Analysis I und II

WS 2004/05 und SS 2005

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:

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

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

oder nur

$$M = \{x \mid x \text{ hat die Eigenschaft } p\}$$

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

[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.

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

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:

$$\begin{aligned}
& A \setminus A = \emptyset \\
& A \setminus \emptyset = A \\
\text{(Idempotenz)} \quad & A \cup A = A \\
\text{(Idempotenz)} \quad & A \cap A = A \\
\text{(Kommutativität)} \quad & A \cup B = B \cup A \\
\text{(Kommutativität)} \quad & A \cap B = B \cap A \\
& A \cup \emptyset = A \\
& A \cap \emptyset = \emptyset
\end{aligned}$$

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

$$\begin{array}{ll}
\text{(Assoziativität)} & A \cap (B \cap C) = (A \cap B) \cap C \\
\text{(Assoziativität)} & A \cup (B \cup C) = (A \cup B) \cup C \\
\text{(Distributivität)} & A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \\
\text{(Distributivität)} & A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \\
\text{(de Morgan)} & (A \cap B)^c = A^c \cup B^c \\
\text{(de Morgan)} & (A \cup B)^c = A^c \cap B^c \\
& (A^c)^c = A \\
& A \subset B \Longleftrightarrow B^c \subset A^c \\
& A \times \emptyset = \emptyset \times A = \emptyset \\
& (A \times C) \cup (B \times C) = (A \cup B) \times C \\
& (A \times C) \cap (B \times C) = (A \cap B) \times C
\end{array}$$

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 verschiedene 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 aber $[a] = [c] = [b]$. 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:

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

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:

$$\text{Bild } f[X] = \{y \in Y \mid \exists x \in X \text{ mit } f(x) = y\}.$$

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

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

$$\forall x, x' \in X \textit{ mit } \textit{ folgt aus } x \neq x' \textit{ auch } f(x) \neq f(x')$$

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

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

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

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

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 X, Y, Z Mengen und $f : X \to Y, x \to f(x)$ und $g : Y \to Z, y \to g(y)$ Abbildungen, dann definiert $g \circ f : X \to Z, x \to g(f(x))$ eine Abbildung.

Satz 1.3. *Die Komposition von Abbildungen ist assoziativ, d.h. für Abbildungen $f : X \to Y, x \mapsto f(x), g : Y \to Z, y \mapsto g(y)$ und $h : Z \to V, z \mapsto h(z)$ gilt $h \circ (g \circ f) = (h \circ g) \circ f$. Sei $f : X \to Y, x \mapsto f(x)$ eine Abbildung und seien $\mathbb{1}_X : X \to X, x \mapsto x$ 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 diese Abbildung f bijektiv, dann gilt auch

$$f \circ f^{-1} = \mathbb{1}_Y, f^{-1} \circ f = \mathbb{1}_X$$

Beweis:

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

1.6 Kategorien*

Eine Kategorie besteht aus 3 Dingen:

(i) Eine Klasse von Objekten.

(ii) Eine Klasse von Morphismen (Abbildungen). Zu jedem Morphismus gehören zwei Objekte aus der Klasse in (i) entsprechend dem Definitionsbereich und dem Wertebereich von Abbildungen. Wir bezeichnen die Morphismen auch als $f : A \to B$, wobei A der Definitionsbereich und B der Wertebereich ist.

(iii) Eine Verknüpfungsregel für Morphismen, die zwei Morphismen $f : A \to B$ und $g : B \to C$ einen dritten Morphismus $g \circ f : A \to C$ zuordnet.

Außerdem gibt es für jedes Objekt A den identischen Morphismus $\mathbb{1}_A$ und es gelten die analogen Aussagen zu dem vorangehenden Satz: Für drei Morphismen $f : A \to B, g : B \to C$ und $h : C \to D$ gilt das Assoziativgesetz:

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

und die Idempotenz

$$\mathbb{1}_B \circ f = f = f \circ \mathbb{1}_A.$$

Die Konzepte von injektiven und bijektiven Abbildungen besitzen verwandte Verallgemeinerungen für die Morphismen einer Kategorie.

Ein Morphismus $f : A \to B$ heißt Monomorphismus, wenn für alle Morphismen $g, h : C \to A$ gilt

$$f \circ g = f \circ h \Leftrightarrow g = h$$

Ein Morphismus $f : A \to B$ heißt Epimorphismen, wenn für alle Morphismen $g, h : B \to C$ gilt

$$g \circ f = h \circ f \Leftrightarrow g = h.$$

Beispiel 1.4. *Die Objekte der Mengen bilden zusammen mit den Abbildungen als Morphismen eine Kategorie. Dabei sind die injektiven Abbildungen Monomophismen und die surjektiven Abbildungen Epimorphismen. Sei nun* 1 *die einelementige Menge* $\{\emptyset\}$. *Dann lassen sich die Elemente einer beliebigen Menge* A *eindeutig den Morphismen* $1 \to A$ *zuordnen, weil das Bild jedes solchen Morphismuses genau aus einem Element der Menge* A *besteht. Dadurch lassen sich alle Aussagen der Mengenlehre in die Sprache der Kategorie übersetzen, ohne das Konzept des Elements einzuführen.*

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) $x + y = y + x$ *für alle* $x, y \in \mathbb{R}$

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

(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 jeder Zahl* $x \in \mathbb{R}$ *gibt es eine Zahl* $-x \in \mathbb{R}$*, so dass* $x + (-x) = 0$.

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) $x \cdot y = y \cdot x$ *für alle* $x, y \in \mathbb{R}$

(ii) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ *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 jeder Zahl* $x \in \mathbb{R} \setminus \{0\}$ *gibt es eine Zahl* $x^{-1} \in \mathbb{R}$*, mit* $x \cdot x^-1 = 1$.

A3. Distributivgesetz 2.3.

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

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:*

$$0+0=0 \quad 0+1=1 \qquad 0 \cdot 0 = 0 \quad 0 \cdot 1 = 0$$

$$1+0=1 \quad 1+1=0 \qquad 1 \cdot 0 = 0 \quad 1 \cdot 1 = 1$$

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:

$$\begin{aligned}
x, y \in \mathbb{R}: \quad x - y &= x + (-y) \\
x \in \mathbb{R} \setminus \{0\} \quad \frac{1}{x} &= x^{-1} \\
x \in \mathbb{R}, y \in \mathbb{R} \setminus \{y\} \quad \frac{x}{y} &= x \cdot y^{-1} \\
x, y \in \mathbb{R} \quad xy &= x \cdot y \\
xyz &= x \cdot (y \cdot z) \\
2x &= x + x = x \cdot (1+1) \\
-xy &= -(xy)
\end{aligned}$$

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

$$\begin{aligned} y &= y + 0 = 0 + y = (-x + x) + y \\ &= -x + (x + y) \\ &= -x + (x + z) \\ &= (-x + x) + z = 0 + z = z + 0 \\ &= z \end{aligned}$$

(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 (ii) $y = -x$.

(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

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

(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 $x^{-1}x = 1$. 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) $x \cdot y = 0 \Leftrightarrow x = 0$ *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) $(-x)(-y) = -(x \cdot (-y)) = -(-xy) = xy$.

(vi) $1 = (xy)^{-1}xy = ((xy)^{-1}x)y = y((xy)^{-1}x) \Rightarrow y^{-1} = (xy)^{-1}x \Rightarrow y^{-1}x^{-1} = (xy)^{-1}xx^{-1} = (xy)^{-1}$.

q.e.d.

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

(i) *Totalität der Ordnung: Für je zwei reelle Zahlen gilt genau eine der drei folgenden Relationen $x < y$ oder $x = y$ oder $y < x$.*

(ii) *Transivität: $x < y$ und $y < z \Rightarrow x < z$*

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

Wir benutzten folgende Abkürzungen:

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

Satz 2.13. *(Folgerungen aus A4)*

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

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

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

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

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

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

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 auch $-x < 0$.
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 $-ax < -ay \Leftrightarrow 0 < ax - ay \Leftrightarrow ay < ax$.

(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 $x^2 = (-x) \cdot (-x) > (-x) \cdot 0 = 0$.

(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*

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

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* $c < d \Rightarrow a + c < b + d$

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

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

(iv) $ab < 0 \Leftrightarrow$ *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

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

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

Aus der Definition folgt

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

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

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

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

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

Beweis:

(ii) Prüfe durch Fallunterscheidung

(iii) Zwei Fälle: $x + y > 0 \Rightarrow |x + y| = x + y \leq |x| + y \leq |x| + |y|$ wegen Monotonie.
$x + y < 0 \Rightarrow |x + y| = -x - y \leq |x| - y \leq |x| + |y|$ wegen Monotonie.

q.e.d.

Folgerung 2.19.

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

Beweis:

$$|x| = |x - y + y| \leq |x - y| + |y| \Rightarrow |x| - |y| \leq |x - y|$$

Vertausche x und $y \Rightarrow |y| - |x| \leq |x - y|$. Also gilt auch $||x| - |y|| \leq |x - y|$ **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* $d(x, y) = 0 \Leftrightarrow x = y$

(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 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 für alle $x \in M$ gilt $x \leq \beta$. β heißt dann obere Schranke von M.*

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

Warnung 2.25. *Das Supremum $\sup M$ bzw. Infimum $\inf M$ braucht nicht zu der Menge M zu gehören. Wenn $s = \sup M$ bzw. $t = \inf M$ zu der Menge M dazugehört, so ist s das Maximum von M bzw. t das Minimum von M (siehe (ii) und (iii) im Satz 2.29).*

Definition 2.26. *(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 $m = \min M$.*

Übungsaufgabe 2.27. *Zeige, dass jede endliche Teilmenge der reellen Zahlen ein Maximum und ein Minimum besitzt.*

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

Später werden wir sehen, dass das Axiom A5 nicht für die rationalen Zahlen $\mathbb{Q}$ gilt. Diese erfüllen aber alle anderen Axiome A1-A4.

Wir führen jetzt folgende Teilmenge der reellen Zahlen ein, die Intervalle genannt werden:

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

Offenbar ist b eine obere Schranke von $[a,b)$. Andererseits gibt es wegen Satz 2.15 für jedes $x < b$ ein $y = \max\{\frac{x+b}{2}, a\}$ 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 das Supremum von $[a,b)$.
Analog gilt:

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

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

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

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 gerade nicht symmetrisch ist, erhalten wir dadurch, dass wir in einer Aussage alle auftauchenden Ordnungsrelationen gerade 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 aber $x < y \Leftrightarrow -y < -x$, 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.

(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.

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

Also gibt es in M keine obere Schranke von M.

(iii) analog zu (ii)

q.e.d.

Also existiert

$$\max[a, b] = \max(a, b] = \max(-\infty, b] = b$$

$$\min[a, b] = \min[a, b) = \min[a, \infty) = a$$

während $[a, b)$ und (a, b) und $(-\infty, b)$	kein Maximum besitzen
und $(a, b]$ und (a, b) und (a, ∞)	kein Minimum.

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

[1]Wenn diese Aussagen aber algebraischen 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.

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:

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

$$x \cdot \infty \begin{cases} \infty \text{ für } x > 0 \\ \text{nicht definiert für } x = 0 \\ -\infty \text{ für } x < \infty. \end{cases}$$

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

Außerdem gilt

$$\begin{aligned} \infty + \infty &= \infty & \quad -\infty - \infty &= -\infty. \\ \infty \cdot \infty &= \infty & \quad \infty(-\infty) &= -\infty \quad (-\infty)(-\infty) = \infty. \end{aligned}$$

Nicht definiert ist

$$\infty - \infty, \quad \frac{\pm\infty}{\pm\infty}, \quad \frac{\pm\infty}{0}, \quad \frac{0}{0}, \quad 0 \cdot (\pm\infty).$$

$\bar{\mathbb{R}}$ ist also kein Körper. Die Definitionen von oberen und unteren Schranken, von Supremum und Infimum, und von Maximum und Infimum übertragen sich aber sofort 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 aber sofort, 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, das wegen (ii) in Satz 2.8 eindeutig ist. Aus (iv) im Satz 2.13 folgt dann mit $1 = 1 \cdot 1$, dass $1 > 0$. Dann folgt aus der Monotonie

$$\begin{array}{rcl} 1+1 & = & 2 > 1 \\ 2+1 & = & 3 > 2 \\ & \vdots & \\ n+1 & > & n \\ & \vdots & \end{array}$$

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

(i) $1 \in M$

(ii) $a \in M \Rightarrow a + 1 \in M$

$\mathbb{R}$ selber ist offenbar induktiv oder auch $[1, \infty)$.

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. *(Prinzip der vollständigen Induktion) Für die Menge $S \subset \mathbb{N}$ gelte*

(i) $1 \in S$

(ii) $a \in S \Rightarrow a + 1 \in S$.

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.**

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 $A(n + 1)$.

Einen solchen Beweis über eine Aussage über alle natürlichen Zahlen nennt man einen Beweis durch vollständige Induktion.

Satz 2.34. *(Bernoulli–Ungleichung) Sei $x > -1$, dann gilt*

$$(1+x)^n \geq 1 + nx \text{ für alle } n \in \mathbb{N}.$$

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

Beweis durch vollständige Induktion: Für $x = 0$ oder $n = 1$ gilt offenbar die Gleichheit. Für jede natürliche Zahl n sei also $A(n)$ die Aussage

$$\text{für alle } x > -1 \text{ mit } x \neq 0 \text{ gilt } (1+x)^{n+1} > 1 + (n+1)x.$$

Wenn $x \neq 0$ dann folgt $(1+x)^2 = 1 + 2x + x^2 > 1 + 2x$. Also gilt $A(1)$.

Es gelte $A(n)$. Dann folgen wir aufgrund der Monotonie:

$$\text{wegen } x > -1 \text{ gilt auch } 1 + x > 0$$

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

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

Satz 2.35. *Sei $n \in \mathbb{N}$. Dann gibt es keine Zahl in $\mathbb{N}$ zwischen $n - 1$ und n.*

Beweis durch vollständige Induktion: Offenbar ist $[1, \infty)$ eine induktive Menge. Also ist $\mathbb{N}$ eine Teilmenge von $[1, \infty)$ und enthält keine Zahl in $(0, 1)$. Wir nehmen jetzt an, dass es für ein beliebiges $n \in \mathbb{N}$ keine natürliche Zahl zwischen $n - 1$ und n gibt. Die Menge $M = \{1\} \cup \{x \in \mathbb{R} \mid x - 1 \in \mathbb{N}\}$ ist aber eine induktive Menge, weil $(1 + 1) - 1 = 1 \in \mathbb{N}$ und für jedes $x - 1 \in \mathbb{N}$ auch $(x + 1) - 1 = (x - 1) + 1 \in \mathbb{N}$. Also gilt $\mathbb{N} \subset \{1\} \cup \{x \in \mathbb{R} \mid x - 1 \in \mathbb{N}\}$. Weil aber $n \geq 1$ kann es dann auch keine Zahl in $\mathbb{N}$ zwischen n und $n + 1$ geben. **q.e.d.**

Satz 2.36. *(Wohlordnungsprinzip). Jede nichtleere Teilmenge M von $\mathbb{N}$ besitzt ein Minimum.*

Beweis: Weil $[1, \infty)$ eine induktive Menge ist, ist 1 eine untere Schranke von $\mathbb{N}$ und damit auch von M. Andererseits ist $\inf M + 1$ keine untere Schranke von M. Also gibt es ein Element $m \in M$ mit $\inf M \leq m < \inf M + 1$. Wegen dem vorangehenden Satz enthält dann aber $\mathbb{N}$ und damit auch M kein Element in dem Intervall $(m - 1, m)$. Weil aber $m - 1 < \inf M$, sind alle Elemente von M größer als $m - 1$. Also gibt es kein Element von M, das kleiner als m ist, und m ist das Minimum von M. **q.e.d.**

Satz 2.37. *(Archimedes-Endoxos)*

(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 muss es ein $n \in \mathbb{N}$ geben mit $n \in (\sup \mathbb{N} - 1, \sup \mathbb{N}]$. Dann gilt aber $n + 1 > \sup \mathbb{N} \geq n + 1$, was ein Widerspruch ist.
(ii): Nach (i) gibt es ein $n > \frac{1}{\epsilon} \Rightarrow \frac{1}{n} < \epsilon$. **q.e.d.**

Definition 2.38. *(ganze Zahlen $\mathbb{Z}$, rationale Zahlen $\mathbb{Q}$)*

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

Wegen dem Satz von Archimedes–Endoxos gibt es viele rationale Zahlen.

Satz 2.39. *Sei $a < b$. Dann existiert $r \in \mathbb{Q}$ mit $a < r < b$.*

Beweis: $a < \frac{m}{n} < b \Leftrightarrow na < m < nb$. Weil $b - a > 0$ gibt es nach dem Satz von Archimedes-Endoxos ein $n \in \mathbb{N}$ mit $n > \frac{1}{b-a}$. Dann gilt $na < nb$ und $nb - na > 1$.

Wir nehmen zunächst an, das $a \geq 0$ ist. Sei m die kleinste natürliche Zahl, die größer ist als na. Die natürlichen Zahlen sind enthalten in $\{1\} \cup \{x \in \mathbb{R} \mid x - 1 \in \mathbb{N}\}$, weil diese Menge induktiv ist. Also ist $m - 1$ entweder eine natütliche Zahl oder gleich Null. Dann folgt aber $m - 1 \leq na$ und damit auch

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

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

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

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

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 deshalb, dass $\mathbb{Q}$ dicht in den reellen Zahlen liegt. Die rationalen Zahlen erfüllen als Unterkörper der reellen Zahlen die Axiome A1-A4. Wir werden gleich aber sehen, dass sie nicht das Vollständigkeitsaxiom erfüllen.

2.4 Wurzeln und Intervallschachtelung

Satz 2.40. *(Quadratwurzeln) Für alle $a > 0$ gibt es genau ein $b > 0$, so dass $b^2 = a$.*

Wir schreiben $b = \sqrt{a} = a^{\frac{1}{2}}$.
Beweis: Eindeutigkeit: Aus $0 < b_1 < b_2$ folgt aufgrund der Monotonie $b_1^2 < b_1 b_2 < b_2^2$. Also gibt es höchstens ein $b > 0$ mit $b^2 = a$.

Existenz: Die Menge $M = \{x \in \mathbb{R}_0^+ \mid x^2 < a\}$ enthält 0 und ist nach oben beschränkt, weil aus $x > a + 1$ folgt

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

Sei $b = \sup M$.

Wenn $b^2 < a$ gilt, gibt es ein $n \in \mathbb{N}$ mit $n > \frac{2b+1}{a-b^2}$. Dann folgt aber

$$\left(b + \frac{1}{n}\right)^2 \le b^2 + \frac{2b}{n} + \frac{1}{n^2} \le b^2 + \frac{2b+1}{n} < a.$$

Also ist $b + \frac{1}{n} \in M$ und $b + \frac{1}{n} \le b$ Widerspruch.

Wenn $b^2 > a$ gilt, gibt es ein $n \in \mathbb{N}$ mit $n > \max\left\{\frac{2b}{b^2-a}, \frac{1}{b}\right\}$. Dann folgt aber

$$\left(b - \frac{1}{n}\right)^2 \ge b^2 - \frac{2b}{n} + \frac{1}{n^2} > b^2 - \frac{2b}{n} > a.$$

Jedes $x > b - \frac{1}{n} > 0$ erfüllt dann aber

$$x^2 > x\left(b - \frac{1}{n}\right) > \left(b - \frac{1}{n}\right)^2 > 0.$$

Also ist $b - \frac{1}{n}$ eine obere Schranke von M. Widerspruch. Also gilt $b^2 = a$. **q.e.d.**

Wir werden in Anwendung 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 2.41. $\sqrt{2} \notin \mathbb{Q}$

Beweis: Hier benutzen wir (ohne Beweis) die

Primfaktorzerlegung 2.42. *Jede Zahl in $\mathbb{N} \setminus \{1\}$ läßt sich in bis auf Permutation der Primfaktoren eindeutiger Weise in ein endliches Produkt von Primzahlen zerlegen (Primzahlen sind Zahlen in $\mathbb{N} \setminus \{1\}$, die nur durch 1 und sich selber teilbar sind).*

Seien $m, n \in \mathbb{N}$ mit $\left(\frac{m}{n}\right)^2 = 2$ oder $m^2 = 2n^2$. Teile beide Zahlen m und n solange durch 2, bis eine von beiden ungerade ist, also die entsprechende Primfaktorzerlegung keine 2 enthält: $m = 2^l\tilde{m}$ und $n = 2^l\tilde{n}$. Danach gilt immer noch $2^{2l}\tilde{m}^2 = 2^{2l+1}\tilde{n}^2 \Leftrightarrow \tilde{m}^2 = 2\tilde{n}^2$. Also ist $\tilde{m}$ gerade und die Primfaktorzerlegung von $\tilde{m}$ enthält mindestends eine 2. Dann gilt aber $4k^2 = 2\tilde{n}^2 \Leftrightarrow 2k^2 = \tilde{n}^2$. Also ist auch $\tilde{n}$ gerade. Widerspruch. **q.e.d.**

Satz 2.43. *(Intervallschatelungsprinzip) 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 $\bigcap\limits_{n\geq 1} I_n = \{x\}$ der Durchschnitt genau ein $x \in \mathbb{R}$.

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

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 $x, y \in \bigcap\limits_{n>1} I_n$. Es gilt sogar $\bigcap\limits_{n>1} I_n = [x, y]$. Andererseits ist für alle $\epsilon > 0$ auch $y - x < \epsilon$. Dann muss aber $0 \leq y - x \leq \inf\{\epsilon | \epsilon > 0\} = 0$ und deshalb $y = x$ gelten. **q.e.d.**

Es gilt auch die Umkehrung, dass aus dem Intervallschachtelungsprinzip und den Axiomen A1-A4 das Vollständigkeitsaxiom folgt (Übungsaufgabe). Deshalb können die reellen Zahlen auch durch die Axiome A1-A4 und das Intervallschachtelungsprinzip charakterisiert werden. Wir werden später noch andere äquivalente Aussagen angeben.

2.5 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. Zwei Mengen mit endlich vielen Elementen sind offenbar genau dann gleichmächtig, wenn sie die gleiche Anzahl an Elementen haben. Also werden die Äquivalenzklassen der endlichen Mengen durch die Elemente von $\mathbb{N}_0$ beschrieben. Deshalb kann man alle diese Äquivalenzklassen auch als eine Erweiterung von $\mathbb{N}_0$ betrachten. Dabei werden manchmal besonders einfache Repräsentanten in den Äquivalenzklassen zur Beschreibung der natürlichen Zahlen (einschließlich der Null) benutzt:

$$\begin{array}{rcccl} 0 & & \emptyset & & \\ 1 & \{ & \emptyset & \} & \\ 2 & \{\{ & \emptyset & \}\} & \\ 3 & \{\{\{ & \emptyset & \}\}\} & \end{array}$$

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 Teilemenge 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 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 $\mathbb{N} \to \{1, 2, \cdots, n\}, m \mapsto \min\{m, n\}$ ist eine surjektive Abbildung, so dass dann auch eine surjektive Abbildung auf A existiert. Wenn A abzählbar ist, gibt es sogar 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 $g : A \to \mathbb{N}, a \to \min f^{-1}[\{a\}]$, 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.

Satz 2.47. (i) *Die Menge* $\mathbb{N} \times \mathbb{N}$ *ist abzählbar.*

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

Beweis:

(i) Wir definieren eine injektive Abbildung f von $\mathbb{N} \times \mathbb{N}$ nach $\mathbb{N}$ durch

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

(Diagonalnummerierung)

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

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

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

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

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

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

Korollar 2.48. $\mathbb{Q}$ *ist abzählbar.*

Beweis: $\mathbb{Z} \times \mathbb{N} \to \mathbb{Q}, (m,n) \mapsto \frac{m}{n}$ 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 2.49. *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 die Intervalle $(I_n)_{n\in\mathbb{N}}$: Wähle für all $n \in \mathbb{N}$ ein abgeschlossenes Teilintervall I_{n+1} von I_n, das nur ein Drittel so groß ist wie I_n und x_n nicht enthält. Dann liegen für alle $n \in \mathbb{N}, x_n$ nicht in den Intervallen I_n. Wegen dem Intervallschachtelungsprinzip ist aber $\bigcap\limits_{n\in\mathbb{N}} I_n$ nicht leer. Also gibt es eine reelle Zahl, die nicht zu der Menge $\{x_n \mid n \in \mathbb{N}\}$ gehört. Widerspruch. **q.e.d.**

Auch die Potenzmenge der natürlichen Zahlen ist nicht abzählbar. Wir können sie identifizieren mit den Folgen, 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 den reellen Zahlen $[0,1]$ identifizieren. Diese sind wiederum gleichmächtig zu den reellen Zahlen.

2.6 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 2.50. *(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 und damit auch die Folgerungen daraus. Dabei ist

$$\begin{aligned} \text{Null} :0_{\mathbb{C}} &= (0,0) \\ \text{Eins} :1_{\mathbb{C}} &= (1,0) \end{aligned}$$

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

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:

$$z + (-z) = z - z = 0$$

$$z^{-1} = \frac{1}{z}. \qquad z \cdot \frac{1}{z} = 1. \qquad \text{usw.}$$

Da ja die komplexen Zahlen eine Erweiterung der reellen Zahlen sein sollen, müssen wir die reellen Zahlen als Teilmenge der komplexen Zahlen auffassen können. Wir definieren also eine injektive Abbildung

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

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

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

Definition 2.51. *(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 2.52. *(komplexe Konjugation) Die Abbildung* $\mathbb{C} \to \mathbb{C}, z = x + \imath y \mapsto \bar{z} = x - \imath y$ *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 2.53. **(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) $(x + \imath y)(x - \imath y) = x^2 + y^2 \geq 0$

Definition 2.54. *(Betrag) Der Betrag einer komplexen Zahl* $z = x + \imath y$ *ist definiert als die reelle nicht negative Wurzel aus* $z \cdot \bar{z} : |\cdot| \mathbb{C} \to \mathbb{R}, z \mapsto |z| = \sqrt{z \cdot \bar{z}}$.

Satz 2.55. *(Eigenschaften des Betrags)*

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

(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 $|\Im(z)| \leq |z|$.

(iii)

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

Aus $|z|+|w|<|z+w|$ folgt aber 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 2.56. *(Abstand) Der Abstand zweier komplexer Zahlen ist die nicht negative Zahl* $d : \mathbb{C}\times\mathbb{C} \to \mathbb{R}, (z,w) \mapsto d(z,w) = |z-w|$

Aus den Eigenschaften des Betrages folgt genau wie im Reellen.

Satz 2.57. *(Eigenschaften des Abstandes)*

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

(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 dann 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 reelen Zahlen als auch für die komplexen Zahlen gelten. Wir benutzen dann das Symbol $\mathbb{K}$ um entweder die reelen 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, so dass es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle natürlichen Zahlen $n \geq N$ gilt $|a_n - a| < \epsilon$. Die Zahl a heißt dann Grenzwert der Folge $(a_n)_{n\in\mathbb{N}}$*

Wir schreiben dann $\lim\limits_{n\to\infty} a_n = a$ 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-Endoxos gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ mit $\frac{1}{N} < \epsilon$. Dann gilt aber 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 Folgen der entsprechenden Realteile und Imaginärteile konvergieren.*

(iii) *Jede konvergente Zahlenfolge ist beschränkt, als Teilmenge von $\mathbb{K}$*

Beweis:

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

$$0 \leq |a - b| \leq |a - a_n| + |a_n - b| < \epsilon.$$

Dann gilt aber 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 der Ungleichung

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

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 für alle $n \geq M$ $|y_n - y| < \frac{\epsilon}{2}$. 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 für alle $n \geq N$ gilt $|a_n - a| < 1$. Dann gilt aber auch für alle $n \geq N$

$$|a_n| \leq |a_n - a + a| < 1 + |a|.$$

Daraus folgt dann für alle $n \in \mathbb{N}$

$$|a_n| \leq \max\{|a| + 1, |a_1|, |a_2|, \cdots, |a_{N-1}|\}.$$

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 $\lim\limits_{n\to\infty} a_n = a$. 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 $\lim\limits_{n\to\infty} a_n = \infty$ bzw. $\lim\limits_{n\to\infty} a_n = -\infty$ Weil in beiden Fällen die Folgen nicht beschränkt sind können sie wegen dem vorangehenden Satz nicht konvergieren. Es gilt offenbar $\lim\limits_{n\to\infty} n = \infty$ und $\lim\limits_{n\to\infty} -n = -\infty$.

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

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

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

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

(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 $\lim\limits_{n\to\infty} x^n = 0$. 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 aber $|x|^n < \frac{1}{n}\frac{|x|}{1-|x|}$. Aus dem Satz von Archimedes-Endoxos 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 aber 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 $\lim\limits_{n\to\infty} 1^n = 1$.

(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 aber für jedes $b > 0$ ein $N \in \mathbb{N}$ mit $N > \frac{b}{x-1}$. Dann folgt aber für alle $n \geq N$:

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

Also gilt $\lim\limits_{n\to\infty} x^n = \infty$.

(ii) Für $|x| > 1$ gilt wegen (iv) $\lim\limits_{n\to\infty} |x^n| = \infty$. Also ist die Folge $(x^n)_{n\in\mathbb{N}}$ 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

$$\max\{|y - x^n|, |y - x^{n+1}|\} \geq \frac{|x-1|}{2}.$$

Also kann es für $x \neq 1$ keinen Grenzwert geben.

q.e.d.

Satz 3.5. *(Rechenregeln) Seien $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ zwei konvergente Zahlenfolgen. Dann 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 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 für alle $n \in \mathbb{N}$ $x_n \leq z_n \leq y_n$ erfüllt,* $\lim\limits_{n\to\infty} z_n = \lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} y_n$.

Beweis:

(i) Sei $x = \lim\limits_{n\to\infty} x_n$ 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 für alle $n \geq N |x_n - x| < \frac{\epsilon}{2}$ und für alle $n \geq M |y_n - y| < \frac{\epsilon}{2}$ gilt. Dann gilt aber für alle $n \geq \max\{N, M\}$
$|x_n + y_n - (x + y)| \leq |x_n - x| + |y_n - y| < \epsilon$. Also konvergiert $(x_n + y_n)_{n\in\mathbb{N}}$ gegen $x + y$.

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

$$|\lambda x_n - \lambda x| \leq |\lambda||x_n - x| < \epsilon$$

(iii) Weil $(x_n)_{n\in\mathbb{N}}$ konvergent ist, 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 für $n \geq N$ gilt $|x_n - x| < \frac{\epsilon}{|y|+1}$ gilt und für alle $n \geq M$ $|y_n - y| < \frac{\epsilon}{2\lambda}$. Dann gilt aber für alle $n \geq \max\{N, M\}$ auch

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

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

$$|x_n| \geq |x| - |x_n - x| > \frac{|x|}{2} \text{ und } \frac{1}{|x_n|} < \frac{2}{|x|}$$

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

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

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

(vi) Für alle $\epsilon > 0$ gibt es zwei natürliche Zahlen N und M, so dass für alle $n \geq N$ gilt $|x_n - x| < \frac{\epsilon}{2}$ und für alle $n \geq M$ $|y_n - y| < \frac{\epsilon}{2}$. 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) Für alle $\epsilon > 0$ gibt es zwei natürliche Zahlen $N, M \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $|x_n - x| < \epsilon$ und für alle $n \geq M$ $|y_n - x| < \epsilon$. Für alle $n \geq \max\{N, M\}$ gilt dann aber

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

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 für $x \neq 1$ und alle $n \in \mathbb{N}_0$:

$$1 + x + \ldots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

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

Satz 3.6.

$$\lim_{n\to\infty} \sum_{m=0}^{n} x^m = \frac{1}{1 - x} \textit{ für } |x| < 1.$$

3.2 Konvergenzprinzipien

Wir wollen 3 Methoden behandeln 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. *(Monotonie Prinzip)*

(i) *Eine monoton wachsende (fallende) beschränkte reelle Folge $(a_n)_{n\in\mathbb{N}}$ ist konvergent. Es gilt*

$$\lim_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} a_n \text{ bzw. } \lim_{n\to\infty} a_n = \inf_{n\in\mathbb{N}} a_n$$

(ii) *Für eine monoton wachsende (fallende) unbeschränkte reelle Folge gilt*

$$\lim_{n\to\infty} a_n = \infty \text{ bzw. } \lim_{n\to\infty} a_n = -\infty.$$

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.**

Anwendung 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*

$$a_0 = 1 + \frac{a-1}{k} \qquad\qquad a_{n+1} = a_n \left(1 + \frac{a - a_n^k}{k \cdot a_n^k}\right) \text{ für alle } n \in \mathbb{N}_0.$$

Für $a = 1$ ist dann $a_n = 1$ für alle $n \in \mathbb{N}$. Für $a \neq 1$ gilt für alle $n \in \mathbb{N}$

$$0 < a_n < a_0, \qquad\qquad a_n < a_{n-1} \text{ und } \qquad\qquad a < a_n^k.$$

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 dann $a_0^k = \left(1 + \frac{a-1}{k}\right)^k > 1 + a - 1 = a$. Daraus folgt aber $-1 < -\frac{1}{k} < \frac{a-a_0^k}{ka_0^k} < 0$. Also gilt auch $0 < a_1 < a_0$ und wegen 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 für $n \in \mathbb{N}$: $0 < a_n < a_0, a_n < a_{n-1}$ und $a < a_n^k$. Dann folgt $-1 < -\frac{1}{k} < \frac{a-a_n^k}{k\cdot a_n^k} < 0$. Daraus folgt aber: $0 < a_{n+1} < a_n$ 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 Monotonie Prinzip konvergiert sie dann. Sei $b = \lim\limits_{n\to\infty} a_n$. 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.*

(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 $(a^r)^{qn} = a^{pn}$ für alle $n \in \mathbb{N}$. Wir zeigen jetzt, dass $0 < a < b$ äquivalent ist zu $0 < a^n < b^n$ mit $n \in \mathbb{N}$. 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 $y^q = a^p$. Für $r < 0$ ist $a^r = \frac{1}{a^{-r}}$ die entsprechende positive Zahl.

(ii) Sei $r = \frac{p}{q} > 0$. Dann ist $0 < a < b$ äquivalent zu $0 < a^p < b^p$ und das wiederum äquivalent zu $0 < a^{\frac{p}{q}} < b^{\frac{p}{q}}$.

(iii) Sei $r = \frac{-p}{q} < 0$. Dann folgt (iii) aus (ii) wegen (vi) Satz 2.13

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 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 Zahlenfolge enthält eine monotone Teilfolge.*

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ eine reelle Folge und A die Menge $A = \{n \in \mathbb{N} | a_n \geq a_m$ für alle natürliche Zahlen $m > n\}$. Wenn A eine unendliche Menge ist, dann sei $n_1 < n_2 < \ldots$ eine Abzählung der Element von A. Die Teilfolge $b_m = a_{n_m}$ ist dann monoton fallend. Wenn A eine endliche Menge ist besitzt es 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 dann 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 Auswahlrpinzip von Bolzano–Weierstraß besitzt dann die Folge der Realteile eine konvergente Teilfolge. Die entsprechende Teilfoge der Imaginärteilte besitzt dann wieder wegen dem Auswahlprinzip eine konvergente Teilfolge. Wegen Satz 3.3 (ii) konvergiert die entsprechende komplexe Teilfolge. **q.e.d.**

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

Satz 3.16. *(Kriterium von Cauchy) Eine Zahlenfolge konvergiert genau dann, wenn sie eine Cauchy–Folge 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 $|a_m - \lim\limits_{n\to\infty} a_n| < \frac{\epsilon}{2}$ erfüllen. Also gilt auch für alle $m, l \geq N$

$$|a_m - a_l| \leq |a_m - \lim_{n\to\infty} a_n| + |\lim_{n\to\infty} a_n - a_l| < \epsilon.$$

Sei umgekehrt $(a_n)_{n\in\mathbb{N}}$ eine Cauchy–Folge. 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 $a = \lim\limits_{n\to\infty} b_n$. 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 aber $(b_n)_{n\in\mathbb{N}}$ eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$ ist, folgt für alle $n \geq \max\{N, M\}$ dann $|a_n - a| \leq |a_n - b_n| + |b_n - a| < \epsilon$. **q.e.d.**

Es gilt sogar unter der Annahme der Axiome A1-A4, dass folgende Aussagen äquivalent sind:

Vollständigkeitsaxiom A5.

(i) aus dem Monotonieprinzip.

Auswahlprinzip von Bolzano–Weierstraß.

Jede Cauchy–Folge konvergiert.

Intervallschachtelungsprinzip.

Wir können die reellen Zahlen also auch auffassen als folgende Äquivalenzklasse von Cauchy–Folgen von rationalen Zahlen:
Zwei Cauchy–Folgen heißen äquivalent, 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 Häufungspunkte einer reellen Folge nicht leer und nach oben (unten) beschränkt, so besitzt sie ein Maximum (Minimum).*

Beweis: Wir nehmen an, dass die Menge der Häufungspunkte der Folge $(a_n)_{n\in\mathbb{N}}$ nach oben beschränkt ist. Sei a 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 für alle $m \in \mathbb{N}$ $|c_m - b_m| < \frac{1}{2m}$ 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 $\overline{\lim} a_n = \lim\limits_{n\to\infty} \sup a_n$ bzw. $\underline{\lim} a_n = \lim\limits_{n\to\infty} \inf a_n$.*

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 ein untere Schranke von mindestens einen Häufungspunkt. Also ist die Charkerisierung der Zahl $\overline{a}$ in (i) äquivalent dazu, dass alle Zahlen, die größer sind als $\overline{a}$ auch obere Schranken der Häufungspunkte von $(a_n)_{n\in\mathbb{N}}$ sind, aber $\overline{a}$ selber ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$ ist. Dann ist $\overline{a}$ aber der maximale Häufungspunkt. **q.e.d.**

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

Weil $\overline{\lim} a_n$ und $\underline{\lim} a_n$ das Maximum und das Minimum der Häufungspunkte sind, ist die Bedingung $\overline{\lim} a_n = \underline{\lim} a_n$ 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 auch $|a_n - a| < \epsilon$ und $(a_n)_{n\in\mathbb{N}}$ konvergiert gegen a.

Wenn umgekehrt $(a_n)_{n\in\mathbb{N}}$ konvergiert, dann konvergiert auch jede Teilfolge gegen $a = \lim\limits_{n\to\infty} a_n$. 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 (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 $\overline{\lim} a_n$ die gegen $\overline{\lim} a_n$) konvergiert. Dann gibt es eine Teilfolge $(c_n)_{n\in\mathbb{N}}$ von $(c_n)_{n\in\mathbb{N}}$ (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 $\lim\limits_{n\to\infty} c_n$ ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$ (bzw. $\lim\limits_{n\to\infty} d_n$ ein Häfungspunkt von $(b_n)_{n\in\mathbb{N}}$). Also gilt $\underline{\lim} a_n \leq \lim\limits_{n\to\infty} c_n \leq \lim\limits_{n\to\infty} d_n \leq \underline{\lim} b_n$. **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 $n \geq N$

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

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–Endoxos, 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*

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

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

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

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

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

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

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

Dann folgt aber für alle $n \geq 2$:

$$n \geq 1 + \frac{n(n-1)}{2} y_n^2 \Leftrightarrow y_n^2 \leq \frac{2}{n} \Leftrightarrow y_n \leq \sqrt{\frac{2}{n}}.$$

Also ist y_n eine Nullfolge. **q.e.d.**

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

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 $\frac{x}{\sqrt[n]{n!}} < 1 \Leftrightarrow x < \sqrt[n]{n!}$. Also folgt $\lim\limits_{n\to\infty} \sqrt[n]{n!} = \infty$. **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 eine Häufungspunkt hat, dann hat die Folge $\frac{a_{n+1}}{a_n}$ auch einen Häufungspunkt. Und wenn gilt $\overline{\lim} \left(\frac{a_{n+1}}{a_n}\right) < \infty$ dann gilt auch $\overline{\lim} \sqrt[n]{a_n} < \infty$.

Beweis: Wenn $\underline{\lim} \left(\frac{a_{n+1}}{a_n}\right) = 0$ ist, ist die Aussage trivial. Wir nehmen also an, dass es ein $a > 0$ gibt, so dass 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 $n > N$

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

Wegen (iii) gilt dann aber $\underline{\lim} \sqrt[n]{a_n} \geq (a-\epsilon)$. Weil dies aber 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 Häufungspunnkt von $\frac{a_{n+1}}{a_n}$. 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 aber für jede positive Folge $(b_n)_{n\in\mathbb{N}}$ $\left(\underline{\lim} b_n^{-1}\right)^{-1} = \overline{\lim} b_n$. Also folgt die zweite Ungleichung aus Satz 2.13 (vi). **q.e.d.**

Die Folge $\left(\sum\limits_{k=0}^{n} \frac{1}{k!}\right)_{n\in\mathbb{N}}$ ist streng monoton wachsend und beschränkt, weil 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 $e \in (\frac{5}{2}, 3)$.

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

Beweis: Wegen der Binomischen Formel gilt

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

Also gilt $\left(1 + \frac{1}{n}\right)^n \le \sum\limits_{k=0}^{n} \frac{1}{k!} \le e$. Offenbar ist aber für alle $k \in \mathbb{N}$

$$\lim_{n\to\infty} \binom{n}{k} \frac{1}{n^k} = \frac{1}{k!}.$$

Also gilt auch für alle $m \in \mathbb{N}$

$$\underline{\lim} \left(1 + \frac{1}{n}\right)^n \ge \sum_{k=0}^{m} \frac{1}{k!}.$$

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

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

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ die Folge $\frac{n^n}{n!}$. Dann gilt $\frac{a_{n+1}}{a_n} = \frac{(n+1)^{n+1}}{(n+1)n^n} = \left(1 + \frac{1}{n}\right)^n$. Es gilt aber $\lim\limits_{n\to\infty} \frac{a_{n+1}}{a_n} = e$ und $\lim\limits_{n\to\infty} \frac{a_n}{a_{n+1}} = \frac{1}{e}$. Dann folgt aus dem vorangehenden Satz $e \le \underline{\lim} \frac{n}{\sqrt[n]{n!}}$ und $\frac{1}{e} \le \underline{\lim} \frac{\sqrt[n]{n!}}{n}$. Daraus folgt dann: $e \le \underline{\lim} \frac{n}{\sqrt[n]{n!}} \le \overline{\lim} \frac{n}{\sqrt[n]{n!}} \le e$ und dann auch $\lim\limits_{n\to\infty} \frac{n}{\sqrt[n]{n!}} \le e$. **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* $\left(\sum a_n\right)_{n\in\mathbb{N}}$.

Wenn die Reihe $\left(\sum a_n\right)_{n\in\mathbb{N}}$ 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* $\left(\sum q^n\right)_{n\in\mathbb{N}_0}$. *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*

$$\text{Für } |q| < 1 \text{ ist } \left(\sum q^n\right)_{n\in\mathbb{N}_0} \text{ konvergent:} \qquad \sum_{n=0}^{\infty} q^n = \frac{1}{1-q}.$$

$$\text{Für } |q| \geq 1 \text{ ist } \left(\sum q^n\right)_{n\in\mathbb{N}_0} \text{ divergent.}$$

$$\text{Für reelles } q \geq 1 \text{ ist } \left(\sum q^n\right)_{n\in\mathbb{N}_0} \text{ divergent:} \qquad \sum_{n=0}^{\infty} q^n = \infty.$$

(ii) *Zeta Funktion* $\zeta(s) = \sum\limits_{n=1}^{\infty} \frac{1}{n^s}$ *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 $\frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{n+n} \geq \frac{n}{2n} \geq \frac{1}{2}$. Also sind für alle $n \in \mathbb{N}_0$ jeweils die Summen $\sum_{j=1}^{2^n} \frac{1}{j} = 1 + \sum_{m=1}^{n} \sum_{j=2^{m-1}+1}^{2^m} \frac{1}{j} \geq 1 + \frac{n}{2}$. Dann gilt aber $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$. **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 $\sum_{n=1}^{\infty} \frac{1}{n(n+1)\cdots(n+k)} = \frac{1}{k \cdot k!}$

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 $\frac{1}{(m+1)\cdots(m+k)} \leq \frac{1}{m^k}$ 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 für alle $m \geq n \geq N$ gilt* $\left|\sum_{j=n}^{m} a_j\right| < \epsilon$.

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* $\sum_{n=1}^{\infty} a_n = \sup_{m\in\mathbb{N}} \sum_{n=1}^{m} a_n$. **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* $\left|\sum_{n=1}^{\infty} a_n\right| \leq \sum_{n=1}^{m} |a_n|$.

Beweis: Aus der Dreieckungleichung folgt für alle $m \geq n$ $\left|\sum_{j=n}^{m} a_j\right| \leq \sum_{j=n}^{m} |a_j|$. 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 für alle* $n \in \mathbb{N}$ $b_n \leq a_n$.

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

(ii) *Wenn außerdem* $\left(\sum b_n\right)_{n\in\mathbb{N}}$ *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* $\left(\sum \frac{1}{n\cdots(n+k)}\right)$ *auch die Konvergenz von* $\left(\sum \frac{1}{n^{k+1}}\right)_{n\in\mathbb{N}}$.

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* $\left(\sum a_n\right)_{n\in\mathbb{N}}$ *absolut.*

(ii) *Falls* $\alpha > 1$, *dann divergiert* $\left(\sum a_n\right)_{n\in\mathbb{N}}$.

Im Fall $\overline{\lim} \sqrt[n]{|a_n|} = 1$ kann die Reihe $\left(\sum a_n\right)_{n\in\mathbb{N}}$ 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 $\sqrt[n]{|a_n|} \leq \beta \Leftrightarrow |a_n| \leq \beta^n$. Weil aber $\left(\sum \beta^n\right)_{n\in\mathbb{N}}$ konvergiert ist dann auch $\left(\sum a_n\right)_{n\in\mathbb{N}}$ 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 ist aber die Reihe $(\sum a_n)_{n\in\mathbb{N}}$ keine Cauchyfolge, also divergent. **q.e.d.**

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

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

Aufgrund des Beispiels (v) im letzten Kapitel ist $\lim\limits_{n\to\infty} \sqrt[n]{n!} = \infty$. 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 der Ungleichung $\overline{\lim} \sqrt[n]{|a_n|} \leq \overline{\lim} \left|\frac{a_n + 1}{a_n}\right|$ folgt (i) aus dem Wurzeltext.

(ii) Aus der Bedingung $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ folgt $|a_{n+1}| = \left|\frac{a_{n+1}}{a_n}\right| \cdot \left|\frac{a_n}{a_{n-1}}\right| \cdots \left|\frac{a_{N+1}}{a_N}\right| \cdot |a_N| \geq |a_N| > 0$. Also ist $|a_n|$ keine Nullfolge und deshalb ist wieder die Reihe $(\sum a_n)_{n\in\mathbb{N}}$ keine Cauchyfolge. **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 $j \in \mathbb{N}$:

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

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 $s > 1$. $\Rightarrow$ 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 $\left(\sum(-1)^n a_n\right)_{n\in\mathbb{N}_0}$*

Beweis: Sei für alle $n \in \mathbb{N}_0$ $s_n = \sum\limits_{m=0}^{n}(-1)^m a_n$. 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 $(s_{2n+1})_{n\in\mathbb{N}}$ monoton wachsend und beschränkt und die Folge $(s_{2n})_{n\in\mathbb{N}_0}$ 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 $\sum\limits_{n=0}^{\infty}\frac{(-1)^n}{n+1} = -\sum\limits_{n=1}^{\infty}\frac{(-1)^n}{n}$ konvergent, während sie nicht absolut konvergiert, weil $\sum\limits_{n=1}^{\infty}\frac{1}{n} = \infty$.

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 $M = \{(z_n)_{n\in\mathbb{N}} | (z_n)_{n\in\mathbb{N}} \text{ konvergiert nicht gegen } p-1\}$. 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 $n \in \mathbb{N}$

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

gilt. Dann folgt aber auch $0 \leq x - \sum\limits_{m=1}^{n} \frac{z_m}{p^m} < \frac{1}{p^n}$. 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

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

muss auch $|z_n - w_n| \leq 1$ gelten. Sei also $z_n = w_n + 1$, dann muss für alle $m > n$ gelten $w_m - z_m = p - 1 \Rightarrow z_m = 0$ und $w_m = p - 1 \Rightarrow \lim\limits_{m\to\infty} w_m = p-1$. 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*

$$\left(\sum (a_n + b_n)\right)_{n\in\mathbb{N}} \quad \textit{und} \qquad \left(\sum \lambda a_n\right)_{n\in\mathbb{N}} \textit{ für alle } \lambda \in \mathbb{K}.$$ **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 $(\sum c_n)_{n\in\mathbb{N}}$ mit $c_n = \sum\limits_{k=0}^{n} a_k b_{n-k}$ für alle $n \in \mathbb{N}$ das (Chauchy-)Produkt der beiden Reihen.*

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

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

Beweis: Für alle $n \in \mathbb{N}$ 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

$$\begin{aligned} C_n &= a_0b_0 + (a_0b_1 + a_1b_0) + \ldots + (a_0b_n + \ldots + a_nb_0) \\ &= a_0B_n + a_1B_{n-1} + \ldots + a_nB_0 \\ &= a_0(B - \beta_n) + a_1(B - \beta_{n-1}) + \ldots + a_n(B - \beta_0) \end{aligned}$$

Hierbei ist $B = \lim\limits_{n\to\infty} B_n = \sum\limits_{n=0}^{\infty} b_n$ 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 $n \in \mathbb{N}_0$

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

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

$$\begin{aligned} |a_0\beta_n + \ldots + a_n\beta_0| &\leq |\beta_0a_n + \ldots + \beta_{N-1}a_{n-N+1}| + |\beta_Na_{n-N} + \ldots + \beta_na_0| \\ &< N\beta \cdot \frac{\epsilon}{2N\beta} + \frac{\epsilon}{2\alpha}\sum_{k=0}^{n-N} |a_k| \leq \epsilon. \end{aligned}$$

Also konvergiert das Produkt der Reihen $(\sum a_n)_{n\in\mathbb{N}}$ und $(\sum b_n)_{n\in\mathbb{N}}$ wenn beide konvergieren und eine absolut konvergiert. Wenn beide Reihen absolut konvergieren, dann konvergiert auch das Produkt der Reihen $(\sum |a_n|)_{n\in\mathbb{N}}$ und $(\sum |b_n|)_{n\in\mathbb{N}}$. Also ist 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 $\sum_{k=0}^{n} \frac{1}{\sqrt{k+1}\sqrt{n-k+1}} \geq \sum_{k=0}^{n} \frac{1}{\sqrt{n+1}\sqrt{n+1}} = 1$. *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* $x \in \mathbb{R}$ *ist* $\exp(x) > 0$.

(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* $r \in \mathbb{Q}$: $\exp(r) = \exp(r \cdot 1) = e^r$.

Beweis:

(i) Wir hatten schon gesehen, dass die Reihe $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ 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 aber

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

Dann folgt $\exp(x)\exp(y) = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} = \exp(x+y)$.

(ii) Wegen (i) gilt für alle $x \in \mathbb{R}$ $\exp(x) = \left(\exp\left(\frac{x}{2}\right)\right)^2 \geq 0$. Wenn für ein $x \in \mathbb{R}$ gilt $\exp(x) = 0$, dann folgt $1 = \exp(0) = \exp(-x) \cdot \exp(x) = 0$. Widerspruch.

(iii) Offenbar ist für alle $n \in \mathbb{N}, m \in \mathbb{Z}$ $\left[\exp\left(\frac{x \cdot m}{n}\right)\right]^n = \exp(x \cdot m) = [\exp(x)]^m$. 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 $\left(\sum a_n\right)_{n\in\mathbb{N}}$.*

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 $\left(\sum a_n\right)_{n\in\mathbb{N}}$ 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 $\left(\sum a_n\right)_{n\in\mathbb{N}}$ 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 aber 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 aber auch $\left(\sum a_{\tau(n)}\right)_{n\in\mathbb{N}}$. Mit denselben N und M in Abhängigkeit von $\epsilon > 0$ gilt dann aber auch

$$\left|\sum_{k=1}^{M} a_{\tau(k)} - \sum_{k=1}^{N} a_k\right| = \left|\sum_{k\in\tau[\{1,\ldots,M\}]\setminus\{1,\ldots,N\}} a_k\right| \leq \sum_{k=\min \tau[\{1,\ldots,M\}]\setminus\{1,\ldots,N\}}^{k=\max \tau[\{1,\ldots,M\}]\setminus\{1,\ldots,N\}} |a_k| < \epsilon.$$

Weil es aber für jedes $\epsilon > 0$ solche N und M gibt, folgt dann dass eine Teilfolge der Reihe $\left(\sum a_{\tau(n)}\right)_{n\in\mathbb{N}}$ gegen den Grenzwert der Reihe $(\sum a_n)_{n\in\mathbb{N}}$ konvergiert. Weil die erste Reihe konvergiert, gilt dann $\sum_{n\in\mathbb{N}} a_n = \sum_{n\in\mathbb{N}} a_{\tau(n)}$. **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 Unordnung $\left(\sum a_{\tau(n)}\right)_{n\in\mathbb{N}}$, 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 die beiden Reihen $(\sum b_n)_{n\in\mathbb{N}}$ und $(\sum c_n)_{n\in\mathbb{N}}$ und es gilt $\sum\limits_{n=0}^{\infty} b_n = \infty$ und $\sum\limits_{n=0}^{\infty} c_n = -\infty$.

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 $\sum\limits_{k=1}^{n} a_{\tau(k)}$ 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 also* $f : \{x \in \mathbb{K} \mid |x| < R\} \to \mathbb{K}, x \mapsto \sum_{n=0}^{\infty} a_n x^n$ *eine Potenzreihenfunktion.* **q.e.d.**

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

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

(iii) $\left(\sum \frac{x^n}{n}\right)_{n\in\mathbb{N}}$ *also* $\alpha = \overline{\lim} \sqrt[n]{\frac{1}{n}} = \frac{1}{\lim_{n\to\infty} \sqrt[n]{n}} = 1 \Rightarrow R = 1$. *Für* $|x| < 1$ *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* $\alpha = \overline{\lim} \sqrt[n]{\frac{1}{n^2}} = \left(\frac{1}{\lim_{n\to\infty} \sqrt[n]{n}}\right)^2 = 1 \Rightarrow R = 1$. *Für* $|x| < 1$ *also konvergent und für* $|x| > 1$ *divergent, aber für* $|x| = 1$ *auch konvergent.*

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

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

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

(iv) *Für alle* $r < R_1$ *gibt es ein* $L(r)$, *so dass für alle* x, y *mit* $|x| \le r$ *und* $|y| \le r$ *gilt*

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

Beweis:

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

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

(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 gilt $\sum_{n=N}^{\infty} |a_n| r^n < \epsilon$. Dann folgt aber 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 $\left(\sum a_n x^n\right)_{n\in\mathbb{N}_0}$ und $\left(\sum n \cdot a_n x^{n-1}\right)_{n\in\mathbb{N}_0}$ den gleichen Konvergenzradius. Also gilt für $|x| \leq r$ und $|y| \leq r$

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

Wähle also $L(r) = \sum_{n=1}^{\infty} n|a_n| r^{n-1} < \infty$. **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* $\sum\limits_{n=0}^{\infty} a_n x^n$ *und* $\sum\limits_{n=0}^{\infty} b_n x^n$ *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 alle $0 < r < R$ und alle $|x| \leq r$

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

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 $|a_N| \leq |x_m|\left(\sum\limits_{n=0}^{\infty} |a_{n+N+1}|r^n\right)$. Also hat die Potenzreihenfunktion keine Nullstelle auf

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

(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 aber auch für alle $n \in \mathbb{N}_0$:

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

Also ist $b_n = \sum\limits_{k=0}^{\infty} a_{n+k}\binom{n+k}{k} x_0^n < \infty$ für alle $n \in \mathbb{N}$. Wegen Satz 4.21 gilt für alle $N \in \mathbb{N}_0$

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

Also ist auch
$$\sum_{n=0}^{\infty} |b_n| s^n \leq \sum_{n=0}^{\infty} |a_n|(r+s)^n < \infty.$$

Also ist $\sum\limits_{n=0}^{\infty} b_n x^n$ für alle $|x| < R - r$ konvergent, und der Konvergenzradius nicht kleiner als $R - r = R - |x_0|$. Für $|x| < R - |x_0|$ und alle $N, M \in \mathbb{N}_0$ gilt dann
$$\left| \sum_{n=0}^{N} x^n \sum_{k=0}^{M} a_{n+k} \binom{n+k}{k} x_0^k - \sum_{n=0}^{N+M} a_n (x+x_0)^n \right| \leq \sum_{n=\min\{N,M\}}^{N+M} |a_n|(|x| + |x_0|)^n.$$

Also folgt im Grenzwert $M \to \infty$
$$\left| \sum_{n=0}^{N} b_n x^n - \sum_{n=0}^{\infty} (a_n (x+x_0)^n \right| \leq \sum_{n=N}^{\infty} |a_n|(|x| + |x_0|)^n < \infty.$$

und im Grenzwert $N \to \infty$ auch $\sum\limits_{n=0}^{\infty} b_n x^n - \sum\limits_{n=0}^{\infty} a_n (x+x_0)^n = 0$.

(iii) Sei $(x_m)_{m\in\mathbb{N}}$ eine Folge von paarweise verschiedenen Nullstellen der Potenzreihenfunktion $x \mapsto \sum\limits_{n=0}^{\infty} (a_n - b_n) x^n$ 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 $\sum\limits_{n=0}^{\infty} (a_n - b_n)(y + x_0)^n$, 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 $\sum\limits_{n=0}^{\infty} a_n x^n$ und $\sum\limits_{n=0}^{\infty} b_n x^n$ auf dem Gebiet $\{x \in \mathbb{K} \mid |x - x_0| < R - r\}$ überein. Also gibt es auch eine Folge $(\tilde{x}_m)_{m\in\mathbb{N}}$ von paarweise verschiedenen Nullstellen von $x \mapsto \sum\limits_{n=0}^{\infty} (a_n - b_n) x^n$, die gegen ein $\tilde{x}_0$ mit $|\tilde{x}_0| \leq \max\{r - (R - r), 0\} < r$ konvergiert. Induktiv folgt dann, dass dassselbe auch für $\tilde{x}_0 = 0$ gilt. 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

$$\cos(x) = \Re(\exp(\imath x)) \qquad \sin(x) = \Im(\exp(\imath x))$$

und für alle $x \in \mathbb{K}$ die **Eulersche Formel**:

$$\exp(\imath x) = \cos(x) + \imath \sin(x).$$

Außerdem gilt für alle $x \in \mathbb{K}$

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

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

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 $x, y \in \mathbb{K}$

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

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 $n \in \mathbb{N}_0$ $\imath^{2n} = (-1)^n$ und $\imath^{2n+1} = \imath(-1)^n$. 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

Stetige Funktionen auf metrischen Räumen

5.1 Metrische Räume

Definition 5.1. *(Metrik auf einer Menge X) Eine Metrik (oder Distanzfunktion) ist eine Abbildung* $d : X \times X \to \mathbb{R}, (x, y) \mapsto d(x, y)$ *mit drei Eigenschaften*

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

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

(iii) $d(x, y) \leq d(x, z) + d(z, y)$ *(Dreiecksungleichung)*

für alle $x, y, z \in X$.

Wegen der Dreiecksungleichung gilt

$$d(x, y) \leq d(x, u) + d(u, y) \leq d(x, u) + d(u, v) + d(v, y)$$

Also gilt auch $d(x, y) - d(u, v) \leq d(x, u) + d(v, y)$. Durch yertauschen $(x, y) \leftrightarrow (u, v)$ und unter Benutzung der Symmetrie erhalten wir

$$d(u, v) - d(x, y) \leq d(x, u) + d(v, y) \Rightarrow |d(x, y) - d(u, v)| \leq d(x, u) + d(v, y).$$

Wenn also x und u dicht beieinander liegen und y und v, dann ist der Abstand zwischen x und y ungefähr gleich dem Abstand zwischen u und v. (Stetigkeit der Metrik)

Beispiel 5.2. **(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.*

(vi) *Die Einschränkung der Metrik (ii) auf die Vereinigung der inversen der natürlichen Zahlen mit* $\{0\}$ *definiert eine Metrik auf* $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\} \simeq \{\frac{1}{n} \mid n \in \mathbb{N}\} \cup \{0\}$*:*

$$d(n,m) = \frac{|n-m|}{nm} \quad d(\infty,n) = d(n,\infty) = \frac{1}{n} \quad d(\infty,\infty) = 0 \quad \textit{für alle } n,m \in \mathbb{N}.$$

Die Menge $V = \mathbb{R}^n$ bzw. $\mathbb{C}^n$ aller reellen bzw. komplexen 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 heißt (reeller bzw. komplexer) Vektorraum.

Definition 5.3. *Eine Norm auf einem reellen bzw. komplexen Vektorraum* V *ist eine Abbildung* $\|\cdot\| : V \to \mathbb{R}, x \mapsto \|x\|$ *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 5.4. *Jede Norm definiert durch* $d : V \times V \to \mathbb{R}, (x,y) \mapsto \|x-y\|$*eine Metrik.*

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) $d(x,y) = \|x-y\| \leq \|x-z+(z-y)\| \leq \|x-z\| + \|z-y\| = d(x,z) + d(z,y)$.

q.e.d.

Definition 5.5. *(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:

$$\sqrt{(x_1+y_1)^2+\ldots+(x_n+y_n)^2} \le \sqrt{x_1^2+\ldots+x_n^2} + \sqrt{y_1^2+\ldots+y_n^2}.$$

Um diese zu beweisen zeigen wir zuerst die

Cauchy–Schwarz'sche Ungleichung 5.6.

$$|x_1y_1|+\ldots+|x_ny_n| \le \sqrt{x_1^2+\ldots+x_n^2}\sqrt{y_1^2+\ldots+y_n^2}$$

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 $y \ne 0$.

$$\begin{aligned}
\left\|\|y\|x - \frac{x_1y_1+\ldots+x_ny_n}{\|y\|}y\right\|^2 &= \sum_{\imath=1}^n\left(\|y\|x_\imath - \frac{x_1y_1+\ldots+x_ny_n}{\|y\|}y_\imath\right)^2 \\
= \sum_{\imath=1}^n\left(\|y\|^2x_\imath^2 + \frac{(x_1y_1+\ldots+x_ny_n)^2}{\|y\|^2}y_\imath^2 - 2(x_1y_1+\ldots+x_ny_n)x_\imath y_\imath\right) \\
= \|y\|^2\|x\|^2 + \frac{(x_1y_1+\ldots+x_ny_n)^2}{\|y\|^2}\|y\|^2 - 2(x_1y_1+\ldots+x_ny_n)^2 \\
= \|y\|^2\|x\|^2 - (x_1y_1+\ldots+x_ny_n)^2.
\end{aligned}$$

Also gilt $x_1y_1+\ldots+x_ny_n \le \|x\|^2\|y\|^2$. **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 5.7. *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 5.8. *(offene Kugel, Umgebung, offene Menge) Eine offene Kugel in* (X, d) *mit Zentrum* $x \in X$ *und Radius* $r > 0$ *ist die Menge* $B(x, r) = \{y \in X | d(x, y) < r\}$. *Eine Umgebung eines Punktes* $x \in X$ *ist eine Menge* $O \subset X$*, die eine Kugel* $B(x, r)$ *mit einem beliebigen* $r > 0$ *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 5.9. *In* $\mathbb{R}$ *besteht die Kugel* $B(x, r)$ *aus dem Intervall* $(x - r, x + r)$. *Im* $\mathbb{R}^n$ *besteht die Kugel* $B(x, r)$ *aus allen Punkten, deren euklidischer Abstand zu* x *kleiner ist als* r.

Alle offenen Kugeln $B(x, r)$ sind offenbar Umgebungen von x. Sei $y \in B(x, r)$. Dann 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 gilt auch $B(y, r - d(x, y)) \subset B(x, r)$. Deshalb ist sind die offenen Kugeln tatsächlich offene Mengen.

Offenbar ist die beliebige Vereinigung von offenen Mengen wieder 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 wieder offen.

Definition 5.10. *(abgeschlossene Mengen, Abschluss) Die Komplemente von offenen 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 der Abschluss einer beliebigen Menge wieder abgeschlossen und der Abschluss einer abgeschlossenen Menge gleich der Menge. Offenbar gehört ein Punkt x 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 der Kugel $B(x, \frac{1}{n}) \cap A$ gibt, oder auch dazu, dass es eine Folge in A gibt, die gegen x konvergiert. Damit haben wir gezeigt:

Lemma 5.11. *Der Abschluss einer Teilmenge eines metrischen Raumes besteht aus allen Grenzwerten von Folgen innerhalb der Teilmenge. 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.**

5.2 Vollständigkeit und Kompaktheit

Zunächst verallgemeinern wir einige Aussagen über Zahlenfolge auf allgemeine Folgen in metrischen Räumen.

Definition 5.12. *(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 $d(x_n, x_m) < \epsilon$.

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 5.13. *In einem metrischen Raum (X,d) ist jede konvergente Folge eine Cauchyfolge.* **q.e.d.**

Definition 5.14. *Ein metrischer Raum (X,d) heißt vollständig, wenn auch jede Cauchyfolge konvergiert.*

Wegen dem Vollstänidkgietsaxiom sind $\mathbb{R}^n$ und $\mathbb{C}^n$ für alle $n \in \mathbb{N}$ vollständige metrische Räume. Wegen Lemma 5.11 ist jede abgeschlossene Teilmenge eines vollständigen metrsischen Raumes wieder ein vollständiger metrischer Raum.

Übungsaufgabe 5.15. *Zeige in mehereren 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*

$$(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}} \quad \Longleftrightarrow \quad \text{die reelle Folge } (d(x_n, y_n))_{n\in\mathbb{N}} \text{ konvergiert gegen Null.}$$

eine Äquivalenzrelation ist. 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 $\tilde{d}((x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}) = \lim_{n\to\infty} d(x_n, y_n)$ 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 ist.*

(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 $\tilde{X}$).*

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 aus den rationalen Zahlen konstruieren als Äquivalenzklassen von rationalen Cauchyfolgen.

Definition 5.16. *(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 5.17. *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 Kugeln 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 $F_n = \{x_m | m \geq n\}$ 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 $\bigcap\limits_{n\in\mathbb{N}} F_n$ 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 Kugeln 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 $x_{m+1} \in$

$X \setminus \bigcup_{m=1}^{n} B(x_m, \epsilon)$. 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.

(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 Kugeln $B(x_n, 2^{-n})$ keine endliche Teilüberdeckung von $(U_\lambda)_{\lambda\in L}$ besitzen und für alle $n \in \mathbb{N}$ die Kugeln $B(x_{n+1}, 2^{-(n+1)})$ und $B(x_n, 2^{-n})$ nicht disjunkt sind. Weil nämlich (X, d) eine endliche Überdeckung von Kugeln 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 eine Kugel vom Radius $2^{-(n+1)}$, die nichtleeren Schnitt mit $B(x_n, 2^{-n})$ hat und keine endliche Teilüberdeckung von $(U_\lambda)_{\lambda\in L}$ besitzt. Weil aber

$$d(x_n, x_{n+1}) \leq \frac{2}{2^n} \text{ und } \qquad \sum_{m=n}^{\infty} \frac{1}{2^m} = \frac{1}{2^n} \cdot \frac{1}{1-\frac{1}{2}} = 2^{-n+1},$$

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 damit auch

$$B(x_m, 2^{-m}) \subset B(x, 2^{-m+2}) \subset B(x, \epsilon).$$

Also besitzt eine Kugel $B(x_m, 2^{-m})$ eine endliche Teilüberdeckung von $(U_\lambda)_{\lambda\in L}$. Widerspruch. **q.e.d.**

Dieser Satz hat einige wichtige Folgerungen:

Korollar 5.18. **(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 mit ihrem Abschluss überein.

(ii) Abgeschlossene Teilmengen einer kompakten Menge erfüllen offenbar wieder die Bedingung (ii) des vorangehenden Satzes. **q.e.d.**

Definition 5.19. *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 Menge der Abstände $\{d(x, y) | y \in A\}$ beschränkt sind, aber nicht uniform in $x \in X$.

Satz 5.20. *(Heine–Borel) Eine Teilmenge A des metrischen Raumes $(\mathbb{R}^n, d)$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Beweis: Offenbar gilt für alle $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$

$$\max\{|x_1|, |x_2|, \ldots, |x_n|\} \leq \|x\| \leq |x_1| + |x_2| + \ldots + |x_n|.$$

Deshalb sind für eine Folge in einer beschränkten Teilmenge $A \subset \mathbb{R}^n$ alle entsprechenden Koordinatenfolgen beschränkt und besitzen konvergente Teilfolgen. Dann besitzt auch die Folge in A eine konvergente Teilfolge. Also erfüllt eine abeschlossene beschränkte Teilmenge $A \subset \mathbb{R}^n$ die Bedingung (ii) von Satz 5.17. Umgekehrt enthält eine unbeschränkte Teilmenge $A \subset \mathbb{R}^n$ eine Folge $(x_n)_{n \in \mathbb{N}}$, so dass für alle $x \in X$ die Folge $d(x, x_n)$ gegen ∞ konvergiert. Eine solche Folge kann keinen Häufungspunkt haben. **q.e.d.**

Beispiel 5.21. **(i)** *Die Intervalle $[a, b]$ sind kompakt. Weil jede beschränkte Teilmenge von $\mathbb{R}$ eine Folge enthält, die gegen das Supremum bzw. Infimum der Menge konvergiert, enthält jede kompakte Menge auch das Supremum und Infimum und besitzt damit ein Minimum und ein Maximum.*

(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.*

5.3 Stetigkeit

Definition 5.22. *Eine Abbildung $f : X \to Y, x \mapsto f(x)$ von einem metrischen Raum (X, d) in den metrischen Raum (Y, d) heißt stetig in dem Punkt $x \in X$, wenn es für jedes $\epsilon > 0$ ein $\delta > 0$ gibt, so dass alle $y \in X$, die $d(x, y) < \delta$ erüllen, auch $d(f(x), f(y)) < \epsilon$ 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 5.23. *Für eine Abbildung $f : X \to Y, x \mapsto f(x)$ 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)$\Leftrightarrow$(ii): Die Umgebungen von x sind gerade die Mengen, die eine δ-Kugel um x enthalten. Also ist (ii) äquivalent zu der Aussage, dass das Urbild jeder ϵ-Kugel um $f(x)$ eine δ-Kugel um x enthält. Diese Aussage ist nur eine Umformulierung von (i).
(ii)$\Leftrightarrow$(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 also auch $(f(x_n))_{n\in\mathbb{N}}$ gegen $f(x)$. Also folgt aus (ii) auch (iii). Wenn es umgekehrt eine ϵ-Kugel von $f(x)$ gibt, deren Urbild keine δ-Kugel 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 dieser ϵ-Kugel 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 5.24. *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 ist (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 5.25. *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

$$(f\circ g)^{-1}[A] = g^{-1}[f^{-1}[A]].$$

Korollar 5.26. *Das Bild einer kompakten Menge unter einer stetigen Abbildung ist kompakt.*

Beweis: Sei $f : X \to Y, x \mapsto f(x)$ eine stetige Abbildung und $A \subset X$ eine kompakte Menge. Dann ist das Urbild einer beliebig offenen Überdeckung von dem Bild

$$f[A] = \{y \in Y | \exists x \in A \text{ mit } f(x) = y\}$$

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 5.27. *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 5.17 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 5.28. **(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) *Mit der entsprechenden Metrik auf $X \times X$ ist $d : X \times X \to \mathbb{R}$ stetig.*

(iv) *Aus den Rechenregeln für Folgen folgt, dass die Abbildungen*

$$+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}, (x, y) \mapsto x + y \qquad \cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}, (x, y) \mapsto x \cdot y$$

stetig sind. Das gilt auch für die Abbildungen

$$- : \mathbb{K} \to \mathbb{K}, x \mapsto -x \text{ und } \qquad ^{-1} : \mathbb{K} \setminus \{0\} \to \mathbb{K} \setminus \{0\}, x \mapsto x^{-1}.$$

(v) *Eine Folge $(a_n)_{n\in\mathbb{N}}$ läßt sich genau dann zu einer stetigen Abbildung von $\bar{\mathbb{N}}$ nach $\mathbb{K}$ fortsetzten, wenn sie konvergiert. ∞ wird dann auf $\lim\limits_{n\to\infty} a_n$ abgebildet.*

Definition 5.29. *(Gleichmässige Stetigkeit, Lipschitz–Stetigkeit) Eine Abbildung $f : X \to Y, x \mapsto f(x)$ zwischen metrischen Räumen 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 $d(x, y) < \delta$ auch $d(f(x), f(y)) < \epsilon$ gilt.*
Die Abbildung heißt Lipschitz–stetig 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 Lipschitz–stetige Abbildung auch gleichmäßig stetig und jede gleichmäßig stetige Abbildung auch stetig. Es gilt auch folgende Umkehrung:

Satz 5.30. *Sei $f : X \to Y, x \mapsto f(x)$ eine stetige Abbildung zwischen metrischen Räumen. Dann ist f auf jeder kompakten Menge A auch gleichmäßig stetig.*

Auf kompakten Mengen ist also eine Abbildung genau dann stetig, wenn sie gleichmäßig stetig ist.
Beweis: Sei also $f : X \to Y, x \mapsto f(x)$ stetig und $A \subset X$ kompakt. Dann gibt es für jedes $\epsilon > 0$ und jedes $x \in A$ ein $\delta(x)$, so dass aus $d(x,y) < \delta(x)$ auch $d(f(x), f(y)) < \frac{\epsilon}{2}$ folgt. Wir wählen also eine endliche Teilüberdeckung von der offenen Überdeckung $\{B(x, \delta(x)/2) | x \in A\}$ von A. Sei δ das Minimum der Radien dieser endlichen Teilüberdeckung. Dann gibt es für alle $y, z \in A$ mit $d(y,z) < \delta$ eine Kugel $B(x, \delta(x)/2)$ der endlichen Teilüberdeckung mit $y \in B(x, \delta(x)/2)$. Dann folgt

$$d(x,z) \leq d(x,y) + d(y,z) < \frac{\delta(x)}{2} + \frac{\delta(x)}{2} = \delta(x).$$

Also gilt $y, z \in B(x, \delta(x))$. Dann folgt aber

$$d(f(y), f(z)) \leq d(f(x), f(y)) + d(f(x), f(z)) < \epsilon.$$

Also ist f gleichmäßig stetig. **q.e.d.**

Zum Abschluss dieses Abschnittes beweisen wir einen der wichtigsten Sätze der Analysis:

Banachscher Fixpunktsatz 5.31. *Sei $f : X \to X$ eine Lipschitz–stetige 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

$$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq L d(x_{n-1}, x_n) \leq L^n d(x_0, x_1).$$

Mit der Dreieckungleichung 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.**

5.4 Stetige Funktionen

In diesem Abschnitt sei (X, d) ein metrischer Raum und $(\mathbb{K}, d)$ entweder der Körper der reellen oder der komplexen Zahlen mit den entsprechenden Metriken. Zunächst betrachten wir die Menge aller $\mathbb{K}$-wertigen (nicht unbedingt stetigen) Funktionen auf X. Solche Funktionen f, g können wir punktweise miteinander addieren und multiplizieren und wir können sie mit Elementen $\lambda \in \mathbb{K}$ multiplizieren:

$$\begin{aligned} f+g :&X \to \mathbb{K}, & x \mapsto& f(x)+g(x), & f \cdot g :&X \to \mathbb{K}, & x \mapsto& f(x) \cdot g(x), \\ \lambda f :&X \to \mathbb{K}, & x \mapsto& \lambda f(x). \end{aligned}$$

Die Addition erfüllt offenbar die Axiome $A1$ und die Multiplikation die Axiome $A2$, bis auf die Existenz der Inversen. Das Inverse einer Funktion f existiert offenbar nur, wenn $f(x) \neq 0$ für alle $x \in X$. Indem wir die Elemente von $\mathbb{K}$ mit den entsprechenden konstanten Funktionen auf X identifizieren, wird die Multiplikation mit Elementen von $\mathbb{K}$ zu einem Spezialfall der Multiplikation von Funktionen. Der Raum aller Funktionen von X nach $\mathbb{K}$ wird dadurch zu einer Algebra.

Definition 5.32. *Eine Folge von Funktionen $(f_n)_{n\in\mathbb{N}}$ auf der Menge X 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 $f : x \to \mathbb{K}, x \mapsto \lim\limits_{n\to\infty} f_n(x)$.*

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

Offenbar ist jede gleichmäßige konvergente Folge (f_n) auch punktweise konvergent, aber nicht umgekehrt.

Beispiel 5.33. *Die Folge von Funktionen $(f_n)_{n\in\mathbb{N}}$ mit $f_n : [0, 1] \to \mathbb{R}, x \mapsto x^n$ konvergiert wegen Satz 3.4 punktweise gegen die Funktion*

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

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

Definition 5.34. *Eine Funktion $f : X \to \mathbb{K}, x \mapsto f(x)$ heißt beschränkt, wenn das Bild $f[X]$ eine beschränkte Menge ist. $B_{\mathbb{K}}(X)$, bezeichne die Menge aller beschränkten Funktionen auf X. Auf $B_{\mathbb{K}}(X)\| \cdot \|_\infty$ bezeichne folgende Abbildung auf $B_{\mathbb{K}}(X)$:*

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

Satz 5.35. *Es gilt*

(i) $B_{\mathbb{K}}(X)$ *ist eine Unteralgebra aller Funktionen von* X *nach* $\mathbb{K}$.

(ii) $\|\cdot\|_\infty$ *ist eine Norm auf dem* $\mathbb{K}$*-Vektorraum* $B_{\mathbb{K}}(X)$.

(iii) $d : B_{\mathbb{K}}(X) \times B_{\mathbb{K}}(X) \to \mathbb{R}, (f,g) \mapsto d(f,g) = \|f-g\|_\infty$ *ist eine Metrik auf* $B_{\mathbb{K}}(X)$.

(iv) $(B_{\mathbb{K}}(X), d)$ *ist ein vollständiger metrischer Raum.*

Beweis:

(i) Für $\lambda, \mu \in \mathbb{K}$ gilt $|\lambda + \mu| \leq |\lambda| + |\mu|$ und $|\lambda \cdot \mu| = |\lambda| \cdot |\mu|$ Dann folgt auch $\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$ und $\|f \cdot g\|_\infty \leq \|f\|_\infty \cdot \|g\|_\infty$. Damit ist die Summe und das Produkt zweier beschränkter Funktionen wieder beschränkt.

(ii) folgt daraus, dass $|\cdot|$ eine Norm auf $\mathbb{K}$ ist.

(iii) folgt aus (ii) genau wie bei $\mathbb{R}$ und $\mathbb{C}$.

(iv) Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $B_{\mathbb{K}}(x)$. Für alle $\epsilon > 0$ gibt es also ein N, so dass für alle $n, m \geq N$ und alle $x \in X$ gilt

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty < \frac{\epsilon}{2}.$$

Dann sind für alle $x \in X$ die Folgen $(f_n(x))_{n\in\mathbb{N}}$ Cauchyfolgen. Also konvergieren sie punktweise gegen eine Funktion $f : X \to \mathbb{K}, x \mapsto f(x)$. Weil auch die Folge $\|f_n\|_\infty$ eine Cauchyfolge ist und jede Cauchyfolge in $\mathbb{K}$ beschränkt ist, ist dann auch $\|f\|_\infty \leq \lim\limits_{n\to\infty} \|f_n\|_\infty$ beschränkt. Für alle $\epsilon > 0$ und alle $x \in X$ gibt es also ein $N(x) \in \mathbb{N}$, so dass für alle $m \geq N(x)$ gilt $|f_n(x) - f(x)| < \frac{\epsilon}{2}$. Damit folgt für $n \geq N$ und $m \geq \max\{N, N(x)\}$

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

Also konvergiert $(f_n)_{n\in\mathbb{N}}$ gleichmäßIg gegen f. **q.e.d.**

Definition 5.36. $C_{\mathbb{K}}(X)$ *sei der Unterraum von* $B_{\mathbb{K}}(X)$ *aller stetigen und beschränkten Funktionen von* X *nach* $\mathbb{K}$.

Satz 5.37. *Es gilt*

(i) *Der Raum aller stetigen Funktionen von* X *nach* $\mathbb{K}$ *ist eine Unteralgebra aller Funktionen von* X *nach* $\mathbb{K}$. *Das Inverse einer nicht verschwindenden stetigen Funktion von* X *nach* $\mathbb{K}$ *ist wieder stetig.*

(ii) *$C_{\mathbb{K}}(X)$ ist eine Unteralgebra von $B_{\mathbb{K}}(X)$.*

(iii) *$C_{\mathbb{K}}(X)$ ist abgeschlossen in $B_{\mathbb{K}}(X)$ und deshalb vollständig.*

Beweis:

(i) Wegen Korollar 5.24 folgt (i) aus den Rechenregeln für konvergente Folgen.

(ii) Folgt aus (i) und dem vorangehenden Satz.

(iii) Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $C_{\mathbb{K}}(x)$. Dann konvergiert $(f_n)_{n\in\mathbb{N}}$ in $B_{\mathbb{K}}(x)$. Wir müssen zeigen, dass der Grenzwert f stetig ist. Sei $x \in X$. Dann gibt es für jedes $\epsilon > 0$ ein $n \in \mathbb{N}$, so dass $\|f_n - f\|_\infty < \frac{\epsilon}{3}$. Weil f_n stetig ist gibt es ein $\delta > 0$, so dass für alle $y \in B(x, \delta)$ auch $|f_n(x) - f_n(y)| < \frac{\epsilon}{3}$ gilt. Dann folgt aber auch

$$\begin{aligned}|f(x) - f(y)| &\le |f(x) - f_n(x) + f_n(x) - f_n(y) + f_n(y) - f(y)\\ &\le |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)|\\ &< \epsilon. \qquad \text{Also ist } f \text{ stetig.}\end{aligned}$$

q.e.d.

Wichtig in dem Beweis war, dass die Folge f_n in x gleichmäßig gegen f konvergiert und nicht nur punktweise. Auf $[0, 1]$ konvergieren die Funktionen $x \mapsto x^n$ gegen die unstetige Funktion $x \to \begin{cases} 0 & \text{für } x \in [0, 1) \\ 1 & \text{für } x = 1 \end{cases}$

Satz 5.38. *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 $f : X \to \mathbb{K}, x \mapsto f(x)$ stetig. Dann ist $f[X]$ kompakt und damit auch beschränkt. Die kompakten Teilmengen von $\mathbb{R}$ sind aber gerade die beschränkten und abgeschlossenen Teilmengen. Weil für jede beschränkte nicht leere Teilmenge $A \subset \mathbb{R}$ und für jedes $n \in \mathbb{N}$ $\sup A + \frac{1}{n}$ keine obere Schranke von A ist und $\inf A + \frac{1}{n}$ keine untere Schranke von A ist gibt es ein

$$a_n \in (\sup A - \frac{1}{n}, \sup A] \cap A \text{ und ein } b_n \in [\inf A, \inf A - \frac{1}{n}) \cap A.$$

Die Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergieren dann gegen $\sup A$ bzw. $\inf A$. Also liegen $\sup A$ und $\inf A$ im Abschluss von A. Deshalb enthält jede nicht leere kompakte Teilmenge von $\mathbb{R}$ ein Minimum und ein Maximum. **q.e.d.**

Dieser Satz hat viele Anwendungen. Wir werden aus ihm den Fundamentalsatz der Algebra folgern. Zum Abschluss wollen wir folgenden Satz beweisen:

Satz 5.39.* *(Satz von Stone–Weierstraß) Sei (X, d) ein kompakter metrischer Raum und $A \subset C_{\mathbb{R}}(X)$ eine Unteralgebra, die die konstanten Funktionen enthält und die Punkte trennt, d.h. für alle $x \neq y \in X$ gibt es ein $f \in A$, so dass $f(x) \neq f(y)$. Dann ist der Abschluss von A gleich $C_{\mathbb{R}}(X)$.*

Lemma 5.40.* *Auf dem Intervall $[0, 1]$ 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 $0 \leq p_n(x)$ und $0 \leq p_n^2(x) \leq x$ für $x \in [0, 1]$ gilt. Beides ist für $p_0 = 0$ 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 $p_n^2(x) \leq x$ folgt $p_n(x) \leq 1$ und damit $1 - \frac{p_n(x)}{2} \geq \frac{1}{2}$ und $\left(1 - \frac{p_n(x)}{2}\right)^2 \geq \frac{1}{4} \geq \frac{x}{4}$. Deshalb ist die Folge $(p_n(x))_{n \in \mathbb{N}}$ für $x \in [0, 1]$ monoton wachsend und $0 \leq p_n^2(x) \leq x$. Dann gilt aber auch $\left(\left(1 - \frac{p_n(x)}{2}\right)^2 - \frac{x}{4}\right) \leq 1 - \frac{x}{4}$ und deshalb auch $0 \leq x - p_n^2(x) \leq x \cdot \left(1 - \frac{x}{4}\right)^n$. Wegen $\frac{1}{1 - \frac{x}{4}} = \sum\limits_{n=0}^{\infty} \left(\frac{x}{4}\right)^n \geq 1 + \frac{x}{4}$ folgt dann aus der Beroulli Ungleichung $\frac{1}{(1 - \frac{x}{4})^n} \geq 1 + \frac{nx}{4}$ und $0 \leq x - p_n^2(x) \leq \frac{x}{1 + \frac{nx}{4}} < \frac{4}{n}$. Also konvergiert $(p_n^2(x))_{n \in \mathbb{N}}$ auf $x \in [0, 1]$ gleichmäßig gegen x. Die Funktion $[0, 1] \to [0, 1], x \mapsto \sqrt{x}$ ist die Umkehrfunktion von $[0, 1] \to [0, 1], x \mapsto x^2$. Weil die zweite Funktion stetig ist, ist dann wegen Korollar 5.27 die erste auch stetig und wegen Satz 5.30 sogar gleichmäßig stetig. Dann folgt aber dass die Folge $\left(p_n(x) = \sqrt{p_n^2(x)}\right)_{n \in \mathbb{N}}$ gleichmäßig gegen $\sqrt{x}$ konvergiert. **q.e.d.**

Beweis des Satzes von Stone–Weierstraß*: Wegen dem Lemma gibt es auf $[0, 1]$ eine Folge von Polynomen, die gleichmäßig gegen $x \to \sqrt{x}$ konvergieren. Daraus folgt dann, dass für jedes $f \in A$ auch $|f| = \|f\|_\infty \sqrt{\left(\frac{f}{\|f\|_\infty}\right)^2}$ zu dem Abschluss $\bar{A}$ von A gehört. Dann gehört für jedes $f, g \in \bar{A}$ aber auch

$$\sup(f, g) = \frac{1}{2}(f + g + |f - g|) \text{ und } \inf(f, g) = \frac{1}{2}(f + g - |f - g|)$$

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 $f \in A$, dass $f(x) = \alpha$ und $f(y) = \beta$ erfüllt. 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 $f \in C_{\mathbb{R}}(x)$ eine fest vorgegebene Funktion und $\epsilon > 0$. Dann gibt es für jedes $x, y \in X$ eine Funktion $g_{x,y} \in \bar{A}$ die bei x und y mit f übereinstimmt. Dann gibt es aber auch ein $\delta_{x,y} > 0$, so dass für alle $z \in B(y, \delta_{x,y})$ gilt $g_{x,y}(z) < f(z) + \epsilon$. Durch Übergang zu einer endlichen Teilüberdeckung und dem Infimum der entsprechenden Funktionen $g_{x,y} \in \bar{A}$ gibt es dann eine Funktion $g_x \in \bar{A}$, die $g_x(x) = f(x)$ und $g_x < f + \epsilon$ erfüllt. Wegen der Stetigkeit gibt es wieder für alle $x \in X$ ein $\delta_x > 0$, so dass für alle $y \in B(x, \delta_x)$ gilt $f(y) - \epsilon < g_x(y)$. Durch Übergang zu einer endlichen Teilüberdeckung und dem Supremum der entsprechenden Funktionen g_x finden wir schließlich eine Funktion g in $\bar{A}$, die auf X $f - \epsilon < g < f + \epsilon$ erfüllt. Weil ϵ beliebig ist folgt dann, dass f in $\bar{A}$ enthalten ist. **q.e.d.**

Beispiel 5.41. (i) *Weil die identische Abbildung auf $\mathbb{K}$ stetig ist, sind wegen Satz 5.37 auch alle Polynome auf $\mathbb{K}$ stetig und alle Quotienten von Polynomen (rationale Funktionen) auf der Teilmenge von $\mathbb{K}$, auf der der Nenner nicht verschwindet.*

(ii) *Wegen Satz 4.26 (iv) sind alle Potenzreihen stetige Funktionen.*

(iii) *Wenn $f : X \to \mathbb{K}, x \mapsto f(x)$ eine injektive Funktion auf einer kompakten Menge ist, dann ist die Umkehrfunktion $f[X] \to X, x \mapsto f^{-1}(x)$ stetig. Dies gilt auch, wenn X eine offene Teilmenge von $\mathbb{K}$ ist, weil dann jedes $x \in X$ in einer kompakten Umgebung enthalten ist.*

Satz 5.42.* *(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 von stetigen reellen Funktionen, die gegen die stetige Funktion f konvergiert. Dann gibt es zu jedem $x \in X$ ein $n(x)$, so dass für alle $m \geq n(x)$ auch gilt $f(x) - f_m(x) < \frac{\epsilon}{3}$. Da $f_{n(x)}$ und f stetig sind gibt es auch ein $\delta(x)$, so dass aus $y \in B(x, \delta(x))$ folgt

$$|f_{n(x)}(x) - f_{n(x)}(y)| < \frac{\epsilon}{3} \text{ und } |f(x) - f(y)| < \frac{\epsilon}{3}.$$

Dann gilt auch $f(y) - f_{n(x)}(y) < \epsilon$. Wähle nun eine endliche Teilüberdeckung von $\{B(x, \delta(x))| \in X\}$. Dann gilt für jedes m, das größer oder gleich dem Maximum der entsprechenden $n(x)$ ist, für alle $y \in B(x, \delta(x))$

$$f(y) - f_m(y) \leq f(y) - f_{n(x)}(y) < \epsilon.$$

Weil diese Mengen X überdecken, folgt die gleichmäßige Konvergenz. **q.e.d.**

Kapitel 6

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

6.1 Umkehrfunktionen

Satz 6.1. *(Zwischenwertsatz) Sei $f : [a, b] \to \mathbb{R}, x \mapsto f(x)$ 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 $f : X \to \mathbb{R}, x \mapsto f(x)$ 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 $f : [a, b] \to \mathbb{R}, x \mapsto f(x)$ 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 aber f^{-1} wegen Korollar 5.27 bei $y = f(x)$ stetig. Weil f surjetiv ist, ist damit f bei allen y im Wertebereich 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 Punkte des Intervalls, an denen f nicht stetig ist, 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 aber 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 ξ stetig. Die Unstetigkeitsstellen bestehen also aus den Stellen, an denen $f(\xi_-) < f(\xi_+)$. In jedem solchen Intervall $(f(\xi_-), f(\xi_+))$ ist aber 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 und damit 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* 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) $x < y \Rightarrow e^x < e^y$

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

Beweis:

(i) und **(ii)** folgen aus der Definition.

(iii) $x < y \Rightarrow y - x > 0$. 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 $[-n, n] \to \mathbb{R}^+, x \mapsto e^x$. **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 und hat wegen Satz 4.19 und dem vorangehenden Satz folgende Eigenschaften.

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

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

Beweis: (i)$\Leftrightarrow$ $e^0 = 1$ und (ii)$\Leftrightarrow$ $e^1 = e$

(iii) $\Leftrightarrow e^{\ln x + \ln(y)} = (e^{\ln x})(e^{\ln y})$.

(iv) sei $r = \frac{p}{q}$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ $\left(e^{\ln(a)\frac{p}{q}}\right)^q = e^{\ln(a)p} = a^p$ und $e^{\ln(a)\frac{p}{q}} > 0$ Wegen der Eindeutigkeit der k-ten Wurzel gilt dann $e^{\ln(a)\frac{p}{q}} = a^{\frac{p}{q}}$.

(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* $x < y \Rightarrow a^x < a^y$.

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

(v) *Für* $a \neq 1$ *ist* $a^\cdot : \mathbb{R} \to \mathbb{R}^+, x \mapsto a^x$ *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 $\mathbb{R} \to \mathbb{R}, x \mapsto \ln(a)x$ bijektiv und stetig, also auch $\mathbb{R} \to \mathbb{R}^+, x \mapsto \exp(\ln(a)x)$. **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) $\log_a : \mathbb{R}^+ \to \mathbb{R}, x \mapsto \log_a(x)$ *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 alle $x \in [-5, 5]$ gilt $1 - \frac{x^2}{2} \leq \cos x \leq 1 - \frac{x^2}{2} + \frac{x^4}{24}$. Gleichheit gilt nur für $x = 0$.*

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

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

(xi) $\cos : [0, \pi] \to [-1, 1], x \mapsto \cos(x)$ *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 $\left(\frac{x^{2k}}{2k!}\right)_{k \in \mathbb{N}\setminus\{1,2\}}$ 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 für alle $x \in [-5, 5]$ gilt $1 - \frac{x^2}{2} \leq \cos(x) \leq 1 - \frac{x^2}{2} + \frac{x^4}{24}$ 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 $\left(\frac{x^{2k}}{(2k+1)!}\right)_{k \in \mathbb{N}\setminus\{1\}}$ 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 $1 - \frac{x^2}{6} \leq \frac{\sin x}{x} \leq 1$ und Gleichheit nur für $x = 0$ gilt.

(iii) Wegen den Additionstheorem gilt: $\cos(x) - \cos(y) = 2\sin(\frac{x+y}{2})\sin(\frac{y-x}{2}) > 0$ 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 $\exp(\imath\frac{\pi}{2}) = i$.

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

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

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

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

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 bestimmt 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 $\frac{y^2}{x^2+y^2} + \frac{x^2}{x^2+y^2} = 1 \Rightarrow y = r \sin(q)$. 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

$$re^{\imath q} = r'e^{\imath q'} \Rightarrow r = |re^{\imath q}| = |r'e^{\imath q}| = r'.$$

$$e^{\imath q}e^{-\imath q'} = e^{\imath(q-q')} = 1 \Rightarrow q - q' = 2\pi n.$$

q.e.d.

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

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 jedes* $z \in \mathbb{C}\setminus\{0\}$ *gibt es genau* n *verschiedene Lösungen* $w_1, \ldots, w_n$ *der Gleichung* $w^n = z$ *für* $n \in \mathbb{N}$.

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, q_m) und $(s, q_{m'})$ 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 $|z| \geq R = 1 + 2|\dfrac{a_{n-1}}{a_n}| + \ldots + 2|\dfrac{a_0}{a_n}| \geq 1$ gilt

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

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 2.29 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 auch

$$|p(z_0 + y)| < |b_0 + b_m y^m| + \frac{|b_m||y|^m}{2}.$$

Sei jetzt 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 für alle $0 < t < \min\left\{1, \frac{r}{|w|}\right\}$

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

Also ist z_0 nicht das Minimum von $|p(z)|$. Widerspruch. Also muss $|p(z)| = 0$ bei dem Minimum gelten. **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)*

$$\tan: \mathbb{R} \setminus \left\{\left(n + \tfrac{1}{2}\right)\pi | n \in \mathbb{Z}\right\} \to \mathbb{R}, x \to \frac{\sin(x)}{\cos(x)}.$$

$$\cot: \mathbb{R} \setminus \{n\pi | n \in \mathbb{Z}\} \to \mathbb{R}, x \to \frac{\cos(x)}{\sin(x)}.$$

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

(ii) *Die Abbildung* $\cot: (0, \pi) \to \mathbb{R}, x \to \cot(x)$ *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 Satz 5.37 stetig. Außerdem ist für alle $n \in \mathbb{Z}$ $\tan(n\pi) = 0$ und $\cot((n + \frac{1}{2})\pi) = 0$. Dann gilt aber 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 dem Zwischenwertsatz auch surjektiv. **q.e.d.**

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

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

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

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

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

6.4 Konvergenz von reellen Funktionenfolgen

Satz 6.22.* *(Dirichlet) Sei X eine Menge, $(f_n)_{n\in\mathbb{N}}$ eine Folge von reellen beschränkten Funktionen auf X und $(a_n)_{n\in\mathbb{N}}$ eine Folge von $\mathbb{K}$-wertigen Funktionen auf X. Wenn folgende Bedingungen erfüllt sind, dann konvergiert die Reihe von Funktionen $(\Sigma a_n f_n)_{n\in\mathbb{N}}$ auf X gleichmäßig.*

(i) *Für alle $x \in X$ sind die Folgen $(f_n(x))_{n\in\mathbb{N}}$ monoton fallend.*

(ii) *$(\|f_n\|_\infty)_{n\in\mathbb{N}}$ konvergiert gegen Null.*

(iii) *Für alle $n \in \mathbb{N}$ ist* $\|\sum_{k=1}^{n} a_k\|_\infty \le C < 0$.

Beweis*: Wir benutzen die sogenannte Abelsche Summation. Sei $A_n = \sum_{k=1}^{n} a_k$, dann gilt

$$\begin{aligned}\sum_{k=1}^{n} a_k f_k &= A_1 f_1 + (A_2 - A_1) f_2 + \ldots + (A_n - A_{n-1}) f_n \\ &= A_1(f_1 - f_2) + \ldots + A_{n-1}(f_{n-1} - f_n) + A_n f_n\end{aligned}$$

Wegen (i) ist $f_k - f_{k+1} \ge 0$ und $f \ge 0$ wegen (ii). Dann folgt wegen (iii)

$$\begin{aligned}\left\|\sum_{k=n+1}^{m} a_k f_k\right\|_\infty &= \left\|A_m f_m - A_n f_n + \sum_{k=n}^{m-1} A_k(f_k - f_{k+1})\right\|_\infty \\ &\le C\left\|f_m + f_n + \sum_{k=n}^{m-1} f_k - f_{k+1}\right\|_\infty \\ &\le 2C\|f_n\|_\infty.\end{aligned}$$

Wegen (ii) folgt die Behauptung dann aus dem Satz 5.35 (iv). **q.e.d.**

Satz 6.23. *(Abel) Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von reellen Funktionen auf X und $(a_n)_{n\in\mathbb{N}}$ eine Folge von $\mathbb{K}$-wertigen Funktionen auf X. Wenn folgende Bedingungen erfüllt sind, konvergiert die Reihe $(\Sigma a_n f_n)_{n\in\mathbb{N}}$ von Funktionen auf X gleichmäßig.*

(i) *Für alle $x \in X$ ist $(f_n(x))$ monoton fallend.*

(ii) $\|f_n\|_\infty \le C$ *für alle $n \in \mathbb{N}$.*

(iii) $(\Sigma a_n)_{n\in\mathbb{N}}$ *konvergiert gleichmäßig auf X.*

Beweis: Wir benutzen wieder die Abelsche Summation. Sei also $A_n = \sum_{k=1}^{n} a_k$ und $A = \lim_{n\to\infty} A_n$.

$$\begin{aligned}\sum_{k=n+1}^{m} a_k f_k &= A_m f_m - A_n f_n + \sum_{k=n}^{m-1} A_k(f_k - f_{k+1}) \\ &= (A_m - A) f_m - (A_n - A) f_n + \sum_{k=n}^{m-1} (A_k - A)(f_k - f_{k+1}) \text{ wegen} \\ 0 &= A\left(f_m - f_n + \sum_{k=n}^{m-1} f_k - f_{k+1}\right).\end{aligned}$$

Wegen (iii) gibt es für jedes $\epsilon > 0$ ein N, so dass alle $k \geq N$ auch $\|A_k - A\|_\infty < \frac{\epsilon}{4C}$ erfüllen. Also gilt für $N \leq n < m$

$$\begin{aligned}\left\| \sum_{k=n+1}^{m} a_k f_k \right\|_\infty &< \frac{\epsilon}{4C}\left(\|f_m\|_\infty + \|f_n\|_\infty + \left\| \sum_{k=n}^{m-1} f_k - f_{k+1} \right\|_\infty \right) \\ &\leq \frac{2\epsilon}{4C}(\|f_m\|_\infty + \|f_n\|_\infty) \leq \epsilon.\end{aligned}$$

Also folgt die Behauptung wieder aus Satz 5.35 (iv). **q.e.d.**

Satz 6.24. *(Abelscher Grenzwertsatz) Wenn die $\mathbb{K}$–wertige Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ für $x = x_0 \neq 0$ konvergiert. Dann konvergiert sie gleichmäßig auf $\{tx_0 \in \mathbb{K} \mid t \in [0,1]\}$*

Beweis: Für $x_0 \neq 0$ sei für alle $n \in \mathbb{N}_0$

$$f_n : \{tx_0 \in \mathbb{K} \mid t \in [0,1]\} \to \mathbb{R}, \qquad x = tx_0 \mapsto \left(\frac{x}{x_0}\right)^n = t^n$$
$$b_n : \{tx_0 \in \mathbb{K} \mid t \in [0,1]\} \to \mathbb{K}, \qquad x = tx_0 \mapsto a_n \cdot x_0^n.$$

Für alle $t \in [0,1]$ ist die Folge $(f_n)_{n\in\mathbb{N}}$ punktweise monton fallend und beschränkt durch 1. Wenn also $(\Sigma a_n x_0^n)_{n\in\mathbb{N}_0}$ konvergiert, erfüllen die Folgen $(f_n)_{n\in\mathbb{N}_0}$ und $(b_n)_{n\in\mathbb{N}_0}$ die Voraussetzungen des vorangehenden Satzes. Also konvergiert $(\Sigma b_n f_n)_{n\in\mathbb{N}_0} = (\Sigma a_n x^n)_{n\in\mathbb{N}_0}$ auf $\{tx_0 \in \mathbb{K} \mid t \in [0,1]\}$ gleichmäßig. **q.e.d.**

Wenn also eine Potenzreihenfunktion mit Konvergenzradius R für $x_0 \in \mathbb{K}$ mit $|x_0| = R$ konvergiert, dann konvergiert sie auf der kompakten Menge $\{tx_0 \in \mathbb{K} \mid t \in [0,1]\}$ gegen eine stetige Funktion. Insbesonders ist dann also der Funktionswert bei x_0 gegeben durch den Grenzwert $\lim_{n\to\infty} f\left(\left(1 - \frac{1}{n}\right) x_0\right)$ der Potenzreihenfunktion. Wir werden später sehen, dass wir dadurch in vielen Fällen diesen Grenzwert bestimmen können.

Beispiel 6.25. (i) *Die Potenzreihe $\left(\sum \frac{x^n}{n}\right)_{n\in\mathbb{N}}$ konvergiert auf $[-1,1)$ gegen eine stetige Funktion.*

(ii) *Die Potenzreihe $\left(\sum \frac{(-1)^n x^{2n+1}}{2n+1}\right)_{n\in\mathbb{N}_0}$ konvergiert auf $[-1,1]$ gegen eine stetige Funktion.*

Kapitel 7

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

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*

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

stetig bei $x = x_0$ ist. Wenn X offen ist und f in jedem Punkt differenzierbar ist, heißt die Funktion $f' : X \to \mathbb{R}, x \mapsto f'(x)$ 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 5.23 **q.e.d.**

Definition 7.3. *Das Differential von f im Punkt x_0 ist die lineare Abbildung $df(x_0) : \mathbb{R} \to \mathbb{R}, h \mapsto f'(x_0)h$. 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) *$f(x) = c \Rightarrow \frac{f(x)-f(x_0)}{x-x_0} = 0$ für alle $x \neq x_0$ also ist f differenzierbar und es gilt $f'(x) = 0$.*

(iii) *$f(x) = x \Rightarrow \frac{f(x)-f(x_0)}{x-x_0} = 1$ 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) *$f(x) = \sin(x)$*

$$\frac{\sin(x)-\sin(x_0)}{x-x_0} = \frac{\sin\left(\frac{x-x_0}{2}+\frac{x+x_0}{2}\right)+\sin\left(\frac{x-x_0}{2}-\frac{x+x_0}{2}\right)}{x-x_0} = \frac{2\sin\left(\frac{x-x_0}{2}\right)\cos\left(\frac{x+x_0}{2}\right)}{x-x_0}$$

Und wegen der Potenzreihe von sin *gilt auch*

$$\frac{2\sin(\frac{x-x_0}{2})}{x-x_0} = \sum_{k=0}^{\infty}\frac{(-1)^k(\frac{x-x_0}{2})^{2k}}{(2k+1)!}$$

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) $f(x) = \cos(x)$

$$\frac{\cos(x) - \cos(x_0)}{x - x_0} = \frac{\cos\left(\frac{x_0+x}{2} - \frac{x_0-x}{2}\right) - \cos\left(\frac{x_0+x}{2} + \frac{x_0-x}{2}\right)}{x - x_0} = \frac{2\sin\left(\frac{x_0+x}{2}\right)\sin\left(\frac{x_0-x}{2}\right)}{x - x_0}.$$

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. *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 Satz 5.37.

Satz 7.6. *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: $\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \cdot \frac{f(x) - f(x_0)}{x - x_0}$. Der ertse Faktor ist aber die Komposition von $x \mapsto f(x)$ mit $y \mapsto \frac{g(y) - g(y_0)}{y - y_0}$ also wegen Satz 5.25 und wegen Satz 7.2 stetig. Also folgt die Behauptung aus Satz 5.37.

Satz 7.7. *(Ableitung der Inversen) 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 $(f^{-1})'(y_0) = \frac{1}{f'(x_0)}$.*

Beweis: $\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)}$ 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 $x \mapsto \begin{cases} \frac{x-x_0}{f(x)-f(x_0)} & \text{für } x \neq x_0 \Leftrightarrow f(x) \neq f(x_0) \\ \frac{1}{f'(x_0)} & \text{für } x = x_0 \end{cases}$. Also folgt der Satz aus Satz 5.25. **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) $\arccos : [-1, 1] \to [0, \pi], x \mapsto \arccos(x)$

$$\arccos'(x) = \frac{-1}{\sin(y)} = \frac{-1}{\sqrt{1 - x^2}} \textit{ mit } \cos(y) = x \textit{ und } x^2 \neq 1.$$

(iv) $\cdot^\alpha : \mathbb{R}^+ \to \mathbb{R}, x \mapsto x^\alpha, \alpha \in \mathbb{R}$.

$$(\cdot^\alpha)'(x) = \exp(\alpha \ln(x))' = \exp(\alpha \ln(x)) \cdot \frac{\alpha}{x} = \alpha x^{\alpha - 1}.$$

(v) $a^\cdot : \mathbb{R} \to \mathbb{R}, x \mapsto a^x, a \in \mathbb{R}^+$.

$$(a^\cdot)'(x) = \exp(x \cdot \ln(a))' = \exp(x \ln(a)) \cdot \ln(a) = \ln(a) \cdot a^x.$$

(vi) Quotientenregel. *Seien f und g in x_0 differenzierbar und $g(x_0) \neq 0$. Dann ist $\frac{f}{g}$ ind 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) $\arctan : \mathbb{R} \to \left(\frac{\pi}{2}, \frac{\pi}{2}\right), x \mapsto \arctan(x)$

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

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

$$\operatorname{arccot}'(x) = \frac{-1}{1 + \cot^2(g)} = \frac{-1}{1 + x^2} \textit{ mit } \cot(y) = x.$$

(xi) $\log_a \mathbb{R}^+ \to \mathbb{R}, x \mapsto \log_a(x) \quad \log_a'(x) = \left(\dfrac{\ln(x)}{\ln(a)}\right)' = \dfrac{1}{x\ln(a)}.$

(xii) $x^x : \mathbb{R}^+ \to \mathbb{R}^+, x \mapsto x^x$

$$(x^x)' = \exp(x \cdot \ln(x))' = \exp(x \cdot \ln(x)) \left(\ln(x) + x\frac{1}{x}\right) = (\ln(x) + 1) \cdot x^x.$$

(xiii) $f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \textit{für } x \neq 0 \\ 0 & \textit{für } x = 0. \end{cases}$

$$f'(x) = \begin{cases} \lim\limits_{x\to 0\pm} \frac{x^2 \sin\left(\frac{1}{x}\right)}{x} = \lim\limits_{x\to 0\pm} x\sin\left(\frac{1}{x}\right) = 0 & \textit{für } x = 0 \\ 2x\sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right) & \textit{für } x \neq 0. \end{cases}.$$

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) 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 5.26 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 dann die Ableitung verschwindet. **q.e.d.**

Satz 7.12. *(verallgemeinerter Mittelwertsatz) Seien* $f, g \cdot [a, b] \to \mathbb{R}$ *stetig und auf* (a, b) *differenzierbar. Dann existiert ein* $x_0 \in (a, b)$ *mit*

$$(f(b) - f(a))g'(x_0) = (g(b) - g(a))f'(x_0).$$

Beweis: Wende den Satz von Rolle auf die Funktion $x \mapsto (f(b) - f(a))g(x) - (g(b) - g(a))f(x)$ an. Offenbar erfüllt sie die Voraussetzungen und ihre Ableitung ist $\mapsto (f(b) - f(a))g'(x) - (g(b) - g(a))f'(x)$. **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 gilt $|f'(x)| \leq L$ für alle $x \in (a, b)$, dann ist f Lipschitz–stetig 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 $f : [a, b] \to \mathbb{R}$ auf (a, b) differenzierbar. Dann gilt*

(i) *$f'(x) = 0$ für alle $x \in (a, b) \Leftrightarrow f$ ist konstant.*

(ii) *$f'(x) \geq 0$ für alle $x \in (a, b) \Leftrightarrow f$ ist monoton steigend.*

(iii) *$f'(x) \leq 0$ für alle $x \in (a, b) \Leftrightarrow f$ 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 $[a, b] \Leftrightarrow f$ 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 $[a, b] \Leftrightarrow f$ 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 $x_1 < x_2 \in [a, b]$. Dann erfüllt $f : [x_1, x_2] \to \mathbb{R}$ die Voraussetzungen des Mittelwertsatzes. Wenn $f'(x) \geq 0$ für alle $x \in (a, b)$, dann folgt also $f(x_2) - f(x_1) \geq 0$ 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 $x_1 < x_2 \in [a, b]$ mit $f(x_1) = f(x_2)$. Dann ist f aber auf $[x_1, x_2']$ konstant und f' verschwindet auf (x_1, x_2). 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 $f : \mathbb{R} \to \mathbb{R}, x \mapsto (x+1)e^{-x}$ 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 $f(0) = 1$ das 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 $f : (a,b) \to \mathbb{R}, x \mapsto f(x)$ 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 5.23 ist das äquivalent dazu, dass für jede Folge $(x_n)_{n \in \mathbb{N}}$ in (a, b), die gegen a konvergiert, die Folge der Funktionswerte $(f(x_n))_{n \in \mathbb{N}}$ 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 außerdem 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 wegen Satz 5.37 auch* $\lim\limits_{x\to a+} \frac{f'(x)}{g'(x)}$ *und es gilt* $\lim\limits_{x\to a+} \frac{f'(x)}{g'(x)} = \frac{\lim\limits_{x\to a+} f'(x)}{\lim\limits_{x\to a+} g'(x)}$.

Beweis: Die Funktion 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 $\dfrac{f(x)}{g(x)} = \dfrac{f(x)-f(a)}{g(x)-g(a)} = \dfrac{f'(x_0)}{g'(x_0)}$. Wenn also der Grenzwert $\lim\limits_{x\to a+} \frac{f'(x)}{g'(x)}$ existiert, dann existiert auch der Grenzwert $\lim\limits_{x\to a+} \frac{f(x)}{g(x)} = \lim\limits_{x\to a+} \frac{f'(x)}{g'(x)}$. **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 $\dfrac{f(x)-f(y)}{g(x)-g(y)} = \dfrac{f'(x_0)}{g'(x_0)}$. 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 $x \in (a,y)$ gilt $\lim\limits_{x_0\to a+} \dfrac{f'(x_0)}{g'(x_0)} - \dfrac{\epsilon}{2} < \dfrac{f(x)-f(y)}{g(x)-g(y)} < \lim\limits_{x_0\to a+} \dfrac{f'(x_0)}{g'(x_0)} + \dfrac{\epsilon}{2}$. Wegen $\lim\limits_{x\to a+} g(x) = \infty$ gibt es ein $y_0 \in (a,y)$ so dass für alle $x \in (a,y_0)$ gilt $g(x) > \min\{(g(y),0\}$. Mit $\alpha = \lim\limits_{x_0\to a+} \dfrac{f'(x_0)}{g'(x_0)}$ folgt dann

$$\left(\alpha - \frac{\epsilon}{2}\right)(g(x)-g(y)) + f(y) \quad < f(x) < \quad \left(\alpha + \frac{\epsilon}{2}\right)(g(x)-g(y)) + f(x) \quad \text{oder auch}$$
$$\left(\alpha - \frac{\epsilon}{2}\right) + \frac{f(y)-g(y)\left(\alpha-\frac{\epsilon}{2}\right)}{g(x)} \quad < \frac{f(x)}{g(x)} < \quad \left(\alpha + \frac{\epsilon}{2}\right) + \frac{f(y)-g(y)\left(\alpha+\frac{\epsilon}{2}\right)}{g(x)}.$$

Wegen $\lim\limits_{x\to a+} g(x) = \infty$ gibt es dann auch ein $y_0 \in (a,y)$, so dass für alle $x \in (a,y_0)$ gilt $\alpha - \epsilon < \dfrac{f(x)}{g(x)} < \alpha + \epsilon$. Also gilt $\lim\limits_{x_0\to a+} \dfrac{f(x_0)}{g(x_0)} = \lim\limits_{x_0\to a+} \dfrac{f'(x_0)}{g'(x_0)}$. **q.e.d.**

Die analogen Aussagen für die Grenzwerte $\lim\limits_{x\to b-}$ gelten natürlich auch. Grenzwerte der Form $\lim\limits_{x\to -\infty+} f(x)$ bzw. $\lim\limits_{x\to\infty-} f(x)$ definieren wir als die Grenzwerte $\lim\limits_{x\to 0-} f(1/x)$ bzw. $\lim\limits_{x\to 0+} f(1/x)$. Wegen der Kettenregel gilt dann $\dfrac{\frac{df(1/x)}{dx}}{\frac{dg(1/x)}{dx}} = \dfrac{f'(1/x)}{g'(1/x)}$. 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*

$$f((1-t)a + tb) \leq (1-t)f(a) + tf(b) \quad \textit{bzw.} \quad f((1-t)a + tb) < (1-t)f(a) + tf(b).$$

Satz 7.23. *Für eine reelle Funktion f auf einem Intervall sind folgende Eigenschaften äquivalent:*

(i) *f ist konvex*

(ii) *Für $a < x < b$ im Definitionsbereich gilt $f(x) \leq f(a) + \dfrac{f(b)-f(a)}{b-a}(x-a) = L(x)$.*

(iii) *Für $a < x < b$ im Definitionsbereich gilt*

$$\frac{f(x)-f(a)}{x-a} \leq \frac{f(b)-f(a)}{b-a} \leq \frac{f(b)-f(x)}{b-x}$$

(iv) *Für $a < x < b$ im Definitionsbereich gilt $\dfrac{f(x)-f(a)}{x-a} \leq \dfrac{f(b)-f(x)}{b-x}$.*

Außerdem gelten die analogen Äquivalenzen zu streng konvex, wenn für dieselben $a < x < b$ die Ungleichungen $\leq$ 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) annehmen $a < b$. 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 $\dfrac{f(x)-f(a)}{x-a} \leq \dfrac{f(c)-f(a)}{c-a} \leq \dfrac{f(y)-f(c)}{y-c}$. Aus dem Grenzwert $x \to a$ und $y \to c$ folgt dann $f'(a) \leq \dfrac{f(c)-f(a)}{c-a} \leq f'(c)$ 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)** $f(x) = x^2 \Rightarrow f'' = 2$. *Also ist f streng konvex.*

(ii) $f : \mathbb{R}_0^+ \to \mathbb{R}_0^+, x \mapsto \sqrt{x} \Rightarrow f'' = \frac{-1}{4x^{3/2}}$. *Also ist f streng konkav.*

(iii) $\exp : \mathbb{R} \to \mathbb{R}^+, \to \exp(x) \Rightarrow \exp'' = \exp$. *Also ist* exp *streng konvex.*

(iv) $\ln : \mathbb{R}^+ \to \mathbb{R}, x \to \ln(x) \Rightarrow \ln''(x) = -\frac{1}{x^2}$. *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

$$-\ln(\lambda_1 x_1 + \ldots + \lambda_n x_n) \leq -\lambda_1 \ln x_1 - \ldots - \lambda_n \ln x_n$$
$$\Leftrightarrow \lambda_1 \ln x_1 + \ldots + \lambda_n \ln x_n \leq \ln(\lambda_1 x_1 + \ldots + \lambda_n x_n)$$

Wegen der Monotonie von exp folgt:

$$x_1^{\lambda_1} \cdots x_n^{\lambda_n} = \exp(\lambda_1 \ln x_1 + \ldots + \lambda_n \ln x_n) \leq \lambda_1 x_1 + \ldots + \lambda_n x_n.$$

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.**

Definition 7.32. *(Norm und Skalarprodukt) Für $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ und $1 \leq p < \infty$ sei*

$$\|x\|_p = (|x_1|^p + \ldots + |x_n|^p)^{1/p}$$

Für $p \to \infty$ hatten wir in einer Übungsaufgabe gesehen, dass $\|x\|_p$ konvergiert gegen

$$\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}.$$

Das Skalarprodukt wird definiert als

$$\langle x, y \rangle = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n \in \mathbb{K}, \text{ für } x, y \in \mathbb{K}^n.$$

Für $y \in \mathbb{R}^n$ setzen wir hierbei $(\bar{y}, \ldots, \bar{y}_n) = \bar{y} = y$.

Satz 7.33. *(Höldersche Ungleichung) Seien $p \geq 1$ und $q \geq 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt für alle $x, y \in \mathbb{K}^n$*

$$|\langle x, y \rangle| \leq |x_1y_1| + \ldots + |x_ny_n| \leq \|x\|_p \cdot \|y\|_q$$

Für $p = q = 2$ erhalten wir wieder die Cauchy-Schwarzsche Ungleichung

$$|x_1y_1| + \ldots + |x_ny_n| \leq \|x\| \cdot \|y\|.$$

Beweis: Wir können annehmen, dass $\|x\|_p \neq 0$ und $\|y\|_q \neq 0$ gilt. Dann folgt aus der Young'schen Ungleichung für alle $k = 1, \ldots, n$

$$\frac{|x_ky_k|}{\|x\|\|y\|_q} = \frac{|x_k|}{\|x\|_p}\frac{|y_k|}{\|y\|_q} \leq \frac{1}{p}\frac{|x_k|^p}{\|x\|^p} + \frac{1}{q}\frac{|y_k|^q}{\|y\|_q}$$

Nach Summation über $k = 1, \ldots, n$ erhalten wir

$$\frac{|x_1y_1| + \ldots + |x_ny_n|}{\|x\|_p\|y\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1$$

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. **q.e.d.**

Satz 7.34. *(Minkowski Ungleichung) Sei $p \geq 1$ und $x, y \in \mathbb{K}^n$, dann gilt*

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p$$

Korollar 7.35. *Für alle $1 \leq p \leq \infty$ ist $\|\cdot\|_p : \mathbb{K} \to \mathbb{R}$ eine Norm.* **q.e.d.**

Beweis der Minkowski Ungleichung: Sei $\frac{1}{p} + \frac{1}{q} = 1 \Leftrightarrow p + q = pq \Leftrightarrow p = (p-1)q$

$$\begin{aligned} |x_1+y_1|^p + \ldots + |x_n+y_n|^p &= |x_1+y_1||x_1+y_1|^{p-1} + \ldots + |x_n+y_n||x_n+y_n|^{p-1} \\ &\leq (|x_1|+|y_1|)|x_1+y_1|^{p-1} + \ldots + (|x_n|+|y_n|)|x_n+y_n|^{p-1} \\ &\leq (\|x\|_p + \|y\|_p)(|x_1+y_1|^{(p-1)q} + \ldots + |x_n+y_n|^{(p-1)q})^{1/q} \\ &= (\|x\|_p + \|y\|_p)\|x+y\|_p^{p/q} \end{aligned}$$

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. Sei also $\|x+y\|_p \neq 0$, dann folgt $\|x+y\|_p = \|x+y\|_p^{p(1-\frac{1}{q})} \leq \|x\|_p + \|y\|_p$. **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 7.36. *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 $C^\infty_{\mathbb{R}}$ 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 7.37. **(i)** $\exp \in C^\infty_{\mathbb{R}}$, *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 7.38. *Für alle* $n \in \mathbb{N}$ *gilt*

(i) $\dfrac{d^n}{dx^n} f \cdot g = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}$ *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 7.39. (i) *Die Komposition von n-mal stetig differenzierbaren Funktionen ist wieder n-mal stetig differenzierbar.*

(ii) *Die inverse Funktion einer n-mal stetig differenzierbaren invertierbaren Funktion ist n-mal stetig differenzierbar, wenn die Ableitung keine Nullstellen hat.* **q.e.d.**

Definition 7.40. *(Taylor-Polynom) Sei* $f : I \to \mathbb{R}$ *in* x_0 *n-mal differenzierbar, also es gibt eine offene Menge* $O < I$, *die* x_0 *enthält, so dass die Einschränkung von* f *auf* O *in* $C^n(O)$ *liegt. 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:

$$p(x) = c_0 + c_1(x-x_0) + \ldots + c_n(x-x_0)^n \Rightarrow p(x_0) = c_0, p^{(1)}(x_0) = c_1, \ldots, p^{(n)}(x_0) = n!c_n.$$

Satz 7.41. *(Taylor-Formel) Sei $f \in C^n_{\mathbb{R}}((a,b))$. Wenn $f^{(n+1)}(x)$ für alle $x \in (a,b)$ existeirt, 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

$$f(x) - T_{n,x_0}(x) = \frac{f^{(n+1)}(\xi)(x-\xi)^n(x-x_0)^{n+1}}{n!(n+1)(x-\xi)^n} = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}.$$ **q.e.d.**

Definition 7.42. *(Taylorreihe) Sei $f \in C^\infty((a,b))$ und $x_0 \in (a,b)$. Dann heißt die Potenzreihe $\sum\limits_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}$ Taylorreihe von f in x_0. Es gibt für jedes $x_0, x \in (a,b)$ 3 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 7.43. *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 für alle $\xi \in (x_0,x)$ bzw. (x,x_0) gilt $\lim\limits_{n\to\infty} \left|\frac{f^n(\xi)}{n!}(x-x_0)^n\right| = 0$.* **q.e.d.**

Beispiel 7.44.

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \textit{für } x > 0 \\ 0 & \textit{für } x \leq 0 \end{cases}$$

$$f^{(n)}(x) = \begin{cases} \textit{Polynom vom Grad 3n von } (\frac{1}{x})e^{-\frac{1}{x^2}} & \textit{für } x > 0 \\ 0 & \textit{für } x \leq 0 \end{cases}$$

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* $x \to \begin{cases} \frac{1}{x^n}e^{-\frac{1}{x^2}} & \textit{für } x > 0 \\ 0 & \textit{für } x \leq 0 \end{cases}$ *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 7.45. *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 $n, m \in \mathbb{N}$

$$|f_n(x) - f_m(x)| \leq |f_n(x_0) - f_m(x_0)| + |x - x_0| \, \|f'_n - f'_m\|_\infty .$$

Also 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 $x, x_0 \in (a, b)$ eine Folge $(\xi_n)_{n\in\mathbb{N}} \in [x, x_0]$ bzw. $[x_0, x]$, so dass für alle $n \in \mathbb{N}$ gilt $f_n(x) - f_n(x_0) = (x - x_0)f'_n(\xi_n)$. Wegen dem Auswahlprinzipien von Bolzano-Weierstraß gibt es eine konvergente Teilfolge von $(\xi_n)_{n\mathbb{N}}$ mit Grenzwert $\xi \in [x, x_0]$ bzw. $[x_0, x]$. Wegen $|f'_n(\xi_n) - g(\xi)| \leq |f_n(\xi_n) - g(\xi_n)| + |g(\xi_n) - g(\xi)|$ und der Stetigkeit von g konvergiert die Folge $(f'_n(\xi_n))_{n\mathbb{N}}$ gegen $g(\xi)$. Also konvergiert $(f_n(x))_{n\in\mathbb{N}}$ gegen $f(x) = f(x_0) + (x - x_0)g(\xi)$. Aus der Stetigkeit von g folgt, dass f bei x_0 differentzierbar ist und $g(x_0)$ die Ableitung $f'(x_0)$ ist. **q.e.d.**

Korollar 7.46. *Sei* $(\sum f'_n)_{n\in\mathbb{N}}$ *gleichmäßig konvergent auf einem beschränktem Intervall* (a, b) *und* $(\sum f(x_0))_{n\in\mathbb{N}}$ *konvergent mit* $x_0 \in (a, b)$. *Dann konvergiert* $(\sum f_n)_{n\in\mathbb{N}}$ *gleichmäßig gegen eine Funktion* $f \in C^1_\mathbb{R}((a, b))$ *und* $(\sum f'_n)_{n\in\mathbb{N}}$ *gegen* f'. **q.e.d.**

Korollar 7.47. *(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 $n\in\mathbb{N}_0$ existiert dann eine Konstante $M_n>0$

$$M_n=\max\{\|h_n\|_\infty,\|h_n'\|_\infty,\ldots,\|h_n^{(n)}\|_\infty\}$$

mit $h_n:\mathbb{R}\to\mathbb{R},x\mapsto x^n\cdot h(x)$. 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 für alle $n>m\in\mathbb{N}_0$

$$\|f_n^{(m)}\|_\infty=\frac{|a_n|C_n^m}{n!C_n^n}\|h_n^{(m)}\|_\infty\le\frac{|a_n|M_nC_n^m}{n!C_n^n}<\frac{C_n^{m+1}}{n!C_n^n}\le\frac{1}{n!}.$$

Also konvergiert für alle $m\in\mathbb{N}_0$ $(\Sigma f_n^{(m)})_{n\in\mathbb{N}_0}$ gleichmäßig. Wegen Korollar 7.46 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 $f=\sum\limits_{n=0}^\infty f_n$ 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 7.48. *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: Wegen $\lim\limits_{n\to\infty}\sqrt[n]{n}=1$ hat die Potenzreihenfunktion $\sum\limits_{n=0}^\infty(a_nx^n)'=\sum\limits_{n=0}^\infty(n+1)a_{n+1}x^n$ offenbar den gleichen Konvergenzradius wie $\left(\sum a_nx^n\right)_{n\in\mathbb{N}_0}$. Wegen Korollar 7.46 folgt dann, dass f differenzierbar ist und f' gegeben ist durch $\sum\limits_{n=0}^\infty(n+1)a_{n+1}x^n$. Dann folgt die Aussage aus dem Identitätssatz für Potenzreihenfunktionen (ii). **q.e.d.**

Definition 7.49. *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 7.50. (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 $\left\| \frac{f(0)-f}{f(0)} \right\|_\infty \leq L < 1$ gilt. Dort konvergiert $\frac{1}{f} = \sum_{n=0}^\infty \frac{1}{f(0)} \left(\frac{f(0)-f}{f(0)} \right)^n$ als Potenzreihenfunktion. Dann folgt aus dem Identitässatz für Potenzreihenfunktionen, 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*

$$x \mapsto f(x) = \sum_{n=0}^\infty \binom{\alpha}{n} x_0^{\alpha-n} x^n \quad \text{mit} \quad \binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n(n-1)\cdots 1}$$

den Konvergenzradius $R = \frac{1}{\lim\limits_{n\to\infty} \frac{\alpha-n}{|x_0|(n+1)}} = |x_0|$. Die Ableitung dieser Potenzreihenfunktion 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 aber $(x_0+x)\sum_{n=0}^\infty \binom{\alpha-1}{n} x_0^{\alpha-1-n} x^n = \sum_{n=0}^\infty \binom{\alpha}{n} x_0^{\alpha-n} x^n$. Dann erfüllt f die Differentialgleichung $f' = \alpha(x_0+x)f$ 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 verschwindet 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

$$\ln\left((-1)^n \binom{-\frac{1}{2}}{n}\right) = \ln\left(\frac{1}{2}\frac{3}{4}\cdots\frac{2n-1}{2n}\right) = \ln\left(\left(1-\frac{1}{2}\right)\left(1-\frac{1}{4}\right)\cdots\left(1-\frac{1}{2n}\right)\right)$$
$$\leq -\frac{1}{2}\left(1+\frac{1}{2}+\ldots+\frac{1}{n}\right) \leq -\frac{1}{2}\ln\left(\frac{2}{1}\frac{3}{2}\cdots\frac{n+1}{n}\right) = \ln\left(\frac{1}{\sqrt{n+1}}\right).$$

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 diese 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 7.51. *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 Riemann–integrable 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 $\|p\| = \max_i \Delta_i$ mit $\Delta_i = x_i - x_{i-1}$ 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*

$$s(p,f) = \sum_{i=1}^{n} m_i \Delta_i \text{ und } S(p,f) = \sum_{i=1}^{n} M_i \Delta x_i$$

die Untersumme und Obersumme von f bezüglich der Partition p.

Offenbar gilt

$$m(b-a) \leq s(p,f) \leq S(p,f) \leq M(b-a).$$

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 $s(p,f) \leq s(p',f)$ und $S(p',f) \leq S(p,f)$.*

(ii) *Für $p, p' \in \mathcal{P}[a,b]$ gilt $s(p,f) \leq S(p',f)$.*

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

$$m(b-a) \leq s(p,f) \text{ und } S(p,f) \leq M(b-a).$$

(ii) Sei $p'' = p \cup p'$. Dann folgen aus (i) die Ungleichungen

$$\begin{aligned} s(p,f) \leq s(p'',f) \leq S(p'',f) \leq S(p,f) \\ s(p',f) \leq s(p'',f) \leq S(p'',f) \leq S(p',f). \end{aligned}$$

Daraus folgt dann (ii). **q.e.d.**

Definition 8.5. *(Unterintegral und Oberintegral) Für $f \in B_{\mathbb{R}}([a,b])$ heißt $\underline{\int} f = \sup\limits_{p \in \mathcal{P}[a,b]} s(p,f)$ Unterintegral und $\overline{\int} f = \inf\limits_{p \in \mathcal{P}[a,b]} S(p,f)$ Oberintegral von f.*

Offenbar gilt

$$\underline{\int} f \leq \overline{\int} f.$$

Definition 8.6. *Eine Funktion $f \in B_{\mathbb{R}}([a,b])$ heißt Riemann–integrabel, wenn gilt $\underline{\int} f = \overline{\int} f$. Diese Zahl heißt dann das Riemannintegral von f über $[a,b]$: $\int\limits_a^b f dx$. Die Menge aller Riemann–integrablen 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 $\int\limits_a^b f(x)dx$ als diesen Flächeninhalt.

8.2 Kriterien von Darboux, Riemann und Lebesgue

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 $\epsilon > 0$ $p', p'' \in \mathcal{P}[a,b]$ mit $\int\limits_a^b f(x)dx - s(p',f) < \frac{\epsilon}{2}$ und $S(p'',f) - \int\limits_a^b f(x)dx < \frac{\epsilon}{2}$. Dann folgt für $p = p' \cup p''$

$$S(p,f) - s(p,f) \leq S(p'',f) - \int\limits_a^b f(x)dx + \int\limits_a^b f(x)dx - s(p',f) < \epsilon.$$

Wenn es umgekehrt 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 für alle $\epsilon > 0$

$$0 \leq \inf_{p \in \mathcal{P}[a,b]} S(p,f) - \sup_{p \in \mathcal{P}[a,b]} s(p,f) \leq S(p_\epsilon, p) - s(p_\epsilon, f) < \epsilon.$$

Also gilt $\underline{\int} f = \overline{\int} f$. **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* $f \notin \mathcal{R}[a,b]$.

Satz 8.9. *Sei* f *eine reelle monotone Funktion auf* $[a,b]$, *dann ist* $f \in \mathcal{R}[a,b]$.

Beweis: Wir nehmen an, dass f monoton steigend ist.

$$\begin{aligned} S(p,f) - s(p,f) &= \sum_{i=1}^n (f(x_i) - f(x_{i-1}))\Delta x_i \\ &\leq \|p\| \sum_{i=1}^n f(x_i) - f(x_{i-1}) = \|p\|(f(b) - f(a)). \end{aligned}$$

Wenn also $\|p\| < \frac{\epsilon}{f(b)-f(a)+1}$ folgt $S(p,f) - s(p,f) < \epsilon$. Dann folgt der Satz aus dem Kriterium von Darboux. **q.e.d.**

Satz 8.10. $C_{\mathbb{R}}([a,b]) \subset \mathcal{R}[a,b]$.

Beweis: Wegen Satz 5.30 ist jede Funktion $f \in C_{\mathbb{R}}([a,b])$ auch gleichmäßig stetig. Dann gibt es 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 $S(p,f) - s(p,f) < \sum\limits_{i=1}^{n} \frac{\epsilon}{b-a}\Delta x_i = \epsilon$. Also folgt die Behauptung aus dem Kriterium von Darboux. **q.e.d.**

Satz 8.11. *Sei $f \in \mathcal{R}[a,b]$. Dann sind auch*

$$\begin{aligned}
&f^+ : [a,b] \to \mathbb{R}, x \mapsto f^+(x) = \max\{f(x), 0\} \text{ und} \\
&f^- : [a,b] \to \mathbb{R}, x \mapsto f^-(x) = \max\{-f(x), 0\} \text{ integrabel.}
\end{aligned}$$

Beweis: Offenbar sind

$$\begin{aligned}
O &\leq \sup_{x\in[x_{i-1},x_i]} f^+(x) - \inf_{x\in[x_{i-1},x_i]} f^+(x) &&\leq \sup_{x\in[x_{i-1},x_i]} f(x) - \inf_{x\in[x_{i-1},x_i]} f(x) \\
O &\leq \sup_{x\in[x_{i-1},x_i]} f^-(x) - \inf_{x\in[x_{i-1},x_i]} f^-(x) &&\leq \sup_{x\in[x_{i-1},x_i]} f(x) - \inf_{x\in[x_{i-1},x_i]} f(x)
\end{aligned}$$

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 der Satz aus dem Kriterium von Darboux. **q.e.d.**

Definition 8.12. *(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 8.13. *(Kriterium von Riemann) Eine beschränkte Funktion f auf einem kompakten Intervall $[a,b]$ ist genau dann Riemann–integrabel, 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*

$$|R(p,f,\xi) - A| < \epsilon. \quad \textit{Wenn das Krierium erfüllt ist, dann gilt } A = \int\limits_a^b f(x)dx.$$

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

$$\delta = \min\left\{\frac{\epsilon}{4n\|f\|_\infty}, \Delta x_1, \ldots, \Delta x_n\right\}.$$

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

$$S(p',f) - s(p',f) < S(p,f) - s(p,f) + 2\|f\|_\infty \cdot n \cdot \delta < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Dann gilt aber auch für alle Zwischenpunkte ξ

$$\left| R(p',f,\xi) - \int_a^b f(x)dx \right| \leq S(p',f) - s(p',f) < \epsilon.$$

q.e.d.

Korollar 8.14. *Sei $f \in \mathcal{R}[a,b]$. Dann gilt für alle $t \in [0,1]$*

$$\int_a^b f(x)dx = \lim_{n\to\infty} \sum_{i=1}^n \frac{b-a}{n} f(a + \frac{i-t}{n}(b-a)).$$

Beweis: Die Partitionen $p_n \in \mathcal{P}[a,b]$

$$p = \left\{x_i = a + \frac{i}{n}(b-a) \mid i = 0, \ldots, n\right\} \qquad \text{mit den Zwischenpunkten}$$

$$\xi_i = a + \frac{i-t}{n}(b-a) \text{ für } i = 1, \ldots, n \qquad \text{erfüllt } \|p_n\| = \frac{b-a}{n}.$$

Also folgt die Aussage aus dem Kriterium von Riemann. **q.e.d.**

Korollar 8.15. *Seien $f, g \in \mathcal{R}[a,b]$. Wenn f und g auf einer dichten Teilmenge von $[a,b]$ (z.B. den rationalen Zahlen in $[a,b]$) 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 8.16. *(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 $\mathcal{R}[a,b] \to \mathbb{R}, f \mapsto \int\limits_a^b f(x)dx$ ist $\mathbb{R}$-linear.*

(ii) *Monotonie: Für $f, g \in \mathcal{R}[a,b]$ folgt aus $f \leq g$ (d.h. $f(x) \leq g(x)$ für alle $x \in [a,b]$) $\int\limits_a^b f(x)dx \leq \int\limits_a^b g(x)dx$. Insbesondere gilt $|\int\limits_a^b f(x)dx| \leq \int\limits_a^b |f(x)|dx \leq (b-a)\|f\|_\infty$.*

(iii) *Normierung: $\int\limits_a^b 1dx = b - a$.*

(iv) *Stetigkeit: $f \in \mathcal{R}[a,b]$ und $g \in C_{\mathbb{R}}(\mathbb{R})$, dann ist $g \circ f \in \mathcal{R}[a,b]$.*

(v) *Intervall Additivität: Für jedes $c \in (a,b)$ gilt:*

$$f \in \mathcal{R}[a,b] \Leftrightarrow f \in \mathcal{R}[a,c] \cap \mathcal{R}[c,b] \text{ und } \int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

(vi) *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 $\lim\limits_{n\to\infty} a_n = a$ und $\lim\limits_{n\to\infty} b_n = b$. Dann konvergiert $\left(\int\limits_{a_n}^{b_n} f(x)dx\right)_{n\in\mathbb{N}}$ gegen $\int\limits_a^b f(x)dx$.*

(vii) *Gleichmäßige Konvergenz und Stetigkeit des Riemannintegrals: $\mathcal{R}[a,b]$ ist abgeschlossen in $B_{\mathbb{R}}([a,b])$ und $\displaystyle\int : \mathcal{R}[a,b] \to \mathbb{R}, f \mapsto \int_a^b f(x)dx$ ist stetig.*

Beweis:

(i) Für $f, g \in B_{\mathbb{R}}([a,b])$ und $p \in [a,b]$ gilt

$$S(p, f+g) \leq S(p,f) + S(p,g) \text{ und } \quad -s(p, f+g) \leq -s(p,f) - s(p,g)$$

Daraus und aus $f(x)g(x)-f(y)g(y) = g(x)(f(x)-f(y))+f(y)(g(x)-g(y))$ folgt

$$S(p, f+g) - s(p, f+g) \leq S(p,f) - s(p,f) + S(p,g) - s(p,g)$$
$$S(p, fg) - s(p, fg) \leq \|g\|_\infty(S(p,f) - s(p,f)) + \|f\|_\infty(S(p,g) - s(p,g))$$

Also folgt wegen dem Darbouxkriterium aus $f, g \in \mathcal{R}[a,b]$ sowohl $f+g \in \mathcal{R}[a,b]$ als auch $fg \in \mathcal{R}[a,b]$. Die stetigen und also auch die konstanten Funktionen sind wegen Satz 8.10 in $\mathcal{R}[a,b]$. Die Linearität des Riemannintegrals folgt aus der Linearität der Riemannsummen und den Rechenregeln für Folgen.

(ii) Aus $f \leq g$ folgt $\int\limits_a^b f(x)dx = \sup\limits_{p \in \mathcal{P}[a,b]} s(p,f) \leq \sup\limits_{p \in \mathcal{P}[a,b]} s(p,g) = \int\limits_a^b g(x)dx.$

(iii) Für $f = 1$(konstant) gilt $S(p,1) - s(p,1) = b - a$ für alle $p \in \mathcal{P}[a,b]$.

(iv) 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

$$g : [-\|f\|_\infty, \|f\|_\infty] \to \mathbb{R}, x \mapsto g(x).$$

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 $\sup\limits_{x \in [x_{i-1}, x_i]} f(x) - \inf\limits_{x \in [x_{i-1}, x_i]} f(x) < \delta$ gilt und die Summe über alle Teilintervalle, auf denen $\sup\limits_{x \in [x_{i-1}, x_i]} f(x) - \inf\limits_{x \in [x_{i-1}, x_i]} f(x) \geq \delta$ 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.

(v) 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.

(vi) folgt aus (v) und (ii).

(vii) Aus dem Beweis von (i) folgt für $f, g \in B_\mathbb{R}([a,b])$ und $p \in \mathcal{P}[a,b]$

$$|S(p,f) - s(p,f) - (S(p,g) - s(p,g))| \leq S(p, f-g) - s(p, f-g) \leq 2(b-a)\|f-g\|_\infty.$$

Deshalb erfüllt jeder Grenzwert in $B_\mathbb{R}([a,b])$ einer Folge in $\mathcal{R}[a,b]$ das Kriterium von Darboux. Also ist $\mathcal{R}[a,b]$ abgeschlossen in $B_\mathbb{R}([a,b])$.

Andererseits folgt für $f \in \mathcal{R}[a,b]$ aus der Monotonie $\left|\int\limits_a^b f(x)dx\right| \leq (b-a)\|f\|_\infty$.

Also ist die Abbildung $\int : \mathcal{R}[a,b] \to \mathbb{R}, f \mapsto \int\limits_a^b f(x)dx$ stetig. **q.e.d.**

8.3 Differentation und Integration

Hauptsatz der Differential- und Integralrechnung 8.17. *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*

$$\int\limits_a^b f(x)dx = \int\limits_a^b F'(x)dx = F(b) - F(a).$$

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

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

q.e.d.

Beispiel 8.18. (i) *Sei $F \in C^1[a,b]$. Dann ist F' Riemann–integrabel und es gilt für alle $x \in [a,b]$*

$$\int\limits_a^x F'(t)dt = F(x) - F(a).$$

(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 Riemann–integrabel 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 Riemann–integrabel. Offenbar gilt* $F(x) = \int\limits_0^x f(t)dt = |x|$. *Also sind nicht alle Integrale von Riemann–integrablen Funktionen differenzierbar.*

Satz 8.19. *Sei* $f \in \mathcal{R}[a,b]$ *im Punkt* $x_0 \in (a,b)$ *stetig. Dann ist* $F(x) = \int\limits_a^x f(t)dt$ *im Punkt* x_0 *differenzierbar und es gilt* $F'(x_0) = f(x_0)$.

Beweis:

$$\left|\frac{F(x)-F(x_0)}{x-x_0} - f(x_0)\right| \leq \frac{1}{|x-x_0|}\int\limits_{x_0}^{x} |f(t)-f(x_0)|dt$$

Weil f im Punkt stetig ist folgt aus den Eigenschaften des Integrals (ii), 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 $F(x) = \int\limits_a^x f(t)dt$ 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 8.20. *Sei* $f \in \mathcal{R}[a,b]$. *Dann ist* $F(x) = \int\limits_a^x f(t)dt$ *Lipschitz–stetig mit Lipschitzkonstante* $\|f\|_\infty$.

Beweis:

$$|F(x)-F(y)| = \left|\int\limits_x^y f(t)dt\right| \leq |y-x|\|f\|_\infty$$

q.e.d.

Definition 8.21. *(Stammfunktion) Eine differenzierbare Funktion* F *mit* $F' = f$ *heißt Stammfunktion von* f. *Die Differenz zweier Stammfunktionen ist eine konstante Funktion. Wir bezeichnen eine Stammfunktion von* f *als* $\int f(x)dx$.

Beispiel 8.22. (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 8.23. *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 also

$$\int f(\phi(t))\phi'(t)dt = \left(\int f(x)dx\right) \circ \phi + C.$$

Beweis: Sei F eine Stammfunktion von f. Dann ist F wegen Satz 8.19 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 8.24. *(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*

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

Für die Stammfunktionen gilt also

$$\int f(x)dx = \left(\int f(\phi(t))\phi'(t)dt\right) \circ \phi^{-1} + C.$$

q.e.d.

Beispiel 8.25. **(i)**

$$\int\limits_a^b f(\alpha t + \beta)dt = \frac{1}{\alpha} \int\limits_{\alpha a+\beta}^{\alpha b+\beta} f(x)dx$$

(ii)

$$\int R\left(x, \sqrt[n]{ax+b}\right) dx = \int R\left(\frac{t^n - b}{a}, t\right) \frac{n}{a} t^{n-1} dt + C$$

mit der Substitution $t = \sqrt[n]{ax+b}$ für $n \in \mathbb{N}, a, b \in \mathbb{R}, a \neq 0$.

$$ax + b = t^n \Rightarrow x = \frac{t^n - b}{a} \text{ und } dx = \frac{nt^{n-1}}{a} dt.$$

(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 $t \in \mathbb{R}^{\pm}$

$$\sqrt{x^2 - 1} = \sinh t \text{ und } dx = \pm \sinh t dt.$$

(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*

$$\begin{aligned} \frac{1-t^2}{1+t^2} &= \frac{\cos^2\left(\frac{x}{2}\right) - \sin^2\left(\frac{x}{2}\right)}{\cos^2\left(\frac{x}{2}\right) + \sin^2\left(\frac{x}{2}\right)} = \cos(x) \text{ und} \\ \frac{2t}{1+t^2} &= \frac{2\cos\left(\frac{x}{2}\right)\sin\left(\frac{x}{2}\right)}{\cos^2\left(\frac{x}{2}\right) + \sin^2\left(\frac{x}{2}\right)} = \sin(x). \end{aligned}$$

Partielle Integration 8.26. *Seien* $f, g \in C_{\mathbb{R}}[a,b]$ *differenzierbar mit* $f', g' \in \mathcal{R}[a,b]$. *Dann gilt*

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

Für die Stammfunktionen gilt also

$$\int f(x)g'(x)dx = fg - \int f'(x)g(x)dx + C.$$

Beweis folgt aus dem Hauptsatz der Differentialrechnung und der Produktregel. **q.e.d.**

Beispiel 8.27. (i) $\int \ln(x)dx = x\ln(x) - \int x\frac{1}{x}dx = x(\ln(x)-1) + C$.

(ii) $\int \sqrt{1-x^2}dx = x\sqrt{1-x^2} + \int \frac{x^2}{\sqrt{1-x^2}}dx + C$
$= x\sqrt{1-x^2} - \int \frac{1-x^2}{\sqrt{1-x^2}}dx + \int \frac{1}{\sqrt{1-x^2}}dx + C$
$\Rightarrow \int \sqrt{1-x^2}dx = \frac{x\sqrt{1-x^2}}{2} + \frac{\arcsin(x)}{2} + C$ *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 8.28. *(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 8.29. *(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 läß 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

$$x \mapsto \frac{P(x)}{q(x)} - (a_n + b_n(x - x_0)) - \ldots - (a_1 + b_1(x - x_0))(x - x_0)^{n-1}(x - x_0)^{n-1}$$

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:

$$\frac{P(x)}{Q(x)} = \frac{p(x)}{q(x)} + \frac{a_1 - b_1 x_0 + b_1 x}{(x - x_0)(x - \bar{x}_0)} + \ldots + \frac{a_n - b_n x_0 + b_n x}{(x - x_0)^n (x - \bar{x}_0)^n}.$$

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. $$\int \frac{dx}{(x - x_0)^l} = \begin{cases} \ln|x - x_0| + C & \text{für } l = 1 \\ \frac{-1}{(l-1)(x-x_0)^{l-1}} + C & \text{sonst.} \end{cases}$$

$$\int \frac{(a + bx)dx}{(x^2 + px + q)^l} = \begin{cases} \frac{b}{2}\ln(x^2 + px + q) + \left(a - \frac{bp}{2}\right) \int \frac{dx}{x^2+px+q} + C & \text{für } l = 1 \\ \frac{-b}{2(l-1)(x^2+px+q)^{l-1}} + \left(a - \frac{bp}{2}\right) \int \frac{dx}{(x^2+px+q)^{l-1}} + C & \text{sonst.} \end{cases}$$

$$\int \frac{dx}{(x^2 + px + q)^l} = \begin{cases} \frac{2}{\sqrt{4q-p^2}} \arctan\left(\frac{2x+p}{\sqrt{4q-p^2}}\right) + C & l = 1 \\ \frac{2x+p}{(l-1)(4q-p^2)(x^2+px+q)^{l-1}} + \frac{2(2l-3)}{(l-1)(4q-p^2)} \int \frac{dx}{(x^2+px+q)^{l-1}} + C & l \neq 1. \end{cases}$$

Damit können wir alle rationalen Funktionen integrieren.

Satz 8.30. *(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

$$\begin{aligned}
f(x) &= \sum_{m=0}^{n} \frac{f^{(m)}(x_0)}{m!}(x-x_0)^m + (x-x_0)^{n+1} \int_0^1 \frac{(1-t)^n}{n!} f^{n+1}(x_0 + t(x-x_0))dt \\
&= \sum_{m=0}^{n} \frac{f^{(m)}(x_0)}{m!}(x-x_0)^m - \frac{(x-x_0)^{n+1}(1-t)^{n+1}}{(n+1)!} f^{n+1}(x_0 + t(x-x_0))\Big|_{t=0}^{t=1} \\
&\quad +(x-x_0)^{n+2} \int_0^1 \frac{(1-t)^{n+1}}{(n+1)!} f^{n+2}(x_0 + t(x-x_0))dt \\
&= \sum_{m=0}^{n+1} \frac{f^{(m)}(x_0)}{m!}(x-x_0)^m + (x-x_0)^{n+2} \int_0^1 \frac{(1-t)^{n+1}}{n+1} f^{n+2}(x_0 + t(x-x_0))dt
\end{aligned}$$

Also gilt die Aussage für $n+1$. **q.e.d.**

Satz 8.31.* *(Mittelwertsatz der Integralrechnung) Seien $f, g \in \mathcal{R}[a,b]$ und $g \geq 0$. 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 $\inf\limits_{x\in[a,b]} f(x) \le f \le \sup\limits_{x\in[a,b]} f(x)$ folgt die erste Aussage aus der Monotonie. Wenn f stetig ist folgt aus dem Mittelwertsatz für $F(x) = \int\limits_a^x f(t)dt$, dass es ein $x_0 \in (a,b)$ gibt mit $f(x_0) = F'(x_0) = \frac{F(b)-F(a)}{b-a}$. Dann folgt $f(x_0) = \frac{1}{b-a}\int\limits_a^b f(x)dx$.
q.e.d.

8.5 Uneigentliches Integral

In diesem Abschnitt erweitern wir den Begriff des Riemannintegrals auf offene und unbeschränkte Intervalle.

Definition 8.32. *Eine Funktion f heisst Riemann–integrabel auf dem offenen (nicht notwendigerweise beschränkten) Intervall $(a,b) \subset \mathbb{R}$, wenn f Riemann–integrabel ist auf allen kompakten Teilintervallen, und wenn für ein $c \in (a,b)$ beide Grenzwerte $\lim\limits_{x\to a+}\int\limits_x^c f(t)dt$ und $\lim\limits_{x\to b-}\int\limits_c^b f(t)dt$ existieren.*

Beispiel 8.33. **(i)** $\displaystyle\int\limits_1^\infty \frac{1}{x^\alpha}dx.$ $\displaystyle\int \frac{1}{x^\alpha}dx = \begin{cases} \frac{-1}{(\alpha-1)x^{\alpha-1}} + C \text{ für } \alpha \neq 1 \\ \ln|x| + C \text{ für } \alpha = 1 \end{cases}$. *Also existiert der Grenzwert $\lim\limits_{x\to\infty-}\int\limits_1^x \frac{1}{t^\alpha}dt$ genau dann, wenn $\alpha > 1$. Dann gilt* $\displaystyle\int\limits_1^\infty \frac{1}{x^\alpha}dx = \frac{1}{\alpha-1}.$

(ii) $\displaystyle\int\limits_0^1 \frac{1}{x^\alpha}dx.$ $\displaystyle\int \frac{1}{x^\alpha}dx = \begin{cases} \frac{-1}{(\alpha-1)x^{\alpha-1}} + C \text{ für } \alpha \neq 1 \\ \ln|x| + C \text{ für } \alpha = 1 \end{cases}$. *Dann existiert der Grenzwert $\lim\limits_{x\to 0+}\int\limits_x^1 \frac{1}{x^\alpha}dx$ genau dann, wenn $\alpha < 1$. In diesem Fall gilt* $\displaystyle\int\limits_0^1 \frac{1}{x^\alpha}dx = \frac{1}{1-\alpha}.$

Wegen (i) folgt dann, dass $\displaystyle\int\limits_0^\infty \frac{1}{x^\alpha}dx$ 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*

$$\lim_{x\to-\infty+} \int\limits_{x}^{0} \frac{1}{1+t^2}dt = \frac{\pi}{2} = \lim_{x\to\infty-} \int\limits_{0}^{x} \frac{1}{1+t^2}dt. \quad \textit{Dann folgt} \int\limits_{-\infty}^{\infty} \frac{1}{1+x^2}dx = \pi.$$

Verschiedenen Kriterien helfen zu entscheiden, ob diese Grenzwerte existieren:

Cauchykriterium: $\int\limits_a^\infty f(x)dx$ existiert genau dann, wenn es zu jedem $\epsilon > 0$ ein $b \in \mathbb{R}$ gibt, so dass für alle $d > c > b$ gilt $\left|\int\limits_c^d f(x)dx\right| < \epsilon$.

Monotoniekriterium: Wenn $f \geq 0$, dann existiert $\int\limits_a^\infty f(x)dx$ genau dann, wenn $F(x) = \int\limits_a^x f(t)dt$ beschränkt ist.

Majorantenkriterium: Wenn $f \geq 0$ und $f \leq g$, dann existiert $\int\limits_a^\infty f(x)dx$, wenn $\int\limits_a^\infty g(x)dx$ existiert.

Definition 8.34. *Eine Funktion f auf einem offenen (unbeschränkten) Intervall heißt absolut Riemann–integrabel, wenn $|f|$ Riemann–integrabel ist.*

Wegen der Dreiecksungleichung gilt dann $\left|\int\limits_a^b f(x)dx\right| \leq \int\limits_a^b |f(x)|dx.$

Alle absolut Riemann–integrablen Funktionen sind also auch Riemann–integrabel.

Satz 8.35. *(Höldersche Ungleichung) Seien $\frac{1}{p} + \frac{1}{q} = 1$ und $|f|^p$ und $|g|^q$ Riemann–integrabel auf einem Intervall I. Dann ist $f \cdot g$ absolut Riemann–integrabel und es gilt*

$$\int\limits_I |fg|dx \leq \left(\int\limits_I |f|^p dx\right)^{1/p} \left(\int\limits_I |g|^q dx\right)^{1/q}.$$

Beweis: Wegen der Monotonie genügt es die Höldersche Ungleichung für nicht negative Funktionen zu zeigen. Für kompakte Intervalle folgt diese Ungleichung aus der Hölderschen Ungleichung für Summen und Korollar 8.14. Für offene Intervalle folgt sie dann aus dem Majorantenkriterium. **q.e.d.**

Satz 8.36. *(Integralkriterium für Reihen) Sei* $f : [1, \infty) \to \mathbb{R}$ *monoton fallend und* $f \geq 0$. *Dann konvergiert* $(\sum f(n))_{n \in \mathbb{N}}$ *genau dann, wenn* $\int\limits_1^\infty f(x)dx$ *existiert und endlich ist.*

Beweis: $p_n \in \mathcal{P}[1, n]$ sei die Partition $\{1, 2, \ldots, n\}$. Dann ist offenbar $s(p_n, f) = \sum\limits_{m=2}^{n} f(m)$ und $S(p_n, f) = \sum\limits_{m=1}^{n-1} f(m)$. Also gilt

$$\int\limits_1^{n+1} f(x)dx \leq \sum_{m=1}^{n} f(m) \leq f(1) + \int\limits_1^{n} f(x)dx.$$

Dann folgt die Aussage aus dem Majorantenkriterium. **q.e.d.**

Beispiel 8.37. **(i)** $\left(\sum \frac{1}{n^s}\right)_{n \in \mathbb{N}}$ *ist genau dann konvergent, wenn* $\int\limits_1^\infty \frac{1}{x^s}dx$ *existiert. Also für* $s > 1$. *Dann gilt*

$$\frac{1}{s-1} < \sum_{n=1}^{\infty} \frac{1}{n^s} < 1 + \frac{1}{s-1}.$$

(ii) $\left(\sum \frac{1}{(n+1)\ln^s(n+1)}\right)_{n \in \mathbb{N}}$ *ist genau dann konvergent, wenn* $\int\limits_2^\infty \frac{1}{x \ln^s(x)} dx = \int\limits_{\ln(2)}^\infty \frac{1}{x^s} dx$ *existiert, also für* $s > 1$.

Kapitel 9

Differentialrechnung von Funktionen mehrerer Veränderlicher

9.1 Normierte Räume und 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 $\mathbb{K}$, der hier entweder für $\mathbb{R}$ oder für $\mathbb{C}$ steht. Wir wiederholen zunächst die Definition eines normierten Vektorraumes.

Definition 9.1. *Ein Vektorraum V ist eine Menge V zusammen mit den Abbildungen:*

$$+ : V \times V \to V, \qquad (v, w) \mapsto v + w \qquad \cdot : \mathbb{K} \times V \to V, (\lambda, v) \qquad \mapsto \lambda v,$$

die die Axiome A1 der Addition und das Distributivgesetz A3 erfüllen und $\lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v$ für alle $\lambda, \mu \in \mathbb{K}, v \in V$. Ein normierter Vektorraum ist ein Vektorraum V zusammen mit einer Norm, also einer Abbildung

$$\| \cdot \| : V \to \mathbb{R}, v \mapsto \|v\|$$

die folgende drei Bedingungen erfüllt:

(i) $\|v\| \geq 0$ *und* $\|v\| = 0$ *nur für* $v = 0$*(Positivität)*

(ii) $\|\lambda v\| = |\lambda| \|v\|$ *für alle* $\lambda \in \mathbb{K}, v \in V$.

(iii) $\|v+w\| \leq \|v\| + \|w\|$ *für alle* $v, w \in V$ *(Dreiecksungleichung)*

Jede Norm definiert durch

$$d : V \times V \to \mathbb{R}, (v, w) \mapsto d(v, w) = \|v - w\|$$

eine Metrik auf V. Zwei Normen $\| \quad \|_1$ und $\| \quad \|_2$ heißen äquivalent, wenn es Konstanten $C, C' > 0'$ gibt, so dass für alle $v \in V$ gilt

$$C\|v\|_1 \leq \|v\|_2 \leq C'\|v\|_1.$$

Beispiel 9.2. *Auf den Vektorräumen* $\mathbb{K}^n$ *haben wir für* $1 \leq p \leq \infty$ *die Normen*

$$\| \cdot \|_p : \mathbb{K}^n \to \mathbb{R}, v \mapsto \|v\|_p = \begin{cases} \left(\sum\limits_{i=1}^{n} |v_i|^p \right)^{1/p} & \textit{für } p < \infty \\ \sup\{|v_1|, \dots |v_n|\} & \textit{für } p = \infty. \end{cases}$$

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.

Satz 9.3. *Auf* $\mathbb{K}^n$ *sind alle Normen paarweise äquivalent.*

Beweis: Offenbar ist die Relation zwischen Normen äquivalent zu sein eine Äquivalenzrelation, also insbesondere transitiv. Denn aus

$$\begin{array}{ccc} C_1\|v\|_1 \leq \|v\|_2 \leq C_2\|v\|_1 \text{ und} & & C_3\|v\|_2 \leq \|v\|_3 \leq C_4\|v\|_2 \\ \text{folgt} & & C_1C_3\|v\|_1 \leq \|v\|_3 \leq C_2C_4\|v\|_1. \end{array}$$

Deshalb genügt es zu zeigen, dass alle Normen äquivalent sind zu $\| \cdot \|_\infty$. Sei also $\| \cdot \|$ eine beliebige Norm und $e_1, \dots, e_n$ die Basis, deren i-tes Element die Koordinaten δ_{ij} hat, d.h. alle Koordinaten bis auf die i-te verschwinden und die i-te Koordinate ist 1. Für jedes $v = (v_1, \dots, v_n) \in \mathbb{K}^n$ gilt dann $v = v_1e_1 + \dots + v_ne_n$. Wegen der Dreiecksungleichung gilt dann für alle $i = 1, \dots, n$

$$|v_i| \cdot \|e_i\| \leq \|v\| \leq |v_1| \cdot \|e_1\| + \dots + |v_n| \cdot \|e_n\|.$$

Also folgt für alle $v \in \mathbb{K}^n$

$$\min\{\|e_1\|, \dots, \|e_n\|\}\|v\|_\infty \leq \|v\| \leq (\|e_1\| + \dots + \|e_n\|)\|v\|_\infty.$$ **q.e.d.**

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 Kugel um die Null der anderen Norm enthalten muß.

Definition 9.4. *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*

$$A(v + w) = Av + Aw \text{ und } \qquad A(\lambda v) = \lambda Av.$$

Satz 9.5. *Seinen 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 $\|Av\| \leq C\|v\|$.*

(v) *A ist auf $B(0, 1)$ beschränkt, d.h. es gibt ein $C > 0$, so dass für alle $\|v\| < 1$ gilt $\|Av\| \leq C$.*

Beweis:

(i)⇒(v): Wenn A in 0 stetig ist, dann muß das Urbild jeder Kugel $B(0, \epsilon) \subset W$ eine Kugel $B(0, \delta) \subset V$ enthalten. Also gibt es ein $\delta > 0$, so dass aus $\|v\| < \delta$ folgt $\|Av\| < 1$. Wegen der Linearität folgt dann aber aus $\|v\| < 1$ auch $\|Av\| = \frac{1}{\delta}\|A\delta v\| < \frac{1}{\delta}$. Also ist (v) erfüllt.

(v)⇒(iv): Wegen der Linearität folgt aus (v), dass für alle $v \in V$ gilt

$$\|Av\| = A\left(2\|v\| \cdot \frac{v}{2\|v\|}\right) = 2\|v\| \cdot A\left(\frac{v}{2\|v\|}\right) \leq 2C\|v\|.$$

(iv)⇒(iii): Seien $v, w \in V$, dann folgt aus (iv) $\|A(v - w)\| \leq C\|v - w\|$. Also ist A sogar Lipschitz–stetig mit Lipschitzkonstante C. Dann gilt auch (iii).

(iii)⇒(ii) und (ii)⇒(i): Sind offensichtlich. **q.e.d.**

Satz 9.6. *Jede lineare Abbildung $A : \mathbb{K}^n \to \mathbb{K}^m$ ist stetig.*

Beweis: Wir benutzen wieder die Basis $e_1, \ldots, e_n$ von $\mathbb{K}^n$. Dann gilt für alle $v \in \mathbb{K}^n$

$$\begin{aligned}\|Av\| &\leq |v_1| \cdot \|Ae_1\| + \ldots + |v_n| \cdot \|Ae_n\| \\ &\leq \|v\|_1 \max\{\|Ae_1\|, \ldots, \|Ae_n\|\}.\end{aligned}$$

Also folgt die Aussage aus Satz 9.3. **q.e.d.**

Definition 9.7. *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:*

$$+ : \mathcal{L}(V, W) \times \mathcal{L}(V, W) \to \mathcal{L}(V, W), (A, B) \mapsto A + B$$

$$\mathit{mit}\ A + B : V \to W, v \mapsto Av + Bv$$

$$\cdot : \mathbb{K} \times \mathcal{L}(V, W) \to \mathcal{L}(V, W), (\lambda, A) \mapsto \lambda A$$

$$\mathit{mit}\ \lambda A : V \to W, v \mapsto \lambda Av$$

$$\|\cdot\| : \mathcal{L}(V, W) \to \mathbb{R}, A \mapsto \|A\| = \sup_{v \in B(0,1) \subset V} \|Av\| = \sup_{v \in \overline{B(0,1)}} \|Av\|.$$

Satz 9.8. *$\mathcal{L}(V, W)$ 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 $\overline{B(0,1)} = \{v \in \mathbb{K}^n | \quad \|v\| \leq 1\}$ 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 $Av = \|v\| \cdot A\left(\frac{v}{\|v\|}\right)$ 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 den beschränkten stetigen Funktionen auf $\overline{B(0,1)} \subset V$. So folgt z.B. aus der Konvergenz einer Folge $(A_n)_{n \in \mathbb{N}}$ in $\mathcal{L}(V, W)$ die punktweise Konvergenz auf $\overline{B(0,1)}$ (und damit sogar die punktweise Konvergenz auf V). Die Konvergenz in $\mathcal{L}(V, W)$ beinhaltet sogar die gleichmäßige Konvergenz auf $\overline{B(0,1)} \subset V$.

Definition 9.9. *Ein vollständiger normierter Vektorraum heißt Banachraum.*

Satz 9.10. *Seien V ein normierter Vektorraum und W ein Banachraum. Dann ist $\mathcal{L}(V, W)$ ein Banachraum.*

Beweis: Wir müssen wegen Satz 9.8 nur noch zeigen, dass $\mathcal{L}(V,W)$ vollständig ist. Sei also $(A_n)_{n\in\mathbb{N}}$ 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 also konvergiert. Wir definieren also als den Grenzwert von $(A_n)_{n\in\mathbb{N}}$ die Abbildung von V nach W, die für alle $v \in V$ durch den Grenzwert

$$A : V \to W, v \mapsto Av = \lim_{n\to\infty} A_n v$$

gegeben ist. Wir müssen dann noch zeigen, dass $(A_n)_{n\in\mathbb{N}}$ gegen A konvergiert und dass A linear stetig ist. Weil $(A_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist, gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $n, m \geq N$ auch $\|A_n - A_m\| < \frac{\epsilon}{2}$ gilt. Für jedes $v \in \overline{B(0,1)} \subset V$ gibt es ein $m \geq N$ mit $\|Av - A_n v\| < \frac{\epsilon}{2}$. Daraus folgt

$$\|(A - A_n)v\| \leq \|(A - A_m)v\| + \|(A_m - A_n)v\| < \frac{\epsilon}{2} + \frac{\epsilon}{2}\|v\| \leq \epsilon.$$

Also konvergiert $(A_n)_{n\in\mathbb{N}}$ gegen A. Aus der Linearität von A_n folgt für alle $n \in \mathbb{N}$

$$\begin{aligned}
&\|A(v+w) - (Av + Aw)\| \leq \\
&\quad \leq \|(A - A_n)(v+w) - (A - A_n)v - (A - A_n)w))\| + \|A_n(v+w) - (A_n v + A_n)\| \\
&\qquad\qquad \leq \|A - A_n\| \left(\|v + w\| + \|v\| + \|w\|\right), \text{ und} \\
&\qquad \|\lambda Av - A(\lambda v)\| \leq \|\lambda(A - A_n)v - (A - A_n)(\lambda v)\| + \|\lambda A_n v - A_n(\lambda v)\| \\
&\qquad\qquad\qquad \leq \|A - A_n\| \left(|\lambda| \cdot \|v\| + \|\lambda v\|\right).
\end{aligned}$$

Im Grenzwert $n \to \infty$ konvergieren die rechten Seiten aber gegen Null, so dass A linear ist. Weil die Konvergenz in $\mathcal{L}(V,W)$ aber die gleichmäßige Konvergenz auf $\overline{B(0,1)} \subset V$ ist, können wir um die Stetigkeit von A zu zeigen wie im Satz 5.37 (iii) wieder den $\epsilon/3$-Trick benutzen. Wir benutzen aber Satz 9.5. Weil jede Cauchyfolge beschränkt ist, gibt es ein $M > 0$, so dass $\|A_n\| \leq M$ für $n \in \mathbb{N}$ gilt. Für alle $v \in V$ folgt

$$\|Av\| \leq \|(A - A_n)v\| + \|A_n v\| \leq \left(\|A - A_n\| + M\right)\|v\|.$$

Jetzt wählen wir ein $n \in \mathbb{N}$, so dass $\|A - A_n\|$ kleiner ist als 1. Dann folgt $\|Av\| \leq (M+1)\|v\|$ und A ist wegen Satz 9.5 stetig. **q.e.d.**

Satz 9.11. *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 $\circ : \mathcal{L}(U,V) \times \mathcal{L}(V,W) \to \mathcal{L}(U,W), (A,B) \mapsto B \circ A$ eine stetige Abbildung von dem kartesischen Produkt der metrischen Räume $\mathcal{L}(U,V)$ und $\mathcal{L}(V,W)$ in den metrischen Raum $\mathcal{L}(U,W)$.*

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 9.5. Für zwei normierte Vektorräume V, W mit Normen $\| \cdot \|_V$ und $\| \cdot \|_W$ ist

$$\| \cdot \|_{V \times W} : V \times W \to \mathbb{R}, (v, w) \mapsto \|v\|_V + \|w\|_W$$

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 aber

$$\begin{aligned}
\|B \circ A - B' \circ A'\| &= \|B \circ A - B \circ A' + B \circ A' - B' \circ A'\| \\
&= \|B \circ (A - A') + (B - B') \circ A'\| \\
&\leq \|B\| \cdot \|A - A'\| + \|B - B'\| \cdot \|A'\| \\
&\leq (\|A - A'\| + \|B - B'\|)(\|B\| + \|A'\|) \\
&\leq (\|A - A'\| + \|B - B'\|)(\|B\| + \|A\| + \|A - A'\|) \\
&\leq (\|(A, B) - (A', B')\|(\|B\| + \|A\| + \|(A, B) - (A', B')\|).
\end{aligned}$$

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 einfach nur als BA.

Definition 9.12. *Sei V ein normierter Vektorraum. Dann ist $\mathcal{L}(V) = \mathcal{L}(V, V)$ eine Algebra, d.h. ein Vektorraum mit einer Abbildung*

$$\circ : \mathcal{L}(V) \times \mathcal{L}(V) \to \mathcal{L}(V), (A, B) \mapsto AB$$

die bilinear ist, d.h. sie erfüllt

$$\begin{aligned}
(A + A')B = AB + A'B \text{ und } (\lambda A)B = \lambda(AB) \\
A(B + B') = AB + AB' \text{ und } A(\lambda B) = \lambda(AB).
\end{aligned}$$

$\mathcal{L}(V)$ ist sogar eine normierte Algebra, d.h. die Norm $\| \cdot \|$ erfüllt $\|AB\| \leq \|A\| \cdot \|B\|$. Wenn V ein Banachraum ist, dann ist $\mathcal{L}(V)$ eine Banachalgebra.

Satz 9.13. *(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 $(\mathbb{1} - A)^{-1} = \sum\limits_{n=0}^{\infty} A^n$.*

Beweis: Weil $\|A^n\| \leq \|A\|^n$ ist für $\|A\| < 1$ die Reihe $\left(\sum A^n\right)_{n \in \mathbb{N}_0}$ eine Cauchyfolge und es gilt

$$\left\| \sum_{n=0}^{N} A^n \right\| \leq \frac{1}{1 - \|A\|}.$$

Wegen Satz 9.10 konvergiert diese Reihe gegen einen Operator $B \in \mathcal{L}(V)$. Offenbar gilt

$$(\mathbb{1} - A)B = \sum_{n=0}^{\infty} A^n - \sum_{n=1}^{\infty} A^n = \mathbb{1}$$

und genauso

$$B(\mathbb{1} - A) = \sum_{n=0}^{\infty} A^n - \sum_{n=1}^{\infty} A^n = \mathbb{1}.$$

Also ist $(\mathbb{1} - A)$ invertierbar und es gilt

$$(\mathbb{1} - A)^{-1} = \sum_{n=0}^{\infty} A^n.$$

Insbesondere gilt $$\left\|(\mathbb{1} - A)^{-1}\right\| \leq \frac{1}{1 - \|A\|}.$$ **q.e.d.**

Alle Potenzreihenfunktionen $f(x) = \sum_{n \in \mathbb{N}_0} a_n x^n$ mit Konvergenzradius $R > 0$ definieren dann offenbar eine Abbildung

$$f : \{A \in \mathcal{L}(V) \mid \|A\| < R\} \to \mathcal{L}(V), A \mapsto f(A) = \sum_{n=0}^{\infty} a_n A^n.$$

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 9.14. *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 9.15. (i) *Zeige, dass für jedes $A \in \mathcal{L}(V)$, die Abbildung*

$$D_A : \mathcal{L}(V) \to \mathcal{L}(V), B \mapsto AB - BA$$

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*

$$\{C \in \mathcal{L}(\mathcal{L}(V)) \mid C(AB) = C(A)C(B) \text{ für alle } A, B \in \mathcal{L}(V)\} \subset \mathcal{L}(\mathcal{L}(V)).$$

(iii) *Sei V ein Banachraum. Zeige, dass für alle $B \in \mathcal{L}(V)$ gilt*

$$\exp(D_A)B = \exp(A) \cdot B \cdot \exp(-A).$$

9.2 Ableitungen von $f : X \to Y$

Definition 9.16. *(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 eine lineare stetige Abbildung $A : X \to Y$ gibt, so dass die Abbildung*

$$U \to \mathbb{R}, \quad x \mapsto \begin{cases} \frac{\|f(x)-f(x_0)-A(x-x_0)\|}{\|x-x_0\|} & \text{für } x = x_0 \\ 0 & \text{für } x = x_0 \end{cases}$$

stetig in x_0 ist.

Die lineare stetige Abbildung A heißt dann Ableitung von f an der Stelle x_0 und wird mit $f'(x_0)$ oder $\frac{df}{dx}(x_0)$ bezeichnet. Wenn A und A' beide diese Bedingung erfüllen, dann gilt $\|(A - A')(x - x_0)\| = 0$ für alle $x - x_0 \in U$. Also ist dann $\|A - A'\| = 0$ und $A = A'$, so dass die Ableitung einer differenzierbaren Funktion eindeutig ist.

Satz 9.17. *Sei $f : U \subset X \to Y$ in $x_0 \in U$ differenzierbar. Dann ist f im Punkt 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 aber f auch stetig. **q.e.d.**

Beispiel 9.18. (i) *Sei f konstant. Dann ist*

$$\frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} = 0 \text{ für } x \neq x_0.$$

Also ist f differenzierbar mit $f'(x_0) = 0$.

(ii) *Sei $f : X \to Y$ eine lineare stetige Abbildung. Dann ist*

$$\frac{\|f(x) - f(x_0) - f(x - x_0)\|}{\|x - x_0\|} = 0 \text{ für } x \neq x_0.$$

Also ist f in x_0 differenzierbar und $f'(x_0) = f$.

(iii) *Sei $f : (a, b) \to \mathbb{R}$ eine reelle Funktion. Dann ist f als Funktion von der Teilmenge (a, b) des normierten Vektorraumes $\mathbb{R}$ in den normierten Vektorraum $\mathbb{R}$ genau dann in x_0 differenzierbar, wenn f als reelle Funktion in x_0 differenzierbar ist, weil gilt*

$$\frac{\|f(x) - f(x_0) - A(x - x_0)\|}{\|x - x_0\|} = \left\| \frac{f(x) - f(x_0)}{x - x_0} - A \right\|$$

für alle $x \neq x_0$.

Satz 9.19. (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*

$$(f+g)'(x_0) + g'(x_0) bzw. (\lambda f)'(x_0) = \lambda f'(x_0).$$

(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*

$$\begin{gathered}(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0) \qquad mit \\ (f'(x_0) \cdot g(x_0))\, x = (f'(x_0)x)\, g(x_0) \;\; und \quad (f(x_0) \cdot g'(x_0))\, x = f(x_0) \cdot (g'(x_0)x)\,.\end{gathered}$$

(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 und es gilt*

$$(g \circ f)'(x_0) = g'(f(x_0)) \circ f'(x_0).$$

Beweis:

(i) Wegen der Dreiecksungleichung gilt

$$\begin{aligned}&\frac{\|f(x) + g(x) - f(x_0) - g(x_0) - f'(x_0)(x - x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|} \\ &\qquad \leq \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} + \frac{\|g(x) - g(x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}\end{aligned}$$

Daraus folgt (i).

(ii) Weil in einer normierten Algebra $\|AB\| \leq \|A\| \cdot \|B\|$ gilt, folgt

$$\begin{aligned}&\frac{\|f(x) \cdot g(x) - f(x_0) \cdot g(x_0) - f'(x_0)(x - x_0) g(x_0) - f(x_0) g'(x_0)(x - x_0)\|}{\|x - x_0\|} \\ &\qquad \leq \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} \|g(x)\| + \|f'(x_0)\| \cdot \|g(x) - g(x_0)\| + \\ &\qquad\qquad + \|f(x_0)\| \frac{\|g(x) - g(x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}\end{aligned}$$

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 9.11 folgt $\frac{\|(g\circ f)(x)-(g\circ f)(x_0)-g'(y_0)f'(x_0)(x-x_0)\|}{\|x-x_0\|}\leq$

$$\leq \frac{\|g(f(x))-g(f(x_0))-g'(f(x_0))(f(x)-f(x_0))\|}{\|f(x)-f(x_0)\|}\cdot\frac{\|f(x)-f(x_0)\|}{\|x-x_0\|} + \frac{\|g'(f(x_0))(f(x)-f(x_0)-f'(x_0)(x-x_0))\|}{\|x-x_0\|}$$

$$\leq \frac{\|g(f(x))-g(f(x_0))-g'(f(x_0))(f(x)-f(x_0))\|}{\|f(x)-f(x_0)\|} \cdot\left(\|f'(x_0)\|+\frac{\|f(x)-f(x_0)+f'(x_0)(x-x_0)\|}{\|x-x_0\|}\right) + \|g'(f(x_0))\|\frac{\|f(x)-(f(x_0))-f'(x_0)(x-x_0))\|}{\|x-x_0\|}$$

Daraus folgt (iii). **q.e.d.**

9.3 Schrankensatz

Wir verallgemeinern in diesem Abschnitt den Schrankensatz auf differenzierbare Abbildungen zwischen normierten Vektorräumen.

Lemma 9.20. *Seien f eine stetige Abbildung von einem kompakten Intervall $[a,b]$ in einem normierten Vektorraum W 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

$$\left\{y \mid \text{ für alle } x\in(a,y) \text{ gilt } \|f(x)-f(a)\|\leq\phi(x)-\phi(a)+\epsilon(x-a)+\epsilon\sum_{x_n<x}2^{-n}\right\}.$$

Offenbar ist A_ϵ ein Intervall von der Form $A_\epsilon=[a,\sup A_\epsilon]$. Aus der Stetigkeit von f und ϕ folgt nämlich, dass auch für $y=\sup A_\epsilon$ gilt

$$\begin{aligned}\|f(y)-f(a)\| &\leq \phi(y)-\phi(a)+\epsilon(y-a)+\epsilon\sup_{x\in(a,y)}\sum_{x_n<x}2^{-n}\\ &\leq \phi(y)-\phi(a)+\epsilon(y-a)+\epsilon\sum_{x_n<y}2^{-n}\end{aligned}$$

Wenn $y \in (a,b)$ und f und ϕ in y differenzierbar sind und gilt $\|f'(y)\| \leq \phi'(y)$, dann gibt es ein $\epsilon' > 0$, so dass für alle $x \in (y-\epsilon', y+\epsilon')$ gilt

$$\left|\frac{\phi(x)-\phi(y)}{x-y} - \phi'(y)\right| < \frac{\epsilon}{2} \text{ und } \frac{\|f(x)-f(y)-f'(y)(x-y)\|}{|x-y|} < \frac{\epsilon}{2}.$$

Dann folgt für dieselben x

$$\begin{aligned}\|f(x)-f(a)\| &\leq \|f(x)-f(y)\| + \|f(y)-f(a)\| \\ &< \left(\|f'(y)\| + \frac{\epsilon}{2}\right)|x-y| + \phi(y)-\phi(a)+\epsilon(y-a)+\epsilon\sum_{x_n<y} 2^{-n} \\ &\leq \left(\phi'(y) + \frac{\epsilon}{2}\right)|x-y| + \phi(y)-\phi(a)+\epsilon(y-a)+\epsilon\sum_{x_n<y} 2^{-n} \\ &< \left(\frac{\phi(x)-\phi(y)}{x-y} + \epsilon\right)|x-y| + \phi(y)-\phi(a)+\epsilon(y-a)+\epsilon\sum_{x_n<y} 2^{-n}\end{aligned}$$

Also gilt für $x \in (y, y+\epsilon')$

$$\begin{aligned}\|f(x)-f(a)\| &< \phi(x)-\phi(a)+\epsilon(x-a)+\sum_{x_n<y} 2^{-n} \\ &\leq \phi(x)-\phi(a)+\epsilon(x-a)+\sum_{x_n<x} 2^{-n}\end{aligned}$$

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 $\epsilon' > 0$, so dass aus $x \in (y, y+\epsilon')$ folgt

$$\|f(x)-f(y)\| - \phi(x) - \phi(y) < 2^{-N} \cdot \epsilon.$$

Dann folgt für dieselben x wieder

$$\begin{aligned}\|f(x)-f(a)\| &\leq \|f(x)-f(y)\| + \|f(y)-f(a)\| \\ &< 2^{-N}\cdot\epsilon + \phi(x)-\phi(a)+\epsilon(y-a)+\epsilon\sum_{x_n<y} 2^{-n} \\ &< \phi(x)-\phi(a)+\epsilon(x-a)+\sum_{x_n\leq y} 2^{-n} \\ &\leq \phi(x)-\phi(a)+\epsilon(x-a)+\sum_{x_n<x} 2^{-n}\end{aligned}$$

Also gilt wieder $y+\epsilon' \in A_\epsilon$, was $y = \sup A_\epsilon$ widerspricht. Dann muß aber $\sup A_\epsilon = b$ gelten. Weil das aber für alle $\epsilon > 0$ gilt, folgt auch

$$\|f(b)-f(a)\| \leq \phi(b)-\phi(a).$$

q.e.d.

Korollar 9.21. *(Schrankensatz) Sei f eine Abbildung von einer offenen Teilmenge U des normierten Vektorraumes V in den normierten Vektorraum W. Wenn im Komplement einer abzählbaren Teilmenge S von $D = \{(1-t)a + tb | t \in [0,1]\} \subset U$ die Abbildung f differenzierbar ist, und die Ableitung auf $D \setminus S$ beschränkt ist, dann gilt*

$$\|f(b) - f(a)\| \leq \|b - a\| \sup_{x \in D \setminus S} \|f'(x)\|.$$

Insbesondere gilt für alle $x_0 \in D \setminus S$:

$$\|f(b) - f(a) - f'(x_0)(b - a)\| \leq \|b - a\| \sup_{x \in D \setminus S} \|f'(x) - f'(x_0)\|.$$

Beweis: Sei für ein $t_0 \in (0,1)$ die Abbildung f in $x_t = (1-t_0)a + t_0 b \in U$ differenzierbar. Dann ist die Abbildung $[0,1] \to W, \quad t \mapsto f(x_t)$ 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 $f'(x_{t_0})(b - a)$. Dann sind die Voraussetzungen von Lemma 9.20 mit den beiden Funktionen

$$[0,1] \to W, \quad t \mapsto f(x_t) \text{ und } \qquad [0,1] \to \mathbb{R}, \quad t \mapsto t\|b - a\| \sup_{x \in D \setminus S} \|f'(x)\|$$

erfüllt und die erste Behauptung folgt aus diesem Lemma. Wenn wir dieses Lemma auf die Funktionen

$$\begin{aligned} &[0,1] \to W, && t \to f(x_t) - t f'(x_0)(b - a) \\ &[0,1] \to \mathbb{R}, && t \to t\|b - a\| \sup_{x \in D \setminus S} \|f'(x) - f'(x_0)\| \end{aligned}$$

anwenden, erhalten wir die zweite Behauptung. **q.e.d.**

9.4 Partielle Ableitungen

Definition 9.22. *(stetig differenzierbar) Sei $U \subset V$ eine offene Teilmenge eines normierten Vektorraumes, dann heisst eine Abbildung $f : U \to W$ von U in einem normierten Vektorraum W stetig differenzierbar, wenn*

(i) *f in allen $x_0 \in U$ differenzierbar ist, und*

(ii) *die Abbildung $f' : U \to \mathcal{L}(V, W), \quad x \mapsto f'(x)$ stetig ist.*

Im Falle von $V = W = \mathbb{R}$ stimmt diese Definition überein mit der Definition von stetig differenzierbaren reellen Funktionen auf offenen Teilmengen von $\mathbb{R}$. Weil die Abbildung $\mathbb{R} \to \mathcal{L}(\mathbb{R}, \mathbb{R}), \quad x \mapsto$ Multiplikation mit x, ein Isomorphismus von normierten Vektorräumen ist, können wir die Ableitungen von differenzierbaren reellen Funktionen auf offenen Teilmengen von $\mathbb{R}$ mit reellen Funktionen auf diesen offenen Teilmengen identifizieren, und die stetig differenzierbaren Funktionen sind solche differenzierbare Funktionen, deren Ableitungen mit stetigen reellen Funktionen identifiziert werden.

Definition 9.23. *(der partiellen Ableitung) Seinen V_1 und V_2 normierte Vektorräume und f eine Funktion von einer offenen Teilmenge $U \subset V_1 \times V_2$ des kartesischen Produktes der beiden normierten Vektorräume in den normierten Vektorraum W. Dann heißt f im Punkt $(x_1, x_2) \in U \subset V_1 \times V_2$ partiell differenzierbar, falls die Abbildungen $x \mapsto f(x, x_2)$ im Punkt x_1 differenzierbar ist und $x \mapsto f(x_1, x)$ im Punkt x_2 differenzierbar ist. Allgemeiner heißt eine Abbildung von einer offenen Menge $U \subset V_1 \times \ldots \times V_n$ eines n-fachen kartesischen Produktes von normierten Vektorräumen in einen normierten Vektorraum W im Punkt $(x_1, \ldots, x_n) \in U \subset V_1 \times \ldots \times V_n$ 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 $A \in \mathcal{L}(V_1 \times V_2, W)$ die Abbildungen

$$A_1 : V_1 \to W, \qquad x_1 \mapsto A(x_1, 0) \text{ und } \qquad A_2 : V_2 \to W, \qquad x_2 \mapsto A(0, x_2)$$

stetige lineare Abbildungen sind und weil für alle $x \in V_1$, mit $(x, x_2) \in U$ gilt

$$\frac{\|f(x, x_2) - f(x_1, x_2) - A((x, x_2) - (x_1, x_2))\|}{\|(x, x_2) - (x_1, x_2)\|} = \frac{\|f(x, x_2) f(x_1, x_2) - A(x - x_1, 0)\|}{\|x - x_1\|}$$

bzw. $x \in V_2$, mit $(x_1, x) \in U$ gilt

$$\frac{\|f(x_1, x) - f(x_1, x_2) - A((x_1, x) - (x_1, x_2))\|}{\|(x_1, x) - (x_1, x_2)\|} = \frac{\|f(x_1, x) f(x_1, x_2) - A(0, x - x_2)\|}{\|x - x_1\|}$$

folgt aus den Definitionen der folgende

Satz 9.24. *Wenn eine Funktion $f : U \to W$ von einer offenen Teilmenge $U \subset V_1 \times V_2$ des kartesischen Produktes zweier normierter Vektorräume in einen normierten Vektorraum W im Punkt $(x_1, x_2) \in U$ differenzierbar ist, dann ist f in x_0 auch partiell differenzierbar.* **q.e.d.**

Die Umkehrung dieses Satzes gilt nicht.

Beispiel 9.25.

$$f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \to \begin{cases} \frac{2xy}{x^2+y^2} & \textit{für } (x, y) \neq 0 \\ 0 & \textit{für } (x, y) = 0 \end{cases}$$

existieren die partiellen Ableitungen

$$\frac{\partial f}{\partial y} = \begin{cases} \frac{2y}{x^2+y^2} - \frac{4x^2y}{(x^2+y^2)^2} & \textit{für } (x, y) \neq 0 \\ 0 & \textit{für } (x, y) = 0 \end{cases}$$

$$\frac{\partial f}{\partial y} = \begin{cases} \frac{2x}{x^2+y^2} - \frac{4y^2x}{(x^2+y^2)^2} & \textit{für } (x, y) \neq 0 \\ 0 & \textit{für } (x, y) = 0 \end{cases}$$

Aber f ist im Punkt $(x, y) = 0$ nicht stetig, weil für alle (r, ϕ) gilt

$$f(r \cos \phi, r \sin \phi) = 2 \sin \phi \cos \phi = \sin(2\phi),$$

so dass für alle ϕ gilt

$$\lim_{r \to 0+} f(r \cos \phi, r \sin \phi) = \sin(2\phi).$$

Aber es gilt folgende Umkehrung.

Satz 9.26. *Sei $f : U \to W$ eine partiell differenzierbare Funktion von einer offenen Teilmenge U des kartesischen Produktes zwei normierter Vektorräume $V_1 \times V_2$ in dem normierten Vektorraum W. Dann sind die partiellen Ableitungen von f*

$$\frac{\partial f}{\partial x_1} : U \to \mathcal{L}(V_1, W) \textit{ und } \qquad \frac{\partial f}{\partial x_2} : U \to \mathcal{L}(V_2, W)$$

genau dann auf U stetig, wenn f auf U stetig differenzierbar ist.

Beweis: Wenn $A_1 \in \mathcal{L}(V_1, W)$ und $A_2 \in \mathcal{L}(V_2, W)$, dann sind auch die Abbildungen

$$V_1 \times V_2 \to W, \quad (x_1, x_2) \to A_1 x_1 \text{ bzw.} \quad V_1 \times V_2 \to W, \quad (x_1, x_2) \to A_2 x_2$$

linear stetige Abbildungen in $\mathcal{L}(V_1 \times V_2, W)$. Also ist

$$A_1 \times A_2 : V_1 \times V_2 \to W, \quad (x_1, x_2) \mapsto A_1 x_1 + A_2 x_2$$

eine stetige lineare Abbildung in $\mathcal{L}(V_1 \times V_2, W)$. Umgekehrt sind für jede stetige lineare Abbildung $A \in \mathcal{L}(V_1 \times V_2, W)$ die Abbildungen $A_1 : V_1 \to W$, $x_1 \mapsto A(x_1, 0)$ und $A_2 : V_2 \to W$, $x_2 \mapsto A(0, x_2)$ stetige lineare Abbildungen in $\mathcal{L}(V_1, W)$ und $\mathcal{L}(V_2, W)$. Wegen

$$\|A(x_1, x_2)\| = \|A(x_1, 0) + A(0, x_2)\| \leq \|A(x_1, 0)\| + \|A(0, x_2)\|$$

folgt
$$\|A\| \le \|A_1\| + \|A_2\|.$$
Aufgrund der Definition der Norm von $\mathcal{L}(V_1 \times V_2, W)$ folgt aber auch
$$\|A_1\| \le \|A\| \text{ und } \|A_2\| \le \|A\|.$$
Also ist die Abbildung
$$\mathcal{L}(V_1, W) \times \mathcal{L}(V_2, W) \to \mathcal{L}(V_1 \times V_2, W), (A_1, A_2) \mapsto A$$
eine bijektive Abbildung von normierten Vektorräumen und die beiden Normen von $\mathcal{L}(V_1, W) \times \mathcal{L}(V_2, W)$ und $\mathcal{L}(V_1 \times V_2, W)$ sind bezüglich dieser Identifikation äquivalent. Daraus folgt, dass für jede stetig differenzierbare Funktion $f : U \to W$, die beiden partiellen Ableitungen $\frac{\partial f}{\partial x_1} : U \to \mathcal{L}(V_1, W)$ und $\frac{\partial f}{\partial x_2} : U \to \mathcal{L}(V_2, W)$ stetig sind.

Wenn umgekehrt $\frac{\partial f}{\partial x_1}$ stetig ist, dann folgt
$$\begin{aligned}\left\| f(y_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\| &\le \\ \le \left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) \right\| &+ \\ + \left\| f(x_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\| &.\end{aligned}$$
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 9.21 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 9.27. (i)

$$f : \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x,y) \neq 0 \\ 0 & \text{für } (x,y) = 0. \end{cases}$$

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} \frac{2y(y^2-x^2)}{(x^2+y^2)^2} & \text{für } (x,y) \neq 0 \\ 0 & \text{für } (x,y) = 0 \end{cases}$$

$$\frac{\partial f}{\partial y}(x,y) = \begin{cases} \frac{2x(x^2-y^2)}{(x^2+y^2)^2} & \text{für } (x,y) \neq 0 \\ 0 & \text{für } (x,y) = 0 \end{cases}$$

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 9.25 (i) gesehen, dass f bei $(0,0)$ nicht stetig und deshalb auch nicht differentzierbar ist.

(ii)

$$f : \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto \begin{cases} \frac{x^3-y^3}{x^2+y^2} & \text{für } (x,y) \neq 0 \\ 0 & \text{für } (x,y) = 0. \end{cases}$$

Offenbar gilt $|f(x,y)| < \frac{|y|^3}{x^2+y^2} + \frac{|y|^3}{x^2+y^2} \leq |x| + |y|$. Also ist f stetig.

$$\frac{\partial f}{\partial x} = \begin{cases} \frac{3x^2(x^2+y^2)-2x(x^3-y^3)}{(x^2+y^2)^2} = \frac{x(x^3+3xy^2+2y^3)}{(x^2+y^2)^2} & \text{für } (x,y) \neq 0 \\ 1 & \text{für } (x,y) = 0 \end{cases}$$

$$\frac{\partial f}{\partial y} = \begin{cases} -\frac{3y^2(x^2+y^2)+2y(x^3-y^3)}{(x^2+y^2)^2} = -\frac{y(y^3+3yx^2+2x^3)}{(x^2+y^2)^2} & \text{für } (x,y) \neq 0 \\ -1 & \text{für } (x,y) = 0. \end{cases}$$

Deshalb ist f partiell differenzierbar. In Beispiel 9.29 (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 9.28. *(Richtungsableitung) Sei $f : U \to W$ eine Funktion von einer offenen Teilmenge U eines normierten Vektorraumes V in einem normierten Vektorraum W. Für $x_0 \in U$ und $v \in V \setminus \{0\}$ heißt die Ableitung in $t_0 = 0$ der Funktion*

$$(-\epsilon, \epsilon) \to W, t \mapsto f(x_0 + tv)$$

wenn sie existiert die Richtungsableitung von f in x_0 in Richtung v.

Beispiel 9.29. **(i)** *Sei $f : U \to W$ in x_0 differenzierbar. Dann ist die Abbildung*

$$(-\epsilon, \epsilon) \to \mathbb{R}, t \mapsto \begin{cases} \frac{\|f(x_0+tv)-f(x_0)-tf'(x_0)v\|}{|t|\|v\|} & \text{für } t \neq 0 \\ 0 & \text{für } t = 0 \end{cases}$$

im Punkt $t = 0$ stetig. Dann ist aber auch die Abbildung

$$(-\epsilon, \epsilon) \to \mathbb{R}, t \mapsto \begin{cases} \frac{\|f(x_0+tv)-f(x_0)-tf'(x_0)v\|}{|t|} & \text{für } t \neq 0 \\ 0 & \text{für } t = 0 \end{cases}$$

stetig. Also existiert die Richtungsableitung und es gilt

$$\frac{df(x_0 + tv)}{dt}\Big|_{t=0} = f'(x_0) \cdot v$$

(ii)

$$f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

Dann gilt für $v = (\cos\phi, \sin\phi)$ und $x_0 = 0$

$$f(tv) = \begin{cases} \sin(2\phi) & \text{für } t \neq 0 \\ 0 & \text{für } t = 0. \end{cases}$$

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 aber die Richtungsableitungen und verschwinden.

(iii)

$$f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto \begin{cases} \frac{x^3-y^3}{x^2+y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

Dann gilt für $v = (\cos\phi, \sin\phi)$ *und* $x_0 = 0$

$$f(tv) = t(\cos^3\phi - \sin^3\phi).$$

Also existieren alle Richtungsableitungen. Die Richtungsableitungen setzten sich also im Punkt $(0,0)$ *zu der Abbildung* f *zusammen. Weil diese Abbildung aber nicht linear ist, ist* f *im Punkt* $(0,0)$ *nicht differenzierbar.*

Wir wollen den wichtigesten Fall von Funktionen $\mathbb{R}^n \to \mathbb{R}^m$ nochmal getrennt betrachten.

Definition 9.30. *(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.*

Satz 9.24 und Satz 9.26 zeigen, dass im für uns besonders wichtigen Fall von Funktionen von Teilmengen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ folgendes gilt:

Korollar 9.31. *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 9.32. *(Gradient) Der Vektor* $(\frac{\partial f}{\partial x_i}, \ldots, \frac{\partial f}{\partial x_n})$ *der partiellen Ableitungen einer partiell differenzierbaren reellen Funktion* f *auf einer offenen Menge* $U \subset \mathbb{R}^n$ *heisst Gradient von* f*. Er wird mit* ∇f *bezeichnet.*

Wenn f in $x_0 \in U$ differenzierbar ist, dann ist f auch in x_0 partiell differenzierbar und die Ableitung ist dann gegeben durch 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 mit der Jacobimatrix identifizieren:

$$\frac{df}{dx}(x_0) = \begin{pmatrix} \frac{\partial f_1(x_0)}{\partial x_1} & \ldots & \frac{\partial f_1(x_0)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m(x_0)}{\partial x_1} & \ldots & \frac{\partial f_m(x_0)}{\partial x_n} \end{pmatrix}$$

und die entsprechende lineare Abbildung

$$\frac{df}{dx}(x_0) : \mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto \frac{df}{dx}(x_0) \cdot x$$

ist dann 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 $a \in \mathbb{R}^n$ das Skalarprodukt des Gradienten $\nabla f(x_0)$ von f an der Stelle x_0 mit dem Vektor a:

$$\frac{d}{dt} f(x_0 + ta)|_{t=0} = a \cdot \nabla f(x_0).$$

Definition 9.33. *Ein Punkt $x_0 \in U$ einer reellen differenzierbaren Funktion f auf einer offenen Menge U, an dem die Ableitung von f verschwindet, heisst kritischer Punkt.*

Satz 9.34. *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 $v \in V$ die entsprechende Abbildung $t \mapsto f(x_0 + tv)$ auf einer Umgebung von $0 \in \mathbb{R}$ differenzierbar und besitzt dort ein lokales Maximum (bzw. Minimum). Also verschwindet die entsprechende Richtungsableitung. Dann verschwindet also auch die Ableitung $\frac{df}{dx}(x_0)$ auf allen v. **q.e.d.**

9.5 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 9.35. *Eine Abbildung $A : V \times V \to W$ heisst bilinear, wenn für alle $v, v', v'' \in V$ und $\lambda \in \mathbb{K}$ gilt*

$$\begin{aligned} A(v+v'',v') &= A(v,v') + A(v'',v) \text{ und} & A(v,v',v'') &= A(v,v') + A(v,v'') \\ A(\lambda v,v') &= \lambda A(v,v') \text{ und} & A(v,\lambda v') &= \lambda A(v,v'). \end{aligned}$$

Das kartesische Produkt $V \times V$ von (normierten) Vektorräumen ist wieder ein normierter Vektorraum. Aber 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 hier nur kurz erwähnen wollen, und den man als das Tensorprodukt von V mit V bezeichnet. Die bilinearen Abbdildungen 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

$$\begin{aligned} \dim(V \times V) &= \dim(V) & + \dim(V) \\ \dim(V \otimes V) &= \dim(V) & \cdot \dim(V). \end{aligned}$$

Die linearen Abbildungen von V nach $\mathcal{L}(V,W)$ lassen sich nun mit den bilinearen Abbdilungen von $V \times V$ nach W identifizieren.

Lemma 9.36. *Eine Abbildung $A : V \times V \to W$ ist genau dann bilinear, wenn die Abbildung*

$$B : V \to \{ \text{ lin. Abbild. } V \to W \}, \quad v \mapsto B(v), \quad B(v) : V \to W, \quad B(v)(v') = A(v,v')$$

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 9.37. *(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. für alle $u, v \in V$ gilt*

$$(f''(x_0)u)v = (f''(x_0)v)u.$$

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 alle $u, v \in B(0, \delta) \subset V$ die Ungleichungen

$$\begin{aligned} &\|f'(x_0 + tu + v) - f'(x_0) - f''(x_0) \cdot (tu + v)\| < \epsilon(\|u\| + \|v\|) \\ &\|f'(x_0 + tu) - f'(x_0) - f''(x_0)tu\| < \epsilon\|u\| \end{aligned}$$

gelten. Daraus folgt

$$\|g'(t) - (f''(x_0) \cdot v) \cdot u\| < \epsilon\|u\|(2\|u\| + \|v\|).$$

Die Anwendung von Korollar 9.21 auf die Funktion $t \mapsto g(t) - t(f''(x_0)v)u$ ergibt dann

$$\|g(1) - g(0) - (f''(x_0)v)u\| \le \sup_{t \in [0,1]} \|g'(t) - (f''(x_0)v)u\| < \epsilon\|u\|(2\|u\| + \|v\|).$$

Weil aber $g(1) - g(0) = f(x_0 + u + v) - f(x_0 + u) - f(x_0 + v) + f(x_0)$ in u und v symmetrisch ist gilt dann auch $\|g(1) - g(0) - (f''(x_0)u)v\| < \epsilon\|v\|(2\|v\| + \|u\|)$. Daraus folgt

$$\|(f''(x_0)v)u - (f''(x_0)u)v\| < 2\epsilon(\|u\|^2 + \|u\| \cdot \|v\| + \|v\|^2).$$

Wegen der Linearität gilt diese Ungleichung aber nicht nur für $u, v \in B(0, \delta)$, sondern für alle $u, v \in V$. Weil aber für jedes ϵ eine solche Ungleichung gilt mit einem geeigneten $\delta > 0$, folgt für alle $u, v \in V$ auch $(f''(x_0)v)u = (f''(x_0)u)v$. **q.e.d.**

Zusammen mit Satz 9.26 erhalten wir

Korollar 9.38. *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*

$$\partial_i \partial_j f_k(x) = \frac{\partial^2 f_k(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f_k(x)}{\partial x_j \partial x_i} = \partial_j \partial_i f_k(x).$$ **q.e.d.**

Durch mehrfaches Anwenden und differenzieren erhalten wir dann auch

Korollar 9.39. *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 9.40. *Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto f(x, y)$ mit $f(x, y) = xy\dfrac{x^2 - y^2}{x^2 + y^2}$ für $(x, y) \neq (0, 0)$ und $f(0, 0) = 0$. Dann ist f zweimal partiell differenzierbar.*

$$\begin{aligned}
\frac{\partial f}{\partial x}(x,y) = \quad & y\frac{(x^2-y^2)(x^2+y^2)+2x^2(x^2+y^2)-2x^2(x^2-y^2)}{(x^2+y^2)^2} = y\frac{x^4+4x^2y^2-y^4}{(x^2+y^2)^2} \\
\frac{\partial f}{\partial y}(x,y) = \quad & x\frac{(x^2-y^2)(x^2+y^2)-2y^2(x^2+y^2)-2y^2(x^2-y^2)}{(x^2+y^2)^2} = x\frac{x^4-4x^2y^2-y^4}{(x^2+y^2)^2}
\end{aligned}$$

$$\frac{\partial^2 f}{\partial y \partial x}(0, y) = -1 \qquad\qquad \frac{\partial^2 f}{\partial x \partial y}(x, 0) = 1$$

Also existieren auf ganz $\mathbb{R}^2$ alle zweiten partiellen Abeleitungen, aber es gilt

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1 \neq 1 = \frac{\partial^2 f}{\partial x \partial y}(x, 0).$$

Zum Abschluss dieses Kapitels 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 die Funktion g n-mal differenzierbar. Wegen der Kettenregel Satz 9.19 (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 9.41. *(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.**

Eine unendlich oft differenzierbare Funktion heisst dann wieder reell analytisch in x_0, wenn die entsprechende Taylorreihe auf einer Umgebung von x_0 gegen die Funktion konvergiert. Das Taylorpolynom und die Taylorreihe lassen sich offenbar auch auf Funktionen mit Werten in einem Banachraum verallgemeinern. So definiert die Exponentialfunktion eine analytische Funktion von der Banachalgebra aller stetigen linearen Abbildungen $\mathcal{L}(V)$ eines Banachraumes auf sich selber.

9.6 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 werden. Wenn A invertierbar ist, erhalten wir die eindeutige Lösung $x = A^{-1}y$. In diesem Abschnitt wollen wir versuchen das Verständnis dieser Gleichungen auch 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 nutzbar zu machen. Dabei werden wir annehmen, dass f differenzierbar ist und deshalb durch lineare Abbildungen angenähert werden kann. Ausgangspunkt ist daher die Beobachtung, dass kleine Störungen von invertierbaren linearen Abbildungen wieder invertierbar sind.

Lemma 9.42. *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 in*

$$B \in B\left(A, \frac{1}{\|A^{-1}\|}\right) \subset \mathcal{L}(V,W)$$

invertierbar erfüllen

$$\|B^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2\|A-B\|}{1 - \|A^{-1}\|\|A-B\|}$$

Beweis: Offenbar ist $B = A - (A-B) = A(\mathbb{1}_V - A^{-1}(A-B))$. Für $B \in B\left(A, \frac{1}{\|A^{-1}\|}\right)$ gilt wegen Satz 9.11 $\|A^{-1}(A-B)\| \leq \|A^{-1}\|\|A-B\| < 1$. 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

$$B^{-1} = \left(\mathbb{1}_V - A^{-1}(A-B)\right)^{-1} A^{-1} \text{ mit } \|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|A-B\|}.$$

Für die Differenz $B^{-1} - A^{-1}$ gilt dann

$$\begin{aligned} B^{-1} - A^{-1} &= ((\mathbb{1}_V - A^{-1}(A-B))^{-1} - \mathbb{1})A^{-1} \\ &= A^{-1}(A-B)(\mathbb{1} - A^{-1}(A-B))^{-1}A^{-1} \\ &= (\mathbb{1}_V - A^{-1}(A-B)^{-1}A^{-1}(A-B)A^{-1} \end{aligned}$$

mit

$$\|B^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2\|A-B\|}{1 - \|A^{-1}\|\|A-B\|}.$$

q.e.d.

Damit bilden die invertierbaren Elemente von $\mathcal{L}(V,W)$ eine offene Teilmenge.

Korollar 9.43. *Seien V und W Banachräume und A ein invertierbares Element von $\mathcal{L}(V,W)$. Dann ist die Abbildung*

$$B\left(A, \frac{1}{\|A^{-1}\|}\right) \to \mathcal{L}(W,V), A \mapsto A^{-1}$$

eine analytische Abbildung. Also insbesondere unendlich oft stetig differenzierbar.

Beweis: Aus Lemma 9.42 und der Neumannschen Reihe folgt für alle $B \in B\left(0, \frac{1}{\|A^{-1}\|}\right)$

$$(A+B)^{-1} = A^{-1}\sum_{n=0}^{\infty}(-BA^{-1})^n = \left(\sum_{n=0}^{\infty}(-A^{-1}B)^n \cdot\right) A^{-1}.$$

Insbesondere ist die Ableitung an der Stelle A die lineare Abbildung $B \mapsto -A^{-1}BA^{-1}$, und damit genauso oft differenzierbar, wie $A \mapsto A^{-1}$. **q.e.d.**

Satz 9.44. *(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 $f'(x_0)$ ind $x_0 \in U$ 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 $x \to (f'(x_0))^{-1} \circ (f(x - x_0) - f(x_0))$ 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 für $x \in \overline{B(0,\delta)} \subset U$ gilt $\|f'(x) - \mathbb{1}_V\| \leq \frac{1}{2}$. Wegen dem Schrankensatz gilt dann für alle $x \in \overline{B(0,\delta)}$

$$\|f(x) - x\| = \|f(x) - x - (f(0) - 0)\| \leq \frac{\|x\|}{2}.$$

Also 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 $\frac{1}{2}$. 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 auch eine Abbildung von $\overline{B(y, \frac{\delta}{2})}$ auf sich selbst. Aus dem Banachschen Fixpunktsatz folgt dann, dass für jedes $y \in \overline{B(y, \frac{\delta}{2})}$ auf $\overline{B(y, \frac{\delta}{2})}$ die Abbildung F_y genau einen Fixpunkt hat. Wegen der Kontraktionseigenschaft gilt sogar, dass F_y in $\overline{B(0,\delta)}$ nur genau einen Fixpunkt hat. 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

$$O = B\left(0, \frac{\delta}{2}\right) \text{ und } U' = \left\{x \in B(0,\delta) \,|\, f(x) \in B\left(0, \frac{\delta}{2}\right)\right\}.$$

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 Lipschitz-stetig 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 9.42 die Abbildung

$$\overline{B(0,\delta)} \to \mathcal{L}(V), x \mapsto (f'(x))^{-1}$$

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 $f : O \to U'$ mit $x \mapsto (f'(x))^{-1}$ die Ableitung von f^{-1}. Dann ist also f^{-1} auch stetig differenzierbar. **q.e.d.**

Beispiel 9.45. *Die Voraussetzung der stetigen Differenzierbarkeit kann nicht abgeschwächt werden zu einfacher Differenzierbarkeit. Die Funktion*

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) = \begin{cases} x + x^2 \sin(\frac{1}{x}) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

ist differenzierbar und bei 0 *gilt* $f'(0) = 0$. *Aber* f *ist in keiner Umgebung der* 0 *injektiv.*

Korollar 9.46. *Die Umkehrabbildung einer bijektiven* n*-mal stetig differenzierbaren Abbildung ist in allen Punkten, in denen* $f'(x)$ *invertierbar ist,* n*-mal stetig differenzierbar.* **q.e.d.**

Korollar 9.47. *(Satz über die implizite Funktion) Seien* V *und* W *Banachräume und* U *und* O *offene Teilmengen von* V *bzw.* W *und* $f : U \times O \to V$ *eine stetig differenzierbare Funktion. Wenn die partielle Ableitung* $\frac{\partial f}{\partial x_1}$ *in* $(v_0, w_0) \in U \times O$ *als Element von* $\mathcal{L}(V)$ *invertierbar ist, dann gibt es offene Umgebungen* $U' \subset W$ *und* $O' \subset O$ *von* $f(v_0, w_0)$ *bzw.* w_0 *und eine stetig differenzierbare Funktion* $g : U' \times O' \to V$, *so dass für alle* $(v, w) \in U' \times O$ *gilt* $f(g(v, w), w) = v$. *Außerdem sind für alle* $v \in U'$ *alle Lösungen der Gleichungen* $f(u, w) = v$ *in einer Umgebung von* (v_0, w_0) *im Graphen der Abbildung* $O' \to V, w \to g(v, w)$ *enthalten.*

Beweis: Die Ableitung der Abbildung $F : U \times O \to V \times W, \quad (v, w) \mapsto (f(v, w), w)$ 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. Die erste Komponente der Umkehrfunktion F^{-1} ist dann die gesuchte Funktion g. **q.e.d.**

Beispiel 9.48. (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 $\frac{\partial f}{\partial x_1}$ (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 die Höhenlinien von f gerade durch die Graphen der Funktionen

$$y \to g(z, y)$$

beschrieben werden mit

$$z \in (f(x_0, y_0) - \epsilon, f(x_0, y_0) + \epsilon)(\text{ der Höhe } z = f(g(z, y), y)).$$

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. Wir nehmen an*

$$\frac{\partial f}{\partial x_1}(x_0) \neq 0.$$

Dann lassen sich lokal die Niveaumengen:

$$\{(x \in \mathbb{R}|f(x) = z\}$$

mit $z \in (f(x_0) - \epsilon, f(x_0 + \epsilon))$ *durch den Graphen einer Funktion* $g : \mathbb{R} \times \mathbb{R}^{n-1} \to \mathbb{R}$ *beschreiben als* $(g(z, y), y) \in \mathbb{R} \times \mathbb{R}^{n-1}$ *mit* (z, y) *in einer Umgebung von* $(f(x_0), y_0)$*, wobei* y_0 *alle bis auf die erste Komponente von* x_0 *enthält. 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*

$$\frac{\partial g}{\partial y}(z, y) \times \mathbb{1}_Y \in \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}) \times \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}^{n-1}) \simeq \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R} \times \mathbb{R}^{n-1})$$

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 9.49. *Eine unendlich oft (stetig) differenzierbare bijektive Abbildung mit unendlich oft (stetig) differenzierbarer Umkehrabbildung heisst Diffeomorphismus.*

Beispiel 9.50. *Polarkoordinaten Die Abbildung*

$$\mathbb{R}^+ \times \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{R}^2 \setminus \{0\}(r, \phi) \mapsto (r \cos\phi, \sin\phi)$$

heisst Polarkoordinaten von $\mathbb{R}^2$*. Offenbar ist diese Abbildung unendlich oft stetig differenzierbar. Die Umkehrabbildung ist dann gegeben durch*

$$(x, y) \mapsto (r, \phi) \text{ mit } r = \sqrt{x^2 + y^2}$$

und

$$\phi = \begin{cases} \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) & \text{für } y \geq 0 \\ 2\pi - \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) & \text{für } y < 0 \end{cases}$$

Also ist diese Abbildung tatsächlich ein Diffeormorphismus von

$$\mathbb{R}^+ \times \mathbb{R}/2\pi\mathbb{Z} \text{ nach } \mathbb{R}^2 \setminus \{0\}.$$

Hier beschreibt $\mathbb{R}/2\pi\mathbb{Z}$ *den Raum aller Äquivalenzklassen von* $\mathbb{R}$*, wobei*

$$x \sim y \Leftrightarrow \frac{x - y}{2\pi} \in \mathbb{Z}.$$

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 auf beliebige Intervalle mit kleinerer Länge als* 2π *der Abbildung* $\mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z}$*, die jedes Element auf die entsprechende Äquivalenzklasse abbildet, ein Diffeomeorphismen.*

Kapitel 10

Das Lebesgueintegral auf dem $\mathbb{R}^d$

10.1 Stufenfunktionen

Zunächst führen wir die Klasse der Mengen von Quader im $\mathbb{R}^d$ ein.

Definition 10.1. *Ein Quader ist ein n-faches kartesisches Produkt von Intervallen im* $\mathbb{R}^d$

$$Q = I_1 \times \ldots \times I_d = \{x \in \mathbb{R}^d | x_1 \in I_1, \ldots, x_d \in I_d\}$$

wobei $I_1, \ldots, I_d$ *Intervalle in* $\mathbb{R}$ *sind. Diese können den linken und rechten Rand enthalten, bzw. nicht enthalten.*

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 10.2. *Eine Teilmenge A des $\mathbb{R}^d$ heisst Nullmenge, wenn es für jedes $\epsilon > 0$ eine Folge von Quadern mit endlichem Volumen $(Q_n)_{n\in\mathbb{N}}$ im $\mathbb{R}^d$ gibt, mit*

$$A \subset \bigcup_{n+1}^{\infty} Q_n \text{ und } \qquad \sum_{n=1}^{\infty} \mu(Q_n) \leq \epsilon.$$

Lemma 10.3. *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 für alle $n \in \mathbb{N}$ Q_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 10.4. *Eine abzählbare Vereinigung von Nullmengen ist wieder eine Nullmenge.*

Beweis: Sei $\bigcup\limits_{n=1}^{\infty} A_n$ eine abzählbare Vereinigung von Nullmengen. Dann besitzt für jedes $\epsilon > 0$ jede Menge A_n eine Überdeckung von Quadern, deren gesamtes Volumen nicht größer ist als $\epsilon \cdot 2^{-n}$. Die Vereinigung aller dieser Quader hat dann ein gesamtes Volumen von $\sum\limits_{n=1}^{\infty} \epsilon 2^{-n} = \epsilon$. Also gibt es eine Überdeckung von $\bigcup\limits_{n=1}^{\infty} A_n$ von Quadern, deren gesamtes Volumen nicht größer ist als ϵ. **q.e.d.**

Definition 10.5. *Eine Stufenfunktion ist eine endliche Linearkombination von charakteristischen Funktionen von Quadern.*

Proposition 10.6. *Jede Stufenfunktion ist eine endliche Linearkombination von charakteristischen Funktionen von paarweise disjunkten Quadern.*

Beweis: Zunächst zeigen wir, dass je zwei Quader Q_1 und Q_2 im $\mathbb{R}^d$ eine disjunkte Vereinigung von höchstens 3^d-Quadern ist. Das folgt daraus, dass zwei Intervalle $\mathbb{R}$ entweder disjunkt sind, oder eine disjunkte Vereinigung von der Schnittmenge mit den relativen Komplementen der Schnittmenge in beiden 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. **q.e.d.**

Als nächstes definieren wir für jede charakteristische Funktion χ_Q eines Quaders das Integral

$$\int \chi_Q d\mu = \mu(Q).$$

Proposition 10.7. *Sei f eine Stufenfunktion und*

$$f = \sum_i c_i \chi_{Q_i} = \sum_j d_j \chi_{R_j}$$

zwei Zerlegungen in endliche Linearkombinationen von charakteristischen Funktionen von paarweise disjunkten Quadern. Dann ist

$$\int f d\mu = \sum_i c_i \mu(Q_i) = \sum_j d_j \mu(R_j).$$

Beweis: Im Fall $d = 1$ ist das Komplement eines beschränkten Intervalles die disjunkte Vereinigung von zwei Intervallen. Für jedes beschränkte Intervall wird damit $\mathbb{R}$ zu einer disjunktem Vereinigung von drei Intervallen. Wählen wir von den endlich vielen beschränkten Intervallen Q_i und R_j jeweils eine der drei entsprechenden Intervalle aus, so bilden die entsprechenden Schnittmengen eine disjunkte Vereinigung von $\mathbb{R}$. Die Teilmenge aller der Schnittmengen, die in der Vereinigung $(\cup_i Q_i) \cup (\cup_j R_j)$ enthalten sind, ergibt eine disjunkte Vereinigung dieser Menge. Wenden wir für $d > 1$ diese Zerlegung auf alle Faktoren des kartesischen Produktes $\mathbb{R}^{\times d}$ an, dann bilden die kartesichen Produkten $I_1 \times \ldots \times I_d$ von allen Kombinationen von Intervallen $I_1, \ldots, I_d$ aus dem kartesischen Produkt dieser Zerlegungen wieder eine disjunkte Vereinigung von $(\cup_i Q_i) \cup (\cup_j R_j)$. Weil die Stufenfunktion f einen eindeutigen Wert auf jedem Quader annimmt, erhalten wir eine eindeutige gemeinsame Zerlegung in eine endliche Linearkombination von charakteristischen Funktionen auf paarweise disjunkten Quadern. Dann genügt es wegen der Linearität die Behauptung für eine solche Zerlegung eines Quaders $I_1 \times \ldots \times I_d$ in ein kartesisches Produkt von Zerlegungen der Intervalle $I_1, \ldots, I_d$ in paareweise disjunkten Intervalle zu zeigen. Wegen dem Distributivgesetz folgt das aus dem Spezialfall mit $d = 1$. Dort folgt es 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 Stufenfunktionen nach $\mathbb{R}$.

Proposition 10.8. *Seien f und g zwei Stufenfunktionen mit $f \geq g$. Dann gilt*

$$\int f d\mu \geq \int g d\mu.$$

Beweis: Wir zerlegen die beiden Vereinigungen von Quadern der Stufenfunktion f und der Stufenfunktion g in eine gemeinsame disjunkte Vereinigung von Quadern. Auf jedem der Quader ist dann aber f größer oder gleich g. Deshalb gilt das dann auch für die entsprechenden Summen, die die Integrale berechnen. **q.e.d.**

10.2 Lebesgue–integrable Funktionen auf dem $\mathbb{R}^d$

Satz 10.9. *Sei $(f_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge von Stufenfunktionen, deren Integrale $\left(\int f_n d\mu\right)_{n\in\mathbb{N}}$ beschränkt sind. Dann ist die Menge aller Punkte $\{x \in \mathbb{R} | (f_n(x))_{n\in\mathbb{N}}$ konvergiert nicht $\}$ eine Nullmenge.*

Beweis: Sei $M > 0$ eine obere Schranke von $\left(\int (f_n - f_1) d\mu\right)_{n\in\mathbb{N}}$:

$$\int f_n d\mu \leq M + \int f_1 d\mu \qquad \text{für alle } n \in \mathbb{N}.$$

Dann ist für alle $\epsilon > 0$ und alle $n \in \mathbb{N}$, die Menge

$$S_{n,\epsilon} = \left\{ x \in \mathbb{R}^d | f_n(x) \geq \frac{M}{\epsilon} + f_1(x) \right\}$$

eine monoton wachsende Folge von endlichen Vereinigungen von Quadraten. Aus der Konstruktion einer gemeinsamen Zerlegung in eine disjunkte Vereinigung von Quadern im Beweis von Proposition 10.6 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 Quadraten. Weil aber $f_n - f_1$ positive Funktionen sind, ist das Gesamtvolumen von $S_{n,\epsilon}$ nicht größer als

$$\int \chi_{S_{n,\epsilon}} d\mu \leq \int \frac{\epsilon}{M}(f_n - f_1) d\mu = \frac{\epsilon}{M} \int (f_n - f_1) d\mu \leq \epsilon.$$

Wegen der Monotonie ist dann auch das Gesamtvolumen der abzählbaren Vereinigung $\bigcup\limits_{n=1}^{\infty} S_{n,\epsilon}$ nicht größer als ϵ. Weil die kritische Menge 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.**

Wir bezeichnen nun die Komplemente von Nullmengen als fast überall. Also besagt der vorangehende Satz, dass jede monotonwachsende Folge von Stufenfunktionen fast überall konvergiert.

Satz 10.10. *Für jede Nullmenge $A \subset \mathbb{R}^d$ gibt es eine monoton wachsende Folge $(f_n)_{n \in \mathbb{N}}$ von Stufenfunktionen 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 monotonwachsende Folge von Stufenfunktionen, 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 monotonwachsende Folge $(f_n)_{n\in\mathbb{N}}$ von Stufenfunktionen 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 $\lim\limits_{n\to\infty} \int f_n d\mu$ definieren. Damit diese Definition aber konsitent nur von der fast überall definierten Funktion f abhängt, benötigen wir noch die folgenden Lemmata.

Lemma 10.11. *Sei $(f_n)_{n\in\mathbb{N}}$ eine monoton fallende Folge von nicht negativen Stufenfunktionen, die fast überall gegen Null konvergieren. Dann ist $\left(\int f_n d\mu\right)_{n\in\mathbb{N}}$ eine Nullfolge.*

Beweis: 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. Wegen $0 \le f_n \le f_1$ ist A_n eine Nullmenge in Q_0. Dann ist auch $A = \bigcup\limits_{n=1}^{\infty} A_n$ 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 dann eine Überdeckung $\bigcup\limits_{m=1}^{\infty} Q_m \supset (A \cup B)$ durch Quader, deren Gesamtvolumen nicht größer ist als ϵ. Indem wir die Kanten aller Quader um den Faktor $(1 + \epsilon')$ vergrößern, dabei aber den Mittelpunkt fixieren, und alle Quader vom Volumen Null weglassen, erhalten wir auch eine solche Überdeckung $\bigcup\limits_{m=1}^{\infty} Q_m \supset (A \cup B)$ durch offene Quader, deren Gesamtvolumen nicht größer ist als ϵ. Für jeden Punkt $x \in Q_0 \setminus (A \cup B)$ gibt es auch ein N, so dass $f_N(x) \le \epsilon$. Weil $(f_n)_{n\in\mathbb{N}}$ monoton fallend ist, gilt für alle $n \ge N$ auch $f_n(x) \le \epsilon$. Weil alle f_N bei den Punkten von $Q_0 \setminus (A \cup B)$ lokal konstant sind, gibt es eine offene Überdeckung von offenen Quadern $(R_m)_{m\in\mathbb{N}}$ von $Q_0 \setminus (A \cup B)$, und eine Folge $(N_m)_{m\in\mathbb{N}}$, so dass auf R_m für $n \ge N_m$ gilt $f_n \le \epsilon$. Dann bilden $(R_m)_{m\in\mathbb{N}}$ 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_m's können wir $\int f_n d\mu$ abschätzen durch

$$\int f_n d\mu \le \epsilon(\max\{f_1(x)|x \in Q_0\} + \mu(Q_0)).$$

Auf den Quadern $(Q_m)_{m\in\mathbb{N}} f$ schätzen wir f ab durch $\max\{f_1(x)|x \in Q_0\}$ und auf den offenen Quadern $(R_m)_{m\in\mathbb{N}}$ durch ϵ. Also konvergiert $\left(\int f_n d\mu\right)_{n\in\mathbb{N}}$ gegen Null. **q.e.d.**

Lemma 10.12. *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 Stufenfunktionen 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 gilt $f \geq g$, dann gilt auch*

$$\lim_{n\to\infty} \int f_n d\mu \geq \lim_{n\to\infty} \int g_n d\mu.$$

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. Weil aber

$$g_m - f_n \leq (g_m - f_n)^+$$

gilt, folgt aus Proposition 10.8 und Lemma 10.11

$$\int g_m d\mu - \lim_{n\to\infty} \int f_n d\mu \leq= 0.$$

Dann gilt aber auch

$$\lim_{m\to\infty} \int g_m d\mu \leq \lim_{n\to\infty} \int f_n d\mu.$$

q.e.d.

Aus Lemma 10.12 folgt, dass wir das Integral auf die Grenzwerte von monoton wachsenden Folgen von Stufenfunktionen 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 Stufenfunktionen 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 10.12 sowohl auf diese Folge, als auch auf die vertrauschten 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 10.13. *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 Stufenfunktionen 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 10.14. *(Eigenschaften der Lebesgue–integrablen Funktionen)*

(i) *$L^1(\mathbb{R}^d)$ ist ein Vektorraum über $\mathbb{R}$ und das Integral über Stufenfunktionen 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 auch*

$$\int f d\mu \geq 0.$$

(iii) *Wenn $f \in L^1(\mathbb{R}^d)$, dann ist auch $|f| \in L^1(\mathbb{R}^d)$ und es gilt*

$$\left|\int f d\mu\right| \leq \int |f|\, d\mu.$$

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 Stufenfunktionen 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 10.12

$$\int (g+\tilde{h})d\mu = \int g d\mu + \int \tilde{h} d\mu = \int \tilde{g} d\mu + \int h d\mu = \int (\tilde{g}+h)d\mu.$$

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 Stufenfunktionen.

(ii) Wenn $(g_n)_{n\in\mathbb{N}}$ und $(h_n)_{n\in\mathbb{N}}$ monoton wachsende Folgen von Stufenfunktionen mit beschränkten Integralen sind, so dass fast überall $\lim\limits_{n\to\infty} g_n - \lim\limits_{n\to\infty} h_n$ nicht negativ ist, dann ist auch fast überall $g = \lim\limits_{n\to\infty} g_n \geq \lim\limits_{n\to\infty} h_n = h$. Aus Lemma 10.12 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 Stufenfunktionen mit beschränkten Integralen sind wieder monoton wachsende Folgen von Stufenfunktionen 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 Stufenfunktionen 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 $(h_n)_{n\in\mathbb{N}} = (\min\{g_n, h_n\}_{n\in\mathbb{N}}$ von Stufenfunktionen mit beschränkten Integralen. Deshalb ist $|f| \in L^1(\mathbb{R}^d)$. Wegen (ii) folgt dann aus

$$-|f| \leq f \leq |f| \text{ auch } -\int |f| d\mu \leq \int f d\mu \leq \int |f| d\mu \text{ bzw. } \left|\int f d\mu\right| \leq \int |f| d\mu.$$

q.e.d.

Satz 10.15. *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 $Q \subset \mathbb{R}^d$, außerhalb dessen die Funktion verschwindet. Für jedes $n \in \mathbb{N}$ teilen wir jede der d-Kanten des Quaders in 2^n gleichlange Abschnitte. Dadurch wird der Quader jeweils eine disjunkte Vereinigung von $2^{d\cdot n}$ Quadern. Dann sei f_n die Stufenfunktion, die auf jedem der 2^{dn} Quader gleich dem Infimum der entsprechenden Funktionswerte von f ist. Offenbar ist $(f_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge von Stufenfunktionen, 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 aber auch ein $N \in \mathbb{N}$, so dass der Durchmesser von Q kleiner ist als $2^N \delta$. Für alle $n \geq N$, ist dann der Teilquader der 2^{dn} 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 10.16.* *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 Stufenfunktionen 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

$$g_n = \max\{\tilde{g}_n, \tilde{h}_n\} \text{ und } h_n = \min\{\tilde{g}_n, \tilde{h}_n\}.$$

Dann gilt fast überall

$$|f(x)| = \lim_{n\to\infty} g_n(x) - \lim_{n\to\infty} h_n(x).$$

Wenn $\int |f| d\mu = 0$ gilt also auch

$$\lim_{n\to\infty} \int g_n d\mu = \lim_{n\to\infty} \int h_n d\mu.$$

Für $\epsilon, \delta > 0$ sei also $N \in \mathbb{N}$, so dass für alle $m \geq N$ gilt

$$\lim_{n\to\infty} \int g_n d\mu - \int h_m d\mu \leq \epsilon\delta.$$

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, wobei

$$g(x) = \lim_{n\to\infty} g_n(x) \text{ (fast überall) }.$$

Sei also

$$\begin{aligned} A_1 &= \{x \in \mathbb{R}^d \mid g_1(x) - h_m(x) > \delta\} && \text{und für } n = 2, \ldots \\ A_n &= \{x \in \mathbb{R}^d \mid g_n(x) - h_m(x) > \delta \text{ und } g_{n-1}(x) - h_m(x) \leq \delta\}. \end{aligned}$$

Dann ist

$$\{x \in \mathbb{R}^d \mid g(x) - h_m(x) > \delta\} = \bigcup_{n=1}^{\infty} A_n \text{ und } \sum_{n=1}^{\infty} \mu(A_n) \leq \epsilon.$$

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.**

10.3 Das Riemann– und das Lebesgueintegral

In diesem Abschnitt wollen wir für $d = 1$ das Riemannintegral mit dem Lebesgueintegral in Beziehung setzen. Zunächst wollen wir die Riemann–integrablen Funktionen charakterisieren.

Satz 10.17. *(Lebesgue-Kriterium) Eine beschränkte Funktion auf einem Intervall $[a, b]$ ist genau dann Riemann–integrabel, wenn ihre Unstetigkeitsstellen eine Nullmenge bilden. Insbesondere sind alle Riemann–integrablen Funktionen auch Lebesgue–integrabel und die beiden Integrale stimmen überein.*

Beweis: Wir zerlegen für alle $n \in \mathbb{N}$ das Intervall $[a, b]$ in die Vereinigung der Intervalle $[a, b] = I_1 \cup \ldots \cup I_n$ mit

$$I_1 = \left[a, a + \frac{b-a}{n}\right], I_2 = \left(a + \frac{b-a}{n}\right] \ldots I_n = \left(a + (n-1)\frac{b-a}{n}, b\right]$$

Das entspricht genau der Partition $p_n \in \mathcal{P}[a, b]$ aus dem Beweis von Korollar 8.14. Für jede Funktion $f \in B_{\mathbb{R}}([a, b])$ seien

$$f_n(x) = \inf \{f(y) \mid y \in I_k \text{ mit } x \in I_k\} \quad F_n(x) = \sup \{f(y) \mid y \in I_k \text{ mit } x \in I_k\}$$

Dann ist $(f_n)_{n\in\mathbb{N}}$ eine monoton wachsende und $(F_n)_{n\in\mathbb{N}}$ eine monoton fallende Folge von Stufenfunktionen mit beschränkten Integralen. Wegen dem Riemann-Kriterium konvergiert $\lim\limits_{n\to\infty} \int f_n d\mu$ genau dann gegen $\lim\limits_{n\to\infty} \int F_n d\mu$, wenn $f \in \mathcal{R}[a, b]$.

Wenn die Menge aller $\{x \in [a, b] \mid \lim\limits_{n\to\infty} F_n(x) > \lim\limits_{n\to\infty} f_n(x)\}$ eine Nullmenge ist, folgt aus Lemma 10.12 $\lim\limits_{n\to\infty} \int F_n d\mu = \lim\limits_{n\to\infty} \int f_n d\mu$. Umgekehrt folgt aus dieser Gleicheit, dass für alle $\epsilon > 0$ die Gesamtvolumen der Mengen

$$S_{n,\epsilon} = \{x \in [a, b] \mid F_n(x) - f_n(x) \geq \epsilon\}$$

im Grenzwert $n \to \infty$ nach Null konvergieren. Also ist dann die Schnittmenge

$$S_\epsilon = \bigcap_{n=1}^{\infty} S_{n,\epsilon} = \left\{x \in [a, b] \mid \lim_{n\to\infty} F_n(x) - \lim_{n\to\infty} f_n(x) \geq \epsilon\right\}$$

aller dieser Mengen eine Nullmenge. Dann ist aber auch die Menge

$$\{x \in [a, b] \mid \lim_{n\to\infty} F_n(x) > \lim_{n\to\infty} f_n(x)\} = \bigcup_{m=1}^{\infty} S_{\frac{1}{m}}$$

eine Nullmenge. Deshalb ist f genau dann Riemann–integrabel, wenn die Menge $\{x \in [a,b] \mid \lim\limits_{n\to\infty} F_n(x) > \lim\limits_{n\to\infty} f_n(x)\}$ eine Nullmenge ist.

Wenn f bei x stetig ist, dann gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass alle $x' \in (x-\delta, x+\delta) \cap [a,b]$ auch $|f(x') - f(x)| < \epsilon/2$ erfüllen. Dann gilt für $n > (b-a)/\delta$

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

Dann gilt auch $\lim\limits_{n\to\infty} F_n(x) = \lim\limits_{n\to\infty} f_n(x)$. Wenn umgekehrt f bei x nicht ststig ist, dann gibt es ein $\epsilon > 0$, so dass für alle $\delta > 0$ gilt

$$\sup\{f(x') \mid x' \in (x-\delta, x+\delta) \cap [a,b]\} - \inf\{f(x') \mid x' \in (x-\delta, x+\delta) \cap [a,b]\} \geq \epsilon.$$

Wenn x außerdem nicht in $\{a + t(b-a) \mid t \in \mathbb{Q} \cap [0,1]\}$ enthalten ist, dann gilt $\lim\limits_{n\to\infty} F_n(x) - \lim\limits_{n\to\infty} f_n(x) \geq \epsilon$, weil dann für alle $n \in \mathbb{N}$ das Teilintervall der obigen Zerlegungen, das x entält, eine Umgebung von x ist. Weil aber die Menge $\{a + t(b-a) \mid t \in \mathbb{Q} \cap [0,1]\}$ abzählbar und damit eine Nullmenge ist, ist dann die Menge $\{x \in [a,b] \mid \lim\limits_{n\to\infty} F_n(x) > \lim\limits_{n\to\infty} f_n(x)\}$ genau dann eine Nullmenge, wenn die Menge der Unstetigkeitsstellen von f eine Nullmenge ist. **q.e.d.**

Beispiel 10.18. (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 $\sum\limits_{n=1}^{\infty} \epsilon 2^{-n} = \epsilon < 1$. Also ist die Folge

$$\left(\prod_{k=1}^{n}(1 - \chi_{I_k})\right)_{n\in\mathbb{N}}$$

eine monoton fallende Folge von Stufenfunktionen, die gegen eine Lebesgue–integrable Funktion konvergiert. Weil die rationalen Zahlen dicht in $[0,1]$ liegen, ist diese Funktion an allen Punkten im Komplement der offenen Menge $\bigcup\limits_{n=1}^{\infty} I_n$ unstetig. Also ist der Grenzwert von $\left(\prod\limits_{k=1}^{n}(1 - \chi_{I_k})\right)_{n\in\mathbb{N}}$ eine Lebesgue–integrable

Funktion, die nicht Riemann–integrabel 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 eine ungerade natürliche Zahl und χ die Funktion $\chi : \mathbb{R} \to \mathbb{R}, x \to \chi(x)$ 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 Folge von Stufenfunktionen auf $[0,1]$. Auf $[0,1]$ ist die Funktion also Lebesgue–integrabel. Sie hat aber offenbar Unstetigkeitsstellen bei allen Zahlen, deren p-adischen Bruchdarstellung aus endlich vielen Ziffern aus $\{0, 2, \ldots, p-1\}$ besteht, und am Ende eine 1 oder 2 hat. Offenbar besteht der Abschluss dieser Menge 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 Riemann–integrabel und Lebesgue–integrabel.

10.4 Der Satz von Fubini

Für jeden Quader Q in $\mathbb{R}^d \times \mathbb{R}$ und jedes $x \in \mathbb{R}^d$ ist die Funktion $\mathbb{R} \to \mathbb{R}, y \mapsto \chi_Q(x,y)$ eine Stufenfunktion auf $\mathbb{R}$. Wenn wir die Funktion integrieren erhalten wir eine Stufenfunktion auf dem $\mathbb{R}^d$:

$$\int \chi_Q(x,y) d\mu(y) = \begin{cases} \text{Länge der Kante in der Dimension } d+1 \text{ von } Q, \\ \text{wenn es ein } y \in \mathbb{R} \text{ gibt mit } (x,y) \in Q. \\ 0 \text{ sonst} \end{cases}$$

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 Stufenfunktionen auf $\mathbb{R}^d\times\mathbb{R}$ in die Stufenfunktionen auf $\mathbb{R}^d$. Und für jede Stufenfunktion f auf $\mathbb{R}^d\times\mathbb{R}$ gilt
$$\int\left(\int f(x,y)d\mu(y)\right)d\mu(x)=\int fd\mu.$$
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
$$\int\left(\int f(x,y)d\mu(y)\right)d\mu(x)=\int fd\mu.$$
Wenn $f\geq g$ zwei Stufenfunktionen auf $\mathbb{R}^d\times\mathbb{R}$ sind, dann erfüllten für jedes $x\in\mathbb{R}^d$ die entsprechenden Stufenfunktionen $f_x:y\to f(x,y)$ bzw. $g_x:y\to g(x,y)$ auch $f_x\geq g_x$. Wegen Proposition 10.8 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 Stufenfunktionen auf $\mathbb{R}^d\times\mathbb{R}$ in die Stufenfunktionen 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 Stufenfunktionen, die fast überall übereinstimmen, auch auf zwei fast überall definierte Grenzwerte von monoton wachsenden Stufenfunktionen abgebildet werden, die fast überall übereinstimmen.

Lemma 10.19. *Sei $S\subset\mathbb{R}^d\times\mathbb{R}$ eine Nullmenge. Dann ist fast überall in $x\in\mathbb{R}^d$, die Menge $S_x=\{y\in\mathbb{R}|(x,y)\in S\}$ eine Nullmenge von $\mathbb{R}$.*

Beweis: Wegen Satz 10.10 gibt es eine monoton wachsende Folge $(f_n)_{n\in\mathbb{N}}$von Stufenfunktionen auf $\mathbb{R}^d\times\mathbb{R}$ mit beschränkten Integralen, die auf S divergiert. Dann sind auch die entsprechenden Integrale $(\int f_nd\mu(y))_{n\in\mathbb{N}}$ über $\mathbb{R}$ monoton wachsende Stufenfunktionen mit beschränkten Integralen auf $\mathbb{R}^d$. Wegen Satz 10.9 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 Stufenfunktionen auf $\mathbb{R}$. Wegen Satz 10.9 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 10.20. *Die Integration über $\mathbb{R}$ induziert eine monotone Abbildung $L^1(\mathbb{R}^d \times \mathbb{R})$ 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 Stufenfunktionen auf $\mathbb{R}^d \times \mathbb{R}$ in die Stufenfunktionen auf $\mathbb{R}^d$ definiert und wegen Lemma 10.19, induziert sie eine Abbildung von den Äquivalenzklassen von den Grenzwerten von monoton wachsenden Folgen von Stufenfunktionen 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 Stufenfunktionen 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 10.21. *(Satz von Fubini) Für alle Funktionen $f \in L^1(\mathbb{R}^d \times \mathbb{R}^{d'})$ gilt*

$$\int \left(\int f(x,y) d\mu(y) \right) d\mu(x) = \int \left(\int f(x,y) d\mu(x) \right) d\mu(y)$$

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 $\int\limits_{-\infty}^{\infty} f(x_1, \ldots, x_d) dx_1$ 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 aber auch beliebig permutieren.

10.5 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 10.22. *(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 Stufenfunktionen 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

$$\left|\int \tilde{h}_{nm}d\mu - \int \tilde{h}_{n,m'}d\mu\right| \leq 2^n.$$

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 für alle $n\in\mathbb{N}$ die Folgen $(h_{nm})_{m\in\mathbb{N}}$ 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. Dann sind für alle $n\in\mathbb{N}$ auch die Integrale $\int g_{nm}d\mu)_{n,m\in\mathbb{N}}$ beschränkt. Für alle $n\in\mathbb{N}$ seien $g_n = \lim_{m\to\infty} g_{nm}$ und $h_n = \lim_{m\to\infty} h_{nm}$ und

$$\tilde{g}_m = \max\{g_{1m},\ldots,g_{mm}\} \text{ und } \tilde{h}_m = \max\{h_{1m},\ldots,h_{mm}\}.$$

Dann gilt für alle $n \in \mathbb{N}$ auch fast überall $f_n = g_n - h_n$. Außerdem ist die Folge $(h_n)_{n\in\mathbb{N}}$ 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 Stufenfunktionen mit beschränkten Integralen. Seien $\tilde{g}$ und $\tilde{h}$ die entsprechenden Grenzwerte. Für $n \leq m$ gilt

$$g_{nm} \leq \tilde{g}_m \text{ und } h_{nm} \leq \tilde{h}_m.$$

Also gilt für die entsprechenden Grenzwerte $m \to \infty$ fast überall

$$g_n \leq \tilde{g} \text{ und } h_n \leq \tilde{h}.$$

Weil aber $(g_n)_{n\in\mathbb{N}}$ und $(h_n)_{n\in\mathbb{N}}$ fast überall monoton wachsende Folgen sind, gilt auch fast überall

$$\tilde{g}_m \leq \max\{g_1, \ldots, g_m\} \leq g_m \text{ und } \qquad \tilde{h}_m \leq \max\{h_1, \ldots, h_m\} \leq h_m.$$

Also gilt auch fast überall $g = \tilde{g}$ und $h = \tilde{h}$ bzw. $f = \tilde{g} - \tilde{h} = g - h$. Dann folgt aus Lemma 10.12

$$\int f d\mu = \lim_{m\to\infty} \int \left(\tilde{g}_m - \tilde{h}_m\right) d\mu = \lim_{n\to\infty} \int (g_n - h_n) d\mu = \lim_{n\to\infty} \int f_n d\mu.$$

q.e.d.

Korollar 10.23. *(Norm $\|\cdot\|_1$) Auf $L^1(\mathbb{R}^d)$ definiert $\|\cdot\| : L^1(\mathbb{R}^d) \to \mathbb{R}, f \mapsto \|f\|_1 = \int |f| d\mu$ 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 10.24. *(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 $|f_n| \leq k$ 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

$$g_{nm} = \min\{f_n, f_{n+1}, \ldots, f_{n+m}\} \text{ und } h_{nm} = \max\{f_n, f_{n+1}, \ldots, f_{n+m}\}.$$

Für alle $n \in \mathbb{N}$ sind wegen den Eigenschaften des Lebesgueintegrals $(g_{nm})_{m\in\mathbb{N}}$ 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. Also konvergieren diese Folgen gegen Funktionen $(g_n)_{n\in\mathbb{N}}$ und $(h_n)_{n\in\mathbb{N}}$ in $L^1(\mathbb{R}^d)$.

$$g_n = \inf\{f_n, f_{n+1}, f_{n+2}, \ldots\} \text{ und } h_n = \sup\{f_n, f_{n+1}, f_{n+2}, \ldots\}$$

sind monotone Folgen in $L^1(\mathbb{R}^d)$ mit beschränktern Integralen. Also konvergieren fast überall $(g_n)_{n\in\mathbb{N}}$ und $(h_n)_{n\in\mathbb{N}}$ gegen f. Dann gilt aber auch

$$\int g_n d\mu \leq \int f_n d\mu \leq \int h_n d\mu \text{ und } \qquad \int f d\mu \leq \lim_{n\to\infty} \int f_n d\mu \leq \int f d\mu.$$

q.e.d.

Korollar 10.25. *(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 $(f_{n_m})_{n\in\mathbb{N}}$, so dass für alle $m \in \mathbb{N}$ gilt $\|f_{n_{m+1}} - f_{n_m}\|_1 \leq 2^{-m}$. Die Reihe $\left(\sum\limits_{m=1}^{n} \left|f_{n_{m+1}} - f_{n_m}\right|\right)_{n\in\mathbb{N}}$ erfüllt dann die Voraussetzungen des Satzes über die Monotone Konvergenz. Also konvergiert sie fast überall gegen eine Funktion $k \in L^1(\mathbb{R}^d)$. Dann konvergiert aber 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 10.26. *(Fatou's Lemma) Sei $(f_{n_m})_{m\in\mathbb{N}}$ 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 $g_{mn} = \min\{f_n, f_{n+1}, \ldots, f_{n+m}\}$. Dann erfüllt für alle $n \in \mathbb{N}$ die Folge $(g_{nm})_{m\in\mathbb{N}}$ die Voraussetzungen des Satzes über die Monotone Konvergenz. Also konvergieren diese Folgen gegen $(g_n)_{m\in\mathbb{N}}$ in $L^1(\mathbb{R}^d)$ mit $g_n = \inf\{f_n, f_{n+1}, \ldots\}$. die Folgen $(g_n)_{m\in\mathbb{N}}$ erfüllen wieder die Voraussetzungen des Satzes über die Monotone Konvergenz und konvergieren fast überall gegen f. Also gilt auch $f \in L^1(\mathbb{R}^d)$ und für alle $m \in \mathbb{N}_0$ $f_{n+m} \geq g_n$. Daraus folgt $\int f d\mu = \lim\limits_{n\to\infty} \int g_n d\mu$ und $\int f d\mu \leq \lim\limits_{n\to\infty} \inf\limits_{m\in\mathbb{N}_0} \int f_{n+m} d\mu$. **q.e.d.**

10.6 Messbare Mengen und Maße

In diesem Abschnitt untersuchen wir, wann wir eine Lebesgue–integrable Funktion über eine Teilmenge integrieren können. Das führt dann zu einer allgemeineren Definition vom Volumen von sogenannten messbaren Mengen. Dieses Volumen heisst Maß.

Definition 10.27. *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 Lebesgue–integrablen Funktionen mit der charakteristsichen Funktion χ_A von A den Teilraum von $L^1(\mathbb{R}^d)$ aller Lebesgue–integrablen Funktionen, die außerhalb von A verschwinden. Als diesen Teilraum definieren wird den Raum $L^1(A)$ aller Lebesgue–integrablen Funktionen auf A

Definition 10.28. *Für messbare Teilmengen A von $\mathbb{R}^d$ sei $L^1(A) \subset L^1(\mathbb{R}^d)$ der Teilraum aller Lebsegue–integrablen Funktionen auf $\mathbb{R}^d$, die ausserhalb von A verschwinden.*

Satz 10.29. (i) *Das Komplement einer messbaren Menge ist messbar.*

(ii) *Die abzählbare Schnittmenge von messbaren Mengen ist messbar.*

(iii) *Jede offene Menge ist messbar.*

Beweis: (i) Weil die charakteristische Funktion des Komplements gerade 1 minus der charakteritischen 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 $\left(f \prod\limits_{k=1}^{n} \chi_{A_k}\right)_{n\in\mathbb{N}}$ 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 $A = \bigcap\limits_{n=1}^{\infty} A_n$.
(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 Zentren und rationalen Kantenlängen ist abzählbar. Jeder Quader ist offenbar messbar. Weil aber jede offene Menge U gleich der Vereinigung aller der offenen Quader ist, deren Zentren und Kantenlängen 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 10.30. *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 $\mathcal{B}$ abgeschlosssen ist. Wenn X ein metrischer Raum ist, dann heissen dei Elemente der kleinsten σ–Algebra, die alle offenen (und abgeschlossenen) Mengen enhält, Borelmengen.*

Definition 10.31. *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 10.32. *(Lebesguemaß) Seien $\mathcal{B}(\mathbb{R}^d)$ die Borelmengen von $\mathbb{R}^n$. Dann definiert die Abbildung*

$$\mu : \mathcal{B}(\mathbb{R}^d) \to \bar{\mathbb{R}}_0^+, \qquad A \mapsto \int \chi_A d\mu$$

ein Maß auf den Borelmengen des $\mathbb{R}^d$, also ein Borelmaß.

Beweis: Wegen Satz 10.29 sind alle Borelmengen messbar. Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge von paarweise disjunkten Mengen. Dann ist für jede nichtnegative Funktion $g \in L^1(\mathbb{R}^d)$ die Folge

$$(gf_n)_{n\in\mathbb{N}} = \left(g\left(1 - \prod_{k=1}^{n}(1-\chi_{A_n})\right)\right)_{n\in\mathbb{N}}$$

eine monoton wachesen 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

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Wenn wir diese Aussage auf die Folge $(g_n)_{n\in\mathbb{N}}$ der charakteristischen Funktionen $g_n = \chi_{B(0,n)}$ von $B(0,n)$ anwenden, dann konvergiert die entsprechenden Folge $(g_f)_{n\mathbb{N}}$ entweder gegen eine Lebesgue–integrable Funktion, und $\mu\left(\bigcup_{n\in\mathbb{N}}\right)$ 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 10.33. *(Zerlegung der Eins) Eine Zerlegung der Eins bezüglich einer offenen Überdeckung, ist eine abzählbare Familie $(f_n)_{n\in\mathbb{N}}$ von glatten Funktionen $f_n : \mathbb{R}^d \to [0,1]$, so dass*

(i) *für jedes $x \in \mathbb{R}^d$ auf einer Umgebung von x nur endlich viele f_n ungleich Null sind.*

(ii) *Für alle $x \in \mathbb{R}^d$ gilt $\sum\limits_{n=1}^{\infty} f_n = 1$.*

(iii) *Jedes f_n außerhalb eines Elementes der offenen Überdeckung verschwindet.*

Satz 10.34. *(Existenz der Zerlegung der Eins) Jede offene Überdeckung des $\mathbb{R}^d$ besitzt eine Zerlegung der Eins.*

Beweis: Weil der $\mathbb{R}^d$ eine abzählbare Vereinigung von kompakten Mengen ist, besitzt jede Überdeckung auch eine abzählbare Teilüberdeckung $(U_n)_{n\in\mathbb{N}}$. Sei $0 < a < b < \infty$ und $f_{a,b} : \mathbb{R} \to [0,1], x \mapsto f_{a,b}(x)$ mit

$$f_{a,b}(x) = \begin{cases} 1 & \text{für } 0 \leq |x| \leq a \\ \exp\left(\frac{1}{x^2-b^2} \exp\left(\frac{-1}{x^2-a^2}\right)\right) & \text{für } a < |x| < b \\ 0 & \text{für } b \leq |x| \end{cases}$$

Dann ist $f_{a,b} \in C^\infty(\mathbb{R})$. Sei $(l_n)_{n\in\mathbb{N}}$ eine Abzählung der Elemente von $\mathbb{Z}^d$. Dann ist die Familie $(f_n)_{n\in\mathbb{N}}$ mit

$$f_n(x) = f_{\sqrt{d},d}(\|x - l_n\|) \prod_{m=1}^{n-1} (1 - f_{\sqrt{d},d}(\|x - l_m\|))$$

eine Zerlegung der Eins bezüglich der Überdeckung $(B(l_n, \sqrt{d}))_{n\in\mathbb{N}}$. Denn für jedes $n \in \mathbb{N}$ gilt

$$f_1(x) + \ldots + f_n(x) + (1 - f_{\sqrt{d},d}(x - l_1) \cdots (1 - f_{\sqrt{d},d}(x - l_n) = 1.$$

Weil die Abschlüsse $\overline{B(l_n, 1)}$ kompakt sind, gibt es für jedes $n \in \mathbb{N}$ eine endliche Teilüberckung von $\overline{B(l_n, 1)}$. Für jeden Punkt x von $\overline{B(l_n, 1)}$ gibt es ein $r(x) > 0$, so dass $B(x, r(x))$ in einer der entsprechenden endlich vielen offenen Mengen enthalten ist. Die offene Überdeckung $\{B(x, r(x)/2) | x \in \overline{B(l_n, 1)}\}$ von $\overline{B(l_n, 1)}$ besitzt eine endliche Teilüberdeckung $\{B(x_i, r_i/2) | i \in I_n\}$. Wir statten die Indexmengen I_n mit einer totalen Ordnung aus, indem wir den Elementen eine Reihenfolge geben. Die Funktionen

$$f_{n,i} = f_n(x) f_{r_{i/2},r_i}(\|x - x_i\|) \prod_{j<i, j\in I_n} (1 - f_{r_{j/2},r_j}(\|x - x_j\|))$$

bilden also eine Zerlegung der Eins. Zuletzt kann man jeweils solche Funktionen zu einer Funktion f_n aufsummieren, die außerhalb der Menge U_n verschwinden. **q.e.d.**

Lemma 10.35. (i) *Die Multiplikation mit einer beschränkten stetigen Funktion f definiert eine stetige lineare Abbildung $L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$, deren Norm beschränkt ist durch $\|f\|_\infty$.*

(ii) *Sei Φ eine bijektive Abbildung von einer offenen Menge $U \subset \mathbb{R}^d$ auf eine offene Menge O. Wenn Φ Lipschitzstetig ist mit Lipschitzkonstante L, dann induziert die Abbildung $f \mapsto f \circ \Phi^{-1}$ eine lineare Abbildung $L^1(U) \to L^1(O)$, deren Norm beschränkt ist durch L^d.*

(iii) *Sei Φ eine bijektive Abbildung von einer offenen Menge $U \subset \mathbb{R}^d$ auf eine offene Menge O. Wenn für alle $x \in U$ und ein $0 < \epsilon < 1$ gilt $\|\Phi(x) - x\| \leq \epsilon \|x\|$, dann gilt für alle $f \in L^1(U)$*

$$\left| \int\limits_O f \circ \Phi^{-1} d\mu - \int\limits_U f d\mu \right| \leq ((1+\epsilon)^d - 1)\|f\|_1.$$

Beweis: (i) Wegen Satz 10.15 ist die Multiplikation einer Stufenfunktion g mit einer stetigen Funktion f eine Lebesgue- integrable Funktion. Wenn f beschränkt ist, gilt $\|fg\|_1 \leq \|f\|_\infty \|g\| - 1$. Weil die Stufenfunktionen dicht in $L^1(\mathbb{R}^d)$ liegen, kann man dann die Abbildung $g \mapsto f \cdot g$ zu einer linearen Abbildung $L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$ fortsetzen, deren Norm beschränkt ist durch $\|f\|_\infty$.
(ii) Weil die Schnittmenge eines offenen Quaders in $\mathbb{R}^d$ mit U als offene Menge eine abzählbare Vereinigung von Quadern ist (siehe Beweis von Satz 10.29 (iii)), ist sie wegen Proposition 10.6 auch eine abzählbare disjunkte Vereinigung von Quadern in U. Also liegen die Stufenfunktionen in $L^1(U)$ dicht. Das Bild $\Phi[Q]$ eines Quaders in U liegt innerhalb des Quaders, dessen Zentrum das Bild des Zentrums von Q ist, und dessen Kantenlängen L mal der Kantenlängen von Q sind. Dieser Quader hat das Volumen $L^d \cdot \mu(Q)$. Wegen Satz 10.29 ist dieses Bild messbar. Für jede Stufenfunktion $f \in L^1(U)$ ist dann $f \circ \Phi^{-1}$ Lebesgue–integrabel und es gilt

$$\|f \circ \Phi^{-1}\|_1 \leq L^d \|f\|_1.$$

Wieder läßt sich die Abbildung $f \mapsto f \circ \Phi^{-1}$ zu einer linearen Abbildung $L^1(U) \to L^1(O)$ fortsetzen, deren Norm beschränkt ist durch L^d.
(iii) Wenn für alle $x \in U$ gilt $\|\Phi(x) - x\| \leq \epsilon \|x\|$, dann ist das Bild $\Phi[Q]$ eines Quaders in U enthalten in dem Quader, dessen Zentrum das Bild des Zentrums von Q ist, und dessen Kantenlängen $(1+\epsilon)$ mal der Kantenlängen von Q sind. Umgekehrt enthält $\Phi[Q]$ den Quader, dessen Zentrum das Bild des Zentrums von Q ist, und dessen Kantenlängen $(1-\epsilon)$ mal der Kantenlängen von Q sind. Also gilt auch

$$\left| \int \chi_{\Phi[Q]} d\mu - \int \chi_Q d\mu \right| \leq \mu(Q) \max\{1 - (1-\epsilon)^d, (1+\epsilon)^d - 1\} \leq \left((1+\epsilon)^d - 1\right) \mu(Q).$$

Wegen der Linearität und der Dreieckungleichung folgt die Behauptung für alle Stufenfunktionen f, und weil diese dicht liegen für alle $f \in L^1(\mathbb{R}^d)$. **q.e.d.**

10.7 Jacobi's Transformation von Maßen

In diesem Abschnitt untersuchen wir, wie sich die Integration unter Koordinatentransformationen verhält. Danach werden wir die Integration über den Rand definieren.

Satz 10.36. *Sei $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ eine invertierbare lineare Abbildung von $\mathbb{R}^d$ auf sich selber. Dann definiert die Abbildung*

$$\pi(A) : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d), f \mapsto \pi(A)f \text{ mit } (\pi(A)f)(x) = f(A^{-1}x)$$

eine stetige lineare Abbildung. Für alle $f \in L^1(\mathbb{R}^d)$ gilt

$$\int \pi(A) f d\mu = |\det(A)| \int f d\mu.$$

Beweis: Wegen Lemma 10.35 (ii) ist die Abbildung

$$\pi(A) : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d), f \mapsto \pi(A)f$$

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 Stufenfunktionen von Quadern mit gleichen Kantenlängen dicht in $L^1(\mathbb{R}^d)$. Für alle diese Funktionen f gilt aber

$$\int f \circ A^{-1} d\mu = \alpha(A) \int f d\mu \text{ mit } \alpha(A) = \int \chi_{A[[0,1]^d]} d\mu.$$

Also gilt das für alle $f \in L^1(\mathbb{R}^d)$. Wenn A und B zwei invertierbare Elemente in $\mathcal{L}(\mathbb{R}^d)$ sind, dann gilt $\pi(A \cdot B) = \pi(A) \cdot \pi(B)$. Also gilt auch $\alpha(A \cdot B) = \alpha(A)\alpha(B)$. Für diagonale Matrizen A gilt offenbar $\alpha(A) = |\det(A)|$. Also gilt das auch für diagonalisierbare Matrizen A. Aufgrund der Jordanschen Normalform, ist jede Matrix A mit paarweise verscheidenen 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 10.35 (iii) ist α stetig. Für alle A gilt dann $\alpha(A) = |\det(A)|$. **q.e.d.**

Satz 10.37. *(Jacobische Transformationsformel) Sei $\Phi : U \to O$ eine stetig differenzierbarere bijektive Abbildung von der offenen Menge $U \subset \mathbb{R}^d$ auf die offene Menge $O \subset \mathbb{R}^d$. Wenn $\frac{d\Phi}{dx}$ auf U in $\mathcal{L}(\mathbb{R}^d)$ invertierbar ist, dann ist die Abbildung $f \to \left|\det\left(\frac{d\Phi}{dx}\right)\right| (f \circ \Phi)$ eine Isometrie von $L^1(O)$ nach $L^1(U)$, d.h. für $f \in L^1(O)$ gilt*

$$\int \left|\det\left(\frac{d\Phi}{dx}\right)\right| (f \circ \Phi) d\mu = \int f d\mu.$$

Beweis: Weil $\frac{d\Phi}{dx}$ auf U in $\mathcal{L}(\mathbb{R}^d)$ invertierbar ist, ist auch Φ^{-1} stetig differenzierbar. Wir überdecken U und O durch offene Mengen, auf denen $\|\frac{d\Phi}{dx}\|$ und $\|\frac{d\Phi^{-1}}{dx}\|$ beschränkt sind. Mit Hilfe einer entsprechenden Zerlegung der Eins genügt es die Aussage für solche U und O zu zeigen, auf denen $\|\frac{d\Phi}{dx}\|$ und $\|\frac{d\Phi^{-1}}{dx}\|$ beschränkt sind. Dann folgt aus Lemma 10.35 (i) und (ii), dass die Abbildung $f \mapsto \left|\det\left(\frac{d\Phi}{dx}\right)\right|(f \circ \Phi)$ eine lineare Abbildung $L^1(U)$ nach $L^1(O)$ induziert. Für jedes $\epsilon > 0$ können wir dann U und O durch offene Mengen überdecken, auf denen $\|\frac{d\Phi}{dx} - A\| \leq \epsilon$ mit $A = \frac{d\Phi}{dx}(x_0)$ und einem geeigneten $x_0 \in U$. Dann folgt aus Lemma 10.35 (iii) und Satz 10.36, dass auf diesen kleinen Mengen O mit $f \in L^1(O)$ gilt

$$\int |\det(A)|(f \circ \Phi)d\mu - \int f d\mu| \leq \|A\| \left((1 + \|A^{-1}\|\epsilon)^d - 1\right) \|f\|_1.$$

Im Grenzwert $\epsilon \to 0$ folgt dann $\displaystyle\int \left|\det\left(\frac{d\Phi}{dx}\right)\right|(f \circ \Phi)d\mu = \int f d\mu.$ **q.e.d.**

Die Mengen, die wir im Gauß'schen Satz betrachten wollen, sollen einen zweimal differenzierbaren Rand haben.

Definition 10.38. *Eine offene Menge $\Omega \subset \mathbb{R}^d$ hat einen k–mal (stetig) differenzierbaren Rand $\partial\Omega = \overline{\Omega} \cap \Omega^c$, wenn es für jeden Punkt $x \in \overline{\Omega}$ eine offene Umgebung $U \subset \mathbb{R}^d$ und eine bijektive, k–mal (stetig) differenzierbare Abbildung $\Phi : U \to O \subset \mathbb{R}^d$ mit k–mal (stetig) differenzierbarer Umkehrabbildung gibt, so dass $\Phi(x) = 0$ und*

$$\begin{aligned}
\Phi[\Omega \cap U] &= O \cap \{x \in \mathbb{R}^d | x_d > 0\} \\
\Phi[\Omega^c \cap U] &= O \cap \{x \in \mathbb{R}^d | x_d \leq 0\} \\
\Phi[\partial\Omega \cap U] &= O \cap \{x \in \mathbb{R}^d | x_d = 0\}
\end{aligned}$$

Ein Vektor y gehört also genau dann im Punkt $x \in \partial\Omega$ zum Tangentialraum an den Rand, wenn gilt

$$\frac{d\Phi(x)}{dx} y \cdot e_d = 0 \quad \Longleftrightarrow \quad y \cdot \left(\frac{d\Phi(x)}{dx}\right)^t e_d = 0$$

Deshalb ist der äußere Normalvektor $N(x)$ im Punkt x gegeben durch

$$N(x) = -\left\|\left(\frac{d\Phi(x)}{dx}\right)^t e_d\right\|^{-1} \left(\frac{d\Phi(x)}{dx}\right)^t e_d.$$

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 10.39. **(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) *Die Bälle $B(y, r)$ im $\mathbb{R}^d$ werden beschrieben durch die 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 Nulsstelle bei $x = y$. Also haben die Bälle $B(y, r)$ einen unendlich oft differenzierbaren Rand.*

Lemma 10.40. *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 von einer offenen Umgebung U von $x \in \partial\Omega$ mit einmal stetig differenzierbarer Umkehrabbildung. Wenn f eine stetige Funktion auf $\partial\Omega$ ist, die ausserhalb von U verschwindet, dann ist das Lebesgueintegral*

$$\int\limits_{\Phi[\partial\Omega \cap U]} \left(\left\| \left(\frac{d\Phi}{dx}\right)^t e_d \right\| \cdot \left| \det\left(\frac{d\Phi}{dx}\right)\right|^{-1} f \right) \circ \Phi^{-1} d\mu_{\mathbb{R}^{d-1}}$$

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 Koordinaten in 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 $\left(\frac{d(\tilde{\Phi}\circ\Phi^{-1})}{dy}\right)^{-1}_{dd}$ von der Matrix $\frac{d(\tilde{\Phi}\circ\Phi^{-1})}{dy}$. Außerdem sind die Determinanten dieser Matrizen jeweils gleich dem Produkt dieser Einträge mit den Unterdeterminanten der entsprechenden $(d - 1) \times (d - 1)$ Matrizen, in denen jeweils die letzte Zeile und die letzte Spalte getrichen wurde. Also sind folgende Vektoren gleich:

$$\begin{aligned} e_d &= \left(\frac{d(\tilde{\Phi} \circ \Phi^{-1})}{dy}\right)^{-1}_{dd} \left(\frac{d(\tilde{\Phi} \circ \Phi^{-1})}{dy}\right)^t e_d \\ \left(\frac{d\tilde{\Phi}}{dx}\right)^t e_d &= \left(\frac{d\Phi}{dx}\right)^t \left(\frac{d(\tilde{\Phi} \circ \Phi^{-1})}{dy}\right)^t e_d = \left(\frac{d(\tilde{\Phi} \circ \Phi^{-1})}{dy}\right)^{-1}_{dd} \left(\frac{d\Phi}{dx}\right)^t e_d. \end{aligned}$$

Weil e_d in beide Mengen $\Phi[\Omega \cap U]$ und $\tilde{\Phi}[\Omega \cap U]$ hineinzeigt, ist der letzte Eintrag der letzten Zeile der Jacobimatrix $\frac{d(\tilde{\Phi} \circ \Phi^{-1})}{dx}$ positiv. Dann ergibt Satz 10.37

$$\int\limits_{\tilde{\Phi}[\partial\Omega \cap U]} \left(\left\| \left(\frac{d\tilde{\Phi}}{dx} \right)^t e_d \right\| \cdot \left| \det \left(\frac{d\tilde{\Phi}}{dx} \right) \right|^{-1} f \right) \circ \tilde{\Phi}^{-1} d\mu_{\mathbb{R}^{d-1}} =$$

$$= \int\limits_{\Phi[\partial\Omega \cap U]} \left\| \left(\frac{d\Phi^{-1}}{dx} \right)^t e_d \right\|^{-1} \left| \det \left(\frac{d\Phi^{-1}}{dx} \right) \right| (f \circ \Phi^{-1}) d\mu_{\mathbb{R}^{d-1}}. \qquad \textbf{q.e.d.}$$

Mit Hilfe dieses Lemma's definieren wir das Integral einer Funktion auf $\partial\Omega$.

Definition 10.41. *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 $\{0\} \times \mathbb{R}^{d-1}$ abbilden und dann das Integral mit Hilfe einer Zerlegung der Eins und dem vorangehenden Lemma auswerten.*

10.8 Der Gaußsche Satz

Satz 10.42. *(Gauß'scher Satz oder Divergenzsatz) Sei $\Omega \subset \mathbb{R}^d$ ein offenes 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 auf Ω Lebesgue–integrabel sind. 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 $f = \sum\limits_{i=1} e_i f_i$ 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 10.38 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 $\tilde{f}(y) = \left(\left|\det\left(\frac{d\Phi}{dx}\right)\right|^{-1}\left(\frac{d\Phi}{dx}f\right)\right) \circ \Phi^{-1}(y)$, wobei wir punktweise die Ableitung $\frac{d\Phi}{dx}$ auf den Vektor f wirken lassen. Dann folgt aus der Definition 10.41

$$\begin{aligned}
\int\limits_{\partial\Omega} f \cdot N d\sigma &= \int\limits_{\Phi[\partial\Omega\cap U]} \left(\left\|\left(\frac{d\Phi}{dx}\right)^t e_d\right\| \left|\det\left(\frac{d\Phi}{dx}\right)\right|^{-1} f \cdot N\right) \circ \Phi^{-1} d\mu_{\mathbb{R}^{d-1}} \\
&= -\int\limits_{\Phi[\partial\Omega\cap U]} \left(\left|\det\left(\frac{d\Phi}{dx}\right)\right|^{-1} f \cdot \left(\frac{d\Phi}{dx}\right)^t e_d\right) \circ \Phi^{-1} d\mu_{\mathbb{R}^{d-1}} \\
&= -\int\limits_{\Phi[\partial\Omega\cap U]} \tilde{f} \cdot e_d d\mu_{\mathbb{R}^{d-1}} = \int\limits_{\Phi[\partial\Omega\cap U]} \tilde{f} \cdot \tilde{N} d\mu_{\mathbb{R}^{d-1}}
\end{aligned}$$

Hierbei haben wir die äußere Normalen $N(x) = -\|\left(\frac{d\Phi}{dx}\right)^t e_d\|^{-1}\left(\frac{d\Phi}{dx}\right)^t e_d$ und $\tilde{N} = -e_d$ eingesetzt. Wenn wir f durch $\tilde{f}$ ausdrücken und Jacobi's Transformation von Maßen benutzen, erhalten wir

$$\int\limits_{\Omega} \nabla \cdot f d\mu = \int\limits_{\Phi[\Omega\cap U]} \left(\left|\det\left(\frac{d\Phi}{dx}\right)\right|^{-1}\left(\nabla_x \cdot \left|\det\left(\frac{d\Phi}{dx}\right)\right|\left(\frac{d\Phi}{dx}\right)^{-1} \tilde{f} \circ \Phi\right)\right) \circ \Phi^{-1} d\mu$$

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 9.19 (iii) und Korllar 9.43 erhalten wir

$$\begin{aligned}
\nabla_x \left(\frac{d\Phi}{dx}\right)^{-1} \tilde{f} \circ \Phi &= \sum_{i,j=1}^{d} \frac{\partial J^{-1}_{ij}}{\partial x_i} \tilde{f}_j \circ \Phi + \sum_{i,j,k=1}^{d} J^{-1}_{ij} J_{ki} \frac{\partial \tilde{f}_j}{\partial y_k} \circ \Phi \\
&= -\sum_{i,j,k,l=1}^{d} J^{-1}_{ij} \frac{\partial J_{jk}}{\partial x_i} J^{-1}_{kl} \tilde{f}_l \circ \Phi + (\nabla_y \cdot \tilde{f}) \circ \Phi.
\end{aligned}$$

Um den Gradienten der Determinante der Jacobimatrix zu berechnen, benötigen wir

Lemma 10.43. *Sei A eine differenzierbare Abbbildung von $(t_0 - \epsilon, t_0 + \epsilon)$ in die $n \times n$–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 $\frac{d}{dt}\det(A + tB)\Big|_{t=0} = \det(A)\,\mathrm{Spur}(A^{-1}B).$ **q.e.d.**

Damit erhalten wir

$$\left(\nabla_x \left|\det\left(\frac{d\Phi}{dx}\right)\right|\right) \cdot \left(\frac{d\Phi}{dx}\right)^{-1} \tilde{f} \circ \Phi \quad = \quad \left|\det\left(\frac{d\Phi}{dx}\right)\right| \sum_{i,j,k,l=1}^{d} J^{-1}_{ij} \frac{\partial J_{ji}}{\partial x_k} J^{-1}_{kl} \tilde{f}_l \circ \Phi.$$

Wegen dem Satz von Schwarz gilt

$$\frac{\partial J_{ji}}{\partial x_k} = \frac{\partial^2 y_j}{\partial x_k \partial x_i} = \frac{\partial^2 y_j}{\partial x_i \partial x_k} = \frac{\partial J_{jk}}{\partial x_i}, \text{ und damit auch } \int_\Omega \nabla \cdot f d\mu = \int_{\Phi[\Omega \cap U]} \nabla_y \cdot \tilde{f} d\mu.$$

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 Definitiopn 10.38 beschrieben ist, genau dann gilt, wenn er für die entsprechenden Funktionen $\tilde{f}$ gilt:

$$\int_{\Phi[\Omega \cap U]} \nabla_y \cdot \tilde{f} d\mu = \int_{\Phi[\partial\Omega \cap U]} \tilde{f} \cdot \tilde{N} d\mu_{\mathbb{R}^{d-1}}.$$

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

$$\int_{\Phi[\Omega \cap U]} \frac{\partial \tilde{f}_d}{\partial y_d} d\mu = -\int_{\Phi[\partial\Omega \cap U]} \tilde{f}_d d\mu_{\mathbb{R}^{d-1}}.$$

Daraus folgt der Gauß'sche Satz. **q.e.d.**

Index