Analysis I und II

HS 2007 und FS 2008

Martin U. Schmidt

Inhaltsverzeichnis

Kapitel 1

Mengen und Abbildungen

1.1 Mengen

Georg Cantor(1845-1918) hat den Begriff der Menge definiert als „eine Zusammenfassung von wohlbestimmten und wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen“. Diese Objekte werden Elemente der Menge genannt. Um ein Objekt a als Element der Menge A zu kennzeichnen, schreiben wir $a \in A$. Ist a dagegen kein Element der Menge A, so schreiben wir $a \notin A$.

Wir können Mengen dadurch beschreiben, dass wir alle ihre Elemente angeben, also z.B. ist

$$A = \{a\}$$

die Menge, die nur das Element a enthält und B

$$B = \{a, b, c\}$$

die Menge, die drei Elemente a, b und c enthält. Dabei kann auch ein Element einer Menge wieder eine Menge sein:

$$C = \{a, \{a\}\}.$$

$$M = \{x \in X \mid x \text{ hat die Eigenschaft } p\} \quad \text{oder nur} \quad M = \{x \mid x \text{ hat die Eigenschaft } p\}$$

bezeichnet die Menge aller Elemente x (von der Menge X), die die Eigenschaft p haben.[1]

A ist eine Teilmenge von B, wenn alle Elemente von A auch Elemente von B sind. In Symbolen $A \subset B$.

[1] Bei dieser Beschreibung muss man allerdings Vorsicht walten lassen, um die Russellsche Antinomie zu vermeiden. Läßt man nämlich die Menge aller Mengen zu, die sich nicht selbst als Element enthalten, so wird nicht entscheidbar, ob diese Menge sich selbst als Element enthält oder nicht. Als Ausweg wird in der axiomatischen Mengenlehre die Frage, ob eine Menge Element einer Menge ist, nicht in allen Fällen als sinnvoll zugelassen.

1.2 Operationen von Mengen

Die Vereinigung zweier Mengen A und B ist die Menge $A \cup B$ aller Elemente, die in mindestens einer der beiden Mengen A und B enthalten sind. Der Durchschnitt zweier Mengen A und B ist die Menge $A \cap B$ aller Elemente, die sowohl Element von A als auch Element von B sind. Die Differenzmenge $A \setminus B$ ist die Menge aller Elemente von A, die nicht Element von B sind. Das kartesische Produkt der Mengen A und B ist die Menge aller geordneten Paare (a, b) von Elementen von A und B.

$$A \times B = \{(a, b) \mid a \in A \text{ und } b \in B\}$$

Die Potenzmenge $\mathcal{P}(M)$ einer Menge M ist die Menge aller Teilmengen von M. Dabei ist die leere Menge $\emptyset$ stets ein Element der Potenzmenge.
z.B. $\mathcal{P}(\{a, \{a\}\}) = \{\emptyset, \{a\}, \{\{a\}\}, \{a, \{a\}\}\}$.

Wenn wir nur Teilmengen einer vorgegebenen Menge M betrachten, dann wird für eine solche Menge $A \in \mathcal{P}(M)$ die Menge $M \setminus A$ auch als das Komplement A^c bezeichnet. Diese Operationen erfüllen die folgenden Regeln:

$$\begin{array}{rcc}
\text{(Idempotenz)} & A \cup A = A & A \cap A = A \\
\text{(Kommutativität)} & A \cup B = B \cup A & A \cap B = B \cap A \\
\text{(Assoziativität)} & A \cap (B \cap C) = (A \cap B) \cap C & A \cup (B \cup C) = (A \cup B) \cup C \\
\text{(Distributivität)} & A \cap (B \cup C) = (A \cap B) \cup (A \cap C) & A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \\
\text{(de Morgan)} & (A \cap B)^c = A^c \cup B^c & (A \cup B)^c = A^c \cap B^c
\end{array}$$

$$\begin{array}{ccc}
A \subset A \cup B & & A \cap B \subset A \\
A \cup \emptyset = A & & A \cap \emptyset = \emptyset \\
A \setminus A = \emptyset & & A \setminus \emptyset = A \\
(A^c)^c = A & & A \subset B \Longleftrightarrow B^c \subset A^c \\
(A \subset B \text{ und } B \subset A) & \Longleftrightarrow & A = B \\
A \cup B = B & \Longleftrightarrow & A \subset B \\
A \cap B = A & \Longleftrightarrow & A \subset B \\
(A \subset C \text{ und } B \subset C) & \Longleftrightarrow & (A \cup B) \subset C \\
(C \subset A \text{ und } C \subset B) & \Longleftrightarrow & C \subset (A \cap B)
\end{array}$$

$$A \times \emptyset = \emptyset \times A = \emptyset$$

$$(A \times C) \cup (B \times C) = (A \cup B) \times C \qquad (A \times C) \cap (B \times C) = (A \cap B) \times C.$$

1.3 Relationen

Als Relationen auf einer Menge bezeichnet man Aussagen über mehrere Elemente der Menge, die entweder wahr oder falsch sind. Wir wollen hier nur Aussagen über zwei Elemente einer Menge A betrachten. Jede solche Relation beschreiben wir durch eine Teilmenge R aller geordneten Paare in $A \times A$, in der wir alle die Paare zusammenfassen, für die die Aussage der Relation wahr ist. Wir sagen dann, dass $(a, b) \in A \times A$ diese Relation erfüllt, wenn (a, b) zu der Teilmenge R gehört. Andernfalls erfüllt (a, b) die Relation nicht. Wir führen jetzt folgende Eigenschaften einer Relation ein:

Reflexivität: für alle $a \in A$ erfüllt (a, a) die Relation.
Symmetrie: falls (a, b) die Relation erfüllt, dann auch (b, a).
Transitivität: falls (a, b) und (b, c) die Relation erfüllen, dann auch (a, c).

Definition 1.1. *Eine Äquivalenzrelation ist eine Relation, die reflexiv, symmetrisch und transitiv ist.*

Eine Äquivalenzrelation definiert dann sogenannte Äquivalenzklassen. Für jedes $a \in A$ ist die Äquivalenzklasse $[a]$ von a die Teilmengen aller Elemente $b \in A$, so dass (a,b) die Relation erfüllt. Die Transitivität impliziert, dass für jedes Element $b \in [a]$ die entsprechende Äquivalenzklasse $[b]$ eine Teilmenge von $[a]$ ist. Wenn zwei Äquivalenzklassen $[a]$ und $[b]$ beide ein Element $c \in A$ enthalten, dann sind wegen der Symmetrie sowohl a als auch b in $[c]$ enthalten. Also ist $[a] \subset [c] \subset [a]$ und $[b] \subset [c] \subset [b]$. Damit gilt $[a] = [c] = [b]$. Also sind zwei Äquivalenzklassen entweder disjunkt oder gleich. Wegen der Reflexivität ist jedes Element in einer Äquivalenzklasse enthalten. Also zerfällt A in eine disjunkte Vereinigung von Äquivalenzklassen, d.h. jedes Element von A gehört zu genau einer Äquivalenzklasse.

1.4 Abbildungen

Eine Abbildung $f : X \to Y$ ist eine Zuordnungsvorschrift, die jedem Element x von X genau ein Element $f(x)$ aus Y zuordnet:

$$f : X \to Y, \quad x \mapsto f(x)$$

Die Menge X der Argumente wird Definitionsbereich genannt und die Menge Y, in denen die Abbilde liegen, Wertebereich. Das Bild ist die Teilmenge aller Elemente y des Wertebereichs Y, die Abbild eines Arguments in X sind:

$$\text{Bild } f[X] = \{y \in Y \mid \exists x \in X \text{ mit } f(x) = y\}. \quad \exists \text{ steht für “es gibt (mindestens) ein”}$$

Definition 1.2. *Eine Abbildung* $f : X \to Y, \quad x \mapsto f(x)$ *heißt*

(i) *injektiv, wenn je zwei verschiedene Elemente* $x, x' \in X$ *auch auf verschiedene Elemente von* Y *abgebildet werden:*

$$\forall x, x' \in X \text{ folgt aus } x \neq x' \text{ auch } f(x) \neq f(x'). \quad \forall \text{ steht für "für alle"}$$

(ii) *surjektiv, wenn das Bild von* f *der ganze Wertebereich* Y *ist.*

$$\forall y \in Y \quad \exists x \in X \text{ mit } f(x) = y.$$

(iii) *bijektiv, wenn sie injektiv und surjektiv ist.*

Für jede Teilmenge $B \subset Y$ ist das Urbild von B unter f die Menge aller Elemente von X, die nach B abgebildet werden:

$$f^{-1}[B] = \{x \in X \mid f(x) \in B\}.$$

Für eine bijektive Abbildung $f : X \to Y, \quad x \mapsto f(x)$ besteht für jedes $y \in Y$ das Urbild $f^{-1}[\{y\}]$ nur aus genau einem Element. Also existiert auch die Umkehrabbildung

$$f^{-1} : Y \to X, \quad y \mapsto f^{-1}(y) \text{ mit } f^{-1}[\{y\}] = \{f^{-1}(y)\}.$$

Offenbar gilt dann $f^{-1}(f(x)) = x$ für alle $x \in X$ und $f(f^{-1}(y)) = y$ für alle $y \in Y$.

1.5 Komposition von Abbildungen

Seien $f : X \to Y, \quad x \mapsto f(x)$ und $g : Y \to Z, \quad y \mapsto g(y)$ Abbildungen, dann definiert $g \circ f : X \to Z, \quad x \mapsto g(f(x))$ die sogenannte Komposition (Verkettung) von f und g.

Satz 1.3. *Die Komposition von Abbildungen ist assoziativ, d.h. für Abbildungen* $f : X \to Y, \quad x \mapsto f(x), g : Y \to Z, \quad y \mapsto g(y)$ *und* $h : Z \to V, \quad z \mapsto h(z)$ *gilt*

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Sei $f : X \to Y, \quad x \mapsto f(x)$ *eine Abbildung und seien* $\mathbb{1}_X : X \to X, \quad x \mapsto x$ *und* $\mathbb{1}_Y : Y \to Y, y \mapsto y$ *die identischen Abbildungen von den Mengen* X *und* Y*. Dann gilt*

$$f \circ \mathbb{1}_X = f = \mathbb{1}_Y \circ f = f$$

Ist die Abbildung f *bijektiv, dann gilt auch* $\qquad f \circ f^{-1} = \mathbb{1}_Y$ *und* $f^{-1} \circ f = \mathbb{1}_X$.

Beweis:

$$\begin{aligned} (h \circ (g \circ f))(x) &= h(g(f(x))) = ((h \circ g) \circ f)(x) && \forall x \in X \\ f \circ \mathbb{1}_X(x) &= f(x) = (\mathbb{1}_Y \circ f)(x) && \forall x \in X \\ f \circ f^{-1}(y) &= f(f^{-1}(y)) = y = \mathbb{1}_Y(y) && \forall y \in Y \\ f^{-1} \circ f(x) &= f^{-1}(f(x)) = x = \mathbb{1}_X(x) && \forall x \in X. \end{aligned}$$

q.e.d.

Kapitel 2

Reelle Zahlen

2.1 Axiome der reellen Zahlen

Die Menge der reellen Zahlen $\mathbb{R}$ wird durch folgende Axiome charakterisiert:

A1. Axiome der Addition
A2. Axiome der Multiplikation
A3. Distributivgesetz
A4. Ordnungsaxiome
A5. Vollständigkeit

A1. Axiome der Addition 2.1. *Es gibt eine Operation*
$+: \quad \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (x, y) \mapsto x + y$ *mit*

(i) *Kommutativgesetz:* $x + y = y + x$ *für alle* $x, y \in \mathbb{R}$.

(ii) *Assoziativgesetz:* $x + (y + z) = (x + y) + z$ *für alle* $x, y, z \in \mathbb{R}$.

(iii) *Existenz der Null: es gibt eine Zahl* $0 \in \mathbb{R}$ *mit* $x + 0 = x$ *für alle* $x \in \mathbb{R}$

(iv) *Existenz des Negativen: zu jedem* $x \in \mathbb{R}$ *gibt es ein* $-x \in \mathbb{R}$ *mit* $x + (-x) = 0$.

A2. Axiome der Multiplikation 2.2. *Es gibt eine Operation*
$\cdot: \quad \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (x, y) \mapsto x \cdot y$ *mit*

(i) *Kommutativgesetz:* $x \cdot y = y \cdot x$ *für alle* $x, y \in \mathbb{R}$.

(ii) *Assoziativgesetz:* $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ *für alle* $x, y, z \in \mathbb{R}$.

(iii) *Existenz der Eins: es gibt eine Zahl* $1 \in \mathbb{R}, 1 \neq 0$ *mit* $x \cdot 1 = x$ *für alle* $x \in \mathbb{R}$

(iv) *Existenz des Inversen: zu jedem* $x \in \mathbb{R} \setminus \{0\}$ *gibt es ein* $x^{-1} \in \mathbb{R}$ *mit* $x \cdot x^{-1} = 1$.

A3. Distributivgesetz 2.3.

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z) = x \cdot y + x \cdot z \text{ für alle } x, y, z \in \mathbb{R}.$$

Definition 2.4. *Allgemein heißt eine Menge $\mathbb{K}$, die die Axiome* **A1-A3** *erfüllt Körper. Für Körper gelten daher auch alle Folgerungen aus* **A1-A3***. Es gibt viele Körper.*

Beispiel 2.5. *Der kleinste Körper besteht aus zwei Elementen $\mathbb{Z}_2 = \{0, 1\}$. Die Operationen $+$ und $\cdot$ sind dann definiert durch:*

$$\begin{array}{llll} 0+0=0 & 0+1=1 & 0\cdot 0=0 & 0\cdot 1=0 \\ 1+0=1 & 1+1=0 & 1\cdot 0=0 & 1\cdot 1=1 \end{array}$$

Zeige dass diese Definitionen von $+$ und $\cdot$ auf $\mathbb{Z}_2$ die Axiome **A1-A3** erfüllen und umgekehrt durch **A1-A3** eindeutig bestimmt sind. In $\mathbb{R}$ soll aber $1 + 1 = 2 \neq 0$ gelten, so dass wir noch weitere Axiome benötigen um den Körper der reellen Zahlen zu charakterisieren.

Wir benützen folgende Abkürzungen:

$$\begin{array}{llll} x - y = x + (-y) & xy = x \cdot y & -xy = -(x \cdot y) & \forall x, y \in \mathbb{R} \\ \dfrac{1}{x} = x^{-1} & \dfrac{y}{x} = y \cdot x^{-1} & & \forall x \in \mathbb{R} \setminus \{0\}, y \in \mathbb{R} \\ x + y + z = x + (y + z) & xyz = x \cdot (y \cdot z) & xy + z = (x \cdot y) + z & \forall x, y, z \in \mathbb{R} \end{array}$$

Satz 2.6. *(Folgerungen aus* **A1***)*

(i) *Falls $x + y = x + z$, dann $y = z$*

(ii) *Falls $x + y = x$, dann $y = 0$*

(iii) *Falls $x + y = 0$, dann $y = -x$*

(iv) $-(-x) = x$

Bemerkung 2.7. *(i) heißt Kürzungsregel. (ii) zeigt, dass die Null eindeutig durch die Eigenschaft $x + 0 = x$ bestimmt ist. (iii) zeigt, dass das Negative $(-x)$ eindeutig durch $x + (-x) = 0$ bestimmt ist.*

Beweis: **(i)** Sei $x + y = x + z$, dann folgern wir

$$\begin{array}{llll} y = y + 0 & = 0 + y & = (x - x) + y & \\ \quad = (-x + x) + y) & = -x + (x + y) & = -x + (x + z) & = (-x + x) + z \\ \quad = (x - x) + z & = 0 + z & = z + 0 & = z. \end{array}$$

(ii) Sei $x + y = x$, dann gilt $x + y = x + 0$. Also folgt aus (i) $y = 0$.
(iii) Sei $x + y = 0$, dann gilt $x + y = x + (-x)$. Also folgt aus (i) $y = -x$.
(iv) $-x + x = x + (-x) = 0$. Also folgt aus (iii) $x = -(-x)$. **q.e.d.**

Satz 2.8. *(Folgerungen aus* **A2***)*

(i) *Falls* $x \neq 0$ *und* $xy = xz$, *dann* $y = z$

(ii) *Falls* $x \neq 0$ *und* $xy = x$, *dann* $y = 1$

(iii) *Falls* $x \neq 0$ *und* $x \cdot y = 1$, *dann* $y = x^{-1}$

(iv) *Falls* $x \neq 0$, *dann* $(x^{-1})^{-1} = x$

Bemerkung 2.9. *Die Folgerungen sind analog zu denen aus* **A1**. *Wieder heißt (i) Kürzungsregel, (ii) impliziert wieder die Eindeutigkeit der Eins und (iii) die Eindeutigkeit des Inversen.*

Beweis: (i) Sei $x \neq 0$ und $xy = xz$, dann folgern wir

$$y = 1y = \left(x\frac{1}{x}\right) y = \left(\frac{1}{x}x\right) y = \frac{1}{x}(xy) = \frac{1}{x}(xz) = \left(\frac{1}{x}x\right) z = \left(x\frac{1}{x}\right) z = 1z = z.$$

(ii) Sei $x \neq 0$ und $xy = x$, dann gilt $xy = x \cdot 1$ Also folgt aus (i) $y = 1$.
(iii) Sei $x \neq 0$ und $xy = 1$, dann gilt $xy = x \cdot x^{-1}$. Also folgt aus (i) $y = x^{-1}$.
(iv) Sei $x \neq 0$. Dann ist $x^{-1}x = xx^{-1} = 1$. Also folgt aus (iii) $x = (x^{-1})^{-1}$. **q.e.d.**

Satz 2.10. *(Folgerungen aus* **A1**-**A3***)*

(i) $x \cdot 0 = 0$ *für alle* x

(ii) $x \cdot y = 0 \Longleftrightarrow x = 0$ *oder* $y = 0$

(iii) $(-x)y = -xy = x(-y)$.

(iv) $(-1) \cdot x = -x$

(v) $(-x)(-y) = xy$

(vi) $x \neq 0, y \neq 0$ *dann* $(xy)^{-1} = y^{-1}x^{-1}$.

Bemerkung 2.11. *Die Null hat kein multiplikatives Inverses, sonst wäre ja* $0 \cdot 0^{-1} = 1$, *was (i) widerspricht.*

Beweis: (i) $x \cdot 0 + x \cdot 0 = x \cdot (0 + 0) = x \cdot 0$. Aus (ii) von Satz 2.6 folgt dann $x \cdot 0 = 0$
(ii) Sei $x \cdot y = 0$ und $x \neq 0$. Dann gilt wegen (i)

$$y = y \cdot 1 = 1 \cdot y = \left(\frac{1}{x} \cdot x\right) \cdot y = \frac{1}{x} \cdot (x \cdot y) = \frac{1}{x} \cdot 0 = 0.$$

Wegen der Kommutativität folgt aus (i) auch die Umkehrung $0 \cdot y = x \cdot 0 = 0$.

(iii) $0 = 0 \cdot y = (x + (-x))y = xy + (-x)y$. Aus (iii) im Satz 2.6 folgt dann

$$(-x)y = -xy = -yx = (-y)x = x(-y).$$

(iv) Setze in (iii) $y = 1$.
(v) Wegen (iii) gilt $(-x)(-y) = -(x \cdot (-y)) = -(-xy) = xy$.
(vi) $1 = xy(xy)^{-1} = (xy)^{-1}xy = ((xy)^{-1}x)y = y((xy)^{-1}x)$.
$\Longrightarrow y^{-1} = (xy)^{-1}x. \Longrightarrow y^{-1}x^{-1} = (xy)^{-1}xx^{-1} = (xy)^{-1}$. **q.e.d.**

A4. Ordnungsaxiome 2.12. *Es gibt eine Relation $<$ in $\mathbb{R}$ mit drei Eigenschaften:*

(i) *Totalität der Ordnung: Für je zwei reelle Zahlen $x, y \in \mathbb{R}$ gilt genau eine der drei folgenden Relationen $x < y$ oder $x = y$ oder $y < x$.*

(ii) *Transitivität: $x < y$ und $y < z \Longrightarrow x < z$*

(iii) *Monotonie:* $x < y \Longrightarrow \begin{cases} x + c < y + c & \text{für alle } c \in \mathbb{R} \\ x \cdot c < y \cdot c & \text{für alle } 0 < c \in \mathbb{R} \end{cases}$

Wir benutzten folgende Abkürzungen:

$$\begin{aligned} x > y &\Longleftrightarrow y < x \\ x \leq y &\Longleftrightarrow (x < y \text{ oder } x = y) \Longleftrightarrow y \geq x \\ \mathbb{R}^+ &= \{x \in \mathbb{R} \mid 0 < x\} \text{ positive Zahlen} \\ \mathbb{R}^- &= \{x \in \mathbb{R} \mid x < 0\} \text{ negative Zahlen} \\ \mathbb{R}_0^+ &= \{x \in \mathbb{R} \mid 0 \leq x\} \text{ nichtnegative Zahlen} \\ \mathbb{R}_0^- &= \{x \in \mathbb{R} \mid x \leq 0\} \text{ nichtpositive Zahlen} \end{aligned}$$

Satz 2.13. *(Folgerungen aus* **A4***)*

(i) $0 < x \Longrightarrow -x < 0$ *und* $x < 0 \Longrightarrow -x > 0$

(ii) $x < y \Longleftrightarrow 0 < y - x$

(iii) $x < y$ *und* $a < 0 \Longrightarrow ay < ax$

(iv) $x \neq 0 \Longrightarrow x \cdot x = x^2 > 0$

(v) $x > 0 \Longrightarrow \frac{1}{x} > 0$

(vi) $0 < x < y \Longrightarrow 0 < \frac{1}{y} < \frac{1}{x}$

Bemerkung 2.14. *Da $1 \cdot 1 = 1$ folgt aus (iv) $1 > 0$. Dann folgt aus (i) $-1 < 0$. Also gilt $-1 < x^2$ für jedes $x \in \mathbb{R}$. Also gibt es keine reelle Zahl x mit $x^2 = -1$.*

Beweis: **(i)** Sei $0 < x$. Dann folgt mit Monotonie $0 + (-x) < x + (-x)$, also $-x < 0$. Sei $x < 0$. Dann folgt mit Monotonie $x + (-x) < -x$ also auch $0 < -x$.
(ii) Sei $x < y$. Dann folgt mit Monotonie $x - x < y - x$, also auch $0 < y - x$.
Sei umgekehrt $0 < y - x$. Dann folgt mit Monotonie $x < y$.
(iii) Sei $x < y$ und $a < 0$. Dann folgt aus (i) $-a > 0$. Also gilt wegen Monotonie $-ax < -ay \iff 0 < ax - ay \iff ay < ax$.
(iv) Sei $x > 0$. Dann folgt wegen Monotonie $x^2 > 0 \cdot x = 0$.
Sei $x < 0$. Dann folgt aus (i) $-x > 0$ und mit Monotonie $x^2 = (-x)\cdot(-x) > 0\cdot(-x) = 0$.
(v) Sei $x > 0$. Dann ist $x \neq 0$. Aus (iv) folgt $\frac{1}{x} \cdot \frac{1}{x} > 0$. Mit Monotonie folgt dann $\frac{1}{x} = x \cdot \frac{1}{x} \cdot \frac{1}{x} > 0$.
(vi) Sei $0 < x < y$. Dann ist $x > 0$ und $y > 0$ also wegen (v) auch $\frac{1}{x} > 0$ und $\frac{1}{y} > 0$. Dann folgt mit Monotonie $\frac{1}{y} = \frac{1}{y} \cdot \frac{1}{x} \cdot x < \frac{1}{y} \cdot \frac{1}{x} \cdot y = \frac{1}{x} \cdot \frac{1}{y} \cdot y = \frac{1}{x}$. **q.e.d.**

Satz 2.15. *(Arithmetisches Mittel) Seien $x, y \in \mathbb{R}$, dann gilt*

$$x < y \Longrightarrow x < \frac{x+y}{2} < y$$

Also liegt zwischen zwei verschiedenen reellen Zahlen immer eine weitere.

Beweis: Aus $x < y$ folgt mit Monotonie $x + x < x + y < y + y$. Weil aber $x + x = x(1 + 1) = 2x$ und $2 = 1 + 1 > 0$ folgt dann $2x < x + y < 2y$ und $x < \frac{x+y}{2} < y$ **q.e.d.**

Übungsaufgabe 2.16. *Es gelten auch folgende Regeln:*

(i) $a < b$ *und* $c < d \Longrightarrow a + c < b + d$

(ii) $0 < a < b$ *und* $0 < c < d \Longrightarrow ac < bd$

(iii) $ab > 0 \iff$ *entweder* $a > 0, b > 0$ *oder* $a < 0, b < 0$

(iv) $ab < 0 \iff$ *entweder* $a > 0, b < 0$ *oder* $a < 0, b > 0$

Definition 2.17. *(Betrag)*
Der Betrag einer reellen Zahl $x \in \mathbb{R}$ ist die nicht negative Zahl

$$|x| = \begin{cases} x & \text{falls } x \geq 0 \quad \iff \quad x \geq -x \\ -x & \text{falls } x < 0 \quad \iff \quad x < -x \end{cases} = \max\{x, -x\}.$$

$$\max\{x, y\} = \begin{cases} x & \text{falls } y \leq x \\ y & \text{falls } y > x \end{cases} \qquad \forall x, y \in \mathbb{R}.$$

Aus der Definition folgt

$$\begin{array}{ccc} |x| \geq 0 & \text{und} & |x| = 0 \iff x = 0. \\ -|x| \leq x \leq |x| & \iff & x \leq |x| \text{ und } -x \leq |x| \\ |-x| = |x| & \text{denn} & |-x| = \max\{-x, x\}. \end{array}$$

Satz 2.18. *(Eigenschaften des Betrags) Für alle $x, y \in \mathbb{R}$ gilt:*

(i) $|x| \geq 0$ *und* $|x| = 0 \Longleftrightarrow x = 0$

(ii) $|x \cdot y| = |x| \cdot |y|$

(iii) $|x + y| \leq |x| + |y|$

Beweis: **(i)** haben wir schon gesehen.
(ii) Wegen Satz 2.10 (iii) und (v) ändern sich beide Seiten nicht wenn wir x durch $-x$ bzw. y duch $-y$ ersetzen. Für $x \geq 0$ und $y \geq 0$ ist die Aussage klar.
(iii) Zwei Fälle: $x + y > 0 \Longrightarrow |x + y| = x + y \leq |x| + y \leq |x| + |y|$ wegen Monotonie.
$x + y < 0 \Longrightarrow |x + y| = -x - y \leq |x| - y \leq |x| + |y|$ wegen Monotonie. **q.e.d.**

Folgerung 2.19. $$||x| - |y|| \leq |x - y|$$

Beweis: $|x| = |x - y + y| \leq |x - y| + |y| \Longrightarrow |x| - |y| \leq |x - y|$.
Vertausche x und $y \Longrightarrow |y| - |x| \leq |y - x| = |x - y|$. Also gilt $||x| - |y|| \leq |x - y|$.**q.e.d.**

Definition 2.20. *(Abstand)*
Der Abstand $d(x, y)$ zweier Zahlen $x, y \in \mathbb{R}$ ist die nicht negative Zahl $d(x, y) = |x - y|$.

Satz 2.21. *(Eigenschaften des Abstands)*
Der Abstand $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (x, y) \mapsto d(x, y)$ hat folgende Eigenschaften:

(i) $d(x, y) \geq 0$ *und* $d(x, y) = 0 \Longleftrightarrow x = y$

(ii) $d(x, y) = d(y, x)$

(iii) $d(x, y) \leq d(x, z) + d(z, y)$ *für alle* $x, y, z \in \mathbb{R}$.

Beweis: folgt aus den Folgerungen der Definition und Satz 2.18.
(iii) $|x - y| = |x - z + z - y| \leq |x - z| + |z - y|$. **q.e.d.**

Definition 2.22. *Sei $M \subset \mathbb{R}$ eine nicht leere Teilmenge von Zahlen.*

(i) *M heißt nach oben beschränkt, falls es eine Zahl $\beta \in \mathbb{R}$ gibt, so dass $x \leq \beta$ für alle $x \in M$ gilt. β heißt dann obere Schranke von M.*

(ii) *M heißt nach unten beschränkt, falls es eine Zahl $\alpha \in \mathbb{R}$ gibt, so dass $\alpha \leq x$ für alle $x \in M$ gilt. α heißt dann untere Schranke .*

(iii) *M heißt beschränkt, wenn M nach oben und unten beschränkt ist.*

Definition 2.23. *(Supremum einer Menge) Sei $M \subset \mathbb{R}$ eine nicht leere nach oben beschränkte Menge. Eine reelle Zahl s heißt die kleinste obere Schranke von M oder das Supremum, wenn sie eine obere Schranke von M ist, und es keine obere Schranke von M gibt, die kleiner ist als s. Wir schreiben dann $s = \sup M$. Wenn das Supremum einer Menge existiert ist es eindeutig, weil jedes Supremum weder kleiner noch größer als ein anderes Supremum von M ist.*

Definition 2.24. *(Infimum einer Menge) Sei $M \subset \mathbb{R}$ eine nicht leere nach unten beschränkte Menge. Eine reelle Zahl t heißt grösste untere Schranke von M oder Infimum von M, wenn es eine untere Schranke von M ist, und es keine untere Schranke von M gibt, die größer ist als t. Wir schreiben dann $t = \inf M$. Auch das Infimum ist eindeutig, wenn es existiert.*

Bemerkung 2.25. *Folgende Charakterisierung erleichtert manche Beweise:*

$$\begin{array}{lcll} s = \sup M & \iff & s \text{ ist obere Schranke und} & \forall \epsilon > 0 \quad \exists x \in M \text{ mit } s - \epsilon < x. \\ t = \inf M & \iff & t \text{ ist untere Schranke und} & \forall \epsilon > 0 \quad \exists x \in M \text{ mit } x < t + \epsilon. \end{array}$$

Warnung 2.26. *Das Supremum* $\sup M$ *bzw. Infimum* $\inf M$ *braucht nicht zu der Menge M zu gehören.*

Definition 2.27. *(Maximum und Minimum) Sei $M \subset \mathbb{R}$ eine nicht leere nach oben beschränkte Menge. Ein Element $m \in M$ von M, das eine obere Schranke von M ist heißt Maximum. Wir schreiben dann $m = \max M$. Analog heißt ein Element m einer nicht leeren nach unten beschränkten Menge M, das eine untere Schranke von M ist Minimum. Wir schreiben dann $m = \min M$.*

A5. Vollständigkeitsaxiom 2.28. *Für jede nicht leere nach oben beschränkte Menge M existiert das Supremum $s = \sup M \in \mathbb{R}$.*

Wir werden sehen, dass die rationalen Zahlen $\mathbb{Q}$ **A1**-**A4** erfüllen aber nicht **A5**. Wir führen jetzt die Intervalle ein, das sind folgende Teilmenge der reellen Zahlen:

$$\begin{array}{rlr} [a,b] &= \{x \in \mathbb{R} \mid a \le x \le b\} & \text{für alle } a \le b \in \mathbb{R} \\ [a,b) &= \{x \in \mathbb{R} \mid a \le x < b\} & \text{für alle } a < b \in \mathbb{R} \\ (a,b] &= \{x \in \mathbb{R} \mid a < x \le b\} & \text{für alle } a < b \in \mathbb{R} \\ (a,b) &= \{x \in \mathbb{R} \mid a < x < b\} & \text{für alle } a < b \in \mathbb{R} \\ (-\infty,b] &= \{x \in \mathbb{R} \mid x \le b\} & \text{für alle } b \in \mathbb{R} \\ (-\infty,b) &= \{x \in \mathbb{R} \mid x < b\} & \text{für alle } b \in \mathbb{R} \\ [a,\infty) &= \{x \in \mathbb{R} \mid a \le x\} & \text{für alle } a \in \mathbb{R} \\ (a,\infty) &= \{x \in \mathbb{R} \mid a < x\} & \text{für alle } a \in \mathbb{R} \end{array}$$

Offenbar ist b eine obere Schranke von (a,b). Andererseits gibt es wegen Satz 2.15 für jedes $x<b$ ein $y=\max\{\frac{x+b}{2},\frac{a+b}{2}\}$ mit $x<y$ und $y\in(a,b)$. Also gibt es keine obere Schranke von (a,b), die kleiner ist als b. Damit ist $b=\sup(a,b)$. Analog gilt:

$$\begin{array}{rl} \inf[a,b]=a & \sup[a,b]=b \\ \inf[a,b)=a & \sup[a,b)=b \\ \inf(a,b]=a & \sup(a,b]=b \\ \inf(a,b)=a & \sup(a,b)=b \end{array}$$

$$\begin{array}{rr} (-\infty,b] \text{ nicht nach unten beschränkt} & \sup(-\infty,b]=b \\ (-\infty,b) \text{ nicht nach unten beschränkt} & \sup(-\infty,b)=b \end{array}$$

$$\begin{array}{rl} \inf[a,\infty)=a & [a,\infty) \text{ nicht nach oben beschränkt} \\ \inf(a,\infty)=a & (a,\infty) \text{ nicht nach oben beschränkt} \end{array}$$

Satz 2.29. (i) *Für jede nicht leere nach unten beschränkte Menge M existiert das Infimum* $\inf M\in\mathbb{R}$.

(ii) *Eine nicht leere nach oben beschränkte Menge M besitzt genau dann ein Maximum, wenn* $\sup M\in M$. *In diesem Fall ist* $\max M=\sup M$.

(iii) *Eine nicht leere nach unten beschränkte Menge M besitzt genau dann ein Minimum, wenn* $\inf M\in M$. *In diesem Fall ist* $\min M=\inf M$.

Beweis: (i) Weil die Ordnungsrelation nicht symmetrisch ist, erhalten wir dadurch, dass wir in einer Aussage alle auftauchenden Ordnungsrelationen umkehren, eine andere Aussage. Zum Beispiel erhalten wir die Definition der unteren Schranke, indem wir in der Definition der oberen Schranke alle Ordnungsrelationen umdrehen. Dann erhalten wir aber auch die Definition des Infimums, indem wir in der Definition des Supremums alle Ordnungsrelationen umkehren. Weil $x<y\iff -y<-x$, sind also Aussagen über Ordnungsrelationen zwischen reellen Zahlen äquivalent zu den analogen Aussagen, in denen wir alle Ordnungsrelationen umdrehen und alle reellen Zahlen durch ihre Negativen ersetzen[1]. Insbesondere besitzt die Menge M genau dann ein Infimum, wenn die Menge $-M=\{x\in\mathbb{R}\mid -x\in M\}$ ein Supremum besitzt und es gilt $\inf M=-\sup -M$. Also folgt (i) aus dem Vollständigkeitsaxiom **A5**.

[1] Wenn diese Aussagen aber algebraische Operationen benutzen, die nicht verträglich sind mit der Abbildung jeder rellen Zahl auf ihre Negative, wie z. B. das Produkt zweier reeller Zahlen, dann müssen bei dieser Ersetzung auch die algebraischen Operationen entsprechend geändert werden, also z. B. das Produkt in das negative des Produktes.

(ii) Wenn $\sup M \in M$, dann ist $\sup M$ eine obere Schranke von M, die Element von M ist. Wenn $\sup M \notin M$, dann gilt für alle $x \in M$ sogar $x < \sup M$. Dann gilt aber für alle oberen Schranken s von M.

$$x < \sup M \le s \text{ für alle } x \in M.$$

Also gibt es in M keine obere Schranke von M.
(iii) analog zu (ii). **q.e.d.**

$$\begin{array}{rr} \text{Also existiert} & \max[a,b] = \max(a,b] = \max(-\infty,b] = b \\ & \min[a,b] = \min[a,b) = \min[a,\infty) = a, \\ \text{während} & [a,b) \text{ und } (a,b) \text{ und } (-\infty,b) \text{ kein Maximum besitzen} \\ \text{und} & (a,b] \text{ und } (a,b) \text{ und } (a,\infty) \text{ kein Minimum.} \end{array}$$

Wenn wir die reellen Zahlen durch $-\infty$ und ∞ erweitern, können wir auch für unbeschränkte Mengen obere und untere Schranken und Suprema und Infima definieren.

2.2 Die erweiterte Zahlengerade $\bar{\mathbb{R}}$

Definition 2.30. *Die erweiterte Zahlengerade besteht aus* $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$.

Mit ∞ bezeichnen wir auch $+\infty$. Die Ordnungsrelation läßt sich auf $\bar{\mathbb{R}}$ durch $-\infty < x < \infty$ für alle $x \in \mathbb{R}$ fortsetzen und erfüllt offensichtlich auch (i) und (ii) des Ordnungsaxioms. Die Operationen $+$ und $\cdot$ lassen sich teilweise auf $\bar{\mathbb{R}}$ fortsetzen:

$$x + \infty = \infty, \qquad x - \infty = -\infty \quad \text{für alle } x \in \mathbb{R}.$$

$$x \cdot \infty = \begin{cases} \infty & \text{für } x > 0 \\ -\infty & \text{für } x < 0, \end{cases} \qquad x \cdot (-\infty) = \begin{cases} -\infty & \text{für } x > 0 \\ \infty & \text{für } x < 0 \end{cases}$$

$$\frac{x}{\infty} = \frac{x}{-\infty} = 0, \qquad \frac{\infty}{x} = \begin{cases} \infty & \text{für } x > 0 \\ -\infty & \text{für } x < 0, \end{cases} \qquad \frac{-\infty}{x} = \begin{cases} -\infty & \text{für } x > 0 \\ \infty & \text{für } x < 0. \end{cases}$$

$$\begin{array}{rlll} \text{Außerdem gilt} & \infty + \infty = \infty & -\infty - \infty = -\infty & \\ & \infty \cdot \infty = \infty & \infty(-\infty) = -\infty & (-\infty)(-\infty) = \infty. \\ \text{Nicht definiert sind} & \infty - \infty, \ \dfrac{\pm\infty}{\pm\infty}, & \dfrac{\pm\infty}{0}, \ \dfrac{0}{0}, & 0 \cdot (\pm\infty). \end{array}$$

$\bar{\mathbb{R}}$ ist also kein Körper. Die Definitionen von oberen und unteren Schranken, von Supremum und Infimum, und von Maximum und Infimum übertragen sich auf die erweiterte

Zahlengerade. Offenbar ist ∞ das Maximum und $-\infty$ das Minimum von $\bar{\mathbb{R}}$. Insbesondere ist ∞ eine obere Schranke und $-\infty$ eine untere Schranke von jeder Teilmenge von $\bar{\mathbb{R}}$. Weil dann aber ∞ die einzige obere Schranke einer Teilmenge M von $\mathbb{R} \subset \bar{\mathbb{R}}$ ist, die in $\mathbb{R}$ keine obere Schranke hat, ist ∞ dann auch das Supremum von M als Teilmenge von $\bar{\mathbb{R}}$. Analog ist $-\infty$ das Infimum einer Teilmenge von $\mathbb{R} \subset \bar{\mathbb{R}}$, die keine untere Schranke in $\mathbb{R}$ hat. Daraus folgt, dass jede Teilmenge von $\bar{\mathbb{R}}$ ein Supremum und ein Infimum hat. Außerdem gilt wieder, dass eine Teilmenge $M \subset \bar{\mathbb{R}}$ genau dann ein Maximum bzw. Infimum hat, wenn $\sup M \in M$ bzw. $\inf M \in M$ gilt. In diesen Fällen ist wieder $\max M = \sup M$ bzw. $\min M = \inf M$. Wir benutzen die Symbole sup, inf, max und min sowohl für Teilmengen von $\mathbb{R}$ als auch für Teilmengen von $\bar{\mathbb{R}}$, so dass aus dem Zusammenhang klar werden muss, was genau gemeint ist.

2.3 Die natürlichen Zahlen $\mathbb{N} \subset \mathbb{R}$

Wir wollen die natürlichen Zahlen als Teilmenge der reellen Zahlen charakterisieren. Aufgrund von **A2** gibt es das Element $1 \in \mathbb{N} \subset \mathbb{R}$, das wegen (ii) in Satz 2.8 eindeutig ist. Aus (iv) im Satz 2.13 folgt $1 > 0$ aus $1 = 1 \cdot 1$. Wegen der Monotonie ist dann die Nachfolgerabbildung monoton:

$$\mathbb{N} \to \mathbb{N}, \qquad n \mapsto n+1 > n$$
$$1 \mapsto 1+1 = 2 > 1, \qquad 2 \mapsto 2+1 = 3 > 2, \qquad \ldots \qquad n \mapsto n+1 > n, \qquad \ldots$$

Definition 2.31. *Eine Menge $M \subset \mathbb{R}$ heißt induktiv, falls*

(i) $1 \in M$ *(Einselement).*

(ii) $a \in M \Longrightarrow a+1 \in M$ *(invariant unter der Nachfolgerabbildung).*

$\mathbb{R}$ selber ist offenbar induktiv oder auch $[1, \infty)$. Offenbar ist der Durchschnitt zweier induktiver Mengen wieder induktiv.

Definition 2.32. *(Natürliche Zahlen) $\mathbb{N}$ ist die kleinste induktive Teilmenge von $\mathbb{R}$, also der Durchschnitt aller induktiven Teilmengen von $\mathbb{R}$.*

Satz 2.33. *Es gilt $\mathbb{N} = \{1\} \cup \{n \in \mathbb{R} \mid n-1 \in \mathbb{N}\} = \{1\} \cup$Bild der Nachfolgerabbildung.*

Beweis: Wegen $(1+1)-1 = 1$, $(n+1)-1 = (n-1)+1$ und weil $\mathbb{N}$ eine induktive Menge ist, ist auch $S = \{1\} \cup \{n \in \mathbb{R} \mid n-1 \in \mathbb{N}\}$ eine induktive Menge. Also folgt $S \supset \mathbb{N}$. Weil $\mathbb{N}$ eine induktive Menge ist, folgt $S \subset \mathbb{N}$ aus $n = (n-1)+1$. **q.e.d.**

Satz 2.34. *(Prinzip der vollständigen Induktion) Eine Teilmenge $S \subset \mathbb{N}$ erfülle*
(i) $1 \in S$ **(ii)** $a \in S \Longrightarrow a+1 \in S$. *Dann ist $S = \mathbb{N}$.*

Beweis: S ist offenbar eine induktive Menge. Also gilt $\mathbb{N} \subset S$. Andererseits ist S eine Teilmenge von $\mathbb{N}$. Also folgt $\mathbb{N} = S$. **q.e.d.**

Beweis durch vollständige Induktion: Um eine Aussage $A(n)$ über alle natürlichen Zahlen $n \in \mathbb{N}$ zu beweisen genügt es also zu zeigen:

(i) die Aussage $A(1)$ ist richtig.

(ii) Falls die Aussage $A(n)$ richtig ist, dann auch $A(n + 1)$.

Ein solcher Beweis heißt Beweis durch vollständige Induktion.

Satz 2.35. *(Bernoulli–Ungleichung) Es gilt*

$$(1 + x)^n \geq 1 + nx \quad \textit{für alle } x \in (-1, \infty) \textit{ und alle } n \in \mathbb{N}.$$

Gleichheit gilt nur für $x = 0$ oder $n = 1$.

Beweis durch vollständige Induktion: Für $x = 0$ oder $n = 1$ gilt offenbar die Gleichheit. Für jede natürliche Zahl $n \in \mathbb{N}$ sei $A(n)$ die Aussage

$$\text{Für alle } x \in (-1, 0) \cup (0, \infty) \quad \text{gilt} \quad (1 + x)^{n+1} > 1 + (n + 1)x.$$

(i) Wenn $x \neq 0$ dann folgt $(1 + x)^2 = 1 + 2x + x^2 > 1 + 2x$. Also gilt $A(1)$.
(ii) Es gelte $A(n)$. Dann folgern wir für $x \in (-1, 0) \cup (0, \infty)$ aufgrund der Monotonie:

$$\text{Wegen } x > -1 \text{ gilt } 1 + x > 0, \quad \text{und wegen } x \neq 0 \text{ folgt daraus}$$

$$(1 + x)(1 + x)^{n+1} > (1 + x)(1 + (n + 1)x) \geq 1 + (n + 2)x + (n + 1)x^2 > 1 + (n + 2)x.$$

Also gilt $A(n + 1)$. **q.e.d.**

Satz 2.36. *Für alle $n \in \mathbb{N}$ gibt es keine Zahl in $\mathbb{N}$ zwischen $n - 1$ und n.*

Beweis durch vollständige Induktion:
(i) $[1, \infty)$ ist eine induktive Menge. Also ist $\mathbb{N} \cap (0, 1) = \mathbb{N} \cap [0, \infty) \cap (0, 1) = \emptyset$.
(ii) Wir nehmen an, dass es für ein $n \in \mathbb{N}$ keine natürliche Zahl in $(n - 1, n)$ gibt. Wenn es eine Zahl $m \in \mathbb{N} \cap (n, n + 1)$ gibt, dann liegt m wegen Satz 2.33 auch in $\{1\} \cup \{x \in \mathbb{R} \mid x - 1 \in \mathbb{N}\}$. Wegen $1 \leq n < m$ ist m nicht gleich 1. Also liegt m dann in $\{n \in \mathbb{R} \mid n - 1 \in \mathbb{N}\}$. Wegen der Monotonie liegt $m - 1$ dann in $\mathbb{N}$ und $(n - 1, n)$, was der Annahme wiederspricht. Also gibt es keine natürliche Zahl in $\mathbb{N} \cap (n, n + 1)$.**q.e.d.**

Für alle $n \in \mathbb{N}$ ist also $\{m \in \mathbb{N} \mid m \leq n + 1\} = \{m \in \mathbb{N} \mid m \leq n\} \cup \{m + 1\}$. Für alle $n \in \mathbb{N}$ schreiben wir deshalb $\{1, \ldots, n\} = \{m \in \mathbb{N} \mid m \leq n\}$.

Satz 2.37. *(Wohlordnungsprinzip). Für jede nichtleere Menge $M \subset \mathbb{N}$ existiert* $\min M$.

Beweis: Weil $[1, \infty)$ eine induktive Menge ist, ist $\mathbb{N}$ nach unten durch 1 beschränkt. Also existiert $\inf M$. Dann gibt es ein Element $m \in M$ mit $m < \inf M + 1$, weil $\inf M + 1$ keine untere Schranke von M ist. Wegen dem vorangehenden Satz sind alle Elemente von $\mathbb{N}$ und damit auch M entweder kleiner oder gleich $m - 1$ oder größer oder gleich m. Wegen der Monotonie gilt $m - 1 < \inf M$. Also sind alle Elemente von M nicht kleiner oder gleich $m - 1$, und damit gößer oder gleich m. Dann ist $m = \min M$.**q.e.d.**

Satz 2.38. *(Archimedes-Eudoxos)*

(i) *Für jede reelle Zahl $x \in \mathbb{R}$ gibt es eine natürliche Zahl $n > x$.*

(ii) *Für jedes $\epsilon > 0$ gibt es ein $n \in \mathbb{N}$ mit $\frac{1}{n} < \epsilon$.*

Beweis: (i) Es reicht zu zeigen, dass $\mathbb{N}$ keine obere Schranke hat. Wenn $\mathbb{N}$ eine obere Schranke hat, dann existiert $\sup \mathbb{N}$, und $\exists n \in \mathbb{N}$ mit $n > \sup \mathbb{N} - 1$, weil $\sup \mathbb{N} - 1$ keine obere Schranke von $\mathbb{N}$ ist. Dann gilt $\mathbb{N} \ni n + 1 > \sup \mathbb{N}$, was ein Widerspruch ist.
(ii) Nach (i) gibt es ein $n > \frac{1}{\epsilon}$. Aus Satz 2.13 (v)-(vi) folgt $\frac{1}{n} < \epsilon$. **q.e.d.**

Definition 2.39. *(ganze Zahlen $\mathbb{Z}$, rationale Zahlen $\mathbb{Q}$)*

$$\begin{aligned}
\mathbb{Z} &= \mathbb{N} \cup \{0\} \cup \{x \in \mathbb{R} \mid -x \in \mathbb{N}\} \\
\mathbb{N}_0 &= \mathbb{N} \cup \{0\} \\
\mathbb{Q} &= \left\{ x \in \mathbb{R} \mid \exists m \in \mathbb{Z} \text{ und } \exists n \in \mathbb{N} \text{ mit } x = \frac{m}{n} \right\}.
\end{aligned}$$

Wegen dem Satz von Archimedes-Eudoxos gibt es viele rationale Zahlen.

Satz 2.40. *Sei $a < b$. Dann existiert $r \in \mathbb{Q}$ mit $a < r < b$.*

Beweis: Für alle $n \in \mathbb{N}$ ist $a < \frac{m}{n} < b$ äquivalent zu $na < m < nb$. Wegen $b - a > 0$ gibt es nach Satz 2.38 (i) ein $n \in \mathbb{N}$ mit $n > \frac{1}{b-a}$. Dann gilt $na < nb$ und $nb - na > 1$.

Wir nehmen zunächst $a \geq 0$ an. Sei m die kleinste natürliche Zahl, die größer ist als na. Wegen Satz 2.33 sind die natürlichen Zahlen enthalten in $\{1\} \cup \{x \in \mathbb{R} \mid x - 1 \in \mathbb{N}\}$. Also ist $m - 1$ entweder eine natürliche Zahl oder gleich Null. Dann folgt

$$m - 1 \leq na < m \quad \iff \quad na < m \leq na + 1 < nb.$$

Als nächstes nehmen wir $b \leq 0$ an. Sei $-m$ die kleinste natürliche Zahl, die größer ist als $-nb$. Wieder ist $-m - 1$ entweder eine natürliche Zahl oder Null. Also folgt

$$-m - 1 \leq -nb < -m \quad \iff \quad na < nb - 1 \leq m < nb.$$

Wenn $a < 0$ und $b > 0$ ist, wählen wir $m = 0$. **q.e.d.**

Wir haben sogar gezeigt, dass $\forall a < b$ mit $b - a > 1$ das Intervall (a, b) eine ganze Zahl enthält. Also enthält jedes Intervall mindestens eine und damit sogar unendlich viele rationale Zahlen. Insbesondere gibt es für jede reelle Zahl und jedes $\epsilon > 0$ eine rationale Zahl $r \in (x - \epsilon, x + \epsilon)$, die dann $d(x, r) < \epsilon$ erfüllt. Wir sagen, dass $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt. Die rationalen Zahlen erfüllen als Unterkörper der reellen Zahlen die Axiome **A1**-**A4**. Wir werden gleich sehen, dass sie nicht das Vollständigkeitsaxiom **A5** erfüllen.

2.4 Die Peano Axiome

Es gibt andere Möglichkeiten die reellen Zahlen einzuführen. Man kann zuerst die natürlichen Zahlen einführen und danach der Reihe nach die ganzen Zahlen, die rationalen Zahlen und die rellen Zahlen. Dabei kann man die natürlichen Zahlen im Rahmen der Mengenlehre einführen oder sie durch die Peano Axiome charakterisieren:

1. Peano Axiom: $1 \in \mathbb{N}$ (Einselement).

2. Peano Axiom: $\forall n \in \mathbb{N} \quad \exists$ genau einen Nachfolger $n' \in \mathbb{N}$ (Nachfolgerabbildung).

3. Peano Axiom: Für alle $n \in \mathbb{N}$ gilt $n' \neq 1$ (Eins $\notin$ Bild der Nachfolgerabbildung).

4. Peano Axiom: Für alle $n, m \in \mathbb{N}$ folgt $n = m$ aus $n' = m'$ (Injektivität der Nachfolgerabbildung).

5. Peano Axiom: Jede Teilmenge $S \subset \mathbb{N}$ mit folgenden Eigenschaften umfaßt die natürlichen Zahlen $\mathbb{N} \subset S$ (Jede Teilmenge, die die Eins enthält und invariant ist unter der Nachfolgerabbildung ist gleich $\mathbb{N}$).

(i) $1 \in S$. **(ii)** Falls $n \in S$, dann ist auch $n' \in S$.

Man kann die natürlichen Zahlen also durch diese Axiome charakterisieren und dann damit die reellen Zahlen einfüren. Dafür muss wesentlich mehr Schlussfolgerungsarbeit geleistet werden, bevor die rellen Zahlen eingeführt sind. Wir wollen in diesem Abschnitt kurz vorstellen, wodurch sich diese zweite Einführung der reellen Zahlen, von der von uns gewählten unterscheidet, und sie in ihren Grundzügen vorstellen.
Die Addition von natürlichen Zahlen $n + m$ wird $\forall n, m \in \mathbb{N}$ so eingeführt, dass **(i)** $n + 1 = n'$ für alle $n \in \mathbb{N}$ und **(ii)** $n + m' = (n + m)'$ für alle $n, m \in \mathbb{N}$ gilt.
Für $n = 1$ gilt $1+1 = 1'$ und für alle $m \in \mathbb{N}$ folgt aus $1+m = m'$ auch $1+m' = (m')' = (1 + m)'$. Wegen dem 5. Peano Axiom ist also $1 + m$ für alle $m \in \mathbb{N}$ durch m' definiert. Wenn für ein $n \in \mathbb{N}$ durch diese Bedingungen $n + m$ für alle $m \in \mathbb{N}$ definiert ist, dann folgt aus (i) $n' + 1 = (n')' = (n + 1)'$. Wenn $n' + m$ für ein $m \in \mathbb{N}$ definiert ist, dann folgt aus (ii) $n' + m' = (n' + m)'$ also ist dann wegen dem 5. Peano Axiom $n' + m$ für alle $m \in \mathbb{N}$ definiert. Also ist wegen dem 5. Peano Axiom die Menge aller $n \in \mathbb{N}$, für die $n + m$ durch diese Bedingungen für alle $m \in \mathbb{N}$ definiert ist gleich $\mathbb{N}$. Deshalb ist dadurch die Addition zwischen allen natürlichen Zahlen definiert. Mit Hilfe der Peano Axiome kann man dann folgern, dass die Addition kommutativ und assoziativ ist.
Die Multiplikation von natürlichen Zahlen nm wird $\forall n, m \in \mathbb{N}$ so eingeführt, dass **(i)** $n1 = n$ für alle $n \in \mathbb{N}$ und **(ii)** $nm' = n \cdot m + n$ für alle $n, m \in \mathbb{N}$ gilt.
Wieder kann man aus den Peano Axiomen folgern, dass diese Bedingungen eine Multiplikation $n \cdot m$ für alle $n, m \in \mathbb{N}$ definiert. Danach zeigt man, dass auch die Multiplikation kommutativ, assoziativ und zusammen mit der Addition distributiv ist.

Die Ordnungsrelation wird auf den natürlichen Zahlen durch $n < n + m$ und $m < n + m \quad \forall n, m \in \mathbb{N}$ eingeführt. Sie erfüllt **A4** wobei alle $n \in \mathbb{N}$ als positiv gelten.
Ganze Zahlen: Die ganzen Zahlen werden als formale Differenzen $n - m$ eingeführt. Wegen $n - m = \tilde{n} - \tilde{m} \Leftrightarrow n + \tilde{m} = \tilde{n} + m$ sind das Äquivalenzklassen in $\mathbb{N} \times \mathbb{N}$

$$(n, m) \sim (\tilde{n}, \tilde{m}) \qquad \Longleftrightarrow \qquad n + \tilde{m} = \tilde{n} + m \qquad \forall (n, m), (\tilde{n}, \tilde{m}) \in \mathbb{N} \times \mathbb{N}$$

Die Addition, Multiplikation und Ordnungsrelation von $\mathbb{N}$ wird auf $\mathbb{Z}$ fortgesetzt.
Rationale Zahlen: Die rationalen Zahlen werden als formale Brüche $\frac{m}{n}$ mit $(m, n) \in \mathbb{Z} \times \mathbb{N}$ eingeführt. Wegen $\frac{m}{n} = \frac{\tilde{m}}{\tilde{n}} \Leftrightarrow m\tilde{n} = \tilde{m}n$ sind das Äquivalenzklassen in $\mathbb{Z} \times \mathbb{N}$

$$(m, n) \sim (\tilde{m}, \tilde{n}) \qquad \Longleftrightarrow \qquad m\tilde{n} = \tilde{m}n \qquad \forall (m, n), (\tilde{m}, \tilde{n}) \in \mathbb{Z} \times \mathbb{N}$$

Addition, Multiplikation, Ordnungsrelation wird auf $\mathbb{Q}$ definiert, so dass **A1-A4** gilt.
Reelle Zahlen werden als Dedekindsche Schnitte eingeführt, also Mengen $A \subset \mathbb{Q}$ mit

(i) $A \neq \emptyset, \mathbb{Q}$ **(ii)** $p \in A$ und $q < p \Rightarrow q \in A$. **(iii)** A hat kein Maximum.
Addition, Multiplikation, Ordnungsrelation wird auf $\mathbb{R}$ definiert, so dass **A1-A5** gilt.

2.5 Wurzeln und Intervallschachtelung

Satz 2.41. *(Quadratwurzeln) Für alle $a > 0$ gibt es genau ein $b > 0$ mit $b^2 = a$.*

Wir schreiben $b = \sqrt{a} = a^{\frac{1}{2}}$.
Beweis der Eindeutigkeit: Aus $0 < x < y$ folgt aufgrund der Monotonie $x^2 < xy < y^2$. Also gibt es höchstens ein $b > 0$ mit $b^2 = a$.
Beweis der Existenz: Die Menge $M = \{x \in \mathbb{R}_0^+ \mid x^2 < a\}$ enthält 0. Aus $x > a + 1$ folgt

$$x^2 > x(a + 1) > (a + 1)^2 = a^2 + 2a + 1 > 2a > a.$$

Also ist $a + 1$ eine obere Schranke von M. Sei $b = \sup M$. Aus $0 \in M$ folgt $0 \leq b$.
Wenn $b^2 < a$, folgt $0 < a - b^2$. Für $h = \min\{1, \frac{a-b^2}{2b+2}\}$ gilt $0 < h \leq 1$ und $h(2b + 1) < h(2b + 2) \leq a - b^2$. Hieraus folgt

$$(b + h)^2 = b^2 + 2bh + h^2 \leq b^2 + 2bh + h = b^2 + h(2b + 1) < b^2 + a - b^2 = a.$$

Also ist $b < b + h \in M$ im Widerspruch zu $b = \sup M$.
Wenn $a < b^2$, folgt $0 < b$, $-b^2 < a - b^2 < 0$ und $-\frac{1}{2} < \frac{a-b^2}{2b^2} < 0$. Also folgt aus der Bernoulli Ungleichung

$$\left(b\left(1 + \frac{a - b^2}{2b^2}\right)\right)^2 = b^2\left(1 + \frac{a - b^2}{2b^2}\right)^2 \geq b^2\left(1 + 2\frac{a - b^2}{2b^2}\right) = a.$$

Dann folgt aus $x > b(1 + \frac{a-b^2}{2b^2})$ wieder $x^2 > b^2(1 + \frac{a-b^2}{2b^2})^2 \geq a$ und damit auch $x \notin M$. Also ist $b(1 + \frac{a-b^2}{2b^2}) < b$ eine obere Schranke von M im Widerspruch zu $b = \sup M$.

Weil also weder $b^2 < a$ noch $a < b^2$ gilt, folgt $b^2 = a$. **q.e.d.**

Wir werden in Beispiel 3.9 für jedes $n \in \mathbb{N}$ zeigen, dass es für jedes $a > 0$ genau ein $b > 0$ gibt mit $b^n = a$. Wir schreiben dann $b = \sqrt[n]{a} = a^{\frac{1}{n}}$. Insbesondere gibt es also genau ein $\sqrt{2} \in \mathbb{R}$, aber wie wir gleich sehen werden gilt $\sqrt{2} \notin \mathbb{Q}$. Also erfüllt $\mathbb{Q}$ tatsächlich nicht das Vollständigkeitsaxiom **A5**.

Lemma 2.42. $\sqrt{2} \notin \mathbb{Q}$

Beweis: Wir zeigen, dass es nicht zwei natürliche Zahlen $n, m \in \mathbb{N}$ geben kann mit $\left(\frac{m}{n}\right)^2 = 2$. Wenn es zwei solche Zahlen gibt, dann enthält die Teilmenge

$$M = \{n \in \mathbb{N} \mid \exists m \in \mathbb{N} \text{ mit } m^2 = 2n^2\}$$

der natürlichen Zahlen ein kleinstes Element n_0. Sei also $m_0^2 = 2n_0^2$. Wenn m_0 ungerade ist, also $m_0 = 2m_1 - 1$ mit $m_1 \in \mathbb{N}$, dann ist auch $m_0^2 = (2m_1-1)^2 = 2(2m_1^2 - 2m_1 + 1) - 1$ ungerade, im Widerspruch zu $m_0^2 = 2n_0^2$. Also ist m_0 gerade mit $m_0 = 2m_2$ und $m_2 \in \mathbb{N}$. Dann folgt $m_0^2 = 4m_2^2 = 2n_0^2$ oder auch $2m_2^2 = n_0^2$. Dann gehört m_2 zu M und es gilt $n_0 \leq m_2$. Das ergibt folgenden Widerspruch $m_2^2 < 2m_2^2 = n_0^2 \leq n_0 m_2 \leq m_2^2$. **q.e.d.**

Satz 2.43. *(Intervallschachtelungsprinzip) Seien $I_n = [a_n, b_n] \subset \mathbb{R}$, abgeschlossene Intervalle $a_n < b_n$ für $n = 1, 2, \cdots$ mit folgenden Eigenschaften.*

(i) *$I_{n+1} \subset I_n$ für alle $n \in \mathbb{N}$*

(ii) *Für jedes $\epsilon > 0$ gibt es ein $n \in \mathbb{N}$, so dass $b_n - a_n < \epsilon$.*

Dann enthält $\bigcap_{n\geq 1} I_n = \{x\}$ der Durchschnitt genau ein $x \in \mathbb{R}$.

Dieser Satz ist falsch für offene Intervalle. So ist z.B. $\bigcap_{n\geq 1} \left(0, \frac{1}{n}\right) = \emptyset$ nach dem Satz von Archimedes-Eudoxos.

Beweis: Wegen (i) gilt

$$a_1 \leq a_2 \leq \cdots \leq a_n \leq a_{n+1} \leq \cdots \leq b_{n+1} \leq b_n \leq \cdots \leq b_2 \leq b_1$$

Also besteht die Menge $A = \{a_1, a_2, \cdots\}$ aus unteren Schranken der Menge $B = \{b_1, b_2, \cdots\}$ und umgekehrt die Menge B aus oberen Schranken der Menge A. Sei also $x = \sup A$ und $y = \inf B$. Dann sind x und y obere Schranken von A und untere Schranken von B. Also gilt $x \leq y$ und $x, y \in \bigcap_{n>1} I_n$. Es gilt sogar $\bigcap_{n>1} I_n = [x, y]$. Andererseits ist wegen (ii) für alle $\epsilon > 0$ und alle $n \in \mathbb{N}$ auch $y - x \leq b_n - a_n < \epsilon$. Dann muss aber $0 \leq y - x \leq \inf\{\epsilon | \epsilon > 0\} = 0$ und deshalb $y = x$ gelten. **q.e.d.**

Umgekehrt folgt aus dem Intervallschachtelungsprinzip, den Axiomen **A1**-**A4** und dem Satz von Archimedes-Eudoxos das Vollständigkeitsaxiom (Übungsaufgabe). Deshalb können die reellen Zahlen auch durch die Axiome **A1**-**A4**, den Sazt von Archimedes-Eudoxos, und das Intervallschachtelungsprinzip charakterisiert werden.

2.6 Mächtigkeit von Mengen

Definition 2.44. *Zwei Mengen heißen gleichmächtig, wenn es eine bijektive Abbildung zwischen ihnen gibt.*

Offensichtlich ist die Relation von Mengen gleichmächtig zu sein eine Äquivalenzrelation, d.h. sie ist reflexiv, symmetrisch und transitiv. Deshalb stellt sich die Frage, die Äquivalenzklassen dieser Relation zu bestimmen. Eine Menge ist genau dann endlich, wenn sie für ein $n \in \mathbb{N}$ gleichmächtig ist zu $\{1, \ldots, n\}$. Für $n \neq m$ sind die Mengen $\{1, \ldots, n\}$ und $\{1, \ldots, m\}$ nicht gleichmächtig. Deshalb entsprechen die Äquivalenzklassen der endlichen Mengen genau den natürlichen Zahlen. Indem wir die Eins mit der Menge $\{\emptyset\}$ indentifizieren und die Nachfolgerabbildung für Mengen durch $A \mapsto A \cup \{A\}$ definieren, erhalten wir die natürtlichen Zahlen als folgende Mengen:

$$
\begin{array}{ccl}
1 & \leftrightarrow & \{\emptyset\} \\
2 & \leftrightarrow & \{\emptyset, \{\emptyset\}\} \\
3 & \leftrightarrow & \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \\
4 & \leftrightarrow & \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\} \\
\vdots & & \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots\}
\end{array}
$$

Alle Äquivalenzklassen kann man als eine Erweiterung von $\mathbb{N}$ betrachten.

Definition 2.45. *Eine nicht leere Menge A heißt*

endlich, *falls es ein $n \in \mathbb{N}$ gibt, so dass A gleichmächtig ist zu $\{1, 2, \cdots, n\}$.*

unendlich, *falls sie nicht endlich ist.*

abzählbar, *falls sie gleichmächtig ist zu $\mathbb{N}$.*

höchstens abzählbar, *wenn sie endlich oder abzählbar ist.*

Satz 2.46. **(i)** *Jede nichtleere Teilmenge von $\mathbb{N}$ ist höchstens abzählbar.*

(ii) *Eine Menge A ist genau dann höchstens abzählbar, wenn es eine surjektive Abbildung von $\mathbb{N}$ auf A gibt.*

Beweis: **(i)** Jede Teilmenge $M \subset \mathbb{N}$ können wir aufgrund des Wohlordnungsprinzip der Größe nach mit dem kleinsten Element anfangend durchnumerieren. Wenn M nach oben unbeschränkt ist, erhalten wir so eine bijektive Abbildung von $\mathbb{N}$ nach M und M ist abzählbar. Andernfalls ist $M \subset \{1, \ldots, n\}$ für ein $n \in \mathbb{N}$ und damit endlich.
(ii) Sei A eine höchstens abzählbare Menge. Wenn A endlich ist, gibt es ein $n \in \mathbb{N}$, so dass A gleichmächtig ist zu $\{1, 2, \cdots, n\}$. Die Abbildung $\mathbb{N} \to \{1, 2, \cdots, n\}, \quad m \mapsto$

$\min\{m,n\}$ ist eine surjektive Abbildung, so dass dann auch eine surjektive Abbildung auf A existiert. Wenn A abzählbar ist, gibt es eine bijektive Abbildung von $\mathbb{N}$ auf A.

Sei umgekehrt A eine Menge und f eine surjektive Abbildung von $\mathbb{N}$ auf A. Dann existiert eine Abbildung $g : A \to \mathbb{N}, \quad a \mapsto \min f^{-1}[\{a\}]$, die offenbar eine bijektive Abbildung von A auf eine Teilmenge von $\mathbb{N}$ definiert. Also ist A gleichmächtig zu einer Teilmenge von $\mathbb{N}$ und damit wegen (i) höchstens abzählbar. **q.e.d.**

Übungsaufgabe 2.47. *Zeige, dass jede endliche Teilmenge der reellen Zahlen ein Maximum und ein Minimum besitzt.*

Satz 2.48. **(i)** *Die Menge* $\mathbb{N} \times \mathbb{N}$ *ist abzählbar.*

(ii) *Eine höchstens abzählbare Vereinigung von höchstens abzählbaren Mengen ist wieder höchstens abzählbar.*

Beweis: **(i)** Wir definieren eine injektive Abbildung f von $\mathbb{N} \times \mathbb{N}$ nach $\mathbb{N}$ durch die sogenannte Diagonalnummerierung:

$$f(x,y) = y + \frac{(x+y-2)(x+y-1)}{2} = y + \sum_{n=0}^{x+y-2} n.$$

Diese Abbildung ist injektiv. Gilt nämlich $x + y < x' + y'$ so folgt aus

$$f(x,y) - y \leq f(x',y') - y' - (x'+y'-2) \quad \text{und } y - y' < y \leq x + y - 1 \leq x' + y' - 2$$

$$f(x,y) \leq f(x',y') + y - y' - (x'+y'-2) < f(x',y').$$

Gilt aber $x + y = x' + y'$ so gilt auch $f(x,y) - f(x',y') = y - y'$.
Also definiert diese Abbildung eine bijektive Abbildung von $\mathbb{N} \times \mathbb{N}$ auf eine Teilmenge von $\mathbb{N}$. Dann folgt (i) aus (i) vom vorangehenden Satz.
(ii) Wegen (ii) im vorangehenden Satz genügt es eine surjektive Abbildung $n \mapsto A_n$ von $\mathbb{N}$ in die höchstens abzählbaren Mengen zu betrachten. Wegen (ii) im vorangehenden Satz gibt es dann für alle $n \in \mathbb{N}$ eine surjektive Abbildung $f_n : \mathbb{N} \to A_n$. Dann ist

$$F : \mathbb{N} \times \mathbb{N} \to \bigcup_{n\in\mathbb{N}} A_n, \quad (n,m) \mapsto f_n(m)$$

eine surjektive Abbildung. Also folgt (ii) aus (i) und dem vorangehenden Satz. **q.e.d.**

Korollar 2.49. $\mathbb{Q}$ *ist abzählbar.*

Beweis: $\mathbb{Z} \times \mathbb{N} \to \mathbb{Q}, \quad (m,n) \mapsto \frac{m}{n}$ ist eine surjektive Abbildung. $\mathbb{Z}$ ist abzählbar, also ist $\mathbb{Q}$ höchstens abzählbar. $\mathbb{Q}$ ist aber nicht endlich und damit abzählbar. **q.e.d.**

Satz 2.50. *Die Menge der reellen Zahlen ist nicht abzählbar.*

Beweis: Wir nehmen an $\mathbb{R}$ ist abzählbar. Sei also $(x_n)_{n\in\mathbb{N}}$ eine Durchnumerierung von $\mathbb{R}$. Wähle $a, b \in \mathbb{R}$ mit $a < b$. Sei $I_1 = [a, b]$. Wir definieren induktiv eine Folge $(I_n)_{n\in\mathbb{N}}$ von Intervallen. Für alle $n \in \mathbb{N}$ wählen wir dafür ein abgeschlossenes Teilintervall I_{n+1} von I_n, das nur ein Drittel so groß ist wie I_n und x_n nicht enthält. Für alle $n \in \mathbb{N}$ liegt dann x_n nicht in dem Intervall I_n. Wegen dem Intervallschachtelungsprinzip ist aber $\bigcap_{n\in\mathbb{N}} I_n$ nicht leer. Also gibt es eine reelle Zahl, die nicht zu der Menge $\{x_n \mid n \in \mathbb{N}\}$ gehört. Das widerspricht der Annahme, dass $\mathbb{R}$ abzählbar ist. **q.e.d.**

Auch die Potenzmenge der natürlichen Zahlen ist nicht abzählbar. Wir können sie mit den Folgen identifizieren, die Werte in $\mathbb{Z}_2 = \{0, 1\}$ annehmen, also der Menge aller Abbildungen von $\mathbb{N}$ nach $\mathbb{Z}_2$. Diese Folgen lassen sich aufgrund der dyadischen Entwicklung mit der Teilmenge $[0, 1] \subset \mathbb{R}$ identifizieren. Diese ist gleichmächtig zu $\mathbb{R}$.

2.7 Der Körper der komplexen Zahlen

Motivation: Wir hatten aus der Ordnungsrelation gefolgert, dass $x^2 > 0$ gilt, falls $x \neq 0$. Deshalb existieren in den reellen Zahlen keine Quadratwurzeln von negativen Zahlen. Erweitert man die reellen Zahlen durch eine Quadratwurzel $\imath$ von -1, so erhält man die komplexen Zahlen, in denen sich dann alle algebraischen Gleichungen lösen lassen.

Definition 2.51. *(Komplexe Zahlen) Die Komplexen Zahlen $\mathbb{C}$ ist die Menge $\mathbb{R} \times \mathbb{R}$ aller geordneten Paare (x, y) von reellen Zahlen zusammen mit den Operationen*

$$\begin{aligned} + : &\quad \mathbb{C} \times \mathbb{C} \to \mathbb{C}, \quad ((x, y), (u, v)) \mapsto (x, y) + (u, v) = (x + u, y + v) \\ \cdot : &\quad \mathbb{C} \times \mathbb{C} \to \mathbb{C}, \quad ((x, y), (u, v)) \mapsto (xu - yv, xv + yu) \end{aligned}$$

Mit dieser Definition erfüllt $\mathbb{C}$ die Körperaxiome **A1-A3**. Dabei ist

$$\text{Null} : \quad 0_{\mathbb{C}} = (0, 0) \qquad \text{Eins} : \quad 1_{\mathbb{C}} = (1, 0)$$

$$\begin{aligned} \text{negatives Element} : &\quad -(x, y) = (-x, -y) \\ \text{inverses Element} : &\quad (x, y)^{-1} = \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2}\right) \qquad \text{für}(x, y) \neq (0, 0). \end{aligned}$$

Wir bezeichnen komplexe Zahlen meistens auch nur durch einen Buchstaben, üblicherweise z. Die Null und die Eins bezeichnen wir auch durch 0 und 1, so dass aus dem Zusammenhang klar werden muss, ob es sich um die Null bzw. Eins der reellen oder der komplexen Zahlen handelt. Außerdem benutzen wir dieselben Abkürzungen wie bei den reellen Zahlen:

$$z + (-z) = z - z = 0 \qquad z^{-1} = \frac{1}{z} \qquad z \cdot \frac{1}{z} = 1 \qquad \text{usw.}$$

Da die komplexen Zahlen eine Erweiterung der reellen Zahlen sind, können wir die reellen Zahlen als Teilmenge der komplexen Zahlen auffassen. Dafür definieren wir eine injektive Abbildung

$$\mathbb{R} \hookrightarrow \mathbb{C}, \quad x \mapsto (x, 0)$$

Offenbar ist diese Abbildung verträglich mit den Operationen $+$ und $\cdot$ von $\mathbb{R}$ und $\mathbb{C}$:

$$\begin{aligned} (x,0)+(y,0) &= (x+y,0) & \qquad (x,0)\cdot(y,0) &= (xy,0) \\ (0,0) &= 0 & (1,0) &= 1 \\ -(x,0) &= (-x,0) & (x,0)^{-1} &= (x^{-1},0) \end{aligned}$$

Definition 2.52. *(imaginäre Einheit)* $\imath = (0,1) \in \mathbb{C}$.

Dann gilt $\imath^2 = (-1,0) = -1$. Wir können also auch schreiben $(x,y) = x + \imath y$. Dann ergeben sich die Operationen $+$ und $\cdot$

$$\begin{aligned} x + \imath y + u + \imath v &= x + u + \imath(y, v) \\ (x+\imath y)(u+\imath v) &= xu + \imath(yu + xv) + \imath^2 yv = xu - yv + \imath(yu + xv) \end{aligned}$$

Für die komplexe Zahl $z = x + \imath y$ heißt

$x = \Re(z)$ Realteil von z		$y = \Im(z)$ Imaginärteil von z
z heißt reell, falls	$\Im(z) = 0$	
z heißt imaginär, falls	$\Re(z) = 0$	

Definition 2.53. *(komplexe Konjugation) Die Abbildung* $\mathbb{C} \to \mathbb{C}, \quad z = x + \imath y \mapsto \bar{z} = x - \imath y$ *heißt komplexe Konjugation oder einfach nur Konjugation*

Die komplexen Zahlen erhalten wir aus den reellen Zahlen, indem wir reelle Vielfache einer Wurzel aus -1 hinzufügen. Weil aber $(-1) \cdot (-1) = 1$ ist das Negative einer Wurzel aus -1 wieder eine Wurzel aus -1. Welche dieser beiden Wurzeln aus -1 wir zur imaginären Einheit machen ist aber eine Konvention. Deshalb ist die Konjugation ein Endomorphismus der komplexen Zahlen, d.h. eine bijektive Abbildung, die mit den Operationen $+$ und $\cdot$ verträglich ist, die die reellen Zahlen invariant läßt.

Satz 2.54. (i) $(\bar{\bar{z}}) = z$

(ii) $\overline{z + w} = \bar{z} + \bar{w}$

(iii) $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$

(iv) $z + \bar{z} = 2\Re z$ *und* $z - \bar{z} = 2\imath\Im(z)$

(v) $z \cdot \bar{z}$ *ist reell und nicht negativ.*

Beweis: (i)-(iv) nachrechnen. **(v)** $(x + \imath y)(x - \imath y) = x^2 + y^2 \geq 0$. **q.e.d.**

Definition 2.55. *(Betrag) Der Betrag einer komplexen Zahl* $z = x + \imath y$ *ist definiert als die reelle nicht negative Wurzel aus* $z \cdot \bar{z} : |\cdot| \mathbb{C} \to \mathbb{R}, \quad z \mapsto |z| = \sqrt{z \cdot \bar{z}}$.

Satz 2.56. *(Eigenschaften des Betrags)*

(i) $|z| \geq 0$ *und* $|z| = 0 \iff z = 0$

(ii) $|z \cdot w| = |z| \cdot |w|$

(iii) $|z + w| \leq |z| + |w|$ *(Dreiecksungleichung)*

(iv) $||z| - |w|| \leq |z - w|$

(v) $|\bar{z}| = |z|$

(vi) $|\Re(z)| \leq |z|$ *und* $|\Im(z)| \leq |z|$

Beweis: (ii) $|z \cdot w|^2 = zw\overline{(zw)} = z\bar{z}w\bar{w} = (|z||w|)^2$ Wegen der Eindeutigkeit der Wurzel folgt dann $|z \cdot w| = |z| \cdot |w|$.
(vi) Für $z = 0$ ist die Aussage offensichtlich. Sei also $z = x + \imath y \neq 0$. Dann gilt $x^2 \leq x^2 + y^2$. Aus $\sqrt{x^2 + y^2} < x$ folgt aber mit Monotonie $x^2 + y^2 < x\sqrt{x^2 + y^2} < x^2$. Also gilt $x \leq \sqrt{x^2 + y^2}$ und damit auch $|\Re(z)| \leq |z|$. Durch vertauschen von x und y erhalten wir $|\Im(z)| \leq |z|$.

$$\textbf{(iii)} \quad \begin{aligned} |z+w|^2 = (z+w)(\bar{z}+\bar{w}) &= z\bar{z} + z\bar{w} + w\bar{z} + w\bar{w} &&= z\bar{z} + 2\mathbb{R}(z\bar{w}) + w\bar{w} \\ \leq |z|^2 + 2|z\bar{w}| + |w|^2 &= |z|^2 + 2|z||w| + |w|^2 &&= (|z| + |w|)^2. \end{aligned}$$

Aus $|z| + |w| < |z + w|$ folgt mit Monotonie $(|z| + |w|)^2 < (|z| + |w|)|z + w| < |z + w|^2$. Also gilt $|z + w| \leq |z| + |w|$.
(iv) folgt aus (i)-(iii) genau wie im reellen Fall.
(v) $|\bar{z}|^2 = \bar{z}z = |z|^2$. Dann folgt wegen der Eindeutigkeit der Wurzel $|\bar{z}| = |z|$. **q.e.d.**

Definition 2.57. *(Abstand) Der Abstand zweier komplexer Zahlen ist die nicht negative Zahl* $d : \mathbb{C} \times \mathbb{C} \to \mathbb{R}, \quad (z, w) \mapsto d(z, w) = |z - w|$

Aus den Eigenschaften des Betrages folgt genau wie im Reellen.

Satz 2.58. *(Eigenschaften des Abstandes)*

(i) $d(z, w) \geq 0$ *und* $d(z, w) = 0 \iff z = w$

(ii) $d(z, w) = d(w, z)$

(iii) $d(z, w) \leq d(z, u) + d(u, w)$. *(Dreiecksungleichung).*

Wir veranschaulichen die komplexen Zahlen in der zweidimensionalen Ebene. Der Abstand ist der euklidische Abstand zwischen den entsprechenden Punkten der Ebene.

Kapitel 3

Zahlenfolgen

3.1 Konvergenz

Im Folgenden werden wir des öfteren Aussagen vorstellen, die sowohl für die reellen Zahlen als auch für die komplexen Zahlen gelten. Wir benutzen dann das Symbol $\mathbb{K}$ um entweder die reellen oder die komplexen Zahlen zusammen mit den entsprechenden Abbildungen und Operationen zu bezeichnen. Die Elemente von $\mathbb{K}$ wollen wir dann einfach Zahlen nennen. Eine Folge ist eine Abbildung $\mathbb{N} \to \mathbb{K}$, $n \mapsto a_n$. Wir bezeichnen sie mit $(a_n)_{n\in\mathbb{N}}$. Wir interessieren uns vorwiegend für die Grenzwerte solcher Zahlenfolgen.

Definition 3.1. *Die Folge $(a_n)_{n\in\mathbb{N}}$ heißt konvergent, wenn es eine Zahl $a \in \mathbb{K}$ gibt und für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $|a_n - a| < \epsilon$ für alle natürlichen Zahlen $n \geq N$ gilt. Die Zahl a heißt dann Grenzwert der Folge $(a_n)_{n\in\mathbb{N}}$*

Wir schreiben dann $\lim_{n\to\infty} a_n = a$ oder $\lim a_n = a$ oder auch $a_n \to a$ für $n \to \infty$.

Beispiel 3.2. $\lim\limits_{n\to\infty} \frac{1}{n} = 0$

Beweis: Nach dem Satz von Archimedes-Eudoxos gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ mit $\frac{1}{N} < \epsilon$. Dann gilt wegen (iv) in Satz 2.13 für alle $n \geq N$ auch $|\frac{1}{n} - 0| < \epsilon$. **q.e.d.**

Satz 3.3. **(i)** *Der Grenzwert einer konvergenten Folge ist eindeutig.*

(ii) *Eine komplexe Folge konvergiert genau dann, wenn die entsprechenden Folgen der Realteile und Imaginärteile konvergieren.*

(iii) *Für jede konvergente Zahlenfolge $(a_n)_{n\in\mathbb{N}}$ ist $\{|a_n| \mid n \in \mathbb{N}\} \subset \mathbb{R}$ beschränkt.*

Beweis: **(i)** Seien a und b zwei Grenzwerte einer konvergenten Folge $(a_n)_{n\in\mathbb{N}}$, so dass $|a_n - a| < \frac{\epsilon}{2}$ für alle $n \geq N$ und $|a_n - b| < \frac{\epsilon}{2}$ für alle $n \geq M$ gilt. Dann folgt

$$0 \leq |a - b| \leq |a - a_n| + |a_n - b| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \text{ für alle } n \geq \max\{N, M\}.$$

Dann gilt auch $0 \leq |a-b| \leq \inf(0,\infty) = 0$. Also ist $|a-b| = 0$ und damit auch $a = b$.
(ii) Die Realteile und Imaginärteile einer komplexen Folge $(z_n)_{n\in\mathbb{N}}$ bilden zwei reelle Folgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ mit $z_n = x_n + iy_n$ für alle $n \in \mathbb{N}$. Sei $z = x + iy$ die Zerlegung einer komplexen Zahl in Realteil und Imaginärteil. Wegen Satz 2.56 (vi)

$$\max\{|x_n - x|, |y_n - y|\} \leq |z_n - z|$$

konvergieren die Real- bzw. Imaginärteile einer konvergenten komplexen Folge gegen den Realteil bzw. Imaginärteil des Grenzwertes. Konvergieren umgekehrt die Folgen $(x_n)_{n\in\mathbb{N}}$ bzw. $(y_n)_{n\in\mathbb{N}}$ gegen x bzw. y, dann gibt es für jedes $\epsilon > 0$ zwei natürliche Zahlen $N, M \in \mathbb{N}$ so dass $|x_n - x| < \frac{\epsilon}{2}$ für alle $n \geq N$ gilt, und $|y_n - y| < \frac{\epsilon}{2}$ für alle $n \geq M$. Dann gilt für alle $n \geq \max\{N, M\}$

$$|x_n + iy_n - x + iy| \leq |x_n - x| + |y_n - y| < \epsilon.$$

Also konvergiert eine komplexe Folge genau dann, wenn ihre Realteile und Imaginärteile konvergieren.
(iii) Sei $(a_n)_{n\in\mathbb{N}}$ eine konvergente Folge mit Grenzwert a. Dann gibt es ein $N \in \mathbb{N}$, so dass $|a_n - a| < 1$ für alle $n \geq N$ gilt und damit auch $|a_n| \leq |a_n - a + a| < 1 + |a|$. Daraus folgt $\quad |a_n| \leq \max\{|a| + 1, |a_1|, |a_2|, \cdots, |a_{N-1}|\} \quad$ für alle $n \in \mathbb{N}$. **q.e.d.**

Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt divergent, wenn sie nicht konvergiert, wenn es also kein $a \in \mathbb{K}$ gibt, so dass $\lim_{n\to\infty} a_n = a$. Wenn es für eine reelle Folge für jedes $b \in \mathbb{R}$ ein $N \in \mathbb{N}$ gibt, so dass für alle natürlichen Zahlen $n \geq N$ gilt $a_n > b$ bzw. $a_n < b$ dann schreiben wir $\lim_{n\to\infty} a_n = \infty$ bzw. $\lim_{n\to\infty} a_n = -\infty$ Weil in beiden Fällen die Folgen nicht beschränkt sind, können sie wegen dem vorangehenden Satz nicht konvergieren. Es gilt offenbar $\lim_{n\to\infty} n = \infty$ und $\lim_{n\to\infty} -n = -\infty$.

Satz 3.4. (i) *Sei $|x| < 1$ dann gilt* $\lim_{n\to\infty} x^n = 0$

(ii) *Sei $|x| > 1$ dann ist $(x^n)_{n\in\mathbb{N}}$ divergent.*

(iii) *Sei $x = 1$ dann gilt* $\lim_{n\to\infty} x^n = 1$

(iv) *Sei $x \in (1,\infty)$ dann gilt* $\lim_{n\to\infty} x^n = \infty$.

(v) *Sei $|x| = 1$ und $x \neq 1$ dann ist $(x^n)_{n\in\mathbb{N}}$ divergent.*

Beweis: (i) Wenn $x = 0$ ist gilt natürlich $\lim_{n\to\infty} x^n = 0$. Sei also $0 < |x| < 1$. Dann ist $1 < \frac{1}{|x|} = 1 + y$ mit $y = \frac{1-|x|}{|x|}$. Dann folgt aus der Bernoulli–Ungleichung $\frac{1}{|x|^n} = (1+y)^n \geq 1 + ny > ny = n\frac{1-|x|}{|x|}$. Dann folgt $|x|^n < \frac{1}{n}\frac{|x|}{1-|x|}$. Aus dem Satz von Archimedes-Eudoxos folgt dann, dass es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $\frac{1}{N} < \epsilon \cdot \frac{1-|x|}{|x|}$. Daraus folgt für alle $n \geq N$

$$|x^n - 0| = |x|^n < \frac{1}{n}\frac{|x|}{1-|x|} \leq \frac{1}{N}\frac{|x|}{1-|x|} < \epsilon.$$

(iii) Wegen $1^n = 1$ ist $\lim_{n\to\infty} 1^n = 1$.
(iv) Sei $x > 1$. Dann ist $y = x - 1 > 0$. Also gilt aufgrund der Bernoulli–Ungleichung

$$x^n = (1+y)^n \geq 1 + ny > ny = n(x-1).$$

Dann gibt es für jedes $b \in \mathbb{R}$ ein $N \in \mathbb{N}$ mit $N > \frac{b}{x-1}$. Dann folgt für alle $n \geq N$:

$$x^n > n(x-1) \geq N(x-1) > b.$$

Also gilt $\lim_{n\to\infty} x^n = \infty$.
(ii) Für $|x| > 1$ gilt wegen (iv) $\lim_{n\to\infty} |x^n| = \infty$. Also ist $(x^n)_{n\in\mathbb{N}}$ nicht beschränkt.
(v) Für alle $y \in \mathbb{K}$ und alle natürlichen Zahlen n gilt

$$|y - x^n| + |y - x^{n+1}| \geq |x^n - x^{n+1}| \geq |x-1| \cdot |x|^n.$$

Für $|x| = 1$ mit $x \neq 1$ gilt also $\qquad \max\{|y - x^n|, |y - x^{n+1}|\} \geq \dfrac{|x-1|}{2}$.
Also kann es für $x \neq 1$ keinen Grenzwert geben. **q.e.d.**

Satz 3.5. *(Rechenregeln) Für konvergente Zahlenfolgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ gilt*

(i) $\lim\limits_{n\to\infty}(x_n + y_n) = \lim\limits_{n\to\infty} x_n + \lim\limits_{n\to\infty} y_n$

(ii) $\lim\limits_{n\to\infty} \lambda x_n = \lambda \lim\limits_{n\to\infty} x_n$ *für alle* $\lambda \in \mathbb{K}$.

(iii) $\lim\limits_{n\to\infty} x_n y_n = (\lim\limits_{n\to\infty} x_n)(\lim\limits_{n\to\infty} y_n)$

(iv) *Wenn $x_n \neq 0$ für alle $n \in \mathbb{N}$ und $\lim\limits_{n\to\infty} x_n \neq 0$, dann gilt auch $\lim\limits_{n\to\infty} \frac{1}{x_n} = \frac{1}{\lim\limits_{n\to\infty} x_n}$.*

(v) $\lim\limits_{n\to\infty} |x_n| = |\lim\limits_{n\to\infty} x_n|$.

(vi) *Wenn zwei reelle Folgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ für alle $n \in \mathbb{N}$ $x_n \leq y_n$ erfüllen, dann gilt auch $\lim\limits_{n\to\infty} x_n \leq \lim\limits_{n\to\infty} y_n$.*

(vii) *Wenn zwei reelle konvergente Folgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ für alle $n \in \mathbb{N}$ $x_n \leq y_n$ erfüllen und den gleichen Grenzwert haben, dann gilt für jede reelle Folge $(z_n)_{n\in\mathbb{N}}$, die $x_n \leq z_n \leq y_n$ für alle $n \in \mathbb{N}$ erfüllt, $\lim\limits_{n\to\infty} z_n = \lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} y_n$.*

Beweis: (i) Seien $x = \lim\limits_{n\to\infty} x_n$ und $y = \lim\limits_{n\to\infty} y_n$. Dann gilt es für jedes $\epsilon > 0$ natürliche Zahlen $N, M \in \mathbb{N}$, so dass $|x_n - x| < \frac{\epsilon}{2}$ für alle $n \geq N$ und $|y_n - y| < \frac{\epsilon}{2}$ für alle $n \geq M$ gilt. Dann gilt $|x_n + y_n - (x+y)| \leq |x_n - x| + |y_n - y| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ für alle $n \geq \max\{N, M\}$. Also konvergiert $(x_n + y_n)_{n\in\mathbb{N}}$ gegen $x + y$.

(ii) Für $\lambda = 0$ gilt $\lim\limits_{n\to\infty} \lambda x_n = \lambda \cdot \lim\limits_{n\to\infty} x_n$. Sei also $\lambda \neq 0$ und damit $0 < |\lambda|$. Dann gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $|x_n - x| < \frac{\epsilon}{|\lambda|}$ für alle $n \geq N$ gilt. Daraus folgt:

$$|\lambda x_n - \lambda x| \leq |\lambda| \cdot |x_n - x| < \epsilon \qquad \text{für alle } n \geq N.$$

(iii) Seien $x = \lim\limits_{n\to\infty} x_n$ und $y = \lim\limits_{n\to\infty} y_n$. Weil $(x_n)_{n\in\mathbb{N}}$ konvergiert, gibt es ein $\lambda > 0$ mit $|x_n| < \lambda$ für alle $n \in \mathbb{N}$. Also gibt es für alle $\epsilon > 0$ zwei natürliche Zahlen $N, M \in \mathbb{N}$, so dass $|x_n - x| < \frac{\epsilon}{2|y|+1}$ für $n \geq N$ gilt und $|y_n - y| < \frac{\epsilon}{2\lambda}$ für alle $n \geq M$. Dann gilt

$$|x_n y_n - xy| = |x_n y_n - x_n y + x_n y - xy| \leq |x_n| \cdot |y_n - y| + |x_n - x| \cdot |y| < \frac{\epsilon}{2} + \frac{\epsilon |y|}{2|y|+1} \leq \epsilon$$

für alle $n \geq \max\{N, M\}$.

(iv) Aufgrund der Voraussetzungen gibt es ein $N \in \mathbb{N}$, so dass $|x_n - x| < \frac{|x|}{2}$ für alle $n \geq N$ gilt. Daraus folgt

$$|x_n| \geq |x| - |x - x_n| > \frac{|x|}{2} \qquad \text{und} \qquad \frac{1}{|x_n|} < \frac{2}{|x|} \qquad \text{für alle } n \geq N.$$

Dann gibt es für jedes $\epsilon > 0$ ein $M \in \mathbb{N}$, so dass $|x_n - x| < \frac{\epsilon \cdot |x|^2}{2}$ für alle $n \geq M$ gilt. Dann gilt für alle $n \geq \max\{N, M\}$ auch

$$\left| \frac{1}{x_n} - \frac{1}{x} \right| = \left| \frac{x - x_n}{x_n x} \right| = \frac{|x - x_n|}{|x_n| \cdot |x|} < \frac{\epsilon \cdot |x|^2}{2} \frac{2}{|x|} \frac{1}{|x|} = \epsilon$$

(v) folgt aus der Ungleichung $\big||x_n| - |x|\big| \leq |x_n - x|$.

(vi) Seien $x = \lim\limits_{n\to\infty} x_n$ und $y = \lim\limits_{n\to\infty} y_n$. Für alle $\epsilon > 0$ gibt es zwei natürliche Zahlen N und M, so dass $|x_n - x| < \frac{\epsilon}{2}$ für alle $n \geq N$ gilt und $|y_n - y| < \frac{\epsilon}{2}$ für alle $n \geq M$. Dann gilt für alle $n \geq \max\{N, M\}$

$$x - y \leq (x - x_n) + (y_n - y) + (x_n - y_n) < \epsilon$$

Also ist $x - y \leq \inf(0, \infty) \leq 0$. Daraus folgt $x \leq y$.

(vii) Seien $x = \lim\limits_{n\to\infty} x_n$ und $y = \lim\limits_{n\to\infty} y_n$. Für alle $\epsilon > 0$ gibt es zwei natürliche Zahlen $N, M \in \mathbb{N}$, so dass $|x_n - x| < \epsilon$ für alle $n \geq N$ gilt und $|y_n - x| < \epsilon$ für alle $n \geq M$. Für alle $n \geq \max\{N, M\}$ gilt dann

$$-\epsilon < x_n - x \leq z_n - x \leq y_n - x < \epsilon$$

Daraus folgt $|z_n - x| < \epsilon$. **q.e.d.**

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

$$(1 + x + x^2 + \ldots + x^n)(1 - x) = (1 + x + \ldots + x^n) - (x + x^2 + \ldots + x^{n+1}) = 1 - x^{n+1}$$

Also gilt $$1 + x + \ldots + x^n = \frac{1 - x^{n+1}}{1 - x} \qquad \text{für } x \neq 1 \text{ und } n \in \mathbb{N}_0.$$

Aus Satz 3.4 folgt, dass die Folge $y_n = 1 + x + \ldots + x^n$ für $|x| < 1$ gegen $\frac{1}{1-x}$ konvergiert.

Satz 3.6. $$\lim_{n\to\infty}\sum_{m=0}^{n} x^m = \frac{1}{1-x} \quad \textit{für } |x|<1.$$ **q.e.d.**

3.2 Konvergenzprinzipien

Wir stellen 3 Methoden vor, um zu entscheiden, ob eine Folge konvergiert oder nicht.

Definition 3.7. *(Monotonie) Eine reelle Folge $(a_n)_{n\in\mathbb{N}}$ heißt:*
monoton wachsend, *wenn $a_{n+1} \geq a_n$ für alle $n \in \mathbb{N}$.*
streng monoton wachsend, *wenn $a_{n+1} > a_n$ für alle $n \in \mathbb{N}$.*
monoton fallend, *wenn $a_{n+1} \leq a_n$ für alle $n \in \mathbb{N}$.*
streng monoton fallend, *wenn $a_{n+1} < a_n$ für alle $n \in \mathbb{N}$.*

Satz 3.8. *(Monotonieprinzip)*

(i) *Eine monoton wachsende (fallende) beschränkte reelle Folge $(a_n)_{n\in\mathbb{N}}$ ist konvergent, und es gilt*
$$\lim_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} a_n \quad \textit{bzw.} \quad \lim_{n\to\infty} a_n = \inf_{n\in\mathbb{N}} a_n.$$

(ii) *Für eine monoton wachsende (fallende) unbeschränkte reelle Folge gilt*
$$\lim_{n\to\infty} a_n = \infty \quad \textit{bzw.} \quad \lim_{n\to\infty} a_n = -\infty.$$

Beweis: **(i)** Sei $(a_n)_{n\in\mathbb{N}}$ eine monoton wachsende (fallende) beschränkte reelle Folge. Dann existiert $\sup_{n\in\mathbb{N}} a_n$ bzw. $\inf_{n\in\mathbb{N}} a_n$. Für alle $\epsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass
$$a_N > \sup_{n\in\mathbb{N}} a_n - \epsilon \text{ bzw. } a_N < \inf_{n\in\mathbb{N}} a_n + \epsilon.$$
Dann gilt aber für alle $m \geq N$:
$$\sup_{n\in\mathbb{N}} a_n - \epsilon < a_N \leq a_m \leq \sup_{n\in\mathbb{N}} a_n \text{ bzw. } \inf_{n\in\mathbb{N}} a_n \leq a_m \leq a_N < \inf_{n\in\mathbb{N}} a_n + \epsilon.$$
Dann gilt aber auch $|a_m - \sup_{n\in\mathbb{N}} a_n| < \epsilon$ bzw. $|a_m - \inf_{n\in\mathbb{N}} a_n| < \epsilon$.
(ii) Sei $(a_n)_{n\in\mathbb{N}}$ eine monoton wachsende (fallende) reelle unbeschränkte Folge. Dann gibt es für jedes $b \in \mathbb{R}$ ein $N \in \mathbb{N}$, so dass $a_N > b$ bzw. $a_N < b$ gilt. Dann gilt aber für alle $m \geq N$ auch $a_m \geq a_N > b$ bzw. $a_m \leq a_N < b$. **q.e.d.**

Beispiel 3.9. *Existenz und Konstruktion der k–ten Wurzel. Für alle $a > 0$ und alle $k \in \mathbb{N} \setminus \{1\}$ definieren wir die Folge $(a_n)_{n\in\mathbb{N}}$ rekursiv durch*
$$a_0 = 1 + \frac{a-1}{k} \qquad a_{n+1} = a_n \left(1 + \frac{a - a_n^k}{k \cdot a_n^k}\right) \quad \textit{für alle } n \in \mathbb{N}_0.$$
Für $a = 1$ ist dann $a_n = 1$ für alle $n \in \mathbb{N}$. Für $a \neq 1$ gilt für alle $n \in \mathbb{N}$
$$0 < a_n < a_0, \qquad a_n < a_{n-1} \textit{ und } \qquad a < a_n^k.$$

Beweis durch vollständige Induktion:
(i) Für $a > 0$ und $a \neq 1$ ist $-1 < -\frac{1}{k} < \frac{a-1}{k} \neq 0$. Wegen der Bernoulli–Ungleichung gilt $a_0^k = \left(1 + \frac{a-1}{k}\right)^k > 1 + a - 1 = a$. Es folgt $-1 < -\frac{1}{k} < \frac{a-a_0^k}{ka_0^k} < 0$. Also gilt $0 < a_1 < a_0$ und mit der Bernoulli–Ungleichung $a_1^k > a_0^k \left(1 + \frac{a - a_0^k}{ka_0^k}\right)^k > a_0^k \left(1 + k\frac{a - a_0^k}{ka_0^k}\right) = a$.
(ii) Wir nehmen an es gilt $0 < a_n < a_0$, $a_n < a_{n-1}$ und $a < a_n^k$ für ein $n \in \mathbb{N}$. Dann folgt $-1 < -\frac{1}{k} < \frac{a-a_n^k}{k\cdot a_n^k} < 0$. Daraus folgt $0 < a_{n+1} < a_n < a_0$ und wegen der Bernoulli–Ungleichung: $a_{n+1}^k > a_n^k \left(1 + \frac{a - a_n^k}{ka_n^k}\right)^k > a_n^k \left(1 + k\frac{a - a_n^k}{ka_n^k}\right) = a$. **q.e.d.**

Die Folge $(a_n)_{n\in\mathbb{N}}$ ist also monoton fallend und beschränkt. Wegen dem Monotonieprinzip konvergiert sie dann. Sei $b = \lim_{n\to\infty} a_n$. Wir formulieren die Rekursionsgleichung um zu

$$a_{n+1} \cdot ka_n^{k-1} = (k-1)a_n^k + a.$$

Bilden wir links und rechts den Grenzwert $n \to \infty$ so erhalten wir

$$kb^k = (k-1)b^k + a \text{ oder auch } b^k = a.$$

Also existiert eine Zahl b mit $b^k = a$. Diese Folge konvergiert sehr schnell. Außerdem sind für rationale a alle Folgenglieder rational.

Satz 3.10. **(i)** *Für jede positive Zahl $a > 0$ und jede rationale Zahl r gibt es genau eine positive Zahl a^r. Für r=0 setzen wir $a^0 = 1$.*

(ii) *Für jede positive rationale Zahl $r > 0$ und $0 < a < b$ gilt auch $0 < a^r < b^r$.*

(iii) *Für jede negative rationale Zahl $r < 0$ und $0 < a < b$ gilt auch $0 < b^r < a^r$.*

Beweis: **(i)** Für $r = \frac{p}{q}$ definieren wir $a^r = \sqrt[q]{a^p}$. Offenbar ist a^r eine positive Lösung der Gleichungen $(a^r)^{qn} = (a^p)^n = a^{pn}$ für alle $n \in \mathbb{N}$. Wir zeigen, dass für $a, b \in \mathbb{R}^+$ und für $n \in \mathbb{N}$ die Ungleichung $a < b$ äquivalent ist zu $a^n < b^n$. Denn für $n \in \mathbb{N} \setminus \{1\}$ gilt

$$b^n - a^n = (b-a)(b^{n-1} + b^{n-2}a + \ldots + ba^{n-2} + a^{n-1}).$$

Also folgt aus $0 < a < b$ auch $a^n < b^n$ und aus $0 < b \leq a$ auch $a^n \leq b^n$. Also ist für $a > 0$ und $b > 0$ die Ungleichung $a < b$ äquivalent zu der Ungleichung $a^n < b^n$. Also gibt es für alle $n \in \mathbb{N}$ und alle $\frac{p}{q} > 0$ genau eine positive Lösung der Gleichung $y^{qn} = a^{pn}$. Für $r < 0$ ist $a^r = \frac{1}{a^{-r}}$ die entsprechende positive Zahl.
(ii) Sei $r = \frac{p}{q} > 0$. Dann ist für $a, b \in \mathbb{R}^+$ die Ungleichung $a < b$ äquivalent zu $a^p < b^p$ und das wiederum äquivalent zu $a^{\frac{p}{q}} < b^{\frac{p}{q}}$.
(iii) Sei $r = \frac{-p}{q} < 0$. Dann folgt (iii) aus (ii) wegen (vi) Satz 2.13. **q.e.d.**

Definition 3.11. *(Teilfolge) Eine Teilfolge einer Folge* $(a_n)_{n\in\mathbb{N}}$ *ist eine Folge* $(b_m)_{m\in\mathbb{N}}$ *von der Form* $b_m = a_{n_m}$*, wobei* $0 < n_1 < n_2 < \ldots$ *eine streng monoton wachsende Folge von natürlichen Zahlen ist.*

Z.B. hat die divergente Folge $a_n = (-1)^n$ zwei konvergente Teilfolgen $b_m = a_{2m} = 1$ und $c_m = a_{2m+1} = -1$.

Satz 3.12. *(monotone Teilfolgen) Jede reelle Folge enthält eine monotone Teilfolge.*

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ eine reelle Folge und A die Menge $A = \{n \in \mathbb{N} \mid a_n \geq a_m \forall m > n\}$. Wenn A eine unendliche Menge ist, dann sei $n_1 < n_2 < \ldots$ eine Abzählung der Elemente von A. Die Teilfolge $(b_m)_{m\in\mathbb{N}} = (a_{n_m})_{m\in\mathbb{N}}$ ist dann monoton fallend.

Wenn A eine endliche Menge ist besitzt sie ein Maximum N. Dann gibt es also zu jedem $n > N$ ein $m > n$, so dass $a_m > a_n$ ist. Also definieren wir induktiv eine Teilfolge $(b_m)_{m\in\mathbb{N}}$, so dass $b_{m+1} > b_m$ gilt für alle $m \in \mathbb{N}$. Diese Folge ist streng monoton steigend. Also enthält $(a_n)_{n\in\mathbb{N}}$ entweder eine monoton fallende oder eine streng monoton steigende Folge. Umgekehrt gilt auch, dass $(a_n)_{n\in\mathbb{N}}$ entweder eine monoton steigende oder eine streng monoton fallende Folge enthält. **q.e.d.**

Satz 3.13. *(Auswahlprinzip von Bolzano–Weierstraß) Jede beschränkte reelle Folge besitzt eine konvergente Teilfolge.*

Beweis: Wegen dem vorangehenden Satz besitzt jede Folge $(a_n)_{n\in\mathbb{N}}$ eine monotone Teilfolge $(b_n)_{n\in\mathbb{N}}$. Wenn die ursprüngliche Folge beschränkt ist, ist auch die Teilfolge beschränkt. Diese konvergiert dann wegen dem Monotonieprinzip. **q.e.d.**

Korollar 3.14. *Jede beschränkte komplexe Folge besitzt eine konvergente Teilfolge.*

Beweis: Wegen $\max\{|x|, |y|\} \leq |x + \imath y|$ sind die reellen Folgen der Realteile und Imaginärteile einer komplexen beschränkten Folge beschränkt. Wegen dem Auswahlprinzip von Bolzano–Weierstraß besitzt die Folge der Realteile eine konvergente Teilfolge, und dann auch die entsprechende Teilfolge der Imaginärteile. Wegen Satz 3.3 (ii) konvergiert die der zweiten Teilfolge entsprechende komplexe Teilfolge. **q.e.d.**

Definition 3.15. *(Cauchyfolge) Eine Folge* $(a_n)_{n\in\mathbb{N}}$ *heißt Cauchyfolge, wenn es für jedes* $\epsilon > 0$ *ein* $N \in \mathbb{N}$ *gibt, so dass* $|a_n - a_m| < \epsilon$ *für alle* $n, m \geq N$ *gilt.*

Satz 3.16. *(Kriterium von Cauchy) Eine Zahlenfolge konvergiert genau dann, wenn sie eine Cauchyfolge ist.*

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ eine konvergente Folge. Dann gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass alle $m \geq N$ die Ungleichung $|a_m - \lim_{n\to\infty} a_n| < \frac{\epsilon}{2}$ erfüllen. Also gilt auch

$$|a_m - a_l| \leq |a_m - \lim_{n\to\infty} a_n| + |\lim_{n\to\infty} a_n - a_l| < \epsilon \quad \text{für alle } m, l \geq N.$$

Sei umgekehrt $(a_n)_{n\in\mathbb{N}}$ eine Cauchyfolge. Dann gibt es ein $N \in \mathbb{N}$, so dass für alle $m \geq N$ gilt $|a_m - a_N| < 1$, und damit auch $|a_m| \leq |a_N| + |a_m - a_N| < |a_N| + 1$. Für alle $n \in \mathbb{N}$ gilt dann aber $|a_n| \leq \max\{|a_1|, \ldots, |a_{N-1}|, |a_N| + 1\}$ Deshalb ist die Folge a_n beschränkt und besitzt wegen dem Auswahlprinzip von Bolzano–Weierstraß eine konvergente Teilfolge $(b_n)_{n\in\mathbb{N}}$. Sei $a = \lim_{n\to\infty} b_n$. Dann gibt es für jedes $\epsilon > 0$ zwei natürliche Zahlen $N, M \in \mathbb{N}$, so dass alle $n, m \geq N$ die Ungleichung $|a_n - a_m| < \frac{\epsilon}{2}$ erfüllen und alle $n \geq M$ die Ungleichung $|b_n - a| < \frac{\epsilon}{2}$. Weil $(b_n)_{n\in\mathbb{N}}$ eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$ ist, folgt $|a_n - a| \leq |a_n - b_n| + |b_n - a| < \epsilon$ für alle $n \geq \max\{N, M\}$. **q.e.d.**

Unter der Annahme der Axiome **A1-A4** und dem Satz von Archimedes-Eudoxos sind folgende Aussagen äquivalent:

Vollständigkeitsaxiom A5.

(i) aus dem Monotonieprinzip.

Auswahlprinzip von Bolzano–Weierstraß.

Jede Cauchyfolge konvergiert.

Intervallschachtelungsprinzip.

Wir können die reellen Zahlen auch als Äquivalenzklasse von Cauchyfolgen von rationalen Zahlen auffassen, wobei zwei Cauchyfolgen äquivalent heißen, wenn ihre Differenz eine Nullfolge ist.

3.3 Häufungspunkte

Definition 3.17. *(Häufungspunkt) Die Grenzwerte von konvergenten Teilfolgen heißen Häufungspunkte. Bei reellen Teilfolgen sind zusätzlich $+\infty$ bzw. $-\infty$ Häufungspunkte, wenn es Teilfolgen mit diesen Grenzwerten gibt.*

Satz 3.18. *(Limes superior und Limes inferior) Ist die Menge der reellen Häufungspunkte einer reellen Folge nicht leer und nach oben (unten) in $\mathbb{R}$ beschränkt, so besitzt sie ein Maximum (Minimum) in $\mathbb{R}$.*

Beweis: Wir nehmen an, dass die Menge der Häufungspunkte der Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ nach oben beschränkt ist. Sei $a \in \mathbb{R}$ das Supremum der Häufungspunkte, dann gibt es für jedes $m \in \mathbb{N}$ einen Häufungspunkt $b_m \in (a - \frac{1}{2m}, a]$. Dann gibt es aber auch eine Teilfolge $(c_m)_{m\in\mathbb{N}}$ von $(a_n)_{n\in\mathbb{N}}$, die $|c_m - b_m| < \frac{1}{2m}$ für alle $m \in \mathbb{N}$ erfüllt. Dann gilt aber für alle $m \in \mathbb{N}$ sogar $|c_m - a| \leq |c_m - b_m| + |b_m - a| < \frac{1}{m}$. Also ist a ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$ und damit ein Maximum aller dieser Häufungspunkte. Der Beweis für das Minimum ist analog. **q.e.d.**

Definition 3.19. *Für eine nach oben (unten) beschränkte reelle Folge heißt das Supremum (bzw. Infimum) der Häufungspunkte Limes superior bzw. inferior. Wir bezeichnen es mit* $\overline{\lim} a_n = \limsup\limits_{n\to\infty} a_n$ *bzw.* $\underline{\lim} a_n = \liminf\limits_{n\to\infty} a_n$.

Satz 3.20. **(i)** $\overline{a}$ *ist genau dann der Limes superior einer nach oben beschränkten reellen Folge* $(a_n)_{n\in\mathbb{N}}$*, wenn für alle* $\epsilon > 0$ *unendlich viele Elemente der Folge* $a_n > \overline{a} - \epsilon$ *erfüllen, aber höchstens endlich viele* $a_n > \overline{a} + \epsilon$.

(ii) $\underline{a}$ *ist genau dann der Limes inferior einer nach unten beschränkten reellen Folge* $(a_n)_{n\in\mathbb{N}}$*, wenn für alle* $\epsilon > 0$ *unendlich viele Elemente* $a_n < \underline{a} + \epsilon$ *erfüllen, aber höchstens endliche viele* $a_n < \underline{a} - \epsilon$.

Beweis: Wir beweisen wieder nur (i), weil (ii) analog zu beweisen ist. Sei $(a_n)_{n\in\mathbb{N}}$ eine nach oben beschränkte Folge. Wegen dem Auswahlprinzip von Bolzano–Weierstraß ist jede untere Schranke einer Teilfolge von $(a_n)_{n\in\mathbb{N}}$ auch eine untere Schranke von mindestens einem Häufungspunkt. Also ist die Charakterisierung der Zahl $\overline{a}$ in (i) äquivalent dazu, dass alle Zahlen, die kleiner sind als $\overline{a}$ keine oberen Schranken der Häufungspunkte von $(a_n)_{n\in\mathbb{N}}$ sind, aber alle Zahlen, die größer sind als $\overline{a}$ obere Schranken der Häufungspunkte von $(a_n)_{n\in\mathbb{N}}$ sind. Dann ist $\overline{a}$ der maximale Häufungspunkt. **q.e.d.**

Korollar 3.21. *Eine reelle Folge konvergiert genau dann, wenn sie beschränkt ist und wenn* $\overline{\lim} a_n = \underline{\lim} a_n$ *gilt, sie also nur einen Häufungspunkt hat.*

Weil $\overline{\lim} a_n$ und $\underline{\lim} a_n$ das Maximum und das Minimum der Häufungspunkte sind, ist $\overline{\lim} a_n = \underline{\lim} a_n$ äquivalent zu der Bedingung, dass es nur einen Häufungspunkt gibt.
Beweis: Wenn $(a_n)_{n\in\mathbb{N}}$ beschränkt ist und $\overline{\lim} a_n = \underline{\lim} a_n = a$ gilt, dann folgt aus dem vorangehenden Satz, dass es für jedes $\epsilon > 0$ ein N gibt, so dass für alle $n \geq N$ gilt $\underline{\lim} a_n - \epsilon \leq a_n \leq \overline{\lim} a_n + \epsilon$. Also gilt $|a_n - a| \leq \epsilon$ und $(a_n)_{n\in\mathbb{N}}$ konvergiert gegen a.
Wenn $(a_n)_{n\in\mathbb{N}}$ konvergiert, dann konvergiert jede Teilfolge gegen $a = \lim_{n\to\infty} a_n$. Also besitzt $(a_n)_{n\in\mathbb{N}}$ nur einen Häufungspunkt und es gilt $\overline{\lim} a_n = \underline{\lim} a_n = a$. **q.e.d.**

Korollar 3.22. *Seien* $(a_n)_{n\in\mathbb{N}}$ *und* $(b_n)_{n\in\mathbb{N}}$ *nach unten (bzw. oben) beschränkte reelle Folgen, die für alle* $n \in \mathbb{N}$ $a_n \leq b_n$ *erfüllen. Dann gilt* $\underline{\lim} a_n \leq \underline{\lim} b_n$ *bzw.* $\overline{\lim} a_n \leq \overline{\lim} b_n$.

Beweis: Sei $(d_n)_{n\in\mathbb{N}}$ eine Teilfolge von $(b_n)_{n\in\mathbb{N}}$, die gegen $\underline{\lim} b_n$ (bzw. $(c_n)_{n\in\mathbb{N}}$ eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$ die gegen $\overline{\lim} a_n$) konvergiert. Dann gibt es eine Teilfolge $(c_n)_{n\in\mathbb{N}}$ von $(a_n)_{n\in\mathbb{N}}$ (bzw. $(d_n)_{n\in\mathbb{N}}$ von $(b_n)_{n\in\mathbb{N}}$), die für alle $n \in \mathbb{N}$ auch $c_n \leq d_n$ erfüllt. Diese Teilfolge ist beschränkt und besitzt eine konvergente Teilfolge. Wir können also ohne Beschränkung der Allgemeinheit annehmen, dass $(c_n)_{n\in\mathbb{N}}$ (bzw. $(d_n)_{n\in\mathbb{N}}$) konvergiert. Dann ist aber $\lim_{n\to\infty} c_n$ ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$ (bzw. $\lim_{n\to\infty} d_n$ ein Häfungspunkt von $(b_n)_{n\in\mathbb{N}}$). Also gilt $\underline{\lim} a_n \leq \lim\limits_{n\to\infty} c_n \leq \lim\limits_{n\to\infty} d_n = \underline{\lim} b_n$
(bzw. $\overline{\lim} a_n = \lim\limits_{n\to\infty} c_n \leq \lim\limits_{n\to\infty} d_n \leq \overline{\lim} b_n$). **q.e.d.**

3.4 Beispiele

(i) Für alle $x \in \mathbb{C}$ ist $\lim\limits_{n\to\infty} \frac{x^n}{n!} = 0$.

Beweis: Wegen dem Satz von Archimedes–Endoxes gibt es ein $N \in \mathbb{N}$ mit $|x| < N$. Dann gilt für alle $n \geq N$

$$\left|\frac{x^n}{n!}\right| = \frac{|x|^{N-1}}{(N-1)!} \cdot \frac{|x|^{n+1-N}}{N \cdots n} \leq \frac{|x|^{N-1}}{(N-1)!}\left(\frac{|x|}{N}\right)^{n-N} \cdot \frac{|x|}{n} < \frac{|x|^N}{(N-1)!} \cdot \frac{1}{n}$$

Weil aber $\frac{1}{n}$ eine Nullfolge ist, konvergiert dann $\left|\frac{x^n}{n!}\right|$ auch gegen Null. **q.e.d.**

(ii) Für alle positiven rationalen Zahlen $r > 0$ gilt $\lim\limits_{n\to\infty} \frac{1}{n^r} = 0$.

Beweis: Nach dem Satz von Archimedes–Eudoxos, gibt es für alle $\epsilon > 0$ ein $N \in \mathbb{N}$ mit $N > \frac{1}{\epsilon^{\frac{1}{r}}}$. Dann gilt aber für alle $n \geq N$ auch $\frac{1}{n^r} \leq \frac{1}{N^r} < \epsilon$. Also konvergiert $\frac{1}{n^r}$ nach Null. **q.e.d.**

(iii) Für alle $x \in \mathbb{R}^+$ gilt $\lim\limits_{n\to\infty} \sqrt[n]{x} = 1$.

Beweis: Sei zunächst $x \geq 1$. Sei also $y_n = \sqrt[n]{x} - 1 \geq 0$. Wegen der Bernoulli–Ungleichung gilt dann $x = (1 + y_n)^n \geq 1 + ny_n$. Daraus folgt $0 \leq y_n \leq \frac{x-1}{n}$. Dann konvergiert y_n aber gegen Null. $\lim\limits_{n\to\infty} \sqrt[n]{x} = 1 + \lim\limits_{n\to\infty} y_n = 1$.

Sei jetzt $0 < x < 1$. Dann ist $\frac{1}{x} > 1$. Also gilt $\lim\limits_{n\to\infty} \sqrt[n]{x} = \dfrac{1}{\lim\limits_{n\to\infty} \sqrt[n]{\frac{1}{x}}} = 1$. **q.e.d.**

Binomische Formel 3.23. *Für alle $n \in \mathbb{N}$ und alle Zahlen $x, y \in \mathbb{K}$ gilt*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}, \text{ wobei } \binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \text{ und } \binom{n}{0} = 1.$$

Beweis: durch vollständige Induktion:
(i) Offenbar ist $\binom{1}{0} = \binom{1}{1} = 1$, und die Binomische Formel ist für $n = 1$ richtig.
(ii) Wenn die Binomische Formel für $n \in \mathbb{N}$ gilt, dann folgt

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)(x+y)^n = (x+y)\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k} \\
&= \sum_{k=1}^{n+1}\binom{n}{k-1}x^k y^{n+1-k} + \sum_{k=0}^{n}\binom{n}{k}x^k y^{n+1-k} \\
&= \sum_{k=0}^{n+1}\left(k\frac{n(n-1)\cdots(n-k+2)}{k!} + \frac{n(n-1)\cdots(n-k+1)}{k!}\right)x^k y^{n+1-k} \\
&= \sum_{k=0}^{n+1}\left(\frac{(n+1)n\cdots(n-k+2)}{k!}\right)x^k y^{n+1-k}.
\end{aligned}$$

Also gilt sie auch für $n+1$. **q.e.d.**

(iv) $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

Beweis: Sei $y_n = \sqrt[n]{n} - 1 \geq 0$. Wegen der Binomischen Formel gilt für alle $n \in \mathbb{N}$

$$n = \left(\sqrt[n]{n}\right)^n = (1+y_n)^n = \sum_{k=0}^{n} \binom{n}{k} y_n^k$$

Für alle $n \geq 2$ folgt $\quad n \geq 1 + \frac{n(n-1)}{2} y_n^2 \iff y_n^2 \leq \frac{2}{n} \iff y_n \leq \sqrt{\frac{2}{n}}.$
Also ist y_n eine Nullfolge. **q.e.d.**

(v) $\lim_{n\to\infty} \sqrt[n]{n!} = \infty$.

Beweis: Wegen (i) gibt es für alle $x \in \mathbb{R}$ ein N, so dass für alle $n \geq N$ gilt $\frac{x^n}{n!} < 1$. Dann gilt aber auch $\frac{x}{\sqrt[n]{n!}} < 1 \iff x < \sqrt[n]{n!}$. Also folgt $\lim_{n\to\infty} \sqrt[n]{n!} = \infty$. **q.e.d.**

Satz 3.24. *Sei $(a_n)_{n\in\mathbb{N}}$ eine positive Folge, dann gilt*

$$\underline{\lim} \left(\frac{a_{n+1}}{a_n}\right) \leq \underline{\lim} \sqrt[n]{a_n} \qquad \overline{\lim} \sqrt[n]{a_n} \leq \overline{\lim} \left(\frac{a_{n+1}}{a_n}\right).$$

Wenn die Folge $\sqrt[n]{a_n}$ also einen reellen Häufungspunkt hat, dann hat auch die Folge $\frac{a_{n+1}}{a_n}$ einen reellen Häufungspunkt. Und wenn gilt $\overline{\lim} \left(\frac{a_{n+1}}{a_n}\right) < \infty$ dann auch $\overline{\lim} \sqrt[n]{a_n} < \infty$.

Beweis: Wenn $\underline{\lim} \left(\frac{a_{n+1}}{a_n}\right) = 0$ ist, ist die erste Aussage trivial. Wir nehmen also an, dass es ein $a > 0$ gibt und für alle $0 < \epsilon < a$ ein $N \in \mathbb{N}$ gibt, so dass alle $n \geq N$ auch $\frac{a_{n+1}}{a_n} \geq a - \epsilon$ erfüllen. Dann gilt für alle $n > N$

$$\frac{a_n}{a_N} = \frac{a_{N+1}}{a_N} \cdots \frac{a_n}{a_{n-1}} \geq (a-\epsilon)^{n-N} \Longrightarrow a_n \geq a_N \frac{(a-\epsilon)^n}{(a-\epsilon)^N} \Longrightarrow \sqrt[n]{a_n} \geq \sqrt[n]{\frac{a_N}{(a-\epsilon)^N}} \cdot (a-\epsilon).$$

Wegen (iii) gilt dann $\underline{\lim} \sqrt[n]{a_n} \geq (a-\epsilon)$. Weil dies für alle $\epsilon > 0$ gilt, folgt auch $\underline{\lim} \sqrt[n]{a_n} \geq a$. Also ist $\underline{\lim} \sqrt[n]{a_n} \geq a$ nicht kleiner als der kleinste Häufungspunkt von $(\frac{a_{n+1}}{a_n})_{n\in\mathbb{N}}$. Und wenn diese Folge $(\frac{a_{n+1}}{a_n})_{n\in\mathbb{N}}$ keinen reellen Häufungspunkt hat, dann hat auch die Folge $(\sqrt[n]{a_n})_{n\in\mathbb{N}}$ keinen reellen Häufungspunkt.

Für die Folge $(a_n^{-1})_{n\in\mathbb{N}}$ gilt entsprechend $\underline{\lim} \frac{a_n}{a_{n+1}} \leq \underline{\lim} \sqrt[n]{a_n^{-1}} = \underline{\lim} \left(\sqrt[n]{a_n}\right)^{-1}$. Wegen Satz 2.13 (vi) und Satz 3.5 (iv) ist für jede positive Folge $(b_n)_{n\in\mathbb{N}}$ $\underline{\lim} b_n^{-1} = \left(\overline{\lim} b_n\right)^{-1}$. Also folgt die zweite Ungleichung aus Satz 2.13 (vi). **q.e.d.**

(vi) $\left(\sum\limits_{k=0}^{n}\frac{1}{k!}\right)_{n\in\mathbb{N}}$ ist streng monoton wachsend und beschränkt. Für $n > 3$ gilt

$$\begin{aligned}\frac{5}{2} < \sum_{k=0}^{n}\frac{1}{k!} &< 2+\sum_{k=1}^{n-1}\frac{1}{k(k+1)} &&= 2+\sum_{k=1}^{n-1}\left(\frac{1}{k}-\frac{1}{k+1}\right)\\ &= 2+\sum_{k=1}^{n-1}\frac{1}{k}-\sum_{k=2}^{n}\frac{1}{k} &&= 2+1-\frac{1}{n}\end{aligned}$$

Also konvergiert diese Folge. Der Grenzwert heißt Eulersche Zahl $e \in (\frac{5}{2}, 3)$.

(vii) $\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$.

Beweis: Wegen der Binomischen Formel gilt für alle $n \in \mathbb{N}$

$$\left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{n}\binom{n}{k}\frac{1}{n^k} = \sum_{k=0}^{n}\frac{n(n-1)\cdots(n-k+1)}{n^k}\cdot\frac{1}{k!} \le \sum_{k=0}^{n}\frac{1}{k!} \le e.$$

Für alle $k \in \mathbb{N}$ gilt $\quad \lim\limits_{n\to\infty}\binom{n}{k}\frac{1}{n^k} = \lim\limits_{n\to\infty}\frac{n(n-1)\cdots(n-k+1)}{n^k}\cdot\frac{1}{k!} = \frac{1}{k!}.$

Also gilt für alle $m \in \mathbb{N}$ $\quad \underline{\lim}\left(1+\frac{1}{n}\right)^n \ge \sum\limits_{k=0}^{m}\lim\limits_{n\to\infty}\binom{n}{k}\frac{1}{n^k} = \sum\limits_{k=0}^{m}\frac{1}{k!}.$

Weil $\sup\limits_{m\in\mathbb{N}}\left(\sum\limits_{k=0}^{m}\frac{1}{k!}\right) = e$ gilt, folgt dann $e \le \underline{\lim}\left(1+\frac{1}{n}\right)^n \le \overline{\lim}\left(1+\frac{1}{n}\right)^n \le e$. **q.e.d.**

(viii)* $\lim\limits_{n\to\infty}\frac{n}{\sqrt[n]{n!}} = e$.

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ die positive Folge $\frac{n^n}{n!}$. Dann gilt $\frac{a_{n+1}}{a_n} = \frac{(n+1)^{n+1}}{(n+1)n^n} = \left(1+\frac{1}{n}\right)^n$. Wegen (vii) folgt $\lim_{n\to\infty}\frac{a_{n+1}}{a_n} = e$ Dann folgt aus Satz 3.24

$$e = \underline{\lim}\frac{a_{n+1}}{a_n} \le \underline{\lim}\frac{n}{\sqrt[n]{n!}} \le \overline{\lim}\frac{n}{\sqrt[n]{n!}} \le \overline{\lim}\frac{a_{n+1}}{a_n} = e.$$

Daraus folgt $\quad \lim\limits_{n\to\infty}\frac{n}{\sqrt[n]{n!}} = e.$ **q.e.d.**

Kapitel 4

Reihen

4.1 Konvergenzkriterien

Definition 4.1. *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge. Dann heißt die Folge $(s_n)_{n\in\mathbb{N}}$ mit $s_n = \sum\limits_{j=1}^{n} a_j$ die zu $(a_n)_{n\in\mathbb{N}}$ gehörende Reihe. Diese Folge bezeichnen wir mit $(\sum a_n)_{n\in\mathbb{N}}$.*

Wenn die Reihe $(\sum a_n)_{n\in\mathbb{N}}$ konvergiert, dann bezeichnen wir den Grenzwert mit $\lim\limits_{n\to\infty} \sum\limits_{j=1}^{n} a_j = \sum\limits_{n=1}^{\infty} a_n = \sum\limits_{n\in\mathbb{N}} a_n$. Analog definieren wir $\sum\limits_{n=m}^{\infty} a_n = \lim\limits_{n\to\infty} \sum\limits_{j=m}^{n} a_j$.

Beispiel 4.2. **(i)** *Geometrische Reihe $(\sum q^n)_{n\in\mathbb{N}_0}$. Für $q \neq 1$ hatten wir berechnet: $\sum\limits_{j=0}^{n} q^j = \frac{1-q^{n+1}}{1-q}$. Dann folgt aus dem letzten Abschnitt*

$$\begin{array}{rll} \text{Für } |q| < 1 \text{ ist } \left(\sum q^n\right)_{n\in\mathbb{N}_0} & \text{konvergent:} & \displaystyle\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}. \\ \text{Für } |q| \geq 1 \text{ ist } \left(\sum q^n\right)_{n\in\mathbb{N}_0} & \text{divergent.} & \\ \text{Für reelles } q \geq 1 \text{ ist } \left(\sum q^n\right)_{n\in\mathbb{N}_0} & \text{divergent:} & \displaystyle\sum_{n=0}^{\infty} q^n = \infty. \end{array}$$

(ii) *Die ζ-Funktion ist definiert als $\zeta(s) = \sum\limits_{n=1}^{\infty} \frac{1}{n^s}$ ist der Grenzwert (wenn er existiert) der Reihe $\left(\sum \frac{1}{n^s}\right)_{n\in\mathbb{N}}$. Zunächst ist diese Reihe nur für alle rationalen Zahlen $s \in \mathbb{Q}$ definiert. Für $s = 1$ ist $\left(\sum \frac{1}{n^s}\right)_{n\in\mathbb{N}}$ divergent.*

Beweis: Diese Reihe ist monoton wachsend. Also ist nur die Frage, ob sie beschränkt ist oder nicht. Für alle $n \in \mathbb{N}$ gilt aber $\dfrac{1}{n+1} + \dfrac{1}{n+2} + \ldots + \dfrac{1}{n+n} \geq \dfrac{n}{2n} \geq \dfrac{1}{2}$. Also sind für alle $n \in \mathbb{N}_0$ jeweils die Summen $\displaystyle\sum_{j=1}^{2^n} \frac{1}{j} = 1 + \sum_{m=1}^{n} \sum_{j=2^{m-1}+1}^{2^m} \frac{1}{j} \geq 1 + \frac{n}{2}$. **q.e.d.**

(iii) Für alle $k \in \mathbb{N}$ ist die Reihe $\left(\sum \frac{1}{n(n+1)\cdots(n+k)}\right)_{n\in\mathbb{N}}$ konvergent und es gilt $\sum\limits_{n=1}^{\infty} \frac{1}{n(n+1)\cdots(n+k)} = \frac{1}{k\cdot k!}$

Beweis:

$$\begin{aligned}
\sum_{n=1}^{m} \frac{1}{n(n+1)\cdots(n+k)} &= \sum_{n=1}^{m} \frac{1}{k} \frac{n+k-n}{n(n+1)\cdots(n+k)} \\
&= \frac{1}{k}\left(\sum_{n=1}^{m} \frac{1}{n\cdots(n+k-1)} - \sum_{n=2}^{m+1} \frac{1}{n\cdots(n+k-1)}\right) \\
&= \frac{1}{k}\left(\frac{1}{k!} - \frac{1}{(m+1)\cdots(m+k)}\right)
\end{aligned}$$

Die Folge $\dfrac{1}{(m+1)\cdots(m+k)} \leq \dfrac{1}{m^k}$ konvergiert aber gegen Null. **q.e.d.**

Wenn wir das Cauchykriterium und das Monotonieprinzip auf Reihen anwenden, so erhalten wir:

Satz 4.3. *(Cauchykriterium für Reihen) Die Reihe $(\sum a_n)_{n\in\mathbb{N}}$ konvergiert genau dann, wenn es für jedes $\epsilon > 0$ ein N gibt, so dass $\left|\sum\limits_{j=n}^{m} a_j\right| < \epsilon$ für alle $m \geq n \geq N$ gilt.* **q.e.d.**

Satz 4.4. *(Monotonieprinzip für Reihen) Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge von nicht negativen Zahlen. Dann konvergiert die Reihe $(\sum a_n)_{n\in\mathbb{N}}$ genau dann, wenn sie beschränkt ist. Für den Grenzwert gilt dann $\displaystyle\sum_{n=1}^{\infty} a_n = \sup_{m\in\mathbb{N}} \sum_{n=1}^{m} a_n$.* **q.e.d.**

Definition 4.5. *(absolut konvergent) Die Reihe $(\sum a_n)_{n\in\mathbb{N}}$ heißt absolut konvergent, wenn die Reihe $(\sum |a_n|)_{n\in\mathbb{N}}$ konvergiert.*

Satz 4.6. *Jede absolut konvergente Reihe konvergiert. Und es gilt $\displaystyle\left|\sum_{n=1}^{\infty} a_n\right| \leq \sum_{n=1}^{\infty} |a_n|$.*

Beweis: Aus der Dreiecksungleichung folgt $\displaystyle\left|\sum_{j=n}^{m} a_j\right| \leq \sum_{j=n}^{m} |a_j|$ für alle $m \geq n$. Also ist die Reihe einer absolut konvergenten Reihe eine Cauchyfolge und konvergiert. Insbesondere gilt für alle $m \in \mathbb{N}$ auch $\left|\sum_{n=1}^{m} a_n\right| \leq \sum_{n=1}^{m} |a_n|$. Dann erfüllen auch die Grenzwerte $\left|\sum_{n=1}^{\infty} a_n\right| \leq \sum_{n=1}^{\infty} |a_n|$. **q.e.d.**

Aus dem Monotonie-Prinzip und Satz 3.5 folgt das

Satz 4.7. *(Majoranten Kriterium) Die Folgen von nicht negativen Zahlen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ erfüllen $b_n \leq a_n$ für alle $n \in \mathbb{N}$.*

(i) *Wenn außerdem $(\sum a_n)_{n\in\mathbb{N}}$ konvergiert, dann konvergiert auch $(\sum b_n)_{n\in\mathbb{N}}$*

(ii) *Wenn außerdem $(\sum b_n)_{n\in\mathbb{N}}$ divergiert, dann gilt $\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} a_n = \infty$.* **q.e.d.**

Beispiel 4.8. *Für alle $n, k \in \mathbb{N}$ ist $\frac{1}{(n+k)^{k+1}} \leq \frac{1}{n\cdots(n+k)}$. Also folgt aus der Konvergenz von $(\sum \frac{1}{n\cdots(n+k)})$ auch die Konvergenz von $(\sum \frac{1}{n^{k+1}})_{n\in\mathbb{N}}$.*

Satz 4.9. *(Wurzeltest) Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge und sei $\alpha = \overline{\lim} \sqrt[n]{|a_n|}$.*

(i) *Falls $\alpha < 1$, dann konvergiert $(\sum a_n)_{n\in\mathbb{N}}$ absolut.*

(ii) *Falls $\alpha > 1$, dann divergieren $(\sum a_n)_{n\in\mathbb{N}}$ und $(\sum |a_n|)_{n\in\mathbb{N}}$.*

Im Fall $\overline{\lim} \sqrt[n]{|a_n|} = 1$ kann die Reihe $(\sum a_n)_{n\in\mathbb{N}}$ sowohl konvergent als auch divergent sein. So ist z.B. $\lim_{n\to\infty} \sqrt[n]{\frac{1}{n}} = 1 = \lim_{n\to\infty} \sqrt[n]{\frac{1}{n^2}}$. Aber die Reihe $\left(\sum \frac{1}{n}\right)_{n\in\mathbb{N}}$ ist divergent, während die Reihe $\left(\sum \frac{1}{n^2}\right)_{n\in\mathbb{N}}$ konvergent ist.

Beweis: (i) Sei $\alpha < 1$. Dann gibt es für jedes $\alpha < \beta < 1$ aufgrund von Satz 3.20 (i) ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $\sqrt[n]{|a_n|} \leq \beta \iff |a_n| \leq \beta^n$. Weil aber $(\sum \beta^n)_{n\in\mathbb{N}}$ konvergiert, ist auch $(\sum a_n)_{n\in\mathbb{N}}$ absolut konvergent.

(ii) Sei $\alpha > 1$. Dann gibt es wieder aufgrund von Satz 3.20 (i) unendlich viele $\sqrt[n]{|a_n|} > 1$. Also kann die Folge $(|a_n|)_{n\in\mathbb{N}}$ nicht gegen Null konvergieren. Dann sind Reihen $(\sum a_n)_{n\in\mathbb{N}}$ und $(\sum |a_n|)_{n\in\mathbb{N}}$ keine Cauchyfolgen, also divergent. **q.e.d.**

Satz 4.10. *(Exponentialfunktion:) Für alle $x \in \mathbb{K}$ definieren wir*

$$\exp(x) : \mathbb{K} \to \mathbb{K}, \quad x \mapsto \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!}$$

Aufgrund des Beispiels (v) im letzten Kapitel ist $\lim_{n\to\infty} \sqrt[n]{n!} = \infty$. Also gilt auch

$$\lim_{n\to\infty} \sqrt[n]{\frac{|x^n|}{n!}} = \lim_{n\to\infty} \frac{|x|}{\sqrt[n]{n!}} = 0.$$

Deshalb konvergiert die Reihe $\left(\sum \frac{x^n}{n!}\right)_{n\in\mathbb{N}_0}$ absolut.

Satz 4.11. *(Quotiententest) Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge und sei $\alpha = \overline{\lim}\left|\frac{a_{n+1}}{a_n}\right|$.*

(i) *Falls $\alpha < 1$, dann konvergiert $(\sum a_n)_{n\in\mathbb{N}}$ absolut.*

(ii) *Falls es ein $N \in \mathbb{N}$ gibt, so dass alle $n \geq N$ auch $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ erfüllen, dann divergieren $(\sum a_n)_{n\in\mathbb{N}}$ und $(\sum |a_n|)_{n\in\mathbb{N}}$.*

Beweis: **(i)** Wegen $\overline{\lim}\sqrt[n]{|a_n|} \leq \overline{\lim}\left|\dfrac{a_n+1}{a_n}\right|$ (Satz 3.24) folgt (i) aus dem Wurzeltest.
(ii) Aus $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ folgt $|a_{n+1}| = \left|\dfrac{a_{n+1}}{a_n}\right| \cdot \left|\dfrac{a_n}{a_{n-1}}\right| \cdots \left|\dfrac{a_{N+1}}{a_N}\right| \cdot |a_N| \geq |a_N| > 0$. Also ist $|a_n|$ keine Nullfolge und die Reihen $(\sum a_n)_{n\in\mathbb{N}}$ und $(\sum |a_n|)_{n\in\mathbb{N}}$ divergent. **q.e.d.**

Satz 4.12.* *(Cauchy's Verdichtungssatz) Sei $(a_n)_{n\in\mathbb{N}}$ eine nicht negative monoton fallende Folge. Dann konvergiert die Reihe $(\sum a_n)_{n\in\mathbb{N}}$ genau dann, wenn die Reihe $(\sum 2^n a_{2^n})_{n\in\mathbb{N}}$ konvergiert.*

Beweis*: Für alle $n \in \mathbb{N}$ sei $s_n = \sum\limits_{j=1}^{n} a_j$ und $t_n = \sum\limits_{j=0}^{n} 2^j a_{2^j}$. Wegen der Monotonie von $(a_n)_{n\in\mathbb{N}}$ gilt für alle $j \in \mathbb{N}$:

$$a_{2^j} + a_{2^j+1} + \ldots + a_{2^{j+1}-1} \leq 2^j a_{2^j} \leq 2(a_{2^{j-1}+1} + a_{2^{j-1}+2} + \ldots + a_{2^j})$$

und für $j = 0$ gilt: $a_1 \leq a_1 \leq 2a_1$. Deshalb gilt für alle $n \in \mathbb{N}$

$$s_{2^{n+1}-1} \leq t_n \leq 2s_{2^n}.$$

Also ist die Folge $(t_n)_{n\in\mathbb{N}}$ genau dann nach oben beschränkt, wenn die Folge $(s_n)_{n\in\mathbb{N}}$ nach oben beschränkt ist. **q.e.d.**

Die Reihe $\left(\sum \frac{1}{n^s}\right)_{n\in\mathbb{N}}$ ist für $s \in \mathbb{Q} \setminus \{0\}$ genau dann konvergent, wenn die Reihe

$$\left(\sum \frac{2^n}{(2^n)^s}\right)_{n\in\mathbb{N}} = \left(\sum 2^{(1-s)\cdot n}\right)_{n\in\mathbb{N}}$$

konvergent ist, also genau dann, wenn $s > 1$. $\Longrightarrow$ Für alle $s \in \mathbb{Q}$ mit $s > 1$ ist $\zeta(s)$ wohl definiert.

Satz 4.13. *(Alternierende Reihe von Leibniz) Sei $(a_n)_{n\in\mathbb{N}_0}$ eine monoton fallende Nullfolge. Dann konvergiert $(\sum(-1)^n a_n)_{n\in\mathbb{N}_0}$*

Beweis: Wegen dem Monotonieprinzip sind alle a_n einer monoton fallenden Nullfolge $(a_n)_{n\in\mathbb{N}_0}$ nicht negativ. Sei für alle $n \in \mathbb{N}_0$ $s_n = \sum\limits_{m=0}^{n} (-1)^m a_m$. Dann folgt aus der Monotonie:

$$s_1 \leq s_3 \leq \ldots \leq s_{2n+1} \leq \ldots \leq s_{2n} \leq \ldots \leq s_2 \leq s_0.$$

Also ist die Folge $(s_{2n+1})_{n\in\mathbb{N}_0}$ monoton wachsend und beschränkt und $(s_{2n})_{n\in\mathbb{N}_0}$ monoton fallend und beschränkt. Dann konvergieren aber beide Folgen und es gilt

$$\lim_{n\to\infty} s_{2n} - \lim_{n\to\infty} s_{2n+1} = -\lim_{n\to\infty}(-1)^{2n+1}a_{2n+1} = \lim_{n\to\infty} a_{2n+1} = 0$$

Also konvergiert die Reihe $\left(\sum(-1)^n a_n\right)_{n\in\mathbb{N}}$ **q.e.d.**

Damit ist also die Reihe $\sum_{n=0}^{\infty}\frac{(-1)^n}{n+1} = -\sum_{n=1}^{\infty}\frac{(-1)^n}{n}$ konvergent, während sie nicht absolut konvergiert, weil $\sum_{n=1}^{\infty}\frac{1}{n} = \infty$.

4.2 Dezimalbruchdarstellung von reellen Zahlen

Als Ziffern wählen wir $Z = \{0,1,\ldots,9\}$ (bzw. $\{0,1,\ldots,p-1\}$). Sei $(z_n)_{n\in\mathbb{N}}$ eine Folge mit Werten in Z. Definiere die entsprechende Zahlenfolge $\left(\sum x_n\right)_{n\in\mathbb{N}}$ mit $x_n = \frac{z_n}{p^n}$ für alle $n \in \mathbb{N}$. Dann ist $\left(\sum x_n\right)_{n\in\mathbb{N}}$ nach dem Majorantenkriterium absolut konvergent, weil $\lim\limits_{n\to\infty}\sqrt[n]{\frac{p-1}{p^n}} = \left(\lim\limits_{n\to\infty}\sqrt[n]{p-1}\right)\cdot\frac{1}{p} = \frac{1}{p}$. Also definiert $x = \sum\limits_{n=1}^{\infty} x_n$ eine reelle Zahl. Sei jetzt M die Menge aller Folgen $M = \{(z_n)_{n\in\mathbb{N}} | (z_n)_{n\in\mathbb{N}}$ konvergiert nicht gegen $p-1\}$. Dann ist die Abbildung $M \to [0,1), (z_n)_{n\in\mathbb{N}} \mapsto \sum\limits_{n=0}^{\infty}\frac{z_n}{p^n}$ surjektiv und injektiv.

Bemerkung 4.14. *Wir hätten auch fordern können, dass $(z_n)_{n\in\mathbb{N}}$ nicht gegen Null konvergiert, so haben nämlich alle reellen Zahlen, deren Ziffernfolge gegen* 0 *konvergiert auch eine Dezimalbruchdarstellung, deren Ziffernfolge gegen* 9 *konvergiert, z.B.* $\frac{1}{2} = 0,5000\ldots = 0,4999\ldots$.

Surjektiv: Sei $x \in [0,1)$. Dann definieren wir $(z_n)_{n\in\mathbb{N}}$ induktiv, so dass für jedes $n \in \mathbb{N}$

$$\frac{z_n}{p^n} \le x - \sum_{m=1}^{n-1}\frac{z_m}{p^m} < \frac{z_n+1}{p^n} \quad\Longleftrightarrow\quad x \in \left[\frac{z_n}{p^n} + \sum_{m=1}^{n-1}\frac{z_m}{p^m}, \frac{z_n+1}{p^n} + \sum_{m=1}^{n-1}\frac{z_m}{p^m}\right)$$

gilt. Dann folgt aber auch $0 \le x - \sum\limits_{m=1}^{n}\frac{z_m}{p^m} < \frac{1}{p^n}$. Also gilt $\sum\limits_{n=1}^{\infty}\frac{z_n}{p^n} = x$. Weil aber für alle $n \in \mathbb{N}$ gilt $\sum\limits_{m=n+1}^{\infty}\frac{p-1}{p^m} = \frac{p-1}{p^{n+1}\left(1-\frac{1}{p}\right)} = \frac{1}{p^n}$, gilt auch $\lim\limits_{n\to\infty} z_n \neq p-1$.

Injektiv: Seien $(z_n)_{n\in\mathbb{N}}$ und $(w_n)_{n\in\mathbb{N}}$ zwei Ziffernfolgen, mit $\sum\limits_{n=1}^{\infty}\frac{z_n}{p^n} = \sum\limits_{n=1}^{\infty}\frac{w_n}{p^n}$. Sei also $n \in \mathbb{N}$ der kleinste Index, so dass $z_n \neq w_n$. Weil aber gilt

$$\sum_{m=n+1}^{\infty}\frac{|z_m - w_m|}{p^m} \le \sum_{m=n+1}^{\infty}\frac{p-1}{p^m} = \frac{p-1}{p^{n+1}}\sum_{m=0}^{\infty}\frac{1}{p^m} = \frac{p-1}{p^{n+1}}\frac{1}{1-\frac{1}{p}} = \frac{1}{p^n},$$

muss auch $|z_n - w_n| \leq 1$ gelten. Sei also $z_n = w_n + 1$, dann muss für alle $m > n$ gelten $w_m - z_m = p - 1 \Longrightarrow z_m = 0$ und $w_m = p - 1 \Longrightarrow \lim_{m\to\infty} w_m = p - 1$. Also gehört $(w_n)_{n\in\mathbb{N}}$ nicht zu M. **q.e.d.**

4.3 Addition, Multiplikation, Umordnung

Aus dem Satz 3.5 folgt

Satz 4.15. *(Rechenregeln für Reihen) Die Reihen $(\sum a_n)_{n\in\mathbb{N}}$ und $(\sum b_n)_{n\in\mathbb{N}}$ seien konvergent, dann konvergieren auch die Reihen*

$$\left(\sum(a_n + b_n)\right)_{n\in\mathbb{N}} \quad \textit{und} \qquad \left(\sum \lambda a_n\right)_{n\in\mathbb{N}} \; \textit{für alle } \lambda \in \mathbb{K}. \qquad \textbf{q.e.d.}$$

Definition 4.16. *Gegeben seien die Reihen $(\sum a_n)_{n\in\mathbb{N}_0}$ und $(\sum b_n)_{n\in\mathbb{N}_0}$. Dann heißt die Reihe $(\sum c_n)_{n\in\mathbb{N}_0}$ mit $c_n = \sum\limits_{k=0}^{n} a_k b_{n-k}$ für alle $n \in \mathbb{N}_0$ das (Cauchy-)Produkt der beiden Reihen.*

Diese Definition kommt von den Potenzreihen, die wir später kennenlernen werden. Das Produkt der beiden Potenzreihen $(\sum a_n x^n)_{n\in\mathbb{N}_0}$ und $(\sum b_n x^n)_{n\in\mathbb{N}_0}$ ist dann nämlich die Potenzreihe $(\sum c_n x^n)_{n\in\mathbb{N}_0}$, d.h. wir haben alle Summanden des Produktes mit gleichen Potenzen zusammengefasst.

Satz 4.17. *(Konvergenz des Produktes) Wenn die Reihe $(\sum a_n)_{n\in\mathbb{N}_0}$ absolut konvergiert und die Reihe $(\sum b_n)_{n\in\mathbb{N}_0}$ konvergiert, dann konvergiert auch das Produkt $(\sum c_n)_{n\in\mathbb{N}_0}$ der beiden Reihen. Wenn auch $(\sum b_n)_{n\in\mathbb{N}_0}$ absolut konvergiert, dann konvergiert auch $(\sum c_n)_{n\in\mathbb{N}_0}$ absolut.*

Beweis: Für alle $n \in \mathbb{N}_0$ sei $c_n = \sum\limits_{k=0}^{n} a_k b_{n-k}$, $A_n = \sum\limits_{k=0}^{n} a_k$, $B_n = \sum\limits_{k=0}^{n} b_k$ und $C_n = \sum\limits_{k=0}^{n} c_k$. Es gilt

$$\begin{aligned} C_n &= a_0 b_0 + (a_0 b_1 + a_1 b_0) + \ldots + (a_0 b_n + \ldots + a_n b_0) \\ &= a_0 B_n + a_1 B_{n-1} + \ldots + a_n B_0 \\ &= a_0(B - \beta_n) + a_1(B - \beta_{n-1}) + \ldots + a_n(B - \beta_0) \end{aligned}$$

Hierbei ist $B = \lim_{n\to\infty} B_n = \sum\limits_{n=0}^{\infty} b_n$ und $\beta_n = B - B_n = \sum\limits_{k=n+1}^{\infty} b_k$. Daraus ergibt sich

$$C_n = A_n \cdot B - (a_0\beta_n + a_1\beta_{n-1} + \ldots + a_n\beta_0).$$

Also genügt es zu zeigen, dass $a_0\beta_n + \ldots + a_n\beta_0$ im Grenzwert $n \to \infty$ gegen Null konvergiert. Aufgrund der Voraussetzungen gibt es positive Zahlen $\alpha, \beta > 0$, so dass für alle $n \in \mathbb{N}_0$

$$|\beta_n| \le \beta \quad \text{und} \quad \sum_{k=0}^{n} |a_k| \le \alpha$$

gilt, und für alle $\epsilon > 0$ natürliche Zahlen N, M, so dass $|\beta_n| < \frac{\epsilon}{2\alpha}$ für alle $n \ge N$ gilt und $|a_n| < \frac{\epsilon}{2N\beta}$ für alle $n \ge M$. Dann gilt für alle $n \ge N + M - 1$:

$$\begin{aligned} |a_0\beta_n + \ldots + a_n\beta_0| &\le |\beta_0 a_n + \ldots + \beta_{N-1}a_{n-N+1}| + |\beta_N a_{n-N} + \ldots + \beta_n a_0| \\ &< N\beta \cdot \frac{\epsilon}{2N\beta} + \frac{\epsilon}{2\alpha}\sum_{k=0}^{n-N} |a_k| \le \epsilon. \end{aligned}$$

Also konvergiert das Produkt der Reihen $(\sum a_n)_{n\in\mathbb{N}_0}$ und $(\sum b_n)_{n\in\mathbb{N}_0}$ wenn beide konvergieren und eine absolut konvergiert. Wenn beide Reihen absolut konvergieren, dann konvergiert auch das Produkt der Reihen $(\sum |a_n|)_{n\in\mathbb{N}_0}$ und $(\sum |b_n|)_{n\in\mathbb{N}_0}$ und ist eine Majorante von $(\sum |c_n|)_{n\in\mathbb{N}_0}$. Also ist dann auch das Produkt absolut konvergent.**q.e.d.**

Beispiel 4.18. *Das Quadrat der Reihe* $\left(\sum \frac{(-1)^n}{\sqrt{n+1}}\right)_{n\in\mathbb{N}_0}$ *ist nicht konvergent, obwohl die Reihe als Beispiel einer alternierenden Reihe nach Leibniz konvergiert, aber nicht absolut konvergiert. Die Koeffizienten des Quadrates sind gegeben durch:*

$$\sum_{k=0}^{n} \frac{(-1)^k(-1)^{n-k}}{\sqrt{k+1}\sqrt{n-k+1}} = (-1)^n \sum_{k=0}^{n} \frac{1}{\sqrt{k+1}\sqrt{n-k+1}}$$

Es gilt aber $\displaystyle\sum_{k=0}^{n} \frac{1}{\sqrt{k+1}\sqrt{n-k+1}} \ge \sum_{k=0}^{n} \frac{1}{\sqrt{n+1}\sqrt{n+1}} = 1$. *Also ist das Quadrat der Reihe keine Cauchyfolge.* **q.e.d.**

Satz 4.19. *(Eigenschaften der Exponentialfunktion)*

(i) *für alle* $x, y \in \mathbb{K}$ *gilt* $\exp(x+y) = \exp(x)\exp(y)$.

(ii) *Für alle* $x \in \mathbb{K}$ *ist* $\exp(x) \neq 0$ *und für alle* $x \in \mathbb{R}$ *sogar* $\exp(x) > 0$.

(iii) *Für alle* $x \in \mathbb{R}$ *und alle* $r \in \mathbb{Q}$ *ist* $\exp(rx) = (\exp(x)^r)$.

(iv) *Für alle* $x \in \mathbb{C}$ *ist* $\exp(\bar{x}) = \overline{\exp(x)}$.

(v) *Für alle* $x \in \mathbb{R}$ *ist* $|\exp(\imath x)| = 1$.

Die Zahl $\exp(1)$ *wird Eulersche Zahl genannt und mit* e *bezeichnet. Wegen (iii) gilt dann für alle* $r \in \mathbb{Q}$: $\exp(r) = \exp(r \cdot 1) = e^r$. *Deshalb schreiben wir auch* e^x *für* $\exp(x)$.

Beweis: (i) Wir hatten schon gesehen, dass die Reihe $\exp(x) = \sum\limits_{n=0}^{\infty} \frac{x^n}{n!}$ für alle $x \in \mathbb{K}$ absolut konvergiert. Dann ist das Produkt von $\exp(x) \cdot \exp(y) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{x^k y^{n-k}}{k!(n-k)!}$. Wegen der Binomischen Formel gilt

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

Dann folgt $\exp(x)\exp(y) = \sum\limits_{n=0}^{\infty} \frac{(x+y)^n}{n!} = \exp(x+y)$.
(ii) Wegen (i) gilt $\exp(x)\exp(-x) = 1$. Also besitzt $\exp(x)$ für alle $x \in \mathbb{K}$ ein Inverses und ist ungleich Null. Wegen (i) gilt für alle $x \in \mathbb{R}$ $\exp(x) = \left(\exp\left(\frac{x}{2}\right)\right)^2 > 0$.
(iii) Offenbar ist für alle $n \in \mathbb{N}, m \in \mathbb{Z}$ $(\exp(\frac{x \cdot m}{n}))^n = \exp(x \cdot m) = (\exp(x))^m$. Also gilt auch wegen (ii) $\exp\left(\frac{xm}{n}\right) = (\exp(x))^{\frac{m}{n}}$.
(iv) $\exp(\bar{x}) = \sum\limits_{n=0}^{\infty} \frac{\bar{x}^n}{n!} = \overline{\sum\limits_{n=0}^{\infty} \frac{x^n}{n!}} = \overline{\exp(x)}$.
(v) Für $x \in \mathbb{R}$ gilt $\exp(\imath x)\overline{\exp(\imath x)} = \exp(\imath x)\exp(-\imath x) = 1$. **q.e.d.**

Wir können jetzt für jede Zahl $y > 0$, für die es ein $x \in \mathbb{R}$ gibt, so dass $y = \exp(x)$ gilt, und für jedes $z \in \mathbb{R}$ die Zahl $y^z = \exp(zx)$ definieren. Wir werden später sehen, dass wir so für alle $y > 0$ und alle $z \in \mathbb{R}$ y^z definieren können.

Definition 4.20. *(Umordnen von Reihen) Sei $\tau : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung von den natürlichen Zahlen auf sich selber. Dann heißt die Reihe $\left(\sum a_{\tau(n)}\right)_{n\in\mathbb{N}}$ eine Umordnung der Reihe $(\sum a_n)_{n\in\mathbb{N}}$.*

Analog könnten wir auch die Umordnung von Folgen bilden. Letztere sind aber weniger interessant, weil jede Umordnung einer konvergenten Folge wieder gegen den gleichen Grenzwert konvergiert (Übungsaufgabe). Dagegen gilt dies bei Reihen nicht.

Satz 4.21. *Konvergiert eine Reihe $(\sum a_n)_{n\in\mathbb{N}}$ absolut, so konvergiert auch jede Umordnung $\left(\sum a_{\tau(n)}\right)_{n\in\mathbb{N}}$ absolut. In diesem Fall gilt dann $\sum\limits_{n=1}^{\infty} a_n = \sum\limits_{n=1}^{\infty} a_{\tau(n)}$.*

Beweis: Sei also τ eine gegebene bijektive Abbildung $\tau : \mathbb{N} \to \mathbb{N}$. Wenn $(\sum a_n)_{n\in\mathbb{N}}$ absolut konvergent, dann gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $N \leq n \leq m$ gilt: $\sum\limits_{k=n}^{m} |a_k| < \epsilon$. Dann gibt es auch ein $M = \max \tau^{-1}[\{1, \ldots, N\}] \in \mathbb{N}$, so dass für alle $m \geq M$ gilt $\tau(m) \geq N$. Dann folgt für alle $m \geq n \geq M$

$$\sum_{k=n}^{m} \left|a_{\tau(k)}\right| \leq \sum_{k=\min \tau[\{n,n+1,\ldots,m\}]}^{\max \tau[\{n,n+1,\ldots,m\}]} |a_k| < \epsilon.$$

Also ist $\left(\sum |a_{\tau(n)}|\right)_{n\in\mathbb{N}}$ eine Cauchyfolge und die Umordnung konvergiert absolut. Dann konvergiert auch $\left(\sum a_{\tau(n)}\right)_{n\in\mathbb{N}}$.

Mit denselben N und M in Abhängigkeit von $\epsilon > 0$ gilt für alle $n \geq N$ and $m \geq M$

$$\begin{aligned}\left|\sum_{k=1}^{m} a_{\tau(k)} - \sum_{k=1}^{n} a_k\right| &\leq \sum_{k\in\tau[\{1,\ldots,m\}]\setminus\{1,\ldots,N\}} |a_k| + \sum_{k=N+1}^{n} |a_k| \\ &\leq \sum_{k=\min\tau[\{1,\ldots,m\}]\setminus\{1,\ldots,N\}}^{k=\max\tau[\{1,\ldots,m\}]\setminus\{1,\ldots,N\}} |a_k| + \sum_{k=N+1}^{n} |a_k| < 2\epsilon.\end{aligned}$$

Also konvergiert $(\sum a_{\tau(n)})_{n\in\mathbb{N}}$ gegen den gleichen Grenzwert wie $(\sum a_n)_{n\in\mathbb{N}}$. **q.e.d.**

Satz 4.22. *(Riemann) Sei $(\sum a_n)_{n\in\mathbb{N}}$ eine konvergente reelle Reihe, die nicht absolut konvergiert, und $\alpha \leq \beta$ zwei reelle Zahlen. Dann gibt es eine Umordnung $\left(\sum a_{\tau(n)}\right)_{n\in\mathbb{N}}$, die als Reihe beschränkt ist und für die α der Limes inferior der Reihe ist und β der Limes superior. Wenn $\alpha \neq \beta$ konvergiert die Reihe also nicht.*

Beweisskizze: Weil $(\sum a_n)_{n\in\mathbb{N}}$ konvergiert, ist die Folge $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge. Wir betrachten im folgenden die beiden Teilfolgen aller nichtnegativen Elemente $(b_n)_{n\in\mathbb{N}}$ und aller negativen Elemente $(c_n)_{n\in\mathbb{N}}$.

1. Schritt: Weil $(\sum a_n)_{n\in\mathbb{N}}$ konvergiert, aber nicht absolut konvergiert, divergieren $(\sum b_n)_{n\in\mathbb{N}}$ und $(\sum c_n)_{n\in\mathbb{N}}$ und es gilt $\sum_{n=0}^{\infty} b_n = \infty$ und $\sum_{n=0}^{\infty} c_n = -\infty$.

2. Schritt: Wir setzen die umgeordnete Folge abwechselnd jeweils der Reihe nach aus den Folgen $(b_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ zusammen. Immer wenn die Summe aller bisherigen Folgenglieder größer ist als β, dann fahren wir fort mit Folgengliedern aus c_n, und wenn die Summe aller bisherigen Folgenglieder kleiner ist als α, dann fahren wir fort mit Folgengliedern aus b_n. Wenn $0 \in [\alpha, \beta]$ starten wir mit Folgengliedern aus b_n.

3. Schritt: Für jedes $\epsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass alle Summen $\sum_{k=1}^{n} a_{\tau(k)}$ für alle $n \geq N$ in $(\alpha - \epsilon, \beta + \epsilon)$ liegen. Die Reihe $\left(\sum a_{\tau(n)}\right)_{n\in\mathbb{N}}$ dieser Umordnung hat als Limes inferior α und als Limes superior β. Die Menge der Häufungspunkte dieser Reihe ist sogar gleich $[\alpha, \beta]$. **q.e.d.**

Definition 4.23. *Sei $(a_n)_{\in\mathbb{N}_0}$ eine Folge, dann heißt $(\sum a_n x^n)_{n\in\mathbb{N}_0}$ die entsprechende Potenzreihe mit Koeffizienten $(a_n)_{n\in\mathbb{N}_0}$.*

Aus dem Wurzeltest folgt sofort

Satz 4.24. *(Konvergenzradius von Potenzreihen) Seien $(a_n)_{n\in\mathbb{N}_0}$ die Koeffizienten der Potenzreihe $(\sum a_n x^n)_{n\in\mathbb{N}_0}$ und sei $\alpha = \overline{\lim}\sqrt[n]{|a_n|}$ und $R = \frac{1}{\alpha}$ ($R = 0$ für $\alpha = \infty$ und $R = \infty$ für $\alpha = 0$).*

(i) *Für $|x| < R$ konvergiert $(\sum a_n x^n)_{n\in\mathbb{N}_0}$ absolut.*

(ii) *Für $|x| > R$ divergiert $(\sum a_n x^n)_{n\in\mathbb{N}_0}$.*

Wenn $\alpha < \infty$ definiert folgende Reihe also eine Potenzreihenfunktion

$$f: \quad \{x \in \mathbb{K} \mid |x| < R\} \to \mathbb{K}, \quad x \mapsto \sum_{n=0}^{\infty} a_n x^n. \qquad \textbf{q.e.d.}$$

Beispiel 4.25. **(i)** $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, $\alpha = \overline{\lim} \sqrt[n]{\frac{1}{n!}} = \frac{1}{\lim_{n\to\infty} \sqrt[n]{n!}} = 0 \Longrightarrow R = \infty$.

(ii) $\left(\sum x^n\right)_{n\in\mathbb{N}_0}$ *also* $\alpha = \overline{\lim} \sqrt[n]{1} = 1 \Longrightarrow R = 1$. *Für* $|x| < 1$ *gilt* $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$.

(iii) $\left(\sum \frac{x^n}{n}\right)_{n\in\mathbb{N}}$ *also* $\alpha = \overline{\lim} \sqrt[n]{\frac{1}{n}} = \frac{1}{\lim_{n\to\infty} \sqrt[n]{n}} = 1 \Longrightarrow R = 1$.

Für $|x| < 1$ *ist die Potenzreihe also konvergent, aber für* $x = 1$ *divergent und für* $x = -1$ *konvergent (alternierende Reihe von Leibniz).*

(iv) $\left(\sum \frac{x^n}{n^2}\right)_{n\in\mathbb{N}}$ *also* $\alpha = \overline{\lim} \sqrt[n]{\frac{1}{n^2}} = \left(\frac{1}{\lim_{n\to\infty} \sqrt[n]{n}}\right)^2 = 1 \Longrightarrow R = 1$.

Für $|x| \leq 1$ *also konvergent und für* $|x| > 1$ *divergent.*

Satz 4.26. *(Eigenschaften von Potenzreihenfunktionen)*

(i) *Seien* $\left(\sum a_n x^n\right)_{n\in\mathbb{N}_0}$ *und* $\left(\sum b_n x^n\right)_{n\in\mathbb{N}_0}$ *Potenzreihen mit Konvergenzradius* R_1 *bzw.* R_2. *Dann konvergieren für* $|x| < \min\{R_1, R_2\}$ *die Summe* $\left(\sum (a_n + b_n) x^n\right)_{n\in\mathbb{N}_0}$ *und das Cauchy-Produkt* $\left(\sum c_n x^n\right)_{n\in\mathbb{N}}$ *der beiden Potenzreihen und es gilt*

$$\sum_{n=0}^{\infty} (a_n + b_n) x^n = \sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n \text{ und } \sum_{n=0}^{\infty} c_n x^n = \left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right).$$

(ii) *Für* $r < R_1$ *gibt es ein* $M(r) \in \mathbb{R}^+$, *so dass* $\left|\sum_{n=0}^{\infty} a_n x^n\right| \leq M(r)$ *für alle* $|x| \leq r$ *gilt.*

(iii) *Für* $r < R_1$ *und für alle* $\epsilon > 0$ *gibt es ein* $N \in \mathbb{N}$, *so dass für alle* $|x| \leq r$ *gilt*

$$\left|\sum_{n=0}^{\infty} a_n x^n - \sum_{n=0}^{N} a_n x^n\right| < \epsilon.$$

(iv) *Für* $r < R_1$ *gibt es ein* $L(r) \in \mathbb{R}^+$, *so dass für alle* x, y *mit* $|x| \leq r$ *und* $|y| \leq r$ *gilt*

$$\left|\sum_{n=0}^{\infty} a_n x^n - \sum_{n=0}^{\infty} a_n y^n\right| \leq L(r)|x - y|.$$

Beweis: **(i)** Folgt aus den Rechenregeln für Reihen und der Konvergenz des Cauchy Produktes.

(ii) Für $|x| \leq r$ gilt $\left|\sum\limits_{n=0}^{\infty} a_n x^n\right| \leq \sum\limits_{n=0}^{\infty} |a_n| r^n = M(r) < \infty$.

(iii) Weil die Reihe $\sum_{n=0}^{\infty} |a_n| r^n$ absolut konvergiert gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $\sum_{n=N}^{\infty} |a_n| r^n < \epsilon$ gilt. Dann folgt für $|x| \leq r$

$$\left|\sum_{n=0}^{\infty} a_n x^n - \sum_{n=0}^{N-1} a_n x^n\right| \leq \sum_{n=N}^{\infty} |a_n| r^n < \epsilon.$$

(iv) $(x^n - y^n) = (x-y)(x^{n-1} + x^{n-2}y + \ldots + xy^{n-2} + y^{n-1})$. Für $|x| \leq r$ und $|y| \leq r$ folgt also $|x^n - y^n| \leq |x-y| n r^{n-1}$. Weil aber gilt

$$\overline{\lim} \sqrt[n]{n|a_n|} = \overline{\lim} \sqrt[n]{n} \cdot \sqrt[n]{|a_n|} = \left(\lim_{n\to\infty} \sqrt[n]{n}\right) \overline{\lim} \sqrt[n]{|a_n|} = \overline{\lim} \sqrt[n]{|a_n|},$$

haben die Reihen $(\sum a_n x^n)_{n\in\mathbb{N}_0}$ und $(\sum n \cdot a_n x^{n-1})_{n\in\mathbb{N}_0}$ den gleichen Konvergenzradius. Also gilt für $|x| \leq r$ und $|y| \leq r$

$$\left|\sum_{n=0}^{\infty} a_n x^n - \sum_{n=0}^{\infty} a_n y^n\right| \leq \sum_{n=0}^{\infty} |a_n| \cdot |x^n - y^n| \leq |x-y| \sum_{n=1}^{\infty} |a_n| \cdot n r^{n-1}.$$

Wähle also $L(r) = \sum\limits_{n=1}^{\infty} n|a_n| r^{n-1} < \infty$. **q.e.d.**

Satz 4.27. *(Identitässatz für Potenzreihenfunktionen)*

(i) *Sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R > 0$, die nicht identisch verschwindet, dann gibt es ein $0 < r < R$, so dass die Potenzreihenfunktion in $\{x \in \mathbb{K} \mid |x| < r\}$ höchstens endlich viele Nullstellen $x_1, \ldots, x_N$ hat.*

(ii) *Sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihenfunktion mit Konvergenzradius $R > 0$. Für alle x_0 mit $|x_0| < R$ und alle $n \in \mathbb{N}_0$ ist dann $b_n = \sum_{k=0}^{\infty} \binom{n+k}{n} a_{n+k} x_0^k < \infty$. Außerdem ist der Konvergenzradius der Potenzreihenfunktion $\sum_{n=0}^{\infty} b_n x^n$ nicht kleiner als $R - |x_0|$ und für alle $|x| < R - |x|_0$ gilt auch $\sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} a_n (x + x_0)^n$.*

(iii) *Seien $\sum_{n=0}^{\infty} a_n x^n$ und $\sum_{n=0}^{\infty} b_n x^n$ zwei Potenzreihenfunktionen, deren Konvergenzradien größer sind als $r > 0$. Falls $\{x \in \mathbb{K} \mid |x| \leq r\}$ unendliche viele verschiedene Elemente enthält, an denen die beiden Potenzreihenfunktionen übereinstimmen, dann gilt $a_n = b_n$ für alle $n \in \mathbb{N}_0$, d.h. sie stimmen als Potenzreihen überein.*

Beweis: **(i)** Sei $N \in \mathbb{N}_0$ der kleinste Index, so dass $a_N \neq 0$. Wenn alle anderen Koeffizienten $(a_n)_{n>N}$ verschwinden, hat die Potenzreihe nur Nullstellen bei $x = 0$. Andernfalls gilt für ein $0 < r < R$ und alle $|x| \leq r$

$$\left|\sum_{n=0}^{\infty} a_n x^n - a_N x^N\right| \leq \sum_{n=N+1}^{\infty} |a_n| \cdot |x|^n \leq |x|^{N+1} \left(\sum_{n=0}^{\infty} |a_{n+N+1}| r^n\right).$$

Also gilt für alle Nullstellen x_m der Potenzreihenfunktionen

$$|a_N| \cdot |x_m|^N \leq |x_m|^{N+1} \left(\sum_{n=0}^{\infty} |a_{n+N+1}| r^n\right).$$

Wenn $|x_m| \neq 0$ ist folgt daraus $0 < |a_N| \leq |x_m| (\sum_{n=0}^{\infty} |a_{n+N+1}| r^n)$. Also hat die Potenzreihenfunktion keine Nullstelle auf

$$\left\{ x \in \mathbb{K} \;\middle|\; 0 < |x| < \min\left\{ r, |a_N| \left(\sum_{n=0}^{\infty} |a_{n+N+1}| r^n\right)^{-1} \right\} \right\}.$$

(ii) Für $x_0 = 0$ trivial. Sei $0 < |x_0| = r < R$. Dann ist für alle $0 < s < R - r$ die Reihe

$$\sum_{n=0}^{\infty} a_n (r+s)^n = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_n \binom{n}{k} r^k s^{n-k} < \infty$$

absolut summierbar. Dann gilt für alle $m \in \mathbb{N}_0$ auch:

$$s^m \sum_{k=0}^{\infty} |a_{m+k}| \binom{m+k}{k} r^k \leq \sum_{n=0}^{\infty} \sum_{k=0}^{n} |a_n| \binom{n}{k} r^{n-k} s^k < \infty.$$

Also konvergiert $b_m = \sum\limits_{k=0}^{\infty} a_{m+k} \binom{m+k}{k} x_0^m$ für alle $m \in \mathbb{N}$. Und es gilt

$$\sum_{m=0}^{M} s^m \sum_{k=0}^{\infty} |a_{m+k}| \binom{m+k}{k} r^k \leq \sum_{n=0}^{\infty} \sum_{k=0}^{n} |a_n| \binom{n}{k} r^{n-k} s^k < \infty$$

für alle $M \in \mathbb{N}_0$. Also ist auch $\displaystyle\sum_{m=0}^{\infty} |b_m| s^m \leq \sum_{n=0}^{\infty} |a_n| (r+s)^n < \infty.$

Also ist $\sum_{m=0}^{\infty} b_m x^m$ für alle $|x| < R - r$ konvergent, und der Konvergenzradius nicht kleiner als $R - r = R - |x_0|$. Für $|x| < R - |x_0|$ und alle $M, K \in \mathbb{N}_0$ gilt dann

$$\left|\sum_{m=0}^{M} x^m \sum_{k=0}^{K} a_{m+k} \binom{m+k}{k} x_0^k - \sum_{n=0}^{M+K} a_n (x+x_0)^n\right| \leq \sum_{n=\min\{M,K\}}^{M+K} |a_n| (|x| + |x_0|)^n.$$

Also folgt im Grenzwert $K \to \infty$

$$\left|\sum_{m=0}^{M} b_m x^m - \sum_{n=0}^{\infty} a_n (x+x_0)^n\right| \leq \sum_{n=M}^{\infty} |a_n|(|x|+|x_0|)^n < \infty.$$

und im Grenzwert $M \to \infty$ auch $\displaystyle\sum_{n=0}^{\infty} b_m x^m - \sum_{n=0}^{\infty} a_n (x+x_0)^n = 0.$

(iii) Sei $(x_m)_{m\in\mathbb{N}}$ eine Folge von paarweise verschiedenen Nullstellen der Potenzreihenfunktion $x \mapsto \sum_{n=0}^{\infty}(a_n - b_n)x^n$ in $\{x \in \mathbb{K} \mid |x| \leq r\}$. Dann hat die Folge $(x_m)_{m\in\mathbb{N}}$ einen Häufungspunkt x_0 mit $|x_0| \leq r$, und in jeder ϵ-Umgebung von x_0 gibt es unendlich viele verschiedene Folgenglieder. Sei R das Minimum der Konvergenzradien von $(\sum a_n)_{n\in\mathbb{N}_0}$ und $(\sum b_n)_{n\in\mathbb{N}_0}$. Dann hat aufgrund der Voraussetzung und wegen (ii) die Potenzreihe $\sum_{n=0}^{\infty}(a_n - b_n)(y + x_0)^n$, als Potenzreihe in y mindestens den Konvergenzradius $R - r > 0$, und für alle $0 < \epsilon \leq R - r$ unendlich viele Nullstellen auf $\{y \in \mathbb{C} \mid |y| < \epsilon\}$. Also verschwindet wegen (i) diese Potenzreihe in y identisch. Dann stimmen wegen (ii) die beiden Potenzreihen $\sum_{n=0}^{\infty} a_n x^n$ und $\sum_{n=0}^{\infty} b_n x^n$ auf dem Gebiet $\{x \in \mathbb{K} \mid |x - x_0| < R - r\}$ überein. Also gibt es eine Folge $(\tilde{x}_m)_{m\in\mathbb{N}}$ von paarweise verschiedenen Nullstellen von $x \mapsto \sum_{n=0}^{\infty}(a_n - b_n)x^n$, die gegen ein $\tilde{x}_0$ mit $|\tilde{x}_0| \leq \max\{r-(R-r), 0\} < r$ konvergiert. Dabei ist $R-|\tilde{x}_0| > R-r$. Sei $\frac{R}{R-r} < N \in \mathbb{N}$. Indem wir die beiden Potenzreihe immerwieder an einer Stelle mit minimalem Radius in dem Bereich entwickeln, in dem wir schon die Gleichheit beider Potenzreihen gezeigt haben, erhalten wir also nach höchstens N Schritten, dass beide Potenzreihen auf einer Nullfoge übereinstimmen. Wegen (i) sind dann die beiden Potenzreihen $(\sum a_n x^n)_{n\in\mathbb{N}_0}$ und $(\sum b_n x^n)_{n\in\mathbb{N}_0}$ identisch. **q.e.d.**

4.4 Sinus und Cosinus

Definition 4.28. *Für alle $x \in \mathbb{K}$ sei*

$$\cos(x) = \frac{\exp(\imath x) + \exp(-\imath x)}{2} \qquad \sin(x) = \frac{\exp(\imath x) - \exp(-\imath x)}{2i}$$

Also gilt für reelle x $\quad \cos(x) = \Re(\exp(\imath x)) \quad \sin(x) = \Im(\exp(\imath x))$
und für alle $x \in \mathbb{K}$ die **Eulersche Formel**: $\quad \exp(\imath x) = \cos(x) + \imath \sin(x)$.
Außerdem gilt für alle $x \in \mathbb{K}$

$$\cos^2(x) + \sin^2(x) = \frac{\exp(2\imath x) + 2 + \exp(-2\imath x)}{4} - \frac{\exp(2\imath x) - 2 + \exp(-2\imath x)}{4} = 1,$$

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

Satz 4.29. *(Additionstheorem) Für alle $x, y \in \mathbb{K}$ gilt:*

$$\begin{aligned} \cos(x+y) &= \cos(x)\cos(y) - \sin(x)\sin(y) \\ \sin(x+y) &= \sin(x)\cos(y) + \cos(x)\sin(y) \end{aligned}$$

Beweis: $\exp(\imath(x+y)) = \exp(\imath x)\exp(\imath y)$.

$$\begin{aligned}\cos(x+y)+\imath\sin(x+y) &= (\cos(x)+\imath\sin(x))(\cos(y)+\imath\sin(y))\\ &= \cos(x)\cos(y)-\sin(x)\sin(y)+\imath(\sin(x)\cos(y)+\cos(x)\sin(y)).\end{aligned}$$

Ersetzen wir x und y durch $-x$ und $-y$ und benutzen $\cos(-x)=\cos(x)$ und $\sin(-x) = -\sin(x)$, dann erhalten wir

$$\cos(x+y)-\imath\sin(x+y) = \cos(x)\cos(y)-\sin(x)\sin y-\imath(\sin(x)\cos(y)+\cos(x)\sin(y)).$$

Die Summe und die Differenz dieser beiden Gleichungen ergibt

$$\begin{aligned}\cos(x+y) &= \cos(x)\cos(y)-\sin(x)\sin(y)\\ \sin(x+y) &= \sin(x)\cos(y)+\cos(x)\sin(y).\end{aligned}$$

q.e.d.

Durch Einsetzen von (x,y) und $(x,-y)$ erhalten wir für alle $x,y\in\mathbb{K}$

$$\begin{aligned}\cos(x+y)+\cos(x-y) &= 2\cos(x)\cos(y)\\ \cos(x-y)-\cos(x+y) &= 2\sin(x)\sin(y)\\ \sin(x+y)+\sin(x-y) &= 2\sin(x)\cos(y)\end{aligned}$$

Satz 4.30. *(Potenzreihen von Sinus und Cosinus)*

$$\cos(x)=\sum_{k=0}^{\infty}(-1)^k\frac{x^{2k}}{(2k)!} \qquad \sin(x)=\sum_{k=0}^{\infty}(-1)^k\frac{x^{2k+1}}{(2k+1)!}$$

Beweis: Weil $\imath^2=-1$ gilt für alle $k\in\mathbb{N}_0$ $\imath^{2k}=(-1)^k$ und $\imath^{2k+1}=\imath(-1)^k$. Also gilt auch für alle $x\in\mathbb{K}$:

$$\begin{aligned}\cos(x) &= \frac{\exp(\imath x)+\exp(-\imath x)}{2} = \sum_{n=0}^{\infty}\frac{1}{2}\left(\frac{(\imath x)^n}{n!}+\frac{(-\imath x)^n}{n!}\right) = \sum_{k=0}^{\infty}(-1)^k\frac{x^{2k}}{(2k)!}\\ \sin(x) &= \frac{\exp(\imath x)-\exp(-\imath x)}{2\imath} = \sum_{n=0}^{\infty}\frac{1}{2\imath}\left(\frac{(\imath x)^n}{n!}-\frac{(-\imath x)^n}{n!}\right) = \sum_{k=0}^{\infty}(-1)^k\frac{x^{2k+1}}{(2k+1)!}\end{aligned}$$

q.e.d.

Kapitel 5

Stetigkeit

5.1 Teilmengen von $\mathbb{K}$

In diesem Abschnitt betrachten wir Teilmengen X von $\mathbb{K}$, also von $\mathbb{R}$ oder $\mathbb{C}$. Der Absolutbetrag $|x-y|$ der Differenz zweier Elemente $x, y \in X$ definiert einen Abstand. Wir hatten in den Sätzen 2.21 und 2.58 folgende Eigenschaften hergeleitet:

(i) Für alle $x, y \in X$ ist $|x-y| \geq 0$ und $|x-y| = 0 \Longleftrightarrow x = y$ (Positivität).

(ii) Für alle $x, y \in X$ ist $|x-y| = |y-x|$ (Symmetrie).

(iii) Für alle $x, y, z \in X$ ist $|x-y| \leq |x-z| + |z-y|$ (Dreiecksungleichung).

Definition 5.1. *(offener Ball, Umgebung, offene Menge)*
Sei $X \subset \mathbb{K}$. Ein offener Ball in X mit Zentrum $x \in X$ und Radius $r > 0$ ist die Menge $B(x, r) = \{y \in X \mid |x-y| < r\}$. Eine Umgebung eines Punktes $x \in X$ ist eine Menge $O \subset X$, die für ein $\epsilon > 0$ den Ball $B(x, \epsilon)$ enthält. Eine offene Menge $O \subset X$ ist eine Teilmenge, die eine Umgebung aller ihrer Punkte ist, d.h. für alle $x \in O$ gibt es ein $\epsilon > 0$, so dass $B(x, \epsilon) \subset O$.

Beispiel 5.2. *In $\mathbb{R}$ besteht der Ball $B(x, r)$ aus dem Intervall $(x-r, x+r)$. In $\mathbb{C}$ besteht der Ball $B(x, r)$ aus allen Punkten, deren euklidischer Abstand zu x kleiner ist als r, also aus einer Kreisscheibe.*

Alle offenen Bälle $B(x, r)$ sind offenbar Umgebungen von x. Sei $y \in B(x, r)$. Dann ist $|x-y| < r$. Sei $z \in B(y, r - |x-y|)$. Dann gilt $|x-z| \leq |x-y| + |y-z| < r$, also auch $B(y, r-|x-y|) \subset B(x, r)$. Deshalb sind die offenen Bälle offene Mengen.

Offenbar ist eine beliebige Vereinigung von offenen Mengen offen. Seien O und O' zwei offene Mengen und $x \subset O \cap O'$. Dann gibt es $r > 0$ und $r' > 0$ so dass $B(x, r) \subset O$ und $B(x, r') \subset O'$. Also ist $B(x, \min\{r, r'\}) \subset B(x, r) \cap B(x, r') \subset O \cap O'$. Also ist $O \cap O'$ offen. Damit ist auch die Schnittmenge von endlich vielen offenen Mengen offen.

Definition 5.3. *(abgeschlossene Mengen, Abschluss)*
Für eine Teilmenge $X \subset \mathbb{K}$ heißen die Komplemente von offenen Teilmengen von X abgeschlossen. Der Abschluss $\bar{A}$ einer Teilmenge $A \subset X$ ist die Schnittmenge aller abgeschlossenen Teilmengen von X, die A enthalten.

Wegen der Regel von de Morgan, sind beliebige Schnitte und endliche Vereinigungen von abgeschlossenen Teilmengen von X abgeschlossen. Deshalb ist der Abschluss einer beliebigen Teilmenge von X abgeschlossen und der Abschluss einer abgeschlossenen Teilmenge gleich der Menge.

Offenbar gehört ein Punkt $x \in X$ genau dann zu dem Abschluss $\bar{A}$, wenn es keine offene Menge in X gibt, die x enthält aber mit A schnittfremd ist. Dies ist wiederum äquivalent dazu, dass für alle $n \in \mathbb{N}$ die offenen Bälle $B(x, \frac{1}{n})$ ein Element a_n aus A enthalten, oder auch dazu, dass es eine Folge $(a_n)_{n\in\mathbb{N}}$ in A gibt, die gegen x konvergiert. Wir sagen, dass eine Folge $(x_n)_{n\in\mathbb{N}}$ in X konvergiert, wenn sie konvergiert, und der Grenzwert in X liegt. Damit haben wir gezeigt:

Lemma 5.4. *Der Abschluss $\bar{A}$ einer Teilmenge $A \subset X$ besteht aus den Grenzwerten von allen Folgen in A, die in X konvergieren. $A \subset X$ ist genau dann abgeschlossen, wenn die Grenzwerte von allen Folgen in A, die in X konvergieren, in A liegen.* **q.e.d.**

5.2 Vollständigkeit und Kompaktheit

Wenn eine Folge $(x_n)_{n\in\mathbb{N}}$ in X konvergiert, dann auch in $\mathbb{K}$. Deshalb gilt

Satz 5.5. *In $X \subset \mathbb{K}$ ist jede konvergente Folge eine Cauchyfolge.* **q.e.d.**

Definition 5.6. *$X \subset \mathbb{K}$ heißt vollständig, wenn jede Cauchyfolge konvergiert.*

Wegen dem Vollständigkeitsaxiom sind $\mathbb{R}$ und $\mathbb{C}$ vollständig. Wegen Lemma 5.4 ist $A \subset \mathbb{K}$ genau dann vollständige, wenn A in $\mathbb{K}$ abgeschlossen ist.

Weil jede reelle Zahl der Grenzwert einer Folge von rationalen Zahlen ist, sind die reellen Zahlen der Abschluss der rationalen Zahlen $\mathbb{Q} \subset \mathbb{R}$. Anstelle unserer axiomatischen Charakterisierung der reellen Zahlen können wir also die reellen Zahlen auch aus den rationalen Zahlen konstruieren als Äquivalenzklassen von rationalen Cauchyfolgen, wobei zwei Cauchyfolgen als äquivalent gelten, wenn ihre Differenz eine Nullfolge ist.

Definition 5.7. *(kompakt) Eine Teilmenge $X \subset \mathbb{K}$ heißt kompakt, wenn jede Folge in X eine in X konvergente Teilfolge besitzt.*

Definition 5.8. *Eine Teilmenge $X \subset \mathbb{K}$ heißt beschränkt, wenn für ein $x \in X$, die Menge der Abstände $\{|x - y| \mid y \in X\}$ beschränkt ist.*

Wegen der Dreiecksungleichung ist diese Bedingung äquivalent dazu, dass für alle $x \in \mathbb{K}$ die Menge der Abstände $\{|x - y| \mid y \in X\}$ beschränkt ist.

Satz 5.9. *(Heine-Borel) Eine Teilmenge $X \subset \mathbb{K}$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Beweis: Offenbar ist jede Folge in einer beschränkten Menge beschränkt. Wegen dem Auswahlprinzip von Bolzano Weierstraß besitzt jede Folge $(x_n)_{n\in\mathbb{N}}$ in einer beschränkten Menge $X \subset \mathbb{K}$ also eine in $\mathbb{K}$ konvergente Teilfolge. Der Grenzwert einer Folge in einer kompakten Teilmenge $X \subset \mathbb{K}$, die in $\mathbb{K}$ konvergiert, muss offenbar in X liegen. Wegen Lemma 5.4 ist also eine beschränkte Teilmenge $X \subset \mathbb{K}$ genau dann kompakt, wenn sie abgeschlossen ist. Umgekehrt enthält eine unbeschränkte Teilmenge $X \subset \mathbb{K}$ eine Folge $(x_n)_{n\in\mathbb{N}}$, so dass für alle $x \in X$ die Folge $|x - x_n|$ gegen ∞ konvergiert. Eine solche Folge kann keinen Häufungspunkt haben. **q.e.d.**

Korollar 5.10. *Teilmengen $A \subset X$ einer kompakten Teilmenge $X \subset \mathbb{K}$ sind genau dann kompakt, wenn sie abgeschlossen sind.*

Beweis: Eine Folge $(x_n)_{n\in\mathbb{N}}$ in einer Teilmenge $A \subset X$ einer kompakten Teilmenge $X \subset \mathbb{K}$ konvergiert wegen Lemma 5.4 genau dann in X, wenn sie in $\mathbb{K}$ konvergiert. Wegen Lemma 5.4 ist A genau dann abgeschlossen in X, wenn sie es in $\mathbb{K}$ ist. **q.e.d.**

Korollar 5.11. *Die kompakten Teilmengen von $\mathbb{R}$ besitzen Minimum und Maximum.*

Beweis: Für jede beschränkte nicht leere Teilmenge $A \subset \mathbb{R}$ und für jedes $n \in \mathbb{N}$ ist $\sup A + \frac{1}{n}$ keine obere Schranke von A und $\inf A + \frac{1}{n}$ keine untere Schranke. Deshalb gibt es ein $a_n \in (\sup A - \frac{1}{n}, \sup A] \cap A$ und ein $b_n \in [\inf A, \inf A - \frac{1}{n}) \cap A$. Die Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergieren gegen $\sup A$ bzw. $\inf A$. Also liegen $\sup A$ und $\inf A$ im Abschluss von A. Kompakte Teilmengen besitzen Minimum und Maximum. **q.e.d.**

Beispiel 5.12. *Die Intervalle $[a, b]$ sind kompakt.*

5.3 Stetigkeit

Definition 5.13. *Seien X und Y jeweils eine Teilmenge entweder von $\mathbb{R}$ oder von $\mathbb{C}$. Eine Abbildung $f : X \to Y, \quad x \mapsto f(x)$ heißt stetig in $x \in X$, wenn es für jedes $\epsilon > 0$ ein $\delta > 0$ gibt, so dass alle $y \in X$, die $|x - y| < \delta$ erfüllen, auch $|f(x) - f(y)| < \epsilon$ erfüllen. Die Abbildung f heißt stetig, wenn sie in allen Punkten von X stetig ist.*

Stetig im Punkt x heißt also, dass alle Punkte, die sehr nahe bei x liegen, auf Werte abgebildet werden, die sehr nahe bei $f(x)$ liegen.

Satz 5.14. *Für eine Abbildung $f : X \to Y, \quad x \mapsto f(x)$ zwischen zwei Teilmengen X und Y jeweils von $\mathbb{R}$ oder $\mathbb{C}$ ist folgendes äquivalent:* **(i)** *f ist stetig in x.*

(ii) *Das Urbild jeder Umgebung von $f(x)$ ist eine Umgebung von x.*

(iii) *Für jede konvergente Folge* $(x_n)_{n\in\mathbb{N}}$ *in* X *konvergiert* $(f(x_n))_{n\in\mathbb{N}}$ *gegen* $f(\lim x_n)$.

Beweis: (i)⇔(ii) Die Umgebungen von x sind gerade die Mengen, die einen δ-Ball um x enthalten. Also ist (ii) äquivalent zu der Aussage, dass das Urbild jedes ϵ-Balles um $f(x)$ einen δ-Ball um x enthält. Das ist nur eine Umformulierung von (i).
(ii)⇔(iii) Die Folgen $(x_n)_{n\in\mathbb{N}}$ und $(f(x_n))_{n\in\mathbb{N}}$ konvergieren genau dann gegen x bzw. $f(x)$, wenn jede Umgebung von x bzw. $f(x)$ alle bis auf endlich viele Folgenglieder enthält. Wenn also $(x_n)_{n\in\mathbb{N}}$ gegen x konvergiert und f (ii) erfüllt, dann konvergiert auch $(f(x_n))_{n\in\mathbb{N}}$ gegen $f(x)$. Also folgt aus (ii) auch (iii). Wenn es umgekehrt einen ϵ-Ball von $f(x)$ gibt, deren Urbild keinen δ-Ball von x enthält, dann gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ von Punkten $x_n \in B\left(x, \frac{1}{n}\right)$, so dass die Folge der entsprechenden Werte $f(x_n)$ im Komplement dieses ϵ-Balles von $f(x)$ liegt: $f(x_n) \notin B(f(x), \epsilon)$. Also konvergiert $(x_n)_{n\in\mathbb{N}}$ gegen x aber $f(x_n)$ nicht gegen $f(x)$. **q.e.d.**

Korollar 5.15. *Für eine Funktion* $f : X \to Y$ *zwischen zwei Teilmengen* X *und* Y *jeweils von* $\mathbb{R}$ *oder* $\mathbb{C}$ *ist folgendes äquivalent:* **(i)** f *ist stetig.*

(ii) *Das Bild* $(f(x_n))_{n\in\mathbb{N}}$ *jeder in* X *konvergenten Folge* $(x_n)_{n\in\mathbb{N}}$ *ist konvergent in* Y *und es gilt* $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$.

(iii) *Das Urbild jeder offenen Teilmenge von* Y *ist offen in* X.

(iv) *Das Urbild jeder abgeschlossenen Teilmenge von* Y *ist abgeschlossen in* X.

Beweis: Wegen dem vorangehenden Satz sind (i) und (ii) äquivalent. Weil eine Menge genau dann offen ist, wenn sie eine Umgebung von allen ihren Punkten ist, zeigt der vorangehende Satz, dass aus (i) bzw. (ii) auch (iii) folgt. Weil jede Umgebung eines Punktes auch eine offene Umgebung des Punktes enthält, folgt wieder wegen dem vorangehenden Satz aus (iii) auch (i) bzw. (ii). Weil nun die abgeschlossenen Mengen gerade die Komplemente der offenen Mengen sind und das Urbild eines Komplementes gerade gleich dem Komplement des Urbildes ist, ist (iii) zu (iv) äquivalent. **q.e.d.**

Korollar 5.16. *Die Komposition zweier stetiger Abbildungen ist stetig. Die analoge punktweise Aussage gilt auch.*

Beweis: Benutze die Äquivalenz zwischen (i) und (iii) im vorangehenden Korollar und die Gleichung $(f \circ g)^{-1}[A] = g^{-1}[f^{-1}[A]]$. **q.e.d.**

Korollar 5.17. *Das Bild einer kompakten Menge unter einer stetigen Abbildung* $f : X \to Y$ *zwischen zwei Teilmengen* X *und* Y *jeweils von* $\mathbb{R}$ *oder* $\mathbb{C}$ *ist kompakt.*

Beweis: Sei $f : X \to Y, \quad x \mapsto f(x)$ eine stetige Abbildung zwischen zwei Teilmengen X und Y jeweils von $\mathbb{R}$ oder $\mathbb{C}$ und $A \subset X$ eine kompakte Menge. Für jede Folge $(y_n)_{n\in\mathbb{N}}$ in $f[A]$ gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ in A mit $f(x_n) = y_n$ für alle $n \in \mathbb{N}$. Weil A kompakt ist, besitzt $(x_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge. Die entsprechende Teilfolge von $(y_n)_{n\in\mathbb{N}} = (f(x_n))_{n\in\mathbb{N}}$ konvergiert wegen der Stetigkeit von f. Also besitzt jede Folge in $f[A]$ eine konvergente Teilfolge. **q.e.d.**

Korollar 5.18. *Sei f eine bijektive stetige Abbildung von einer kompakten Teilmenge X auf eine Teilmenge Y jeweils von $\mathbb{R}$ oder $\mathbb{C}$. Dann ist die Umkehrabbildung stetig.*

Beweis: Wegen dem vorangehenden Korollar ist das Bild $f[X] = Y$ kompakt. Weil aber wegen Korollar 5.10 eine Teilmenge eines kompakten metrischen Raumes genau dann abgeschlossen ist, wenn sie kompakt ist, folgt die Aussage aus dem vorangehenden Korollar und der Charakterisierung (iv) im Korollar über stetige Abbildungen. **q.e.d.**

Beispiel 5.19. (i) *Auf jedem metrischen Raum ist die identische Abbildung $\mathbb{1}_X$ stetig.*
(ii) *Die konstante Abbildung, die alle $x \in X$ auf einen Punkt y abbildet ist stetig.*
(iii) *Aus den Rechenregeln für Folgen folgt, dass für jede Teilmenge X von $\mathbb{R}$ oder $\mathbb{C}$ und alle stetigen Abbildungen $f, g : X \to \mathbb{K}$ auch die Abbildungen*

$$f + g : X \to \mathbb{K}, \qquad x \mapsto f(x) + g(x) \qquad f \cdot g : X \to \mathbb{K}, \qquad x \mapsto f(x) \cdot g(x)$$

stetig sind. Das gilt auch für die Abbildungen

$$-f : X \to \mathbb{K}, \quad x \mapsto -f(x) \text{ und } \quad \frac{1}{f} : X \setminus f^{-1}[\{0\}] \to \mathbb{K} \setminus \{0\}, \quad x \mapsto \frac{1}{f(x)}.$$

Definition 5.20. *(Gleichmässige Stetigkeit, Lipschitzstetigkeit) Eine Abbildung $f : X \to Y, \quad x \mapsto f(x)$ zwischen den beiden Teilmengen X und Y jeweils von $\mathbb{R}$ oder $\mathbb{C}$ heißt gleichmäßig stetig auf einer Teilmenge $A \subset X$, wenn es für alle $\epsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x, y \in A$ mit $|x - y| < \delta$ auch $|f(x) - f(y)| < \epsilon$ gilt.*
Die Abbildung heißt lipschitzstetig auf A, wenn es eine Konstante $L > 0$(Lipschitzkonstante) gibt, so dass $|f(x) - f(y)| \leq L|x - y|$ für alle $x, y \in A$ gilt.

Beispiel 5.21. *Wegen Satz 4.26 ist jede Potenzreihenfunktion $x \mapsto \sum_{n\in\mathbb{N}} a_n x^n$ für jedes r, das kleiner ist als der Konvergenzradius R auf der Teilmenge $\{x \in \mathbb{K} \mid |x| \leq r\}$ lipschitzstetig. Insbesondere ist sie auf dem ganzen Konvergenzbereich stetig.*

Offenbar ist jede lipschitzstetige Abbildung auch gleichmäßig stetig und jede gleichmäßig stetige Abbildung auch stetig. Es gilt auch folgende Umkehrung:

Satz 5.22. *Eine Abbildung $f : X \to Y$ von einer kompakten Teilmenge X von $\mathbb{R}$ oder $\mathbb{C}$ auf eine Teilmenge $Y \subset \mathbb{K}$ ist genau dann stetig, wenn sie gleichmäßig stetig ist.*

Beweis: Sei also $f : X \to Y, \quad x \mapsto f(x)$ stetig und X kompakt. Wenn f nicht gleichmäßig stetig ist, dann gibt es ein $\epsilon > 0$ so dass für alle $n \in \mathbb{N}$ zwei Punkte x_n und y_n in X existieren, mit $|x_n - y_n| < \frac{1}{n}$ und $|f(x_n) - f(y_n)| \geq \epsilon$. Die Folge $(x_n - y_n)_{n\in\mathbb{N}}$ ist also eine Nullfolge. Weil $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ zwei Folgen in der kompakten Menge sind, gibt es eine konvergente Teilfoge von $(x_n)_{n\in\mathbb{N}}$. Dann konvergiert auch die entsprechende Teilfolge von $(y_n)_{n\in\mathbb{N}}$ gegen den gleichen Grenzwert. Wegen der Stetigkeit von f konvergieren dann auch die entsprechenden Teilfolgen von $(f(x_n))_{n\in\mathbb{N}}$ und $(f(y_n))_{n\in\mathbb{N}}$ gegen den gleichen Grenzwert. Das widerspricht $|f(x_n) - f(y_n)| \geq \epsilon$, und damit auch der Annahme, dass f auf X nicht gleichmäßig stetig ist. **q.e.d.**

Definition 5.23. *Eine Folge von Funktionen $(f_n)_{n\in\mathbb{N}}$ von X nach $\mathbb{K}$ heißt*

punktweise konvergent, *wenn die Folgen $(f_n(x))_{n\in\mathbb{N}}$ für jedes $x \in X$ konvergieren. Die Grenzwerte definieren eine Funktion $f : X \to \mathbb{K}, \quad x \mapsto \lim_{n\to\infty} f_n(x)$.*

gleichmäßig konvergent, *wenn es eine Funktion $f : X \to \mathbb{K}, \quad x \mapsto f(x)$ gibt, und für alle $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $|f_n(x) - f(x)| < \epsilon$ für $n \geq N$ und $x \in X$ gilt.*

Gleichmäßig konvergente Folgen sind punktweise konvergent, aber nicht umgekehrt.

Beispiel 5.24. *Die Folge von Funktionen $(f_n)_{n\in\mathbb{N}}$ mit $f_n : [0,1] \to \mathbb{R}, \quad x \mapsto x^n$ konvergiert wegen Satz 3.4 punktweise gegen die Funktion*

$$f : [0,1] \to \mathbb{R}, \quad x \mapsto \begin{cases} 0 & \text{für } x \in [0,1) \\ 1 & \text{für } x = 1 \end{cases}$$

Für alle $n \in \mathbb{N}$ gilt $\sup_{x\in[0,1]} |f_n(x) - f(x)| = \sup_{x\in[0,1)} x^n = 1$. Also konvergiert die Folge $(f_n)_{n\in\mathbb{N}}$ nicht gleichmäßig gegen f.

Definition 5.25. *Sei X eine Teilmenge von $\mathbb{R}$ oder $\mathbb{C}$. Die Menge aller stetigen Funktionen von X nach $\mathbb{K}$ bezeichnen wir mit $C_{\mathbb{K}}(X)$. Eine Funktion $f : X \to \mathbb{K}, \quad x \mapsto f(x)$ heißt beschränkt, wenn das Bild $f[X]$ eine beschränkte Menge ist. $B_{\mathbb{K}}(X)$, bezeichne die Menge aller beschränkten Funktionen auf X. $C_{\mathbb{K}}(X)$ und $B_{\mathbb{K}}(X)$ sind offenbar $\mathbb{K}$-Algebren. Auf $B_{\mathbb{K}}(X)$ bezeichne $\|\cdot\|_\infty$ folgende Abbildung:*

$$\|\cdot\|_\infty : B_{\mathbb{K}}(X) \to \mathbb{R}, \quad f \mapsto \|f\|_\infty = \sup_{x\in X} |f(x)|.$$

Satz 5.26. *Alle stetigen reellen Funktionen auf einer kompakten Teilmenge $X \subset \mathbb{K}$ sind beschränkt. Das Bild einer reellen stetigen Funktion auf einer kompakten Teilmenge $X \subset \mathbb{K}$ besitzt ein Minimum und ein Maximum.*

Beweis: Wegen Satz 5.17 ist das Bild jeder kompakten Teilmenge $X \subset \mathbb{K}$ unter einer stetigen Abbildung kompakt. Wegen Heine-Borel besitzt dann das Bild $f[X]$ einer stetigen reellen Funktion ein Maximum und ein Minimum. **q.e.d.**

Dieser Satz hat viele Anwendungen, z.B. den Fundamentalsatz der Algebra.

Satz 5.27. *Sei X eine Teilmenge von $\mathbb{R}$ oder $\mathbb{C}$. Dann ist der Grenzwert einer gleichmäßig konvergenten Folge $(f_n)_{n\in\mathbb{N}}$ in $C_{\mathbb{K}}(X)$ auch stetig.*

Beweis: Für alle $\epsilon > 0$ gibt es ein N, so dass $|f(y) - f_n(y)| \leq \frac{\epsilon}{3}$ für alle $n \geq N$ und alle $y \in Y$ gilt. Dann gibt es für alle $x \in X$ ein $\delta > 0$, so dass $|f_N(x) - f_N(y)| < \frac{\epsilon}{3}$ für alle $y \in X$ mit $|x - y| < \delta$ gilt. Dann folgt für diese y

$$|f(x) - f(y)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < \epsilon.$$ **q.e.d.**

Kapitel 6

Stetige Funktionen $\mathbb{R} \to \mathbb{R}$

6.1 Umkehrfunktionen

Satz 6.1. *(Zwischenwertsatz) Sei $f : [a,b] \to \mathbb{R}, \quad x \mapsto f(x)$ stetig und $f(a) \neq f(b)$. Dann enthält das Bild $f[[a,b]]$ das abgeschlossene Intervall*

$$[\min\{f(a), f(b)\}, \max\{f(a), f(b)\}] .$$

Beweis: Wir nehmen an $f(a) < f(b)$, andernfalls ist die Argumentation analog. Sei also $y_0 \in (f(a), f(b))$. Sei $A = \{x \in [a,b] \mid f(x) \leq y_0\}$. Weil $a \in A$ ist A eine nicht leere beschränkte Teilmenge von $\mathbb{R}$. Außerdem ist A abgeschlossen. Also ist A kompakt und besitzt ein Maximum $x_0 = \max A$ mit $f(x_0) \leq y_0$. Weil $y_0 \neq f(b)$ ist $x_0 < b$ und es gilt für alle $x > x_0$ auch $f(x) > y_0$. Sei also $(x_n)_n$ eine Folge in $(x_0, b]$, die gegen x_0 konvergiert. Dann gilt $\lim_{n\to\infty} f(x_n) = f(x_0)$ und damit auch $f(x_0) \geq y_0$ und $f(x_0) = y_0$. Also ist $f(x_0) = y_0$ und das Bild von f enthält $[f(a), f(b)]$. **q.e.d.**

Mit diesem Satz läßt sich von vielen stetigen Funktionen zeigen, dass sie surjektiv sind, bzw. ihr Bild bestimmen. Die Injektivität von stetigen Funktionen auf Intervallen ist dagegen äquivalent zu ihrer Monotonie.

Definition 6.2. *(Monotonie) Eine stetige Funktion $f : X \to \mathbb{R}, \quad x \mapsto f(x)$ auf einer Teilmenge X von $\mathbb{R}$ heißt*
monoton wachsend, *wenn aus $x, x' \in X, x \leq x'$ folgt $f(x) \leq f(x')$*
streng monoton wachsend, *wenn aus $x, x' \in X, x < x'$ folgt $f(x) < f(x')$.*
monoton fallend, *wenn aus $x, x' \in X, x \leq x'$ folgt $f(x) \geq f(x')$.*
streng monoton fallend, *wenn aus $x, x' \in X, x < x'$ folgt $f(x) > f(x')$.*

Satz 6.3. *Eine stetige reelle Funktion f auf einem Intervall ist genau dann injektiv, wenn f entweder streng monoton wachsend oder streng monoton fallend ist.*

Beweis: Wir zeigen zunächst dass jede injektive stetige reelle Funktion $f : [a,b] \to \mathbb{R}, \quad x \mapsto f(x)$ das Bild $[\min\{f(a), f(b)\}, \max\{f(a), f(b)\}]$ hat. Wenn es andernfalls ein

$y_0 \in f[[a,b]]$ gibt, dass nicht zu dieser Menge gehört, dann folgt aus dem Zwischenwertsatz, dass jeder Wert in $(\min\{y_0, f(a)\}, \max\{y_0, f(a)\}) \cap (\min\{y_0, f(b)\}, \max\{y_0, f(b)\})$ einmal auf (a, y_0) und einmal auf (y_0, b) angenommen wird, was der Injektivität widerspricht. Falls $f(a) < f(b)$ sind also für alle $x \in (a,b)$ die Bilder $f[(a,x)]$ gleich $(f(a), f(x))$ und falls $f(a) > f(b)$ gleich $(f(x), f(a))$. Im ersten Fall ist f streng monoton wachsend und im zweiten Fall streng monoton fallend. Weil aber alle Paare von Punkten eines beliebigen Intervalles (das mehr als einen Punkt enthält) in einem abgeschlossenen Intervall enthalten sind, das zwei Referenzpunkte enthält, die dann festlegen ob f streng monoton fallend oder streng monoton steigend ist, folgt, dass jede injektive stetige Funktion auf einem Intervall entweder streng monoton wachsend oder streng monoton fallend ist. Umgekehrt ist jede streng monotone Funktion auf einem Intervall auch injektiv. **q.e.d.**

Korollar 6.4. *Die Umkehrfunktion einer bijektiven stetigen Funktion von einem Intervall auf ein Intervall ist stetig.*

Beweis: Offenbar besitzt jeder Punkt x eines Intervalls, das mehr als einen Punkt enthält, eine Umgebung in diesem Intervall, die ein abgeschlossenes beschränktes Intervall ist. Das Bild solcher kompakten Intervalle ist wieder kompakt und wegen dem vorangehenden Satz wieder eine Umgebung von $f(x)$. Dann ist f^{-1} wegen Korollar 5.18 bei $y = f(x)$ stetig. Weil f surjetiv ist, ist damit f^{-1} stetig. **q.e.d.**

Satz 6.5.* *Sei f eine monoton wachsende (fallende) Funktion von einem Interval I nach $\mathbb{R}$. Dann ist die Menge aller Unstetigkeitsstellen von f höchstens abzählbar.*

Beweis*: Wir betrachten monoton wachsende Funktionen. Für monoton fallende Funktionen verläuft der Beweis analog. Für jeden inneren Punkt ξ von I sei $f(\xi_-) = \sup\{f(x) \mid x \in I, x < \xi\}$ und $f(\xi_+)\} = \inf\{f(x) \mid x \in I, \xi < x\}$. Wenn $f(\xi_-) = f(\xi_+)$, dann gibt es für jedes $\epsilon > 0$ $x_-, x_+ \in I$ mit $x_- < \xi < x_+$ so dass

$$f(x_+) - \epsilon < f(\xi_+) = f(\xi_-) < f(x_-) + \epsilon.$$

Dann gilt für alle $x \in [x_-, x_+]$ auch

$$-\epsilon < f(x_-) - f(\xi_-) \leq f(x) - f(\xi_-) = f(x) - f(\xi_+) \leq f(x_+) - f(\xi - +) < \epsilon.$$

Wegen der Monotonie gilt $f(\xi_-) \leq f(\xi) \leq f(\xi_+)$. Also ist f bei solchen ξ stetig. Die Unstetigkeitsstellen bestehen aus den ξ mit $f(\xi_-) < f(\xi_+)$. In jedem solchen Intervall $(f(\xi_-), f(\xi_+))$ ist eine rationale Zahl enthalten. Also gibt es eine Abbildung von den Unstetigkeitsstellen auf die rationalen Zahlen die injektiv sind, weil alle diese offenen Intervalle wegen der Monotonie disjunkt sind. Damit sind die Unstetigkeitsstellen gleichmächtig zu einer Teilmenge der rationalen Zahlen also höchstens abzählbar.**q.e.d.**

Beispiel 6.6. *$\mathbb{R}_0^+ \to \mathbb{R}_0^+, x \to x^k$ ist streng monoton wachsend, also ist die Umkehrabbildung $\mathbb{R}_0^+ \to \mathbb{R}_0^+, x \to x^{\frac{1}{k}}$ stetig und streng monoton wachsend. Dasselbe gilt dann auch für die Abbildungen $\mathbb{R}_0^+ \to \mathbb{R}_0^+, x \to x^{\frac{p}{q}}$ mit der Umkehrabbildung $\mathbb{R}_0^+ \to \mathbb{R}_0^+, x \to x^{\frac{q}{p}}, p, q \in \mathbb{N}$.*

6.2 Die reellen Funktionen $e^x, \ln x, a^x, \log_a x$.

Satz 6.7. *(Eigenschaften von* exp*)*

(i) $e^0 = \exp(0) = 1$

(ii) $e^x > \sum\limits_{k=0}^{n} \frac{x^k}{k!}$ *für jedes* $n \in \mathbb{N}$ *und* $x > 0$

(iii) $x < y \Longrightarrow e^x < e^y$

(iv) $\exp : \mathbb{R} \to \mathbb{R}^+, x \to e^x$ *ist bijektiv.*

Beweis:(i) und **(ii)** folgen aus der Definition.
(iii) $x < y \Longrightarrow y - x > 0$. Dann gilt wegen (ii) $e^{y-x} > 1$. Wegen Satz 4.19 (i) und (ii) gilt dann $e^y - e^x = (e^{y-x} - 1)e^x > 0$. Also folgt $e^x < e^y$.
(iv) Offenbar ist die Funktion wegen (iii) injektiv. Wegen Satz 3.4 gibt es für jedes $y \in \mathbb{R}^+$ ein $n \in \mathbb{R}$, so dass $e^{-n} < y < e^n$. Wegen dem Zwischenwertsatz gehört dann y zum Bild von $[-n, n] \to \mathbb{R}^+, \quad x \mapsto e^x$. **q.e.d.**

Definition 6.8. *(des natürlichen Logarithmus). Die Umkehrfunktion von* $\exp : \mathbb{R} \to \mathbb{R}^+ x \mapsto e^x$ *heißt natürlicher Logarithmus:* $\ln : \mathbb{R}^+ \to \mathbb{R}, x \mapsto \ln x$.

Wegen Korollar 6.4 ist der Logarithmus stetig.

Satz 6.9. *(Eigenschaften von* ln*)*

(i) $\ln(1) = 0$

(ii) $\ln(e) = 1$

(iii) $\ln(xy) = \ln(x) + \ln(y)$ *für alle* $x, y \in \mathbb{R}^+$

(iv) $a^r = e^{\ln(a)\cdot r}$ *für alle* $r \in \mathbb{Q}$ *und* $a \in \mathbb{R}^+$

(v) $\ln\left(e^{ln(a)x}\right) = x\ln(a)$ *für alle* $a \in \mathbb{R}^+, x \in \mathbb{R}$

(vi) $x, y \in \mathbb{R}^+, x < y \Longrightarrow \ln(x) < \ln(y)$

(vii) $\ln : \mathbb{R}^+ \to \mathbb{R}, x \to \ln(x)$ *ist bijektiv*

Beweis: **(i)** $\Leftrightarrow e^0 = 1$ und **(ii)** $\Leftrightarrow e^1 = e$ und **(iii)** $\Longleftrightarrow e^{\ln x + \ln(y)} = (e^{\ln x})(e^{\ln y})$.
(iv) Für $r = \frac{p}{q}$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ gilt $(e^{\ln(a)\frac{p}{q}})^q = e^{\ln(a)p} = a^p$ und $e^{\ln(a)\frac{p}{q}} > 0$. Wegen der Eindeutigkeit der k-ten Wurzel folgt $e^{\ln(a)\frac{p}{q}} = a^{\frac{p}{q}}$.
(v) ist offensichtlich.
(vi) folgt aus (iii) des vorhergehenden Satzes.
(vii) folgt aus (iv) des vorhergehenden Satzes. **q.e.d.**

Definition 6.10. *Für alle $a > 0$ und alle $x \in \mathbb{R}$ sei $a^x = e^{x \ln(a)}$*

Satz 6.11. *(Eigenschaften von a^x)*

(i) $a^{x+y} = a^x a^y$ *für alle* $a \in \mathbb{R}^+, x, y \in \mathbb{R}$.

(ii) $(a^x)^y = a^{xy}$ *für alle* $a \in \mathbb{R}^+, x, y \in \mathbb{R}$.

(iii) *Für $a > 1$ und $x, y \in \mathbb{R}$ gilt $x < y \Longrightarrow a^x < a^y$.*

(iv) *Für $a < 1$ und $x, y \in \mathbb{R}$ gilt $x < y \Longrightarrow a^x > a^y$.*

(v) *Für $a \neq 1$ ist $a^{\cdot} : \mathbb{R} \to \mathbb{R}^+, \quad x \mapsto a^x$ bijektiv und stetig.*

Beweis:(i) $a^{x+y} = e^{x \ln a + y \ln(a)} = e^{x \ln a} e^{y \cdot \ln a} = a^x a^y$.
(ii) $(a^x)^y = e^{y \ln(e^{x \ln a})} = e^{y \cdot x \ln a} = a^{xy}$.
(iii) Für $a > 1$ ist $\ln(a) > 0$. Also folgt aus $x < y$ auch $x \ln a < y \ln a$ und $a^x < a^y$.
(iv) Für $a < 1$ ist $\ln(a) < 0$. Also folgt aus $x < y$ auch $x \ln(a) > y \ln(a)$ und $a^x > a^y$.
(v) Für $x \neq 1$ ist $\ln(a) \neq 0$. Also ist $\mathbb{R} \to \mathbb{R}, \quad x \mapsto \ln(a)x$ bijektiv und stetig, also auch $\mathbb{R} \to \mathbb{R}^+, \quad x \mapsto \exp(\ln(a)x)$. **q.e.d.**

Definition 6.12. *(des Logarithmus zur Basis a) Für alle $a \neq 1$ sei $\log_a : \mathbb{R}^+ \to \mathbb{R}, x \mapsto \log_a(x) = \frac{\ln(x)}{\ln(a)}$ die Umkehrfunktion von $a^{\cdot}$.*

Satz 6.13. *(Eigenschaften des Logarithmus zur Basis a)*

(i) $\log_a(1) = 0$

(ii) $\log_a(a) = 1$

(iii) $\log_a(xy) = \log_a(x) + \log_a(y)$

(iv) *Für $a > 1$ und $x, y \in \mathbb{R}^+$ folgt aus $x < y$ auch $\log_a(x) < \log_a(y)$*

(v) *Für $a < 1$ und $x, y \in \mathbb{R}^+$ folgt aus $x < y$ auch $\log_a(x) > \log_a(y)$*

(vi) $\log_a : \mathbb{R}^+ \to \mathbb{R}, \quad x \mapsto \log_a(x)$ *ist bijektiv und stetig.*

Beweis analog zum Beweis der Eigenschaften von ln. **q.e.d.**

6.3 Die reellen Funktionen $\sin, \cos, \arcsin, \arccos$

Satz 6.14. (i) *Für $x \in [-5,5] \setminus \{0\}$ gilt $1 - \frac{x^2}{2} < \cos x < 1 - \frac{x^2}{2} + \frac{x^4}{24}$.*

(ii) *Für alle $x \in [-4,4]$ gilt $1 - \frac{x^2}{6} \leq \frac{\sin x}{x} \leq 1$. Gleichheit gilt nur für $x = 0$.*

(iii) $\cos : [0, \sqrt{6}] \to \mathbb{R}, x \mapsto \cos(x)$ *ist streng monoton fallend.*

(iv) $\cos$ *hat auf* $[0,2]$ *genau eine Nullstelle, die wir mit* $\frac{\pi}{2}$ *bezeichnen.*

(v) $\sin\left(\frac{\pi}{2}\right) = 1, \quad \exp\left(\imath\frac{\pi}{2}\right) = \imath.$

(vi) $\cos(n\pi) = (-1)^n \quad \sin(n\pi) = 0$ *für alle* $n \in \mathbb{N}$.

(vii) $\cos\left(\left(n + \frac{1}{2}\right)\pi\right) = 0 \quad \sin\left(\left(n + \frac{1}{2}\right)\pi\right) = (-1)^n$ *für alle* $n \in \mathbb{N}$.

(viii) $\cos(x + n\pi) = (-1)^n \cos(x) \quad \sin(x + n\pi) = (-1)^n \sin(x)$

(ix) $\cos\left(x + \left(n - \frac{1}{2}\right)\pi\right) = (-1)^n \sin(x) \quad \sin\left(x + \left(n + \frac{1}{2}\right)\pi\right) = (-1)^n \cos(x).$

(x) $\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1,1], \quad x \mapsto \sin(x)$ *ist streng monoton steigend und bijektiv.*

(xi) $\cos : [0,\pi] \to [-1,1], \quad x \mapsto \cos(x)$ *ist streng monton fallend und bijektiv.*

Beweis:(i) Für $x \in [-5,5]$ und $k \in \mathbb{N} \setminus \{1\}$ gilt $\frac{x^2}{(2k+1)(2k+2)} < 1$. Also ist für alle $x \in [-5,5]$ die Folge $(\frac{x^{2k}}{2k!})_{k \in \mathbb{N}\setminus\{1,2\}}$ monoton fallend und für $x \neq 0$ sogar streng monoton fallend und konvergiert gegen Null (Beispiel (i) im Abschnitt 3.4). Dann folgt aus dem Beweis zu Satz 4.13, dass $1 - \frac{x^2}{2} \leq \cos(x) \leq 1 - \frac{x^2}{2} + \frac{x^4}{24} \leq 1$ für alle $x \in [-5,5]$ und Gleichheit nur für $x = 0$ gilt.
(ii) Für $x \in [-4,4]$ und $k \in \mathbb{N} \setminus \{1\}$ gilt $\frac{x^2}{2k(2k+1)} < 1$. Also ist für alle $x \in [-4,4]$ die Folge $(\frac{x^{2k}}{(2k+1)!})_{k \in \mathbb{N}\setminus\{1\}}$ streng monoton fallend und konvergiert gegen Null (Beispiel (i) in Abschnitt 3.4). Wieder folgt aus dem Beweis von Satz 4.13, dass für alle $x \in [-4,4]$ gilt $1 - \frac{x^2}{6} \leq \frac{\sin x}{x} \leq 1 - \frac{x^2}{6} + \frac{x^4}{120} \leq 1$ und Gleichheit nur für $x = 0$ gilt.
(iii) Wegen den Additionstheorem gilt: $\cos(x) - \cos(y) = \cos(\frac{x+y}{2} - \frac{y-x}{2}) - \cos(\frac{x+y}{2} + \frac{y-x}{2}) = 2\sin(\frac{x+y}{2})\sin(\frac{y-x}{2}) > 0$ wegen (ii) für $x, y \in [0, \sqrt{6}]$ und $x < y$.
(iv) $\sin$ und $\cos$ sind wegen Beispiel (ii) aus dem Abschnitt über stetige Funktionen stetig auf ganz $\mathbb{R}$. Wegen (i) ist $\cos(2) \leq 1 - 2 + \frac{2}{3} = -\frac{1}{3}$. Dann folgt aus dem Zwischenwertsatz, dass es eine Nullstelle in $[0,2]$ gibt. Wegen (iii) kann es höchstens eine Nullstelle geben.
(v) Wegen $\sin^2(x) + \cos^2(x) = 1$ folgt aus (iv) $\sin^2(\frac{\pi}{2}) = 1$ und aus (ii) $\sin(\frac{\pi}{2}) > 0$. Also gilt $\sin(\frac{\pi}{2}) = 1$. Dann folgt aus der Eulerschen Formel $\exp(\imath\frac{\pi}{2}) = \imath$.
(vi) Wegen (v) folgt aus der Eulerschen Formel $\exp(n\imath\pi) = (-1)^n$, also $\cos(n\pi) = (-1)^n$ und $\sin(n\pi) = 0$.

(vii) Wegen (v) folgt aus der Eulerschen Formel: $\exp((n+\frac{1}{2})\imath\pi) = (-1)^n\imath$ also $\cos((n+\frac{1}{2})\pi) = 0$ und $\sin((n+\frac{1}{2})\pi) = (-1)^n$.
(viii) Aus dem Additionstheorem und (vi) folgt (viii).
(ix) Aus dem Additionstheorem und (vii) folgt (ix).
(x) Aus (ix) folgt $\sin(x) = \begin{cases} -\cos(x+\frac{\pi}{2}) & \text{für } x \in [-\frac{\pi}{2}, 0] \\ -\sin(-x) = \cos(\frac{\pi}{2} - x) & \text{für } x \in [0, \frac{\pi}{2}]. \end{cases}$
Dann folgt (x) aus (iii).
(xi) Aus (viii) folgt $\cos(x) = \begin{cases} \cos(x) & \text{für } x \in [0, \frac{\pi}{2}] \\ \cos(-x) = -\cos(\pi - x) & \text{für } x \in [\frac{\pi}{2}, \pi]. \end{cases}$
Dann folgt (xi) aus (iii). **q.e.d.**

Die Umkehrfunktion von $\cos : [0,\pi] \to [-1,1], \quad x \mapsto \cos(x)$ heißt

$$\text{Arcuscosinus} \quad \arccos : [-1,1] \to [0,\pi], \quad x \mapsto \arccos(x).$$

Sie ist wegen (xi) streng monoton fallend und wegen Korollar 6.4 stetig. Die Umkehrfunktion von $\sin : \big[-\frac{\pi}{2}, \frac{\pi}{2}\big] \to [-1,1], \quad x \mapsto \sin(x)$ heißt

$$\text{Arcussinus} \quad \arcsin : [-1,1] \to \Big[-\frac{\pi}{2}, \frac{\pi}{2}\Big], \quad x \mapsto \arcsin(x).$$

Sie ist wegen (x) streng monoton steigend und wegen Korollar 6.4 stetig.

Satz 6.15. *(Polardasrstellung von $z \in \mathbb{C}$) Jede komplexe Zahl hat die Darstellung:*

$$z = r \cdot e^{\imath q} \quad r = |z| \text{ und } q \in \mathbb{R}.$$

Für $z \neq 0$ ist q bis auf Addition von $2\pi n$ eindeutig und heißt Argument von z.

Beweis: Sei $z = x + \imath y$ mit $(x,y) \in \mathbb{R}^2$. Wenn $y \geq 0$ sei $q = \arccos(\frac{x}{\sqrt{x^2+y^2}}) \in [0,\pi]$ und $r = \sqrt{x^2+y^2}$. Dann gilt offenbar $x = r \cdot \cos(q)$ und $r\sin(q) \geq 0$. Außerdem gilt $\frac{y^2}{x^2+y^2} + \frac{x^2}{x^2+y^2} = 1 \Longrightarrow y = r\sin(q)$. Wegen der Eulerschen Formel gilt dann

$$z = r \cdot e^{\imath q} = r\cos(q) + \imath r \sin q = x + \imath y.$$

Wenn $y < 0$ sei $q = \arccos(\frac{-x}{\sqrt{x^2+y^2}}) + \pi$ und $r = \sqrt{x^2+y^2}$. Dann folgt wieder

$$z = re^{\imath q} = r\cos q + \imath r \sin q = x + \imath y.$$

Seien (r,q) und (r',q') mit $re^{\imath q} = r'e^{\imath q'} \Rightarrow r = |re^{\imath q}| = |r'e^{\imath q}| = r'$ und für
$z \neq 0 \quad \Longleftrightarrow \quad r \neq 0$ auch $e^{\imath q}e^{-\imath q'} = e^{\imath(q-q')} = 1 \Rightarrow q - q' = 2\pi n.$ **q.e.d.**

Korollar 6.16. $\exp : \mathbb{C} \to \mathbb{C} \setminus \{0\}$ *ist surjektiv und* $\exp(z) = \exp(z') \Leftrightarrow z - z' \in 2\pi\imath\mathbb{Z}$.

Beweis: Seien $z = x + iy$ mit $(x, y) \in \mathbb{R}^2$. Dann gilt $e^z = e^x e^{iy}$. Also folgt das Korollar aus dem Satz 6.15 und weil $\exp : \mathbb{R} \to \mathbb{R}^+$ bijektiv ist. **q.e.d.**

Korollar 6.17. *Für $z \in \mathbb{C}\setminus\{0\}$ und $n \in \mathbb{N}$ gibt es genau n $w_1, \ldots, w_n \in \mathbb{C}$ mit $w^n = z$.*

Beweis: Seien (r, q) die Polarkoordinaten von z. Dann müssen die Polarkoordinaten (s, p) der Lösungen von $w^n = z$ Die Gleichungen $np = q + 2\pi m$ für $m \in \mathbb{Z}$ erfüllen und $s^n = r$. Also sind die Lösungen gegeben durch $s = \sqrt[n]{r}$ und $p_m = \frac{q}{n} + \frac{2\pi m}{n}$, wobei zwei Lösungen (s, p_m) und $(s, p_{m'})$ genau dann übereinstimmen, wenn $\frac{m-m'}{n} \in \mathbb{Z}$. Also ergeben $m = 0, \ldots, n-1$ alle Lösungen. **q.e.d.**

Satz 6.18. *(Fundamentalsatz der Algebra) Jedes komplexe Polynom $p(z) = a_n z^n + \ldots + a_0$ mit $a_n \neq 0$ und $n \in \mathbb{N}$ hat mindestens eine Nullstelle auf $z \in \mathbb{C}$.*

Beweis: Für $|z| \geq R = 1 + 2\left|\frac{a_{n-1}}{a_n}\right| + \ldots + 2\left|\frac{a_0}{a_n}\right| \geq 1$ gilt

$$\left|\frac{p(z)}{z^n}\right| = \left|\frac{p(z)}{z^n}\right| + \left|-\frac{a_{n-1}}{z} - \ldots - \frac{a_0}{z^n}\right| - \left|-a_n\left(\frac{a_{n-1}}{a_n z} + \ldots + \frac{a_0}{a_n z^n}\right)\right|$$
$$\geq |a_n| - |a_n|\left|\frac{a_{n-1}}{a_n z} + \ldots + \frac{a_0}{a_n z^n}\right| \geq |a_n|\left(1 - \frac{\left|\frac{a_{n-1}}{a_n}\right| + \ldots + \left|\frac{a_0}{a_n}\right|}{|z|}\right) > |a_n|\frac{1}{2}.$$

Also ist $|p(z)| > \frac{|a_n|}{2}|z|^n \geq \frac{|a_n|}{2}|z| > \frac{|a_n|}{2}2\frac{|a_0|}{|a_n|} = |a_0|$. Auf der kompakten Menge $\overline{B(0, R)}$ nimmt $z \mapsto |p(z)|$ wegen Satz 5.26 das Minimum bei einem z_0 an. Dieses liegt in $B(0, R)$ und ist das Minimum auf ganz $z \in \mathbb{C}$, weil außerhalb von $z \in B(0, R)$ gilt $|(p(z)| > |a_0| = |p(0)|$. Wir schreiben jetzt $p(y + z_0) = b_n y^n + \ldots + b_0$ als Polynom in $y = z - z_0$. Dann gilt $b_n = a_n \neq 0$. Wenn $b_0 \neq 0$ gilt $|p(z_0)| = |b_0| > 0$. Dann sei m das kleinste $m \in \mathbb{N}$ mit $b_m \neq 0$. Für $0 < |y| \leq r = \dfrac{1}{1 + 2\left|\frac{b_{m+1}}{b_m}\right| + \ldots + 2\left|\frac{b_n}{b_m}\right|} \leq 1$ gilt dann

$$\left|b_{m+1}y^{m+1} + \ldots + b_n y^n\right| \leq |b_m||y|^m\left(\left|\frac{b_{m+1}}{b_m}\right||y| + \ldots + \left|\frac{b_n}{b_m}\right||y|\right) < \frac{|b_m||y|^m}{2}.$$

Also gilt $$|p(z_0 + y)| < |b_0 + b_m y^m| + \frac{|b_m||y|^m}{2}.$$
Sei w eine Lösung der Gleichung $w^m b_m = -b_0$. Dann gilt für alle $t \in \mathbb{C}$ mit $0 < |tw| \leq r$

$$|p(z_0 + tw)| < |b_0||1 - t^m| + \frac{|b_0|}{2}|t|^m.$$

Insbesondere gilt $|p(z_0 + tw)| < |b_0|\left(1 - \frac{t^m}{2}\right) < |b_0|$ für alle $0 < t < \min\left\{1, \frac{r}{|w|}\right\}$.
Dann ist z_0 nicht das Minimum von $|p(z)|$, im Widerspruch zu $p(z_0) = b_0 \neq 0$. **q.e.d.**

Korollar 6.19. *Jedes komplexe Polynom vom Grade $n \in \mathbb{N}$ zerfällt in ein Produkt von Polynomen ersten Grades.*

Beweis durch vollständige Induktion:

(i) für $n = 1$ ist die Aussage trivial.

(ii) Die Aussage gelte für $n \in \mathbb{N}$. Sei p ein beliebiges Polynom $(n+1)$-ten Grades. Wegen dem Fundamentalsatz der Algebra hat p eine Nullstelle bei $z_0 \in \mathbb{C}$. Wenn wir p als Polynom in $z - z_0$ schreiben, erhalten wir p als Produkt von $(z - z_0)$ mit einem Polynom n-ten Grades. Wegen der Induktionsvoraussetzung zerfällt dieses in ein Produkt von Polynomen ersten Grades, also auch p. **q.e.d.**

Definition 6.20. *(von Tangens und Cotangens)*

$$\tan : \mathbb{C} \setminus \left(\frac{\pi}{2} + \mathbb{Z}\pi\right) \to \mathbb{C}, \quad x \to \frac{\sin(x)}{\cos(x)} \qquad \cot : \mathbb{C} \setminus (\mathbb{Z}\pi) \to \mathbb{C}, \quad x \to \frac{\cos(x)}{\sin(x)}.$$

Beachte $\tan(x+\pi) = \frac{\sin(x+\pi)}{\cos(x+\pi)} = \frac{-\sin(x)}{-\cos(x)} = \tan(x)$ und $\cot(x+\pi) = \cot(x)$.

Satz 6.21. **(i)** $\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$ *ist streng monoton steigend, stetig und bijektiv.*

(ii) $\cot : (0,\pi) \to \mathbb{R}, \quad x \mapsto \cot(x)$ *ist streng monoton fallend, stetig und bijektiv.*

Beweis: (i) auf $x \in [0, \frac{\pi}{2}]$ ist sin streng monoton steigend und cos streng monoton fallend. Also ist tan streng monoton steigend. Wegen $\tan(-x) = -\tan(x)$ folgt dann auch, dass tan auf $(-\frac{\pi}{2}, 0]$ streng monoton steigend ist.
(ii) $\cot(x) = \frac{1}{\tan(x)}$ für $x \in (0, \frac{\pi}{2})$ also ist cot auf $(0, \frac{\pi}{2})$ streng monoton fallen und analog auf $(\frac{\pi}{2}, \pi)$. Für alle $n \in \mathbb{Z}$ sind tan und cot auf $(n\pi, (n+\frac{1}{2})\pi)$ positiv auf $(n-\frac{1}{2})\pi, n\pi)$ negativ. Sie sind beide wegen Beispiel 5.19 (iii) stetig. Außerdem ist für alle $n \in \mathbb{Z}$ $\tan(n\pi) = 0$ und $\cot((n+\frac{1}{2})\pi) = 0$. Dann gilt auch

$$\lim_{n\to\infty} \tan\left(-\frac{\pi}{2} + \frac{1}{n}\right) = -\infty \quad \lim_{n\to\infty} \tan\left(\frac{\pi}{2} - \frac{1}{n}\right) = \infty$$
$$\lim_{n\to\infty} \cot\left(\frac{1}{n}\right) = \infty \quad \lim_{n\to\infty} \cot\left(\pi - \frac{1}{n}\right) = -\infty$$

Also sind $\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$ und $\cot : (0,\pi) \to \mathbb{R}$ wegen Satz 6.1 surjektiv. **q.e.d.**

Die Umkehrfunktion von $\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$ heißt

$$\text{ArcusTangens} \quad \arctan : \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \quad x \mapsto \arctan(x).$$

Die Umkehrfunktion von $\cot : (0,\pi) \to \mathbb{R}$ heißt

$$\text{Arcuscotangens} \quad \operatorname{arccot} : \mathbb{R} \to (0,\pi), \quad x \mapsto \operatorname{arccot}(x).$$

Diese beiden Umkehrfunktionen sind wegen Satz 6.21 streng monoton und wegen Korollar 6.4 stetig.

Kapitel 7

Differenzierbare Funktionen $f : \mathbb{R} \to \mathbb{R}$

7.1 Definition der Ableitung

Definition 7.1. *Sei $f : X \to \mathbb{R}$ eine reelle Funktion auf einer Teilmenge X von $\mathbb{R}$, die eine Umgebung von $x_0 \in \mathbb{R}$ enthält. Dann heißt f im Punkt x_0 differenzierbar, wenn es ein $f'(x_0) \in \mathbb{R}$ gibt, so dass die reelle Funktion*

$$X \to \mathbb{R}, \quad x \mapsto \begin{cases} \frac{f(x)-f(x_0)}{x-x_0} & \textit{für } x \neq x_0 \\ f'(x_0) & \textit{für } x = x_0 \end{cases}$$

stetig bei $x = x_0$ ist. Wenn X offen ist und f in jedem Punkt differenzierbar ist, heißt die Funktion $f' : X \to \mathbb{R}, \quad x \mapsto f'(x)$ die Ableitung von f.

Wir bezeichnen $f'(x)$ auch durch $\frac{df}{dx}(x)$.

Satz 7.2. *Sei f im Punkt x_0 differenzierbar, dann ist f im Punkt x_0 auch stetig.*

Beweis:

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

Also folgt die Aussage aus den Rechenregeln für Folgen und daraus, dass $x \mapsto (x - x_0)$ stetig ist. Hierbei benutzen wir das Kriterium (iii) aus Satz 5.14 **q.e.d.**

Definition 7.3. *Das Differential von f im Punkt x_0 ist die lineare Abbildung $df(x_0) : \mathbb{R} \to \mathbb{R}, \quad h \mapsto f'(x_0)h$. Die Gerade $\{(x, y) \in \mathbb{R}^2 \mid y = f(x_0) + (x - x_0)f'(x_0)\}$ heißt Tangente an den Graphen von f im Punkt $(x_0, f(x_0))$.*

$$\mathit{Graph}(f) = \{(x, y) \in \mathbb{R}^2 \mid y = f(x)\}.$$

Die Sekante durch zwei Punkte $(x_0, f(x_0))$ *und* $(x_1, f(x_1))$ *des Graphen ist gegeben durch*

$$\{(x,y) \in \mathbb{R}^2 \mid y = f(x_0) + \frac{x - x_0}{x_1 - x_0}(f(x_1) - f(x_0))\}.$$

Im Grenzwert $x_1 \to x_0$ konvergiert die Sekante durch $(x_0, f(x_0))$ und $(x_1, f(x_1))$ gegen die Tangente an den Graphen von f im Punkt $(x_0.f(x_0))$.

Beispiel 7.4. (i) $f(x) = |x|$. *Für* $x_0 = 0$ *ist* $\frac{f(x) - f(0)}{x - 0} = \begin{cases} 1 \text{ für } x > 0 \\ -1 \text{ für } x < 0. \end{cases}$ *Also ist* f *im Punkt* $x_0 = 0$ *stetig aber nicht differenzierbar.*

(ii) $f(x) = c \Longrightarrow \frac{f(x)-f(x_0)}{x-x_0} = 0$ *für alle* $x \neq x_0$ *also ist* f *differenzierbar und es gilt* $f'(x) = 0$.

(iii) $f(x) = x \Longrightarrow \frac{f(x)-f(x_0)}{x-x_0} = 1$ *für alle* $x \neq x_0$ *also ist* f *differenzierbear und es gilt* $f'(x) = 1$.

(iv) $f(x) = x^n$ *für* $n \in \mathbb{N}$.

$$\frac{f(x) - f(x_0)}{x - x_0} = x^{n-1} + x_0 x^{n-2} + \ldots + x_0^{n-1} \text{ für alle } x \neq x_0$$

also ist f *differenzierbar und es gilt* $f'(x) = nx^{n-1}$.

(v) $f(x) = \exp(x)$

$$\frac{f(x) - f(x_0)}{x - x_0} = \left(\frac{\exp(x - x_0) - 1}{x - x_0}\right)\exp(x_0).$$

Aufgrund der Definition der Exponentialfunktion gilt: $\frac{\exp(x-x_0)-1}{x-x_0} = \sum_{n=0}^{\infty} \frac{(x-x_0)^n}{(n+1)!}$. *Diese Potenzreihenfunktion ist stetig und bei* $x - x_0 = 0$ *gleich* 1. *Also folgt*

$$f'(x) = \exp(x)$$

(vi) $f(x) = \sin(x)$

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

$$= \frac{\sin\left(\frac{x-x_0}{2} + \frac{x+x_0}{2}\right) + \sin\left(\frac{x-x_0}{2} - \frac{x+x_0}{2}\right)}{x - x_0} = \frac{2\sin\left(\frac{x-x_0}{2}\right)\cos\left(\frac{x+x_0}{2}\right)}{x - x_0}$$

Und wegen der Potenzreihe von sin *gilt* $\quad \frac{2\sin(\frac{x-x_0}{2})}{x - x_0} = \sum_{k=0}^{\infty} \frac{(-1)^k(\frac{x-x_0}{2})^{2k}}{(2k+1)!}$.

Diese Potenzreihenfunktion ist stetig und bei $x - x_0 = 0$ *gleich* 1. *Also folgt*

$$f'(x) = \cos\left(\frac{x + x}{2}\right) = \cos(x).$$

(vii) $f(x) = \cos(x)$

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

$$= \frac{\cos\left(\frac{x_0+x}{2} - \frac{x_0-x}{2}\right) - \cos\left(\frac{x_0+x}{2} + \frac{x_0-x}{2}\right)}{x - x_0} = \frac{2\sin\left(\frac{x_0+x}{2}\right)\sin\left(\frac{x_0-x}{2}\right)}{x - x_0}.$$

Wegen $\dfrac{2\sin\left(\frac{x_0-x}{2}\right)}{x - x_0} = -\dfrac{2\sin\left(\frac{x-x_0}{2}\right)}{x - x_0}$ *folgt* $f'(x) = -\sin(x)$.

7.2 Rechenregeln der Ableitung

Satz 7.5. *(Leibnizregel) Seien $f, g : I \to \mathbb{R}$ in x_0 differenzierbar. Dann sind auch die Funktionen $\lambda f, f + g$ und $f \cdot g$ in x_0 differenzierbar und es gilt*

$$(\lambda f)'(x_0) = \lambda f'(x_0) \qquad (f + g)'(x_0) = f'(x_0) + g'(x_0)$$
$$(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0) \cdot g'(x_0).$$

Wenn $f(x) \neq 0$ für $x \in I$, dann ist auch $\dfrac{1}{f} : I \to \mathbb{R}, x \to \dfrac{1}{f(x)}$ in x_0 differenzierbar und es gilt $\left(\dfrac{1}{f}\right)'(x_0) = -\dfrac{f'(x_0)}{f^2(x_0)}$.

Beweis: Es gilt

$$\frac{\lambda f(x) - \lambda f(x_0)}{x - x_0} = \lambda \frac{f(x) - f(x_0)}{x - x_0} \qquad \text{und}$$
$$\frac{f(x) + g(x) - (f(x_0) + g(x_0))}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x_0 - x_0} \qquad \text{und}$$
$$\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} g(x) + \frac{g(x) - g(x_0)}{x - x_0} f(x_0) \qquad \text{und}$$
$$\left(\frac{1}{f(x)} - \frac{1}{f(x_0)}\right)\frac{1}{x - x_0} = -\frac{f(x) - f(x_0)}{f(x)f(x_0)(x - x_0)}.$$

Also folgt die Aussage aus Beispiel 5.19.

Satz 7.6. *(Kettenregel) Seien f und g reelle Funktionen und der Definitionsbereich von f eine Umgebung von x_0 und der Definitionsbereich von g eine Umgebung von $y_0 = f(x_0)$. Wenn f in x_0 differenzierbar ist und g in y_0, dann ist $g \circ f$ in x_0 differenzierbar und es gilt*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Beweis: $\dfrac{g(f(x))-g(f(x_0))}{x-x_0} = \dfrac{g(f(x))-g(f(x_0))}{f(x)-f(x_0)} \cdot \dfrac{f(x)-f(x_0)}{x-x_0}$. Der ertse Faktor ist aber die Komposition von $x \mapsto f(x)$ mit $y \mapsto \dfrac{g(y)-g(y_0)}{y-y_0}$ also wegen Satz 5.16 und wegen Satz 7.2 stetig. Also folgt die Behauptung aus Beispiel 5.19.

Satz 7.7. *(Ableitung der Umkehrfunktion) Sei f eine bijektive Funktion von $X \to Y$ mit $X, Y \subset \mathbb{R}$ und X eine Umgebung von x_0 und Y eine Umgebung von $y_0 = f(x_0)$. Wenn f in x_0 differenzierbar ist und $f'(x_0) \neq 0$ und f^{-1} in y_0 stetig ist, dann ist auch f^{-1} in y_0 differenzierbar und es gilt $(f^{-1})'(y_0) = \dfrac{1}{f'(x_0)}$.*

Beweis: $\dfrac{f^{-1}(y)-f^{-1}(y_0)}{y-y_0} = \dfrac{x-x_0}{f(x)-f(x_0)}$ für $y=f(x)$ und $y_0=f(x_0)$. Die Funktion $y \mapsto \begin{cases} \frac{f^{-1}(y)-f^{-1}(y_0)}{y-y_0} & \text{für } y \neq y_0 \\ \frac{1}{f'(x_0)} & \text{für } y = y_0 \end{cases}$ ist die Komposition der Funktion $y \to f^{-1}(y)$ mit der Funktion $x \mapsto \begin{cases} \frac{x-x_0}{f(x)-f(x_0)} & \text{für } x \neq x_0 \Longleftrightarrow f(x) \neq f(x_0) \\ \frac{1}{f'(x_0)} & \text{für } x = x_0 \end{cases}$. Also folgt der Satz aus Satz 5.16. **q.e.d.**

Beispiel 7.8. **(i)** $\ln \mathbb{R}^+ \to \mathbb{R}, x \mapsto \ln(x)$

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

(ii) $\arcsin : [-1,1] \to [-\frac{\pi}{2}, \frac{\pi}{2}], x \mapsto \arcsin(x)$

$$\arcsin'(x) = \frac{1}{\cos(y)} = \frac{1}{\sqrt{1-x^2}} \textit{ mit } \sin(y) = x \textit{ und } x^2 \neq 1.$$

(iii) $\arccos : [-1,1] \to [0,\pi], \quad x \mapsto \arccos(x)$

$$\arccos'(x) = \frac{-1}{\sin(y)} = \frac{-1}{\sqrt{1-x^2}} \textit{ mit } \cos(y) = x \textit{ und } x^2 \neq 1.$$

(iv) $\cdot^\alpha : \mathbb{R}^+ \to \mathbb{R}, \quad x \mapsto x^\alpha, \alpha \in \mathbb{R}.$

$$(\cdot^\alpha)'(x) = \exp(\alpha \ln(x))' = \exp(\alpha \ln(x)) \cdot \frac{\alpha}{x} = \alpha x^{\alpha-1}.$$

(v) $a^\cdot : \mathbb{R} \to \mathbb{R}, \quad x \mapsto a^x, a \in \mathbb{R}^+.$

$$(a^\cdot)'(x) = \exp(x \cdot \ln(a))' = \exp(x \ln(a)) \cdot \ln(a) = \ln(a) \cdot a^x.$$

(vi) Quotientenregel. *Seien f und g in x_0 differenzierbar und $g(x_0) \neq 0$. Dann ist $\frac{f}{g}$ in x_0 differenzierbar und es gilt*

$$\left(\frac{f}{g}\right)'(x_0) = \left(f \cdot \frac{1}{g}\right)'(x_0) = \frac{f'(x_0)}{g(x_0)} - \frac{f(x_0)}{g^2(x_0))} g'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}.$$

(vii) $x \mapsto \tan(x) = \dfrac{\sin(x)}{\cos(x)}$

$$\tan'(x) = \frac{\cos(x)\cos(x) - \sin(x)(-\sin(x))}{\cos^2(x)} = 1 + \tan^2(x) = \frac{1}{\cos^2(x)}.$$

(viii) $x \mapsto \cot(x) = \dfrac{\cos(x)}{\sin(x)}$

$$\cot'(x) = \frac{-\sin(x)\sin(x) - \cos(x) cos(x)}{\sin^2(x)} = -1 - \cot^2(x) = \frac{-1}{\sin^2(x)}.$$

(ix) $\arctan : \mathbb{R} \to \left(\frac{\pi}{2}, \frac{\pi}{2}\right), \quad x \mapsto \arctan(x)$

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

(x) $\operatorname{arccot} : \mathbb{R} \to (0, \pi), \quad x \mapsto \operatorname{arccot}(x)$

$$\operatorname{arccot}'(x) = \frac{-1}{1 + \cot^2(g)} = \frac{-1}{1 + x^2} \text{ mit } \cot(y) = x.$$

(xi) $\log_a \mathbb{R}^+ \to \mathbb{R}, \quad x \mapsto \log_a(x) \quad \log_a'(x) = \left(\dfrac{\ln(x)}{\ln(a)}\right)' = \dfrac{1}{x \ln(a)}.$

(xii) $x^x : \mathbb{R}^+ \to \mathbb{R}^+, \quad x \mapsto x^x$

$$(x^x)' = \exp(x \cdot \ln(x))' = \exp(x \cdot \ln(x)) \left(\ln(x) + x\frac{1}{x}\right) = (\ln(x) + 1) \cdot x^x.$$

(xiii) $f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0. \end{cases}$

$$f'(x) = \begin{cases} \lim\limits_{x \to 0\pm} \frac{x^2 \sin\left(\frac{1}{x}\right)}{x} = \lim\limits_{x \to 0\pm} x \sin\left(\frac{1}{x}\right) = 0 & \text{für } x = 0 \\ 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right) & \text{für } x \neq 0. \end{cases}.$$

Diese Funktion ist zwar differenzierbar, aber f' ist im Punkt $x = 0$ nicht stetig.

7.3 Mittelwertsatz und Monotonie

Wenn $f'(x_0)$ einer differenzierbaren Funktion positiv ist, dann gibt es ein $\epsilon > 0$, so dass aus $|x - x_0| < \epsilon$ folgt $\frac{f(x)-f(x_0)}{x-x_0} > 0$. Dann gilt für $x \in (x_0 - \epsilon, x_0)$ auch $f(x) < f(x_0)$ und für $x \in (x_0, x_0 + \epsilon)$ auch $f(x) > f(x_0)$. Analoges gilt für negatives $f'(x_0)$.

Definition 7.9. *(relative Maxima und Minima) $f : (a, b) \to \mathbb{R}$, $x \mapsto f(x)$ hat bei $x_0 \in (a, b)$ ein lokales Maximum (Minimum), falls es ein $\epsilon > 0$ gibt, so dass aus $|x - x_0| < \epsilon$ folgt $f(x) \leq f(x_0)$ bzw. $f(x) \geq f(x_0)$.*

Eine differenzierbare Funktion kann also nur an den Nullstellen der Ableitung relative Extremwerte besitzen.

Definition 7.10. *(kritischer Punkt und kritischer Wert) Eine Nullstelle der Ableitung einer differenzierbaren Funktion heißt kritischer Punkt. Der entsprechende Funktionswert heißt kritischer Wert.*

Kanditaten für die Minima und Maxima einer stetigen Funktion $f : [a, b] \to \mathbb{R}$ sind

(i) Kritische Punkte

(ii) Randpunkte

(iii) Punkte an denen f nicht differenzierbar ist.

Satz 7.11. *(Satz von Rolle) Sei $f : [a, b] \to \mathbb{R}$ stetig und auf (a, b) differenzierbar. Falls $f(a) = f(b)$, dann existiert ein $x_0 \in (a, b)$ mit $f'(x_0) = 0$.*

Beweis: Wegen Korollar 5.17 gibt es $x_1, x_2 \in [a, b]$ mit $f(x_1) \leq f(x) \leq f(x_2)$ für alle $x \in [a, b]$. Wenn x_1 und x_2 beide am Rand liegen $x_1, x_2 \in \{a, b\}$ dann muss f konstant gleich $f(a) = f(b)$ sein. Also gilt dann $f'(x) = 0$ für alle $x \in (a, b)$. Andernfalls muss es einen lokalen Extremwert in (a, b) geben, an dem die Ableitung verschwindet. **q.e.d.**

Satz 7.12. *(Verallgemeinerter Mittelwertsatz) Seien $f, g : [a, b] \to \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann existiert ein $x_0 \in (a, b)$ mit*

$$(f(b) - f(a))g'(x_0) = (g(b) - g(a))f'(x_0) \quad \Leftrightarrow \quad \frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x_0)}{g'(x_0)} \text{ für } g(b) \neq g(a).$$

Die Tangente an den (f, g)-Graphen in $(f(x_0), g(x_0))$ verläuft also parallel zu der Verbindungsgeraden der Endpunkte $(f(a), g(a))$ und $(f(b), g(b))$.

Beweis: Wende Satz 7.11 auf die Funktion $x \mapsto (f(b) - f(a))g(x) - (g(b) - g(a))f(x)$ an. Sie erfüllt die Vorraussetzungen und ihre Ableitung ist $x \mapsto (f(b) - f(a))g'(x) - (g(b) - g(a))f'(x)$. **q.e.d.**

Satz 7.13. *(Mittelwertsatz) Sei $f:[a,b]\to\mathbb{R}$, $x\mapsto f(x)$ stetig und auf (a,b) differenzierbar. Dann existiert ein $x_0\in(a,b)$ mit $f'(x_0)=\frac{f(b)-f(a)}{b-a}$.*

Beweis: Wende den verallgemeinerten Mittelwertsatz auf f und $\mathbb{1}_{[a,b]}$ an. **q.e.d.**

Satz 7.14. *(Schrankensatz) Sei $f:[a,b]\to\mathbb{R}$ stetig und auf (a,b) differenzierbar. Wenn $|f'(x)|\leq L$ für alle $x\in(a,b)$ gilt, dann ist f lipschitzstetig mit Lipschitzkonstante L.*

Beweis: Seien $x<y\in[a,b]$. Dann erfüllt $f:[x,y]\to\mathbb{R}$ die Voraussetzungen des Mittelwertsatzes. Also gibt es $x_0\in(x,y)$ mit $f(y)-f(x)=f'(x_0)(y-x)$. Dann folgt aber $|f(y)-f(x)|=|f'(x_0)||y-x|\leq L|y-x|$. **q.e.d.**

Satz 7.15. *(Ableitung und Monotonie) Sei $f:[a,b]\to\mathbb{R}$ stetig und auf (a,b) differenzierbar. Dann gilt*

(i) *$f'(x)=0$ für alle $x\in(a,b)\iff f$ ist konstant.*

(ii) *$f'(x)\geq 0$ für alle $x\in(a,b)\iff f$ ist monoton steigend.*

(iii) *$f'(x)\leq 0$ für alle $x\in(a,b)\iff f$ ist monoton fallend.*

(iv) *$f'(x)\geq 0$ für alle $x\in(a,b)$ und der Abschluss der Menge $\{x\in(a,b)\mid f'(x)>0\}$ ist $[a,b]\iff f$ ist streng monoton steigend.*

(v) *$f'(x)\leq 0$ für alle $x\in(a,b)$ und der Abschluss der Menge $\{x\in(a,b)\mid f'(x)<0\}$ ist $[a,b]\iff f$ ist streng monoton fallend.*

Beweis: Weil eine Funktion genau dann konstant ist, wenn sie monoton steigend und monoton fallend ist, folgt (i) aus (ii) und (iii). Wir beweisen nun (ii) und (iv). Seien $x<y\in[a,b]$. Dann erfüllt $f:[x,y]\to\mathbb{R}$ die Voraussetzungen des Mittelwertsatzes. Wenn $f'(x_0)\geq 0$ für alle $x_0\in(a,b)$, dann folgt also $f(y)-f(x)\geq 0$ und f ist monoton wachsend. Umgekehrt folgt aus $f'(x_0)<0$, dass für ein $\epsilon>0$ gilt $f(x_0-\epsilon)>f(x_0+\epsilon)$, f also nicht monoton steigend sein kann. Das zeigt (ii). Wenn f monoton wachsend, aber nicht streng monoton wachsend ist, dann gibt es $x<y\in[a,b]$ mit $f(x)=f(y)$. Dann ist f aber auf $[x,y]$ konstant und f' verschwindet auf (x,y). Weil jede offene Menge ein offenes Intervall enthält folgt damit (iv). **q.e.d.**

Korollar 7.16. *(isolierte kritische Punkte) Sei $f:(a,b)\to\mathbb{R}$ differenzierbar und x_0 ein kritischer Punkt.*

(i) *Wenn es ein $\epsilon>0$ gibt, so dass $f'(x)<0$ für $x\in(x_0-\epsilon,x_0)$ und $f'(x)>0$ für $x\in(x_0,x_0+\epsilon)$, dann ist x_0 ein lokales Minimum.*

(ii) *Wenn es ein $\epsilon>0$ gibt, so dass $f'(x)>0$ für $x\in(x_0-\epsilon,x_0)$ und $f'(x)<0$ für $x\in(x_0,x_0+\epsilon)$, dann ist x_0 ein lokales Maximum.*

(iii) *Wenn f' bei x_0 differenzierbar ist und $f''(x_0) > 0$, dann ist x_0 ein lokales Minimum.*

(iv) *Wenn f' bei x_0 differenzierbar ist und $f''(x_0) < 0$, dann ist x_0 ein lokales Maximum.* **q.e.d.**

Beispiel 7.17. *Die Funktion $f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto (x+1)e^{-x}$ hat die Ableitung $f'(x) = (1-(x+1))e^{-x} = -xe^{-x}$. Also ist sie auf $(-\infty, 0]$ streng monoton wachsend und auf $[0, \infty)$ streng monoton fallend. Insbesondere ist $f(0) = 1$ ein globales Maximum. Also gilt für alle $x \in \mathbb{R}$ auch $e^x \geq (1+x)$.*

7.4 Regel von de L'Hopital

Definition 7.18. *(Grenzwerte von Funktionswerten) Für eine Funktion $f : (a,b) \to \mathbb{R}, \quad x \mapsto f(x)$ existiert der Grenzwert $\lim\limits_{x \to a+} f(x)$ genau dann, wenn es eine Zahl $f(a)$ gibt, so dass auf $[a,b)$ die Funktion $x \mapsto \begin{cases} f(x) \text{ für } x \in (a,b) \\ f(a) \text{ für } x = a \end{cases}$ stetig bei $x = a$ ist. Wir schreiben dann $\lim\limits_{x \to a+} f(x) = f(a)$.*

Der analoge Grenzwert $x \to b$ wird mit $\lim\limits_{x \to b-} f(x)$ bezeichnet. Aufgrund der Definition der Stetigkeit existiert der Grenzwert $\lim\limits_{x \to a+} f(x)$ also genau dann, wenn es eine Zahl $f(a)$ gibt, so dass für jedes $\epsilon > 0$ ein $\delta > 0$ gibt mit den aus $|x - a| < \delta$ folgt $|f(x) - f(a)| < \epsilon$. Wegen Satz 5.14 ist das äquivalent dazu, dass für jede Folge $(x_n)_{n \in \mathbb{N}}$ in (a,b), die gegen a konvergiert, die Folge $(f(x_n))_{n \in \mathbb{N}}$ gegen $f(a)$ konvergiert.

Satz 7.19. *(1. Regel von de L'Hopital) Seien $\infty < a < b < \infty$ und f und g auf (a,b) differenzierbare Funktionen, so dass $\lim\limits_{x \to a+} f(x) = 0 = \lim\limits_{x \to a+} g(x)$. Wenn der Grenzwert $\lim\limits_{x \to a+} \frac{f'(x)}{g'(x)}$ existsiert, dann existiert auch $\lim\limits_{x \to a+} \frac{f(x)}{g(x)}$ und es gilt $\lim\limits_{x \to a+} \frac{f(x)}{g(x)} = \lim\limits_{x \to a+} \frac{f'(x)}{g'(x)}$.*

Bemerkung 7.20. *Wenn die Grenzwerte $\lim\limits_{x \to a+} f'(x)$ und $\lim\limits_{x \to a+} g'(x)$ existieren und der zweite nicht verschwindet, dann existiert auch $\lim\limits_{x \to a+} \frac{f'(x)}{g'(x)}$ mit $\lim\limits_{x \to a+} \frac{f'(x)}{g'(x)} = \frac{\lim\limits_{x \to a+} f'(x)}{\lim\limits_{x \to a+} g'(x)}$.*

Beweis: Die auf $[a,b)$ stetig fortgesetzten Funktionen f und g erfüllen die Voraussetzungen des Verallgemeinerten Mittelwertsatzes. Deshalb gibt es für jedes $x \in (a,b)$ ein $x_0 \in (a,x)$ so dass gilt $\frac{f(x)}{g(x)} = \frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f'(x_0)}{g'(x_0)}$. Wenn also der Grenzwert $\lim_{x \to a+} \frac{f'(x)}{g'(x)}$ existiert, dann existiert auch der Grenzwert $\lim_{x \to a+} \frac{f(x)}{g(x)} = \lim_{x \to a+} \frac{f'(x)}{g'(x)}$. **q.e.d.**

Satz 7.21. *(2. Regel von de L'Hopital) Unter derselben Voraussetzung wie bei der 1.Regel von de L'Hopital, nur gelte $\lim\limits_{x \to a+} f(x) = \lim\limits_{x \to a+} g(x) = \infty$ statt $\lim\limits_{x \to a+} f(x) = \lim\limits_{x \to a+} g(x) = 0$, gilt dieselbe Schlußfolgerung.*

Beweis: Für jedes $a < x < y < b$ gibt es wegen dem verallgemeinerten Mittelwertsatz ein $x_0 \in (x,y)$ so dass gilt $\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(x_0)}{g'(x_0)}$. Wenn f und g die Vorraussetzungen der 2. Regel von de L'Hopital erfüllen, dann gibt es also für jedes $\epsilon > 0$ ein $y \in (a,b)$, so dass es für alle $x_0 \in (a,y)$ gilt $|\frac{f'(x_0)}{g'(x_0)} - \alpha| < \frac{\epsilon}{2}$ mit $\alpha = \lim_{x\to a+} \frac{f'(x)}{g'(x)}$. Dann folgt $\alpha - \frac{\epsilon}{2} < \frac{f(x)-f(y)}{g(x)-g(y)} < \alpha + \frac{\epsilon}{2}$ für alle $x \in (a,y)$. Wegen $\lim_{x\to a+} g(x) = \infty$ gibt es ein $y_0 \in (a,y)$ so dass für alle $x \in (a,y_0)$ gilt $g(x) > \min\{(g(y),0\}$. Daraus folgt

$$\left(\alpha - \frac{\epsilon}{2}\right)(g(x)-g(y)) + f(y) \quad < f(x) < \quad \left(\alpha + \frac{\epsilon}{2}\right)(g(x)-g(y)) + f(x) \quad \text{oder}$$
$$\left(\alpha - \frac{\epsilon}{2}\right) + \frac{f(y) - g(y)\left(\alpha - \frac{\epsilon}{2}\right)}{g(x)} \quad < \frac{f(x)}{g(x)} < \quad \left(\alpha + \frac{\epsilon}{2}\right) + \frac{f(y) - g(y)\left(\alpha + \frac{\epsilon}{2}\right)}{g(x)}.$$

Wegen $\lim_{x\to a+} g(x) = \infty$ gibt es dann auch ein $y_1 \in (a,y_0)$, so dass für alle $x \in (a,y_1)$ gilt $\alpha - \epsilon < \frac{f(x)}{g(x)} < \alpha + \epsilon$. Also gilt $\lim_{x_0\to a+} \frac{f(x_0)}{g(x_0)} = \lim_{x_0\to a+} \frac{f'(x_0)}{g'(x_0)}$. **q.e.d.**

Die analogen Aussagen für die Grenzwerte $\lim_{x\to b-}$ gelten natürlich auch. Grenzwerte der Form $\lim_{x\to -\infty+} f(x)$ bzw. $\lim_{x\to\infty-} f(x)$ definieren wir als die Grenzwerte $\lim_{x\to 0-} f(1/x)$ bzw. $\lim_{x\to 0+} f(1/x)$. Wegen der Kettenregel gilt dann

$$\frac{\frac{df(\frac{1}{x})}{dx}}{\frac{dg(\frac{1}{x})}{dx}} = \frac{\frac{-1}{x^2}f'(\frac{1}{x})}{\frac{-1}{x^2}g'(\frac{1}{x})} = \frac{f'(\frac{1}{x})}{g'(\frac{1}{x})}.$$

Deshalb gelten die analogen Aussagen auch für diese Grenzwerte.

7.5 Konvexität und Ableitungen

Definition 7.22. *Eine reelle Funktion auf einem Intervall heißt (streng) konvex, wenn für alle a,b im Definitionsbereich und alle $t \in (0,1)$ gilt*

$$f((1-t)a + tb) \le (1-t)f(a) + tf(b) \quad \textit{bzw.} \quad f((1-t)a+tb) < (1-t)f(a) + tf(b).$$

Satz 7.23. *Für eine reelle Funktion f auf einem Intervall I ist folgendes äquivalent:*

(i) *f ist konvex*

(ii) *Für $[a,b] \subset I$ und $x \in (a,b)$ gilt* $\quad f(x) \le f(a) + \frac{f(b)-f(a)}{b-a}(x-a).$

(iii) *Für $[a,b] \subset I$ und $x \in (a,b)$ gilt* $\quad \frac{f(x)-f(a)}{x-a} \le \frac{f(b)-f(a)}{b-a} \le \frac{f(b)-f(x)}{b-x}.$

(iv) *Für $[a,b] \subset I$ und $x \in (a,b)$ gilt* $\quad \frac{f(x)-f(a)}{x-a} \le \frac{f(b)-f(x)}{b-x}.$

Außerdem gelten die analogen Äquivalenzen zu streng konvex, wenn für $[a,b] \subset I$ und $x \in (a,b)$ die Ungleichungen $\le$ durch $<$ ersetzt werden.

Beweis:(i)⇒(ii) Sei also $a < x < b$ im Definitionsbereich. Definiere $t = \frac{x-a}{b-a}$ dann ist $t \in (0,1)$ und $(1-t)a + tb = x$. Also folgt aus (i)

$$f(x) \leq \left(1 - \frac{x-a}{b-a}\right) f(a) + \left(\frac{x-a}{b-a}\right) f(b) \qquad = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$$

(ii)⇒(iii) Die erste Ungleichung in (iii) folgt sofort aus (ii). Außerdem folgt aus (ii)

$$f(b) - f(x) \geq f(b) - f(a) - \frac{f(b)-f(a)}{b-a}(x-a) = \frac{f(b)-f(a)}{b-a}(b-x)$$

und damit folgt auch die zweite Ungleichung in (iii) aus (ii).
(iii)⇒(iv) ist offensichtlich.
(iv)⇒(i) Wir können wegen der Symmetrie $(a,b,t) \leftrightarrow (b,a,1-t)$ in (i) $a < b$ annehmen. Dann sei $x = (1-t)a + tb \in (a,b)$. Wegen (iv) gilt

$$(b-x)(f(x)-f(a)) \leq (x-a)(f(b)-f(x)) \quad \text{also auch}$$
$$(b-a)f(x) \leq (b-x)f(a) + (x-a)f(b) = ((b-a)-(x-a))f(a) + (x-a)f(b).$$

Es gilt aber $x - a = t(b-a)$. Also folgt $f((1-t)a + tb) \leq (1-t)f(a) + tf(b)$.

Die analogen Aussagen für streng konvex lassen sich genauso beweisen, wenn wir alle Ungleichungen $\leq$ durch $<$ ersetzen. **q.e.d.**

Korollar 7.24. *Für eine stetige relle Funktion auf einem Intervall, die im Inneren des Intervalls differenzierbar ist, ist folgendes äquivalent:*

(i) *f ist (streng) konvex*

(ii) *f' ist (streng) monoton wachsend*

Beweis: (i)⇒(ii): Seien $a < b < c$ und $x < b < y$ im Definitionsbereich. Wegen Satz 7.23 (iii) folgt $\frac{f(x)-f(a)}{x-a} \leq \frac{f(c)-f(a)}{c-a} \leq \frac{f(y)-f(c)}{y-c}$. Aus dem Grenzwert $x \to a$ und $y \to c$ folgt dann $f'(a) \leq \frac{f(c)-f(a)}{c-a} \leq f'(c)$ und damit (ii). Umgekehrt folgt aus (ii) wegen dem Mittelwertsatz die Bedingung (iv) von Satz 7.23. **q.e.d.**

Korollar 7.25. *Für eine stetige reelle Funktion auf einem Intervall, die im Inneren zweimal differenzierbar ist, ist folgendes äquivalent*

(i) *f ist (streng) konvex.*

(ii) *$f''(x) \geq 0$ im Inneren des Intervalls (der Abschluss der Menge $\{x \mid f''(x) > 0\}$ ist das ganze Intervall).*

Dieses Korollar folgt sofort aus Korollar 7.24 und Satz 7.15. **q.e.d.**

Wenn wir die Ungleichungen alle umdrehen, so erhalten wir die analogen Aussagen für konkave Funktionen. Also ist eine Funktion f genau dann (streng) konkav, wenn die negative Funktion $-f$ (streng) konvex ist.

Übungsaufgabe 7.26. *Zeige, dass die Umkehrfunktion einer konvexen bijektiven monoton wachsenden Funktion konkav ist.*

Beispiel 7.27. **(i)** $f(x) = x^2 \Longrightarrow f'' = 2$. *Also ist* f *streng konvex.*

(ii) $f : \mathbb{R}_0^+ \to \mathbb{R}_0^+, \quad x \mapsto \sqrt{x} \Longrightarrow f'' = \frac{-1}{4x^{3/2}}$. *Also ist* f *streng konkav.*

(iii) $\exp : \mathbb{R} \to \mathbb{R}^+, \to \exp(x) \Longrightarrow \exp'' = \exp$. *Also ist* exp *streng konvex.*

(iv) $\ln : \mathbb{R}^+ \to \mathbb{R}, x \to \ln(x) \Longrightarrow \ln''(x) = -\frac{1}{x^2}$. *Also ist* ln *streng konkav.*

7.6 Konvexität und Ungleichungen

Satz 7.28.* *(Ungleichung von Jensen) Sei* f *eine reelle konvexe Funktion auf einem Intervall. Seien* $x_1, \ldots, x_n$ *im Definitionsbereich und* $\lambda_1, \ldots, \lambda_n$ *positive Zahlen, die* $\lambda_1 + \ldots + \lambda_n = 1$ *erfüllen. Dann gilt*

$$f(\lambda_1 x_1 + \ldots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \ldots + \lambda_n f(x_n).$$

Wenn f *streng konvex ist, dann gilt Gleichheit nur für* $x_1 = x_2 = \ldots = x_n$.

Beweis*: durch vollständige Induktion:

(i) Für $n = 1$ muss $\lambda_1 = 1$ sein, so dass die Aussage klar ist.

(ii) Die Aussage gelte für $n \in \mathbb{N}$. Seien $x_1, \ldots, x_{n+1}$ und $\lambda_1, \ldots, \lambda_{n+1}$ wie gefordert. Dann definieren wir $\lambda = \lambda_1 + \ldots + \lambda_n$ und $x = \frac{\lambda_1}{\lambda} x_1 + \ldots + \frac{\lambda_n}{\lambda} x_n$. Also gilt $\lambda_{n+1} = 1 - \lambda$ und $\frac{\lambda_1}{\lambda} + \ldots + \frac{\lambda_n}{\lambda} = 1$. Dann folgt aus der Induktionsvoraussetzung $f(x) \leq \frac{\lambda_1}{\lambda} f(x_1) + \ldots + \frac{\lambda_n}{\lambda} f(x_n)$. Wenn f streng konvex ist, dann gilt Gleichhheit nur für $x_1 = \ldots = x_n$. Weil f konvex ist folgt aber $f(\lambda x + (1 - \lambda)x_{n+1}) \leq \lambda f(x) + (1 - \lambda) f(x_{n+1}) \leq \lambda_1 f(x_1) + \ldots + \lambda_{n+1} f(x_{n+1})$. Wenn f streng konvex ist, dann gilt Gleichheit wieder nur für $x_{n+1} = x = x_1 = \ldots = x_n$. **q.e.d.**

Korollar 7.29. *(Ungleichung arithmetisches-geometrisches Mittel) Seien* $\lambda_1, \ldots, \lambda_n$ *positive Zahlen mit* $\lambda_1, \ldots, \lambda_n = 1$. *Dann gilt für positive Zahlen* $x_1, \ldots, x_n$

$$x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_n^{\lambda_n} \leq \lambda_1 x_1 + \ldots + \lambda_n x_n.$$

Insbesondere gilt $\sqrt[n]{x_1 \cdots x_n} \leq \frac{x_1 + \ldots + x_n}{n}$. *Gleichheit gilt nun für* $x_1 = x_2 \ldots = x_n$.

Beweis*: $-\ln$ ist streng konvex. Also folgt aus Jensen's Ungleichung

$$\begin{aligned} -\ln(\lambda_1 x_1 + \ldots + \lambda_n x_n) &\leq -\lambda_1 \ln x_1 - \ldots - \lambda_n \ln x_n \\ \Longleftrightarrow \lambda_1 \ln x_1 + \ldots + \lambda_n \ln x_n &\leq \ln(\lambda_1 x_1 + \ldots + \lambda_n x_n) \end{aligned}$$

Wegen der Monotonie von exp folgt:

$$x_1^{\lambda_1}\cdots x_n^{\lambda_n} = \exp(\lambda_1 \ln x_1 + \ldots + \lambda_n \ln x_n) \leq \lambda_1 x_1 + \ldots + \lambda_n x_n. \qquad \textbf{q.e.d.}$$

Ersetzen wir $x_1,\ldots,x_n$ durch $y_1^{1/\lambda_1},\ldots,y_n^{1/\lambda_n}$ so erhalten wir

Korollar 7.30.* *Seien $\lambda_1,\ldots,\lambda_n$ positive Zahlen mit $\lambda_1+\ldots+\lambda_n = 1$ und $y_1,\ldots,y_n$ positive Zahlen. Dann gilt*

$$y_1\ldots y_n \leq \lambda_1 y_1^{1/\lambda_1} + \ldots + \lambda_n y^{1/\lambda_n}.$$

Gleichheit gilt nur für $y_1^{1/\lambda_1} = y_2^{1/\lambda_2} = \ldots = y_n^{1/\lambda_n}$. **q.e.d.**

Korollar 7.31.* *(Young'sche Ungleichung) Seien $p>0, q>0$ mit $\frac{1}{p}+\frac{1}{q}=1$. Dann gilt für alle $x>0$ und $y>0$*

$$xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q.$$

Gleichheit gilt nur für $x^p = y^q$. **q.e.d.**

7.7 Taylorreihen

Auf offenen Intervallen I (Teilmenge von $\mathbb{R}$) ist die Ableitung f' einer differenzierbaren Funktion f wieder eine Funktion auf I. Die Bildung der Ableitung ist also ein linearer Operater $\frac{d}{dx}$, der differenzierbaren Funktionen auf I, Funktionen auf I zuordnet. Wenn die Ableitung wieder differenzierbar ist, können wir diesen Operator nochmal anwenden und erhalten $(\frac{d}{dx})^2 f = f''$ die zweite Ableitung von f. Durch n–faches Anwenden erhalten wir gegebenenfalls dann die n-te Ableitung $f^{(n)}$.

Definition 7.32. *Für alle $n\in\mathbb{N}_0$ sei $C^n_{\mathbb{R}}(I)$ die Menge aller n-mal stetig differenzierbaren reellen Funktionen auf I, und $C^\infty_{\mathbb{R}}(I)$ die Menge aller beliebig oft differenzierbaren reellen Funktionen auf I.*

$$C_{\mathbb{R}}(I) = C^0_{\mathbb{R}}(I) \supset C^1_{\mathbb{R}}(I) \supset \ldots \supset C^n_{\mathbb{R}}(I) \supset \ldots \supset C^\infty_{\mathbb{R}}(I)$$

Beispiel 7.33. **(i)** $\exp \in C^\infty_{\mathbb{R}}(\mathbb{R})$, *weil* $\exp^{(n)} = \exp$.

(ii) *für alle $n\in\mathbb{N}_0$ ist $x\mapsto x^n \in C^\infty_{\mathbb{R}}(\mathbb{R})$, weil $(x^n)^{(n)} = n!$ und $(x^n)^{(m)} = 0$ für $m>n$.*

(iii) *für alle $n\in\mathbb{N}$ ist $x\mapsto x^{-n}\in C^\infty_{\mathbb{R}}(\mathbb{R}\setminus\{0\})$, weil $(x\mapsto x^{-n})^{(m)} =$*

$$x\mapsto \frac{(-n)(-n-1)\ldots(-n-m+1)}{x^{n+m}} = (-1)^m\frac{(n+m-1)(n+m-2)\ldots n}{x^{n+m}}.$$

(iv) $\ln \in C^\infty_{\mathbb{R}}(\mathbb{R}^+)$ *weil* $\ln^{(m)}(x) = \dfrac{(-1)^{m-1}(m-1)!}{x^m}$ *für $m\geq 1$ und mit $0! = 1$.*

Satz 7.34. *Für alle $n \in \mathbb{N}$ gilt*

(i) $\frac{d^n}{dx^n} f \cdot g = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}$ *für alle $f, g \in C^n_{\mathbb{R}}(I)$.*

(ii) *$C^n_{\mathbb{R}}(I)$ ist eine Unteralgebra von $C_{\mathbb{R}}(I)$.*

Beweis durch vollständige Induktion:

(i) Für $n = 1$ folgen (i) und (ii) aus den Rechenregeln für differenzierbare Funktionen.

(ii) Wir nehmen an, dass (i) und (ii) für $n \in \mathbb{N}$ gelten. Für $f, g \in C^{n+1}_{\mathbb{R}}(I)$ folgt dann aus Satz 7.5

$$\frac{d}{dx}\frac{d^n}{dx^n}(f \cdot g) = \frac{d}{dx}\sum_{k=0}^{n}\binom{n}{k} f^{(k)} g^{(n-k)} = \sum_{k=0}^{n}\binom{n}{k}\left(f^{(k+1)} g^{(n-k)} + f^{(k)} g^{(n+1-k)}\right)$$

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

$$\text{weil} \quad \binom{n}{k-1} + \binom{n}{k} = \frac{n!}{k!(n+1-k)!}(k + n + 1 - k) = \binom{n+1}{k}.$$

Also gilt (i) und (ii) auch für $(n + 1)$. **q.e.d.**

Aus der Rechenregel und der Kettenregel folgt auch

Korollar 7.35. **(i)** *Die Komposition von n-mal stetig differenzierbaren Funktionen ist wieder n-mal stetig differenzierbar.*

(ii) *Die Umkehrfunktion einer n-mal stetig differenzierbaren bijektiven Funktion ist n-mal stetig differenzierbar, wenn die Ableitung keine Nullstellen hat.* **q.e.d.**

Definition 7.36. *(Taylor-Polynom) Sei $f : I \to \mathbb{R}$ in x_0 n-mal differenzierbar, d.h. es gibt eine offene Umgebung $O \subset I$ von x_0, so dass die Einschränkung von f auf O in $C^{n-1}(O)$ liegt, und $f^{(n-1)}$ in x_0 differenzierbar ist. Dann heißt*

$$T_{n,x_0}(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2} f^{(2)}(x_0) + \ldots + \frac{(x - x_0)^n}{n!} f^{(n)}(x_0)$$

das Taylorpolynom von f der Ordnung n in x_0.

Offenbar hat das Taylorpolynom der Ordnung n in x_0 die gleichen Ableitungen bis zur Ordnung n wie f an dem Punkt x_0. Es ist das eindeutig bestimmte Polynom vom Grad n, das an der Stelle x_0 die Ableitungen $f(x_0), f^{(1)}(x_0), \ldots, f^{(n)}(x_0)$ besitzt:

$$p(x) = c_0 + c_1(x - x_0) + \ldots + c_n(x - x_0)^n \Longrightarrow p(x_0) = c_0, p^{(1)}(x_0) = c_1, \ldots, p^{(n)}(x_0) = n!c_n.$$

Satz 7.37. *(Taylor-Formel) Sei $f \in C^n_{\mathbb{R}}((a,b))$. Wenn $f^{(n+1)}(x)$ für alle $x \in (a,b)$ existiert, dann gibt es für jedes $x_0 \neq x \in (a,b)$ ein $\xi \in (x_0,x)$ bzw. $\xi \in (x,x_0)$, so dass*

$$f(x) = T_{n,x_0}(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1} \quad \textit{gilt.}$$

Beweis: Sei $x \in (a,b)$ und definiere $g(t) = \sum\limits_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k$ für $t \in (a,b)$. Dann gilt

$$g'(t) = \sum_{k=0}^{n} \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \sum_{k=1}^{n} \frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1} = \frac{f^{(n+1)}(t)}{n!}(x-t)^n.$$

Außerdem sei $h(t) = (x-t)^{n+1}$ und $h'(t) = -(n+1)(x-t)^n$. Dann folgt aus dem verallgemeinerten Mittelwertsatz, dass es ein $\xi \in (x_0,x)$ bzw. $\xi \in (x_0,x)$ gibt mit

$$(g(x)-g(x_0))h'(\xi) = (h(x)-h(x_0))g'(\xi).$$

Es gilt aber $g(x)-g(x_0) = f(x) - T_{n,x_0}(x)$ und $h(x)-h(x_0) = -(x-x_0)^{n+1}$. Also folgt $f(x) - T_{n,x_0}(x) = \dfrac{f^{(n+1)}(\xi)(x-\xi)^n(x-x_0)^{n+1}}{n!(n+1)(x-\xi)^n} = \dfrac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$. **q.e.d.**

Definition 7.38. *(Taylorreihe) Für $f \in C^\infty((a,b))$ und $x_0 \in (a,b)$ heißt die Potenzreihe $\sum\limits_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$ Taylorreihe von f in x_0.*

Für $f \in C^\infty((a,b))$ und $x_0, x \in (a,b)$ gibt es drei Möglichkeiten:

(i) Die Taylorreihe von f in x_0 konvergiert an dem Punkt x gegen $f(x)$.

(ii) Die Taylorreihe von f in x_0 konvergiert an dem Punkt x, aber nicht gegen $f(x)$.

(iii) Die Taylorreihe von f in x_0 konvergiert an dem Punkt x nicht.

Korollar 7.39. *Sei $f \in C^\infty_{\mathbb{R}}((a,b))$ und $x_0 \in (a,b)$. Dann konvergiert die Taylorreihe von f in x_0 an dem Punkt $x \in (a,b)$ gegen $f(x)$, wenn auf $\xi \in (x_0,x)$ bzw. (x,x_0) die Folge $\left(\dfrac{|f^n(\xi)|}{n!}|x-x_0|^n\right)_{n\in\mathbb{N}_0}$ gleichmäßig gegen Null konvergiert.* **q.e.d.**

Beispiel 7.40.

$$\begin{aligned} f(x) &= \begin{cases} e^{-\frac{1}{x^2}} & \textit{für } x>0 \\ 0 & \textit{für } x \le 0 \end{cases} \\ f^{(n)}(x) &= \begin{cases} e^{-\frac{1}{x^2}} \cdot \textit{ Polynom vom Grad } 3n \textit{ von } \frac{1}{x} & \textit{für } x>0 \\ 0 & \textit{für } x \le 0 \end{cases} \end{aligned}$$

Wegen Beispiel 7.17 gilt $1+x \leq e^x$ für alle $x \in \mathbb{R}$. Dann folgt für alle $x \in \mathbb{R}\setminus\{0\}$

$$|x| \leq e^{|x|-1} \quad \Rightarrow \quad \frac{1}{|x|^n} \leq \exp\left(\frac{n}{|x|} - n\right) \quad \Rightarrow \quad \frac{\exp\left(\frac{-1}{x^2}\right)}{|x|^n} \leq \exp\left(\frac{-1+n|x|-nx^2}{x^2}\right).$$

Für alle $n \in \mathbb{N}$ ist dann $x \to \begin{cases} \frac{1}{x^n} e^{-\frac{1}{x^2}} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0 \end{cases}$ stetig, und $f \in C^\infty_{\mathbb{R}}(\mathbb{R})$. Und alle Ableitungen von f verschwinden bei $x_0 = 0$. Also verschwindet die Taylorreihe von f bei $x_0 = 0$ identisch.

Satz 7.41. *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen in $C^1_{\mathbb{R}}((a,b))$ eines beschränkten Intervalles (a,b), die für ein $x_0 \in (a,b)$ punktweise konvergiert. Wenn die Folge $(f'_n)_{n\in\mathbb{N}}$ außerdem gleichmäßig gegen g konvergiert, dann konvergiert $(f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen eine Funktion $f \in C^1_{\mathbb{R}}((a,b))$ und es gilt $f' = g$.*

Beweis: Wegen dem Mittelwertsatz gilt für alle $x, x_0 \in (a,b)$ und alle $n, m \in \mathbb{N}$

$$|f_n(x) - f_m(x)| \leq |f_n(x_0) - f_m(x_0)| + |x - x_0| \sup_{y\in(a,b)} |f'_n(y) - f'_m(y)|.$$

Wegen Satz 5.27 konvergiert $(f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen eine Funktion $f \in C_{\mathbb{R}}((a,b))$. Wegen dem Mittelwertsatz gibt es für alle $x, y \in (a,b)$ eine Folge $(\xi_n)_{n\in\mathbb{N}} \in (x,y)$ bzw. (y,x), so dass für alle $n \in \mathbb{N}$ gilt $f_n(x) - f_n(y) = (x-y)f'_n(\xi_n)$. Wegen dem Auswahlprinzipien von Bolzano-Weierstraß gibt es eine konvergente Teilfolge von $(\xi_n)_{n\mathbb{N}}$ mit Grenzwert $\xi \in [x,y]$ bzw. $[y,x]$. Wegen

$$|f'_n(\xi_n) - g(\xi)| \leq |f_n(\xi_n) - g(\xi_n)| + |g(\xi_n) - g(\xi)|,$$

und weil g wegen Satz 5.27 stetig ist, konvergiert die Folge $(f'_n(\xi_n))_{n\mathbb{N}}$ gegen $g(\xi)$. Also konvergiert $(f_n(x))_{n\in\mathbb{N}}$ gegen $f(x) = f(y) + (x-y)g(\xi)$. Aus der Stetigkeit von g folgt, dass f bei y differenzierbar ist und $g(y)$ die Ableitung $f'(y)$ ist. **q.e.d.**

Korollar 7.42. *Sei $(\sum f'_n)_{n\in\mathbb{N}}$ eine gleichmäßig konvergente Reihe in $C_{\mathbb{R}}(I)$ auf einem beschränktem Intervall I und $(\sum f(x_0))_{n\in\mathbb{N}}$ konvergent mit $x_0 \in I$. Dann konvergiert $(\sum f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen eine Funktion $f \in C^1_{\mathbb{R}}(I)$ und $(\sum f'_n)_{n\in\mathbb{N}}$ gegen f'.* **q.e.d.**

Korollar 7.43. *(Satz von Borel) Sei $(a_n)_{n\in\mathbb{N}_0}$ eine beliebige Folge in $\mathbb{R}$. Dann gibt es eine Funktion $f \in C^\infty_{\mathbb{R}}(\mathbb{R})$ mit kompaktem Träger in $(-2,2)$, deren Taylorreihe bei $x_0 = 0$ gleich $\sum\limits_{n=0}^{\infty} \frac{a_n}{n!} x^n$ ist. Im Allgemeinen konvergiert die Taylorreihe für $x \neq 0$ nicht.*

Beweis*:

$$\text{Sei} \quad h(x) = \begin{cases} \exp\left(\exp\left(\frac{-1}{(|x|-1)^2}\right) \cdot \frac{-1}{(|x|-2)^2}\right) & \text{für } 1 < |x| < 2 \\ 1 & \text{für } |x| < 1 \\ 0 & \text{für } 2 < |x| \end{cases}$$

Dann ist $h \in C^\infty_\mathbb{R}(\mathbb{R})$ eine 'Hutfunktion', also eine Funktion mit kompaktem Träger in $[-2, 2]$, die auf $[-1, 1]$ identisch gleich 1 ist. Für alle $n \in \mathbb{N}$ existiert dann eine Konstante $M_n > 0$

$$M_n = \max\{\|h_n\|_\infty, \|h_n'\|_\infty, \ldots, \|h_n^{(n-1)}\|_\infty\}$$

mit $h_n : \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^n \cdot h(x)$. Dann sei für eine beliebige reelle Folge $(a_n)_{n\in\mathbb{N}_0}$

$$C_n = |a_n|M_n + 1 \quad \text{und} \qquad f_n(x) = \frac{a_n}{n!C_n^n} h_n(C_n \cdot x) \qquad \text{für alle } n \in \mathbb{N}_0.$$

Für alle $n, m \in \mathbb{N}_0$ gilt dann $f_n^{(m)}(0) = \begin{cases} 0 & \text{für } m \neq n \\ a_n & \text{für } m = n. \end{cases}$ Außerdem gilt

$$\|f_n^{(m)}\|_\infty = \frac{|a_n|C_n^m}{n!C_n^n}\|h_n^{(m)}\|_\infty \leq \frac{|a_n|M_nC_n^m}{n!C_n^n} < \frac{C_n^{m+1}}{n!C_n^n} \leq \frac{1}{n!} \qquad \text{für alle } n > m \in \mathbb{N}_0.$$

Also konvergiert für alle $m \in \mathbb{N}_0$ $(\Sigma f_n^{(m)})_{n\in\mathbb{N}_0}$ gleichmäßig. Wegen Korollar 7.42 konvergieren also für alle $m \in \mathbb{N}_0$ die Reihen $(\Sigma f_n)_{n\in\mathbb{N}_0}, (\Sigma f_n')_{n\in\mathbb{N}_0}, \ldots, (\Sigma f_n^{(m)})_{n\in\mathbb{N}_0}$ gleichmäßig gegen $f, f', \ldots, f^{(m)}$. Also ist der Grenzwert $f = \sum_{n=0}^\infty f_n$ eine Funktion in $C^\infty_\mathbb{R}(\mathbb{R})$ und es gilt $f^{(m)}(0) = a_m$ für alle $m \in \mathbb{N}_0$. **q.e.d.**

Korollar 7.44. *Für jede Potenzreihenfunktion $f(x) = \sum\limits_{n=0}^\infty a_nx^n$ mit Konvergenzradius $R > 0$ und jedes $|x_0| < R$ hat die Taylorreihe von $f(x)$ in x_0 einen Konvergenzradius nicht kleiner als $R - |x_0|$, und konvergiert auf dem Bereich $|x - x_0| < R - |x_0|$ gegen f.*

Beweis: Die Potenzreihenfunktion $(\sum_{n=0}^\infty a_nx^n)' = \sum_{n=0}^\infty(n + 1)a_{n+1}x^n$ hat wegen $\lim_{n\to\infty} \sqrt[n]{n} = 1$ den gleichen Konvergenzradius wie $(\sum a_nx^n)_{n\in\mathbb{N}_0}$. Wegen Korollar 7.42 folgt dann, dass f differenzierbar ist und f' gegeben ist durch $\sum_{n=0}^\infty(n+1)a_{n+1}x^n$. Dann folgt die Aussage aus dem Identitätssatz für Potenzreihenfunktionen (ii). **q.e.d.**

Satz 7.45. *(Abelscher Grenzwertsatz) Wenn die Potenzreihe $(\sum a_nx^n)_{n\in\mathbb{N}_0}$ für ein $x \in \mathbb{K}$ konvergiert, dann konvergiert die Potenzreihenfunktion $t \mapsto \sum_{n\in\mathbb{N}_0} a_n(tx)^n$ auf $t \in [0, 1]$ gleichmäßig gegen eine stetige Funktion auf $t \in [0, 1]$.*

Beweis: Indem wir a_n durch a_nx^n ersetzen können wir x weglassen. Zur Abkürzung setzen wir $S_{m,k} = \sum_{n=m+1}^{m+k} a_n$. Wegen dem Cauchykriterium für Reihen gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ so dass $|S_{m,k}| < \epsilon$ für alle $m \geq N$ und alle $k \in \mathbb{N}$ gilt. Dann folgt

$$\sum_{n=m+1}^{m+k} a_nt^n = S_{m,1}t^{m+1} + (S_{m,2} - S_{m,1})t^{m+2} + \ldots + (S_{m,k} - S_{m,k-1})t^{m+k}$$
$$= S_{m,1}(t^{m+1} - t^{m+2}) + S_{m,2}(t^{m+2} - t^{m+3}) + \ldots + S_{m,k-1}(t^{m+k-1} - t^{m+k}) + S_{m,k}t^{m+k}.$$

Für $t \in [0,1]$ sind die hinteren Faktoren $t^{m+1}-t^{m+2}, t^{m+2}-t^{m+3}, \ldots, t^{m+k-1}-t^{m+k}, t^{m+k}$ alle nicht negativ und ihre Summe gleich $t^{m+1} \leq 1$. Aus $|S_{m,k}| < \epsilon$ folgt also

$$\left|\sum_{n=m+1}^{m+k} a_n t^k\right| < \epsilon(t^{m+1}-t^{m+2}+t^{m+2}-\ldots-t^{m+k}+t^{m+k}) = \epsilon t^{m+1} \leq \epsilon \quad \text{für alle } t \in [0,1].$$

Also konvergiert die Potenzreihe auf $t \in [0,1]$ gleichmäßig, und damit wegen Satz 5.27 gegen eine stetige Funktion. **q.e.d.**

Definition 7.46. *Eine Funktion $f \in C^\infty((a,b))$ heißt reellanalytisch bei x_0, falls die Taylorreihe bei x_0 einen Konvergenzradius größer als Null hat und auf einer Umgebung von x_0 gegen $f(x)$ konvergiert.*

Also sind alle Potenzreihenfunktionen im Inneren ihres Kovergenzbereiches reellanalytisch. Umgekehrt sind alle reellanalytischen Funktionen Potenzreihenfunktionen.

Beispiel 7.47. (i) *Die Funktionen* $\exp, \sin, \cos, x \mapsto a^x$ *und alle Polynome sind reellanalytische Funktionen auf ganz $\mathbb{R}$.*
(ii) *Wegen Satz 4.26 (iv) gibt es für jede Potenzreihenfunktion f mit Konvergenzradius $R > 0$, die bei $x = 0$ nicht verschwindet, eine Umgebung von 0, auf der $|\frac{f(0)-f(x)}{f(0)}| \leq L < 1$ gilt. Dort konvergiert $\frac{1}{f} = \sum_{n=0}^\infty \frac{1}{f(0)}(\frac{f(0)-f}{f(0)})^n$ als Potenzreihenfunktion. Aus dem Identitässatz für Potenzreihenfunktionen folgt, dass der Quotient zweier Potenzreihenfunktionen reellanalytisch ist, solange beide Potenzreihenfunktionen absolut konvergieren und der Nenner nicht verschwindet. Also sind* tan *und* cot *und alle rationalen Funktionen auf dem Definitionsbereich reellanalytisch.*
(iii) *Für $\alpha \in \mathbb{R}$ und $x_0 \in \mathbb{R}^+$ (für $\alpha \in \mathbb{Z}$ auch $x_0 \in \mathbb{R}^-$) hat die Potenzreihenfunktion*

$$x \mapsto f(x) = \sum_{n=0}^\infty \binom{\alpha}{n} x_0^{\alpha-n} x^n \quad \textit{mit} \qquad \binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n(n-1)\cdots 1}$$

wegen dem Quotiententest den Konvergenzradius $R = \frac{1}{\lim\limits_{n\to\infty} \frac{|\alpha-n|}{|x_0|(n+1)}} = |x_0|$. *Die Ableitung ist* $x \mapsto \sum_{n=0}^\infty \binom{\alpha}{n} n x_0^{\alpha-n} x^{n-1} = \alpha \sum_{n=0}^\infty \binom{\alpha-1}{n} x_0^{\alpha-1-n} x^n$. *Wegen*

$$\binom{\alpha-1}{n} + \binom{\alpha-1}{n-1} = \frac{(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!}(\alpha-n+n) = \binom{\alpha}{n}$$

ist $(x_0+x)\sum_{n=0}^\infty \binom{\alpha-1}{n} x_0^{\alpha-1-n} x^n = \sum_{n=0}^\infty \binom{\alpha}{n} x_0^{\alpha-n} x^n$. *Dann erfüllt f die Differentialgleichung $(x_0+x)f' = \alpha f$ mit $f(0) = x_0^\alpha$. Also verschwindet die Ableitung der*

Funktion $x \mapsto \ln\left(\frac{f(x)}{(x+x_0)^\alpha}\right)$ *und diese Funktion selber bei* $x = 0$. *Dann folgt aus Satz 7.15, dass für alle* $|x| < x_0$ *gilt* $f(x) = (x+x_0)^\alpha$. *Also sind für* $\alpha \in \mathbb{R}$ *die Funktionen* $x \mapsto x^\alpha$ *auf* $\mathbb{R}^+$ *und für* $\alpha \in \mathbb{Z}$ *auf* $\mathbb{R} \setminus \{0\}$ *reellanalytisch.*

(iv) *Für alle* $x_0 \in \mathbb{R}^+$ *hat die Potenzreihenfunktion* $x \mapsto -\sum_{n=1}^{\infty} \frac{(-x)^n}{nx_0^n}$ *im Konveregemzbereich* $|x| < x_0$ *die Ableitung* $f'(x) = \sum_{n=0}^{\infty} \frac{(-x)^n}{x_0^{n+1}} = \frac{1}{x+x_0}$. *Also stimmt sie mit der Funktion* $\ln(x+x_0) - \ln(x_0)$ *überein. Deshalb sind sowohl* $\ln$ *also auch* $\log_a$ *auf* $\mathbb{R}^+$ *reellanalytisch. Insbesondere folgt aus dem Abelschen Grenzwertsatz* $\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -\ln(2)$.

(v)] *Die Ableitungen der Umkehrfunktionen der trigonometrischen Funktionen sind wegen (iii) im Inneren ihrer Definitionsbereiche alle reellanalytisch. Wegen Satz 7.15 sind sie selber dann auch reellanalytisch. Für alle* $|x| < 1$ *gilt*

$$\arcsin(x) = \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} \frac{(-1)^n x^{2n+1}}{2n+1} \qquad \arccos(x) = \frac{\pi}{2} - \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} \frac{(-1)^n x^{2n+1}}{2n+1}$$
$$\arctan(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1} \qquad \operatorname{arccot}(x) = \frac{\pi}{2} - \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1}$$

Wegen Beispiel 7.17 gilt für $x > -1$ *auch* $x \geq \ln(1+x)$. *Dann folgt*

$$\begin{aligned} \ln\left((-1)^n \binom{-\frac{1}{2}}{n}\right) &= \ln\left(\frac{1}{2}\frac{3}{4}\cdots\frac{2n-1}{2n}\right) = \ln\left(\left(1-\frac{1}{2}\right)\left(1-\frac{1}{4}\right)\cdots\left(1-\frac{1}{2n}\right)\right) \\ &\leq -\frac{1}{2}\left(1+\frac{1}{2}+\ldots+\frac{1}{n}\right) \leq -\frac{1}{2}\ln\left(\frac{2}{1}\frac{3}{2}\cdots\frac{n+1}{n}\right) = \ln\left(\frac{1}{\sqrt{n+1}}\right). \end{aligned}$$

Dann konvergieren aber die ersten beiden Potenzreihen auch für $x = \pm 1$ *und die letzten beiden wegen der alternierenden Reihe von Leibniz. Wegen dem Abelschen Grenzwertsatz gelten die obigen Gleichungen dann auch für* $x = \pm 1$. *Insbesondere ist*

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

(vi) *Die Funktionen* $\mathbb{R} \to \mathbb{R}, x \mapsto \begin{cases} e^{-\frac{1}{x^2}} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0 \end{cases}$ *ist reellanalytisch auf* $\mathbb{R} \setminus \{0\}$, *aber nicht bei* $x_0 = 0$.

Aus dem Identitätssatz für Potenzreihenfunktionen folgt nun

Korollar 7.48. *Zwei reellanalytische Funktionen* $f, g \in C^\infty((a,b))$ *stimmen auf* (a,b) *überein, wenn ihre Taylorreihen für ein* $x_0 \in (a,b)$ *übereinstimmen.* **q.e.d.**

Kapitel 8

Das Riemannintegral

8.1 Riemannintegrable Funktionen

In diesem Abschnitt betrachten wir nur beschränkte Funktionen $f \in B_{\mathbb{R}}([a,b])$ auf einem beschränkten abeschlossenen Intervall. Das Ziel ist für solche Funktionen den Flächeninhalt zwischen den Graphen der Funktion und der x-Achse zu definieren. Dabei werden wir diese Fläche durch eine disjunkte Vereinigung von Rechtecken annähern.

Definition 8.1. *(Partition) Eine Partition p von $[a,b]$ ist eine endliche geordnete Menge $\{x_0,\ldots,x_n\}$ von Punkten $a = x_0 < x_1 < \ldots < x_n = b$ in $[a,b]$. Die Feinheit der Partition p ist dann $\|p\| = \max\{\Delta x_1,\ldots,\Delta x_n\}$ mit $\Delta x_i = x_i - x_{i-1}$ für alle $i = 1,\ldots,n$. $\mathcal{P}[a,b]$ bezeichnet die Menge aller Partitionen von $[a,b]$.*

Für eine Funktion $f \in B_{\mathbb{R}}([a,b])$ und eine Partition $p \in \mathcal{P}[a,b]$ seien

$$\begin{aligned} m &= \inf\{f(x) \mid x \in [a,b]\} \\ M &= \sup\{f(x) \mid x \in [a,b]\} \\ m_i &= \inf\{f(x) \mid x \in [x_{i-1},x_i]\} \text{ für alle } i = 1,\ldots,n \\ M_i &= \sup\{f(x) \mid x \in [x_{i-1},x_i]\} \text{ für alle } i = 1,\ldots,n \end{aligned}$$

Definition 8.2. *(Untersummen und Obersummen) Dann heißen*

$$s(p,f) = \sum_{i=1}^{n} m_i \Delta x_i \text{ und } S(p,f) = \sum_{i=1}^{n} M_i \Delta x_i$$

die Untersumme und Obersumme von f bezüglich der Partition p.

Offenbar gilt $$m(b-a) \le s(p,f) \le S(p,f) \le M(b-a).$$

Definition 8.3. *(Verfeinerung) $p' \in \mathcal{P}[a,b]$ heißt Verfeinerung von $p \in \mathcal{P}[a,b]$, wenn $p' \supset p$. Offenbar gibt es für endlich viele Partitionen $p_1,\ldots,p_n \in \mathcal{P}[a,b]$ eine gemeinsame Verfeinerung $p' = p_1 \cup \ldots \cup p_n \in \mathcal{P}[a,b]$.*

Lemma 8.4. **(i)** *Wenn* $p \subset p'$ *gilt* $s(p,f) \le s(p',f)$ *und* $S(p',f) \le S(p,f)$.

(ii) *Für* $p, p' \in \mathcal{P}[a,b]$ *gilt* $s(p,f) \le S(p',f)$.

Beweis:(i) Die Verfeinerung p' von p besteht aus einer Partition jedes Teilintervalles $[x_{i-1}, x_i]$ von p. Dann folgt (i) aus den Ungleichungen

$$m(b-a) \le s(p,f) \quad \text{und} \quad S(p,f) \le M(b-a).$$

(ii) Sei $p'' = p \cup p'$. Dann folgen aus (i) die Ungleichungen

$$s(p,f) \le s(p'',f) \le S(p'',f) \le S(p,f) \quad s(p',f) \le s(p'',f) \le S(p'',f) \le S(p',f).$$

Daraus folgt dann (ii). **q.e.d.**

Definition 8.5. *(Unterintegral und Oberintegral) Für* $f \in B_{\mathbb{R}}([a,b])$ *heißt* $\underline{\int} f = \sup_{p \in \mathcal{P}[a,b]} s(p,f)$ *Unterintegral und* $\overline{\int} f = \inf_{p \in \mathcal{P}[a,b]} S(p,f)$ *Oberintegral von* f.

Offenbar gilt $$\underline{\int} f \le \overline{\int} f.$$

Definition 8.6. *Eine Funktion* $f \in B_{\mathbb{R}}([a,b])$ *heißt riemannintegrabel, wenn gilt* $\underline{\int} f = \overline{\int} f$. *Diese Zahl heißt dann das Riemannintegral von* f *über* $[a,b]$: $\int_a^b f dx$. *Die Menge aller riemannintegrablen Funktionen auf* $[a,b]$ *bezeichnen wir mit* $\mathcal{R}[a,b]$.

Aufgrund der Definition von $s(p,f)$ und $S(p,f)$ liegt der Flächeninhalt zwischen dem Graphen von f und der x-Achse in dem Intervall $[s(p,f), S(p,f)]$. Deshalb interpretieren wir für $f \in \mathcal{R}[a,b]$ das Riemannintegral $\int_a^b f(x)dx$ als diesen Flächeninhalt.

8.2 Kriterien von Darboux und Riemann

Satz 8.7. *(Darboux)* $f \in \mathcal{R}[a,b]$ *genau dann, wenn es für jedes* $\epsilon > 0$ *ein* $p \in \mathcal{P}[a,b]$ *gibt mit* $S(p,f) - s(p,f) < \epsilon$.

Beweis: Sei $f \in \mathcal{R}[a,b]$. Dann gibt es für jedes $\epsilon > 0$ Partitionen $p', p'' \in \mathcal{P}[a,b]$ mit $\int_a^b f(x)dx - s(p',f) < \frac{\epsilon}{2}$ und $S(p'',f) - \int_a^b f(x)dx < \frac{\epsilon}{2}$. Dann folgt für $p = p' \cup p''$

$$S(p,f) - s(p,f) \le S(p'',f) - \int\limits_a^b f(x)dx + \int\limits_a^b f(x)dx - s(p',f) < \epsilon.$$

Wenn es für jedes $\epsilon > 0$ ein $p_\epsilon \in \mathcal{P}[a,b]$ gibt mit $S(p_\epsilon, f) - s(p_\epsilon, f) < \epsilon$ dann folgt

$$0 \le \inf_{p \in \mathcal{P}[a,b]} S(p,f) - \sup_{p \in \mathcal{P}[a,b]} s(p,f) \le S(p_\epsilon, p) - s(p_\epsilon, f) < \epsilon \quad \text{für alle } \epsilon > 0.$$

Also gilt dann auch $$\underline{\int} f = \overline{\int} f.$$ **q.e.d.**

Beispiel 8.8. *Sei* $f(x) = \begin{cases} 1 & \text{für } x \in [a,b] \cap \mathbb{Q} \\ 0 & \text{für } x \in [a,b] \text{ und } x \notin \mathbb{Q}. \end{cases}$
Dann gilt für alle $p \in \mathcal{P}[a,b]$

$$S(p,f) - s(p,f) = \sum_{i=1}^{n} (M_i - m_i)\Delta x_i = \sum_{i=1}^{n} \Delta x_i = b - a > 0.$$

Also ist $\underline{\int} f = 0$ *und* $\overline{\int} f = b - a$ *und* $f \notin \mathcal{R}[a,b]$.

Definition 8.9. *(Riemannsummen) Für* $p \in \mathcal{P}[a,b]$ *sei* $\xi = (\xi_1, \ldots, \xi_n)$ *eine Wahl von Zwischenpunkten* $\xi_i \in [x_{i-1}, x_i]$ *für alle* $i = 1, \ldots, n$. *Dann heißt*

$$R(p,f,\xi) = \sum_{i=1}^{n} f(\xi_i)\Delta x_i$$

Riemannsumme von f *bezüglich der Partition* p *und der Zwischenpunkte* ξ.

Satz 8.10. *(Kriterium von Riemann) Eine beschränkte Funktion* f *auf einem kompakten Intervall* $[a,b]$ *ist genau dann riemannintegrabel, wenn es ein* $A \in \mathbb{R}$ *gibt, so dass für jedes* $\epsilon > 0$ *es ein* $\delta > 0$ *gibt mit der Eigenschaft: Für alle* $p \in \mathcal{P}[a,b]$ *mit* $\|p\| < \delta$ *und alle entsprechenden Zwischenpunkte* ξ *gilt*

$$|R(p,f,\xi) - A| < \epsilon. \quad \text{Wenn das Krierium erfüllt ist, dann gilt } A = \int_a^b f(x)dx.$$

Beweis: Offenbar erfüllt jede Funktion, die das Kriterium von Riemann erfüllt auch das Kriterium von Darboux. Also genügt es zu zeigen, dass auch jede Funktion das Kriterium von Riemann erfüllt, die das von Darboux erfüllt. Sei also f eine Funktion, die das Kriterium von Darboux erfüllt. Dann gibt es für jedes $\epsilon > 0$ ein $p \in \mathcal{P}[a,b]$, so dass gilt $S(p,f) - s(p,f) < \frac{\epsilon}{2}$. Sei n die Anzahl der Teilintervalle von p und $\|f\|_\infty$ das Supremum von $|f(x)|$ auf $x \in [a,b]$, und

$$\delta = \min\left\{\frac{\epsilon}{4n\|f\|_\infty + 1}, \Delta x_1, \ldots, \Delta x_n\right\}.$$

Jedes Teilintervall einer Partition $p' \in \mathcal{P}[a,b]$ mit $\|p'\| < \delta$ ist entweder in einem Teilintervall von p enthalten, oder in der Vereinigung von zwei benachbarten Teilintervallen von p. Also gibt es höchstens n Teilintervalle von p', die nicht in einem Teilintervall von p enthalten sind. Dann folgt aber

$$S(p',f) - s(p',f) \le S(p,f) - s(p,f) + 2\|f\|_\infty \cdot n \cdot \delta < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Dann gilt aber auch für alle Zwischenpunkte ξ

$$\left|R(p', f, \xi) - \int_a^b f(x)dx\right| \leq S(p', f) - s(p', f) < \epsilon. \qquad \textbf{q.e.d.}$$

Korollar 8.11. *Sei $f \in \mathcal{R}[a, b]$. Dann gilt für alle $t \in [0, 1]$*

$$\int_a^b f(x)dx = \lim_{n\to\infty} \frac{b-a}{n} \sum_{i=1}^{n} f\left(a + \frac{i-t}{n}(b-a)\right).$$

Beweis: Die Partitionen $p_n \in \mathcal{P}[a, b]$ mit $p_n = \left\{x_i = a + \frac{i}{n}(b - a) \mid i = 0, \ldots, n\right\}$ mit den Zwischenpunkten $\xi_i = a + \frac{i-t}{n}(b - a)$ für $i = 1, \ldots, n$ erfüllt $\|p_n\| = \frac{b-a}{n}$. Also folgt die Aussage aus dem Kriterium von Riemann. **q.e.d.**

Korollar 8.12.* *Seien $f, g \in \mathcal{R}[a, b]$. Wenn f und g auf einer dichten Teilmenge von $[a, b]$ (z.B. $\mathbb{Q} \cap [a, b]$) übereinstimmen, dann gilt $\int_a^b f(x)dx = \int_a^b g(x)dx$.*

Beweis*: Weil jedes Teilintervall einer beliebigen Partition $p \in \mathcal{P}[a, b]$ immer Elemente einer dichten Teilmenge von $[a, b]$ enthält, können die Zwischenpunkte immer aus einer dichten Teilmenge gewählt werden. **q.e.d.**

Satz 8.13. *(Eigenschaften des Riemannintegrals)*

(i) *$\mathcal{R}[a, b]$ ist eine Unteralgebra von $B_{\mathbb{R}}([a, b])$ die $C_{\mathbb{R}}([a, b])$ enthält. Die Abbildung $\mathcal{R}[a, b] \to \mathbb{R}, \quad f \mapsto \int_a^b f(x)dx$ ist $\mathbb{R}$-linear.*

(ii) *$\mathcal{R}[a, b]$ enthält die monotonen Funktionen, und mit $f \in \mathcal{R}[a, b]$ auch $|f| \in \mathcal{R}[a, b]$.*

(iii) *Monotonie: Für $f, g \in \mathcal{R}[a, b]$ folgt aus $f \leq g$ (d.h. $f(x) \leq g(x)$ für alle $x \in [a, b]$) $\int_a^b f(x)dx \leq \int_a^b g(x)dx$. Insbesondere gilt $|\int_a^b f(x)dx| \leq \int_a^b |f(x)|dx \leq (b - a)\|f\|_\infty$.*

(iv) *Normierung: $\int_a^b 1dx = b - a$.*

(v) *Stetigkeit: $f \in \mathcal{R}[a, b]$ und $g \in C_{\mathbb{R}}(\mathbb{R})$, dann ist $g \circ f \in \mathcal{R}[a, b]$.*

(vi) *Intervall Additivität: Für jedes $c \in (a, b)$ gilt:*

$$f \in \mathcal{R}[a, b] \Longleftrightarrow f \in \mathcal{R}[a, c] \cap \mathcal{R}[c, b] \text{ und } \int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

(vii) *Sei $f \in \mathcal{R}[a,b]$ und $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ Folgen in $[a,b]$ mit $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$. Dann konvergiert $(\int_{a_n}^{b_n} f(x)dx)_{n\in\mathbb{N}}$ gegen $\int_a^b f(x)dx$.*

(viii) *Gleichmäßige Konvergenz und Stetigkeit des Riemannintegrals: Der Grenzwert f einer gleichmäßig konvergenten Folge $(f_n)_{n\in\mathbb{N}}$ in $\mathcal{R}[a,b]$ liegt auch in $\mathcal{R}[a,b]$ und die Folge $(\int_a^b f_n(x)dx)_{n\in\mathbb{N}}$ konvergiert dann gegen $\int_a^b f(x)dx$.*

Beweis:(i) Für $f, g \in B_{\mathbb{R}}([a,b])$ und $p \in [a,b]$ gilt

$$S(p, f+g) \leq S(p,f) + S(p,g) \text{ und } \qquad -s(p,f+g) \leq -s(p,f) - s(p,g)$$

Daraus und aus $f(x)g(x) - f(y)g(y) = g(x)(f(x) - f(y)) + f(y)(g(x) - g(y))$ folgt

$$\begin{aligned} S(p,f+g) - s(p,f+g) &\leq S(p,f) - s(p,f) + S(p,g) - s(p,g) \\ S(p,fg) - s(p,fg) &\leq \|g\|_\infty(S(p,f) - s(p,f)) + \|f\|_\infty(S(p,g) - s(p,g)) \end{aligned}$$

Aus $f, g \in \mathcal{R}[a,b]$ folgt also wegen dem Darbouxkriterium $f+g, fg \in \mathcal{R}[a,b]$.

Wegen Satz 5.22 ist jede Funktion $f \in C_{\mathbb{R}}([a,b])$ gleichmäßig stetig, d.h. es gibt für jedes $\epsilon > 0$ ein $\delta > 0$, so dass aus $|x - x'| < \delta$ folgt $|f(x) - f(x')| < \frac{\epsilon}{b-a}$. Dann gilt für alle $p \in \mathcal{P}[a,b]$ mit $\|p\| < \delta$ auch $S(p,f) - s(p,f) < \sum_{i=1}^n \frac{\epsilon}{b-a}\Delta x_i = \epsilon$. Also folgt aus dem Kriterium von Darboux $f \in \mathcal{R}[a,b]$.

Die Linearität des Riemannintegrals folgt aus der Linearität der Riemannsummen und den Rechenregeln für Folgen.

(ii) Sei f monoton steigend ist. Dann gilt $S(p,f) - s(p,f) =$

$$= \sum_{i=1}^n (f(x_i) - f(x_{i-1}))\Delta x_i \leq \|p\| \sum_{i=1}^n f(x_i) - f(x_{i-1}) = \|p\|(f(b) - f(a)).$$

Wenn also $\|p\| < \frac{\epsilon}{f(b)-f(a)+1}$ folgt $S(p,f) - s(p,f) < \epsilon$. Wegen dem Kriterium von Darboux gehört dann f zu $\mathcal{R}[a,b]$. Analoges gilt für monoton fallende f.

Für $f \in \mathcal{R}[a,b]$ seien $f^+ = \max\{f, 0\}$ und $f^- = \max\{-f, 0\}$. Dann folgt

$$\begin{aligned} 0 &\leq \sup_{x\in[x_{i-1},x_i]} f^+(x) - \inf_{x\in[x_{i-1},x_i]} f^+(x) &&\leq \sup_{x\in[x_{i-1},x_i]} f(x) - \inf_{x\in[x_{i-1},x_i]} f(x) \\ 0 &\leq \sup_{x\in[x_{i-1},x_i]} f^-(x) - \inf_{x\in[x_{i-1},x_i]} f^-(x) &&\leq \sup_{x\in[x_{i-1},x_i]} f(x) - \inf_{x\in[x_{i-1},x_i]} f(x) \end{aligned}$$

Also gilt für alle $p \in \mathcal{P}[a,b]$ auch $S(p, f^\pm) - s(p, f^\pm) \leq S(p,f) - s(p,f)$. Dann folgt $f^\pm \in \mathcal{R}[a,b]$ und damit auch $|f| = f^+ + f^- \in \mathcal{R}[a,b]$ aus dem Kriterium von Darboux.

(iii) Aus $f \leq g$ folgt $\displaystyle\int_a^b f(x)dx = \sup_{p\in\mathcal{P}[a,b]} s(p,f) \leq \sup_{p\in\mathcal{P}[a,b]} s(p,g) = \int_a^b g(x)dx.$

(iv) Für $f = 1$(konstant) gilt $S(p,1) = s(p,1) = b - a$ für alle $p \in \mathcal{P}[a,b]$.

(v) Sei $f \in \mathcal{R}[a,b]$ und $g \in C_{\mathbb{R}}(\mathbb{R})$. Dann ist g auf $[-\|f\|_\infty, \|f\|_\infty]$ gleichmäßig stetig. Also gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass aus $|x - x'| < \delta$ mit $x, x' \in [-\|f\|_\infty, \|f\|_\infty]$ folgt $|g(x) - g(x')| < \frac{\epsilon}{2(b-a)}$. Sei $\|g\|_\infty$ die Supremumsnorm der Funktion

$$g : [-\|f\|_\infty, \|f\|_\infty] \to \mathbb{R}, \quad x \mapsto g(x).$$

Weil $f \in \mathcal{R}[a,b]$ gibt es nach dem Darbouxkriterium ein $p \in \mathcal{P}[a,b]$, so dass gilt $S(p,f) - s(p,f) < \frac{\epsilon \cdot \delta}{4\|g\|_\infty}$. Wir zerlegen die Summe $S(p, g \circ f) - s(p, g \circ f)$ in die Summe über alle Teilintervalle, auf denen $\sup_{x \in [x_{i-1}, x_i]} f(x) - \inf_{x \in [x_{i-1}, x_i]} f(x) < \delta$ gilt und die Summe über alle Teilintervalle, auf denen $\sup_{x \in [x_{i-1}, x_i]} f(x) - \inf_{x \in [x_{i-1}, x_i]} f(x) \geq \delta$ gilt. Aufgrund der Wahl von δ folgt dann, dass die erste Summe kleiner ist als $\frac{\epsilon}{2(b-a)} \cdot (b-a) = \frac{\epsilon}{2}$ und die zweite Summe nicht größer als $(S(p,f) - s(p,f))\frac{2\|g\|_\infty}{\delta} < \frac{\epsilon}{2}$. Also gilt $S(p, g \circ f) - s(p, g \circ f) < \epsilon$ und $g \circ f$ erfüllt das Darbouxkriterium.
(vi) Jede Partition von $[a,b]$ besitzt eine Verfeinerung, die aus zwei Partitionen von $[a,c]$ und $[c,b]$ besteht. Dann folgt (v) aus dem Darbouxkriterium.
(vii) folgt aus (vi) und (iii).
(viii) Aus dem Beweis von (i) folgt für $f, f_n \in \mathcal{R}[a,b]$ und $p \in \mathcal{P}[a,b]$

$$|S(p,f) - s(p,f) - (S(p,f_n) - s(p,f_n))| \leq S(p, f - f_n) - s(p, f - f_n) \leq 2(b-a)\|f - f_n\|_\infty.$$

Für ein $\epsilon > 0$ wählen wir zuerst n so groß, dass $\|f - f_n\|_\infty < \frac{\epsilon}{4(b-a)}$ gilt, und dann p so dass $S(p,f_n) - s(p,f_n) < \frac{\epsilon}{2}$ gilt. Dann erfüllt f das Kriterium von Darboux.

Andererseits folgt für $f, f_n \in \mathcal{R}[a,b]$ aus der Monotonie $\left|\int_a^b f(x)dx - \int_a^b f_n(x)dx\right| \leq (b-a)\|f - f_n\|_\infty$. Also konvergiert die Folge von Integralen $(\int_a^b f_n(x)dx)_{n \in \mathbb{N}}$ gegen $\int_a^b f(x)dx$. **q.e.d.**

8.3 Differentiation und Integration

Hauptsatz der Differential- und Integralrechnung 8.14. *Sei $f \in \mathcal{R}[a,b]$ und F eine stetige Funktion auf $[a,b]$, die auf (a,b) differenzierbar ist mit $F' = f$. Dann gilt*

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

Beweis: Sei $p \in \mathcal{P}[a,b]$, dann gibt es wegen dem Mittelwertsatz eine Wahl von Zwischenpunkten ξ, so dass gilt $R(p,f,\xi) = F(b) - F(a)$. Wegen $s(p,f) \leq R(p,f,\xi) \leq S(p,f)$ folgt dann $\int_a^b f(x)dx = F(b) - F(a)$. **q.e.d.**

Beispiel 8.15. **(i)** *Sei $F \in C^1[a,b]$. Dann ist F' riemannintegrabel und es gilt für alle $x \in [a,b]$*

$$\int_a^x F'(t)dt = F(x) - F(a).$$

(ii) *Sei $1 < \alpha < 2$ und $F(x) = \begin{cases} |x|^\alpha \sin \frac{1}{x} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0. \end{cases}$ Dann ist F für $x \in \mathbb{R} \setminus \{0\}$ differenzierbar und*

$$F'(x) = \alpha \frac{|x|^\alpha}{x} \sin\left(\frac{1}{x}\right) + \frac{|x|^\alpha}{x^2} \cos\left(\frac{1}{x}\right).$$

Wegen $\frac{F(x)-F(0)}{|x|} = |x|^{\alpha-1} \sin\left(\frac{1}{x}\right)$ ist f auch bei $x = 0$ differenzierbar und dort gilt $F'(0) = 0$. Also gibt es differenzierbare Funktionen, deren Ableitungen auf einer kompakten Teilmenge nicht beschränkt sind, also dort auch nicht riemannintegrabel sind.

(iii) *Sei $f(x) = \begin{cases} 1 & \text{für } x \geq 0 \\ -1 & \text{für } x < 0. \end{cases}$ Dann ist f auf allen kompakten Intervallen riemannintegrabel. Offenbar gilt $F(x) = \int_0^x f(t)dt = |x|$. Also sind nicht alle Integrale von riemannintegrablen Funktionen differenzierbar.*

Satz 8.16. *Sei $f \in \mathcal{R}[a,b]$ im Punkt $x_0 \in (a,b)$ stetig. Dann ist $F(x) = \int_a^x f(t)dt$ im Punkt x_0 differenzierbar und es gilt $F'(x_0) = f(x_0)$.*

Beweis:

$$\left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| \leq \frac{1}{|x - x_0|} \int_{x_0}^x |f(t) - f(x_0)| dt$$

Weil f im Punkt stetig ist folgt aus den Eigenschaften des Integrals (iii), dass es für jedes $\epsilon > 0$ ein $\delta > 0$ gibt, so dass aus $|x - x_0| < \delta$ folgt

$$\left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| < \epsilon.$$ **q.e.d.**

Also sind die Integrale $F(x) = \int_a^x f(t)dt$ aller stetigen Funktionen $f \in C_{\mathbb{R}}([a,b])$ differenzierbar und es gilt $F'(x) = f(x)$ für alle $x \in (a,b)$.

Satz 8.17.* *Sei $f \in \mathcal{R}[a,b]$. Dann ist $F(x) = \int_a^x f(t)dt$ lipschitzstetig mit Lipschitzkonstante $L = \|f\|_\infty$.*

Beweis*: $$|F(x)-F(y)| = \left|\int_x^y f(t)dt\right| \le |y-x|\|f\|_\infty.$$ **q.e.d.**

Definition 8.18. *(Stammfunktion) Eine differenzierbare Funktion F mit $F' = f$ heißt Stammfunktion von f. Die Differenz zweier Stammfunktionen von f ist eine konstante Funktion. Wir bezeichnen eine Stammfunktion von f als $\int f(x)dx$.*

Beispiel 8.19. (i) $\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C$ *für* $\alpha \neq -1$ *und entweder* $\alpha \in \mathbb{N}$ *oder* $x \in \mathbb{R}^+$.

(ii) $\int \frac{1}{x} dx = \ln|x| + C$ *für* $x \neq 0$.

(iii) $\int e^x dx = e^x + C$.

(iv) $\int a^x dx = \frac{a^x}{\ln(a)} + C$ *für* $a \in \mathbb{R}^+ \setminus \{1\}$.

(v) $\int \cos(x)dx = \sin(x) + C$.

(vi) $\int \sin(x)dx = -\cos(x) + C$.

(vii) $\int \tan(x)dx = -\ln|\cos(x)| + C$ *für* $x \notin \{(n+\frac{1}{2})\pi \mid n \in \mathbb{Z}\}$.

(viii) $\int \cot(x)dx = \ln|\sin(x)| + C$ *für* $x \notin \{n\pi \mid n \in \mathbb{Z}\}$.

(ix) $\int \frac{1}{1+x^2}dx = \arctan(x) + C$.

(x) $\int \frac{1}{\sqrt{1-x^2}}dx = \arcsin(x) + C$ *für* $x \in [-1,1]$.

8.4 Technik des Integrierens

Substitutionsregel 8.20. *Sei $f \in C_{\mathbb{R}}([a,b])$ und $\phi : [\alpha,\beta] \to [a,b]$ eine differenzierbare Funktion, so dass $\phi' \in \mathcal{R}[\alpha,\beta]$. Dann gilt*

$$\int_{\phi(\alpha)}^{\phi(\beta)} f(x)dx = \int_\alpha^\beta f(\phi(t))\phi'(t)dt.$$

Für die Stammfunktionen gilt $$\int f(\phi(t))\phi'(t)dt = \left(\int f(x)dx\right) \circ \phi + C.$$

Beweis: Sei F eine Stammfunktion von f. Dann ist F wegen Satz 8.16 stetig differenzierbar und es gilt $F' = f$. Also ist $(F \circ \phi)' = (F' \circ \phi) \cdot \phi'$. Wegen den Eigenschaften des Riemannintegrals (i) und (iv) ist $(F' \circ \phi) \cdot \phi' = (f \circ \phi) \cdot \phi' \in \mathcal{R}[a,b]$. Dann folgt die Aussage aus dem Hauptsatz der Differential- und Integralrechnung. **q.e.d.**

Die Voraussetzung, dass das Bild von ϕ gleich $[a,b]$ sein muss kann abgeschwächt werden zu der Voraussetzung, dass f auf dem Bild von ϕ definiert und stetig sein muss.

Korollar 8.21. *(Transformation der Variabeln) Sei $\phi : [\alpha, \beta] \to [a, b]$ eine differenzierbare bijektive Funktion mit $\phi' \in \mathcal{R}[\alpha, \beta]$ und $f \in C_{\mathbb{R}}([a, b])$. Dann gilt*

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

Also gilt
$$\int f(x)dx = \left(\int f(\phi(t))\phi'(t)dt\right) \circ \phi^{-1} + C.$$
q.e.d.

Beispiel 8.22. **(i)**

$$\int_a^b f(\alpha t + \beta)dt = \frac{1}{\alpha} \int_{\alpha a+\beta}^{\alpha b+\beta} f(x)dx$$

(ii)

$$\int R\left(x, \sqrt[n]{ax + b}\right) dx = \int R\left(\frac{t^n - b}{a}, t\right) \frac{n}{a} t^{n-1}dt + C$$

für $n \in \mathbb{N}, a, b \in \mathbb{R}, a \neq 0$ und einer rationale Funktion $R(\cdot, \cdot)$ in zwei Variablen. Wir substituieren $t = \sqrt[n]{ax + b} \Longrightarrow ax + b = t^n \Longrightarrow x = \frac{t^n-b}{a}$ und $dx = \frac{nt^{n-1}}{a}dt$.

(iii)

$$\int R\left(x, \sqrt{x^2 + 1}\right) dx = \int R(\sinh t, \cosh t) \cosh t dt + C$$

mit der Substitution $x = \sinh t, \sqrt{x^2 + 1} = \cosh t$ und $dx = \cosh t dt$.

(iv)

$$\int R\left(x, \sqrt{x^2 - 1}\right) dx = \pm \int R(\pm \cosh t, \sinh t) \sinh t dt + C$$

mit der Substitution $x = \pm \cosh t$, je nachdem ob $x \in \mathbb{R}^{\pm}$. Dann gilt $\sqrt{x^2 - 1} = \sinh t$ und $dx = \pm \sinh t dt$.

(v)

$$\int R\left(x, \sqrt{1 - x^2}\right) dx = \mp \int R(\pm \cos t, \sin t) \sin t dt + C.$$

mit der Substitution $x = \pm \cos t, \sqrt{1 - x^2} = \sin t$ und $dx = \mp \sin t dt$.

(vi)

$$\int R(\cos x, \sin x)dx = \int R\left(\frac{1 - t^2}{1 + t^2}, \frac{2t}{1 + t^2}\right) \frac{2dt}{1 + t^2} + C$$

mit der Substition $t = \tan\left(\frac{x}{2}\right), x = 2\arctan(t)$ *und* $dx = \frac{2dt}{1+t^2}$*, so dass gilt*

$$\frac{1-t^2}{1+t^2} = \frac{\cos^2\left(\frac{x}{2}\right) - \sin^2\left(\frac{x}{2}\right)}{\cos^2\left(\frac{x}{2}\right) + \sin^2\left(\frac{x}{2}\right)} = \cos(x) \text{ und } \frac{2t}{1+t^2} = \frac{2\cos\left(\frac{x}{2}\right)\sin\left(\frac{x}{2}\right)}{\cos^2\left(\frac{x}{2}\right) + \sin^2\left(\frac{x}{2}\right)} = \sin(x).$$

(vii)

$$\int R(\cosh x, \sinh x)dx = \int R\left(\frac{t^2+1}{2t}, \frac{t^2-1}{2t}\right)\frac{dt}{t} + C$$

mit der Substition $t = e^x, x = \ln(t)$ *und* $dx = \frac{dt}{t}$.

Partielle Integration 8.23. *Seien* $f, g \in C_{\mathbb{R}}[a,b]$ *differenzierbar mit* $f', g' \in \mathcal{R}[a,b]$. *Dann gilt*

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

Also gilt

$$\int f(x)g'(x)dx = fg - \int f'(x)g(x)dx + C.$$

Beweis folgt aus dem Hauptsatz der Differentialrechnung und der Leibnizregel. **q.e.d.**

Beispiel 8.24. (i) $\int \ln(x)dx = x\ln(x) - \int x\frac{1}{x}dx = x(\ln(x) - 1) + C$.

(ii) $\int \sqrt{1-x^2}dx = x\sqrt{1-x^2} + \int \frac{x^2}{\sqrt{1-x^2}}dx + C$
$= x\sqrt{1-x^2} - \int \frac{1-x^2}{\sqrt{1-x^2}}dx + \int \frac{1}{\sqrt{1-x^2}}dx + C$
$\Longrightarrow \int \sqrt{1-x^2}dx = \frac{x\sqrt{1-x^2}}{2} + \frac{\arcsin(x)}{2} + C$ *also* $\int_{-1}^{1} \sqrt{1-x^2}dx = \frac{\pi}{2}$.

(iii) $\int x^n e^x dx = x^n e^x - n\int x^{n-1}e^x dx$

(iv) $\int x^n \cos x dx = x^n \sin x - n\int x^{n-1}\sin x dx$.

(v) $\int x^n \sin x dx = -x^n \cos x + n\int x^{n-1}\cos x dx$.

Partialbruchzerlegung 8.25. *(Integration von rationalen Funktionen)*
1. Faktorisierung des Nenners. *In* $\mathbb{C}$ *lässt sich wegen dem Fundamentalsatz der Algebra das Nennerpolynom* Q *einer rationalen Funktion in ein Produkt von Polynome ersten Grades zerlegen. Wenn das Polynom reelle Koeffizienten hat, dann sind die Nullstellen entweder reell oder sie treten in Paaren von komplex konjugierten Nullstellen auf. Indem wir die Paare zu Polynomen zweiten Grades zusammenfassen zerlegen wir Polynome mit reellen Koeffizienten in ein Produkt von Polynomen ersten und zweiten Grades mit reellen Koeffizienten.*
2. Polynomdivision. *Zerlege eine komplexe, rationale Funktion in eine Summe eines Polynoms und Summanden von der Form* $\frac{c}{(x-x_0)^l}$.

Lemma 8.26. *(Abspaltung des Hauptteiles) Sei $R(x) = \frac{P(x)}{Q(x)}$ eine rationale Funktion mit komplexen Koeffizienten, dessen Zählerpolynom P und Nennerpolynom Q keine gemeinsamen Nullstellen haben. Wenn das Nennerpolynom $Q(x)$ an der Stelle x_0 eine Nullstelle der Ordnung n hat, d.h. $Q(x) = (x - x_0)^n q(x)$ mit $q(x_0) \neq 0$, dann gibt es komplexe Koeffizienten $c_1, \ldots, c_n$ und ein Polynom $p(x)$, so dass gilt*

$$\frac{P(x)}{Q(x)} = \frac{c_1}{x - x_0} + \ldots + \frac{c_n}{(x - x_0)^n} + \frac{p(x)}{q(x)}.$$

Beweis: *Weil jedes komplexe Polynom für jede komplexe Zahl x_0 auch ein Polynom in $x - x_0$ ist, können wir annehmen, dass $x_0 = 0$ ist. Seien*

$$c_l = \frac{1}{(n-l)!} \left(\frac{d}{dx}\right)^{n-l} \frac{P(x)}{q(x)}\Bigg|_{x=x_0} \qquad \text{für } l = 1, \ldots, n.$$

Weil $q(x_0) \neq 0$ sind diese Ableitungen wohl definiert. Dann verschwinden die 0–te bis zur $(n-1)$–ten Ableitungen der rationalen Funktion

$$x \mapsto \frac{P(x)}{q(x)} - c_n - c_{n-1}(x - x_0) - \ldots - c_1(x - x_0)^{n-1}.$$

Deshalb läßt sich diese rationale Funktion schreiben als

$$\frac{P(x)}{q(x)} - c_n - c_{n-1}(x - x_0) - \ldots - c_1(x - x_0)^{n-1} = (x - x_0)^n \frac{p(x)}{q(x)} \text{ mit einem Polynom } p.$$

Daraus folgt $$\frac{P(x)}{Q(x)} = \frac{p(x)}{q(x)} + \frac{c_1}{x - x_0} + \ldots + \frac{c_n}{(x - x_0)^n}.$$ **q.e.d.**

Im Fall von Paaren von komplex konjugierten Nullstellen des Nennerpolynoms einer reellen rationalen Funktion können wir das Verfahren so modifizieren, dass wir keine komplexen Zahlen benötigen. Seien also x_0 und $\bar{x}_0$ zwei komplex konjugierte Nullstellen von dem Polynom Q mit reellen Koeffizienten, die wieder jeweils n–fach auftreten, d.h. $Q(x) = (x - x_0)^n (x - \bar{x}_0)^n q(x)$ mit einem Polynom q mit reellen Koeffizienten, das keine Nullstellen bei x_0 und $\bar{x}_0$ hat. Dann hat

$$x \mapsto \frac{P(x)}{q(x)} - (a_n + b_n(x - x_0)) - \ldots - (a_1 + b_1(x - x_0))(x - x_0)^{n-1}(x - \bar{x}_0)^{n-1}$$

mit den Koeffizienten

$$a_l = \frac{\left(\frac{d}{dx}\right)^{n-l} \frac{P(x)}{q(x)}\Big|_{x=x_0}}{(n-l)!(x_0 - \bar{x}_0)^{n-l}} \qquad b_l = \frac{\left(\frac{d}{dx}\right)^{n-l} \frac{P(x)}{q(x)}\Big|_{x=\bar{x}_0}}{(n-l)!(\bar{x}_0 - x_0)^{n+1-l}} \qquad \text{für } l = 1, \ldots, n$$

eine n–fache Nullstelle sowohl bei x_0 als auch bei $\bar{x}_0$. Diese Funktion ist also das Produkt von $(x-x_0)^n(x-\bar{x}_0)^n$ mal einer rationalen Funktion $\frac{p(x)}{q(x)}$. Weil $\frac{P(x)}{Q(x)}$ bei komplex konjugierten Variablen komplex konjugierte Werte annimmt, gilt das auch für $\frac{p(x)}{q(x)}$, das dann wieder eine rationale, reelle Funktion ist. Das ergibt folgende Relation von rationalen reellen Funktionen:

$$\frac{P(x)}{Q(x)} = \frac{p(x)}{q(x)} + \frac{a_1 - b_1x_0 + b_1x}{(x-x_0)(x-\bar{x}_0)} + \ldots + \frac{a_n - b_nx_0 + b_nx}{(x-x_0)^n(x-\bar{x}_0)^n}.$$

Deshalb lässt sich jede rational Funktion mit reellen Koffizienten zerlegen in ein Summe eines Polynoms mit rellen Koeffizienten und Summanden von der Form $\frac{c}{(x-x_0)^l}$ und $\frac{a+bx}{(x^2+px+q)^l}$ mit $a,b,c,p,q,x_0 \in \mathbb{R}$ und $l \in \mathbb{N}$ und $4q-p^2>0$.

3. Termweise Integration. $$\int \frac{dx}{(x-x_0)^l} = \begin{cases} \ln|x-x_0| + C & \textit{für } l=1 \\ \frac{-1}{(l-1)(x-x_0)^{l-1}} + C & \textit{sonst.} \end{cases}$$

$$\int \frac{(a+bx)dx}{(x^2+px+q)^l} = \begin{cases} \frac{b}{2}\ln(x^2+px+q) + \left(a - \frac{bp}{2}\right)\int \frac{dx}{x^2+px+q} + C & \textit{für } l=1 \\ \frac{-b}{2(l-1)(x^2+px+q)^{l-1}} + \left(a-\frac{bp}{2}\right)\int \frac{dx}{(x^2+px+q)^l} + C & \textit{sonst.} \end{cases}$$

$$\int \frac{dx}{(x^2+px+q)^l} = \begin{cases} \frac{2}{\sqrt{4q-p^2}} \arctan\left(\frac{2x+p}{\sqrt{4q-p^2}}\right) + C & l=1 \\ \frac{2x+p}{(l-1)(4q-p^2)(x^2+px+q)^{l-1}} + \frac{2(2l-3)}{(l-1)(4q-p^2)}\int \frac{dx}{(x^2+px+q)^{l-1}} + C & l \neq 1. \end{cases}$$

Damit können wir alle rationalen Funktionen integrieren.

Satz 8.27. *(Restglied der Taylorformel in Integralform) Sei $f \in C^n_{\mathbb{R}}([a,b])$ und $f^{(n+1)} \in \mathcal{R}[a,b]$. Dann gilt*

$$f(x) = \sum_{m=0}^{n} \frac{f^{(m)}(x_0)}{m!}(x-x_0)^m + (x-x_0)^{n+1}\int_0^1 \frac{(1-t)^n}{n!} f^{n+1}(x_0+t(x-x_0)dt$$

Beweis: durch vollständige Induktion und partielle Integration.

(i) für $n=0$ folgt die Aussage aus dem Hauptsatz der Differential- und Integralrechnung:

$$f(x) - f(x_0) = (x-x_0)\int_0^1 f'(x_0+t(x-x_0))dt = \int_{x_0}^{x} f'(t)dt.$$

(ii) Die Aussage gelte für $n \in \mathbb{N}$, und f sei $(n+1)$ mal differenzierbar mit $f^{(n+1)} \in \mathcal{R}[a,b]$. Dann folgt mit einer partiellen Integration

$$f(x) = \sum_{m=0}^{n} \frac{f^{(m)}(x_0)}{m!}(x-x_0)^m + (x-x_0)^{n+1}\int_0^1 \frac{(1-t)^n}{n!} f^{n+1}(x_0+t(x-x_0))dt$$

$$= \sum_{m=0}^{n} \frac{f^{(m)}(x_0)}{m!}(x-x_0)^m - \frac{(x-x_0)^{n+1}(1-t)^{n+1}}{(n+1)!} f^{n+1}(x_0+t(x-x_0))\Big|_{t=0}^{t=1}$$

$$+(x-x_0)^{n+2} \int_0^1 \frac{(1-t)^{n+1}}{(n+1)!} f^{n+2}(x_0+t(x-x_0))dt$$

$$= \sum_{m=0}^{n+1} \frac{f^{(m)}(x_0)}{m!}(x-x_0)^m + (x-x_0)^{n+2} \int_0^1 \frac{(1-t)^{n+1}}{n+1} f^{n+2}(x_0+t(x-x_0))dt$$

Also gilt die Aussage für $n+1$. **q.e.d.**

Satz 8.28.* *(Mittelwertsatz der Integralrechnung) Seien $f \in \mathcal{R}[a,b]$. Dann ist*

$$\frac{1}{b-a} \int_a^b f(x)dx \in [\inf_{x\in[a,b]} f(x), \sup_{x\in[a,b]} f(x)]$$

Wenn $f \in C_{\mathbb{R}}([a,b])$, dann gibt es ein $x_0 \in (a,b)$ mit $\frac{1}{b-a}\int_a^b f(x)dx = f(x_0)$.

Beweis:* Wegen $\inf_{x\in[a,b]} f(x) \leq f \leq \sup_{x\in[a,b]} f(x)$ folgt die erste Aussage aus der Monotonie. Wenn f stetig ist folgt aus dem Mittelwertsatz für $F(x) = \int_a^x f(t)dt$, dass es ein $x_0 \in (a,b)$ gibt mit $f(x_0) = F'(x_0) = \frac{F(b)-F(a)}{b-a} = \frac{1}{b-a}\int_a^b f(x)dx$. **q.e.d.**

8.5 Uneigentliches Integral

In diesem Abschnitt erweitern wir den Begriff des Riemannintegrals auf offene und unbeschränkte Intervalle.

Definition 8.29. *Eine Funktion f heißt riemannintegrabel auf dem offenen (nicht notwendigerweise beschränkten) Intervall $(a,b) \subset \mathbb{R}$, wenn f riemannintegrabel ist auf allen kompakten Teilintervallen, und wenn für ein $c \in (a,b)$ beide Grenzwerte $\lim_{x\to a+}\int_x^c f(t)dt$ und $\lim_{x\to b-}\int_c^b f(t)dt$ existieren.*

Beispiel 8.30. (i) $\displaystyle\int_1^\infty \frac{1}{x^\alpha}dx$. $\displaystyle\int \frac{1}{x^\alpha}dx = \begin{cases} \frac{-1}{(\alpha-1)x^{\alpha-1}} + C & \text{für } \alpha \neq 1 \\ \ln|x| + C & \text{für } \alpha = 1 \end{cases}$. *Also existiert der Grenzwert $\lim_{x\to\infty-}\int_1^x \frac{1}{t^\alpha}dt$ nur für $\alpha > 1$. Dann gilt $\int_1^\infty \frac{1}{x^\alpha}dx = \frac{1}{\alpha-1}$.*

(ii) $\displaystyle\int_0^1 \frac{1}{x^\alpha}dx$. $\displaystyle\int \frac{1}{x^\alpha}dx = \begin{cases} \frac{-1}{(\alpha-1)x^{\alpha-1}} + C & \text{für } \alpha \neq 1 \\ \ln|x| + C & \text{für } \alpha = 1 \end{cases}$. *Dann existiert der Grenzwert $\lim_{x\to 0+}\int_x^1 \frac{1}{x^\alpha}dx$ genau dann, wenn $\alpha < 1$. In diesem Fall gilt $\int_0^1 \frac{1}{x^\alpha}dx = \frac{1}{1-\alpha}$. Wegen (i) folgt dann, dass $\int_0^\infty \frac{1}{x^\alpha}dx$ für kein α existiert.*

(iii) $\int\limits_{-\infty}^{\infty} \frac{1}{1+x^2}dx$. $\int \frac{1}{1+x^2}dx = \arctan(x) + C$. *Also gilt* $\lim\limits_{x\to-\infty+} \int_x^0 \frac{1}{1+t^2}dt = \frac{\pi}{2} = \lim_{x\to\infty-} \int_0^x \frac{1}{1+t^2}dt$. *Dann folgt* $\int_{-\infty}^{\infty} \frac{1}{1+x^2}dx = \pi$.

Verschiedenen Kriterien helfen zu entscheiden, ob diese Grenzwerte existieren. Hier einige Kriterien für uneigentliche Integrale $\lim_{x\to b-} \int_a^x f(t)dt$.

Cauchykriterium: $\lim_{x\to b-} \int_a^x f(t)dt$ existiert genau dann, wenn es zu jedem $\epsilon > 0$ ein $c \in (a,b)$ gibt, so dass für alle $a < c < d < e < b$ gilt $\left|\int_d^e f(x)dx\right| < \epsilon$.

Monotoniekriterium: Wenn $f \geq 0$, dann existiert $\lim_{x\to b-} \int_a^x f(t)dt$ genau dann, wenn $F(x) = \int_a^x f(t)dt$ auf $x \in (a,b)$ beschränkt ist.

Majorantenkriterium: Wenn $f \geq 0$ und $f \leq g$, dann existiert $\lim_{x\to b-} \int_a^x f(t)dt$, wenn $\lim_{x\to b-} \int_a^x g(t)dt$ existiert.

Definition 8.31. *Eine Funktion f auf einem offenen (unbeschränkten) Intervall heißt absolut riemannintegrabel, wenn $|f|$ riemannintegrabel ist.*

Wegen der Dreiecksungleichung gilt dann $\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx$. Alle absolut riemannintegrablen Funktionen sind also auch riemannintegrabel.

Satz 8.32. *(Integralkriterium für Reihen) Sei $f : [1,\infty) \to \mathbb{R}^+$ monoton fallend mit dem Grenzwert $f(x) \to 0$ für $x \to \infty$. Dann ist die Folge*

$$\left(\sum_{k=1}^{n} f(k) - \int\limits_1^n f(x)dx\right)_{n\in\mathbb{N}}$$

eine monoton fallende konvergente Folge positiver Zahlen. Für alle $m < n \in \mathbb{N}$ gilt nämlich

$$f(n) - f(m) \leq \sum_{k=m+1}^{n} f(k) - \int\limits_m^n f(x)dx \leq 0.$$

Die Reihe $(\sum f(n))_{n\in\mathbb{N}}$ konvergiert genau dann, wenn $\int_1^\infty f(x)dx < \infty$. Dann gilt:

$$\int\limits_1^\infty f(x)dx \leq \sum_{n=1}^{\infty} f(n) \leq f(1) + \int\limits_1^\infty f(x)dx.$$

Beweis: Für $m < n \in \mathbb{N}$ sei $p_{m,n} \in \mathcal{P}[m,n]$ die Partition $\{m, m+1, \ldots, n\}$. Dann ist offenbar $s(p_{m,n}, f) = \sum_{k=m+1}^{n} f(k)$ und $S(p_{m,n}, f) = \sum_{k=m}^{n-1} f(k)$. Also gilt

$$\sum_{k=m+1}^{n} f(k) \leq \int\limits_m^n f(x)dx \leq \sum_{k=m}^{n-1} f(k).$$

Daraus folgt sofort

$$f(n) - f(m) = \sum_{k=m+1}^{n} f(k) - \sum_{k=m}^{n-1} f(k) \leq \sum_{k=m+1}^{n} f(k) - \int\limits_{m}^{n} f(x)dx \leq 0.$$

Also ist die Folge $(\sum_{k=1}^{n} f(k) - \int_1^n f(x)dx)_{n\in\mathbb{N}}$ eine monoton fallende beschränkte Folge, die wegen dem Monotonieprinzip konvergiert. Mit $m = 1$ folgt

$$\int\limits_{1}^{n} f(x)dx \leq \int\limits_{1}^{n+1} f(x)dx \leq \sum_{k=1}^{n} f(k) \leq f(1) + \int\limits_{1}^{n} f(x)dx.$$

Dann folgt die Aussage aus dem Majorantenkriterium. **q.e.d.**

Beispiel 8.33. (i) *$(\sum \frac{1}{n^s})_{n\in\mathbb{N}}$ ist genau dann konvergent, wenn $\int_1^\infty \frac{1}{x^s}dx$ existiert. Also für $s > 1$. Dann gilt*

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

Der Grenzwert heißt Riemannsche ζ-Funktion.

(ii) *Die Folge $\sum_{k=1}^{n} \frac{1}{k} - \ln(n)$ ist eine monoton fallende konvergente Folge positiver Zahlen. Der Grenzwert $\gamma = \lim_{n\to\infty}(\sum_{k=1}^{n} \frac{1}{k}) - \ln(n)$ wird Eulersche Konstante genannt. Bis heute ist nicht bekannt, ob er rational oder irrational ist.*

(iii) *Wegen (i) ist die Funktion $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ für $s \in (1,\infty)$ konvergent. Die Folge von Funktionen*

$$f_n(s) = \sum_{k=1}^{n} \frac{1}{k^s} - \int\limits_{1}^{n} \frac{dx}{x^s}$$

ist wegen dem vorangehenden Satz auf $s \in (0,\infty)$ eine monoton fallende Folge von Funktionen. Aus dem Beispiel 7.17 folgt für $0 < s < t$ und $x \in [1,\infty)$

$$0 < \frac{1}{x^s} - \frac{1}{x^t} = \frac{1}{x^s}(1 - \exp((s-t)\ln(x))) \leq \frac{1}{x^s}(t-s)\ln(x).$$

Deshalb ist das eine Folge von stetigen Funktionen auf $s \in \mathbb{R}^+$. Wegen dem vorgangehenden Satz und Satz 5.27 konvergiert sie für alle $\epsilon > 0$ auf $[\epsilon,\infty)$ gleichmäßig gegen eine stetige Funktion auf $[\epsilon,\infty)$. Auf $s \in (1,\infty)$ ist wegen (i) der Grenzwert gleich

$$\zeta(s) - \int\limits_{1}^{\infty} \frac{dx}{x^s} = \zeta(s) - \frac{1}{s-1}.$$

Und für $s = 1$ ist der Grenzwert gleich

$$\lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \int_{1}^{n}\frac{dx}{x}\right) = \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \ln(n)\right) = \gamma.$$

Also folgt
$$\lim_{s\to 1+}\left(\zeta(s) - \frac{1}{s-1}\right) = \gamma.$$

(iv) $\left(\sum \frac{1}{(n+1)\ln^s(n+1)}\right)_{n\in\mathbb{N}}$ *ist genau dann konvergent, wenn* $\int_2^\infty \frac{1}{x\ln^s(x)}dx = \int_{\ln(2)}^\infty \frac{1}{x^s}dx$ *existiert, also für* $s > 1$.

(v) *Nach Euler ist die* Γ*-Funktion definiert durch*

$$\Gamma(x) = \int_{0}^{\infty} e^{-t}t^{x-1}dt.$$

Dieses Integral ist an beiden Grenznen uneigentlich. Deshalb zerlegen wir es in eine Summe von zwei Integralen $\int_0^\infty = \int_0^1 + \int_1^\infty$. *Auf* $t \in (0,1]$ *ist der Integrand beschränkt durch* $e^{-1}t^{x-1} \leq e^{-t}t^{x-1} \leq t^{x-1}$. *Deshalb konvergiert das erste Integral für* $x - 1 > -1 \iff x > 0$. *Wegen* $e^{-t}t^{x-1} = \exp(-t + (x-1)\ln(t))$ *und weil für alle* $\epsilon > 0$ *im Grenzwert* $\lim_{t\to\infty}$ *der Ausdruck*$-\epsilon t + x - 1\ln(t)$ *negativ ist, konvergiert das zweite Intgeral für alle* $x \in \mathbb{R}$. *Also ist* $\Gamma(x)$ *für alle* $x \in \mathbb{R}^+$ *definiert. Durch eine partielle Integration erhalten wir*

$$\int_{\epsilon}^{R} e^{-t}t^{x}dt = -e^{-t}t^{x}\Big|_{t=\epsilon}^{t=R} + x\int_{\epsilon}^{R} e^{-t}t^{x-1}dt.$$

Im Grenzwert $\epsilon \to 0$ *und* $R \to \infty$ *erhalten wir folgende Funktionalgleichung:*

$$\Gamma(x+1) = x\Gamma(x).$$

Mit $\Gamma(1) = \int_0^\infty e^{-t}dt = 1$ *folgt induktiv* $\Gamma(n) = (n-1)!$.

Kapitel 9

Metrische Räume und Banachräume

9.1 Metrik und Norm

Definition 9.1. *(Metrik auf einer Menge X) Eine Metrik (oder Abstandsfunktion) ist eine Abbildung* $d : X \times X \to \mathbb{R}, \quad (x,y) \mapsto d(x,y)$ *mit drei Eigenschaften*

(i) $d(x,y) \geq 0$ *für alle* $x,y \in X$ *und* $d(x,y) = 0 \Leftrightarrow x = y$ *(Positivität).*

(ii) $d(x,y) = d(y,x)$ *(Symmetrie).*

(iii) $d(x,y) \leq d(x,z) + d(z,y)$ *für alle* $x,y,z \in X$ *(Dreiecksungleichung).*

Beispiel 9.2. (i) *auf jeder Menge X definiert* $d(x,y) = \begin{cases} 0 & \text{für } x = y \\ 1 & \text{für } x \neq y \end{cases}$

die sogenannte diskrete Metrik.

(ii) *Auf* $\mathbb{R}$ *definiert* $d(x,y) = |x-y|$ *eine Metrik.*

(iii) *Auf* $\mathbb{C}$ *definiert* $d(x,y) = |x-y|$ *eine Metrik.*

(iv) *Auf jeder nicht leeren Teilmenge* $A \subset X$ *eines metrischen Raumes* (X,d) *definiert die Einschränkung von* d *auf* $A \times A \to \mathbb{R}$ *eine Metrik.*

(v) *Auf dem kartesischen Produkt zweier metrischer Räume definiert die Summe beider Metriken eine Metrik. Sie heißt Metrik des kartesischen Produktes.*

(vi) *Die Einschränkung der Metrik (ii) auf die Vereinigung der inversen der natürlichen Zahlen mit* $\{0\}$ *definiert eine Metrik auf* $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\} \simeq \{\frac{1}{n} \mid n \in \mathbb{N}\} \cup \{0\}$*:*

$$d(n,m) = \frac{|n-m|}{nm} \quad d(\infty,n) = d(n,\infty) = \frac{1}{n} \quad d(\infty,\infty) = 0 \quad \text{für alle } n,m \in \mathbb{N}.$$

Die Menge $V = \mathbb{R}^n$ bzw. $\mathbb{C}^n$ aller reellen bzw. komplexen Zahlen erfüllt mit $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$ und $0 = (0, \ldots, 0)$ die Axiome $A1$. Außerdem besitzt sie eine Skalarmultiplikation

$$\mathbb{R} \times V \to V \text{ bzw. } \mathbb{C} \times V \to V (\lambda, (x_1, \ldots, x_n)) \to \lambda \cdot (x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n).$$

Mit $(\lambda \cdot \mu) \cdot (x_1, \ldots, x_n) = \lambda(\mu \cdot (x_1, \ldots, x_n))$. Eine Menge mit Addition und Skalarmultiplikation, die das erfüllt, heißt (reeller bzw. komplexer) Vektorraum.

Definition 9.3. *Ein Vektorraum V ist eine Menge V zusammen mit den Abbildungen:*

$$+ : V \times V \to V, \qquad (v, w) \mapsto v + w \qquad \cdot : \mathbb{K} \times V \to V, \qquad (\lambda, v) \mapsto \lambda v,$$

die die Axiome $A1$ der Addition und das Distributivgesetz $A3$ erfüllen und

$$\lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v \quad \textit{für alle } \lambda, \mu \in \mathbb{K}, v \in V.$$

Definition 9.4. *Eine Norm auf einem reellen bzw. komplexen Vektorraum V ist eine Abbildung $\|\cdot\| : V \to \mathbb{R}, \quad x \mapsto \|x\|$ mit folgenden 3 Eigenschaften:*

(i) $\|x\| \geq 0$ *und* $\|x\| = 0 \Leftrightarrow x = 0$ *für alle* $x \in V$

(ii) $\|\lambda \cdot x\| = |\lambda| \cdot \|x\|$ *für* $\lambda \in \mathbb{R}$ *bzw.* $\mathbb{C}$ *und* $x \in V$

(iii) $\|x + y\| \leq \|x\| + \|y\|$ *für alle* $x, y \in V$

Satz 9.5. *Jede Norm induziert eine Metrik $d : V \times V \to \mathbb{R}, (x, y) \mapsto d(x, y) = \|x - y\|$.*

Beweis:(i) folgt aus (i) der Definition einer Norm.
(ii) $d(y, x) = \|y - x\| = \|(-1)(x - y)\| = |-1|\|x - y\| = \|x - y\| = d(x, y)$
(iii) $d(x, y) = \|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$.**q.e.d.**

Im Folgenden werden wir mit der Vorgabe einer Norm auf einem Vektorraum immer auch die entsprechende induzierte Metrik auf diesem Vektorraum betrachten. Deshalb fassen wir jeden normierten Vektorraum auch als metrischen Raum auf.

Definition 9.6. *(Euklidische Norm und Metrik) Auf $\mathbb{R}^n$ definiert*

$$\|(x_1, \ldots, x_n)\| = \sqrt{|x_1|^2 + \ldots + |x_n|^2} = \sqrt{x_1^2 + \ldots + x_n^2}$$

die euklidische Norm. Die entsprechende Metrik heißt dann euklidische Metrik.

Die Eigenschaft (iii) heißt dabei Minkowski–Ungleichung:

$$\sqrt{(x_1 + y_1)^2 + \ldots + (x_n + y_n)^2} \leq \sqrt{x_1^2 + \ldots + x_n^2} + \sqrt{y_1^2 + \ldots + y_n^2}.$$

Um diese zu beweisen zeigen wir zuerst die

Cauchy–Schwarz'sche Ungleichung 9.7.

$$|x_1y_1| + \ldots + |x_ny_n| \le \sqrt{x_1^2 + \ldots + x_n^2}\sqrt{y_1^2 + \ldots + y_n^2}$$

Beweis: Wegen $|x_iy_i| = |x_i| \cdot |y_i|, x_i^2 = |x_i|^2$ und $y_i^2 = |y_i|^2$ genügt es offenbar

$$x_1y_1 + \ldots + x_ny_n \le \sqrt{x_1^2 + \ldots + x_n^2}\sqrt{y_1^2 + \ldots + y_n^2}$$

zu zeigen. Das ist aber äquivalent zu

$$(x_1y_1 + \ldots + x_ny_n)^2 \le (x_1^2 + \ldots + x_n^2)(y_1^2 + \ldots + y_n^2).$$

Für $y = (y_1, \ldots, y_n) = 0$ ist die Aussage trivial. Sei also $y \neq 0$.

$$\begin{aligned}
\left\|\|y\|x - \frac{x_1y_1 + \ldots + x_ny_n}{\|y\|}y\right\|^2 &= \sum_{\imath=1}^n \left(\|y\|x_\imath - \frac{x_1y_1 + \ldots + x_ny_n}{\|y\|}y_\imath\right)^2 \\
= \sum_{\imath=1}^n \left(\|y\|^2x_\imath^2 + \frac{(x_1y_1 + \ldots + x_ny_n)^2}{\|y\|^2}y_\imath^2 - 2(x_1y_1 + \ldots + x_ny_n)x_\imath y_\imath\right) \\
= \|y\|^2\|x\|^2 + \frac{(x_1y_1 + \ldots + x_ny_n)^2}{\|y\|^2}\|y\|^2 - 2(x_1y_1 + \ldots + x_ny_n)^2 \\
= \|y\|^2\|x\|^2 - (x_1y_1 + \ldots + x_ny_n)^2.
\end{aligned}$$

Weil diese Ausdrücke nicht negativ sind folgt $(x_1y_1 + \ldots + x_ny_n) \le \|x\|^2\|y\|^2$. **q.e.d.**

Beweis der Minkowski Ungleichung:

$$\begin{aligned}
(x_1 + y_1)^2 + \ldots + (x_n + y_n)^2 &\le \|x\|^2 + \|y\|^2 + 2(x_1y_1 + \ldots + x_ny_n) \\
&\le \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| \le (\|x\| + \|y\|)^2.
\end{aligned}$$

q.e.d.

Analog definiert $\|(x_1, \ldots, x_n)\| = \sqrt{x_1\bar{x}_1 + \ldots + x_n\bar{x}_n}$ eine Norm auf $\mathbb{C}^n$. Identifizieren wir die komplexen Zahlen $\mathbb{C}$ mit $\mathbb{R}^2$ (Realteil und Imaginärteil), dann ist $\mathbb{C}^n \simeq (\mathbb{R}^2)^n \simeq \mathbb{R}^{2n}$.

Übungsaufgabe 9.8. *Die euklidische Norm auf $\mathbb{R}^{2n}$ induziert durch diese Identifikation auf $\mathbb{C}^n$ die Norm $\|(x_1, \ldots, x_n)\| = \sqrt{x_1\bar{x}_1 + \ldots + x_n\bar{x}_n}$.*

Definition 9.9. *(Norm und Skalarprodukt) Für $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ und $1 \le p < \infty$ sei*

$$\|x\|_p = (|x_1|^p + \ldots + |x_n|^p)^{1/p}$$

Wir hatten in einer Übungsaufgabe gesehen, dass $\|x\|_p$ im Grenzwert $p \to \infty$ gegen

$$\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$$

konvergiert. Das Skalarprodukt wird definiert als

$$\langle x, y\rangle = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n \in \mathbb{K}, \quad \text{für } x, y \in \mathbb{K}^n.$$

Für $y \in \mathbb{R}^n$ setzen wir hierbei $(\bar{y}_1, \ldots, \bar{y}_n) = \bar{y} = y$.

Satz 9.10. *(Höldersche Ungleichung) Seien $p \geq 1$ und $q \geq 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$.*

Dann gilt $$|\langle x, y\rangle| \leq |x_1y_1| + \ldots + |x_ny_n| \leq \|x\|_p \cdot \|y\|_q \quad \text{für alle } x, y \in \mathbb{K}^n.$$

Für $p = q = 2$ erhalten wir wieder die Cauchy-Schwarzsche Ungleichung

$$|x_1y_1| + \ldots + |x_ny_n| \leq \|x\| \cdot \|y\|.$$

Beweis: Wenn $p = 1$ und $q = \infty$ gilt

$$|x_1y_1| + \ldots + |x_ny_n| \leq (|x_1| + \ldots + |x_n|)\max\{|y_1|, \ldots, |y_n|\}.$$

Den Fall $p = \infty$ und $q = 1$ erhalten wir durch vertauschen von x und y. Wir können also $1 < p, q < \infty$ annehmen, und dass $\|x\|_p \neq 0 \neq \|y\|_q$ gilt, weil die Ungleichung sonst offensichtlich ist. Dann folgt aus der Young'schen Ungleichung für alle $k = 1, \ldots, n$

$$\frac{|x_ky_k|}{\|x\|_p\|y\|_q} = \frac{|x_k|}{\|x\|_p}\frac{|y_k|}{\|y\|_q} \leq \frac{1}{p}\frac{|x_k|^p}{\|x\|_p^p} + \frac{1}{q}\frac{|y_k|^q}{\|y\|_q^q}$$

Nach Summation über $k = 1, \ldots, n$ erhalten wir

$$\frac{|x_1y_1| + \ldots + |x_ny_n|}{\|x\|_p\|y\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1.$$ **q.e.d.**

Satz 9.11. *(Minkowski Ungleichung) Sei $p \geq 1$ und $x, y \in \mathbb{K}^n$, dann gilt*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

Korollar 9.12. *Für alle $1 \leq p \leq \infty$ und $n \in \mathbb{N}$ ist $\|\cdot\|_p : \mathbb{K}^n \to \mathbb{R}$ eine Norm.* **q.e.d.**

Beweis der Minkowski Ungleichung: Für $p = 1$ oder $p = \infty$ folgt sie aus der Dreiecksungleichung. Sei also $1 < p, q < \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1 \Leftrightarrow p + q = pq \Leftrightarrow p = (p-1)q$.

$$\begin{aligned}
|x_1 + y_1|^p + \ldots + |x_n + y_n|^p &= |x_1 + y_1||x_1 + y_1|^{p-1} + \ldots + |x_n + y_n||x_n + y_n|^{p-1} \\
&\leq (|x_1| + |y_1|)|x_1 + y_1|^{p-1} + \ldots + (|x_n| + |y_n|)|x_n + y_n|^{p-1} \\
&\leq (\|x\|_p + \|y\|_p)(|x_1 + y_1|^{(p-1)q} + \ldots + |x_n + y_n|^{(p-1)q})^{1/q} \\
&= (\|x\|_p + \|y\|_p)\|x + y\|_p^{p/q}
\end{aligned}$$

In der dritten Zeile haben wir dabei die Höldersche Ungleichung benutzt. Also erhalten wir $\|x+y\|_p^p \leq (\|x\|_p + \|y\|_p)\|x+y\|_p^{p/q}$. Wenn $\|x+y\|_p = 0$ ist die Aussage trivial. Aus $\|x + y\|_p \neq 0$ folgt $\|x + y\|_p = \|x + y\|^{p(1-\frac{1}{q})} = \|x + y\|_p^{(p-\frac{p}{q})} \leq \|x\|_p + \|y\|_p$. **q.e.d.**

Definition 9.13. *(offener Ball, Umgebung, offene Menge) Ein offener Ball in (X, d) mit Zentrum $x \in X$ und Radius $r > 0$ ist die Menge $B(x, r) = \{y \in X \mid d(x, y) < r\}$. Eine Umgebung eines Punktes $x \in X$ ist eine Menge $O \subset X$, die für ein $r > 0$ einen Ball $B(x, r)$ enthält. Eine offene Menge $O \subset X$ ist eine Teilmenge, die eine Umgebung aller ihrer Punkte ist, d.h. für alle $x \in O$ gibt es ein $\epsilon > 0$, so dass $B(x, \epsilon) \subset O$.*

Beispiel 9.14. *In $\mathbb{R}$ besteht der Ball $B(x, r)$ aus $(x - r, x + r)$. Im $\mathbb{R}^n$ besteht der Ball $B(x, r)$ aus allen Punkten, deren euklidischer Abstand zu x kleiner ist als r.*

Alle offenen Bälle $B(x, r)$ sind offenbar Umgebungen von x. Für $y \in B(x, r)$ ist $d(x, y) < r$. Sei $z \in B(y, r - d(x, y))$. Dann gilt $d(x, z) \leq d(x, y) + d(y, z) < r$, also auch $B(y, r - d(x, y)) \subset B(x, r)$. Deshalb sind die offenen Bälle tatsächlich offen.

Offenbar ist die beliebige Vereinigung von offenen Mengen wieder offen. Für zwei offene Mengen O und O' und $x \in O \cap O'$ gibt es $r > 0$ und $r' > 0$ mit $B(x, r) \subset O$ und $B(x, r') \subset O'$. Also ist $B(x, \min\{r, r'\}) \subset B(x, r) \cap B(x, r') \subset O \cap O'$, und $O \cap O'$ offen. Damit ist auch die Schnittmenge von endlich vielen offenen Mengen wieder offen.

Definition 9.15. *(abgeschlossene Mengen, Abschluss) Die Komplemente von offenen Mengen heißen abgeschlossen. Der Abschluss $\bar{A}$ eine Menge A ist die Schnittmenge aller abgeschlossenen Mengen, die A enthalten.*

Wegen der Regel von de Morgan, sind beliebige Schnitte und endliche Vereinigungen von abgeschlossenen Mengen wieder abgeschlossen. Deshalb ist eine Menge genau dann abgeschlossen, wenn sie mit ihrem Abschluss übereinstimmt.

Definition 9.16. *Zwei Normen $\| \ \|_1$ und $\| \ \|_2$ heißen äquivalent, wenn es Konstanten $C_1, C_2 > 0$ gibt, so dass für alle $v \in V$ gilt*

$$\|v\|_1 \leq C_1 \|v\|_2 \quad \textit{und} \quad \|v\|_2 \leq C_2 \|v\|_1.$$

Offenbar ist die Relation zwischen Normen äquivalent zu sein eine Äquivalenzrelation, also insbesondere transitiv. Denn aus

$$\begin{array}{llll} \|v\|_1 \leq C_1\|v\|_2 & \|v\|_2 \leq C_2\|v\|_1 & \|v\|_2 \leq C_3\|v\|_3 & \|v\|_3 \leq C_4\|v\|_2 \\ \text{folgt} & \|v\|_1 \leq C_1C_3\|v\|_3 \quad \text{und} & \|v\|_3 \leq C_4C_2\|v\|_1. & \end{array}$$

Beispiel 9.17. *Auf den Vektorräumen $\mathbb{K}^n$ haben wir für $1 \leq p \leq \infty$ die Normen*

$$\| \cdot \|_p : \mathbb{K}^n \to \mathbb{R}, \quad v \mapsto \|v\|_p = \begin{cases} \left(\sum_{i=1}^n |v_i|^p \right)^{1/p} & \textit{für } p < \infty \\ \sup\{|v_1|, \ldots |v_n|\} & \textit{für } p = \infty. \end{cases}$$

eingeführt. Offenbar gilt für alle $1 \leq p < \infty$ und alle $v \in \mathbb{K}^n$

$$\|v\|_\infty \leq \|v\|_p \leq n^{1/p} \|v\|_\infty.$$

Also sind alle diese Normen äquivalent.

Äquivalente Normen besitzen offenbar die gleichen offenen Mengen, so dass eine Folge bezüglich einer Norm genau dann konvergiert, wenn sie bezüglich einer äquivalenten Norm konvergiert. Außerdem stimmen die entsprechenden stetigen Funktionen von äquivalenten Normen überein. Wenn umgekehrt die offenen Mengen von zwei Normen übereinstimmen, dann müssen sie äquivalent sein, weil dann jede Kugel um die Null der einen Norm einen Ball um die Null der anderen Norm enthalten muß.

9.2 Vollständigkeit und Kompaktheit

Zunächst verallgemeinern wir einige Aussagen über Zahlenfolgen auf allgemeine Folgen in metrischen Räumen.

Definition 9.18. *(Folgen und Cauchyfolgen) Eine Folge* $(x_n)_{n\in\mathbb{N}}$ *in einem metrischen Raum* (X, d) *ist eine Abbildung von* $\mathbb{N}$ *nach* X*, mit* $n \mapsto x_n$*. Eine Folge* $(x_n)_{n\in\mathbb{N}}$ *in einem metrischen Raum* (X, d) *konvergiert gegen* $x \in X$*, wenn es für jedes* $\epsilon > 0$ *ein* $N \in \mathbb{N}$ *gibt, so dass für alle* $n \geq N$ *gilt* $d(x_n, x) < \epsilon$*.*
Eine Folge $(x_n)_{n\in\mathbb{N}}$ *heißt Cauchyfolge, wenn es für jedes* $\epsilon > 0$ *ein* $N \in \mathbb{N}$ *gibt, so dass für alle* $n, m \geq N$ *gilt* $d(x_n, x_m) < \epsilon$*.*

Offenbar konvergiert eine Folge $(x_n)_{n\in\mathbb{N}}$ genau dann gegen einen Punkt a, wenn jede Umgebung von a alle bis auf endlich viele Folgenglieder enthält. Deshalb hängt der Begriff der Konvergenz nur von der Wahl der offenen Mengen ab.

Ein Punkt x gehört genau dann zu dem Abschluss $\bar{A}$, wenn alle offenen Mengen, die x enthalten, einen nicht leeren Schnitt mit A haben. Dies ist wiederum äquivalent dazu, dass es für jedes $n \in \mathbb{N}$ ein Element a_n in dem Ball $B(x, \frac{1}{n}) \cap A$ gibt, oder auch dazu, dass es eine Folge in A gibt, die gegen x konvergiert. Damit haben wir gezeigt:

Lemma 9.19. *Der Abschluss einer Teilmenge eines metrischen Raumes besteht aus allen Grenzwerten von Folgen innerhalb der Teilmenge, die in dem metrischen Raum konvergieren. Und eine Teilmenge ist genau dann abgeschlossen, wenn die Grenzwerte von allen konvergenten Folgen in der Teilmenge auch zu der Menge gehören.* **q.e.d.**

Wenn die Folge $(x_n)_{n\in\mathbb{N}}$ gegen x konvergiert, dann gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $n, m \geq N$ gilt $d(x_n, x) < \frac{\epsilon}{2}$ und $d(x_m, x) < \frac{\epsilon}{2}$. Dann gilt aber auch $d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \epsilon$. Damit haben wir gezeigt:

Satz 9.20. *In einem metrischen Raum sind konvergente Folgen Cauchyfolgen.* **q.e.d.**

Definition 9.21. *Ein metrischer Raum* (X, d) *heißt vollständig, wenn auch jede Cauchyfolge konvergiert.*

Definition 9.22. *Ein vollständiger normierter Vektorraum heißt Banachraum.*

Wegen dem Vollständigkeitsaxiom sind $\mathbb{R}^n$ und $\mathbb{C}^n$ für alle $n \in \mathbb{N}$ Banachräume. Wegen Lemma 9.19 ist eine Teilmenge eines vollständigen metrischen Raumes genau dann ein vollständiger metrischer Raum, wenn sie abgeschlossen ist.

Definition 9.23. *(kompakt) Eine Teilmenge der offenen Mengen von (X, d), die X überdeckt, (d.h. jedes Element von X ist in mindestens einer der offenen Mengen enthalten) heißt offene Überdeckung von X. Der metrische Raum heißt kompakt, wenn jede offene Überdeckung eine endliche Teilüberdeckung besitzt.*

Satz 9.24. *Für einen metrischen Raum (X, d) sind folgende Aussagen äquivalent:*

(i) *(X, d) ist kompakt.*

(ii) *Jede Folge in (X, d) besitzt eine konvergente Teilfolge.*

(iii) *(X, d) ist vollständig und für jedes $\epsilon > 0$ besitzt (X, d) eine endliche Überdeckung mit offenen Bällen vom Radius ϵ.*

Beweis: (i)⇒(ii): Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge ohne Häufungspunkt. Dann sind für alle $n \in \mathbb{N}$ die Mengen $F_n = \{x_m \mid m \geq n\}$ abgeschlossen, weil der Abschluss von jedem F_n gerade aus der Vereinigung von F_n mit den Häufungspunkten von $(x_n)_{n\in\mathbb{N}}$ besteht. Weil aber der Schnitt $\bigcap_{n\in\mathbb{N}} F_n$ nur aus Häufungspunkten von $(x_n)_{n\in\mathbb{N}}$ bestehen kann, bilden die Mengen $(X \setminus F_n)_{n\in\mathbb{N}}$ eine offene Überdeckung von X. Offenbar ist aber der Schnitt von endlich vielen Mengen der Mengen $(F_n)_{n\in\mathbb{N}}$ nicht leer. Also besitzt die Überdeckung $(X \setminus F_n)_{n\in\mathbb{N}}$ keine endliche Teilüberdeckung.

(ii)⇒(iii): Sei also (X, d) ein metrischer Raum, der (ii) erfüllt. Dann besitzt aber jede Cauchyfolge $(x_n)_{n\in\mathbb{N}}$ einen Häufungspunkt x. Dann gibt es für jedes $\epsilon > 0$ eine natürliche Zahl $N \in \mathbb{N}$, so dass alle $m, n \geq N$ auch $d(x_n, x_m) < \frac{\epsilon}{2}$ erfüllen. Weil x ein Häufungspunkt ist, gibt es aber ein $m \geq N$, so dass $d(x_m, x) < \frac{\epsilon}{2}$ gilt. Also gilt für alle $n \geq N$ auch

$$d(x_n, x) \leq d(x_n, x_m) + d(x_m, x) < \epsilon.$$

Also konvergiert $(x_n)_{n\in\mathbb{N}}$ gegen x und damit ist (X, d) vollständig. Sei $\epsilon > 0$ so gewählt, dass es keine endliche Überdeckung von X mit Bällen vom Radius ϵ gibt. Dann können wir für jedes $n \in \mathbb{N}$ eine Folge $(x_m)_{m\in\mathbb{N}}$ induktiv so definieren, dass $x_{n+1} \in X \setminus \bigcup_{m=1}^{n} B(x_m, \epsilon)$. Die Folge $(x_m)_{m\in\mathbb{N}}$ erfüllt also für alle $n \neq m \in \mathbb{N}$ $d(x_n, x_m) \geq \epsilon$. Also besitzt sie keine Teilfolge, die eine Cauchyfolge ist, und damit auch keinen Häufungspunkt im Widerspruch zu (ii).

(iii)⇒(i): Wir nehmen an (X, d) erfüllt Bedingung (iii) und $(U_\lambda)_{\lambda\in L}$ sei eine Überdeckung von (X, d), die keine endliche Teilüberdeckung besitzt. Dann definieren wir induktiv eine Folge $(x_n)_{n\in\mathbb{N}}$, so dass die Bälle $B(x_n, 2^{-n})$ keine endliche Teilüberdeckung von $(U_\lambda)_{\lambda\in L}$ besitzen und für alle $n \in \mathbb{N}$ die Bälle $B(x_{n+1}, 2^{-(n+1)})$ und $B(x_n, 2^{-n})$ nicht disjunkt sind. Weil nämlich (X, d) eine endliche Überdeckung von Bällen vom Radius

$2^{-(n+1)}$ besitzt, und weil $B(x_n, 2^{-n})$ keine endliche Teilüberdeckung von $(U_\lambda)_{\lambda\in L}$ besitzt, gibt es mindestens einen Ball vom Radius $2^{-(n+1)}$, der nichtleeren Schnitt mit $B(x_n, 2^{-n})$ hat und keine endliche Teilüberdeckung von $(U_\lambda)_{\lambda\in L}$ besitzt. Weil aber

$$d(x_n, x_{n+1}) \leq \frac{3}{2 \cdot 2^n} \quad \text{und} \quad \sum_{m=n}^{\infty} \frac{1}{2^m} = \frac{1}{2^n} \cdot \frac{1}{1-\frac{1}{2}} = 2^{-n+1},$$

ist die Folge $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge. Der Grenzwert gehört dann zu einer Menge $(U_\lambda)_{\lambda\in L}$, und es gibt ein $\epsilon > 0$, so dass $B(x,\epsilon) \subset U_\lambda$. Für genügend großes m ist dann $\frac{1}{2^{m-2}} \leq \epsilon$, und deshalb auch $d(x_m, x) \leq 3 \cdot 2^{-m}$ und

$$B(x_m, 2^{-m}) \subset B(x, 2^{-m+2}) \subset B(x, \epsilon).$$

Also besitzt ein Ball $B(x_m, 2^{-m})$ eine endliche Teilüberdeckung von $(U_\lambda)_{\lambda\in L}$im Widerspruch zu der Annahme, dass $(U_\lambda)_{\lambda\in L}$ keine endliche Teilüberdeckung besitzt. **q.e.d.**

Dieser Satz hat einige wichtige Folgerungen:

Korollar 9.25. **(i)** *Kompakte Mengen eines metrischen Raums sind abgeschlossen.*

(ii) *Abgeschlossene Teilmengen einer kompakten Menge sind wieder kompakt.*

Beweis:(i) Kompakte Mengen sind vollständig, und stimmen wegen Lemma 9.19 mit ihrem Abschluss überein.
(ii) Abgeschlossene Teilmengen einer kompakten Menge erfüllen offenbar wieder die Bedingung (ii) des vorangehenden Satzes. **q.e.d.**

Definition 9.26. *Eine Teilmenge A eines metrischen Raumes (X,d) heißt beschränkt, wenn für ein $x \in X$, die Menge der Abstände $\{d(x,y)|y \in A\}$ beschränkt ist.*

Wegen der Dreiecksungleichung ist diese Bedingung äquivalent dazu, dass für alle $x \in X$ die Mengen $\{d(x,y) \mid y \in A\}$ beschränkt sind, aber nicht uniform in $x \in X$.

Offenbar sind alle kompakten Teilmengen eines metrischen Raumes beschränkt und abgeschlossen. In $\mathbb{K}^n$ gilt auch die Umkehrung. Dort konvergiert eine Folge nämlich genau dann, wenn für alle $i = 1, \ldots, n$ die jeweilige Folge der i-ten Komponenten konvergieren. Deshalb überträgt sich der Beweis des Satzes 5.9 sofort auf den

Satz 9.27. *Eine Teilmenge $A \subset \mathbb{K}^n$ ist bezüglich einer der äquivalenten Normen $\|\cdot\|_p$ mit $p \in [1,\infty]$ genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Beispiel 9.28. **(i)** *Die Intervalle $[a,b]$ sind kompakt.*
(ii) *$\bar{\mathbb{N}}$ aus Beispiel (vi) ist kompakt.*
(iii) *Sei $(a_n)_{n\in\mathbb{N}}$ konvergent mit Grenzwert a. Dann ist $\{a\} \cup \{a_n|n \in \mathbb{N}\}$ kompakt.*

9.3 Stetigkeit

Definition 9.29. *Eine Abbildung $f : X \to Y, \quad x \mapsto f(x)$ von einem metrischen Raum (X, d) in den metrischen Raum (Y, d) heißt stetig in $x \in X$, wenn es für jedes $\epsilon > 0$ ein $\delta > 0$ gibt, so dass alle $y \in B(x, \delta) \subset X$ auch $f(y) \in B(f(x), \epsilon) \subset Y$ erfüllen. Die Abbildung f heißt stetig, wenn sie in allen Punkten von X stetig ist.*

Stetig im Punkt x heißt also, dass alle Punkte, die sehr nahe bei x liegen, auf Werte abgebildet werden, die sehr nahe bei $f(x)$ liegen.

Satz 9.30. *Für eine Abbildung $f : X \to Y, \quad x \mapsto f(x)$ zwischen den metrischen Räumen (X, d) und (Y, d) ist folgendes äquivalent:*

(i) *f ist stetig in x.*

(ii) *Das Urbild jeder Umgebung von $f(x)$ ist eine Umgebung von x.*

(iii) *Für jede Folge $(x_n)_{n\in\mathbb{N}}$ in (X, d), die gegen x konvergiert, konvergiert auch die Folge $(f(x_n))_{n\in\mathbb{N}}$ gegen $f(x)$.*

Beweis:(i)⇔(ii) Die Umgebungen von x sind gerade die Mengen, die einen δ-Ball um x enthalten. Also ist (ii) äquivalent zu der Aussage, dass das Urbild jedes ϵ-Balles um $f(x)$ einen δ-Ball um x enthält. Diese Aussage ist nur eine Umformulierung von (i).
(ii)⇔(iii) Die Folgen $(x_n)_{n\in\mathbb{N}}$ und $(f(x_n))_{n\in\mathbb{N}}$ konvergieren genau dann gegen x bzw. $f(x)$, wenn jede Umgebung von x bzw. $f(x)$ alle bis auf endlich viele Folgenglieder enthält. Wenn also $(x_n)_{n\in\mathbb{N}}$ gegen x konvergiert und f (ii) erfüllt, dann konvergiert auch $(f(x_n))_{n\in\mathbb{N}}$ gegen $f(x)$. Also folgt aus (ii) auch (iii). Wenn es umgekehrt einen ϵ-Ball von $f(x)$ gibt, dessen Urbild keinen δ-Ball von x enthält, dann gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ von Punkten $x_n \in B\left(x, \frac{1}{n}\right)$, so dass die Folge der entsprechenden Werte $f(x_n)$ im Komplement dieses ϵ-Balles von $f(x)$ liegt: $f(x_n) \not\in B(f(x), \epsilon)$. Also konvergiert $(x_n)_{n\in\mathbb{N}}$ gegen x aber $f(x_n)$ nicht gegen $f(x)$. **q.e.d.**

Korollar 9.31. *Für eine Funktion $f : X \to Y$ zwischen den metrischen Räumen (X, d) und (Y, d) ist folgendes äquivalent:*

(i) *f ist stetig.*

(ii) *Das Bild $(f(x_n))_{n\in\mathbb{N}}$ jeder konvergenten Folge $(x_n)_{n\in\mathbb{N}}$ ist konvergent und es gilt*

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right).$$

(iii) *Das Urbild jeder offenen Menge ist offen.*

(iv) *Das Urbild jeder abgeschlossenen Menge ist abgeschlossen.*

Beweis: Wegen dem vorangehenden Satz sind (i) und (ii) äquivalent. Weil eine Menge genau dann offen ist, wenn sie eine Umgebung von allen ihren Punkten ist, zeigt der vorangehende Satz, dass aus (i) bzw. (ii) auch (iii) folgt. Weil jede Umgebung eines Punktes auch eine offene Umgebung des Punktes enthält, folgt wieder wegen dem vorangehenden Satz aus (iii) auch (i) bzw. (ii). Weil nun die abgeschlossenen Mengen gerade die Komplemente der offenen Mengen sind und das Urbild eines Komplementes gerade gleich dem Komplement des Urbildes ist, ist (iii) zu (iv) äquivalent. **q.e.d.**

Korollar 9.32. *Die Komposition zweier stetiger Abbildungen ist stetig. Die analoge punktweise Aussage gilt auch.*

Beweis: Benutze die Äquivalenz zwischen (i) und (iii) im vorangehenden Korollar und die Gleichung

$$(f \circ g)^{-1}[A] = g^{-1}[f^{-1}[A]].$$

Beispiel 9.33. **(i)** *Auf jedem metrischen Raum ist die identische Abbildung $\mathbb{1}_X$ stetig.*

(ii) *Die konstante Abbildung, die alle $x \in X$ auf einen Punkt y abbildet ist stetig.*

(iii) *Wegen der Dreiecksungleichung gilt*

$$d(x,y) \le d(x,u) + d(u,y) \le d(x,u) + d(u,v) + d(v,y).$$

Also gilt auch $d(x,y) - d(u,v) \le d(x,u) + d(v,y)$. Durch vertauschen $(x,y) \leftrightarrow (u,v)$ und unter Benutzung der Symmetrie erhalten wir

$$d(u,v) - d(x,y) \le d(x,u) + d(v,y) \Rightarrow |d(x,y) - d(u,v)| \le d(x,u) + d(v,y).$$

Mit der Metrik aus dem Beispiel 9.2 (v) auf $X \times X$ ist $d : X \times X \to \mathbb{R}$ also stetig.

(iv) *In Bezug auf die induzierte Metrik ist wegen*

$$|\|v\| - \|w\|| \le \|v - w\| \quad \text{für alle } v, w \in V$$

auf jedem normierten Vektorraum V die Norm $\|\cdot\|$ eine stetige reelle Funktion.

(v) *Wegen der Dreiecksungleichung sind für jeden normierten Vektorraum V*

$$+ : V \times V \to V, \qquad (v,w) \mapsto v + w \qquad \cdot : \mathbb{K} \times V \to \mathbb{K}, \qquad (\lambda, v) \mapsto \lambda \cdot v$$

stetige Abbildungen. Das gilt auch für die Abbildungen

$$- : \mathbb{K} \to \mathbb{K}, \qquad x \mapsto -x \qquad \text{und} \qquad {}^{-1} : \mathbb{K} \setminus \{0\} \to \mathbb{K} \setminus \{0\}, \qquad x \mapsto x^{-1}.$$

(vi) *Eine Folge $(x_n)_{n \in \mathbb{N}}$ in einem metrischen Raum (X,d) läßt sich genau dann zu einer stetigen Abbildung von $\bar{\mathbb{N}}$ nach X fortsetzen, wenn sie konvergiert. ∞ wird dann auf $\lim\limits_{n \to \infty} x_n$ abgebildet.*

Korollar 9.34. *Das Bild einer kompakten Menge unter einer stetigen Abbildung ist kompakt.*

Beweis: Sei $f : X \to Y, \quad x \mapsto f(x)$ eine stetige Abbildung und $A \subset X$ eine kompakte Menge. Dann ist das Urbild einer beliebig offenen Überdeckung von dem Bild

$$f[A] = \{y \in Y \mid \exists x \in A \text{ mit } f(x) = y\}$$

eine offene Überdeckung von A. Diese besitzt, wenn A kompakt ist, eine endliche Teilüberdeckung. Also besitzt jede offene Überdeckung von $f[A]$ eine endliche Teilüberdeckung und $f[A]$ ist kompakt. **q.e.d.**

Korollar 9.35. *Alle stetigen Funktionen auf einem kompakten metrischen Raum sind beschränkt. Das Bild einer reellen stetigen Funktion auf einem kompakten metrischen Raum besitzt ein Minimum und ein Maximum.*

Beweis: Sei (X, d) kompakt und $f : X \to \mathbb{K}, \quad x \mapsto f(x)$ stetig. Dann ist $f[X]$ kompakt und damit auch beschränkt. Wegen Korollar 5.11 besitzen die kompakten Teilmengen von $\mathbb{R}$ sowohl ein Minimum als auch ein Maximum. **q.e.d.**

Korollar 9.36. *Sei f eine bijektive stetige Abbildung von einem kompakten metrischen Raum (X, d) auf einen metrischen Raum (Y, d). Dann ist die Umkehrabbildung stetig.*

Beweis: Wegen dem vorangehenden Korollar ist das Bild $f[X] = Y$ kompakt. Weil aber wegen Korollar 9.25 eine Teilmenge eines kompakten metrischen Raumes genau dann abgeschlossen ist, wenn sie kompakt ist, folgt die Aussage aus dem vorangehenden Korollar und der Charakterisierung (iv) im Korollar über stetige Abbildungen. **q.e.d.**

Satz 9.37. *Auf $\mathbb{K}^n$ sind alle Normen paarweise äquivalent.*

Beweis: Es genügt zu zeigen, dass alle Normen äquivalent sind zu $\|\cdot\|_1$. Sei also $\|\cdot\|$ eine beliebige Norm. Sei $e_1, \ldots, e_n$ die Basis von $\mathbb{K}^n$, deren i–tes Element nur an der i–ten Stelle eine nicht verschwindende Komponente hat, die dann jeweils gleich Eins ist. Wegen der Dreiecksungleichung gilt dann für jedes $v = (v_1, \ldots, v_n) \in \mathbb{K}^n$

$$\|v\| \leq |v_1| \cdot \|e_1\| + \ldots + |v_n| \cdot \|e_n\| \leq \max\{\|e_1\|, \ldots, \|e_n\|\}\|v\|_1.$$

Aus der Dreiecksungleichung folgt dann für alle $v, w \in \mathbb{K}^n$

$$|\|v\| - \|w\|| \leq \|v - w\| \leq \max\{\|e_1\|, \ldots, \|e_n\|\}\|v - w\|_1.$$

Also ist die Abbildung $v \mapsto \|v\|$ stetig bezüglich der Norm $\|\cdot\|_1$, und wegen dem Satz 9.27 von Heine-Borel ist die Menge $\{v \in \mathbb{K}^n \mid \|v\|_1 = 1\}$ kompakt. Wegen Korollar 9.35 nimmt diese Funktion auf dieser Menge das Minimum C an. Wegen der Positivität von $\|\cdot\|$ gilt $C > 0$. Daraus folgt

$$C\|v\|_1 \leq \left\|\frac{v}{\|v\|_1}\right\| \|v\|_1 = \|v\| \leq \max\{\|e_1\|, \ldots, \|e_n\|\}\|v\|_1 \text{ für alle } v \in \mathbb{K}^n \setminus \{0\}.$$ **q.e.d.**

Definition 9.38. *(Gleichmässige Stetigkeit, Lipschitzstetigkeit) Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen heißt gleichmäßig stetig, wenn es für alle $\epsilon > 0$ ein $\delta > 0$ gibt, so dass $d(f(x), f(y)) < \epsilon$ für alle $x, y \in X$ mit $d(x, y) < \delta$ gilt.*

Die Abbildung heißt lipschitzstetig auf A, wenn es eine Konstante $L > 0$ (Lipschitzkonstante) gibt, so dass für alle $x, y \in A$ gilt $d(f(x), f(y)) \leq Ld(x, y)$.

Offenbar ist jede lipschitzstetige Abbildung auch gleichmäßig stetig und jede gleichmäßig stetige Abbildung auch stetig. Es gilt auch folgende Umkehrung:

Satz 9.39. *Sei $f : X \to Y, \quad x \mapsto f(x)$ eine stetige Abbildung zwischen metrischen Räumen. Dann ist f auf jeder kompakten Menge A auch gleichmäßig stetig.*

Auf kompakten Mengen sind also gleichmäßige und einfache Stetigkeit äquivalent.
Beweis: Sei also $f : X \to Y, \quad x \mapsto f(x)$ stetig und $A \subset X$ kompakt. Dann gibt es für jedes $\epsilon > 0$ und jedes $x \in A$ ein $\delta(x)$, so dass $f(y) \in B(f(x), \frac{\epsilon}{2})$ aus $y \in B(x, \delta(x))$ folgt. Wir wählen eine endliche Teilüberdeckung von der offenen Überdeckung $\{B(x, \frac{\delta(x)}{2}) \mid x \in A\}$ von A. Sei δ das Minimum der Radien dieser endlichen Teilüberdeckung. Dann gibt es für alle $y, z \in A$ mit $d(y, z) < \delta$ einen Ball $B(x, \frac{\delta(x)}{2})$ der endlichen Teilüberdeckung mit $y \in B(x, \frac{\delta(x)}{2})$. Dann folgt

$$d(x, z) \leq d(x, y) + d(y, z) < \frac{\delta(x)}{2} + \frac{\delta(x)}{2} = \delta(x) \quad \text{also } z \in B(x, \delta(x)).$$

Daraus folgt $$d(f(y), f(z)) \leq d(f(x), f(y)) + d(f(x), f(z)) < \epsilon.$$ **q.e.d.**

Übungsaufgabe 9.40. *Zeige in mehreren Schritten, dass sich jeder metrische Raum (X, d) auf eindeutige Weise vervollständigen läßt.*

(i) *Auf dem Raum aller Cauchyfolgen in (X, d) ist die Relation*

$$(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}} \iff \textit{die reelle Folge } (d(x_n, y_n))_{n\in\mathbb{N}} \textit{ konvergiert gegen Null.}$$

eine Äquivalenzrelation. Die Menge der entsprechenden Äquivalenzklassen bezeichnen wir mit $\tilde{X}$.

(ii) *Für Cauchyfolgen $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ existiert der Grenzwert $\tilde{d}((x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}) = \lim\limits_{n\to\infty} d(x_n, y_n)$ und hängt nur von den Äquivalenzklassen der Cauchyfolgen ab.*

(iii) *Die entsprechende Abbildung $\tilde{d} : \tilde{X} \times \tilde{X} \to \mathbb{R}$ definiert eine Metrik.*

(iv) *$(\tilde{X}, \tilde{d})$ ist ein vollständiger metrischer Raum.*

(v) *Die konstanten Folgen definieren eine isometrische (mit beiden Metriken verträgliche) Abbildung $X \to \tilde{X}$, und das Bild dieser Abbildung liegt dicht in $\tilde{X}$ (d.h. der Abschluss von dem Bild ist gleich $\tilde{X}$).*

(vi) *Zeige, dass sich jede gleichmäßig stetige Abbildung f von (X, d) in einen vollständigen metrischen Raum (Y, d) eindeutig zu einer stetigen Abbildung $\tilde{f}$ von $(\tilde{X}, \tilde{d})$ nach (Y, d) fortsetzen läßt. Daraus folgt dann, dass $\tilde{X}$ sich isometrisch und bijektiv auf den Abschluss des Bildes jeder isometrischen Abbildung von X in einen vollständigen metrischen Raum abbilden läßt.*

Weil jede reelle Zahl der Grenzwert einer Folge von rationalen Zahlen ist, sind die reellen Zahlen die Vervollständigung des metrischen Raums der rationalen Zahlen. Anstelle unserer axiomatischen Charakterisierung der reellen Zahlen können wir also die reellen Zahlen auch als die Vervollständigung der rationalen Zahlen konstruieren.

9.4 Funktionenräume

In diesem Abschnitt sei (X, d) ein metrischer Raum und (Y, d) ein metrischer Raum, ein normierter Vektorraum oder eine normierte Algebra:

Definition 9.41. *Eine normierte Algebra ist ein normierter Vektorraum V mit einer assoziativen und distributiven Multiplikation $\cdot : V \times V \to V$, die Folgendes erfüllt:*

$$\begin{aligned} &(v + v'') \cdot v' = v \cdot v' + v'' \cdot v && (\lambda v) \cdot v' = \lambda(v \cdot v') && \|v \cdot v'\| \leq \|v\| \cdot \|v'\| \\ &v \cdot (v' + v'') = v \cdot v' + v \cdot v'' && v \cdot (\lambda v') = \lambda(v \cdot v') && \text{für alle } v, v', v'' \in V, \lambda \in \mathbb{K}. \end{aligned}$$

Wenn V vollständig ist, heißt V Banachalgebra.

Wir betrachten in diesem Abschnitt Mengen von Abbildungen von X nach Y. Wenn Y ein normierter Vektorraum ist können wir solche Abbildungen punktweise miteinander addieren und mit Elemente von $\mathbb{K}$ multiplizieren, und wenn Y eine Algbera ist, auch punktweise miteinander multiplizieren:

$$\begin{aligned} f + g :& X \to Y, & x \mapsto& f(x) + g(x), & \lambda f :& X \to Y, & x \mapsto& \lambda f(x) \\ f \cdot g :& X \to Y, & x \mapsto& f(x) \cdot g(x). \end{aligned}$$

Die Addition erfüllt offenbar die Axiome $A1$ und mit der Skalarmultiplikation das Distributivgesetz. Dadurch wird die Menge aller Abbildungen in einen Vektorraum Y zu einem Vektorraum, und zu einer Algebra, wenn Y eine Algebra ist. Das Inverse einer Funktion f in eine Algebra mit Eins $\mathbb{1} \in Y$ existiert nur, wenn $f(x)$ für alle $x \in X$ invertierbar ist. Indem wir die Elemente von $\mathbb{K}$ mit den entsprechenden Vielfachen der Eins identifizieren, wird die Skalarmultiplikation zu einem Spezialfall der Multiplikation.

Definition 9.42. *Eine Folge von Funktionen $(f_n)_{n \in \mathbb{N}}$ von X nach Y heißt*

punktweise konvergent, *wenn die Folgen $(f_n(x))_{n \in \mathbb{N}}$ für jedes $x \in X$ konvergieren. Die Grenzwerte definieren wieder eine Funktion $f : x \to Y, \quad x \mapsto \lim_{n \to \infty} f_n(x)$.*

gleichmäßig konvergent, *wenn es eine Funktion* $f : X \to Y, \quad x \mapsto f(x)$ *gibt, und für alle* $\epsilon > 0$ *ein* $N \in \mathbb{N}$*, so dass* $d(f_n(x), f(x)) < \epsilon$ *für* $n \geq N$ *und* $x \in X$ *gilt.*

Offenbar ist jede gleichmäßige konvergente Folge (f_n) auch punktweise konvergent, aber nicht umgekehrt (siehe Beispiel 5.24).

Definition 9.43. *Eine Funktion* $f : X \to Y, \quad x \mapsto f(x)$ *heißt beschränkt, wenn das Bild* $f[X]$ *eine beschränkte Menge in* Y *ist.* $B(X, Y)$*, bezeichne die Menge aller beschränkten Abbildungen von* X *nach* Y*. Auf* $B(X, Y)$ *bezeichne* d *folgende Abbildung:*

$$d : B(X, Y) \times B(X, Y) \to \mathbb{R}, \quad (f, g) \mapsto d(f, g) = \sup\{d(f(x), g(x)) \mid x \in X\}.$$

Wenn Y *ein normierter Vektorraum ist, dann bezeichne* $\|\cdot\|_\infty$ *folgende Abbildung:*

$$\|\cdot\|_\infty : B(X, Y) \to \mathbb{R}, \quad f \mapsto \|f\|_\infty = \sup_{x \in X} \|f(x)\|.$$

Satz 9.44. **(i)** *Für metrische Räume* X *und* Y *ist* d *eine Metrik auf* $B(X, Y)$*.*

(ii) *Wenn* Y *ein normierter Vektorraum (Algebra) ist, ist* $B(X, Y)$ *ein normierter Vektorraum (Algebra) mit Norm* $\|\cdot\|_\infty$*, die die Metrik aus (i) induziert.*

(iii) *Wenn* Y *ein vollständiger metrischer Raum ist, dann auch* $B(X, Y)$*.*

Beweis:(i) und **(ii)** folgen daraus, dass d eine Metrik bzw. $\|\cdot\|$ eine Norm auf Y ist, und weil wegen der Dreiecksungleichung und wegen $\|v \cdot w\| \leq \|v\| \cdot \|w\|$ die Summe und das Produkt zweier beschränkter Funktionen wieder beschränkt ist.
(iii) Sei $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in $B_{\mathbb{K}}(X, Y)$. Für alle $\epsilon > 0$ gibt es ein $N \in \mathbb{N}$ mit

$$d(f_n(x), f_m(x)) \leq d(f_n, f_m) < \frac{\epsilon}{2} \quad \text{für alle } n, m \geq N \text{ und alle } x \in X.$$

Dann sind für alle $x \in X$ die Folgen $(f_n(x))_{n \in \mathbb{N}}$ Cauchyfolgen. Also konvergieren sie punktweise gegen eine Funktion $f : X \to Y, \quad x \mapsto f(x)$. Wegen Beispiele 9.33 (iii) sind dann für alle $g \in B(X, Y)$ auch die Folgen $d(g, f_n)$ Cauchyfolgen in $\mathbb{R}$ und damit beschränkt. Für alle $\epsilon > 0$ und alle $x \in X$ gibt es ein $N(x) \in \mathbb{N}$, so dass für alle $m \geq N(x)$ gilt $d(f_n(x), f(x)) < \frac{\epsilon}{2}$. Damit folgt für $n \geq N$ und $m \geq \max\{N, N(x)\}$

$$d(f_n(x), f(x)) \leq d(f_n(x), f_m(x)) + d(f_m(x), f(x)) < \epsilon.$$

Also konvergiert $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f. **q.e.d.**

Definition 9.45. $C_b(X, Y)$ *sei der Unterraum von* $B(X, Y)$ *aller stetigen und beschränkten Funktionen von* X *nach* Y*.*

Satz 9.46. **(i)** *Für metrische Räume* X *und* Y *ist* $C_b(X, Y)$ *abgeschlossen in* $B(X, Y)$*.*

(ii) *Wenn Y ein normierter Vektorraum (Algebra) ist, dann auch $C_b(X,Y)$.*

(iii) *Wenn Y vollständig ist, dann auch $C_b(X,Y)$.*

Beweis: Wegen dem vorangehenden Satz und Lemma 9.19 genügt es zu zeigen, dass für jede Folge in $C_b(X,Y)$, die als Folge in $B(X,Y)$ konvergiert, der Grenzwert in $C_b(X,Y)$ liegt. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge in $C_b(X,Y))$, die in $B(X,Y)$ gegen f konvergiert. Dann gibt es für jedes $\epsilon > 0$ ein $n \in \mathbb{N}$ mit $d(f_n, f) < \frac{\epsilon}{3}$. Weil f_n stetig bei $x \in X$ ist gibt es ein $\delta > 0$, so dass $d(f_n(x), f_n(y)) < \frac{\epsilon}{3}$ für alle $y \in B(x,\delta)$ gilt. Dann folgt

$$\begin{aligned} d(f(x), f(y)) &\leq d(f(x), f_n(x)) + d(f_n(x), f_n(y)) + d(f_n(y), f(y)) \\ &\leq d(f(x), f_n(x)) + d(f_n(x), f_n(y)) + d(f_n(y), f(y)) \\ &< \epsilon. \qquad \text{Also ist } f \text{ bei } x \in X \text{ stetig.} \end{aligned}$$

q.e.d.

Die gleichmäßige Konvergenz der Folge $(f_n)_{n\in\mathbb{N}}$ ist notwendig (siehe Beispiel 5.24). Wenn Y ein Banachraum ist, dann sind sowohl $B(X,Y)$ als auch $C_b(X,Y)$ Banachräume. Wenn Y eine Banachalgebra ist wie z.B. $\mathbb{K}$, dann sind auch $B(X,Y)$ und $C_b(X,Y)$ Banachalgebren. Der Fall $Y = \mathbb{K}$ wird im Folgenden noch öfter vorkommen. Jetzt können wir die Vervollständigungen aller metrischen Räume leicht konstruieren:

Satz 9.47. *Sei X ein metrischer Raum und $x_0 \in X$. Für alle $x \in X$ gehört dann*

$$\begin{array}{lll} I(x) : X \to \mathbb{R} & y \mapsto d(x,y) - d(x_0,y) & \textit{zu } C_b(X,\mathbb{R}). \\ \textit{Die Abbildung} \quad I : X \to C_b(X,\mathbb{R}) & x \mapsto I(x) & \end{array}$$

ist eine isometrische Abbildung von X nach $C_b(X,\mathbb{R})$, d.h. es gilt $d(I(x), I(y)) = d(x,y)$ für alle $x, y \in X$. Für jede gleichmäßig stetige Abbildung f von X in einen vollständigen metrischen Raum Y, gibt es eine stetige Abbildung g von dem Abschluss des Bildes $I[X] \subset C_b(X,\mathbb{R})$ nach Y, so dass f gleich $g \circ I$ ist (vergleiche Übungsaufgabe 9.40).

Beweis: Wegen Beispiel 9.33 (iii) sind die reellen Funktionen $I(x)$ für alle $x \in X$ stetig. Wegen der Dreiecksungleichung gilt für alle $y \in X$

$$I(x)(y) = d(x,y) - d(x_0,y) \leq d(x,x_0) + d(x_0,y) - d(x_0,y) = d(x,x_0).$$

Also gehören diese Funktionen zu $C_b(X,\mathbb{R})$. Für $x, y \in X$ gilt

$$d(x,y) = d(x,y) - d(y,y) \leq d(I(x), I(y)) = \sup_{z\in X} |d(x,z) - d(y,z)| \leq d(x,y).$$

Also ist I eine isometrische Abbildung. Sei $h \in C_b(X,\mathbb{R})$ ein Element im Abschluss von $I[X]$ und $f : X \to Y$ eine gleichmäßig stetige Abbildung in einen vollständigen metrischen Raum Y. Für jedes $\epsilon > 0$ gibt es ein $\delta > 0$, so dass $d(f(x), f(y)) < \epsilon$ aus $d(x,y) < \delta$ folgt. Insbesondere haben alle Elemente von $\{f(x) \in Y \mid x \in X \text{ mit } d(I(x), h) < \frac{\delta}{2}\}$ paarweise einen Abstand kleiner als ϵ. Deshalb werden alle Folgen in X, deren Bilder unter I gegen h konvergieren, auf Cauchyfolgen in Y abgebildet, die alle gegen das gleiche Element von Y konvergieren. Indem wir h durch g auf diesen Grenzwert in Y abbilden, erhalten wir eine Abbildung g mit den gewünschten Eigenschaften. **q.e.d.**

Satz 9.48. *(Satz von Stone–Weierstraß) Sei (X, d) ein kompakter metrischer Raum und $A \subset C_b(X, \mathbb{R})$ eine Unteralgebra, die die konstanten Funktionen enthält und die Punkte trennt, d.h. für alle $x \neq y \in X$ gibt es $f \in A$ mit $f(x) \neq f(y)$. Dann ist der Abschluss von A gleich $C_b(X, \mathbb{R})$.*

Lemma 9.49. *Auf $[0, 1] \subset \mathbb{R}$ konvergiert die induktiv definierte Folge von Polynomen $p_{n+1}(x) = p_n(x) + \frac{1}{2}(x - p_n^2(x))$ mit $p_0 = 0$, gleichmäßig gegen die Funktion $x \mapsto \sqrt{x}$.*

Beweis : Wir zeigen zunächst mit vollständiger Induktion, dass $0 \le p_n(x)$ und $0 \le p_n^2(x) \le x$ für $x \in [0, 1]$ und alle $n \in \mathbb{N}_0$ gilt. Beides ist für $n = 0$ offensichtlich.

$$\begin{aligned} x - p_{n+1}^2(x) &= x - p_n^2(x) - p_n(x)\left(x - p_n^2(x)\right) - \frac{1}{4}\left(x - p_n^2(x)\right)^2 \\ &= (x - p_n^2(x))\left(1 - p_n(x) - \frac{1}{4}(x - p_n^2(x))\right) \\ &= (x - p_n^2(x))\left(\left(1 - \frac{p_n(x)}{2}\right)^2 - \frac{x}{4}\right) \end{aligned}$$

Aus $p_n^2(x) \le x \le 1$ folgt $p_n(x) \le 1$ und damit $1 - \frac{p_n(x)}{2} \ge \frac{1}{2}$ und $(1 - \frac{p_n(x)}{2})^2 \ge \frac{1}{4} \ge \frac{x}{4}$. Deshalb ist die Folge $(p_n(x))_{n\in\mathbb{N}}$ für $x \in [0, 1]$ monoton wachsend und $0 \le p_n^2(x) \le x$. Dann gilt auch $((1-\frac{p_n(x)}{2})^2 - \frac{x}{4}) \le 1 - \frac{x}{4}$ und deshalb auch $0 \le x - p_n^2(x) \le x \cdot \left(1 - \frac{x}{4}\right)^n$. Wegen $\frac{1}{1-\frac{x}{4}} = \sum_{n=0}^{\infty} (\frac{x}{4})^n \ge 1 + \frac{x}{4}$ folgt dann aus der Bernoulli Ungleichung $\frac{1}{(1-\frac{x}{4})^n} \ge 1 + \frac{nx}{4}$ und $0 \le x - p_n^2(x) \le \frac{x}{1+\frac{nx}{4}} < \frac{4}{n}$. Also konvergiert $(p_n^2(x))_{n\in\mathbb{N}}$ auf $x \in [0, 1]$ gleichmäßig gegen x. Die Funktion $[0, 1] \to [0, 1], \quad x \mapsto \sqrt{x}$ ist die Umkehrfunktion von $[0, 1] \to [0, 1], \quad x \mapsto x^2$. Weil die zweite Funktion stetig ist, ist wegen Korollar 5.18 die erste stetig und wegen Satz 5.22 sogar gleichmäßig stetig. Dann konvergiert die Folge $(p_n(x) = \sqrt{p_n^2(x)})_{n\in\mathbb{N}}$ gleichmäßig gegen $\sqrt{x}$. **q.e.d.**

Beweis des Satzes von Stone–Weierstraß: Wegen Lemma 9.49 gibt es eine Folge von Polynomen, die auf $[0, 1]$ gleichmäßig gegen $x \to \sqrt{x}$ konvergieren. Deshalb gehört $|f| = \|f\|_\infty \sqrt{(\frac{f}{\|f\|_\infty})^2}$ für jedes $f \in A$ zu dem Abschluss $\bar{A}$ von A. Für $f, g \in \bar{A}$ gehören

$$\sup(f, g) = \frac{1}{2}(f + g + |f - g|) \quad \text{und} \quad \inf(f, g) = \frac{1}{2}(f + g - |f - g|)$$

zu $\bar{A}$. Weil A die Punkte von X trennt, gibt es für alle $x \neq y \in X$ und alle $\alpha, \beta \in \mathbb{R}$ ein Element $f \in A$ mit $f(x) = \alpha$ und $f(y) = \beta$. Sei nämlich g eine Funktion mit $g(x) \neq g(y)$. Dann ist $f = \alpha + \frac{\beta-\alpha}{g(y)-g(x)}(g - g(x))$ eine solche Funktion.

Sei jetzt $f \in C_b(X, \mathbb{R})$ eine fest vorgegebene Funktion und $\epsilon > 0$. Dann gibt es für alle $x, y \in X$ eine Funktion $g_{x,y} \in \bar{A}$ die bei x und y mit f übereinstimmt. Dann gibt es ein $\delta_{x,y} > 0$, so dass $g_{x,y}(z) < f(z) + \epsilon$ für alle $z \in B(y, \delta_{x,y})$ gilt. Durch Übergang zu einer endlichen Teilüberdeckung von $\{B(y, \delta_{x,y}) \mid y \in X\}$ und dem Infimum der entsprechenden Funktionen $g_{x,y} \in \bar{A}$ gibt es eine Funktion $g_x \in \bar{A}$, die $g_x(x) = f(x)$ und $g_x < f + \epsilon$ erfüllt. Wegen der Stetigkeit von f und g_x gibt es für alle $x \in X$ ein $\delta_x > 0$, so dass $f(y) - \epsilon < g_x(y)$ für alle $y \in B(x, \delta_x)$ gilt. Durch Übergang zu einer endlichen Teilüberdeckung von $\{B(x, \delta_x) \mid x \in X\}$ und dem Supremum der entsprechenden Funktionen g_x finden wir schließlich eine Funktion g in $\bar{A}$, die $f - \epsilon < g < f + \epsilon$ auf X erfüllt. Weil ϵ beliebig ist folgt $f \in \bar{A}$. **q.e.d.**

Satz 9.50.* *(Satz von Dini) Auf einem kompakten metrischen Raum (X, d) konvergiert eine monotone Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ von stetigen reellen Funktionen gleichmäßig, wenn sie punktweise gegen eine stetige Funktion f konvergiert.*

Beweis*: Sei $(f_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge in $C_b(X, \mathbb{R})$, die punktweise gegen $f \in C_b(X, \mathbb{R})$ konvergiert. Dann gibt es zu jedem $\epsilon > 0$ und $x \in X$ ein $n(x) \in \mathbb{N}$, so dass $f(x) - f_{n(x)}(x) < \frac{\epsilon}{3}$ gilt. Da $f_{n(x)}$ und f stetig sind gibt es ein $\delta(x)$, so dass

$$|f_{n(x)}(x) - f_{n(x)}(y)| < \frac{\epsilon}{3} \quad \text{und} \quad |f(x) - f(y)| < \frac{\epsilon}{3} \quad \text{für alle } y \in B(x, \delta(x)) \text{ gilt.}$$

Dann gilt dort auch $f(y) - f_{n(x)}(y) < \epsilon$. Wähle eine endliche Teilüberdeckung von $\{B(x, \delta(x)) \mid x \in X\}$. Dann gilt für $m \geq$ Maximum der entsprechenden $n(x)$

$$f(y) - f_m(y) \leq f(y) - f_{n(x)}(y) < \epsilon$$

auf den Mengen der Teilüberdeckung. Das zeigt die gleichmäßige Konvergenz. **q.e.d.**

Definition 9.51. *(relativkompakt) Eine Teilmenge eines metrischen Raumes heißt relativkompakt, wenn der Abschluss kompakt ist.*

Lemma 9.52. *Eine Teilmenge A eines metrischen Raumes (X, d) ist genau dann relativkompakt, wenn jede Folge in A eine in X konvergente Teilfolge besitzt.*

Beweis: Wenn A relativkompakt ist, dann besitzt wegen Satz 9.24 jede Folge in A eine konvergente Teilfolge, deren Grenzwert im Abschluss $\bar{A}$ liegt. Hat umgekehrt jede Folge in A eine konvergente Teilfolge, dann gibt es wegen Lemma 9.19 für jede Folge $(x_n)_{n\in\mathbb{N}}$ im Abschluss von A auch eine Folge $(a_n)_{n\in\mathbb{N}}$ in A mit $d(x_n, a_n) < \frac{1}{n}$. Dann konvergiert die jeder konvergenten Teilfolge von $(a_n)_{n\in\mathbb{N}}$ entsprechende Teilfolge von $(x_n)_{n\in\mathbb{N}}$ gegen den gleichen Grenzwert wie die entsprechende Teilfoge von $(a_n)_{n\in\mathbb{N}}$. Wegen Satz 9.24 ist dann der Abschluss von A kompakt. **q.e.d.**

Satz 9.53.* *(Arzela–Ascoli) Sei X ein kompakter und Y ein vollständiger metrischer Raum. Eine Teilmenge $\mathcal{F} \subset C_b(X, Y)$ ist genau dann relativkompakt, wenn*

(i) *für jedes $x \in X$ die Menge $\{f(x) \mid f \in \mathcal{F}\}$ relativkompakt ist, und*

(ii) *für jedes $x \in X$ die Menge $\mathcal{F}$ gleichgradig stetig ist in x, d.h. für jedes $x \in X$ und jedes $\epsilon > 0$ gibt es ein $\delta > 0$ mit $f(x') \in B(f(x), \epsilon)$ für alle $x' \in B(x, \delta), f \in \mathcal{F}$.*

Beweis*: Zunächst nehmen wir an, dass die Menge $\mathcal{F}$ die Bedingungen (i)-(ii) erfüllt. Wir zeigen dann, dass jede Folge $(f_n)_{n\in\mathbb{N}}$ in $\mathcal{F}$ eine in $C_b(X,Y)$ konvergente Teilfolge $(g_n)_{n\in\mathbb{N}}$ besitzt. Dafür zeigen wir zuerst, dass $\mathcal{F}$ auf X sogar gleichmäßig gleichgradig stetig ist. Für jedes $\epsilon > 0$ und jedes $y \in X$ gibt es wegen (ii) ein $\delta_y > 0$, so dass $d(f(x), f(y)) < \frac{\epsilon}{2}$ für alle $f \in \mathcal{F}$ aus $d(x,y) < 2\delta_y$ folgt. Wegen der Kompaktheit von X hat die Überdeckung $\{B(y,\delta_y) \mid y \in X\}$ eine endliche Teilüberdeckung $X = B(y_1,\delta_1) \cup \ldots \cup B(y_N,\delta_N)$. Sei δ das Minimum von $\delta_1,\ldots,\delta_N$. Dann enthält für alle Paare $x, x' \in X$ mit $d(x,x') < \delta$ einer der Bälle $B(y_1,\delta_1),\ldots,B(y_N,\delta_N)$ den einen Punkt x. Damit sind beide in einem der Bälle $B(y_1,2\delta_1),\ldots,B(y_N,2\delta_N)$ enthalten. Daraus folgt $d(f(x), f(x')) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ für alle $f \in \mathcal{F}$. Also ist $\mathcal{F}$ gleichgradig stetig auf ganz X.

Sei $(x_m)_{m\in\mathbb{N}}$ eine Folge, die in X dicht liegt. Wegen (i) ist für alle $m \in \mathbb{N}$ der Abschluss A_m der Menge $\{f_n(x_m) \mid n \in \mathbb{N}\}$ eine kompakte Teilmenge von Y. Wir definieren induktiv eine Teilfolge von $(g_n)_{n\in\mathbb{N}}$ von $(f_n)_{n\in\mathbb{N}}$ und eine Folge $(a_m)_{m\in\mathbb{N}}$ in Y, so dass $d(g_n(x_m), a_m) < \frac{1}{n}$ für alle $m \in \mathbb{N}$ und alle $n \geq m$ gilt. Dafür wählen wir zunächst einen Häufungspunkt a_1 von $(f_n(x_1))_{n\in\mathbb{N}}$ und eine Teilfolge $(g_n)_{n\in\mathbb{N}}$ von $(f_n)_{n\in\mathbb{N}}$, so dass $d(g_n(x_1), a_1) \leq \frac{1}{n}$ für alle $n \in \mathbb{N}$ gilt. Induktiv wählen wir für jedes $M \in \mathbb{N}\setminus\{1\}$ einen Häufungspunkt a_M von $(g_n(x_M))_{n\in\mathbb{N}}$ und ersetzen alle Folgenglieder von $(g_n)_{n\in\mathbb{N}}$ mit Indizes$\geq M$ durch eine Teilfolge von $(g_n)_{n\geq M}$, so dass $d(g_n(x_M), a_M) < \frac{1}{n}$ für alle $n \geq M$ gilt. Dann gilt $d(g_n(x_m), a_m) < \frac{1}{n}$ für alle $m = 1,\ldots,M$ und $n \geq m$.

Dann gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass aus $x, x' \in X$ mit $d(x,x') < \delta$ für alle $n \in \mathbb{N}$ folgt $d(g_n(x), g_n(x')) < \frac{\epsilon}{3}$. Die Überdeckung $(B(x_m,\delta))_{m\in\mathbb{N}}$ von X besitzt eine endliche Teilüberdeckung. Also gibt es ein $M \in \mathbb{N}$, so dass alle $l, n \geq M$ an den Zentren der Bälle der Teilüberdeckung $d(g_l(x_m), g_n(x_m)) < \frac{\epsilon}{3}$ erfüllen. Dann folgt für alle $x \in X$ und alle $l, n \geq M$

$$d(g_l(x), g_n(x)) \leq d(g_l(x), g_l(x_m)) + d(g_l(x_m), g_n(x_m)) + d(g_n(x_m), g_n(x)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Also ist $(g_n)_{n\in\mathbb{N}}$ in $C_b(X,Y)$ eine Cauchyfolge. Wegen der Bedingung (i) konvergiert sie dann in $B(X,Y)$. Wegen Satz 9.46 liegt der Grenzwert in $C_b(X,Y)$.

Wenn umgekehrt $\mathcal{F}$ relativkompakt ist, dann besitzt wegen Lemma 9.52 mit jeder Folge in $\mathcal{F}$ für jedes $x \in X$ auch die Folge der entsprechenden Funktionswerte eine konvergente Teilfolge. Also erfüllt $\mathcal{F}$ die Bedingung (i).

Außerdem gibt es für jedes $x \in X$ und $\epsilon > 0$ endlich viele $f_1,\ldots,f_k$ im Abschluss von $\mathcal{F}$, so dass $B(f_1,\epsilon/3) \cup \ldots \cup B(f_k,\epsilon/3)$ den Abschluss von $\mathcal{F}$ überdeckt. Weil $f_1,\ldots,f_k$ stetig sind, gibt es $\delta_1,\ldots,\delta_k > 0$, so dass für alle $i = 1,\ldots,k$ aus $x' \in B(x,\delta_i)$ folgt $f_i(x') \in B(f(x),\epsilon/3)$. Dann gibt es für alle $f \in \mathcal{F}$ ein f_i so dass für alle $x' \in B(x,\delta)$

$$d(f(x'), f(x)) \leq d(f(x'), f_i(x')) + d(f_i(x'), f_i(x)) + d(f_i(x), f(x)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

gilt mit $\delta = \min\{\delta_1,\ldots,\delta_k\}$. **q.e.d.**

9.5 Lineare Operatoren

Die Ableitung einer Funktion von mehreren Veränderlichen ist eine lineare Abbildung. Zur Vorbereitung der Differentialrechnung von Funktionen mehrerer Veränderlicher behandeln wir in diesem Abschnitt solche linearen Abbildungen zwischen normierten Vektorräumen. Dabei betrachten wir wieder Vektorräume über dem Körper $\mathbb{K}$.

Definition 9.54. *Eine Abbildung $A : V \to W$ von einem Vektorraum V in einen Vektorraum W heißt linear, wenn für alle $v, w \in V$ und $\lambda \in \mathbb{K}$ gilt*

$$A(v + w) = Av + Aw \quad \text{und} \quad A(\lambda v) = \lambda Av.$$

Satz 9.55. *Seien V und W normierte Vektorräume und $A : V \to W$ eine lineare Abbildung. Dann ist folgendes äquivalent:*

(i) *A ist stetig in 0.*

(ii) *A ist stetig.*

(iii) *A ist gleichmäßig stetig.*

(iv) *Es gibt ein $C > 0$, so dass für alle $v \in V$ gilt $\|Av\| \le C\|v\|$.*

(v) *A ist auf $B(0, 1)$ beschränkt, d.h. $\|Av\| \le C$ für alle $\|v\| < 1$ mit $0 < C < \infty$.*

Beweis:(i)⇒(v): Wenn A in 0 stetig ist, dann enthält das Urbild jedes Balles $B(0, \epsilon) \subset W$ einen Ball $B(0, \delta) \subset V$. Also gibt es ein $\delta > 0$, so dass $\|Av\| < 1$ aus $\|v\| < \delta$ folgt. Wegen der Linearität folgt dann $\|Av\| = \frac{1}{\delta}\|A\delta v\| < \frac{1}{\delta}$ aus $\|v\| < 1$. Also ist (v) erfüllt.
(v)⇒(iv): Wegen der Linearität folgt aus (v), dass für alle $v \in V$ gilt

$$\|Av\| = A\left(2\|v\| \cdot \frac{v}{2\|v\|}\right) = 2\|v\| \cdot A\left(\frac{v}{2\|v\|}\right) \le 2C\|v\|.$$

(iv)⇒(iii): Für $v, w \in V$ folgt $\|A(v - w)\| \le C\|v - w\|$ aus (iv). Also ist A sogar lipschitzstetig mit Lipschitzkonstante C. Dann gilt auch (iii).
(iii)⇒(ii) und (ii)⇒(i): Sind offensichtlich. **q.e.d.**

Satz 9.56. *Jede lineare Abbildung A von $\mathbb{K}^n$ in einen normierten Vektorraum ist stetig.*

Beweis: Wir benutzen wieder die Basis $e_1, \ldots, e_n$ von $\mathbb{K}^n$. Dann gilt für alle $v \in \mathbb{K}^n$

$$\|Av\| \le |v_1| \cdot \|Ae_1\| + \ldots + |v_n| \cdot \|Ae_n\| \quad \le \|v\|_1 \max\{\|Ae_1\|, \ldots, \|Ae_n\|\}.$$

Also folgt die Aussage aus Satz 9.37. **q.e.d.**

Definition 9.57. *Seien V, W normierte Vektorräume. Dann sei $\mathcal{L}(V, W)$ die Menge aller linearen stetigen Abbildungen von V nach W zusammen mit den Abbildungen:*

$$\begin{aligned} + : \quad \mathcal{L}(V,W) \times \mathcal{L}(V,W) &\to \mathcal{L}(V,W), \quad (A,B) \mapsto A + B : \quad V \to W, \quad v \mapsto Av + Bv \\ \cdot : \quad \mathbb{K} \times \mathcal{L}(V,W) &\to \mathcal{L}(V,W), \quad (\lambda, A) \mapsto \lambda A : \quad V \to W, \quad v \mapsto \lambda A v \end{aligned}$$

$$\| \cdot \| : \mathcal{L}(V,W) \to \mathbb{R}, A \mapsto \|A\| = \sup\{\|Av\| \mid v \in B(0,1)\} = \sup\{\|Av\| \mid v \in \overline{B(0,1)}\}.$$

Satz 9.58. *$\mathcal{L}(V, W)$ ist ein normierter Vektorraum.*

Beweis: Aus der Linearität von A und B folgt die Linearität von $A + B$ und $\lambda \cdot A$. Wegen der Dreiecksungleichung folgt aus der Stetigkeit von A und B auch die Stetigkeit von $A + B$. Und schließlich folgt aus der Linearität und der Stetigkeit von A auch die Stetigkeit von $\lambda \cdot A$. Weil W ein Vektorraum ist, ist dann auch $\mathcal{L}(V, W)$ ein Vektorraum. Wegen der Linearität der Elemente von $\mathcal{L}(V, W)$ und weil W ein normierter Vektorraum ist, ist auch $\mathcal{L}(V, W)$ ein normierter Vektorraum. **q.e.d.**

Wenn $V = \mathbb{K}^n$, dann ist der Abschluss der Einheitskugel $\overline{B(0,1)} = \{v \in \mathbb{K}^n \mid \|v\| \leq 1\}$ kompakt. Deshalb gibt es also für jedes $A \in \mathcal{L}(\mathbb{K}^n, W)$ ein $v \in \mathbb{K}^n \setminus \{0\}$, so dass gilt $\|A\| = \|A\frac{v}{\|v\|}\| = \frac{\|Av\|}{\|v\|}$. Weil für jeden linearen Operator $A \in \mathcal{L}(V, W)$ gilt $Av = \|v\| \cdot A(\frac{v}{\|v\|})$ ist jeder lineare Operator A durch seine Werte auf $\overline{B(0,1)}$ eindeutig bestimmt. Die Norm von $\mathcal{L}(V, W)$ ist dann einfach die Supremumsnorm der stetigen Abbildung von $\overline{B(0,1)}$ nach W. Deshalb ist der normierte Vektorraum $\mathcal{L}(V, W)$ ein Unterraum von $C_b(\overline{B(0,1)}, W)$. So folgt z.B. aus der Konvergenz einer Folge $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{L}(V, W)$ die gleichmäßige Konvergenz auf $\overline{B(0,1)}$ (und sogar die punktweise Konvergenz auf V).

Satz 9.59. *Seien V ein normierter Vektorraum und W ein Banachraum. Dann ist $\mathcal{L}(V, W)$ ein Banachraum.*

Beweis: Wir müssen wegen Satz 9.58 nur zeigen, dass $\mathcal{L}(V, W)$ vollständig ist. Sei $(A_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $\mathcal{L}(V, W)$. Für jedes $v \in V$ ist wegen $\|(A_n - A_m)v\| \leq \|A_n - A_m\| \cdot \|v\|$ die Folge $(A_n v)_{n\in\mathbb{N}}$ eine Cauchyfolge in W, die konvergiert. Wir definieren als den Grenzwert von $(A_n)_{n\in\mathbb{N}}$ folgende Abbildung von V nach W:

$$A : V \to W, \quad v \mapsto Av = \lim_{n\to\infty} A_n v \quad \text{für alle } v \in V.$$

Wir müssen noch $A \in \mathcal{L}(V, W)$ zeigen, und dass $(A_n)_{n\in\mathbb{N}}$ gegen A konvergiert. Weil $(A_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist, gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ mit $\|A_n - A_m\| < \frac{\epsilon}{2}$ für alle $n, m \geq N$. Für jedes $v \in V$ gibt es ein $m \geq N$ mit $\|Av - A_n v\| < \frac{\epsilon}{2}\|v\|$. Es folgt

$$\|(A - A_n)v\| \leq \|(A - A_m)v\| + \|(A_m - A_n)v\| < \left(\frac{\epsilon}{2} + \frac{\epsilon}{2}\right)\|v\| = \epsilon\|v\|.$$

Also konvergiert $(A_n)_{n\in\mathbb{N}}$ gegen A. Aus der Linearität von A_n folgt für alle $n \in \mathbb{N}$

$$\begin{aligned}
&\|A(v+w)-(Av+Aw)\| \leq \\
&\leq \|(A-A_n)(v+w)-(A-A_n)v-(A-A_n)w))\| + \|A_n(v+w)-(A_nv+A_nw)\| \\
&\qquad\qquad \leq \|A-A_n\|\,(\|v+w\|+\|v\|+\|w\|)\,, \text{ und} \\
&\|\lambda Av - A(\lambda v)\| \leq \|\lambda(A-A_n)v-(A-A_n)(\lambda v)\| + \|\lambda A_nv - A_n(\lambda v)\| \\
&\qquad\qquad \leq \|A-A_n\|\,(|\lambda|\cdot\|v\|+\|\lambda v\|)\,.
\end{aligned}$$

Im Grenzwert $n \to \infty$ konvergieren die rechten Seiten gegen Null, so dass A linear ist. Weil die Konvergenz in $\mathcal{L}(V,W)$ die gleichmäßige Konvergenz auf $\overline{B(0,1)} \subset V$ ist, folgt aus Satz 9.44 (iii), dass der Grenzwert A auf $\overline{B(0,1)} \subset V$ beschränkt ist, und damit wegen Satz 9.55 stetig. **q.e.d.**

Satz 9.60. *Seien U, V und W normierte Vektorräume und $A \in \mathcal{L}(U,V)$ und $B \in \mathcal{L}(V,W)$, dann ist $B\circ A \in \mathcal{L}(U,W)$ und es gilt $\|B\circ A\| \leq \|B\|\cdot\|A\|$. Insbesondere ist die Abbildung $\circ : \mathcal{L}(U,V)\times\mathcal{L}(V,W) \to \mathcal{L}(U,W)$, $\quad (A,B)\mapsto B\circ A$ stetig.*

Beweis: Für alle $v \in U$ gilt $\|(B\circ A)u\| \leq \|B\|\cdot\|Au\| \leq \|B\|\cdot\|A\|\cdot\|u\|$. Also folgt die Ungleichung $\|B\circ A\| \leq \|B\|\cdot\|A\|$ aus Satz 9.55. Für zwei normierte Vektorräume V, W mit Normen $\|\cdot\|_V$ und $\|\cdot\|_W$ ist

$$\|\cdot\|_{V\times W} : V\times W \to \mathbb{R}, \quad (v,w)\mapsto \|v\|_V + \|w\|_W$$

eine Norm auf $V\times W$ und induziert die Metrik des kartesischen Produktes der metrischen Räume V und W. Für $(A,B),(A',B') \in \mathcal{L}(U,V)\times\mathcal{L}(V,W)$ gilt dann

$$\begin{aligned}
\|B\circ A - B'\circ A'\| &= \|B\circ A - B\circ A' + B\circ A' - B'\circ A'\| \\
&= \|B\circ(A-A') + (B-B')\circ A'\| \\
&\leq \|B\|\cdot\|A-A'\| + \|B-B'\|\cdot\|A'\| \\
&\leq (\|A-A'\|+\|B-B'\|)(\|B\|+\|A'\|) \\
&\leq (\|A-A'\|+\|B-B'\|)(\|B\|+\|A\|+\|A-A'\|) \\
&\leq (\|(A,B)-(A',B')\|(\|B\|+\|A\|+\|(A,B)-(A',B')\|).
\end{aligned}$$

Also ist diese Abbildung im Punkt $(A,B)\in\mathcal{L}(U,V)\times\mathcal{L}(V,W)$ stetig. **q.e.d.**

Wir bezeichnen die Komposition $B\circ A$ von linearen Operatoren auch mit BA.

Definition 9.61. *Auf einem normierten Vektorraum V ist $\mathcal{L}(V) = \mathcal{L}(V,V)$ mit*

$$\circ : \mathcal{L}(V)\times\mathcal{L}(V)\to\mathcal{L}(V), \quad (A,B)\mapsto AB \quad \textit{und} \quad \|\cdot\| : \mathcal{L}(V)\to\mathbb{R}, \quad A\mapsto\|A\|$$

eine normierte Algebra. Für einen Banachraum V ist $\mathcal{L}(V)$ eine Banachalgebra.

Satz 9.62. *(Neumannsche Reihe) Sei V ein Banachraum und $A \in \mathcal{L}(V)$ ein Operator mit $\|A\| < 1$. Dann ist $\mathbb{1} - A$ invertierbar und es gilt $(\mathbb{1} - A)^{-1} = \sum_{n=0}^{\infty} A^n$.*

Beweis: Wegen $\|A^n\| \leq \|A\|^n$ ist $\left(\sum A^n\right)_{n\in\mathbb{N}_0}$ für $\|A\| < 1$ eine Cauchyfolge mit

$$\left\| \sum_{n=0}^{N} A^n \right\| \leq \frac{1}{1 - \|A\|}.$$

Wegen Satz 9.59 konvergiert diese Reihe gegen ein $B \in \mathcal{L}(V)$. Offenbar gilt

$$(\mathbb{1} - A)B = \sum_{n=0}^{\infty} A^n - \sum_{n=1}^{\infty} A^n = \mathbb{1} \quad \text{und genauso} \quad B(\mathbb{1} - A) = \sum_{n=0}^{\infty} A^n - \sum_{n=1}^{\infty} A^n = \mathbb{1}.$$

Also ist $(\mathbb{1} - A)$ invertierbar und es gilt $\qquad (\mathbb{1} - A)^{-1} = \sum_{n=0}^{\infty} A^n.$

Insbesondere gilt $\qquad \left\|(\mathbb{1} - A)^{-1}\right\| \leq \dfrac{1}{1 - \|A\|}.$ **q.e.d.**

Jede Potenzreihenfunktion $f(x) = \sum_{n\in\mathbb{N}_0} a_n x^n$ mit Konvergenzradius $R > 0$ definiert also eine Abbildung $\quad f : \{A \in \mathcal{L}(V) \mid \|A\| < R\} \to \mathcal{L}(V), A \mapsto f(A) = \sum_{n=0}^{\infty} a_n A^n.$ Viele der Aussagen, die wir für Potenzreihenfunktionen auf $\mathbb{K}$ gezeigt haben, lassen sich jetzt auf Potenzreihenfunktionen auf $\mathcal{L}(V)$ ausdehnen. Aber weil im Allgemeinen $AB \neq BA$ für $A, B \in \mathcal{L}(V)$, gilt im Allgemeinen auch $\exp(A)\exp(B) \neq \exp(A + B)$.

Definition 9.63. *Eine Derivation einer Algebra $\mathcal{L}(V)$ ist ein Operator $D \in \mathcal{L}(\mathcal{L}(V))$, der die Bedingung $D(AB) = D(A) \cdot B + A \cdot D(B)$ erfüllt.*

Übungsaufgabe 9.64. (i) *Zeige, dass für jedes $A \in \mathcal{L}(V)$, die Abbildung*

$$D_A : \mathcal{L}(V) \to \mathcal{L}(V), \quad B \mapsto AB - BA \quad \textit{eine Derivation ist.}$$

(ii) *Sei V ein Banachraum und D eine Derivation von $\mathcal{L}(V)$. Zeige dass $\exp(D)$ ein Algebraisomorphismus ist, d.h. ein invertierbares Element von*

$$\{C \in \mathcal{L}(\mathcal{L}(V)) \mid C(AB) = C(A)C(B) \textit{ für alle } A, B \in \mathcal{L}(V)\} \subset \mathcal{L}(\mathcal{L}(V)).$$

(iii) *Zeige $\exp(D_A)B = \exp(A) \cdot B \cdot \exp(-A) \quad \forall A, B \in \mathcal{L}(V)$ eines Banachraums V.*

Kapitel 10

Differentialrechnung von Funktionen mehrerer Veränderlicher

10.1 Ableitungen von $f : X \to Y$

Definition 10.1. *(Ableitung) Seien X und Y normierte Vektorräume. Eine Abbildung f von einer offenen Menge $U \subset X$ nach Y heißt im Punkt $x_0 \in X$ differenzierbar, wenn es ein $A \in \mathcal{L}(X,Y)$ gibt, so dass die folgende Abbildung in x_0 stetig ist:*

$$U \to \mathbb{R}, \quad x \mapsto \begin{cases} \frac{\|f(x)-f(x_0)-A(x-x_0)\|}{\|x-x_0\|} & \text{für } x \neq x_0 \\ 0 & \text{für } x = x_0. \end{cases}$$

A heißt Ableitung von f bei x_0 und wird mit $f'(x_0)$ oder $\frac{df}{dx}(x_0)$ bezeichnet. Wenn A und A' beide diese Bedingung erfüllen, dann gibt es für alle $\epsilon > 0$ ein $\delta > 0$ mit

$$\frac{\|(A-A')(x-x_0)\|}{\|x-x_0\|} \leq \frac{\|f(x)-f(x_0)-A(x-x_0)\|}{\|x-x_0\|} + \frac{\|f(x)-f(x_0)-A'(x-x_0)\|}{\|x-x_0\|} \leq \epsilon \quad \text{für alle } x \in B(x_0,\delta).$$

Also ist $A = A'$ und damit die Ableitung, wenn sie existiert, eindeutig.

Satz 10.2. *Sei $f : U \subset X \to Y$ in $x_0 \in U$ differenzierbar. Dann ist f in x_0 stetig.*

Beweis: Weil $f(x) - f(x_0) = A(x - x_0) + (f(x) - f(x_0) - A(x - x_0))$ folgt aus der Differenzierbarkeit von f in x_0, dass es ein $\delta > 0$ gibt, so dass $\|f(x) - f(x_0)\| \leq (\|A\| + 1)\|x - x_0\|$ gilt für alle $\|x - x_0\| < \delta$. Dann ist f auch stetig. **q.e.d.**

Beispiel 10.3. **(i)** *Sei f konstant. Dann ist* $$\frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} = 0 \quad \text{für } x \neq x_0.$$ *Also ist f differenzierbar mit $f'(x_0) = 0$.*

(ii) *Für $f \in \mathcal{L}(X,Y)$ und $x \neq x_0 \in X$ gilt* $\frac{\|f(x) - f(x_0) - f(x - x_0)\|}{\|x - x_0\|} = 0.$
Also ist f in x_0 differenzierbar und $f'(x_0) = f$.
(iii) *Die Abbildung $\mathbb{R} \to \mathcal{L}(\mathbb{R},\mathbb{R})$, $\quad x \mapsto$ Multiplikation mit x besitzt offenbar die Umkehrabbildung $\mathcal{L}(\mathbb{R},\mathbb{R}) \to \mathbb{R}$, $\quad A \mapsto A(1)$ und ist ein isometrischer Isomorphismus von normierten Vektorräumen. Deshalb können wir die Ableitungen von differenzierbaren reellen Funktionen f auf offenen Intervallen $(a,b) \subset \mathbb{R}$ mit reellen Funktionen auf diesen Intervallen identifizieren, die wir auch mit f' bezeichnen. Für $x \neq x_0$ gilt*

$$\frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} = \left\| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right\|.$$

Deshalb ist f als Funktion von der Teilmenge (a,b) des normierten Vektorrraumes $\mathbb{R}$ in den normierten Vektorraum $\mathbb{R}$ genau dann in x_0 differenzierbar, wenn f als reelle Funktion in x_0 im Sinne von Definition 7.1 differenzierbar ist.

Satz 10.4. **(i)** *Sei $f,g : U \subset X \to Y$ in $x_0 \in U$ differenzierbar. Dann sind $f + g$ und $\lambda \cdot f$ für alle $\lambda \in \mathbb{K}$ in x_0 differenzierbar und es gilt*

$$(f+g)'(x_0) + g'(x_0) \quad \text{bzw.} \quad (\lambda f)'(x_0) = \lambda f'(x_0).$$

(ii) *Seien $f,g : U \subset X \to Y$ in x_0 differenzierbar und Y eine normierte Algebra. Dann ist $f \cdot g$ in x_0 differenzierbar und es gilt*

$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0) \quad \text{mit}$$
$$(f'(x_0) \cdot g(x_0))x = (f'(x_0)x) \cdot g(x_0) \quad \text{und} \quad (f(x_0) \cdot g'(x_0))x = f(x_0) \cdot (g'(x_0)x).$$

(iii) *Seien $f : U \subset X \to Y$ und $g : V \subset Y \to Z$ differenzierbar in $x_0 \in U$ bzw. $y_0 = f(x_0) \in V$ mit $f[U] \subset V$. Dann ist $g \circ f$ im Punkt x_0 differenzierbar mit*

$$(g \circ f)'(x_0) = g'(f(x_0)) \circ f'(x_0).$$

Beweis:(i) Wegen der Dreiecksungleichung gilt

$$\frac{\|f(x) + g(x) - f(x_0) - g(x_0) - f'(x_0)(x - x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}$$
$$\leq \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} + \frac{\|g(x) - g(x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}.$$

(ii) Weil in einer normierten Algebra $\|AB\| \leq \|A\| \cdot \|B\|$ gilt, folgt

$$\frac{\|f(x) \cdot g(x) - f(x_0) \cdot g(x_0) - f'(x_0)(x - x_0)g(x_0) - f(x_0)g'(x_0)(x - x_0)\|}{\|x - x_0\|}$$
$$\leq \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|}\|g(x)\| + \|f'(x_0)\| \cdot \|g(x) - g(x_0)\| +$$
$$+ \|f(x_0)\| \frac{\|g(x) - g(x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}.$$

Weil g in x_0 stetig ist, gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass aus $\|x - x_0\| < \delta$ folgt $\|g(x) - g(x_0)\| < \epsilon$ bzw. $\|g(x)\| \leq \|g(x_0)\| + \epsilon$. Dann folgt (ii).

(iii) Aus Satz 9.60 folgt $\dfrac{\|(g \circ f)(x) - (g \circ f)(x_0) - g'(y_0)f'(x_0)(x - x_0)\|}{\|x - x_0\|} \leq$

$$\leq \frac{\|g(f(x)) - g(f(x_0)) - g'(f(x_0))(f(x) - f(x_0))\|}{\|f(x) - f(x_0)\|} \cdot \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} + \frac{\|g'(f(x_0))(f(x) - f(x_0) - f'(x_0)(x - x_0))\|}{\|x - x_0\|}$$

$$\leq \frac{\|g(f(x)) - g(f(x_0)) - g'(f(x_0))(f(x) - f(x_0))\|}{\|f(x) - f(x_0)\|} \cdot \left(\|f'(x_0)\| + \frac{\|f(x) - f(x_0) + f'(x_0)(x - x_0)\|}{\|x - x_0\|}\right) + \|g'(f(x_0))\| \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0))\|}{\|x - x_0\|}$$

Daraus folgt (iii). **q.e.d.**

10.2 Schrankensatz

Wir verallgemeinern in diesem Abschnitt den Schrankensatz auf differenzierbare Abbildungen zwischen normierten Vektorräumen.

Lemma 10.5. *Seien f eine stetige Abbildung von einem kompakten Intervall $[a, b]$ mit $a < b$ in einem normierten Vektorraum Y und ϕ eine stetige reelle Funktion auf $[a, b]$. Wenn im Komplement einer abzählbaren Teilmenge von (a, b) sowohl f als auch ϕ differenzierbar sind und dort gilt $\|f'\| \leq \phi'$, dann gilt auch $\|f(b) - f(a)\| \leq \phi(b) - \phi(a)$.*

Beweis: Sei $(x_n)_{n \in \mathbb{N}}$ eine Abzählung der Punkte in (a, b), an denen entweder f oder ϕ nicht differenzierbar ist oder nicht gilt $\|f'\| \leq \phi'$. Sei $\epsilon > 0$ und $A_\epsilon \subset [a, b]$ die Menge

$$\left\{ y \,\middle|\, \text{für alle } x \in (a, y) \text{ gilt } \|f(x) - f(a)\| \leq \phi(x) - \phi(a) + \epsilon(x - a) + \epsilon \sum_{x_n < x} 2^{-n} \right\}.$$

Aus der Stetigkeit von f und ϕ folgt, dass auch für $y = \sup A_\epsilon$ gilt

$$\begin{aligned} \|f(y) - f(a)\| &\leq \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sup_{x \in (a,y)} \sum_{x_n < x} 2^{-n} \\ &\leq \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sum_{x_n < y} 2^{-n}. \end{aligned}$$

Deshalb ist A_ϵ ein Intervall von der Form $A_\epsilon = [a, \sup A_\epsilon]$. Wenn $y \in (a, b)$ und f und ϕ in y differenzierbar sind und $\|f'(y)\| \leq \phi'(y)$ gilt, dann gibt es ein $\delta > 0$, so dass

$$\left|\frac{\phi(x) - \phi(y)}{x - y} - \phi'(y)\right| < \frac{\epsilon}{2} \text{ und } \frac{\|f(x) - f(y) - f'(y)(x - y)\|}{|x - y|} < \frac{\epsilon}{2}$$

für alle $x \in (y - \delta, y + \delta)$ gilt. Dann folgt für dieselben x

$$\begin{aligned}
\|f(x) - f(a)\| &\leq \|f(x) - f(y)\| + \|f(y) - f(a)\| \\
&< \left(\|f'(y)\| + \frac{\epsilon}{2}\right)|x - y| + \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sum_{x_n < y} 2^{-n} \\
&\leq \left(\phi'(y) + \frac{\epsilon}{2}\right)|x - y| + \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sum_{x_n < y} 2^{-n} \\
&< \left(\frac{|\phi(x) - \phi(y)|}{|x - y|} + \epsilon\right)|x - y| + \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sum_{x_n < y} 2^{-n}
\end{aligned}$$

Aus $\phi'(y) \geq 0$ folgt $\phi(x) \geq \phi(y)$ für $x \in (y, y + \delta)$ und damit auch

$$\|f(x) - f(a)\| < \phi(x) - \phi(a) + \epsilon(x - a) + \sum_{x_n < y} 2^{-n} \leq \phi(x) - \phi(a) + \epsilon(x - a) + \sum_{x_n < x} 2^{-n}.$$

Woraus folgt $x \in A_\epsilon$, im Widerspruch zu $x > \sup A_\epsilon$. Wenn es andererseits ein $N \in \mathbb{N}$ gibt mit $x_N = y$, dann gibt es ein $\delta > 0$, so dass aus $x \in (y, y + \delta)$ folgt

$$\|f(x) - f(y)\| - \phi(x) - \phi(y) < 2^{-N} \cdot \epsilon.$$

Dann folgt für dieselben x wieder $\|f(x) - f(a)\| \leq \|f(x) - f(y)\| + \|f(y) - f(a)\|$

$$< 2^{-N} \cdot \epsilon + \phi(x) - \phi(a) + \epsilon(y - a) + \epsilon \sum_{x_n < y} 2^{-n} \leq \phi(x) - \phi(a) + \epsilon(x - a) + \sum_{x_n < x} 2^{-n}.$$

Also gilt wieder $y + \delta \in A_\epsilon$, was $y = \sup A_\epsilon$ widerspricht. Dann muß aber $\sup A_\epsilon = b$ gelten. Weil das für alle $\epsilon > 0$ gilt, folgt auch $\|f(b) - f(a)\| \leq \phi(b) - \phi(a)$. **q.e.d.**

Korollar 10.6. *(Schrankensatz) Sei f eine Abbildung von einer offenen Teilmenge U des normierten Vektorraumes X in den normierten Vektorraum Y. Wenn im Komplement einer abzählbaren Teilmenge S von $D = \{(1 - t)a + tb \mid t \in [0, 1]\} \subset U$ die Abbildung f differenzierbar ist, und die Ableitung auf $D \setminus S$ beschränkt ist, dann gilt*

$$\|f(b) - f(a)\| \leq \|b - a\| \sup_{x \in D \setminus S} \|f'(x)\| \qquad \textit{und}$$

$$\|f(b) - f(a) - A(b - a)\| \leq \|b - a\| \sup_{x \in D \setminus S} \|f'(x) - A\| \qquad \textit{für } A \in \mathcal{L}(X, Y).$$

Beweis: Sei für ein $t_0 \in (0,1)$ die Abbildung f in $x_t = (1-t)a + tb \in U$ differenzierbar. Dann ist die Abbildung $[0,1] \to Y, \quad t \mapsto f(x_t)$ im Punkt t_0 differenzierbar, und es gilt

$$t \to \begin{cases} \frac{\|f(x_t) - f(x_{t_0}) - (t-t_0) f'(x_{t_0})(b-a)\|}{|t-t_0|} & \text{für } t \neq t_0 \\ 0 & \text{für } t = t_0 \end{cases}$$

ist stetig. Also ist die Ableitung von dieser Funktion in t_0 gleich $s \mapsto sf'(x_{t_0})(b-a) \in \mathcal{L}(\mathbb{R}, Y)$. Dann folgt die erste Behauptung aus Lemma 10.5 mit den beiden Funktionen

$$[0,1] \to Y, \quad t \mapsto f(x_t) \quad \text{und} \quad [0,1] \to \mathbb{R}, \quad t \mapsto t\|b-a\| \sup_{x \in D \setminus S} \|f'(x)\|.$$

Die zweite Behauptung folgt aus diesem Lemma mit den Funktionen

$$t \mapsto f(x_t) - tA(b-a) \in Y \qquad t \mapsto t\|b-a\| \sup_{x \in D \setminus S} \|f'(x) - A\| \in \mathbb{R}.$$

q.e.d.

10.3 Partielle Ableitungen

Definition 10.7. *(stetig differenzierbar) Sei $U \subset X$ eine offene Teilmenge eines normierten Vektorraumes, dann heißt eine Abbildung $f : U \to Y$ von U in einen normierten Vektorraum Y stetig differenzierbar, wenn*

(i) *f in allen $x_0 \in U$ differenzierbar ist, und*

(ii) *die Abbildung $f' : U \to \mathcal{L}(X, Y), \quad x \mapsto f'(x)$ stetig ist.*

Wegen Beispiel 10.3 (iii) stimmt im Falle von $X = Y = \mathbb{R}$ diese Definition mit der Definition von stetig differenzierbaren reellen Funktionen auf Intervallen in $\mathbb{R}$ überein.

Definition 10.8. *(partielle Ableitung) Seien X_1 und X_2 normierte Vektorräume und f eine Funktion von einer offenen Teilmenge $U \subset X_1 \times X_2$ des kartesischen Produktes der beiden normierten Vektorräume in den normierten Vektorraum Y. Dann heißt f im Punkt $(x_1, x_2) \in U \subset X_1 \times X_2$ partiell differenzierbar, falls die Abbildung $x \mapsto f(x, x_2)$ im Punkt $x = x_1$, und die Abbildung $x \mapsto f(x_1, x)$ im Punkt $x = x_2$ differenzierbar ist. Die Ableitungen heißen partielle Ableitungen an der Stelle (x_1, x_2) und werden mit $\frac{\partial f}{\partial x_1}(x_1, x_2)$ und $\frac{\partial f}{\partial x_2}(x_1, x_2)$ bezeichnet. Allgemeiner heißt eine Abbildung von einer offenen Menge $U \subset X_1 \times \ldots \times X_n$ eines n-fachen kartesischen Produktes von normierten Vektorräumen in einen normierten Vektorraum Y im Punkt $(x_1, \ldots, x_n) \in U \subset X_1 \times \ldots \times X_n$ partiell differenzierbar, wenn für alle $i = 1, \ldots, n$ die Abbildungen $x \to f(x_1, \ldots, x_i, x, x_{i+1}, \ldots, x_n)$ im Punkt $x = x_i$ differenzierbar sind.*

Die wichtigsten Beispiele sind Funktionen auf offenen Teilmengen des n-fachen kartesischen Produktes $\mathbb{R}^n$ von $\mathbb{R}$ mit sich selbst. Weil für jede lineare Abbildung $A \in \mathcal{L}(X_1 \times X_2, Y)$ die Abbildungen

$$A_1 : X_1 \to Y, \qquad x_1 \mapsto A(x_1, 0) \text{ und } \qquad A_2 : X_2 \to Y, \qquad x_2 \mapsto A(0, x_2)$$

stetige lineare Abbildungen sind, und weil für alle $x \in X_1$, mit $(x, x_2) \in U$

$$\frac{\|f(x,x_2) - f(x_1,x_2) - A((x,x_2) - (x_1,x_2))\|}{\|(x,x_2) - (x_1,x_2)\|} = \frac{\|f(x,x_2)f(x_1,x_2) - A(x - x_1, 0)\|}{\|x - x_1\|}$$

gilt, bzw. für alle $x \in X_2$, mit $(x_1, x) \in U$

$$\frac{\|f(x_1,x) - f(x_1,x_2) - A((x_1,x) - (x_1,x_2))\|}{\|(x_1,x) - (x_1,x_2)\|} = \frac{\|f(x_1,x)f(x_1,x_2) - A(0, x - x_2)\|}{\|x - x_1\|}$$

gilt, folgt aus den Definitionen der folgende

Satz 10.9. *Wenn eine Funktion $f : U \to Y$ von einer offenen Teilmenge $U \subset X_1 \times X_2$ des kartesischen Produktes zweier normierter Vektorräume in einen normierten Vektorraum Y im Punkt $(x_1, x_2) \in U$ differenzierbar ist, dann ist f in (x_1, x_2) auch partiell differenzierbar.* **q.e.d.**

Die Umkehrung dieses Satzes gilt nicht.

Beispiel 10.10.

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x,y) \neq 0 \\ 0 & \text{für } (x,y) = 0 \end{cases}$$ *besitzt die partiellen Ableitungen*

$$\frac{\partial f}{\partial x} = \begin{cases} \frac{2y}{x^2+y^2} - \frac{4x^2y}{(x^2+y^2)^2} & \text{für } (x,y) \neq 0 \\ 0 & \text{für } (x,y) = 0 \end{cases} \quad \frac{\partial f}{\partial y} = \begin{cases} \frac{2x}{x^2+y^2} - \frac{4y^2x}{(x^2+y^2)^2} & \text{für } (x,y) \neq 0 \\ 0 & \text{für } (x,y) = 0 \end{cases}$$

Aber f ist im Punkt $(x, y) = 0$ nicht stetig, weil für alle $r \in (0, \infty)$ und alle $\phi \in \mathbb{R}$ gilt

$$f(r\cos\phi, r\sin\phi) = 2\sin\phi\cos\phi = \sin(2\phi),$$

so dass für alle ϕ gilt $\lim\limits_{r\to 0+} f(r\cos\phi, r\sin\phi) = \sin(2\phi)$.

Aber es gilt folgende Umkehrung.

Satz 10.11. *Sei $f : U \to Y$ eine partiell differenzierbare Funktion von einer offenen Teilmenge U des kartesischen Produktes zwei normierter Vektorräume $X_1 \times X_2$ in dem normierten Vektorraum Y. Dann sind die partiellen Ableitungen von f*

$$\frac{\partial f}{\partial x_1} : U \to \mathcal{L}(X_1, Y) \quad \text{und} \quad \frac{\partial f}{\partial x_2} : U \to \mathcal{L}(X_2, Y)$$

genau dann auf U stetig, wenn f auf U stetig differenzierbar ist.

Beweis: Wenn $A_1 \in \mathcal{L}(X_1, Y)$ und $A_2 \in \mathcal{L}(X_2, Y)$, dann sind auch die Abbildungen

$$X_1 \times X_2 \to Y, \qquad (x_1, x_2) \to A_1 x_1 \text{ bzw.} \qquad X_1 \times X_2 \to Y, \qquad (x_1, x_2) \to A_2 x_2$$

linear stetige Abbildungen in $\mathcal{L}(X_1 \times X_2, Y)$. Also ist

$$A_1 \times A_2 : X_1 \times X_2 \to Y, \quad (x_1, x_2) \mapsto A_1 x_1 + A_2 x_2$$

eine stetige lineare Abbildung in $\mathcal{L}(X_1 \times X_2, Y)$. Umgekehrt sind für jede stetige lineare Abbildung $A \in \mathcal{L}(X_1 \times X_2, Y)$ die Abbildungen $A_1 : X_1 \to Y,\ x_1 \mapsto A(x_1, 0)$ und $A_2 : X_2 \to Y,\ x_2 \mapsto A(0, x_2)$ stetige lineare Abbildungen in $\mathcal{L}(X_1, Y)$ und $\mathcal{L}(X_2, Y)$. Wegen

$$\|A(x_1, x_2)\| = \|A(x_1, 0) + A(0, x_2)\| \leq \|A(x_1, 0)\| + \|A(0, x_2)\|$$

folgt

$$\|A\| \leq \|A_1\| + \|A_2\|.$$

Aufgrund der Definition der Norm von $\mathcal{L}(X_1 \times X_2, Y)$ folgt aber auch

$$\|A_1\| \leq \|A\| \text{ und } \|A_2\| \leq \|A\|.$$

Also ist die Abbildung $\quad \mathcal{L}(X_1, Y) \times \mathcal{L}(X_2, Y) \to \mathcal{L}(X_1 \times X_2, Y), \quad (A_1, A_2) \mapsto A$ eine bijektive Abbildung von normierten Vektorräumen und die beiden Normen von $\mathcal{L}(X_1, Y) \times \mathcal{L}(X_2, Y)$ und $\mathcal{L}(X_1 \times X_2, Y)$ sind bezüglich dieser Identifikation äquivalent. Daraus folgt, dass für jede stetig differenzierbare Funktion $f : U \to Y$, die beiden partiellen Ableitungen $\frac{\partial f}{\partial x_1} : U \to \mathcal{L}(X_1, Y)$ und $\frac{\partial f}{\partial x_2} : U \to \mathcal{L}(X_2, Y)$ stetig sind.

Wenn umgekehrt $\frac{\partial f}{\partial x_1}$ stetig ist, dann folgt

$$\begin{aligned} \left\| f(y_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\| &\leq \\ \leq \left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) \right\| &+ \\ + \left\| f(x_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\| &. \end{aligned}$$

Wegen der Stetigkeit von $\frac{\partial f}{\partial x_1}$ gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass für $z_1 \in B(x_1, \delta)$ und $z_2 \in B(x_2, \delta)$ gilt

$$\left\| \frac{\partial f}{\partial x_1}(z_1, z_2) - \frac{\partial f}{\partial x_1}(x_1, x_2) \right\| < \frac{\epsilon}{2}.$$

Aus dem Schrankensatz Korollar 10.6 folgt dann für $y_1 \in B(x_1, \delta)$ und $y_2 \in B(x_1, \delta)$

$$\left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) \right\| < \frac{\epsilon}{2} \|y_1 - x_1\|.$$

Weil $\frac{\partial f}{\partial x_2}$ in (x_1, x_2) existiert, gibt es auch ein $\delta' > 0$, so dass für $y_2 \in B(x_2, \delta')$ folgt

$$\left\| f(x_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\| < \frac{\epsilon}{2}\|y_2 - x_2\|.$$

Dann folgt für $y_1 \in B(x_1, \delta)$ und $y_2 \in B(x_2, \min\{\delta, \delta'\})$ auch

$$\left\| f(y_1, y_2) - f(x_1, x_2) - \frac{df(x_1, x_2)}{dx}((y_1, y_2) - (x_1, x_2)) \right\| < \epsilon \left\| (y_1 - x_1, y_2 - x_2) \right\|.$$

Also ist f auch differenzierbar. Weil dann die Ableitung durch die partiellen Ableitungen gegeben ist, ist dann f auch stetig differenzierbar. **q.e.d.**

Durch mehrmaliges Anwenden erhalten wir dann auch die entsprechende Aussage für Abbildungen von offenen Teilmengen U des n-fachen kartesischen Produktes von normierten Vektorräumen in einen normierten Vektorraum. Unsere wichtigsten Beispiele sind wieder Funktionen von offenen Teilmengen von $\mathbb{R}^n$ nach $\mathbb{R}$.

Beispiel 10.12. **(i)**
$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \textit{für } (x, y) \neq 0 \\ 0 & \textit{für } (x, y) = 0. \end{cases}$$

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

Für $y = 0$ *ist* $\frac{\partial f}{\partial x} = 0$ *und für* $x = 0$ *ist* $\frac{\partial f}{\partial y} = 0$, *so dass diese partiellen Ableitungen für alle* $(x, y) \in \mathbb{R}^2$ *existieren. Allerdings sind sie in keiner Umgbung von* $(x, y) = (0, 0)$ *beschränkt, und deshalb auch nicht stetig. Wir hatten schon im Beispiel 10.10 gesehen, dass* f *bei* $(0, 0)$ *nicht stetig und deshalb auch nicht differenzierbar ist.*

(ii)
$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{x^3-y^3}{x^2+y^2} & \textit{für } (x, y) \neq 0 \\ 0 & \textit{für } (x, y) = 0. \end{cases}$$

Offenbar gilt $|f(x, y)| \leq \frac{|x|^3}{x^2+y^2} + \frac{|y|^3}{x^2+y^2} \leq |x| + |y|$. *Also ist* f *stetig.*

$$\frac{\partial f}{\partial x} = \begin{cases} \frac{3x^2(x^2+y^2)-2x(x^3-y^3)}{(x^2+y^2)^2} = \frac{x(x^3+3xy^2+2y^3)}{(x^2+y^2)^2} & \textit{für } (x, y) \neq 0 \\ 1 & \textit{für } (x, y) = 0 \end{cases}$$

$$\frac{\partial f}{\partial y} = \begin{cases} -\frac{3y^2(x^2+y^2)+2y(x^3-y^3)}{(x^2+y^2)^2} = -\frac{y(y^3+3yx^2+2x^3)}{(x^2+y^2)^2} & \textit{für } (x, y) \neq 0 \\ -1 & \textit{für } (x, y) = 0. \end{cases}$$

Deshalb ist f *partiell differenzierbar. In Beispiel 10.14 (iii) werden wir sehen, dass* f *in* $(0, 0)$ *nicht differenzierbar ist.*

(iii) *Alle Polynome in* x *und* y *sind partiell unendlich oft stetig differenzierbar, und deshalb auch differenzierbar.*

Definition 10.13. *(Richtungsableitung) Sei $f : U \to Y$ eine Funktion von einer offenen Teilmenge U eines normierten Vektorraumes X in einem normierten Vektorraum Y. Für $x_0 \in U$ und $x \in X \setminus \{0\}$ heißt die Ableitung in $t_0 = 0$ der Funktion*

$$(-\epsilon, \epsilon) \to Y, \quad t \mapsto f(x_0 + tx),$$

wenn sie existiert, die Richtungsableitung von f in x_0 in Richtung x.

Beispiel 10.14. (i) *Sei $f : U \to Y$ in x_0 differenzierbar. Für $x \in X \setminus \{0\}$ gibt es ein Intervall $(-\epsilon, \epsilon)$ im Urbild von U unter $t \mapsto x_0 + tx$. Bei $t = 0$ sind die Abbildungen*

$$t \mapsto \begin{cases} \frac{\|f(x_0+tx)-f(x_0)-tf'(x_0)x\|}{|t|\|x\|} & \text{für } t \neq 0 \\ 0 & \text{für } t = 0 \end{cases} \quad t \mapsto \begin{cases} \frac{\|f(x_0+tx)-f(x_0)-tf'(x_0)x\|}{|t|} & \text{für } t \neq 0 \\ 0 & \text{für } t = 0 \end{cases}$$

stetig. Also existiert die Richtungsableitung und es gilt $\dfrac{df(x_0+tx)}{dt}\Big|_{t=0} = (f'(x_0))(x)$.

(ii) $f : \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x,y) \neq 0 \\ 0 & \text{für } (x,y) = 0. \end{cases}$

Für $t \neq 0$ ist dann $f(t\cos(\phi), t\sin(\phi)) = \sin(2\phi)$ und $f(t\cos(\phi), t\sin(\phi)) = 0$ für $t = 0$. Also ist f in $t = 0$ für $\phi \notin \frac{\pi}{2}\mathbb{Z}$ nicht stetig und auch nicht differenzierbar. Für $\phi \in \frac{\pi}{2}\mathbb{Z}$ existieren die Richtungsableitungen und verschwinden.

(iii) $f : \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto \begin{cases} \frac{x^3-y^3}{x^2+y^2} & \text{für } (x,y) \neq 0 \\ 0 & \text{für } (x,y) = 0. \end{cases}$

Dann gilt $f(tx) = t(\cos^3\phi - \sin^3\phi)$ für $x = (\cos\phi, \sin\phi)$ und $x_0 = 0$. Also existieren alle Richtungsableitungen und setzen sich im Punkt $(0,0)$ zu f zusammen. Weil diese Abbildung nicht linear ist, ist f im Punkt $(0,0)$ nicht differenzierbar.

Wir wollen den wichtigsten Fall von Funktionen $\mathbb{R}^n \to \mathbb{R}^m$ genauer betrachten.

Definition 10.15. *(Partielle Ableitungen in $\mathbb{R}^n$) Sei $f : U \to \mathbb{R}^m$ eine Funktion von einer offenen Teilmenge $U \subset \mathbb{R}^n$ in den $\mathbb{R}^m$. Dann sind die Komponenten $(f_1, \ldots, f_m)$ von f offenbar reelle Funktionen auf U. Die Funktion f ist in $(x_1, \ldots, x_n) \in U$ genau dann partiell differenzierbar, wenn für alle $i = 1, \ldots, n$ und $j = 1, \ldots, m$ die Funktionen*

$$x \to f_j(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)$$

bei $x = x_i$ differenzierbar sind. Die entsprechenden Ableitungen heißen partielle Ableitungen von f und werden mit $\frac{\partial f_j}{\partial x_i}(x_1, \ldots, x_n)$ bezeichnet. Wenn diese partiellen Ableitungen für alle $x \in U$ existieren, heißt f auf U partiell differenzierbar.

Definition 10.16. *(Vektorfeld, Gradient, Divergenz und Rotation) Eine Abbildung von einer offenen Menge $U \subset \mathbb{R}^n$ nach $\mathbb{R}^n$ wird Vektorfeld genannt. Das Vektorfeld*

$$\operatorname{grad} f = \nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

der partiellen Ableitungen einer partiell differenzierbaren reellen Funktion f auf einer offenen Menge $U \subset \mathbb{R}^n$ heißt Gradient von f. Wenn f ein partiell differenzierbares Vektofeld ist, dann ist die Divergenz von f folgende reelle Funktion

$$\operatorname{div} f = \nabla \cdot f = \frac{\partial f_1}{\partial x_1} + \ldots + \frac{\partial f_n}{\partial x_n}.$$

Die lineare Abbildung $\triangle : \quad f \mapsto \triangle f = \operatorname{div} \operatorname{grad} f = \frac{\partial^2 f}{\partial x_1^2} + \ldots + \frac{\partial^2 f}{\partial x_n^2}$

auf den zweimal partiell differenzierbaren reellen Funktionen heißt Laplaceoperator. Im Fall von $n = 3$ ist die Rotation eines differenzierbaren Vektorfeldes f definiert durch

$$\operatorname{rot} f = \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right).$$

Wenn die reelle Funktion f in $x_0 \in U$ differenzierbar ist, dann ist f auch in x_0 partiell differenzierbar und die partiellen Ableitungen sind die Richtungsableitungen in Richtung der kanonischen Einheitsvektoren $e_1, \ldots, e_n$ aus dem Beweis von Satz 9.37. Wegen der Linearität der Ableitung ist die Ableitung die lineare Abbildung:

$$\frac{df}{dx}(x_0) : \mathbb{R}^n \to \mathbb{R}, \quad (x_1, \ldots, x_n) \mapsto x_1 \frac{\partial f}{\partial x_1}(x_0) + \ldots + x_n \frac{\partial f}{\partial x_n}(x_0).$$

Wenn wir den $\mathbb{R}^n$ mit den Spaltenvektoren bezeichnen, können wir diese Abbildung durch das Matrixprodukt des Zeilenvektors ∇f mit dem Spaltenvektor x darstellen:

$$\frac{df}{dx}(x_0) : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto (\nabla f(x_0)) \cdot x.$$

Oder allgemeiner, für eine $\mathbb{R}^m$-wertige Funktion können wir die Ableitung $\frac{df}{dx}(x_0)$ von f an der Stelle x_0 als lineare Abbildung mit der Jacobimatrix identifizieren:

$$\frac{df}{dx}(x_0) = \begin{pmatrix} \frac{\partial f_1(x_0)}{\partial x_1} & \cdots & \frac{\partial f_1(x_0)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m(x_0)}{\partial x_1} & \cdots & \frac{\partial f_m(x_0)}{\partial x_n} \end{pmatrix} \qquad \frac{df}{dx}(x_0) : \mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto \frac{df}{dx}(x_0) \cdot x.$$

Die lineare Abbildung ist einfach die Matrixmultiplikation der Spaltenvektoren in $\mathbb{R}^n$ mit der Jacobimatrix, einer $m \times n$-Matrix. Insbesondere ist also die Richtungsableitung einer reellen Funktion f auf U an der Stelle $x_0 \in U$ in Richtung eines Vektors $x \in \mathbb{R}^n$ das Skalarprodukt des Gradienten $\nabla f(x_0)$ von f an der Stelle x_0 mit dem Vektor x:

$$\frac{d}{dt} f(x_0 + tx)|_{t=0} = x \cdot \nabla f(x_0).$$

Satz 10.9 und Satz 10.11 zeigen insbesondere, dass folgendes gilt:

Korollar 10.17. *Sei $f : U \to \mathbb{R}^m$ eine Funktion auf einer offenen Teilmenge $U \subset \mathbb{R}^n$.*

(i) *Wenn f in $x_0 \in U$ differenzierbar ist, dann existieren in x_0 alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}(x_0)$ und setzen sich zu der Jacobimatrix zusammen.*

(ii) *Wenn f auf U stetig differenzierbar ist, dann existieren auf U die partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}$ und setzen sich zusammen zu einer stetigen Funktion von U in die $m \times n$-Matrizen in $\mathbb{R}^{m\times n}$. Diese Matrizen heißen Jacobimatrizen von f.*

(iii) *Wenn umgekehrt f auf U partiell differenzierbar ist, und alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}$ auf U stetig sind, dann ist f auf U stetig differenzierbar.* **q.e.d.**

Definition 10.18. *Eine Nullstelle $x_0 \in U$ der Ableitung f' einer auf einer offenen Menge U reellen differenzierbaren Funktion f heißt kritischer Punkt.*

Satz 10.19. *Jedes relative Maximum (bzw. Minimum) einer differenzierbaren reellen Funktion auf einer offenen Menge ist ein kritischer Punkt.*

Beweis: Sei x_0 ein solches lokales Maximum (bzw. Minimum). Dann ist für alle $x \in X$ die entsprechende Abbildung $t \mapsto f(x_0 + tx)$ auf einer Umgebung von $0 \in \mathbb{R}$ differenzierbar und besitzt dort ein lokales Maximum (bzw. Minimum). Also verschwindet die entsprechende Richtungsableitung. Dann verschwindet auch $\frac{df}{dx}(x_0)$ auf allen x. **q.e.d.**

10.4 Höhere Ableitungen

Sei $f : U \to W$ eine auf einer offenen Teilmenge U eines Banachraumes V differenzierbare Funktion in den Banachraum W. Wenn f zweimal differenzierbar ist, dann ist die Ableitung stetig. Also wollen wir in diesem Abschnitt annehmen, dass f auf U stetig differenzierbar ist. Die Ableitung f' ist dann eine stetige Abbildung von U nach $\mathcal{L}(V, W)$. Die zweite Abbildung $f''(x_0)$ an einer Stelle $x_0 \in U$ ist dann (wenn sie existiert) eine lineare Abbildung von V nach $\mathcal{L}(V, W)$.

Definition 10.20. *Eine Abbildung $A : V \times V \to W$ heißt bilinear, wenn für alle $v, v', v'' \in V$ und $\lambda \in \mathbb{K}$ gilt*

$$\begin{aligned} A(v + v'', v') &= A(v, v') + A(v'', v) \quad \text{und} \quad & A(v, v' + v'') &= A(v, v') + A(v, v'') \\ A(\lambda v, v') &= \lambda A(v, v') \quad \text{und} & A(v, \lambda v') &= \lambda A(v, v'). \end{aligned}$$

Das kartesische Produkt $V \times V$ von (normierten) Vektorräumen ist wieder ein normierter Vektorraum. Die bilinearen Abbildungen von $V \times V$ nach W unterscheiden sich von den linearen Abbildungen von $V \times V$ nach W. Es gibt einen anderen Vektorraum $V \otimes V$, den wir kurz erwähnen wollen, und den man das Tensorprodukt von V mit V nennt. Die bilinearen Abbildungen von $V \times V$ nach W sind dann genau die lineare

Abbildungen von $V \otimes V$ nach W. Allerdings ist es nicht so klar wie $V \otimes V$ zu einem normierten Vektorraum gemacht wird. Für die Dimensionen dieser Vektorräume gilt

$$\dim(V \times V) = \dim(V) + \dim(V) \qquad \dim(V \otimes V) = \dim(V) \cdot \dim(V).$$

Die linearen Abbildungen von V nach $\mathcal{L}(V, W)$ lassen sich nun mit den bilinearen Abbilungen von $V \times V$ nach W identifizieren.

Lemma 10.21. *Eine Abbildung $A : V \times V \to W$ ist genau dann bilinear, wenn*

$$B : V \to \{\textit{Abbildungen } V \to W\}, \quad v \mapsto B(v), \quad B(v) : V \to W, \quad B(v)(v') = A(v, v')$$

eine lineare Abbildung von V in die linearen Abbildungen von V nach W definiert.

Der Beweis folgt aus der Definition von bilinearen Abbildungen. **q.e.d.**

Satz 10.22. *(Satz von Schwarz) Sei f eine differenzierbare Abbildung von einer offenen Teilmenge $U \subset V$ eines normierten Vektorraumes V in den normierten Vektorraum W. Wenn f im Punkt x_0 zweimal differenzierbar ist, dann ist die der zweiten Ableitung entsprechende bilineare Abbildung $f''(x_0) : V \times V \to W$ symmetrisch, d.h.*

$$(f''(x_0)u)v = (f''(x_0)v)u \quad \textit{für alle } u, v \in V.$$

Beweis: Für $t \in [0, 1]$ und kleine $u, v \in V$ sei $g(t) = f(x_0 + tu + v) - f(x_0 + tu)$. Dann ist g differenzierbar und es gilt

$$\begin{aligned} g'(t) &= f'(x_0 + tu + v)u - f'(x_0 + tu)u \\ &= ((f'(x_0 + tu + v) - f'(x_0)) - (f'(x_0 + tu) - f'(x_0)))u \end{aligned}$$

Weil f in x_0 zweimal differenzierbar ist, gibt es zu jedem $\epsilon > 0$ ein $\delta > 0$, so dass $B(x_0, 2\delta) \subset U$ und außerdem für $u, v \in B(0, \delta) \subset V$ und $t \in [0, 1]$ die Ungleichungen

$$\begin{aligned} &\|f'(x_0 + tu + v) - f'(x_0) - f''(x_0)(tu + v)\| \le \epsilon(\|u\| + \|v\|) \\ &\|f'(x_0 + tu) - f'(x_0) - f''(x_0)tu\| \le \epsilon\|u\| \end{aligned}$$

gelten. Daraus folgt für $t \in [0, 1]$ $\qquad |g'(t) - (f''(x_0)v)u\| \le \epsilon\|u\|(2\|u\| + \|v\|)$.
Die Anwendung von Korollar 10.6 auf die Funktion $t \mapsto g(t) - t(f''(x_0)v)u$ ergibt dann

$$\|g(1) - g(0) - (f''(x_0)v)u\| \le \sup_{t \in [0,1]} \|g'(t) - (f''(x_0)v)u\| \le \epsilon\|u\|(2\|u\| + \|v\|).$$

Weil $g(1) - g(0) = f(x_0 + u + v) - f(x_0 + u) - f(x_0 + v) + f(x_0)$ in u und v symmetrisch ist gilt dann auch $\|g(1) - g(0) - (f''(x_0)u)v\| \le \epsilon\|v\|(2\|v\| + \|u\|)$. Daraus folgt

$$\|(f''(x_0)v)u - (f''(x_0)u)v\| \le 2\epsilon(\|u\|^2 + \|u\| \cdot \|v\| + \|v\|^2).$$

Diese Ungleichung gilt wegen der Linearität nicht nur für $u, v \in B(0, \delta)$, sondern für $u, v \in V$. Im Grenzwert $\epsilon \to 0$ folgt $(f''(x_0)v)u = (f''(x_0)u)v$ für alle $u, v \in V$. **q.e.d.**

Zusammen mit Satz 10.11 erhalten wir

Korollar 10.23. *Sei $f : U \to \mathbb{R}^m$ eine zweimal stetig partiell differenzierbare Funktion von einer offenen Teilmenge $U \subset \mathbb{R}^n$ nach $\mathbb{R}^m$. Dann vertauschen die partiellen Ableitungen, d.h. für alle $i, j = 1, \ldots, n$ und $k = 1, \ldots, m$ gilt*

$$\partial_i \partial_j f_k = \frac{\partial^2 f_k}{\partial x_i \partial x_j} = \frac{\partial^2 f_k}{\partial x_j \partial x_i} = \partial_j \partial_i f_k \text{ und } \operatorname{rot}\operatorname{grad} f = 0 \text{ für } n = 3, m = 1.$$ **q.e.d.**

Durch mehrfaches Anwenden und differenzieren erhalten wir dann auch

Korollar 10.24. *Sei f eine n-mal differenzierbare Abbildung von einer offenen Teilmenge $U \subset V$ eines normierten Vektorraumes V in den normierten Vektorraum W. Dann ist für jedes $x_0 \in U$ die n-fache Ableitung von f eine multilineare symmetrische Abbildung von $V \times V \times \ldots \times V$ nach W. D.h. für jede Permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ und $v_1, \ldots, v_n \in V$ gilt*

$$(\ldots((f^{(n)}(x_0)v_1)v_2)\ldots)v_n = (\ldots((f^{(n)}(x_0)v_{\sigma(1)})v_{\sigma(2)})\ldots)v_{\sigma(n)}.$$ **q.e.d.**

Beispiel 10.25. *Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto f(x, y)$ mit $f(x, y) = xy\frac{x^2 - y^2}{x^2 + y^2}$ für $(x, y) \neq (0, 0)$ und $f(0, 0) = 0$. Dann ist f zweimal partiell differenzierbar.*

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

$$\frac{\partial f}{\partial y}(x, y) = x\frac{(x^2 - y^2)(x^2 + y^2) - 2y^2(x^2 + y^2) - 2y^2(x^2 - y^2)}{(x^2 + y^2)^2} = x\frac{x^4 - 4x^2y^2 - y^4}{(x^2 + y^2)^2}.$$

$$\frac{\partial^2 f}{\partial y \partial x}(0, y) = -1 \qquad\qquad \frac{\partial^2 f}{\partial x \partial y}(x, 0) = 1$$

Also existieren auf ganz $\mathbb{R}^2$ alle zweiten partiellen Ableitungen, aber es gilt

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1 \neq 1 = \frac{\partial^2 f}{\partial x \partial y}(x, 0).$$

Definition 10.26. *Sei $f : U \to W$ eine reelle Funktion auf einer offenen Teilmenge U des normierten Vektorraumes V, die bei $x_0 \in U$ zweimal differenzierbar ist. Dann definiert die zweite Ableitung eine symmetrische Bilinearform auf V:*

$$f''(x_0) : V \times V \to \mathbb{R}, \qquad\qquad (v, w) \mapsto f''(x_0)(v, w).$$

Für $V = \mathbb{R}^n$ identifizieren wir die Elemente von V wieder mit den Spaltenvektoren. Dann ist diese Bilinearform durch die sogenannte Hessematrix gegeben:

$$f''(x_0)(v, w) = w^t \cdot \frac{df}{dx}(x_0) \cdot v \qquad\qquad = \sum_{i,j=1}^{n} w_j \frac{\partial^2 f}{\partial x_j \partial x_i}(x_0) v_i.$$

Satz 10.27. *Sei $f : U \to \mathbb{R}$ eine auf einer offenen Menge zweimal differenzierbare reelle Funktion f. Dann ist die zweite Ableitung bei allen relativen Minima (Maxima) eine nicht negative (nicht positive) symmetrische Bilinearform. Gibt es umgekehrt einen kritischen Punkt $x_0 \in U$ und ein $\delta > 0$ mit*

$$(f''(x_0))(x,x) \geq \delta\|x\|^2 \quad \textit{bzw.} \quad (f''(x_0))(x,x) \leq -\delta\|x\|^2 \quad \textit{für alle } x \in X,$$

dann ist der kritische Punkt ein lokales Minimum bzw. Maximum.

Beweis: Wenn x_0 ein lokales Maximum bzw. Minimum von f ist, dann für alle $x \in X$ auch $t = 0$ von $t \mapsto f(x_0 + tx)$. Deshalb folgt die erste Aussage aus Korollar 7.16.

Umgekehrt folgt aus den obigen Ungleichungen, dass es ein $\epsilon > 0$ gibt, so dass

$$(f'(x_0 + x))(x) \geq \frac{\delta}{2}\|x\|^2 \quad \text{bzw.} \quad (f'(x_0 + x))(x) \leq -\frac{\delta}{2}\|x\|^2 \quad \text{für alle } x \in B(0,\epsilon)$$

gilt. Dann folgt auch die zweite Aussage aus Korollar 7.16. **q.e.d.**

Zum Abschluss wollen wir den Satz von Taylor auf reelle Funktionen f auf offenen konvexen Teilmengen $U \subset V$ eines normierten Vektorraumes verallgemeinern. Wenn $x_0, x \in U$ in einer solchen konvexen offenen Teilmenge U liegen, dann ist

$$g : [0,1] \to \mathbb{R}, t \mapsto f(x_0 + t(x - x_0))$$

eine reelle Funktion. Wenn f auf U n-mal differenzierbar ist, dann ist auch g n-mal differenzierbar. Wegen der Kettenregel Satz 10.4 (iii) ist die n-te Ableitung von g gleich

$$g^{(m)}(t) = (\ldots(f^{(m)}(x_0 + t(x - x_0)) \cdot (x - x_0))\ldots)(x - x_0),$$

also die n-lineare symmetrische Form zu $f^{(n)}(x_0)$ ausgewertet auf $((x - x_0), \ldots, (x - x_0)) \in V^{xm}$. Dann erhalten wir nach dem Satz von Taylor:

Satz 10.28. *(von Taylor in höheren Dimensionen) Sei $f : U \to \mathbb{R}$ eine auf einer offenen konvexen Teilmenge eines normierten Vektorraumes definierte $(m+1)$-mal differenzierbare Funktion. Dann gibt es für jedes $x, x_0 \in U$ ein $\xi \in (0,1)$, so dass gilt*

$$f(x) = \sum_{k=0}^{m} \frac{f^{(k)}(x_0)((x - x_0), \ldots, (x - x_0))}{k!} + \\ + \frac{f^{(m+1)}(x_0 + \xi(x - x_0))((x - x_0), \ldots, (x - x_0))}{(m+1)!}.$$

Hierbei bezeichnen wir mit $f^{(k)}(x_0)$ bzw. $f^{(m+1)}(x_0 + \xi(x - x_0))$ die entsprechende multilineare Abbildung von $V^{\times k}$ bzw. $V^{\times(m+1)}$ nach $\mathbb{R}$. Dabei heißt der erste Term wieder Taylorpolynom und der zweite Term Restglied. **q.e.d.**

Das Taylorpolynom und entsprechend die Taylorreihe ist auch für glatte Funktionen in einen normierten Vektorraum definiert. Eine unendlich oft differenzierbare Funktion heißt wieder reell analytisch in x_0, wenn die entsprechende Taylorreihe auf einer Umgebung von x_0 gegen die Funktion konvergiert. So definiert auf einem Banachraum die Exponentialfunktion eine analytische Funktion von $\mathcal{L}(V)$ auf sich selber.

Kapitel 11

Nichtlineare Analysis

11.1 Der Banachsche Fixpunktsatz

In diesem Kapitel bieten wir eine kurze Einführung in die nichtlineare Analysis. Der bei weitem wichtigste Satz der nichtlinearen Analysis ist der sogenannte Banachsche Fixpunktsatz. Wir werden gleich mehrere Anwendungen kennenlernen.

Banachscher Fixpunktsatz 11.1. *Sei $f : X \to X$ eine lipschitzstetige Abbildung eines vollständigen metrischen Raumes auf sich selber mit Lipschitzkonstante $L < 1$. Dann hat f genau einen Fixpunkt: $x \in X$ mit $f(x) = x$.*

Beweis: Sei $x_0 \in X$ beliebig und für alle $n \in \mathbb{N}$ x_n induktiv definiert durch $x_n = f(x_{n-1})$. Dann gilt

$$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq Ld(x_{n-1}, x_n) \leq L^n d(x_0, x_1).$$

Mit der Dreiecksungleichung folgt dann für $n \leq m$

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + \ldots + d(x_{m-1}, x_m) \\ &\leq (L^n + \ldots + L^{m-1})d(x_0, x_1) \\ &= L^n \frac{1 - L^{m-n}}{1 - L} d(x_0, x_1) \leq \frac{L^n}{1 - L} d(x_0, x_1) \end{aligned}$$

Also ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge und es existiert $x = \lim\limits_{n\to\infty} x_n$. Aus der Stetigkeit von f folgt dann $f(x) = \lim\limits_{n\to\infty} f(x_n) = \lim\limits_{n\to\infty} x_{n+1} = x$. Also ist x ein Fixpunkt. Ist $y \in X$ ein zweiter Fixpunkt, so gilt $d(x, y) = d(f(x), f(y)) \leq L \cdot d(x, y)$. Wegen $0 \leq L < 1$ folgt dann $(1 - L)d(x, y) \leq 0$ oder auch $d(x, y) = 0$. Also gilt $x = y$. **q.e.d.**

Eine Anwendung ist z.B. der Satz von Picard Lindelöf über die Existenz und Eindeutigkeit von Anfangswertproblemen von gewöhnlichen Differentialgleichungen. Eine gewöhnliche Differentialgleichung ist eine Gleichung, von der Form

$$\dot{u}(t) = f(t, u(t)) \quad \text{mit } u : I \to V \quad \text{und } f : I \times U \to V.$$

Hierbei ist $I \subset \mathbb{R}$ ein Intervall, V ein normierter Vektorraum und $U \subset V$ eine offene Teilmenge. Der Punkt bezeichnet die Ableitung nach t, die in vielen Anwendungen für die Zeit steht. Das Anfangswertproblem besteht aus der Suche nach einer differenzierbaren Funktion $u : I \to U$, die die Differentialgleichung erfüllt und an einem Punkt $t_0 \in I$ den Anfansgwert $u(t_0) = u_0 \in U$ annimmt.

Definition 11.2. *Eine Funktion f von einem metrischen Raum X in den metrischen Raum Y heißt lokal lipschitzstetig, wenn es für jedes $x_0 \in X$ eine Umgebung $U \subset X$ von x_0 gibt und eine Lipschitzkonstante $L > 0$, so dass für alle $x, x' \in U$ gilt*

$$d(f(x), f(x')) \leq Ld(x, x').$$

Satz 11.3. *(Lokale Existenz und Eindeutigkeit) Sei I ein offenes Intervall, $U \subset \mathbb{R}^n$ eine offene Teilmenge und $f : I \times U \to \mathbb{R}^n$ eine stetige Abbildung, die bezüglich der zweiten Variablen lokal lipschitzstetig ist, d.h. für jedes $(t_0, u_0) \in I \times U$ gibt es ein $\delta > 0$ und ein $L > 0$, so dass für alle $(t, u), (t, \tilde{u}) \in (t_0 - \delta, t_0 + \delta) \times B(u_0, \delta)$ gilt*

$$\|f(t, u) - f(t, \tilde{u})\| \leq L\|u - \tilde{u}\|.$$

Dann gibt es für jedes $(t_0, u_0) \in I \times U$ ein $\epsilon > 0$, so dass das Anfangswertproblem $\dot{u}(t) = f(t, u(t))$ mit $u(t_0) = u_0$ auf $(t_0 - \epsilon, t_0 + \epsilon)$ genau eine Lösung besitzt.

Beweis: Wegen der lokalen Lipschitzstetigkeit gibt es $\delta > 0$ und $L > 0$, so dass für alle $(t, u), (t, \tilde{u}) \in [t_0 - \delta, t_0 + \delta] \times \overline{B(u_0, \delta)}$ auch $\|f(t, u) - f(t, \tilde{u})\| \leq L\|u - \tilde{u}\|$ gilt. Wegen der Stetigkeit von f ist die Abbildung

$$F : u \mapsto F(u) \text{ mit } F(u)(t) = u_0 + \int_{t_0}^{t} f(s, u(s))ds$$

eine stetige Abbildung von $C([t_0 - \delta, t_0 + \delta], \overline{B(u_0, \delta)})$ nach $C([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$. Sei

$$\|f(\cdot, u_0)\|_\infty = \sup_{s \in [t_0 - \delta, t_0 + \delta]} \|f(s, u_0)\|.$$

Wenn $\epsilon \leq \delta$ und $\epsilon\,(\|f(\cdot, u_0)\|_\infty + L\delta) \leq \delta$, dann gilt für alle $u \in C([t_0 - \epsilon, t_0 + \epsilon], \overline{B(u_0, \delta)})$ und alle $t \in [t_0 - \epsilon, t_0 + \epsilon]$

$$\|F(u)(t) - u_0\| \leq \left\| \int_{t_0}^{t} (f(s, u_0) + f(s, u(s)) - f(s, u_0))ds \right\| \leq \epsilon\,(\|f(\cdot, u_0)\|_\infty + L\delta) \leq \delta$$

Also bildet F den vollständigen metrischen Raum $C([t_0 - \epsilon, t_0 + \epsilon], \overline{B(u_0, \delta)})$ auf sich selber ab. Für $u, \tilde{u} \in C([t_0 - \epsilon, t_0 + \epsilon], \overline{B(u_0, \delta)})$ und $t \in [t_0 - \epsilon, t_0 + \epsilon]$ gilt

$$\|F(u)(t) - F(\tilde{u})(t)\| \leq \int_{t_0}^{t} \|f(s, u(s) - f(s, \tilde{u}(s))\|\, ds \leq \epsilon L\|u - \tilde{u}\|_\infty.$$

Sei also ϵ kleiner als $\quad \epsilon < \min\left\{\delta, \frac{\delta}{\|f(\cdot,u_0)\|_\infty + L\delta}, \frac{1}{L}\right\} = \min\left\{\delta, \frac{\delta}{\|f(\cdot,u_0)\|_\infty + L\delta}\right\}.$

Dann definiert die Abbildung F eine lipschitzstetige Abbildung mit Lipschitzkonstante $\epsilon \cdot L < 1$ von dem vollständigen metrischen Raum $C([t_0 - \epsilon, t_0 + \epsilon], \overline{B(u_0, \delta)})$ auf sich selber. Jeder Fixpunkt ist wegen dem Hauptsatz der Differential– und Integralrechnung stetig differenzierbar und es gilt $\dot{u}(t) = f(t, u)$ für alle $t \in (t_0 - \epsilon, t_0 + \epsilon)$ mit $u(t_0) = u_0$. Also löst u dieses Anfangswertproblem auf $(t_0 - \epsilon, t_0 + \epsilon)$. Wenn u umgekehrt auf $(t_0 - \epsilon, t_0 + \epsilon)$ dieses Anfangswertproblem löst, dann ist die Ableitung von $F(u) - u$ gleich Null, und beide Funktionen $F(u)$ und u sind bei $t = t_0$ gleich u_0. Also stimmen beide Funktionen überein und jede Lösung des obigen Anfangswertproblems ist ein Fixpunkt von F. Also folgt die Existenz und Eindeutigkeit dieses Anfangswertproblems auf $(t_0 - \epsilon, t_0 + \epsilon)$ aus dem Banachschen Fixpunktsatz. **q.e.d.**

Satz 11.4. *(Globale Existenz und Eindeutigkeit) Sei $O \subset \mathbb{R} \times \mathbb{R}^n$ eine offene Teilmenge und $f : O \to \mathbb{R}^n$ eine stetige Abbildung, die wie bei der lokalen Existenz und Eindeutigkeit lokal lipschitzstetig ist. Dann gibt es für jedes $(t_0, u_0) \in O$ genau ein maximales Intervall $(a, b) \subset \mathbb{R}$, das t_0 enthält, und auf dem das Anfangswertproblem*

$$\dot{u}(t) = f(t, u) \quad \textit{mit} \quad u(t_0) = u_0$$

genau eine Lösung u enthält. Das Intervall ist in dem Sinne maximal, dass an beiden Rändern, also bei a und b, eine der folgenden Bedingungen erfüllt ist:

(i) *$a = -\infty$ (bzw. $b = \infty$).*

(ii) *$t \mapsto \|f(t, u(t))\|$ ist für alle $\epsilon > 0$ auf $(a, a + \epsilon)$ (bzw. $(b - \epsilon, b)$) unbeschränkt.*

(iii) *Die Lösung u läßt sich stetig auf $[a, b)$ (bzw. $(a, b]$) fortsetzen, der Graph der Fortsetzung liegt aber nicht in O, d.h. $\lim_{t \downarrow a}(t, u(t)) \notin O$ (bzw. $\lim_{t \uparrow b}(t, u(t)) \notin O$).*

Beweis: Zunächst bemerken wir, dass für jedes Intervall (a, b), das t_0 enthält, und auf dem das Anfangswertproblem

$$\dot{u}(t) = f(t, u(t)) \text{ mit } u(t_0) = u_0$$

eine Lösung $\tilde{u}$ besitzt, so dass sich $\tilde{u}$ auf $[a, b)$ oder $(a, b]$ stetig fortsetzen läßt, und der Graph der Fortsetzung in O liegt, das neue Anfangswertproblem

$$\dot{u}(t) = f(t, u(t)) \text{ mit } u(a) = \lim_{t \to a+} \tilde{u}(t) \text{ bzw. } u(b) = \lim_{t \to b-} \tilde{u}(t)$$

wegen dem vorangehenden Satz eine Lösung in einer Umgebung von a bzw. b besitzt, die dann auf $[a, a + \epsilon)$ bzw. $(b - \epsilon, b]$ mit $\tilde{u}$ übereinstimmt. Also existiert ein maximales Intervall (a, b), auf dem das Anfangswertproblem eine eindeutige Lösung besitzt. Wenn am linken bzw. rechten Rand die Bedingungen (i) und (ii) nicht erfüllt sind, dann ist die Ableitung der Lösung auf einer offenen Menge $(a, a+\epsilon)$ bzw. $(b-\epsilon, b)$ beschränkt und deshalb ist die Lösung dort lipschitzstetig. Dann konvergiert für jede Folge $(t_n)_{n\in\mathbb{N}}$, die gegen a bzw. b konvergiert auch die Folge $((t_n, u(t_n))_{n\in\mathbb{N}}$ in $\mathbb{R}\times\mathbb{R}^n$. Der Grenzwert kann dann aber nicht in O liegen, weil sonst die Lösung eine Fortsetzung auf eine Umgebung von a bzw. b hätte. **q.e.d.**

Bemerkung 11.5. *Wenn (ii) erfüllt ist, kann $t \mapsto f(t, u(t))$ nicht stetig auf $[a, a+\epsilon)$ bzw. $(b-\epsilon, b]$ fortgesetzt werden. Also können u und f nicht so stetig auf größere Definitionsbereiche fortgesetzt werden, dass a (bzw. b) im Definitionsbereich von u und $(a, u(a))$ (bzw. $(b, u(b))$) im Definitionsbereich von f liegt.*

11.2 Das Lösen von nichtlinearen Gleichungen

Die Lösungen der Gleichungen von der Form

$$Ax = y, \quad A \in \mathcal{L}(V, W), \quad x \in V \text{ und } \quad y \in W$$

in einem (endlichdimensionalen) Vektorraum V sind in der linearen Algebra untersucht worden. Wenn A invertierbar ist, dann ist $x = A^{-1}y$ die eindeutige Lösung. In diesem Abschnitt nutzen wir das Verständnis dieser Gleichungen für Gleichungen von der Form

$$f(x) = y, \quad f : V \to W, \quad x \in V \text{ und } \quad y \in W$$

mit nichtlinearen Abbildungen f. Dabei nehmen wir an, dass f differenzierbar ist, und durch lineare Abbildungen angenähert werden kann. Ausgangspunkt ist die Beobachtung, dass kleine Störungen von invertierbaren linearen Abbildungen invertierbar sind.

Lemma 11.6. *Seien V und W Banachräume und A ein invertierbares Element von $\mathcal{L}(V; W)$. D.h. es gibt ein Element $A^{-1} \in \mathcal{L}(W, V)$ mit $AA^{-1} = \mathbb{1}_W$ und $A^{-1}A = \mathbb{1}_V$. Dann sind alle Elemente des folgenden Balles um A invertierbar:*

$$B \in B\left(A, \frac{1}{\|A^{-1}\|}\right) \subset \mathcal{L}(V, W) \quad \mathit{mit} \quad \|B^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2\|A - B\|}{1 - \|A^{-1}\|\|A - B\|}$$

Beweis: Offenbar ist $B = A - (A - B) = A(\mathbb{1}_V - A^{-1}(A - B))$. Wegen Satz 9.60 gilt $\|A^{-1}(A - B)\| \leq \|A^{-1}\| \cdot \|A - B\| < 1$ für $B \in B\left(A, \frac{1}{\|A^{-1}\|}\right)$. Dann folgt aus der Neumannschen Reihe, dass $\mathbb{1}_V - A^{-1}(A - B)$ invertierbar ist in $\mathcal{L}(V)$ und der inverse Operator beschränkt ist durch $\frac{1}{1-\|A^{-1}\|\cdot\|A-B\|}$. Also ist auch B invertierbar und es gilt

$$B^{-1} = \left(\mathbb{1}_V - A^{-1}(A - B)\right)^{-1} A^{-1} \quad \text{mit} \quad \|B^{-1}\| \leq \frac{\|A^{-1}\|}{\mathbb{1}_V - \|A^{-1}\|\|A - B\|}.$$

Für die Differenz $B^{-1} - A^{-1}$ gilt dann

$$\begin{aligned} B^{-1} - A^{-1} &= ((\mathbb{1}_V - A^{-1}(A-B))^{-1} - \mathbb{1})A^{-1} \\ &= A^{-1}(A-B)(\mathbb{1}_V - A^{-1}(A-B))^{-1}A^{-1} \\ &= (\mathbb{1}_V - A^{-1}(A-B))^{-1}A^{-1}(A-B)A^{-1} \quad \text{und deshalb} \end{aligned}$$

$$\|B^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2\|A-B\|}{1 - \|A^{-1}\| \cdot \|A-B\|}.$$ **q.e.d.**

Damit bilden die invertierbaren Elemente von $\mathcal{L}(V, W)$ eine offene Teilmenge.

Korollar 11.7. *Seien V und W Banachräume und A ein invertierbares Element von $\mathcal{L}(V, W)$. Dann ist die Abbildung*

$$B\left(A, \frac{1}{\|A^{-1}\|}\right) \to \mathcal{L}(W, V), \quad B \mapsto B^{-1}$$

eine analytische Abbildung. Also insbesondere unendlich oft stetig differenzierbar.

Beweis: Aus Lemma 11.6 und der Neumannschen Reihe folgt für alle $B \in \mathcal{L}(V, W)$

$$(A + tB)^{-1} = A^{-1}\sum_{n=0}^{\infty} t^n(-BA^{-1})^n = \left(\sum_{n=0}^{\infty} t^n(-A^{-1}B)^n\right)A^{-1} \text{ für } t \in B\left(0, \frac{\|B\|}{\|A^{-1}\|}\right).$$

Insbesondere ist diese Abbildung analytisch mit der Ableitung $B \mapsto -A^{-1}BA^{-1}$ an der Stelle A, und damit genauso oft differenzierbar, wie $A \mapsto A^{-1}$. **q.e.d.**

Satz 11.8. *(Satz über die inverse Funktion) Seien V, W Banachräume, $f : U \to W$ eine stetig differenzierbare Abbildung von einer offenen Teilmenge $U \subset V$ nach W. Wenn $f'(x_0)$ bei $x_0 \in U$ invertierbar ist in $\mathcal{L}(V, W)$, dann gibt es offene Umgebungen $U' \subset U$ und $O \subset W$ von x_0 bzw. $f(x_0)$, so dass die Einschränkung $f : U' \to O$ bijektiv ist mit stetig differenzierbarer Umkehrabbildung $f^{-1} : O \to U'$ mit Ableitung $y \to (f'(f^{-1}(y)))^{-1}$.*

Beweis Indem wir zu der Abbildung $x \to (f'(x_0))^{-1} \circ (f(x_0 + x) - f(x_0))$ übergehen, können wir annehmen, dass $W = V$ ist und x_0 und $f(x_0)$ gleich Null sind und $f'(x_0) = \mathbb{1}_V$ ist. Weil f stetig differenzierbar ist, gibt es ein $\delta > 0$, so dass $\|f'(x) - \mathbb{1}_V\| \leq \frac{1}{2}$ für $x \in \overline{B(0, \delta)} \subset U$ gilt. Wegen dem Schrankensatz ist für jedes $y \in V$ die Abbildung

$$F_y : x \mapsto y + x - f(x)$$

eine lipschitzstetige Abbildung von $\overline{B(0, \delta)}$ nach $\overline{B(y, \frac{\delta}{2})}$ mit Lipschitzkonstante $\frac{1}{2}$.Außerdem gilt auch wegen dem Schrankensatz für alle $x \in \overline{B(0, \delta)}$

$$\|F_y(x) - y\| = \|f(x) - x\| = \|f(x) - x - (f(0) - 0)\| \leq \frac{\delta}{2}.$$

Wenn y in $\overline{B(0,\frac{\delta}{2})}$ liegt, dann liegt $\overline{B(y,\frac{\delta}{2})}$ in $\overline{B(0,\delta)}$. Also definiert F_y dann eine Abbildung von $\overline{B(0,\delta)}$ auf sich selbst. Aus dem Banachschen Fixpunktsatz folgt, dass für jedes $y \in \overline{B(y,\frac{\delta}{2})}$ die Abbildung F_y auf $\overline{B(0,\delta)}$ genau einen Fixpunkt hat und der Fixpunkt in $\overline{B(y,\frac{\delta}{2})}$ liegt. Weil aber x genau dann ein Fixpunkt von F_y ist, wenn $f(x) = y$ ist, gibt es für alle $y \in \overline{B(y,\frac{\delta}{2})}$ auf $\overline{B(0,\delta)}$ genau eine Lösung der Gleichung $f(x) = y$. Sei also

$$O = B\left(0, \frac{\delta}{2}\right) \qquad \text{und} \qquad U' = \left\{x \in B\left(0,\delta\right) | f(x) \in B\left(0, \frac{\delta}{2}\right)\right\}.$$

Dann ist O offen und U' als Urbild einer offenen Menge unter einer stetigen Abbildung auch offen und die Abbildung $f : U' \to O$ bijektiv. Weil aber die Abbildung F_0 auf $\overline{B(0,\delta)}$ lipschitzstetig ist mit Lipschitzkonstante $\frac{1}{2}$, gilt für alle $x, x' \in \overline{B(0,\delta)}$ auch

$$\|x - x'\| = \|f(x) - f(x') - F_0(x) - F_0(x'))\| \leq \|f(x) - f(x')\| + \frac{1}{2}\|x - x'\|$$
$$\text{oder auch} \qquad \|x - x'\| \leq 2\|f(x) - f(x')\|.$$

Also ist f^{-1} sogar lipschitzstetig mit Lipschitzkonstante 2. Wegen der Neumann'schen Reihe ist für alle $x \in \overline{B(0,\delta)}$ dann $f'(x)$ in $\mathcal{L}(V)$ invertierbar, und wegen dem Lemma 11.6 die Abbildung

$$\overline{B(0,\delta)} \to \mathcal{L}(V), \quad x \mapsto (f'(x))^{-1}$$

stetig. Die Komposition von $f^{-1} : O \to U'$ mit dieser Abbildung ist dann auch stetig. Für $x, x_0 \in O$ folgt aus $\|f(x) - f(x_0) - f'(x_0)(x - x_0)\| < \epsilon\|x - x_0\|$

$$\begin{aligned}\|x - x_0 - (f'(x_0))^{-1}(f(x) - f(x_0))\| \leq \\ &\|(f'(x_0))^{-1}\|\|(f(x) - f(x_0) - f'(x_0)(x - x_0)\| \\ &\leq \|(f'(x_0))^{-1}\|\epsilon\|x - x_0\| \leq \|(f'(x_0))^{-1}\| \cdot 2\epsilon\|f(x) - f(x_0)\|.\end{aligned}$$

Also ist die Komposition von $f^{-1} : O \to U'$ mit $x \mapsto (f'(x))^{-1}$ die Ableitung von f^{-1}. Dann ist also f^{-1} auch stetig differenzierbar. **q.e.d.**

Beispiel 11.9. *Die Voraussetzung der stetigen Differenzierbarkeit kann nicht abgeschwächt werden zu einfacher Differenzierbarkeit. Die Funktion*

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) = \begin{cases} x + x^2 \sin(\frac{1}{x}) & \textit{für } x \neq 0 \\ 0 & \textit{für } x = 0 \end{cases}$$

ist differenzierbar und bei 0 *gilt* $f'(0) = 1$. *Aber* f *ist in keiner Umgebung der* 0 *injektiv.*

Korollar 11.10. *Die Umkehrabbildung einer bijektiven n-mal stetig differenzierbaren Abbildung ist bei allen Punkten x mit invertierbarer Ableitung $f'(x)$ auch n-mal stetig differenzierbar.* **q.e.d.**

Korollar 11.11. *(Satz über die implizite Funktion) Seien V und W Banachräume, U eine offene Teilmenge von $V \times W$ und $f : U \to V$ eine stetig differenzierbare Funktion. Wenn die partielle Ableitung $\frac{\partial f}{\partial v}$ in $(v_0, w_0) \in U$ als Element von $\mathcal{L}(V)$ invertierbar ist, dann gibt es offene Umgebungen O von $f(v_0, w_0)$ in V, O' von w_0 in W und U' von (v_0, w_0) in U und eine stetig differenzierbare Funktion $g : O \times O' \to V$, so dass für alle $(u, w) \in O \times O'$ gilt $f(g(u, w), w) = u$. Außerdem sind für alle $u \in O$ alle Lösungen $(v, w) \in U'$ der Gleichungen $f(v, w) = u$ im Graphen der Abbildung $O' \to V, \quad w \mapsto g(u, w)$ enthalten.*

Beweis: Die Ableitung der Abbildung $F : U \to V \times W, \quad (v, w) \mapsto (f(v, w), w)$ ist gegeben durch

$$F'(v, w) : V \times W \to V \times W, \qquad (x, y) \mapsto \left(\frac{\partial f(v, w)}{\partial v} x + \frac{\partial f(v, w)}{\partial w} y, y \right)$$

Wenn $\frac{\partial f(v,w)}{\partial v}$ invertierbar ist, dann ist der inverse Operator gegeben durch

$$(F'(v, w))^{-1} : V \times W \to V \times W, \quad (x, y) \mapsto \left(\left(\frac{(\partial f(v, w)}{\partial v} \right)^{-1} \left(x - \frac{\partial f(v, w)}{\partial w} y \right), y \right)$$

Also erfüllt sie die Voraussetzungen des Satzes über die inverse Funktion. Deshalb gibt es Umgebungen O von $f(v_0, w_0)$ in V, O' von w_0 in W und U' von (v_0, w_0) in U, so dass die Abbildung $U \to O \times O', \quad (v, w) \mapsto (f(f(v, w), w)$ bijektiv ist und eine Umkehrabbildung besitzt. Diese Umfehrabbildung muss aber wegen der Gestalt von F von der Form $O \times O' \to U', \quad (u, w) \mapsto (g(u, w), w)$ sein, mit einer stetig differenzierbaren Funktion g. Insbesondere sind für alle $u \in O$ alle Lösungen $(v, w) \in U'$ von $f(v, w) = u$ im Graphen von $O \to V, \quad w \mapsto g(u, w)$ enthalten. **q.e.d.**

Beispiel 11.12. *(i) Höhenlinien: Sei $f : \mathbb{R}^2 \to \mathbb{R}$ eine stetig differenzierbare Funktion, die z.B. in Abhängigkeit von Längen- und Breitengraden die Höhe über dem Meeresspiegel beschreibt. Wenn die partielle Ableitung $\frac{\partial f}{\partial x}$ (oder eine andere partielle Ableitung) in einem Punkt $(x_0, y_0) \in \mathbb{R}$ nicht verschwindet, dann gibt es eine stetig differenzierbare Funktion*

$$g : (f(x_0, y_0) - \epsilon, f(x_0, y_0) + \epsilon) \times (y_0 - \epsilon, y_0 + \epsilon) \to \mathbb{R},$$

so dass für alle $z \in (f(x_0, y_0) - \epsilon, f(x_0, y_0) + \epsilon)$ die Höhenlinien zur Höhe z von f gerade durch die Graphen der Funktionen

$$(y_0 - \epsilon, y_0 + \epsilon) \to \mathbb{R}, \quad y \mapsto g(z, y)$$

beschrieben werden mit Für festes z zeigen also die Richtungen

$$\left(\frac{\partial g(z,y)}{\partial y}, 1\right)$$

in Richtung der Höhenlinien und stehen senkrecht auf dem Gradienten von f.
(ii) Hyperflächen: *Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine stetig differenzierbare Funktion, deren Ableitung in eimem Punkt x_0 nicht verschwindet. Dann verschwindet auch mindestens eine partielle Ableitung nicht. Nach einer geegneten Permutation der Varibaeln, können wir $\frac{\partial f}{\partial x_1}(x_0) \neq 0$ annehmen. Sei $y_0 \in \mathbb{R}^{n-1}$ der Vektor der letzten $n-1$ Koordinaten von x_0 Dann lassen sich lokal die Niveaumengen: $\{(x \in \mathbb{R} \mid f(x) = z\}$ mit $z \in (f(x_0) - \epsilon, f(x_0 + \epsilon))$ durch den Graphen einer Funktion $g : (f(x_0 - \epsilon, f(x_0 + \epsilon \times B(y_0, \delta) \to \mathbb{R}$ beschreiben als $(g(z,y), y) \in \mathbb{R} \times B(y_0, \delta)$ mit $(z,y) \in B(f(x_0), \epsilon) \times B(y_0, \delta)$. Lokal werden die Niveauflächen also von $y \in \mathbb{R}^{n-1}$ parametrisiert. Für alle (z,y) in dieser Umgebung von $(f(x_0), y_0)$ ist das Bild der partiellen Ableitungen*

$$\frac{\partial g}{\partial y}(z,y) \times \mathbb{1}_Y \in \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}) \times \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}^{n-1}) \simeq \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R} \times \mathbb{R}^{n-1})$$

$$\text{dann der Kern von } f'(g(z,y), y) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}).$$

Dieser Kern wird auch Tangentialraum an die Niveauflächen genannt.

Definition 11.13. *Eine unendlich oft (stetig) differenzierbare bijektive Abbildung mit unendlich oft (stetig) differenzierbarer Umkehrabbildung heißt Diffeomorphismus.*

Beispiel 11.14. *Polarkoordinaten Die Abbildung*

$$\mathbb{R}^+ \times \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{R}^2 \setminus \{0\}, \quad (r, \phi) \mapsto (r\cos\phi, \sin\phi)$$

heißt Polarkoordinaten von $\mathbb{R}^2$. Offenbar ist diese Abbildung unendlich oft stetig differenzierbar. Die Umkehrabbildung ist dann gegeben durch

$$(x,y) \mapsto (r,\phi) \qquad \text{mit} \qquad r = \sqrt{x^2+y^2}$$

$$\text{und} \qquad \phi = \begin{cases} \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) & \text{für } y \geq 0 \\ 2\pi - \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) & \text{für } y < 0 \end{cases}$$

Also ist diese Abbildung ein Diffeomorphismus von $\mathbb{R}^+ \times \mathbb{R}/2\pi\mathbb{Z}$ nach $\mathbb{R}^2 \setminus \{0\}$. Hier beschreibt $\mathbb{R}/2\pi\mathbb{Z}$ den Raum aller Äquivalenzklassen von $\mathbb{R}$, wobei

$$x \sim y \Leftrightarrow x - y \in 2\pi\mathbb{Z}.$$

Dieser Raum ist offenbar lokal diffeomorph zu $\mathbb{R}$, weil in jedem Intervall dessen Unterlänge kleiner ist als 2π, verschiedene Elemente verschiedene Äquivalenzklassen repräsentieren. Deshalb sind die Einschränkungen der Abbildung $\mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z}$ auf beliebige offene Intervalle mit Längen nich t gößer als 2π, die jedes Element auf die entsprechende Äquivalenzklasse abbilden, Diffeomorphismen.

11.3 Lagrangemultiplikatoren

Ziel dieses Abschnittes ist es ein Verfahren vorzustellen, mit dem man die lokalen Extremwerte von Funktionen auf solchen Teilmengen eines Banachraumes bestimmen kann, die die Nullstellen von endlich vielen reellen diffenzierbaren Funktionen bilden. Wir sprechen dann von Zwangsbedingungen, wegen denen nur die Punkte in diesen Nullstellenmengen in Betracht kommen. Diese Situation ist recht allgemein und kommt in sehr vielen Anwendungen gerade der Wirtschaftswissenschaften vor. Dieses Verfahren ist die Grundlage für die nichtlineare Optimierung, in der man nach Extremwerten auf Teilmengen eines Banachraumes sucht. Darauf aufbauend wird in der konvexen Analysis nach Bedingungen gesucht, die die Existenz und Eindeutigkeit von solchen Extremwertproblemen garantiert.

Definition 11.15. *Sei $U \subset X$ eine offene Teilmenge eines Banachraumes X und $g = (g_1, \ldots, g_m) : U \to \mathbb{R}^m$ ein Funktion von U nach $\mathbb{R}^m$. Dann definiert für jedes $x_0 \in U$*

$$A = x \in U \mid g(x) = g(x_0)$$

eine abgeschlossene Menge A von U auf der die reellen Funktionen $g_1, \ldots, g_m(x_0)$ konstant sind. Wir nennen solche Mengen Niveaumengen zu den Zwangsbedingungen $g_1, \ldots, g_n$,

Typischerweise werden die Zwangsbedingungen glatte Funktionen sein. Aber selbst in diesem Fall sind die Niveaumengen nicht immer glatt. Sie können Singularitäten besitzen. Wenn allerdings der Rang der linearen Abbildung

$$g'(x_0) = (g_1'(x_0), \ldots, g_m'(x_0)) : X \to \mathbb{R}^m$$

auf einer offenen Menge konstant ist, dann wissen wir wegen dem Satz der implizierten Funktion, dass dort die Niveaumengen glatt sind. In der folgenden Diskussion wollen wir uns zunächst auf solche Punkte der Niveaumenge beschränken, an denen der Rang von g' gleich m ist. Wenn g stetig differenzierbar ist, und der Rang der Jacobimatrix $g'(x_0)$ an einem Punkt $x_0 \in U$ gleich m ist, dann gibt es auch eine Umgebung von x_0, auf dem der Rang der Jacobimatrix gleich m ist. Um in diesem Fall den Satz der impliziten Funktion anzuwenden müssen wir X so in ein kartesisches Produkt von $\mathbb{R}^m$ mit einem normierten Vektorraum zerlegen, dass die erste partielle Ableitung als Element von $\mathcal{L}(\mathbb{R}^m)$ invertierbar ist. In diesem Fall folgt aus dem Satz der implizierten Funktion, dass wir die Niveaumengen lokal in einer Umgebubg von x_0 durch offene Teilmengen eines normierten Vektorraumes parametrisieren können. Wir nennen alle Funktionen f auf A, die Einschränkungen von glatten Funktionen f auf offenen Umgebungen von A sind, glatte Funktionen auf A.

Beispiel 11.16. **(i)** $$g : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto g(x, y) = x^2 + y^2.$$

Der Gradient $\nabla g(x,y) = (2x, 2y)$ von g verschwindet nur bei $(x,y) = 0$. Also sind alle Niveaumengen $g(x,g) = g_0$ mit $g_0 \neq g(0,0) = 0$ glatte 1-dimensionale Teilmengen von $\mathbb{R}^2$. Es sind jeweils die Kreise mit Radius $\sqrt{g_0}$ um den Nullpunkt. Für $g_0 = 0$ besteht die Niveaumenge allerdings nur aus dem Nullpunkt. Er ist eine Singularität der Niveaumenge.

(ii) $$g : \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto g(x,y) = x^2 - y^2.$$
Der Gradient $\nabla g(x,y) = (2x, -2y)$ verschwindet wieder nur bei $(x,y) = (0,0)$. Also sind alle Niveaumengen $g(x,y) = g_0$ mit $g_0 \neq g(0,0) = 0$ glatte eindimensionale Teilmengen von $\mathbb{R}^2$. Es sind jewils zwei Hyperebenen. Die Teilmenge $g(x,y) = x^2 - y^2 = (x-y)(x+y) = 0$ besteht aber aus zwei Geraden $y = x$ und $y = -x$, die sich im Nullpunkt schneiden. Diese Niveaumenge hat also im Nullpunkt eine Singularität, weil sich dort zwei glatte Teilmengen schneiden. Man spricht von einem Doppelpunkt.

(iii) $$g : \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto g(x,y) = y^2 - x^3.$$
Der Gradient $\nabla g(x,y) = (-3x^2, 2y)$ verschwindet wieder nur im Nullpunkt. Also sind wieder alle Niveaumengen $g(x,y) = g_0$ mit $g_0 \neq g(0,0) = 0$ glatte eindimensionale Teilmengen von $\mathbb{R}^2$. Die Niveaumenge $g(x,y) = y^2 - x^3 = 0$ besteht aus zwei Lösungen $y = \pm\sqrt{x^3}$ mit $x \geq 0$, die sich bei $(x,y) = 0$ einer gemeinsamen Halbgeraden parallel zu der x-Achse annähern. Man nennt deshalb die Singularität im Nullpunkt eine Spitze.

Satz 11.17. *(Kritische Punkte auf Niveaumengen) Sei $U \subset X$ eine offene Teilmenge eines Banachraumes und $g = (g_1, \ldots, g_m) : U \to \mathbb{R}^m$ eine stetig differenzierbare Funktionen auf U. Für ein $x_0 \in U$ sei A die entsprechende Niveaumenge $A = \{x \in U \mid g(x) = g(x_0)\}$. Sei f eine glatte Funktion auf einer Umgebung von x_0 in X. Wenn die Niveaumenge A bei x_0 keine Singularität hat, also die Ableitungen $g_1'(x_0), \ldots, g_m'(x_0)$ linear unabhängig sind in $\mathcal{L}(X, \mathbb{R})$, dann ist x_0 genau dann ein kritischer Punkt von der Einschränkung $f|_A$ von f auf die Niveaumenge A, wenn es Zahlen $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ gibt, so dass in $\mathcal{L}(X, \mathbb{R})$*

$$f'(x_0) = \lambda_1 g_1'(x_0) + \ldots + \lambda_m g_m'(x_0)$$

gilt. Die reellen Zahlen $\lambda_1, \ldots, \lambda_m$ heißen Lagrangemultiplikatoren.

Beweis: Weil $g_1'(x_0), \ldots, g_m'(x_0)$ in $\mathcal{L}(X, \mathbb{R})$ linear unabhängig sind, gibt es einen m-dimensionalen Unterraum $Y \subset X$ von X, so dass die lineare Abbildung

$$g'(x_0) = (g_1'(x_0), \ldots, g_m'(x_0)) : Y \to \mathbb{R}^m$$

eine bijektive lineare Abbildung ist, und damit wegen Satz 9.56 eine stetige lineare Abbildung mit stetiger linearer Umkehrabbildung von Y nach $\mathbb{R}^m$ ist. Sei $Z \subset X$ der Unterraum

$$Z = \{x \in X \mid g'(x_0)x = 0\}.$$

Er ist das Urbild von $\{0\} \subset \mathbb{R}$ unter $g'(x_0)$ und damit abgeschlossen. Offenbar besitzt Y eine Basis von Elementen $x_1, \ldots, x_m$, so dass

$$g'_j(x_0)x_i = \delta_{ij} = \begin{cases} 0 \text{ für } j \neq i \\ 1 \text{ für } j = 1 \end{cases} \qquad \text{mit } i, j \in \{1, \ldots, m\}$$

gilt. Dann sind die Abbildungen

$$X \to Y, \quad x \mapsto (g'_1(x_0)x)x_1 + \ldots + (g'_m(x_0)x)x_m$$

und

$$X \to Z, \quad x \mapsto x - (g'_1(x_0)x)x_1 - \ldots - (g'_m(x_0)x)x_m$$

stetige lineare Abbildungen. Die Verkettung von der kannonischen Einbettung $Y \hookrightarrow X$ mit der ersten ergibt offenbar die identische Abbildung von Y. Deshalb setzen sich beide Abbildungen zu der Umkehrabbildung von

$$Y \times Z \to X, \quad (y, z) \mapsto y + z$$

zusammen. Also sind die normierten Vektorräume X und $Y \times Z$ isomorph mit äquivalenten Normen. Wegen dem Satz der implizierten Funktion gibt es dann offene Umgebungen O von $(g_1(x_0), \ldots, g_m(x_0))$ in $\mathbb{R}^m$, O' von z_0 in Z und U' von x_0 in X, wobei $x_0 = (y_0, z_0) \in Y \times Z$ entspricht, und eine stetig differenzierbare Abbildung $h : O \times O' \to X$, so dass für alle $u \in O$ alle Lösungen der Gleichung $g(x) = u$ in U' zum Graphen der Funktion $z \mapsto h(u, z)$ gehören, also von der Form $h(u, z) + z$ sind mit $z \in O'$. Wir haben den Unterraum $Z \subset X$ gerade so definiert, dass die partielle Ableitung $\frac{\partial g}{\partial z}$ an der Stelle $x_0 = y_0 + z_0$ vreschwindet. Deshalb verschwindet auch die partielle Ableitung $\frac{\partial h}{\partial z}(g(x_0), z_0)$. Also ist das Bild von Z unter der Ableitung von $z \mapsto h(g(x_0), z) + z$ an der Stelle $z = z_0$ gleich Z. Dann ist x_0 genau dann ein kritischer Punkt, wenn $f'(x_0)$ auf Z verschwindet. Das ist wegen der Definition von $x_1, \ldots, x_m$ äquivalent zu

$$\begin{aligned} f'(x_0)x &= f'(x_0)\left((g'_1(x_0)x)x_1 + \ldots + (g'_m(x_0)x)x_m\right) \qquad \text{für alle } x \in X \\ &= (f'(x_0)x_1)g'_1(x_0)x + \ldots + (f'(x_0)x_m)g'_m(x_0)x \\ &= \lambda_1 g'_1(x_0)x + \ldots + \lambda_m g'_m(x_0)x \quad \text{ mit} \\ \lambda_1 &= f'(x_0)x_1, \quad \ldots, \quad \lambda_m = f'(x_0)x_m. \end{aligned}$$

Daraus folgt sofort $\qquad f'(x_0) = \lambda_1 g'_1(x_0) + \ldots + \lambda_m g'_m(x_0).$ **q.e.d.**

Die Singularitäten von den Niveaumengen können allerdings auch lokale Extemwerte sein. Deshalb muss man im Allgemeinen erst alle Singularitäten von den Niveaumengen bestimmen und dann noch alle glatten kritischen Punkte, um alle möglichen lokalen Extremwerte von der Einschränkung von f auf die Niveaumenge zu bestimmen.

Korollar 11.18. *Sei $g = (g_1, \dots, g_m) : U \to \mathbb{R}^m$ eine stetig differenzierbare Funktion auf einer offenen Teilmenge eines Banachraumes X. Wenn bei allen $x \in U'$ aus einer offenen Teilmenge U' von U die Ableitungen $g'_1 \dots, g'_{m-1}$ linear unabhängig sind und die Ableitung von g_m linear von $g_1, \dots, g_{m-1}$ abhängt, dann ist für alle $x_0 \in U'$ g_m loakl auf folgender Niveaumenge A' konstant:*

$$A' = \{x \in U' \mid g_1(x) = g_1(x_0), \dots, g_{m-1}(x) = g_{m-1}(x_0)\}.$$

Beweis: Mit der analogen Anwendung des Satzes der impliziten Funktion im vorangehenden Satz auf die Nivequmenge A' folgt, dass A' sich auf U' lokal mit stetig differenzierbaren Abbildungen bijektiv auf offene Teilmengen eines Banachraumes abbilden läßt, also glatt ist, und dass alle Punkte dieser Niveaumenge kritische Punkte von g_m sind. Dann ist g_m lokal konstant. **q.e.d.**

Durch mehrfaches Anwenden und Umordnen der Funktionen $g_1, \dots, g_m$ können wir für jede nicht konstante stetig differenzierbare Funktion $g = (g_1, \dots, g_m) : U \to \mathbb{R}^m$ erreichen, dass es einen Punkt $x_1 \in U$ und ein $1 \leq l \leq m$ gibt, so dass $g'_1(x_1), \dots, g'_l(x_1)$ linear unabhängig sind, aber $g'_{l+1}, \dots, g'_m$ an allen Punkten von U linear von $g'_1, \dots, g'_l$ abhängen. Weil g' stetig ist folgt aus Lemma 11.6, dass die Menge aller x_1, bei denen $g'_1, \dots, g'_l$ linear unabhängig sind, sogar eine offene Teilmenge von U ist. Danach bestimmen wir für ein gegebenes $x_0 \in U$ alle Singularitäten der entsprechenden Niveaumenge

$$\{x \in U \mid g_1(x) = g_1(x_0), \dots, g(x) = g_l(x_0)\}.$$

Zuletzt bestimmen wir alle kritischen glatten Punkte von einer differenzierbaren Funktion f auf dieser Niveaumenge. Dadurch haben wir alle möglichen lokalen Extremwerte bestimmt.

Kapitel 12

Das Lebesgueintegral auf dem $\mathbb{R}^d$

12.1 Treppenfunktionen

Zunächst führen wir die Klasse der Mengen von Quader im $\mathbb{R}^d$ ein.

Definition 12.1. *Ein Quader ist ein d-faches kartesisches Produkt von Intervallen*

$$Q = I_1 \times \ldots \times I_d = \{x \in \mathbb{R}^d \mid x_1 \in I_1, \ldots, x_d \in I_d\} \subset \mathbb{R}^d,$$

wobei $I_1, \ldots, I_d \subset \mathbb{R}$ Intervalle sind. Sie können den linken bzw. rechten Rand enthalten und nicht enthalten. Wenn alle Intervalle beschränkt sind heißt der Quader endlich.

Für jeden solchen Quader definieren wir das Volumen als das Produkt der Längen aller Intervalle $I_1, \ldots, I_d$. Wenn die Intervalle alle beschränkt sind, sind alle ihre Längen endlich und das Volumen des entsprechenden Quaders ist dann auch endlich. Das Volumen bezeichnen wir mit $\mu(Q)$.

Definition 12.2. *Eine Teilmenge A des $\mathbb{R}^d$ heißt Nullmenge, wenn es für jedes $\epsilon > 0$ eine Folge von endlichen Quadern $(Q_n)_{n\in\mathbb{N}}$ im $\mathbb{R}^d$ gibt, mit*

$$A \subset \bigcup_{n+1}^{\infty} Q_n \quad \textit{und} \quad \sum_{n=1}^{\infty} \mu(Q_n) \leq \epsilon.$$

Lemma 12.3. *Jede abzählbare Teilmenge von $\mathbb{R}^d$ ist eine Nullmenge.*

Beweis: Sei $(x_n)_{n\in\mathbb{N}}$ eine beliebige Folge und $\epsilon > 0$. Sei Q_n für alle $n \in \mathbb{N}$ der Quader mit Zentrum x_n, dessen Kantenlängen alle gleich $\sqrt[d]{\epsilon \cdot 2^{-n}}$ sind. Dann überdeckt die Folge $(Q_n)_{n\in\mathbb{N}}$ die Folge $(x_n)_{n\in\mathbb{N}}$ im $\mathbb{R}^d$. Wegen der geometrischen Reihe gilt aber

$$\sum_{n=1}^{\infty} \mu(Q_n) = \sum_{n=1}^{\infty} \epsilon \cdot 2^{-n} = \epsilon \sum_{n=1}^{\infty} 2^{-n} = \epsilon.$$

Also gibt es für jedes $\epsilon > 0$ eine Folge $(Q_n)_{n\in\mathbb{N}}$, die $(x_n)_{n\in\mathbb{N}}$ überdeckt, und deren Gesamtvolumen nicht größer als ϵ ist. **q.e.d.**

Lemma 12.4. *Eine höchstens abzählbare Vereinigung von Nullmengen ist wieder eine Nullmenge.*

Beweis: Sei $\bigcup_{n=1}^{\infty} A_n$ eine höchstens abzählbare Vereinigung von Nullmengen. Dann besitzt für jedes $\epsilon > 0$ jedes A_n eine Überdeckung von Quadern, deren gesamtes Volumen nicht größer ist als $\epsilon \cdot 2^{-n}$. Die höchstens abzählbare Vereinigung dieser jeweils höchstens abzählbar vielen Quader ist wegen Satz 2.48 eine höchstens abzählbare Menge von Quader mit einem Volumen nicht gößer als $\sum_{n=1}^{\infty} \epsilon 2^{-n} = \epsilon$. Also wird $\bigcup_{n=1}^{\infty} A_n$ von abzählbar vielen Quadern überdeckt, deren Volumen nicht größer ist als ϵ. **q.e.d.**

Definition 12.5. *Eine Treppenfunktion ist eine endliche Linearkombination von charakteristischen Funktionen von endlichen Quadern, d.h. solcher Funktionen, die gleich* 1 *sind, auf Elementen des Quaders und sonst gleich* 0.

Proposition 12.6. *Endlich viele Quader lassen sich in endlich viele paarweise disjunkte Quader zerlegen. Jede Treppenfunktion ist eine endliche Linearkombination von charakteristischen Funktionen von paarweise disjunkten endlichen Quadern.*

Beweis: Es genügt zu zeigen, dass je zwei Quader Q_1 und Q_2 im $\mathbb{R}^d$ eine disjunkte Vereinigung von höchstens 3^d-Quadern ist. In einer Dimension folgt das daraus, dass zwei Intervalle in $\mathbb{R}$ entweder disjunkt sind, oder die Schnittmenge und die beiden Komplemente jeweils im anderen Intervall beides Intervalle sind, oder das eine Intervall in dem anderen enthalten ist und das Komplement dieses Intervalls in dem anderen Intervall eine disjunkte Vereinigung von höchstens zwei Intervallen ist. Wenn wir das auf alle Faktoren $\mathbb{R}$ im kartesischen Produkt anwenden, lassen sich zwei Quader in eine disjunkte Vereinigung von höchstens 3^d-Quadern zerlegen.

Wenn wir das auf alle endlichen Quader einer Treppenfunktion anwenden, dann nimmt diese Treppenfunktion auf jedem der paarweise disjunkten endlichen Quader genau einen Wert an und läßt sich als eine endliche Linearkombination von chrakteristischen Funktionen von paarweise disjunkten endlichen Quadern schreiben. **q.e.d.**

Als nächstes wollen wir das Integral von Treppenfunktionen definieren. Zunächst definieren wir für jede charakteristische Funktion χ_Q eines Quaders das Integral

$$\int \chi_Q d\mu = \mu(Q).$$

Proposition 12.7. *Sei f eine Treppenfunktion und seien*

$$f = \sum_i c_i \chi_{Q_i} = \sum_j d_j \chi_{R_j}$$

zwei Zerlegungen in endliche Linearkombinationen von charakteristischen Funktionen von endlichen Quadern. Dann ist

$$\int f d\mu = \sum_i c_i \mu(Q_i) = \sum_j d_j \mu(R_j).$$

Beweis: Wir zerlegen alle diese endlichen Quader Q_i und R_j in endlich viele paarweise disjunkte endliche Quader. Dabei zerlegen wir wie im vorangehenden Beweis beschrieben jeden der d Faktoren im kartesischen Produkt in eine disjunkte Vereinigung von endlich vielen beschränkten Intervallen. Auf jedem Quader dieser Zerlegung nimmt f die Summe aller c_i bzw. d_j an, deren entsprechnende Quader Q_i bzw. R_j diesen Quader enthalten. Dann genügt es offenbar zu zeigen, dass die Maße $\mu(Q_i)$ bzw. $\mu(R_j)$ jeweils gleich der Summe der Maße aller der Quader der Zerlegung sind, die in Q_i bzw. R_j enthalten sind. Wegen dem Distributivgesetz folgt das aus dem Spezialfall mit $d = 1$. Der folgt daraus, dass die Gesamtlänge einer disjunkten Vereinigung von Intervallen gleich der Summe der Intervalllängen ist. **q.e.d.**

Wegen dieser Proposition definiert das Integral $f \mapsto \int f d\mu$ eine lineare Abbildung von dem Raum aller Treppenfunktionen nach $\mathbb{R}$.

Proposition 12.8. *Seien f und g zwei Treppenfunktionen mit $f \geq g$. Dann gilt*

$$\int f d\mu \geq \int g d\mu.$$

Beweis: Wir zerlegen die beiden Vereinigungen von Quadern der Treppenfunktion f und der Treppenfunktion g in eine gemeinsame disjunkte Vereinigung von Quadern. Auf jedem der Quader ist f größer oder gleich g. Deshalb gilt das auch für die Summen, die die entsprechenden Integrale berechnen. **q.e.d.**

12.2 Lebesgueintegrable Funktionen auf dem $\mathbb{R}^d$

Satz 12.9. *Sei $(f_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge von Treppenfunktionen, deren Integrale $\left(\int f_n d\mu\right)_{n\in\mathbb{N}}$ beschränkt sind. Dann ist fogende Menge eine Nullmenge:*

$$\{x \in \mathbb{R} \mid (f_n(x))_{n\in\mathbb{N}} \textit{ konvergiert nicht } \}.$$

Beweis: Sei $M > 0$ eine obere Schranke von $\left(\int (f_n - f_1) d\mu\right)_{n\in\mathbb{N}}$:

$$\int (f_n - f_1) d\mu \leq M \qquad \text{für alle } n \in \mathbb{N}.$$

Dann ist für alle $\epsilon > 0$ und alle $n \in \mathbb{N}$, die Menge

$$S_{n,\epsilon} = \left\{ x \in \mathbb{R}^d \;\middle|\; f_n(x) - f_1(x) \geq \frac{M}{\epsilon} \right\} = \left\{ x \in \mathbb{R}^d \;\middle|\; f_n(x) \geq \frac{M}{\epsilon} + f_1(x) \right\}$$

eine monoton wachsende Folge von endlichen Vereinigungen von Quadern. Aus der Konstruktion einer gemeinsamen Zerlegung in eine disjunkte Vereinigung von Quadern im Beweis von Proposition 12.6 folgt, dass das relative Komplement eines Quaders in einem anderen Quader wieder eine disjunkte Vereinigung von Quadern ist. Dann ist

$$\bigcup_{n=1}^{\infty} S_{n,\epsilon} = S_{1,\epsilon} \cup (S_{2,\epsilon} \setminus S_{1,\epsilon}) \cup (S_{3,\epsilon} \setminus S_{2,\epsilon})$$

eine abzählbare Vereinigung von disjunkten Quadern. Weil $f_n - f_1$ nichtnegative Funktionen sind, ist $\frac{\epsilon}{M}(f_n - f_1)$ gößer oder gleich $\chi_{S_{n,\epsilon}}$. Also gilt für das Maß von $S_{n,\epsilon}$

$$\int \chi_{S_{n,\epsilon}} d\mu \leq \int \frac{\epsilon}{M}(f_n - f_1) d\mu = \frac{\epsilon}{M} \int (f_n - f_1) d\mu \leq \epsilon.$$

Wegen der Monotonie ist dann auch das Gesamtvolumen der abzählbaren Vereinigung $\bigcup_{n=1}^{\infty} S_{n,\epsilon}$ nicht größer als ϵ. Weil die kritische Menge S gleich der Schnittmenge

$$S = \{x \in \mathbb{R}^d | (f_n(x))_{n\in\mathbb{N}} \text{ konvergiert nicht } \} = \bigcap_{\epsilon>0} \left(\bigcup_{n=1}^{\infty} S_{n,\epsilon} \right)$$

ist, folgt, dass diese Menge S eine Nullmenge ist. **q.e.d.**

Die Komplemente von Nullmengen werden fast überall genannt. Also konvergiert jede monoton wachsende Folge von Treppenfunktionen fast überall.

Satz 12.10. *Für jede Nullmenge $A \subset \mathbb{R}^d$ gibt es eine monoton wachsende Folge $(f_n)_{n\in\mathbb{N}}$ von Treppenfunktionen mit beschränkten Integralen $(\int f_n d\mu)_{n\in\mathbb{N}}$, so dass A in der Menge enthalten ist, auf der die Folge $(f_n)_{n\in\mathbb{N}}$ nicht konvergiert.*

Beweis: Sei A eine Nullmenge. Dann gibt es für jedes $n \in \mathbb{N}$ eine Überdeckung von A mit abzählbar vielen Quadern, deren Gesamtvolumen nicht größer ist als 2^{-n}. Sei nun $(Q_n)_{n\in\mathbb{N}}$ eine Abzählung der Vereinigung aller dieser Quader. Dann gehört jeder Punkt von A zu unendlich vielen Quadern. Also definiert die Reihe $(\sum \chi_{Q_n})_{n\in\mathbb{N}}$ eine monoton wachsende Folge von Treppenfunktionen, die auf A nicht konvergiert. Die Integrale $(\sum \int \chi_{Q_n} d\mu)_{n\in\mathbb{N}}$ sind beschränkt durch $\sum_{n\in\mathbb{N}} 2^{-n} = 1$. **q.e.d.**

Für jede monoton wachsende Folge $(f_n)_{n\in\mathbb{N}}$ von Treppenfunktionen mit beschränkten Integralen $\left(\int f_n d\mu\right)_{n\in\mathbb{N}}$ können wir jetzt den Grenzwert fast überall definieren:

$$f(x) = \begin{cases} \lim\limits_{n\to\infty} f_n(x) & \text{wenn} (f_n(x))_{n\in\mathbb{N}} \text{ beschränkt ist} \\ 0 & \text{wenn } (f_n(x))_{n\in\mathbb{N}} \text{ nicht beschränkt ist.} \end{cases}$$

Dann wollen wir $\int f d\mu$ als den Grenzwert $\lim_{n\to\infty} \int f_n d\mu$ definieren. Damit diese Definition aber konsistent nur von der fast überall definierten Funktion f abhängt, benötigen wir noch die folgenden Lemmata.

Lemma 12.11. *Sei $(f_n)_{n\in\mathbb{N}}$ eine monoton fallende Folge von Treppenfunktionen, die fast überall gegen Null konvergiert. Dann ist $\left(\int f_n d\mu\right)_{n\in\mathbb{N}}$ eine Nullfolge.*

Beweis: Kein $\int f_n d\mu$ kann kleiner Null sein, weil sonst f_n auf einem Quader mit positiven Maß negativ ist und $(f_n)_{n\in\mathbb{N}}$ dort nicht gegen Null konvergiert. Offenbar gibt es einen kompakten Quader Q_0 außerhalb dessen f_1 verschwindet. Für jedes $n \in \mathbb{N}$ sei A_n die Menge der Unstetigkeitsstellen von f_n, also der Punkte, an denen f_n nicht lokal konstant ist. Dann ist A_n abgeschlossen und in Q_0 enthalten, und als eine endliche Vereinigung von $d-1$-dimensionalen Quadern eine Nullmenge. Dann ist auch $A = \bigcup_{n=1}^{\infty} A_n$ eine Nullmenge. Sei B die Nullmenge aller Punkte, an denen $(f_n)_{n\in\mathbb{N}}$ nicht gegen Null konvergiert. Für jedes $\epsilon > 0$ gibt es eine Überdeckung $\bigcup_{m=1}^{\infty} Q_m \supset (A \cup B)$ durch Quader, deren Gesamtvolumen nicht größer ist als $\frac{\epsilon}{2}$. Indem wir die Kanten aller Quader um ein hinreichend kleines ϵ' verlängern, dabei aber den Mittelpunkt festhalten, erhalten wir auch eine solche Überdeckung $\bigcup_{m=1}^{\infty} Q_m \supset (A \cup B)$ durch offene Quader, deren Gesamtvolumen nicht größer ist als ϵ. Für jeden Punkt $x \in Q_0 \setminus (A \cup B)$ gibt es ein N_x mit $f_{N_x}(x) \le \epsilon$. Weil $(f_n)_{n\in\mathbb{N}}$ monoton fallend ist, gilt $f_n(x) \le \epsilon$ für alle $n \ge N$. Weil alle f_{N_x} bei den Punkten von $Q_0 \setminus (A \cup B)$ lokal konstant sind, gibt es eine offene Überdeckung von offenen Quadern $(R_x)_{x\in Q_0\setminus(A\cup B)}$ von $Q_0 \setminus (A \cup B)$, und entsprechende Zahlen $(N_x)_{x\in Q_0\setminus(A\cup B)}$, so dass auf R_x für $n \ge N_x$ gilt $f_n \le \epsilon$. Dann bilden $(R_x)_{x\in Q_0\setminus(A\cup B)}$ zusammen mit $(Q_m)_{m\in\mathbb{N}}$ eine offene Überdeckung von Q_0. Weil Q_0 kompakt ist, gibt es eine endliche Teilüberdeckung. Wenn n größer ist als die entsprechenden endlich vielen N_x's können wir $\int f_n d\mu$ abschätzen durch

$$0 \le \int f_n d\mu \le \epsilon(\max\{f_1(x) \mid x \in Q_0\} + \mu(Q_0)).$$

Auf den Quadern $(Q_m)_{m\in\mathbb{N}}$ schätzen wir dabei f durch $\max\{f_1(x)|x \in Q_0\}$ ab und auf den Quadern $(R_x)_{x\in Q_0\setminus(A\cup B)}$ durch ϵ. Also konvergiert $\left(\int f_n d\mu\right)_{n\in\mathbb{N}}$ gegen Null. **q.e.d.**

Lemma 12.12. *Seien f und g fast überall definierte Grenzwerte von monoton wachsenden Folgen $(f_n)_{n\in\mathbb{N}}$ und $(g_n)_{n\in\mathbb{N}}$ von Treppenfunktionen mit beschränkten $\left(\int f_n d\mu\right)_{n\in\mathbb{N}}$ und $\left(\int g_n d\mu\right)_{n\in\mathbb{N}}$. Wenn fast überall $f \ge g$ gilt, dann gilt auch*

$$\lim_{n\to\infty} \int f_n d\mu \ge \lim_{n\to\infty} \int g_n d\mu.$$

Beweis: Für jedes feste $m \in \mathbb{N}$ erfüllen die Funktionen

$$((g_m - f_n)^+)_{n\in\mathbb{N}} = \left(\frac{1}{2}(g_m - f_n + |g_m - f_n|)\right)_{n\in\mathbb{N}}$$

die Voraussetzungen von dem vorangehenden Lemma. Deshalb konvergieren die entsprechenden Integrale gegen Null. Wegen $g_m - f_n \le (g_m - f_n)^+$ folgt aus Proposition 12.8 und Lemma 12.11

$$\int g_m d\mu - \lim_{n\to\infty} \int f_n d\mu \le 0 \quad \text{und damit auch} \quad \lim_{m\to\infty} \int g_m d\mu \le \lim_{n\to\infty} \int f_n d\mu.\textbf{q.e.d.}$$

Aus Lemma 12.12 folgt, dass wir das Integral auf die Grenzwerte von monoton wachsenden Folgen von Treppenfunktionen mit beschränkten Integralen konsistent fortsetzen können. Seien nämlich $(f_n)_{n\in\mathbb{N}}$ und $(g_n)_{n\in\mathbb{N}}$ monoton wachsenden Folgen von Treppenfunktionen mit beschränkten Integralen $\left(\int f_n d\mu\right)_{n\in\mathbb{N}}$ und $\left(\int g_n d\mu\right)_{n\in\mathbb{N}}$, deren Grenzwerte fast überall übereinstimmen, dann können wir Lemma 12.12 sowohl auf diese Folge, als auch auf die vertauschten Folgen anwenden und erhalten

$$\lim_{n\to\infty}\int f_n d\mu \geq \lim_{n\to\infty}\int g_n d\mu \geq \lim_{n\to\infty}\int f_n d\mu.$$

Definition 12.13. *Sei $L^1(\mathbb{R}^d)$ die Menge der Äquivalenzklassen von fast überall definierten Funktionen f, für die es monoton wachsende Folgen $(g_n)_{n\in\mathbb{N}}$ und $(h_n)_{n\in\mathbb{N}}$ von Treppenfunktionen mit beschränkten Integralen $\left(\int g_n d\mu\right)_{n\in\mathbb{N}}$ und $\left(\int h_n d\mu\right)_{n\in\mathbb{N}}$ gibt, so dass fast überall gilt*

$$f = \lim_{n\to\infty} g_n - \lim_{n\to\infty} h_n.$$

Hierbei werden zwei Funktionen miteinander identifiziert, wenn sie fast überall miteinander übereinstimmen.

Satz 12.14. *(Eigenschaften der lebesgueintegrablen Funktionen)*

(i) *$L^1(\mathbb{R}^d)$ ist ein Vektorraum über $\mathbb{R}$ und das Integral über Treppenfunktionen induziert eine lineare Abbildung*

$$\int : L^1(\mathbb{R}^d) \to \mathbb{R}, f \to \int f d\mu$$

(ii) *Wenn $f \in L^1(\mathbb{R}^d)$ fast überall nicht negativ ist, dann gilt* $\displaystyle\int f d\mu \geq 0.$

(iii) *Wenn $f \in L^1(\mathbb{R}^d)$, dann ist auch $|f| \in L^1(\mathbb{R}^d)$ mit* $\displaystyle\left|\int f d\mu\right| \leq \int |f|\, d\mu.$

Beweis: (i) Seien $(g_n)_{n\in\mathbb{N}}$, $(h_n)_{n\in\mathbb{N}}$, $(\tilde{g}_n)_{n\in\mathbb{N}}$ und $(\tilde{h}_n)_{n\in\mathbb{N}}$ monoton wachsende Folgen von Treppenfunktionen mit beschränkten Integralen. Wenn die Grenzwerte

$$g(x) - h(x) = \lim_{n\to\infty} g_n(x) - \lim_{n\to\infty} h_n(x)$$

fast überall mit den Grenzwerten von

$$\tilde{g}(x) - \tilde{h}(x) = \lim_{n\to\infty} \tilde{g}_n(x) - \lim_{n\to\infty} \tilde{h}_n(x)$$

übereinstimmen, dann stimmen auch die Funktionen $g(x) + \tilde{h}(x)$ und $\tilde{g}(x) + h(x)$ fast überall überein und sind fast überall auch die Grenzwerte von

$$(g_n + \tilde{h}_n)_{n\in\mathbb{N}} \text{ bzw. } (\tilde{g}_n + h_n)_{n\in\mathbb{N}}.$$

Dann folgt aus Lemma 12.12

$$\int (g+\tilde{h})d\mu = \int g d\mu + \int \tilde{h} d\mu = \int \tilde{g} d\mu + \int h d\mu = \int (\tilde{g}+h) d\mu.$$

Daraus folgt wegen der Linearität des Integrals

$$\int (g-h)d\mu = \int g d\mu - \int h d\mu = \int \tilde{g} d\mu - \int \tilde{h} d\mu = \int (\tilde{g}-\tilde{h}) d\mu.$$

Deshalb definiert $\int$ eine Abbildung von $L^1(\mathbb{R})$ nach $\mathbb{R}$. Die Linearität folgt aus den Rechenregeln für Folgen und der Linearität des Integrals auf Treppenfunktionen.
(ii) Wenn $(g_n)_{n\in\mathbb{N}}$ und $(h_n)_{n\in\mathbb{N}}$ monoton wachsende Folgen von Treppenfunktionen mit beschränkten Integralen sind, so dass fast überall $\lim_{n\to\infty} g_n - \lim_{n\to\infty} h_n$ nicht negativ ist, dann ist auch fast überall $g = \lim_{n\to\infty} g_n \geq \lim_{n\to\infty} h_n = h$. Aus Lemma 12.12 folgt dann $\int g d\mu \geq \int h d\mu$ bzw. $\int (g-h) d\mu \geq 0$.
(iii) Sowohl die Minima als auch die Maxima von zwei monoton wachsenden Folgen von Treppenfunktionen mit beschränkten Integralen sind wieder monoton wachsende Folgen von Treppenfunktionen mit beschränkten Integralen. Wenn f fast überall die Differenz $g-h$ der Grenzwerte der monoton wachsenden Folgen $(g_n)_{n\in\mathbb{N}}$ und $(h_n)_{n\in\mathbb{N}}$ von Treppenfunktionen mit beschränkten Integralen ist, dann ist $|f|$ fast überall die Differenz $\tilde{g}-\tilde{h}$ der Grenzwerte der monoton wachsenden Folgen $(\tilde{g}_n)_{n\in\mathbb{N}} = (\max\{g_n, h_n\})_{n\in\mathbb{N}}$ und $(\tilde{h}_n)_{n\in\mathbb{N}} = (\min\{g_n, h_n\}_{n\in\mathbb{N}}$ von Treppenfunktionen mit beschränkten Integralen. Deshalb ist $|f| \in L^1(\mathbb{R}^d)$. Wegen (ii) folgt dann aus $-|f| \leq f \leq |f|$

$$-\int |f| d\mu \leq \int f d\mu \leq \int |f| d\mu \quad \text{bzw.} \quad \left| \int f d\mu \right| \leq \int |f| d\mu. \qquad \textbf{q.e.d.}$$

Satz 12.15. *Eine beschränkte Funktion, die außerhalb einer beschränkten Menge verschwindet und deren Unstetigkeitsstellen eine Nullmenge bildet, gehört zu $L^1(\mathbb{R}^d)$.*

Beweis: Wir wählen einen Quader $Q_0 = [a_1, b_1] \times \ldots \times [a_d, b_d]$, außerhalb dessen die Funktion verschwindet. Wir zerlegen für alle $n \in \mathbb{N}$ und alle $i = 1, \ldots, d$ das Intervall $[a_i, b_i]$ in die Vereinigung der Intervalle

$$[a_i, b_i] = \left[a_i, a_i + \tfrac{b_i - a_i}{2^n}\right] \cup \left(a_i + \tfrac{b_i - a_i}{2^n}, a_i + 2\tfrac{b_i - a_i}{2^n}\right] \cup \ldots \cup \left(a_i + (2^n - 1)\tfrac{b_i - a_i}{2^n}, b_i\right].$$

Die kartesischen Produkte dieser Zerlegungen ergeben eine Zerlegung $\mathcal{P}_n$ von Q_0 in eine Vereinigung von 2^{nd} paarweise disjunkten Quadern. Dann sei f_n die Treppenfunktion, die auf jedem der 2^{nd} Quader gleich dem Infimum der entsprechenden Funktionswerte von f ist. Offenbar ist $(f_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge von Treppenfunktionen, deren Integrale durch $\|f\|_\infty \cdot \mu(Q)$ beschränkt sind. An allen Punkten $x_0 \in \mathbb{R}^d$, an denen f stetig ist, gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass aus $x \in B(x_0, \delta)$ folgt $f(x) \in B(f(x_0), \epsilon)$. Dann gibt es auch ein $N \in \mathbb{N}$, so dass der Durchmesser von Q

kleiner ist als $N\delta$. Für alle $n \geq N$ ist dann der Teilquader der 2^{nd} Teilquader von Q, der x_0 enthält, in $B(x_0, \delta)$ enthalten. Deshalb gilt dann

$$f(x_0) - \epsilon < f_n(x_0) \leq f(x_0).$$

Also konvergiert $(f_n(x_0))$ gegen $f(x_0)$. Weil aber die Menge der Unstetigkeitsstellen von f eine Nullmenge ist, konvergiert $(f_n)_{n\in\mathbb{N}}$ dann fast überall gegen f. **q.e.d.**

Satz 12.16.* *Sei $f \in L^1(\mathbb{R}^d)$ mit $\int |f| d\mu = 0$. Dann ist f fast überall gleich Null.*

Beweis:* Seien $(\tilde{g}_n)_{n\in\mathbb{N}}$ und $(\tilde{h}_n)_{n\in\mathbb{N}}$ monoton wachsende Folgen von Treppenfunktionen mit beschränkten Integralen, so dass fast überall gilt

$$f(x) = \lim_{n\to\infty} \tilde{g}_n(x) - \lim_{n\to\infty} \tilde{h}_n(x).$$

Für alle $n \in \mathbb{N}$ seien $\quad g_n = \max\{\tilde{g}_n, \tilde{h}_n\}$ und $h_n = \min\{\tilde{g}_n, \tilde{h}_n\}$.

Dann gilt fast überall $\quad |f(x)| = \lim_{n\to\infty} g_n(x) - \lim_{n\to\infty} h_n(x).$

Wenn $\int |f| d\mu = 0$ gilt also auch $\quad \lim_{n\to\infty} \int g_n d\mu = \lim_{n\to\infty} \int h_n d\mu.$

Für $\epsilon, \delta > 0$ sei $N \in \mathbb{N}$ so gewählt, dass $\quad \lim_{n\to\infty} \int g_n d\mu - \int h_m d\mu \leq \epsilon\delta.$

für alle $m \geq N$ gilt. Wir definieren g als den fast überall definierten Grenzwert $\lim_{n\to\infty} g_n$. Weil $(g_n)_{n\in\mathbb{N}}$ und $(h_n)_{n\in\mathbb{N}}$ monoton wachsend sind, ist die Menge

$\{x \in \mathbb{R}^d \mid |f(x)| > \delta\}$ in den Mengen

$$\{x \in \mathbb{R}^d \mid g(x) - h_m(x) > \delta\} = \{x \in \mathbb{R}^d \mid g_n(x) - h_m(x) > \delta \text{ für ein } n \in \mathbb{N}\}$$

enthalten. Sei also $\quad A_1 = \{x \in \mathbb{R}^d \mid g_1(x) - h_m(x) > \delta\}$ und für $n = 2, \ldots$

$A_n = \{x \in \mathbb{R}^d \mid g_n(x) - h_m(x) > \delta \text{ und } g_{n-1}(x) - h_m(x) \leq \delta\}$. Dann ist

$$\{x \in \mathbb{R}^d \mid g(x) - h_m(x) > \delta\} = \bigcup_{n=1}^{\infty} A_n \quad \text{und} \quad \sum_{n=1}^{\infty} \mu(A_n) \leq \lim_{n\to\infty} \frac{1}{\delta} \int (g_n - h_m) d\mu \leq \epsilon.$$

Also ist für alle $\delta > 0$ die Menge $\{x \in \mathbb{R}^d \mid g(x) - \lim_{m\to\infty} h_m(x) > \delta\}$ eine Nullmenge. Weil die abzählbare Vereinigung von Nullmengen wieder eine Nullmenge ist, folgt, dass die Menge $\bigcup_{n\in\mathbb{N}}\{x \in \mathbb{R}^d \mid g(x) - \lim_{m\to\infty} h_m(x) > \frac{1}{n}\}$ eine Nullmenge ist. Also ist fast überall $g(x) \leq h(x) = \lim_{m\to\infty} h_m(x)$. Aufgrund der Konstruktion gilt aber fast überall $g(x) \geq h(x)$. Also ist fast überall $|f(x)| = g(x) - h(x) = 0$. **q.e.d.**

12.3 Das Riemann- und das Lebesgueintegral

In diesem Abschnitt wollen wir das Riemannintegral mit dem Lebesgueintegral in Beziehung setzen. Für $d > 1$ sind die riemannintegrablen Funktionen f auf einem endlichen Quader Q_0 dadurch charakterisiert, dass das Unterintegral, also das Supremum aller Integrale von Treppenfunktionen nicht größer als f, mit dem Oberintegral, also dem Infimum aller Integrale von Treppenfunktionen nicht kleiner als f, übereinstimmt. Zunächst wollen wir die riemannintegrablen Funktionen charakterisieren.

Satz 12.17. *(Lebesguekriterium) Eine beschränkte Funktion auf einem endlichen Quader $Q_0 = [a_1, b_1] \times \ldots \times [a_d, b_d]$ ist genau dann riemannintegrabel, wenn ihre Unstetigkeitsstellen eine Nullmenge bilden. Insbesondere sind alle riemannintegrablen Funktionen auch lebesgueintegrabel und die beiden Integrale stimmen überein.*

Beweis: Wir zerlegen für alle $n \in \mathbb{N}$ und alle $i = 1, \ldots, d$ das Intervall $[a_i, b_i]$ in die Vereinigung der Intervalle

$$[a_i, b_i] = \left[a_i, a_i + \tfrac{b_i - a_i}{2^n}\right] \cup \left(a_i + \tfrac{b_i - a_i}{2^n}, a_i + 2\tfrac{b_i - a_i}{2^n}\right] \cup \ldots \cup \left(a_i + (2^n - 1)\tfrac{b_i - a_i}{2^n}, b_i\right].$$

Das entspricht für $d = 1$ der Partition $p_n \in \mathcal{P}[a, b]$ aus dem Beweis von Korollar 8.11. Die kartesischen Produkte dieser Zerlegungen ergeben eine Zerlegung $\mathcal{P}_n$ von Q_0 in eine Vereinigung von 2^{nd} paarweise disjunkte Quadern. Für alle $f \in B_{\mathbb{R}}(Q_0)$ seien

$$\begin{aligned} f_n(x) &= \inf\{f(y) \mid y \in Q \text{ mit } Q \in \mathcal{P}_n \text{ und } x \in Q\} \\ F_n(x) &= \sup\{f(y) \mid y \in Q \text{ mit } Q \in \mathcal{P}_n \text{ und } x \in Q\} \end{aligned}$$

Es sind $(f_n)_{n\in\mathbb{N}}$ und $(F_n)_{n\in\mathbb{N}}$ monoton wachsende bzw. monoton fallende Folgen von Treppenfunktionen mit beschränkten Integralen. Wenn die Folgen $(\int f_n d\mu)_{n\in\mathbb{N}}$ und $(\int F_n d\mu)_{n\in\mathbb{N}}$ gegen den gleichen Grenzwert konvergieren, dann ist f riemannintegrabel.

Wenn die Menge $\{x \in Q_0 \mid \lim_{n\to\infty} F_n(x) > \lim_{n\to\infty} f_n(x)\}$ eine Nullmenge ist, folgt aus Lemma 12.12 $\lim_{n\to\infty} \int F_n d\mu = \lim_{n\to\infty} \int f_n d\mu$. Umgekehrt folgt aus dieser Gleicheit, dass für alle $\epsilon > 0$ die Gesamtvolumen der Mengen

$$S_{n,\epsilon} = \{x \in Q_0 \mid F_n(x) - f_n(x) \geq \epsilon\}$$

im Grenzwert $n \to \infty$ nach Null konvergieren. Also ist dann die Schnittmenge

$$S_\epsilon = \bigcap_{n=1}^{\infty} S_{n,\epsilon} = \left\{x \in Q_0 \mid \lim_{n\to\infty} F_n(x) - \lim_{n\to\infty} f_n(x) \geq \epsilon\right\}$$

aller dieser Mengen eine Nullmenge. Dann ist auch die Menge

$$\{x \in Q_0 \mid \lim_{n\to\infty} F_n(x) > \lim_{n\to\infty} f_n(x)\} = \bigcup_{m=1}^{\infty} S_{\frac{1}{m}}$$

eine Nullmenge. Deshalb gilt genau dann $\lim_{n\to\infty}\int F_n d\mu = \lim_{n\to\infty}\int f_n d\mu$, wenn die Menge $\{x \in Q_0 \mid \lim_{n\to\infty} F_n(x) > \lim_{n\to\infty} f_n(x)\}$ eine Nullmenge ist.

Wenn f bei x stetig ist, dann gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass alle $x' \in B(x,\delta)\cap Q_0$ auch $|f(x') - f(x)| < \epsilon/2$ erfüllen. Dann gilt für $n > \frac{\|b-a\|}{\delta}$

$$|F_n(x) - f_n(x)| \le |F_n(x) - f(x)| + |f(x) - f_n(x)| \le \epsilon.$$

Dann gilt auch $\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} f_n(x)$. Also ist f riemannintegrabel, wenn die Unstetigkeitsstellen von f eine Nullmenge bilden.

Wenn umgekehrt f riemannintegrabel ist, dann folgt für $d = 1$ aus dem Kriterium von Riemann, dass die Integrale $(\int F_n d\mu)_{n\in\mathbb{N}}$ und $(\int f_n d\mu)_{n\in\mathbb{N}}$ gegen die gleiche Zahl konvergieren. Für $d > 1$ gibt es zwei Folgen von Treppenfunktionen nicht kleiner bzw. nicht größer als f, deren Integrale gegen die gleiche Zahl konvergieren. Die Minima bzw. Maxima der jeweils ersten n Folgenglieder dieser beiden Folgen definieren eine monoton fallende $(F_n)_{n\in\mathbb{N}}$ bzw. wachsende Folge $(f_n)_{n\in\mathbb{N}}$ von Treppenfunktionen mit beschränkten Integralen, deren Integrale gegen die gleiche Zahl konvergieren. Die Differenz $(F_n - f_n)_{n\in\mathbb{N}}$ ist eine monoton fallende Folge von nichtnegativen Treppenfunkltionen, deren Integrale gegen Null konvergieren. Wegen Satz 12.16 (siehe auch Korollar 12.23) konvergieren dann beide Folgen fast überall gegen die gleiche Funktion f. Die Unstetigkeitsstellen aller dieser Treppenfunktionen der beiden Folgen sind abzählbare Vereinigungen von Nullmengen und damit Nullmengen. Von jedem Punkt im Komplement dieser Nullmenge sind alle Quader der Treppenfunktionen, die den Punkt enthalten, eine Umgebung. Wenn f bei x im Komplement dieser Nullmenge unstetig ist, dann gibt es ein $\epsilon > 0$, so dass für alle $\delta > 0$ gilt:

$$\sup\{f(x') \mid x' \in B(x,\delta)\cap Q_0\} - \inf\{f(x') \mid x' \in B(x,\delta)\cap Q_0\} \ge \epsilon.$$

Deshalb konvergieren diese beiden Folgen nur dann fast überall gegen die gleiche Funktion, wenn die Unstetigkeitsstellen von f eine Nullmenge bilden. **q.e.d.**

Beispiel 12.18. (i) *Sei* $(r_n)_{n\in\mathbb{N}}$ *eine Abzählung aller rationalen Zahlen in* $(0,1)$. *Dann ist für* $0 < \epsilon < 1$ *das Komplement der Teilmenge*

$$\bigcup_{n=1}^{\infty}\left((r_n - 2^{-(n+1)}\epsilon, r_n + 2^{-(n+1)}\epsilon)\cap[0,1]\right)$$

von $[0,1]$ *keine Nullmenge, weil alle offenen Intervalle*

$$I_n = \left(r_n - 2^{-(n+1)}\epsilon, r_n + 2^{-(n+1)}\epsilon\right)\cap[0,1]$$

höchstens das Maß $\epsilon 2^{-n}$ *haben und* $\sum_{n=1}^{\infty}\epsilon 2^{-n} = \epsilon < 1$. *Also ist die Folge*

$$\left(\prod_{k=1}^{n}(1-\chi_{I_k})\right)_{n\in\mathbb{N}}$$

eine monoton fallende Folge von Treppenfunktionen, die gegen eine lebesgueintegrable Funktion konvergiert. Weil die rationalen Zahlen dicht in $[0,1]$ *liegen, ist diese Funktion an allen Punkten im Komplement der offenen Menge* $\bigcup_{n=1}^{\infty} I_n$ *unstetig. Also ist der Grenzwert von* $(\prod_{k=1}^{n}(1-\chi_{I_k}))_{n\in\mathbb{N}}$ *eine lebesgueintegrable Funktion, die nicht riemannintegrabel ist. Diese offene Menge ist also ein Beispiel für eine offene Menge, deren Rand positives Lebesguemaß hat, also eine charakteristische Funktion mit nicht verschwindendem Lebesgueintegral.*

(ii) *Sei* $p \in \mathbb{N} \setminus \{1,2\}$ *und* χ *die Funktion* $\chi : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \chi(x)$ *mit*

$$\chi(x) = \begin{cases} 0 & \text{wenn es eine ganze Zahl } q \in \mathbb{Z} \text{ gibt mit } x - q \cdot p \in (1,2) \\ 1 & \text{sonst.} \end{cases}$$

Dann definiert die Folge $\left(\prod_{k=1}^{n} \chi(p^k \cdot x)\right)_{n\in\mathbb{N}}$ *eine monoton fallende Folge von Treppenfunktionen auf* $[0,1]$ *mit beschränkten Integralen. Auf* $[0,1]$ *ist die Funktion also lebesgueintegrabel. Sie hat aber offenbar Unstetigkeitsstellen bei allen Zahlen, deren p-adische Bruchdarstellung aus endlich vielen Ziffern aus* $\{0,2,\ldots,p-1\}$ *besteht, und am Ende eine* 1 *oder* 2 *hat. Der Abschluss dieser Menge besteht aus allen Zahlen aus den Komplementen der Vereinigung von den offenen Mengen mit p-adischen Büchen*

$$(0.z_1\ldots z_n 1, \quad 0.z_1\ldots z_n 2) \qquad\qquad z_1,\ldots,z_n \in \{0,2,\ldots,p-1\}.$$

Das Maß dieser disjunkten Vereinigung von offenen Menge ist

$$\sum_{n=0}^{\infty} \frac{1}{p}\left(\frac{p-1}{p}\right)^n = 1.$$

Also ist das Komplement dieser offenen Menge eine Nullmenge. Diese Menge ist aber gleichmächtig zu der Menge aller Folgen mit Werten in $\{0,2,\ldots,p-1\}$*. Wenn* $p > 2$ *ist diese Menge nicht abzählbar. Aber der Grenzwert* $\lim_{n\to\infty}\left(\prod_{k=1}^{\infty}\chi(p^k x)\right)$ *ist riemannintegrabel und lebesgueintegrabel.*

12.4 Der Satz von Fubini

Für jeden Quader $Q = I_1 \times \ldots \times I_d \times I_{d+1} \subset \mathbb{R}^d \times \mathbb{R}$ und jedes $x \in \mathbb{R}^d$ ist die Funktion $\mathbb{R} \to \mathbb{R}, \quad y \mapsto \chi_Q(x,y)$ eine Treppenfunktion auf $\mathbb{R}$. Wenn wir diese Funktion integrieren erhalten wir eine Treppenfunktion auf dem $\mathbb{R}^d$:

$$\int \chi_Q(x,y)d\mu(y) = \begin{cases} \text{Länge von } I_{d+1} & \text{wenn } x \in I_1 \times \ldots \times I_d, \\ 0 & \text{wenn } x \notin I_1 \times \ldots \times I_d. \end{cases}$$

Also ist

$$\int \left(\int \chi_Q(x,y)d\mu(y)\right) d\mu(x) = \mu(Q).$$

Wegen der Linearität des Integrals definiert die Abbildung $\int d\mu(y)$ also eine lineare Abbildung von den Treppenfunktionen auf $\mathbb{R}^d \times \mathbb{R}$ in die Treppenfunktionen auf $\mathbb{R}^d$. Und für jede Treppenfunktion f auf $\mathbb{R}^d \times \mathbb{R}$ gilt

$$\int \left(\int f(x,y) d\mu(y) \right) d\mu(x) = \int f d\mu.$$

In diesem Abschnitt wollen wir zeigen, dass diese Abbildung eine Abbildung

$$\int d\mu(y) : L^1(\mathbb{R}^d \times \mathbb{R}) \to L^1(\mathbb{R}^d)$$

induziert, so dass für alle $f \in L^1(\mathbb{R}^d \times \mathbb{R})$ gilt

$$\int \left(\int f(x,y) d\mu(y) \right) d\mu(x) = \int f d\mu.$$

Wenn $f \geq g$ zwei Treppenfunktionen auf $\mathbb{R}^d \times \mathbb{R}$ sind, dann erfüllen für jedes $x \in \mathbb{R}^d$ die entsprechenden Treppenfunktionen $f_x : y \to f(x,y)$ bzw. $g_x : y \to g(x,y)$ auch $f_x \geq g_x$. Wegen Proposition 12.8 gilt für die Integrale auch

$$\int f(x,y) d\mu(y) \geq \int g(x,y) d\mu(y).$$

Also definiert $\int d\mu(y)$ eine lineare monotone Abbildung von den Treppenfunktionen auf $\mathbb{R}^d \times \mathbb{R}$ in die Treppenfunktionen auf $\mathbb{R}^d$. Damit diese Abbildungen eine Abbildung von $L^1(\mathbb{R}^d \times \mathbb{R})$ nach $L^1(\mathbb{R}^d)$ induziert, müssen zwei fast überall definierte Grenzwerte von monoton wachsenden Treppenfunktionen, die fast überall übereinstimmen, auch auf zwei fast überall definierte Grenzwerte von monoton wachsenden Treppenfunktionen abgebildet werden, die fast überall übereinstimmen.

Lemma 12.19. *Sei $S \subset \mathbb{R}^d \times \mathbb{R}$ eine Nullmenge. Dann ist fast überall in $x \in \mathbb{R}^d$, die Menge $S_x = \{y \in \mathbb{R} \mid (x,y) \in S\}$ eine Nullmenge von $\mathbb{R}$.*

Beweis: Wegen Satz 12.10 gibt es eine monoton wachsende Folge $(f_n)_{n\in\mathbb{N}}$von Treppenfunktionen auf $\mathbb{R}^d \times \mathbb{R}$ mit beschränkten Integralen, die auf S divergiert. Dann sind auch die entsprechenden Integrale $(\int f_n d\mu(y))_{n\in\mathbb{N}}$ über $\mathbb{R}$ monoton wachsende Treppenfunktionen mit beschränkten Integralen auf $\mathbb{R}^d$. Wegen Satz 12.9 konvergieren die entsprechenden Integrale dann fast überall auf $x \in \mathbb{R}^d$. Für alle $x \in \mathbb{R}^d$, für die die Integrale konvergieren, sind die entsprechenden Einschränkungen auf $\{x\} \times \mathbb{R}$ monoton wachsende Folgen von Treppenfunktionen auf $\mathbb{R}$ mit beschränkten Integralen. Wegen Satz 12.9 sind also für alle $x \in \mathbb{R}^d$, so dass die Integrale über $\mathbb{R}$ konvergieren, die Mengen S_x Nullmengen. **q.e.d.**

Proposition 12.20. *Die Integration über $\mathbb{R}$ induziert eine lineare monotone Abbildung von $L^1(\mathbb{R}^d \times \mathbb{R})$ nach $L^1(\mathbb{R}^d)$, so dass für alle $f \in L^1(\mathbb{R}^d \times \mathbb{R})$ gilt*

$$\int \left(\int f(x,y) d\mu(y) \right) d\mu(x) = \int f d\mu.$$

Beweis: Weil die Integration über $\mathbb{R}$ eine monotone lineare Abbildung von den Treppenfunktionen auf $\mathbb{R}^d \times \mathbb{R}$ in die Treppenfunktionen auf $\mathbb{R}^d$ definiert und wegen Lemma 12.19, induziert sie eine Abbildung von den Äquivalenzklassen von den Grenzwerten von monoton wachsenden Folgen von Treppenfunktionen mit beschränkten Integralen auf $\mathbb{R}^d \times \mathbb{R}$ in die entsprechenden Äquivalenzklassen auf $\mathbb{R}^d$. Wegen der Konstruktion des Lebesgueintegrals induziert sie also auch eine Abbildung von $L^1(\mathbb{R}^d \times \mathbb{R})$ nach $L^1(\mathbb{R}^d)$. Weil für alle Treppenfunktionen f auf $\mathbb{R}^d \times \mathbb{R}$ gilt

$$\int \left(\int f(x,y) d\mu(y) \right) d\mu(x) = \int f d\mu.$$

gilt das auch für alle $f \in L^1(\mathbb{R}^d \times \mathbb{R})$. **q.e.d.**

Die Argumente zeigen die analoge Aussage auch für die vertauschten Faktoren $\mathbb{R} \times \mathbb{R}^d$. Wenn wir die mehrfach anwenden erhalten wir also

Korollar 12.21. *(Satz von Fubini) Für alle Funktionen $f \in L^1(\mathbb{R}^d \times \mathbb{R}^{d'})$ gilt*

$$\int \left(\int f(x,y) d\mu(y) \right) d\mu(x) = \int \left(\int f(x,y) d\mu(x) \right) d\mu(y). \qquad \textbf{q.e.d.}$$

Mit dem Satz von Fubini und dem Lebesguekriterium können wir jetzt auch Integrale auf dem $\mathbb{R}^d$ ausrechnen. Als erstes können wir für fast alle $(x_2, \ldots, x_s) \in \mathbb{R}^d$ das Integral $\int_{-\infty}^{\infty} f(x_1, \ldots, x_d) dx_1$ ausrechnen. Dabei können wir die Methoden der eindimensionalen Integration, wie wir sie bei dem Riemannintegral kennen, benutzen. Dann integrieren wir genauso über $dx_2, \ldots, dx_d$ bis wir schließlich haben

$$\int\limits_{\mathbb{R}^d} f(x) d\mu = \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} f(x_1, \ldots, x_d) dx_1 \ldots dx_d.$$

Wir können die Reihenfolge dieser eindimensionalen Integrale beliebig vertauschen.

12.5 Konvergenzsätze

In diesem Abschnitt werden wir drei Aussagen darüber beweisen, wann Grenzwertbildungen mit der Integration vertauschen. Als erstes werden wir die Konvergenz von monotonen Folgen mit beschränkten Integralen beweisen.

Satz 12.22. *(Satz der monotonen Konvergenz von Beppo Levi) Sei $(f_n)_{n\in\mathbb{N}}$ eine monotone Folge in $L^1(\mathbb{R}^d)$ mit beschränkten Integralen. Dann konvergiert $(f_n)_{n\in\mathbb{N}}$ fast überall gegen eine Funktion f in $L^1(\mathbb{R}^d)$ und es gilt*

$$\lim_{n\to\infty}\int f_n d\mu = \int f d\mu.$$

Beweis: Sei $(f_n)_{n\in\mathbb{N}}$ eine monotone Folge in $L^1(\mathbb{R}^d)$ mit beschränkten Integralen. Durch Übergang zu $(\pm f_n)_{n\in\mathbb{N}}$ können wir annehmen, dass $(f_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge von Funktionen mit beschränkten Integralen ist. Für alle $n\in\mathbb{N}$ seinen $(\tilde{g}_{nm})_{m\in\mathbb{N}}$ und $(\tilde{h}_{nm})_{m\in\mathbb{N}}$ monoton wachsende Folgen von Treppenfunktionen mit beschränkten Integralen, so dass fast überall gilt

$$f_n = \lim_{m\to\infty}\tilde{g}_{nm} - \lim_{m\to\infty}\tilde{h}_{nm}.$$

Die entsprechenden Folgen der Integrale $(\int \tilde{g}_{nm}d\mu)_{m\in\mathbb{N}}$ und $(\int \tilde{h}_{nm}d\mu)_{m\in\mathbb{N}}$ konvergieren. Für alle $n\in\mathbb{N}$ sei $M(n)\in\mathbb{N}$ so gewählt, dass für alle $m,m'\geq M(n)$ gilt

$$\left|\int \tilde{h}_{nm}d\mu - \int \tilde{h}_{nm'}d\mu\right| \leq 2^{-n}.$$

Für alle $n\in\mathbb{N}$ seien h_{nm} und g_{nm} induktiv definiert durch

$$h_{nm} = \begin{cases} h_{n-1m} & \text{für } m < M(n) \\ \tilde{h}_{nm} - \tilde{h}_{nM(n)} + h_{n-1m} & \text{für } m \geq M(n)\end{cases} \quad \text{und} \quad g_{nm} = \tilde{g}_{nm} - \tilde{h}_{nM(n)} + h_{n-1m}$$

mit $h_{0m}=0$ für alle $m\in\mathbb{N}$. Weil für alle $n\in\mathbb{N}$ die Folge $(\tilde{h}_{nm})_{m\in\mathbb{N}}$ monoton wachsend ist, bestehen die Folgen $(h_{nm})_{m\in\mathbb{N}}$ induktiv für alle $n\in\mathbb{N}$ nur aus nicht negativen Funktionen. Weil auch die Folgen $(\tilde{g}_{nm})_{m\in\mathbb{N}}$ monoton wachsend sind, sind für alle $n\in\mathbb{N}$ auch die Folgen $(g_{nm})_{m\in\mathbb{N}}$ und $(h_{nm})_{n\in\mathbb{N}}$ monoton wachsend. Aufgrund der Wahl von $M(n)$ sind die Integrale $(\int h_{nm}d\mu - \int h_{n-1m}d\mu)_{m\in\mathbb{N}}$ beschränkt durch 2^{-n}. Also sind alle Integrale $(\int h_{nm}d\mu)_{n,m\in\mathbb{N}}$ beschränkt durch 1. Weil $(\int f_n d\mu)_{n\in\mathbb{N}}$ beschränkt sind, sind auch die Integrale $(\int g_{nm}d\mu)_{n,m\in\mathbb{N}}$ beschränkt. Für alle $n\in\mathbb{N}$ bzw. $m\in\mathbb{N}$ seien

$$g_n = \lim_{m\to\infty} g_{nm} \quad \text{und} \quad h_n = \lim_{m\to\infty} h_{nm},$$
$$\tilde{g}_m = \max\{g_{1m},\ldots,g_{mm}\} \quad \text{und} \quad \tilde{h}_m = \max\{h_{1m},\ldots,h_{mm}\}.$$

Dann gilt für alle $n\in\mathbb{N}$ auch fast überall $f_n = g_n - h_n$. Außerdem ist die Folge $(h_n)_{n\in\mathbb{N}}$ auf Grund der Definition von h_{nm} fast überall monoton wachsend und $(g_n)_{n\in\mathbb{N}} = (f_n + h_n)_{n\in\mathbb{N}}$ auch. Offenbar sind $(\tilde{g}_m)_{m\in\mathbb{N}}$ und $(\tilde{h}_m)_{m\in\mathbb{N}}$ monoton wachsende Folgen von Treppenfunktionen mit beschränkten Integralen. Seien $\tilde{g}$ und $\tilde{h}$ die entsprechenden Grenzwerte. Für $n\leq m$ gilt

$$g_{nm} \leq \tilde{g}_m \quad \text{und} \quad h_{nm} \leq \tilde{h}_m.$$

Also gilt für die entsprechenden Grenzwerte $m \to \infty$ fast überall

$$g_n \le \tilde{g} \quad \text{und} \quad h_n \le \tilde{h}.$$

Weil $(g_n)_{n\in\mathbb{N}}$ und $(h_n)_{n\in\mathbb{N}}$ fast überall monoton wachsende Folgen sind, gilt fast überall

$$\tilde{g}_m \le \max\{g_1,\dots,g_m\} = g_m \quad \text{und} \qquad \tilde{h}_m \le \max\{h_1,\dots,h_m\} = h_m.$$

Also gilt fast überall $g = \tilde{g}$ und $h = \tilde{h}$ bzw. $f = \tilde{g} - \tilde{h} = g - h$. Aus Lemma 12.12 folgt

$$\int f d\mu = \lim_{m\to\infty} \int \left(\tilde{g}_m - \tilde{h}_m\right) d\mu = \lim_{n\to\infty} \int (g_n - h_n) d\mu = \lim_{n\to\infty} \int f_n d\mu. \qquad \textbf{q.e.d.}$$

Korollar 12.23. *(Norm $\|\cdot\|_1$) Auf $L^1(\mathbb{R}^d)$ definiert*

$$\|\cdot\|_1 : \quad L^1(\mathbb{R}^d) \to \mathbb{R}, \quad f \mapsto \|f\|_1 = \textstyle\int |f| d\mu \qquad \textit{eine Norm.}$$

Beweis: Die Dreiecksungleichung und die Eigenschaft

$$\|\lambda f\|_1 = \int |\lambda| \cdot |f| d\mu = |\lambda| \cdot \int |f| d\mu = |\lambda| \|f\|_1$$

folgt aus der Monotonie und der Linearität des Lebesgueintegrals. Zu zeigen bleibt noch, dass aus $\|f\|_1 = 0$ folgt $f = 0$ fast überall. Sei also $\int |f| d\mu = 0$. Dann konvergiert wegen dem Satz der monotonen Konvergenz die Folge $(n|f|)_{n\in\mathbb{N}}$ fast überall. Also gilt auch fast überall $|f| = 0$. **q.e.d.**

Korollar 12.24. *(Lebesgue's Satz der beschränkten Konvergenz) Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge in $L^1(\mathbb{R}^d)$ und $k \in L^1(\mathbb{R}^d)$, so dass fast überall gilt $|f_n| \le k$ für alle $n \in \mathbb{N}$. Wenn $(f_n)_{n\in\mathbb{N}}$ fast überall gegen f konvergiert, dann ist $f \in L^1(\mathbb{R})$ und $\left(\int f_n d\mu\right)_{n\in\mathbb{N}}$ konvergiert gegen $\int f d\mu$.*

Beweis: Seien $(g_{nm})_{m\in\mathbb{N}}$ und $(h_{nm})_{m\in\mathbb{N}}$ definiert durch

$$g_{nm} = \min\{f_n, f_{n+1},\dots,f_{n+m}\} \quad \text{und} \quad h_{nm} = \max\{f_n, f_{n+1},\dots,f_{n+m}\}.$$

Für alle $n \in \mathbb{N}$ sind wegen den Eigenschaften der lebesgueintegrablen Funktionen $(g_{nm})_{m\in\mathbb{N}}$ monoton fallende Folgen in $L^1(\mathbb{R}^d)$ mit durch $\int k d\mu$ beschränkten Integralen, und $(h_{nm})_{m\in\mathbb{N}}$ monoton wachsende Folgen von Funktionen mit durch $\int k d\mu$ beschränkten Integralen. Wegen dem Satz der monotonen Konvergenz konvergieren diese Folgen fast überall gegen Funktionen $(g_n)_{n\in\mathbb{N}}$ und $(h_n)_{n\in\mathbb{N}}$ in $L^1(\mathbb{R}^d)$. Fast überall sind

$$g_n = \inf\{f_n, f_{n+1}, f_{n+2},\dots\} \quad \text{und} \quad h_n = \sup\{f_n, f_{n+1}, f_{n+2},\dots\}$$

monotone Folgen in $L^1(\mathbb{R}^d)$ mit beschränkten Integralen. Also konvergieren $(g_n)_{n\in\mathbb{N}}$ und $(h_n)_{n\in\mathbb{N}}$ fast überall gegen f. Dann gilt auch

$$\int g_n d\mu \le \int f_n d\mu \le \int h_n d\mu \quad \text{und} \quad \int f d\mu \le \lim_{n\to\infty} \int f_n d\mu \le \int f d\mu. \qquad \textbf{q.e.d.}$$

Korollar 12.25. *(Vollständigkeit von $L^1(\mathbb{R}^d)$, Satz von Riesz-Fischer)* $L^1(\mathbb{R}^d)$ *ist mit* $\|\cdot\|_1$ *ein Banachraum.*

Beweis: Sei $(f_n)_{m\in\mathbb{N}}$ eine Cauchyfolge in $L^1(\mathbb{R}^d)$. Dann gibt es eine Teilfolge $(f_{n_m})_{m\in\mathbb{N}}$, so dass für alle $m \in \mathbb{N}$ gilt $\|f_{n_{m+1}} - f_{n_m}\|_1 \leq 2^{-m}$. Die Reihe $\left(\sum_{m=1}^{n} \left|f_{n_{m+1}} - f_{n_m}\right|\right)_{n\in\mathbb{N}}$ erfüllt dann die Voraussetzungen des Satzes der monotonen Konvergenz. Also konvergiert sie fast überall gegen eine Funktion $k \in L^1(\mathbb{R}^d)$. Dann konvergiert auch die Folge $(f_{n_m} - f_{n_1})_{m\in\mathbb{N}}$ fast überall und erfüllt mit $k \in L^1(\mathbb{R}^d)$ die Voraussetzungen von Lebesgue's Satz der beschränkten Konvergenz. Dann konvergiert auch die Teilfolge gegen einen Grenzwert $f \in L^1(\mathbb{R}^d)$. Weil $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist, konvergiert $(\|f_n - f\|_1)_{n\in\mathbb{N}}$ gegen Null, und damit auch $(f_n)_{n\in\mathbb{N}}$ gegen f. **q.e.d.**

Satz 12.26. *(Fatou's Lemma) Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge in $L^1(\mathbb{R}^d)$ von fast überall nicht negativen Funktionen, die fast überall gegen f konvergieren. Wenn $\left(\int f_n d\mu\right)_{n\in\mathbb{N}}$ beschränkt ist, dann ist $f \in L^1(\mathbb{R}^d)$ und es gilt*

$$\int f d\mu \leq \lim_{n\to\infty} \inf_{m\in\mathbb{N}_0} \int f_{n+m} d\mu.$$

Beweis: Sei wieder $g_{nm} = \min\{f_n, f_{n+1}, \ldots, f_{n+m}\}$. Dann erfüllen für alle $n \in \mathbb{N}$ die Folgen $(g_{nm})_{m\in\mathbb{N}}$ die Voraussetzungen des Satzes der monotonen Konvergenz. Also konvergieren für alle $n \in \mathbb{N}$ diese Folgen gegen $g_n \in L^1(\mathbb{R}^d)$ mit $g_n = \inf\{f_n, f_{n+1}, \ldots\}$. Die Folgen $(g_n)_{n\in\mathbb{N}}$ erfüllen wieder die Voraussetzungen des Satzes der monotonen Konvergenz und konvergieren fast überall gegen f. Also gilt auch $f \in L^1(\mathbb{R}^d)$ und fast überall $f_{n+m} \geq g_n$ für alle $m \in \mathbb{N}_0$. Es folgt $\int f d\mu = \lim_{n\to\infty} \int g_n d\mu$ und $\int f d\mu \leq \lim_{n\to\infty} \inf_{m\in\mathbb{N}_0} \int f_{n+m} d\mu$. **q.e.d.**

12.6 Messbare Mengen und Maße

In diesem Abschnitt untersuchen wir, wann wir eine lebesgueintegrable Funktion über eine Teilmenge integrieren können. Das führt dann zu einer allgemeineren Definition vom Volumen von sogenannten messbaren Mengen. Dieses Volumen heißt Maß.

Definition 12.27. *Eine Menge $A \subset \mathbb{R}^d$ heißt messbar, wenn für jede nicht negative Funktion $f \in L^1(\mathbb{R}^d)$ das Produkt $f \cdot \chi_A$ mit der charakteristischen Funktion von A lebesgueintegrabel ist.*

Für messbare Mengen A bilden die Produkte von lebesgueintegrablen Funktionen mit der charakteristischen Funktion χ_A von A den Teilraum von $L^1(\mathbb{R}^d)$ aller lebesgueintegrablen Funktionen, die außerhalb von A verschwinden. Als diesen Teilraum definieren wird den Raum $L^1(A)$ aller lebesgueintegrablen Funktionen auf A.

Definition 12.28. *Für messbare Teilmengen A von $\mathbb{R}^d$ sei $L^1(A) \subset L^1(\mathbb{R}^d)$ der Teilraum aller lebesgueintegrablen Funktionen auf $\mathbb{R}^d$, die außerhalb von A verschwinden.*

Satz 12.29. **(i)** *Das Komplement einer messbaren Menge ist messbar.*

(ii) *Die Schnittmenge von abzählbar vielen messbaren Mengen ist messbar.*

(iii) *Jede offene Menge ist messbar.*

Beweis: **(i)** Weil die charakteristische Funktion des Komplements gerade 1 minus der charakteristischen Funktion ist, folgt (i) daraus, dass $L^1(\mathbb{R}^d)$ ein Vektorraum ist.
(ii) Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge von messbaren Mengen und f eine nicht negative Funktion in $L^1(\mathbb{R}^d)$. Dann ist die Folge $(f\prod_{k=1}^n \chi_{A_k})_{n\in\mathbb{N}}$ eine monoton fallende Folge in $L^1(\mathbb{R}^d)$ mit beschränkten Integralen. Wegen dem Satz der monotonen Konvergenz konvergiert sie in $L^1(\mathbb{R}^d)$. Der Grenzwert stimmt fast überall mit $f\chi_A$ überein, wobei $A=\bigcap_{n=1}^\infty A_n$.
(iii) Jeder Quader ist messbar, weil die Multiplikation einer Stufenfunktion mit der charakteristischen Funktion eines Quaders eine Stufenfunktion ergibt. Die Menge aller offenen Quader mit rationalen Koordinaten ist abzählbar. Weil jede offene Menge U gleich der Vereinigung aller der offenen Quader ist, deren Koordinaten in $\mathbb{Q}^d$ liegen, und die in U liegen, folgt (iii) aus (i) und (ii) und den de Morganschen Regeln. **q.e.d.**

Definition 12.30. *Eine Teilmenge $\mathcal{B}$ der Potenzmenge $\mathcal{P}(X)$ einer Menge X heißt σ–Algebra, wenn diese Teilmenge $\mathcal{B}\subset\mathcal{P}(X)$ unter Komplementbildung und dem Schnitt von abzählbar vielen Elementen von $\mathcal{B}$ abgeschlossen ist. Wenn X ein metrischer Raum ist, dann heißen die Elemente der kleinsten σ–Algebra, die alle offenen (und abgeschlossenen) Mengen enthält, Borelmengen.*

Definition 12.31. *Sei $\mathcal{B}$ eine σ–Algebra auf der Menge X. Dann heißt $\mu:\mathcal{B}\to\bar{\mathbb{R}}_0^+$ Maß, wenn für alle Folgen $(A_n)_{n\in\mathbb{N}}$ von paarweise disjunkten Mengen in $\mathcal{B}$ gilt*

$$\mu\left(\bigcup_{n=1}^\infty A_n\right)=\sum_{n=1}^\infty \mu(A_n).$$

Satz 12.32. *(Lebesguemaß) Seien $\mathcal{B}(\mathbb{R}^d)$ die Borelmengen von $\mathbb{R}^d$. Dann definiert*

$$\mu:\mathcal{B}(\mathbb{R}^d)\to\bar{\mathbb{R}}_0^+,\quad A\mapsto\lim_{n\to\infty}\int\chi_{Q_n}\chi_A d\mu\quad \textit{mit}\quad Q_n=(-n,n)^d\quad \textit{für alle } n\in\mathbb{N}$$

ein Maß auf den Borelmengen des $\mathbb{R}^d$, also ein Borelmaß. Es heißt Lebesguemaß.

Beweis: Wegen Satz 12.29 sind alle Borelmengen messbar. Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge von paarweise disjunkten Borelmengen. Für jede nichtnegative Funktion $g\in L^1(\mathbb{R}^d)$ ist

$$(gf_n)_{n\in\mathbb{N}}\quad\text{mit}\quad f_n=1-\prod_{k=1}^n(1-\chi_{A_k})\quad\text{für alle } n\in\mathbb{N}$$

eine monoton wachsende Folge von Funktionen in $L^1(\mathbb{R}^d)$ mit beschränkten Integralen. Also konvergiert die Folge $(gf_n)_{n\mathbb{N}}$ gegen eine Funktion gf in $L^1(\mathbb{R}^d)$, und es gilt

$$\lim_{n\to\infty}\int gf_n d\mu = \int gf d\mu.$$

Wenn wir diese Aussage auf die Folge $(\chi_{Q_n})_{n\in\mathbb{N}}$ der charakteristischen Funktionen von den Quadern Q_n anwenden, dann konvergiert die entsprechende Folge $(\chi_{Q_n}f)_{n\mathbb{N}}$ entweder gegen eine lebesgueintegrable Funktion, und $\mu(\bigcup_{n\in\mathbb{N}} A_n)$ ist endlich, oder das Maß der disjunkten Vereinigung der Borelmengen $(A_n)_{n\in\mathbb{N}}$ ist undendlich. In beiden Fällen folgt die σ–Additivität aus den Rechenregeln für Folgen. **q.e.d.**

Definition 12.33. *Eine Äquivalenzklasse von fast überall definierten rellen Funktionen auf $\mathbb{R}^d$ heißt messbar, wenn es eine Folge von Treppenfunktionen gibt, die fast überall gegen einen Repräsentanten der Äquivalenzklasse konvergiert.*

Satz 12.34. **(i)** *Die messbaren Funktionen bilden mit der punktweisen Addition und Multiplikation und der Skalarmultiplikation eine Algebra, die $L^1(\mathbb{R}^d)$ enthält.*

(ii) *Wenn f und g messbar sind, dann auch $|f|$, $\max\{f,g\}$ und $\min\{f,g\}$. Ist außerdem fast überall f ungleich Null, dann ist auch $\frac{1}{f}$ messbar.*

(iii) *Der Grenzwert einer fast überall konvergenten Folge von messbaren Funktionen ist messbar.*

(iv) *Die Äquivalenzklassen von stetigen Funktionen sind messbar.*

(v) *Eine messbare Funktion ist genau dann lebesgueintegrabel, wenn ihr Betrag durch eine lebesgueintegrable Funktion beschränkt ist.*

(vi) *Eine Menge $A \subset \mathbb{R}^d$ ist genau dann messbar, wenn χ_A messbar ist.*

Beweis (i): Die Treppenfunktionen bilden eine Algebra, und damit auch die messbaren Funktionen. Alle lebesgueintegrablen Funktionen sind nach unserer Definition messbar.
(ii): Der Betrag, das Maximum und das Minimum von Treppenfunktionen sind wieder Treppenfunktionen. Wenn eine Folge $(f_n)_{n\in\mathbb{N}}$ von Treppenfunktionen fast überall gegen eine nicht verschwindende Funktion f konvergiert, dann konvergiert die Folge

$$(g_n)_{n\in\mathbb{N}} \quad \text{mit} \quad g_n : \mathbb{R}^d \to \mathbb{R}, \quad x \mapsto g_n(x) = \begin{cases} \frac{1}{f_n(x)} & \text{wenn } f_n(x) \neq 0 \\ 0 & \text{wenn } f_n(x) = 0 \end{cases}$$

von Treppenfunktionen fast überall gegen $\frac{1}{f}$.
(iii): Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von messbaren Funktionen, die fast überall gegen f konvergiert. Sei h folgende positive lebesgueintegrable Hilfsfunktion auf $\mathbb{R}^d$:

$$h : x \mapsto h(x) = \frac{2^{-n-d}}{n^d - (n-1)^d} \text{ mit } n \in \mathbb{N} \text{ so dass } x \in (-n,n)^d \setminus (1-n,n-1)^d.$$

Die Folge $(g_n)_{n\in\mathbb{N}}$ von messbaren Funktionen

$$g_n : \mathbb{R}^d \to \mathbb{R}, \quad x \mapsto g_n(x) = \frac{h(x)f_n(x)}{h(x) + |f_n(x)|} < h(x) \quad \text{für alle } n \in \mathbb{N}$$

konvergiert fast überall gegen $g = \frac{h\cdot f}{h+|f|}$. Wegen Lebesgue's Satz der beschränkten Konvergenz sind alle $(g_n)_{n\in\mathbb{N}}$ und dann auch g lebesgueintegrabel und damit auch messbar. Dann ist wegen (i) und (ii) auch folgende Funktion messbar:

$$\frac{hg}{h - |g|} = \frac{h\frac{hf}{h+|f|}}{h - \frac{h|f|}{h+|f|}} = \frac{h^2 f}{h^2 + h|f| - h|f|} = f.$$

(iv): Für jede stetige Funktion f und jeden Quader Q ist wegen Satz 12.15 das Produkt von f mit der charakteristischen Funktion von Q lebesgueintegrabel. Die Folge $(f\chi_{Q_n})_{n\in\mathbb{N}}$ der Produkte mit den charakteritischen Funktionen der Quader $Q_n = (-n, n)^d$ konvergiert punktweise gegen f. Also folgt (iv) aus (i) und (iii).
(v): Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Treppenfunktionen, die fast überall gegen f konvergiert und $|f| \leq k$ für ein $k \in L^1(\mathbb{R}^d)$. Dann konvergiert auch $(\max\{-k, \min\{k, f_n\}\})_{n\in\mathbb{N}}$ fast überall gegen f. Aus Lebesgue's Satz der beschränkten Konvergenz folgt (v).
(iv): Für eine messbare Menge $A \subset \mathbb{R}^d$ konvergieren die Produkte $(\chi_A\chi_{Q_n})_{n\in\mathbb{N}}$ der charakteristischen Funktionen mit denen der Quader $Q_n = (-n, n)^d$ punktweise gegen χ_A. Wegen (iii) ist also χ_A messbar. Die Umkehrung folgt aus (i) und (v). **q.e.d.**

Satz 12.35. **(i)** *Eine Äquivalenzklasse f von fast überall definierten reellen Funktionen auf $\mathbb{R}^d$ ist genau dann messbar, wenn*

(ii) *Für alle $a \in \mathbb{R}$ sind die Mengen $\{x \in \mathbb{R}^d \mid f(x) \leq a\}$ messbar.*

(iii) *Für alle $a \in \mathbb{R}$ sind die Mengen $\{x \in \mathbb{R}^d \mid f(x) \geq a\}$ messbar.*

(iv) *Für alle $a \in \mathbb{R}$ sind die Mengen $\{x \in \mathbb{R}^d \mid f(x) < a\}$ messbar.*

(v) *Für alle $a \in \mathbb{R}$ sind die Mengen $\{x \in \mathbb{R}^d \mid f(x) > a\}$ messbar.*

(vi) *Für alle $k \geq 0$ in $L^1(\mathbb{R}^d)$ liegt die Funktion $\max\{-k, \min\{k, f\}\}$ in $L^1(\mathbb{R}^d)$.*

Beweis (i)⇒(ii) folgt aus Satz 12.34, weil die Folge $n\,(\max(a + 1/n, f) - \max(a, f))$ gegen die charakteristische Funktion der Menge $\{x \in \mathbb{R}^d \mid f(x) \leq a\}$ konvergiert.
(ii)⇔(iii)⇔(iv)⇔(v): Wegen Satz 12.29 sind (ii) und (iv) bzw. (iii) und (v) äquivalent. Wegen $\{x \in \mathbb{R}^d \mid f(x) < a\} = \bigcup_{n\in\mathbb{N}}\{x \in \mathbb{R}^d \mid f(x) \leq a - \frac{1}{n}\}$ und $\{x \in \mathbb{R}^d \mid f(x) \leq a\} = \bigcap_{n\in\mathbb{N}}\{x \in \mathbb{R}^d \mid f(x) < a + \frac{1}{n}\}$ sind auch (ii) und (iv) äquivalent.
(v)⇒(vi): Wenn eine Funktion f die äquivalenten Bedingungen (ii)–(v) erfüllt, dann sind folgende Funktionen für alle $n \in \mathbb{N}$ messbar:

$$f_n : \mathbb{R}^d \to \mathbb{R}, \quad x \mapsto \frac{1}{n^2}\sum_{l=1}^{n^2} \chi_{\left\{x\in\mathbb{R}^d \mid f(x) \geq \frac{l}{n^2}\right\}} - \chi_{\left\{x\in\mathbb{R}^d \mid f(x) \leq -\frac{l}{n^2}\right\}}.$$

Die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert gegen f. Aus Satz 12.34 (iii) und (v) folgt (vi).
(vi)⇒(i): Sei Q_n für alle $n \in \mathbb{N}$ der Quader $Q_n = (-n,n)^d$. Wegen (vi) besteht die Folge $(\max\{-n\chi_{Q_n}, \min\{n\chi_{Q_n}, f\}\})_{n\in\mathbb{N}}$ aus lebsgueintegrablen Funktionen und konvergiert punktweise gegen f. Also folgt (i) aus Satz 12.34 (i) und (iii). **q.e.d.**

12.7 Die Räume $L^p(\mathbb{R}^d)$

Eine komplexwertige Funktion ist genau dann lebesgueintegrabel bzw. messbar, wenn sowohl der Realteil als auch der Imaginärteil lebesgueintegrabel bzw. messbar sind. Eine messbare komplexe Funktion ist offenbar genau dann lebesgueintegrabel, wenn der Absolutbetrag lebesgueintegrabel ist. Aus Satz 12.35 folgt, dass für jede messbare Funktion f und $p \in \mathbb{R}$ auch die Funktion $|f|^p$ messbar ist.

Definition 12.36. *Für alle $1 < p < \infty$ sei $L^p(\mathbb{R}^d)$ die Menge aller Äquivalenzklassen von fast überall definierten messbaren Funktionen von $\mathbb{R}^d$ nach $\mathbb{K}$, für die die Funktion $|f|^p$ lebesgueintegrabel ist. Der Raum $L^\infty(\mathbb{R}^d)$ ist definiert als die Menge aller Äquivalenzklassen von messbaren beschränkten Funktionen von $\mathbb{R}^d$ nach $\mathbb{K}$.*

Für $\mathbb{K} = \mathbb{R}$ und $p = 1$ stimmt diese Definition mit Definition 12.13 überein.

Satz 12.37. *Für alle $1 \le p \le \infty$ ist $L^p(\mathbb{R}^d)$ zusammen mit der Abbildung*

$$\|\cdot\|: \quad L^p(\mathbb{R}^d) \to \mathbb{R}, \quad f \mapsto \|f\|_p = \begin{cases} \left(\int |f|^p d\mu\right)^{\frac{1}{p}} & \text{für } 1 \le p < \infty \\ \inf\{C \in \mathbb{R}_0^+ \mid |f| \le C \text{ a.e.}\} & \text{für } p = \infty. \end{cases}$$

ein normierter Vektorraum. Für $1 \le p,q,r \le \infty$ mit $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ und $f \in L^p(\mathbb{R}^d)$ und $g \in L^q(\mathbb{R}^d)$ gilt $fg \in L^r(\mathbb{R}^d)$ und die Höldersche Ungleichung: $\|fg\|_r \le \|f\|_p \|g\|_q$.

Die Dreiecksungleichung der Norm $\|\cdot\|_p$ wird wieder Minkowski–Ungleichung genannt: für alle $f, g \in L^p(\mathbb{R}^d)$ gilt $f+g \in L^p(\mathbb{R}^d)$ und $\|f+g\|_p \le \|f\|_p + \|g\|_p$.
Beweis: Wir beweisen zuerst die Höldersche Ungleichung. Indem wir zu den Funktionen $|f|^r$ und $|g|^r$ übergehen anstatt der Funktionen f und g, und den Exponenten 1, $\frac{r}{p}$ und $\frac{r}{q}$ anstatt r, p und q, genügt es den Fall $1 = \frac{1}{p} + \frac{1}{q}$ für nicht–negative reelle Funktionen zu betrachten. Es gilt nämlich $\|f\|_p^r = \||f|^r\|_{\frac{p}{r}}$ bzw. $\|g\|_q^r = \||g|^r\|_{\frac{q}{r}}$. Für $p = 1$, $q = \infty$ bzw. $p = \infty$, $q = 1$ folgt die Höldersche Ungleichung aus den Eigenschaften der lebesguesintegrablen Funktionen. Für $f = 0$ oder $g = 0$ ist die Aussage offensichtlich. Sei also $f \ne 0$, $g \ne 0$ und $1 < p, q < \infty$ mit $1 = \frac{1}{p} + \frac{1}{q}$. Die Youngsche Ungleichung ergibt für die Funktionen $\frac{|f|}{\|f\|_p}$ und $\frac{|g|}{\|g\|_q}$

$$\frac{|f|}{\|f\|_p}\frac{|g|}{\|g\|_q} \le \frac{|f|^p}{\|f\|_p^p} + \frac{|g|^q}{\|g\|_q^q} \text{ fast überall.}$$

Wegen $f \in L^p(\mathbb{R}^d)$ und $g \in L^q(\mathbb{R}^d)$ ist die rechte Seite lebesguesintegrabel und das Integral gleich 1. Also ist auch die linke Seite lebesguesintegrabel und das Integral kleiner oder gleich 1. Daraus folgt $fg \in L^1(\mathbb{R}^d)$, $\|fg\|_1 \le \|f\|_p\|g\|_q$.

Wegen Korollar 12.23 erfüllen die Abbildungen $f \mapsto \|f\|_p$ die Positivität. Die Linearität ist offensichtlich. Für $p = 1$ und $p = \infty$ ist die Dreieckungleichung schon gezeigt. Sei also $1 < p < \infty$ und $f, g \in L^p(\mathbb{R}^d)$. Offenbar gilt für $f, g \in L^p(\mathbb{R}^d)$ fast überall $|f+g|^p \le 2^p \max^p\{|f|, |g|\} \le 2^p(|f|^p + |g|^p)$. Also liegt $f + g$ in $L^p(\mathbb{R}^d)$. Die Höldersche Ungleichung ergibt für die Funktionen $|f+g|^p = |f+g||f+g|^{p-1} \le (|f|+|g|)|f+g|^{p-1}$ mit $f, g \in L^p(\mathbb{R}^d)$ und $1 = \frac{1}{p} + \frac{1}{q} \Leftrightarrow (p-1)q = p$

$$\|f+g\|_p^p \le \|f \cdot |f+g|^{p-1}\|_1 + \|g \cdot |f+g|^{p-1}\|_1 \le (\|f\|_p + \|g\|_p)\|f+g\|_p^{p-1}.$$

Daraus folgt die Minkowski Ungleichung. **q.e.d.**

Satz 12.38. *(Satz von Riesz Fischer) Für alle $1 \le p \le \infty$ ist $L^p(\mathbb{R}^d)$ ein Banachraum.*

Beweis: Für $p = \infty$ folgt die Aussage aus den Sätzen 9.44 (iii) und 12.34 (iii). Sei $1 \le p < \infty$ und $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $L^p(\mathbb{R}^d)$. Sei $(f_{n_m})_{m\in\mathbb{N}}$ eine Teilfolge mit $\|f_{n_{m+1}} - f_{n_m}\|_p \le 2^{-m}$ für alle $m \in \mathbb{N}$. Die monoton wachsende Folge $g_m = |f_{n_1}| + \sum_{l=1}^m |f_{n_{l+1}} - f_{n_l}|$ für alle $m \in \mathbb{N}$ erfüllt wegen der Minkowskiungleichung $\|g_m\|_p \le 1 + \|f_{n_1}\|_p$ für alle $m \in \mathbb{N}$. Wegen dem Satz der monotonen Konvergenz, konvergiert $(g_m^p)_{m\in\mathbb{N}}$ gegen eine Funktion $k \in L^1(\mathbb{R}^d)$. Wegen dem Monotonieprinzip konvergiert $(g_m)_{m\in\mathbb{N}}$ fast überall gegen eine messbare Funktion g. Dann konvergiert $(f_{n_m})_{m\in\mathbb{N}}$ fast überall gegen eine messbare Funktionen f. Wegen Lebesgue's Satz der beschränkten Konvergenz ist $|f|^p$ lebesgueintegrabel und damit $f \in L^p(\mathbb{R}^d)$. Außerdem gilt fast überall $|f - f_{n_{m+1}}| \le |g - g_m|$ und $\|g - g_m\|_p \le 2^{-m}$. Also konvergiert $(f_{n_m})_{m\in\mathbb{N}}$ in $L^p(\mathbb{R}^d)$ gegen f. Weil $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist konvergiert sogar $(f_n)_{n\in\mathbb{N}}$ in $L^p(\mathbb{R}^d)$ gegen f. **q.e.d.**

Definition 12.39. *Für jede messbare Menge $A \subset \mathbb{R}^d$ und alle $1 \le p \le \infty$ sei $L^p(A)$ der Unterraum von $L^p(\mathbb{R}^d)$ aller Äquivalenzklassen von Funktionen, die außerhalb von A verschwinden.*

Für messbare Mengen A und $1 \le p \le \infty$ ist offenbar $L^p(A) \subset L^p(\mathbb{R}^d)$ abgeschlossen.

Korollar 12.40. *Für A messbar und $1 \le p \le \infty$ ist $L^p(A)$ ein Banachraum.* **q.e.d.**

Satz 12.41. *Für alle messbaren Mengen $A \subset \mathbb{R}^d$ und alle $1 \le p, q \le \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1$ ist folgende Abbildung eine Isometrie:*

$$j : L^q(A) \to \mathcal{L}(L^p(A), \mathbb{K}), \quad g \mapsto j(g) \text{ mit } \quad j(g) : L^p(A) \to \mathbb{K}, \quad f \mapsto \int\limits_A fg d\mu.$$

Beweis: Für messbaren Mengen $A \subset \mathbb{R}^d$ und $1 \leq p \leq \infty$ ist j wegen der Hölderschen Ungleichung lipschitzstetig mit $L = 1$. Wir definieren f so, dass im Beweis der Hölderschen Ungleichung die Youngsche Ungleichung eine Gleichheit wird. Sei also

$$f = \frac{|g|^{\frac{q}{p}+1}}{g} = \frac{|g|^{q\left(\frac{1}{p}+\frac{1}{q}\right)}}{g} = \frac{|g|^q}{g} \qquad \text{mit } f(x) = 0 \text{ für } g(x) = 0.$$

Für $1 \leq q < \infty$ ist f messbar und liegt in $L^p(A)$. Es gilt $\|f\|_p^p = \|g\|_q^q$ und $\|fg\|_1 = \|g\|_q^q$. Daraus folgt

$$\|g\|_q = \|g\|_q^{q-q\left(1-\frac{1}{q}\right)} = \frac{\|g\|_q^q}{\|g\|_q^{\frac{q}{p}}} = \frac{\|gf\|_1}{\|f\|_p} \leq \|j(g)\| \leq \|g\|_q \text{ für alle } g \in L^q(A) \setminus \{0\}.$$

Also ist für $1 < p \leq \infty$ die Abbildung j eine Isometrie. Für $p = 1$, $q = \infty$ und jedes $\epsilon > 0$ ist $\{x \in \mathbb{R}^d \mid |g(x)| \geq \|g\|_\infty - \epsilon\}$ keine Nullmenge und messbar. Auf allen Funktionen $f \in L^1(A)$, die außerhalb dieser Menge verschwinden, nimmt $|j(g)|$ größere Werte als $(\|g\|_\infty - \epsilon)\|f\|_1$ an. Also gilt für jedes $\epsilon > 0$ und jedes $g \in L^\infty(A) \setminus \{0\}$ auch $\|g\|_\infty - \epsilon \leq \|j(g)\| \leq \|g\|_\infty$. Dann ist j auch für $p = 1$ eine Isometrie. **q.e.d.**

12.8 Jacobi's Transformation von Maßen

In diesem Abschnitt untersuchen wir, wie sich die Integration unter Koordinatentransformationen verhält. Wir verallgemeinern also die Substitutionsregel auf $\mathbb{R}^d$. Wir benutzen in diesem Abschnitt auf $\mathbb{R}^d$ die Norm $\|\cdot\|_\infty$ aus Definition 9.9. Dann sind nämlich die offenen Bälle genau die offenen Quader mit gleichen Kantenlängen.

Lemma 12.42. **(i)** *Die Multiplikation mit einer beschränkten messbaren Funktion f ist eine lineare Abbildung in $\mathcal{L}(L^1(\mathbb{R}^d))$, deren Norm beschränkt ist durch $\|f\|_\infty$.*

(ii) *Sei $\Phi : U \to O$ eine bijektive lipschitzstetige Abbildung zwischen den offenen Mengen $U, O \subset \mathbb{R}^d$ mit Lipschitzkonstante L. Dann ist $f \mapsto f \circ \Phi^{-1}$ eine lineare stetige Abbildung $L^1(U) \to L^1(O)$, deren Norm beschränkt ist durch L^d.*

(iii) *Sei $\Phi : U \to O$ eine bijektive Abbildung zwischen offenen Mengen $U, O \subset \mathbb{R}^d$ mit $\|\Phi(x) - \Phi(y) - (x - y)\|_\infty \leq \epsilon\|x - y\|_\infty$ für $x, y \in U$ und ein $\epsilon \in (0, 1)$. Dann gilt*

$$\left| \int\limits_U f \circ \Phi d\mu - \int\limits_O f d\mu \right| \leq ((1-\epsilon)^{-d} - 1)\|f\|_1 \qquad \textit{für alle } f \in L^1(O).$$

Beweis: **(i)** folgt aus der Hölderschen Ungleichung in Satz 12.37: $\|fg\|_1 \leq \|f\|_\infty \|g\|_1$.
(ii) Im Beweis von Satz 12.29 (iii) haben wir gesehen, dass die Schnittmenge eines offenen Quaders $Q \subset \mathbb{R}^d$ mit U als offene Menge eine Vereinigung von den abzählbar

vielen offenen endlichen Quadern $(Q_n)_{n\in\mathbb{N}}$ mit rationalen Koordinaten ist. Wir können sogar erreichen, dass die Abschlüsse dieser Quader in U liegen. Wegen Proposition 12.6 ist $Q_{n+1} \setminus (Q_1 \cup \ldots \cup Q_n)$ für alle $n \in \mathbb{N}$ eine disjunkte Vereinigung von endlich vielen Quadern. Induktiv erhalten wir $Q \cap U$ als eine abzählbare Vereinigung von paarweise disjunkten endlichen Quadern mit rationalen Koordinaten, deren Abschlüsse in U liegen. Also liegen die Linearkombinationen von charakteristischen Funktionen von Quadern mit rationalen Koordinaten, deren Abschlüsse in U liegen, dicht in $L^1(U)$. Für jeden Quader Q in U ist $\chi_Q \circ \Phi^{-1} = \chi_{\phi[U]}$. Jeder Quader mit rationalen Koordinaten ist eine disjunkte Vereinigung von endlich vielen Quadern mit gleichen Kantenlängen, also bis auf eine Nullmenge eine disjunkte Vereinigung von endlich vielen offenen Bällen bezüglich der Norm $\|\cdot\|_\infty$ aus Definition 9.9. Wegen Satz 12.29 ist $\Phi[U]$ messbar. Wenn Q gleiche Kantenlängen hat liegt das Bild $\Phi[Q]$ innerhalb des Quaders mit dem Volumen $L^d \cdot \mu(Q)$, dessen Zentrum das Bild des Zentrums von Q ist, und dessen Kantenlängen L mal der Kantenlängen von Q sind. Für jede Treppenfunktion $f \in L^1(U)$ ist dann $f \circ \Phi^{-1}$ lebesgueintegrabel und es gilt

$$\|f \circ \Phi^{-1}\|_1 \leq L^d \|f\|_1.$$

Die Abbildung $f \mapsto f \circ \Phi^{-1}$ setzt sich zu einer linearen stetigen Abbildung $L^1(U) \to L^1(O)$ fort, deren Norm beschränkt ist durch L^d.
(iii) Wie in (ii) genügt es die Behauptung für chrakteristischen Funktionen von offenen Bällen bezüglich $\|\cdot\|_\infty$ zu beweisen. Aus $\|\Phi(x) - \Phi(y) - (x-y)\|_\infty \leq \epsilon \|x-y\|_\infty$ folgt

$$(1-\epsilon)\|x-y\|_\infty \leq \|\Phi(x) - \Phi(y)\|_\infty \leq (1+\epsilon)\|x-y\|_\infty.$$

Aus $\Phi(x) \in B(\Phi(y), r)$ folgt $x \in B(y, \frac{r}{1-\epsilon})$ und $\Phi(x) \in B(\Phi(y), r)$ aus $x \in B(y, \frac{r}{1+\epsilon})$. Wegen $1 - (1+\epsilon)^{-d} \leq (1-\epsilon)^{-d} - 1$ folgt dann

$$\left| \int \chi_Q \circ \Phi d\mu - \int \chi_Q d\mu \right| \leq \left((1-\epsilon)^{-d} - 1 \right) \mu(Q).$$

Wegen der Linearität und der Dreiecksungleichung folgt die Behauptung für alle Treppenfunktionen f, und weil diese dicht liegen für alle $f \in L^1(O)$. **q.e.d.**

Satz 12.43. *Sei $A \in \mathcal{L}(\mathbb{R}^d)$ ein invertierbare lineare Abbildung. Dann ist*

$$\pi(A) : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d), \quad f \mapsto \pi(A)f \quad \text{mit } (\pi(A)f)(x) = f(A^{-1}x)$$

eine stetige lineare Abbildung, deren Norm beschränkt ist durch $\|A\|^d$.Es gilt

$$\int \pi(A) f d\mu = |\det(A)| \int f d\mu \quad \text{für alle } f \in L^1(\mathbb{R}^d).$$

Beweis: Wegen Lemma 12.42 (ii) ist die Abbildung

$$\pi(A) : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d), \quad f \mapsto \pi(A)f$$

beschränkt durch $\|A\|^d$. Alle Quader mit rationalen Kanten sind disjunkte Vereinigungen von Quadern mit gleichen Kantenlängen. Weil die charakteristischen Funktionen von Quadern mit rationalen Kantenlängen dicht liegen in den charakteristischen Funktionen von allen Quadern, liegen die Treppenfunktionen von Quadern mit gleichen Kantenlängen dicht in $L^1(\mathbb{R}^d)$. Für diagonale Matrizen A gilt wegen dem Satz von Fubini $\alpha(A) = |\det(A)|$. Dann gilt für alle A

$$\int f \circ A^{-1} d\mu = \alpha(A) \int f d\mu \quad \text{mit} \quad \alpha(A) = \int \chi_{A[[0,1]^d]} d\mu \quad \text{für alle } f \in L^1(\mathbb{R}^d).$$

Für zwei invertierbare Elemente $A, B \in \mathcal{L}(\mathbb{R}^d)$ gilt $\pi(A \cdot B) = \pi(A) \cdot \pi(B)$. Also gilt $\alpha(A \cdot B) = \alpha(A)\alpha(B)$. Aus $\alpha(\mathbb{1}) = 1$ folgt $\alpha(A^{-1}) = \alpha^{-1}(A)$. Dann gilt $\alpha(A) = |\det(A)|$ für alle diagonalisierbaren Matrizen A. Aufgrund der Jordanschen Normalform, ist jede Matrix A mit paarweise verschiedenen Eigenwerten diagonalisierbar. Weil die Diskriminante des charakteristischen Polynoms einer Matrix A ein Polynom in den Einträgen von A ist, liegt die Teilmenge aller Matrizen, deren Diskrimante nicht verschwindet, dicht in allen Matrizen. Also liegen die diagonalisierbaren Matrizen dicht in allen invertierbaren Matrizen. Wegen Lemma 12.42 (iii) folgt $|\alpha(A^{-1}) - \alpha(\mathbb{1})| \leq ((1-\epsilon)^{-d} - 1)\alpha(\mathbb{1})$ aus $\|A^{-1} - \mathbb{1}\| \leq \epsilon$. Dann folgt $|\alpha(A) - \alpha(B)| \leq ((1 - \|A^{-1}\|\epsilon)^{-d} - 1)|\alpha(A)|$ aus $\|A - B\| \leq \epsilon$. Also ist α stetig. Für alle A gilt dann $\alpha(A) = |\det(A)|$. **q.e.d.**

Satz 12.44. *(Jacobische Transformationsformel) Sei $\Phi : U \to O$ eine stetig differenzierbare bijektive Abbildung von der offenen Menge $U \subset \mathbb{R}^d$ auf die offene Menge $O \subset \mathbb{R}^d$. Wenn Φ' auf U in $\mathcal{L}(\mathbb{R}^d)$ invertierbar ist, dann ist die Abbildung $f \to |\det(\Phi')| (f \circ \Phi)$ eine Isometrie von $L^1(O)$ nach $L^1(U)$, d.h. für $f \in L^1(O)$ gilt*

$$\int_U |\det(\Phi')| (f \circ \Phi) d\mu = \int_O f d\mu.$$

Beweis: Im Beweis von Lemma 12.42 (ii) haben wir gesehen, dass die Treppenfunktionen mit kompaktem Träger in O dicht in $L^1(O)$ liegen. Deshalb genügt es

$$\int_U |\det(\Phi')| (\chi_Q \circ \Phi) d\mu = \mu(Q)$$

für alle charakteristischen Funktionen χ_Q von kompakten Quadern in $Q \subset O$ zu zeigen. Weil Φ' auf U in $\mathcal{L}(\mathbb{R}^d)$ invertierbar ist, ist wegen dem Satz der inversen Funktion Φ^{-1} stetig differenzierbar und die Ableitung $\frac{d\Phi}{dx}$ auf $\Phi[Q]$ wegen Korollar 9.35 beschränkt und wegen Satz 9.39 gleichmäßig stetig. Dann gibt es für jedes $\epsilon > 0$ eine Zerlegung $Q =$

$Q_1 \cup \ldots \cup Q_N$ in paarweise disjunkte Quader, auf deren Urbildern $\Phi^{-1}[Q_1], \ldots, \Phi^{-1}[Q_N]$ jeweils $\|\Phi' - A\| < \|A\|\epsilon$ oder auch $\|\Phi' \circ A^{-1} - \mathbb{1}\| < \|A\| \cdot \|A^{-1}\|\epsilon \leq \epsilon$ gilt, wobei A der Wert von Φ' an einem Punkt x_0 des Urbildes des jeweiligen Quaders ist. Aus dem Schrankensatz folgt $\|(\Phi \circ A^{-1})(x) - (\Phi \circ A^{-1})(y) - (x - y)\|_\infty \leq \epsilon\|x - y\|_\infty$. Aus Lemma 12.42 (iii) folgt für $n = 1, \ldots, N$

$$\left|\int \chi_{Q_n} \circ \Phi \circ A^{-1} d\mu - \mu\left(Q_n\right)\right| \leq \left((1-\epsilon)^{-d} - 1\right) \mu\left(Q_n\right).$$

Auf den Quadern $x \in Q_1, \ldots, Q_N$ liegen jeweils alle Eigenwerte von $\Phi'(x) \circ A^{-1}$ in $(1 - \epsilon, 1 + \epsilon)$. Es folgt $|\det(\Phi') - \det(A)| \leq |\det(A)|((1 + \epsilon)^d - 1)$. Satz 12.43 ergibt

$$\left|\int \left|\det\left(\Phi'\right)\right| \left(\chi_{Q_n} \circ \Phi\right) d\mu - \int \chi_{Q_n} \circ \Phi \circ A^{-1} d\mu\right| \leq \left((1+\epsilon)^d - 1\right) (1-\epsilon)^{-d} \mu(Q_n).$$

Im Grenzwert $\epsilon \to 0$ folgt dann $\displaystyle\int \left|\det\left(\Phi'\right)\right| \left(\chi_Q \circ \Phi\right) d\mu = \mu(Q)$. **q.e.d.**

12.9 Integration über den Rand einer Menge

In diesem Abschnitt definieren wir die Integration über den Rand einer offenen Teilmenge $\Omega \subset \mathbb{R}^d$. Dafür muss dieser Rand allerdings stetig differenzierbar sein.

Definition 12.45. *(Zerlegung der Eins) Eine glatte Zerlegung der Eins bezüglich einer offenen Überdeckung einer offenen Teilmenge $\Omega \subset \mathbb{R}^d$, ist eine abzählbare Familie $(f_n)_{n\in\mathbb{N}}$ von glatten Funktionen $f_n : \Omega \to [0, 1]$, so dass*

(i) *für jedes $x \in \Omega$ auf einer Umgebung von x nur endlich viele f_n ungleich Null sind.*

(ii) *Für alle $x \in \Omega$ gilt $\sum_{n=1}^{\infty} f_n(x) = 1$.*

(iii) *Jedes f_n außerhalb eines Elementes der offenen Überdeckung verschwindet.*

Satz 12.46. *(Existenz der Zerlegung der Eins) Jede offene Überdeckung $\mathcal{U}$ einer offenen Teilmenge $\Omega \subset \mathbb{R}^d$ besitzt eine glatte Zerlegung der Eins.*

Beweis: Für $0 < a < b < \infty$ sei $f_{a,b} \in C^\infty(\mathbb{R})$ definiert durch

$$f_{a,b} : \mathbb{R} \to [0, 1], \quad x \mapsto f_{a,b}(x) = \begin{cases} 1 & \text{für } 0 \leq |x| \leq a \\ \exp\left(\frac{1}{x^2 - b^2} \exp\left(\frac{-1}{x^2 - a^2}\right)\right) & \text{für } a < |x| < b \\ 0 & \text{für } b \leq |x| \end{cases}$$

Jeder Punkt x einer offenen Menge $\Omega \subset \mathbb{R}^d$ ist im Inneren eines abgschlossenen Balles $\overline{B(y, r)} \subset \Omega$ mit rationalem Zentrum $y \in \Omega \cap \mathbb{Q}^d$ und rationalem $r \in \mathbb{Q}$ enthalten.

Deshalb ist jede offene Menge $\Omega \subset \mathbb{R}^d$ eine abzählbare Vereinigung von relativkompakten offenen Teilmengen von Ω. Weil der Abschluss von endlich vielen Teilmengen dieser abzählbaren Überdeckung jeweils kompakt ist und deshalb von endlich vielen Elementen der Überdeckung überdeckt wird, gibt es eine Folge von offenen Teilmengen $(O_n)_{n\in\mathbb{N}}$ von Ω, die Ω überdecken, und deren Abschlüsse $\bar{O}_n$ für alle $n \in \mathbb{N}$ in O_{n+1} enthalten sind. Wir setzen $O_0 = O_{-1} = \emptyset$. Für jedes $n \in \mathbb{N}$ ist jeder Punkt $x \in \bar{O}_n \setminus O_{n-1}$ in einem Ball $B(y, r)$ enthalten, für den $B(y, 2r)$ in $O_{n+1} \setminus \bar{O}_{n-2}$ und in einer der offenen Mengen von $\mathcal{U}$ enthalten ist. Also wird die kompakte Menge $\bar{O}_n \setminus O_{n-1}$ von endlich vielen solchen Bällen überdeckt. Wir erhalten induktiv eine Folge von offenen Bällen $(B(y_n, r_n))_{n\in\mathbb{N}}$, die Ω überdecken, und deren Bälle $(B(y_n, 2r_n))_{n\in\mathbb{N}}$ jeweils in einer der offenen Mengen von $\mathcal{U}$ enthalten sind, und die jeweils nur mit endlich vielen anderen Bällen von $(B(y_n, 2r_n))_{n\in\mathbb{N}}$ nicht schnittfremd sind. Dann ist die Familie $(f_n)_{n\in\mathbb{N}}$

$$\begin{aligned} f_n : \Omega \to [0,1], \quad x \mapsto f_n(x) &= f_{r_n,2r_n}(\|x - y_n\|) \prod_{m=1}^{n-1} (1 - f_{r_m,2r_m}(\|x - y_m\|)) \\ &= \prod_{m=1}^{n-1} (1 - f_{r_m,2r_m}(\|x - y_m\|)) - \prod_{m=1}^{n} (1 - f_{r_m,2r_m}(\|x - y_m\|)). \end{aligned}$$

eine glatte Zerlegung der Eins bezüglich der Überdeckung $\mathcal{U}$. Denn induktiv gilt

$$f_1(x) + \ldots + f_n(x) + (1 - f_{r_1,2r_1}(\|x - y_1\|) \cdots (1 - f_{r_n,2r_n}(\|x - y_n\|) = 1 \qquad \text{für jedes } n \in \mathbb{N},$$

und für alle $n \in \mathbb{N}$ verschwindet $(1 - f_{r_n,2r_n}(\|x - y_n\|)$ auf $B(y_n, r_n)$. **q.e.d.**

Die Mengen, die wir im Gauß'schen Satz betrachten wollen, sollen einen zweimal differenzierbaren Rand haben.

Definition 12.47. *Eine offene Menge $\Omega \subset \mathbb{R}^d$ hat einen k–mal (stetig) differenzierbaren Rand $\partial\Omega = \overline{\Omega} \cap \Omega^c$, wenn jeder Punkt $x \in \overline{\Omega}$ eine offene Umgebung $U \subset \mathbb{R}^d$ und eine bijektive, k–mal (stetig) differenzierbare Abbildung $\Phi : U \to O \subset \mathbb{R}^d$ auf eine offene Teilmenge mit k–mal (stetig) differenzierbarer Umkehrabbildung besitzt, so dass*

$$\begin{aligned} \Phi[\Omega \cap U] &= B(0, r) \cap \{x \in \mathbb{R}^d | x_d > 0\} \\ \Phi[\Omega^c \cap U] &= B(0, r) \cap \{x \in \mathbb{R}^d | x_d \leq 0\} \\ \Phi[\partial\Omega \cap U] &= B(0, r) \cap \{x \in \mathbb{R}^d | x_d = 0\} \end{aligned}$$

Ein Vektor y gehört genau dann bei $x \in \partial\Omega$ zum Tangentialraum an $\partial\Omega$, wenn gilt

$$\Phi'(x)y \cdot e_d = 0 \qquad \Longleftrightarrow \qquad y \cdot (\Phi'(x))^t e_d = 0 \qquad \text{mit } e_d = (0, \ldots, 0, 1).$$

Deshalb ist der äußere Normalvektor $N(x)$ im Punkt x gegeben durch

$$N(x) = -\left\|(\Phi'(x))^t e_d\right\|^{-1} (\Phi'(x))^t e_d.$$

Unter stetig differenzierbaren Abbildungen Φ mit stetig differenzierbaren Umkehrabbildungen transformieren sich der Tangentialraum an den Rand durch die Jacobimatrix und der Normalenvektor durch die transponierte der Jacobimatrix.

Beispiel 12.48. (i): *Sei f eine einmal stetig differenzierbare Funktion auf $\mathbb{R}^d$, deren Gradient ∇f keine gemeinsamen Nullstellen mit f hat. Dann hat die Menge $\Omega = \{x \in \mathbb{R}^d \mid f(x) < 0\}$ einen stetig differenzierbaren Rand. Der Rand $\partial\Omega$ ist die Hyperfläche, auf der f verschwindet, die wegen dem Satz der impliziten Funktion lokal das Bild einer stetig differenzierbaren Abbildung von offenen Teilmengen von $\mathbb{R}^{d-1}$ nach $\mathbb{R}^d$ ist. Auf dem Rand $\partial\Omega$ ist ∇f ein Vektor, der orthogonal auf dem Tangentialraum an den Rand steht, und nach außen zeigt. Deshalb ist $\frac{\nabla f}{\|\nabla f\|}$ die äußere Normale.*
(ii): *Der Ball $B(y, r)$ mit $y \in \mathbb{R}^d$ und $r > 0$ im $\mathbb{R}^d$ ist gleich der Menge $B(y, r) = \{x \in \mathbb{R}^d \mid (x - y)^2 - r^2 < 0\}$. Die Funktion $f(x) = (x - y)^2 - r^2$ ist offenbar unendlich oft differenzierbar, und der Gradient $\nabla f(x) = 2(x - y)$ hat nur eine Nullstelle bei $x = y$. Also hat der Ball $B(y, r)$ einen unendlich oft differenzierbaren Rand.*

Lemma 12.49. *Sei $\Omega \subset \mathbb{R}^d$ eine offene Menge mit stetig differenzierbarem Rand und $\Phi : U \to O \subset \mathbb{R}^d$ eine einmal stetig differenzierbare Abbildung mit den Eigenschaften wie in Definition 12.47. Sei f eine stetige Funktion auf $\partial\Omega$ mit kompaktem Träger die außerhalb von U verschwindet. Dann ist das Lebesgueintegral*

$$\int\limits_{\Phi[\partial\Omega\cap U]} \left(\left\|(\Phi')^t\, e_d\right\| \cdot \left|\det\left(\Phi'\right)\right|^{-1} f\right) \circ \Phi^{-1} d\mu_{\mathbb{R}^{d-1}}$$

unabhängig von der Wahl von Φ. Hierbei ist $e_d = (0, \ldots, 0, 1)$.

Beweis: Seien Φ und $\tilde{\Phi}$ zwei solche stetig differenzierbare Abbildungen von der offenen Teilmenge U nach O bzw. $\tilde{O}$. Dann ist $\tilde{\Phi} \circ \Phi^{-1}$ eine stetig differenzierbare Abbildung von O nach $\tilde{O}$. Im Folgenden seien $y = \Phi(x)$ die entsprechenden Koordinate von O und $\tilde{\Phi}(x)$ von $\tilde{O}$. Dann ist die letzte Komponente der Koordinaten von Φ bzw. $\tilde{\Phi}$ auf dem Rand $\partial\Omega$ identisch gleich Null. Also verschwinden auf dem Rand $\partial\Omega$ alle Einträge der letzten Zeile bis auf das letzte $(\tilde{\Phi} \circ \Phi^{-1})'_{dd}$ von der Jacobimatrix $(\tilde{\Phi} \circ \Phi^{-1})'$. Außerdem ist die Determinante dieser Jacobimatrix gleich dem Produkt von $(\tilde{\Phi} \circ \Phi^{-1})'_{dd}$ mit der Unterdeterminante der entsprechenden $(d - 1) \times (d - 1)$ Matrix, in der die letzte Zeile und die letzte Spalte gestrichen wurde. Also sind folgende Vektoren gleich:

$$\begin{aligned}
(\tilde{\Phi} \circ \Phi^{-1})'_{dd} e_d &= ((\tilde{\Phi} \circ \Phi^{-1})')^t e_d \\
(\tilde{\Phi}')^t e_d &= (\Phi')^t ((\tilde{\Phi} \circ \Phi^{-1})')^t e_d = (\tilde{\Phi} \circ \Phi^{-1})'_{dd} (\Phi')^t e_d.
\end{aligned}$$

Weil e_d in beide Mengen $\Phi[\Omega \cap U]$ und $\tilde{\Phi}[\Omega \cap U]$ hineinzeigt, ist $(\tilde{\Phi} \circ \Phi^{-1})'_{dd}$ positiv.

Dann ergibt Satz 12.44

$$\int\limits_{\tilde{\Phi}[\partial\Omega\cap U]} \left(\left\|(\tilde{\Phi}')^t e_d\right\| \cdot \left|\det(\tilde{\Phi}')\right|^{-1} f\right) \circ \tilde{\Phi}^{-1} d\mu_{\mathbb{R}^{d-1}} =$$

$$\int\limits_{\tilde{\Phi}[\partial\Omega\cap U]} \frac{|\det((\tilde{\Phi}\circ\Phi^{-1})')|}{(\tilde{\Phi}\circ\Phi^{-1})'_{dd}} \left(\left\|(\tilde{\Phi}')^t e_d\right\| \cdot \left|\det(\tilde{\Phi}')\right|^{-1} f\right) \circ \tilde{\Phi}^{-1} \circ (\tilde{\Phi}\circ\Phi^{-1}) d\mu_{\mathbb{R}^{d-1}} =$$

$$= \int\limits_{\Phi[\partial\Omega\cap U]} \left(\left\|(\Phi')^t e_d\right\| \left|\det(\Phi')\right|^{-1} f\right) \circ \Phi^{-1}) d\mu_{\mathbb{R}^{d-1}}. \qquad \textbf{q.e.d.}$$

Mit Hilfe dieses Lemma's definieren wir das Integral einer Funktion auf $\partial\Omega$.

Definition 12.50. *Sei $\Omega \subset \mathbb{R}^d$ eine offene Teilmenge mit stetig differenzierbarem Rand, und f eine stetige Funktion auf $\partial\Omega$. Dann definieren wir das Integral $\int_{\partial\Omega} f d\sigma$ indem wir den Rand $\partial\Omega$ lokal auf Teilmengen von $\mathbb{R}^{d-1} \times \{0\}$ abbilden und es dann mit einer entsprechenden Zerlegung der Eins in eine Summe über die entsprechenden Integrale aus dem vorangehenden Lemma auswerten. Wegen dem vorangehenden Lemma ist das entsprechende Integral, wenn es existiert, unabhängig von den lokalen Abbildungen.*

12.10 Der Gaußsche Satz

Satz 12.51. *(Gauß'scher Satz oder Divergenzsatz) Sei $\Omega \subset \mathbb{R}^d$ ein offenes beschränktes Gebiet mit zweimal differenzierbarem Rand und f eine auf $\overline{\Omega}$ stetig $\mathbb{R}^d$-wertige Funktion, die auf Ω stetig differenzierbar ist, und deren partielle Ableitungen sich stetig auf $\bar{\Omega}$ forstsetzen lassen. Dann gilt:*

$$\int\limits_{\Omega} \nabla \cdot f d\mu = \int\limits_{\partial\Omega} f \cdot N d\sigma$$

Hierbei ist N die äußere Normale auf dem Rand $\partial\Omega$.

Beweis: Offenbar sind die linke und die rechte Seite der Gleichung linear in f. Deshalb können wir mit Hilfe einer Zerlegung der Eins die Funktion f in eine Summe von Funktionen mit Trägern in kleinen offenen Mengen zerlegen. Außerdem können wir f in eine Summe $f = \sum_{i=1} e_i f_i$ mit der kanonischen Basis $e_1, \ldots, e_d$ von $\mathbb{R}^d$ zerlegen. Zunächst möchten wir solche f betrachten, die außerhalb einer kompakten Teilmenge von Ω verschwinden. Für jedes $i = 1, \ldots, d$ gibt es dann ein Intervall I_i, so dass f_i außerhalb der Menge $\{x \in \mathbb{R}^d \mid x_i \in I_i\}$ verschwindet. Wegen dem Hauptsatz verschwindet dann für jedes $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) \in \mathbb{R}^{d-1}$ das Integral $\int_{I_i} \frac{\partial f}{\partial x_i}(x_1, \ldots, x_d) dx_i$ über das

Intervall I_i. Wegen dem Satz von Fubini folgt dann $\int_\Omega \nabla \cdot e_i f_i d\mu = 0$ Also erfüllen alle Funktionen mit kompaktem Träger in Ω den Gaußschen Satz.

Zuletzt betrachten wir Funktionen f die außerhalb einer Umgebung $U \subset \mathbb{R}^d$ eines Randpunktes verschwinden, wie sie in Definition 12.47 beschrieben ist. Es existiert also eine zweimal differenzierbare Abbildung $\Phi : U \to O \subset \mathbb{R}^d$ mit zweimal differenzierbarer Umkehrabbildung, so dass Φ den Teil $U \cap \partial\Omega$ vom Rand in U auf eine offene Teilmenge von $\mathbb{R}^{d-1} \times \{0\}$ abbildet. Wir definieren $\tilde{f} = \left(|\det(\Phi')|^{-1}(\Phi' f)\right) \circ \Phi^{-1}$, wobei wir punktweise die Ableitung Φ' auf den Vektor f wirken lassen. Dann folgt aus der Definition 12.50

$$\begin{aligned}
\int\limits_{\partial\Omega} f \cdot N d\sigma &= \int\limits_{\Phi[\partial\Omega\cap U]} \left(\left\|(\Phi')^t e_d\right\| |\det(\Phi')|^{-1} f \cdot N\right) \circ \Phi^{-1} d\mu_{\mathbb{R}^{d-1}} \\
&= -\int\limits_{\Phi[\partial\Omega\cap U]} \left(|\det(\Phi')|^{-1} f \cdot (\Phi')^t e_d\right) \circ \Phi^{-1} d\mu_{\mathbb{R}^{d-1}} \\
&= -\int\limits_{\Phi[\partial\Omega\cap U]} \tilde{f} \cdot e_d d\mu_{\mathbb{R}^{d-1}} = \int\limits_{\Phi[\partial\Omega\cap U]} \tilde{f} \cdot \tilde{N} d\mu_{\mathbb{R}^{d-1}}
\end{aligned}$$

Hierbei haben wir die äußere Normalen $N(x) = -\|(\Phi')^t e_d\|^{-1} (\Phi')^t e_d$ und $\tilde{N} = -e_d$ eingesetzt. Wenn wir f durch $\tilde{f}$ ausdrücken und Jacobi's Transformation von Maßen benutzen, erhalten wir

$$\int\limits_{\Omega} \nabla \cdot f d\mu = \int\limits_{\Phi[\Omega\cap U]} \left(|\det(\Phi')|^{-1}\left(\nabla_x \cdot |\det(\Phi')| (\Phi')^{-1} \tilde{f} \circ \Phi\right)\right) \circ \Phi^{-1} d\mu$$

Im Folgenden bezeichnen wir mit y die Koordinaten auf O und mit J die Jacobimatrix $J_{ij} = \frac{\partial \Phi_i}{\partial x_j} = \frac{\partial y_i}{\partial x_j}$. Mit Satz 10.4 (iii) und Korollar 11.7 erhalten wir

$$\begin{aligned}
\nabla_x (\Phi')^{-1} \tilde{f} \circ \Phi &= \sum_{i,j=1}^d \frac{\partial J^{-1}_{ij}}{\partial x_i} \tilde{f}_j \circ \Phi + \sum_{i,j,k=1}^d J^{-1}_{ij} \frac{\partial \tilde{f}_j}{\partial y_k} J_{ki} \circ \Phi \\
&= -\sum_{i,j,k,l=1}^d J^{-1}_{ij} \frac{\partial J_{jk}}{\partial x_i} J^{-1}_{kl} \tilde{f}_l \circ \Phi + (\nabla_y \cdot \tilde{f}) \circ \Phi.
\end{aligned}$$

Um den Gradienten der Determinante der Jacobimatrix zu berechnen, benötigen wir

Lemma 12.52. *Sei A eine differenzierbare Abbbildung von $(t_0 - \epsilon, t_0 + \epsilon)$ in die $d \times d$–Matrizen. Wenn $A(t_0)$ invertierbar ist, dann gilt*

$$\frac{d \det(A(t_0))}{dt} = \det(A(t_0)) \operatorname{Spur}\left(A^{-1}(t_0) \frac{dA(t_0)}{dt}\right).$$

Beweis: Für zwei $n \times n$–Matrizen A und B gilt

$$\det(A + tB) = \det(A)\det(\mathbb{1} + tA^{-1}B).$$

Offenbar ist $\det(\mathbb{1} + tA^{-1}B)$ ein Polynom in t vom Grad n, und die Koeffizienten sind Polynome in den Koeffizienten von $A^{-1}B$. Weil aber die Unterdeterminanten von $\mathbb{1}$ genau dann nicht verschwinden, wenn die genausovielte Spalte wie Zeile gestrichen wird und dann die Unterdeterminanten gleich Eins sind, gilt

$$\det(\mathbb{1} + tA^{-1}B) = 1 + t\,\mathrm{Spur}(A^{-1}B) + \text{Terme höherer Ordnung}.$$

Damit folgt $\left.\frac{d}{dt}\det(A + tB)\right|_{t=0} = \det(A)\,\mathrm{Spur}(A^{-1}B)$. **q.e.d.**

Damit erhalten wir

$$(\nabla_x |\det(\Phi')|) \cdot (\Phi')^{-1} \tilde{f} \circ \Phi \quad = \quad |\det(\Phi')| \sum_{i,j,k,l=1}^{d} J^{-1}_{ij} \frac{\partial J_{ji}}{\partial x_k} J^{-1}_{kl} \tilde{f}_l \circ \Phi.$$

Wegen dem Satz von Schwarz gilt

$$\frac{\partial J_{ji}}{\partial x_k} = \frac{\partial^2 y_j}{\partial x_k \partial x_i} = \frac{\partial^2 y_j}{\partial x_i \partial x_k} = \frac{\partial J_{jk}}{\partial x_i}, \text{ und damit auch } \int\limits_{\Omega} \nabla \cdot f d\mu = \int\limits_{\Phi[\Omega \cap U]} \nabla_y \cdot \tilde{f} d\mu.$$

Damit haben wir gezeigt, dass der Gaußsche Satz für Funktionen f, die außerhalb einer offenen Umgebung $U \subset \mathbb{R}^d$ verschwinden, wie sie in Definition 12.47 beschrieben ist, genau dann gilt, wenn er für die entsprechenden Funktionen $\tilde{f}$ gilt:

$$\int\limits_{\Phi[\Omega \cap U]} \nabla_y \cdot \tilde{f} d\mu = \int\limits_{\Phi[\partial\Omega \cap U]} \tilde{f} \cdot \tilde{N} d\mu_{\mathbb{R}^{d-1}}.$$

Für die Funktionen $\tilde{f}$ gilt für alle $i = 1, \ldots, d-1$ wieder $\int_\Omega \frac{\partial \tilde{f}_i}{\partial x_i} d\mu = 0$. Für $i = d$ benutzen wir den Satz von Fubini und den Hauptsatz der Differential– und Integralrechnung und erhalten

$$\int\limits_{\Phi[\Omega \cap U]} \frac{\partial \tilde{f}_d}{\partial y_d} d\mu = - \int\limits_{\Phi[\partial\Omega \cap U]} \tilde{f}_d d\mu_{\mathbb{R}^{d-1}}.$$

Daraus folgt der Gauß'sche Satz. **q.e.d.**

Index