# 1 Topologie in metrischen Räumen

## 1.1 Metrische und normierte Räume

Das Ziel dieses Abschnitts ist es, die Erkenntnisse über Folgen (und Reihen) reeller Zahlen aus dem vergangenen Semester zu verallgemeinern. Insbesondere wollen wir Aussagen über die Konvergenz von Folgen im $\mathbb{R}^n$ formulieren und beweisen können. Wir wählen jedoch den allgemeineren Ansatz des metrischen Raumes, einer abstrakten mathematischen Struktur auf der ein Abstandsbegriff definiert ist. Die Theorie wird dadurch nicht komplizierter (nur abstrakter) aber wir können später mit den Werkzeugen die wir uns erarbeiten werden auch zum Beispiel die Konvergenz von Funktionenfolgen untersuchen.

**Definition 1.1 (Metrischer Raum).** Ein metrischer Raum $(X, d)$ ist ein Paar bestehend aus einer Menge $X$ zusammen mit einer Abbildung $d : X \times X \to \mathbb{R}$, die als Metrik bezeichnet wird und die folgenden Bedingungen für alle $x, y, z \in X$ erfüllt:

$$\begin{aligned} d(x,y) &\geq 0 \text{ (Positivität)} \\ d(x,y) = 0 &\Longleftrightarrow x = y \text{ (Definitheit)} \\ d(x,y) &= d(y,x) \text{ (Symmetrie)} \\ d(x,y) &\leq d(x,z) + d(z,y) \text{ (Dreiecksungleichung).} \end{aligned}$$

Die Elemente eines metrischen Raumes werden oft als Punkte bezeichnet. Die Bezeichnungen „Raum“ und „Punkt“ rühren daher, dass es sich hier um Objekte der Geometrie (in einem sehr weit gefassten Sinn des Wortes) handelt. Ihnen kommt keine tiefere Bedeutung zu. Ein einzelner Punkt eines abstrakten metrischen Raumes besitzt keine besonderen Eigenschaften außer der Element eben dieses Raumes zu sein. Interessante Eigenschaften treten erst durch die Struktur (die Metrik mit ihren Eigenschaften) auf, die verschiedene Punkte des Raumes zueinander in Beziehung setzt.

In vielen für uns wichtigen Beispielen besitzt die Menge $X$ außer der metrischen noch eine damit kompatible Vektorraumstruktur:

**Definition 1.2 (Normierter Raum).** Ein normierter Raum $(V, \|\cdot\|)$ ist ein Paar bestehend aus einem Vektorraum $V$ über dem Körper $\mathbb{K}$ der reellen oder komplexen Zahlen und einer Abbildung $\|\cdot\| : V \to \mathbb{R}$, die als Norm bezeichnet wird und die folgenden Bedingungen für alle $x, y \in V$ erfüllt:

$$\begin{aligned} \|x\| &\geq 0 \text{ (Positivität)} \\ \|x\| = 0 &\Longleftrightarrow x = 0 \text{ (Definitheit)} \\ \|\lambda x\| &= |\lambda| \, \|x\| \text{ (absolute Homogenität)} \\ \|x + y\| &\leq \|x\| + \|y\| \text{ (Dreiecksungleichung).} \end{aligned}$$

Jeder normierte Raum erhält die Struktur eines metrischen Raumes indem wir definieren $d(x, y) := \|x - y\|$. Man überprüft leicht, dass die so definierte Abbildung eine Metrik ist, sie wird als die von der Norm induzierte Metrik bezeichnet. Das für uns bei Weitem wichtigste Beispiel:

*Beispiel 1.3 ($\mathbb{R}^n$ mit euklidischer Norm).* Definiere die euklidische Norm

$$\begin{aligned} \|\cdot\| &: \mathbb{R}^n \to \mathbb{R} \\ \|(x_1, \ldots, x_n)\| &:= \sqrt{x_1^2 + \cdots + x_n^2}. \end{aligned}$$

Dann ist $(\mathbb{R}^n, \|\cdot\|)$ ein normierter Raum.

Im Fall des komplexen Vektorraums $\mathbb{C}^n$ müssen wir die Definition der Norm leicht verändern:

$$\|(x_1, \ldots, x_n)\| := \sqrt{\overline{x_1} x_1 + \cdots + \overline{x_n} x_n}.$$

*Beispiel 1.4.* (i) Die Abbildung

$$(x_1, \ldots, x_n) \mapsto \max\{|x_1|, \ldots, |x_n|\}$$

definiert ebenfalls eine Norm auf dem $\mathbb{R}^n$. Bereits auf dem $\mathbb{R}^n$ sind also verschiedene Normen möglich.

(ii) Sei $M$ eine beliebige Menge. Dann wird durch

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

eine Metrik definiert, die als diskrete Metrik bezeichnet wird.

(iii) Seien $(X, d_X)$ und $(Y, d_Y)$ metrische Räume. Dann ist $(X \times Y, d_{X \times Y})$ mit der Metrik

$$d_{X\times Y}((x_1, y_1), (x_2, y_2)) := d_X(x_1, x_2) + d_Y(y_1, y_2) \tag{1.1}$$

ein metrischer Raum.

Analog ist $\|(v, w)\|_{V \times W} := \|v\|_V + \|w\|_W$ eine Norm auf dem direkten Produkt $V \times W$.

*Bemerkung 1.5.*

- Sei $(X, d)$ ein metrischer Raum und $Y \subset X$ dann ist $(Y, d|_{Y \times Y})$ ebenfalls ein metrischer Raum.
- Analog sind Untervektorräume (nicht beliebige Teilmengen) normierter Vektorräume ebenfalls normierte Vektorräume.
- Zu einer vorgegebenen Menge gibt es keine „natürliche" Metrik. Man kann vielmehr auf ein und der selben Menge unterschiedliche Metriken (im Fall von Vektorräumen auch unterschiedliche Normen) mit ganz unterschiedlichen Eigenschaften betrachten. Im konkreten Fall wählt man sich dann die Metrik, die dem Problem am besten angemessen ist.
- Die „Minkowski-Metrik" die in der Relativitätstheorie eine zentrale Rolle spielt ist keine Metrik im hier eingeführten Sinne.

**Definition 1.6 (Kugeln und Beschränktheit).** Sei $(X, d)$ ein metrischer Raum, $x \in X$ und $r > 0$. Die Menge

$$B_r(x) := \{y \in X |\ d(y, x) < r\}$$

heißt offene Kugel mit Radius $r$ um $x$ (manchmal auch $r$-Kugel um $x$) und die Menge

$$K_r(x) := \{y \in X |\ d(y, x) \leq r\}$$

heißt abgeschlossene Kugel mit Radius $r$ um $x$.

Eine Teilmenge $M$ von $X$ heißt beschränkt, wenn sie in irgendeiner Kugel enthalten ist.

Die Bezeichnungen „offen“ und „abgeschlossen“ werden sich später erschließen.

## 1.2 Folgen, Reihen und Grenzwerte

Mittels der soeben eingeführten Begriffe können wir nun Konvergenz von Folgen definieren.

**Definition 1.7 (Folgen und Konvergenz).** Sei $(X, d)$ ein metrischer Raum. Eine Folge in $X$ ist eine Abbildung von $\mathbb{N}$ nach $X$. Wir schreiben dafür $(x_n)_{n\in\mathbb{N}}$ oder noch kürzer $(x_n)$.

Die Folge $(x_n)_{n\in\mathbb{N}}$ heißt konvergent gegen $x \in X$, wenn für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass

$$d(x_n, x) < \epsilon \text{ für alle } n > N.$$

Der Punkt $x$ heißt in diesem Fall der Grenzwert der Folge $(x_n)$.

Eine Folge die nicht konvergent ist heißt divergent.

*Bemerkung 1.8.* Auf den reellen Zahlen stimmt die euklidische Metrik mit der Betragsfunktion überein. Man überzeuge sich davon, dass in diesem Fall die neue Definition (Konvergenz von Folgen im metrischen Raum) mit der alten Definition (Konvergenz von reellen Zahlenfolgen) übereinstimmt. Die obige Definition ist also eine echte Verallgemeinerung der Konvergenz von reellen Zahlenfolgen.

**Satz 1.9.** *Sei $(x_n)$ eine Folge in einem metrischen Raum $(X, d)$ und $x \in X$. Dann sind äquivalent:*

*(i) Die Folge $(x_n)$ konvergiert gegen $x$.*

*(ii) Für jedes $\epsilon > 0$ enthält $B_\epsilon(x)$ alle bis auf endlich viele Folgenglieder.*

*(iii) Die Folge (reeller Zahlen) $(d(x_n, x))$ ist eine Nullfolge.*

*Beweis.* Die Aussage (ii) ist lediglich eine Umformulierung der Definition der Konvergenz. Wir zeigen $(i) \iff (iii)$. Es gilt $d(x_n, x) = |d(x_n, x) - 0|$. Damit ist $d(x_n, x) < \epsilon$, genau dann wenn $|d(x_n, x) - 0| < \epsilon$. Die Folge $(x_n)$ konvergiert gegen $x$, genau dann wenn $d(x_n, x)$ gegen 0 konvergiert. □

**Satz 1.10.** *Sei $(x^{(n)}) = (x_1^{(n)}, \ldots, x_d^{(n)})$ eine Folge in $\mathbb{R}^d$ und $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$. Dann konvergiert $(x^{(n)})$ gegen $x$, genau dann wenn für jedes $i \in \{1, \ldots, d\}$ die Folge $(x_i^{(n)})$ gegen $x_i$ konvergiert.*

*Beweis.* Man überzeugt sich leicht, dass $|y_i| \leq \|y\|$ für alle $y \in \mathbb{R}^n$ und alle $i \in \{1, \cdots, d\}$ gilt. Wenn also $(x^{(n)})$ gegen $x$ konvergiert, dann konvergiert nach Satz 1.9 $\left\|x^{(n)} - x\right\|$ gegen 0 und damit auch $\left|x_i^{(n)} - x_i\right| \leq \left\|x^{(n)} - x\right\|$.

Es gelte andererseits $x_i^{(n)} \to x_i$ für alle $i \in \{1, \ldots, d\}$. Sei $\epsilon > 0$ Dann existiert $N \in \mathbb{N}$, so dass $\left|x_i^{(n)} - x_i\right| \leq \frac{\epsilon}{\sqrt{d}}$ für alle $n > N$ und alle $i \in \{1, \ldots d\}$. Es gilt folglich für $n > N$

$$\left\|x^{(n)} - x\right\| = \sqrt{\left(x_1^{(n)} - x_1\right)^2 + \cdots + \left(x_d^{(n)} - x_d\right)^2} \leq \sqrt{\frac{\epsilon^2}{d} + \cdots + \frac{\epsilon^2}{d}} = \epsilon. \qquad \square$$

**Satz 1.11.** *(i) Jede Folge besitzt höchstens einen Grenzwert.*

*(ii) Jede konvergente Folge ist beschränkt.*

*Beweis.* Angenommen $x$ und $y$ seien zwei verschiedene Grenzwerte der Folge $(x_n)$. Aus der Definitheit der Norm folgt dann $d(x, y) > 0$. Da $(x_n)$ sowohl gegen $x$ als auch gegen $y$ konvergiert existiert $N$, so dass $d(x_n, x) < d(x, y)/2$ und $d(x_n, y) < d(x, y)/2$ für alle $n > N$. Damit gilt für jedes $n > N$, dass $d(x_n, x) + d(x_n, y) < d(x, y)$ was jedoch im Widerspruch zur Dreiecksungleichung steht.

Zur Beschränktheit: Wenn $(x_n)$ eine Folge ist, die gegen $x$ konvergiert, so existiert per definitionem ein $N \in \mathbb{N}$, so dass $d(x_n, x) < 1$ für alle $n > N$. Dann gilt $d(x_n, x) \leq \max\{d(x_1, x), \dots, d(x_N, x), 1\}$ für jedes $n \in \mathbb{N}$. □

**Definition 1.12 (Teilfolge und Häufungspunkt).** Sei $(X, d)$ ein metrischer Raum, $(x_n)$ eine Folge in $X$ und $n : \mathbb{N} \to \mathbb{N}$, $k \mapsto n_k$ eine streng monoton wachsende Folge natürlicher Zahlen. Dann heißt $(x_{n_k})_{k\in\mathbb{N}}$ eine Teilfolge von $(x_n)$.

Ein Punkt $x \in X$ heißt Häufungspunkt von $(x_n)$, wenn es eine Teilfolge $(x_{n_k})$ gibt, die gegen $x$ konvergiert.

**Satz 1.13.** *Ersatzlos gestrichen.*

**Satz 1.14.** *Sei $(V, \|\cdot\|)$ ein normierter Raum über dem Vektorraum $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$), $(x_n), (y_n)$ Folgen in $V$ mit Grenzwerten $x$ und $y$ und $(\lambda_n)$ eine Folge in $\mathbb{K}$ mit Grenzwert $\lambda$. Dann konvergiert die Folge $(x_n + y_n)$ gegen $x + y$ und die Folge $\lambda_n x_n$ gegen $\lambda x$.*

*Beweis.* Übung □

**Definition 1.15 (Cauchy-Folge und Vollständigkeit).** Eine Folge $(x_n)$ im metrischen Raum $(X, d)$ heißt Cauchy-Folge (alternativ besitzt die Cauchy-Eigenschaft oder erfüllt die Cauchy-Bedingung), wenn für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass

$$d(x_n, x_m) < \epsilon \text{ für alle } n, m > N.$$

Ein metrischer Raum $(X, d)$ heißt vollständig, wenn jede Cauchy-Folge konvergent ist.

**Satz 1.16.** *Jede konvergente Folge ist eine Cauchy-Folge.*

*Beweis.* Sei $(x_n)$ eine konvergente Folge im metrischen Raum $(X, d)$. Sei $\epsilon > 0$. Da $(x_n)$ konvergent ist, existiert ein Grenzpunkt $x \in X$ und ein $N \in \mathbb{N}$, so dass $d(x_n, x) < \epsilon/2$ für alle $n > N$. Dann gilt für alle $n, m > N$, dass

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

Die Folge $(x_n)$ ist eine Cauchy-Folge. □

*Bemerkung 1.17.*

- Um an Hand der Definition der Konvergenz zu überprüfen ob eine Folge konvergent ist, muss man ihren Grenzpunkt bereits kennen. Um zu überprüfen ob eine Folge eine Cauchy-Folge ist muss kein Grenzpunkt bekannt sein. Darin liegt die Bedeutung der Vollständigkeit, denn sie erlaubt es auf die Existenz eines Grenzpunktes zu schließen ohne diesen zu kennen.
- Ein vollständiger normierter Raum heißt Banachraum.

**Satz 1.18.** *Der $\mathbb{R}^d$ versehen mit der euklidischen Norm ist vollständig (also ein Banachraum).*

*Beweis.* Wir betrachten zunächst den Fall $d = 1$. Sei $(x_n)$ eine Cauchy-Folge reeller Zahlen. Zunächst zeigen wir, dass $(x_n)$ beschränkt ist. Wähle dazu mittels der Cauchy-Eigenschaft $N \in \mathbb{N}$, so dass $|x_n - x_m| < 1$ für alle $n, m \geq N$. Dann gilt für $m \geq N$

$$|x_m| \leq |x_m - x_N| + |x_N| < 1 + |x_N|$$

und damit ist

$$|x_m| \leq \max\{|x_1|, \cdots, |x_{N-1}|, |x_N| + 1\}$$

für alle $m$.

Sei nun $x$ ein Häufungspunkt von $x_n$ (existiert nach dem Satz von Bolzano-Weierstraß aus dem vergangenen Semester da die Folge $x_n$ beschränkt ist). Wir zeigen $x_n \to x$. Sei dazu $\epsilon > 0$ und $(x_{n_k})$ eine Teilfolge die gegen $x$ konvergiert. Wähle wiederum mittels der Cauchy-Eigenschaft ein $N \in \mathbb{N}$, so dass $|x_n - x_m| < \frac{\epsilon}{2}$ für alle $n, m > N$. Wähle nun $k \in \mathbb{N}$, so dass $n_k > N$ und $|x_{n_k} - x| < \frac{\epsilon}{2}$. Dann gilt für jedes $n > n_k$

$$|x_n - x| < |x_n - x_{n_k}| + |x_{n_k} - x| < \epsilon,$$

also konvergiert $(x_n)$ gegen $x$.

Betrachte schließlich den Fall $d > 1$. Für beliebiges $y = (y_1, \ldots, y_d) \in \mathbb{R}^d$ gilt $|y_i| \leq \|y\|$ für jedes $i \in \{1, \ldots, d\}$. Daraus folgt, dass für eine Cauchy-Folge $(x^{(n)})$ in $\mathbb{R}^d$ jede der Koordinatenfolgen $(x_i^{(n)})$ eine Cauchy-Folge in $\mathbb{R}$ ist. Aus der Vollständigkeit der reellen Zahlen folgt damit, dass die Koordinatenfolgen konvergent sind und damit ist nach Satz 1.10 die Folge $x^{(n)}$ konvergent. $\square$

*Bemerkung 1.19.* Im Beweis haben wir die folgenden Aussagen gezeigt (nur für reelle Folgen, sie gelten aber allgemein): Cauchy-Folgen sind beschränkt. Cauchy-Folgen mit einem Häufungspunkt sind konvergent.

**Definition 1.20 (Reihen und (absolute) Konvergenz).** Sei $(V, \|\cdot\|)$ ein normierter Vektorraum und $(x_n)$ eine Folge in $V$. Bezeichne mit

$$s_n := \sum_{k=1}^{n} x_k$$

die $n$-te Partialsumme. Dann heißt die Folge $(s_n)$ die zu $(x_n)$ gehörige Reihe. Ist die Folge $(s_n)$ konvergent, so heißt

$$\sum_{k=1}^{\infty} x_k := \lim_{n\to\infty} s_n$$

der Grenzwert der Reihe (auch Wert oder Summe der Reihe). Häufig verwendet man (formal nicht ganz korrekt) die Notation $\sum_{k=1}^{\infty} x_n$ für die Reihe selbst (egal ob diese konvergiert oder divergiert).

Die Reihe $\sum_{k=1}^{\infty} x_n$ heißt absolut konvergent, wenn $\sum_{k=1}^{\infty} \|x_n\|$ konvergiert.

**Satz 1.21.** *In Banachräumen sind absolut konvergente Reihen konvergent.*

*Beweis.* Sei $(V, \|\cdot\|)$ ein Banachraum und die Reihe $\sum_{n=1}^{\infty} x_n$ in $V$ absolut konvergent. Es reicht zu zeigen, dass die Partialsummenfolge $(s_n)$ eine Cauchy-Folge ist. Sei dazu $\epsilon > 0$. Bezeichne mit

$$t_n := \sum_{k=1}^{n} \|x_k\|$$

die Partialsummenfolge der Reihe $\sum_{n=1}^{\infty} \|x_n\|$. Diese ist monoton wachsend und konvergiert nach Voraussetzung. Insbesondere ist sie also Cauchy-Folge, das heißt es existiert $N \in \mathbb{N}$, so dass $|t_n - t_m| < \epsilon$ für alle $n, m > N$.

Dann gilt für $n > m > N$

$$\|s_n - s_m\| = \left\| \sum_{k=m+1}^{n} x_k \right\| \leq \sum_{k=m+1}^{n} \|x_k\| = t_n - t_m = |t_n - t_m| < \epsilon.$$

Die Partialsummenfolge $(s_n)$ ist also eine Cauchy-Folge und damit konvergent. □

**Theorem 1.22 (Umordnungssatz).** *Sei $(V, \|\cdot\|)$ ein Banachraum. Sei $\sum_{n=1}^{\infty} x_n$ eine absolut konvergente Reihe und $\sigma : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Dann ist die umgeordnete Reihe $\sum_{n=1}^{\infty} x_{\sigma(n)}$ ebenfalls konvergent mit Grenzwert $\sum_{n=1}^{\infty} x_n$.*

*Beweis.* Sei $\epsilon > 0$. Sei $(s_n)$ die zu $(x_n)$ und $(t_n)$ die zu $(\|x_n\|)$ gehörige Partialsummenfolge. Nach Voraussetzung ist $(t_n)$ konvergent, also insbesondere Cauchy-Folge, das heißt es existiert $N$, so dass $\|t_n - t_m\| < \frac{\epsilon}{2}$ für $n, m \geq N$. Nach Satz 1.21 konvergiert $(s_n)$, es existiert also $s$ und ein (eventuell größeres) $N$, so dass $\|s_n - s\| < \frac{\epsilon}{2}$ für $n \geq N$. Setze

$A := \sigma^{-1}\{1, \ldots, N\}$ und $M := \max A$. Dann gilt für jedes $m > M$

$$\begin{aligned}
\left\| \sum_{k=1}^{m} x_{\sigma(k)} - s \right\| &= \left\| \sum_{k=1, k\in A}^{m} x_{\sigma(k)} + \sum_{k=1, k\notin A}^{m} x_{\sigma(k)} - s \right\| \\
&\leq \left\| \sum_{l=1}^{N} x_l - s \right\| + \sum_{k=1, k\notin A}^{m} \left\| x_{\sigma(k)} \right\| \\
&\leq \| s_N - s \| + \sum_{l=N+1}^{L} \| x_l \| \\
&< \frac{\epsilon}{2} + t_L - t_N < \epsilon.
\end{aligned}$$

Wobei $L \geq \max \sigma(\{1, \ldots m\})$ sein soll. Damit ist die Konvergenz gezeigt. □

*Bemerkung 1.23.* Da wir obigen Satz insbesondere auf die Reihe $\sum_{n=1}^{\infty} \left\| x_{\sigma(n)} \right\|$ anwenden können, ist die umgeordnete Reihe auch absolut konvergent.

## 1.3 Offene und abgeschlossene Mengen

**Definition 1.24 (Innere Punkte und Randpunkte).** Sei $(X, d)$ ein metrischer Raum, $A \subset X$ und $x \in X$.

(i) Der Punkt $x$ heißt innerer Punkt von $A$, wenn es ein $\epsilon > 0$ gibt, so dass $B_\epsilon(x) \subset A$,

(ii) Der Punkt $x$ heißt Randpunkt von $A$, wenn für jedes $\epsilon > 0$ die Kugel $B_\epsilon(x)$ mindestens je einen Punkt aus $A$ sowie aus $X \backslash A$ enthält.

(iii) Die Menge aller inneren Punkte von $A$ heißt das Innere von $A$ und wird mit $\mathring{A}$ bezeichnet.

(iv) Die Menge aller Randpunkte von $A$ heißt Rand von $A$ und wird mit $\partial A$ bezeichnet.

(v) Die Menge $A \cup \partial A$ heißt Abschluss von $A$ und wird mit $\overline{A}$ bezeichnet

*Bemerkung 1.25.*

- Aus der Definition ist ersichtlich, dass ein innerer Punkt von $A$ stets selbst in $A$ liegt es gilt also $\mathring{A} \subset A$.
- Randpunkte von $A$ können in $A$ liegen, müssen es jedoch nicht.
- Jeder Punkt in $A$ ist entweder ein innerer Punkt oder ein Randpunkt von $A$ (niemals beides). Es gilt also $A \subset \mathring{A} \cup \partial A$ und $\mathring{A} \cap \partial A = \emptyset$.
- Man überzeugt sich leicht, dass der Rand von $A$ und der Rand von $X \backslash A$ übereinstimmen.

**Satz 1.26.** *Seit $(X, d)$ ein metrischer Raum, $A \subset X$ und $x \in X$ innerer Punkt oder Randpunkt von $A$. Dann gibt es eine Folge $(x_n)$ in $A$ die gegen $x$ konvergiert.*

*Beweis.* Falls $x \in A$ (also insbesondere falls $x$ innerer Punkt von $A$ ist), so liegt die konstante Folge $(x)_{n\in\mathbb{N}}$ in $A$ (und ist natürlich konvergent gegen $x$). Sei also $x \notin A$ (und damit Randpunkt von $A$). Für jedes $n \in \mathbb{N}$ wähle einen Punkt $x_n \in B_{\frac{1}{n}}(x) \cap A$. Diese Menge ist nicht leer, da $x$ Randpunkt von $A$ ist. Wir zeigen nun $x_n \to x$. Sei dazu $\epsilon > 0$. Dann existiert $N \in \mathbb{N}$, so dass $\frac{1}{N} < \epsilon$. Sei $n > N$ dann gilt nach der Definition der Folge $(x_n)$, dass $x_n \in B_{\frac{1}{n}}(x) \subset B_{\frac{1}{N}}(x)$ also

$$d(x_n, x) < \frac{1}{N} < \epsilon.$$

Damit konvergiert die Folge gegen $x$. □

**Definition 1.27 (Offene und abgeschlossene Mengen).** Sei $(X, d)$ ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt

(i) offen wenn jeder ihrer Punkte innerer Punkt von $A$ ist,

(ii) abgeschlossen wenn die Menge $A$ jeden ihrer Randpunkte enthält.

**Definition 1.28 (Umgebung).** Sei $(X, d)$ ein metrischer Raum und $x \in X$. Eine Menge $U \subset X$ heißt Umgebung von $x$ genau dann wenn ein $\epsilon > 0$ existiert, so dass $B_\epsilon(x) \subset U$ ist.

*Bemerkung 1.29.* Aus der Definition ist ersichtlich, dass $U$ genau dann Umgebung von $x$ ist, wenn $x$ innerer Punkt von $U$ ist. Insbesondere sind offene Teilmengen Umgebungen für jedes ihrer Elemente.

**Satz 1.30.** *Eine Teilmenge $A$ eines metrischen Raumes $(X, d)$ ist offen genau dann, wenn ihr Komplement $X \backslash A$ abgeschlossen ist.*

*Beweis.* Per definitionem ist die Menge $A$ offen, genau dann wenn sie nur aus inneren Punkten von $A$ besteht. Da jeder Punkt in $A$ entweder innerer Punkt oder Randpunkt von $A$ ist, ist das der Fall genau dann wenn $A$ keinen Randpunkt von $A$ enthält. Da der Rand von $A$ und der Rand von $X \backslash A$ übereinstimmen, gilt das genau dann wenn $A$ keinen Randpunkt von $X \backslash A$ enthält also genau dann wenn $X \backslash A$ alle Randpunkte von $X \backslash A$ enthält. Das ist per definitionem der Fall genau dann wenn $X \backslash A$ abgeschlossen ist. □

*Beispiel 1.31.* Sei $(X, d)$ ein metrischer Raum, $x \in X$ und $r > 0$. Die offene Kugel $B_r(x) = \{y \in X |\ d(x, y) < r\}$ ist eine offene, die abgeschlossene Kugel $\{y \in X |\ d(x, y) \leq r\}$ eine abgeschlossene Menge.

*Beweis.* Übung □

*Bemerkung 1.32.* Ein Spezialfall des vorhergehenden Beispiels sind Intervalle der Form $(a, b)$ – offene Intervalle – und $[a, b]$ – abgeschlossene Intervalle – für $a < b$ Elemente der reellen Zahlen.

**Satz 1.33.** *Sei $(X, d)$ ein metrischer Raum und $A \subset X$. Die Menge $A$ ist abgeschlossen genau dann, wenn für jede konvergente Folge $(x_n)$ in $A$ auch ihr Grenzwert in $A$ liegt.*

*Beweis.* Sei zunächst $A$ abgeschlossen, $(x_n) \subset A$ eine Folge und $x$ ihr Grenzwert. Für beliebiges $\epsilon > 0$ existiert $n \in \mathbb{N}$, so dass $x_n \in B_\epsilon(x)$. Das heißt, in jeder $\epsilon$-Kugel liegt mindestens ein Punkt von $A$, $x$ ist also innerer Punkt von $A$ oder Randpunkt von $A$. Innere Punkte von $A$ liegen stets in $A$, Randpunkte von $A$ liegen ebenfalls in $A$ da $A$ abgeschlossen ist.

Erfülle nun $A$ die Eigenschaft, dass der Grenzwert jeder konvergenten Folge in $A$ wieder in $A$ liegt. Sei $x \in \partial A$. Dann existiert nach Satz 1.26 eine Folge $(x_n)$ in $A$ die gegen $x$ konvergiert. Nach Voraussetzung ist dann $x \in A$. Damit haben wir $\partial A \subset A$ gezeigt, $A$ ist also abgeschlossen. □

**Satz 1.34.** *Sei $(X, d)$ ein metrischer Raum und $\tau$ die Menge aller offenen Teilmengen von $X$. Es gelten:*

*(i) die leere Menge $\emptyset$ und der ganze Raum $X$ sind in $\tau$,*

*(ii) die Vereinigung beliebig vieler Mengen aus $\tau$ ist wiederum in $\tau$,*

*(iii) der Schnitt endlich vieler Mengen aus $\tau$ ist wiederum in $\tau$.*

*Beweis.* Übung □

*Bemerkung 1.35.*

- Das System von Teilmengen $\tau$ in obigem Satz bezeichnet man als die von der Metrik $d$ induzierte Topologie. (Topologieen sind Strukturen – noch allgemeiner als metrische Räume – die es erlauben über Begriffe wie Konvergenz und Stetigkeit zu reden)
- Der obige Satz impliziert für abgeschlossene Mengen mittels Satz 1.30 und der De Morganschen Regeln:
  (i) die leere Menge $\emptyset$ und der ganze Raum $X$ sind abgeschlossen,
  (ii) der Schnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen,
  (iii) die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.
- Unendliche Schnitte offener Mengen müssen nicht mehr offen, unendliche Vereinigungen abgeschlossener Mengen nicht mehr abgeschlossen sein wie die folgenden Beispiele zeigen
$$\bigcap_{n\in\mathbb{N}} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}$$
$$\bigcup_{n\in\mathbb{N}} [1/n, 2] = (0, 2]\,.$$

**Satz 1.36.** *Sei $(X, d)$ ein metrischer Raum und $A \subset X$. Die Menge $\mathring{A}$ ist offen, die Mengen $\partial A$ und $\overline{A}$ sind abgeschlossen.*

*Beweis.* Sei zunächst $x \in \mathring{A}$. Wir wissen, dass $x$ innerer Punkt von $A$ ist und wir wollen zeigen, dass es innerer Punkt von $\mathring{A}$ ist. Nach Voraussetzung existiert $\epsilon > 0$, so dass $B_\epsilon(x) \subset A$. Sei nun $y \in B_\epsilon(x)$. Nach Beispiel 1.31 ist $B_\epsilon(x)$ offen, das heißt es existiert $\delta > 0$, so dass $B_\delta(y) \subset B_\epsilon(x) \subset A$ ist. Damit ist $y$ innerer Punkt von $A$ und wir haben gezeigt, dass $B_\epsilon(x) \subset \mathring{A}$. Dann ist aber $x$ innerer Punkt von $\mathring{A}$ und $\mathring{A}$ ist offen.

Setze nun $B := X \backslash A$. Dann gilt (Bemerkung 1.25) $\partial A = \partial B$. Damit erhält man

$$X = A \cup B = \mathring{A} \cup \partial A \cup \mathring{B} \cup \partial B = \mathring{A} \cup \partial A \cup \mathring{B} \tag{1.2}$$

Da $A$ und $B$ disjunkt sind, gilt das auch für $\mathring{A} \subset A$ und $\mathring{B} \subset B$. Außerdem ist $\partial A = \partial B$ disjunkt zu $\mathring{A}$ und $\mathring{B}$ (nochmals Bemerkung 1.25). Wir haben also die disjunkte Zerlegung $X = \mathring{A} \cup \partial A \cup \mathring{B}$. Damit sind die Mengen $\mathring{A} \cup \partial A = X \backslash \mathring{B}$ und $\partial A = X \backslash (\mathring{A} \cup \mathring{B})$ abgeschlossen. □

**Satz 1.37.** *Sei $(X, d)$ ein metrischer Raum und $A \subset X$. Es gelten*

$$\bigcup_{O \text{ offen}, O \subset A} O = \mathring{A} \text{ und}$$

$$\bigcap_{C \text{ abgeschlossen}, A \subset C} C = \overline{A}$$

*das heißt $\mathring{A}$ ist die größte offene Teilmenge von $A$, $\overline{A}$ ist die kleinste abgeschlossene Teilmenge von $X$ die $A$ enthält.*

*Beweis.* Da $\mathring{A}$ offen und in $A$ enthalten ist gilt in jedem Fall

$$\bigcup_{O \subset A \text{ offen}} O \supset \mathring{A}.$$

Sei nun $O \subset A$ offen und $x \in O$. Dann existiert $\epsilon > 0$, so dass $B_\epsilon(x) \subset O \subset A$ also ist $x$ innerer Punkt von $A$. Da das für jedes $x \in O$ gilt haben wir gezeigt, dass $O \subset \mathring{A}$ ist. Da das für jedes offene $O \subset A$ gilt ist auch

$$\bigcup_{O \subset A \text{ offen}} O \subset \mathring{A}.$$

Die zweite Aussage erhält man nun durch Komplementbildung. □

**Satz 1.38.** *Sei $(X, d)$ ein vollständiger metrischer Raum und $A \subset X$ abgeschlossen. Dann ist $(A, d)$ ebenfalls vollständig.*

*Beweis.* Übung □

*Bemerkung 1.39.* Es ist wichtig sich bewusst zu machen, dass topologische Eigenschaften – Konvergenz einer Folge $(x_n)$, Offenheit/Abgeschlossenheit einer Menge $A$ (später Stetigkeit von Funktionen, Kompaktheit von Mengen ...) – nicht nur von der Folge $(x_n)$ oder Menge $A$ abhängen, sondern auch vom umgebenden metrischen Raum $(X, d)$.

So ist die Folge $\left(\frac{1}{n}\right)$ im Raum $(\mathbb{R}, |\cdot|)$ konvergent, in $((0,1), |\cdot|)$ jedoch nicht (der Grenzwert liegt nicht im Raum).

In ähnlicher Weise ist $(0, 1)$ im Raum $((0,1), |\cdot|)$ abgeschlossen (der ganze Raum ist stets abgeschlossen), nicht jedoch im Raum $(\mathbb{R}, |\cdot|)$.

## 1.4 Grenzwerte von Abbildungen

**Definition 1.40 (Häufungspunkt, isolierter Punkt).** Sei $(X, d)$ ein metrischer Raum und $x \in X$. Der Punkt $x$ heißt

(i) isolierter Punkt, wenn es ein $\epsilon > 0$ gibt, so dass $B_\epsilon(x) = \{x\}$,

(ii) Häufungspunkt, wenn er kein isolierter Punkt ist.

**Definition 1.41 (Punktierte Umgebung, punktierte Kugel).** Sei $(X, d_X)$ ein metrischer Raum und $x \in X$ ein Häufungspunkt. Eine Menge $\dot{U}$ heißt punktierte Umgebung von $x$, wenn $\dot{U} \cup \{x\}$ eine Umgebung von $x$ ist. Sei $r > 0$ dann ist die punktierte Kugel definiert als $\dot{B}_r(x) = B_r(x) \setminus \{x\}$.

*Bemerkung 1.42.* Durch die Voraussetzung, dass $x$ ein Häufungspunkt von $X$ ist, enhält jede $\epsilon$-Kugel um $x$ noch mindestens einen weiteren Punkt. Die punktierten Kugeln um $x$, und damit die punktierten Umgebungen um $x$, die jeweils eine punktierte Kugel enthalten müssen, sind also nicht leer.

**Definition 1.43 (Grenzwert von Abbildungen).** Seien $(X, d_X)$, $(Y, d_Y)$ metrische Räume, $x \in X$ ein Häufungspunkt und $\dot{U}$ eine punktierte Umgebung von $x$. Sei $f : \dot{U} \to Y$ eine Abbildung. Dann konvergiert $f$ in $x$ gegen $y \in Y$ (und $y$ heißt Grenzwert von $f$ im Punkt $x$), wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so dass $f(x') \in B_\epsilon(y)$ für alle $x' \in \dot{B}_\delta(x)$. Wir schreiben $y = \lim_{x' \to x} f(x')$.

*Beispiel 1.44.* Definiere $f, g : \mathbb{R} \setminus \{0\} \to \mathbb{R}$

$$f(x) := x^2$$
$$g(x) := \frac{1}{x}.$$

Dann ist $f$ in 0 konvergent (Wogegen?), $g$ jedoch nicht (Warum?).

**Satz 1.45.** *Seien $(X, d_X)$, $(Y, d_Y)$ metrische Räume, $x \in X$ ein Häufungspunkt, $\dot{U}$ eine punktierte Umgebung von $x$ und $f : \dot{U} \to Y$ eine Abbildung. Die Abbildung $f$ konvergiert in $x$ gegen $y \in Y$, genau dann wenn für jede Folge $(x_n)$ in $\dot{U}$ mit $x_n \to x$ gilt, dass $(f(x_n))$ gegen $y$ konvergiert.*

*Beweis.* Es gelte zunächst $\lim_{x' \to x} f(x') = y$ und $(x_n)$ sei eine gegen $x$ konvergente Folge. Sei $\epsilon > 0$ dann existiert nach Voraussetzung $\delta > 0$, so dass $f(x') \in B_\epsilon(y)$ für alle $x' \in \dot{B}_\delta(x)$. Da $x_n \to x$ existiert $N \in \mathbb{N}$, so dass $x_m \in \dot{B}_\delta(x)$ für alle $m > N$. Dann gilt aber $f(x_m) \in B_\epsilon(y)$ für alle $m > N$ und wir haben gezeigt, dass $(f(x_n))$ gegen $y$ konvergiert.

Es gelte nun dass $f(x_n) \to y$, wann immer die Folge $(x_n)$ aus $\dot{U}$ gegen $x$ konvergiert. Angenommen $f$ konvergiert in $x$ nicht gegen $y$, das heißt es existiert ein $\epsilon > 0$ und für jedes $n \in \mathbb{N}$ ein $x_n \in B_{\frac{1}{n}}(x)$ mit $f(x_n) \notin B_\epsilon(y)$. Die Folge $x_n$ konvergiert gegen $x$ da $d_X(x_n, x) < \frac{1}{n} \to 0$. Die Folge $f(x_n)$ konvergiert nicht gegen $y$, da $d_Y(f(x_n), y) \geq \epsilon$. Da das im Widerspruch zur Voraussetzung steht, muss $f$ in $x$ gegen $y$ konvergieren. $\square$

**Korollar 1.46.** *Seien* $(X, d)$ *ein metrischer Raum,* $x \in X$ *ein Häufungspunkt,* $\dot{U}$ *eine punktierte Umgebung von* $x$ *und* $f, g : \dot{U} \to \mathbb{K}$ *(*$\mathbb{K} = \mathbb{R}$ *oder* $\mathbb{K} = \mathbb{C}$*) Funktionen. Es gelte* $\lim_{x' \to x} f(x') = y_1$ *und* $\lim_{x' \to x} g(x') = y_2$*. Dann konvergiert* $f + g$ *in* $x$ *gegen* $y_1 + y_2$ *und* $f \cdot g$ *gegen* $y_1 y_2$*. Ist* $y_1 \neq 0$*, dann ist* $1/f$ *auf einer punktierten Umgebung von* $x$ *definiert und* $1/f$ *konvergiert gegen* $1/y_1$*.*

*Beweis.* Sei $(x_n)$ eine beliebige Folge in $X$ mit $x_n \to x$. Dann gilt nach Voraussetzung $f(x_n) \to y_1$ und $g(x_n) \to y_2$. Daraus folgt mittels der Konvergenzregeln für Zahlenfolgen $f(x_n) + g(x_n) \to y_1 + y_2$ und $f(x_n)g(x_n) \to y_1 y_2$. Damit gilt nach Satz 1.45, dass die Funktionen $\lim_{x' \to x} f + g = y_1 + y_2$, sowie $\lim_{x' \to x} f \cdot g = y_1 y_2$.

Sei nun zusätzlich $y_1 \neq 0$, das heißt $|y_1| > 0$. Da $f$ in $x$ gegen $y_1$ konvergiert, existiert $\delta > 0$, so dass $f(x') \in B_{\frac{|y_1|}{2}}(y_1)$ für alle $x' \in \dot{B}_\delta(x)$. Für solche $x'$ gilt also insbesondere (umgekehrte Dreiecksungleichung) $|f(x')| > |y_1| - |f(x') - y_1| > \frac{y_1}{2}$. Damit ist die Funktion $1/f$ auf der punktierten Umgebung $\dot{B}_\delta(x)$ von $x$ definiert. Es gilt jetzt wieder für jede gegen $x$ konvergente Folge $(x_n)$, dass $f(x_n) \to y_1$ und damit $\frac{1}{f(x_n)} \to \frac{1}{y_1}$. Wir verwenden nochmals Satz 1.45 und erhalten $\lim_{x' \to x} \frac{1}{f(x')} = \frac{1}{y_1}$. $\square$

## 1.5 Stetigkeit

Der Begriff der Stetigkeit formalisiert die Eigenschaft einer Funktion $f$, dass sich die Funktionswerte $f(x)$ nicht zu stark ändern, wenn sich das Argument $x$ nicht zu stark ändert. Je nach dem was genau man unter „zu stark ändern“ versteht ergeben sich unterschiedliche Stetigkeitsbegriffe die in unterschiedlichen Situationen nützlich sein können und von denen wir einige im Folgenden untersuchen wollen.

Viele funktionale Zusammenhänge die in der Natur vorkommen (Position eines Objektes in Abhängigkeit von der Zeit, Lufttemperatur in Abhängigkeit vom Ort, ...) sind stetig.

**Definition 1.47 (Stetigkeit).** Seien $(X, d_X)$ und $(Y, d_Y)$ metrische Räume. Eine Abbildung $f : X \to Y$ heißt stetig im Punkt $x_0 \in X$, wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so dass $f(x) \in B_\epsilon(f(x_0))$ für alle $x \in B_\delta(x_0)$. Wir sagen $f$ ist stetig (ohne angabe eines Punktes), wenn $f$ in jedem Punkt von $X$ stetig ist.

*Beispiel 1.48.* (i) Sei $(X, d)$ ein metrischer Raum. Die Identitätsabbildung $\mathrm{id} : X \to X$, $x \mapsto x$ ist stetig in jedem Punkt.

(ii) Die Funktion $\chi_\mathbb{Q} : \mathbb{R} \to \mathbb{R}$ definiert durch

$$\chi_\mathbb{Q}(x) := \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & \text{sonst} \end{cases}$$

ist in keinem Punkt stetig.

*Bemerkung 1.49.* (i) Leicht umformuliert bedeutet die Stetigkeit von $f$ in $x$, dass für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so dass $f(B_\delta(x)) \subset B_\epsilon(f(x))$.

(ii) Stetigkeit (im Punkt $x_0$) ist eine sogenannte lokale Eigenschaft, das heißt ob $f$ in $x_0$ stetig ist oder nicht kann man entscheiden, wenn man $f$ nur in einer beliebig kleinen Umgebung von $x_0$ kennt. Formal lautet die Aussage: Seien $f, g$ Funktionen von $(X, d_X)$ nach $(Y, d_Y)$, $U$ eine Umgebung von $x_0$ und $f|_U = g|_U$ Dann ist $f$ in $x_0$ stetig, genau dann wenn $g$ in $x_0$ stetig ist.

Um das zu zeigen, sei $f$ stetig und $\epsilon > 0$. Dann existiert $\delta > 0$, so dass $f(B_\delta(x_0)) \subset B_\epsilon(f(x_0))$. Da $U$ Umgebung von $x$ ist, existiert $0 < \rho \leq \delta$, so dass $B_\rho(x_0) \subset U$ ist. Dann gilt

$$g(B_\rho(x_0)) = f(B_\rho(x_0)) \subset f(B_\delta(x_0)) \subset B_\epsilon(f(x_0)) = B_\epsilon(g(x_0)).$$

Damit ist $g$ stetig in $x_0$.

Durch Vertauschen von $f$ und $g$ erhält man „genau dann wenn“.

**Satz 1.50.** *Seien $(X, d_X)$, $(Y, d_Y)$ und $(Z, d_Z)$ metrische Räume, $x_0 \in X$, $g : X \to Y$ und $f : Y \to Z$ Abbildungen. Wenn $g$ stetig in $x_0$ ist und $f$ stetig in $g(x_0)$, dann ist $f \circ g$ stetig in $x_0$.*

*Beweis.* Sei $\epsilon > 0$. Nach Voraussetzung existiert $\delta > 0$, so dass $f(y) \in B_\epsilon(f(g(x_0)))$ für alle $y \in B_\delta(g(x_0))$. Außerdem existiert $\rho > 0$, so dass $g(x) \in B_\delta(g(x_0))$ für alle $x \in B_\rho(x_0)$. Dann gilt für jedes $x \in B_\rho(x_0)$, dass $g(x) \in B_\delta(g(x_0))$ und damit auch $f \circ g(x) = f(g(x)) \in B_\epsilon(f \circ g(x_0))$. □

**Definition 1.51 (Lipschitz-Stetigkeit, Kontraktion, Isometrie).** Seien $(X, d_X)$ und $(Y, d_Y)$ metrische Räume. Eine Abbildung $f : X \to Y$ heißt Lipschitz-stetig, wenn es ein $L > 0$ gibt, so dass

$$d_Y(f(x), f(y)) \leq L \cdot d_X(x, y)$$

für alle $x, y \in X$ gilt. Die Konstante $L$ heißt dann Lipschitz-Konstante für die Funktion $f$. Die Abbildung $f$ heißt Kontraktion falls $d_Y(f(x), f(y)) \leq d_X(x, y)$ und Isometrie, falls $d_Y(f(x), f(y)) = d_X(x, y)$.

*Beispiel 1.52.* Betrachte die Funktionen $f, g : [0, 1] \to \mathbb{R}$

$$\begin{aligned} f(x) &:= x^2 \\ g(x) &:= \sqrt{x}. \end{aligned}$$

Die Funktion $f$ ist Lipschitz-stetig da

$$|f(x) - f(y)| = \left|x^2 - y^2\right| = |x - y|\,|x + y| \leq 2\,|x - y|\,.$$

Die Funktion $g$ ist nicht Lipschitz-stetig, da

$$\frac{|g(x) - g(0)|}{|x - 0|} = \frac{\sqrt{x}}{x} = \frac{1}{\sqrt{x}} \to \infty$$

für $x$ gegen 0. Wie wir noch sehen werden, ist $g$ stetig.

**Satz 1.53.** *Jede Lipschitz-stetige Funktion ist stetig.*

*Beweis.* Seien $(X, d_X), (Y, d_Y)$ metrische Räume, $x \in X$ und $f : X \to Y$ eine Lipschitz-stetige Funktion mit Lipschitz-Konstante $L$. Sei $\epsilon > 0$. Dann gilt für jedes $x' \in B_{\frac{\epsilon}{L}}(x)$

$$d_Y(f(x'), f(x)) \leq L \cdot d_X(x', x) < \epsilon.$$

Damit ist die Stetigkeit gezeigt. □

*Beispiel 1.54.* (i) Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Die Abbildung $p_i : \mathbb{K}^n \to \mathbb{K}$

$$x = (x_1, \ldots, x_n) \mapsto x_i$$

ist eine Kontraktion und damit stetig.

(ii) Sei $(X, d)$ ein metrischer Raum und $(X \times X, d_{X\times X})$ das Produkt von $X$ mit sich selbst und der Produktmetrik (Beispiel 1.4 (iii)). Wir zeigen, dass die Metrik, also die Abbildung

$$\begin{aligned} d : X \times X &\to \mathbb{R} \\ (x, y) &\mapsto d(x, y) \end{aligned}$$

eine Kontraktion von $(X \times X, d_{X\times X})$ nach $\mathbb{R}$ ist (insbesondere also stetig).

Für $x_1, x_2, y_1, y_2 \in X$ gilt nach der Dreiecksungleichung

$$d(x_1, y_1) \leq d(x_1, x_2) + d(x_2, y_1) \leq d(x_1, x_2) + d(y_2, y_1) + d(x_2, y_2).$$

Damit erhalten wir

$$|d(x_1, y_1) - d(x_2, y_2)| \leq d(x_1, x_2) + d(y_1, y_2) = d_{X\times X}((x_1, x_2), (y_1, y_2)).$$

(iii) Analog gilt mit der umgekehrten Dreiecksungleichung

$$|\|x\| - \|y\|| \leq \|x - y\|$$

und $\|\cdot\|$ ist eine Kontraktion von $(V, \|\cdot\|)$ nach $\mathbb{R}$ also insbesondere stetig.

**Satz 1.55.** *Seien $(X, d_X)$ und $(Y, d_Y)$ metrische Räume und $f : X \to Y$. Die Abbildung $f$ ist stetig in $x \in X$ genau dann, wenn entweder $x$ isolierter Punkt von $X$ ist, oder $\lim_{x'\to x} f(x') = f(x)$ gilt.*

*Beweis.* Sei zunächst $f$ stetig. Wenn $x$ isolierter Punkt ist, ist nichts zu zeigen. Sei also $x$ Häufungspunkt von $X$. Sei $\epsilon > 0$, dann existiert wegen der Stetigkeit $\delta > 0$, so dass $f(x') \in B_\epsilon(f(x))$ für alle $x' \in B_\delta(x)$ also insbesondere für alle $x' \in \dot{B}_\delta(x)$. Damit gilt $\lim_{x'\to x} f(x') = f(x)$.

Es sei nun $x$ isolierter Punkt von $X$, das heißt es existiert $\delta > 0$, so dass $B_\delta(x) = \{x\}$. Dann ist $f(B_\delta(x)) = \{f(x)\} \subset B_\epsilon(f(x))$ für jedes $\epsilon > 0$ und damit $f$ stetig in $x$.

Sei schließlich $x$ Häufungspunkt von $X$ und es gelte $\lim_{x'\to x} f(x') = f(x)$. Die Grenzwertbedingung bedeutet gerade, dass für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so dass $f(\dot{B}_\delta(x)) \subset B_\epsilon(f(x))$. Dann gilt aber auch $f(B_\delta(x)) \subset B_\epsilon(f(x))$ also ist $f$ stetig in $x$. □

Mit dem obigen Satz können wir nun Aussagen über Grenzwerte von Funktionen in entsprechende Aussagen über stetige Funktionen übersetzen. So zum Beispiel die folgenden Korollare deren Beweise dem Leser zur eigenständigen Übung überlassen werden. Dazu macht man sich zunächst bewusst, dass eine Folge $(x_n)$ genau dann gegen einen isolierten Punkt $x$ konvergiert, wenn sie ab einem bestimmten Index konstant gleich $x$ ist.

**Korollar 1.56.** *Seien $(X, d_X)$ und $(Y, d_Y)$ metrische Räume und $f : X \to Y$ eine Abbildung. Dann ist $f$ stetig in $x \in X$, genau dann wenn für jede Folge $(x_n)$ in $X$, die gegen $x$ konvergiert auch $(f(x_n))$ gegen $f(x)$ konvergiert.*

**Korollar 1.57.** *Seien* $f_1, \dots, f_n : \mathbb{R}^m \to \mathbb{R}$ *Funktionen. Dann ist die Abbildung* $f : \mathbb{R}^m \to \mathbb{R}^n$, $x \mapsto (f_1(x), \dots, f_n(x))$ *stetig im Punkt* $x_0$*, genau dann wenn* $f_i$ *für jedes* $i \in \{1, \dots, n\}$ *stetig im Punkt* $x_0$ *ist.*

**Korollar 1.58.** *Seien* $(X, d)$ *ein metrischer Raum und* $f, g : X \to \mathbb{K}$ *(*$\mathbb{K} = \mathbb{R}$ *oder* $\mathbb{K} = \mathbb{C}$*) Funktionen die stetig in* $x \in X$ *sind, dann sind die Funktionen* $f + g$ *und* $f \cdot g$ *stetig in* $x$*. Ist* $f(x) \neq 0$*, dann ist* $1/f$ *auf einer Umgebung von* $x$ *definiert und in* $x$ *stetig.*

*Beispiel 1.59.* Für $\eta = (\eta_1, \dots, \eta_n) \in \mathbb{N}_0^n$ und $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) definieren wir

$$x^\eta := x_1^{\eta_1} x_2^{\eta_2} \cdots x_n^{\eta_n}.$$

Dabei bezeichnen wir $\eta$ als Multiindex. Wir setzen außerdem $|\eta| := \eta_1 + \cdots + \eta_n$. Seien $c_\eta \in \mathbb{K}$ Konstanten für jedes $\eta$ mit $|\eta| \leq N$, dann heißt die Funktion

$$\begin{aligned} f : \mathbb{K}^n &\to \mathbb{K} \\ x &\mapsto \sum_{\eta, |\eta| \leq N} c_\eta x^\eta \end{aligned}$$

Polynom in $n$ Variablen (über dem Körper $\mathbb{K}$). Solche Polynome sind stetig.

*Bemerkung 1.60.* Im Kontext von Polynomen (und später Potenzreihen) interpretieren wir den Ausdruck $0^0$, der auftreten kann wenn $\eta_i = x_i = 0$ ist, als 1.

Der folgende Satz ist eine Charakterisierung stetiger Funktionen, die insbesondere in Beweisen oft hilfreich ist.

**Satz 1.61.** *Seien* $(X, d_X)$ *und* $(Y, d_Y)$ *metrische Räume und* $f : X \to Y$ *eine Abbildung.*

*(i) Die Abbildung* $f$ *ist stetig in* $x \in X$ *genau dann wenn* $f^{-1}(U)$ *eine Umgebung von* $x$ *ist für jede Umgebung* $U$ *von* $f(x)$*.*

*(ii) Die Abbildung* $f$ *ist stetig (in jedem Punkt), genau dann wenn* $f^{-1}(O)$ *offen ist für jede offene Menge* $O \subset Y$*.*

*(iii) Die Abbildung* $f$ *ist stetig (in jedem Punkt), genau dann wenn* $f^{-1}(C)$ *abgeschlossen ist für jede abgeschlossene Menge* $C \subset Y$*.*

*Beweis.* Übung □

*Beispiel 1.62.* Betrachte die Funktion

$$\begin{aligned} f : \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \|x\| . \end{aligned}$$

Wir wissen bereits, dass diese Funktion stetig ist. Sei $r > 0$. Dann folgt mit dem vorhergehenden Satz, dass die Sphäre mit Radius $r$

$$f^{-1}(\{r\}) = \{x \in \mathbb{R}^n | \ \|x\| = r\}$$

und die abgeschlossene Kugel mit Radius $r$

$$f^{-1}([0,r]) = \{x \in \mathbb{R}^n | \ \|x\| \leq r\} = K_r(0)$$

abgeschlossen sind, und die offene Kugel

$$f^{-1}((-1,r)) = \{x \in \mathbb{R}^n | \ \|x\| < r\} = B_r(0)$$

offen ist.

**Theorem 1.63 (Zwischenwertsatz).** *Seien $a, b \in \mathbb{R}$, $a < b$ und $f : [a,b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) \leq f(b)$. Dann existiert zu jedem $y \in [f(a), f(b)]$ ein $x \in [a,b]$, so dass $f(x) = y$.*

*Beweis.* Sei $y \in [f(a), f(b)]$. Setze

$$x := \sup \left\{x' \in [a,b] \middle| \ f(x') \leq y\right\}$$

und wähle eine Folge $(x_n)$ in $\{x' \in [a,b] | \ f(x') \leq y\}$ die gegen $x$ konvergiert. Da $f$ stetig ist konvergiert $(f(x_n))$ gegen $f(x)$. Wir zeigen $f(x) = y$. Zunächst gilt nach der Definition von $x$, dass $f(x_n) \leq y$ für alle $n \in \mathbb{N}$ und damit gilt für den Grenzwert ebenfalls $f(x) \leq y$. Falls $x = b$, so gilt $f(x) = f(b) \leq y$ und $y \in [f(a), f(b)]$ also $f(x) = f(b) = y$. Falls $x < b$, so konvergiert die Folge $x + \frac{1}{n}$ gegen $x$. Nach der Definition von $x$ ist $f(x + \frac{1}{n}) > y$ (für hinreichend große $n$). Wir verwenden nochmals die Stetigkeit von $f$ und erhalten $f(x) = \lim_{n\to\infty} f(x + \frac{1}{n}) \geq y$. Damit ist $f(x) = y$. $\square$

**Satz 1.64.** *Sei $A \subset \mathbb{R}^n$ eine beschränkte und abgeschlossene Teilmenge und $f : A \to \mathbb{R}^m$ eine stetige Abbildung. Dann ist das Bild von $f$, $f(A)$, beschränkt und abgeschlossen.*

*Beweis.* Wir wollen das Folgenkriterium für Abgeschlossenheit (Satz 1.33) verwenden. Sei dazu $(y_n)$ eine Folge in $f(A)$, die in $\mathbb{R}^m$ gegen $y$ konvergiert. Wir müssen zeigen, dass der Grenzwert $y$ ebenfalls in $f(A)$ liegt. Da $(y_n) \subset f(A)$ existieren $x_n$, so dass $f(x_n) = y_n$ für jedes $n \in \mathbb{N}$. Nach dem Satz von Bolzano-Weierstraß existiert eine konvergente Teilfolge $(x_{n_k})$ von $(x_n)$. Da die Teilfolge $(x_{n_k})$ in $A$ liegt und $A$ abgeschlossen ist, liegt auch ihr Grenzwert $x$ in $A$. Wegen der Stetigkeit von $f$ konvergiert die Folge $(f(x_{n_k}))$ gegen $f(x)$. Andererseits ist $(f(x_{n_k}))$ Teilfolge der konvergenten Folge $f(x_n)$ und konvergiert daher gegen deren Grenzwert $y$. Da Grenzwerte (im metrischen Raum) stets eindeutig bestimmt sind, muss $f(x) = y$ gelten und wir haben gezeigt, dass $y \in f(A)$.

Es bleibt noch die Beschränktheit von $f(A)$ zu zeigen. Angenommen $f(A)$ wäre unbeschränkt, das heißt es ist in keiner noch so großen Kugel um den Koordinatenursprung enthalten. Dann existiert für jedes $n \in \mathbb{N}$ ein $x_n \in A$, so dass $\|f(x_n)\| > n$.

Wir verwenden nochmals Bolzano-Weierstraß um aus $(x_n)$ eine konvergente Teilfolge $(x_{n_k})$ auszuwählen. Wie oben liegt deren Grenzwert in $A$ und wegen der Stetigkeit von $f$ gilt $\lim_{k\to\infty} f(x_{n_k}) = f(x)$. Für jedes $N \in \mathbb{N}$ und jedes $n \in \mathbb{N}$ mit $n > N$ gilt dann $\|f(x_n)\| > n > N$. Damit gilt wegen der Stetigkeit der Norm

$$\|f(x)\| = \lim_{k\to\infty} \|f(x_{n_k})\| > N.$$

Da $f(x) \in \mathbb{R}^m$ ist, kann diese Ungleichung unmöglich für alle $N \in \mathbb{N}$ gelten, $f(A)$ muss also beschränkt sein. □

*Bemerkung 1.65.* (i) Beschränkte und abgeschlossene Mengen im $\mathbb{R}^n$ heißen kompakt. Wir werden uns später eingehender mit Kompaktheit in allgemeineren metrischen Räumen beschäftigen und um bis dahin keine Verwirrung zu stiften, werden wir den Begriff zunächst noch meiden.

(ii) Stetige Abbildungen bilden im Allgemeinen nicht beschränkte Mengen auf beschränkte Mengen ab und auch nicht abgeschlossene auf abgeschlossene.

**Korollar 1.66.** *Sei $A \subset \mathbb{R}^n$ eine beschränkte und abgeschlossene Teilmenge und $f : A \to \mathbb{R}$ stetig, dann nimmt $f$ auf $A$ sein Maximum an, das heißt es existiert $x \in A$, so dass $f(y) \leq f(x)$ für alle $y \in A$.*

*Beweis.* Wir zeigen, dass eine beschränkte, abgeschlossene Teilmenge von $K \subset \mathbb{R}$ ein Maximum hat. Sei dazu $M = \sup K$ und wähle eine Folge $(x_n) \subset K$, so dass $x_n \to M$. Da $K$ abgeschlossen ist liegt der Grenzwert $M$ dieser Folge in $K$.

Wenden wir dass auf die Menge $f(A)$ an, die nach dem vorhergehenden Satz abgeschlossen und beschränkt ist, so folgt die Aussage des Korollars. □

**Satz 1.67.** *Seien $A \subset \mathbb{R}^n$ und $B \subset \mathbb{R}^m$ beschränkte, abgeschlossene Teilmengen und $f : A \to B$ eine stetige, bijektive Funktion. Dann ist die Umkehrfunktion $f^{-1} : B \to A$ ebenfalls stetig.*

*Beweis.* Da $f$ bijektiv ist, besitzt es eine Umkehrabbildung. Setze nun $g = f^{-1}$. Die Abbildung $g$ ist bijektiv da $f$ bijektiv ist. Sei $C \subset A$ abgeschlossen (da $A$ selbst abgeschlossen ist, ist $C$ abgeschlossen in $\mathbb{R}^n$, genau dann wenn es in $A$ abgeschlossen ist). Aus der Beschränktheit von $A$ folgt, dass $C$ ebenfalls beschränkt ist. Wegen der Bijektivität von $g$ gilt

$$\begin{aligned} f(C) &= \{f(x) \in B \mid x \in C\} \\ &= \{f(g(y)) \in B \mid g(y) \in C\} \\ &= \{y \in B \mid g(y) \in C\} = g^{-1}(C). \end{aligned}$$

Nach Satz 1.64 ist $f(C) = g^{-1}(C)$ abgeschlossen (und beschränkt). Da $C$ eine beliebige abgeschlossene Teilmenge von $A$ war, folgt damit aus Satz 1.61 die Stetigkeit von $g = f^{-1}$. □

*Beispiel 1.68.* Sei $n \in \mathbb{N}$. Dann ist die Funktion

$$\begin{aligned} f : [0, \infty) &\to [0, \infty) \\ x &\mapsto \sqrt[n]{x} \end{aligned}$$

stetig.

*Beweis.* Da $[0, \infty)$ nicht beschränkt ist, können wir Satz 1.67 nicht unmittelbar anwenden. Wir nutzen hier aus, dass Stetigkeit eine lokale Eigenschaft ist (Bemerkung 1.49). Sei also $x_0 \in [0, \infty)$.

Definiere

$$\begin{aligned} \tilde{f} : [0, x_0 + 1] &\to \left[0, \sqrt[n]{x_0 + 1}\right] \\ x &\mapsto \sqrt[n]{x}. \end{aligned}$$

Dann ist $\tilde{f}$ die Umkehrfunktion von

$$\begin{aligned} g : \left[0, \sqrt[n]{x_0 + 1}\right] &\to [0, x_0 + 1] \\ y &\mapsto y^n. \end{aligned}$$

Die Funktion $g$ ist stetig und streng monoton (also insbesondere bijektiv). Damit folgt nach Satz 1.67 die Stetigkeit von $\tilde{f}$, insbesondere im Punkt $x_0$. Da $f$ und $\tilde{f}$ in einer Umgebung von $x_0$ übereinstimmen, ist also auch $f$ in $x_0$ stetig. Da $x_0 \in [0, \infty)$ beliebig war, ist $f$ überall stetig. $\square$

## 1.6 Konvergenz von Funktionenfolgen

**Definition 1.69 (Punktweise Konvergenz).** Sei $M$ eine Menge, $f_n : M \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) Funktionen für jedes $n \in \mathbb{N}$ und $f : M \to \mathbb{K}$ eine weitere Funktion. Die Folge $(f_n)$ heißt punktweise konvergent gegen $f$, wenn gilt

$$\lim_{n\to\infty} f_n(x) = f(x) \text{ für alle } x \in M.$$

Punktweise Konvergenz von Funktionenfolgen ist oft vergleichsweise einfach zu zeigen, aber oft auch nicht sehr nützlich, da viele wichtige Eigenschaften von Funktionen bei punktweiser Konvergenz nicht erhalten bleiben. Die punktweise Konvergenz läßt sich im Allgemeinen – konkreter falls $M$ überabzählbar ist – auch nicht durch eine Metrik charakterisieren.

*Beispiel 1.70.* Seien $f_n : [0,1] \to \mathbb{R}$ die folgenden Funktionen

$$f_n(x) := \begin{cases} 1 - nx & x \in [0, 1/n] \\ 0 & \text{sonst.} \end{cases}$$

Dann ist $f_n$ stetig für jedes $n \in \mathbb{N}$ und die Folge $(f_n)$ konvergiert punktweise gegen $f : [0,1] \to \mathbb{R}$

$$f(x) := \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases}$$

welches nicht mehr stetig ist.

Wir können nun den Formalismus des normierten Raumes nutzen, um einen anderen Konvergenzbegriff für Funktionenfolgen zu definieren, der viele gute Eigenschaften von Funktionen erhält.

*Beispiel 1.71.* Sei $M$ eine Menge und $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Betrachte den Raum der beschränkten Funktionen über $M$, das heißt

$$\mathbf{B}(M) = \{f : M \to \mathbb{K} \mid \exists K \in \mathbb{R} \text{ so dass } |f(x)| \leq K \text{ für alle } x \in M\}. \tag{1.3}$$

Mit der punktweisen Addition und skalaren Multiplikation, das heißt für $f, g \in \mathbf{B}(M)$, $\lambda \in \mathbb{K}$ und für alle $x \in M$

$$(f+g)(x) := f(x) + g(x)$$
$$(\lambda f)(x) := \lambda f(x)$$

erhält $\mathbf{B}(M)$ die Struktur eines Vektorraumes.

Dann ist

$$\|\cdot\|_\infty : \mathbf{B}(M) \to \mathbb{R}$$
$$f \mapsto \sup_{x\in M} |f(x)|$$

eine Norm auf $\mathbf{B}(M)$.

*Beweis.* Wir ziegen, dass $\|\cdot\|_\infty$ eine Norm ist. Die Positivität und absolute Homogenität sind einfach nachzuprüfen. Um die Definitheit zu zeigen sei $f \in \mathbf{B}(M)$ so, dass $\|f\|_\infty = 0$ gilt. Dann ist $|f(x)| = 0$, also $f(x) = 0$ für alle $x \in M$ und damit ist $f$ die Nullfunktion.

Es bleibt noch die Dreiecksungleichung zu zeigen. Zunächst stellen wir fest, dass nach der Definition der Norm $|f(x)| \leq \|f\|_\infty$ für jedes $x \in M$ gilt. Dann folgt für $f, g \in \mathbf{B}(M)$ mit der Dreiecksungleichung für den Betrag

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty .$$

Da das für alle $x \in M$ gilt, ist dann auch

$$\|f + g\|_\infty = \sup_{x \in M} |f(x) + g(x)| \leq \|f\|_\infty + \|g\|_\infty . \qquad \square$$

**Definition 1.72 (Gleichmäßige Konvergenz).** Eine Folge von beschränkten Funktionen $(f_n)$, $f_n : M \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$), die bezüglich der Norm $\|\cdot\|_\infty$ gegen $f : M \to \mathbb{K}$ konvergiert heißt gleichmäßig konvergent. Wenn von dem Funktionenraum $\mathbf{B}(M)$ (und später $\mathbf{C}(M)$) die Rede ist werden wir die Norm $\|\cdot\|_\infty$ oft nicht explizit erwähnen.

*Bemerkung 1.73.* In formaler Notation bedeutet die punktweise Konvergenz einer Funktionenfolge $f_n$ gegen $f$, dass

$$\forall \epsilon > 0 \quad \forall x \in M \quad \exists N \in \mathbb{N} \quad \forall n > N \quad |f_n(x) - f(x)| < \epsilon$$

erfüllt ist, bei gleichmäßiger Konvergenz hingegen gilt

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall x \in M \quad \forall n > N \quad |f_n(x) - f(x)| < \epsilon.$$

Die Definitionen unterscheiden sich also nur darin, ob $N$ von $x$ abhängen darf oder nicht. Insbesondere ist jede gleichmäßig konvergente Folge auch punktweise konvergent.

**Satz 1.74.** *Der Funktionenraum $\mathbf{B}(M)$ ist vollständig, also ein Banachraum.*

*Beweis.* Sei $(f_n) \subset \mathbf{B}(M)$ eine Cauchy-Folge. Wähle $x \in M$ fest. Dann ist die Folge $(f_n(x))$ ebenfalls eine Cauchy-Folge, da

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty$$

gilt. Da $\mathbb{K}$ (also $\mathbb{R}$ oder $\mathbb{C}$) vollständig sind, konvergiert also die Folge $(f_n(x))$. Wir setzen jetzt

$$f(x) := \lim_{n \to \infty} f_n(x).$$

Damit haben wir einen Grenzwert $f$ für die Folge $(f_n)$ gefunden, allerdings nur bezüglich der punktweisen Konvergenz.

Da jede Cauchy-Folge beschränkt ist (Bemerkung 1.19) existiert $K \in \mathbb{R}$, so dass $\|f_n\|_\infty \leq K$ für alle $n \in \mathbb{N}$. Damit folgt für jede $x \in M$

$$|f(x)| = \lim_{n\to\infty} |f_n(x)| \leq \limsup_{n\to\infty} \|f_n\|_\infty \leq K,$$

unsere Grenzfunktion $f$ ist also beschränkt.

Sei nun $\epsilon > 0$. Wir wählen mit der Cauchy-Eigenschaft $N \in \mathbb{N}$ so, dass $\|f_n - f_m\|_\infty < \epsilon$ für alle $n, m \geq N$. Die Wahl von $N$ ist nicht von $x$ abhängig. Dann gilt für jedes $x \in M$ und jedes $n, m > N$

$$|f_m(x) - f_n(x)| \leq \|f_m - f_n\|_\infty \leq \epsilon.$$

Da die rechte Seite nicht mehr von $m$ abhängt, können wir den Grenzwert $m \to \infty$ bilden und erhalten, da die Betragsfunktion stetig ist

$$|f(x) - f_n(x)| \leq \epsilon.$$

Da die rechte Seite auch von $x$ unabhängig ist können wir das Supremum bilden und erhalten

$$\sup_{x\in M} |f(x) - f_n(x)| = \|f - f_n\|_\infty \leq \epsilon.$$

Damit haben wir die Konvergenz von $(f_n)$ gegen $f$ bezüglich der Norm gezeigt. $\square$

**Definition 1.75.** Sei $(X, d)$ ein metrischer Raum und $\mathbf{B}(X)$ der Raum der beschränkten Funktionen auf $X$. Der Teilraum der stetigen, beschränkten Funktionen wird mit $\mathbf{C}_b(X)$ bezeichnet.

*Bemerkung 1.76.* Falls $X$ eine beschränkte, abgeschlossene Teilmenge von $\mathbb{R}^n$ ist, dann ist nach Satz 1.64 jede stetige Funktion beschränkt. In diesem Fall enthält der Raum $\mathbf{C}_b(X)$ alle stetigen Funktionen auf $X$ und wir schreiben $\mathbf{C}(X)$.

Im Gegensatz zur punktweisen Konvergenz erhält die gleichmäßige Konvergenz viele gute Eigenschaften von Funktionen.

**Satz 1.77.** *Sei $(X, d)$ ein metrischer Raum. Der Unterraum $\mathbf{C}_b(X)$ ist abgeschlossen in $\mathbf{B}(X)$. Insbesondere ist $\mathbf{C}_b(X)$ vollständig.*

*Beweis.* Wir wollen das Folgenkriterium für Abgeschlossenheit Satz 1.33 verwenden. Sei also $(f_n) \subset \mathbf{C}_b(X)$ eine Folge stetiger Funktionen, die (gleichmäßig) gegen $f \in \mathbf{B}(X)$ konvergiert. Sei $x \in X$ und $\epsilon > 0$. Da $(f_n)$ gleichmäßig gegen $f$ konvergiert, existiert $N \in \mathbb{N}$ so, dass $\|f - f_n\|_\infty < \frac{\epsilon}{3}$ für jedes $n > N$. Wähle jetzt $n > N$ fest. Da $f_n$ stetig ist, existiert $\delta > 0$ so, dass $|f_n(x) - f_n(y)| < \frac{\epsilon}{3}$ falls $d(x, y) < \delta$. Dann gilt für jedes $y \in B_\delta(x)$

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &\leq 2\|f - f_n\|_\infty + |f_n(x) - f_n(y)| < \epsilon. \end{aligned}$$

Damit ist $f$ stetig in $x$ und da $x \in X$ beliebig war, ist $f$ stetig.

Da die Folge $(f_n)$ beliebig war, ist damit auch die Abgeschlossenheit von $\mathbf{C}_b(X)$ gezeigt. Die Vollständigkeit dieses Raumes folgt dann aus Satz 1.38. $\square$

**Definition 1.78.** Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, $x_0 \in \mathbb{K}$ und $(a_n)_{n\in\mathbb{N}_0} \subset \mathbb{K}$. Dann heißt $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine (formale) Potenzreihe um den Punkt $x_0$ über dem Körper $\mathbb{K}$.

Die Zahl

$$\sup\left\{ |x-x_0| \;\middle|\; \sum_{n=0}^{\infty} a_n(x-x_0)^n \text{ konvergiert} \right\}$$

in $[0,\infty) \cup \{\infty\}$ heißt Konvergenzradius der Potenzreihe.

**Satz 1.79.** *Sei $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine Potenzreihe über dem Körper $\mathbb{K}$ und $r \in [0,\infty) \cup \{\infty\}$ ihr Konvergenzradius. Für $|x-x_0| < r$ konvergiert die Reihe absolut, für $|x-x_0| > r$ divergiert sie.*

*Es gilt die Formel*

$$\frac{1}{r} = \limsup_{n\to\infty} \sqrt[n]{|a_n|}. \tag{1.4}$$

*Wobei wir $\frac{1}{\infty} = 0$ und $\frac{1}{0} = \infty$ setzen.*

*Beweis.* Um die Notation einfach zu halten, führen wir den Beweis nur für $x_0 = 0$. Dass die Reihe für $|x| > r$ divergiert folgt aus der Definition des Konvergenzradius. Wir zeigen zunächst die absolute Konvergenz der Reihe falls $|x| < r$. Falls $r = 0$ ist, so ist nichts zu zeigen. Sei also $r > 0$ und $|x| < r$. Nach der Definition des Konvergenzradius existert $y$, so dass $|x| < |y| \le r$ und $\sum_{n=0}^{\infty} a_n y^n$ konvergiert. Insbesondere muss also $(a_n y^n)$ eine Nullfolge sein und so gilt $|a_n|\,|y|^n \le C$ für ein $C \in [0,\infty)$ und alle $n \in \mathbb{N}_0$. Damit gilt

$$|a_n|\,|y|^n \left(\frac{|x|}{|y|}\right)^n \le C \left(\frac{|x|}{|y|}\right)^n$$

die Reihe

$$\sum_{n=0}^{\infty} |a_n|\,|y|^n \left(\frac{|x|}{|y|}\right)^n$$

wird also durch eine konvergente geometrische Reihe majorisiert und ist somit selbst konvergent. Damit ist die Reihe

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n y^n \left(\frac{x}{y}\right)^n \tag{1.5}$$

absolut konvergent.

Die Formel für den Konvergenzradius ergibt sich aus dem Wurzelkriterium. Setze

$$b := \limsup_{n\to\infty} \sqrt[n]{|a_n|}.$$

Sei zunächst $b \in (0,\infty)$. Dann gilt

$$\limsup_{n\to\infty} \sqrt[n]{|a_n x^n|} = |x| \limsup_{n\to\infty} \sqrt[n]{|a_n|} = |x|\, b.$$

Damit ist die Reihe konvergent falls $|x| < \frac{1}{b}$ – also gilt $r \ge \frac{1}{b}$ – und divergent falls $|x| > \frac{1}{b}$ also ist $r \le \frac{1}{b}$. Schließlich sieht man leicht, dass die Potenzreihe lediglich in $x = 0$ konvergiert falls $b = \infty$ und für jedes $x \in \mathbb{K}$ konvergiert falls $b = 0$ ist. $\square$

*Bemerkung 1.80.* (i) Aus dem obigen Beweis ist ersichtlich, dass

$$r = \sup \left\{ a \in [0,\infty) \,\middle|\, \sum_{n=0}^{\infty} |a_n|\, a^n \text{ konvergiert} \right\}.$$

(ii) Falls der folgende Grenzwert existiert (also insbesondere $a_n \neq 0$ ist ab einem bestimmten Index $n$), so gilt auch

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

*Beispiel 1.81.* Die Potenzreihe

$$\sum_{n=0}^{\infty} x^n$$

ist konvergent auf $(-1,1)$ (beziehungsweise auf $B_1(0) \subset \mathbb{C}$) und konvergiert dort gegen die Funktion $f(x) = \frac{1}{1-x}$.

Man beachte dabei, dass der Grenzwert der Reihe nur auf dem Konvergenzkreis mit $f$ übereinstimmt (außerhalb dieses Kreises existiert der Grenzwert nicht) obwohl der maximale Definitionsbereich von $f$ darüber hinaus geht.

**Satz 1.82.** *Sei $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine Potenzreihe über dem Körper $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) mit Konvergenzradius $r > 0$. Sei $0 < a < r$. Dann konvergiert die Potenzreihe auf $K_a(x_0)$ gleichmäßig.*

*Insbesondere ist die Funktion*

$$\begin{aligned} f : B_r(x_0) &\to \mathbb{K} \\ x &\mapsto \sum_{n=0}^{\infty} a_n(x-x_0)^n \end{aligned}$$

*stetig.*

*Beweis.* Wir betrachten wiederum nur den Fall $x_0 = 0$. Sei $0 < a < r$. Wir wissen bereits, dass die Reihe $\sum_{n=0}^{\infty} a_n x^n$ auf $K_a(0)$ (sogar auf $B_r(0)$) punktweise konvergiert, wir wollen jedoch die Konvergenz bezüglich der Supremumnorm $\|\cdot\|_\infty$ zeigen. Setze

$$\begin{aligned} f_n : K_a(0) &\to \mathbb{K} \\ x &\mapsto x^n \end{aligned}$$

für alle $n \in \mathbb{N}_0$. Dann gilt

$$\|f_n\|_\infty = \sup_{x\in K_a(0)} |x^n| = a^n$$

und wir erhalten

$$\sum_{n=0}^{\infty} \|a_n f_n\|_\infty = \sum_{n=0}^{\infty} |a_n|\, a^n < \infty \tag{1.6}$$

da $a < r$. Die Reihe $\sum_{n=0}^{\infty} a_n f_n$ ist also absolut konvergent in $\mathbf{B}(K_a(0))$. Da dieser Raum vollständig ist, folgt daraus die Konvergenz der Reihe in $\mathbf{B}(K_a(0))$ (Satz 1.21) also die gleichmäßige Konvergenz.

Da die gleichmäßige Konvergenz die punktweise Konvergenz impliziert (Bemerkung 1.73), muss der gleichmäßige Grenzwert mit dem punktweisen übereinstimmen also mit der Funktion $x \mapsto \sum_{n=0}^{\infty} a_n x^n$. Da die Partialsummen $\sum_{n=0}^{N} a_n x^n$ für jedes $N$ Polynome und somit stetig sind, ist nach Satz 1.77 auch der Grenzwert stetig (auf $K_a(0)$).

Wir wollen nun noch die Stetigkeit auf $B_r(0)$ zeigen. Wir setzen

$$\begin{aligned} s : B_r(0) &\to \mathbb{K} \\ x &\mapsto \sum_{n=0}^{\infty} a_n x^n \end{aligned}$$

Sei nun $y \in B_r(0)$ und wähle $a$ so, dass $|y| < a < r$. Dann ist nach dem soeben gezeigten $s|_{K_a(0)}$ stetig, also insbesondere stetig in $y \in K_a(0)$. Da $K_a(0)$ eine Umgebung von $y$ enthält, stimmen $s$ und $s|_{K_a(0)}$ auf dieser Umgebung überein und da Stetigkeit eine lokale Eigenschaft ist, folgt daraus die Stetigkeit von $s$ in $y$. Da das für jedes $y \in B_r(0)$ gilt, ist $s$ stetig. □

*Beispiel 1.83.* (i) Wir definieren die Exponentialfunktion

$$e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Der Konvergenzradius ist

$$r = \lim_{n\to\infty} \frac{\frac{1}{n!}}{\frac{1}{(n+1)!}} = \lim_{n\to\infty} n + 1 = \infty,$$

die Reihe also auf ganz $\mathbb{R}$ (oder ganz $\mathbb{C}$) konvergent. Nach obigem Satz ist die Exponentialfunktion stetig.

(ii) Ebenfalls definieren wir für $x \in \mathbb{R}$

$$\begin{aligned} \sin(x) &:= \Im e^{\mathrm{i}x} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \\ \cos(x) &:= \Re e^{\mathrm{i}x} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}. \end{aligned}$$

Beide Reihen konvergieren auf ganz $\mathbb{R}$ und sind daher stetige Funktionen.

*Bemerkung 1.84.* Die Ergebnisse dieses Abschnitts können – inklusive der Beweise – ohne Weiteres auf $\mathbb{R}^n$-wertige Funktionen oder sogar auf Funktionen mit Werten in einem beliebigen Banachraum verallgemeinert werden.

# 2 Differenzierbarkeit

## 2.1 Differenzierbarkeit von Funktionen einer reellen Variablen

Wie bisher sei $\mathbb{K}$ einer der Zahlkörper $\mathbb{R}$ oder $\mathbb{C}$.

**Definition 2.1.** Sei $U \subset \mathbb{R}$ offen, $f : U \to \mathbb{K}$ eine Funktion und $x \in U$. Die Funktion $f$ heißt differenzierbar in $x$, wenn der Grenzwert

$$f'(x) := \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

existiert. Dann heißt $f'(x)$ die Ableitung der Funktion $f$ an der Stelle $x$.

Die Funktion $f$ heißt differenzierbar, falls sie in jedem Punkt des Definitionsbereiches $U$ differenzierbar ist.

*Bemerkung 2.2.* (i) Genauso können wir Differenzierbarkeit und Ableitung für Kurven im $\mathbb{R}^n$ (oder sogar in einem beliebigen Banachraum $V$), also für Abbildungen $f : U \to \mathbb{R}^n$ ($f : U \to V$) definieren. Für $\mathbb{R}^n$-wertige Abbildungen entspricht die Differenzierbarkeit genau der Differenzierbarkeit der einzelnen Komponentenfunktionen.

(ii) Ist eine Funktion in $x$ differenzierbar, so gilt

$$\lim_{h \to 0} f(x+h) - f(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} h = 0$$

also ist $f$ in $x$ stetig.

(iii) Differenzierbarkeit ist eine lokale Eigenschaft, hängt also nur vom Verhalten der Funktion in einer beliebig kleinen Umgebung ab.

(iv) Die Funktion $f$ ist differenzierbar in $x$, genau dann wenn

$$f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x}$$

existiert.

**Satz 2.3.** *Seien $U \subset \mathbb{R}$ offen, $x \in U$ und $f, g : U \to \mathbb{K}$ in $x$ differenzierbare Funktionen. Dann sind auch $f + g$ und $f \cdot g$ differenzierbar in $x$ und es gilt*

$$(f+g)'(x) = f'(x) + g'(x),$$
$$(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x).$$

*Ist $f(x) \neq 0$, dann ist $\frac{1}{f}$ auf einer (offenen) Umgebung von $x$ definiert und es gilt*

$$\left(\frac{1}{f}\right)'(x) = -\frac{f'(x)}{f(x)^2}.$$

*Beweis.* Übung □

**Satz 2.4.** *Seien $U, V \subset \mathbb{R}$ offen, $g : U \to V$ differenzierbar in $x$ und $f : V \to \mathbb{K}$ differenzierbar in $g(x)$, dann ist $f \circ g$ differenzierbar in $x$ und es gilt*

$$(f \circ g)'(x) = f'(g(x))g'(x).$$

*Beweis.* Erfolgt weiter unten allgemeiner für Funktionen mehrerer Variablen. □

**Definition 2.5.** Sei $(X, d)$ ein metrischer Raum, $f : X \to \mathbb{R}$ eine Funktion und $x \in X$. Die Funktion $f$ hat in $x$ ein lokales Maximum, wenn es $\epsilon > 0$ gibt, so dass

$$f(y) \leq f(x) \text{ für alle } y \in B_\epsilon(x).$$

Die Funktion $f$ hat in $x$ ein lokales Minimum, wenn $-f$ in $x$ ein lokales Maximum hat. Ein lokales Extremum ist entweder ein lokales Maximum oder ein lokales Minimum.

**Satz 2.6.** *Sei $U \subset \mathbb{R}$ offen, $x \in U$ und $f : U \to \mathbb{R}$ differenzierbar in $x$. Falls $f$ in $x$ ein lokales Extremum hat, dann ist $f'(x) = 0$.*

*Beweis.* Die Funktion $f$ habe oBdAin $x$ ein lokales Maximum (andernfalls betrachte $-f$). Dann gibt es $\epsilon > 0$, so dass $f(y) \leq f(x)$ für alle $y \in B_\epsilon(x)$. Dann gilt

$$f(x+h) - f(x) \leq 0$$

für alle $h \in B_\epsilon(0)$. Damit gilt für $h_n = \frac{1}{n}$

$$\lim_{n\to\infty} \frac{f(x+h_n) - f(x)}{h_n} \leq 0 \text{ und}$$
$$\lim_{n\to\infty} \frac{f(x-h_n) - f(x)}{-h_n} \geq 0.$$

Da $f$ in $x$ differenzierbar ist, müssen beide Grenzwerte mit

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = f'(x)$$

übereinstimmen, es muss also $f'(x) = 0$ gelten. □

*Bemerkung 2.7.* Das Verschwinden der Ableitung ist für die Existenz eines lokalen Extremums nur eine notwendige, keine hinreichende Bedingung.

**Theorem 2.8 (Mittelwertsatz der Differentialrechnung).** *Seien $a, b \in \mathbb{R}$, $a < b$ und $f : [a, b] \to \mathbb{R}$ eine stetige Funktion die auf $(a, b)$ differenzierbar ist. Dann existiert $\xi \in (a, b)$, so dass*

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

*Beweis.* Wir betrachten zunächst den Fall $f(a) = f(b)$. Dann ist zu zeigen, dass $\xi \in (a, b)$ existiert, so dass $f'(\xi) = 0$. Da $[a, b]$ beschränkt und abgeschlossen ist, nimmt die Funktion $f$ dort sowohl ihr Maximum als auch ihr Minimum an (Korollar 1.66). Falls sowohl das Maximum, als auch das Minimum in $a$ und $b$ angenommen werden, dann ist $f$ konstant und $f'(\xi) = 0$ für jedes $\xi \in (a, b)$. Andernfalls bezeichne $\xi$ eine der Extremstellen in $(a, b)$. Da ein globales Extremum insbesondere ein lokales Extremum ist, gilt nach Satz 2.6 $f'(\xi) = 0$.

Wir betrachten nun allgemeine $f$ und definieren

$$\begin{aligned} g : [a, b] &\to \mathbb{R} \\ x &\mapsto f(x) - \frac{f(b) - f(a)}{b - a}(x - a). \end{aligned}$$

Dann ist $g$ ebenfalls stetig auf $[a, b]$ und differenzierbar auf $(a, b)$ und es gilt $g(a) = f(a) = g(b)$. Dann existiert nach dem ersten Teil des Beweises $\xi \in (a, b)$ mit

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

**Korollar 2.9.** *Seien $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$, $a < b$, $f : (a, b) \to \mathbb{R}$ differenzierbar und es gelte $f'(x) = 0$ für alle $x \in (a, b)$. Dann ist $f$ konstant.*

*Beweis.* Seien $x, y \in (a, b)$ und $y > x$. Dann gibt es nach dem Mittelwertsatz $\xi \in (x, y)$, so dass

$$0 = f'(\xi) = \frac{f(y) - f(x)}{y - x},$$

also gilt $f(x) = f(y)$. Insbesondere gilt also für festes $x \in (a, b)$, dass $f(x) = f(y)$ für alle $y \in (a, b)$, also ist $f$ konstant. □

**Korollar 2.10.** *Seien $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$, $a < b$, $f : (a, b) \to \mathbb{R}$ differenzierbar. Die Funktion $f$ ist monoton wachsend, genau dann wenn $f'(x) \geq 0$ für alle $x \in (a, b)$. Wenn $f'(x) > 0$ ist, so ist $f$ streng monoton wachsend.*

*Beweis.* Es gelte zunächst $f'(x) \geq 0$ ($f'(x) > 0$) für alle $x \in (a, b)$. Seien wiederum $x, y \in (a, b)$, $y > x$. Nach dem Mittelwertsatz gibt es $\xi \in (x, y)$, so dass

$$f'(\xi) = \frac{f(y) - f(x)}{y - x}.$$

Damit ist $f(y) \geq f(x)$ ($f(y) > f(x)$). Da $x$ und $y$ beliebig waren, ist $f$ (streng) monoton wachsend.

Sei andererseits $f$ monoton wachsend, dann gilt für jedes $x$

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h} = \lim_{n \to \infty} \frac{f\left(x + \frac{1}{n}\right) - f(x)}{\frac{1}{n}} \geq 0.$$ □

*Bemerkung 2.11.* (i) In Korollar 2.9 ist entscheidend, dass die Funktion auf einem Intervall definiert ist. Die Ableitung der Funktion

$$\begin{aligned} f : \mathbb{R}\setminus\{0\} &\to \mathbb{R} \\ x &\mapsto \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases} \end{aligned}$$

verschwindet zwar auf ihrem Definitionsbereich, aber die Funktion ist offensichtlich nicht konstant.

(ii) Die Umkehrung der zweiten Aussage in Korollar 2.10 gilt nicht. Die Funktion $x \mapsto x^3$ ist streng monoton wachsend aber es gilt $f'(0) = 0$.

**Korollar 2.12 (Erweiterter Mittelwertsatz).** *Seien $a, b \in \mathbb{R}$, $a < b$, $f, g : [a, b] \to \mathbb{R}$ stetige Funktionen die auf $(a, b)$ differenzierbar sind. Dann existiert $\xi \in (a, b)$, so dass*

$$f'(\xi)(g(b) - g(a)) = g'(\xi)(f(b) - f(a)).$$

*Beweis.* Falls $g(b) = g(a)$ ist, so existiert nach dem Mittelwertsatz $\xi$ mit $g'(\xi) = 0$ und damit ist die Aussage gezeigt.

Sei also $g(b) \neq g(a)$. Wir wenden den Mittelwertsatz auf die Hilfsfunktion

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

an. Es gilt $h(b) = f(a) = h(a)$, also existiert $\xi \in (a, b)$, so dass

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

Multiplikation mit $g(b) - g(a)$ ergibt die gewünschte Aussage. □

**Satz 2.13 (Regel von de l'Hospital).** *Sei $U \subset \mathbb{R}$ offen, $x \in U$ und $f, g : U \to \mathbb{R}$ differenzierbare Funktionen. Falls $f(x) = g(x) = 0$, $g' \neq 0$ in einer punktierten Umgebung von $x$ und*

$$\lim_{y \to x} \frac{f'(y)}{g'(y)}$$

*existiert, dann gilt*

$$\lim_{y \to x} \frac{f(y)}{g(y)} = \lim_{y \to x} \frac{f'(y)}{g'(y)}.$$

*Beweis.* Wir zeigen zunächst, dass $x$ eine isolierte Nullstelle von $g$ ist. Angenommen $x$ ist Häufungspunkt von Nullstellen von $g$, das heißt es existieren Nullstellen $x_n \neq x$ für alle $n \in \mathbb{N}$, so dass $x_n \to x$. Wir können oBdAannehmen, dass $(x_n)$ streng monoton ist (sonst wähle eine Teilfolge aus). Sei $(x_n)$ streng monoton wachsend. Nach dem Mittelwertsatz existiert für jedes $n \in \mathbb{N}$ ein $\xi_n \in (x_n, x_{n+1})$ mit $g'(\xi_n) = 0$. Da $\xi_n \to x$ konvergieren, steht das im Widerspruch zur Voraussetzung $g' \neq 0$ auf einer punktierten Umgebung von $x$. Für den Fall streng monoton fallender Folge $(x_n)$ liefert ein analoges Argument ebenfalls einen Widerspruch.

Sei nun $\dot{V}$ eine Umgebung von $x$, so dass $g \neq 0$ und $g' \neq 0$ auf $\dot{V}$. Wähle eine beliebige Folge $x_n \in V$, so dass $x_n \to x$. Für jedes $n \in \mathbb{N}$ existiert nach dem erweiterten Mittelwertsatz $\xi_n \in (\min\{x, x_n\}, \max\{x, x_n\})$, so dass

$$\frac{f'(\xi_n)}{g'(\xi_n)} = \frac{f(x) - f(x_n)}{g(x) - g(x_n)} = \frac{f(x_n)}{g(x_n)}.$$

Es gilt $\lim_{n\to\infty} \xi_n \to x$. Nach Voraussetzung existiert

$$\lim_{y\to x} \frac{f'(y)}{g'(y)} = \lim_{n\to\infty} \frac{f'(\xi_n)}{g'(\xi_n)} = \lim_{n\to\infty} \frac{f(x_n)}{g(x_n)}.$$

Da die Folge $(x_n)$ beliebig war, folgt mit dem Folgenkriterium für Konvergenz von Funktionen die Existenz von

$$\lim_{y\to x} \frac{f(y)}{g(y)} = \lim_{y\to x} \frac{f'(y)}{g'(y)}.$$ □

*Beispiel 2.14.* Wir wollen mittels der Regel von l'Hospital den Grenzwert

$$\lim_{x\to 0} \frac{1 - \cos x}{x^2}$$

bestimmen. Es sind also $f(x) = 1 - \cos x$, $f'(x) = \sin x$ und $f''(x) = \cos x$ sowie $g(x) = x^2$, $g'(x) = 2x$ und $g''(x) = 2$. Da $f'(0) = g'(0) = 0$ und wegen der Stetigkeit des cos

$$\lim_{x\to 0} \frac{f''(x)}{g''(x)} = \frac{1}{2}$$

ist, existiert nach l'Hospital auch

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = \frac{1}{2}.$$

Da auch $f(x) = g(x) = 0$ ist und der obige Grenzwert existiert, gilt dann auch

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = \frac{1}{2}.$$

**Definition 2.15.** Sei $U \subset \mathbb{R}$ offen und $f : U \to \mathbb{R}$ eine Funktion. Die Funktion heißt $n+1$ mal differenzierbar wenn die $n$-te Ableitung $f^{(n)}$ differenzierbar ist und die $n+1$-te Ableitung ist die Funktion ${f^{(n)}}'$. Die Funktion heißt beliebig differenzierbar oder glatt falls sie $n$ mal differenzierbar ist für jedes $n \in \mathbb{N}$.

Die Funktion heißt $n$ mal stetig differenzierbar, falls $f^{(n)}$ stetig ist.

**Definition 2.16.** Sei $U \subset \mathbb{R}$ offen und $f : U \to \mathbb{R}$ eine in $x$ differenzierbare Funktion. Dann heißt

$$T_n(y) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k$$

das $n$-te Taylor-Polynom um den Entwicklungspunkt $x$.

**Theorem 2.17 (Satz von Taylor).** *Sei $(a,b) \subset \mathbb{R}$ ein offenes Intervall und $f : (a,b) \to \mathbb{R}$ eine $n+1$ mal differenzierbare Funktion. Seien $x, y \in (a,b)$ zwei verschiedene Punkte und $I$ das abgeschlossene, $\mathring{I}$ das offene von $x$ und $y$ aufgespannte Intervall. Sei $h : I \to \mathbb{R}$ eine stetige auf $\mathring{I}$ differenzierbare Funktion, so dass $h' \neq 0$ auf $\mathring{I}$. Dann existiert $\xi \in \mathring{I}$, so dass*

$$f(y) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k + \frac{f^{(n+1)}(\xi)}{n!}(y-\xi)^n \frac{h(y)-h(x)}{h'(\xi)}. \tag{2.1}$$

*Beweis.* Wir wenden den verallgemeinerten Mittelwertsatz auf

$$\begin{aligned} g : (a,b) &\to \mathbb{R} \\ t &\mapsto \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(y-t)^k \end{aligned}$$

und $h$ an, es existiert also $\xi \in \mathring{I}$, so dass

$$\frac{g'(\xi)}{h'(\xi)}(h(y)-h(x)) = g(y) - g(x) = f(y) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k.$$

Wir müssen nun noch $g'$ ausrechnen:

$$\begin{aligned} g'(t) &= \sum_{k=0}^{n} \left( \frac{f^{(k+1)}(t)}{k!}(y-t)^k - \frac{f^{(k)}(t)}{k!} k (y-t)^{k-1} \right) \\ &= \frac{f^{(n+1)}(t)}{n!}(y-t)^n \end{aligned}$$

wobei die letzte Umformung aus der Tatsache folgt, dass es sich um eine Teleskopsumme handelt. Einsetzen von $\xi$ liefert dann die gesuchte Gleichung. □

*Bemerkung 2.18.* (i) Je nach Wahl der Funktion $h$ erhält man nun verschiedene Formen des Restgliedes

$$R_{n+1}(y,x) := f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k.$$

(ii) Für $h(t) = (y-t)^{n+1}$ also $h'(t) = -(n+1)(y-t)^n$ ergibt sich die Lagrange-Form des Restgliedes:

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

**Korollar 2.19.** *Sei $(a,b) \subset \mathbb{R}$ ein intervall und $f : (a,b) \to \mathbb{R}$ mindestens $n$ mal stetig differenzierbar. Sei $x \in (a,b)$ und es gelte $0 = f'(x) = f''(x) = \cdots = f^{(n-1)}(x)$ und $f^{(n)}(x) \neq 0$. Ist $n$ ungerade so liegt bei $x$ kein lokales Extremum vor, ist $n$ gerade, so hat $f$ bei $x$ ein lokales Minimum falls $f^{(n)}(x) > 0$ und ein lokales Maximum falls $f^{(n)}(x) < 0$.*

*Beweis.* Wir setzten oBdAvoraus, dass $f^{(n)}(x) > 0$ ist (sonst betrachte $-f$). Wir verwenden die Taylor-Formel mit Lagrange-Restglied. Nach Voraussetzung verschwinden in der Entwicklung (2.1) die Terme 1 bis $n-1$. Da nach Voraussetzung $f^{(n)}(x)$ stetig ist, existiert $\epsilon > 0$, so dass $f^{(n)}(x) > 0$ auf ganz $(x-\epsilon, x+\epsilon)$. Dann existiert nach der Taylor-Formel für jedes $y \in (x-\epsilon, x+\epsilon)$ ein $\xi_y \in (x-\epsilon, x+\epsilon)$, so dass

$$f(y) = f(x) + \frac{f^{(n)}(\xi_y)}{n!}(y-x)^n.$$

Falls $n$ ungerade ist, so ist $f(x-\delta) < f(x) < f(x+\delta)$ für jedes $\delta > 0$. Es liegt also kein lokales Extremum vor. Ist $n$ gerade, so ist der zweite Term stets nichtnegativ, es gilt also $f(y) \geq f(x)$ für alle $y \in (x-\epsilon, x+\epsilon)$, an $x$ liegt also ein lokales Minimum vor. □

**Satz 2.20.** *Sei $U \subset \mathbb{R}$ offen und $x \in U$. Eine Funktion $f : U \to \mathbb{K}$ ist differenzierbar in $x$, genau dann wenn $a \in \mathbb{K}$ und $\varphi : U \to \mathbb{K}$ existieren, so dass*

$$f(y) = f(x) + a(y-x) + \varphi(y)\,|y-x| \tag{2.2}$$

*für alle $y \in U$ und $\lim_{y\to x} \varphi(y) = 0$.*

*Beweis.* Seien zunächst $a \in \mathbb{K}$ und $\varphi : U \to \mathbb{K}$ so, dass $\lim_{y\to x} \varphi(y) = 0$ und Gleichung (2.2) gilt. Dann gilt

$$\frac{f(y) - f(x)}{y-x} = a + \varphi(y)\frac{|y-x|}{y-x}$$

und die rechte Seite dieser Gleichung konvergiert für $y \to x$ nach Voraussetzung gegen $a$. Damit ist $f$ in $x$ differenzierbar und $f'(x) = a$.

Sei andererseits $f$ in $x$ differenzierbar und definiere

$$\varphi(y) := \begin{cases} \left(\frac{f(y)-f(x)}{y-x} - f'(x)\right)\frac{y-x}{|y-x|} & y \neq x \\ 0 & y = x \end{cases}.$$

Dann gilt Gleichung (2.2) (für $a = f'(x)$) und $\lim_{y\to x} \varphi(y) = 0$. □

## 2.2 Differenzierbarkeit von Funktionen mehrerer Veränderlicher

### Wiederholung lineare Abbildungen und Matrizen

Im Folgenden seien $V, W$ stets Vektorräume über einem Körper $\mathbb{K}$.

**Definition 2.21.** Eine Abbildung $\varphi : V \to W$ heißt linear, wenn

$$\varphi(v + \lambda\tilde{v}) = \varphi(v) + \lambda\varphi(\tilde{v})$$

für alle $v, \tilde{v} \in V$ und alle $\lambda \in \mathbb{K}$.

*Bemerkung 2.22.* (i) Die Menge $\mathcal{L}(V|W)$ aller linearen Abbildungen zwischen festen Räumen $V$ und $W$ ist selbst ein Vektorraum über $\mathbb{K}$.

(ii) Seien $e_1, \ldots, e_k$ und $f_1, \ldots, f_l$ Basen für $V$ beziehungsweise $W$, dann ist durch

$$\varphi(e_j) = \sum_{i=1}^{l} a_{ij} f_i \tag{2.3}$$

für alle $j \in \{1, \ldots, k\}$ eindeutig eine Matrix $A_\varphi = (a_{ij})$ definiert, die sogenannte Darstellungs- oder Abbildungsmatrix der linearen Abbildung $\varphi$ bezüglich der gegebenen Basen.

Umgekehrt wird durch jede $l \times k$-Matrix $A = (a_{ij})$ mittels der Gleichungen (2.3) eindeutig eine lineare Abbildung $\varphi_A$ definiert.

(iii) Für fest gewählte Basen von $V$ und $W$ hängt die Darstellungsmatrix, also die Abbildung

$$\begin{aligned} \mathcal{L}(V|W) &\to \mathbb{R}^{l\times k} \\ \varphi &\mapsto A_\varphi \end{aligned}$$

linear von $\varphi$ ab, es gilt also für lineare Abbildungen $\varphi, \psi : V \to W$ und $\lambda \in \mathbb{K}$

$$A_{\varphi+\lambda\psi} = A_\varphi + \lambda A_\psi.$$

(iv) Seien $v = v_1e_1 + \cdots v_ke_k \in V$ und $w = \varphi(v) = w_1f_1 + \cdots w_lf_l$, dann gilt

$$\begin{pmatrix} w_1 \\ \vdots \\ w_l \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{l1} & \cdots & a_{lk} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix}$$

die Anwendung von $\varphi$ auf $v$ entspricht also der Matrizenmultiplikation der Matrixdarstellungen von $\varphi$ und $v$.

(v) Falls $U$ ein weiterer Vektorraum mit fest gewählter Basis ist und $\varphi : V \to W$ und $\psi : W \to U$ lineare Abbildungen sind, so gilt für die zugehörigen Darstellungsmatrizen

$$A_{\psi\circ\varphi} = A_\psi A_\varphi,$$

die Hintereinanderausführung der Abbildungen entspricht also der Matrixmultiplikation der Darstellungsmatrizen.

(vi) Falls $V = \mathbb{R}^k$ und $W = \mathbb{R}^l$ gilt, so gibt es auf jedem der Räume die Standardbasis. Wird nichts anderes gesagt, so beziehen sich alle Darstellungsmatrizen auf diese Standardbasen und wir werden die lineare Abbildung $\varphi$ mit ihrer Abbildungsmatrix $A_\varphi$ identifizieren.

(vii) Auf $\mathbb{R}^n$ existiert ebenfalls das Standardskalarprodukt $\langle \cdot \,|\, \cdot \rangle$. Für

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \qquad \tilde{v} = \begin{pmatrix} \tilde{v}_1 \\ \vdots \\ \tilde{v}_n \end{pmatrix}$$

erhalten wir

$$\langle v|\,\tilde{v}\rangle = \sum_{i=1}^{n} v_i\tilde{v}_i = (v_1, \ldots, v_n) \begin{pmatrix} \tilde{v}_1 \\ \vdots \\ \tilde{v}_n \end{pmatrix}.$$

**Definition 2.23.** Sei $U \subset \mathbb{R}^n$ offen, $x = (x_1, \cdots, x_n) \in U$ und $f : U \to \mathbb{R}^m$. Die Abbildung $f$ heißt partiell differenzierbar in $x$ bezüglich $x_i$, wenn die Funktionen $t \mapsto f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots x_n)$ differenzierbar ist, das heißt wenn der Grenzwert

$$\partial_i f(x) := \lim_{h\to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots x_n) - f(x_1, \ldots, x_n)}{h}$$

existiert. Der Grenzwert $\partial_i f(x) \in \mathbb{R}^m$ heißt die partielle Ableitung von $f$ an der Stelle $x$ bezüglich $x_i$ und wird auch als $\frac{\partial f}{\partial x_i}(x)$ geschrieben.

Die Abbildung $f$ heißt partiell differenzierbar in $x$, wenn sie partiell differenzierbar bezüglich $x_i$ für jedes $i \in \{1, \ldots n\}$ ist.

Die Existenz nur der partiellen Ableitungen ist in vielen Fällen eine zu schwache Bedingung wie das folgende Beispiel zeigt.

*Beispiel 2.24.* Die Funktion

$$\begin{aligned} f : \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto \begin{cases} 1 & x = 0 \text{ oder } y = 0 \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

ist im Punkt $(0, 0)$ partiell differenzierbar, aber nicht stetig.

Über lediglich partiell differenzierbare Funktionen können wir kaum Aussagen treffen, es bedarf also eines stärkeren Differenzierbarkeitsbegriffes. Dazu bietet sich die Verallgemeinerung der in Satz 2.20 eingeführten Bedingung an.

**Definition 2.25.** Sei $U \subset \mathbb{R}^n$ offen, $x \in U$ und $f : U \to \mathbb{R}^m$. Die Abbildung $f$ heißt in $x$ differenzierbar oder total differenzierbar, falls eine lineare Abbildung $A \in \mathcal{L}(\mathbb{R}^n|\mathbb{R}^m)$ und $\varphi : U \to \mathbb{R}^m$ existieren, so dass

$$f(y) = f(x) + A(y - x) + \varphi(y)\,\|y - x\| \tag{2.4}$$

für alle $y \in U$ und $\lim_{y\to x} \varphi(y) = 0$.

Genauso können wir Differenzierbarkeit von Abbildungen zwischen beliebigen normierten Vektorräumen $V$ und $W$ definieren.

**Satz 2.26.** *Sei $U \subset \mathbb{R}^n$ offen, $x \in U$ und $f : U \to \mathbb{R}^m$. Seien $f_1, \dots, f_m : U \to \mathbb{R}$ die Komponentenfunktionen von $f$, das heißt $f(x) = (f_1(x), \dots, f_m(x))$ für alle $x \in X$. Wenn die Abbildung $f$ total differenzierbar ist, so ist sie partiell differenzierbar und die Darstellungsmatrix der Abbildung $A$ aus Definition 2.25 ist gegeben durch $A_{ji} = \partial_i f_j(x)$.*

*Beweis.* Seien $A$ und $\varphi$ wie in Definition 2.25. Seien $e_1, \dots, e_n$ und $e'_1, \dots, e'_m$ die Standardbasen von $\mathbb{R}^n$ beziehungsweise $\mathbb{R}^m$. Wir setzen $y = x + he_i$ in (2.4) ein und erhalten

$$\begin{aligned} f(x + he_i) &= f(x) + hA(e_i) + \varphi(x + he_i)\,|h| \\ \frac{f(x + he_i) - f(x)}{h} &= A(e_i) + \varphi(x + he_i)\frac{|h|}{h}. \end{aligned}$$

Für $h \to 0$ konvergiert diese Gleichung gegen

$$\sum_{j=1}^{m} \partial_i f_j(x) e'_j = \partial_i f(x) = A(e_i) = \sum_{j=1}^{m} A_{ji} e'_j,$$

also existiert die partielle Ableitung von $f$ bezüglich $x_i$. Aus der Eindeutigkeit der Basisdarstellung erhält man durch Koeffizientenvergleich die Gleichung $\partial_i f_j(x) = A_{ji}$. □

Damit ist insbesondere gezeigt, dass $A$ in (2.4) eindeutig bestimmt ist, wir können also von dieser Matrix (oder vielmehr von der linearen Abbildung die sie darstellt) als der totalen Ableitung von $f$ sprechen.

**Definition 2.27.** Die Matrix $(\partial_i f_j(x))_{ji}$ heißt Jacobi-Matrix (auch Funktionalmatrix oder Ableitungsmatrix) von $f$ an der Stelle $x$. Wir werden sowohl diese Matrix, als auch die lineare Abbildung die sie repräsentiert, mit $\mathrm{D}f(x)$ bezeichnen auch dann, wenn die Funktion nicht total differenzierbar ist.

Für reellwertige $f$, also $f : \mathbb{R}^n \to \mathbb{R}$, hat die Jakobi-Matrix lediglich eine Spalte und wir bezeichnen sie auch als Gradient von $f$ und schreiben $\operatorname{grad} f := \mathrm{D}f$.

*Beispiel 2.28.* (i) Seien $V$ und $W$ (endlichdimensionale) Vektorräume und $A : V \to W$ eine lineare Abbildung. Dann ist $A$ total differenzierbar und es gilt $\mathrm{D}A(x) = A$ für alle $x \in V$. Um das zu sehen schreiben wir

$$A(y) = A(x) + A(y - x)$$

womit (2.4) für $\varphi \equiv 0$ erfüllt ist.

(ii) Betrachte die Abbildung

$$\begin{aligned} f : (0, \infty) \times (0, 2\pi) &\to \mathbb{R}^2 \\ (r, \varphi) &\mapsto (r \cos \varphi, r \sin \varphi). \end{aligned}$$

Dann ist $f$ partiell differenzierbar und es gilt

$$Df(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

*Bemerkung 2.29.* (i) Eine Abbildung $f$ ist also total differenzierbar in $x$ genau dann, wenn die Jacobi-Matrix $\mathrm{D}f(x)$ existiert und

$$\lim_{y \to x} \frac{f(x) - f(y) - \mathrm{D}f(x)(y - x)}{\|y - x\|} = 0.$$

(ii) Aus dieser Formulierung ist ersichtlich, dass eine Abbildung $f = (f_1, \ldots f_m)$, mit $f_1, \ldots, f_m : U \to \mathbb{R}$ total differenzierbar ist, genau dann wenn $f_i$ total differenzierbar ist für jedes $i \in \{1, \ldots m\}$.

(iii) Wie im eindimensionalen Fall folgt aus der totalen Differenzierbarkeit insbesondere die Stetigkeit von $f$.

(iv) Sowohl partielle, als auch totale Differenzierbarkeit sind lokale Eigenschaften.

(v) Um Verwirrung zu vermeiden: Wenn $f$ auf einer offenen Menge $U$ differenzierbar ist, dann ist die Abbildung $h \mapsto \mathrm{D}f(x)h$ linear für jedes $x \in U$. Die Abbildung $x \mapsto \mathrm{D}f(x)$ ist in den meisten Fällen nicht linear.

**Definition 2.30.** Seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ normierte Vektorräume. Dann definiert

$$\|\varphi\|_{\mathcal{L}(V|W)} := \sup \{ \|\varphi(v)\|_W | \, v \in V, \|v\|_V = 1\}$$

eine Norm auf $\mathcal{L}(V|W)$.

*Bemerkung 2.31.* Der Beweis der Normeigenschaften verläuft analog zu Aufgabe 1 vom Übungsblatt 0,5. Die obige Norm ist auf $\mathcal{L}(V|W)$ für uns die Standardnorm, wir werden also häufig nur $\|\varphi\|$ schreiben wenn durch das Argument klar ist von welcher Norm die Rede ist. Es gilt die Ungleichung

$$\|\varphi(x)\| \leq \|\varphi\| \, \|x\| \,.$$

Damit folgt insbesondere die Lipschitz-Stetigkeit (und damit Stetigkeit) linearer Abbildungen:

$$\|\varphi(y) - \varphi(x)\| = \|\varphi(y-x)\| \leq \|\varphi\| \, \|y-x\| \,.$$

**Satz 2.32.** *Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offen, $x \in U$, $g : U \to V$ total differenzierbar in $x$ und $f : V \to \mathbb{R}^k$ total differenzierbar in $g(x)$. Dann ist $f \circ g$ total differenzierbar in $x$ und es gilt*

$$\mathrm{D}(f \circ g)(x) = \mathrm{D}f(g(x))\mathrm{D}g(x).$$

*Beweis.* Differenzierbarkeit von $g$ im Punkt $x$ bedeutet, dass $\varphi_g : U \to \mathbb{R}^m$ existiert, so dass $\lim_{y\to x} \varphi_g(y) = 0$ und

$$g(y) - g(x) = \mathrm{D}g(x)(y-x) + \varphi_g(y) \, \|y-x\| \,.$$

Insbesondere gilt also

$$\|g(y) - g(x)\| \leq \|\mathrm{D}g(x)(y-x)\| + \|y-x\| \, \|\varphi_g(y)\| \leq (\|\mathrm{D}g(x)\| + \|\varphi_g(y)\|) \, \|y-x\| \,.$$

Da $\varphi_g(y)$ für $y \to x$ gegen 0 konvergiert, ist es insbesondere beschränkt in einer Umgebung von $x$. Es gibt also ein $K \in \mathbb{R}$, so dass in einer punktierten Umgebung von $x$

$$\frac{\|g(y) - g(x)\|}{\|y-x\|} \leq K \tag{2.5}$$

gilt.

Differenzierbarkeit von $f$ im Punkt $g(x)$ bedeutet, dass es ein $\varphi_f : V \to \mathbb{R}^k$ gibt, so dass $\lim_{z\to g(x)} \varphi_f(z) = 0$ und

$$f(z) = f(g(x)) + \mathrm{D}f(g(x))(z - g(x)) + \varphi_f(z) \, \|z - g(x)\| \,.$$

Wir setzen $z = g(y)$ und benutzen die obige Formel für $g(y) - g(x)$

$$\begin{aligned} f(g(y)) &= f(g(x)) + \mathrm{D}f(g(x))(g(y) - g(x)) + \varphi_f(g(y)) \, \|g(y) - g(x)\| \\ (f \circ g)(y) &= (f \circ g)(x) + \mathrm{D}f(g(x))\mathrm{D}g(x)(y-x) \\ &\quad + \left( \varphi_f(g(y)) \frac{\|g(y) - g(x)\|}{\|y-x\|} + \mathrm{D}f(g(x))\varphi_g(y) \right) \|y-x\| \,. \end{aligned}$$

Um den Beweis abzuschließen muss gezeigt werden, dass für $y \to x$ der Ausdruck in Klammern gegen 0 konvergiert. Da $g$ in $x$ differenzierbar ist, ist es insbesondere stetig und so gilt $\lim_{y\to x} g(y) = g(x)$. Dann gilt $\lim_{y\to x} \varphi_f(g(y)) = 0$ und wegen (2.5) konvergiert der erste Summand in der Klammer gegen 0. Nach Voraussetzung konvergiert $\varphi_g(y)$ genen 0 und da $D(f(g(x)))$ linear und somit stetig ist konvergiert auch der zweite Summmand in der Klammer gegen 0. $\square$

**Satz 2.33.** *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$. Wenn für jedes $x \in U$ und jedes $i \in \{1, \dots, n\}$ die partielle Ableitung $\partial_i f(x)$ existiert und stetig ist, dann ist $f$ für jedes $x$ total differenzierbar.*

**Lemma 2.34.** *Sei $x \in U$, $r > 0$ und $f : B_r(x) \to \mathbb{R}$ partiell differenzierbar. Sei $h = (h_1, \dots, h_n) \in B_r(0)$. Dann existieren $\xi_1, \dots, \xi_n \in K_{\|h\|}(x)$, so dass*

$$f(x+h) - f(x) = \partial_1 f(\xi_1)h_1 + \dots + \partial_n f(\xi_n)h_n.$$

*Beweis.* Sei $e_1, \dots, e_n$ die Standardbasis des $\mathbb{R}^n$ und definiere $x_0 = x$, $x_i = x_{i-1} + h_i e_i$. Dann ist $x_n = x + h$. Weiter gilt

$$\|x_i - x\| = \|h_1 e_1 + \dots + h_i e_i\| = \sqrt{|h_1|^2 + \dots + |h_i|^2} \leq \sqrt{|h_1|^2 + \dots + |h_n|^2} = \|h\|$$

die Punkte $x_i$ sind also alle in $K_{\|h\|}(x) \subset B_r(x)$ enthalten.

Für ein festes $i \in \{1, \dots, n\}$ verfahren wir wie folgt, falls $h_i \neq 0$ ist. Im Folgenden betrachten wir den Fall $h_i > 0$, der Fall $h_i < 0$ kann jedoch analog behandelt werden. Wähle $\epsilon > 0$, so dass $B_\epsilon(x_{i-1}), B_\epsilon(x_i) \subset B_r(x)$. Dann gilt für die Funktion

$$\begin{aligned} g : (-\epsilon, h_i + \epsilon) &\to \mathbb{R} \\ t &\mapsto f(x_{i-1} + te_i), \end{aligned}$$

dass

$$\lim_{s \to 0} \frac{g(t+s) - g(t)}{s} = \lim_{s \to 0} \frac{f(x_{i-1} + (t+s)e_i) - f(x_{i-1} + te_i)}{s} = \partial_i f(x_{i-1} + te_i).$$

Damit ist $g$ differenzierbar und es gilt $g'(t) = \partial_i f(x_{i-1} + te_i)$. Insbesondere ist $g$ also stetig auf $[0, h_i]$. Jetzt können wir den Mittelwertsatz anwenden um $\theta \in (0, h_i)$ zu erhalten, so dass

$$\partial_i f(x_{i-1} + \theta e_i)h_i = g'(\theta)h_i = g(h_i) - g(0) = f(x_i) - f(x_{i-1}).$$

Für $\xi_i := x_{i-1} + \theta e_i$ erhalten wir also

$$\partial_i f(\xi_i)h_i = f(x_i) - f(x_{i-1}).$$

Da die Kugel $K_{\|h\|}(x)$ konvex ist und $x_i$ und $x_{i-1}$ enthält, ist auch $\xi_i \in K_{\|h\|}(x)$.

Schließlich erhalten wird

$$f(x+h) - f(x) = \sum_{i=1}^n f(x_i) - f(x_{i-1}) = \sum_{i=1}^n \partial_i f(\xi_i)h_i.$$

Dabei ist zu beachten, dass falls $h_i = 0$ ist, sowohl $f(x_i) - f(x_{i-1})$ als auch $f(\xi_i)h_i$ verschwinden, unabhängig vom Wert von $\xi_i$. Wir können für diese Fälle also willkürlich z.B. $\xi_i = x_i$ setzen. □

*Beweis von Satz 2.33.* Da totale Differenzierbarkeit komponentenweise festgestellt werden kann, reicht es den Fall $m = 1$ zu betrachten. Wähle $x \in U$ fest und sei $\epsilon > 0$, so dass $B_\epsilon(x) \subset U$. Nach dem vorhergehenden Lemma existieren für jedes $h = (h_1, \dots, h_n) \in B_\epsilon(0)$ Punkte $\xi_1(h), \dots, \xi_n(h) \in K_{\|h\|}(x)$, so dass

$$f(x+h) - f(x) = \partial_1 f(\xi_1(h))h_1 + \cdots + \partial_n f(\xi_n(h))h_n.$$

Wir setzen

$$a(h) := (\partial_1 f(\xi_1(h)) - \partial_1 f(x), \cdots, \partial_n f(\xi_n(h)) - \partial_n f(x))$$

und erhalten

$$\begin{aligned} &f(x+h) - f(x) - \mathrm{D}f(x)h \\ &\qquad = (\partial_1 f(\xi_1(h)) - \partial_1 f(x))h_1 + \cdots + (\partial_n f(\xi_n(h)) - \partial_n f(x))h_n = \langle a(h) | h \rangle, \end{aligned}$$

wobei $\langle \cdot | \cdot \rangle$ das Standardskalarprodukt im $\mathbb{R}^n$ ist.

Dann folgt mit der Cauchy-Schwarz-Ungleichung

$$|f(x+h) - f(x) - \mathrm{D}f(x)h| = |\langle a(h) | h \rangle| \leq \|a(h)\| \, \|h\| \, .$$

Da $\|\xi_i(h) - x\| \leq \|h\|$ gilt $\lim_{h\to 0} \xi_i(h) = x$ für jedes $i \in \{1, \dots, n\}$. Nach Voraussetzung ist $\partial_i f$ stetig, also gilt $\lim_{h\to 0} \partial_i f(\xi_i(h)) = \partial_i f(x)$. Also konvergiert $a(h)$ gegen 0 und es gilt

$$\lim_{h\to 0} \frac{|f(x+h) - f(x) - \mathrm{D}f(x)h|}{\|h\|} \leq \lim_{h\to 0} \|a(h)\| = 0.$$

Damit ist $f$ in $x$ total differenzierbar und da $x \in U$ beliebig war ist $f$ total differenzierbar auf ganz $U$. $\square$

*Beispiel 2.35.* Seien $N \in \mathbb{N}$ und $c_\eta \in \mathbb{K}$ Koeffizienten für jeden Multiindex $\eta \in \mathbb{N}_0^n$ mit $|\eta| \leq \mathbb{N}$. Dann ist das Polynom

$$\begin{aligned} p : \mathbb{R}^n &\to \mathbb{K} \\ x &\mapsto \sum_{\eta, |\eta| < N} c_\eta x^\eta \end{aligned}$$

total differenzierbar.

Dazu sehen wir uns die partielle Ableitung eines Monoms an:

$$\partial_i x^\eta = \partial_i x_1^{\eta_1} \cdots x_{i-1}^{\eta_{i-1}} x_i^{\eta_i} x_{i+1}^{\eta_{i+1}} \cdots x_n^{\eta_n} = \eta_i x_1^{\eta_1} \cdots x_{i-1}^{\eta_{i-1}} x_i^{\eta_i - 1} x_{i+1}^{\eta_{i+1}} \cdots x_n^{\eta_n}$$

und stellen fest, dass diese ein Polynom ist. Damit ist jede partielle Ableitung eines Polynoms wiederum ein Polynom also insbesondere stetig. Dann folgt mit Satz 2.33, dass $p$ total differenzierbar ist.

**Korollar 2.36.** *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$. Dann ist $f$ stetig differenzierbar, das heißt die Abbildung $x \mapsto \mathrm{D}f(x)$ ist stetig, genau dann wenn es partiell differenzierbar ist und alle partiellen Ableitungen stetig sind.*

*Beweis.* Das folgt sofort, da Stetigkeit der Abbildung $x \mapsto \mathrm{D}f(x)$ äquivalent zur Stetigkeit der einzelnen Komponentenfunktionen, also der partiellen Ableitungen ist. □

**Satz 2.37 (Mittelwertsatz für reellwertige Funktionen).** *Sei $U \subset \mathbb{R}^n$, $x, y \in U$ und $S_{xy} := \{x + t(y - x) |\ t \in [0, 1]\} \subset U$. Sei $f : U \to \mathbb{R}$ differenzierbar auf $S_{xy}$ und stetig in $x$ und $y$, dann existiert $\xi \in S_{xy}$, so dass*

$$f(y) - f(x) = \mathrm{D}f(\xi)(y - x).$$

*Beweis.* Wir definieren die Hilsfunktion

$$\begin{aligned} g : [0, 1] &\to \mathbb{R} \\ t &\mapsto f(x + t(y - x)). \end{aligned}$$

Nach der Kettenregel und der Voraussetzung ist $g$ differenzierbar auf $(0, 1)$ und stetig auf $[0, 1]$, also können wir den (eindimensionalen) Mittelwertsatz anwenden. Wir erhalten also $\theta \in (0, 1)$, so dass $g(1) - g(0) = g'(\theta)$. Daraus folgt mit der Kettenregel

$$f(y) - f(x) = g(1) - g(0) = g'(\theta) = \mathrm{D}f(x + \theta(y - x))(y - x)$$

also für $\xi = x + \theta(y - x)$ die gesuchte Aussage. □

**Satz 2.38 (Allgemeiner Mittelwertsatz).** *Sei $U \subset \mathbb{R}^n$, $x, y \in U$ und*

$$S_{xy} := \{x + t(y - x) |\ t \in [0, 1]\} \subset U.$$

*Ist $f : U \to \mathbb{R}^m$ differenzierbar dann gilt*

$$\|f(y) - f(x)\| \leq \sup_{t \in [0,1]} \|\mathrm{D}f(x + t(y - x))\| \, \|y - x\| \, .$$

Ohne Beweis.

**Korollar 2.39.** *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ differenzierbar. Falls $\mathrm{D}f \equiv 0$ auf $U$ gilt, so ist $f$ lokal konstant, das heißt für jedes $x \in U$ existiert eine Umgebung $V$, so dass $f$ auf $V$ konstant ist.*

*Beweis.* Sei $x \in U$ und wähle $\epsilon > 0$ so, dass $B_\epsilon(x) \subset U$. Dann gilt für alle $y \in B_\epsilon(x)$ nach dem allgemeinen Mittelwertsatz

$$\|f(y) - f(x)\| \leq \sup_{t \in [0,1]} \|\mathrm{D}f(x + t(y - x))\| \, \|y - x\| = 0$$

also $f(y) = f(x)$. Damit ist $f$ auf $B_\epsilon(x)$ konstant. □

## 2.3 Multilineare Abbildungen und höhere Ableitungen

**Definition 2.40.** Seien $V_1, \dots, V_n, W$ Vektorräume. Die Abbildung $\varphi : V_1 \times \cdots \times V_n \to W$ heißt multilinear falls sie linear in jedem ihrer Argumente ist, das heißt die Abbildung von $V_i \to V$

$$v \mapsto \varphi(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_n)$$

ist linear für jedes $i \in \{1, \dots, n\}$ und alle $v_k \in V_k$ für $k \in \{1, \dots, n\}$ und $k \neq i$.

Wir bezeichnen den Raum der multilinearen Abbildungen von $\varphi : V_1 \times \cdots \times V_n \to W$ mit $\mathcal{L}(V_1, \dots, V_n|W)$. Falls $V_1 = V_2 = \cdots = V_n = V$ ist, so schreiben wir $\mathcal{L}^n(V|W)$.

*Beispiel 2.41.* (i) Multilineare Abbildungen sind Verallgemeinerungen von linearen Abbildungen.

(ii) Die Determinante ist eine multilineare Abbildung von $\mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$.

(iii) Ist $V$ ein reeller Vektorraum, dann ist jedes Skalarprodukt $\langle \cdot | \cdot \rangle$ eine multilineare Abbildung von $V \times V \to \mathbb{R}$.

(iv) Für Vektorräume $V$ und $W$ ist die Abbildung

$$\begin{aligned} \mathcal{L}(V|W) \times V &\to W \\ (A, v) &\mapsto A(v) \end{aligned}$$

eine multilineare Abbildung.

*Bemerkung 2.42.* (i) Multilineare Abbildungen der Form $V \times V \times \cdots \times V \to \mathbb{R}$, sogenannte Multilinearformen, sind (für endlichdimensionales $V$) eine Möglichkeit (kovariante) Tensoren über dem Vektorraum $V$ einzuführen.

(ii) Genau wie für lineare Abbildungen, ist auch eine multilineare Abbildungen $\varphi$ eindeutig durch ihre Werte bezüglich festgeleter Basen von $V_1, \dots, V_n$ und $W$ bestimmt. Wir betrachten den Fall $V_1 = V_2 = \cdots = V_n = V$. Es seien Basen $e_1, \dots, e_k$ von $V$ und $f_1, \dots, f_l$ von $W$ gegeben. Für vorgegebenes $\varphi : V^n \to W$ definieren die Gleichungen

$$\varphi(e_{i_1}, e_{i_2}, \cdots, e_{i_n}) = \sum_{j=1}^{l} a_{j i_1 i_2 \cdots i_n} f_j$$

für $i_1, \dots, i_n \in \{1, \dots, k\}$ die Komponenten $a_{j i_1 i_2 \cdots i_n}$ der multilinearen Abbildung $\varphi$. Für vorgegebene Komponenten $a_{j i_1 i_2 \cdots i_n}$, $i_1, \dots, i_n \in \{1, \dots, k\}$ und $j \in \{1, \dots, l\}$ definieren obige Gleichungen eindeutig eine multilineare Abbildung $\varphi$.

(iii) Insbesondere ist für $V = W = \mathbb{R}$ jede multilineare Abbildung durch genau eine Zahl bestimmt.

**Satz 2.43.** *Seien $V, W$ Vektorräume, dann ist die Abbildung $\Phi : \mathcal{L}(V|\mathcal{L}(V|W)) \to \mathcal{L}(V, V|W)$, definiert durch*

$$[\Phi(\varphi)](v_1, v_2) := [\varphi(v_1)](v_2)$$

*ein Isomorphismus von Vektorräumen.*

*Beweis.* Zunächst überprüfen wir die Linearität. Seien dazu $\varphi, \tilde{\varphi} \in \mathcal{L}(V|\mathcal{L}(V|W))$ und $\lambda \in \mathbb{R}$. Dann gilt für $v_1, v_2 \in V$

$$\begin{aligned}\Phi(\varphi + \lambda\tilde{\varphi})(v_1, v_2) &= (\varphi + \lambda\tilde{\varphi})(v_1)(v_2) = \varphi(v_1)(v_2) + \lambda\tilde{\varphi}(v_1)(v_2)\\ &= \Phi(\varphi)(v_1, v_2) + \lambda\Phi(\varphi)(v_1, v_2),\end{aligned}$$

also $\Phi(\varphi + \lambda\tilde{\varphi}) = \Phi(\varphi) + \lambda\Phi(\tilde{\varphi})$.

Wir müssen nun noch Injektivität und Surjektivität zeigen. Sei $\Phi(\varphi) = 0$, das heißt es gilt $0 = \Phi(\varphi)(v_1, v_2) = \varphi(v_1)(v_2)$ für alle $v_1, v_2 \in V$. Dann ist aber $\varphi(v_1) = 0$ für alle $v_1 \in V$ und es folgt auch $\varphi = 0$. Die Abbildung $\Phi$ ist also injektiv.

Sei nun $\theta \in \mathcal{L}(V, V|W)$. Wir setzen $\varphi(v_1)(v_2) := \theta(v_1, v_2)$. Da $\theta$ im zweiten Argument linear ist, ist $\varphi(v_1) \in \mathcal{L}(V|W)$ für jedes $v_1 \in V$. Da $\theta$ auch im ersten Argument linear ist, ist die Abbildung $v_1 \mapsto \varphi(v_1)$ ebenfalls linear, also $\varphi \in \mathcal{L}(V|\mathcal{L}(V|W))$. Es gilt nach Definition von $\varphi$

$$\Phi(\varphi)(v_1, v_2) = \varphi(v_1)(v_2) = \theta(v_1, v_2)$$

für alle $v_1, v_2 \in V$ und somit ist $\Phi(\varphi) = \theta$. Da $\theta$ beliebig war, ist damit die Surjektivität von $\Phi$ gezeigt. □

Analog zu obigem Resultat erhalten wir:

**Korollar 2.44.** *Seien $V, W$ Vektorräume. Dann ist $\mathcal{L}^n(V|\mathcal{L}^m(V|W))$ isomorph zu $\mathcal{L}^{n+m}(V|W)$. Setze $W_0 = W$ und $W_{n+1} = \mathcal{L}(V|W_n)$. Dann ist $W_n$ isomorph zu $\mathcal{L}^n(V|W)$ für $n \in \mathbb{N}$.*

**Definition 2.45.** Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n$. Die Abbildung $f$ heißt $k+1$ mal differenzierbar falls $\mathrm{D}^k f : U \to \mathcal{L}^k(R^n|R^m)$ differenzierbar ist und $\mathrm{D}^{k+1} f(x) = \mathrm{D}\mathrm{D}^k(f)(x)$. Die Abbildung heißt beliebig differenzierbar falls sie $k$ mal differenzierbar ist für jedes $k \in \mathbb{N}$. Die Abbildung heißt $k$ mal stetig differenzierbar falls $\mathrm{D}^k f$ stetig ist. Mit $\mathbf{C}^k(U, \mathbb{R}^n)$ bezeichnen wir den Vektorraum der $k$ mal stetig differenzierbaren Abbildungen von $U$ nach $\mathbb{R}^n$.

*Bemerkung 2.46.*

(i) In obiger Definition haben wir die Isomorphismen aus Korollar 2.44 benutzt und wir werden sie im weiteren Verlauf oft nutzen ohne jedes mal explizit darauf hinzuweisen.

(ii) Aus Satz 2.33 erhalten wir induktiv, dass eine Abbildung $f$ $k$ mal stetig differenzierbar ist, genau dann wenn alle partiellen Ableitungen vom Grad $k$ existieren und stetig sind.

**Lemma 2.47.** *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ mindestens $k+1$ mal differenzierbar und seien $v_1, \dots, v_k \in \mathbb{R}^n$. Dann ist die Abbildung*

$$\begin{aligned} g : U &\to \mathbb{R}^m \\ x &\mapsto \mathrm{D}^k f(x)(v_1, \dots, v_k) \end{aligned}$$

*differenzierbar und es gilt* $\mathrm{D}g(x) = \mathrm{D}^{k+1} f(x)(v_1, \dots, v_k, \,\cdot\,)$.

*Beweis.* Definiere die Abbildung

$$\begin{aligned} I : \mathcal{L}^k(\mathbb{R}^n|\mathbb{R}^m) &\to \mathbb{R}^m \\ A &\mapsto A(v_1, \dots, v_k). \end{aligned}$$

Dann gilt $g = I \circ \mathrm{D}^k f$. Die Abbildung $\mathrm{D}^k f$ ist nach Voraussetzung differenzierbar, die Abbildung $I$ ist linear und damit insbesondere differenzierbar und es gilt $\mathrm{D}I(A) = I$ für alle $A \in \mathcal{L}^k(\mathbb{R}^n|\mathbb{R}^m)$ (Beispiel 2.28). Dann ist nach der Kettenregel auch $g$ differenzierbar und es gilt

$$\mathrm{D}g(x)(v) = (\mathrm{D}I)(\mathrm{D}^k f(x))\mathrm{D}\mathrm{D}^k f(x)(v) = I\mathrm{D}^{k+1} f(x)(v) = \mathrm{D}^{k+1} f(x)(v_1, \dots, v_n, v). \quad \square$$

**Satz 2.48.** *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ mindestens $k$ mal differenzierbar und $f = (f_1, \dots, f_m)$. Sei $e_1, \dots, e_n$ die Standardbasis von $\mathbb{R}^n$. Dann gilt für die Komponenten von* $\mathrm{D}^k f(x)$

$$\left(\mathrm{D}^k f(x)\right)_{ji_1 \dots i_k} = \partial_{i_k} \cdots \partial_{i_1} f_j(x)$$

*für alle $j \in \{1, \dots, m\}$ und $i_1, \dots, i_k \in \{1, \dots, n\}$.*

*Beweis.* Für $k = 1$ ist die Aussage bereits bekannt. Wir führen nun eine Induktion über $k$ aus. Angenommen wir wissen bereits, dass die Aussage für $1 \leq l < k$ stimmt, das heißt es gilt

$$\left(\mathrm{D}^l f(x)\right)(e_{i_1}, \cdots, e_{i_l}) = \sum_{j=1}^{m} (\mathrm{D}^l f(x))_{ji_1 \cdots i_k} = \sum_{j=1}^{m} \partial_{i_l} \cdots \partial_{i_1} f_j(x).$$

Dann gilt nach dem vorhergehenden Lemma

$$\begin{aligned} \mathrm{D}^{l+1} f(x)(e_{i_1}, \cdots, e_{i_l}, e_{i_{l+1}}) &= \mathrm{D}\left(\sum_{j=1}^{m} \partial_{i_l} \cdots \partial_{i_1} f_j(x)\right)(e_{i_{l+1}}) \\ &= \sum_{j=1}^{m} \mathrm{D}\left(\partial_{i_l} \cdots \partial_{i_1} f_j\right)(x)(e_{i_{l+1}}) \\ &= \sum_{j=1}^{m} \partial_{i_{l+1}} \partial_{i_l} \cdots \partial_{i_1} f_j(x). \end{aligned} \quad \square$$

**Satz 2.49 (Satz von Schwarz).** *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ zweimal stetig differenzierbar, dann gilt für alle $x \in U$*

$$\mathrm{D}^2 f(x)(u,v) = \mathrm{D}^2 f(x)(v,u).$$

*Insbesondere gilt für die partiellen Ableitungen*

$$\partial_{i_1}\partial_{i_2} f(x) = \partial_{i_2}\partial_{i_1} f(x)$$

*für alle $i_1, i_2 \in \{1, \ldots, n\}$.*

*Beweis.* Es genügt den Fall $m = 1$ zu betrachten, da alle Ableitungen komponentenweise berechnet werden können.

Sei $x \in U$ fest und $\epsilon > 0$ so, dass $B_\epsilon(x) \subset U$. Seien $u, v \in \mathbb{R}^n$ und $0 < t < c = \frac{\epsilon}{2\max\{\|u\|,\|v\|\}}$. Wir wenden den (eindimensionalen) Mittelwertsatz auf die Hilfsfunktion

$$\begin{aligned} g_1 : [0,t] &\to \mathbb{R} \\ s &\mapsto f(x + tv + su) - f(x + su) \end{aligned}$$

an und finden $\xi \in (0,t)$ so, dass

$$g_1(t) - g_1(0) = g_1'(\xi)t = (\mathrm{D}f(x + tv + \xi u)(u) - \mathrm{D}f(x + \xi u)(u))t.$$

Dabei haben wir für die letzte Umformung die Kettenregel verwendet. Wir wenden nun erneut den Mittelwertsatz an, diesmal auf die Hilfsfunktion

$$\begin{aligned} g_2 : [0,t] &\to \mathbb{R} \\ s &\mapsto \mathrm{D}f(x + sv + \xi u)(u) \end{aligned}$$

und erhalten $\eta \in (0,t)$ so, dass

$$g_2(t) - g_2(0) = g_2'(\eta)t = \mathrm{D}^2 f(x + \eta v + \xi u)(u,v)t.$$

Für die letzte Umformung haben wir wiederum die Kettenregel sowie Lemma 2.47 benutzt.

Zusammen haben wir für jedes $t \in (0,c)$ Zahlen $\xi, \eta \in (0,t)$ gefunden, so dass

$$\begin{aligned} f(x + tv + tu) &- f(x + tv) - f(x + tu) + f(x) \\ &= g_1(t) - g_1(0) = (\mathrm{D}f(x + tv + \xi u)(u) - \mathrm{D}f(x + \xi u)(u))t \\ &= (g_2(t) - g_2(0))t = \mathrm{D}^2 f(x + \eta v + \xi u)(u,v)t^2. \end{aligned}$$

Da die zweite Ableitung als stetig vorausgesetzt wurde und $\xi$ und $\eta$ für $t \to 0$ gegen 0 konvergieren, erhalten wir

$$\lim_{t\to 0^+} \frac{f(x + tv + tu) - f(x + tv) - f(x + tu) + f(x)}{t^2} = \mathrm{D}^2 f(x)(u,v). \tag{2.6}$$

Diese Gleichung gilt für jedes $u, v \in \mathbb{R}^n$. Die linke Seite der Gleichung ist symmetrisch unter Vertauschung von $u$ und $v$, also muss die rechte Seite es auch sein und es gilt $\mathrm{D}^2 f(x)(u,v) = \mathrm{D}^2 f(x)(v,u)$.

Setzen wir für $u$ und $v$ Elemente der Standardbasis ein, also $u = e_{n_1}$ und $v = e_{n_2}$, so folgt aus Satz 2.48 die Vertauschbarkeit der partiellen Ableitungen. $\square$

*Bemerkung 2.50.* (i) Mittels Induktion folgt, dass $\mathrm{D}^k f(x)(u_1, \dots, u_k)$ seinen Wert unter beliebigen Vertauschungen der $u_i$ nicht ändert vorausgesetzt $\mathrm{D}^k f$ ist stetig. Man sagt, die Multilinearform $\mathrm{D}^k f(x)$ ist symmetrisch. Die Reihenfolge der Anwendung von höchstens $k$ partiellen Ableitungen ist also unerheblich falls $f \in \mathbf{C}^k(\mathbb{R}^n, \mathbb{R}^m)$ ist.

(ii) Ohne Stetigkeit von $\mathrm{D}^2 f$ können partielle Ableitungen von der Reihenfolge abhängen.

(iii) Gleichung (2.6) stellt die zweite Ableitung mittels eines einfachen Grenzwertes dar. Eine solche Darstellung kann zum Beispiel für die Diskretisierung von Differenzialgleichungen hilfreich sein. Es wurde jedoch zweimal stetig differenzierbares $f$ vorausgesetzt. Es gibt Funktionen, für die der Grenzwert auf der linken Seite existiert, die aber nicht zweimal differenzierbar sind.

**Satz 2.51 (Taylorformel).** *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ eine $l+1$ mal differenzierbare Funktion und sei $h \in \mathbb{R}^n$, so dass $x+th \in U$ für jedes $t \in [0,1]$. Dann gilt für ein $\theta \in (0,1)$*

$$f(x+h) = \sum_{k=0}^{l} \frac{1}{k!} \mathrm{D}^k f(x)(h, \dots, h) + \frac{1}{(l+1)!} \mathrm{D}^{l+1} f(x+\theta h)(h, \dots, h).$$

*Beweisidee.* Wende die eindimensionale Taylorformel mit Lagrange-Restglied auf die Hilfsfunktion

$$\begin{aligned} g : [0,1] &\to \mathbb{R} \\ t &\mapsto f(x+th) \end{aligned}$$

an. □

*Bemerkung 2.52.* (i) Stellen wir $h$ bezüglich der Standardbasis dar $h = \sum_{i=1}^n h_i e_i$, dann können wir Koordinatendarstellungen der einzelnen Terme bestimmen wie wir hier exemplarisch für den Term zweiten Grades tun wollen:

$$\mathrm{D}^2 f(x)(h,h) = \sum_{i,j=1}^{n} h_i h_j \mathrm{D}^2 f(x)(e_i, e_j) = \sum_{i,j=1}^{n} h_i h_j \partial_i \partial_j f(x).$$

Analog zum eindimensionalen Fall approximiert die Taylorformel eine $l+1$ mal differenzierbare Funktion also durch ein Polynom (in $n$ Variablen) dessen Koeffizienten partielle Ableitungen von $f$ an der Stelle $x$ sind.

(ii) Analog zu Korollar 2.19 ist es im Prinzip möglich lokale Extrema von hinreichend differenzierbaren Funktionen mittels der Ableitungen zu charakterisieren. Es gilt zum Beispiel: Falls $\mathrm{D}f(x) = 0, \cdots, \mathrm{D}^{l-1} f(x) = 0$ und $\mathrm{D}^l f(x) \neq 0$ so liegt in $x$

- ein Minimum vor, falls $l$ gerade und $\mathrm{D}^l f(x)$ positiv definit ist, das heißt $\mathrm{D}^l f(x)(h, \dots, h) > 0$ für alle $h \in \mathbb{R}^n \setminus \{0\}$,

- ein Maximum vor, falls $l$ gerade und $\mathrm{D}^l f(x)$ negativ definit ist,
- kein Extremum vor, falls $l$ ungerade oder $\mathrm{D}^l f(x)$ indefinit ist, das heißt es existieren $h_1, h_2 \in \mathbb{R}^n$, so dass $\mathrm{D}^l(h_1, \dots, h_1) > 0$ und $\mathrm{D}^l(h_2, \dots, h_2) < 0$.

Im Allgemeinen ist diese Charakterisierung jedoch weniger hilfreich als im eindimensionalen Fall. Zum Einen ist es insbesondere für $l > 2$ relativ kompliziert zu bestimmen ob eine multilineare Abbildung positiv definit ist. Zum Anderen können die Ableitungen auch (positiv) semidefinit sein, also $\mathrm{D}^l(h, \dots, h) \geq 0$ für alle $h \in \mathbb{R}^n \setminus \{0\}$ aber $\mathrm{D}^l(h, \dots, h) = 0$ für bestimmte $h$.

## 2.4 Funktionenfolgen und Differenzierbarkeit

*Beispiel 2.53.* Betrachte die Funktionenfolge $(f_n)$ gegeben durch

$$\begin{aligned} f_n &: \mathbb{R} \to \mathbb{R} \\ x &\mapsto \frac{1}{n} e^{\mathrm{i}nx}. \end{aligned}$$

Es gilt

$$\|f_n\|_\infty = \frac{1}{n}$$

also konvergiert die Folge $(f_n)$ gleichmäßig gegen 0.

Die Folge der Ableitungen

$$f_n'(x) = \mathrm{i}e^{\mathrm{i}nx}$$

konvergiert jedoch nur für $x = 0$.

Das obige Beispiel suggeriert, dass Bedingungen an die Konvergenz der Funktionenfolge nicht ausreichen um gutes Verhalten bezüglich der Ableitungen zu erreichen. Das liegt daran, dass selbst Funktionen mit beliebig kleinen Normen sehr stark oszillieren können. Wir müssen also zusätzlich Bedinungen an die Konvergenz der Folge der Ableitungen stellen.

*Bemerkung 2.54.* Wir haben in Abschnitt 1.6 die Supremumnorm für zahlwertige Funktionen eingeführt und damit die gleichmäßige Konvergenz definiert. Falls $f : X \to V$ eine Abbildung von einem metrischen Raum in einen normierten Raum $(V, \|\cdot\|)$ ist, dann können wir analog definieren

$$\|f\|_\infty = \sup \{ \|f(x)\| | \, x \in X \} .$$

Auch die Räume der beschränkten $\mathbf{B}(X, V)$ beziehungsweise stetigen beschränkten $\mathbf{C}_b(X, V)$ Funktionen können damit analog definiert werden. Gleichmäßige Konvergenz einer Folge solcher Funktionen ist dann die Konvergenz bezüglich dieser Norm.

Satz 1.74 und Satz 1.77 gelten dann entsprechend, vorausgesetzt der Raum $V$ ist vollständig. Dazu muss in den entsprechenden Beweisen nur jeweils der Betrag durch die Norm $\|\cdot\|$ ersetzt werden.

**Satz 2.55.** *Sei $U \subset \mathbb{R}^n$ und $f_n : U \to \mathbb{R}^m$ stetig differenzierbar für jedes $n \in \mathbb{N}$. Falls $f_n \to f$ und $\mathrm{D}f_n \to g$ gleichmäßig konvergieren, dann ist $f$ differenzierbar und es gilt $\mathrm{D}f = g$.*

*Beweis.* Es genügt wiederum den Fall $m = 1$ zu betrachten, da wir im allgemeinen Fall den Satz komponentenweise anwenden können. Da die $f_n$ stetig differenzierbar sind und $\mathrm{D}f_n$ gleichmäßig gegen $g$ konvergieren, folgt aus Satz 1.77 die Stetigkeit von $g$.

Sei nun $x \in U$ fest und $\epsilon > 0$. Dann gibt es $\delta > 0$ so, dass $\|g(y) - g(x)\| < \frac{\epsilon}{3}$ falls $\|y - x\| < \delta$. Außerdem existiert $N \in \mathbb{N}$, so dass für alle $n > N$ gilt $\|\mathrm{D}f_n - g\|_\infty \leq \frac{\epsilon}{3}$.

Sei $h \in B_\delta(x)$. Dann existiert nach dem Mittelwertsatz für jedes $n$ ein $\xi_n \in \{x + th |\, t \in (0,1)\}$, so dass $f_n(x+h) - f_n(x) = \mathrm{D}f_n(\xi_n)(h)$. Dann gilt für $n > N$

$$\begin{aligned}\frac{1}{\|h\|} |f_n(x+h) - f_n(x) - \mathrm{D}f_n(x)(h)| &= \frac{1}{\|h\|} |(\mathrm{D}f_n(\xi_n) - \mathrm{D}f_n(x))(h)| \\ &\leq \|\mathrm{D}f_n(\xi_n) - \mathrm{D}f_n(x)\| \\ &\leq \|\mathrm{D}f_n(\xi_n) - g(\xi_n)\| + \|g(\xi_n) - g(x)\| + \|g(x) - \mathrm{D}f_n(x)\| \leq \epsilon.\end{aligned}$$

Für $n \to \infty$ erhalten wir

$$\frac{1}{\|h\|} |f(x+h) - f(x) - g(x)(h)| \leq \epsilon.$$

Da $\epsilon$ beliebig positiv gewählt war, konvergiert die linke Seite für $h \to 0$ gegen 0. Damit ist $f$ in $x$ (total) differenzierbar mit Ableitung $\mathrm{D}f(x) = g(x)$. □

*Bemerkung 2.56.* Wir können nun auf $\mathbf{C}^1_b(\mathbb{R}^n, \mathbb{R}^m)$, dem Raum der differenzierbaren, beschränkten Funktionen mit beschränkten Ableitungen, eine Norm wie folgt definieren:

$$\|f\|_{\mathbf{C}^1} := \|f\|_\infty + \|Df\|_\infty . \tag{2.7}$$

Konvergenz bezüglich dieser Norm bedeutet dann genau gleichmäßige Konvergenz und gleichmäßige Konvergenz der Ableitungen. Der obige Satz zeigt dann, dass $\mathbf{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ vollständig ist.

Wir formulieren die folgenden Resultate über Potenzreihen lediglich für Entwicklungen um 0, sie gelten aber sinngemäß für beliebige Entwicklungspunkte.

**Satz 2.57.** *Sei $f(x) = \sum_{n=0}^\infty a_n x^n$ eine Potenzreihe mit positivem Konvergenzradius $r$. Dann ist $f$ auf $B_r(0)$ differenzierbar und es gilt $f'(x) = \sum_{n=0}^\infty n a_n x^{n-1}$.*

*Beweis.* Wir untersuchen zunächst den Konvergenzradius der formal abgeleiteten Reihe $\sum_{n=0}^\infty n a_n x^{n-1}$:

$$\limsup_{n\to\infty} \sqrt[n]{(n+1)a_{n+1}} = \lim_{n\to\infty} \sqrt[n]{n+1} \limsup_{n\to\infty} \sqrt[n]{a_{n+1}} = \frac{1}{r}.$$

Diese Reihe hat also den selben Konvergenzradius wie die ursprüngliche. Setze nun für $x \in B_r(0)$

$$g(x) := \sum_{n=0}^\infty n a_n x^{n-1}.$$

Sei nun $x \in B_r(0)$ fest und wähle $a > 0$, so dass $|x| < a < r$. Nach Satz 1.82 konvergieren die Partialsummenfolgen

$$f_N := \sum_{n=0}^{N} a_n x^n$$
$$g_N := \sum_{n=0}^{N} n a_n x^{n-1}$$

auf $B_a(0)$ (sogar auf $K_a(0)$) gleichmäßig gegen $f$ beziehungsweise $g$. Offensichtlich gilt $f_N' = g_N$. Damit folgt aus Satz 2.55, dass die Grenzfunktion $f$ differenzierbar auf $B_a(0)$ ist und $f' = g$.

Da Differenzierbarkeit eine lokale Eigenschaft ist und $x$ innerer Punkt von $B_a(0)$ ist gilt $f'(x) = g(x)$. Da $x \in B_r(0)$ beliebig war, gilt $f' = g$ auf $B_r(0)$. □

**Korollar 2.58.** *Sei $f(x) = \sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit positivem Konvergenzradius. Dann ist $f$ beliebig differenzierbar, und es gilt $a_k = \frac{f^{(k)}(0)}{k!}$.*

*Beweis.* Die beliebige Differenzierbarkeit folgt induktiv aus dem vorhergehenden Satz. Ebenfalls induktiv folgt, dass wir auch die höheren Ableitungen gliedweise berechnen dürfen. Es gilt also

$$f^{(k)}(x) = \sum_{n=0}^{\infty} n(n-1)\cdots(n-k)a_n x^{n-k}.$$

Für $x = 0$ folgt $f^{(k)}(0) = k! a_k$. □

*Beispiel 2.59.* Mit dieser Vorbereitung können wir jetzt die Ableitung der Exponentialfunktion berechnen:

$$(e^x)' = \sum_{n=0}^{\infty} \left(\frac{x^n}{n!}\right) = \sum_{n=0}^{\infty} \frac{nx^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = e^x.$$

Wir erhalten damit auch die Ableitungen von Sinus und Cosinus. Da beide Funktionen durch auf ganz $\mathbb{R}$ konvergente Potenzreihen gegeben sind, wissen wir bereits dass sie beliebig differenzierbar sind.

$$\begin{aligned}(e^{\mathrm{i}x})' &= (\cos(x) + \mathrm{i}\sin(x))' = (\cos(x))' + \mathrm{i}(\sin(x))' \\ &= \mathrm{i}e^{\mathrm{i}x} = \mathrm{i}\cos(x) - \sin(x).\end{aligned}$$

Vergleich von Real- und Imaginärteil ergibt

$$(\cos(x))' = -\sin(x) \qquad (\sin(x))' = \cos(x).$$

**Definition 2.60.** Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und seien $a_\eta \in \mathbb{K}$ Koeffizienten für jedes $\eta \in \mathbb{N}_0^d$. Dann ist

$$\sum_{\eta \in \mathbb{N}_0^d} a_\eta x^\eta$$

eine (formale) Potenzreihe in $d$ Variablen (wobei wir Multiindexnotation verwendet haben).

Eine Funktion $f : U \to \mathbb{K}$ für eine offene Nullumgebung $U \subset \mathbb{R}^d$ heißt analytisch in 0, wenn sie in eine Potenzreihe um 0 entwickelt werden kann, das heißt es existieren $\epsilon > 0$ und $a_\eta \in \mathbb{K}$ für jedes $\eta \in \mathbb{N}_0^d$, so dass

$$f(x) = \sum_{\eta \in \mathbb{N}_0^d} a_\eta x^\eta$$

für alle $x \in B_\epsilon(0)$.

*Bemerkung 2.61.* (i) Die Konvergenz der auftretenden Reihen kann zum Beispiel wie folgt definiert werden: wir nennen $\sum_{\eta \in \mathbb{N}_0^d} a_\eta x^\eta$ konvergent gegen $S$, falls für jedes $\epsilon > 0$ eine endliche Teilmenge $A \subset \mathbb{N}_0^d$ existiert, so dass für jede endliche Teilmenge $A \subset B \subset N_0^d$ gilt

$$\left| \sum_{\eta \in B} a_\eta x^\eta - S \right| < \epsilon.$$

(ii) Der Konvergenzbereich einer Potenzreihe in mehr als einer Variablen ist nicht mehr so einfach zu charakterisieren wie im eindimensionalen Fall (siehe Beispiel 2.62). Es gilt aber analog zum eindimensionalen Fall: falls $y = (y_1, \cdots, y_d) \in \mathbb{R}^d$ ein Punkt ist mit $y_1 \neq 0, \cdots, y_d \neq 0$, so dass

$$\sum_{\eta \in \mathbb{N}_0^d} a_\eta y^\eta$$

konvergiert, dann konvergiert die Reihe für jedes $x$ im offenen Quader

$$\left\{ x \in \mathbb{R}^d \middle| \ |x_i| < |y_i| \text{ für alle } i \in \{1, \ldots, d\} \right\}$$

absolut, das heißt es gilt

$$\sum_{\eta \in \mathbb{N}_0^d} |a_\eta| \left( |x_1|, \cdots, |x_2| \right)^\eta < \infty.$$

(iii) Konvergiert eine Potenzreihe auf einer offenen Umgebung $U$ von 0, so ist sie dort beliebig differenzierbar und es gilt

$$a_\eta = \frac{\partial^\eta f(0)}{\eta!}.$$

Dabei setzen wir

$$\partial^\eta f = \partial_1^{\eta_1} \cdots \partial_d^{\eta_d} f \text{ und } \eta! := \eta_1! \cdots \eta_d!.$$

(iv) Aus der vorhergehenden Bemerkung folgt insbesondere, dass die Koeffizienten der Potenzreihendarstellung einer Funktion (falls diese existiert) eindeutig bestimmt sind (vergleiche dazu auch Übungsaufgabe 3 vom Blatt 6). Die obige Gleichung ist selten nützlich um die Potenzreihendarstellung einer Funktion zu finden (falls diese existiert). Umgekehrt kann sie jedoch hilfreich sein, falls man viele (partielle) Ableitungen einer Funktion gleichzeitig berechnen will um zum Beispiel Extrema zu untersuchen oder die Taylorformel aufzustellen. Wir kennen bereits einige Reihendarstellungen, z.B. für $e^x, \sin(x), \cos(x), \frac{1}{1-x}$ sowie alle Polynome. In vielen Fällen kann man nun Reihendarstellungen für Funktionen finden, indem man sie mittels geschickter Umformungen als Summen oder Produkte (manchmal auch Verkettungen) dieser Funktionen darstellt:

$$f(x) = xe^{x^2} = x\sum_{n=0}^{\infty} \frac{(x^2)^n}{n!} = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{n!}.$$

(v) Es gilt

$$\mathbf{C}(U) \supset \mathbf{C}^1(U) \supset \mathbf{C}^2(U) \supset \cdots \supset \mathbf{C}^k(U) \supset \cdots \supset \mathbf{C}^\infty(U) \supset \mathbf{C}^\omega(U).$$

Dabei bezeichne $\mathbf{C}^\infty(U)$ den Raum der beliebig differenzierbaren Funktionen und $\mathbf{C}^\omega(U)$ den Raum der analytischen Funktionen. Alle Inklusionen in obiger Kette sind echt.

(vi) Die analytischen Funktionen sind abgeschlossen unter Summen, Produkten und Verkettungen. Eine Potenzreihe ist auf ihrem Konvergenzkreis analytisch. Die Beweise für diese Aussagen sind nicht trivial.

*Beispiel 2.62.* Die Potenzreihe in zwei reellen Variablen

$$\sum_{n=0}^{\infty} (xy)^n$$

konvergiert auf $\{(x,y) |\ |xy| < 1\}$ gegen $\frac{1}{1-xy}$. Diese Funktion ist also analytisch in 0.

## 2.5 Der Banachsche Fixpunktsatz und der Satz über implizite Funktionen

**Theorem 2.63 (Banachscher Fixpunktsatz).** *Sei $(X, d)$ ein vollständiger metrischer Raum und $\varphi : X \to X$ eine streng kontraktive Abbildung, das heißt es gibt $c \in [0, 1)$, so dass*

$$d(\varphi(x), \varphi(y)) \leq cd(x, y)$$

*für alle $x, y \in X$ gilt. Dann hat $\varphi$ genau einen Fixpunkt, das heißt es existiert genau ein $x \in X$, so dass $\varphi(x) = x$.*

*Beweis.* Zunächst stellen wir fest, dass $\varphi$ insbesondere Lipschitz-stetig, also stetig ist. Sei $x_0 \in X$ und definiere rekursiv die Folge $x_n := \varphi(x_{n-1})$ für $n \in \mathbb{N}$. Dann gilt auf Grund der strengen Kontraktivität

$$d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1})) \leq cd(x_n, x_{n-1})$$

und mit Induktion folgt

$$d(x_{n+k}, x_{n+k-1}) \leq c^k d(x_n, x_{n-1})$$

für alle $k, n \in \mathbb{N}$. Insbesondere folgt

$$d(x_n, x_{n-1}) \leq c^{n-1} d(x_1, x_0)$$

Aus der Dreiecksungleichung folgt dann

$$\begin{aligned} d(x_{n+k}, x_{n-1}) &\leq d(x_{n+k}, x_{n+k-1}) + d(x_{n+k-1}, x_{n+k-2}) + \cdots + d(x_n, x_{n-1}) \\ &\leq \left(c^k + c^{k-1} + \cdots + c + 1\right) d(x_n, x_{n-1}) \\ &= \frac{1 - c^{k+1}}{1 - c} d(x_n, x_{n-1}) \leq \frac{1 - c^{k+1}}{1 - c} c^{n-1} d(x_1, x_0) \\ &\leq \frac{c^{n-1}}{1 - c} d(x_1, x_0). \end{aligned}$$

Da der letzte Term unabhängig von $k$ ist und für $n \to \infty$ gegen 0 konvergiert, haben wir damit die Cauchy-Eigenschaft der Folge $(x_n)$ gezeigt. Nach Voraussetzung ist $X$ vollständig, also konvergiert $(x_n)$ gegen ein $x \in X$. Dann gilt

$$x = \lim_{n \to \infty} x_n = \lim_{n \to \infty} \varphi(x_{n-1}) = \varphi\left(\lim_{n \to \infty} x_{n-1}\right) = \varphi(x)$$

wobei wir im vorletzten Schritt die Stetigkeit von $\varphi$ verwendet haben.

Es bleibt noch die Eindeutigkeit des Fixpunktes zu zeigen. Seien $x, y \in X$ Fixpunkte von $\varphi$. Dann gilt

$$d(x, y) = d(\varphi(x), \varphi(y)) \leq cd(x, y).$$

Da $c < 1$ folgt daraus $d(x, y) = 0$ und aus der Definitheit der Metrik $x = y$. $\square$

*Bemerkung 2.64.* (i) Eine Folgerung aus dem Banachschen Fixpunktsatz: läßt man in diesem Hörsaal einen Stadtplan von Leipzig zufällig auf den Boden fallen, dann gibt es genau einen Punkt auf dem Stadtplan, der genau über dem entsprechenden Punkt der Stadt liegt.

(ii) Aus dem Beweis ist auch ein Verfahren zum Bestimmen des Fixpunktes ersichtlich: man wendet $\varphi$ wiederholt auf einen beliebigen Startpunkt an. Läßt man in einer der obigen Abschätzungen $k \to \infty$ gehen, so erhält man

$$d(x, x_{n-1}) \leq \frac{1}{1-c} d(x_n, x_{n-1}),$$

also eine Abschätzung für den Fehler.

(iii) Viele Probleme lassen sich in Fixpunktprobleme überführen. Die Schwierigkeit liegt dann häufig darin, einen passenden metrischen Raum zu finden, so dass man den Satz anwenden kann.

*Beispiel 2.65.* Betrachte die Gleichung $x - y^2 = 0$. Für $x \in (0, \infty)$ gibt es die beiden Lösungen

$$y = \sqrt{x} \text{ beziehungsweise } y = -\sqrt{x}.$$

Für jeden Lösungspunkt $(\xi, \eta)$ der Gleichung gibt es also eine Umgebung $U$ und eine Funktion $f$, so dass alle in $U$ liegenden Lösungen der Gleichung durch $(x, f(x))$ gegeben sind. Darüber hinaus ist $f$ sogar stetig differenzierbar.

*Bemerkung 2.66.* In vielen Fällen wird es nicht mehr möglich sein, Formeln für die Lösungen von Gleichungssystemen anzugeben. Der Satz über implizite Funktionen erlaubt es unter bestimmten Voraussetzungen zumindest die Existenz und bestimmte Eigenschaften solcher Lösungen nachzuweisen.

Sei $F : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^n$ eine Abbildung. Betrachten wir $x_1, \ldots, x_p$ als unabhängige Variablen und $y_1, \ldots, y_q$ als abhängige Variablen und setzen $x = (x_1, \ldots, x_p)$, $y = (y_1, \ldots, y_q)$, dann ist durch

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

ein Gleichungssystem (mit $n$ Gleichungen) gegeben.

Ist so ein $F$ gegeben, so bezeichnen wir mit $D_y F(x, \eta)$ die Ableitung der Funktion $\mathbb{R}^q \ni y \mapsto F(x, y)$ an der Stelle $\eta$, also für festes $x$. Analog ist $D_x F(\xi, y)$ die Ableitung der Funktion $\mathbb{R}^p \ni x \mapsto F(x, y)$ an der Stelle $\xi$ für festes $y$.

**Theorem 2.67 (Satz über implizite Funktionen).** *Seien $U \subset \mathbb{R}^p$ und $V \subset \mathbb{R}^q$ offen. Sei $F : U \times V \to \mathbb{R}^q$ stetig differenzierbar und es seien $\xi \in U$ und $\eta \in V$ so, dass $F(\xi, \eta) = 0$ und $\mathrm{D}_y F(\xi, \eta)$ invertierbar ist.*

*Dann existieren Umgebungen $\tilde{U} \subset U$ von $\xi$ und $\tilde{V} \subset V$ von $\eta$ und eine stetige Funktion $f : \tilde{U} \to \tilde{V}$, so dass $f(\xi) = \eta$ und $F(x, f(x)) = 0$ für alle $x \in \tilde{U}$.*

*Beweis.* Wir setzen $D = \mathrm{D}_y F(\xi, \eta)$. Wir betrachten das Fixpunktproblem (auf Funktionen $g$)

$$g(x) - D^{-1}F(x, g(x)) = g(x) \qquad \forall x \in \tilde{U}.$$

Eine Funktion $f$ ist Lösung dieses Fixpunktproblems genau dann wenn $F(x, f(x)) = 0$ (da $D$ nach Voraussetzung invertierbar ist). Wir suchen also Fixpunkte der Abbildung

$$(\varphi(g))(x) := g(x) - D^{-1}F(x, g(x)).$$

Auf dieses Fixpunktproblem wollen wir den Banachschen Fixpunktsatz anwenden und müssen nun noch einen geeigneten Definitionsbereich für $\varphi$ suchen.

Bezeichne $I : \mathbb{R}^q \to \mathbb{R}^q$ die identische Abbildung. Die Funktion

$$(x, y) \mapsto \left\| I - D^{-1}\mathrm{D}_y F(x, y) \right\|$$

ist nach Voraussetzung stetig und verschwindet im Punkt $(\xi, \eta)$. Es existieren also $\delta, \epsilon > 0$ so, dass $B_\delta(\xi) \subset U$, $B_\epsilon(\eta) \subset V$ und

$$\left\| I - D^{-1}\mathrm{D}_y F(x, y) \right\| \leq \frac{1}{2}$$

für alle $x \in B_\delta(\xi)$ und $y \in \tilde{V} := B_\epsilon(\eta)$. Auf Grund der Stetigkeit von $x \mapsto \left\| D^{-1}F(x, \eta) \right\|$ können wir ein (möglicherweise kleineres) $\delta > 0$ wählen, so dass

$$\left\| D^{-1}F(x, \eta) \right\| \leq \frac{\epsilon}{4}$$

für alle $x \in \tilde{U} = B_\delta(\xi)$.

Sei nun $X$ die Menge all der Funktionen $g : \tilde{U} \to \mathbb{R}^q$ die

(i) $g$ ist stetig,

(ii) $g(\xi) = \eta$ und

(iii) $\|g(x) - \eta\| \leq \frac{\epsilon}{2}$ für alle $x \in \tilde{U}$

erfüllen. Die dritte Bedingung impliziert insbesondere, dass $g(x) \in \tilde{V}$ für alle $g \in X$ und $x \in \tilde{U}$. Damit ist $X$ Teilmenge des Banachraumes $\mathbf{B}(\tilde{U}, \mathbb{R}^q)$ mit der Norm

$$\|g\|_\infty = \sup\left\{ \|g(x)\| \,\middle|\, x \in \tilde{U} \right\}.$$

Darüber hinaus ist $X$ in diesem Raum abgeschlossen, da die drei obigen Bedingungen unter gleichmäßigen Grenzwerten erfüllt bleiben (siehe Bemerkung 2.54) also nach Satz 1.38 vollständig.

Betrachte nun für festes $x \in \tilde{U}$ die Abbildung

$$\begin{aligned} \Phi : \tilde{V} &\to \mathbb{R}^q \\ y &\mapsto y - D^{-1}F(x, y). \end{aligned}$$

Dann folgt mit dem allgemeinen Mittelwertsatz Satz 2.38 für alle $y, z \in V$

$$\|\Phi(y) - \Phi(z)\| \leq \sup_{v \in \tilde{V}} \left\|I - D^{-1} \mathrm{D}_y F(x, y)\right\| \|y - z\| \leq \frac{1}{2} \|y - z\| .$$

Insbesondere gilt für $g_1, g_2 \in X$ und $x \in \tilde{U}$

$$\|(\varphi(g_1))(x) - (\varphi(g_2))(x)\| = \|\Phi(g_1(x)) - \Phi(g_2(x))\| \leq \frac{1}{2} \|g_1(x) - g_2(x)\|$$

und damit ist $\varphi$ eine strenge Kontraktion

$$\|\varphi(g_1) - \varphi(g_2)\|_\infty \leq \frac{1}{2} \|g_1 - g_2\|_\infty .$$

Für $g \in X$ ist $\varphi(g)$ stetig und $(\varphi(g))(\xi) = \eta$. Schließlich gilt

$$\|(\varphi(g))(x) - \eta\| \leq \|(\varphi(g))(x) - (\varphi(\eta))(x)\| + \|(\varphi(\eta))(x) - \eta\| \leq \frac{1}{2} \|g(x) - \eta\| + \frac{\epsilon}{4} \leq \frac{\epsilon}{2}.$$

Damit ist gezeigt, dass $\varphi$ die Menge $X$ in sich selbst abbildet.

Jetzt sind alle Voraussetzungen des Banachschen Fixpunktsatzes erfüllt. Die Abbildung $\varphi$ hat also genau einen Fixpunkt $f \in X$ und für dieses $f$ gilt $F(x, f(x)) = 0$ für alle $x \in \tilde{U}$. □

Für die folgenden Aussagen behalten wir die Voraussetzungen sowie die Notation des Theorems und seines Beweises bei.

*Bemerkung 2.68.* Die Eindeutigkeitsaussage des Banachschen Fixpunktsatzes besagt, dass es für fest gewähltes $\tilde{U}$ und $\tilde{V}$ genau eine Funktion in $X$ gibt, die $F(x, f(x))$ erfüllt. Es gilt jedoch etwas stärker: für festes $x \in \tilde{U}$ ist $f(x)$ die einzige Lösung des Gleichungssystems $F(x, y) = 0$. Für jede Lösung $y$ gilt

$$\|y - f(x)\| = \|\Phi(y) - \Phi(f(x))\| \leq \frac{1}{2} \|y - f(x)\| ,$$

was nur durch $y = f(x)$ zu erfüllen ist.

**Satz 2.69.** *Es gibt eine (möglicherweise kleinere) Umgebung $\tilde{\tilde{U}}$ von $\xi$ auf der $f$ stetig differenzierbar ist. Die Ableitung ist durch*

$$\mathrm{D} f(x) = -\left(\mathrm{D}_y F(x, f(x))\right)^{-1} \mathrm{D}_x F(x, f(x))$$

*gegeben.*

*Ohne Beweis.*

**Korollar 2.70 (Satz über die Umkehrfunktion).** *Sei $U \subset \mathbb{R}^n$ und $f : U \to \mathbb{R}^n$ stetig differenzierbar. Falls $\mathrm{D}f(\xi)$ invertierbar ist für ein $\xi \in U$, dann gibt es eine Umgebung $\tilde{U}$ von $\xi$ und eine Umgebung $\tilde{V}$ von $\eta = f(\xi)$, so dass $f$ die Umgebung $\tilde{U}$ bijektiv auf $\tilde{V}$ abbildet und dass die Umkehrfunktion*

$$\begin{aligned} g : \tilde{V} &\to \tilde{U} \\ y &\mapsto f^{-1}(y) \end{aligned}$$

*stetig differenzierbar ist. Es gilt*

$$\mathrm{D}g(\eta) = (\mathrm{D}f(\xi))^{-1}.$$

*Beweisidee.* Wende den Satz über implizite Funktionen an, um das Gleichungssystem

$$F(x, y) = f(x) - y = 0$$

(lokal) nach $x$ aufzulösen. □

*Beispiel 2.71.* Betrachte die Exponentialfunktion

$$\begin{aligned} f : \mathbb{R} &\to (0, \infty) \\ x &\mapsto e^x. \end{aligned}$$

Sie ist streng monoton wachsend (Warum eigentlich?), also injektiv und stetig, also auf Grund des Zwischenwertsatzes auch surjektiv. Es existiert also die Umkehrfunktion

$$\begin{aligned} g : (0, \infty) &\to \mathbb{R} \\ y &\mapsto \ln(y). \end{aligned}$$

Es gilt $f'(x) = f(x) > 0$ für jedes $x \in \mathbb{R}$. Dann sagt der Satz über die Umkehrfunktion, dass $f$ lokal invertierbar ist (das wissen wir bereits, sie ist ja sogar global invertierbar), dass die Inverse $g$ differenzierbar ist und für die Ableitung gilt für $\eta = e^\xi$:

$$g'(\eta) = \frac{1}{f'(\xi)} = \frac{1}{e^\xi} = \frac{1}{\eta}.$$

# 3 Maße und Integrale

## 3.1 Inhalte und Maße

**Definition 3.1.** Sei $\Omega$ eine Menge. Ein System von Teilmengen ($I$ eine beliebige Indexmenge) $(A_i)_{i\in I} \subset \mathcal{P}(\Omega)$ heißt paarweise disjunkt, wenn $A_i \cap A_j = \emptyset$ wann immer $i \neq j$.

*Bemerkung 3.2.* (i) Sei $\mathcal{A} \subset \mathcal{P}(\mathbb{R}^n)$ System von Teilmengen. Eine Abbildung $\mu : \mathcal{A} \to [0,\infty]$ heißt Inhalt, falls für alle $A_1, \dots, A_n \in \mathcal{A}$ paarweise disjunkt gilt: falls $A_1 \cup \dots \cup A_n \in \mathcal{A}$, dann gilt

$$\mu(A_1 \cup \dots \cup A_n) = \sum_{i=1}^{n} \mu(A_i) \qquad \text{(Additivität)}.$$

Der Inhaltsbegriff verallgemeinert die Begriffe Länge (1-dim.), Fläche (2-dim.) und Volumen (3-dim.).

(ii) Im Kontext von Maßen und Inhalten (und später Integralen) definieren wir

$$\begin{aligned} c + \infty &= \infty \quad \forall c \in \mathbb{R} \cup \infty \\ 0 \cdot \infty &= 0. \end{aligned}$$

(iii) Ziel wird es sein, das Mengensystem $\mathcal{A}$ möglichst groß zu wählen. Ideal wäre $\mathcal{A} = \mathcal{P}(\mathbb{R}^n)$, aber das sogenannte Banach-Tarski-Paradoxon besagt: Es existiert eine disjunkte Zerlegung

$$A_1 \cup \dots \cup A_p \cup B_1 \cup \dots \cup B_q = B_1(0) \subset \mathbb{R}^3$$

und Bewegungen (Kombinationen aus Drehungen und Verschiebungen) $D_1, \dots, D_p$ sowie $T_1, \dots, T_q$, so dass

$$D_1 A_1 \cup \dots \cup D_p A_p = B_1(0) = T_1 B_1 \cup \dots \cup T_q B_q.$$

Es kann also keinen unter Bewegungen invarianten Inhalt auf $\mathcal{P}\left(\mathbb{R}^3\right)$ geben.

**Definition 3.3.** Sei $\Omega$ eine Menge. Ein System $\mathcal{A}$ von Teilmengen von $\Omega$ heißt $\sigma$-Algebra, wenn

(i) $\emptyset \in \mathcal{A}$,

(ii) $A \in \mathcal{A} \Longrightarrow A^c \in \mathcal{A}$,

(iii) für eine abzählbare Teilmenge $\{A_1, A_2, \dots\} \subset \mathcal{A}$ folgt

$$\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{A}.$$

Eine Abbildung $\mu : \mathcal{A} \to [0,\infty]$ heißt ein Maß, falls $\mu(\emptyset) = 0$ und für paarweise disjunkte Teilmengen $(A_i)_{i\in\mathbb{N}} \subset \mathcal{A}$ gilt

$$\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}} \mu(A_i) \qquad \sigma\text{-Additivität.}$$

Eine Grundmenge zusammen mit einer $\sigma$-Algebra $(\Omega, \mathcal{A})$ heißt messbarer Raum oder Messraum.

Ist zusätzlich ein Maß vorgegeben, so heißt das Tripel $(\Omega, \mathcal{A}, \mu)$ Maßraum.

*Beispiel 3.4.* (i) Sei $\Omega$ eine beliebige Menge mit der disjunkten Zerlegung

$$\Omega = A_1 \cup \dots \cup A_n.$$

Dann ist

$$\left\{ \bigcup_{i\in I} A_i \,\middle|\, I \subset \{1,\dots,n\} \right\}$$

eine $\sigma$-Algebra.

(ii) Sei $\Omega$ beliebig und $x \in \Omega$. Dann ist

$$\begin{aligned} \delta_x : \mathcal{P}(\Omega) &\to [0,\infty] \\ A &\mapsto \begin{cases} 1 & x \in A \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

ein Maß.

(iii) Sei $\Omega$ beliebig, dann ist

$$\begin{aligned} \# : \mathcal{P}(\Omega) &\to [0,\infty] \\ A &\mapsto \begin{cases} \text{Anzahl der Elemente in } A & A \text{ endlich} \\ \infty & \text{sonst} \end{cases} \end{aligned}$$

ein Maß, das sogenannte Zählmaß. Dieses Maß ist besonders für endliche oder abzählbare Mengen nützlich.

(iv) Sei $\Omega$ abzählbar und $(a_\omega)_{\omega\in\Omega} \subset [0,\infty]$, dann ist

$$\begin{aligned} \mu : \mathcal{P}(\Omega) &\to [0,\infty] \\ A &\mapsto \sum_{\omega\in A} a_\omega \end{aligned}$$

ein Maß. Man kann zeigen, dass jedes Maß auf einer abzählbaren Menge $\Omega$ von dieser Form ist.

(v) Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $A \in \mathcal{A}$. Definiere die auf $A$ eingeschränkte $\sigma$-Algebra

$$\mathcal{A}|_A := \{B \cap A |\; B \in \mathcal{A}\}.$$

Dann ist $(A, \mathcal{A}|_A, \mu)$ ein Maßraum.

*Bemerkung 3.5.* Für abzählbar viele Teilmengen $\{A_1, A_2, \ldots\} \subset \mathcal{A}$ einer $\sigma$-Algebra $\mathcal{A}$ gilt

$$\bigcap_{i\in\mathbb{N}} A_i = \left(\bigcup_{i\in\mathbb{N}} A^c\right)^c \in \mathcal{A}.$$

Falls $A, B \in \mathcal{A}$, dann gilt auch $A \backslash B = A \cap B^c \in \mathcal{A}$.

Ein Maß ist stets monoton, das heißt falls $A, B \in \mathcal{A}$ und $A \subset B$, dann gilt

$$\mu(B) = \mu(B \backslash A) + \mu(A) \geq \mu(A).$$

**Definition 3.6.** Eine Abbildung $\mu : \mathcal{P}(\Omega) \to [0, \infty]$ heißt äußeres Maß, falls $\mu(\emptyset) = 0$ und

$$A \subset \bigcup_{i\in\mathbb{N}} A_i \Longrightarrow \mu(A) \leq \sum_{i\in\mathbb{N}} \mu(A_i)$$

gilt.

Genauso wie Maße, sind auch äußere Maße monoton.

Sei $\mathcal{I}$ die Menge der beschränkten Intervalle

$$\mathcal{I} = \bigcup_{x,y\in\mathbb{R}, x<y} \{[x, y], [x, y), (x, y], (x, y)\}.$$

Wir definieren auf $\mathcal{I}$ die Länge als

$$l([x, y]) := l([x, y)) := l((x, y]) := l((x, y)) = y - x.$$

**Satz 3.7.** *Die Abbildung*

$$\begin{aligned} \lambda : \mathcal{P}(\mathbb{R}) &\to [0, \infty] \\ A &\mapsto \inf\left\{ \sum_{i\in\mathbb{N}} l(I_i) \,\middle|\, A \subset \bigcup_{i\in\mathbb{N}} I_i, I_i \in \mathcal{I}\; \forall i \in \mathbb{N} \right\} \end{aligned}$$

*definiert ein äußeres Maß auf den reellen Zahlen.*

*Beweis.* Es gilt $\lambda(\emptyset) \leq l([0,\epsilon]) = \epsilon$ für jedes $\epsilon > 0$ also $\lambda(\emptyset) = 0$.

Sei nun $A \subset \bigcup_{k\in\mathbb{N}} A_k$. Wir wollen zeigen, dass

$$\lambda(A) \leq \sum_{k\in\mathbb{N}} \lambda(A_k)$$

erfüllt ist. Falls die Summe auf der rechten Seite (bestimmt) divergiert, ist nichts zu zeigen, wir können also $\lambda(A_k) < \infty$ annehmen. Sei $\epsilon > 0$, dann existieren für jedes $k \in \mathbb{N}$ Intervalle $(I_{ki})$, so dass

$$A_k \subset \bigcup_{i\in\mathbb{N}} I_{ki} \text{ und } \sum_{i\in\mathbb{N}} l(I_{ki}) < \lambda(A_k) + \frac{\epsilon}{2^k}.$$

Dann gilt

$$A \subset \bigcup_{k\in\mathbb{N}} A_k \subset \bigcup_{k,i\in\mathbb{N}} I_{ki}$$

und

$$\lambda(A) \leq \sum_{k,i\in\mathbb{N}} l(I_{ki}) < \sum_{k\in\mathbb{N}} \left(\lambda(A_k) + \frac{\epsilon}{2^k}\right) = \sum_{k\in\mathbb{N}} \lambda(A_k) + \epsilon.$$

Da das für jedes $\epsilon > 0$ gilt, folgt die gesuchte Ungleichung. $\square$

Das äußere Maß ist noch nicht als unser gesuchter Volumenbegriff geeignet, da es im Allgemeinen nicht endlich additiv ist. Wir möchten daher den Definitionsbereich einschränken, so dass wir (sogar $\sigma$-) Additivität erhalten.

**Satz 3.8.** *Sei $\mu$ ein äußeres Maß auf $(\Omega, \mathcal{P}(\Omega))$. Dann ist das System der messbaren Mengen*

$$\mathcal{A} = \{A \subset \Omega |\ \mu(E) \geq \mu(E\cap A) + \mu(E \cap A^c)\ \forall E \in \mathcal{P}(\Omega)\}$$

*eine $\sigma$-Algebra und $\mu|_{\mathcal{A}}$ ist ein Maß.*

*Beweis.* Zunächst folgt aus der Subadditivität stets $\mu(E) \leq \mu(E\cap A) + \mu(E\cap A^c)$, eine Menge ist also messbar genau dann wenn

$$\mu(E) = \mu(E\cap A) + \mu(E\cap A^c)$$

für alle $E \in \mathcal{P}(\Omega)$.

Offenbar ist $\emptyset$ messbar und $A$ messbar genau dann wenn $A^c$ messbar ist.

Es gilt für jedes $E \in \mathcal{P}(\Omega)$

$$E\cap(A\cup B) = (E\cap A)\cup(E\cap B) = (E\cap A)\cup(E\cap B\cap A^c)$$

und so folgt für $A, B$ messbar mit der Subadditivität von $\mu$

$$\begin{aligned}\mu(E) &= \mu(E\cap A) + \mu(E\cap A^c) = \mu(E\cap A) + \mu(E\cap A^c\cap B) + \mu(E\cap A^c\cap B^c)\\ &\geq \mu(E\cap(A\cup B)) + \mu(E\cap(A\cup B)^c) \geq \mu(E),\end{aligned}$$

Damit ist $A\cup B$ messbar und die obigen Ungleichungen sind in Wahrheit Gleichungen, es gilt also

$$\mu(E\cap A) + \mu(E\cap A^c\cap B) = \mu(E\cap(A\cup B)).$$

Aus der letzten Gleichung folgt falls $A, B$ zusätzlich disjunkt sind

$$\mu(E\cap(A\cup B)) = \mu(E\cap A) + \mu(E\cap B).$$

Seien nun $(A_i)_{i\in\mathbb{N}}$ paarweise disjunkte, messbare Mengen und setze

$$S_n = \bigcup_{i=1}^{n} A_i \text{ und } S = \bigcup_{i=1}^{\infty} A_i.$$

Dann folgt aus dem oben gezeigten induktiv, dass $S_n$ messbar ist und für jedes $E\in\mathcal{P}(\Omega)$ gilt

$$\mu(E\cap S_n) = \sum_{i=1}^{n} \mu(E\cap A_i).$$

Die Messbarkeit impliziert dann auch

$$\mu(E) \geq \mu(E\cap S_n) + \mu(E\cap S_n^c) \geq \sum_{i=1}^{n} + \mu(E\cap S^c)$$

wobei wir die Monotonie des äußeren Maßes benutzt haben. Lassen wir nun $n\to\infty$ gehen und benutzen nochmals die Subadditivität dann folgt

$$\mu(E) \geq \sum_{i=1}^{\infty} \mu(E\cap A_i) + \mu(E\cap S^c) \geq \mu(E\cap S) + \mu(E\cap S^c) \geq \mu(E).$$

Daraus ist wiederum ersichtlich, dass $S$ messbar ist und

$$\mu(E\cap S) = \sum_{i=1}^{n} \mu(E\cap A_j)$$

gilt. Insbesondere folgt, für $E=\Omega$ die $\sigma$-Additivität von $\mu$ für messbare Mengen.

Es bleibt noch zu zeigen, dass beliebige abzählbare Vereinigungen messbarer Mengen wieder messbar sind. Das folgt jedoch leicht, da für messbare $(A_i)_{i\in\mathbb{N}}$ nach dem bereits gezeigten die Mengen

$$B_i := A_i \backslash \left( \bigcup_{j=1}^{i-1} A_i \right)$$

messbar und paarweise disjunkt sind und es gilt

$$\bigcup_{i\in\mathbb{N}} A_i = \bigcap_{i\in\mathbb{N}} B_i. \qquad \square$$

**Definition 3.9.** Wenden wir den vorhergehenden Satz auf das äußere Maß in Satz 3.7 an, so erhalten wir die $\sigma$-Algebra der Lebesgue-messbaren Mengen und das Lebesgue-Maß $\lambda$ auf den reellen Zahlen.

Wir wissen im Moment noch nicht, dass diese $\sigma$-Algebra außer $\emptyset$ und $\Omega$ noch weitere Mengen enthält und die Begriffe damit sinnvoll sind.

*Bemerkung 3.10.* Für eine Teilmenge $A \subset \mathbb{R}$ und $x \in \mathbb{R}$ bezeichne

$$A + x := \{y + x |\ y \in A\}.$$

Ein Maß $\mu$ auf einer $\sigma$-Algebra $\mathcal{A}$ über $\mathbb{R}$ heißt translationsinvariant, wenn $\mu(A + x) = \mu(A)$ gilt für alle $A \in \mathcal{A}$ und alle $x \in \mathbb{R}$. Dazu muss insbesondere gelten, dass die $\sigma$-Algebra translationsinvariant ist, das heißt $A + x \in \mathcal{A}$ für alle $A \in \mathcal{A}$ und alle $x \in \mathbb{R}$.

Da Translationen von Intervallen wieder Intervalle sind, überzeugt man sich leicht, dass das äußere Lebesgue-Maß, und damit auch das Lebesgue-Maß translationsinvariant sind.

Man kann darüber hinaus zeigen, dass $\lambda$ das einzige translationsinvariante Maß ist, dass $\lambda([0, 1]) = 1$ erfüllt.

**Satz 3.11.** *Sei $\mathcal{E} \subset \mathcal{P}(\Omega)$ ein beliebiges Mengensystem, dann gibt es eine kleinste $\sigma$-Algebra $\sigma(\mathcal{E})$, die $\mathcal{E}$ enthält. Es gilt*

$$\sigma(\mathcal{E}) = \bigcap_{\mathcal{A}\ \sigma\text{-Algebra},\ \mathcal{E}\subset\mathcal{A}} \mathcal{A}.$$

*Beweis.* Zunächst ist die Potenzmenge $\mathcal{P}(\Omega)$ eine $\sigma$-Algebra die $\mathcal{E}$ enthält. Man zeigt dann, dass der Durchschnitt beliebig vieler $\sigma$-Algebren wieder eine $\sigma$-Algebra ist (Übung). Damit ist die rechte Seite in obiger Gleichung eine $\sigma$-Algebra, enthält $\mathcal{E}$ und ist offensichtlich die kleinste solche $\sigma$-Algebra. $\square$

*Bemerkung 3.12.* Wir werden sehr oft die folgende Schlussweise benutzen: Wenn $\mathcal{E} \subset \mathcal{A}$ für eine $\sigma$-Algebra $\mathcal{A}$, dann gilt $\sigma(\mathcal{E}) \subset \mathcal{A}$.

**Definition 3.13.** Die $\sigma$-Algebra $\sigma(\mathcal{E})$ aus dem vorhergehenden Satz bezeichnet man als die von $\mathcal{E}$ erzeugte $\sigma$-Algebra.

Ein Mengensystem $\mathcal{E}$ heißt Erzeugendensystem für eine $\sigma$-Algebra $\mathcal{A}$ wenn gilt $\sigma(\mathcal{E}) = \mathcal{A}$.

Eine wesentliche Schwierigkeit der Maßtheorie liegt darin, dass die Elemente der erzeugten $\sigma$-Algebra im Allgemeinen sehr schwer zu charakterisieren sind.

*Beispiel 3.14.* Sei $(X,d)$ ein metrischer Raum und $\mathcal{E} = \{O \subset X |\ O \text{ offen}\}$. Dann heißt $\sigma(\mathcal{E})$ die Borelsche $\sigma$-Algebra von $(X,d)$. Ein auf dieser $\sigma$-Algebra definiertes Maß heißt Borelmaß.

*Beispiel 3.15.* Sei $\mathcal{A}$ die Borelsche $\sigma$-Algebra auf $\mathbb{R}$. Dann wird $\mathcal{A}$ von $\mathcal{I}$ erzeugt.

*Beweis.* Sei $O \subset \mathbb{R}$ offen. Wir zeigen

$$O = \bigcup_{l,r\in\mathbb{Q},(l,r)\subset O} (l,r)\,. \tag{3.1}$$

Die Inklusion „$\supset$“ ist offensichtlich. Sei $x \in O$, dann existiert $\epsilon > 0$, so dass $(x-\epsilon, x+\epsilon) \subset O$. Da $\mathbb{Q} \subset \mathbb{R}$ dicht liegt, gibt es $r, l \in \mathbb{Q}$, so dass $x - \epsilon < l < x < r < x + \epsilon$, so dass $x \in (l,r) \subset O$. Damit ist $x$ in der rechten Seite von Gleichung (3.1) enthalten und da $x \in O$ beliebig war gilt auch die Inklusion „$\subset$“.

Gleichung (3.1) stellt $O$ als abzählbare Vereinigung von Intervallen dar. Falls $\mathcal{B}$ also eine $\sigma$-Algebra ist, die $\mathcal{E}$ enthält, dann enthält sie auch $O$.

Da $O$ beliebig war, enthält $\mathcal{B}$ also alle offenen Teilmengen von $\mathbb{R}$ und da es eine $\sigma$-Algebra ist gilt $\mathcal{A} \subset \mathcal{B}$. Damit gilt auch

$$\mathcal{A} \subset \bigcup_{\mathcal{B}\ \sigma\text{-Algebra},\mathcal{E}\subset\mathcal{B}} \mathcal{B} = \sigma(\mathcal{E}).$$

Der Beweis für die umgekehrte Inklusion geht analog unter Verwendung von

$$\{x\} = \bigcap_{n\in\mathbb{N}} \left(x - \frac{1}{n}, x + \frac{1}{n}\right).$$

□

*Bemerkung 3.16.* Weitere Erzeugendensysteme für die Borelsche $\sigma$-Algebra auf $\mathbb{R}$ sind

$$\begin{aligned} \{\,[x,y)|\ x,y \in \mathbb{R}, x<y\} \\ \{\,[x,y)|\ x,y \in \mathbb{Q}, x<y\} \\ \{\,[x,\infty)|\ x \in \mathbb{R}\} \\ \{\,(-\infty,x]|\ x \in \mathbb{R}\}\,. \end{aligned}$$

**Satz 3.17.** *Intervalle sind Lebesgue-messbar und damit sind auch alle Borelmengen Lebesgue-messbar.*

*Beweis.* Sei $A \subset \mathbb{R}$ ein beschränktes Intervall. Wir zerlegen $\mathbb{R}$ in drei disjunkte Intervalle

$$\mathbb{R} = I_l \cup A \cup I_r.$$

Für $I \in \mathcal{I}$ überzeugt man sich leicht, dass $I \cap I_l, I \cap A$ und $I \cap I_r$ beschränkte Intervalle sind und

$$l(I) = l(I \cap I_l) + l(I \cap A) + l(I \cap I_r)$$

gilt. Sei nun $E \in \mathcal{P}(\mathbb{R})$ und

$$E \subset \bigcup_{i\in\mathbb{N}} I_i$$

eine Überdeckung von $E$ durch Intervalle. Dann sind

$$E \cap A \subset \bigcup_{i\in\mathbb{N}} I_i \cap A \qquad E \cap A^c = \bigcup_{i\in\mathbb{N}} (I_i \cap I_l) \cup (I_i \cap I_r)$$

Überdeckungen von $E \cap A$ und $E \cap A^c$ durch abzählbar viele Intervalle und es gilt

$$\sum_{i\in\mathbb{N}} l(I_i) = \sum_{i\in\mathbb{N}} l(I_i \cap A) + \sum_{i\in\mathbb{N}} (l(I_i \cap I_l) + l(I_i \cap I_r)) \geq \lambda(E \cap A) + \lambda(E \cap A^c).$$

Da das für jede Überdeckung durch Intevalle von $E$ gilt, folgt

$$\lambda(E) \geq \lambda(E \cap A) + \lambda(E \cap A^c)$$

und da $E$ beliebig war, ist damit die Messbarkeit von $A$ gezeigt.

Damit enthält die $\sigma$-Algebra der Lebesgue-messbaren Mengen alle Intervalle und somit auch die von diesen erzeugte Borelsche $\sigma$-Algebra. □

*Bemerkung 3.18.* Die Borelsche $\sigma$-Algebra ist echt kleiner, als die $\sigma$-Algebra der Lebesgue-messbaren Mengen.

**Definition 3.19.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Eine Menge $A \subset \Omega$ heißt Nullmenge, wenn es $B \in \mathcal{A}$ gibt, so dass $A \subset B$ und $\mu(B) = 0$.

Der Maßraum $(\Omega, \mathcal{A}, \mu)$ heißt vollständig, wenn $\mathcal{A}$ alle Nullmengen enthält.

Sei $E$ eine von $\omega \in \Omega$ abhängige Eigenschaft. Wir sagen $E$ gilt fast überall (f.ü.), wenn die Menge der $\omega$ für die $E$ nicht erfüllt ist eine Nullmenge ist.

*Beispiel 3.20.* Der Maßraum $(\mathbb{R}, \mathcal{A}, \lambda)$ mit $\mathcal{A}$ der Menge der Lebesgue-messbaren Mengen ist vollständig. Der Maßraum $(\mathbb{R}, \mathcal{B}, \lambda)$ mit $\mathcal{B}$ der Borelschen $\sigma$-Algebra ist nicht vollständig.

Für eine Funktion $f : \mathbb{R} \to \mathbb{R}$ bedeutet $f > 0$ fast überall (bezüglich des Lebesgue-Maßes), dass

$$\lambda(\{x \in \mathbb{R} | \ f(x) \leq 0\}) = 0.$$

**Satz 3.21.** *Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum Für eine monoton wachsende Folge $(A_n) \subset \mathcal{A}$, das heißt $A_n \subset A_{n+1}$ für $n \in \mathbb{N}$, gilt*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n) = \sup_{n\in\mathbb{N}} \mu(A_n).$$

*Für eine monoton fallende Folge $(B_n) \subset \mathcal{A}$ gilt*

$$\mu\left(\bigcap_{n\in\mathbb{N}} B_n\right) = \lim_{n\to\infty} \mu(B_n) = \inf_{n\in\mathbb{N}} \mu(A_n),$$

*falls $\mu(B_n) < \infty$ für ein $n \in \mathbb{N}$.*

*Beweis.* Falls $\mu(A_n) = \infty$ für ein $n \in \mathbb{N}$, dann folgt die erste Gleichheit sofort. Sei also $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$. Setze $A_0 = \emptyset$ und $C_n = A_n \backslash A_{n-1}$ für $n \in \mathbb{N}$. Dann sind die $C_n$ paarweise disjunkt und daher gilt

$$\begin{aligned}\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mu\left(\bigcup_{n\in\mathbb{N}} C_n\right) &= \sum_{n\in\mathbb{N}} \mu(C_n) = \sum_{n\in\mathbb{N}} (\mu(A_n) - \mu(A_{n-1})) \\ &= \lim_{n\to\infty} \mu(A_n) - \mu(A_0) = \lim_{n\to\infty} \mu(A_n).\end{aligned}$$

Es sein nun $\mu(B_N)$ endlich (und damit auch $\mu(B_n)$ für alle $n \geq N$). Setze

$$B = \bigcap_{n\in\mathbb{N}} B_n.$$

Dann ist die Folge $D_n = B_N \backslash B_n$ für $n \geq N$ monoton wachsend und es gilt

$$\bigcup_{n=N}^{\infty} D_n = \bigcup_{n=N}^{\infty} B_N \cap B_n^c = B_N \cap \left(\bigcap_{n=N}^{\infty} B_n\right)^c = B_N \backslash B.$$

Nach dem eben gezeigten ist dann

$$\mu(B_N) - \mu(B) = \mu\left(B_N \backslash B\right) = \lim_{n\to\infty} \mu(B_N \backslash B_n) = \mu(B_N) - \lim_{n\to\infty} \mu(B_n).$$

Die übrigen Gleichheiten folgen dann aus der Monotonie der Folgen $\mu(A_n)$ beziehungsweise $\mu(B_n)$. $\square$

*Bemerkung 3.22.* Die Einschränkung $\mu(B_n) < \infty$ für ein $n \in \mathbb{N}$ ist notwendig, wie das Beispiel

$$0 = \lambda(\emptyset) = \lambda\left(\bigcap_{n\in\mathbb{N}} [n, \infty)\right) \neq \lim_{n\to\infty} \lambda([n, \infty)) = \infty$$

zeigt.

## 3.2 Integrale

**Definition 3.23.** Seien $(\Omega, \mathcal{A})$ und $(\Theta, \mathcal{B})$ messbare Räume und $f : \Omega \to \Theta$. Die Funktion $f$ heißt messbar, wenn für jedes $B \in \mathcal{B}$ gilt

$$f^{-1}(B) \in \mathcal{A}.$$

*Bemerkung 3.24.* (i) Im für uns praktisch wichtigsten Fall ist $\Theta = \mathbb{R}$ und $\mathcal{B}$ die Borelsche $\sigma$-Algebra.

(ii) An dieser Stelle müsste eigentlich ein Abschnitt über messbare Funktionen und ihre Eigenschaften folgen. Wir fassen stattdessen die wichtigsten Ergebnisse stichpunktartig zusammen.

(iii) Um zu zeigen, dass eine reellwertige Funktion $f$ messbar ist, reicht es aus zu zeigen, dass $f^{-1}((-\infty, a])$ für jedes $a \in \mathbb{R}$ messbar ist.

(iv) Stetige (auch stückweise stetige) Funktionen sind messbar. Charakteristische Funktionen messbarer Mengen sind messbar (siehe unten).

(v) Summen, Produkte und Quotienten (falls letztere existieren) von messbaren Funktionen sind messbar.

(vi) Falls $(f_n)$ eine Folge messbarer Funktionen ist, dann sind auch

$$\sup_{n\in\mathbb{N}} f_n$$
$$\limsup_{n\to\infty} f_n$$
$$\lim_{n\to\infty} f_n$$

messbar (und analog für inf und lim inf). Für diese entscheidenden Aussagen ist die „$\sigma$-Eigenschaft“ der $\sigma$-Algebren notwendig.

(vii) Zusammenfassend kann man sagen, dass die Funktionen mit denen man üblicherweise konfrontiert ist, messbar sind. Ähnlich wie bei nicht messbaren Mengen, muss man besondere Anstrengungen unternehmen, um nicht messbare Funktionen zu konstruieren.

Von nun an sei stets $(\Omega, \mathcal{A}, \mu)$ ein Maßraum.

**Definition 3.25.** Die Mengen $A_1, \ldots, A_n \in \mathcal{P}(\Omega)$ bilden einen Zerlegung von $\Omega$, wenn sie paarweise disjunkt sind und

$$\Omega = A_1 \cup \cdots \cup A_n$$

erfüllen.

**Definition 3.26.** Sei $A \subset \Omega$ eine Teilmenge, dann heißt

$$\begin{aligned} \mathbb{1}_A : \Omega &\to \mathbb{R} \\ \omega &\mapsto \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases} \end{aligned}$$

die charakteristische Funktion der Menge $A$. Wir nennen $A$ den Träger der charakteristischen Funktion.

Mit $\mathbf{X}(\Omega, \mathcal{A})$, oder nur $\mathbf{X}(\Omega)$ falls klar ist von welcher $\sigma$-Algebra die Rede ist, bezeichnen wir den folgenden, von charakteristischen Funktionen aufgespannten Vektorraum

$$\mathbf{X}(\Omega) = \left\{ \sum_{i=1}^{n} a_i \mathbb{1}_{A_i} \;\middle|\; n \in \mathbb{N}, A_i \in \mathcal{A}, a_i \in \mathbb{R}, i \in \{1, \ldots, n\} \right\}.$$

Die Elemente von $\mathbf{X}(\Omega, \mathcal{A})$ bezeichnen wir als einfache Funktionen.

Mit $\mathbf{X}^+(\Omega)$ bezeichnen wir die Menge der nichtnegativen einfachen Funktionen.

*Bemerkung 3.27.* (i) Für $A, B \in \mathcal{P}(\Omega)$ gilt

$$\mathbb{1}_A \mathbb{1}_B = \mathbb{1}_{A \cap B}.$$

(ii) Falls $A_1, \ldots, A_n$ eine Zerlegung von $\Omega$ bilden, dann gilt

$$\mathbb{1} := \mathbb{1}_\Omega = \sum_{i=1}^{n} \mathbb{1}_{A_i}.$$

(iii) Die Darstellung einer einfachen Funktion als Linearkombination charakteristischer Funktionen ist nicht eindeutig wie das folgende Beispiel zeigt:

$$\mathbb{1}_{[0,2]} + \mathbb{1}_{(2,3]} = \mathbb{1}_{[0,1)} + \mathbb{1}_{[1,3]}.$$

(iv) Man überzeugt sich leicht, dass einfache Funktionen nur endlich viele verschiedene Werte annehmen. Damit erhalten wir die sogenannte kanonische Darstellung

$$f = \sum_{y \in f(\Omega)} y \mathbb{1}_{f^{-1}(\{y\})}.$$

Die Träger der hier auftretenden charakteristischen Funktionen sind eine Zerlegung von $\Omega$ und die auftretenden Koeffizienten sind verschieden.

Wir werden nun mit der Definition von Integralen beginnen. Wir definieren das Integral zunächst für einfache Funktionen und erweitern die Definition dann Stück für Stück.

**Definition 3.28.** Sei $f \in \mathbf{X}^+(\Omega, \mathcal{A})$ und

$$f = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$$

die kanonische Darstellung von $f$. Definiere

$$\int f \mathrm{d}\mu = \sum_{i=1}^{n} a_i \mu(A_i).$$

*Bemerkung 3.29.* Man sieht leicht, dass $\int f \mathrm{d}\mu \geq 0$ für $f \in \mathbf{X}^+(\Omega, \mathcal{A})$.

Die Mengen $B_1, \dots, B_m \in \mathcal{A}$ seien eine Zerlegung von $\Omega$ und $b_1, \dots, b_m \in \mathbb{R}^+$ Koeffizienten, so dass

$$f = \sum_{j=1}^{m} b_j \mathbb{1}_{B_j}.$$

Sei

$$f = \sum_{j=1}^{n} a_i \mathbb{1}_{A_i}$$

die kanonische Darstellung von $f$. Dann gilt, falls $A_i \cap B_j \neq \emptyset$ für $x \in A_i \cap B_j$

$$f(x) = a_i = b_j,$$

und damit

$$\begin{aligned}\int f \mathrm{d}\mu &= \sum_{i=1}^{n} a_i \mu(A_i) = \sum_{i=1}^{n} a_i \mu\left(A_i \cap \bigcup_{j=1}^{m} B_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i \mu(A_i \cap B_j) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} b_j \mu(A_i \cap B_j) = \sum_{j=1}^{n} b_j \mu(B_j).\end{aligned}$$

**Satz 3.30.** *Es gilt*

$$\int (f + \alpha g) \mathrm{d}\mu = \int f \mathrm{d}\mu + \alpha \int g \mathrm{d}\mu$$

*für alle $f, g \in \mathbf{X}^+(\Omega, \mathcal{A})$ und alle $\alpha \geq 0$.*

*Falls $f \leq g$, so ist auch*

$$\int f \mathrm{d}\mu \leq \int g \mathrm{d}\mu.$$

*Beweis.* Seien

$$f = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i} \text{ und } g = \sum_{j=1}^{m} b_j \mathbb{1}_{B_j}$$

die kanonischen Darstellungen von $f$ und $g$. Dann können wir $f + \alpha g$ wie folgt darstellen

$$\begin{aligned} f + \alpha g &= \sum_{i=1}^{n} a_i \mathbb{1}_{A_i} + \alpha \sum_{j=1}^{m} b_i \mathbb{1}_{B_i} = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i} \sum_{j=1}^{m} \mathbb{1}_{B_j} + \alpha \sum_{j=1}^{m} b_i \mathbb{1}_{B_i} \sum_{i=1}^{n} \mathbb{1}_{A_i} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} (a_i + \alpha b_j) \mathbb{1}_{A_i \cap B_j} \end{aligned}$$

Die Mengen $A_i \cap B_j$, $i \in \{1, \dots, n\}$ und $j \in \{1, \dots, m\}$ bilden eine Zerlegung von $\Omega$ und es gilt nach der vorhergehenden Bemerkung

$$\begin{aligned} \int (f + \alpha g) \mathrm{d}\mu &= \sum_{i=1}^{n} \sum_{j=1}^{m} (a_i + \alpha b_j) \mu(A_i \cap B_j) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} a_i \mu(A_i \cap B_j) + \alpha \sum_{i=1}^{n} \sum_{j=1}^{m} b_j \mu(A_i \cap B_j) \\ &= \sum_{i=1}^{n} a_i \mu(A_i) + \sum_{j=1}^{m} b_j \mu(B_j) = \int f \mathrm{d}\mu + \alpha \int g \mathrm{d}\mu. \end{aligned}$$

Falls $f \leq g$, so ist $g - f \in \mathbf{X}^+(\Omega, \mathcal{A})$ und es gilt

$$\int g \mathrm{d}\mu = \int (g - f + f) \mathrm{d}\mu = \int (g - f) \mathrm{d}\mu + \int f \mathrm{d}\mu \geq \int f \mathrm{d}\mu. \qquad \square$$

**Definition 3.31.** Wir definieren $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ und $\overline{\mathbb{R}}^+ = \mathbb{R}^+ \cup \{\infty\}$ mit den folgenden Rechenregeln für alle $x, a \in \mathbb{R}$ und $a > 0$

$$\begin{aligned} x + \infty &= \infty + x = \infty & -\infty + x &= x - \infty = -\infty \\ \infty + \infty &= \infty & -\infty - \infty &= -\infty \\ 0 \cdot \infty &= \infty \cdot 0 = 0 & 0 \cdot (-\infty) &= -\infty \cdot 0 = 0 \\ a \cdot \infty &= (-a) \cdot (-\infty) = \infty & a \cdot (-\infty) &= (-a) \cdot \infty = -\infty \\ \infty \cdot \infty &= (-\infty) \cdot (-\infty) = \infty & \infty \cdot (-\infty) &= (-\infty) \cdot \infty = -\infty \\ -\infty &< x < \infty. \end{aligned}$$

*Bemerkung 3.32.* Für Funktionen $f, g : \Omega \to \overline{\mathbb{R}}$ und $y \in \mathbb{R}$ definieren wir

$$[f > y] := \{\omega \in \Omega |\ f(\omega) > y\}.$$

Analog werden auch Ausdrücke mit mehr als einer Funktion oder Bedingung definiert, zum Beispiel

$$\left[f = 0, g < g^2\right] = \left\{\omega \in \Omega \middle|\ f(\omega) = 0, g(\omega) < g(\omega)^2\right\}.$$

**Satz 3.33.** *Sei $f : \Omega \to \overline{\mathbb{R}}^+$ eine messbare Funktion. Dann gibt es eine monoton wachsende Folge einfacher Funktionen $f_n$, so dass $f_n \to f$ punktweise.*

*Beweis.* Definiere die folgenden, disjunkten Elemente von $\mathcal{A}$:

$$A_{j,n} = \left[\frac{j}{2^n} \leq f < \frac{j+1}{2^n}\right] = f^{-1}\left(\left[\frac{j}{2^n}, \frac{j+1}{2^n}\right)\right) \text{ und } B_n = f^{-1}([n, \infty])$$

und die Funktionen

$$f_n = \sum_{j=0}^{n2^n-1} \frac{j}{2^n} \mathbb{1}_{A_{j,n}} + n \mathbb{1}_{B_n}.$$

Sei nun $\omega \in \Omega$. Falls $n + 1 \leq f(\omega)$ so gilt $f_n(\omega) = n \leq n + 1 = f_{n+1}(\omega)$. Falls $n \leq f(\omega) < n + 1$ dann existiert $j \in \left\{n2^{n+1}, \ldots, (n+1)2^{n+1} - 1\right\}$, so dass $f_n(\omega) = n \leq \frac{j}{2^{n+1}} = f_{n+1}(\omega)$. Falls $f(\omega) < n$, so existiert $j \in \{0, \ldots, n2^n - 1\}$, so dass $f_n(\omega) = \frac{j}{2^n} = \frac{2j}{2^{n+1}} \leq f_{n+1}(\omega)$, da

$$\left[\frac{j}{2^n}, \frac{j+1}{2^n}\right) = \left[\frac{2j}{2^{n+1}}, \frac{2j+1}{2^{n+1}}\right) \cup \left[\frac{2j+1}{2^{n+1}}, \frac{2j+2}{2^{n+1}}\right),$$

die Funktionenfolge ist also monoton.

Man sieht leicht, dass stets $f_n \leq f$ gilt. Sei nun $\omega \in \Omega$. Falls $f(\omega) < \infty$, so gilt für alle $n > f(\omega)$, dass $f(\omega) - f_n(\omega) < \frac{1}{2^n}$. Da die rechte Seite für $n \to \infty$ verschwindet gilt $f_n(\omega) \to f(\omega)$. Falls $f(\omega) = \infty$, so gilt $f_n(\omega) = n$ für jedes $n \in \mathbb{N}$, also ebenfalls $f_n(\omega) \to f(\omega)$. □

**Lemma 3.34.** *Sei $g \in \mathbf{X}^+(\Omega, \mathcal{A})$ und $(f_n) \subset \mathbf{X}^+(\Omega, \mathcal{A})$ eine monoton wachsende Funktionenfolge, so dass*

$$g \leq \lim_{n\to\infty} f_n,$$

*dann gilt*

$$\int g \mathrm{d}\mu \leq \lim_{n\to\infty} \int f_n \mathrm{d}\mu.$$

*Beweis.* Sei

$$g = \sum_{i=1}^{m} a_i \mathbb{1}_{A_i}$$

die kanonische Darstellung von $g$ und $c > 1$ eine Konstante. Wir definieren die Mengen

$$B_n = [cf_n \geq g].$$

Da die Funktionenfolge $(f_n)$ monoton wachsend ist, ist auch die Folge $(B_n)$ monoton und da $\lim_{n\to\infty} f_n \geq g$ ist, gilt

$$\bigcup_{n\in\mathbb{N}} B_n = \Omega.$$

Dann gilt nach Satz 3.21

$$\mu(A_i) = \mu\left(\bigcup_{n\in\mathbb{N}} A_i \cap B_n\right) = \lim_{n\to\infty} \mu(A_i \cap B_n).$$

Für die Integrale impliziert dass

$$\begin{aligned}
\int g \mathrm{d}\mu = \sum_{i=1}^{m} a_i \mu(A_i) &= \lim_{n\to\infty} \sum_{i=1}^{m} a_i \mu(A_i \cap B_n) = \lim_{n\to\infty} \sum_{i=1}^{m} a_i \int \mathbb{1}_{A_i \cap B_n} \mathrm{d}\mu \\
&= \lim_{n\to\infty} \int \left(\sum_{i=1}^{m} a_i \mathbb{1}_{A_i}\right) \mathbb{1}_{B_n} \mathrm{d}\mu = \lim_{n\to\infty} \int g \mathbb{1}_{B_n} \mathrm{d}\mu \leq \lim_{n\to\infty} \int c f_n \mathbb{1}_{B_n} \mathrm{d}\mu \\
&\leq c \lim_{n\to\infty} \int f_n \mathrm{d}\mu.
\end{aligned}$$

Da die Konstante $c > 1$ beliebig war, folgt die gesuchte Ungleichung. □

**Satz 3.35.** *Sei $f : \Omega \to \overline{\mathbb{R}}^+$ eine messbare Funktion und $(f_n), (g_n) \subset \mathbf{X}^+(\Omega, \mathcal{A})$ monoton wachsende Funktionenfolgen, die punktweise gegen $f$ konvergieren. Dann gilt*

$$\lim_{n\to\infty} \int f_n \mathrm{d}\mu = \lim_{n\to\infty} \int g_n \mathrm{d}\mu.$$

*Insbesondere existieren die beiden Grenzwerte (eventuell als uneigentliche Grenzwerte).*

*Beweis.* Die Existenz der Grenzwerte folgt aus der Monotonie der Funktionenfolgen und des Integrals (Satz 3.30). Für festes $n \in \mathbb{N}$ gilt

$$g_n \leq f = \lim_{n\to\infty} f_n$$

und daher nach dem vorhergehenden Lemma

$$\int g_n \mathrm{d}\mu \leq \lim_{n\to\infty} \int f_n \mathrm{d}\mu.$$

Da das für alle $n \in \mathbb{N}$ gilt folgt

$$\lim_{n\to\infty} \int g_n \mathrm{d}\mu \leq \lim_{n\to\infty} \int f_n \mathrm{d}\mu.$$

Vertauschen wir nun $g_n$ und $f_n$, so erhalten wir auch die umgekehrte Ungleichung. □

**Definition 3.36.** Sei $f : \Omega \to \overline{\mathbb{R}}^+$ messbar und $(f_n) \subset \mathbf{X}^+(\Omega, \mathcal{A})$ eine monoton wachsende Folge die punktweise gegen $f$ konvergiert, dann definieren wir

$$\int f \mathrm{d}\mu = \lim_{n\to\infty} \int f_n \mathrm{d}\mu.$$

*Bemerkung 3.37.* Wir haben oben gezeigt, dass diese Definition eindeutig ist (also nicht von der konkreten Folge $(f_n)$ abhängt. Das so definierte Integral kann, wie bei einfachen Funktionen, den Wert $\infty$ annehmen, aber es können keine unbestimmten Ausdrücke (zum Beispiel $\infty - \infty$) entstehen, da alle auftretenden Zahlen nichtnegativ sind.

*Beispiel 3.38.* Wir wollen das Lebesgue-Integral der Funktion $f(y) = y \cdot \mathbb{1}_{[0,x]}(y)$ berechnen. Dazu überzeugt man sich, dass die Funktionenfolge

$$f_n = \sum_{k=0}^{2^n-1} \frac{kx}{2^n} \mathbb{1}_{\left[k\frac{x}{2^n},(k+1)\frac{x}{2^n}\right)} + x\mathbb{1}_{\{x\}}$$

monoton wachsend ist und gegen $f$ konvergiert. Damit gilt

$$\int f_n \mathrm{d}\lambda = \sum_{k=0}^{2^n-1} \frac{kx}{2^n} \mu\left(\left[k\frac{x}{2^n}, (k+1)\frac{x}{2^n}\right)\right) = \frac{x}{2^n}\frac{x}{2^n} \sum_{k=0}^{2^n-1} k = \frac{x^2}{2^{2n}} \frac{(2^n-1)2^n}{2}$$

also

$$\int f \mathrm{d}\lambda = \lim_{n\to\infty} \int f_n \mathrm{d}\lambda = \frac{x^2}{2} \lim_{n\to\infty} \frac{2^n - 1}{2^n} = \frac{x^2}{2}.$$

**Satz 3.39.** *Es gilt*

$$\int (f + \alpha g) \mathrm{d}\mu = \int f \mathrm{d}\mu + \alpha \int g \mathrm{d}\mu$$

*für alle $f, g : \Omega \to \overline{\mathbb{R}}^+$ und alle $\alpha \geq 0$.*

*Falls $f \leq g$, so ist auch*

$$\int f \mathrm{d}\mu \leq \int g \mathrm{d}\mu.$$

*Beweis.* Seien $(f_n), (g_n) \subset \mathbf{X}^+(\Omega, \mathcal{A})$ monoton wachsende Folgen, so dass $\lim_{n\to\infty} f_n = f$ und $\lim_{n\to\infty} g_n = g$. Die positive Linearität folgt sofort aus der Definition, da die Folge $f_n + \alpha g_n$ monoton wachsend ist und gegen $f + \alpha g$ konvergiert.

Die Monotonie folgt ähnlich wie für einfache Funktionen. Ist für beliebiges $h : \Omega \to \overline{\mathbb{R}}^+$ das Integral der Grenzwert einer monoton wachsenden Folge nichtnegativer Zahlen, es gilt also

$$\int h \mathrm{d}\mu \geq 0.$$

Schließlich folgt aus der Linearität

$$\int g \mathrm{d}\mu = \int (g - f) \mathrm{d}\mu + \int f \mathrm{d}\mu \geq \int f \mathrm{d}\mu.$$

□

**Satz 3.40.** *Sei* $f : \Omega \to \overline{\mathbb{R}}^+$ *messbar. Dann ist* $\int f \mathrm{d}\mu = 0$ *genau dann wenn* $[f > 0]$ *eine Nullmenge ist.*

*Beweis.* Setze $A = [f > 0]$. Sei zunächst $A$ Nullmenge, dann konvergiert die Funktionenfolge $n \cdot \mathbb{1}_A$ monoton wachsend gegen $g = \infty \cdot \mathbb{1}_A$ und wir erhalten

$$\int g \mathrm{d}\mu = \lim_{n\to\infty} \int n \cdot \mathbb{1}_A \mathrm{d}\mu = 0.$$

Da $f \leq g$ folgt aus der Monotonie auch $\int f \mathrm{d}\mu = 0$.

Sei nun $\int f \mathrm{d}\mu = 0$. Dann gilt für jedes $n$

$$\frac{1}{n} \cdot \mathbb{1}_{[f \geq \frac{1}{n}]} \leq f,$$

also wegen der Monotonie des Integrals

$$\int \frac{1}{n} \cdot \mathbb{1}_{[f \geq \frac{1}{n}]} \mathrm{d}\mu = \frac{1}{n} \mu \left( \left[ f \geq \frac{1}{n} \right] \right) \leq \int f \mathrm{d}\mu = 0.$$

Daraus folgt, dass $\left[f \geq \frac{1}{n}\right]$ eine Nullmenge ist, für jedes $n \in \mathbb{N}$. Darüber hinaus wachsen die Mengen monoton mit $n$ und so folgt aus Satz 3.21

$$\mu([f > 0]) = \mu \left( \bigcup_{n\in\mathbb{N}} \left[ f \geq \frac{1}{n} \right] \right) = \lim_{n\to\infty} \mu \left( \left[ f \geq \frac{1}{n} \right] \right) = 0. \qquad \square$$

*Beispiel 3.41.* Da das Integral über einen Grenzwertprozess definiert ist, muss man beim Vertauschen mit anderen Grenzwertprozessen vorsichtig sein. Zum Beispiel konvergiert die Funktionenfolge

$$n \cdot \mathbb{1}_{\left[\frac{1}{n}, \frac{1}{n}\right]}$$

punktweise gegen $\infty \cdot \mathbb{1}_{\{0\}}$, die Folge ist aber nicht monoton. Für die Integrale gilt nach dem vorhergehenden Satz

$$\lim_{n\to\infty} \int n \cdot \mathbb{1}_{\left[\frac{1}{n}, \frac{1}{n}\right]} \mathrm{d}\mu = \lim_{n\to\infty} n \frac{2}{n} = 2 \neq 0 = \int \infty \cdot \mathbb{1}_{\{0\}} \mathrm{d}\mu.$$

**Definition 3.42.** Sei $f : \Omega \to \overline{\mathbb{R}}$ messbar. Wir definieren den positiven Teil $f_+$ von $f$ sowie den negativen Teil $f_-$ von $f$ wie folgt:

$$f_+ = f \cdot \mathbb{1}_{[f \geq 0]} \text{ und } f_- = -f \cdot \mathbb{1}_{[f \leq 0]}.$$

*Bemerkung 3.43.* Man überzeugt sich leicht, dass dann die Gleichungen

$$f = f_+ - f_-$$
$$|f| = f_+ + f_-$$

gelten.

**Definition 3.44.** Sei $f : \Omega \to \overline{\mathbb{R}}$ messbar. Die Funktion $f$ heißt integrierbar, falls

$$\int |f| \, \mathrm{d}\mu < \infty.$$

Für integrierbare Funktionen definieren wir

$$\int f \mathrm{d}\mu := \int f_+ \mathrm{d}\mu - \int f_- \mathrm{d}\mu.$$

Mit $\mathscr{L}(\Omega, \mathcal{A})$ bezeichnen wir den Raum der integrierbaren Funktionen.

*Bemerkung 3.45.* Aus der positiven Linearität des Integrals (für nichtnegative Funktionen) folgt

$$\int |f| \, \mathrm{d}\mu = \int f_+ \mathrm{d}\mu + \int f_- \mathrm{d}\mu.$$

Die Funktion $f$ ist also integrierbar, genau dann wenn sowohl $\int f_+ \mathrm{d}\mu$ als auch $\int f_- \mathrm{d}\mu$ endlich sind. Insbesondere ist $\int f \mathrm{d}\mu$ für integrierbare Funktionen tatsächlich definiert und liegt in $\mathbb{R}$.

**Satz 3.46.** *Der Raum $\mathscr{L}(\Omega, \mathcal{A})$ der integrierbaren Funktionen ist ein Vektorraum und das Integral ist ein lineares Funktional.*

*Für integrierbare Funktionen $f \le g$ folgt*

$$\int f \mathrm{d}\mu \le \int g \mathrm{d}\mu.$$

*Es gilt die verallgemeinerte Dreiecksungleichung*

$$\left| \int f \mathrm{d}\mu \right| \le \int |f| \, \mathrm{d}\mu.$$

*Beweis.* Seien $f, g$ integrierbar und $\alpha \in \mathbb{R}$. Dann folgt aus der Monotonie und positiven Linearität des Integrals (für nichtnegative Funktionen)

$$\int |f + \alpha g| \, \mathrm{d}\mu \le \int (|f| + |\alpha| \, |g|) \, \mathrm{d}\mu = \int |f| \, \mathrm{d}\mu + \alpha \int |g| \, \mathrm{d}\mu < \infty,$$

also die Integrierbarkeit von $f + \alpha g$.

Wir zeigen nun zunächst, dass für beliebige nichtnegative integrierbare Funktionen $g, h$ gilt

$$\int (g - h) \mathrm{d}\mu = \int g \mathrm{d}\mu - \int h \mathrm{d}\mu.$$

Sei dazu $f = g - h$, dann gilt

$$f = g - h = f_+ - f_- \text{ also } g + f_- = h + f_+.$$

Damit folgt wiederum mittels der positiven Linearität

$$\int g \mathrm{d}\mu + \int f_- \mathrm{d}\mu = \int (g + f_-) \mathrm{d}\mu = \int (h + f_+) \mathrm{d}\mu = \int h \mathrm{d}\mu + \int f_+ \mathrm{d}\mu,$$

also

$$\int f \mathrm{d}\mu = \int f_+ \mathrm{d}\mu - \int f_- \mathrm{d}\mu = \int g \mathrm{d}\mu - \int h \mathrm{d}\mu.$$

Seien nun $f, g$ integrierbar. Falls $\alpha \in \mathbb{R}^+$, dann sind $f_+ + \alpha g_+$ und $f_- + \alpha g_-$ nichtnegative, integrierbare Funktionen und es gilt

$$f + g = (f_+ + \alpha g_+) - (f_- + \alpha g_-)$$

Gemäß dem oben gezeigten gilt dann

$$\begin{aligned} \int (f + g) \mathrm{d}\mu &= \int (f_+ + \alpha g_+) \mathrm{d}\mu - \int (f_- + \alpha g_-) \mathrm{d}\mu \\ &= \int f_+ \mathrm{d}\mu - \int f_- \mathrm{d}\mu + \alpha \left( \int g_+ \mathrm{d}\mu - \int g_- \mathrm{d}\mu \right) = \int f \mathrm{d}\mu + \alpha \int g \mathrm{d}\mu. \end{aligned}$$

Analog sind für $\alpha < 0$ die Funktionen $f_+ - \alpha g_-$ und $f_- - \alpha g_+$ nichtnegativ und integrierbar und es gilt

$$\int (f + g) \mathrm{d}\mu = \int (f_+ - \alpha g_-) \mathrm{d}\mu - \int (f_- - \alpha g_+) \mathrm{d}\mu = \int f \mathrm{d}\mu + \alpha \int g \mathrm{d}\mu.$$

Falls nun $f \leq g$ ist, für integrierbare $f, g$, dann ist $g - f$ nichtnegativ und daher

$$\int (g - f) \mathrm{d}\mu \geq 0.$$

Daraus folgt mit der eben bewiesenen Linearität

$$\int g \mathrm{d}\mu \geq \int f \mathrm{d}\mu.$$

Es gilt offensichtlich

$$-|f| = -f_+ - f_- \leq f_+ - f_- \leq f_+ + f_- = |f|$$

und daraus folgt mit der Monotonie

$$-\int |f| \mathrm{d}\mu \leq \int f \mathrm{d}\mu \leq \int |f| \mathrm{d}\mu.$$

Diese Ungleichungen können zu

$$\left| \int f \mathrm{d}\mu \right| \leq \int |f| \mathrm{d}\mu$$

zusammengefasst werden. □

*Bemerkung 3.47.* Häufig wird nicht über den gesamten Raum $\Omega$ integriert. Wir hatten bereits erwähnt (Beispiel 3.4), dass die $\sigma$-Algebra und das Maß sich auf messbare Teilmengen $A \subset \Omega$ einschränken lassen. Für $f : A \to \overline{\mathbb{R}}$ integrierbar ist dann das Integral

$$\int f \mathrm{d}\, \mu|_A$$

definiert.

Für $g : \Omega \to \overline{\mathbb{R}}$ eine messbare Funktion und $A \subset \mathcal{A}$, erhält man

$$\int g|_A \,\mathrm{d}\, \mu|_A = \int g \mathbb{1}_A \mathrm{d}\mu.$$

Dabei ist zu beachten, dass $g\mathbb{1}_A$ integrierbar sein kann, auch dann wenn $g$ es nicht ist.

Wir hätten also auch umgekehrt oben $f$ zu einer Funktion

$$\begin{aligned} \tilde{f} &: \Omega \to \overline{\mathbb{R}} \\ \omega &\mapsto \begin{cases} f(\omega) & \omega \in A \\ 0 & \omega \notin A \end{cases} \end{aligned}$$

auf ganz $\Omega$ fortsetzen können um zu erhalten

$$\int f \mathrm{d}\, \mu|_A = \int \tilde{f} \mathrm{d}\mu.$$

Wir werden in diesen Fällen, die Einschränkung des Maßes beziehungsweise Fortsetzung der Funktion nicht explizit hervorheben, und

$$\int_A f \mathrm{d}\mu \text{ beziehungsweise } \int_A g \mathrm{d}\mu$$

schreiben.

*Beispiel 3.48.* Sei $\Omega = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$ und $\mu$ das Zählmaß (Beispiel 3.4). Dann gilt $\mu(\{n\}) = 1$. Für eine nichtnegative Funktion $f : \mathbb{N} \to \overline{\mathbb{R}}^+$ ist dann

$$f_N = f \mathbb{1}_{\{1,\dots,N\}} = \sum_{n=1}^{N} f(n) \mathbb{1}_{\{n\}}$$

eine monoton wachsende Folge einfacher Funktionen und es gilt

$$\int f \mathrm{d}\mu = \lim_{N \to \infty} \sum_{n=1}^{n} f(n) \mu(\{n\}) = \sum_{n=1}^{\infty} f(n).$$

Eine beliebige Funktion $f : \mathbb{N} \to \overline{\mathbb{R}}^+$ ist dann integrierbar genau dann wenn

$$\int |f| \,\mathrm{d}\mu = \sum_{n=1}^{\infty} |f(n)| < \infty$$

und für solche Funktionen gilt

$$\int f \mathrm{d}\mu = \sum_{n=1}^{\infty} f(n).$$

Das Integral ist also eine echte Verallgemeinerung des Wertes einer Reihe.

*Bemerkung 3.49.* Sei $f : \Omega \to \overline{\mathbb{R}}$ eine messbare Funktion, so dass $[f \neq 0]$ eine Nullmenge ist. Dann sind auch $[f_+ > 0]$ und $[f_- > 0]$ Nullmengen und es gilt nach Satz 3.40

$$\int f \mathrm{d}\mu = \int f_+ \mathrm{d}\mu - \int f_- \mathrm{d}\mu = 0.$$

Sei nun $g : \Omega \to \overline{\mathbb{R}}$ eine weitere messbare Funktion, dann gilt

$$\int (g + f) \mathrm{d}\mu = \int g \mathrm{d}\mu + \int f \mathrm{d}\mu = \int g \mathrm{d}\mu.$$

Wir können $g$ also auf einer Nullmenge beliebig verändern, ohne den Wert des Integrals zu beeinflussen. Davon werden wir in Zukunft häufig Gebrauch machen.

## 3.3 Integrale über den reellen Zahlen

Wir betrachten nun Integration von Funktionen von $\mathbb{R}$ nach $\mathbb{R}$.

**Definition 3.50.** Seien $a, b \in \overline{\mathbb{R}}$ mit $a < b$ und $f : [a, b] \to \overline{\mathbb{R}}$ integrierbar. Mit dem Lebesgue-Maß $\lambda$ setzen wir

$$\int_a^b f(x)\mathrm{d}x := \int_{(a,b)} f \mathrm{d}\lambda.$$

Außerdem definieren wir

$$\int_b^a f(x)\mathrm{d}x := -\int_a^b f(x)\mathrm{d}x.$$

*Bemerkung 3.51.* Man rechnet leicht nach, dass mit diesen Definitionen für beliebiges $c \in \mathbb{R}$ gilt

$$\int_a^b f(x)\mathrm{d}x = \int_a^c f(x)\mathrm{d}x + \int_c^b f(x)\mathrm{d}x$$

**Theorem 3.52 (Mittelwertsatz der Integralrechnung).** *Seien $a, b \in \mathbb{R}$, $a < b$ und $f, g : \mathbb{R} \to \mathbb{R}$ stetig und $g(x) \geq 0$ auf $[a, b]$. Dann existiert $\xi \in (a, b)$, so dass*

$$\int_a^b f(x)g(x)\mathrm{d}x = f(\xi) \int_a^b g(x)\mathrm{d}x.$$

*Insbesondere existiert $\eta \in (a, b)$ so, dass*

$$\int_a^b f(x)\mathrm{d}x = f(\eta)(b - a).$$

*Beweis.* Da $f$ stetig ist und $[a, b]$ beschränkt und abgeschlossen, existieren nach Korollar 1.66

$$c = \min_{a \leq x \leq b} f(x) \text{ und } C = \max_{a \leq x \leq b} f(x)$$

und Punkte $x_m, x_M \in [a, b]$, so dass $f(x_m) = c$ und $f(x_M) = C$. Setze $\tilde{a} = \min\{x_m, x_M\}$ und $\tilde{b} = \max\{x_m, x_M\}$. Es gilt dann für alle $x \in [a, b]$

$$cg(x) \leq f(x)g(x) \leq Cg(x).$$

Aus der Monotonie des Integrals folgt mit der Abkürzung $I = \int_a^b g(x)\mathrm{d}x$

$$cI \leq \int_a^b f(x)g(x)\mathrm{d}x \leq CI.$$

Falls $I = 0$ ist, oder $c = C$, so folgt bereits die erste Behauptung. Andernfalls ist $\tilde{a} < \tilde{b}$ und auf Grund des Zwischenwertsatzes (Theorem 1.63) existiert $\xi \in \left(\tilde{a}, \tilde{b}\right) \subset (a, b)$, so dass

$$f(\xi) = \frac{1}{I} \int_a^b f(x)g(x)\mathrm{d}x.$$

Damit ist die erste Behauptung des Satzes gezeigt.

Die zweite Behauptung folgt als Spezialfall für $g \equiv 1$. □

**Definition 3.53.** Seien $a, b \in \mathbb{R}$, $a < b$ und $f : (a, b) \to \mathbb{R}$ eine Funktion. Dann heißt $F : [a, b] \to \mathbb{R}$ Stammfunktion von $f$, wenn es stetig und auf $(a, b)$ differenzierbar ist und $F' = f$ auf $(a, b)$ gilt.

*Bemerkung 3.54.* Seien $F$ und $G$ Stammfunktionen von $f$, dann gilt auf $(a, b)$

$$(F - G)'(x) = F'(x) - G'(x) = f(x) - f(x) = 0.$$

und deshalb existiert eine Konstante $c \in \mathbb{R}$, so dass $F = G + c$ auf $(a, b)$ (siehe Korollar 2.9). Wegen der Stetigkeit von $F$ und $G$ gilt die Gleichheit dann auch auf $[a, b]$.

**Theorem 3.55 (Hauptsatz der Integral und Differentialrechnung).** *Seien $a, b \in \mathbb{R}$, $a < b$ und $f : [a, b] \to \mathbb{R}$ stetig. Dann ist für beliebiges $x_0 \in [a, b]$ die Funktion*

$$x \mapsto \int_{x_0}^x f(y)\mathrm{d}y$$

*eine Stammfunktion von $f$.*

*Sei $G$ eine beliebige Stammfunktion von $f$, dann gilt*

$$\int_a^b f(y)\mathrm{d}y = G(b) - G(a).$$

*Beweis.* Wir definieren

$$F(x) = \int_{x_0}^x f(y)\mathrm{d}y.$$

Für festes $x \in [a, b]$ und für $h$ so dass $x + h \in [a, b]$ gilt dann

$$F(x + h) - F(x) = \int_{x_0}^{x+h} f(y)\mathrm{d}y - \int_{x_0}^x f(y)\mathrm{d}y = \int_x^{x+h} f(y)\mathrm{d}y = f(\xi_h)h$$

wobei wir in der letzen Umformung den Mittelwertsatz verwendet haben um $\xi_h$ zwischen $x$ und $x+h$ zu finden. Für $h \to 0$ konvergiert $\xi_h$ gegen $x$ und $f(\xi_h)$ gegen $f(x)$ (Stetigkeit von $f$). Daraus folgt für beliebiges $x \in [a, b]$ die Stetigkeit von $F$ im Punkt $x$, da

$$\lim_{h \to 0}(F(x + h) - F(x)) = 0.$$

Für $x \in (a, b)$ gilt sogar

$$\lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \to 0} f(\xi_h) = f(x).$$

Damit ist $F$ differenzierbar in $x$ mit $F'(x) = f(x)$.

Sei schließlich $G$ eine beliebige Stammfunktion von $F$, dann existiert $c \in \mathbb{R}$ so, dass $G = F + c$ und es gilt

$$\int_a^b f(y)\mathrm{d}y = \int_a^{x_0} f(y)\mathrm{d}y + \int_{x_0}^b f(y)\mathrm{d}y = F(b) - F(a) = G(b) - G(a). \qquad \square$$

*Bemerkung 3.56.* Häufig wird die Stammfunktion von $f$, so sie denn existiert, mit

$$\int f(x)\mathrm{d}x$$

bezeichnet. Diese Bezeichnung erschließt sich im Lichte des Hauptsatzes, ist aber insofern etwas unglücklich, dass „die Stammfunktion“ ja nur bis auf eine beliebige additive Konstante genau bestimmt ist, und der obige Ausdruck also keine Funktion, sondern eine ganze Schaar von Funktionen bezeichnet.

**Korollar 3.57 (Partielle Integration).** *Seien $a, b \in \mathbb{R}$, $a < b$ und $f, g : [a, b] \to \mathbb{R}$ stetig differenzierbar. Dann gilt*

$$\int f'(x)g(x)\mathrm{d}x = f(x)g(x) - \int f(x)g'(x)\mathrm{d}x.$$

*Für die Werte von Integralen ergibt sich entsprechend*

$$\int_a^b f'(x)g(x)\mathrm{d}x = f(b)g(b) - f(a)g(a) - \int_a^b f(x)g'(x)\mathrm{d}x.$$

*Beweis.* Sei $H : [a, b]$ eine Stammfunktion von $fg'$, dann ist $fg - H$ stetig und auf $(a, b)$ differenzierbar und es gilt auf Grund der Produktregel

$$(fg - H)' = f'g + fg' - H' = f'g.$$

Die Funktione $fg - H$ ist also eine Stammfunktion von $f'g$. Dies entspricht der ersten Aussage des Korollars.

Für das Integral folgt dann mit dem Hauptsatz

$$\begin{aligned}\int_a^b f'(x)g(x)\mathrm{d}x &= (fg - H)(b) - (fg - H)(a) = f(b)g(b) - f(a)g(a) - (H(b) - H(a)) \\ &= f(b)g(b) - f(a)g(a) - \int_a^b f(x)g'(x)\mathrm{d}x. \qquad \square\end{aligned}$$

**Korollar 3.58 (Substitutionsregel).** *Seien* $a, b \in \mathbb{R}$, $a < b$ *und* $g : [a, b] \to \mathbb{R}$ *stetig differenzierbar. Setze* $\xi = \min_{x\in[a,b]} g(x)$ *und* $\eta = \max_{x\in[a,b]} g(x)$. *Sei* $f : [\xi, \eta] \to \mathbb{R}$ *stetig. Dann gilt*

$$\int f(g(x))g'(x)\mathrm{d}x = \int f(y)\mathrm{d}y$$

*für* $y = g(x)$.

*Für das bestimmte Integral bedeutet das*

$$\int_a^b f(g(x))g'(x)\mathrm{d}x = \int_{g(a)}^{g(b)} f(y)\mathrm{d}y.$$

*Beweis.* Wir setzen $f$ zunächst zu einer stetigen Funktion $\tilde{f} : [\xi - 1, \eta + 1]$ fort. (Warum geht das immer?) Sei $F$ eine Stammfunktion von $\tilde{f}$, dann ist $F \circ g$ stetig und es gilt auf Grund der Kettenregel

$$(F \circ g)'(x) = F'(g(x))g'(x) = \tilde{f}(g(x))g'(x) = f(g(x))g'(x)$$

für alle $x \in (a, b)$ also ist $F \circ g$ eine Stammfunktion von $(f \circ g)g'$.

Für das Integral erhalten wir

$$\begin{aligned}\int_a^b f(g(x))g'(x)\mathrm{d}x &= (F \circ g)(b) - (F \circ g)(a) = F(g(b)) - F(g(a)) \\ &= \int_{g(a)}^{g(b)} f(y)\mathrm{d}y.\end{aligned}$$

□

# 4 Gewöhnliche Differenzialgleichungen

## 4.1 Lösungsmethoden

**Definition 4.1 (Gewöhnliche Differenzialgleichung).** Eine gewöhnliche Differenzialgleichung (GDGl) ist eine Gleichung der Form

$$F(x, y, y', \ldots, y^{(n)}) = 0$$

für ein $F : \mathbb{R}^{n+2} \to \mathbb{R}$. Dabei ist $n$ die Ordnung der Differenzialgleichung.

Sei $I$ ein offenes Intervall. Eine Funktion $y : I \to \mathbb{R}$ heißt Lösung dieser Differenzialgleichung wenn sie mindestens $n$ mal differenzierbar ist und wenn gilt

$$F(x, y(x), y'(x), \ldots, y^{(n)}(x)) = 0$$

für alle $x \in I$.

*Beispiel 4.2.* (i) Die Differenzialgleichung

$$y'' = -\frac{1}{y^2} \text{ beziehungsweise } y'' + \frac{1}{y^2} = 0$$

beschreibt die Bewegung einer Probemasse im Gravitationsfeld außerhalb einer radialsymmetrischen Massenverteilung entlang einer Geraden durch das Zentrum der Massenverteilung.

(ii) Die Differenzialgleichung

$$y'' = -\sin(y) \text{ oder } y'' + \sin(y) = 0$$

beschreibt die Bewegung eines mathematischen Pendels.

*Bemerkung 4.3.* (i) Auch wenn in unserer Definition die rechte Seite der Gleichung stets 0 ist, also alle Terme auf einer Seite stehen, werden wir natürlich auch mit anderen Differenzialgleichungen arbeiten, die sich immer einfach in diese Form überführen lassen.

(ii) Häufig ist $F$ auch nur auf einer Teilmenge von $\mathbb{R}^{n+2}$ definiert und die Definitionen gelten sinngemäß.

(iii) Das Lösen von (selbst einfach aussehenden) Differenzialgleichungen ist im Allgemeinen schwierig. Wir werden Verfahren für einige Arten von GDGLen erarbeiten. In vielen (auch praktisch relevanten) Fällen erfordert jedoch jede Gleichung eine eigene Theorie.

(iv) Auch wenn sich eine Differenzialgleichung nicht explizit lösen lässt, kann man dennoch versuchen die folgenden Aspekte zu untersuchen:

- Existenz von Lösungen
- Eindeutigkeit der Lösungen

- Stabilität
- Näherungsweise und numerische Lösungen

(v) In der obigen Definition spielt $y$ zwei verschiedene Rollen. Es ist die (funktionswertige) Variable der Differenzialgleichung sowie der Name der Lösungsfunktion. Diese Doppeldeutigkeit ist etwas unglücklich aber weit verbreitet.

*Beispiel 4.4.* (i) Sei $I$ ein offenes Intervall und $f : I \to \mathbb{R}$ stetig. Dann sind die Lösungen der GDGl

$$y' = f(x)$$

genau die Stammfunktionen von $f$.

Sei nun $x_0 \in I$ dann folgt aus dem Hauptsatz (Theorem 3.55) und aus Bemerkung 3.54, dass die Lösungen genau die Form

$$y(x) = \int_{x_0}^{x} f(t)\mathrm{d}t + c$$

für beliebiges $c \in \mathbb{R}$ haben.

(ii) Für die Gleichung $y' = y$ überzeugt man sich leicht, dass $x \mapsto ce^x$ für jedes $c \in \mathbb{R}$ eine Lösung auf ganz $R$ ist. Um zu zeigen, dass jede Lösung diese Form hat, sei $y$ eine beliebige Lösung der Gleichung und betrachte $u(x) = y(x)e^{-x}$. Dann gilt

$$u'(x) = y'(x)e^{-x} - y(x)e^{-x} = 0$$

woraus $u(x) = c$ für eine Konstante $c$ folgt.

Wie in den obigen Fällen haben Differenzialgleichungen üblicherweise ganze Schaaren von Funktionen als Lösungen. Es können also zusätzliche Bedingungen an die Lösungen gestellt werden.

**Definition 4.5 (Anfangswertproblem).** Seien $y_0, \dots, y_{n-1} \in \mathbb{R}$ und $F : \mathbb{R}^{n+2} \to \mathbb{K}$. Das System von Gleichungen

$$\begin{aligned} F(x, y, y', \dots, y^{(n)}) &= 0 \\ y(0) &= y_0 \\ &\vdots \\ y^{(n-1)}(0) &= y_{n-1} \end{aligned}$$

heißt Anfangswertproblem.

*Beispiel 4.6.* Das Anfangswertproblem (AWP)

$$\begin{aligned} y'' &= -\frac{1}{y^2} \\ y(0) &= y_0 \\ y'(0) &= y_1 \end{aligned}$$

legt für die Bewegung des Massenpunktes den Anfangsort $y_0$ sowie die Anfangsgeschwindigkeit $y_1$ fest.

*Beispiel 4.7.* Sei $I$ ein offenes Intervall, dass die 0 enthält. Wir betrachten das Anfangswertproblem

$$\begin{aligned} y' &= -y^2 \\ y(0) &= 1 \end{aligned}$$

Angenommen $y : I \to \mathbb{R}$ ist eine Lösung und es gilt $y(x) > 0$ für $x \in I$. Integrieren wir die Gleichung

$$1 = -\frac{1}{y(t)^2} y'(t)$$

von 0 bis $x$, so erhalten wir

$$x = -\int_0^x \frac{1}{y(t)^2} y'(t) \mathrm{d}t = -\int_1^{y(x)} \frac{1}{y^2} \mathrm{d}y = \frac{1}{y(x)} - 1,$$

also

$$y(x) = \frac{1}{x+1}.$$

Man überprüft leicht, dass das tatsächlich eine Lösung des Anfangswertproblems ist. Wir haben hier vorausgesetzt, dass $y(x)$ positiv ist, und $I$ muss die 0 enthalten. Daher ist $I = (-1, \infty)$ das größte mögliche Intervall auf dem die Lösung definiert ist.

Auf ähnliche Weise kann verfahren werden, wenn das Anfangswertproblem die Form

$$\begin{aligned} y' &= f(y)g(x) \\ y(x_0) &= y_0 \end{aligned}$$

hat. Formal können wir wie oben schreiben

$$\int_{y(x_0)}^{y(x)} \frac{1}{f(y)} \mathrm{d}y = \int_{x_0}^{x} g(x).$$

Es stellt sich aber die Frage, ob die Integrale existieren und ob die entstehende Gleichung nach $y(x)$ aufgelöst werden kann.

**Satz 4.8 (Trennung der Variablen).** *Seien $I, J$ offene Intervalle, $f : I \to \mathbb{R}$ und $g : J \to \mathbb{R}$ stetig mit $0 \notin f(I)$. Seien $x_0 \in J$ und $y_0 \in I$.*

*Dann existiert ein offenes Intervall $I_2 \subset J$, so dass das Anfangswertproblem*

$$\begin{aligned} y' &= f(y)g(x) \\ y(x_0) &= y_0 \end{aligned}$$

*genau eine Lösung $y$ mit Definitionsbereich $I_2$ besizt. Setze*

$$F(y) = \int_{y_0}^{y} \frac{1}{f(t)} \mathrm{d}t.$$

*Dann ist $y : I_2 \to I$ eindeutig durch*

$$F(y(x)) = \int_{x_0}^{x} g(t)\mathrm{d}t$$

*bestimmt.*

*Beweis.* Nach Voraussetzung hat $f$ keine Nullstellen. Wir nehmen hier $f > 0$ an, der Fall $f < 0$ wir analog behandelt. Da die Ableitung von $F'(y) = \frac{1}{f(y)}$ positiv ist, ist $F$ steng monoton wachsend (Korollar 2.10) also insbesondere injektiv. Es besitzt also eine Umkehrfunktion $H : F(I) \to I$ und nach dem Satz über die Umkehrfunktion (Korollar 2.70) ist diese stetig differenzierbar und es gilt

$$H'(z) = \frac{1}{F'(H(z))}$$

für alle $z \in F(I)$.

Setze jetzt

$$G(x) := \int_{x_0}^{x} g(t)\mathrm{d}t,$$

für $x \in J$. Das Bild $F(I)$ ist ein offenes Intervall (Warum?), das die 0 enthält. Wähle ein offenes Intervall $I_2$, so dass $x_0 \in I_2$ und $G(I_2) \subset F(I)$. (Warum existiert das?) Dann gilt für $y(x) = H(G(x))$ $(x \in I_2)$:

$$y'(x) = H'(G(x))G'(x) = \frac{1}{F'(H(G(x)))}g(x) = \frac{1}{F'(y(x))}g(x) = f(y(x))g(x),$$

und $y'(x_0) = H(0) = y_0$, also löst $y$ das Anfangswertproblem.

Sei andererseits $\tilde{y} : I_2 \to \mathbb{R}$ irgendeine Lösung des Anfangswertproblems, dann gilt für jedes $x \in I_2$

$$\int_{x_0}^{x} g(x)\mathrm{d}x = \int_{x_0}^{x} \frac{\tilde{y}'(x)}{f(\tilde{y}(x))}\mathrm{d}x = \int_{y_0}^{\tilde{y}(x)} \frac{1}{f(\tilde{y})}\mathrm{d}\tilde{y} = F(\tilde{y}(x)).$$

Nach Voraussetzung ist die linke Seite im Definitionsbereich von $H$ und damit gilt $\tilde{y}(x) = H(G(x)) = y(x)$. $\square$

*Bemerkung 4.9.* Die Wahl von $I_2$ ist natürlich nicht eindeutig. Wir können jedoch das maximale Intervall wählen, das heißt die Vereinigung aller Intervalle die die gestellten Bedingungen erfüllen. In diesem Sinne gibt es für jedes $x_0$ und $y_0$ eine eindeutige Lösung der Differenzialgleichung mit maximalem Definitionsbereich.

Analog zum Bestimmen von Stammfunktionen und Integralen mittels Substitution kann man auch Differenzialgleichungen durch Substitution auf bereits bekannte Fälle zurückführen. Ähnlich wie bei Integralen gibt es kein „Rezept“ zum Auffinden geeigneter Substitutionen.

**Satz 4.10.** *Sei $f : \mathbb{R} \to \mathbb{R}$ eine Funktion, $a, b, c \in \mathbb{R}$ und $I$ ein offenes Intervall. Dann ist $y : I \to \mathbb{R}$ eine Lösung der Differenzialgleichung*

$$y' = f(ax + by + c)$$

*genau dann wenn $u(x) = ax + by(x) + c$ eine Lösung von*

$$u' = a + bf(u)$$

*ist. Letztere Gleichung hat getrennte Variablen.*

*Der Beweis der Aussage folgt direkt aus der Gleichung*

$$u'(x) = a + by'(x).$$

*Beispiel 4.11 (eulerhomogene Differenzialgleichung).* Sei $f : \mathbb{R} \to \mathbb{R}$ eine Funktion und $I$ ein offenes Intervall das die 0 nicht enthält. Dann ist $y : I \to \mathbb{R}$ eine Lösung der Differenzialgleichung

$$y' = f\left(\frac{y}{x}\right)$$

genau dann wenn $u(x) = \frac{y(x)}{x}$ die Differenzialgleichung

$$u' = \frac{f(u) - u}{x}$$

löst.

*Beweis.* Übung □

*Beispiel 4.12.* Sei $f : \mathbb{R} \to \mathbb{R}$ eine Funktion, $a_1, a_2, b_1, b_2, c_1, c_2 \in \mathbb{R}$ so, dass

$$\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \neq 0.$$

Seien $\tilde{x}$ und $\tilde{y}$ die eindeutig bestimmten (Warum?) Lösungen des Gleichungssystems

$$\begin{aligned} a_1\tilde{x} + b_1\tilde{y} + c_1 &= 0 \\ a_2\tilde{x} + b_2\tilde{y} + c_2 &= 0. \end{aligned}$$

Sei $I$ ein offenes Intervall, das die 0 nicht enthält. Dann ist $y : I \to \mathbb{R}$ eine Lösung der Differenzialgleichung

$$y' = f\left(\frac{a_1 x + b_1 y + c_1}{a_2 x + b_2 y + c_2}\right)$$

genau dann wenn

$$\begin{aligned} u : I - \tilde{x} &\to \mathbb{R} \\ x &\mapsto y(x + \tilde{x}) - \tilde{y} \end{aligned}$$

eine Lösung von

$$u' = f\left(\frac{a_1 + b_1 \frac{u(x)}{x}}{a_2 + b_2 \frac{u(x)}{x}}\right)$$

ist. Die letzte Gleichung ist eine eulerhomogene Differenzialgleichung.

Wir zeigen nur die eine Richtung der Substitution. Die Rücksubstitution verläuft analog.

$$\begin{aligned} u'(x) = y'(x + \tilde{x}) &= f\left(\frac{a_1(x + \tilde{x}) + b_1 y(x + \tilde{x}) + c_1}{a_2(x + \tilde{x}) + b_2 y(x + \tilde{x}) + c_2}\right) \\ &= f\left(\frac{a_1 x + b_1 u(x) + a_1 \tilde{x} + b_1 \tilde{y} + c_1}{a_2 x + b_2 u(x) + a_2 \tilde{x} + b_2 \tilde{y} + c_2}\right) \\ &= f\left(\frac{a_1 + b_1 \frac{u(x)}{x}}{a_2 + b_2 \frac{u(x)}{x}}\right) \end{aligned}$$

**Definition 4.13 (exakte Differenzialgleichung).** Seien $D \subset \mathbb{R}^2$ offen und $p, q : D \to \mathbb{R}$ stetige Funktionen. Die Differenzialgleichung

$$p(x, y) + q(x, y) y' = 0$$

heißt exakt, wenn es eine stetig differenzierbare Funktion $H : D \to \mathbb{R}$ gibt, so dass $\partial_1 H = p$ und $\partial_2 H = q$ gelten. Die Funktion $H$ heißt Stammfunktion oder Potentialfunktion der exakten Differenzialgleichung.

**Satz 4.14.** *Seien $D \subset \mathbb{R}^2$ offen und $p, q : D \to \mathbb{R}$ stetige Funktionen, die Differenzialgleichung*

$$p(x, y) + q(x, y) y' = 0$$

*exakt und $H$ eine Potentialfunktion. Sei $I$ ein offenes Intervall und $y : I \to \mathbb{R}$ eine Funktion, so dass der Graph $(x, y(x))$ der Funktion in $D$ enthalten ist. Dann ist $y$ eine Lösung der Differenzialgleichung genau dann wenn es ein $c \in \mathbb{R}$ gibt, so dass $H(x, y(x)) = c$ für alle $x \in I$ ist.*

*Beweis.* Die Aussage folgt unmittelbar aus der Gleichung

$$\frac{\mathrm{d}}{\mathrm{d}x}H(x,y(x)) = \partial_1 H(x,y(x)) + \partial_2 H(x,y(x))y'(x) = p(x,y(x)) + q(x,y(x))y'(x). \quad \square$$

*Bemerkung 4.15.* Die obige Aussage führt die Lösung einer Differenzialgleichung auf die Lösung einer gewöhnlichen Gleichung zurück (zu beachten ist jedoch, dass die Lösung der Gleichung gewisse Bedingungen erfüllen muss um auch eine Lösung der Differenzialgleichung zu sein). Allerdings ist dazu die Kenntnis der Potentialfunktion notwendig. Eine systematische Methode zur Bestimmung von Stammfunktionen steht uns bisher nicht zur Verfügung. Der folgende Satz gibt uns aber wenigstens ein notwendiges Kriterium zum Erkennen von exakten Differenzialgleichungen an die Hand.

**Satz 4.16.** *Seien $D \subset \mathbb{R}^2$ offen und $p, q : D \to \mathbb{R}$ stetig differenzierbare Funktionen. Falls die Differenzialgleichung*

$$p(x,y) + q(x,y)y' = 0$$

*exakt ist, gilt $\partial_2 p(x,y) = \partial_1 q(x,y)$.*

*Beweis.* Sei $H$ eine Potentialfunktion, dann ist $H$ zweimal stetig differenzierbar und nach dem Satz von Schwarz gilt

$$\partial_2 p(x,y) = \partial_2\partial_1 H(x,y) = \partial_1\partial_2 H(x,y) = \partial_1 q(x,y). \quad \square$$

*Bemerkung 4.17.* Unter zusätzlichen Voraussetzungen an den Definitionsbereich $D$ ist die obige „Integrabilitätsbedingung“ auch hinreichend. Wir werden später darauf zurückkommen.

*Beispiel 4.18.* Wir betrachten das Anfangswertproblem

$$\begin{aligned} 2x + y^2 + 2xyy' &= 0 \\ y(1) &= 1. \end{aligned}$$

Es sind $p(x,y) = 2x + y^2$ und $q(x,y) = 2xy$ und es gilt $\partial_2 p(x,y) = 2y = \partial_1 q(x,y)$, die notwendige Bedingung für Exaktheit ist also erfüllt.

Für eine Potentialfunktion $H$ muss gelten $\partial_1 H(x,y) = 2x + y^2$, also muss $H(x,y) = x^2 + xy^2 + G(y)$ sein, für eine differenzierbare Funktion $G$. Außerdem muss gelten

$$\partial_2 H(x,y) = 2xy + G'(y) = 2xy.$$

Also muss $G$ konstant sein und eine mögliche Potentialfunktion ist $H(x,y) = x^2 + xy^2$. Für die gesuchte Lösung gilt also $H(x,y(x)) = x^2 + xy^2 = H(1,1) = 2$. Die letzte Gleichung hat die beiden Lösungen

$$y(x) = \pm\sqrt{\frac{2}{x} - x}.$$

von denen jedoch nur die mit positivem Vorzeichen die Anfangsbedingung erfüllt. Diese Funktion ist auf $(-\infty, -2] \cup (0, 2]$ definiert, der maximale (sinnvolle) Definitionsbereich der Lösung ist jedoch $(0, 2)$.

*Bemerkung 4.19.* Falls eine Differenzialgleichung der Form

$$p(x, y) + q(x, y)y' = 0$$

nicht exakt ist, so kann man versuchen, einen integrierenden Faktor zu finden, das heißt eine nichtverschwindende Funktion $h(x, y)$, so dass

$$h(x, y)p(x, y) + h(x, y)q(x, y)y' = 0$$

exakt ist. In manchen Fällen ist es möglich, einen integrierenden Faktor zu raten, oder durch bestimmte Ansätze zu ermitteln (zum Beispiel ist es üblich anzunehmen, dass $h$ nur von $x$ oder nur von $y$ abhängt).

Allgemein können wir mit Satz 4.16 die folgende notwendige Bedingung für einen integrierenden Faktor herleiten:

$$p(x, y)\partial_2 h(x, y) + h(x, y)\partial_2 p(x, y) = q(x, y)\partial_1 h(x, y) + h(x, y)\partial_1 q(x, y).$$

Da es sich hier um eine partielle Differenzialgleichung handelt, wird sich die Lösung des Problems dadurch im Allgemeinen nicht vereinfachen.

**Definition 4.20 (Gewöhnliche Differenzialgleichungssysteme).** Ein gewöhnliches Differenzialgleichungssystem (GDGS) ist eine Gleichung der Form

$$F(x, y, y', \ldots, y^{(n)}) = 0$$

für ein $F : \mathbb{R} \times \mathbb{R}^l \times \mathbb{R}^l \times \cdots \times \mathbb{R}^l \to \mathbb{R}^m$.

Eine mindestens $n$ mal differenzierbare Funktion $y : I \to \mathbb{R}^l$ – $I$ offenes Intervall – heißt Lösung des Differnzialgleichungssystems, wenn für $x \in I$ gilt

$$F(x, y(x), y'(x), \ldots, y^{(n)}(x)) = 0.$$

*Beispiel 4.21.* (i) Das Differenzialgleichungssystem ($z = (z_1, z_2, z_3)$)

$$z'' = -\frac{z}{\|z\|^3} = -\frac{1}{\|z\|^2}\frac{z}{\|z\|}$$

beschreibt die Bewegung einer Punktmasse im Gravitationsfeld außerhalb einer radialsymmetrischen Massenverteilung.

(ii) Die Lotka-Volterra-Gleichungen sind ein Modell für die Entwicklung der Populationsgröße von Räubern $r$ sowie deren Beutetieren $b$:

$$\begin{aligned} b' &= \alpha_1 b - \gamma_1 br \\ r' &= \alpha_2 r + \gamma_2 br. \end{aligned}$$

*Bemerkung 4.22.* Das Differenzialgleichungssystem

$$F(x, y, y', y'', \dots, y^{(n-1)}, y^{(n)}) = 0$$

ist äquivalent (hat also genau die selben Lösungen) wie das Differenzialgleichungssystem erster Ordnung

$$\begin{aligned} F(x, y, y_1, y_2, \dots, y_{n-1}, y'_{n-1}) &= 0 \\ y' &= y_1 \\ y'_1 &= y_2 \\ &\vdots \\ y'_{n-2} &= y_{n-1}. \end{aligned}$$

Diese Substitution bringt uns einer Lösung natürlich nicht näher, wir können uns aber in der theoretischen Behandlung von Differenzialgleichungssystemen auf Systeme erster Ordnung beschränken.

## 4.2 Der Satz von Picard-Lindelöf

*Beispiel 4.23.* Betrachte die Differenzialgleichung

$$y' = 2\sqrt{|y|}.$$

Man überzeugt sich leicht, dass $y(x) = (x-c)^2$ für $x > c$, $y(x) = -(x-c)^2$ für $x < c$ und $y(x) = 0$ für $x \in \mathbb{R}$ Lösungen sind. Aus diesen können wir jetzt allerdings weitere Lösungen konstruieren. Seien $a, b \in \mathbb{R}$ und $a \leq b$ und

$$y_1(x) = \begin{cases} (x-a)^2 & x \in (-\infty, a) \\ 0 & x \in [a, b] \\ (x-b)^2 & x \in (b, \infty) \end{cases}$$

Dann ist die Funktionen $y_1$ Lösungen der Differenzialgleichung für beliebige Werte von $a$ und $b$ (Warum gilt das auch in $a$ und $b$?).

Insbesondere lösen diese Funktionen das Anfangswertproblem

$$\begin{aligned} y' &= \sqrt{|y|} \\ y(0) &= 0 \end{aligned}$$

falls $a \leq 0 \leq b$ ist.

Unser Ziel ist nun ein Satz, der sicherstellt, dass ein gegebenes Anfangswertproblem genau eine Lösung hat. Dazu benötigen wir zunächst noch einige Vorbereitungen.

**Definition 4.24.** Sei $D \subset \mathbb{R} \times \mathbb{R}^n$, $(x_0, y_0) \in D$ und $f : D \to \mathbb{R}^n$. Wir sagen, dass $f$ eine lokale Lipschitzbedingung in $y$ (oder bezüglich $y$) erfüllt, wenn es eine Umgebung $U$ von $(x_0, y_0)$ gibt und eine Konstante $L \in \mathbb{R}$ gibt, so dass

$$\|f(x, y) - f(x, z)\| \leq L \|y - z\|$$

für alle $(x, y), (x, z) \in U$ erfüllt ist. Falls wir $U = D$ wählen können, so sagen wir, dass $f$ eine globale Lipschitzbedingung bezüglich $y$ erfüllt.

Wir werden im Zusammenhang mit Differenzialgleichungssystemen auf Ausdrücke der Form

$$\int_a^b f(x)\mathrm{d}x$$

stoßen für vektorwertige $f : [a, b] \to \mathbb{R}^n$. Solche Integrale definieren wir komponentenweise.

**Definition 4.25.** Sei also $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $f : \Omega \to \mathbb{R}^n$ eine messbare Funktion und $f_1, \dots, f_n$ die Komponentenfunktionen von $f$, das heißt $f(\omega) = (f_1(\omega), \dots, f_n(\omega))$. Dann heißt $f$ integrierbar, falls $f_1, \dots, f_n$ integrierbar sind und in diesem Falle definieren wir

$$\int f \mathrm{d}\mu := \left( \int f_1 \mathrm{d}\mu, \dots, \int f_n \mathrm{d}\mu \right).$$

**Satz 4.26.** *Sei $(\Omega, \Sigma, \mu)$ ein Maßraum, $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$ und $f : \Omega \to \mathbb{R}^n$ eine messbare Funktion. Die Funktion $f$ ist integrierbar genau dann, wenn $\|f\|$ integrierbar ist und es gilt*

$$\left\| \int f \mathrm{d}\mu \right\| \leq \int \|f\| \, \mathrm{d}\mu.$$

Ohne Beweis.

**Lemma 4.27.** *Sei $D \subset \mathbb{R} \times \mathbb{R}^n$, $(x_0, y_0) \in D$ und $f : D \to \mathbb{R}^n$ stetig. Sei $I$ ein offenes Intervall, $y : I \to \mathbb{R}^n$ stetig differenzierbar und $(x, y(x)) \in D$ für jedes $x \in I$. Dann ist $y$ Lösung des Anfangswertproblems*

$$\begin{aligned} y' &= f(x, y) \\ y(x_0) &= y_0 \end{aligned}$$

*genau dann wenn es die Integralgleichung*

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t)) \mathrm{d}t$$

*erfüllt.*

*Beweis.* Falls $y$ das Anfangswertproblem löst, dann gilt nach dem Hauptsatz

$$y(x) - y_0 = y(x) - y(x_0) = \int_{x_0}^{x} y'(t) \mathrm{d}t = \int_{x_0}^{x} f(t, y(t)) \mathrm{d}t.$$

Ist die Integralgleichung erfüllt, so folgt

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

ebenfalls aus dem Hauptsatz. □

**Theorem 4.28 (Satz von Picard-Lindelöf).** *Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen und $(x_0, y_0) \in D$. Sei $f : D \to \mathbb{R}^n$ stetig und erfülle eine lokale Lipschitzbedingung in $(x_0, y_0)$ bezüglich $y$. Dann existiert $\epsilon > 0$ so, dass das Anfangswertproblem*

$$\begin{aligned} y' &= f(x, y) \\ y(x_0) &= y_0 \end{aligned}$$

*auf $(x_0 - \epsilon, x_0 + \epsilon)$ genau eine Lösung hat.*

*Beweis.* Wir haben oben die Lösung des Anfangswertproblems zu einem Fixpunktproblem umformuliert und wollen darauf nun den Banachschen Fixpunktsatz anwenden. Sei dazu $U \subset D$ eine Umgebung von $(x_0, y_0)$ und $L > 0$, so dass

$$\|f(x,y) - f(x,z)\| \leq L\,\|y - z\|$$

für alle $(x,y),(x,z) \in U$. Wähle $a, r > 0$ so, dass $\tilde{D} = [x_0 - a, x_0 + a] \times K_r(y_0) \subset U$ (Warum geht das?). Da $\tilde{D}$ abgeschlossen und beschränkt ist, existiert $M > 0$ so, dass $\|f(x,y)\| \leq M$ für $(x,y) \in \tilde{D}$ (Warum existiert das?). Wir wählen jetzt $a \geq \epsilon > 0$ so, dass $\epsilon M < r$ und $\epsilon L < 1$ und definieren $I = [x_0 + \epsilon, x_0 - \epsilon]$.

Sei nun $X \subset C(I)$ die Menge der Funktionen $y$, die die Bedingung $y(x) \in K_r(y_0)$ für alle $x \in I$ erfüllen. Dann ist $X$ eine abgeschlossene Teilmenge (Warum?) des vollständigen Raumes $C(I)$ und somit selbst vollständig (Satz 1.38). Wir definieren die Abbildung $T : X \to X$ durch

$$T(y)(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\mathrm{d}t.$$

Dazu müssen wir zunächst zeigen, dass $T(y)$ tatsächlich wiederum in $X$ liegt. Die Stetigkeit folgt dabei aus dem Hauptsatz. Für $x \geq x_0$ (der Fall $x < x_0$ wird analog behandelt) folgt weiter

$$\begin{aligned}\|T(y)(x) - y_0\| &= \left\|\int_{x_0}^{x} f(t,y(t))\mathrm{d}t\right\| \leq \int_{x_0}^{x} \|f(t,y(t))\|\,\mathrm{d}t \\ &\leq \int_{x_0}^{x} M\mathrm{d}t = (x - x_0)M \leq \epsilon M < r,\end{aligned}$$

das heißt $T(y)(x) \in K_r(y_0)$ für alle $x \in I$.

Schließlich gilt für $y, \tilde{y} \in X$ und $x \in I$ (wir betrachten erneut nur $x \geq x_0$)

$$\begin{aligned}\|T(y)(x) - T(\tilde{y})(x)\| &= \left\|\int_{x_0}^{x} (f(t,y(t)) - f(t,\tilde{y}(t))\mathrm{d}t\right\| \\ &\leq \int_{x_0}^{x} \|f(t,y(t)) - f(t,\tilde{y}(t)\|)\mathrm{d}t \\ &\leq \int_{x_0}^{x} L\,\|y(t) - \tilde{y}(t)\|\,\mathrm{d}t \leq L \int_{x_0}^{x} \|y - \tilde{y}\|_\infty\,\mathrm{d}t \leq \epsilon L\,\|y - \tilde{y}\|_\infty\,.\end{aligned}$$

Bilden wir nun das Supremum über alle $x \in I$, so folgt

$$\|T(y) - T(\tilde{y})\|_\infty \leq \epsilon L\,\|y - \tilde{y}\|_\infty\,,$$

da $\epsilon L < 1$ ist, haben wir gezeigt, dass $T$ eine strenge Kontraktion ist.

Jetzt folgt aus dem Banachschen Fixpunktsatz (Theorem 2.63) die Existenz eines eindeutigen Fixpunktes für die Gleichung $T(y) = y$. Aus Lemma 4.27 folgt dann, dass das Anfangswertproblem auf $I$ eine eindeutige Lösung hat. $\square$

*Bemerkung 4.29.* Aus dem Banachschen Fixpunktsatz folgt auch ein Verfahren zum Bestimmen der Lösung (Bemerkung 2.64). Wir wenden die Abbildung $T$ iterativ auf einen beliebigen Startpunkt in $X$, also zum Beispiel auf die konstante Funktion $\varphi(x) = y_0$ an:

$$\varphi_0 = \varphi \text{ und } \varphi_i(x) = y_0 + \int_{x_0}^{x} \varphi_{i-1}(x)\mathrm{d}t.$$

Dann konvergieren die sogenannten Picard-Iterierten $\varphi_i$ in $X$, also gleichmäßig, gegen die gesuchte Lösung. Dieses Verfahren eignet sich prinzipiell auch zum numerischen Lösen von Differenzialgleichungssystemen, der praktische Nutzen hängt jedoch von der Größe des Intervalls $(x_0 - \epsilon, x_0 + \epsilon)$ ab.

*Beispiel 4.30.* Wir betrachten erneut die Gleichung $y' = f(y) := \sqrt{|y|}$. Es gilt (vergleiche Beispiel 1.52)

$$\lim_{y \to 0} \frac{|f(y) - f(0)|}{|y - 0|} = \lim_{y \to 0} \frac{1}{\sqrt{|y|}} = \infty$$

also erfüllt $f$ in keinem Punkt $(x_0, 0)$ eine lokale Lipschitzbedingung.

**Satz 4.31.** *Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f : D \to \mathbb{R}^n$ stetig differenzierbar, dann erfüllt $f$ eine lokale Lipschitzbedingung bezüglich $y$.*

*Beweis.* Sei $(x_0, y_0) \in D$. Wähle $r > 0$ so, dass $K_r(x_0, y_0) \subset D$. Da $f$ stetig differenzierbar ist, ist $D_y f$ eine stetige Funktion und damit auf $K_r(x_0, y_0)$ (beschränkt und abgeschlossen) beschränkt (Satz 1.64), es gibt also ein $L > 0$ so, dass $\|D_y f(x, y)\| \leq L$ für alle $(x, y) \in K_r(x_0, y_0)$. Mit dem allgemeinen Mittelwertsatz (Satz 2.38) erhalten wir für $(x, y), (x, z) \in K_r(x_0, y_0)$ (der Wert von $x$ ist hier zunächst fest)

$$\|f(x, y) - f(x, z)\| \leq \sup_{t \in [0,1]} \|D_y f(x, y + t(z - y))\| \, \|y - z\| \leq L \|y - z\| . \qquad \square$$

*Beispiel 4.32.* Betrachte die Differenzialgleichung

$$y'' = -\frac{y}{\|y\|^3}$$

beziehungsweise das äquivalente Differenzialgleichungssystem

$$(y, z)' = \left(z, -\frac{y}{\|y\|^3}\right).$$

Die Funktion

$$\begin{aligned} f : R \times (\mathbb{R}^3 \setminus \{0\}) \times \mathbb{R}^3 &\to \mathbb{R}^3 \times \mathbb{R}^3 \\ (x, y, z) &\mapsto \left(z, -\frac{y}{\|y\|^3}\right) = \left(z, -\frac{1}{\sqrt{y_1^2 + y_2^2 + y_3^2}^3}(y_1, y_2, y_3)\right) \end{aligned}$$

ist stetig differenzierbar auf ihrem gesamten Definitionsbereich und $f$ erfüllt damit eine lokale Lipschitzbedingung. Das Anfangswertproblem für die Planeten- beziehungsweise Satelitenbewegung hat also eine eindeutige Lösung.

**Definition 4.33.** Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen und $(x_0, y_0) \in D$. Eine Lösung $\tilde{y} : I \to \mathbb{R}^n$ des Anfangswertproblems

$$\begin{aligned} y' &= f(x,y) \\ y(x_0) &= y_0 \end{aligned}$$

heißt (echte) Fortsetzung der Lösung $y : J \to \mathbb{R}^n$ falls $J \subset I$ ($J \subsetneq I$) und $y(x) = \tilde{y}(x)$ für alle $x \in J$ gilt. Die Lösung $y$ heißt maximale Lösung, falls sie keine echten Fortsetzungen besitzt.

**Satz 4.34.** *Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen und $(x_0, y_0) \in D$. Sei $f : D \to \mathbb{R}^n$ stetig und erfülle eine lokale Lipschitzbedingung (in allen Punkten von $D$) bezüglich $y$. Dann besitzt das Anfangswertproblem*

$$\begin{aligned} y' &= f(x,y) \\ y(x_0) &= y_0 \end{aligned} \tag{4.1}$$

*eine eindeutige, maximale Lösung.*

*Beweis.* Wir zeigen zunächst: falls $y : I \to \mathbb{R}^n$ und $\tilde{y} : J \to \mathbb{R}^n$ Lösungen des Anfangswertproblems sind, dann gilt $y(x) = \tilde{y}(x)$ für alle $x \in (a,b) := I \cup J$. Sei

$$c = \sup \{\tilde{c} \in [x_0, b) |\ y = \tilde{y} \text{ auf } [x_0, \tilde{c})\}.$$

Dann gibt es eine Folge $(c_n) \subset (x_0, c)$ so, dass $c_n < c$ – also $y(c_n) = \tilde{y}(c_n)$ für alle $n \in \mathbb{N}$ und $c_n \to c$. Falls $c < b$ ist, folgt aus der Stetigkeit von $y$ und $\tilde{y}$ dann $y(c) = \tilde{y}(c)$. Da $f$ im Punkt $(c, y(c))$ eine lokale Lipschitzbedingung erfüllt, existiert $\epsilon > 0$ so, dass das Anfangswertproblem

$$\begin{aligned} u' &= f(x,u) \\ u(c) &= y(c) \end{aligned}$$

auf $(c - \epsilon, c + \epsilon)$ eine eindeutige Lösung hat. Da sowohl $y$, als auch $\tilde{y}$ dieses Anfangswertproblem erfüllen, muss $y = \tilde{y}$ auf $[x_0, c + \epsilon)$ gelten, was jedoch der Definition von $c$ widerspricht. Es muss also $c = b$ und $y = \tilde{y}$ auf $[x_0, b)$ sein. Analog kann man nun mit dem Intervall $(a, x_0]$ verfahren.

Sei nun $I_{\max}$ die Vereinigung aller offenen Intervalle $I$, die Definitionsbereich einer Lösung des Anfangswertproblems (4.1) sind. Für $x \in I_{\max}$ wähle eine Lösung $y : I \to \mathbb{R}^n$ mit $x \in I$ und setze $y_{\max}(x) := y(x)$. Nach dem ersten Teil des Beweises ist $y(x)$ nicht von der Wahl der Lösung $y$ abhängig, und $y_{\max}$ daher wohldefiniert. Gemäß seiner Definition stimmt $y_{\max}$ in einer Umgebung jedes Punktes $x$ mit einer Lösung des Anfangswertproblems überein und ist daher selbst eine Lösung dieses Anfangswertproblems. Offensichtlich ist $y_{\max}$ eine maximale Lösung. Falls $y : I \to \mathbb{R}^n$ eine weitere maximale Lösung ist, so gilt nach Konstruktion $I \subset I_{\max}$, nach dem ersten Teil des Beweises $y = y_{\max}$ auf $I$ und schließlich $I = I_{\max}$ aufgrund der Maximalität von $y$. $\square$

*Beispiel 4.35.* (i) Das Anfangswertproblem

$$\begin{aligned} y' &= e^{-y} \\ y(1) &= 0 \end{aligned}$$

hat die Lösung $y : (0, \infty) \to \mathbb{R}$, $y(x) = \ln x$ welche nach dem Satz von Picard-Lindelöf eindeutig ist. Sie ist auch maximal, da sie nach rechts offensichlich nicht fortgesetzt werden kann, und eine hypotetische Fortsetzung nach links in 0 nicht stetig sein könnte.

(ii) Das Anfangswertproblem

$$\begin{aligned} y' &= -\mathrm{i}\frac{y}{x^2} \\ y\left(\frac{1}{2\pi}\right) &= 1 \end{aligned}$$

hat die Lösung $y : (0, \infty) \to \mathbb{C}$, $y(x) = e^{\frac{\mathrm{i}}{x}}$. Da die Funktion

$$\begin{aligned} f : (\mathbb{R}\backslash 0) \times \mathbb{R} &\to \mathbb{R} \\ (x, y) &\mapsto -\frac{y}{x^2} \end{aligned}$$

stetig differenzierbar ist, ist diese Lösung nach dem Satz von Picard-Lindelöf wiederum eindeutig. Sie ist ebenfalls maximal, da sie dem Rand des Definitionsbereiches von $f$ beliebig nahe kommt.

Im folgenden Satz benutzen wir den Abstand zu einer Menge. Für einen metrischen Raum $(X, d)$, $x \in X$ und $A \subset X$ ist

$$\operatorname{dist}(x, A) := \inf \{d(x, y) |\ y \in A\} .$$

**Satz 4.36.** *Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen und $(x_0, y_0) \in D$. Sei $f : D \to \mathbb{R}^n$ stetig und erfülle eine lokale Lipschitzbedingung bezüglich $y$. Seien $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$ so, dass $-\infty \leq a < x_0 < b \leq \infty$ und $y : (a, b) \to \mathbb{R}^n$ eine Lösung des Anfangswertproblems*

$$\begin{aligned} y' &= f(x, y) \\ y(x_0) &= y_0. \end{aligned}$$

*Dann ist $y$ die maximale Lösung, genau dann wenn eine der Bedingungen*

*(i)* $b = \infty$,

*(ii)* $\lim_{x \to b} \|y(x)\| = \infty$,

*(iii)* $\lim_{x \to b} \operatorname{dist}((x, y(x)), \partial D) = 0$

*und eine der Bedinungen*

*(i)* $a = -\infty$,

*(ii)* $\lim_{x \to a} \|y(x)\| = \infty$,

*(iii)* $\lim_{x \to a} \operatorname{dist}((x, y(x)), \partial D) = 0$

*erfüllt ist.*

Ohne Beweis.

## 4.3 Lineare Systeme

**Definition 4.37.** Sei $I$ ein offenes Intervall und $A : I \to \mathbb{R}^{n\times n}$ und $b : I \to \mathbb{R}^n$ Funktionen. Dann heißt das Differenzialgleichungssystem

$$y' = A(x)y + b(x)$$

lineares Differenzialgleichungssystem.

Falls $b$ die Nullfunktion ist, so nennen wir das System homogen, anderenfalls inhomogen. Falls $A(x)$ konstant ist, sagen wir, dass das System konstante Koeffizienten hat.

*Bemerkung 4.38.* (i) Mittels der Substitution in Bemerkung 4.22 können wir die lineare Differenzialgleichung

$$y^{(n)} = a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y + b(x)$$

in das System

$$\begin{aligned} y'_{n-1} &= a_{n-1}(x)y_{n-2} + \cdots + a_1(x)y_1 + a_0(x)y + b(x) \\ y_1 &= y' \\ &\vdots \\ y_{n-1} &= y'_{n-2} \end{aligned}$$

überführen, dass von obiger Form ist.

(ii) Seien $y$ und $z$ Lösungen des inhomogenen Systems

$$y' = A(x)y + b(x),$$

dann ist $y - z$ eine Lösung des zugehörigen homogenen Systems

$$y' = A(x)y.$$

Das folgt unmittelbar aus

$$(y-z)'(x) = A(x)y(x) + b(x) - A(x)z(x) - b(x) = A(x)(y-z)(x).$$

**Lemma 4.39 (Lemma von Grönwall).** *Sei $I$ ein Intervall, $x_0 \in I$, $y : I \to \mathbb{R}$ eine stetige, nicht-negative Funktion, $a, b \geq 0$ und es gelte*

$$y(x) \leq a + b\left|\int_{x_0}^{x} y(t)\mathrm{d}t\right|. \tag{4.2}$$

*Dann gilt $y(x) \leq ae^{|x-x_0|}$.*

*Beweis.* Wir führen den Beweis für $t \geq t_0$, der andere Fall wird analog behandelt. Sei $\epsilon > 0$ beliebig und setze

$$z(x) := a + \epsilon + b \int_{x_0}^{x} y(t)\mathrm{d}t. \tag{4.3}$$

Dann gilt $z'(x) = by(x) \leq bz(t)$. Da $z(t) \geq a + \epsilon > 0$ ist, erhalten wir

$$\ln z(x) - \ln z(x_0) = \int_{x_0}^{x} \frac{z'(t)}{z(t)}\mathrm{d}t \leq \int_{x_0}^{x} b\mathrm{d}t = b(x - x_0)$$

also dank der Monotonie der Exponentialfunktion $z(x) \leq z(x_0)e^{b(x-x_0)} = (a+\epsilon)e^{b(x-x_0)}$. Dann gilt auch $y(x) \leq z(x) \leq (a+\epsilon)e^{b(x-x_0)}$ und da diese Ungleichung für beliebige $\epsilon > 0$ gilt, folgt die gesuchte Ungleichung.. □

Für den Rest dieses Unterkapitels seien stets $I$ ein offenes Intervall, $A : I \to \mathbb{R}^{n\times n}$ und $b : I \to \mathbb{R}^n$ stetige Funktionen, $x_0 \in I$ und $y_0 \in \mathbb{R}^n$.

**Korollar 4.40.** *Des Anfangswertproblem*

$$\begin{aligned} y' &= A(x)y + b(x) \\ y(x_0) &= y_0 \end{aligned}$$

*hat eine eindeutige maximale Lösung die auf ganz $I$ definiert ist.*

*Beweis.* Wir zeigen zunächst, dass

$$\begin{aligned} f : I \times \mathbb{R}^n &\to \mathbb{R}^n \\ (x, y) &\mapsto A(x)y + b(x) \end{aligned}$$

eine lokale Lipschitzbedingung erfüllt. Sei dazu $(x_1, y_1) \in I \times \mathbb{R}^n$. Wähle ein abgeschlossenes Intervall $I_1$ so, dass $x_1 \in I_1 \subset I$. Dann existiert $L > 0$ so, dass $\|A(x)\| \leq L$ für alle $x \in I_1$ und somit gilt

$$\|f(x, y) - f(x, z)\| = \|A(x)(y - z)\| \leq \|A(x)\| \|y - z\| \leq L \|y - z\|$$

für alle $x \in I_1$ und $y, z \in \mathbb{R}^n$.

Es folgt also aus Satz 4.34 die Existenz und Eindeutigkeit einer maximalen Lösung. Es seien $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$ so, dass $y : (a, b) \to \mathbb{R}^n$ diese maximale Lösung ist. Angenommen $b$ liegt in $I$ (ist also insbesondere endlich). Dann gibt es $M, K > 0$ so, dass $\|A(x)\| \leq M$ und $\|b(x)\| \leq K$ für alle $x \in [x_0, b]$ und es gilt

$$\begin{aligned} \|y(x)\| &= \left\| y_0 + \int_{x_0}^{x} (A(t)y(t) + b(t))\mathrm{d}t \right\| \\ &\leq \|y_0\| + \int_{x_0}^{x} \|A(t)\| \|y(t)\|\mathrm{d}t + \int_{x_0}^{x} \|b(t)\| \mathrm{d}t \\ &\leq \|y_0\| + K(b - x_0) + M \int_{x_0}^{x} \|y(t)\| \mathrm{d}t. \end{aligned}$$

Aus dem Lemma von Grönwall (für $\|y\|$) folgt dann, dass

$$\|y(x)\| \leq (\|y_0\| + K(b - x_0))e^{M(x-x_0)} \leq (\|y_0\| + K(b - x_0))e^{M(b-x_0)}.$$

Damit kann keiner der Fälle aus Satz 4.36 erfüllt sein, was im Widerspruch zur Annahme steht, dass $y$ die maximale Lösung ist. Analog zeigt man, dass $a \in I$ zu einem Widerspruch führt. Es muss also $a \notin I$ und $b \notin I$ gelten, was $(a, b) = I$ impliziert. $\square$

*Bemerkung 4.41.* Man kann zeigen (und auf diese Weise einen alternativen Beweis für obigen Satz angeben), dass die Picard-Iterierten mit Startfunktion $\varphi_0(x) = y_0$ für lineare Differenzialgleichungen auf ganz $I$ konvergieren.

**Korollar 4.42.** *Seien $y, z : I \to \mathbb{R}^n$ Lösungen der Differenzialgleichung $y' = A(x)y + b(x)$, dann sind äquivalent*

*(i)* $y(x) = z(x)$ *für alle* $x \in I$,

*(ii)* $y(x_0) = z(x_0)$,

*(iii)* $y(x) = z(x)$ *für ein* $x \in I$.

*Beweis.* Die Implikationen von oben nach unten sind offensichtlich. Sei nun $x_1 \in I$ so, dass $y(x_1) = z(x_1) =: y_1$. Dann sind $y$ und $z$ Lösungen des Anfangswertproblems

$$\begin{aligned} y' &= A(x)y + b(x) \\ y(x_1) &= y_1. \end{aligned}$$

Da dieses Anfangswertproblem eine eindeutige maximale Lösung hat muss gelten $y(x) = z(x)$ für alle $x \in I$. $\square$

*Bemerkung 4.43.* Gestrichen.

**Satz 4.44.** *Die Lösungsmenge der Differenzialgleichung $y' = A(x)y$, genauer die Menge*

$$V := \left\{ y : I \to \mathbb{R}^n \,\middle|\, y'(x) = A(x)y(x) \text{ für alle } x \in I \right\}$$

*ist ein $n$-dimensionaler Untervektorraum von $\mathbf{C}^1(I, \mathbb{R}^n)$.*

*Beweis.* Man überprüft leicht, dass für zwei Lösungen $y, z : I \to \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ auch $y + \lambda z$ eine Lösung der Differenzialgleichung ist.

Sei nun $e_1, \dots, e_n$ eine Basis für $\mathbb{R}^n$ und seien $y_i : I \to \mathbb{R}^n$ die eindeutig bestimmten Lösungen der Anfangswertprobleme

$$\begin{aligned} y' &= A(x)y \\ y(x_0) &= e_i \end{aligned}$$

für $i \in \{1, \dots, n\}$. Dann ist $y_1, \dots, y_n$ eine Basis für $V$.

Denn falls

$$\alpha_1 y_1 + \cdots + \alpha_n y_n = 0$$

ist, so gilt insbesondere

$$\alpha_1 y_1(x_0) + \cdots + \alpha_n y_n(x_0) = \alpha_1 e_1 + \cdots + \alpha_n e_n = 0$$

woraus $\alpha_1 = \cdots = \alpha_n = 0$ folgt, die Funktionen $y_1, \ldots, y_n$ sind also linear unabhängig.

Sei nun $z \in V$ und wähle $\alpha_1, \ldots, \alpha_n$ so, dass

$$\alpha_1 e_1 + \cdots + \alpha_n e_n = z(x_0).$$

Dann sind $z$ und $\alpha_1 y_1 + \cdots + \alpha_n y_n$ maximale Lösungen des Anfangswertproblems die in $x_0$ übereinstimmen, also folgt aus Korollar 4.42

$$z = \alpha_1 y_1 + \cdots + \alpha_n y_n.$$

Die Funktionen $y_1, \ldots, y_n$ spannen den Raum $V$ auf. □

*Bemerkung 4.45.* (i) Der obige Satz ist die mathematisch exakte Formulierung von „die allgemeine Lösung hängt von $n$ unbestimmten Parametern ab“. Er gilt in dieser Form nur für lineare Differenzialgleichungen.

(ii) Mit Bemerkung 4.38 erhalten wir, dass die Lösungsmenge des inhomogenen Differenzialgleichungssystems $y' = A(x)y + b(x)$ durch $y_i + V$ gegeben ist, wobei $y_i$ eine beliebige Lösung der inhomogenen Gleichung ist.

**Definition 4.46.** Eine Basis $y_1, \cdots, y_n$ des Vektorraume $V$ aus Satz 4.44 heißt Fundamentalsystem der Differenzialgleichung $y' = A(x)y$.

Nach Bemerkung 4.38 sind lineare Differenzialgleichungen $n$-ter Ordnung äquivalent zu linearen Differenzialgleichungssystemen erster Ordnung (mit $n$ Gleichungen). Wir werden also auch $n$ linear unabhängige Lösungen der Gleichung

$$y^{(n)} = a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y + b(x)$$

als Fundamentalsystem bezeichnen.

*Beispiel 4.47.* Betrachte die Differenzialgleichung $y' = \sin(x)y + \sin(x)\cos(x)$. Die zugehörige homogene Differenzialgleichung $y' = \sin(x)y$ hat getrennte Variablen und wir lösen sie heuristisch wie folgt:

$$\begin{aligned} \frac{y'}{y} &= \sin(x) \\ \ln y(x) &= \int \frac{y'(x)}{y(x)} = -\cos(x) + c \\ y(x) &= Ke^{-\cos(x)}. \end{aligned}$$

Man überprüft leicht, dass die erhaltene Funktion für beliebiges $K \in \mathbb{R}$ tatsächlich eine Lösung der homogenen Differenzialgleichung ist. Darüber hinaus stellt die Menge aller solcher Funktionen offensichtlich einen eindimensionalen Vektorraum dar und damit haben wir bereits alle Lösungen gefunden.

Um nun die inhomogene Gleichung zu lösen verwenden wir das Verfahren der Variation der Konstanten. Falls $y(x) = C(x)e^{-\cos(x)}$ eine Lösung der inhomogenen Differenzialgleichung ist, so muss gelten

$$y'(x) = C'(x)e^{-\cos(x)} + C(x)\sin(x)e^{-\cos(x)} = \sin(x)C(x)e^{-\cos(x)} + \sin(x)\cos(x)$$

Es gilt also

$$\begin{aligned} C'(x) &= \sin(x)\cos(x)e^{\cos(x)} \\ C(x) &= (1-\cos(x))e^{\cos(x)}. \end{aligned}$$

Es genügt eine Lösung der inhomogenen Gleichung zu finden (Bemerkung 4.38).

Das im Beispiel illustrierte Verfahren kann zur Lösung beliebiger inhomogener Differenzialgleichungssysteme benutzt werden, vorausgesetzt man kennt bereits ein Fundamentalsystem des zugehörigen homogenen Systems.

**Satz 4.48.** *Sei $y_1, \dots y_n$ ein Fundamentalsystem der Differenzialgleichung $y' = A(x)y$. Definiere die $n \times n$-Matrix $W(x) := (y_1(x), \cdots, y_n(x))$ (Fundamentalmatrix). Dann ist $W(x)$ für alle $x \in I$ invertierbar und $z : I \to \mathbb{R}^n$*

$$z(x) := W(x) \int_{x_0}^{x} W(t)^{-1} b(t) \mathrm{d}t$$

*ist eine Lösung des inhomogenen Differenzialgleichungssystems.*

*Beweis.* Nach Voraussetzung sind die Funktionen $y_1, \dots, y_n$ linear unabhängig. Für $x \in I$ sind dann auch $y_1(x), \dots, y_n(x)$ linear unabhängig, denn falls

$$\alpha_1 y_1(x) + \ldots + \alpha_n y_n(x) = 0$$

ist, dann muss nach Korollar 4.42

$$\alpha_1 y_1 + \ldots + \alpha_n y_n = 0$$

gelten, was lediglich für $\alpha_1 = \cdots = \alpha_n = 0$ möglich ist. Damit ist die Matrix $W(x)$ für jedes $x$ invertierbar.

Dann ist $z$ wohldefiniert. Auf Grund der Definition von $W$ gilt $W'(x) = A(x)W(x)$ (die $i$-te Spalte dieser Matrixgleichung lautet $y_i'(x) = A(x)y_i(x)$) und wir erhalten

$$z'(x) = W'(x) \int_{x_0}^{x} W(t)^{-1} b(t) \mathrm{d}t + W(x)W(x)^{-1}b(x) = A(x)z(x) + b(x).$$

Für die Anwendung des Hauptsatzes wie oben muss der Integrand stetig sein. Man kann zeigen, dass die Abbildung $A \mapsto A^{-1}$ definiert auf der Menge der invertierbaren Matrizen stetig ist (ohne Beweis). $\square$

*Beispiel 4.49.* Betrachte das System

$$\begin{aligned} u' &= v + \sin(x) \\ v' &= -u + \cos(x) \end{aligned}$$

Man überzeugt sich leicht, dass

$$y_1(x) := \begin{pmatrix} \sin(x) \\ \cos(x) \end{pmatrix} \text{ und } y_2 := \begin{pmatrix} \cos(x) \\ -\sin(x) \end{pmatrix}$$

linear unabhängige Lösungen des zugehörigen homogenen Systems

$$\begin{pmatrix} u \\ v \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} =: A \begin{pmatrix} u \\ v \end{pmatrix}$$

sind. Um eine Lösung des inhomogenen Systems zu finden machen wir den Ansatz

$$z(x) := C_1(x)y_1(x) + C_2(x)y_2(x) = \begin{pmatrix} \sin(x) & \cos(x) \\ \cos(x) & -\sin(x) \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}.$$

Die Koeffizientenmatrix ist dabei die Matrix $W$ aus dem vorhergehenden Satz. Soll $z$ eine Lösung sein, so muss gelten

$$\begin{aligned} z'(x) &= C_1'(x)y_1(x) + C_2'(x)y_2 + C_1(x)y_1'(x) + C_2(x)y_2'(x) \\ &= C_1(x)A(x)y_1(x) + C_2(x)A(x)y_2(x) + b(x), \end{aligned}$$

Um nun $C_1'$ und $C_2'$ zu erhalten muss also das Gleichungssystem

$$C_1'(x)y_1(x) + C_2'(x)y_2(x) = b(x)$$

gelöst werden, konkret

$$\begin{aligned} C_1'(x)\sin(x) + C_2'(x)\cos(x) &= \sin(x) \\ C_1'(x)\cos(x) - C_2'(x)\sin(x) &= \cos(x). \end{aligned}$$

Löst man das lineare Gleichungssystem (zum Beispiel mit dem Gaussverfahren) erhält man

$$\begin{aligned} C_1'(x) &= C_1'(x)(\sin^2(x) + \cos^2(x)) = \sin^2(x) + \cos^2(x) = 1 \\ C_2'(x) &= C_2'(x)(\cos^2(x) + \sin^2(x)) = \sin(x)\cos(x) - \cos(x)\sin(x) = 0 \end{aligned}$$

und damit $C_1(x) = x$ und $C_2(x) = 0$. Die Lösung der homogenen Gleichung ist also

$$y_h(x) = \begin{pmatrix} x\sin(x) \\ x\cos(x) \end{pmatrix}$$

Wir wenden uns nun Systemen mit konstanten Koeffizienten zu, wollen also $y' = Ay$ lösen für $A \in \mathbb{R}^{n\times n}$. Wir werden sehen, dass die „naive“ Verallgemeinerung der Lösung $e^{Ax}$ für $n = 1$ zum Erfolg führt.

Die folgenden Aussagen gelten sinngemäß auch für $A \in \mathbb{C}^{n\times n}$, wir werden sie hier jedoch nur für $A \in \mathbb{R}^{n\times n}$ ausformulieren.

*Bemerkung 4.50.* Sei $A \in \mathbb{R}^{n\times n}$ und betrachte die Funktion

$$e^{Ax} = \sum_{k=0}^{\infty} \frac{1}{k!} A^k x^k. \tag{4.4}$$

Wir wissen, dass für die Konvergenz der Reihe auf der rechten Seite, ihre absolute Konvergenz hinreichend ist. Da

$$\sum_{k=0}^{\infty} \left\| \frac{1}{k!} A^k x^k \right\| \leq \sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k |x|^k = e^{\|A\||x|}$$

gilt, ist $e^{Ax}$ also für alle $x \in \mathbb{R}$ definiert.

Konvergenz in $\mathbb{R}^{n\times n}$ entspricht der koordinatenweisen Konvergenz, das heißt die Koordinatenfunktionen der Reihe (4.4) sind Potenzreihen mit Konvergenzradius $\infty$. Wir können also die Sätze über Potenzreihen (Satz 2.57) koordinatenweise anwenden und erhalten

$$\frac{\mathrm{d}}{\mathrm{d}x} e^{Ax} = \sum_{k=1} \frac{1}{k!} A^k k x^{k-1} = Ae^{Ax}.$$

**Satz 4.51.** *Sei $A \in \mathbb{R}^{n\times n}$. Das Anfangswertproblem mit konstanten Koeffizienten*

$$\begin{aligned} y' &= Ay \\ y(x_0) &= y_0. \end{aligned}$$

*wird genau von $y(x) = e^{A(x-x_0)} y_0$ gelöst.*

*Beweis.* Das folgt unmittelbar aus der vorhergehenden Bemerkung. □

*Bemerkung 4.52.* (i) Mit dem vorhergehenden Satz ist das Lösen der Differenzialgleichung auf das Berechnen der Matrixexponentialfunktion zurückgeführt. Letztere Aufgabe läßt sich algorithmisch gut lösen wie wir bald sehen werden.

(ii) Es gilt im Allgemeinen *nicht*

$$\frac{\mathrm{d}}{\mathrm{d}x} e^{A(x)} = A'(x) e^{A(x)},$$

ebensowenig wie $e^{A+B} = e^A e^B$.

(iii) Ein Fall, in dem wir $e^{Ax}v$ einfach ausrechnen können, ist der, dass $v$ ein Eigenvektor der Matrix $A$ ist, es also ein skalares $\lambda$ gibt, so dass $Av = \lambda v$ gilt. Dann folgt

$$e^{Ax}v = \left(\sum_{k=0}^{\infty} \frac{1}{k!} A^k x^k\right) v = \sum_{k=0}^{\infty} \frac{1}{k!} x^k A^k v = \sum_{k=0}^{\infty} \frac{1}{k!} x^k \lambda^k v = e^{\lambda x} v.$$

*Beispiel 4.53.* Wir suchen Lösungen für das Anfangswertproblem

$$\begin{aligned} y' &= z \\ z' &= y \\ y(0) &= 1 \\ z(0) &= 0. \end{aligned}$$

Wir müssen also $e^{Ax}y_0$ bestimmen für

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ und } y_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

In diesem Falle kommt uns zu Gute, dass die Matrix $A$ diagonalisierbar ist. Sie hat die Eigenwerte 1 und $-1$ und die zugehörigen Eigenvektoren

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ und } v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Mit der obigen Bemerkung erhalten wir dann

$$e^{Ax}y_0 = e^{Ax}(v_1 + v_2) = e^x v_1 + e^{-x} v_2,$$

also die Lösung $y(x) = e^x - e^{-x}$ und $z(x) = e^x + e^{-x}$.

Das hier illustrierte Verfahren führt stets zum Erfolg, falls die Matrix $A$ diagonalisierbar ist.

**Satz 4.54.** *Sei $A \in \mathbb{R}^n \times \mathbb{R}^n$ eine diagonalisierbare Matrix, $v_1, \dots, v_n$ eine Basis aus Eigenvektoren und $\lambda_1, \dots, \lambda_n$ die zugehörigen Eigenwerte. Dann bilden die Funktionen $y_i(x) = e^{\lambda_i x} v_i$ für $i \in \{1, \dots, n\}$ ein Fundamentalsystem für das Differenzialgleichungssystem $y' = Ay$.*

*Beweis.* Für beliebiges $i \in \{1, \dots, n\}$ gilt $e^{Ax}v_i = e^{\lambda_i x} v_i$, die Funktionen sind also tatsächlich Lösungen des Differenzialgleichungssystems. Im Punkt $x = 0$ sind die Funktionswerte $y_1(0) = v_1, \dots, y_n(0) = v_n$ linear unabhängig und damit auch die Funktionen $y_1, \dots, y_n$ (Übungsaufgabe 1 vom Blatt 4). □

*Bemerkung 4.55.* (i) Im obigen Fall ist das Lösen der Differenzialgleichung also auf die Lösung des Eigenwertproblems $Av = \lambda v$ zurückgeführt, wofür wir bereits Methoden kennen.

(ii) Falls $A \in \mathbb{R}^{n \times n}$ nicht im Reellen diagonalisierbar ist, so können wir versuchen $A$ als Matrix in $\mathbb{C}^{n \times n}$ zu interpretieren und dort zu diagonalisieren. Für einen (komplexen) Eigenwert $\lambda = \lambda_r + \mathrm{i}\lambda_i$ und den zugehörigen Eigenvektor $v = v_r + \mathrm{i}v_i$ ($\lambda_r, \lambda_i \in \mathbb{R}$ und $v_r, v_i \in \mathbb{R}^n$) sind dann

$$e^{\lambda_r x}(v_r \sin(\lambda_i x) + v_i \cos(\lambda_i x)) \text{ und } e^{\lambda_r x}(v_r \cos(\lambda_i x) - v_i \sin(\lambda_i x))$$

linear unabhängige reelle Lösungen des Differenzialgleichungssystems.

(iii) Falls $A$ nicht diagonalisierbar ist, so ist es dennoch möglich $e^{At}$ effektiv zu berechnen. Man bringt dafür die Matrix auf Jordansche Normalform. In jedem Fall kann ein homogenes System linearer Differenzialgleichungen mit konstanten Koeffizienten mit Methoden der linearen Algebra vollständig gelöst werden.

*Beispiel 4.56.* Betrachte das Anfangswertproblem

$$\begin{aligned} y' &= z \\ z' &= -y \\ y(0) &= 1 \\ z(0) &= 0. \end{aligned}$$

Die zugehörige Matrix

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{4.5}$$

ist im Reellen nicht diagonalisierbar, sie besitzt jedoch im Komplexen die Eigenwerte i und $-\mathrm{i}$ mit den zugehörigen Eigenvektoren

$$v_1 = \begin{pmatrix} 1+\mathrm{i} \\ -1+\mathrm{i} \end{pmatrix} \text{ und } v_2 = \begin{pmatrix} 1-\mathrm{i} \\ -1-\mathrm{i} \end{pmatrix}.$$

Soll die allgemeine (komplexe) Lösung $c_1 e^{\mathrm{i}x} v_1 + c_2 e^{-\mathrm{i}x} v_2$ die Anfangsbedingungen erfüllen, so muss gelten

$$\begin{aligned} (1+\mathrm{i})c_1 + (1-\mathrm{i})c_2 &= 1 \\ (-1+\mathrm{i})c_1 + (-1-\mathrm{i})c_2 &= 0, \end{aligned}$$

also $c_1 = \frac{1}{4}(1-\mathrm{i})$ und $c_2 = \frac{1}{4}(1+\mathrm{i})$.

Damit ergibt sich als Lösung

$$\begin{aligned} y(x) &= \frac{1}{4}(1-\mathrm{i})e^{\mathrm{i}x}(1+\mathrm{i}) + \frac{1}{4}(1+i)e^{-\mathrm{i}x}(1-\mathrm{i}) = \frac{1}{2}e^{\mathrm{i}x} + \frac{1}{2}e^{-\mathrm{i}x} = \cos(x) \\ z(x) &= \frac{1}{4}(1-\mathrm{i})e^{\mathrm{i}x}(-1+\mathrm{i}) + \frac{1}{4}(1+i)e^{-\mathrm{i}x}(-1-\mathrm{i}) = \frac{\mathrm{i}}{2}e^{\mathrm{i}x} - \frac{\mathrm{i}}{2}e^{-\mathrm{i}x} = -\sin(x). \end{aligned}$$

Als wichtigen Spezialfall wollen wir zuletzt noch die lineare Differenzialgleichungen höherer Ordnung mit konstanten Koeffizienten behandeln.

**Satz 4.57.** *Seien $a_0, \dots, a_{n-1} \in \mathbb{C}$, $\lambda_1, \dots, \lambda_k$ die Nullstellen des Polynoms*

$$a_0 + a_1\lambda + \dots + a_{n-1}\lambda^{n-1} + \lambda^n$$

*und $\nu_1, \dots, \nu_k$ die zugehörigen Vielfachheiten. Dann bilden die Funktionen $x \mapsto x^l e^{\lambda_i x}$ mit $i \in \{1, \dots, k\}$ und $l \in \{0, \dots, \nu_i - 1\}$ ein Fundamentalsystem für die Differenzialgleichung*

$$a_0 y + a_1 y' + \dots + a_{n-1} y^{(n-1)} + y^{(n)} = 0.$$

*Ohne Beweis.*

# 5 Maße und Integrale (Fortsetzung)

## 5.1 Der Satz von Lebesgue

In diesem Abschnitt geht es um die Frage, unter welchen Bedingungen Integrale mit anderen Grenzwertprozessen vertauscht werden können. Die Notwendigkeit solcher Betrachtungen ergibt sich aus dem folgenden Beispiel.

*Beispiel 5.1.* Sei $\lambda$ das Lebesgue-Maß. Betrachte die Funktionenfolge $f_n : \mathbb{R} \to \mathbb{R}$, $f_n = \mathbb{1}_{[n,n+1)}$. Dann konvergiert $f_n$ offensichtlich punktweise gegen die Nullfunktion, es gilt aber

$$\lim_{n\to\infty} \int f_n \mathrm{d}\lambda = 1 \neq 0.$$

Die Funktionenfolge $g_n = \frac{1}{n}\mathbb{1}_{[n,2n)}$ konvergiert sogar gleichmäßig gegen die Nullfunktion und dennoch gilt

$$\lim_{n\to\infty} \int g_n \mathrm{d}\lambda = 1.$$

Es sei für den ganzen Abschnitt $(\Omega, \mathcal{A}, \mu)$ ein fest gewählter Maßraum. Wird nichts anderes gesagt, dann bilden Funktionen von $\Omega$ nach $\mathbb{R}$ ab.

**Satz 5.2 (Satz von Beppo Levi).** *Sei $(f_n)$ eine monoton wachsende Folge messbarer Funktionen und $f$ ihr (punktweiser) Grenzwert. Dann gilt*

$$\int f \mathrm{d}\mu = \lim_{n\to\infty} \int f_n \mathrm{d}\mu.$$

*Insbesondere ist $f$ integrierbar, genau dann wenn*

$$\lim_{n\to\infty} \int f_n \mathrm{d}\mu < \infty.$$

*Beweis.* Da die Funktionenfolge monoton wächst gilt $f_n \leq f$ für jedes $n \in \mathbb{N}$ und aus der Monotonie des Integrals folgt

$$\lim_{n\to\infty} \int f_n \mathrm{d}\mu \leq \int f \mathrm{d}\mu.$$

Sei nun $c > 1$. Definiere die Mengen

$$B_n = [cf_n \geq f] = [cf_n - f \geq 0] = (cf_n - f)^{-1} [0, \infty] .$$

Die letzte Darstellung zeigt, dass $B_n$ messbar ist (da die Funktion $cf_n - f$ messbar ist). Es gilt $B_n \subset B_{n+1}$ und (Warum?)

$$\bigcup_{n\in\mathbb{N}} B_n = \Omega.$$

Für die zugehörigen charakteristischen Funktionen heißt das, das die Folge $(\mathbb{1}_{B_n})$ monoton wachsend ist und (punktweise) gegen die $\mathbb{1}$ konvergiert.

Sei nun $(g_n) \subset \mathbf{X}^+(\Omega, \mathcal{A})$ eine monoton wachsende Folge einfacher Funktionen die gegen $f$ konvergiert. Dann ist auch $(g_n \mathbb{1}_{B_n})$ eine monoton wachsende Folge einfacher Funktionen die gegen $f$ konvergiert und es gilt

$$g_n \mathbb{1}_{B_n} \leq f \mathbb{1}_{B_n} \leq c f_n \mathbb{1}_{B_n} \leq c f_n.$$

Für die Integrale bedeutet das

$$\int f \mathrm{d}\mu = \lim_{n\to\infty} \int g_n \mathbb{1}_{B_n} \mathrm{d}\mu \leq c \lim_{n\to\infty} \int f_n \mathrm{d}\mu.$$

Das erste Gleichheitszeichen entspricht dabei der Definition des Integrals. Da $c > 1$ beliebig ist, folgt die Aussage des Satzes. □

*Beispiel 5.3.* Die Funktion

$$\begin{aligned} f : [1, \infty) &\to \mathbb{R} \\ x &\mapsto x^\kappa \end{aligned}$$

ist integrierbar (bezüglich des Lebesgue-Maßes) genau dann, wenn $\kappa < -1$.

*Beweis.* Die Funktion $f$ ist integrierbar genau dann, wenn

$$\int_{[1,\infty)} |f| \, \mathrm{d}\lambda = \int_{[1,\infty)} f \mathrm{d}\lambda < \infty.$$

Betrachte die nichtnegativen Funktionen $f_n = f \mathbb{1}_{[1,n]}$. Die Folge $f_n$ ist monoton wachsend und konvergiert punktweise gegen $f$. Die Funktion $f$ besitzt die Stammfunktion

$$F(x) = \begin{cases} \frac{1}{\kappa+1} x^{\kappa+1} & \kappa \neq -1 \\ \ln(x) & \kappa = -1. \end{cases}$$

Die Integrale können wir nun mit Hilfe des Hauptsatzes berechnen:

$$\int_{[1,\infty]} f \mathbb{1}_{[1,n]} \mathrm{d}\lambda = \int_1^n x^\kappa \mathrm{d}x = F(n) - F(1)$$

und mit dem Satz von Beppo Levi folgt

$$\int_{[1,\infty)} f \mathrm{d}\lambda = \lim_{n\to\infty} \int f \mathbb{1}_{[1,n]} \mathrm{d}\lambda = \lim_{n\to\infty} F(n) - F(1).$$

Die rechte Seite konvergiert für $\kappa < -1$ gegen $-F(1) = -\frac{1}{\kappa+1}$ und divergiert andernfalls bestimmt gegen $+\infty$. □

*Beispiel 5.4.* Die Funktion $\frac{1}{\ln(x)}$ ist auf $[2, \infty)$ nicht integrierbar. Da $x \geq \ln x$ ist, gilt

$$\infty = \int_2^\infty \frac{1}{x} \mathrm{d}x \leq \int_2^\infty \frac{1}{\ln(x)} \mathrm{d}x.$$

**Korollar 5.5.** *Sei $f_n$ eine Folge nichtnegativer Funktionen, dann gilt*

$$\int \sum_{n=0}^{\infty} f_n \mathrm{d}\mu = \sum_{n=0}^{\infty} \int f_n \mathrm{d}\mu.$$

*Beweis.* Die Partialsummen

$$s_n := \sum_{k=0}^{n} f_k$$

bilden eine monoton wachsende Folge nichtnegativer Funktionen. Dann gilt

$$\int \sum_{n=0}^{\infty} f_n \mathrm{d}\mu = \int \lim_{n\to\infty} s_n \mathrm{d}\mu = \lim_{n\to\infty} \int s_n \mathrm{d}\mu = \lim_{n\to\infty} \sum_{k=0}^{n} \int f_n \mathrm{d}\mu = \sum_{n=0}^{\infty} f_n \mathrm{d}\mu. \qquad \square$$

**Theorem 5.6 (Satz von Lebesgue).** *Sei $(f_n)$ eine Folge messbarer Funktionen die (fast überall) punktweise gegen $f$ konvergiert. Sei $g$ eine integrierbare Funktion und es gelte $|f_n| \leq g$ für alle $n \in \mathbb{N}$ (fast überall). Dann ist*

$$\lim_{n\to\infty} \int f_n \mathrm{d}\mu = \int f \mathrm{d}\mu.$$

*Beweis.* Wir definieren die Hilfsfunktionen

$$g_n = \inf_{k\geq n} f_k \text{ und } h_n = \sup_{k\geq n} f_k.$$

Dann ist die Folge $(g_n)$ monoton wachsend und $(h_n)$ monoton fallend. Außerdem gilt $g_n \to f$ und $h_n \to f$. Sei dazu $\omega \in \Omega$. Wir zeigen die Aussage für $f(\omega) \in \mathbb{R}$. Dann gilt für $\epsilon > 0$, dass ein $N \in \mathbb{N}$ existiert so, dass $f_k(\omega) \in [f(\omega) - \epsilon, f(\omega) + \epsilon]$ für $k > N$. Dann gilt für $n > N$ auch $g_n(\omega), h_n(\omega) \in [f(\omega) - \epsilon, f(\omega) + \epsilon]$.

Punktweise gilt die Ungleichung

$$-g \leq g_n \leq f_n \leq h_n \leq g.$$

Insbesondere sind $|g_n| \leq g$ und $|h_n| \leq g$. Damit sind $f_n, g_n$ und $h_n$ integrierbar. Wegen

$$|f| = \lim_{n\to\infty} |f_n| \leq g$$

ist auch $f$ integrierbar.

Wir wenden jetzt den Satz von Beppo Levi auf die Funktionenfolgen $(g - h_n)$ beziehungsweise $(g + g_n)$ an. Dann folgt

$$\int g \mathrm{d}\mu - \lim_{n\to\infty} \int h_n \mathrm{d}\mu = \lim_{n\to\infty} \int (g - h_n) \mathrm{d}\mu = \int (g - f) \mathrm{d}\mu = \int g \mathrm{d}\mu - \int f \mathrm{d}\mu$$
$$\int g \mathrm{d}\mu + \lim_{n\to\infty} \int g_n \mathrm{d}\mu = \lim_{n\to\infty} \int (g + g_n) \mathrm{d}\mu = \int (g + f) \mathrm{d}\mu = \int g \mathrm{d}\mu + \int f \mathrm{d}\mu$$

also

$$\int f \mathrm{d}\mu = \lim_{n\to\infty} \int h_n \mathrm{d}\mu = \lim_{n\to\infty} \int g_n \mathrm{d}\mu.$$

Andererseits folgt aus der Monotonie des Integrals

$$\lim_{n\to\infty} \int g_n \mathrm{d}\mu \le \lim_{n\to\infty} \int f_n \mathrm{d}\mu \le \lim_{n\to\infty} \int h_n \mathrm{d}\mu.$$

Falls $f_n \to f$ und $|f_n| \le g$ nur fast überall gilt, dann sei $\Omega' \subset \Omega$ eine messbare Menge mit $\mu(\Omega \backslash \Omega') = 0$ so, dass $f_n(\omega) \to f(\omega)$ und $|f_n(\omega)| \le g(\omega)$ für alle $\omega \in \Omega'$ und alle $n \in \mathbb{N}$ gilt. Wir wenden dann den bisher bewiesenen Satz auf die Funktionenfolge $\tilde{f}_n := f_n \mathbb{1}_{\Omega'}$ und $\tilde{f} := f \mathbb{1}_{\Omega'}$ an. Dann gilt nach Bemerkung 3.49

$$\lim_{n\to\infty} \int f_n \mathrm{d}\mu = \lim_{n\to\infty} \int \tilde{f}_n \mathrm{d}\mu = \int \tilde{f} \mathrm{d}\mu = \int f \mathrm{d}\mu. \qquad \square$$

*Bemerkung 5.7.* Anstatt der Existenz einer integrierbaren Funktion $g$ mit $|f_n| \le g$ kann man auch konkreter fordern, dass die Funktion

$$\sup_{n\in\mathbb{N}} |f_n| \tag{5.1}$$

integrierbar ist.

Da sich die auftretenden Grenzwerte nicht ändern reicht es aus, $|f_n| \le g$ für hinreichend große $n$ zu fordern.

**Korollar 5.8.** *Sei $a \in \mathbb{R}$, $f : [a, \infty) \to \mathbb{R}$ eine stetige Funktion und $F$ eine Stammfunktion von $f$. Falls $f$ nichtnegativ oder integrierbar ist, so gilt*

$$\int_{[a,\infty)} f \mathrm{d}\lambda = \lim_{n\to\infty} F(n) - F(a).$$

*Beweis.* Für alle $n > a$ betrachte $f_n = f\mathbb{1}_{[a,n)}$. Falls $f$ nichtnegativ ist, dann ist $(f_n)$ monoton wachsend. In jedem Fall gilt $|f_n| \le |f|$. Für nichtnegative $f$ können wir also den Satz von Beppo Levi, für integrierbare $f$ den Satz von Lebesgue anwenden und erhalten

$$\int_{[a,\infty)} f \mathrm{d}\lambda = \lim_{n\to\infty} \int_a^n f(x)\mathrm{d}x = \lim_{n\to\infty} F(n) - F(a). \qquad \square$$

*Bemerkung 5.9.* Die Existenz des Grenzwertes $\lim_{n\to\infty} F(n) - F(a)$ allein ist nicht ausreichend.

**Korollar 5.10.** *Sei $f_n$ eine Folge von Funktionen. Falls*

$$\int \sum_{n=0}^{\infty} |f_n| \,\mathrm{d}\mu < \infty \text{ oder } \sum_{n=0}^{\infty} \int |f_n| \,\mathrm{d}\mu < \infty,$$

*dann gilt*

$$\int \sum_{n=0}^{\infty} f_n \mathrm{d}\mu = \sum_{n=0}^{\infty} \int f_n \mathrm{d}\mu.$$

*Beweis.* Nach Korollar 5.5 gilt

$$\int \sum_{n=0}^{\infty} |f_n| \,\mathrm{d}\mu = \sum_{n=0}^{\infty} \int |f_n| \,\mathrm{d}\mu < \infty$$

und für die Folge der Partialsummen $s_n = \sum_{k=0}^{n} f_k$ gilt

$$|s_n| \le \sum_{k=0}^{n} |f_k| \le \sum_{k=0}^{\infty} |f_k| \,.$$

Die Funktion auf der rechten Seite ist nach Voraussetzung integrierbar und wir können den Satz von Lebesgue anwenden:

$$\int \lim_{n\to\infty} s_n \mathrm{d}\mu = \lim_{n\to\infty} \sum_{k=0}^{n} \int f_n \mathrm{d}\mu = \sum_{n=0}^{\infty} \int f_n \mathrm{d}\mu. \qquad \square$$

*Beispiel 5.11.* Betrachte die Funktion

$$\begin{aligned} F : (0,\infty) &\to \mathbb{R} \\ a &\mapsto \int_0^{\infty} \frac{e^{-x}}{x+a} \mathrm{d}x. \end{aligned}$$

Ist $F$ stetig? Da wir das Integral nicht explizit ausrechnen können, ist diese Frage zunächst schwierig zu beantworten.

**Korollar 5.12.** *Sei $X$ ein metrischer Raum, $f : \Omega \times X \to \mathbb{R}$ und $\tilde{a} \in X$. Sei $f(\,\cdot\,, a)$ integrierbar für jedes $a \in X$ und $f(\omega, \cdot\,)$ stetig in $\tilde{a}$ für alle $\omega \in \Omega$. Sei $U$ eine Umgebung von $\tilde{a}$ und $g$ eine integrierbare Funktion (unabhängig von $a \in X$) so, dass $|f(\omega, a)| \le g(\omega)$ für alle $\omega \in \Omega$ und alle $a \in U$, dann ist die Funktion*

$$\begin{aligned} F : X &\to \mathbb{R} \\ a &\mapsto \int f(\omega, a) \mathrm{d}\mu(\omega) \end{aligned}$$

*stetig in $\tilde{a}$.*

*Beweis.* Sei $(a_n)$ eine Folge mit $a_n \to \tilde{a}$. Setze $f_n := f(\,\cdot\,, a_n)$. Für hinreichend große $n$ ist dann $a_n \in U$ und es gilt $|f_n| = |f(\,\cdot\,, a_n)| \le g$. Nach Voraussetzung gilt für $\omega \in \Omega$, dass

$$\lim_{n\to\infty} f_n(\omega) = \lim_{n\to\infty} f(\omega, a_n) = f(\omega, \tilde{a}).$$

Nach dem Satz von Lebesgue gilt also

$$\lim_{n\to\infty} F(a_n) = \lim_{n\to\infty} \int f_n(\omega) \mathrm{d}\mu(\omega) = \int \lim_{n\to\infty} f_n(\omega) \mathrm{d}\mu(\omega) = \int f(\omega, \tilde{a}) \mathrm{d}\mu = F(\tilde{a}).$$

Da die Folge $(a_n)$ beliebig war folgt

$$\lim_{a\to\tilde{a}} F(a) = F(\tilde{a}),$$

also die Stetigkeit von $F$ im Punkt $\tilde{a}$. $\square$

*Beispiel 5.13.* Im obigen Beispiel sei jetzt also $\tilde{a} \in (0, \infty)$. Es gilt für alle $a \in \left(\frac{\tilde{a}}{2}, \infty\right)$ und für alle $x \in [0, \infty]$

$$\frac{e^{-x}}{x+a} \leq \frac{2}{\tilde{a}} e^{-x}.$$

Da die Funktion auf der rechten Seite über $[0, \infty)$ integrierbar ist und $\left(\frac{\tilde{a}}{2}, \infty\right)$ eine Umgebung von $\tilde{a}$ ist, sind die Voraussetzungen des Korollars erfüllt und somit die Funktion $F$ stetig in $\tilde{a}$.

Da $\tilde{a}$ beliebig war, ist die Funktion überall stetig.

**Korollar 5.14.** *Sei $X \subset \mathbb{R}^n$ offen, $f : \Omega \times X \to \mathbb{R}$ und $\tilde{a} \in X$. Sei $f(\,\cdot\,, a)$ integrierbar für jedes $a \in X$ Sei $U$ eine Umgebung von $\tilde{a}$ und $f(\omega, \cdot\,)$ differenzierbar für alle $\omega \in \Omega$ in jedem Punkt $a \in U$. Sei $g$ eine integrierbare Funktion (unabhängig von $a \in X$) so, dass*

$$\|\mathrm{D}_a f(\omega, a)\| \leq g(\omega)$$

*für alle $\omega \in \Omega$ und alle $a \in U$, dann ist die Funktion*

$$\begin{aligned} F : X &\to \mathbb{R} \\ a &\mapsto \int f(\omega, a) \mathrm{d}\mu(\omega) \end{aligned}$$

*differenzierbar in $\tilde{a}$ und es gilt*

$$\mathrm{D}F(\tilde{a}) = \int \mathrm{D}_a F(\omega, \tilde{a}) \mathrm{d}\mu(\omega).$$

*Ohne Beweis.*

*Beispiel 5.15.* Wir betrachten weiter das vorhergehende Beispiel. Der Integrand ist nach $a$ differenzierbar. Sei $\tilde{a} \in (0, \infty)$, dann gilt für $x \in [0, \infty]$ und $a \in \left(\frac{\tilde{a}}{2}, \infty\right)$

$$\left| \frac{\mathrm{d}}{\mathrm{d}a} \frac{e^{-x}}{x+a} \right| = \frac{e^{-x}}{(x+a)^2} \leq \frac{4}{\tilde{a}^2} e^{-x}.$$

Damit sind die Voraussetzungen des Korollars erfüllt. Die Funktion $F$ ist also in $\tilde{a}$ differenzierbar und es gilt

$$F'(\tilde{a}) = -\int_0^\infty \frac{e^x}{(x+\tilde{a})^2} \mathrm{d}x.$$

Da $\tilde{a}$ beliebig war, ist $F$ für jedes $a \in (0, \infty)$ differenzierbar.

## 5.2 Produktmaße und der Satz von Fubini

Wir beschäftigen uns zunächst mit der Konstruktion von Produktmaßen, insbesondere des Lebesgue-Maßes auf $\mathbb{R}^n$.

**Definition 5.16 (endliche und $\sigma$-endliche Maßräume).** Ein Maßraum $(\Omega, \mathcal{A}, \mu)$ heißt endlich, falls $\mu(\Omega) < \infty$ und $\sigma$-endlich falls es eine Folge $(A_n) \subset \mathcal{A}$ gibt, so dass $\mu(A_n) < \infty$ für jedes $n \in \mathbb{N}$ und

$$\bigcup_{n\in\mathbb{N}} A_n = \Omega.$$

*Beispiel 5.17.* (i) Das Intervall $[0, 1]$ mit dem Lebesgue-Maß ist endlich.

(ii) Ganz $\mathbb{R}$ mit dem Lebesgue-Maß ist nicht endlich, aber $\sigma$-endlich.

(iii) Jede Abzählbare Menge (z.B. $\mathbb{N}$) mit dem Zählmaß ist $\sigma$-endlich.

(iv) Jede Überabzählbare Menge (z.B. $\mathbb{R}$) mit dem Zählmaß ist nicht $\sigma$-endlich.

**Definition 5.18.** Seien $(\Omega, \mathcal{A})$ und $(\Phi, \mathcal{B})$ Messräume. Dann heißt

$$\mathcal{A} \otimes \mathcal{B} := \sigma\left(\{A \times B |\ A \in \mathcal{A} \text{ und } B \in \mathcal{B}\}\right) \subset \mathcal{P}\left(\Omega \times \Phi\right)$$

die Produkt-$\sigma$-algebra.

Seien $\mu$ auf $(\Omega, \mathcal{A})$ und $\nu$ auf $(\Phi, \mathcal{B})$ Maße. Dann heißt ein Maß $\theta$ auf $(\Omega \times \Phi, \mathcal{A} \otimes \mathcal{B})$ Produktmaß von $\mu$ und $\nu$ falls für alle $A \in \mathcal{A}$ und $B \in \mathcal{B}$

$$\theta(A \times B) = \mu(A)\mu(B).$$

Für den Rest dieses Abschnitts seien $(\Omega, \mathcal{A}, \mu)$ und $(\Phi, \mathcal{B}, \nu)$ $\sigma$-endliche Maßräume.

**Satz 5.19.** *Auf $(\Omega \times \Phi, \mathcal{A} \otimes \mathcal{B})$ existiert ein eindeutiges Produktmaß welches wir mit $\mu \otimes \nu$ bezeichnen.*

*Beweisskizze.* Der Beweis verläuft analog zur Konstruktion des Lebesgue-Maßes. Zunächst definiere

$$\begin{aligned} \theta : \mathcal{P}\left(\Omega \times \Phi\right) &\to \overline{\mathbb{R}}^+ \\ M &\mapsto \inf\left\{ \sum_{i\in\mathbb{N}} \mu(A_i)\nu(B_i) \,\middle|\, M \subset \bigcup_{i\in\mathbb{N}} A_i \times B_i, A_i \in \mathcal{A}, B_i \in \mathcal{B} \forall i \in \mathbb{N} \right\} \end{aligned}$$

und zeige, dass $\theta$ ein äußeres Maß ist. Dann folgt aus Satz 3.8, dass

$$\mathcal{D} := \{M \subset \Omega \times \Phi |\ \theta(E) \geq \theta(E \cap M) + \theta(E \cap M^c)\ \forall E \in \mathcal{P}\left(\Omega \times \Phi\right)\}$$

eine $\sigma$-Algebra und die Einschränkung von $\theta$ auf diese Menge ein Maß ist.

Man zeigt nun, dass für $A \in \mathcal{A}$ und $B \in \mathcal{B}$ die Menge $A \times B$ in $\mathcal{D}$ enthalten ist. Damit gilt $\mathcal{A} \otimes \mathcal{B} \subset \mathcal{D}$. Schließlich ist zu zeigen, dass $\theta(A \times B) = \mu(A)\nu(B)$.

Um zu zeigen, dass das Produktmaß eindeutig bestimmt ist sei $\tilde{\theta}$ ein weiteres Produktmaß. Dann gilt für $A \in \mathcal{A}$ und $B \in \mathcal{B}$, dass $\theta(A \times B) = \mu(A)\nu(B) = \tilde{\theta}(A \times B)$. Man zeigt dann, dass die Menge

$$\left\{ M \subset \mathcal{A} \otimes \mathcal{B} \,\middle|\, \theta(M) = \tilde{\theta}(M) \right\}$$

eine $\sigma$-Algebra ist (Aufgabe 4 vom Blatt 6) und damit gleich $\mathcal{A} \otimes \mathcal{B}$. □

*Bemerkung 5.20.* (i) Man kann zeigen, dass $(\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C} = \mathcal{A} \otimes (\mathcal{B} \otimes \mathcal{C})$ für $\sigma$-Algebren $\mathcal{A}$, $\mathcal{B}$ und $\mathcal{C}$ sowie $(\mu \otimes \nu) \otimes \theta = \mu \otimes (\nu \otimes \theta)$ für Maße $\mu$, $\nu$ und $\theta$ gilt. Damit können wir auch mehrfache Produktmaßräume definieren.

(ii) Insbesondere erhalten wir durch wiederholte Anwendung des obigen Satzes aus dem eindimensionalen Lebesgue-Maß $\lambda$ die mehrdimensionalen Varianten, die wir mit $\lambda^n = \lambda \otimes \cdots \otimes \lambda$ bezeichnen.

(iii) Die Borelsche $\sigma$-Algebra (Beispiel 3.14) des $\mathbb{R}^n$ ist das $n$-fache Produkt der Borelschen $\sigma$-Algebra von $\mathbb{R}$. Es sind also insbesondere alle offenen und abgeschlossenen Mengen $\lambda^n$-messbar (Satz 3.17).

(iv) Analog zum eindimensionalen Fall ist $\lambda^n$ das einzige translationsinvariante Maß auf $\mathbb{R}^n$, das $\lambda^n([0,1)^n) = 1$ erfüllt.

(v) Im folgenden werden oft Mehrfachintegrale von Funktionen mehrerer Veränderlicher auftauchen. In diesen Fällen kann die variablenfreie Schreibweise von Integralen

$$\int f \mathrm{d}\mu$$

zu Uneindeutigkeiten führen und wir werden daher die Schreibweise

$$\int f(\omega) \mathrm{d}\mu(\omega)$$

verwenden. In anderer Literatur wir teilweise auch die Schreibweise

$$\int f(\omega) \mu(\mathrm{d}\omega)$$

verwendet.

*Beispiel 5.21.* Betrachte die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f = \sum_{n=0}^{\infty} \left( \mathbb{1}_{[n,n+1)^2} - \mathbb{1}_{[n+1,n+2) \times [n,n+2)} \right).$$

Dann gilt

$$\int \int f(x,y)\mathrm{d}x\mathrm{d}y = 0 \text{ aber } \int \int f(x,y)\mathrm{d}y\mathrm{d}x = 1.$$

Bezüglich $\lambda^2$ ist die Funktion gar nicht integrierbar da offensichtlich

$$\int |f(x,y)| \,\mathrm{d}\lambda^2(x,y) = \infty$$

gilt.

**Satz 5.22 (Satz von Tonelli).** *Sei $f : \Omega \times \Phi \to \overline{\mathbb{R}}^+$ messbar. Dann sind die Funktionen $\omega \mapsto f(\omega,\varphi)$ für fast alle $\varphi \in \Phi$ und $\varphi \mapsto f(\omega,\varphi)$ für fast alle $\omega \in \Omega$ messbar. Außerdem sind die Funktionen*

$$\varphi \mapsto \int f(\omega,\varphi)\mathrm{d}\mu(\omega) \text{ und } \omega \mapsto \int f(\omega,\varphi)\mathrm{d}\nu(\varphi)$$

*messbar. Es gilt*

$$\int f(\omega,\varphi)\mathrm{d}(\mu\times\nu)(\omega,\varphi) = \int\int f(\omega,\varphi)\mathrm{d}\mu(\omega)\mathrm{d}\nu(\varphi) = \int\int f(\omega,\varphi)\mathrm{d}\nu(\varphi)\mathrm{d}\mu(\omega)$$

*Insbesondere ist $f$ integrierbar falls entweder*

$$\int\int f(\omega,\varphi)\mathrm{d}\mu(\omega)\mathrm{d}\nu(\varphi) \text{ oder } \int\int f(\omega,\varphi)\mathrm{d}\nu(\varphi)\mathrm{d}\mu(\omega)$$

*endlich sind.*

*Ohne Beweis.*

**Korollar 5.23 (Prinzip des Cavalieri).** *Sei $A \in \mathcal{A}\otimes\mathcal{B}$. Bezeichne mit $A_\omega$ die Schnitte von $A$, das heißt*

$$A_\omega = \{\varphi \in \Phi |\ (\omega,\varphi) \in A\}.$$

*Dann ist $\omega \mapsto \nu(A_\omega)$ messbar und es gilt*

$$(\mu\otimes\nu)(A) = \int \nu(A_\omega)\mathrm{d}\mu(\omega).$$

*Beweis.* Nach der Definition von $A_\omega$ gilt, dass $(\omega,\varphi) \in A$ ist, genau dann wenn $\varphi \in A_\omega$ ist also gilt $\mathbb{1}_A(\omega,\varphi) = \mathbb{1}_{A_\omega}(\varphi)$. Wir wenden den Satz von Tonelli auf die Funktion $\mathbb{1}_A$ an. Dann folgt

$$\begin{aligned}(\mu\otimes\nu)(A) &= \int \mathbb{1}_A(\omega,\varphi)\mathrm{d}(\mu\otimes\nu)(\omega,\varphi) = \int\int \mathbb{1}_A(\omega,\varphi)\mathrm{d}\nu(\varphi)\mathrm{d}\mu(\omega)\\ &= \int\int \mathbb{1}_{A_\omega}(\varphi)\mathrm{d}\nu(\varphi)\mathrm{d}\mu(\omega) = \int \nu(A_\omega)\mathrm{d}\mu(\omega).\end{aligned}$$ □

*Beispiel 5.24.* Sei $V$ ein höchstens $n-1$ dimensionaler Untervektorraum von $\mathbb{R}^n$. Dann ist $\lambda^n(V) = 0$.

*Beweis.* Es gibt mindestens einen Vektor der Standardbasis, der nicht in $V$ enthalten ist (Warum?). Durch Umsortieren können wir ohne Einschränkung der Allgemeinheit annehmen, dass $e_1$ dieser Vektor ist und zerlegen $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$. Die Schnitte von $V$ bezüglich dieser Zerlegung, also die Mengen

$$V_{(x_2,\ldots,x_n)} = \{x_1 \in \mathbb{R} |\ (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n\}$$

enthalten für beliebige Werte von $(x_2, \ldots, x_n)$ höchstens einen Punkt (Warum?) und haben damit $\lambda$-Maß 0. Damit gilt

$$\lambda^n(V) = \int \lambda(V_{(x_2,\ldots,x_n)}) \mathrm{d}\lambda^{n-1}(x_2, \ldots, x_n) = 0 \qquad \square$$

**Korollar 5.25.** *Sei $f : \Omega \to \overline{\mathbb{R}}^+$ eine messbare Funktion und*

$$A_f := \{(\omega, y) \in \Omega \times \mathbb{R} |\ 0 \leq y \leq f(\omega)\} \subset \Omega \times \mathbb{R}.$$

*Dann ist $A_f$ messbar und es gilt*

$$(\mu \otimes \lambda)(A_f) = \int f(\omega) \mathrm{d}\mu(\omega).$$

*Beweis.* Es gilt $\mathbb{1}_{A_f}(\omega, y) = \mathbb{1}_{[0,f(\omega)]}(y)$ und damit

$$\begin{aligned}(\mu \otimes \lambda)(A_f) &= \int \mathbb{1}_{A_f}(\omega, y) \mathrm{d}(\mu \otimes \lambda)(\omega, y) = \int \int \mathbb{1}_{[0,f(\omega)]} \mathrm{d}\lambda(y) \mathrm{d}\mu(\omega) \\ &= \int f(\omega) \mathrm{d}\mu(\omega). \qquad \square\end{aligned}$$

Das Integral über $f$ entspricht also tatsächlich der Fläche unter dem Graphen der Funktion.

**Theorem 5.26 (Satz von Fubini).** *Sei $f : \Omega \times \Phi$ eine messbare Funktion, die bezüglich $\mu \otimes \nu$ integrierbar ist, dann sind die Funktionen $\omega \mapsto f(\omega, \varphi)$ messbar und integrierbar für $\mu$-fast alle $\varphi \in \Phi$, die Funktionen $\varphi \mapsto f(\omega, \varphi)$ messbar und integrierbar für $\nu$-fast alle $\omega \in \Omega$. Die Funktionen*

$$\omega \mapsto \int f(\omega, \varphi) \mathrm{d}\nu(\varphi) \text{ und } \varphi \mapsto \int f(\omega, \varphi) \mathrm{d}\mu(\omega)$$

*sind messbar und integrierbar und es gilt*

$$\int f(\omega, \varphi) \mathrm{d}(\mu \otimes \nu)(\omega, \varphi) = \int \int f(\omega, \varphi) \mathrm{d}\nu(\varphi) \mu(\omega) = \int \int f(\omega, \varphi) \mathrm{d}\mu(\omega) \mathrm{d}\nu(\varphi).$$

*Ohne Beweis.*

*Bemerkung 5.27.* Um die Integrierbarkeit bezüglich des Produktmaßes zu überprüfen wird häufig der Satz von Tonelli verwendet. Auch die Aussage des Satzes von Fubini lässt sich unmittelbar auf Produkte von mehr als zwei Maßen und entsprechende Integrale verallgemeinern.

**Korollar 5.28.** *Seien $a_i, b_i \in \mathbb{R}$ und $a_i < b_i$ für $i \in \{1, \dots, n\}$ und setze $D = [a_1, b_1] \times \cdots \times [a_n, b_n]$. Sei $f : D \to \mathbb{R}$ eine stetige Funktion, dann gilt*

$$\int_D f \mathrm{d}\lambda^n = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \dots, x_n) \mathrm{d}x_n \cdots \mathrm{d}x_1.$$

*Die Integrale können dabei in beliebiger Reihenfolge ausgeführt werden.*

*Beweis.* Da $D$ beschränkt und abgeschlossen ist und $f$ stetig, ist $f$ durch ein $K \in \mathbb{R}$ beschränkt und es gilt

$$\int_D |f| \, \mathrm{d}\lambda^n = \int_D K \mathrm{d}\lambda^n = K\lambda^n(D) < \infty.$$

Damit ist $f$ integrierbar und wir können wiederholt den Satz von Fubini anwenden und erhalten

$$\begin{aligned} \int_D f \mathrm{d}\lambda^n &= \int f(x_1, \dots, x_n) \mathbb{1}_D(x_1, \dots, x_n) \mathrm{d}\lambda^n \\ &= \int \cdots \int f(x_1, \dots, x_n) \mathbb{1}_{[a_1, b_1]}(x_1) \cdots \mathbb{1}_{[a_n, b_n]}(x_n) \mathrm{d}\lambda(x_n) \cdots \mathrm{d}\lambda(x_1) \\ &= \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \dots, x_n) \mathrm{d}x_n \cdots \mathrm{d}x_1 \end{aligned}$$

□

## 5.3 Die Transformationsformel

**Definition 5.29.** Seien $(X, m_X), (Y, m_Y)$ metrische Räume. Dann heißt $f : X \to Y$ gleichmäßig stetig, wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert so, dass $m_Y(f(x), f(y)) < \epsilon$ für alle $x, y \in X$ mit $m_X(x, y) < \delta$.

*Bemerkung 5.30.* Stetigkeit auf ganz $X$:

$$\forall \epsilon > 0 \forall x \in X \exists \delta > 0 \forall y \in X : m_X(x, y) < \delta \Longrightarrow m_Y(f(x), f(y)) < \epsilon.$$

Gleichmäßige Stetigkeit auf $X$:

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in X \forall y \in X : m_X(x, y) < \delta \Longrightarrow m_Y(f(x), f(y)) < \epsilon.$$

**Satz 5.31.** *Sei $D \subset \mathbb{R}^n$ abgeschlossen und beschränkt, $(Y, m)$ ein metrischer Raum und $f : D \to Y$ stetig, dann ist $f$ gleichmäßig stetig.*

*Beweis.* Angenommen $f$ wäre nicht gleichmäßig stetig. Dann existiert ein $\epsilon > 0$ so, dass es für jedes $\delta > 0$ Punkte $x, y \in D$ gibt so, dass $\|x - y\| < \delta$ aber $m(f(x), f(y)) \geq \epsilon$. Insbesondere existieren Folgen $(x_n), (y_n) \subset D$ so, dass $\|x_n - y_n\| < \frac{1}{n}$ und $m(f(x), f(y)) \geq \epsilon$. Da $D$ beschränkt und abgeschlossen ist, existiert eine Teilfolge $(x_{n_k})$ von $(x_n)$ die gegen $x$ konvergiert. Dann konvergiert auch $y_{n_k} = x_{n_k} + y_{n_k} - x_{n_k}$ gegen $x$. Wegen der Stetigkeit von $f$ konvergieren dann $(f(x_{n_k}))$ und $(f(y_{n_k}))$ gegen $f(x)$, was aber im Widerspruch zu $m(f(x_{n_k}), f(y_{n_k})) \geq \epsilon$ für jedes $k \in \mathbb{N}$ steht. □

**Definition 5.32.** Seien $U, V \subset \mathbb{R}^n$ offen. Eine Abbildung $T : U \to V$ heißt Diffeomorphismus wenn sie bijektiv ist und sowohl $T$ als auch $T^{-1}$ stetig differenzierbar sind. Falls $T$ und $T^{-1}$ darüber hinaus $r$-fach oder beliebig oft stetig differenzierbar sind, sprechen wir von $C^r$-Diffeomorphismen beziehungsweise $C^\infty$-Diffeomorphismen.

*Bemerkung 5.33.* (i) Diffeomorphismen, insbesondere $C^\infty$-Diffeomorphismen, beschreiben Koordinatentransformationen. Man sagt $T$ erhält die differenzierbare Struktur von $X$, das heißt vereinfacht formuliert: Eigenschaften einer Funktion die mit Hilfe ihrer Ableitungen fomuliert sind werden von $f$ genau dann erfüllt, wenn sie von $f \circ T$ erfüllt werden.

(ii) In der Physik werden die Funktionen $f$ und $f \circ T$ häufig mit dem selben Symbol bezeichnet, da sie die selbe physikalische Größe lediglich in einem anderen Koordinatensystem beschreiben. Diese Notation ist jedoch mit Vorsicht zu behandeln, da der funktionale Zusammenhang sich natürlich im Allgemeinen ändert.

(iii) Wir leiten die Gleichung $T \circ T^{-1} = \mathrm{id}_V$ ab und erhalten mittels der Kettenregel

$$\mathrm{D}T(T^{-1}(y))\mathrm{D}T^{-1}(y) = I.$$

Da $T^{-1}$ surjektiv ist besitzt $\mathrm{D}T(x)$ für jedes $x \in U$ eine Rechtsinverse. Verfährt man analog mit $T^{-1} \circ T = \mathrm{id}_U$ so folgt, dass $\mathrm{D}T(x)$ eine invertierbare Matrix ist für jedes $x \in U$. Andererseits ist nach dem Satz von der Umkehrfunktion die Funktion $T^{-1}$ stetig differenzierbar falls $\mathrm{D}T(x)$ invertierbar ist für jedes $x \in U$.

(iv) Offensichtlich ist die Umkehrfunktion $T^{-1}$ eines Diffeomorphismus wieder ein Diffeomorphismus.

*Beispiel 5.34.* Die Abbildung

$$\begin{aligned} T : (0,\infty) \times (0,2\pi) &\to \mathbb{R}^2 \setminus ((0,\infty) \times \{0\}) \\ (r,\varphi) &\mapsto (r\cos(\varphi), r\sin(\varphi)) \end{aligned}$$

ist ein $C^\infty$-Diffeomorphismus und beschreibt den Übergang zu Polarkoordinaten.

**Satz 5.35.** *Sei $A$ eine invertierbare Matrix und $F : \mathbb{R}^n \to \mathbb{R}^n$ definiert durch $F(x) = Ax$. Dann gilt $\lambda^n(F(U)) = \det(A)\lambda^n(U)$ für alle Borel-messbaren Mengen $U$.*

*Beweisskizze.* Betrachte zunächst die folgenden Abbildung auf der Borelschen $\sigma$-Algeba $\mathcal{B}$

$$\begin{aligned} \nu : \mathcal{B} &\to \overline{\mathbb{R}}^+ \\ U &\mapsto \lambda^n(F(U)). \end{aligned}$$

Man überzeugt sich leicht, dass $\nu$ ein Maß ist. Außerdem gilt für $U \in \mathcal{B}$ und $a \in \mathbb{R}^n$

$$\nu(U+a) = \lambda^n(F(U+a)) = \lambda^n(F(U) + F(a)) = \lambda^n(F(U)) = \nu(U),$$

das Maß $\nu$ ist also Translationsinvariant. Da das Lebesgue-Maß (bis auf Normierung) das einzige translationsinvariante Maß auf $\mathbb{R}^n$ ist, muss $\nu = \kappa\lambda^n$ gelten für ein $\kappa > 0$. Es bleibt also zu zeigen, dass $\kappa = \det(A)$ ist. Dazu ist es ausreichend, jeweils das Maß von einer geschickt gewählten Menge und ihrem Bild unter $F$ zu betrachten.

Wir betrachten nun zunächst den Fall, dass $A$ orthogonal ist. Dann gilt für die Euklidische Norm $\|F(x)\| = \|Ax\| = \|x\|$ für alle $x \in \mathbb{R}^n$, insbesondere ist $F(B_1(0)) = B_1(0)$. In diesem Fall gilt $\kappa\lambda^n(B_1(0)) = \nu(B_1(0)) = \lambda^n(B_1(0))$, also $\kappa = 1 = |\det A|$ da die Einheitskugel nichtverschwindendes Maß hat.

Als zweites betrachten wir $A = \operatorname{diag}(a_1, \ldots, a_n)$ mit $a_1, \ldots, a_n > 0$. Dann gilt $F([0,1]^n) = [0,a_1] \times \cdots \times [0,a_n]$ also $\kappa = a_1 \cdots a_n = |\det A|$.

Schließlich können wir für eine beliebige invertierbare Matrix schreiben $A = O_1 D O_2$ mit orthogonalen Matrizen $O_1, O_2$ und einer Diagonalmatrix mit positiven Einträgen $D$. Dann gilt nach den oben behandelten Fällen $\nu = |\det D|\,\lambda^n$ und auf Grund der Multiplikativität der Determinante $|\det A| = |\det D|$ also wiederum $\kappa = |\det A|$. □

Als wichtigen Spezialfall erhält man, dass für die Abbildung $T(x) = \rho x$ für $\rho \in \mathbb{R}$ gilt $\lambda^n(T(U)) = \rho^n \lambda^n(U)$.

**Theorem 5.36 (Transformationsformel).** *Seien $U, V \subset \mathbb{R}^n$ offen und $T : U \to V$ ein Diffeomorphismus. Dann ist $f : V \to \overline{\mathbb{R}}$ genau dann integrierbar wenn $f \circ T\,|\det \mathrm{D}T|$ über $U$ integrierbar ist und es gilt*

$$\int_V f(y)\mathrm{d}\lambda^n(y) = \int_U f \circ T(x)\,|\det \mathrm{D}T(x)|\,\mathrm{d}\lambda^n(x). \tag{5.2}$$

*Beispiel 5.37.* Wir wollen $K := \int e^{-x^2}\mathrm{d}x$ berechnen. Wir schreiben dazu

$$\begin{aligned} K^2 &= \int e^{-x^2}\mathrm{d}x \int e^{-y^2}\mathrm{d}y = \int\int e^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y \\ &= \int e^{-(x^2+y^2)}\mathrm{d}\lambda^2(x,y) = \int f(x,y)\mathrm{d}\lambda^2(x,y) \end{aligned}$$

für $f(x,y) = e^{-(x^2+y^2)}$. Da die Integranden nichtnegativ sind, konnten wir oben den Satz von Tonelli verwenden.

Wir transformieren das Integral nun auf Polarkoordinaten (Beispiel 5.34). Setze also $U = (0,\infty) \times (0, 2\pi)$ und $V = \mathbb{R}^2 \backslash ((0,\infty) \times \{0\})$ und $T(r,\varphi) = (r\cos(\varphi), r\sin(\varphi))$ für $(r,\varphi) \in U$. Man rechnet leicht nach, dass $\det \mathrm{D}T(r,\varphi) = r$ gilt. Dann gilt

$$\begin{aligned} \int_R f(x,y)\mathrm{d}\lambda^2(x,y) &= \int_D f \circ T(r,\varphi) \left|\det \mathrm{D}T(r,\varphi)\right| \mathrm{d}\lambda^2(r,\varphi) = \int e^{-r^2} r \mathrm{d}\lambda^2(r,\varphi) \\ &= \int_0^{2\pi} \mathrm{d}\varphi \int_0^\infty r e^{-r^2}\mathrm{d}r = 2\pi \lim_{n\to\infty} \left(-\frac{1}{2}e^{-n^2} + \frac{1}{2}\right) = \pi. \end{aligned}$$

Dabei haben wir nochmals den Satz von Tonelli und Korollar 5.8 verwendet. Da $((0,\infty) \times \{0\})$ eine Nullmenge bezüglich $\lambda^2$ ist gilt $K^2 = \pi$.

Wir verwenden im Folgenden die Notation des Theorems. Mit $\mathcal{B}$ bezeichnen wir die Borel-$\sigma$-Algebra auf $U$.

**Lemma 5.38.** *Sei $Q$ ein abgeschlossener Quader in $U$. Dann gilt*

$$\lambda^n(T(Q)) \leq \int_Q \left|\det \mathrm{D}T(x)\right| \mathrm{d}\lambda^n(x)$$

*Beweis.* Wir führen hier den Beweis für einen Würfel der Kantenlänge $d$, der sich jedoch ohne Weiteres auf den allgemeinen Fall verallgemeinern lässt. Die Strategie für den Beweis ist es, den Würfel in $k := m^n$ hinreichend kleine Teilwürfel $Q_1, \ldots, Q_k$ mit Kantenlänge $\frac{d}{m}$ zu zerlegen und auf diesen die Abbildung $T$ durch eine passende lineare Abbildung abzuschätzen. Wir wählen die Teilwürfel abgeschlossen. Setze

$$K = \max_{x \in Q} \left\|(\mathrm{D}T(x))^{-1}\right\|$$

wobei wir die Stetigkeit der Abbildung $A \mapsto A^{-1}$ benutzt haben.

Sei jetzt $\epsilon > 0$. Die Funktion $x \mapsto \mathrm{D}T(x)$ ist stetig auf der beschränkten und abgeschlossenen Menge $Q$ und somit gleichmäßig stetig (Satz 5.31), es gibt also $\delta > 0$ so, dass

$$\|\mathrm{D}T(x) - \mathrm{D}T(y)\| < \frac{\epsilon}{K}$$

gilt, falls $\|x - y\| < \delta$. Verkleinere nun $\delta$ und wähle $m$ so, dass der Durchmesser (Diagonale) der Teilwürfel $Q_i$ gleich $\delta$ ist, also $\sqrt{n}\frac{d}{m} = \delta$.

Wir betrachten jetzt für ein festes $i \in \{1, \dots, k\}$ den Teilwürfel $Q_i$. Wähle $x_i \in Q_i$ so, dass

$$|\det(\mathrm{D}T(x_i))| = \min_{x \in Q_i} |\det(\mathrm{D}T(x))| .$$

Für beliebiges $x \in Q_i$ setzen wir $\sigma(x) = T(x) - T(x_i) - \mathrm{D}T(x_i)(x - x_i)$ und wenden auf die Abbildung $h(x) = T(x) - \mathrm{D}T(x_i)x$ den verallgemeinerten Mittelwertsatz an um zu erhalten

$$\|\sigma(x)\| = \|h(x) - h(x_i)\| \leq \|x - x_i\| \max_{\tilde{x} \in Q_i} \|\mathrm{D}T(\tilde{x}) - \mathrm{D}T(x_i)\| \leq \delta \frac{\epsilon}{K}$$

$$\left\|\mathrm{D}T^{-1}(x_i)\sigma(x)\right\| \leq \left\|\mathrm{D}T^{-1}(x_i)\right\| \|\sigma(x)\| \leq \delta\epsilon.$$

Dann folgt für $x \in Q_i$

$$T(x) = T(x_i) + \mathrm{D}T(x_i)(x - x_i) + \sigma(x) = T(x_i) + \mathrm{D}T(x_i)\left(x + \mathrm{D}T(x_i)^{-1}\sigma(x) - x_i\right)$$

$$T(Q_i) \subset T(x_i) + \mathrm{D}T(x_i)\left(Q_i + B_{\delta\epsilon}(0) - x_i\right) \subset T(x_i) + \mathrm{D}T(x_i)\left(\tilde{Q}_i - x_i\right),$$

wobei $\tilde{Q}_i$ der Würfel mit dem selben Zentrum wie $Q_i$ und mit Kantenlänge $\frac{d}{m} + 2\delta\epsilon$ ist. Aus der Translationsinvarianz des Lebesgue-Maßes sowie aus Satz 5.35 folgt dann

$$\begin{aligned}\lambda^n(T(Q_i)) &\leq |\det(\mathrm{D}T(x_i))|\, \lambda^n(\tilde{Q}_i) = |\det(\mathrm{D}T(x_i))| \left(\frac{d}{m} + 2\delta\epsilon\right)^n \\ &= |\det(\mathrm{D}T(x_i))| \left(\frac{d}{m}\right)^n \left(1 + \frac{2\delta\epsilon m}{d}\right)^n = |\det(\mathrm{D}T(x_i))|\, \lambda^n(Q_i)(1 + 2\sqrt{n}\epsilon)^n.\end{aligned}$$

Insgesamt folgt dann aus der Subadditivität des Maßes

$$\begin{aligned}\lambda^n(T(Q)) &\leq \sum_{i=1}^k \lambda^n(T(Q_i)) \leq (1 + 2\sqrt{n}\epsilon)^n \sum_{i=1}^k |\det \mathrm{D}T(x_i)|\, \lambda^n(Q_i) \\ &\leq (1 + 2\sqrt{n}\epsilon)^n \sum_{i=1}^k \int |\det(\mathrm{D}T(x))|\, \mathbb{1}_{Q_i} \mathrm{d}\lambda^n(x) \\ &= (1 + 2\sqrt{n}\epsilon)^n \int |\det(\mathrm{D}T(x))|\, \mathbb{1}_Q \mathrm{d}\lambda^n(x).\end{aligned}$$

Dabei wurde für die letzte Umformung benutzt, dass die Menge auf der sich irgendwelche der $Q_i$ überlappen eine Nullmenge ist. Für $\epsilon \to 0$ folgt nun die Behauptung. □

**Lemma 5.39.** *Sei $A \subset U$ messbar, dann gilt $\lambda^n(T(A)) \leq \int_A |\det \mathrm{D}T(x)|\, \mathrm{d}\lambda^n(x)$.*

*Beweisskizze.* Man überzeugt sich leicht, dass die Abbildungen $\mu(A) := \lambda^n(T(A))$ und $\nu(A) := \int_A |\det \mathrm{D}T(x)|\, \mathrm{d}\lambda^n(x)$ Maße sind. Man zeigt dann, dass das Mengensystem

$$\mathcal{A} = \{A \subset \mathcal{A} |\ \nu(A) \leq \mu(A)\}$$

eine $\sigma$-Algebra ist. Nach dem vorhergehenden Lemma enthält es auch alle abgeschlossenen Quader und da diese die Borelsche $\sigma$-Algebra erzeugen gilt $\mathcal{B} \subset \mathcal{A}$, also die gewünschte Beziehung. □

**Lemma 5.40.** *Für $f : V \to \overline{\mathbb{R}}^+$ gilt*

$$\int_V f(y) \mathrm{d}\lambda^n(y) = \int_U f \circ T(x) \left|\det \mathrm{D}T(x)\right| \mathrm{d}\lambda^n(x).$$

*Beweis.* Wir zeigen zunächst

$$\int_U f(y) \mathrm{d}\lambda^n(y) \leq \int f \circ T(x) \left|\det \mathrm{D}T(x)\right| \mathrm{d}\lambda^n(x). \tag{5.3}$$

Sei $B \subset V$ messbar und $A = T^{-1}(B)$. Dann gilt $\mathbb{1}_B \circ T = \mathbb{1}_{T^{-1}(B)}$. Nach dem vorhergehenden Lemma folgt also

$$\begin{aligned}\int_V \mathbb{1}_B \mathrm{d}\lambda^n = \lambda^n B = \lambda^n T(T^{-1}(B)) &= \int_{T^{-1}(B)} \left|\det \mathrm{D}T(x)\right| \mathrm{d}\lambda^n(x) \\ &= \int_U \mathbb{1}_B \circ T(x) \left|\det \mathrm{D}T(x)\right| \mathrm{d}\lambda^n(x)\end{aligned}$$

also (5.3) für characteristische Funktionen.

Da beide Seiten der Ungleichung (5.3) linear in $f$ sind, gilt sie also auch für einfache Funktionen $f \in \mathbf{X}^+(V)$.

Für beliebiges $f$ sei $(f_n) \subset \mathbf{X}^+(V)$ eine monotone Folge mit $f_n \to f$ punktweise. Dann ist auch die Folge $(f_n \circ T \left|\det \mathrm{D}T\right|)$ monoton und konvergiert punktweise gegen $f \circ T \left|\det \mathrm{D}T\right|$. Dann gilt

$$\begin{aligned}\int_V f(y) \mathrm{d}\lambda^n(y) = \lim_{n\to\infty} \int_V f_n(y) \mathrm{d}\lambda^n(y) &\leq \int_U f_n \circ T(x) \left|\det \mathrm{D}T(x)\right| \mathrm{d}\lambda^n(x) \\ &= \lim_{n\to\infty} \int_U f \circ T(x) \left|\det \mathrm{D}T(x)\right| \mathrm{d}\lambda^n(x).\end{aligned}$$

Da auch $T^{-1} : V \to U$ ein Diffeomorphismus ist gilt für beliebiges $g : U \to \overline{\mathbb{R}}^+$

$$\int_U g(x) \mathrm{d}\lambda^n(x) \leq \int_V g \circ T^{-1}(y) \left|\det \mathrm{D}T^{-1}(y)\right| \mathrm{d}\lambda^n(y).$$

Wir wenden diese Beziehung auf $g = f \circ T \left|\det \mathrm{D}T\right|$ an und erhalten

$$\int_U f \circ T(x) \left|\det \mathrm{D}T(x)\right| \mathrm{d}\lambda^n(x) \leq \int_V f(y) \left|\det \mathrm{D}T^{-1}(T^{-1}(y)\right| \left|\det \mathrm{D}T^{-1}(y)\right| \mathrm{d}\lambda^n(y).$$

Nach der Kettenregel gilt

$$I = \mathrm{D}(T \circ T^{-1})(y) = \mathrm{D}T(T^{-1}(y))\mathrm{D}T^{-1}(y)$$

und damit

$$1 = \left|\det \mathrm{D}T^{-1}(T^{-1}(y)\right| \left|\det \mathrm{D}T^{-1}(y)\right|. \qquad \square$$

*Beweis von Theorem 5.36.* Aus dem vorhergehenden Lemma folgt sofort, dass $f$ über $V$ integrierbar ist genau dann, wenn $f \circ T\,|\det \mathrm{D}T|$ über $U$ integrierbar ist. Zerlege $f$ in Positiv- und Negativteil und wende auf beide das vorhergehende Lemma an. $\square$

*Beispiel 5.41 (Integrierbarkeit von radialsymmetrischen Funktionen).* Sei $f : [0,\infty] \to \overline{\mathbb{R}}$ messbar und $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ die Funktion $F(x) = f(\|x\|)$ wobei $\|\cdot\|$ die euklidische Norm bezeichnet. Unter welchen Bedingungen an $f$ ist $F$ dann $\lambda^n$ integrierbar? Dazu können wir Kugelkoordinaten in $\mathbb{R}^n$ folgendermaßen definieren:

$$\begin{aligned}
D &:= (0,\infty) \times D_\varphi := (0,\infty) \times (0,\pi)^{n-2} \times (0,2\pi) \\
T &: D \to \mathbb{R}^n \\
(r,\varphi) &\mapsto \begin{array}{l}
(r\cos(\varphi_1), \\
\ r\sin(\varphi_1)\cos(\varphi_2), \\
\ r\sin(\varphi_1)\sin(\varphi_2)\cos(\varphi_3), \\
\ \ldots, \\
\ r\sin(\varphi_1)\cdots\sin(\varphi_{n-3})\cos(\varphi_{n-2}), \\
\ r\sin(\varphi_1)\cdots\sin(\varphi_{n-2})\cos(\varphi_{n-1}), \\
\ r\sin(\varphi_1)\cdots\sin(\varphi_{n-2})\sin(\varphi_{n-1})).
\end{array}
\end{aligned}$$

Dabei ist $\varphi = (\varphi_1, \ldots, \varphi_{n-1})$. Man kann zeigen (Übung) dass $T$ injektiv ist und

$$|\det \mathrm{D}T(r,\varphi)| = r^{n-1}\sin(\varphi_1)^{n-2}\sin(\varphi_2)^{n-3}\cdots\sin(\varphi_{n-2}) =: r^{n-1}A(\varphi)$$

gilt. Damit ist insbesondere $\mathrm{D}T$ auf $D$ invertierbar und $T : D \to T(D)$ ein Diffeomorphismus (Bemerkung 5.33). Außerdem überzeugt man sich, dass $\|T(r,\varphi)\| = r$ gilt.

Dann gilt mit dem Transformationssatz und dem Satz von Tonelli

$$\begin{aligned}
\int |F(x)|\,\mathrm{d}\lambda^n(x) &= \int |F \circ T(r,\varphi)|\,|\det \mathrm{D}T(r,\varphi)|\,\mathrm{d}\lambda^n(r,\varphi) \\
&= \int_{D_\varphi}\int_0^\infty |f(r)|\,r^{n-1}A(\varphi)\mathrm{d}r\mathrm{d}\lambda^{n-1}(\varphi) \\
&= \int_0^\infty r^{n-1}\,|f(r)|\,\mathrm{d}r \int_{D_\varphi} A(\varphi)\mathrm{d}\lambda^{n-1}(\varphi).
\end{aligned}$$

Das Integral über $\varphi$ kann man exakt berechnen, hier ist jedoch nur entscheidend, dass $A$ beschränkt ist und $D_\varphi$ endliches $\lambda^{n-1}$ Maß hat. Man erhält also, dass $F$ integrierbar ist, genau dann wenn $r \mapsto r^{n-1}\,|f(r)|$ integrierbar über $(0,\infty)$ ist.

Insbesondere ist $e^{-\|x\|}$ integrierbar für beliebigen Dimensionen $n$. Die Funktion $\frac{1}{\|x\|^\kappa}$ ist über $\mathbb{R}^n \backslash B_\epsilon(0)$ (für beliebiges $\epsilon > 0$) integrierbar genau dann wenn $\kappa > n$.

# 6 Mannigfaltigkeiten und Integralsätze

## 6.1 Kurvenintegrale

**Definition 6.1.** Seien $I$ ein Intervall mit $\mathring{I} \neq \emptyset$. Eine stetige Abbildung $\gamma : I \to \mathbb{R}^n$ heißt Kurve. Sie heißt regulär, falls $\gamma$ auf $\mathring{I}$ stetig differenzierbar ist und dort $\gamma' \neq 0$ gilt. Sie heißt stückweise regulär, falls es eine disjunkte Zerlegung

$$I = I_1 \cup \cdots \cup I_k$$

in Teilintervalle $I_1, \ldots, I_k$ gibt, so dass $\gamma$ auf $I_k$ glatt ist für jedes $i \in \{1, \ldots, k-1\}$.

Im Folgenden wird jede Kurve als stückweise regulär vorausgesetzt.

*Bemerkung 6.2.* Häufig ist das eigentliche Objekt von Interesse das Bild $\gamma(I)$, welches wir auch als Kurve bezeichnen. Für $A \subset \mathbb{R}^n$ heißt die Aussage „$A$ ist eine Kurve“ also, dass es $\gamma : I \to \mathbb{R}^n$ stückweise regulär gibt, so dass $\gamma(I) = A$. In diesem Fall heißt $\gamma$ eine Parametrisierung von $A$.

**Definition 6.3.** Seien $I, J \subset \mathbb{R}$ Intervalle mit $\mathring{I} \neq \emptyset$ und $\mathring{J} \neq \emptyset$. Eine reguläre Kurve $\theta : J \to I$ heißt Reparametrisierung. Sie heißt orientierungserhaltend, falls $\theta' > 0$ und orientierungsumkehrend falls $\theta' < 0$ ist.

Falls $\gamma : I \to \mathbb{R}^n$ eine Kurve ist, so ist auch $\gamma \circ \theta$ eine Kurve mit dem selben Bild: $\gamma(I) = \gamma \circ \theta(J)$.

*Beispiel 6.4.* (i) Seien $\alpha, \kappa > 0$. Dann ist die Schraubenlinie

$$\begin{aligned} \gamma : \mathbb{R} &\to \mathbb{R}^3 \\ t &\mapsto (\cos(\alpha t), \sin(\alpha t), \kappa t) \end{aligned}$$

eine reguläre Kurve.

(ii) Das Rechteck $\left\{(x, y) \in \mathbb{R}^2 \middle| \; |x| = 1 \text{ oder } |y| = 1\right\}$ ist eine (stückweise reguläre) Kurve.

**Definition 6.5 (Differenzialform 1. Grades).** Sei $D \subset \mathbb{R}^n$. Eine Abbildung

$$\omega : D \to \mathcal{L}(\mathbb{R}^n | \mathbb{R})$$

heißt Differenzialform 1. Grades oder 1-Form (Pfaffsche Form, kovariantes Vektorfeld (Kovektor)).

*Beispiel 6.6.* (i) Sei $E : \mathbb{R}^n \to \mathbb{R}^n$ ein (kontravariantes) Vektorfeld, dann ist $\omega_E : \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n | \mathbb{R})$ definiert durch $\omega_E(x)v = \langle E(x) | v \rangle$ eine 1-Form.

(ii) Sei $I$ ein Intervall und $\gamma : I \to \mathbb{R}^n$ eine injektive, reguläre Kurve. Dann ist $\omega_\gamma(\gamma(t)) := \frac{1}{\|\gamma'(t)\|} \langle \gamma'(t) | \cdot \rangle$ eine 1-Form auf dem Bild $\gamma(I)$.

Sei $\theta : J \to I$ eine orientierungserhaltende Reparametrisierung, $s \in J$ und $t = \theta(s)$. Dann gilt

$$(\gamma \circ \theta)'(s) = \gamma'(\theta(s))\theta'(s)$$

und

$$\begin{aligned}\omega_{\gamma\circ\theta}(\gamma(t)) = \omega_{\gamma\circ\theta}(\gamma\circ\theta(s)) &= \frac{1}{|\theta'(s)|\,\|\gamma'(\theta(s))\|}\big\langle \theta'(s)\gamma'(\theta(s))\big|\,\cdot\big\rangle \\ &= \frac{1}{\|\gamma'(t)\|}\big\langle\gamma'(t)\big|\,\cdot\big\rangle = \omega_\gamma(\gamma(t)).\end{aligned}$$

Das gilt für jedes $s \in J$, also jedes $t \in I$. Die 1-Form $\omega_\gamma$ ist also von der Parametrisierung unabhängig und hängt nur von dem geometrischen Objekt $\gamma(I) = \gamma \circ \theta(J)$ ab. Man sagt, $\omega_\gamma$ ist eine geometrische Invariante.

(iii) Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine differenzierbare Funktion, dann ist $\mathrm{d}f = \mathrm{D}f$ eine 1-Form, die das Differenzial von $f$ genannt wird.

(iv) Einen wichtigen Spezialfall des vorhergehenden Punktes erhält man für $\pi_i : \mathbb{R}^n \to \mathbb{R}$, $\pi_i(x_1, \ldots, x_n) = x_i$. Für das Differenzial dieser Funktion schreibt man $\mathrm{d}x_i$ (anstatt $\mathrm{d}\pi_i$). Für $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ gilt

$$\mathrm{d}x_i(x)v = \mathrm{D}\pi_i(x)v = v_i.$$

(v) Seien $\omega_1, \omega_2$ 1-Formen und $f : \mathbb{R}^n \to \mathbb{R}$. Dann sind auch

$$\omega_1 + \omega_2 \text{ sowie } f\omega_1$$

1-Formen.

*Bemerkung 6.7.* Sei $f : \mathbb{R}^n \to \mathbb{R}$. Es gilt

$$\mathrm{d}f(x)v = \mathrm{D}f(x)v = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)v_i = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)\mathrm{d}x_i(x)v$$

für alle $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$, also

$$\mathrm{d}f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\mathrm{d}x_i.$$

Allgemeiner gibt es für jede 1-Form $\omega$ auf $D$ Funktionen $\omega_1, \ldots, \omega_n : D \to \mathbb{R}$ so, dass

$$\omega = \sum_{i=1}^{n} \omega_i \mathrm{d}x_i.$$

*Beispiel 6.8.* Die 1-Form

$$\omega(x,y) = \frac{1}{2\pi}\left(-\frac{y}{x^2+y^2}\mathrm{d}x + \frac{x}{x^2+y^2}\mathrm{d}y\right)$$

auf $\mathbb{R}^2 \setminus \{0\}$ heißt Windungsform.

Die zum Gravitationsfeld gehörige 1-Form

$$\frac{1}{(x^2+y^2+z^2)^{\frac{3}{2}}}(x\mathrm{d}x + y\mathrm{d}y + z\mathrm{d}z)$$

heißt Gravitationsform.

**Satz 6.9.** *Sei $\gamma : I \to \mathbb{R}^n$ eine Kurve, $\omega$ eine Differenzialform auf $\gamma(I)$ und $\theta : J \to I$ eine orientierungserhaltende Reparametrisierung. Setze $\rho := \gamma \circ \theta$. Dann gilt, dass $t \mapsto \omega(\gamma(t))\gamma'(t)$ über $I$ integrierbar ist genau dann wenn $s \mapsto \omega(\rho(s))\rho'(s)$ über $J$ integrierbar ist. In diesem Fall gilt*

$$\int_I \omega(\gamma(t))\gamma'(t)\mathrm{d}t = \int_J \omega(\rho(s))\rho'(s)\mathrm{d}s.$$

*Beweis.* Es ist ausreichend, den Fall für $I, J$ offen zu betrachten, da die Endpunkte keine Relevanz für die Integrale haben. In diesem Fall ist $\theta$ ein Diffeomorphismus zwischen $J$ und $I$ mit $|\det \mathrm{D}\theta| = \theta'$ und wir erhalten mit der Transformationsformel, dass $\omega(\gamma(t))\gamma'(t)$ integrierbar über $I$ ist, genau dann wenn $\omega(\gamma \circ \theta(s))\gamma' \circ \theta(s)\theta'(s)$ über $J$ integrierbar ist und falls eine der Bedingungen erfüllt ist gilt

$$\int_I \omega(\gamma(t))\gamma'(t)\mathrm{d}t = \int_J \omega(\gamma \circ \theta(s))\gamma' \circ \theta(s)\theta'(s)\mathrm{d}s = \int_J \omega(\rho(s))\rho'(s)\mathrm{d}s. \qquad \square$$

**Definition 6.10.** Sei $\gamma : I \to \mathbb{R}^n$ eine Kurve und $\omega$ eine 1-Form auf $\gamma(I)$. Dann heißt $\omega$ integrierbar über $\gamma(I)$ wenn $t \mapsto \omega(\gamma(t))\gamma'(t)$ integrierbar über $I$ ist. In diesem Fall definieren wir das Integral von $\omega$ über $\gamma(I)$ mittels

$$\int_{\gamma(I)} \omega := \int_I \omega(\gamma(t))\gamma'(t)\mathrm{d}t.$$

*Beispiel 6.11.* Wir wollen das Integral der Windungsform

$$\omega(x,y) = \frac{1}{2\pi}\left(-\frac{y}{x^2+y^2}\mathrm{d}x + \frac{x}{x^2+y^2}\mathrm{d}y\right)$$

über den Einheitskreis berechnen, den wir wie folgt parametrisieren können

$$\begin{aligned}\gamma : [0,1] &\to \mathbb{R}^2\\ t &\mapsto (\cos(2\pi t), \sin(2\pi t)).\end{aligned}$$

Die Ableitung der Kurve ist $\gamma'(t) = 2\pi(-\sin(2\pi t), \cos(2\pi t))$. Dann gilt

$$\begin{aligned}\int_\gamma \omega &= \frac{1}{2\pi}\int_0^1 \big(-\sin(2\pi t)\mathrm{d}x(\gamma'(t)) + \cos(2\pi t)\mathrm{d}y(\gamma'(t))\big)\,\mathrm{d}t\\ &= \int_0^1 (\sin^2(2\pi t) + \cos^2(2\pi t))\mathrm{d}t = 1.\end{aligned}$$

*Beispiel 6.12.* (i) Für ein Vektorfeld $E : \mathbb{R}^n \to \mathbb{R}^n$ gilt

$$\int_\gamma \omega_E = \int_I \big\langle E(\gamma(t)) \big| \gamma'(t) \big\rangle \mathrm{d}t.$$

Wirkt auf eine Punktmasse mit Masse $m$ die sich entlang $\gamma$ bewegt das (zeitunabhängige) Kraftfeld $F$, dann berechnet sich die verrichtete Arbeit als

$$m\int_\gamma \omega_F.$$

Die Unabhängigkeit unter Reparametrisierung bedeutet dabei, dass die verrichtete Arbeit unabhängig von der Geschwindigkeit der Punktmasse ist.

(ii) Für eine reguläre Kurve $\gamma : I \to \mathbb{R}^n$ ist

$$\int_I \omega_\gamma = \int_I \frac{1}{\|\gamma'(t)\|}\big\langle \gamma'(t) \big| \gamma'(t) \big\rangle \mathrm{d}t = \int_I \big\|\gamma'(t)\big\|\,\mathrm{d}t$$

die Bogenlänge der Kurve $\gamma$. Da sowohl $\omega$ als auch das Integral nicht von der Parametrisierung abhängen, ist die Bogenlänge ebenfalls davon unabhängig.

(iii) Allgemeiner ist für eine Funktion $f : \gamma(I) \to \mathbb{R}$ das Integral von $f$ über $\gamma$ gegeben durch

$$\int_\gamma f\omega_\gamma = \int_I f(\gamma(t))\,\big\|\gamma'(t)\big\|\,\mathrm{d}t$$

Stellt beispielsweise $\rho : \gamma(I) \to \mathbb{R}$ die lineare Dichte (Masse/Länge) eines Drahtes mit vernachlässigbarer Dicke dar, dann ist

$$\int_\gamma \rho\omega_\gamma$$

die Gesamtmasse des Drahtes.

(iv) Für differenzierbare $f : \mathbb{R}^n \to \mathbb{R}$ und $\gamma : [a, b] \to \mathbb{R}^n$ regulär gilt

$$\int_\gamma \mathrm{d}f = \int_I \mathrm{D}f(\gamma(t))\gamma'(t)\mathrm{d}t = \int_a^b (f\circ\gamma)'(t)\mathrm{d}t = f(\gamma(b)) - f(\gamma(a)).$$

In diesem fall läßt sich das Integral also einfach bestimmen und ist darüber hinaus vom Weg unabhängig.

*Bemerkung 6.13.* (i) Für stückweise reguläre Kurven kann man das Integral als Summe der Integrale der regulären Teilstücke definieren. Alle obigen Formeln behalten ihre Gültigkeit.

(ii) Es gilt $\int_\gamma(\omega_1 + \lambda\omega_2) = \int_\gamma \omega_1 + \lambda \int_\gamma \omega_2$, das Kurvenintegral ist also linear in $\omega$.

(iii) Die Abbildung

$$\begin{aligned} \theta : [0,1] &\to [a,b] \\ s &\mapsto a + s(b-a) \end{aligned}$$

ist eine orientierungserhaltende Reparametrisierung. Wir können statt einer Kurve $\gamma : [a,b] \to \mathbb{R}^n$ also stets die reparametrisierte Kurve $\gamma \circ \theta : [0,1] \to \mathbb{R}^n$ betrachten.

(iv) Für zwei Kurven $\gamma, \rho : [0,1] \to \mathbb{R}^n$ mit $\gamma(1) = \rho(0)$ definieren wir die Kurven

$$\begin{aligned} \gamma^{-1} : [0,1] &\to \mathbb{R}^n \\ t &\mapsto \gamma(1-t) \\ \gamma\rho : [0,1] &\to \mathbb{R}^n \\ t &\mapsto \begin{cases} \gamma(2t) & 0 \le t < \frac{1}{2} \\ \rho(2t-1) & \frac{1}{2} \le t \le 1 \end{cases}. \end{aligned}$$

Man überzeugt sich leicht, dass dann

$$\int_{\gamma^{-1}} \omega = -\int_\gamma \omega \text{ und } \int_{\gamma\rho} \omega = \int_\gamma \omega + \int_\rho \omega$$

gilt.

*Bemerkung 6.14.* Sei $\gamma : I \to \mathbb{R}^n$ regulär und $f : \gamma(I) \to \mathbb{R}$ stetig. Seien $a, b \in \mathring{I}$ und $a < b$. Eine Zerlegung $Z$ ist gegeben durch Punkte $a = t_0 < t_1 < \cdots < t_n = b$. Wir nennen die $t_i$ Stützstellen. Die Feinheit einer Zerlegung sei

$$m(Z) = \max_{0 \le i < n} (t_{i+1} - t_i). \tag{6.1}$$

Setze

$$I(Z) := \sum_{i=0}^{n-1} f(\gamma t_i) \left\|\gamma(t_{i+1}) - \gamma(t_i)\right\| = \int_{[a,b]} \sum_{i=0}^{n-1} f(\gamma t_i) \left\| \frac{\gamma(t_{i+1}) - \gamma(t_i))}{t_{i+1} - t_i} \right\| \mathbb{1}_{[t_i, t_{i+1})}(t) \mathrm{d}t.$$

Bezeichne den Integranden im letzen Ausdruck mit $g_Z$. Mit dem verallgemeinerten Mittelwertsatz kann man abschätzen

$$|g_Z(t)| \le \max_{t \in [a,b]} |f(t)| \max_{t \in [a,b]} \left\|\gamma'\right\| =: K < \infty.$$

Sei nun $(Z_j)$ eine Folge von Zerlegungen mit $m(Z_j) \to 0$ und $t$ sei nicht Stützstelle einer Zerlegung $Z_j$. Dann gibt es eine eindeutig bestimmte Folge von Stützstellen $t_{j,i_j}$, so dass $t \in \left[t_{j,i_j}, t_{j,i_j+1}\right)$. Es gilt $\lim_{j\to\infty} t_{j,i_j} = t = \lim_{j\to\infty} t_{j,i_j+1}$ und damit wegen der Stetigkeit von $f$ und der Differenzierbarkeit von $\gamma$

$$\lim_{j\to\infty} g_{Z_j}(t) = f(\gamma t) \left\|\gamma'(t)\right\|.$$

Da die Menge aller Stützstellen abzählbar, also eine Nullmenge ist, gilt $\lim_{j\to\infty} g_{Z_j} = f\left\|\gamma'\right\|$ fast überall und wegen (6.14) sind die Funktionen $g_{Z_j}$ integrierbar beschränkt. Dann folgt aus dem Satz von Lebesgue

$$\lim_{j\to\infty} I(Z_j) = \int_a^b f(\gamma t) \left\|\gamma'(t)\right\| \mathrm{d}t.$$

Insbesondere folgt für $f$ die Einsfunktion die Formel für die Bogenlänge.

**Definition 6.15.** Sei $U \subset \mathbb{R}^n$ offen. Eine 1-Form $\omega : U \to \mathcal{L}(\mathbb{R}^n|\mathbb{R})$ heißt exakt, falls es eine differenzierbare Funktion $f : U \to \mathcal{L}(\mathbb{R}^n|\mathbb{R})$ gibt mit $\omega = \mathrm{d}f$. Die Funktion $f$ heißt dann Stammfunktion oder Potential von $\omega$.

*Bemerkung 6.16.* (i) Für eine exakte 1-Form

$$\omega = \omega_1 \mathrm{d}x_1 + \cdots + \omega_n \mathrm{d}x_n$$

mit $\omega_1, \ldots, \omega_n$ stetig differenzierbar und Potential $f$ muss gelten

$$\omega = \mathrm{d}f = \frac{\partial f}{\partial x_1}\mathrm{d}x_1 + \cdots + \frac{\partial f}{\partial x_n}\mathrm{d}x_n.$$

also $\omega_i = \frac{\partial f}{\partial x_i}$. Da die Koeffizientenfunktionen $\omega_i$ stetig differenzierbar sind, folgt aus dem Satz von Schwarz, dass

$$\frac{\partial \omega_i}{\partial x_j} = \frac{\partial \omega_j}{\partial x_i}$$

für alle $i, j \in \{1, \ldots, n\}$ gilt. Diese Bedingung ist also notwendig für die Exaktheit einer Form.

(ii) Die Differenzialgleichung $p(x, y) + q(x, y)y'$ ist exakt, genau dann wenn die 1-Form $p(x, y)\mathrm{d}x + q(x, y)\mathrm{d}y$ exakt ist, da $f$ ein Potential für die Form ist, genau dann wenn es ein Potential für die Differenzialgleichung ist.

(iii) Kennen wir ein Potential für eine exakte 1-Form, dann können wir Kurvenintegrale über $\omega$ nach Beispiel 6.12 (iv) leicht ausrechnen.

**Satz 6.17.** *Eine stetige* 1*-Form* $\omega : U \to \mathcal{L}(\mathbb{R}^n|\mathbb{R})$ *ist genau dann exakt, wenn für alle geschlossenen Kurven* $\gamma : [0, 1] \to U$ *(das heißt* $\gamma(0) = \gamma(1)$*) gilt*

$$\int_\gamma \omega = 0.$$

*Beweis.* Sei $\omega$ exakt und $f$ ein Potential, dann gilt nach Beispiel 6.12 (iv) für geschlossene Kurven $\gamma : [0,1] \to U$

$$\int_\gamma \omega = \int_\gamma \mathrm{d}f = f(\gamma(1)) - f(\gamma(0)) = 0.$$

Es gelte nun $\int_\gamma \omega = 0$ für beliebige geschlossene $\gamma$. Wir nehmen (ohne Einschränkung der Allgemeinheit) an, dass $U$ wegzusammenhängend ist, das heißt, dass es für $x, y \in U$ eine Kurve $\gamma : [0,1] \to U$ gibt so, dass $\gamma(0) = x$ und $\gamma(1) = y$. Wähle eine festen Punkt $x_0 \in U$. Sei $x \in U$ und $\gamma, \rho : [0,1] \to U$ Kurven mit $\gamma(0) = \rho(0) = x_0$ und $\gamma(1) = \rho(1) = x$. Dann ist $\gamma\rho^{-1}$ eine geschlossene Kurve und es gilt

$$0 = \int_{\gamma\rho^{-1}} \omega = \int_\gamma \omega - \int_\rho \omega.$$

Das heißt, das Kurvenintegral von $x_0$ nach $x$ hängt nicht vom Integrationsweg ab. Wir definieren jetzt

$$f(x) = \int_{x_0}^x \omega$$

wobei das Integral rechts entlang irgendeines Integrationsweges von $x_0$ nach $x$ ausgeführt wird.

Wir zeigen jetzt die Differenzierbarkeit von $f$. Zunächst berechnen wir die Richtungsableitung von $f$ im Punkt $x$ in Richtung $h$. Wähle $r$ so, dass $B_r(x) \subset U$. Für alle $s \in \mathbb{R}$ mit $|s| < \frac{r}{\|h\|}$ gilt

$$f(x+sh) - f(x) = \int_{x_0}^{x+sh} \omega - \int_{x_0}^{x} \omega = \int_{x}^{x+sh} \omega = \int_0^s \omega(x+th)h\mathrm{d}h.$$

Dabei haben wir einen beliebigen Integrationsweg von $x_0$ bis $x$ und die Kurve $t \mapsto x+th$ von $x$ bis $x+sh$ gewählt. Nach dem Mittelwertsatz der Integralrechnung existiert $\xi_s \in (0,s)$ so, dass

$$\int_0^s \omega(x+th)h\mathrm{d}h = s\omega(x+\xi_s h)h$$

gilt und es folgt für die Richtungsableitung

$$\lim_{s\to 0} \frac{f(x+sh) - f(x)}{s} = \lim_{s\to 0} \omega(x+\xi_s h)h = \omega(x)h,$$

wobei wir die Stetigkeit von $\omega$ benutzt haben.

Insbesondere folgt (für $h = e_i$) die Stetigkeit der partiellen Ableitungen von $f$ und damit aus Satz 2.33 die stetige Differenzierbarkeit von $f$. Wir erhalten dann auch

$$\mathrm{d}f(x)h = \mathrm{D}f(x)h = \omega(x)h$$

und da das für beliebige $h \in \mathbb{R}^n$ und beliebiges $x \in U$ gilt, ist $\mathrm{d}f = \omega$. □

*Bemerkung 6.18.* Die Voraussetzung des Satzes ist schwierig nachzuprüfen. Er liefert aber ein Kriterium wann Formen (und Differenzialgleichungen) nicht exakt sind und ein Verfahren um Potentiale zu bestimmen.

## 6.2 Eingebettete Mannigfaltigkeiten

Anschaulich sind Untermannigfaltigkeiten (mehr oder weniger) glatte Teilmengen des $\mathbb{R}^m$. Etwas präziser ist eine Untermannigfaltigkeit eine Teilmenge von $\mathbb{R}^m$, die lokal wie ein Untervektorraum aussieht.

**Definition 6.19 (Mannigfaltigkeit).** Eine Teilmenge $M \subset \mathbb{R}^m$ heißt Untermannigfaltigkeit (eingebettete Mannigfaltigkeit) der Dimension $n \leq m$, falls für jedes $x \in M$ eine offene Umgebung $U \subset \mathbb{R}^m$, eine offene Teilmenge $V \subset \mathbb{R}^m = \mathbb{R}^n \times \mathbb{R}^{m-n}$ und ein Diffeomorphismus $\Phi : U \to V$ existiert so, dass $\Phi(M \cap U) \subset \mathbb{R}^n \times \{0\}$. Die Zahl $n$ heißt Dimension der Mannigfaltigkeit.

Bezeichne $\pi : \mathbb{R}^m \to \mathbb{R}^n$ die kanonische Projektion, dann heißt die Abbildung $\varphi := \pi \circ \Phi : M \cap U \to \mathbb{R}^n$ (die zu $\Phi$ gehörige) Karte von $M \cap U$.

*Beispiel 6.20.* (i) Die Menge $\mathbb{R}^n \times \{0\} \subset \mathbb{R}^m = \mathbb{R}^n \times \mathbb{R}^{m-n}$ ist mit $\Phi = \mathrm{id}$ eine $n$-dimensionale Untermannigfaltigkeit.

(ii) Sei $U \subset \mathbb{R}^n$ offen und $\tilde{U} = U \times \mathbb{R}^{m-n}$. Dann ist $U$ eine $n$-Dimensionale Untermannnigfaltigkeit mit $\Phi : \tilde{U} \to \tilde{U}$ ebenfalls die Identität.

(iii) Sei $M = \mathbb{S}^1 := \mathbb{T}^1 := \left\{(x, y) \in \mathbb{R}^2 \middle| \, x^2 + y^2 = 1\right\}$ der Einheitskreis. Setze $U = \left\{(x, y) \in \mathbb{R}^2 \middle| \, \frac{1}{2} < x^2 + y^2 < 2\right\}$ und

$$\begin{aligned} U_1 &= \{(x, y) \in U | \, x > 0\} \quad & U_2 &= \{(x, y) \in U | \, x < 0\} \\ U_3 &= \{(x, y) \in U | \, y > 0\} \quad & U_4 &= \{(x, y) \in U | \, y < 0\} \,. \end{aligned}$$

Die obigen Mengen sind offen und es gilt $M \subset U_1 \cup U_2 \cup U_3 \cup U_4$. Dann sind die Abbildungen $\Phi_i : U_i \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times \left(\frac{1}{2}, 2\right)$

$$\begin{aligned} \Phi_{1,2}(x, y) &= \left(\arctan\left(\frac{y}{x}\right), \sqrt{x^2 + y^2} - 1\right) \\ \Phi_{3,4}(x, y) &= \left(\arctan\left(\frac{x}{y}\right), \sqrt{x^2 + y^2} - 1\right) \end{aligned}$$

Diffeomorphismen mit $\Phi_i(M \cap U_i) = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times \{0\}$.

(iv) Die abgeschlossene Einheitskugel $K_1(0)$ oder die Menge $(\mathbb{R} \times \{0\}) \cup (\{0\} \times \mathbb{R})$ sind keine Mannigfaltigkeiten.

*Bemerkung 6.21.* (i) Häufig spielt der umgebende $\mathbb{R}^m$ keine Rolle. Wir werden daher Mannigfaltigkeit statt Untermannigfaltigkeit sagen. Es ist möglich, die Theorie der Mannigfaltigkeiten abstrakt, ohne umgebenden Raum $\mathbb{R}^m$, zu formulieren.

(ii) Ein Karte $\varphi : M \cap U \to \mathbb{R}^n$ bildet $M \cap U$ bijektiv auf $\varphi(M \cap U) = V \cap (\mathbb{R}^n \times \{0\})$ ab und $\varphi^{-1}$ ist differenzierbar. Das folgt sofort aus den entsprechenden Eigenschaften der Abbildung $\Phi$.

Für $U_1, U_2$ offene Teilmengen von $\mathbb{R}^m$ und Diffeomorphismen $\Phi_i : U_i \to V_i$. Setze $U = U_1 \cap U_2$, dann sind $\Phi_1(U)$ und $\Phi_2(U)$ offen (Warum?). Außerdem ist $\Phi_1 \circ$

$\Phi_2^{-1} : \Phi_2(U) \to \Phi_1(U)$ ein Diffeomorphismus (Warum?). Für die zu $\Phi_1$ und $\Phi_2$ gehörigen Karten folgt daraus das $\varphi_1 \circ \varphi_2^{-1} : \varphi_2(U \cap M) \to \varphi_1(U \cap M)$ ebenfalls ein Diffeomorphismus ist.

Die Abbildung $\varphi_1 \circ \varphi_2^{-1}$ heißt Übergangsabbildung oder Kartenwechsel.

(iii) Der oben definierte Begriff beschreibt $\mathbf{C}^1$-Untermannigfaltigkeiten. Fordert man von den Abbildungen $\Phi$ und $\Phi^{-1}$ lediglich Stetigkeit, dann erhält man sogenannte topologische Untermannigfaltigkeiten ($\mathbf{C}^0$-Mannigfaltigkeiten).

Fordert man $r$-fache (oder beliebige) stetige Differenzierbarkeit dann erhält man $\mathbf{C}^r$-Mannigfaltigkeiten (oder glatte Mannigfaltigkeiten).

(iv) Da Verschiebungen Diffeomorphismen sind, können wir für $x \in M$, $\Phi$ (und damit $\varphi$) stets so wählen, dass $\Phi(x) = 0$ ($\varphi(x) = 0$).

(v) Falls $M$ und $N$ Mannigfaltigkeiten sind, dann ist $M \times N$ ebenfalls eine Mannigfaltigkeit.

Beispielsweise ist $\mathbb{T}^2 = \mathbb{T}^1 \times \mathbb{T}^1 = \left\{(x, y, z, w) \in \mathbb{R}^4 \middle|\, x^2 + y^2 = 1 \text{ und } z^2 + w^2 = 1\right\}$ eine Mannigfaltigkeit. Bereits hier scheitert die Anschauung am einbettenden Raum $\mathbb{R}^4$. Die Mannigfaltigkeit $\mathbb{T}^2$ läßt sich zwar auch in $\mathbb{R}^3$ einbetten (als Oberfläche eine Doughnuts), diese Einbettung verzerrt jedoch die Abstände.

Häufig reicht es aus, zu wissen dass eine Menge $M$ eine Mannigfaltigkeit ist ohne die Abbildungen $\Phi$ oder $\varphi$ konkret angeben zu müssen. Dann liefert der folgende Satz ein hilfreiches Kriterium.

**Satz 6.22 (Satz vom regulären Wert).** *Sei $U \subset \mathbb{R}^m$ offen und $f : U \to \mathbb{R}^k$ eine stetig differenzierbare Abbildung und $y \in \mathbb{R}^k$. Dann ist die Menge*

$$M = \{x \in \mathbb{R}^m |\, f(x) = y \text{ und } \mathrm{D}f(x) \text{ surjektiv}\}$$

*eine Untermannigfaltigkeit der Dimension $m - k$.*

*Ohne Beweis.*

*Beispiel 6.23.* (i) Betrachte

$$\begin{aligned} f : \mathbb{R}^3 &\to \mathbb{R} \\ (x, y, z) &\mapsto x^2 + y^2 + z^2. \end{aligned}$$

Es gilt $\mathrm{D}f(x, y, z)(v_1, v_2, v_3) = 2xv_1 + 2yv_2 + 2zv_3$. Die lineare Abbildung $\mathrm{D}f(x, y, z)$ ist also surjektiv, genau dann wenn $(x, y, z) \neq 0$ ist. Damit ist

$$\mathbb{S}^2 = \left\{(x, y, z) \in \mathbb{R}^3 \middle|\, x^2 + y^2 + z^2 = 1\right\} = f^{-1}(\{1\})$$

surjektiv und die Kugeloberfläche $\mathbb{S}^2$ damit eine 2-dimensionale Untermannigfaltigkeit.

Ein analoges Argument zeigt, dass die $n$-Sphäre

$$\mathbb{S}^n(r) = \left\{ x \in \mathbb{R}^{n+1} \middle| \; \|x\| = r \right\}$$

für beliebiges $r > 0$ eine $n$-dimensionale Untermannigfaltigkeit ist.

(ii) Für

$$\begin{aligned} f : \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto xy \end{aligned}$$

ist $f^{-1}(\{0\}) = (\mathbb{R} \times \{0\}) \cup (\{0\} \times \mathbb{R})$ keine Mannigfaltigkeit. Die Abbildung $\mathrm{D}f(x, y) = (y, x)$ ist surjektiv falls $(x, y) \neq 0$ und damit ist

$$(\mathbb{R} \times \{0\}) \cup (\{0\} \times \mathbb{R}) = f^{-1}(\{0\}) \cap \left\{ (x, y) \in \mathbb{R}^2 \middle| \, \mathrm{D}f(x, y) \text{ surjektiv} \right\}$$

eine 1-dimensionale Mannigfaltigkeit.

(iii) Mit dem obigen Kriterium kann man in vielen Fällen zeigen, dass der Konfigurationsraum eines mechanischen Systems mit holonomen Zwangsbedingungen eine differenzierbare Mannigfaltigkeit ist. Der Konfigurationsraum eines ebenen Doppelpendels mit Pendellängen $l, r > 0$ läßt sich schreiben als

$$M := \left\{ (x_1, x_2, y_1, y_2) \in \mathbb{R}^4 \middle| \, x_1^2 + x_2^2 = l^2, (x_1 - y_1)^2 + (x_2 - y_2)^2 = r^2 \right\}.$$

Mit der Definition

$$\begin{aligned} f : \mathbb{R}^4 &\to \mathbb{R}^2 \\ (x_1, x_2, y_1, y_2) &\mapsto (x_1^2 + x_2^2, (x_1 - y_1)^2 + (x_2 - y_2)^2) \end{aligned}$$

gilt $M = f^{-1}(\{l^2, r^2\})$. Wann hat

$$\begin{aligned} &\mathrm{D}f(x_1, x_2, y_1, x_2)(v_1, v_2, w_1, w_2) \\ &= \begin{pmatrix} 2x_1 v_1 + 2x_2 v_2 \\ 2(x_1 - y_1)v_1 + 2(x_2 - y_2)v_2 - 2(x_1 - y_1)w_1 - 2(x_2 - y_2)w_2 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \end{aligned}$$

für beliebiges $a, b \in \mathbb{R}$ eine Lösung? Das ist der Fall wenn $(x_1, x_2) \neq 0$ und $(x_1 - y_1, x_2 - y_2) \neq 0$, was auf $M$ erfüllt ist. Damit ist $M$ nach dem Satz eine 2-dimensionale Untermannigfaltigkeit.

Die Wahl einer Karte $\varphi$ entspricht in diesem Fall der Wahl von (lokalen) generalisierte Koordinaten $(q_1, q_2)$.

**Satz 6.24.** *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine stetig differenzierbare Funktion. Dann ist der Graph von $f$, also die Menge*

$$G(f) = \left\{ (x, f(x)) \in \mathbb{R}^{n+1} \middle| \, x \in U \right\}$$

*eine Mannigfaltigkeit.*

*Beweis.* Setze $\tilde{U} = U \times \mathbb{R} \subset \mathbb{R}^{n+1}$ und

$$\begin{aligned} \Phi : \tilde{U} &\to \tilde{U} \\ (x, z) &\mapsto (x, z - f(x)). \end{aligned}$$

Dann ist $\Phi$ stetig differenzierbar, die Umkehrabbildung $(x, z) \mapsto (x, z + f(x))$ ebenfalls und $\Phi(G(f)) = U \times \{0\}$. $\square$

Man nennt $n$-dimensionale Untermannigfaltigkeiten des $\mathbb{R}^{n+1}$ auch Hyperflächen.

**Definition 6.25 (Tangentialraum).** Sei $M \subset \mathbb{R}^m$ eine $n$-dimensionale Mannigfaltigkeit. Die Menge

$$T_x M := \left\{ \gamma'(0) \middle| \, \gamma : I \to M \text{ differenzierbar}, I \text{ offen}, 0 \in I, \gamma(0) = x \right\}$$

heißt Tangentialraum von $M$ im Punkt $x$.

**Satz 6.26.** *Die Menge $T_x M$ ist ein $n$-dimensionaler Untervektorraum von $\mathbb{R}^m$.*

*Beweis.* Sei $U \subset \mathbb{R}^m$ eine offene Umgebung von $x$ und $\Phi : U \to V$ ein Diffeomorphismus so, dass $\Phi(M \cap U) \subset \mathbb{R}^n \times \{0\}$ und $\Phi(x) = 0$. Die Abbildung $\mathrm{D}\Phi^{-1}(0)$ ist eine invertierbare (bijektive) lineare Abbildung (Warum?). Wir betrachten die lineare Abbildung

$$\begin{aligned} A : \mathbb{R}^n \times \{0\} &\to T_x M \\ w &\mapsto \mathrm{D}\Phi^{-1}(0) w \end{aligned}$$

Zunächst müssen wir zeigen, dass $A$ tatsächlich nach $T_x M$ abbildet. Sei dazu $w \in \mathbb{R}^n \times \{0\}$ und setze

$$\begin{aligned} \gamma : (-\epsilon, \epsilon) &\to M \\ t &\mapsto \Phi^{-1}(tw) \end{aligned}$$

für $\epsilon > 0$ klein genug. Dann gilt $\gamma(0) = x$ und

$$\gamma'(0) = \mathrm{D}\Phi^{-1}(0) w = Aw \in T_x M.$$

Die Abbildung $A$ ist injektiv (Warum?) und wir zeigen, dass sie auch surjektiv ist. Sei dazu $v \in T_x M$, das heißt es existiert $\gamma : (-1, 1) \to M$ so, dass $\gamma(0) = x$ und $\gamma'(0) = v$. Für $\epsilon > 0$ klein genug liegt $\gamma((-\epsilon, \epsilon))$ in $M \cap U$. Dann ist $\Phi \circ \gamma(t) \subset \mathbb{R}^n \times \{0\} \cap V$ und damit ist

$$w := (\Phi \circ \gamma)'(0) = \mathrm{D}\Phi(\gamma(0))\gamma'(0) = \mathrm{D}\Phi(x) v$$

ebenfalls in $\mathbb{R}^n \times \{0\}$. Es gilt dann $Aw = \mathrm{D}\Phi^{-1}(0)\mathrm{D}\Phi(x) v = \mathrm{D}(\Phi^{-1} \circ \Phi)(x) v = v$, also ist $A$ surjektiv.

Damit ist $T_x M$ als Bild der linearen Abbildung $A$ ein Untervektorraum und da $A$ bijektiv ist, bildet es jede Basis von $\mathbb{R}^n \times \{0\}$ auf eine Basis von $T_x M$ ab. Der Raum $T_x M$ hat also Dimension $n$. $\square$

*Bemerkung 6.27.* (i) Falls die Untermannigfaltigkeit $M = V$ ein Untervektorraum ist, dann sieht man leicht, dass $T_xV = V$ gilt.

(ii) Im Allgemeinen hängt der Tangentialraum $T_xM$ vom Punkt $x$ ab. Wir können also immer nur von Tangentenvektoren in einem Punkt reden.

(iii) Wir haben im Beweis gesehen, dass die Ableitung $\mathrm{D}\Phi^{-1}(0)$ den Tangentialraum von $V \cap \mathbb{R}^n \times \{0\}$ im Punkt $\Phi(x)$, also $\mathbb{R}^n \times \{0\}$, bijektiv auf $T_xM$ abbildet.

(iv) Die Menge

$$TM := \{(x,v) \in \mathbb{R}^m \times \mathbb{R}^m |\; x \in M \text{ und } v \in T_xM\}$$

heißt Tangentialbündel von $M$ und ist eine Untermannigfaltigkeit von $\mathbb{R}^m \times \mathbb{R}^m$.

(v) Solche Tangentialbündel sind der mathematische Hintergrund für die Lagrange-Formulierung der klassischen Mechanik. Bewegt sich eine Punktmasse (durch Zwangsbedingungen eingeschränkt) auf einer Mannigfaltigkeit $M$, dann charakterisieren die (lokalen) generalisierten Koordinaten $q_i$ Punkte in $M$, die generalisierten Geschwindigkeiten $\dot{q}_i$ Punkte im Tangentialraum $T_xM$ und das Tupel $(q_i, \dot{q}_i)$ Punkte in $TM$. Für eine Kurve $\gamma : [0,1] \to M$ erhalten wir eine zugehörige Kurve $\tilde{\gamma} : [a,b] \to TM$ indem wir setzen $\tilde{\gamma}(t) = (\gamma(t), \gamma'(t))$. Die Lagrange-Funktion $L$ ist eine Funktion auf dem Tangentialbündel und die Bahnen physikalischer Teilchen sind stationären Punkte der Abbildung

$$\gamma \mapsto \int_a^b L(\tilde{\gamma}(t))\mathrm{d}t.$$

(vi) Der Tangentialraum lässt sich für beliebige Teilemengen von $\mathbb{R}^m$ definieren, ist jedoch dann im Allgemeinen kein $n$-dimensionaler Vektorraum mehr. So kann man beispielsweise sehen, dass $K_1(0)$ und $(\mathbb{R} \times \{0\}) \cup (\{0\} \times \mathbb{R})$ keine Mannigfaltigkeiten sind.

Bezeichne mit $H_n$ den $n$-dimensionalen Halbraum, also

$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n |\; x_n \leq 0\}.$$

**Definition 6.28 (Mannigfaltigkeit mit Rand).** Eine Teilmenge $M \subset \mathbb{R}^m$ heißt ($n$-dimensionale) Mannigfaltigkeit mit Rand, falls für jedes $x \in M$ eine Umgebung $U$ von $x$, $V \subset \mathbb{R}^m$ offen und ein Diffeomorphismus $\Phi : U \to V$ existiert so, dass $\Phi(M) \subset H_n \times \{0\}$. Die Punkte $x \in M$ mit $\Phi(x) \in \partial H_n \times \{0\} = \mathbb{R}^{n-1} \times \{0\}$ heißen Randpunkte. Die Menge aller Randpunkte wird mit $\partial M$ bezeichnet

*Beispiel 6.29.* (i) Das abgeschlossene Intervall $[a,b]$ ist eine Mannigfaltigkeit mit Rand mit $\partial\,[a,b] = \{a,b\}$.

(ii) Die abgeschlossen Einheitskugel $K_1(0) \subset \mathbb{R}^n$ ist eine Mannigfaltigkeit mit Rand und ihr Rand ist $\partial K_1(0) = S^{n-1}$.

(iii) Jede (gewöhnliche) Mannigfaltigkeit $M$ ist eine Mannigfaltigkeit mit Rand $\partial M = \emptyset$. Beispielsweise gilt $\partial S^n = \emptyset$.

(iv) Der Einheitswürfel $\left\{(x,y) \subset \mathbb{R}^2 \middle|\ |x| \leq 1, |y| \leq 1\right\}$ und das Koordinatenkreuz $(\mathbb{R}^n \times \{0\}) \cup (\{0\} \times \mathbb{R}^n)$ sind keine Mannigfaltigkeiten mit Rand.

*Bemerkung 6.30.* (i) Ob ein Punkt $x$ von $\Phi$ auf den Rand $\partial H_n \times \{0\}$ abgebildet wird hängt nicht von $\Phi$ ab (Warum?).

(ii) Der Rand einer Mannigfaltigkeit $M$ mit Rand ist eine $n-1$-dimesionale Mannigfaltigkeit. Sei dazu $x \in \partial M$. Dann gibt es $U$ offene Umgebung von $x$, $V \subset \mathbb{R}^m$ offen und einen Diffeomorphismus $\Phi : U \to V$ so, dass $M \subset H_n \times \{0\}$. Dann gilt nach Definition des Randes $\Phi(\partial M \cup U) \subset \mathbb{R}^{n-1} \times \{0\}$.

(iii) Eine Mannigfaltigkeit $M$ mit Rand ist eine Mannigfaltigkeit genau dann wenn $\partial M = \emptyset$.

(iv) Der Rand im oben definierten Sinne stimmt meist nicht mit dem topologischen Rand von $M$ als Teilmenge von $\mathbb{R}^m$ überein. Für den Halbkreis

$$M := \left\{(x,y) \in \mathbb{R}^2 \middle|\ x^2 + y^2 = 1, x \geq 0\right\}$$

besteht der Rand im Mannigfaltigkeitensinne aus den Punkten $(1,0), (-1,0)$. Der topologische Rand ist jedoch ganz $M$.

## 6.3 Oberflächenintegrale

**Definition 6.31.** Sei $M$ eine $n$-dimensionale Mannigfaltigkeit und $V \subset \mathbb{R}^n$ offen. Eine injektive, stetig differenzierbare Abbildung $\varphi : V \to M$ heißt Parametrisierung falls $\mathrm{D}\varphi(t)$ injektiv ist für alle $t \in V$.

Eine Familie von Parametrisierungen $(\varphi_\alpha)_{\alpha \in I}$ ($I$ beliebige Indexmenge) heißt Atlas von $M$, falls

$$M \subset \varphi_\alpha(V_\alpha).$$

Wir nennen $M$ endlich parametrisierbar, falls es einen endlichen Atlas gibt.

Im Zusammenhang mit Parametrisierungen werden wir Diffeomorphismen $T : \tilde{V} \to V$ auch als Reparametrisierungen bezeichnen. Wir nennen $T$ orientierungserhaltend falls $\det(\mathrm{D}T) > 0$ andernfalls orientierungsumkehrend.

*Bemerkung 6.32.* (i) Die Umkehrabbildungen der Karten aus Definition 6.19 sind Parametrisierungen, jede Mannigfaltigkeit besitzt also einen Atlas.

(ii) Die Abbildung $\mathrm{D}\varphi(t)$ bildet den Tangentialraum $T_t V$ in den Tangentialraum $T_{\varphi(t)} M$ ab (Warum?). Da diese Räume gleiche Dimension haben, ist $\mathrm{D}\varphi(t)$ injektiv, genau dann wenn es es surjektiv auf $T_{\varphi(t)} M$ abbildet. Um festzustellen, ob eine differnzierbare Abbildung $\varphi : V \to M$ eine Parametrisierung ist kann man also alternativ überprüfen ob die Vektoren $\partial_1 \varphi, \ldots, \partial_n \varphi$ linear unabhängig sind.

(iii) Seien $\varphi_1, \varphi_2$ Parametrisierungen von $M$. Setze $U = \varphi_1(V_1) \cap \varphi_2(V_2)$. Dann ist $\varphi_1^{-1} \circ \varphi_2$ ein Diffeomorphismus von $\varphi_2^{-1}(U)$ nach $\varphi_1^{-1}(U)$.

Ohne Beweis.

**Definition 6.33 (Kreuzprodukt).** Seien $v, w \in \mathbb{R}^3$ und $n \in \mathbb{R}^3$ normiert und orthogonal zu $v$ und $w$. Dann ist

$$v \times w := n \det(n, v, w).$$

*Bemerkung 6.34.* (i) Die Definition ist eindeutig, da falls $v, w$ linear unabhängig sind, $n$ bis auf das Vorzeichen eindeutig definiert ist. Falls $v, w$ linear abhängig sind, verschwindet die rechte Seite unabhängig von $n$.

(ii) Aus der Definition ist ersichtlich, dass das Kreuzprodukt linear in $v$ und $w$ und antisymmetrisch unter Vertauschen von $v$ und $w$ ist.

(iii) Für $v = (v_1, v_2, v_3)$ und $w = (w_1, w_2, w_3)$ rechnet man leicht nach, dass

$$u = (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1)$$

orthogonal zu $v$ und $w$ ist und dass $\det(u, v, w) = \|u\|^2$ gilt. Damit folgt

$$v \times w = \frac{u}{\|u\|} \det\left(\frac{u}{\|u\|}, v, w\right) = u \frac{\det(u, v, w)}{\|u\|^2} = u.$$

Von nun an sei stets $M \subset \mathbb{R}^3$ eine 2-dimensionale Untermannigfaltigkeit also eine Fläche.

**Definition 6.35 (Oberflächenelement).** Sei $\varphi : V \to M$ eine Parametrisierung. Dann heißt

$$\begin{aligned}\sigma_\varphi : V &\to \mathbb{R}^3 \\ t &\mapsto \partial_1\varphi(t) \times \partial_2\varphi(t)\end{aligned}$$

das vektorielle und

$$\|\sigma_\varphi\|(t) = \|\partial_1\varphi(t) \times \partial_2(t)\varphi\|$$

das skalare Oberflächenelement (Flächenelement) von $\varphi$.

*Bemerkung 6.36.* Das Oberflächenelement läßt sich für beliebige differenzierbare Abbildungen $\varphi : V \to \mathbb{R}^3$ definieren. Die Abbildung $\varphi$ ist dann eine Parametrisierung genau dann, wenn $\sigma_\varphi$ (oder $\|\sigma_\varphi\|$) nirgends verschwindet, da dann die Vektoren $\partial_1\varphi(t)$ und $\partial_2\varphi(t)$ überall linear unabhängig sind.

*Beispiel 6.37.* (i) Für die Einheitssphäre sei

$$\begin{aligned}\psi : (0, 2\pi) \times (0, \pi) &\to \mathbb{S}^2 \\ (\varphi, \theta) &\mapsto (\sin(\theta)\cos(\varphi), \sin(\theta)\sin(\varphi), \cos(\theta))\end{aligned}$$

gegeben. Wir berechnen

$$\begin{aligned}\partial_1\psi(\varphi,\theta) &= (-\sin(\theta)\sin(\varphi), \sin(\theta)\cos(\varphi), 0), \\ \partial_2\psi(\varphi,\theta) &= (\cos(\theta)\cos(\varphi), \cos(\theta)\sin(\varphi), -\sin(\theta)) \text{ und} \\ \partial_1\psi(\varphi,\theta) \times \partial_2\psi(\varphi,\theta) &= (-\sin(\theta)^2\cos(\varphi), -\sin(\theta)^2\sin(\varphi), -\sin(\theta)\cos(\theta)).\end{aligned}$$

Da der letzte Ausdruck nirgends verschwindet, ist $\psi$ eine Parametrisierung. Diese deckt nicht die gesamte Kugel ab (das ist mit nur einer Parametrisierung nicht möglich). Schließlich können wir das skalare Flächenelement berechnen

$$\|\sigma_\psi\|(\varphi,\theta) = \sin(\theta).$$

(ii) Sei $U \subset \mathbb{R}^2$ offen, $f : U \to \mathbb{R}$ differenzierbar und $M$ der Graph von $f$. Dann ist $\varphi(x, y) \mapsto (x, y, f(x, y))$ eine Parametrisierung die ganz $M$ abdeckt. Das sieht man daran, dass $\partial_1\varphi = (1, 0, \partial_1 f(x, y))$ und $\partial_2\varphi = (0, 1, \partial_2 f(x, y))$ linear unabhängig sind. Dann ergeben sich die Flächenelemente

$$\begin{aligned}\sigma_\varphi(x,y) &= (-\partial_1 f(x,y), -\partial_2 f(x,y), 1) \\ \|\sigma_\varphi\|(x,y) &= \sqrt{(\partial_1 f(x,y))^2 + (\partial_1 f(x,y))^2 + 1}.\end{aligned}$$

**Satz 6.38.** *Sei $T : \tilde{V} \to V$ eine Reparametrisierung und $\varphi : V \to M$ eine Parametrisierung. Setze $\psi = \varphi \circ T$. Dann gilt*

$$\sigma_\psi(s) = \det(\mathrm{D}T(s))\sigma_\varphi(T(s)).$$

*Beweis.* Sei $s \in \tilde{V}$ und $n$ ein Einheitsvektor orthogonal zu $T_{\varphi(T(s))}M$. Dann gilt nach der Kettenregel $\mathrm{D}\psi(s) = \mathrm{D}\varphi(T(s))\mathrm{D}T(s)$. Es gilt

$$(n, \mathrm{D}\psi(s)) = (n, \mathrm{D}\varphi(T(s))) \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{D}T(s) \end{pmatrix}$$

wobei wir Blockmatrixschreibweise benutzt haben. Damit gilt

$$\begin{aligned} \sigma_\psi(s) &= \partial_1(\psi)(t) \times \partial_2(\psi)(t) = n \det(n, \mathrm{D}\psi(s)) \\ &= \det(n, \mathrm{D}\varphi(T(s)) \det \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{D}T(s) \end{pmatrix} = \det(\mathrm{D}T(s))\sigma_\varphi(T(s)). \end{aligned}$$ □

**Definition 6.39 (Skalares Oberflächenintegral).** Sei $\varphi : V \to M$ eine Parametrisierung und $f : M \to \mathbb{R}$ eine Funktion. Das Oberflächenintegral von $f$ über $\varphi(M)$ ist

$$\int_\varphi f \mathrm{d}\left\|\sigma\right\| := \int_V f(\varphi(t)) \left\|\sigma_\varphi\right\|(t) \mathrm{d}\lambda^2(t),$$

falls das Integral auf der rechten Seite existiert.

*Bemerkung 6.40.* Falls $T : \tilde{V} \to V$ eine Reparametrisierung ist und $\psi = \varphi \circ T$, dann folgt aus der Transformationsformel

$$\begin{aligned} \int_\varphi f \mathrm{d}\left\|\sigma\right\| &= \int_V f(\varphi(t)) \left\|\sigma_\varphi\right\|(t) \mathrm{d}\lambda^2(t) = \int_{\tilde{V}} f(\varphi(T(s))) \left\|\sigma_\varphi\right\|(T(s)) \left|\det \mathrm{D}T(s)\right| \mathrm{d}\lambda^2(s) \\ &= \int_{\tilde{V}} f(\psi(s)) \left\|\sigma_\psi\right\|(s) \mathrm{d}\lambda^2(s) = \int_\psi f \mathrm{d}\left\|\sigma\right\|. \end{aligned}$$

Nach der Transformationsformel existiert das eine Integral genau dann, wenn das andere existiert. Das Integral (auch seine Existenz) hängt also nicht von der gewählten Parametrisierung sondern lediglich vom parametrisierten Teil von $M$ also von $\varphi(V) = \psi(\tilde{V})$ ab.

**Definition 6.41.** Sei $f : \mathbb{R}^m \to \mathbb{R}$ eine Funktion. Dann heißt die Menge

$$\operatorname{supp} f = \overline{\{x \in \mathbb{R}^m \mid f(x) \neq 0\}}$$

der Träger von $f$. Wir sagen dass $f$ kompakten Träger hat, falls $\operatorname{supp} f$ kompakt ist. Die Menge solcher Funktionen wird mit $\mathbf{C}_c$ (stetig), $\mathbf{C}_c^r$ ($r$-fach differenzierbar) bezeichnet.

**Definition 6.42 (Zerlegung der Eins).** Sei $M \subset \mathbb{R}^m$ eine Menge und sei $(U_\alpha)_{\alpha \in I}$ ($I$ Indexmenge) eine offene Überdeckung von $M$, das heißt die $U_\alpha$ sind offene Mengen so, dass

$$M \subset \bigcup_{\alpha \in I} U_\alpha.$$

Eine Familie von nichtnegativen Funktionen $(f_\beta)_{\beta \in J}$ heißt Zerlegung der Eins für $(U_\alpha)$, falls

(i) für jedes $\beta \in J$ existiert $\alpha \in I$ so, dass $\operatorname{supp} f_\beta \subset U_\alpha$,

(ii)

$$\sum_{\beta \in J} f_\beta(x) = 1$$

gilt für alle $x \in M$ und

(iii) für jedes $x \in \mathbb{R}^m$ existiert eine Umgebung $U$ so, dass $\{\beta \in J | \operatorname{supp} f \cap U \neq \emptyset\}$ endlich ist.

**Satz 6.43.** *Sei $M \subset \mathbb{R}^m$ abgeschlossen und $(U_\alpha)_{\alpha \in I}$ eine offene Überdeckung von $M$. Dann existiert eine glatte Zerlegung der Eins für $(U_\alpha)$.*

*Beweisskizze.* Nur für beschränkes also kompaktes $M$. Wir verschaffen uns zunächst glatte Funktionen mit beliebig kleinen Trägern. Die Funktion

$$\begin{aligned} g : \mathbb{R} &\to \mathbb{R} \\ t &\mapsto \begin{cases} 0 & t \leq 0 \\ e^{-\frac{1}{t}} & t > 0 \end{cases} \end{aligned}$$

ist glatt. Dann ist

$$\begin{aligned} \delta : \mathbb{R}^m &\to \mathbb{R} \\ x &\mapsto g(1 - \|x\|^2) \end{aligned}$$

eine glatte Funktion (Verkettung zweier glatter Funktionen) mit $\operatorname{supp} \delta = K_1(0)$. Für $x_0 \in \mathbb{R}^m$ und $\epsilon > 0$ ist dann die Funktion

$$\begin{aligned} \delta_{\epsilon, x_0} : \mathbb{R}^m &\to \mathbb{R} \\ x &\mapsto \delta\left(\frac{x - x_0}{\epsilon}\right) \end{aligned}$$

glatt und hat Träger $K_\epsilon(x_0)$.

Wähle nun endlich viele $x_i \in M$, $\alpha_i \in I$ sowie $\epsilon_i > 0$ so, dass $K_{\epsilon_i}(x_i) \subset U_{\alpha_i}$ und

$$M \subset \bigcup_{i=1}^{n} B_{\epsilon_i}(x_i).$$

Das ist möglich, da $M$ kompakt ist. Dann ist

$$f = \sum_{i=1}^{n} \delta_{\epsilon_i, x_i}$$

auf $M$ stets größer Null und die Funktionen $\delta_{\epsilon_i,x_i}/f$ sind eine Zerlegung der Eins für $(U_\alpha)$. □

**Definition 6.44.** Sei $M$ endlich parametrisierbar, $\varphi_1, \ldots, \varphi_l$ ein Atlas und $\rho_1, \ldots, \rho_l$ eine zugehörige Zerlegung der Eins auf $M$, das heißt $\operatorname{supp} \rho_i \subset \varphi_i(V_i)$. Dann wird das Integral von $f$ über $M$ wie folgt definiert

$$\int_M f \mathrm{d} \|\sigma\| := \sum_{i=1}^{l} \int_{\varphi_i} \rho_i f \mathrm{d} \|\sigma\|,$$

falls alle auftretenden Integrale existieren. In diesem Fall heißt $f$ integrierbar über $M$.

*Bemerkung 6.45.* (i) Die Definition hängt nicht von der Wahl des Atlas und der Zerlegung der Eins ab, dann falls $\psi_1, \ldots, \psi_k$ eine weiterer Atlas ist und $\theta_1, \ldots, \theta_k$ eine zugehörige Zerlegung der Eins, dann gilt falls die auftretenden Integrale existieren

$$\begin{aligned}\sum_{i=1}^{l} \int_{\varphi_i} \rho_i f \mathrm{d} \|\sigma\| &= \sum_{i=1}^{l} \sum_{j=1}^{k} \int_{\varphi_i} \rho_i \theta_j f \mathrm{d} \|\sigma\| = \sum_{j=1}^{k} \sum_{i=1}^{l} \int_{\psi_j} \rho_i \theta_j f \mathrm{d} \|\sigma\| \\ &= \sum_{j=1}^{k} \int_{\psi_j} \theta_j f \mathrm{d} \|\sigma\|.\end{aligned}$$

Für nichtnegative $f$ gilt die Gleichung unabhängig von der Existenz der Integrale. Angewandt auf $|f|$ bedeutet das, dass auch die Integrierbarkeit von $f$ über $M$ vom Atlas und der Zerlegung unabhängig ist.

(ii) Auf die endliche Parametrisierbarkeit kann verzichtet werden, nur ist dann mehr Vorsicht bei Operationen mit unendlichen Summen geboten.

(iii) Es ist für die Definition nicht notwendig eine glatte Zerlegung der Eins (wie in Satz 6.43) zu verwenden. Häufig ist es jedoch hilfreich Stetigkeits- oder Differenzierbarkeitseigenschaften von $f$ zu erhalten.

**Definition 6.46 (Vektorielles Oberflächenintegral).** Sei $E : M \to \mathbb{R}^3$ ein Vektorfeld und $\varphi : V \to M$ eine Parametrisierung. Das Oberflächenintegral von $E$ über $\varphi(V)$ ist

$$\int_{\varphi} E \mathrm{d}\sigma = \int_V \langle E(\varphi(t)) | \sigma_\varphi(t) \rangle \mathrm{d}\lambda^2(t)$$

falls das Integral existiert.

*Bemerkung 6.47.* Um festzustellen, ob ein Vektorfeld integrierbar ist, ist oft die aus der Cauchy-Schwarz-Ungleichung folgende Abschätzung hilfreich.

$$\left|\int_\varphi E\mathrm{d}\sigma\right| \leq \int_V |\langle E(\varphi(t))|\,\sigma_\varphi(t)\rangle|\,\mathrm{d}\lambda^2(t) \leq \int_V \|E(\varphi(t))\|\,\|\sigma_\varphi(t)\|\,\mathrm{d}\lambda^2(t) = \int_\varphi \|E\|\,\mathrm{d}\,\|\sigma\|$$

*Bemerkung 6.48.* Analog zu Bemerkung 6.40 ist das vektorielle Oberflächenintegral invariant unter einer Reparametrisierung $T$ vorausgesetzt $T$ ist orientierungserhaltend. Andernfalls wechselt das Integral das Vorzeichen. Daraus ergeben sich Probleme bei der globalen Integraldefinition.

**Definition 6.49 (Orientierbare Fläche).** Die Mannigfaltigkeit $M$ heißt orientierbar, falls es ein stetiges Normalenvektorfeld gibt, das heißt eine stetige Abbildung $n : M \to \mathbb{R}^3$ so, dass $n(x)$ orthogonal zu $T_xM$ und $\|n\| = 1$ ist für alle $x \in M$. In diesem Fall heißt $n$ eine Orientierung für $M$.

Eine Parametrisierung $\varphi : V \to M$ heißt orientiert (bezüglich $n$), falls $\langle \sigma_\varphi(t)|\,n(\varphi(t))\rangle > 0$ ist für alle $t \in V$. Ein Atlas $(\varphi_\alpha)_{\alpha\in I}$ heißt orientiert, falls $\varphi_\alpha$ orientiert ist für jedes $\alpha \in I$.

*Bemerkung 6.50.* (i) Falls $n$ eine Orientierung ist, so ist $-n$ ebenfalls eine Orientierung. Falls $M$ (weg)zusammenhängend ist gibt es keine weiteren Orientierungen. Eine Parametrisierung ist orientiert, genau dann wenn $\sigma_\varphi = n\,\|\sigma_\varphi\|$ gilt.

(ii) Sei $\varphi : V \to M$ eine orientierte Parametrisierung und $T : \tilde{V} \to V$ eine Reparametrisierung. Dann ist $T$ orientierungserhaltend genau dann, wenn $\varphi \circ T$ orientiert ist.

**Definition 6.51.** Sei $M$ endlich parametrisierbar und orientierbar, $n : M \to \mathbb{R}^3$ eine Orientierung, $\varphi_1, \dots, \varphi_n$ ein orientierter Atlas und $\rho_1, \dots, \rho_n$ eine zugehörige Zerlegung der Eins. Sei $E : M \to \mathbb{R}^3$ ein Vektorfeld. Dann definieren wir

$$\int_M E\mathrm{d}\sigma = \sum_{i=1}^n \int_{\varphi_i} \rho_i E\mathrm{d}\sigma$$

falls alle auftretenden Integrale existieren.

*Bemerkung 6.52.* (i) Wie im skalaren Fall zeigt man, dass diese Definition nicht von der Wahl des Atlas oder der Zerlegung abhängt. Für die entgegengesetzte Orientierung $-n$ wechselt das Integral das Vorzeichen.

(ii) Es gilt

$$\left|\int_M E\mathrm{d}\sigma\right| \leq \int_M \|E\|\,\mathrm{d}\,\|\sigma\|\,.$$

## 6.4 Integralsätze

**Definition 6.53.** Sei $U \subset \mathbb{R}^3$ offen. Wir definieren die Differentialoperatoren $\operatorname{grad}: \mathbf{C}^1(U) \to \mathbf{C}(U, \mathbb{R}^3)$, $\operatorname{div}: \mathbf{C}^1(U, \mathbb{R}^3) \to \mathbf{C}(U)$, $\operatorname{rot}: \mathbf{C}^1(U, \mathbb{R}^3) \to \mathbf{C}(U, \mathbb{R}^3)$ und $\Delta: \mathbf{C}^2(U) \to \mathbf{C}(U)$ durch

$$\begin{aligned} \operatorname{grad} f &:= (\partial_1 f, \partial_2 f, \partial_3 f) \\ \operatorname{div} E &:= \partial_1 E_1 + \partial_2 E_2 + \partial 3 E_3 \\ \operatorname{rot} E &:= (\partial_2 E_3 - \partial_3 E_2, \partial_3 E_1 - \partial_1 E_3, \partial_1 E_2 - \partial_2 E_1) \\ \Delta g &:= \partial_1^2 g + \partial_2^2 g + \partial_3^2 g. \end{aligned}$$

für $f \in \mathbf{C}^1(U)$, $g \in \mathbf{C}^2(U)$ und $E \in \mathbf{C}^1(U, \mathbb{R}^3)$.

*Bemerkung 6.54.* (i) Aus der Definition ist ersichtlich, dass die Differentialoperatoren linear sind. Sie wirken üblicherweise auf alles was rechts von ihnen steht bis zum nächsten $+$ oder $-$ -Zeichen. Der Ausdruck

$$f \Delta g h + l$$

wird also als

$$f(\Delta(gh)) + l$$

interpretiert.

(ii) Die Operatoren $\operatorname{grad}$, $\operatorname{div}$ und $\Delta$ lassen sich ohne Weiteres auf $\mathbb{R}^n$ für beliebiges $n$ erweitern.

(iii) Häufig führt man formal den vektorwertigen Nabla-Operator $\nabla := (\partial_1, \partial_2, \partial_3)$ ein, um die obigen Definitionen formal zusammenzufassen:

$$\begin{aligned} \operatorname{grad} f &= \nabla f \\ \operatorname{div} E &= \langle \nabla | E \rangle \\ \operatorname{rot} E &= \nabla \times E \\ \Delta f &= \langle \nabla | \nabla \rangle. \end{aligned}$$

Die obigen Skalar- und Kreuzprodukte sind nicht definiert, die Formeln dienen also primär als Gedächtnisstütze.

(iv) Anhand der Definitionen rechnet man leicht die folgenden Rechenregeln nach:

$$\begin{aligned} \operatorname{div} \operatorname{rot} E &= 0 \\ \operatorname{rot} \operatorname{grad} E &= 0 \\ \operatorname{div} \operatorname{grad} f &= \Delta f \\ \operatorname{grad}(fg) &= (\operatorname{grad} f) g + f \operatorname{grad} g \\ \operatorname{div}(fE) &= \langle \operatorname{grad} f | E \rangle + f \operatorname{div} E \\ \operatorname{rot}(fE) &= (\operatorname{grad} f) \times E + f \operatorname{rot} E \\ \Delta(fg) &= (\Delta f) g + 2 \langle \operatorname{grad} f | \operatorname{grad} g \rangle + f \Delta g \end{aligned}$$

*Bemerkung 6.55.* Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^n$ sei stetig mit kompaktem Träger. Dann ist $f$ beschränkt und $\lambda^n$-integrierbar.

*Beweis.* Nach Voraussetzung ist $\operatorname{supp} f$ beschränkt und abgeschlossen. Da $f$ stetig ist, gilt nach Satz 1.64 $|f| \leq K$ für passendes $K$ auf $\operatorname{supp} f$, also auch auf ganz $U$. Da $\operatorname{supp} f$ beschränkt ist, gilt dann

$$\int_U |f| \,\mathrm{d}\lambda^n \leq \lambda^n(\operatorname{supp} f)K,$$

also ist $f$ integrierbar. □

**Satz 6.56 (Satz von Green).** *Sei $U \subset \mathbb{R}^2$ offen und $\gamma : I \to \mathbb{R}^2$ eine stückweise reguläre Kurve, die $U$ in zwei Teile zerschneidet. Sei $D \subset U$ der (offene) Teil von $U$, der links der Kurve liegt (man sagt $\gamma$ ist bezüglich $D$ positiv orientiert). Seien $F, G : U \to \mathbb{R}$ stetig differenzierbare Funktionen mit kompakten Trägern* $\operatorname{supp} F, \operatorname{supp} G \subset U$. *Dann gilt*

$$\int_D (\partial_1 G(x,y) - \partial_2 F(x,y)) \,\mathrm{d}\lambda^2(x,y) = \int_\gamma (F\mathrm{d}x + G\mathrm{d}y). \tag{6.2}$$

*Beweisskizze.* Wir betrachten zunächst den Spezialfall wo $\gamma$ der Graph einer Funktion ist. Da beide Seiten der zu zeigenden Gleichung translationsinvariant sind, können wir oBdA voraussetzen, dass $f < 0$ und $\operatorname{supp} F, \operatorname{supp} G$ in der unteren Halbebene enthalten sind. Sei also $U = (0,1) \times \mathbb{R}$, $f : [0,1] \to \mathbb{R}$ stetig differenzierbar und $\gamma(t) = (t, f(t))$. Bei den im folgenden auftretenden Integralen sind die Integranden stets stetig und haben kompakten Träger. Daher dürfen wir nach Bemerkung 6.55 den Satz von Fubini anwenden. Wir berechnen zunächst mit Hilfe des Hauptsatzes

$$-\int_D \partial_2 F(x,y)\mathrm{d}\lambda^2(x,y) = -\int_0^1 \int_{f(x)}^0 \partial_2 F(x,y)\mathrm{d}y\mathrm{d}x = \int_0^1 F(x, f(x))\mathrm{d}x.$$

Andererseits gilt

$$\int_\gamma F\mathrm{d}x = \int_0^1 F(t, f(t))\mathrm{d}x(1, f'(t))\mathrm{d}t = \int_0^1 F(t, f(t))\mathrm{d}t$$

und wir haben eine „Hälfte“ der Gleichung (6.2) gezeigt.

Betrachte nun für $s \in \mathbb{R}$ die Kurve $\gamma_s(t) = (t, sf(t))$ und

$$I(s) := \int_{\gamma_s} G\mathrm{d}y = \int_0^1 G(t, sf(t))\mathrm{d}y(1, sf'(t))\mathrm{d}t = \int_0^1 sf'(t)G(t, sf(t))\mathrm{d}t.$$

Der Integrand erfüllt für $s \in [-2,2]$ und $t \in [0,1]$

$$\frac{\mathrm{d}}{\mathrm{d}s}(sf'(t)G(t, sf(t))) = f'(t)G(t, sf(t)) + sf(t)f'(t)\partial_2 G(t, sf(t)).$$

Dieser Ausdruck ist (unabhängig von $s$ und $t$) beschränkt, also integrierbar beschränkt. Damit können wir Korollar 5.14 anwenden und erhalten, dass $I(s)$ differenzierbar ist und es gilt

$$\begin{aligned}
I(1) = I(1) - I(0) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s} I(s) \mathrm{d}s \\
&= \int_0^1 \int_0^1 \big(f'(t)G(t, sf(t)) + sf(t)f'(t)\partial_2 G(t, sf(t))\big) \,\mathrm{d}t\mathrm{d}s.
\end{aligned}$$

Andererseits ist mit der Substitution $y = f(x)z$

$$\int_D \partial_1 G(x,y)\mathrm{d}\lambda^2(x,y) = \int_0^1 \int_{f(x)}^0 \partial_1 G(x,y)\mathrm{d}y\mathrm{d}x = \int_0^1 \int_1^0 f(x)\partial_1 G(x, f(x)z)\mathrm{d}x\mathrm{d}z$$

Es gilt

$$\begin{aligned}
0 &= \int_0^1 (f(1)G(1, f(1)z) - f(0)G(0, f(0)z))\mathrm{d}z = \int_0^1 \int_0^1 \frac{\mathrm{d}}{\mathrm{d}x} \left(f(x)G(x, f(x)z)\right) \mathrm{d}x\mathrm{d}z \\
&= \int_0^1 \int_0^1 \big(f'(x)G(x, f(x)z) + f(x)\partial_1 G(x, f(x)z) + zf(x)f'(x)\partial_2 G(x, zf(x))\big) \,\mathrm{d}z\mathrm{d}x
\end{aligned}$$

und damit

$$\int_D \partial_1 G(x,y)\mathrm{d}\lambda^2(x,y) = \int_0^1 \int_0^1 \big(f'(x)G(x, f(x)z) + zf(x)f'(x)\partial_2 G(x, zf(x))\big) \,\mathrm{d}x\mathrm{d}z.$$

Das zeigt die zweite „Hälfte“ der Gleichung (6.2) und beendet den Beweis des betrachteten Spezialfalles. Analog behandelt man den Fall $f : [a,b] \to \mathbb{R}$ sowie, durch Vertauschen der Variablen den Fall $\gamma(t) = (f(t), t)$.

Insbesondere erhalten wir, dass für $D$ offen und $\operatorname{supp} E, \operatorname{supp} F \subset D$ kompakt gilt $\int_D (\partial_1 G(x,y) - \partial_2 F(x,y)) \,\mathrm{d}\lambda^2(x,y) = 0$, indem wir $\operatorname{supp} E \cup \operatorname{supp} F$ in eine hinreichend große Box einschließen.

Für den allgemeinen Fall überdecken wir $(\operatorname{supp} E \cup \operatorname{supp} F) \cap \gamma(I)$ durch endlich viele offene Rechtecke $Q_1, \ldots, Q_n$ für die die Voraussetzungen des Spezialfalles erfüllt sind. Dann ist $Q_1, \ldots, Q_n, D$ eine offene Überdeckung der abgeschlossenen Menge $\overline{D}$ und wir können eine zugehörige glatte Zerlegung der Eins $\rho_1, \ldots, \rho_n, \rho_{n+1}$ wählen. Dann gilt mit den bereits gezeigten Spezialfällen

$$\begin{aligned}
\int_D (\partial_1 G - \partial_2 F)\mathrm{d}\lambda^2 &= \int_D (\partial_1 \sum_{k=1}^{n+1} \rho_k G - \partial_2 \sum_{k=1}^{n+1} \rho_k F)\mathrm{d}\lambda^2 = \sum_{k=1}^{n} \int_{D \cap Q_k} (\partial_1 \rho_k G - \partial_2 \rho_k F)\mathrm{d}\lambda^2 \\
&= \sum_{k=1}^{n} \int_\gamma (\rho_k F\mathrm{d}x + \rho_k G\mathrm{d}y) = \int (F\mathrm{d}x + G\mathrm{d}y).
\end{aligned}$$

Für dieses Argument ist entscheidend, dass die Teilung der Eins glatt ist (mindestens $\mathbf{C}^1$). $\square$

Im Folgenden seien $U$, $\gamma$ und $D$ stets so wie im vorhergehenden Satz.

*Bemerkung 6.57.* Häufig setzt man nicht $\operatorname{supp} F, \operatorname{supp} G \subset U$ voraus, sondern dass $\gamma$ geschlossen ist und $D$ die von $\gamma$ umschlossene Fläche. Wähle dazu eine stetig differenzierbare Funktion $\rho$ die auf $\overline{D}$ identisch 1 ist und kompakten Träger in $U$ hat und wende den obigen Satz auf $\rho F$ und $\rho G$ an. Durch die Multiplikation mit $\rho$ ändert sich keines der Integrale in (6.2).

**Korollar 6.58 (Satz von Gauß in 2D).** *Sei $E \in \mathbf{C}^1(U, \mathbb{R}^2)$ ein Vektorfeld mit kompaktem Träger und setze*

$$\sigma(t) = (\gamma_2(t), -\gamma_1(t)).$$

*Dann gilt*

$$\int_D \operatorname{div} E = \int_I \langle E(\gamma(t)) | \sigma(t) \rangle \mathrm{d}t.$$

*Beweis.* Setze $G = E_1$ und $F = -E_2$, dann erhalten wir

$$\begin{aligned}\int_D \operatorname{div} E \mathrm{d}\lambda^2 &= \int_I (-E_2(\gamma(t)) \mathrm{d}x(\gamma'(t)) + E_1(\gamma(t)) \mathrm{d}y(\gamma'(t))) \mathrm{d}t \\ &= \int_I \left(-E_2(\gamma(t))\gamma_1'(t) + E_1(\gamma(t))\gamma_2'(t)\right) \mathrm{d}t = \int_I \langle E(\gamma(t)) | \sigma(t) \rangle \mathrm{d}t. \qquad \square\end{aligned}$$

**Korollar 6.59 (Satz von Stokes in der $x$-$y$-Ebene).** *Sei $E \in \mathbf{C}^1(U, \mathbb{R}^3)$ ein Vektorfeld mit kompaktem Träger. Dann gilt*

$$\int_D E \mathrm{d}\sigma = \int_\gamma \omega_E.$$

*Beweis.* Die Abbildung $(x, y) \mapsto (x, y, 0)$ ist eine Parametrisierung der Mannigfaltigkeit $D \times \{0\}$ für die $\sigma(x, y) = (0, 0, 1)$ gilt und damit

$$\begin{aligned}\int_D E \mathrm{d}\sigma &= \int_D \langle \operatorname{rot} E(x, y, 0) | \sigma(x, y) \rangle \mathrm{d}x \mathrm{d}y = \int_D (\partial_1 E_2 - \partial_2 E_1) \mathrm{d}\lambda^2(x, y) \\ &= \int_I (E_1(\gamma(t), 0) \mathrm{d}x + E_2(\gamma(t), 0) \mathrm{d}y)(\gamma'(t)) \mathrm{d}t = \int_I \langle E | \gamma'(t) \rangle \mathrm{d}t = \int_I \omega_E \mathrm{d}t. \qquad \square\end{aligned}$$

*Bemerkung 6.60.* Auch für die beiden Korrollare gilt die entsprechende Version ohne kompakten Träger aber dafür mit geschlossener Kurve $\gamma$.

**Satz 6.61 (Satz von Gauß 3D).** *Sei $U \subset \mathbb{R}^3$ offen, $M \subset \mathbb{R}^3$ eine orientierbare, zweidimensionale Mannigfaltigkeit, die $U$ in zwei Teile zerschneidet und $D$ (offen) einer dieser Teile. Wähle die aus $D$ herauszeigende Orientierung. Sei $E \in \mathbf{C}^1(U, \mathbb{R}^3)$ ein Vektorfeld mit kompaktem Träger in $U$. Dann gilt*

$$\int_D \operatorname{div} E \mathrm{d}\lambda^3 = \int_M E \mathrm{d}\sigma.$$

*Ohne Beweis.*

**Satz 6.62.** *Sei $U \subset \mathbb{R}^3$ offen, $M \subset U$ eine orientierbare, zweidimensionale Mannigfaltigkeit mit Rand. Sei $n$ eine Orientierung von $M$ und $\gamma : I \to \mathbb{R}^3$ eine reguläre Parametrisierung des Randes von $M$ so orientiert, dass $M$ sich, aus Richtung $-n$ betrachtet, links von $\gamma$ befindet. Sei $E \in \mathbf{C}^1(U, \mathbb{R}^3)$ ein Vektorfeld mit kompaktem Träger in $U$. Dann gilt*

$$\int_M \operatorname{rot} E \mathrm{d}\sigma = \int_\gamma \omega_E = \int_I \big\langle E(\gamma(t)) \big| \gamma'(t) \big\rangle \mathrm{d}t.$$

*Ohne Beweis.*

*Bemerkung 6.63.* (i) Für die obigen Sätze gibt es jeweils auch eine Version ohne kompakte Träger, aber dafür mit geschlossener Randkurve beziehungsweise kompakter Oberfläche.

(ii) Der Satz von Gauß gilt analog in beliebigen Dimensionen.

**Korollar 6.64 (Greensche Formeln).** *Seien $U$, $M$ und $D$ wie in Satz 6.61 und $f, g \in \mathbf{C}^2(U, \mathbb{R})$ Funktionen mit kompaktem Träger. Sei $n$ die gewählte Orientierung. Dann gelten*

$$\int_D (f\Delta g + \langle \operatorname{grad} f | \operatorname{grad} g \rangle) \mathrm{d}\lambda^3 = \int_M f \operatorname{grad} g \mathrm{d}\sigma \text{ und}$$

$$\int_D (f\Delta g - g\Delta f) = \int_M (f \operatorname{grad} g - g \operatorname{grad} f) \mathrm{d}\sigma = \int_M (f\partial_n g - g\partial_n f) \mathrm{d} \|\sigma\|$$

*wobei $\partial_n$ die Richtungsableitung in Richtung $n$ bezeichnet.*

*Beweis.* Wir wenden Satz 6.61 auf das Vektorfeld $f \operatorname{grad} g$ an. Dann gilt mit Bemerkung 6.54

$$\int_D \operatorname{div} (f \operatorname{grad} g) \mathrm{d}\lambda^3 = \int_D (f\Delta g + \langle \operatorname{grad} f | \operatorname{grad} g \rangle) \mathrm{d}\lambda^3 = \int_M f \operatorname{grad} g \mathrm{d}\sigma.$$

Vertauschen wir $f$ und $g$ und subtrahieren das Ergebnis von dieser Gleichung, so erhalten wir

$$\int_D (f\Delta g - g\Delta f) = \int_M (f \operatorname{grad} g - g \operatorname{grad} f) \mathrm{d}\sigma.$$

Für die letzte Umformung sei $\varphi : V \to M$ eine orientierte Parametrisierung eines Teils von $M$. Dann gilt

$$\begin{aligned}\int_\varphi f \operatorname{grad} g \mathrm{d}\sigma &= \int_V \langle f(\varphi(t)) \operatorname{grad} g(\varphi(t)) | \sigma_\varphi(t) \rangle \mathrm{d}\lambda^2(t) \\ &= \int_V f(\varphi(t)) \langle \operatorname{grad} g(\varphi(t)) | n(\varphi(t)) \rangle \|\sigma_\varphi(t)\| \mathrm{d}\lambda^2(t) = \int_\varphi (f\partial_n g) \mathrm{d} \|\sigma\| .\end{aligned}$$

Zusammensetzen mittels Teilung der Eins ergibt die gesuchte Umformung. □

# 7 Partielle Differenzialgleichungen

## 7.1 Einführung

**Definition 7.1.** Eine partielle Differentialgleichung ist eine Gleichung in einer unbekannten Funktion $f : U \to \mathbb{R}$ ($U \subset \mathbb{R}^n$ offen), in der $f$ und Ableitungen von $f$ vorkommen. Die Ordnung der Gleichung ist die höchste vorkommende Ableitung. Eine partielle Differentialgleichung der Form

$$\sum_{\alpha \in \mathbb{N}^n, |\alpha| < k} a_\alpha \partial^\alpha f = b$$

mit $a_\alpha, b : U \to \mathbb{R}$ heißt linear.

*Beispiel 7.2.* (i) $\Delta f = 0$ Laplacegleichung: beschreibt das Gravitations- bzw. elektrostatische Potential ohne Massen/Ladungen.

(ii) $-\Delta f = \rho$ Poissongleichung: beschreibt das Gravitations- bzw. elektrostatische Potential bei vorgegebener Massen-/Ladungsdichte $\rho$.

(iii) $\partial_t f - \Delta f = \rho$ Wärmeleitungsgleichung: Zeitentwicklung einer Temperatur-oder Konzentrationsverteilung durch Wärmeleitung/Diffusion

(iv) $\mathrm{i}\partial_t f = -\Delta f + Vf$ Schrödingergleichung: Wellenfunktion eines quantenmechanischen Teilchens im Potential $V$

(v) $\partial_t^2 f - \Delta f = \rho$ Wellengleichung: Ausbreitung von elektromagnetischen oder Druckwellen

(vi)

$$\begin{aligned} \operatorname{rot} B &= \partial_t E + j & \qquad \operatorname{rot} E &= -\partial_t B \\ \operatorname{div} E &= \rho & \qquad \operatorname{div} B &= 0 \end{aligned}$$

Maxwellgleichungen: ($E, B : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$ gesucht) Elektromagnetisches Feld bei vorgegebener Stromdichte $j$ und Ladungsdichte $\rho$

(vii) Navier-Stokes-Gleichung: Beschreibung der Bewegung von Fluiden

(viii) Einstein-Gleichungen: Zusammenhang zwischen materiellem Inhalt und Geometrie der Raumzeit

*Beispiel 7.3.* Wir betrachten die Wellengleichung in $1 + 1$ Dimensionen, das heißt die Gleichung für $f \in \mathbf{C}^2(\mathbb{R} \times \mathbb{R}, \mathbb{R})$

$$\partial_t^2 f - \partial_x^2 f = 0.$$

Angenommen $f$ sei eine Lösung. Wir führen eine Variablentransformation durch: $\tilde{f}(u,v) := f(u-v,u+v)$. Dann gilt

$$\begin{aligned}\partial_u\partial_v\tilde{f}(u,v) &= \partial_u(-\partial_1 f(u-v,u+v) + \partial_2 f(u-v,u+v)) \\ &= -\partial_1^2 f(u-v,u+v) - \partial_2\partial_1 f(u-v,u+v) \\ &\qquad + \partial_1\partial_2 f(u-v,u+v) + \partial_2^2 f(u-v,u+v) = 0.\end{aligned}$$

Daraus folgt, dass es $\tilde{g} \in \mathbf{C}^1(\mathbb{R})$ gibt, so dass $\partial_v\tilde{f}(u,v) = \tilde{g}(v)$. Wähle $g$ eine Stammfunktion von $\tilde{g}$, dann muss es $h \in \mathbf{C}^2(\mathbb{R})$ geben so, dass $\tilde{f}(u,v) = g(v) + h(u)$. Rücksubstitution liefert

$$f(t,x) = g\left(\frac{1}{2}(x-t)\right) + h\left(\frac{1}{2}(x+t)\right).$$

Andererseits überzeugt man sich leicht, dass der obige Ausdruck für beliebige $g, h \in \mathbf{C}^2(\mathbb{R})$ die partielle Differentialgleichung löst.

*Bemerkung 7.4.* (i) Das Beispiel illustriert, dass der Lösungsraum von partiellen Differentialgleichungen im Allgemeinen sehr groß ist. Um ein „gut gestelltes Problem“ zu erhalten muss man zusätzliche Anfangs- und/oder Randbedingungen stellen. Diese werden entsprechend durch das zu modellierende System vorgegeben. Im obigen Fall sind die „richtigen“ Zusatzbedingungen z. B. durch Angabe von $f(0,x)$ und $\partial_1 f(0,x)$ gegeben und man spricht von einem Anfangswertproblem.

(ii) Die Wahl des richtigen Koordinatensystems kann die Behandlung von partiellen Differentialgleichungen stark vereinfachen. Häufig kann man sich dabei von Symmetrien des Systems leiten lassen.

*Bemerkung 7.5.* Häufig macht man einen Separationsansatz. Hat man Beispielsweise die Laplace-Gleichung in 2 Dimensionen

$$\partial_1^2 f + \partial_2^2 f = 0,$$

so macht man den Ansatz $f(x,y) = g(x)h(y)$ und formt zu

$$\frac{g''(x)}{g(x)} + \frac{h''(y)}{h(y)} = 0$$

um. Da $x$ und $y$ unabhängig voneinander variiert werden können, müssen die beiden Summanden in der Gleichung separat konstant sein, und das Problem zerfällt in die gewöhnlichen Differentialgleichungen

$$g''(x) = \lambda g(x) \text{ und } h''(y) = -\lambda h(y).$$

Der gewählte Ansatz ist allerdings eine starke Einschränkung und führt nur dann zum Erfolg, wenn die gesuchte Lösung „zufällig“ die vermutete Gestalt hat. Viele Anfangs- und Randbedingungen werden sich so aber nicht erfüllen lassen und auch Eindeutigkeitsaussagen für Lösungen kann man so nicht erhalten.