbb{N}$.
(10) $m, n \in \mathbb{N}^* \Longrightarrow m \cdot n \in \mathbb{N}^*$.
(11) $\forall m, n \in \mathbb{N}\ \forall k \in \mathbb{N}^*$ *gilt* $m \leq n \iff mk \leq nk$ *und* $m < n \iff mk < nk$.

Die erste wichtige Konsequenz des Induktionsaxioms ist der folgende Wohlordnungssatz.

Satz I.1.3. *$\mathbb{N}$ ist* wohlgeordnet, *d.h., jede nicht leere Teilmenge $M$ von $\mathbb{N}$ besitzt ein Minimum.*

Beweis. Sei $M \subset \mathbb{N}, M \neq \emptyset$. Wir nehmen an, $M$ besäße kein Minimum. Dann gilt

$$0 \in N = \{n \in \mathbb{N} : n < m\ \forall m \in M\}.$$

Sei $n \in N$. Dann ist $n + 1 \leq m$ für alle $m \in M$. Da $M$ kein Minimum besitzt, gilt $n + 1 \notin M$. Also ist $n + 1 \in N$.

Aus dem Induktionsaxiom folgt $N = \mathbb{N}$. Dies ist ein Widerspruch. Denn nach Voraussetzung gibt es ein $m \in M$. Dann gilt aber $m+1 \notin N$. □

Definition I.1.4. Für $n \in \mathbb{N}$ sei

$$\mathbb{N}_n = \{0, 1, \ldots, n\}$$

und für $n \in \mathbb{N}^*$ sei

$$\mathbb{N}_n^* = \{1, \ldots, n\}.$$

Satz I.1.5 (Induktionsprinzip, 1. Fassung). *Für alle $n \in \mathbb{N}$ sei $A(n)$ eine Aussage. Es gelte:*

(1) (Induktionsanfang; IA) *$A(0)$ ist richtig.*
(2) (Induktionsschritt; IS) *$A(n)$ ist richtig $\Longrightarrow$ $A(n+1)$ ist richtig.*

*Dann gilt:*

$$A(n) \textit{ ist für alle } n \in \mathbb{N} \textit{ richtig.}$$

Beweis. Sei

$$M := \{n \in \mathbb{N} : A(n) \text{ ist richtig}\}.$$

Dann gilt:

(1) $0 \in M$.
(2) $n \in M \Longrightarrow n+1 \in M$.

Damit folgt aus dem Induktionsaxiom $M = \mathbb{N}$. □

Wir wollen zwei einfache Beispiele für das Induktionsprinzip geben. Dazu benötigen wir eine Bezeichnung.

Definition I.1.6. Seien $m \leq n$ ganze Zahlen und $a_m, \ldots, a_n$ reelle Zahlen. Dann ist

$$\sum_{k=m}^{n} a_k = a_m + a_{m+1} + \ldots + a_n$$
$$\prod_{k=m}^{n} a_k = a_m \cdot a_{m+1} \cdot \ldots \cdot a_n.$$

Falls $m > n$ ist, definieren wir

$$\sum_{k=m}^{n} a_k = 0 \qquad \text{(leere Summe)}$$
$$\prod_{k=m}^{n} a_k = 1 \qquad \text{(leeres Produkt)}.$$

Beispiel I.1.7. Für alle $n \in \mathbb{N}$ gilt

(a)
$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2}$$

(b) $$\sum_{k=0}^{n}(2k+1) = (n+1)^2.$$

Beweis. ad (a):

(IA) $$\sum_{k=0}^{0} k = 0 = 0 \cdot \frac{1}{2}$$

(IS) $$\begin{aligned}\sum_{k=0}^{n+1} k &= \sum_{k=0}^{n} k + (n+1)\\ &= \frac{n(n+1)}{2} + (n+1)\\ &= \frac{(n+1)(n+2)}{2}.\end{aligned}$$

ad (b):

(IA) $$\sum_{k=0}^{0}(2k+1) = 1 = (0+1)^2$$

(IS) $$\begin{aligned}\sum_{k=0}^{n+1}(2k+1) &= \sum_{k=0}^{n}(2k+1) + [2(n+1)+1]\\ &= (n+1)^2 + 2n + 3\\ &= n^2 + 2n + 1 + 2n + 3\\ &= n^2 + 4n + 4\\ &= (n+2)^2.\end{aligned}$$

□

Die folgende Abwandlung des Induktionsprinzips erlaubt den Beweis von Aussagen, die für alle natürlichen Zahlen $n \geq n_0$ gelten.

Satz I.1.8 (Induktionsprinzip, 2. Fassung). *Es sei* $n_0 \in \mathbb{N}$ *und für alle* $n \geq n_0$ *sei* $A(n)$ *eine Aussage. Es gelte:*

(1) (Induktionsanfang; IA) $A(n_0)$ *ist richtig.*
(2) (Induktionsschritt; IS) $n \geq n_0$ *und* $A(n)$ *ist richtig* $\Longrightarrow$ $A(n+1)$ *ist richtig.*

*Dann gilt:*

$$A(n) \text{ ist für alle } n \geq n_0 \text{ richtig.}$$

Beweis. Wir nehmen an, dass es ein $m \geq n_0$ gibt, so dass $A(m)$ falsch ist, d.h.

$$M := \{n \in \mathbb{N} : n \geq n_0 \text{ und } A(n) \text{ ist falsch}\} \neq \emptyset.$$

Gemäß Satz I.1.3 existiert

$$m_0 = \min(M).$$

Aus (1) folgt $m_0 > n_0$. Also ist $A(n)$ richtig für alle $n$ mit $n_0 \leq n \leq m_0 - 1$. Aus (2) folgt, dass $A(m_0)$ richtig ist im Widerspruch zur Konstruktion von $m_0$. □

Beispiel I.1.9. Für alle natürlichen Zahlen $n \geq 5$ gilt

$$n^2 < 2^n.$$

Beweis.

(IA) $$2^5 = 32 > 25 = 5^2$$

(IS) $$\begin{aligned} 2^{n+1} &= 2 \cdot 2^n \\ &> 2n^2 = n^2 + n^2 \\ &= n^2 + 2n + 1 + n^2 - 2n - 1 \\ &= (n+1)^2 + (n-1)^2 - 2 \\ &\geq (n+1)^2 + 16 - 2 \qquad \text{(wegen } n \geq 5\text{)} \\ &> (n+1)^2. \end{aligned}$$

□

Eine andere wichtige Konsequenz des Induktionsaxioms ist das Prinzip der rekursiven Definition.

Satz I.1.10 (Prinzip der rekursiven Definition). *Gegeben sei eine Abbildung $g : M \to M$ einer Menge $M$ in sich und ein Element $a \in M$. Dann gibt es genau eine Abbildung $f : \mathbb{N} \to M$ mit folgenden Eigenschaften:*

(1) $f(0) = a$
(2) $f(n+1) = g(f(n)) \quad \forall n \in \mathbb{N}$.

Beweis. 1. Schritt: Zu jedem $n \in \mathbb{N}$ gibt es genau eine Abbildung $f_n : \mathbb{N}_n \to M$ mit

(i) $f_n(0) = a$
(ii) $f_n(m+1) = g(f_n(m)) \quad \forall 0 \leq m < n$.

Beweis des 1. Schrittes durch Induktion:
(IA): $f_0(0) = a$ legt $f_0$ eindeutig fest.
(IS): Definiere $f_{n+1}$ durch

$$\begin{aligned} f_{n+1}(m) &= f_n(m) \quad \forall 0 \leq m \leq n \\ f_{n+1}(n+1) &= g(f_n(n)). \end{aligned}$$

Offensichtlich erfüllt $f_{n+1}$ die Eigenschaften (i) und (ii). Sei nun $f'_{n+1}$ eine weitere Abbildung mit diesen Eigenschaften. Aus der Induktionsvoraussetzung folgt dann

$$f_{n+1}(m) = f'_{n+1}(m) \quad \forall 0 \leq m \leq n.$$

Wegen (ii) gilt dann auch

$$f_{n+1}(n+1) = f'_{n+1}(n+1).$$

Damit ist der 1. Schritt bewiesen.
2. Schritt: Wir definieren eine Abbildung $f : \mathbb{N} \to M$ durch

$$f(n) := f_n(n).$$

Dann gilt nach Schritt 1

$$\begin{aligned} f(0) &= f_0(0) = a \\ f(n+1) &= f_{n+1}(n+1) = g(f_n(n)) = g(f(n)). \end{aligned}$$

Also leistet $f$ das Gewünschte.
3. Schritt: Sei $f' : \mathbb{N} \to M$ eine weitere Abbildung mit den Eigenschaften (1) und (2). Dann folgt

$$f(0) = a = f'(0)$$

und

$$\begin{aligned} & f(n) = f'(n) \\ \Longrightarrow & f(n+1) = g(f(n)) = g(f'(n)) = f'(n+1). \end{aligned}$$

Aus dem Induktionsprinzip folgt, dass $f$ und $f'$ übereinstimmen. □

Mit Hilfe von Satz I.1.10 kann man zeigen, dass unsere intuitive Definition I.1.6 sinnvoll ist. Ebenso folgt, dass die Summen und Produkte nicht von der Reihenfolge der Summanden bzw. Faktoren abhängen. Wir werden später noch weitere Anwendungen des Rekursionsprinzips kennen lernen.

Definition I.1.11. Für $n \in \mathbb{N}$ und $k \in \mathbb{N}_n$ definieren wir

$$n! = \prod_{k=1}^{n} k \qquad \text{(Fakultät)}$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad \text{(Binomialkoeffizient)}.$$

Satz I.1.12. *Für $n \in \mathbb{N}^*$ ist $n!$ gleich der Anzahl aller möglichen Anordnungen einer $n$-elementigen Menge $\{A_1, \ldots, A_n\}$.*

Beweis. (IA): $n = 1$ ist klar.
(IS): Für $1 \le k \le n+1$ sei $K_k$ die Klasse aller Anordnungen von $\{A_1, \ldots, A_{n+1}\}$, die $A_k$ an erster Stelle haben. Offensichtlich sind die $K_k$ disjunkt, und jede Anordnung von $\{A_1, \ldots, A_{n+1}\}$ gehört einer Klasse $K_k$ an. Nach Induktionsvoraussetzung hat jedes $K_k$ genau $n!$ Elemente. Also gibt es

$$(n+1) \cdot n! = (n+1)!$$

Anordnungen von $\{A_1, \ldots, A_{n+1}\}$. □

Lemma I.1.13. (1) *Für $n \in \mathbb{N}$ und $k \in \mathbb{N}_n$ ist*

$$\binom{n}{k} = \binom{n}{n-k}.$$

(2) *Für* $n \in \mathbb{N}^*$ *und* $k \in \mathbb{N}^*_{n-1}$ *ist*

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

Beweis. ad (1): Folgt aus Definition 1.11.
ad (2):

$$\begin{aligned}\binom{n-1}{k-1} + \binom{n-1}{k} &= \frac{(n-1)!}{(k-1)!(n-k)!} + \frac{(n-1)!}{k!(n-k-1)!}\\ &= \frac{k(n-1)!}{k!(n-k)!} + \frac{(n-k)(n-1)!}{k!(n-k)!}\\ &= \frac{n(n-1)!}{k!(n-k)!}\\ &= \binom{n}{k}.\end{aligned}$$

□

Satz I.1.14. *Für* $n \in \mathbb{N}^*$ *und* $k \in \mathbb{N}_n$ *ist* $\binom{n}{k}$ *die Zahl aller* $k$*-elementigen Teilmengen einer* $n$*-elementigen Menge* $\{A_1, \ldots, A_n\}$.

Beweis. Sei $C^n_k$ die Zahl aller $k$-elementigen Teilmengen einer $n$-elementigen Menge. Wir zeigen durch Induktion nach $n$, dass gilt $C^n_k = \binom{n}{k}$.
(IA): $C^1_0 = 1 = \binom{1}{0} = \binom{1}{1} = C^1_1$.
(IS): Wegen $C^{n+1}_0 = 1 = \binom{n+1}{0} = \binom{n+1}{n+1} = C^{n+1}_{n+1}$ können wir $1 \le k \le n$ voraussetzen.
Sei $K_0$ die Menge aller $k$-elementigen Teilmengen von $\{A_1, \ldots, A_{n+1}\}$, die $A_{n+1}$ nicht enthalten, und $K_1$ die Menge aller $k$-elementigen Teilmengen von $\{A_1, \ldots, A_{n+1}\}$, die $A_{n+1}$ enthalten. Offensichtlich sind $K_0$ und $K_1$ disjunkt, und jede $k$-elementige Teilmenge von $\{A_1, \ldots, A_{n+1}\}$ ist in $K_0$ oder $K_1$ enthalten. Nach Induktionsvoraussetzung ist

$$|K_0| = C^n_k = \binom{n}{k}$$
$$|K_1| = C^n_{k-1} = \binom{n}{k-1}.$$

Damit folgt die Behauptung aus Lemma I.1.13. □

Satz I.1.14 zeigt, dass die Zahlen $\binom{n}{k}$ natürliche Zahlen sind, was nicht direkt aus der Definition folgt.
Es gibt

$$\binom{49}{6} = 13'983'816$$

6-elementige Teilmengen einer 49-elementigen Menge. Also ist die Chance, im Lotto 6 Richtige zu haben, ca. 1 : 14 Millionen.

Satz I.1.15 (Binomischer Lehrsatz). *Für alle reellen Zahlen $x, y$ und alle $n \in \mathbb{N}$ gilt*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Beweis. Wir erinnern daran, dass für jede reelle Zahl $z$ gilt

$$z^0 = 1.$$

Außerdem benutzen wir die Beziehung

$$\sum_{k=m}^{n} a_k = \sum_{k=m+1}^{n+1} a_{k-1}.$$

(IA): $n = 0$ klar.
(IS):

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)(x+y)^n \\
&= (x+y) \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \\
&= \sum_{k=0}^{n} \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k} \\
&= \sum_{k=1}^{n+1} \binom{n}{k-1} x^k y^{n+1-k} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k}
\end{aligned}$$

wegen Lemma I.1.13. □

Satz I.1.16 (Summenformel für die geometrische Reihe). *Für jede reelle Zahl $x \neq 1$ und jedes $n \in \mathbb{N}$ gilt*

$$\sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x}.$$

Beweis. (IA): $n = 0$ klar
(IS):

$$\sum_{k=0}^{n+1} x^k = \sum_{k=0}^{n} x^k + x^{n+1} = \frac{1 - x^{n+1}}{1 - x} + x^{n+1} = \frac{1 - x^{n+2}}{1 - x}.$$

□

Zum Abschluss dieses Paragraphen kehren wir zum Problem des Abzählens zurück. Intuitiv ist eine endliche Menge dadurch charakterisiert, dass wir ihre Elemente aufzählen können und irgendwann damit fertig werden. Mathematisch heißt dies:

Definition I.1.17. Eine Menge $M$ heißt endlich, wenn es ein $m \in \mathbb{N}$ und eine surjektive Abbildung $f : \mathbb{N}_m \to M$ gibt.

Satz I.1.18. $\mathbb{N}$ *ist nicht endlich.*

Beweis. Angenommen, $\mathbb{N}$ wäre endlich. Dann gibt es ein $m \in \mathbb{N}$ und eine surjektive Abbildung $f : \mathbb{N}_m \to \mathbb{N}$. Definiere

$$\begin{aligned} k_0 &= f(0) \\ k_{n+1} &= \max\{k_n, f(n+1)\} \quad 0 \le n \le m-1. \end{aligned}$$

Offensichtlich gilt

$$k_m \ge f(n) \quad \forall n \in \mathbb{N}_m.$$

Mithin ist $k_m + 1 \notin f(\mathbb{N}_m)$; Widerspruch. □

$\mathbb{N}$ ist zwar nicht endlich, aber wir können die Elemente von $\mathbb{N}$ immerhin auf- bzw. abzählen. Dies führt zu:

Definition I.1.19. Eine Menge $M$ heißt abzählbar, wenn es eine surjektive Abbildung $f : \mathbb{N} \to M$ gibt.

Bemerkung I.1.20. Eine endliche Menge $M$ ist abzählbar. Setze $f : \mathbb{N}_m \to M$ durch

$$f(n) = f(m) \quad \forall n \ge m$$

auf ganz $\mathbb{N}$ fort.

Der folgende Satz zeigt, dass es nicht abzählbare Mengen gibt. Sein Beweis verläuft nach dem Prinzip „Der Barbier, der alle Leute rasiert, die sich nicht selbst rasieren", das auf B. Russell (1872-1970) zurückgeht.

Satz I.1.21. *Die* Potenzmenge $\mathcal{P}(\mathbb{N})$ *von* $\mathbb{N}$ *ist nicht abzählbar.*

Beweis. Wir nehmen an $\mathcal{P}(\mathbb{N})$ wäre abzählbar. Sei $f : \mathbb{N} \to \mathcal{P}(\mathbb{N})$ eine surjektive Abbildung. Definiere

$$M = \{n \in \mathbb{N} : n \notin f(n)\}.$$

Dann gibt es ein $m \in \mathbb{N}$ mit $M = f(m)$ (Surjektivität). Wir erhalten

$$\begin{aligned} m \in M &\Longrightarrow m \notin f(m) = M \\ m \notin M &\Longrightarrow m \in f(m) \Longrightarrow m \in M; \end{aligned}$$

in jedem Fall ein Widerspruch. □

## I.2. Die ganzen Zahlen

In $\mathbb{N}$ können wir addieren, multiplizieren und Zahlen miteinander vergleichen. Es gibt aber z.B. keine Zahl $n \in \mathbb{N}$ mit

$$n + 5 = 3.$$

Um dieses Manko zu beheben, definieren wir negative Zahlen durch

(1) $-m$ ist für $m \in \mathbb{N}^*$, die Zahl mit $m + (-m) = 0$.

(2) $-0 = 0$.

Man kann zeigen, dass diese Definition sinnvoll ist. Die GANZEN ZAHLEN $\mathbb{Z}$ sind dann gegeben durch

$$\mathbb{Z} = \mathbb{N} \cup \{-m : m \in \mathbb{N}^*\}.$$

Als nächstes setzen wir die Addition, Multiplikation und die Relation $\leq$ auf $\mathbb{Z}$ wie folgt fest:

$$\begin{aligned}
(-m) + (-n) &= -(m+n) \\
m + (-n) &= (-n) + m \\
&= \begin{cases} l & \text{falls } m \geq n, m = n + l, \\ -(n + (-m)) & \text{falls } m < n. \end{cases} \\
m \cdot (-n) &= (-n) \cdot m \\
&= -(n \cdot m) \\
(-m) \cdot (-n) &= m \cdot n \\
-m &< 0 \quad \forall m \in \mathbb{N}^* \\
-m &\leq -n \iff n \leq m.
\end{aligned}$$

Man kann zeigen:

SATZ I.2.1. (1) *Die Addition auf* $\mathbb{Z}$ *ist kommutativ und assoziativ.* 0 *ist das neutrale Element. Für jedes* $p \in \mathbb{Z}$ *gilt:*

$$p + (-p) = 0.$$

$(\mathbb{Z}, +)$ *ist eine* ABELSCHE GRUPPE.
(2) *Die Multiplikation auf* $\mathbb{Z}$ *ist kommutativ und assoziativ mit neutralem Element* 1. $(\mathbb{Z}, \cdot)$ *ist eine* KOMMUTATIVE HALBGRUPPE.
(3) *Es gilt das Distributivgesetz.* $(\mathbb{Z}, +, \cdot)$ *ist ein* RING.
(4) *Die Ordnung* $\leq$ *ist reflexiv, transitiv und antisymmetrisch.* $(\mathbb{Z}, \leq)$ *ist totalgeordnet.*
(5)

$$\begin{aligned}
p \leq q &\iff \exists\, l \in \mathbb{N} : p + l = q \\
p < q &\iff \exists\, l \in \mathbb{N}^* : p + l = q
\end{aligned}$$

(6) $\forall p \in \mathbb{Z} \ \nexists q \in \mathbb{Z}$ *mit*

$$p < q < p + 1.$$

(7) $p \leq q \iff p + r \leq q + r \quad \forall r \in \mathbb{Z}$
(8) $\forall p, q \in \mathbb{Z} \quad \forall k \in \mathbb{N}^*$ *gilt*

$$\begin{aligned}
p \leq q &\iff kp \leq kq \iff (-k)q \leq (-k)p \\
p < q &\iff kp < kq \iff (-k)q < (-k)p.
\end{aligned}$$

Satz I.1.3 (S. 6) kann in $\mathbb{Z}$ nicht mehr gelten, da

$$M = \{p \in \mathbb{Z} : p \leq 0\}$$

offensichtlich kein Minimum besitzt. Um eine adäquate Variante von Satz I.1.3 zu beweisen, benötigen wir eine Definition.

Definition I.2.2. Eine Teilmenge $M$ von $\mathbb{Z}$ heißt nach oben bzw. nach unten beschränkt, wenn es ein $p \in \mathbb{Z}$ gibt mit

$$p > q \quad \forall q \in M$$

bzw.

$$p < q \quad \forall q \in M.$$

Satz I.2.3. *Jede nach oben bzw. nach unten beschränkte Teilmenge $M$ von $\mathbb{Z}$ besitzt ein Maximum bzw. Minimum.*

Beweis. Sei $M$ nach oben beschränkt. Dann ist

$$N = \{n \in \mathbb{N} : n > q \quad \forall q \in M\}$$

nicht leer und besitzt nach Satz I.1.3 (S. 6) ein Minimum $n_0$.
Fall 1: $n_0 \geq 1$: Dann ist offensichtlich $n_0 - 1$ das gesuchte Maximum von $M$.
Fall 2: $n_0 = 0$: Dann ist

$$N' = \{n \in \mathbb{N} : (-n) \in M\}$$

nicht leer und besitzt nach Satz I.1.3 (S. 6) ein Minimum $n_1$. Wegen $n_0 = 0$ ist $(-n_1)$ das gesuchte Maximum von $M$.
Falls $M$ nach unten beschränkt ist, folgt die Behauptung aus dem Gezeigten und der leicht zu beweisenden Beziehung

$$\min M = -\max\{-m : m \in M\}.$$

□

Wir haben $\mathbb{Z}$ in gewissem Sinn durch „Verdoppelung“ von $\mathbb{N}$ erhalten. Der folgende Satz zeigt, dass $\mathbb{N}$ und $\mathbb{Z}$ die gleiche Mächtigkeit haben.

Satz I.2.4. *$\mathbb{Z}$ ist abzählbar.*

Beweis. Definiere $f : \mathbb{N} \to \mathbb{Z}$ durch

$$\begin{aligned} f(0) &= 0, \\ f(2n) &= n, \\ f(2n-1) &= -n. \end{aligned}$$

Offensichtlich ist $f$ surjektiv. □

## I.3. Die rationalen Zahlen

Die ganzen Zahlen erlauben uns die Lösung von Gleichungen der Form
$$p + q = r.$$
Jedoch gibt es keine ganze Zahl p mit
$$2p = 1.$$
Denn es gilt:
$$\begin{aligned}
&p = 0 \Longrightarrow 2p = 0 \neq 1.\\
&p > 0 \Longrightarrow p \geq 1 \Longrightarrow 2p \geq 2 > 1.\\
&p < 0 \Longrightarrow p \leq -1 \Longrightarrow 2p \leq -2 < 1.
\end{aligned}$$
Um dieses Manko zu beheben, definieren wir für $p \in \mathbb{Z}$ und $q \in \mathbb{Z}^* = \mathbb{Z}\backslash\{0\}$ als $\frac{p}{q}$ die Zahl $x$, die die Gleichung
$$qx = p$$
löst, und setzen
$$\mathbb{Q} = \left\{ \frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{Z}^* \right\}.$$
Bei dieser Definition muss man natürlich alle Brüche der Form $\frac{(rp)}{(rq)}$ mit $r \in \mathbb{Z}^*$ mit $\frac{p}{q}$ identifizieren. Wir gehen hier auf diese Details aber nicht ein, da wir nur an der prinzipiellen Konstruktion interessiert sind.
Wir setzen die Addition, Multiplikation und die Ordnung $\leq$ wie folgt auf $\mathbb{Q}$ fort:
$$\begin{aligned}
\frac{p}{q} + \frac{r}{s} &= \frac{ps + rq}{qs}\\
\frac{p}{q} \cdot \frac{r}{s} &= \frac{pr}{qs}\\
\frac{p}{q} \leq \frac{r}{s} &\quad \Longleftrightarrow \quad sp \leq rq
\end{aligned}$$
Man kann zeigen:

Satz I.3.1. (1) *Die Addition auf* $\mathbb{Q}$ *ist kommutativ und assoziativ mit neutralem Element* $0$. *Zu jedem* $x \in \mathbb{Q}$ *gibt es ein* $y \in \mathbb{Q}$ *mit* $x + y = 0$. $(\mathbb{Q}, +)$ *ist eine* Abelsche Gruppe.
(2) *Die Multiplikation auf* $\mathbb{Q}$ *ist kommutativ und assoziativ mit neutralem Element* $1$. *Zu jedem* $x \in \mathbb{Q}^* = \mathbb{Q}\backslash\{0\}$ *gibt es ein* $y \in \mathbb{Q}^*$ *mit* $x \cdot y = 1$. $(\mathbb{Q}^*, \cdot)$ *ist eine* Abelsche Gruppe.
(3) *Es gilt das Distributivgesetz.* $(\mathbb{Q}, +, \cdot)$ *ist ein* Körper.
(4) *Die Relation* $\leq$ *ist reflexiv, transitiv und symmetrisch. Für je zwei* $x, y \in \mathbb{Q}$ *ist genau eine der Beziehungen*
$$x < y \ , \ x = y \ , \ y < x$$

*erfüllt. Es gilt*

(a) $$x < y \Longrightarrow x + z < y + z \quad \forall z \in \mathbb{Q}$$

(b) $$x > 0, y > 0 \Longrightarrow x \cdot y > 0.$$

$(\mathbb{Q}, +, \cdot, \leq)$ *ist ein* ANGEORDNETER KÖRPER.

Der folgende Satz zeigt, dass $\mathbb{Q}$ die gleiche Mächtigkeit wie $\mathbb{Z}$ und $\mathbb{N}$ hat. Die Beweisidee ist als sog. DIAGONALVERFAHREN bekannt.

SATZ I.3.2. $\mathbb{Q}$ *ist abzählbar.*

BEWEIS. Gemäß Satz I.2.4 (S. 15) ist $\mathbb{Z}$ abzählbar, d.h., wir können $\mathbb{Z}$ in der Form

$$\mathbb{Z} = \{p_i : i \in \mathbb{N}\}$$

darstellen, wobei wir o.E. $p_0 = 0$ wählen können. Wir ordnen nun die Elemente von $\mathbb{Q}$ in folgendem quadratischen Schema an und zählen sie wie angedeutet ab:

$$
\begin{array}{ccccccccc}
0 & & 1 & & 2 & & 6 & & \\
0 & \longrightarrow & \frac{p_1}{p_1} & \longrightarrow & \frac{p_2}{p_1} & & \frac{p_3}{p_1} & \longrightarrow & \cdots \\
 & & 3 & \swarrow & 5 & \nearrow & & & \cdots \\
 & & \frac{p_1}{p_2} & & \frac{p_2}{p_2} & & \frac{p_3}{p_2} & & \cdots \\
 & & \downarrow & \nearrow & & & & & \\
 & & 4 & & & & & & \\
 & & \frac{p_1}{p_3} & & \frac{p_2}{p_3} & & \frac{p_3}{p_3} & & \cdots \\
 & & \vdots & & \vdots & & \vdots & & \cdots
\end{array}
$$

□

## I.4. Die reellen Zahlen

SATZ I.4.1. *Es gibt keine rationale Zahl $x$ mit*

$$x^2 = 2.$$

BEWEIS. Angenommen, $x \in \mathbb{Q}$ löse obige Gleichung. Dann ist $x \neq 0$. O.E. können wir $x > 0$ voraussetzen, sonst gehen wir zu $-x$ über. Wir stellen $x$ dar als

$$x = \frac{p}{q}$$

mit teilerfremden Zahlen $p, q \in \mathbb{Z}^*$. Dann folgt

$$x^2 = 2 \Longrightarrow \frac{p^2}{q^2} = 2 \Longrightarrow p^2 = 2q^2.$$

Da das Quadrat einer ungeraden Zahl stets ungerade ist, folgt $p = 2n$ mit $n \in \mathbb{N}^*$. Also

$$2q^2 = p^2 = 4n^2 \Longrightarrow q^2 = 2n^2.$$

Mithin ist $q = 2m$ mit $m \in \mathbb{N}^*$, im Widerspruch zu Teilerfremdheit von $p$ und $q$. □

Andererseits können wir eine Lösung der Gleichung $x^2 = 2$ beliebig gut durch rationale Zahlen „approximieren“. Um dies einzusehen, definieren wir Zahlen $a_n, b_n, n \in \mathbb{N}$, rekursiv wie folgt (INTERVALLSCHACHTELUNG):

$$a_0 = 1, b_0 = 2$$
$$\begin{cases} a_{n+1} = \frac{1}{2}(a_n + b_n),\ b_{n+1} = b_n, & \text{falls } \frac{(a_n+b_n)^2}{4} < 2 \\ a_{n+1} = a_n,\ b_{n+1} = \frac{1}{2}(a_n + b_n), & \text{falls } \frac{(a_n+b_n)^2}{4} > 2. \end{cases}$$

(Man beachte, dass wegen Satz I.4.1 der Fall $\frac{(a_n+b_n)^2}{4} = 2$ nicht auftreten kann.)
Durch Induktion zeigt man leicht:

$$a_n, b_n \in \mathbb{Q} \quad \forall n \in \mathbb{N} \tag{1}$$
$$a_n < b_n \quad \forall n \in \mathbb{N} \tag{2}$$
$$a_n^2 < 2 \quad \forall n \in \mathbb{N} \tag{3}$$
$$b_n^2 > 2 \quad \forall n \in \mathbb{N} \tag{4}$$
$$b_n - a_n \leq 2^{-n} \quad \forall n \in \mathbb{N} \tag{5}$$

Wenn wir uns die rationalen Zahlen auf einer Geraden angeordnet denken, zeigt obige Konstruktion, dass diese Gerade „Lücken“ hat. Um diese Beobachtung zu präzisieren, benötigen wir einige Bezeichnungen.

DEFINITION I.4.2. Sei $(K, +, \cdot, \leq)$ ein angeordneter Körper und $M \subset K$ nicht leer.
(1) Die Menge $M$ heißt NACH OBEN bzw. NACH UNTEN BESCHRÄNKT, wenn es ein $x \in K$ gibt mit

$$y \leq x \quad \text{bzw.} \quad x \leq y$$

für alle $y \in M$. Jedes derartige $x$ heißt OBERE bzw. UNTERE SCHRANKE von $M$.
(2) Falls $M$ nach oben bzw. nach unten beschränkt ist, heißt ein $x \in K$ SUPREMUM bzw. INFIMUM von $M$, kurz $x = \sup M$ bzw. $x = \inf M$, wenn $x$ eine obere bzw. untere Schranke von $M$ ist und für jede andere obere bzw. untere Schranke $x^*$ von $M$ gilt

$$x \leq x^* \quad \text{bzw.}\ x^* \leq x.$$

BEMERKUNG I.4.3. Infimum und Supremum sind eindeutig.

DEFINITION I.4.4. Ein angeordneter Körper heißt ORDNUNGSVOLLSTÄNDIG, wenn jede nach oben beschränkte Teilmenge ein Supremum hat.

LEMMA I.4.5. *Sei $K$ ein angeordneter Körper. Folgende Aussagen sind äquivalent:*
(1) *$K$ ist ordnungsvollständig.*

(2) *Jede nach unten beschränkte Teilmenge von $K$ besitzt ein Infimum.*
(3) *Für je zwei nicht leere Teilmengen $A, B$ von $K$ mit*

$$a \leq b \quad \forall a \in A, b \in B$$

*gibt es ein $c \in K$ mit*

$$a \leq c \leq b \quad \forall a \in A, b \in B.$$

Beweis. (1) $\Longrightarrow$ (2): Sei $A \subset K, A \neq \emptyset$ nach unten beschränkt. Definiere

$$B = \{x \in K : x \leq a \quad \forall a \in A\}.$$

Dann ist B nicht leer und nach oben beschränkt. Wegen (1) existiert $b = \sup B$. Konstruktionsgemäß ist $b = \inf A$.
(2) $\Longrightarrow$ (3): Aus (2) folgt, dass $m = \inf B$ existiert. Wegen $a \leq b$ für alle $a \in A, b \in B$ folgt

$$a \leq m \leq b \quad \forall a \in A, b \in B.$$

(3) $\Longrightarrow$ (1): Sei $A \subset K, A \neq \emptyset$ nach oben beschränkt. Definiere

$$B = \{x \in K : x \geq a \quad \forall a \in A\}.$$

Aus (3) folgt, dass es ein $c \in K$ gibt mit

$$a \leq c \leq b \quad \forall a \in A, b \in B.$$

Aus Definition I.4.2 folgt

$$c = \sup A.$$

□

Wir können nun unsere oben gewonnene intuitive Vorstellung von der Lückenhaftigkeit von $\mathbb{Q}$ präzisieren.

Satz I.4.6. *$\mathbb{Q}$ ist nicht ordnungsvollständig.*

Beweis. Wir nehmen das Gegenteil an und definieren

$$A = \{x \in \mathbb{Q} : x > 0 \ \text{ und } x^2 < 2\}$$
$$B = \{x \in \mathbb{Q} : x > 0 \ \text{ und } x^2 > 2\}.$$

Dann ist $1 \in A$ und $2 \in B$, und für $a \in A, b \in B$ gilt

$$b - a = \frac{b^2 - a^2}{b + a} > 0 \Longrightarrow b > a.$$

Wegen unserer Annahme und wegen Lemma I.4.5 gibt es ein $c \in \mathbb{Q}$ mit $a \leq c \leq b$ für alle $a \in A, b \in B$. Definiere

$$x = c - \frac{c^2 - 2}{c + 2}.$$

Dann folgt

$$x = \frac{2c + 2}{c + 2} > 0 \text{ (wegen } c > 0)$$

$$x^2 - 2 = \frac{2c^2 - 4}{(c+2)^2} = \frac{2(c^2-2)}{(c+2)^2}.$$

1. Fall $c^2 > 2$: $\Longrightarrow x^2 > 2$ und $x < c \Longleftrightarrow x \in B$ und $x < c$ Widerspruch!
2. Fall $c^2 < 2$: $\Longrightarrow x > c$ und $x^2 < 2 \Longleftrightarrow x \in A$ und $x > c$ Widerspruch!
Also ist $c^2 = 2$ im Widerspruch zu Satz I.4.1. $\square$

Wir kommen nun zum Hauptsatz über die reellen Zahlen. Sein Beweis stammt von R. Dedekind (1832-1916). Wir skizzieren hier nur die wesentlichen Punkte der Dedekindschen Konstruktion und verweisen für Details auf das Buch von Landau.

Satz I.4.7 (Dedekindscher Hauptsatz). *Es gibt (bis auf Isomorphie) genau einen ordnungsvollständigen Körper, den Körper $\mathbb{R}$ der* reellen Zahlen, *der $\mathbb{Q}$ enthält und auf $\mathbb{Q}$ die natürliche Ordnung induziert.*

Beweisidee. (1) Wir nennen eine nicht leere Teilmenge $S$ von $\mathbb{Q}$ einen Schnitt, wenn folgende Bedingungen erfüllt sind:

(i) $s < q \quad \forall s \in S, q \in \mathbb{Q}\backslash S$
(ii) $\nexists\, s \in S$ mit $s \geq t \quad \forall t \in S$.

Schnitte sind also nach oben beschränkte Teilmengen von $\mathbb{Q}$, die ihr Supremum (sofern überhaupt existent) nicht enthalten.
(2) Zwei Schnitte heißen gleich, wenn Sie im mengentheoretischen Sinn gleich sind.
(3) Man definiert $S < T$, falls es ein $t \in T$ gibt mit $t > s$ für alle $s \in S$.
(4) $S + T = \{s + t : s \in S, t \in T\}$ ist ein Schnitt.
(5) $S \cdot T = \{s \cdot t : s \in S, t \in T\}$ ist ein Schnitt.
(6) Man zeigt, dass die Menge aller Schnitte zusammen mit $+, \cdot$ und $\leq$ ein angeordneter Körper ist.
(7) $p \in \mathbb{Q} \Longrightarrow P = \{q \in \mathbb{Q} : q < p\}$ ist ein Schnitt. Die Schnitte „umfassen" also $\mathbb{Q}$.
(8) Die Menge der Schnitte ist ordnungsvollständig. $\square$

Bemerkung I.4.8. Wir haben folgende Kette natürlicher Inklusionen

$$\begin{array}{ccccccc} \mathbb{N} & \subset & \mathbb{Z} & \subset & \mathbb{Q} & \subset & \mathbb{R} \\ & & \text{Ring} & & \text{Körper} & & \text{ordnungsvollst.} \\ & & & & & & \text{Körper.} \end{array}$$

Für das Weitere ist es sinnvoll, einige Bezeichnungen einzuführen.

Definition I.4.9. (1) $\mathbb{R}^* = \mathbb{R}\backslash\{0\}$, $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$, $\mathbb{R}_+^* = \mathbb{R}_+\backslash\{0\}$.
(2) Wir definieren zwei Symbole $-\infty$ und $+\infty$ mit der Eigenschaft

$$-\infty < x < +\infty \quad \forall x \in \mathbb{R}.$$

(3) $\sup \emptyset = -\infty$, $\inf \emptyset = +\infty$.
(4) Sei $M \subset \mathbb{R}$ nicht leer. Dann setzen wir

$$\inf M = -\infty \iff M \text{ ist nicht nach unten beschränkt.}$$
$$\sup M = +\infty \iff M \text{ ist nicht nach oben beschränkt.}$$

SATZ I.4.10 (SATZ VON ARCHIMEDES). (1) $\mathbb{N}$ *ist in* $\mathbb{R}$ *nicht nach oben beschränkt, d.h., zu jedem* $x \in \mathbb{R}$ *gibt es ein* $n \in \mathbb{N}$ *mit* $n > x$.
(2) *Zu jedem* $x \in \mathbb{R}_+^*$ *gibt es ein* $n \in \mathbb{N}$ *mit*

$$\frac{1}{n} < x.$$

BEWEIS. AD (1): Offensichtlich ist nur für $x > 0$ etwas zu zeigen. Definiere

$$A = \{n \in \mathbb{N} : n \leq x\}.$$

Wegen $A \neq \emptyset$ existiert $s = \sup A$. Aus der Definition des Supremums folgt, dass es ein $a \in A$ mit $a > s - \frac{1}{2}$ gibt. Setze $m = a + 1 \in \mathbb{N}$. Wegen

$$m > s + \frac{1}{2}$$

ist $m \notin A$ und somit

$$x < m.$$

AD (2): Wegen $x > 0$ folgt die Behauptung aus Teil (a) angewandt auf $\frac{1}{x}$. □

SATZ I.4.11. *Die rationalen und die irrationalen Zahlen, d.h.* $\mathbb{Q}$ *und* $\mathbb{R}\backslash\mathbb{Q}$*, liegen dicht in* $\mathbb{R}$*, d.h., zu* $a, b \in \mathbb{R}$ *mit* $a < b$ *gibt es ein* $q \in \mathbb{Q}$ *und ein* $x \in \mathbb{R}\backslash\mathbb{Q}$ *mit*

$$a < q < b \quad \text{und} \quad a < x < b.$$

BEWEIS. Seien $a, b \in \mathbb{R}$ mit $a < b$. Wegen Satz I.4.10 gibt es ein $n \in \mathbb{N}$ mit

$$n > \frac{1}{b-a}.$$

Also ist

$$nb > na + 1.$$

Ebenso folgt aus Satz I.4.10, dass es ein $m \in \mathbb{Z}$ gibt mit

$$m - 1 \leq na < m.$$

Damit folgt

$$na < m \leq na + 1 < nb$$

also

$$a < \frac{m}{n} < b.$$

Aus dem soeben Gezeigten folgt die Existenz von zwei rationalen Zahlen $r_1, r_2$ mit

$$a < r_1 < r_2 < b.$$

Definiere

$$x = r_1 + \frac{r_2 - r_1}{\sqrt{2}} > r_1.$$

Dann ist $x \in \mathbb{R} \backslash \mathbb{Q}$. Denn andernfalls wäre

$$\sqrt{2} = \frac{r_2 - r_1}{x - r_1}$$

aus $\mathbb{Q}$ im Widerspruch zu Satz I.4.1. Außerdem ist

$$r_2 - x = (r_2 - r_1)(1 - \frac{1}{\sqrt{2}}) > 0.$$

Also ist

$$a < r_1 < x < r_2 < b.$$

□

SATZ I.4.12. *Für alle $a \in \mathbb{R}_+$ und alle $n \in \mathbb{N}^*$ existiert genau ein $x \in \mathbb{R}_+$ mit $x^n = a$.*

BEWEIS. EINDEUTIGKEIT: Seien $x, y$ zwei Lösungen. O.E. ist $x < y$. Durch Induktion über $n$ folgt $x^n < y^n$.
EXISTENZ: Die Fälle $n = 1$ und $a \in \{0, 1\}$ sind offensichtlich. Sei also $n \geq 2$ und $a \neq 0, a \neq 1$.
FALL $a > 1$: Definiere

$$A = \{x \in \mathbb{R}_+ : x^n \leq a\}.$$

Aus $a > 1$ folgt

$$1 \in A$$

und

$$x \geq a \Longrightarrow x^n \geq a^n > a$$

und somit

$$x \leq a \quad \forall x \in A.$$

Also existiert

$$s = \sup A.$$

Wir wollen zeigen: $s^n = a$.
ANN. $s^n < a$: Definiere

$$b = \sum_{k=0}^{n-1} \binom{n}{k} s^k > 0$$

und wähle

$$0 < \varepsilon < \min\{1, \frac{a - s^n}{b}\}.$$

Dann folgt

$$(s + \varepsilon)^n = \sum_{k=0}^{n-1} \binom{n}{k} s^k \varepsilon^{n-k} + s^n$$

$$\begin{aligned}&\leq \varepsilon \sum_{k=0}^{n-1} \binom{n}{k} s^k + s^n \\ &= \varepsilon b + s^n \\ &< a\end{aligned}$$

Im Widerspruch zu $s = \sup A$.
Ann. $s^n > a$: Definiere

$$c = \sum_{\substack{j \\ 0<2j-1\leq n}} \binom{n}{2j-1} s^{n-2j+1} > 0$$

und wähle

$$0 < \varepsilon < \min\{1, \frac{s^n - a}{c}\}.$$

Dann folgt

$$\begin{aligned}(s-\varepsilon)^n &= s^n + \sum_{k=0}^{n-1} (-1)^{n-k} \binom{n}{k} s^k \varepsilon^{n-k} \\ &= s^n + \sum_{l=1}^{n} (-1)^l \underbrace{\binom{n}{n-l}}_{=\binom{n}{l}} s^{n-l} \varepsilon^l \\ &= s^n + \sum_{\substack{j \\ 0<2j\leq n}} \binom{n}{2j} s^{n-2j} \varepsilon^{2j} - \sum_{\substack{j \\ 0<2j-1\leq n}} \binom{n}{2j-1} s^{n-2j+1} \varepsilon^{2j-1} \\ &\geq s^n - \varepsilon \sum_{\substack{j \\ 0<2j-1\leq n}} \binom{n}{2j-1} s^{n-2j+1} \\ &= s^n - \varepsilon c \\ &> a\end{aligned}$$

im Widerspruch zu $s = \sup A$.
Also ist $s^n = a$.
Fall $a < 1$: Gemäß Obigem gibt es genau ein $y$ mit $y^n = \frac{1}{a}$. $x = \frac{1}{y}$ ist die gesuchte Lösung. □

Bemerkung I.4.13. (1) Für $a \in \mathbb{R}_+^*$ und $n = 2k$ mit $k \in \mathbb{N}^*$ hat die Gleichung $x^n = a$ genau zwei Lösungen in $\mathbb{R}$. Nämlich $x = y$ und $x = -y$, wobei $y \in \mathbb{R}_+^*$ die eindeutige Lösung in $\mathbb{R}_+$ ist.
(2) Für $n = 2k+1$ mit $k \in \mathbb{N}$ hat die Gleichung $x^n = a$ für jedes $a \in \mathbb{R}$ eine eindeutige Lösung in $\mathbb{R}$. Für $a \in \mathbb{R}_+$ folgt dies aus Satz I.4.12, für $a < 0$ aus Satz I.4.12 angewandt auf $-a$.
(3) Sei $a \in \mathbb{R}_+$ und $x \in \mathbb{R}_+$ die eindeutige Lösung von $x^n = a$. Dann schreibt man $x = a^{\frac{1}{n}} = \sqrt[n]{a}$. Für $m \in \mathbb{Z}$ ist

$$a^{\frac{m}{n}} = (a^{\frac{1}{n}})^m.$$

Bemerkung I.4.14. Eine Zahl $x \in \mathbb{R}$ heißt algebraisch, wenn es ein $n \in \mathbb{N}$ und rationale Zahlen $a_0, \ldots, a_n$ gibt mit

$$a_0 + a_1 x + \ldots + a_n x^n = 0.$$

Insbesondere sind alle rationalen Zahlen algebraisch. $\sqrt{2}$ ist algebraisch, aber nicht rational. Interessanterweise gibt es aber auch reelle Zahlen, die nicht algebraisch sind. Solche Zahlen heißen transzendent. $\pi$ ist z.B. transzendent.

Der folgende Satz zeigt, dass die Dedekindsche Konstruktion die Menge $\mathbb{Q}$ wesentlich vergrößert.

Satz I.4.15. *$\mathbb{R}$ ist nicht abzählbar.*

Beweis. Es reicht zu zeigen, dass die Menge

$$[0,1] = \{x \in \mathbb{R} : 0 \le x \text{ und } x \le 1\}$$

nicht abzählbar ist.
Wir nehmen das Gegenteil an. Dann ist

$$[0,1] = \{x_0, x_1, \ldots\},$$

wobei o.E. $x_0 = 0$ gilt. Wir können jedes $x_n$ als (unendlichen) Dezimalbruch in der Form

$$x_n = 0.a_{n1}a_{n2}\ldots$$

mit $a_{nk} \in \mathbb{N}_9$ und $a_{nk} \neq 0$ für ein $k \in \mathbb{N}^*$ darstellen. Wir definieren eine neue Zahl

$$c = 0.c_1c_2\ldots$$

wie folgt

$$c_n = \begin{cases} 5, & \text{falls } a_{nn} \neq 5 \\ 4, & \text{falls } a_{nn} = 5 \end{cases}$$

Offensichtlich ist $0 < c$ und $c < 1$ und $c \neq x_n$ für alle $n \in \mathbb{N}$. Widerspruch. □

Zum Abschluss dieses Paragraphen führen wir noch einige nützliche Bezeichnungen ein.

Definition I.4.16. Der Betrag einer reellen Zahl $x$ ist definiert als

$$|x| = \max\{x, -x\} = \begin{cases} x & \text{falls } x \ge 0 \\ -x & \text{falls } x < 0 \end{cases}$$

Satz I.4.17. *Es seien $x, y \in \mathbb{R}$ und $\varepsilon \in \mathbb{R}_+^*$. Dann gilt:*

(1) $|x| = |-x|$.
(2) $|x| \ge 0, |x| = 0 \iff x = 0$.
(3) $|x + y| \le |x| + |y|$ (Dreiecksungleichung).
(4) $||x| - |y|| \le |x - y|$.
(5) $|x - y| \le \varepsilon \iff y - \varepsilon \le x \le y + \varepsilon$.
(6) $|x \cdot y| = |x| \cdot |y|$.

Beweis. Übungsaufgabe. $\square$

Definition I.4.18. Seien $a, b \in \mathbb{R}$ mit $a < b$. Dann definieren wir:

(1) $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ (abgeschlossenes Intervall).
(2) $(a, b) = [a, b]\backslash\{a, b\}$ (offenes Intervall).
(3) $[a, b) = [a, b]\backslash\{b\}$.
(4) $(a, b] = [a, b]\backslash\{a\}$.

## I.5. Die komplexen Zahlen

Da für alle $x \in \mathbb{R}$ gilt $x^2 \geq 0$, besitzt die Gleichung

$$x^2 = -1$$

keine reelle Lösung. Wir wollen einen möglichst kleinen Oberkörper von $\mathbb{R}$ konstruieren, in dem diese Gleichung lösbar ist. Dazu gehen wir wie folgt vor: Wir definieren formal die imaginäre Einheit $i$ durch die Beziehung

$$i^2 = -1$$

und definieren die komplexen Zahlen $\mathbb{C}$ durch

$$\mathbb{C} = \{x + iy : x, y \in \mathbb{R}\}.$$

Dabei vereinbaren wir sinnvollerweise

$$0 \cdot i = 0,$$

so dass $\mathbb{R}$ in $\mathbb{C}$ enthalten ist. Wir setzen die Addition und Multiplikation auf $\mathbb{C}$ durch

$$\begin{aligned}(x + iy) + (u + iv) &= (x + u) + i(y + v)\\ (x + iy) \cdot (u + iv) &= (xu - yv) + i(xv + uy)\end{aligned}$$

fort. Man kann zeigen:

Satz I.5.1. (1) *$(\mathbb{C}, +, \cdot)$ ist ein Oberkörper von $\mathbb{R}$, in dem die Gleichung $z^2 = -1$ genau die Lösungen $z = \pm i$ hat.*
(2) *Jeder Oberkörper von $\mathbb{R}$, in dem die Gleichung $z^2 = -1$ lösbar ist, ist auch ein Oberkörper von $\mathbb{C}$.*

Der Beweis von Teil (1) ist langweilig. Beim Nachweis der Körpereigenschaften, muss man u.a. zeigen, dass für $x, y \in \mathbb{R}^*$ gilt

$$\frac{1}{x + iy} = \frac{x - iy}{x^2 + y^2},$$

so dass wir die folgende Divisionsregel erhalten

$$\frac{u + iv}{x + iy} = \frac{xu + yv}{x^2 + y^2} + i\frac{xv - yu}{x^2 + y^2}.$$

Bemerkung I.5.2. Stellen wir komplexe Zahlen $z = x + iy$ als Vektoren in der $x, y$-Ebene dar, so entspricht die Addition zweier komplexer Zahlen der Addition der entsprechenden Vektoren. Die Multiplikation $w \cdot z$ entspricht einer Drehung um den von $z$ und der $x$-Achse eingeschlossenen Winkel gefolgt von einer Streckung um den Faktor $\sqrt{x^2 + y^2}$ (vgl. Abb. I.5.1).

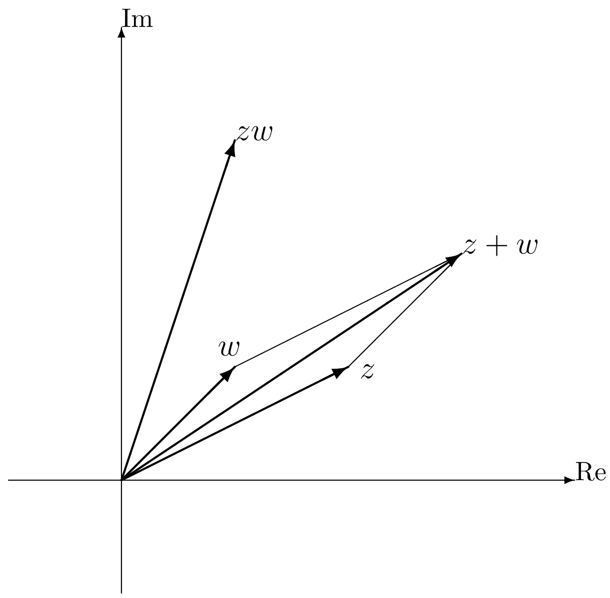


Abbildung I.5.1. Addition und Multiplikation von $z = 2 + i$ und $w = 1 + i$

Bemerkung I.5.3. $\mathbb{C}$ kann nicht angeordnet werden. Denn sonst würde für jedes $z \in \mathbb{C}$ gelten

$$z^2 \geq 0 \neq -1.$$

Definition I.5.4. Sei $z = x + iy, x, y \in \mathbb{R}$, eine komplexe Zahl. Dann heißen

$$\begin{aligned} \text{Re } z = x &\quad \text{der Realteil von } z, \\ \text{Im } z = y &\quad \text{der Imaginärteil von } z \\ |z| = \sqrt{x^2 + y^2} &\quad \text{der Betrag von } z, \\ \overline{z} = x - iy &\quad \text{die zu } z \text{ konjugierte Zahl.} \end{aligned}$$

Satz I.5.5. *Es gelten folgende Rechenregeln* ($z = x + iy, w = u + iv, x, y, u, v \in \mathbb{R}$)*:*

(1) $\text{Re } z = \frac{1}{2}(z + \overline{z})$, $\text{Im } z = \frac{1}{2i}(z - \overline{z})$
(2) $z \in \mathbb{R} \iff z = \overline{z}$
(3) $\overline{(\overline{z})} = z$
(4) $\overline{z + w} = \overline{z} + \overline{w}, \overline{z \cdot w} = \overline{z} \cdot \overline{w}$

(5) $|z|^2 = z \cdot \overline{z}$
(6) $|z| \geq 0, |z| = 0 \Longleftrightarrow z = 0$
(7) $z \in \mathbb{R} \Longrightarrow |z|_{\mathbb{C}} = |z|_{\mathbb{R}}$
(8) $|z + w| \leq |z| + |w|$ (Dreiecksungleichung)
(9) $|z \cdot w| = |z| \cdot |w|$
(10) $||z| - |w|| \leq |z - w|$
(11) $|\operatorname{Re}\ z| \leq |z|, |\operatorname{Im}\ z| \leq |z|$
(12) $z \neq 0 \Longrightarrow \frac{1}{z} = \frac{\overline{z}}{|z|^2}$

Beweis. Übungsaufgabe. □

Übereinkunft: Im Folgenden bezeichnen wir mit $\mathbb{K}$ stets $\mathbb{R}$ oder $\mathbb{C}$.

KAPITEL II

# Folgen und Reihen

In diesem Kapitel betrachten wir Folgen und Reihen in $\mathbb{K}$. Wesentlich ist der Grenzwertbegriff und die Vollständigkeit von $\mathbb{K}$.
Die Ergebnisse sind einerseits von eigenständigem Interesse, indem sie wichtige Beweistechniken einführen. Andererseits sind sie Einleitung zu späteren Kapiteln, indem sie z.B. nichttriviale Beispiele normierter Vektorräume liefern oder den Begriff der Vollständigkeit motivieren.

## II.1. Konvergenz von Folgen

DEFINITION II.1.1. Eine FOLGE (in $\mathbb{K}$) ist eine Abbildung von $\mathbb{N}$ in $\mathbb{K}$; jedem $n \in \mathbb{N}$ ist ein Wert $x_n \in \mathbb{K}$ zugeordnet. Wir schreiben hierfür $(x_0, x_1, \ldots)$ oder kürzer $(x_n)_{n\in\mathbb{N}}$.

BEMERKUNG II.1.2. Die Folge $(x_n)_{n\in\mathbb{N}}$ ist von der Punktmenge $\{x_n : n \in \mathbb{N}\}$ zu unterscheiden.

DEFINITION II.1.3. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $a \in \mathbb{K}$. Dann heißt $a$ HÄUFUNGSPUNKT (kurz HP) von $(x_n)_{n\in\mathbb{N}}$, wenn es zu jedem $\varepsilon > 0$ eine unendliche Teilmenge $N_\varepsilon$ von $\mathbb{N}$ gibt mit

$$|x_n - a| < \varepsilon \quad \forall n \in N_\varepsilon.$$

Der folgende Satz gibt eine äquivalente, häufig praktischere Definition des HP. Sein Beweis ist offensichtlich.

SATZ II.1.4. *Folgende Aussagen sind äquivalent:*

(1) *$a$ ist HP von $(x_n)_{n\in\mathbb{N}}$.*
(2) $\forall \varepsilon > 0\ \forall n \in \mathbb{N}\ \exists m \geq n : |x_m - a| < \varepsilon$.

BEISPIEL II.1.5. (1) $(x_n)_{n\in\mathbb{N}}$ mit $x_n = a$ für alle $n \in \mathbb{N}$ hat den HP $a$.
(2) $(x_n)_{n\in\mathbb{N}}$ mit $x_n = \frac{1}{n}$ hat den HP 0.
(3) $((-1)^n)_{n\in\mathbb{N}}$ hat die HPe 1 und $-1$.
(4) $(i^n)_{n\in\mathbb{N}}$ hat die HPe $1, -1, i, -i$.
(5) $(x_n)_{n\in\mathbb{N}}$ mit $x_n = n$ hat keinen HP.
(6) $(x_n)_{n\in\mathbb{N}}$ mit $x_0 = 1, x_1 = 1$ und

$$x_n = x_{n-1} + x_{n-2} \quad \forall n \geq 2$$

hat keinen HP. (FIBONACCI FOLGE)
(7) $(x_n)_{n\in\mathbb{N}}$ sei eine Abzählung von $\mathbb{Q}$. Dann ist jedes $a \in \mathbb{R}$ ein HP.

Beweis. (1) Ist offensichtlich.
(2) Folgt aus Satz I.4.10 (S. 21).
(3)–(5) Trivial.
(6) Man zeigt durch Induktion $x_n \geq n$ für alle $n \geq 1$.
(7) Angenommen $a \in \mathbb{R}$ ist kein HP. Dann gibt es ein $\varepsilon_0 > 0$ und ein $n_0 \in \mathbb{N}$ mit

$$|x_n - a| \geq \varepsilon_0 \quad \forall n > n_0.$$

Insbesondere ist

$$\delta = \min\{\varepsilon_0, \min\{|x_n - a| : n \leq n_0, x_n \neq a\}\}$$

wohldefiniert und positiv. Dann enthält $(a, a + \frac{\delta}{2})$ keine rationale Zahl im Widerspruch zu Satz I.4.11 (S. 21). □

Definition II.1.6. Eine Folge $(x_n)_{n\in\mathbb{N}}$ heißt konvergent, wenn es ein $a \in \mathbb{K}$ gibt mit:

$$\forall \varepsilon > 0 \, \exists \, n_\varepsilon \in \mathbb{N} \quad \forall n \geq n_\varepsilon : |x_n - a| < \varepsilon.$$

$a$ heißt dann der Grenzwert von $(x_n)_{n\in\mathbb{N}}$. Wir schreiben kurz

$$x_n \underset{n\to\infty}{\longrightarrow} a \text{ oder } a = \lim_{n\to\infty} x_n.$$

Eine Folge, die keinen Grenzwert hat, heißt divergent.

Beispiel II.1.5 (Forts.). (1) $\lim\limits_{n\to\infty} x_n = a$
(2) $\lim\limits_{n\to\infty} x_n = 0$.
(3)–(7) Die Folgen sind divergent.

Definition II.1.7. Eine Folge $(x_n)_{n\in\mathbb{N}}$ heißt beschränkt, wenn es ein $c \in \mathbb{R}_+$ gibt mit

$$|x_n| \leq c \quad \forall n \in \mathbb{N}.$$

Satz II.1.8. *Jede konvergente Folge ist beschränkt.*

Beweis. Sei $a = \lim\limits_{n\to\infty} x_n$. Dann existiert $n_1 \in \mathbb{N}$ mit

$$|x_n - a| \leq 1 \quad \forall n \geq n_1.$$

Also ist

$$|x_n| \leq |a| + |x_n - a| \leq |a| + 1 \quad \forall n \geq n_1.$$

Setze

$$c = \max\{|a| + 1, \max\{|x_n| : n < n_1\}\}.$$

Dann folgt

$$|x_n| \leq c \quad \forall n \in \mathbb{N}.$$

□

Beispiel II.1.9. Beispiele II.1.5 (3) und (4) zeigen, dass die Umkehrung von Satz II.1.8 i. a. falsch ist.

Satz II.1.10. *Es gelte* $x_n \underset{n\to\infty}{\longrightarrow} a$. *Dann ist* $a$ *der einzige HP von* $(x_n)_{n\in\mathbb{N}}$. *Insbesondere ist der Grenzwert einer konvergenten Folge eindeutig bestimmt.*

Beweis. Aus den Definitionen II.1.3 und II.1.6 folgt, dass $a$ HP von $(x_n)_{n\in\mathbb{N}}$ ist. Sei $b \in \mathbb{K}\backslash\{a\}$. Setze $\varepsilon = \frac{1}{2}|b-a|$. Dann gibt es ein $n_0 \in \mathbb{N}$ mit

$$|x_n - a| < \varepsilon \quad \forall n \geq n_0.$$

Also ist

$$\begin{aligned} |x_n - b| &\geq ||b-a| - |x_n - a|| \\ &\geq |b-a| - |x_n - a| \\ &> 2\varepsilon - \varepsilon \\ &= \varepsilon \qquad \forall n \geq n_0. \end{aligned}$$

Mithin ist $b$ kein HP von $(x_n)_{n\in\mathbb{N}}$. □

Bemerkung II.1.11. Die Folge $(\frac{1}{2}, 2, \frac{1}{3}, 3, \frac{1}{4}, 4, \ldots)$ zeigt, dass die Umkehrung von Satz II.1.10 i. a. falsch ist.

Definition II.1.12. Seien $(x_n)_{n\in\mathbb{N}}$ eine Folge und $\varphi : \mathbb{N} \to \mathbb{N}$ eine streng monoton wachsende Abbildung. Dann heißt

$$(x_{n_k})_{k\in\mathbb{N}} = (x_{\varphi(k)})_{k\in\mathbb{N}}$$

eine Teilfolge von $(x_n)_{n\in\mathbb{N}}$.

Beispiel II.1.13. Die Folge $((-1)^n)_{n\in\mathbb{N}}$ hat die konstanten Teilfolgen $((-1)^{2k})_{k\in\mathbb{N}}$ und $((-1)^{2k+1})_{k\in\mathbb{N}}$.

Satz II.1.14. *Folgende Aussagen sind äquivalent:*

(1) $x_n \underset{n\to\infty}{\longrightarrow} a$
(2) $x_{n_k} \underset{k\to\infty}{\longrightarrow} a$ *für jede Teilfolge* $(x_{n_k})_{k\in\mathbb{N}}$ *von* $(x_n)_{n\in\mathbb{N}}$.

Beweis. (1) $\Longrightarrow$ (2): Sei $(x_{n_k})_{k\in\mathbb{N}}$ eine beliebige Teilfolge und $\varepsilon > 0$. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit

$$|x_n - a| < \varepsilon \quad \forall n \geq n_\varepsilon.$$

Wegen der Monotonie von $\varphi$ gibt es ein $k_\varepsilon \in \mathbb{N}$ mit

$$n_k \geq n_\varepsilon \quad \forall k \geq k_\varepsilon.$$

Damit folgt die Behauptung.
(2) $\Longrightarrow$ (1): Wähle für $\varphi$ die identische Abbildung. □

Satz II.1.15. *Folgende Aussagen sind äquivalent:*

(1) $a$ *ist HP von* $(x_n)_{n\in\mathbb{N}}$.
(2) *Es gibt eine Teilfolge* $(x_{n_k})_{k\in\mathbb{N}}$ *von* $(x_n)_{n\in\mathbb{N}}$ *mit* $x_{n_k} \underset{k\to\infty}{\longrightarrow} a$.

Beweis. (1) $\Longrightarrow$ (2): Wir konstruieren eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ rekursiv. Setze $n_0 = 0$. Es seien $n_0 < n_1 < \ldots < n_{k-1}$ gewählt. Da $a$ HP von $(x_n)_{n\in\mathbb{N}}$ ist, existiert ein $n \geq n_{k-1} + 1$ mit

$$|x_n - a| < \frac{1}{k}.$$

Definiere

$$n_k = \min\{n : n > n_{k-1} \text{ und } |x_n - a| < \frac{1}{k}\}.$$

Offensichtlich ist die so konstruierte Abbildung $k \to n_k$ streng monoton wachsend.
Sei nun $\varepsilon > 0$. Dann existiert ein $k_\varepsilon \in \mathbb{N}$ mit

$$\frac{1}{k_\varepsilon} < \varepsilon.$$

Konstruktionsgemäß gilt

$$|x_{n_k} - a| < \frac{1}{k} \leq \frac{1}{k_\varepsilon} < \varepsilon \quad \forall k \geq k_\varepsilon.$$

Also gilt $x_{n_k} \underset{k\to\infty}{\longrightarrow} a$.
(2) $\Longrightarrow$ (1): Folgt direkt aus den Definitionen II.1.3, II.1.6 und II.1.12.
□

Als nächstes leiten wir einige sehr nützliche Regeln für das Rechnen mit konvergenten Folgen her. Dazu benötigen wir folgende Definition:

Definition II.1.16. Eine Folge $(x_n)_{n\in\mathbb{N}}$ heißt Nullfolge genau dann, wenn gilt

$$\lim_{n\to\infty} x_n = 0.$$

Lemma II.1.17. (1) *$(x_n)_{n\in\mathbb{N}}$ ist Nullfolge $\Longleftrightarrow$ $(|x_n|)_{n\in\mathbb{N}}$ ist Nullfolge.*
(2) *$x_n \underset{n\to\infty}{\longrightarrow} a \Longleftrightarrow (x_n - a)_{n\in\mathbb{N}}$ ist Nullfolge.*
(3) *$(x_n)_{n\in\mathbb{N}}$ sei eine Folge in $\mathbb{K}$ und $(r_n)_{n\in\mathbb{N}}$ sei eine Nullfolge in $\mathbb{R}_+$. Weiter gebe es ein $n_0 \in \mathbb{N}$ mit*

$$|x_n| \leq r_n \quad \forall n \geq n_0.$$

*Dann ist $(x_n)_{n\in\mathbb{N}}$ auch eine Nullfolge.*

Beweis. (1) und (2) sind klar.
ad(3): Sei $\varepsilon > 0$. Dann gibt es ein $n'_\varepsilon$ mit

$$|r_n| = r_n < \varepsilon \quad \forall n \geq n'_\varepsilon.$$

Setze $n_\varepsilon = \max\{n_0, n'_\varepsilon\}$. Dann folgt

$$|x_n| < r_n < \varepsilon \quad \forall n \geq n_\varepsilon.$$

□

Satz II.1.18. *Seien $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ zwei Folgen und $\alpha \in \mathbb{K}$ eine Zahl. Dann gilt:*

(1) $x_n \xrightarrow[n\to\infty]{} a \Longrightarrow \alpha x_n \xrightarrow[n\to\infty]{} \alpha a$

(2) $x_n \xrightarrow[n\to\infty]{} a$ *und* $y_n \xrightarrow[n\to\infty]{} b \Longrightarrow x_n + y_n \xrightarrow[n\to\infty]{} a + b$

(3) $x_n \xrightarrow[n\to\infty]{} 0$ *und* $(y_n)_{n\in\mathbb{N}}$ *beschränkt* $\Longrightarrow x_n \cdot y_n \xrightarrow[n\to\infty]{} 0$

(4) $x_n \xrightarrow[n\to\infty]{} a$ *und* $y_n \xrightarrow[n\to\infty]{} b \Longrightarrow x_n \cdot y_n \xrightarrow[n\to\infty]{} a \cdot b$

(5) $x_n \xrightarrow[n\to\infty]{} a$ *und* $a \neq 0 \Longrightarrow$ *es existiert ein* $n_0 \in \mathbb{N}$ *mit* $x_n \neq 0$ *für alle* $n \geq n_0$ *und* $\dfrac{1}{x_n} \xrightarrow[\substack{n\to\infty \\ n\geq n_0}]{} \dfrac{1}{a}$.

(6) $x_n \xrightarrow[n\to\infty]{} a \Longrightarrow |x_n| \xrightarrow[n\to\infty]{} |a|$.

Beweis. (1) O.E. ist $\alpha \neq 0$. Sei $\varepsilon > 0$. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit

$$|x_n - a| < \frac{\varepsilon}{|\alpha|} \quad \forall n \geq n_\varepsilon.$$

Dann folgt

$$|\alpha x_n - \alpha a| = |\alpha||x_n - a| < \varepsilon \quad \forall n \geq n_\varepsilon.$$

(2) Sei $\varepsilon > 0$. Dann gibt es $n_{\varepsilon_1} \in \mathbb{N}$ und $n_{\varepsilon_2} \in \mathbb{N}$ mit

$$|x_n - a| < \frac{\varepsilon}{2} \quad \forall n \geq n_{\varepsilon_1}$$
$$|y_n - b| < \frac{\varepsilon}{2} \quad \forall n \geq n_{\varepsilon_2}.$$

Setze $n_\varepsilon = \max\{n_{\varepsilon_1}, n_{\varepsilon_2}\}$. Dann folgt

$$|(x_n + y_n) - (a + b)| \leq |x_n - a| + |y_n - b| < \varepsilon \quad \forall n \geq n_\varepsilon.$$

(3) Es existiert ein $c > 0$ mit

$$|y_n| \leq c \quad \forall n \in \mathbb{N}.$$

Sei $\varepsilon > 0$. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit

$$|x_n| < \frac{\varepsilon}{c} \quad \forall n \geq n_\varepsilon.$$

Dann folgt

$$|x_n \cdot y_n| = |x_n| \cdot |y_n| \leq c|x_n| < \varepsilon \quad \forall n \geq n_\varepsilon.$$

(4) Es ist

$$x_n \cdot y_n - a \cdot b = (x_n - a)y_n + a(y_n - b) \quad \forall n \in \mathbb{N}.$$

Die Behauptung folgt somit aus Satz II.1.8 ($y_n$ ist beschränkt), Lemma II.1.17 Teil (3) und Teil (2).

(5) Es gibt ein $n_0 \in \mathbb{N}$ mit

$$|x_n - a| < \frac{|a|}{2} \quad \forall n \geq n_0.$$

Daraus folgt

$$|x_n| \geq |a| - |x_n - a| > \frac{|a|}{2} > 0 \quad \forall n \geq n_0.$$

Weiter folgt für alle $n \geq n_0$

$$|\frac{1}{x_n} - \frac{1}{a}| = \frac{|x_n - a|}{|x_n \cdot a|} \leq \frac{2}{|a|^2}|x_n - a|.$$

Damit folgt die Behauptung aus Lemma II.1.17.
(6) Die Behauptung folgt aus Definition II.1.6 und

$$||x_n| - |a|| \leq |x_n - a|.$$

□

Abschließend noch ein wichtiges Vergleichskriterium für konvergente Folgen.

Satz II.1.19. (1) *Seien $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ zwei konvergente Folgen in $\mathbb{R}$ und $a = \lim_{n\to\infty} x_n$, $b = \lim_{n\to\infty} y_n$. Für unendlich viele $n \in \mathbb{N}$ gelte*

$$x_n \leq y_n.$$

*Dann ist*

$$a \leq b.$$

(2) *Seien $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ zwei konvergente Folgen in $\mathbb{R}$ mit $\lim_{n\to\infty} x_n = a = \lim_{n\to\infty} y_n$ und $(z_n)_{n\in\mathbb{N}}$ eine dritte Folge in $\mathbb{R}$. Es gebe ein $n_0 \in \mathbb{N}$ mit*

$$x_n \leq z_n \leq y_n \quad \forall n \geq n_0.$$

*Dann folgt, dass $(z_n)_{n\in\mathbb{N}}$ konvergent ist und*

$$a = \lim_{n\to\infty} z_n.$$

Beweis. (1) Sei $\varepsilon > 0$. Dann gibt es ein $n_{\varepsilon_1} \in \mathbb{N}$ und ein $n_{\varepsilon_2} \in \mathbb{N}$ mit

$$\begin{aligned} & |x_n - a| < \varepsilon \quad \forall n \geq n_{\varepsilon_1} \\ \Longrightarrow \quad & a - \varepsilon < x_n \quad \forall n \geq n_{\varepsilon_1} \end{aligned}$$

und

$$\begin{aligned} & |y_n - b| < \varepsilon \quad \forall n \geq n_{\varepsilon_2} \\ \Longrightarrow \quad & y_n < b + \varepsilon \quad \forall n \geq n_{\varepsilon_2}. \end{aligned}$$

Weiter gibt es ein $n_\varepsilon \geq \max\{n_{\varepsilon_1}, n_{\varepsilon_2}\}$ mit

$$x_{n_\varepsilon} \leq y_{n_\varepsilon}.$$

Damit folgt

$$a - \varepsilon < x_{n_\varepsilon} \leq y_{n_\varepsilon} < b + \varepsilon \quad \Longrightarrow \quad a - b \leq 2\varepsilon.$$

Da $\varepsilon$ beliebig ist, folgt hieraus die Behauptung.
(2) Sei $\varepsilon > 0$. Dann gibt es $n_{\varepsilon_1} \in \mathbb{N}$ und $n_{\varepsilon_2} \in \mathbb{N}$ mit

$$\begin{aligned} |x_n - a| &< \varepsilon \quad \forall n \geq n_{\varepsilon_1}, \\ |y_n - a| &< \varepsilon \quad \forall n \geq n_{\varepsilon_2}. \end{aligned}$$

Setze $n_\varepsilon = \max\{n_{\varepsilon_1}, n_{\varepsilon_2}, n_0\}$. Dann folgt

$$a - \varepsilon < x_n \le z_n \le y_n < a + \varepsilon \quad \forall n \ge n_\varepsilon.$$

Hieraus folgt die Behauptung. □

Bemerkung II.1.20. Die Folgen $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ mit $x_n = -\frac{1}{n}$ und $y_n = \frac{1}{n}$ zeigen, dass aus

$$x_n < y_n \quad \forall n \in \mathbb{N}$$

i. a. nicht folgt

$$\lim_{n\to\infty} x_n < \lim_{n\to\infty} y_n.$$

## II.2. Vollständigkeit

Definition II.2.1. Eine Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ heißt monoton wachsend bzw. monoton fallend, wenn für alle $n \in \mathbb{N}$ gilt

$$x_{n+1} \ge x_n \qquad \text{bzw. } x_{n+1} \le x_n.$$

Satz II.2.2. *Jede monoton wachsende bzw. monoton fallende, beschränkte Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ ist konvergent, und es gilt*

$$\lim_{n\to\infty} x_n = \sup\{x_n : n \in \mathbb{N}\}$$

*bzw.*

$$\lim_{n\to\infty} x_n = \inf\{x_n : n \in \mathbb{N}\}).$$

Beweis. Wir beweisen die Behauptung nur für monoton wachsende Folgen. Der Beweis für monoton fallende Folgen verläuft völlig analog.
Sei also $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ monoton wachsend und beschränkt. Wegen Satz I.4.7 (S. 20) existiert

$$s = \sup\{x_n : n \in \mathbb{N}\}.$$

Sei $\varepsilon > 0$. Aus der Definition des Supremums folgt, dass es ein $n_\varepsilon \in \mathbb{N}$ gibt mit

$$s - \varepsilon < x_{n_\varepsilon} \le s.$$

Wegen der Monotonie gilt diese Ungleichung für alle $n \ge n_\varepsilon$. □

Beispiel II.2.3. (1) Sei $a \in \mathbb{K}$. Dann gilt

$$\begin{aligned} a^n &\underset{n\to\infty}{\longrightarrow} 0, && \text{falls } |a| < 1 \\ a^n &\underset{n\to\infty}{\longrightarrow} 1, && \text{falls } a = 1 \\ (a^n)_{n\in\mathbb{N}} & && \text{divergiert, falls } |a| \ge 1 \text{ und } a \ne 1. \end{aligned}$$

(2) Sei $a \in \mathbb{K}$ mit $|a| < 1$ und $r \in \mathbb{N}^*$. Dann gilt

$$n^r a^n \underset{n\to\infty}{\longrightarrow} 0.$$

(3) $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$.
(4) Sei $a \in \mathbb{R}_+^*$. Dann ist $\lim\limits_{n\to\infty} \sqrt[n]{a} = 1$.
(5) Die Folge $\left(\left(1+\frac{1}{n}\right)^n\right)_{n\in\mathbb{N}}$ ist konvergent. Der Grenzwert heißt Eulersche Zahl und wird mit $e$ bezeichnet. Es ist

$$2 < e \le 3.$$

(6) $\lim\limits_{n\to\infty} \sum\limits_{k=0}^{n} \frac{1}{k!} = e$.

Beweis. (1) Angenommen $(a^n)_{n\in\mathbb{N}}$ konvergiert. Sei $\alpha = \lim\limits_{n\to\infty} a^n$. Für die Teilfolge $(a^{n_k})_{k\in\mathbb{N}}$ mit $n_k = k+1$ gilt dann $\alpha = \lim\limits_{k\to\infty} a^{k+1}$. Aus den Rechenregeln für Grenzwerte folgt

$$\begin{aligned} \alpha &= \lim_{k\to\infty} a^{k+1} \\ &= \lim_{k\to\infty} (a \cdot a^k) \\ &= a \cdot \lim_{k\to\infty} a^k \\ &= \alpha a. \end{aligned}$$

Also ist $\alpha = 0$ oder $a = 1$.
Sei $|a| < 1$. Dann ist $(|a|^n)_{n\in\mathbb{N}}$ monoton fallend und beschränkt, also konvergent. Aus Obigem folgt

$$\lim_{n\to\infty} |a|^n = 0.$$

Lemma II.1.17(3) (S. 32) liefert $a^n \underset{n\to\infty}{\longrightarrow} 0$.
Sei $|a| = 1$ und $a \ne 1$. Angenommen $(a^n)_{n\in\mathbb{N}}$ konvergiert. Dann ist $\lim\limits_{n\to\infty} a^n = 0$ und somit $\lim\limits_{n\to\infty} |a|^n = 0$ im Widerspruch zu $|a|^n = 1$ für alle $n \in \mathbb{N}$.
Sei $|a| > 1$. Dann folgt $\left(\frac{1}{a}\right)^n \underset{n\to\infty}{\longrightarrow} 0$. Also gibt es zu jedem $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ mit

$$\frac{1}{|a^n|} < \varepsilon \quad \forall n \ge n_\varepsilon.$$

Also ist $(a^n)_{n\in\mathbb{N}}$ unbeschränkt. Damit folgt die Behauptung aus Satz II.1.8 (S. 30).
(2) O.E. ist $a \ne 0$. Setze $\alpha = |a|$ und

$$x_n = n^r \alpha^n.$$

Dann gilt

$$x_n > 0 \quad \forall n \in \mathbb{N}^*$$

und

$$\frac{x_{n+1}}{x_n} = \frac{(n+1)^r}{n^r} \frac{\alpha^{n+1}}{\alpha^n} = (1+\frac{1}{n})^r \alpha \underset{n\to\infty}{\longrightarrow} \alpha.$$

Sei $\beta \in (\alpha, 1)$. Dann gibt es ein $n_0 \in \mathbb{N}^*$ mit

$$\frac{x_{n+1}}{x_n} \leq \beta \qquad \forall n \geq n_0,$$

d.h.

$$0 < x_{n+1} \leq \beta x_n \qquad \forall n \geq n_0.$$

Damit folgt durch Induktion

$$x_n \leq x_{n_0} \beta^{n-n_0} \qquad \forall n \geq n_0.$$

Also gilt

$$|n^r a^n| = x_n \leq x_{n_0} \beta^{n-n_0} \underset{\substack{n\to\infty \\ n\geq n_0}}{\longrightarrow} 0.$$

Hieraus folgt die Behauptung zusammen mit Lemma II.1.17 (S. 32).
(3): Sei $\varepsilon > 0$. Wegen $0 < \frac{1}{1+\varepsilon} < 1$ folgt aus Teil (2):

$$\lim_{n\to\infty} n(\frac{1}{1+\varepsilon})^n = 0.$$

Also gibt es ein $n_0 \in \mathbb{N}^*$ mit

$$\begin{aligned} & 0 < n(\frac{1}{1+\varepsilon})^n < 1 \quad \forall n \geq n_0 \\ \iff \quad & 1 \leq n < (1+\varepsilon)^n \quad \forall n \geq n_0 \\ \iff \quad & 1 \leq \sqrt[n]{n} < 1+\varepsilon \quad \forall n \geq n_0. \end{aligned}$$

Hieraus folgt die Behauptung mit Satz II.1.19 (S. 34).
(4): Wegen $a > 0$ gibt es ein $n_0 \in \mathbb{N}$ mit

$$\frac{1}{n} < a < n \quad \forall n \geq n_0.$$

Hieraus folgt

$$\frac{1}{\sqrt[n]{n}} < \sqrt[n]{a} < \sqrt[n]{n} \quad \forall n \geq n_0.$$

Damit folgt die Behauptung aus Teil (3) und Satz II.1.19 (S. 34).
(5): Definiere

$$e_n = (1 + \frac{1}{n})^n \quad n \in \mathbb{N}^*.$$

Aus der Binomischen Formel folgt

$$\begin{aligned} e_n &= \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} \\ &= 1 + \sum_{k=1}^{n} \binom{n}{k} \frac{1}{n^k} \\ &= 1 + \sum_{k=1}^{n} \frac{1}{k!} \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{n^k} \\ &\leq 1 + \sum_{k=1}^{n} \frac{1}{k!} \end{aligned}$$

$$\begin{aligned}
&\leq 1 + \sum_{k=1}^{n} 2^{-(k-1)} \quad \text{wegen } k! \geq 2^{k-1} \\
&= 1 + \frac{1-2^{-n}}{1-\frac{1}{2}} \\
&\leq 1 + 2 \\
&= 3.
\end{aligned}$$

Dies zeigt insbesondere

$$(*) \qquad e_n \leq 1 + \sum_{k=1}^{n} \frac{1}{k!} \leq 3.$$

Aus der Bernoullischen Ungleichung

$$(1+x)^n \geq 1 + nx \quad \forall x > -1$$

folgt weiter

$$\begin{aligned}
\frac{e_n}{e_{n-1}} &= \left(\frac{n+1}{n}\right)^n \cdot \left(\frac{n-1}{n}\right)^{n-1} \\
&= \left(\frac{n+1}{n} \cdot \frac{n-1}{n}\right)^n \cdot \frac{n}{n-1} \\
&= \left(1 - \frac{1}{n^2}\right)^n \frac{n}{n-1} \\
&\geq (1 - \frac{1}{n})\frac{n}{n-1} \\
&= 1 \qquad \forall n \geq 2.
\end{aligned}$$

Also ist $(e_n)_{n\in\mathbb{N}}$ monoton wachsend und durch 2 bzw. 3 nach unten bzw. oben beschränkt. Hieraus folgt die Behauptung.
(6): Sei

$$x_n = \sum_{k=0}^{n} \frac{1}{k!}.$$

$(x_n)_{n\in\mathbb{N}}$ ist monoton wachsend und wegen $(*)$ nach oben durch 3 beschränkt. Mithin existiert

$$e' = \lim_{n\to\infty} x_n.$$

Weiter folgt aus $(*)$

$$e_n \leq x_n \qquad \forall n \in \mathbb{N}$$

und somit wegen Satz II.1.19 (S. 34)

$$e \leq e'.$$

Sei nun $m \in \mathbb{N}$ beliebig. Für $n > m$ folgt

$$e_n = \left(1 + \frac{1}{n}\right)^n$$

$$
\begin{aligned}
&= \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} \\
&> \sum_{k=0}^{m} \binom{n}{k} \frac{1}{n^k} \\
&= \sum_{k=0}^{m} \frac{1}{k!} \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{n^k} \\
&= \sum_{k=0}^{m} \frac{1}{k!} 1 \cdot (1 - \frac{1}{n}) \cdot \ldots \cdot (1 - \frac{k-1}{n}) \\
&\underset{n\to\infty}{\longrightarrow} \sum_{k=0}^{m} \frac{1}{k!} \\
&= x_m.
\end{aligned}
$$

Also gilt

$$e \geq x_m \quad \forall m \in \mathbb{N}$$

und somit

$$e \geq e'.$$

Damit ist die Behauptung bewiesen. □

Der folgende Satz ist das zentrale Ergebnis dieses Abschnitts. Es geht auf B. Bolzano (1781-1848) und K. Weierstrass (1815-1897) zurück.

Satz II.2.4 (Satz von Bolzano-Weierstrass). *Jede beschränkte Folge in $\mathbb{K}$ besitzt mindestens einen Häufungspunkt.*

Beweis. Sei zunächst $\mathbb{K} = \mathbb{R}$ und $(x_k)_{k\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$. Wir definieren rekursiv zwei Folgen $(A_n)_{n\in\mathbb{N}}$ und $(B_n)_{n\in\mathbb{N}}$ mit folgenden Eigenschaften:

(1) $A_n \leq B_n$ für alle $n \in \mathbb{N}$.
(2) Jedes Intervall $[A_n, B_n]$ enthält unendlich viele Glieder der Folge $(x_k)_{k\in\mathbb{N}}$.
(3) $(A_n)_{n\in\mathbb{N}}$ ist monoton wachsend.
(4) $(B_n)_{n\in\mathbb{N}}$ ist monoton fallend.
(5) $B_n - A_n \leq 2^{-n}(B_0 - A_0)$ für alle $n \in \mathbb{N}$.

Aus der Beschränktheit von $(x_k)_{k\in\mathbb{N}}$ folgt die Existenz von $A_0$ und $B_0$. $A_n$ und $B_n$ seien konstruiert.
Fall 1: Es gibt unendlich viele $k \in \mathbb{N}$ mit

$$x_k = \frac{1}{2}(A_n + B_n).$$

Dann ist $c = \frac{1}{2}(A_n + B_n)$ offensichtlich ein HP von $(x_k)_{k\in\mathbb{N}}$ und wir sind fertig.

Fall 2: Unendlich viele Folgenglieder von $(x_k)_{k\in\mathbb{N}}$ liegen in $[A_n, \frac{1}{2}(A_n+B_n)]$. Dann setzen wir

$$A_{n+1} = A_n$$
$$B_{n+1} = \frac{1}{2}(A_n + B_n).$$

Fall 3: Unendlich viele Folgenglieder von $(x_k)_{k\in\mathbb{N}}$ liegen in $(\frac{1}{2}(A_n + B_n), B_n]$. Dann setzen wir

$$A_{n+1} = \frac{1}{2}(A_n + B_n)$$
$$B_{n+1} = B_n.$$

Wegen der Eigenschaften (1), (3) und (4) und Satz II.2.2 sind die Folgen $(A_n)_{n\in\mathbb{N}}$ und $(B_n)_{n\in\mathbb{N}}$ konvergent. Wegen der Eigenschaft (5) stimmen ihre Grenzwerte überein. Sei also

$$c = \lim_{n\to\infty} A_n.$$

Wir müssen noch zeigen, dass $c$ HP der Folge $(x_k)_{k\in\mathbb{N}}$ ist. Sei dazu $\varepsilon > 0$ beliebig. Dann gibt es ein $n_{\varepsilon_1} \in \mathbb{N}$ und ein $n_{\varepsilon_2} \in \mathbb{N}$ mit

$$c - \varepsilon < A_n \le c \quad \forall n \ge n_{\varepsilon_1}$$
$$c \le B_n < c + \varepsilon \quad \forall n \ge n_{\varepsilon_2}.$$

Setze $n_\varepsilon = \max\{n_{\varepsilon_1}, n_{\varepsilon_2}\}$. Konstruktionsgemäß gibt es unendlich viele $k$ mit

$$c - \varepsilon < A_{n_\varepsilon} \le x_k \le B_{n_\varepsilon} < c + \varepsilon.$$

Also ist $c$ HP von $(x_k)_{k\in\mathbb{N}}$.

Sei nun $\mathbb{K} = \mathbb{C}$ und $(x_k)_{k\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{C}$. Definiere für $k \in \mathbb{N}$

$$u_k = \mathrm{Re}\ x_k$$
$$v_k = \mathrm{Im}\ x_k$$

wegen

$$|\operatorname{Re} z| \le |z|, \ |\operatorname{Im} z| \le |z| \quad \forall z \in \mathbb{C}$$

sind $(u_k)_{k\in\mathbb{N}}$ und $(v_k)_{k\in\mathbb{N}}$ beschränkte Folgen in $\mathbb{R}$. Gemäß Obigem gibt es ein $a \in \mathbb{R}$ und eine Teilfolge $(u_{k_l})_{l\in\mathbb{N}}$ von $(u_k)_{k\in\mathbb{N}}$, die gegen $a$ konvergiert. Dann gibt es aber auch eine Teilfolge $(v_{k_{l_m}})_{m\in\mathbb{N}}$ von $(v_{k_l})_{l\in\mathbb{N}}$ und ein $b \in \mathbb{R}$ mit

$$b = \lim_{m\to\infty} v_{k_{l_m}}.$$

Wegen

$$|z| \le |\operatorname{Re} z| + |\operatorname{Im} z|$$

folgt

$$\lim_{m\to\infty} x_{k_{l_m}} = a + ib.$$

Damit ist die Behauptung bewiesen. □

Der folgende, für praktische Anwendungen äußerst wichtige Begriff geht auf A.L. Cauchy (1789-1857) zurück.

Definition II.2.5. Eine Folge $(x_k)_{k\in\mathbb{N}}$ heißt Cauchyfolge, kurz CF, falls es zu jedem $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ gibt mit

$$|x_n - x_m| < \varepsilon \quad \forall n, m \geq n_\varepsilon.$$

Satz II.2.6. *Jede konvergente Folge ist eine Cauchyfolge.*

Beweis. Sei $\varepsilon > 0$. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit

$$|x_n - a| < \frac{\varepsilon}{2} \quad \forall n \geq n_\varepsilon,$$

wobei $a = \lim_{n\to\infty} x_n$ ist. Für $n, m \geq n_\varepsilon$ folgt

$$|x_n - x_m| \leq |x_n - a| + |x_m - a| < \varepsilon.$$

□

Satz II.2.7. *Jede Cauchyfolge ist beschränkt.*

Beweis. Nach Voraussetzung gibt es ein $n_1 \in \mathbb{N}$ mit

$$|x_n - x_m| < 1 \quad \forall n, m \geq n_1.$$

Definiere

$$c = \max\{|x_n| + 1 : n \leq n_1\}.$$

Für $n \leq n_1$ gilt trivialerweise

$$|x_n| \leq c.$$

Für $n \geq n_1$ folgt

$$\begin{aligned}|x_n| &\leq |x_{n_1}| + |x_n - x_{n_1}| \\ &< |x_{n_1}| + 1 \\ &\leq c.\end{aligned}$$

□

Satz II.2.8. *Jede Cauchyfolge in $\mathbb{K}$ ist konvergent.*

Beweis. Aus Satz II.2.4 und Satz II.2.7 folgt, dass die Cauchyfolge $(x_k)_{k\in\mathbb{N}}$ einen HP $a$ hat. Sei $\varepsilon > 0$. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit

$$|x_n - x_m| < \frac{\varepsilon}{2} \quad \forall n, m \geq n_\varepsilon.$$

Außerdem gibt es ein $k_\varepsilon \geq n_\varepsilon$ mit

$$|x_{k_\varepsilon} - a| < \frac{\varepsilon}{2}.$$

Damit folgt

$$|x_n - a| \leq |x_{k_\varepsilon} - a| + |x_n - x_{k_\varepsilon}| < \varepsilon \quad \forall n \geq n_\varepsilon.$$

□

Bemerkung II.2.9. Die Aussage von Satz II.2.8 ist in $\mathbb{Q}$ i. a. falsch. Um dies einzusehen, erinnern wir an die beiden Folgen $(a_n)_{n\in\mathbb{N}} \subset \mathbb{Q}$ und $(b_n)_{n\in\mathbb{N}} \subset \mathbb{Q}$, die wir zu Beginn von Paragraph I.4 konstruiert haben. $(a_n)_{n\in\mathbb{N}}$ war monoton wachsend, $(b_n)_{n\in\mathbb{N}}$ war monoton fallend, es galt

$$|a_n - b_n| \leq 2^{-n} \quad \forall n \in \mathbb{N}$$

und

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \sqrt{2} \notin \mathbb{Q}.$$

Wegen Satz II.2.6 sind $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ zwei Cauchyfolgen in $\mathbb{Q}$, die in $\mathbb{Q}$ nicht konvergieren. Man sagt auch, $\mathbb{Q}$ ist nicht vollständig.
Wir haben die reellen Zahlen als ordnungsvollständige Erweiterung von $\mathbb{Q}$ konstruiert. Aus der Ordnungsvollständigkeit von $\mathbb{R}$ folgt dann die Vollständigkeit, d.h. die Konvergenz von Cauchyfolgen. Man kann $\mathbb{R}$ äquivalent auch als Vervollständigung von $\mathbb{Q}$, d.h. als Obermenge, in der alle Cauchyfolgen konvergieren, konstruieren. Die Ordnungsvollständigkeit folgt dann aus der Vollständigkeit. Schematisch können wir diesen Sachverhalt wie folgt darstellen:

$$\begin{array}{cccl} & \leq & K_1 & \text{ordnungsvollständige Erweiterung} \\ & \nearrow & \downarrow\cong & \\ \mathbb{Q} & & \mathbb{R} & \\ & \searrow & \uparrow\cong & \\ & \text{CF} & K_2 & \text{vollständige Erweiterung} \end{array}$$

## II.3. Uneigentliche Konvergenz

Wir erinnern an Definition I.4.9 (S. 20), in der wir zwei Symbole $+\infty$ und $-\infty$ eingeführt haben mit

$$-\infty < x < +\infty \quad \forall x \in \mathbb{R}.$$

Definition II.3.1. Wir definieren $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ und setzen Addition und Multiplikation auf $\overline{\mathbb{R}}$ wie folgt fort:

$$\begin{aligned} x + \infty &= +\infty \quad \forall x > -\infty \\ x - \infty &= -\infty \quad \forall x < +\infty \\ x \cdot (+\infty) &= \begin{cases} +\infty & \text{falls } x > 0 \\ -\infty & \text{falls } x < 0 \end{cases} \\ x \cdot (-\infty) &= -(x \cdot (+\infty)) \\ 0 \cdot \infty &= 0 \\ \frac{x}{+\infty} &= \frac{x}{-\infty} = 0 \quad \forall x \in \mathbb{R} \\ \frac{x}{0} &= \begin{cases} +\infty & \text{falls } x > 0 \\ -\infty & \text{falls } x < 0. \end{cases} \end{aligned}$$

Aus Definition I.4.9 (S. 20) folgt:

Bemerkung II.3.2. Jede Teilmenge von $\mathbb{R}$ hat ein Infimum und Supremum in $\overline{\mathbb{R}}$.

Definition II.3.3. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$. Wir sagen, dass $x_n$ gegen $+\infty$ bzw. gegen $-\infty$ konvergiert, wenn es zu jedem $R \in \mathbb{R}_+$ ein $n_R \in \mathbb{N}$ gibt mit

$$x_n > R \quad \forall n \geq n_R$$

bzw.

$$x_n < -R \quad \forall n \geq n_R.$$

Beispiel II.3.4. (1) $\lim\limits_{n\to\infty} n = +\infty$
(2) $\lim\limits_{n\to\infty} (-n) = -\infty$
(3) Die Folge $((-1)^n n)_{n\in\mathbb{N}}$ konvergiert nicht in $\overline{\mathbb{R}}$. Sie hat die beiden HP $+\infty$ und $-\infty$.

Satz II.3.5. *Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$. Dann gilt:*

$$\text{(1)} \qquad x_n \xrightarrow[n\to\infty]{} +\infty \text{ oder } x_n \xrightarrow[n\to\infty]{} -\infty \Longrightarrow \frac{1}{x_n} \xrightarrow[n\to\infty]{} 0.$$

$$\text{(2)} \qquad x_n \xrightarrow[n\to\infty]{} 0 \text{ und } x_n > 0 \text{ für fast alle } n \Longrightarrow \frac{1}{x_n} \xrightarrow[n\to\infty]{} +\infty.$$

$$x_n \xrightarrow[n\to\infty]{} 0 \text{ und } x_n < 0 \text{ für fast alle } n \Longrightarrow \frac{1}{x_n} \xrightarrow[n\to\infty]{} -\infty.$$

Beweis. (1) Es gelte $x_n \xrightarrow[n\to\infty]{} +\infty$. Sei $\varepsilon > 0$ beliebig. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit

$$x_n > \frac{1}{\varepsilon} \quad \forall n \geq n_\varepsilon.$$

Dann folgt $x_n > 0$ für alle $n \geq n_\varepsilon$ und

$$\frac{1}{x_n} < \varepsilon \quad \forall n \geq n_\varepsilon.$$

Also ist $\lim\limits_{n\to\infty} \frac{1}{x_n} = 0$. Der Fall $x_n \xrightarrow[n\to\infty]{} -\infty$ wird analog behandelt.
(2) Es gelte $x_n \xrightarrow[n\to\infty]{} 0$ und $x_n > 0$ für fast alle $n \in \mathbb{N}$. Dann gibt es ein $n_0 \in \mathbb{N}$ mit

$$x_n > 0 \quad \forall n \geq n_0.$$

Sei $R \in \mathbb{R}^*$ beliebig. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit

$$|x_n| < \frac{1}{R} \quad \forall n \geq n_\varepsilon.$$

Setze $n_R = \max\{n_0, n_\varepsilon\}$. Dann gilt

$$\frac{1}{x_n} > R \quad \forall n \geq n_R.$$

Der Fall $x_n < 0$ für fast alle $n$ wird analog behandelt. $\square$

SATZ II.3.6. *Jede monoton wachsende bzw. monoton fallende Folge* $(x_n)_{n\in\mathbb{N}}$ *in* $\overline{\mathbb{R}}$ *konvergiert in* $\overline{\mathbb{R}}$ *und es gilt*

$$\lim_{n\to\infty} x_n = \sup\{x_n : n \in \mathbb{N}\}$$

*bzw.*

$$\lim_{n\to\infty} x_n = \inf\{x_n : n \in \mathbb{N}\}.$$

BEWEIS. Wir betrachten nur monoton wachsende Folgen. Der Beweis für monoton fallende Folgen ist völlig analog.
FALL 1 $x_n = -\infty \quad \forall n \in \mathbb{N}$: Dann ist natürlich

$$-\infty = \lim_{n\to\infty} x_n \quad \text{und} \quad -\infty = \sup\{x_n : n \in \mathbb{N}\}.$$

FALL 2 $\exists\ n_0 \in \mathbb{N}$ mit $x_{n_0} > -\infty$: Es reicht, die Folge $(x_n)_{n\geq n_0}$ zu betrachten.
Angenommen $(x_n)_{n\geq n_0}$ ist beschränkt in $\mathbb{R}$. Dann folgt die Behauptung aus Satz II.2.2 (S. 35).
Angenommen $(x_n)_{n\geq n_0}$ ist nicht beschränkt in $\mathbb{R}$. Sei $R \in \mathbb{R}^*$. Dann gibt es ein $n_R \geq n_0$ mit

$$x_{n_R} > R.$$

Da $(x_n)_{n\in\mathbb{N}}$ monoton wachsend ist, folgt

$$x_n > R \quad \forall n \geq n_R.$$

Damit ist die Behauptung bewiesen. □

Sei nun $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$. Definiere für $n \in \mathbb{N}$

$$y_n = \sup\{x_k : k \geq n\} \in \overline{\mathbb{R}}$$
$$z_n = \inf\{x_k : k \geq n\} \in \overline{\mathbb{R}}.$$

Dann ist $(y_n)_{n\in\mathbb{N}}$ eine monoton fallende Folge in $\overline{\mathbb{R}}$ und $(z_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge in $\overline{\mathbb{R}}$. Gemäß Satz II.3.6 sind daher beide Folgen in $\overline{\mathbb{R}}$ konvergent. Mithin ist die folgende Definition sinnvoll.

DEFINITION II.3.7. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$. Dann ist

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup\{x_k : k \geq n\} \ (\text{LIMES SUPERIOR})$$
$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf\{x_k : k \geq n\} \ (\text{LIMES INFERIOR}).$$

BEMERKUNG II.3.8. Wegen

$$\sup\{x_k : k \geq n\} \geq \inf\{x_k : k \geq n\} \qquad \forall n \in \mathbb{N}$$

gilt stets

$$\limsup_{n\to\infty} x_n \geq \liminf_{n\to\infty} x_n.$$

SATZ II.3.9. *Sei* $(x_n)_{n\in\mathbb{N}}$ *eine Folge in* $\mathbb{R}$. *Dann ist* $\limsup\limits_{n\to\infty} x_n$ *bzw.* $\liminf\limits_{n\to\infty} x_n$ *der größte bzw. kleinste HP der Folge in* $\overline{\mathbb{R}}$.

Beweis. Wir beweisen die Behauptung nur für den Limes superior. Der Beweis für den Limes inferior verläuft analog. Sei $x = \limsup_{n\to\infty} x_n$ und $y_n = \sup\{x_k : k \geq n\}$ für alle $n \in \mathbb{N}$.
Fall 1 $x = +\infty$: Da $(y_n)_{n\in\mathbb{N}}$ monoton fallend ist, folgt

$$y_n = +\infty \quad \forall n \in \mathbb{N}$$

und somit

$$\forall R > 0 \quad \forall n \in \mathbb{N} \quad \exists\, k_n \geq n : x_{k_n} > R.$$

Also ist $+\infty$ HP von $(x_n)_{n\in\mathbb{N}}$. Für jeden anderen HP $z$ gilt trivialerweise $z \leq x$.
Fall 2 $x \in \mathbb{R}$: Sei $\varepsilon > 0$. Da $(y_n)_{n\in\mathbb{N}}$ monoton fallend ist, folgt

$$\begin{aligned} &\exists\, n_\varepsilon \in \mathbb{N} \quad \forall n \geq n_\varepsilon : y_n < x + \varepsilon \\ \Longrightarrow &\operatorname{Anz}\{k \in \mathbb{N} : x_k \geq x + \varepsilon\} < \infty. \end{aligned}$$

Also gilt für jeden HP $z$ von $(x_n)_{n\in\mathbb{N}}$

$$z < x + \varepsilon.$$

Da $\varepsilon$ beliebig ist, gilt für jeden HP $z$ von $(x_n)_{n\in\mathbb{N}}$

$$z \leq x.$$

Weiter gilt

$$\begin{aligned} &\forall \varepsilon > 0 \quad \exists\, n_\varepsilon \in \mathbb{N} \quad \forall n \geq n_\varepsilon : x \leq y_n < x + \varepsilon \\ \Longrightarrow &\forall \varepsilon > 0 \quad \forall n \in \mathbb{N} \quad \exists\, k \geq n : x - \varepsilon \leq x_k < x + \varepsilon. \end{aligned}$$

Also ist $x$ auch HP von $(x_n)_{n\in\mathbb{N}}$.
Fall 3 $x = -\infty$: Sei $R \in \mathbb{R}_+$. Dann folgt

$$\begin{aligned} &\exists\, n_R \in \mathbb{N} \quad \forall n \geq n_R : y_n < -R \\ \Longrightarrow\ &\operatorname{Anz}\{k \in \mathbb{N} : x_k \geq -R\} < \infty. \end{aligned}$$

Da $R$ beliebig ist, folgt wieder, dass für jeden HP $z$ von $(x_n)_{n\in\mathbb{N}}$ gilt

$$z = -\infty.$$

Außerdem folgt

$$\begin{aligned} &\forall R \in \mathbb{R}_+ \quad \exists\, n_R \in \mathbb{N} \quad \forall n \geq n_R : y_n < -R \\ \Longrightarrow &\forall R \in \mathbb{R}_+ \quad \forall n \in \mathbb{N} \quad \exists\, k \geq n : x_k < -R \end{aligned}$$

Also ist $-\infty$ auch ein HP der Folge $(x_n)_{n\in\mathbb{N}}$. □

Beispiel II.3.10. $x_n = (-1)^n \frac{n}{n+1}$. Dann ist

$$\limsup_{n\to\infty} x_n = 1 \quad , \quad \liminf_{n\to\infty} x_n = -1.$$

Satz II.3.11. *Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$. Dann sind folgende Aussagen äquivalent:*

(1)
$$x_n \xrightarrow[n\to\infty]{} a \in \overline{\mathbb{R}}$$

(2)
$$\limsup_{n\to\infty} x_n \leq a \leq \liminf_{n\to\infty} x_n.$$

Beweis. (1) $\Longrightarrow$ (2): Folgt aus Satz II.3.9, da $a$ einziger HP von $(x_n)_{n\in\mathbb{N}}$ ist.
(2) $\Longrightarrow$ (1): Aus Bemerkung II.3.8 folgt

$$\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n = a \in \overline{\mathbb{R}}.$$

Wegen Satz II.3.9 ist $a$ der einzige HP. Hieraus folgt

$$\forall \varepsilon > 0 : \text{Anz}\{n \in \mathbb{N} : x_n < a - \varepsilon\} < \infty$$
$$\forall \varepsilon > 0 : \text{Anz}\{n \in \mathbb{N} : x_n > a + \varepsilon\} < \infty.$$

Also ist $a = \lim\limits_{n\to\infty} x_n$. □

## II.4. Reihen

Definition II.4.1. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$. Wir definieren die $n$-te Partialsumme $s_n$ durch

$$s_n = \sum_{k=0}^{n} x_k.$$

Die Folge $(s_n)_{n\in\mathbb{N}}$ heißt Reihe und wird zur Abkürzung mit dem Symbol

$$\sum x_n$$

bezeichnet. Die Reihe heißt konvergent bzw. divergent, wenn die Folge $(s_n)_{n\in\mathbb{N}}$ konvergent bzw. divergent ist. Falls $s_n \xrightarrow[n\to\infty]{} s \in \mathbb{R}$ gilt, schreiben wir kurz

$$s = \sum_{n=0}^{\infty} x_n.$$

Beispiel II.4.2. (1) $\sum \frac{1}{n!}$ ist konvergent, $e = \sum\limits_{n=0}^{\infty} \frac{1}{n!}$.

(2) $\sum\limits_{n\geq 1} \frac{1}{n^2}$ ist konvergent.

(3) Die harmonische Reihe $\sum\limits_{n\geq 1} \frac{1}{n}$ ist divergent.

(4) Die geometrische Reihe $\sum a^n$ ist genau dann konvergent, wenn gilt $|a| < 1$. In diesem Fall ist

$$\sum_{n=0}^{\infty} a^n = \frac{1}{1-a}.$$

Beweis. ad (1): Beispiel II.2.3(6) (S. 35).
ad (2): $(s_n)_{n\in\mathbb{N}}$ mit $s_n = \sum_{k=1}^{n} \frac{1}{k^2}$ ist monoton wachsend. Weiter gilt für $n \geq 2$

$$\begin{aligned} s_n &= 1 + \sum_{k=2}^{n} \frac{1}{k^2} \\ &\leq 1 + \sum_{k=2}^{n} \frac{1}{k(k-1)} \\ &= 1 + \sum_{k=2}^{n} (\frac{1}{k-1} - \frac{1}{k}) \\ &= 1 + \sum_{l=1}^{n-1} \frac{1}{l} - \sum_{k=2}^{n} \frac{1}{k} \\ &= 1 + 1 - \frac{1}{n} \\ &\leq 2. \end{aligned}$$

Damit folgt die Behauptung aus Satz II.2.2 (S. 35).
ad (3): Sei $n \in \mathbb{N}^*$. Dann gilt

$$\begin{aligned} s_{2n} - s_n &= \sum_{k=n+1}^{2n} \frac{1}{k} \\ &\geq \sum_{k=n+1}^{2n} \frac{1}{2n} \\ &= \frac{1}{2}. \end{aligned}$$

Also ist $(s_n)_{n\in\mathbb{N}}$ keine Cauchyfolge und die Behauptung folgt aus Satz II.2.6 (S. 41).
ad (4): Für $a = 1$ ist

$$\sum_{k=0}^{n} a^k = n + 1.$$

Für $a \neq 1$ ist gemäß Satz I.1.16 (S. 12)

$$\sum_{k=0}^{n} a^k = \frac{1 - a^{n+1}}{1 - a}.$$

Damit folgt die Behauptung aus Beispiel II.2.3(1) (S. 35). □

Satz II.4.3. *Ist die Reihe $\sum x_n$ konvergent, so ist $(x_n)_{n\in\mathbb{N}}$ eine Nullfolge.*

Beweis. Ist $\sum x_n$ konvergent, so ist definitionsgemäß $(s_n)_{n\in\mathbb{N}}$ konvergent und damit eine Cauchyfolge. Daher gilt

$$\forall \varepsilon > 0 \quad \exists\, n_\varepsilon \in \mathbb{N} \quad \forall n \geq n_\varepsilon : |x_n| = |s_n - s_{n-1}| < \varepsilon.$$

Hieraus folgt die Behauptung. □

Bemerkung II.4.4. Beispiel II.4.2 (3) zeigt, dass die Umkehrung von Satz II.4.3 i. a. falsch ist.

Aus Definition II.4.1 und Satz II.1.18 (S. 32) folgen unmittelbar folgende Rechenregeln für Reihen.

Satz II.4.5. *Seien $\sum x_n$ und $\sum y_n$ zwei konvergente Reihen und $\alpha \in \mathbb{K}$. Dann gilt:*

(1) $\sum(\alpha x_n)$ *ist konvergent und* $\displaystyle\sum_{n=0}^{\infty}(\alpha x_n) = \alpha \sum_{n=0}^{\infty} x_n$.

(2) $\sum(x_n + y_n)$ *ist konvergent und*

$$\sum_{n=0}^{\infty}(x_n + y_n) = \sum_{n=0}^{\infty} x_n + \sum_{n=0}^{\infty} y_n.$$

Als nächstes geben wir einige Konvergenzkriterien für Reihen an.

Satz II.4.6 (Cauchykriterium). *$\sum x_n$ ist genau dann konvergent, wenn es zu jedem $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ gibt mit*

$$\Big|\sum_{k=m+1}^{n} x_k\Big| < \varepsilon \quad \forall n > m \geq n_\varepsilon.$$

Beweis. Die Behauptung folgt wegen

$$|s_n - s_m| = \Big|\sum_{k=m+1}^{n} x_k\Big|$$

aus den Sätzen II.2.6 (S. 41) und II.2.8 (S. 41). □

Satz II.4.7. *Sei $\sum x_n$ eine Reihe in $\mathbb{R}$ mit $x_n \geq 0$ für alle $n \in \mathbb{N}$. Dann ist die Reihe genau dann konvergent, wenn die Folge der Partialsummen beschränkt ist.*

Beweis. Wegen $x_n \geq 0$ für alle $n \in \mathbb{N}$ ist $(s_n)_{n\in\mathbb{N}}$ monoton wachsend. Damit folgt die Behauptung aus den Sätzen II.1.8 (S. 30) und II.2.2 (S. 35). □

Satz II.4.8 (Leibnizkriterium). *Sei $(x_n)_{n\in\mathbb{N}}$ eine monoton fallende Folge in $\mathbb{R}$ mit $x_n \geq 0$ für alle $n \in \mathbb{N}$. Dann ist die Reihe $\sum(-1)^n x_n$ genau dann konvergent, wenn $(x_n)_{n\in\mathbb{N}}$ eine Nullfolge ist.*

Beweis. Die Notwendigkeit folgt aus Satz II.4.3. Sei also $(x_n)_{n\in\mathbb{N}}$ eine monoton fallende Nullfolge mit nicht negativen Gliedern. Dann folgt für $n \in \mathbb{N}$

$$s_{2n+2} - s_{2n} = x_{2n+2} - x_{2n+1} \leq 0$$
$$s_{2n+3} - s_{2n+1} = -x_{2n+3} + x_{2n+2} \geq 0.$$

Also sind $(s_{2n})_{n\in\mathbb{N}}$ und $(s_{2n+1})_{n\in\mathbb{N}}$ monoton fallende bzw. monoton wachsende Folgen. Weiter gilt

$$s_{2n+1} - s_{2n} = -x_{2n+1} \leq 0 \quad \forall n \in \mathbb{N}.$$

Also sind $(s_{2n})_{n\in\mathbb{N}}$ durch $s_1$ nach unten und $(s_{2n+1})_{n\in\mathbb{N}}$ durch $s_0$ nach oben beschränkt. Gemäß Satz II.2.2 (S. 35) sind sie konvergent. Sei

$$s = \lim_{n\to\infty} s_{2n} \ , \ t = \lim_{n\to\infty} s_{2n+1}.$$

Dann folgt aus Satz II.1.18 (S. 32)

$$s - t = \lim_{n\to\infty} (s_{2n} - s_{2n+1}) = \lim_{n\to\infty} x_{2n+1} = 0.$$

Sei nun $\varepsilon > 0$ beliebig. Dann gibt es ein $n_1 \in \mathbb{N}$ und ein $n_2 \in \mathbb{N}$ mit

$$|s_{2n} - s| < \varepsilon \quad \forall n \geq n_1.$$

und

$$|s_{2n+1} - s| < \varepsilon \quad \forall n \geq n_2.$$

Sei $n_\varepsilon = \max\{2n_1, 2n_2 + 1\}$. Dann folgt für $n \geq n_\varepsilon$

$$|s_n - s| < \varepsilon.$$

□

Beispiel II.4.9 (alternierende harmonische Reihe).

(1) $$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots = \log 2$$

(2) $$\sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots = \frac{\pi}{4}.$$

## II.5. Absolute Konvergenz

Definition II.5.1. Eine Reihe $\sum x_n$ heißt absolut konvergent, wenn die Reihe $\sum |x_n|$ konvergent ist. Eine konvergente Reihe, die nicht absolut konvergent ist, heißt bedingt konvergent.

Satz II.5.2. *Eine absolut konvergente Reihe ist stets konvergent.*

Beweis. Sei $\sum x_n$ absolut konvergent und $\varepsilon > 0$. Gemäß Satz II.4.6 (S. 48) gibt es ein $n_\varepsilon \in \mathbb{N}$ mit

$$\sum_{k=m+1}^{n} |x_k| < \varepsilon \quad \forall n > m \geq n_\varepsilon$$

$$\Longrightarrow \Big|\sum_{k=m+1}^{n} x_k\Big| \leq \sum_{k=m+1}^{n} |x_k| < \varepsilon \quad \forall n > m \geq n_\varepsilon.$$

Also ist $\sum x_n$ gemäß Satz II.4.6 (S. 48) konvergent. □

Bemerkung II.5.3. Die Umkehrung von Satz II.5.2 ist i. a. falsch. Die alternierende harmonische Reihe ist nur bedingt konvergent.

Als nächstes geben wir einige wichtige Kriterien für die absolute Konvergenz von Reihen an.

Satz II.5.4 (Majorantenkriterium). *Sei $(\alpha_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}_+$, derart dass $\sum \alpha_n$ konvergiert. Es gebe ein $n_0 \in \mathbb{N}$ mit*

$$|x_n| \leq \alpha_n \quad \forall n \geq n_0.$$

*Dann ist $\sum x_n$ absolut konvergent.*

Beweis. $(\sum_{k=0}^{n} |x_k|)_{n\in\mathbb{N}}$ ist monoton wachsend und durch $\sum_{k=0}^{n_0-1} |x_k| + \sum_{k=0}^{\infty} \alpha_k$ nach oben beschränkt. Damit folgt die Behauptung aus Satz II.4.7 (S. 48). □

Beispiel II.5.5. Sei $k \geq 2$ fest. Wegen Beispiel II.4.2(2) (S. 46) und Satz II.5.4 ist $\sum_{n\geq 1} \frac{1}{n^k}$ absolut konvergent.

Satz II.5.6 (Wurzelkriterium). *Sei*

$$\alpha = \limsup_{n\to\infty} \sqrt[n]{|x_n|}.$$

*Dann gilt:*

(1) *Falls $\alpha < 1$ ist, ist $\sum x_n$ absolut konvergent.*
(2) *Falls $\alpha > 1$ ist, ist $\sum x_n$ divergent.*
(3) *Für $\alpha = 1$ ist keine Aussage möglich.*

Beweis. ad (1): Sei $\alpha < 1$ und $q \in (\alpha, 1)$. Wegen Satz II.3.9 (S. 44) gibt es ein $n_q \in \mathbb{N}$ mit

$$\sqrt[n]{|x_n|} \leq q \quad \forall n \geq n_q$$
$$\Longrightarrow |x_n| \leq q^n \quad \forall n \geq n_q.$$

Also ist die geometrische Reihe mit $a = q$ eine konvergente Majorante.
ad (2): Sei $\alpha > 1$. Dann folgt

$$\mathrm{Anz}\{n \in \mathbb{N} : \sqrt[n]{|x_n|} \geq 1\} = \infty$$
$$\Longrightarrow \mathrm{Anz}\{n \in \mathbb{N} : |x_n| \geq 1\} = \infty.$$

Also ist $(x_n)_{n\in\mathbb{N}}$ keine Nullfolge. Damit folgt die Behauptung aus Satz II.4.3 (S. 47).
AD (3): $\sum \frac{1}{n^2}$ ist absolut konvergent und

$$\lim_{n\to\infty} \sqrt[n]{\frac{1}{n^2}} = \left(\lim_{n\to\infty} \sqrt[n]{n}\right)^{-2} = 1.$$

$\sum \frac{1}{n}$ ist divergent und

$$\lim_{n\to\infty} \sqrt[n]{\frac{1}{n}} = \left(\lim_{n\to\infty} \sqrt[n]{n}\right)^{-1} = 1.$$

□

SATZ II.5.7 (QUOTIENTENKRITERIUM). *Es gebe ein $k_0 \in \mathbb{N}$ mit $x_n \neq 0$ für alle $n \geq k_0$.*

(1) *Gilt*

$$\limsup_{\substack{n\to\infty\\ n\geq k_0}} \left|\frac{x_{n+1}}{x_n}\right| < 1,$$

*so ist $\sum x_n$ absolut konvergent.*

(2) *Gibt es ein $k_1 \geq k_0$ mit*

$$\left|\frac{x_{n+1}}{x_n}\right| \geq 1 \quad \forall n \geq k_1,$$

*so ist $\sum x_n$ divergent.*

BEWEIS. AD (1): Wegen Satz II.3.9 (S. 44) gibt es ein $q \in (0,1)$ und ein $n_q \geq k_0$ mit

$$\left|\frac{x_{n+1}}{x_n}\right| \leq q \quad \forall n \geq n_q.$$

Durch Induktion folgt

$$|x_n| \leq |x_{n_q}| q^{n-n_q} = \frac{|x_{n_q}|}{q^{n_q}} q^n \quad \forall n \geq n_q.$$

Also ist $\sum cq^n$ mit $c = \frac{|x_{n_q}|}{q^{n_q}}$ eine konvergente Majorante.
AD (2): Für $n \geq k_1$ folgt

$$|x_n| \geq |x_{k_1}| > 0.$$

Also ist $(x_n)_{n\in\mathbb{N}}$ keine Nullfolge. □

BEISPIEL II.5.8. (1) $\sum \frac{n^2}{2^n}$ ist absolut konvergent.
(2) $\sum (\frac{1}{2})^{n+(-1)^n}$ ist absolut konvergent.
(3) $\sum \frac{z^n}{n!}$ ist für jedes $z \in \mathbb{C}$ absolut konvergent. Die Funktion $\exp : \mathbb{C} \to \mathbb{C}$, die jedem $z \in \mathbb{C}$ die Zahl $\displaystyle\sum_{n=0}^{\infty} \frac{z^n}{n!}$ zuordnet, heißt EXPONENTIALFUNKTION. Es gilt $\exp(\mathbb{R}) \subset \mathbb{R}$.

Beweis. ad (1):

$$\left|\frac{x_{n+1}}{x_n}\right| = \frac{(n+1)^2}{2^{n+1}}\frac{2^n}{n^2} = \frac{1}{2}\left(\frac{n+1}{n}\right)^2 \underset{n\to\infty}{\longrightarrow} \frac{1}{2}.$$

ad (2):

$$\begin{aligned}\left|\frac{x_{n+1}}{x_n}\right| &= 2^{-(n+1)-(-1)^{n+1}+n+(-1)^n} \\ &= 2^{-1+2(-1)^n} \\ &= \begin{cases} 2, & \text{falls } n \text{ gerade} \\ \frac{1}{8}, & \text{falls } n \text{ ungerade} \end{cases}\end{aligned}$$

Daher ist das Quotientenkriterium nicht anwendbar. Wegen

$$\begin{aligned}\sqrt[n]{|x_n|} &= 2^{-1-\frac{1}{n}(-1)^n} \\ &= \begin{cases} \frac{1}{2}\sqrt[n]{\frac{1}{2}}, & \text{falls } n \text{ gerade} \\ \frac{1}{2}\sqrt[n]{2}, & \text{falls } n \text{ ungerade} \end{cases}\end{aligned}$$

folgt aber $\limsup_{n\to\infty} \sqrt[n]{|x_n|} < 1$, so dass das Wurzelkriterium die behauptete Konvergenz liefert.
ad (3): Sei o.E. $z \neq 0$. Dann folgt

$$\left|\frac{x_{n+1}}{x_n}\right| = \frac{|z|^{n+1}}{(n+1)!}\frac{n!}{|z|^n} = \frac{|z|}{n+1} \underset{n\to\infty}{\longrightarrow} 0.$$

□

Definition II.5.9. Sei $\sigma : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Dann heißt $\sum x_{\sigma(n)}$ eine Umordnung von $\sum x_n$.

Beispiel II.5.10. Betrachte die alternierende harmonische Reihe

$$s = \sum_{n=1}^{\infty} (-1)^{n+1}\frac{1}{n}$$

und die Umordnung $\sigma : \mathbb{N}^* \to \mathbb{N}^*$ definiert durch

$$\sigma(n) = \begin{cases} 2m+1 & \text{falls } n = 3m+1, m \geq 0, \\ 4m+2 & \text{fals } n = 3m+2, m \geq 0, \\ 4m & \text{falls } n = 3m, m \geq 1. \end{cases}$$

Dann folgt

$$\begin{aligned}\sum_{n=1}^{\infty} x_{\sigma(n)} &= 1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \frac{1}{7}\dots \\ &= (1-\frac{1}{2}) - \frac{1}{4} + (\frac{1}{3} - \frac{1}{6}) - \frac{1}{8} + (\frac{1}{5} - \frac{1}{10}) - \frac{1}{12} + \dots\end{aligned}$$

$$= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \dots$$
$$= \frac{1}{2}s.$$

Also ist die Addition bei unendlich vielen Summanden i.a. nicht kommutativ.

SATZ II.5.11. *Die Reihe $\sum x_n$ sei absolut konvergent. Dann ist jede Umordnung auch absolut konvergent und es gilt*

$$\sum_{n=0}^{\infty} x_n = \sum_{n=0}^{\infty} x_{\sigma(n)} \quad \forall \sigma : \mathbb{N} \to \mathbb{N} \quad \textit{bijektiv.}$$

BEWEIS. Sei $\sigma : \mathbb{N} \to \mathbb{N}$ eine beliebige bijektive Abbildung und im Folgenden fest. Sei $\varepsilon > 0$. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit

$$\sum_{k=m+1}^{n} |x_k| < \varepsilon \quad \forall n > m \geq n_\varepsilon.$$

Sei

$$r_\varepsilon = \max\{k : \sigma(k) \leq n_\varepsilon\}.$$

Dann gilt

$$\mathbb{N}_{n_\varepsilon} \subset \sigma(\mathbb{N}_{r_\varepsilon}).$$

Für $r \geq r_\varepsilon$ und $n \geq n_\varepsilon$ folgt daher mit $N = \max\{r, \sigma(0), \dots, \sigma(r)\}$

$$|\sum_{k=0}^{r} x_{\sigma(k)} - \sum_{k=0}^{n} x_k| \leq \sum_{k=n_\varepsilon+1}^{N} |x_k| < \varepsilon.$$

Also gilt für $r \geq r_\varepsilon$

$$|\sum_{k=0}^{r} x_{\sigma(k)} - \sum_{k=0}^{\infty} x_k| < \varepsilon.$$

Mithin konvergiert $\sum x_{\sigma(n)}$ und

$$\sum_{k=0}^{\infty} x_{\sigma(n)} = \sum_{n=0}^{\infty} x_n.$$

Analog folgt für $r \geq r_\varepsilon$ und $n \geq n_\varepsilon$

$$|\sum_{k=0}^{r} |x_{\sigma(k)}| - \sum_{k=0}^{n} |x_k|| \leq \sum_{k=n_\varepsilon+1}^{N} |x_k| < \varepsilon$$

und somit

$$|\sum_{k=0}^{r} |x_{\sigma(k)}| - \sum_{k=0}^{\infty} |x_k|| < \varepsilon.$$

Also ist $\sum x_{\sigma(n)}$ auch absolut konvergent. □

Bemerkung II.5.12 (Satz von Riemann). Sei $\sum x_n$ bedingt konvergent in $\mathbb{R}$ und $z \in \mathbb{R}$ beliebig. Man kann zeigen, dass es dann eine Umordnung von $\sum x_n$ gibt mit

$$\sum_{n=0}^{\infty} x_{\sigma(n)} = z.$$

Als nächstes wenden wir uns dem Problem zu, das Produkt zweier konvergenter Reihen zu berechnen. Für die Partialsummen erhalten wir

$$\begin{aligned}(\sum_{k=0}^{n} x_k) \cdot (\sum_{l=0}^{n} y_l) &= \sum_{k=0}^{n} \sum_{l=0}^{n} (x_k y_l) \\ (*) \qquad\qquad &= \sum_{m=0}^{2n} (\sum_{\substack{0 \le k,l \le n \\ k+l=m}} x_k y_l) \\ &= \sum_{m=0}^{n} (\sum_{k=0}^{m} x_k y_{m-k}) + \sum_{m=n+1}^{2n} (\sum_{\substack{0 \le k,l \le n \\ k+l=m}} x_k y_l).\end{aligned}$$

Formal erhalten wir somit

$$(\sum x_n) \cdot (\sum y_k) = \sum z_n$$

mit

$$z_n = \sum_{\substack{0 \le k,l \le n \\ k+l=n}} x_k y_l = \sum_{k=0}^{n} x_k y_{n-k} = \sum_{k=0}^{n} x_{n-k} y_k.$$

Der Ausdruck $z_n$ heißt Cauchyprodukt.
Der folgende Satz präzisiert unsere formale Argumentation.

Satz II.5.13. *Die Reihen $\sum x_n$ und $\sum y_n$ seien absolut konvergent. Dann ist die Reihe $\sum z_n$ mit*

$$z_n = \sum_{k=0}^{n} x_k y_{n-k}$$

*ebenfalls absolut konvergent und es gilt*

$$(\sum_{n=0}^{\infty} x_n) \cdot (\sum_{n=0}^{\infty} y_n) = \sum_{n=0}^{\infty} z_n.$$

Beweis. Durch Anwenden von $(*)$ auf $\sum |x_n|$ und $\sum |y_n|$ erhalten wir

$$\begin{aligned}\sum_{m=0}^{n} |z_m| &\le \sum_{m=0}^{n} \sum_{\substack{0 \le k,l \le n \\ k+l=m}} |x_k||y_l| \\ &\le (\sum_{k=0}^{n} |x_k|) \cdot (\sum_{l=0}^{n} |y_l|)\end{aligned}$$

$$\leq (\sum_{k=0}^{\infty} |x_k|) \cdot (\sum_{l=0}^{\infty} |y_l|).$$

Hieraus und aus Satz II.2.2 (S. 35) folgt die absolute Konvergenz von $\sum z_n$.
Sei

$$a = \sum_{k=0}^{\infty} |x_k| \quad , \quad b = \sum_{k=0}^{\infty} |y_k|.$$

O.E. ist $a > 0$ und $b > 0$. Sei $\varepsilon > 0$. Dann gibt es ein $n_1 \in \mathbb{N}$ und ein $n_2 \in \mathbb{N}$ mit

$$\sum_{k=m+1}^{n} |x_k| < \frac{\varepsilon}{2b} \quad \forall n > m \geq n_1$$

und

$$\sum_{k=m+1}^{n} |y_k| < \frac{\varepsilon}{2a} \quad \forall n > m \geq n_2.$$

Setze $n_\varepsilon = \max\{n_1, n_2\} + 1$. Dann folgt für $n \geq 2n_\varepsilon$

$$\begin{aligned}
&|\sum_{k=0}^{n} z_k - (\sum_{k=0}^{n} x_k) \cdot (\sum_{k=0}^{n} y_k)| \\
=&|\sum_{m=n+1}^{2n} (\sum_{\substack{0 \leq k,l \leq n \\ k+l=m}} x_k y_l)| \quad \text{wegen } (*) \\
\leq &\sum_{m=n+1}^{2n} \sum_{\substack{0 \leq k,l \leq n \\ k+l=m}} |x_k||y_l| \\
\leq &\sum_{m=n+1}^{2n} \sum_{k=0}^{n} |x_k||y_{m-k}| \\
= &\sum_{k=0}^{n} |x_k| (\sum_{l=n+1-k}^{2n-k} |y_l|) \\
\leq &\sum_{k=0}^{n_\varepsilon} |x_k| (\sum_{l=n_\varepsilon+1}^{2n} |y_l|) + \sum_{k=n_\varepsilon+1}^{n} |x_k| (\sum_{l=n+1-k}^{2n-k} |y_l|) \\
\leq &a \cdot \sum_{l=n_2+1}^{2n} |y_l| + b \cdot \sum_{k=n_1+1}^{2n} |x_k| \\
< &\varepsilon.
\end{aligned}$$

Also ist

$$\left(\sum_{k=0}^{n} z_k - (\sum_{k=0}^{n} x_k) \cdot (\sum_{k=0}^{n} y_k)\right)_{n\in\mathbb{N}}$$

eine Nullfolge. Damit folgt die Behauptung aus Lemma II.1.17 (S. 32) und Satz II.1.18 (S. 32). □

Bemerkung II.5.14. (1) Wenn die Reihen $\sum x_n$ und $\sum y_n$ nur bedingt konvergent sind, ist Satz II.5.13 i. a. falsch. Um dies einzusehen, betrachte

$$x_n = y_n = \frac{(-1)^n}{\sqrt{n}} \quad \forall n \in \mathbb{N}^*.$$

Wegen Satz II.4.8 (S. 48) ist $\sum \frac{(-1)^n}{\sqrt{n}}$ konvergent. Wegen

$$\frac{1}{\sqrt{n}} \geq \frac{1}{n} \quad \forall n \in \mathbb{N}^*$$

und Beispiel II.4.2(3) (S. 46) ist $\sum \frac{(-1)^n}{\sqrt{n}}$ nur bedingt konvergent. Für dieses Beispiel ist

$$\begin{aligned}
|z_n| &= |\sum_{k=1}^{n-1} \frac{(-1)^k}{\sqrt{k}} \frac{(-1)^{n-k}}{\sqrt{n-k}}| \\
&= \sum_{k=1}^{n-1} \frac{1}{\sqrt{k(n-k)}} \\
&\geq \sum_{k=1}^{n-1} \frac{2}{k+n-k} \qquad \text{wegen} \quad \sqrt{ab} \leq \frac{a+b}{2} \\
&= 2\frac{n-1}{n} \\
&\geq 1 \quad \forall n \geq 2.
\end{aligned}$$

Also ist $\sum z_n$ nicht konvergent.
(2) Man kann zeigen, dass Satz II.5.13 gültig bleibt, wenn eine der Reihen absolut und die andere bedingt konvergent ist.

Beispiel II.5.15. Für die Exponentialfunktion aus Beispiel II.5.8 (3) gilt

$$\begin{aligned}
e^{z_1+z_2} &= \exp(z_1 + z_2) \\
&= \exp(z_1)\exp(z_2) \\
&= e^{z_1}e^{z_2} \quad \forall z_1, z_2 \in \mathbb{C}
\end{aligned}$$

Insbesondere gilt

$$e^{-z} = \frac{1}{e^z} \quad \forall z \in \mathbb{C}.$$

Beweis. Aus Satz II.5.13 und der Binomischen Formel folgt

$$\begin{aligned} e^{z_1} \cdot e^{z_2} &= \left(\sum_{n=0}^{\infty} \frac{z_1^n}{n!}\right) \cdot \left(\sum_{n=0}^{\infty} \frac{z_2^n}{n!}\right) \\ &= \sum_{n=0}^{\infty} \left\{\sum_{k=0}^{n} \frac{z_1^k}{k!} \frac{z_2^{n-k}}{(n-k)!}\right\} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left\{\sum_{k=0}^{n} \binom{n}{k} z_1^k z_2^{n-k}\right\} \\ &= e^{z_1+z_2}. \end{aligned}$$

Insbesondere gilt für jedes $z \in \mathbb{C}$

$$1 = e^0 = e^{z-z} = e^z \cdot e^{-z}.$$

Also ist $e^z \neq 0$ für alle $z \in \mathbb{C}$ und

$$e^{-z} = \frac{1}{e^z}.$$

□

## II.6. Potenzreihen

Definition II.6.1. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $z_0 \in \mathbb{K}$. Dann heißt die Reihe $\sum a_n(z-z_0)^n$ mit $z \in \mathbb{K}$ eine Potenzreihe mit Koeffizienten $a_n$ und Entwicklungspunkt $z_0$.

Beispiel II.6.2. (1) $\sum \frac{z^n}{n!}$; $a_n = \frac{1}{n!}$, $z_0 = 0$. Die Potenzreihe konvergiert für alle $z \in \mathbb{C}$ absolut.
(2) $\sum z^n$; $a_n = 1$ ,$z_0 = 0$. Die Potenzreihe konvergiert für alle $z \in \mathbb{C}$ mit $|z| < 1$ absolut.
(3) $\sum n!z^n$; $a_n = n!$, $z_0 = 0$. Wegen

$$\left|\frac{(n+1)!z^{n+1}}{n!z^n}\right| = (n+1)|z| \underset{n\to\infty}{\longrightarrow} \infty \quad \forall z \neq 0$$

konvergiert die Potenzreihe nur für $z = 0$.

Satz II.6.3. *Sei $\sum a_n(z-z_0)^n$ eine Potenzreihe und*

$$\rho = \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}} \in [0,\infty].$$

*Dann gilt:*

(1) *Die Potenzreihe konvergiert absolut für alle $z \in \mathbb{K}$ mit $|z - z_0| < \rho$.*
(2) *Die Potenzreihe divergiert für alle $z \in \mathbb{K}$ mit $|z-z_0| > \rho$.*

*Die Zahl $\rho$ heißt* Konvergenzradius *der Potenzreihe; die Menge $\{z \in \mathbb{K} : |z - z_0| < \rho\}$ heißt* Konvergenzkreis *der Potenzreihe.*

Beweis. Wegen

$$\begin{aligned}&\limsup_{n\to\infty}\sqrt[n]{|a_n||z-z_0|^n}\\=&|z-z_0|\limsup_{n\to\infty}\sqrt[n]{|a_n|}\\=&\frac{|z-z_0|}{\rho}\end{aligned}$$

folgt die Behauptung aus Satz II.5.6 (S. 50). □

Satz II.6.4. *Sei $\sum a_n(z-z_0)^n$ eine Potenzreihe und es existiere $\lim\limits_{n\to\infty}\frac{|a_n|}{|a_{n+1}|}$ in $\overline{\mathbb{R}}$. Dann gilt für den Konvergenzradius der Potenzreihe*

$$\rho=\lim_{n\to\infty}\frac{|a_n|}{|a_{n+1}|}.$$

Beweis. Sei $\alpha=\lim\limits_{n\to\infty}\frac{|a_n|}{|a_{n+1}|}$. Dann folgt

$$\left|\frac{a_{n+1}(z-z_0)^{n+1}}{a_n(z-z_0)^n}\right|=\frac{|a_{n+1}|}{|a_n|}|z-z_0|\underset{n\to\infty}{\longrightarrow}\frac{|z-z_0|}{\alpha}.$$

Damit folgt die Behauptung aus Satz II.5.7 (S. 51). □

Beispiel II.6.5.

(1) $\displaystyle\sum\frac{z^n}{n!}$

$$\Longrightarrow\frac{|a_n|}{|a_{n+1}|}=\frac{(n+1)!}{n!}=n+1\underset{n\to\infty}{\longrightarrow}\infty$$

$$\Longrightarrow\rho=\infty.$$

(2) $\displaystyle\sum n^m z^n \quad m\in\mathbb{N}$ fest

$$\Longrightarrow\frac{|a_n|}{|a_{n+1}|}=\frac{n^m}{(n+1)^m}=\left(\frac{n}{n+1}\right)^m\underset{n\to\infty}{\longrightarrow}1$$

$$\Longrightarrow\rho=1.$$

(3) $\displaystyle\sum\frac{z^{n^2}}{n!}$

$$\Longrightarrow a_n=\begin{cases}\frac{1}{j!} & n=j^2\\ 0 & \text{sonst}\end{cases}$$

$$\Longrightarrow\limsup_{n\to\infty}\sqrt[n]{|a_n|}=\limsup_{j\to\infty}\left(\frac{1}{j!}\right)^{1/j^2}$$

$$1\le j!\le j^j\Longrightarrow\frac{1}{j^j}\le\frac{1}{j!}\le 1$$

$$\Longrightarrow 1\ge\limsup_{j\to\infty}\left(\frac{1}{j!}\right)^{1/j^2}\ge\limsup_{j\to\infty}\left(\left(\frac{1}{j^j}\right)^{1/j}\right)^{1/j}=\limsup_{j\to\infty}\frac{1}{\sqrt[j]{j}}=1$$

$\Longrightarrow \rho = 1.$

Bemerkung II.6.6. Über das Konvergenzverhalten einer Potenzreihe auf dem Rand ihres Konvergenzkreises kann keine allgemeine Aussage getroffen werden:

(1) $\sum z^n \Longrightarrow \rho = 1$ und Divergenz für alle $z$ mit $|z| = 1$.
(2) $\sum \frac{z^n}{n} \Longrightarrow \rho = 1$ und Konvergenz für $z = -1$ und Divergenz für $z = 1$.
(3) $\sum \frac{z^n}{n^2} \Longrightarrow \rho = 1$ und absolute Konvergenz für alle $z$ mit $|z| = 1$.

Die folgenden Rechenregeln sind eine unmittelbare Konsequenz der Sätze II.4.5 (S. 48) und II.5.13 (S. 54).

Satz II.6.7. *Seien $\alpha \in \mathbb{K}$ und $\sum a_n(z-z_0)^n$ und $\sum b_n(z-z_0)^n$ zwei Potenzreihen mit gleichem Entwicklungspunkt $z_0$ und Konvergenzradien $\rho_1$ und $\rho_2$. Dann gilt:*

(1)
$$\sum_{n=0}^{\infty} \alpha a_n (z-z_0)^n = \alpha \sum_{n=0}^{\infty} a_n (z-z_0)^n \quad \forall |z-z_0| < \rho_1$$

(2)
$$\sum_{n=0}^{\infty} (a_n + b_n)(z-z_0)^n = \sum_{n=0}^{\infty} a_n (z-z_0)^n + \sum_{n=0}^{\infty} b_n (z-z_0)^n \quad \forall |z-z_0| < \min\{\rho_1, \rho_2\}$$

(3)
$$\sum_{n=0}^{\infty} \Big(\sum_{k=0}^{n} a_k b_{n-k}\Big)(z-z_0)^n = \Big(\sum_{n=0}^{\infty} a_n (z-z_0)^n\Big)\Big(\sum_{n=0}^{\infty} b_n (z-z_0)^n\Big) \quad \forall |z-z_0| < \min\{\rho_1, \rho_2\}.$$

Sei $\sum a_n(z-z_0)^n$ eine Potenzreihe mit Konvergenzradius $\rho > 0$. Dann wird durch

$$z \mapsto f(z) = \sum_{n=0}^{\infty} a_n (z-z_0)^n$$

eine Funktion $\{z \in \mathbb{K} : |z-z_0| < \rho\} \to \mathbb{K}$ definiert. Die Potenzreihe stellt die Funktion $f$ auf dem Konvergenzkreis dar.
Die Funktion

$$f(z) = \frac{1}{1-z}$$

ist auf $\mathbb{C}\backslash\{1\}$ definiert. Auf $\{z \in \mathbb{C} : |z| < 1\}$ wird sie durch die Potenzreihe $\sum z^n$ dargestellt.

Darstellungen von Funktionen als Potenzreihen werden in späteren Abschnitten eine wichtige Rolle spielen.

KAPITEL III

# Stetige Funktionen

Im Zentrum dieses Kapitels steht der Begriff der Stetigkeit von Funktionen zwischen normierten Vektorräumen. Zunächst führen wir den Begriff eines normierten Vektorraumes ein und leiten einige topologische Grundbegriffe her. Danach führen wir den Begriff der Stetigkeit ein und leiten Folgerungen daraus ab. Dann betrachten wir Funktionen auf zusammenhängenden und kompakten Mengen und daraus resultierende Konsequenzen. Zum Abschluss stellen wir einige spezielle Eigenschaften von stetigen Funktionen von $\mathbb{K}$ nach $\mathbb{K}$ bzw. $\mathbb{R}$ nach $\mathbb{R}$ zusammen.

## III.1. Normierte Vektorräume

Definition III.1.1. Eine Menge $X$ heißt $\mathbb{K}$-Vektorraum, wenn es zwei Abbildungen

$$+ : X \times X \to X \quad \text{und} \cdot : \mathbb{K} \times X \to X$$

mit folgenden Eigenschaften gibt:

(VR1) $(X, +)$ ist eine Abelsche Gruppe mit neutralem Element $o$.

(VR2) Für alle $\alpha, \beta \in \mathbb{K}$ und alle $x, y \in X$ gelten die Distributivgesetze

$$\begin{aligned} \alpha \cdot (x + y) &= (\alpha \cdot x) + (\alpha \cdot y) \\ (\alpha + \beta) \cdot x &= (\alpha \cdot x) + (\beta \cdot x) \\ \alpha \cdot (\beta \cdot x) &= (\alpha\beta) \cdot x \end{aligned}$$

(VR3) $1 \cdot x = x$ für alle $x \in X$.

Bemerkung III.1.2. Aus obigen Axiomen folgt insbesondere

$$\begin{aligned} 0 \cdot x &= o \quad \forall x \in X \\ (-1) \cdot x &= \widetilde{x} \quad \forall x \in X, \end{aligned}$$

wobei $\widetilde{x}$ das inverse Element zu $x$ in $(X, +)$ ist.

Beispiel III.1.3. Beispiele für Vektorräume:

(1) $\mathbb{C}$ ist ein $\mathbb{R}$-Vektorraum.
(2) $\mathbb{K}^m = \underbrace{\mathbb{K} \times \ldots \times \mathbb{K}}_{m-\text{ mal}}, m \in \mathbb{N}^*$, ist ein $\mathbb{K}$-Vektorraum.
(3) Der Raum $\mathbb{K}^{\mathbb{N}}$ der Folgen in $\mathbb{K}$ ist ein $\mathbb{K}$-Vektorraum.

(4) Der Raum $\ell_\infty$ aller beschränkten Folgen in $\mathbb{K}$ ist ein $\mathbb{K}$-Vektorraum.
(5) Der Raum $\ell_1$ aller $(x_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ derart, dass $\sum x_n$ absolut konvergiert, ist ein $\mathbb{K}$-Vektorraum.
(6) Der Raum $\ell_2$ aller $(x_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ derart, dass $\sum |x_n|^2$ konvergiert, ist ein $\mathbb{K}$-Vektorraum.

Beweis. ad (1)–(4): Sind klar.
ad (5): Folgt aus Satz II.4.5 (S. 48), der Dreiecksungleichung in $\mathbb{K}$ und Satz II.5.4 (S. 50).
ad (6): Folgt aus Satz II.4.5 (S. 48), der Abschätzung

$$|a+b|^2 \le 2(|a|^2 + |b|^2) \quad \forall a, b \in \mathbb{K}$$

und Satz II.5.4 (S. 50). □

Definition III.1.4. Sei $X$ ein $\mathbb{K}$-Vektorraum. Eine Abbildung $\|\cdot\| : X \to \mathbb{R}$ heißt Norm, wenn gilt:

(N1) $\quad \|x\| \ge 0 \quad \forall x \in X$ und $\|x\| = 0 \iff x = 0$,

(N2) $\quad \|\alpha x\| = |\alpha| \|x\| \quad \forall x \in X, \alpha \in \mathbb{K}$,

(N3) $\quad \|x+y\| \le \|x\| + \|y\| \quad \forall x, y \in X.$ (Dreiecksungleichung)

$(X, \|\cdot\|)$ heißt normierter $\mathbb{K}$ Vektorraum.

Bemerkung III.1.5. Aus der Dreiecksungleichung folgt die sog. umgekehrte Dreiecksungleichung

$$|\|x\| - \|y\|| \le \|x - y\|.$$

Denn:

$$\begin{aligned}
&\|x\| = \|y + (x-y)\| \le \|y\| + \|x-y\| \\
&\|y\| = \|x + (y-x)\| \le \|x\| + \|x-y\| \\
\Longrightarrow &\|x\| - \|y\| \le \|x-y\| \ \text{ und } \|y\| - \|x\| \le \|x-y\| \\
\Longrightarrow &|\|x\| - \|y\|| = \max\{\|x\| - \|y\|, \|y\| - \|x\|\} \le \|x-y\|.
\end{aligned}$$

Beispiel III.1.6. Beispiele für Normen:
(1) $|\cdot|$ ist eine Norm auf $\mathbb{R}$ bzw. $\mathbb{C}$.
(2) $\|x\|_\infty = \max\limits_{1\le k\le n} |x_k|$, $\|x\|_1 = \sum\limits_{k=1}^{n} |x_k|$, $\|x\|_2 = \{\sum\limits_{k=1}^{n} |x_k|^2\}^{1/2}$ sind Normen auf $\mathbb{K}^n$. Für alle $x \in \mathbb{K}^n$ gilt:
   (a) $\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty$,
   (b) $\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty$,
   (c) $\frac{1}{\sqrt{n}}\|x\|_1 \le \|x\|_2 \le \sqrt{n}\|x\|_1$.
(3) $\|x\|_\infty = \sup\limits_{n\in\mathbb{N}} |x_n|$ ist eine Norm auf $\ell_\infty$.
(4) $\|x\|_1 = \sum\limits_{n=0}^{\infty} |x_n|$ ist eine Norm auf $\ell_1$.

(5) $\|x\|_2 = \{\sum_{n=0}^{\infty} |x_n|^2\}^{1/2}$ ist eine Norm auf $\ell_2$.

Beweis. ad (1): Ist klar.
ad (2): Die Normeigenschaften von $\|\cdot\|_1$ und $\|\cdot\|_\infty$ sind klar, ebenso (N1) und (N2) für $\|\cdot\|_2$. Die Dreiecksungleichung folgt mittels der Cauchy-Schwarzschen Ungleichung

$$\sum_{k=1}^{n} a_k b_k \le \{\sum_{k=1}^{n} a_k^2\}^{1/2} \{\sum_{k=1}^{n} b_k^2\}^{1/2}$$

für alle $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}_+$. Abschätzungen (a) und (b) sind offensichtlich. Die erste Ungleichung (c) folgt aus der Cauchy-Schwarzschen Ungleichung, die zweite Ungleichung (c) folgt aus (a) und (b).
ad (3): Ist klar.
ad (4), (5): (N1) und (N2) sind klar. (N3) folgt wie bei (2) zunächst für die Partialsummen und dann durch Grenzübergang. □

Im Folgenden sei stets $(X, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum. Die folgende Vereinbarung erleichtert die Notationen.

Definition III.1.7. Für $x_0 \in X$ und $r \in \mathbb{R}_+^*$ heißt

$$B(x_0; r) = \{x \in X : \|x - x_0\| < r\}$$

der (offene) Ball um $x_0$ mit Radius $r$. $B(0; 1)$ heißt auch Einheitsball.

In Abbildung III.1.1 sind die Einheitsbälle in $\mathbb{R}^2$ für die Normen $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_\infty$ skizziert.

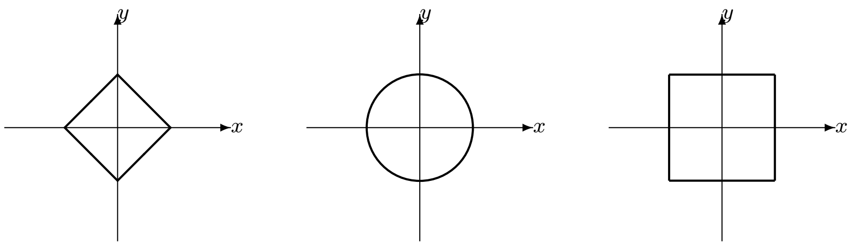


Abbildung III.1.1. Einheitsbälle in $\mathbb{R}^2$ für die Normen $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_\infty$ (v.l.n.r)

Bemerkung III.1.8. In einem normierten Vektorraum kann man ganz analog zu $\mathbb{K}$ die Begriffe Folge, konvergente Folge, Häufungspunkt einer Folge, Cauchyfolge, Reihe, absolute Konvergenz einer Reihe definieren. Man muss überall den Betrag $|\cdot|$ durch die Norm $\|\cdot\|$ ersetzen. So erhält man z.B.

$$(x_n)_{n\in\mathbb{N}} \subset X \text{ konvergiert gegen } x \in X$$
$$\iff \forall \varepsilon > 0 \ \exists n_\varepsilon \in \mathbb{N} \ \forall n \ge n_\varepsilon : \|x_n - x\| < \varepsilon$$

oder

$$\begin{aligned}&(x_n)_{n\in\mathbb{N}} \subset X \text{ ist eine Cauchyfolge}\\ \iff &\forall \varepsilon > 0\ \exists n_\varepsilon \in \mathbb{N}\ \forall n, m \geq n_\varepsilon : \|x_n - x_m\| < \varepsilon.\end{aligned}$$

Beispiel III.1.6 zeigt, dass man auf einen Vektorraum verschiedene Normen definieren kann. Dann stellt sich natürlich die Frage, ob z.B. eine Folge, die bezüglich der einen Norm konvergiert, auch bezüglich einer anderen Norm konvergiert. Dies ist sicherlich der Fall, wenn zwischen den Normen eine Abschätzung analog zu denen in Beispiel III.1.6 (2) gilt. Dies führt auf folgende Definition:

Definition III.1.9. Seien $X$ ein $\mathbb{K}$-Vektorraum und $\|\cdot\|_a$ und $\|\cdot\|_b$ zwei Normen auf $X$. Dann heißen $\|\cdot\|_a$ und $\|\cdot\|_b$ Äquivalent, wenn es zwei Konstanten $\underline{c}$ und $\overline{c}$ mit $\underline{c} > 0$, $\overline{c} > 0$ und

$$\underline{c}\|x\|_a \leq \|x\|_b \leq \overline{c}\|x\|_a \quad \forall x \in X$$

gibt.

Bemerkung III.1.10. (1) Beispiel III.1.6 (2) zeigt, dass die Normen $\|\cdot\|_1, \|\cdot\|_2$ und $\|\cdot\|_\infty$ auf $\mathbb{K}^n$ äquivalent sind.
(2) Sind $\|\cdot\|_a$ und $\|\cdot\|_b$ zwei äquivalente Normen auf dem $\mathbb{K}$-Vektorraum $X$, so ist eine Folge in $X$ genau dann bzgl. $\|\cdot\|_a$ konvergent, wenn sie bzgl. $\|\cdot\|_b$ konvergent ist. Gleiches gilt natürlich auch für Cauchyfolgen, Häufungspunkte usw.

Bemerkung III.1.10 zeigt, wie wesentlich der folgende Satz ist.

Satz III.1.11. *Sei $n \in \mathbb{N}^*$. Auf $\mathbb{K}^n$ sind alle Normen äquivalent.*

Beweis. Wie man sich leicht überlegt, reicht es zu zeigen, dass jede Norm $\|\cdot\|$ auf $\mathbb{K}^n$ zu $\|\cdot\|_1$ äquivalent ist. Sei also $\|\cdot\|$ eine beliebige Norm auf $\mathbb{K}^n$.
Für $1 \leq i \leq n$ sei

$$e_i = (0, \ldots, 0, \underbrace{1}_{i\text{-te Komponente}}, 0, \ldots, 0)$$

der $i$-te Einheitsvektor. Dann ist

$$x = \sum_{i=1}^n x_i e_i \quad \forall x \in \mathbb{K}^n.$$

Definiere

$$\begin{aligned}C_i &= \|e_i\| \quad 1 \leq i \leq n\\ C &= \max\{C_1, \ldots, C_n\}.\end{aligned}$$

Dann folgt aus der Dreiecksungleichung

$$\|x\| = \|\sum_{i=1}^n x_i e_i\|$$

$$\begin{aligned} &\leq \sum_{i=1}^{n} |x_i| \|e_i\| \\ &\leq C\|x\|_1. \end{aligned}$$

Also ist $\overline{c} = C$.
Sei nun

$$M = \{\|x\| : x \in \mathbb{K}^n, \|x\|_1 = 1\} \subset \mathbb{R}.$$

$M$ ist nicht leer ($C_i \in M, 1 \leq i \leq n$) und durch 0 nach unten beschränkt. Also existiert

$$\rho = \inf M$$

und es ist $\rho \geq 0$.
Angenommen, es wäre $\rho = 0$. Dann existiert eine Folge $(x_k)_{k\in\mathbb{N}} \subset \mathbb{K}^n$ mit

$$\|x_k\| \leq \frac{1}{k} \ , \ \|x_k\|_1 = 1 \quad \forall k \in \mathbb{N}^*.$$

Da für die $i$-te Komponente $x_{k,i}$ von $x_k$ gilt

$$|x_{k,i}| \leq \|x_k\|_1 = 1,$$

folgt aus Satz II.2.4 (S. 39), dass es ein $x^* \in \mathbb{K}^n$ und eine Teilfolge $(x_{k_l})_{l\in\mathbb{N}}$ von $(x_k)_{k\in\mathbb{N}}$ gibt mit

$$x_{k_l,i} \underset{l\to\infty}{\longrightarrow} x_i^* \quad \forall 1 \leq i \leq n.$$

Also gilt

$$\begin{aligned} \|x^*\| &\leq \|x_{k_l}\| + \|x^* - x_{k_l}\| \\ &\leq \frac{1}{k_l} + C\|x^* - x_{k_l}\|_1 \\ &\underset{l\to\infty}{\longrightarrow} 0. \end{aligned}$$

Mithin ist $\|x^*\| = 0$ und somit $x^* = 0$. Andererseits gilt

$$\begin{aligned} \|x^*\|_1 &\geq \|x_{k_l}\|_1 - \|x^* - x_{k_l}\|_1 \\ &= 1 - \|x^* - x_{k_l}\|_1 \\ &\underset{l\to\infty}{\longrightarrow} 1. \end{aligned}$$

Dies ist ein Widerspruch. Also ist $\rho > 0$.
Sei nun $x \in \mathbb{K}^n, x \neq 0$, beliebig. Dann folgt $\|\frac{1}{\|x\|_1}x\| \in M$ und somit

$$\frac{1}{\|x\|_1}\|x\| = \|\frac{1}{\|x\|_1}x\| \geq \rho > 0$$

und somit

$$(*) \qquad \rho\|x\|_1 \leq \|x\| \quad \forall x \in \mathbb{K}^n, x \neq 0.$$

Da $(*)$ offensichtlich für 0 gilt, ist die Behauptung damit bewiesen. □

Bemerkung III.1.12. Satz III.1.11 gilt in unendlich dimensionalen Räumen i. a. nicht. So ist z.B. $\|\cdot\|_\infty$ auch eine Norm auf $\ell_1$. Für die Folgen $(x_n^{(m)})_{n\in\mathbb{N}}$ mit

$$x_n^{(m)} = \begin{cases} 1 & n \le m \\ 0 & n > m \end{cases}$$

gilt

$$\|x^{(m)}\|_\infty = 1 \ \text{ und } \|x^{(m)}\|_1 = m+1.$$

Daher können $\|\cdot\|_1$ und $\|\cdot\|_\infty$ auf $\ell_1$ nicht äquivalent sein.

Wir haben in Satz II.2.8 (S. 41) gezeigt, dass in $\mathbb{K}$ jede Cauchyfolge konvergent ist. Diese Eigenschaft ist so zentral, dass normierte Vektorräume mit dieser Eigenschaft besonders bezeichnet werden.

Definition III.1.13. Ein normierter $\mathbb{K}$-Vektorraum $(X, \|\cdot\|)$ heißt vollständig oder kurz ($\mathbb{K}$)-Banachraum, wenn jede Cauchyfolge in $X$ konvergent ist.

Satz III.1.14. *Die folgenden normierten* $\mathbb{K}$-*Vektorräume sind vollständig:*

(1) $\mathbb{K}$,
(2) $\mathbb{K}^n, n \in \mathbb{N}^*$,
(3) $(\ell_1, \|\cdot\|_1)$,
(4) $(\ell_2, \|\cdot\|_2)$,
(5) $(\ell_\infty, \|\cdot\|_\infty)$.

Beweis. ad (1): Satz II.2.8 (S. 41).
ad (2): Wegen Satz III.1.11 brauchen wir die Behauptung nur für die Norm $\|\cdot\|_\infty$ zu zeigen. Sei also $(x_k)_{k\in\mathbb{N}}$ eine Cauchyfolge in $(\mathbb{K}^n, \|\cdot\|_\infty)$. Dann ist für jedes $i \in \mathbb{N}_n^*$ die Folge $(x_{k,i})_{k\in\mathbb{N}}$ der $i$-ten Komponenten eine Cauchyfolge in $\mathbb{K}$ und konvergiert gemäß Satz II.2.8 (S. 41) gegen ein $x_i^* \in \mathbb{K}$. Sei $x^* = (x_1^*, \ldots, x_n^*) \in \mathbb{K}^n$ und $\varepsilon > 0$ beliebig. Dann gibt es Zahlen $k_{\varepsilon_i} \in \mathbb{N}, 1 \le i \le n$, mit

$$|x_i^* - x_{k,i}| < \varepsilon \quad \forall k \ge k_{\varepsilon_i}, 1 \le i \le n.$$

Sei

$$k_\varepsilon = \max_{1\le i\le n} k_{\varepsilon_i}.$$

Dann gilt für alle $k \ge k_\varepsilon$

$$\|x^* - x_k\|_\infty = \max_{1\le i\le n} |x_i^* - x_{k,i}| < \varepsilon.$$

Damit ist die Behauptung bewiesen.
ad (3): Sei $(x_k)_{k\in\mathbb{N}} \subset \ell_1$ eine Cauchyfolge bzgl. $\|\cdot\|_1$. Wir bezeichnen mit $x_{k,n}$ die $n$-te Komponente von $x_k, n \in \mathbb{N}$. Dann gilt:

$$\forall \varepsilon > 0 \, \exists \, k_\varepsilon \in \mathbb{N} \, \forall k, l \ge k_\varepsilon : \|x_k - x_l\|_1 = \sum_{n=0}^{\infty} |x_{k,n} - x_{l,n}| < \varepsilon$$

$$\Longrightarrow \forall \varepsilon > 0 \; \exists \; k_\varepsilon \in \mathbb{N} \; \forall k, l \geq k_\varepsilon \; \forall n \in \mathbb{N} : |x_{k,n} - x_{l,n}| < \varepsilon.$$

Wegen Satz II.2.8 (S. 41) gibt es also zu jedem $n \in \mathbb{N}$ ein $x_n^* \in \mathbb{K}$ mit

$$x_{k,n} \underset{k \to \infty}{\longrightarrow} x_n^*.$$

Sei $x^* = (x_n^*)_{n \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$. Wir wollen zunächst zeigen, dass $x^* \in \ell_1$ ist. Aus Bemerkung III.1.5 folgt, dass $(\|x_k\|_1)_{k \in \mathbb{N}}$ eine Cauchyfolge in $\mathbb{R}$ und somit beschränkt ist. Also gilt

$$C = \sup_{k \in \mathbb{N}} \|x_k\|_1 < \infty.$$

Sei nun $N \in \mathbb{N}$ beliebig. Dann gibt es Zahlen $k_{\varepsilon_0}, \ldots, k_{\varepsilon_N} \in \mathbb{N}$ mit

$$|x_i^* - x_{k,i}| < \frac{1}{N+1} \qquad \forall k \geq k_{\varepsilon_i}, 0 \leq i \leq N.$$

Sei $k_\varepsilon = \max\limits_{1 \leq i \leq N} k_{\varepsilon_i}$. Dann folgt

$$\begin{aligned}
\sum_{i=0}^{N} |x_i^*| &\leq \sum_{i=0}^{N} |x_i^* - x_{k_\varepsilon,i}| + \sum_{i=0}^{N} |x_{k_\varepsilon,i}| \\
&\leq \sum_{i=0}^{N} \frac{1}{N+1} + \|x_{k_\varepsilon}\|_1 \\
&\leq 1 + C.
\end{aligned}$$

Da $C$ nicht von $N$ abhängt und $N$ beliebig war, folgt, dass $\sum x_i^*$ absolut konvergiert, d.h., $x^* \in \ell_1$.
Wir müssen noch zeigen: $\|x^* - x_k\|_1 \underset{k \to \infty}{\longrightarrow} 0$. Angenommen, dies gelte nicht. Dann gibt es ein $\varepsilon_0 > 0$, so dass gilt

(i) $$\forall k \in \mathbb{N} \; \exists \; m \geq k : \|x^* - x_m\|_1 > \varepsilon_0.$$

Da $(x_k)_{k \in \mathbb{N}} \subset \ell_1$ eine Cauchyfolge ist, gibt es ein $k_0 \in \mathbb{N}$ mit

(ii) $$\|x_k - x_l\|_1 < \frac{\varepsilon_0}{4} \quad \forall k, l \geq k_0.$$

Zu $k_0$ gibt es wegen (i) ein $m_0 \geq k_0$ mit

(iii) $$\|x^* - x_{m_0}\|_1 > \varepsilon_0.$$

Da $\sum |x_i^* - x_{m_0,i}|$ absolut konvergiert, gibt es gemäß Satz II.4.6 (S. 48) ein $n_0 \in \mathbb{N}$ mit

(iv) $$\sum_{i=n_0}^{\infty} |x_i^* - x_{m_0,i}| < \frac{\varepsilon_0}{4}.$$

Wegen $x_{k,i} \to x_i^*$ für $i \in \mathbb{N}$ gibt es ein $l_0 \geq k_0$ mit

(v) $$|x_{l_0,i} - x_i^*| < \frac{\varepsilon_0}{4n_0} \quad \forall 0 \leq i \leq n_0 - 1.$$

Aus (i)–(v) folgt

$$\begin{aligned}\varepsilon_0 < \|x^* - x_{m_0}\|_1 &\le \sum_{i=0}^{n_0-1} |x_i^* - x_{m_0,i}| + \frac{\varepsilon_0}{4}\\ &\le \sum_{i=0}^{n_0-1} |x_i^* - x_{l_0,i}| + \sum_{i=0}^{n_0-1} |x_{l_0,i} - x_{m_0,i}| + \frac{\varepsilon_0}{4}\\ &\le \sum_{i=0}^{n_0-1} \frac{\varepsilon_0}{4n_0} + \|x_{l_0} - x_{m_0}\|_1 + \frac{\varepsilon_0}{4}\\ &\le \frac{3}{4}\varepsilon_0.\end{aligned}$$

Dies ist wegen $\varepsilon_0 > 0$ ein Widerspruch. Also gilt $\|x^* - x_k\|_1 \underset{k\to\infty}{\longrightarrow} 0$. Damit ist die Vollständigkeit von $(\ell_1, \|\cdot\|_1)$ gezeigt.
AD (4),(5): Übungsaufgabe, Beweis analog zu (3). □

## III.2. Topologische Grundbegriffe

Im Folgenden bezeichnet $(X, \|\cdot\|)$ stets einen normierten $\mathbb{K}$-Vektorraum.

DEFINITION III.2.1. (1) Sei $M \subset X, M \neq \emptyset$. Ein Punkt $x \in M$ heißt INNERER PUNKT VON $M$, wenn es ein $\varepsilon > 0$ gibt mit $B(x;\varepsilon) \subset M$.
(2) $M \subset X$ heißt OFFEN, falls jedes $x \in M$ innerer Punkt von $M$ ist.
(3) Sei $x \in X$ und $U \subset X$ offen mit $x \in U$. Dann heißt $U$ eine UMGEBUNG von $x$. Die Menge aller Umgebungen von $x$ wird mit $\mathcal{U}(x)$ bezeichnet.

BEISPIEL III.2.2. Der Ball $B(x_0;r)$ mit $x_0 \in X, r > 0$, ist offen.

BEWEIS. Sei $x \in B(x_0;r)$. Dann ist

$$\varepsilon = \frac{1}{2}(r - \|x - x_0\|) > 0.$$

Für $y \in B(x;\varepsilon)$ folgt

$$\begin{aligned}\|x_0 - y\| &\le \|x_0 - x\| + \|x - y\|\\ &< \|x_0 - x\| + \frac{1}{2}(r - \|x_0 - x\|)\\ &= \frac{1}{2}r + \frac{1}{2}\|x_0 - x\|\\ &< r.\end{aligned}$$

□

BEMERKUNG III.2.3. (1) Der Begriff „offen“ hängt vom umgebenden Vektorraum ab. So ist z.B. das Intervall $(0,1)$ offen in $\mathbb{R}$ aber nicht in $\mathbb{R}^2$ (in $\mathbb{R}^2$ wird $(0,1)$ dabei mit $(0,1) \times \{0\}$ identifiziert).

(2) Sind $\|\cdot\|_a$ und $\|\cdot\|_b$ zwei äquivalente Normen auf $X$, so ist jede bzgl. $\|\cdot\|_a$ offene Menge auch bzgl. $\|\cdot\|_b$ offen und umgekehrt.

SATZ III.2.4. *Sei $\mathcal{O} \subset \mathcal{P}(X)$ die Menge aller offenen Mengen in $X$. Dann gilt:*

(1) $\emptyset \in \mathcal{O}$, $X \in \mathcal{O}$.
(2) $\mathcal{O}_\alpha \in \mathcal{O}$, $\alpha \in A$ *(A beliebige Indexmenge)* $\Longrightarrow \bigcup_{\alpha \in A} \mathcal{O}_\alpha \in \mathcal{O}$.
*(Beliebige Vereinigungen offener Mengen sind offen.)*
(3) $O_1, O_2 \in \mathcal{O} \longrightarrow O_1 \cap O_2 \in \mathcal{O}$.
*(Endliche Durchschnitte offener Mengen sind offen.)*

BEWEIS. Ist offensichtlich. □

BEMERKUNG III.2.5. Sei $X \neq \emptyset$ eine beliebige Menge und $\mathcal{T} \subset \mathcal{P}(X)$ eine Menge mit den Eigenschaften (1)–(3) von Satz III.2.4. Dann heißt $\mathcal{T}$ eine TOPOLOGIE auf $X$ und $(X, \mathcal{T})$ ein TOPOLOGISCHER RAUM. Viele der Eigenschaften, die wir für normierte Vektorräume beweisen, gelten allgemein für topologische Räume.

BEMERKUNG III.2.6. Der abzählbare Durchschnitt offener Mengen ist i. a. nicht offen. So ist z.B.

$$\bigcap_{n \in \mathbb{N}^*} B(0; \frac{1}{n}) = \{0\}$$

nicht offen.

DEFINITION III.2.7. $M \subset X$ heißt ABGESCHLOSSEN genau dann, wenn $M^c = X \backslash M$ offen ist.

SATZ III.2.8. *Es gilt:*

(1) $\emptyset$, $X$ *sind abgeschlossen.*
(2) *Beliebige Durchschnitte abgeschlossener Mengen sind wieder abgeschlossen.*
(3) *Endliche Vereinigungen abgeschlossener Mengen sind wieder abgeschlossen.*

BEWEIS. AD (1) Ist klar.
AD (2) $(\bigcap_{\alpha \in A} M_\alpha)^c = \bigcup_{\alpha \in A} (M_\alpha)^c$
AD (3) $(M_1 \cup M_2)^c = M_1^c \cap M_2^c$. □

BEMERKUNG III.2.9. Beliebige Vereinigungen abgeschlossener Mengen sind i. a. nicht abgeschlossen. So ist z.B.

$$\bigcup_{n \in \mathbb{N}^*} B(x; \frac{1}{n})^c = X \backslash \{x\}$$

nicht abgeschlossen.

Definition III.2.10. Sei $M \subset X$. Ein Punkt $x \in X$ heißt Berührungspunkt bzw. Häufungspunkt von $M$, wenn für jedes $U \in \mathcal{U}(x)$ gilt

$$U \cap M \neq \emptyset \quad \text{bzw. } M \cap U \backslash \{x\} \neq \emptyset.$$

Wir definieren

$$\overline{M} = \{x \in X : x \text{ ist Berührungspunkt von } M\}.$$

Beispiel III.2.11. (1) $M = \{\frac{1}{n} : n \in \mathbb{N}\} \subset \mathbb{R}$. Dann ist $\overline{M} = M \cup \{0\}$ und 0 ist einziger Häufungspunkt von $M$.
(2) $M = B(0; 1)$. Dann ist $\overline{M} = \{x \in X : \|x\| \leq 1\}$ und jeder Berührungspunkt ist auch Häufungspunkt.

Beweis. ad (1) Ist klar.
ad (2)

$$\begin{aligned}
&\|x\| < 1 : \Longrightarrow x \in M \Longrightarrow x \in \overline{M} \\
&\|x\| = 1 : \varepsilon > 0 \text{ beliebig } \Longrightarrow x_\varepsilon = (1 - \frac{\varepsilon}{2})x \in M \\
&\qquad\qquad \text{und } \|x - x_\varepsilon\| = \frac{\varepsilon}{2} < \varepsilon. \\
&\|x\| > 1 : \text{Sei } y \in B(x; r) \text{ mit } r = \frac{1}{2}(\|x\| - 1) > 0. \\
&\qquad\qquad \Longrightarrow \|y\| \geq \|x\| - \|x - y\| \\
&\qquad\qquad\qquad > \|x\| - \frac{1}{2}(\|x\| - 1) \\
&\qquad\qquad\qquad = \frac{1}{2}\|x\| + \frac{1}{2} \\
&\qquad\qquad\qquad > 1 \\
&\qquad\qquad \Longrightarrow B(x; r) \cap M = \emptyset.
\end{aligned}$$

□

Satz III.2.12. (1) *Es ist* $M \subset \overline{M}$.
(2) *Es ist* $M = \overline{M}$ *genau dann, wenn* $M$ *abgeschlossen ist.*

Beweis. ad (1): Ist klar.
ad (2): „$\Longrightarrow$“:

$$\begin{aligned}
x \in M^c = (\overline{M})^c &\Longrightarrow \exists U \in \mathcal{U}(x) \quad U \cap M = \emptyset \\
&\Longrightarrow U \subset M^c \\
&\Longrightarrow M^c \text{ offen} \\
&\Longrightarrow M = (M^c)^c \text{ abgeschlossen}
\end{aligned}$$

„$\Longleftarrow$“:

$$\begin{aligned}
M \text{ abgeschlossen} &\Longrightarrow M^c \text{ offen} \\
&\Longrightarrow \forall x \in M^c\ \exists\, U \in \mathcal{U}(x) \quad U \subset M^c \\
&\Longrightarrow \forall x \in M^c\ \exists\, U \in \mathcal{U}(x) \quad U \cap M = \emptyset
\end{aligned}$$

$$\begin{aligned} &\Longrightarrow M^c \subset (\overline{M})^c \\ &\Longrightarrow \overline{M} \subset M \\ &\Longrightarrow \overline{M} = M \quad \text{wegen (1).} \end{aligned}$$

□

Satz III.2.13. *Es gilt:*

(1) *$x$ ist Häufungspunkt von $M$ genau dann, wenn es eine Folge $(x_n)_{n\in\mathbb{N}} \subset M$ gibt mit $x_n \neq x$ für alle $n \in \mathbb{N}$ und $x_n \underset{n\to\infty}{\longrightarrow} x$.*
(2) *$x$ ist Berührungspunkt von $M$ genau dann, wenn es eine Folge $(x_n)_{n\in\mathbb{N}} \subset M$ gibt mit $x_n \underset{n\to\infty}{\longrightarrow} x$.*

Beweis. ad (1): „$\Longrightarrow$“: Da $x$ Häufungspunkt von $M$ ist, gibt es für jedes $n \in \mathbb{N}^*$ ein $x_n$ in $B(x; \frac{1}{n}) \cap M \backslash \{x\}$. $(x_n)_{n\in\mathbb{N}}$ leistet das Gewünschte.
„$\Longleftarrow$“: Sei $U \in \mathcal{U}(x)$. Dann gibt es ein $k \in \mathbb{N}$ mit $x_k \in U$ und $x_k \neq x$. Also ist $x$ Häufungspunkt.
ad (2): Folgt aus (1). □

Satz III.2.14. *Folgende Aussagen sind äquivalent:*

(1) *$M \subset X$ ist abgeschlossen.*
(2) *$M$ enthält alle seine Häufungspunkte.*
(3) *Für jede konvergente Folge $(x_n)_{n\in\mathbb{N}} \subset M$ gilt $\lim\limits_{n\to\infty} x_n \in M$.*

Beweis. (1) $\Longrightarrow$ (2): Aus Definition III.2.10 folgt

$$\begin{aligned} &\overline{M} = M \cup M' \quad \text{mit} \\ &M' = \{x \in X : x \text{ ist Häufungspunkt von } M\}. \end{aligned}$$

Damit folgt die Behauptung aus Satz III.2.12.
(2) $\Longrightarrow$ (3): Folgt aus Satz III.2.13 (1).
(3) $\Longrightarrow$ (1): Folgt aus Satz III.2.13 (1) und Satz III.2.12. □

Satz III.2.15. *$\overline{M}$ ist die kleinste abgeschlossene Obermenge von $M$, d.h.*

$$\overline{M} = \bigcap \{A : A \supset M, A \textit{ ist abgeschlossen}\}.$$

Beweis. Sei

$$B = \bigcap \{A : A \supset M, A \text{ ist abgeschlossen}\}.$$

Gemäß Satz III.2.8 (2) ist $B$ abgeschlossen. Außerdem ist $M \subset B$. Sei $x \in B^c$. Dann gibt es ein $U \in \mathcal{U}(x)$ mit $U \subset B^c$. Also ist

$$U \cap B = \emptyset \Longrightarrow U \cap M = \emptyset.$$

Mithin ist $x \notin \overline{M}$. Dies zeigt

$$B^c \subset (\overline{M})^c \Longrightarrow \overline{M} \subset B.$$

Sei nun umgekehrt $x \in (\overline{M})^c$. Dann gibt es ein $U \in \mathcal{U}(x)$ mit $U \cap M = \emptyset$. $U^c$ ist abgeschlossen, $x \notin U^c$ und $M \subset U^c$. Also ist $x \notin B$. Dies zeigt

$$(\overline{M})^c \subset B^c \Longrightarrow B \subset \overline{M}.$$

$\square$

Satz III.2.16. *Es gilt:*

(1) $A \subset B \Longrightarrow \overline{A} \subset \overline{B}$,
(2) $\overline{(\overline{A})} = \overline{A}$,
(3) $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

Beweis. ad (1): $\overline{B} \supset B$ ist abgeschlossene Obermenge von $A$. Damit folgt die Behauptung aus Satz III.2.15.
ad (2): $\overline{A}$ ist gemäß Satz III.2.15 abgeschlossen. Damit folgt die Behauptung aus Satz III.2.12.
ad (3): $A \cup B \supset A$, $A \cup B \supset B \Longrightarrow$ (mit (1)) $\overline{A \cup B} \supset \overline{A} \cup \overline{B}$.
$\overline{A} \cup \overline{B}$ abgeschlossen, $A \cup B \subset \overline{A} \cup \overline{B} \Longrightarrow \overline{A \cup B} \subset \overline{A} \cup \overline{B}$. $\square$

Definition III.2.17. Wir definieren für $M \subset X$

$$\overset{\circ}{M} = \{x \in X : x \text{ ist innerer Punkt von } M\}.$$

Aus der Definition von „offen“ folgt unmittelbar.

Satz III.2.18. (1) $\overset{\circ}{M} \subset M$.
(2) $\overset{\circ}{M} = M \iff M$ *ist offen.*

In Analogie zu den Sätzen III.2.15 und III.2.16 erhalten wir.

Satz III.2.19. $\overset{\circ}{M}$ *ist die größte offene Teilmenge von* $M$, *d.h.*

$$\overset{\circ}{M} = \bigcup \{O : O \subset M, O \textit{ ist offen}\}.$$

Beweis. Gemäß Satz III.2.4 (2) ist

$$B = \bigcup \{O : O \subset M, O \text{ ist offen}\}$$

eine offene Teilmenge von $M$. Sei $x \in B$. Dann ist $x$ innerer Punkt von $B$ und wegen $B \subset M$ auch innerer Punkt von $M$. Also ist

$$B \subset \overset{\circ}{M}.$$

Sei $x \in \overset{\circ}{M}$. Dann gibt es ein $U \in \mathcal{U}(x)$ mit $U \subset M$. Konstruktionsgemäß ist $U \subset B$. Also ist auch

$$\overset{\circ}{M} \subset B.$$

$\square$

Satz III.2.20. *Es gilt:*

(1) $A \subset B \Longrightarrow \overset{\circ}{A} \subset \overset{\circ}{B}$.

(2) $(\mathring{A})^\circ = \mathring{A}$.
(3) $(A \cap B)^\circ = \mathring{A} \cap \mathring{B}$.

BEWEIS. AD (1): Folgt aus Satz III.2.19.
AD (2): Folgt aus Satz III.2.18 und III.2.19.
AD (3): $A \cap B \subset A \Longrightarrow$ (mit (1)) $(A \cap B)^\circ \subset \mathring{A}$
$A \cap B \subset B \Longrightarrow$ (mit (1)) $(A \cap B)^\circ \subset \mathring{B}$
Satz III.2.4 $\Longrightarrow \mathring{A} \cap \mathring{B}$ offene Teilmenge von $A \cap B$
$\Longrightarrow \mathring{A} \cap \mathring{B} \subset (A \cap B)^\circ$. □

DEFINITION III.2.21. Der RAND $\partial M$ einer Menge $M$ ist definiert durch

$$\partial M = \overline{M} \backslash \mathring{M}.$$

SATZ III.2.22. *Es gilt:*
(1) *$\partial M$ ist abgeschlossen.*
(2) *Es ist $x \in \partial M$ genau dann, wenn es für jedes $U \in \mathcal{U}(x)$ gilt*

$$U \cap M \neq \emptyset \quad \textit{und}\ U \cap M^c \neq \emptyset.$$

BEWEIS. AD (1): $\partial M = \overline{M} \cap (\mathring{M})^c$.
AD (2): $x \notin \mathring{M} \iff \forall U \in \mathcal{U}(x) : U \cap M^c \neq \emptyset$
$x \in \overline{M} \iff \forall U \in \mathcal{U}(x) : U \cap M \neq \emptyset$. □

BEISPIEL III.2.23. (1) $M = B(0;1)$. Dann ist

$$\begin{aligned} M &= \mathring{M} \\ \overline{M} &= \{x \in X : \|x\| \leq 1\}, \\ \partial M &= \{x \in X : \|x\| = 1\}. \end{aligned}$$

(2) $M = B(0;1)\backslash\{0\}$. Dann ist

$$\begin{aligned} M &= \mathring{M} \\ \overline{M} &= \{x \in X : \|x\| \leq 1\}, \\ \partial M &= \{x \in X : \|x\| = 1\} \cup \{0\}. \end{aligned}$$

Das folgende HAUSDORFFSCHE TRENNUNGSAXIOM ist für normierte Vektorräume trivial. Für allgemeine topologische Räume muss es separat als Axiom gefordert werden. Ohne seine Gültigkeit sind viele wesentliche Sätze falsch.

SATZ III.2.24 (HAUSDORFFSCHES TRENNUNGSAXIOM). *Sei $x, y \in X$ mit $x \neq y$. Dann gibt es zwei Umgebungen $U_x \in \mathcal{U}(x)$ und $U_y \in \mathcal{U}(y)$ mit*

$$U_x \cap U_y = \emptyset.$$

BEWEIS. Sei $r = \|x - y\| > 0$. Setze

$$U_x = B(x; \frac{r}{2}) \ , \ U_y = B(y; \frac{r}{2}).$$

Dann gilt für $z \in U_y$

$$\begin{aligned} \|x - z\| &\geq \|x - y\| - \|y - z\| \\ &> \|x - y\| - \frac{r}{2} \\ &= \frac{r}{2}. \end{aligned}$$

Also ist $U_x \cap U_y = \emptyset$. □

Eine einfache Konsequenz aus Satz III.2.24 ist:

SATZ III.2.25. *Sei $x \in X$. Dann ist $\{x\}$ abgeschlossen.*

Zum Abschluss führen wir noch einen, für das Folgende sehr hilfreichen Begriff ein.

DEFINITION III.2.26. Sei $M \subset X$, $M \neq \emptyset$, und $A \subset M$. Dann heißt $A$ RELATIV OFFEN bzw. RELATIV ABGESCHLOSSEN in $M$, wenn es eine offene bzw. abgeschlossene Menge $B$ in $X$ gibt mit $A = B \cap M$.

BEISPIEL III.2.27. Sei $X = \mathbb{R}$ und $M = [0, 1)$. Dann ist $[0, \frac{1}{2})$ relativ offen in $M$ und $[\frac{1}{2}, 1)$ relativ abgeschlossen in $M$.

## III.3. Stetigkeit

Im Folgenden seien $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ und $(Z, \|\cdot\|_Z)$ normierte Vektorräume.

DEFINITION III.3.1. Sei $D \subset X$ nicht leer und $f : D \to Y$ eine Funktion. $f$ heißt STETIG IN $x_0 \in D$, wenn es zu jedem $V \in \mathcal{U}(f(x_0))$ ein $U \in \mathcal{U}(x_0)$ gibt mit $f(U \cap D) \subset V$. $f$ heißt STETIG IN $D' \subset D$, $D' \neq \emptyset$, wenn $f$ stetig ist in jedem Punkt $x \in D'$. Ist insbesondere $f$ stetig in $D$, so sagen wir auch kurz „$f$ IST STETIG“.

Bevor wir Beispiele stetiger und nicht stetiger Abbildungen geben, wollen wir einige andere äquivalente Definitionen der Stetigkeit angeben.

SATZ III.3.2. *Sei $f : D \to Y, D \subset X$, eine Funktion und $x_0 \in D$. Dann sind folgende Aussagen äquivalent:*

(1) *$f$ ist stetig in $x_0$.*
(2) $\forall \varepsilon > 0 \ \exists \, \delta > 0 : x \in D, \|x - x_0\| < \delta \Longrightarrow \|f(x) - f(x_0)\|_Y < \varepsilon$.
(3) $\forall (x_n)_{n\in\mathbb{N}} \subset D$ *mit* $\lim_{n\to\infty} x_n = x_0$ *gilt*

$$\lim_{n\to\infty} f(x_n) = f(x_0).$$

BEWEIS. (1) $\Longrightarrow$ (2): Sei $\varepsilon > 0$ und $V = B(f(x_0); \varepsilon) \in \mathcal{U}(f(x_0))$. Dann gibt es ein $U \in \mathcal{U}(x_0)$ mit $f(U \cap D) \subset V$. Da $U$ offen ist, gibt es ein $\delta > 0$ mit $B(x_0; \delta) \subset U$. Wegen $f(B(x_0; \delta) \cap D) \subset f(U \cap D) \subset V$ folgt die Behauptung.
(2) $\Longrightarrow$ (3): Sei $(x_n)_{n\in\mathbb{N}} \subset D$ mit $\lim\limits_{n\to\infty} x_n = x_0$ und $\varepsilon > 0$. Sei $\delta$ wie in Teil (2). Dann gibt es ein $n_\delta \in \mathbb{N}$ mit $\|x_n - x_0\|_X < \delta$ für alle $n \geq n_\delta$. Dann gilt aber auch

$$\|f(x_n) - f(x_0)\|_Y < \varepsilon \quad \forall n \geq n_\delta.$$

Also ist $f(x_0) = \lim\limits_{n\to\infty} f(x_n)$.
(3) $\Longrightarrow$ (1): Angenommen, $f$ ist nicht stetig in $x_0$. Dann folgt

$$\begin{aligned}
&\exists\, V \in \mathcal{U}(f(x_0)) \quad \forall U \in \mathcal{U}(x_0) : f(U \cap D) \not\subset V \\
\Longrightarrow &\exists\, V \in \mathcal{U}(f(x_0)) \quad \forall n \in \mathbb{N}^* : f(B(x_0; \frac{1}{n}) \cap D) \not\subset V \\
\Longrightarrow &\exists\, V \in \mathcal{U}(f(x_0)) \quad \forall n \in \mathbb{N}^* : \exists\, x_n \in B(x_0; \frac{1}{n}) \cap D : f(x_n) \notin V \\
\Longrightarrow &\exists\, (x_n)_{n\in\mathbb{N}} \subset D, \lim_{n\to\infty} x_n = x_0 \;\text{ und } f(x_n) \not\to f(x_0)
\end{aligned}$$

Widerspruch. □

SATZ III.3.3. *Sei $f : D \to Y$, $D \subset X$, $D \neq \emptyset$, eine Funktion und $D' \subset D$ nicht leer. Dann sind folgende Aussagen äquivalent:*

(1) *$f$ ist stetig in $D'$.*
(2) *Für jede offene Teilmenge $V$ von $Y$ ist $f^{-1}(V)$ relativ offen in $D'$.*
(3) *Für jede abgeschlossene Teilmenge $A$ von $Y$ ist $f^{-1}(A)$ relativ abgeschlossen in $D'$.*

BEWEIS. (1) $\Longrightarrow$ (2): O.E. ist $D' \cap f^{-1}(V) \neq \emptyset$. Sei $x \in D' \cap f^{-1}(V)$. Dann ist $V \in \mathcal{U}(f(x))$. Also gibt es ein $U \in \mathcal{U}(x)$ mit

$$f(D' \cap U) \subset f(D \cap U) \subset V$$

d.h. $D' \cap U \subset f^{-1}(V)$. Also ist $x$ relativ innerer Punkt von $D' \cap f^{-1}(V)$.
(2) $\Longrightarrow$ (3): Sei $A \subset Y$ abgeschlossen. Dann ist $V = Y \backslash A$ offen. Wegen

$$f^{-1}(V) = f^{-1}(Y \backslash A) = D \backslash f^{-1}(A)$$

ist $D \backslash f^{-1}(A)$ relativ offen in $D'$ und damit $f^{-1}(A)$ relativ abgeschlossen in $D'$.
(3) $\Longrightarrow$ (1): Sei $x \in D'$ und $V \subset \mathcal{U}(f(x))$. Dann ist $Y \backslash V$ abgeschlossen und damit nach Voraussetzung

$$f^{-1}(Y \backslash V) = D \backslash f^{-1}(V)$$

relativ abgeschlossen in $D'$. Also ist $f^{-1}(V)$ relativ offen in $D'$, d.h., es gibt ein $U \in \mathcal{U}(x)$ mit

$$f^{-1}(V) \cap D' = U \cap D'.$$

Also ist

$$f(U \cap D') \subset V.$$

Damit ist die Behauptung bewiesen. □

Beispiel III.3.4. (1) $f : X \to \mathbb{R}_+$ mit $f(x) = \|x\|$ ist stetig.
(2) $f : \mathbb{R}_+ \to \mathbb{R}_+$ mit $f(x) = \sqrt{x}$ ist stetig.
(3) $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \lceil x \rceil = \max\{k \in \mathbb{Z} : k \leq x\}$ ist stetig für alle $x \in \mathbb{R} \backslash \mathbb{Z}$.
(4) $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

ist in keinem Punkt stetig.
(5) $f : \mathbb{K}^m \to \mathbb{K}$ mit $f((x_1, \ldots, x_m)) = x_j, 1 \leq j \leq m$, ist stetig. Dies gilt für allgemeine Produkträume $X_1 \times X_2 \times \ldots \times X_m$.
(6) Die Funktionen $z \to \operatorname{Re}(z)$ und $z \to \operatorname{Im}(z)$ von $\mathbb{C}$ nach $\mathbb{R}$ sind stetig.
(7) Sei $Y = X$ und $\|\cdot\|_Y$ zu $\|\cdot\|_X$ äquivalent. Dann ist die identische Abbildung $x \to x$ von $X \to Y$ stetig.
(8) Sei $f : X \to \mathbb{R}$ stetig und $r \in \mathbb{R}$. Dann sind die Mengen

$$\{x \in X : f(x) < r\} \ , \ \{x \in X : f(x) > r\}$$

offen und die Mengen

$$\{x \in X : f(x) \leq r\} \ , \ \{x \in X : f(x) \geq r\}$$

abgeschlossen.

Beweis. ad (1): Folgt aus Bemerkung III.1.5 (S. 62).
ad (2): (a) „$x_0 = 0$“: Sei $\varepsilon < 0$ gegeben, setze $\delta = \delta(\varepsilon) = \varepsilon^2$. Dann gilt

$$0 \leq x < \delta \Longrightarrow |\sqrt{x_0} - \sqrt{x}| = \sqrt{x} < \sqrt{\delta} = \varepsilon,$$

(b) „$x_0 > 0$“: Sei $\varepsilon > 0$ gegeben, setze $\delta = \delta(\varepsilon, x_0) = \varepsilon\sqrt{x_0}$. Dann gilt

$$|x - x_0| < \delta \Longrightarrow |\sqrt{x_0} - \sqrt{x}| = \frac{|x_0 - x|}{\sqrt{x_0} + \sqrt{x}} \leq \frac{|x_0 - x|}{\sqrt{x_0}} < \varepsilon.$$

ad (3):

$$\begin{aligned} x \notin \mathbb{Z} \Longrightarrow & \exists\, \delta > 0 : (x - \delta, x + \delta) \subset \mathbb{R} \backslash \mathbb{Z} \\ \Longrightarrow & f|_{(x-\delta, x+\delta)} \text{ ist konstant,} \\ x \in \mathbb{Z} \Longrightarrow & f(y) \leq x - 1 \quad \forall y < x \\ & f(y) \geq x \qquad\quad \forall y \geq x. \end{aligned}$$

ad (4): Folgt aus Satz I.4.11 (S. 21).
ad (5): Folgt aus Teil (7) und

$$|x_j - y_j| \leq \{\sum_{i=1}^{n} |x_i - y_i|^2\}^{1/2} = \|x - y\|_2.$$

AD (6): Folgt aus Satz I.5.5 (S. 26).
AD (7): Folgt aus der Definition III.1.9 (S. 64)

$$\|x-y\|_X \le \bar{c}\|x-y\|_Y \quad \forall x,y \in X.$$

AD (8): Folgt aus Satz III.3.3. □

SATZ III.3.5. *Seien $D_f \subset X$, $D_g \subset X$, $x_0 \in D_f \cap D_g$, $f : D_f \to Y$ und $g : D_g \to Y$ in $x_0$ stetig, sowie $\alpha \in \mathbb{K}$. Dann gilt:*

(1) $\alpha f : x \mapsto \alpha f(x)$ *ist stetig in* $x_0$.
(2) $f+g : x \mapsto f(x)+g(x)$ *ist stetig in* $x_0$.
(3) $Y = \mathbb{K}$; $f \cdot g : x \mapsto f(x) \cdot g(x)$ *ist stetig in* $x_0$.
(4) $Y = \mathbb{K}$ *und* $g(x_0) \ne 0$; $\frac{f}{g} : x \mapsto \frac{f(x)}{g(x)}$ *ist stetig in* $x_0$.

BEWEIS. Die Behauptung folgt aus Satz III.3.2 und Satz II.1.18 (S. 32), sowie Bemerkung III.1.8 (S. 63). □

Wegen Satz III.3.5 ist die folgende Bezeichnung sinnvoll.

DEFINITION III.3.6. Sei $M \subset X$, $M \ne \emptyset$. Dann bezeichnet $C(M,Y)$ den Vektorraum aller stetigen Funktionen von $M$ in $Y$. Ist insbesondere $Y = \mathbb{K}$, so schreiben wir kurz $C(M)$ statt $C(M,\mathbb{K})$.

Aus Satz III.3.5 folgt unmittelbar:

SATZ III.3.7. *Alle* POLYNOME, *d.h. Funktionen der Form*

$$f(x) = \sum_{k=0}^{n} a_k x^k, \quad a_k \in \mathbb{K},\ n \in \mathbb{N},$$

*und alle* RATIONALEN FUNKTIONEN, *d.h. Funktionen der Form*

$$f(x) = \frac{p(x)}{q(x)}, \quad p,q \text{ Polynome},$$

*sind auf ihrem Definitionsbereich stetig.*

SATZ III.3.8. *Seien $f : D_f \to Y$, $D_f \subset X$, und $g : D_g \to Z$, $D_g \subset Y$, zwei Funktionen mit $f(D_f) \subset D_g$. Dann ist die* KOMPOSITION *oder* VERKNÜPFUNG *$g \circ f : D_f \to Z$ definiert durch*

$$g \circ f : x \mapsto g(f(x)).$$

*Ist $f$ in $x_0 \in D_f$ stetig und $g$ in $y_0 = f(x_0) \in D_g$ stetig, so ist $g \circ f$ in $x_0$ stetig.*

BEWEIS. Sei $W \in \mathcal{U}(g(y_0))$. Dann folgt

$$g \text{ stetig in } y_0 \Longrightarrow \exists\, V \in \mathcal{U}(y_0) : g(D_g \cap V) \subset W$$
$$f \text{ stetig in } x_0 \Longrightarrow \exists\, U \in \mathcal{U}(x_0) : f(D_f \cap U) \subset V.$$

Wegen

$$g \circ f(D_f \cap U) \subset g(D_g \cap V) \subset W$$

ist die Behauptung damit bewiesen. □

Beispiel III.3.9. (1) Sei $f : D \to Y$, $D \subset X$, stetig. Dann ist $\|f\|_Y : D \to \mathbb{R}_+$ mit $x \mapsto \|f(x)\|_Y$ stetig.
(2) Sei $f : D \to \mathbb{K}$, $D \subset X$, stetig. Dann ist $f^2 : D \to \mathbb{K}$ mit $x \mapsto f(x)^2$ stetig.
(3) Sei $f : D \to \mathbb{R}_+$, $D \subset X$, stetig. Dann ist $\sqrt{f} : D \to \mathbb{R}_+$ mit $x \mapsto \sqrt{f(x)}$ stetig.

Satz III.3.10. *(1) $f = (f_1 \ldots, f_n) : D \to \mathbb{K}^n$, $D \subset X$, ist genau dann stetig in $x_0 \in D$, wenn alle Komponentenfunktionen $f_i$, $1 \le i \le n$, in $x_0$ stetig sind.*
*(2) $f : D \to \mathbb{C}$, $D \subset X$, ist genau dann in $x_0 \in D$ stetig, wenn $\operatorname{Re} f$ und $\operatorname{Im} f$ in $x_0$ stetig sind.*

Beweis. ad (1): „$\Longrightarrow$“ folgt aus Satz III.3.8 und Beispiel III.3.4 (5).
„$\Longleftarrow$“ Sei $\varepsilon > 0$. Dann gibt es nach Voraussetzung Zahlen $\delta_j(\varepsilon) > 0, 1 \le j \le n$, mit

$$|f_j(x) - f_j(x_0)| < \frac{\varepsilon}{n} \quad \forall \|x - x_0\|_X < \delta_j.$$

Setze $\delta(\varepsilon) = \min\limits_{1 \le j \le n} \delta_j(\varepsilon)$. Dann folgt für $x \in B(x_0; \delta)$

$$\|f(x) - f(x_0)\|_{\mathbb{K}^n} = \{\sum_{j=1}^{n} |f(x_j) - f(x_0)|^2\}^{1/2} < \varepsilon.$$

ad (2): Folgt wegen $\mathbb{C} \cong \mathbb{R}^2$ mit $f_1 = \operatorname{Re} f$, $f_2 = \operatorname{Im} f$ aus Teil (1). □

Definition III.3.11. Sei $f : D \to Y$, $D \subset X$, eine Funktion. Zu $x_0 \in X$ gebe es eine Folge $(x_n)_{n\in\mathbb{N}} \subset D$ mit $\lim\limits_{n\to\infty} x_n = x_0$. Dann definieren wir

$$y = \lim_{x \to x_0} f(x),$$

wenn für *jede* Folge $(z_n)_{n\in\mathbb{N}} \subset D$ mit $z_n \underset{n\to\infty}{\longrightarrow} x_0$ gilt $f(z_n) \underset{n\to\infty}{\longrightarrow} y$.

Satz III.3.12. *Folgende Aussagen sind äquivalent:*
(1) $y = \lim\limits_{x \to x_0} f(x)$.
(2) $\forall V \in \mathcal{U}(y)\ \exists\, U \in \mathcal{U}(x_0) : f(D \cap U) \subset V$.

Beweis. (1) $\Longrightarrow$ (2): Angenommen, es existiert ein $V \in \mathcal{U}(y)$ mit $f(D \cap U) \not\subset V$ für alle $U \in \mathcal{U}(x_0)$. Dann gilt

$$\begin{aligned} &\forall n \in \mathbb{N}^* \quad f(D \cap B(x_0; \frac{1}{n})) \not\subset V \\ \Longrightarrow &\forall n \in \mathbb{N}^*\ \exists\, x_n \in D \cap B(x_0; \frac{1}{n}) : f(x_n) \notin V \\ \Longrightarrow &\exists (x_n)_{n\in\mathbb{N}} \subset D \ \text{ mit } x_n \to x_0 \ \text{ und } f(x_n) \not\to y_0. \end{aligned}$$

Widerspruch.
(2) $\Longrightarrow$ (1): Sei $(x_n)_{n\in\mathbb{N}} \subset D$ mit $\lim\limits_{n\to\infty} x_n = x_0$ und $\varepsilon > 0$. Dann

gibt es ein $U \in \mathcal{U}(x_0)$ mit $f(D \cap U) \subset B(y_0; \varepsilon)$. Zu $U$ gibt es aber ein $n_U \in \mathbb{N}$ mit $x_n \in U$ für alle $n \geq n_U$. Also gilt

$$f(x_n) \in B(y_0; \varepsilon) \quad \forall n \geq n_U.$$

Mithin konvergiert $f(x_n)$ gegen $y_0$. $\square$

Bemerkung III.3.13. (1) Sei $f : D \to Y$, $D \subset X$, und $x_0 \in D$. Dann gilt $\lim\limits_{x \to x_0} f(x) = f(x_0)$ genau dann, wenn $f$ in $x_0$ stetig ist.
(2) Sei $f : D \to Y$, $D \subset X$, $x_0 \notin D$ und es existiere $y = \lim\limits_{x \to x_0} f(x)$. Dann können wir eine Erweiterung $\tilde{f} : D \cup \{x_0\} \to Y$ von $f$ durch

$$\tilde{f}(x) = \begin{cases} f(x) & \text{falls } x \neq x_0 \\ y & \text{falls } x = x_0 \end{cases}$$

definieren. $\tilde{f}$ ist dann in $x_0$ stetig und heißt daher stetige Ergänzung von $f$ in $x_0$.

Beispiel III.3.14. (1) $X = Y = \mathbb{R}$, $D = X \backslash \{1\}$ und

$$f(x) = \frac{x^n - 1}{x - 1}, \quad n \in \mathbb{N}^*.$$

Aus Satz I.1.16 (S. 12) folgt

$$f(x) = \sum_{k=0}^{n-1} x^k \quad \forall x \in X \backslash \{1\}$$

und damit

$$\lim_{x \to 1} f(x) = n.$$

$\tilde{f} : \mathbb{R} \to \mathbb{R}$ mit

$$\tilde{f}(x) = \begin{cases} \frac{x^n-1}{x-1} & \text{falls } x \neq 1 \\ n & \text{falls } x = 1 \end{cases}$$

ist die stetige Ergänzung von $f$.
(2) $X = Y = \mathbb{C}$, $D = \mathbb{C} \backslash \{0\}$ und

$$f(z) = \frac{e^z - 1}{z}.$$

Da für alle $z \in \mathbb{C}$ mit $|z| < 1$ gilt

$$\begin{aligned} |f(z) - 1| &= |\sum_{n=1}^{\infty} \frac{z^{n-1}}{n!} - 1| \\ &= |\sum_{m=1}^{\infty} \frac{1}{(m+1)!} z^m| \\ &\leq \sum_{m=1}^{\infty} \frac{1}{(m+1)!} |z|^m \end{aligned}$$

$$\leq \sum_{m=1}^{\infty} \frac{1}{2}|z|^m$$
$$\leq \frac{1}{2}|z|\frac{1}{1-|z|},$$

ist

$$\lim_{z\to 0} f(z) = 1.$$

Die stetige Ergänzung von $f$ ist also

$$\tilde{f}(z) = \begin{cases} \frac{e^z-1}{z} & z \neq 0 \\ 1 & z = 0. \end{cases}$$

Wir kommen nun zu einer Besonderheit stetiger Funktionen auf $\mathbb{R}$, die mit der Anordnung von $\mathbb{R}$ zusammenhängt.

Definition III.3.15. Die Funktion $f : D \to Y$, $D \subset \mathbb{R}$, heißt linksseitig stetig bzw. rechtsseitig stetig in $x_0 \in D$ genau dann, wenn es zu jedem $V \in \mathcal{U}(f(x_0))$ ein $\delta > 0$ gibt mit

$$f(D \cap (x_0 - \delta, x_0]) \subset V \quad \text{bzw.} \quad f(D \cap [x_0, x_0 + \delta)) \subset V.$$

Beispiel III.3.16. Die Funktion $\lceil x \rceil$ aus Beispiel III.3.4 (3) ist rechtsseitig stetig in jedem $x \in \mathbb{R}$.

Satz III.3.17. *Sei $f : D \to Y$, $D \subset \mathbb{R}$, eine Funktion und $x_0 \in D$. Dann sind folgende Aussagen äquivalent,*

(1) *$f$ ist linksseitig bzw. rechtsseitig stetig in $x_0$.*
(2) $\forall \varepsilon > 0 \, \exists \, \delta > 0 \, \forall x_0 - \delta < x \leq x_0 : \|f(x) - f(x_0)\|_Y < \varepsilon$ *bzw.*
$\forall \varepsilon > 0 \, \exists \, \delta > 0 \, \forall x_0 \leq x < x_0 + \delta : \|f(x) - f(x_0)\|_Y < \varepsilon.$
(3) *$\forall (x_n)_{n\in\mathbb{N}} \subset D$ mit $x_n \leq x_0$ für alle $n \in \mathbb{N}$ bzw. $x_n \geq x_0$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} x_n = x_0$ gilt*
$$\lim_{n\to\infty} f(x_n) = f(x_0).$$

Beweis. Übungsaufgabe (analog zum Beweis von Satz III.3.2). □

Satz III.3.18. *Die Funktion $f : D \to Y$, $D \subset \mathbb{R}$, ist genau dann in $x_0 \in D$ stetig, wenn sie in $x_0$ linksseitig und rechtsseitig stetig ist.*

Beweis. „$\Longrightarrow$“: Folgt aus Satz III.3.2 und III.3.17.
„$\Longleftarrow$“: Sei $V \in \mathcal{U}(f(x_0))$. Dann gibt es nach Voraussetzung ein $\delta_1 > 0$ mit

$$f(D \cap (x_0 - \delta_1, x_0]) \subset V$$

und ein $\delta_2 > 0$ mit

$$f(D \cap [x_0, x_0 + \delta_2)) \subset V.$$

Für $\delta = \min\{\delta_1, \delta_2\}$ folgt

$$f(D \cap B(x_0; \delta)) = f(D \cap (x_0 - \delta, x + \delta)) \subset V.$$

Damit ist die Behauptung bewiesen. $\square$

Bei der Definition der Stetigkeit hängt das $\delta$ neben $\varepsilon$ auch von dem Punkt $x$, der gerade betrachtet wird, ab. Für viele praktische und theoretische Ergebnisse ist der Fall von Bedeutung, dass $\delta$ unabhängig von $x$ gewählt werden kann. Dies führt zu folgender Definition.

Definition III.3.19. Sei $f : D \to Y$, $D \subset X$, eine Funktion und $D' \subset D$ nicht leer. Dann heißt $f$ auf $D'$ gleichmässig stetig, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt mit

$$\|f(x) - f(y)\|_Y < \varepsilon \quad \forall x, y \in D' \text{ mit } \|x - y\|_X < \delta.$$

Bemerkung III.3.20. Eine gleichmäßig stetige Funktion ist auch stetig.

Beispiel III.3.21. (1) $f : X \to \mathbb{R}_+$ mit $f(x) = \|x\|$ ist gleichmäßig stetig.
(2) $f : \mathbb{R}^* \to \mathbb{R}^*$ mit $f(x) = \frac{1}{x}$ ist nicht gleichmäßig stetig.
(3) $f : [-1, 1] \to \mathbb{R}$ mit $f(x) = x^2$ ist gleichmäßig stetig.
(4) $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$ ist nicht gleichmäßig stetig.

Beweis. ad (1): Folgt aus Bemerkung III.1.5 (S. 62).
ad (2): Wegen Satz III.3.5 ist $f$ stetig. Sei $\delta > 0$ und $n \in \mathbb{N}^*$ mit $n > \frac{1}{2\delta}$. Dann folgt

$$|\frac{1}{n} - \frac{1}{2n}| = \frac{1}{2n} < \delta$$

und

$$|f(\frac{1}{n}) - f(\frac{1}{2n})| = n \geq 1.$$

Also kann $f$ nicht gleichmäßig stetig sein.
ad (3): Für $x, y \in [-1, 1]$ gilt

$$|x^2 - y^2| = |x - y||x + y| \leq 2|x - y|.$$

Also ist $f$ gleichmäßig stetig auf $[-1, 1]$.
ad (4): Sei $\delta > 0$ und $n \in \mathbb{N}^*$ mit $n > \frac{1}{\delta}$. Dann folgt für $x \geq n$ und $y = x + \frac{\delta}{2}$ :

$$|x - y| < \delta$$

und

$$|f(x) - f(y)| = \delta x + \frac{\delta^2}{4} > \delta n > 1.$$

Also ist $f$ nicht gleichmäßig stetig auf $\mathbb{R}$. $\square$

## III.4. Kompaktheit

Im Folgenden seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ stets normierte Vektorräume.

Definition III.4.1. Sei $M \subset X$.

(1) Eine Menge $\{O_\alpha : \alpha \in A\} \subset \mathcal{P}(x)$ heißt offene Überdeckung von $M$, wenn gilt $M \subset \bigcup_{\alpha \in A} O_\alpha$ und alle $O_\alpha$ sind offen.
(2) $M$ heißt kompakt, wenn *jede* offene Überdeckung $\{O_\alpha : \alpha \in A\}$ von $M$ eine endliche Teilüberdeckung enthält, d.h., es gibt ein $n \in \mathbb{N}$ und $\alpha_0, \ldots, \alpha_n \in A$ mit $M \subset \bigcup_{0 \le i \le n} O_{\alpha_i}$.

Beispiel III.4.2. (1) $\emptyset$ ist kompakt.
(2) $\{0\} \cup \{\frac{1}{n} : n \in \mathbb{N}^*\} \subset \mathbb{R}$ ist kompakt.
(3) $\{\frac{1}{n} : n \in \mathbb{N}^*\} \subset \mathbb{R}$ ist nicht kompakt.

Beweis. ad (1): Ist offensichtlich.
ad (2): Setze $M = \{0\} \cup \{\frac{1}{n} : n \in \mathbb{N}^*\}$. Sei $\{O_\alpha : \alpha \in A\}$ eine offene Überdeckung von $M$. Dann gibt es ein $\alpha_0 \in A$ mit $0 \in O_{\alpha_0}$. Wegen $\frac{1}{n} \underset{n\to\infty}{\longrightarrow} 0$ gibt es ein $n_0 \in \mathbb{N}$ mit $\frac{1}{n} \in O_{\alpha_0}$ für alle $n \ge n_0$. Schließlich gibt es Zahlen $\alpha_1, \ldots, \alpha_{n_0} \in A$ mit $\frac{1}{k} \in O_{\alpha_k}, 1 \le k \le n_0$. Dann ist $\{O_{\alpha_k} : 0 \le k \le n_0\}$ eine endliche Teilüberdeckung von $M$.
(3) $\{B(\frac{1}{n}; \frac{1}{2n^2}) : n \in \mathbb{N}^*\}$ ist eine offene Überdeckung von $\{\frac{1}{n} : n \in \mathbb{N}^*\}$ durch paarweise disjunkte Kugeln und enthält daher keine endliche Teilüberdeckung. $\square$

Satz III.4.3. *$M \subset X$ sei kompakt. Dann gilt:*

(1) *$M$ ist beschränkt, d.h., es gibt ein $R > 0$ mit $M \subset B(0; R)$.*
(2) *$M$ ist abgeschlossen.*

Beweis. ad (1): $\{B(0; r) : r \in \mathbb{R}_+^*\}$ ist eine offene Überdeckung von $M$. Damit folgt die Behauptung unmittelbar aus der Definition der Kompaktheit.
ad (2): Sei $x \in M^c$. Dann ist $\{B(y; \frac{1}{2}\|y - x\|_X) : y \in M\}$ eine offene Überdeckung von $M$. Sei $\{B(y_i; \frac{1}{2}\|y_i - x\|_X) : 0 \le i \le n\}$ eine endliche Teilüberdeckung. Definiere

$$r = \min_{0 \le i \le n} \frac{1}{2}\|y_i - x\|_X.$$

Dann ist $r > 0$ und

$$B(x; r) \cap B(y_i; r) = \emptyset \quad \forall 0 \le i \le n.$$

Also ist $B(x; r) \subset M^c$ und somit $x$ innerer Punkt von $M^c$. $\square$

Satz III.4.4. *Folgende Aussagen sind äquivalent:*

(1) *$M \subset X$ ist kompakt.*

(2) *Jede Folge* $(x_n)_{n\in\mathbb{N}} \subset M$ *in* $M$ *besitzt einen Häufungspunkt in* $M$*.*

Beweis. (1) $\Longrightarrow$ (2): Angenommen (2) wäre falsch. Dann gibt es eine Folge $(x_n)_{n\in\mathbb{N}} \subset M$, die keinen HP in $M$ hat. D.h., zu jedem $x \in M$ gibt es ein $\varepsilon(x) > 0$, so dass $B(x;\varepsilon(x))$ nur endlich viele Folgenglieder enthält. $\{B(x;\varepsilon(x)) : x \in M\}$ ist eine offene Überdeckung von $M$. Sei $\{B(y_i;\varepsilon(y_i)) : 0 \le i \le n\}$ eine endliche Teilüberdeckung. Da jedes $B(y_i;\varepsilon(y_i))$ nur endlich viele Folgenglieder enthält, enthält $\bigcup_{0\le i\le n} B(y_i;\varepsilon(y_i))$ auch nur endlich viele Folgenglieder. Widerspruch.
(2) $\Longrightarrow$ (1): Wir zeigen zunächst, dass für jedes $r \in \mathbb{R}_+^*$ die offene Überdeckung

$$B_r = \{B(x;r) : x \in M\}$$

von $M$ eine endliche Teilüberdeckung enthält.
Angenommen dies sei nicht der Fall. Dann gibt es ein $R > 0$, so dass jede endliche Teilmenge von $B_R$ die Menge $M$ nicht überdeckt. Sei $x_0 \in M$ beliebig. Dann gibt es ein $x_1 \in M\backslash B(x_0;R)$. Weiter gibt es ein $x_2 \in M\backslash[B(x_0;R) \cup B(x_1;R)]$. Durch Induktion zeigen wir, dass es eine Folge $(x_n)_{n\in\mathbb{N}} \subset M$ gibt mit

$$x_{n+1} \in M\backslash \bigcup_{0\le i\le n} B(x_i;R).$$

Gemäß (2) besitzt $(x_n)_{n\in\mathbb{N}}$ einen HP $x^* \in M$. Dann enthält $B(x^*;\frac{R}{2})$ unendlich viele Folgenglieder $(x_{n_i})_{i\in\mathbb{N}}$. Für je zwei solche Folgenglieder $x_{n_1}$ und $x_{n_2}$ gilt

$$\|x_{n_1} - x_{n_2}\|_X \le \|x_{n_1} - x^*\|_X + \|x_{n_2} - x^*\|_X < R$$

im Widerspruch zur Konstruktion.
Sei nun $\{O_\alpha : \alpha \in A\}$ eine beliebige offene Überdeckung von $M$. Wir nehmen an, dass sie keine endliche Teilüberdeckung enthält. Wie wir oben gezeigt haben, enthält $B_{\frac{1}{n}}$ für jedes $n \in \mathbb{N}^*$ eine endliche Teilüberdeckung. Da $\{O_\alpha, \alpha \in A\}$ keine endliche Teilüberdeckung enthält, gibt es somit für jedes $n \in \mathbb{N}^*$ ein $x_n \in M$, so dass $B(x_n;\frac{1}{n}) \cap M$ nicht durch endlich viele $O_\alpha$'s überdeckt werden kann. Die Folge $(x_n)_{n\in\mathbb{N}}$ besitzt gemäß (2) einen HP $x^*$ in $M$. Dann gibt es ein $\alpha^* \in A$ und ein $\varepsilon^* \in \mathbb{R}_+^*$ mit

$$x^* \in O_{\alpha^*} \text{ und } B(x^*;\varepsilon^*) \subset O_{\alpha^*}.$$

Da $x^*$ HP von $(x_n)_{n\in\mathbb{N}}$ ist, gibt es zu $N > (\frac{\varepsilon^*}{2})^{-1}$ ein $m \ge N$ mit

$$\|x_m - x^*\|_X < \frac{\varepsilon^*}{2}.$$

Für $x \in B(x_m;\frac{1}{m})$ folgt

$$\begin{aligned}\|x - x^*\|_X &\le \|x - x_m\|_X + \|x_m - x^*\| \\ &< \frac{1}{m} + \frac{\varepsilon^*}{2}\end{aligned}$$

$$\leq \frac{1}{N} + \frac{\varepsilon^*}{2}$$
$$< \varepsilon^*.$$

Also ist

$$B(x_m; \frac{1}{m}) \subset B(x^*; \varepsilon^*) \subset O_{\alpha^*}.$$

Widerspruch. □

Bemerkung III.4.5. Eine Menge $M$ mit der Eigenschaft (2) aus Satz III.4.4 nennt man auch folgenkompakt. Satz III.4.4 sagt also, dass in einem normierten Vektorraum eine Menge genau dann kompakt ist, wenn sie folgenkompakt ist. Diese Aussage ist in allgemeinen topologischen Räumen i. a. falsch; dort kann man Beispiele folgenkompakter, nicht kompakter Mengen konstruieren.

Wir kommen nun zu einer wesentlichen topologischen Charakterisierung des $\mathbb{K}^m$. Sie geht auf E. Heine (1821-1881) und E. Borel (1871-1956) zurück.

Satz III.4.6 (Satz von Heine-Borel). *Eine Teilmenge $M$ des $\mathbb{K}^m$ ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.*

Beweis. „$\Longrightarrow$“: Satz III.4.3.
„$\Longleftarrow$“: Sei $M \subset \mathbb{K}^m$ beschränkt und abgeschlossen und $(x_n)_{n\in\mathbb{N}}$ eine Folge in $M$. Gemäß Satz II.1.15 (S. 31) und Satz II.2.4 (S. 39) gibt es ein $x^* \in \mathbb{K}^m$ und eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ mit

$$x_{n_k,i} \underset{k\to\infty}{\longrightarrow} x_i^* \quad \forall 1 \leq i \leq m,$$

d.h., $x_{n_k} \underset{k\to\infty}{\longrightarrow} x^*$ in $\mathbb{K}^m$. Da $M$ abgeschlossen ist, gilt $x^* \in M$. Damit folgt die Behauptung aus Satz III.4.4. □

Bemerkung III.4.7. Der Satz von Heine-Borel gilt für unendlich dimensionale Vektorräume nicht. So sind z.B. die abgeschlossenen Einheitskugeln $B(0; 1)$ in $\ell_1$, $\ell_2$ und $\ell_\infty$ nicht kompakt. Denn für die Folge $(e_n)_{n\in\mathbb{N}^*}$ der „Einheitsvektoren“ mit

$$e_{n,k} = \delta_{k,n} = \begin{cases} 1 & \text{für } k = n, \\ 0 & \text{für } k \neq n, \end{cases}$$

gilt

$$\|e_n\|_1 = \|e_n\|_2 = \|e_n\|_\infty = 1$$

und

$$\|e_n - e_m\|_1 = 2$$
$$\|e_n - e_m\|_2 = \sqrt{2}$$
$$\|e_n - e_m\|_\infty = 1$$

für alle $n \neq m$, so dass sie keinen HP besitzen kann.

Als nächstes befassen wir uns mit dem Zusammenhang zwischen Kompaktheit und Stetigkeit.

Satz III.4.8. *Sei $M \subset X$, $M \neq \emptyset$, kompakt und $f \in C(M, Y)$. Dann ist $f$ gleichmäßig stetig.*

Beweis. Sei $\varepsilon > 0$ beliebig. Da $f$ stetig ist, gibt es zu jedem $x \in M$ ein $\delta(x) > 0$ mit

$$f(M \cap B(x; \delta(x))) \subset B(f(x); \frac{\varepsilon}{2}).$$

$\{B(x; \frac{1}{2}\delta(x)) : x \in M\}$ ist eine offene Überdeckung von $M$. Sei $\{B(x_i; \frac{1}{2}\delta(x_i)) : 0 \leq i \leq n\}$ eine endliche Teilüberdeckung. Dann ist

$$\delta = \min_{0 \leq i \leq n} \frac{1}{2}\delta(x_i) > 0.$$

Seien nun $x, y \in M$ mit $\|x - y\|_X < \delta$ beliebig. Dann gibt es ein $x_i$, $0 \leq i \leq n$, mit $x \in B(x_i, \frac{1}{2}\delta(x_i))$. Es folgt

$$\begin{aligned} \|y - x_i\|_X &\leq \|y - x\|_X + \|x_i - x\|_X \\ &< \delta + \frac{1}{2}\delta(x_i) \\ &\leq \delta(x_i) \end{aligned}$$

und daher

$$\begin{aligned} \|f(x) - f(y)\|_Y &\leq \|f(x) - f(x_i)\|_Y + \|f(x_i) - f(y)\|_Y \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig war, ist hiermit die gleichmäßige Stetigkeit von $f$ gezeigt. □

Satz III.4.9. *Sei $M \subset X$, $M \neq \emptyset$, kompakt und $f \in C(M, Y)$. Dann ist $f(M)$ kompakt.*

Beweis. Sei $\{V_\alpha : \alpha \in A\}$ eine offene Überdeckung von $f(M)$. Da $V_\alpha$ offen und $f$ stetig ist, gibt es zu jedem $V_\alpha$ eine offene Menge $U_\alpha$ mit

$$f^{-1}(V_\alpha) = U_\alpha \cap M.$$

Da $\{V_\alpha : \alpha \in A\}$ die Bildmenge $f(M)$ überdeckt, ist $\{U_\alpha : \alpha \in A\}$ eine offene Überdeckung der Urbildmenge $M$. Also gibt es eine endliche Teilüberdeckung $\{U_{\alpha_i} : 0 \leq i \leq n\}$ von $M$. Dann ist aber $\{V_{\alpha_i} : 0 \leq i \leq n\}$ eine endliche Teilüberdeckung von $f(M)$. □

Eine einfache Konsequenz der Sätze III.4.3 und III.4.9 ist der folgende Satz:

Satz III.4.10. *Sei $M \subset X$, $M \neq \emptyset$, kompakt und $f \in C(M, \mathbb{R})$. Dann nimmt $f$ auf $M$ sein Minimum und Maximum an, d.h., es gibt ein $\underline{x} \in M$ und ein $\overline{x} \in M$ mit*

$$f(\underline{x}) \leq f(x) \leq f(\overline{x}) \quad \forall x \in M.$$

Beweis. $f(M) \subset \mathbb{R}$ ist kompakt und damit beschränkt und abgeschlossen. Also gilt

$$-\infty < \rho = \inf f(M) \leq \sup f(M) = R < +\infty.$$

Aus der Definition des Supremums folgt, dass es eine Folge $(x_n)_{n\in\mathbb{N}} \subset M$ gibt mit

$$R = \lim_{n\to\infty} f(x_n).$$

Da $M$ kompakt ist, besitzt $(x_n)_{n\in\mathbb{N}}$ einen HP $\overline{x}$ in $M$. Da $f$ stetig ist, gilt

$$f(\overline{x}) = R.$$

Analog folgt die Existenz eines $\underline{x} \in M$ mit

$$f(\underline{x}) = \rho.$$

□

Bemerkung III.4.11. Sei $M \subset X$, $M \neq \emptyset$, kompakt und $f \in C(M, Y)$. Wegen Beispiel III.3.9(1) (S. 78) und Satz III.4.10 ist

$$\|f\|_\infty = \|f\|_{C(M,Y)} = \max_{x\in M} \|f(x)\|_Y$$

wohldefiniert. Wie man leicht nachrechnet, ist $\|\cdot\|_{C(M,Y)}$ eine Norm auf $C(M, Y)$.

Eine weitere wichtige Konsequenz aus Satz III.4.10 ist:

Satz III.4.12 (Fundamentalsatz der Algebra). *Jedes nicht konstante Polynom besitzt eine Nullstelle in $\mathbb{C}$.*

Beweis. Sei

$$p(z) = \sum_{k=0}^{n} a_k z^k, \quad a_k \in \mathbb{C},$$

ein nicht konstantes Polynom. O.E. können wir $a_n = 1$ und $n \geq 2$ annehmen. Sei

$$R = 1 + \sum_{k=0}^{n} |a_k|.$$

Für $|z| \geq R$ folgt

$$\begin{aligned} |p(z)| &\geq |z|^n - \sum_{k=0}^{n-1} |a_k||z|^k \\ &\geq |z|^n - \sum_{k=0}^{n-1} |a_k||z|^{n-1} \end{aligned}$$

$$\begin{aligned}
&= |z|^{n-1}\{|z| - (R-1)\} \\
&\geq |z|^{n-1} \\
&\geq |R|^{n-1} \\
&\geq R.
\end{aligned}$$

Weiter ist

$$|p(0)| = |a_0| < R.$$

Also gilt wegen Satz III.4.10

$$\begin{aligned}
\inf_{z\in\mathbb{C}} |p(z)| &= \inf_{z\in\overline{B(0;R)}} |p(z)| \\
&= \min_{z\in\overline{B(0;R)}} |p(z)| = |p(z_0)|
\end{aligned}$$

mit $|z_0| \leq R$. Wir nehmen an, dass $p(z_0) \neq 0$ ist. Dann ist

$$q(z) = \frac{1}{p(z_0)} p(z_0 + z)$$

ein Polynom vom Grade $n$ mit

$$(*) \qquad |q(z)| \geq 1 \quad \forall z \in \mathbb{C}$$

und

$$q(0) = 1.$$

Also können wir $q$ in der Form

$$q(z) = 1 + bz^k + z^{k+1}\rho(z)$$

mit $k \geq 1$, $b \in \mathbb{C}^*$ und einem Polynom $\rho$ schreiben. Wie wir in Paragraph III.7 genauer sehen werden, können wir $-\frac{1}{b}$ in der Form

$$-\frac{1}{b} = r\cos\omega + ir\sin\omega$$

mit $r \in \mathbb{R}_+^*$ und $\omega \in [0, 2\pi)$ darstellen, und für

$$z^* = \sqrt[k]{r}\cos\frac{\omega}{k} + i\sqrt[k]{r}\sin\frac{\omega}{k}$$

gilt dann

$$(z^*)^k = -\frac{1}{b}.$$

Die Funktion

$$t \mapsto |(z^*)^{k+1}\rho(tz^*)|$$

nimmt gemäß Satz III.4.10 auf dem Intervall $[0, 1]$ ihr Maximum an. Sei $m$ dieses Maximum. O.E. ist $m \geq 1$. Für $t \in (0, \frac{1}{2m})$ folgt dann

$$\begin{aligned}
|q(tz^*)| &\leq |1 + b(tz^*)^k| + |(tz^*)^{k+1}\rho(tz^*)| \\
&\leq |1 - t^k| + t^{k+1}m \\
&= 1 - t^k(1 - tm)
\end{aligned}$$

$$\leq 1 - \frac{1}{2}t^k$$
$$< 1.$$

Dies ist ein Widerspruch zu $(*)$. □

## III.5. Zusammenhang

Im Folgenden seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Vektorräume.

Definition III.5.1. Eine Teilmenge $M$ von $X$ heißt zusammenhängend, wenn es keine zwei in $M$ relativ offenen Mengen $O_1, O_2$ gibt mit

$$O_1 \neq \emptyset, O_2 \neq \emptyset, O_1 \cap O_2 = \emptyset, O_1 \cup O_2 = M.$$

Satz III.5.2. *Folgende Aussagen sind äquivalent:*

(1) *$M \subset X$ ist zusammenhängend.*
(2) *Für jede Teilmenge $O$ von $M$, die gleichzeitig relativ offen und relativ abgeschlossen in $M$ ist, gilt $O = \emptyset$ oder $O = M$.*

Beweis. (1) $\Longrightarrow$ (2): Sei $O \subset M$ relativ offen und relativ abgeschlossen in $M$. Setze $O' = M \setminus O$. Dann sind $O$ und $O'$ relativ offen in $M$ und erfüllen

$$O \cap O' = \emptyset, O \cup O' = M.$$

Da $M$ zusammenhängend ist, gilt $O = \emptyset$ oder $O' = \emptyset$, d.h., $O = M$.
(2) $\Longrightarrow$ (1): Angenommen $M$ ist nicht zusammenhängend. Dann gibt es zwei nicht leere, in $M$ relativ offene Mengen $O_1, O_2$ mit

$$O_1 \cap O_2 = \emptyset \text{ und } O_1 \cup O_2 = M.$$

Also ist $O_1 = M \setminus O_2$ und somit auch relativ abgeschlossen in $M$. Aus (2) folgt $O_1 = \emptyset$ oder $O_1 = M$, d.h., $O_2 = \emptyset$; Widerspruch. □

Der Begriff des Zusammenhangs führt u.a. zu einem wichtigen Beweisprinzip:

> Sei $P(x)$ eine Eigenschaft, die man für alle $x$ aus einer Menge $M$ beweisen will. Dann zeigt man sukzessive:
> (1) $M$ ist zusammenhängend,
> (2) $\{x \in M : P(x) \text{ gilt}\} \neq \emptyset$,
> (3) $\{x \in M : P(x) \text{ gilt}\}$ ist offen in $M$,
> (4) $\{x \in M : P(x) \text{ gilt}\}$ ist abgeschlossen in $M$.
> Wegen Satz III.5.2 gilt dann $P(x)$ für alle $x \in M$.

Als nächstes geben wir eine wesentliche topologische Charakterisierung der Intervalle in $\mathbb{R}$.

Satz III.5.3. *Folgende Aussagen sind äquivalent:*

(1) *$M \subset \mathbb{R}$ ist zusammenhängend.*
(2) *$M$ ist ein Intervall.*

Beweis. (1) $\Longrightarrow$ (2): Sei

$$a = \inf M \in \overline{\mathbb{R}}\ ,\ b = \sup M \in \overline{\mathbb{R}}.$$

Dann ist

$$M \subset [a, b].$$

Wir müssen noch

$$(a, b) \subset M$$

nachweisen. Angenommen, dies sei nicht der Fall. Dann gibt es ein $c$ mit

$$a < c < b \ \text{ und } c \notin M.$$

Definiere

$$O_1 = [a, c) \cap M\ ,\ O_2 = (c, b] \cap M.$$

Dann sind $O_1$ und $O_2$ relativ offen in $M$. Wegen der Definition von $a, b, c$ und wegen $c \notin M$ gilt

$$O_1 \cup O_2 = M, O_1 \cap O_2 = \emptyset, O_1 \neq \emptyset, O_2 \neq \emptyset,$$

im Widerspruch zum Zusammenhang von $M$.
(2) $\Longrightarrow$ (1): Angenommen $M$ ist nicht zusammenhängend. Dann existieren in $M$ relativ offene Mengen $O_1$ und $O_2$ mit

$$O_1 \neq \emptyset, O_2 \neq \emptyset, O_1 \cap O_2 = \emptyset, O_1 \cup O_2 = M.$$

Seien $x \in O_1$ und $y \in O_2$. O.E. ist $x < y$. Sei

$$z = \sup(O_1 \cap [x, y]) \in \mathbb{R}.$$

Ann. $z \in O_1 :\Longrightarrow\ \exists\, \delta > 0 : [z, z + \delta] \subset O_1 \cap [x, y] \Longrightarrow$ Widerspruch.
Ann. $z \in O_2 :\Longrightarrow\ \exists\, \delta > 0 : [z - \delta, z] \subset O_2 \cap [x, y] \Longrightarrow$ Widerspruch.
Also ist $z \notin O_1 \cup O_2 = M$. Andererseits gilt

$$M \ni x \leq z \leq y \in M$$

im Widerspruch zur Voraussetzung, dass $M$ ein Intervall ist. $\square$

Als nächstes beschäftigen wir uns mit den Auswirkungen des Zusammenhangs auf stetige Funktionen.

Satz III.5.4. *Sei $M \subset X$, $M \neq \emptyset$, zusammenhängend und $f \in C(M, Y)$. Dann ist $f(M)$ zusammenhängend.*

Beweis. Sei $V \subset f(M)$ nicht leer und in $f(M)$ relativ offen und relativ abgeschlossen. Da $f$ stetig ist, ist $U = f^{-1}(V)$ in $M$ relativ offen und relativ abgeschlossen. Wegen $\emptyset \neq V \subset f(M)$ ist $U \neq \emptyset$. Da $M$ zusammenhängend ist, folgt $U = M$. Also ist

$$V = f(U) = f(M).$$

Damit folgt die Behauptung aus Satz III.5.2. $\square$

Satz III.5.5 (Zwischenwertsatz). *Sei $M \subset X$, $M \neq \emptyset$, zusammenhängend, $f \in C(M, \mathbb{R})$ und $x, y \in M$ mit $f(x) < f(y)$. Dann nimmt $f$ jeden Wert zwischen $f(x)$ und $f(y)$ an, d.h., zu jedem $a \in [f(x), f(y)]$ gibt es ein $z \in M$ mit $f(z) = a$.*

Beweis. Wegen Satz III.5.3 und Satz III.5.4 ist $f(M)$ ein Intervall. Daher gilt $[f(x), f(y)] \subset f(M)$. □

Eine Konsequenz des Zwischenwertsatzes ist:

Satz III.5.6. *Sei $p(x) = \sum_{k=0}^{n} a_k x^k$ ein Polynom mit reellen Koeffizienten und ungeradem Grad $n$, d.h., $a_k \in \mathbb{R}$, $a_n \neq 0$ und $n = 2\ell + 1$, $\ell \in \mathbb{N}$. Dann besitzt $p$ eine reelle Nullstelle.*

Beweis. O.E. ist $n \geq 3$ und $a_n = 1$. Sei

$$R = 1 + \sum_{k=0}^{n-1} |a_k|.$$

Dann folgt

$$\begin{aligned}
p(R) &\geq R^n - \sum_{k=0}^{n-1} |a_k| R^k \\
&\geq R^n - \sum_{k=0}^{n-1} |a_k| R^{n-1} \\
&= R^{n-1}(R - (R-1)) \\
&\geq R
\end{aligned}$$

und

$$\begin{aligned}
p(-R) &\leq (-R)^n + \sum_{k=0}^{n-1} |a_k| R^k \\
&\leq -R^n + \sum_{k=0}^{n-1} |a_k| R^{n-1} \\
&= -R^{n-1}(R - (R-1)) \\
&= -R^{n-1} \\
&< 0.
\end{aligned}$$

Damit folgt die Behauptung aus Satz III.5.5. □

Bemerkung III.5.7. Satz III.5.6 ist für Polynome mit geradem Grad i. a. falsch, wie das Beispiel $p(x) = x^2 + 1$ zeigt.

Wir kommen nun zu einigen Abwandlungen des Zusammenhangbegriffes.

Definition III.5.8. (1) Eine Funktion $\omega \in C([0,1],X)$ heißt (stetiger) Weg in $X$. Die Punkte $\omega(0) \in X$ und $\omega(1) \in X$ heißen Anfangs- und Endpunkt des Weges. Die Menge $\omega([0,1]) \subset X$ heißt Spur des Weges.
(2) Eine nicht leere Teilmenge $M$ von $X$ heißt wegzusammenhängend, wenn es zu jedem $x \in M$ und jedem $y \in M$ einen Weg mit Anfangspunkt $x$ und Endpunkt $y$ gibt, der ganz in $M$ verläuft, d.h., $\omega([0,1]) \subset M$.

Aus Satz III.5.4 folgt:

Bemerkung III.5.9. Die Spur eines Weges ist zusammenhängend.

Satz III.5.10. *Jede wegzusammenhängende Menge $M$ ist zusammenhängend.*

Beweis. Wir nehmen an $M \subset X$, $M \neq \emptyset$, sei wegzusammenhängend, aber nicht zusammenhängend. Dann gibt es zwei nicht leere, disjunkte, in $M$ offene Mengen $O_1, O_2$ mit $O_1 \cup O_2 = M$. Seien $x \in M \cap O_1$ und $y \in M \cap O_2$. Dann gibt es einen Weg $\omega$, der ganz in $M$ verläuft mit Anfangspunkt $x$ und Endpunkt $y$. Sei $S$ die Spur von $\omega$. Dann folgt

$$\begin{aligned}
&(O_1 \cap S) \cup (O_2 \cap S) = M \cap S = S \\
&(O_1 \cap S) \cap (O_2 \cap S) = \emptyset \\
&O_1 \cap S \neq \emptyset, O_2 \cap S \neq \emptyset \\
&O_i \cap S \text{ ist relativ offen in } S, i = 1,2.
\end{aligned}$$

Dies ist ein Widerspruch zu Bemerkung III.5.9. □

Definition III.5.11. Eine nicht leere Teilmenge $M$ von $X$ heißt konvex, wenn für alle $x, y \in M$ gilt (vgl. Abbildung III.5.1)

$$\{tx + (1-t)y : 0 \leq t \leq 1\} \subset M.$$

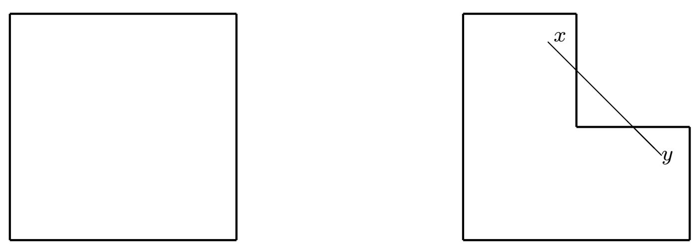


Abbildung III.5.1. Beispiel einer konvexen (links) und einer nicht konvexen (rechts) Menge

Bemerkung III.5.12. (1) Eine konvexe Menge ist wegzusammenhängend und damit zusammenhängend.
(2) $B(a;r)$ und $\overline{B(a;r)}$ sind für jedes $a \in X$ und $r \in \mathbb{R}_+^*$ konvex.

Beweis. ad (1): $\omega \in C([0,1], M)$ mit $\omega(t) = tx + (1-t)y$ ist ein Weg mit Anfangspunkt $y$ und Endpunkt $x$.
ad (2): Für $x, y \in B(a; r)$ und $t \in [0,1]$ gilt

$$\begin{aligned}\|a - (tx + (1-t)y)\|_X &= \|ta - tx + (1-t)a - (1-t)y\|_X \\ &\le t\|a - x\| + (1-t)\|a - y\| \\ &< tr + (1-t)r \\ &= r.\end{aligned}$$

Analog für $x, y \in \overline{B(a; r)}$. □

Lemma III.5.13. *Durch*

> *„$x \sim y$ für $x, y \in M$ genau dann, wenn es einen Weg in $M$ mit Anfangspunkt $x$ und Endpunkt $y$ gibt“*

*wird eine Äquivalenzrelation $\sim$ auf $M$ definiert.*

Beweis. (1) $x \sim x : \omega(t) = x$ für alle $t \in [0,1]$ ist stetig.
(2) $x \sim y \Longrightarrow y \sim x : \omega \in C([0,1], M)$ $\omega(0) = x, \omega(1) = y \Longrightarrow w^*(t) = \omega(1-t)$ leistet das Gewünschte.
(3) $x \sim y, y \sim z \Longrightarrow x \sim z :$

$$\begin{aligned}x \sim y &\Longrightarrow\ \exists\, \omega_1 \in C([0,1], M) \ \text{ mit } \omega_1(0) = x, \omega_1(1) = y \\ y \sim z &\Longrightarrow\ \exists\, \omega_2 \in C([0,1], M) \ \text{ mit } \omega_2(0) = y, \omega_2(1) = z.\end{aligned}$$

$$\Longrightarrow \omega(t) = \begin{cases} \omega_1(2t) & 0 \le t \le \frac{1}{2} \\ \omega_2(2t-1) & \frac{1}{2} < t \le 1 \end{cases}$$

leistet das Gewünschte. □

Satz III.5.14. *$M \subset X$, $M \neq \emptyset$, sei offen. Dann sind folgende Aussagen äquivalent:*

(1) *$M$ ist zusammenhängend.*
(2) *$M$ ist wegzusammenhängend.*

Beweis. (1) $\Longrightarrow$ (2): Sei $a \in M$ beliebig und

$$U = \{x \in M : x \sim a\}.$$

Dann ist $a \in U$.
Sei $x \in U$. Da $M$ offen ist, gibt es ein $\varepsilon > 0$ mit $B(x; \varepsilon) \subset M$. Sei $y \in B(x; \varepsilon)$. Gemäß Bemerkung III.5.12 (2) gilt $y \sim x$ und wegen $x \sim a$ auch $y \sim a$. Also ist $B(x; \varepsilon) \subset U$. Mithin ist $U$ offen in $M$.
Sei $x \in M \backslash U$. Da $M$ offen ist, gibt es wieder ein $\varepsilon > 0$ mit $B(x; \varepsilon) \subset M$. Dann ist auch $B(x; \varepsilon) \cap U = \emptyset$. Denn sonst gäbe es ein $y \in B(x; \varepsilon)$ mit $y \sim a$; wegen $y \sim x$ folgte $x \sim a$ ein Widerspruch zu $x \notin U$. Mithin ist $U$ auch abgeschlossen in $M$.
Da $M$ zusammenhängend ist, folgt $U = M$. Also ist $M$ wegzusammenhängend.
(2) $\Longrightarrow$ (1): Satz III.5.10. □

Bemerkung III.5.15. Auf die Voraussetzung „$M$ offen“ kann in Satz III.5.14 nicht verzichtet werden. Denn sei $X = \mathbb{R}^2$

$$X' = \{(x,y) \in \mathbb{R}^2 : 0 \le x \le 1, 0 \le y < 1\}$$
$$Y_n = \{(x,y) \in \mathbb{R}^2 : x = \frac{1}{n}, 0 \le y < 1 - \frac{1}{n}\}$$

und

$$M = X' \setminus \bigcup_{n \in \mathbb{N}^*} X_n.$$

Die Menge $M$ (vgl. Abbildung III.5.2) ist zusammenhängend aber nicht wegzusammenhängend.

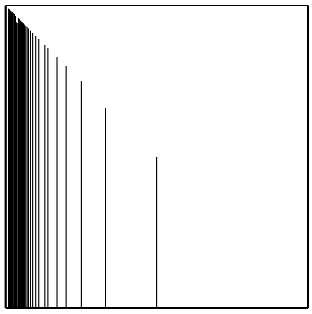

Abbildung III.5.2. Skizze der Menge $M$ aus Bemerkung III.5.15. Dicke Linien gehören zu $M$, dünne Linien nicht.

Beweis. (1): $(0,0) \not\sim (1,0)$. Angenommen $\omega = (\omega_1, \omega_2) \in C([0,1], M)$ erfüllt $\omega(0) = (0,0), \omega(1) = (1,0)$. Aus Satz III.4.10 (S. 86) folgt

$$m = \max_{0 \le t \le 1} \omega_2(t) < 1.$$

Sei $n \in \mathbb{N}$ so, dass $1 - \frac{1}{n} > m$ ist. Aus Satz III.5.5 folgt, dass es ein $t^* \in (0,1)$ gibt mit

$$\omega_1(t^*) = \frac{1}{n}.$$

Dann ist

$$\omega(t^*) \in Y_n \iff \omega(t^*) \notin M.$$

Dies ist ein Widerspruch.
(2): $(0, y_1) \sim (0, y_2)$ für alle $y_1, y_2 \in [0,1)$.
(3): $(x_1, y_1) \sim (x_2, y_2)$ für alle $(x_1, y_1), (x_2, y_2) \in M$ mit $x_1 > 0, x_2 > 0$.
Sei nun $N \subset M$ nicht leer und in $M$ relativ offen und in $M$ relativ abgeschlossen.
(4): Bemerkung III.5.7 impliziert:

$$(x_1, y_1) \sim (x_2, y_2) \text{ in } M \text{ und } (x_1, y_1) \in N \Longrightarrow (x_2, y_2) \in N.$$

(5): $N$ relativ offen in $M \Longrightarrow \exists\, (x,y) \in N$ mit $x > 0$.
(6): (3), (4), (5) $\Longrightarrow \forall x > 0 \quad \forall y \in [0,1) : (x,y) \in N$.

(7): (6) $\Longrightarrow (\frac{\sqrt{2}}{n}, \frac{1}{2}) \in N \quad \forall n \geq 2.$
$N$ relativ abgeschlossen in $M$ und

$$\lim_{n\to\infty}(\frac{\sqrt{2}}{n}, \frac{1}{2}) = (0, \frac{1}{2}) \in M \Longrightarrow (0, \frac{1}{2}) \in N$$

(8): (2), (4), (7) $\Longrightarrow (0, y) \in N \quad \forall y \in [0, 1)$
Also ist $N = M$ und $M$ somit zusammenhängend. □

Definition III.5.16. Eine nicht leere, offene und zusammenhängende Teilmenge von $X$ heißt Gebiet.

## III.6. Funktionen in $\mathbb{R}$

Im Folgenden sei $I \subset \mathbb{R}$ stets ein nicht leeres Intervall.

Definition III.6.1. Eine Funktion $f : I \to \mathbb{R}$ heißt

(a) monoton wachsend,
(b) streng monoton wachsend,
(c) monoton fallend,
(d) streng monoton fallend,

wenn für alle $x, y \in I$ mit $x < y$ gilt

(a) $f(x) \leq f(y)$,
(b) $f(x) < f(y)$,
(c) $f(x) \geq f(y)$,
(d) $f(x) > f(y)$.

Eine monoton wachsende oder monoton fallende Funktion heißt monoton.

Beispiel III.6.2. (1) $f(x) = x^2$ ist streng monoton fallend auf $(-\infty, 0]$ und streng monoton wachsend auf $[0, +\infty)$. Falls $0 \in \mathring{I}$ ist, ist $f$ auf $I$ nicht monoton.
(2) Die Funktion $f(x) = \lceil x \rceil$ aus Beispiel III.3.4(3) (S. 76) ist monoton wachsend.
(3) Die Funktion

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

aus Beispiel III.3.4(4) (S. 76) ist nicht monoton.

Satz III.6.3. *Sei $f : I \to \mathbb{R}$ monoton und $\alpha = \inf I \in \overline{\mathbb{R}}$, $\beta = \sup I \in \overline{\mathbb{R}}$. Dann existieren in $\overline{\mathbb{R}}$ die Grenzwerte*

$$f(\alpha + 0) = \lim_{x\to\alpha+0} f(x) = \lim_{\substack{x\to\alpha \\ x>\alpha}} f(x)$$

$$f(\beta - 0) = \lim_{x\to\beta-0} f(x) = \lim_{\substack{x\to\beta \\ x<\beta}} f(x)$$

*und es gilt*

$$f(\alpha+0)=\begin{cases}\inf\{f(x):x\in(\alpha,\beta]\} & \textit{falls } f \textit{ wachsend,}\\ \sup\{f(x):x\in(\alpha,\beta]\} & \textit{falls } f \textit{ fallend,}\end{cases}$$

$$f(\beta-0)=\begin{cases}\sup\{f(x):x\in[\alpha,\beta)\} & \textit{falls } f \textit{ wachsend,}\\ \inf\{f(x):x\in[\alpha,\beta)\} & \textit{falls } f \textit{ fallend.}\end{cases}$$

Beweis. Sei $f$ monoton wachsend und

$$b=\sup f(I)\in\overline{\mathbb{R}}.$$

Sei $(x_n)_{n\in\mathbb{N}}\subset I$ eine Folge mit $\lim\limits_{n\to\infty}x_n=\beta$ und $x_n<\beta$ für alle $n\in\mathbb{N}$. Dann gibt es zu jedem $\eta<b$ ein $\zeta\in I$, $\zeta<\beta$, mit $f(\zeta)=\eta$. Zu $\zeta$ gibt es ein $n_\zeta\in\mathbb{N}$ mit

$$\zeta\le x_n<\beta\quad\forall n\ge n_\zeta.$$

Da $f$ monoton wachsend ist, folgt

$$\eta=f(\zeta)\le f(x_n)\le b\quad\forall n\ge n_\zeta.$$

Also gilt $\lim\limits_{n\to\infty}f(x_n)=b$. Also existiert $f(\beta-0)$ und ist gleich $b$. Ganz analog zeigt man die Existenz von $f(\alpha+0)$ und geht für monoton fallendes $f$ vor. □

Definition III.6.4. Seien $D\subset\mathbb{R}$, $D\neq\emptyset$, $Y$ ein normierter Vektorraum und $f:D\to Y$ ein Funktion. Ein Punkt $x_0\in\mathbb{R}$ mit $x_0\in\overline{D\cap(-\infty,x_0)}\cap\overline{D\cap(x_0,+\infty)}$ heißt Sprungstelle von $f$, wenn die Grenzwerte $\lim\limits_{x\to x_0\pm 0}f(x)=f(x_0\pm 0)$ existieren und verschieden sind.

Beispiel III.6.5. Für die Funktion $f$ aus Beispiel III.6.2 (2) ist jedes $k\in\mathbb{Z}$ eine Sprungstelle. Es ist

$$f(k+0)-f(k-0)=1\quad\forall k\in\mathbb{Z}.$$

Satz III.6.3 zeigt, dass für eine monotone Funktion $f:I\to\mathbb{R}$ für jedes $x_0\in\overset{\circ}{I}$ die Grenzwerte $f(x_0\pm 0)$ existieren. Der folgende Satz zeigt, dass eine solche Funktion bis auf abzählbar viele Sprungstellen stetig ist. Beispiel III.6.5 zeigt, dass umgekehrt auch abzählbar viele Sprungstellen auftreten können.

Satz III.6.6. *Sei $f:I\to\mathbb{R}$ monoton. Dann ist $f$ bis auf abzählbar viele Sprungstellen stetig.*

Beweis. O.E. ist $f$ monoton wachsend. Sonst gehen wir zu $-f$ über. Definiere

$$M=\{x\in\overset{\circ}{I}:f(x+0)\neq f(x-0)\}.$$

Wir müssen zeigen, dass $M$ abzählbar ist. Dazu genügt es, eine injektive Abbildung von $M$ in eine abzählbare Menge zu konstruieren. Sei dazu

$x \in M$. Da $f$ monoton wachsend ist, ist $f(x+0) > f(x-0)$. Also gibt es ein

$$q_x \in \mathbb{Q} \cap (f(x-0), f(x+0)).$$

$x \mapsto q_x$ ist eine Abbildung von $M$ in $\mathbb{Q}$. Seien nun $x, y \in M$ mit $x < y$. Da $f$ monoton wachsend ist, folgt

$$f(x+0) \le f(y-0).$$

Also gilt

$$q_x < f(x+0) \le f(y-0) < q_y.$$

Mithin ist die Abbildung injektiv. Damit folgt die Behauptung aus Satz I.3.2 (S. 17). □

Der folgende Satz zeigt, dass stetige, monotone Funktionen sehr gutartig sind. Er wird uns im nächsten Paragraphen gute Dienste leisten.

Satz III.6.7. *Sei $f : I \to \mathbb{R}$ stetig und streng monoton wachsend bzw. fallend. Dann gilt*

(1) *$J = f(I)$ ist ein Intervall.*
(2) *$f : I \to J$ ist bijektiv.*
(3) *Die Umkehrabbildung $f^{-1} : J \to I$ ist stetig und streng monoton wachsend bzw. fallend.*

Beweis. ad (1): Folgt aus Satz III.5.3 (S. 88) und III.5.4 (S. 89).
ad (2): Die Surjektivität folgt aus der Definition von $J$, die Injektivität aus der strengen Monotonie.
ad (3): Sei $g = f^{-1} : J \to \mathbb{R}$. O.E. setzen wir voraus, dass $f$ streng monoton wachsend ist. Andernfalls gehen wir zu $-f$ über.
Sei $y_1, y_2 \in J$ mit $y_1 < y_2$. Angenommen es ist $g(y_1) \ge g(y_2)$. Da $f$ streng monoton wachsend ist, folgt

$$y_2 = f(g(y_2)) \le f(g(y_1)) = y_1.$$

Also ist $g$ streng monoton wachsend.
Sei nun $y_0 \in \overset{\circ}{I}$. Gemäß Satz III.6.3 existieren die Grenzwerte

$$g(y_0+0) \ \text{ und } g(y_0-0).$$

Wir müssen zeigen, dass sie mit $g(y_0)$ übereinstimmen. Sei $(y_k)_{k\in\mathbb{N}} \subset J$ eine monoton wachsende Folge mit

$$\lim_{n\to\infty} y_k = y_0 \ \text{ und } y_k < y_0 \quad \forall k \in \mathbb{N}.$$

Dann ist $(x_k)_{k\in\mathbb{N}}$ mit $x_k = g(y_k)$ eine monoton wachsende Folge mit

$$x_k < x_0 = g(y_0) \quad \forall k \in \mathbb{N}.$$

Sei

$$\overline{x} = \lim_{k\to\infty} x_k.$$

Dann ist

$$\overline{x} \leq x_0 \text{ und } \overline{x} \in I.$$

Angenommen es wäre, $\overline{x} < x_0$. Dann folgt aus der Stetigkeit und der strengen Monotonie von $f$

$$\begin{aligned} y_0 &= \lim_{k\to\infty} y_k = \lim_{k\to\infty} f(x_k) \\ &= f(\overline{x}) \\ &< f(x_0) = y_0 \end{aligned}$$

Also ist $\overline{x} = x_0$ und somit

$$g(y_0 - 0) = g(y_0).$$

Ganz analog folgt $g(y_0 + 0) = g(y_0)$. Also ist $g$ stetig. □

Beispiel III.6.8. Sei $n \geq 2$ und $f : \mathbb{R}_+ \to \mathbb{R}_+$ definiert durch

$$f(x) = x^n.$$

Für $0 \leq x < y$ gilt

$$\begin{aligned} f(y) - f(x) &= y^n - x^n \\ &= y^n(1 - (\frac{x}{y})^n) \\ &= \underbrace{y^n}_{>0} \underbrace{(1 - (\frac{x}{y}))}_{>0} \sum_{k=0}^{n-1} \underbrace{(\frac{x}{y})^k}_{>0} \\ &> 0. \end{aligned}$$

Also ist $f$ streng monoton wachsend.
Wegen $0 \leq f(x) = x^n \leq x$ für alle $0 \leq x \leq 1$ gilt

$$\lim_{x\to 0+} f(x) = 0.$$

Wegen $f(x) = x^n \geq x$ für alle $x \geq 1$ gilt

$$\lim_{x\to +\infty} f(x) = +\infty.$$

Also ist gemäß Satz III.6.7 $f : \mathbb{R}_+ \to \mathbb{R}_+$ bijektiv und es gibt eine stetige, streng monoton wachsende Umkehrfunktion $g : \mathbb{R}_+ \to \mathbb{R}_+$. Es ist

$$g(x) = \sqrt[n]{x}.$$

## III.7. Exponentialfunktion und Verwandte

Definition III.7.1. Durch

$$\exp(z) = e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

$$\cos(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}$$

$$\sin(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}$$

werden drei Funktionen von $\mathbb{C}$ in $\mathbb{C}$, genannt EXPONENTIALFUNKTION, COSINUS und SINUS, definiert.

SATZ III.7.2. *Es gilt:*

(1) *Die Potenzreihen in Definition III.7.1 haben alle den Konvergenzradius* $\infty$.
(2) $\exp, \sin, \cos$ *sind reellwertig auf* $\mathbb{R}$.
(3) $e^{z_1+z_2} = e^{z_1} \cdot e^{z_2}$ *für alle* $z_1, z_2 \in \mathbb{C}$.
(4) $e^z \neq 0$ *für alle* $z \in \mathbb{C}; e^{-z} = \frac{1}{e^z}$ *für alle* $z \in \mathbb{C}$
(5) $e^{iz} = \cos z + i \sin z$ *für alle* $z \in \mathbb{C}$.
(6) $\cos$ *ist gerade, d.h.,* $\cos(-z) = \cos(z)$ *für alle* $z \in \mathbb{C}$.
$\sin$ *ist ungerade, d.h.,* $\sin(-z) = -\sin(z)$ *für alle* $z \in \mathbb{C}$.
(7) $\cos(z) = \frac{1}{2}(e^{iz} + e^{-iz})$ *für alle* $z \in \mathbb{C}$.
$\sin(z) = \frac{1}{2i}(e^{iz} - e^{-iz})$ *für alle* $z \in \mathbb{C}$.
(8) $\cos(x) = \mathrm{Re}(e^{ix})$ *für alle* $x \in \mathbb{R}$.
$\sin(x) = \mathrm{Im}(e^{ix})$ *für alle* $x \in \mathbb{R}$.
(9) $\exp, \cos, \sin$ *sind stetig.*

BEWEIS. AD (1): Folgt aus Satz II.6.4 (S. 58).
AD (2): Ist klar.
AD (3): Beispiel II.5.15 (S. 56).
AD (4): Beispiel II.5.15 (S. 56).
AD (5):

$$\begin{aligned} e^{iz} &= \sum_{n=0}^{\infty} \frac{(iz)^n}{n!} \\ &= \sum_{k=0}^{\infty} (i)^{2k} \frac{z^{2k}}{(2k)!} + \sum_{k=0}^{\infty} (i)^{2k+1} \frac{z^{2k+1}}{(2k+1)!} \qquad \text{wg. abs. Konv. mögl.} \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} \\ &= \cos(z) + i \sin(z). \end{aligned}$$

AD (6): Ist offensichtlich.
AD (7):

$$\begin{aligned} \frac{1}{2}(e^{iz} + e^{-iz}) &= \frac{1}{2}(\cos(z) + i\sin(z) + \cos(-z) + i\sin(-z)) \\ &= \cos(z) \\ \frac{1}{2i}(e^{iz} - e^{-iz}) &= \frac{1}{2i}(\cos(z) + i\sin(z) - \cos(-z) - i\sin(-z)) \\ &= \sin(z). \end{aligned}$$

AD (8): Folgt aus (2) und (5).
AD (9): Gemäß Beispiel III.3.14(2) (S. 79) gibt es ein $\rho > 0$ mit

$$\frac{|e^z - 1|}{|z|} \leq 2 \qquad \forall |z| < \rho.$$

Sei nun $z_0 \in \mathbb{C}$ und $\varepsilon > 0$. Wähle

$$\delta = \min\{\rho, \frac{\varepsilon}{2|e^{z_0}|}\} > 0.$$

Dann folgt für $z \in \mathbb{C}$ mit $|z - z_0| < \delta$

$$\begin{aligned} |e^z - e^{z_0}| &= |e^{z_0}| \cdot \frac{|e^{(z-z_0)} - 1|}{|z - z_0|} |z - z_0| \\ &\leq 2|e^{z_0}||z - z_0| \\ &< \varepsilon. \end{aligned}$$

Also ist exp stetig. Wegen Satz III.3.8 (S. 77) ist dann auch

$$z \mapsto \exp(-z)$$

stetig. Damit folgt die Stetigkeit von Sinus und Cosinus aus (7) und Satz III.3.5 (S. 77). □

SATZ III.7.3 (ADDITIONSTHEOREME). *Für alle $z, w \in \mathbb{C}$ gilt:*
(1) $\cos(z \pm w) = \cos(z)\cos(w) \mp \sin(z)\sin(w)$.
$\sin(z \pm w) = \sin(z)\cos(w) \pm \cos(z)\sin(w)$.
(2) $\sin(z) - \sin(w) = 2\cos(\frac{z+w}{2})\sin(\frac{z-w}{2})$.
$\cos(z) - \cos(w) = -2\sin(\frac{z+w}{2})\sin(\frac{z-w}{2})$.
(3) $\cos^2(z) + \sin^2(z) = 1$.

BEWEIS. AD (1):

$$\begin{aligned} \cos(z + w) &= \frac{1}{2}(e^{i(z+w)} + e^{-(z+w)}) \\ &= \frac{1}{2}[e^{iz}e^{iw} + e^{-iz}e^{-iw}] \\ &= \frac{1}{4}[(e^{iz} + e^{-iz})(e^{iw} + e^{-iw}) \\ &\qquad + (e^{iz} - e^{-iz})(e^{iw} - e^{-iw})] \\ &= \cos(z)\cos(w) - \sin(z)\sin(w) \end{aligned}$$

Die Identität für $\cos(z - w)$ folgt aus $\cos(z - w) = \cos(z + (-w))$ und Satz III.7.2 (6). Der Beweis der Identität für $\sin(z \pm w)$ ist völlig analog.
AD (2): Setze $u = \frac{z-w}{2}$, $v = \frac{z+w}{2}$. Dann ist

$$z = u + v, \ \ w = v - u.$$

Damit folgt aus (1) und Satz III.7.2 (6)

$$\sin(z) - \sin(w) = \sin(u + v) + \sin(u - v)$$

$$
\begin{aligned}
&= 2\sin(u)\cos(v) \\
&= 2\sin(\frac{z-w}{2})\cos(\frac{z+w}{2})
\end{aligned}
$$

Die zweite Identität folgt analog.
AD (3): Aus (1) folgt

$$\cos^2(z) + \sin^2(z) = \cos(z-z) = \cos(0) = 1.$$

□

Wegen $e^{x+iy} = e^x \cdot e^{iy}$ ist das Verhalten der Exponentialfunktion auf $\mathbb{C}$ festgelegt durch ihr Verhalten auf $\mathbb{R} \subset \mathbb{C}$ und $i\mathbb{R} \subset \mathbb{C}$. Wir betrachten zunächst die reelle Exponentialfunktion.

SATZ III.7.4. *Es gilt:*
(1) $0 < e^x < 1$ *für alle* $x < 0$;
$e^x > 1$ *für alle* $x > 0$.
(2) $\exp : \mathbb{R} \to \mathbb{R}_+^*$ *ist streng monoton wachsend.*
(3) $\lim\limits_{x\to+\infty} x^{-n}e^x = +\infty$ *für alle* $n \in \mathbb{N}$,
$\lim\limits_{x\to-\infty} e^x = 0$.

BEWEIS. AD (1): Für $x > 0$ folgt

$$e^x = 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!} > 1.$$

Für $x < 0$ folgt

$$e^x = e^{-(-x)} = \frac{1}{e^{-x}} \in (0,1).$$

AD (2): Für $x < y$ folgt aus (1)

$$e^y = e^x e^{y-x} > e^x.$$

AD (3): Für $n \in \mathbb{N}$ und $x > 0$ folgt

$$x^{-n}e^x > x^{-n}\frac{x^{n+1}}{(n+1)!} = \frac{x}{(n+1)!} \underset{x\to+\infty}{\longrightarrow} +\infty.$$

Schließlich gilt

$$\lim_{x\to-\infty} e^x = \lim_{x\to+\infty} e^{-x} = \lim_{x\to+\infty} \frac{1}{e^x} = 0.$$

□

Wegen Satz III.7.4 und Satz III.6.7 (S. 96) ist die folgende Definition sinnvoll.

DEFINITION III.7.5. Die Umkehrfunktion der reellen Exponentialfunktion heißt (NATÜRLICHER) LOGARITHMUS und wird mit ln bezeichnet;

$$\ln = (\exp|_{\mathbb{R}})^{-1} : \mathbb{R}_+^* \to \mathbb{R}.$$

SATZ III.7.6. (1) *Der Logarithmus ist eine stetige, streng monoton wachsende Funktion mit*

$$\lim_{x\to+\infty} \ln(x) = +\infty, \quad \lim_{x\to 0+} \ln(x) = -\infty.$$

(2) *Für alle $x, y \in \mathbb{R}_+^*$ gilt*

$$\ln(x \cdot y) = \ln(x) + \ln(y)$$
$$\ln(\frac{x}{y}) = \ln(x) - \ln(y).$$

BEWEIS. AD (1): Folgt aus Satz III.6.7 (S. 96) und Satz III.7.4. AD (2): Sei $a = \ln(x)$, $b = \ln(y)$. Dann folgt aus Satz III.7.2 und der Definition des Logarithmus

$$x \cdot y = e^a \cdot e^b = e^{a+b}$$
$$\Longrightarrow \ln(x) + \ln(y) = a + b = \ln(x \cdot y)$$

Analog für $\ln(\frac{x}{y})$. □

Sei $a > 0$. Dann folgt

$$a = e^{\ln(a)}$$

und damit

$$a^n = (e^{\ln(a)})^n = e^{n\ln(a)} \quad \forall n \in \mathbb{N}$$
$$a^{-n} = \frac{1}{a^n} = \frac{1}{e^{n\ln(a)}} = e^{-n\ln(a)} \quad \forall n \in \mathbb{N}.$$

Sei $n \in \mathbb{N}^*$ und $x = e^{\frac{1}{n}\ln(a)} \in \mathbb{R}_+^*$. Dann folgt

$$x^n = e^{n\frac{1}{n}\ln(a)} = e^{\ln(a)} = a.$$

Also gilt

$$\sqrt[n]{a} = e^{\frac{1}{n}\ln(a)} \quad \forall n \in \mathbb{N}^*.$$

Hieraus folgt für $r = \frac{p}{q} \in \mathbb{Q}$ mit $p \in \mathbb{Z}$, $q \in \mathbb{N}^*$

$$a^{\frac{p}{q}} = e^{\frac{p}{q}\ln(a)}.$$

Dies führt zu folgender Definition.

DEFINITION III.7.7. Für $a > 0$ und $x \in \mathbb{R}$ sei

$$a^x = e^{x\ln(a)}.$$

SATZ III.7.8. *Seien $a, b \in \mathbb{R}_+^*$ und $x, y \in \mathbb{R}$ dann gilt:*

(1) $a^x a^y = a^{x+y}$, $\frac{a^x}{a^y} = a^{x-y}$.
(2) $a^x b^x = (ab)^x$, $\frac{a^x}{b^x} = (\frac{a}{b})^x$.
(3) $(a^x)^y = a^{x\cdot y}$.
(4) $\ln(a^x) = x\ln(a)$.

BEWEIS. Übungsaufgabe. □

Der folgende Satz sagt etwas aus über das asymptotische Verhalten des Logarithmus für $x \to 0+$ und $x \to +\infty$.

Satz III.7.9. *Für $\alpha \in \mathbb{R}_+^*$ gilt*

$$\lim_{x\to+\infty} x^{-\alpha}\ln(x) = 0, \quad \lim_{x\to 0+} x^{\alpha}\ln(x) = 0.$$

Beweis. Da aus $x \to +\infty$ folgt $y = \ln(x) \to +\infty$ erhalten wir mit Satz III.7.4 (3)

$$\begin{aligned}\lim_{x\to+\infty} x^{-\alpha}\ln(x) &= \lim_{y\to+\infty} \frac{y}{e^{\alpha y}}\\ &= \frac{1}{\alpha}\lim_{y\to+\infty}\frac{\alpha y}{e^{\alpha y}}\\ &= \frac{1}{\alpha}\lim_{z\to+\infty}\frac{z}{e^{z}} \qquad (z=\alpha y)\\ &= 0.\end{aligned}$$

Ebenso folgt

$$\begin{aligned}\lim_{x\to 0+} x^{\alpha}\ln(x) &= \lim_{y\to+\infty} y^{-\alpha}\ln(\frac{1}{y}) \qquad (x=\frac{1}{y})\\ &= -\lim_{y\to+\infty} y^{-\alpha}\ln(y)\\ &= 0.\end{aligned}$$

□

Als nächstes untersuchen wir das Verhalten der Exponentialfunktion auf $i\mathbb{R} \subset \mathbb{C}$. Um Schreibarbeit zu sparen, definieren wir für $x \in \mathbb{R}$

$$\operatorname{ixp}(x) = e^{ix} \quad \forall x \in \mathbb{R}.$$

Satz III.7.10. $\operatorname{ixp}(\mathbb{R}) = S^1 = \{z \in \mathbb{C} : |z| = 1\}$.

Beweis. Wegen Satz III.7.3 (3) ist $\operatorname{ixp}(\mathbb{R}) \subset S^1$. Um die Gleichheit zu zeigen, beweisen wir zunächst:
Beh.: $\cos(\mathbb{R}) = [-1, 1]$.
Bew. der Beh.: Wegen Satz III.5.3 (S. 88) und Satz III.5.4 (S. 89) ist $I = \cos(\mathbb{R})$ ein Intervall. Wegen $\cos(0) = 1$ ist $1 \in I$. Wegen

$$\begin{aligned}\cos 1 &= 1 + \sum_{k=1}^{\infty}\{\frac{(-1)^{2k-1}}{(4k-2)!} + \frac{(-1)^{2k}}{(4k)!}\}\\ &= 1 - \sum_{k=1}^{\infty}\{\frac{1}{(4k-2)!} - \frac{1}{(4k)!}\}\\ &< 1\end{aligned}$$

ist $I \neq \{1\}$ und

$$a = \inf I < 1.$$

Wegen Satz III.7.3 (3) ist $a \ge -1$. Angenommen $a > -1$. Dann ist $\alpha = \frac{a+1}{2} \in I$ und es gibt ein $x$ mit $\cos(x) = \alpha$. Mit Satz III.7.3 folgt

$$\begin{aligned}\cos(2x) &= \cos^2(x) - \sin^2(x) \\ &= 2\cos^2(x) - 1 \\ &= 2(\frac{a+1}{2})^2 - 1 \\ &= a - \frac{1-a^2}{2} \\ &< a.\end{aligned}$$

Also ist $a = -1$. Insbesondere ist $0 \in I$ und es gibt ein $x$ mit $\cos(x) = 0$. Mit Satz III.7.3 folgt wieder

$$\cos(2x) = 2\cos^2(x) - 1 = -1.$$

Damit ist Beh. bewiesen.
Sei nun $z \in S^1$. Dann gibt es gemäß Behauptung ein $x \in \mathbb{R}$ mit

$$\cos(x) = \operatorname{Re} z \in [-1, 1].$$

Dann gilt aber $e^{ix} = z$ oder $\overline{(e^{ix})} = e^{-ix} = z$, d.h.,

$$\operatorname{ixp}(x) = z \;\text{ oder }\; \operatorname{ixp}(-x) = z.$$

Damit ist die Behauptung bewiesen. □

Satz III.7.11. *Die Menge $M = \{x \in \mathbb{R}_+^* : \operatorname{ixp}(x) = 1\}$ besitzt ein positives Minimum. Wir definieren*

$$\pi = \frac{1}{2}\min\{x \in \mathbb{R}_+^* : \operatorname{ixp}(x) = 1\}.$$

Beweis. Wegen Satz III.7.10 gibt es ein $x$ mit

$$\operatorname{ixp}(x) = -1.$$

Wegen $\operatorname{ixp}(0) = 1$ ist $x \neq 0$. Wegen

$$\operatorname{ixp}(-x) = [\operatorname{ixp}(x)]^{-1} = -1$$

ist o.E. $x > 0$. Wegen

$$\operatorname{ixp}(2x) = [\operatorname{ixp}(x)]^2 = 1$$

ist $M \neq \emptyset$. Für $x \in (0,1)$ gilt weiter

$$\begin{aligned}\sin(x) &= \sum_{k=0}^{\infty}\{\frac{(-1)^{2k}x^{4k+1}}{(4k+1)!} + \frac{(-1)^{2k+1}x^{4k+3}}{(4k+3)!}\} \\ &= \sum_{k=0}^{\infty} x^{4k+1}[\frac{1}{(4k+1)!} - \frac{x^2}{(4k+3)!}]\end{aligned}$$

$$\begin{aligned}&\geq \sum_{k=0}^{\infty} x^{4k+1}[\frac{1}{(4k+1)!} - \frac{1}{(4k+3)!}]\\ &> 0.\end{aligned}$$

Also ist

$$M = \varphi^{-1}(\{1\})$$

mit $\varphi = \mathrm{ixp}\,|_{[1,\infty)}$. Hieraus folgt die Behauptung. □

SATZ III.7.12. *Es gilt:*

(1) $e^z = 1 \iff z \in 2\pi i\mathbb{Z} \iff \exists\, k \in \mathbb{Z} : z = 2k\pi i.$
(2) $e^z = -1 \iff z \in \pi i + 2\pi i\mathbb{Z} \iff \exists\, k \in \mathbb{Z} : z = (2k+1)\pi i.$

BEWEIS. AD (1): „$\Longleftarrow$“: $e^{2k\pi i} = (e^{2\pi i})^k = 1^k = 1$
„$\Longrightarrow$“: $z = x + iy, x, y \in \mathbb{R}, e^z = 1$

$$\begin{aligned}&\Longrightarrow 1 = |e^z| = |e^x||e^{iy}| = e^x \quad \Longrightarrow \quad x = 0\\ &\quad y = 2k\pi + r \;\text{ mit } 0 \leq r < 2\pi\\ &\Longrightarrow 1 = e^z = e^{2k\pi i}e^{ir} = e^{ir} \quad \Longrightarrow \quad r = 0 \quad \text{nach Def. von } \pi.\end{aligned}$$

AD (2): „$\Longleftarrow$“: $e^{(2k+1)\pi i} = e^{\pi i} = a$

$$a^2 = e^{2\pi i} = 1 \Longrightarrow a = \pm 1 \Longrightarrow a = -1 \text{ nach Def. von } \pi.$$

„$\Longrightarrow$“:

$$\begin{aligned}e^z = -1 \Longrightarrow e^{2z} = 1 &\Longrightarrow 2z \in 2\pi i\mathbb{Z}\\ &\Longrightarrow z \in \pi i\mathbb{Z}\\ &\Longrightarrow z = (2\ell+1)\pi i \;\text{ mit } \ell \in \mathbb{Z} \text{ wegen } (1).\end{aligned}$$

□

Aus Satz III.7.12 folgt

$$e^z = e^{z+2k\pi i} \quad \forall z \in \mathbb{C}, k \in \mathbb{Z}.$$

Falls es für eine Funktion $f : \mathbb{K} \to \mathbb{K}$ ein $p \in \mathbb{K}^*$ gibt mit

$$f(x+p) = f(x) \quad \forall x \in \mathbb{K},$$

sagt man, $f$ sei $p$-PERIODISCH. exp ist also $2\pi i$ periodisch.

SATZ III.7.13. *Für jedes $a \in \mathbb{R}$ gilt*

$$\mathrm{ixp} : [a, a+2\pi) \to S^1$$

*und*

$$\mathrm{ixp} : (a, a+2\pi] \to S^1$$

*sind bijektiv.*

Beweis. Injektiv:

$$\begin{aligned}
\operatorname{ixp}(x)=\operatorname{ixp}(y) &\iff \operatorname{ixp}(x-y)=1\\
&\iff x-y=2k\pi \text{ für ein } k\in\mathbb{Z} \text{ (Satz III.7.12)}\\
&\iff x=y \text{ da } |x-y|<2\pi\ \forall x,y\in[a,a+2\pi)\\
&\qquad\qquad \text{bzw. } \forall x,y\in(a,a+2\pi].
\end{aligned}$$

Surjektiv: Zu $z\in S^1$ gibt es nach Satz III.7.10 ein $x\in\mathbb{R}$ mit $\operatorname{ixp}(x)=z$. $x$ lässt sich in der Form $x=a+2k\pi+r$ mit $k\in\mathbb{Z}$ und $r\in[0,2\pi)$ darstellen. Damit folgt

$$\begin{aligned}
z=\operatorname{ixp}(x)&=\operatorname{ixp}(a+r)\cdot\operatorname{ixp}(2k\pi)\\
&=\operatorname{ixp}(a+r).
\end{aligned}$$

Wegen $a+r\in[a,a+2\pi)$ folgt die Surjektivität. □

Satz III.7.14. *Es gilt:*

(1) $\cos(z+2k\pi)=\cos(z)$, $\sin(z+2k\pi)=\sin(z)$ *für alle* $z\in\mathbb{C}$, $k\in\mathbb{Z}$.
(2) $\cos(z)=0 \iff z\in\frac{\pi}{2}+\pi\mathbb{Z}$,
$\sin(z)=0 \iff z\in\pi\mathbb{Z}$.
(3) $\sin(x)>0$ *für alle* $x\in(0,\pi)$,
$\sin(x)$ *ist streng monoton wachsend auf* $(0,\frac{\pi}{2})$.
(4) $\cos(z+\pi)=-\cos(z)$, $\sin(z+\pi)=-\sin(z)$ *für alle* $z\in\mathbb{C}$.
(5) $\cos(z)=\sin(\frac{\pi}{2}-z)$, $\sin(z)=\cos(\frac{\pi}{2}-z)$ *für alle* $z\in\mathbb{C}$.
(6) $\cos(\mathbb{R})=\sin(\mathbb{R})=[-1,1]$.

Beweis. ad (1): Folgt aus Satz III.7.2 (7) und Satz III.7.12.
ad (2):

$$\begin{aligned}
\cos(z)=0 &\iff e^{iz}=-e^{-iz}\\
&\iff e^{2iz}=-1\\
&\iff 2zi\in\pi i+2\pi i\mathbb{Z} \text{ wegen Satz III.7.12}\\
&\iff z\in\frac{\pi}{2}+\pi\mathbb{Z}\\
\sin(z)=0 &\iff e^{iz}=e^{-iz}\\
&\iff e^{2iz}=1\\
&\iff 2iz\in2\pi i\mathbb{Z} \text{ wegen Satz III.7.12}\\
&\iff z\in\pi\mathbb{Z}.
\end{aligned}$$

ad (3): Im Beweis von Satz III.7.11 wurde gezeigt

$$\sin(x)>0\quad\forall x\in(0,1).$$

Damit folgt aus Teil (2) und Satz III.5.5 (S. 90)

$$\sin(x)>0\quad\forall x\in(0,\pi).$$

Mit dem gleichen Argument folgt aus $\cos(0) = 1$

$$\cos(x) > 0 \quad \forall x \in (0, \frac{\pi}{2}).$$

Sei nun $0 < x < y < \frac{\pi}{2}$. Dann folgt aus Satz III.7.3

$$\sin(y) - \sin(x) = 2\cos(\frac{x+y}{2})\sin(\frac{x-y}{2}) > 0.$$

AD (4): Folgt aus Satz III.7.12 und Satz III.7.3.
AD (5): Folgt aus (2), (3) und Satz III.7.3.
AD (6): Aus Satz III.7.10 folgt $\cos(\mathbb{R}) = [-1, 1]$. Aus (5) folgt dann $\sin(\mathbb{R}) = [-1, 1]$. □

BEMERKUNG III.7.15. (1) Die reellen Funktionen cos und sin sind durch die Werte von $\sin(x)$ für $x \in [0, \frac{\pi}{2}]$ eindeutig festgelegt.
(2) Die Zahl $\pi$ ist irrational, ja sie ist sogar transzendent (Beweis von LINDEMANN 1882), d.h., es gibt kein Polynom $p$ mit rationalen Koeffizienten und $p(\pi) = 0$.
(3) Die Zahl $\frac{\pi}{2}$ ist die kleinste positive Nullstelle des Cosinus. Um sie numerisch zu berechnen, benutzen wir die Identität

$$\cos(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} + r_{2n+2}(x)$$

mit

$$r_{2n+2}(x) = \sum_{k=n+1}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$$

und die (zu beweisende!) Abschätzung

$$|r_{2n+2}(x)| \le \frac{|x|^{2n+2}}{(2n+2)!} \quad \forall |x| \le 2n+3.$$

Für $0 < x \le 5$ folgt hieraus

$$\begin{aligned} \cos(x) &= 1 - \frac{x^2}{2} + r_4 \\ &\le 1 - \frac{x^2}{2} + \frac{x^4}{24} \\ &= 1 - \frac{x^2}{2}(1 - \frac{x^2}{12}) \\ &< 1 \qquad\qquad \forall\, 0 < x < \sqrt{12}. \end{aligned}$$

Also ist $2\pi \ge \sqrt{12}$ und damit $\pi \ge \sqrt{3} > 1.5$. Für $x = 1.5$ und $x = 1.6$ erhalten wir

$$\begin{aligned} \cos(1.5) &= 0.070'737'201'63 \pm 20 \cdot 10^{-11} > 0 \\ \cos(1.6) &= -0.029'199'522'39 \pm 20 \cdot 10^{-11} < 0. \end{aligned}$$

Also ist $1.5 < \frac{\pi}{2} < 1.6$. Sei

$$a = 1.5 + 0.1\frac{\cos(1.5)}{\cos(1.5) - \cos(1.6)} = 1.57078\ldots \quad .$$

$a$ ist der Schnittpunkt der Geraden durch $(1.5, \cos(1.5))$ und $(1.6, \cos(1.6))$ mit der $x$-Achse. Wir erhalten

$$\begin{aligned} \cos(1.5707) &= 0.000'096'326'73 \pm 20 \cdot 10^{-11} > 0 \\ \cos(1.5708) &= -0.000'003'673'26 \pm 20 \cdot 10^{-11} < 0. \end{aligned}$$

Wegen

$$\begin{aligned} \cos y - \cos x &= -2\sin(\frac{y+x}{2})\sin(\frac{x-y}{2}) \\ &< 0 \qquad\qquad\qquad\qquad \forall\, 0 < x < y < \pi \end{aligned}$$

ist der Cosinus auf $(0, \pi)$ streng monoton fallend. Daher folgt aus obiger Abschätzung

$$1.5707 < \frac{\pi}{2} < 1.5708.$$

Sei

$$\begin{aligned} b &= 1.5707 + 0.0001\frac{\cos(1.5707)}{\cos(1.5707) - \cos(1.5708)} \\ &= 1.570'796'326'\ldots \quad . \end{aligned}$$

Wir erhalten

$$\begin{aligned} \cos(1.570'796'326) &= 0.000'000'000'73 \pm 20 \cdot 10^{-11} > 0 \\ \cos(1.570'796'327) &= -0.000'000'000'27 \pm 20 \cdot 10^{-11} < 0. \end{aligned}$$

Also gilt

$$1.570'796'326 < \frac{\pi}{2} < 1.507'796'327$$

d.h.

$$\pi = 3.141'592'653 \pm 10^{-9}.$$

Satz III.7.16. *Für jedes $a \in \mathbb{R}$ gilt*

$$\begin{aligned} \exp : \mathbb{R} \times [a, a + 2\pi)i &\to \mathbb{C}^* \\ \exp : \mathbb{R} \times (a, a + 2\pi]i &\to \mathbb{C}^* \end{aligned}$$

*ist bijektiv.*

Beweis. Injektiv: Für $z = x + iy$, $w = u + iv$ mit $x, y, u, v \in \mathbb{R}$ folgt

$$\begin{aligned} e^z = e^w &\Longleftrightarrow e^x e^{iy} = e^u e^{iv} \\ &\Longrightarrow |e^x e^{iy}| = |e^u e^{iv}| \\ &\Longrightarrow e^x = e^u \Longrightarrow x = u \end{aligned}$$

$$\begin{aligned}
&\Longrightarrow e^{iy} = e^{iv} \\
&\Longrightarrow e^{i(y-v)} = 1 \\
&\Longrightarrow y - v \in 2\pi\mathbb{Z} \\
&\Longrightarrow y = v.
\end{aligned}$$

SURJEKTIV: Sei $z \in \mathbb{C}^*$. Dann ist $|z| > 0$ und $\frac{z}{|z|} \in S^1$. Also gibt es nach Satz III.7.6 und Satz III.7.10 ein $x \in \mathbb{R}$ und ein $y \in [a, a+2\pi)$ bzw. $y \in (a, a+2\pi]$ mit

$$e^x = |z|, \ e^{iy} = \frac{z}{|z|}.$$

Hieraus folgt die Behauptung. □

Hieraus folgt unmittelbar die POLARKOORDINATENDARSTELLUNG der komplexen Zahlen.

SATZ III.7.17 (POLARKOORDINATENDARSTELLUNG DER KOMPLEXEN ZAHLEN). *Zu jedem $z \in \mathbb{C}^*$ gibt es genau ein $\alpha \in [0, 2\pi)$ mit*

$$z = |z|e^{i\alpha}.$$

*$\alpha$ heißt das* ARGUMENT *von $z$, kurz $\alpha = \arg(z)$.*

BEMERKUNG III.7.18. Seien $z, u \in \mathbb{C}^*$ und

$$z = re^{i\varphi}, \ u = se^{i\psi}$$

mit $r, s \in \mathbb{R}_+^*$, $\varphi, \psi \in [0, 2\pi)$. Dann folgt

$$z \cdot u = (r \cdot s)e^{i(\varphi+\psi)}.$$

Also

$$\arg(z \cdot u) = \arg(z) + \arg(u) \mod 2\pi.$$

Zwei komplexe Zahlen werden also miteinander multipliziert, indem ihre Beträge multipliziert und ihre Argumente modulo $2\pi$ addiert werden.

SATZ III.7.19 (KOMPLEXE WURZELN). *Für jedes $a \in \mathbb{C}^*$ und jedes $n \in \mathbb{N}^*$ besitzt die Gleichung $z^n = a$ genau die $n$ komplexen Lösungen*

$$z_k = |a|^{1/n} e^{i(\arg(a)+2\pi k)/n} \quad 0 \le k \le n-1.$$

BEWEIS. Sei $z$ eine Lösung der Gleichung $z^n = a$. Dann ist $z \in \mathbb{C}^*$ und somit

$$z = re^{i\varphi} \ \text{ mit } r \in \mathbb{R}_+^* \ \text{ und } \varphi \in [0, 2\pi).$$

Es folgt

$$\begin{aligned}
&z^n = r^n e^{in\varphi} = a = |a|e^{i\arg(a)} \\
\Longleftrightarrow\ &r^n = |a| \\
&\Longleftrightarrow \ r = |a|^{1/n}
\end{aligned}$$

und

$$\begin{aligned}
&n\varphi - \arg(a) \in 2\pi\mathbb{Z} \\
\iff\ &\varphi = \frac{1}{n}(\arg(a) + 2k\pi) \quad 0 \le k \le n-1.
\end{aligned}$$

□

Zum Abschluss beweisen wir noch eine andere Darstellung der Exponentialfunktion.

Satz III.7.20. *Für alle $z \in \mathbb{C}$ ist*

$$e^z = \lim_{n\to\infty} (1 + \frac{z}{n})^n.$$

Beweis. Für $n \in \mathbb{N}^*$ und $z \in \mathbb{C}$ gilt

$$\begin{aligned}
e^z - (1 + \frac{z}{n})^n &= (e^{\frac{z}{n}})^n - (1 + \frac{z}{n})^n \\
&= [e^{\frac{z}{n}} - (1 + \frac{z}{n})] \sum_{k=0}^{n-1} e^{\frac{z}{n}k}(1 + \frac{z}{n})^{n-1-k} \\
&= A \cdot B.
\end{aligned}$$

Weiter ist

$$\begin{aligned}
A &= e^{\frac{z}{n}} - (1 + \frac{z}{n}) \\
&= \frac{z}{n}[\frac{e^{\frac{z}{n}} - 1}{\frac{z}{n}} - 1] \\
&= \frac{z}{n} r_n(z).
\end{aligned}$$

Gemäß Beispiel III.3.14(2) (S. 79) gilt

$$\lim_{n\to\infty} r_n(z) = 0.$$

Weiter gilt

$$\begin{aligned}
|e^{\frac{z}{n}}| &\le e^{\frac{|z|}{n}} \\
|1 + \frac{z}{n}| &\le 1 + \frac{|z|}{n} \\
&\le e^{\frac{|z|}{n}}.
\end{aligned}$$

Damit folgt

$$\begin{aligned}
|B| &\le \sum_{k=0}^{n-1} e^{\frac{|z|}{n}k} |1 + \frac{z}{n}|^{n-1-k} \\
&\le \sum_{k=0}^{n-1} e^{\frac{|z|}{n}(k+n-1-k)} \\
&\le n e^{|z|\frac{n-1}{n}}
\end{aligned}$$

$$\leq ne^{|z|}.$$

Insgesamt erhalten wir

$$\begin{aligned}
|e^{\frac{z}{n}} - (1 + \frac{z}{n})^n| &\leq \frac{|z|}{n}|r_n(z)|ne^{|z|} \\
&= |z|e^{|z|}|r_n(z)| \\
&\underset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

$\square$

KAPITEL IV

# Differentialrechnung einer Veränderlichen

Zunächst führen wir den Begriff der Differenzierbarkeit einer Funktion von $M \subset \mathbb{K}$ in einen normierten Vektorraum ein und betrachten darauf aufbauende Begriffe. Danach leiten wir Mittelwertsätze her und befassen uns mit Konsequenzen wie z.B. der Regel von de l'Hôpital. Anschließend beweisen wir Taylorformeln und ziehen daraus Konsequenzen.
Als Anwendung betrachten wir die numerische Lösung von Gleichungen. Wir beweisen den Banachschen Fixpunktsatz, leiten das Newtonverfahren her und beweisen seine Konvergenz.

## IV.1. Differenzierbarkeit

Im Folgenden sei, sofern nicht anders vermerkt, $(Y, \|\cdot\|_Y)$ ein normierter Raum, $M \subset \mathbb{K}$ eine nicht leere Menge, $x_0 \in M$ ein Häufungspunkt von $M$ und $f : M \to Y$ eine Funktion.

Definition IV.1.1. Die Funktion $f : M \to Y$ heißt im Punkt $x_0 \in M$ differenzierbar, wenn der Grenzwert

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

in $Y$ existiert. Falls $f$ in $x_0$ differenzierbar ist, nennen wir

$$\begin{aligned} f'(x_0) = Df(x_0) &= \frac{df}{dx}(x_0) \\ &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \\ &= \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \in Y \end{aligned}$$

die Ableitung von $f$ im Punkte $x_0$.

Bevor wir Eigenschaften differenzierbarer Funktionen untersuchen einige Beispiele.

Beispiel IV.1.2. (1) $f : \mathbb{C} \to \mathbb{C}$, $f(z) = z^n$, $n \in \mathbb{N}^*$ ist differenzierbar für jedes $z \in \mathbb{C}$ und $f'(z) = nz^{n-1}$.
(2) $f : \mathbb{C} \to \mathbb{C}$, $f(z) = e^z$ ist differenzierbar für jedes $z \in \mathbb{C}$ und $f'(z) = e^z$.
(3) $f : \mathbb{R}^* \to \mathbb{R}$, $f(x) = \frac{1}{x}$ ist differenzierbar für jedes $x \in \mathbb{R}^*$ und $f'(x) = -\frac{1}{x^2}$.

(4) $f : \mathbb{R} \to \mathbb{R}^2$, $f(x) = (x, x^2)^t$ ist differenzierbar für jedes $x \in \mathbb{R}$ und $f'(x) = (1, 2x)^t$.
(5) $f : \mathbb{R} \to \mathbb{R}_+$, $f(x) = |x|$ ist in $x_0 = 0$ nicht differenzierbar.
(6) $f : \mathbb{C} \to \mathbb{R}$, $f(z) = \operatorname{Re} z$ ist in keinem Punkt $z \in \mathbb{C}$ differenzierbar.

Beweis. ad (1):

$$\begin{aligned} \frac{z^n - z_0^n}{z - z_0} &= \sum_{k=0}^{n-1} z^k z_0^{n-k-1} \\ &\underset{z \to z_0}{\longrightarrow} \sum_{k=0}^{n-1} z_0^k z_0^{n-k-1} = n z_0^{n-1}. \end{aligned}$$

ad (2):

$$\frac{e^z - e^{z_0}}{z - z_0} = e^{z_0} \frac{e^{z - z_0} - 1}{z - z_0} \underset{z \to z_0}{\longrightarrow} e^{z_0}.$$

ad (3):

$$\frac{\frac{1}{x_0 + h} - \frac{1}{x_0}}{h} = \frac{-h}{h x_0 (x_0 + h)} = -\frac{1}{x_0(x_0 + h)} \underset{h \to 0}{\longrightarrow} -\frac{1}{x_0^2}.$$

ad (4):

$$\begin{aligned} \left\| \frac{f(x_0 + h) - f(x_0)}{h} - (1, 2x_0)^t \right\| &= \|(1, 2x_0 + h)^t - (1, 2x_0)^t\| \\ &= \sqrt{(1-1)^2 + (2x_0 + h - 2x_0)^2} \\ &= |h| \\ &\underset{h \to 0}{\longrightarrow} 0 \end{aligned}$$

ad (5):

$$\begin{aligned} \lim_{n \to \infty} \frac{f(\frac{1}{n}) - f(0)}{\frac{1}{n}} &= \lim_{n \to \infty} \frac{\frac{1}{n}}{\frac{1}{n}} = 1 \\ \lim_{n \to \infty} \frac{f(-\frac{1}{n}) - f(0)}{-\frac{1}{n}} &= \lim_{n \to \infty} \frac{\frac{1}{n}}{-\frac{1}{n}} = -1. \end{aligned}$$

Für $x \in \mathbb{R}^*$ folgt analog, dass $f$ differenzierbar ist mit

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

ad (6): Sei $z = x + iy$ mit $x, y \in \mathbb{R}$. Dann ist

$$\begin{aligned} \frac{f(z + \frac{1}{n}) - f(z)}{\frac{1}{n}} &= \frac{\frac{1}{n}}{\frac{1}{n}} = 1 \quad \forall\, n \in \mathbb{N}^* \\ \frac{f(z + \frac{i}{n}) - f(z)}{\frac{i}{n}} &= 0 \qquad\quad \forall\, n \in \mathbb{N}^*. \end{aligned}$$

□

Satz IV.1.3. *Folgende Aussagen sind äquivalent.*

(1) *$f$ ist differenzierbar in $x_0$.*
(2) *Es gibt ein $c \in Y$ mit*

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

(3) *Es gibt ein $c \in Y$ und eine Abbildung $r : M \to Y$, die in $x_0$ stetig ist und $r(x_0) = 0$ erfüllt, mit*

$$f(x) = f(x_0) + c(x - x_0) + r(x)(x - x_0) \quad \forall x \in M.$$

*In den Fällen (2) und (3) ist $c$ eindeutig bestimmt und $c = f'(x_0)$.*

Beweis. (1) $\Longrightarrow$ (2): Ist offensichtlich.
(2) $\Longrightarrow$ (3): Definiere $r : M \to Y$ durch

$$r(x) = \begin{cases} \frac{f(x) - f(x_0) - c(x - x_0)}{x - x_0} & x \neq x_0 \\ 0 & x = x_0. \end{cases}$$

Aus (2) folgt dann, dass $r$ in $x_0$ stetig ist.
(3) $\Longrightarrow$ (1): Ist offensichtlich. □

Aus Satz IV.1.3 folgt unmittelbar:

Satz IV.1.4. *Ist $f$ in $x_0$ differenzierbar, so ist $f$ in $x_0$ stetig.*

Beweis. Gemäß Satz IV.1.3 (3) ist

$$f(x) = f(x_0) + c(x - x_0) + r(x)(x - x_0)$$

mit einer in $x_0$ stetigen Funktion $r$. Also ist

$$\lim_{x \to x_0} f(x) = f(x_0).$$

□

Bemerkung IV.1.5. (1) Beispiele IV.1.2 (5) und (6) zeigen, dass die Umkehrung von Satz IV.1.4 i. a. falsch ist.
(2) Eine Funktion der Form $x \to f(x_0) + f'(x_0)(x - x_0)$ ist affin. Mithin besagt Satz IV.1.3, dass eine in $x_0$ differenzierbare Funktion $f$ in der Nähe von $x_0$ durch eine affine Funktion bis auf einen Fehler $R(x)$ mit

$$(*) \qquad \lim_{x \to x_0} \frac{R(x)}{x - x_0} = 0,$$

d.h. einen Fehler höherer Ordnung, approximiert werden kann. Sei nun umgekehrt $h : M \to Y$ eine affine Funktion, $R : M \to Y$ eine Funktion, die $(*)$ erfüllt, und

$$f(x) = h(x) + R(x) \quad \forall x \in M.$$

Da $h$ affin ist, gilt für alle $x \in M$

$$h(x) = h(0) + cx$$

mit $c \in Y$. Daher gilt

$$\begin{aligned} f(x) - f(x_0) &= h(x) + R(x) - h(x_0) - \underbrace{R(x_0)}_{=0} \\ &= h(0) + cx + R(x) - h(0) - cx_0 \\ &= c(x - x_0) + R(x) \end{aligned}$$

$$\Longrightarrow \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = c.$$

Also ist $f$ in $x_0$ differenzierbar.
Mithin ist also eine Funktion genau dann in $x_0$ differenzierbar, wenn sie dort linear approximierbar, d.h. durch eine affine Funktion bis auf Terme höherer Ordnung darstellbar ist. Geometrisch bedeutet dies, dass der Graph von $f$ durch eine Gerade approximiert werden kann.
(3) Ist $f : M \subset \mathbb{C} \to \mathbb{C}$ in $x_0 \in M \cap \mathbb{R}$ differenzierbar, ist weiter $x_0$ ein Häufungspunkt von $M \cap \mathbb{R}$ und ist $f(M \cap \mathbb{R}) \subset \mathbb{R}$, so ist $f|_{M \cap \mathbb{R}}$ als Funktion von $\mathbb{R}$ in $\mathbb{R}$ in $x_0$ differenzierbar. Beispiel IV.1.2 (6) zeigt, dass die Umkehrung i. a. falsch ist.
(4) Ist $f : M \subset \mathbb{R} \to \mathbb{R}^2$ in $x_0$ differenzierbar, so ist $f$ aufgefasst als Abbildung von $\mathbb{R}$ in $\mathbb{C}$ auch in $x_0$ differenzierbar und die Ableitungen stimmen unter Berücksichtigung der Konvention $\mathbb{R}^2 \ni (x, y) \sim x + iy \in \mathbb{C}$ überein.

Satz IV.1.6. *$f = (f_1, \ldots, f_m)^t : M \to \mathbb{K}^m$ ist genau dann in $x_0$ differenzierbar, wenn jede Komponentenfunktion $f_j$, $1 \le j \le m$, in $x_0$ differenzierbar ist. In diesem Fall gilt*

$$f'(x_0) = (f_1'(x_0), \ldots, f_m'(x_0))^t.$$

Beweis. Wegen

$$\frac{f(x) - f(x_0)}{x - x_0} = \left( \frac{f_1(x) - f_1(x_0)}{x - x_0}, \ldots, \frac{f_m(x) - f_m(x_0)}{x - x_0} \right)^t$$

folgt die Behauptung aus Satz III.3.10 (S. 78) und Bemerkung III.3.13 (S. 79). □

Satz IV.1.7. *Seien $f : M \to Y$, $g : M \to Y$ in $x_0$ differenzierbar und $\alpha, \beta \in \mathbb{K}$. Dann gilt:*

(1) *$\alpha f + \beta g$ ist in $x_0$ differenzierbar und*

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0),$$

*d.h., die in $x_0$ differenzierbaren Funktionen bilden einen $\mathbb{K}$-Vektorraum.*

(2) *Ist $Y = \mathbb{K}$, so ist auch $f \cdot g$ in $x_0$ differenzierbar und*

$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0).$$

(3) *Ist* $Y = \mathbb{K}$ *und* $g(x_0) \neq 0$ *so ist auch* $\frac{f}{g}$ *in* $x_0$ *differenzierbar und*

$$(\frac{f}{g})'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

BEWEIS. AD (1): Ist offensichtlich.
AD (2):

$$\begin{aligned}&\frac{f(x)\cdot g(x) - f(x_0)\cdot g(x_0)}{x-x_0}\\ =&\frac{f(x)-f(x_0)}{x-x_0}\cdot g(x) + f(x_0)\frac{g(x)-g(x_0)}{x-x_0}\\ \xrightarrow[x\to x_0]{}& f'(x_0)g(x_0) + f(x_0)g'(x_0) \quad \text{wg. Satz IV.1.4}\end{aligned}$$

AD (3): Wegen Satz IV.1.4 gibt es eine Umgebung $U \in \mathcal{U}(x_0)$ mit $g(x) \neq 0$ für alle $x \in U$. Für $x \in U$ folgt

$$\begin{aligned}&\frac{\frac{f(x)}{g(x)} - \frac{f(x_0)}{g(x_0)}}{x-x_0}\\ =&\frac{f(x)g(x_0) - f(x_0)g(x)}{g(x)g(x_0)(x-x_0)}\\ =&\frac{1}{g(x)g(x_0)}\left[\frac{f(x)-f(x_0)}{x-x_0}g(x_0) - f(x_0)\frac{g(x)-g(x_0)}{x-x_0}\right]\\ \xrightarrow[x\to x_0]{}&\frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.\end{aligned}$$

□

SATZ IV.1.8 (KETTENREGEL). *$f : M \to \mathbb{K}$ sei in $x_0$ differenzierbar, $y_0 = f(x_0)$ sei ein Häufungspunkt von $N \subset \mathbb{K}$ und $g : N \to Y$ sei differenzierbar in $y_0$. Dann ist $g \circ f$ in $x_0$ differenzierbar und es gilt*

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

BEWEIS. Es ist gemäß Satz IV.1.3

$$\begin{aligned}f(x) &= f(x_0) + f'(x_0)(x-x_0) + r(x)(x-x_0) \quad \forall x \in M\\ g(y) &= g(y_0) + g'(y_0)(y-y_0) + s(y)(y-y_0) \quad \forall y \in N\end{aligned}$$

mit in $x_0$ bzw. $y_0$ stetigen Funktionen $r$ und $s$ und $r(x_0) = 0$, $s(y_0) = 0$. Damit folgt

$$\begin{aligned}g \circ f(x) = g \circ f(x_0) &+ g'(f(x_0))(f(x) - f(x_0))\\ &+ s \circ f(x)(f(x) - f(x_0))\\ =\ & g \circ f(x_0) + g'(f(x_0)) \cdot f'(x_0)(x-x_0)\\ &+ [g'(f(x_0))r(x)\\ &\quad + s \circ f(x) \cdot f'(x_0) + s \circ f(x)r(x)](x-x_0).\end{aligned}$$

Da die Funktion $\varphi : M \to Y$ mit

$$\varphi(x) = g'(f(x_0))r(x) + s \circ f(x) \cdot f'(x_0) + s \circ f(x) \cdot r(x)$$

gemäß Satz III.3.5 (S. 77) und Satz III.3.8 (S. 77) in $x_0$ stetig ist und $\varphi(x_0) = 0$ erfüllt, folgt die Behauptung. □

Satz IV.1.9 (Differenzierbarkeit der Umkehrfunktion). *Sei $f : M \to \mathbb{K}$ in $x_0$ differenzierbar und injektiv. Weiter sei $f^{-1} : N \to \mathbb{K}$ mit $N = f(M)$ in $y_0 = f(x_0)$ stetig. Dann ist $f^{-1}$ in $y_0$ differenzierbar genau dann, wenn gilt $f'(x_0) \neq 0$. In diesem Fall ist*

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

Beweis. Gemäß Satz IV.1.4 ist $f$ stetig in $x_0$. Sei $(x_n)_{n\in\mathbb{N}} \subset M$ eine Folge mit $x_n \neq x_0$ für alle $n \in \mathbb{N}$ und $x_n \underset{n\to\infty}{\longrightarrow} x_0$. Dann gilt $f(x_n) \neq f(x_0)$ für alle $n \in \mathbb{N}$ und $f(x_n) \underset{n\to\infty}{\longrightarrow} f(x_0)$. Also ist $y_0$ ein Häufungspunkt von $N$, so dass es sinnvoll ist, von Differenzierbarkeit in $y_0$ zu sprechen.
„$\Longrightarrow$“: Für alle $x \in M$ gilt

$$f^{-1} \circ f(x) = x.$$

Damit folgt aus Satz IV.1.8

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

Hieraus folgt $f'(x_0) \neq 0$ und $(f^{-1})'(y_0) = \frac{1}{f'(x_0)}$.
„$\Longleftarrow$“:

$$\begin{aligned}\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} &\underset{y=f(x)}{=} \frac{x - x_0}{f(x) - f(x_0)} \\ &= \frac{1}{\frac{f(x)-f(x_0)}{x-x_0}} \\ &\underset{\substack{y\to y_0 \\ (\Longleftrightarrow x\to x_0)}}{\longrightarrow} \frac{1}{f'(x_0)}.\end{aligned}$$

Dabei wurde ausgenutzt, dass wegen der Injektivität von $f$ gilt

$$y \neq y_0 \iff x \neq x_0,$$

und dass wegen der Stetigkeit von $f^{-1}$ in $y_0$ gilt

$$y \to y_0 \iff f^{-1}(y) \to f^{-1}(y_0) = x_0.$$

□

Bemerkung IV.1.10. Sei $M \neq \emptyset$, $M \neq \{x_0\}$, $M \subset \mathbb{R}$ ein Intervall, $f : M \to \mathbb{R}$ streng monoton, stetig, in $x_0$ differenzierbar und $f'(x_0) \neq 0$. Dann ist $f^{-1}$ in $y_0 = f(x_0)$ differenzierbar und $(f^{-1})'(y_0) = \frac{1}{f'(x_0)}$.

Beweis. Die Behauptung folgt aus Satz III.6.7 (S. 96) und Satz IV.1.9. □

Beispiel IV.1.11. $\ln : \mathbb{R}_+^* \to \mathbb{R}$ ist für jedes $y_0 \in \mathbb{R}_+^*$ differenzierbar und

$$\frac{d}{dx}\ln(y_0) = \frac{1}{y_0} \quad \forall y_0 \in \mathbb{R}_+^*.$$

Beweis. Wegen Satz III.7.4 (S. 100) und Beispiel IV.1.2 (2) sind die Voraussetzungen von Satz IV.1.9 erfüllt. Damit folgt

$$\begin{aligned}\frac{d}{dx}\ln(y_0) &= \frac{1}{\exp(x_0)} \quad (\exp(x_0) = y_0)\\ &= \frac{1}{y_0}.\end{aligned}$$

□

Definition IV.1.12. Eine Teilmenge $M$ eines normierten Vektorraumes heißt perfekt, wenn jeder Punkt von $M$ Häufungspunkt von $M$ ist.

Beispiel IV.1.13. (1) $M$ offen $\Longrightarrow M$ perfekt.
(2) $M \subset \mathbb{R}$ Intervall. $M \neq \emptyset, M \neq \{x\} \Longrightarrow M$ perfekt.
(3) $\{x\} \subset \mathbb{R}$ ist nicht perfekt.

Definition IV.1.14. Sei $M \subset \mathbb{K}$ perfekt.

(1) $f : M \to Y$ heißt differenzierbar (in $M$), wenn $f$ in jedem Punkt $x \in M$ differenzierbar ist.
(2) $f : M \to Y$ heißt stetig differenzierbar (in $M$), wenn $f$ differenzierbar und $f'$ stetig ist.
(3) $f^{(0)} = f$, $f^{(1)} = Df = f'$,
$f^{(n+1)} = D^{n+1}f = (f^{(n)})' \quad \forall n \in \mathbb{N}$, sofern definiert.
(4) $f$ heißt $n$-mal stetig differenzierbar, $n \in \mathbb{N}^*$, wenn $f^{(n)}$ existiert und stetig ist.
(5) $C^n(M,Y) = \{f \in C(M,Y) : f$ ist $n$-mal stetig differenzierbar$\}$, $n \in \mathbb{N}^*$,
$C^\infty(M,Y) = \bigcap\limits_{n\in\mathbb{N}} C^n(M,Y)$.

Bemerkung IV.1.15. (1) $C^n(M,Y)$ ist ein $\mathbb{K}$-Vektorraum.
(2) $C(M,Y) \underset{\neq}{\supset} C^1(M,Y) \underset{\neq}{\supset} C^2(M,Y) \cdots$.
(3) Falls $M$ perfekt und kompakt ist, ist

$$\|f\|_{C^n(M,Y)} = \max_{0\le k\le n} \|f^{(k)}\|_{C(M,Y)}$$

wohldefiniert und eine Norm auf $C^n(M,Y)$.

Beweis. ad (1): Folgt durch Induktion über $n$ aus Satz IV.1.7.
ad (2): $C^n(M,Y) \subset C^{n-1}(M,Y)$ für alle $n \in \mathbb{N}^*$ folgt durch Induktion aus Satz IV.1.3. $|x|x^{n-1}$ ist in $C^{n-1}(\mathbb{R},\mathbb{R})$ aber nicht in $C^n(\mathbb{R},\mathbb{R})$, $n \in \mathbb{N}^*$.
ad (3): Übungsaufgabe. □

Satz IV.1.16 (Leibnizregel). *Sei $n \in \mathbb{N}^*$ und $f, g \in C^n(M, \mathbb{K})$ mit $M \subset \mathbb{K}$ perfekt. Dann ist $f \cdot g \in C^n(M, \mathbb{K})$ und*

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

Beweis. Induktion über $n$.
$n = 1$ : Satz IV.1.7 (2).
$n \to n+1$ :

$$\begin{aligned}
(f \cdot g)^{(n+1)} &= \Big((f \cdot g)^{(n)}\Big)' \\
&= \Big(\sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}\Big)' \\
&= \sum_{k=0}^{n} \binom{n}{k} [f^{(k+1)} g^{(n-k)} + f^{(k)} g^{(n+1-k)}] \\
&= \sum_{l=1}^{n+1} \binom{n}{l-1} f^{(l)} g^{(n+1-l)} + \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n+1-k)} \\
&= f^{(n+1)} g^{(0)} + f^{(0)} g^{(n+1)} \\
&\quad + \sum_{k=1}^{n} [\binom{n}{k-1} + \binom{n}{k}] f^{(k)} g^{(n+1-k)} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} f^{(k)} g^{(n+1-k)}.
\end{aligned}$$

□

Beispiel IV.1.17. (1) $M \subset \mathbb{R}, f : M \to \mathbb{C}$. Dann ist $f = g + ih$ mit $g, h : M \to \mathbb{R}$. Es gilt: $f$ differenzierbar in $x_0$ $\iff$ $g$ und $h$ differenzierbar in $x_0$.

$$f'(x_0) = g'(x_0) + ih'(x_0).$$

(2) $p = \sum_{k=0}^{n} a_k x^k$ mit $a_k \in \mathbb{K}$. Dann ist $p \in C^\infty(\mathbb{K}, \mathbb{K})$ mit

$$p' = \sum_{k=0}^{n} k a_k x^{k-1} = \sum_{l=0}^{n-1} a_{l+1}(l+1)x^l$$

und

$$p^{(m)} = 0 \quad \forall m > n.$$

(3) $p = \sum_{k=0}^{n} a_k x^k$, $q = \sum_{k=0}^{m} b_k x^k$ mit $a_k, b_k \in \mathbb{K}$ und $q(x) \neq 0$ für alle

$x \in M$, $M$ perfekt. Dann ist $\frac{p}{q} \in C^\infty(M, \mathbb{K})$.
(4) $\exp \in C^\infty(\mathbb{K}, \mathbb{K}), \exp' = \exp$.
(5) $\sin, \cos \in C^\infty(\mathbb{K}, \mathbb{K})$

$$\cos' = -\sin\ ,\ \sin' = \cos\,.$$

(6) $\ln \in C^\infty(\mathbb{R}^*_+, \mathbb{R})$,

$$\ln^{(k)}(x) = \frac{(-1)^{k-1}(k-1)!}{x^k}, k \in \mathbb{N}^*.$$

(7) $a > 0$. Dann ist $x \mapsto a^x \in C^\infty(\mathbb{R}, \mathbb{R})$

$$(a^x)' = a^x \cdot \ln(a).$$

(8)

$$f(x) = \begin{cases} x^2 \sin(\frac{1}{x}) & , x \in \mathbb{R}^* \\ 0 & , x = 0 \end{cases}$$

ist auf $\mathbb{R}$ differenzierbar, aber $f'$ ist nicht stetig.

Beweis. ad (1): Folgt aus Satz IV.1.6 und Bemerkung IV.1.5 (4).
ad (2): Folgt aus Satz IV.1.7 und Beispiel IV.1.2 (1).
ad (3): Folgt aus (2) und Satz IV.1.7.
ad (4): Beispiel IV.1.2 (2).
ad (5): Folgt aus (4) und Satz III.7.2(7) (S. 98).
ad (6): Folgt aus Beispiel IV.1.11 und Satz IV.1.7.
ad (7): Folgt aus Satz IV.1.8 und $a^x = \exp(x \ln(a))$.
ad (8): Aus

$$|\frac{f(x)}{x}| = |x||\sin(\frac{1}{x})| \leq |x|$$

folgt die Differenzierbarkeit in 0 und $f'(0) = 0$. Für $x \neq 0$ folgt aus Satz IV.1.7, Satz IV.1.8 und Teil (5)

$$f'(x) = 2x \sin(\frac{1}{x}) - \cos(\frac{1}{x}).$$

Damit folgt

$$f'(\frac{1}{2n\pi}) = \frac{1}{n\pi}\sin(2n\pi) - \cos(2n\pi) = -1.$$

Also ist $f'$ in 0 nicht stetig. □

Definition IV.1.18. Sei $M \subset \mathbb{R}$ und $x_0$ ein Häufungspunkt von $M$. Dann heißt $f$ in $x_0$ linksseitig bzw. rechtsseitig differenzierbar, falls der Grenzwert

$$D_- f(x_0) = \lim_{x \to x_0 - 0} \frac{f(x) - f(x_0)}{x - x_0}$$

bzw.

$$D_+f(x_0) = \lim_{x \to x_0+0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert.

Beispiel IV.1.19. Sei $f(x) = |x|$. Dann ist

$$D_-f(0) = -1, \; D_+f(0) = 1.$$

Satz IV.1.20. *Sei $M \subset \mathbb{R}$. Dann ist $f : M \to Y$ genau dann in $x_0 \in M$ differenzierbar, wenn es links- und rechtsseitig differenzierbar ist und*

$$D_+f(x_0) = D_-f(x_0)$$

*ist.*

Beweis. „$\Longrightarrow$:“ Ist offensichtlich.
„$\Longleftarrow$:“ Definiere $\varphi : M \to Y$ durch

$$\varphi(x) = \begin{cases} \frac{f(x)-f(x_0)}{x-x_0} & , x \neq x_0, \\ D_+f(x_0) & , x = x_0. \end{cases}$$

Wegen $D_+f(x_0) = D_-f(x_0)$ ist $\varphi$ in $x_0$ stetig. Damit folgt die Behauptung aus der Definition der Differenzierbarkeit. □

Beispiel IV.1.21. Die Funktion

$$f(x) = \begin{cases} 0 & \text{für } x \leq 0, \\ e^{-\frac{1}{x}} & \text{für } x > 0 \end{cases}$$

ist in $C^\infty(\mathbb{R}, \mathbb{R})$ und es gilt

$$f^{(n)}(0) = 0 \quad \forall n \in \mathbb{N}.$$

Beweis. Offensichtlich ist $f \in C^\infty((-\infty, 0], \mathbb{R})$ und

$$f^{(n)}(x) = 0 \quad \forall n \in \mathbb{N}, x \leq 0.$$

Beh.: $\forall n \in \mathbb{N} \; \exists \; p_{2n}$ Polynom vom Grad $\leq 2n$ mit:

$$f^{(n)}(x) = p_{2n}(\frac{1}{x})e^{-\frac{1}{x}} \quad \forall x > 0.$$

Bew. der Beh.: Induktion über $n$.
$n = 0 : p_0(y) = 1$.
$n \to n+1 :$

$$\begin{aligned} [p_{2n}(\frac{1}{x})e^{-\frac{1}{x}}]' &= -\frac{1}{x^2}p'_{2n}(\frac{1}{x})e^{-\frac{1}{x}} + \frac{1}{x^2}p_{2n}(\frac{1}{x})e^{-\frac{1}{x}} \\ &= p_{2n+2}(\frac{1}{x})e^{-\frac{1}{x}} \end{aligned}$$

mit

$$p_{2n+2}(y) = y^2p_{2n}(y) - y^2p'_{2n}(y).$$

Aus Beh. folgt $f \in C^\infty(\mathbb{R}_+^*, \mathbb{R})$.
Wegen Satz IV.1.20 müssen wir noch zeigen, dass für $n \in \mathbb{N}$ gilt

$$\lim_{x \to 0+} f^{(n)}(x) = 0.$$

Wegen Beh. und Satz III.7.4 (S. 100) gilt aber

$$\begin{aligned} \lim_{x \to 0+} f^{(n)}(x) &= \lim_{x \to 0+} p_{2n}(\frac{1}{x}) e^{-\frac{1}{x}} \\ &= \lim_{\substack{y \to +\infty \\ y = \frac{1}{x}}} p_{2n}(y) e^{-y} \\ &= \sum_{k=0}^{2n} a_{2n,k} \lim_{y \to \infty} \frac{y^k}{e^y} \qquad (\text{mit } p_{2n} = \sum a_{2n,k} x^k) \\ &= 0. \end{aligned}$$

□

## IV.2. Mittelwertsätze

Definition IV.2.1. Sei $f : M \to \mathbb{R}$ eine Funktion der nicht leeren Teilmenge $M$ des normierten Vektorraumes $(X, \|\cdot\|_X)$ in die reellen Zahlen. Dann hat $f$ in $x_0 \in M$ ein lokales Minimum bzw. lokales Maximum, wenn es eine Umgebung $U \in \mathcal{U}(x_0)$ gibt mit

$$f(x) \geq f(x_0) \quad \forall x \in U \cap M$$

bzw.

$$f(x) \leq f(x_0) \quad \forall x \in U \cap M.$$

Ein lokales Minimum oder Maximum nennen wir auch lokales Extremum.

Satz IV.2.2. *Sei $M \subset \mathbb{R}$ und $x_0 \in \overset{\circ}{M}$. Die Funktion $f : M \to \mathbb{R}$ habe in $x_0$ ein lokales Extremum und sei in $x_0$ differenzierbar. Dann gilt*

$$f'(x_0) = 0.$$

Beweis. O.E. habe $f$ in $x_0$ ein lokales Minimum. Sonst gehen wir zu $-f$ über. Dann gibt es ein $\varepsilon > 0$ mit

$$f(x) \geq f(x_0) \quad \forall x \in (x_0 - \varepsilon, x_0 + \varepsilon).$$

Daher ist

$$D_+ f(x_0) = \lim_{x \to x_0 + 0} \frac{f(x) - f(x_0)}{x - x_0} \geq 0$$

und

$$D_- f(x_0) = \lim_{x \to x_0 - 0} \frac{f(x) - f(x_0)}{x - x_0} \leq 0.$$

Damit folgt die Beh. aus Satz IV.1.20 (S. 120). □

Im Weiteren setzen wir, soweit nicht anders gesagt, voraus, dass

$$\text{(V)} \qquad f \in C([a,b],\mathbb{R}) \ \text{ in } (a,b) \ \text{ differenzierbar}$$

ist mit $-\infty < a < b < +\infty$.

Bemerkung IV.2.3. (1) Falls $(V)$ gilt, ist

$$\max_{a\le x\le b} f(x) = \max\{f(a), f(b), f(x) : f'(x) = 0, x \in (a,b)\}$$
$$\min_{a\le x\le b} f(x) = \min\{f(a), f(b), f(x) : f'(x) = 0, x \in (a,b)\}.$$

(2) Falls $x_0 \notin \overset{\circ}{M}$, ist, ist Satz IV.2.2 i. a. falsch, wie das Beispiel $f(x) = x$, $M = [0,1]$, $x_0 = 1$ zeigt.

Der folgende, fundamentale Satz geht auf M. Rolle (1652-1719) zurück.

Satz IV.2.4 (Satz von Rolle). *Es gelte $(V)$ und $f(a) = f(b)$. Dann gibt es ein $\zeta \in (a,b)$ mit*

$$f'(\zeta) = 0.$$

Beweis. O.E. ist $f$ nicht konstant. Gemäß Satz III.4.10 (S. 86) nimmt $f$ sein Minimum $m$ und sein Maximum $M$ in einem Punkt $x_m \in [a,b]$, bzw. $x_M \in [a,b]$ an. Da $f$ nicht konstant ist und $f(a) = f(b)$ gilt, ist mindestens einer der beiden Punkte ein innerer Punkt von $[a,b]$. Damit folgt die Behauptung aus Satz IV.2.2. □

Der folgende Satz ist für das Weitere zentral.

Satz IV.2.5 (Mittelwertsatz). *Es gelte $(V)$. Dann gibt es ein $\zeta \in (a,b)$ mit*

$$f(b) = f(a) + f'(\zeta)(b-a).$$

Beweis. Definiere $g \in C([a,b],\mathbb{R})$ durch

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

Dann ist $g$ in $(a,b)$ differenzierbar und erfüllt

$$g(a) = f(a) = g(b).$$

Damit folgt die Behauptung aus Satz IV.2.4. □

Satz IV.2.6. *Sei $I \subset \mathbb{R}$ ein perfektes Intervall, $f \in C(I,\mathbb{R})$ in $\overset{\circ}{I}$ differenzierbar und $f'(x) = 0$ für alle $x \in \overset{\circ}{I}$. Dann ist $f$ auf $I$ konstant.*

Beweis. Seien $a, b \in I$ mit $a < b$. Dann ist $(a,b) \subset \overset{\circ}{I}$. Damit folgt die Behauptung aus Satz IV.2.5. □

Satz IV.2.7. *Sei $I \subset \mathbb{R}$ ein perfektes Intervall und $f \in C(I, \mathbb{R})$ in $\overset{\circ}{I}$ differenzierbar. Dann gilt*

$$
\begin{aligned}
&(1) && f \text{ ist monoton fallend (wachsend)} \\
& && \iff f'(x) \le 0 \quad (\text{bzw. } f'(x) \ge 0) \quad \forall x \in \overset{\circ}{I}, \\
&(2) && f'(x) < 0 \quad (\text{bzw. } f'(x) > 0) \quad \forall x \in \overset{\circ}{I} \\
& && \Longrightarrow \text{ist streng monoton fallend (wachsend).}
\end{aligned}
$$

Beweis. ad (1): „$\Longrightarrow$:" O.E. ist $f$ monoton fallend, sonst gehe zu $-f$ über. Sei $x_0 \in \overset{\circ}{I}$ und $x \in I$. Dann gilt

$$
\begin{aligned}
x > x_0 &\Longrightarrow \frac{f(x) - f(x_0)}{x - x_0} \le 0 \\
&\Longrightarrow D_+ f(x_0) \le 0 \\
x < x_0 &\Longrightarrow \frac{f(x) - f(x_0)}{x - x_0} \le 0 \\
&\Longrightarrow D_- f(x_0) \le 0.
\end{aligned}
$$

Hieraus folgt die Behauptung.
„$\Longleftarrow$:" Seien $x, y \in I$ mit $x < y$. Dann ist $(x, y) \subset \overset{\circ}{I}$ und wegen Satz IV.2.5 gibt es ein $\zeta \in (x, y)$ mit

$$f(y) = f(x) + f'(\zeta)(y - x) \le f(x).$$

Hieraus folgt die Behauptung.
(2): Folgt wie „$\Longleftarrow$" in (1). □

Bemerkung IV.2.8. $f : [-1, 1] \to \mathbb{R}$ mit $f(x) = x^3$ ist streng monoton wachsend und in $(-1, 1)$ differenzierbar. Aber es ist $f'(0) = 0$. Dies zeigt, dass in Satz IV.2.7 (2) die Umkehrung i. a. nicht gilt.

Satz IV.2.9. *Sei $I \subset \mathbb{R}$ ein perfektes Intervall, $f \in C(I, \mathbb{R})$ in $\overset{\circ}{I}$ differenzierbar und $f'(x) \neq 0$ für alle $x \in \overset{\circ}{I}$. Dann ist $f$ injektiv.*

Beweis. Angenommen, $f$ ist nicht injektiv. Dann gibt es $a, b \in I$ mit $f(a) = f(b)$. O.E. ist $a < b$. Dann ist $(a, b) \subset \overset{\circ}{I}$ und aus Satz IV.2.4 folgt

$$f'(\zeta) = 0$$

für ein $\zeta \in (a, b)$; Widerspruch. □

Bemerkung IV.2.10. Satz IV.2.9 ist für Funktionen $f : I \to \mathbb{C}$ i. a. falsch. Z. B. ist $\mathrm{ixp} : [0, 2\pi] \to \mathbb{C}$ mit

$$\mathrm{ixp}(x) = e^{ix}$$

nicht injektiv und erfüllt

$$\mathrm{ixp}'(x) = ie^{ix} \neq 0 \quad \forall x \in (0, 2\pi).$$

Satz IV.2.11. *Sei $I \subset \mathbb{R}$ ein perfektes Intervall und $f \in C(I, \mathbb{R})$ in $\overset{\circ}{I}$ differenzierbar mit $f'(x) \neq 0$ für alle $x \in \overset{\circ}{I}$. Dann gilt:*

(1) *$f$ ist streng monoton,*
(2) *$J = f(I)$ ist ein perfektes Intervall,*
(3) *$f^{-1} : J \to I$ ist differenzierbar in $\overset{\circ}{J}$ und*

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

Beweis. Gemäß Satz IV.2.9 ist $f$ injektiv. Wegen Satz III.5.5 (S. 90) ist $J$ ein Intervall. Angenommen $f$ wäre nicht streng monoton. Dann gibt es $x, y, z \in I$ mit $x < y < z$ und

$$f(x) < f(y) \text{ und } f(y) > f(z)$$

oder

$$f(x) > f(y) \text{ und } f(y) < f(z).$$

Daher hat $f$ in $(x, z)$ ein Extremum. Wegen Satz IV.2.2 ist dies ein Widerspruch zur Annahme $f'(u) \neq 0$ für alle $u \in \overset{\circ}{I}$. Also ist $f$ streng monoton. Da $f$ streng monoton ist, besteht $J$ nicht aus einem Punkt, ist also perfekt. Der Rest folgt aus Bemerkung IV.1.10 (S. 116). $\square$

Beispiel IV.2.12. Es ist

$$\cos'(x) = -\sin(x) \neq 0 \text{ auf } (0, \pi)$$
$$\sin'(x) = \cos(x) \neq 0 \text{ auf } (-\frac{\pi}{2}, \frac{\pi}{2}).$$

Also existieren

$$\arccos = (\cos|_{(0,\pi)})^{-1} : (-1, 1) \to (0, \pi)$$
$$\arcsin = (\sin|_{(-\frac{\pi}{2},\frac{\pi}{2})})^{-1} : (-1, 1) \to (-\frac{\pi}{2}, \frac{\pi}{2}).$$

Es gilt

$$\begin{aligned}
\arccos'(x) &\underset{x=\cos y}{=} \frac{1}{\cos'(y)} \\
&= \frac{-1}{\sin(y)} \\
&= \frac{-1}{\sqrt{1-\cos^2(y)}} \\
&= \frac{-1}{\sqrt{1-x^2}} \qquad \forall x \in (-1, 1) \\
\arcsin'(x) &\underset{x=\sin y}{=} \frac{1}{\sin'(y)} \\
&= \frac{1}{\cos(y)}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{1-\sin^2(y)}} \\
&= \frac{1}{\sqrt{1-x^2}} \qquad \forall x \in (-1,1).
\end{aligned}
$$

Durch Induktion folgt hieraus

$$
\begin{aligned}
\arccos &\in C^\infty((-1,1),(0,\pi)), \\
\arcsin &\in C^\infty((-1,1),(-\frac{\pi}{2},\frac{\pi}{2})).
\end{aligned}
$$

DEFINITION IV.2.13. Sei $(X, \|\cdot\|_X)$ ein normierter Vektorraum, $M \subset X$ konvex und $f : M \to \mathbb{R}$ eine Funktion. Dann heißt

(1) $f$ KONVEX,
(2) $f$ STRIKT KONVEX,
(3) $f$ KONKAV,
(4) $f$ STRIKT KONKAV,

wenn gilt

(1) $f((1-t)x+ty) \le (1-t)f(x)+tf(y)$,
(2) $f((1-t)x+ty) < (1-t)f(x)+tf(y)$,
(3) $f((1-t)x+ty) \ge (1-t)f(x)+tf(y)$,
(4) $f((1-t)x+ty) > (1-t)f(x)+tf(y)$

für alle $x, y \in M, x \neq y$, und alle $t \in (0,1)$.

BEMERKUNG IV.2.14. (1) $f$ ist genau dann (strikt) konvex, wenn $-f$ (strikt) konkav ist.
(2) Ist $f$ konvex, so ist die Menge

$$\{(x,u) : x \in M, u \in \mathbb{R}, u \ge f(x)\}$$

konvex.
(3) Ist $M \subset \mathbb{R}$, so ist $f$ genau dann (strikt) konvex, wenn für alle $a, b \in M$ mit $a < b$ und alle $x \in (a,b)$ gilt

$$f(x) \underset{(<)}{\le} f(a) + \frac{f(b)-f(a)}{b-a}(x-a).$$

(4) Ist $M \subset \mathbb{R}$, so ist $f$ genau dann (strikt) konvex, wenn für alle $a, b \in M$ mit $a < b$ und alle $x \in (a,b)$ gilt

$$\frac{f(x)-f(a)}{x-a} \underset{(<)}{\le} \frac{f(b)-f(a)}{b-a} \underset{(<)}{\le} \frac{f(b)-f(x)}{b-x}.$$

BEWEIS. AD (1), (2): Sind klar.
AD (3): Sei $a, b \in M, a < b$, und $x \in (a,b)$. Dann ist $x = (1-t)a + tb$ mit $t = \frac{x-a}{b-a} \in (0,1)$. Damit folgt

$$f(x) = f((1-t)a+tb) \underset{(<)}{\le} (1-t)f(a) + tf(b)$$

$$= f(a) + \frac{f(b) - f(a)}{b - a}(x - a).$$

AD (4): Die erste Ungleichung ist offensichtlich äquivalent zu (3). Sei $a, b \in M$, $a < b$, und $x \in (a, b)$. Dann ist $x = (1 - t)b + ta$ mit $t = \frac{b-x}{b-a} \in (0, 1)$. Damit folgt

$$\begin{aligned} f(x) &= f((1 - t)b + ta) \\ &\underset{(<)}{\leq} f(b) + \frac{f(a) - f(b)}{b - a}(b - x). \end{aligned}$$

Das ist offensichtlich äquivalent zur zweiten Ungleichung. □

SATZ IV.2.15. (1) *Sei $I \subset \mathbb{R}$ ein perfektes Intervall und $f$ in $\overset{\circ}{I}$ differenzierbar. Dann ist $f$ genau dann (strikt) konvex, wenn $f'$ (streng) monoton wachsend ist.*
(2) *Sei $I \subset \mathbb{R}$ ein perfektes Intervall und $f$ in $\overset{\circ}{I}$ zweimal differenzierbar. Dann ist $f$ genau dann konvex, wenn $f''(x) \geq 0$ ist für alle $x \in \overset{\circ}{I}$. Gilt $f''(x) > 0$ für alle $x \in \overset{\circ}{I}$, so ist $f$ strikt konvex.*

BEWEIS. AD (1): „$\Longrightarrow$:“ Sei $a, b \in \overset{\circ}{I}$ mit $a < b$. Dann folgt aus Bemerkung IV.2.14 (4)

$$\begin{aligned} \frac{f(x) - f(a)}{x - a} &\leq \frac{f(b) - f(a)}{b - a} \leq \frac{f(b) - f(x)}{b - x} \\ \Longrightarrow \quad f'(a) = D_+(a) &\leq \frac{f(b) - f(a)}{b - a} \\ &\leq D_- f(b) = f'(b). \end{aligned}$$

Also ist $f'$ monoton wachsend. Sei nun $f$ strikt konvex.
Ann.: $f'$ ist nicht streng monoton wachsend. Da $f'$ monoton wachsend ist, gibt es dann $x, y \in \overset{\circ}{I}$ mit $x < y$ und

$$f'(x) = f'(z) = f'(y) \quad \forall z \in [x, y].$$

Sei $z \in (x, y)$. Dann folgt aus der strikten Konvexität von $f$ und Satz IV.2.5

$$\begin{aligned} f(z) &< f(x) + \frac{f(y) - f(x)}{y - x}(z - x) \\ &= f(x) + f'(\eta_1)(z - x) \end{aligned}$$

mit $\eta_1 \in (x, y)$. Andererseits gibt es wegen Satz IV.2.5 ein $\eta_2 \in (x, y)$ mit

$$f(z) = f(x) + f'(\eta_2)(z - x).$$

Wegen $f'(\eta_1) = f'(\eta_2)$ ist dies ein Widerspruch.
„$\Longleftarrow$:“ Seien $a, b \in I$ mit $a < b$ und $x \in (a, b)$. Dann ist $x \in \overset{\circ}{I}$ und

gemäß Satz IV.2.5 gibt es ein $\zeta \in (a,x)$ und ein $\eta \in (x,b)$ mit

$$\frac{f(x)-f(a)}{x-a} = f'(\zeta)$$

und

$$\frac{f(b)-f(x)}{b-x} = f'(\eta).$$

Wegen der (strengen) Monotonie von $f'$ folgt

$$\begin{aligned} \frac{f(x)-f(a)}{x-a} &\underset{(<)}{\leq} \frac{f(b)-f(x)}{b-x} \\ \Longrightarrow\ f(x)\frac{b-a}{(x-a)(b-x)} &\underset{(<)}{\leq} \frac{f(a)}{x-a} + \frac{f(b)}{b-x} \\ \Longrightarrow\ f(x) &\underset{(<)}{\leq} f(a)\frac{b-x}{b-a} + f(b)\frac{x-a}{b-a}. \end{aligned}$$

Also gilt für $t \in (0,1)$

$$f((1-t)a+tb) \underset{(<)}{\leq} (1-t)f(a) + tf(b).$$

Mithin ist $f$ (strikt) konvex.
AD (2): Folgt aus Teil (1) und Satz IV.2.7. □

BEISPIEL IV.2.16. (1) $\exp : \mathbb{R} \to \mathbb{R}$ ist streng monoton wachsend und strikt konvex.
(2) $\ln : \mathbb{R}_+^* \to \mathbb{R}$ ist streng monoton wachsend und strikt konkav.
(3) $x \mapsto x^\alpha : \mathbb{R}_+^* \to \mathbb{R}$ ist streng monoton wachsend und konvex für $\alpha > 1$, streng monoton wachsend und konkav für $0 < \alpha < 1$, streng monoton fallend und konvex für $\alpha < 0$.
(4) Sei $1 < p < \infty$ und $p' = \frac{p}{p-1}$. Dann gilt für alle $x, y \in \mathbb{R}_+^*$

$$x \cdot y \leq \frac{1}{p}x^p + \frac{1}{p'}y^{p'}. \quad \text{(YOUNGSCHE UNGLEICHUNG)}$$

(5) Sei $1 < p < \infty$ und $p' = \frac{p}{p-1}$. Für $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ definiere

$$\|x\|_p = \{\sum_{k=1}^n |x_k|^p\}^{1/p}.$$

Dann gilt für $x, y \in \mathbb{K}^n$

$$\sum_{k=1}^n |x_k||y_k| \leq \|x\|_p \|y\|_{p'}. \quad \text{(HÖLDERSCHE UNGLEICHUNG)}$$

Insbesondere ist $\|\cdot\|_p$ eine Norm auf $\mathbb{K}^n$.

Beweis. ad (1): $\exp = \exp' = \exp'' > 0$ auf $\mathbb{R}$.
ad (2):

$$\ln' = \frac{1}{x} > 0 \quad \text{auf} \quad \mathbb{R}_+^*,$$
$$\ln'' = -\frac{1}{x^2} < 0 \quad \text{auf} \quad \mathbb{R}_+^*.$$

ad (3):

$$(x^\alpha)' = \alpha x^{\alpha-1} \quad \begin{cases} > 0 \ \text{auf } \mathbb{R}_+^* & , \text{ für } \alpha > 0, \\ < 0 \ \text{auf } \mathbb{R}_+^* & , \text{ für } \alpha < 0 \end{cases}$$

$$(x^\alpha)'' = \alpha(\alpha-1)x^{\alpha-2} \quad \begin{cases} > 0 \quad \text{auf } \mathbb{R}_+^* & , \text{ für } \alpha > 1 \text{ und } \alpha < 0 \\ < 0 \quad \text{auf } \mathbb{R}_+^* & , \text{ für } 0 < \alpha < 1 \end{cases}$$

ad (4): Aus (2) folgt

$$\begin{aligned} \ln(\frac{1}{p}x^p + \frac{1}{p'}y^{p'}) &= \ln(\frac{1}{p}x^p + (1-\frac{1}{p})y^{p'}) \\ &\geq \frac{1}{p}\ln(x^p) + (1-\frac{1}{p})\ln(y^{p'}) \\ &= \ln(x) + \ln(y) \\ &= \ln(x \cdot y). \end{aligned}$$

ad (5): O.E. ist $\|x\|_p \neq 0$, $\|y\|_{p'} \neq 0$. Dann folgt aus (4)

$$\begin{aligned} \sum_{k=1}^n \frac{|x_k|}{\|x\|_p} \frac{|y_k|}{\|y\|_{p'}} &\leq \frac{1}{p} \sum_{k=1}^n \frac{|x_k|^p}{\|x\|_p^p} + \frac{1}{p'} \sum_{k=1}^n \frac{|y_k|^{p'}}{\|y\|_{p'}^{p'}} \\ &= \frac{1}{p} + \frac{1}{p'} \\ &= 1 \end{aligned}$$

und somit

$$\sum_{k=1}^n |x_k||y_k| \leq \|x\|_p \|y\|_{p'}.$$

Offensichtlich muss man von den Norm-Eigenschaften nur noch die Dreiecksungleichung beweisen. Für $x, y \in \mathbb{K}^n$ folgt

$$\begin{aligned} \|x+y\|_p^p = & \sum_{k=1}^n |x_k + y_k|^p \\ \leq & \sum_{k=1}^n |x_k + y_k|^{p-1}[|x_k| + |y_k|] \\ = & \sum_{k=1}^n |x_k + y_k|^{p-1}|x_k| + \sum_{k=1}^n |x_k + y_k|^{p-1}|y_k| \end{aligned}$$

$$\begin{aligned} &\underset{\text{Hölder}}{\leq} \|x\|_p\{\sum_{k=1}^n |x_k+y_k|^{(p-1)p'}\}^{1/p'} \\ &\qquad + \|y\|_p\{\sum_{k=1}^n |x_k+y_k|^{(p-1)p'}\}^{1/p'} \\ &= \quad [\|x\|_p + \|y\|_p]\|x+y\|_p^{p-1}. \end{aligned}$$

Hieraus folgt die Behauptung. □

Bemerkung IV.2.10 zeigt, dass der Mittelwertsatz IV.2.5 für Funktionen mit Werten in $\mathbb{C}$ oder $\mathbb{K}^n, n \geq 2$, nicht gelten kann. Man kann allerdings eine Abschwächung beweisen.

Satz IV.2.17. *Sei $M \subset \mathbb{K}$ konvex und $f : M \to \mathbb{K}^n$, $n \in \mathbb{N}^*$, in $\overset{\circ}{M}$ differenzierbar. Dann gibt es für alle $a, b \in M$ ein $\theta \in (0,1)$ mit*

$$\|f(b) - f(a)\|_2 \leq \|f'((1-\theta)a + \theta b)\|_2 |b-a|,$$

*wobei $\|\cdot\|_2$ die euklidische Norm auf $\mathbb{K}^n$ ist.*

Beweis. Seien $a, b \in M$. Wir definieren einen Vektor $v \in \mathbb{K}^n$ mit $\|v\|_2 = 1$ wie folgt

$$v_k = \begin{cases} \frac{1}{\sqrt{n}} & \text{, falls } f(b) = f(a), \\ \frac{\overline{f_k(b) - f_k(a)}}{\|f(b)-f(a)\|_2} & \text{, falls } f(b) \neq f(a), \end{cases}$$

$1 \leq k \leq n$. Außerdem definieren wir eine Funktion $\varphi : [0,1] \to \mathbb{R}$ durch

$$\varphi(t) = \operatorname{Re}\left(\sum_{k=1}^n f_k((1-t)a + tb)v_k\right).$$

Dann gilt

$$\begin{aligned} \varphi(1) - \varphi(0) &= \operatorname{Re}\left(\sum_{k=1}^n [f_k(b) - f_k(a)]v_k\right) \\ &= \|f(b) - f(a)\|_2. \end{aligned}$$

Offensichtlich ist $\varphi$ in $(0,1)$ differenzierbar mit

$$\varphi'(t) = \operatorname{Re}\left(\sum_{k=1}^n (b-a) f_k'((1-t)a + tb)v_k\right).$$

Gemäß Satz IV.2.5 gibt es ein $\theta \in (0,1)$ mit

$$\begin{aligned} \|f(b) - f(a)\|_2 &= \varphi(1) - \varphi(0) \\ &= \varphi'(\theta) \\ &= \operatorname{Re}\left(\sum_{k=1}^n (b-a) f_k'((1-\theta)a + \theta b)v_k\right) \end{aligned}$$

$$
\begin{aligned}
&\leq |b-a| \sum_{k=1}^{n} |f_k'((1-\theta)a+\theta b)||v_k| \\
&\leq |b-a| \|f'((1-\theta)a+\theta b)\|_2.
\end{aligned}
$$

□

Bemerkung IV.2.18. Man kann die Aussage von Satz IV.2.17 auch für beliebige andere Normen auf $\mathbb{K}^n$ und für Funktionen mit Werten in beliebigen normierten Vektorräumen beweisen.

Zum Abschluss beweisen wir noch eine Variante des Mittelwertsatzes, die mit der Regel von de l'Hôpital (1661-1704) eine besonders wichtige praktische Konsequenz hat.

Satz IV.2.19. *Seien $I \subset \mathbb{R}$ ein perfektes Intervall und $f, g \in C(I, \mathbb{R})$ in $\overset{\circ}{I}$ differenzierbar mit $g'(x) \neq 0$ für alle $x \in \overset{\circ}{I}$. Dann gibt es für alle $a, b \in I$ mit $a < b$ ein $\zeta \in (a, b)$ mit*

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

Beweis. Wegen Satz IV.2.4 ist $g(b) \neq g(a)$. Daher ist

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

in $[a, b]$ stetig und in $(a, b)$ differenzierbar. Wegen

$$h(a) = f(a) = h(b)$$

folgt die Behauptung aus Satz IV.2.4. □

Satz IV.2.20 (Regel von de l'Hôpital). *Seien $a, b \in \mathbb{R}$ mit $a < b$ und $f, g \in C([a, b], \mathbb{R})$ in $(a, b)$ differenzierbar mit $f(a) = g(a) = 0$ und $g'(x) \neq 0$ für alle $a < x < b$. Existiert*

$$\lim_{x \to a+0} \frac{f'(x)}{g'(x)},$$

*so existiert auch $\lim\limits_{x \to a+0} \dfrac{f(x)}{g(x)}$ und es gilt*

$$\lim_{x \to a+0} \frac{f(x)}{g(x)} = \lim_{x \to a+0} \frac{f'(x)}{g'(x)}.$$

Beweis. Wegen Satz IV.2.4 ist $g(x) \neq 0$ für alle $x \in (a, b)$. Wegen Satz IV.2.19 existiert zu jedem $x \in (a, b)$ ein $\zeta(x) \in (a, x)$ mit

$$\frac{f(x)}{g(x)} = \frac{f'(\zeta(x))}{g'(\zeta(x))}.$$

Wegen $\zeta(x) \to a$ für $x \to a$ folgt hieraus die Behauptung. □

Bemerkung IV.2.21. (1) Die Voraussetzungen von Satz IV.2.20 sind insbesondere erfüllt, wenn $f, g \in C^1([a,b],\mathbb{R})$ sind mit $g'(a) \neq 0$.
(2) Eine analoge Aussage gilt auch für die linksseitigen Grenzwerte $x \to b - 0$.

Beispiel IV.2.22.

(1)
$$\begin{aligned}\lim_{x\to a} \frac{x^n - a^n}{x^m - a^m} &= \lim_{x\to a} \frac{nx^{n-1}}{mx^{m-1}} \\ &= \frac{n}{m} a^{n-m}.\end{aligned}$$

(2)
$$\begin{aligned}&\lim_{x\to\infty} \left( \sqrt[n]{x^n + a_1 x^{n-1} + \ldots + a_{n-1}x + a_n} - x \right) \\ &= \lim_{x\to\infty} \frac{\sqrt[n]{1 + a_1 x^{-1} + \ldots + a_{n-1}x^{-(n-1)} + a_n x^{-n}} - 1}{\frac{1}{x}} \\ &\underset{y=\frac{1}{x}}{=} \lim_{y\to 0} \frac{\sqrt[n]{1 + a_1 y + \ldots + a_n y^n} - 1}{y} \\ &= \lim_{y\to 0} \frac{1}{n}(1 + a_1 y + \ldots + a_n y^n)^{-1+\frac{1}{n}}(a_1 + 2a_2 y + \ldots + n a_n y^{n-1}) \\ &= \frac{a_1}{n}.\end{aligned}$$

## IV.3. Taylorformeln

Sei $I \subset \mathbb{R}$ ein perfektes Intervall, $f \in C^n(I,\mathbb{R})$ und $x_0 \in \overset{\circ}{I}$. Wir möchten $f$ „möglichst gut“ durch ein Polynom $p$ der Form

$$p(x) = \sum_{k=0}^{n} c_k (x - x_0)^k$$

approximieren. Dabei stellen wir uns vor, dass die Annäherung gut ist, wenn $f$ und $p$ in $x_0$ mit allen Ableitungen bis und mit Ordnung $n$ übereinstimmen. Dies liefert

$$\begin{aligned} f(x_0) = p(x_0) &\iff c_0 = f(x_0) \\ f'(x_0) = p'(x_0) &\iff c_1 = f'(x_0) \\ f^{(k)}(x_0) = p^{(k)}(x_0) &\iff k! c_k = f^{(k)}(x_0), 1 \leq k \leq n, \end{aligned}$$

d.h., $p$ ist notwendigerweise von der Form

$$p(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k.$$

Offensichtlich können wir diese Argumentation auch durchführen, wenn $I$ eine konvexe, perfekte Teilmenge von $\mathbb{K}$ ist und wenn $f$ Werte in einem normierten Vektorraum $(Y, \|\cdot\|_Y)$ hat. Dies führt auf folgende Definition.

DEFINITION IV.3.1. Sei $M \subset \mathbb{K}$ konvex und perfekt und $(Y, \|\cdot\|_Y)$ ein normierter Vektorraum.
(1) Sei $f \in C^n(M, Y), n \in \mathbb{N}^*$. Dann heißt

$$T_n(f, x_0)(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k$$

das $n$-te TAYLORPOLYNOM von $f$ mit Entwicklungspunkt $x_0$.
(2) Sei $f \in C^\infty(M, Y)$. Dann heißt

$$T(f, x_0)(x) = \sum \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k$$

die TAYLORREIHE von $f$ mit Entwicklungspunkt $x_0$.
Man spricht in beiden Fällen auch kurz von der TAYLORENTWICKLUNG von $f$ um $x_0$.

Wir wenden uns zunächst der Frage zu, wie gut $f$ durch seine Taylorpolynome approximiert wird.

SATZ IV.3.2 (TAYLORSCHE FORMEL, B. TAYLOR (1685-1731)).
(1) *Sei $I \subset \mathbb{R}$ ein perfektes Intervall, $f \in C^n(I, \mathbb{R})$ und $f^{(n)}$ in $\overset{\circ}{I}$ differenzierbar. Dann gibt es zu jedem $x, x_0 \in I$ ein $\zeta \in \overset{\circ}{I}$ mit*

$$f(x) = T_n(f, x_0)(x) + \frac{f^{(n+1)}(\zeta)}{(n+1)!}(x - x_0)^{n+1}.$$

(2) *Sei $M \subset \mathbb{K}$ konvex und perfekt, $f \in C^n(M, \mathbb{K}^m)$ und $f^{(n)}$ in $\overset{\circ}{M}$ differenzierbar. Dann gibt es zu jedem $x, x_0 \in M$ ein $\theta \in (0, 1)$ mit*

$$\begin{aligned} &\|f(x) - T_n(f, x_0)(x)\|_2 \\ \le &\frac{1}{(n+1)!}\|f^{(n+1)}((1-\theta)x_0 + \theta x)\|_2 |x - x_0|^{n+1}. \end{aligned}$$

BEWEIS. AD (1): Seien $x, x_0 \in I$. O.E. ist $x \neq x_0$. Sei

$$\rho = [f(x) - T_n(f, x_0)(x)] \frac{(n+1)!}{(x - x_0)^{n+1}}$$

und

$$\begin{aligned} \varphi(t) = f(x) - &\sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0 + t(x - x_0))(1 - t)^k (x - x_0)^k \\ &- \rho(1 - t)^{n+1} \frac{(x - x_0)^{n+1}}{(n+1)!}. \end{aligned}$$

Dann gilt

$$\begin{aligned} \varphi(0) &= f(x) - T_n(f, x_0)(x) - \rho \frac{(x - x_0)^{n+1}}{(n+1)!} = 0 \\ \varphi(1) &= f(x) - f(x) = 0 \end{aligned}$$

und

$$\begin{aligned}\varphi'(t) = -\sum_{k=0}^{n} \frac{1}{k!}[f^{(k+1)}(x_0+t(x-x_0))(1-t)^k(x-x_0)^{k+1} \\ - f^{(k)}(x_0+t(x-x_0))k(1-t)^{k-1}(x-x_0)^k] \\ + \rho(1-t)^n\frac{(x-x_0)^{n+1}}{n!} \\ = -\frac{1}{n!}f^{(n+1)}(x_0+t(x-x_0))(1-t)^n(x-x_0)^{n+1} \\ + \rho\frac{1}{n!}(1-t)^n(x-x_0)^{n+1}.\end{aligned}$$

Gemäß Satz IV.2.4 (S. 122) gibt es ein $\theta \in (0,1)$ mit

$$\begin{aligned}&\varphi'(\theta) = 0 \\ \Longrightarrow &\rho = f^{(n+1)}(x_0+\theta(x-x_0)).\end{aligned}$$

Hieraus folgt die Behauptung mit $\zeta = x_0 + \theta(x-x_0)$.
AD (2): Seien $x, x_0 \in M$ und o.E. $x \neq x_0$. Wir definieren $\rho$ wie in Teil (1). Man beachte, dass jetzt $\rho \in \mathbb{K}^m$ ist. Wir definieren den Vektor $v \in \mathbb{K}^m$ durch

$$v_k = \begin{cases} \frac{1}{\sqrt{m}} & \text{falls } \rho = 0, \\ \frac{\overline{\rho}_k}{\|\rho\|_2} & \text{falls } \rho \neq 0, \end{cases}$$

$1 \le k \le m$. Sei

$$\begin{aligned}\psi(t) = \mathrm{Re}\Bigg(\sum_{l=1}^{m} v_l\Big\{f_l(x) - \sum_{k=0}^{n}\frac{1}{k!}f_l^{(k)}(x_0+t(x-x_0))\cdot(1-t)^k(x-x_0)^k \\ -\rho_l\frac{1}{(n+1)!}(1-t)^{n+1}(x-x_0)^{n+1}\Big\}\Bigg).\end{aligned}$$

Dann ist $\psi : [0,1] \to \mathbb{R}$. Wie in Teil (1) folgt

$$\begin{aligned}&\psi(0) = 0 \\ &\psi(1) = 0 \\ &\psi'(t) = \mathrm{Re}\left(\sum_{l=1}^{n} v_l\Big\{\rho_l - f_l^{(n+1)}(x_0+t(x-x_0))\Big\}\frac{1}{n!}(x-x_0)^{n+1}\right).\end{aligned}$$

Gemäß Satz IV.2.4 (S. 122) gibt es wieder ein $\theta \in (0,1)$ mit $\psi'(\theta) = 0$. Wegen der Definition von $v$ folgt

$$\|\rho\|_2 = \mathrm{Re}\left(\sum_{l=1}^{n} \rho_l v_l\right)$$

$$= \operatorname{Re}\left(\sum_{l=1}^{n} f_l^{(n+1)}(x_0+\theta(x-x_0))v_l\right)$$
$$\le \|f^{(n+1)}(x_0+\theta(x-x_0))\|_2$$

und somit

$$\begin{aligned}\|f(x)-T_n(f,x_0)x\|_2 &= \frac{1}{(n+1)!}|x-x_0|^{n+1}\|\rho\|_2 \\ &\le \frac{1}{(n+1)!}|x-x_0|^{n+1}\|f^{(n+1)}(x_0+\theta(x-x_0))\|_2.\end{aligned}$$

Damit ist die Behauptung bewiesen. □

Bemerkung IV.3.3. Ebenso wie Satz IV.2.17 (S. 129) kann man Satz IV.3.2 (2) auch für jede andere Norm des $\mathbb{K}^m$ und für Funktionen mit Werten in einem beliebigen normierten Vektorraum $(Y, \|\cdot\|_Y)$ beweisen.

Bemerkung IV.3.4. (1) Sei $I \subset \mathbb{R}$ ein perfektes Intervall und $f \in C^2(I,\mathbb{R})$ konvex. Dann gilt für alle $x_0, x \in I$

$$f(x) \ge f(x_0) + f'(x_0)(x-x_0).$$

(2) Sei $I \subset \mathbb{R}$ ein perfektes Intervall, $f \in C^n(I,\mathbb{R})$ und $x_0 \in \overset{\circ}{I}$ mit

$$f^{(1)}(x_0) = \ldots = f^{(n-1)}(x_0) = 0, f^{(n)}(x_0) \ne 0, \quad n \ge 2.$$

Dann gilt:

(a) Ist $n$ ungerade, so ist $x_0$ kein lokales Extremum von $f$.
(b) Ist $n$ gerade, so ist $x_0$ ein lokales Extremum von $f$ und zwar ein lokales Minimum, falls $f^{(n)}(x_0) > 0$ ist, und ein lokales Maximum, falls $f^{(n)}(x_0) < 0$ ist.

Beweis. ad (1): Wegen Satz IV.2.15(2) (S. 126) und Satz IV.3.2 gilt für $x, x_0 \in I$ mit einem $\zeta \in \overset{\circ}{I}$

$$\begin{aligned}f(x) &= f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2}f''(\zeta)(x-x_0)^2 \\ &\ge f(x_0) + f'(x_0)(x-x_0).\end{aligned}$$

ad (2): O.E. ist $f^{(n)}(x_0) > 0$; sonst gehen wir zu $-f$ über. Da $f^{(n)}$ stetig ist, gibt es ein $\delta > 0$ mit $(x_0-\delta, x_0+\delta) \subset I$ und

$$f^{(n)}(x) > 0 \quad \forall x \in (x_0-\delta, x_0+\delta).$$

Wegen Satz IV.3.2 gibt es zu jedem $x \in (x_0-\delta, x_0+\delta)$ ein $\zeta \in (x_0-\delta, x_0+\delta)$ mit

$$f(x) = f(x_0) + \frac{1}{n!}f^{(n)}(\zeta)(x-x_0)^n.$$

Falls $n$ ungerade ist, folgt

$$f(x) > f(x_0) \quad \forall x \in (x_0, x_0+\delta)$$

$$f(x) < f(x_0) \quad \forall x \in (x_0 - \delta, x_0).$$

Falls $n$ gerade ist, folgt

$$f(x) > f(x_0) \quad \forall x \in (x_0 - \delta, x_0 + \delta).$$

□

Bemerkung IV.3.5. (1) Die Taylorreihe von $f$ ist eine Potenzreihe und konvergiert in ihrem Konvergenzkreis. Für ein $x \neq x_0$ aus dem Konvergenzkreis gilt $f(x) = T(f, x_0)(x)$ genau dann, wenn gilt

$$\lim_{n\to\infty} |f(x) - T_n(f, x_0)(x)| = 0.$$

(2) Für die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 0 & \text{für } x \le 0 \\ e^{-1/x} & \text{für } x > 0 \end{cases}$$

aus Beispiel IV.1.21 (S. 120) gilt

$$T(f, 0)(x) = 0 \quad \forall x \in \mathbb{R}.$$

Die Taylorreihe stellt also die Funktion in keinem Punkt $x \in \mathbb{R}_+^*$ dar, obwohl sie den Konvergenzradius $\infty$ hat.

Beispiel IV.3.6. Es gilt:

(1) $\ln(1 + x) = \displaystyle\sum_{n=0}^{\infty} (-1)^{n-1} \frac{x^n}{n}$ für alle $x \in [-\frac{1}{2}, 1]$.

(2) $\ln(1 - x) = -\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{n}$ für alle $x \in [-1, \frac{1}{2}]$.

(3) $\ln(\frac{1+x}{1-x}) = 2\displaystyle\sum_{k=0}^{\infty} \frac{x^{2k+1}}{2k+1}$ für alle $x \in [-\frac{1}{2}, \frac{1}{2}]$.

(4) Möchte man $\ln(2)$ mit Hilfe der Darstellungen (1) oder (2) auf 10-Stellen genau berechnen, so benötigt man ca. $10^{10}$ Summanden. Benutzt man dagegen die Darstellung (3) mit $x = \frac{1}{3}$, so kommt man mit ca. 11 Summanden aus.

Beweis. ad (1): Gemäß Beispiel IV.1.17(6) (S. 118) gilt für

$$\begin{aligned} \varphi(x) &= \ln(1 + x) \quad , x \in (-1, +\infty) \\ \varphi(0) &= 0 \\ \varphi^{(k)}(0) &= (-1)^{k-1}(k - 1)! \quad \forall k \in \mathbb{N}^*. \end{aligned}$$

Also ist

$$T(\ln(1 + x), 0) = \sum_{n\ge1} (-1)^{n-1} \frac{x^n}{n}.$$

Aus Satz IV.3.2 folgt weiter

$$|\ln(1 + x) - \sum_{n=1}^{m} (-1)^{n-1} \frac{x^n}{n}| = \frac{1}{m+1} (\frac{|x|}{1+\zeta})^{m+1}$$

mit $x < \zeta < 0$ falls $x \le 0$ und $0 < \zeta < x$ sonst. Für $x \in [-\frac{1}{2}, 1]$ gilt daher

$$\begin{aligned} \frac{|x|}{1+\zeta} &\le \frac{|x|}{1-|x|} \le 1 && \text{falls } x < 0 \\ \frac{x}{1+\zeta} &\le x \le 1 && \text{falls } x \ge 0. \end{aligned}$$

AD (2): Folgt aus Teil (1) durch Ersetzen von $x$ durch $-x$.
AD (3): Folgt durch Subtraktion von (2) von (1).
AD (4): Folgt aus

$$|\ln(2) - \sum_{n=1}^{m} (-1)^{n-1} \frac{1}{n}| \le \frac{1}{m+1}$$

bzw.

$$\begin{aligned} |\ln(2) - 2\sum_{k=0}^{m-1} \frac{1}{2k+1} (\frac{1}{3})^{2k+1}| &\le \frac{2}{3} \cdot 9^{-m} \cdot \frac{1}{1-\frac{1}{9}} \\ &= \frac{3}{4} \cdot 9^{-m}. \end{aligned}$$

□

## IV.4. Numerische Lösung von Gleichungen

Im Folgenden sei stets $(X, \|\cdot\|_X)$ ein Banachraum, $M \subset X$ nicht leer und $f : M \to X$ eine Funktion.

DEFINITION IV.4.1. (1) Ein Punkt $a \in M$ heißt FIXPUNKT von $f$, wenn gilt $f(a) = a$.
(2) Die Funktion $f$ heißt eine KONTRAKTION in $M$, wenn gilt $f(M) \subset M$ und wenn es ein $\kappa \in (0, 1)$ gibt mit

$$\|f(x) - f(y)\|_X \le \kappa \|x - y\|_X \quad \forall x, y \in M.$$

Die Zahl $\kappa$ heißt KONTRAKTIONSRATE von $f$.

BEMERKUNG IV.4.2. (1) Eine Kontraktion ist stets stetig. Die Umkehrung ist natürlich i. a. falsch.
(2) Falls $M \subset \mathbb{K}$ und $f \in C^1(M, M)$ ist mit

$$|f'(x)| \le \kappa < 1 \quad \forall x \in M,$$

so ist $f$ eine Kontraktion in $M$.

Der folgende Satz hat vielfältige Anwendungen, die weit über die Beispiele dieses Paragraphen hinausgehen. So werden wir z. B. später mit seiner Hilfe die Lösbarkeit von Differentialgleichungen beweisen.

Satz IV.4.3 (Banachscher Fixpunktsatz). *$M$ sei abgeschlossen und $f$ sei eine Kontraktion in $M$ mit Kontraktionsrate $\kappa$. Dann besitzt $f$ genau einen Fixpunkt $x^*$ in $M$. Für beliebiges $x_0 \in M$ konvergiert die Folge $(x_n)_{n\in\mathbb{N}}$ mit*

$$(*) \qquad x_n = f(x_{n-1}) \quad \forall n \in \mathbb{N}^* \text{ (Fixpunktiteration)}$$

*gegen $x^*$ und es gelten die Fehlerabschätzungen*

$$\|x_n - x^*\|_X \le \frac{\kappa^n}{1-\kappa}\|x_1 - x_0\|_X \text{ (a priori Abschätzung)}$$

*und*

$$\|x_n - x^*\|_X \le \frac{\kappa}{1-\kappa}\|x_n - x_{n-1}\|_X. \text{ (a posteriori Abschätzung)}$$

Beweis. Eindeutigkeit: Seien $x^*, y^* \in M$ zwei Fixpunkte von $f$. Dann folgt

$$\begin{aligned} &\|x^* - y^*\|_X = \|f(x^*) - f(y^*)\|_X \le \kappa\|x^* - y^*\|_X \\ \Longrightarrow\, &x^* = y^*. \end{aligned}$$

Existenz: Sei $x_0 \in M$ beliebig und $(x_n)_{n\in\mathbb{N}} \subset M$ gemäß $(*)$ definiert. Dann folgt

$$\begin{aligned} \|x_{n+k} - x_{n+k+1}\|_X &= \|f(x_{n+k-1}) - f(x_{n+k})\|_X \\ &\le \kappa\|x_{n+k-1} - x_{n+k}\|_X. \end{aligned}$$

und somit

$$\text{(i)} \qquad \|x_{n+k} - x_{n+k+1}\|_X \le \kappa^k \|x_n - x_{n+1}\|_X.$$

Aus (i) folgt weiter für $m \in \mathbb{N}^*$

$$\begin{aligned} \|x_{n+m} - x_n\|_X &\le \sum_{k=0}^{m-1} \|x_{n+k+1} - x_{n+k}\|_X \\ &\le \sum_{k=0}^{m-1} \kappa^k \|x_n - x_{n+1}\|_X \\ &\le \frac{1}{1-\kappa}\|x_n - x_{n+1}\|_X \\ \text{(ii)} \qquad &\le \frac{\kappa}{1-\kappa}\|x_n - x_{n-1}\|_X \\ \text{(iii)} \qquad &\le \frac{\kappa^n}{1-\kappa}\|x_1 - x_0\|_X. \end{aligned}$$

Da $\kappa \in (0,1)$ ist, folgt aus (iii), dass $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist. Da $M \subset X$ abgeschlossen ist, konvergiert $(x_n)_{n\in\mathbb{N}}$ gegen ein $x^* \in M$. Aus der Stetigkeit von $f$ folgt

$$x^* = \lim_{n\to\infty} x_n = \lim_{n\to\infty} f(x_{n-1}) = f(x^*).$$

Die Fehlerabschätzungen folgen schließlich, indem man in (ii) bzw. (iii) $n$ festhält und $m$ gegen $\infty$ streben lässt. □

Als nächstes zeigen wir eine für Anwendungen besonders praktische Variante des Banachschen Fixpunktsatzes.

SATZ IV.4.4. *Zu der Abbildung $f : X \to X$ gebe es ein $x_0 \in X$ und zwei Zahlen $\kappa \in (0, 1)$ und $r \in \mathbb{R}_+^*$ mit*

$$\|f(x_0) - x_0\|_x \le (1 - \kappa)r$$

*und*

$$\|f(x) - f(y)\|_X \le \kappa\|x - y\|_X \quad \forall x, y \in \overline{B(x_0, r)}.$$

*Dann besitzt $f$ einen eindeutigen Fixpunkt $x^*$ in $\overline{B(x_0, r)}$, die Folge $(x_n)_{n\in\mathbb{N}}$ mit*

$$x_n = f(x_{n-1}) \quad \forall n \in \mathbb{N}^*$$

*konvergiert gegen $x^*$ und es gelten die Fehlerabschätzungen von Satz IV.4.3.*

BEWEIS. Offensichtlich reicht es zu zeigen, dass $f$ die Menge $M = \overline{B(x_0, r)}$ in sich abbildet. Für $x \in M$ folgt

$$\begin{aligned}\|f(x) - x_0\|_X &\le \|f(x) - f(x_0)\|_X + \|f(x_0) - x_0\|_X \\ &\le \kappa\|x - x_0\|_X + (1 - \kappa)r \\ &\le \kappa r + (1 - \kappa)r \\ &= r.\end{aligned}$$

□

BEISPIEL IV.4.5. Die Strahlungsintensität eines schwarzen Körpers bei der absoluten Temperatur $T$ und der Wellenlänge $\lambda$ beträgt

$$J(\lambda) = \frac{c^2 h}{\lambda^5(e^{ch/\lambda kT} - 1)},$$

wobei $c$ die Lichtgeschwindigkeit, $h$ die Plancksche und $k$ die Boltzmannsche Konstante ist. Gesucht ist die maximale Strahlungsintensität und die zugehörige Wellenlänge. Wegen

$$J(\lambda) = \frac{k^5 T^5}{c^3 h^4} f(\frac{kT}{ch}\lambda)$$

müssen wir das Maximum von

$$f(x) = \frac{1}{x^5(e^{1/x} - 1)}$$

auf $\mathbb{R}_+^*$ bestimmen.
Wegen Satz III.7.4(3) (S. 100) ist

$$\begin{aligned}\lim_{x\to 0+} f(x) &\underset{y=\frac{1}{x}}{=} \lim_{y\to\infty} \frac{1}{y^{-5}(e^y-1)} \\ &= 0.\end{aligned}$$

Wegen Satz IV.2.20 (S. 130) ist

$$\begin{aligned}\lim_{x\to +\infty} f(x) &\underset{y=\frac{1}{x}}{=} \lim_{y\to 0+} \frac{y^5}{e^y-1} \\ &= \lim_{y\to 0+} \frac{5y^4}{e^y} \\ &= 0.\end{aligned}$$

Wegen $f(1) > 0$ besitzt $f$ mindestens ein Maximum $x^*$ in $\mathbb{R}^*$. Dort gilt wegen Satz IV.2.2 (S. 121)

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

Wegen

$$f'(x) = -\frac{5x^4(e^{\frac{1}{x}}-1) - x^3 e^{\frac{1}{x}}}{x^{10}(e^{\frac{1}{x}}-1)^2}$$

gilt

$$\begin{aligned}f'(x^*) = 0 &\iff 5x^*(e^{1/x^*}-1) = e^{1/x^*} \\ &\iff 5\frac{1}{t^*}(e^{t^*}-1) = e^{t^*}, \quad t^* = \frac{1}{x^*} \\ &\iff 5(1-e^{-t^*}) = t^*, \quad t^* = \frac{1}{x^*}.\end{aligned}$$

Also haben wir unser Problem darauf reduziert, einen Fixpunkt von

$$g(t) = 5(1-e^{-t})$$

zu finden. Wegen

$$g'(t) = 5e^{-t}$$

gilt

$$\begin{aligned}g'(t) > 1 \quad &\text{für} \quad 0 < t < \ln(5) \\ g'(t) < 1 \quad &\text{für} \quad t > \ln(5).\end{aligned}$$

Also ist $g(t) - t$ auf $(0, \ln(5))$ streng monoton wachsend und auf $(\ln(5), +\infty)$ streng monoton fallend. Wegen $g(0) = 0$ hat $g$ daher höchstens einen Fixpunkt $t^*$ in $\mathbb{R}_+^*$. Wegen

$$g(4) = 4.90\ldots > 4 \quad \text{und} \quad g(5) = 4.96\ldots < 5$$

gilt $4 < t^* < 5$.
Wir wollen Satz IV.4.4 mit $t_0 = 5$ und $r = 1$ anwenden. Wegen

$$|g'(t)| = 5e^{-t} \le 5e^{-4} = 0.09157\ldots \quad \forall t \in [4,6]$$

und Bemerkung IV.4.2 (S. 136) ist $\kappa \leq 0.092$. Wegen

$$|g(5) - 5| = 0.04\ldots < 1 - \kappa$$

sind die Voraussetzungen von Satz IV.4.4 erfüllt. Die a posteriori Fehlerabschätzung liefert

$$|t^* - t_n| \leq \frac{\kappa}{1-\kappa}|t_n - t_{n-1}| \leq |t_n - t_{n-1}|.$$

Bei Rechnung mit 6 Dezimalstellen erhalten wir folgende Tabelle:

| $n$ | $t_n$ |
|---|---|
| 0 | 5.0 |
| 1 | 4.96631 |
| 2 | 4.965155 |
| 3 | 4.965115 |
| 4 | 4.965114 |
| 5 | 4.965114 |

Also ist

$$t^* = 4.965114 \pm 10^{-6}.$$

Für die gesuchte Wellenlänge ergibt sich

$$\lambda^* = \frac{1}{t^*}\frac{ch}{kT} = (0.2014052 \pm 10^{-7})\frac{ch}{kT}.$$

Die entsprechende Strahlungsintensität ist

$$J(\lambda^*) = 21.201442\frac{k^5T^5}{c^3h^4}.$$

Sei nun $f : \mathbb{K} \to \mathbb{K}$. Gesucht ist eine Nullstelle $x^*$ von $f$. Für jedes $a \in \mathbb{K}^*$ gilt

$$\begin{aligned} f(x^*) = 0 &\iff af(x^*) = 0 \\ &\iff x^* - af(x^*) = x^*. \end{aligned}$$

Das Problem ist also äquivalent dazu, einen Fixpunkt von $g(x) = x - af(x)$ zu bestimmen. Falls $f$ und damit $g$ differenzierbar ist, wissen wir auf Grund von Bemerkung IV.4.2, dass wir $a$ so wählen sollten, dass zumindest in der Nähe von $x^*$

$$|g'(x)| = |1 - af'(x)|$$

möglichst klein ist. Dies ist offensichtlich für $a = \frac{1}{f'(x^*)}$ der Fall. Da wir aber $x^*$ nicht kennen, ist dieses Verfahren nicht praktikabel. Allerdings könnten wir versuchen, bei der Berechnung der $n+1$-ten Iterierten $a = \frac{1}{f'(x_n)}$ zu setzen.

Dieser Ansatz wird auch durch folgende Überlegung gestützt: In der Nähe der $n$-ten Iterierten $x_n$ gilt

$$f(x) \approx f(x_n) + f'(x_n)(x - x_n).$$

Setzen wir diese Approximation in die Bedingung $f(x_{n+1}) = 0$ ein, erhalten wir

$$f(x_n) + f'(x_n)(x_{n+1} - x_n) = 0 \quad \Longleftrightarrow \quad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Diese Überlegungen führen uns zu folgendem Verfahren:

Definition IV.4.6 (Newtonverfahren). Sei $M \subset \mathbb{K}$ perfekt und $f \in C^1(M, \mathbb{K})$. Für $x_0 \in M$ ist das Newtonverfahren mit Startwert $x_0$ gegeben durch

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad n \in \mathbb{N},$$

solange $f'(x_n) \neq 0$ ist.

Bemerkung IV.4.7. Sei $\mathbb{K} = \mathbb{R}$. Dann ist $x_{n+1}$ der Schnittpunkt mit der $x$-Achse der Geraden durch $(x_n, f(x_n))$ mit Steigung $f'(x_n)$.

Satz IV.4.8. *Sei $M \subset \mathbb{K}$ perfekt und $f \in C^2(M, \mathbb{K})$. Es gebe ein $x^* \in \overset{\circ}{M}$ mit $f(x^*) = 0$ und $f'(x^*) \neq 0$. Dann gibt es ein $\delta > 0$, so dass $\overline{B(x^*, \delta)} \subset \overset{\circ}{M}$ ist und dass das Newtonverfahren für jeden Startwert $x_0 \in \overline{B(x^*, \delta)}$ durchführbar ist und gegen $x^*$ konvergiert. Die Konvergenz ist* quadratisch, *d.h., es gibt eine Konstante $c > 0$ mit*

$$|x_{n+1} - x^*| \leq c|x_n - x^*|^2 \quad \forall n \in \mathbb{N}.$$

Beweis. Wegen $f'(x^*) \neq 0$ gibt es ein $\delta_1 > 0$ mit $\overline{B(x^*; \delta_1)} \subset \overset{\circ}{M}$ und

$$|f'(x)| \geq \frac{1}{2}|f'(x^*)| \quad \forall x \in \overline{B(x^*; \delta_1)}.$$

Setze

$$m = \frac{1}{2}|f'(x^*)|.$$

Auf $\overline{B(x^*; \delta_1)}$ definieren wir eine Funktion $g$ durch

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

Es ist offensichtlich

$$g(x^*) = x^*.$$

Außerdem ist $g$ differenzierbar mit

$$g'(x) = 1 - \frac{f'(x)}{f'(x)} + \frac{f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2}.$$

Gemäß Satz III.4.10 (S. 86) existieren

$$M_1 = \max_{x \in \overline{B(x^*; \delta_1)}} |f'(x)|$$

und

$$M_2 = \max_{x \in \overline{B(x^*;\delta_1)}} |f''(x)|.$$

Wegen $f(x^*) = 0$ folgt mit Satz IV.2.17 (S. 129) für alle $x \in \overline{B(x^*;\delta_1)}$

$$|g'(x)| \le \frac{M_2}{m^2}|f(x) - f(x^*)| \le \frac{M_1 M_2}{m^2}|x - x^*|.$$

Wähle

$$\delta = \min\{\delta_1, \frac{m^2}{2M_1, M_2}\}.$$

Dann gilt gemäß Bemerkung IV.4.2 für alle $x, y \in \overline{B(x^*;\delta)}$

$$\begin{aligned} |g(x) - g(y)| &\le \frac{1}{2}|x - y| \\ |g(x) - x^*| &= |g(x) - g(x^*)| \\ &\le \frac{1}{2}|x - x^*|. \end{aligned}$$

Also erfüllen $g$ und $\overline{B(x^*;\delta)}$ die Voraussetzungen von Satz IV.4.3. Hieraus folgt, dass das Newtonverfahren für jeden Startwert $x_0 \in \overline{B(x^*;\delta)}$ durchführbar ist und gegen $x^*$ konvergiert. Aus Satz IV.2.17 (S. 129) folgt schließlich für ein $\theta \in (0,1)$:

$$\begin{aligned} |x_{n+1} - x^*| &= |g(x_n) - g(x^*)| \\ &\le |g'((1-\theta)x^* + \theta x_n)||x_n - x^*| \\ &\le \frac{M_1 M_2}{m^2}\theta |x_n - x^*|^2 \\ &\le \frac{M_1 M_2}{m^2}|x_n - x^*|^2. \end{aligned}$$

□

Bemerkung IV.4.9. Die Konvergenz des Newtonverfahrens ist i. a. nur lokal gegeben. Betrachte z. B. die Funktion $f = \arctan : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$ mit der eindeutigen Nullstelle $x^* = 0$. Es ist

$$f'(x) = \frac{1}{1+x^2}.$$

Wegen

$$\lim_{|x| \to +\infty} |\arctan(x)| = \frac{\pi}{2}$$

und

$$\lim_{|x| \to +\infty} \frac{3|x|}{1+x^2} = 0$$

gibt es ein $R \in \mathbb{R}_+^*$ mit

$$|\arctan(x)| \geq \frac{3|x|}{1+x^2} \quad \forall |x| \geq R.$$

Sei $|x_0| \geq R$. Dann ist

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

und somit

$$\begin{aligned} |x_1| &\geq -|x_0| + |\arctan(x_0)|(1+x_0^2) \\ &\geq -|x_0| + 3|x_0| \\ &= 2|x_0|. \end{aligned}$$

Also divergiert das Newtonverfahren mit Startwert $x_0$.
Für den Startwert $x_0 = 2$ erhalten wir z. B. folgende Ergebnisse:

| $n$ | $x_n$ |
|---|---:|
| 0 | 2.000000 |
| 1 | -3.535747 |
| 2 | 13.950959 |
| 3 | -279.344067 |
| 4 | 122016.998918 |

Unter bestimmten Voraussetzungen lässt sich jedoch die globale Konvergenz des Newtonverfahrens beweisen. Hierzu ein Beispiel.

Satz IV.4.10. *Sei $I \subset \mathbb{R}$ ein perfektes Intervall und $f \in C^2(I, \mathbb{R})$ strikt konvex bzw. strikt konkav. Es gebe eine Nullstelle $x^* \in \overset{\circ}{I}$ von $f$. Dann konvergiert das Newtonverfahren für jeden Startwert $x_0 \in I$ mit $f(x_0) > 0$ bzw. mit $f(x_0) < 0$ monoton gegen $x^*$.*

Beweis. O.E. sei $f$ konvex, sonst gehen wir zu $-f$ über. Wegen Satz IV.2.4 (S. 122) hat $f$ höchstens zwei Nullstellen $z_1, z_2$ in $I$.
Fall 1: $x^*$ ist einzige Nullstelle. Sei $x_0 \in I$ mit $f(x_0) > 0$.
Fall 1a: $x_0 > x^*$: Sei $\overline{x}$ eine Nullstelle von $f'$. Aus Bemerkung IV.3.4 (S. 134) folgt

$$0 = f(x^*) \geq f(\overline{x}) + f'(\overline{x})(x^* - \overline{x}) = f(\overline{x}).$$

Wegen $f(x_0) > 0$ folgt aus Satz IV.2.15 (S. 126)

$$\overline{x} \leq x^*.$$

Also gilt

$$\begin{aligned} f'(x) > 0 \quad \forall x > x^* \\ f(x) > 0 \quad \forall x > x^*. \end{aligned}$$

Damit folgt

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} < x_0$$

und

$$\begin{aligned} x_1 - x^* &= \frac{f(\zeta)f''(\zeta)}{f'(\zeta)^2}(x_0 - x^*) \quad (\text{mit } x^* < \zeta < x_0) \\ &> 0. \end{aligned}$$

Also konvergiert das Newtonverfahren monoton fallend gegen $x^*$.
FALL 1B: $x_0 < x^*$: Dann folgt wie oben

$$\begin{aligned} f'(x) < 0 \quad \forall x < x^* \\ f(x) > 0 \quad \forall x < x^* \end{aligned}$$

und somit

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} > x_0$$

und

$$\begin{aligned} x_1 - x^* &= \frac{f(\zeta)f''(\zeta)}{f'(\zeta)^2}(x_0 - x^*) \quad (\text{mit } x_0 < \zeta < x^*) \\ &< 0. \end{aligned}$$

Also konvergiert das Newtonverfahren monoton wachsend gegen $x^*$.
FALL 2: ES GIBT ZWEI NULLSTELLEN $z_1 < z_2$: Dann gibt es gemäß Satz IV.2.4 (S. 122) ein $z_0 \in (z_1, z_2)$ mit $f'(z_0) = 0$. Es folgt

$$\begin{aligned} f'(x) > 0 \quad &\forall x > z_0 \\ f'(x) < 0 \quad &\forall x < z_0 \\ f(x) < 0 \quad &\forall x \in (z_1, z_2) \\ f(x) > 0 \quad &\forall x > z_2 \ \text{ oder } x < z_1. \end{aligned}$$

Falls $x_0 > z_2$ ist, folgt die Behauptung wie in FALL 1A, andernfalls wie in FALL 1B. □

BEISPIEL IV.4.11. (1) Wir wenden das Newtonverfahren auf Beispiel IV.4.5 an. Es ist

$$f(t) = t - 5(1 - e^{-t}).$$

Wegen $f''(t) = 5e^{-t}$ sind die Voraussetzungen von Satz IV.4.10 erfüllt. Mit dem Startwert $t_0 = 5$ erhalten wir folgende Tabelle:

| $n$ | $t_n$ |
|---|---|
| 0 | 5.0 |
| 1 | 4.9651356 |
| 2 | 4.9651142 |
| 3 | 4.9651142 |

(2) (DIVISIONSFREIE BERECHNUNG VON $\frac{1}{a}$). Sei $a \in \mathbb{R}_+^*$ und $f(x) = \frac{1}{x} - a$. Wegen $f''(x) = \frac{2}{x^3}$ sind die Voraussetzungen von Satz IV.4.10

auf $\mathbb{R}_+^*$ erfüllt. Die Iterationsvorschrift des Newtonverfahrens lautet

$$\begin{aligned} x_{n+1} &= x_n - \frac{\frac{1}{x_n} - a}{-\frac{1}{x_n^2}} \\ &= x_n + (x_n - a x_n^2) \\ &= 2x_n - a x_n^2 \\ &= x_n(2 - a x_n). \end{aligned}$$

Sie erlaubt somit die divisionsfreie Berechnung von $\frac{1}{a}$.

(3) (Verfahren von Heron). Sei $a \in \mathbb{R}_+^*, k \geq 2$ und $f(x) = x^k - a$. Wegen $f''(x) = k(k-1)x^{k-2}$ sind die Voraussetzungen von Satz IV.4.10 auf $\mathbb{R}_+^*$ erfüllt. Die Nullstelle von $f$ ist offensichtlich $\sqrt[k]{a}$. Die Iterationsvorschrift des Newtonverfahrens lautet

$$\begin{aligned} x_{n+1} &= x_n - \frac{x_n^k - a}{k x_n^{k-1}} \\ &= \frac{k-1}{k} x_n + \frac{a}{k x_n^{k-1}}. \end{aligned}$$

Sie erlaubt somit die Berechnung von $\sqrt[k]{a}$ mittels der 4 Grundrechenarten.

Für $k = 2$, $a = 2$, $x_0 = 2$ erhalten wir z. B. folgende Tabelle:

| $n$ | $x_n$ |
|---|---|
| 0 | 2.0 |
| 1 | 1.5 |
| 2 | $1.41\overline{6}$ |
| 3 | 1.4142156 |
| 4 | 1.4142135 |
| 5 | 1.4142135 |

KAPITEL V

# Funktionenfolgen

In diesem Abschnitt beschäftigen wir uns mit Funktionenfolgen und deren Konvergenzverhalten. Im Mittelpunkt stehen Sätze über das Vertauschen von Grenzprozessen, d.h. Sätze der Form „$f_n$ konvergiert gegen $f$, $f_n$ ist stetig $\Longrightarrow$ $f$ ist stetig“. Als Anwendung kommen wir auf Potenzreihen zurück und betrachten analytische Funktionen.
Zur Vorbereitung des nächsten Kapitels betrachten wir schließlich Treppenfunktionen, d.h. stückweise konstante Funktionen, und ihre Grenzwerte.

## V.1. Gleichmässige Konvergenz

Definition V.1.1. Sei $M \subset X$, $M \neq \emptyset$, und $f, f_n, n \in \mathbb{N}$, Abbildungen von $M$ in $Y$. Wir sagen $f_n$ konvergiert punktweise gegen $f$, kurz $f_n \xrightarrow[n\to\infty]{\text{pkt}} f$, wenn für jedes $x \in M$ gilt

$$\lim_{n\to\infty} f_n(x) = f(x).$$

Wir sagen $f_n$ konvergiert gleichmässig gegen $f$, kurz $f_n \xrightarrow[n\to\infty]{\text{glm}} f$, wenn gilt

$$\forall \varepsilon > 0 \; \exists N_\varepsilon \in \mathbb{N} \quad \forall n \geq N_\varepsilon \quad \forall x \in M : \|f_n(x) - f(x)\|_Y < \varepsilon.$$

Bemerkung V.1.2. Es gilt:

(1) $f_n \xrightarrow[n\to\infty]{\text{pkt}} f \Longrightarrow (f_n(x))_{n\in\mathbb{N}}$ ist für jedes $x \in M$ eine Cauchyfolge.

(2) $f_n \xrightarrow[n\to\infty]{\text{pkt}} f \iff \forall \varepsilon > 0 \quad \forall x \in M \quad \exists N_{\varepsilon,x} \in \mathbb{N} \quad \forall n \geq N_{\varepsilon,x}$: $\|f_n(x) - f(x)\|_Y < \varepsilon$.

(3) $f_n \xrightarrow[n\to\infty]{\text{glm}} f \Longrightarrow f_n \xrightarrow[n\to\infty]{\text{pkt}} f$.

Beispiel V.1.3 zeigt, dass die Umkehrung i.a. falsch ist.

Beweis. Ist offensichtlich. □

Beispiel V.1.3. (1) $M = [0, 1]$, $f_n(x) = x^n$,

$$f(x) = \begin{cases} 0 & 0 \leq x < 1 \\ 1 & x = 1. \end{cases}$$

Es gilt $f_n \xrightarrow[n\to\infty]{\text{pkt}} f$, aber $f_n \underset{n\to\infty}{\overset{\text{glm}}{\not\longrightarrow}} f$.

(2) $M = [0,1]$ und

$$f_n(x) = \begin{cases} 2nx & 0 \le x \le \frac{1}{2n} \\ 2 - 2nx & \frac{1}{2n} < x \le \frac{1}{n} \\ 0 & \frac{1}{n} < x \le 1 \end{cases}$$

und

$$f(x) = 0.$$

Dann gilt

$$f_n \xrightarrow[n\to\infty]{\text{pkt}} f \quad \text{aber} \quad f_n \underset{n\to\infty}{\overset{\text{glm}}{\not\longrightarrow}} f.$$

(3) $M = \mathbb{R}$ und

$$f_n = \begin{cases} \frac{1}{n} & n \le x < n+1 \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt $f_n \xrightarrow[n\to\infty]{\text{glm}} 0$.

(4) $M = \mathbb{R}_+^*$ und $f_n(x) = \frac{1}{nx}$. Dann gilt

(1) $f_n \xrightarrow[n\to\infty]{\text{pkt}} 0$,

(2) $f_n \xrightarrow[n\to\infty]{\text{glm}} 0$ auf $[c, \infty)$ für alle $c > 0$,

(3) $f_n \underset{n\to\infty}{\overset{\text{glm}}{\not\longrightarrow}} 0$ auf $\mathbb{R}_+^*$.

Beweis. ad (1): Folgt aus Beispiel II.2.3(1) (S. 35) und der Identität

$$\sup_{0\le x\le 1} |f_n(x) - f(x)| = \sup_{0\le x\le 1} |x^n| = 1 \quad \forall n \in \mathbb{N}.$$

ad (2): Es ist $f_n(0) = 0$ für alle $n \in \mathbb{N}$ und für $x > 0$ gilt

$$\exists n_x \in \mathbb{N} \quad \frac{1}{n_x} < x \iff f_n(x) = 0 \quad \forall n \ge n_x.$$

Hieraus folgt $f_n \xrightarrow[n\to\infty]{\text{pkt}} 0$.

Für alle $n \in \mathbb{N}$ ist aber

$$\sup_{x\in[0,1]} |f_n(x) - f(x)| = 1.$$

Also gilt $f_n \underset{n\to\infty}{\overset{\text{glm}}{\not\longrightarrow}} f$.

ad (3): Wähle zu $\varepsilon > 0$ ein $N_\varepsilon$ so, dass $\frac{1}{N_\varepsilon} < \varepsilon$ ist. Dann gilt

$$|f_n(x) - f(x)| \le \frac{1}{n} < \varepsilon \quad \forall x \in \mathbb{R}, n \ge N_\varepsilon.$$

AD (4):

(1) Ist klar.

(2) Für $\varepsilon > 0$ und $c > 0$ wähle $N_\varepsilon$ so, dass $\frac{1}{N_\varepsilon} < c\varepsilon$ ist. Dann folgt

$$|f_n(x)| \leq \frac{1}{nc} < \varepsilon \quad \forall n \geq N_\varepsilon, x \in [c, \infty).$$

(3) $f_n(\frac{1}{n}) = 1 \Longrightarrow f_n \overset{\text{glm}}{\not\longrightarrow} 0.$

□

DEFINITION V.1.4. $B(M, Y)$ bezeichnet die Menge aller beschränkten Abbildungen von $M$ in $Y$. Für $f \in B(M, Y)$ setze

$$\|f\|_\infty = \sup_{x \in M} \|f(x)\|_Y.$$

BEMERKUNG V.1.5. (1) $B(M, Y)$ ist eine Verallgemeinerung von $\ell_\infty$.

(2) $f_n \xrightarrow[n\to\infty]{\text{glm}} f \iff \|f_n - f\|_\infty \xrightarrow[n\to\infty]{} 0.$

(3) Es kann gelten $f_n \xrightarrow[n\to\infty]{\text{glm}} f$ und $f_n, f \notin B(M, Y)$. Betrachte dazu z.B. $M = Y = \mathbb{R}$ und

$$f_n(x) = x + \frac{1}{n}, \quad f(x) = x.$$

Die Funktionen $f, f_n$ sind nicht beschränkt, aber $f_n$ konvergiert gleichmäßig gegen $f$.

SATZ V.1.6. *$Y$ sei ein Banachraum, dann ist $(B(M, Y), \|\cdot\|_\infty)$ ebenfalls ein Banachraum.*

BEWEIS. Offensichtlich ist $B(M, Y)$ ein Vektorraum und $\|\cdot\|_\infty$ eine Norm auf $B(M, Y)$. Sei $(f_n)_{n\in\mathbb{N}} \subset B(M, Y)$ eine Cauchyfolge in $B(M, Y)$. Dann ist $(f_n(x))_{n\in\mathbb{N}} \subset Y$ für jedes $x \in M$ eine Cauchyfolge in $Y$. Da $Y$ vollständig ist, existiert somit

$$f(x) = \lim_{n\to\infty} f_n(x) \quad \forall x \in M.$$

Wegen Bemerkung III.1.5 (S. 62) ist $(\|f_n\|_\infty)_{n\in\mathbb{N}} \subset \mathbb{R}$ auch eine Cauchyfolge und somit beschränkt. Sei

$$C = \sup_{n\in\mathbb{N}} \|f_n\|_\infty.$$

Zu $x \in M$ gibt es ein $n_x \in \mathbb{N}$ mit

$$\|f(x) - f_{n_x}(x)\|_Y \leq 1.$$

Damit folgt

$$\begin{aligned}\|f(x)\|_Y &\leq \|f_{n_x}(x)\|_Y + 1 \\ &\leq \|f_{n_x}\|_\infty + 1 \\ &\leq C + 1.\end{aligned}$$

Also ist $f \in B(M,Y)$. Sei nun $\varepsilon > 0$. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit $\|f_n - f_m\|_\infty < \varepsilon$ für alle $n, m \geq n_\varepsilon$. Damit folgt für alle $x \in M$ und $n \geq n_\varepsilon$

$$\begin{aligned}
\|f(x) - f_n(x)\|_Y &= \|\lim_{m\to\infty} f_m(x) - f_n(x)\|_Y \\
&= \lim_{m\to\infty} \|f_m(x) - f_n(x)\|_Y \\
&\leq \limsup_{m\to\infty} \|f_m - f_n\|_\infty \\
&< \varepsilon.
\end{aligned}$$

Also gilt auch $\lim\limits_{n\to\infty} \|f - f_n\|_\infty = 0$. □

Aus Satz V.1.6 folgt das folgende wichtige Konvergenzkriterium. Man beachte, dass dabei die Abbildungen $f_n$ und $f$ nicht in $B(M,Y)$ liegen müssen.

Satz V.1.7 (Cauchykriterium für gleichmässige Konvergenz). *Folgende Aussagen sind äquivalent:*

(1) *Die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig.*
(2) $\forall \varepsilon > 0 \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall n, m \geq N_\varepsilon : \|f_n - f_m\|_\infty < \varepsilon.$

Beweis. (1) $\Longrightarrow$ (2): Es gelte $f_n \xrightarrow[n\to\infty]{\text{glm}} f$. Dann ist $(f_n - f)_{n\in\mathbb{N}}$ eine Nullfolge und damit eine Cauchyfolge in $B(M,Y)$. Wegen

$$(f_n - f) - (f_m - f) = f_n - f_m$$

folgt hieraus die Behauptung.
(2) $\Longrightarrow$ (1): Es gibt ein $N_1 \in \mathbb{N}$ mit

$$\|f_n - f_{N_1}\|_\infty \leq 1 \quad \forall n \geq N_1.$$

Setze zur Abkürzung $\hat{f} = f_{N_1}$. Dann ist $(f_n - \hat{f})_{n\in\mathbb{N}}$ eine Cauchyfolge in $B(M,Y)$. Wegen Satz V.1.6 konvergiert $f_n - \hat{f}$ gleichmäßig gegen ein $\tilde{f} \in B(M,Y)$. Damit folgt

$$f_n \xrightarrow[n\to\infty]{\text{glm}} \tilde{f} + \hat{f}.$$

□

Definition V.1.8. Die Funktionenreihe $\sum f_n$ heißt

(1) punktweise konvergent,
(2) absolut konvergent,
(3) gleichmässig konvergent,
(4) normal konvergent,

wenn gilt

(1) $\sum f_n(x)$ konvergiert für jedes $x \in M$,
(2) $\sum \|f_n(x)\|_Y$ konvergiert für jedes $x \in M$,

(3) die Folge $s_n = \sum_{k=0}^{n} f_k$ der Partialsummen konvergiert gleichmäßig,

(4) $\sum_{n=0}^{\infty} \|f_n\|_\infty < \infty$.

Bemerkung V.1.9. Es gilt:

(1) $\sum f_n$ absolut konvergent $\Longrightarrow \sum f_n$ punktweise konvergent.

(2) $\sum f_n$ gleichmäßig konvergent $\not\Longrightarrow \sum f_n$ absolut konvergent,
$\sum f_n$ absolut konvergent $\not\Longrightarrow \sum f_n$ gleichmäßig konvergent.

(3) $\sum f_n$ absolut und gleichmäßig konvergent $\not\Longrightarrow \sum f_n$ normal konvergent.

Beweis. ad (1): Folgt aus Satz II.5.2 (S. 49).
ad (2): Sei $f_n(x) = \frac{(-1)^n}{n}$ für alle $n \in \mathbb{N}$ und alle $x \in M$, dann gilt $\sum f_n$ konvergiert gleichmäßig aber nicht absolut.
Sei $M = (-1,1)$, $Y = \mathbb{R}$ und $f_n(x) = x^n$. Dann ist $\sum f_n$ absolut konvergent, aber

$$\begin{aligned} \|\sum_{k=n+1}^{m} f_k\|_\infty &= \sup_{-1<x<1} |\sum_{k=n+1}^{m} x^k| \\ &= \sup_{-1<x<1} |\frac{1-x^{m-n}}{1-x} x^{n+1}| \\ &= \lim_{x\to 1-0} |x^{n+1}\frac{1-x^{m-n}}{1-x}| \\ &= m-n. \end{aligned}$$

Also ist $\sum f_n$ nicht gleichmäßig konvergent.
ad (3): Sei $M = Y = \mathbb{R}$ und

$$f_n(x) = \begin{cases} \frac{1}{n} & n \le x < n+1 \\ 0 & \text{sonst.} \end{cases}$$

$\sum f_n$ ist gleichmäßig und absolut konvergent, aber $\sum \|f_n\|_\infty = \sum \frac{1}{n}$ ist divergent. □

Satz V.1.10. *Jede normal konvergente Funktionenreihe konvergiert absolut und gleichmäßig.*

Beweis. Sei $\sum f_n$ normal konvergent. Dann ist für jedes $x \in M$ $\sum \|f_n\|_\infty$ eine konvergente Majorante zu $\sum \|f_n(x)\|_Y$. Wegen Satz II.5.4 (S. 50) ist $\sum f_n$ absolut konvergent. Sei nun $\varepsilon > 0$. Da $\sum \|f_n\|_\infty$ konvergiert, gibt es ein $N_\varepsilon \in \mathbb{N}$ mit

$$\sum_{k=n+1}^{m} \|f_k\|_\infty < \varepsilon \quad \forall n, m \ge N_\varepsilon.$$

Wegen

$$\|\sum_{k=n+1}^{m} f_k\|_\infty \leq \sum_{k=n+1}^{m} \|f_k\|_\infty$$

folgt damit die gleichmäßige Konvergenz von $\sum f_n$ aus Satz V.1.7. $\square$

Beispiel V.1.11. $\sum \frac{\cos(nx)}{n^2}$ konvergiert gleichmäßig und absolut auf $\mathbb{R}$, da $\sum \frac{1}{n^2}$ eine konvergente Majorante ist.

Satz V.1.12. *Sei $\sum a_k(x-x_0)^k$ eine Potenzreihe in $\mathbb{K}$ mit Konvergenzradius $\rho > 0$. Dann konvergiert $\sum a_k(x-x_0)^k$ für jedes $r \in (0, \rho)$ normal und damit gleichmäßig und absolut in $\overline{B(x_0; r)}$.*

Beweis. Sei $r \in (0, \rho)$, $M = \overline{B(x_0; r)}$, $Y = \mathbb{K}$ und $f_k = a_k(x-x_0)^k$. Dann ist

$$\|f_k\|_\infty = |a_k| r^k.$$

Wegen Satz II.6.3 (S. 57) ist

$$\limsup_{k\to\infty} \sqrt[k]{\|f_k\|_\infty} = r \limsup_{k\to\infty} \sqrt[k]{|a_k|} = \frac{r}{\rho} < 1.$$

Damit folgt die Behauptung aus Satz II.5.6 (S. 50). $\square$

## V.2. Vertauschen von Grenzprozessen

Satz V.2.1. *Seien $f, f_n$ Funktionen von $M \subset X$, $M \neq \emptyset$, in $Y$ und $x_0 \in M$. Es gelte*

(1) $f_n \xrightarrow[n\to\infty]{glm} f$,

(2) *$f_n$ ist für jedes $n \in \mathbb{N}$ in $x_0$ stetig.*

*Dann ist $f$ in $x_0$ stetig.*

Beweis. Sei $\varepsilon > 0$ beliebig. Gemäß Bemerkung V.1.5(2) (S. 149) gibt es ein $N_\varepsilon \in \mathbb{N}$ mit

$$\|f_{N_\varepsilon} - f\|_\infty < \frac{\varepsilon}{3}.$$

Zu $f_{N_\varepsilon}$ gibt es eine Umgebung $U \in \mathcal{U}(x_0)$ mit

$$\|f_{N_\varepsilon}(x) - f_{N_\varepsilon}(x_0)\|_Y < \frac{\varepsilon}{3} \quad \forall x \in U \cap M.$$

Damit folgt

$$\begin{aligned}
\|f(x) - f(x_0)\|_Y &\leq \|f(x) - f_{N_\varepsilon}(x)\|_Y + \|f_{N_\varepsilon}(x) - f_{N_\varepsilon}(x_0)\|_Y \\
&\quad + \|f_{N_\varepsilon}(x_0) - f(x_0)\|_Y \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&= \varepsilon \qquad\qquad\qquad \forall x \in U \cap M.
\end{aligned}$$

$\square$

Bemerkung V.2.2. Beispiel V.1.3(1) (S. 147) zeigt, dass die Aussage des Satzes falsch ist, wenn $f_n$ nur punktweise gegen $f$ konvergiert.

Definition V.2.3. Wir sagen $f_n$ konvergiert lokal gleichmässig gegen $f$, wenn es zu jedem $x \in M$ eine Umgebung $U \in \mathcal{U}(x)$ gibt, so dass $f_{n|U\cap M}$ gleichmäßig auf $U \cap M$ gegen $f$ konvergiert.

Bemerkung V.2.4. Es gilt:

(1) $f_n \underset{n\to\infty}{\xrightarrow{\text{glm}}} f \Longrightarrow f_n \underset{n\to\infty}{\xrightarrow{\text{lokal glm}}} f$.

(2) $f_n \underset{n\to\infty}{\xrightarrow{\text{lokal glm}}} f$ und $M$ kompakt $\Longrightarrow f_n \underset{n\to\infty}{\xrightarrow{\text{glm}}} f$.

Beweis. ad (1): Ist klar.
ad (2): Sei $\varepsilon > 0$ beliebig. Zu jedem $x \in M$ gibt es dann ein $U_x \in \mathcal{U}(x)$ und ein $N_{\varepsilon,x} \in \mathbb{N}$ mit

$$\|f_n(y) - f(y)\|_Y < \varepsilon \quad \forall n \geq N_{\varepsilon,x}, y \in U_x \cap M.$$

Da $\{U_x : X \in M\}$ eine offene Überdeckung von $M$ und $M$ kompakt ist, gibt es eine endliche Teilüberdeckung $\{U_{x_i} : 1 \leq i \leq k\}$. Sei

$$N_\varepsilon = \max_{1\leq i\leq k} N_{\varepsilon,x_i}.$$

Dann gilt

$$\|f_n(y) - f(y)\|_Y < \varepsilon \quad \forall n \geq N_\varepsilon, y \in M.$$

□

Satz V.2.5. *Die Folge $(f_n)_{n\in\mathbb{N}} \subset C(M,Y)$ konvergiere lokal gleichmäßig gegen $f$. Dann ist $f \in C(M,Y)$.*

Beweis. Folgt aus Satz V.2.1, da die Stetigkeit eine lokale Eigenschaft ist. □

Bemerkung V.2.6. (1) Seien $f$ und $f_n$ wie in Beispiel V.1.3(2) (S. 147). Dann gilt $f_n, f \in C(M,Y)$ und $f_n \underset{n\to\infty}{\xrightarrow{\text{pkt}}} f$, aber $f_n$ konvergiert in keiner Umgebung von 0 gleichmäßig gegen $f$.
(2) Seien $f_n \in C(M,Y)$, $\sum f_n$ konvergiere lokal gleichmäßig. Dann ist $\sum_{n=0}^{\infty} f_n \in C(M,Y)$.
(3) Seien $f_n \in C(M,Y)$, $\sum f_n$ konvergiere lokal gleichmäßig und es sei $x_0 \in M$. Dann gilt

$$\begin{aligned}\lim_{x\to x_0}\Big(\sum_{n=0}^{\infty} f_n(x)\Big) &= \sum_{n=0}^{\infty}\Big(\lim_{x\to x_0} f_n(x)\Big)\\ &= \sum_{n=0}^{\infty} f_n(x_0).\end{aligned}$$

(4) Jede Potenzreihe stellt im Innern ihres Konvergenzkreises eine stetige Funktion dar.

BEWEIS. AD (1): Ist klar.
AD (2): Folgt aus Satz V.2.5.
AD (3): $f = \sum_{n=0}^{\infty} f_n$ ist stetig.
AD (4): Wegen Satz V.1.12 (S. 152) konvergiert jede Potenzreihe im Innern ihres Konvergenzkreises lokal gleichmäßig. □

DEFINITION V.2.7. Definiere $BC(M, Y) = B(M, Y) \cap C(M, Y)$ und $\|\cdot\|_{BC(M,Y)} = \|\cdot\|_\infty$.

Aus Satz V.2.5 folgt:

SATZ V.2.8. (1) *$BC(M, Y)$ ist ein abgeschlossener Untervektorraum von $B(M, Y)$.*
(2) *Falls $M$ kompakt ist, gilt $BC(M, Y) = C(M, Y)$.*

BEWEIS. AD (1): Dass $BC(M, Y)$ ein Untervektorraum von $B(M, Y)$ ist, ist offensichtlich. Dass $BC(M, Y)$ in $B(M, Y)$ abgeschlossen ist, folgt aus Satz V.2.5 und Satz III.2.14 (S. 71).
AD (2): Aus Satz III.4.10 (S. 86) folgt $C(M, Y) \subset B(M, Y)$. □

Als nächstes möchten wir zeigen, dass unter geeigneten Voraussetzungen der Grenzwert einer Folge differenzierbarer Funktionen ebenfalls differenzierbar ist. Dazu benötigen wir folgenden Hilfssatz.

LEMMA V.2.9. *Sei $-\infty < a < b < +\infty$ und $\varphi : [a, b] \to Y$ differenzierbar. Dann gilt*

$$\|\varphi(b) - \varphi(a) - \varphi'(a)(b - a)\|_Y \leq \sup_{0 \leq x \leq b} \|\varphi'(x) - \varphi'(a)\|_Y |b - a|.$$

BEWEIS. Definiere $\psi : [a, b] \to Y$ durch

$$\psi(x) = \varphi(x) - \varphi'(a)(x - a).$$

Dann ist $\psi$ differenzierbar und

$$\psi'(x) = \varphi'(x) - \varphi'(a)$$
$$\psi(b) - \psi(a) = \varphi(b) - \varphi(a) - \varphi'(a)(b - a).$$

Damit folgt die Behauptung aus Satz IV.2.17 (S. 129) und Bemerkung IV.2.18 (S. 130). □

SATZ V.2.10. *Es sei $M \subset \mathbb{K}$ offen und $(f_n)_{n \in \mathbb{N}} \subset C^1(M, Y)$. Es gelte:*

(1) $f_n \xrightarrow[n\to\infty]{pkt} f$,
(2) $f'_n \xrightarrow[n\to\infty]{lokal\ glm} g$.

*Dann gilt:*

(1) $f \in C^1(M, Y)$,
(2) $f' = g$,

(3) $f_n \xrightarrow[n\to\infty]{\textit{lokal glm}} f$.

Beweis. Sei $x_0 \in M$ und $r > 0$, so dass $B(x_0; r) \subset M$ ist und $f_n'$ auf $B(x_0; r)$ gleichmäßig gegen $g$ konvergiert. Sei $x \in B(x_0; r)$ beliebig. Definiere

$$\varphi_n(t) = f_n(x_0 + t(x - x_0)).$$

Dann ist $\varphi_n \in C^1([0,1], Y)$ und

$$\begin{aligned}
\varphi_n(0) &= f_n(x_0) \\
\varphi_n(1) &= f_n(x) \\
\varphi_n'(t) &= f_n'(x_0 + t(x - x_0))(x - x_0).
\end{aligned}$$

Aus Lemma V.2.9 folgt

$$\begin{aligned}
&\|f_n(x) - f_n(x_0) - f_n'(x_0)(x - x_0)\|_Y \\
=&\|\varphi_n(1) - \varphi_n(0) - \varphi_n'(0)\|_Y \\
\le& \sup_{0\le t\le 1} \|\varphi_n'(t) - \varphi_n'(0)\|_Y \\
=& \sup_{0\le t\le 1} \|f_n'(x_0 + t(x - x_0)) - f_n'(x_0)\|_Y |x - x_0|.
\end{aligned}$$

Wegen $f_n \xrightarrow{\text{pkt}} f$ und $f_n' \xrightarrow{\text{glm}} g$ auf $B(x_0; r)$ folgt

$$\begin{aligned}
&\|f(x) - f(x_0) - g(x_0)(x - x_0)\|_Y \\
\le& \sup_{0\le t\le 1} \|g(x_0 + t(x - x_0)) - g(x_0)\|_Y |x - x_0|.
\end{aligned}$$

Wegen $f_n' \in C(M, Y)$ und $f_n' \xrightarrow{\text{lokal glm}} g$ folgt aus Satz V.2.5 $g \in C(M, Y)$. Daher gilt

$$\begin{aligned}
&\lim_{x - x_0} \frac{\|f(x) - f(x_0) - g(x_0)(x - x_0)\|_Y}{|x - x_0|} \\
\le& \lim_{x\to x_0} \sup_{0\le t\le 1} \|g(x_0 + t(x - x_0)) - g(x_0)\|_Y \\
=&0.
\end{aligned}$$

Also ist $f$ in $x_0$ differenzierbar und $f'(x_0) = g(x_0)$. Da $x_0 \in M$ beliebig war, sind damit (1) und (2) bewiesen.
Weiter gilt für $x_0 \in M$, $r \in \mathbb{R}_+^*$ mit $B(x_0; r) \subset M$ und $x \in B(x_0; r)$

$$\begin{aligned}
&\|f_n(x) - f(x)\|_Y \\
\le&\|f_n(x) - f(x) - f_n(x_0) + f(x_0)\|_Y + \|f_n(x_0) - f(x_0)\|_Y \\
\le& \sup_{0\le t\le 1} \|f_n'(x_0 + t(x - x_0)) - f'(x_0 + t(x - x_0))\|_Y |x - x_0| \\
&+ \|f_n(x_0) - f(x_0)\|_Y \\
\le& r\|f_n' - f'\|_{\infty, B(x_0;r)} + \|f_n(x_0) - f(x_0)\|_Y,
\end{aligned}$$

Wegen $f' = g$ und $f_n' \xrightarrow{\text{lokal glm}} g$ und $f_n \xrightarrow{\text{pkt}} f$ folgt hieraus (3). $\square$

Ganz analog gilt:

SATZ V.2.11. *Sei $M \subset \mathbb{K}$ offen und $(f_n)_{n\in\mathbb{N}} \subset C^1(M,Y)$. Es gelte:*
(1) *$\sum f_n$ konvergiert punktweise,*
(2) *$\sum f_n'$ konvergiert lokal gleichmäßig.*

*Dann gilt:*
(1) *$\sum f_n$ konvergiert lokal gleichmäßig,*
(2) *$f = \sum_{n=0}^{\infty} f_n \in C^1(M,Y)$,*
(3) *$f' = \sum_{n=0}^{\infty} f_n'$.*

BEWEIS. Folgt aus Satz V.2.10 angewandt auf die Folge der Partialsummen. □

BEMERKUNG V.2.12. (1) $f_n, f \in C^1(M,Y)$ und $f_n \xrightarrow{\text{glm}} f$, impliziert nicht die Konvergenz von $f_n'$ gegen $f'$. Dies zeigt das Beispiel $M = Y = \mathbb{R}$

$$f_n(x) = \frac{1}{n}\sin(nx), \;\; f(x) = 0.$$

(2) $f_n \in C^1(M,Y)$ und $\sum f_n$ gleichmäßig konvergent impliziert nicht die gleichmäßige Konvergenz von $\sum f_n'$. Dies zeigt das Beispiel $M = Y = \mathbb{R}$

$$f_n(x) = \frac{1}{n^2}\sin(nx).$$

(3) Satz V.2.10 gilt auch unter folgenden Voraussetzungen:
(1) $M$ ist ein Gebiet,
(2) $f_n : M \to Y$ differenzierbar,
(3) $f_n' \xrightarrow{\text{lokal glm}} g$.

(4) Satz V.2.10 gilt auch, wenn $M$ konvex und perfekt ist, z.B. für $M = [a,b)$ mit $a < b$.

## V.3. Analytische Funktionen

Wir beginnen mit einer Aussage über die gliedweise Differentiation von Potenzreihen.

SATZ V.3.1. *Sei $\sum a_n(x-x_0)^n$ eine Potenzreihe in $\mathbb{K}$ mit Konvergenzradius $\rho > 0$. Dann hat die Potenzreihe $\sum na_n(x-x_0)^{n-1}$ ebenfalls den Konvergenzradius $\rho$ und es gilt*

$$\left(\sum_{n=0}^{\infty} a_n(x-x_0)^n\right)' = \sum_{n=0}^{\infty} na_n(x-x_0)^{n-1}, \quad \forall x \in B(x_0,\rho),$$

*d.h., Potenzreihen dürfen im Innern des Konvergenzkreises gliedweise differenziert werden.*

Beweis. Sei $\rho'$ der Konvergenzradius von $\sum n a_n(x-x_0)^n$. Dann folgt aus Satz II.6.3 (S. 57) und Beispiel II.2.3(3) (S. 35)

$$\rho' = \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{n|a_n|}} = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{n}} \cdot \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}} = \rho.$$

Damit folgt die Behauptung aus Bemerkung V.2.6(4) (S. 153) und Satz V.2.11 (S. 156). □

Aus Satz V.3.1 folgt unmittelbar:

Satz V.3.2. *Sei $\sum a_n(x-x_0)^n$ eine Potenzreihe mit Konvergenzradius $\rho > 0$ und*

$$f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n \quad \forall x \in B(x_0;\rho).$$

*Dann ist $f \in C^\infty(B(x_0;\rho),\mathbb{K})$ und $\sum a_n(x-x_0)^n = T(f,x_0)$, d.h., die Potenzreihe mit Entwicklungspunkt $x_0$ stimmt mit der Taylorreihe mit Entwicklungspunkt $x_0$ der durch sie dargestellten Funktion überein.*

Definition V.3.3. Sei $M \subset \mathbb{K}$, $M \neq \emptyset$, offen. Dann heißt eine Funktion $f : M \to \mathbb{K}$ $\mathbb{K}$-analytisch, wenn es zu jedem $x_0 \in M$ eine Potenzreihe mit Entwicklungspunkt $x_0$ und Konvergenzradius $\rho > 0$ gibt, die $f$ in einer Umgebung von $x_0$ darstellt. Die Gesamtheit aller $\mathbb{K}$-analytischen Funktionen von $M$ in $\mathbb{K}$ wird mit $C^\omega(M,\mathbb{K})$ bezeichnet.

Bemerkung V.3.4. (1) Sei $f \in C^\omega(M,\mathbb{K})$ und $x_0 \in M$. Dann ist die $f$ darstellende Potenzreihe mit Entwicklungspunkt $x_0$ eindeutig bestimmt und stimmt mit der Taylorreihe $T(f,x_0)$ überein.
(2) $C^\omega(M,\mathbb{K}) \underset{\neq}{\subset} C^\infty(M,\mathbb{K})$. Die Funktion $f$ aus Beispiel IV.1.21 (S. 120) erfüllt

$$f \in C^\infty(\mathbb{R},\mathbb{R}) \backslash C^\omega(\mathbb{R},\mathbb{R}).$$

(3) Die Analytizität ist eine lokale Eigenschaft, d.h., es ist $f \in C^\omega(M,\mathbb{K})$ genau dann, wenn es zu jedem $x_0 \in M$ eine Umgebung $U \in \mathcal{U}(x_0)$ gibt mit $f|_U \in C^\omega(U,\mathbb{K})$.
(4) $C^\omega(M,\mathbb{K})$ ist ein $\mathbb{K}$-Vektorraum. Für $f,g \in C^\omega(M,\mathbb{K})$ ist auch $f \cdot g \in C^\omega(M,\mathbb{K})$ und $\frac{f}{g} \in C^\omega(M,\mathbb{K})$, sofern zusätzlich $g(x) \neq 0$ ist für alle $x \in M$.

Beispiel V.3.5. (1) Polynome sind analytisch.
(2) $f(x) = \frac{1}{x} \in C^\omega(\mathbb{K}^*,\mathbb{K})$.

Beweis. ad (1): Sei $p(x) = \sum\limits_{k=0}^{n} a_k x^k$ und $x_0$ beliebig. Dann folgt

$$p(x) = \sum_{k=0}^{n} a_k(x-x_0+x_0)^k$$

$$= \sum_{k=0}^{n} \sum_{l=0}^{k} a_k \binom{k}{l} (x - x_0)^l x_0^{k-l}$$
$$= \sum_{l=0}^{n} \left\{ \sum_{k=l}^{n} a_k \binom{k}{l} x_0^{k-l} \right\} (x - x_0)^l.$$

AD (2): Sei $x_0 \in \mathbb{K}^*$. Dann folgt für $x \in \mathbb{K}$ mit $|x - x_0| < |x_0|$

$$\frac{1}{x} = \frac{1}{x - x_0 + x_0} = \frac{1}{x_0} \frac{1}{1 + \frac{x - x_0}{x_0}}$$
$$= \frac{1}{x_0} \sum_{k=0}^{\infty} (-1)^k x_0^{-k} (x - x_0)^k$$
$$= \sum_{k=0}^{\infty} (-1)^k x_0^{-1-k} (x - x_0)^k.$$

□

SATZ V.3.6. *Konvergente Potenzreihen stellen im Innern des Konvergenzkreises analytische Funktionen dar.*

BEWEIS. Sei $\sum a_n (x - x_0)^n$ eine Potenzreihe mit Konvergenzradius $\rho > 0$ und

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n \quad \forall x \in B(x_0; \rho).$$

Sei $y_0 \in B(x_0; \rho)$ beliebig und $r = \frac{1}{2}(\rho - |x_0 - y_0|) > 0$. Dann gilt

$$\frac{1}{n!} f^{(n)}(y_0) = \frac{1}{n!} \sum_{k=n}^{\infty} a_k k \cdot (k-1) \cdot \ldots \cdot (k - n + 1)(y_0 - x_0)^{k-n}$$
$$= \sum_{k=n}^{\infty} a_k \binom{k}{n} (y_0 - x_0)^{k-n}.$$

Also ist

$$T(f, y_0)(x) = \sum \left\{ \sum_{k=n}^{\infty} a_k \binom{k}{n} (y_0 - x_0)^{k-n} \right\} (x - y_0)^n.$$

Wir zeigen zunächst, dass $T(f, y_0)$ in $\overline{B(y_0; r)}$ normal konvergiert. Es ist

$$\sum_{n=0}^{\infty} \| \{ \sum_{k=n}^{\infty} a_k \binom{k}{n} (y_0 - x_0)^{k-n} \} (x - y_0)^n \|_{\infty, \overline{B(y_0; r)}}$$
$$\leq \sum_{n=0}^{\infty} \{ \sum_{k=n}^{\infty} |a_k| \binom{k}{n} |y_0 - x_0|^{k-n} \} r^n$$
$$(*) \qquad = \sum_{k=0}^{\infty} |a_k| \left\{ \sum_{n=0}^{k} \binom{k}{n} |y_0 - x_0|^{k-n} r^n \right\}$$

$$
\begin{aligned}
&= \sum_{k=0}^{\infty} |a_k|(r + |y_0 - x_0|)^k \\
&= \sum_{k=0}^{\infty} |a_k|(\underbrace{\frac{1}{2}(\rho + |y_0 - x_0|)}_{<\rho})^k \\
&< \infty.
\end{aligned}
$$

Dabei haben wir an der Stelle $(*)$ ausgenutzt, dass eine absolut konvergente Doppelreihe beliebig umgeordnet werden kann (Beweis siehe Heuser Satz 45.1 bzw. 45.2).
Da $T(f, y_0)(x)$ insbesondere für jedes $x \in B(y_0; r)$ absolut konvergiert, erhalten wir mit dem gleichen Argument

$$
\begin{aligned}
T(f, y_0)(x) &= \sum_{n=0}^{\infty} \left\{ \sum_{k=0}^{\infty} a_k \binom{k}{n} (y_0 - x_0)^{k-n} \right\} (x - y_0)^n \\
&= \sum_{k=0}^{\infty} a_k \left\{ \sum_{n=0}^{k} \binom{k}{n} (y_0 - x_0)^{k-n} (x - y_0)^n \right\} \\
&= \sum_{k=0}^{\infty} a_k (x - x_0)^k \\
&= f(x).
\end{aligned}
$$

Also konvergiert $T(f, y_0)$ in $B(y_0; r)$ und stellt dort $f$ dar. Da $y_0 \in B(x_0; \rho)$ beliebig war folgt hieraus die Behauptung. □

Bemerkung V.3.7. Aus Satz V.3.6 folgt:

(1) $\exp, \sin, \cos \in C^\omega(\mathbb{K}, \mathbb{K})$.
(2) $f \in C^\omega(M, \mathbb{K}) \Longrightarrow f' \in C^\omega(M, \mathbb{K})$.

Definition V.3.8. Sei $M \subset \mathbb{K}$ offen und $f : M \to Y$ eine Abbildung. Dann heißt eine Abbildung $F : M \to Y$ Stammfunktion von $f$ genau dann, wenn gilt $F' = f$.

Bemerkung V.3.9. (1) Ist $M \subset \mathbb{K}$ ein Gebiet, so unterscheiden sich zwei Stammfunktionen von $f$ höchstens um eine additive Konstante.
(2) Ist

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

in $B(x_0; \rho)$, so ist

$$
F(x) = \sum_{n=0}^{\infty} \frac{1}{n+1} a_n (x - x_0)^{n+1}
$$

eine Stammfunktion von $f$ in $B(x_0; \rho)$.

Beweis. ad (1): Seien $F_1$ und $F_2$ zwei Stammfunktionen von $f$. Dann gilt

$$(F_1 - F_2)' = 0 \text{ in } M.$$

Sei $x_0 \in M$ beliebig,

$$c = F_1(x_0) - F_2(x_0)$$

und

$$N = \{x \in M : F_1(x) - F_2(x) = c\}.$$

Dann ist $N \neq \emptyset$. Sei $x \in N$ und $r \in \mathbb{R}_+^*$ so, dass $B(x;r) \subset M$ ist. Dann folgt für $y \in B(x;r)$ und

$$\varphi(t) = F_1(x + t(y-x)) - F_2(x + t(y-x)) \quad \forall t \in [0,1]$$

die Identität

$$\varphi'(t) = 0 \quad \forall t \in [0,1].$$

Aus Satz IV.2.17 (S. 129) und Bemerkung IV.2.18 (S. 130) folgt

$$\begin{aligned}\varphi(1) &= F_1(y) - F_2(y) \\ &= \varphi(0) \\ &= F_1(x) - F_2(x) \\ &= c.\end{aligned}$$

Also ist $N$ offen in $M$. Da $F_1 - F_2$ stetig ist, ist $N$ aber auch abgeschlossen in $M$. Da $M$ zusammenhängend ist, folgt $N = M$, d.h., $F_1 - F_2$ ist konstant auf $M$.
ad (2): Folgt direkt aus Satz V.3.1. □

Satz V.3.10. *Ist $f \in C^\omega(M,\mathbb{K})$ und $F$ eine Stammfunktion von $f$, so ist $F \in C^\omega(M,\mathbb{K})$.*

Beweis. Sei $x_0 \in M$ und

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n \quad \text{in} \quad B(x_0;r)$$

mit $r > 0$. Dann folgt aus Bemerkung V.3.9

$$F(x) = c + \sum_{n=0}^{\infty} \frac{1}{n+1} a_n (x-x_0)^{n+1}$$

mit einem geeigneten $c \in \mathbb{K}$. Hieraus folgt die Behauptung. □

Beispiel V.3.11. (1) $\ln \in C^\omega(\mathbb{R}_+^*, \mathbb{R})$ und

$$\ln(x) = \ln(x_0) + \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)x_0^{n+1}} (x-x_0)^{n+1} \quad \forall x_0 \in \mathbb{R}_+^*, |x - x_0| < x_0.$$

(2) Für $\alpha \in \mathbb{R}$ sei $\binom{\alpha}{0} = 1$ und für $n \in \mathbb{N}^*$

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}.$$

Es ist

(Binomialreihe) $$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \quad \forall |x| < 1.$$

Insbesondere ist $x^\alpha \in C^\omega(\mathbb{R}_+^*, \mathbb{R})$ und

$$x^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x_0^{\alpha-n}(x-x_0)^n \quad \forall x_0 \in \mathbb{R}_+^*, |x-x_0| < x_0.$$

(3) Sei $\arcsin : (-1,1) \to (-\frac{\pi}{2}, \frac{\pi}{2})$ die Umkehrfunktion von $\sin|_{(-\frac{\pi}{2}, \frac{\pi}{2})}$. Dann ist $\arcsin \in C^\omega((-1,1) : \mathbb{R})$ und

$$\arcsin(x) = \sum_{n=0}^{\infty} \frac{1 \cdot 3 \cdot \dots \cdot (2n-1)}{2 \cdot 4 \cdot \dots \cdot (2n)} \frac{x^{2n+1}}{2n+1} \quad \forall |x| < 1.$$

Beweis. ad (1): Folgt aus Beispiel V.3.5 (2) und Satz V.3.10 zusammen mit $\ln' = \frac{1}{x}$.
ad (2): Es ist

$$\lim_{n\to\infty} \left| \frac{\binom{\alpha}{n}}{\binom{\alpha}{n+1}} \right| = \lim_{n\to\infty} \frac{n+1}{|\alpha - n|} = 1.$$

Also hat die Reihe $\sum \binom{\alpha}{n} x^n$ gemäß Satz II.6.4 (S. 58) den Konvergenzradius 1. Für $|x| < 1$ sei

$$g(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n.$$

Aus Satz V.3.1 folgt für $|x| < 1$

$$\begin{aligned}
g'(x) &= \sum_{n=0}^{\infty} \binom{\alpha}{n} n x^{n-1} \\
&= \sum_{n=1}^{\infty} \binom{\alpha}{n} n x^{n-1} \\
&= \sum_{m=0}^{\infty} \binom{\alpha}{m+1} (m+1) x^m \\
&= \sum_{m=0}^{\infty} \frac{\alpha \cdot (\alpha-1) \cdot \dots \cdot (\alpha - m)}{m!} x^m \\
&= \alpha \sum_{m=0}^{\infty} \binom{\alpha-1}{m} x^m.
\end{aligned}$$

Damit folgt

$$\begin{aligned}(1+x)g'(x) &= \alpha \sum_{m=0}^{\infty} \binom{\alpha-1}{m} x^m (1+x) \\ &= \alpha \left[ \sum_{m=0}^{\infty} \binom{\alpha-1}{m} x^m + \sum_{k=1}^{\infty} \binom{\alpha-1}{k-1} x^k \right] \\ &= \alpha \left[ 1 + \sum_{m=1}^{\infty} \left[ \binom{\alpha-1}{m} + \binom{\alpha-1}{m-1} \right] x^m \right] \\ &= \alpha \left[ 1 + \sum_{m=1}^{\infty} \binom{\alpha}{m} x^m \right] \\ &= \alpha g(x)\end{aligned}$$

und somit

$$\begin{aligned}0 &= (1+x)^{-\alpha-1}[(1+x)g'(x) - \alpha g(x)] \\ &= (1+x)^{-\alpha} g'(x) - \alpha(1+x)^{-\alpha-1} g(x) \\ &= \Big((1+x)^{-\alpha} g(x)\Big)'.\end{aligned}$$

Also ist $(1+x)^{-\alpha} g(x)$ auf $(-1,1)$ konstant. Wegen

$$1^\alpha = 1 = g(0)$$

folgt

$$g(x) = (1+x)^\alpha.$$

Für $x_0 \in \mathbb{R}_+^*$ und $|x - x_0| < x_0$ folgt nun

$$\begin{aligned}x^\alpha &= (x - x_0 + x_0)^\alpha \\ &= (\frac{x-x_0}{x_0} + 1)^\alpha x_0^\alpha \\ &= x_0^\alpha \sum_{n=0}^{\infty} \binom{\alpha}{n} (\frac{x-x_0}{x_0})^n \\ &= \sum_{n=0}^{\infty} \binom{\alpha}{n} x_0^{\alpha-n} (x-x_0)^n.\end{aligned}$$

AD (3): Da sin auf $(-\frac{\pi}{2}, \frac{\pi}{2})$ streng monoton wachsend ist, existiert arcsin. Wegen

$$\sin' = \cos = \sqrt{1 - \sin^2} \quad \text{auf} \quad (-\frac{\pi}{2}, \frac{\pi}{2})$$

folgt

$$\arcsin' = \frac{1}{\sqrt{1-x^2}}.$$

Aus Teil (2) folgt mit $\alpha = -\frac{1}{2}$ für $|x| < 1$

$$(1+x)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} x^n.$$

Nun ist für $n \in \mathbb{N}^*$

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

Substitution $x \to -x^2$ liefert

$$\frac{1}{\sqrt{1-x^2}} = \sum_{n=0}^{\infty} \frac{1 \cdot 3 \cdot \ldots \cdot (2n-1)}{2 \cdot 4 \cdot \ldots \cdot (2n)} x^{2n}.$$

Zusammen mit Satz V.3.10 folgt hieraus die Behauptung. □

Satz V.3.12 (Identitätssatz für analytische Funktionen). *Sei $M \subset \mathbb{K}$ ein Gebiet und $f, g \in C^\omega(M, \mathbb{K})$. Es gebe eine nicht leere, offene Teilmenge $U$ von $M$ mit $f|_U = g|_U$. Dann ist $f = g$ auf $M$.*

Beweis. Sei $h = f - g$. Dann ist $h \in C^\omega(M, \mathbb{K})$ und $h|_U = 0$. Wir müssen zeigen $h = 0$. Sei

$$N = \{x \in M : \exists\, U \in \mathcal{U}(x) : h|_U = 0\}.$$

Nach Voraussetzung ist $N \neq \emptyset$.
Konstruktionsgemäß ist $N$ offen.
Um zu zeigen, dass $N$ abgeschlossen ist, sei $x_0 \in M$ ein Häufungspunkt von $N$ und $(x_n)_{n\in\mathbb{N}} \subset M \backslash \{x_0\}$ eine Folge in $N$ mit $\lim\limits_{n\to\infty} x_n = x_0$. Wegen $h \in C^\omega(M, \mathbb{K})$ gibt es ein $\rho > 0$ mit

$$h(x) = \sum_{n=0}^{\infty} a_k (x - x_0)^k \quad \forall x \in B(x_0; \rho).$$

O.E. ist $(x_n)_{n\in\mathbb{N}} \subset B(x_0; \rho)$. Wir nehmen an, es sei $\{k \in \mathbb{N} : a_k \neq 0\} \neq \emptyset$. Sei $K = \min\{k : a_k \neq 0\}$. Dann gilt für $x \in B(x_0; \rho) \backslash \{x_0\}$

$$\begin{aligned}\frac{h(x)}{(x-x_0)^K} &= \sum_{k=K}^{\infty} a_k (x - x_0)^{K-k} \\ &= \sum_{l=0}^{\infty} a_{K+l} (x - x_0)^l.\end{aligned}$$

Da $\sum a_{k+l}(x-x_0)^l$ in $B(x_0;\rho)$ eine analytische Funktion darstellt, also insbesondere stetig ist, folgt

$$a_K = \lim_{n\to\infty} \frac{h(x_n)}{(x_n - x_0)^K} = 0.$$

Also ist $h|_{B(x_0;\rho)} = 0$ und damit $x_0 \in N$.
Da $M$ zusammenhängend ist, folgt $N = M$. $\square$

Bemerkung V.3.13. (1) Aus dem Beweis von Satz V.3.12 ergibt sich folgende Variante:
Ist $M \subset \mathbb{K}$ ein Gebiet, $f, g \in C^\omega(M,\mathbb{K})$ und gibt es $x_0 \in M$ und $(x_n)_{n\in\mathbb{N}} \subset M\backslash\{x_0\}$ mit

$$x_n \underset{n\to\infty}{\longrightarrow} x_0 \quad \text{und} \quad f(x_n) = g(x_n) \quad \forall n \in \mathbb{N},$$

dann ist $f = g$.
(2) Beispiel IV.1.21 (S. 120) zeigt, dass Satz V.3.12 für $C^\infty$-Funktionen nicht gilt.

Zum Abschluss dieses Paragraphen beweisen wir noch einige Eigenschaften komplex-analytischer Funktionen, die für reell-analytische Funktionen nicht gelten. Weitere, wie z.B. die Beziehung

$$C^1(M,\mathbb{C}) = C^\omega(M,\mathbb{C}) \quad \forall M \subset \mathbb{C}, M \text{ Gebiet},$$

werden wir in den Kapiteln VII und VIII kennen lernen.

Satz V.3.14 (Maximumprinzip). *Sei $M \subset \mathbb{C}$ ein Gebiet und $f \in C^\omega(M,\mathbb{C})$. Es gebe ein $z_0 \in M$ mit $|f(z_0)| \geq |f(z)|$ für alle $z \in M$. Dann ist $f$ konstant.*

Beweis. Es gibt ein $\rho > 0$ mit $B(z_0;\rho) \subset M$ und

$$f(z) = \sum_{n=0}^{\infty} a_n(z-z_0) \quad \forall z \in B(z_0;\rho).$$

Wir nehmen an, dass $\{n \in \mathbb{N}^* : a_n \neq 0\} \neq \emptyset$ ist. Sei $k = \min\{n \in \mathbb{N}^* : a_n \neq 0\}$. Dann ist

$$f(z) = a_0 + a_k(z-z_0)^k + (z-z_0)^{k+1}g(z) \quad \forall z \in B(z_0;\rho)$$

mit

$$g(z) = \sum_{l=0}^{\infty} a_{k+1+l}(z-z_0)^l \quad \forall z \in B(z_0;\rho).$$

Sei $c = \max\limits_{z\in \overline{B(z_0;\frac{\rho}{2})}} |g(z)|$. Dann gilt für $z \in B(z_0;\frac{\rho}{2})$

$$|f(z)| \geq |a_0 + a_k(z-z_0)^k| - c|z-z_0|^{k+1}.$$

Sei

$$a_0 = r_0 e^{i\alpha}, \; a_k = r_k e^{i\beta}$$

mit $r_0, r_k, \alpha, \beta \in \mathbb{R}$ und $r_k = |a_k| > 0$. Sei $s \in (0, \frac{\rho}{2})$ und

$$z = z_0 + se^{i\frac{\alpha-\beta}{k}}.$$

Dann ist $z \in B(z_0; \frac{\rho}{2})$ und

$$\begin{aligned}|f(z)| &\geq |r_0 e^{i\alpha} + r_k e^{i\beta} s^k e^{i(\alpha-\beta)}| - cs^{k+1} \\ &= |r_0 + r_k s^k| - cs^{k+1} \\ &\geq r_0 + r_k s^k - cs^{k+1} \\ &= r_0 + s^k(r_k - cs).\end{aligned}$$

Für

$$0 < s \leq \min\{\frac{\rho}{2}, \frac{r_k}{2c}\}$$

folgt

$$|f(z)| \geq r_0 + s^k \frac{1}{2} r_k > r_0 = |f(z_0)|.$$

Dies ist ein Widerspruch.
Also ist $f$ auf $B(z_0; \rho)$ konstant und damit gemäß Satz V.3.12 auf ganz $M$ konstant. □

Satz V.3.15. *Sei $M \subset \mathbb{C}$ ein Gebiet, $f \in C^\omega(M, \mathbb{C})$ und $K \subset M$ nicht leer und kompakt. Dann nimmt $f|_K$ das Maximum seines Betrages auf dem Rand $\partial K$ an.*

Beweis. Gemäß Satz III.4.10 (S. 86) gibt es ein $z_0 \in K$ mit

$$|f(z_0)| \geq |f(z)| \quad \forall z \in K.$$

Falls $z_0 \in \partial K$ ist, sind wir fertig.
Sei also $z_0 \in \overset{\circ}{K}$. Dann ist gemäß Satz III.4.10 (S. 86)

$$(**) \qquad s = \min\{|z - z_0| : z \in \partial K\} > 0.$$

Es gilt

$$B(z_0; s) \subset \overset{\circ}{K} \quad \text{und} \quad \overline{B(z_0; s)} \subset K.$$

Wegen Satz V.3.14 ist $f$ auf $B(z_0; s)$ konstant gleich $f(z_0)$. Wegen der Stetigkeit gilt dies auch auf $\overline{B(z_0; s)}$. Da das Minimum in $(**)$ an einem Punkt $z_1 \in \partial K$ angenommen wird, folgt $f(z_0) = f(z_1)$. □

Satz V.3.16 (Minimumprinzip). *Sei $M \subset \mathbb{C}$ ein Gebiet und $f \in C^\omega(M, \mathbb{C})$.*

(1) *Es gebe ein $z_0 \in M$ mit*

$$|f(z)| \geq |f(z_0)| > 0 \quad \forall z \in M.$$

*Dann ist $f$ konstant.*

(2) *Ist $K \subset M$ nicht leer und kompakt und $f(z) \neq 0$ für alle $z \in M$, so nimmt $f|_K$ das Minimum seines Betrages auf $\partial K$ an.*

Beweis. Folgt aus den Sätzen V.3.14 und V.3.15 angewandt auf $\frac{1}{f}$. $\square$

Satz V.3.17 (Satz von der Gebietstreue). *Sei $M \subset \mathbb{C}$ ein Gebiet und $f \in C^\omega(M, \mathbb{C})$ nicht konstant. Dann ist $f(M)$ ein Gebiet.*

Beweis. Wegen Satz III.5.4 (S. 89) müssen wir noch zeigen, dass $f(M)$ offen ist. Sei also $z_0 \in M$ und $w_0 = f(z_0)$. Wegen Bemerkung V.3.13 (1) gibt es ein $r \in \mathbb{R}_+^*$ mit $\overline{B(z_0; r)} \subset M$ und

$$f(z) \neq f(z_0) \quad \forall z \in \overline{B(z_0; r)} \backslash \{z_0\}.$$

Daher ist

$$\rho = \inf\{|f(z) - w_0| : z \in \partial B(z_0; r)\} > 0.$$

Sei nun $w \notin f(M)$. Wegen Satz V.3.16 ist

$$\begin{aligned} |w - w_0| &= |f(z_0) - w| \\ &\geq \inf\{|f(z) - w| : z \in \overline{B(z_0; r)})\} \\ &= |f(z_1) - w| \end{aligned}$$

mit einem geeigneten $z_1 \in \partial B(z_0; r)$. Weiter gilt

$$\begin{aligned} \rho &\leq |f(z_1) - w_0| \\ &\leq |f(z_1) - w| + |w - w_0| \\ &\leq 2|w - w_0|, \end{aligned}$$

also $|w - w_0| \geq \frac{1}{2}\rho$. Mithin ist $B(w_0; \frac{1}{2}\rho) \subset f(M)$. $\square$

## V.4. Sprungstetige Funktionen

Im Folgenden sei stets $I \subset \mathbb{R}$ ein perfektes Intervall mit $\alpha = \inf I \in \overline{\mathbb{R}}$, $\beta = \sup I \in \overline{\mathbb{R}}$ und $(Y, \|\cdot\|_Y)$ ein Banachraum.

Definition V.4.1. (1) Eine Menge $Z = (\alpha_0, \ldots, \alpha_n)$, $n \in \mathbb{N}^*$, mit $\alpha = \alpha_0 < \alpha_1 < \ldots < \alpha_n = \beta$ heißt Zerlegung von $I$.
(2) Eine Funktion $f : I \to Y$ heißt Treppenfunktion, wenn es eine Zerlegung $Z = (\alpha_0, \ldots, \alpha_n)$ von $I$ gibt, so dass $f|_{(\alpha_{i-1}, \alpha_i)}$ für jedes $1 \leq i \leq n$ konstant ist.
(3) $T(I, Y)$ ist die Menge aller Treppenfunktionen.
(4) Eine Funktion $f : I \to Y$ heißt sprungstetig, wenn für jedes $x \in I$ die links- und rechtsseitigen Grenzwerte $f(x - 0)$ und $f(x + 0)$ existieren.
(5) $S(I, Y)$ ist die Menge aller sprungstetigen Funktionen.

Beispiel V.4.2. (1) Die Funktion

$$f(x) = \begin{cases} -1 & -1 < x < 0 \\ 0 & x = 0 \\ 1 & 0 < x < 1 \end{cases}$$

ist in $T((-1,1),\mathbb{R})$.
(2) Die Funktion

$$f(x) = \begin{cases} 1-x^2 & -1 \le x < 0 \\ \frac{1}{2} & x = 0 \\ x^2 & 0 < x \le 1 \end{cases}$$

ist in $S([-1,1],\mathbb{R})$.
(3) Jede monotone Funktion $f : I \to \mathbb{R}$ ist sprungstetig.
(4) Die Funktion

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R}\backslash\mathbb{Q} \end{cases}$$

ist nicht sprungstetig.

Bemerkung V.4.3. (1) $T(I,Y)$ ist ein Untervektorraum von $B(I,Y)$. Es ist $T(I,Y) \subset S(I,Y)$.
(2) $S(I,Y)$ ist ein Vektorraum. Es ist $C(I,Y) \subset S(I,Y)$.
(3) Falls $I$ kompakt ist, ist $S(I,Y) \subset B(I,Y)$.

Beweis. ad (1): $T(I,Y) \subset S(I,Y)$ ist offensichtlich. Ebenso, dass aus $f \in T(I,Y)$, $\lambda \in \mathbb{K}$ folgt $\lambda f \in T(I,Y)$. Seien also $f, g \in T(I,Y)$ und $Z_f = (\alpha_0 \ldots, \alpha_n)$ und $Z_g = (\beta_0, \ldots, \beta_m)$ Zerlegungen von $I$ mit

$$f|_{(\alpha_{i-1},\alpha_i)}, g|_{(\beta_{j-1},\beta_i)} \quad 1 \le i \le n, 1 \le j \le m$$

sind konstant. Sei $Z = Z_f \cup Z_g = (\gamma_0, \ldots, \gamma_K)$ mit der natürlichen Anordnung. Dann folgt

$$(f+g)|_{(\gamma_{l-1},\gamma_l)} \quad \text{ist konstant.}$$

ad (2): $C(I,Y) \subset S(I,Y)$ ist klar. Die Vektorraumeigenschaft folgt aus der Linearität der einseitigen Grenzwerte.
ad (3): Sei $I$ kompakt und $f \in S(I,Y)$. Aufgrund der Definition der einseitigen Grenzwerte gibt es zu jedem $x \in I$ ein $\alpha(x) < x$ und ein $\beta(x) > x$ mit

$$\|f(s) - f(t)\|_Y \le 1 \quad \forall s, t \in (\alpha(x), x) \text{ oder } s, t \in (x, \beta(x)).$$

Da $I$ kompakt ist, gibt es ein $n \in \mathbb{N}^*$ und $x_1, \ldots, x_n \in I$ mit

$$I \subset \bigcup_{1 \le i \le n} (\alpha(x_i), \beta(x_i)).$$

Damit folgt

$$\begin{aligned} \|f\|_\infty \le \max_{1 \le i \le n} \max\Big\{ &\|f(x_i)\|_Y, \\ &1 + \|f(\frac{x_i + \alpha(x_i)}{2})\|_Y, \\ &1 + \|f(\frac{x_i + \beta(x_i)}{2})\|_Y \Big\} \end{aligned}$$

$$< \infty.$$

□

Wir kommen nun zum Hauptsatz dieses Paragraphen.

Satz V.4.4. *Sei $I$ kompakt. Dann sind folgende Aussagen äquivalent:*

(1) $f \in S(I,Y)$.
(2) $\exists (f_n)_{n\in\mathbb{N}} \subset T(I,Y) : \|f_n - f\|_\infty \underset{n\to\infty}{\longrightarrow} 0$.

Beweis. (1) $\Longrightarrow$ (2): Wie im Beweis von Bemerkung V.4.3 (3) folgt, dass es zu jedem $n \in \mathbb{N}$ endlich viele Punkte $x_1, \ldots, x_m \in I$ und Zahlen $\alpha(x_i) < x_i < \beta(x_i), 1 \le i \le m$, gibt mit

$$I \subset \bigcup_{1\le i\le m} (\alpha(x_i), \beta(x_i))$$

und

$$\begin{aligned}\|f(s) - f(t)\|_Y \le \frac{1}{n} \quad &\forall s,t \in (\alpha(x_i), x_i) \text{ oder} \\ &s,t \in (x_i, \beta(x_i)), 1 \le i \le m.\end{aligned}$$

Also gibt es eine Zerlegung $Z = (\gamma_0, \ldots, \gamma_k)$ von $I$ mit

$$\|f(s) - f(t)\|_Y \le \frac{1}{n} \quad \forall s,t \in (\gamma_{i-1}, \gamma_i), 1 \le i \le k.$$

Definiere $f_n \in T(I,Y)$ durch

$$f_n(x) = \begin{cases} f(\frac{\gamma_{i-1}+\gamma_i}{2}) & \gamma_{i-1} < x < \gamma_i, 1 \le i \le k \\ f(x) & x = \gamma_i, 1 \le i \le k. \end{cases}$$

Dann folgt

$$\|f_n - f\|_\infty \le \frac{1}{n}.$$

(2) $\Longrightarrow$ (1): Sei $(f_n)_{n\in\mathbb{N}} \subset T(I,Y)$ mit $\|f_n - f\|_\infty \underset{n\to\infty}{\longrightarrow} 0$. Sei $\varepsilon > 0$. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit

$$\|f_{n_\varepsilon} - f\|_\infty \le \frac{\varepsilon}{2}.$$

Sei $x \in I$. Dann gibt es Zahlen $\alpha, \beta \in I$, $\alpha < \beta$ mit

$$f_{n_\varepsilon}(s) = f_{n_\varepsilon}(t) \quad \forall s,t \in (\alpha, x) \text{ oder } s,t \in (x, \beta).$$

Damit folgt

$$\begin{aligned}\|f(s) - f(t)\|_Y &\le \|f(s) - f_{n_\varepsilon}(s)\|_Y + \|f_{n_\varepsilon}(t) - f(t)\|_Y \\ &\le \varepsilon \qquad \forall s,t \in (\alpha, x) \text{ oder } s,t \in (x, \beta).\end{aligned}$$

Sei nun $(x_n)_{n\in\mathbb{N}} \subset I$ mit $x_n < x$ für alle $n \in \mathbb{N}$ und $x_n \underset{n\to\infty}{\longrightarrow} x$. Dann gibt es ein $m \in \mathbb{N}$ mit $x_k \in (\alpha, x)$ für alle $k \ge m$ und somit

$$\|f(x_k) - f(x_l)\|_Y \le \varepsilon \quad \forall k,l \ge m.$$

Also ist $(f(x_n))_{n\in\mathbb{N}} \subset Y$ eine Cauchyfolge und damit konvergent. Setze $z = \lim\limits_{n\to\infty} f(x_n)$. Sei nun $(y_n)_{n\in\mathbb{N}} \subset I$ eine andere Folge mit $y_n < x$ für alle $n \in \mathbb{N}$ und $x = \lim\limits_{n\to\infty} y_n$. Dann folgt wieder, dass $(f(y_n))_{n\in\mathbb{N}} \subset Y$ konvergiert. Sei $z' = \lim\limits_{n\to\infty} f(y_k)$. Dann folgt

$$\|z - z'\|_Y \le \varepsilon.$$

Also ist $z = z'$ und $f(x-0)$ existiert. Analog zeigt man die Existenz von $f(x+0)$. Also ist $f \in S(I,Y)$. □

Bemerkung V.4.5. Aus dem Beweis von Satz V.4.4 ergibt sich, dass man für $f \in S(I,\mathbb{R})$ mit $f \ge 0$ eine Folge $(f_n)_{n\in\mathbb{N}} \subset T(I,\mathbb{R})$ mit $\|f_n - f\|_\infty \to 0$ und $f_n \ge 0$ für alle $n \in \mathbb{N}$ finden kann.

Definition V.4.6. Seien $A, B$ Teilmengen eines normierten Vektorraumes $(X, \|\cdot\|_X)$ mit $A \subset B$. Dann heißt $A$ dicht in $B$, wenn gilt $B \subset \overline{A}$.

Bemerkung V.4.7. (1) Ist $A$ dicht in $B$, so ist jeder Punkt von $B$ Häufungspunkt von $A$, kann also beliebig gut durch Punkte aus $A$ approximiert werden.
(2) Falls $B$ abgeschlossen ist, gilt

$$A \text{ dicht in } B \iff \overline{A} = B.$$

(3) $\mathbb{Q}$ und $\mathbb{R}\backslash\mathbb{Q}$ sind dicht in $\mathbb{R}$.

Wir können Satz V.4.4 somit auch folgendermaßen formulieren:

Satz V.4.8. *Sei $I$ kompakt. Dann ist $S(I,Y)$ ein in der Norm von $B(I,Y)$ abgeschlossener Unterraum von $B(I,Y)$ und $T(I,Y)$ ist dicht in $S(I,Y)$.*

Bemerkung V.4.9. (1) $S(I,Y)$ ist ein Banachraum bzgl. $\|\cdot\|_\infty$. Insbesondere sind gleichmäßige Grenzwerte sprungstetiger Funktionen wieder sprungstetig.
(2) $T(I,Y)$ ist ein Beispiel eines nicht vollständigen normierten Vektorraumes.
(3) Jede stetige Funktion auf einem kompakten Intervall ist gleichmäßiger Grenzwert von Treppenfunktionen.

KAPITEL VI

# Integralrechnung einer Variablen

In diesem Kapitel führen wir das (Riemann-) Integral für Funktionen einer (reellen) Variablen ein. Dies geschieht zunächst für Treppenfunktionen und danach mit Hilfe eines allgemeinen Satzes über die Fortsetzung linearer Operatoren für sprungstetige Funktionen. Danach leiten wir einige Eigenschaften des Integrals und wichtige Integrationsregeln her und erweiterten diese Ergebnisse auf sog. uneigentliche Integrale, d.h., Integrale mit unbeschränktem Integrationsbereich und/oder singulären Integranden. Als wichtiges Beispiel für uneigentliche Integrale untersuchen wir abschließend eingehend die Eulersche Gammafunktion.

## VI.1. Das (Riemann-) Integral

Im Folgenden sei stets $(X, \|\cdot\|)$ ein $\mathbb{K}$-Banachraum und $I = [\alpha, \beta]$ ein kompaktes perfektes Intervall, d.h., $\alpha < \beta$.

Als erstes definieren wir das Integral einer Treppenfunktion.

Definition VI.1.1. Seien $f \in T(I, X)$ eine Treppenfunktion und $Z = (\alpha_0, \ldots, \alpha_n)$ eine zugehörige Zerlegung. Dann ist das Integral von $f$ bzgl. $Z$ definiert durch

$$\int_{(Z)} f = \sum_{i=1}^{n} (\alpha_i - \alpha_{i-1}) \cdot f|_{(\alpha_{i-1}, \alpha_i)}.$$

Lemma VI.1.2. *Sei $f \in T(I, X)$. Dann ist $\int_{(Z)} f$ unabhängig von der speziellen Zerlegung $Z$.*

Beweis. Seien $f \in T(I, X)$ und $Z = (\alpha_0, \ldots, \alpha_n)$ eine zugehörige Zerlegung.

1. Schritt: Sei $Z'$ eine Verfeinerung von $Z$. O.E. ist

$$Z' = (\alpha_0, \ldots, \alpha_k, \gamma, \alpha_{k+1}, \ldots, \alpha_n).$$

Dann folgt

$$\begin{aligned}\int_{(Z)} f = &\sum_{i=1}^{k} (\alpha_i - \alpha_{i-1}) f|_{(\alpha_{i-1}, \alpha_i)} \\ &+ (\alpha_{k+1} - \gamma + \gamma - \alpha_k) f|_{(\alpha_k, \alpha_{k+1})} \\ &+ \sum_{i=k+2}^{n} (\alpha_i - \alpha_{i-1}) f|_{(\alpha_{i-1}, \alpha_i)}\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{k} (\alpha_i - \alpha_{i-1}) f|_{(\alpha_{i-1},\alpha_i)} \\
&\quad + (\gamma - \alpha_k) f|_{(\alpha_k,\gamma)} + (\alpha_{k+1} - \gamma) f|_{(\gamma,\alpha_{k+1})} \\
&\quad + \sum_{i=k+2}^{n} (\alpha_i - \alpha_{i-1}) f|_{(\alpha_{i-1},\alpha_i)} \\
&= \int_{(Z')} f.
\end{aligned}
$$

2. Schritt: Sei nun $Z'$ eine beliebige, für $f$ zulässige Zerlegung. Dann ist $Z'' = Z \cup Z'$ mit der natürlichen Anordnung eine Verfeinerung von $Z$ und von $Z'$. Mit Schritt 1 folgt

$$
\int_{(Z)} f = \int_{(Z'')} f = \int_{(Z')} f.
$$

□

Wegen Lemma VI.1.2 ist die folgende Definition sinnvoll.

Definition VI.1.3. Sei $f \in T(I, X)$. Dann ist das (Riemann-)Integral von $f$ definiert durch

$$
\int_I f = \int_\alpha^\beta f = \int_{(Z)} f,
$$

wobei $Z$ eine beliebige, für $f$ zulässige Zerlegung von $I$ ist.

Eine erste, für das Folgende wichtige Eigenschaft des Integrals liefert:

Lemma VI.1.4. *Die Abbildung* $\displaystyle\int_\alpha^\beta : T(I, X) \to X$ *mit* $f \mapsto \displaystyle\int_\alpha^\beta f$ *ist linear. Weiter gilt*

$$
\| \int_\alpha^\beta f \| \le (\beta - \alpha) \|f\|_\infty \quad \forall f \in T(I, X).
$$

Beweis. Die Linearität des Integrals folgt unmittelbar aus seiner Definition.
Sei $f \in T(I, X)$ und $Z = (\alpha_0, \ldots, \alpha_n)$ eine zugehörige Zerlegung von $I$. Dann folgt

$$
\begin{aligned}
\| \int_\alpha^\beta f \| &= \| \int_{(Z)} f \| \\
&= \| \sum_{i=1}^{n} (\alpha_i - \alpha_{i-1}) f|_{(\alpha_{i-1},\alpha_i)} \| \\
&\le \sum_{i=1}^{n} (\alpha_i - \alpha_{i-1}) \| f|_{(\alpha_{i-1},\alpha_i)} \|
\end{aligned}
$$

$$\leq (\beta - \alpha)\|f\|_\infty.$$

□

Als nächstes wollen wir das Integral für sprungstetige Funktionen definieren. Dazu benötigen wir einige Vorbereitungen.

Definition VI.1.5. Seien $(Y, \|\cdot\|_Y)$ und $(Z, \|\cdot\|_Z)$ normierte $\mathbb{K}$-Vektorräume.

(1) Eine lineare Abbildung $A : Y \to Z$ heißt beschränkt, wenn es ein $\alpha \in \mathbb{R}_+$ gibt mit

$$\|Ax\|_Z \leq \alpha\|x\|_Y \quad \forall x \in Y.$$

(2) $\mathcal{L}(Y, Z) = \{A : Y \to Z : A \text{ ist beschränkt}\}$ ist ein $\mathbb{K}$-Vektorraum.

(3) Durch

$$\begin{aligned}\|A\| &= \|A\|_{\mathcal{L}(Y,Z)} \\ &= \inf\{\alpha \in \mathbb{R}_+ : \|Ax\|_Z \leq \alpha\|x\|_Y \ \forall x \in Y\}\end{aligned}$$

wird eine Norm auf $\mathcal{L}(Y, Z)$ definiert.

Bemerkung VI.1.6. Es gilt:

(1) Sei $A \in \mathcal{L}(Y, Z)$. Dann ist

$$\begin{aligned}\|A\|_{\mathcal{L}(Y,Z)} &= \sup_{x \in Y \setminus \{0\}} \frac{\|Ax\|_Z}{\|x\|_Y} \\ &= \sup_{\substack{x \in Y \\ \|x\|_Y \leq 1}} \|Ax\|_Z \\ &= \sup_{\substack{x \in Y \\ \|x\|_Y = 1}} \|Ax\|_Z.\end{aligned}$$

(2) Sei $A \in \mathcal{L}(Y, Z)$. Dann gilt

$$\|Ax\|_Z \leq \|A\|_{\mathcal{L}(Y,Z)}\|x\|_Y \quad \forall x \in Y.$$

(3) Sei $A : Y \to Z$ linear. Dann gilt

$$\begin{aligned}&A \text{ ist gleichmäßig stetig} \\ \iff &A \text{ ist stetig} \\ \iff &A \text{ ist in } 0 \in Y \text{ stetig} \\ \iff &A \in \mathcal{L}(Y, Z)\end{aligned}$$

(4) $\displaystyle\int_\alpha^\beta \in \mathcal{L}(T(I, X), X)$ und

$$\Big\| \int_\alpha^\beta \Big\|_{\mathcal{L}(T(I,X),X)} \leq \beta - \alpha.$$

BEWEIS. AD(1):

$$\begin{aligned}
\|A\|_{\mathcal{L}(Y,Z)} &= \inf\{\alpha \in \mathbb{R}_+ : \|Ax\|_Z \le \alpha \|x\|_Y \ \forall x \in Y\} \\
&= \inf\{\alpha \in \mathbb{R}_+ : \frac{\|Ax\|_Z}{\|x\|_Y} \le \alpha \ \forall x \in Y\backslash\{0\}\} \\
&= \sup_{x\in Y\backslash\{0\}} \frac{\|Ax\|_Z}{\|x\|_Y} \\
&= \sup_{x\in Y\backslash\{0\}} \|A(\frac{x}{\|x\|_Y})\|_Z \\
&= \sup_{\substack{y\in Y \\ \|y\|_Y=1}} \|Ay\|_Z \\
&\le \sup_{\substack{y\in Y \\ \|y\|_Y\le 1}} \|Ay\|_Z .
\end{aligned}$$

Sei andererseits $x \in Y$ mit $0 < \|x\|_Y < 1$. Dann folgt

$$\begin{aligned}
\|Ax\|_Z &\le \frac{1}{\|x\|_Y}\|Ax\|_Z \\
&= \|A(\frac{x}{\|x\|_Y})\|_Z \\
&\le \sup_{\substack{y\in Y \\ \|y\|_Y=1}} \|Ay\|_Z
\end{aligned}$$

also

$$\sup_{\substack{x\in Y \\ \|x\|_Y\le 1}} \|Ax\|_Z \le \sup_{\substack{y\in Y \\ \|y\|_Y=1}} \|Ay\|_Z .$$

Hieraus folgt die Behauptung.
AD(2): Folgt direkt aus Teil (1).
AD(3):

$$\begin{aligned}
&A \text{ gleichmäßig stetig} \\
\Longrightarrow &A \text{ stetig} \\
\Longrightarrow &A \text{ stetig in } 0 \\
\Longrightarrow &\exists \delta > 0 \quad \forall x \in Y \quad \|x\|_Y \le \delta : \|Ax\|_Z \le 1 \\
\Longrightarrow &\exists \delta > 0 \quad \forall x \in Y \quad \|x\|_Y = 1 : \|Ax\|_Z \le \frac{1}{\delta} \\
\Longrightarrow &A \in \mathcal{L}(Y,Z) \text{ und } \|A\|_{\mathcal{L}(Y,Z)} \le \frac{1}{\delta} \\
\Longrightarrow &\|Ax - Ay\|_Z = \|A(x-y)\|_Z \le \|A\|_{\mathcal{L}(X,Z)}\|x-y\|_Y \\
\Longrightarrow &A \text{ gleichmäßig stetig.}
\end{aligned}$$

AD(4): Folgt aus der Definition VI.1.5 und Hilfssatz VI.1.4. □

Lemma VI.1.7. *Sei $(Y, \|\cdot\|_Y)$ ein normierter $\mathbb{K}$-Vektorraum und $U \subset Y$ ein Untervektorraum. Dann ist $\overline{U}$ auch ein Untervektorraum.*

Beweis. Seien $u, v \in \overline{U}$ und $\alpha \in \mathbb{K}$. Dann gibt es Folgen $(u_n)_{n\in\mathbb{N}}$ und $(v_n)_{n\in\mathbb{N}}$ in $U$ mit

$$u = \lim_{n\to\infty} u_n, \quad v = \lim_{n\to\infty} v_n.$$

Dann folgt

$$\begin{aligned} u + v &= \lim_{n\to\infty} u_n + \lim_{n\to\infty} v_n \\ &= \lim_{n\to\infty} (u_n + v_n) \in \overline{U} \\ \alpha u &= \alpha \lim_{n\to\infty} u_n \\ &= \lim_{n\to\infty} (\alpha u_n) \in \overline{U}. \end{aligned}$$

Also ist $\overline{U}$ ein Untervektorraum von $Y$. □

Satz VI.1.8 (Über die Fortsetzung stetiger linearer Operatoren). *Seien $(Y, \|\cdot\|_Y)$ ein normierter $\mathbb{K}$-Vektorraum, $U \subset Y$ ein Untervektorraum, $(Z, \|\cdot\|_Z)$ ein $\mathbb{K}$-Banachraum und $A \in \mathcal{L}(U, Z)$. Dann gibt es genau ein $\overline{A} \in \mathcal{L}(\overline{U}, Z)$ mit*

(1) $\overline{A}y = Ay$ *für alle* $y \in U$,
(2) $\|\overline{A}\|_{\mathcal{L}(\overline{U},Z)} = \|A\|_{\mathcal{L}(U,Z)}$,
(3) $\overline{A}y = \lim_{n\to\infty} Ay_n$ *für alle* $(y_n)_{n\in\mathbb{N}} \subset U$ *mit* $\lim_{n\to\infty} y_n = y$.

$\overline{A}$ *heißt* Forsetzung *von* $A$.

Beweis. Eindeutigkeit: Seien $B, C \in \mathcal{L}(\overline{U}, Z)$ zwei lineare Operatoren, die (1) erfüllen. Dann folgt aus ihrer Stetigkeit für jedes $y \in \overline{U}$ mit $y = \lim_{n\to\infty} y_n$ und $(y_n)_{n\in\mathbb{N}} \subset U$:

$$\begin{aligned} By &= B(\lim_{n\to\infty} y_n) \\ &= \lim_{n\to\infty} By_n \\ &= \lim_{n\to\infty} Ay_n \\ &= \lim_{n\to\infty} Cy_n \\ &= C(\lim_{n\to\infty} y_n) \\ &= Cy. \end{aligned}$$

Also ist $B = C$.
Existenz: Sei $y \in \overline{U}$ und $(y_n)_{n\in\mathbb{N}} \subset U$ mit $y = \lim_{n\to\infty} y_n$. Dann ist $(y_n)_{n\in\mathbb{N}}$ eine Cauchyfolge und

$$\|Ay_n - Ay_m\|_Z \le \|A\|_{\mathcal{L}(U,Z)} \|y_n - y_m\|_Y.$$

Also ist $(Ay_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $Z$. Mithin existiert ein $z \in Z$ mit
$$z = \lim_{n\to\infty} Ay_n.$$
Definiere
$$\overline{A}y = z.$$
Sei nun $(\widehat{y}_n)_{n\in\mathbb{N}} \subset U$ eine weitere Folge mit $y = \lim\limits_{n\to\infty} \widehat{y}_n$. Sei
$$\widehat{z} = \lim_{n\to\infty} A\widehat{y}_n.$$
Dann folgt
$$\begin{aligned}\|z-\widehat{z}\|_Z &\le \|z - Ay_n\|_Z + \|Ay_n - A\widehat{y}_n\|_Z + \|\widehat{z} - A\widehat{y}_n\|_Z \\ &\le \|z - Ay_n\|_Z + \|A\|_{\mathcal{L}(U,Z)}\|y_n - \widehat{y}_n\|_Z \\ &\quad + \|\widehat{z} - A\widehat{y}_n\|_Z \\ &\underset{n\to\infty}{\longrightarrow} 0.\end{aligned}$$
Mithin ist $\overline{A}$ wohldefiniert.
Seien nun $u, v \in \overline{U}, \alpha \in \mathbb{K}$ und $(u_n)_{n\in\mathbb{N}} \subset U$, $(v_n)_{n\in\mathbb{N}} \subset U$ mit
$$u = \lim_{n\to\infty} u_n \quad , \quad v = \lim_{n\to\infty} v_n.$$
Dann folgt
$$\begin{aligned}\overline{A}(u+v) &= \overline{A}(\lim_{n\to\infty}(u_n+v_n)) \\ &= \lim_{n\to\infty} A(u_n+v_n) \\ &= \lim_{n\to\infty}[Au_n + Av_n] \\ &= \lim_{n\to\infty} Au_n + \lim_{n\to\infty} Av_n \\ &= \overline{A}u + \overline{A}v \\ \alpha\overline{A}u &= \alpha \lim_{n\to\infty} Au_n \\ &= \lim_{n\to\infty} A(\alpha u_n) \\ &= \overline{A}(\alpha u).\end{aligned}$$
Also ist $\overline{A}$ linear.
Sei $u \in U$. Dann ist $u = \lim\limits_{n\to\infty} u_n$ mit $u_n = u$ für alle $n \in \mathbb{N}$. Daraus folgt
$$\overline{A}u = Au \quad \forall u \in U.$$
Sei schließlich $y \in \overline{U}$ mit $y = \lim\limits_{n\to\infty} y_n$ und $(y_n)_{n\in\mathbb{N}} \subset U$. Dann folgt
$$\begin{aligned}\|\overline{A}y\|_Z &= \|\lim_{n\to\infty} Ay_n\|_Z \\ &= \lim_{n\to\infty} \|Ay_n\|_Z \\ &\le \lim_{n\to\infty} \|A\|_{\mathcal{L}(U,Z)}\|y_n\|_Z \\ &= \|A\|_{\mathcal{L}(U,Z)}\|y\|_Y\end{aligned}$$

Also ist

$$\|\overline{A}\|_{\mathcal{L}(U,Z)} \le \|A\|_{\mathcal{L}(U,Z)}.$$

Wegen (1) gilt aber andererseits

$$\|\overline{A}\|_{\mathcal{L}(\overline{U},Z)} \ge \sup_{\substack{u\in U\\ \|u\|_Y=1}} \|\overline{A}u\|_Z = \|A\|_{\mathcal{L}(U,Z)}.$$

Damit ist die Behauptung gezeigt. □

Gemäß Satz V.4.8 (S. 169) ist $S(I,X)$ ein abgeschlossener Unterraum von $B(I,X)$ und $S(I,X) = \overline{T(I,X)}$ bzgl. $\|\cdot\|_\infty$. Wir können daher Satz VI.1.8 auf $Y = B(I,X)$, $U = T(I,X)$, $\overline{U} = S(I,X)$, $Z = X$ und $A = \int_\alpha^\beta$ anwenden. Dies liefert:

Definition VI.1.9. Das (Riemann-) Integral ist die eindeutig definierte Fortsetzung von $\int_\alpha^\beta$ von $\mathcal{L}(T(I,X),X)$ auf $\mathcal{L}(S(I,X),X)$ und wird wieder mit $\int_\alpha^\beta$ bezeichnet. Für $f \in S(I,X)$ verwenden wir simultan die folgenden Bezeichnungen

$$\int_\alpha^\beta f = \int_I f = \int_\alpha^\beta f(x)dx = \int_\alpha^\beta f dx.$$

Aus Bemerkung VI.1.6 und Satz VI.1.8 folgt unmittelbar:

Bemerkung VI.1.10. (1) $\|\int_\alpha^\beta f\| \le (\beta-\alpha)\|f\|_\infty$ für alle $f \in S(I,X)$.
(2) $\int_\alpha^\beta f = \lim_{n\to\infty} \int_\alpha^\beta f_n$, wobei $(f_n)_{n\in\mathbb{N}} \subset T(I,X)$ eine beliebige Folge von Treppenfunktionen mit $\lim_{n\to\infty} \|f - f_n\|_\infty = 0$ ist.

Der folgende Satz ist für die folgenden Untersuchungen des Riemann-Integrals und für die praktische Berechnung des Integrals wesentlich.

Satz VI.1.11. *Sei $f \in S(I,X)$. Dann gibt es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon, f) > 0$, so dass für jede Zerlegung $Z = (x_0,\ldots,x_n)$ von $I$ der Feinheit $\Delta_Z = \max_{1\le i\le n}(x_i - x_{i-1}) < \delta$ und je $n$ Zwischenpunkte $\zeta_j \in [x_{j-1}, x_j], 1 \le j \le n$, gilt*

$$\|\int_\alpha^\beta f - \sum_{j=1}^n f(\zeta_j)(x_j - x_{j-1})\| < \varepsilon.$$

Beweis. 1. Schritt: Sei $f \in T(I,X)$, $\widehat{Z} = (\widehat{x}_0, \ldots, \widehat{x}_{\widehat{n}})$ eine beliebige, aber im Folgenden feste, zu $f$ gehörige Zerlegung von $I$, $\varepsilon > 0$ beliebig und

$$\delta = \delta(\varepsilon, f) = \frac{\varepsilon}{4\widehat{n}\|f\|_\infty}.$$

Seien $Z = (x_0, \ldots, x_n)$ eine beliebige Zerlegung von $I$ und $\zeta_j \in [x_{j-1}, x_j]$, $1 \le j \le n$, beliebige Zwischenpunkte mit

$$\Delta_Z = \max_{1 \le j \le n} (x_j - x_{j-1}) < \delta.$$

Sei

$$Z' = \widehat{Z} \cup Z = (y_0, \ldots, y_m).$$

Definiere

$$\begin{aligned} &e_j' = f|_{(y_{j-1}, y_j)} \quad 1 \le j \le m \\ &e_j'' = f(\zeta_k) \quad \text{falls } [y_{j-1}, y_j] \subset [x_{k-1}, x_k] \ 1 \le j \le m, 1 \le k \le n. \end{aligned}$$

Offensichtlich ist $e_j' \ne e_j''$ für ein $j \in \mathbb{N}_m^*$ höchstens dann, wenn für das entsprechende $\zeta_k$ gilt $\zeta_k = \widehat{x}_l$ für ein $l \in \mathbb{N}_{\widehat{n}}$. Weiter gilt

$$\begin{aligned} &\int_\alpha^\beta f - \sum_{k=1}^n f(\zeta_k)(x_k - x_{k-1}) \\ &= \sum_{j=1}^m (y_j - y_{j-1}) f|_{(y_{j-1}, y_j)} - \sum_{k=1}^n f(\zeta_k)(x_k - x_{k-1}) \\ &= \sum_{j=1}^m e_j'(y_j - y_{j-1}) - \sum_{j=1}^m e_j''(y_j - y_{j-1}) \\ (*) \qquad &= \sum_{j=1}^m (e_j' - e_j'')(y_j - y_{j-1}). \end{aligned}$$

Wegen obiger Beobachtung sind in $(*)$ höchstens $2\widehat{n}$ Summanden von 0 verschieden. Außerdem gilt für jeden Summanden

$$\begin{aligned} \|(y_j - y_{j-1})(e_j' - e_j'')\| &\le |y_j - y_{j-1}|\{\|e_j'\| + \|e_j''\|\} \\ &< \delta 2\|f\|_\infty \\ &\le \frac{\varepsilon}{2\widehat{n}}. \end{aligned}$$

Hieraus folgt

$$\| \int_\alpha^\beta f - \sum_{k=1}^n f(\zeta_k)(x_k - x_{k-1}) \| < \varepsilon.$$

2. Schritt: Sei nun $f \in S(I,X)$ und $\varepsilon > 0$. Dann gibt es ein $g \in T(I,X)$ mit

$$\|f - g\|_\infty < \frac{\varepsilon}{3(\beta - \alpha)}.$$

Gemäß Schritt 1 gibt es ein $\delta = \delta(\frac{\varepsilon}{3}, g)$, so dass für $g$ und $\frac{\varepsilon}{3}$ die Behauptung gilt. Seien nun $Z = (x_0, \dots, x_n)$ eine Zerlegung von $I$ mit Feinheit $\Delta_Z < \delta$ und $\zeta_j \in [x_{j-1}, x_j]$, $1 \le j \le n$, beliebige Zwischenpunkte. Dann folgt

$$\begin{aligned}
&\| \int_\alpha^\beta f - \sum_{j=1}^n f(\zeta_j)(x_j - x_{j-1}) \| \\
\le &\| \int_\alpha^\beta f - \int_\alpha^\beta g \| + \| \sum_{j=1}^n g(\zeta_j)(x_j - x_{j-1}) - \int_\alpha^\beta g \| \\
&+ \| \sum_{j=1}^n [g(\zeta_j) - f(\zeta_j)](x_j - x_{j-1}) \| \\
\le &(\beta - \alpha) \| f - g \|_\infty + \frac{\varepsilon}{3} + \sum_{j=1}^n \| f(\zeta_j) - g(\zeta_j) \| (x_j - x_{j-1}) \\
< &(\beta - \alpha) \frac{\varepsilon}{3(\beta - \alpha)} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3(\beta - \alpha)} (\beta - \alpha) \\
= &\varepsilon.
\end{aligned}$$

Damit ist die Behauptung bewiesen. □

Bemerkung VI.1.12. Sei $f : I \to X$ eine Funktion. Sie heißt Riemann-integrierbar, wenn es ein $x^* \in X$ gibt, so dass zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon, f) > 0$ existiert, so dass für jede Zerlegung $Z = (x_0, \dots, x_n)$ der Feinheit $\Delta_Z < \delta$ und je $n$ Zwischenpunkte $\zeta_i \in [x_{i-1}, x_i]$, $1 \le i \le n$, gilt

$$\| x^* - \sum_{i=1}^n f(\zeta_i)(x_i - x_{i-1}) \| < \varepsilon.$$

Die Größe $x^*$ ist, sofern sie existiert, eindeutig bestimmt und heißt Riemann-Integral von $f$. Die „approximierenden Summen“

$$\sum_{i=1}^n f(\zeta_i)(x_i - x_{i-1})$$

heißen auch Riemann-Summen, und man schreibt häufig symbolisch

$$\int_\alpha^\beta f = \lim_{\Delta_Z \to 0} \sum_{i=1}^n f(\zeta_i)(x_i - x_{i-1}).$$

Satz VI.1.11 zeigt, dass jede sprungstetige Funktion Riemann-integrierbar ist und dass der in diesem Paragraphen eingeführte Integralbegriff mit dem oben definierten für sprungstetige Funktionen übereinstimmt.

Es gibt jedoch Riemann-integrierbare Funktionen, die nicht sprungstetig sind. Ein Beispiel ist $f : [-1,1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 1 & \text{falls } x = \frac{1}{n}, n \in \mathbb{N}^* \\ -1 & \text{falls } x = -\frac{1}{n}, n \in \mathbb{N}^* \\ 0 & \text{sonst.} \end{cases}$$

## VI.2. Eigenschaften des Integrals

Im Folgenden bezeichnet wieder $(X, \|\cdot\|)$ einen $\mathbb{K}$-Vektorraum und $I = [\alpha, \beta]$ mit $\alpha < \beta$ ein kompaktes perfektes Intervall.

Wir beginnen mit einem Satz über die gliedweise Integration gleichmäßig konvergenter Folgen und Reihen.

SATZ VI.2.1. (1) *Sei $(f_n)_{n\in\mathbb{N}} \subset S(I,X)$ gleichmäßig konvergent gegen $f$. Dann ist $f \in S(I,X)$ und*

$$\int_\alpha^\beta f = \lim_{n\to\infty} \int_\alpha^\beta f_n.$$

(2) *Sei $(f_n)_{n\in\mathbb{N}} \subset S(I,X)$ und $\sum f_n$ gleichmäßig konvergent. Dann ist $\sum_{n=0}^{\infty} f_n \in S(I,X)$ und*

$$\int_\alpha^\beta \Big(\sum_{n=0}^{\infty} f_n\Big) = \sum_{n=0}^{\infty} \Big(\int_\alpha^\beta f_n\Big).$$

BEWEIS. AD(1): Wegen Bemerkung V.1.5 (S. 149) gilt $\|f_n - f\|_\infty \to 0$ und wegen Satz V.1.6 (S. 149) ist $f \in B(I,X)$. Da nach Satz V.4.8 (S. 169) $S(I,X)$ ein abgeschlossener Unterraum von $B(I,X)$ ist, folgt $f \in S(I,X)$. Aus Bemerkung VI.1.10 (S. 177) folgt

$$\Big\|\int_\alpha^\beta f - \int_\alpha^\beta f_n\Big\| \le (\beta-\alpha)\|f - f_n\|_\infty \underset{n\to\infty}{\longrightarrow} 0.$$

Dies beweist die Behauptung.
AD(2): Folgt aus Teil (1) angewandt auf

$$s_n = \sum_{k=0}^{n} f_k \quad \text{und} \quad s = \sum_{k=0}^{\infty} f_k.$$

□

BEMERKUNG VI.2.2. Satz VI.2.1 gilt nicht bei punktweiser Konvergenz. Betrachte z. B. $I = [0,1]$, $X = \mathbb{R}$ und $f_n \in T(I,X)$ mit

$$f_n(x) = \begin{cases} 0 & \text{für } x = 0 \\ n & \text{für } 0 < x \le \frac{1}{n} \\ 0 & \text{für } \frac{1}{n} < x \le 1 \end{cases} \qquad , n \in \mathbb{N}^*.$$

Dann konvergiert $f_n$ punktweise gegen 0, aber es ist

$$\int_0^1 f_n = 1 \qquad \forall n \in \mathbb{N}^*.$$

Satz VI.2.3. *Sei $f \in S(I,X)$. Dann ist $\|f\| \in S(I,\mathbb{R})$ und*

$$\| \int_\alpha^\beta f \| \le \int_\alpha^\beta \|f\| \le (\beta - \alpha)\|f\|_\infty.$$

Beweis. Sei $(f_n)_{n\in\mathbb{N}} \subset T(I,X)$ eine Folge von Treppenfunktionen, die gleichmäßig gegen $f$ konvergiert. Dann ist $\|f_n\| \in T(I,\mathbb{R})$ für alle $n \in \mathbb{N}$. Für jedes $x \in I$ gilt

$$\begin{aligned} \Big| \|f_n\|(x) - \|f\|(x) \Big| &= \Big| \|f_n(x)\| - \|f(x)\| \Big| \\ &\le \|f_n(x) - f(x)\| \\ &\le \|f_n - f\|_\infty. \end{aligned}$$

Also konvergiert $\|f_n\|$ gleichmäßig gegen $\|f\|$. Mithin ist $\|f\| \in S(I,\mathbb{R})$. Sei $Z_n = (x_{0,n}, \ldots, x_{m,n})$ eine zulässige Zerlegung von $I$ für $f_n$. Dann folgt für jedes $n \in \mathbb{N}$

$$\begin{aligned} \| \int_\alpha^\beta f_n \| &= \| \sum_{k=1}^m f_n|_{(x_{k-1,n}, x_{k,n})} (x_{k,n} - x_{k-1,n}) \| \\ &\le \sum_{k=1}^m \|f_n|_{(x_{k-1,n}, x_{k,n})}\| (x_{k,n} - x_{k-1,n}) \\ &= \int_\alpha^\beta \|f_n\| \\ (*) \qquad &\le (\beta - \alpha) \|f_n\|_\infty. \end{aligned}$$

Wegen $f_n \xrightarrow{\text{glm}} f$, $\|f_n\| \xrightarrow{\text{glm}} \|f\|$ und $\|f_n\|_\infty \to \|f\|_\infty$ folgt die Behauptung mit Satz VI.2.1 aus $(*)$ durch Grenzübergang $n \to \infty$. □

Als nächstes definieren wir das orientierte Integral und zeigen seine Additivität bzgl. der disjunkten Vereinigung von Intervallen.

Definition VI.2.4. Sei $f \in S(I,X)$. Dann definieren wir

$$\int_\beta^\alpha f = - \int_\alpha^\beta f \qquad \text{und} \qquad \int_\alpha^\alpha f = 0.$$

Satz VI.2.5. *Seien $f \in S(I,X)$ und $a,b,c \in I$. Dann ist*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Beweis. O.E. sei $a \le b \le c$. Sei $(f_n)_{n\in\mathbb{N}} \subset T(I,X)$ mit $f_n \xrightarrow{\text{glm}} f$. Dann gilt für jedes kompakte Teilintervall $J$ von $I$

$$T(J,X) \ni f_n|_J \xrightarrow{\text{glm}} f|_J \in S(J,X).$$

Wie man leicht nachrechnet, gilt für jedes $n \in \mathbb{N}$

$$\int_a^c f_n = \int_a^b f_n + \int_b^c f_n.$$

Mit Satz VI.2.1 folgt hieraus

$$\begin{aligned} \int_a^c f &= \int_a^b f + \int_b^c f \\ \Longrightarrow \int_a^b f &= \int_a^c f - \int_b^c f = \int_a^c f + \int_c^b f. \end{aligned}$$

□

Als nächstes untersuchen wir Monotonie-Eigenschaften des Integrals.

Satz VI.2.6. (1) *Sei $f \in S(I, \mathbb{R})$ mit $f \geq 0$. Dann ist*

$$\int_\alpha^\beta f \geq 0.$$

(2) *Seien $f, g \in S(I, \mathbb{R})$ mit $f \geq g$. Dann ist*

$$\int_\alpha^\beta f \geq \int_\alpha^\beta g.$$

(3) *Sei $f \in S(I, \mathbb{R})$ mit $f \geq 0$. Weiter gebe es ein $a \in I$ mit: $f(a) > 0$ und $f$ ist stetig in $a$. Dann gilt*

$$\int_\alpha^\beta f > 0.$$

Beweis. ad(1): Gemäß Bemerkung V.4.5 (S. 169) gibt es eine Folge $(f_n)_{n\in\mathbb{N}} \subset T(I, \mathbb{R})$ mit $f_n \xrightarrow{\text{glm}} f$ und $f_n \geq 0$ für alle $n \in \mathbb{N}$. Offensichtlich ist

$$\int_\alpha^\beta f_n \geq 0 \quad \forall n \in \mathbb{N}.$$

Hieraus folgt die Behauptung durch Grenzübergang $n \to \infty$.
ad(2): Folgt aus Teil (1) angewandt auf $f - g$.
ad(3): Wegen der Stetigkeit von $f$ in $a$ und wegen $f(a) > 0$ gibt es ein $\delta > 0$ mit

$$f(x) \geq \frac{1}{2} f(a) > 0 \quad \forall x \in I \cap (a - \delta, a + \delta).$$

Damit folgt aus Satz VI.2.5 und den Teilen (1) und (2)

$$\begin{aligned} \int_\alpha^\beta f &= \int_\alpha^{a-\delta} f + \int_{a-\delta}^{a+\delta} f + \int_{a+\delta}^\beta f \\ &\geq \int_{a-\delta}^{a+\delta} \frac{1}{2} f(a) \end{aligned}$$

$$
\begin{aligned}
&= \delta f(a)\\
&> 0
\end{aligned}
$$

mit den offensichtlichen Modifikationen für $a \in \{\alpha, \beta\}$. $\square$

Als nächstes betrachten wir die Integration komplex- und vektorwertiger Funktionen.

SATZ VI.2.7. (1) *Sei $f : I \to \mathbb{K}^n$. Dann ist $f \in S(I, \mathbb{K}^n)$ genau dann, wenn für jede Komponente $f_i$, $1 \le i \le n$, gilt $f_i \in S(I, \mathbb{K})$. Weiter ist*

$$\int_\alpha^\beta f = (\int_\alpha^\beta f_1, \dots, \int_\alpha^\beta f_n).$$

(2) *Sei $f = g + ih : I \to \mathbb{C}$ mit $g, h : I \to \mathbb{R}$. Dann ist $f \in S(I, \mathbb{C})$ genau dann, wenn gilt $g, h \in S(I, \mathbb{R})$. Weiter ist*

$$\int_\alpha^\beta f = \int_\alpha^\beta g + i \int_\alpha^\beta h.$$

BEWEIS. AD(1): Es gilt

$$
\begin{aligned}
& f \in S(I, \mathbb{K}^n)\\
\iff & \exists (f_m) \subset T(I, \mathbb{K}^n) : f_m \underset{m\to\infty}{\overset{\text{glm}}{\longrightarrow}} f\\
\iff & \exists (f_m) \subset T(I, \mathbb{K}^n) : f_{i,m} \underset{m\to\infty}{\overset{\text{glm}}{\longrightarrow}} f_i, 1 \le i \le n\\
\iff & f_i \in S(I, \mathbb{K}) \quad \forall 1 \le i \le n.
\end{aligned}
$$

Die Aussage über die Integrale ist für Treppenfunktionen offensichtlich. Für sprungstetige Funktionen folgt sie durch Grenzübergang.
AD(2): Folgt aus Teil (1) mit $\mathbb{C} \cong \mathbb{R}^2$. $\square$

Als Vorbereitung auf den Hauptsatz der Differential- und Integralrechnung betrachten wir nun die stetige und differenzierbare Abhängigkeit des Integrals von den Integrationsgrenzen.

SATZ VI.2.8 (STETIGE ABHÄNGIGKEIT VON DEN INTEGRATIONSGRENZEN). *Sei $f \in S(I, X)$ und $c \in I$. Dann ist die Funktion*

$$F(x) = \int_c^x f \qquad , \forall x \in I,$$

*aus $C(I, X)$, und es gilt*

$$\|F(x) - F(y)\| \le |x - y| \|f\|_\infty \quad \forall x, y \in I.$$

BEWEIS. Aus Satz VI.2.5 folgt für $x, y \in I$

$$
\begin{aligned}
F(x) - F(y) &= \int_c^x f - \int_c^y f\\
&= \int_y^x f
\end{aligned}
$$

und damit aus Bemerkung VI.1.10 (S. 177)

$$\|F(x) - F(y)\| \le |x - y| \|f\|_\infty.$$

Hieraus folgt die Stetigkeit von $F$. □

SATZ VI.2.9 (DIFFERENZIERBARE ABHÄNGIGKEIT VON DEN INTEGRATIONSGRENZEN). *Sei $f \in S(I, X)$ in $a \in I$ stetig. Dann ist die Funktion*

$$F(x) = \int_\alpha^x f \qquad , \forall x \in I,$$

*in $a$ differenzierbar, und es gilt*

$$F'(a) = f(a).$$

BEWEIS. Gemäß Satz VI.2.8 ist $F$ stetig. Sei $\varepsilon > 0$ beliebig. Dann gibt es ein $\delta > 0$ mit

$$\|f(\zeta) - f(a)\| < \varepsilon \quad \forall \zeta \in I, |\zeta - a| < \delta.$$

Sei $h \in \mathbb{R}^*$ mit $a + h \in I$ und $|h| < \delta$. Dann folgt

$$\begin{aligned}
\|\frac{F(a+h) - F(a)}{h} - f(a)\| &= \|\frac{1}{h} \int_a^{a+h} f(\zeta) d\zeta - f(a)\| \\
&= \|\frac{1}{h} \int_a^{a+h} [f(\zeta) - f(a)] d\zeta\| \\
&\le \frac{1}{|h|} |h| \varepsilon \\
&= \varepsilon.
\end{aligned}$$

Hieraus folgt die Behauptung. □

DEFINITION VI.2.10. Seien $f, F : I \to X$ zwei Funktionen. $F$ heißt STAMMFUNKTION von $f$, wenn $F$ in jedem Punkt von $I$ differenzierbar ist und wenn gilt

$$F'(x) = f(x) \quad \forall x \in I.$$

Die Menge aller Stammfunktionen von $f$ wird mit $\int f$ oder $\int f dx$ bezeichnet und heißt UNBESTIMMTES INTEGRAL von $f$.

Aus Satz IV.2.6 folgt:

BEMERKUNG VI.2.11. Seien $F, G$ zwei Stammfunktionen von $f$. Dann ist $F - G$ auf $I$ konstant.

SATZ VI.2.12 (HAUPTSATZ DER DIFFERENTIAL- UND INTEGRALRECHNUNG). *Jedes $f \in C(I, X)$ besitzt eine Stammfunktion und für jede Stammfunktion $F$ von $f$ gilt*

$$F(x) = F(\alpha) + \int_\alpha^x f \quad \forall x \in I.$$

*Insbesondere gilt für jede Stammfunktion $F$ von $f$*

$$\int_\alpha^\beta f = F(\beta) - F(\alpha) =: F|_\alpha^\beta.$$

BEWEIS. Sei

$$G(x) = \int_\alpha^x f.$$

Gemäß Satz VI.2.9 ist $G$ eine Stammfunktion von $f$. Aus Bemerkung VI.2.11 folgt für jede Stammfunktion $F$ von $f$:

$$\begin{aligned}
&F(x) = G(x) + c \quad , \text{ mit } c \in X, \quad \forall x \in I \\
\Longrightarrow &F(\alpha) = G(\alpha) + c = c \\
\Longrightarrow &F(x) = F(\alpha) + G(x) = F(\alpha) + \int_\alpha^x f \quad \forall x \in I.
\end{aligned}$$

□

TABELLE 1. Stammfunktionen einiger wichtiger Funktionen

| $f$ | $\int f$ |
|---|---|
| $x^\alpha$, $\alpha \neq -1$ | $\frac{1}{\alpha+1} x^{\alpha+1} + c$ |
| $\frac{1}{x}$ | $\ln x + c$ |
| $a^x$, $a > 0$, $a \neq 1$ | $\frac{1}{\ln a} a^x + c$ |
| $e^{ax}$, $a \in \mathbb{C}^*$ | $\frac{1}{a} e^{ax} + c$ |
| $\cos x$ | $\sin x + c$ |
| $\sin x$ | $-\cos x + c$ |
| $\frac{1}{\cos^2 x}$ | $\tan x + c$ |
| $\frac{1}{\sin^2 x}$ | $-\cot x + c$ |
| $\frac{1}{\sqrt{1-x^2}}$ | $\arcsin x + c$ |
| $\frac{1}{1+x^2}$ | $\arctan x + c$ |
| $\sum_{k=0}^\infty a_k (x-x_0)^k$ | $\sum_{k=0}^\infty \frac{1}{k+1} a_k (x-x_0)^{k+1} + c$ |

Beispiel VI.2.13. (1) Tabelle 1 gibt die Stammfunktionen für einige wichtige Funktionen an.
(2) $f \in C^1(I, \mathbb{R})$, $f(x) \neq 0$ für alle $x \in I \Longrightarrow \displaystyle\int \frac{f'}{f} = \ln|f| + c$.

Beweis. ad(1): Folgt durch Differentiation der rechten Seite.
ad(2): Wegen des Zwischenwertsatzes III.5.5 (S. 90) gilt für alle $x \in I$ entweder

$$(a)\ f(x) > 0 \quad \text{oder} \quad (b)\ f(x) < 0.$$

In Fall $(a)$ folgt

$$(\ln|f|)' = (\ln f)' = \frac{f'}{f}$$

und in Fall $(b)$

$$(\ln|f|)' = (\ln(-f))' = \frac{-f'}{-f} = \frac{f'}{f}.$$

□

Zum Abschluss beweisen wir ein Analogon des Mittelwertsatzes IV.2.5 (S. 122).

Satz VI.2.14 (Mittelwertsatz der Integralrechnung). *Seien $f, \varphi \in C(I, \mathbb{R})$ und $\varphi \geq 0$ auf $I$. Dann gibt es ein $\zeta \in I$ mit*

$$\int_\alpha^\beta (f \cdot \varphi) = f(\zeta) \int_\alpha^\beta \varphi.$$

*Insbesondere gibt es ein $\eta \in I$ mit*

$$\int_\alpha^\beta f = f(\eta)(\beta - \alpha).$$

Beweis. Sei

$$\int_\alpha^\beta \varphi = 0.$$

Aus Satz VI.2.6 folgt $\varphi = 0$. Also ist die Behauptung richtig.
Sei nun

$$\int_\alpha^\beta \varphi > 0.$$

Definiere

$$m = \inf_{x \in I} f(x) \quad , \quad M = \sup_{x \in I} f(x).$$

Dann folgt

$$m\varphi(x) \leq f(x)\varphi(x) \leq M\varphi(x) \quad \forall x \in I$$

und wegen Satz VI.2.6

$$m \int_\alpha^\beta \varphi \leq \int_\alpha^\beta (f \cdot \varphi) \leq M \int_\alpha^\beta \varphi$$

$$\Longrightarrow m \leq \frac{\int_\alpha^\beta f \cdot \varphi}{\int_\alpha^\beta \varphi} \leq M.$$

Wegen des Zwischenwertsatzes III.5.5 (S. 90) gibt es daher ein $\zeta \in I$ mit

$$f(\zeta) = \frac{\int_\alpha^\beta f \cdot \varphi}{\int_\alpha^\beta \varphi}.$$

Damit ist die erste Behauptung bewiesen.
Die zweite Behauptung folgt aus der ersten mit der speziellen Wahl $\varphi = 1$. □

## VI.3. Integrationstechniken

In diesem Paragraphen betrachten wir einige, für praktische Anwendungen wichtige Integrationstechniken. Dabei bezeichnet wieder $(X, \|\cdot\|)$ einen $\mathbb{K}$-Banachraum und $I = [\alpha, \beta]$ ein kompaktes, perfektes Intervall.

Wir beginnen mit der Substitutionsregel.

Satz VI.3.1 (Substitutionsregel). *Seien $f \in C(I, X)$, $a, b \in \mathbb{R}$ mit $a < b$, und $\varphi \in C^1([a, b], \mathbb{R})$ mit $\varphi([a, b]) \subset I$. Dann gilt*

$$\int_a^b f(\varphi(x))\varphi'(x)dx = \int_{\varphi(a)}^{\varphi(b)} f(y)dy.$$

Beweis. Gemäß Satz VI.2.12 (S. 184) besitzt $f$ eine Stammfunktion $F \in C^1(I, X)$. Dann ist $F \circ \varphi \in C^1([a, b], X)$ mit

$$\begin{aligned}(F \circ \varphi)' &= F' \circ \varphi \cdot \varphi' \\ &= f \circ \varphi \cdot \varphi'.\end{aligned}$$

Damit folgt

$$\begin{aligned}\int_a^b f(\varphi(x))\varphi'(x)dx &= F(\varphi(b)) - F(\varphi(a)) \\ &= \int_{\varphi(a)}^{\varphi(b)} f(y)dy.\end{aligned}$$

□

Bemerkung VI.3.2. Sei $M \subset \mathbb{R}$ perfekt und $\varphi : M \to \mathbb{R}$ differenzierbar. Dann heißt

$$d\varphi = \varphi' dx$$

das Differential von $\varphi$ (vgl. Abbildung VI.3.1). Speziell ist $dx$ das Differential der Funktion $x \mapsto x$. Anschaulich ist das Differential der inkrementelle Zuwachs von $\varphi$.

In der Sprache der Differentiale können wir die Substitutionsregel auch so schreiben:

$$\int_a^b f \circ \varphi d\varphi = \int_{\varphi(a)}^{\varphi(b)} f(y)dy.$$

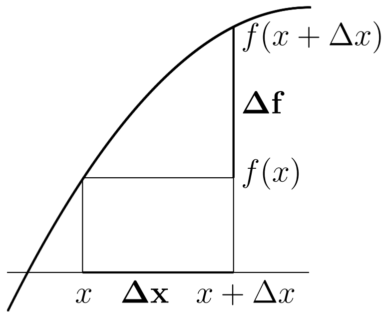


Abbildung VI.3.1. Differential $df = f'(x)dx \approx \Delta f = f(x + \Delta x) - f(x)$

Beispiel VI.3.3. (1) Sei $a \neq 0$.

$$\begin{aligned}
\int_\alpha^\beta \sin(ax+b)dx &= \frac{1}{a}\int_\alpha^\beta \sin(ax+b)adx \quad y = ax+b, dy = adx \\
&= \frac{1}{a}\int_{a\alpha+b}^{a\beta+b} \sin(y)dy \\
&= \frac{1}{a}(-\cos(y))|_{a\alpha+b}^{a\beta+b} \\
&= \frac{1}{a}[\cos(a\alpha+\beta) - \cos(a\beta+b)]
\end{aligned}$$

(2) Sei $n \geq 2$.

$$\begin{aligned}
\int_0^1 x^{n-1}\sin(x^n)dx &= \frac{1}{n}\int_0^1 nx^{n-1}\sin(x^n)dx \quad y = x^n, dy = nx^{n-1}dx \\
&= \frac{1}{n}\int_0^1 \sin(y)dy \\
&= \frac{1}{n}(-\cos(y))|_0^1 \\
&= \frac{1-\cos(1)}{n}.
\end{aligned}$$

(3) Sei $a, b \in \mathbb{R}$ mit $d = b - \frac{a^2}{4} > 0$.

$$\begin{aligned}
\int_\alpha^\beta \frac{1}{x^2+ax+b}dx &= \int_\alpha^\beta \frac{1}{(x+\frac{a}{2})^2+d}dx \quad y = \frac{x+\frac{a}{2}}{\sqrt{d}}, dy = \frac{1}{\sqrt{d}}dx \\
&= \sqrt{d}\int_\alpha^\beta \frac{1}{d(\frac{x+\frac{a}{2}}{\sqrt{d}})^2+d}\frac{1}{\sqrt{d}}dx
\end{aligned}$$

$$\begin{aligned} &= \sqrt{d} \int_{\frac{\alpha+\frac{a}{2}}{\sqrt{d}}}^{\frac{\beta+\frac{a}{2}}{\sqrt{d}}} \frac{1}{d(y^2+1)} dy \\ &= \frac{1}{\sqrt{d}} \arctan(y)\Big|_{\frac{\alpha+\frac{a}{2}}{\sqrt{d}}}^{\frac{\beta+\frac{a}{2}}{\sqrt{d}}} \\ &= \frac{1}{\sqrt{d}}[\arctan(\frac{\beta+\frac{a}{2}}{\sqrt{d}}) - \arctan(\frac{\alpha+\frac{a}{2}}{\sqrt{d}})]. \end{aligned}$$

Insbesondere folgt

$$\int \frac{1}{x^2+ax+b} = \frac{1}{\sqrt{d}} \arctan(\frac{x+\frac{a}{2}}{\sqrt{d}}) + c.$$

Als nächstes betrachten wir die partielle Integration.

Satz VI.3.4 (Partielle Integration). *Seien $u, v \in C^1(I, \mathbb{K})$. Dann ist*

$$\int_\alpha^\beta uv'dx = (u \cdot v)|_\alpha^\beta - \int_\alpha^\beta u'vdx$$

*oder in Differentialschreibweise*

$$\int_\alpha^\beta udv = (uv)|_\alpha^\beta - \int_\alpha^\beta vdu.$$

Beweis. Aus der Produktregel folgt

$$\begin{aligned} (uv)|_\alpha^\beta &= \int_\alpha^\beta (uv)'dx \\ &= \int_\alpha^\beta (u'v+uv')dx. \end{aligned}$$

□

Beispiel VI.3.5. (1)

$$\begin{aligned} \int_\alpha^\beta x\sin x dx &= -x\cos x|_\alpha^\beta + \int_\alpha^\beta \cos x dx \quad u=x,\ v=-\cos x \\ &= (-x\cos x + \sin x)|_\alpha^\beta \\ &= \sin\beta - \beta\cos\beta - \sin\alpha + \alpha\cos\alpha. \end{aligned}$$

Insbesondere ist

$$\int x\sin x = \sin x - x\cos x + c.$$

(2) Sei $F_R$ der Flächeninhalt eines Kreises um 0 mit Radius $R > 0$ und $H_R$ der Flächeninhalt des entsprechenden Halbkreises. Dann ist

$$F_R = 2H_R$$

und

$$H_R = \int_{-R}^{R} \sqrt{R^2 - x^2} dx \quad \alpha = -R, \beta = R$$

$$x = R\cos t, t \in [-\pi, 0]$$

$$dx = -R\sin t dt$$

$$\begin{aligned}
&= \int_{-\pi}^{0} \sqrt{R^2(1-\cos^2 t)}(-R\sin t)dt \\
&= \int_{-\pi}^{0} (-R\sin t)(-R\sin t)dt \\
&= R^2 \int_{-\pi}^{0} \sin^2 t dt \qquad u = \sin t, v = -\cos t \\
&= R^2(-\cos t \sin t)|_{-\pi}^{0} + R^2 \int_{-\pi}^{0} \cos^2 t dt \\
&= R^2 \int_{-\pi}^{0} 1 - \sin^2 t dt \\
&= \pi R^2 - H_R.
\end{aligned}$$

Hieraus folgt

$$H_R = \frac{1}{2}\pi R^2$$

und

$$F_R = \pi R^2.$$

(3)

$$I_n = \int \sin^n x \ , n \in \mathbb{N}$$

$$\Longrightarrow I_0 = x + c$$

$$I_1 = -\cos x + c.$$

Sei $n \geq 2$. Dann gilt

$$\begin{aligned}
I_n &= \int \sin^{n-1} x \sin x \qquad u = \sin^{n-1} x,\ v = -\cos x \\
&= -\sin^{n-1} x \cos x + (n-1) \int \sin^{n-2} x \cos^2 x \\
&= -\sin^{n-1} x \cos x + (n-1) \int \sin^{n-2} x (1 - \sin^2 x) \\
&= -\sin^{n-1} x \cos x + (n-1) I_{n-2} - (n-1) I_n
\end{aligned}$$

$$\Longrightarrow I_n = -\frac{1}{n}\sin^{n-1} x\cos x + \frac{n-1}{n} I_{n-2}.$$

(4) Für $n \in \mathbb{N}$ sei

$$A_n = \int_0^{\frac{\pi}{2}} \sin^n x dx.$$

Aus (3) folgt

$$\begin{aligned}
A_0 &= \frac{\pi}{2} \\
A_1 &= 1 \\
A_n &= \frac{n-1}{n} A_{n-2} \quad n \geq 2. \\
\Longrightarrow A_{2n} &= \frac{2n-1}{2n} A_{2n-2} \\
&= \frac{2n-1}{2n}\frac{2n-3}{2n-2} A_{2n-4} \\
&= \frac{2n-1}{2n}\frac{2n-3}{2n-2}\cdots\frac{1}{2}\frac{\pi}{2} \\
A_{2n+1} &= \frac{2n}{2n+1} A_{2n-1} \\
&= \frac{2n}{2n+1}\frac{2n-2}{2n-1} A_{2n-3} \\
&= \frac{2n}{2n+1}\frac{2n-2}{2n-1}\cdots\frac{2}{3}\cdot 1.
\end{aligned}$$

Wegen

$$\sin^{2n+2} x \leq \sin^{2n+1} x \leq \sin^{2n} x \quad \forall x \in [0, \frac{\pi}{2}]$$

folgt

$$A_{2n+2} \leq A_{2n+1} \leq A_{2n}.$$

Wegen

$$\frac{A_{2n+2}}{A_{2n}} = \frac{2n+1}{2n+2} \underset{n\to\infty}{\longrightarrow} 1$$

folgt

$$\lim_{n\to\infty} \frac{A_{2n+1}}{A_{2n}} = 1.$$

Andererseits gilt

$$\begin{aligned}
\frac{A_{2n+1}}{A_{2n}} &= \frac{2n}{2n+1}\frac{2n}{2n-1}\frac{2n-2}{2n-1}\frac{2n-2}{2n-3}\cdots\frac{4}{3}\cdot\frac{2}{3}\cdot\frac{2}{1}\cdot\frac{2}{\pi} \\
&= \prod_{k=1}^{n} \frac{(2k)^2}{(2k)^2-1}\cdot\frac{2}{\pi}.
\end{aligned}$$

Hieraus folgt:

$$\begin{aligned}\frac{\pi}{2} &= \prod_{k=1}^{\infty} \frac{(2k)^2}{(2k)^2-1} \\ &= \lim_{n\to\infty} \prod_{k=1}^{n} \frac{(2k)^2}{(2k)^2-1} \\ &= \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots .\end{aligned}$$

Diese Darstellung von $\frac{\pi}{2}$ heißt das Wallissche Produkt.

Als nächstes betrachten wir die Methode der Partialbruchzerlegung, die geeignet ist, Integrale rationaler Funktionen zu berechnen. Wir erklären das Prinzip an Hand eines Beispiels.

Beispiel VI.3.6 (Partialbruchzerlegung). Gesucht

$$\begin{aligned}\int \frac{p}{q} &= \int \frac{x^5-2x^4+x^3-2x^2+5x-8}{x^4-x^2-2x+2} \\ p &= x^5-2x^4+x^3-2x^2+5x-8 \\ q &= x^4-x^2-2x+2.\end{aligned}$$

1. Schritt: Bestimme mit dem Euklidischen Algorithmus Polynome $r$ und $s$ mit $p = r + sq$ und $\partial r < \partial q$.

$$\begin{array}{cccccccl} (x^5 & -2x^4 & +x^3 & -2x^2 & +5x & -8) & : (x^4-x^2 & -2x+2) \\ x^5 & & -x^3 & -2x^2 & +2x & & & = x-2 \\ \hline & -2x^4 & +2x^3 & & +3x & -8 & & \\ & -2x^4 & & +2x^2 & +4x & -4 & & \\ \hline & & 2x^3 & -2x^2 & -x & -4 & & \end{array}$$

$$\begin{aligned}\Longrightarrow\ s &= x-2, \\ r &= 2x^3-2x^2-x-4 \\ \int \frac{p}{q} &= \int s + \int \frac{r}{q}.\end{aligned}$$

2. Schritt: Zerlege das Nennerpolynom $q$ in der Form

$$\begin{aligned}q = \prod_{i=1}^{m} (x-a_i)^{\alpha_i} &\cdot \prod_{j=1}^{n} (x^2+b_jx+c_j)^{\beta_j} \\ &\text{mit } a_i, b_j, c_j \in \mathbb{R},\ c_j - \frac{b_j^2}{4} > 0, \\ &\alpha_i, \beta_j \in \mathbb{N}^*, \\ &\sum_{i=1}^{m} \alpha_i + \sum_{j=1}^{n} 2\beta_j = \partial q.\end{aligned}$$

Dabei ist

$$\begin{aligned} q(a_i) &= 0, \\ q(z_j) &= q(\bar{z}_j) \\ &= 0 \\ &\qquad \text{mit } b_j = -2\operatorname{Re} z_j,\ c_j = |z_j|^2 \neq 0. \end{aligned}$$

Es ist $q(1) = 0$. Daher ist $q$ durch $x - 1$ teilbar. Der Euklidische Algorithmus liefert

$$\begin{array}{lllll} (x^4 & & -x^2 & -2x & +2) \quad : (x-1) = x^3 + x^2 - 2. \\ \ \ x^4 & -x^3 & & & \\ \hline & \ \ x^3 & -x^2 & & \\ & \ \ x^3 & -x^2 & & \\ \hline & & & -2x & +2 \\ & & & -2x & +2 \\ \hline & & & & \ \ 0 \end{array}$$

Das Polynom $x^3 + x^2 - 2$ hat eine Nullstelle bei $x = 1$ und ist daher durch $x - 1$ teilbar. Der Euklidische Algorithmus liefert

$$\begin{array}{lllll} (x^3 & +x^2 & & -2) \quad : (x-1) = x^2 + 2x + 2. \\ \ \ x^3 & -x^2 & & \\ \hline & 2x^2 & & \\ & 2x^2 & -2x & \\ \hline & & 2x & -2 \\ & & 2x & -2 \\ \hline & & & \ \ 0 \end{array}$$

Also ist

$$q = (x-1)^2(x^2 + 2x + 2).$$

3. Schritt: Zerlege $\frac{r}{q}$ in der Form

$$\frac{r}{q} = \sum_{i=1}^{m} \sum_{\mu=1}^{\alpha_i} \frac{\gamma_{i,\mu}}{(x - a_i)^\mu} + \sum_{j=1}^{n} \sum_{\nu=1}^{\beta_j} \frac{\delta_{j,\nu} x + \varepsilon_{j,\nu}}{(x^2 + b_j x + c_j)^\nu}.$$

Dies liefert die Bedingung

$$\frac{2x^3 - 2x^2 - x - 4}{x^4 - x^2 - 2x + 2} \overset{!}{=} \frac{a}{x-1} + \frac{b}{(x-1)^2} + \frac{cx + d}{x^2 + 2x + 2}.$$

Multiplikation beider Seiten mit $q$ ergibt die Bedingung

$$\begin{aligned} 2x^3 - 2x^2 - x - 4 &\overset{!}{=} a(x-1)(x^2 + 2x + 2) \\ &\quad + b(x^2 + 2x + 2) \\ &\quad + (cx + d)(x-1)^2 \\ &= a(x^3 + x^2 - 2) \\ &\quad + b(x^2 + 2x + 2) \end{aligned}$$

$$
\begin{aligned}
&+ (cx+d)(x^2-2x+1)\\
&= x^3(a+c)\\
&\quad + x^2(a+b-2c+d)\\
&\quad + x(2b+c-2d)\\
&\quad + (-2a+2b+d)
\end{aligned}
$$

Koeffizientenvergleich liefert das Gleichungssystem

$$
\begin{aligned}
a+c &= 2 & a+c &= 2\\
a+b-2c+d &= -2 \quad\Longrightarrow & b-3c+d &= -4\\
2b+c-2d &= -1 & 2b+c-2d &= -1\\
-2a+2b+d &= -4 & 2b+2c+d &= 0
\end{aligned}
$$

$$
\begin{aligned}
a+c &= 2 & a+c &= 2\\
\Longrightarrow \quad b-3c+d &= -4 \quad\Longrightarrow & b-3c+d &= -4\\
7c-4d &= 7 & 7c-4d &= 7\\
8c-d &= 8 & \tfrac{25}{7}d &= 0
\end{aligned}
$$

Also ist

$$d = 0,\ c = 1,\ b = -1,\ a = 1$$

und

$$\frac{p}{q} = x - 2 + \frac{1}{x-1} - \frac{1}{(x-1)^2} + \frac{x}{x^2+2x+2}.$$

4. Schritt: Integration der einzelnen Terme liefert schließlich

$$
\begin{aligned}
&\int \frac{x^5-2x^4+x^3-2x^2+5x-8}{x^4-x^2-2x+2}\\
=&\int \{x-2+\frac{1}{x-1}-\frac{1}{(x-1)^2}+\frac{x+1}{x^2+2x+2}-\frac{1}{x^2+2x+2}\}\\
=&\frac{1}{2}x^2-2x+\ln|x-1|+\frac{1}{x-1}+\frac{1}{2}\ln|x^2+2x+2|\\
&-\arctan(x+1)+c.
\end{aligned}
$$

Zum Abschluss gehen wir noch kurz auf die numerische Integration ein. In vielen Fällen ist nämlich die Stammfunktion nicht geschlossen darstellbar oder von einer praktisch kaum auswertbaren Form, so dass numerische Methoden benützt werden. Wir beschränken uns hier auf eines der einfachsten, aber auch vielfältigsten Verfahren, die sog. Trapezregel.

Definition VI.3.7. Seien $f \in C(I, X)$ und $n \in \mathbb{N}^*$, sowie $h = \frac{\beta-\alpha}{n}$. Dann ist die zusammengesetzte Trapezregel angewandt auf $f$

gegeben durch

$$T_n(f) = \frac{h}{2}\{f(\alpha) + 2\sum_{i=1}^{n-1} f(\alpha + ih) + f(\beta)\}.$$

Wir wollen eine Fehlerabschätzung für $T_n(f) - \int_\alpha^\beta f$ herleiten. Dazu benötigen wir:

LEMMA VI.3.8. *Sei $\varphi \in C^2([0,1],\mathbb{R})$. Dann gibt es ein $\zeta \in [0,1]$ mit*

$$\int_0^1 \varphi = \frac{1}{2}(\varphi(0) + \varphi(1)) - \frac{1}{12}\varphi''(\zeta).$$

BEWEIS. Durch zweimalige partielle Integration erhalten wir mit Satz VI.2.14 (S. 186):

$$\begin{aligned}
&\int_0^1 \varphi(x)dx \qquad u = \varphi(x), v = (x - \frac{1}{2}) \\
=&(x - \frac{1}{2})\varphi(x)|_0^1 - \int_0^1 (x - \frac{1}{2})\varphi'(x)dx \\
&\qquad\qquad u = \varphi'(x), v = \frac{1}{2}(x - \frac{1}{2})^2 - \frac{1}{8} \\
=&\frac{1}{2}[\varphi(1) + \varphi(0)] - [\frac{1}{2}(x - \frac{1}{2})^2 - \frac{1}{8}]\varphi'(x)|_0^1 \\
&\qquad + \int_0^1 [\frac{1}{2}(x - \frac{1}{2})^2 - \frac{1}{8}]\varphi''(x)dx \\
=&\frac{1}{2}[\varphi(0) + \varphi(1)] + \frac{1}{2}\int_0^1 x(x-1)\varphi''(x)dx \\
=&\frac{1}{2}[\varphi(0) + \varphi(1)] + \frac{1}{2}\varphi''(\zeta)\int_0^1 x(x-1)dx \\
=&\frac{1}{2}[\varphi(0) + \varphi(1)] - \frac{1}{12}\varphi''(\zeta).
\end{aligned}$$

□

SATZ VI.3.9. *Sei $f \in C^2(I,\mathbb{R})$. Dann gilt*

$$|\int_\alpha^\beta f - T_n(f)| \le \frac{\beta - \alpha}{12} h^2 \|f''\|_\infty.$$

BEWEIS. Aus Definition VI.3.7 und Lemma VI.3.8 folgt:

$$\begin{aligned}
\Big|\int_0^1 f(x)dx - T_n(f)\Big| = \Big|\sum_{i=1}^n \Big\{ &\int_{\alpha+(i-1)h}^{\alpha+ih} f(x)dx \\
&- \frac{h}{2}[f(\alpha + (i-1)h) + f(\alpha + ih)]\Big\}\Big|
\end{aligned}$$

$$
\begin{aligned}
&= \Big|\sum_{i=1}^{n}\Big\{\int_0^1 \underbrace{hf(\alpha+(i-1)h+th)}_{=:\varphi(t)}\,dt \\
&\qquad\qquad - \frac{1}{2}[hf(\alpha+(i-1)h)+hf(\alpha+ih)]\Big\}\Big| \\
&= \Big|\sum_{i=1}^{n}\Big\{-\frac{1}{12}h^3 f''(\alpha+(i-1)h+\zeta_i h)\Big\}\Big| \\
&\le \frac{h^3}{12}\sum_{i=1}^{n}|f''(\alpha+(i-1)h+\zeta_i h)| \\
&\le \frac{h^3}{12}n\|f''\|_\infty \\
&= \frac{h^2}{12}(\beta-\alpha)\|f''\|_\infty .
\end{aligned}
$$

Dies beweist die Behauptung. □

Als Anwendung beweisen wir die Stirlingsche Formel.

Satz VI.3.10 (Stirlingsche Formel). *Es ist*

$$n! = \sqrt{2\pi n}\; n^n e^{-n} e^{R_n}$$

*mit* $0 < R_n \le \frac{1}{12(n-1)}$.

Beweis. Mit Lemma VI.3.8 folgt

$$
\begin{aligned}
n\ln(n) - n + 1 &= (x\ln x - x)\big|_1^n \\
&= \int_1^n \ln(x)dx \\
&= \sum_{k=1}^{n-1}\int_k^{k+1}\ln(x)dx \\
&= \sum_{k=1}^{n-1}\{\frac{1}{2}[\ln(k+1)+\ln(k)] + r_k\} \\
&\quad \text{mit } r_k = -\frac{1}{12}\ln''(\zeta_k) = \frac{1}{12}\zeta_k^{-2}, k \le \zeta_k \le k+1 \\
&= \sum_{k=1}^{n-1}\ln(k) + \frac{1}{2}\ln(n) + \sum_{k=1}^{n-1} r_k \\
&= \sum_{k=1}^{n}\ln(k) - \frac{1}{2}\ln(n) + \sum_{k=1}^{n-1} r_k \\
&= \ln(n!) - \frac{1}{2}\ln(n) + \sum_{k=1}^{n-1} r_k .
\end{aligned}
$$

Hieraus folgt

$$\begin{aligned}\ln(n!) &= (n+\frac{1}{2})\ln(n) - n + 1 - \sum_{k=1}^{n-1} r_k \\ &= (n+\frac{1}{2})\ln(n) - n + c + \sum_{k=n}^{\infty} r_k \\ &= (n+\frac{1}{2})\ln(n) - n + c + R_n\end{aligned}$$

mit

$$c = 1 - \sum_{k=1}^{\infty} r_k$$

und

$$R_n = \sum_{k=n}^{\infty} r_k.$$

Weiter gilt

$$\begin{aligned}0 < R_n &= \sum_{k=n}^{\infty} \frac{1}{12}\zeta_k^{-2} \\ &\le \frac{1}{12}\sum_{k=n}^{\infty} \frac{1}{k^2} \\ &\le \frac{1}{12}\sum_{k=n}^{\infty} \frac{1}{k(k-1)} \\ &= \frac{1}{12(n-1)}.\end{aligned}$$

Mit $\gamma = e^c$ folgt

$$n! = \gamma\sqrt{n}\; n^n e^{-n} e^{R_n}$$

mit $0 < R_n \le \frac{1}{12(n-1)}$.
Also müssen wir noch zeigen, dass $\gamma = \sqrt{2\pi}$ ist.
Gemäß Beispiel VI.3.5 (4) ist

$$2\pi = 4 \lim_{n\to\infty} \prod_{k=1}^{n} \frac{(2k)^2}{(2k)^2 - 1}.$$

Aus Obigem folgt aber

$$
\begin{aligned}
4\prod_{k=1}^{n}\frac{(2k)^2}{(2k)^2-1} &= 4\frac{\prod\limits_{k=1}^{n}(2k)^2}{\prod\limits_{k=1}^{n}(2k-1)\prod\limits_{k=1}^{n}(2k+1)} \\
&= 4\frac{\prod\limits_{k=1}^{n}(2k)^4}{\prod\limits_{k=1}^{2n}k\prod\limits_{k=1}^{2n+1}k} \\
&= 4\frac{2^{4n}(n!)^4}{(2n)!(2n+1)!} \\
&= \frac{4\,2^{4n}\gamma^4 n^2 n^{4n}e^{-4n}e^{4R_n-R_{2n}-R_{2n+1}}}{\gamma^2\sqrt{2n}(2n)^{2n}e^{-2n}\sqrt{2n+1}(2n+1)^{2n+1}e^{-2n-1}} \\
&= \gamma^2\frac{ee^{4R_n-R_{2n}-R_{2n+1}}}{(1+\frac{1}{2n})^{2n}(1+\frac{1}{2n})\sqrt{1+\frac{1}{2n}}} \\
&\underset{n\to\infty}{\longrightarrow} \gamma^2.
\end{aligned}
$$

Hieraus folgt die Behauptung. $\square$

## VI.4. Uneigentliche Integrale

Bisher können wir nur sprungstetige Funktionen auf kompakten Intervallen integrieren. In diesem Paragraphen untersuchen wir die Grenzfälle, bei denen der Integrand an den Integrationsgrenzen eventuell nicht sprungstetig und/oder der Integrationsbereich unbeschränkt ist. Im Folgenden ist stets $(X, \|\cdot\|)$ ein $\mathbb{K}$-Banachraum und $-\infty \le a < b \le +\infty$.

Definition VI.4.1. Die Funktion $f : (a, b) \to X$ heißt zulässig, wenn für alle $a < \alpha < \beta < b$ gilt

$$f|_{[\alpha,\beta]} \in S([\alpha, \beta], X).$$

Bemerkung VI.4.2. Es gilt:

(1) $f \in C((a, b), X) \Longrightarrow f$ zulässig.
(2) $-\infty < a < b < +\infty$ und $f \in S([a, b], X) \Longrightarrow f$ zulässig.

Definition VI.4.3. Sei $f : (a, b) \to X$ zulässig. Falls die Grenzwerte

$$\lim_{\alpha\to a+0}\int_{\alpha}^{c} f \quad\text{und}\quad \lim_{\beta\to b-0}\int_{c}^{\beta} f$$

für ein $c \in (a,b)$ existieren, heißt $f$ UNEIGENTLICH AUF $(a,b)$ INTEGRIERBAR und

$$\int_a^b f = \lim_{\alpha \to a+0} \int_\alpha^c f + \lim_{\beta \to b-0} \int_c^\beta f$$

das UNEIGENTLICHE INTEGRAL von $f$.

SATZ VI.4.4. (1) *Die Definition von* $\displaystyle\int_a^b f$ *ist unabhängig von* $c$.
(2) *Ist* $-\infty < a < b < +\infty$ *und* $f \in S([a,b],X)$, *so stimmt das uneigentliche Integral von* $f$ *mit dem gewöhnlichen Integral überein.*

BEWEIS. AD (1): Seien $c, c' \in (a,b)$ und $a < \alpha < \min\{c,c'\} \le \max\{c,c'\} < \beta < b$. Dann gilt

$$\int_\alpha^\beta f = \int_\alpha^c f + \int_c^\beta f$$
$$\int_\alpha^\beta f = \int_\alpha^{c'} f + \int_{c'}^\beta f.$$

Hieraus folgt die Behauptung.
AD (2): Gemäß Satz VI.2.8 (S. 183) sind die Funktionen

$$\alpha \mapsto \int_\alpha^c f \quad \text{und} \quad \beta \mapsto \int_c^\beta f$$

stetig. Hieraus folgt die Behauptung. □

BEISPIEL VI.4.5. Es gilt:

(1) $$\int_a^\infty x^{-s} dx = \frac{a^{1-s}}{s-1} \quad \forall a > 0, s > 1.$$

(2) $$\int_0^b x^{-s} dx = \frac{b^{1-s}}{1-s} \quad \forall b > 0, s < 1.$$

(3) $$\int_{-\infty}^{+\infty} x dx \text{ existiert nicht, aber } \int_{-\gamma}^\gamma x dx = 0 \quad \forall \gamma > 0.$$

(4) $$\int_0^\infty \frac{1}{1+x^2} dx = \frac{\pi}{2}.$$

(5) $$\int_{-1}^0 \frac{1}{\sqrt{1-x^2}} dx = \pi.$$

BEWEIS. AD (1): Sei $s \neq 1$. Dann ist

$$\int_a^\beta x^{-s} dx = \frac{1}{1-s}(\beta^{1-s} - a^{1-s}).$$

Der Grenzwert existiert für $\beta \to +\infty$ genau dann, wenn $s > 1$ ist.
Sei $s = 1$. Dann ist

$$\int_a^\beta \frac{1}{x} dx = \ln(\beta) - \ln(a).$$

Der Grenzwert existiert nicht.
AD (2): Folgt wie Teil (1).
AD (3): Ist offensichtlich.
AD (4): Es gilt

$$\begin{aligned}\int_0^\beta \frac{1}{1+x^2}dx &= \arctan(\beta) - \arctan(0) \\ &= \arctan(\beta) \\ &\underset{\beta\to+\infty}{\longrightarrow} \frac{\pi}{2}.\end{aligned}$$

AD (5): Es gilt

$$\begin{aligned}\int_\alpha^0 \frac{1}{\sqrt{1-x^2}}dx &= \arcsin(0) - \arcsin(\alpha) \\ &= -\arcsin(\alpha) \\ &\underset{\alpha\to-1+}{\longrightarrow} \frac{\pi}{2}.\end{aligned}$$

$\square$

Aus der Definition des uneigentlichen Integrals folgt ein Konvergenzkriterium für Reihen.

SATZ VI.4.6 (INTEGRATIONSKRITERIUM FÜR REIHEN). *Sei $f \in S([1,+\infty),\mathbb{R}_+)$ monoton fallend. Dann existiert $\sum_{n=1}^{\infty} f(n)$ genau dann, wenn $\int_1^\infty f$ existiert.*

BEWEIS. Für alle $n \geq 2$ gilt

$$\begin{aligned} f(n) \leq f(x) \leq f(n-1) \quad \forall x \in [n-1,n] \\ \Longrightarrow\ f(n) \leq \int_{n-1}^n f(x)dx \leq f(n-1) \\ \Longrightarrow\ \sum_{n=2}^N f(n) \leq \int_1^N f dx \leq \sum_{n=1}^{N-1} f(n) \quad \forall N \geq 2.\end{aligned}$$

„$\Longrightarrow$“: Aus obiger Abschätzung folgt

$$\int_1^N f dx \leq \sum_{n=1}^{\infty} f(n) < \infty.$$

Also ist die Funktion $t \mapsto \int_1^t f(x)dx$ monoton wachsend und beschränkt. Wegen Satz III.6.3 (S. 94) existiert daher $\int_1^\infty f dx$.

„$\Longleftarrow$“: Aus obiger Abschätzung folgt jetzt

$$\sum_{n=2}^{N} f(n) \leq \int_1^\infty f dx < \infty \quad \forall N \geq 2.$$

Also existiert $\sum_{n=1}^{\infty} f(n)$. □

Als Anwendung ergibt sich:

Beispiel VI.4.7. Es gilt:

(1) $$\sum_{n=1}^{\infty} n^{-s} \text{ existiert für alle } s > 1.$$

(2) $$\sum_{n=2}^{\infty} \frac{1}{n[\ln(n)]^s} \text{ existiert } \iff s > 1.$$

Beweis. ad (1): Folgt aus Satz VI.4.6 und Beispiel VI.4.5 (1).
ad (2): Für $s \leq 0$ ist die harmonische Reihe eine divergente Minorante. Sei also $s > 0$. Dann ist $\frac{1}{x(\ln x)^s}$ monoton fallend und

$$\int_2^\beta \frac{1}{x(\ln x)^s} dx = \begin{cases} \ln(\ln(x))|_2^\beta & \text{, für } s = 1, \\ \frac{1}{-s+1} \ln(x)^{1-s}|_2^\beta & \text{, für } s \neq 1. \end{cases}$$

Damit folgt die Behauptung aus Satz VI.4.6. □

Man beachte, dass wegen Satz III.7.9 (S. 102) für jedes $s > 0$ und jedes $\varepsilon > 0$ die Reihe $\sum \frac{1}{n(\ln(n))^s}$ eine Majorante zu $\sum \frac{1}{n^{1+\varepsilon}}$ und eine Minorante zu $\sum \frac{1}{n}$ ist.

Die folgende Definition ist das Analogon zur absoluten Konvergenz von Reihen, vgl. Definition II.5.1 (S. 49).

Definition VI.4.8. $f : (a, b) \to X$ sei zulässig. Dann heißt $\int_a^b f$ absolut konvergent, genau dann wenn $\int_a^b \|f\|$ existiert.

Der folgende Satz ist das Analogon zum Majorantenkriterium für Reihen, vgl. Satz II.5.4 (S. 50).

Satz VI.4.9 (Majorantenkriterium). *Es gelte*

(1) $f : (a, b) \to X$ *ist zulässig,*
(2) $g : (a, b) \to \mathbb{R}_+$ *ist zulässig,*
(3) $\|f(x)\| \leq g(x)$ *für alle* $x \in (a, b)$,
(4) $\int_a^b g$ *existiert.*

*Dann ist* $\int_a^b f$ *absolut konvergent.*

Beweis. Beh.: $\int_a^b f$ existiert.

Bew. von Beh.: Sei $\varepsilon > 0$. Dann gibt es ein $\delta > 0$ mit

$$|\int_{\beta_1}^{\beta_2} g| < \varepsilon \qquad \forall b - \delta < \beta_1 \leq \beta_2 < b.$$

Sei $c \in (a, b - \delta)$. Dann folgt

$$\begin{aligned} \|\int_c^{\beta_1} f - \int_c^{\beta_2} f\| &= \|\int_{\beta_1}^{\beta_2} f\| \\ &\leq \int_{\beta_1}^{\beta_2} \|f\| \\ &\leq \int_{\beta_1}^{\beta_2} g \\ &< \varepsilon \qquad \forall b - \delta < \beta_1 \leq \beta_2 < b. \end{aligned}$$

Sei nun $(\beta_n)_{n\in\mathbb{N}} \subset (a, b)$ eine Folge mit $\beta_n \uparrow b$. Dann folgt aus Obigem, dass $(\int_c^{\beta_n} f)_{n\in\mathbb{N}}$ eine Cauchyfolge ist und somit gegen ein $e \in X$ konvergiert. Man überlegt sich sofort, dass $e$ von der Folge $(\beta_n)_{n\in\mathbb{N}}$ unabhängig ist. Also existiert $\lim\limits_{\beta \to b-0} \int_c^{\beta} f$. Analog argumentiert man für $\lim\limits_{\alpha \to a+0} \int_\alpha^c f$. Also existiert $\int_a^b f$.

Beweis des Satzes: Da $f$ zulässig ist, ist auch $\|f\|$ zulässig. Aus Beh. folgt, dass $\int_a^b \|f\|$ existiert. □

Bemerkung VI.4.10. Aus dem Beweis von Satz VI.4.9 folgt:

$$\int_a^b f \text{ absolut konvergent } \Longrightarrow \int_a^b f \text{ existiert.}$$

Beispiel VI.4.11. (1) $X = \mathbb{R}$, $f \geq 0$. Dann gilt

$$\int_a^b f \;\text{ konvergiert absolut } \iff \int_a^b f \;\text{ existiert.}$$

(2) $\displaystyle\int_0^\infty \frac{\sin x}{1 + x^2} dx$ konvergiert absolut.

(3) $f : (a, b) \to X$ sei zulässig.

(i) Sei $a > 0$, $b = +\infty$. Dann gilt

$$\begin{aligned} &\exists \varepsilon > 0 \; \exists M > 0 \; \exists c \geq a : |f(x)| \leq \frac{M}{x^{1+\varepsilon}} \; \forall x \geq c \\ &\Longrightarrow \int_a^\infty f \text{ konvergiert absolut.} \end{aligned}$$

(ii) Sei $a = 0$, $0 < b < +\infty$. Dann gilt

$$\exists \varepsilon > 0 \; \exists M > 0 \; \exists c \in (0, b) : |f(x)| \leq \frac{M}{x^{1-\varepsilon}} \; \forall x \in (0, c)$$

$$\Longrightarrow \int_0^b f \text{ konvergiert absolut.}$$

(4) Das Eulersche Beta-Integral

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt$$

existiert genau dann, wenn gilt $x > 0$ und $y > 0$. Es gilt für alle $x > 0, y > 0$:

(i) $$B(x, y) = B(y, x)$$

(ii) $$B(x + 1, y) = \frac{x}{x + y} B(x, y).$$

Beweis. ad (1): Klar wegen $f = \|f\|$.
ad (2): Folgt wegen $|\frac{\sin x}{1+x^2}| \leq \frac{1}{1+x^2}$ für alle $x \in \mathbb{R}_+$ aus Beispiel VI.4.5 (4) und Satz VI.4.9.
ad (3): Folgt aus Satz VI.4.9 und Beispiel VI.4.5 (1,2).
ad (4): Sei $y \in \mathbb{R}$ fest. Dann gibt es Zahlen $0 < m \leq M$ mit

$$m \leq (1-t)^{y-1} \leq M \quad \forall t \in [0, \frac{1}{2}].$$

Daher existiert

$$\int_0^{\frac{1}{2}} t^{x-1}(1-t)^{y-1} dt$$

genau dann, wenn

$$\int_0^{\frac{1}{2}} t^{x-1} dt$$

existiert. Gemäß Beispiel VI.4.5 (2) ist dies genau dann der Fall, wenn $x > 0$ ist.
Sei nun $x \in \mathbb{R}$ fest. Dann gibt es Zahlen $0 < k \leq K$ mit

$$k \leq t^{x-1} \leq K \quad \forall t \in [\frac{1}{2}, 1].$$

Also existiert

$$\int_{\frac{1}{2}}^1 t^{x-1}(1-t)^{y-1} dt$$

genau dann, wenn

$$\int_{\frac{1}{2}}^1 (1-t)^{y-1} dt$$

existiert. Wegen

$$\int_{\frac{1}{2}}^{\alpha} (1-t)^{y-1} dt = -\int_{\frac{1}{2}}^{1-\alpha} s^{y-1} ds \quad s = 1 - t$$

$$= \int_{1-\alpha}^{\frac{1}{2}} s^{y-1} ds \quad \forall \frac{1}{2} < \alpha < 1$$

und Beispiel VI.4.5 (2) ist dies genau dann der Fall, wenn $y > 0$ ist.
Sei nun $x, y \in \mathbb{R}_+^*$ und $0 < \alpha < \beta < 1$. Dann folgt

$$\begin{aligned}
\int_\alpha^\beta t^{x-1}(1-t)^{y-1} dt &= -\int_{1-\alpha}^{1-\beta} (1-s)^{x-1} s^{y-1} ds \quad s = 1-t \\
&= \int_{1-\beta}^{1-\alpha} (1-s)^{x-1} s^{y-1} ds \\
&\underset{\substack{\alpha\to 0+ \\ \beta \to 1-}}{\longrightarrow} \int_0^1 (1-t)^{x-1} t^{y-1} dt.
\end{aligned}$$

Hieraus folgt (i).
Weiter folgt für $x, y \in \mathbb{R}_+^*$ und $0 < \alpha < \beta < 1$:

$$\begin{aligned}
\int_\alpha^\beta t^x (1-t)^{y-1} dt = \int_\alpha^\beta (\frac{t}{1-t})^x (1-t)^{x+y-1} dt \quad & u = (\frac{t}{1-t})^x \\
, \quad & v = -\frac{1}{x+y}(1-t)^{x+y} \\
= -\frac{1}{x+y}(\frac{t}{1-t})^x (1-t)^{x+y}|_\alpha^\beta & \\
+ \int_\alpha^\beta \frac{x}{x+y}(\frac{t}{1-t})^{x-1} \frac{1}{(1-t)^2} \cdot (1-t)^{x+y} dt & \\
= -\frac{1}{x+y} t^x (1-t)^y |_\alpha^\beta & \\
+ \frac{x}{x+y} \int_\alpha^\beta t^{x-1}(1-t)^{y-1} dt & \\
\underset{\substack{\alpha\to 0+ \\ be \to 1-}}{\longrightarrow} \frac{x}{x+y} \int_0^1 t^{x-1}(1-t)^{y-1} dt. &
\end{aligned}$$

Hieraus folgt (ii). $\square$

Bemerkung VI.4.12. Satz VI.2.1 (S. 180) gilt für uneigentliche Integrale nicht. Betrachte z. B.

$$f_n(x) = \frac{1}{n} e^{-\frac{x}{n}} \in C(\mathbb{R}_+^*, \mathbb{R}).$$

Dann gilt $f_n \underset{n\to\infty}{\overset{\text{glm}}{\longrightarrow}} 0$ und

$$\int_0^\infty f_n(x) dx = \lim_{\beta \to +\infty} \int_0^\beta \frac{1}{n} e^{-\frac{x}{n}} dx$$

$$= \lim_{\beta\to+\infty}(1-e^{-\frac{\beta}{n}})$$
$$= 1 \qquad \forall n\in\mathbb{N}.$$

## VI.5. Die Eulersche Gammafunktion

Zuerst definieren wir die Eulersche Gammafunktion.

Satz VI.5.1 (Eulersche Gammafunktion). *Für alle $x \in \mathbb{R}_+^*$ existiert das uneigentliche Integral*

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

*Die Funktion $\Gamma$ heißt* Eulersche Gammafunktion. *Es gilt*

(i) $$\Gamma(1) = 1,$$

(ii) $$\Gamma(x+1) = x\Gamma(x) \qquad \forall x\in\mathbb{R}_+^*,$$

(iii) $$\Gamma(n+1) = n! \qquad \forall n\in\mathbb{N}.$$

Beweis. Sei $0 < t \le 1$. Dann ist

$$0 < t^{x-1}e^{-t} \le t^{x-1}.$$

Damit folgt aus Beispiel VI.4.5(2) (S. 199), dass $\int_0^1 t^{x-1}e^{-t}dt$ existiert. Sei nun $t \ge 1$ und $m \in \mathbb{N}$ so, dass gilt $x < m$. Dann folgt

$$\frac{1}{m!}t^m < \sum_{n=0}^\infty \frac{1}{n!}t^n = e^t$$
$$\Longrightarrow\ t^{x-1}e^{-t} < t^{x-1}m!t^{-m} = \frac{m!}{t^{1+m-x}}.$$

Wegen Beispiel VI.4.5(1) (S. 199) existiert daher auch $\int_1^\infty t^{x-1}e^{-t}dt$. Es ist

$$\begin{aligned}\Gamma(1) &= \int_0^\infty e^{-t}dt\\ &= \lim_{\beta\to+\infty}\int_0^\beta e^{-t}dt\\ &= \lim_{\beta\to+\infty}(1-e^{-\beta})\\ &= 1.\end{aligned}$$

Für $x \in \mathbb{R}_+^*$ und $0 < \alpha < \beta < +\infty$ gilt weiter

$$\int_\alpha^\beta t^x e^{-t}dt = -t^x e^{-t}\Big|_\alpha^\beta + x\int_\alpha^\beta t^{x-1}e^{-t}dt.$$

Wegen

$$\lim_{\alpha\to0+}\alpha^x e^{-\alpha} = 0 \ \text{ und } \ \lim_{\beta\to+\infty}\beta^x e^{-\beta} = 0$$

folgt hieraus (ii).
Die Beziehung (iii) folgt direkt aus (ii) und (i). □

Für die weitere Analyse der Eulerschen Gammafunktion benötigen wir den folgenden Begriff.

Definition VI.5.2. Sei $J \subset \mathbb{R}$ ein Intervall und $f : J \to \mathbb{R}_+^*$ eine Funktion. $f$ heißt logarithmisch konvex, falls die Funktion $\ln(f) : J \to \mathbb{R}$ konvex ist.

Bemerkung VI.5.3. Es gelten folgende Eigenschaften logarithmisch konvexer Funktionen:

(1) $f$ logarithmisch konvex $\iff f(tx + (1-t)y) \leq f(x)^t f(y)^{1-t}$ für alle $x, y \in J, t \in [0, 1]$.
(2) $f, g$ logarithmisch konvex $\Longrightarrow f \cdot g$ logarithmisch konvex.
(3) $f_n \xrightarrow[n\to\infty]{\text{pkt}} f$, $f_n$ logarithmisch konvex für alle $n \Longrightarrow f$ logarithmisch konvex.
(4) $f : J \to \mathbb{R}_+^*$ sei zweimal differenzierbar. Dann gilt: $f$ logarithmisch konvex $\iff f \cdot f'' - (f')^2 \geq 0$.
(5) $f, g : J \to \mathbb{R}_+^*$ seien zweimal differenzierbar und logarithmisch konvex. Dann ist $f + g$ logarithmisch konvex.
(6) Sei $-\infty \leq a < b \leq +\infty$ und $f : (a, b) \times J \to \mathbb{R}_+^*$. Für jedes $t \in (a, b)$ sei $f(t, .) : J \to \mathbb{R}_+^*$ zweimal differenzierbar und logarithmisch konvex. Für alle $x \in J$ existiere
$$F(x) = \int_a^b f(t, x)dt.$$
Dann ist $F$ logarithmisch konvex.
(7) Sei $-\infty \leq a < b \leq +\infty$ und $\varphi \in C((a, b), \mathbb{R}_+^*)$. Dann ist die Funktion $x \mapsto \int_a^b \varphi(t)t^{x-1}dt$ auf ihrem Definitionsbereich logarithmisch konvex.
(8) $f : J \to \mathbb{R}_+^*$ sei logarithmisch konvex. Dann ist für jedes $a \in \mathbb{R}$ die Funktion $x \mapsto f(x + a)$ logarithmisch konvex auf ihrem Definitionsbereich.

Beweis. ad (1): $f$ logarithmisch konvex
$$\begin{aligned} &\iff \ln(f) \quad \text{konvex} \\ &\iff \ln f(tx + (1-t)y) \leq t \ln f(x) + (1-t) \ln f(y) \\ &\iff f(tx + (1-t)y) \leq f(x)^t f(y)^{1-t} \quad \forall x, y \in J, t \in [0, 1]. \end{aligned}$$
ad (2): Folgt direkt aus (1).
ad (3): Folgt direkt aus (1).
ad (4): $\ln(f)$ ist zweimal differenzierbar und
$$(\ln(f))'' = (\frac{f'}{f})' = \frac{f''f - (f')^2}{f^2}.$$

Damit folgt die Behauptung aus Satz IV.2.15 (S. 126).
AD (5): Folgt aus

$$\begin{aligned}&(f''+g'')(f+g)-(f'+g')^2\\ =&f''f+f''g+fg''+g''g-(f')^2-2f'g'-(g')^2\\ =&f''f-(f')^2+g''g-(g')^2+f''g+fg''-2f'g'\\ =&f''f-(f')^2+g''g-(g')^2+\frac{1}{fg}(fg'-f'g)^2\\ &+\frac{g}{f}(f''f-(f')^2)+\frac{f}{g}(g''g-(g')^2)\\ \geq&0.\end{aligned}$$

AD (6): Wegen

$$F(x)=\lim_{\substack{\alpha\to a+\\ \beta\to b-}}\int_\alpha^\beta f(t,x)dt$$

und Teil (3) reicht es zu zeigen, dass

$$x\mapsto\int_\alpha^\beta f(t,x)dt$$

für alle $a<\alpha<\beta<b$ logarithmisch konvex ist. Wegen Satz VI.1.11 (S. 177) ist aber

$$\int_\alpha^\beta f(t,x)dt=\lim_{n\to\infty}\sum_{k=0}^{n-1}f(\alpha+k\frac{\beta-\alpha}{n},x)\frac{\beta-\alpha}{n}.$$

Damit folgt die Behauptung aus Teil (3) und (5).
AD (7): Folgt aus (6) angewandt auf

$$f(t,x)=\varphi(t)t^{x-1},$$

da nach Teil (1) $x\mapsto t^x$ logarithmisch konvex ist.
AD (8): Sei $g(x)=f(x+a)$. Die Behauptung folgt aus Teil (1) und

$$g(tx+(1-t)y)=f(t(x+a)+(1-t)(y+a)).$$

□

Wir kommen nun zur Charakterisierung der Gammafunktion.

SATZ VI.5.4. *Die Eulersche Gammafunktion ist die einzige Funktion $f:\mathbb{R}_+^*\to\mathbb{R}_+^*$ mit den folgenden drei Eigenschaften:*

(1) $f(1)=1$,
(2) $f(x+1)=xf(x)$ *für alle* $x\in\mathbb{R}_+^*$,
(3) $f$ *ist logarithmisch konvex.*

BEWEIS. Aus Satz VI.5.1 und Bemerkung VI.5.3 (7) folgt, dass $\Gamma$ die Eigenschaften (1)–(3) erfüllt. Sei nun $f:\mathbb{R}_+^*\to\mathbb{R}_+^*$ eine beliebige Funktion mit den Eigenschaften (1)–(3). Dann folgt durch Induktion

$$f(n+1)=n!\quad\forall n\in\mathbb{N}$$

$$f(x+n) = f(x)x \cdot (x+1) \cdots (x+n-1) \quad \forall n \in \mathbb{N}^*, x \in \mathbb{R}_+^*.$$

Also ist $f$ eindeutig bestimmt durch seine Werte auf $(0,1)$.
Sei nun $0 < x < 1$. Aus

$$n + x = (1-x)n + x(n+1)$$

und (3) folgt

$$\begin{aligned} f(n+x) &\le f(n)^{1-x} f(n+1)^x \\ &= f(n)^{1-x} f(n)^x n^x \\ &= f(n) n^x \\ &= (n-1)! n^x. \end{aligned}$$

Aus

$$n + 1 = x(n+x) + (1-x)(n+1+x)$$

folgt

$$\begin{aligned} n! = f(n+1) &\le f(n+x)^x f(n+1+x)^{1-x} \\ &= f(n+x)^x f(n+x)^{1-x} (n+x)^{1-x} \\ &= f(n+x)(n+x)^{1-x}. \end{aligned}$$

Also gilt

$$\begin{aligned} n!(n+x)^{x-1} &\le f(n+x) \\ &\le (n-1)! n^x \\ \Longrightarrow & \frac{n!(n+x)^{x-1}}{x \cdot (x+1) \cdots (x+n-1)} \\ &\le f(x) \\ &\le \frac{(n-1)! n^x}{x(x+1) \cdots (x+n-1)} \\ \Longrightarrow & (1 + \frac{x}{n})^x \frac{n! n^x}{x(x+1) \cdots (x+n-1)(x+n)} \\ &\le f(x) \\ &\le \frac{n! n^x}{x(x+1) \cdots (x+n-1)(x+n)} (1 + \frac{x}{n}) \\ \Longrightarrow \quad f(x) \underset{n \to \infty}{\longleftarrow} & (\frac{n}{n+x}) f(x) \\ &\le \frac{n! n^x}{x(x+1) \cdots (x+n)} \\ &\le (1 + \frac{x}{n})^{-x} f(x) \\ &\underset{n \to \infty}{\longrightarrow} f(x). \end{aligned}$$

Also gilt für alle $x \in (0,1)$

$$f(x) = \lim_{n\to\infty} \frac{n!n^x}{x(x+1)\cdots(x+n)}.$$

Damit ist die Eindeutigkeit von $f$ gezeigt. □

Bemerkung VI.5.5. Aus dem Beweis von Satz VI.5.4 folgt die Gaussche Darstellung der Gammafunktion

$$\Gamma(x) = \lim_{n\to\infty} \frac{n!n^x}{x(x+1)\cdots(x+n)}.$$

Der folgende Satz stellt einen Zusammenhang her zwischen dem Eulerschen Beta-Integral und der Eulerschen Gammafunktion.

Satz VI.5.6. *Für alle $x, y \in \mathbb{R}_+^*$ gilt*

$$\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = B(x,y).$$

Beweis. Sei $y \in \mathbb{R}_+^*$ beliebig und

$$f(x) = B(x,y)\frac{\Gamma(x+y)}{\Gamma(y)}.$$

Dann folgt aus Beispiel VI.4.11(4) (S. 202)

$$\begin{aligned} f(1) &= B(1,y)\frac{\Gamma(y+1)}{\Gamma(y)} \\ &= yB(1,y) \\ &= yB(y,1) \\ &= y \lim_{\alpha\to 0+} \int_\alpha^1 t^{y-1}dt \\ &= \lim_{\alpha\to 0+} (1-\alpha^y) \\ &= 1 \end{aligned}$$

und

$$\begin{aligned} f(x+1) &= B(x+1,y)\frac{\Gamma(x+1+y)}{\Gamma(y)} \\ &= \frac{x}{x+y}B(x,y)(x+y)\frac{\Gamma(x+y)}{\Gamma(y)} \\ &= xf(x). \end{aligned}$$

Wegen Bemerkung VI.5.3 (8) ist $x \mapsto \Gamma(x+y)$ logarithmisch konvex. Wegen Bemerkung VI.5.3 (7) ist $x \mapsto B(x,y)$ logarithmisch konvex. Wegen Bemerkung VI.5.3 (2) ist daher $f$ logarithmisch konvex. Damit folgt die Behauptung aus Satz VI.5.4. □

Das folgende Ergebnis für die Stochastik wesentlich.

SATZ VI.5.7. *Es gilt:*

(1) $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

(2) GAUSSSCHES FEHLERINTEGRAL: $\displaystyle\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$.

BEWEIS. AD (1): Aufgrund unserer vorherigen Ergebnisse gilt

$$\begin{aligned}
[\Gamma(\frac{1}{2})]^2 &= \frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(1)} \\
&= B(\frac{1}{2}, \frac{1}{2}) \\
&= \int_0^1 \frac{1}{\sqrt{t(1-t)}} dt \\
&= \lim_{\substack{\alpha \to 0+ \\ \beta \to 1-}} \int_\alpha^\beta \frac{1}{\sqrt{t(1-t)}} dt \\
&\qquad\qquad t = \sin^2 \varphi \\
&\qquad\qquad dt = 2 \sin\varphi \cos\varphi d\varphi \\
&= \lim_{\substack{\alpha \to 0+ \\ \beta \to 1-}} \int_{\arcsin\sqrt{\alpha}}^{\arcsin\sqrt{\beta}} \frac{2\sin\varphi\cos\varphi}{\sqrt{\sin^2\varphi\cos^2\varphi}} d\varphi \\
&= 2\int_0^{\frac{\pi}{2}} d\varphi \\
&= \pi.
\end{aligned}$$

AD (2): Es ist

$$\begin{aligned}
\Gamma(\frac{1}{2}) &= \int_0^\infty t^{-\frac{1}{2}} e^{-t} dt \\
&= \lim_{\substack{\alpha \to 0+ \\ \beta \to +\infty}} \int_\alpha^\beta t^{-\frac{1}{2}} e^{-t} dt \qquad t = s^2 \quad dt = 2sds \\
&= 2 \lim_{\substack{\alpha \to 0+ \\ \beta \to +\infty}} \int_{\sqrt{\alpha}}^{\sqrt{\beta}} e^{-s^2} ds \\
&= \lim_{\substack{\alpha \to 0+ \\ \beta \to +\infty}} \{\int_{\sqrt{\alpha}}^{\sqrt{\beta}} e^{-s^2} ds + \underbrace{\int_{\sqrt{\alpha}}^{\sqrt{\beta}} e^{-s^2} ds}_{s=-r}\} \\
&= \lim_{\substack{\alpha \to 0+ \\ \beta \to +\infty}} \{\int_{\sqrt{\alpha}}^{\sqrt{\beta}} e^{-s^2} ds - \int_{-\sqrt{\alpha}}^{-\sqrt{\beta}} e^{-r^2} dr\} \\
&= \int_{-\infty}^{\infty} e^{-x^2} dx.
\end{aligned}$$

$\square$

KAPITEL VII

# Differentialrechnung mehrerer Variablen

Als Vorbereitung stellen wir zunächst einige Eigenschaften stetiger linearer Abbildungen zusammen. Anschließend definieren wir die Ableitung von Funktionen zwischen beliebigen Vektorräumen als geeignete Approximation durch lineare Abbildungen. Als Spezialfall erhalten wir Richtungs- und partielle Ableitungen, sowie Differenzierbarkeitskriterien. Danach stellen wir einige Rechenregeln wie die Produkt- und Kettenregel zusammen.
Als Vorbereitung auf höhere Ableitungen betrachten wir multilineare Abbildungen. Nach der Einführung höherer Ableitungen betrachten wir Taylorformeln und als Anwendung lokale Extrema von Funktionen mehrerer Variablen.
Zum Abschluss betrachten wir umkehrbare Abbildungen und beweisen den Satz über implizite Funktionen.

## VII.1. Stetige lineare Abbildungen

Im Folgenden seien stets $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ zwei normierte $\mathbb{K}$-Vektorräume. Wir untersuchen in diesem Abschnitt den Raum $\mathcal{L}(X, Y)$ der stetigen linearen Abbildungen von $X$ in $Y$ (vgl. Definition VI.1.5 (S. 173)).

Satz VII.1.1. *Ist $Y$ ein Banachraum, so ist $(\mathcal{L}(X, Y), \|\cdot\|_{\mathcal{L}(X,Y)})$ ebenfalls ein Banachraum.*

Beweis. Sei $(A_n)_{n\in\mathbb{N}} \subset \mathcal{L}(X, Y)$ eine Cauchyfolge Dann ist für jedes $x \in X$ die Bildfolge $(A_n x)_{n\in\mathbb{N}} \subset Y$ eine Cauchyfolge und damit konvergent. Durch

$$(*) \qquad Ax = \lim_{n\to\infty} A_n x$$

wird die Abbildung $A : X \to Y$ definiert. Aus den Rechenregeln für Grenzwerte folgt, dass $A$ linear ist. Mit $(A_n)_{n\in\mathbb{N}}$ ist auch die Folge der Normen $(\|A_n\|_{\mathcal{L}(X,Y)})_{n\in\mathbb{N}}$ eine Cauchyfolge und damit beschränkt. Aus $(*)$ folgt

$$\|A\|_{\mathcal{L}(X,Y)} \le \sup_{n\in\mathbb{N}} \|A_n\|_{\mathcal{L}(X,Y)} < \infty.$$

Also ist $A \in \mathcal{L}(X, Y)$. Sei schließlich $\varepsilon > 0$ beliebig. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit

$$\|A_n - A_m\|_{\mathcal{L}(X,Y)} < \varepsilon \qquad \forall n, m \ge n_\varepsilon.$$

Daher gilt für jedes $x \in X$ mit $\|x\|_X = 1$

$$\begin{aligned}
&\|A_n x - A_m x\|_Y \le \|A_n - A_m\|_{\mathcal{L}(X,Y)} < \varepsilon \qquad \forall n, m \ge n_\varepsilon \\
\underset{m\to\infty}{\Longrightarrow} &\|A_n x - Ax\|_Y \le \varepsilon \quad \forall n \ge n_\varepsilon, x \in X \ \|x\|_X = 1 \\
\Longrightarrow &\|A_n - A\|_{\mathcal{L}(X,Y)} \le \varepsilon \quad \forall n \ge n_\varepsilon.
\end{aligned}$$

Also konvergiert $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{L}(X,Y)$ gegen $A$. □

Definition VII.1.2. (1) Eine bijektive, stetige, lineare Abbildung $A : X \to Y$ mit stetiger Inverser $A^{-1}$ heißt topologischer Isomorphismus oder kurz Isomorphismus.
(2) Die Menge der topologischen Isomorphismen zwischen $X$ und $Y$ wird mit $\mathrm{Isom}(X, Y)$ bezeichnet.
(3) $A \in \mathrm{Isom}(X, Y)$ heißt Isometrie, wenn gilt

$$\|Ax\|_Y = \|x\|_X \quad \forall x \in X.$$

(4) Die Räume $X$ und $Y$ heißen topologisch isomorph, kurz $X \cong Y$, wenn $\mathrm{Isom}(X, Y) \neq \emptyset$ ist.

Bemerkung VII.1.3. (1) $\cong$ ist eine Äquivalenzrelation.
(2) $A : X \to Y$ sei bijektiv und linear. Dann ist $A \in \mathrm{Isom}(X, Y)$ genau dann, wenn es zwei Konstanten $0 < \underline{c} \le \overline{c} < \infty$ gibt mit

$$\underline{c}\|x\|_X \le \|Ax\|_Y \le \overline{c}\|x\|_Y \quad \forall x \in X.$$

Der folgende Satz zeigt, dass alle Räume mit gleicher endlicher Dimension isomorph sind.

Satz VII.1.4. *Sei $m = \dim X < \infty$. Dann ist $X \cong \mathbb{K}^m$.*

Beweis. Wir versehen $\mathbb{K}^m$ mit der euklidischen Norm $\|\cdot\|_2$. Sei $(b_1, \ldots, b_m)$ eine Basis von $X$. Zu jedem $x \in X$ gibt es genau ein $\zeta = (\zeta_1, \ldots, \zeta_m) \in \mathbb{K}^m$ mit

$$x = \sum_{i=1}^m \zeta_i b_i.$$

Die Abbildung $A : x \mapsto \zeta$ ist linear mit der Inversen

$$A^{-1} : \zeta \mapsto \sum_{i=1}^m \zeta_i b_i.$$

Für $\zeta \in \mathbb{K}^m$ gilt wegen der Cauchy-Schwarzschen Ungleichung

$$\begin{aligned}
\|A^{-1}\zeta\|_X &= \|\sum_{i=1}^m \zeta_i b_i\|_X \\
&\le \sum_{i=1}^m |\zeta_i| \|b_i\|_X
\end{aligned}$$

$$\leq \left\{ \sum_{i=1}^{m} \|b_i\|_X^2 \right\}^{1/2} \|\zeta\|_2.$$

Also ist $A^{-1} \in \mathcal{L}(\mathbb{K}^m, X)$. Für $\zeta \in \mathbb{K}^m$ sei

$$\|\zeta\|_* = \|A^{-1}\zeta\|_X.$$

Wie man leicht nachprüft, ist $\|\cdot\|_*$ eine Norm auf $\mathbb{K}^m$. Wegen Satz III.1.11 (S. 64) gibt es ein $\beta \in \mathbb{R}_+^*$ mit

$$\|\zeta\|_2 \leq \beta\|\zeta\|_* \qquad \forall \zeta \in \mathbb{K}^m.$$

Damit folgt für $x \in X$

$$\|Ax\|_2 \leq \beta\|Ax\|_* = \beta\|x\|_X.$$

Also ist auch $A \in \mathcal{L}(X, \mathbb{K}^m)$. Damit ist die Behauptung bewiesen. □

Aus Satz VII.1.4 folgt:

Satz VII.1.5. *Sei* $\dim X < \infty$. *Dann sind alle Normen auf* $X$ *äquivalent und* $X$ *ist vollständig.*

Beweis. Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei Normen auf $X$ und $A \in \mathrm{Isom}(\mathbb{K}^m, X)$. Dann sind

$$|\zeta|_1 = \|A\zeta\|_1 \text{ und } |\zeta|_2 = \|A\zeta\|_2$$

zwei Normen auf $\mathbb{K}^m$. Wegen Satz III.1.11 (S. 64) gibt es zwei Konstanten $0 < \alpha \leq \beta < +\infty$ mit

$$\begin{aligned} &\alpha|\zeta|_1 \leq |\zeta|_2 \leq \beta|\zeta|_1 \quad \forall \zeta \in \mathbb{K}^m \\ \Longrightarrow\ &\alpha\|x\|_1 = \alpha|A^{-1}x|_1 \\ &\qquad \leq |A^{-1}x|_2 = \|x\|_2 \\ &\qquad\qquad \leq \beta|A^{-1}x|_1 = \beta\|x\|_1 \ \forall x \in X. \end{aligned}$$

Also sind $\|.\|_1$ und $\|.\|_2$ äquivalent.
Sei nun $(x_n)_{n\in\mathbb{N}} \subset X$ eine Cauchyfolge Dann ist $(\zeta_n)_{n\in\mathbb{N}} \subset \mathbb{K}^m$ mit $\zeta_n = A^{-1}x_n$ eine Cauchyfolge in $\mathbb{K}^m$ und wegen Satz III.1.14 (S. 66) konvergent. Sei $\zeta = \lim\limits_{n\to\infty} \zeta_n$ und $x = A\zeta$. Wegen der Stetigkeit von $A$ folgt

$$x = \lim_{n\to\infty} A\zeta_n = \lim_{n\to\infty} x_n.$$

Also ist $X$ vollständig. □

Bemerkung VII.1.6. Bemerkung III.1.12 (S. 66) und Bemerkung V.4.9 (S. 169) zeigen, dass Satz VII.1.5 für unendlich dimensionale Vektorräume nicht gilt.

Lineare Abbildungen auf endlich dimensionalen Räumen sind immer stetig.

Satz VII.1.7. *Sei* $\dim X < \infty$ *und* $Y$ *beliebig. Dann ist jede lineare Abbildung* $X \to Y$ *stetig.*

Beweis. Sei $A : X \to Y$ eine lineare Abbildung, $m = \dim X$ und $b_1, \ldots, b_m$ eine Basis von $X$. Dann gilt für jedes

$$x = \sum_{i=1}^{m} \zeta_i b_i \in X$$

die Abschätzung

$$\begin{aligned}\|Ax\|_Y &= \|\sum_{i=1}^{m} \zeta_i Ab_i\|_Y \\ &\leq \sum_{i=1}^{m} |\zeta_i| \|Ab_i\|_Y \\ &\leq \alpha \|\zeta\|_2\end{aligned}$$

mit

$$\alpha = \left\{ \sum_{i=1}^{m} \|Ab_i\|_Y^2 \right\}^{1/2}$$

und

$$\zeta = (\zeta_1, \ldots, \zeta_m) \in \mathbb{K}^m.$$

Da gemäß dem Beweis von Satz VII.1.4 die Abbildung $x \mapsto \zeta$ aus $\mathcal{L}(X, \mathbb{K}^m)$ ist, gibt es ein $\beta > 0$ mit

$$\|\zeta\|_2 \leq \beta \|x\|_X \quad \forall x \in X.$$

Damit folgt

$$\|Ax\|_Y \leq \alpha\beta \|x\|_X \quad \forall x \in X.$$

□

Bemerkung VII.1.8. (1) Sei $(X, \|\cdot\|_X) = (\ell_1, \|\cdot\|_\infty)$ und $(Y, \|\cdot\|_Y) = (\ell_1, \|\cdot\|_1)$ und $A : X \to Y$ die Identität. Wegen Beispiel III.1.12 (S. 66) ist $A$ nicht stetig. Dies zeigt, dass Satz VII.1.7 für unendlich dimensionale Räume nicht gilt.
(2) Sei $(X, \|\cdot\|_X) = (C^1([0,1], \mathbb{R}), \|\cdot\|_\infty)$ und $(Y, \|\cdot\|_Y) = (C([0,1], \mathbb{R}), \|\cdot\|_\infty)$ und $A : X \to Y$ definiert durch $f \mapsto f'$. Dann ist $A$ nicht stetig. Denn sei

$$f_n(x) = \frac{1}{n} \sin(n^2 x) \quad \forall x \in [0,1], n \in \mathbb{N}^*.$$

Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen 0, aber $Af_n = f_n' = n \cos(n^2 x)$ konvergiert nicht.

Als nächstes stellen wir einen Zusammenhang her zwischen linearen Abbildungen auf endlich dimensionalen Räumen und Matrizen.

DEFINITION VII.1.9. Wir bezeichnen mit

$$M_{m,n}(\mathbb{K}) = \mathbb{K}^{m\times n} = \{(a_{ij})_{1\le i\le m, 1\le j\le n} : a_{ij} \in \mathbb{K}\}$$

die Menge aller $m \times n$ Matrizen. $M_{m,n}(\mathbb{K})$ ist ein Vektorraum mit der üblichen komponentenweise Addition und Multiplikation mit Skalaren. Durch

$$\|(a_{ij})_{1\le i\le m, 1\le j\le n}\|_2 = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2}$$

wird eine Norm, die sog. FROBENIUS-NORM, auf $M_{m,n}(\mathbb{K})$ definiert, und es ist

$$M_{m,n}(\mathbb{K}) \cong \mathbb{K}^{m\cdot n}.$$

SATZ VII.1.10. *Seien $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ und $(Z, \|\cdot\|_Z)$ endlich dimensionale Vektorräume und $l = \dim X$, $m = \dim Y$, $n = \dim Z$. Dann ist*

$$\mathcal{L}(X, Y) \cong M_{m,l}(\mathbb{K})$$

*und die Verknüpfung von Abbildungen entspricht der Multiplikation von Matrizen.*

BEWEIS. Seien $(e_1, \ldots, e_l)$ eine Basis von $X$, $(f_1, \ldots, f_m)$ eine Basis von $Y$ und $(g_1, \ldots, g_n)$ eine Basis von $Z$, sowie $A \in \mathcal{L}(X, Y)$, $B \in \mathcal{L}(Y, Z)$. Dann ist

$$Ae_i = \sum_{j=1}^{m} a_{ji} f_j \quad \forall 1 \le i \le l$$

und

$$\begin{aligned} Ax &= A(\sum_{i=1}^{l} \zeta_i e_i) \\ &= \sum_{i=1}^{l}\sum_{j=1}^{m} \zeta_i a_{ji} f_j \\ &= \sum_{j=1}^{m} \{\sum_{i=1}^{l} a_{ji}\zeta_i\} f_j. \end{aligned}$$

Die Zuordnung $A \mapsto (a_{ji})_{1\le j\le m, 1\le i\le l} \in M_{m,l}(\mathbb{K})$ ist offensichtlich bijektiv und linear. Insbesondere folgt

$$\dim(\mathcal{L}(X, Y)) = \dim(M_{m,l}(\mathbb{K})) = m \cdot l.$$

Wegen Satz VII.1.7 ist diese Zuordnung ein topologischer Isomorphismus. Dies beweist

$$\mathcal{L}(X, Y) \cong M_{m,l}(\mathbb{K}).$$

Weiter gilt

$$Bf_j = \sum_{k=1}^{n} b_{kj} g_k \quad \forall 1 \le j \le m.$$

Damit folgt für $x \in X$

$$\begin{aligned} B \circ Ax &= B\left(\sum_{j=1}^{m} \Big\{\sum_{i=1}^{l} a_{ji}\zeta_i\Big\} f_j\right) \\ &= \sum_{j=1}^{m} \left\{\Big\{\sum_{i=1}^{l} a_{ji}\zeta_i\Big\} \sum_{k=1}^{n} b_{kj} g_k\right\} \\ &= \sum_{k=1}^{n} \left\{\sum_{i=1}^{l} \Big[\sum_{j=1}^{m} b_{kj} a_{ji}\Big] \zeta_i\right\} g_k. \end{aligned}$$

Dies zeigt, dass die Hintereinanderschaltung von Abbildungen der Matrizenmultiplikation entspricht. □

Bemerkung VII.1.11. (1) Wir verwenden im $\mathbb{K}^m$ immer die Standardbasis $e_1, \ldots, e_m$ mit

$$e_i = (0, \ldots, 0, \underset{\substack{\uparrow \\ i}}{1}, 0, \ldots, 0).$$

(2) Man beachte, dass eine lineare Abbildung $X \to Y$ ein universelles Gebilde ist, während ihre Matrixdarstellung von den speziell gewählten Basen von $X$ und $Y$ abhängt. Daher sind universelle Eigenschaften einer linearen Abbildung wie z.B. Bijektivität oder Stetigkeit von der Wahl der Basis unabhängig.

## VII.2. Differenzierbarkeit

Im Folgenden bezeichnen $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ $\mathbb{K}$-Banachräume und $U \subset X, U \neq \emptyset$, eine offene Menge.

Zuerst definieren wir den Begriff der Differenzierbarkeit und zeigen, dass die neue Definition für Funktionen einer Variablen mit der alten aus Paragraph IV.1 übereinstimmt.

Definition VII.2.1. Eine Funktion $f : U \to Y$ heißt in $x_0$ differenzierbar, wenn es ein $A_{x_0} \in \mathcal{L}(X, Y)$ gibt mit

$$\lim_{x \to x_0} \frac{1}{\|x - x_0\|_X} \{f(x) - f(x_0) - A_{x_0}(x - x_0)\} = 0.$$

Satz VII.2.2. (1) *Die Funktion $f : U \to Y$ ist genau dann in $x_0 \in U$ differenzierbar, wenn es ein $A_{x_0} \in \mathcal{L}(X, Y)$ und eine in $x_0$ stetige Funktion $r_{x_0} : U \to Y$ mit $r_{x_0}(x_0) = 0$ und*

$$f(x) = f(x_0) + A_{x_0}(x - x_0) + r_{x_0}(x)\|x - x_0\|_X \quad \forall x \in U$$

*gibt.*
(2) *Ist $f : U \to Y$ in $x_0 \in U$ differenzierbar, so ist $f$ in $x_0$ stetig.*
(3) *Ist $f : U \to Y$ in $x_0 \in U$ differenzierbar, so ist $A_{x_0}$ eindeutig definiert.*

Beweis. ad (1): Definiere $r_{x_0} : U \to Y$ durch

$$r_{x_0}(x) = \begin{cases} 0 & \text{für } x = x_0, \\ \frac{1}{\|x-x_0\|_X}[f(x) - f(x_0) - A_{x_0}(x - x_0)] & \text{für } x \neq x_0. \end{cases}$$

Dann ist $r_{x_0}$ genau dann in $x_0$ stetig, wenn $f$ in $x_0$ differenzierbar ist.
ad (2): Folgt aus Teil (1).
ad (3): Es gelte für alle $x \in U$

$$\begin{aligned} f(x) &= f(x_0) + A(x - x_0) + s(x)\|x - x_0\|_X \\ &= f(x_0) + B(x - x_0) + r(x)\|x - x_0\|_X \end{aligned}$$

mit $A, B \in \mathcal{L}(X, Y)$, $s(x_0) = r(x_0) = 0$ und $s$ und $r$ in $x_0$ stetig. Sei $y \in X\backslash\{0\}$ und $\varepsilon \in \mathbb{R}_+^*$. Dann folgt

$$\begin{aligned} (A - B)y &= \frac{1}{\varepsilon}(A - B)(\varepsilon y) \\ &= \frac{1}{\varepsilon}[r(x_0 + \varepsilon y) - s(x_0 + \varepsilon y)]\|\varepsilon y\|_X \\ &= [r(x_0 + \varepsilon y) - s(x_0 + \varepsilon y)]\|y\|_X \\ &\underset{\varepsilon \to 0}{\longrightarrow} 0. \end{aligned}$$

Also ist $(A - B)y = 0$ für alle $y \in X\backslash\{0\}$ und somit $A = B$. □

Wegen Satz VII.2.2 ist folgende Definition sinnvoll.

Definition VII.2.3. (1) Sei $f : U \to Y$ in $x_0 \in U$ differenzierbar. Dann ist

$$Df(x_0) = f'(x_0) = A_{x_0} \in \mathcal{L}(X, Y)$$

die Ableitung von $f$ in $x_0$.
(2) $f : U \to Y$ heißt differenzierbar, wenn $f$ in $x_0$ differenzierbar ist für jedes $x_0 \in U$.
(3) $f : U \to Y$ heißt stetig differenzierbar, wenn $f$ differenzierbar und $Df \in C(U, \mathcal{L}(X, Y))$ ist. Die Menge der stetig differenzierbaren Abbildungen $U \to Y$ wird mit $C^1(U, Y)$ bezeichnet.

Bemerkung VII.2.4. (1) Die Differenzierbarkeit und die Ableitungen sind unabhängig von den speziellen, äquivalenten Normen auf $X$ und $Y$.
(2) Sei $X = Y = \mathbb{K}$. Dann ist

$$\mathcal{L}(X, Y) \cong M_{1,1}(\mathbb{K}) \cong \mathbb{K}$$

und $f : U \to Y$ ist genau dann in $x_0 \in U$ differenzierbar, wenn es ein $f'(x_0) \in \mathbb{K}$ gibt mit

$$\begin{aligned}
&\lim_{x\to x_0} \frac{1}{|x-x_0|}[f(x)-f(x_0)-f'(x_0)(x-x_0)] = 0 \\
\iff &\frac{|f(x)-f(x_0)-f'(x_0)(x-x_0)|}{|x-x_0|} \underset{x\to x_0}{\longrightarrow} 0 \\
\iff &\frac{f(x)-f(x_0)}{x-x_0} \underset{x\to x_0}{\longrightarrow} f'(x_0).
\end{aligned}$$

Also stimmt für $X = Y = \mathbb{K}$ obige Definition der Differenzierbarkeit mit der alten Definition IV.1.1 (S. 111) überein.
(3) Sei $X = \mathbb{K}$ und $Y$ beliebig. Sei $A \in \mathcal{L}(\mathbb{K}, Y)$. Dann ist

$$A(x) = A(x \cdot 1) = xA(1) = xv$$

mit

$$v = A(1)$$

und

$$\begin{aligned}
&\|A(x)\|_Y = |x|\|v\|_Y \\
\Longrightarrow &\|A\|_{\mathcal{L}(\mathbb{K},Y)} = \|v\|_Y = \|A(1)\|_Y.
\end{aligned}$$

Also ist die Zuordnung $A \mapsto A(1)$ eine Isometrie von $\mathcal{L}(\mathbb{K}, Y)$ auf $Y$, vermittels derer wir $\mathcal{L}(\mathbb{K}, Y)$ und $Y$ identifizieren können. Hieraus folgt, dass obige Definition der Differenzierbarkeit für $X = \mathbb{K}$, $Y$ beliebig mit der alten Definition IV.1.1 (S. 111) übereinstimmt.

Beispiel VII.2.5. Sei $A \in \mathcal{L}(X, Y)$. Wegen

$$\begin{aligned}
Ax &= A(x_0 + x - x_0) \\
&= Ax_0 + A(x - x_0)
\end{aligned}$$

ist $A \in C^1(X, Y)$ und

$$A'(x_0) = A \quad \forall x_0 \in Y.$$

Der folgende Begriff der Richtungsableitung stimmt für Funktionen einer reellen Variablen mit dem Begriff der Differenzierbarkeit überein. Wie wir sehen werden, gibt es für Funktionen mehrerer Variablen zwischen den beiden Begriffen wesentliche Unterschiede.

Definition VII.2.6. Sei $f : U \to Y$, $x_0 \in U$ und $v \in X\backslash\{0\}$. Dann heißt die Größe

$$D_v f(x_0) = \lim_{t\to 0} \frac{1}{t}[f(x_0 + tv) - f(x_0)],$$

sofern sie existiert, die Richtungsableitung von $f$ in $x_0$ in Richtung $v$.

Satz VII.2.7. *Sei $f : U \to Y$ in $x_0 \in U$ differenzierbar. Dann existiert $D_v f(x_0)$ für alle $v \in X \backslash \{0\}$ und*

$$D_v f(x_0) = Df(x_0)v \qquad \forall v \in X\backslash\{0\}.$$

Beweis. Sei $v \in X\backslash\{0\}$. Dann folgt

$$\begin{aligned} &\|\frac{1}{t}[f(x_0 + tv) - f(x_0)] - Df(x_0)v\|_Y \\ =&\|v\|_X \|\frac{1}{\|tv\|_X}[f(x_0 + tv) - f(x_0) - Df(x_0)(tv)]\|_Y \\ \xrightarrow[t\to 0]{}& 0. \end{aligned}$$

Hieraus folgt die Behauptung. □

Bemerkung VII.2.8. (1) Die Umkehrung von Satz VII.2.7 ist falsch. Betrachte z. B. die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

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

Sei $v = (\alpha, \beta) \in \mathbb{R}^2\backslash\{0\}$. Dann ist

$$\frac{1}{t}[f(tv) - f(0)] = \frac{1}{t}\frac{t^3\alpha^2\beta}{t^2(\alpha^2 + \beta^2)} = f(v) \quad \forall t \in \mathbb{R}.$$

Also ist

$$D_v f(0) = f(v),$$

und ist somit insbesondere keine lineare Funktion von $v$. Daher kann $f$ nicht in 0 differenzierbar sein.
(2) Sei $f : U \to Y$, $x_0 \in U$ und $v \in X\backslash\{0\}$. Dann gibt es ein $\varepsilon > 0$ mit $B(x_0, \varepsilon) \subset U$. Definiere die Funktion $\varphi : (-\varepsilon, \varepsilon) \to Y$ durch

$$\varphi(t) = f(x_0 + tv).$$

Dann ist $\varphi$ genau dann in 0 differenzierbar, wenn $D_v f(x_0)$ existiert, und in diesem Fall ist

$$\varphi'(0) = D_v f(x_0).$$

Als nächstes betrachten wir speziell die Richtungsableitungen in Richtung der Einheitsvektoren $e_j$ in $\mathbb{R}^m$.

Definition VII.2.9. Sei $X = \mathbb{R}^m$, $f : U \to Y$ und $x_0 \in U$. Dann heißt die Größe

$$D_j f(x_0) = \frac{\partial f}{\partial x_j}(x_0) = D_{e_j} f(x_0), \quad 1 \leq j \leq m,$$

sofern sie existiert, die $j$-te partielle Ableitung von $f$ im Punkt $x_0$. $f$ heißt in $x_0$ partiell differenzierbar, wenn alle partiellen Ableitungen $D_j f(x_0)$ existieren.

Bemerkung VII.2.10. (1) Es ist

$$\begin{aligned} D_j f(x_0) &= \lim_{t\to 0} \frac{1}{t}[f(x_{0,1},\dots,x_{0,j-1},x_{0,j}+t,x_{0,j+1},\dots,x_{0,m}) \\ &\qquad - f(x_{0,1},\dots,x_{0,m})] \\ &= \lim_{\zeta\to x_{0,j}} \frac{1}{\zeta - x_{0,j}}[f(x_{0,1},\dots,x_{0,j-1},\zeta,x_{0,j+1},\dots,x_{0,m}) \\ &\qquad - f(x_{0,1},\dots,x_{0,m})]. \end{aligned}$$

(2) Ist $f$ in $x_0$ differenzierbar, so ist $f$ auch in $x_0$ partiell differenzierbar.
(3) Die Funktion $f:\mathbb{R}^2\to\mathbb{R}$ mit

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

ist in $(0,0)$ partiell differenzierbar, aber nicht stetig.

Beweis. ad (1): Folgt aus der Definition der Richtungsableitung.
ad (2): Folgt aus Satz VII.2.7.
ad (3): Wegen $f(x,x)=\frac{1}{4x^2}$ ist $f$ in $(0,0)$ nicht stetig. Wegen $f(t,0)=f(0,t)=0$ für alle $t\in\mathbb{R}$ ist $f$ in $(0,0)$ partiell differenzierbar mit $D_1f(0,0)=D_2f(0,0)=0$. □

Satz VII.2.11. (1) *Sei $X=\mathbb{R}^m$, $Y$ beliebig, $f:U\to Y, x_0\in U$ und $f$ in $x_0$ differenzierbar. Dann gilt für jedes $h=(h_1,\dots,h_m)\in\mathbb{R}^m$*

$$Df(x_0)h = \sum_{j=1}^{m} D_jf(x_0)h_j.$$

(2) *Sei $Y=\mathbb{R}^n$, $X$ beliebig, $f=(f_1,\dots,f_n):U\to\mathbb{R}^n, x_0\in U$. Dann ist $f$ genau dann in $x_0$ differenzierbar, wenn alle Komponentenfunktionen $f_1,\dots,f_n$ in $x_0$ differenzierbar sind. In diesem Fall gilt für alle $h\in X$*

$$Df(x_0)h = (Df_1(x_0)h,\dots,Df_n(x_0)h).$$

Beweis. ad (1): Sei $h=(h_1,\dots,h_m)\in\mathbb{R}^m$. Dann folgt aus der Linearität der Ableitung

$$\begin{aligned} h &= \sum_{j=1}^{m} h_je_j \\ \Longrightarrow\quad Df(x_0)h &= \sum_{j=1}^{m} h_jDf(x_0)e_j \\ &= \sum_{j=1}^{m} D_jf(x_0)h_j. \end{aligned}$$

AD (2): Es ist $A \in \mathcal{L}(X, \mathbb{R}^n)$ genau dann, wenn $A = (A_1, \ldots, A_n)$ ist mit $A_i \in \mathcal{L}(X, \mathbb{R}), 1 \le i \le n$. Gemäß Satz III.3.10 (S. 78) gilt

$$\lim_{h \to 0} \frac{1}{\|h\|_X}[f(x_0 + h) - f(x_0) - Df(x_0)h] = 0$$
$$\iff \lim_{h \to 0} \frac{1}{\|h\|_X}[f_i(x_0 + h) - f_i(x_0) - Df_i(x_0)h] = 0 \quad \forall 1 \le i \le n.$$

Aus Satz VII.2.2 (3) folgt daher

$$Df(x_0)h = (Df_1(x_0)h, \ldots, Df_n(x_0)h).$$

□

BEMERKUNG VII.2.12. Sei $X = \mathbb{R}^m$, $Y = \mathbb{R}^n$ und $f = (f_1, \ldots, f_n) : U \to Y$ in $x_0$ differenzierbar. Wegen Satz VII.2.11 und Bemerkung VII.2.10 sind dann die Komponentenfunktionen $f_1, \ldots, f_n$ in $x_0$ partiell differenzierbar. Für $h = (h_1, \ldots, h_m) \in \mathbb{R}^m$ gilt

$$\begin{aligned} Df(x_0)h &= \sum_{i=1}^{n} Df_i(x_0)he_i \\ &= \sum_{i=1}^{n} e_i \left\{ \sum_{j=1}^{m} D_j f_i(x_0)h_j \right\}, \end{aligned}$$

d.h., $Df(x_0) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ wird durch die Matrix

$$\begin{pmatrix} D_1f_1(x_0) & D_2f_1(x_0) & \ldots & D_mf_1(x_0) \\ D_1f_2(x_0) & D_2f_2(x_0) & \ldots & D_mf_2(x_0) \\ \vdots & & & \\ D_1f_n(x_0) & D_2f_n(x_0) & \ldots & D_mf_n(x_0) \end{pmatrix} \in M_{n,m}(\mathbb{R})$$

dargestellt. Diese heißt JACOBIMATRIX oder FUNKTIONALMATRIX von $f$ in $x_0$.

Aus Bemerkung VII.2.10 und Satz VII.2.2 folgt, dass eine differenzierbare Funktion stets partiell differenzierbar ist, und dass die Umkehrung i. a. nicht gilt. Der folgende Satz zeigt, dass die Umkehrung gilt, wenn die partiellen Ableitungen zusätzlich stetig sind.

SATZ VII.2.13 (DIFFERENZIERBARKEITSKRITERIUM). *Sei $X = \mathbb{R}^m$ und $Y$ beliebig. Dann ist $f \in C^1(U, Y)$ genau dann, wenn für alle $x \in U$ die Funktion $f$ in $x$ partiell differenzierbar ist und wenn gilt*

$$D_j f \in C(U, Y) \quad \forall 1 \le j \le m.$$

BEWEIS. „$\Longrightarrow$“ Folgt aus Bemerkung VII.2.10 und Satz VII.2.11. „$\Longleftarrow$“ Für $x \in U$ definieren wir $A_x \in \mathcal{L}(X, Y)$ durch

$$A_x h = \sum_{j=1}^{m} D_j f(x)h_j \quad \forall h = (h_1, \ldots, h_m) \in \mathbb{R}^m.$$

Da die Zuordnung $x \mapsto A_x$ wegen $D_j f \in C(U, Y)$ stetig ist, müssen wir noch zeigen

$$\lim_{h \to 0} \frac{1}{\|h\|_1}[f(x+h) - f(x) - A_x h] = 0 \quad \forall x \in U,$$

wobei $\|\cdot\|_1$ die $\ell^1$-Norm auf $\mathbb{K}^n$ ist.
Sei $x \in U$ und $\varepsilon > 0$. Dann gibt es ein $\delta > 0$ mit $x + h \in U$ für alle $h \in \mathbb{R}^m$ mit $\|h\|_1 < \delta$ und

$$\|D_j f(x+h) - D_j f(x)\|_Y < \varepsilon \quad \forall 1 \le j \le m,\ h \in \mathbb{R}^m,\ \|h\|_1 < \delta.$$

Sei $h \in \mathbb{R}^m$ mit $\|h\|_1 < \delta$. Dann folgt mit $e_0 = 0$, $h_0 = 0$

$$\begin{aligned}
&f(x+h) - f(x) - A_x h \\
=&\sum_{j=1}^m [f(x + h_1 e_1 + \ldots + h_j e_j) \\
&\qquad - f(x + h_0 e_0 + \ldots + h_{j-1} e_{j-1})] \\
&- \sum_{j=1}^m D_j f(x) h_j \\
=&\sum_{j=1}^m \int_0^1 \frac{d}{dt} f(x + h_0 e_0 + \ldots + h_{j-1} e_{j-1} + t h_j e_j) dt \\
&- \sum_{j=1}^m D_j f(x) h_j \\
=&\sum_{j=1}^m \int_0^1 [D_j f(x + h_0 e_0 + \ldots + h_{j-1} e_{j-1} + t h_j e_j) h_j \\
&\qquad\qquad - D_j f(x) h_j] dt
\end{aligned}$$

und somit

$$\begin{aligned}
&\| \frac{1}{\|h\|_1}[f(x+h) - f(x) - A_x h] \|_Y \\
\le& \frac{1}{\|h\|_1} \sum_{j=1}^m \int_0^1 \|D_j f \underbrace{(x + h_0 e_0 + \ldots + h_{j-1} e_{j-1} + t h_j e_j)}_{= x + \tilde{h}, \|\tilde{h}\|_1 < \delta} h_j \\
&\qquad\qquad - D_j f(x) h_j \|_Y dt \\
\le& \frac{1}{\|h\|_1} \sum_{j=1}^m \int_0^1 \varepsilon |h_j| dt
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\|h\|_1}\varepsilon\|h\|_1 \\
&= \varepsilon.
\end{aligned}
$$

Da $x \in U$ und $\varepsilon > 0$ beliebig waren, folgt hieraus die Behauptung. □

Bemerkung VII.2.14. Sei $X = \mathbb{R}^m$ und $Y = \mathbb{R}^n$. Dann ist wegen Satz VII.2.13 und Satz VII.2.11 $f \in C^1(U, Y)$ genau dann, wenn für jedes $1 \le j \le m$ und jedes $1 \le i \le n$ die partielle Ableitung $D_j f_i \in C(U, \mathbb{R})$ ist.

Beispiel VII.2.15. Sei $f : \mathbb{R}^3 \to \mathbb{R}^2$ definiert durch

$$f(x, y, z) = (e^x \cos y, \sin(xz)).$$

Dann ist

$$
\begin{aligned}
D_1 f_1(x, y, z) &= e^x \cos y \\
D_2 f_1(x, y, z) &= -e^x \sin y \\
D_3 f_1(x, y, z) &= 0 \\
D_1 f_2(x, y, z) &= z \cos(xz) \\
D_2 f_2(x, y, z) &= 0 \\
D_3 f_2(x, y, z) &= x \cos(xz).
\end{aligned}
$$

Also ist $f \in C^1(\mathbb{R}^3, \mathbb{R}^2)$.

Bemerkung VII.2.16. Sei $X = \mathbb{R}^m$, $Y = \mathbb{R}$ und $f : U \to \mathbb{R}$ in $x_0 \in U$ differenzierbar. Dann ist für jedes $h = (h_1, \ldots, h_m) \in \mathbb{R}^m$

$$
\begin{aligned}
Df(x_0)h &= \sum_{j=1}^{m} D_j f(x_0) h_j \\
&= (\nabla f(x_0), h),
\end{aligned}
$$

wobei

$$(x, y) = \sum_{j=1}^{m} x_j y_j$$

das euklidische Skalarprodukt auf $\mathbb{R}^m$ und

$$\nabla f(x_0) = (D_1 f(x_0), \ldots, D_m f(x_0))$$

den Gradienten von $f$ bezeichnet.
Sei $\nabla f(x_0) \neq 0$. Dann ist $v = \frac{1}{\|\nabla f(x_0)\|_2} \nabla f(x_0) \in \mathbb{R}^m$ mit $\|v\|_2 = 1$. Für jedes $h \in \mathbb{R}^m$ mit $\|h\|_2 \le 1$ gilt wegen der Cauchy-Schwarzschen Ungleichung

$$|Df(x_0)h| \le \|\nabla f(x_0)\|_2 \|h\|_2 \le \|\nabla f(x_0)\|_2$$

und speziell für $h = v$

$$Df(x_0)v = \frac{1}{\|\nabla f(x_0)\|_2}(\nabla f(x_0), \nabla f(x_0)) = \|\nabla f(x_0)\|_2.$$

Also gilt

$$Df(x_0)v = \sup_{\|h\|_2=1} |Df(x_0)h|,$$

d.h., die Richtungsableitung ist in Richtung des Gradienten maximal, oder anders ausgedrückt:

> Der Gradient gibt die Richtung des steilsten Anstieges von $f$ an.

Zum Abschluss dieses Paragraphen untersuchen wir den Zusammenhang zwischen der Differenzierbarkeit von Funktionen $f : \mathbb{C} \to \mathbb{C}$ im Sinne von Definition IV.1.1 (S. 111) und der ihnen kanonisch entsprechenden Funktionen $F : \mathbb{R}^2 \to \mathbb{R}^2$ im Sinne von Definition VII.2.1.

Satz VII.2.17. *Sei $U \subset \mathbb{C}$ und $W = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + ix_2 \in U\}$. Sei $f : U \to \mathbb{C}$ eine Funktion und $F : W \to \mathbb{R}^2$ definiert durch*

$$F(x_1, x_2) = (\operatorname{Re} f(x_1 + ix_2), \operatorname{Im} f(x_1 + ix_2)).$$

*Dann sind folgende Aussagen äquivalent:*

(1) *$f$ ist in $z_0 = x_1 + ix_2 \in U$ komplex differenzierbar im Sinne von Definition IV.1.1 (S. 111).*

(2) *$F$ ist in $x_0 = (x_1, x_2) \in W$ differenzierbar im Sinne von Definition VII.2.1 und die partiellen Ableitungen im Punkt $x_0$ erfüllen die* Cauchy-Riemannschen Differentialgleichungen

$$D_1F_1(x_0) = D_2F_2(x_0)$$
$$D_1F_2(x_0) = -D_2F_1(x_0).$$

Beweis. (1) $\Longrightarrow$ (2): Sei $f'(z_0) = \alpha + i\beta$. Definiere $A \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ durch

$$Ah = (\alpha h_1 - \beta h_2, \beta h_1 + \alpha h_2) \quad \forall h = (h_1, h_2) \in \mathbb{R}^2.$$

Sei $h = (h_1, h_2) \in \mathbb{R}^2$ und $w = h_1 + ih_2$, so dass $z_0 + w \in U$ ist. Dann folgt

$$\begin{aligned} &F(x_0 + h) - F(x_0) - Ah \\ =&\begin{pmatrix} \operatorname{Re} f(z_0 + w) & - & \operatorname{Re} f(z_0) & - & \operatorname{Re}(f'(z_0) \cdot w) \\ \operatorname{Im} f(z_0 + w) & - & \operatorname{Im} f(z_0) & - & \operatorname{Im}(f'(z_0) \cdot w) \end{pmatrix} \end{aligned}$$

und

$$\begin{aligned} &\frac{\|F(x_0 + h) - F(x_0) - Ah\|_2}{\|h\|_2} \\ =&\frac{|f(z_0 + w) - f(z_0) - f'(z_0)w|}{|w|} \end{aligned}$$

$$\underset{\substack{w\to 0\\ \iff h\to 0}}{\longrightarrow} 0.$$

Hieraus folgt die Behauptung.
(2) $\implies$ (1): Setze zur Abkürzung

$$\alpha = D_1F_1(x_0) \quad \beta = D_2F_1(x_0)$$
$$\gamma = D_1F_2(x_0) \quad \delta = D_2F_2(x_0).$$

Definiere die Abbildung $B : \mathbb{C} \to \mathbb{C}$ durch

$$w = h_1 + ih_2 \mapsto \alpha h_1 + \beta h_2 + i(\gamma h_1 + \delta h_2).$$

Wie im Beweis von „(1) $\implies$ (2)“ folgt

$$\frac{|f(z_0 + w) - f(z_0) - B(w)|}{|w|}$$
$$= \frac{\|F(x_0 + h) - F(x_0) - DF(x_0)h\|_2}{\|h\|_2}$$
$$\underset{\substack{h\to 0\\ \iff w\to 0}}{\longrightarrow} 0.$$

Also müssen wir noch zeigen, dass $B$ $\mathbb{C}$-linear ist. Da $B$ $\mathbb{R}$-linear ist, d.h.,

$$B(\alpha w) = \alpha B(w) \quad \forall \alpha \in \mathbb{R}, w \in \mathbb{C},$$

ist dies genau dann der Fall, wenn gilt

$$B(iw) = iB(w) \quad \forall w \in \mathbb{C}.$$

Wegen

$$B(iw) = -\alpha h_2 + \beta h_1 + i(-\gamma h_2 + \delta h_1)$$
$$iB(w) = -\gamma h_1 - \delta h_2 + i(\alpha h_1 + \beta h_2) \quad \forall w = h_1 + ih_2$$

ist dies äquivalent zu

$$\alpha = \delta, \quad \beta = -\gamma,$$

d.h., zu den Cauchy-Riemannschen Differentialgleichungen. □

## VII.3. Eigenschaften differenzierbarer Funktionen

Im Folgenden seien $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$ $\mathbb{K}$-Banachräume und $U \subset X$, $U \neq \emptyset$, offen.

Satz VII.3.1. *Seien $f, g : U \to Y$ in $x_0 \in U$ differenzierbar und $\alpha, \beta \in \mathbb{K}$. Dann ist $\alpha f + \beta g : U \to Y$ in $x_0$ differenzierbar und*

$$D(\alpha f + \beta g)(x_0) = \alpha Df(x_0) + \beta Dg(x_0).$$

Beweis. Es ist

$$f(x) = f(x_0) + Df(x_0)(x - x_0) + r(x)\|x - x_0\|_X$$
$$g(x) = g(x_0) + Dg(x_0)(x - x_0) + s(x)\|x - x_0\|_X$$

mit $r, s$ stetig in $x_0$ und $r(x_0) = s(x_0) = 0$. Dann ist

$$\begin{aligned}\alpha f(x) + \beta g(x) = \alpha f(x_0) &+ \beta g(x_0)\\ &+ [\alpha Df(x_0) + \beta Dg(x_0)](x - x_0)\\ &+ [\alpha r(x) + \beta s(x)]\|x - x_0\|_X.\end{aligned}$$

Da $\alpha Df(x_0) + \beta Dg(x_0) \in \mathcal{L}(X, Y)$ und $\alpha r + \beta s$ in $x_0$ stetig ist mit $\alpha r(x_0) + \beta s(x_0) = 0$ folgt hieraus die Behauptung. □

Bemerkung VII.3.2. $C^1(U, Y)$ ist ein Untervektorraum von $C(U, Y)$ und die Abbildung $D : C^1(U, Y) \to C(U, \mathcal{L}(X, Y))$ mit $f \mapsto Df$ ist linear. Ist speziell $X = \mathbb{K}$, so ist $\mathcal{L}(X, Y) \cong Y$ und $D : C^1(U, Y) \to C(U, Y)$ ist linear (vgl. Satz IV.1.7 (S. 114)).

Satz VII.3.3 (Kettenregel). *Sei $f : U \to Y$ und $g : V \to Z$ mit $V \subset Y$, $V \neq \emptyset$, offen. Es gelte:*

1. *$f(U) \subset V$,*
2. *$f$ ist differenzierbar in $x_0$,*
3. *$g$ ist differenzierbar in $y_0 = f(x_0)$.*

*Dann ist $g \circ f : U \to Z$ in $x_0$ differenzierbar und*

$$(*) \qquad D(g \circ f)(x_0) = Dg(f(x_0)) \cdot Df(x_0).$$

*Dabei bezeichnet $\cdot$ die Verknüpfung von linearen Abbildungen in $\mathcal{L}(Y, Z)$ und $\mathcal{L}(X, Y)$.*

Beweis. Es ist

$$f(x) = f(x_0) + Df(x_0)(x - x_0) + r(x)\|x - x_0\|_X$$

mit $r : U \to Y$ stetig in $x_0$ und $r(x_0) = 0$ und

$$g(y) = g(y_0) + Dg(y_0)(y - y_0) + s(y)\|y - y_0\|_Y$$

mit $s : V \to Z$ stetig in $y_0$ und $s(y_0) = 0$. Damit folgt

$$\begin{aligned}g \circ f(x) &= g(f(x))\\ &= g(f(x_0)) + Dg(f(x_0))(f(x) - f(x_0))\\ &\quad + s(f(x))\|f(x) - f(x_0)\|_Y\\ &= g(f(x_0)) + Dg(f(x_0))(Df(x_0)(x - x_0))\\ &\quad + Dg(f(x_0))(r(x)\|x - x_0\|_X)\\ &\quad + s(f(x))\Big\|Df(x_0)(x - x_0) + r(x)\|x - x_0\|_X\Big\|_Y\\ &= g(f(x_0)) + Dg(f(x_0))(Df(x_0)(x - x_0))\\ &\quad + z(x)\|x - x_0\|_X\end{aligned}$$

mit $z(x_0) = 0$ und

$$z(x) = Dg(f(x_0))r(x)$$

$$+ s(f(x)) \left\| \frac{1}{\|x - x_0\|_X} Df(x_0)(x - x_0) + r(x) \right\|_Y$$

für $x \neq x_0$. Aus den Voraussetzungen folgt, dass $z$ in $x_0$ stetig ist. Hieraus folgt die Behauptung. □

Bemerkung VII.3.4. Sei $X = \mathbb{R}^l$, $Y = \mathbb{R}^m$, $Z = \mathbb{R}^n$. Ansonsten seien die Bezeichnungen und Voraussetzungen wie in Satz VII.3.3. Sei $h = g \circ f$. Da die Hintereinanderschaltung linearer Abbildungen aus $\mathcal{L}(\mathbb{R}^l, \mathbb{R}^m)$ und $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ der Multiplikation der entsprechenden Matrizen aus $M_{m,l}(\mathbb{R})$ bzw. $M_{n,m}(\mathbb{R})$ entspricht, erhalten wir folgende Koordinatendarstellung der Kettenregel

$$D_i h_k(x_0) = \sum_{j=1}^{m} D_j g_k(f(x_0)) D_i f_j(x_0) \quad 1 \leq i \leq l,\ 1 \leq k \leq n.$$

Beispiel VII.3.5. Betrachte

$$\begin{aligned} f : \mathbb{R}^2 \to \mathbb{R}^3 &\quad f(x,y) = (x^2, xy, xy^2) \\ g : \mathbb{R}^3 \to \mathbb{R}^2 &\quad g(u,v,w) = (\sin u, \cos(uvw)) \\ h = g \circ f : \mathbb{R}^2 \to \mathbb{R}^2 &\quad h(x,y) = (\sin(x^2), \cos(x^4 y^3)). \end{aligned}$$

Dann ist

$$Dh(x,y) = \begin{pmatrix} 2x\cos(x^2) & 0 \\ -4x^3y^3\sin(x^4y^3) & -3y^2x^4\sin(x^4y^3) \end{pmatrix}$$

$$Df(x,y) = \begin{pmatrix} 2x & 0 \\ y & x \\ y^2 & 2xy \end{pmatrix}$$

$$Dg(u,v,w) = \begin{pmatrix} \cos u & 0 & 0 \\ -vw\sin(uvw) & -uw\sin(uvw) & -uv\sin(uvw) \end{pmatrix}$$

$$\begin{aligned} Dg(f(x)) \cdot Df(x) &= \begin{pmatrix} \cos(x^2) & 0 & 0 \\ -x^2y^3\sin(x^4y^3) & -x^3y^2\sin(x^4y^3) & -x^3y\sin(x^4y^3) \end{pmatrix} \cdot \\ &\quad \cdot \begin{pmatrix} 2x & 0 \\ y & x \\ y^2 & 2xy \end{pmatrix} \\ &= \begin{pmatrix} 2x\cos(x^2) & 0 \\ -4x^3y^3\sin(x^4y^3) & -3x^4y^2\sin(x^4y^3) \end{pmatrix}. \end{aligned}$$

Satz VII.3.6 (Produktregel). *Seien $f, g : U \to \mathbb{R}$ in $x_0 \in U$ differenzierbar. Dann ist $f \cdot g : U \to \mathbb{R}$ in $x_0$ differenzierbar und*

$$D(f \cdot g)(x_0) = g(x_0)Df(x_0) + f(x_0)Dg(x_0).$$

*Ist insbesondere $X = \mathbb{R}^m$, so ist*

$$\nabla(f \cdot g)(x_0) = g(x_0)\nabla f(x_0) + f(x_0)\nabla g(x_0).$$

Beweis. Definiere $F : U \to \mathbb{R}^2$ durch

$$F(x) = (f(x), g(x))$$

und $\varphi : \mathbb{R}^2 \to \mathbb{R}$ durch

$$\varphi(x_1, x_2) = x_1 \cdot x_2.$$

Dann sind $F$ in $x_0$ und $\varphi$ in ganz $\mathbb{R}^2$ differenzierbar und für $v \in X\backslash\{0\}$ und $h = (h_1, h_2) \in \mathbb{R}^2\backslash\{0\}$ gilt

$$\begin{aligned} DF(x_0)v &= (Df(x_0)v, Dg(x_0)v) \\ D\varphi(x_1, x_2)h &= x_2 h_1 + x_1 h_2. \end{aligned}$$

Also ist gemäß Satz VII.3.3 $f \cdot g = \varphi \circ F$ in $x_0$ differenzierbar und

$$\begin{aligned} D(f \cdot g)(x_0)v &= D(\varphi \circ F)(x_0)v \\ &= D\varphi(F(x_0))(DF(x_0)v) \\ &= D\varphi(F(x_0))h \quad \text{mit } h = (Df(x_0)v, Dg(x_0)v) \\ &= g(x_0)Df(x_0)v + f(x_0)Dg(x_0)v \\ &= [g(x_0)Df(x_0) + f(x_0)Dg(x_0)]v. \end{aligned}$$

Hieraus folgt die Behauptung wegen

$$f(x_0)Dg(x_0) + g(x_0)Df(x_0) \in \mathcal{L}(X, \mathbb{R}).$$

□

Satz VII.3.7 (Mittelwertsatz). *Seien $x, y \in U$, $x \neq y$, mit $\{x + t(y - x) : 0 \leq t \leq 1\} \subset U$. Dann gilt:*

(1) *Ist $f : U \to Y$ in $U$ differenzierbar, so ist*

$$\|f(x) - f(y)\|_Y \leq \sup_{0 \leq t \leq 1} \|Df(x + t(y - x))\|_{\mathcal{L}(X,Y)} \cdot \|x - y\|_X.$$

(2) *Ist $f \in C^1(U, Y)$, so ist*

$$f(y) - f(x) = \int_0^1 Df(x + t(y - x))(y - x)dt.$$

Beweis. ad (1): Definiere $\varphi : [0, 1] \to Y$ durch

$$\varphi(t) = f(x + t(y - x)).$$

Dann ist $\varphi$ auf $[0, 1]$ differenzierbar und

$$\varphi'(t) = Df(x + t(y - x))(y - x) \quad \forall t \in [0, 1].$$

Aus Satz IV.2.17 (S. 129) und Bemerkung IV.2.18 (S. 130) folgt

$$\begin{aligned} \|f(y) - f(x)\|_Y &= \|\varphi(1) - \varphi(0)\|_Y \\ &\leq \sup_{0 \leq t \leq 1} \|\varphi'(t)\|_Y \\ &= \sup_{0 \leq t \leq 1} \|Df(x + t(y - x))(y - x)\|_Y \\ &\leq \sup_{0 \leq t \leq 1} \|Df(x + t(y - x))\|_{\mathcal{L}(X,Y)} \|x - y\|_X. \end{aligned}$$

AD (2): Sei $\varphi$ wie im Beweis von (1). Da $\varphi' \in C([0,1], Y)$ ist, folgt aus Satz VI.2.12 (S. 184)

$$\begin{aligned} f(y) - f(x) &= \varphi(1) - \varphi(0) \\ &= \int_0^1 \varphi'(t)dt \\ &= \int_0^1 Df(x + t(y-x))(y-x)dt. \end{aligned}$$

□

SATZ VII.3.8. (1) *Sei $U$ konvex und $f : U \to Y$ differenzierbar mit*

$$\|Df(x)\|_{\mathcal{L}(X,Y)} \le \alpha \quad \forall x \in U.$$

*Dann gilt*

$$\|f(y) - f(x)\|_Y \le \alpha \|y - x\|_X \quad \forall x, y \in U.$$

(2) *Sei $U$ ein Gebiet und $f \in C^1(U, Y)$ mit $Df(x) = 0$ für alle $x \in U$. Dann ist $f$ konstant.*

BEWEIS. AD (1): Folgt direkt aus Satz VII.3.7.
AD (2): Sei $y \in f(U)$ beliebig und $V = f^{-1}(\{y\})$. Dann ist $V \neq \emptyset$ und abgeschlossen. Sei $x \in V$. Dann gibt es ein $\delta > 0$ mit $B(x, \delta) \subset U$. Da $B(x, \delta)$ konvex ist, folgt aus Teil (1)

$$f(z) = f(x) \quad \forall z \in B(x, \delta),$$

d.h.

$$B(x, \delta) \subset V.$$

Also ist $V$ auch offen. Da $U$ zusammenhängend ist, folgt $V = U$. Damit ist die Behauptung bewiesen. □

SATZ VII.3.9 (GLIEDWEISE DIFFERENTIATION VON FUNKTIONENFOLGEN). *Sei $(f_n)_{n\in\mathbb{N}} \subset C^1(U, Y)$ und $f : U \to Y$. Es gelte*

(1) $f_n \xrightarrow[n\to\infty]{pkt} f$,

(2) $\exists g \in C(U, \mathcal{L}(X, Y)) : Df_n \xrightarrow[n\to\infty]{lok.\ glm} g$.

*Dann ist $f \in C^1(U, Y)$ und $g = Df$, d.h.,*

$$\lim_{n\to\infty} Df_n = Df.$$

BEWEIS. Wörtlich der selbe wie der von Satz V.2.10 (S. 154). □

SATZ VII.3.10 (NOTWENDIGE BEDINGUNG FÜR LOKALE EXTREMA). *Es sei $f : U \to \mathbb{R}$ und $x_0 \in U$. $f$ habe in $x_0$ ein lokales Extremum und alle Richtungsableitungen von $f$ mögen in $x_0$ existieren. Dann gilt*

$$D_v f(x_0) = 0 \quad \forall v \in X \backslash \{0\}.$$

BEWEIS. Es gibt ein $\varepsilon > 0$ mit $B(x_0, \varepsilon) \subset U$. Sei $v \in X \backslash \{0\}$ beliebig. Dann ist $\varphi : (-\varepsilon/\|v\|_X, \varepsilon/\|v\|_X) \to \mathbb{R}$ definiert durch

$$\varphi(t) = f(x_0 + tv)$$

in 0 differenzierbar mit

$$\varphi'(0) = D_v f(x_0)$$

und hat in 0 ein lokales Extremum. Damit folgt die Behauptung aus Satz IV.2.2 (S. 121). □

Satz VII.3.10 legt folgende Definition nahe.

DEFINITION VII.3.11. Sei $f : U \to \mathbb{R}$ in $x_0 \in U$ differenzierbar. Dann heißt $x_0$ ein KRITISCHER PUNKT von $f$, wenn gilt $Df(x_0) = 0$.

BEMERKUNG VII.3.12. (1) Ist $f : U \to \mathbb{R}$ in $x_0 \in U$ differenzierbar und $x_0$ ein lokales Extremum, so ist $x_0$ ein kritischer Punkt von $f$.
(2) Ist $X = \mathbb{R}^m$ und $f$ in $x_0 \in U$ differenzierbar, so ist $x_0$ genau dann ein kritischer Punkt von $f$, wenn $\nabla f(x_0) = 0$ ist.
(3) Die Umkehrung von (1) ist falsch. Betrachte z.B. $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = x^2 - y^2$. Dann ist $\nabla f(0, 0) = 0$, d.h., $(0, 0)$ ist ein kritischer Punkt von $f$, aber $(0, 0)$ ist offensichtlich kein lokales Extremum von $f$.

## VII.4. Höhere Ableitungen

Im Folgenden seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ $\mathbb{K}$-Banachräume und $U \subset X$, $U \neq \emptyset$, offen.

DEFINITION VII.4.1. (1) Die Funktion $f : U \to X$ sei in einer Umgebung $V \subset U$ von $x_0 \in U$ differenzierbar und die Funktion $Df : V \to \mathcal{L}(X, Y)$ sei in $x_0$ differenzierbar. Dann ist die ZWEITE ABLEITUNG $D^2 f(x_0) \in \mathcal{L}(X, \mathcal{L}(X, Y))$ von $f$ in $x_0$ definiert durch

$$D^2 f(x_0) = D(Df)(x_0).$$

In diesem Fall heißt $f$ in $x_0$ ZWEIMAL DIFFERENZIERBAR.
(2) Die Funktion $f : U \to X$ heißt ZWEIMAL DIFFERENZIERBAR in $U$, wenn sie in jedem Punkt $x \in U$ zweimal differenzierbar ist.
(3) Sei $m \in \mathbb{N}^*$. Die Funktion $f : U \to X$ sei in einer Umgebung $V \subset U$ von $x_0 \in U$ $m$-mal differenzierbar und die $m$-te Ableitung

$$D^m f : V \to \underbrace{\mathcal{L}(X, \mathcal{L}(X, \ldots, \mathcal{L}(X, Y)) \ldots)}_{m-\text{mal } \mathcal{L}}$$

sei in $x_0$ differenzierbar. Dann ist die $(m+1)$-TE ABLEITUNG

$$D^{m+1} f(x_0) \in \underbrace{\mathcal{L}(X, \mathcal{L}(X, \ldots, \mathcal{L}(X, Y)) \ldots)}_{(m+1)-\text{mal } \mathcal{L}}$$

von $f$ in $x_0$ definiert durch

$$D^{m+1} f(x_0) = D(D^m f)(x_0).$$

In diesem Fall heißt $f$ in $x_0$ $(m+1)$-MAL DIFFERENZIERBAR.
(4) Sei $m \in \mathbb{N}^*$. Die Funktion $f : U \to X$ heißt $m$-MAL DIFFERENZIERBAR in $U$, wenn sie in jedem Punkt $x \in U$ $m$-mal differenzierbar ist.

BEISPIEL VII.4.2. Sei $A \in M_{n,n}(\mathbb{R})$ eine symmetrische Matrix und $f : \mathbb{R}^n \to \mathbb{R}$ definiert durch

$$f(x) = \frac{1}{2} x^t A x = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} x_i A_{ij} x_j.$$

Eine einfache Rechnung liefert

$$Df(x) = x^t A \in \mathbb{R}^n \cong \mathcal{L}(\mathbb{R}^n, \mathbb{R}).$$

Also ist

$$\begin{aligned} D^2 f(x) = A \in M_{n,n}(\mathbb{R}) &\cong \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \\ &\cong \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R})). \end{aligned}$$

Die Räume $\mathcal{L}(X, \mathcal{L}(X, \ldots, \mathcal{L}(X, Y) \ldots)$ sind für praktische Rechnungen ziemlich unhandlich. Daher betrachten wir nun multilineare Abbildungen.

DEFINITION VII.4.3. Seien $(X_i, \|\cdot\|_{X_i})$, $1 \le i \le m$, normierte $\mathbb{K}$-Vektorräume. Eine Abbildung $\varphi : X_1 \times \ldots \times X_m \to Y$ heißt MULTILINEAR oder $m$-LINEAR, wenn für jeden Punkt $(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_m) \in X_1 \times \ldots \times X_{k-1} \times X_{k+1} \times \ldots \times X_m$ und jedes $k \in \mathbb{N}_m^*$ die Abbildung

$$\varphi(x_1, \ldots, x_{k-1}, \cdot, x_{k+1}, \ldots, x_m) : X_k \to Y$$

linear ist.

SATZ VII.4.4. *Sei $\varphi : X_1 \times \ldots \times X_m \to Y$ $m$-linear. Dann sind folgende Aussagen äquivalent:*

(1) *$\varphi$ ist stetig.*
(2) *$\exists \alpha \in \mathbb{R}^* \; \forall (x_1, \ldots, x_m) \in X_1 \times \ldots \times X_m$:*

$$\|\varphi(x_1, \ldots, x_m)\|_Y \le \alpha \|x_1\|_{X_1} \ldots \|x_m\|_{X_m}.$$

BEWEIS. (1) $\Longrightarrow$ (2): Durch

$$\|(x_1, \ldots, x_m)\| = \max_{1 \le i \le m} \|x_i\|_{X_i}$$

wird eine Norm auf $X_1 \times \ldots \times X_m$ definiert (Übungsaufgabe Analysis I). Da $\varphi$ insbesondere in $(0, \ldots, 0) \in X_1 \times \ldots \times X_m$ stetig ist, gibt es ein $\delta > 0$ mit

$$\|\varphi(x_1, \ldots, x_m)\|_Y \le 1 \quad \forall \|(x_1, \ldots, x_m)\| \le \delta.$$

Seien $x_i \in X_i \backslash \{0\}$, $1 \le i \le m$. Wegen der Multilinearität von $\varphi$ folgt

$$\|\varphi(x_1, \ldots, x_m)\|_Y$$

$$
\begin{aligned}
&= \left\| \varphi\left(\frac{\|x_1\|_{X_1}}{\delta}\delta\frac{x_1}{\|x_1\|_{X_1}}, \ldots, \frac{\|x_m\|_{X_m}}{\delta}\delta\frac{x_m}{\|x_m\|_{X_m}}\right)\right\|_Y \\
&= \delta^{-m}\|x_1\|_{X_1}\ldots\|x_m\|_{X_m}\left\|\varphi\left(\delta\frac{x_1}{\|x_1\|_{X_1}}, \ldots, \delta\frac{x_m}{\|x_m\|_{X_m}}\right)\right\|_Y \\
&\leq \delta^{-m}\|x_1\|_{X_1}\ldots\|x_m\|_{X_m}.
\end{aligned}
\tag{VII.4.1}
$$

Da Gleichung (VII.4.1) offensichtlich richtig ist, wenn für ein $i \in \mathbb{N}_m^*$ gilt $x_i = 0$, folgt damit (2).
(2) $\Longrightarrow$ (1): Für $(x_1, \ldots, x_m), (y_1, \ldots, y_m) \in X_1 \times \ldots \times X_m$ folgt wegen der Multilinearität von $\varphi$

$$
\begin{aligned}
&\|\varphi(x_1, \ldots, x_m) - \varphi(y_1, \ldots, y_m)\|_Y \\
=&\|\varphi(x_1 - y_1, x_2, \ldots, x_m) \\
&+ \varphi(y_1, x_2 - y_2, x_3, \ldots, x_m) + \ldots + \\
&+ \varphi(y_1, y_2, \ldots, y_{m-1}, x_m - y_m)\| \\
\leq&\alpha\{\|x_1 - y_1\|_{X_1}\|x_2\|_{X_2}\ldots\|x_m\|_{X_m} \\
&\quad + \|y_1\|_{X_1}\|x_2 - y_2\|_{X_2}\|x_3\|_{X_3}\ldots\|x_m\|_{X_m} \\
&\quad + \ldots + \|y_1\|_{X_1}\ldots\|y_{m-1}\|_{X_{m-1}}\|x_m - y_m\|_{X_m}\} \\
\leq&\alpha\max\{\|(x_1, \ldots, x_m)\|^{m-1}, \|(y_1, \ldots, y_m)\|^{m-1}\}\cdot \\
&\|(x_1 - y_1, \ldots, x_m - y_m)\|.
\end{aligned}
$$

Hieraus folgt (1). □

Definition VII.4.5. (1) $\mathcal{L}(X_1, \ldots, X_m, Y) = \{\varphi : X_1 \times \ldots \times X_m \to Y : \varphi$ ist $m$-linear und stetig$\}$.
(2) $\mathcal{L}^m(X, Y) = \mathcal{L}(\underbrace{X, \ldots, X}_{m-\text{mal}}, Y)$.
(3) Für $\varphi \in \mathcal{L}(X_1, \ldots, X_m, Y)$ sei

$$
\begin{aligned}
&\|\varphi\|_{\mathcal{L}(X_1, \ldots, X_m, Y)} \\
&= \inf\{\alpha \in \mathbb{R}^* : \|\varphi(x_1, \ldots, x_m)\|_Y \leq \alpha\|x_1\|_{X_1} \cdot \ldots \cdot \|x_m\|_{X_m} \\
&\qquad \forall (x_1, \ldots, x_m) \in X_1 \times \ldots \times X_m\}.
\end{aligned}
$$

(4) $\|\cdot\|_{\mathcal{L}^m(X,Y)} = \|\cdot\|_{\mathcal{L}(X,\ldots,X,Y)}$.

Satz VII.4.6. (1) *Sei $\varphi \in \mathcal{L}(X_1, \ldots, X_m, Y)$. Dann ist*

$$
\begin{aligned}
&\|\varphi\|_{\mathcal{L}(X_1, \ldots, X_m, Y)} \\
&= \sup\{\|\varphi(x_1, \ldots, x_m)\|_Y : \|x_j\|_{X_j} \leq 1 \ \forall 1 \leq j \leq m\}
\end{aligned}
$$

*und*

$$
\begin{aligned}
\|\varphi(x_1, \ldots, x_m)\|_Y \leq \|\varphi\|_{\mathcal{L}(X_1, \ldots, X_m, Y)}\|x_1\|_{X_1} \cdot \ldots \cdot \|x_m\|_{X_m} \\
\forall (x_1, \ldots, x_m) \in X_1 \times \ldots \times X_m.
\end{aligned}
$$

(2) $(\mathcal{L}(X_1,\ldots,X_m,Y),\|\cdot\|_{\mathcal{L}(X_1,\ldots,X_m,Y)})$ *ist ein normierter* $\mathbb{K}$-*Vektorraum.*
(3) *Falls* $Y$ *vollständig ist, ist auch* $\mathcal{L}(X_1,\ldots,X_m,Y)$ *vollständig.*

Beweis. Völlig analog zum Beweis von Bemerkung VI.1.6 (S. 173) und Satz VII.1.1 (S. 211). □

Satz VII.4.7. $\mathcal{L}(X_1,\ldots,X_m,Y) \cong \mathcal{L}(X_1,\mathcal{L}(X_2,\ldots,\mathcal{L}(X_m,Y))$.

Beweis. Der Übersichtlichkeit halber führen wir den Beweis nur für $m=2$ aus. Der allgemeine Fall folgt ganz analog.
Sei $T \in \mathcal{L}(X_1,\mathcal{L}(X_2,Y))$. Definiere $\varphi_T : X_1 \times X_2 \to Y$ durch

$$\varphi_T(x_1,x_2) = (Tx_1)x_2.$$

Offensichtlich ist $\varphi_T$ bilinear. Weiter gilt

$$\begin{aligned}\|\varphi_T(x_1,x_2)\|_Y &\le \|Tx_1\|_{\mathcal{L}(X_2,Y)}\|x_2\|_{X_2}\\ &\le \|T\|_{\mathcal{L}(X_1,\mathcal{L}(X_2,Y))}\|x_1\|_{X_1}\|x_2\|_{X_2}.\end{aligned}$$

Also ist $\varphi_T \in \mathcal{L}(X_1,X_2,Y)$ und

$$\|\varphi_T\|_{\mathcal{L}(X_1,X_2,Y)} \le \|T\|_{\mathcal{L}(X_1,\mathcal{L}(X_2,Y))}.$$

Sei nun $\psi \in \mathcal{L}(X_1,X_2,Y)$. Definiere $T_\psi$ durch

$$(T_\psi x_1)(x_2) = \psi(x_1,x_2).$$

Dann folgt

$$\|(T_\psi x_1)(x_2)\|_Y \le \|\psi\|_{\mathcal{L}(X_1,X_2,Y)}\|x_1\|_{X_1}\|x_2\|_{X_2}.$$

Hieraus folgt $T_\psi \in \mathcal{L}(X_1,\mathcal{L}(X_2,Y))$ und

$$\|T_\psi\|_{\mathcal{L}(X_1,\mathcal{L}(X_2,Y))} \le \|\psi\|_{\mathcal{L}(X_1,X_2,Y)}.$$

Damit ist die Behauptung bewiesen. □

Mit diesen Vorbereitungen erhalten wir eine für die Praxis handlichere Interpretation der höheren Ableitungen.

Definition VII.4.8. (1) Die Funktion $f : U \to Y$ sei in $x_0 \in U$ $m$-mal differenzierbar. Dann identifizieren wir vermöge Satz VII.4.7 die $m$-te Ableitung $D^m f(x_0)$ von $f$ in $x_0$ mit einem Element von $\mathcal{L}^m(X,Y)$.
(2) Die Funktion $f : U \to Y$ heißt in $U$ $m$-mal stetig differenzierbar, wenn sie in $U$ $m$-mal differenzierbar ist und wenn die Funktion $D^m f : U \to \mathcal{L}^m(X,Y)$ stetig ist.
(3) $C^m(U,Y)$ bezeichnet den Vektorraum der $m$-mal stetig differenzierbaren Funktionen von $U$ in $Y$.
(4) $C^\infty(U,Y) = \bigcap\limits_{m\in\mathbb{N}} C^m(U,Y)$.

Bemerkung VII.4.9. Es gilt

$$C^\infty(U,Y) \underset{\neq}{\subset} \ldots \underset{\neq}{\subset} C^{m+1}(U,Y) \underset{\neq}{\subset} C^m(U,Y) \underset{\neq}{\subset} \ldots \underset{\neq}{\subset} C(U,Y).$$

Beweis. Folgt aus Satz VII.2.2 (S. 216) und Bemerkung IV.1.15 (S. 117), die ein Spezialfall von Bemerkung VII.4.9 ist. □

Definition VII.4.10. Sei $\varphi : X\times\ldots\times X \longrightarrow Y$ $m$-linear. Dann heißt $\varphi$ symmetrisch, wenn für alle $x_1, \ldots, x_m \in X$ und alle Permutationen $\sigma : \mathbb{N}^*_m \longrightarrow \mathbb{N}^*_m$ gilt

$$\varphi(x_1, \ldots, x_m) = \varphi(x_{\sigma(1)}, \ldots, x_{\sigma(m)}).$$

Satz VII.4.11. *Sei $m \geq 2$ und $f \in C^m(U, Y)$. Dann ist $D^m f(x) \in \mathcal{L}^m(X, Y)$ für jedes $x \in U$ symmetrisch.*

Beweis. Wir beweisen die Behauptung durch Induktion über $m$.
„$m = 2$“: Sei $x \in U$ beliebig und $r \in \mathbb{R}^*_+$, so dass $B(x, r) \subset U$ ist. Seien $h, k \in X$ mit

$$\|h\|_X \leq \frac{r}{2}\ ,\ \|k\|_X \leq \frac{r}{2}.$$

Aus Satz VII.3.7 (S. 228) folgt

$$\begin{aligned}
&f(x + h + k) - f(x + k) - f(x + h) + f(x)\\
=&\int_0^1 [Df(x + k + th) - Df(x + th)]hdt\\
=&\int_0^1 \int_0^1 D^2 f(x + sk + th)(k, h)dsdt\\
=&D^2 f(x)(k, h)\\
&+\int_0^1 \int_0^1 [D^2 f(x + sk + th) - D^2 f(x)](k, h)dsdt\\
=&D^2 f(x)(k, h) + \varphi(x, k, h)
\end{aligned}$$

und

$$\begin{aligned}
&f(x + h + k) - f(x + h) - f(x + k) + f(x)\\
=&\int_0^1 [Df(x + h + tk) - Df(x + tk)]kdt\\
=&\int_0^1 \int_0^1 D^2 f(x + sh + tk)(h, k)dsdt\\
=&D^2 f(x)(h, k)\\
&+\int_0^1 \int_0^1 [D^2 f(x + sh + tk) - D^2 f(x)](h, k)dsdt\\
=&D^2 f(x)(h, k) + \psi(x, h, k).
\end{aligned}$$

Also gilt

$$\begin{aligned}
&\|D^2 f(x)(k, h) - D^2 f(x)(h, k)\|_Y\\
=&\|\varphi(x, k, h) - \psi(x, h, k)\|_Y
\end{aligned}$$

$$\begin{aligned}&\leq\|\varphi(x,k,h)\|_Y+\|\psi(x,h,k)\|_Y\\&\leq 2\|h\|_X\|k\|_X\sup_{\substack{0\leq s\leq 1\\0\leq t\leq 1}}\|D^2f(x+sh+tk)-D^2f(x)\|_{\mathcal{L}^2(X,Y)}.\end{aligned}$$

Ersetzen von $h, k$ durch $\varepsilon h, \varepsilon k$ mit $\varepsilon > 0$ liefert

$$\begin{aligned}&\|D^2f(x)(k,h)-D^2f(x)(h,k)\|_Y\\=&\frac{1}{\varepsilon^2}\|D^2f(x)(\varepsilon k,\varepsilon h)-D^2f(x)(\varepsilon h,\varepsilon k)\|_Y\\\leq&2\frac{1}{\varepsilon^2}\|\varepsilon k\|_X\|\varepsilon h\|_X\sup_{\substack{0\leq s\leq 1\\0\leq t\leq 1}}\|D^2f(x+s\varepsilon h+t\varepsilon k)-D^2f(x)\|_{\mathcal{L}^2(X,Y)}\\\leq&2\|h\|_X\|k\|_X\sup_{\substack{0\leq s\leq 1\\0\leq t\leq 1}}\|D^2f(x+s\varepsilon h+t\varepsilon k)-D^2f(x)\|_{\mathcal{L}^2(X,Y)}\\\underset{\varepsilon\to 0}{\longrightarrow}&0\end{aligned}$$

wegen der Stetigkeit von $D^2 f$. Damit ist die Behauptung für $m = 2$ gezeigt.
„$m \to m+1$“: Seien $x \in U$ und $r \in \mathbb{R}_+^*$ wie oben und $h_1, \ldots, h_{m+1} \in X$ mit $\|h_k\|_X \leq \frac{r}{m+1}$, $1 \leq k \leq m+1$. Sei $\sigma \in \mathbb{N}_m^* \to \mathbb{N}_m^*$ eine beliebige Permutation von $\mathbb{N}_m^*$. Dann folgt aus der Induktionsvoraussetzung für jedes $t \in \mathbb{R}_+^*$

$$\begin{aligned}&\frac{1}{t}[D^mf(x+th_1)(h_2,\ldots,h_{m+1})-D^mf(x)(h_2,\ldots,h_{m+1})]\\=&\frac{1}{t}[D^mf(x+th_1)(h_{\sigma(2)},\ldots,h_{\sigma(m+1)})-D^mf(x)(h_{\sigma(2)},\ldots,h_{\sigma(m+1)})].\end{aligned}$$

Grenzübergang $t \to 0$ liefert

$$D^{m+1}f(x)(h_1,h_2,\ldots,h_{m+1})=D^{m+1}f(x)(h_1,h_{\sigma(2)},\ldots,h_{\sigma(m+1)}).$$

Wegen $D^{m+1}f = D^2(D^{m-1}f)$ folgt weiter

$$D^{m+1}f(x)(h_1,h_2,\ldots,h_{m+1})=D^{m+1}f(x)(h_2,h_1,\ldots,h_{m+1}).$$

Da jede Permutation von $\mathbb{N}_{m+1}^*$ dargestellt werden kann als $\sigma \circ \tau \circ \rho$, wobei $\tau$ die Indizes 1 und 2 vertauscht und $\sigma$ und $\rho$ Permutationen von $\mathbb{N}_m^*$ sind, folgt hieraus die Behauptung. □

Satz VII.4.12. *Sei $(Z, \|\cdot\|_Z)$ ein Banachraum und $f \in C^m(U, Y)$, $g \in C^m(V, Z)$ mit $f(U) \subset V$. Dann ist $g \circ f \in C^m(U, Z)$.*

Beweis. Die Behauptung folgt durch Induktion aus der Kettenregel, Satz VII.3.3 (S. 226). □

Die folgende Definition ist eine Verallgemeinerung von Definition IV.3.1 (S. 132).

Definition VII.4.13. Seien $f \in C^m(U, Y)$, $x_0 \in U$ und $h \in X$. Dann heißt

$$T_m(f, x_0)(h) = f(x_0) + \sum_{k=1}^{m} \frac{1}{k!} D^k f(x_0)(\underbrace{h, \ldots, h}_{k-\text{mal}})$$

das $m$-te Taylorpolynom von $f$ mit Entwicklungspunkt $x_0$.

Satz VII.4.14 (Taylorsche Formel). *Seien $f \in C^m(U, Y), x \in U$ und $h \in X$, derart dass $\{x + th : 0 \leq t \leq 1\} \subset U$ ist. Dann gilt:*

(1) $f(x + h) = T_{m-1}(f, x)(h) + R_m(f, x, h)$ *mit*

$$R_m(f, x, h) = \int_0^1 \frac{1}{(m-1)!}(1-t)^{m-1} D^m f(x + th)(\underbrace{h, \ldots, h}_{m-mal})dt.$$

(2) $f(x + h) = T_m(f, x)(h) + \tilde{R}_m(f, x, h)$ *mit*

$$\tilde{R}_m(f, x, h) = \int_0^1 \frac{1}{(m-1)!}(1-t)^{m-1}[D^m f(x + th) - D^m f(x)](\underbrace{h, \ldots, h}_{m-mal})dt.$$

Beweis. ad (1): Wir zeigen die Behauptung durch Induktion über $m$.
„$m = 1$“: Dies ist Satz VII.3.7(2) (S. 228).
„$m \to m + 1$“: Partielle Integration liefert

$$\begin{aligned} R_m(f, x, h) &= \int_0^1 \frac{1}{(m-1)!}(1-t)^{m-1} D^m f(x + th)(h, \ldots, h)dt \\ &= [-\frac{1}{m!}(1-t)^m D^m f(x + th)(h, \ldots, h)]_{t=0}^{t=1} \\ &\quad + \int_0^1 \frac{1}{m!}(1-t)^m D^{m+1} f(x + th)(\underbrace{h, \ldots, h}_{(m+1)-\text{mal}})dt \\ &= \frac{1}{m!} D^m f(x)(\underbrace{h, \ldots, h}_{m-\text{mal}}) + R_{m+1}(f, x, h). \end{aligned}$$

Hieraus folgt die Behauptung.
ad (2): Wegen

$$\int_0^1 \frac{1}{(m-1)!}(1-t)^{m-1}dt = \frac{1}{m!}$$

folgt die Behauptung aus Teil (1). □

Bemerkung VII.4.15. Die Voraussetzungen und Bezeichnungen seien wie in Satz VII.4.14. Dann folgt

$$\|\tilde{R}_m(f, x, h)\|_Y$$

$$\leq \frac{1}{(m-1)!} \|h\|_X^m \sup_{0 \leq t \leq 1} \|D^m f(x+th) - D^m f(x)\|_{\mathcal{L}^m(X,Y)}.$$

Mithin gilt

$$\lim_{\|h\|_X \to 0} \|h\|_X^{-m} \|\tilde{R}_m(f,x,h)\|_Y = 0.$$

Als nächstes wenden wir die bisher erzielten Ergebnisse auf Funktionen auf dem $\mathbb{R}^k$ an.

Definition VII.4.16. Sei $U \subset \mathbb{R}^k$, $U \neq \emptyset$, offen und $f \in C^m(U,Y)$ und $j_1, \ldots, j_m \in \mathbb{N}_k^*$. Dann heißt der Ausdruck

$$D_{j_m} \ldots D_{j_1} f(x) = \frac{\partial}{\partial x_{j_m}} \cdots \frac{\partial}{\partial x_{j_1}} f(x) = \frac{\partial^m}{\partial x_{j_m} \ldots \partial x_{j_1}} f(x)$$

eine partielle Ableitung $m$-ter Ordnung von $f$.

Beispiel VII.4.17. Für

$$f(x,y) = \frac{xy}{1+x^4+y^4}$$

erhalten wir

$$\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{y(1-3x^4+y^4)}{(1+x^4+y^4)^2} \\
\frac{\partial f}{\partial y} &= \frac{x(1+x^4-3y^4)}{(1+x^4+y^4)^2} \\
\frac{\partial^2 f}{\partial x \partial x} &= \frac{x^3 y(-20-20y^4+12x^4)}{(1+x^4+y^4)^3} \\
\frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial^2 f}{\partial y \partial x} \\
&= \frac{1-2x^4-2y^4-3x^8-3y^8+26x^4y^4}{(1+x^4+y^4)^3} \\
\frac{\partial^2 f}{\partial y \partial y} &= \frac{xy^3(-20-20x^4+12y^4)}{(1+x^4+y^4)^3}.
\end{aligned}$$

Satz VII.4.18. *Sei $U \subset \mathbb{R}^k, U \neq \emptyset$ offen.*

(1) *Es ist $f \in C^m(U,Y)$ genau dann, wenn alle partiellen Ableitungen der Ordnung $l$ für alle $1 \leq l \leq m$ existieren und stetig sind.*
(2) *Ist $f \in C^m(U,Y)$, so existieren alle partiellen Ableitungen der Ordnung $\leq m$ und sind von der Reihenfolge der Differentiation unabhängig.*

Beweis. ad (1): „$\Longrightarrow$“ Folgt aus

$$D_{j_l} \ldots D_{j_1} f(x) = D^l f(x)(e_{j_1}, \ldots, e_{j_l}) \tag{VII.4.2}$$

für alle $j_1, \dots, j_l \in \mathbb{N}_k^*$ und $1 \le l \le m$.
“$\Longleftarrow$“ Folgt durch Induktion aus Satz VII.2.13 (S. 221).
AD (2): Folgt aus Gleichung (VII.4.2) und Satz VII.4.11. $\square$

Sei $\underline{\alpha} = (\alpha_1, \dots, \alpha_k) \in \mathbb{N}^k$ ein Multiindex. Dann definieren wir

$$\begin{aligned}
|\underline{\alpha}| &= \sum_{i=1}^{k} \alpha_i \\
\underline{\alpha}! &= \prod_{i=1}^{k} (\alpha_i!) \\
x^{\underline{\alpha}} &= \prod_{i=1}^{k} x_i^{\alpha_i} \quad \forall x = (x_1, \dots, x_k) \in \mathbb{R}^k \\
D^{\underline{\alpha}} f &= D_1^{\alpha_1} \dots D_k^{\alpha_k} f \quad \forall f \in C^{|\underline{\alpha}|}(U, Y), U \in \mathbb{R}^k \text{ offen} \\
D^{\underline{0}} f &= f.
\end{aligned}$$

Sei nun $h = (h_1, \dots, h_k) \in \mathbb{R}^k$. Dann ist für $m \in \mathbb{N}^*$

$$\begin{aligned}
&D^m f(x)(\underbrace{h, \dots, h}_{m-\text{mal}}) \\
&= D^m f(x)\Big(\sum_{i_1=1}^{k} h_{i_1} e_{i_1}, \dots, \sum_{i_m=1}^{k} h_{i_m} e_{i_m}\Big) \\
&= \sum_{i_1=1}^{k} \dots \sum_{i_m=1}^{k} D_{i_1} \dots D_{i_m} f(x) h_{i_1} \dots h_{i_m}. \qquad \text{(VII.4.3)}
\end{aligned}$$

Die Zahl aller $m$-Tupel $(j_1, \dots, j_m)$ mit $j_i \in \mathbb{N}_k^*$ und den Eigenschaften

der Index 1 kommt $\alpha_1$-mal vor,

$\vdots$

der Index $k$ kommt $\alpha_k$-mal vor,

$$\sum_{i=1}^{k} \alpha_i = m$$

beträgt

$$\frac{m!}{\alpha_1! \dots \alpha_k!} = \frac{m!}{\underline{\alpha}!}.$$

Hieraus und aus Gleichung (VII.4.3) folgt

$$\frac{1}{m!} D^m f(x)(\underbrace{h, \dots, h}_{m-\text{ mal}}) = \sum_{|\underline{\alpha}|=m} \frac{1}{\underline{\alpha}!} D^{\underline{\alpha}} f(x) h^{\underline{\alpha}}.$$

Hiermit können wir für Funktionen auf $\mathbb{R}^k$ die Taylorschen Formeln aus Satz VII.4.14 wie folgt schreiben:

Satz VII.4.19 (Taylorsche Formel für Funktionen auf $\mathbb{R}^k$). *Sei $U \subset \mathbb{R}^k$, $U \neq \emptyset$, offen, $f \in C^m(U, Y)$, $x \in U$ und $h \in \mathbb{R}^k$, so, dass $\{x + th : 0 \leq t \leq 1\} \subset U$ ist. Dann gilt*

$$\begin{aligned} f(x+h) &= \sum_{|\underline{\alpha}| \leq m-1} \frac{1}{\underline{\alpha}!} D^{\underline{\alpha}} f(x) h^{\underline{\alpha}} + R_m(f, x, h) \\ &= \sum_{|\underline{\alpha}| \leq m} \frac{1}{\underline{\alpha}!} D^{\underline{\alpha}} f(x) h^{\underline{\alpha}} + \tilde{R}_m(f, x, h) \end{aligned}$$

*mit*

$$R_m(f, x, h) = \int_0^1 \sum_{|\underline{\alpha}| = m} \frac{m}{\underline{\alpha}!} (1-t)^{m-1} D^{\underline{\alpha}} f(x + th) h^{\underline{\alpha}} dt$$

*und*

$$\tilde{R}_m(f, x, h) = \int_0^1 \sum_{|\underline{\alpha}| = m} \frac{m}{\underline{\alpha}!} (1-t)^{m-1} [D^{\underline{\alpha}} f(x + th) - D^{\underline{\alpha}} f(x)] h^{\underline{\alpha}} dt.$$

Als nächstes betrachten wir Funktionen auf $\mathbb{R}^k$ mit Werten in $\mathbb{R}$.

Bemerkung VII.4.20. Sei $U \subset \mathbb{R}^k$, $U \neq \emptyset$, offen und $f \in C^2(U, \mathbb{R})$. Dann gilt für $v = (v_1, \ldots, v_k)$, $w = (w_1, \ldots, w_k)$ aus $\mathbb{R}^k$

$$\begin{aligned} D^2 f(x)(v, w) &= D^2 f(x)\Big(\sum_{i=1}^k v_i e_i, \sum_{j=1}^k w_j e_j\Big) \\ &= \sum_{i=1}^k \sum_{j=1}^k v_i w_j D_i D_j f(x) \\ &= v^t [D_i D_j f(x)]_{1 \leq i,j \leq h} w. \end{aligned}$$

Die Matrix $[D_i D_j f(x)]_{1 \leq i,j \leq k} \in M_{k,k}(\mathbb{R})$ heißt Hesse-Matrix von $f$. Wegen Satz VII.4.18 ist sie symmetrisch.

Definition VII.4.21. Sei $A \in M_{k,k}(\mathbb{R})$ symmetrisch. Dann heißt

(1) $A$ positiv semi-definit
$\iff \quad v^t A v \geq 0 \quad \forall v \in \mathbb{R}^k$,
(2) $A$ positiv definit
$\iff \quad v^t A v > 0 \quad \forall v \in \mathbb{R}^k, v \neq 0$,
(3) $A$ negativ (semi-)definit
$\iff \quad -A$ positiv (semi-) definit,
(4) $A$ indefinit
$\iff \quad \exists v, w \in \mathbb{R}^k \backslash \{0\} : v^t A v > 0,\ w^t A w < 0$.

Aus der linearen Algebra sind folgende Eigenschaften symmetrischer Matrizen bekannt.

Bemerkung VII.4.22. (1) $A \in M_{k,k}(\mathbb{R})$ sei symmetrisch. Dann gibt es Zahlen $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ und Vektoren $x_1, \ldots, x_k \in \mathbb{R}^k$ mit

$$Ax_i = \lambda_i x_i \quad 1 \leq i \leq k$$
$$x_i^t x_j = \delta_{ij} \quad 1 \leq i, j \leq k.$$

Die Zahlen $\lambda_1, \ldots, \lambda_k$ heißen Eigenwerte der Matrix $A$ und die Vektoren $x_1, \ldots, x_k$ die zugehörigen Eigenvektoren. O.E. gilt

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_k.$$

$A$ ist positiv semi-definit $\iff \lambda_1 \geq 0$,
$A$ ist positiv definit $\iff \lambda_1 > 0$,
$A$ ist indefinit $\iff \lambda_1 < 0 < \lambda_k$.
Falls $A$ positiv definit ist, gilt für alle $v \in \mathbb{R}^k$ mit $v \neq 0$

$$v^t A v \geq \lambda_1 \|v\|_2^2.$$

(2) $A \in M_{k,k}(\mathbb{R})$ sei symmetrisch. Dann ist $A$ genau dann positiv definit, wenn alle Hauptunterdeterminanten positiv sind.

Mit Hilfe der Hesse-Matrix und dem Begriff „positiv definit“ können wir nun hinreichende Bedingungen für lokale Extrema herleiten.

Satz VII.4.23 (Hinreichende Bedingungen für lokale Extrema). *Sei $U \subset \mathbb{R}^k$, $U \neq \emptyset$, offen, $x_0 \in U$, $f \in C^2(U, \mathbb{R})$ und $x_0$ ein kritischer Punkt von $f$, d.h. $Df(x_0) = 0$. Dann gilt:*

(1) *$D^2 f(x_0)$ positiv definit $\Longrightarrow$ $f$ hat in $x_0$ ein lokales (isoliertes) Minimum.*
(2) *$D^2 f(x_0)$ negativ definit $\Longrightarrow$ $f$ hat in $x_0$ ein lokales (isoliertes) Maximum.*
(3) *$D^2 f(x_0)$ indefinit $\Longrightarrow$ $x_0$ ist kein lokales Extremum von $f$.*

Beweis. ad (1): Aus Satz VII.4.14 folgt für $h \in \mathbb{R}^k \backslash \{0\}$ mit $x_0 + h \in U$:

$$f(x_0 + h) = f(x_0) + \frac{1}{2} h^t D^2(x_0) h + \tilde{R}_2(f, x_0, h).$$

Sei $\lambda_1 > 0$ der kleinste Eigenwert von $D^2 f(x_0)$. Wegen Bemerkung VII.4.15 gibt es ein $\delta > 0$ mit

$$x_0 + h \in U \text{ und } |\tilde{R}_2(f, x_0, h)| \leq \frac{\lambda_1}{4} \|h\|_2^2 \; \forall h \in \mathbb{R}^k \backslash \{0\} \text{ mit } \|h\|_2 \leq \delta.$$

Für $h \in \mathbb{R}^k \backslash \{0\}$ mit $\|h\|_2 \leq \delta$ folgt dann

$$\begin{aligned} f(x_0 + h) &\geq f(x_0) + \frac{\lambda_1}{2} \|h\|_2^2 - \frac{\lambda_1}{4} \|h\|_2^2 \\ &= f(x_0) + \frac{\lambda_1}{4} \|h\|_2^2 \\ &> f(x_0). \end{aligned}$$

Also ist $x_0$ ein lokales (isoliertes) Minimum.
AD (2): Folgt aus Teil (1) angewandt auf $-f$.
AD (3): Seien $v, w \in \mathbb{R}^k \setminus \{0\}$ mit

$$v^t D^2(x_0)v > 0 \text{ und } w^t D^2(x_0)w < 0.$$

Für hinreichend kleines $t > 0$ folgt dann aus Satz VII.4.14 und Bemerkung VII.4.15

$$\begin{aligned} f(x_0 + tv) &= f(x_0) + \frac{1}{2} v^t D^2 f(x_0) v t^2 + \tilde{R}_2(f, x_0, tv) \\ &= f(x_0) + t^2 [\frac{1}{2} v^t D^2 f(x_0) v + \underbrace{t^{-2} \tilde{R}_2(f, x_0, tv)}_{\substack{\longrightarrow 0 \\ t \to 0}}] \\ &> f(x_0) \end{aligned}$$

und

$$\begin{aligned} f(x_0 + tw) &= f(x_0) + \frac{1}{2} t^2 w^t D^2 f(x_0) w + \tilde{R}_2(f, x_0, tw) \\ &= f(x_0) + t^2 [\frac{1}{2} w^t D^2 f(x_0) w + \underbrace{t^{-2} \tilde{R}_2(f, x_0, tw)}_{\substack{\longrightarrow 0 \\ t \to 0}}] \\ &< f(x_0). \end{aligned}$$

Hieraus folgt die Behauptung. □

BEISPIEL VII.4.24. (1) Für $f(x, y) = c + x^2 + y^2$, mit $c \in \mathbb{R}$ erhalten wir

$$Df(x, y) = 2(x, y) \Longrightarrow (0,0) \text{ ist einziger kritischer Punkt}$$
$$D^2 f(x, y) = 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Longrightarrow (0,0) \text{ ist lokales Minimum.}$$

(2) Für $f(x, y) = c - x^2 - y^2$, mit $c \in \mathbb{R}$ erhalten wir

$$Df(x, y) = -2(x, y) \Longrightarrow (0,0) \text{ ist einziger kritischer Punkt}$$
$$D^2 f(x, y) = -2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Longrightarrow (0,0) \text{ ist lokales Maximum.}$$

(3) Für $f(x, y) = c + x^2 - y^2$, mit $c \in \mathbb{R}$ erhalten wir

$$Df(x, y) = 2(x, -y) \Longrightarrow (0,0) \text{ ist einziger kritischer Punkt}$$
$$D^2 f(x, y) = 2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \Longrightarrow (0,0) \text{ ist kein lokales Extremum.}$$

(4) Für $f(x, y) = c + x^2 + y^4$ erhalten wir

$$Df(x, y) = (2x, 4y^3) \Longrightarrow (0,0) \text{ ist einziger kritischer Punkt}$$
$$D^2 f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \text{ ist positiv semi-definit}$$

$(0,0)$ ist lokales Minimum.

(5) Für $f(x,y) = x^2$ erhalten wir

$$Df(x,y) = (2x,0) \Longrightarrow (0,y), y \in \mathbb{R} \text{ sind kritische Punkte}$$

$$D^2 f(0,y) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \text{ ist positiv semi-definit}$$

$(0,y)$ sind lokale Minima, aber nicht isoliert.

(6) Für $f(x,y) = x^2 + y^3$ erhalten wir

$$Df(x,y) = (2x,3y^2) \Longrightarrow (0,0) \text{ ist einziger kritischer Punkt}$$

$$D^2 f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \text{ ist positiv semi-definit}$$

$(0,0)$ ist kein lokales Extremum.

(7) Betrachte

$$f(x,y) = \frac{xy}{1+x^4+y^4}.$$

Wegen

$$\lim_{\|(x,y)\|\to\infty} f(x,y) = 0$$

und

$$f(1,1) > 0$$
$$f(1,-1) < 0$$

hat $f$ ein Minimum und Maximum. Aus Beispiel VII.4.17 folgt

$$Df(x,y) = \frac{1}{(1+x^4+y^4)^2} \begin{pmatrix} y(1-3x^4+y^4) \\ x(1+x^4-3y^4) \end{pmatrix}.$$

Eine leichte Rechnung zeigt, dass

$$(0,0), (\frac{1}{\sqrt[4]{2}}, \frac{1}{\sqrt[4]{2}}), (-\frac{1}{\sqrt[4]{2}}, -\frac{1}{\sqrt[4]{2}}), (\frac{1}{\sqrt[4]{2}}, -\frac{1}{\sqrt[4]{2}}), (-\frac{1}{\sqrt[4]{2}}, \frac{1}{\sqrt[4]{2}})$$

die einzigen kritischen Punkte sind. Aus Beispiel VII.4.17 folgt

$$D^2 f(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ indefinit,}$$

$$\begin{aligned} D^2 f(\frac{1}{\sqrt[4]{2}}, \frac{1}{\sqrt[4]{2}}) &= D^2 f(-\frac{1}{\sqrt[4]{2}}, -\frac{1}{\sqrt[4]{2}}) \\ &= \begin{pmatrix} -\frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{3}{2} \end{pmatrix} \text{ negativ definit,} \\ D^2 f(-\frac{1}{\sqrt[4]{2}}, \frac{1}{\sqrt[4]{2}}) &= D^2 f(\frac{1}{\sqrt[4]{2}}, -\frac{1}{\sqrt[4]{2}}) \end{aligned}$$

$$= \begin{pmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} \end{pmatrix} \text{ positiv definit.}$$

Also ist $(0,0)$ kein lokales Extremum, $(\frac{1}{\sqrt[4]{2}}, \frac{1}{\sqrt[4]{2}})$ und $(-\frac{1}{\sqrt[4]{2}}, -\frac{1}{\sqrt[4]{2}})$ sind (isolierte) lokale Maxima, $(-\frac{1}{\sqrt[4]{2}}, \frac{1}{\sqrt[4]{2}})$ und $(\frac{1}{\sqrt[4]{2}}, -\frac{1}{\sqrt[4]{2}})$ sind (isolierte) lokale Minima.

Daher gilt für alle $x, y \in \mathbb{R}$

$$-\frac{1}{\sqrt{8}} \le f(x,y) \le \frac{1}{\sqrt{8}}.$$

## VII.5. Umkehrabbildungen

Wir erinnern an Satz IV.2.11 (S. 124): Sei $I \subset \mathbb{R}$ ein perfektes Intervall, $f \in C(I, \mathbb{R})$ in $I$ differenzierbar mit $f'(x) \neq 0$ für alle $x \in I$. Dann ist $J = f(I)$ ein Intervall, $f : I \to J$ bijektiv und $f^{-1} : J \to I$ differenzierbar und

$$(f^{-1})'(y) = \frac{1}{f'(x)} \quad \text{mit } y = f(x).$$

In diesem Paragraphen wollen wir ein analoges Resultat für Funktionen mehrerer Veränderlicher beweisen.

Im Folgenden sind $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ und $(Z, \|\cdot\|_Z)$ Banachräume.

Definition VII.5.1. $\mathcal{GL}(X) = \mathrm{Isom}(X, X) \subset \mathcal{L}(X, X)$ heißt die Automorphismengruppe von $X$.

Satz VII.5.2. *Sei $A \in \mathcal{L}(X, X)$ mit $\|A\|_{\mathcal{L}(X,X)} < 1$. Dann ist $I - A \in \mathcal{GL}(X)$, wobei $I$ die Identität auf $X$ bezeichnet. Weiter gilt*

$$\|(I - A)^{-1}\|_{\mathcal{L}(X,X)} \le \frac{1}{1 - \|A\|_{\mathcal{L}(X,X)}}.$$

Beweis. Wegen

$$\|A^n\|_{\mathcal{L}(X,X)} \le \|A\|^n_{\mathcal{L}(X,X)} \qquad \forall n \in \mathbb{N}^*$$

und $\|A\|_{\mathcal{L}(X,X)} < 1$ konvergiert die Reihe $\sum A^n$ in $\mathcal{L}(X, X)$. Für $n \in \mathbb{N}^*$ folgt

$$\begin{aligned}(I - A) \cdot (\sum_{k=0}^{n} A^k) &= \sum_{k=0}^{n} A^k - \sum_{k=0}^{n} A^{k+1} \\ &= I - A^{n+1} \\ &= (\sum_{k=0}^{n} A^k)(I - A).\end{aligned}$$

Grenzübergang $n \to \infty$ liefert

$$\sum_{k=0}^{\infty} A^k = (I - A)^{-1}.$$

Schließlich folgt

$$\begin{aligned}
\|(I-A)^{-1}\|_{\mathcal{L}(X,X)} &= \lim_{n\to\infty} \|\sum_{k=0}^{n} A^k\|_{\mathcal{L}(X,X)} \\
&\le \lim_{n\to\infty} \sum_{k=0}^{n} \|A\|_{\mathcal{L}(X,X)}^k \\
&= \frac{1}{1 - \|A\|_{\mathcal{L}(X,X)}}.
\end{aligned}$$

□

Aus Satz VII.5.2 folgt:

Satz VII.5.3. *(1)* $\mathrm{Isom}(X,Y)$ *ist offen in* $\mathcal{L}(X,Y)$.
*(2) Die Funktion*

$$\begin{aligned}
f : \mathrm{Isom}(X,Y) &\longrightarrow \mathcal{L}(X,Y) \\
A &\longrightarrow A^{-1}
\end{aligned}$$

*ist* $C^\infty$ *und*

$$Df(A)B = -A^{-1}BA^{-1} \quad \forall A \in \mathrm{Isom}(X,Y),\ B \in \mathcal{L}(X,Y).$$

Beweis. ad (1): Sei $A_0 \in \mathrm{Isom}(X,Y)$ und

$$\varepsilon = \|A_0^{-1}\|_{\mathcal{L}(Y,X)}^{-1}.$$

Sei $A \in \mathcal{L}(X,Y)$ mit

$$\|A - A_0\|_{\mathcal{L}(X,Y)} < \varepsilon.$$

Dann folgt

$$A = A_0 + A - A_0 = A_0[I - A_0^{-1}(A_0 - A)]. \tag{VII.5.1}$$

Wegen

$$\|A_0^{-1}(A_0 - A)\|_{\mathcal{L}(X,X)} \le \|A_0^{-1}\|_{\mathcal{L}(Y,X)} \|A - A_0\|_{\mathcal{L}(X,Y)} < 1$$

und Satz VII.5.2 ist $I - A_0^{-1}(A_0 - A) \in \mathcal{GL}(X)$. Zusammen mit Gleichung (VII.5.1) folgt hieraus $A \in \mathrm{Isom}(X,Y)$.
ad (2): 1. Schritt: $f$ ist stetig.
Seien $A, A_0 \in \mathrm{Isom}(X,Y)$ mit

$$\|A - A_0\|_{\mathcal{L}(X,Y)} < \|A_0^{-1}\|_{\mathcal{L}(Y,X)}^{-1}.$$

Zusammen mit Teil (1) folgt

$$\begin{aligned}
&f(A) - f(A_0) \\
&= A^{-1} - A_0^{-1}
\end{aligned}$$

$$
\begin{aligned}
&= (A^{-1}A_0 - I)A_0^{-1} \\
&= ([A - A_0 + A_0]^{-1}A_0 - I)A_0^{-1} \\
&= ([A_0(A_0^{-1}\{A - A_0\} + I)]^{-1}A_0 - I)A_0^{-1} \\
&= ([I - A_0^{-1}\{A_0 - A\}]^{-1} - I)A_0^{-1} \\
&= [I - A_0^{-1}(A_0 - A)]^{-1}[I - (I - A_0^{-1}(A_0 - A))]A_0^{-1} \\
&= [I - A_0^{-1}(A_0 - A)]^{-1}A_0^{-1}(A_0 - A)A_0^{-1} \qquad \text{(VII.5.2)}
\end{aligned}
$$

und wegen Satz VII.5.2

$$
\begin{aligned}
&\|f(A) - f(A_0)\|_{\mathcal{L}(Y,X)} \\
&\le \frac{\|A_0^{-1}\|^2_{\mathcal{L}(Y,X)}}{1 - \|A_0^{-1}\|_{\mathcal{L}(Y,X)}\|A_0 - A\|_{\mathcal{L}(X,Y)}}\|A_0 - A\|_{\mathcal{L}(X,Y)}.
\end{aligned}
$$

Also ist $f$ stetig.
2. Schritt: $f$ ist differenzierbar.
Sei $A_0 \in \operatorname{Isom}(X, Y)$ und $B \in \mathcal{L}(X, Y)$ mit

$$\|B\|_{\mathcal{L}(X,Y)} < \|A_0^{-1}\|^{-1}_{\mathcal{L}(Y,X)}.$$

Aus Gleichung (VII.5.2) folgt

$$
\begin{aligned}
&f(A_0 + B) - f(A_0) + A_0^{-1}BA_0^{-1} \\
&= -[I + A_0^{-1}B]^{-1}A_0^{-1}BA_0 + A_0^{-1}BA_0^{-1} \\
&= [I + A_0^{-1}B]^{-1}\{-A_0^{-1}BA_0 + (I + A_0^{-1}B)A_0^{-1}BA_0^{-1}\} \\
&= [I + A_0^{-1}B]^{-1}A_0^{-1}BA_0^{-1}BA_0^{-1}
\end{aligned}
$$

und somit

$$
\begin{aligned}
&\frac{1}{\|B\|_{\mathcal{L}(X,Y)}}\|f(A_0 + B) - f(A_0) + A_0^{-1}BA_0^{-1}\|_{\mathcal{L}(Y,X)} \\
&\le \frac{\|A_0^{-1}\|^3_{\mathcal{L}(Y,X)}\|B\|_{\mathcal{L}(X,Y)}}{1 - \|A_0^{-1}\|_{\mathcal{L}(Y,X)}\|B\|_{\mathcal{L}(X,Y)}} \\
&\underset{\|B\|_{\mathcal{L}(X,Y)} \to 0}{\longrightarrow} 0.
\end{aligned}
$$

Also ist $f$ differenzierbar und $Df(A_0)B = -A_0^{-1}BA_0^{-1}$.
3. Schritt: $f$ ist $C^\infty$.
Definiere für $S \in \mathcal{L}(X, Y)$ die Abbildung $g_S : \mathcal{L}(Y, X) \times \mathcal{L}(Y, X) \to \mathcal{L}(Y, X)$ durch

$$g_S(T_1, T_2) = T_1 S T_2 \quad \forall T_1, T_2 \in \mathcal{L}(Y, X).$$

Offensichtlich ist $g$ stetig und bilinear und somit aus $C^\infty$. Aus Schritt 2 folgt für $A \in \operatorname{Isom}(X, Y)$ und $B \in \mathcal{L}(X, Y)$

$$Df(A)B = g_B(f(A), f(A)).$$

Hieraus folgt die Behauptung durch Induktion. □

Wir kommen nun zu einem ersten Hauptsatz dieses Abschnittes.

Satz VII.5.4 (Satz über die Umkehrabbildung). *Sei $U \subset X$, $U \neq \emptyset$, offen, $f \in C^m(U,Y)$, $m \in \mathbb{N}^*$, $x_0 \in U$, $y_0 = f(x_0)$ und $Df(x_0) \in \mathrm{Isom}(X,Y)$. Dann gilt:*

(1) $\exists V \in \mathcal{U}(x_0), W \in \mathcal{U}(y_0) : f|_V : V \to W$ *ist bijektiv,*
(2) $f^{-1} = (f|_V)^{-1} \in C^m(W,V)$,
(3) $Df^{-1}(y) = Df(x)^{-1}$ *für alle* $y = f(x), x \in V, y \in W$.

Beweis. 1. Schritt: Reduktion auf den Fall $X = Y$, $x_0 = y_0 = 0$, $Df(x_0) = I$.
Definiere

$$\begin{aligned}\tilde{U} &= \{x - x_0 : x \in U\}\\ \tilde{f}(x) &= Df(x_0)^{-1}[f(x+x_0) - f(x_0)] \quad \forall x \in \tilde{U}\end{aligned}$$

Dann ist $\tilde{f} \in C^m(\tilde{U}, X)$, $\tilde{f}(0) = 0$ und

$$D\tilde{f}(0) = I.$$

Weiter gilt

$$f(x) = y \iff \tilde{f}(x - x_0) = Df(x_0)^{-1}[y - f(x_0)].$$

Daher können wir im Folgenden stets annehmen, dass gilt $X = Y$, $x_0 = y_0 = 0$ und $Df(x_0) = I$.
2. Schritt: $\exists V, W \in \mathcal{U}(0) : f|_V : V \to W$ ist bijektiv.
Wegen $f \in C^1(U,X)$ und $Df(0) = I$ gibt es ein $\varepsilon > 0$ mit

$$\|I - Df(x)\|_{\mathcal{L}(X,X)} \le \frac{1}{2} \quad \forall x \in B(0,\varepsilon).$$

Definiere

$$W = B(0, \frac{\varepsilon}{4}).$$

Für $y \in W$ definieren wir $\Phi_y : X \to X$ durch

$$\Phi_y(x) = x - f(x) + y.$$

Offensichtlich gilt

$$\Phi_y(x) = x \iff f(x) = y.$$

Für $x_1, x_2 \in \overline{B(0, \frac{\varepsilon}{2})}$ gilt

$$\begin{aligned}\|\Phi_y(x_1) - \Phi_y(x_2)\|_X &= \|x_1 - x_2 - f(x_1) + f(x_2)\|_X\\ &= \| \int_0^1 [I - Df(x_1 + t(x_2 - x_1))](x_1 - x_2) dt\|_X\\ &\le \frac{1}{2}\|x_1 - x_2\|_X.\end{aligned}$$

Wegen

$$\|\Phi_y(0)\|_X = \|y\|_X < \frac{\varepsilon}{4} = \frac{1}{2} \cdot \frac{\varepsilon}{2}$$

folgt hieraus mit Satz IV.4.4 (S. 138), dass $\Phi_y$ einen eindeutigen Fixpunkt in $\overline{B(0,\frac{\varepsilon}{2})} \subset B(0,\varepsilon)$ hat.
Definiere

$$V = \{x \in B(0,\varepsilon) : f(x) \in B(0,\frac{\varepsilon}{4})\}.$$

Dann ist $V \in \mathcal{U}(0)$, und aus Obigem folgt, dass $f|_V : V \to W$ bijektiv ist.
3. Schritt: $f^{-1} = (f|_V)^{-1}$ ist stetig.
Seien $y_1, y_2 \in W$ und $x_i = f^{-1}(y_i)$, $1 \le i \le 2$. Dann folgt

$$\begin{aligned}
&\|f^{-1}(y_1) - f^{-1}(y_2)\|_X \\
&= \|x_1 - f(x_1) + y_1 - x_2 + f(x_2) - y_2\|_X \\
&\le \|\int_0^1 [I - Df(x_1 + t(x_2 - x_1))](x_1 - x_2)dt\|_X + \|y_1 - y_2\|_X \\
&\le \frac{1}{2}\|x_1 - x_2\|_X + \|y_1 - y_2\|_X \\
&= \frac{1}{2}\|f^{-1}(y_1) - f^{-1}(y_2)\|_X + \|y_1 - y_2\|_X
\end{aligned}$$

und daher

$$\|f^{-1}(y_1) - f^{-1}(y_2)\|_X \le 2\|y_1 - y_2\|_X.$$

Also ist $f^{-1}$ stetig.
4. Schritt: $f^{-1}$ ist differenzierbar und

$$Df^{-1}(y) = Df(x)^{-1} \quad \forall y = f(x), x \in V, y \in W.$$

Seien $y_1, y_2 \in W$ und $x_1, x_2 \in V$ wie in Schritt 3. Dann folgt

$$\begin{aligned}
&\|f^{-1}(y_2) - f^{-1}(y_1) - Df(x_1)^{-1}(y_2 - y_1)\|_X \\
&= \|x_2 - x_1 - Df(x_1)^{-1}(f(x_2) - f(x_1))\|_X \\
&= \|\int_0^1 [I - Df(x_1)^{-1}Df(x_1 + t(x_2 - x_1))](x_2 - x_1)dt\|_X \\
&\le \|Df(x_1)^{-1}\|_{\mathcal{L}(X,X)} \\
&\quad \cdot \underbrace{\sup_{0 \le t \le 1} \|Df(x_1) - Df(x_1 + t(x_2 - x_1))\|_{\mathcal{L}(X,X)}}_{\to 0 \text{ für } x_2 \to x_1} \\
&\quad \cdot \|x_2 - x_1\|_X.
\end{aligned}$$

Hieraus folgt die Behauptung.
5. Schritt: $f^{-1} \in C^m(W,V)$.
Gemäß Schritt 4 ist

$$Df^{-1} = ((Df) \circ f^{-1})^{-1}. \qquad \text{(VII.5.3)}$$

Da gemäß Schritt 3 $f^{-1}$ stetig ist, folgt $f^{-1} \in C^1(W,V)$. Der Rest folgt durch Induktion aus Gleichung (VII.5.3) mit der Kettenregel und Satz VII.5.3. Damit ist der Satz bewiesen. □

DEFINITION VII.5.5. (1) Seien $U \subset X$, $X \neq \emptyset$, und $V \subset Y$, $V \neq \emptyset$, offen. Dann heißt die Funktion $f : U \to V$ ein $C^m$-DIFFEOMORPHISMUS, $m \in \mathbb{N} \cup \{\infty, \omega\}$ wenn gilt

(1) $f : U \to V$ ist bijektiv,
(2) $f \in C^m(U, V)$,
(3) $f^{-1} \in C^m(V, U)$.

Im Falle $m = 0$ spricht man auch von einem HOMÖOMORPHISMUS.
(2) Sei $U \subset X$, $U \neq \emptyset$, offen. Dann heißt die Funktion $f : U \to Y$ ein LOKALER $C^m$-DIFFEOMORPHISMUS, $m \in \mathbb{N} \cup \{\infty, \omega\}$ wenn es zu jedem $x_0 \in U$, ein $V \in \mathcal{U}(x_0)$ und ein $W \in \mathcal{U}(f(x_0))$ gibt, so dass $f|_V : V \to W$ ein $C^m$- Diffeomorphismus ist. Im Fall $m = 0$ spricht man auch von einem LOKALEN HOMÖOMORPHISMUS bzw. sagt, $f$ sei LOKAL TOPOLOGISCH.

BEMERKUNG VII.5.6. (1) $f$ (lokaler) $C^{m+1}$-Diffeomorphismus $\Longrightarrow$ $f$ (lokaler) $C^m$-Diffeomorphismus.
(2) $f \in C^{m+1}(U, Y)$, $f$ $C^m$-Diffeomorphismus $\not\Longrightarrow$ $f$ $C^{m+1}$-Diffeomorphismus.
(3) $f : U \to Y$ lokal topologisch $\Longrightarrow$ $f$ ist offen, d.h., Bilder offener Mengen sind offen.
(4) $f \in C^m(U, Y)$. Dann gilt:
$f$ lokaler $C^m$-Diffeomorphismus $\iff$ $Df(x) \in \operatorname{Isom}(X, Y)$ für alle $x \in U$.
(5) $U \subset \mathbb{R}^k$, $U \neq \emptyset$, offen und $f \in C^m(U, \mathbb{R}^k)$. Dann gilt:
$f$ lokaler $C^m$-Diffeomorphismus $\iff$ $\det(Df(x)) \neq 0$ für alle $x \in U$.
(6) Unter den Voraussetzungen von Satz VII.5.4 braucht $f$ kein GLOBALER $C^m$-Diffeomorphismus zu sein.

BEWEIS. AD (1): Ist klar.
AD (2): $X = Y = \mathbb{R}$, $f(x) = x^3$. Dann ist $f \in C^\omega(\mathbb{R}, \mathbb{R})$ topologisch, aber $f^{-1}(y) = \sqrt[3]{y}$ ist in $y = 0$ nicht differenzierbar.
AD (3): Ist klar.
AD (4): „$\Longrightarrow$" Satz VII.5.4.
„$\Longleftarrow$": $f \circ f^{-1} = id_Y$, $f^{-1} \circ f = id_X$. Mit der Kettenregel folgt hieraus für alle $y = f(x)$

$$Df(x) \cdot Df^{-1}(y) = I_{\mathcal{L}(Y,Y)}$$
$$Df^{-1}(y) \cdot Df(x) = I_{\mathcal{L}(X,X)}.$$

AD (5): In diesem Fall kann man $Df(x)$ mit einer Matrix identifizieren und es ist $Df(x) \in \operatorname{Isom}(\mathbb{R}^k, \mathbb{R}^k)$ genau dann, wenn gilt $\det(Df(x)) \neq 0$.
AD (6): $X = Y = \mathbb{C}$, $f(z) = e^z$. $f$ ist $2\pi i$-periodisch, also nicht injektiv, aber ein lokaler $C^\omega$-Diffeomorphismus. $\square$

BEISPIEL VII.5.7. Es gibt ein $\varepsilon > 0$, so dass das Gleichungssystem

$$\tan(x + y) = u$$

$$\ln(1+y+xy) = v$$

für alle $u, v \in \mathbb{R}$ mit $\sqrt{u^2+v^2} < \varepsilon$ mindestens eine Lösung hat.
Dies folgt aus Satz VII.5.4 und Bemerkung VII.5.6 (5) angewandt auf die Funktion $F : \mathbb{R}^2 \to \mathbb{R}^2$ mit

$$F(x,y) = \begin{pmatrix} \tan(x+y) \\ \ln(1+y+xy) \end{pmatrix}$$

wegen

$$DF(x,y) = \begin{pmatrix} \dfrac{1}{\cos^2(x+y)} & \dfrac{1}{\cos^2(x+y)} \\ \dfrac{y}{1+y+xy} & \dfrac{1+x}{1+y+xy} \end{pmatrix}$$

und

$$\begin{aligned} F(0,0) &= (0,0) \\ \det(DF(0,0)) &= \det \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \\ &= 1. \end{aligned}$$

Wir kommen nun zu dem zweiten Hauptsatz dieses Abschnittes. Dazu benötigen wir folgende Notation.

Definition VII.5.8. Seien $X_1, X_2$ Untervektorräume von $X$, so dass $X$ dargestellt werden kann als $X = X_1 \times X_2$. Sei $U \subset X$, $U \neq \emptyset$, offen und $f : U \to Y$ in $U$ differenzierbar. Dann stellen wir jeden Punkt $x \in X$ in der Form $x = (x_1, x_2)$, $x_1 \in X_1$, $x_2 \in X_2$ dar und bezeichnen mit

$$D_{x_1} f \in \mathcal{L}(X_1, Y) \text{ und } D_{x_2} f \in \mathcal{L}(X_2, Y)$$

die Ableitungen der Funktionen

$$\begin{aligned} f(\cdot, x_2^*) : &\{x_1 \in X_1 : (x_1, x_2^*) \in U\} \to Y \\ &x_1 \mapsto f(x_1, x_2^*) \quad (x_2^* \text{ fest}) \end{aligned}$$

bzw.

$$\begin{aligned} f(x_1^*, \cdot) : &\{x_2 \in X_2 : (x_1^*, x_2) \in U\} \to Y \\ &x_2 \mapsto f(x_1^*, x_2) \quad (x_1^* \text{ fest}). \end{aligned}$$

Satz VII.5.9 (Satz über implizite Funktionen). *Sei $U \subset X$, $U \neq \emptyset$, offen, $m \in \mathbb{N}^*$, und $f \in C^m(U, Y)$. $X$ lasse sich in der Form $X = X_1 \times X_2$ mit Untervektorräumen $X_1, X_2$ darstellen. Es gebe ein $x^* = (x_1^*, x_2^*) \in U$ mit*

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

(b) $$D_{x_2} f(x^*) \in \operatorname{Isom}(X_2, Y).$$

*Dann gibt es Umgebungen $V_1$ und $W$ von $x_1^*$ in $X_1$ und $V_2$ von $x_2^*$ in $X_2$ und ein $g \in C^m(W, X_2)$ mit folgenden Eigenschaften:*

(1) $D_{x_2} f(x_1, x_2) \in \mathrm{Isom}(X_2, Y)$ *für alle* $x_1 \in V_1, x_2 \in V_2$.
(2) $(x_1, x_2) \in V_1 \times V_2$ *und* $f(x_1, x_2) = 0 \iff x_1 \in W$ *und* $x_2 = g(x_1)$, *d.h., die Nullstellenmenge* $f^{-1}(0)$ *lässt sich in einer Umgebung von* $x^*$ *als Graph der Funktion* $g$ *darstellen, bzw. die Gleichung* $f(x_1, x_2) = 0$ *kann in einer Umgebung von* $x^*$ *nach* $x_2$ *aufgelöst werden.*
(3) $Dg(x_1) = -[D_{x_2} f(x_1, g(x_1))]^{-1} D_{x_1} f(x_1, g(x_1))$ *für alle* $x_1 \in W$.

Beweis. 1. Schritt: Reduktion auf Normalform.
Definiere $\tilde{f} : U \to X_2$ durch

$$\tilde{f}(x_1, x_2) = [D_{x_2} f(x_1^*, x_2^*)]^{-1} f(x_1, x_2).$$

Dann ist $\tilde{f} \in C^m(U, X_2)$ und

$$\tilde{f}(x_1, x_2) = 0 \iff f(x_1, x_2) = 0$$
$$D_{x_2} \tilde{f}(x_1^*, x_2^*) = I_{X_2}.$$

Also können wir im Folgenden o.E. annehmen, dass $X_2 = Y$ und $D_{x_2} f(x_1^*, x_2^*) = I_{X_2}$ ist.
2. Schritt: Konstruktion von $V = V_1 \times V_2$.
O.E. können wir annehmen, dass $U = U_1 \times U_2$ ist mit offenen Mengen $U_1 \subset X_1$, $U_2 \subset X_2$. Definiere $\varphi : U_1 \times U_2 \to X_1 \times X_2$ durch

$$\varphi(x_1, x_2) = (x_1, f(x_1, x_2)).$$

Dann ist $\varphi \in C^m(U_1 \times U_2, X_1 \times X_2)$ und

$$D\varphi(x_1, x_2)(u_1, u_2) = (u_1, D_{x_1} f(x_1, x_2) u_1 + D_{x_2} f(x_1, x_2) u_2) \quad \forall u_1 \in X_1, u_2 \in X_2.$$

Definiere $B \in \mathcal{L}(X_1 \times X_2, X_1 \times X_2)$ durch

$$B(u_1, u_2) = (u_1, -D_{x_1} f(x_1^*, x_2^*) u_1 + u_2) \quad \forall u_1 \in X_1, u_2 \in X_2.$$

Dann folgt für $u_1 \in X_1$, $u_2 \in X_2$

$$\begin{aligned} B(D\varphi(x_1^*, x_2^*)(u_1, u_2)) &= B(u_1, D_{x_1} f(x_1^*, x_2^*) u_1 + u_2) \\ &= (u_1, u_2) \end{aligned}$$

und

$$\begin{aligned} D\varphi(x_1^*, x_2^*)(B(u_1, u_2)) &= D\varphi(x_1^*, x_2^*)(u_1, -D_{x_1} f(x_1^*, x_2^*) u_1 + u_2) \\ &= (u_1, u_2). \end{aligned}$$

Also ist $D\varphi(x_1^*, x_2^*) \in \mathrm{Isom}(X_1 \times X_2, X_1 \times X_2)$ und $B = D\varphi(x_1^*, x_2^*)^{-1}$. Gemäß Satz VII.5.4 gibt es eine Umgebung $V$ von $x^*$ in $X$ und eine Umgebung $\tilde{W}$ von $(x_1^*, 0) = \varphi(x_1^*, x_2^*)$ in $X$, so dass $\varphi$ ein $C^m$-Diffeomorphismus von $V$ auf $\tilde{W}$ ist. O.E. können wir annehmen, dass $V = V_1 \times V_2$ ist mit Umgebungen $V_1$ von $x_1^*$ in $X_1$ und $V_2$ von $x_2^*$ in $X_2$.

3. SCHRITT: Konstruktion von $g$ und $W$.
Sei $\psi = (\varphi|_V)^{-1}$. Dann gilt

$$\psi(u_1, u_2) = (x_1, x_2) \iff \varphi(x_1, x_2) = (u_1, u_2).$$

Aus der Konstruktion von $\varphi$ folgt, dass $\psi$ geschrieben werden kann als

$$\psi(u_1, u_2) = (u_1, h(u_1, u_2))$$

mit $h \in C^m(\tilde{W}, V_2)$. Insbesondere gilt

$$\begin{aligned} &(x_1, x_2) \in V \text{ und } f(x_1, x_2) = v \\ \iff &(x_1, x_2) \in \tilde{W} \text{ und } h(x_1, v) = x_2. \end{aligned}$$

Definiere

$$W = \{x_1 \in X_1 : (x_1, 0) \in \tilde{W}\} = \tilde{W} \cap (X_1 \times \{0\})$$

und $g : W \to X_2$ durch

$$g(x_1) = h(x_1, 0).$$

Dann ist $W$ eine Umgebung von $x_1^*$ in $X_1$ und $g \in C^m(W, X_2)$. Außerdem gilt

$$\begin{aligned} &(x_1, x_2) \in V \text{ und } f(x_1, x_2) = 0 \\ \iff &x_1 \in W \text{ und } x_2 = g(x_1). \end{aligned}$$

Damit ist die Behauptung (2) gezeigt.
Die Behauptung (1) folgt aus der Darstellung von $D\varphi$ und der Tatsache, dass für alle $(x_1, x_2) \in V$, gilt $D\varphi(x_1, x_2) \in \mathrm{Isom}(X_1 \times X_2, X_1 \times X_2)$.
4. SCHRITT: Bestimmung der Ableitung von $g$.
Definiere $F : W \to X_2$ durch

$$F(x_1) = f(x_1, g(x_1)).$$

Dann ist $F \in C^m(W, X_2)$ und für jedes $x_1 \in W$ gilt

$$\begin{aligned} F(x_1) &= 0 \\ 0 &= DF(x_1) \\ &= D_{x_1} f(x_1, g(x_1)) + D_{x_2} f(x_1, g(x_1)) \cdot Dg(x_1). \end{aligned}$$

Da $D_{x_2} f(x_1, g(x_1)) \in \mathrm{Isom}(X_2, X_2)$ ist für jedes $x_1 \in W$, folgt hieraus die Behauptung (3). □

BEISPIEL VII.5.10. Sei $k \geq 2$ und $f \in C^\infty(\mathbb{R}^k, \mathbb{R})$ definiert durch

$$f(x) = \sum_{i=1}^{k} x_i^2 - 1.$$

$S^{k-1} = f^{-1}(0)$ heißt $(k-1)$-SPHÄRE. Offensichtlich ist $S^{k-1} = \partial B(0, 1)$ und $0 \notin S^{k-1}$, $S^{k-1} \neq \emptyset$. Sei $x^* \in S^{k-1}$. Dann ist

$$Df(x^*) = 2(x_1^*, \ldots, x_2^*) \neq 0.$$

Also gibt es ein $i_0 \in \mathbb{N}^k$ mit $x^*_{i_0} \neq 0$. Wir können Satz VII.5.9 anwenden mit

$$\begin{aligned} X &= \mathbb{R}^k, \\ Y &= \mathbb{R}, \\ X_1 &= \{(x_1, \ldots, x_{i_0-1}, 0, x_{i_0+1}, \ldots, x_k)\} \cong \mathbb{R}^{k-1}, \\ X_2 &= \{(0, \ldots, 0, x_{i_0}, 0, \ldots, 0) : x_{i_0} \in \mathbb{R}\} \cong \mathbb{R} \end{aligned}$$

und erhalten, dass $S^{k-1}$ in einer Umgebung von $x^*$ als Graph einer Funktion

$$g : \mathbb{R}^{k-1} \to \mathbb{R}$$

in der Form

$$\{(x_1, \ldots, x_{i_0-1}, g(x_1, \ldots, x_{k-1}), x_{i_0}, \ldots, x_{k-1}) : (x_1, \ldots, x_{k-1}) \in \mathbb{R}^{k-1}\}$$

dargestellt werden kann.

Wir wollen Satz VII.5.9 auf Funktionen auf dem $\mathbb{R}^k$ mit Werten in $\mathbb{R}^l$ anwenden. Dazu erinnern wir an die Lineare Algebra.

Bemerkung VII.5.11. Sei $\dim X = k$, $\dim Y = l$ und $A \in \mathcal{L}(X, Y)$. Dann ist

$$\ker(A) = \{x \in X : Ax = 0\}$$

ein Untervektorraum von $X$ und

$$\mathrm{im}(A) = \{y \in Y : \exists\, x \in X : Ax = y\}$$

ein Untervektorraum von $Y$.

$$\mathrm{Rang}(A) = \dim(\mathrm{im}(A))$$

heißt der Rang von $A$. Es gilt

$$\dim X = \dim(\ker(A)) + \mathrm{Rang}(A).$$

Nach Einführen von Basen auf $X$ und $Y$ können wir $A$ mit einer Matrix aus $M_{l,k}(\mathbb{R})$ identifizieren. $\mathrm{Rang}(A)$ ist unabhängig von der Wahl der Basen. Es gilt:

> $\mathrm{Rang}(A) = n \iff$ Es gibt eine $n$-reihige Unterdeterminante, die nicht verschwindet, und jede $(n+1)$-reihige Unterdeterminante verschwindet.

Definition VII.5.12. Sei $U \subset \mathbb{R}^k$, $U \neq \emptyset$, offen und $f \in C^1(U, \mathbb{R}^l)$. Dann heißt $x_0 \in U$ regulärer Punkt von $f$ genau dann, wenn $Df(x_0)$ surjektiv ist. Letzteres ist genau dann der Fall, wenn

$$\mathrm{Rang}(Df(x_0)) = l$$

ist.

Mit diesen Begriffen lassen sich unsere bisherigen Ergebnisse auch folgendermaßen formulieren:

Bemerkung VII.5.13. Sei $U \subset \mathbb{R}^k$, $U \neq \emptyset$, offen und $f \in C^m(U, \mathbb{R}^l)$ mit $m \in \mathbb{N}^*$. Dann gilt:

(1) Besitzt $f$ einen regulären Punkt, so gilt $k \geq l$.

(2) Sei $k > l$ und $x_0$ ein regulärer Punkt von $f$ mit $f(x_0) = 0$. Dann besitzt $Df(x_0)$ genau $l$ linear unabhängige Spaltenvektoren. O.E. seien dies die letzten $l$ Spalten, sonst führe eine Koordinatentransformation durch. Dann kann die Lösung des Gleichungssystems

$$f(x) = 0$$

in einer Umgebung von $x_0$ als $C^m$-Funktion der ersten $k - l$ Variablen dargestellt werden. D.h., $M = f^{-1}(0)$ kann in einer Umgebung von $x_0$ als Graph einer $C^m$-Funktion von $k - l$ Variablen dargestellt werden.

(3) Eine nicht leere Teilmenge $M$ des $\mathbb{R}^k$ heißt $n$-dimensionale $C^m$-Untermannigfaltigkeit des $\mathbb{R}^k$, $0 \leq n < k$, $m \in \mathbb{N}^*$, wenn es zu jedem $x_0 \in M$ ein $V \subset \mathcal{U}(x_0)$ gibt, so dass $M \cap V$ als Graph einer $C^m$-Funktion von $n$-Variablen dargestellt werden kann. Ist speziell $n = k - 1$, so spricht man von einer Hyperfläche im $\mathbb{R}^k$.

(4) $S^{k-1}$ ist eine Hyperfläche im $\mathbb{R}^k$, $k \geq 2$.

(5) Es sei $M = f^{-1}(0) \neq \emptyset$ und jedes $x \in M$ sei regulärer Punkt von $f$. Dann ist $M$ eine $k-l$ dimensionale Untermannigfaltigkeit von $\mathbb{R}^k$.

Als Anwendung des Satzes über implizite Funktionen betrachten wir die Berechnung von Extrema unter Nebenbedingungen.

Satz VII.5.14. *Sei $U \subset \mathbb{R}^k$, $U \neq \emptyset$, offen, $1 \leq l < k$ und $f \in C^1(U, \mathbb{R})$, $g \in C^1(U, \mathbb{R}^l)$. Für den Punkt $x^* \in U$ gelte*

(a) $g(x^*) = 0$,

(b) *$x^*$ ist regulärer Punkt von $g$,*

(c) *$x^*$ ist ein lokales Extremum von $f|_M$ mit $M = g^{-1}(0)$.*

*Dann gibt es $l$ Zahlen $\lambda_1^*, \ldots, \lambda_l^* \in \mathbb{R}$, die sog.* Langrange-Multiplikatoren, *so dass $(x_1^*, \ldots, x_k^*, \lambda_1^*, \ldots, \lambda_l^*) \in \mathbb{R}^k \times \mathbb{R}^l$ kritischer Punkt der* Langrange-Funktion

$$F : U \times \mathbb{R}^l \to \mathbb{R}$$

$$(x, \lambda) \mapsto f(x) + \sum_{i=1}^{l} \lambda_i g_i(x)$$

*ist.*

Beweis. Gemäß Bemerkung VII.5.13 (2) besitzt $Dg(x^*)$ genau $l$ linear unabhängige Spaltenvektoren. O.E. sind dies die letzten $l$ Spalten, sonst führen wir eine Koordinatentransformation durch, die die Spalten von $Dg(x^*)$ geeignet vertauscht. Setze zur Abkürzung $n = k - l$

und $y^* = (x_1^*, \dots, x_n^*) \in \mathbb{R}^n$. Gemäß Satz VII.5.9 gibt es Umgebungen $V$ von $x^*$ in $\mathbb{R}^k$ und $W$ von $y^*$ in $\mathbb{R}^n$ und eine Funktion $\varphi \in C^1(W, \mathbb{R}^l)$, so dass gilt

$$M \cap V = \{(y, \varphi(y)) : y \in W\}.$$

Definiere die Funktion $\tilde{f} : W \to \mathbb{R}$ durch

$$\tilde{f}(y) = f(y, \varphi(y)) = f(y_1, \dots, y_n, \varphi_1(y), \dots, \varphi_l(y)).$$

Dann ist $\tilde{f} \in C^1(W, \mathbb{R})$, $\tilde{f} = f|_{M \cap V}$ und $y^*$ ist ein lokales Extremum von $\tilde{f}$. Also ist $y^*$ kritischer Punkt von $\tilde{f}$, d.h., für $1 \le i \le n$ gilt

$$\begin{aligned} 0 &= \frac{\partial}{\partial y_i} \tilde{f}(y^*) \\ &= \frac{\partial}{\partial x_i} f(y^*, \varphi(y^*)) + \sum_{j=1}^{l} \frac{\partial}{\partial x_{n+j}} f(y^*, \varphi(y^*)) \frac{\partial}{\partial y_i} \varphi_j(y^*) \\ &= \frac{\partial}{\partial x_i} f(x^*) + \sum_{j=1}^{l} \frac{\partial}{\partial x_{n+j}} f(x^*) \frac{\partial}{\partial y_i} \varphi_j(y^*). \end{aligned}$$

Setze zur Abkürzung

$$A = \Big( \frac{\partial}{\partial x_{n+j}} g_i(x^*) \Big)_{1 \le i,j \le l}.$$

Dann folgt aus Satz VII.5.9 für $1 \le i \le n$, $1 \le j \le l$

$$\begin{aligned} \frac{\partial}{\partial y_i} \varphi_j(y^*) &= (-A^{-1} D_{x_1} g(x^*))_{j,i} \\ &= - \sum_{\mu=1}^{l} (A^{-1})_{j\mu} \frac{\partial}{\partial x_i} g_\mu(x^*). \end{aligned}$$

Damit folgt für $1 \le i \le n$

$$\begin{aligned} 0 &= \frac{\partial}{\partial x_i} f(x^*) - \sum_{j=1}^{l} \sum_{\mu=1}^{l} (A^{-1})_{j\mu} \frac{\partial}{\partial x_{n+j}} f(x^*) \frac{\partial}{\partial x_i} g_\mu(x^*) \\ &= \frac{\partial}{\partial x_i} [f(x^*) + \sum_{\mu=1}^{l} \lambda_\mu^* g_\mu(x^*)] \\ &= \frac{\partial}{\partial x_i} F(x^*, \lambda^*) \end{aligned}$$

mit

$$\lambda_\mu^* = - \sum_{j=1}^{l} (A^{-1})_{j\mu} \frac{\partial}{\partial x_{n+j}} f(x^*) \quad 1 \le \mu \le l.$$

Schließlich folgt für $1 \le \nu \le l$

$$\begin{aligned}\frac{\partial}{\partial x_{n+\nu}} F(x^*, \lambda^*) &= \frac{\partial}{\partial x_{n+\nu}} f(x^*) + \sum_{\mu=1}^{l} \lambda^*_\mu \frac{\partial}{\partial x_{n+\nu}} g_\mu(x^*) \\ &= \frac{\partial}{\partial x_{n+\nu}} f(x^*) \\ &\quad - \sum_{\mu=1}^{l} \sum_{j=1}^{l} (A^{-1})_{j\mu} \frac{\partial}{\partial x_{n+j}} f(x^*) \underbrace{\frac{\partial}{\partial x_{n+\nu}} g^*_\mu}_{=A_{\mu\nu}} \\ &= \frac{\partial}{\partial x_{n+\nu}} f(x^*) - \sum_{j=1}^{l} \delta_{j\nu} \frac{\partial}{\partial x_{n+j}} f(x^*) \\ &= 0.\end{aligned}$$

Damit ist die Behauptung bewiesen. □

Beispiel VII.5.15. Sei

$$\begin{aligned}\mathcal{E} &= \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\} \\ \mathcal{E}_+ &= \{(x, y, z) \in \mathbb{R}^3 : x + y + z \ge 0\}.\end{aligned}$$

Gesucht sind die Mimima und Maxima der Funktion $f \in C^\infty(\mathbb{R}^3, \mathbb{R})$ mit

$$f(x, y, z) = \exp(xyz),$$

sowie die Punkte, an denen diese angenommen werden, auf den Mengen

(a) $S^2$,
(b) $S^2 \cap \mathcal{E}$,
(c) $S^2 \cap \mathcal{E}_+$.

Da die genannten Mengen kompakt sind, ist die Aufgabe wohlgestellt.
ad (a): Die Lagrange-Funktion lautet

$$F(x, y, z, \lambda) = e^{xyz} + \lambda(x^2 + y^2 + z^2 - 1).$$

Wir erhalten als Bedingung für einen kritischen Punkt:

$$\begin{aligned} yze^{xyz} + 2\lambda x &= 0 \\ xze^{xyz} + 2\lambda y &= 0 \\ xye^{xyz} + 2\lambda z &= 0 \\ x^2 + y^2 + z^2 - 1 &= 0.\end{aligned}$$

Multiplikation der ersten drei Gleichungen mit $x$ bzw. $y$ bzw. $z$ und Addition liefert wegen der vierten Gleichung

$$2\lambda + 3xyze^{xyz} = 0.$$

Einsetzen ergibt

$$e^{xyz} yz(1 - 3x^2) = 0$$

$$e^{xyz}xz(1-3y^2) = 0$$
$$e^{xyz}xy(1-3z^2) = 0,$$
$$(x,y,z) \in S^2.$$

Dies liefert die Lösungen

$$(\pm 1,0,0),\ (0,\pm 1,0),\ (0,0,\pm 1)$$

und

$$(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}).$$

Definiere

$$\begin{aligned} N &= \{(\pm 1,0,0),(0,\pm 1,0),(0,0,\pm 1)\}, \\ M &= \{(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}) : \\ &\qquad \text{- kommt eine gerade Anzahl Male vor}\}, \\ m &= \{(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}) : \\ &\qquad \text{- kommt eine ungerade Anzahl Male vor}\}. \end{aligned}$$

Dann folgt

$$\begin{aligned} f(x,y,z) &= 1 \quad \forall (x,y,z) \in N, \\ f(x,y,z) &= e^{\frac{1}{3\sqrt{3}}} \quad \forall (x,y,z) \in M, \\ f(x,y,z) &= e^{-\frac{1}{3\sqrt{3}}} \quad \forall (x,y,z) \in m. \end{aligned}$$

Also gilt

$$e^{-\frac{1}{3\sqrt{3}}} \le f(x,y,z) \le e^{\frac{1}{3\sqrt{3}}} \quad \forall (x,y,z) \in S^2$$

und das Maximum und Minimum wird genau in den Punkten von $M$ bzw. $m$ angenommen.

AD (B): In diesem Fall lautet die Lagrange-Funktion

$$F(x,y,z,\lambda,\mu) = e^{xyz} + \lambda(x^2+y^2+z^2-1) + \mu(x+y+z).$$

Die Bedingungen für einen kritischen Punkt lauten:

$$\begin{aligned} yze^{xyz} + 2\lambda x + \mu &= 0 \\ xze^{xyz} + 2\lambda y + \mu &= 0 \\ xye^{xyz} + 2\lambda z + \mu &= 0 \\ x^2+y^2+z^2-1 &= 0 \\ x+y+z &= 0. \end{aligned}$$

Wie in Teil (A) folgt wegen der letzten beiden Gleichungen

$$2\lambda + 3xyze^{xyz} = 0.$$

Einsetzen ergibt

$$\begin{aligned} yze^{xyz}(1-3x^2)+\mu &= 0\\ xze^{xyz}(1-3y^2)+\mu &= 0\\ xye^{xyz}(1-3z^2)+\mu &= 0\\ (x,y,z) &\in S^2\cap\mathcal{E}. \end{aligned}$$

Subtraktion der ersten Gleichung von der zweiten und dritten liefert

$$\begin{aligned} e^{xyz}z(y-x)(1+3xy) &= 0\\ e^{xyz}y(z-x)(1+3xz) &= 0. \end{aligned}$$

Wobei die zweite der obigen Gleichungen aus der ersten durch Vertauschen von $y$ und $z$ hervorgeht.
Wie man sich leicht überlegt, führen die Annahmen

$$z=0 \quad \text{oder} \quad y=0 \quad \text{oder} \quad x=0$$

zu einem Widerspruch zu der Bedingung $(x,y,z)\in S^2\cap\mathcal{E}$.
Aus $x=y$ und $(x,y,z)\in S^2\cap\mathcal{E}$ folgt

$$x=\pm\frac{1}{\sqrt{6}}, y=\pm\frac{1}{\sqrt{6}}, z=\mp\frac{2}{\sqrt{6}}.$$

Dies ist eine Lösung. Analog erhält man die Lösungen

$$\begin{gathered} (\mp\frac{2}{\sqrt{6}},\pm\frac{1}{\sqrt{6}},\pm\frac{1}{\sqrt{6}})\\ (\pm\frac{1}{\sqrt{6}},\mp\frac{2}{\sqrt{6}},\pm\frac{1}{\sqrt{6}}) \end{gathered}$$

und überzeugt sich, dass dies alle Lösungen sind. Setze

$$\begin{aligned} \tilde{M} &= \{(\frac{2}{\sqrt{6}},-\frac{1}{\sqrt{6}},-\frac{1}{\sqrt{6}}),(-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},-\frac{1}{\sqrt{6}}),(-\frac{1}{\sqrt{6}},-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}}\}\\ \tilde{m} &= \{(-\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}}),(\frac{1}{\sqrt{6}},-\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}}),(\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}},-\frac{2}{\sqrt{6}}\}. \end{aligned}$$

Dann folgt

$$\begin{aligned} f(x,y,z) &= e^{\frac{1}{3\sqrt{6}}} \quad \forall(x,y,z)\in\tilde{M}\\ f(x,y,z) &= e^{-\frac{1}{3\sqrt{6}}} \quad \forall(x,y,z)\in\tilde{m}. \end{aligned}$$

Also gilt

$$e^{-\frac{1}{3\sqrt{6}}}\le f(x,y,z)\le e^{\frac{1}{3\sqrt{6}}} \quad \forall(x,y,z)\in S^2\cap\mathcal{E}$$

und das Maximum und Minimum wird genau in den Punkten von $\tilde{M}$ bzw. $\tilde{m}$ angenommen.
AD (C): Sei $(x,y,z)\in S^2\cap\mathcal{E}_+$ ein lokales Extremum von $f$. Dann gilt entweder $x+y+z=0$ oder $x+y+z>0$. Im ersten Fall ist $(x,y,z)$ auch ein Extremum von $f$ auf $S^2\cap\mathcal{E}$; im zweiten Fall ist es auch ein

Extremum von $f$ auf $S^2$.
Wegen

$$M \cap S^2 \cap \mathcal{E}_+ = \{(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})\}$$

$$m \cap S^2 \cap \mathcal{E}_+ = \{(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}), (\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}), (-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})\}$$

folgt aus Teil (A) und (B)

$$\max_{(x,y,z)\in S^2\cap\mathcal{E}_+} f(x,y,z) = \max\{e^{\frac{1}{3\sqrt{3}}}, e^{\frac{1}{3\sqrt{6}}}\} = e^{\frac{1}{3\sqrt{3}}}$$

und

$$\min_{(x,y,z)\in S^2\cap\mathcal{E}_+} f(x,y,z) = \min\{e^{-\frac{1}{3\sqrt{3}}}, e^{-\frac{1}{3\sqrt{6}}}\} = e^{-\frac{1}{3\sqrt{3}}}$$

und das Maximum und Minimum wird in den Punkten von $M \cap S^2 \cap \mathcal{E}_+$ bzw. $m \cap S^2 \cap \mathcal{E}_+$ angenommen.

KAPITEL VIII

# Kurven und Kurvenintegrale

Im Mittelpunkt dieses Kapitels steht das Stammfunktionenproblem für Funktionen mehrerer Veränderlicher, d.h. die Frage, wann es zu einem gegebenen Vektorfeld $f \in C(U, \mathbb{R}^n), U \subset \mathbb{R}^n$, eine Funktion $F \in C^1(U, \mathbb{R})$ gibt mit $f = DF$. Um diese Frage zu beantworten, betrachten wir zunächst Kurven im $\mathbb{R}^n$ und Integrale von Vektorfeldern entlang solcher Kurven.
Anschließend wenden wir die gewonnenen Ergebnisse auf den Spezialfall komplexer Funktionen an und beweisen die zentralen Sätze der Funktionentheorie (Cauchyscher Integralsatz, Satz von Liouville, Residuensatz usw.).

## VIII.1. Kurven und ihre Länge

In diesem Paragraphen sei stets $I = [a, b]$ ein kompaktes, perfektes Intervall, $m \in \mathbb{N}$ und $n \in \mathbb{N}^*$, $\|\cdot\|$ bezeichne stets die euklidische Norm auf dem $\mathbb{R}^n$.

Wir knüpfen an Definition III.5.8 (S. 91) an.

Definition VIII.1.1. Eine Funktion $\gamma \in C^m(I, \mathbb{R}^n)$ heißt $C^m$-Weg in $\mathbb{R}^n$ mit Anfangspunkt $\gamma(a)$ und Endpunkt $\gamma(b)$. Der Weg heißt geschlossen, wenn gilt $\gamma(a) = \gamma(b)$. Die Menge

$$\operatorname{Spur}(\gamma) = \gamma(I)$$

heißt die Spur des Weges $\gamma$.

Bemerkung VIII.1.2. Verschiedene Wege können die gleiche Spur haben. Betrachte z.B. $\gamma_1, \gamma_2 : [0, 2\pi] \to \mathbb{R}^2$ mit

$$\gamma_1(t) = (\cos t, \sin t)$$
$$\gamma_2(t) = (\cos 2t, \sin 2t)$$

Dann ist

$$\operatorname{Spur}(\gamma_1) = \operatorname{Spur}(\gamma_2) = S^1.$$

Die Spur des Weges $\gamma$ sei ein Polygonzug im $\mathbb{R}^n$. Dann gibt es offensichtlich Zahlen $a = a_0 < a_1 < \ldots < a_k = b$, so dass $\gamma([a_{i-1}, a_i])$ für alle $1 \le i \le k$ eine Strecke ist. Die Länge des Polygonzuges ist offensichtlich $\sum_{i=1}^{k} \|\gamma(a_i) - \gamma(a_{i-1})\|$. Anschaulich können wir die Spur eines beliebigen Weges beliebig gut durch einen Polygonzug approximieren. Dies führt auf folgende Definition.

Definition VIII.1.3. Ein Weg $\gamma \in C(I, \mathbb{R}^n)$ heißt rektifizierbar, wenn gilt

$$L(\gamma) = \sup\Big\{\sum_{i=1}^{k} \|\gamma(a_{i-1}) - \gamma(a_i)\| : \mathcal{Z} = (a_0, \ldots, a_k) \text{ Zerlegung von } I\Big\} \\ < \infty.$$

In diesem Fall heißt $L(\gamma)$ die Länge von $\gamma$.

Bemerkung VIII.1.4. Es gibt stetige, nicht rektifizierbare Wege. Definiere z.B. $\gamma : [0, 1] \to \mathbb{R}$ durch

$$\gamma(t) = \begin{cases} t \cos \frac{\pi}{2t} & \text{für } 0 < t \le 1, \\ 0 & \text{für } t = 0. \end{cases}$$

Dann ist $\gamma \in C([0, 1], \mathbb{R})$. Für $k \in \mathbb{N}^*$ sei $\mathcal{Z}_k = (t_{0,k}, \ldots, t_{4k,k})$ mit $t_{0,k} = 0$ und

$$t_{j,k} = \frac{1}{4k + 1 - j} \qquad 1 \le j \le 4k.$$

Dann folgt

$$\begin{aligned} \gamma(t_{j,k}) &= t_{j,k} \cos \frac{\pi}{2}(4k + 1 - j) \\ &= \begin{cases} 0 & \text{für } 1 \le j \le 4k, j \text{ gerade}, \\ (-1)^l t_{j,k} & \text{für } 1 \le j \le 4k, j = 2l + 1. \end{cases} \end{aligned}$$

und somit

$$\begin{aligned} \sum_{j=1}^{4k} \|\gamma(t_{j-1,k}) - \gamma(t_{j,k})\| &= 2 \sum_{l=0}^{2k-1} \|\gamma(t_{2j+1,k})\| \\ &= 2 \sum_{l=0}^{2k} \frac{1}{4k - 2l} \\ &= \sum_{m=1}^{2k} \frac{1}{m} \\ &\underset{k\to\infty}{\longrightarrow} \infty. \end{aligned}$$

Satz VIII.1.5. *Sei $\gamma \in C^1(I, \mathbb{R}^n)$. Dann ist $\gamma$ rektifizierbar und*

$$\begin{aligned} L(\gamma) &= \int_a^b \|\gamma'(t)\| dt \\ &= \int_a^b \sqrt{\gamma_1'(t)^2 + \cdots + \gamma_n'(t)^2} dt. \end{aligned}$$

Beweis. Sei $\mathcal{Z} = (a_0, \ldots, a_k)$ eine Zerlegung von $I$. Dann folgt mit Satz VI.2.3 (S. 181)

$$\begin{aligned}\sum_{i=1}^{k} \|\gamma(a_i) - \gamma(a_{i-1})\| &= \sum_{i=1}^{k} \| \int_{a_{i-1}}^{a_i} \gamma'(t)dt\| \\ &\leq \sum_{i=1}^{k} \int_{a_{i-1}}^{a_i} \|\gamma'(t)\|dt \\ &= \int_a^b \|\gamma'(t)\|dt.\end{aligned}$$

Da $\mathcal{Z}$ beliebig war, folgt, dass $\gamma$ rektifizierbar ist und dass gilt

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

Da $I$ kompakt ist, ist $\gamma'$ gleichmäßig stetig auf $I$. Sei $\varepsilon > 0$ beliebig. Dann gibt es ein $k \in \mathbb{N}^*$ mit

$$\|\gamma'(t) - \gamma'(s)\| < \frac{\varepsilon}{2(b-a)} \quad \forall t, s \in I \text{ mit } |t-s| \leq \frac{b-a}{k}.$$

Damit folgt

$$\begin{aligned}&\int_a^b \|\gamma'(t)\|dt \\ &= \sum_{j=1}^{k} \int_{a+\frac{j-1}{k}(b-a)}^{a+\frac{j}{k}(b-a)} \|\gamma'(t)\|dt \\ &\leq \sum_{j=1}^{k} \Big\{ \int_{a+\frac{j-1}{k}(b-a)}^{a+\frac{j}{k}(b-a)} \|\gamma'(t) - \gamma'(a + \frac{2j-1}{2k}(b-a))\|dt \\ &\qquad + \|\gamma(a + \frac{j}{k}(b-a)) - \gamma(a + \frac{j-1}{k}(b-a)) \\ &\qquad\qquad - \gamma'(a + \frac{2j-1}{2k}(b-a)) \cdot \frac{b-a}{k}\| \\ &\qquad + \|\gamma(a + \frac{j-1}{k}(b-a)) - \gamma(a + \frac{j}{k}(b-a))\| \Big\} \\ &= \sum_{j=1}^{k} \|\gamma(a + \frac{j-1}{k}(b-a)) - \gamma(a + \frac{j}{k}(b-a))\| \\ &\quad + \sum_{j=1}^{k} \Big\{ \int_{a+\frac{j-1}{k}(b-a)}^{a+\frac{j}{k}(b-a)} \|\gamma'(t) - \gamma'(a + \frac{2j-1}{2k}(b-a))\|dt \\ &\qquad + \| \int_{a+\frac{j-1}{k}(b-a)}^{a+\frac{j}{k}(b-a)} [\gamma'(s) - \gamma'(a + \frac{2j-1}{2k}(b-a))]ds\| \Big\}\end{aligned}$$

$$\begin{aligned} &\le L(\gamma) + 2(b-a)\frac{\varepsilon}{2(b-a)} \\ &= L(\gamma) + \varepsilon. \end{aligned}$$

Da $\varepsilon$ beliebig war, folgt die Behauptung. □

Anschaulich stellt die Spur eines Weges eine „Kurve“ im $\mathbb{R}^n$ dar. Wie wir gesehen haben, kann dabei eine „Kurve“ durch verschiedene Wege beschrieben werden. Wir interessieren uns im Folgenden für Eigenschaften, die von der speziellen Darstellung der „Kurve“ unabhängig sind. Dazu benötigen wir folgende Definition.

Definition VIII.1.6. Seien $I$ und $\widetilde{I}$ kompakte, perfekte Intervalle und $\gamma \in C^m(I, \mathbb{R}^n)$, $\widetilde{\gamma} \in C^m(\widetilde{I}, \mathbb{R}^n)$ zwei $C^m$-Wege. Dann heißt $\widetilde{\gamma}$ eine $C^m$-Umparametrisierung von $\gamma$, kurz $\widetilde{\gamma} \sim_m \gamma$, wenn es einen streng monoton wachsenden $C^m$-Diffeomorphismus $\varphi$ von $I$ auf $\widetilde{I}$ gibt mit

$$\gamma = \widetilde{\gamma} \circ \varphi.$$

Bemerkung VIII.1.7. (1) Ist $\widetilde{\gamma}$ eine Umparametrisierung von $\gamma$, so haben beide Wege die gleichen Anfangs- und Endpunkte und die gleiche Spur.
(2) $\sim_m$ ist eine Äquivalenzrelation.
(3) Gilt $\gamma \sim_m \widetilde{\gamma}$, so gilt auch $\gamma \sim_k \widetilde{\gamma}$ für alle $0 \le k \le m$.

Satz VIII.1.8. *Seien $I$ und $\widetilde{I}$ kompakte, perfekte Intervalle und $\gamma \in C(I, \mathbb{R}^n)$, $\widetilde{\gamma} \in C(\widetilde{I}, \mathbb{R}^n)$ zwei stetige Wege mit $\gamma \sim_0 \widetilde{\gamma}$. Dann gilt*

$$L(\gamma) = L(\widetilde{\gamma}).$$

Beweis. Sei $I = [a, b]$, $\widetilde{I} = [\widetilde{a}, \widetilde{b}]$ und $\mathcal{Z} = (a_0, \ldots, a_k)$ eine Zerlegung von $I$. Dann ist $\widetilde{\mathcal{Z}} = (\widetilde{a}_0, \ldots, \widetilde{a}_k)$ mit $\widetilde{a}_j = \varphi(a_j)$ eine Zerlegung von $\widetilde{I}$, wobei $\varphi \in C(I, \widetilde{I})$ der Homöomorphismus aus der Definition von $\gamma \sim_0 \widetilde{\gamma}$ ist. Dann folgt

$$\begin{aligned} \sum_{i=1}^{k} \|\gamma(a_i) - \gamma(a_{i-1})\| &= \sum_{i=1}^{k} \|\widetilde{\gamma}(\varphi(a_i)) - \widetilde{\gamma}(\varphi(a_{i-1}))\| \\ &= \sum_{i=1}^{k} \|\widetilde{\gamma}(\widetilde{a}_i) - \widetilde{\gamma}(\widetilde{a}_{i-1})\| \\ &\le L(\widetilde{\gamma}) \end{aligned}$$

und somit

$$L(\gamma) \le L(\widetilde{\gamma}).$$

Vertauschen von $\gamma$ und $\widetilde{\gamma}$ liefert die Behauptung. □

Definition VIII.1.9. Unter einer $C^m$–Kurve $\Gamma$ in $\mathbb{R}^n$ verstehen wir eine Äquivalenzklasse von $C^m$-Wegen bzgl. $\sim_m$. Wir schreiben

$\Gamma : I \to \mathbb{R}^n$, $t \mapsto \gamma(t)$, wenn $\gamma$ ein Element der zu $\Gamma$ gehörenden Äquivalenzklasse ist. Die Kurve $\Gamma$ heißt REKTIFIZIERBAR, wenn ein (und damit alle) Element $\gamma$ aus der zugehörigen Äquivalenzklasse rektifizierbar ist. Die Zahl

$$L(\Gamma) = L(\gamma)$$

heißt dann die BOGENLÄNGE oder kurz LÄNGE von $\Gamma$.

BEMERKUNG VIII.1.10. (1) Wegen Satz VIII.1.8 ist die Länge einer rektifizierbaren Kurve wohldefiniert.
(2) Wegen Satz VIII.1.8 und Bemerkung VIII.1.4 gibt es stetige, nicht rektifizierbare Kurven (vgl. Beispiel VIII.1.12 (7)).
(3) Wegen Bemerkung VIII.1.7 hat jede Kurve eine eindeutige Orientierung, die durch die Orientierung der Wege aus der entsprechenden Äquivalenzklasse festgelegt ist.

Aus Satz VIII.1.5 folgt unmittelbar:

SATZ VIII.1.11. *Jede $C^1$-Kurve ist rektifizierbar.*

BEISPIEL VIII.1.12. (1) (GRAPH EINER $C^1$-FUNKTION) Sei $f \in C^1(I, \mathbb{R})$ und $\Gamma$ der Graph von $f$ in $\mathbb{R}^2$. Dann ist $\gamma : t \mapsto (t, f(t))$ eine Parametrisierung von $\Gamma$ und

$$L(\Gamma) = L(\gamma) = \int_a^b \sqrt{1 + f'(t)^2}\, dt.$$

(2) (EBENE KURVE IN POLARKOORDINATENDARSTELLUNG) Betrachte $\Gamma : I \to \mathbb{R}^2$, $t \mapsto \gamma(t) = (r(t)\cos t, r(t)\sin t)$ mit $r \in C^1(I, \mathbb{R})$. Dann ist

$$\begin{aligned} \|\gamma'(t)\| &= \|(r'(t)\cos t - r(t)\sin t, r'(t)\sin t + r(t)\cos t)\| \\ &= \sqrt{r'(t)^2 + r(t)^2} \end{aligned}$$

und

$$L(\Gamma) = \int_a^b \sqrt{r'(t)^2 + r(t)^2}\, dt.$$

(3) (KREISLINIE MIT RADIUS $R$) Dies ist ein Spezialfall von (2) mit $I = [0, 2\pi]$ und $r(t) = R$ für alle $t \in I$. Es ist

$$L(\Gamma) = 2\pi R.$$

(4) (ARCHIMEDISCHE SPIRALE) Dies ist ein Spezialfall von (2) mit $I = [0, b]$, $b > 0$, und $r(t) = t\rho$ für alle $t \in I$ mit $\rho > 0$ fest. Es ist

$$\begin{aligned} L(\Gamma) &= \int_0^b \sqrt{\rho^2 + \rho^2 t^2}\, dt \\ &= \rho \int_0^b \sqrt{1 + t^2}\, dt \end{aligned}$$

$$\begin{aligned}
&= \frac{\rho}{2}[t\sqrt{1+t^2} + \ln(t+\sqrt{1+t^2})]_{t=0}^{t=b} \\
&= \frac{\rho}{2}[b\sqrt{1+b^2} + \ln(b+\sqrt{1+b^2})] \\
&\underset{b\to+\infty}{\longrightarrow} +\infty.
\end{aligned}$$

(5) (Logarithmische Spirale) Dies ist ein Spezialfall von (2) mit $I = [a,b]$ und $r(t) = \rho e^{\lambda t}$ für alle $t \in I$ mit $\rho > 0$, $\lambda > 0$ fest. Es ist

$$\begin{aligned}
L(\Gamma) &= \int_a^b \sqrt{\rho^2 e^{2\lambda t} + \rho^2\lambda^2 e^{2\lambda t}}\, dt \\
&= \rho\sqrt{1+\lambda^2}\int_a^b e^{\lambda t} dt \\
&= \frac{\rho}{\lambda}\sqrt{1+\lambda^2}[e^{\lambda b} - e^{\lambda a}] \\
&\underset{a\to-\infty}{\longrightarrow} \frac{\rho}{\lambda}\sqrt{1+\lambda^2}e^{\lambda b}.
\end{aligned}$$

(6) (Schraubenlinie) Betrachte $\Gamma : I \to \mathbb{R}^3$, $t \mapsto \gamma(t) = (R\cos t, R\sin t, ht)$ mit $R > 0, h > 0$. $R$ heißt Radius; $h$ heißt Ganghöhe. Es ist

$$\begin{aligned}
\|\gamma'(t)\| &= \|(-R\sin t, R\cos t, h)\| \\
&= \sqrt{R^2+h^2}
\end{aligned}$$

und

$$L(\Gamma) = (b-a)\sqrt{R^2+h^2}.$$

(7) (nicht rektifizierbare stetige Kurve) Betrachte $\Gamma : [0,1] \to \mathbb{R}^2$, $t \mapsto \gamma(t) = (t\cos\frac{1}{t}, t\sin\frac{1}{t})$. Für $t > 0$ ist

$$\begin{aligned}
\|\gamma'(t)\| &= \|(\cos\frac{1}{t} + \frac{1}{t}\sin\frac{1}{t}, \sin\frac{1}{t} - \frac{1}{t}\cos\frac{1}{t})\| \\
&= \sqrt{1+\frac{1}{t^2}}.
\end{aligned}$$

Für $0 < a < 1$ folgt

$$\begin{aligned}
\int_a^1 \|\gamma'(t)\| &= \int_a^1 \sqrt{1+\frac{1}{t^2}}\, dt \\
&\geq \int_a^1 \frac{1}{t} dt \\
&= -\ln(a) \\
&\underset{a\to 0+}{\longrightarrow} +\infty.
\end{aligned}$$

## VIII.2. Tangente und Krümmung

In diesem Paragraphen sei wieder $I$ ein kompaktes, perfektes Intervall, $m \in \mathbb{N}$ und $n \in \mathbb{N}^*$, $\|\cdot\|$ bezeichne wieder die euklidische Norm auf dem $\mathbb{R}^n$.

Wir wollen den Begriff des Tangentenvektors einführen. Dazu benötigen wir:

Definition VIII.2.1. Sei $\gamma \in C^1(I, \mathbb{R}^n)$ ein $C^1$-Weg. Dann heißt $t_0 \in I$ ein regulärer Punkt von $\gamma$, wenn gilt $\gamma'(t_0) \neq 0$. $\gamma$ heißt regulärer Weg, wenn jedes $t \in I$ regulärer Punkt von $\gamma$ ist. Eine $C^1$-Kurve $\Gamma$ heißt regulär, wenn es in der entsprechenden Äquivalenzklasse einen regulären Weg gibt.

$C^1$-Umparametrisierungen beeinflussen nicht die Regularität eines Weges. Anders verhält es sich bei $C^0$-Umparametrisierungen. Dies zeigen die folgende Bemerkung und das folgende Beispiel.

Bemerkung VIII.2.2. Seien $I$ und $\widetilde{I}$ kompakte, perfekte Intervalle und $\gamma \in C^1(I, \mathbb{R}^n)$, $\widetilde{\gamma} \in C^1(\widetilde{I}, \mathbb{R}^n)$ Wege mit $\gamma \sim_1 \widetilde{\gamma}$. Dann ist $\gamma$ genau dann regulär, wenn $\widetilde{\gamma}$ regulär ist.

Beweis. Sei $\varphi \in C^1(I, \widetilde{I})$ ein streng wachsender $C^1$-Diffeomorphismus mit $\gamma = \widetilde{\gamma} \circ \varphi$. Dann ist

$$\gamma'(t) = \widetilde{\gamma}'(\varphi(t)) \cdot \varphi'(t) \quad \forall t \in I.$$

Wegen $\varphi'(t) > 0$ für alle $t \in I$ folgt hieraus die Behauptung. □

Beispiel VIII.2.3. Die Wege $\gamma \in C^1([-1, 1], \mathbb{R}^2)$ mit

$$\gamma(t) = (t^{\frac{4}{3}}, t)$$

und $\widetilde{\gamma} \in C^1([-1, 1], \mathbb{R}^2)$ mit

$$\widetilde{\gamma}(t) = (t^4, t^3)$$

haben die gleichen Anfangs- und Endpunkte und die gleiche Spur. $\gamma$ ist regulär, $\widetilde{\gamma}$ ist nicht regulär. Es ist $\gamma = \widetilde{\gamma} \circ \varphi$ mit

$$\varphi : [-1, 1] \to [-1, 1] \quad , \quad t \mapsto \sqrt[3]{t}.$$

$\varphi$ ist ein Homöomorphismus, aber kein $C^1$-Diffeomorphismus.

Definition VIII.2.4. Sei $\gamma \in C^1(I, \mathbb{R}^n)$ ein regulärer Weg. Dann heißt der Vektor

$$e_1(t) = \frac{\gamma'(t)}{\|\gamma'(t)\|} \quad \forall t \in I$$

der Tangenteneinheitsvektor an $\gamma$ im Punkt $\gamma(t)$.

Bemerkung VIII.2.5. (1) Sind $I$, $\widetilde{I}$ kompakte, perfekte Intervalle, $\gamma \in C^1(I,\mathbb{R}^n)$, $\widetilde{\gamma} \in C^1(\widetilde{I},\mathbb{R}^n)$ regulär und $\varphi \in C^1(I,\widetilde{I})$ ein streng wachsender $C^1$-Diffeomorphismus mit $\gamma = \widetilde{\gamma} \circ \varphi$, so ist

$$\begin{aligned} e_1(t) &= \frac{\gamma'(t)}{\|\gamma'(t)\|} \\ &= \frac{\widetilde{\gamma}'(\varphi(t)) \cdot \varphi'(t)}{\|\widetilde{\gamma}'(\varphi(t)) \cdot \varphi'(t)\|} \\ &= \frac{\widetilde{\gamma}'(\varphi(t))}{\|\widetilde{\gamma}'(\varphi(t))\|} \\ &= \widetilde{e}_1(\varphi(t)) \quad \forall t \in I. \end{aligned}$$

Der Tangenteneinheitsvektor ist also invariant unter Parameterwechseln.
(2) Es ist

$$\gamma(t_0 + h) = \gamma(t_0) + h\|\gamma'(t_0)\|e_1(t_0) + hr(t_0, h) \quad \forall t_0 \in I, h \in \mathbb{R}$$

mit

$$\lim_{h\to 0} \|r(t_0, h)\| = 0.$$

Der Weg $\gamma$ wird also im Punkt $\gamma(t_0)$ in erster Ordnung von der Tangente

$$s \mapsto \gamma(t_0) + se_1(t_0)$$

approximiert.

Definition VIII.2.6. $\gamma \in C^1(I,\mathbb{R}^n)$ heißt nach der Bogenlänge parametrisiert, wenn gilt

$$\|\gamma'(t)\| = 1 \quad \forall t \in I.$$

Wir kennzeichnen die Bogenlänge stets durch $s$ und die Ableitung nach der Bogenlänge durch $\dot{\gamma}$.

Bemerkung VIII.2.7. (1) Ist $\gamma$ nach der Bogenlänge parametrisiert, so ist $\gamma$ regulär.
(2) Jede reguläre Kurve $\Gamma$ kann nach der Bogenlänge parametrisiert werden.

Beweis. ad (1): Ist klar.
ad (2): Sei $\gamma \in C^1(I,\mathbb{R}^n)$ eine reguläre Parametrisierung von $\Gamma$. Sei $L = L(\gamma)$ und $\widetilde{I} = [0, L]$. Definiere $\varphi : I \to \widetilde{I}$ durch

$$\varphi(t) = \int_a^t \|\gamma'(\tau)\|d\tau.$$

Dann ist $\varphi$ ein streng wachsender $C^1$-Diffeomorphismus. Für $\widetilde{\gamma} = \gamma \circ \varphi^{-1} \in C^1(\widetilde{I},\mathbb{R}^n)$ folgt

$$\dot{\widetilde{\gamma}}(s) = \frac{d}{ds}\widetilde{\gamma}(s)$$

$$= \gamma'(\varphi^{-1}(s)) \cdot \frac{1}{\varphi'(\varphi^{-1}(s))}$$
$$= \frac{\gamma'(\varphi^{-1}(s))}{\|\gamma'(\varphi^{-1}(s))\|} \quad \forall s \in \widetilde{I}$$

und somit

$$\|\dot{\tilde{\gamma}}(s)\| = 1 \quad \forall s \in \tilde{I}.$$

□

Satz VIII.2.8. *Sei $\gamma \in C^2(I, \mathbb{R}^n)$ nach der Bogenlänge parametrisiert. Dann gilt*

$$\langle \dot{\gamma}(s), \ddot{\gamma}(s) \rangle = 0 \quad \forall s \in I,$$

*wobei*

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

*das euklidische Skalarprodukt bezeichnet.*

Beweis. Für alle $s \in I$ gilt

$$\|\dot{\gamma}(s)\|^2 = \sum_{i=1}^{n} \dot{\gamma}_i(s)^2 = 1.$$

Durch Differentiation folgt

$$0 = \frac{d}{ds} \|\dot{\gamma}(s)\|^2 = 2 \sum_{i=1}^{n} \dot{\gamma}_i(s) \ddot{\gamma}_i(s) = 2 \langle \dot{\gamma}(s), \ddot{\gamma}(s) \rangle.$$

□

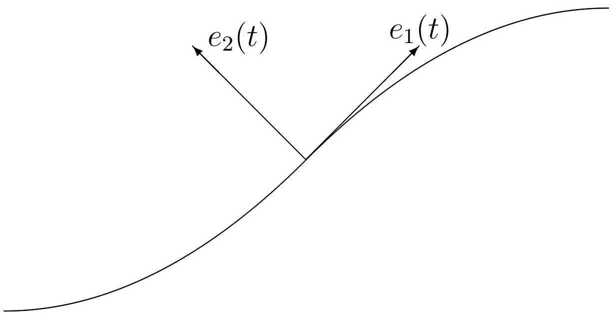


Abbildung VIII.2.1. Vektoren $e_1(t)$ und $e_2(t)$

Für den Rest des Paragraphen betrachten wir ebene Kurven, d.h., den Fall $n = 2$ (vgl. Abbildung VIII.2.1). Zu $e_1(t)$ gibt es dann einen eindeutig bestimmten Vektor $e_2(t)$ mit den Eigenschaften

$$\begin{aligned} \|e_2(t)\| &= 1, \\ \langle e_1(t), e_2(t) \rangle &= 0, \\ \det[e_1(t), e_2(t)] &= 1. \end{aligned} \tag{VIII.2.1}$$

DEFINITION VIII.2.9. Sei $\gamma \in C^2(I, \mathbb{R}^2)$ eine nach der Bogenlänge parametrisierte ebene Kurve. Sei $e_1(s)$ der Tangenteneinheitsvektor zu $\gamma$ und $e_2(s)$ der durch Gleichung (VIII.2.1) bestimmte Vektor. Dann ist

$$\ddot{\gamma}(s) = \kappa(s) e_2(s) \quad \forall s \in I.$$

Die Größe $\kappa(s)$ heißt die KRÜMMUNG der Kurve im Punkt $\gamma(s)$.

BEMERKUNG VIII.2.10. Ist $\kappa(s) > 0$ bzw. $\kappa(s) < 0$, so dreht sich der Tangenteneinheitsvektor $e_1(s)$ bei Durchlaufen der Kurve in mathematisch positiver Richtung bzw. mathematisch negativer Richtung.

Der folgende Satz beschreibt einen fundamentalen Zusammenhang zwischen den Vektoren $e_1$ und $e_2$ und erlaubt die Berechnung der Krümmung von Wegen, die nicht nach der Bogenlänge parametrisiert sind.

SATZ VIII.2.11. (1) *Sei $\gamma \in C^2(I, \mathbb{R}^2)$ nach der Bogenlänge parametrisiert. Dann gelten die* FRENETSCHEN FORMELN

$$\begin{aligned} \dot{e}_1 &= \kappa e_2, \\ \dot{e}_2 &= -\kappa e_1. \end{aligned}$$

(2) *Sei $\gamma \in C^2(I, \mathbb{R}^2)$ regulär. Dann gilt für die Krümmung*

$$\kappa(t) = \frac{x'(t) y''(t) - x''(t) y'(t)}{[x'(t)^2 + y'(t)^2]^{3/2}} \quad \forall t \in I,$$

*wobei $\gamma(t) = (x(t), y(t))$ für alle $t \in I$ ist.*

BEWEIS. AD (1): Die erste Formel folgt direkt aus der Definition der Krümmung. Aus Gleichung (VIII.2.1) folgt durch Differentiation

$$\begin{aligned} & 0 = \langle e_2(s), \dot{e}_2(s) \rangle \quad \forall s \in I \\ \Longrightarrow \quad & \dot{e}_2(s) = \alpha(s) e_1(s) \quad \forall s \in I \end{aligned}$$

und

$$\begin{aligned} 0 &= \langle \dot{e}_1(s), e_2(s) \rangle + \langle e_1(s), \dot{e}_2(s) \rangle \\ &= \langle \kappa(s) e_2(s), e_2(s) \rangle + \langle e_1(s), \alpha(s) e_1(s) \rangle \\ &= \kappa(s) + \alpha(s). \end{aligned}$$

Hieraus folgt die zweite Formel.
AD (2): Seien $\widetilde{\gamma}$ und $\varphi$ wie im Beweis von Bemerkung VIII.2.7 (2). Setze $\psi = \varphi^{-1}$. Dann folgt

$$\begin{aligned} \dot{\widetilde{\gamma}}(s) &= \gamma'(\psi(s)) \cdot \dot{\psi}(s) \\ \ddot{\widetilde{\gamma}}(s) &= \gamma''(\psi(s)) \cdot \dot{\psi}(s)^2 + \gamma'(\psi(s)) \ddot{\psi}(s) \\ \dot{\psi}(s) &= \|\gamma'(\psi(s))\|^{-1}. \end{aligned}$$

Aus der Definition der Krümmung und aus Gleichung (VIII.2.1) folgt

$$\begin{aligned}\kappa(s) &= \det(\dot{\tilde{\gamma}}(s), \ddot{\tilde{\gamma}}(s)) \\ &= \dot{\psi}(s)^3 \det(\gamma'(\psi(s)), \gamma''(\psi(s))) \\ &\quad + \dot{\psi}(s)\ddot{\psi}(s) \det(\gamma'(\psi(s)), \gamma'(\psi(s))) \\ &= \dot{\psi}(s)^3 \det(\gamma'(\psi(s)), \gamma''(\psi(s))). \quad \forall s \in \widetilde{I}.\end{aligned}$$

Also ist

$$\begin{aligned}\kappa(t) &= \|\gamma'(t)\|^{-3} \det[\gamma'(t), \gamma''(t)] \\ &= \frac{x'(t)y''(t) - x''(t)y'(t)}{[x'(t)^2 + y'(t)^2]^{3/2}} \quad \forall t \in I.\end{aligned}$$

□

Beispiel VIII.2.12. (1) (Graph einer $C^2$-Funktion) Sei $f \in C^2(I, \mathbb{R})$ und $\Gamma$ der Graph von $f$. Die Krümmung ist gemäß Satz VIII.2.11 (2)

$$\kappa(t) = \frac{f''(t)}{[1 + f'(t)^2]^{3/2}}.$$

(2) (Ebene Kurve in Polarkoordinatendarstellung) Für die Krümmung ergibt sich

$$\kappa(t) = \frac{r(t)^2 + 2r'(t)^2 - r(t)r''(t)}{[r(t)^2 + r'(t)^2]^{3/2}}.$$

(3) (Kreislinie mit Radius $R$) Aus (2) folgt mit $r(t) = R$

$$\kappa = \frac{1}{R}.$$

(4) (Archimedische Spirale) Aus (2) folgt mit $r(t) = \rho t$

$$\kappa(t) = \frac{2 + t^2}{\rho(1 + t^2)^{3/2}} \quad \forall t \in [0, b].$$

(5) (Logarithmische Spirale) Aus (2) folgt mit $r(t) = \rho e^{\lambda t}$

$$\kappa(t) = \frac{1}{\rho e^{\lambda t}\sqrt{1 + \lambda^2}}.$$

Der folgende Satz charakterisiert die ebenen Kurven konstanter Krümmung.

Satz VIII.2.13. *Sei $\gamma \in C^2(I, \mathbb{R})$ regulär. Dann gilt:*

(1) *$\gamma$ ist die Parametrisierung einer Geraden genau dann, wenn gilt*

$$\kappa(t) = 0 \quad \forall t \in I.$$

(2) *$\gamma$ ist die Parametrisierung eines Kreisbogens, d.h. es gibt ein $x_0 \in \mathbb{R}^2$ und ein $R \in \mathbb{R}_+^*$ mit*

$$\|\gamma(t) - x_0\| = R \quad \forall t \in I$$

*genau dann, wenn gilt*

$$|\kappa(t)| = \frac{1}{R} \quad \forall t \in I.$$

Beweis. ad (1): O.E. ist $\gamma$ nach der Bogenlänge parametrisiert. Dann folgt

$$\begin{aligned}
&\kappa(s) = 0 \quad \forall s \in I \\
\iff &\ddot{\gamma}(s) = 0 \quad \forall s \in I \\
\iff &\exists b \in \mathbb{R}^2 : \|b\| = 1 \text{ und } \dot{\gamma}(s) = b \quad \forall s \in I \\
\iff &\exists a, b \in \mathbb{R}^2 : \|b\| = 1 \text{ und } \gamma(s) = sb + a \quad \forall s \in I.
\end{aligned}$$

ad (2): O.E. ist $\gamma$ wieder nach der Bogenlänge parametrisiert.
„$\Longrightarrow$“: Aus $\|\gamma(s) - x_0\|^2 = R^2$ für alle $s \in I$ folgt

$$\langle \gamma(s) - x_0, \dot{\gamma}(s) \rangle = 0$$

Also ist

$$\gamma(s) - x_0 = \varepsilon R e_2(s) \quad \text{mit } \varepsilon \in \{-1, 1\}.$$

Aus den Frenetschen Formeln folgt

$$e_1(s) = \dot{\gamma}(s) = \varepsilon R \dot{e}_2(s) = -\varepsilon R \kappa e_1(s)$$

und somit

$$\kappa(s) = \frac{-\varepsilon}{R} \quad \forall s \in I.$$

„$\Longleftarrow$“: Sei $\varepsilon \in \{-1, 1\}$, so dass gilt

$$\kappa(s) = \frac{\varepsilon}{R} \quad \forall s \in I.$$

Aus den Frenetschen Formeln folgt

$$\begin{aligned}
&(\gamma + \varepsilon R e_2)^{\cdot} = e_1 + \varepsilon R \dot{e}_2 = 0 \quad \forall s \in I \\
\iff &\gamma + \varepsilon R e_2 = x_0 \quad \text{für ein } x_0 \in \mathbb{R}^2 \\
\iff &\|\gamma - x_0\| = |\varepsilon| R \|e_2\| = R \quad \forall s \in I.
\end{aligned}$$

□

## VIII.3. Kurvenintegrale

In diesem Paragraphen wenden wir uns dem Stammfunktionenproblem für Funktionen mehrerer Veränderlicher zu. Dabei bezeichnet $G \subset \mathbb{R}^n$, $G \neq \emptyset$, $n \in \mathbb{N}^*$, stets ein Gebiet, $I \subset \mathbb{R}$ ein kompaktes, perfektes Intervall sowie $\| \cdot \|$ und $\langle \cdot, \cdot \rangle$ die euklidische Norm bzw. das euklidische Skalarprodukt auf $\mathbb{R}^n$.

Definition VIII.3.1. Seien $f \in C(G, \mathbb{R}^n)$ und $\gamma \in C^1(I, \mathbb{R}^n)$ mit $\gamma(I) \subset G$. Dann ist das Wegintegral von $f$ längs $\gamma$ definiert durch

$$\int_\gamma f dx = \int_a^b \langle f \circ \gamma(t), \gamma'(t) \rangle dt.$$

Bemerkung VIII.3.2. (1) Physikalisch beschreibt $\displaystyle\int_\gamma f dx$ die Arbeit, die aufgebracht werden muss, um einen Körper in dem Kraftfeld $f$ längs des Weges $\gamma$ von $\gamma(a)$ nach $\gamma(b)$ zu bewegen.
(2) Seien $I, \widetilde{I}$ kompakte perfekte Intervalle, $\gamma \in C(I, \mathbb{R}^n)$, $\widetilde{\gamma} \in C(\widetilde{I}, \mathbb{R}^n)$ zwei Wege mit $\gamma(I) \subset G$, $\widetilde{\gamma}(\widetilde{I}) \subset G$ und $\varphi \in C^1(I, \widetilde{I})$ ein $C^1$-Diffeomorphismus mit $\gamma = \widetilde{\gamma} \circ \varphi$. Dann folgt

$$\begin{aligned}\int_\gamma f dx &= \int_a^b \langle f(\widetilde{\gamma} \circ \varphi(t)), \widetilde{\gamma}' \circ \varphi(t) \cdot \varphi'(t) \rangle dt \\ &= \int_{\varphi(a)}^{\varphi(b)} \langle f \circ \widetilde{\gamma}(s), \widetilde{\gamma}'(s) \rangle ds \\ &= \int_{\widetilde{\gamma}} f dx.\end{aligned}$$

D.h., das Wegintegral hängt nicht von der Parametrisierung des Weges ab.

Wegen Bemerkung VIII.3.2 (2) ist folgende Definition sinnvoll.

Definition VIII.3.3. Sei $\Gamma$ eine $C^1$-Kurve, die ganz in $G$ verläuft, und $f \in C(G, \mathbb{R}^n)$. Dann ist das Kurvenintegral von $f$ längs $\Gamma$ definiert durch

$$\int_\Gamma f dx = \int_\gamma f dx,$$

wobei $\gamma$ ein $C^1$-Weg aus der zu $\Gamma$ gehörenden Äquivalenzklasse ist.

Beispiel VIII.3.4. (1) Sei $G = \mathbb{R}^2 \backslash \{0\}$, $\Gamma = \partial B(0, R)$ mit $R \in \mathbb{R}_+^*$ und

$$f(x, y) = (-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2}).$$

Sei $\gamma : [0, 2\pi] \to \mathbb{R}^2$ definiert durch

$$t \mapsto (R \cos t, R \sin t).$$

Dann folgt

$$\begin{aligned}\int_\Gamma f dx &= \int_\gamma f dx \\ &= \int_0^{2\pi} \langle (-\frac{1}{R} \sin t, \frac{1}{R} \cos t), (-R \sin t, R \cos t) \rangle dt \\ &= \int_0^{2\pi} dt\end{aligned}$$

$$= 2\pi.$$

(2) Sei $f = DF$ mit $F \in C^1(G, \mathbb{R})$ und $\gamma$ ein beliebiger $C^1$-Weg, der ganz in $G$ verläuft. Dann ist

$$\begin{aligned}\int_\gamma f dx &= \int_a^b \langle f \circ \gamma(t), \gamma'(t) \rangle dt \\ &= \int_a^b \sum_{i=1}^n D_i F(\gamma(t)) \gamma_i'(t) dt \\ &= F(\gamma(b)) - F(\gamma(a)).\end{aligned}$$

D.h., $\displaystyle\int_\gamma f dx$ hängt nur von dem Anfangs- und Endpunkt von $\gamma$ ab. Insbesondere ist $\displaystyle\int_\gamma f dx = 0$ für jeden geschlossenen $C^1$-Weg, der ganz in $G$ verläuft.

Für unsere Zwecke müssen wir den Begriff des Kurvenintegrals noch erweitern, so dass wir z.B. auch entlang von Polygonzügen integrieren können.

Definition VIII.3.5. (1) Sei $\gamma \in C(I, \mathbb{R}^n)$. Dann heißt $\gamma^- \in C(I, \mathbb{R}^n)$ mit

$$\gamma^-(t) = \gamma(b + a - t) \quad \forall t \in I$$

der zu $\gamma$ inverse Weg. $\gamma$ stellt die Kurve $\Gamma$ dar, $\gamma^-$ die zu $\Gamma$ inverse Kurve $-\Gamma$.
(2) Seien $\gamma_i \in C(I_i, \mathbb{R}^n)$, $I_i = [a_i, b_i], i = 1, 2$, mit $\gamma_1(b_1) = \gamma_2(a_2)$. Dann heißt $\gamma_1 \oplus \gamma_2 \in C([0, 1], \mathbb{R}^n)$ mit

$$(\gamma_1 \oplus \gamma_2)(t) = \begin{cases} \gamma_1(a_1 + 2t(b_1 - a_1)) & \text{für } 0 \le t \le \frac{1}{2}, \\ \gamma_2(a_2 + (2t - 1)(b_2 - a_2)) & \text{für } \frac{1}{2} \le t \le 1, \end{cases}$$

der zugehörige Summenweg. Stellen $\gamma_1, \gamma_2$ die Kurven $\Gamma_1$ und $\Gamma_2$ dar, so stellt $\gamma_1 \oplus \gamma_2$ die Kurve $\Gamma_1 + \Gamma_2$ dar.
(3) $\gamma \in C(I, \mathbb{R}^n)$ heißt stückweise $m$-mal stetig differenzierbar, wenn es Zahlen $a = a_0 < a_1 < \ldots < a_k = b$ gibt mit

$$\gamma_i = \gamma|_{[a_{i-1}, a_i]} \in C^m([a_{i-1}, a_i], \mathbb{R}^n) \quad \forall 1 \le i \le k.$$

Dann ist $\gamma = \gamma_1 \oplus \ldots \oplus \gamma_k$. Die durch $\gamma$ dargestellte Kurve $\Gamma$ heißt entsprechend stückweise $m$-mal stetig differenzierbar.
(4) Sei $f \in C(G, \mathbb{R}^n)$ und $\Gamma = \Gamma_1 + \ldots + \Gamma_k$ eine stückweise stetig differenzierbare Kurve, die ganz in $G$ verläuft. Dann ist das Kurvenintegral von $f$ längs $\Gamma$ definiert durch

$$\int_\Gamma f dx = \int_{\Gamma_1} f dx + \ldots + \int_{\Gamma_k} f dx.$$

Das Kurvenintegral hat folgende Eigenschaften:

SATZ VIII.3.6. *Seien $f, f_1, f_2 \in C(G, \mathbb{R}^n)$, $\lambda_1, \lambda_2 \in \mathbb{R}$ und $\Gamma, \Gamma_1, \Gamma_2$ stückweise stetig differenzierbare Kurven, die ganz in $G$ verlaufen. Dann gilt*

(LINEARITÄT) $$\int_\Gamma (\lambda_1 f_1 + \lambda_2 f_2) dx = \lambda_1 \int_\Gamma f_1 dx + \lambda_2 \int_\Gamma f_2 dx$$

(ADDITIVITÄT) $$\int_{\Gamma_1 + \Gamma_2} f dx = \int_{\Gamma_1} f + \int_{\Gamma_2} f dx$$

(ORIENTIERUNG) $$\int_\Gamma f dx = - \int_{-\Gamma} f dx$$

$$|\int_\Gamma f dx| \leq \max_{x \in \Gamma} \|f(x)\| L(\Gamma).$$

BEWEIS. AD (1), (2): Folgen direkt aus der Definition des Kurvenintegrals und den Eigenschaften des (Riemann-) Integrals.
AD (3): O.E. ist $\Gamma$ eine $C^1$-Kurve und $\gamma$ ein $C^1$-Weg aus der entsprechenden Äquivalenzklasse. Dann gilt

$$\begin{aligned}
\int_\Gamma f dx &= \int_\gamma f dx \\
&= \int_a^b \langle f \circ \gamma(t), \gamma'(t) \rangle dt \\
&= - \int_a^b \langle f \circ \gamma^-(a + b - t), \gamma^{-'}(a + b - t) \rangle dt \\
&= \int_b^a \langle f \circ \gamma^-(s), \gamma^{-'}(s) \rangle ds \quad (s = a + b - t) \\
&= - \int_a^b \langle f \circ \gamma^-(s), \gamma^{-'}(s) \rangle ds \\
&= - \int_{\gamma^-} f dx \\
&= - \int_{-\Gamma} f dx.
\end{aligned}$$

AD (4): O.E. wird $\Gamma$ von dem $C^1$-Weg $\gamma$ dargestellt, sonst gehen wir zu den Teilwegen über. Dann folgt

$$\begin{aligned}
|\int_\Gamma f dx| &= |\int_\gamma f dx| \\
&\leq \int_a^b \|f \circ \gamma(t)\| \|\gamma'(t)\| dt \\
&\leq L(\gamma) \max_{x \in \Gamma} \|f(x)\|.
\end{aligned}$$

□

Wir kommen nun zu der eigentlichen Fragestellung dieses Paragraphen: Wann gibt es zu gegebenem $f \in C(G, \mathbb{R}^n)$ ein $F \in C^1(G, \mathbb{R})$ mit $f = DF$?

Definition VIII.3.7. Eine Funktion $f \in C(G, \mathbb{R}^n)$ heißt Gradientenfeld, wenn es ein $F \in C^1(G, \mathbb{R})$ gibt mit

$$f = DF.$$

Jedes solche $F$ heißt eine Stammfunktion von $f$.

Bemerkung VIII.3.8. (1) Ist $f$ ein Gradientenfeld, so ist die Stammfunktion von $f$ wegen Satz VII.3.8(2) (S. 229) bis auf eine additive Konstante eindeutig bestimmt.
(2) Ist $f \in C^1(G, \mathbb{R}^n)$ ein Gradientenfeld, so ist wegen Satz VII.4.18(2) (S. 237) die Funktionalmatrix $Df$ symmetrisch.

Satz VIII.3.9. *Sei $f \in C(G, \mathbb{R}^n)$. Dann sind folgende Aussagen äquivalent:*

(1) *$f$ ist Gradientenfeld.*
(2) *$\int_\Gamma f dx = 0$ für jede geschlossene, stückweise $C^1$-Kurve, die ganz in $G$ verläuft.*

Beweis. (1) $\Longrightarrow$ (2): Folgt direkt aus Beispiel VIII.3.4 (2) angewandt auf die $C^1$-Teilkurven von $\Gamma$.
(2) $\Longrightarrow$ (1): Sei $x_0 \in G$ beliebig. Aus dem Beweis von Satz III.5.14 (S. 92) folgt, dass es zu jedem $x \in G$ einen Weg $\gamma_x$ gibt, der stückweise stetig differenzierbar ist, ganz in $G$ verläuft und Anfangspunkt $x_0$ und Endpunkt $x$ hat. Definiere

$$F(x) = \int_{\gamma_x} f dx.$$

Aus (2) folgt, dass die Definition von $F$ nicht von der Wahl des Weges $\gamma_x$ abhängt. Sei nun $x \in G$ und $\varepsilon \in \mathbb{R}_+^*$ mit $\overline{B(x, \varepsilon)} \subset G$. Für $i \in \mathbb{N}_n^*$ und $h \in (-\varepsilon, \varepsilon)$ folgt dann aus (2) und Satz VIII.3.6

$$F(x + he_i) - F(x) = \int_{[x, x+he_i]} f dx,$$

wobei $[x, x+he_i]$ die Strecke von $x$ nach $x+he_i$ bezeichnet. Damit folgt

$$\begin{aligned}
|\frac{1}{h}[F(x + he_i) - F(x)] - f_i(x)| &= |\frac{1}{h} \int_0^1 \langle f(x + the_i), he_i \rangle dt - f_i(x)| \\
&= | \int_0^1 [f_i(x + the_i) - f_i(x)] dt| \\
&\le \max_{y \in \overline{B(x, |h|)}} |f_i(y) - f_i(x)| \\
&\underset{h \to 0}{\longrightarrow} 0.
\end{aligned}$$

Zusammen mit dem Differenzierbarkeitskriterium Satz VII.2.13 (S. 221) folgt hieraus, dass $F$ eine Stammfunktion von $f$ ist. □

Satz VIII.3.9 löst unser Problem vollständig und erlaubt auch mit der im Beweis benutzten Technik die praktische Berechnung einer Stammfunktion. Umgekehrt ist aber die Bedingung (2) in der Praxis nur schwer nachprüfbar. Wir möchten im Folgenden ein einfacheres Kriterium herleiten. Und zwar möchten wir zeigen, dass unter einer zusätzlichen Voraussetzung an $G$ ein $f \in C^1(G, \mathbb{R}^n)$ genau dann eine Stammfunktion besitzt, wenn $Df$ symmetrisch ist. Wegen Bemerkung VIII.3.8 (2) ist letzteres eine notwendige Bedingung. Beispiel VIII.3.4 (1) zeigt, dass sie i. a. aber nicht hinreichend ist.

Satz VIII.3.10 (Lemma von Goursat). *Sei $G$ konvex, $f \in C^1(G, \mathbb{R}^n)$ und $Df$ symmetrisch. Für je drei paarweise verschiedene Punkte $x_1, x_2, x_3 \in G$ bezeichne $[x_1, x_2, x_3]$ den Streckenzug $[x_1, x_2] \oplus [x_2, x_3] \oplus [x_3, x_1]$, wobei $[x_i, x_j]$ die Strecke von $x_i$ nach $x_j$ ist. Dann gilt*

$$\int_{[x_1,x_2,x_3]} f dx = 0 \quad \forall x_1, x_2, x_3 \in G \; x_1 \neq x_2 \neq x_3.$$

Beweis. Seien $x_1, x_2, x_3$ beliebige, im Folgenden feste, paarweise verschiedene Punkte in $G$. Wir definieren drei Folgen $(x_{i,m})_{m\in\mathbb{N}}$, $1 \leq i \leq 3$, wie folgt:

(1) $x_{i,0} = x_i \quad 1 \leq i \leq 3$.

(2) $x_{i,m}$ seien bekannt, sei

$$x'_{i,m} = \frac{1}{2}(x_{i,m} + x_{i+1,m}), \quad 1 \leq i \leq 3,$$

wobei die Indizes modulo 3 zu rechnen sind. Dann bestimmen wir $x_{i,m+1} \in \{x_{i,m}, x'_{i+1,m}, x'_{i+2,m}\}$, $1 \leq i \leq 3$, so dass gilt

$$\begin{aligned}
&|\int_{[x_{1,m+1},x_{2,m+1},x_{3,m+1}]} f dx| \\
&= \max\Big\{|\int_{[x_{1,m},x'_{1,m},x'_{3,m}]} f dx|, |\int_{[x'_{1,m},x_{2,m},x'_{2,m}]} f dx|, \\
&\qquad |\int_{[x'_{2,m},x_{3,m},x'_{3,m}]} f dx|, |\int_{[x'_{1,m},x'_{2,m},x'_{3,m}]} f dx|\Big\}.
\end{aligned}$$

Wegen

$$\begin{aligned}
&\int_{[x_{1,m},x_{2,m},x_{3,m}]} f dx \\
&= \int_{[x_{1,m},x'_{1,m},x'_{3,m}]} f dx + \int_{[x'_{1,m},x_{2,m},x'_{2,m}]} f dx \\
&\quad + \int_{[x'_{3,m},x'_{2,m},x_{3,m}]} f dx + \int_{[x'_{1,m},x'_{2,m},x'_{3,m}]} f dx
\end{aligned}$$

folgt

$$|\int_{[x_{1,m},x_{2,m},x_{3,m}]} f dx| \leq 4| \int_{[x_{1,m+1},x_{2,m+1},x_{3,m+1}]} f dx|$$

und somit durch Induktion

$$|\int_{[x_1,x_2,x_3]} f dx| \leq 4^m| \int_{[x_{1,m},x_{2,m},x_{3,m}]} f dx| \quad \forall m \in \mathbb{N}.$$

Offensichtlich gilt für die Länge der Wege

$$L([x_{1,m+1}, x_{2,m+1}, x_{3.m+1}]) = \frac{1}{2} L([x_{1,m}, x_{2,m}, x_{3,m}]) \quad \forall m \in \mathbb{N}$$

und damit

$$L([x_{1,m}, x_{2,m}, x_{3,m}]) = 2^{-m} L([x_1, x_2, x_3]) \quad \forall m \in \mathbb{N}.$$

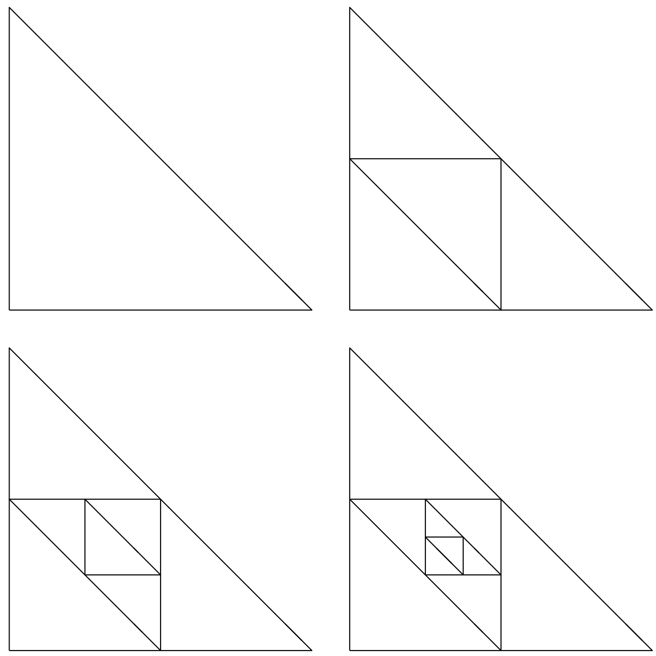

Abbildung VIII.3.1. Beispiel für Dreiecke $\Delta_m, \ldots, \Delta_{m+3}$

Bezeichne mit $\Delta_m$ das abgeschlossene Dreieck mit den Eckpunkten $x_{1,m}, x_{2,m}, x_{3,m}$ (vgl. Abb. VIII.3.1). Dann folgt

$$\ldots \subset \Delta_{m+1} \subset \Delta_m \subset \Delta_{m-1} \subset \ldots \subset \Delta_0 \subset G$$

und

$$\text{Fläche } (\Delta_{m+1}) = \frac{1}{4} \text{ Fläche } (\Delta_m).$$

Hieraus folgt

$$\bigcap_{m \in \mathbb{N}} \Delta_m = \{x^*\}$$

für ein $x^* \in G$ (Beweis: Übungsaufgabe!). Für $h \in \mathbb{R}^n$ mit $\|h\|$ hinreichend klein, gilt wegen der Differenzierbarkeit von $f$

$$f(x^* + h) = f(x^*) + Df(x^*)h + r(x^*, h)$$

mit

$$\lim_{\|h\| \to 0} \frac{1}{\|h\|} \|r(x^*, h)\| = 0.$$

Sei $\varepsilon > 0$ beliebig. Dann gibt es ein $\delta > 0$ mit $\overline{B(x^*, \delta)} \subset G$ und

$$\sup_{\substack{h \in \mathbb{R}^n \\ \|h\| \le \delta}} \frac{1}{\|h\|} \|r(x^*, h)\| \le \varepsilon.$$

Weiter gibt es ein $m \in \mathbb{N}$ mit $\Delta_m \subset \overline{B(x^*, \delta)}$. Da die Funktion

$$h \mapsto f(x^*) + Df(x^*)h$$

wegen der Symmetrie von $Df(x^*)$ die Stammfunktion

$$h \mapsto \langle f(x^*), h \rangle + \frac{1}{2} \langle h, Df(x^*)h \rangle$$

hat, folgt

$$\begin{aligned}
&|\int_{[x_1, x_2, x_3]} f dx| \\
&\le 4^m |\int_{[x_{1,m}, x_{2,m}, x_{3,m}]} [f(x^*) + Df(x^*)(x - x^*) + r(x^*, x - x^*)] dx| \\
&= 4^m |\int_{[x_{1,m}, x_{2,m}, x_{3,m}]} r(x^*, x - x^*) dx| \\
&\le 4^m L([x_{1,m}, x_{2,m}, x_{3,m}]) \max_{y \in [x_{1,m}, x_{2,m}, x_{3,m}]} \|r(x^*, y - x^*)\| \\
&\le \varepsilon 4^m L([x_{1,m}, x_{2,m}, x_{3,m}]) \max_{y \in [x_{1,m}, x_{2,m}, x_{3,m}]} \|y - x^*\| \\
&\le \varepsilon 4^m L([x_{1,m}, x_{2,m}, x_{3,m}])^2 \\
&= \varepsilon L([x_1, x_2, x_3])^2.
\end{aligned}$$

Da $\varepsilon$ beliebig war, ist damit die Behauptung gezeigt. □

SATZ VIII.3.11. *Sei $G$ konvex und $f \in C^1(G, \mathbb{R}^n)$. Dann sind folgende Aussagen äquivalent:*

(1) *$f$ ist Gradientenfeld.*
(2) *$\int_\Gamma f dx = 0$ für jede geschlossene, stückweise $C^1$-Kurve, die ganz in $G$ verläuft.*
(3) *$Df$ ist symmetrisch.*

Beweis. Wegen Satz VIII.3.9 und Bemerkung VIII.3.8 (2) müssen wir nur noch die Implikation „(3) $\Longrightarrow$ (1)“ zeigen. Sei dazu $x_0 \in G$ beliebig. Dann definieren wir die Funktion $F : G \to \mathbb{R}$ durch

$$F(x) = \int_{[x_0,x]} f dx \quad \forall x \in G,$$

Sei $x \in G$ und $\varepsilon > 0$ so, dass gilt $\overline{B(x,\varepsilon)} \subset G$. Aus Satz VIII.3.10 folgt dann für jedes $i \in \mathbb{N}_n^*$ und $h \in (-\varepsilon, \varepsilon)$

$$F(x + he_i) - F(x) = \int_{[x,x+he_i]} f dx.$$

Hieraus folgt aber wie im Beweis von Satz VIII.3.9, dass $F$ eine Stammfunktion von $f$ ist. □

Satz VIII.3.11 gibt uns einfaches Kriterium zur Lösung des Stammfunktionenproblems an die Hand. Allerdings ist die Voraussetzung „$G$ konvex“ noch zu stark. Um sie bestmöglich abzuschwächen, benötigen wir folgende Definition.

Definition VIII.3.12. (1) Seien $\gamma_1, \gamma_2 \in C(I, U)$, $U \subset \mathbb{R}^n$, zwei Wege mit gleichem Anfangs- und Endpunkt

$$x = \gamma_1(a) = \gamma_2(a) \quad , \quad y = \gamma_1(b) = \gamma_2(b).$$

Dann heißen $\gamma_1$ und $\gamma_2$ homotop in $U$, kurz $\gamma_1 \sim \gamma_2$, wenn es eine Homotopie $H \in C([0,1] \times I, U)$ gibt mit

$$\begin{aligned} H(0,t) &= \gamma_1(t) \quad \forall t \in I, \\ H(1,t) &= \gamma_2(t) \quad \forall t \in I, \\ H(s,a) &= x \quad \forall s \in [0,1], \\ H(s,b) &= y \quad \forall s \in [0,1]. \end{aligned}$$

(2) Ein geschlossener Weg $\gamma \in C(I,U), U \subset \mathbb{R}^n$ offen, heißt nullhomotop in $U$, kurz $\gamma \sim 0$, wenn er in $U$ homotop ist zu dem konstanten Weg $t \mapsto \gamma(a)$.
(3) Ein Gebiet $G$ heißt einfach zusammenhängend, wenn jeder geschlossene Weg in $G$ nullhomotop ist in $G$.

Bemerkung VIII.3.13. (1) Ist $H$ eine Homotopie, so ist $\gamma_s = H(s, .)$ für jedes $s \in [0,1]$ ein Weg mit Anfangspunkt $x$ und Endpunkt $y$.
(2) Eine konvexe Menge ist einfach zusammenhängend.

Beweis. ad (1): Folgt aus der Definition.
ad (2): Sei $\gamma \in C(I,G)$ ein geschlossener Weg in $G$. Dann leistet $H \in C([0,1] \times I, G)$ mit

$$H(s,t) = (1-s)\gamma(t) + s\gamma(a)$$

das Gewünschte. □

Die folgenden Beispiele sind anschaulich klar. Ein exakter Beweis ist jedoch zum Teil mühselig und wird hier nicht geführt. Die Aussage des zweiten Beispiels folgt aus Beispiel VIII.3.4 (1) und Satz VIII.3.16.

Beispiel VIII.3.14. (1) $\mathbb{R}^2 \backslash \{(x, 0) : x \in \mathbb{R}_+^*\}$ ist einfach zusammenhängend.
(2) $\mathbb{R}^2 \backslash \{0\}$ ist nicht einfach zusammenhängend.
(3) $\mathbb{R}^n \backslash \{0\}$ ist für $n \geq 3$ einfach zusammenhängend.

Der folgende Satz bereitet unser endgültiges Ergebnis vor, ist aber auch von eigenständigem praktischen Interesse.

Satz VIII.3.15. *Sei $G$ ein Gebiet, $f \in C^1(G, \mathbb{R}^n)$ und $Df$ symmetrisch. Dann gilt für je zwei in $G$ homotope, stückweise stetig differenzierbare Wege $\gamma_0, \gamma_1$*

$$\int_{\gamma_0} f dx = \int_{\gamma_1} f dx.$$

Beweis. O.E. ist $I = [0, 1]$. Sei $H \in C([0, 1] \times I, G)$ eine Homotopie zwischen $\gamma_0$ und $\gamma_1$. Da $H([0, 1] \times I) \subset G$ kompakt und $G$ offen ist, gibt es ein $\varepsilon > 0$ und ein $\delta > 0$ mit

(1) $\|H(s, t) - y\| \geq \varepsilon$ für alle $s \in [0, 1], t \in I, y \notin G$ und
(2) $\|H(s, t) - H(s', t')\| < \frac{\varepsilon}{2}$ für alle $s, s' \in [0, 1]$, $t, t' \in I$, mit $|s - s'| + |t - t'| \leq \delta$.

Sei $m \in \mathbb{N}^*$ mit $\frac{1}{m} \leq \frac{\delta}{2}$. Für $s \in [0, 1]$ bezeichnen wir dann mit $\widetilde{\gamma}_s$ den Polygonzug mit den Ecken

$$H(s, 0), H(s, \frac{1}{m}), \ldots, H(s, \frac{m-1}{m}), H(s, 1).$$

Wegen (1) und (2) gilt

$$\widetilde{\gamma}_s(I) \subset G \quad \forall s \in [0, 1].$$

Wir zeigen nun zunächst

$$\int_{\widetilde{\gamma}_{\frac{i-1}{m}}} f dx = \int_{\widetilde{\gamma}_{\frac{i}{m}}} f dx \quad \forall 1 \leq i \leq m. \tag{3}$$

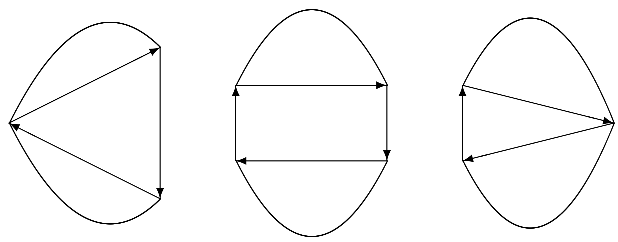

Abbildung VIII.3.2. Beispiel für Wege $\widetilde{\gamma}_{i1}$, $\widetilde{\gamma}_{ij}$, $2 \leq j \leq m-1$, und $\widetilde{\gamma}_{im}$

Sei $1 \le i \le m$ und $1 \le j \le m$. Definiere den Weg $\widetilde{\gamma}_{ij}$ wie folgt (vgl. Abb. VIII.3.2

$$\widetilde{\gamma}_{ij} = \begin{cases} [H(\frac{i-1}{m}, 0), H(\frac{i-1}{m}, \frac{1}{m})] \\ \oplus [H(\frac{i-1}{m}, \frac{1}{m}), H(\frac{i}{m}, \frac{1}{m})] \\ \oplus [H(\frac{i}{m}, \frac{1}{m}), H(\frac{i}{m}, 0)] & \text{für } j = 1, \\ [H(\frac{i-1}{m}, \frac{j-1}{m}), H(\frac{i-1}{m}, \frac{j}{m})] \\ \oplus [H(\frac{i-1}{m}, \frac{j}{m}), H(\frac{i}{m}, \frac{j}{m})] \\ \oplus [H(\frac{i}{m}, \frac{j}{m}), H(\frac{i-1}{m}, \frac{j}{m})] \\ \oplus [H(\frac{i-1}{m}, \frac{j}{m}), H(\frac{i-1}{m}, \frac{j-1}{m})] & \text{für } 2 \le j \le m-1, \\ [H(\frac{i-1}{m}, \frac{m-1}{m}), H(\frac{i-1}{m}, 1)] \\ \oplus [H(\frac{i}{m}, 1), H(\frac{i}{m}, \frac{m-1}{m})] \\ \oplus [H(\frac{i}{m}, \frac{m-1}{m}), H(\frac{i-1}{m}, \frac{m-1}{m})] & \text{für } j = m. \end{cases}$$

Wegen (2) gilt

$$\operatorname{Spur}(\widetilde{\gamma}_{ij}) \subset B(H(\frac{i-1}{m}, \frac{j-1}{m}), \frac{2}{3}\varepsilon) \subset G.$$

Damit folgt aus Satz VIII.3.11

$$\int_{\widetilde{\gamma}_{ij}} f dx = 0 \quad \forall 1 \le i, j \le m$$

und daher

$$0 = \sum_{j=1}^{m} \int_{\widetilde{\gamma}_{ij}} f = \int_{\widetilde{\gamma}_{\frac{i-1}{m}}} f dx - \int_{\widetilde{\gamma}_{\frac{i}{m}}} f dx, \quad 1 \le i \le m.$$

Damit ist (3) gezeigt.
Wir zeigen nun

$$(4) \qquad \int_{\widetilde{\gamma}_0} f dx = \int_{\gamma_0} f dx \, , \int_{\widetilde{\gamma}_1} f dx = \int_{\gamma_1} f dx.$$

Dazu definieren wir für $\mu = 0, 1$ und $1 \le j \le m$ die Wege $\widehat{\gamma}_{\mu j}$ durch

$$\widehat{\gamma}_{\mu j} = H(\mu, t)|_{[\frac{j-1}{m}, \frac{j}{m}]} \oplus [H(\mu, \frac{j}{m}), H(\mu, \frac{j-1}{m})].$$

Wegen (2) ist dann wieder

$$\operatorname{Spur}(\widehat{\gamma}_{\mu j}) \subset B(H(\mu, \frac{j-1}{m}), \frac{2}{3}\varepsilon) \subset G \quad \forall 1 \le j \le m, \mu = 0, 1,$$

so dass aus Satz VIII.3.11 folgt

$$\int_{\widehat{\gamma}_{\mu j}} f dx = 0 \quad 1 \le j \le m, \mu = 0, 1$$

Damit erhalten wir

$$0 = \sum_{j=1}^{m} \int_{\widehat{\gamma}_{\mu j}} f dx = \int_{\gamma_\mu} f dx - \int_{\widehat{\gamma}_\mu} f dx, \quad \mu = 0, 1.$$

Damit ist auch (4) bewiesen.
Aus (3) und (4) folgt offensichtlich die Behauptung. □

Nun können wir die Frage nach der Stammfunktion endgültig beantworten.

Satz VIII.3.16 (Satz von Poincaré). *Sei $G$ einfach zusammenhängend und $f \in C^1(G, \mathbb{R}^n)$. Dann sind folgende Aussagen äquivalent:*

(1) *$f$ ist Gradientenfeld.*
(2) *$\int_\Gamma f dx = 0$ für jede geschlossene, stückweise $C^1$-Kurve, die ganz in $G$ verläuft.*
(3) *$Df$ ist symmetrisch.*

Beweis. Wegen Satz VIII.3.9 und Bemerkung VIII.3.8 (2) müssen wir nur noch die Implikation „(3) $\Longrightarrow$ (2)“ zeigen. Sei also $\Gamma$ eine geschlossene, stückweise $C^1$-Kurve, die ganz in $G$ verläuft. Dann ist $\Gamma$ in $G$ homotop zu dem konstanten Weg $\widetilde{\gamma} : t \mapsto \Gamma(a)$. Aus Satz VIII.3.15 folgt

$$\int_\Gamma f dx = \int_{\widetilde{\gamma}} f dx = \int_a^b \langle f \circ \widetilde{\gamma}(t), \underbrace{\widetilde{\gamma}'(t)}_{=0} \rangle dt = 0.$$

Damit ist die Behauptung bewiesen. □

## VIII.4. Komplexe Kurvenintegrale

In diesem Paragraphen beweisen wir einige der wichtigsten Sätze der Funktionentheorie. Wir knüpfen dabei an Ergebnisse der Paragraphen V.3 und VII.2 an und nutzen Ergebnisse des letzten Paragraphen.

Im Folgenden bezeichnet $G \subset \mathbb{C}$ stets ein Gebiet in $\mathbb{C}$.

Definition VIII.4.1. Sei $f \in C(G, \mathbb{C})$ und $\gamma \in C^1(I, G)$ ein $C^1$-Weg in $G$. Dann ist das Wegintegral von $f$ längs $\gamma$ definiert durch

$$\int_\gamma f dz = \int_a^b f(\gamma(t)) \cdot \gamma'(t) dt.$$

Bemerkung VIII.4.2. Seien $f \in C(G, \mathbb{C})$ und $\gamma \in C^1(I, G)$. Dann sei $W = \{(x, y) \in \mathbb{R}^2 : x + iy \in G\}$ und

$$\begin{aligned}
u(x, y) &= \operatorname{Re} f(x + iy) \in C(W, \mathbb{R}),\\
v(x, y) &= \operatorname{Im} f(x + iy) \in C(W, \mathbb{R}),\\
F(x, y) &= (u(x, y), -v(x, y)) \in C(W, \mathbb{R}^2),\\
\overline{F}(x, y) &= (v(x, y), u(x, y)) \in C(W, \mathbb{R}^2),\\
\gamma_1(t) &= \operatorname{Re} \gamma(t) \in C^1(I, \mathbb{R}),\\
\gamma_2(t) &= \operatorname{Im} \gamma(t) \in C^1(I, \mathbb{R}),\\
\widetilde{\gamma}(t) &= (\gamma_1(t), \gamma_2(t)) \in C^1(I, W).
\end{aligned}$$

Dann ist

$$\begin{aligned}
&f(\gamma(t)) \cdot \gamma'(t) \\
&= [u(\widetilde{\gamma}(t)) + iv(\widetilde{\gamma}(t))] \cdot [\gamma_1'(t) + i\gamma_2'(t)] \\
&= [u(\widetilde{\gamma}(t))\gamma_1'(t) - v(\widetilde{\gamma}(t))\gamma_2'(t)] + i[u(\widetilde{\gamma}(t))\gamma_2'(t) + v(\widetilde{\gamma}(t))\gamma_1'(t)] \\
&= \langle [u(\widetilde{\gamma}(t)), -v(\widetilde{\gamma}(t))], \widetilde{\gamma}(t)\rangle + i\langle [v(\widetilde{\gamma}(t)), u(\widetilde{\gamma}(t))], \widetilde{\gamma}(t)\rangle.
\end{aligned}$$

Also ist

$$\int_\gamma f dz = \int_{\widetilde{\gamma}} F dx + i \int_{\widetilde{\gamma}} \overline{F} dx.$$

Auf diese Weise können wir das komplexe Wegintegral $\int_\gamma f dz$ mit den beiden reellen Wegintegralen $\int_{\widetilde{\gamma}} F dx$ und $\int_{\widetilde{\gamma}} \overline{F} dx$ identifizieren. Wir werden dies im Folgenden häufiger tun, wobei wir der Einfachheit halber wieder $G$ statt $W$ und $\gamma$ statt $\widetilde{\gamma}$ schreiben. Die Bedeutung ist dann aus dem Zusammenhang klar. Insbesondere können wir auf diese Art und Weise die Begriffe (z.B. Integral längs eines stückweise $C^1$-Weges usw.) und Sätze des vorigen Abschnittes übertragen.

Definition VIII.4.3. Die Funktion $f : G \to \mathbb{C}$ heißt holomorph in $G$, wenn sie in ganz $G$ stetig komplex differenzierbar ist im Sinne von Definition IV.1.1 (S. 111).

Satz VIII.4.4 (Cauchyscher Integralsatz). *Sei $G$ einfach zusammenhängend und $f$ in $G$ holomorph. Dann ist*

$$\int_\gamma f dz = 0$$

*für jeden geschlossenen, stückweise $C^1$-Weg in $G$. Insbesondere besitzt $f$ in $G$ eine Stammfunktion.*

Beweis. Wegen der Cauchy-Riemannschen Differentialgleichungen, Satz VII.2.17 (S. 224), sind die Funktionalmatrizen $DF$ und $D\overline{F}$ der Funktionen $F$ und $\overline{F}$ aus Bemerkung VIII.4.2 symmetrisch. Damit folgt die Behauptung aus Satz VIII.3.16 (S. 281). □

Satz VIII.4.5 (Cauchysche Integralformel). *Sei $f$ in $G$ holomorph, $z_0 \in G$ und $R \in \mathbb{R}_+^*$ mit $K = \overline{B(z_0, R)} \subset G$. Dann folgt für alle $z \in \mathring{K}$*

$$f(z) = \frac{1}{2\pi i} \int_{\partial K} \frac{f(\eta)}{\eta - z} d\eta,$$

*wobei $\partial K$ im mathematisch positiven Sinn durchlaufen wird.*

Beweis. Sei $z \in \mathring{K}$ beliebig. Definiere $g : G \to \mathbb{C}$ durch

$$g(w) = \begin{cases} \frac{f(w) - f(z)}{w - z} & \text{falls } w \neq z, \\ f'(z) & \text{falls } w = z. \end{cases}$$

Dann ist $g$ auf $G$ stetig und auf $G\backslash\{z\}$ holomorph. Sei $r = \frac{1}{2}(R-|z-z_0|)$ und

$$M = \max_{w\in \overline{B(z,r)}} |g(w)|.$$

Da $\partial K$ in $K\backslash\{z\}$ homotop ist zu $\partial B(z,\varepsilon)$ für jedes $0 < \varepsilon \le r$, folgt aus Satz VIII.3.15 (S. 279) und Satz VIII.3.6 (S. 273)

$$|\int_{\partial K} g d\eta| = |\int_{\partial B(z,\varepsilon)} g d\eta| \le M 2\pi\varepsilon \quad \forall 0 < \varepsilon \le r.$$

Also ist

$$0 = \int_{\partial K} g d\eta = \int_{\partial K} \frac{f(\eta) - f(z)}{\eta - z} d\eta$$

und somit

$$\frac{1}{2\pi i} \int_{\partial K} \frac{f(\eta)}{\eta - z} d\eta = \frac{f(z)}{2\pi i} \int_{\partial K} \frac{1}{\eta - z} d\eta.$$

Da die Funktion $w \mapsto \frac{1}{w-z}$ auf $K\backslash\{z\}$ holomorph ist, folgt wieder aus Satz VIII.3.15 (S. 279)

$$\begin{aligned}
\frac{1}{2\pi i} \int_{\partial K} \frac{1}{\eta - z} d\eta &= \frac{1}{2\pi i} \int_{\partial B(z,r)} \frac{1}{\eta - z} d\eta \\
&= \frac{1}{2\pi i} \int_0^{2\pi} i dt \\
&= 1.
\end{aligned}$$

Damit ist die Behauptung bewiesen. □

Satz VIII.4.6. *Sei $f$ in $G$ holomorph. Dann ist $f$ auf $G$ analytisch.*

Beweis. Seien $z_0 \in G$ und $R \in \mathbb{R}_+^*$ mit $K = \overline{B(z_0, R)} \subset G$. Definiere

$$M = \max_{w\in\partial K} |f(w)|.$$

Sei $z \in \overset{\circ}{K}$ beliebig und $r_0 = |z - z_0| < R$. Wegen

$$\begin{aligned}
|\frac{f(w)}{(w - z_0)^{k+1}} (z - z_0)^k| &\le M \frac{r_0^k}{R^{k+1}} \\
&= \frac{M}{R} (\frac{r_0}{R})^k \quad \forall w \in \partial K,\ k \in \mathbb{N}
\end{aligned}$$

konvergiert die Reihe

$$\sum \frac{f(w)}{(w - z_0)^{k+1}} (z - z_0)^k$$

auf $\partial K$ gleichmäßig. Damit folgt aus Satz VI.2.1 und Satz VIII.4.5 mit

$$a_k = \frac{1}{2\pi i} \int_{\partial K} \frac{f(\eta)}{(\eta - z_0)^{k+1}} d\eta \quad \forall k \in \mathbb{N}$$

die Beziehung

$$\begin{aligned}\sum_{k=0}^{\infty} a_k(z-z_0)^k &= \sum_{k=0}^{\infty}[\frac{1}{2\pi i}\int_{\partial K}\frac{f(\eta)}{(\eta-z_0)^{k+1}}d\eta](z-z_0)^k \\ &= \frac{1}{2\pi i}\int_{\partial K}[\sum_{k=0}^{\infty}\frac{f(\eta)}{\eta-z_0}(\frac{z-z_0}{\eta-z_0})^k]d\eta \\ &= \frac{1}{2\pi i}\int_{\partial K}\frac{f(\eta)}{\eta-z_0}\sum_{k=0}^{\infty}(\frac{z-z_0}{\eta-z_0})^k d\eta \\ &= \frac{1}{2\pi i}\int_{\partial K}\frac{f(\eta)}{\eta-z_0}\frac{1}{1-\frac{z-z_0}{\eta-z_0}}d\eta \\ &= \frac{1}{2\pi i}\int_{\partial K}\frac{f(\eta)}{\eta-z}d\eta \\ &= f(z).\end{aligned}$$

Wegen Satz VIII.3.6 (S. 273) gilt

$$|a_k| \le \frac{1}{2\pi}2\pi RMR^{-(k+1)} = MR^{-k}.$$

Mithin hat die Reihe $\sum a_k(z-z_0)^k$ den Konvergenzradius $R$. Damit ist die Behauptung bewiesen. □

Bemerkung VIII.4.7. (1) Für komplexe Funktionen gilt also $C^1(G,\mathbb{C}) = C^\omega(G,\mathbb{C})$ im Gegensatz zu reellen Funktionen.
(2) Aus dem Beweis von Satz VIII.4.6 folgt direkt die Cauchysche Ableitungsformel

$$f^{(k)}(z_0) = \frac{k!}{2\pi i}\int_{\partial B(z_0,R)}\frac{f(\eta)}{(\eta-z_0)^{k+1}}d\eta \quad \forall k \in \mathbb{N}.$$

Satz VIII.4.8 (Satz von Liouville). *Die einzigen auf ganz $\mathbb{C}$ holomorphen und beschränkten Funktionen sind die Konstanten.*

Beweis. Sei $f$ auf $\mathbb{C}$ holomorph und beschränkt, d.h. es gibt ein $M > 0$ mit

$$|f(z)| \le M \quad \forall z \in \mathbb{C}.$$

Sei $z_0 \in \mathbb{C}$ beliebig und $n \in \mathbb{N}^*$. Aus Bemerkung VIII.4.7 (2) folgt für jedes $r \in \mathbb{R}_+^*$

$$\begin{aligned}|f^{(n)}(z_0)| &= \frac{n!}{2\pi}|\int_{\partial B(z_0,r)}\frac{f(\eta)}{(\eta-z_0)^{n+1}}d\eta| \\ &\le \frac{n!}{2\pi}M2\pi r\frac{1}{r^{n+1}} \\ &= Mn!r^{-n} \\ &\underset{r\to+\infty}{\longrightarrow} 0.\end{aligned}$$

Hieraus folgt die Behauptung. □

Beispiel VIII.4.9. Es gilt

$$\int_0^\infty \cos(2t^2)dt = \int_0^\infty \sin(2t^2)dt = \frac{\sqrt{\pi}}{4}.$$

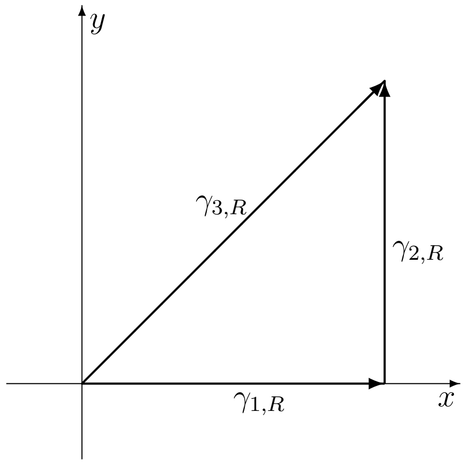


Abbildung VIII.4.1. Wege $\gamma_{1,R}$, $\gamma_{2,R}$, $\gamma_{3,R}$

Beweis. Die Funktion $e^{-z^2}$ ist holomorph auf $\mathbb{C}$. Sei $R > 0$ und (vgl. Abbildung VIII.4.1)

$$\begin{aligned}\gamma_{1,R} &= [0, R],\\ \gamma_{2,R} &= [R, R(1+i)],\\ \gamma_{3,R} &= [0, R(1+i)].\end{aligned}$$

Aus Satz VIII.4.4 folgt

$$\int_{\gamma_{3,R}} e^{-z^2}dz = \int_{\gamma_{1,R}} e^{-z^2}dz + \int_{\gamma_{2,R}} e^{-z^2}dz.$$

Es ist

$$\begin{aligned}\int_{\gamma_{3,R}} e^{-z^2}dz &= \int_0^R e^{-(t(1+i))^2}(1+i)dt\\ &= (1+i)\int_0^R e^{-2it^2}dt\\ &= [\int_0^R \cos(2t^2)dt + \int_0^R \sin(2t^2)dt]\\ &\quad + i[\int_0^R \cos(2t^2)dt - \int_0^R \sin(2t^2)dt]\\ \int_{\gamma_{1,R}} e^{-z^2}dz &= \int_0^R e^{-t^2}dt\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\int_{-R}^{R} e^{-t^2} dt \\
&\underset{R\to+\infty}{\longrightarrow} \frac{\sqrt{\pi}}{2} \quad \text{gemäß Satz VI.5.7 (S. 210)}
\end{aligned}
$$

$$
\begin{aligned}
\Big|\int_{\gamma_{2,R}} e^{-z^2} dz\Big| &= \Big|\int_0^1 e^{-[R(1+it)]^2} iR dt\Big| \\
&= \Big|iR\int_0^1 e^{-R^2(1-t^2)} e^{-2iR^2t} dt\Big| \\
&\le R\int_0^1 e^{-R^2(1-t^2)} dt \\
&= Re^{-R^2}\int_0^1 e^{R^2t} e^{-R^2t(1-t)} dt \\
&\le Re^{-R^2}\int_0^1 e^{R^2t} dt \\
&= Re^{-R^2}\frac{1}{R^2}[e^{R^2}-1] \\
&\underset{R\to+\infty}{\longrightarrow} 0.
\end{aligned}
$$

Also existiert $\lim\limits_{R\to+\infty}\int_{\gamma_{3,R}} e^{-z^2} dz$ und erfüllt

$$
\begin{aligned}
\frac{\sqrt{\pi}}{2} &= \lim_{R\to+\infty}\int_{\gamma_{1,R}} e^{-z^2} dz \\
&= \lim_{R\to+\infty}\int_{\gamma_{3,R}} e^{-z^2} dz \\
&= \Big[\int_0^\infty \cos(2t^2) dt + \int_0^\infty \sin(2t^2) dt\Big] \\
&\quad + i\Big[\int_0^\infty \cos(2t^2) dt - \int_0^\infty \sin(2t^2) dt\Big].
\end{aligned}
$$

Hieraus folgt die behauptete Gleichheit. □

Satz VIII.4.6 erlaubt die Berechnung der Potenzreihenentwicklung einer holomorphen Funktion in einem Punkt ihres Definitionsbereiches. Nun wollen wir Funktionen in „Potenzreihen“ um Punkte außerhalb ihres Definitionsbereiches entwickeln.

Satz VIII.4.10. *Seien $z_0 \in G$ mit $\overline{B(z_0, R)} \subset G$ und $f$ in $G\backslash\{z_0\}$ holomorph. Dann kann $f$ um $z_0$ in eine* Laurent-Reihe

$$\sum_{n\in\mathbb{Z}} a_n(z-z_0)^n$$

*entwickelt werden, die für jedes $r \in (0, R)$ in der Kreisscheibe $\{z \in \mathbb{C} : r < |z-z_0| < R\}$ gleichmäßig gegen $f(z)$ konvergiert. Für die*

*Koeffizienten* $a_n$ *gilt*

$$a_n = \frac{1}{2\pi i}\int_{\partial B(z_0,s)} \frac{f(\eta)}{(\eta - z_0)^{n+1}} d\eta \quad \forall n \in \mathbb{Z},\ 0 < s < R.$$

*Der Koeffizient* $a_{-1}$ *heißt* RESIDUUM *von* $f$ *in* $z_0$ *und wird mit* $\operatorname{Res}_{z_0} f$ *bezeichnet.*

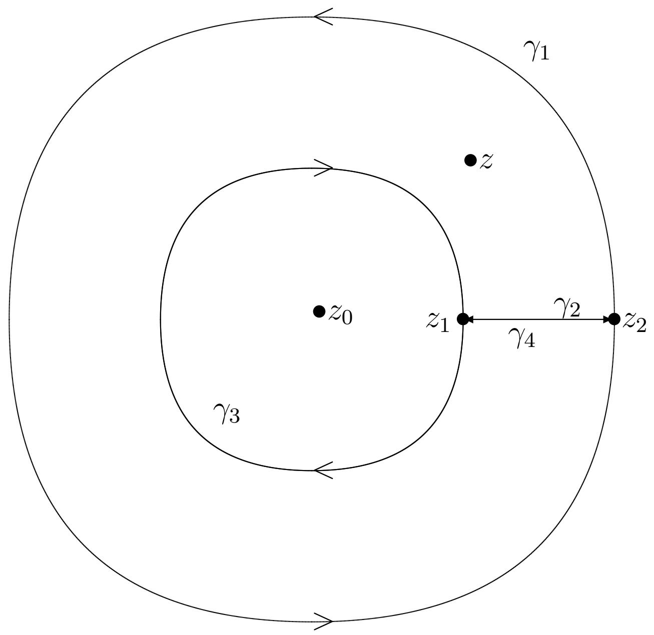


ABBILDUNG VIII.4.2. Wege $\gamma_1$, $\gamma_2$, $\gamma_3$, $\gamma_4$

BEWEIS. Sei $z \in B(z_0, R)\backslash\{z_0\}$ beliebig und $0 < r_1 < |z - z_0| < r_2 < R$. Weiter sei $z_2 \in \partial B(z_0, r_2)$, so dass $z$ nicht auf der Geraden durch $z_0$ und $z_2$ liegt. Bezeichne mit $z_1$ den Schnittpunkt der Geraden durch $z_0$ und $z_2$ mit $\partial B(z_0, r_1)$. Sei (vgl. Abbildung VIII.4.2)

$\gamma_1$: der positiv orientierte Kreisbogen mit Mittelpunkt $z_0$, Radius $r_2$ und Anfangspunkt $z_2$,
$\gamma_2$: die Strecke von $z_2$ nach $z_1$,
$\gamma_3$: der negativ orientierte Kreisbogen mit Mittelpunkt $z_0$, Radius $r_1$ und Anfangspunkt $z_1$,
$\gamma_4$: die Strecke von $z_1$ nach $z_2$

und

$$\gamma = \gamma_1 \oplus \gamma_2 \oplus \gamma_3 \oplus \gamma_4.$$

Da $\gamma$ in $B(z_0, R)\backslash\{z_0\}$ nullhomotop ist, folgt mit Satz VIII.4.5 und Satz VIII.3.15 (S. 279)

$$f(z) = \frac{1}{2\pi i}\int_\gamma \frac{f(\eta)}{\eta - z} d\eta$$

$$= \frac{1}{2\pi i} \int_{\gamma_1} \frac{f(\eta)}{\eta - z} d\eta + \frac{1}{2\pi i} \int_{\gamma_3} \frac{f(\eta)}{\eta - z} d\eta.$$

Wie im Beweis von Satz VIII.4.6 folgt

$$\frac{1}{2\pi i} \int_{\gamma_1} \frac{f(\eta)}{\eta - z} d\eta = \sum_{n=0}^{\infty} \alpha_n (z - z_0)^n$$

mit

$$\alpha_n = \frac{1}{2\pi i} \int_{\partial B(z_0, r_2)} \frac{f(\eta)}{(\eta - z_0)^{n+1}} d\eta$$

und, dass die Reihe $\sum \alpha_n (z - z_0)^n$ in $B(z_0, r_2)$ gleichmäßig konvergiert. Sei

$$M = \max_{w \in \partial B(z_0, r_1)} |f(w)|.$$

Dann gilt

$$\begin{aligned} \left|\frac{f(w)}{(w - z_0)^{-k+1}} (z - z_0)^{-k}\right| &\leq \frac{M}{r_1^{-k+1}} |z - z_0|^{-k} \\ &= \frac{M}{r_1} \left(\frac{r_1}{|z - z_0|}\right)^k \quad \forall w \in \partial B(z_0, r_1),\ k \in \mathbb{N}^*. \end{aligned}$$

Mithin konvergiert die Reihe

$$\sum_{k \in \mathbb{N}^*} \frac{f(w)}{(w - z_0)^{-k+1}} (z - z_0)^{-k}$$

auf $\partial B(z_0, r_1)$ gleichmäßig. Damit folgt aus Satz VI.2.1 (S. 180) mit

$$\beta_k = \frac{1}{2\pi i} \int_{\partial B(z_0, r_1)} \frac{f(\eta)}{(\eta - z_0)^{-k+1}} d\eta \ , k \in \mathbb{N}^*$$

die Beziehung

$$\begin{aligned} \sum_{k=1}^{\infty} \beta_k (z - z_0)^{-k} &= \sum_{k=1}^{\infty} \left[\frac{1}{2\pi i} \int_{\partial B(z_0, r_1)} \frac{f(\eta)}{(\eta - z_0)^{-k+1}} d\eta\right] (z - z_0)^{-k} \\ &= \frac{1}{2\pi i} \int_{\partial B(z_0, r_1)} \left[\sum_{k=1}^{\infty} \frac{f(\eta)}{\eta - z_0} \left(\frac{\eta - z_0}{z - z_0}\right)^k\right] d\eta \\ &= \frac{1}{2\pi i} \int_{\partial B(z_0, r_1)} \frac{f(\eta)}{\eta - z_0} \left[\frac{1}{1 - \frac{\eta - z_0}{z - z_0}} - 1\right] d\eta \\ &= -\frac{1}{2\pi i} \int_{\partial B(z_0, r_1)} \frac{f(\eta)}{\eta - z} d\eta \\ &= \int_{\gamma_3} \frac{f(\eta)}{\eta - z} d\eta. \end{aligned}$$

Wegen

$$|\beta_k (z - z_0)^{-k}| \leq M r_1 r_1^{k-1} |z - z_0|^{-k}$$

$$= M(\frac{r_1}{|z - z_0|})^k \quad \forall k \in \mathbb{N}^*$$

konvergiert die Reihe $\sum \beta_k (z - z_0)^{-k}$ auf $\mathbb{C} \backslash \overline{B(z_0, r_1)}$ gleichmäßig. Insgesamt erhalten wir

$$f(z) = \sum_{k=1}^{\infty} \beta_k (z - z_0)^{-k} + \sum_{n=0}^{\infty} \alpha_n (z - z_0)^n,$$

wobei die Reihen auf der rechten Seite gleichmäßig auf $\{z \in \mathbb{C} : r_1 < |z - z_0| < r_2\}$ konvergieren. Da die Funktionen $z \mapsto (z - z_0)^n$ für alle $n \in \mathbb{Z}$ auf $\mathbb{C} \backslash \{z_0\}$ holomorph sind, folgt aus Satz VIII.3.15 (S. 279)

$$\alpha_n = \frac{1}{2\pi i} \int_{\partial B(z_0, s)} \frac{f(\eta)}{(\eta - z_0)^{n+1}} d\eta \quad \forall n \in \mathbb{N}$$

$$\beta_k = \frac{1}{2\pi i} \int_{\partial B(z_0, s)} \frac{f(\eta)}{(\eta - z_0)^{-k+1}} d\eta \quad \forall k \in \mathbb{N}^*$$

und alle $s \in (0, R)$. Damit ist die Behauptung bewiesen. □

Bemerkung VIII.4.11. (1) Die Reihe

$$\sum_{k=1}^{\infty} \beta_k (z - z_0)^{-k}$$

heißt Hauptteil der Laurent-Entwicklung von $f$ um $z_0$. Aus dem Beweis von Satz VIII.4.10 folgt, dass er für jedes $r \in \mathbb{R}_+^*$ auf $\mathbb{C} \backslash \overline{B(z_0, r)}$ gleichmäßig konvergiert. Mit den gleichen Argumenten wie beim Beweis der Sätze V.3.1 (S. 156) und V.3.2 (S. 157) folgt hieraus, dass der Hauptteil auf $\mathbb{C} \backslash \{z_0\}$ holomorph ist.
(2) Gilt $\beta_k = 0$ für alle $k \in \mathbb{N}^*$, so heißt $z_0$ eine hebbare Singularität von $f$. Gibt es ein $m \in \mathbb{N}^*$ mit $\beta_m \neq 0$ und $\beta_k = 0$ für alle $k > m$, so heißt $z_0$ eine Polstelle der Ordnung $m$ von $f$. Ist schließlich $z_0$ weder eine hebbare Singularität noch eine Polstelle, so heißt $z_0$ eine wesentliche Singularität von $f$.
(3) Ist $z_0$ eine Polstelle der Ordnung $m$ von $f$, so ist

$$\operatorname{Res}_{z_0} f = \frac{1}{(m-1)!} \frac{d^{m-1}}{dz^{m-1}} [(z - z_0)^m f(z)](z_0).$$

Beispiel VIII.4.12. (1) Sei

$$f(z) = \exp(\frac{1}{z}).$$

$z_0 = 0$ ist eine wesentliche Singularität und

$$\operatorname{Res}_0 f = 1.$$

(2) Sei

$$f(z) = \frac{1}{(z-1)^2 (z+1)}.$$

$z_0 = 1$ ist eine Polstelle 2-ter Ordnung und

$$\operatorname{Res}_1 f = -\frac{1}{4}.$$

$z_0 = -1$ ist eine Polstelle 1-ter Ordnung und

$$\operatorname{Res}_{-1} f = \frac{1}{4}.$$

Satz VIII.4.13 (Riemannscher Hebbarkeitssatz). *Sei $z_0 \in G$ und $f$ in $G\backslash\{z_0\}$ holomorph. Dann sind folgende Aussagen äquivalent:*

(1) *$f$ kann in $z_0$ holomorph fortgesetzt werden.*
(2) *$f$ ist in einer Umgebung von $z_0$ beschränkt.*

Beweis. (1) $\Longrightarrow$ (2): Ist klar, da die Fortsetzung insbesondere stetig ist.
(2) $\Longrightarrow$ (1): Dann gibt es ein $r \in \mathbb{R}_+^*$, so dass $f$ auf $B(z_0, r)\backslash\{z_0\}$ beschränkt ist. Sei

$$M = \sup_{z \in B(z_0,r)\backslash\{z_0\}} |f(z)|.$$

Dann folgt für $k \in \mathbb{N}^*$ und $s \in (0, r)$

$$\left|\frac{1}{2\pi i}\int_{\partial B(z_0,s)} \frac{f(\eta)}{(\eta - z_0)^{-k+1}} d\eta\right| \le M s^k \xrightarrow[s\to 0]{} 0.$$

Damit folgt die Behauptung aus Satz VIII.4.10 und Bemerkung VIII.4.11 (2). □

Satz VIII.4.14 (Residuensatz). *Sei $G$ einfach zusammenhängend, $z_1, \ldots, z_m \in G$ und $f$ in $G\backslash\{z_1, \ldots, z_m\}$ holomorph. Dann gilt für jeden geschlossenen Weg $\gamma$, der für jedes $i \in \mathbb{N}_m^*$ in $G\backslash\{z_i\}$ zu $\partial B(z_i, \varepsilon)$ mit hinreichend kleinem $\varepsilon > 0$ homotop ist,*

$$\frac{1}{2\pi i}\int_\gamma f(z)dz = \sum_{j=1}^m \operatorname{Res}_{z_j} f.$$

Beweis. Sei $h_j$, $1 \le j \le m$, der Hauptteil der Laurent-Entwicklung von $f$ um $z_j$. Dann ist $f - \sum_{j=1}^m h_j$ in $G$ holomorph. Damit folgt aus Satz VIII.4.4

$$\frac{1}{2\pi i}\int_\gamma f(z)dz = \sum_{j=1}^m \frac{1}{2\pi i}\int_\gamma h_j(z)dz.$$

Sei nun $1 \le j \le m$ beliebig. Da $\gamma$ in $G\backslash\{z_j\}$ zu einem positiv orientierten, einfach durchlaufenen Kreis um $z_j$ homotop ist, folgt aus Satz VIII.3.15 (S. 279) für $\varepsilon \in \mathbb{R}_+^*$ mit $\overline{B(z_j, \varepsilon)} \subset G$

$$\frac{1}{2\pi i}\int_\gamma h_j(z)dz = \frac{1}{2\pi i}\int_{\partial B(z_j,\varepsilon)} h_j(z)dz$$

$$
\begin{aligned}
&= \sum_{k=1}^{\infty} \beta_{k,j} \frac{1}{2\pi i} \int_{\partial B(z_j,\varepsilon)} (z-z_j)^{-k} dz \\
&= \sum_{k=1}^{\infty} \beta_{k,j} \frac{1}{2\pi i} \int_0^{2\pi} e^{-ikt} i e^{it} dt \\
&= \beta_{1,j} \\
&= \operatorname{Res}_{z_j} f.
\end{aligned}
$$

Damit ist die Behauptung bewiesen. □

Beispiel VIII.4.15. (1) Seien $\alpha, \beta \in \mathbb{R}_+^*$ und $n \in \mathbb{N}^*$. Dann ist

$$
\int_{-\infty}^{\infty} \frac{dx}{(ax^2+b)^n} = \frac{4}{(4b)^n} \sqrt{\frac{b}{a}} \frac{(2n-1)!}{[(n-1)!]^2}.
$$

Denn: Sei

$$
f(z) = \frac{1}{(az^2+b)^n} = \frac{1}{a^n (z - i\sqrt{\frac{b}{a}})^n (z + i\sqrt{\frac{b}{a}})^n}.
$$

Es ist

$$
\begin{aligned}
\operatorname{Res}_{i\sqrt{\frac{b}{a}}} f &= \frac{1}{(n-1)!} \frac{d^{n-1}}{dz^{n-1}} \Big[ \frac{1}{a^n (z + i\sqrt{\frac{b}{a}})^n} \Big]_{z = i\sqrt{\frac{b}{a}}} \\
&= \frac{1}{a^n} \frac{(2n-1)!(-1)^{n-1}}{[(n-1)!]^2} \frac{1}{(2i\sqrt{\frac{b}{a}})^{2n-1}} \\
&= \frac{-2i}{(4b)^n} \sqrt{\frac{b}{a}} \frac{(2n-1)!}{[(n-1)!]^2}.
\end{aligned}
$$

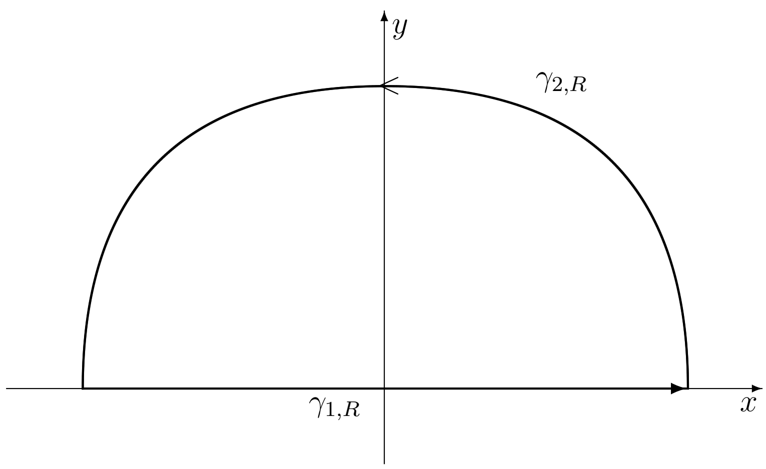


Abbildung VIII.4.3. Wege $\gamma_{1,R}$, $\gamma_{2,R}$

Sei $R > \sqrt{\frac{b}{a}}$ und $\gamma_{1,R}$ die Strecke von $-R$ nach $R$ und $\gamma_{2,R}$ der Halbkreis um 0 von $R$ nach $-R$ (vgl. Abbildung VIII.4.3). Dann folgt aus Satz VIII.4.14

$$\begin{aligned}\frac{4\pi}{(4b)^n}\sqrt{\frac{b}{a}}\frac{(2n-1)!}{[(n-1)!]^2} &= 2\pi i \operatorname{Res}_{i\sqrt{\frac{b}{a}}} f \\ &= \int_{\gamma_{1,R}} f(z)dz + \int_{\gamma_{2,R}} f(z)dz \\ &= \int_{-R}^{R} \frac{dx}{(ax^2+b)^n} + \int_{\gamma_{2,R}} f(z)dz \\ &= 2\int_0^R \frac{dx}{(ax^2+b)^n} + \int_{\gamma_{2,R}} f(z)dz.\end{aligned}$$

Weiter ist

$$\begin{aligned}|\int_{\gamma_{2,R}} f(z)dz| &\le \pi R \max_{z\in\gamma_{2,R}} |f(z)| \\ &= \pi R \max_{z\in\gamma_{2,R}} \frac{1}{|az^2+b|^n} \\ &\le \frac{\pi R}{a^n(R^2-\frac{b}{a})^n} \\ &\underset{R\to+\infty}{\longrightarrow} 0.\end{aligned}$$

Also existiert $\lim\limits_{R\to+\infty}\int_0^R \frac{dx}{(ax^2+b)^n}$. Damit folgt

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{dx}{(ax^2+b)^n} &= \lim_{\alpha\to-\infty}\int_\alpha^0 \frac{dx}{(ax^2+b)^n} + \lim_{\beta\to+\infty}\int_0^\beta \frac{dx}{(ax^2+b)^n} \\ &= \lim_{\alpha\to-\infty}\int_0^{-\alpha} \frac{dx}{(ax^2+b)^n} + \lim_{\beta\to+\infty}\int_0^\beta \frac{dx}{(ax^2+b)^n} \\ &= \frac{1}{2}\lim_{\alpha\to\infty}[\int_{\gamma_{1,\alpha}} f(z)dz + \int_{\gamma_{2,\alpha}} f(z)dz] \\ &\quad + \frac{1}{2}\lim_{\beta\to\infty}[\int_{\gamma_{1,\beta}} f(z)dz + \int_{\gamma_{2,\beta}} f(z)dz] \\ &= \frac{4\pi}{(4b)^n}\sqrt{\frac{b}{a}}\frac{(2n-1)!}{[(n-1)!]^2}.\end{aligned}$$

(2) Sei $0 < a < 1$. Dann ist

$$\int_0^{2\pi} \frac{\sin^2 x}{1-2a\cos x+a^2}dx = \pi.$$

DENN: Wir wollen den Integranden als $\int_{\partial B(0,1)} f(z)dz$ mit einer komplexen Funktion darstellen. Dies liefert für $f$ die Bedingung

$$\begin{aligned} f(e^{ix}) &= \frac{1}{ie^{ix}} \cdot \frac{\sin^2 x}{1 - 2a\cos x + a^2} \\ &= \frac{1}{ie^{ix}} \frac{-\frac{1}{4}(e^{ix} - e^{-ix})^2}{1 - a(e^{ix} + e^{-ix}) + a^2} \\ &= -\frac{1}{4i} e^{-ix} \frac{e^{2ix} - 2 + e^{-2ix}}{1 - a(e^{ix} + e^{-ix}) + a^2} \\ &= -\frac{1}{4i} \frac{e^{4ix} - 2e^{2ix} + 1}{e^{2ix}[-ae^{2ix} - a + (a^2+1)e^{ix}]} \\ &= \frac{1}{4ai} \frac{e^{4ix} - 2e^{2ix} + 1}{e^{2ix}[e^{2ix} - (a + \frac{1}{a})e^{ix} + 1]}. \end{aligned}$$

Sei also

$$\begin{aligned} f(z) &= \frac{1}{4ai} \frac{z^4 - 2z^2 + 1}{z^2[z^2 - (a + \frac{1}{a})z + 1]} \\ &= \frac{1}{4ai} \frac{z^4 - 2z^2 + 1}{z^2[z - a][z - \frac{1}{a}]}. \end{aligned}$$

$f$ hat die Polstellen 0, $a$ und $\frac{1}{a}$. Weiter ist

$$\begin{aligned} \operatorname{Res}_0 f &= \frac{d}{dz}(z^2 f(z))|_{z=0} \\ &= \frac{1}{4ai} \frac{(-1)}{(-a)^2(-\frac{1}{a})^2}(-a - \frac{1}{a}) \\ &= \frac{1}{4i}(1 + \frac{1}{a^2}) \\ \operatorname{Res}_a f &= \frac{1}{4ai} \frac{(a^2-1)^2}{a^2(a - \frac{1}{a})} \\ &= \frac{1}{4a^2 i}(a^2 - 1) \\ &= \frac{1}{4i}(1 - \frac{1}{a^2}). \end{aligned}$$

Damit folgt aus Satz VIII.4.14

$$\begin{aligned} \int_0^{2\pi} \frac{\sin^2 x}{1 - 2a\cos x + a^2} dx &= \int_{\partial B(0,1)} f(z)dz \\ &= 2\pi i[\operatorname{Res}_0 f + \operatorname{Res}_a f] \\ &= \pi. \end{aligned}$$

KAPITEL IX

# Integralrechnung mehrerer Veränderlicher

In diesem Kapitel entwickeln wir die Lebesguesche Integrationstheorie auf dem $\mathbb{R}^n$. Dazu führen wir zunächst den Begriff einer Nullmenge ein. Dies sind Mengen, die durch „beliebig kleine $n$-dimensionale Intervalle" überdeckt werden können. Danach definieren wir das Integral für Treppenfunktionen und setzen es fort auf Funktionen, deren positiver und negativer Teil bis auf eine Nullmenge als Grenzwert einer monoton wachsenden Folge von Treppenfunktionen dargestellt werden kann.
Diese Definition unterscheidet sich von derjenigen des Riemann-Integrals in Kapitel VI dadurch, dass die gleichmäßige Konvergenz durch punktweise monotone Konvergenz außerhalb einer Nullmenge ersetzt wird. Wir gehen dann auch kurz auf die Gemeinsamkeiten und Unterschiede des Riemann und Lebesgue Integrals ein.
Als nächstes untersuchen wir wichtige Eigenschaften des Lebesgue-Integrals: den Satz von Fubini, der die praktische Rechnung erleichtert, und die Konvergenzsätze.
Danach betrachten wir messbare Funktionen und Mengen. Dabei ist eine Menge, vereinfacht gesprochen, messbar, wenn ihre charakteristische Funktion messbar ist. Mit Hilfe dieses Begriffes können wir dann den wichtigen Transformationssatz beweisen.
Zum Abschluss untersuchen wir noch die $L^p$-Räume.

## IX.1. Nullmengen

Zuerst verallgemeinern wir den Begriff des Intervalles.

Definition IX.1.1. (1) Für $x, y \in \mathbb{R}$ schreiben wir $x \prec y$, wenn eine der Beziehungen $x < y$ oder $x \leq y$ gilt.
(2) Für $x, y \in \mathbb{R}^n$, $n \geq 1$, schreiben wir $x \prec y$, wenn für alle $1 \leq i \leq n$ gilt $x_i \prec y_i$.
(3) Für $x, y \in \mathbb{R}^n$, $n \geq 1$, heißt die Menge (vgl. Abb. IX.1.1)

$$\prec x, y \succ = \{z \in \mathbb{R}^n : x \prec z \prec y\}$$

ein beschränktes ($n$-dimensionales) Intervall. Gilt $x_i - y_i = x_1 - y_1$ für alle $1 \leq i \leq n$, so nennen wir es auch einen Würfel.
(4) Ist $I = \prec x, y \succ$ ein beschränktes Intervall, so heißen die $2n$ Hyperebenen

$$\{z \in I : z_i = x_i\} \quad , \quad \{z \in I : z_i = y_i\} \quad , \quad 1 \leq i \leq n$$

die Flächen von $I$.
(5) Seien $x, y \in \mathbb{R}^n$, $n \geq 1$, mit $x \prec y$. Dann ist das ($n$-dimensionale) Lebesgue-Mass oder Volumen von $\prec x, y \succ$ definiert durch

$$\lambda_n(\prec x, y \succ) = \prod_{i=1}^{n} (y_i - x_i).$$

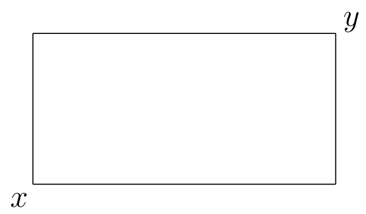


Abbildung IX.1.1. Zweidimensionales Intervall $\prec x, y \succ$

Bemerkung IX.1.2. (1) $I = \prec x, y \succ$ ist offen bzw. abgeschlossen genau dann, wenn in Definition IX.1.1 (3) $\prec$ immer für $<$ bzw. $\leq$ steht. Es gilt

$$\begin{aligned} \mathring{I} &= \{z \in \mathbb{R}^n : x_i < z_i < y_i \quad \forall 1 \leq i \leq n\} \\ \overline{I} &= \{z \in \mathbb{R}^n : x_i \leq z_i \leq y_i \quad \forall 1 \leq i \leq n\}. \end{aligned}$$

(2) Sei $I = \prec x, y \succ$ nicht leer. Dann gilt

$$\lambda_n(I) \neq 0 \quad \iff \quad \mathring{I} \neq \emptyset.$$

Beweis. Folgt unmittelbar aus der Definition. □

Der Begriff der Nullmenge ist für das Folgende grundlegend. Anschaulich ist eine Nullmenge dadurch charakterisiert, dass sie durch abzählbar viele Intervalle mit beliebig kleinem Maß überdeckt werden kann.

Definition IX.1.3. $N \subset \mathbb{R}^n$, $n \geq 1$, heißt eine Nullmenge, wenn es zu jedem $\varepsilon > 0$ eine Folge $(I_k)_{k \in \mathbb{N}}$ von beschränkten, offenen Intervallen gibt mit

(1) $N \subset \bigcup_{k \in \mathbb{N}} I_k$ und
(2) $\sum_{k \in \mathbb{N}} \lambda_n(I_k) \leq \varepsilon$.

Beispiel IX.1.4. (1) Sei $x \in \mathbb{R}^n$. Dann ist $\{x\}$ eine Nullmenge.
(2) Sei $I = \prec x, y \succ$ nichtleer. Dann sind die Flächen von $I$ Nullmengen.
(3) $\mathbb{R} \times \{0\} = \{(x, 0) : x \in \mathbb{R}\} \subset \mathbb{R}^2$ ist eine Nullmenge.

Beweis. ad (1): Sei $0 < \varepsilon < 1$ und

$$I_\varepsilon = \{z \in \mathbb{R}^n : \|z - x\|_\infty < \frac{\varepsilon}{2}\}$$

der Würfel mit Mittelpunkt $x$ und Kantenlänge $\varepsilon$. Dann ist

$$\lambda_n(I_\varepsilon) = \varepsilon^n < \varepsilon.$$

Hieraus folgt die Behauptung.
ad (2): Für $k \in \mathbb{N}^*$ sei

$$\begin{aligned} I_k &= (x_1 - \frac{1}{2k}, x_1 + \frac{1}{2k}) \times (x_2 - \frac{1}{2k}, y_2 + \frac{1}{2k}) \times \\ &\quad \cdots \times (x_n - \frac{1}{2k}, y_n + \frac{1}{2k}) \\ &= \{z \in \mathbb{R}^n : x_1 - \frac{1}{2k} < z_1 < x_1 + \frac{1}{2k}, \\ &\qquad x_j - \frac{1}{2k} < z_j < y_j + \frac{1}{2k},\ 2 \le j \le n\}. \end{aligned}$$

Dann gilt

$$\{z \in \mathbb{R}^n : z_1 = x_1, x_j \prec z_j \prec y_j, 2 \le j \le n\} \subset I_k$$

und

$$\lambda_n(I_k) = \frac{1}{k} \prod_{j=2}^n (y_j - x_j + \frac{1}{k}) \underset{k \to \infty}{\longrightarrow} 0.$$

Hieraus folgt die Behauptung.
ad (3): Sei $0 < \varepsilon < 1$ und $(q_k)_{k \in \mathbb{N}}$ eine Abzählung von $\mathbb{Q}$. Definiere

$$I_k = (q_k - \frac{\varepsilon}{2}, q_k + \frac{\varepsilon}{2}) \times (-\frac{\varepsilon}{2^{k+2}}, \frac{\varepsilon}{2^{k+2}}).$$

Da gemäß Satz I.4.11 (S. 21) $\mathbb{Q}$ dicht ist in $\mathbb{R}$ gilt

$$\mathbb{R} \times \{0\} \subset \bigcup_{k \in \mathbb{N}} I_k.$$

Weiter ist

$$\sum_{k \in \mathbb{N}} \lambda_2(I_k) = \sum_{k \in \mathbb{N}} \varepsilon^2 2^{-k-1} = \varepsilon^2 < \varepsilon.$$

Hieraus folgt die Behauptung. □

Teilmengen von Nullmengen und abzählbare Vereinigungen von Nullmengen sind wieder Nullmengen.

Satz IX.1.5. (1) *Sei $N \subset \mathbb{R}^n$, $n \ge 1$, eine Nullmenge und $M \subset N$. Dann ist $M$ auch eine Nullmenge.*
(2) *Seien $N_k \subset \mathbb{R}^n$, $n \ge 1$, $k \in \mathbb{N}$, Nullmengen. Dann ist $M = \bigcup_{k \in \mathbb{N}} N_k$ auch eine Nullmenge.*

Beweis. ad (1): Die Behauptung folgt unmittelbar aus der Beobachtung, dass jede Überdeckung von $N$ auch eine Überdeckung von $M$ ist.
ad (2): Sei $\varepsilon > 0$ und für $k \in \mathbb{N}$ $(I_{kl})_{l\in\mathbb{N}}$ eine Folge von beschränkten offenen Intervallen mit

$$N_k \subset \bigcup_{l\in\mathbb{N}} I_{kl} \quad \text{und} \quad \sum_{l\in\mathbb{N}} \lambda_n(I_{kl}) \leq \varepsilon 2^{-k-1}.$$

Dann gilt offensichtlich

$$M \subset \bigcup_{k,l\in\mathbb{N}} I_{kl} \quad \text{und} \quad \sum_{k,l\in\mathbb{N}} \lambda_n(I_{kl}) \leq \varepsilon \sum_{k\in\mathbb{N}} 2^{-k-1} = \varepsilon.$$

Da $\mathbb{N} \times \mathbb{N}$ abzählbar ist (vgl. Beweis von Satz I.3.2 (S. 17)), folgt die Behauptung. □

Aus Beispiel IX.1.4 (1) und Satz IX.1.5 folgt unmittelbar:

Bemerkung IX.1.6. Jede abzählbare Teilmenge von $\mathbb{R}^n$ ist eine Nullmenge. Insbesondere ist $\mathbb{Q} \subset \mathbb{R}$ eine Nullmenge in $\mathbb{R}$.

Das folgende Beispiel zeigt, dass die Umkehrung von Bemerkung IX.1.6 nicht gilt.

Beispiel IX.1.7 (Cantorsches Diskontinuum). Sei $A_0 = [0, 1] \subset \mathbb{R}$. $A_{n+1}$ entstehe aus $A_n$, $n \in \mathbb{N}$, indem man alle abgeschlossenen Intervalle, deren disjunkte Vereinigung $A_n$ ist, jeweils in drei abgeschlossene Intervalle gleicher Länge zerlegt und das Innere des mittleren Intervalles entfernt; also:

$$A_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$$
$$A_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1] \qquad \text{usw.}$$

Definiere

$$C = \bigcap_{k\in\mathbb{N}} A_k.$$

$C$ heißt Cantorsches Diskontinuum. Durch Induktion folgt sofort

$$\lambda_1(A_k) = (\frac{2}{3})^k \quad \forall k \in \mathbb{N}.$$

Da gilt $C \subset A_k$ für alle $k \in \mathbb{N}$, folgt hieraus, dass $C$ eine Nullmenge ist. Wir wollen zeigen, dass $C$ nicht abzählbar ist. Wie man leicht nachprüft, gilt

$$C = \Big\{ x \in \mathbb{R} : x = \sum_{n=1}^{\infty} a_n 3^{-n}, a_n \in \{0, 2\} \Big\}.$$

Wir nehmen an, $(c_k)_{k\in\mathbb{N}}$ sei eine Abzählung von $C$. Dann ist

$$c_k = \sum_{n=1}^{\infty} a_{kn} 3^{-n} \quad \forall k \in \mathbb{N}.$$

Definiere

$$\overline{c} = \sum_{n=1}^{\infty} \overline{a}_n 3^{-n}$$

durch

$$\overline{a}_n = \begin{cases} 0 & \text{falls } a_{nn} = 2 \\ 2 & \text{falls } a_{nn} = 0. \end{cases}$$

Dann ist $\overline{c} \in C$ und $\overline{c} \neq c_k$ für alle $k \in \mathbb{N}$. Dies ist der gewünschte Widerspruch.

Die folgende Definition ist für das Weitere wesentlich.

Definition IX.1.8. Sei $X \subset \mathbb{R}^n$, $n \geq 1$, eine Menge und $A(x)$, $x \in X$, eine Aussage über jedes $x \in X$. Wir sagen $A(x)$ gilt fast überall (f.ü.) in $X$, wenn es eine Nullmenge $N \subset \mathbb{R}^n$ gibt, so dass $A(x)$ für alle $x \in X \backslash N$ gilt.

Beispiel IX.1.9. Seien $X \subset \mathbb{R}^n$, $n \geq 1$, eine Menge und $f$, $f_k$, $k \in \mathbb{N}$, Funktionen auf $X$ mit Werten in $\mathbb{R}$. Dann konvergiert $(f_k)_{k \in \mathbb{N}}$ f.ü. (punktweise) gegen $f$, wenn es eine Nullmenge $N \subset \mathbb{R}^n$ gibt mit

$$f_k(x) \underset{k \to \infty}{\longrightarrow} f(x) \quad \forall x \in X \backslash N.$$

## IX.2. Das Lebesgue-Integral

Sei $X \subset \mathbb{R}^n$, $n \geq 1$, eine Menge und $f : X \to \mathbb{R}$ eine Funktion auf $X$. Dann setzen wir $f$ außerhalb von $X$ durch 0 fort. In diesem Sinne sind, sofern nicht anders vermerkt, alle Funktionen in diesem Abschnitt auf ganz $\mathbb{R}^n$ definiert.

Die folgende Definition von Treppenfunktionen verallgemeinert Definition V.4.1 (S. 166):

Definition IX.2.1. (1) Sei $X \subset \mathbb{R}^n$, $n \geq 1$. Dann ist die charakteristische Funktion $\chi_X$ von $X$ definiert durch

$$\chi_X(x) = \begin{cases} 1 & \text{falls } x \in X, \\ 0 & \text{falls } x \notin X. \end{cases}$$

(2) Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ heißt Treppenfunktion, wenn es ein $k \in \mathbb{N}^*$, $k$ beschränkte Intervalle $I_1, \ldots, I_k$ und $k$ reelle Zahlen $d_1, \ldots, d_k$ gibt mit

$$f = \sum_{i=1}^{k} d_i \chi_{I_i}.$$

(3) Die Menge der Treppenfunktionen bezeichnen wir mit $T(\mathbb{R}^n, \mathbb{R})$ oder kurz $T$.

Bemerkung IX.2.2. (1) Aus der Definition von $T$ folgt unmittelbar, dass $T$ ein Vektorraum ist.
(2) Die Darstellung einer Treppenfunktion als Linearkombination von endlich vielen beschränkten Intervallen ist nicht eindeutig. So ist z.B.

$$\chi_{[0,1]} = \chi_{[0,0.5]} + \chi_{(0.5,1]} = \chi_{[0,0.5]} + \chi_{[0.5,1]} - \chi_{[0.5,0.5]}.$$

Die folgenden beiden Sätze bereiten die Definition des Lebesgue-Integrals für Treppenfunktionen vor.

Satz IX.2.3. *Jede Treppenfunktion kann dargestellt werden als Linearkombination von endlich vielen charakteristischen Funktionen zu disjunkten, beschränkten Intervallen.*

Beweis. Sei

$$f = \sum_{i=1}^{k} d_i \chi_{I_i}$$

eine beliebige Treppenfunktion. Sei

$$E_j = \{x \in \mathbb{R}^n : x_j = 0\}, \quad 1 \le j \le n$$

die $j$-te Koordinatenhyperebene und $n_j$ die Zahl der verschiedenen Flächen der $I_i$, die parallel sind zu $E_j$, $1 \le j \le n$. Diese Flächen zerlegen den $\mathbb{R}^n$ in $\prod_{j=1}^{n}(2n_j + 1)$ disjunkte Intervalle, von denen $\prod_{j=1}^{n}(2n_j - 1)$ beschränkt sind. Jedes $I_i$, $1 \le i \le k$, ist die Vereinigung von endlich vielen dieser Intervalle, so dass $\chi_{I_i}$ die Summe der entsprechenden charakteristischen Funktionen ist. Hieraus folgt die Behauptung. $\square$

Satz IX.2.4. *Seien*

$$\sum_{i=1}^{k} c_i \chi_{I_i} = f = \sum_{j=1}^{l} d_j \chi_{J_j}$$

*zwei verschiedene Darstellungen der gleichen Treppenfunktion. Dann gilt*

$$\sum_{i=1}^{k} c_i \lambda_n(I_i) = \sum_{j=1}^{l} d_j \lambda_n(J_j).$$

Beweis. 1. Schritt: Seien $a, b \in \mathbb{R}$ und $I$ ein beschränktes Intervall. Dann ist offensichtlich

$$(a + b)\chi_I = a\chi_I + b\chi_I$$

und

$$(a + b)\lambda_n(I) = a\lambda_n(I) + b\lambda_n(I).$$

2. Schritt: Sei $I$ ein beschränktes Intervall und $H_i = \{x \in \mathbb{R}^n : x_i = \alpha\}$, $\alpha \in \mathbb{R}$, $1 \le i \le n$, eine Hyperebene mit $I \cap H_i \neq \emptyset$. Definiere (vgl. Abb. IX.2.1)

$$\begin{aligned} I_- &= \{x \in I : x_i < \alpha\} \\ I_0 &= I \cap H_i = \{x \in I : x_i = \alpha\} \\ I_+ &= \{x \in I : x_i > \alpha\}. \end{aligned}$$

Dann ist

$$\chi_I = \chi_{I_-} + \chi_{I_0} + \chi_{I_+}$$

und wegen $\lambda_n(I_0) = 0$ gilt

$$\lambda_n(I) = \lambda_n(I_-) + \lambda_n(I_+) + \lambda_n(I_0).$$

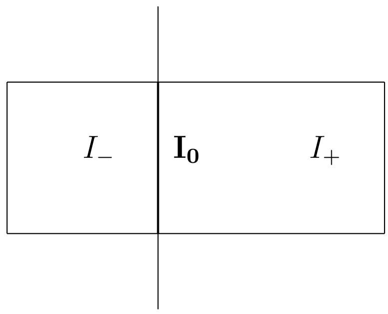


Abbildung IX.2.1. Aufteilung von $I \subset \mathbb{R}^2$ durch eine Hyperfläche $H_1$

3. Schritt: Seien $I_1, \ldots, I_k$, $J_1, \ldots, J_l$ wie im Satz. Wie im Beweis von Satz IX.2.3 zerlegen die Flächen von $I_1, \ldots, I_k$, $J_1, \ldots, J_l$ den $\mathbb{R}^n$ in disjunkte, beschränkte Intervalle, so dass wir eine gemeinsame Verfeinerung

$$f = \sum_{\mu=1}^{m} e_\mu \chi_{K_\mu}$$

von $f$ mit disjunkten Intervallen erhalten. Dieser Verfeinerungsprozess kann durchgeführt werden als eine Folge von Schritten 1 und 2. Da die Behauptung für diese Schritte gezeigt ist, folgt hieraus die behauptete Gleichheit. □

Wegen Satz IX.2.4 ist die folgende Definition sinnvoll.

Definition IX.2.5. Sei

$$f = \sum_{i=1}^{k} d_i \chi_{I_i}$$

eine Treppenfunktion. Dann ist das LEBESGUE-INTEGRAL von $f$ definiert durch

$$\int f = \int f dx = \int f(x)dx = \sum_{i=1}^{k} d_i \lambda_n(I_i).$$

Das Lebesgue-Integral hat folgende wichtige Eigenschaften:

SATZ IX.2.6. (1) *Seien $f, g \in T$ und $\alpha, \beta \in R$. Dann gilt*

(LINEARITÄT) $$\int (\alpha f + \beta g) = \alpha \int f + \beta \int g.$$

(2) *Sei $f \in T$ und $f \geq 0$. Dann ist*

(MONOTONIE) $$\int f \geq 0.$$

(3) *Sei $f \in T$. Dann ist $|f| \in T$ und*

$$|\int f| \leq \int |f|.$$

BEWEIS. AD (1): Folgt unmittelbar aus Definition IX.2.5 und Satz IX.2.4.
AD (2): Sei $f \in T$. Dann kann gemäß Satz IX.2.3 $f$ dargestellt werden als

$$f = \sum_{i=1}^{k} d_i \chi_{I_i}$$

mit disjunkten, beschränkten Intervallen. Aus $f \geq 0$ folgt dann

$$d_i \geq 0 \quad \forall 1 \leq i \leq k.$$

Hieraus folgt

$$\int f = \sum_{i=1}^{k} d_i \lambda_n(I_i) \geq 0.$$

AD (3): Sei wieder

$$f = \sum_{i=1}^{k} d_i \chi_{I_i}$$

mit disjunkten beschränkten Intervallen. Dann folgt

$$|f| = \sum_{i=1}^{k} |d_i| \chi_{I_i}.$$

Also ist $|f| \in T$. Weiter folgt

$$\begin{aligned} |\int f| &= |\sum_{i=1}^{k} d_i \lambda_n(I_i)| \\ &\leq \sum_{i=1}^{k} |d_i| \lambda_n(I_i) \end{aligned}$$

$$= \int |f|.$$

□

Die folgenden Sätze bereiten die Ausweitung des Integralbegriffes auf eine größere Funktionenklasse vor.

Satz IX.2.7. (1) *Sei $(f_k)_{k\in\mathbb{N}} \subset T$ eine monoton wachsende Folge von Treppenfunktionen, für die $(\int f_k)_{k\in\mathbb{N}} \subset \mathbb{R}$ konvergiert. Dann konvergiert $(f_k)_{k\in\mathbb{N}}$ fast überall in $\mathbb{R}^n$.*
(2) *Sei $N \subset \mathbb{R}^n$ eine Nullmenge. Dann gibt es eine monoton wachsende Folge $(g_k)_{k\in\mathbb{N}} \subset T$ von Treppenfunktionen, derart dass $(\int g_k)_{k\in\mathbb{N}}$ konvergiert und die Folge $(g_k(x))_{k\in\mathbb{N}}$ für jedes $x \in N$ divergiert.*

Beweis. ad (1): O.E. ist $f_k \geq 0$ für alle $k \in \mathbb{N}$. Sonst betrachte die Folge $(f_k - f_0)_{k\in\mathbb{N}}$. Dann gibt es ein $K > 0$ mit

$$0 \leq \int f_k \leq K \quad \forall k \in \mathbb{N}.$$

Sei $\varepsilon > 0$ beliebig. Definiere

$$S_k^\varepsilon = \{x \in \mathbb{R}^n : f_k(x) \geq \frac{K}{\varepsilon}\}.$$

Da die Folge $(f_k)_{k\in\mathbb{N}}$ monoton wachsend ist, gilt

$$S_k^\varepsilon \subset S_{k+1}^\varepsilon \quad \forall k \in \mathbb{N}.$$

Da $f_k$ eine Treppenfunktion ist, folgt weiter, dass $S_k^\varepsilon$ die Vereinigung von endlich vielen, disjunkten, beschränkten Intervallen ist. Daher ist

$$\lambda_n(S_k^\varepsilon) = \int \chi_{S_k^\varepsilon} \quad \forall k \in \mathbb{N}.$$

Wegen $f_k \geq 0$ gilt schließlich

$$\frac{K}{\varepsilon}\chi_{S_k^\varepsilon} \leq f_k \quad \forall k \in \mathbb{N}.$$

Damit folgt aus Satz IX.2.6

$$\frac{K}{\varepsilon}\lambda_n(S_k^\varepsilon) = \int \frac{K}{\varepsilon}\chi_{S_k^\varepsilon} \leq \int f_k \leq K$$

und somit

$$\lambda_n(S_k^\varepsilon) \leq \varepsilon \quad \forall k \in \mathbb{N}.$$

Definiere

$$S^\varepsilon = \bigcup_{k\in\mathbb{N}} S_k^\varepsilon \quad , \quad S = \bigcap_{\varepsilon>0} S^\varepsilon.$$

Dann ist $S$ die Menge aller derjenigen $x \in \mathbb{R}^n$, für die die Folge $(f_k(x))_{k\in\mathbb{N}}$ divergiert. Es reicht also zu zeigen, dass $S$ eine Nullmenge ist. Wegen $S_k^\varepsilon \subset S_{k+1}^\varepsilon$ für alle $k \in \mathbb{N}$ gilt

$$S^\varepsilon = S_0^\varepsilon \cup \bigcup_{k\in\mathbb{N}^*} (S_k^\varepsilon \backslash S_{k-1}^\varepsilon).$$

Also ist $S^\varepsilon$ die Vereinigung von abzählbar vielen beschränkten Intervallen. Da für $m \in \mathbb{N}^*$ gilt

$$S_m^\varepsilon = S_0^\varepsilon \cup \bigcup_{1 \le k \le m} (S_k^\varepsilon \backslash S_{k-1}^\varepsilon)$$

folgt

$$\lambda_n(S_0^\varepsilon \cup \bigcup_{1 \le k \le m} (S_k^\varepsilon \backslash S_{k-1}^\varepsilon)) \le \varepsilon \quad \forall m \in \mathbb{N}^*$$

und somit

$$\lambda_n(S^\varepsilon) \le \varepsilon.$$

Also ist $S$ eine Nullmenge.

AD (2): Da $N$ eine Nullmenge ist, gibt es zu jedem $k \in \mathbb{N}^*$ eine Folge $(I_{kl})_{l \in \mathbb{N}^*}$ von beschränkten offenen Intervallen mit

$$N \subset \bigcup_{l \in \mathbb{N}^*} I_{kl} \quad \text{und} \quad \sum_{l \in \mathbb{N}^*} \lambda_n(I_{kl}) \le 2^{-k}.$$

Wir ordnen diese Intervalle in einem quadratischen Schema an und zählen sie wie angedeutet ab (vgl. Beweis von Satz I.3.2 (S. 17)):

$$\begin{array}{ccccccc}
1 & & 2 & & 4 & & 7 \\
I_{11} & & I_{12} & & I_{13} & & \dots \\
 & & & & & & \\
3 & \swarrow & 5 & \swarrow & 8 & & \dots \\
I_{21} & & I_{22} & & I_{23} & & \dots \\
 & \swarrow & & & & & \\
6 & & 9 & & 13 & & \\
I_{31} & & I_{32} & & I_{33} & & \dots \\
\vdots & & \vdots & & \vdots & & \dots
\end{array}$$

Wir erhalten so eine Folge $(J_m)_{m \in \mathbb{N}^*}$ beschränkter offener Intervalle. Definiere

$$f_m = \sum_{i=1}^{m} \chi_{J_i} \quad \forall m \in \mathbb{N}^*.$$

Dann ist $(f_m)_{m \in \mathbb{N}^*}$ eine monoton wachsende Folge von Treppenfunktionen. Da für jedes $m \in \mathbb{N}^*$ gilt

$$\int f_m = \sum_{i=1}^{m} \lambda_n(J_i) \le \sum_{i=1}^{m} 2^{-i} < 1,$$

ist die Folge $(\int f_m)_{m \in N}$ konvergent. Sei nun $x \in N$ beliebig. Dann gibt es zu jedem $k \in \mathbb{N}^*$ ein $l_k \in \mathbb{N}^*$ mit

$$x \in I_{kl_k}.$$

Sei $K \in \mathbb{N}^*$ beliebig. Dann gibt es ein $m_K \in \mathbb{N}^*$, so dass unter $J_1, \dots,$ $J_{m_K}$ mindestens $K$ der $I_{kl_k}$ vorkommen. Das bedeutet aber, dass

$$f_{m_K}(x) \ge K$$

ist. Da $K \in \mathbb{N}^*$ beliebig war, folgt, dass $(f_k(x))_{k\in\mathbb{N}}$ für $x \in N$ divergiert. Damit ist die Behauptung gezeigt. □

Satz IX.2.7 legt folgende Definition nahe.

Definition IX.2.8. Mit $L^{\mathrm{inc}}(\mathbb{R}^n, \mathbb{R})$ oder kurz $L^{\mathrm{inc}}$ bezeichnen wir die Menge aller Funktionen $f : \mathbb{R}^n \to \mathbb{R}$, zu denen es eine monoton wachsende Folge $(f_k)_{k\in\mathbb{N}} \subset T$ von Treppenfunktionen gibt, derart dass $(\int f_k)_{k\in\mathbb{N}} \subset \mathbb{R}$ beschränkt ist und $f_k \underset{k\to\infty}{\longrightarrow} f$ f.ü. in $\mathbb{R}^n$ gilt.

Beispiel IX.2.9. Sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} 0 & \text{für } x \le 0, \\ \frac{1}{\sqrt{x}} & \text{für } 0 < x \le 1, \\ 0 & \text{für } x > 1. \end{cases}$$

Da $-f$ nicht nach unten beschränkt ist, folgt sofort $-f \notin L^{\mathrm{inc}}$. Um zu sehen, dass $f \in L^{\mathrm{inc}}$ ist, definiere für $k \in \mathbb{N}^*$

$$f_k = \sum_{l=1}^{2^k} \sqrt{\frac{2^k}{l}} \chi_{(\frac{l-1}{2^k}, \frac{l}{2^k}]}.$$

Dann ist $(f_k)_{k\in\mathbb{N}^*}$ eine monoton wachsende Folge von Treppenfunktionen, die auf $\mathbb{R}$ punktweise (aber nicht gleichmäßig!) gegen $f$ konvergiert. Weiter gilt für $k \in \mathbb{N}^*$

$$\begin{aligned}
\int f_k &= \sum_{l=1}^{2^k} \sqrt{\frac{2^k}{l}} 2^{-k} \\
&= \frac{1}{\sqrt{2^k}} \Big\{ 1 + \sum_{i=1}^{k} \sum_{j=2^{i-1}+1}^{2^i} \frac{1}{\sqrt{j}} \Big\} \\
&\le \frac{1}{\sqrt{2^k}} \Big\{ 1 + \sum_{i=1}^{k} \frac{2^i - 2^{i-1}}{\sqrt{2^{i-1}+1}} \Big\} \\
&\le \frac{1}{\sqrt{2^k}} \Big\{ 1 + \sum_{l=0}^{k-1} \sqrt{2}^l \Big\} \\
&\le \frac{1}{\sqrt{2^k}} \Big\{ 1 + \frac{\sqrt{2}^k - 1}{\sqrt{2} - 1} \Big\} \\
&\le 1 + \frac{1}{\sqrt{2} - 1} \\
&= \frac{\sqrt{2}}{\sqrt{2} - 1}.
\end{aligned}$$

SATZ IX.2.10. (1) *Sei $(f_k)_{k\in\mathbb{N}} \subset T$ eine monoton fallende Folge von Treppenfunktionen mit*

$$f_k \underset{k\to\infty}{\longrightarrow} 0 \text{ f.ü. in } \mathbb{R}^n.$$

*Dann gilt*

$$\lim_{k\to\infty} \int f_k = 0.$$

(2) *Seien $f, g \in L^{inc}$ fast überall bestimmt durch die monoton wachsenden Folgen $(f_k)_{k\in\mathbb{N}}$ bzw. $(g_k)_{k\in\mathbb{N}}$ von Treppenfunktionen. Gilt*

$$f \geq g \;\text{ f.ü. in } \mathbb{R}^n$$

*bzw.*

$$f = g \;\text{ f.ü. in } \mathbb{R}^n,$$

*so gilt*

$$\lim_{k\to\infty} \int f_k \geq \lim_{k\to\infty} \int g_k$$

*bzw.*

$$\lim_{k\to\infty} \int f_k = \lim_{k\to\infty} \int g_k.$$

BEWEIS. AD (1): Sei $\varepsilon > 0$ beliebig. Sei $I$ ein beschränktes, abgeschlossenes Intervall, so dass $f_0$ außerhalb von $I$ verschwindet. Sei $K > 0$, so dass gilt $f_0 \leq K$. Dann gilt gleiches für alle $f_k$, $k \in \mathbb{N}$.
Für jedes $k \in \mathbb{N}$ ist die Menge der Unstetigkeitsstellen von $f_k$ die Vereinigung von endlich vielen Flächen von beschränkten Intervallen und somit gemäß Beispiel IX.1.4(2) (S. 296) und Satz IX.1.5(2) (S. 297) eine Nullmenge. Sei $A$ die Vereinigung aller Unstetigkeitsstellen aller $f_k$. Gemäß Satz IX.1.5(2) (S. 297) ist $A$ eine Nullmenge. Sei $B$ die Menge aller Punkte $x$, für die $(f_k(x))_{k\in\mathbb{N}}$ nicht gegen 0 konvergiert. Nach Voraussetzung ist $B$ eine Nullmenge. Wegen Satz IX.1.5(2) (S. 297) ist dann $C = A \cup B$ eine Nullmenge und kann daher durch eine Folge $(I_k)_{k\in\mathbb{N}}$ offener beschränkter Intervalle, deren Gesamtmaß $\leq \varepsilon$ ist, überdeckt werden.
Sei nun $x \in I \backslash C$. Dann gibt es ein $k_x \in \mathbb{N}$ mit $f_{k_x}(x) \leq \frac{\varepsilon}{2}$. Da $x$ keine Unstetigkeitsstelle von $f_{k_x}$ ist, gibt es ein beschränktes offenes Intervall $J_x$, so dass gilt

$$f_{k_x}(y) \leq \varepsilon \quad \forall y \in J_x.$$

Da die Folge $(f_k)_{k\in\mathbb{N}}$ monoton fallend ist, gilt

$$f_k(y) \leq \varepsilon \quad \forall k \geq k_x,\; y \in J_x.$$

Da $I$ kompakt ist, gibt es Zahlen $r, s \in \mathbb{N}$, Punkte $x_1, \ldots, x_s$ aus $I \backslash C$ und Zahlen $k_1, \ldots, k_r$ aus $\mathbb{N}$ mit

$$I \subset \bigcup_{j=1}^{r} I_{k_j} \cup \bigcup_{j=1}^{s} J_{x_j}.$$

O.E. ist $s \geq 1$. Sei

$$M = \max_{1 \leq j \leq s} k_{x_j}.$$

Dann folgt

$$f_k(x) \leq \varepsilon \quad \forall k \geq M, x \in \bigcup_{j=1}^{s} J_{x_j}.$$

Sei

$$S = I \cap \bigcup_{j=1}^{r} I_{k_j} \quad , \quad T = I \cap \bigcup_{j=1}^{s} J_{x_j}.$$

$S$ und $T$ können als die Vereinigung von endlich vielen disjunkten Intervallen dargestellt werden. Weiter ist

$$\lambda_n(S) \leq \varepsilon \quad , \quad \lambda_n(T) \leq \lambda_n(I) = c.$$

Konstruktionsgemäß gilt

$$f_k \leq K\chi_S + \varepsilon\chi_T \quad \forall k \geq M.$$

Damit folgt

$$\int f_k \leq \varepsilon K + \varepsilon c \quad \forall k \geq M.$$

Da $\varepsilon$ beliebig war, folgt hieraus die Behauptung.

AD (2): Der zweite Teil der Behauptung folgt offensichtlich aus dem ersten Teil durch vertauschen von $f$ und $g$.

Zum Beweis des ersten Teiles der Behauptung sei $m \in \mathbb{N}$ beliebig. Dann ist $(g_m - f_k)_{k \in \mathbb{N}}$ eine monoton fallende Folge von Treppenfunktionen mit

$$g_m - f_k \underset{k \to \infty}{\longrightarrow} g_m - f \quad \text{f.ü.}$$

und

$$g_m - f \leq 0 \quad \text{f.ü.}$$

wegen $f \geq g$ f.ü.. Daher erfüllt

$$((g_m - f_k)_+)_{k \in \mathbb{N}} = (\max\{g_m - f_k, 0\})_{k \in \mathbb{N}}$$

die Voraussetzungen von Teil (1). Also gilt

$$\int g_m - \lim_{k \to \infty} \int f_k = \lim_{k \to \infty} \int (g_m - f_k) \leq \lim_{k \to \infty} \int (g_m - f_k)_+ = 0.$$

Da $m \in \mathbb{N}$ beliebig war, folgt die Behauptung durch Grenzübergang $m \to \infty$. □

Wegen Satz IX.2.10 ist die folgende Definition sinnvoll.

Definition IX.2.11. Sei $f \in L^{\text{inc}}$ und $(f_k)_{k\in\mathbb{N}} \in T$ eine monoton wachsende Folge von Treppenfunktionen mit $f_k \underset{k\to\infty}{\longrightarrow} f$ f.ü. und $(\int f_k)_{k\in\mathbb{N}}$ ist beschränkt. Dann definieren wir das Lebesgue-Integral von $f$ durch

$$\int f = \int f dx = \int f(x)dx = \lim_{k\to\infty} \int f_k.$$

Beispiel IX.2.9 zeigt, dass $L^{\text{inc}}$ kein Vektorraum ist. Um einen Vektorraum zu erhalten, müssen wir Funktionen der Form $g - h$ mit $g, h \in L^{\text{inc}}$ betrachten. Damit wir für solche Funktionen das Integral sinnvoll definieren können, benötigen wir folgenden Satz.

Satz IX.2.12. (1) *Seien $g, h \in L^{inc}$. Dann ist $g + h \in L^{inc}$ und*

$$\int (g + h) = \int g + \int h.$$

(2) *Seien $g_i, h_i \in L^{inc}$, $i = 1, 2$ mit*

$$g_1 - h_1 = g_2 - h_2.$$

*Dann gilt*

$$\int g_1 - \int h_1 = \int g_2 - \int h_2.$$

Beweis. ad (1): Seien $g, h \in L^{\text{inc}}$ und $(g_k)_{k\in\mathbb{N}}, (h_k)_{k\in\mathbb{N}} \subset T$ monoton wachsende Folgen von Treppenfunktionen mit:

$$g_k \underset{k\to\infty}{\longrightarrow} g \text{ f.ü.}\ ,\ h_k \underset{k\to\infty}{\longrightarrow} h \text{ f.ü.}$$
$$(\int g_k)_{k\in\mathbb{N}}\ ,\ (\int h_k)_{k\in\mathbb{N}} \text{ sind beschränkt.}$$

Dann ist $(g_k + h_k)_{k\in\mathbb{N}}$ eine monoton wachsende Folge von Treppenfunktionen mit:

$$g_k + h_k \underset{k\to\infty}{\longrightarrow} g + h \text{ f.ü.}$$
$$(\int (g_k + h_k))_{k\in\mathbb{N}} = (\int g_k + \int h_k)_{k\in\mathbb{N}} \text{ ist beschränkt}$$

und

$$\lim_{k\to\infty} \int (g_k + h_k) = \lim_{k\to\infty} \int g_k + \lim_{k\to\infty} \int h_k.$$

Hieraus folgt die Behauptung.
ad (2): Aus

$$g_1 - h_1 = g_2 - h_2$$

folgt

$$g_1 + h_2 = g_2 + h_1.$$

Aus Teil (1) folgt

$$\begin{aligned}\int g_1 + \int h_2 &= \int (g_1 + h_2)\\ &= \int (g_2 + h_1)\\ &= \int g_2 + \int h_1.\end{aligned}$$

Hieraus folgt die Behauptung. □

Wegen Satz IX.2.12 ist die folgende Definition sinnvoll. Damit sind wir am Ende unserer Konstruktion.

DEFINITION IX.2.13. (1) Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ heißt LEBESGUE-INTEGRIERBAR (auf $\mathbb{R}^n$), wenn es zwei Funktionen $g, h \in L^{\text{inc}}$ gibt mit $f = g - h$. In diesem Fall ist das LEBESGUE-INTEGRAL von $f$ definiert durch

$$\int f = \int f dx = \int f(x)dx = \int g - \int h.$$

Die Menge der Lebesgue-integrierbaren Funktionen bezeichnen wir mit $L^1(\mathbb{R}^n, \mathbb{R})$ oder kurz $L^1$.
(2) Sei $X \subset \mathbb{R}^n$ eine beliebige Teilmenge. Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ heißt LEBESGUE-INTEGRIERBAR AUF $X$, wenn $f \cdot \chi_X \in L^1$ ist. In diesem Fall ist das LEBESGUE-INTEGRAL VON $f$ AUF $X$ definiert durch

$$\int_X f = \int (f \cdot \chi_X).$$

Die Menge der Lebesgue-integrierbaren Funktionen auf $X$ bezeichnen wir mit $L^1(X, \mathbb{R})$.
(3) Eine Funktion $f = (f_1, \ldots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ heißt LEBESGUE-INTEGRIERBAR, wenn alle Komponentenfunktionen $f_k \in L^1$, $1 \le k \le m$, sind. In diesem Fall ist das LEBESGUE-INTEGRAL von $f$

$$\int f = (\int f_1, \ldots, \int f_m).$$

Die Menge der Lebesgue integrierbaren Funktionen $f : \mathbb{R}^n \to \mathbb{R}^m$ bezeichnen wir mit $L^1(\mathbb{R}^n, \mathbb{R}^m)$.
(4) Ist $f : \mathbb{C}^n \to \mathbb{C}^m$, so identifizieren wir auf kanonische Weise $\mathbb{C}^n$ mit $\mathbb{R}^{2n}$ und $\mathbb{C}^m$ mit $\mathbb{R}^{2m}$ und definieren auf diese Weise die Begriffe „Lebesgue-integrierbar“ und „Lebesgue-Integral“.

Für das Lebesgue-Integral gelten folgende wichtige Eigenschaften (vgl. Satz IX.2.6):

SATZ IX.2.14. (1) *Seien* $f_1, f_2 \in L^1$ *und* $\alpha_1, \alpha_2 \in \mathbb{R}$. *Dann ist* $\alpha_1 f_1 + \alpha_2 f_2 \in L^1$ *und*

(LINEARITÄT)
$$\int (\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 \int f_1 + \alpha_2 \int f_2.$$

(2) *Sei* $f \in L^1$ *und* $f \geq 0$ *f.ü.. Dann ist*

(Monotonie)
$$\int f \geq 0.$$

(3) *Sei* $f \in L^1$*. Dann ist* $|f| \in L^1$ *und*

$$|\int f| \leq \int |f|.$$

Beweis. ad (1): Die Behauptung folgt offensichtlich aus den folgenden drei Hilfsbehauptungen:

(i) $f_1, f_2 \in L^1 \Longrightarrow f_1 + f_2 \in L^1$ und $\int (f_1 + f_2) = \int f_1 + \int f_2$.
(ii) $f \in L^1, \alpha \in \mathbb{R}_+ \Longrightarrow \alpha f \in L^1$ und $\int (\alpha f) = \alpha \int f$.
(iii) $f \in L^1 \Longrightarrow -f \in L^1$ und $\int (-f) = - \int f$.

Behauptungen (i)–(iii) folgen aber direkt aus Satz IX.2.12 (1) und Definition IX.2.13.

ad (2): Sei $f = g - h \in L^1$ mit $g, h \in L^{\text{inc}}$. Dann folgt aus $f \geq 0$ f.ü. $g \geq h$ f.ü.. Damit folgt die Behauptung aus Satz IX.2.10 (2), Definition IX.2.11 und Definition IX.2.13.

ad (3): Die Behauptung folgt aus Teil (2) und den Ungleichungen

$$|f| - f \geq 0 \text{ f.ü. }, \ |f| + f \geq 0 \text{ f.ü..}$$

□

Satz IX.2.15. *Sei* $f_1 \in L^1$ *und* $f_2 = f_1$ *f.ü.. Dann ist* $f_2 \in L^1$ *und*

$$\int f_2 = \int f_1.$$

Beweis. Wegen Satz IX.2.10 (2) und Definition IX.2.11 ist die Behauptung für Funktionen aus $L^{\text{inc}}$ erfüllt. Sei weiter $f_1 = g_1 - h_1$ mit $g_1, h_1 \in L^{\text{inc}}$. Definiere

$$g_2 = g_1 \ , \ h_2 = h_1 + (f_1 - f_2).$$

Wegen $h_2 = h_1$ f.ü. folgt $g_2, h_2 \in L^{\text{inc}}$. Weiter ist

$$g_2 - h_2 = g_1 - h_1 - f_1 + f_2 = f_2.$$

Also ist $f_2 \in L^1$ und

$$\begin{aligned} \int f_2 &= \int g_2 - \int h_2 \\ &= \int g_1 - \int h_1 \quad \text{(wegen Satz IX.2.10 (2))} \\ &= \int f_1. \end{aligned}$$

□

Der folgende Satz gibt eine sehr praktische hinreichende Bedingung für die Lebesgue-Integrierbarkeit einer Funktion.

SATZ IX.2.16. (1) *Sei* $I \subset \mathbb{R}^n$ *ein beschränktes Intervall und* $f : \mathbb{R}^n \to \mathbb{R}$ *auf* $I$ *beschränkt und fast überall stetig. Dann ist* $f \in L^1(I, \mathbb{R})$.
(2) *Sei* $I \subset \mathbb{R}^n$ *ein beschränktes, abgeschlossenes Intervall und* $f \in C(I, \mathbb{R})$. *Dann ist* $f \in L^1(I, \mathbb{R})$.

BEWEIS. AD (1): O.E. können wir $f(x) = 0$ für alle $x \notin I$ annehmen. Wegen Satz IX.2.15 ist o.E. $I$ von der Form

$$(*) \qquad I = \{z \in \mathbb{R}^n : x_i \le z_i < y_i \quad \forall 1 \le i \le n\}.$$

Sei $K \in \mathbb{R}_+^*$ mit

$$|f(x)| \le K \quad \forall x \in I.$$

Wir konstruieren Intervalle $I_{kl}$, $k \in \mathbb{N}$, $1 \le l \le 2^{nk}$, wie folgt

(1) $I_{01} = I$.
(2) Die Intervalle $I_{k+1,l}$, $1 \le l \le 2^{n(k+1)}$, entstehen, indem man jedes Intervall $I_{k,m}$, $1 \le m \le 2^{nk}$, durch Halbieren der Kanten in $2^n$ gleich große disjunkte Intervalle der Form $(*)$ aufteilt.

Die Intervalle $I_{k,l}$, $1 \le l \le 2^{nk}$, sind dann für jedes $k \in \mathbb{N}$ paarweise disjunkt und erfüllen

$$I = \bigcup_{l=1}^{2^{nk}} I_{kl}.$$

Da $f$ beschränkt ist, existiert

$$c_{kl} = \inf_{x \in I_{kl}} f(x) \quad \forall k \in \mathbb{N}, 1 \le l \le 2^{nk}.$$

Definiere

$$f_k = \sum_{l=1}^{2^{nk}} c_{kl} \chi_{I_{kl}} \in T.$$

Dann ist $(f_k)_{k \in \mathbb{N}}$ eine monoton wachsende Folge von Treppenfunktionen mit

$$f_k \le f \le K\chi_I \quad \forall k \in \mathbb{N}.$$

Daher ist die Folge $(\int f_k)_{k \in \mathbb{N}}$ durch $K\lambda_n(I)$ nach oben beschränkt und damit konvergent. Sei nun $x \in I$, derart, dass $f$ in $x$ stetig ist. Sei $\varepsilon > 0$ beliebig. Dann gibt es ein $\delta > 0$ mit

$$|f(y) - f(x)| < \varepsilon \quad \forall y \in I \text{ mit } \|y - x\|_\infty < \delta.$$

Weiter gibt es ein $k_\varepsilon \in \mathbb{N}$ und ein $l_\varepsilon \in \mathbb{N}^*_{2^{nk_\varepsilon}}$ mit $I_{k_\varepsilon l_\varepsilon} \subset B_{\|\cdot\|_\infty}(x, \delta)$. Aus der Definition von $f_{k_\varepsilon}$ folgt dann

$$f(x) - \varepsilon \le f_{k_\varepsilon}(x) \le f(x).$$

Da $(f_k(x))_{k \in \mathbb{N}}$ monoton wachsend und durch $f(x)$ nach oben beschränkt ist, folgt hieraus

$$f(x) - \varepsilon \le f_k(x) \le f(x) \quad \forall k \ge k_\varepsilon.$$

Also konvergiert $(f_k(x))_{k \in \mathbb{N}}$ gegen $f(x)$. Da nach Voraussetzung $f$ auf $I$ f.ü. stetig ist, folgt hieraus $f \in L^{\text{inc}} \subset L^1$.

AD (2): Da $I$ gemäß Satz III.4.6 (S. 84) kompakt ist, ist wegen Satz III.4.9 (S. 85) und Satz III.4.3 (S. 82) $f$ beschränkt. Damit folgt die Behauptung aus Teil (1). $\square$

Zum Abschluss dieses Abschnittes wollen wir noch kurz auf den Zusammenhang zwischen dem Riemann- und dem Lebesgue-Integral für Funktionen auf $\mathbb{R}$ eingehen.

SATZ IX.2.17. (1) *Seien $a, b \in \mathbb{R}$ mit $a < b$ und $f \in S([a,b],\mathbb{R})$. Dann ist $f \in L^1([a,b],\mathbb{R})$ und das Riemann- und das Lebesgue-Integral von $f$ stimmen überein, d.h.*

$$\int_a^b f(x)dx = \int_{[a,b]} f.$$

(2) *Seien $a, b \in \overline{\mathbb{R}}$ mit $a < b$ und $f : (a,b) \to \mathbb{R}$ zulässig. Das uneigentliche Riemann-Integral $\int_a^b f(x)dx$ konvergiere absolut. Dann ist $f \in L^1((a,b),\mathbb{R})$ und das uneigentliche Riemann-Integral und das Lebesgue-Integral stimmen überein, d.h.*

$$\int_a^b f(x)dx = \int_{(a,b)} f.$$

BEWEIS. AD (1): Da in jedem Punkt $x \in [a,b]$ die links- und rechtsseitigen Grenzwerte $f(x-0)$ und $f(x+0)$ existieren, gilt für die Menge $S$ der Punkte, in denen $f$ unstetig ist,

$$S = \bigcup_{n\in\mathbb{N}^*} S_n$$

mit

$$S_n = \{x \in [a,b] : |f(x-0) - f(x+0)| > \frac{1}{n}\}.$$

Aus dem Beweis von Satz V.4.4 (S. 168) folgt aber, dass $S_n$ für jedes $n \in \mathbb{N}^*$ endlich, also eine Nullmenge ist. Wegen Satz IX.1.5(2) (S. 297) ist daher $S$ eine Nullmenge, und aus Satz IX.2.16 (1) folgt $f \in L^1([a,b],\mathbb{R})$. Wir benutzen nun die gleiche Notation wie im Beweis von Satz IX.2.16 (1). Da $f \in S([a,b],\mathbb{R})$ ist, gibt es zu jedem $I_{kl}$ ein $\zeta_{kl} \in I_{kl}$ mit $c_{kl} = f(\zeta_{kl})$. Also ist für jedes $k \in \mathbb{N}$

$$(**) \qquad \int_{[a,b]} f_k = \sum_{l=1}^{2^k} f(\zeta_{kl})\lambda_1(I_{kl}).$$

Die linke Seite von $(**)$ konvergiert gegen das Lebesgue-Integral $\int_{[a,b]} f$. Die rechte Seite von $(**)$ konvergiert nach Satz VI.1.11 (S. 177) gegen das Riemann-Integral $\int_a^b f(x)dx$. Hieraus folgt die Behauptung.
AD (2): Sei

$$f_+ = \max\{f, 0\} \quad , \quad f_- = \max\{-f, 0\}.$$

Dann ist

$$f = f_+ - f_- \quad , \quad |f| = f_+ + f_-.$$

Da die Funktion $x \mapsto \max\{x, 0\}$ stetig auf $\mathbb{R}$ ist, sind $f_+$ und $f_-$ zulässig. Seien $c \in (a, b)$ und $(a_k)_{k\in\mathbb{N}} \subset (a, c]$ streng monoton fallend und $(b_k)_{k\in\mathbb{N}} \subset [c, b)$ streng monoton wachsend mit

$$\lim_{k\to\infty} a_k = a \text{ und } \lim_{k\to\infty} b_k = b.$$

Da $\int_a^b f(x)dx$ absolut konvergiert, existieren die Grenzwerte

$$\lim_{k\to\infty} \int_{a_k}^c (f_+(x) - f_-(x))dx,$$
$$\lim_{k\to\infty} \int_{a_k}^c (f_+(x) + f_-(x))dx,$$
$$\lim_{k\to\infty} \int_c^{b_k} (f_+(x) - f_-(x))dx,$$
$$\lim_{k\to\infty} \int_c^{b_k} (f_+(x) + f_-(x))dx.$$

Daher existieren auch die Grenzwerte

$$\lim_{k\to\infty} \int_{a_k}^c f_+(x)dx,$$
$$\lim_{k\to\infty} \int_{a_k}^c f_-(x)dx,$$
$$\lim_{k\to\infty} \int_c^{b_k} f_+(x)dx,$$
$$\lim_{k\to\infty} \int_c^{b_k} f_-(x)dx.$$

Die Behauptung des Satzes folgt nun unmittelbar aus der folgenden Hilfsbehauptung.

Beh.: Es ist

$$f_+|_{(a,c]}, f_-|_{(a,c]} \in L^1((a, c], \mathbb{R})$$
$$f_+|_{[c,b)}, f_-|_{[c,b)} \in L^1([c, b), \mathbb{R})$$

und

$$\int_{(a,c]} f_\pm = \int_a^c f_\pm(x)dx = \lim_{k\to\infty} \int_{a_k}^c f_\pm(x)dx$$
$$\int_{[c,b)} f_\pm = \int_c^b f_\pm(x)dx = \lim_{k\to\infty} \int_a^{b_k} f_\pm(x)dx.$$

Wir zeigen diese Hilfsbehauptung nur für $f_+|_{(a,c]}$. Die anderen drei Fälle folgen ganz analog.

Seien $S$ und $S_k$ die Menge der Unstetigkeitsstellen von $f_+$ auf $(a, c]$

bzw. $[a_k, c]$. Dann ist $S = \bigcup\limits_{k \in \mathbb{N}} S_k$, und aus Teil (1) und Satz IX.1.5(2) (S. 297) folgt, dass $S$ eine Nullmenge ist. Also ist

$$f_+|_{(a,c]} \in L^1((a,c], \mathbb{R})$$

und

$$f_+|_{[a_k,c]} \in L^1([a_k, c], \mathbb{R}) \quad \forall k \in \mathbb{N}.$$

Wegen

$$f_+\chi_{[a_k,c]} \le f_+\chi_{(a,c]}$$

folgt aus Teil (1) und Satz IX.2.14

$$\int_{a_k}^{c} f_+(x)dx = \int_{[a_k,c]} f_+ \le \int_{(a,c]} f_+ \quad \forall k \in \mathbb{N}.$$

Also gilt

$$\int_{a}^{c} f_+(x)dx \le \int_{(a,c]} f_+.$$

Andererseits überlegt man sich leicht, dass es Zahlen $r_k \in \mathbb{N}$, $k \in \mathbb{N}$, und Intervalle $I_{kl}$, $1 \le l \le r_k$, $k \in \mathbb{N}$ mit folgenden Eigenschaften gibt:

(i) $I_{kl}$ ist von der Form $(x, y]$ für alle $k, l$.
(ii) Für festes $k$ sind die $I_{kl}$ paarweise disjunkt.
(iii) $(a_k, c] = \bigcup\limits_{l=1}^{r_k} I_{kl}$ für alle $k \in \mathbb{N}$.
(iv) Jedes $I_{k+1,l}$ ist entweder in $(a_{k+1}, a_k]$ oder in einem $I_{k,m}$ enthalten.
(v) $\lambda_1(I_{kl}) \le 2^{-k}$ für alle $k, l$.

Sei

$$c_{kl} = \inf_{x \in I_{kl}} f_+(x)$$

und

$$f_k = \sum_{l=1}^{r_k} c_{kl}\chi_{I_{kl}} \quad \forall k \in \mathbb{N}.$$

Dann ist $(f_k)_{k \in \mathbb{N}}$ eine monoton wachsende Folge von Treppenfunktionen, die fast überall gegen $f_+|_{(a,c]}$ konvergiert. Für jedes $k \in \mathbb{N}$ gilt

$$\begin{aligned}
\int_{(a,c]} f_k &= \int_{(a_k,c]} f_k \\
&= \int_{[a_k,c]} f_k \\
&\le \int_{[a_k,c]} f_+ \\
&= \int_{a_k}^{c} f_+(x)dx.
\end{aligned}$$

Also ist

$$\int_{(a,c]} f_+ \leq \int_a^c f_+(x)dx.$$

Damit ist die Hilfsbehauptung für $f_+|_{(a,c]}$ gezeigt. □

Das folgende Beispiel zeigt, dass es Funktionen gibt, die Lebesgue- aber nicht Riemann-integrierbar sind, und dass es auch umgekehrt Funktionen gibt, die (uneigentlich) Riemann- aber nicht Lebesgue-integrierbar sind.

Beispiel IX.2.18. (1) Es ist $\chi_{\mathbb{Q}\cap[0,1]} \in L^1([0,1],\mathbb{R})$ und

$$\int \chi_{\mathbb{Q}\cap[0,1]} = 0,$$

da $\mathbb{Q}\cap[0,1]$ eine Nullmenge ist.
$\chi_{\mathbb{Q}\cap[0,1]}$ ist nicht Riemann-integrierbar.
Denn sei $\delta > 0$ beliebig und $0 = x_0 < x_1 < \ldots < x_n = 1$ eine beliebige Zerlegung von $[0,1]$ der Feinheit $\delta$. Sei

$$\zeta_k \in \mathbb{Q}\cap(x_{k-1},x_k), \eta_k \in (\mathbb{R}\backslash\mathbb{Q})\cap(x_{k-1},x_k),\ 1 \leq k \leq n$$

beliebig. Dann folgt

$$\sum_{k=1}^{n} \chi_{\mathbb{Q}\cap[0,1]}(\zeta_k)(x_k - x_{k-1}) = 1$$
$$\sum_{k=1}^{n} \chi_{\mathbb{Q}\cap[0,1]}(\eta_k)(x_k - x_{k-1}) = 0.$$

(2) Sei $f$ auf $[1,\infty)$ definiert durch

$$f(x) = \frac{\sin x}{x}.$$

Gemäß Übungsaufgabe (Aufgabe 2, Blatt 4, Analysis II) ist $f$ uneigentlich Riemann-integrierbar, aber das uneigentliche Riemann-Integral konvergiert nicht absolut.
$f$ ist nicht Lebesgue-integrierbar.
Denn wäre $f \in L^1([1,\infty),\mathbb{R})$, so wäre gemäß Satz IX.2.14 auch $|f| \in L^1([1,\infty),\mathbb{R})$. Dann gilt aber für jede Folge $(b_k)_{k\in\mathbb{N}} \subset [1,\infty)$ mit $\lim\limits_{k\to\infty} b_k = \infty$

$$\int_1^{b_k} |f(x)|dx = \int_{[1,b_k]} |f| < \int_{[1,\infty)} |f| < \infty.$$

Dies ist ein Widerspruch.

## IX.3. Eigenschaften des Lebesgue-Integrals

Zunächst wollen wir den Satz von Fubini beweisen. Er führt die Berechnung eines Lebesgue-Integrales über $\mathbb{R}^n$ zurück auf die Berechnung einer Folge geeigneter Lebesgue-Integrale über $\mathbb{R}$. Letztere können wegen Satz IX.2.17 (S. 312) in der Regel mit den Methoden aus Kapitel VI berechnet werden.

Im Folgenden seien $m, n \in \mathbb{N}^*$. Wir identifizieren auf kanonische Weise den $\mathbb{R}^{m+n}$ mit $\mathbb{R}^m \times \mathbb{R}^n$; jedes $x \in \mathbb{R}^{m+n}$ zerlegen wir in der Form $x = (x_1, x_2)$ mit $x_1 \in \mathbb{R}^m$, $x_2 \in \mathbb{R}^n$.

Lemma IX.3.1. *Sei $f \in T(\mathbb{R}^{m+n}, \mathbb{R})$. Dann ist für jedes $x_1 \in \mathbb{R}^m$ $f(x_1, .) \in T(\mathbb{R}^n, \mathbb{R})$. Die Funktion*

$$g : \mathbb{R}^m \to \mathbb{R} \text{ mit } x_1 \mapsto \int_{\mathbb{R}^n} f(x_1, x_2) dx_2$$

*ist aus $T(\mathbb{R}^m, \mathbb{R})$. Es gilt*

$$\int_{\mathbb{R}^{m+n}} f = \int_{\mathbb{R}^m} g = \int_{\mathbb{R}^m} \Big( \int_{\mathbb{R}^n} f(x_1, x_2) dx_2 \Big) dx_1.$$

Beweis. Sei zunächst $f = \chi_I$ mit einem nicht leeren, beschränkten Intervall $I$. Sei

$$\begin{aligned} I &= \prec a, b \succ, \ a, b \in \mathbb{R}^{m+n}, \\ I_1 &= \{x \in \mathbb{R}^m : a_i \prec x_i \prec b_i, 1 \le i \le m\}, \\ I_2 &= \{y \in \mathbb{R}^n : a_{m+j} \prec y_j \prec b_{m+j}, 1 \le j \le n\}. \end{aligned}$$

Dann ist

$$\chi_I(x_1, x_2) = \chi_{I_1}(x_1) \cdot \chi_{I_2}(x_2).$$

Daher ist für jedes $x_1 \in \mathbb{R}^m$ $f(x_1, .) \in T(\mathbb{R}^n, \mathbb{R})$ und

$$g(x_1) = \chi_{I_1}(x_1) \lambda_n(I_2)$$

sowie

$$\begin{aligned} \int_{\mathbb{R}^m} g &= \lambda_n(I_2) \int_{\mathbb{R}^m} \chi_{I_1} \\ &= \lambda_n(I_2) \lambda_m(I_1) \\ &= \prod_{j=1}^{n} (b_{m+j} - a_{m+j}) \prod_{i=1}^{m} (b_i - a_i) \\ &= \lambda_{m+n}(I) \\ &= \int_{\mathbb{R}^{m+n}} f. \end{aligned}$$

Hieraus folgt die Behauptung im allgemeinen Fall zusammen mit Satz IX.2.6 (S. 302). □

Lemma IX.3.2. *Sei $N \subset \mathbb{R}^{m+n}$ eine Nullmenge. Dann ist für fast alle $x_1 \in \mathbb{R}^m$ die Menge*

$$N_{x_1} = \{x_2 \in \mathbb{R}^n : (x_1, x_2) \in N\}$$

*eine Nullmenge in $\mathbb{R}^n$.*

Beweis. Gemäß Satz IX.2.7(2) (S. 303) gibt es eine monoton wachsende Folge $(f_k)_{k\in\mathbb{N}} \subset T(\mathbb{R}^{m+n}, \mathbb{R})$ von Treppenfunktionen, derart dass $(\int_{\mathbb{R}^{m+n}} f_k)_{k\in\mathbb{N}}$ beschränkt ist und $(f_k(x))_{k\in\mathbb{N}}$ für alle $x \in N$ divergiert. Gemäß Lemma IX.3.1 wird für jedes $k \in \mathbb{N}$ durch

$$g_k(x_1) = \int_{\mathbb{R}^n} f_k(x_1, x_2) dx_2 \quad \forall x_1 \in \mathbb{R}^m$$

eine Treppenfunktion auf $\mathbb{R}^m$ definiert. Wie man sich leicht überzeugt, ist die Folge $(g_k)_{k\in\mathbb{N}} \subset T(\mathbb{R}^m, \mathbb{R})$ monoton wachsend. Aus Lemma IX.3.1 folgt, dass $(\int_{\mathbb{R}^m} g_k)_{k\in\mathbb{N}} = (\int_{\mathbb{R}^{m+n}} f_k)_{k\in\mathbb{N}}$ beschränkt ist. Wegen Satz IX.2.7(1) (S. 303) konvergiert daher $(g_k)_{k\in\mathbb{N}}$ f.ü. in $\mathbb{R}^m$. Sei $x_1^* \in \mathbb{R}^m$ so, dass $(g_k(x_1^*))_{k\in\mathbb{N}} = (\int_{\mathbb{R}^n} f(x_1^*, x_2) dx_2)_{k\in\mathbb{N}}$ konvergiert. Aus Satz IX.2.7(1) (S. 303) folgt, dass $(f_k(x_1^*, .))_{k\in\mathbb{N}}$ f.ü. auf $\mathbb{R}^n$ konvergiert. Da aber $(f_k(x_1^*, x_2))_{k\in\mathbb{N}}$ für alle $x_2 \in N_{x_1^*}$ divergiert, folgt hieraus, dass $N_{x_1^*}$ eine Nullmenge in $\mathbb{R}^n$ ist. □

Satz IX.3.3 (Satz von Fubini). *Sei $f \in L^1(\mathbb{R}^{m+n}, \mathbb{R})$. Dann ist für fast jedes $x_1 \in \mathbb{R}^m$ und fast jedes $x_2 \in \mathbb{R}^n$ $f(x_1, \cdot) \in L^1(\mathbb{R}^n, \mathbb{R})$ und $f(\cdot, x_2) \in L^1(\mathbb{R}^m, \mathbb{R})$. Die Funktionen $g : \mathbb{R}^m \to \mathbb{R}$ und $h : \mathbb{R}^n \to \mathbb{R}$ mit*

$$g(x_1) = \int_{\mathbb{R}^n} f(x_1, x_2) dx_2 \quad \textit{für fast jedes } x_1 \in \mathbb{R}^m$$
$$h(x_2) = \int_{\mathbb{R}^m} f(x_1, x_2) dx_1 \quad \textit{für fast jedes } x_2 \in \mathbb{R}^n$$

*sind aus $L^1(\mathbb{R}^m, \mathbb{R})$ bzw. $L^1(\mathbb{R}^n, \mathbb{R})$. Es gilt*

$$\begin{aligned}\int_{\mathbb{R}^{m+n}} f &= \int_{\mathbb{R}^m} g \\ &= \int_{\mathbb{R}^m} \Big( \int_{\mathbb{R}^n} f(x_1, x_2) dx_2 \Big) dx_1 \\ &= \int_{\mathbb{R}^n} h \\ &= \int_{\mathbb{R}^n} \Big( \int_{\mathbb{R}^m} f(x_1, x_2) dx_1 \Big) dx_2.\end{aligned}$$

Beweis. 1. Schritt: Es genügt, die Behauptung für $f(x_1, \cdot)$ und $g$ zu zeigen, da dann der Rest durch Vertauschen der Rollen von $\mathbb{R}^m$ und $\mathbb{R}^n$ folgt.
2. Schritt: Die Aussage gilt gemäß Lemma IX.3.1 für alle Treppenfunktionen.
3. Schritt: Sei nun $f \in L^{\mathrm{inc}}(\mathbb{R}^{m+n}, \mathbb{R})$ und $(f_k)_{k\in\mathbb{N}} \subset T(\mathbb{R}^{m+n}, \mathbb{R})$

eine monoton wachsende Folge von Treppenfunktionen, derart dass $(\int_{\mathbb{R}^{m+n}} f_k)_{k\in\mathbb{N}}$ konvergiert und $f_k \underset{k\to\infty}{\longrightarrow} f$ f.ü. in $\mathbb{R}^{m+n}$ gilt. Für $k \in \mathbb{N}$ sei

$$g_k(x_1) = \int_{\mathbb{R}^n} f_k(x_1, x_2)dx_2 \quad \forall x_1 \in \mathbb{R}^m.$$

Gemäß Lemma IX.3.1 ist $g_k$ wohldefiniert und aus $T(\mathbb{R}^m, \mathbb{R})$. Außerdem ist $(g_k)_{k\in\mathbb{N}}$ monoton wachsend und

$$(\int_{\mathbb{R}^m} g_k)_{k\in\mathbb{N}} = (\int_{\mathbb{R}^{m+n}} f_k)_{k\in\mathbb{N}}$$

konvergent.
Gemäß Satz IX.2.7(1) (S. 303) konvergiert $(g_k)_{k\in\mathbb{N}}$ f.ü. in $\mathbb{R}^m$. Sei $N$ die Menge aller $x \in \mathbb{R}^{m+n}$, für die $(f_k(x))_{k\in\mathbb{N}}$ divergiert. Dann ist $N$ eine Nullmenge. Sei $x_1^* \in \mathbb{R}^m$ ein Punkt, für den $(g_k(x_1^*))_{k\in\mathbb{N}}$ konvergiert und $N_{x_1^*}$ eine Nullmenge ist. Wegen Lemma IX.3.2 gilt dies für fast jedes $x_1^*$ in $\mathbb{R}^m$. Dann gilt

$$f_k(x_1^*, x_2) \underset{k\to\infty}{\longrightarrow} f(x_1^*, x_2) \quad \text{f.ü. in } \mathbb{R}^m$$

und

$$\Big(\int_{\mathbb{R}^n} f_k(x_1^*, x_2)dx_2\Big)_{k\in\mathbb{N}} = (g_k(x_1^*))_{k\in\mathbb{N}}$$

konvergiert. Also existiert $\int_{\mathbb{R}^n} f(x_1^*, x_2)dx_2$ und ist gleich $\lim_{k\to\infty} g_k(x_1^*)$. Mithin gilt $g_k \underset{k\to\infty}{\longrightarrow} g$ f.ü. in $\mathbb{R}^m$ und

$$\lim_{k\to\infty} \int_{\mathbb{R}^m} g_k(x_1)dx_1 = \lim_{k\to\infty} \int_{\mathbb{R}^{m+n}} f_k = \int_{\mathbb{R}^{m+n}} f.$$

Also ist $g \in L^{\text{inc}}(\mathbb{R}^m, \mathbb{R})$ und

$$\int_{\mathbb{R}^m} g = \int_{\mathbb{R}^{m+n}} f.$$

4. Schritt: Sei nun $f = \varphi - \psi$ mit $\varphi, \psi \in L^{\text{inc}}(\mathbb{R}^{m+n}, \mathbb{R})$. Dann folgt die Behauptung aus dem dritten Schritt zusammen mit der Linearität des Integrals. □

Beispiel IX.3.4. (1) Sei $B_R \subset \mathbb{R}^2$ der abgeschlossene Kreis um Null mit Radius $R > 0$. Da $\partial B_R$ eine Nullmenge in $\mathbb{R}^2$ ist (Übungsaufgabe), ist gemäß Satz IX.2.16(1) (S. 311) $\chi_{B_R} \in L^1(\mathbb{R}^2, \mathbb{R})$. Mit dem Satz von Fubini folgt

$$\begin{aligned}
\int_{\mathbb{R}^2} \chi_{B_R} &= \int_{-R}^{R} \Big(\int_{-\sqrt{R^2-x_1^2}}^{\sqrt{R^2-x_1^2}} 1 dx_2\Big) dx_1 \\
&= 2\int_{-R}^{R} \sqrt{R^2 - x_1^2} dx_1 \\
&= 2\int_0^{\pi} \sqrt{R^2 - R^2\cos^2\varphi} R\sin\varphi d\varphi \quad x_1 = R\cos\varphi
\end{aligned}$$

$$= 2R^2 \int_0^\pi \sin^2 \varphi d\varphi$$
$$= \pi R^2.$$

(2) Für $\sigma \in \mathbb{R}$ sei

$$\operatorname{sgn}(\sigma) = \begin{cases} 1 & \text{falls } \sigma > 0, \\ 0 & \text{falls } \sigma = 0, \\ -1 & \text{falls } \sigma < 0. \end{cases}$$

Sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x_1, x_2) = \begin{cases} e^{x_1+x_2} \operatorname{sgn}(x_1 - x_2) & \text{falls } (x_1, x_2) \in [-1, 1] \times [-1, 1], \\ 0 & \text{sonst.} \end{cases}$$

Da $\{x \in \mathbb{R}^2 : |x_1| \leq 1, |x_2| \leq 1 \text{ und } x_1 = x_2\}$ eine Nullmenge ist (Übungsaufgabe), ist $f \in L^1(\mathbb{R}^2, \mathbb{R})$. Der Satz von Fubini liefert

$$\begin{aligned}
\int_{\mathbb{R}^2} f &= \int_{-1}^1 \Big( \int_{-1}^1 e^{x_1+x_2} \operatorname{sgn}(x_1 - x_2) dx_2 \Big) dx_1 \\
&= \int_{-1}^1 \Big\{ \int_{-1}^{x_1} e^{x_1+x_2} dx_2 - \int_{x_1}^1 e^{x_1+x_2} dx_2 \Big\} dx_1 \\
&= \int_{-1}^1 \{ e^{2x_1} - e^{x_1-1} - e^{x_1+1} + e^{2x_1} \} dx_1 \\
&= [e^{2x_1} - e^{x_1-1} - e^{x_1+1}]_{x_1=-1}^{x_1=+1} \\
&= e^2 - 1 - e^2 - (e^{-2} - e^{-2} - 1) \\
&= 0.
\end{aligned}$$

Als nächstes untersuchen wir, wie sich das Lebesgue-Integral bei Grenzprozessen verhält, d.h., wann Grenzprozesse mit der Integration vertauscht werden können.

Lemma IX.3.5. *Sei $(f_k)_{k \in \mathbb{N}} \subset L^{inc}(\mathbb{R}^n, \mathbb{R})$ eine monoton wachsende Folge, derart dass $(\int f_k)_{k \in \mathbb{N}}$ beschränkt ist. Dann konvergiert $f_k$ fast überall gegen eine Funktion $f \in L^{inc}$. Es gilt*

$$\int f = \lim_{k \to \infty} \int f_k.$$

Beweis. Für $k \in \mathbb{N}$ sei $(g_{kl})_{l \in \mathbb{N}} \subset T$ eine monoton wachsende Folge von Treppenfunktionen mit

$$g_{kl} \underset{l \to \infty}{\longrightarrow} f_k \quad \text{f.ü. und} \quad \Big(\int g_{kl}\Big)_{l \in \mathbb{N}} \quad \text{ist beschränkt.}$$

Dann ist insbesondere

$$\int f_k = \lim_{l \to \infty} \int g_{kl} \quad \forall k \in \mathbb{N}.$$

Sei

$$h_m = \max\{g_{kl} : 0 \leq k, l \leq m\} \quad \forall m \in \mathbb{N}.$$

Dann ist $(h_m)_{m\in\mathbb{N}}$ eine monoton wachsende Folge von Treppenfunktionen. Weiter gilt

$$h_m \leq f_m \quad \text{f.ü.}$$

und damit

$$\int h_m \leq \sup_{k\in\mathbb{N}} \int f_k = K < \infty \quad \forall m \in \mathbb{N}.$$

Wegen Satz IX.2.7 (S. 303) konvergiert $h_m$ fast überall gegen eine Funktion $f \in L^{\text{inc}}$ und

$$\int f = \lim_{m\to\infty} \int h_m.$$

Konstruktionsgemäß gilt für $m \in \mathbb{N}$

$$g_{ml} \leq h_l \quad \forall l \geq m$$

und somit

$$f_m = \lim_{l\to\infty} g_{ml} \leq \lim_{l\to\infty} h_l = f \quad \text{f.ü.},$$

d.h.

$$S_m = \{x \in \mathbb{R}^n : f_m(x) > f(x)\}$$

ist eine Nullmenge. Daher ist

$$S = \bigcup_{m\in\mathbb{N}} S_m$$

auch eine Nullmenge, d.h., es gilt fast überall

$$f_m \leq f \quad \forall m \in \mathbb{N}.$$

Also gilt fast überall

$$h_m \leq f_m \leq f \quad \forall m \in \mathbb{N}.$$

Wegen $h_m \underset{m\to\infty}{\longrightarrow} f$ f.ü. folgt $f_m \underset{m\to\infty}{\longrightarrow} f$ f.ü.. Ebenso folgt aus der Monotonie des Integrals

$$\int h_m \leq \int f_m \leq \int f \quad \forall m \in \mathbb{N}.$$

Wegen $\displaystyle\int f = \lim_{m\to\infty} \int h_m$ folgt hieraus schließlich

$$\int f = \lim_{m\to\infty} \int f_m.$$

□

LEMMA IX.3.6. *Sei $f \in L^1$. Dann gibt es zu jedem $\varepsilon > 0$ Funktionen $g, h \in L^{inc}$ mit*

(1) $f = g - h$,
(2) $h \geq 0$,
(3) $\int h < \varepsilon$.

BEWEIS. Konstruktionsgemäß gibt es Funktionen $g_1, h_1 \in L^{\text{inc}}$ mit $f = g_1 - h_1$. Sei $(\varphi_k)_{k\in\mathbb{N}} \subset T$ eine monoton wachsende Folge von Treppenfunktionen mit $\varphi_k \underset{k\to\infty}{\longrightarrow} h_1$ f.ü. und $\int \varphi_k \underset{k\to\infty}{\longrightarrow} \int h_1$. Sei $\varepsilon > 0$ beliebig. Dann gibt es ein $k_\varepsilon \in \mathbb{N}$ mit

$$0 \le \int h_1 - \int \varphi_{k_\varepsilon} < \varepsilon.$$

Dann ist $h_2 = h_1 - \varphi_{k_\varepsilon}$ fast überall nicht negativ und aus $L^{\text{inc}}$ und erfüllt $\int h_2 < \varepsilon$. Ebenso ist $g_2 = g_1 - \varphi_{k_\varepsilon}$ aus $L^{\text{inc}}$ und $f = g_2 - h_2$. Sei schließlich $h = h_2^+ = \max\{h_2, 0\}$ und $g = f + h$. Wegen $h_2 \ge 0$ f.ü. gilt $g = g_2$ f.ü. und $h = h_2$ f.ü.. Daher erfüllen $g$ und $h$ die Bedingungen (1)–(3). □

SATZ IX.3.7 (SATZ VON DER MONOTONEN KONVERGENZ). *Sei $(f_k)_{k\in\mathbb{N}} \subset L^1(\mathbb{R}^n, \mathbb{R})$ eine monotone Folge, derart dass die Folge der Integrale $(\int f_k)_{k\in\mathbb{N}}$ beschränkt ist. Dann konvergiert $f_k$ f.ü. gegen eine Funktion $f \in L^1(\mathbb{R}^n, \mathbb{R})$ und*

$$\int f = \lim_{k\to\infty} \int f_k.$$

BEWEIS. O.E. ist $(f_k)_{k\in\mathbb{N}}$ monoton wachsend, sonst betrachte $(-f_k)_{k\in\mathbb{N}}$. O.E. ist $f_k \ge 0$ f.ü. für alle $k \in \mathbb{N}^*$, sonst betrachte $(f_k - f_0)_{k\in\mathbb{N}}$. Definiere

$$g_1 = f_1, \quad g_k = f_k - f_{k-1} \quad \forall k \ge 2.$$

Dann gilt

$$g_k \ge 0 \quad \text{f.ü.} \quad \forall k \in \mathbb{N}^*$$

und

$$f_k = \sum_{l=1}^{k} g_l \quad \forall k \in \mathbb{N}.$$

Lemma IX.3.6 angewandt auf $g_k$, $k \in \mathbb{N}^*$, mit $\varepsilon = 2^{-k}$ liefert die Existenz von Funktionen $\varphi_k, \psi_k \in L^{\text{inc}}$ mit

$$g_k = \varphi_k - \psi_k \quad \forall k \in \mathbb{N}^*$$
$$\psi_k \ge 0 \quad \forall k \in \mathbb{N}^*$$
$$\varphi_k \ge 0 \quad \forall k \in \mathbb{N}^*$$
$$\int \psi_k < 2^{-k} \quad \forall k \in \mathbb{N}^*.$$

(Man beachte, dass $\varphi_k \ge 0$ aus den entsprechenden Ungleichungen für $\psi_k$ und $g_k$ folgt.)
Definiere für $k \in \mathbb{N}^*$.

$$u_k = \sum_{l=1}^{k} \varphi_l \in L^{\text{inc}},$$

$$v_k = \sum_{l=1}^{k} \psi_l \in L^{\mathrm{inc}}.$$

Dann sind die Folgen $(u_k)_{k\in\mathbb{N}^*}$, $(v_k)_{k\in\mathbb{N}^*}$ monoton wachsend und

$$\begin{aligned}
\int v_k &\leq \sum_{l=1}^{k} 2^{-l} < 1, \quad \forall k \in \mathbb{N}^* \\
\int u_k &= \sum_{l=1}^{k} \int (g_l + \psi_l) \\
&= \int f_k + \int v_k \\
&< 1 + \sup_{l\in\mathbb{N}^*} \int f_l \\
&< \infty \quad \forall k \in \mathbb{N}^*.
\end{aligned}$$

Gemäß Lemma IX.3.5 gibt es daher Funktionen $u, v \in L^{\mathrm{inc}}$ mit

$$\begin{aligned}
u_k &\xrightarrow[k\to\infty]{} u \text{ f.ü.}, \\
v_k &\xrightarrow[k\to\infty]{} v \text{ f.ü.}
\end{aligned}$$

und

$$\begin{aligned}
\int u_k &\xrightarrow[k\to\infty]{} \int u, \\
\int v_k &\xrightarrow[k\to\infty]{} \int v.
\end{aligned}$$

Sei $f = u - v \in L^1$. Wegen

$$f_k = u_k - v_k \quad \forall k \in \mathbb{N}^*$$

gilt dann

$$f_k \xrightarrow[k\to\infty]{} f \quad \text{f.ü. und} \quad \int f_k \xrightarrow[k\to\infty]{} \int f.$$

□

Eine einfache, aber sehr wichtige Anwendung von Satz IX.3.7 ist:

SATZ IX.3.8. *Sei $f \in L^1(\mathbb{R}^n, \mathbb{R})$ mit $f \geq 0$ f.ü.. Gilt $\int f = 0$, so ist $f = 0$ f.ü..*

BEWEIS. Setze $f_k = kf$, $k \in \mathbb{N}^*$. Dann ist $(f_k)_{k\in\mathbb{N}} \subset L^1$ monoton wachsend und $\int f_k = 0$ für alle $k \in \mathbb{N}^*$. Gemäß Satz IX.3.7 konvergiert $f_k$ fast überall. $(f_k(x))_{k\in\mathbb{N}}$ kann aber nur konvergieren, wenn $f(x) = 0$ ist. Damit ist die Behauptung gezeigt. □

Eine weitere wichtige Anwendung von Satz IX.3.7 ist:

SATZ IX.3.9. *Sei* $(I_k)_{k\in\mathbb{N}}$ *eine Folge von Intervallen in* $\mathbb{R}^n$ *mit* $I_k \subset I_{k+1}$ $\forall k \in \mathbb{N}$. *Sei* $I = \bigcup\limits_{k\in\mathbb{N}} I_k$. *Sei* $f \in L^1(I_k, \mathbb{R})$ *für alle* $k \in \mathbb{N}$ *und* $(\int_{I_k} |f|)_{k\in\mathbb{N}}$ *beschränkt. Dann ist* $f \in L^1(I, \mathbb{R})$ *und*

$$\int_I f = \lim_{k\to\infty} \int_{I_k} f.$$

BEWEIS. Sei zunächst $f \geq 0$ f.ü.. Definiere $f_k = f \cdot \chi_{I_k}$. Dann ist $(f_k)_{k\in\mathbb{N}} \subset L^1(\mathbb{R}^n, \mathbb{R})$ monoton wachsend, konvergiert f.ü. gegen $f$ auf $I$ und $(\int f_k)_{k\in\mathbb{N}}$ ist beschränkt. Aus Satz IX.3.7 folgt $f \in L^1(I, \mathbb{R})$ und

$$\int_I f = \lim_{k\to\infty} \int f_k = \lim_{k\to\infty} \int_{I_k} f.$$

Im allgemeinen Fall zerlegen wir $f = f^+ - f^-$ mit

$$\begin{aligned} f^+ &= \max\{f, 0\}, \\ f^- &= \max\{-f, 0\}. \end{aligned}$$

Wegen

$$\int_{I_k} f^\pm \leq \int_{I_k} |f|$$

gilt dann $f^\pm \in L^1(I, \mathbb{R})$ und

$$\lim_{k\to\infty} \int_{I_k} f_k^\pm = \int_I f^\pm$$

Hieraus folgt die Behauptung zusammen mit der Linearität des Integrals. □

Als weitere Anwendung von Satz IX.3.7 beweisen wir eine Verallgemeinerung der partiellen Integration von Satz VI.3.4 (S. 189).

SATZ IX.3.10. *Sei* $a, b \in \mathbb{R}$, $a < b$, *und* $f, g \in L^1([a, b], \mathbb{R})$. *Dann existieren für alle* $x \in [a, b]$

$$\begin{aligned} F(x) &= \int_{[a,x]} f, \\ G(x) &= \int_{[a,x]} g \end{aligned}$$

*und es gilt*

$$\int_{[a,b]} Fg + \int_{[a,b]} fG = F(b)G(b) - F(a)G(a).$$

BEWEIS. 1. SCHRITT: Wenn $f, g$ auf $[a, b]$ konstant sind, folgt die Behauptung aus Satz IX.2.17 (S. 312) und Satz VI.3.4 (S. 189).
2. SCHRITT: Wenn $f, g$ Treppenfunktionen sind, können wir $[a, b]$ in endlich viele disjunkte Intervalle, auf denen $f$ und $g$ konstant sind, zerlegen. Die Behauptung folgt dann aus dem 1. Schritt und der Linearität

des Integrals.
3. SCHRITT: Seien $f, g$ positive Funktionen in $L^{\text{inc}}$ und $(f_k)_{k\in\mathbb{N}}, (g_k)_{k\in\mathbb{N}}$ monoton wachsende Folgen von Treppenfunktionen mit

$$\begin{aligned} f_k &\underset{k\to\infty}{\longrightarrow} f \text{ f.ü.}, \\ g_k &\underset{k\to\infty}{\longrightarrow} g \text{ f.ü.}, \\ \int_{[a,b]} f_k &\underset{k\to\infty}{\longrightarrow} \int_{[a,b]} f, \\ \int_{[a,b]} g_k &\underset{k\to\infty}{\longrightarrow} \int_{[a,b]} g. \end{aligned}$$

O.E. gilt $f_k \geq 0$ f.ü., $g_k \geq 0$ f.ü. für alle $k \in \mathbb{N}$, sonst gehen wir zu $f_k^+$, $g_k^+$ über. Gemäß dem 2. Schritt existieren für alle $x \in [a, b]$ und alle $k \in \mathbb{N}$

$$\begin{aligned} F_k(x) &= \int_{[a,x]} f_k, \\ G_k(x) &= \int_{[a,x]} g_k. \end{aligned}$$

Die Folgen $(F_k)_{k\in\mathbb{N}}, (G_k)_{k\in\mathbb{N}}$ sind monoton wachsend und erfüllen

$$\begin{aligned} \int_{[a,b]} F_k g_k + \int_{[a,b]} f_k G_k &= F_k(b)G_k(b) - F_k(a)G_k(a) \\ &= \Big(\int_{[a,b]} f_k\Big)\Big(\int_{[a,b]} g_k\Big) \\ &< \infty \quad \forall k \in \mathbb{N}. \end{aligned}$$

Damit folgt die Behauptung aus Schritt 2 und Satz IX.3.7.
4. SCHRITT: Sind $f, g \in L^1$, so können wir sie gemäß Lemma IX.3.6 als Differenz positiver Funktion aus $L^{\text{inc}}$ darstellen, und die Behauptung folgt aus Schritt 3 und der Linearität des Integrals. □

Ein anderer wichtiger Grenzwertsatz für Lebesgue-Integrale ist:

SATZ IX.3.11 (SATZ VON LEBESGUE ÜBER DIE MAJORISIERTE KONVERGENZ). *Sei $(f_k)_{k\in\mathbb{N}} \subset L^1(\mathbb{R}^n, \mathbb{R})$ mit $f_k \underset{k\to\infty}{\longrightarrow} f$ f.ü.. Es gebe ein $g \in L^1(\mathbb{R}^n, \mathbb{R})$ mit*

$$|f_k| \leq g \text{ f.ü.} \quad \forall k \in \mathbb{N}.$$

*Dann ist $f \in L^1(\mathbb{R}^n, \mathbb{R})$ und*

$$\int f = \lim_{k\to\infty} \int f_k.$$

BEWEIS. Für $k \in \mathbb{N}$ sei

$$l_k(x) = \begin{cases} \inf\limits_{m\geq k} f_m(x) & \text{falls } (f_l(x))_{l\in\mathbb{N}} \text{ konvergiert,} \\ 0 & \text{sonst.} \end{cases}$$

$$u_k(x) = \begin{cases} \sup\limits_{m \ge k} f_m(x) & \text{falls } (f_l(x))_{l\in\mathbb{N}} \text{ konvergiert} \\ 0 & \text{sonst.} \end{cases}$$

Wir wollen zeigen

$$l_k, u_k \in L^1 \quad \forall k \in \mathbb{N}.$$

Seien dazu $\varphi, \psi \in L^1$. Wegen

$$\min\{\varphi, \psi\} = \frac{1}{2}(\varphi + \psi) - \frac{1}{2}|\varphi - \psi|$$
$$\max\{\varphi, \psi\} = \frac{1}{2}(\varphi + \psi) + \frac{1}{2}|\varphi - \psi|$$

folgt aus Satz IX.2.14 (S. 309)

$$\min\{\varphi, \psi\} \in L^1, \quad \max\{\varphi, \psi\} \in L^1.$$

Durch Induktion folgt hieraus, dass für alle $k \in \mathbb{N}$ und $m \in \mathbb{N}$ gilt

$$l_{km} = \min\{f_k, f_{k+1}, \ldots, f_{k+m}\} \in L^1$$
$$u_{km} = \max\{f_k, f_{k+1}, \ldots, f_{k+m}\} \in L^1.$$

Für festes $k$ erhalten wir somit zwei monotone Folgen $(l_{km})_{m\in\mathbb{N}}$ und $(u_{km})_{m\in\mathbb{N}}$ in $L^1$, die fast überall gegen $l_k$ bzw. $u_k$ konvergieren. Wegen

$$|f_m| \le g \quad \forall m \in \mathbb{N} \text{ f.ü.}$$

gilt

$$|\int l_{km}| \le \int g \quad , \quad |\int u_{km}| \le \int g \quad \forall k, m \in \mathbb{N}.$$

Damit folgt aus Satz IX.3.7

$$l_k, u_k \in L^1 \quad \forall k \in \mathbb{N}.$$

Die Folgen $(l_k)_{k\in\mathbb{N}}$, $(u_k)_{k\in\mathbb{N}}$ sind monoton wachsend bzw. fallend, konvergieren f.ü. gegen $f$ und erfüllen

$$l_k \le f_k \le u_k \quad \forall k \in \mathbb{N}$$

und somit

$$(*) \qquad \int l_k \le \int f_k \le \int u_k \quad \forall k \in \mathbb{N}.$$

Außerdem gilt für alle $k \in \mathbb{N}$

$$\begin{aligned} |\int l_k| &= \lim_{m\to\infty} |\int l_{km}| \\ &\le \int g \\ |\int u_k| &= \lim_{m\to\infty} |\int u_{km}| \\ &\le \int g. \end{aligned}$$

Damit folgt aus Satz IX.3.7 $f \in L^1$ und

$$\int f = \lim_{k\to\infty} \int l_k = \lim_{k\to\infty} \int u_k.$$

Wegen $(*)$ folgt hieraus

$$\int f = \lim_{k\to\infty} \int f_k.$$

□

Eine einfache, aber wichtige Konsequenz von Satz IX.3.11 ist:

Satz IX.3.12 (Satz von Lebesgue über die beschränkte Konvergenz). *Sei $I \subset \mathbb{R}^n$ ein beschränktes Intervall und $(f_k)_{k\in\mathbb{N}} \subset L^1(I,\mathbb{R})$ eine Folge mit $f_k \underset{k\to\infty}{\longrightarrow} f$ f.ü. in $I$. Es gebe ein $K > 0$ mit*

$$|f_k| \le K \text{ f.ü. in } I \quad \forall k \in \mathbb{N}.$$

*Dann ist $f \in L^1(I,\mathbb{R})$ und*

$$\int_I f = \lim_{k\to\infty} \int_I f_k.$$

Beweis. Mit $g = K\chi_I$ folgt die Behauptung aus Satz IX.3.11 angewandt auf die Folge $(f_k\chi_I)_{k\in\mathbb{N}}$. □

Ein anderes wichtiges Konvergenzkriterium ist:

Satz IX.3.13 (Lemma von Fatou). *Sei $(f_k)_{k\in\mathbb{N}} \subset L^1(\mathbb{R}^n,\mathbb{R})$ mit $f_k \ge 0$ f.ü. für alle $k \in \mathbb{N}$ und $f_k \underset{k\to\infty}{\longrightarrow} f$ f.ü.. Es gebe ein $K > 0$ mit*

$$\int f_k \le K \quad \forall k \in \mathbb{N}.$$

*Dann ist $f \in L^1(\mathbb{R}^n,\mathbb{R})$ und*

$$\int f \le K.$$

Beweis. Wie im Beweis von Satz IX.3.11 sei für $k, m \in \mathbb{N}$

$$\begin{aligned} l_{km} &= \min\{f_k, f_{k+1}, \ldots, f_{k+m}\} \\ l_k &= \inf_{m\ge k} f_m. \end{aligned}$$

Dann ist für festes $k \in \mathbb{N}$ $(l_{km})_{m\in\mathbb{N}} \subset L^1(\mathbb{R}^n,\mathbb{R})$ eine monoton fallende Folge positiver Funktionen mit

$$0 \le \int l_{km} \le K \quad \forall k, m.$$

Gemäß Satz IX.3.7 ist $l_k \in L^1(\mathbb{R}^n,\mathbb{R})$ für jedes $k \in \mathbb{N}$ und

$$0 \le \int l_k = \lim_{m\to\infty} \int l_{km} \le K \quad \forall k \in \mathbb{N}.$$

Da die monoton wachsende Folge $(l_k)_{k\in\mathbb{N}}$ f.ü. gegen $f$ konvergiert, folgt damit die Behauptung aus Satz IX.3.7. □

Als eine Anwendung von Satz IX.3.12 beweisen wir den Hauptsatz der Differential- und Integralrechnung (Satz VI.2.12 (S. 184)) unter abgeschwächten Voraussetzungen.

SATZ IX.3.14. *Sei $F \in C([a,b],\mathbb{R})$, $a,b \in \mathbb{R}$, $a < b$, differenzierbar und $f = F'$ beschränkt. Dann ist $f \in L^1([a,b],\mathbb{R})$ und*

$$\int_{[a,b]} f = F(b) - F(a).$$

BEWEIS. Sei $K = \sup\limits_{a\le x\le b} |f(x)| < \infty$. Definiere $\widetilde{F}$ auf $[a,b+1]$ durch

$$\widetilde{F}(x) = \begin{cases} F(x) & \text{für } a \le x \le b, \\ F(b) + (x-b)f(b) & \text{für } b < x \le b+1. \end{cases}$$

Dann ist $\widetilde{F} \in C([a,b+1],\mathbb{R})$ auf $[a,b+1]$ differenzierbar mit

$$|\widetilde{F}'(x)| \le K \quad \forall x \in [a,b+1].$$

Für $k \in \mathbb{N}^*$ sei

$$f_k(x) = k\{\widetilde{F}(x+\frac{1}{k}) - \widetilde{F}(x)\} \quad \forall x \in [a,b].$$

Dann ist $f_k \in C([a,b],\mathbb{R}) \subset L^1([a,b],\mathbb{R})$ für alle $k \in \mathbb{N}^*$ und

$$f(x) = \lim_{k\to\infty} f_k(x) \quad \forall x \in [a,b].$$

Wegen des Mittelwertsatzes gilt mit $\theta \in [0,1]$

$$|f_k(x)| = |\widetilde{F}'(x+\theta\frac{1}{k})| \le K \quad \forall k \in \mathbb{N}^*,\ x \in [a,b].$$

Aus Satz IX.3.12 folgt $f \in L^1([a,b],\mathbb{R})$ und

$$\begin{aligned} \int_{[a,b]} f &= \lim_{k\to\infty} \int_{[a,b]} f_k \\ &= \lim_{k\to\infty} k \int_a^b [\widetilde{F}(x+\frac{1}{k}) - \widetilde{F}(x)]dx \\ &= \lim_{k\to\infty} k\Big\{\int_{a+\frac{1}{k}}^{b+\frac{1}{k}} \widetilde{F}(y)dy - \int_a^b \widetilde{F}(x)dx\Big\} \\ &= \lim_{k\to\infty} k\Big\{\int_b^{b+\frac{1}{k}} \widetilde{F}(y)dy - \int_a^{a+\frac{1}{k}} \widetilde{F}(x)dx\Big\} \\ &= F(b) - F(a), \end{aligned}$$

wobei wir im letzten Schritt Satz VI.2.12 (S. 184) auf die stetige Funktion $\widetilde{F}$ angewandt haben. □

Als letzte Anwendung von Satz IX.3.11 beweisen wir den folgenden Satz über das Vertauschen von Differentiation und Integration.

Satz IX.3.15. *Seien $a, b \in \mathbb{R}$, $a < b$, und $f : (a, b) \times \mathbb{R}^n \to \mathbb{R}$ eine Funktion mit folgenden Eigenschaften:*

(1) *$f(t, \cdot) \in L^1(\mathbb{R}^n, \mathbb{R})$ für alle $t \in (a, b)$.*
(2) *$f(\cdot, x)$ ist differenzierbar für alle $x \in \mathbb{R}^n$.*
(3) *Es gibt ein $g \in L^1(\mathbb{R}^n, \mathbb{R})$ mit*

$$|\frac{\partial}{\partial t} f(t, x)| \le g(x) \quad \forall t \in (a, b),\ x \in \mathbb{R}^n.$$

*Dann ist $\frac{\partial}{\partial t} f(t, \cdot) \in L^1(\mathbb{R}^n, \mathbb{R})$ für alle $t \in (a, b)$ und*

$$\int_{\mathbb{R}^n} \frac{\partial}{\partial t} f(t, x) dx = \frac{d}{dt} \int_{\mathbb{R}^n} f(t, x) dx.$$

Beweis. Definiere zur Abkürzung $F : (a, b) \to \mathbb{R}$ durch

$$F(t) = \int_{\mathbb{R}^n} f(t, x) dx.$$

Sei $t^* \in (a, b)$ beliebig und $(\tau_k)_{k \in \mathbb{N}} \subset \mathbb{R}^*$ eine beliebige Nullfolge. Definiere für $k \in \mathbb{N}$ die Funktion $f_k \in L^1(\mathbb{R}^n, \mathbb{R})$ durch

$$f_k(x) = \begin{cases} \frac{1}{\tau_k}[f(t^* + \tau_k, x) - f(t^*, x)] & \text{falls } t^* + \tau_k \in (a, b), \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt

$$\lim_{k \to \infty} f_k(x) = \frac{\partial}{\partial t} f(t^*, x) \quad \forall x \in \mathbb{R}^n.$$

Weiter gibt es ein $k^* \in \mathbb{N}$ mit $t^* + \tau_k \in (a, b)$ für alle $k \ge k^*$. Aus dem Mittelwertsatz folgt mit $\theta \in [0, 1]$

$$|f_k(x)| = |\frac{\partial}{\partial t} f(t^* + \theta \tau_k, x)| \le g(x) \quad \forall x \in \mathbb{R}^n, k \ge k^*.$$

Also ist wegen Satz IX.3.11 $\frac{\partial}{\partial t} f(t^*, \cdot) \in L^1(\mathbb{R}^n, \mathbb{R})$ und

$$\begin{aligned} \int_{\mathbb{R}^n} \frac{\partial}{\partial t} f(t^*, x) dx &= \lim_{k \to \infty} \int_{\mathbb{R}^n} f_k(x) dx \\ &= \lim_{k \to \infty} \frac{1}{\tau_k}[F(t^* + \tau_k) - F(t^*)]. \end{aligned}$$

Da $\tau_k$ beliebig war, folgt, dass $F$ an der Stelle $t^*$ differenzierbar ist mit

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^n} f(t^*, x) dx &= \frac{d}{dt} F(t^*) \\ &= \int_{\mathbb{R}^n} \frac{\partial}{\partial t} f(t^*, x) dx. \end{aligned}$$

Damit ist die Behauptung gezeigt. □

## IX.4. Messbare Funktionen und Mengen

Definition IX.4.1. (1) Seien $f : \mathbb{R}^n \to \mathbb{R}$ und $g : \mathbb{R}^n \to \mathbb{R}_+$ zwei Funktionen. Dann ist die Funktion $f|_g : \mathbb{R}^n \to \mathbb{R}$ definiert durch (vgl. Abb. IX.4.1)

$$\begin{aligned} f|_g &= \max\{-g, \min\{f, g\}\} \\ &= \min\{g, \max\{f, -g\}\} \\ &= \begin{cases} -g(x) & \text{falls } f(x) < -g(x), \\ f(x) & \text{falls } -g(x) \le f(x) \le g(x), \\ g(x) & \text{falls } f(x) > g(x). \end{cases} \end{aligned}$$

(2) Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ heißt MESSBAR, wenn die Funktion $f|_g \in L^1(\mathbb{R}^n, \mathbb{R})$ ist für jedes $g \in L^1(\mathbb{R}^n, \mathbb{R}_+)$.
(3) Die Menge der messbaren Funktionen auf $\mathbb{R}^n$ mit Werten in $\mathbb{R}$ bezeichnen wir mit $M(\mathbb{R}^n, \mathbb{R})$.

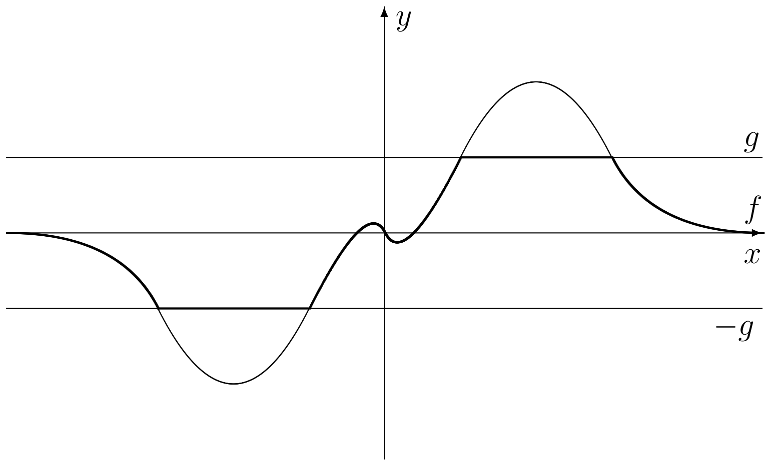


Abbildung IX.4.1. Funktion $f|_g$, $g = 1$

Satz IX.4.2. $L^1(\mathbb{R}^n, \mathbb{R}) \subset M(\mathbb{R}^n, \mathbb{R})$.

Beweis. Die Behauptung folgt aus Satz IX.2.14 (S. 309), Definition IX.4.1 und den Darstellungen

$$\begin{aligned} \max\{\varphi, \psi\} &= \frac{1}{2}(\varphi + \psi) + \frac{1}{2}|\varphi - \psi|, \\ \min\{\varphi, \psi\} &= \frac{1}{2}(\varphi + \psi) - \frac{1}{2}|\varphi - \psi| \end{aligned}$$

für beliebige Funktionen $\varphi, \psi$. □

Der folgende Satz gibt eine für die Praxis leichter handhabbare Charakterisierung messbarer Funktionen.

Satz IX.4.3. *Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine Funktion. Dann sind folgende Aussagen äquivalent:*

(1) $f \in M(\mathbb{R}^n, \mathbb{R})$.
(2) $f|_g \in L^1(\mathbb{R}^n, \mathbb{R}) \quad \forall g \in T(\mathbb{R}^n, \mathbb{R}_+)$.
(3) $f|_{K\chi_I} \in L^1(\mathbb{R}^n, \mathbb{R})$ *für alle $K \in \mathbb{R}_+$ und alle beschränkten Intervalle $I$.*

Beweis. (1) $\Longrightarrow$ (2) : Folgt direkt aus Definition IX.4.1.
(2) $\Longrightarrow$ (3) : Ist offensichtlich.
(3) $\Longrightarrow$ (2) : Sei $g \in T(\mathbb{R}^n, \mathbb{R}_+)$. Dann gibt es ein $K \in \mathbb{R}_+$ und ein beschränktes Intervall $I$ mit $g \le K\chi_I$. Dann ist nach Voraussetzung $\widetilde{f} = f|_{K\chi_I} \in L^1(\mathbb{R}^n, \mathbb{R})$. Hieraus folgt

$$f|_g = \widetilde{f}|_g \in L^1(\mathbb{R}^n, \mathbb{R}).$$

(2) $\Longrightarrow$ (1) : Sei $g \in L^1(\mathbb{R}^n, \mathbb{R}_+)$. Dann gibt es eine Folge $(g_k)_{k\in\mathbb{N}} \subset T(\mathbb{R}^n, \mathbb{R}_+)$ mit $g_k \underset{k\to\infty}{\longrightarrow} g$ f.ü.. Definiere für $k \in \mathbb{N}$

$$\widetilde{f_k} = f|_{g_k} \in L^1(\mathbb{R}^n, \mathbb{R})$$
$$f_k = \widetilde{f_k}|_g \in L^1(\mathbb{R}^n, \mathbb{R}).$$

Dann gilt $f_k \underset{k\to\infty}{\longrightarrow} f|_g$ f.ü. und

$$|f_k| \le g \in L^1(\mathbb{R}^n, \mathbb{R}_+) \quad \forall k \in \mathbb{N}.$$

Damit folgt die Behauptung aus Satz IX.3.11 (S. 324). □

Satz IX.4.4. (1) $C(\mathbb{R}^n, \mathbb{R}) \subset M(\mathbb{R}^n, \mathbb{R})$.
(2) *Seien $f, g \in M(\mathbb{R}^n, \mathbb{R})$ und $\alpha, \beta \in \mathbb{R}$. Dann gilt*

$$\alpha f + \beta g, |f|, f^+, f^-, \max\{f, g\}, \min\{f, g\} \in M(\mathbb{R}^n, \mathbb{R}).$$

(3) *Sei $(f_k)_{k\in\mathbb{N}} \subset M(\mathbb{R}^n, \mathbb{R})$ mit $f_k \underset{k\to\infty}{\longrightarrow} f$ f.ü.. Dann ist $f \in M(\mathbb{R}^n, \mathbb{R})$.*

Beweis. ad (1): Ist $f \in C(\mathbb{R}^n, \mathbb{R})$ und $I$ ein beschränktes Intervall, so ist gemäß Satz IX.2.16 (S. 311) $f \cdot \chi_I \in L^1(\mathbb{R}^n, \mathbb{R})$ und damit $f|_{K\chi_I} = (f \cdot \chi_I)|_{K\chi_I} \in L^1(\mathbb{R}^n, \mathbb{R})$ für jedes $K \in \mathbb{R}_+$.
ad (2): Sei $\varphi \in T(\mathbb{R}^n, \mathbb{R}_+)$. Dann ist

$$(|f|)|_\varphi = |(f|_\varphi)| \in L^1(\mathbb{R}^n, \mathbb{R}_+).$$

Also ist $|f| \in M(\mathbb{R}^n, \mathbb{R})$. Für $k \in \mathbb{N}$ sei

$$f_k = f|_{(k\varphi)} \quad , \quad g_k = g|_{(k\varphi)}.$$

O.E. ist $\varphi > 0$ f.ü.. Dann gilt $f_k \underset{k\to\infty}{\longrightarrow} f$ f.ü. und $g_k \underset{k\to\infty}{\longrightarrow} g$ f.ü. und somit

$$(\alpha f_k + \beta g_k)|_\varphi \underset{k\to\infty}{\longrightarrow} (\alpha f + \beta g)|_\varphi \text{ f.ü..}$$

Wegen

$$|(\alpha f_k + \beta g_k)|_\varphi| \le \varphi \quad \forall k \in \mathbb{N}$$

folgt hieraus und aus Satz IX.3.11 (S. 324) $\alpha f + \beta g \in M(\mathbb{R}^n, \mathbb{R})$. Die anderen Behauptungen folgen aus dem soeben Bewiesenen und den Beziehungen

$$\begin{aligned} f^+ &= \frac{1}{2}(|f| + f) \\ f^- &= \frac{1}{2}(|f| - f) \\ \max\{f, g\} &= \frac{1}{2}(f + g) + \frac{1}{2}|f - g| \\ \min\{f, g\} &= \frac{1}{2}(f + g) - \frac{1}{2}|f - g|. \end{aligned}$$

AD (3): Sei $g \in L^1(\mathbb{R}^n, \mathbb{R}_+)$ und $\widetilde{f}_k = f_k|_g \in L^1(\mathbb{R}^n, \mathbb{R})$ für alle $k \in \mathbb{N}$. Dann gilt

$$\widetilde{f}_k \underset{k\to\infty}{\longrightarrow} f|_g \text{ f.ü.}$$

und

$$|\widetilde{f}_k| \le g \quad \forall k \in \mathbb{N}.$$

Damit folgt die Behauptung aus Satz IX.3.11 (S. 324). □

Der folgende Satz stellt den Zusammenhang zwischen Messbarkeit und Integrierbarkeit einer Funktion her.

SATZ IX.4.5. (1) *Sei $f \in M(\mathbb{R}^n, \mathbb{R})$ und $g \in L^1(\mathbb{R}^n, \mathbb{R}_+)$ mit $|f| \le g$ f.ü.. Dann ist $f \in L^1(\mathbb{R}^n, \mathbb{R})$.*
(2) *Sei $f \in M(\mathbb{R}^n, \mathbb{R})$ und $|f| \in L^1(\mathbb{R}^n, \mathbb{R})$. Dann ist $f \in L^1(\mathbb{R}^n, \mathbb{R})$.*

BEWEIS. AD (1): Wegen $|f| \le g$ f.ü. ist $f|_g = f$ f.ü.. Hieraus folgt die Behauptung.
AD (2): Folgt aus Teil (1) mit $g = |f|$. □

Der folgende Satz ist eine teilweise Umkehrung des Satzes von Fubini und ist für die praktische Rechnung sehr wichtig.

SATZ IX.4.6 (SATZ VON TONELLI). *Seien $k, l \in \mathbb{N}^*$ und $f \in M(\mathbb{R}^{k+l}, \mathbb{R})$. Weiter existiere mindestens eines der Integrale*

(1) $$\int_{\mathbb{R}^k} \Big( \int_{\mathbb{R}^l} |f(x_1, x_2)| dx_2 \Big) dx_1$$

*oder*

(2) $$\int_{\mathbb{R}^l} \Big( \int_{\mathbb{R}^k} |f(x_1, x_2)| dx_1 \Big) dx_2.$$

*Dann ist $f \in L^1(\mathbb{R}^{k+l}, \mathbb{R})$ und*

$$\int_{\mathbb{R}^{k+l}} f = \int_{\mathbb{R}^k} \Big( \int_{\mathbb{R}^l} f(x_1, x_2) dx_2 \Big) dx_1$$

$$= \int_{\mathbb{R}^l} \Big( \int_{\mathbb{R}^k} f(x_1, x_2) dx_1 \Big) dx_2.$$

BEWEIS. Seien $I_n = \overline{B_{\|\cdot\|_\infty}(0, n)}$ und

$$g_n = (|f|)|_{n\chi_{I_n}}, \qquad n \in \mathbb{N}^*.$$

Dann ist $g_n$ eine monoton wachsende Folge in $L^1(\mathbb{R}^{k+l}, \mathbb{R})$, die fast überall gegen $|f|$ konvergiert. Wegen Satz IX.3.3 (S. 317) ist

$$\begin{aligned}\int_{\mathbb{R}^{k+l}} g_n &= \int_{\mathbb{R}^k} \Big( \int_{\mathbb{R}^l} g_n(x_1, x_2) dx_2 \Big) dx_1 \\ &= \int_{\mathbb{R}^l} \Big( \int_{\mathbb{R}^k} g_n(x_1, x_2) dx_1 \Big) dx_2 \quad \forall n \in \mathbb{N}^*.\end{aligned}$$

Daher ist $(\int_{\mathbb{R}^{k+l}} g_n)_{n \in \mathbb{N}}$ durch eines der Integrale (1) oder (2) beschränkt. Wegen Satz IX.3.7 (S. 321) ist daher $|f| \in L^1(\mathbb{R}^{k+l}, \mathbb{R})$. Damit folgt die Behauptung aus Satz IX.4.5 (2) und Satz IX.3.3. □

Wir kommen nun zum Begriff der messbaren Menge.

DEFINITION IX.4.7. Eine Teilmenge $X \subset \mathbb{R}^n$ heißt MESSBAR, wenn ihre charakteristische Funktion $\chi_X$ messbar ist. Ist $\chi_X$ sogar integrierbar, so definieren wir das $n$-DIMENSIONALE LEBESGUE-MASS von $X$ durch

$$\lambda_n(X) = \int \chi_X.$$

Ist $\chi_X$ messbar, aber nicht integrierbar, so definieren wir

$$\lambda_n(X) = +\infty.$$

Die Menge der messbaren Mengen bezeichnen wir mit $\mathcal{M}_n$.

SATZ IX.4.8. *Es gelten folgende Eigenschaften messbarer Mengen:*

(1) *Sind $X, Y \in \mathcal{M}_n$, so gilt*

$$X \cup Y, \quad X \cap Y, \quad X \backslash Y, \quad X \Delta Y \in \mathcal{M}_n.$$

*Weiter gilt*

(a) $\lambda_n(X) = 0 \iff X$ *ist Nullmenge.*

(b) $X \subset Y \Longrightarrow \lambda_n(X) \le \lambda_n(Y)$. (MONOTONIE)

(c) $\lambda_n(X \cup Y) + \lambda_n(X \cap Y) = \lambda_n(X) + \lambda_n(Y)$.

(2) *Ist $(X_k)_{k \in \mathbb{N}} \subset \mathcal{M}_n$ mit $X_k \subset X_{k+1}$ für alle $k \in \mathbb{N}$ und $X = \bigcup_{k \in \mathbb{N}} X_k$, so ist $X \in \mathcal{M}_n$ und*

$$\lambda_n(X) = \lim_{k \to \infty} \lambda_n(X_k).$$

(3) *Ist $(X_k)_{k \in \mathbb{N}} \subset \mathcal{M}_n$, so ist $X = \bigcup_{k \in \mathbb{N}} X_k \in \mathcal{M}_n$ und*

$$\lambda_n(X) \le \sum_{k \in \mathbb{N}} \lambda_n(X_k).$$

*Sind zusätzlich die Mengen $X_k$ paarweise disjunkt, so ist*

$$\lambda_n(X) = \sum_{k\in\mathbb{N}} \lambda_n(X_k).$$

Beweis. ad (1): Der erste Teil der Aussage folgt aus Satz IX.4.4 und den Beziehungen

$$\begin{aligned}
\chi_{X\cup Y} &= \max\{\chi_X, \chi_Y\},\\
\chi_{X\cap Y} &= \min\{\chi_X, \chi_Y\},\\
\chi_{X\setminus Y} &= (\chi_X - \chi_Y)^+,\\
\chi_{X\Delta Y} &= |\chi_X - \chi_Y|.
\end{aligned}$$

Aussage (a) folgt aus der Definition von Maß und Integral („$\Longleftarrow$“) bzw. aus Satz IX.3.8 (S. 322) („$\Longrightarrow$“).
Aussage (b) folgt aus der Beziehung

$$\chi_X \le \chi_Y \iff X \subset Y$$

und der Monotonie des Integrals.
Aussage (c) schließlich folgt aus der Beziehung

$$\chi_{X\cup Y} + \chi_{X\cap Y} = \chi_X + \chi_Y$$

und der Linearität des Integrals.
ad (2): Folgt aus Satz IX.4.4 angewandt auf die monoton wachsende Folge $(\chi_{X_k})_{k\in\mathbb{N}}$.
ad (3): Folgt aus Teil (2) angewandt auf die Mengen

$$Y_k = \bigcup_{l=0}^{k} X_l \quad \forall k \in \mathbb{N}$$

und aus Aussage (c) von Teil (1). □

Bemerkung IX.4.9. Seien $X \neq \emptyset$ eine beliebige Menge. $\mathcal{A} \subset \mathcal{P}(X)$ und $\mu : \mathcal{A} \to \mathbb{R}_+ \cup \{+\infty\}$ eine Funktion. $\mathcal{A}$ heißt $\sigma$-Algebra, wenn gilt:

S1: $\emptyset, X \in \mathcal{A}$,
S2: $A \in \mathcal{A} \iff X\setminus A \in \mathcal{A}$,
S3: $(A_k)_{k\in\mathbb{N}} \subset \mathcal{A} \Longrightarrow \bigcup_{k\in\mathbb{N}} A_k \in \mathcal{A}$.

$\mu$ heißt ein Mass, wenn $\mathcal{A}$ eine $\sigma$-Algebra ist und wenn gilt

M1: $\mu(\emptyset) = 0$,
M2: $(A_k)_{k\in\mathbb{N}} \subset \mathcal{A}$ paarweise disjunkt $\Rightarrow \mu(\bigcup_{k\in\mathbb{N}} A_k) = \sum_{k\in\mathbb{N}} \mu(A_k)$.

Ein Tripel $(X, \mathcal{A}, \mu)$ mit diesen Eigenschaften heißt Massraum. Satz IX.4.8 zeigt, dass $(\mathbb{R}^n, \mathcal{M}_n, \lambda_n)$ ein Maßraum ist.

Der folgende Satz zeigt, dass eine große Klasse von Mengen messbar sind.

Satz IX.4.10. (1) *Sei* $O \subset \mathbb{R}^n$, $O \neq \emptyset$ *offen. Dann ist* $O$ *in* $\mathcal{M}_n$.
(2) *Sei* $A \subset \mathbb{R}^n$, $A \neq \emptyset$, *abgeschlossen. Dann ist* $A \in \mathcal{M}_n$.
(3) *Sei* $K \subset \mathbb{R}^n$, $K \neq \emptyset$, *kompakt. Dann ist* $K \in \mathcal{M}_n$ *und* $\lambda_n(K) < +\infty$.

Beweis. ad (1): Sei $\{q_k\}_{k\in\mathbb{N}}$ eine Abzählung von $O \cap \mathbb{Q}^n$. Für $k, l \in \mathbb{N}$ sei

$$I_{kl} = B_{\|\cdot\|_\infty}(q_k, 2^{-l}) \in \mathcal{M}_n.$$

Wegen Satz IX.4.8 (3) ist dann für jedes $l \in \mathbb{N}$

$$I_l = \bigcup\{I_{kl} : I_{kl} \subset O\} \in \mathcal{M}_n.$$

Offensichtlich ist $\bigcup\limits_{l\in\mathbb{N}} I_l \subset O$. Sei nun $x \in O$. Dann gibt es ein $\varepsilon > 0$ mit

$$B_{\|\cdot\|_\infty}(x, \varepsilon) \subset O.$$

Sei $l \in \mathbb{N}$ mit $2^{-l+1} < \varepsilon$. Da $\mathbb{Q}$ in $\mathbb{R}$ und damit $\mathbb{Q}^n$ in $\mathbb{R}^n$ dicht ist, gibt es ein $k \in \mathbb{N}$ mit

$$\|x - q_k\|_\infty < 2^{-l},$$

d.h. $x \in I_{kl}$. Für $y \in B_{\|\cdot\|_\infty}(q_k, 2^{-l})$ gilt weiter

$$\|y - x\|_\infty \leq \|y - q_k\|_\infty + \|q_k - x\|_\infty \leq 2 \cdot 2^{-l} < \varepsilon.$$

Also ist $I_{kl} \subset O$ und damit $x \in I_l$. Dies zeigt $O = \bigcup\limits_{l\in\mathbb{N}} I_l$. Damit folgt die Behauptung aus Satz IX.4.8 (3).
ad (2): Gemäß Teil (1) sind $A^c$ und $\mathbb{R}^n$ messbar. Wegen $A = (A^c)^c = \mathbb{R}^n \backslash A^c$ folgt damit die Behauptung aus Satz IX.4.8.
ad (3): Gemäß Teil (2) ist $K \in \mathcal{M}_n$ Gemäß Satz III.4.3 (S. 82) gibt es ein $R > 0$ mit $\overline{B_{\|\cdot\|_\infty}(0, R)} \supset K$. Aus Satz IX.4.8 folgt

$$\lambda_n(K) \leq \lambda_n(\overline{B_{\|\cdot\|_\infty}(0, R)}) = (2R)^n.$$

□

Beispiel IX.4.11. Wir wollen eine nicht messbare Teilmenge von $\mathbb{R}$ konstruieren. Dazu definieren wir

$$x \sim y \iff x - y \in \mathbb{Q}.$$

Wie man leicht nachprüft ist $\sim$ eine Äquivalenzrelation. Bezeichne mit $[x]$ die zu $x$ gehörige Äquivalenzklasse. Für jedes $x \in (0, 1)$ wähle einen Repräsentanten aus $[x] \cap (0, 1)$ aus. Die Menge aller so ausgewählten Elemente von $(0, 1)$ bezeichne mit $M$.
Sei $x \in (0, 1)$. Dann gibt es ein $y \in M$ mit $x \sim y$. Also existiert ein $q \in \mathbb{Q}$ mit $x - y = q$ und $|q| < 1$. Daher ist

$$(*) \qquad (0, 1) \subset \bigcup_{q\in\mathbb{Q}\cap(-1,1)} (M + q) \subset (-1, 2)$$

mit

$$M + q = \{x + q : x \in M\}.$$

Seien nun $q, r \in \mathbb{Q}$ und $x \in (M+q) \cap (M+r)$. Dann folgt

$$x = y + q = z + r$$

mit $y, z \in M$. Also ist $y - z = r - q \in \mathbb{Q}$ und damit $y = z$ und daher $r = q$. Also sind die Mengen $M + q$ für verschiedene $q$ disjunkt.
Wir nehmen an, $M$ wäre messbar. Dann ist $\lambda_1(M+q) = \lambda_1(M)$ für alle $q \in \mathbb{Q}$. Wäre $\lambda_1(M) \neq 0$, so folgte aus $(*)$ $\lambda_1((-1,2)) = +\infty$. Wäre $\lambda_1(M) = 0$, so folgte dagegen aus $(*)$ $\lambda_1((0,1)) = 0$. Wir erhalten also in jedem Fall einen Widerspruch.
$\chi_M$ ist ein Beispiel für eine nicht-messbare Funktion.

Zum Abschluss gehen wir noch auf den Zusammenhang zwischen messbaren Mengen und Funktionen näher ein.

Satz IX.4.12. *Es ist $f \in M(\mathbb{R}^n, \mathbb{R})$ genau dann, wenn für jedes $c \in \mathbb{R}$ die Menge*

$$A_c = \{x \in \mathbb{R}^n : f(x) \geq c\}$$

*messbar ist.*

Beweis. „$\Longrightarrow$“: Definiere für $c \in \mathbb{R}$ und $k \in \mathbb{N}^*$

$$f_k(x) = k[\min\{f(x), c\} - \min\{f(x), c - \frac{1}{k}\}] \quad \forall x \in \mathbb{R}^n.$$

Dann ist $f_k \in M(\mathbb{R}^n, \mathbb{R})$ für alle $k \in \mathbb{N}^*$, $c \in \mathbb{R}$. Sei $x \in \mathbb{R}^n$ und $f(x) \geq c$. Dann ist $f_k(x) = 1$ für alle $k \in \mathbb{N}^*$. Ist dagegen $f(x) < c$, so gibt es ein $k_* \in \mathbb{N}$ mit $f(x) \leq c - \frac{1}{k_*}$. Für $k \geq k_*$ folgt $f_k(x) = 0$. Also gilt $f_k \underset{k\to\infty}{\longrightarrow} \chi_{A_c}$. Damit folgt die Behauptung aus Satz IX.4.4 (3).
„$\Longleftarrow$“: Aus Satz IX.4.8 folgt, dass für alle $c, d \in \mathbb{R}$ mit $c < d$ die Menge

$$Z_{c,d} = \{x \in \mathbb{R}^n : c \leq f(x) < d\}$$

messbar ist. Dann ist für $k \in \mathbb{N}^*$

$$f_k = \sum_{l=0}^{2k^2-1} (-k + \frac{l}{k}) \chi_{Z_{-k+\frac{l}{k}, -k+\frac{l+1}{k}}}$$

aus $M(\mathbb{R}^n, \mathbb{R})$. Wie man leicht nachprüft, gilt $f_k \xrightarrow[k\to\infty]{\text{pkt}} f$. Damit folgt die Behauptung aus Satz IX.4.4 (3). $\square$

Das folgende Resultat ist für die Praxis besonders wichtig.

Satz IX.4.13. *Seien $f, g \in M(\mathbb{R}^n, \mathbb{R})$ und $h \in C(\mathbb{R}^2, \mathbb{R})$. Dann ist $F : \mathbb{R}^n \to \mathbb{R}$ mit $x \mapsto h(f(x), g(x))$ messbar. Insbesondere ist $f \cdot g \in M(\mathbb{R}^n, \mathbb{R})$.*

Beweis. Wie im Beweis von Satz IX.4.12 sei für $k \in \mathbb{N}^*$

$$f_k = \sum_{l=0}^{2k^2-1} (-k + \frac{l}{k}) \chi_{\{x \in \mathbb{R}^n : -k + \frac{l}{k} \leq f(x) < -k + \frac{l+1}{k}\}}$$

$$g_k = \sum_{l=0}^{2k^2-1} (-k + \frac{l}{k})\chi_{\{x\in\mathbb{R}^n: -k+\frac{l}{k} \le g(x) < -k+\frac{l+1}{k}\}}.$$

Definiere $F_k : \mathbb{R}^n \to \mathbb{R}$ durch

$$F_k(x) = h(f_k(x), g_k(x)).$$

Dann ist

$$F_k = \sum_{i,j=0}^{2k^2-1} h(-k+\frac{i}{k}, -k+\frac{j}{k})\chi_{\{x\in\mathbb{R}^n: -k+\frac{i}{k} \le f(x) < -k+\frac{i+1}{k}, -k+\frac{j}{k} \le g(x) < -k+\frac{j+1}{k}\}}$$

wegen Satz IX.4.12 und Satz IX.4.4 messbar für jedes $k \in \mathbb{N}^*$. Wegen $f_k \underset{k\to\infty}{\longrightarrow} f$, $g_k \underset{k\to\infty}{\longrightarrow} g$ und $h \in C(\mathbb{R}^2, \mathbb{R})$ folgt $F_k \underset{k\to\infty}{\longrightarrow} F$. Damit folgt die Behauptung aus Satz IX.4.4 (3). □

## IX.5. Der Transformationssatz

Wir erinnern an die Substitutionsregel Satz VI.3.1 (S. 187). Sei $I \subset \mathbb{R}$ ein perfektes Intervall, $f \in C(I, \mathbb{R})$, $[a, b] \subset \mathbb{R}$, $a < b$ und $\varphi \in C^1([a, b], \mathbb{R})$ mit $\varphi([a, b]) \subset I$. Dann ist

$$\int_a^b f(\varphi(x))\varphi'(x)dx = \int_{\varphi(a)}^{\varphi(b)} f(y)dy.$$

Setzen wir zusätzlich $\varphi'(x) \ne 0$ für alle $x \in (a, b)$ voraus, so sind zwei Fälle möglich:
Fall 1: $\varphi'(x) > 0$ für alle $x \in (a, b)$. Dann ist

$$\int_{\varphi(a)}^{\varphi(b)} f(y)dy = \int_{\varphi([a,b])} f dy.$$

Fall 2: $\varphi'(x) < 0$ für alle $x \in (a, b)$. Dann ist

$$\int_{\varphi(a)}^{\varphi(b)} f(y)dy = -\int_{\varphi([a,b])} f dy.$$

Also können wir die Substitutionsregel unter der Annahme $\varphi'(x) \ne 0$ für alle $x \in (a, b)$ in der Form schreiben

$$\int_{[a,b]} f \circ \varphi |\varphi'| dx = \int_{\varphi([a,b])} f dy.$$

Ziel dieses Abschnittes ist es, ein Analogon dieser Formel für $C^1$-Diffeomorphismen auf $\mathbb{R}^n$ zu beweisen. Dabei wird $|\varphi'|$ durch $|\det D\varphi|$ ersetzt werden.

Im Folgenden versehen wir den $\mathbb{R}^n$ stets mit der kanonischen Basis und identifizieren so lineare Abbildungen mit Matrizen. Wenn wir die Norm $\|\cdot\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)}$ benützen, setzen wir stets voraus, dass $\mathbb{R}^n$ mit der Maximum-Norm versehen ist.

Zur Erinnerung: In diesem Fall gilt

$$\|A\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)} = \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |A_{ij}|.$$

Wir bereiten den Transformationssatz mit einer Reihe von Sätzen vor.

Satz IX.5.1. *Seien $U \subset \mathbb{R}^n$ offen, $U \neq \emptyset$, $\varphi \in C(U, \mathbb{R}^n)$ und $I \subset U$ ein beschränktes Intervall. Dann ist $\varphi(I)$ messbar und $\lambda_n(\varphi(I)) < \infty$.*

Beweis. Sei $I = \prec a, b \succ$ mit $a, b \in \mathbb{R}^n$. Für $1 \le l \le n$ und $k \in \mathbb{N}^*$ sei

$$I_{kl} = \begin{cases} [a_l, b_l] & \text{falls } \prec a_l, b_l \succ = [a_l, b_l] \\ [a_l + \frac{1}{k}, b_l] & \text{falls } \prec a_l, b_l \succ = (a_l, b_l] \\ [a_l, b_l - \frac{1}{k}] & \text{falls } \prec a_l, b_l \succ = [a_h, b_l) \\ [a_l + \frac{1}{k}, b_l - \frac{1}{k}] & \text{falls } \prec a_l, b_l \succ = (a_l, b_l) \end{cases}$$

und $I_k = I_{k1} \times \ldots \times I_{kn}$. Dann ist jedes $I_k$ kompakt und $I_k \subset I_{k+1}$ für alle $k \in \mathbb{N}$ und $I = \bigcup_{k\in\mathbb{N}^*} I_k$. Daher ist $\varphi(I_k)$ für jedes $k \in \mathbb{N}$ kompakt und damit messbar und $\varphi(I_k) \subset \varphi(I_{k+1})$ für alle $k \in \mathbb{N}^*$ und $\varphi(I) = \bigcup_{k\in\mathbb{N}^*} \varphi(I_k)$. Wegen Satz IX.4.8 (S. 332) ist somit $\varphi(I)$ messbar. Da $\varphi(I) \subset \varphi(\overline{I})$ und $\varphi(\overline{I})$ kompakt ist, ist $\lambda_n(\varphi(I)) < \infty$. □

Im folgenden Satz benützen wir zwei spezielle Arten von Matrizen:

Typ $P$: Für $1 \le i, j \le n$, $i \neq j$, geht $P_{(ij)}$ aus der Einheitsmatrix durch Vertauschen der Zeilen $i$ und $j$ hervor. Die entsprechende lineare Abbildung ist eine Spiegelung an der Hyperebene $x_i = x_j$.

Typ $S$: Für $1 \le i, j \le n$, $i \neq j$ und $t \in \mathbb{R}$, geht $S_{(ij)}(t)$ aus der Einheitsmatrix hervor, indem das $t$-fache der $i$-ten Zeile zur $j$-ten Zeile addiert wird. Die zugehörige lineare Abbildung ist eine Scherung in der $(i, j)$-Ebene parallel zur $j$-Achse.

Die Matrizen der Typen $P$ und $S$ sind invertierbar. Wie man leicht nachrechnet, gilt

$$(P_{(ij)})^{-1} = P_{(ij)},$$
$$(S_{(ij)}(t))^{-1} = S_{(ij)}(-t).$$

Die Multiplikation einer Matrix von links bzw. rechts mit einer Matrix $P_{(ij)}$ oder $S_{(ij)}(t)$ entspricht dem Vertauschen der $i$-ten und $j$-ten Zeile bzw. Spalte oder der Addition des $t$-fachens der $i$-ten Zeile bzw. Spalte zur $j$-ten Zeile bzw. Spalte.

Satz IX.5.2. *Sei $L \in \mathcal{GL}(\mathbb{R}^n)$. Dann gibt es eine Diagonalmatrix $D$ mit nicht verschwindenden Diagonalelementen und Matrizen $A$, $B$,* *die Produkte von Matrizen der Typen $P$ und $S$ sind, mit*

$$L = ADB.$$

Beweis. Wir führen folgenden Algorithmus aus, der eine Modifikation des Gaußschen Eliminationsverfahrens ist:

(0) Setze $i = 1$, $L^{(0)} = L$.

(1) Bestimme $j_i$ mit $i \le j_i \le n$ und

$$|L^{(i-1)}_{j_i i}| = \max_{i \le k \le n} |L^{(i-1)}_{ki}|.$$

(2) Vertausche die Zeilen $i$ und $j_i$:

$$P^{(i)} = P_{(i,j_i)}, \; \widetilde{L}^{(i-1)} = P^{(i)} L^{(i-1)}.$$

(3) Eliminiere die Elemente unterhalb der Diagonalen in der $i$-ten Spalte und rechts von der Diagonalen in der $i$-ten Zeile:

$$\begin{aligned}
A^{(i)} &= S_{(ni)}\Big(-\frac{\widetilde{L}^{(i-1)}_{ni}}{\widetilde{L}^{(i-1)}_{ii}}\Big) \cdot \ldots \cdot S_{(i+1i)}\Big(-\frac{\widetilde{L}^{(i-1)}_{i+1i}}{\widetilde{L}^{(i-1)}_{ii}}\Big)\\
B^{(i)} &= S_{(ii+1)}\Big(-\frac{\widetilde{L}^{(i-1)}_{ii+1}}{\widetilde{L}^{(i-1)}_{ii}}\Big) \cdot \ldots \cdot S_{(in)}\Big(-\frac{\widetilde{L}^{(i-1)}_{in}}{\widetilde{L}^{(i-1)}_{ii}}\Big)\\
L^{(i)} &= A^{(i)} \widetilde{L}^{(i-1)} B^{(i)}.
\end{aligned}$$

(4) Falls $i = n-1$ ist, stop. Sonst erhöhe $i$ um 1 und gehe nach 1 zurück.

Dann ist

$$D = L^{(n-1)}$$

die gesuchte Diagonalmatrix und

$$\begin{aligned}
A^{-1} &= A^{(n-1)} P^{(n-1)} A^{(n-2)} P^{(n-2)} \ldots A^{(1)} P^{(1)}\\
B^{-1} &= B^{(1)} B^{(2)} \ldots B^{(n-1)}.
\end{aligned}$$

Man beachte, dass wegen $L \in \mathcal{GL}(\mathbb{R}^n)$ für $1 \le i \le n-1$ gilt

$$\max_{i \le k \le n} |L^{(i-1)}_{ki}| > 0.$$

□

Satz IX.5.3. *Seien $a \in \mathbb{R}^n$, $L \in \mathcal{GL}(\mathbb{R}^n)$, $\varphi(x) = a + Lx$ für alle $x \in \mathbb{R}^n$ und $I$ ein beschränktes Intervall. Dann ist*

$$\lambda_n(\varphi(I)) = |\det L| \lambda_n(I).$$

Beweis. Wegen Satz IX.5.2 ist $\varphi$ die Komposition einer Translation $x \mapsto a + x$, einer linearen Abbildung mit Diagonalmatrix und endlich vieler linearer Abbildungen vom Typ $P$ und $S$. Da $\det L$ das Produkt der Determinanten der entsprechenden Matrizen ist, reicht es, die Aussage für diese elementaren Abbildungen zu beweisen. Sei dazu $I = \prec u, v \succ$ mit $u, v \in \mathbb{R}^n$.

1. Schritt: Betrachte die Translation $\tau : x \mapsto x + a$. Dann ist $\tau(I) = \prec u + a, v + a \succ$ und

$$\begin{aligned}\lambda_n(\tau(I)) &= \prod_{i=1}^{n}(v_i - u_i) \\ &= \lambda_n(I) \\ &= |\det Id|\lambda_n(I).\end{aligned}$$

2. Schritt: Betrachte die lineare Abbildung $\delta : x \mapsto Dx$ mit einer Diagonalmatrix $D = \operatorname{diag}(d_1, \ldots, d_n)$ und $d_i \in \mathbb{R}^*$, $1 \le i \le n$. Dann ist

$$\delta(I) = \prod_{k=1}^{n} \prec \min\{d_k u_k, d_k v_k\}, \max\{d_k u_k, d_k v_k\} \succ$$

und

$$\begin{aligned}\lambda_n(\delta(I)) &= \prod_{k=1}^{n} |d_k v_k - d_k u_k| \\ &= |\prod_{k=1}^{n} d_k|\lambda_n(I) \\ &= |\det D|\lambda_n(I).\end{aligned}$$

3. Schritt: Betrachte die lineare Abbildung $\rho : x \mapsto P_{(ij)}x$ mit $1 \le i, j \le n$ und $i \ne j$. Dann ist O.E. $i < j$ und

$$\begin{aligned}\rho(I) = &\prec u_1, v_1 \succ \times \ldots \times \prec u_{i-1}, v_{i-1} \succ \times \prec u_j, v_j \succ \times \\ &\times \prec u_{i+1}, v_{i+1} \succ \times \ldots \times \prec u_{j-1}, v_{j-1} \succ \times \prec u_i, v_i \succ \times \\ &\times \prec u_{j+1}, v_{j+1} \succ \times \ldots \times \prec u_n, v_n \succ\end{aligned}$$

und daher

$$\begin{aligned}\lambda_n(\rho(I)) &= \lambda_n(I) \\ &= |\det P_{ij}|\lambda_n(I).\end{aligned}$$

4. Schritt: Betrachte die lineare Abbildung $\sigma : x \mapsto S_{(ij)}(t)x$ mit $1 \le i, j \le n$, $i \ne j$, $t \in \mathbb{R}$. Aus Satz IX.5.1 und Satz IX.3.3 (S. 317) folgt

$$\begin{aligned}\lambda_n(\sigma(I)) &= \int_{\mathbb{R}^n} \chi_{\sigma(I)} \\ &= \int_{\mathbb{R}^{n-1}} [\int_{\mathbb{R}} \chi_{\sigma(I)} dx_j] dx_1 \ldots dx_{j-1} dx_{j+1} \ldots dx_n \\ &= \int_{\mathbb{R}^{n-1}} \chi_{\prec u_1, v_1 \succ \times \ldots \times \prec u_{j-1}, v_{j-1} \succ \times \prec u_{j+1}, v_{j+1} \succ \times \ldots \times \prec u_n, v_n \succ} \\ &\qquad [\int_{\mathbb{R}} \chi_{\prec u_j + t x_i, v_j + t x_i \succ} dx_j] dx_1 \ldots dx_{j-1} dx_{j+1} \ldots dx_n\end{aligned}$$

$$\begin{aligned} &= \lambda_n(I) \\ &= |\det S_{(ij)}(t)|\lambda_n(I). \end{aligned}$$

□

Satz IX.5.4. *Seien $a \in \mathbb{R}^n$, $L \in \mathcal{GL}(\mathbb{R}^n)$, $\varphi(x) = a + Lx$ für alle $x \in \mathbb{R}^n$ und $f \in L^1(\mathbb{R}^n, \mathbb{R})$. Dann ist $f \circ \varphi^{-1} \in L^1(\mathbb{R}^n, \mathbb{R})$ und*

$$\int f \circ \varphi^{-1} = |\det L| \int f.$$

Beweis. 1. Schritt: Sei $I$ ein beschränktes Intervall und $f = \chi_I$. Dann ist $f \circ \varphi^{-1} = \chi_{\varphi(I)}$. Damit folgt die Behauptung aus Satz IX.5.3.
2. Schritt: Sei $f \in T(\mathbb{R}^n, \mathbb{R})$. Dann folgt die Behauptung aus dem 1. Schritt und der Linearität des Integrals.
3. Schritt: Sei $N$ eine Nullmenge. Wir zeigen, dass $\varphi(N)$ ebenfalls eine Nullmenge ist.
Gemäß Satz IX.2.7(2) (S. 303) gibt es eine monoton wachsende Folge $(f_k)_{k\in\mathbb{N}} \subset T(\mathbb{R}^n, \mathbb{R})$, derart dass $(f_k)_{k\in\mathbb{N}}$ auf $N$ divergiert und $(\int f_k)_{k\in\mathbb{N}}$ beschränkt ist. Wegen Schritt 2 ist $(f_k \circ \varphi^{-1})_{k\in\mathbb{N}} \subset L^1(\mathbb{R}^n, \mathbb{R})$ monoton wachsend, auf $\varphi(N)$ divergent und $(\int f_k \circ \varphi^{-1})_{k\in\mathbb{N}} = (|\det L| \int f_k)_{k\in\mathbb{N}}$ beschränkt. Aus Satz IX.3.7 (S. 321) folgt, dass $(f_k \circ \varphi^{-1})_{k\in\mathbb{N}}$ f.ü. konvergiert. Daher ist $\varphi(N)$ eine Nullmenge.
4. Schritt: Sei $f \in L^{\text{inc}}$ und $(f_k)_{k\in\mathbb{N}} \subset T(\mathbb{R}^n, \mathbb{R})$ eine monoton wachsende Folge mit $(\int f_k)_{k\in\mathbb{N}}$ beschränkt, die f.ü. gegen $f$ konvergiert. Wegen Schritt 2 und 3 konvergiert $(f_k \circ \varphi^{-1})_{k\in\mathbb{N}} \subset L^1(\mathbb{R}^n, \mathbb{R})$ f.ü. gegen $f \circ \varphi^{-1}$, ist monoton wachsend und erfüllt, $(\int f_k \circ \varphi^{-1})_{k\in\mathbb{N}}$ ist beschränkt. Aus Satz IX.3.7 (S. 321) folgt die Behauptung für $f$.
5. Schritt: Sei $f \in L^1(\mathbb{R}^n, \mathbb{R})$. Dann folgt die Behauptung aus dem 4. Schritt und der Linearität des Integrals. □

Satz IX.5.5. *Seien $U, V \subset \mathbb{R}^n$ offene, nicht leere Teilmengen, $\varphi : U \to V$ ein $C^1$-Diffeomorphismus von $U$ auf $V$ und $I \subset \mathbb{R}^n$ ein beschränktes Intervall mit $\mathring{I} \neq \emptyset$ und $\overline{I} \subset U$. Dann ist*

$$\lambda_n(\varphi(I)) = \int_I |\det D\varphi(x)| dx.$$

Beweis. Sei $I = \prec u, v \succ$ mit $u, v \in \mathbb{R}^n$. Definiere

$$\begin{aligned} K_0 &= \max\{1, \lambda_n(I)\}, \\ L_{\max} &= \max_{1\le k\le n} |v_k - u_k| \\ L_{\min} &= \min_{1\le k\le n} |v_k - u_k| > 0. \end{aligned}$$

Da $\overline{I}$ kompakt ist und $(D\varphi)^{-1}$ und $\det(D\varphi)$ auf $\overline{I}$ stetig sind, existieren

$$\begin{aligned} K_1 &= \max\{1, \max_{x\in\overline{I}} \|(D\varphi(x))^{-1}\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)}\} \\ K_2 &= \max\{1, \max_{x\in\overline{I}} |\det(D\varphi(x))|\}. \end{aligned}$$

Sei $0 < \varepsilon < 1$ beliebig. Da $D\varphi$ und $\det D\varphi$ auf $\overline{I}$ gleichmäßig stetig sind, gibt es ein $\delta > 0$ mit

$$\begin{aligned} |\det(D\varphi(x)) - \det(D\varphi(y))| &< \frac{\varepsilon}{2K_0} \\ \|D\varphi(x) - D\varphi(y)\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)} &< \varepsilon \frac{L_{\min}}{L_{\max} K_0 K_1 K_2 n 2^n} \end{aligned}$$

für alle $x, y \in \overline{I}$ mit $\|x - y\|_\infty < \delta$. Sei $k \in \mathbb{N}$ so, dass

$$2^{-k} L_{\max} < \delta$$

ist. Durch sukzessives Halbieren der Kanten können wir $I$ in $2^{nk}$ disjunkte Intervalle $I_l$, $1 \le l \le 2^{nk}$, mit maximaler Kantenlänge $2^{-k}L_{\max}$ und minimaler Kantenlänge $2^{-k}L_{\min}$ unterteilen. Dann sind die $\varphi(I_l)$ paarweise disjunkt und $\varphi(I) = \bigcup_{l=1}^{2^{nk}} \varphi(I_l)$. Damit folgt

$$\begin{aligned} &\Big|\lambda_n(\varphi(I)) - \int_I |\det D\varphi(x)| dx\Big| \\ &= \Big|\sum_{l=1}^{2^{nk}} \{\lambda_n(\varphi(I_l)) - \int_{I_l} |\det D\varphi(x)| dx\}\Big| \\ &\le \sum_{l=1}^{2^{nk}} \Big|\lambda_n(\varphi(I_l)) - \int_{I_l} |\det D\varphi(x)| dx\Big|. \end{aligned} \tag{IX.5.1}$$

Sei $1 \le l \le 2^{nk}$ beliebig. Dann ist $I_l = \prec \tilde{u}, \tilde{v} \succ$. Definiere

$$\begin{aligned} z &= \frac{1}{2}(\tilde{u} + \tilde{v}) \\ \psi(x) &= \varphi(z) + D\varphi(z)(x - z) \quad \forall x \in \mathbb{R}^n \\ \rho &= \psi^{-1} \circ \varphi. \end{aligned}$$

Mit Satz IX.5.3 angewandt auf $\psi^{-1}$ und $\varphi(I_l)$ folgt

$$\begin{aligned} &\Big|\lambda_n(\varphi(I_l)) - \int_{I_l} |\det D\varphi(x)| dx\Big| \\ &\le \Big|\int_{I_l} \{|\det D\varphi(x)| - |\det D\varphi(z)|\} dx\Big| \\ &\quad + \Big|\lambda_n(I_l) |\det D\varphi(z)| - \lambda_n(\varphi(I_l))\Big| \\ &\le \int_{I_l} |\det D\varphi(x) - \det D\varphi(z)| dx \\ &\quad + K_2 \Big|\lambda_n(I_l) - \underbrace{|\det D\varphi(z)|^{-1} \lambda_n(\varphi(I_l))}_{=\lambda_n(\rho(I_l))}\Big| \\ &\le \frac{\varepsilon}{2K_0} \lambda_n(I_l) + K_2 |\lambda_n(I_l) - \lambda_n(\rho(I_l))|. \end{aligned} \tag{IX.5.2}$$

Sei $x \in I_l$. Dann gilt wegen $\rho(z) = z$

(IX.5.3)
$$
\begin{aligned}
&\|\rho(x) - x\|_\infty \\
&= \|\rho(x) - \rho(z) - (x - z)\|_\infty \\
&= \| \int_0^1 [D\rho(z + t(x - z)) - I](x - z) dt \|_\infty \\
&\leq \int_0^1 \|(D\varphi(z))^{-1}\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)} \|D\varphi(z + t(x - z)) - D\varphi(z)\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)} \\
&\qquad \|x - z\|_\infty dt \\
&\leq K_1 \varepsilon \frac{L_{\min}}{L_{\max} K_0 K_1 K_2 n 2^n} L_{\max} 2^{-k-1} \\
&= \varepsilon \frac{L_{\min}}{K_0 K_2 n 2^{n+1}} 2^{-k} \\
&=: \eta.
\end{aligned}
$$

Da $\rho$ ein Homöomorphismus ist, ist $\rho(I_l)$ zusammenhängend und

$$
\rho(\overset{\circ}{I_l}) = \overset{\circ}{\overbrace{\rho(I_l)}} \quad , \quad \rho(\overline{I}_l) = \overline{\rho(I_l)} \quad , \quad \rho(\partial I_l) = \partial(\rho(I_l)).
$$

Hieraus folgt mit Gleichung (IX.5.3)

$$
\prec u + \eta e, v - \eta e \succ \subset \rho(I_l) \subset \prec u - \eta e, v + \eta e \succ,
$$

wobei $e = (1, \ldots, 1) \in \mathbb{R}^n$ ist. Damit folgt

$$
\begin{aligned}
\lambda_n(\rho(I_l)) - \lambda_n(I_l) &\leq \prod_{i=1}^n (v_i - u_i + 2\eta) - \prod_{i=1}^n (v_i - u_i) \\
&= \lambda_n(I_l) \Big\{ \prod_{i=1}^n (1 + \frac{2\eta}{v_i - u_i}) - 1 \Big\} \\
&\leq \lambda_n(I_l) \Big\{ (1 + \frac{2\eta}{L_{\min} 2^{-k}})^n - 1 \Big\} \\
&= \lambda_n(I_l) \Big\{ (1 + \frac{\varepsilon}{K_0 K_2 n 2^n})^n - 1 \Big\} \\
&\leq \lambda_n(I_l) \frac{\varepsilon}{K_0 K_2 n 2^n} n 2^{n-1} \\
&= \lambda_n(I_l) \frac{\varepsilon}{2 K_0 K_2}
\end{aligned}
$$

und

$$
\lambda_n(I_l) - \lambda_n(\rho(I_l)) \leq \prod_{i=1}^n (v_i - u_i) - \prod_{i=1}^n (v_i - u_i - 2\eta)
$$

$$\begin{aligned}
&= \lambda_n(I_l)\Big\{1 - \prod_{i=1}^{n}(1 - \frac{2\eta}{v_i - u_i})\Big\} \\
&\le \lambda_n(I_l)\Big\{1 - (1 - \frac{2\eta}{2^{-k}L_{\min}})^n\Big\} \\
&= \lambda_n(I_l)\Big\{1 - (1 - \frac{\varepsilon}{K_0 K_2 n 2^n})^n\Big\} \\
&\le \lambda_n(I_l)\frac{\varepsilon}{K_0 K_2 n 2^n} n \\
&\le \lambda_n(I_l)\frac{\varepsilon}{2K_0 K_2}
\end{aligned}$$

also

$$|\lambda_n(\rho(I_l)) - \lambda_n(I_l)| \le \lambda_n(I_l)\frac{\varepsilon}{2K_0K_2}. \tag{IX.5.4}$$

Aus den Gleichungen (IX.5.1), (IX.5.2) und (IX.5.4) folgt schließlich

$$\begin{aligned}
\Big|\lambda_n(\varphi(I)) - \int_I |\det D\varphi(x)|dx\Big| &\le \frac{\varepsilon}{2K_0}\sum_{l=1}^{2^{nk}}\lambda_n(I_l) + \frac{\varepsilon}{2K_0}\sum_{l=1}^{2^{nk}}\lambda_n(I_l) \\
&= \frac{\varepsilon}{K_0}\lambda_n(I) \\
&\le \varepsilon.
\end{aligned}$$

Da $\varepsilon$ beliebig war, folgt hieraus die Behauptung. □

Nach diesen Vorbereitungen können wir nun den Transformationssatz beweisen. Man beachte, dass Satz IX.5.4 ein Spezialfall hiervon ist.

Satz IX.5.6 (Transformationssatz). *Seien $U, V \subset \mathbb{R}^n$ offene, nicht leere Teilmengen und $\varphi : U \to V$ ein $C^1$-Diffeomorphismus von $U$ auf $V$. Dann ist $f \in L^1(V, \mathbb{R})$ genau dann, wenn $f \circ \varphi|\det D\varphi| \in L^1(U, \mathbb{R})$ ist. In diesem Fall gilt*

$$\int_V f dy = \int_U f(\varphi(x))|\det D\varphi(x)|dx.$$

Beweis. 1. Schritt: Es genügt, die Implikation „$\Longrightarrow$" zu zeigen. Die Implikation „$\Longleftarrow$" folgt dann durch Anwenden des Gezeigten auf $\varphi^{-1} : V \to U$.

2. Schritt: Sei $I \subset V$ ein beschränktes, offenes Intervall. Dann ist $\mathcal{O} = \varphi^{-1}(I) \subset U$ offen. Sei $(z_k)_{k\in\mathbb{N}}$ eine Abzählung von $\mathbb{Z}^n$. Für $k, l \in \mathbb{N}$ sei

$$I_{kl} = \prod_{i=1}^{n}(2^{-l}z_{k,i} - 2^{-l-1}, 2^{-l}z_{k,i} + 2^{-l-1}]$$

und

$$I_l = \bigcup\{I_{kl} : \overline{I}_{kl} \subset \mathcal{O}\}.$$

Dann sind für festes $l$ die $I_{kl}$ disjunkt und

$$I_l \subset I_{l+1} \quad \forall l \in \mathbb{N}.$$

Wie im Beweis von Satz IX.4.10(1) (S. 334) folgt

$$\mathcal{O} = \bigcup_{l \in \mathbb{N}} I_l.$$

Dann ist $I = \varphi(\mathcal{O}) = \bigcup_{l \in \mathbb{N}} \varphi(I_l)$, $\varphi(I_l) = \bigcup \{\varphi(I_{kl}) : \overline{I}_{kl} \subset \mathcal{O}\}$ und die $\varphi(I_{kl})$ sind für festes $l$ disjunkt. Damit folgt aus Satz IX.5.5 und Satz IX.4.8 (S. 332) $\varphi(I_l) \in \mathcal{M}_n$ und

$$\begin{aligned}
\lambda_n(I) &\geq \lambda_n(\varphi(I_l)) \\
&= \sum_{\overline{I}_{kl} \subset \mathcal{O}} \lambda_n(\varphi(I_{kl})) \\
&= \sum_{\overline{I}_{kl} \subset \mathcal{O}} \int_{I_{kl}} |\det D\varphi(x)| dx \\
&= \int_{I_l} |\det D\varphi(x)| dx \quad \forall l \in \mathbb{N}.
\end{aligned}$$

Wiederum mit Satz IX.4.8 (S. 332) und Satz IX.3.7 (S. 321) folgt

$$\begin{aligned}
\int_V \chi_I &= \lambda_n(I) \\
&= \lim_{l \to \infty} \lambda_n(\varphi(I_l)) \\
&= \lim_{l \to \infty} \int_{I_l} |\det D\varphi(x)| dx \\
&= \int_{\mathcal{O}} |\det D\varphi(x)| dx \\
&= \int_U \chi_{\mathcal{O}} |\det D\varphi(x)| dx \\
&= \int_U \chi_I \circ \varphi(x) |\det D\varphi(x)| dx.
\end{aligned}$$

Damit ist die Behauptung für $f = \chi_I$ gezeigt.
3. Schritt: Sei $I \subset V$ ein beschränktes Intervall. Da jedes beschränkte Intervall in $\mathbb{R}$ in der Form $A \backslash (B \cup C)$ mit offenen, beschränkten Intervallen $A$, $B$, $C$ und $B \cup C \subset A$ dargestellt werden kann, ist $\chi_I$ Linearkombination von charakteristischen Funktionen beschränkter, offener Intervalle in $V$. Damit folgt die Behauptung für $f = \chi_I$ aus dem 2. Schritt und der Linearität des Integrals.
4. Schritt: Aus dem 3. Schritt und der Linearität des Integrals folgt die Behauptung für alle $f \in T(V, \mathbb{R})$.
5. Schritt: Sei $N \subset V$ eine Nullmenge. Wir zeigen, dass $\varphi^{-1}(N) \subset U$ ebenfalls eine Nullmenge ist.

Nach Satz IX.2.7 (S. 303) gibt es eine monoton wachsende Folge $(f_k)_{k\in\mathbb{N}} \subset T(V,\mathbb{R})$ derart, dass $(\int_V f_k)_{k\in\mathbb{N}}$ beschränkt ist und $(f_k)_{k\in\mathbb{N}}$ auf $N$ divergiert. Gemäß Schritt 4 ist $f_k \circ \varphi |\det D\varphi| \in L^1(U,\mathbb{R})$ für alle $k \in \mathbb{N}$ und $\int_U f_k \circ \varphi |\det D\varphi| = \int_V f_k$ für alle $k \in \mathbb{N}$. Die Folge $(f_k \circ \varphi |\det D\varphi|)_{k\in\mathbb{N}}$ ist monoton wachsend und divergiert auf $\varphi^{-1}(N)$. Wegen Satz IX.3.7 (S. 321) ist daher $\varphi^{-1}(N)$ eine Nullmenge.

6. Schritt: Sei nun $f \in L^{\mathrm{inc}}(V,\mathbb{R})$ und $(f_k)_{k\in\mathbb{N}} \subset T(V,\mathbb{R})$ eine monoton wachsende Folge, derart dass $f_k \underset{k\to\infty}{\longrightarrow} f$ f.ü. in $V$ und $(\int_V f_k)_{k\in\mathbb{N}}$ beschränkt ist. Wegen der Schritte 4 und 5 ist $(f_k \circ \varphi |\det D\varphi|)_{k\in\mathbb{N}} \subset L^1(U,\mathbb{R})$ monoton wachsend, konvergiert f.ü. in $U$ gegen $f \circ \varphi |\det D\varphi|$ und $(\int_U f_k \circ \varphi |\det D\varphi|)_{k\in\mathbb{N}}$ ist beschränkt. Damit folgt die Behauptung für $f$ aus Satz IX.3.7 (S. 321).

7. Schritt: Aus dem 6. Schritt und der Linearität des Integrals folgt schließlich die Behauptung. □

Beispiel IX.5.7. (1) (Volumen von Kugeln in $\mathbb{R}^n$) Für $x \in \mathbb{R}^n$, $R \in \mathbb{R}_+^*$ sei

$$\begin{aligned} B_n(x,\mathbb{R}) &= \{y \in \mathbb{R}^n : \|y-x\|_2 \le R\} \\ B_n &= B_n(0,1) \\ \omega_n &= \lambda_n(B_n). \end{aligned}$$

Mit $\varphi(y) = x + Ry$ für alle $y \in \mathbb{R}^n$ folgt

$$B_n(x,R) = \varphi(B_n).$$

Es ist

$$D\varphi(y) = RId_{\mathbb{R}^n} \quad \forall y \in \mathbb{R}^n \Longrightarrow \det D\varphi(y) = R^n \quad \forall y \in \mathbb{R}^n.$$

Mit Satz IX.5.6 folgt

$$\lambda_n(B_n(x,R)) = R^n\omega_n \quad \forall x \in \mathbb{R}^n,\ R \in \mathbb{R}_+^*.$$

Aus Satz IX.3.3 (S. 317) und Satz IX.5.6 folgt weiter

$$\begin{aligned} \omega_n &= \int_{\mathbb{R}^n} \chi_{B_n} \\ &= \int_{\mathbb{R}} [\int_{\mathbb{R}^{n-1}} \chi_{B_n} dx_1 \dots dx_{n-1}] dx_n \\ &= \int_{-1}^{1} \lambda_{n-1}(B_{n-1}(x_n, \sqrt{1-x_n^2})) dx_n \\ &= \omega_{n-1} \int_{-1}^{1} (1-t^2)^{\frac{(n-1)}{2}} dt \quad (t = -\cos x) \\ &= \omega_{n-1} \int_0^{\pi} \sin^n x dx \\ &= 2\omega_{n-1} \int_0^{\frac{\pi}{2}} \sin^n x dx. \end{aligned}$$

Gemäß Beispiel VI.3.5(4) (S. 189) gilt für $A_n = \int_0^{\frac{\pi}{2}} \sin^n x dx$:

$$\begin{aligned} A_0 &= \frac{\pi}{2} \\ A_1 &= 1 \\ A_{2m} &= \frac{\pi}{2} \prod_{k=1}^{m} \frac{2k-1}{2k} \\ A_{2m+1} &= \prod_{k=1}^{m} \frac{2k}{2k+1}. \end{aligned}$$

Hieraus folgt

$$A_{n-1} A_n = \frac{\pi}{2n} \quad \forall n \in \mathbb{N}^*$$

und damit

$$\begin{aligned} \omega_n &= 2\omega_{n-1} A_n \\ &= 4\omega_{n-2} A_{n-1} A_n \\ &= \frac{2\pi}{n} \omega_{n-2} \quad \forall n \geq 3. \end{aligned}$$

Weiter ist

$$\begin{aligned} \omega_1 &= \lambda_n((-1,1)) \\ &= 2 \\ \omega_2 &= 2\omega_1 A_2 \\ &= 2 \cdot 2 \cdot \frac{\pi}{2} \cdot \frac{1}{2} \\ &= \pi. \end{aligned}$$

Hieraus folgt durch Induktion

$$\begin{aligned} \omega_{2m} &= \frac{\pi^m}{m!} \quad \forall m \geq 1 \\ \omega_{2m+1} &= \frac{2^{m+1}}{1 \cdot 3 \cdot \ldots \cdot (2m+1)} \pi^m \quad \forall m \geq 0. \end{aligned}$$

Gemäß Satz VI.5.1 (S. 205) und Satz VI.5.7 (S. 210) ist aber

$$\begin{aligned} \Gamma(m+1) &= m! \quad \forall m \geq 1 \\ \Gamma(m+\frac{3}{2}) &= (m+\frac{1}{2})\Gamma(m+\frac{1}{2}) \\ &= \Gamma(\frac{1}{2}) \prod_{k=0}^{m} \frac{2k+1}{2} \\ &= \sqrt{\pi} \prod_{k=0}^{m} \frac{2k+1}{2}. \end{aligned}$$

Damit folgt schließlich

$$\omega_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)} \quad \forall n \in \mathbb{N}^*.$$

(2) (Ebene Polarkoordinaten) Sei $R \in \mathbb{R}_+^*$ beliebig. Definiere

$$\begin{aligned} U_R &= (0,R) \times (0,2\pi) \\ V_R &= B_2(0,R) \backslash [\mathbb{R}_+ \times \{0\}] \\ \psi(r,\varphi) &= (r\cos\varphi, r\sin\varphi). \end{aligned}$$

Dann gilt

$$\begin{aligned} \psi &: U_R \to V_R \text{ ist bijektiv} \\ \psi(\overline{U}_R) &= \overline{V}_R \\ &= \overline{B_2(0,R)} \\ D\psi &= \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} \\ \det D\psi &= r > 0 \quad \forall (r,\varphi) \in U_R. \end{aligned}$$

Also ist $\psi$ ein $C^1$-Diffeomorphismus von $U_R$ auf $V_R$. Da $\mathbb{R}_+ \times \{0\}$ eine Nullmenge in $B_2(0,R)$ ist, folgt aus Satz IX.5.6

$$f \in L^1(B_2(0,R),\mathbb{R}) \iff rf(r\cos\varphi, r\sin\varphi) \in L^1(U_R,\mathbb{R}).$$

Für

$$f(x_1,x_2) = e^{x_1^2+x_2^2}$$

erhalten wir so z.B.

$$\begin{aligned} \int_{B_2(0,R)} f dx &= \int_{U_R} rf(r\cos\varphi, r\sin\varphi) dr d\varphi \\ &= \int_0^R \int_0^{2\pi} re^{r^2} d\varphi dr \\ &= 2\pi[\frac{1}{2}e^{r^2}]_0^R \\ &= \pi(e^{R^2}-1). \end{aligned}$$

(3) (Zylinderkoordinaten) Seien $R \in \mathbb{R}_+^*$, $a,b \in \mathbb{R}$ mit $a < b$ beliebig. Definiere

$$\begin{aligned} U &= (0,R) \times (0,2\pi) \times (a,b) \\ Z &= \{x \in \mathbb{R}^3 : a < x_3 < b, x_1^2 + x_2^2 < R\} \\ V &= Z \backslash [\mathbb{R}_+ \times \{0\} \times (a,b)] \\ \psi(r,\varphi,z) &= (r\cos\varphi, r\sin\varphi, z). \end{aligned}$$

Dann gilt

$$\begin{aligned} \psi &: U \to V \text{ ist bijektiv} \\ \psi(\overline{U}) &= \overline{V} \end{aligned}$$

$$
\begin{aligned}
&= \overline{Z} \\
D\psi &= \begin{pmatrix} \cos\varphi & -r\sin\varphi & 0 \\ \sin\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
\det D\psi &= r.
\end{aligned}
$$

Also ist $\psi$ ein $C^1$-Diffeomorphismus von $U$ auf $V$. Da $\mathbb{R}_+ \times \{0\} \times (a,b)$ eine Nullmenge in $Z$ ist, folgt aus Satz IX.5.6

$$
f \in L^1(Z,\mathbb{R}) \iff rf(r\cos\varphi, r\sin\varphi, z) \in L^1(U,\mathbb{R}).
$$

Für

$$
f(x) = x_3 e^{x_1^2+x_2^2}
$$

erhalten wir z.B.

$$
\begin{aligned}
\int_Z f dx &= \int_U rf(r\cos\varphi, r\sin\varphi, z) dr d\varphi dz \\
&= \int_0^R \int_0^{2\pi} \int_a^b zre^{r^2} dz d\varphi dr \\
&= \frac{1}{2}(b^2-a^2)\pi(e^{R^2}-1).
\end{aligned}
$$

(4) (Kugelkoordinaten) Sei $R \in \mathbb{R}_+^*$ beliebig. Definiere

$$
\begin{aligned}
U_R &= (0,R) \times (0,2\pi) \times (-\frac{\pi}{2}, \frac{\pi}{2}) \\
V_R &= B_3(0,R) \backslash [\mathbb{R}_+ \times \{0\} \times \mathbb{R}] \\
\psi(r,\varphi,\theta) &= (r\cos\varphi\cos\theta, r\sin\varphi\cos\theta, r\sin\theta).
\end{aligned}
$$

Dann gilt

$$
\begin{aligned}
&\psi : U_R \to V_R \text{ ist bijektiv} \\
&\psi(\overline{U}_R) = \overline{V}_R \\
&\qquad\quad = \overline{B_3(0,R)} \\
D\psi &= \begin{pmatrix} \cos\varphi\cos\theta & -r\sin\varphi\cos\theta & -r\cos\varphi\sin\theta \\ \sin\varphi\cos\theta & r\cos\varphi\cos\theta & -r\sin\varphi\sin\theta \\ \sin\theta & 0 & r\cos\theta \end{pmatrix} \\
\det D\psi &= \sin\theta[r^2\sin^2\varphi\sin\theta\cos\theta + r^2\cos^2\varphi\sin\theta\cos\theta] \\
&\quad + r\cos\theta[r\cos^2\varphi\cos^2\theta + r\sin^2\varphi\cos^2\theta] \\
&= r^2\sin^2\theta\cos\theta + r^2\cos^3\theta \\
&= r^2\cos\theta \\
&> 0 \quad \forall (r,\varphi,\theta) \in U_R.
\end{aligned}
$$

Also ist $\psi$ ein $C^1$-Diffeomorphismus von $U_R$ auf $V_R$. Da $[\mathbb{R}_+ \times \{0\} \times \mathbb{R}] \cap B_3(0,R)$ eine Nullmenge in $B_3(0,R)$ ist, folgt aus Satz IX.5.6

$$
f \in L^1(B_3(0,R),\mathbb{R})
$$

$$\iff r^2 \cos\theta f(r\cos\varphi\cos\theta, r\sin\varphi\cos\theta, r\sin\theta) \in L^1(U_R, \mathbb{R}).$$

Für

$$\begin{aligned} f(x) &= \frac{1}{\|x\|_2} e^{\|x\|_2^2} \\ &= \frac{e^{x_1^2+x_2^2+x_3^2}}{\sqrt{x_1^2+x_2^2+x_3^2}} \end{aligned}$$

erhalten wir z.B.

$$r^2 \cos\theta f(r\cos\varphi\cos\theta, r\sin\varphi\cos\theta, r\sin\theta) = r\cos\theta e^{r^2} \in L^1(U_R, \mathbb{R})$$

und

$$\begin{aligned} \int_{B_3(0,R)} f dx &= \int_0^R \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} r\cos\theta e^{r^2} d\theta d\varphi dr \\ &= 2\pi[\frac{1}{2}e^{r^2}]_0^R [\sin\theta]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \\ &= 2\pi(e^{R^2} - 1). \end{aligned}$$

## IX.6. Die $L^p$-Räume

In diesem Abschnitt ist stets $1 \le p < \infty$. Ist $p > 1$, so ist $p'$ gegeben durch

$$\frac{1}{p} + \frac{1}{p'} = 1$$

d.h.

$$p' = \frac{p}{p-1}.$$

Im Folgenden identifizieren wir Funktionen, die fast überall übereinstimmen. Genauer: Durch

$$f \sim g \iff f - g = 0 \text{ f.ü.}$$

wird auf $M(\mathbb{R}^n, \mathbb{R})$ eine Äquivalenzrelation definiert. Wenn wir im Folgenden von einer Funktion $f \in M(\mathbb{R}^n, \mathbb{R})$ sprechen, meinen wir dann immer die entsprechende Äquivalenzklasse, ohne diese genauer zu bezeichnen.

Definition IX.6.1. Wir definieren

$$L^p(\mathbb{R}^n, \mathbb{R}) = L^p = \{f \in M(\mathbb{R}^n, \mathbb{R}) : |f|^p \in L^1(\mathbb{R}^n, \mathbb{R})\}$$

und setzen für $f \in L^p$

$$\|f\|_p = \|f\|_{L^p(\mathbb{R}^n,\mathbb{R})} = \left\{ \int |f|^p \right\}^{1/p}.$$

Wir wollen zeigen, dass $(L^p, \|\cdot\|_p)$ ein Banachraum ist. Dazu benötigen wir das folgende Hilfsresultat.

Satz IX.6.2 (Höldersche Ungleichung). *Seien $p > 1$, $f \in L^p$ und $g \in L^{p'}$. Dann ist $f \cdot g \in L^1$ und*

$$\int |f \cdot g| \le \|f\|_p \|g\|_{p'}.$$

Beweis. Sei $0 < \alpha < 1$ und $f_\alpha : \mathbb{R}_+ \to \mathbb{R}$ definiert durch

$$f_\alpha(x) = (1 - \alpha) + \alpha x - x^\alpha.$$

Wie man leicht nachprüft, ist $f_\alpha$ auf $[0, 1]$ monoton fallend, auf $[1, \infty)$ monoton wachsend und $f_\alpha(1) = 0$. Also gilt

$$f_\alpha(x) \ge 0 \quad \forall x \in \mathbb{R}_+.$$

Seien nun $a, b \in \mathbb{R}_+^*$ und $\alpha = \frac{1}{p} \in (0, 1)$. Dann folgt

$$\begin{aligned} 0 &\le f_{\frac{1}{p}}(\frac{a^p}{b^{p'}}) \\ &= (1 - \frac{1}{p}) + \frac{1}{p}\frac{a^p}{b^{p'}} - \frac{a}{b^{p'/p}} \\ &= \frac{1}{p'} + \frac{1}{p}\frac{a^p}{b^{p'}} - \frac{ab}{b^{p'}} \\ \iff \quad ab &\le \frac{1}{p}a^p + \frac{1}{p'}b^{p'}. \end{aligned} \tag{$*$}$$

Die Ungleichung $(*)$ gilt offensichtlich auch für $a = 0$ oder $b = 0$.
O.E. sind $f \ne 0$, $g \ne 0$, da sonst die Behauptung trivial ist. Aus $(*)$ folgt mit

$$\begin{aligned} a &= \frac{|f(x)|}{\|f\|_p}, \\ b &= \frac{|g(x)|}{\|g\|_{p'}} \\ \frac{|f(x) \cdot g(x)|}{\|f\|_p \|g\|_{p'}} &\le \frac{1}{p}\frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{p'}\frac{|g(x)|^{p'}}{\|g\|_{p'}^{p'}} \quad \forall x \in \mathbb{R}^n. \end{aligned}$$

Aus Satz IX.4.5 (S. 331) und Satz IX.4.13 (S. 335) folgt $f \cdot g \in L^1$ und

$$\begin{aligned} \frac{1}{\|f\|_p \|g\|_{p'}} \int |f \cdot g| &\le \frac{1}{p} \cdot \frac{1}{\|f\|_p^p} \int |f|^p + \frac{1}{p'}\frac{1}{\|g\|_{p'}^{p'}} \int |g|^{p'} \\ &= \frac{1}{p} + \frac{1}{p'} \\ &= 1. \end{aligned}$$

□

Satz IX.6.3. *$(L^p, \| \cdot \|_p)$ ist ein Banachraum.*

Beweis. 1. Schritt: $L^p$ ist ein Vektorraum. Für $p = 1$ ist dies gerade Satz IX.2.14 (S. 309). Sei also $p > 1$. Die Eigenschaft

$$f \in L^p, \alpha \in \mathbb{R} \Longrightarrow \alpha f \in L^p$$

ist trivial. Seien also $f, g \in L^p$. Da die Funktion $x \mapsto x^p$ auf $\mathbb{R}_+^*$ konvex ist, folgt für alle $x \in \mathbb{R}^n$

$$\begin{aligned} |f(x) + g(x)|^p &= 2^p \Big| \frac{1}{2}(f(x) + g(x)) \Big|^p \\ &\leq 2^p \Big[ \frac{1}{2} |f(x)| + |g(x)| \Big]^p \\ &\leq 2^{p-1} \Big[ |f(x)|^p + |g(x)|^p \Big]. \end{aligned}$$

Hieraus und aus Satz IX.4.5 (S. 331) folgt $f + g \in L^p$.
2. Schritt: $\| \cdot \|_p$ ist eine Norm. Die Eigenschaften

$$\begin{aligned} \|f\|_p &\geq 0 \quad \forall f \in L^p \\ \|f\|_p &= 0 \iff f = 0 \\ \|\alpha f\|_p &= |\alpha| \|f\|_p \quad \forall f \in L^p, \ \alpha \in \mathbb{R} \end{aligned}$$

sind trivial. (Man beachte unsere Vereinbarung bzgl. Funktionen, die f.ü. übereinstimmen.) Die Dreiecksungleichung ist für $p = 1$ trivial. Sei also $p > 1$ und $f, g \in L^p$. O.E. ist $f \neq 0$, $g \neq 0$. Aus Satz IX.6.2 folgt

$$\begin{aligned} \|f + g\|_p^p &= \int |f + g|^p \\ &= \int |f + g|^{p-1} |f + g| \\ &\leq \int |f + g|^{p-1} [|f| + |g|] \\ &\leq \|f\|_p \Big\{ \int |f + g|^{(p-1)p'} \Big\}^{1/p'} \\ &\quad + \|g\|_p \Big\{ \int |f + g|^{(p-1)p'} \Big\}^{1/p'} \\ &= \Big[ \|f\|_p + \|g\|_p \Big] \|f + g\|_p^{p-1}. \end{aligned}$$

Hieraus folgt die Dreiecksungleichung.
3. Schritt: $(L^p, \| \cdot \|_p)$ ist vollständig. Sei $(f_k)_{k \in \mathbb{N}} \subset L^p$ eine Cauchyfolge. Dann gibt es zu jedem $k \in \mathbb{N}^*$ ein $N(k) \in \mathbb{N}$ mit

$$\begin{aligned} N(k+1) &> N(k) \\ \|f_m - f_{N(k)}\|_p &< 2^{-k} \quad \forall m \geq N(k). \end{aligned}$$

Definiere

$$\begin{aligned} g_1 &= f_{N(1)} \\ g_k &= f_{N(k)} - f_{N(k-1)} \quad \forall k \geq 2. \end{aligned}$$

und

$$h_k = \sum_{l=1}^{k} |g_l|.$$

Dann ist $(h_k)_{k\in\mathbb{N}^*}$ eine monoton wachsende Folge positiver Funktionen in $L^p$ mit

$$\begin{aligned}\|h_k\|_p &\le \sum_{l=1}^{k} \|g_l\|_p \\ &\le \|f_{N(1)}\|_p + \sum_{l=2}^{k} 2^{-l+1} \\ &< 1 + \|f_{N(1)}\|_p.\end{aligned}$$

Wegen Satz IX.3.7 (S. 321) konvergiert $(h_k^p)_{k\in\mathbb{N}}$ f.ü. gegen eine positive Funktion $\widetilde{h} \in L^1$ mit

$$\int \widetilde{h} = \int |\widetilde{h}| \le [1 + \|f_{N(1)}\|_p]^p.$$

Definiere $h = \widetilde{h}^{1/p}$. Dann gilt $h \in L^p$, $\|h\|_p \le 1 + \|f_{N(1)}\|_p$ und $h_k \underset{h\to\infty}{\longrightarrow} h$ f.ü.. Daher konvergiert $\big(\sum_{l=1}^{k} g_l\big)_{k\in\mathbb{N}}$ f.ü. gegen eine messbare Funktion $f$ mit $|f| \le h$. Also ist $f \in L^p$. Wegen

$$\sum_{l=1}^{k} g_l - f \underset{k\to\infty}{\longrightarrow} 0 \text{ f.ü.}$$

und

$$|\sum_{l=1}^{k} g_l - f| \le h_k + |f| \le 2h$$

folgt aus Satz IX.3.11 (S. 324)

$$\int |\sum_{l=1}^{k} g_l - f|^p \underset{k\to\infty}{\longrightarrow} 0,$$

d.h.

$$\|f - f_{N(k)}\|_p = \|f - \sum_{l=1}^{k} g_l\|_p \underset{k\to\infty}{\longrightarrow} 0.$$

Sei nun $\varepsilon > 0$ beliebig. Dann gibt es ein $k \in \mathbb{N}^*$ mit

$$2^{-k} < \frac{\varepsilon}{2}$$

und

$$\|f - f_{N(k)}\|_p < \frac{\varepsilon}{2}.$$

Für $m \geq N(k)$ folgt

$$\begin{aligned}\|f - f_m\|_p &\leq \|f - f_{N(k)}\|_p + \|f_{N(k)} - f_m\|_p \\ &< \frac{\varepsilon}{2} + 2^{-k} \\ &\leq \varepsilon.\end{aligned}$$

Also konvergiert $(f_k)_{k\in\mathbb{N}}$ in $L^p$ gegen $f$. Damit ist die Behauptung gezeigt. □

Aus Teil (3) des obigen Beweises folgt sofort das folgende Resultat:

Satz IX.6.4. *Seien $f \in L^p$ und $(f_k)_{k\in\mathbb{N}} \subset L^p$ mit*

$$\lim_{k\to\infty} \|f_k - f\|_p = 0.$$

*Dann gibt es eine Teilfolge $(f_{k_l})_{l\in\mathbb{N}}$ von $(f_k)_{k\in\mathbb{N}}$ mit $f_{k_l} \underset{l\to\infty}{\longrightarrow} f$ f.ü..*

Bemerkung IX.6.5. Sei $X$ ein Vektorraum. Auf $X$ gebe es ein Skalarprodukt, d.h. eine symmetrische Bilinearform $(\cdot,\cdot) : X \times X \to \mathbb{R}$ mit den Eigenschaften

(S1) $(u,u) \geq 0 \quad \forall u \in X$,
(S2) $(u,u) = 0 \iff u = 0$.

Dann wird durch

$$\|u\| = (u,u)^{1/2} \quad \forall u \in X$$

eine Norm auf $X$ definiert (Beweis!). Ist $X$ mit dieser Norm vollständig, so heißt $X$ ein Hilbertraum. Satz IX.6.2 und Satz IX.6.3 zeigen, dass $L^2$ mit

$$(f,g) = \int f \cdot g \quad \forall f,g \in L^2$$

ein Hilbertraum ist. Andere Beispiele für Hilberträume sind $\mathbb{R}^n$ mit

$$(x,y) = \sum_{i=1}^{n} x_i y_i \quad \forall x,y \in \mathbb{R}^n$$

und $\ell^2$ mit

$$(u,v) = \sum_{k=1}^{\infty} u_k v_k \quad \forall u,v \in \ell^2.$$

Wir wollen zeigen, dass Funktionen in $L^p$ „beliebig gut“ durch „glatte Funktionen“ approximiert werden können. Dazu benötigen wir einige Begriffe und Resultate, die uns im nächsten Kapitel von Nutzen sein werden.

Definition IX.6.6. (1) Seien $A$, $B$ Teilmengen eines normierten Raumes $(X, \|\cdot\|)$. Dann schreiben wir $A \Subset B$, wenn $\overline{A}$ kompakt und $\overline{A} \subset B$ ist.
(2) Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume, $U \subset X$ und $f : U \to Y$ eine Funktion. Dann heißt

$$\operatorname{supp}(f) = \overline{\{x \in U : f(x) \neq 0\}}$$

der TRÄGER von $f$.
(3) Sei $U \subset \mathbb{R}^n$ offen, nicht leer und $k \in \mathbb{N} \cup \{+\infty\}$. Dann ist

$$C_0^k(U, \mathbb{R}) = \{f \in C^k(U, \mathbb{R}) : \operatorname{supp}(f) \Subset U\}.$$

SATZ IX.6.7. *Sei $U \subset \mathbb{R}^n$ offen und $K \subset U$, $K \neq \emptyset$, kompakt. Dann existiert eine* ABSCHNEIDEFUNKTION $\varphi \in C_0^\infty(\mathbb{R}^n, \mathbb{R})$ *mit*

(1) $\operatorname{supp}(\varphi) \Subset U$,
(2) $0 \leq \varphi(x) \leq 1$ *für alle* $x \in \mathbb{R}^n$,
(3) $\varphi(x) = 1$ *für alle* $x \in K$.

BEWEIS. Da $K$ kompakt ist, gibt es ein $m \in \mathbb{N}$, Punkte $x_0, \ldots, x_m \in K$ und Zahlen $\delta_0, \ldots, \delta_m \in \mathbb{R}_+^*$ mit

$$\overline{B_{|.|_2}(x_i, \delta_i)} \subset U \quad \forall 0 \leq i \leq m$$
$$K \subset \bigcup_{i=0}^{m} B_{|.|_2}(x_i, \delta_i),$$

wobei $|.|_2$ die euklidische Norm auf $\mathbb{R}^n$ bezeichnet. Sei $\delta = \min\limits_{0 \leq i \leq m} \delta_i$ und $0 < \varepsilon \leq \frac{\delta}{4}$. Definiere

$$K_1 = \{x \in \mathbb{R}^n : \min_{y \in K} |x - y|_2 \leq \varepsilon\}$$
$$K_2 = \{x \in \mathbb{R}^n : \min_{y \in K} |x - y|_2 \leq 3\varepsilon\}.$$

Dann sind $K_1$ und $K_2$ kompakt und

$$K \subset K_1 \subset K_2 \subset U.$$

Definiere die Funktion $\widetilde{\psi} : \mathbb{R}^n \to \mathbb{R}$ durch

$$\widetilde{\psi}(x) = \begin{cases} \exp\{-\frac{1}{1-|x|_2^2}\} & \text{falls } |x|_2 < 1, \\ 0 & \text{sonst.} \end{cases}$$

Wegen Beispiel IV.1.21 (S. 120) und Satz VII.4.12 (S. 235) ist $\widetilde{\psi} \in C_0^\infty(\mathbb{R}^n, \mathbb{R})$ mit $\operatorname{supp}(\widetilde{\psi}) = \overline{B_{|.|_2}(0, 1)}$. Insbesondere ist $\widetilde{\psi} \in L^1(\mathbb{R}^n, \mathbb{R})$. Definiere

$$\psi = \left[\int_{\mathbb{R}^n} \widetilde{\psi}\right]^{-1} \cdot \widetilde{\psi}.$$

Aus den Eigenschaften von $\psi$ und Satz IX.3.15 (S. 328) folgt, dass die Funktion $\varphi : \mathbb{R}^n \to \mathbb{R}$ mit

$$\varphi(x) = \int_{\mathbb{R}^n} \varepsilon^{-n} \psi(\frac{y - x}{\varepsilon}) \chi_{K_1}(y) dy \quad \forall x \in \mathbb{R}^n$$

aus $C^\infty(\mathbb{R}^n, \mathbb{R})$ ist. Aus Satz IX.5.6 (S. 343) folgt

$$\varphi(x) = \int_{\mathbb{R}^n} \psi(z) \chi_{K_1}(x + \varepsilon z) dz \quad \forall x \in \mathbb{R}^n.$$

Ist $x \notin K_2$, so ist $\overline{B_{|.|_2}(x,\varepsilon)} \cap K_1 = \emptyset$ und daher $\varphi(x) = 0$. Also ist $\mathrm{supp}(\varphi) \subset K_2 \subset U$. Ist dagegen $x \in K$, so ist $\overline{B_{|.|_2}(x,\varepsilon)} \subset K_1$ und daher

$$\varphi(x) = \int_{\mathbb{R}^n} \psi(z)dz = 1.$$

Schließlich gilt für alle $x \in \mathbb{R}^n$

$$0 \le \varphi(x) \le \int_{\mathbb{R}^n} \psi(z)dz = 1.$$

Damit leistet $\varphi$ das Gewünschte. □

Bemerkung IX.6.8. Aus Beispiel IV.1.21 (S. 120), Satz VII.4.12 (S. 235) und Satz IX.3.15 (S. 328) folgt, dass es für jedes $k \in \mathbb{N}^*$ ein $C_k \in \mathbb{R}_+$ gibt mit

$$\sup_{x\in\mathbb{R}^n} \|D^k\varphi(x)\|_{\mathcal{L}^k(\mathbb{R}^n,\mathbb{R}^n)} \le c_k\varepsilon^{-k}.$$

Satz IX.6.9. *Seien $K \subset \mathbb{R}^n$, $K \neq \emptyset$, kompakt und $U_0, \ldots, U_m \subset \mathbb{R}^n$ offen mit $K \subset \bigcup_{i=0}^m U_i$. Dann gibt es offene Mengen $V_0, \ldots, V_m$ mit*

$$V_i \Subset U_i \quad \forall 0 \le i \le m$$
$$K \subset \bigcup_{i=0}^m V_i.$$

Beweis. Zu jedem $x \in K$ gibt es ein $0 \le i \le m$ und ein $\varepsilon_x \in \mathbb{R}_+^*$ mit $x \in U_i$ und $\overline{B_{|.|_2}(x,\varepsilon_x)} \subset U_i$. Da $K$ kompakt ist, gibt es ein $k \in \mathbb{N}$ und Punkte $x_0, \ldots, x_k \in K$ mit

$$K \subset \bigcup_{j=0}^k B_{|.|_2}(x_j, \varepsilon_{x_j}).$$

Für $0 \le i \le m$ sei

$$V_i = \bigcup \{B_{|.|_2}(x_j, \varepsilon_{x_j}) : \overline{B_{|.|_2}(x_j, \varepsilon_{x_j})} \subset U_i\}.$$

Dann ist $V_i \Subset U_i$ für alle $0 \le i \le m$ und $K \subset \bigcup_{i=0}^m V_i$. □

Satz IX.6.10. *Seien $K \subset \mathbb{R}^n$, $K \neq \emptyset$, kompakt und $U_0, \ldots, U_m \subset \mathbb{R}^n$ offen mit $K \subset \bigcup_{i=0}^m U_i$. Dann gibt es eine* Partition der Eins *auf $K$, die der Überdeckung $U_0, \ldots, U_m$ untergeordnet ist, d.h., es existieren Funktionen $\varphi_0, \ldots, \varphi_m \in C_0^\infty(\mathbb{R}^n, \mathbb{R})$ mit*

(1) $\mathrm{supp}(\varphi_i) \Subset U_i$ *für alle* $0 \le i \le m$,
(2) $0 \le \varphi_i \le 1$ *für alle* $0 \le i \le m$,

(3) $\sum_{i=0}^{m} \varphi_i = 1$ *auf* $K$.

Beweis. Seien $V_0, \dots, V_m$ wie in Satz IX.6.9. Gemäß Satz IX.6.7 gibt es Funktionen $\psi_0, \dots, \psi_m \in C_0^\infty(\mathbb{R}^n, \mathbb{R})$ mit

(a) $\operatorname{supp}(\psi_i) \Subset U_i$ für alle $0 \le i \le m$.
(b) $0 \le \psi_i \le 1$ für alle $0 \le i \le m$,
(c) $\psi_i = 1$ auf $\overline{V}_i$, für alle $0 \le i \le m$.

Definiere

$$\begin{aligned} \varphi_0 &= \psi_0 \\ \varphi_j &= \psi_j \prod_{i=0}^{j-1} (1 - \psi_i) \quad 1 \le j \le m. \end{aligned}$$

Offensichtlich erfüllen $\varphi_0, \dots, \varphi_m$ die Bedingungen (1) und (2). Durch Induktion folgt

$$\sum_{j=0}^{m} \varphi_j = 1 - \prod_{i=0}^{m} (1 - \psi_i).$$

Hieraus folgt die Bedingung (3). $\square$

Wir kommen nun zu dem angekündigten Ergebnis über glatte Funktionen in $L^p$.

Satz IX.6.11. $C_0^\infty(\mathbb{R}^n, \mathbb{R})$ *ist dicht in* $L^p(\mathbb{R}^n, \mathbb{R})$.

Beweis. 1. Schritt: $T(\mathbb{R}^n, \mathbb{R})$ ist dicht in $L^p(\mathbb{R}^n, \mathbb{R})$. Sei $f \in L^p$ beliebig. Wir müssen zeigen, dass es zu jedem $\varepsilon > 0$ ein $\varphi \in T(\mathbb{R}^n, \mathbb{R})$ gibt mit $\|f - \varphi\|_p < \varepsilon$. Wegen $f = f^+ - f^-$ und $f^+, f^- \in L^p$, $f^+ \ge 0$, $f^- \ge 0$ können wir o.E. annehmen, dass $f \ge 0$ ist.
Wegen Lemma IX.3.6 gibt es zu jedem $k \in \mathbb{N}^*$ Funktionen $\widetilde{g}_k, \widetilde{h}_k \in L^{\mathrm{inc}}$ mit

$$\begin{aligned} \widetilde{g}_k &\ge 0, \\ \widetilde{h}_k &\ge 0, \\ f^p &= \widetilde{g}_k - \widetilde{h}_k \\ \int \widetilde{h}_k &< \frac{1}{k}, \end{aligned}$$

d.h.

$$\|f^p - \widetilde{g}_k\|_1 \underset{k\to\infty}{\longrightarrow} 0.$$

Nach Übergang zu einer Teilfolge, die wir wieder mit $(\widetilde{g}_k)_{k\in\mathbb{N}}$ bezeichnen, folgt aus Satz IX.6.4

$$\widetilde{g}_k \underset{k\to\infty}{\longrightarrow} f^p \text{ f.ü..}$$

Sei $g_k = \widetilde{g}_k^{1/p}$. Dann gilt

$$g_k \underset{k\to\infty}{\longrightarrow} f \text{ f.ü.}$$

also

$$|g_k - f|^p \underset{k\to\infty}{\longrightarrow} 0 \text{ f.ü.}$$

und

$$\begin{aligned}\int |g_k - f|^p &= \|g_k - f\|_p^p \\ &\le [\|g_k\|_p + \|f\|_p]^p \\ &= [\|\widetilde{g}_k\|_1^{1/p} + \|f\|_p]^p \\ &\le [\{\|f\|_p^p + \frac{1}{k}\}^{1/p} + \|f\|_p]^p.\end{aligned}$$

Also ist $(\int |g_k - f|^p)_{k\in\mathbb{N}}$ beschränkt und aus Satz IX.3.11 (S. 324) folgt

$$\lim_{k\to\infty} \|g_k - f\|_p = 0.$$

Wir können also ohne Einschränkung weiter annehmen, dass $f^p \in L^{\text{inc}}$ ist. Dann gibt es eine monoton wachsende Folge $(\widetilde{\varphi}_k)_{k\in\mathbb{N}} \subset T(\mathbb{R}^n, \mathbb{R})$ mit

$$\widetilde{\varphi} \underset{k\to\infty}{\longrightarrow} f^p \text{ f.ü.}$$

$$\int \widetilde{\varphi}_k \underset{k\to\infty}{\longrightarrow} \int f^p.$$

Indem wir gegebenenfalls zu $\widetilde{\varphi}_k^+$ übergehen, können wir o.E. $\widetilde{\varphi}_k \ge 0$ für alle $k \in \mathbb{N}$ annehmen. Sei $\varphi_k = \widetilde{\varphi}_k^{1/p} \in T(\mathbb{R}^n, \mathbb{R})$. Dann gilt

$$\varphi_k \underset{k\to\infty}{\longrightarrow} f \text{ f.ü.} \quad \text{und} \quad \int \varphi_k^p \underset{k\to\infty}{\longrightarrow} \int f^p.$$

Mit dem gleichen Argument wie oben folgt

$$\lim_{k\to\infty} \|f - \varphi_k\|_p = 0.$$

2. Schritt: $C_0^\infty(\mathbb{R}^n, \mathbb{R})$ ist dicht in $T(\mathbb{R}^n, \mathbb{R})$. Offensichtlich reicht es zu zeigen, dass es zu jedem beschränkten Intervall $I$ und jedem $\varepsilon > 0$ ein $\varphi \in C_0^\infty(\mathbb{R}^n, \mathbb{R})$ gibt mit

$$\|\varphi - \chi_I\|_p < \varepsilon.$$

Sei also $I$ ein beschränktes Intervall und $\varepsilon > 0$. Da $\partial I$ eine Nullmenge ist, gibt es offene Intervalle $I_k$ mit

$$\partial I \subset \bigcup_{k\in\mathbb{N}} I_k \quad \text{und} \quad \sum_{k\in\mathbb{N}} \lambda_n(I_k) < \varepsilon^p.$$

Sei $U = \overline{I} \cup \bigcup_{k\in\mathbb{N}} I_k$. Dann ist $U$ offen und

$$\lambda_n(U\backslash\overline{I}) \leq \sum_{k\in\mathbb{N}} \lambda_n(I_k) < \varepsilon^p.$$

Gemäß Satz IX.6.7 gibt es ein $\varphi \in C_0^\infty(\mathbb{R}^n, \mathbb{R})$ mit

$$\begin{aligned} \operatorname{supp}(\varphi) &\Subset U \\ 0 \leq \varphi &\leq 1 \\ \varphi &= 1 \quad \text{auf } \overline{I}. \end{aligned}$$

Dann folgt

$$\|\chi_I - \varphi\|_p \leq \lambda_n(U\backslash\overline{I})^{1/p} < \varepsilon.$$

Damit ist die Behauptung gezeigt. □

KAPITEL X

# Analysis auf Mannigfaltigkeiten

Anknüpfend an den Satz über implizite Funktionen VII.5.9 (S. 249) definieren wir eine Mannigfaltigkeit als eine Menge, die lokal als Nullstellenmenge einer glatten Funktion mit surjektiver Funktionalmatrix dargestellt werden kann. Sie ist dann lokal auch darstellbar als Graph einer glatten Funktion. Ausgehend von diesen Eigenschaften definieren wir den Tangentialraum und die Orientierung einer Mannigfaltigkeit.
Anschließend definieren wir das Integral von Funktionen auf Mannigfaltigkeiten, indem wir es mit Hilfe der lokalen Darstellungen auf ein Lebesgue Integral im $\mathbb{R}^n$ zurückführen.
Als Anwendung berechnen wir verschiedene Oberflächen und zeigen den Integralsatz von Gauß zusammen mit einigen wichtigen Anwendungen aus der Physik.
Danach wenden wir uns dem Differentialformen-Kalkül zu. Differentialformen sind alternierende Multilinearformen auf den Tangentialräumen von Mannigfaltigkeiten. Als Spezialfall der hier gezeigten allgemeinen Ergebnisse, erhalten wir die Resultate aus Kapitel VIII über Kurvenintegrale und Gradientenfelder zurück. Als Hauptresultat der Theorie beweisen wir den Stokesschen Integralsatz in seiner allgemeinen Form und interpretieren ihn danach in der „klassischen Sprache" der Vektoranalysis. Auf diese Weise erhalten wir einige für die Anwendungen wichtige Ergebnisse.

## X.1. Mannigfaltigkeiten

Sofern nicht anders vermerkt, ist im Folgenden stets $1 \leq k < n$, $\alpha \in \mathbb{N}^*$ und $M \subset \mathbb{R}^n$ nicht leer. Wir erinnern an die Ergebnisse aus Abschnitt VII.5, insbesondere an Satz VII.5.9 (S. 249) und Bemerkung VII.5.13 (S. 253).

Definition X.1.1. Eine Teilmenge $M$ des $\mathbb{R}^n$ heißt $k$-dimensionale (Unter-) Mannigfaltigkeit der Klasse $C^\alpha$, wenn es zu jedem $x_0 \in M$ eine Umgebung $U \in \mathcal{U}(x_0)$ im $\mathbb{R}^n$ und ein $f \in C^\alpha(U, \mathbb{R}^{n-k})$ gibt mit

(1) $M \cap U = f^{-1}(\{0\})$,
(2) $x_0$ ist ein regulärer Punkt von $f$, d.h., $\operatorname{Rang}(Df(x_0)) = n - k$.

Beispiel X.1.2. (1) Die Sphäre

$$S^{n-1} = \{x \in \mathbb{R}^n : \sum_{i=1}^{n} x_i^2 = 1\}$$

ist eine $(n-1)$-dimensionale Mannigfaltigkeit der Klasse $C^\infty$ in $\mathbb{R}^n$.
(2) Seien $a_1, \dots, a_n \in \mathbb{R}_+^*$. Das Ellipsoid

$$\mathcal{E} = \{x \in \mathbb{R}^n : \sum_{i=1}^{n} (\frac{x_i}{a_i})^2 = 1\}$$

ist eine $(n-1)$-dimensionale Mannigfaltigkeit der Klasse $C^\infty$ in $\mathbb{R}^n$.
(3) Sei $c \in \mathbb{R}_+$. Das Hyperboloid

$$\mathcal{H} = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 - x_3^2 = c^2\}$$

ist für $c > 0$ eine 2-dimensionale Mannigfaltigkeit im $\mathbb{R}^3$. Für $c = 0$ ist $\mathcal{H}$ keine Mannigfaltigkeit.
(4) Seien $0 < r < R$. Der Torus

$$\mathcal{T} = \{x \in \mathbb{R}^3 : (\sqrt{x_1^2 + x_2^2} - R)^2 + x_3^2 = r^2\}$$

ist eine 2-dimensionale Mannigfaltigkeit der Klasse $C^\infty$ in $\mathbb{R}^3$.

Beweis. ad (1): Beispiel VII.5.10 (S. 251).
ad (2): Es ist $\mathcal{E} = f^{-1}(\{0\})$ mit

$$f(x) = \sum_{i=1}^{n} (\frac{x_i}{a_i})^2 - 1$$

und

$$Df(x) = (2\frac{x_1}{a_1}, \dots, 2\frac{x_n}{a_n}) \neq 0 \quad \forall x \in \mathcal{E}.$$

ad (3): Es ist $\mathcal{H} = f^{-1}(\{0\})$ mit

$$f(x) = x_1^2 + x_2^2 - x_3^2 - c^2$$

und

$$Df(x) = (2x_1, 2x_2, -2x_3) \neq 0 \quad \forall x \in \mathcal{H} \text{ sofern } c > 0.$$

ad (4): Es ist $\mathcal{T} = f^{-1}(\{0\})$ mit

$$f(x) = (\sqrt{x_1^2 + x_2^2} - R)^2 + x_3^2 - r^2 \quad \forall x \neq 0$$

und

$$\begin{aligned} &Df(x) \\ &= \Big(2(\sqrt{x_1^2 + x_2^2} - R)\frac{x_1}{\sqrt{x_1^2 + x_2^2}}, 2(\sqrt{x_1^2 + x_2^2} - R)\frac{x_2}{\sqrt{x_1^2 + x_2^2}}, 2x_3\Big) \\ &\neq 0 \quad \forall x \in T. \end{aligned}$$

□

Wenn wir im Folgenden ohne weiteren Zusatz von einer Mannigfaltigkeit (kurz MFGKT) sprechen, so meinen wir stets eine Untermannigfaltigkeit der Klasse $C^1$. Im Folgenden geben wir verschiedene Charakterisierungen und Darstellungen von Mfgkten an. Dabei werden wir später diejenige wählen, die für den konkreten Fall am geeignetesten ist. Die folgende Charakterisierung ist bereits in Bemerkung VII.5.13 (S. 253) gegeben.

SATZ X.1.3. *$M \subset \mathbb{R}^n$ ist genau dann eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$, wenn es zu jedem $x_0 \in M$ nach eventueller Umnummerierung der Koordinaten Umgebungen $U_1$ von $x_0' = (x_{01}, \dots, x_{0k})$ in $\mathbb{R}^k$ und $U_2$ von $x_0'' = (x_{0k+1}, \dots, x_{0n})$ in $\mathbb{R}^{n-k}$ und eine Funktion $g \in C^\alpha(U_1, \mathbb{R}^{n-k})$ gibt mit*

(1) $g(U_1) = U_2$,
(2) $M \cap (U_1 \times U_2) = \{(x', x'') \in U_1 \times U_2 : x'' = g(x')\}$.

BEWEIS. „$\Longrightarrow$“: Folgt aus dem Satz über implizite Funktionen, Satz VII.5.9 (S. 249) (vgl. Bemerkung VII.5.13 (S. 253)).
„$\Longleftarrow$“: Folgt mit

$$f(x', x'') = x'' - g(x') \quad \forall (x', x'') \in U_1 \times U_2.$$

□

Mfgkten der Dimension $k$ können also lokal als Graph einer $C^\alpha$-Funktion von $k$-Variablen dargestellt werden. Wir erinnern daran (Bemerkung VII.5.13 (S. 253)), dass wir im Spezialfall $k = n - 1$ auch von HYPERFLÄCHEN reden. Die folgende Charakterisierung zeigt, dass sich $k$-dimensionale Mfgkten lokal wie $k$-dimensionale Ebenen im $\mathbb{R}^n$ verhalten.

SATZ X.1.4. *Es sei*

$$E_k = \{x \in \mathbb{R}^n : x_{k+1} = \ldots = x_n = 0\}$$

*die $k$-dimensionale Ebene im $\mathbb{R}^n$. Dann ist $M \subset \mathbb{R}^n$ genau dann eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$, wenn es zu jedem $x_0 \in M$ eine Umgebung $U$ von $x_0$ in $\mathbb{R}^n$ und eine Umgebung $V$ von $0$ in $\mathbb{R}^n$ und einen $C^\alpha$-Diffeomorphismus $F : U \to V$ von $U$ auf $V$ gibt mit*

$$F(M \cap U) = V \cap E_k.$$

BEWEIS. „$\Longrightarrow$“: Seien $x_0 \in M$ und $x_0'$, $x_0''$, $U_1$, $U_2$ und $g$ wie in Satz X.1.3. Dann leisten $U = U_1 \times U_2$ und $F : U \to \mathbb{R}^n$ mit

$$F(x', x'') = (x', x'' - g(x')) \quad \forall (x', x'') \in U_1 \times U_2$$

das Gewünschte.
„$\Longleftarrow$“: Definiere $f : U \to \mathbb{R}^{n-k}$ durch

$$f(x) = (F_{k+1}(x), \dots, F_n(x)) \quad \forall x \in U.$$

Dann ist $M \cap U = f^{-1}(\{0\})$ und

$$\mathrm{Rang}(Df(x_0)) = n - k,$$

da $\mathrm{Rang}(DF(x_0)) = n$ ist. Hieraus folgt die Behauptung. $\square$

DEFINITION X.1.5. Sei $U \subset \mathbb{R}^k$ offen. Eine Abbildung $\varphi \in C^\alpha(U, \mathbb{R}^n)$ heißt IMMERSION (DER KLASSE $C^\alpha$), wenn jedes $x \in U$ regulärer Punkt von $\varphi$ ist. Ist $\varphi \in C^\alpha(U, \mathbb{R}^n)$ eine Immersion und $\varphi : U \to V = \varphi(U)$ ein Homöomorphismus, so schreiben wir kurz: $\varphi : U \overset{\alpha}{\hookrightarrow} V$.

SATZ X.1.6. *$M \subset \mathbb{R}^n$ ist genau dann eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$, wenn es zu jedem $x_0 \in M$ eine Umgebung $V$ von $x_0$ in $M$, eine offene Menge $U \subset \mathbb{R}^k$ und eine Immersion $\varphi : U \overset{\alpha}{\hookrightarrow} V$ gibt.*

BEWEIS. „$\Longrightarrow$“: Seien $x_0, x_0', x_0'', U_1, U_2$ und $g$ wie in Satz X.1.3. Dann leisten $U = U_1, V = M \cap (U_1 \times U_2)$ und

$$\varphi(t) = (t, g(t)) \quad \forall t \in U$$

das Gewünschte.
„$\Longleftarrow$“: Sei $t_0 = \varphi^{-1}(x_0)$. Da $t_0$ ein regulärer Punkt von $\varphi$ ist, können wir o.E. annehmen, dass $(D_i\varphi_j(t_0))_{1 \le i,j \le k}$ regulär ist, sonst führen wir eine Koordinatentransformation durch. Gemäß Satz VII.5.4 (S. 246) gibt es Umgebungen $\widetilde{U} \subset U$ von $t_0$ und $\widetilde{V}$ von $x_0' = (x_{01}, \ldots, x_{0k})$ in $\mathbb{R}^k$, so dass $(\varphi_1, \ldots, \varphi_k) : \widetilde{U} \to \widetilde{V}$ ein $C^\alpha$-Diffeomorphismus ist. Definiere $F : \widetilde{U} \times \mathbb{R}^{n-k} \to \mathbb{R}^n$ durch

$$\begin{aligned} F_i(z_1, \ldots, z_n) &= \varphi_i(z_1, \ldots, z_k) && 1 \le i \le k, \\ F_j(z_1, \ldots, z_n) &= z_j + \varphi_j(z_1, \ldots, z_k) && k+1 \le j \le n. \end{aligned}$$

Dann ist $F : \widetilde{U} \times \mathbb{R}^{n-k} \to \widetilde{V} \times \mathbb{R}^{n-k}$ ein $C^\alpha$-Diffeomorphismus mit $F((\widetilde{U} \times \mathbb{R}^{n-k}) \cap E_k) = (\widetilde{V} \times \mathbb{R}^{n-k}) \cap M$. Damit folgt die Behauptung aus Satz X.1.4 angewandt auf $F^{-1}$. $\square$

DEFINITION X.1.7. Seien $\varphi, U$ und $V$ wie in Satz X.1.6. Dann heißt $(U, \varphi, V)$ eine LOKALE PARAMETERDARSTELLUNG oder KARTE von $M$. Eine Menge $\mathcal{A} = \{(U_\lambda, \varphi_\lambda, V_\lambda) : \lambda \in \Lambda\}$ von Karten von $M$ mit $M \subset \bigcup_{\lambda \in \Lambda} V_\lambda$ heißt ATLAS von $M$.

BEMERKUNG X.1.8. Da $\mathbb{Q}^n$ abzählbar und dicht in $\mathbb{R}^n$ ist, besitzt jede Mfgkt einen Atlas aus höchstens abzählbar vielen Karten.

BEISPIEL X.1.9. (1) Definiere $\varphi : \mathbb{R} \to \mathbb{R}^2$ durch

$$\varphi(t) = (\cos t, \sin t).$$

Dann ist $\varphi$ eine Immersion und

$$\left\{ \left( (0, \frac{3}{2}\pi), \varphi|_{(0,\frac{3}{2}\pi)}, \varphi((0, \frac{3}{2}\pi)) \right), \left( (\pi, \frac{5}{2}\pi), \varphi|_{(\pi,\frac{5}{2}\pi)}, \varphi((\pi, \frac{5}{2}\pi)) \right) \right\}$$

ein Atlas von $S^1$.
(2) Definiere $\varphi : \mathbb{R} \times (-1,1) \to \mathbb{R}^3$ durch

$$\varphi(t,r) = \Big(\cos(2t)(2 - r\sin t), \sin(2t)(2 - r\sin t), r\cos t\Big).$$

$\mathcal{M} = \varphi(\mathbb{R} \times (-1,1))$ heißt Möbiusband. Wie man leicht nachprüft, ist für jedes $t_0 \in \mathbb{R}$ $\varphi|_{(t_0-\frac{\pi}{2},t_0+\frac{\pi}{2})\times(-1,1)}$ ein Homöomorphismus. Weiter gilt für jedes $(t,r) \in \mathbb{R} \times (-1,1)$

$$\begin{aligned}
&\operatorname{Rang}(D\varphi(t,r)) \\
&= \operatorname{Rang}\begin{pmatrix} -2\sin(2t)(2 - r\sin t) - r\cos t\cos(2t) & -\sin t\cos(2t) \\ 2\cos(2t)(2 - r\sin t) - r\cos t\sin(2t) & -\sin t\sin(2t) \\ -r\sin t & \cos t \end{pmatrix} \\
&= 2.
\end{aligned}$$

Also ist $\mathcal{M}$ eine 2-dimensionale Mfgkt der Klasse $C^\infty$ in $\mathbb{R}^3$.

Der folgende Satz beschreibt das Verhalten bei Kartenwechseln und ist für spätere Anwendungen wichtig.

Satz X.1.10. *Seien $M \subset \mathbb{R}^n$ eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$ und $(U_1, \varphi_1, V_1)$, $(U_2, \varphi_2, V_2)$ zwei Karten von $M$ mit $V = V_1 \cap V_2 \neq \emptyset$. Dann sind $W_1 = \varphi_1^{-1}(V) \subset U_1$ und $W_2 = \varphi_2^{-1}(V) \subset U_2$ offen, und $\tau = \varphi_2^{-1} \circ \varphi_1 : W_1 \to W_2$ ist ein $C^\alpha$-Diffeomorphismus.*

Beweis. Da $V$ in $V_j$, $j = 1,2$, offen und $\varphi_j$ stetig ist, ist $W_j$ in $U_j$, $j = 1,2$, offen. Konstruktionsgemäß ist $\tau$ ein Homöomorphismus. Sei $x_1^* \in W_1$ beliebig und

$$y^* = \varphi_1(x_1^*) \quad , \quad x_2^* = \varphi_2^{-1}(y^*) = \tau(x_1^*) \in W_2.$$

Nach Satz X.1.4 existieren eine Umgebung $\widetilde{U}$ von $y^*$ in $\mathbb{R}^n$, eine offene Menge $\widetilde{V}$ in $\mathbb{R}^n$ und ein $C^\alpha$-Diffeomorphismus $F : \widetilde{U} \to \widetilde{V}$ mit

$$F(M \cap \widetilde{U}) = \widetilde{V} \cap E_k.$$

O.E. können wir $\widetilde{U} \subset V$ annehmen. Sei $\widetilde{W}_j = \varphi_j^{-1}(M \cap \widetilde{U})$, $j = 1,2$. Auf $\widetilde{W}_1$ bzw. $\widetilde{W}_2$ gilt

$$F \circ \varphi_1 = (g_1, \ldots, g_k, 0, \ldots, 0)$$

bzw.

$$F \circ \varphi_2 = (h_1, \ldots, h_k, 0, \ldots, 0).$$

Da $\operatorname{Rang}(D\varphi_j) = k$ und $DF$ invertierbar ist, folgt

$$\operatorname{Rang}(D(F \circ \varphi_j)) = k \quad j = 1,2,$$

so dass

$$\begin{aligned}
g &= (g_1, \ldots, g_k) : \widetilde{W}_1 \to E_k \cap \widetilde{V} \\
h &= (h_1, \ldots, h_k) : \widetilde{W}_2 \to E_k \cap \widetilde{V}
\end{aligned}$$

$C^\alpha$-Diffeomorphismen sind. (Dabei betrachten wir $E_k \cap \widetilde{V}$ als offene Teilmenge des $\mathbb{R}^k$.) Auf $\widetilde{W}_1$ gilt aber

$$\tau = \varphi_2^{-1} \circ \varphi_1 = (F \circ \varphi_2)^{-1} \circ (F \circ \varphi_1) = h^{-1} \circ g.$$

Also ist $\tau : \widetilde{W}_1 \to \widetilde{W}_2$ ein $C^\alpha$-Diffeomorphismus. Da $x_1^* \in W_1$ beliebig war, folgt hieraus die Behauptung. □

## X.2. Tangentialraum und Orientierung

Seien $a, b \in \mathbb{R}$ mit $a < b$ und $\varphi \in C^1((a,b), \mathbb{R}^n)$ ein regulärer Weg. Dann ist $\mathrm{Spur}(\varphi)$ eine 1-dimensionale Mannigfaltigkeit in $\mathbb{R}^n$. Durch $\varphi$ werden auf kanonische Weise Tangentenvektoren und eine Orientierung für diese Mfgkt definiert. Wir wollen in diesem Abschnitt diese Begriffe auf allgemeine Mfgkten übertragen.

Sofern nicht anders bemerkt, ist im Folgenden stets $1 \le k < n$ und $\alpha \in \mathbb{N}^*$. $(\cdot, \cdot)$ bezeichnet das euklidische Skalarprodukt auf dem $\mathbb{R}^n$

$$(u, v) = \sum_{i=1}^{n} u_i v_i \quad \forall u, v \in \mathbb{R}^n.$$

Definition X.2.1. Seien $M \subset \mathbb{R}^n$ eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$ und $x_0 \in M$.
(1) $v \in \mathbb{R}^n$ heißt Tangentialvektor an $M$ im Punkte $x_0$, wenn es ein $\varepsilon \in \mathbb{R}_+^*$ und ein $\gamma \in C^1((-\varepsilon, \varepsilon), M)$ gibt mit $\gamma(0) = x_0$ und $\gamma'(0) = v$.
(2) $T_{x_0}M = \{v \in \mathbb{R}^n : v \text{ ist Tangentialvektor an } M \text{ im Punkte } x_0\}$ heißt der Tangentialraum von $M$ im Punkte $x_0$.
(3) $u \in \mathbb{R}^n$ heißt Normalenvektor an $M$ im Punkte $x_0$, wenn gilt

$$(u, v) = 0 \quad \forall v \in T_{x_0}M.$$

(4) $N_{x_0}M = \{u \in \mathbb{R}^n : u \text{ ist Normalenvektor zu } M \text{ im Punkte } x_0\}$ heißt der Normalraum von $M$ im Punkte $x_0$.

Satz X.2.2. *Seien $M \subset \mathbb{R}^n$ eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$ und $x_0 \in M$. Dann gilt:*

1. *$T_{x_0}M$ ist ein $k$-dimensionaler Untervektorraum des $\mathbb{R}^n$.*
2. *$N_{x_0}M$ ist ein $n-k$-dimensionaler Untervektorraum des $\mathbb{R}^n$.*
3. *Sei $(U, \varphi, V)$ eine Karte von $M$ mit $x_0 \in V$ und $y_0 = \varphi^{-1}(x_0)$. Dann ist*
$$D_1\varphi(y_0), \dots, D_k\varphi(y_0)$$
*eine Basis von $T_{x_0}M$.*
4. *Sei $U$ eine Umgebung von $x_0$ in $\mathbb{R}^n$ und $f \in C^1(U, \mathbb{R}^{n-k})$ derart, dass $x_0$ ein regulärer Punkt von $f$ und $M \cap U = f^{-1}(\{0\})$ ist. Dann ist*
$$T_{x_0}M = \{v \in \mathbb{R}^n : (\nabla f_j(x_0), v) = 0 \; \forall 1 \le j \le n-k\}.$$

Beweis. Seien

$$\begin{aligned}T_1 &= \operatorname{span}\{D_1\varphi(y_0), \dots, D_k\varphi(y_0)\}\\ T_2 &= \{v \in \mathbb{R}^n : (\nabla f_j(x_0), v) = 0\ \forall 1 \le j \le n-k\}.\end{aligned}$$

Dann sind $T_1$ und $T_2$ $k$-dimensionale Untervektorräume des $\mathbb{R}^n$. Wir zeigen: $T_1 \subset T_{x_0}M \subset T_2$. Hieraus folgen dann sofort die Behauptungen (1), (3) und (4). Wegen (4) ist dann

$$N_{x_0}M = \operatorname{span}\{\nabla f_1(x_0)^t, \dots, \nabla f_{n-k}(x_0)^t\},$$

so dass auch die Behauptung (2) folgt.
„$T_1 \subset T_{x_0}M$:“ Sei

$$v = \sum_{i=1}^{n} \alpha_i D_i\varphi(y_0) \in T_1$$

beliebig. Dann gibt es ein $\varepsilon \in \mathbb{R}_+^*$, so dass gilt

$$y_0 + t\sum_{i=1}^{k} \alpha_i e_i \in U \quad \forall t \in (-\varepsilon, \varepsilon),$$

wobei $e_i$ der $i$-te Einheitsvektor in $\mathbb{R}^k$ ist. Definiere $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ durch

$$\gamma(t) = \varphi(y_0 + t\sum_{i=1}^{k} \alpha_i e_i).$$

Dann ist $\gamma \in C^1((-\varepsilon, \varepsilon), M)$ und

$$\begin{aligned}\gamma(0) &= \varphi(y_0) = x_0\\ \gamma'(0) &= \sum_{i=1}^{k} \alpha_i D_i\varphi(y_0) = v.\end{aligned}$$

Also ist $v \in T_{x_0}M$.
„$T_{x_0}M \subset T_2$:“ Sei $v \in T_{x_0}M$, d.h., $v = \psi'(0)$ für einen $C^1$-Weg $\psi \in C^1((-\varepsilon, \varepsilon), M)$ mit $\psi(0) = x_0$. Dann gilt

$$f_j(\psi(t)) = 0 \quad \forall 1 \le j \le n-k, t \in (-\varepsilon, \varepsilon)$$

und somit

$$\begin{aligned}0 &= \frac{d}{dt} f_j(\psi(t))|_{t=0}\\ &= (\nabla f_j(\psi(0)), \psi'(0))\\ &= (\nabla f_j(x_0), v) \quad \forall 1 \le j \le n-k.\end{aligned}$$

Also ist $v \in T_2$. □

Wir können uns den Tangentialraum $T_{x_0}M$ im Punkte $x_0 \in M$ „angeheftet“ denken. Variiert $x_0$, so auch $T_{x_0}M$. Dies führt zu folgender Definition.

Definition X.2.3. Sei $M$ eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$. Dann heißt

$$TM = \{(x,v) \in \mathbb{R}^{2n} : x \in M, v \in T_x M\}$$

der Tangentialraum von $M$.

Beispiel X.2.4. (1) Es ist $S^{n-1} = f^{-1}(\{0\})$ mit

$$f(x) = \sum_{i=1}^{n} x_i^2 - 1.$$

Also ist

$$N_x S^{n-1} = \operatorname{span}\{(x_1, \ldots, x_n)^t\}$$

und

$$T_x S^{n-1} = \{v \in \mathbb{R}^n : (x,v) = 0\}.$$

Sei speziell $n = 3$. Wir wählen als Karte eine Polarkoordinatendarstellung

$$x = \psi(\varphi, \theta) = (\cos\varphi\cos\theta, \sin\varphi\cos\theta, \sin\theta).$$

Dann ist

$$\begin{aligned} T_x S^2 = \operatorname{span}\{&(-\sin\varphi\cos\theta, \cos\varphi\cos\theta, 0)^t\ , \\ &(-\cos\varphi\sin\theta, -\sin\varphi\sin\theta, \cos\theta)^t\}. \end{aligned}$$

(2) Es sit $\mathcal{T} = f^{-1}(\{0\})$ mit

$$f(x) = (\sqrt{x_1^2 + x_2^2} - R)^2 + x_3^2 - r^2.$$

Also ist

$$T_x\mathcal{T} = \{v \in \mathbb{R}^3 : (\sqrt{x_1^2 + x_2^2} - R)\frac{x_1 v_1 + x_2 v_2}{\sqrt{x_1^2 + x_2^2}} + x_3 v_3 = 0\}.$$

Als Karte kann man eine „Polarkoordinatendarstellung“ wählen

$$x = \psi(\varphi, \theta) = ((R + r\cos\theta)\cos\varphi, (R + r\cos\theta)\sin\varphi, r\sin\theta)^t.$$

Dann ist

$$\begin{aligned} T_x\mathcal{T} = \operatorname{span}\{&(-(R + r\cos\theta)\sin\varphi, (R + r\cos\theta)\cos\varphi, 0)^t, \\ &(-r\sin\theta\cos\varphi, -r\sin\theta\sin\varphi, r\cos\theta)^t\}. \end{aligned}$$

Wir kommen nun zum Begriff der Orientierbarkeit. Dazu benötigen wir:

Definition X.2.5. Seien $U, V \subset \mathbb{R}^n$ offen und $\varphi : U \to V$ ein $C^1$-Diffeomorphismus von $U$ auf $V$. Dann heißt $\varphi$ orientierungstreu bzw. orientierungsumkehrend, wenn gilt

$$\det(D\varphi(x)) > 0 \quad \forall x \in U$$

bzw.

$$\det(D\varphi(x)) < 0 \quad \forall x \in U.$$

Bemerkung X.2.6. (1) Ist $U$ zusammenhängend und $\varphi : U \to V$ ein $C^1$-Diffeomorphismus, so ist $\varphi$ entweder orientierungstreu oder orientierungsumkehrend.
(2) Ist $n = 1$, so ist ein $C^1$-Diffeomorphismus $\varphi$ genau dann orientierungstreu bzw. orientierungsumkehrend, wenn $\varphi$ streng monoton wachsend bzw. streng monoton fallend ist.

Definition X.2.7. $M \subset \mathbb{R}^n$ sei eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$.

(1) Zwei Karten $(U_i, \varphi_i, V_i)$, $i = 1, 2$, von $M$ mit $V_1 \cap V_2 \neq \emptyset$ heißen gleichorientiert, wenn der $C^\alpha$-Diffeomorphismus $\tau = \varphi_2^{-1} \circ \varphi_1 : W_1 = \varphi_1^{-1}(V_1 \cap V_2) \to W_2 = \varphi_2^{-1}(V_1 \cap V_2)$ orientierungstreu ist.
(2) Ein Atlas $\mathcal{A}$ von $M$ heißt orientiert, wenn je zwei Karten von $\mathcal{A}$ gleichorientiert sind.
(3) Eine Orientierung $\mathcal{O}$ von $M$ wird durch einen orientierten Atlas von $M$ gegeben.
(4) $M$ heißt orientierbar, wenn ein orientierter Atlas von $M$ existiert.
(5) Seien $(M, \mathcal{O})$ eine orientierbare Mannigfaltigkeit mit einer zugehörigen Orientierung $\mathcal{O}$ und $(U, \varphi, V)$ eine Karte von $M$. Die Karte heißt positiv orientiert bzgl. $\mathcal{O}$, wenn sie mit jeder Karte des zu $\mathcal{O}$ gehörenden Atlasses gleichorientiert ist.

Bemerkung X.2.8. (1) Genau genommen ist eine Orientierung $\mathcal{O}$ einer Mfgkt $M$ eine Äquivalenzklasse orientierter Atlanten von $M$. Dabei sind zwei orientierte Atlanten $\mathcal{A}, \mathcal{A}'$ von $M$ äquivalent, wenn jede Karte von $\mathcal{A}$ zu jeder Karte von $\mathcal{A}'$ gleichorientiert ist.
(2) Sei $(M, \mathcal{O})$ eine orientierbare Mannigfaltigkeit und $\mathcal{A} = \{(U_\lambda, \varphi_\lambda, V_\lambda) : \lambda \in \Lambda\}$ ein zu $\mathcal{O}$ gehöriger Atlas. Definiere $i : \mathbb{R}^k \to \mathbb{R}^k$ durch

$$i(x_1, \ldots, x_k) = (x_1, \ldots, x_{k-1}, -x_k).$$

Dann ist $\widetilde{\mathcal{A}} = \{(i(U_\lambda), \varphi_\lambda \circ i^{-1}, V_\lambda) : \lambda \in \Lambda\}$ ein orientierter Atlas von $M$, dessen Orientierung von $\mathcal{O}$ verschieden ist. Die Orientierung von $\widetilde{\mathcal{A}}$ wird mit $-\mathcal{O}$ bezeichnet und heißt die zu $\mathcal{O}$ entgegengesetzte Orientierung.
(3) Ist $(M, \mathcal{O})$ eine zusammenhängende, orientierbare Mfgkt, kann man zeigen, dass auf $M$ nur die Orientierungen $\mathcal{O}$ und $-\mathcal{O}$ existieren.

Beispiel X.2.9. Definiere $\varphi, \psi : \mathbb{R} \to \mathbb{R}^2$ durch

$$\varphi(t) = (\cos t, \sin t)$$
$$\psi(t) = (\sin t, \cos t).$$

Dann sind

$$\mathcal{A} = \Big\{ \Big( (0, \frac{3}{2}\pi), \varphi|_{(0,\frac{3}{2}\pi)}, \varphi((0, \frac{3}{2}\pi)) \Big),$$
$$\Big( (\pi, \frac{5}{2}\pi), \varphi|_{(\pi,\frac{5}{2}\pi)}, \varphi((\pi, \frac{5}{2}\pi)) \Big) \Big\}$$

und

$$\mathcal{A}' = \Big\{ \Big( (0, \frac{3}{2}\pi), \psi|_{(0,\frac{3}{2}\pi)}, \psi((0, \frac{3}{2}\pi)) \Big),$$
$$\Big( (\pi, \frac{5}{2}\pi), \psi|_{(\pi,\frac{5}{2}\pi)}, \psi((\pi, \frac{5}{2}\pi)) \Big) \Big\}$$

Zwei entgegengesetzt orientierte Atlanten von $S^1$.

Eine Orientierung von $M$ induziert auf kanonische Weise für jedes $x \in M$ eine Orientierung von $T_xM$.

Definition X.2.10. Seien $(M, \mathcal{O})$ eine orientierbare Mfgkt und $x_0 \in M$. Eine Basis $\{v_1, \ldots, v_k\}$ von $T_{x_0}M$ heißt positiv orientiert bzgl. $\mathcal{O}$, wenn für jede bzgl. $\mathcal{O}$ positiv orientierte Karte $(U, \varphi, V)$ mit $x_0 \in V$ gilt

$$\det(A) > 0,$$

wobei die Matrix $A$ definiert ist durch

$$v_i = \sum_{j=1}^{k} A_{ij} D_j\varphi(y_0), \quad y_0 = \varphi^{-1}(x_0), \quad 1 \le i \le k.$$

Ist $M \subset \mathbb{R}^n$ eine Hyperfläche, so ist für jedes $x \in M$ der Normalraum $N_xM$ eindimensional, enthält also genau zwei Einheitsvektoren. Wir wollen zeigen, dass eine Hyperfläche $M$ genau dann orientierbar ist, wenn von diesen Vektoren einer ausgewählt werden kann, so dass die entstehende Abbildung $M \to \mathbb{R}^n$ stetig ist. Dies wird präzisiert durch:

Definition X.2.11. Sei $M \subset \mathbb{R}^n$ eine $C^1$-Hyperfläche.

(1) Eine Abbildung $\nu \in C(M, \mathbb{R}^n)$ mit den Eigenschaften
$$\nu(x) \in N_xM,$$
$$\|\nu(x)\|_2 = 1$$
für alle $x \in M$ heißt Einheitsnormalenfeld auf $M$.

(2) Ist $M$ orientierbar und $\nu$ ein Einheitsnormalenfeld auf $M$, so heißt $\nu$ bzgl. der Orientierung $\mathcal{O}$ positiv orientiert, wenn für jedes $x \in M$ und jede bzgl. $\mathcal{O}$ positiv orientierte Basis $\{v_1, \ldots, v_{n-1}\}$ von $T_xM$ gilt
$$\det(\nu, v_1, \ldots, v_{n-1}) > 0.$$

Bemerkung X.2.12. Wegen Definition X.2.7 hängen die Definitionen X.2.10 und X.2.11 (2) nicht von der Wahl der Karte bzw. Basis von $T_x M$ ab.

Satz X.2.13. *Eine $C^1$-Hyperfläche $M \subset \mathbb{R}^n$ ist genau dann orientierbar, wenn auf $M$ ein Einheitsnormalenfeld $\nu$ existiert.*

Beweis. „$\Longrightarrow$“: Sei $x \in M$ und $\{v_1, \ldots, v_{n-1}\}$ eine positiv orientierte Basis von $T_x M$. Da $N_x M$ eindimensional ist, gibt es genau ein $\nu(x) \in \mathbb{R}^n$ mit

$$\begin{aligned} \|\nu(x)\|_2 &= 1, \\ \nu(x) &\in N_x M, \\ \det(\nu(x), v_1, \ldots, v_{n-1}) &> 0. \end{aligned}$$

Wir müssen zeigen, dass die so konstruierte Funktion $\nu : M \to \mathbb{R}^n$ stetig ist. Sei dazu $x_0 \in M$ beliebig und $U \in \mathcal{U}(x_0)$, $f \in C^1(U, \mathbb{R})$ wie in Definition X.1.1 (S. 359), d.h., $M \cap U = f^{-1}(\{0\})$ und $x_0$ ist regulärer Punkt von $f$. Indem wir $U$ gegebenenfalls verkleinern, wird durch

$$\widetilde{\nu}(x) = \frac{1}{\|\nabla f(x)\|_2} \nabla f(x) \quad \forall x \in U$$

eine stetige Funktion $\widetilde{\nu} : U \to \mathbb{R}^n$ definiert mit

$$\begin{aligned} \|\widetilde{\nu}(x)\|_2 &= 1, \\ \widetilde{\nu}(x) &\in N_x M \end{aligned}$$

für alle $x \in U$. Indem wir nötigenfalls zu $-f$ übergehen, können wir $\widetilde{\nu}(x_0) = \nu(x_0)$ annehmen. Sei nun $(\widetilde{U}, \varphi, \widetilde{V})$ eine positiv orientierte Karte mit $x_0 \in \widetilde{V}$ und $y_0 = \varphi^{-1}(x_0)$, $\widetilde{W} = \varphi^{-1}(\widetilde{V} \cap U)$. Indem wir gegebenenfalls $\widetilde{U}$ verkleinern, können wir annehmen, dass $\widetilde{W}$ zusammenhängend ist. Durch

$$y \mapsto \Delta(y) = \det\Big(\widetilde{\nu}(\varphi(y)), D_1\varphi(y), \ldots, D_{n-1}\varphi(y)\Big)$$

wird eine stetige Funktion $\Delta : \widetilde{W} \to \mathbb{R}$ mit $\Delta(y_0) > 0$ definiert. Wegen Satz X.2.2 (4), Definition X.1.7 (S. 362) und Definition X.1.5 (S. 362) gilt $\Delta(y) \neq 0$ für alle $y \in \widetilde{W}$. Da $\widetilde{W}$ zusammenhängend ist, gilt $\Delta(y) > 0$ für alle $y \in \widetilde{W}$. Hieraus und der Definition von $\nu$ folgt

$$\widetilde{\nu}(\varphi(y)) = \nu(\varphi(y)) \quad \forall y \in \widetilde{W}.$$

Also ist $\nu$ in $x_0$ stetig.
„$\Longleftarrow$“: Sei nun $\nu$ ein Einheitsnormalenfeld auf $M$. Sei $x \in M$ und $(U, \varphi, V)$ eine Karte von $M$ mit $x \in V$. Indem wir $U$ nötigenfalls verkleinern, können wir annehmen, dass $U$ zusammenhängend ist. Dann ist

$$\Delta(y) = \det\Big(\nu(\varphi(y)), D_1\varphi(y), \ldots, D_{n-1}\varphi(y)\Big)$$

eine stetige Funktion auf $U$, die nirgends verschwindet. Indem wir nötigenfalls die Transformation $(y_1, \dots, y_{n-1}) \mapsto (y_1, \dots, -y_{n-1})$ vorschalten, können wir

$$\Delta(y) > 0, \quad \forall y \in U$$

annehmen. Sei $\mathcal{A}$ der so konstruierte Atlas von $M$. Wir müssen zeigen, dass $\mathcal{A}$ orientiert ist. Seien dazu $(U_i, \varphi_i, V_i)$, $i = 1, 2$, zwei Karten von $\mathcal{A}$ mit $V = V_1 \cap V_2 \neq \emptyset$. Sei $W_i = \varphi_i^{-1}(V), i = 1, 2$, und $\tau = \varphi_2^{-1} \circ \varphi_1$, d.h.

$$\varphi_1 = \varphi_2 \circ \tau.$$

Für $y \in W_1$ gilt dann

$$D\varphi_1(y) = D\varphi_2(\tau(y)) \cdot D\tau(y)$$

und

$$\begin{aligned} 0 &< \det\Big(\nu(\varphi_1(y)), D_1\varphi_1(y), \dots, D_{n-1}\varphi_1(y)\Big) \\ 0 &< \det\Big(\nu(\varphi_2(\tau(y))), D_1\varphi_2(\tau(y)), \dots, D_{n-1}\varphi_2(\tau(y))\Big). \end{aligned}$$

Also ist

$$\det D\tau(y) > 0,$$

d.h., die Karten $(U_1, \varphi_1, V_1)$ und $(U_2, \varphi_2, V_2)$ sind gleich orientiert. □

Beispiel X.2.14. (1) $\nu(x) = x$ ist ein Einheitsnormalenfeld auf $S^{n-1}$.
(2) Die Funktion

$$\nu(x) = \frac{1}{r}\Big(\big(\sqrt{x_1^2 + x_2^2} - R\big)\frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \big(\sqrt{x_1^2 + x_2^2} - R\big)\frac{x_2}{\sqrt{x_1^2 + x_2^2}}, x_3\Big)^t$$

ist ein Einheitsnormalenfeld auf dem Torus $\mathcal{T}$.
(3) Das Möbiusband ist nicht orientierbar. Denn: Sei $\varphi$ wie in Beispiel X.1.9(2) (S. 362). Eine leichte Rechnung liefert

$$\begin{aligned} T_{\varphi(t,0)}\mathcal{M} = \operatorname{span}\{&(-\sin(2t), \cos(2t), 0)^t, \\ &(-\sin t\cos(2t), -\sin t\sin(2t), \cos t)^t\}. \end{aligned}$$

Also ist

$$N_{\varphi(t,0)}M = \operatorname{span}\{(\cos t\cos(2t), \cos t\sin(2t), \sin t)^t\}.$$

Setze zur Abkürzung

$$\mu(t) = (\cos t\cos(2t), \cos t\sin(2t), \sin t)^t.$$

Gäbe es ein Einheitsnormalenfeld $\nu$ auf $M$, so auch eines mit

$$\nu(\varphi(0,0)) = \mu(0) = e_1.$$

Stetige Fortsetzung längs $\varphi(t, 0)$, $-\frac{\pi}{2} \le t \le \frac{\pi}{2}$, liefert dann

$$\nu(\varphi(t,0)) = \mu(t) \quad \forall -\frac{\pi}{2} \le t \le \frac{\pi}{2}.$$

Hieraus folgt aber

$$-e_3 = \mu(-\frac{\pi}{2}) = \nu(-2e_1) = \mu(\frac{\pi}{2}) = e_3.$$

Dies ist ein Widerspruch.

## X.3. Integration auf Mannigfaltigkeiten

Sofern nicht anders vermerkt, ist im Folgenden stets $1 \le k < n$ und $\alpha \in \mathbb{N}^*$. $(\cdot,\cdot)$ bezeichnet das euklidische Skalarprodukt auf $\mathbb{R}^n$. Funktionen, die auf Teilmengen des $\mathbb{R}^n$ definiert sind, werden durch 0 auf ganz $\mathbb{R}^n$ fortgesetzt. Wir wollen Funktionen auf $k$-dimensionalen Mfgkten im $\mathbb{R}^n$ integrieren. Insbesondere sollte das Integral der Funktion $f = 1$ das $k$-dimensionale Volumen der Mfgkt liefern. Außerdem sollte das Integral über die Mfgkt mit Hilfe von Karten auf ein Lebesgue-Integral in $\mathbb{R}^k$ zurückgeführt werden.

Zur Motivation betrachten wir folgende Situation. Sei $U \subset \mathbb{R}^k$ offen und $\varphi : U \overset{\alpha}{\hookrightarrow} M = \varphi(U) \subset \mathbb{R}^n$ eine Immersion. Sei $x_0 \in U$ und

$$v_0 = \varphi(x_0) \quad , \quad v_i = D_i\varphi(x_0) \quad 1 \le i \le k.$$

Dann wird für $\varepsilon > 0$ die Mfgkt $M$ in der Nähe von $v_0$ durch die „Tangentialebene“

$$E = \left\{ y = v_0 + \sum_{i=1}^{k} t_i v_i : |t_i| \le \frac{\varepsilon}{2}, 1 \le i \le k \right\}$$

approximiert. Seien $v_{k+1}, \ldots, v_n \in \mathbb{R}^n$ paarweise orthogonale Einheitsvektoren mit

$$(v_i, v_j) = 0 \quad \forall 1 \le i \le k < j \le n.$$

Sei

$$P = \left\{ y = v_0 + \sum_{i=1}^{n} t_i v_i : |t_i| \le \frac{\varepsilon}{2},\ 1 \le i \le k,\ 0 \le t_j \le 1,\ k+1 \le j \le n \right\}$$

das Parallelepiped der Höhe 1 über $E$. Dann ist $\lambda_n(P)$ ein Maß für das $k$-dimensionale Volumen von $E$. Aus Satz IX.5.6 (S. 343) folgt

$$\begin{aligned}
\lambda_n(P) &= \varepsilon^k |\det(v_1, \ldots, v_n)| \\
&= \varepsilon^k \sqrt{\det(v_1, \ldots, v_n)^2} \\
&= \varepsilon^k \sqrt{\det((v_1, \ldots, v_n)^t (v_1, \ldots, v_n))} \\
&= \varepsilon^k \sqrt{\det((v_i, v_j)_{1 \le i,j \le k})}.
\end{aligned}$$

Also ist

$$g_\varphi(x_0) = \det((v_i, v_j)_{1 \le i,j \le k})$$

ein guter Kandidat für das Volumenelement von $M$ und $\int_U \sqrt{g_\varphi(x)}dx$ ein Kandidat für das $k$-dimensionale Volumen von $M$. Diese Ideen werden wir nun präzisieren.

Definition X.3.1. Seien $U \subset \mathbb{R}^k$ offen und $\varphi \in C^\alpha(U,\mathbb{R}^n)$ eine Immersion. Dann heißen

$$G_\varphi(x) = (\ (D_i\varphi(x), D_j\varphi(x))\ )_{1\le i,j\le k} \in C^{\alpha-1}(U,\mathbb{R}^{k\times k})$$

und

$$g_\varphi(x) = \det G_\varphi(x) \in C^{\alpha-1}(U,\mathbb{R})$$

der Masstensor und die Gramsche Determinante von $\varphi$.

Bemerkung X.3.2. Wegen $(\ (D_i\varphi(x), D_j\varphi(x))\ )_{1\le i,j\le k} = D\varphi(x)^t \cdot D\varphi(x)$ ist $G_\varphi(x)$ positiv semi-definit. Da $D\varphi(x)$ den Rang $k$ hat, ist $G_\varphi(x)$ sogar positiv definit und damit $g_\varphi(x) > 0$ für alle $x \in U$.

Satz X.3.3. (1) *Seien $U, V \subset \mathbb{R}^k$ offen, $\varphi \in C^\alpha(U,\mathbb{R}^n)$ eine Immersion, $\tau : V \to U$ ein $C^\alpha$-Diffeomorphismus und $\psi = \varphi\circ\tau \in C^\alpha(V,\mathbb{R}^n)$. Dann ist*

$$g_\psi(x) = g_\varphi(\tau(x))|\det D\tau(x)|^2 \quad \forall x \in V.$$

(2) *Sei $M \subset \mathbb{R}^n$ eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$, $f : M \to \mathbb{R}$ eine Funktion mit kompaktem Träger, $\mathcal{A} = \{(U_i,\varphi_i,V_i) : 1 \le i \le m\}$ und $\mathcal{A}' = \{(U'_j,\varphi'_j,V'_j) : 1 \le j \le m'\}$ zwei Atlanten von $\operatorname{supp}(f)\cap M$ und $\{\alpha_i : 1 \le i \le m\}$ und $\{\alpha'_j : 1 \le j \le m'\}$ zwei Partitionen der Eins auf $\operatorname{supp}(f)\cap M$, die $\mathcal{A}$ bzw. $\mathcal{A}'$ untergeordnet sind. Dann gilt*

$$\alpha_i\circ\varphi_i \cdot f\circ\varphi_i \cdot \sqrt{g_{\varphi_i}} \in L^1(U_i,\mathbb{R}) \quad \forall 1 \le i \le m$$
$$\iff \alpha'_j\circ\varphi'_j \cdot f\circ\varphi'_j \cdot \sqrt{g_{\varphi'_j}} \in L^1(U'_j,\mathbb{R}) \quad \forall 1 \le j \le m'.$$

*In diesem Fall ist*

$$\int_{\mathbb{R}^k}\sum_{i=1}^m \alpha_i\circ\varphi_i \cdot f\circ\varphi_i \cdot \sqrt{g_{\varphi_i}}$$
$$= \int_{\mathbb{R}^k}\sum_{j=1}^{m'} \alpha'_j\circ\varphi'_j \cdot f\circ\varphi'_j \cdot \sqrt{g_{\varphi'_j}}.$$

Beweis. ad (1): Aus der Kettenregel, Satz VII.3.3 (S. 226), folgt

$$\begin{aligned} g_\psi(x) &= \det(D\psi(x)^t D\psi(x)) \\ &= \det([D\varphi(\tau(x))\cdot D\tau(x)]^t[D\varphi(\tau(x))\cdot D\tau(x)]) \\ &= (\det D\tau(x))^2 \det(D\varphi(\tau(x))^t D\varphi(\tau(x))) \\ &= |\det D\tau(x)|^2 g_\varphi(\tau(x)) \quad \forall x \in V. \end{aligned}$$

ad (2): Seien $1 \le i \le m$ und $1 \le j \le m'$ so, dass $V = V_i \cap V'_j \neq \emptyset$ ist. Sei $W_i = \varphi_i^{-1}(V), W'_j = {\varphi'}_j^{-1}(V)$ und

$$\tau_{ij} = \varphi_i^{-1}\circ\varphi'_j.$$

Aus Teil (1) und Satz IX.5.6 (S. 343) folgt

$$\alpha_i\circ\varphi_i \cdot f\circ\varphi_i \cdot \sqrt{g_{\varphi_i}} \in L^1(U_i,\mathbb{R})$$

$$\begin{aligned}&\Longrightarrow \alpha_i\circ\varphi_i\cdot f\circ\varphi_i\cdot\alpha'_j\circ\varphi_i\cdot\sqrt{g_{\varphi_i}}\in L^1(W_i,\mathbb{R})\\ &\Longrightarrow \alpha_i\circ\varphi_i\circ\tau_{ij}\cdot f\circ\varphi_i\circ\tau_{ij}\cdot\alpha'_j\circ\varphi_i\circ\tau_{ij}\cdot\\ &\qquad\cdot\sqrt{g_{\varphi_i}\circ\tau_{ij}}\cdot|\det D\tau_{ij}|\in L^1(W_j,\mathbb{R})\\ &\Longrightarrow \alpha_i\circ\varphi'_j\cdot\alpha'_j\circ\varphi'_j\cdot f\circ\varphi'_j\cdot\sqrt{g_{\varphi'_j}}\in L^1(W_j,\mathbb{R}).\end{aligned}$$

Summation über $i$ liefert

$$\alpha'_j\circ\varphi'_j\cdot f\circ\varphi'_j\cdot\sqrt{g_{\varphi'_j}}\in L^1(U'_j,\mathbb{R}),$$

da $\sum\limits_{i=1}^{m}\alpha_i\circ\varphi'_j = 1$ auf $U'_j$ ist. Durch Vertauschen der Rollen von $\mathcal{A}$, $\{\alpha_i\}$ und $\mathcal{A}'$, $\{\alpha'_j\}$ folgt die behauptete Äquivalenz und die Gleichheit der Integrale, da

$$\sum_{i=1}^{m}(\alpha_i\circ\varphi'_j\cdot\alpha_j\circ\varphi'_j)=\alpha_j\circ\varphi'_j\quad\forall 1\le j\le m'$$

und

$$\sum_{j=1}^{m'}(\alpha_j\circ\varphi_i\cdot\alpha_i\circ\varphi_i)=\alpha_i\circ\varphi_i\quad\forall 1\le i\le m$$

ist. □

Wegen Satz X.3.3 ist folgende Definition sinnvoll.

Definition X.3.4. Sei $M\subset\mathbb{R}^n$ eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$.

(1) Eine Funktion $f: M\to\mathbb{R}$ mit kompaktem Träger heißt integrierbar auf $M$, wenn es einen Atlas $\{(U_i,\varphi_i,V_i): 1\le i\le m\}$ von $\operatorname{supp}(f)\cap M$ und eine dem Atlas untergeordnete Zerlegung der Eins $\{\alpha_i: 1\le i\le m\}$ auf $\operatorname{supp}(f)\cap M$ gibt mit

$$\alpha_i\circ\varphi_i\cdot f\circ\varphi_i\cdot\sqrt{g_{\varphi_i}}\in L^1(U_i,\mathbb{R})\quad\forall 1\le i\le m.$$

In diesem Fall heißt

$$\begin{aligned}\int_M f=\int_M f dS&=\int_M f(x)dS(x)\\ &=\sum_{i=1}^{m}\int_{\mathbb{R}^k}\alpha_i\circ\varphi_i\cdot f\circ\varphi_i\cdot\sqrt{g_{\varphi_i}}\end{aligned}$$

das Integral von $f$ auf $M$. Die Menge aller auf $M$ integrierbaren Funktionen wird mit $L^1(M,\mathbb{R})$ bezeichnet.

(2) Eine beschränkte Teilmenge $A\subset M$ heißt integrierbar, wenn $\chi_A\in L^1(M,\mathbb{R})$ ist. In diesem Fall heißt

$$\sigma_k(A)=\int_M\chi_A$$

das $k$-DIMENSIONALE VOLUMEN von $A$. Ist $\sigma_k(A) = 0$, so heißt $A$ eine $k$-DIMENSIONALE NULLMENGE in $M$.

BEMERKUNG X.3.5. (1) Die Einschränkung „supp($f$) kompakt" ist rein technischer Natur und kann wesentlich abgeschwächt werden. Insbesondere ist sie hinfällig, wenn $M$ bis auf eine $k$-dimensionale Nullmenge durch eine Karte überdeckt werden kann.
(2) Ist $A \subset M$ integrierbar und als Teilmenge des $\mathbb{R}^n$ messbar, so muss man deutlich zwischen $\sigma_k(A)$ und $\lambda_n(A)$ unterscheiden. In der Regel ist

$$\lambda_n(A) = 0 \quad \text{aber} \quad \sigma_k(A) > 0.$$

(3) Ist $I \subset \mathbb{R}$ offen und beschränkt und $\varphi \in C^1(I, \mathbb{R}^n)$ ein regulärer Weg, so ist

$$g_\varphi = \|\varphi'\|_2^2,$$

und die Länge des Weges stimmt mit dem eindimensionalen Volumen von Spur($\varphi$) überein.

Man kann die Ergebnisse aus Kapitel IX direkt auf das Integral auf Mfgkten übertragen. Wir geben hier einen Teil an.

SATZ X.3.6. *Sei $M \subset \mathbb{R}^n$ eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$.*

(1) $L^1(M, \mathbb{R})$ *ist ein Vektorraum. Sind $f, g \in L^1$ und $\alpha, \beta \in \mathbb{R}$, so gilt*
   (a) $\int_M (\alpha f + \beta g) = \alpha \int_M f + \beta \int_M g$,
   (b) $f \geq 0 \Longrightarrow \int_M f \geq 0$,
   (c) $|f| \in L^1(M, \mathbb{R})$ *und* $|\int_M f| \leq \int_M |f|$.
(2) *Seien $(f_l)_{l \in \mathbb{N}} \subset L^1(M, \mathbb{R})$ und $f : M \to \mathbb{R}$ eine Funktion mit $f_l \underset{l \to \infty}{\longrightarrow} f$ f.ü. auf $M$ (f.ü. bzgl. $\sigma_k$!). Es gebe ein $g \in L^1(M, \mathbb{R}_+)$ mit*

$$|f_l| \leq g \quad \forall l \in \mathbb{N} \text{ f.ü. in } M \text{ (f.ü. bzgl. } \sigma_k!).$$

*Dann ist $f \in L^1(M, \mathbb{R})$ und*

$$\int_M f = \lim_{l \to \infty} \int_M f_l.$$

BEWEIS. Folgt direkt aus Definition X.3.4, Satz IX.2.14 (S. 309) und Satz IX.3.11 (S. 324). □

BEISPIEL X.3.7. (1) Definiere $\psi : (0, 2\pi) \times (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}^3$ durch

$$\psi(\varphi, \theta) = (\cos\varphi \cos\theta, \sin\varphi \cos\theta, \sin\theta).$$

Dann ist

$$D_1\psi = (-\sin\varphi \cos\theta, \cos\varphi \cos\theta, 0)$$
$$D_2\psi = (-\cos\varphi \sin\theta, -\sin\varphi \sin\theta, \cos\theta)$$

$$((D_i\psi, D_j\psi))_{1\le i,j\le 2} = \begin{pmatrix} \cos^2\theta & 0 \\ 0 & 1 \end{pmatrix}$$

und

$$g_\psi = \cos^2\theta.$$

Wie man leicht nachrechnet, ist $S^2\backslash\psi((0,2\pi)\times(-\frac{\pi}{2},\frac{\pi}{2}))$ eine 2-dimensionale Nullmenge. Daher gilt für jedes $f \in L^1(S^2,\mathbb{R})$

$$\int_{S^2} f = \int_0^{2\pi}\Big(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f\circ\psi(\varphi,\theta)\cos\theta d\theta\Big)d\varphi.$$

Insbesondere ist

$$\begin{aligned}\sigma_2(S^2) &= \int_0^{2\pi}\Big(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos\theta d\theta\Big)d\varphi \\ &= 4\pi.\end{aligned}$$

(2) Sei $I \subset \mathbb{R}$ ein beschränktes, offenes Intervall und $f \in C^1(I,\mathbb{R}_+^*)$. Die Rotationsfläche

$$M = \{x \in \mathbb{R}^3 : x_3 \in I, \sqrt{x_1^2+x_2^2} = f(x_3)\}$$

ist eine zweidimensionale Mfgkt. Bis auf eine zweidimensionale Nullmenge wird sie durch die Karte $(U,\varphi,V)$ mit

$$\begin{aligned}U &= I\times(0,2\pi)\\ \varphi(r,t) &= (f(r)\cos t, f(r)\sin t, r)\\ V &= \varphi(U)\end{aligned}$$

dargestellt. Es ist

$$\begin{aligned}D_1\varphi &= (f'\cos t, f'\sin t, 1)\\ D_2\varphi &= (-f\sin t, f\cos t, 0)\\ ((D_i\varphi, D_j\varphi))_{1\le i,j\le 2} &= \begin{pmatrix} 1+f'^2 & 0 \\ 0 & f^2 \end{pmatrix}\end{aligned}$$

und

$$\sqrt{g_\varphi(r,t)} = f(r)\sqrt{1+f'(r)^2}.$$

Daher ist

$$\begin{aligned}\sigma_2(M) &= \int_0^{2\pi}\Big(\int_I f(r)\sqrt{1+f'(r)^2}dr\Big)dt\\ &= 2\pi\int_I f(r)\sqrt{1+f'(r)^2}dr,\end{aligned}$$

sofern das letzte Integral endlich ist (z.B., wenn $f \in C^1(\overline{I}, \mathbb{R}_+^*)$ ist).
(3) Sei $U \subset \mathbb{R}^{n-1}$ offen und $F \in C^1(U, \mathbb{R})$. Der Graph von $F$

$$M = \{x \in \mathbb{R}^n, (x_1, \dots, x_{n-1}) \in U, x_n = F(x_1, \dots, x_{n-1})\}$$

ist eine Hyperfläche in $\mathbb{R}^n$. $M$ wird durch die Karte $(U, \varphi, M)$ mit

$$\varphi(t) = (t_1, \dots, t_{n-1}, F(t_1, \dots, t_{n-1})) \quad \forall t \in U$$

dargestellt. Es ist

$$D\varphi = \begin{pmatrix} \mathbb{E}_{n-1} \\ \nabla F \end{pmatrix},$$

wobei $\mathbb{E}_{n-1}$ die $(n-1)$-reihige Einheitsmatrix ist. Damit folgt

$$g_\varphi = 1 + \|\nabla F\|_2^2.$$

Also ist

$$\sigma_{n-1}(M) = \int_U \sqrt{1 + \|\nabla F\|_2^2},$$

sofern das letzte Integral endlich ist.
(4) Sei

$$M = \{x \in \mathbb{R}^n : \|x\|_2 = r, x_n > 0\}$$

die obere Halbsphäre mit Radius $r > 0$. Sie ist Graph der Funktion $F : U \to \mathbb{R}$ mit

$$U = \{t \in \mathbb{R}^{n-1} : \|t\|_2 < r\} = B_{n-1}(0, r)$$

$$F(t) = \sqrt{r^2 - \sum_{i=1}^{n-1} t_i^2}.$$

Wegen

$$D_i F(t) = -\frac{t_i}{F(t)} \quad \forall t \in U$$

folgt für die Immersion $\varphi$ aus Teil (3)

$$\sqrt{g_\varphi(t)} = \sqrt{1 + \frac{\|t\|_2^2}{F(t)^2}} = \frac{r}{F(t)} \quad \forall t \in U.$$

Daher gilt für jedes $f \in L^1(M, \mathbb{R})$

$$\begin{aligned}\int_M f &= \int_U f(t, \sqrt{r^2 - \|t\|_2^2}) \frac{r}{\sqrt{r^2 - \|t\|_2^2}}\, dt \\ &= \int_{B_{n-1}(0,1)} f(rz, r\sqrt{1 - \|z\|_2^2}) \frac{r^{n-1}}{\sqrt{1 - \|z\|_2^2}}\, dz,\end{aligned}$$

wobei wir die Transformation $t = rz$ auf $\mathbb{R}^{n-1}$ angewandt haben.

Seien $U, V \in \mathbb{R}^n$ offen, $F : U \to V$ ein $C^\alpha$-Diffeomorphismus und $M \subset U$ eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$. Dann ist $F(M)$ ebenfalls eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$ (Übungsaufgabe). Einen Transformationssatz analog zu Satz IX.5.6 (S. 343) kann man für die Integrale auf $M$ und $F(M)$ nicht in der dortigen einfachen Form und Allgemeinheit angeben. Ist allerdings $F$ die Komposition von Translationen, Rotationen oder Homothetien, so gibt es ein einfaches Analogon zu Satz IX.5.6 (S. 343).

Satz X.3.8. *Sei $M \subset \mathbb{R}^n$ eine $k$-dimensionale Mfgkt der Klasse $C^\alpha$. Die Diffeomorphismen $\tau_a, \rho, \Theta_r : \mathbb{R}^n \to \mathbb{R}^n$ seien definiert durch*

(Translation) $\quad \tau_a(x) = a + x \quad (a \in \mathbb{R}^n$ *fest*)

(Rotation) $\quad \rho(x) = Qx \quad (Q \in \mathcal{O}(n)$ *fest*)

(Homothetie) $\quad \Theta_r(x) = rx \quad (r \in \mathbb{R}_+^*$ *fest*).

*Dann gilt*

$$\begin{aligned} f \in L^1(M, \mathbb{R}) &\iff f \circ \tau_a^{-1} \in L^1(\tau_a(M), \mathbb{R}) \\ &\iff f \circ \rho^{-1} \in L^1(\rho(M), \mathbb{R}) \\ &\iff f \circ \Theta_r^{-1} \in L^1(\Theta_r(M), \mathbb{R}). \end{aligned}$$

*In diesem Fall ist*

$$\begin{aligned} \int_{\tau_a(M)} f \circ \tau_a^{-1} &= \int_M f, \\ \int_{\rho(M)} f \circ \rho^{-1} &= \int_M f, \\ \int_{\Theta_r(M)} f \circ \Theta_r^{-1} &= r^k \int_M f. \end{aligned}$$

Beweis. Sei $(U, \varphi, V)$ eine Karte von $M$. Dann sind $(U, \tau_a \circ \varphi, \tau_a(V))$, $(U, \rho \circ \varphi, \rho(V))$ und $(U, \Theta_r \circ \varphi, \Theta_r(V))$ Karten von $\tau_a(M)$, $\rho(M)$ und $\Theta_r(M)$. Weiter ist

$$\begin{aligned} g_{\tau_a \circ \varphi} &= g_\varphi \\ g_{\rho \circ \varphi} &= \det(D(\rho \circ \varphi)^t D(\rho \circ \varphi)) \\ &= \det(D\varphi^t Q^t Q D\varphi) \\ &= g_\varphi \\ g_{\Theta_r \circ \varphi} &= \det(D(\Theta_r \circ \varphi)^t D(\Theta_r \circ \varphi)) \\ &= \det(r^2 D\varphi^t D\varphi) \\ &= r^{2k} g_\varphi. \end{aligned}$$

Damit folgt die Behauptung aus Definition X.3.4 und Satz IX.5.6 (S. 343). □

Der folgende Satz ist eine interessante Variante des Satzes von Fubini, Satz IX.3.3.

SATZ X.3.9. *Sei $f \in L^1(\mathbb{R}^n, \mathbb{R})$. Dann ist für $\lambda_1$ fast alle $r \in \mathbb{R}_+^*$ die Funktion $f$ auf der Sphäre $\partial B(0,r)$ integrierbar und es gilt*

$$\int_{\mathbb{R}^n} f(x)dx = \int_0^\infty \Big\{ \int_{\partial B(0,r)} f(y)ds(y) \Big\} dr$$
$$= \int_0^\infty \Big\{ \int_{\partial B(0,1)} f(ry)ds(y) \Big\} r^{n-1} dr.$$

BEWEIS. Seien

$$H^\pm = \{x \in \mathbb{R}^n : \pm x_n > 0\}$$

der obere bzw. untere Halbraum. Da $\mathbb{R}^{n-1} \times \{0\}$ eine Nullmenge in $\mathbb{R}^n$ ist, reicht es, die Behauptung für $f\chi_{H^\pm}$ zu beweisen. Wir betrachten nur $g = f\chi_{H^+}$. Der andere Fall ist völlig analog.
Sei $U = B_{n-1}(0,1) = \{u \in \mathbb{R}^{n-1} : \|u\|_2 < 1\}$. Wie man leicht nachrechnet, ist $\Phi : U \times \mathbb{R}_+^* \to H^+$ mit

$$\Phi(u,r) = (ru_1, \ldots, ru_{n-1}, r\sqrt{1 - \|u\|_2^2})$$

ein $C^\infty$-Diffeomorphismus mit

$$D\Phi(u,r) = \begin{pmatrix} r\mathbb{E}_{n-1} & u \\ \frac{ru}{\sqrt{1-\|u\|_2^2}} & \sqrt{1 - \|u\|_2^2} \end{pmatrix}$$

und

$$\det D\Phi(u,r) = r^{n-1}\sqrt{1 - \|u\|_2^2} + r^{n-1}\frac{\|u\|_2^2}{\sqrt{1 - \|u\|_2^2}}$$
$$= \frac{r^{n-1}}{\sqrt{1 - \|u\|_2^2}}.$$

Aus Satz IX.3.3, Satz IX.5.6 (S. 343), Satz X.3.8 und Beispiel X.3.7 (4) folgt daher

$$\int_{H^+} g = \int_0^\infty \Big\{ \int_U g(ru, r\sqrt{1 - \|u\|_2^2}) \frac{r^{n-1}}{\sqrt{1 - \|u\|_2^2}} du \Big\} dr$$
$$= \int_0^\infty \Big\{ \int_{\partial B(0,r) \cap H^+} g(y)ds(y) \Big\} dr$$
$$= \int_0^\infty \Big\{ \int_{\partial B(0,1) \cap H^+} g(rz)ds(z) \Big\} r^{n-1} dr.$$

□

BEISPIEL X.3.10. (1) Aus Satz X.3.8 folgt für $x \in \mathbb{R}^n, r \in \mathbb{R}_+^*$

$$\sigma_{n-1}(\partial B(x,r)) = r^{n-1}\sigma_{n-1}(S^{n-1}).$$

Sei $\tau_n = \sigma_{n-1}(S^{n-1})$ und $\omega_n = \lambda_n(B_n(0,1))$. Aus Beispiel IX.5.7(1) (S. 345) und Satz X.3.9 folgt

$$\begin{aligned}\omega_n &= \int_0^1 \sigma_{n-1}(S^{n-1}) r^{n-1} dr \\ &= \frac{1}{n}\tau_n\end{aligned}$$

und somit

$$\begin{aligned}\tau_n &= n\omega_n \\ &= n\frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} \\ &= \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}.\end{aligned}$$

(2) Sei $F \in L^1(\mathbb{R}^n, \mathbb{R})$ ROTATIONSSYMMETRISCH, d.h.

$$F(x) = f(\|x\|_2) \quad \forall x \in \mathbb{R}^n$$

mit $f : \mathbb{R}_+ \to \mathbb{R}$. Aus Satz X.3.9 folgt

$$\begin{aligned}\int_{\mathbb{R}^n} F &= \int_0^\infty \tau_n f(r) r^{n-1} dr \\ &= \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})} \int_{\mathbb{R}} f(r) r^{n-1} dr.\end{aligned}$$

Im zweiten Teil dieses Abschnittes wollen wir den GAUSSSCHEN INTEGRALSATZ beweisen. Er verbindet das Integral über eine beschränkte Teilmenge $M$ des $\mathbb{R}^n$ mit einem Integral über dem Rand $\partial M$ von $M$. Für seine genaue Formulierung und den Beweis benötigen wir einige Notationen und Hilfsresultate.

DEFINITION X.3.11. Sei $K \subset \mathbb{R}^n$ kompakt. Es gebe zwei disjunkte Teilmengen $\partial_R K$ und $\partial_S K$ von $\partial K$ mit folgenden Eigenschaften:

(1) $\partial_R K$ ist relativ offen in $\partial K$.
(2) Zu jedem $x_0 \in \partial_R K$ existiert eine Umgebung $U$ von $x_0$ in $\mathbb{R}^n$ und eine Funktion $\psi \in C^1(U, \mathbb{R})$, derart dass jedes $x \in U$ regulärer Punkt von $\psi$ ist und dass gilt

$$K \cap U = \{x \in U : \psi(x) \le 0\}.$$

(3) $\lim_{\varepsilon \to 0+} \varepsilon^{-1} \lambda_n(\bigcup_{x \in \partial_S K} B(x;\varepsilon)) = 0.$
(4) $\partial K = \partial_R K \cup \partial_S K$.

Dann heißt $K$ STÜCKWEISE GLATT BERANDET. Ist $\partial_S K = \emptyset$, so heißt $K$ GLATT BERANDET. $\partial_R K$ und $\partial_S K$ heißen der REGULÄRE bzw. SINGULÄRE RAND von $K$.

Beispiel X.3.12. (1) Die Kugel $\overline{B(0,1)} \subset \mathbb{R}^n$ ist glatt berandet; die Funktion

$$\psi(x) = \sum_{i=1}^{n} x_i^2 - 1$$

leistet das Gewünschte.
(2) Der Zylinder $Z = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq 1, 0 \leq x_3 \leq 1\}$ ist stückweise glatt berandet mit

$$\begin{aligned}\partial_R Z = \{x \in \mathbb{R}^3 &: x_1^2 + x_2^2 = 1, 0 < x_3 < 1\} \\ &\cup \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 < 1, x_3 = 0\} \\ &\cup \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 < 1, x_3 = 1\}\end{aligned}$$

und

$$\begin{aligned}\partial_S Z = \{x \in \mathbb{R}^3 &: x_1^2 + x_2^2 = 1, x_3 = 0\} \\ &\cup \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 = 1, x_3 = 1\}.\end{aligned}$$

Die Funktionen

$$\begin{aligned}\psi_1(x) &= x_1^2 + x_2^2 - 1 \\ \psi_2(x) &= -x_3 \\ \psi_3(x) &= x_3 - 1\end{aligned}$$

leisten das in Bedingung (2) Gewünschte. Für $\varepsilon > 0$ ist

$$\begin{aligned}\lambda_3\Big(\bigcup_{x \in \partial_S Z} B(x,\varepsilon)\Big) &= 2\lambda_3(\{x \in \mathbb{R}^3 : (\sqrt{x_1^2 + x_2^2} - 1)^2 + x_3^2 \leq \varepsilon^2\}) \\ &= 2\int_0^{\varepsilon}\int_0^{2\pi}\int_0^{2\pi} r(1 + r\cos\theta)d\theta d\varphi dr \\ &= 2 \cdot (2\pi)^2 \int_0^{\varepsilon} rdr \\ &= 4\pi^2\varepsilon^2\end{aligned}$$

und daher

$$\begin{aligned}\varepsilon^{-1}\lambda_3\Big(\bigcup_{x \in \partial_S z} B(x,\varepsilon)\Big) &= 4\pi^2\varepsilon \\ &\underset{\varepsilon \to 0}{\longrightarrow} 0.\end{aligned}$$

Satz X.3.13. *Sei $K \subset \mathbb{R}^n$ kompakt und stückweise glatt berandet. Dann ist $\partial_R K$ eine Hyperfläche.*

Beweis. Seien $x_0 \in \partial_R K$ beliebig und $U$, $\psi$ wie in Definition X.3.11. O.E. können wir $U \cap \partial_S K = \emptyset$ voraussetzen, da $\partial_R K$ relativ offen in $\partial K$ ist. Wir zeigen, dass $\partial_R K \cap U = \psi^{-1}\backslash(\{0\})$ ist. Hieraus folgt dann die Behauptung.
Sei zunächst $x \in U$ mit $\psi(x) < 0$. Dann gibt es eine Umgebung $V$ von

$x$ mit $V \subset U$ und $\psi(y) < 0$ für alle $y \in V$. Also ist $V \subset K$ und damit $x$ innerer Punkt von $K$. Dies zeigt $\partial_R K \cap U \subset \psi^{-1}(\{0\})$. Sei nun $x \in U$ mit $\psi(x) = 0$ und $v = \operatorname{grad} \psi(x) \neq 0$. Für hinreichend kleines $t$ gilt dann

$$\begin{aligned}\psi(x + tv) &= \psi(x) + tD\psi(x)v + r(tv) \\ &= t\|v\|_2^2 + r(tv)\end{aligned}$$

mit $\lim\limits_{t \to 0} t^{-1} r(tv) = 0$. Also gibt es ein $\varepsilon > 0$ mit

$$\begin{aligned}&\psi(x + tv) > 0 \quad \forall 0 < t < \varepsilon \\ &\psi(x + tv) < 0 \quad \forall -\varepsilon < t < 0\end{aligned}$$

und somit

$$\begin{aligned}&x + tv \notin K \quad \forall 0 < t < \varepsilon \\ &x + tv \in K \quad \forall -\varepsilon < t < 0.\end{aligned}$$

Also ist $x \in \partial K$ und wegen $U \cap \partial_S K = \emptyset$ aus $\partial_R K$. Mithin ist $\psi^{-1}(\{0\}) \subset \partial_R K \cap U$. □

SATZ X.3.14. *Sei $K \subset \mathbb{R}^n$ kompakt und stückweise glatt berandet. Zu jedem $x_0 \in \partial_R K$ existiert genau ein Vektor $\nu(x_0) \in \mathbb{R}^n$ mit den folgenden Eigenschaften:*

(1) $\nu(x_0) \in N_{x_0} \partial_R K$.
(2) $\|\nu(x_0)\|_2 = 1$.
(3) *Es gibt ein $\varepsilon > 0$ mit*

$$x_0 + t\nu(x_0) \notin K \quad \forall 0 < t < \varepsilon.$$

*Der Vektor $\nu(x_0)$ heißt* ÄUSSERER NORMALEN-EINHEITSVEKTOR *von $K$ im Punkt $x_0$.*

BEWEIS. EXISTENZ: Seien $x_0 \in \partial_R K$ beliebig und $U$ und $\psi$ wie in Bedingung (2) von Definition X.3.11 mit $U \cap \partial_S K = \emptyset$. Aus dem Beweis von Satz X.3.13 folgt, dass $\nu(x_0) = \frac{\nabla\psi(x_0)}{\|\nabla\psi(x_0)\|_2}$ das Gewünschte leistet.
EINDEUTIGKEIT: Da nach Satz X.2.2 (S. 364)

$$N_{x_0} \partial_R K = \operatorname{span}\{\nabla\psi(x_0)\}$$

ist, folgt

$$\nu(x_0) = \lambda \nabla\psi(x_0) \quad \text{mit } \lambda \in \mathbb{R}.$$

Wegen (2) ist $|\lambda| = \|\nabla\psi(x_0)\|_2^{-1}$. Wegen (3) folgt $\lambda > 0$. Also ist $\nu(x_0)$ eindeutig bestimmt. □

BEMERKUNG X.3.15. Sei $K \subset \mathbb{R}^n$ kompakt und stückweise glatt berandet. Aus dem Beweis von Satz X.3.14 folgt, dass die Zuordnung $x \mapsto \nu(x)$ auf $\partial_R K$ stetig ist. Insbesondere ist $\partial_R K$ orientierbar.

SATZ X.3.16. *Sei $U \subset \mathbb{R}^n$ offen, $U \neq \emptyset$, und $f \in C_0^1(U,\mathbb{R})$. Dann gilt für alle $1 \le i \le n$*

$$\int_U \frac{\partial}{\partial x_i} f(x)dx = 0.$$

BEWEIS. Sei $R > 0$ so, dass $\mathrm{supp}(f) \subset B_{\|\cdot\|_\infty}(0,R)$ ist. Indem wir $f$ außerhalb $U$ durch 0 fortsetzen, erhalten wir eine Funktion aus $C_0^1(\mathbb{R}^n,\mathbb{R})$ mit gleichem Träger. Für $1 \le i \le n$ folgt dann aus Satz X.3.3

$$\begin{aligned}
&\int_U \frac{\partial}{\partial x_i} f(x)dx \\
&= \int_{B_{\|\cdot\|_\infty}(0,R)} \frac{\partial}{\partial x_i} f(x)dx \\
&= \int_{-R}^{R} \cdots \int_{-R}^{R} \Big\{ \int_{-R}^{R} \frac{\partial}{x_i} f(x)dx_i \Big\} dx_1 \ldots dx_{i-1} dx_{i+1} \ldots dx_n \\
&= 0.
\end{aligned}$$

□

SATZ X.3.17. *Seien $U' \subset \mathbb{R}^{n-1}$ offen, $I = (\alpha,\beta)$, $\alpha < \beta$, $g \in C^1(U',\mathbb{R})$ mit $g(U') \subset I$ und*

$$\begin{aligned}
A &= \{(x',x_n) \in U' \times I : x_n \le g(x')\} \\
M &= \{(x',x_n) \in U' \times I : x_n = g(x')\}.
\end{aligned}$$

*Dann gilt für alle $f \in C_0^1(U' \times I,\mathbb{R})$ und alle $1 \le i \le n$*

$$\int_A \frac{\partial}{\partial x_i} f(x)dx = \int_M f(x)\nu_i(x)dS(x),$$

*wobei*

$$\begin{aligned}
\nu_i(x) &= -(1 + \|\nabla g(x')\|_2^2)^{-1/2} \frac{\partial g(x')}{\partial x_i}, \quad 1 \le i \le n-1 \\
\nu_n(x) &= (1 + \|\nabla g(x')\|_2^2)^{-1/2}
\end{aligned}$$

*ist.*

BEWEIS. Wir bemerken zunächst, dass gemäß Beispiel X.3.7 (3) $M$ eine Hyperfläche und die Gramsche Determinante gleich $1 + \|\nabla g(x')\|_2^2$ ist.
FALL $1 \le i \le n-1$: Definiere $F : U' \times I \to \mathbb{R}$ durch

$$F(x',z) = \int_\alpha^z f(x',x_n)dx_n.$$

Dann gilt

$$\frac{\partial F}{\partial z}(x',z) = f(x',z)$$

$$\frac{\partial F}{\partial x_i}(x', z) = \int_\alpha^z \frac{\partial}{\partial x_i} f(x', z) dx_n$$

und daher

$$\begin{aligned}
\frac{\partial}{\partial x_i} \int_\alpha^{g(x')} f(x', x_n) dx_n &= \frac{\partial}{\partial x_i} F(x', g(x')) \\
&= \frac{\partial}{\partial x_i} F(x', g(x')) \\
&\quad + \frac{\partial}{\partial z} F(x', g(x')) \frac{\partial g}{\partial x_i}(x') \\
&= \int_\alpha^{g(x')} \frac{\partial}{\partial x_i} f(x', x_n) dx_n \\
&\quad + f(x', g(x')) \frac{\partial g}{\partial x_i}(x').
\end{aligned}$$

Da die Funktion $x' \mapsto F(x', g(x'))$ kompakten Träger in $U'$ hat, folgt aus Satz X.3.16

$$\int_{U'} \frac{\partial}{\partial x_i} F(x', g(x')) dx' = 0.$$

Damit folgt aus Satz X.3.3

$$\begin{aligned}
\int_A \frac{\partial f}{\partial x_i}(x) dx &= \int_{U'} \Big\{ \int_\alpha^{g(x')} \frac{\partial}{\partial x_i} f(x', x_n) dx_n \Big\} dx' \\
&= \int_{U'} \frac{\partial}{\partial x_i} F(x', g(x')) dx' \\
&\quad - \int_{U'} f(x', g(x')) \frac{\partial g}{\partial x_i}(x') dx' \\
&= \int_M f(x) \nu_i(x) dS(x).
\end{aligned}$$

Fall $i = n$: Es ist wegen Satz X.3.3 und $\operatorname{supp}(f) \Subset U' \times I$

$$\begin{aligned}
\int_A \frac{\partial}{\partial x_n} f(x) dx &= \int_{U'} \Big\{ \int_\alpha^{g(x')} \frac{\partial}{\partial x_n} f(x', x_n) dx_n \Big\} dx' \\
&= \int_{U'} \Big\{ f(x', g(x')) - \underbrace{f(x', \alpha)}_{=0} \Big\} dx' \\
&= \int_{U'} f(x', g(x')) dx' \\
&= \int_M f(x) \nu_n(x) dS(x).
\end{aligned}$$

$\square$

Definition X.3.18. Sei $U \subset \mathbb{R}^n$ offen und $f \in C^1(U, \mathbb{R}^n)$. Dann ist die Divergenz von $f$ definiert durch

$$\nabla \cdot f = \operatorname{div} f = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}.$$

Satz X.3.19 (Gaussscher Integralsatz). *Sei $K \subset \mathbb{R}^n$ kompakt und stückweise glatt berandet mit $\sigma_{n-1}(\partial_R K) < \infty$. Dann gilt für jede offene Menge $U \subset \mathbb{R}^n$ mit $K \subset U$ und jedes $f \in C^1(U, \mathbb{R}^n)$*

$$\int_K \operatorname{div} f dx = \int_{\partial_R K} (f, \nu) dS(x),$$

*wobei $\nu$ das äußere Einheits-Normalenfeld an $\partial_R K$ ist.*

Beweis. Sei $\varepsilon > 0$ beliebig. Da $K$ kompakt ist, gibt es Zahlen $1 < m_0 < m_R < m_S$ und offene Mengen $U_i, 1 \le i \le m_S$, mit folgenden Eigenschaften:

(1) $K \subset \bigcup_{i=1}^{m_S} U_i$.
(2) $U_i \subset \overset{\circ}{K}$ für alle $1 \le i \le m_0$.
(3) $U_i \cap \partial_S K = \emptyset$ für alle $m_0 < i \le m_R$ und nach einer eventuellen Koordinatentransformation hat $U_i$ die Gestalt $U_i = U' \times I$ mit $U' \subset \mathbb{R}^{n-1}$ offen, $I = (\alpha, \beta) \subset \mathbb{R}$, $\alpha < \beta$ und

$$U_i \cap K = \{(x', x_n) \in U' \times I : x_n \le g(x')\}$$

für ein $g \in C^1(U', \mathbb{R})$ mit $g(U') \subset I$.
(4) $U_i = B(x_i, \varepsilon)$ mit $x_i \in \partial_S K$ für alle $m_R < i \le m_S$.

Sei $\alpha_1, \ldots, \alpha_{m_S}$ eine Partition der Eins auf $K$, die $U_1, \ldots, U_{m_S}$ untergeordnet ist. Definiere

$$\varphi_\varepsilon = \sum_{i=m_R+1}^{m_S} \alpha_i.$$

Für $1 \le j \le m_0$ folgt aus Satz X.3.16 angewandt auf die Komponenten von $\alpha_j f$

$$\int_K \operatorname{div}(\alpha_j f) dx = 0 = \int_{\partial_R K} (\alpha_j f, \nu) dS.$$

Für $m_0 < j \le m_R$ folgt aus Satz X.3.17 angewandt auf die Komponenten von $\alpha_j f$

$$\int_K \operatorname{div}(\alpha_j f) dx = \int_{\partial_R K} (\alpha_j f, \nu) dS.$$

Summation über $j = 1, \ldots, m_R$ liefert wegen

$$\sum_{j=1}^{m_R} \alpha_j = 1 - \varphi_\varepsilon \text{ auf } K$$

die Identität

$$(*) \qquad \int_K \operatorname{div}((1-\varphi_\varepsilon)f)dx = \int_{\partial_R K} ((1-\varphi_\varepsilon)f,\nu)dS.$$

Für $\varepsilon \longrightarrow 0$ konvergiert $((1-\varphi_\varepsilon)f,\nu)$ punktweise auf $\partial_R K$ gegen $(f,\nu)$ und ist für jedes $\varepsilon > 0$ beschränkt durch die integrierbare Funktion $\|f\|_{C^0(K,\mathbb{R}^n)}\chi_{\partial_R K}$. Damit folgt aus Satz X.3.6

$$\lim_{\varepsilon\to 0}\int_{\partial_R K} ((1-\varphi_\varepsilon)f,\nu)dS = \int_{\partial_R K}(f,\nu)dS.$$

Aus Bemerkung IX.6.8 (S. 355) und Eigenschaft (3) von Definition X.3.11 folgt andererseits

$$\begin{aligned}
&|\int_K \operatorname{div} f dx - \int_K \operatorname{div}((1-\varphi_\varepsilon)f)dx| \\
&= |\int_K \varphi_\varepsilon \operatorname{div} f dx + \int_K (f,\nabla\varphi_\varepsilon)dx| \\
&\le \|\operatorname{div} f\|_{C^0(K,\mathbb{R})}\lambda_n(\bigcup_{x\in\partial_S K} B(x,\varepsilon)) \\
&\quad + c_1\|f\|_{C^0(K,\mathbb{R}^n)}\varepsilon^{-1}\lambda_n(\bigcup_{x\in\partial_S K} B(x,\varepsilon)) \\
&\underset{\varepsilon\to 0}{\longrightarrow} 0.
\end{aligned}$$

Damit folgt die Behauptung durch Grenzübergang $\varepsilon \longrightarrow 0$ in $(*)$. □

Bemerkung X.3.20. (1) Die Voraussetzungen an $K$ und $f$ in Satz X.3.19 können wesentlich abgeschwächt werden.
(2) Sei $U \subset \mathbb{R}^n$ offen, $f \in C^1(U,\mathbb{R}^n)$ und $x_0 \in U$. Aus Satz X.3.19 folgt

$$\operatorname{div} f(x_0) = \lim_{\varepsilon\to 0+}\frac{1}{\lambda_n(B(x_0,\varepsilon))}\int_{\partial B(x_0,\varepsilon)}(f,\nu)dS.$$

Die Divergenz von $f$ in $x_0$ ist also physikalisch der Fluss pro Volumeneinheit durch die Oberfläche einer beliebig kleinen Kugel um $x_0$.

Beispiel X.3.21. (1) Sei $Z = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 \le 1, 0 \le x_3 \le 1\}$ und

$$f(x) = (x_2x_3, x_1x_3, x_1^2 + x_2^2).$$

Dann ist

$$\operatorname{div} f(x) = 0 \quad \forall x \in \mathbb{R}^3$$

und

$$\int_Z \operatorname{div} f = 0.$$

Auf dem Mantel von $Z$ ist $\nu(x) = (x_1, x_2, 0)$ und

$$\begin{aligned}\int_{\text{Mantel}} (f, \nu) dS &= \int_{\text{Mantel}} 2x_1x_2x_3 dS \\ &= \int_0^1 \Big\{ \int_0^{2\pi} 2z \cos\varphi \sin\varphi d\varphi \Big\} dz \\ &= 0.\end{aligned}$$

Auf dem Boden von $Z$ ist $\nu(x) = (0, 0, -1)$ und

$$\begin{aligned}\int_{\text{Boden}} (f, \nu) &= - \int_0^1 \Big\{ \int_0^{2\pi} r^2 r d\varphi \Big\} dr \\ &= -\frac{\pi}{2}.\end{aligned}$$

Auf dem Deckel von $Z$ ist schließlich $\nu(x) = (0, 0, 1)$ und

$$\begin{aligned}\int_{\text{Deckel}} (f, \nu) &= \int_0^1 \Big\{ \int_0^{2\pi} r^2 r d\varphi \Big\} dr \\ &= \frac{\pi}{2}.\end{aligned}$$

(2) Für $K = \overline{B(0,1)} \subset \mathbb{R}^n$ und $f(x) = x$ ergibt sich

$$\begin{aligned}n\lambda_n(\overline{B(0,1)}) &= \int_{\overline{B(0,1)}} \operatorname{div} f dx \\ &= \int_{\partial B(0,1)} (f, \nu) dS \\ &= \sigma_{n-1}(S^{n-1}).\end{aligned}$$

Damit haben wir die Formel aus Beispiel X.3.10 (1) auf andere Weise bewiesen.

(3) Der Körper $K$ sei ganz in eine Flüssigkeit der konstanten Dichte $\rho > 0$, deren Oberfläche mit der Ebene $x_3 = 0$ zusammenfalle, eingetaucht. $K$ erfülle die Voraussetzungen von Satz X.3.19. Die Flüssigkeit übt im Punkt $x \in \partial_R K$ auf $K$ den Druck $\rho x_3 \nu(x)$ aus, wobei $\nu(x)$ der äußere Normalen-Einheitsvektor ist (der Druck ist also nach innen gerichtet!). Die Kraft $F = (F_1, F_2, F_3)$, die auf $K$ wirkt, ist dann gegeben durch

$$\begin{aligned}F_i &= \int_{\partial_R K} \rho x_3 \nu_i(x) dS(x) \\ &= \int_K \rho \frac{\partial x_3}{\partial x_i} dx \\ &= \begin{cases} 0 & \text{falls } i = 1, 2 \\ \rho \lambda_3(K) & \text{falls } i = 3. \end{cases}\end{aligned}$$

Der Körper erfährt also einen Auftrieb in $x_3$-Richtung, dessen Betrag gleich dem Gewicht der verdrängten Flüssigkeit ist (Archimedisches Prinzip).

Definition X.3.22. (1) Für $a, b \in \mathbb{R}^3$ ist das Vektorprodukt $a \times b$ definiert durch

$$a \times b = (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1).$$

(2) Seien $U \subset \mathbb{R}^n$ offen, $K \subset U$ kompakt mit stückweise glattem Rand und $\varphi \in C^1(U, \mathbb{R})$, $f \in C^2(U, \mathbb{R})$. Dann heißt

$$\frac{\partial \varphi}{\partial \nu}(x) = (\nabla \varphi(x), \nu(x)) \quad \forall x \in \partial_R K$$

die Normalenableitung von $\varphi$ auf $\partial_R K$ und

$$\Delta f = \operatorname{div}(\operatorname{grad} f) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}$$

der Laplace von $f$.
(3) Sei $U \subset \mathbb{R}^3$ offen und $f \in C^1(U, \mathbb{R}^3)$. Dann heißt

$$\operatorname{rot} f = \nabla \times f = (\partial_2 f_3 - \partial_3 f_2, \partial_3 f_1 - \partial_1 f_3, \partial_1 f_2 - \partial_2 f_1)$$

die Rotation von $f$.

Mit Hilfe des Gaußschen Integralsatzes kann man verschiedene Formeln für die Ableitungen von Funktionen herleiten, von denen wir einige im folgenden Satz angeben.

Satz X.3.23. *Seien $U \subset \mathbb{R}^n$ offen und $K \subset U$ kompakt mit stückweise glattem Rand und $\sigma_{n-1}(\partial_R K) < \infty$.*

(1) *Für $f \in C^2(U, \mathbb{R})$ und $g \in C^1(U, \mathbb{R})$ gilt*

$$\int_K \Delta f g = -\int_K (\nabla f, \nabla g) + \int_{\partial_R K} \frac{\partial f}{\partial \nu} g.$$

(2) *Für $f, g \in C^2(U, \mathbb{R})$ gilt die* Greensche Formel

$$\int_K \{\Delta f g - f \Delta g\} = \int_{\partial_R K} \{\frac{\partial f}{\partial \nu} g - \frac{\partial g}{\partial \nu} f\}.$$

(3) *Für $n = 3$ und $f \in C^1(U, \mathbb{R}^3)$ gilt*

$$\int_K \operatorname{rot} f = -\int_{\partial_R K} f \times \nu.$$

Beweis. ad (1): Aus Satz X.3.19 folgt

$$\begin{aligned}\int_{\partial_R K} \frac{\partial f}{\partial \nu} g &= \int_{\partial_R K} (g \nabla f, \nu) \\ &= \int_K \operatorname{div}(g \nabla f) \\ &= \int_K (\nabla f, \nabla g) + \int_K \Delta f g.\end{aligned}$$

AD (2): Vertausche $f$ und $g$ in (1) und subtrahiere die beiden Gleichungen.
AD (3): Aus Satz X.3.19 folgt

$$\begin{aligned}\int_K (\operatorname{rot} f)_1 &= \int_K \partial_2 f_3 - \int_K \partial_3 f_2 \\ &= \int_K \operatorname{div}(f_3 e_2) - \int_K \operatorname{div}(f_2 e_3) \\ &= \int_{\partial_R K} f_3 \nu_2 - \int_{\partial_R K} f_2 \nu_3 \\ &= -\int_{\partial_R K} (f \times \nu)_1.\end{aligned}$$

Analog für die anderen Komponenten. □

## X.4. Multilineare Algebra

Im Folgenden seien $V$, $W$ stets $\mathbb{R}$-Vektorräume. Wir knüpfen an die Abschnitte VII.1 und VII.4 an, in denen wir stetige, lineare bzw. multilineare Abbildungen betrachtet haben. Zur Abkürzung definieren wir

$$V^* = \mathcal{L}(V, \mathbb{R}),$$

den DUALRAUM von $V$, und bezeichnen mit

$$\begin{aligned}\langle \cdot, \cdot \rangle : &V^* \times V \to \mathbb{R} \\ &(\varphi, r) \mapsto \langle \varphi, r \rangle = \varphi(r)\end{aligned}$$

die DUALE PAARUNG zwischen $V$ und $V^*$. Ist $n = \dim V < \infty$, so folgt aus Satz VII.1.10 (S. 215)

$$V^* \cong \mathbb{R}^n.$$

Ist insbesondere $e_1, \ldots, e_n$ eine Basis von $V$, so ist $\varepsilon_1, \ldots, \varepsilon_n \in V^*$ mit

$$\langle \varepsilon_i, \sum_{j=1}^n \alpha_j e_j \rangle = \alpha_i \qquad \forall v = \sum_{j=1}^n \alpha_j e_j \in V$$

eine Basis von $V^*$. Wir nennen $\varepsilon_1, \ldots, \varepsilon_n$ die DUALBASIS zu $e_1, \ldots, e_n$. Insbesondere ist

$$\langle \varepsilon_i, e_j \rangle = \delta_{ij}.$$

Schließlich erinnern wir an die Definition VII.4.5 (S. 232) des Raumes $\mathcal{L}^r(V, \mathbb{R})$ der stetigen, $r$-linearen Abbildungen von $V$ in $\mathbb{R}$.

DEFINITION X.4.1. (1) Sei $r \geq 2$. Eine stetige, $r$-lineare Abbildung $\alpha \in \mathcal{L}^r(V, \mathbb{R})$ mit der Eigenschaft

(∗) $\quad \alpha(v_1, \ldots, v_r) = 0 \quad$ falls $v_i = v_j$ für ein Paar $(i, j)$ mit $i \neq j$

heißt ALTERNIERENDE $r$-FORM.
(2) Definiere

$$\Lambda^0(V) = \mathbb{R},$$
$$\Lambda^1(V) = V^*,$$
$$\Lambda^r(V) = \{\alpha \in \mathcal{L}^r(V, \mathbb{R}) : \alpha \text{ ist alternierende } r - \text{Form}\},\ r \geq 2.$$

(3) $\sigma_r$ bezeichnet die Gruppe der Permutationen von $\{1, \dots, r\}$.

BEISPIEL X.4.2. Durch

$$v_1, \dots, v_n \in \mathbb{R}^n \mapsto \det(v_1, \dots, v_n) \in \mathbb{R}$$

wird eine alternierende $n$-Form auf $\mathbb{R}^n$ definiert.

BEMERKUNG X.4.3. (1) Für $r \geq 2$ und $\alpha \in \mathcal{L}^r(V, \mathbb{R})$ sind folgende Aussagen äquivalent:

(i) $\alpha \in \Lambda^r(V, \mathbb{R})$.

(ii) $\alpha(v_{\sigma(1)}, \dots, v_{\sigma(r)}) = \operatorname{sgn}(\sigma)\alpha(v_1, \dots, v_r)$
$\forall v_1, \dots, v_r \in V, \sigma \in \sigma_r$.

(iii) $\alpha(v_1, \dots, v_i, \dots, v_k, \dots, v_r) = -\alpha(v_1, \dots, v_k, \dots, v_i, \dots, v_r)$
$\forall v_1, \dots, v_r \in V, i \neq k$.

(2) Ist $\alpha \in \Lambda^r(V, \mathbb{R})$ und sind $v_1, \dots, v_r \in V$ linear abhängig, $r \geq 2$, so ist

$$\alpha(v_1, \dots, v_r) = 0.$$

Insbesondere ist $\Lambda^r(V) = \{0\}$, falls $r > \dim V$ ist.
(3) $\Lambda^r(V)$ ist ein Untervektorraum von $\mathcal{L}^r(V, \mathbb{R})$.

BEWEIS. AD (1): „(i) $\Longrightarrow$ (iii)“: Seien $v_1, \dots, v_r \in V$ und $1 \leq i < k \leq r$. Dann gilt

$$\begin{aligned}
0 &= \alpha(v_1, \dots, v_{i-1}, v_i + v_k, v_{i+1}, \dots, v_{k-1}, v_i + v_k, v_{k+1}, \dots, v_r) \\
&= \alpha(v_1, \dots, v_i, \dots, v_i, \dots, v_r) \\
&\quad + \alpha(v_1, \dots, v_i, \dots, v_k, \dots, v_r) \\
&\quad + \alpha(v_1, \dots, v_k, \dots, v_i, \dots, v_r) \\
&\quad + \alpha(v_1, \dots, v_k, \dots, v_k, \dots, v_r) \\
&= \alpha(v_1, \dots, v_i, \dots, v_k, \dots, v_r) \\
&\quad + \alpha(v_1, \dots, v_k, \dots, v_i, \dots, v_r).
\end{aligned}$$

„(iii) $\Longrightarrow$ (i)“: Ist offensichtlich.
„(ii) $\Longleftrightarrow$ (iii)“: Ist offensichtlich, da jede Permutation Komposition von endlich vielen Vertauschungen ist.
AD (2): Folgt direkt aus der Linearität von $\alpha$ und der Eigenschaft $(*)$.
AD (3): Ist offensichtlich. $\square$

Definition X.4.4. Seien $\varphi_1, \dots, \varphi_r \in V^*, r \geq 2$. Dann wird durch

$$\varphi_1 \wedge \dots \wedge \varphi_r(v_1, \dots, v_r) = \det\Big( (\langle \varphi_i, v_j \rangle)_{1 \leq i,j \leq r} \Big) \quad \forall v_1, \dots, v_r \in V$$

eine alternierende $r$-Form definiert. $\varphi_1 \wedge \dots \wedge \varphi_r$ heißt das Äussere Produkt von $\varphi_1, \dots, \varphi_r$.

Satz X.4.5. *Sei $n = \dim V < \infty$ und $e_1, \dots, e_n$ eine Basis von $V$. Sei $\varepsilon_1, \dots, \varepsilon_n$ die Dualbasis zu $e_1, \dots, e_n$ und $1 \leq r \leq n$. Dann ist*

$$\{\varepsilon_{j_1} \wedge \dots \wedge \varepsilon_{j_r} : 1 \leq j_1 < j_2 < \dots < j_r \leq n\}$$

*eine Basis von $\Lambda^r(V)$. Insbesondere ist*

$$\dim \Lambda^r(V) = \binom{n}{r}.$$

Beweis. Definiere zur Abkürzung

$$I_r = \{(j) = (j_1, \dots, j_r) : 1 \leq j_1 < \dots < j_r \leq n\}$$
$$\varepsilon_{(j)} = \varepsilon_{j_1} \wedge \dots \wedge \varepsilon_{j_r} \qquad \forall (j) \in I_r.$$

1. Schritt: Seien $\alpha \in \Lambda^r(V)$ und $v_1, \dots, v_r \in V$. Dann ist

$$v_i = \sum_{j=1}^{n} \langle \varepsilon_j, v_i \rangle e_j \quad \forall 1 \leq i \leq r.$$

Wegen der Multilinearität von $\alpha$ und Bemerkung X.4.3 (1) folgt

$$\begin{aligned}
&\alpha(v_1, \dots, v_r) \\
&= \sum_{k_1=1}^{n} \dots \sum_{k_r=1}^{n} \Big\{ \langle \varepsilon_{k_1}, v_1 \rangle \cdot \ldots \cdot \langle \varepsilon_{k_r}, v_r \rangle \cdot \alpha(e_{k_1}, \dots, e_{k_r}) \Big\} \\
&= \sum_{(j) \in I_r} \alpha(e_{j_1}, \dots, e_{j_r}) \sum_{\sigma \in \sigma_r} \operatorname{sgn}(\sigma) \langle \varepsilon_{i_{\sigma(1)}}, v_1 \rangle \cdot \ldots \cdot \langle \varepsilon_{j_{\sigma(r)}}, v_r \rangle \\
&= \sum_{(j) \in I_r} \alpha(e_{j_1}, \dots, e_{j_r}) \det\Big( (\langle \varepsilon_{j_\rho}, v_\tau \rangle)_{1 \leq \rho, \tau \leq r} \Big) \\
&= \sum_{(j) \in I_r} \alpha(e_{j_1}, \dots, e_{j_r}) \varepsilon_{(j)}(v_1, \dots, v_r).
\end{aligned}$$

Also ist

$$\alpha = \sum_{(j) \in I_r} \alpha_{(j)} \varepsilon_{(j)}$$

mit

$$\alpha_{(j)} = \alpha(e_{j_1}, \dots, e_{j_r}).$$

2. Schritt: Sei $\alpha = \sum_{(j) \in I_r} b_{(j)} \varepsilon_{(j)}$ mit $b_{(j)} \in \mathbb{R}$. Wir müssen zeigen:

$$b_{(j)} = \alpha(e_{j_1}, \dots, e_{j_r}).$$

Sei dazu $(k) \in I_r$. Dann folgt

$$\begin{aligned}
&\alpha(e_{k_1}, \ldots, e_{k_r}) \\
&= \sum_{(j)\in I_r} b_{(j)} \varepsilon_{(j)}(e_{k_1}, \ldots, e_{k_r}) \\
&= \sum_{(j)\in I_r} b_{(j)} \det\Big( (\langle \varepsilon_{j_\rho}, e_{k_\tau} \rangle)_{1\le\rho,\tau\le r} \Big) \\
&= \sum_{(j)\in I_r} b_{(j)} \det\Big( (\delta_{j_\rho k_\tau})_{1\le\rho,\tau\le r} \Big) \\
&= b_{(k)}.
\end{aligned}$$

Dies beweist die Behauptung. □

Satz X.4.6. *Sei $n = \dim V < \infty$ und $r, s, t \in \mathbb{N}^*$.*

(1) *Es gibt genau eine Abbildung*

$$\begin{aligned}
\wedge : \Lambda^r(V) \times \Lambda^s(V) &\to \Lambda^{r+s}(V) \\
(\alpha, \beta) &\mapsto \alpha \wedge \beta,
\end{aligned}$$

*genannt* Äusseres Produkt, *mit folgenden Eigenschaften:*

(a) *$\wedge$ ist bilinear.*

(b) *Für $\varphi_1, \ldots, \varphi_r, \psi_1, \ldots, \psi_s \in V^*$ ist*

$$\begin{aligned}
&(\varphi_1 \wedge \ldots \wedge \varphi_r) \wedge (\psi_1 \wedge \ldots \wedge \psi_s) \\
&= \varphi_1 \wedge \ldots \wedge \varphi_r \wedge \psi_1 \wedge \ldots \wedge \psi_s.
\end{aligned}$$

(2) *Ist $e_1, \ldots, e_n$ eine Basis von $V$ und $\varepsilon_1, \ldots, \varepsilon_n$ die zugehörige Dualbasis, so gilt für $\alpha = \sum_{(j)\in I_r} \alpha_{(j)} \varepsilon_{(j)} \in \Lambda^r(V)$ und $\beta = \sum_{(k)\in I_s} \beta_{(k)} \varepsilon_{(k)} \in \Lambda^s(V)$:*

$$\alpha \wedge \beta = \sum_{\substack{(j)\in I_r \\ (k)\in I_s}} \alpha_{(j)} \beta_{(k)} \varepsilon_{(j)} \wedge \varepsilon_{(k)}.$$

(3) *Für $\alpha \in \Lambda^r(V), \beta \in \Lambda^s(V)$ ist*

$$\alpha \wedge \beta = (-1)^{rs} \beta \wedge \alpha.$$

(4) *Für $\alpha \in \Lambda^r(V)$, $\beta \in \Lambda^s(V)$, $\gamma \in \Lambda^t(V)$ ist*

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma).$$

Beweis. ad (1): Wegen Satz X.4.5 wird $\wedge$ durch die Eigenschaften (a) und (b) eindeutig bestimmt.
ad (2): Folgt direkt aus (a) und (b).
ad (3): Folgt aus Bemerkung X.4.3 (2), da die Permutation $\sigma \in \sigma_{r+s}$ mit

$$(1, \ldots, r, r+1, \ldots, r+s) \mapsto (r+1, \ldots, r+s, 1, \ldots, r)$$

das Signum $(-1)^{rs}$ hat.
AD (4): Folgt direkt aus (a) und (b). □

BEMERKUNG X.4.7. (1) Es ist nützlich, das äußere Produkt auch im Fall $r = 0$ oder $s = 0$ zu definieren. Dazu setzt man für $\alpha \in \mathbb{R}$ und $\beta \in \Lambda^k(V)$

$$\alpha \wedge \beta = \beta \wedge \alpha = \alpha\beta.$$

(2) $\Lambda(V) = \bigoplus_{r \geq 0} \Lambda^r(V)$ heißt ÄUSSERE oder GRASSMANNSCHE ALGEBRA von $V$. $\wedge$ wird auf kanonische Weise auf $\Lambda(V)$ fortgesetzt. $\Lambda(V)$ ist ein $\mathbb{R}$-Vektorraum der Dimension $2^n$ mit $n = \dim V$. $(\Lambda(V), \wedge)$ ist eine Algebra mit Eins.

DEFINITION X.4.8. Seien $h \in \mathcal{L}(V, W)$, $\alpha \in \Lambda^r(W)$ und $r \geq 0$.

(1) Für $r = 0$ definieren wir $h * \alpha \in \Lambda^0(V)$ durch

$$h * \alpha = \alpha.$$

(2) Für $r > 0$ definieren wir $h * \alpha \in \Lambda^r(V)$ durch

$$h * \alpha(v_1, \dots, v_r) = \alpha(h(v_1), \dots, h(v_r)) \; \forall v_1, \dots, v_r \in V.$$

$h * \alpha$ heißt der PULL-BACK oder RÜCKTRANSPORT von $\alpha$ mittels $h$.

BEMERKUNG X.4.9. (1) Es ist $h* \in \mathcal{L}(\Lambda^r(W), \Lambda^r(V))$ und es gilt

$$V \xrightarrow{h} W$$

$$\Lambda^r(V) \xleftarrow{h*} \Lambda^r(W).$$

Im Fall $r = 1$ ist $h*$ die zu $h$ adjungierte Abbildung:

$$h * \alpha(v) = \langle \alpha, h(v) \rangle = \langle h * \alpha, v \rangle \quad \forall v \in V, \alpha \in W^*.$$

(2) Es gilt

$$\begin{aligned} id* &= id, \\ (k \circ h)* &= (h*) \circ (k*), \\ h \in \mathrm{Isom}(V, W) &\Longrightarrow h* \in \mathrm{Isom}(\Lambda^r(W), \Lambda^r(V)), \\ (h*)^{-1} &= (h^{-1}) * . \end{aligned}$$

(3) Es gilt

$$h * (\alpha \wedge \beta) = (h * \alpha) \wedge (h * \beta) \qquad \forall h \in \mathcal{L}(V, W), \alpha \in \Lambda^r(W), \beta \in \Lambda^s(W).$$

BEWEIS. Die Behauptung folgt direkt aus der Definition des pullbacks. □

SATZ X.4.10. *Seien* $n = \dim(V) < \infty$, $h \in \mathcal{L}(V, V)$ *und* $\alpha \in \Lambda^n(V)$. *Dann gilt*

$$h * \alpha = \det(h)\alpha.$$

Beweis. Sei $e_1, \dots, e_n$ eine Basis von $V$ und $\varepsilon_1, \dots, \varepsilon_n$ die zugehörige Dualbasis. Sei $(h_{jk})_{1\le j,k\le n}$ die Matrix von $h$ bzgl. $e_1, \dots, e_n$, d.h.,

$$h(e_k) = \sum_{j=1}^{n} h_{jk} e_j \quad \forall 1 \le k \le n.$$

Dann folgt mit Bemerkung X.4.3 (1)

$$\begin{aligned} & h * \alpha(e_1, \dots, e_n) \\ &= \alpha(h(e_1), \dots, h(e_n)) \\ &= \sum_{j_1=1}^{n} \dots \sum_{j_n=1}^{n} h_{j_1 1} \dots h_{j_n n} \alpha(e_{j_1}, \dots, e_{j_n}) \\ &= \sum_{\sigma \in \sigma_n} h_{\sigma(1)1} \dots h_{\sigma(n)n} \operatorname{sgn}(\sigma) \alpha(e_1, \dots, e_n) \\ &= \det(h) \alpha(e_1, \dots, e_n). \end{aligned}$$

Hieraus folgt die Behauptung wegen $\dim \Lambda^n(V) = \binom{n}{n} = 1$. □

## X.5. Differentialformen

Im Folgenden ist stets $n \in \mathbb{N}^*$, $r, s \in \mathbb{N}$, $k \in \mathbb{N} \cup \{\infty\}$ und $U \subset \mathbb{R}^n$ offen und nicht leer. Wir bezeichnen mit $e_1, \dots, e_n$ die kanonische Basis des $\mathbb{R}^n$ und mit $\varepsilon_1, \dots, \varepsilon_n$ die zugehörige Dualbasis.

Definition X.5.1. (1) Eine Abbildung $\alpha \in C^k(U, \Lambda^r(\mathbb{R}^n))$ heißt Differentialform vom Grade $r$ auf $U$ der Klasse $C^k$, kurz $r$-Form.
(2) Wir setzen

$$\begin{aligned} \Omega_k^r(U) &= \{\alpha : \alpha \text{ ist } r\text{-Form der Klasse } C^k \text{ auf } U\}, \\ \Omega^r(U) &= \Omega_0^r(U). \end{aligned}$$

(3) $\wedge : \Omega_k^r(U) \times \Omega_k^s(U) \to \Omega_k^{r+s}(U)$ mit $(\alpha, \beta) \mapsto \alpha \wedge \beta$ und $(\alpha \wedge \beta)(z) = \alpha(z) \wedge \beta(z)$ für alle $z \in U$ heißt Äussere Multiplikation oder Äusseres Produkt.

Bemerkung X.5.2. (1) Definition X.5.1 ist sinnvoll, da gemäß Abschnitt X.4 $\Lambda^r(\mathbb{R}^n)$ ein normierter Vektorraum ist.
(2) Es ist $\Omega_k^0(U) = C^k(U)$ und $\Omega_k^r(U) = \{0\}$ für alle $r > n$.
(3) $\Omega_k^r(U)$ ist ein $\mathbb{R}$-Vektorraum und ein $C^k(U)$ Modul bzgl. der Multiplikation $(f, \alpha) \mapsto f\alpha = f \wedge \alpha$.
(4) Es ist $\alpha \wedge \beta = (-1)^{rs} \beta \wedge \alpha$ für alle $\alpha \in \Omega^r(U)$, $\beta \in \Omega^s(U)$.
(5) $\Omega_k(U) = \bigoplus_{r \ge 0} \Omega_k^r(U)$ heißt die Algebra der $C^k$-Differentialformen, wobei $\wedge$ die Algebra-Multiplikation ist.
(6) Ist $f \in C^{k+1}(U, \mathbb{R})$, so ist $Df \in C^k(U, \mathcal{L}(\mathbb{R}^n, \mathbb{R})) \cong C^k(U, \mathbb{R}^{n*}) = C^k(U, \Lambda^1(\mathbb{R}^n))$. Daher kann $Df$ als 1-Form aufgefasst werden. In diesem Fall schreiben wir zur Unterscheidung $df$ und nennen dies das totale

Differential von $f$. Ist insbesondere $f = pr_j$ die Projektion auf die $j$-te Variable, d.h.,

$$f(z) = z_j \quad \forall z \in U, 1 \le j \le n,$$

so schreiben wir

$$dx_j = dpr_j, \quad 1 \le j \le n.$$

Wegen

$$Dpr_j(z) = v_j \quad \forall z \in U, v \in \mathbb{R}^n, 1 \le j \le n$$

ist

$$dx_j(z) = \varepsilon_j \quad \forall z \in U, 1 \le j \le n,$$

so dass wegen Satz X.4.5 (S. 390) $dx_1, \ldots, dx_n$ eine Basis von $\Omega^1(U)$ bilden. Für $f \in C^{k+1}(U)$ ist dann

$$df = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} dx_j.$$

Physikalisch beschreibt $df$ das totale Inkrement der Funktion $f$. Man vergleiche hierzu Bemerkung VI.3.2 (S. 187).

Satz X.5.3. *Sei* $1 \le r \le n$. $\{dx_{j_1} \wedge \ldots \wedge dx_{j_r} : 1 \le j_1 < \ldots < j_r \le n\}$ *ist eine Basis von* $\Omega^r(U)$, *d.h. jedes* $\alpha \in \Omega^r(U)$ *besitzt eine eindeutige Darstellung*

$$\alpha = \sum_{1 \le j_1 < \ldots < j_r \le n} a_{j_1 \ldots j_r} dx_{j_1} \wedge \ldots \wedge dx_{j_r}$$

*mit* $a_{j_1 \ldots j_r} \in C(U, \mathbb{R})$. *Insbesondere gilt*

$$\alpha \in \Omega^r_k(U) \iff a_{j_1 \ldots j_r} \in C^k(U, \mathbb{R}) \quad \forall 1 \le j_1 < \ldots < j_r \le n.$$

Beweis. Folgt wegen

$$dx_j(z) = \varepsilon_j \quad \forall z \in U, 1 \le j \le n$$

aus Satz X.4.5 (S. 390). □

Beispiel X.5.4. (1) Sei $n = 2$. Dann gilt

$$\begin{aligned}
\alpha \in \Omega^1(U) &\iff \alpha = a_1 dx_1 + a_2 dx_2, \quad a_1, a_2 \in C(U, \mathbb{R}), \\
\alpha \in \Omega^2(U) &\iff \alpha = a_{12} dx_1 \wedge dx_2, \quad a_{12} \in C(U, \mathbb{R}).
\end{aligned}$$

(2) Sei $n = 3$. Dann gilt

$$\begin{aligned}
\alpha \in \Omega^1(U) &\iff \alpha = a_1 dx_1 + a_2 dx_2 + a_3 dx_3, \\
&\qquad a_1, a_2, a_3 \in C(U, \mathbb{R}), \\
\alpha \in \Omega^2(U) &\iff \alpha = a_1 dx_2 \wedge dx_3 + a_2 dx_3 \wedge dx_1 + a_3 dx_1 \wedge dx_2, \\
&\qquad a_1, a_2, a_3 \in C(U, \mathbb{R}),
\end{aligned}$$

$$\alpha \in \Omega^3(U) \iff \alpha = a dx_1 \wedge dx_2 \wedge dx_3, \quad a \in C(U, \mathbb{R}).$$

(3) Sei $n \geq 2$. Dann gilt

$$\alpha \in \Omega^1(U) \iff \alpha = \sum_{i=1}^{n} a_i dx_i, \quad a_i \in C(U, \mathbb{R}),$$

$$\alpha \in \Omega^{n-1}(U) \iff \alpha = \sum_{i=1}^{n} a_i (-1)^{i-1} dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n,$$
$$a_i \in C(U, \mathbb{R}),$$

wobei $dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n$ bedeutet, dass der Term $dx_i$ fehlt. Wir verwenden im Folgenden stets $\{(-1)^{i-1} dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n : 1 \leq i \leq n\}$ als Basis von $\Omega^{n-1}(U)$.
(4) Sei $n \geq 2$ und $f \in C(U, \mathbb{R})$, $u \in C(U, \mathbb{R}^n)$. Dann entsprechen $f$ die 0- und die $n$-Form

$$\alpha_{0,f} = f,$$
$$\alpha_{n,f} = f dx_1 \wedge \ldots \wedge dx_n$$

und $u$ die 1- und die $(n-1)$-Form

$$\alpha_{1,u} = \sum_{i=1}^{n} u_i dx_i,$$
$$\alpha_{n-1,u} = \sum_{i=1}^{n} u_i (-1)^{i-1} dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n.$$

Definition X.5.5. Seien $V \subset \mathbb{R}^m$ offen, $h \in C^{k+1}(U, \mathbb{R}^m)$ mit $h(U) \subset V$ und $\alpha \in \Omega_k^r(V)$. Dann definieren wir den pull-back oder Rücktransport $h * \alpha \in \Omega_k^r(U)$ von $\alpha$ mittels $h$ durch

$$\begin{aligned}
&(h * \alpha)(z)(v_1, \ldots, v_r) \\
&= [Dh(z)] * \alpha(h(z))(v_1, \ldots, v_r) \\
&= \alpha(h(z))(Dh(z)v_1, \ldots, Dh(z)v_r) \quad \forall z \in U, v_1, \ldots, v_r \in \mathbb{R}^n.
\end{aligned}$$

Aus Definition X.5.5, Definition X.4.8 (S. 392), Bemerkung X.4.9 (S. 392) und Satz X.4.10 (S. 392) folgt unmittelbar:

Bemerkung X.5.6. (1) $h * (\alpha \wedge \beta) = (h * \alpha) \wedge (h * \beta)$.
(2) Es gilt

$$h * dx_j = dh_j \quad , 1 \leq j \leq n,$$
$$h * f = f \circ h.$$

(3) Es gilt

$$\begin{aligned}
&h * (dx_{j_1} \wedge \ldots \wedge dx_{j_r}) \\
&= dh_{j_1} \wedge \ldots \wedge dh_{j_r} \quad , 1 \leq j_1 < \ldots < j_r \leq m,
\end{aligned}$$

und

$$\begin{aligned} h*(&\sum_{1\le j_1<\ldots j_r\le m} a_{j_1\ldots j_r} dx_{j_1}\wedge\ldots\wedge dx_{j_r}) \\ =&\sum_{1\le j_1<\ldots j_r\le m} a_{j_1\ldots j_r}\circ h\ \ dh_{j_1}\wedge\ldots\wedge dh_{j_r}. \end{aligned}$$

(4) Falls $n=m$ ist, gilt

$$\begin{aligned} h*(dx_1\wedge\ldots\wedge dx_n) &= dh_1\wedge\ldots\wedge dh_n \\ &= \det(Dh)dx_1\wedge\ldots\wedge dx_n. \end{aligned}$$

(5) Es gilt

$$\begin{aligned} &h\in C^{k+1}(U,\mathbb{R}^m), h(U)\subset V \\ \Longrightarrow &h*\in\mathcal{L}(\Omega_k^r(V),\Omega_k^r(U))\ \text{ falls } r\ge 1, \\ &h*\in\mathcal{L}(C^{k+1}(V),C^{k+1}(U))\ \text{ falls } r=0. \end{aligned}$$

(6) Es ist $(k\circ h)*=(h*)\circ(k*)$ und $id_{\mathbb{R}^n}*=id_{\Omega(U)}$.

Beispiel X.5.7. Sei $h:\mathbb{R}^2\to\mathbb{R}^2$ definiert durch

$$h(r,\varphi)=(r\cos\varphi, r\sin\varphi).$$

Dann ist

$$\begin{aligned} dh_1 &= \cos\varphi dr - r\sin\varphi d\varphi \\ dh_2 &= \sin\varphi dr + r\cos\varphi d\varphi. \end{aligned}$$

Seien $V\subset\mathbb{R}^2$ offen, $U=h^{-1}(V)$ und

$$\begin{aligned} \alpha &= f dx_1 + g dx_2 \in\Omega_k^1(V) \\ \beta &= F dx_1\wedge dx_2 \quad \in\Omega_k^2(V). \end{aligned}$$

Dann folgt

$$\begin{aligned} h*\alpha &= [\cos\varphi\cdot f\circ h+\sin\varphi\cdot g\circ h]dr \\ &\qquad + [r\cos\varphi\cdot g\circ h - r\sin\varphi\cdot f\circ h]d\varphi, \\ h*\beta &= r\cdot F\circ h\ \ dr\wedge d\varphi. \end{aligned}$$

Definition X.5.8. Seien $r\ge 1$ und

$$\alpha=\sum_{1\le j_1<\ldots<j_r\le n} a_{j_1\ldots j_r}dx_{j_1}\wedge\ldots\wedge dx_{j_r}\in\Omega_{k+1}^r(U).$$

Dann heißt

$$d\alpha=\sum_{1\le j_1<\ldots<j_r\le n} da_{j_1\ldots j_r}\wedge dx_{j_1}\wedge\ldots\wedge dx_{j_r}\in\Omega_k^{r+1}(U)$$

die Äussere Ableitung oder das Differential von $\alpha$.

Beispiel X.5.9. Wir verwenden die Bezeichnungen von Beispiel X.5.4. Seien $f \in C^1(U,\mathbb{R})$ und $u \in C^1(U,\mathbb{R}^n)$. Dann gilt

$$\begin{aligned} df &= d\alpha_{0,f} \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i \\ &= \alpha_{1,\operatorname{grad} f} \\ d\alpha_{n-1,u} &= \sum_{i=1}^n (-1)^{i-1} du_i \wedge dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n \\ &= \sum_{i=1}^n (-1)^{i-1} \Big\{ \sum_{j=1}^n \frac{\partial u_i}{\partial x_j} dx_j \wedge dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n \Big\} \\ &= \sum_{i=1}^n (-1)^{i-1} \frac{\partial u_i}{\partial x_i} dx_i \wedge dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n \\ &= \Big\{ \sum_{i=1}^n \frac{\partial u_i}{\partial x_i} \Big\} dx_1 \wedge \ldots \wedge dx_n \\ &= \alpha_{n,\operatorname{div} u}. \end{aligned}$$

Ist speziell $n = 3$, so folgt schließlich

$$\begin{aligned} d\alpha_{1,u} &= d(u_1 dx_1 + u_2 dx_2 + u_3 dx_3) \\ &= du_1 \wedge dx_1 + du_2 \wedge dx_2 + du_3 \wedge dx_3 \\ &= \frac{\partial u_1}{\partial x_2} dx_2 \wedge dx_1 + \frac{\partial u_1}{\partial x_3} dx_3 \wedge dx_1 \\ &\quad + \frac{\partial u_2}{\partial x_1} dx_1 \wedge dx_2 + \frac{\partial u_2}{\partial x_3} dx_3 \wedge dx_2 \\ &\quad + \frac{\partial u_3}{\partial x_1} dx_1 \wedge dx_3 + \frac{\partial u_3}{\partial x_2} dx_2 \wedge dx_3 \\ &= \Big( \frac{\partial u_3}{\partial x_2} - \frac{\partial u_2}{x_3} \Big) dx_2 \wedge dx_3 \\ &\quad + \Big( \frac{\partial u_1}{\partial x_3} - \frac{\partial u_3}{\partial x_1} \Big) dx_3 \wedge dx_1 \\ &\quad + \Big( \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2} \Big) dx_1 \wedge dx_2 \\ &= \alpha_{2,\operatorname{rot} u}. \end{aligned}$$

Satz X.5.10. *Die Abbildung* $d : \Omega^r_{k+1}(U) \to \Omega^{r+1}_k(U)$ *hat folgende Eigenschaften:*

(1) $d \in \mathcal{L}(\Omega^r_{k+1}(U), \Omega^{r+1}_k(U))$.
(2) $d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^r \alpha \wedge (d\beta)$ *für alle* $\alpha \in \Omega^r_1(U)$, $\beta \in \Omega^s_1(U)$ (Produktregel).
(3) $d^2 = d \circ d = 0$.

(4) *Ist* $h \in C^2(U, \mathbb{R}^m)$, *so ist*

$$(h*) \circ d = d \circ (h*).$$

BEWEIS. AD (1): Ist offensichtlich.
AD (2): Seien $(j) = (j_1, \ldots, j_r)$ mit $1 \leq j_1 < \ldots < j_r \leq n$ und $(k) = (k_1, \ldots, k_s)$ mit $1 \leq k_1 < \ldots < k_s \leq n$ und

$$\begin{aligned} \alpha &= adx_{(j)} = adx_{j_1} \wedge \ldots \wedge dx_{j_r} \\ \beta &= bdx_{(k)} = bdx_{k_1} \wedge \ldots \wedge dx_{k_s}. \end{aligned}$$

FALL 1: $(j)$ und $(k)$ haben einen Index gemeinsam. Aus Satz X.4.6 (S. 391) folgt

$$\alpha \wedge \beta = abdx_{(j)} \wedge dx_{(k)} = 0$$

und

$$\begin{aligned} (d\alpha) \wedge \beta &= bda \wedge dx_{(j)} \wedge dx_{(k)} \\ &= 0 \\ \alpha \wedge (d\beta) &= adx_{(j)} \wedge db \wedge dx_{(k)} \\ &= 0. \end{aligned}$$

FALL 2: $(j)$ und $(k)$ sind disjunkt.
Aus Bemerkung X.5.2 und Satz X.5.3 folgt dann

$$\begin{aligned} d(\alpha \wedge \beta) &= d(abdx_{(j)} \wedge dx_{(k)}) \\ &= d(ab) \wedge dx_{(j)} \wedge dx_{(k)} \\ &= [b(da) + a(db)] \wedge dx_{(j)} \wedge dx_{(k)} \\ &= bda \wedge dx_{(j)} \wedge dx_{(k)} \\ &\quad + (-1)^r adx_{(j)} \wedge db \wedge dx_{(k)} \\ &= (d\alpha) \wedge \beta + (-1)^r \alpha \wedge (d\beta). \end{aligned}$$

Hieraus folgt die Behauptung zusammen mit der Linearität von $d$.
AD (3): Sei $f \in C^2(U, \mathbb{R})$. Aus Satz VII.4.18(2) (S. 237) folgt

$$\begin{aligned} d^2 f &= d\Big(\sum_{i=1}^{n} D_i f dx_i\Big) \\ &= \sum_{i=1}^{n} d(D_i f) \wedge dx_i \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} D_j D_i f dx_j \wedge dx_i \\ &= \sum_{1 \leq i < j \leq n} (D_i D_j f - D_j D_i f) dx_i \wedge dx_j \\ &= 0. \end{aligned}$$

Sei nun $1 \leq j_1 < \ldots < j_r \leq n$ und $a \in C^2(U, \mathbb{R})$. Aus Obigem und Teil (2) folgt

$$\begin{aligned} & d^2(adx_{j_1} \wedge \ldots \wedge dx_{j_r}) \\ &= d(da \wedge dx_{j_1} \wedge \ldots \wedge dx_{j_r}) \\ &= d^2 a \wedge dx_{j_1} \wedge \ldots \wedge dx_{j_r} \\ &\quad - da \wedge (d1) \wedge dx_{j_1} \wedge \ldots \wedge dx_{j_r} \\ &= 0. \end{aligned}$$

Hieraus folgt die Behauptung mit der Linearität von $d$.
AD (4): Wir beweisen die Behauptung durch Induktion über den Grad $r$ der Differentialformen.
„$r = 0$ :“ Seien $f \in C^1(U, \mathbb{R})$ und $h \in C^2(U, \mathbb{R}^m)$ mit $h(U) \subset V$. Dann ist nach Satz VII.3.3 (S. 226)

$$\begin{aligned} d(h * f) &= d(f \circ h) \\ &= \sum_{i=1}^{n} D_i(f \circ h) dx_i \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} (D_j f) \circ h \cdot D_i h_j dx_i \\ &= \sum_{j=1}^{m} (D_j f) \circ h dh_j \\ &= h * (df). \end{aligned}$$

„$r \to r + 1$ :“ Wegen der Linearität von $d$ brauchen wir nur den Fall $\alpha = \beta \wedge dx_j$ mit $\beta \in \Omega_1^r(V)$, $1 \leq j \leq m$, und $h \in C^2(U, \mathbb{R}^m)$ mit $h(U) \subset V$ zu betrachten. Aus den Teilen (2) und (3) und der Induktionsvoraussetzung folgt dann

$$\begin{aligned} h * (d(\beta \wedge dx_j)) &= h * (d\beta \wedge dx_j + (-1)^r \beta \wedge d^2 x_j) \\ &= h * (d\beta \wedge dx_j) \\ &= (h * d\beta) \wedge h * dx_j \\ &= d(h * \beta) \wedge dh_j \\ &= d(h * \beta) \wedge dh_j + (-1)^r (h * \beta) \wedge d^2 h_j \\ &= d((h * \beta) \wedge dh_j) \\ &= d(h * (\beta \wedge dx_j)). \end{aligned}$$

□

BEMERKUNG X.5.11. (1) $d \circ (h*) = (h*) \circ d$ bedeutet, dass $d$ von der speziellen Koordinatendarstellung unabhängig ist.
(2) Für $\alpha \in \Omega_{k+1}(U)$ kann auch $d\alpha \in \Omega_{k+1}(U)$ gelten.

Definition X.5.12. $\alpha \in \Omega^r(U)$ heißt geschlossen, wenn $d\alpha = 0$ ist. $\alpha \in \Omega^r(U)$ heißt exakt, wenn es ein $\beta \in \Omega^{r-1}(U)$ gibt mit $\alpha = d\beta$.

Bemerkung X.5.13. (1) Sei $u \in C(U, \mathbb{R}^n)$. Dann ist $u$ genau dann ein Gradientenfeld, wenn $\alpha_{1,u}$ exakt ist.
(2) Sei $u \in C^1(U, \mathbb{R}^n)$. Dann ist $\alpha_{1,u}$ genau dann geschlossen, wenn $Du$ symmetrisch ist.
(3) Jede $n$-Form ist wegen $\Omega^{n+1}(U) = \{0\}$ geschlossen.
(4) Jede exakte Form ist wegen $d^2 = 0$ geschlossen.
(5) Die Differentialform

$$\alpha = -\frac{y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy \in \Omega^1_\infty(\mathbb{R}^2\backslash\{0\})$$

ist geschlossen, aber nicht exakt. Denn andernfalls wäre

$$f(x,y) = (-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}) \in C^\infty(\mathbb{R}^2\backslash\{0\}, \mathbb{R}^2)$$

ein Gradientenfeld im Widerspruch zu Beispiel VIII.3.4 (S. 271).

Wir wollen einen Zusammenhang zwischen geschlossenen und exakten Differentialformen herstellen. Dazu benötigen wir folgenden Begriff.

Definition X.5.14. $U$ heißt sternförmig, wenn es ein $x_0 \in U$ gibt, derart dass für jedes $x \in U$ die Strecke $x_0 t + (1-t)x$, $t \in [0,1]$, ganz in $U$ verläuft.

Bemerkung X.5.15. (1) Ist $U$ konvex, so ist $U$ sternförmig.
(2) $U = [-1,1]^2\backslash[0,1]^2$ ist sternförmig, aber nicht konvex (vgl. Abb. X.5.1).
(3) Ist $U$ sternförmig, so ist $U$ einfach zusammenhängend.
(4) $\mathbb{R}^3\backslash\{0\}$ ist einfach zusammenhängend, aber nicht sternförmig.

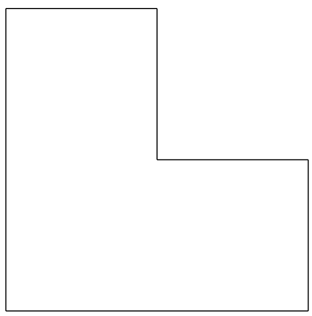

Abbildung X.5.1. Beispiel für ein sternförmiges, nicht konvexes Gebiet

Der folgende Satz ist eine Verallgemeinerung von Satz VIII.3.16 (S. 281).

Satz X.5.16 (Satz von Poincaré). *Ist $U$ sternförmig, so ist jede geschlossene Differentialform auf $U$ exakt.*

Beweis. O.E. können wir annehmen, dass der Punkt $x_0$ aus Definition X.5.14 der Nullpunkt ist. Andernfalls führen wir die Koordinatentransformation $x \mapsto x - x_0$ aus. Sei $\alpha \in \Omega^r(U)$ geschlossen. Definiere $\varphi \in C^\infty(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n)$ durch

$$\varphi(t, x) = tx.$$

Dann ist

$$[0, 1] \times U \subset V = \varphi^{-1}(U).$$

Definiere

$$\beta = \varphi * \alpha \in \Omega^r(V).$$

Wegen Satz X.5.10 ist $\beta$ geschlossen. Definiere $\psi_0, \psi_1 \in C^\infty(U, \mathbb{R} \times \mathbb{R}^n)$ durch

$$\psi_0(x) = (0, x) \quad , \quad \psi_1(x) = (1, x) \quad \forall x \in U.$$

Dann ist $\psi_0(U) \subset V$, $\psi_1(U) \subset V$. Daher ist

$$\gamma = \psi_1 * \beta - \psi_0 * \beta \in \Omega^r(U).$$

Wir setzen

$$\begin{aligned} I_s &= \{(j_1, \ldots, j_s) = 1 \le j_1 < \ldots < j_s \le n\}, 1 \le s \le n, \\ dx_{(j)} &= dx_{j_1} \wedge \ldots \wedge dx_{j_s}, (j) \in I_s \end{aligned}$$

und stellen $\beta$ in der Form

$$\beta = \sum_{(j) \in I_r} f_{(j)} dx_{(j)} + \sum_{(k) \in I_{r-1}} g_{(k)} dt \wedge dx_{(k)}$$

dar mit Funktion $f_{(j)}, g_{(k)} \in C^1(V, \mathbb{R})$. Es folgt

$$\begin{aligned} \psi_1 * \beta &= \sum_{(j) \in I_r} f_{(j)}(1, x) dx_{(j)} \\ \psi_0 * \beta &= \sum_{(j) \in I_r} f_{(j)}(0, x) dx_{(j)} \\ d\beta &= \sum_{(j) \in I_r} D_t f_{(j)} dt \wedge dx_{(j)} \\ &\quad + \sum_{(j) \in I_r} \sum_{i=1}^n D_i f_{(j)} dx_i \wedge dx_{(j)} \\ &\quad - \sum_{(k) \in I_{r-1}} \sum_{i=1}^n D_i g_{(k)} dt \wedge dx_i \wedge dx_{(k)}. \end{aligned}$$

Da $d\beta = 0$ ist, folgt

$$(*) \qquad \sum_{(j) \in I_r} D_t f_{(j)} dx_{(j)} = \sum_{(k) \in I_{r-1}} \sum_{i=1}^n D_i g_{(k)} dx_i \wedge dx_{(k)}.$$

Indem wir beide Seiten von $(*)$ bzgl. $t$ von 0 bis 1 integrieren, erhalten wir

$$\begin{aligned}
\gamma &= \psi_1 * \beta - \psi_0 * \beta \\
&= \sum_{(j)\in I_r} [f_{(j)}(1,x) - f_{(j)}(0,x)]dx_{(j)} \\
&= \sum_{(j)\in I_r} \Big\{\int_0^1 D_t f_{(j)}(t,x)dt\Big\}dx_{(j)} \\
&= \sum_{(k)\in I_{r-1}} \sum_{i=1}^{n} \Big\{\int_0^1 D_i g_{(k)}(t,x)dt\Big\}dx_i \wedge dx_{(k)} \\
&= \sum_{(k)\in I_{r-1}} \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\Big\{\int_0^1 g_{(k)}(t,x)dt\Big\}dx_i \wedge dx_{(k)} \\
&= d\eta
\end{aligned}$$

mit

$$\eta = \sum_{(k)\in I_{r-1}} \Big\{\int_0^1 g_{(k)}(t,x)dt\Big\}dx_{(k)}.$$

Andererseits ist

$$\varphi \circ \psi_1 = id_U \quad , \quad \varphi \circ \psi_0 = 0$$

und somit gemäß Bemerkung X.5.6

$$\begin{aligned}
\psi_1 * \beta &= \psi_1 * (\varphi * \alpha) \\
&= (\varphi \circ \psi_1) * \alpha \\
&= \alpha \\
\psi_0 * \beta &= \psi_0 * (\varphi * \alpha) \\
&= (\varphi \circ \psi_0) * \alpha \\
&= 0.
\end{aligned}$$

Also ist

$$\alpha = d\eta.$$

□

Zum Abschluss interpretieren wir Satz X.5.16 in der Sprache der klassischen Vektoranalysis.

Satz X.5.17. (1) *Seien $U \subset \mathbb{R}^3$ und $f \in C^2(U,\mathbb{R})$, $u \in C^2(U,\mathbb{R}^3)$. Dann gilt*

$$\begin{aligned}
\operatorname{rot}(\operatorname{grad} f) &= 0, \\
\operatorname{div}(\operatorname{rot} u) &= 0.
\end{aligned}$$

(2) *Sei* $U \subset \mathbb{R}^3$ *sternförmig und* $u \in C^1(U, \mathbb{R}^3)$*. Dann gilt*

$$\begin{aligned} \operatorname{rot} u = 0 &\iff u = \operatorname{grad} f && \textit{für ein } f \in C^2(U, \mathbb{R}), \\ \operatorname{div} u = 0 &\iff u = \operatorname{rot} v && \textit{für ein } v \in C^2(U, \mathbb{R}^3). \end{aligned}$$

Beweis. ad (1): Folgt aus Beispiel X.5.9 und $d^2 = 0$.
ad (2): Folgt aus Beispiel X.5.4, Beispiel X.5.9 und Satz X.5.16. □

## X.6. Integration von Differentialformen

Im Folgenden sei stets $n \geq 2$ und $1 \leq k < n$. Zunächst definieren wir das Integral von $n$-Formen auf offenen Mengen im $\mathbb{R}^n$. Wegen $\Omega^n(\mathbb{R}^n) \cong C(\mathbb{R}^n, \mathbb{R})$ können wir dabei auf die Integration von Funktionen im $\mathbb{R}^n$ zurückgreifen.

Definition X.6.1. Seien $U \subset \mathbb{R}^n$ offen, $M \subset U$ und $\alpha = f dx_1 \wedge \ldots \wedge dx_n \in \Omega^n(U)$. $\alpha$ heißt integrierbar über $M$, wenn $f \in L^1(M, \mathbb{R})$ ist. In diesem Fall definieren wir

$$\int_M \alpha = \int_M f dx.$$

Der folgende Satz ist eine Übertragung des Transformationssatzes IX.5.6 (S. 343). Wir erinnern für die Bezeichnungen an Definition X.2.5 (S. 366).

Satz X.6.2 (Transformationssatz). *Seien* $U, V \subset \mathbb{R}^n$ *offen,* $\varphi : U \to V$ *ein* $C^1$ *-Diffeomorphismus,* $M \subset U$ *und* $\alpha \in \Omega^n(V)$*. Dann gilt*

$$\begin{aligned} \int_{\varphi(M)} \alpha &= \int_M \varphi * \alpha && \textit{, falls } \varphi \textit{ orientierungstreu,} \\ \int_{\varphi(M)} \alpha &= -\int_M \varphi * \alpha && \textit{, falls } \varphi \textit{ orientierungsumkehrend.} \end{aligned}$$

Beweis. Sei $\alpha = f dx_1 \wedge \ldots \wedge dx_n$. Aus Bemerkung X.5.6(4) (S. 395) folgt

$$\varphi * \alpha = f \circ \varphi \det D\varphi dx_1 \wedge \ldots \wedge dx_n.$$

Damit ergibt sich die Behauptung direkt aus Satz IX.5.6 (S. 343). □

Wir wollen als nächstes das Integral von $k$-Formen über $k$-dimensionale Mgfkten definieren. Analog zum Vorgehen in Paragraph X.3 soll dies mit Hilfe von Karten und Partitionen der Eins auf die Integration von $k$-Formen im $\mathbb{R}^k$ zurückgeführt werden. Dazu benötigen wir das folgende Analogon zu Satz X.3.3 (S. 372).

Satz X.6.3. (1) *Seien* $U, V \subset \mathbb{R}^k$ *offen,* $\varphi \in C^1(U, \mathbb{R}^n)$ *eine Immersion,* $\tau : V \to U$ *ein orientierungstreuer* $C^1$*-Diffeomorphismus,*

*$\psi = \varphi \circ \tau$ und $\alpha \in \Omega^k(\varphi(U))$. Dann ist $\varphi * \alpha \in \Omega^k(U)$ genau dann integrierbar, wenn $\psi * \alpha \in \Omega^k(V)$ integrierbar ist. In diesem Fall gilt*

$$\int_U \varphi * \alpha = \int_V \psi * \alpha.$$

(2) *Seien $U \subset \mathbb{R}^n$ offen, $M \subset U$ eine orientierbare $k$-dimensionale Mfgkt, $\alpha \in \Omega^k(U)$ mit* $\operatorname{supp}(\alpha) \Subset U$, *$\mathcal{A} = \{(U_i, \varphi_i, V_i) : 1 \le i \le m\}$ und $\mathcal{A}' = \{(U'_j, \varphi'_j, V'_j) : 1 \le j \le m'\}$ zwei positiv orientierte Atlanten von $M \cap \operatorname{supp}(\alpha)$ und $\{\psi_i : 1 \le i \le m\}$, $\{\psi'_j : 1 \le j \le m'\}$ zwei Partitionen der Eins auf $M \cap \operatorname{supp}(\alpha)$, die $\mathcal{A}$ bzw. $\mathcal{A}'$ untergeordnet sind. Dann ist für jedes $1 \le i \le m$ die Differentialform $\varphi_i * (\psi_i \alpha) \in \Omega^k(U_i)$ integrierbar, genau dann, wenn für jedes $1 \le j \le m'$ die Differentialform $\varphi'_j * (\psi'_j \alpha) \in \Omega^k(U'_j)$ integrierbar ist. In diesem Fall gilt*

$$\sum_{i=1}^{m} \int_{U_i} \varphi_i * (\psi_i \alpha) = \sum_{j=1}^{m'} \int_{U'_j} \varphi'_j * (\psi_j \alpha).$$

Beweis. ad (1): Folgt aus Satz X.6.2 und $\psi * \alpha = \tau * (\varphi * \alpha)$. ad (2): Folgt aus Teil (1) mit den gleichen Argumenten wir im Beweis von Teil (2) des Satzes X.3.3 (S. 372). □

Wegen Satz X.6.3 ist folgende Definition sinnvoll.

Definition X.6.4. Seien $U \subset \mathbb{R}^n$ offen, $M \subset U$ eine orientierbare $k$-dimensionale Mfgkt und $\alpha \in \Omega^k(U)$ mit $\operatorname{supp}(\alpha) \Subset U$. Dann heißt $\alpha$ auf $M$ integrierbar, wenn es einen positiv orientierten Atlas $\mathcal{A} = \{(U_i, \varphi_i, V_i) : 1 \le i \le m\}$ von $\operatorname{supp}(\alpha) \cap M$ und eine $\mathcal{A}$ untergeordnete Partition der Eins $\{\psi_i : 1 \le i \le m\}$ auf $M$ gibt, derart dass für jedes $1 \le i \le m$ die Differentialform $\varphi_i * (\psi_i \alpha)$ integrierbar ist. In diesem Fall heißt

$$\int_M \alpha = \sum_{i=1}^{m} \int_{U_i} \varphi_i * (\psi_i \alpha)$$

das Integral von $\alpha$ über $M$.

Bemerkung X.6.5. (1) Streng genommen müsste man bei $\int_M \alpha$ zusätzlich die gewählte Orientierung angeben. Ist $-\mathcal{O}$ die gemäß Bemerkung X.2.8 (S. 367) zur Orientierung von $M$ entgegengesetzte Orientierung, so gilt wegen Satz X.6.2

$$\int_{(M,-\mathcal{O})} \alpha = -\int_{(M,\mathcal{O})} \alpha.$$

(2) Seien $[a,b] \subset \mathbb{R}$ und $\gamma \in C^1([a,b], \mathbb{R}^n)$ ein regulärer $C^1$-Weg sowie

$$\alpha = \sum_{i=1}^{n} f_i dx_i \in \Omega^1(U)$$

mit $\gamma([a,b]) \subset U$. Dann ist $M = \gamma([a,b])$ eine eindimensionale Mfgkt und

$$\begin{aligned}\int_M \alpha &= \int_{[a,b]} \gamma * \alpha \\ &= \int_a^b \sum_{i=1}^n f_i \circ \gamma \gamma_i' dt \\ &= \int_a^b (f \circ \gamma, \gamma') dt \\ &= \int_\gamma f.\end{aligned}$$

D.h., die Kurvenintegrale aus Kapitel VIII sind ein Spezialfall von Definition X.6.4.

Beispiel X.6.6. Sei $M = \{x \in S^2 : x_i > 0, 1 \leq i \leq 3\}$ und

$$\alpha = x_1 dx_1 \wedge dx_2 \in \Omega^2(\mathbb{R}^3).$$

$M$ wird durch die Karte $(U, \psi, M)$ dargestellt mit

$$\begin{aligned}U &= (0, \frac{\pi}{2}) \times (0, \frac{\pi}{2}) \\ \psi(\varphi, \theta) &= (\cos\varphi\cos\theta, \sin\varphi\cos\theta, \sin\theta).\end{aligned}$$

Es ist

$$\begin{aligned}\psi * \alpha &= \cos\varphi\cos\theta[-\sin\varphi\cos\theta d\varphi - \cos\varphi\sin\theta d\theta] \\ &\quad \wedge [\cos\varphi\cos\theta d\varphi - \sin\varphi\sin\theta d\theta] \\ &= \cos\varphi\cos\theta[\sin^2\varphi\cos\theta\sin\theta + \cos^2\varphi\cos\theta\sin\theta] d\varphi \wedge d\theta \\ &= \cos\varphi\cos^2\theta\sin\theta d\varphi \wedge d\theta\end{aligned}$$

und somit

$$\begin{aligned}\int_M \alpha &= \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} \cos\varphi\cos^2\theta\sin\theta d\varphi d\theta \\ &= [\sin\varphi]_0^{\frac{\pi}{2}} [-\frac{1}{3}\cos^3\theta]_0^{\frac{\pi}{2}} \\ &= \frac{1}{3}.\end{aligned}$$

Der folgende Satz stellt einen Zusammenhang her zwischen der Integration von Funktionen und von Differentialformen auf Mfgkten.

Satz X.6.7. *Seien $U \subset \mathbb{R}^n$ offen, $M \subset U$ eine Hyperfläche, die durch das Einheitsnormalenfeld $\nu$ orientiert sei, und $f \in C(U, \mathbb{R}^n)$. Dann gilt für jede kompakte Menge $K \subset M$*

$$\int_K \alpha_{n-1,f} = \int_K \sum_{i=1}^n f_i (-1)^{i-1} dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n$$

$$= \int_K (f, \nu) dS.$$

Beweis. Indem wir $\alpha_{n-1,f}$ gegebenenfalls mit einer hinreichend feinen Partition der Eins multiplizieren, können wir o.E. annehmen, dass $K \cap \operatorname{supp}(\alpha_{n-1,f})$ in einer offenen Menge enthalten ist, in der $M$ Graph einer Funktion von $n-1$ Variablen ist. Indem wir evtl. noch eine Koordinatentransformation durchführen, können wir o.E. annehmen, dass gilt:

$$U = U' \times I \quad \text{mit } U' \subset \mathbb{R}^{n-1} \text{ Gebiet, } I \text{ offenes Intervall,}$$
$$M = \{(x', x_n) \in \mathbb{R}^n : x_n = g(x')\} \text{ mit } g \in C^1(U', \mathbb{R}).$$

Dann wird $M$ durch die Karte $(U', \varphi, M)$ mit

$$\varphi(x') = (x', g(x'))$$

dargestellt. Wie man nicht nachrechnet, ist

$$\tilde{\nu}(x) = \frac{1}{\sqrt{1 + \|\nabla g(x')\|^2}} (-\nabla g(x'), 1) \quad \forall x = (x', x_n) \in M$$

ein Einheitsnormalenfeld von $M$. Also gilt

$$\nu = \varepsilon \tilde{\nu} \quad \text{mit } \varepsilon = \pm 1.$$

Weiter ist

$$\begin{aligned}
&\det(\nu, D_1\varphi, \ldots, D_{n-1}\varphi) \\
&= \frac{\varepsilon}{\sqrt{1 + \|\nabla g(x')\|^2}} \det \begin{pmatrix} -\nabla g(x')^t & \mathbb{I}_{n-1} \\ 1 & \nabla g(x') \end{pmatrix} \\
&= \varepsilon (-1)^{n-1} \sqrt{1 + \|\nabla g(x')\|^2}.
\end{aligned}$$

Also ist die Karte $(U', \varphi, M)$ positiv bzw. negativ orientiert, je nachdem ob $\varepsilon(-1)^{n-1}$ positiv oder negativ ist. Daher ist

$$\int_K \alpha_{n-1,f} = \varepsilon (-1)^{n-1} \int_{\varphi^{-1}(K)} \varphi * \alpha_{n-1,f}.$$

Weiter ist

$$\begin{aligned}
\varphi * \alpha_{n-1,f} &= \sum_{i=1}^{n} f_i \circ \varphi (-1)^{i-1} \varphi * (dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n) \\
&= \sum_{i=1}^{n-1} f_i \circ \varphi (-1)^{i-1} dx'_1 \wedge \ldots \wedge \widehat{dx'_i} \wedge \ldots \wedge dg(x') \\
&\quad + f_n \circ \varphi (-1)^{n-1} dx'_1 \wedge \ldots \wedge dx'_{n-1} \\
&= (-1)^{n-1} \Big\{ - \sum_{i=1}^{n-1} f_i \circ \varphi D_i g + f_n \circ \varphi \Big\} dx'_1 \wedge \ldots \wedge dx'_{n-1}
\end{aligned}$$

und somit

$$\int_K \alpha_{n-1,f} = \varepsilon \int_{\varphi^{-1}(K)} F(x')dx'$$

mit

$$F = -\sum_{i=1}^{n-1} f_i \circ \varphi D_i g + f_n \circ \varphi.$$

Andererseits folgt aus Beispiel X.3.7(3) (S. 374)

$$\begin{aligned}\int_K (f,\nu)dS &= \int_{\varphi^{-1}(K)} (f\circ\varphi, \nu\circ\varphi)\sqrt{1+\|\nabla g\|^2}dx' \\ &= \varepsilon \int_{\varphi^{-1}(K)} (f\circ\varphi, \tilde{\nu}\circ\varphi)\sqrt{1+\|\nabla g\|^2}dx' \\ &= \varepsilon \int_{\varphi^{-1}(K)} F(x')dx'.\end{aligned}$$

Hieraus folgt die Behauptung. □

Als nächstes formulieren wir den Gaußschen Integralsatz in der Sprache der Differentialformen.

Satz X.6.8 (Gaussscher Integralsatz). *Seien $U \subset \mathbb{R}^n$ offen, $K \subset U$ kompakt mit glattem Rand und $\alpha \in \Omega_1^{n-1}(U)$. Dann ist*

$$\int_K d\alpha = \int_{\partial K} \alpha,$$

*wobei $\partial K$ durch das äußere Normalenfeld orientiert ist.*

Beweis. Da $K$ kompakt und $\partial_S K = \emptyset$ ist, ist $\partial_R K = \partial K$ kompakt und damit $\sigma_{n-1}(\partial_R K) < \infty$. Sei $f \in C^1(U, \mathbb{R}^n)$ derart, dass $\alpha = \alpha_{n-1,f}$ ist. Dann folgt aus Satz X.3.19 (S. 384) und Satz X.6.7

$$\begin{aligned}\int_K d\alpha &= \int_K \operatorname{div} f dx \\ &= \int_{\partial K} (f,\nu)dS(x) \\ &= \int_{\partial K} \alpha.\end{aligned}$$

□

Beispiel X.6.9. (1) Sei

$$\alpha = \sum_{i=1}^n x_i(-1)^{i-1} dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n.$$

Wegen

$$d\alpha = n dx_1 \wedge \ldots \wedge dx_n$$

folgt für jedes Kompaktum $K \subset \mathbb{R}^n$ mit glattem Rand

$$\lambda_n(K) = \frac{1}{n} \int_{\partial K} \sum_{i=1}^{n} (-1)^{i-1} x_i dx_1 \wedge \ldots \wedge \widehat{dx_1} \wedge \ldots \wedge dx_n.$$

Speziell ergibt sich in $\mathbb{R}^2$ die Leibnizsche Sektorformel

$$\lambda_2(K) = \frac{1}{2} \int_{\partial K} \{xdy - ydx\}.$$

Zur geometrischen Interpretation betrachte man das Dreieck $\Delta$ mit den Eckpunkten $(0,0), (x,y), (x+\delta x, y+\delta y)$, wobei $\delta x$, $\delta y$ klein seien. Dann ist

$$\begin{aligned} \lambda_2(\Delta) &= \frac{1}{2} \det \begin{pmatrix} x & x+\delta x \\ y & y+\delta y \end{pmatrix} \\ &= \frac{1}{2}(x\delta y - y\delta x). \end{aligned}$$

Die Fläche $K$ kann man sich approximativ durch solche Dreiecke zusammengesetzt denken.
(2) Seien $U \subset \mathbb{R}^2$ offen, $f, g \in C^1(U, \mathbb{R})$ und $K \subset U$ kompakt mit glattem Rand. Dann folgt aus Satz X.6.8 die Green-Riemannsche Formel

$$\int_{\partial K} \{fdx + gdy\} = \int_K \{\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\} dxdy.$$

Satz X.6.8 hat folgende praktische Konsequenz:

Satz X.6.10. *Seien $U \subset \mathbb{R}^n$ offen, $x^* \in U$ und $\alpha \in \Omega_1^{n-1}(U \backslash \{x^*\})$ geschlossen. Weiter seien $K_1, K_2 \subset U$ zwei Kompakta mit glattem Rand und $x^* \in \mathring{K}_1 \cap \mathring{K}_2$. Dann ist*

$$\int_{\partial K_1} \alpha = \int_{\partial K_2} \alpha,$$

*wobei $\partial K_1, \partial K_2$ bzgl. der äußeren Normalen orientiert sind.*

Beweis. Sei $\varepsilon > 0$ so, dass

$$B_\varepsilon = \overline{B(x^*, \varepsilon)} \subset \mathring{K}_1 \cap \mathring{K}_2$$

ist. Setze

$$K_{i,\varepsilon} = K_i \backslash \mathring{B}_\varepsilon \quad , i = 1, 2.$$

Dann sind $K_{1,\varepsilon}, K_{2,\varepsilon}$ Kompakta mit glattem Rand, die in $U \backslash \{x^*\}$ enthalten sind. Wegen $d\alpha = 0$ folgt aus Satz X.6.8

$$\int_{\partial K_{1,\varepsilon}} \alpha = 0 = \int_{\partial K_{2,\varepsilon}} \alpha.$$

Da der Rand von $K_{i,\varepsilon}$, $i = 1, 2$, aus dem positiv orientierten Rand von $K_i$ und dem negativ orientierten Rand von $B_\varepsilon$ besteht, gilt

$$\int_{\partial K_{i,\varepsilon}} \alpha = \int_{\partial K_i} \alpha - \int_{\partial B_\varepsilon} \alpha \quad , i = 1, 2.$$

Hieraus folgt die Behauptung. □

Im Folgenden wollen wir den Integralsatz von Stokes beweisen. Er ist eine Verallgemeinerung von Satz X.6.8, indem $U$ durch eine Mfgkt $M$ ersetzt wird. Zu seiner Formulierung und seinem Beweis benötigen wir einige Vorbereitungen. Wir bezeichnen mit

$$H_k = \{x \in \mathbb{R}^k : x_1 \leq 0\}$$

den Standard-Halbraum in $\mathbb{R}^k$. Das äußere Einheitsnormalenfeld zu $\partial H_k$ ist

$$\nu = e_1 = (1, 0, \dots, 0).$$

$\partial H_k$ besitzt die globale Karte $(\mathbb{R}^{k-1}, \beta, \partial H_k)$ mit

$$\beta(t_1, \dots, t_{k-1}) = (0, t_1, \dots, t_{k-1}).$$

$\partial H_k$ ist durch diese Karte orientiert. Dann ist das äußere Einheitsnormalenfeld $\nu = e_1$ positiv orientiert.

Satz X.6.11. *Sei $\alpha \in \Omega_1^{k-1}(\mathbb{R}^k)$ mit kompaktem Träger. Dann ist*

$$\int_{H_k} d\alpha = \int_{\partial H_k} \alpha.$$

Beweis. Im Fall $k = 1$ ist $H_1 = \mathbb{R}_- = (-\infty, 0]$ und $\partial H_1 = \{0\}$. Die Aussage des Satzes besagt dann

$$\int_{-\infty}^{0} df = f(0) \quad \forall f \in C_0^1(\mathbb{R}, \mathbb{R})$$

und folgt somit aus dem Hauptsatz der Differential- und Integralrechnung.

Sei also $k \geq 2$ und

$$\alpha = \sum_{i=1}^{k} f_i (-1)^{i-1} dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_k$$

mit

$$f_i \in C_0^1(\mathbb{R}^k, \mathbb{R}), \quad 1 \leq i \leq k.$$

Mit der oben definierten globalen Karte $\beta$ von $\partial H_k$ folgt

$$\begin{aligned}
\int_{\partial H_k} \alpha &= \int_{\mathbb{R}^{k-1}} \beta * \alpha \\
&= \int_{\mathbb{R}^{k-1}} \sum_{i=1}^{k} f_i(0, t_1, \dots, t_{k-1})(-1)^{i-1} \\
&\qquad\qquad \beta * (dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_k) \\
&= \int_{\mathbb{R}^{k-1}} f_1(0, t_1 \ldots t_{k-1}) dt_1 \ldots dt_{k-1}.
\end{aligned}$$

Andererseits folgt aus $f_i \in C_0^1(\mathbb{R}^k, \mathbb{R})$ und Satz IX.3.3 (S. 317)

$$\begin{aligned}\int_{H_k} d\alpha &= \int_{H_k} \sum_{i=1}^{n} D_i f_i \\ &= \int_{\mathbb{R}^{k-1}} \underbrace{\Big(\int_{\mathbb{R}_-} D_1 f_1(x_1, \dots, x_k) dx_1\Big)}_{=f_1(0,x_2,\dots,x_k)} dx_2 \dots dx_k \\ &\quad + \sum_{i=2}^{k} \int_{\mathbb{R}_- \times \mathbb{R}^{k-2}} \underbrace{\Big(\int_{\mathbb{R}} D_i f_i(x_1, \dots, x_k) dx_i\Big)}_{=0} dx_1 \dots \widehat{dx_i} \dots dx_k \\ &= \int_{\mathbb{R}^{k-1}} f_1(0, x_2, \dots, x_k) dx_2 \dots dx_k.\end{aligned}$$

Hieraus folgt die Behauptung. □

Lemma X.6.12. *Seien $U, U' \subset \mathbb{R}^k$ offen und $\varphi : U \to U'$ ein orientierungstreuer $C^1$-Diffeomorphismus mit $\varphi(H_k \cap U) = H_k \cap U'$. Dann gilt*

$$\det\Big((D_i\varphi_j(x))_{2 \le i,j \le k}\Big) > 0 \quad \forall x \in \partial H_k \cap U.$$

Beweis. Da $\varphi$ und $\varphi^{-1}$ stetig sind, ist

$$\varphi(\partial H_k \cap U) = \partial H_k \cap U'.$$

Also gilt für alle $x = (0, x') \in \partial H_k \cap U$, $x' \in \mathbb{R}^{k-1}$,

$$\varphi_1(x) = \varphi_1(0, x') = 0$$

und daher

$$D_j\varphi_1(x) = 0 \quad \forall 2 \le j \le k, x \in \partial H_k \cap U.$$

Wegen $\varphi(H_k \cap U) = H_k \cap U'$ gilt weiter für alle $x' \in \mathbb{R}^{k-1}$ und $h \in \mathbb{R}$ mit $(h, x') \in U$

$$\begin{aligned}\varphi_1(h, x') &\le 0 \quad \text{falls } h \le 0 \\ \varphi_1(h, x') &> 0 \quad \text{falls } h > 0.\end{aligned}$$

Also ist

$$(*) \qquad \begin{aligned}D_1\varphi_1(0, x') &= \lim_{h \to 0} \frac{\varphi_1(h, x') - \varphi_1(0, x')}{h} \\ &= \lim_{h \to 0} \frac{\varphi_1(h, x')}{h} \\ &\ge 0.\end{aligned}$$

Damit folgt für jedes $x \in \partial H_k \cap U$

$$0 < \det\Big((D_i\varphi_j(x))_{1 \le i,j \le k}\Big)$$

$$= \det \begin{pmatrix} D_1\varphi_1(x) & 0 \\ * & (D_i\varphi_j(x))_{2\le i,j\le k} \end{pmatrix}$$
$$= D_1\varphi_1(x) \det((D_i\varphi_j(x))_{2\le i,j\le k}).$$

Zusammen mit $(*)$ folgt hieraus die Behauptung. □

Definition X.6.13. (1) Seien $M \subset \mathbb{R}^n$ und $K \subset M$. Dann heißt

$$\partial_M K = \{x \in M : \ \forall U \in \mathcal{U}(x) : U \cap K \neq \emptyset, U \cap K^c \neq \emptyset\}$$

der Rand von $K$ relativ zu $M$.
(2) Sei $M \subset \mathbb{R}^n$ eine $k$-dimensionale Mfgkt und $K \subset M$ kompakt. $K$ hat glatten Rand (relativ zu $M$), wenn es zu jedem $x \in \partial_M K$ eine Karte $(U, \varphi, V)$ von $M$ mit $x \in V$ und

(a) $\varphi(H_k \cap U) = K \cap V$,
(b) $\varphi(\partial H_k \cap U) = \partial_M K \cap V$

gibt.

Bemerkung X.6.14. (1) Wenn wir uns darauf verständigen, unter einer $n$-dimensionalen Mfgkt in $\mathbb{R}^n$ eine offene Menge zu verstehen, ist Definition X.6.13 eine Verallgemeinerung von Definition X.3.11 (S. 379) für Kompakta mit glattem Rand.
(2) Man kann auch den Begriff eines Kompaktums mit stückweise glattem Rand übertragen. Dazu muss man $\partial_M K$ in der Form $\partial_M K = \partial_{M,R} K \cup \partial_{M,S} K$ mit $\partial_{M,R} K \cap \partial_{M,S} K = \emptyset$ schreiben und fordern, dass $\partial_{M,R} K$ relativ offen in $\partial_M K$ ist und die Bedingungen (a) und (b) erfüllt und dass es zu jedem $x \in \partial_{M,S} K$ eine Karte $(U, \varphi, V)$ von $M$ mit $x \in V$ und

$$\lim_{\varepsilon \to 0} \varepsilon^{-1} \lambda_k\Big( \bigcup_{u \in \varphi^{-1}(V \cap \partial_{M,S} K)} B(u, \varepsilon)\Big) = 0$$

gibt.
(3) Es ist stets $\partial_M K \subset \partial K$; i. a. gilt jedoch $\partial_M K \neq \partial K$.

Beispiel X.6.15. (1) Sei $M = S^2$ und

$$K = \{x \in S^2 : x_3 \geq 0\}.$$

Dann ist $\partial_M K$ der Äquator, und $K$ hat glatten Rand relativ zu $M$.
(2) Sei $M = S^2$ und

$$K = \{x \in S^2 : x_i \geq 0 \ \ , \ \ 1 \leq i \leq 3\}.$$

Dann ist

$$\partial_M K = \bigcup_{i=1}^{3} \{x \in S^2 : x_i = 0, x_j \geq 0 \ j \neq i\}$$

und

$$\partial_{M,S} K = \{(1,0,0), (0,1,0), (0,0,1)\}.$$

Satz X.6.16. *Sei $M \subset \mathbb{R}^n$ eine $k$-dimensionale Mfgkt und $K \subset M$ kompakt mit glattem Rand relativ zu $M$. Dann ist $\partial_M K$ eine $(k-1)$-dimensionale Mfgkt. Jede Orientierung von $M$ induziert eine Orientierung von $\partial_M K$.*

Beweis. Der Fall $k = 1$ ist trivial (eine 0-dimensionale Mfgkt ist eine endliche Punktmenge). Sei also $k \geq 2$ und $\mathcal{A} = \{(U, \varphi, V)\}$ ein Atlas von $M$. O.E. können wir annehmen, dass $\mathcal{A}$ rand-adaptiert ist bzgl. $K$, d.h., jede Karte von $\mathcal{A}$ erfüllt die Bedingungen (a) und (b) von Definition X.6.13. Insbesondere ist dann $U \cap \partial H_k = \emptyset$, wenn $V \cap \partial_M K = \emptyset$ ist.
Sei nun $(U, \varphi, V)$ eine Karte von $\mathcal{A}$ mit $V \cap \partial_M K \neq \emptyset$. Sei $\beta$ die globale Karte von $\partial H_k$. Definiere

$$\begin{aligned} U_0 &= \beta^{-1}(\partial H_k \cap U), \\ V_0 &= \partial_M K \cap V, \\ \psi &= \varphi \circ \beta. \end{aligned}$$

Da $\varphi$ ein Homöomorphismus mit $\varphi(\partial H_k \cap U) = \partial_M K \cap V$ ist, ist $\psi : U_0 \to V_0$ ein Homöomorphismus. Weiter ist

$$D\psi(u) = (D_i \varphi_j(0, u))_{\substack{2 \leq i \leq k \\ 1 \leq j \leq n}} \quad \forall u \in U_0.$$

Da $D\varphi$ den Rang $k$ hat, hat somit $D\psi$ den Rang $k - 1$.
Sei $\mathcal{A}_0$ die Menge aller Karten, die sich so aus den Karten von $\mathcal{A}$ ergeben. Aus Satz X.1.6 (S. 362) folgt, dass $\partial_M K$ eine $(k-1)$-dimensionale Mfgkt ist.
Sei nun $\mathcal{A}$ wie oben zusätzlich orientiert. Wir wollen zeigen, dass der wie oben konstruierte Atlas $\mathcal{A}_0$ orientiert ist. Seien dazu $(U_i, \varphi_i, V_i)$, $i = 1, 2$, zwei Karten von $\mathcal{A}$ mit $V_1 \cap V_2 \cap \partial_M K \neq \emptyset$. Dann ist $\tau = \varphi_2^{-1} \circ \varphi_1 : U_1 \to U_2$ ein orientierungstreuer $C^1$-Diffeomorphismus. Sei $\tilde{\tau} = (\varphi_2 \circ \beta)^{-1} \circ (\varphi_1 \circ \beta) : U_{10} \to U_{20}$ die Transformation zwischen den abgeleiteten Karten von $\partial_M K$. Dann ist

$$\tilde{\tau}(u) = (\tau_2(0, u), \ldots, \tau_k(0, u)) \quad \forall u \in U_{10}.$$

Damit folgt aus Lemma X.6.12, dass $\tilde{\tau}$ orientierungstreu ist. Also ist $\mathcal{A}_0$ orientiert. □

Beispiel X.6.17. Seien $M$ und $K$ wie in Beispiel X.6.15 (1). $M$ sei durch die äußere Normale orientiert. Dann ist die induzierte Orientierung von $\partial_M K$ so, dass die Kurve $\partial_M K$ im mathematisch positiven Sinn durchlaufen wird, d.h., die obere Halbkugel liegt beim Durchlaufen von $\partial_M K$ links.

Wir kommen nun zum Hauptergebnis dieses Abschnittes.

Satz X.6.18 (Stokesscher Integralsatz). *Seien $U \subset \mathbb{R}^n$ offen, $M \subset U$ eine orientierbare $k$-dimensionale Mfgkt, $k \geq 2$, $K \subset M$*

*kompakt mit glattem Rand relativ zu $M$ und $\alpha \in \Omega_1^{k-1}(U)$. Dann ist*

$$\int_K d\alpha = \int_{\partial_M K} \alpha,$$

*wobei $\partial_M K$ die durch $M$ induzierte Orientierung hat.*

Beweis. Da $K$ kompakt ist, gibt es ein $m \in \mathbb{N}^*$ und positiv orientierte Karten $(U_i, \varphi_i, V_i)$, $1 \le i \le m$, von $M$ mit

(1) $K \subset \bigcup_{i=1}^{m} V_i$,
(2) $\varphi_i(H_k \cap U_i) = K \cap V_i$,
(3) $\varphi_i(\partial H_k \cap U_i) = \partial_M K \cap V_i$.

Sei $\{\psi_i : 1 \le i \le m\}$ eine $\{V_i : 1 \le i \le m\}$ untergeordnete Partition der Eins auf $K$. Wegen

$$\begin{aligned}
\int_K d\alpha &= \int_K d\Big(\sum_{i=1}^m \psi_i\alpha\Big) \\
&= \sum_{i=1}^m \Big(\int_K d(\psi_i\alpha)\Big) \\
\int_{\partial_M K} \alpha &= \int_{\partial_M K} \sum_{i=1}^m \psi_i\alpha \\
&= \sum_{i=1}^m \Big(\int_{\partial_M K} \psi_i\alpha\Big)
\end{aligned}$$

reicht es, die Behauptung für die $(k-1)$-Formen $\psi_i\alpha$ zu zeigen. Wegen $\operatorname{supp}(\varphi_i * (\psi_i\alpha)) \Subset U_i$ kann $\varphi_i * (\psi_i\alpha)$ zu einer stetig differenzierbaren $(k-1)$ Form $\tilde{\alpha}_i$ auf $\mathbb{R}^k$ mit kompaktem Träger fortgesetzt werden. Damit folgt aus Satz X.5.10(4) (S. 397) und Satz X.6.11

$$\begin{aligned}
\int_K d(\psi_i\alpha) &= \int_{K\cap V_i} d(\psi_i\alpha) \\
&= \int_{H_k\cap U_i} \varphi_i * d(\psi_i\alpha) \\
&= \int_{H_k\cap U_i} d(\varphi_i * (\psi_i\alpha)) \\
&= \int_{H_k} d\tilde{\alpha}_i \\
&= \int_{\partial H_k} \tilde{\alpha}_i \\
&= \int_{\partial H_k\cap U_i} \varphi_i * (\psi_i\alpha)
\end{aligned}$$

$$= \int_{\partial_M K \cap V_i} \psi_i \alpha$$
$$= \int_{\partial_M K} \psi_i \alpha.$$

Hieraus folgt die Behauptung. □

Bemerkung X.6.19. (1) Wenn wir offene Mengen im $\mathbb{R}^n$ als $n$-dimensionale Mfgkten bezeichnen, ist Satz X.6.8 der Spezialfall $k = n$ von Satz X.6.18.
(2) Wenn wir eine endliche Punktmenge als 0-dimensionale Mfgkt bezeichnen, ist der Hauptsatz der Differential- und Integralrechnung der Spezialfall $k = 1$ von Satz X.6.18, d.h.

$$\int_\gamma df = f(\gamma(b)) - f(\gamma(a))$$

für jeden $C^1$-Weg $\gamma \in C^1([a,b], \mathbb{R}^n)$ und jedes $f \in C^1(U, \mathbb{R})$ mit $\gamma([a,b]) \subset U$.
(3) Mit einigem technischem Mehraufwand kann man Satz X.6.18 auch für Kompakta mit stückweise glattem Rand beweisen.

Falls die Mfgkt $M$ selber kompakt ist, kann man $K = M$ in Satz X.6.18 wählen. Da dann $\partial_M K = \emptyset$ ist, folgt sofort:

Satz X.6.20. *Seien $U \subset \mathbb{R}^n$ offen, $M \subset U$ eine kompakte orientierbare k-dimensionale Mfgkt und $\alpha \in \Omega_1^{k-1}(U)$. Dann ist*

$$\int_M d\alpha = 0.$$

Beispiel X.6.21. Sei $n \geq 2$ und

$$\alpha = \sum_{i=1}^n \|x\|_2^{-n} x_i (-1)^{i-1} dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n \in \Omega_\infty^{n-1}(\mathbb{R}^n \backslash \{0\}).$$

Wegen

$$\operatorname{div}(\|x\|_2^{-n} x) = 0$$

ist $\alpha$ geschlossen. $\alpha$ ist nicht exakt. Dann wäre $\alpha = d\beta$ mit $\beta \in \Omega_1^{n-2}(\mathbb{R}^n \backslash \{0\})$, so folgte aus Satz X.6.7 und Satz X.6.20

$$0 = \int_{S^{n-1}} d\beta$$
$$= \int_{S^{n-1}} \alpha$$
$$= \int_{S^{n-1}} (\|x\|_2^{-n} x, \nu) dS$$
$$= \sigma_{n-1}(S^{n-1}).$$

Im Fall $n = 3$ ist daher

$$u = \|x\|_2^{-3} x = \frac{1}{(x_1^2 + x_2^2 + x_3^2)^{3/2}} (x_1, x_2, x_3)$$

ein Beispiel für ein divergenzfreies Vektorfeld, das nicht Rotation eines Vektorfeldes ist.

Als eine Konsequenz von Satz X.6.18 erhalten wir den klassischen Stokesschen Satz.

Satz X.6.22 (Stokesscher Integralsatz im $\mathbb{R}^3$). *Seien $U \subset \mathbb{R}^3$ offen, $M \subset U$ eine durch das Einheitsnormalenfeld $\nu$ orientierte Hyperfläche, $K \subset M$ kompakt mit glattem Rand relativ zu $M$ und $u \in C^1(U, \mathbb{R}^3)$. Dann ist*

$$\int_K (\operatorname{rot} u, \nu) dS = \int_{\partial_M K} (u, \tau) ds,$$

*wobei $ds$ die Bogenlänge von $\partial_M K$ und $\tau$ das durch die induzierte Orientierung von $\partial_M K$ definierte Einheitstangentenfeld ist.*

Beweis. Wende Satz X.6.18 auf

$$\alpha_{1,u} = \sum_{i=1}^{3} u_i dx_i$$

an. Dann folgt die Behauptung aus Satz X.6.7, Beispiel X.5.9 (S. 397), Bemerkung X.6.5 (2) und Bemerkung VIII.2.7 (S. 266), Definition VIII.2.4 (S. 265) und Definition VIII.3.1 (S. 271). □

Beispiel X.6.23. (1) Seien $M, K$ wie in Beispiel X.6.15 (1) und

$$f(x) = (-x_2, x_1, 0).$$

Dann ist

$$\operatorname{rot} f(x) = 2e_3 = (0, 0, 2)$$

und wegen Satz X.6.7 und Beispiel X.3.7 (S. 374)

$$\begin{aligned} \int_K (\operatorname{rot} f, \nu) dS &= \int_0^{2\pi} \int_0^{\frac{\pi}{2}} 2 \sin\theta \cos\theta d\theta d\varphi \\ &= 2\pi. \end{aligned}$$

Andererseits ist

$$\begin{aligned} \int_{\partial_M K} (f, \tau) ds &= \int_{S^1} (-x_2, x_1) \\ &= \int_0^{2\pi} (-\sin t)(-\sin t) + \cos t \cos t dt \\ &= 2\pi. \end{aligned}$$

(2) Seien $x^* \in \mathbb{R}^3$ und $v \in \mathbb{R}^3$ mit $\|v\|_2 = 1$. Sei $M = \{x \in \mathbb{R}^3 : (x - x^*, v) = 0\}$ die Ebene durch $x^*$ senkrecht zu $v$. Wir orientieren $M$ so, dass $v$ ein positiv orientierter Normalenvektor ist. Für $\varepsilon > 0$ sei

$$K_\varepsilon = \{x \in M : \|x - x^*\|_2 \le \varepsilon\}.$$

Sei $U \subset \mathbb{R}^3$ eine Umgebung von $x^*$ und $u \in C^1(U, \mathbb{R}^3)$. Dann folgt aus Satz X.6.22

$$\begin{aligned}(\operatorname{rot} u(x^*), v) &= \lim_{\varepsilon \to 0} \frac{1}{\pi\varepsilon^2} \int_{K_\varepsilon} (\operatorname{rot} u, v) dS \\ &= \lim_{\varepsilon \to 0} \frac{1}{\pi\varepsilon^2} \int_{\partial K_\varepsilon} (u, \tau) ds.\end{aligned}$$

Physikalisch beschreibt also die Rotation $\operatorname{rot} u(x^*)$ den Fluss pro Flächeneinheit durch einen beliebig kleinen Kreis mit Mittelpunkt $x^*$.
(3) Seien $U \subset \mathbb{R}^3$ offen und $I \subset \mathbb{R}$ ein Intervall. Das elektrische und das magnetische Feld sind zeitabhängige Vektorfelder $E, B \in C^1(I \times U, \mathbb{R}^3)$. Sie genügen u.a. der Differentialgleichung

$$\text{(D)} \qquad \operatorname{rot} E = -\frac{\partial B}{\partial t}.$$

Seien $M$ eine durch das Einheitsnormalenfeld $\nu$ orientierte Hyperfläche in $U$ und $K \subset M$ kompakt mit glattem Rand relativ zu $M$. Dann ist

$$\int_K (B(t, x), \nu(x)) dS(x)$$

der magnetische Fluss durch $K$ zur Zeit $t$. Aus (D) und Satz X.6.22 folgt

$$\text{(I)} \qquad \begin{aligned}\frac{d}{dt} \int_K (B(t, x), \nu(x)) dS(x) &= \int_K (\frac{\partial}{\partial t} B(t, x), \nu(x)) dS(x) \\ &= - \int_K (\operatorname{rot} E(t, x), \nu(x)) dS(x) \\ &= - \int_{\partial_M K} (E(t, x), \tau(x)) ds(x).\end{aligned}$$

Die zeitliche Änderung des magnetischen Flusses durch $K$ ist also gleich dem negativen elektrischen Strom durch $\partial_M K$.
Mit Hilfe von Teil (2) kann man umgekehrt aus (I) die Differentialgleichung (D) herleiten.

## X.7. Die Fixpunktsätze von Brouwer und Schauder

Satz X.7.1 (Brouwerscher Fixpunktsatz). *Sei $\overline{B}_n = \{x \in \mathbb{R}^n : \|x\|_2 \le 1\}$ die abgeschlossene Einheitskugel in $\mathbb{R}^n$ und $f \in C(\overline{B}_n, \mathbb{R}^n)$ mit $f(\overline{B}_n) \subset \overline{B}_n$. Dann besitzt $f$ mindestens einen Fixpunkt $x^*$ in $\overline{B}_n$, d.h.*

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

Beweis. Für $n = 1$ besagt Satz X.7.1, dass jedes $f \in C([-1,1],\mathbb{R})$ mit $f([-1,1]) \subset [-1,1]$ einen Fixpunkt besitzt und ist damit eine Konsequenz des Zwischenwertsatzes angewandt auf $x - f(x)$.
Sei also $n \geq 2$.
1. Schritt: Es gebe ein $g \in C^1(\mathbb{R}^n,\mathbb{R}^n)$ mit $g|_{\overline{B}_n} = f$.
Ann.: $f$ besitzt keinen Fixpunkt.
Dann gibt es wegen der Stetigkeit von $g$ ein $r > 1$ mit

$$x - g(x) \neq 0 \quad \forall x \in B(0,r) \subset \mathbb{R}^n.$$

Sei $\varphi_1 = x - g(x)$ und $\alpha \in \Omega_\infty^{n-1}(\mathbb{R}^n\backslash\{0\})$ wie in Beispiel X.6.21 (S. 414). Dann ist $\varphi_1 * \alpha$ in $B(0,r)$ geschlossen und damit wegen Satz X.5.16 (S. 400) exakt. Also gilt nach Satz X.6.20 (S. 414)

$$\int_{S^{n-1}} \varphi_1 * \alpha = 0. \tag{X.7.1}$$

Definiere $\Phi : \mathbb{R} \times B(0,r) \longrightarrow \mathbb{R}^n$ durch

$$\Phi(t,x) = x - tg(x).$$

Für $t \in [0,1]$ und $x \in S^{n-1}$ gilt dann $\Phi(t,x) \neq 0$. Also ist $V = \Phi^{-1}(\mathbb{R}^n\backslash\{0\})$ offen und $[0,1] \times S^{n-1} \subset V$. Dann gibt es auch eine offene Menge $U \subset \mathbb{R}^n$ mit $S^{n-1} \subset U \subset B(0,r)$ und $[0,1] \times U \subset V$. Definiere $\psi_0, \psi_1 : U \to V$ durch

$$\psi_0(x) = (0,x),$$
$$\psi_1(x) = (1,x).$$

Da $\Phi * \alpha$ geschlossen ist, folgt wie im Beweis von Satz X.5.16 (S. 400), dass es ein $\eta \in \Omega_1^{n-2}(U)$ gibt mit

$$\psi_1 * \Phi * \alpha - \psi_0 * \Phi * \alpha = d\eta.$$

Für $x \in U$ ist aber

$$\Phi \circ \psi_1(x) = \Phi(1,x) = x - g(x) = \varphi_1(x)$$
$$\Phi \circ \psi_0(x) = \Phi(0,x) = x$$

und daher

$$d\eta = \varphi_1 * \alpha - \alpha.$$

Damit folgt aus Satz X.6.20 (S. 414) und Gleichung (X.7.1)

$$\begin{aligned}\sigma_{n-1}(S^{n-1}) &= \int_{S^{n-1}} \alpha \\ &= \int_{S^{n-1}} (\varphi_1 * \alpha - d\eta) \\ &= 0.\end{aligned}$$

Dies ist ein Widerspruch. Also besitzt $f$ doch einen Fixpunkt.
2. Schritt: Sei nun $f \in C(\overline{B}_n,\mathbb{R}^n)$ mit $f(\overline{B}_n) \subset \overline{B}_n$ beliebig.

ANN.: $f$ besitzt keinen Fixpunkt.
Da $\overline{B}_n$ kompakt ist, gilt

$$\varepsilon = \inf_{x \in \overline{B}_n} \|x - f(x)\|_2 > 0.$$

Definiere $\tilde{f} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ durch

$$\tilde{f}(x) = \begin{cases} f(x) & \text{falls } x \in \overline{B}_n, \\ f(\frac{x}{\|x\|_2}) & \text{falls } x \notin \overline{B}_n. \end{cases}$$

Dann ist $\tilde{f} \in C(\mathbb{R}^n, \mathbb{R}^n)$ mit $\tilde{f}(\mathbb{R}^n) \subset \overline{B}_n$. Sei $\psi \in C_0^\infty(\mathbb{R}^n, \mathbb{R})$ mit

$$\begin{aligned} &0 \le \psi(x) \le 1 \quad \forall x \in \mathbb{R}^n \\ &\operatorname{supp} \psi \subset \overline{B}_n \\ &\int_{\mathbb{R}^n} \psi = 1 \end{aligned}$$

wie im Beweis von Satz IX.6.7 (S. 354). Da $\overline{B(0,2)}$ kompakt und $\tilde{f}$ stetig ist, gibt es ein $\delta > 0$ mit

$$\|\tilde{f}(x) - \tilde{f}(y)\|_2 \le \frac{\varepsilon}{2} \quad \forall x, y \in \overline{B(0,2)} \text{ mit } \|x - y\|_2 \le \delta.$$

Definiere $g : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ durch

$$\begin{aligned} g(x) &= \int_{\mathbb{R}^n} \delta^{-n} \psi(\frac{y - x}{\delta}) \tilde{f}(x) dx \\ &= \int_{\mathbb{R}^n} \psi(z) \tilde{f}(x + \delta z) dz \quad \forall x \in \mathbb{R}^n. \end{aligned}$$

Dann ist $g \in C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ und

$$\begin{aligned} \|g(x)\|_2 &\le \int_{\mathbb{R}^n} \psi(z) \|\tilde{f}(x + \delta z)\|_2 dz \\ &\le 1 \quad \forall x \in \mathbb{R}^n. \end{aligned}$$

Insbesondere ist $g(\overline{B}_n) \subset \overline{B}_n$.
Weiter ist für $x \in \overline{B}_n$

$$\begin{aligned} \|f(x) - g(x)\|_2 &= \| \int_{\mathbb{R}^n} \psi(z) [f(x) - \tilde{f}(x + \delta z)] dz \|_2 \\ &= \| \int_{\overline{B}_n} \psi(z) [\tilde{f}(x) - \tilde{f}(x + \delta z)] dz \|_2 \\ &\le \frac{\varepsilon}{2}. \end{aligned}$$

Daher gilt für alle $x \in \overline{B}_n$

$$\begin{aligned} \|x - g(x)\|_2 &\ge \|x - f(x)\|_2 - \|f(x) - g(x)\|_2 \\ &\ge \frac{\varepsilon}{2} \end{aligned}$$

im Widerspruch zum 1. Schritt. Also besitzt $f$ doch einen Fixpunkt. □

BEMERKUNG X.7.2. (1) Das Beispiel $n = 1$, $f(x) = x^2$ zeigt, dass der Fixpunkt aus Satz X.7.1 im allgemeinen nicht eindeutig ist.
(2) Sei $v \in S^{n-1}$. Definiere $f \in C(\mathbb{R}^n, \mathbb{R}^n)$ durch

$$f(x) = \frac{1}{2}(v + x).$$

Dann ist $f(B_n) \subset B_n$ und $f$ besitzt keinen Fixpunkt in $B_n$. Dieses Beispiel zeigt, dass man in Satz X.7.1 $\overline{B}_n$ nicht durch $B_n$ ersetzen kann.

Satz X.7.1 hat eine einfache, aber sehr praktische Konsequenz.

SATZ X.7.3. *Sei $f \in C(\mathbb{R}^n, \mathbb{R}^n)$. Es gebe ein $r > 0$ mit*

$$(f(x), x) > 0 \quad \forall x \in \mathbb{R}^n \text{ mit } \|x\|_2 = r. \tag{X.7.2}$$

*Dann besitzt $f$ eine Nullstelle $x$ in $B(0, r)$.*

BEWEIS. ANN.: $f$ besitzt keine Nullstelle in $\overline{B(0, r)}$. Definiere $g : \overline{B}_n \to \mathbb{R}^n$ durch

$$g(x) = -\|f(rx)\|_2^{-1} f(rx).$$

Dann ist $g$ stetig und $g(\overline{B}_n) \subset \overline{B}_n$. Wegen Satz X.7.1 gibt es ein $x^* \in \overline{B}_n$ mit

$$g(x^*) = x^*.$$

Wegen $g(\overline{B}_n) \subset S^{n-1}$ ist $\|x^*\|_2 = 1$. Also gilt

$$\begin{aligned} 1 &= \|x^*\|_2^2 \\ &= \frac{1}{r}(rx^*, x^*) \\ &= \frac{1}{r}(rx^*, g(x^*)) \\ &= -(r\|f(rx^*)\|_2)^{-1}(rx^*, f(rx^*)) \\ &< 0. \end{aligned}$$

Dies ist ein Widerspruch. Also besitzt $f$ eine Nullstelle in $\overline{B(0, r)}$. Wegen Gleichung (X.7.2) kann diese aber nicht auf $\partial B(0, r)$ liegen. □

Satz X.7.3 hat eine interessante Konsequenz.

SATZ X.7.4. *$S^{n-1}$ und $\overline{B}_{n-1}$ sind nicht homöomorph.*

BEWEIS. ANN.: Es gibt einen Homöomorphismus $\varphi$ von $S^{n-1}$ auf $\overline{B}_{n-1}$.
Definiere $\psi : \mathbb{R}^n \to \mathbb{R}^n$ durch

$$\psi(x) = \begin{cases} 0 & \text{falls } x = 0, \\ \|x\|_2 \varphi(\frac{x}{\|x\|_2}) & \text{falls } x \neq 0. \end{cases}$$

Dann ist $\psi$ stetig und damit auch $f : \overline{B}_n \to S^{n-1}$ mit $f = \varphi^{-1} \circ \psi$. Für $x \in S^{n-1}$ gilt

$$f(x) = x$$

und somit

$$(f(x), x) = 1 > 0 \quad \forall x \in S^{n-1}.$$

Wegen Satz X.7.3 gibt es ein $x^* \in B_n$ mit $f(x^*) = 0$. Dies ist ein Widerspruch zu $f(\overline{B}_n) \subset S^{n-1}$. □

Wir wollen den Brouwerschen Fixpunktsatz verallgemeinern. Dazu benötigen wir ein approximationstheoretisches Ergebnis, das von eigenständigem Interesse ist.

Satz X.7.5. *Seien $H$ ein Hilbertraum mit Skalarprodukt $(\cdot,\cdot)$ und Norm $\|\cdot\|$ und $M \subset H$ abgeschlossen und konvex. Dann besitzt jedes $x \in H$ eine eindeutige beste Approximation $P_M x \in M$ in $M$, d.h.*

$$\|x - P_M x\| = \inf_{y \in M} \|x - y\|.$$

*Weiter gilt*

$$\|P_M x - P_M y\| \le \|x - y\| \quad \forall x, y \in H.$$

Beweis. Vorbemerkung: Da $(\cdot,\cdot)$ ein Skalarprodukt ist, gilt die Parallelogrammregel

$$\|x + y\|^2 + \|x - y\|^2 = 2[\|x\|^2 + \|y\|^2]. \tag{X.7.3}$$

Eindeutigkeit: Seien $x \in H$ beliebig und $y, z \in M$ zwei beste Approximationen an $x$. Dann folgt aus Gleichung (X.7.3)

$$\begin{aligned}\|y - z\|^2 &= 2\|x - y\|^2 + 2\|x - z\|^2 - \|2x - y - z\|^2 \\ &= 2\{\|x - y\|^2 + \|x - z\|^2 - 2\|x - \underbrace{\tfrac{1}{2}(y + z)}_{\in M}\|^2\} \\ &\le 0.\end{aligned}$$

Also ist $y = z$.
Existenz: Sei $x \in H$ beliebig und $d = \inf_{y \in H} \|x - y\|$. O.E. ist $d > 0$, sonst ist die Behauptung trivial. Sei $(y_n)_{n \in \mathbb{N}}$ eine Minimalfolge, d.h.

$$\lim_{n \to \infty} \|x - y_n\| = d.$$

Dann folgt aus Gleichung (X.7.3) für $n, m \in \mathbb{N}$

$$\begin{aligned}\|y_n - y_m\|^2 &= 2\{\|x - y_n\|^2 + \|x - y_m\|^2 - 2\|x - \underbrace{\tfrac{1}{2}(y_n + y_m)}_{\in M}\|^2\} \\ &\le 2\{\|x - y_n\|^2 + \|x - y_m\|^2 - 2d\}.\end{aligned}$$

Also ist $(y_n)_{n\in\mathbb{N}}$ eine Cauchyfolge und konvergiert damit gegen ein $y^* \in M$. Konstruktionsgemäß gilt $\|x - y^*\| = d$.
STETIGKEIT VON $P_M$: Seien $x, y \in H$ und $t \in (0,1)$ beliebig. Dann folgt

$$\begin{aligned}
0 &\le \|x - \underbrace{(1-t)P_Mx - tP_My}_{\in M}\|^2 - \|x - P_Mx\|^2 \\
&\quad + \|y - \underbrace{(1-t)P_My - tP_Mx}_{\in M}\|^2 - \|y - P_My\|^2 \\
&= \|x - P_Mx + t[P_Mx - P_My]\|^2 - \|x - P_Mx\|^2 \\
&\quad + \|y - P_My - t[P_Mx - P_My]\|^2 - \|y - P_My\|^2 \\
&= 2t(x - P_Mx, P_Mx - P_My) \\
&\quad - 2t(y - P_My, P_Mx - P_My) \\
&\quad + 2t^2\|P_Mx - P_My\|^2 \\
&= 2t(x - y, P_Mx - P_My) - 2t(1-t)\|P_Mx - P_My\|^2
\end{aligned}$$

und somit

$$\begin{aligned}
\|P_Mx - P_My\|^2 &\le \frac{1}{1-t}(x - y, P_Mx - P_My) \\
&\le \frac{1}{1-t}\|x - y\|\|P_Mx - P_My\| \quad \forall 0 < t < 1.
\end{aligned}$$

Hieraus folgt die Behauptung durch Grenzübergang $t \to 0$. □

SATZ X.7.6. *Seien $X$ ein endlich dimensionaler Banachraum, $M \subset X$ konvex und kompakt und $f \in C(M, M)$. Dann besitzt $f$ einen Fixpunkt in $M$.*

BEWEIS. Sei $n = \dim X$ und $\Phi : \mathbb{R}^n \to X$ ein Isomorphismus. Dann ist $N = \Phi^{-1}(M) \subset \mathbb{R}^n$ konvex und kompakt. Insbesondere gibt es ein $R > 0$ mit $N \subset B(0, R)$. Sei $P_N : \mathbb{R}^n \to N$ der Operator, der gemäß Satz X.7.5 jedem $x \in \mathbb{R}^n$ seine beste Approximation in $N$ zuordnet. Definiere $\psi_R : \mathbb{R}^n \to \mathbb{R}^n$ durch

$$\psi_R(x) = Rx$$

und $g : \overline{B}_n \to \mathbb{R}^n$ durch (vgl. Abb. X.7.1)

$$g = \psi_R^{-1} \circ \Phi^{-1} \circ f \circ \Phi \circ P_N \circ \psi_R.$$

Die Funktion $g$ ist stetig, und wegen $f(M) \subset M$ gilt $g(\overline{B}_n) \subset \overline{B}_n$. Also besitzt gemäß Satz X.7.1 $g$ einen Fixpunkt $z^* \in \overline{B}_n$. Sei $y^* = \psi_R(z^*)$. Dann gilt

$$\begin{aligned}
\Phi^{-1} \circ f \circ \Phi \circ P_N(y^*) &= \psi_R(g(z^*)) \\
&= y^*.
\end{aligned}$$

Also ist $y^* \in N$ und damit $P_N y^* = y^*$. Sei $x^* = \Phi(y^*) \in M$. Dann folgt

$$\begin{aligned} f(x^*) &= \Phi(\Phi^{-1} \circ f \circ \Phi \circ P_N(y^*)) \\ &= x^*. \end{aligned}$$

□

$$\begin{tikzcd}
X \supset M \arrow[r, "f"] & M \subset X \arrow[d, "\Phi^{-1}"] \\
\mathbb{R}^n \supset N \arrow[u, "\Phi"] & N \subset \mathbb{R}^n \arrow[d, "id"] \\
B(0,R) \supset N \arrow[u, "P_N"] & N \subset B(0,R) \arrow[d, "\psi_R^{-1}"] \\
B_n \arrow[u, "\psi_R"] \arrow[r, "g"] & B_n
\end{tikzcd}$$

Abbildung X.7.1. Definition der Funktion $g$

Bemerkung X.7.7. Satz X.7.6 folgt gemäß obigem aus Satz X.7.1. Umgekehrt ist offensichtlich Satz X.7.1 ein Spezialfall von Satz X.7.6. Also sind die beiden Sätze in Wahrheit äquivalente Formulierungen eines Sachverhaltes.

Definition X.7.8. Seien $X$ ein Vektorraum und $x_1, \ldots, x_n \in X$, $n \in \mathbb{N}^*$. Dann heißt

$$\operatorname{conv}[x_1, \ldots, x_n] = \Big\{ x = \sum_{i=1}^n \alpha_i x_i : 0 \le \alpha_i \le 1, \sum_{i=1}^n \alpha_i \le 1 \Big\}$$

die konvexe Hülle von $x_1, \ldots, x_n$.

Satz X.7.9. *Seien $X$ ein normierter Vektorraum und $x_1, \ldots, x_n \in X$, $n \in \mathbb{N}^*$. Dann ist* $\operatorname{conv}[x_1, \ldots, x_n]$ *kompakt und konvex.*

Beweis. Die Konvexität folgt aus der Konstruktion. Die Kompaktheit folgt aus $\operatorname{conv}[x_1, \ldots, x_n] = \Phi(S)$ mit

$$\begin{aligned} \Phi : \mathbb{R}^n &\to X \\ (\alpha_1, \ldots, \alpha_n) &\mapsto \sum_{i=1}^n \alpha_i x_i \end{aligned}$$

und

$$S = \Big\{ \alpha \in \mathbb{R}^n : 0 \le \alpha_i \le 1, \sum_{i=1}^n \alpha_i \le 1 \Big\}.$$

□

Satz X.7.10 (Erster Schauderscher Fixpunktsatz). *Seien $X$ ein normierter Vektorraum, $K \subset X$ konvex, $C \subset K$ kompakt und $f \in C(K, C)$. Dann besitzt $f$ einen Fixpunkt in $C$.*

Beweis. 1. Schritt: Zu jedem $\varepsilon > 0$ gibt es ein $x_\varepsilon \in K$ mit $\|f(x_\varepsilon) - x_\varepsilon\| \leq \varepsilon$.
Sei $\varepsilon > 0$ beliebig. Da $C$ kompakt ist, gibt es ein $n \in \mathbb{N}^*$ und Punkte $x_1, \ldots, x_n \in C$ mit $C \subset \bigcup_{i=1}^{n} B(x_i, \varepsilon)$. Sei

$$K_0 = \operatorname{conv}[x_1, \ldots, x_n] \subset K$$

und

$$\begin{aligned} \varphi_i(x) &= [\varepsilon - \|x - x_i\|]_+ \\ &= \max\{\varepsilon - \|x - x_i\|, 0\} \quad \forall x \in X, 1 \leq i \leq n. \end{aligned}$$

Dann sind die $\varphi_i$ stetig und

$$\varphi(x) = \sum_{i=1}^{n} \varphi_i(x) > 0 \quad \forall x \in C.$$

Daher ist $\psi_i : C \to \mathbb{R}$, $1 \leq i \leq n$, mit

$$\psi_i(x) = \frac{\varphi_i(x)}{\varphi(x)} \quad \forall x \in C$$

stetig und erfüllt

$$0 \leq \psi_i(x) \leq 1 \quad \forall x \in C, 1 \leq i \leq n$$
$$\sum_{i=1}^{n} \psi_i(x) = 1 \quad \forall x \in C.$$

Definiere $g : C \to K_0$ durch

$$g(x) = \sum_{i=1}^{n} \psi_i(x) x_i.$$

$g$ ist stetig, und für jedes $x \in C$ gilt

$$\begin{aligned} \|g(x) - x\| &= \| \sum_{i=1}^{n} \psi_i(x)[x_i - x] \| \\ &= \| \sum_{\substack{1 \leq j \leq n \\ x \in B(x_j, \varepsilon)}} \psi_j(x)[x_j - x] \| \\ &\leq \sum_{\substack{1 \leq j \leq n \\ x \in B(x_j, \varepsilon)}} \psi_j(x) \|x_j - x\| \\ &\leq \varepsilon. \end{aligned}$$

Definiere schließlich $h \in C(K, K_0)$ durch $h = g \circ f$. Aus Satz X.7.6 angewandt auf $h$ und $\operatorname{span}\{x_1, \dots, x_n\}$ (Man beachte: endlich dimensionale normierte Vektorräume sind vollständig!) folgt, dass es ein $x_\varepsilon \in K_0$ gibt mit

$$h(x_\varepsilon) = x_\varepsilon.$$

Hieraus folgt aber

$$\|f(x_\varepsilon) - x_\varepsilon\| = \|f(x_\varepsilon) - g(f(x_\varepsilon))\| \le \varepsilon.$$

2. Schritt: Gemäß dem 1. Schritt gibt es zu jedem $n \in \mathbb{N}^*$ ein $x_n \in K$ mit

$$\|f(x_n) - x_n\| \le \frac{1}{n}. \tag{X.7.4}$$

Da $f(K) \subset C$ und $C$ kompakt ist, gibt es eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ und ein $x^* \in C$ mit

$$f(x_{n_k}) \underset{k\to\infty}{\longrightarrow} x^*.$$

Aus Gleichung (X.7.4) folgt

$$x_{n_k} \underset{k\to\infty}{\longrightarrow} x^*.$$

Aus der Stetigkeit von $f$ ergibt sich schließlich

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

□

Wir wollen eine andere, äquivalente Formulierung des ersten Schauderschen Fixpunktsatzes angeben. Dazu benötigen wir:

Definition X.7.11. Seien $X$ ein normierter Vektorraum und $K \subset X$. $K$ heißt relativ kompakt, wenn $\overline{K}$ kompakt ist.

Bemerkung X.7.12. (1) Definition X.7.11 gilt auch für allgemeine topologische Räume.
(2) Eine relativ kompakte Menge ist notwendig beschränkt.
(3) Da in normierten Räumen Kompaktheit und Folgen-Kompaktheit übereinstimmen, ist $K$ genau dann relativ kompakt, wenn jede Folge in $K$ eine konvergente Teilfolge besitzt.

Satz X.7.13 (Zweiter Schauderscher Fixpunktsatz). *Seien $X$ ein normierter Vektorraum, $K \subset X$ konvex und $f \in C(K, K)$. $f$ besitzt sicher dann einen Fixpunkt in $K$, wenn eine der folgenden Voraussetzungen erfüllt ist:*

(1) *$K$ ist kompakt*

*oder*

(2) *$K$ ist abgeschlossen und $f(K)$ ist relativ kompakt.*

Beweis. Die Behauptung folgt aus Satz X.7.10 mit $C = K$ im Fall (1) und $C = \overline{f(K)} \subset K$ im Fall (2). □

Wir wollen Satz X.7.13 benutzen, um die lokale Existenz von Lösungen gewöhnlicher Differentialgleichungen zu beweisen. Dazu benötigen wir ein Hilfsergebnis das von eigenständigem Interesse ist.

SATZ X.7.14 (SATZ VON ARZELA-ASCOLI). *Seien* $(X, \|\cdot\|_X)$ *ein Banachraum und* $K \subset X$ *kompakt. Dann sind folgende Aussagen äquivalent:*

(1) $A \subset C(K, \mathbb{R}^n)$ *ist relativ kompakt.*
(2) $A$ *ist beschränkt und* GLEICHGRADIG STETIG, *d.h.*
  (a) $\sup_{f \in A} \|f\|_{C(K,\mathbb{R}^n)} \leq C$ *und*
  (b) $\sup_{f \in A} \|f(x) - f(y)\|_{\mathbb{R}^n} \to 0$ *für* $\|x - y\|_X \to 0$.

BEWEIS. „(1) $\Longrightarrow$ (2)“: Sei $\varepsilon > 0$ beliebig. Dann gibt es ein $n_\varepsilon \in \mathbb{N}^*$ und $f_{1,\varepsilon}, \ldots, f_{n_\varepsilon,\varepsilon} \in A$ mit $A \subset \bigcup_{i=1}^{n_\varepsilon} B(f_{i,\varepsilon}, \varepsilon)$, wobei $B(f, \varepsilon)$ die offene Kugel um $f$ mit Radius $\varepsilon$ bzgl. $\|\cdot\|_{C(K,\mathbb{R}^n)}$ ist. Für $f \in A$ gibt es dann ein $i_0 \in \mathbb{N}^*_{n_\varepsilon}$ mit $f \in B(f_{i_0,\varepsilon}, \varepsilon)$. Hieraus folgt

$$\begin{aligned} \|f\|_{C(K,\mathbb{R}^n)} &\leq \varepsilon + \|f_{i_0,\varepsilon}\|_{C(K,\mathbb{R}^n)} \\ &\leq \varepsilon + \max_{1 \leq i \leq n_\varepsilon} \|f_{i,\varepsilon}\|_{C(K,\mathbb{R}^n)}. \end{aligned}$$

Hieraus folgt (a). Weiter gibt es ein $\delta > 0$ mit

$$\max_{1 \leq i \leq n_\varepsilon} \|f_{i,\varepsilon}(x) - f_{i,\varepsilon}(y)\|_{\mathbb{R}^n} < \varepsilon$$

für alle $x, y \in K$ mit $\|x - y\|_X < \delta$ (Man beachte: die $f_{i,\varepsilon}$ sind gleichmäßig stetig!). Hieraus folgt für $f \in A$

$$\begin{aligned} \|f(x) - f(y)\|_{\mathbb{R}^n} &\leq 2\varepsilon + \|f_{i_0,\varepsilon}(x) - f_{i_0,\varepsilon}(y)\|_{\mathbb{R}^n} \\ &\leq 3\varepsilon \quad \forall x, y \in K, \|x - y\|_X < \delta. \end{aligned}$$

Dies zeigt (b).
„(2) $\Longrightarrow$ (1)“: Sei $\varepsilon > 0$ beliebig. Dann gibt es $k, m \in \mathbb{N}^*$ und $\eta_1, \ldots, \eta_k \in \mathbb{R}^n$, $x_1, \ldots, x_m \in K$ mit

$$\mathbb{R}^n \supset B(0, C) \subset \bigcup_{i=1}^{k} B(\eta_i, \varepsilon)$$

und

$$X \supset K \subset \bigcup_{j=1}^{m} B(x_j, \varepsilon).$$

Für jede Abbildung $\pi : \mathbb{N}^*_m \to \mathbb{N}^*_k$ sei

$$A_\pi = \{f \in A : \max_{1 \leq j \leq m} \|f(x_j) - \eta_{\pi(j)}\|_{\mathbb{R}^n} \leq \varepsilon\}.$$

Falls $A_\pi \neq \emptyset$ ist, wähle ein $f_\pi \in A_\pi$.
Zu jedem $f \in A$ gibt es dann ein $\pi$ mit $f \in A_\pi$. Sei nun $x \in K$. Dann gibt es ein $j \in \mathbb{N}^*_m$ mit $x \in B(x_j, \varepsilon)$. Daher gilt

$$\begin{aligned} \|f(x) - f_\pi(x)\|_{\mathbb{R}^n} \leq{}& \|f(x) - f(x_j)\|_{\mathbb{R}^n} + \|f(x_j) - \eta_{\pi(j)}\|_{\mathbb{R}^n} \\ &+ \|\eta_{\pi(j)} - f_\pi(x_j)\|_{\mathbb{R}^n} + \|f_\pi(x_j) - f_\pi(x)\|_{\mathbb{R}^n} \end{aligned}$$

$$\begin{aligned}&\leq 2\sup_{f\in A}\sup_{\substack{x,y\in K\\ \|z-y\|\leq\varepsilon}}\|f(z)-f(y)\|_{\mathbb{R}^n}+2\varepsilon\\ &= r_\varepsilon.\end{aligned}$$

Da $A$ gleichgradig stetig ist, gilt $r_\varepsilon \to 0$ für $\varepsilon \to 0$.
Aus obiger Abschätzung folgt

$$A \subset \bigcup_\pi B(f_\pi, 2r_\varepsilon),$$

wobei die (endliche) Anzahl von Kugeln, deren Vereinigung gebildet wird, von $\varepsilon$ abhängt. Wie in Beweis von Satz III.4.4 (S. 82) folgt hieraus, dass $\overline{A}$ kompakt ist. □

Satz X.7.15 (Existenzssatz von Peano). *Seien $I \subset \mathbb{R}$ ein offenes Intervall, $U \subset \mathbb{R}^n$ offen und $f \in C(I \times U, \mathbb{R})$. Dann gibt es zu jedem $\tau \in I$ und jedem $\eta \in U$ ein $\varepsilon_0 > 0$, so dass das* Anfangswertproblem

$$\begin{aligned}\frac{d}{dt}x(t) &= f(t,x(t)) \quad \forall \tau-\varepsilon_0 < t < \tau+\varepsilon_0\\ x(\tau) &= \eta\end{aligned} \tag{X.7.5}$$

*eine Lösung hat.*

Beweis. Offensichtlich ist $x \in C^1((\tau-\varepsilon_0,\tau+\varepsilon_0),\mathbb{R}^n)$ eine Lösung von Gleichung (X.7.5), genau dann wenn $x \in C((\tau-\varepsilon_0,\tau+\varepsilon_0),\mathbb{R}^n)$ ist und die Beziehung

$$x(t) = \eta + \int_\tau^t f(s,x(s))ds \quad \forall t-\varepsilon_0 < t < \tau+\varepsilon_0 \tag{X.7.6}$$

erfüllt.
Seien nun $\tau \in I$ und $\eta \in U$ beliebig. Dann gibt es ein $\varepsilon_1 > 0$ mit

$$Q = [\tau-\varepsilon_1,\tau+\varepsilon_1] \times \overline{B_{\mathbb{R}^n}(\eta,\varepsilon_1)} \subset I \times U.$$

Sei

$$M = \max\{1, \sup_{(t,y)\in Q}\|f(t,y)\|_{\mathbb{R}^n}\}$$

und

$$\varepsilon_0 = \frac{\varepsilon_1}{M}.$$

Definiere

$$\begin{aligned}X &= C([\tau-\varepsilon_0,\tau+\varepsilon_0],\mathbb{R}^n)\\ \|\cdot\|_X &= \|\cdot\|_{C([\tau-\varepsilon_0,\tau+\varepsilon_0],\mathbb{R}^n)}\\ K &= \{x \in X : \|x-\widetilde{\eta}\|_X \leq \varepsilon_1\},\end{aligned}$$

wobei $\widetilde{\eta}$ die konstante Funktion mit Wert $\eta$ bezeichnet.
$X$ ist ein Banachraum und $K \subset X$ ist konvex und abgeschlossen.

Definiere $L : K \to X$ durch

$$Lx(t) = \eta + \int_\tau^t f(s, x(s))ds.$$

Für $x \in K$ folgt

$$\begin{aligned}\|Lx - \widetilde{\eta}\|_X &= \sup_{|t-\tau|\le\varepsilon_0} \| \int_\tau^t f(s, x(s))ds\|_{\mathbb{R}^n} \\ &\le \varepsilon_0 M \\ &\le \varepsilon_1.\end{aligned}$$

Also ist $L(K) \subset K$.
Sei $\varepsilon > 0$ beliebig. Da $f$ auf der kompakten Menge $Q$ gleichmäßig stetig ist, gibt es ein $\delta > 0$ mit

$$\|f(s, u) - f(t, v)\|_{\mathbb{R}^n} \le \frac{\varepsilon}{\varepsilon_0} \qquad \forall (s, u), (t, v) \in Q$$
$$\text{mit } \max\{|s - t|, \|u - v\|_{\mathbb{R}^n}\} \le \delta.$$

Für $x, y \in K$ mit $\|x - y\|_X \le \delta$ folgt dann

$$\begin{aligned}\|Lx - Ly\|_X &= \sup_{|t-\tau|\le\varepsilon_0} \| \int_\tau^t [f(s, x(s)) - f(s, y(s))]ds\|_{\mathbb{R}^n} \\ &\le \varepsilon_0 \cdot \frac{\varepsilon}{\varepsilon_0} \\ &= \varepsilon.\end{aligned}$$

Also ist $L \in C(K, K)$.
Für $t_1, t_2 \in [\tau - \varepsilon_0, \tau + \varepsilon_0]$ gilt schließlich

$$\begin{aligned}\|Lx(t_1) - Lx(t_2)\|_{\mathbb{R}^n} &= \| \int_{t_1}^{t_2} f(s, x(s))ds\|_{\mathbb{R}^n} \\ &\le M|t_2 - t_1|\end{aligned}$$

und somit

$$\sup_{x\in K} \|Lx(t_1) - Lx(t_2)\|_{\mathbb{R}^n} \le M|t_1 - t_2|.$$

Damit folgt aus Satz X.7.14, dass $L(K)$ relativ kompakt ist und dass $X$, $K$, $L$ die Voraussetzungen von Satz X.7.13 erfüllen. Mithin hat Gleichung (X.7.6) und damit auch Gleichung (X.7.5) eine Lösung. □

Bemerkung X.7.16. (1) Seien $I = U = \mathbb{R}$ und $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(t, x) = \sqrt{|x|} \quad \forall (t, x) \in \mathbb{R}.$$

Für $a \in \mathbb{R}_+$ sei

$$\begin{aligned}u_a(t) &= \frac{1}{4}[(t - a)_+]^2 \\ &= \frac{1}{4} \max\{(t - a), 0\}^2.\end{aligned}$$

Dann ist $u_a$ Lösung des Anfangswertproblems

$$\begin{aligned}\frac{d}{dt}u_a(t) &= f(t, u_a(t)) \quad \forall t \in \mathbb{R}\\ u_a(0) &= 0.\end{aligned}$$

Dies zeigt, dass man unter den Voraussetzungen von Satz X.7.15 i. a. keine Eindeutigkeit der Lösung des Anfangswertproblems erwarten kann.

(2) Seien $I = U = \mathbb{R}$ und $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(t, x) = x^2 \quad \forall (t, x) \in \mathbb{R}.$$

Dann ist $u : (-\infty, 1) \to \mathbb{R}$ mit

$$u(t) = \frac{1}{1-t}$$

die (eindeutige) Lösung des Anfangswertproblems

$$\begin{aligned}\frac{d}{dt}u(t) &= f(t, u(t)) \quad \forall t \in (-\infty, 1)\\ u(0) &= 1.\end{aligned}$$

Dies zeigt, dass man unter den Voraussetzungen von Satz X.7.15 i. a. nur die lokale Existenz einer Lösung des Anfangswertproblems erwarten kann.

(3) Erfüllt $f$ aus Satz X.7.15 zusätzlich eine Lipschitz Bedingung bzgl. des zweiten Argumentes, d.h., gibt es ein $\lambda \in \mathbb{R}_+^*$ mit

$$\|f(t, x) - f(t, y)\|_{\mathbb{R}^n} \le \lambda \|x - y\|_{\mathbb{R}^n} \quad \forall t \in I, x, y \in U,$$

so kann man zeigen, dass das Anfangswertproblem eine eindeutige Lösung besitzt (Satz von Picard-Lindelöf) (s. Satz XI.1.7). Der Beweis beruht auf dem Banachschen Fixpunktsatz angewandt auf den Operator $L$ aus dem Beweis von Satz X.7.15.

KAPITEL XI

# Gewöhnliche Differentialgleichungen

In diesem Kapitel geben wir einen kurzen Einblick in die Theorie gewöhnlicher Differentialgleichungen. Zunächst beweisen wir einige einfache Existenz-, Eindeutigkeits- und Fortsetzbarkeitssätze. Dann stellen wir an Hand von Beispielen die wichtigsten elementaren Lösungsmethoden vor. Anschließend zeigen wir die stetige bzw. differenzierbare Abhängigkeit der Lösung eines Anfangswertproblems von den Anfangswerten. Zum Abschluss geben wir einen kurzen Einblick in die Stabilitätstheorie, d.h. über das Langzeitverhalten von Lösungen von Differentialgleichungen unter dem Einfluss kleiner Störungen.

## XI.1. Existenz- und Eindeutigkeitssätze

Im Folgenden bezeichnen stets $I \subset \mathbb{R}$ ein nicht leeres offenes Intervall und $U \subset \mathbb{R}^n$, $n \in \mathbb{N}^*$, eine nicht leere offene Menge. $\|\cdot\|$ ist eine beliebige Norm auf $\mathbb{R}^n$.

Definition XI.1.1. (1) Seien $k \in \mathbb{N}^*$, $V \subset \mathbb{R}^{kn}$ eine nicht leere offene Menge und $f : I \times V \to \mathbb{R}^n$ eine Funktion. Dann heißt das Problem

(D)
$$\begin{aligned}&\text{Finde } y \in C^k(I,\mathbb{R}^n) \text{ mit}\\ &y^{(k)}(t) = f\Big(t, y(t), \ldots, y^{(k-1)}(t)\Big), \quad t \in I\end{aligned}$$

eine gewöhnliche Differentialgleichung $k$-ter Ordnung auf $I$. Ist speziell $k = 1$, so sprechen wir einfach von einer gewöhnlichen Differentialgleichung (gDgl).
(2) Seien $t_0 \in I$ und $v_0 = (v_{0,0}, \ldots, v_{0,k-1}) \in V$. Dann heißt $(D)$ zusammen mit den Bedingungen

(A)
$$y(t_0) = v_{0,0}, \quad \ldots \quad , y^{(k-1)}(t_0) = v_{0,k-1}$$

ein Anfangswertproblem (AWP).
(3) Eine gDgl bzw. ein AWP heißen autonom, wenn die Funktion $f$ nicht von der Variablen $t$ abhängt.

Bemerkung XI.1.2. (1) Eine gDgl $k$-ter Ordnung, $k \geq 2$, kann stets in eine äquivalente gDgl erster Ordnung auf $\mathbb{R}^{nk}$ umformuliert werden. Um dies einzusehen, setze

$$z(t) = \Big(z_0(t), \ldots, z_{k-1}(t)\Big) = \Big(y(t), y'(t), \ldots, y^{(k-1)}(t)\Big) \in \mathbb{R}^{nk}$$

und

$$F(t,z(t)) = \Big(z_1(t),\dots,z_{k-1}(t), f(t,z_0(t),\dots,z_{k-1}(t))\Big) \in \mathbb{R}^{nk}.$$

Dann ist offensichtlich $y$ genau dann eine Lösung von $(D)$, wenn $z$ eine Lösung von

$$z'(t) = F(t,z(t)) \quad ,t \in I$$

ist. Aus diesem Grunde betrachten wir im Folgenden nur gewöhnliche Differentialgleichungen erster Ordnung.

(2) Eine gDgl kann stets in eine äquivalente autonome gDgl umformuliert werden. Um dies anzusehen, setze

$$z(s) = (y(s), s) \in \mathbb{R}^{n+1}$$

und

$$F(z) = F\Big((y(s),s)\Big) = \Big(f(s,y(s)),1\Big) \in \mathbb{R}^{n+1}.$$

Dann ist $y$ genau dann eine Lösung von $(D)$, wenn $z$ eine Lösung von

$$z'(s) = F(z(s)), \quad s \in I$$

ist.

Beispiel XI.1.3. (1) (Populationsdynamik) $y(t)$ beschreibe die Größe einer Population zur Zeit $t > 0$. Die relative zeitliche Änderung der Populationsgröße sei eine bekannte Funktion $r$ der Zeit und der aktuellen Populationsgröße. Diese Annahmen führen auf die gDgl

$$y'(t) = r(t,y(t))y(t) \quad ,t > 0.$$

Wichtige Spezialfälle sind die des unbeschränkten Wachstums

$$r(t,y(t)) = \alpha > 0 \quad \forall t > 0$$

und des beschränkten Wachstums

$$r(t,y(t)) = \alpha(y^* - y(t)) \quad \forall t > 0$$

mit $\alpha > 0, y^* > 0$. $y^*$ spielt die Rolle einer Grenzpopulation, deren Überschreiten zum Absterben der Population führt (vgl. Beispiel XI.2.5 (S. 444)).

(2) (Räuber-Beute-Modell) $x(t)$ und $y(t)$ beschreiben die Größe einer Räuber- bzw. Beutepopulation zur Zeit $t > 0$. Die Räuberpopulation lebe ausschließlich von der Beutepopulation, ihre Reproduktionsrate sei proportional zur Größe der Beutepopulation und bei fehlender Beute sei ihre relative Sterberate konstant. Die Beutepopulation habe bei fehlenden Räubern eine konstante relative Wachstumsrate und ihre Sterberate sei proportional zur aktuellen Größe der Räuberpopulation. Diese Annahmen führen zu der gDgl

$$\begin{aligned} x' &= -\alpha_1 x + \beta_1 xy \\ y' &= \alpha_2 y - \beta_2 xy \end{aligned}$$

mit $\alpha_1, \beta_1, \alpha_2, \beta_2 \in \mathbb{R}_+^*$ (vgl. Beispiel XI.2.2 (S. 440)).
(3) (Chemische Reaktionskinetik) Zwei Ausgangsstoffe $A$ und $B$ mögen zu einem Endprodukt $C$ reagieren, $u(t)$ beschreibe die Konzentration von $C$ zur Zeit $t > 0$. Die Änderung dieser Konzentration sei proportional zu der aktuellen Konzentration von $A$ und $B$. Wenn $a$ und $b$ die Anfangskonzentrationen von $A$ und $B$ bezeichnen, führen diese Annahmen auf das AWP

$$\begin{aligned} u'(t) &= r(a - u(t))(b - u(t)) \quad , t > 0 \\ u(0) &= 0 \end{aligned}$$

mit $r > 0$ (vgl. Beispiel XI.2.4 (S. 443)).
(4) (Federschwingung) Die vertikale Schwingung $z(t)$ einer Feder mit Masse $m > 0$ wird nach den Newtonschen Kraftgesetzen durch die gDgl 2. Ordnung

$$mz''(t) = -kz(t) - rz'(t) \quad , t > 0$$

beschrieben. Dabei ist $-kz(t), k > 0$, die Rückstellkraft der Feder und $-rz'(t), r \geq 0$, die Reibungskraft. Die äquivalente gDgl 1. Ordnung lautet

$$\begin{aligned} z'(t) &= v(t) \\ v'(t) &= -\frac{k}{m} z(t) - \frac{r}{m} v(t). \end{aligned}$$

Dabei ist $v(t)$ die Geschwindigkeit der Auslenkung.
(5) (Einfluss des Luftwiderstandes) Ein Fahrzeug der Masse $m > 0$ werde durch die konstante Kraft $K > 0$ beschleunigt. Der Luftwiderstand sei proportional zum Quadrat der Geschwindigkeit. Dann führen die Newtonschen Kraftgesetze auf die gDgl 2. Ordnung

$$mx''(t) = K - rx'(t)^2$$

mit $r > 0$. Bezeichnet $v(t)$ die Geschwindigkeit, lautet die äquivalente gDgl 1. Ordnung

$$\begin{aligned} x'(t) &= v(t) \\ mv'(t) &= K - rv(t)^2 \end{aligned}$$

(vgl. Beispiel XI.2.6 (S. 444)).

Wir wollen einen Existenz- und Eindeutigkeitssatz für die Lösungen von AWPen beweisen. Dazu benötigen wir einige Vorbereitungen.

Definition XI.1.4. Die Funktion $f \in C(I \times U, \mathbb{R}^n)$ heißt gleichmässig Lipschitz-stetig auf $I \times U$ bzgl. $U$, wenn es ein $L \in \mathbb{R}_+$, die sog. Lipschitz-Konstante, gibt mit

$$\|f(t, x) - f(t, y)\| \leq L\|x - y\| \quad \forall t \in I, x, y \in U.$$

$f$ heißt LIPSCHITZ-STETIG auf $I \times U$ bzgl. $U$, wenn es zu jedem $(t_0, x_0) \in I \times U$ eine Umgebung $J \times V \in \mathcal{U}\Big((t_0, x_0)\Big)$ gibt, derart dass $f$ auf $J \times V$ gleichmäßig Lipschitz-stetig ist bzgl. $V$.

BEMERKUNG XI.1.5. (1) Ist $f \in C(I \times U, \mathbb{R}^n)$ bzgl. der Variablen $x$ differenzierbar und $D_x f \in C(I \times U, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n))$, so ist $f$ auf $I \times U$ bzgl. $U$ Lipschitz-stetig.
(2) Ist $f \in C(I \times U, \mathbb{R}^n)$ Lipschitz-stetig auf $I \times U$ bzgl. $U$ und $J \times K \subset I \times U$ kompakt, so ist $f$ auf $J \times K$ gleichmäßig Lipschitz-stetig bzgl. $K$.
(3) Ist $f \in C^1(U, \mathbb{R}^n)$, so ist $f$ Lipschitz-stetig (auf $U$ bzgl. $U$). Ist $f \in C(U, \mathbb{R}^n)$ Lipschitz-stetig und $K \subset U$ kompakt, so ist $f$ gleichmäßig Lipschitz-stetig auf $K$.

BEWEIS. AD (1): Sei $(t_0, x_0) \in I \times U$ beliebig und $\varepsilon > 0$ so, dass $[t_0 - \varepsilon, t_0 + \varepsilon] \times \overline{B(x_0, \varepsilon)} \subset I \times U$ ist. Da $[t_0 - \varepsilon, t_0 + \varepsilon] \times \overline{B(x_0, \varepsilon)}$ kompakt und $D_x f \in C(I \times U, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n))$ ist, existiert

$$M = \max_{|t-t_0| \leq \varepsilon} \max_{x \in \overline{B(x_0, \varepsilon)}} \|D_x f(t, x)\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)}.$$

Für alle $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ und alle $x, y \in B(x_0, \varepsilon)$ folgt dann mit Satz VII.3.7(1) (S. 228)

$$\|f(t, x) - f(t, y)\| \leq M \|x - y\|.$$

AD (2): Ist offensichtlich.
AD (3): Ist offensichtlich. □

SATZ XI.1.6 (LEMMA VON GRONWALL). *Seien $\alpha, \beta, u \in C(I, \mathbb{R}_+)$ und $t_0 \in I$ mit*

$$u(t) \leq \alpha(t) + \Big| \int_{t_0}^t \beta(s) u(s) ds \Big| \quad \forall t \in I.$$

*Dann gilt*

$$u(t) \leq \alpha(t) + \Big| \int_{t_0}^t \alpha(s) \beta(s) e^{|\int_s^t \beta(\sigma) d\sigma)|} ds \Big| \quad \forall t \in I.$$

BEWEIS. Sei

$$v(t) = \int_{t_0}^t \beta(s) u(s) ds \quad \forall t \in I.$$

Aus den Voraussetzungen folgt, dass $v$ stetig differenzierbar ist und die Ungleichung

$$\begin{aligned} v'(t) &= \beta(t) u(t) \\ &\leq \beta(t)\alpha(t) + \beta(t) \Big| \int_{t_0}^t \beta(s) u(s) ds \Big| \\ &= \beta(t)\alpha(t) + \operatorname{sgn}(t - t_0) \beta(t) v(t) \quad \forall t \in I \end{aligned}$$

erfüllt. Multiplikation dieser Ungleichung mit

$$\gamma(t) = e^{-|\int_{t_0}^t \beta(s)ds|} = e^{-\int_{t_0}^t \operatorname{sgn}(s-t_0)\beta(s)ds}$$

liefert

$$\begin{aligned}\gamma(t)v'(t) &\le \alpha(t)\beta(t)\gamma(t) + \operatorname{sgn}(t-t_0)\beta(t)\gamma(t)v(t)\\ &= \alpha(t)\beta(t)\gamma(t) - \gamma'(t)v(t) \quad \forall t \in I\end{aligned}$$

und somit

$$(\gamma v)' \le \alpha\beta\gamma \quad \text{auf } I.$$

Integration von $t_0$ bis $t \in I$ liefert wegen $v(t_0) = 0$

$$\operatorname{sgn}(t-t_0)\gamma(t)v(t) \le \operatorname{sgn}(t-t_0)\int_{t_0}^t \alpha(s)\beta(s)\gamma(s)ds$$

und somit

$$\begin{aligned}\left|\int_{t_0}^t \beta(s)u(s)ds\right| &= \operatorname{sgn}(t-t_0)v(t)\\ &\le \operatorname{sgn}(t-t_0)\int_{t_0}^t \alpha(s)\beta(s)\gamma(s)\gamma(t)^{-1}ds\\ &= \left|\int_{t_0}^t \alpha(s)\beta(s)e^{|\int_s^t \beta(\sigma)d\sigma|}ds\right| \quad \forall t \in I.\end{aligned}$$

Hieraus folgt die Behauptung. □

Satz XI.1.7 (Satz von Picard-Lindelöf). *Sei $f \in C(I\times U, \mathbb{R}^n)$ auf $I\times U$ bzgl. $U$ Lipschitz-stetig. Dann gibt es zu jedem $(t_0, x_0) \in I\times U$ ein $\varepsilon > 0$, so dass das AWP*

$$\begin{aligned}x' &= f(t,x) \quad \textit{auf } (t_0-\varepsilon, t_0+\varepsilon)\\ x(t_0) &= x_0\end{aligned}$$

*eine eindeutige Lösung $x \in C^1((t_0-\varepsilon, t_0+\varepsilon), \mathbb{R}^n)$ hat.*

Beweis. 1. Variante: Aus dem Existenzsatz von Peano, Satz X.7.15 (S. 426), folgt, dass das AWP mindestens eine Lösung hat. Seien $x, y \in C^1((t_0-\varepsilon, t_0+\varepsilon), \mathbb{R}^n)$ zwei Lösungen des AWP. Indem wir $\varepsilon$ nötigenfalls verkleinern, folgt aus der Lipschitz-Stetigkeit von $f$

$$\begin{aligned}\|x(t)-y(t)\| &= \left\|\int_{t_0}^t f(s,x(s)) - f(s,y(s))ds\right\|\\ &\le \left|\int_{t_0}^t \|f(s,x(s)) - f(s,y(s))\|ds\right|\\ &\le \left|\int_{t_0}^t L\|x(s)-y(s)\|ds\right| \quad \forall |t-t_0| < \varepsilon.\end{aligned}$$

Diese Abschätzung und Satz XI.1.6 mit $\alpha = 0$ und $\beta = L$ beweisen die behauptete Eindeutigkeit.

2. Variante: Seien $\varepsilon_0 > 0$ und $R > 0$, so dass $[t_0 - \varepsilon_0, t_0 + \varepsilon_0] \times \overline{B(x_0, R)} \subset I \times U$ ist und $f$ auf $[t_0 - \varepsilon_0, t_0 + \varepsilon_0] \times \overline{B(x_0, R)}$ gleichmäßig Lipschitz-stetig ist bzgl. $\overline{B(x_0, R)}$. Sei $L$ die Lipschitz-Konstante von $f$ und

$$M = \max_{|t-t_0| \le \varepsilon_0} \max_{x \in \overline{B(x_0,R)}} \|f(t, x)\|.$$

Setze

$$\varepsilon = \min\Big\{\varepsilon_0, \frac{R}{M}\Big\}.$$

Sei $X = C([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n)$ und $K = B_{\|\cdot\|_\infty}(\overline{x}_0, R)$. Dabei bezeichnet $\|\cdot\|_\infty$ die Maximumsnorm auf $X$ und $\overline{x}_0$ die konstante Funktion mit Wert $x_0$. Durch

$$\|x\|_L = \max_{|t-t_0| \le \varepsilon} \Big\| e^{-L|t-t_0|} x(t) \Big\|$$

wird dann eine Norm auf $X$ definiert, bzgl. derer $X$ vollständig und $K$ abgeschlossen ist (Beweis Übungsaufgabe!). Definiere die Abbildung $\Phi : X \to X$ wie im Beweis von Satz X.7.15 (S. 426) durch

$$\Phi x(t) = x_0 + \int_{t_0}^{t} f(s, x(s)) ds \quad \forall |t - t_0| \le \varepsilon.$$

Offensichtlich ist $x$ ein Fixpunkt von $\Phi$ genau dann, wenn $x$ das AWP löst. Für $x, y \in X$ folgt

$$\begin{aligned} \|\Phi x(t) - x_0\| &= \Big\| \int_{t_0}^{t} f(s, x(s)) ds \Big\| \\ &\le M|t - t_0| \\ &\le M\varepsilon \\ &\le R \end{aligned}$$

und

$$\begin{aligned} \Big\| e^{-L|t-t_0|}[\Phi x(t) - \Phi y(t)] \Big\| &= \Big\| e^{-L|t-t_0|} \int_{t_0}^{t} f(s, x(s)) - f(s, y(s)) ds \Big\| \\ &\le e^{-L|t-t_0|} \operatorname{sgn}(t - t_0) \\ &\quad \cdot \int_{t_0}^{t} \|f(s, x(s)) - f(s, y(s))\| ds \\ &\le e^{-L|t-t_0|} \operatorname{sgn}(t - t_0) \int_{t_0}^{t} L\|x(s) - y(s)\| ds \\ &\le \operatorname{sgn}(t - t_0) \int_{t_0}^{t} e^{-L|t-t_0|} L e^{L|s-t_0|} \|x - y\|_L ds \\ &= (1 - e^{-L|t-t_0|}) \|x - y\|_L \\ &\le (1 - e^{-L\varepsilon}) \|x - y\|_L. \end{aligned}$$

Also ist $\Phi$ eine Kontraktion auf $K$ mit Kontraktionsrate $\kappa = 1 - e^{-L\varepsilon}$. Damit folgt die Behauptung aus dem Banachschen Fixpunktsatz, Satz IV.4.3 (S. 137). □

Bemerkung XI.1.8. Bemerkung X.7.16(1) (S. 427) zeigt, dass man ohne die vorausgesetzte Lipschitz-Stetigkeit i. a. nicht die Eindeutigkeit der Lösung des AWP erwarten kann. Bemerkung X.7.16(2) (S. 427) zeigt, dass man unter den Voraussetzungen von Satz XI.1.7 i. a. nur die lokale Existenz der Lösung des AWP erwarten kann. Satz XI.1.7 zeigt, dass es nicht immer ratsam ist, eine nicht autonome gDgl in die äquivalente autonome gDgl umzuformen. Für die nicht autonome gDgl benötigen wir nämlich nur Lipschitz-Stetigkeit bzgl. $x$, für die äquivalente autonome gDgl dagegen Lipschitz-Stetigkeit bzgl $x$ und $t$.

Satz XI.1.9 (Globaler Existenz- und Eindeutigkeitssatz). *Sei $f \in C(I \times U, \mathbb{R}^n)$ auf $I \times U$ bzgl. $U$ Lipschitz-stetig. Dann existiert für jedes $(t_0, x_0) \in I \times U$ genau eine nicht fortsetzbare Lösung $u(\cdot; t_0, x_0) \in C^1(I(t_0, x_0), U)$ des AWP*

$$(*) \qquad \begin{aligned} x' &= f(t,x) \\ x(t_0) &= x_0. \end{aligned}$$

*Das maximale Existenzintervall $I(t_0, x_0)$ ist offen, d.h.*

$$I(t_0, x_0) = (t^-(t_0, x_0), t^+(t_0, x_0)),$$

*und es gilt entweder*

$$t^- = t^-(t_0, x_0) = \inf I \quad \textit{bzw.} \quad t^+ = t^+(t_0, x_0) = \sup I$$

*oder*

$$\lim_{t \to t^\pm \mp 0} \min\{\operatorname{dist}(u(t; t_0, x_0), \partial U), \|u(t; t_0, x_0)\|^{-1}\} = 0.$$

*Dabei ist*

$$\operatorname{dist}(z, \emptyset) = +\infty$$

*und für eine nicht leere abgeschlossene Menge $A$*

$$\operatorname{dist}(z, A) = \inf\{\|z - y\| : y \in A\}.$$

Beweis. 1. Schritt: Sei $(t_0, x_0) \in I \times U$ beliebig und im Folgenden fest. Wegen Satz XI.1.7 existiert ein $\varepsilon_1 > 0$, so dass das AWP $(*)$ eine eindeutige Lösung $u$ auf $I_1 = [t_0 - \varepsilon_1, t_0 + \varepsilon_1]$ hat. Wegen Satz XI.1.7 existiert weiter ein $\varepsilon_2 > 0$, so dass das AWP

$$\begin{aligned} x' &= f(t,x) \\ x(t_0 + \varepsilon_1) &= u(t_0 + \varepsilon_1) \end{aligned}$$

eine eindeutige Lösung $v$ auf $I_{1,2} = [t_0 + \varepsilon_1 - \varepsilon_2, t_0 + \varepsilon_1 + \varepsilon_2]$ besitzt. Aus dem Beweis von Satz XI.1.7 folgt, dass auf $I_1 \cap I_{1,2}$ gilt $u = v$.

Folglich ist die durch

$$u_1 = \begin{cases} u & \text{auf } I_1, \\ v & \text{auf } I_{1,2}, \end{cases}$$

auf $I_1 \cup I_{1,2}$ definierte Funktion eine Lösung des AWP $(*)$, die $u$ echt fortsetzt. Da wir ganz analog für $t_0 - \varepsilon_1$ argumentieren können, zeigt dies, dass $u$ echt nach links und rechts fortgesetzt werden kann.

2. Schritt: Wegen des 1. Schrittes ist folgende Definition sinnvoll

$$t^+ = t^+(t_0, x_0) = \sup\{\beta \in I : (*) \text{ hat eine Lösung auf } [t_0, \beta]\},$$
$$t^- = t^-(t_0, x_0) = \inf\{\beta \in I : (*) \text{ hat eine Lösung auf } [\beta, t_0]\}.$$

Dann existiert genau eine Lösung $u = u(\cdot; t_0, x_0) \in C^1((t^-, t^+), U)$ von $(*)$ und $u$ kann nicht fortgesetzt werden. Das maximale Existenzintervall $I(t_0, x_0) = (t^-, t^+)$ ist offen, da wir sonst das Fortsetzungsargument aus Schritt 1 auf $(t^+, u(t^+; t_0, x_0))$ bzw. $(t^-, u(t^-; t_0, x_0))$ anwenden könnten.

3. Schritt: Wir nehmen an, dass $t^+ < \sup I$ ist, und wollen zeigen, dass

$$(**) \qquad \lim_{t \to t^+ - 0} \min\{\operatorname{dist}(u(t), \partial U), \|u(t)\|^{-1}\} = 0$$

ist. Wir nehmen dazu an, dass $(**)$ falsch ist. Dann gibt es ein $\varepsilon > 0$ und eine Folge $(t_n)_{n\in\mathbb{N}}$ mit $t_n < t^+$ für alle $n \in \mathbb{N}$, $\lim\limits_{n\to\infty} t_n = t^+$ und

$$\|u(t_n)\| \leq \frac{1}{2\varepsilon} \quad \text{und} \quad \operatorname{dist}(u(t_n); \partial U) \geq 2\varepsilon \quad \forall n \in \mathbb{N}.$$

Dabei ist o.E. $\varepsilon^2 \leq \frac{1}{2}$. Sei

$$M = \max\{\|f(t,x)\| : t_0 \leq t \leq t^+, \|x - x_0\| \leq \frac{1}{\varepsilon}, \operatorname{dist}(x, \partial U) \geq \varepsilon\}$$

und $0 < \delta < \frac{\varepsilon}{M}$.

Beh.: Für alle $n \in \mathbb{N}$ und alle $s \in [0, \min\{\delta, t^+ - t_n\}]$ gilt

$$\|u(t_n + s)\| < \frac{1}{\varepsilon} \quad \text{und} \quad \operatorname{dist}(u(t_n + s), \partial U) > \varepsilon.$$

Bew. der Beh.: Angenommen, die Beh. ist falsch. Dann gibt es ein $k \in \mathbb{N}^*$ und ein $\beta \in (0, \min\{\delta, t^+ - t_k\}]$ mit $\|u(t_k + s)\| \leq \frac{1}{\varepsilon}$ und $\operatorname{dist}(u(t_k + s), \partial U) \geq \varepsilon$ für alle $0 \leq s \leq \beta$ und

$$\|u(t_k + \beta)\| = \frac{1}{\varepsilon} \quad \text{oder} \quad \operatorname{dist}(u(t_k + \beta), \partial U) = \varepsilon.$$

Dann folgt

$$\|f(t_k + s, u(t_k + s))\| \leq M \quad \forall 0 \leq s \leq \beta$$

und somit

$$\begin{aligned}\|u(t_k + \beta) - u(t_k)\| &= \left\| \int_{t_k}^{t_k+\beta} f(s, u(s))ds \right\| \\ &\leq M\beta\end{aligned}$$

$$\begin{aligned}&\leq \delta M\\&< \varepsilon.\end{aligned}$$

Folglich gilt

$$\|u(t_k+\beta)\| \leq \varepsilon + \|u(t_k)\| \leq \varepsilon + \frac{1}{2\varepsilon} < \frac{1}{\varepsilon}$$

wegen $\varepsilon^2 \leq \frac{1}{2}$ und

$$\begin{aligned}\operatorname{dist}(u(t_k+\beta), \partial U) &\geq \operatorname{dist}(u(t_k), \partial U) - \|u(t_k+\beta) - u(t_k)\|\\&> 2\varepsilon - \varepsilon\\&= \varepsilon.\end{aligned}$$

Dies ist ein Widerspruch zur Wahl von $\varepsilon$.
Wegen der Beh. gilt für alle $k \in \mathbb{N}$ mit $t^+ - t_k \leq \delta$ und alle $s, t \in [t_k, t^+)$ die Abschätzung

$$\begin{aligned}\|u(t) - u(s)\| &= \left\| \int_s^t f(\tau, u(\tau)) d\tau \right\|\\&\leq M|t-s|.\end{aligned}$$

Ist also $(t'_n)_{n\in\mathbb{N}}$ eine beliebige Folge mit $t'_n < t^+$ für alle $n \in \mathbb{N}$ und $\lim\limits_{n\to\infty} t'_n = t^+$, so zeigt dies, dass $(u(t'_n))_{n\in\mathbb{N}}$ eine Cauchyfolge ist und somit gegen ein $y \in \mathbb{R}^n$ konvergiert. Da $\operatorname{dist}(u(t), \partial U) \geq \varepsilon$ ist für alle $t$ hinreichend nahe bei $t^+$, folgt $y \in U$. Ebenso folgt aus obiger Abschätzung, dass

$$\lim_{n\to\infty} \int_{t_0}^{t'_n} f(s, u(s)) ds$$

existiert. Sei nun $(t''_n)_{n\in\mathbb{N}}$ eine andere Folge mit $t''_n < t^+$ für alle $n \in \mathbb{N}$ und $\lim\limits_{n\to\infty} t''_n = t^+$. Dann folgt mit den gleichen Argumenten

$$u(t''_n) \underset{n\to\infty}{\longrightarrow} z \in U.$$

Aus obiger Abschätzung folgt

$$\begin{aligned}\|y - z\| &= \lim_{n\to\infty} \|u(t'_n) - u(t''_n)\|\\&\leq M \lim_{n\to\infty} |t'_n - t''_n|\\&= 0.\end{aligned}$$

Also gilt

$$y = \lim_{t\to t^+-0} u(t).$$

Ganz analog folgt, dass

$$\lim_{t\to t^+-0} \int_{t_0}^{t} f(s, u(s)) ds = \int_{t_0}^{t^+} f(s, u(s)) ds$$

gilt, d.h. dass rechts stehende uneigentliche Integral existiert. Definiere

$$v(t) = \begin{cases} u(t) & \text{für } t^- < t < t^+, \\ y & \text{für } t = t^+. \end{cases}$$

Dann folgt $v \in C((t^-, t^+], U)$ und

$$v(t) = x_0 + \int_{t_0}^{t} f(s, v(s))ds \quad \forall t^- < t \le t^+.$$

Also ist $v$ eine Lösung des AWP $(*)$, die $u$ fortsetzt. Dies ist ein Widerspruch. Also gilt $t^+ = \sup I$ oder

$$\lim_{t \to t^+ - 0} \min\{\operatorname{dist}(u(t), \partial U), \|u(t)\|^{-1}\} = 0.$$

Ganz analog verfahren wir für $t^-$. □

Bemerkung XI.1.10. Die Bezeichnungen und Voraussetzungen seien wie in Satz XI.1.9. Definiere

$$\gamma^+(t_0, x_0) = \{u(t; t_0, x_0) : t \in [t_0, t^+(t_0, x_0))\},$$
$$\gamma^-(t_0, x_0) = \{u(t; t_0, x_0) : t \in (t^-(t_0, x_0), t_0]\}.$$

Dann gilt:

(1) Ist $\gamma^\pm$ beschränkt, so ist $t^+ = \sup I$ bzw. $t^- = \inf I$ oder es gilt $\operatorname{dist}(u(t; t_0, x_0), \partial U) \to 0$ für $t \to t^\pm \mp 0$. D.h., die Lösung existiert entweder für alle Zeiten oder sie läuft gegen den Rand von $U$.
(2) Ist $\gamma^\pm$ in einer kompakten Menge enthalten, so ist $t^+ = \sup I$ bzw. $t^- = \inf I$.

Bemerkung XI.1.11. Die Funktion $f$ sei linear beschränkt, d.h. es gebe $\alpha, \beta \in C(I, \mathbb{R}_+)$ mit

$$\|f(t, x)\| \le \alpha(t)\|x\| + \beta(t) \quad \forall (t, x) \in I \times U.$$

Dann folgt aus dem Lemma von Gronwall, Satz XI.1.6, dass jede Lösung von $x' = f(t, x)$ beschränkt ist auf beschränkten Intervallen. Ist insbesondere $U = \mathbb{R}^n$, so besitzt das AWP $x' = f(t, x)$, $x(t_0) = x_0$ für alle $t_0 \in I, x_0 \in \mathbb{R}^n$ eine eindeutige globale Lösung.

## XI.2. Elementare Lösungsmethoden

Beispiel XI.2.1 (Exakte gDgl, integrierender Faktor). Sei $U \subset \mathbb{R}^2$ offen, nicht leer, $P, Q \in C^1(U, \mathbb{R})$ und $(x_0, y_0) \in U$ mit $Q(x_0, y_0) \ne 0$. Wir betrachten das AWP

$$\begin{aligned} y'(x) &= -\frac{P(x, y(x))}{Q(x, y(x))} \\ y(x_0) &= y_0 \end{aligned} \tag{XI.2.1}$$

oder äquivalent

$$\begin{aligned} Q(x,y(x))y'(x) + P(x,y(x)) &= 0 \\ y(x_0) &= y_0. \end{aligned}$$

Sei

$$\alpha = Pdx + Qdy.$$

Wir nehmen zunächst an, $\alpha$ sei geschlossen, d.h.

$$d\alpha = 0 \iff \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

Dann gibt es in einer Umgebung $V$ von $(x_0, y_0)$ in $U$ eine Funktion $F \in C^2(V, \mathbb{R}^2)$ mit

$$\alpha = dF \iff P = \frac{\partial F}{\partial x}, Q = \frac{\partial F}{\partial y}.$$

O.E. können wir $F$ so wählen, dass $F(x_0, y_0) = 0$ ist ($F$ ist eindeutig bis auf eine additive Konstante!). Wegen $Q(x_0, y_0) = \frac{\partial F}{\partial y}(x_0, y_0) \neq 0$ folgt aus dem Satz über implizite Funktionen, Satz VII.5.9 (S. 249), dass die Menge

$$N_0 = \{(x,y) \in V : F(x,y) = 0\}$$

in einer Umgebung von $(x_0, y_0)$ als Graph einer Funktion $y$ von $x$ dargestellt werden kann. Es gibt also ein $\varepsilon > 0$ und eine Funktion $y \in C^1((x_0 - \varepsilon, x_0 + \varepsilon), \mathbb{R})$ mit

$$F(x, y(x)) = 0 \quad \forall |x - x_0| < \varepsilon \text{ und } y(x_0) = y_0.$$

Aus der Kettenregel folgt, dass $y$ das AWP (XI.2.1) löst. Andererseits erfüllt Problem (XI.2.1) die Voraussetzungen von Satz XI.1.7 (S. 433), so dass $y$ die eindeutige Lösung ist. Die Lösung von Problem (XI.2.1) ist also durch die Niveaulinie einer Stammfunktion $F$ von $\alpha$ bestimmt. Da $\alpha = dF$, d.h. $\alpha$ exakt ist, nennt man Problem (XI.2.1) eine EXAKTE GDGL.

Wir betrachten nun den Fall $d\alpha \neq 0$. Wir nehmen an, dass es eine Funktion $M \in C^1(U, \mathbb{R})$ gibt, so dass $\alpha M$ geschlossen ist. $M$ heißt dann ein INTEGRIERENDER FAKTOR. $M$ muss die partielle Differentialgleichung

$$\begin{aligned} &\frac{\partial}{\partial y}(MP) = \frac{\partial}{\partial x}(MQ) \\ \iff\ & Q\frac{\partial M}{\partial x} - P\frac{\partial M}{\partial y} = M\Big\{\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\Big\} \end{aligned}$$

erfüllen. Da $\alpha M$ geschlossen ist, gibt es wieder eine Umgebung $V$ von $(x_0, y_0)$ in $U$ und eine Stammfunktion $\widetilde{F}$ von $\alpha M$, d.h.

$$\alpha M = d\widetilde{F} \iff MP = \frac{\partial \widetilde{F}}{\partial x}, MQ = \frac{\partial \widetilde{F}}{\partial y}.$$

Wie oben folgt daher, dass die Lösung von Problem (XI.2.1) durch die Niveaulinie von $\widetilde{F}$, auf der $(x_0, y_0)$ liegt, gegeben ist.

Beispiel XI.2.2 (Räuber-Beute-Modell). Wir betrachten das AWP aus Beispiel XI.1.3(2) (S. 430)

$$\begin{aligned} x' &= -\alpha_1 x + \beta_1 xy \\ y' &= \alpha_2 y - \beta_2 xy \\ x(0) &= x_0 \\ y(0) &= y_0 \end{aligned} \tag{XI.2.2}$$

mit $\alpha_1, \alpha_2, \beta_1, \beta_2, x_0, y_0 \in \mathbb{R}_+^*$. Die Funktion $f : \mathbb{R}_+^* \times \mathbb{R}_+^* \to \mathbb{R}^2$ mit $f(x,y) = (-\alpha_1 x + \beta_1 xy, \alpha_2 y - \beta_2 xy)$ ist offensichtlich $C^\infty$ und erfüllt wegen Bemerkung XI.1.5 (S. 432) die Voraussetzungen von Satz XI.1.7 (S. 433). Also gibt es ein $\varepsilon > 0$, so dass Gleichung (XI.2.2) eine eindeutige Lösung $(x,y) \in C^1((0,\varepsilon), \mathbb{R}_+^* \times \mathbb{R}_+^*)$ besitzt. Multiplizieren wir die erste Gleichung von Gleichung (XI.2.2) mit $\alpha_2 y - \beta_2 xy$ und die zweite mit $\alpha_1 x - \beta_1 xy$ und addieren beide Gleichungen, so folgt

$$y(\alpha_2 - \beta_2 x)x' + x(\alpha_1 - \beta_1 y)y' = 0.$$

Die Differentialform $\alpha = y(\alpha_2 - \beta_2 x)dx + x(\alpha_1 - \beta_1 y)dy$ ist nicht geschlossen. Wie man leicht nachprüft, ist aber $\frac{1}{xy}$ ein integrierender Faktor und

$$(xy)^{-1}\alpha = (\alpha_2 x^{-1} - \beta_2)dx + (\alpha_1 y^{-1} - \beta_1)dy = df$$

mit

$$f(x,y) = \alpha_2 \ln x - \beta_2 x + \alpha_1 \ln y - \beta_1 y.$$

Aus Beispiel XI.2.1 folgt daher, dass die Lösung von Gleichung (XI.2.2) auf einer geeigneten Niveaulinie von $f$ liegt. Da $\exp : \mathbb{R} \to \mathbb{R}_+^*$ ein Homöomorphismus ist, können wir äquivalent die Niveaulinien von

$$F = \exp \circ f = \frac{x^{\alpha_2}}{e^{\beta_2 x}} \cdot \frac{y^{\alpha_1}}{e^{\beta_1 y}} = \varphi(x) \cdot \psi(y)$$

betrachten. Sei also $c \in \mathbb{R}_+^*$ und

$$N_c = \{(x,y) \in \mathbb{R}^2 : F(x,y) = c\}.$$

Offensichtlich haben die Funktionen $\varphi$ und $\psi$ folgende Eigenschaften:

(1) $\varphi(0) = \psi(0) = 0$,
(2) $\lim\limits_{z\to\infty} \varphi(z) = \lim\limits_{z\to\infty} \psi(z) = 0$,
(3) $M_\varphi = \sup\limits_{z\in\mathbb{R}_+} \varphi(z) = \varphi(\dfrac{\alpha_2}{\beta_2})$,
(4) $M_\psi = \sup\limits_{z\in\mathbb{R}_+} \psi(z) = \psi(\dfrac{\alpha_1}{\beta_1})$,
(5) zu jedem $\mu \in (0, M_\varphi)$ bzw. $\mu \in (0, M_\psi)$ gibt es genau zwei Zahlen $z_{\varphi,\mu}^- < \frac{\alpha_2}{\beta_2} < z_{\varphi,\mu}^+$ bzw. $z_{\psi,\mu}^- < \frac{\alpha_1}{\beta_1} < z_{\psi,\mu}^+$ mit $\varphi(z_{\varphi,\mu}^\pm) = \mu$ bzw. $\psi(z_{\psi,\mu}^\pm) = \mu$.

Hieraus folgt unmittelbar:

(1) Ist $c > M_\varphi M_\psi$, so ist $N_c = \emptyset$.
(2) Ist $c = M_\varphi M_\psi$, so ist $N_c = \{(\frac{\alpha_2}{\beta_2}, \frac{\alpha_1}{\beta_1})\}$.
(3) Ist $c < M_\varphi M_\psi$, so ist $N_c$ eine geschlossene ellipsenförmige Kurve, die in dem Rechteck

$$\left[z^-_{\varphi,cM_\psi^{-1}}, z^+_{\varphi,cM_\psi^{-1}}\right] \times \left[z^-_{\psi,cM_\varphi^{-1}}, z^+_{\psi,cM_\varphi^{-1}}\right]$$

liegt.

Für die Lösung von Gleichung (XI.2.2) bedeutet dies wegen Bemerkung XI.1.10(2) (S. 438):

(1) Ist $(x_0, y_0) = (\frac{\alpha_2}{\beta_2}, \frac{\alpha_1}{\beta_1})$, so ist $(x(t), y(t))$ für alle $t \geq 0$ konstant.
(2) Ist $(x_0, y_0) \neq (\frac{\alpha_2}{\beta_2}, \frac{\alpha_1}{\beta_1})$, so existiert die Lösung von Gleichung (XI.2.2) für alle $t \geq 0$ und liegt auf der Niveaulinie $N_{F(x_0,y_0)}$.

Wir wollen zeigen, dass im Fall (2) die Lösung des AWP (XI.2.2) periodisch ist, d.h., dass es ein $T > 0$ gibt mit

$$(*) \qquad (x(t+T), y(t+T)) = (x(t), y(t)) \quad \forall t \in \mathbb{R}.$$

Dazu setzen wir $C = F(x_0, y_0)$ und bezeichnen mit

$$\gamma_0 = \{(x(t), y(t)) : t \in \mathbb{R}\}$$

die Trajektorie der Lösung von Gleichung (XI.2.2) zum Anfangswert $(x_0, y_0)$. Dann ist gemäß obigem $\gamma_0 \subset N_C$. Aus Satz XI.1.7 (S. 433) folgt, dass $\gamma_0$ relativ offen ist in $N_C$. Sei nun $(x_1, y_1) \in N_C \backslash \gamma_0$. Dann besitzt das AWP (XI.2.2) eine eindeutige Lösung zum Anfangswert $(x_1, y_1)$. Für die zugehörige Trajektorie $\gamma_1$ gilt wegen Satz XI.1.7 (S. 433) $\gamma_1 \cap \gamma_0 = \emptyset$ und $\gamma_1$ ist relativ offen in $N_C$. Also ist $\gamma_0$ auch relativ abgeschlossen in $N_C$. Da $N_C$ zusammenhängend ist, folgt $\gamma_0 = N_C$. Also gibt es ein $T > 0$ mit

$$(x(T), y(T)) = (x_0, y_0) = (x(0), y(0)).$$

Hieraus und aus Satz XI.1.7 (S. 433) folgt dann die Beziehung $(*)$.
Mit Hilfe der Periode $T$ können wir den Mittelwert der Räuber- bzw. Beutepopulation durch

$$\overline{x} = \frac{1}{T}\int_0^T x(t)dt \quad , \quad \overline{y} = \frac{1}{T}\int_0^T y(t)dt$$

definieren. Dividieren wir die erste bzw. zweite Gleichung von Problem (XI.2.2) durch $x$ bzw. $y$ und integrieren wir beide Gleichungen von 0 bis $T$, so erhalten wir wegen der Periodizität

$$\begin{aligned} 0 &= \ln x(T) - \ln x(0) \\ &= \int_0^T \frac{x'(t)}{x(t)} dt \\ &= \int_0^T (-\alpha_1 + \beta_1 y(t)) dt \end{aligned}$$

$$= -\alpha_1 T + \beta_1 T \overline{y}$$

und

$$\begin{aligned} 0 &= \ln y(T) - \ln y(0) \\ &= \int_0^T \frac{y'(t)}{y(t)} dt \\ &= \int_0^T (\alpha_2 - \beta_2 x(t)) dt \\ &= \alpha_2 T - \beta_2 T \overline{x}. \end{aligned}$$

Also gilt für die Mittelwerte der Populationen

$$\overline{x} = \frac{\alpha_2}{\beta_2}, \quad \overline{y} = \frac{\alpha_1}{\beta_1},$$

d.h., die mittlere Populationsdichte der Räuberspezies (Beutespezies) hängt nur von dem Verhältnis der relativen Geburts- und Sterberaten der Beutespezies (Räuberspezies) ab. Dies hat eine interessante Konsequenz: Nehmen wir an, wir wollen die Beutespezies künstlich durch ein Gift reduzieren, das - als Nebenwirkung - auch die Räuberspezies angreift. Dann müssen wir in Gleichung (XI.2.2) $\alpha_1$ und $\alpha_2$ durch $\alpha_1 + \varepsilon_1$ bzw. $\alpha_2 - \varepsilon_2$ mit $\varepsilon_1, \varepsilon_2 > 0$ ersetzen. Falls das Gift so schwach ist, dass $\alpha_2 - \varepsilon_2 > 0$ ist, können wir unsere bisherigen Überlegungen anwenden und erhalten die neuen Mittelwerte

$$\overline{x} = \frac{\alpha_2 - \varepsilon_2}{\beta_2} \quad , \quad \overline{y} = \frac{\alpha_1 + \varepsilon_1}{\beta_1}.$$

D.h., unser Eingriff ist kontraproduktiv und die Beutespezies nimmt im Mittel sogar zu.

Beispiel XI.2.3 (Trennung der Variablen). Seien $I \subset \mathbb{R}$, $J \subset \mathbb{R}$ offene, nicht leere Intervalle, $f \in C(I, \mathbb{R})$, $g \in C(J, \mathbb{R})$ und $x_0 \in I$, $y_0 \in J$ mit $g(y_0) \neq 0$. Dann betrachten wir das AWP

$$\begin{aligned} y' &= f(x)g(y) \quad \text{auf } I \\ y(x_0) &= y_0. \end{aligned} \tag{XI.2.3}$$

Dies ist ein Spezialfall von Beispiel XI.2.1 mit $U = I \times J$ und

$$P(x,y) = f(x) \quad , \quad Q(x,y) = -\frac{1}{g(y)}.$$

Offensichtlich ist die zugehörige Differentialform

$$\alpha = f(x)dx - \frac{1}{g(y)}dy$$

geschlossen und hat die Stammfunktion

$$F(x,y) = \int_{x_0}^x f(u)du - \int_{y_0}^y \frac{1}{g(v)}dv$$

mit $F(x_0, y_0) = 0$. Daher ist die Lösung von Gleichung (XI.2.3) in einer Umgebung von $x_0$ gegeben durch die Niveaulinie

$$N_0 = \{(x, y) \in \mathbb{R}^2 : F(x, y) = 0\}.$$

Wegen Satz VII.5.9 (S. 249) und Satz IV.2.11 (S. 124) können wir die Lösung in einer Umgebung von $x_0$ darstellen als

$$y(x) = \varphi^{-1} \circ \psi(x)$$

mit

$$\begin{aligned}\psi(x) &= \int_{x_0}^{x} f(u)du,\\ \varphi(y) &= \int_{y_0}^{y} \frac{1}{g(v)} dv.\end{aligned}$$

Beispiel XI.2.4 (Chemische Reaktionskinetik). Wir betrachten das AWP

$$\text{(XI.2.4)} \qquad \begin{aligned} u' &= r(a-u)(b-u) \quad , t > 0\\ u(0) &= 0\end{aligned}$$

aus Beispiel XI.1.3(3) (S. 430) mit $r, a, b \in \mathbb{R}_+^*$. Dies ist eine Gleichung mit getrennten Veränderlichen wie in Beispiel XI.2.3 mit

$$f(x) = r \quad , \quad g(y) = (a-y)(b-y).$$

Wir betrachten zunächst den Fall $a \neq b$. Durch elementare Integration bzw. Partialbruchzerlegung folgt

$$\begin{aligned}\psi(x) &= \int_0^x r du\\ &= rx\\ \varphi(y) &= \int_0^y \frac{1}{(a-v)(b-v)} dv\\ &= \frac{1}{a-b} \int_0^y \Big[\frac{1}{b-v} - \frac{1}{a-v}\Big] dv\\ &= \frac{1}{a-b} \Big[\ln\Big(\frac{a-y}{b-y}\Big) - \ln\Big(\frac{a}{b}\Big)\Big].\end{aligned}$$

Also lautet die Lösung von Gleichung (XI.2.4)

$$u(t) = ab \frac{1 - e^{r(a-b)t}}{b - ae^{r(a-b)t}}.$$

Insbesondere folgt durch elementare Rechnung

$$\lim_{t\to\infty} u(t) = \min\{a, b\}.$$

Betrachte nun den Fall $a = b$. Dann folgt

$$\psi(x) = rx$$

$$\begin{aligned}\varphi(y) &= \int_0^y \frac{1}{(a-v)^2}dv \\ &= \frac{1}{a-v} - \frac{1}{a}.\end{aligned}$$

Also lautet die Lösung von Gleichung (XI.2.4)

$$u(t) = a - \frac{1}{\frac{1}{a} + rt},$$

und es gilt wieder

$$\lim_{t\to\infty} u(t) = a.$$

Beispiel XI.2.5 (Populationsdynamik). Wir betrachten das AWP

$$\begin{aligned} y' &= \alpha y(y^* - y) \quad ,t > 0 \\ y(0) &= y_0 \end{aligned} \tag{XI.2.5}$$

aus Beispiel XI.1.3(1) (S. 430) mit $\alpha, y^*, y_0 \in \mathbb{R}_+^*$. Beispiel XI.2.3 liefert mit $f(u) = \alpha$ und $g(v) = v(y^* - v)$:

$$\begin{aligned}\psi(x) &= \int_0^x \alpha du \\ &= \alpha x \\ \varphi(y) &= \int_{y_0}^y \frac{1}{v(y^*-v)}dv \\ &= \frac{1}{y^*}\int_{y_0}^y \Big[\frac{1}{y^*-v} + \frac{1}{v}\Big]dv \\ &= \frac{1}{y^*}\Big[\ln\Big|\frac{y}{y^*-y}\Big| - \ln\Big|\frac{y_0}{y^*-y_0}\Big|\Big].\end{aligned}$$

Also lautet die Lösung von Gleichung (XI.2.5)

$$y(t) = y^* \frac{y_0 e^{\alpha y^* t}}{y^* - y_0 + y_0 e^{\alpha y^* t}}.$$

Insbesondere konvergiert $y(t)$ für $t \to \infty$ gegen die Grenzpopulation $y^*$, und zwar monoton wachsend, wenn $y_0 < y^*$ ist, und monoton fallend, wenn $y_0 > y^*$ ist.

Beispiel XI.2.6 (Einfluss des Luftwiderstandes). Das AWP

$$\begin{aligned} v' &= \frac{K}{m} - \frac{r}{m}v^2 \\ v(0) &= 0 \end{aligned} \tag{XI.2.6}$$

aus Beispiel XI.1.3(5) (S. 430) kann wegen

$$\frac{K}{m} - \frac{r}{m}v^2 = \Big(\sqrt{\frac{K}{m}} - \sqrt{\frac{r}{m}}v\Big)\Big(\sqrt{\frac{K}{m}} + \sqrt{\frac{r}{m}}v\Big)$$

$$= -\frac{r}{m}\Big(\sqrt{\frac{K}{r}} - v\Big)\Big(-\sqrt{\frac{K}{r}} - v\Big)$$

mit den gleichen Methoden gelöst werden wie Beispiel XI.2.4. Wir erhalten die Lösung

$$v(t) = \sqrt{\frac{K}{r}}\frac{1 - e^{-2\sqrt{\frac{Kr}{m}}t}}{1 + e^{-2\sqrt{\frac{Kr}{m}}t}}$$

und

$$\lim_{t\to\infty} v(t) = \sqrt{\frac{K}{r}},$$

d.h., um eine Verdoppelung der Grenzgeschwindigkeit zu erhalten, muss man die Antriebskraft vervierfachen oder den Luftwiderstand vierteln.

Beispiel XI.2.7 (Variation der Konstanten). Seien $a, b \in C(\mathbb{R}, \mathbb{R})$ und $t_0, x_0 \in \mathbb{R}$. Wir betrachten das AWP

$$\begin{aligned} x' &= a(t)x + b(t) \\ x(t_0) &= x_0. \end{aligned} \tag{XI.2.7}$$

Wegen Satz XI.1.7 (S. 433) und Bemerkung XI.1.11 (S. 438) besitzt dieses AWP eine eindeutige Lösung $x \in C^1(\mathbb{R}, \mathbb{R})$.
Zur Konstruktion dieser Lösung betrachten wir zunächst den Fall $b = 0$. Da wegen Satz XI.1.7 (S. 433) das AWP

$$\begin{aligned} y' &= a(t)y \\ y(t_1) &= 0 \end{aligned}$$

für beliebiges $t_1$ stets die eindeutige Lösung $y = 0$ besitzt, folgt für die Lösung von Gleichung (XI.2.7) mit $b = 0$:

(1) $x_0 < 0 \iff x(t) < 0 \quad \forall t \in \mathbb{R}$,
(2) $x_0 = 0 \iff x(t) = 0 \quad \forall t \in \mathbb{R}$,
(3) $x_0 > 0 \iff x(t) > 0 \quad \forall t \in \mathbb{R}$.

Sei o.E. $x_0 > 0$. Dann können wir Gleichung (XI.2.7) mit $b = 0$ durch $x(t)$ dividieren und erhalten

$$\begin{aligned} a(t) &= \frac{x'(t)}{x(t)} \\ &= \frac{d}{dt}\ln x(t) \end{aligned}$$

und somit

$$\ln x(t) - \ln x_0 = \int_{t_0}^{t} a(s)ds$$

also

$$x(t) = x_0 \exp\Big\{\int_{t_0}^{t} a(s)ds\Big\}.$$

Offensichtlich ist dies auch die eindeutige Lösung in Fällen $x_0 = 0$ und $x_0 < 0$.
Wir betrachten nun den Fall $b \neq 0$. Motiviert durch das Ergebnis im Fall $b = 0$ machen wir den folgenden Ansatz, genannt VARIATION DER KONSTANTEN:

$$x(t) = c(t) \exp\Big\{ \int_{t_0}^{t} a(s) ds \Big\}$$

mit $c \in C^1(\mathbb{R}, \mathbb{R})$. Einsetzen in Gleichung (XI.2.7) liefert

$$x_0 = x(t_0) = c(t_0)$$

und

$$\begin{aligned} b(t) &= x' - a(t)x \\ &= c' \exp\Big\{ \int_{t_0}^{t} a(s) ds \Big\}. \end{aligned}$$

Hieraus folgt

$$\begin{aligned} c' &= b(t) \exp\Big\{ - \int_{t_0}^{t} a(s) ds \Big\} \\ c(t_0) &= x_0 \end{aligned}$$

und daher

$$c(t) = x_0 + \int_{t_0}^{t} b(s) \exp\Big\{ - \int_{t_0}^{s} a(\tau) d\tau \Big\} ds.$$

Also lautet die Lösung von Gleichung (XI.2.7)

$$x(t) = x_0 \exp\Big\{ \int_{t_0}^{t} a(s) ds \Big\} + \int_{t_0}^{t} b(s) \exp\Big\{ \int_{s}^{t} a(\tau) d\tau \Big\} ds.$$

BEISPIEL XI.2.8 (AUTONOME LINEARE DIFFERENTIALGLEICHUNGEN). Seien $A \in M_{n,n}(\mathbb{R})$ und $x_0 \in \mathbb{R}^n$. Wir betrachten das AWP

$$\begin{aligned} x' &= Ax \\ x(0) &= x_0. \end{aligned} \tag{XI.2.8}$$

Aus Satz XI.1.7 (S. 433) und Bemerkung XI.1.11 (S. 438) folgt, dass es eine eindeutige globale Lösung $x \in C^1(\mathbb{R}, \mathbb{R}^n)$ besitzt.
Zur Konstruktion dieser Lösung überlegen wir uns zunächst, dass wir o.E. eine komplexe Lösung von Gleichung (XI.2.8) betrachten können. Sei nämlich $z \in C^1(\mathbb{R}, \mathbb{C}^n)$ eine Lösung von Gleichung (XI.2.8). Dann folgt für $u = \operatorname{Re} z$ und $v = \operatorname{Im} z$:

$$\begin{aligned} u, v &\in C^1(\mathbb{R}, \mathbb{R}^n), \\ u' &= Au, \\ u(0) &= x_0, \\ v' &= Av, \end{aligned}$$

$$v(0) = 0.$$

Aus Satz XI.1.7 (S. 433) und Bemerkung XI.1.11 (S. 438) folgt $v(t) = 0$ für alle $t \in \mathbb{R}$, und $u$ ist die eindeutige reelle Lösung von Gleichung (XI.2.8).

Sei also $z \in C^1(\mathbb{R}, \mathbb{C}^n)$ die eindeutige Lösung von Gleichung (XI.2.8). Aus der Linearen Algebra ist bekannt, dass wir $A$ auf Jordansche Normalform transformieren können. D.h., es gibt natürliche Zahlen $k, n_1, \dots,$
$n_k \in \mathbb{N}^*$ mit $n_1 + \ldots + n_k = n$ und komplexe Zahlen $\lambda_1, \ldots \lambda_k \in \mathbb{C}$ und eine reguläre Matrix $U \in M_{n,n}(\mathbb{C})$ mit

$$UAU^{-1} = J = \operatorname{diag}(J(\lambda_1), \dots, J(\lambda_k))$$

und

$$J(\lambda_i) = \begin{cases} \lambda_i & \text{falls } n_i = 1, \\ \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix} \in M_{n_i,n_i}(\mathbb{C}), & \text{falls } n_i > 1. \end{cases}$$

Die Zahlen $\lambda_1, \dots, \lambda_k$ sind die Eigenwerte von $A$. Es ist $n_i > 1$ genau dann, wenn die algebraische Vielfachheit des Eigenwertes $\lambda_i$ größer ist als seine geometrische Vielfachheit. Setze

$$\widetilde{z} = Uz \ , \ \widetilde{z}_0 = Uz_0.$$

Dann folgt

$$\begin{aligned} \widetilde{z}' &= Uz' = UAz \\ &= J\widetilde{z} \\ \widetilde{z}(0) &= \widetilde{z}_0. \end{aligned} \tag{XI.2.9}$$

Da $J$ Block-diagonal ist, zerfällt das AWP (XI.2.9) in $k$ voneinander unabhängige (entkoppelte) AWPe in $\mathbb{R}^m$ der Form

$$\begin{aligned} w' &= J(\lambda)w \\ w(0) &= w_0 \end{aligned} \tag{XI.2.10}$$

mit $\lambda = \lambda_i$, $m = n_i$ für ein $i \in \mathbb{N}_k^*$.

Sei zunächst $m = 1$. Dann folgt für die Lösung von Gleichung (XI.2.10)

$$w(t) = w_0 e^{\lambda t}.$$

Insbesondere gilt:

(1) $\operatorname{Re}\lambda < 0 \Longrightarrow \lim\limits_{t\to\infty} w(t) = 0$,
(2) $\operatorname{Re}\lambda = 0 \Longrightarrow |w(t)|$ ist konstant für alle $t \in \mathbb{R}$,
(3) $\operatorname{Re}\lambda > 0 \Longrightarrow \lim\limits_{t\to\infty} |w(t)| = \infty$ (sofern $w_0 \neq 0$).

Sei nun $m > 1$. Dann gilt für die Komponenten $w_1, \dots, w_m$ der Lösung von Gleichung (XI.2.10):

$$\begin{aligned} w_m' &= \lambda w_m, \\ w_m'(0) &= w_{0,m}, \\ w_l' &= \lambda w_l + w_{l+1}, \\ w_l(0) &= w_{0,l}, \quad l = m-1, \dots, 1. \end{aligned}$$

Aus Beispiel XI.2.7 folgt daher rekursiv:

$$\begin{aligned} w_m(t) &= w_{0,m} e^{\lambda t}, \\ w_l(t) &= w_{0,l} e^{\lambda t} + \int_0^t w_{l+1}(s) e^{\lambda(t-s)} ds \quad , l = m-1, \dots, 1. \end{aligned}$$

Hieraus folgt durch Induktion für $l \in \mathbb{N}_m^*$

$$w_l(t) = e^{\lambda t} \sum_{j=0}^{m-l} w_{0,l+j} \frac{1}{j!} t^j.$$

Also gilt:

(4) $\operatorname{Re} \lambda < 0 \Longrightarrow \lim_{t\to\infty} w_l(t) = 0 \quad \forall l \in \mathbb{N}_m^*$,

(5) $\operatorname{Re} \lambda \geq 0 \Longrightarrow \lim_{t\to\infty} \left\{ \sum_{l=1}^m |w_l(t)|^2 \right\}^{1/2} = \infty$ (sofern $w_0 \neq 0$).

Setzen wir die Lösungen der AWPe (XI.2.10) geeignet zusammen und transformieren zurück, erhalten wir die Lösung $x = U^{-1}\tilde{z}$ von Gleichung (XI.2.8). Aus den Eigenschaften (1) – (5) folgt für das asymptotische Verhalten von $x$ unabhängig vom Anfangswert $x_0$:

- $\lim_{t\to\infty} x(t) = 0$, falls alle Eigenwerte von $A$ negativen Realteil haben,
- $\|x(t)\|$ ist beschränkt, falls alle Eigenwerte von $A$ nicht-positiven Realteil haben und bei den Eigenwerten mit Realteil 0 die geometrische und algebraische Vielfachheit übereinstimmen.

Insbesondere gilt stets $\lim_{t\to\infty} x(t) = 0$, falls $A$ symmetrisch und negativ definit ist.

## XI.3. Stetigkeits- und Differenzierbarkeitssätze

Seien $I \subset \mathbb{R}$ ein offenes, nicht leeres Intervall, $U \subset \mathbb{R}^n, n \in \mathbb{N}^*$, eine offene, nicht leere Menge und $f \in C(I \times U, \mathbb{R}^n)$ auf $I \times U$ Lipschitz-stetig bzgl. $U$. Aufgrund Satz XI.1.7 (S. 433) besitzt das AWP

$$\begin{aligned} x' &= f(t, x(t)) \\ x(t_0) &= x_0 \end{aligned}$$

für jedes $(t_0, x_0) \in I \times U$ in einer Umgebung von $t_0$ eine eindeutige Lösung $x = x(t; t_0, x_0)$. Wir wollen zeigen, dass $x(\cdot; t_0, x_0)$ stetig und –

unter zusätzlichen Annahmen an $f$ – differenzierbar von $x_0$ abhängt. Dazu benutzen wir zunächst, dass aus dem Beweis von Satz XI.1.7 (S. 433) folgt, dass es zu jedem $(t_0, x_0) \in I \times U$ zwei Umgebungen $J = J(t_0) \in \mathcal{U}(t_0)$ und $V = V(x_0) \in \mathcal{U}(x_0)$ mit den folgenden Eigenschaften gibt:

(1) $f$ ist gleichmäßig Lipschitz-stetig auf $J \times V$.
(2) Zu jedem $x_1 \in V$ besitzt das AWP

$$\begin{aligned} x' &= f(t, x(t)) \\ x(t_0) &= x_1 \end{aligned}$$

eine eindeutige Lösung $x(\cdot; t_0, x_1) \in C^1(J, V)$ auf $J$.

Daher können wir durch $\varphi(z) = x(\cdot; t_0, z)$ eine Abbildung $\varphi : V \to C^1(J, V)$ definieren. Dann können wir unser Ziel wie folgt formulieren: $\varphi \in C(V, C(J, V))$ bzw. – unter zusätzlichen Bedingungen an $f$ – $\varphi \in C^1(V, C(J, V))$. Dazu nehmen wir zur Vereinfachung an, dass $J$ beschränkt ist.

Satz XI.3.1 (Stetige Abhängigkeit von den Anfangswerten). *Die Funktion $\varphi$ ist gleichmäßig Lipschitz-stetig auf $V$:*

$$\|\varphi(z_1) - \varphi(z_2)\|_{C(J,V)} \le e^{Ld} \|z_1 - z_2\|_{\mathbb{R}^n} \quad \forall z_1, z_2 \in V.$$

*Dabei ist $d = \sup\{|s - t| : s, t \in J\}$ die Länge von $J$ und $L$ die Lipschitz-Konstante von $f$.*

Beweis. Seien $z_1, z_2 \in V$ und $t \in J$. Dann gilt

$$\begin{aligned} x(t; t_0, z_1) - x(t; t_0, z_2) &= z_1 - z_2 \\ &\quad + \int_{t_0}^{t} \Big[ f(s, x(s; t_0, z_1)) - f(s, x(s; t_0, z_2)) \Big] ds, \end{aligned}$$

und aus der Lipschitz-Stetigkeit von $f$ folgt

$$\begin{aligned} \|x(t; t_0, z_1) - x(t; t_0, z_2)\| &\le \|z_1 - z_2\| \\ &\quad + \Big| \int_{t_0}^{t} L \|x(t; t_0, z_1) - x(t; t_0, z_2)\| ds \Big|. \end{aligned}$$

Zusammen mit dem Lemma von Gronwall, Satz XI.1.6 (S. 432), liefert dies

$$\begin{aligned} \|x(t; t_0, z_1) - x(t; t_0, z_2)\| &\le \|z_1 - z_2\| \\ &\quad + \Big| \int_{t_0}^{t} \|z_1 - z_2\| L e^{L|t-s|} ds \Big| \\ &= \|z_1 - z_2\| e^{L|t-t_0|} \end{aligned}$$

und somit

$$\|\varphi(z_1) - \varphi(z_2)\|_{C(J,V)} = \sup_{t \in J} \|x(t; t_0, z_1) - x(t; t_0, z_2)\|$$

$$\leq e^{Ld}\|z_1 - z_2\|.$$

$\square$

Satz XI.3.2 (Differenzierbare Abhängigkeit von den Anfangswerten). *Die Funktion $f(t,x)$ sei zusätzlich auf $J \times V$ stetig differenzierbar bzgl. der Variablen $x$ und $D_x f$ Lipschitz-stetig bzgl. $V$. Dann ist $\varphi$ stetig differenzierbar:*

$$(D\varphi(z)w)(t) = Z(t; t_0, z)w \quad \forall t \in J, z \in V, w \in \mathbb{R}^n.$$

*Dabei ist $Z(\cdot; t_0, z) \in C^1(J, M_{n,n}(\mathbb{R}))$ die eindeutige Lösung des linearen AWP*

$$(*) \qquad \begin{aligned} Z' &= D_x f(t, x(t; t_0, z))Z \\ Z(t_0) &= Id_{\mathbb{R}^n}. \end{aligned}$$

Beweis. Aus den Voraussetzungen an $f$ und aus Satz XI.1.7 (S. 433) folgt, dass das AWP $(*)$ eine eindeutige Lösung hat. Durch Verkleinern von $J$ und $V$ können wir erreichen, dass gilt:

(1) $f$ ist gleichmäßig Lipschitz-stetig auf $J \times V$ mit Lipschitz-Konstanter $L$,
(2) $D_x f$ ist gleichmäßig Lipschitz-stetig auf $J \times V$ mit Lipschitz-Konstanter $K$,
(3) $\sup_{t\in J} \sup_{v \in V} \|D_x f(t,v)\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)} \leq M$.

Seien $\delta > 0$ und $\varepsilon > 0$ so gewählt, dass gilt

$$[t_0 - \delta, t_0 + \delta] \subset J \quad \text{und} \quad \overline{B(x_0, \varepsilon)} \subset V.$$

Für jedes $z_1, z_2 \in \overline{B(x_0, \varepsilon)}$ und jedes $t \in [t_0 - \delta, t_0 + \delta]$ folgt dann aus Satz XI.3.1

$$\|x(t; t_0, z_1) - x(t; t_0, z_2)\| \leq e^{L\delta}\|z_1 - z_2\|.$$

Indem wir $\delta$ hinreichend klein wählen, können wir somit erreichen, dass gilt

$$x(t; t_0, z_i) \in V \quad \forall |t - t_0| \leq \delta, i = 1, 2.$$

Hieraus und aus $(*)$ folgt

$$\begin{aligned} &\|Z(t; t_0, z_1)\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)} \\ &\leq 1 + \left|\int_{t_0}^t \|D_x f(s, x(s; t_0, z_1))\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)} \|Z(s; t_0, z_1)\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)} ds\right| \\ &\leq 1 + \left|\int_{t_0}^t M\|Z(s; t_0, z_1))\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)} ds\right|. \end{aligned}$$

Das Lemma von Gronwall, Satz XI.1.6 (S. 432), liefert daher die Schranke

$$\|Z(t; t_0, z_1)\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)} \leq Me^{M\delta} \quad \forall |t - t_0| \leq \delta.$$

Weiter gilt

$$x(t; t_0, z_2) - x(t; t_0, z_1) - Z(t; t_0, z_1)(z_2 - z_1)$$

$$
\begin{aligned}
&= z_2 + \int_{t_0}^{t} f(s, x(s; t_0, z_2))ds \\
&\quad - z_1 - \int_{t_0}^{t} f(s, x(s; t_0, z_1))ds \\
&\quad - (z_2 - z_1) - \int_{t_0}^{t} D_x f(s, x(s; t_0, z_1)) Z(s; t_0, z_1)(z_2 - z_1)ds \\
&= \int_{t_0}^{t} \int_0^1 \Big[ D_x f(s, x(s; t_0, z_1) + \theta(x(s; t_0, z_2) - x(s; t_0, z_1))) \Big] \\
&\qquad\qquad \Big[ x(s; t_0, z_2) - x(s; t_0, z_1) - Z(s; t_0, z_1)(z_2 - z_1) \Big] d\theta ds \\
&\quad + \int_{t_0}^{t} \int_0^1 \Big[ D_x f(s, x(s; t_0, z_1) + \theta(x(s; t_0, z_2) - x(s; t_0, z_1))) \\
&\qquad\qquad - D_x f(s, x(s; t_0, z_1)) \Big] Z(s; t_0, z_1)(z_2 - z_1) d\theta ds.
\end{aligned}
$$

Hieraus und den schon bewiesenen Abschätzungen folgt

$$
\begin{aligned}
&\|x(t; t_0, z_2) - x(t; t_0, z_1) - Z(t; t_0, z_1)(z_2 - z_1)\| \\
&\le \Big| \int_{t_0}^{t} \int_0^1 M \|x(s; t_0, z_2) - x(s; t_0, z_1) - Z(s; t_0, z_1)(z_2 - z_1)\| d\theta ds \Big| \\
&\quad + \Big| \int_{t_0}^{t} \int_0^1 \theta K \|x(s; t_0, z_2) - x(s; t_0, z_1)\| \|Z(s; t_0, z_1)\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)} \cdot \\
&\qquad\qquad \cdot \|z_2 - z_1\| d\theta ds \Big| \\
&\le \Big| \int_{t_0}^{t} M \|x(s; t_0, z_2) - x(s; t_0, z_1) - Z(s; t_0, z_1)(z_2 - z_1) ds \Big| \\
&\quad + \Big| \int_{t_0}^{t} K e^{L\delta} M e^{M\delta} \|z_2 - z_1\|^2 ds \Big| \\
&\le \delta K M e^{(L+M)\delta} \|z_2 - z_1\|^2 \\
&\quad + \Big| \int_{t_0}^{t} M \|x(s; t_0, z_2) - x(s; t_0, z_1) - Z(s; t_0, z_1)(z_2 - z_1)\| ds \Big|.
\end{aligned}
$$

Das Lemma von Gronwall, Satz XI.1.6 (S. 432), liefert daher die Abschätzung

$$
\begin{aligned}
&\sup_{|t - t_0| \le \delta} \|x(t; t_0, z_2) - x(t; t_0, z_1) - Z(t; t_0, z_1)(z_2 - z_1)\| \\
&\le M e^{M\delta} \delta K M e^{(L+M)\delta} \|z_2 - z_1\|^2,
\end{aligned}
$$

d.h.

$$
\lim_{\|z_2 - z_1\| \to 0} \frac{1}{\|z_2 - z_1\|} \|\varphi(z_2) - \varphi(z_1) - Z(\cdot; t_0, z_1)(z_2 - z_1)\|_{C(J,V)} = 0.
$$

Hieraus und aus der Definition der Differenzierbarkeit folgt die Behauptung. □

BEMERKUNG XI.3.3. Die Ergebnisse der Sätze XI.3.1 und XI.3.2 erlauben die Lösung von RANDWERTPROBLEMEN, d.h. von Dglen

$$x' = f(t, x) \quad \text{auf } [a, b] \tag{XI.3.1}$$

unter der RANDBEDINGUNG

$$r(x(a), x(b)) = 0 \tag{XI.3.2}$$

mit $r \in C(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^n)$. Mit der Notation dieses Paragraphen ist nämlich die Lösung von Gleichungen (XI.3.1), (XI.3.2) äquivalent dazu, einen Anfangswert $z \in U$ mit

$$r(z, x(b; a, z)) = 0 \tag{XI.3.3}$$

zu finden. Diese Gleichung kann mit einer Fixpunktiteration gelöst werden. Falls sogar $r$ differenzierbar ist und $f$ die Voraussetzungen von Satz XI.3.2 erfüllt, kann Problem (XI.3.3) mit dem Newtonverfahren gelöst werden. In jedem Newtonschritt ist dann ein lineares AWP der Form $(*)$ zu lösen.
Die Frage der Existenz und Eindeutigkeit einer Lösung des Randwertproblems (XI.3.1), (XI.3.2) kann nicht so einfach beantwortet werden wie bei AWPen. Um dies einzusehen, betrachte das Beispiel

$$\begin{aligned} u' &= v \\ v' &= -\omega^2 u \\ u(0) &= 0 \\ u(1) &= 1 \end{aligned}$$

mit $\omega \in \mathbb{R}^*$. (Die äquivalente gDgl 2. Ordnung lautet offensichtlich $u'' = -\omega^2 u$.) Für einen beliebigen Anfangswert $z = (\alpha, \beta) \in \mathbb{R}^2$ erhalten wir mit Beispiel XI.2.8 (S. 446) die Lösung

$$x(t; 0, z) = \begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\omega \sin(\omega t) & \omega \cos(\omega t) \end{pmatrix} \begin{pmatrix} \alpha \\ \frac{\beta}{\omega} \end{pmatrix}.$$

Also ergibt Problem (XI.3.3) das lineare Gleichungssystem

$$\begin{aligned} \alpha &= 0 \\ \alpha \cos \omega + \frac{\beta}{\omega} \sin \omega &= 1, \end{aligned}$$

woraus die Bedingung

$$\beta \sin \omega = \omega$$

folgt. Falls $\omega \in \pi\mathbb{Z}\backslash\{0\}$ ist, besitzt diese Gleichung offensichtlich keine Lösung. Ist dagegen $\omega \notin \pi\mathbb{Z}\backslash\{0\}$, so besitzt sie die eindeutige Lösung $\beta = \frac{\omega}{\sin \omega}$.
Betrachtet man die gleiche gDgl mit den Randbedingungen

$$u(0) = u(1) = 0,$$

so folgt mit den gleichen Argumenten, dass das Randwertproblem unendlich viele Lösungen hat, falls $\omega \in \pi\mathbb{Z}\backslash\{0\}$ ist, und nur die triviale Lösung $u = 0$ hat, falls $\omega \notin \pi\mathbb{Z}$ ist.

## XI.4. Stabilität

Sei $U \subset \mathbb{R}^n$, $n \in \mathbb{N}^*$, eine offene, nicht leere Menge und $f \in C(\mathbb{R} \times U, \mathbb{R}^n)$ auf $\mathbb{R} \times U$ Lipschitz-stetig bzgl. $U$. Für beliebiges $t_0 \in \mathbb{R}$ und $x_0 \in U$ betrachten wir in diesem Abschnitt das AWP

$$\begin{aligned} x' &= f(t, x(t)) \\ x(t_0) &= x_0. \end{aligned} \tag{XI.4.1}$$

Wir bezeichnen die gemäß Satz XI.1.9 (S. 435) existierende eindeutige globale Lösung von Gleichung (XI.4.1) mit $x(\cdot; t_0, x_0)$ und setzen voraus, dass sie für alle Zeiten existiert, d.h., dass gilt $t^+(t_0, x_0) = +\infty$ für alle $t_0 \in \mathbb{R}, x_0 \in U$. Ist speziell $t_0 = 0$, so lassen wir im Folgenden das Argument $t_0$ weg. Gemäß Satz XI.3.1 (S. 449) hängt $x(\cdot; t_0, x_0)$ stetig von $x_0$ ab, genauer: Für alle $x_1, x_2 \in U$, $t_0, t \in \mathbb{R}$ gilt

$$\|x(t; t_0, x_1) - x(t; t_0, x_2)\| \le e^{c|t-t_0|}\|x_1 - x_2\|$$

mit einem geeigneten $c \in \mathbb{R}_+^*$. Diese Ungleichung sagt aber nur über die kurzzeitige Abhängigkeit etwas Sinnvolles aus, da der Vorfaktor exponentiell mit $|t - t_0|$ wächst. Andererseits zeigt Beispiel XI.2.8 (S. 446), dass u. U. stets gilt

$$\lim_{t\to\infty} \|x(t; t_0, x_1) - x(t; t_0, x_2)\| = 0.$$

Dieses Langzeitverhalten wollen wir nun genauer untersuchen.

Definition XI.4.1. Die Lösung $x(\cdot; t_0, x_0)$ von Gleichung (XI.4.1) heißt Ljapunov-stabil, kurz stabil, wenn es ein $R > 0$ und $C > 0$ gibt, so dass für alle $x_1 \in \overline{B(x_0, R)}$ und alle $t \ge t_0$ gilt

$$\|x(t; t_0, x_1) - x(t; t_0, x_0)\| \le C\|x_0 - x_1\|.$$

Andernfalls heißt die Lösung $x(\cdot; t_0, x_0)$ instabil. Die Lösung $x(\cdot; t_0, x_0)$ heißt asymptotisch stabil, wenn es ein $R > 0$ und eine Funktion $\gamma \in C([t_0, \infty), \mathbb{R}_+)$ mit $\lim\limits_{t\to\infty} \gamma(t) = 0$ gibt, so dass für alle $x_1 \in \overline{B(x_0, R)}$ und alle $t \ge t_0$ gilt

$$\|x(t; t_0, x_1) - x(t; t_0, x_0)\| \le \gamma(t)\|x_0 - x_1\|.$$

Der folgende Satz ist eine Umformulierung und teilweise Verschärfung der Ergebnisse von Beispiel XI.2.8 (S. 446).

Satz XI.4.2. *Sei $A \in M_{n,n}(\mathbb{R})$. Für $x_0 \in \mathbb{R}^n$ bezeichne $x(\cdot; x_0)$ die Lösung des linearen autonomen AWP*

$$\begin{aligned} x' &= Ax \\ x(0) &= x_0. \end{aligned} \tag{XI.4.2}$$

*Dann gilt:*

(1) *Die Realteile der Eigenwerte von $A$ seien alle nicht positiv. Zusätzlich sei für alle Eigenwerte mit verschwindendem Realteil die algebraische gleich der geometrischen Vielfachheit. Dann ist jede Lösung von Gleichung* (XI.4.2) *stabil.*
(2) *Die Realteile der Eigenwerte $A$ seien alle negativ. Dann ist jede Lösung von Gleichung* (XI.4.2) *asymptotisch stabil.*
(3) *$A$ habe einen Eigenwert mit positivem Realteil. Dann ist jede Lösung von Gleichung* (XI.4.2) *instabil.*

Beweis. Seien $x_1, x_2 \in \mathbb{R}^n$ beliebig und

$$y = x(\cdot; x_2) - x(\cdot; x_1).$$

Dann gilt

$$\begin{aligned} y' &= Ay \\ y(0) &= x_2 - x_1. \end{aligned} \tag{XI.4.3}$$

Also reicht es in allen drei Fällen, die Stabilität bzw. die Instabilität der Nulllösung zu untersuchen.
ad (1): Wegen obiger Vorbemerkung ist die Behauptung eine Umformulierung des Ergebnisses von Beispiel XI.2.8 (S. 446).
ad (2): Wie wir uns bereits in Beispiel XI.2.8 (S. 446) überlegt haben, können wir o. E. komplexe Lösung von Gleichung (XI.4.3) betrachten. Außerdem seien $U$ und $J$ wie in Beispiel XI.2.8 (S. 446), d.h. $UAU^{-1} = J$ ist die Jordansche Normalform von $A$. Sei

$$\begin{aligned} \alpha &= \min\Big\{-\frac{1}{2}\operatorname{Re}\lambda_i : 1 \le i \le k\Big\} \\ &= -\frac{1}{2}\max\Big\{\operatorname{Re}\lambda_i : 1 \le i \le k\Big\} \\ &= -\frac{1}{2}\max\Big\{\operatorname{Re}\lambda : \lambda \text{ ist Eigenwert von } A\Big\}. \end{aligned}$$

Nach Voraussetzung ist $\alpha > 0$, und für jeden Eigenwert $\lambda$ von $A$ gilt

$$\operatorname{Re}\lambda \le -2\alpha.$$

Setze

$$D_\alpha = \operatorname{diag}(D_1, \dots, D_k)$$

mit

$$D_i = \begin{cases} 1 & \text{falls } n_i = 1, \\ \begin{pmatrix} 1 & & & \\ & \alpha & & \\ & & \ddots & \\ & & & \alpha^{n_i-1} \end{pmatrix} \in M_{n_i,n_i}(\mathbb{R}) & \text{falls } n_i > 1, \end{cases}$$

und

$$V = D_\alpha^{-1} U.$$

Dann ist

$$\begin{aligned} VAV^{-1} &= D_\alpha^{-1} J D_\alpha \\ &= J_\alpha \\ &= \operatorname{diag}(J_\alpha(\lambda_1), \ldots, J_\alpha(\lambda_k)) \end{aligned}$$

mit

$$J_\alpha(\lambda_i) = \begin{cases} \lambda_i & \text{falls } n_i = 1, \\ \begin{pmatrix} \lambda_i & \alpha & & \\ & \ddots & \ddots & \\ & & \ddots & \alpha \\ & & & \lambda_i \end{pmatrix} & \text{falls } n_i > 1. \end{cases}$$

Für jedes $i \in \mathbb{N}_k^*$ und jedes $w \in \mathbb{C}^{n_i}$ folgt dann mit der Cauchy-Schwarzschen Ungleichung

$$\begin{aligned} & w^H J_\alpha(\lambda_i) w + w^H J_\alpha(\lambda_i)^H w \\ &= \sum_{k=1}^{n_i} \overline{w}_k w_k (\lambda_i + \overline{\lambda}_i) + \alpha \sum_{k=1}^{n_i - 1} (\overline{w}_k w_{k+1} + \overline{w}_{k+1} w_k) \\ &= 2 \operatorname{Re} \lambda_i \|w\|^2 + \alpha \sum_{k=1}^{n_i - 1} (\overline{w}_k w_{k+1} + \overline{w}_{k+1} w_k) \\ &\leq 2 \operatorname{Re} \lambda_i \|w\|^2 + 2\alpha \|w\|^2 \\ &\leq -2\alpha \|w\|^2. \end{aligned}$$

Also gilt für jedes $z \in \mathbb{C}^n$

$$z^H J_\alpha z + z^H J_\alpha^H z \leq -2\alpha \|z\|^2. \tag{XI.4.4}$$

Durch

$$\|z\|_V = \|Vz\|$$

wird eine Norm auf $\mathbb{C}^n$ definiert. Da auf $\mathbb{C}^n$ alle Normen äquivalent sind, Satz III.1.11 (S. 64), gibt es ein $c \geq 1$ mit

$$\frac{1}{c} \|z\| \leq \|z\|_V \leq c \|z\| \quad \forall z \in \mathbb{C}^n.$$

Setze

$$\begin{aligned} \varphi(t) &= \|y(t)\|_V^2 = \|Vy(t)\|^2 \\ &= y(t)^H V^H V y(t). \end{aligned}$$

Dann gilt

$$\begin{aligned}\varphi' &= y^H V^H V y' + y'^H V^H V y \\ &= y^H V^H V y' + (V y')^H V y\end{aligned} \tag{XI.4.5}$$

und

$$Vy' = VAy = VAV^{-1}Vy = J_\alpha V y.$$

Zusammen mit Gleichung (XI.4.4) liefert dies

$$\begin{aligned}\varphi' &= y^H V^H J_\alpha V y + y^H V^H J_\alpha^H V y \\ &\le -2\alpha \|Vy\|^2 \\ &= -2\alpha\varphi.\end{aligned}$$

Division durch $\varphi$ und Integration von 0 bis $t$ liefert

$$\ln\varphi(t) - \ln\varphi(0) = \int_0^t \frac{\varphi'(s)}{\varphi(s)} ds \le -2\alpha t$$

und somit

$$\begin{aligned}\|y(t)\|^2 &\le c^2 \|y(t)\|_V^2 \\ &= c^2 \varphi(t) \\ &\le c^2 e^{-2\alpha t} \varphi(0) \\ &= c^2 e^{-2\alpha t} \|y(0)\|_V^2 \\ &\le c^4 e^{-2\alpha t} \|y(0)\|^2 \\ &= c^4 e^{-2\alpha t} \|x_1 - x_2\|^2.\end{aligned}$$

Also gilt

$$\|x(t; x_2) - x(t; x_1)\| \le c^2 e^{-\alpha t} \|x_2 - x_1\|.$$

Dies ist die behauptete asymptotische Stabilität mit $\gamma(t) = c^2 e^{-\alpha t}$.
AD (3): O.E. ist $\operatorname{Re}\lambda_1 > 0$. Sei $w$ ein Eigenvektor von $A$ zum Eigenwert $\lambda_1$. Dann folgt

$$x(t; w) = e^{\lambda_1 t} w$$

und somit

$$\|x(t; w)\| = e^{\operatorname{Re}\lambda_1 t} \|w\| \underset{t\to\infty}{\longrightarrow} \infty.$$

Wegen unserer Vorüberlegung beweist dies die behauptete Instabilität der Lösungen von Gleichung (XI.4.2). □

Die folgenden beiden Sätze sagen etwas aus über das Stabilitätsverhalten der Lösungen von Gleichung (XI.4.1), wenn Problem (XI.4.1) eine „kleine Störung“ einer linearen autonomen gDgl ist.

SATZ XI.4.3. *Seien $A \in M_{n,n}(\mathbb{R})$, $g \in C(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n)$ und $t_0 \in \mathbb{R}$. Es gelte:*

(1) *Alle Eigenwerte von $A$ haben negativen Realteil.*

(2) *Es gibt eine Funktion* $k \in C(\mathbb{R}, \mathbb{R}_+)$ *mit* $\lim_{t\to\infty} k(t) = 0$ *und*

$$\|g(t,u) - g(t,v)\| \le k(t)\|u - v\| \quad \forall t \ge t_0, u, v \in \mathbb{R}^n.$$

*Dann ist jede Lösung des AWP* (XI.4.1) *mit*

$$f(t,x) = Ax + g(t,x)$$

*asymptotisch stabil.*

Beweis. Durch eine Translation der Zeitvariablen können wir erreichen, dass o.E. $t_0 = 0$ ist. Entsprechend lassen wir im Folgenden stets das Argument $t_0$ fort. Weiter können wir $g$ durch

$$g(t,z) = g(t, \operatorname{Re} z) \in \mathbb{C}^n$$

zu einer Funktion aus $C(\mathbb{R} \times \mathbb{C}^n, \mathbb{C}^n)$ fortsetzen. Mit dem gleichen Argument wie in Beispiel XI.2.8 (S. 446) können wir daher o.E. komplexe Lösungen von Gleichung (XI.4.1) betrachten.
Seien nun $x_1, x_2 \in \mathbb{R}^n$ beliebig und

$$y = x(t;x_2) - x(t;x_1).$$

Dann gilt

$$\begin{aligned} y' &= Ay + g(t,x(t;x_2)) - g(t,x(t;x_1)) \\ y(0) &= x_2 - x_1. \end{aligned}$$

Wir benutzen die gleichen Notationen wie im Beweis von Satz XI.4.2. Dann gilt

$$\begin{aligned} &\|\{V[g(t,x(t;x_2)) - g(t,x(t;x_1))]\}^H\| \\ &= \|V[g(t,x(t;x_2)) - g(t,x(t;x_1))]\| \\ &\le \|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)}\|g(t,x(t;x_2)) - g(t,x(t;x_1))\| \\ &\le \|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)} k(t)\|x(t;x_2) - x(t;x_1)\| \\ &= \|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)} k(t)\|y\|. \end{aligned}$$

Zusammen mit Gleichungen (XI.4.4) und (XI.4.5) folgt hieraus

$$\begin{aligned} \varphi' &= y^H V^H V y' + (Vy')^H V y \\ &= y^H V^H J_\alpha V y + y^H V^H V[g(t,x(t;x_2)) - g(t,x(t;x_1))] \\ &\quad + y^H V^H J_\alpha^H V y + \{V[g(t,x(t;x_2)) - g(t,x(t;x_1))]\}^H V y \\ &\le -2\alpha\varphi + 2\|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)} k(t)\|y\|^2 \\ &\le [-2\alpha + 2\|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)} k(t)c^2]\varphi. \end{aligned} \tag{XI.4.6}$$

Wegen der Voraussetzung (2) gibt es ein $t_1 > 0$, derart dass für alle $t \ge t_1$ gilt

$$-2\alpha + 2\|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)} k(t)c^2 \le -\alpha.$$

Daher folgt aus Gleichung (XI.4.6) durch Division durch $\varphi$ und Integration von $t_1$ bis $t > t_1$

$$\varphi(t) \le \varphi(t_1) \cdot e^{-\alpha(t-t_1)} \quad \forall t \ge t_1.$$

Wegen Satz XI.3.1 (S. 449) gibt es andererseits ein $\beta \geq 0$ mit

$$\varphi(t) \leq e^{\beta t}\varphi(0) \quad \forall t \leq t_1.$$

Also gilt für $t \geq t_1$

$$\begin{aligned}\varphi(t) &\leq e^{-\alpha(t-t_1)}\varphi(t_1) \\ &\leq e^{-\alpha t}e^{\alpha t_1}e^{\beta t_1}\varphi(0)\end{aligned}$$

und für $t \leq t_1$

$$\begin{aligned}\varphi(t) &\leq e^{-\beta t}\varphi(0) \\ &= e^{-\alpha t}e^{(\beta+\alpha)t}\varphi(0) \\ &\leq e^{-\alpha t}e^{(\alpha+\beta)t_1}\varphi(0).\end{aligned}$$

Wie im Beweis von Satz XI.4.2 (2) folgt hieraus die behauptete asymptotische Stabilität mit

$$\gamma(t) = c^2 e^{\frac{1}{2}(\alpha+\beta)t_1}e^{-\frac{1}{2}\alpha t}.$$

□

Satz XI.4.4. *Seien $A \in M_{n,n}(\mathbb{R})$ und $g \in C(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n)$. Es gelte:*

(1) *Alle Eigenwerte von $A$ haben negativen Realteil.*
(2) *Es gibt eine monoton wachsende Funktion $k \in C(\mathbb{R}_+, \mathbb{R}_+)$ mit $k(0) = 0$ und*

$$\|g(t,x)\| \leq k(\|x\|)\|x\| \quad \forall t \in \mathbb{R}, x \in \mathbb{R}^n.$$

*Dann ist die Lösung des AWP (XI.4.1) mit*

$$f(t,x) = Ax + g(t,x)$$

*und $x_0 = 0$ asymptotisch stabil.*

Beweis. Wegen der Voraussetzung (2) gilt

$$g(t,0) = 0 \quad \forall t \in \mathbb{R}.$$

Daher ist

$$x(t;0) = 0 \quad \forall t \in \mathbb{R}.$$

Sei nun $x_1 \in \mathbb{R}^n$ zunächst beliebig und

$$y = x(t;x_1).$$

Mit den Bezeichnungen der vorigen Beweise gilt

$$\begin{aligned}\|[Vg(t,y(t))]^H\| &= \|Vg(t,y(t))\| \\ &\leq \|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)}k(\|y(t)\|)\|y(t)\|.\end{aligned}$$

Zusammen mit Gleichungen (XI.4.4) und (XI.4.5) folgt hieraus

$$\begin{aligned}\varphi' &= y^H V^H V y' + (Vy')^H V y \\ &= y^H V^H J_\alpha V y + y^H V^H V g(t, y(t)) \\ &\quad + y^H V^H J_\alpha^H V y + [V g(t, y(t))]^H V y \\ &\le -2\alpha\varphi + 2\|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)} k(\|y(t)\|)\|y\|^2 \\ &\le [-2\alpha + 2\|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)} c^2 k(\|y(t)\|)]\varphi \\ &\le [-2\alpha + 2\|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)} c^2 k(c\|y(t)\|_V)]\varphi \\ &= [-2\alpha + 2\|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)} c^2 \widetilde{k}(\varphi(t))]\varphi\end{aligned} \tag{XI.4.7}$$

mit

$$\widetilde{k}(u) = k(cu) \quad \forall u \in \mathbb{R}_+.$$

Wegen der Voraussetzung (2) ist $\widetilde{k}$ stetig, monoton wachsend und $\widetilde{k}(0) = 0$. Insbesondere gibt es ein $R > 0$ mit

$$2\|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)} c^2 \widetilde{k}(u) \le \alpha \quad \forall 0 \le u \le R.$$

Sei nun $x_1$ so gewählt, dass gilt

$$\varphi(0) = \|x_1\|_V^2 \le \frac{1}{2} R.$$

Wegen Satz XI.3.1 (S. 449) gibt es ein $T > 0$, so dass gilt

$$\|x(t; z)\|_V^2 \le R \quad \forall 0 \le t \le T, \|z\|_V^2 \le \frac{1}{2} R.$$

Für alle $0 \le t \le T$ ist dann

$$-2\alpha + 2\|V\|_{\mathcal{L}(\mathbb{C}^n,\mathbb{C}^n)} c^2 \widetilde{k}(\varphi(t)) \le -\alpha,$$

so dass aus Gleichung (XI.4.7) folgt

$$\varphi(t) \le e^{-\alpha t}\varphi(0) \quad \forall 0 \le t \le T. \tag{XI.4.8}$$

Insbesondere ist

$$\varphi(T) \le e^{-\alpha T}\varphi(0) \le \varphi(0) \le \frac{1}{2} R.$$

Daher können wir obiges Argument mit $y(T)$ an Stelle von $x_1$ wiederholen. Wegen

$$y(T + t) = x(T + t; x_1) = x(t; y(T)) \quad \forall t \ge 0$$

erhalten wir die Abschätzung

$$\begin{aligned}\varphi(T + t) &= \|y(T + t)\|_V^2 \\ &\le e^{-\alpha t}\varphi(T) \\ &\le e^{-\alpha(T+t)}\varphi(0) \quad \forall 0 \le t \le T,\end{aligned}$$

d.h., die Abschätzung (XI.4.8) gilt für alle $0 \le t \le 2T$. Durch Induktion folgt dann, dass Gleichung (XI.4.8) für alle $t \in \mathbb{R}_+$ gilt. Hieraus folgt wie in den vorigen Beweisen die behauptete asymptotische Stabilität mit $\gamma(t) = c^2 e^{-\frac{1}{2}\alpha t}$. □

Bemerkung XI.4.5. Mit etwas zusätzlichem Aufwand kann man folgende teilweise Umkehrung der Sätze XI.4.3 und XI.4.4 beweisen: Besitzt $A$ einen Eigenwert mit positivem Realteil und ist die Bedingung (2) aus Satz XI.4.3 bzw. XI.4.4 erfüllt, so ist für $f(t,x) = Ax + g(t,x)$ jede Lösung von Problem (XI.4.1) bzw. die Lösung von Problem (XI.4.1) zum Anfangswert $x_0 = 0$ instabil.

Zum Abschluss betrachten wir das autonome AWP

$$\begin{aligned} x' &= F(x) \\ x(0) &= x_0 \end{aligned} \tag{XI.4.9}$$

mit $F \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. Ist $x_0$ eine Nullstelle von $F$, d.h. $F(x_0) = 0$, so folgt sofort, dass die Lösung von Gleichung (XI.4.9) konstant ist:

$$x(t; x_0) = x_0 \quad \forall t \in \mathbb{R}.$$

Ist umgekehrt $x(t; x_0)$ eine stationäre Lösung von Gleichung (XI.4.9), d.h., gibt es ein $t_0 \in \mathbb{R}$ mit $x'(t_0; x_0) = 0$, so gilt

$$F(x(t_0; x_0)) = 0,$$

und aus dem Eindeutigkeitssatz folgt

$$x(t; x_0) = x(t_0; x_0) \quad \forall t \in \mathbb{R}.$$

Stationäre Lösungen von Gleichung (XI.4.9) sind also genau diejenigen Lösungen, die zu einem Anfangswert $x_0$ mit $F(x_0) = 0$ gehören. Daher heißen die Nullstellen von $F$ auch Ruhepunkte des AWP (XI.4.9). Der folgende Satz charakterisiert die Stabilität solcher Ruhepunkte.

Satz XI.4.6. *Sei $x_0$ ein Ruhepunkt von Gleichung (XI.4.9). Die Jacobi Matrix $DF(x_0)$ habe lauter Eigenwerte mit negativem Realteil. Dann ist die stationäre Lösung $x(t; x_0) = x_0$ von Gleichung (XI.4.9) asymptotisch stabil.*

Beweis. Sei $x_1 \in \mathbb{R}^n$ beliebig und

$$y(t) = x(t; x_1) - x(t; x_0) = x(t; x_1) - x_0.$$

Dann gilt

$$\begin{aligned} y' &= x'(t; x_1) \\ &= F(x(t; x_1)) \\ &= F(x(t; x_1)) - F(x_0) \\ &= DF(x_0)(x(t; x_1) - x_0) \\ &\quad + F(x(t; x_1)) - F(x_0) - DF(x_0)(x(t; x_1) - x_0) \\ &= DF(x_0)y + g(y) \end{aligned}$$

mit

$$g(u) = F(x_0 + u) - F(x_0) - DF(x_0)u.$$

Offensichtlich erfüllt $g$ die Voraussetzung (2) von Satz XI.4.4. Weiter erfüllt $A = DF(x_0)$ die Voraussetzung (1) von Satz XI.4.4. Damit folgt die Behauptung aus Satz XI.4.4. $\square$

Bemerkung XI.4.7. (1) Besitzt $DF(x_0)$ einen Eigenwert mit positivem Realteil, so folgt aus obigem Beweis und Bemerkung XI.4.5, dass die stationäre Lösung $x(t; x_0) = x_0$ von Gleichung (XI.4.9) instabil ist.
(2) Besitzt $DF(x_0)$ einen rein imaginären Eigenwert, so ist keine allgemein gültige Aussage möglich.
Betrachte z.B. Gleichung (XI.4.9) mit

$$F(x) = (-x_2 + x_1^3, x_1 + x_2^3).$$

Offensichtlich ist $x_0 = 0$ die einzige Nullstelle, und

$$DF(x_0) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

hat die Eigenwerte $\pm i$. Sei $x(t)$ eine Lösung von Gleichung (XI.4.9) zu einem Anfangswert $\widetilde{x}_0 \neq 0$. Für

$$r(t)^2 = \|x(t)\|^2 = x_1(t)^2 + x_2(t)^2$$

folgt dann

$$\begin{aligned} rr' &= x_1 x_1' + x_2 x_2' \\ &= x_1(-x_2 + x_1^3) + x_2(x_1 + x_2^3) \\ &= x_1^4 + x_2^4 \\ &> 0. \end{aligned}$$

Also ist $r'(t) > 0$ und die Lösung „läuft von $x_0 = 0$ weg“, d.h., $x_0$ ist instabil.
Betrachte nun Gleichung (XI.4.9) mit

$$F(x) = (-x_2 - x_1^3, x_1 - x_2^3).$$

Wieder ist $x_0 = 0$ die einzige Nullstelle, und

$$DF(x_0) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

hat wieder die Eigenwerte $\pm i$. Für $x(t)$ und $r(t)$ wie oben folgt jetzt

$$\begin{aligned} rr' &= x_1 x_1' + x_2 x_2' \\ &= x_1(-x_2 - x_1^3) + x_2(x_1 - x_2^3) \\ &= -x_1^4 - x_2^4 \\ &< 0. \end{aligned}$$

Also ist $r'(t) < 0$ und die Lösung „läuft in $x_0$ hinein“, d.h., $x_0$ ist stabil.

Beispiel XI.4.8. (1) Betrachte das gedämpfte mathematische Pendel

$$x'' + 2\alpha x' + \lambda \sin x = 0$$

mit $\alpha, \lambda \in \mathbb{R}_+^*$. Die zugehörige gDgl 1. Ordnung ist

$$\begin{aligned} x' &= v \\ v' &= -2\alpha v - \lambda \sin x. \end{aligned}$$

Dies ist von der Form von Gleichung (XI.4.9) mit

$$F(x, v) = (v, -\lambda \sin x - 2\alpha v).$$

Die Nullstellen sind offensichtlich genau die Punkte

$$(x_k, v_k) = (k\pi, 0), \quad k \in \mathbb{N}.$$

Die Jacobi Matrix ist

$$DF(x_k, v_k) = \begin{pmatrix} 0 & 1 \\ -\lambda \cos x_k & -2\alpha \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -(-1)^k \lambda & -2\alpha \end{pmatrix}$$

und hat die Eigenwerte $-\alpha \pm \sqrt{\alpha^2 - \lambda(-1)^k}$. Mithin sind die Ruhepunkte zu geradem $k$ asymptotisch stabil und die zu ungeradem $k$ instabil.

(2) Betrachte Problem (XI.4.9) mit $n = 2$ und

$$F(x) = \begin{pmatrix} -x_1^2 - x_2 \\ -x_1 + x_2^2 \end{pmatrix}.$$

Aus $F(x) = 0$ folgt

$$x_1 = x_2^2$$

und

$$0 = -x_2^4 - x_2 = -x_2(x_2^3 + 1).$$

Also sind die Nullstellen

$$(0, 0) \text{ und } (1, -1).$$

Für die Jacobi Matrix ergibt sich

$$DF(x) = \begin{pmatrix} -2x_1 & -1 \\ -1 & 2x_2 \end{pmatrix}$$

und somit

$$\begin{aligned} DF(0,0) &= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \quad \text{mit den Eigenwerten } \pm 1, \\ DF(1,-1) &= \begin{pmatrix} -2 & -1 \\ -1 & -2 \end{pmatrix} \quad \text{mit den Eigenwerten } -3, -1. \end{aligned}$$

Also ist $(0, 0)$ instabil und $(1, -1)$ asymptotisch stabil.

(3) Betrachte Problem (XI.4.9) mit $n = 2$ und

$$F(x) = \begin{pmatrix} x_1^2 - x_2^2 \\ 2x_1 x_2 \end{pmatrix}.$$

Offensichtlich ist 0 die einzige Nullstelle und $DF(0) = 0$. Also erlauben Satz XI.4.6 und Bemerkung XI.4.7 keine Aussage über die Stabilität oder Instabilität der Nulllösung.
Setze $z(t) = x_1(t) + ix_2(t)$. Dann gilt

$$z' = x_1' + ix_2' = x_1^2 - x_2^2 + 2x_1x_2i = z^2.$$

Also lautet die Lösung von Gleichung XI.4.9 zum Anfangswert $(x_0, y_0)$

$$z(t) = \frac{1}{z_0 - t}$$

mit

$$z_0 = (x_0 + iy_0)^{-1}.$$

Hieraus folgt sofort:

(1) Ist $z_0 \in \mathbb{C}\backslash\mathbb{R}_+$, so konvergiert $z(t)$ für $t \to \infty$ gegen Null.
(2) Ist $z_0 \in \mathbb{R}_+$, so explodiert die Lösung in endlicher Zeit. Insbesondere ist die Nulllösung instabil.

# Zusammenfassung

## I. Aufbau des Zahlsystems

1. Die natürlichen Zahlen
   Peano Axiome; Induktionsprinzip; Prinzip der rekursiven Definition; Fakultät und Binomialkoeffizienten; Binomischer Lehrsatz; geometrische Reihe; endliche, abzählbare, überabzählbare Mengen; Potenzmenge $\mathcal{P}(\mathbb{N})$ nicht abzählbar
2. Die ganzen Zahlen
   Definition und Eigenschaften; Abzählbarkeit von $\mathbb{Z}$
3. Die rationalen Zahlen
   Definition und Eigenschaften; Abzählbarkeit von $\mathbb{Q}$
4. Die reellen Zahlen
   $\sqrt{2}$ nicht rational; Infimum und Supremum; Ordnungsvollständigkeit; $\mathbb{Q}$ ist nicht ordnungsvollständig; Dedekindscher Hauptsatz; Satz von Archimedes; $\mathbb{Q}$ und $\mathbb{R}\backslash\mathbb{Q}$ dicht in $\mathbb{R}$; Existenz $n$-ter Wurzeln; Überabzählbarkeit von $\mathbb{R}$; Betrag; Intervalle
5. Die komplexen Zahlen
   Definition und Rechenregeln; geometrische Interpretation; Real- und Imaginärteil, Betrag und konjugiert komplexe Zahl

## II. Folgen und Reihen

1. Konvergenz von Folgen
   Folgen; Häufungspunkt; Grenzwert; konvergente, divergente und beschränkte Folgen; konvergent $\Rightarrow$ beschränkt; Grenzwert ist eindeutig; Teilfolgen; Charakterisierung von Grenzwert und Häufungspunkten durch Konvergenz von Teilfolgen; Nullfolgen; Rechenregeln für Limites
2. Vollständigkeit
   monotone Folgen; monoton und beschränkt $\Rightarrow$ konvergent; Satz von Bolzano-Weierstraß (beschränkt $\Rightarrow$ ex. Häufungspunkt); Cauchyfolge; konvergent $\Rightarrow$ Cauchyfolge $\Rightarrow$ beschränkt; Vollständigkeit; $\mathbb{K}$ vollständig; $\mathbb{Q}$ nicht vollständig
3. Uneigentliche Konvergenz
   Definition $\overline{\mathbb{R}}$; Konvergenz in $\overline{\mathbb{R}}$; Rechenregeln für Limites in $\overline{\mathbb{R}}$; Definition $\limsup$, $\liminf$; $\limsup =$ größter Häufungspunkt; $\liminf =$ kleinster Häufungspunkt; konvergent genau dann wenn $\limsup = \liminf$
4. Reihen
   Partialsummen; Reihen; Konvergenz und Grenzwerte von Reihen; harmonische und geometrische Reihe; $\sum x_n$ konvergent $\Rightarrow x_n$ Nullfolge; Cauchykriterium; Leibnizkriterium; alternierende harmonische Reihe

5. Absolute Konvergenz
   absolute und bedingte Konvergenz; Majoranten-, Wurzel- und Quotientenkriterium; Umordnung absolut konvergenter Reihen; Cauchyprodukt absolut konvergenter Reihen
6. Potenzreihen
   Definition; Konvergenzradius; Wurzel- und Quotientenkriterium für Konvergenzradius; Exponential- und geometrische Reihe; Rechenregeln für Potenzreihen

## III. Stetige Funktionen

1. Normierte Vektorräume
   Vektorräume und Normen; Beispiele: $\mathbb{R}, \mathbb{C}, \mathbb{K}^n, \ell_1, \ell_2, \ell_\infty$; Grenzwerte, Häufungspunkte und Cauchyfolgen in normierten Vektorräumen; äquivalente Normen; Äquivalenz der Normen auf $\mathbb{K}^n$; Banachräume
2. Topologische Grundbegriffe
   innere Punkte; offene Mengen; Rechnen mit offenen Mengen; abgeschlossene Mengen; Rechnen mit abgeschlossenen Mengen; Häufungs- und Berührungspunkte von Mengen; Zusammenhang mit Folgen; Abschluss, Inneres und Rand einer Menge; Hausdorffsches Trennungsaxiom; relativ offene und abgeschlossene Mengen
3. Stetigkeit
   Umgebungs-, $\epsilon$–$\delta$- und Folgendefinition der Stetigkeit; Rechenregeln für stetige Funktionen; Stetigkeit von Polynomen und rationalen Funktionen; $C(X, Y)$; Stetigkeit von Funktionen mit Werten in $\mathbb{K}^n$, links- und rechtsseitige Stetigkeit; gleichmäßige Stetigkeit; gleichmäßig stetig $\Rightarrow$ stetig
4. Kompaktheit
   Definition; kompakt $\Rightarrow$ beschränkt und abgeschlossen; in normierten Vektorräumen: kompakt = folgenkompakt; Satz von Heine-Borel (kompakt = beschränkt und abgeschlossen im $\mathbb{K}^n$); Heine-Borel gilt nicht in allgemeinen normierten Vektorräumen; stetige Funktionen auf kompakten Mengen sind gleichmäßig stetig; stetige Bilder kompakter Mengen sind wieder kompakt; stetige, reellwertige Funktionen auf kompakten Mengen nehmen Maximum und Minimum an; Fundamentalsatz der Algebra
5. Zusammenhang
   Definition; $M \subset \mathbb{R}$ zusammenhängend $\Leftrightarrow$ $M$ Intervall; stetige Bilder zusammenhängender Mengen sind wieder zusammenhängend; Zwischenwertsatz; reelle Polynome ungeraden Grades besitzen mindestens eine reelle Nullstelle; Wege; Wegzusammenhang; konvexe Mengen; $M$ offen: $M$ wegzusammenhängend $\Leftrightarrow$ $M$ zusammenhängend; Gebiete
6. Funktionen in $\mathbb{R}$
   monotone Funktionen; Sprungstellen; monotone Funktionen sind stetig bis auf abzählbar viele Sprungstellen; streng monotone stetige Funktionen sind bijektiv mit stetiger Umkehrabbildung
7. Exponentialfunktion und Verwandte
   Definition von exp, sin, cos; Eigenschaften und Rechenregeln; Additionstheoreme für sin und cos; Wachstumsverhalten von exp; Definition von ln; Rechenregeln; Wachstumsverhalten von ln; Verhalten von exp für imaginäres Argument; Polarkoordinatendarstellung komplexer Zahlen und Anwendung auf Multiplikation; Eulersche Darstellung von exp

## IV. Differentialrechnung einer Veränderlichen

1. Differenzierbarkeit
   Definitionen für Differenzierbarkeit in einem Punkt; differenzierbar $\Rightarrow$ stetig; Differenzierbarkeit von Funktionen mit Werten in $\mathbb{K}^n$; Rechenregeln; Kettenregel; Differenzierbarkeit der Umkehrfunktion; perfekte Mengen; Räume $C^n(M,\mathbb{K})$, $C^\infty(M,\mathbb{K})$; Leibnizregel; Ableitung von exp, sin, cos, ln, $a^x$, $x^2\sin(1/x)$; links- und rechtsseitige Differenzierbarkeit; Eigenschaften von $\exp(1/x)$
2. Mittelwertsätze
   Satz von Rolle; Mittelwertsatz; Charakterisierung monotoner Funktionen; Umkehrfunktionen differenzierbarer, monotoner Funktionen sind wieder differenzierbar; arccos, arcsin; Charakterisierung und Eigenschaften konvexer und konkaver Funktionen; Youngsche und Höldersche Ungleichung; Regel von de l'Hôpital
3. Taylorformeln
   Taylorpolynom und Taylorreihe; Abschätzung des Restgliedes; Darstellbarkeit einer Funktion durch ihre Taylorreihe; Taylorreihe von ln
4. Numerische Lösung von Gleichungen
   Fixpunkt; Kontraktion; Banachscher Fixpunktsatz; Newtonverfahren; geometrische Interpretation; lokale, quadratische Konvergenz des Newtonvefahrens; globale Konvergenz des Newtonverfahrens bei strikt konvexen oder konkaven Funktionen; divisionsfreie Berechnung von $1/a$; Heronverfahren für $a^{1/n}$

## V. Funktionenfolgen

1. Gleichmäßige Konvergenz
   punktweise und gleichmäßige Konvergenz; gleichmäßig konvergent $\Rightarrow$ punktweise konvergent; Raum $B(M,Y)$; Cauchykriterium für gleichmäßige Konvergenz; punktweise, absolute, gleichmäßige, normale Konvergenz von Funktionenreihen; Beziehungen zwischen den Konvergenzbegriffen; Anwendung auf Potenzreihen
2. Vertauschen von Grenzprozessen
   Stetigkeit der Grenzfunktion; lokal gleichmäßige Konvergenz; Differenzierbarkeit der Grenzfunktion
3. Analytische Funktionen
   Differentiation von Potenzreihen; $\mathbb{K}$-analytische Funktionen; konvergente Potenzreihen stellen analytische Funktionen dar; Stammfunktionen analytischer Funktionen; Binomialreihe; Reihendarstellung von arcsin; Identitätssatz für analytische Funktionen; Maximum- und Minimumprinzip; komplex-analytische Funktionen sind gebietstreu
4. Sprungstetige Funktionen
   Treppenfunktionen $T(I,Y)$; sprungstetige Funktionen $S(I,Y)$; monotone Funktionen sind sprungstetig; Beziehungen zwischen Räumen $T(I,Y)$, $S(I,Y)$, $B(I,Y)$ und $C(I,Y)$; $I$ kompakt $\Rightarrow T(I,Y)$ dicht in $S(I,Y)$ bzgl. $\|\cdot\|_\infty$

## VI. Integralrechnung einer Variablen

1. Das (Riemann-) Integral
   Integral einer Treppenfunktion; Linearität und Beschränktheit des Integrals; die Räume $\mathcal{L}(X,Y)$ und ihre Eigenschaften; Fortsetzung stetiger, linearer Abbildungen auf den Abschluss eines Unterraumes;

Integral sprungstetiger Funktionen; Approximation durch Riemannsche Summen; Zusammenhang mit dem Riemann-Integral

2. Eigenschaften des Integrals

gliedweise Integration von Funktionenfolgen; punktweise Konvergenz nicht ausreichend; Stetigkeit und Linearität des Integrals; orientiertes Integral; Additivität und Monotonie des Integrals; Integrale komplex- und vektorwertiger Funktionen; stetige und differenzierbare Abhängigkeit des Integrals von den Integrationsgrenzen; Stammfunktion und unbestimmtes Integral; Hauptsatz der Differential- und Integralrechnung; Mittelwertsatz der Integralrechnung

3. Integrationstechniken

Substitutionsregel; Darstellung mittels Differentialen; partielle Integration; Partialbruchzerlegung; Trapezregel und Fehlerabschätzung; Stirlingsche Formel

4. Uneigentliche Integrale

zulässige Funktionen; uneigentliches Integral zulässiger Funktionen; Zusammenhang mit dem gewöhnlichen Integral; Integralkriterium für Reihen; absolut konvergente Integrale; Majorantenkriterium; Eulersches Betaintegral; uneigentliche Integrale und gleichmäßig konvergente Funktionenfolgen

5. Die Eulersche Gammafunktion

Definition; Eigenschaften; logarithmisch konvexe Funktionen und ihre Eigenschaften; Charakterisierung der Gammafunktion; Gaußsche Darstellung der Gammafunktion; Zusammenhang mit dem Eulerschen Betaintegral

## VII. Differentialrechnung mehrerer Variablen

1. Stetige lineare Abbildungen

$Y$ vollständig $\Rightarrow \mathcal{L}(X, Y)$ vollständig; Isomorphismus, Isometrie; $\dim X < \infty \Rightarrow X \cong \mathbb{K}^{\dim X}$; $\dim X < \infty \Rightarrow X$ ist vollständig und alle Normen auf $X$ sind äquivalent; $\dim X < \infty$ und $Y$ beliebig $\Rightarrow$ jede lineare Abbildung $X \to Y$ ist stetig; Aussagen gelten nicht für $\dim X = \infty$; Zusammenhang zwischen linearen Abbildungen und Matrizen

2. Differenzierbarkeit

Differenzierbarkeit; Zusammenhang mit Funktionen einer Variablen; differenzierbar $\Rightarrow$ stetig; Richtungsableitung; differenzierbar $\Rightarrow$ alle Richtungsableitungen existieren und hängen linear von der Richtung ab, Umkehrung gilt nicht; partielle Ableitungen; differenzierbar $\Rightarrow$ alle partiellen Ableitungen existieren, Umkehrung gilt nicht; Jacobi-/Funktionalmatrix, Gradient; Differenzierbarkeitskriterium; Zusammenhang mit komplexer Differenzierbarkeit und Cauchy-Riemannsche Differentialgleichungen

3. Eigenschaften differenzierbarer Funktionen

Linearität der Ableitung; Kettenregel und ihre Matrixdarstellung; Produktregel; Mittelwertsatz; gliedweise Differentiation von Funktionenfolgen; notwendige Bedingung für lokale Extrema, kritische Punkte

4. Höhere Ableitungen

$m$-te Ableitung; stetige, multilineare Abbildungen; $m$-te Ableitung als stetige, $m$-lineare Abbildung; Symmetrie höherer Ableitungen;

Kettenregel; Taylorsche Formel; partielle Ableitungen höherer Ordnung; Kriterium für Existenz höherer Ableitungen; Taylorsche Formel für Funktionen auf $\mathbb{R}^k$; Hessesche Matrix; positiv definite Matrizen und ihre Eigenwerte; hinreichende Bedingungen für lokale Extrema

5. Umkehrabbildungen

Automorphismengruppe von $X$; $\mathrm{Isom}(X, Y)$ ist offen in $\mathcal{L}(X, Y)$ und Abbildung $A \to A^{-1}$ ist $C^\infty$; Satz über die Umkehrabbildung; Diffeomorphismen; Satz über implizite Funktionen; reguläre Punkte; Anwendung auf Funktionen $\mathbb{R}^k \to \mathbb{R}^\ell$; Berechnung lokaler Extrema unter Nebenbedingungen mittels Lagrangescher Multiplikatoren

## VIII. Kurven und Kurvenintegrale

1. Kurven und ihre Länge

$C^m$-Wege, rektifizierbare Wege; stetige Wege sind nicht notwendig rektifizierbar; $C^1$-Wege sind rektifizierbar; Umparametrisierungen; Invarianz der Länge eines Weges unter Umparametrisierungen; $C^m$-Kurven; Beispiele

2. Tangente und Krümmung

reguläre Wege; Tangenteneinheitsvektor; Parametrisierung nach der Bogenlänge; Orthogonalität der ersten und zweiten Ableitung nach der Bogenlänge; Krümmung eines ebenen Weges; Frenetsche Formeln

3. Kurvenintegrale

Wegintegral; Invarianz des Wegintegrals unter Umparametrisierungen; Kurvenintegral; stückweise $C^m$-Kurven; Kurvenintegral für stückweise $C^m$-Kurven; Eigenschaften; Gradientenfeld und Stammfunktion; $f$ Gradientenfeld $\Leftrightarrow$ das Kurvenintegral verschwindet für jede geschlossene, stückweise $C^1$-Kurve; Lemma von Goursat; $f \in C^1(G, \mathbb{R}^n)$, $G$ konvex: $f$ Gradientenfeld $\Leftrightarrow Df$ symmetrisch; Homotopie von Wegen; einfach zusammenhängend; Invarianz des Wegintegrals unter Homotopie; Satz von Poincaré

4. Komplexe Kurvenintegrale

komplexe Kurvenintegrale; holomorphe Funktionen; Cauchyscher Integralsatz; Cauchysche Integralformel; $f$ holomorph $\Leftrightarrow f$ komplexanalytisch; Cauchysche Ableitungsformel; Satz von Liouville; Laurent-Reihen; Residuum und Hauptteil; Riemannscher Hebbarkeitssatz; Residuensatz

## IX. Integralrechnung mehrerer Veränderlicher

1. Nullmengen

Lebesgue-Maß mehrdimensionaler Intervalle; Nullmengen; Beispiele; Teilmengen und abzählbare Vereinigungen von Nullmengen sind Nullmengen; Cantorsches Diskontinuum; punktweise Konvergenz fast überall

2. Das Lebesgue-Integral

Treppenfunktionen und ihre Darstellung; Lebesgue-Integral von Treppenfunktionen; Eigenschaften des Integrals; monotone Folgen von Treppenfunktionen und deren Integrale; Definition des Lebesgue-Integrals; Eigenschaften des Lebesgue-Integrals; Integrierbarkeitskriterien; Zusammenhang mit eigentlichem und uneigentlichem Riemann-Integral

3. Eigenschaften des Lebesgue-Integrals

   Satz von Fubini; Anwendungen; Satz von der monotonen Konvergenz; Anwendungen; Satz von Lebesgue über die majorisierte Konvergenz; Satz von Lebesgue über die beschränkte Konvergenz; Lemma von Fatou; Vertauschen von Differentiation und Integration

4. Messbare Funktionen und Mengen

   messbare Funktionen; Charakterisierungen; Eigenschaften; Zusammenhang mit Integrierbarkeit; Satz von Tonelli; messbare Mengen und ihr Lebesgue-Maß; Eigenschaften; offene und abgeschlossene Mengen sind messbar; Beispiel einer nicht messbaren Menge; Zusammenhang zwischen Messbarkeit von Mengen und Funktionen; Verknüpfung messbarer Funktionen

5. Der Transformationssatz

   Messbarkeit transformierter Mengen; Transformationssatz für lineare Transformationen; Approximation durch lineare Transformationen; Transformationssatz; Anwendung auf Koordinatentransformationen

6. Die $L^p$-Räume

   Definition von $L^p$, $1 \leq p < \infty$; Höldersche Ungleichung; Banachraum-Struktur; Abschneidefunktionen; Schrumpfungslemma; Partitionen der Eins; $C_0^\infty$ ist dicht in $L^p$

## X. Analysis auf Mannigfaltigkeiten

1. Mannigfaltigkeiten

   Charakterisierung von Mannigfaltigkeiten als lokale Nullstellenmengen, lokale Graphen, lokale Urbilder von Ebenen, lokale Bilder offener Mengen in $\mathbb{R}^k$; Karten und Atlanten; Kartenwechsel; Beispiele

2. Tangentialraum und Orientierung

   Definition von Tangential- und Normalraum; orientierungserhaltende Diffeomorphismen; orientierte Atlanten; Orientierbarkeit einer Mannigfaltigkeit; Einheitsnormalenfeld; Orientierbarkeit von Hyperflächen; Beispiele

3. Integration auf Mannigfaltigkeiten

   Maßtensor und Gramsche Determinante; Definition des Integrals auf Mannigfaltigkeiten; Eigenschaften des Integrals; Beispiele; Verhalten unter Transformationen; Zerlegung eines $n$-dimensionalen Integrals in Radial- und Winkelanteil; Integral rotationssymmetrischer Funktionen; Kompakta mit stückweise glattem Rand; Eigenschaften des glatten Randteils; Integralsatz von Gauß; Interpretation der Divergenz; Greensche Formeln

4. Multilineare Algebra

   Definition und Charakterisierung alternierender $r$-Formen; äußeres Produkt; Eigenschaften; pull-back; Eigenschaften

5. Differentialformen

   Definition und Eigenschaften von $r$-Formen; totales Differential; äußeres Produkt; pull-back; äußere Ableitung; Eigenschaften; geschlossene und exakte Formen; Satz von Poincaré für Vektorfelder in $\mathbb{R}^3$

6. Integration von Differentialformen

   Integration von $n$-Formen in $\mathbb{R}^n$; Transformationssatz; Integration von $k$-Formen auf $k$-dimensionalen Mannigfaltigkeiten; Beispiele; Zusammenhang mit Integration von Funktionen auf Mannigfaltigkeiten; Integralsatz von Gauß; Leibnizsche Sektorformel; Green-Riemannsche Formel; Integration über Halbräume; Kompakta mit glattem Rand relativ zu einer Mannigfaltigkeit; Eigenschaften; Integralsatz von Stokes; klassische Form im Spezialfall $\mathbb{R}^3$; Beispiele

7. Die Fixpunktsätze von Brouwer und Schauder

Brouwerscher Fixpunktsatz; Nullstellen von Funktionen in $\mathbb{R}^n$; $B_n$ und $S^n$ nicht homöomorph; beste Approximation durch konvexe Teilmengen eines Hilbertraumes; Brouwerscher Fixpunktsatz in endlich dimensionalen Vektorräumen; konvexe Hülle; 1. Schauderscher Fixpunktsatz; relativ kompakt; 2. Schauderscher Fixpunktsatz; Satz von Arzela-Ascoli; Existenzsatz von Peano für Lösungen gewöhnlicher Differentialgleichungen

## XI. Gewöhnliche Differentialgleichungen

1. Existenz- und Eindeutigkeitssätze

gewöhnliche Differentialgleichungen; Anfangswertprobleme; autonome Differentialgleichungen; Beispiele; Lipschitz-Stetigkeit; Lemma von Gronwall; Satz von Picard-Lindelöf; globale Existenz und Eindeutigkeit; maximale Existenzintervalle

2. Elementare Lösungsmethoden

Exakte Differentialgleichungen; integrierende Faktoren; Räuber-Beute Modell; Trennung der Variablen; Chemische Reaktionskinetik; Populationsdynamik; Luftwiderstand; Variation der Konstanten; autonome lineare Differentialgleichungen und asymptotisches Verhalten

3. Stetigkeits- und Differenzierbarkeitssätze

Stetige Abhängigkeit von den Anfangswerten; differenzierbare Abhängigkeit von den Anfangswerten; Randwertprobleme

4. Stabilität

(Lyapunov-) stabil, instabil, asymptotisch stabil; Stabilität der Lösungen linearer Differentialgleichungen; Störungen linearer Differentialgleichungen; Stabilität von Ruhepunkten autonomer Systeme; Beispiele

# Index