Jürgen Müller
Analysis I

Skriptum zur Vorlesung
Wintersemester 2005/2006

Universität Trier
Fachbereich IV
Mathematik/Analysis

Dank an Elke Gawronski und Judith Wahlen für die Mithilfe bei der Erstellung

Inhaltsverzeichnis

1 Mengen und Abbildungen 1

2 Körper und das Prinzip der vollständigen Induktion 7

3 Geordnete Körper und reelle Zahlen 18

4 Komplexe Zahlen 31

5 Folgen 34

6 Reihen 49

7 Die Exponentialfunktion (elementare Funktionen I) 65

8 Metrische Räume 73

9 Topologische Grundbegriffe 84

10 Stetige Funktionen zwischen metrischen Räumen 90

11 Stetige Funktionen auf kompakten Mengen 98

12 Reellwertige Funktionen einer reellen Variablen 106

13 Differenzialrechnung von Funktionen einer reellen Variablen 114

14 Der Taylorsche Satz und Anwendungen 128

15 Funktionenfolgen und Funktionenreihen 137

16 Potenzreihen 149

17 Integralrechnung von Funktionen einer reellen Veränderlichen 156

18 Uneigentliche Integrale 170

19 Differenzialrechnung von Funktionen mehrerer Variablen 181

20 Mittelwert- und Taylorsatz für Funktionen mehrerer Variablen 191

21 Hauptsätze der mehrdimensionalen Analysis 204

22 DGLn: Beispiele und elementare Lösungsmethoden 217

23 Lösungstheorie für allgemeine gewöhnliche DGLn 225

24 Allgemeine lineare Differenzialgleichungen 237

25 Lineare Differenzialgleichungen mit konstanten Koeffizienten 253

26 Wege und Kurven 269

27 Lebesgue-Integral 285

28 Fourier-Reihen 298

29 Holomorphe Funktionen 314

30 Cauchyscher Integralsatz und Cauchysche Integralformel 324

31 Anwendungen der Cauchyschen Integralformel 340

32 Isolierte Singularitäten und Laurent-Reihen 349

33 Der Residuenkalkül 360

34 Der Rungesche Approximationssatz mit Anwendungen 374

35 Harmonische Funktionen 386

1 Mengen und Abbildungen

Wir starten mit einigen einführenden Definitionen und Ergebnissen aus der Theorie der Mengen und Abbildungen, die nicht nur Grundlage der Analysis sondern der gesamten Mathematik sind.
Unsere Darstellung gründet auf den von G. Cantor geprägten (sog. naiven) Mengenbegriff.

> Eine *Menge* M ist eine "Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen".

Ein solches Objekt x heißt *Element* der Menge M (Schreibweise: $x \in M$; ist x nicht Element von M, so schreiben wir $x \notin M$). Es gibt prizipiell zwei Möglichkeiten der Darstellung von Mengen: die aufzählende Schreibweise (etwa $M = \{1, 3, 5, 7\}$) und die beschreibende Schreibweise. Die beschreibende Schreibweise hat allgemein die Form $M = \{x : x \text{ hat die Eigenschaft } E\}$, wobei E irgendeine Eigenschaft ist (also im obigen Fall etwa $M = \{x : x \text{ ungerade natürliche Zahl kleiner als } 9\}$).

Definition 1.1 Es seien A, B Mengen.

1. A heißt *Teilmenge* von B, falls für jedes $x \in A$ auch $x \in B$ gilt. (Schreibweise: $A \subset B$).
2. A und B heißen *gleich* (Schreibweise $A = B$), falls $A \subset B$ und $B \subset A$.
3. Die Menge $B \setminus A := \{x : x \in B \,\text{und}\, x \notin A\}$ heißt *Differenz* von B und A. Ist $A \subset B$, so heißt $A^c := C_B(A) := B \setminus A$ *Komplement* von A (bzgl. B).
4. Die Menge ohne Elemente heißt *leere Menge* (Schreibweise: $\emptyset$).
5. Die Menge $A \cup B := \{x : x \in A \,\text{oder}\, x \in B\}$ heißt *Vereinigung* von A und B.
6. Die Menge $A \cap B := \{x : x \in A \,\text{und}\, x \in B\}$ heißt *Schnitt* von A und B.

Definition 1.2 Es seien A und B Mengen. Dann heißt

$$A \times B := \{(a, b) : a \in A, b \in B\} ,$$

also die Menge der geordneten Paare von Elementen aus A und B, das *Produkt* oder die *Produktmenge* von A und B. (Ein geordnetes Paar ist formal definiert als $(a, b) := \{\{a\}, \{a, b\}\}$; insbesondere gilt also $(a, b) = (\tilde{a}, \tilde{b})$ genau dann, wenn $a = \tilde{a}$ und $b = \tilde{b}$ ist.)

Beispiel 1.3 Ist $A = \{1, 2\}$ und $B = \{3\}$, so ist

$$A \times B = \{(1,3), (2,3)\} \text{ und } B \times A = \{(3,1), (3,2)\} .$$

Man beachte, dass $A \times B$ nicht mit $B \times A$ übereinstimmt.

Satz 1.4 *Es seien A_1, A_2, A_3 Mengen. Dann gilt*

1. $A_1 \cup A_2 = A_2 \cup A_1$,
   $A_1 \cap A_2 = A_2 \cap A_1$.
2. $A_1 \cup (A_2 \cup A_3) = (A_1 \cup A_2) \cup A_3$,
   $A_1 \cap (A_2 \cap A_3) = (A_1 \cap A_2) \cap A_3$.
   Wir schreiben deshalb auch kurz $A_1 \cup A_2 \cup A_3$ und $A_1 \cap A_2 \cap A_3$.
3. $A_1 \cap (A_2 \cup A_3) = (A_1 \cap A_2) \cup (A_1 \cap A_3)$,
   $A_1 \cup (A_2 \cap A_3) = (A_1 \cup A_2) \cap (A_1 \cup A_3)$.

Beweis.
1. und 2. folgen sofort aus Definition 1.2. Wir beweisen die erste Aussage von 3.
"$\subset$": Es sei

$$x \in A_1 \cap (A_2 \cup A_3) .$$

Dann ist $x \in A_1$ und $x \in A_2 \cup A_3$.

1. Fall: $x \in A_1$ und $x \in A_2$. Dann ist $x \in A_1 \cap A_2$, also auch $x \in (A_1 \cap A_2) \cup (A_1 \cap A_3)$.
2. Fall: $x \in A_1$ und $x \in A_3$. Dann ist $x \in A_1 \cap A_3$, also auch $x \in (A_1 \cap A_2) \cup (A_1 \cap A_3)$.

Also ist in jedem Fall $x \in (A_1 \cap A_2) \cup (A_1 \cap A_3)$.
Damit gilt $A_1 \cap (A_2 \cup A_3) \subset (A_1 \cap A_2) \cup (A_1 \cap A_3)$.
"$\supset$": Umgekehrt sei $x \in (A_1 \cap A_2) \cup (A_1 \cap A_3)$. Dann ist $x \in A_1 \cap A_2$ oder $x \in A_1 \cap A_3$. In beiden Fällen ist dann $x \in A_1 \cap (A_2 \cup A_3)$. Also folgt $(A_1 \cap A_2) \cup (A_1 \cap A_3) \subset A_1 \cap (A_2 \cup A_3)$.
Die zweite Aussage von 3. als [Ü]. $\square$

Satz 1.5 (Regeln von de Morgan) *Es seien A_1, A_2 Mengen, und es sei B eine Menge mit $A_1 \subset B$ und $A_2 \subset B$. Dann gilt*

1. $C_B(A_1 \cup A_2) = C_B(A_1) \cap C_B(A_2)$.
2. $C_B(A_1 \cap A_2) = C_B(A_1) \cup C_B(A_2)$.

Beweis.
1. "$\subset$": Es sei $x \in C_B(A_1 \cup A_2)$. Dann ist $x \in B$ und $x \notin A_1 \cup A_2$, also $x \in B$ und $x \notin A_1$ sowie $x \notin A_2$. Damit ist $x \in C_B(A_1)$ und $x \in C_B(A_2)$, d. h. $x \in C_B(A_1) \cap C_B(A_2)$.
"$\supset$": Es sei $x \in C_B(A_1) \cap C_B(A_2)$. Dann ist $x \in C_B(A_1)$ und $x \in C_B(A_2)$. Also ist $x \in B$ und $x \notin A_1$ sowie $x \notin A_2$. Dann ist $x \in B$ und $x \notin A_1 \cup A_2$, also $x \in C_B(A_1 \cup A_2)$.

2. [Ü]. $\square$

Definition 1.6 Es seien X und Y Mengen. Eine Teilmenge R von $X \times Y$ heißt *Relation* (zwischen X und Y). Ist speziell $X = Y$, so heißt R *Relation in* X. Eine Relation R zwischen X und Y heißt *Abbildung (von X nach Y)* bzw. *Funktion (von X nach Y)* falls gilt:

a) Für alle $x \in X$ existiert ein $y \in Y$ mit $(x, y) \in R$.

und

b) Sind $(x, y) \in R$ und $(x, \tilde{y}) \in R$ so gilt $y = \tilde{y}$.

Bemerkung und Definition 1.7 Ist R eine Abbildung von X nach Y, so ist jedem Wert $x \in X$ genau ein Wert $f(x)$ mit $(x, f(x)) \in R$ zugeordnet. Wir identifizieren R dann auch mit dieser Zuordnungsvorschrift f und schreiben $f : X \to Y$ oder $x \mapsto f(x)$. Weiter heißen X der *Definitionsbereich*, Y der *Zielbereich* und

$$W(f) := \{f(x) : x \in X\} = \{y \in Y : \exists\, x \in X \text{ mit } y = f(x)\}$$

der *Wertebereich (von f)*. Ferner setzen wir für $B \subset Y$

$$f^{-1}(B) := \{x \in X : f(x) \in B\}$$

($f^{-1}(B)$ heißt *Urbild von B unter f*) und für $A \subset X$

$$f(A) := \{f(x) : x \in A\} = \{y \in Y : \exists\, x \in A \text{ mit } y = f(x)\}$$

($f(A)$ heißt *Bild von A unter f*).

Ist $f : X \to Y$ und ist $X_0 \subset X$, so heißt $f_{|X_0} : X_0 \to Y$, definiert durch $f_{|X_0}(x) := f(x)$ für alle $x \in X_0$, *Einschränkung* von f auf X_0.

Satz 1.8 *Es seien X, Y Mengen und $f : X \to Y$. Dann gilt für $A_1, A_2 \subset X$ und $B_1, B_2 \subset Y$*

1. $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$,
2. $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$,

3. $f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$,

4. $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$.

Beweis.
1. "$\subset$": Es sei $y \in f(A_1 \cup A_2)$. Dann existiert ein $x \in A_1 \cup A_2$ mit $f(x) = y$. Ist $x \in A_1$, so ist $y = f(x) \in f(A_1) \subset f(A_1) \cup f(A_2)$. Entsprechend ist $y \in f(A_2) \subset f(A_1) \cup f(A_2)$ im Falle $x \in A_2$.
"$\supset$": Nach Definition gilt $f(A_1) \subset f(A_1 \cup A_2)$ und $f(A_2) \subset f(A_1 \cup A_2)$ also $f(A_1) \cup f(A_2) \subset f(A_1 \cup A_2)$.

2. "$\subset$": Es sei $x \in f^{-1}(B_1 \cup B_2)$. Dann ist $f(x) \in B_1 \cup B_2$.
Ist $f(x) \in B_1$, so ist $x \in f^{-1}(B_1)$ also auch $x \in f^{-1}(B_1) \cup f^{-1}(B_2)$. Entsprechend ist $x \in f^{-1}(B_2) \subset f^{-1}(B_1) \cup f^{-1}(B_2)$, falls $f(x) \in B_2$.
"$\supset$": Nach Definition gilt

$$f^{-1}(B_1) \subset f^{-1}(B_1 \cup B_2) \text{ und } f^{-1}(B_2) \subset f^{-1}(B_1 \cup B_2) \ ,$$

also $f^{-1}(B_1) \cup f^{-1}(B_2) \subset f^{-1}(B_1 \cup B_2)$.

3. Es sei $y \in f(A_1 \cap A_2)$. Dann existiert ein $x \in A_1 \cap A_2$ mit $f(x) = y$. Da $x \in A_1$ und $x \in A_2$ ist, folgt $y \in f(A_1) \cap f(A_2)$.

4. [Ü] □

Beispiel 1.9 Es seien

$$\mathbb{N} := \{1, 2, 3, \ldots\} = \{\text{natürliche Zahlen}\}$$

und

$$\mathbb{Z} := \{\text{ganze Zahlen}\} = \{0, \pm 1, \pm 2, \ldots\} \ .$$

Weiter seien $X = Y = \mathbb{Z}$ und $f : \mathbb{Z} \to \mathbb{Z}$ definiert durch

$$f(x) := \begin{cases} x, & \text{falls} \quad x \geq 0 \\ -x, & \text{falls} \quad x < 0 \end{cases} .$$

Dann gilt $W(f) = \mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Weiter ist etwa

$$f^{-1}(\{1, \ldots, n\}) = f^{-1}(\{1, \ldots, n\} \cup \{-1, -2, -3, \ldots\}) = \{-n, \ldots, -1, 1, \ldots, n\}$$

und $f^{-1}(\{-1, -2, -3, \ldots\}) = \emptyset$ sowie $f(\mathbb{N}) = \mathbb{N} = f(\mathbb{Z} \setminus \{0\})$. Ist $\tilde{f} : \mathbb{N}_0 \to \mathbb{Z}$ definiert durch

$$\tilde{f}(x) := x \qquad (x \in \mathbb{N}_0) \ ,$$

so ist zwar $\tilde{f}(x) = f(x)$ für alle $x \in \mathbb{N}_0$, aber $\tilde{f} \neq f$. Es gilt aber $f_{|\mathbb{N}_0} = \tilde{f}$.

Definition 1.10 Es seien X, Y Mengen. Eine Abbildung $f : X \to Y$ heißt

1. *surjektiv* (oder Abbildung von X *auf* Y), falls $W(f) = Y$ ist,
2. *injektiv* (oder *eineindeutige* Abbildung), falls für alle $y \in W(f)$ die Menge $f^{-1}(\{y\})$ einelementig ist (d. h. sind $x_1, x_2 \in X$ mit $f(x_1) = f(x_2)$, so ist $x_1 = x_2$),
3. *bijektiv*, falls f injektiv und surjektiv ist.

Beispiel 1.11 Es seien f und $\tilde{f}$ wie im B 1.9 Dann ist f weder surjektiv noch injektiv (es gilt $f^{-1}(\{(n\}) = \{n, -n\}$ für alle $n \in \mathbb{N}$); $\tilde{f}$ ist injektiv.

Definition 1.12 Es seien X, Y, Z Mengen und $f : X \to Y$ sowie $g : Y \to Z$ Abbildungen. Dann heißt $g \circ f : X \to Z$, definiert durch

$$(g \circ f)(x) := g(f(x)) \qquad (x \in X)$$

Verknüpfung von g und f (oder *Hintereinanderausführung von f und g*).

Satz 1.13 *Es seien X, Y, Z, U Mengen und $f : X \to Y, g : Y \to Z$ und $h : Z \to U$ Abbildungen. Dann gilt*

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

Beweis.
Es gilt $h \circ (g \circ f) : X \to U$ sowie $(h \circ g) \circ f : X \to U$ und für $x \in X$ ist

$$\begin{aligned}(h \circ (g \circ f))(x) &= h((g \circ f)(x)) = h(g(f(x))) = (h \circ g)(f(x)) = \\ &= ((h \circ g) \circ f)(x)\ .\end{aligned}$$

Definition 1.14 Es seien X, Y Mengen und es sei $f : X \to Y$ bijektiv. Wir definieren

$$f^{-1}(y) := x \qquad (y \in Y)\ ,$$

wobei $y = f(x)$. Die Abbildung $f^{-1} : Y \to X$ heißt *Umkehrabbildung von f*. Es gilt dabei

$$f^{-1} \circ f : X \to X\ , (f^{-1} \circ f)(x) = x \qquad (x \in X)\ ,$$

d. h. $f^{-1} \circ f = \mathrm{id}_X$, wobei $\mathrm{id}_X : X \to X$, definiert durch $\mathrm{id}_X(x) := x$ $(x \in X)$, die sog. identische Abbildung auf X bezeichnet. Genauso gilt $f \circ f^{-1} = \mathrm{id}_Y$ und außerdem ist auch $f^{-1} : Y \to X$ bijektiv.

Definition 1.15 Es sei $I \neq \emptyset$ eine Menge, und es seien A_α Mengen für alle $\alpha \in I$. (I nennt man dann "Indexmenge".) Dann heißt

$$\bigcup_{\alpha \in I} A_\alpha := \{x : x \in A_\alpha \text{ für ein } \alpha \in I\}$$

Vereinigung der Mengen A_α (über $\alpha \in I$).
Weiter heißt

$$\bigcap_{\alpha \in I} A_\alpha := \{x : x \in A_\alpha \text{ für alle } \alpha \in I\}$$

Durchschnitt der Mengen A_α (über $\alpha \in I$).
Ist speziell I endlich, etwa $I = \{j_1, \ldots, j_n\}$ so schreiben wir auch

$$A_{j_1} \cup \ldots \cup A_{j_n} = \bigcup_{k=1}^{n} A_{j_k} := \bigcup_{j \in I} A_j$$

und

$$A_{j_1} \cap \ldots \cap A_{j_n} = \bigcap_{k=1}^{n} A_{j_k} := \bigcap_{j \in I} A_j \ .$$

Beispiel 1.16 Es sei $\mathbb{P} = \{p : p \text{ Primzahl}\}$ und

$$A_p := \{kp : k \in \mathbb{N}\} \qquad (p \in \mathbb{P}) \ .$$

Dann ist

$$\bigcup_{p \in \mathbb{P}} A_p = \bigcup_{p \in \mathbb{P}} \{kp : k \in \mathbb{N}\} = \mathbb{N} \setminus \{1\}$$

und

$$\bigcap_{p \in \mathbb{P}} A_p = \emptyset \ .$$

2 Körper und das Prinzip der vollständigen Induktion

Den geeigneten Rahmen für die Fragen der Differential- und Integralrechnung, denen wir uns später zuwenden wollen bilden die reellen Zahlen. Wir werden die reellen Zahlen als den (i. w. einzigen) "vollständigen geordneten Körper" kennenlernen.

Definition 2.1 Es sei K eine Menge mit mindestens zwei Elementen, und es seien $+ : K \times K \to K$ und $\cdot : K \times K \to K$ Abbildungen. Dann heißt $K = (K, +, \cdot)$ *Körper*, falls folgende Rechenaxiome gelten:

(K.1) Für alle $x, y \in K$ ist $x + y = y + x$ und $x \cdot y = y \cdot x$
(Kommutativgesetze)

(K.2) Für alle $x, y, z \in K$ ist $(x + y) + z = x + (y + z)$ und $(x \cdot y) \cdot z = x \cdot (y \cdot z)$
(Assoziativgesetze)

(K.3) Es existiert ein Element $0 = 0_K \in K$ mit $x + 0 = x$ für alle $x \in K$
(Existenz einer Null)

(K.4) Es existiert ein Element $1 = 1_K \in K \setminus \{0\}$ mit $x \cdot 1 = x$ für alle $x \in K \setminus \{0\}$
(Existenz einer Eins)

(K.5) Für alle $x \in K$ existiert ein Element $-x \in K$ mit $x + (-x) = 0$ und
für alle $x \in K \setminus \{0\}$ existiert ein Element $x^{-1} \in K$ mit $x \cdot x^{-1} = 1$
(Existenz von inversen Elementen)

(K.6) Für alle $x, y, z \in K$ gilt $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$
(Distributivgesetz)

Statt $x \cdot y$ schreiben wir im folgenden auch kurz xy. Außerdem schreiben wir kurz $x \cdot y + z$ statt $(x \cdot y) + z$ (Punktrechnung vor Strichrechnung). Schließlich schreiben wir kurz $x - y$ statt $x + (-y)$ und $\frac{x}{y}$ (oder x/y) statt xy^{-1}.

Bemerkung 2.2 Wichtig für Axiom (K.5) ist, dass "die Null und die Eins eindeutig bestimmt sind", d. h. es existiert nur ein $0 \in K$ mit $x + 0 = x$ für alle $x \in X$ und nur ein $1 \in K \setminus \{0\}$ mit $x \cdot 1 = x$ für alle $x \in K \setminus \{0\}$.
(Denn: Es seien 0 und $0' \in K$ mit $x + 0 = x + 0' = x$ für alle $x \in K$.
Dann gilt
$$0' \overset{(K.3)}{=} 0' + 0 \overset{(K.1)}{=} 0 + 0' = 0 \; .$$
Entsprechendes gilt für 1)

Die Eindeutigkeit der inversen Elemente ist ein Spezialfall des folgenden Satzes

Satz 2.3 *Es sei $(K, +, \cdot)$ ein Körper, und es sei $a \in K$.*

1. *Für jedes $b \in K$ hat die Gleichung*

$$a + x = b$$

   genau eine Lösung, nämlich $x = b - a$.

2. *Ist $a \neq 0$, so hat für jedes $b \in K$ die Gleichung*

$$a \cdot x = b$$

   genau eine Lösung, nämlich $x = b/a$.

Beweis. Wir beweisen, nur Aussage 1. Der Beweis von 2. verläuft ähnlich.

1. Wir zeigen, dass $x = b - a$ die Gleichung löst: Es gilt

$$\begin{aligned} a + (b - a) &= a + (b + (-a)) \overset{(K.1)}{=} a + ((-a) + b) \overset{(K.2)}{=} (a + (-a)) + b \\ &\overset{(K.5)}{=} 0 + b \overset{(K.1)}{=} b + 0 \overset{(K.3)}{=} b\,. \end{aligned}$$

2. Wir zeigen, dass die Lösung eindeutig bestimmt ist: Ist $x' \in K$ mit $a + x' = b$, so gilt

$$\begin{aligned} x' &\overset{(K.3)}{=} x' + 0 \overset{(K.5)}{=} x' + (a + (-a)) \overset{(K.2)}{=} (x' + a) + (-a) \\ &\overset{(K.1)}{=} (a + x') + (-a) = b + (-a) = b - a\,. \end{aligned}$$

   Also ist $x' = b - a$.

$\square$

Wir stellen noch einige Rechenregeln zusammen, deren Beweis sich aus den Axiomen (K.1) bis (K.6) und 2.3 ergibt.

Satz 2.4 *Für alle $x, y \in K$ gilt*

1. $x \cdot 0 = 0$
2. $-(-x) = x$
3. $(x^{-1})^{-1} = x \quad (x \neq 0)$
4. *Aus $x \cdot y = 0$ folgt $x = 0$ oder $y = 0$*

5. $-(xy) = (-x)y = x(-y)$
6. $(xy)^{-1} = x^{-1}y^{-1} \quad (x, y \neq 0)$

Beweis. [Ü] □

Beispiel 2.5 1. Es sei $\mathbb{Q} := \{p/q : p \in \mathbb{Z}, q \in \mathbb{Z} \setminus \{0\}\} = \{\text{rationale Zahlen}\}$
(Dabei werden p/q und p'/q' als gleich angesehen, falls

$$pq' = qp'$$

gilt. Eine Eindeutigkeit der Darstellung erhält man etwa durch die Forderung

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

wobei $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ teilerfremd sind.)
Wir definieren wie üblich für $x = p/q$ und $y = r/s \in \mathbb{Q}$

$$x + y = \frac{p}{q} + \frac{r}{s} := \frac{ps + rq}{qs}$$

und

$$x \cdot y = \frac{p}{q} \cdot \frac{r}{s} := \frac{pr}{qs}$$

(man sieht leicht, dass diese Definitionen unabhängig von den gewählten Darstellungen für x und y sind).

2. Es sei $K = \{oh, ei\}$ mit den Rechenoperationen

$+$	oh	ei
oh	oh	ei
ei	ei	oh

$\cdot$	oh	ei
oh	oh	oh
ei	oh	ei

Dann ist $(K, +, \cdot)$ ein Körper mit $oh = 0_K$ und $ei = 1_K$ (Beweis: [Ü]).

3. $(\mathbb{N}, +, \cdot)$ und $(\mathbb{Z}, +, \cdot)$ mit der üblichen Addition und Multiplikation bilden keine Körper.

Definition 2.6 Wir definieren nun Summen und Produkte für mehr als zwei Summanden bzw. Faktoren: Sind $x_1, \ldots, x_n \in K$ für ein $n \in \mathbb{N}$, so setzen wir

$$\sum_{\nu=1}^{1} x_\nu := x_1 \quad \text{und} \quad \sum_{\nu=1}^{k+1} x_\nu := \left(\sum_{\nu=1}^{k} x_\nu\right) + x_{k+1} \text{ für } k = 1, \ldots, n-1$$

und
$$\prod_{\nu=1}^{1} x_\nu := x_1 \quad \text{und} \quad \prod_{\nu=1}^{k+1} x_\nu := \left(\prod_{\nu=1}^{k} x_\nu\right) \cdot x_{k+1} \text{ für } k = 1, \ldots, n-1.$$
Ist speziell $x_1 = \ldots = x_2 =: x$, so schreiben wir
$$\sum_{\nu=1}^{n} x_\nu = \sum_{\nu=1}^{n} x =: nx$$
und
$$\prod_{\nu=1}^{n} x_\nu = \prod_{\nu=1}^{n} x =: x^n \ .$$
Schließlich setzen wir noch
$$\begin{aligned} (-n)x &:= -(nx) && \text{für } n \in \mathbb{N} \\ 0 \cdot x &:= 0_K && \text{wobei 0 die Null in } \mathbb{Z} \text{ bezeichnet} \end{aligned}$$
und
$$\begin{aligned} x^{-n} &:= (x^{-1})^n \quad (x \neq 0,\, n \in \mathbb{N}) \\ x^0 &:= 1_K \ . \end{aligned}$$
Eng verbunden mit dem eben verwendeten Prinzip der rekursiven oder induktiven Definition ist das Beweisverfahren der vollständigen Induktion:
Für alle $n \in \mathbb{N}$ sei eine Aussage $A(n)$ gegeben. Zum Beweis der Behauptung

“Für alle $n \in \mathbb{N}$ gilt $A(n)$”

geht man oft folgendermaßen vor:

1. Man zeigt, dass $A(1)$ richtig ist (*Induktionsanfang*).
2. a) Man nimmt an, dass $A(k)$ (oder auch $A(1), \ldots, A(k)$) für ein $k \in \mathbb{N}$ richtig ist (*Induktionsannahme*).
   b) Man zeigt, dass aus der Richtigkeit von $A(k)$ (bzw. $A(1), \ldots, A(k)$), d. h. aus der Induktionsannahme, die Richtigkeit von $A(k+1)$ folgt (*Induktionsschritt*).

Dieses Beweisschema nennt man *Induktionsbeweis* oder *vollständige Induktion.* Aus 1. und 2. ergibt sich, dass $A(n)$ für alle $n \in \mathbb{N}$ richtig ist, denn es ist ja

$n = 1$: $A(1)$ richtig nach 1.
$n = 2$: $A(2)$ richtig nach 2., wenn dort $k = 1$ gesetzt wird
$n = 3$: $A(3)$ richtig nach 2., wenn dort $k = 2$ gesetzt wird

Manchmal möchte man statt $A(n)$ für alle $n \in \mathbb{N}$ auch $A(n)$ für alle $n \in \mathbb{N}_0, n \geq N$ für ein $N \in \mathbb{N}_0$ zeigen. Dann macht man den Induktionsanfang nicht für $n = 1$, sondern füer $n = N$ und den Induktionsschritt von k auf $k+1$ für beliebiges $k \geq N$.

Ein typischer Induktionsbeweis ist der Beweis zu

Satz 2.7 *Für alle $n \in \mathbb{N}$ gilt*

$$\sum_{\nu=1}^{n} \nu = \frac{n(n+1)}{2} \ .$$

Beweis.

1. Induktionsanfang: Für $n = 1$ gilt $\sum\limits_{\nu=1}^{1} \nu = \frac{1 \cdot 2}{2}$ (d. h. $A(1)$ gilt).

2. a) Induktionsannahme: Für ein $k \in \mathbb{N}$ gelte $\sum\limits_{\nu=1}^{k} \nu = \frac{k(k+1)}{2}$ (d. h. $A(k)$ gelte).

   b) Wir zeigen: aus a) folgt $\sum\limits_{\nu=1}^{k+1} \nu = \frac{(k+1)(k+2)}{2}$ (d. h. $A(k+1)$ folgt).

   Es gilt:

$$\begin{aligned} \sum_{\nu=1}^{k+1} \nu &= \left(\sum_{\nu=1}^{k} \nu \right) + (k+1) = \frac{k(k+1)}{2} + (k+1) = \frac{k(k+1) + 2(k+1)}{2} \\ &= \frac{(k+2)(k+1)}{2} \ . \end{aligned}$$

□

Wir kommen noch einmal auf das Summen- und das Produktzeichen zu sprechen. Dabei kann man natürlich auch allgemeinere Grenzen betrachten: Für $x_{m+1}, x_{m+2}, \ldots, x_n \in K$ setzen wir $\sum\limits_{\nu=m+1}^{n} x_\nu := \sum\limits_{\mu=1}^{n-m} x_{\mu+m}$ und $\prod\limits_{\nu=m+1}^{n} x_\nu := \prod\limits_{\mu=1}^{n-m} x_{\mu+m}$. Es gelten u. a. folgende Rechenregeln, die die Körperaxiome (K.1), (K.2) und (K.6) verallgemeinern (wobei $x_1, \ldots, x_n \in K,\ y_1, \ldots, y_m \in K,\ x \in K$)

1. $\sum\limits_{\nu=1}^{n} x_\nu = \sum\limits_{\nu=1}^{p} x_\nu + \sum\limits_{\nu=p+1}^{n} x_\nu$ und $\prod\limits_{\nu=1}^{n} x_\nu = \prod\limits_{\nu=1}^{p} x_\nu \cdot \prod\limits_{\nu=p+1}^{n} x_\nu \quad (1 \leq p \leq n-1)$,

2. $\sum\limits_{\nu=1}^{n} (x_\nu + y_\nu) = \sum\limits_{\nu=1}^{n} x_\nu + \sum\limits_{\nu=1}^{n} y_\nu \qquad$ (falls $n = m$),

3. $\sum\limits_{\nu=1}^{n} x x_\nu = x \sum\limits_{\nu=1}^{n} x_\nu$,

4. $\left(\sum\limits_{\nu=1}^{n} x_\nu \right) \left(\sum\limits_{\mu=1}^{m} y_\mu \right) = \sum\limits_{\nu=1}^{n} \sum\limits_{\mu=1}^{m} x_\nu y_\mu \left(= \sum\limits_{\mu=1}^{m} \sum\limits_{\nu=1}^{n} x_\nu y_\mu \right)$.

Nach 1. ist auch die Schreibweise

$$\sum_{\nu=1}^{n} x_\nu = x_1 + \ldots + x_n \quad \text{und} \quad \prod_{\nu=1}^{n} x_\nu = x_1 \cdot \ldots \cdot x_n$$

gerechtfertigt.
Weiter kann man zeigen, dass für $m, m_1, m_2 \in \mathbb{Z}$ und $x_1, x_2, x \neq 0$ folgende Potenzgesetze gelten:

$$\begin{aligned} x^{m_1} x^{m_2} &= x^{m_1+m_2} , \\ x_1^m x_2^m &= (x_1 x_2)^m , \\ (x^{m_1})^{m_2} &= x^{m_1 m_2} . \end{aligned}$$

Schließlich sei J eine beliebige endliche Indexmenge (mit n Elementen) und $\varphi : \{1, \ldots, n\} \to J$ bijektiv. Wir setzen dann ganz allgemein für $x_j \in K$ ($j \in J$):

$$\sum_{j \in J} x_j := \sum_{\nu=1}^{n} y_\nu \text{ und } \prod_{j \in J} x_j := \prod_{\nu=1}^{n} y_\nu$$

wobei $y_\nu := x_{\varphi(\nu)}$ (wichtig dabei: die Reihenfolge ist beliebig).

Alle obigen Aussagen ergeben sich per Induktion aus den Körperaxiomen ([Ü]).

Eine der wichtigsten Formeln für Summen und Potenzen in Körpern ist die

Satz 2.8 *(geometrische Summenformel) Für alle $x \in K \setminus \{1\}$ und alle $n \in \mathbb{N}$ gilt:*

$$\frac{1 - x^n}{1 - x} = \sum_{\nu=0}^{n-1} x^\nu \tag{2.1}$$

Beweis. Wir zeigen: $1 - x^n = (1 - x) \sum\limits_{\nu=0}^{n-1} x^\nu$. Da $x \neq 1$ und damit $1 - x \neq 0$ ist, ergibt sich hieraus die Behauptung.
Es gilt

$$\begin{aligned} (1-x)\sum_{\nu=0}^{n-1} x^\nu &= \sum_{\nu=0}^{n-1} x^\nu - x \sum_{\nu=0}^{n-1} x^\nu = \\ &= \sum_{\nu=0}^{n-1} x^\nu - \sum_{\nu=0}^{n-1} x \cdot x^\nu = \sum_{\nu=0}^{n-1} x^\nu - \sum_{\nu=0}^{n-1} x^{\nu+1} \\ &= \sum_{\nu=0}^{n-1} x^\nu - \sum_{\mu=1}^{n} x^\mu = 1 - x^n . \end{aligned}$$

□

Neben der geometrischen Summenformel gibt es eine weitere Rechenformel in Körpern, die von fundamentaler Bedeutung ist, die sog. binomische Formel. Hierauf wollen wir nun lossteuern. Es handelt sich dabei um eine Summenformel für die Ausdrücke $(a+b)^n$, wobei $a, b \in K$ und $n \in \mathbb{N}$ ist. Bekanntlich gilt

$$\begin{aligned} (a+b)^1 &= a+b \\ (a+b)^2 &= a^2+2ab+b^2 \\ (a+b)^3 &= a^3+3a^2b+3ab^2+b^3 \;. \end{aligned}$$

Um eine allgemeine Formel angeben zu können, brauchen wir

Definition 2.9 1. Wir definieren $n!$ (*"n-Fakultät"*) für $n \in \mathbb{N}_0$ durch

$$0! = 1$$

und

$$n! := \prod_{\nu=1}^{n} \nu \;= 1 \cdot 2 \cdot \ldots \cdot n \qquad (n \in \mathbb{N}) \;.$$

2. Für $n, \nu \in \mathbb{N}_0$ setzen wir

(i) $$\binom{n}{\nu} := \frac{\prod_{k=1}^{\nu}(n-k+1)}{\nu!} \;= \frac{n(n-1)\cdot \ldots \cdot (n-\nu+1)}{\nu!} \qquad (\nu \in \mathbb{N})$$

(ii) $$\binom{n}{0} := 1$$

Die Zahlen $\binom{n}{\nu}$ heißen *Binomialkoeffizienten.*

Es gilt also etwa

$$\begin{aligned} 6! &= 1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6 = 720 \;, \qquad 10! = 3.628.800 \;, \\ \binom{7}{5} &= \frac{7\cdot 6\cdot 5\cdot 4\cdot 3}{5!} = 21 \\ \binom{n}{n} &= \frac{n(n-1)\cdots(n-n+1)}{n!} = 1 = \binom{n}{0} \;, \qquad \text{wobei } n \in \mathbb{N} \;. \end{aligned}$$

Wir stellen einige Eigenschaften der Binomialkoeffizienten zusammen.

Satz 2.10 *Es seien $n, \nu \in \mathbb{N}_0$. Dann gilt*

1. $$\binom{n}{\nu} = \frac{n!}{\nu!(n-\nu)!} \qquad , \quad \textit{falls} \;\; \nu \le n$$

2. $\binom{n}{\nu} = 0 \qquad , \quad \textit{falls} \;\; \nu > n$

3. $\binom{n}{\nu} = \binom{n}{n-\nu} \qquad , \quad \textit{falls} \;\; \nu \leq n$

Beweis.

1. Es gilt für $0 < \nu \leq n$

$$\binom{n}{\nu} = \frac{n(n-1)\cdots(n-\nu+1)}{\nu!} = \frac{n(n-1)\cdots(n-\nu+1)}{\nu!} \cdot \frac{(n-\nu)!}{(n-\nu)!} = \frac{n!}{\nu!(n-\nu)!}$$

und für $\nu = 0$ ist

$$\binom{n}{0} = 1 = \frac{n!}{0!n!} \, .$$

2. Für $\nu > n$ ist $n - \nu + 1 \leq 0$, also

$$\binom{n}{\nu} = \frac{n(n-1)\cdots(n-\nu+1)}{\nu!} = 0$$

da ein Faktor im Zähler $= 0$ ist.

3. Nach 1. gilt

$$\binom{n}{\nu} = \frac{n!}{\nu!(n-\nu)!} = \frac{n!}{(n-(n-\nu))!(n-\nu)!} = \binom{n}{n-\nu} \, .$$

□

Besonders wichtig ist folgende Rekursionsformel:

Satz 2.11 *Für $n, \nu \in \mathbb{N}$ gilt*

$$\binom{n+1}{\nu} = \binom{n}{\nu-1} + \binom{n}{\nu}$$

Beweis. Nach 2.10.1 gilt für $\nu \in \{1, \ldots, n\}$

$$\begin{aligned}
\binom{n}{\nu-1} + \binom{n}{\nu} &= \frac{n!}{(\nu-1)!(n-\nu+1)!} + \frac{n!}{\nu!(n-\nu)!} = \\
&= \frac{n!}{\nu!(n+1-\nu)!} \left\{ \frac{\nu!}{(\nu-1)!} + \frac{(n+1-\nu)!}{(n-\nu)!} \right\} = \\
&= \frac{n!}{\nu!(n+1-\nu)!} \{\nu + (n+1-\nu)\} = \frac{(n+1)!}{\nu!(n+1-\nu)!} = \binom{n+1}{\nu} \, .
\end{aligned}$$

Für $\nu = n + 1$ ist nach 2.10.2

$$\binom{n}{\nu-1} + \binom{n}{\nu} = \binom{n}{n} + 0 = 1 = \binom{n+1}{\nu}$$

und für $\nu > n + 1$ sind beide Seiten $= 0$. □

Ordnet man die Binomialkoeffizienten $\binom{n}{\nu}$ in einem dreieckigen Schema an, wobei in der n-ten Zeile die Koeffizienten $\binom{n}{0}, \dots, \binom{n}{n}$ stehen, so entsteht das sog. *Pascal'sche Dreieck*:

$$\begin{array}{c}
\binom{0}{0} \\
\binom{1}{0} \ \binom{1}{1} \\
\binom{1}{0} \ \binom{2}{1} \ \binom{2}{2} \\
\vdots \quad \vdots \quad \vdots \\
\vdots \qquad \vdots \qquad \vdots \\
\binom{n}{0} \ \binom{n}{1} \cdots \binom{n}{\nu-1} \ \binom{n}{\nu} \cdots \binom{n}{n} \\
\binom{n+1}{0} \ \binom{n+1}{1} \ \cdots \binom{n+1}{\nu} \cdots \ \binom{n+1}{n+1}
\end{array}$$

Die ersten Zeilen berechnen sich etwa unter Ausnutzung von S. 2.11 zu

$$\begin{array}{c}
1 \\
1 \quad 1 \\
1 \quad 2 \quad 1 \\
1 \quad 3 \quad 3 \quad 1 \\
1 \quad 4 \quad 6 \quad 4 \quad 1 \\
1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1 \\
1 \quad 6 \quad 15 \quad 20 \quad 15 \quad 6 \quad 1
\end{array}$$

Damit gilt

Satz 2.12 *Es sei K ein Körper, und es seien $x, y \in K$ sowie $n \in \mathbb{N}_0$. Dann gilt*

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

Beweis.

1. Wir betrachten zunächst den Spezialfall $y = 1$ und führen den Beweis durch vollständige Induktion nach n.

   (i) Für $n = 0$ gilt $(1+x)^0 = 1 = \sum_{\nu=0}^{0} \binom{0}{\nu} x^\nu$.

   (ii) Für ein $k \in \mathbb{N}_0$ gelte
$$(1+x)^k = \sum_{\nu=0}^{k} \binom{k}{\nu} x^\nu \, .$$

   Dann gilt
$$\begin{aligned}
(1+x)^{k+1} &= (1+x)(1+x)^k = (1+x)\sum_{\nu=0}^{k} \binom{k}{\nu} x^\nu = \\
&= \sum_{\nu=0}^{k} \binom{k}{\nu} x^\nu + \sum_{\nu=0}^{k} \binom{k}{\nu} x^{\nu+1} = \\
&= \sum_{\nu=0}^{k} \binom{k}{\nu} x^\nu + \sum_{\nu=1}^{k+1} \binom{k}{\nu-1} x^{\nu} = \\
&= \binom{k+1}{0} x^0 + \sum_{\nu=1}^{k} \binom{k+1}{\nu} x^\nu + \binom{k+1}{k+1} x^{k+1} \\
&= \sum_{\nu=0}^{k+1} \binom{k+1}{\nu} x^\nu \, .
\end{aligned}$$

2. Für $y \neq 0$ gilt nach 1.
$$\begin{aligned}
(x+y)^n &= y^n \left(\frac{x}{y} + 1\right)^n = y^n \sum_{\nu=0}^{n} \binom{n}{\nu} (x/y)^\nu = \\
&= \sum_{\nu=0}^{n} \binom{n}{\nu} y^n x^\nu / y^\nu = \sum_{\nu=0}^{n} \binom{n}{\nu} x^\nu y^{n-\nu} \, .
\end{aligned}$$

   Ist $y = 0$ so gilt $(x+y)^n = x^n$ und $\sum_{\nu=0}^{n} \binom{n}{\nu} x^\nu y^{n-\nu} = \binom{n}{n} x^n = x^n$.

□

Beispiel 2.13 Es gilt etwa
$$\begin{aligned}
(x+y)^6 = \sum_{\nu=0}^{6} \binom{6}{\nu} x^\nu y^{n-\nu} &= 1 \cdot y^6 + 6 \cdot xy^5 + 15x^2y^4 + 20x^3y^3 + \\
&+ 15x^4y^2 + 6x^5y + 1 \cdot x^6 \, .
\end{aligned}$$

Bemerkung 2.14 Als Spezialfälle aus S. 2.12 ergeben sich interessante Beziehungen für das Pascal'sche Dreieck:
Für $x = 1, y = 1$ ergibt sich

$$2^n = (1+1)^n = \sum_{\nu=0}^{n} \binom{n}{\nu} 1^\nu 1^{n-\nu} = \sum_{\nu=0}^{n} \binom{n}{\nu} ,$$

d. h. die Summe der Binomialkoeffizienten in der n-ten Zeile des Pascal'schen Dreiecks ergibt stets 2^n.
Für $x = -1, y = 1$ ergibt sich für $n \in \mathbb{N}$

$$0 = 0^n = (-1+1)^n = \sum_{\nu=0}^{n} \binom{n}{\nu} (-1)^\nu 1^{n-\nu} = \sum_{\nu=0}^{n} \binom{n}{\nu} (-1)^\nu ,$$

d. h. versieht man die Binomialkoeffizienten in der n-ten Zeile jeweils abwechselnd mit dem Vorzeichen $+$ und $-$, so erhält man als Summe 0.
Für $n = 6$ gilt etwa

$$1 + 6 + 15 + 20 + 15 + 6 + 1 = 64 = 2^6$$

und

$$1 - 6 + 15 - 20 + 15 - 6 + 1 = 0 .$$

3 Geordnete Körper und reelle Zahlen

Wir betrachten jetzt Körper, die neben den algebraischen Strukturen "+" und "·" eine Ordnungsstruktur haben.

Definition 3.1 Es sei $K = (K, +, \cdot)$ ein Körper. Dann heißt $K = (K, +, \cdot, <)$ *geordnet*, wenn auf K eine Relation $<$ gegeben ist, die folgenden Ordnungsaxiomen genügt.

(O.1) Für alle $x, y \in K$ gilt genau eine der Beziehungen
$x = y$ oder $x < y$ oder oder $y < x$
(Trichotomiegesetz).

(O.2) Aus $x < y$ und $y < z$ folgt $x < z$ (Transitivgesetz).

(O.3) Aus $x < y$ folgt $x + z < y + z$ für alle $z \in K$ (1. Monotoniegesetz).

(O.4) Aus $x < y$ und $0 < z$ folgt $xz < yz$ (2. Monotoniegesetz).

Für $x < y$ schreiben wir auch $y > x$. Außerdem bedeutet $x \leq y$, dass entweder $x = y$ oder $x < y$ gilt. Dann schreibt man auch $y \geq x$. Wir nennen $x \in K$ *positiv*, falls $x > 0$ gilt und *negativ*, falls $x < 0$ gilt.

Beispiel 3.2 1. Es sei $\mathbb{Q} = (\mathbb{Q}, +, \cdot)$ mit der üblichen Addition und Multiplikation (B. 2.5.1). Wir setzen

$$\frac{p_1}{q_1} < \frac{p_2}{q_2} \qquad (p_1, p_2 \in \mathbb{Z}, q_1, q_2 \in \mathbb{N})$$

falls $p_1 q_2$ kleiner als $p_2 q_1$ in den ganzen Zahlen $\mathbb{Z}$. Dann ist $(\mathbb{Q}, +, \cdot, <)$ ein geordneter Körper.
2. Ist $K = \{oh, ei\}$ wie in B. 2.5.2, so existiert keine Relation $<$ auf K so, K geordnet ist ([Ü]).

Satz 3.3 *Es sei $K = (K, +, \cdot, <)$ ein geordneter Körper. Dann gilt*

1. *Es ist $x > 0$ genau dann, wenn $-x < 0$ ist,*
2. *Aus $x < 0$ und $y < 0$ folgt $xy > 0$,*
3. *Für $x \neq 0$ ist $x^2 > 0$, insbesondere also $1 > 0$,*
4. *Aus $x > 0$ folgt $x^{-1} > 0$.*

Beweis.

1. Aus $0 < x$ folgt mit (O.3)

$$-x = 0 + (-x) < x + (-x) = 0\,,$$

d. h. $-x < 0$. Entsprechend folgt aus $-x < 0$ auch $0 = x + (-x) < x + 0 = x$.

2. Aus $x < 0$ folgt $-x > 0$ nach 1. Wegen $y < 0$ ergibt sich mit (O.4)
$$-(xy) = y(-x) < 0(-x) = 0 \,,$$
also $xy > 0$ mit 1.

3. Ist $x > 0$, so folgt $0 < xx = x^2$ aus (O.4).
Ist $x < 0$, so folgt $x^2 = x \cdot x > 0$ aus 2.

4. Angenommen, es ist $x^{-1} < 0$ (beachte $x^{-1} \neq 0$). Dann folgt mit (O.4) $1 = xx^{-1} < x0 = 0$ im Widerspruch zu 3.

□

Das folgende Ergebnis hat insbesondere die interessante Konsequenz, dass ein geordneter Körper stets unendlich viele Elemente enthält. Es gibt also keine endlichen geordneten Körper!

Satz 3.4 *Es sei K ein geordneter Körper und es seien $x, y \in K$ mit $x < y$. Dann existiert ein $z \in K$ mit $x < z < y$*

Beweis. Aus $x < y$ folgt
$$x + x < x + y < y + y \,.$$
Also ist
$$x(2 \cdot 1_K) < x + y < y(2 \cdot 1_K) \,.$$
Nach S. 3.3.3 ist $1_K > 0$ also auch $2 \cdot 1_K > 0$ und nach S. 3.3.4 damit auch $(2 \cdot 1_K)^{-1} > 0$. Durch Multiplikation mit $(2 \cdot 1_K)^{-1}$ ergibt sich
$$x < (x + y)(2 \cdot 1_K)^{-1} < y \,.$$
Mit $z := (x + y)(2 \cdot 1_K)^{-1}$ ergibt sich die Behauptung. □

Als nächstes beschäftigen wir uns mit dem Betrag eines Körperelementes

Definition 3.5 Es sei K ein geordneter Körper, und es sei $x \in K$.
Dann heißt
$$|x| := \begin{cases} x, & \text{falls} \quad x \geq 0 \\ -x, & \text{falls} \quad x < 0 \end{cases}$$
der *(absolute) Betrag* von x.

Satz 3.6 *Es sei K ein geordneter Körper, und es seien $x, y \in K$.*
Dann gilt

1. $|x| = |-x| \geq 0$ *und* $|x| = 0$ *genau dann, wenn* $x = 0$,

2. $x \leq |x|\ , -x \leq |x|$,

3. $|xy| = |x||y|\ ,\quad |x^2| = |x|^2 = x^2$,

4. *(Dreiecksungleichung)*
   $|x + y| \leq |x| + |y|\ ,\quad |x - y| \leq |x| + |y|$.

Beweis. Die Aussagen 1., 2., 3. ergeben sich direkt aus D. 3.5. Wir zeigen 4. Dazu seien $x, y \in K$.
1. Fall: $x + y \geq 0$. Dann gilt mit 2. und (O.3)

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

2. Fall: $x + y < 0$. Dann gilt $|x + y| = -(x + y) = -x - y$, also wieder mit 2. und (O.3)

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

Also ist stets $|x + y| \leq |x| + |y|$. Der zweite Teil folgt sofort aus

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

□

Bemerkung 3.7 Induktiv erhält man aus S. 3.6.4 leicht die allgemeine Dreiecksungleichung: Es sei K ein geordneter Körper. Dann gilt für alle $n \in \mathbb{N}$ und für alle $x_1, \ldots, x_n \in K$

$$\left|\sum_{j=1}^{n} x_j\right| \leq \sum_{j=1}^{n} |x_j|\ .$$

Satz 3.8 *(Bernoullische Ungleichung)*
Es sei K ein geordneter Körper, und es seien $x \in K, x \geq -1$ sowie $n \in \mathbb{N}$. Dann gilt

$$(1 + x)^n \geq 1 + nx\ .$$

Beweis. Es sei $x \geq -1$ fest. Wir führen den Beweis per Induktion nach n.

1. Für $n = 1$ ist die Behauptung klar.

2. Für ein $k \in \mathbb{N}$ gelte $(1+x)^k \geq 1+kx$. Dann folgt mit (O.4)

$$\begin{aligned}(1+x)^{k+1} &= (1+x)(1+x)^k \geq (1+x)(1+kx) = 1+kx+x+x(kx) \\ &= 1+(k+1)x+kx^2 \geq 1+(k+1)x\end{aligned}$$

(man beachte: es ist $x^2 \geq 0$ und damit auch $kx^2 \geq 0$).

□

Es stellt sich heraus, dass jeder geordnete Körper in gewisser Weise den Körper der rationalen Zahlen als "Teilkörper" enthält. Um dies zu präzisieren brauchen wir den Begriff der Einbettung

Definition 3.9 Es seien K und L Körper. L heißt *eingebettet* in K, falls eine injektive Abbildung $\varphi : L \to K$ existiert mit

$$\begin{aligned}\varphi(x_1 + x_2) &= \varphi(x_1) + \varphi(x_2) \\ \varphi(x_1 \cdot x_2) &= \varphi(x_1) \cdot \varphi(x_2)\end{aligned}$$

für alle $x_1, x_2 \in L$ (d. h. φ erhält die Körperstruktur). Weiter heißen K und L *isomorph*, falls eine bijektive Abbildung mit diesen Eigenschaften existiert.
Sind L und K geordnete Körper, so heißt L *ähnlich eingebettet* in K, falls eine injektive Abbildung φ wie oben existiert, die zusätzlich

$$\varphi(x_1) < \varphi(x_2) \quad \text{für alle} \quad x_1, x_2 \quad \text{mit} \quad x_1 < x_2$$

erfüllt (d. h. φ erhält die Ordnungsstruktur). Schließlich heißen K und L *(ähnlich) isomorph*, falls eine bijektive Abbildung mit diesen Eigenschaften existiert.

Es gilt damit

Satz 3.10 *Es sei $K = (K, +, \cdot, <)$ ein geordneter Körper. Dann ist $\mathbb{Q}$ ähnlich eingebettet in K.*

Ist φ die Abbildung, die $\mathbb{Q}$ in K einbettet, so sind bezüglich der Körpereigenschaften $\mathbb{Q}$ und $\varphi(\mathbb{Q})$ nicht unterschiedlich. Deshalb identifiziert man $\mathbb{Q}$ und $\varphi(\mathbb{Q})$ und schreibt auch kurz $\mathbb{Q} \subset K$.

Beweisskizze Es sei 1_K das Einselement in K. Dann gilt für alle $n, m \in \mathbb{Z}$:

$$(n+m)1_K = n1_K + m1_K \quad \text{und} \quad (nm)1_K = n1_K \cdot m1_K$$

Ist $n < m$, so folgt außerdem
$$n1_K < m1_K$$
also insbesondere auch $n1_K \neq 0$ im Falle $n \neq 0$. (Beweis per Induktion; [Ü].)
Wir betrachten die Abbildung $\varphi : \mathbb{Q} \to K$ mit
$$\varphi(x) = \varphi\left(\frac{p}{q}\right) := \frac{p1_K}{q1_K} \qquad (x = \frac{p}{q},\ p \in \mathbb{Z}, q \in \mathbb{N}) \ .$$
Dann ist φ wohldefiniert (man beachte: $q1_K \neq 0$ und ist $x = p/q = p'/q'$, so folgt $p1_K/q1_K = p'1_K/q'1_K$). Außerdem ist φ injektiv und erfüllt die Bedingungen aus D. 3.9, wie man nachrechnen kann. □

Um eine vernünftige Grundlage für das Betreiben von Analysis zu haben, brauchen wir (geordnete) Körper, die eine weitere Struktureigenschaft erfüllen, die sog. Vollständigkeit. Leider erfüllt der Körper $\mathbb{Q}$ der rationalen Zahlen diese Bedingung nicht. Dies ist eine Konsequenz aus dem folgenden, wohl auch aus der Schule bekannten Sachverhalt.

Satz 3.11 *Es gibt kein $x \in \mathbb{Q}$ mit $x^2 = 2$.*

Beweis.

1. Zunächst gilt: Ist $m \in \mathbb{Z}$, und ist m^2 gerade, so ist auch m gerade. (Denn: Wäre m ungerade, also $m = 2m_0 + 1$ für ein $m_0 \in \mathbb{Z}$, so folgte
$$m^2 = (2m_0 + 1)^2 = 4m_0^2 + 4m_0 + 1 \ ,$$
also wäre auch m^2 ungerade.)

2. Angenommen, es existiert ein $x = p/q \in \mathbb{Q}$ mit $x^2 = 2$. O.E. seien p, q so, dass $q \in \mathbb{N}$ und p und q keine gemeinsamen Teiler haben. Aus
$$\left(\frac{p}{q}\right)^2 = 2$$
folgt $p^2 = 2q^2$, d. h. p^2 ist gerade. Nach 1. ist dann auch p gerade, d. h. $p = 2p_0$ für ein $p_0 \in \mathbb{Z}$. Dann ist
$$2q^2 = p^2 = 4p_0^2$$
d. h. $q^2 = 2p_0^2$, also q^2 und damit auch q gerade. Also haben p und q den gemeinsamen Faktor 2 im Widerspruch zur Teilerfremdheit von p und q. Damit ist die Annahme falsch d. h. es existiert kein $x \in \mathbb{Q}$ mit $x^2 = 2$.

□

Bevor wir definieren, was wir unter vollständigen geordneten Körpern verstehen, betrachten wir Schranken für Teilmengen von geordneten Körpern.

Definition 3.12 Es sei K ein geordneter Körper, und es sei $M \subset K$.

1. M heißt *nach oben beschränkt*, wenn ein $\overline{s} \in K$ existiert mit
$$x \leq \overline{s} \quad \text{für alle} \quad x \in M \ .$$
Ein solches $\overline{s}$ heißt dann *obere Schranke* von M.

2. M heißt *nach unten beschränkt*, wenn ein $\underline{s} \in K$ existiert mit
$$x \geq \underline{s} \quad \text{für alle} \quad x \in M \ .$$
Ein solches $\underline{s}$ heißt dann *untere Schranke* von M.

3. M heißt *beschränkt* wenn M nach oben und nach unten beschränkt ist.

Beispiel 3.13 Es sei $K = \mathbb{Q}$ und
$$M := \{x \in \mathbb{Q} : \quad x \geq 0, \quad x^2 < 2\} \ .$$
Dann ist M beschränkt, denn $\underline{s} = 0$ ist eine untere Schranke und $\overline{s} = 3/2$ ist eine obere Schranke von M (Ist $x > 3/2$, so folgt $x^2 > (3/2)^2 = 9/4 > 2$, d. h. $x \notin M$).

Mit einer oberen Schranke $\overline{s}$ von M ist natürlich jedes $\overline{\overline{s}} \in K$ mit $\overline{\overline{s}} > \overline{s}$ ebenfalls eine obere Schranke für M. Es stellt sich in natürlicher Weise die Frage nach "kleinsten" oberen Schranken.

Definition 3.14 Es sei K ein geordneter Körper, und es sei $M \subset K$.

1. Ein $\overline{\xi} \in K$ heißt *kleinste obere Schranke* (oder *Supremum*) von M, falls

   a) $\overline{\xi}$ obere Schranke von M ist

   und

   b) für jede obere Schranke $\overline{s}$ von M gilt $\overline{s} \geq \overline{\xi}$.

   Wir schreiben dann $\overline{\xi} = \sup M$. Weiter nennen wir $\overline{\xi}$ *Maximum* von M (und schreiben $\overline{\xi} = \max M$), falls zusätzlich $\overline{\xi} \in M$ gilt.

2. Ein $\underline{\xi} \in K$ heißt *größte untere Schranke* (oder *Infimum*) von M, falls

a) $\underline{\xi}$ untere Schranke von M ist

und

b) für jede untere Schranke $\underline{s}$ von M gilt $\underline{s} \leq \underline{\xi}$.

Wir schreiben dann $\underline{\xi} := \inf M$. Weiter nennen wir $\underline{\xi}$ *Minimum* von M (und schreiben $\underline{\xi} = \min M$), falls zusätzlich $\underline{\xi} \in M$ gilt.

Bemerkung 3.15 Aus der Definition ergibt sich sofort, dass für jedes M höchstens ein Supremum und ein Infimum existieren.

Beispiel 3.16 Es sei $K = \mathbb{Q}$.

1. Ist $M = \{x \in \mathbb{Q} : 0 \leq x \leq 1\}$, so gilt

$$0 = \inf M (= \min M) \quad \text{und} \quad 1 = \sup M (= \max M) \ .$$

2. Ist $M = \{x \in \mathbb{Q} : 0 < x < 1\}$, so gilt ebenfalls

$$0 = \inf M \quad \text{und} \quad 1 = \sup M \ .$$

Man sieht, dass i.a. $\inf M$ und $\sup M$ nicht in M liegen müssen!

3. Es sei $M = \{x \in \mathbb{Q} : x \geq 0, \quad x^2 < 2\}$.
Hier ist $\inf M = 0$, es existiert aber kein Supremum von M.

(Denn: Angenommen, es existiert eine kleinste obere Schranke $\overline{\xi}$ von M in $\mathbb{Q}$. Dann ist entweder $\overline{\xi}^2 > 2$ oder $\overline{\xi}^2 = 2$ oder $\overline{\xi}^2 < 2$.
1. Fall: $\overline{\xi}^2 > 2$. Dann ist

$$s := \overline{\xi} - \frac{\overline{\xi}^2 - 2}{2\overline{\xi}} < \overline{\xi} \ , \qquad s \in \mathbb{Q}$$

und

$$\begin{aligned} s^2 &= \left(\overline{\xi} - \frac{\overline{\xi}^2 - 2}{2\overline{\xi}}\right)^2 = \overline{\xi}^2 - 2\overline{\xi}\frac{\overline{\xi}^2 - 2}{2\overline{\xi}} + \left(\frac{\overline{\xi}^2 - 2}{2\overline{\xi}}\right)^2 = \\ &= 2 + \left(\frac{\overline{\xi}^2 - 2}{2\overline{\xi}}\right)^2 > 2 \ . \end{aligned}$$

Also ist auch s eine obere Schranke, die kleiner ist als $\overline{\xi}$. Widerspruch!
2. Fall: $\overline{\xi}^2 < 2$. Wir betrachten

$$s := \overline{\xi} + h \in \mathbb{Q} \qquad \text{mit} \quad h := \frac{2 - \overline{\xi}^2}{2\overline{\xi} + 1} \ .$$

Da $\overline{\xi} \geq 1$ ist, folgt $0 < h \leq 1/3$ und damit $h^2 < h$. Folglich ist

$$\begin{aligned} s^2 &= (\overline{\xi} + h)^2 = \overline{\xi}^2 + 2h\overline{\xi} + h^2 < \xi^2 + 2h\overline{\xi} + h = \\ &= \overline{\xi}^2 + (2\overline{\xi} + 1)h = \overline{\xi}^2 + (2 - \overline{\xi}^2) = 2 \, . \end{aligned}$$

Also ist $s \in M$ und $s > \overline{\xi}$, d. h. $\overline{\xi}$ ist keine obere Schranke von M. Widerpruch!
3. Fall: $\overline{\xi}^2 = 2$. Dies widerspricht S. 3.11. Also haben wir in allen Fällen einen Widerspruch.)

Definition 3.17 Es sei K ein geordneter Körper. Dann heißt K *vollständig*, wenn K das Vollständigkeitsaxiom erfüllt.

(V) Ist $M \subset K$, $M \neq \emptyset$ und nach oben beschränkt, so existiert $\sup M$.

(Man sieht leicht ([Ü]), dass dies äquivalent ist zur Forderung, dass für jede nichtleere, nach unten beschränkte Menge $M \subset K$ das Infimum $\inf M$ existiert.)

Während es zahlreiche, wesentlich verschiedene geordnete Körper gibt kann man zeigen, dass zwei vollständige geordnete Körper "i.w. gleich" (d. h. ähnlich isomorph) sind.
Nach B. 3.13 und B. 3.16.3 ist $\mathbb{Q}$ nicht vollständig. Eines der fundamentalen Ergebnisse der Analysis ist die Aussage, dass es einen vollständigen geordneten Körper gibt.

Satz 3.18 *Es gibt einen vollständigen geordneten Körper und jeder vollständige geordnete Körper ist zu diesem ähnlich isomorph. Wir nennen diesen Körper $\mathbb{R}$ (Körper der reellen Zahlen) und dessen Elemente bezeichnen wir als reelle Zahlen.*

Beweisskizze Wir wollen uns darauf beschränken, den Existenzbeweis zu skizzieren. Die Elemente von $\mathbb{R}$ werden als gewisse Teilmengen von $\mathbb{Q}$ definiert (sog. Dedekind'sche Schnitte):
$\alpha \subset \mathbb{Q}$ heißt (Dedekind'scher) Schnitt, falls gilt:

(i) $\alpha \neq \emptyset \, , \quad \alpha \neq \mathbb{Q}$.

(ii) Ist $p \in \alpha$ so gilt $q \in \alpha$ für alle $q < p$.

(iii) Ist $p \in \alpha$ so existiert ein $r \in \alpha$ mit $p < r$.

Sind α, β Schnitte, so definiert man

$$\alpha + \beta := \{r + s : r \in \alpha, s \in \beta\} \, .$$

Dann ist auch $\alpha + \beta$ ein Schnitt, d. h. $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$.
(Denn: offenbar ist $\alpha + \beta \neq \emptyset$.
Sind $r' \notin \alpha, s' \notin \beta$, so ist $r' + s' > r + s$ für alle $r \in \alpha, s \in \beta$. Also ist $r' + s' \notin \alpha + \beta$ und insbesondere $\alpha + \beta \neq \mathbb{Q}$. Es sei $p \in \alpha + \beta$. Dann existieren $r \in \alpha, s \in \beta$ mit $p = r + s$. Ist $q < p$, so folgt $q - s < r$, also $q - s \in \alpha$ und damit $q = (q - s) + s \in \alpha + \beta$, d. h. (ii) gilt.
Schließlich sei $t \in \alpha$ so, dass $t > r$ ist. Dann gilt $p < t + s$ und $t + s \in \alpha + \beta$. Also gilt auch (iii).)

(K.1): Aus $r + s = s + r$ für alle $r \in \alpha, s \in \beta$ folgt
$\alpha + \beta = \{r + s : r \in \alpha; s \in \beta\} = \{s + r : s \in \beta, r \in \alpha\} = \beta + \alpha$

(K.2): analog

(K.3): Wir definieren $0_{\mathbb{R}} := \{r \in \mathbb{Q} : r < 0\}$. Dann ist $0_{\mathbb{R}}$ ein Schnitt.
Wir zeigen: $\alpha + 0_{\mathbb{R}} = \alpha$ für alle Schnitte α.
Es sei α ein Schnitt und $r \in \alpha, s \in 0_{\mathbb{R}}$. Dann ist $r + s < r$,
also $r + s \in \alpha$, d. h. $\alpha + 0_{\mathbb{R}} \subset \alpha$.
Umgekehrt sei $p \in \alpha$ beliebig. Wählt man ein $r > p, r \in \alpha$ so gilt
$p - r \in 0_{\mathbb{R}}$ und $p = r + (p - r) \in \alpha + 0_{\mathbb{R}}$. Also ist $\alpha \subset \alpha + 0_{\mathbb{R}}$.

(K.5): Es sei α ein Schnitt. Wir setzen
$$\beta := \{p \in \mathbb{Q} : -p - r \notin \alpha \quad \text{für ein } r > 0\}.$$
Man kann zeigen, dass dann $\alpha + \beta = 0_{\mathbb{R}}$ gilt.

Wir definieren nun eine Relation $<$ auf den Schnitten (also auf $\mathbb{R}$) durch

$$\alpha < \beta :\Leftrightarrow \alpha \subset \beta,\ \alpha \neq \beta\ .$$

Man rechnet damit nach, dass die Ordnungsaxiome $(O.1), (O.2)$ und $(O.3)$ erfüllt sind.

Nun kann man auch eine Multiplikation in $\mathbb{R}$ definieren. (Dies erweist sich als etwas schwieriger als die Definition der Addition, insbesondere weil Produkte negativer rationaler Zahlen positiv sind). Zunächst definiert man daher eine Multiplikation auf $\mathbb{R}_+ := \{\alpha \in \mathbb{R} : \alpha > 0_{\mathbb{R}}\}$ durch

$$\alpha \cdot \beta := \{p \in \mathbb{Q} : \exists r \in \alpha, s \in \beta \text{ mit } r, s > 0 \text{ und } p < rs\}$$

(wobei $\alpha, \beta > 0_{\mathbb{R}}$). Anschließend definieren wir

$$\alpha 0_{\mathbb{R}} := 0_{\mathbb{R}} \alpha := 0_{\mathbb{R}}$$

und

$$\alpha\beta := \begin{cases} (-\alpha)(-\beta) & , \text{ falls } \alpha < 0_{\mathbb{R}} \ , \ \beta < 0_{\mathbb{R}} \\ -[(-\alpha)\beta] & , \text{ falls } \alpha < 0_{\mathbb{R}} \ , \ \beta > 0_{\mathbb{R}} \\ -[\alpha(-\beta)] & , \text{ falls } \alpha > 0_{\mathbb{R}} \ , \ \beta < 0_{\mathbb{R}} \end{cases}$$

Hiermit wird $(\mathbb{R}, +, \cdot, <)$ zu einem geordneten Körper.

Wir zeigen, dass dieser vollständig ist:
Dazu sei $M \subset \mathbb{R}, M \neq \emptyset$ nach oben beschränkt. Wir definieren

$$\beta := \bigcup_{\alpha \in M} \alpha \,.$$

Dann kann man zeigen, dass β ein Schnitt, also $\beta \in \mathbb{R}$ ist. Natürlich gilt $\alpha \leq \beta$ für alle $\alpha \in M$ und ist $\gamma < \beta$, so existiert ein $s \in \beta$ mit $s \notin \gamma$. Wegen $s \in \beta$ ist $s \in \alpha$ für ein $\alpha \in M$. Also ist $\gamma < \alpha$, d. h. γ ist keine obere Schranke von M. Folglich ist $\beta = \sup M$.)

Auf den Beweis der zweiten Aussage, dass je zwei vollständige geordnete Körper ähnlich isomorph zueinander sind, wollen wir nicht eingehen. (Literatur etwa: Ebbinghaus u.a., *Zahlen*, 2. Auflage, Springer Verlag, Berlin, 1988) □

Wie in S3.10 gesehen ist $\mathbb{Q}$ eingebettet in $\mathbb{R}$. Indem wir $\mathbb{Q}$ mit dieser Einbettung identifizieren, haben wir dann auch $\mathbb{Q} \subset \mathbb{R}$ (und damit $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{R}$).
Wir stellen nun einige wichtige Ergebnisse zusammen, die sich aus der Vollständigkeit von $\mathbb{R}$ ergeben.

Satz 3.19 *Zu jeder Zahl $x \in \mathbb{R}$ existiert ein $n \in \mathbb{N}$ mit $n > x$.*

Beweis. Angenommen, es gibt ein $x \in \mathbb{R}$ mit $n \leq x$ für alle $n \in \mathbb{N}$. Dann ist x obere Schranke von $\mathbb{N}$. Da $\mathbb{N} \neq \emptyset$ ist, existiert nach (V) eine kleinste obere Schranke $\overline{\xi} \in \mathbb{R}$ von $\mathbb{N}$ (d. h. $\overline{\xi} = \sup \mathbb{N}$). Für diese gilt $n \leq \overline{\xi}$ für alle $n \in \mathbb{N}$.
Die reelle Zahl $\overline{\xi} - 1 < \overline{\xi}$ kann keine obere Schranke von $\mathbb{N}$ sein (da $\overline{\xi}$ kleinste obere Schranke ist), d. h. es existiert ein $n_0 \in \mathbb{N}$ mit

$$\overline{\xi} - 1 < n_0 \qquad \text{bzw.} \qquad \overline{\xi} < n_0 + 1$$

da $n_0 + 1 \in \mathbb{N}$ ist, ist $\overline{\xi}$ keine obere Schranke von $\mathbb{N}$, also Widerspruch! □

Satz 3.20 *(Wohlordnungsprinzip für $\mathbb{N}$)*
Es sei $A \subset \mathbb{N}, A \neq \emptyset$. Dann ist $\inf A \in A$ (d. h. es gilt $\inf A = \min A$).

Beweis. Da A nach unten beschränkt ist, existiert $n_0 := \inf A \in \mathbb{R}$ nach (V). Wir zeigen: $n_0 \in A$.
Da $n_0 + 1/2$ keine untere Schranke von A ist, existiert ein $n_1 \in A$ mit

$$n_0 \leq n_1 < n_0 + 1/2$$

Angenommen, $n_0 \notin A$. Da zwischen n_0 und $n_0 + 1/2$ keine weitere natürliche Zahl außer n_1 liegt, ist $n_1 \leq n$ für alle $n \in A$. Also ist n_1 untere Schranke von A im Widerspruch zu $n_0 = \inf A$. Folglich ist $n_0 \in A$. □

Der folgende Satz zeigt insbesondere, dass $\mathbb{R}$ tatsächlich eine echte Obermenge von $\mathbb{Q}$ ist.

Satz 3.21 *Es sei $c \geq 0$, und es sei $n \in \mathbb{N}$. Dann existiert genau ein $x \in \mathbb{R}$ mit $x \geq 0$ und $x^n = c$.*

Beweis. Ist $c = 0$ oder $n = 1$, so ist die Behauptung klar. Wir nehmen daher $c > 0$ und $n \geq 2$ an.

1. Wir betrachten die Menge $M := \{t \in \mathbb{R} : t \geq 0 \ , \ t^n \leq c\}$. Wegen $0 \in M$ ist $M \neq \emptyset$. Aus $1 < 1 + c$ folgt $c < 1 + c < (1 + c)^2 < (1 + c)^3 < \cdots < (1 + c)^{n-1} < (1 + c)^n$. Also ist $1 + c$ eine obere Schranke von M.

2. Nach dem Vollständigkeitsaxiom (V) besitzt M eine kleinste obere Schranke $x := \sup M$. Es gilt dabei $x \geq 0$. Wir zeigen: $x^n = c$.
   Zunächst gilt für $0 \leq a < b$
   $$b^n - a^n < n(b - a)b^{n-1} \tag{$*$}$$
   (Denn: Nach der geometrischen Summenformel (beachte: $a/b \neq 1$) ist
   $$\begin{aligned} b^n - a^n &= b^n \left(1 - \left(\frac{a}{b}\right)^n\right) = \left(1 - \frac{a}{b}\right) \cdot b \cdot \sum_{\nu=0}^{n-1} a^\nu b^{n-\nu-1} \\ &= (b - a)\left(b^{n-1} + ab^{n-2} + \cdots + a^{n-1}\right) < n(b - a)b^{n-1} \ . \ ) \end{aligned}$$

   a) Angenommen, es ist $x^n > c$. Für $h := \frac{x^n - c}{nx^{n-1}}$ gilt dann
   $$0 < h = \frac{x}{n} - \frac{c}{nx^{n-1}} < \frac{x}{n} \leq x \ .$$
   Ist $t \geq x - h$, so folgt aus $(*)$ mit $b = x, \ a = x - h$
   $$x^n - t^n \leq x^n - (x - h)^n < n \cdot h \cdot x^{n-1} = x^n - c \ .$$
   Also ist $t^n > c$, d. h. $t \notin M$. Damit ist $x - h$ obere Schranke von M im Widerspruch dazu, dass x kleinste obere Schranke ist.

   b) Angenommen, es ist $x^n < c$. Dann ist
   $$0 < \frac{c - x^n}{n(x + 1)^{n-1}} \ .$$

Nach S 3.19 existiert ein $k \in \mathbb{N}$ mit $k > \frac{n(x+1)^{n-1}}{c-x^n}$, d. h.

$$0 < \frac{1}{k} < \frac{c - x^n}{n(x+1)^{n-1}} .$$

Aus $(*)$ mit $b = x + 1/k,\ a = x$ folgt

$$\left(x + \frac{1}{k}\right)^n - x^n < \frac{n}{k}\left(x + \frac{1}{k}\right)^{n-1} < \frac{n}{k}(x+1)^{n-1} < c - x^n ,$$

d. h. $(x + 1/k)^n < c$. Also ist $x + 1/k \in M$ und damit ist x keine obere Schranke von M. Widerspruch!

Aus a) und b) folgt $x^n = c$.

3. Es gibt nur ein $x \geq 0$ mit $x^n = c$, denn gilt $0 \leq x_1 < x_2$, so ist $x_1^n < x_2^n$.

□

Bemerkung und Definition 3.22 1. Die Lösung $x \geq 0$ der Gleichung $x^n = c$ wird als *n-te Wurzel* aus c bezeichnet. Man schreibt

$$x = \sqrt[n]{c} \qquad \text{oder} \quad x = c^{1/n} .$$

Im Fall $n = 2$ schreibt man kurz $x = \sqrt{c}$. Unter Ausnutzung der Potenzgesetze aus Abschnitt 2 ergibt sich hiermit für $n, m \in \mathbb{N}$

$$\sqrt[n]{\sqrt[m]{c}} = \sqrt[nm]{c} \ \left(= \sqrt[m]{\sqrt[n]{c}}\right)$$

und für $c, d \geq 0$

$$\sqrt[n]{cd} = \sqrt[n]{c}\,\sqrt[n]{d} .$$

2. Für $c > 0$ und $r \in \mathbb{Q}$, $r = p/q$ mit $p \in \mathbb{Z}, q \in \mathbb{N}$, setzen wir

$$c^r := c^{p/q} := \sqrt[q]{c^p} .$$

Es gilt dabei ([Ü]): Ist $r = p'/q'$ mit $p' \in \mathbb{Z}, q' \in \mathbb{N}$, so ist $\sqrt[q]{c^p} = \sqrt[q']{c^{p'}}$, d. h. die Definition ist unabhängig von der Darstellung von r. Außerdem ist $\sqrt[q]{c^p} = (\sqrt[q]{c})^p$.

Wir stellen noch einige Bezeichnungen zusammen: Für $a, b \in \mathbb{R}, a < b$ setzen wir

$$\begin{array}{lll} [a,b] & := & \{x \in \mathbb{R} : a \leq x \leq b\} \\ (a,b) & := & \{x \in \mathbb{R} : a < x < b\} \\ [a,b) & := & \{x \in \mathbb{R} : a \leq x < b\} \\ (a,b] & := & \{x \in \mathbb{R} : a < x \leq b\} \end{array} \qquad \begin{array}{lll} (-\infty,b] & := & \{x \in \mathbb{R} : x \leq b\} \\ (-\infty,b) & := & \{x \in \mathbb{R} : x < b\} \\ [a,\infty) & := & \{x \in \mathbb{R} : x \geq a\} \\ (a,\infty) & := & \{x \in \mathbb{R} : x > a\} \\ (\infty,\infty) & := & \mathbb{R}\,,\ (-\infty,0) = \mathbb{R}_-\,,\ (0,\infty) = \mathbb{R}_+ \end{array}$$

Diese Mengen heißen *Intervalle*.
Der folgende Satz zeigt, dass jedes Intervall rationale Zahlen enthält.

Satz 3.23 *Zu jedem $x \in \mathbb{R}$ und jedem $\varepsilon > 0$ existiert ein $r \in \mathbb{Q}$ mit*

$$x < r < x + \varepsilon\ .$$

Beweis. Es seien $x \in \mathbb{R}$ und $\varepsilon > 0$ gegeben. Wir wählen ein $q \in \mathbb{N}$ mit $1/q < \varepsilon$. Ist $p \in \mathbb{Z}$ so, dass $p - 1 \leq xq < p$, so folgt

$$\frac{p-1}{q} \leq x < \frac{p}{q}$$

und damit

$$x < \frac{p}{q} = \frac{p-1}{q} + \frac{1}{q} < x + \varepsilon\ .$$

□

4 Komplexe Zahlen

Wie wir in Abschnitt 3 gesehen haben, hat in $\mathbb{R}$ jede Gleichung $x^n = c$ für $n \in \mathbb{N}$ und $c \geq 0$ eine Lösung. Leider gilt dies nicht mehr im Falle $c < 0$ und n gerade (da $x^n \geq 0$ für gerades n und beliebiges $x \in \mathbb{R}$). Unser Ziel ist es nun, den Körper der reellen Zahlen so zu “erweitern”, dass $x^n = c$ auch für $c < 0$ (also etwa $x^2 = -1$) lösbar ist.

Bemerkung und Definition 4.1 Wir setzen

$$\mathbb{C} := \{(x, y) : x, y \in \mathbb{R}\} \qquad (= \mathbb{R} \times \mathbb{R})$$

und für $z_1 = (x_1, y_1), z_2 = (x_2, y_2) \in \mathbb{C}$

$$z_1 + z_2 = (x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2)$$

sowie

$$z_1 z_2 = (x_1, y_1) \cdot (x_2, y_2) := (x_1 x_2 - y_1 y_2 \ , \ x_1 y_2 + x_2 y_1) \ .$$

Man rechnet leicht nach, dass dann $(\mathbb{C}, +, \cdot)$ ein Körper ist. $\mathbb{C} = (\mathbb{C}, +, \cdot)$ heißt *Körper der komplexen Zahlen* und $z \in \mathbb{C}$ heißt *komplexe Zahl*.
Dabei ist die Null in $\mathbb{C}$ gegeben durch $0 = 0_{\mathbb{C}} = (0, 0)$ und die Eins in $\mathbb{C}$ ist gegeben durch $1 = 1_{\mathbb{C}} = (1, 0)$. Weiter sieht man: Ist $z = (x, y) \in \mathbb{C}$, so gilt

$$-z = (-x, -y) \qquad \text{und} \quad z^{-1} = \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right) \quad \text{für } z \neq 0 \ .$$

Wir nennen für $z = (x, y) \in \mathbb{C}$

$$\text{Re}\, z := x \qquad \textit{Realteil} \text{ von } z$$

und

$$\text{Im}\, z := y \qquad \textit{Imaginärteil} \text{ von } z \ .$$

Beispiel 4.2 Es sei $z_1 = (3, -1), \ z_2 = (2, 4)$.
Dann gilt $z_1 + z_2 = (5, 3), \ z_1 - z_2 = (3, -1) + (-2, -4) = (1, -5)$ und

$$z_1 \cdot z_2 = (3, -1) \cdot (2, 4) = (6 - (-4), 12 - 2) = (10, 10) \ .$$

Bemerkung 4.3 Wir können $\mathbb{R}$ als “Teilkörper” in $\mathbb{C}$ wiederfinden. (d. h. $\mathbb{R}$ ist eingebettet in $\mathbb{C}$): Es sei $\varphi : \mathbb{R} \to \mathbb{C}$ definiert durch

$$\varphi(x) := (x, 0) \qquad (x \in \mathbb{R}) \ .$$

Dann sieht man leicht, dass φ injektiv ist und die Körperstruktur erhält (also die Bedingungen aus D. 3.9 erfüllt). Damit können wir $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ auffassen. Den

reellen Zahlen entsprechen die komplexen Zahlen mit Imaginärteil $= 0$. *Wir schreiben dann auch kurz x statt* $(x, 0)$.
Man nennt weiterhin

$$i := (0, 1) \in \mathbb{C}$$

die *imaginäre Einheit* in $\mathbb{C}$. Für i gilt

$$i^2 = (0, 1) \cdot (0, 1) = (-1, 0) = -1\,.$$

Mit diesen Bezeichnungen können wir jedes $z = (x, y) \in \mathbb{C}$ in der Form

$$z = (x, y) = (x, 0) + (0, 1)(y, 0) = x + iy \ (= \ \operatorname{Re} z + i \operatorname{Im} z)$$

schreiben. Diese Darstellung heißt *Normaldarstellung* von z.
So gilt etwa

$$\begin{aligned} z_1 = (3, -1) &= 3 + i(-1) \quad (= 3 - i) \\ z_2 = (2, 4) &= 2 + i4 \qquad\; (= 2 + 4i) \end{aligned}$$

Bemerkung 4.4 In $(\mathbb{C}, +, \cdot)$ ist es nicht möglich, eine Ordnungsrelation $<$ (mit den Eigenschaften aus D. 3.1) zu definieren!
(Denn: Angenommen, doch. Dann gilt $1 > 0$ nach S. 3.3.3, also $-1 < 0$ nach S. 3.3.1. Für $z = i$ gilt mit S. 3.3.3

$$0 < i^2 = -1 < 0$$

also Widerspruch!)

Definition 4.5 Es sei $z = x + iy$ eine komplexe Zahl.

1. Die komplexe Zahl $\overline{z} := x - iy$ heißt zu z *konjugiert komplex.*

2. Die Zahl $|z| := \sqrt{x^2 + y^2} \in [0, \infty)$ heißt *Betrag* von z.

Geometrisch entsteht $\overline{z}$ durch Spiegelung von z an der reellen Achse. Der Betrag $|z|$ gibt anschaulich die Länge der Strecke von 0 zu z wieder (Pythagoras!)

Satz 4.6 *Für $z, z_1, z_2 \in \mathbb{C}$ gilt*

1. $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$,

2. $\overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2}$,

3. $\overline{(\overline{z})} = z$,

4. $|z|^2 = z\overline{z}$.

Beweis. [Ü] □

Für das Rechnen mit Beträgen gelten folgende Regeln

Satz 4.7 *Für* $z, z_1, z_2 \in \mathbb{C}$ *gilt*

1. $|z| \geq 0$ *und* $|z| = 0$ *genau dann, wenn* $z = 0$ *ist,*
2. $|z| = |\overline{z}|$, $\quad \operatorname{Re} z \leq |z|$, $\quad \operatorname{Im} z \leq |z|$,
3. $|z_1 z_2| = |z_1||z_2|$,
4. $|z_1 + z_2| \leq |z_1| + |z_2|$ $\qquad$ *("Dreiecksungleichung in* $\mathbb{C}$*").*

Beweis. 1. und 2. als [Ü].
3. Es gilt

$$|z_1 z_2|^2 = (z_1 z_2)(\overline{z_1 z_2}) = (z_1 \overline{z_1})(z_2 \overline{z_2}) = |z_1|^2 |z_2|^2 = (|z_1||z_2|)^2 .$$

Durch Wurzelziehen folgt die Behauptung.
4. Es gilt für $w \in \mathbb{C}$, $w = u + iv$

$$2 \operatorname{Re}(w) = (u + iv) + (u - iv) = w + \overline{w} .$$

Also folgt

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)(\overline{z_1} + \overline{z_2}) = z_1 \overline{z_1} + z_1 \overline{z_2} + z_2 \overline{z_1} + z_2 \overline{z_2} = \\ &= |z_1|^2 + 2 \operatorname{Re}(z_1 \overline{z_2}) + |z_2|^2 \overset{2.}{\leq} |z_1|^2 + 2|z_1 \overline{z_2}| + |z_2|^2 \\ &\overset{2./3.}{=} |z_1|^2 + 2|z_1||z_2| + |z_2|^2 = (|z_1| + |z_2|)^2 \end{aligned}$$

Durch Wurzelziehen folgt die Behauptung. □

Beispiel 4.8 Es gilt für $z_1 = (3, -1) = 3 - i, z_2 = (2, 4) = 2 + 4i$

$$\begin{aligned} |z_1| &= \sqrt{9 + 1} = \sqrt{10} , \qquad |z_2| = \sqrt{4 + 16} = \sqrt{20} \\ \overline{z_1} &= 3 - (-i) = 3 + i \\ z_1 \overline{z_1} &= (3 - i)(3 + i) = 9 - 3i + 3i - i^2 = 9 + 1 \left(= |z_1|^2\right) \end{aligned}$$

Definition 4.9 In Verallgemeinerung von D. 2.9 setzen wir noch für $z \in \mathbb{C}$ und $\nu \in \mathbb{N}$

$$\binom{z}{\nu} := \frac{z(z-1)\cdots(z-\nu+1)}{\nu!}$$

und

$$\binom{z}{0} := 1 .$$

Die Zahlen $\binom{z}{\nu} \in \mathbb{C}$ heißen ebenfalls *Binomialkoeffizienten.*

5 Folgen

Es sei $a : \mathbb{N} \to Y$, wobei Y eine beliebige Menge ist. Wir schreiben dann üblicherweise n als "Index", d. h.

$$a_n := a(n)$$

und statt a schreiben wir $(a_n)_{n\in\mathbb{N}}$ oder $(a_n)_{n=1}^{\infty}$ oder kurz (a_n). Mann nennt dann $(a_n)_{n\in\mathbb{N}}$ eine *Folge*. Allgemeiner betrachtet man auch Indexmengen $J \subset \mathbb{N}$ oder $J \subset \mathbb{Z}$ (mit ∞ vielen Elementen) und schreibt dann $(a_n)_{n\in J}$ statt $a : J \to Y$. Solche Objekte heißen auch *Folgen*.
Meist betrachten wir $Y = \mathbb{R}$ (Folgen reeller Zahlen) oder $Y = \mathbb{C}$ (Folgen komplexer Zahlen). Um die beiden Fälle $\mathbb{R}$ und $\mathbb{C}$ einheitlich bezeichnen zu können, schreiben wir $\mathbb{K} := K$, falls $K \in \{\mathbb{R}, \mathbb{C}\}$.
Folgen können entweder durch explizite Angabe von a_n für jedes $n \in \mathbb{N}$ (oder $n \in J$) gegeben sein (etwa $a_n = 1/n$ $(n \in \mathbb{N})$) oder induktiv oder rekursiv, d. h. durch eine sog. Rekursionsformel:

$$a_1 := a; \qquad a_{n+1} := \varphi(a_n)$$

wobei $\varphi : Y \to Y$ eine Funktion und $a \in Y$ ist (etwa: $\varphi(y) = \sqrt{y}$ mit $Y = [0, \infty)$ und $a = 2$, d. h. $a_1 := 2$, $a_{n+1} = \sqrt{a_n}$).

Definition 5.1 Eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ heißt

1. *beschränkt*, falls ein $M > 0$ existiert mit $|a_n| \le M$ $(n \in \mathbb{N})$. (Anderenfalls heißt (a_n) *unbeschränkt*.)

2. *konvergent*, falls ein $a \in \mathbb{K}$ so existiert, dass zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ existiert mit

   $$|a_n - a| < \varepsilon \qquad \text{für alle } n \ge N_\varepsilon.$$

   Die Zahl a heißt dann *Grenzwert* von (a_n) und wir schreiben

   $$a = \lim_{n\to\infty} a_n \qquad \text{oder kurz} \qquad a_n \to a \quad (n \to \infty)\ .$$

   (Eine Folge, die nicht konvergent ist, heißt *divergent*.)

Bemerkung 5.2 Man sieht leicht, dass jede Folge höchstens einen Grenzwert hat ([Ü]).

Beispiel 5.3 1. Die Folge (a_n) in $\mathbb{R}$ mit $a_n = n$ ist unbeschränkt.
2. Die Folge (a_n) in $\mathbb{R}$ mit $a_n = 1/n$ ist konvergent zum Grenzwert $a = 0$, d. h.

$$\frac{1}{n} \to 0 \quad (n \to \infty)\ .$$

(Denn: Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $N_\varepsilon \in \mathbb{N}$ mit $n > 1/\varepsilon$ für alle $n \geq N_\varepsilon$ (S. 3.19). Also gilt für alle $n \geq N_\varepsilon$

$$|a_n - 0| = \frac{1}{n} < \varepsilon\ .\ )$$

3. Die Folge (a_n) in $\mathbb{R}$ mit $a_n = (-1)^n$ ist beschränkt (da $|a_n| = 1 (n \in \mathbb{N})$ aber nicht konvergent. (Denn: Angenommen, $a_n \to a\,(n \to \infty)$.
Ist $a \geq 0$, so gilt für $\varepsilon = 1$

$$|a_{2k+1} - a| = |-1 - a| = a + 1 \geq 1 = \varepsilon$$

für alle $k \in \mathbb{N}$, d. h. es existiert kein $N_\varepsilon \in \mathbb{N}$ mit $|a_n - a| < \varepsilon$ für alle $n \geq N_\varepsilon$. Entsprechend argumentiert man im Fall $a < 0$.)

Satz 5.4 *Jede konvergente Folge (a_n) in $\mathbb{K}$ ist beschränkt.*

Beweis. Es sei $a = \lim\limits_{n\to\infty} a_n$. Dann existiert zu $\varepsilon = 1$ ein $N \in \mathbb{N}$ mit

$$|a_n - a| < 1 \qquad (n \geq N)\ ,$$

also $|a_n| = |a_n - a + a| \leq |a_n - a| + |a| < |a| + 1$ für alle $n \geq N$. Mit

$$M := \max\{|a_1|, \ldots, |a_{N-1}|, |a| + 1\}$$

gilt dann

$$|a_n| \leq M \qquad \text{für alle } n \in \mathbb{N}\,.$$

□

Der folgende Satz ist sehr "praktisch", da er es in vielen Fällen gestattet, Grenzwerte zu bestimmen ohne auf die "ε-N_ε-Definition" zurückzugreifen.

Satz 5.5 *Die Folgen (a_n) und (b_n) seien konvergent mit*

$$a = \lim_{n\to\infty} a_n\ , \qquad b = \lim_{n\to\infty} b_n\ .$$

Dann gilt

1. *Die Folge $(a_n + b_n)$ konvergiert gegen $a + b$, d. h.*

$$\lim_{n\to\infty} (a_n + b_n) = a + b \left(= \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n\right)\ .$$

2. *Die Folge $(a_n \cdot b_n)$ konvergiert gegen $a \cdot b$, d. h.*

$$\lim_{n\to\infty} (a_n \cdot b_n) = a \cdot b \left(= \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n\right)\ .$$

3. *Ist $b_n \neq 0$ für alle $n \in \mathbb{N}$* **und** *$b \neq 0$, so konvergiert (a_n/b_n) gegen a/b d. h.*

$$\lim_{n\to\infty}(a_n/b_n) = a/b \left(= \lim_{n\to\infty} a_n / \lim_{n\to\infty} b_n\right) .$$

Beweis.

1. Es sei $\varepsilon > 0$ gegeben. Dann existieren ein $N_\varepsilon^{(1)} \in \mathbb{N}$ und ein $N_\varepsilon^{(2)} \in \mathbb{N}$ mit

$$|a_n - a| < \varepsilon/2 \; (n \geq N_\varepsilon^{(1)}) \qquad \text{und} \qquad |b_n - b| < \varepsilon/2 \; (n \geq N_\varepsilon^{(2)}) .$$

Also gilt für $n \geq N_\varepsilon := \max(N_\varepsilon^{(1)}, N_\varepsilon^{(2)})$:

$$|a_n + b_n - (a+b)| = |a_n - a + b_n - b| \leq |a_n - a| + |b_n - b| < \varepsilon/2 + \varepsilon/2 = \varepsilon .$$

2. Nach S. 5.4 ist (b_n) beschränkt, d. h. es existiert ein $M > 0$ mit $|b_n| \leq M$ für alle $n \in \mathbb{N}$.
Es sei $\varepsilon > 0$ gegeben. Dann existieren ein $N_\varepsilon^{(1)} \in \mathbb{N}$ mit

$$|a_n - a| < \frac{\varepsilon}{2M} \qquad (n \geq N_\varepsilon^{(1)})$$

und ein $N_\varepsilon^{(2)} \in \mathbb{N}$ mit

$$|a||b_n - b| < \frac{\varepsilon}{2} \qquad (n \geq N_\varepsilon^{(2)}) .$$

Für alle $n \geq N_\varepsilon := \max(N_\varepsilon^{(1)}, N_\varepsilon^{(2)})$ gilt dann

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - a b_n + a b_n - ab| \leq \\ &\leq |b_n||a_n - a| + |a||b_n - b| \leq M|a_n - a| + |a||b_n - b| \\ &< M \cdot \frac{\varepsilon}{2M} + \frac{\varepsilon}{2} = \varepsilon . \end{aligned}$$

3. a) Wir zeigen zunächst:

$$\lim_{n\to\infty} \frac{1}{b_n} = \frac{1}{b} .$$

Da $b \neq 0$ ist, existiert ein $N \in \mathbb{N}$ mit

$$|b_n - b| < |b|/2 \qquad (n \geq N) .$$

Also gilt (umgekehrte Dreiecksungleichung!)

$$|b_n| = |b + b_n - b| \geq |b| - |b_n - b| > |b| - |b|/2 = |b|/2 > 0 \qquad (n \geq N) .$$

Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $N'_\varepsilon \in \mathbb{N}$ mit

$$|b_n - b| < \varepsilon \cdot \frac{b^2}{2} \qquad (n \geq N'_\varepsilon) .$$

Für alle $n \geq N_\varepsilon := \max(N, N'_\varepsilon)$ folgt

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b_n - b|}{|b_n b|} < \frac{2}{b^2}|b_n - b| < \varepsilon .$$

Also gilt $1/b_n \to 1/b$ $(n \to \infty)$.

b) Mit 2. und a) ergibt sich

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \left(a_n \cdot \frac{1}{b_n} \right) = a \cdot \frac{1}{b} \,.$$

□

Beispiel 5.6 Es sei $a_n = \frac{3n^2-4n}{2n^2+n}$. Da für Folgen (c_n) mit $c_n = c$ $(n \in \mathbb{N})$ offenbar $\lim\limits_{n\to\infty} c_n = c$ gilt, folgt mit S. 5.5 und B. 5.3.2

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{3-4/n}{2+1/n} = \frac{3-4\lim\limits_{n\to\infty} 1/n}{2+\lim\limits_{n\to\infty} 1/n} = \frac{3-4\cdot 0}{2+0} = \frac{3}{2} \,.$$

Wir betrachten nun speziell Folgen reeller Zahlen, wobei wir entscheidend von der Ordnungsstruktur in $\mathbb{R}$ Gebrauch machen. Als erstes beweisen wir ein nützliches Konvergenzkriterium.

Satz 5.7 *(Einschließungskriterium)*
Es seien $(a_n), (b_n), (c_n)$ Folgen in $\mathbb{R}$ mit $a_n \le b_n \le c_n$ Weiter seien (a_n) und (c_n) konvergent mit $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} c_n = a$. Dann ist auch (b_n) konvergent mit $\lim\limits_{n\to\infty} b_n = a$.

Beweis. Es sei $\varepsilon > 0$ gegeben. Dann existieren ein $N_\varepsilon^{(1)} \in \mathbb{N}$ mit

$$-\varepsilon < a_n - a < \varepsilon \qquad (n \ge N_\varepsilon^{(1)})$$

und ein $N_\varepsilon^{(2)} \in \mathbb{N}$ mit

$$-\varepsilon < c_n - a < \varepsilon \qquad (n \ge N_\varepsilon^{(2)})$$

Aus $a_n \le b_n \le c_n$ folgt damit für alle $n \ge N_\varepsilon := \max(N_\varepsilon^{(1)}, N_\varepsilon^{(2)})$

$$-\varepsilon < a_n - a \le b_n - a \le c_n - a < \varepsilon$$

d. h. $|b_n - a| < \varepsilon$. □

Als wichtiges Anwendungsbeispiel erhalten wir

Beispiel 5.8 Wir betrachten für $q \in \mathbb{R}$ die Folge $(q^n)_n$. Dann gilt
1. Für $-1 < q \le 1$ konvergiert (q^n), und es gilt

$$\lim_{n\to\infty} q^n = \begin{cases} 0, & \text{falls} \quad -1 < q < 1 \\ 1, & \text{falls} \quad q = 1 \end{cases}$$

2. Für alle anderen q divergiert (q^n).
(Denn: Für $q = 0$ und $q = 1$ ist die Behauptung klar. Für $q = -1$ liegt nach B. 5.3.3 Divergenz vor.
Es sei $0 < |q| < 1$. Dann ist mit einem $a > 0$

$$|q| = \frac{1}{1+a} \ .$$

Aus der Bernoullischen Ungleichung (S. 3.8) folgt $(1+a)^n \geq 1 + na > na$ und daher

$$|q^n| = |q|^n = \frac{1}{(1+a)^n} < \frac{1}{na} \qquad (n \in \mathbb{N})$$

also

$$-\frac{1}{na} < q^n < \frac{1}{na} \qquad (n \in \mathbb{N}) \ .$$

Nach dem Einschließungskriterium gilt $q^n \to 0$ $(n \to \infty)$.
Nun sei $|q| > 1$, d. h. $|q| = 1 + b$ mit einem $b > 0$. Mit S. 3.8 gilt

$$|q^n| = (1+b)^n \geq 1 + nb \qquad (n \in \mathbb{N}) \ ,$$

d. h. (q^n) ist unbeschränkt und damit nach S. 5.4 divergent.).

Bemerkung und Definition 5.9 Manchmal erweist es sich als nützlich, für reelle Folgen (a_n) auch erweiterte Grenzwerte $\pm\infty$ zu betrachten. Wir sagen deshalb

1. (a_n) heißt *bestimmt divergent* (oder auch *konvergent*) gegen ∞, falls zu jedem $R > 0$ ein $N_R \in \mathbb{N}$ existiert mit $a_n > R$ für alle $n \geq N_R$. Wir schreiben dann $a_n \to \infty$ $(n \to \infty)$

2. (a_n) heißt *bestimmt divergent* (oder auch *konvergent*) gegen $-\infty$, falls zu jedem $R > 0$ ein $N_R \in \mathbb{N}$ existiert mit $a_n < -R$ für alle $n \geq N_R$. Wir schreiben dann $a_n \to -\infty$ $(n \to \infty)$

So gilt etwa

$$q^n \to \infty \quad (n \to \infty)$$

im Falle $q > 1$, aber für $q < -1$ ist (q^n) weder bestimmt divergent gegen ∞ noch bestimmt divergent gegen $-\infty$.

Wir untersuchen jetzt eine wichtige Klasse von Folgen in $\mathbb{R}$, die monotonen Folgen.

Definition 5.10 Es sei $(a_n)_{n\in\mathbb{N}}$ eine reelle Folge. Dann heißt (a_n)

1. *monoton wachsend*, falls $a_{n+1} \geq a_n$ $(n \in \mathbb{N})$ (Schreibweise: $a_n \nearrow$),

2. *streng monoton wachsend*, falls $a_{n+1} > a_n$ $(n \in \mathbb{N})$ (Schreibweise: a_n streng $\nearrow$),

3. *monoton fallend*, falls $a_{n+1} \leq a_n$ $(n \in \mathbb{N})$ (Schreibweise: $a_n \searrow$),

4. *streng monoton fallend*, falls $a_{n+1} < a_n$ $(n \in \mathbb{N})$ (Schreibweise: a_n streng $\searrow$).

Beispiel 5.11 Es gilt

1. $\left(\frac{1}{n}\right)_n$ ist streng monoton fallend,

2. $(2^n)_n$ ist streng monoton wachsend,

3. $\Big((-1)^n\Big)_n$ ist weder monoton wachsend, noch monoton fallend,

4. $(1)_n$ ist monoton wachsend und fallend, aber nicht streng monoton wachsend oder fallend.

Eine fundamentale Folgerung aus der Vollständigkeit von $\mathbb{R}$ ist

Satz 5.12 *(Hauptsatz über monotone Folgen)*
Es sei (a_n) eine beschränkte Folge in $\mathbb{R}$. Ist (a_n) monoton wachsend oder monoton fallend, so ist (a_n) konvergent.

Beweis. 1. Es sei (a_n) monoton wachsend. Wir betrachten die Menge

$$A := \{x : x = a_n \text{ für ein } n \in \mathbb{N}\} \; .$$

Dann ist $A \neq \emptyset$ und (nach oben) beschränkt. Nach dem Vollständigkeitsaxiom existiert

$$a := \sup A \; .$$

Wir zeigen: $a = \lim\limits_{n\to\infty} a_n$.
Da a obere Schranke von A ist, gilt $a_n \leq a$ für alle $n \in \mathbb{N}$. Es sei $\varepsilon > 0$ gegeben. Dann ist (nach Definition von $\sup A$) $a - \varepsilon$ keine obere Schranke von A, d. h. es existiert ein $N \in \mathbb{N}$ mit $a_N > a - \varepsilon$. Da (a_n) monoton wachsend ist, gilt $a_n \geq a_N > a - \varepsilon$ für alle $n \geq N$, also insgesamt

$$a - \varepsilon < a_n \leq a \qquad (n \geq N)$$

d. h.

$$0 \leq a - a_n < \varepsilon \qquad (n \geq N) \; .$$

Da $\varepsilon > 0$ beliebig war, folgt $a_n \to a$ $(n \to \infty)$.
2. Ist (a_n) monoton fallend, so geht der Beweis entsprechend. □

Der Hauptsatz über monotone Folgen ist eines der wichtigsten Kriterien für die Konvergenz von Folgen. Man beachte, dass der Grenzwert bei der Konvergenzuntersuchung nicht eingeht (und dass auch keine Aussage über den Grenzwert gemacht wird).

Beispiel 5.13 1. Wir betrachten die Folge (a_n) in $\mathbb{R}$ mit

$$a_n = \left(1 + \frac{1}{n}\right)^n \qquad (n \in \mathbb{N}) .$$

Dann ist (a_n) konvergent.
(Denn: (a_n) ist monoton wachsend, da für alle $n \geq 2$ gilt

$$\begin{aligned} \frac{a_n}{a_{n-1}} &= \frac{\left(\frac{n+1}{n}\right)^n}{\left(\frac{n}{n-1}\right)^{n-1}} = \frac{n+1}{n} \cdot \left(\frac{(n+1)}{n} \cdot \frac{(n-1)}{n}\right)^{n-1} = \\ &= \frac{n+1}{n}\left(1 - \frac{1}{n^2}\right)^{n-1} \overset{S.3.8}{\geq} \frac{n+1}{n}\left(1 - \frac{n-1}{n^2}\right) = \frac{n^3+1}{n^3} > 1 . \end{aligned}$$

Außerdem ist (a_n) beschränkt, denn $a_n \leq \left(1 + \frac{1}{n}\right)^{n+1}$ und $\left(\left(1 + \frac{1}{n}\right)^{n+1}\right)_n$ ist monoton fallend ([Ü])).
Nach dem Hauptsatz über monotone Folgen ist (a_n) konvergent.
Wir setzen

$$e := \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$$

(e heißt *Eulersche Zahl*).

2. (Babylonisches Wurzelziehen)
Es sei $x > 0$ gegeben. Ein sehr effizientes Verfahren zur näherungsweisen Berechnung von $\sqrt{x}$ ist das sog. Babylonische Wurzelziehen:
Wir betrachten die Folgen (a_n) in $\mathbb{R}$ mit

$$a_{n+1} := \frac{1}{2}\left(a_n + \frac{x}{a_n}\right) \qquad \text{und } a_0 \geq \sqrt{x} .$$

(Der Startwert a_0 kann eine grobe Näherung an $\sqrt{x}$ sein.)

Behauptung: (a_n) ist konvergent mit $\lim_{n\to\infty} a_n = \sqrt{x}$.
Denn: Zunächst beachte man, dass (a_n) wohldefiniert ist (aus $a_n > 0$ folgt $a_{n+1} > 0$)
1. Wir zeigen: $a_n \geq \sqrt{x}$ $(n \in \mathbb{N}_0)$.
Für $n = 0$ gilt nach Voraussetzung $a_0 \geq \sqrt{x}$. Ist $a_n \geq \sqrt{x}$ für ein $n \in \mathbb{N}$, so folgt

$$\begin{aligned} a_{n+1} - \sqrt{x} &= \frac{1}{2}\left(a_n + \frac{x}{a_n}\right) - \sqrt{x} = \frac{1}{2a_n}(a_n^2 - 2a_n\sqrt{x} + x) = \\ &= \frac{1}{2a_n}(a_n - \sqrt{x})^2 \geq 0 , \end{aligned}$$

also ist auch $a_{n+1} \geq \sqrt{x}$.

2. Wir zeigen: (a_n) ist monoton fallend
Es gilt

$$a_n - a_{n+1} = a_n - \frac{1}{2}\left(a_n + \frac{x}{a_n}\right) = \frac{1}{2}\left(a_n - \frac{x}{a_n}\right) .$$

Aus $a_n \geq \sqrt{x}$ folgt $a_n^2 \geq x$, also $a_n \geq \frac{x}{a_n}$ und damit $a_n - a_{n+1} \geq 0$. Folglich ist (a_n) monoton fallend.
3. Nach dem Hauptsatz über monotone Folgen ist (a_n) konvergent. Zur Bestimmung des Grenzwertes kann man folgendermaßen vorgehen:
Aus $\lim\limits_{n\to\infty} a_n = a$ folgt $\lim\limits_{n\to\infty} a_{n+1} = a$, und damit (da $a \geq \sqrt{x} > 0$)

$$a \longleftarrow a_{n+1} = \frac{1}{2}\left(a_n + \frac{x}{a_n}\right) \longrightarrow \frac{1}{2}\left(a + \frac{x}{a}\right) \qquad (n \to \infty) .$$

Aus der Eindeutigkeit des Grenzwertes folgt

$$a = \frac{1}{2}\left(a + \frac{x}{a}\right) ,$$

woraus man $a = \sqrt{x}$ berechnet (beachte $a > 0$).

Man kann also die Folgeglieder a_n als Näherungen für $\sqrt{x}$ verwenden. Wie sieht es dabei mit dem Fehler aus, wenn man a_n statt $\sqrt{x}$ verwendet? Wir schätzen den Fehler nach oben ab.
Dazu sei

$$a_n = \sqrt{x}(1 + f_n)$$

($f_n = \frac{a_n - \sqrt{x}}{\sqrt{x}}$ heißt "relativer Fehler"). Dann gilt $f_n \geq 0$ und

$$1 + f_{n+1} = \frac{1}{2}\left(1 + f_n + \frac{1}{1 + f_n}\right) ,$$

also

$$f_{n+1} = \frac{1}{2}\frac{f_n^2}{1 + f_n} \leq \frac{1}{2}\min(f_n, f_n^2)(= \frac{1}{2}f_n^2\,,\ \text{falls } f_n < 1) .$$

Hat man nach n Schritten für a_n einen Fehler $f_n \leq 10^{-m}$, so ist der Fehler f_{n+1} dem nächsten Schritt $\leq \frac{1}{2}(10^{-m})^2 = \frac{1}{2}10^{-2m}$; die Anzahl der "exakten Stellen" verdoppelt sich im Wesentlichen.

Wir definieren nun eine der wichtigsten Funktionen für die Analysis, die Logarithmusfunktion. Vorbereitend zeigen wir

Satz 5.14 *Für $x > 0$ seien die Folgen (a_n) und (b_n) in $\mathbb{R}$ definiert durch*

$$\begin{array}{lcll} a_n := a_n(x) & := & 2^n\,(\sqrt[2^n]{x} - 1) & \\ b_n := b_n(x) & := & 2^n\left(1 - \frac{1}{\sqrt[2^n]{x}}\right) & \left(= \frac{1}{\sqrt[2^n]{x}}a_n\right) \end{array} \qquad (n \in \mathbb{N}_0) .$$

Dann gilt:

1. *Für alle* $n \in \mathbb{N}_0$ *ist* $a_{n+1} \leq a_n, b_n \leq b_{n+1}$ *und* $b_n \leq a_n$.

2. (a_n) *und* (b_n) *sind konvergent mit* $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$.

Beweis. 1. Zunächst gilt für beliebiges $y > 0$

$$2(y-1) \leq y^2 - 1 \qquad (*)$$

und

$$y + 1/y \geq 2 \qquad (**)$$

(Denn: Es ist $y^2 - 2y + 1 = (y-1)^2 \geq 0$. Also folgt

$$y^2 - 1 \geq 2y - 2$$

und

$$y^2 + 1 \geq 2y\ .)$$

Nach $(*)$ gilt (mit $x_n := \sqrt[2^n]{x}$)

$$\begin{aligned} a_{n+1} &= 2^{n+1}(x_{n+1}-1) = 2^n 2(x_{n+1}-1) \leq 2^n(x_{n+1}^2 - 1) = \\ &= 2^n(x_n - 1) = a_n \end{aligned}$$

und

$$\begin{aligned} b_{n+1} &= 2^{n+1}\left(1 - \frac{1}{x_{n+1}}\right) = -2^n 2\left(\frac{1}{x_{n+1}} - 1\right) \geq \\ &\geq -2^n\left(\frac{1}{x_{n+1}^2} - 1\right) = 2^n\left(1 - \frac{1}{x_n}\right) = b_n\ . \end{aligned}$$

Aus $(**)$ ergibt sich $x_n - 1 \geq 1 - 1/x_n$ und damit auch $a_n \geq b_n$ $(n \in \mathbb{N}_0)$.

2. Man sieht leicht [Ü], dass $\sqrt[m]{x} \to 1$ $(m \to \infty)$ gilt, also folgt insbesondere $x_n \to 1$ $(n \to \infty)$.
Aus dem Hauptsatz über monotone Folgen und 1. ergibt sich die Konvergenz von (a_n) und (b_n) (beachte: $a_n \geq b_0$ und $b_n \leq a_0$) und aus $x_n \to 1$ folgt $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$. $\square$

Definition 5.15 Für $x > 0$ setzen wir

$$\log x := \ln x := \lim_{n\to\infty} a_n \left(= \lim_{n\to\infty} b_n\right)\ .$$

($\log x$ heißt (*natürlicher*) *Logarithmus* von x).

Im folgenden Satz stellen wir einige Eigenschaften der Logarithmusfunktion $\log : (0, \infty) \to \mathbb{R}$ zusammen:

Satz 5.16 *1. Es sei $x > 0$. Dann gilt*

$$1 - \frac{1}{x} \le \log x \le x - 1$$

und

$$2\left(1 - \frac{1}{\sqrt{x}}\right) \le \log x \le 2\left(\sqrt{x} - 1\right) .$$

2. Für $x, y > 0$ gilt

$$\log(xy) = \log(x) + \log(y) \qquad \textit{und} \qquad \log(1/x) = -\log(x) .$$

3. Für $x, y > 0$ mit $x < y$ gilt

$$\log(x) < \log(y) .$$

Beweis. 1. Aus $b_n \le \lim\limits_{n\to\infty} b_n = \log x = \lim\limits_{n\to\infty} a_n \le a_n$ für alle $n \in \mathbb{N}_0$ sowie

$$a_0 = x - 1,\ b_0 = 1 - \frac{1}{x},\ a_1 = 2(\sqrt{x} - 1),\ b_1 = 2\left(1 - \frac{1}{\sqrt{x}}\right)$$

ergibt sich sofort 1.
2. Es gilt

$$\begin{aligned} \log(xy) &= \lim_{n\to\infty} 2^n \left(\sqrt[2^n]{xy} - 1\right) = \lim_{n\to\infty} 2^n \left(\sqrt[2^n]{x}\,\sqrt[2^n]{y} - 1\right) \\ &= \lim_{n\to\infty} \left\{2^n \cdot \sqrt[2^n]{x}(\sqrt[2^n]{y} - 1) + 2^n(\sqrt[2^n]{x} - 1)\right\} \\ &= \lim_{n\to\infty} \sqrt[2^n]{x} \cdot \lim_{n\to\infty} a_n(y) + \lim_{n\to\infty} a_n(x) = 1 \cdot \log(y) + \log(x) \end{aligned}$$

und damit auch

$$0 = \log 1 = \log(x/x) = \log(x) + \log(1/x) .$$

3. Aus $x/y < 1$ ergibt sich mit 1. und 2.

$$\log(y) - \log(x) = \log(y/x) \ge 1 - \frac{x}{y} > 0 .$$

□

Satz 5.17 *Ist $x > 0$ und (x_n) eine Folge mit $x_n \to x$, $x_n > 0$, so gilt*

$$\log(x_n) \to \log(x) \qquad (n \to \infty) .$$

Beweis. Aus $x_n \to x > 0$ folgt $x_n/x \to 1$ $(n \to \infty)$.
Also erhalten wir mit S. 5.16

$$0 \leftarrow 1 - \frac{x}{x_n} \le \log(x_n/x) \le \frac{x_n}{x} - 1 \to 0 \qquad (n \to \infty)$$

d.h.

$$\log(x_n) - \log(x) = \log(x_n/x) \to 0 \qquad (n \to \infty)$$

nach dem Einschließungskriterium. $\square$

Definition 5.18 Es sei (a_n) eine beschränkte Folge in $\mathbb{R}$. Dann heißt

$$\overline{a} := \lim_{n\to\infty} \sup\{a_k : k \geq n\}$$

Limes superior von (a_n) (Schreibweise: $\overline{a} =: \overline{\lim_{n\to\infty}} a_n$ oder auch $\limsup\limits_{n\to\infty} a_n$) und

$$\underline{a} := \lim_{n\to\infty} \inf\{a_k : k \geq n\}$$

Limes inferior von (a_n) (Schreibweise: $\underline{a} = \underline{\lim_{n\to\infty}}\, a_n$ oder auch $\liminf\limits_{n\to\infty} a_n$).

(Man beachte dabei: Da (a_n) beschränkt ist, ist auch $B_n := \{a_k : k \geq n\}$ beschränkt für alle $n \in \mathbb{N}$. Also existieren

$$\overline{b_n} := \sup B_n \quad \text{und} \quad \underline{b_n} := \inf B_n \qquad (n \in \mathbb{N})\ .$$

Nach Definition von sup und inf sind außerdem die Folgen $(\overline{b_n})$ bzw. $(\underline{b_n})$ monoton fallend bzw. monoton wachsend (und beschränkt), also konvergent nach dem Hauptsatz über monotone Folgen. Damit existieren $\overline{\lim} a_n$ und $\underline{\lim} a_n$ für jede beschränkte Folge.)

Beispiel 5.19 Es sei $a_n = (-1)^n + \frac{n}{n+1}$. Dann gilt

$$\overline{b_n} = \sup\{a_k : k \geq n\} = 2 \qquad (n \in \mathbb{N})$$

und

$$\underline{b_n} = \inf\{a_k : k \geq n\} = \begin{cases} -1 + \frac{n}{n+1} & , \text{ falls } \ n \text{ ungerade} \\ -1 + \frac{n+1}{n+2} & , \text{ falls } \ n \text{ gerade} \end{cases} .$$

Also gilt

$$\overline{b_n} \to 2\ (= \overline{\lim}\ a_n)$$

und

$$\underline{b_n} \to 0\ (= \underline{\lim}\ a_n)\ .$$

Satz 5.20 *Es sei (a_n) eine beschränkte Folge in $\mathbb{R}$, und es sei $a \in \mathbb{R}$. Dann gilt*

1. *Es ist $a = \overline{\lim} a_n$ genau dann, wenn folgende beiden Bedingungen gelten:*

   a) *Für alle $\varepsilon > 0$ existiert ein $N_\varepsilon \in \mathbb{N}$ mit $a_n < a + \varepsilon\, (n \geq N_\varepsilon)$.*

b) Für alle $\varepsilon > 0$ und alle $n \in \mathbb{N}$ existiert ein $k \geq n$ mit $a - \varepsilon < a_k$.

2. *Es ist $a = \underline{\lim} a_n$ genau dann, wenn folgende beiden Bedingungen gelten:*

   a) Für alle $\varepsilon > 0$ existiert ein $N_\varepsilon \in \mathbb{N}$ mit $a_n > a - \varepsilon \, (n \geq N_\varepsilon)$.

   b) Für alle $\varepsilon > 0$ und alle $n \in \mathbb{N}$ existiert ein $k \geq n$ mit $a + \varepsilon > a_k$.

Beweis. 1. "$\Longrightarrow$" Es sei $a = \overline{\lim} a_n = \lim\limits_{n\to\infty} \sup\{a_k : k \geq n\}$, und es sei $\varepsilon > 0$ gegeben. Dann existiert ein N_ε mit $a - \varepsilon < \sup\{a_k : k \geq n\} < a + \varepsilon$ $(n \geq N_\varepsilon)$. Aus der zweiten Ungleichung ergibt sich insbesondere $a_n < a + \varepsilon \, (n \geq N_\varepsilon)$. Aus der ersten Ungleichung folgt, dass für jedes $n (\geq N_\varepsilon)$ ein $k \geq n$ existiert mit $a - \varepsilon < a_k$. Damit gilt auch b).
"$\Longleftarrow$ " Es sei $\varepsilon > 0$ gegeben.
Aus a) folgt für $\overline{b_n} := \sup\{a_k : k \geq n\}$

$$\overline{b_{N_\varepsilon}} \leq a + \varepsilon \, .$$

Da $(\overline{b_n})$ monoton fallend ist, gilt auch $\overline{b_n} \leq a + \varepsilon$ $(n \geq N_\varepsilon)$. Also ist

$$\overline{\lim_{n\to\infty}} a_n = \lim_{n\to\infty} \overline{b_n} \leq a + \varepsilon \, .$$

Da $\varepsilon > 0$ beliebig war, ist $\overline{\lim\limits_{n\to\infty}} a_n \leq a$.
Aus b) folgt $\overline{b_n} > a - \varepsilon$ für alle $n \in \mathbb{N}$. Also folgt

$$\overline{\lim_{n\to\infty}} a_n = \lim_{n\to\infty} \overline{b_n} \geq a - \varepsilon \, .$$

Da $\varepsilon > 0$ beliebig war, ist $\overline{\lim\limits_{n\to\infty}} a_n \geq a$, also insgesamt $\overline{\lim\limits_{n\to\infty}} a_n = a$.

2. Der Beweis für $\underline{\lim} a_n$ verläuft analog. $\square$

Satz 5.21 *Es sei (a_n) eine beschränkte Folge in $\mathbb{R}$. Dann gilt*

1. $\underline{\lim\limits_{n\to\infty}} a_n \leq \overline{\lim\limits_{n\to\infty}} a_n$.

2. *(a_n) ist genau dann konvergent, wenn $\overline{\lim\limits_{n\to\infty}} a_n = \underline{\lim\limits_{n\to\infty}} a_n$ gilt (und in diesem Fall ist $\lim\limits_{n\to\infty} a_n = \overline{\lim\limits_{n\to\infty}} a_n = \underline{\lim\limits_{n\to\infty}} a_n$).*

Beweis. 1. Teil 1. ergibt sich sofort aus

$$\underline{b_n} = \inf\{a_k : k \geq n\} \leq \overline{b_n} = \sup\{a_k : k \geq n\}$$

und

$$\underline{\lim} a_n = \lim \underline{b_n} \, , \qquad \overline{\lim} a_n = \lim \overline{b_n} \, .$$

2. "$\Longrightarrow$" Es sei (a_n) konvergent, $\lim a_n =: a$, und es sei $\varepsilon > 0$ gegeben. Dann existiert ein $N_\varepsilon \in \mathbb{N}$ so, dass

$$a - \varepsilon < a_n < a + \varepsilon \qquad (n \geq N_\varepsilon) \ .$$

Also sind a) und b) aus S. 5.20.1 und 2. erfüllt und folglich gilt mit S. 5.20

$$\lim a_n = \overline{\lim} a_n = \underline{\lim} a_n .$$

"$\Longleftarrow$" Gilt $\overline{\lim} a_n = \underline{\lim} a_n =: a$, so folgt aus S. 5.20 dass für alle $\varepsilon > 0$ ein $N_\varepsilon^{(1)} \in \mathbb{N}$ mit

$$a_n < a + \varepsilon \qquad (n \geq N_\varepsilon^{(1)})$$

und ein $N_\varepsilon^{(2)} \in \mathbb{N}$ mit

$$a_n > a - \varepsilon \qquad (n \geq N_\varepsilon^{(2)})$$

existieren. Also gilt für $n \geq N_\varepsilon := \max(N_\varepsilon^{(1)}, N_\varepsilon^{(2)})$

$$a - \varepsilon < a_n < a + \varepsilon \ .$$

Da $\varepsilon > 0$ beliebig war, ist (a_n) konvergent mit $\lim a_n = a$. $\square$

Definition 5.22 Es sei (a_n) eine beliebige Folge. Ist (n_k) eine streng monoton wachsende Folge in $\mathbb{N}$, so heißt $(a_{n_k})_{k\in\mathbb{N}}$ eine *Teilfolge* von $(a_n)_{n\in\mathbb{N}}$.

Es gilt damit

Satz 5.23 *Es sei (a_n) eine beschränkte Folge in $\mathbb{R}$. Dann gilt:*

1. *Es existiert eine Teilfolge (a_{n_k}) mit $a_{n_k} \to \overline{\lim\limits_{n\to\infty}} a_n$.*
2. *Es existiert eine Teilfolge (a_{m_k}) mit $a_{m_k} \to \underline{\lim\limits_{n\to\infty}} a_n$.*
3. *Ist (a_{ℓ_k}) eine konvergente Teilfolge von (a_n) so ist*

$$\underline{\lim_{n\to\infty}} a_n \leq \lim_{n\to\infty} a_{\ell_k} \leq \overline{\lim_{n\to\infty}} a_n \ .$$

Beweis. 1. Wir definieren (n_k) induktiv. Dazu setzen wir $n_1 := 1$. Sind $n_1, \ldots, n_k$ bereits definiert, so wählen wir ein $n_{k+1} \in \mathbb{N}$ mit $n_{k+1} > n_k$ und so, dass

$$|a_{n_{k+1}} - \overline{\lim_{n\to\infty}} a_n| < 1/(k+1)$$

gilt (existiert nach S. 5.20). Dann gilt $a_{n_k} \to \overline{\lim} a_n \ (k \to \infty)$.
2. Analog

3. Es sei (a_{ℓ_k}) eine konvergente Teilfolge von (a_n) und $a := \lim\limits_{k\to\infty} a_{\ell_k}$. Ist $\varepsilon > 0$, so existiert ein N_ε mit $a_n < \overline{\lim} a_n + \varepsilon \, (n \geq N_\varepsilon)$ also auch $a_{\ell_k} \leq \overline{\lim} a_n + \varepsilon$ für $k \geq N_\varepsilon$ (beachte: $\ell_k \geq k$). Damit ist auch $a \leq \overline{\lim} a_n + \varepsilon$. Da $\varepsilon > 0$ beliebig war, gilt $a \leq \overline{\lim} a_n$. Entsprechend sieht man, dass $\underline{\lim} a_n \leq a$ gilt. □

Als Konsequenz erhalten wir insbesondere folgenden zentralen Satz

Satz 5.24 *(Bolzano-Weierstraß)*
Jede beschränkte Folge in $\mathbb{K}$ besitzt eine konvergente Teilfolge.

Beweis. 1. Es sei $\mathbb{K} = \mathbb{R}$. Dann wähle man etwa (n_k) wie in S. 5.23.1.
2. Es sei $\mathbb{K} = \mathbb{C}$, und es sei (a_n) eine beschränkte Folge in $\mathbb{C}$. Ist $a_n = \alpha_n + i\beta_n$ die Normaldarstellung von a_n, so sind die Folgen (α_n) und (β_n) in $\mathbb{R}$ beschränkt (es gilt $|\alpha_n| \leq |a_n|$ und $|\beta_n| \leq |a_n|$). Nach 1. existieren eine Teilfolge (α_{n_k}) von (α_n) und ein $\alpha \in \mathbb{R}$ mit $\alpha_{n_k} \to \alpha$ $(k \to \infty)$. Da auch $(\gamma_k) := (\beta_{n_k})$ beschränkt ist, existieren wieder nach 1. eine Teilfolge $(\beta_{n_{k_\ell}}) = (\gamma_{k_\ell})$ von (γ_k) und ein $\beta \in \mathbb{R}$ mit $\beta_{n_{k_\ell}} \to \beta \, (\ell \to \infty)$. Nach S. 5.5 gilt also

$$\alpha_{n_{k_\ell}} + i\beta_{n_{k_\ell}} \to \alpha + i\beta \qquad (\ell \to \infty) \, .$$

(Man beachte: Konvergiert eine Folge, so konvergiert auch jede Teilfolge, und zwar gegen den gleichen Grenzwert). □

Zum Abschluss dieses Abschnitts betrachten wir sog. Cauchy-Folgen

Definition 5.25 Es sei (a_n) eine Folge in $\mathbb{K}$. Dann heißt (a_n) *Cauchy-Folge*, falls zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in N$ existiert mit

$$|a_n - a_m| < \varepsilon \qquad \text{für alle} \quad n, m \geq N_\varepsilon \, .$$

Damit gilt folgendes wichtige Konvergenzkriterium (für $\mathbb{R}$ und $\mathbb{C}$).

Satz 5.26 *(Cauchy'sches Konvergenzkriterium)*
Es sei (a_n) eine Folge in $\mathbb{K}$. Genau dann ist (a_n) konvergent wenn (a_n) eine Cauchy-Folge ist.

Beweis. "$\Longrightarrow$" Es sei (a_n) konvergent mit $a := \lim a_n$. Dann existiert zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ mit $|a_n - a| < \varepsilon/2 \, (n \geq N_\varepsilon)$. Also gilt

$$|a_n - a_m| \leq |a_n - a| + |a - a_m| < \varepsilon \qquad \text{für alle } n, m \geq N_\varepsilon \, .$$

"$\Longleftarrow$" 1. Es sei (a_n) eine Cauchy-Folge in $\mathbb{K}$. Wir zeigen als erstes, dass (a_n) beschränkt ist.
(Denn: Es existiert nach Definition ein $N \in \mathbb{N}$ mit

$$|a_n - a_m| < 1 \qquad (n, m \geq N)$$

also insbesondere

$$|a_n - a_N| < 1 \qquad (n \geq N)\ .$$

Damit ist

$$|a_n| \leq |a_n - a_N| + |a_N| < |a_N| + 1 \qquad (n \geq N)\ .$$

Also gilt für $M := \max(|a_1|, \ldots, |a_{N-1}|, |a_N| + 1)$

$$|a_n| \leq M \qquad (n \in \mathbb{N})\ .)$$

2. Es sei zunächst (a_n) eine **reelle** Folge.
Nach S. 5.21.2 genügt es nun zu zeigen $\underline{\lim} a_n = \overline{\lim} a_n$. Angenommen, die ist nicht der Fall. Dann ist $\underline{\lim} a_n < \overline{\lim} a_n$ nach S. 5.21.1. Es sei $\varepsilon := (\overline{\lim} a_n - \underline{\lim} a_n)/3$.
Dann existiert ein $N_\varepsilon \in \mathbb{N}$ mit

$$|a_n - a_m| < \varepsilon \qquad (n, m \geq N_\varepsilon)\ . \qquad (*)$$

Außerdem existieren nach S. 5.20 ein $n_0 \geq N_\varepsilon$ mit

$$a_{n_0} > \overline{\lim} a_n - \varepsilon$$

und ein $m_0 \geq N_\varepsilon$ mit

$$a_{m_0} < \underline{\lim} a_n + \varepsilon\ .$$

Also folgt

$$a_{n_0} - a_{m_0} > \overline{\lim} a_n - \varepsilon - \underline{\lim} a_n - \varepsilon = 3\varepsilon - 2\varepsilon = \varepsilon$$

im Widerspruch zu $(*)$.
3. Nun sei (a_n) eine beliebige komplexe Folge, und es sei $a_n = \alpha_n + i\beta_n$ die Normaldarstellung von a_n. Dann sind (α_n) und (β_n) Folgen in $\mathbb{R}$ und es gilt

$$\left.\begin{array}{l} |\alpha_n - \alpha_m| \\ |\beta_n - \beta_m| \end{array}\right\} \leq |a_n - a_m| \qquad (n, m \in \mathbb{N})\ .$$

Also sind (α_n) und (β_n) Cauchy-Folgen in $\mathbb{R}$. Nach 2. existieren $\alpha, \beta \in \mathbb{R}$ mit $\alpha_n \to \alpha$ und $\beta_n \to \beta (n \to \infty)$.
Also folgt mit S. 5.5

$$\alpha_n + i\beta_n \to \alpha + i\beta \qquad (n \to \infty)\ .$$

$\square$

6 Reihen

Als Motivation betrachten wir die Folge $(q^\nu)_{\nu=0}^\infty$ für ein $q \in \mathbb{R}$ (oder auch $\mathbb{C}$) mit $|q| < 1$. Betrachten wir die Summe der ersten $(n+1)$ Folgelieder, so gilt nach der geometrischen Summenformel

$$\sum_{\nu=0}^{n} q^\nu = \frac{1-q^{n+1}}{1-q} \ .$$

Wegen $q^{n+1} \to 0 \quad (n \to \infty)$ folgt

$$\sum_{\nu=0}^{n} q^\nu \to \frac{1}{1-q} \qquad (n \to \infty) \ ,$$

die Folge der Summen konvergiert also gegen $\frac{1}{1-q}$. Man verwendet dann kurz das **Symbol** $\sum\limits_{\nu=0}^{\infty} q^\nu$ für den Grenzwert $\frac{1}{1-q}$.

Definition 6.1 Es sei $(a_\nu)_{\nu=0}^\infty$ eine Folge in $\mathbb{K}$.
1. Die der Folge (a_ν) zugeordnete Folge $(s_n)_{n=0}^\infty$ der *Partial-* oder *Teilsummen*

$$s_n := \sum_{\nu=0}^{n} a_\nu \qquad (n \in \mathbb{N}_0)$$

heißt (*die mit* (a_ν) *gebildete*) *Reihe* und wird mit $\sum\limits_{\nu=0}^{\infty} a_\nu$ bezeichnet. Die a_ν heißen dann *Reihenglieder*.
2. Die Reihe $\sum\limits_{\nu=0}^{\infty} a_\nu$ heißt *konvergent* (*zur Summe* s), falls $\lim\limits_{n\to\infty} s_n = s$ gilt. Wir schreiben dann $s =: \sum\limits_{\nu=0}^{\infty} a_\nu$.

!!! Man beachte, dass das Sympbol $\sum\limits_{\nu=0}^{\infty} a_\nu$ also zwei Bedeutungen hat: Erstens steht es für die Folge (s_n) der Teilsummen und zweitens (im Falle der Konvergenz!) für deren Grenzwert.
Dass die Summation bei 0 beginnt ist unwesentlich. Genauso kann man Reihen $\sum\limits_{\nu=n_0}^{\infty} a_\nu$ für ein $n_0 \in \mathbb{N}$ (oder auch $\mathbb{Z}$) betrachten.

Beispiel 6.2 1. Es sei $a_\nu = \frac{1}{\nu(\nu+1)}$ $(\nu \in \mathbb{N})$. Dann gilt

$$s_n = \sum_{\nu=1}^{n} \frac{1}{\nu(\nu+1)} = \sum_{\nu=1}^{n} \left(\frac{1}{\nu} - \frac{1}{\nu+1} \right) = \sum_{\nu=1}^{n} \frac{1}{\nu} - \sum_{\nu=2}^{n+1} \frac{1}{\nu} = 1 - \frac{1}{n+1} \ .$$

Also folgt $\lim\limits_{n\to\infty} s_n = 1$, d. h. $\sum\limits_{\nu=1}^{\infty} \frac{1}{\nu(\nu+1)}$ ist konvergent mit

$$\sum_{\nu=1}^{\infty} \frac{1}{\nu(\nu+1)} = 1 \, .$$

2. Es sei $a_\nu = q^\nu$ für ein $q \in \mathbb{K}$, $|q| < 1$. Dann ist (s. o.) $\sum\limits_{\nu=0}^{\infty} q^\nu$ konvergent mit

$$\sum_{\nu=0}^{\infty} q^\nu = \frac{1}{1-q} \, .$$

Diese Reihe heißt *geometrische Reihe*. Speziell ergibt sich etwa für $q = 1/2$

$$\sum_{\nu=0}^{\infty} \frac{1}{2^\nu} = 2$$

oder auch

$$\sum_{\nu=1}^{\infty} \frac{1}{2^\nu} = 1 \, .$$

Für das "Rechnen" mit konvergenten Reihen gilt

Satz 6.3 *Es seien* $\sum\limits_{\nu=0}^{\infty} a_\nu$ *und* $\sum\limits_{\nu=0}^{\infty} b_\nu$ *konvergente Reihen in* $\mathbb{K}$*, und es seien* $\alpha, \beta \in \mathbb{K}$*. Dann ist auch* $\sum\limits_{\nu=0}^{\infty} (\alpha a_\nu + \beta b_\nu)$ *konvergent mit*

$$\sum_{\nu=0}^{\infty} (\alpha a_\nu + \beta b_\nu) = \alpha \sum_{\nu=0}^{\infty} a_\nu + \beta \sum_{\nu=0}^{\infty} b_\nu \, .$$

Beweis. Ergibt sich leicht durch Anwendung von S. 5.5. □

Damit gilt etwa

$$\begin{aligned} \sum_{\nu=0}^{\infty} \frac{2 \cdot 3^\nu + 4}{5^\nu} &= 2 \cdot \sum_{\nu=0}^{\infty} \frac{3^\nu}{5^\nu} + 4 \cdot \sum_{\nu=0}^{\infty} \frac{1}{5^\nu} \\ &= 2 \cdot \frac{1}{1 - 3/5} + 4 \cdot \frac{1}{1 - 1/5} = 10 \, . \end{aligned}$$

Durch Übertragung des Cauchy'schen Konvergenzkriteriums für Folgen ergibt sich unmittelbar

Satz 6.4 *Es sei* $(a_\nu)_{\nu=0}^{\infty}$ *eine Folge in* $\mathbb{K}$*.*

1. *(Cauchy-Kriterium für Reihen)*
   Genau dann ist die Reihe $\sum\limits_{\nu=0}^{\infty} a_\nu$ *konvergent, wenn zu jedem* $\varepsilon > 0$ *ein* $N_\varepsilon \in \mathbb{N}$ *existiert mit*
$$\left|\sum_{\nu=m+1}^{n} a_\nu\right| < \varepsilon \qquad (n > m \geq N_\varepsilon)\ .$$

2. *Ist* $\sum\limits_{\nu=0}^{\infty} a_\nu$ *konvergent, so gilt stets*
$$a_n \longrightarrow 0 \qquad (n \to \infty)\ ,$$
   d. h. eine notwendig für die Konvergenz einer Reihe ist die Bedingung, dass die Reihenglieder eine Nullfolge bilden.

Beweis. 1. Ist $s_n := \sum\limits_{\nu=0}^{n} a_\nu$, so ist (s_n) (und damit $\sum\limits_{\nu=0}^{\infty} a_\nu$) nach S. 5.26 genau dann konvergent, wenn zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ existiert mit
$$|s_n - s_m| < \varepsilon \qquad (n, m \geq N_\varepsilon)$$
Dabei kann ohne Einschränkung $n > m$ vorausgesetzt werden. Mit
$$s_n - s_m = \sum_{\nu=0}^{n} a_\nu - \sum_{\nu=0}^{m} a_\nu = \sum_{\nu=m+1}^{n} a_\nu$$
ergibt sich die Behauptung.
2. Ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ konvergent, so folgt aus 1. insbesondere: Zu jedem $\varepsilon > 0$ existiert ein $N_\varepsilon \in \mathbb{N}$ mit
$$|a_n| = \left|\sum_{\nu=n}^{n} a_\nu\right| < \varepsilon \qquad (n > N_\varepsilon)$$
(wähle $m = n - 1$). Also gilt $a_n \to 0 \quad (n \to \infty)$. □

Beispiel 6.5 1. Es sei $q \in \mathbb{C}$ mit $|q| \geq 1$. Dann konvergiert nach B. 5.8 die Folge q^n nicht gegen 0. Also ist $\sum\limits_{\nu=0}^{\infty} q^\nu$ nach S. 6.4 divergiert. Insbesondere divergiert also etwa $\sum\limits_{\nu=0}^{\infty} 1$ und $\sum\limits_{\nu=0}^{\infty} (-1)^\nu$ oder auch $\sum\limits_{\nu=0}^{\infty} i^\nu$.
2. (harmonische Reihe) Die Reihe $\sum\limits_{\nu=0}^{\infty} \frac{1}{\nu}$ ist divergent.
[Dies zeigt insbesondere, dass die Bedingung “$a_n \to 0$” nicht hinreichend für die Konvergenz von $\sum\limits_{\nu=0}^{\infty} a_\nu$ ist!]

(Denn: Wäre $\sum\limits_{\nu=1}^{\infty} \frac{1}{\nu}$ konvergent, so würde nach S. 6.4 zu $\varepsilon = 1/2$ ein $N \in \mathbb{N}$ existieren mit

$$\sum_{\nu=n+1}^{2n} \frac{1}{\nu} < 1/2 \quad (n \geq N) .$$

Andererseits gilt aber für alle $n \in \mathbb{N}$

$$\sum_{\nu=n+1}^{2n} \frac{1}{\nu} \geq n \cdot \frac{1}{2n} = \frac{1}{2} .$$

Widerspruch!)

Für Reihen mit nichtnegativen Gliedern a_ν gilt folgendes sehr einfache Kriterium

Satz 6.6 *Es sei* $\sum\limits_{\nu=0}^{\infty} a_\nu$ *eine Reihe mit* $a_\nu \geq 0$ *für alle* $\nu \in \mathbb{N}_0$*. Dann ist* $\sum\limits_{\nu=0}^{\infty} a_\nu$ *genau dann konvergent, wenn die Folge* (s_n) *der Teilsummen beschränkt ist.*

Beweis. Für die Teilsummen $s_n = \sum\limits_{\nu=0}^{n} a_\nu$ gilt

$$s_n - s_{n-1} = a_n \geq 0 \qquad (n \in \mathbb{N}) .$$

Also ist (s_n) monoton wachsend.

Ist (s_n) beschränkt, so ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ konvergent nach dem Hauptsatz über monotone Folgen. Ist andererseits (s_n) unbeschränkt, so ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ divergent nach S. 5.4. □

Beispiel 6.7 Es sei $p \in \mathbb{N}, p > 1$. Dann ist $\sum\limits_{\nu=1}^{\infty} \frac{1}{\nu^p}$ konvergent.

Denn: Es gilt für $n \in \mathbb{N}$ (vgl. B. 6.2)

$$s_n = \sum_{\nu=1}^{n} \frac{1}{\nu^p} \leq \sum_{\nu=1}^{n} \frac{1}{\nu^2} \leq 1 + \sum_{\nu=2}^{n} \frac{1}{\nu(\nu-1)} = 1 + \sum_{\nu=1}^{n-1} \frac{1}{\nu(\nu+1)} = 2 - \frac{1}{n} < 2 .$$

Also ist die Teilsummenfolge (s_n) beschränkt. Nach S. 6.6 ist $\sum\limits_{\nu=1}^{\infty} \frac{1}{\nu^p}$ konvergent (und es gilt $\sum\limits_{\nu=1}^{\infty} \frac{1}{\nu^p} \leq 2$).

Viel schwieriger ist die Frage nach dem Reihenwert $\sum\limits_{\nu=1}^{\infty} \frac{1}{\nu^p}$ zu beantworten. Man kann zeigen (etwa mit Methoden der "Fourier-Analysis")

$$\sum_{\nu=1}^{\infty} \frac{1}{\nu^2} = \frac{\pi^2}{6} , \qquad \sum_{\nu=1}^{\infty} \frac{1}{\nu^4} = \frac{\pi^4}{90} .$$

Satz 6.8 *(Cauchy'scher Verdichtungssatz)*
Es sei (a_n) eine monoton fallende Folge mit $a_n \geq 0$. Dann konvergiert $\sum\limits_{\nu=1}^{\infty} a_\nu$ genau dann, wenn die Reihe $\sum\limits_{k=0}^{\infty} 2^k a_{2^k}$ konvergiert.

Beweis. Wir setzen

$$s_n := \sum_{\nu=1}^{n} a_\nu \qquad \text{und} \quad \sigma_n := \sum_{k=0}^{n} 2^k a_{2^k} \; .$$

Dann sind (s_n) und (σ_n) monoton wachsend.
1. Es sei $\sum\limits_{k=0}^{\infty} 2^k a_{2^k}$ konvergent. Dann ist die Folge (σ_n) beschränkt. Da (a_ν) monoton fallend ist, ergibt sich für $n \in \mathbb{N}$

$$\begin{aligned} s_n = \sum_{\nu=1}^{n} a_\nu \;\; &\leq \;\; \sum_{\nu=1}^{2^{n+1}-1} a_\nu \\ &= \;\; a_1 + (a_2 + a_3) + (a_4 + \cdots + a_7) + \cdots + (a_{2^n} + \cdots + a_{2^{n+1}-1}) \\ &\leq \;\; 2^0 a_1 + 2^1 \cdot a_2 + 2^2 \cdot a_4 + \cdots + 2^n a_{2^n} = \sigma_n \; . \end{aligned}$$

Also ist auch (s_n) beschränkt.
2. Es sei $\sum\limits_{\nu=1}^{\infty} a_\nu$ konvergent. Dann ist (s_n) beschränkt. Da (a_ν) monoton fallend ist, ergibt sich für $n \in \mathbb{N}$

$$\begin{aligned} \sigma_n \;\; &= \;\; \sum_{k=0}^{n} 2^k a_{2^k} = 2^0 a_1 + 2^1 a_2 + 2^2 a_4 + 2^3 a_8 + \cdots + 2^n a_{2^n} \\ &= \;\; 2\left\{\frac{1}{2} a_1 + 2^0 a_2 + 2^1 a_4 + 2^2 a_8 + \cdots + 2^{n-1} a_{2^n}\right\} \\ &\leq \;\; 2\left\{a_1 + a_2 + (a_3 + a_4) + (a_5 + \cdots + a_8) + \cdots + (a_{2^{n-1}+1} + \cdots + a_{2^n})\right\} \\ &= \;\; 2\sum_{\nu=1}^{2^n} a_\nu = 2 s_{2^n} \; . \end{aligned}$$

Also ist auch (σ_n) beschränkt und folglich $\sum\limits_{k=0}^{\infty} 2^k a_{2^k}$ konvergent. □

Beispiel 6.9 Es sei $\alpha \in \mathbb{Q}$. Dann ist $\sum\limits_{\nu=1}^{\infty} \frac{1}{\nu^\alpha}$ genau dann konvergent, wenn $\alpha > 1$ ist.
(Denn: Ist $\alpha \leq 0$ so ist $\left(\frac{1}{n^\alpha}\right)$ keine Nullfolge, also ist $\sum\limits_{\nu=1}^{\infty} \frac{1}{\nu^\alpha}$ divergent nach S. 6.4.2. Es sei also $\alpha > 0$. Dann ist $a_n = 1/n^\alpha$ eine monoton fallende Nullfolge. Nach S. 6.8

konvergiert $\sum\limits_{\nu=1}^{\infty} a_\nu$ genau dann, wenn $\sum\limits_{k=0}^{\infty} 2^k a_{2^k}$ konvergiert.
Es gilt aber

$$\sum_{k=0}^{\infty} 2^k a_{2^k} = \sum_{k=0}^{\infty} 2^k/(2^k)^\alpha = \sum_{k=0}^{\infty} (2^k)^{1-\alpha} = \sum_{k=0}^{\infty} (2^{1-\alpha})^k$$

und diese geometrische Reihe ist nach B. 6.2.2 und B. 6.5.1 genau dann konvergent, wenn $2^{1-\alpha} < 1$, also $\alpha > 1$ ist.)

Ein weiteres klassisches Konvergenzkriterium gilt für so genannte "alternierende Reihen":

Satz 6.10 *(Leibniz'sches Konvergenzkriterium für alternierende Reihen)*
Es sei (a_n) eine monotone fallende Folge (nichtnegativer) Zahlen a_n mit $\lim\limits_{n\to\infty} a_n = 0$. Dann konvergiert die Reihe $\sum\limits_{\nu=0}^{\infty} (-1)^\nu a_\nu$.

Beweis. Wir betrachten die Teilsummenfolge $(s_{2n})_n$. Es gilt

$$s_{2n} = \sum_{\nu=0}^{2n} (-1)^\nu a_\nu = \underbrace{(a_0 - a_1)}_{\geq 0} + \underbrace{(a_2 - a_3)}_{\geq 0} + \cdots + \underbrace{(a_{2n-2} - a_{2n-1})}_{\geq 0} + \underbrace{a_{2n}}_{\geq 0} \geq 0\ .$$

Außerdem gilt für $k \in \mathbb{N}_0$

$$\begin{aligned} s_{k+2} - s_k &= \sum_{\nu=0}^{k+2} (-1)^\nu a_\nu - \sum_{\nu=0}^{k} (-1)^\nu a_\nu = \\ &= (-1)^{k+2} a_{k+2} + (-1)^{k+1} a_{k+1} = (-1)^k \underbrace{(a_{k+2} - a_{k+1})}_{\leq 0}\ , \end{aligned}$$

d. h. $(s_{2n})_{n=0}^{\infty}$ ist monoton fallend und $(s_{2n+1})_{n=0}^{\infty}$ ist monoton wachsend.
Insbesondere ist nach dem Hauptsatz über monotone Folgen (s_{2n}) konvergent. Ist $s := \lim\limits_{n\to\infty} s_{2n}$, so gilt auch

$$s_{2n+1} = s_{2n} - \underbrace{a_{2n+1}}_{\to 0} \to s \quad (n \to \infty)\ .$$

Nun sei $\varepsilon > 0$ gegeben. Dann existieren ein $N_\varepsilon^{(1)} \in \mathbb{N}$ mit

$$|s_{2k} - s| < \varepsilon \quad (k \geq N_\varepsilon^{(1)})$$

und ein $N_\varepsilon^{(2)} \in \mathbb{N}$ mit

$$|s_{2k+1} - s| < \varepsilon \quad (k \geq N_\varepsilon^{(2)})\ .$$

Also gilt für $n \geq N_\varepsilon := \max\left(2N_\varepsilon^{(1)}, 2N_\varepsilon^{(2)} + 1\right)$

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

□

Beispiel 6.11 (alternierende harmonische Reihe)
Die Reihe $\sum\limits_{\nu=1}^{\infty}(-1)^\nu/\nu$ konvergiert nach S. 6.10 (denn: $a_n \searrow 0\ (n \to \infty)$).
Während also $\sum\limits_{\nu=1}^{\infty}\frac{1}{\nu}$ (nach B. 6.5.2) divergiert, führt das "Anbringen" abwechselnder Vorzeichen zur Konvergenz. Wir untersuchen nun Reihen, bei denen dieser Unterschied nicht auftritt.

Definition 6.12 Es sei $\sum\limits_{\nu=0}^{\infty} a_\nu$ eine Reihe in $\mathbb{K}$. Dann heißt $\sum\limits_{\nu=0}^{\infty} a_\nu$ *absolut konvergent*, falls $\sum\limits_{\nu=0}^{\infty} |a_\nu|$ konvergent ist. Ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ konvergent und $\sum\limits_{\nu=0}^{\infty} |a_\nu|$ divergent, so heißt $\sum\limits_{\nu=0}^{\infty} a_\nu$ *bedingt konvergent.*

Beispiel 6.13 1. Wie oben gesehen, ist $\sum\limits_{\nu=1}^{\infty}\frac{(-1)^\nu}{\nu}$ bedingt konvergent.
2. Die Reihe $\sum\limits_{\nu=1}^{\infty}\frac{(-1)^\nu}{\nu^p}$ ist für jedes feste $p \in \mathbb{N}$ mit $p > 1$ absolut konvergent (vgl. B. 6.7).

Einfach zu beweisen ist

Satz 6.14 *Ist* $\sum\limits_{\nu=0}^{\infty} a_\nu$ *absolut konvergent, so ist* $\sum\limits_{\nu=0}^{\infty} a_\nu$ *auch konvergent.*

Beweis. Es sei $\varepsilon > 0$. Dann existiert ein $N_\varepsilon \in \mathbb{N}$ mit

$$\sum_{\nu=m+1}^{n} |a_\nu| < \varepsilon \quad (n > m \geq N_\varepsilon)\ .$$

(S. 6.4.1). Also gilt auch

$$\left|\sum_{\nu=m+1}^{n} a_\nu\right| \leq \sum_{\nu=m+1}^{n} |a_\nu| < \varepsilon \quad (n > m \geq N_\varepsilon)\ .$$

Nach S. 6.4.1 ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ konvergent. □

Eines der wichtigsten Kriterien für (absolute) Konvergenz ist

Satz 6.15 *(Majorantenkriterium)*
Es sei (a_ν) eine Folge in $\mathbb{K}$ und es sei (b_ν) eine Folge mit $|a_\nu| \le b_\nu$ $(\nu \ge N)$ für ein $N \in \mathbb{N}$. Ist $\sum\limits_{\nu=0}^{\infty} b_\nu$ konvergent, so ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ absolut konvergent. (Die Reihe $\sum\limits_{\nu=0}^{\infty} b_\nu$ heißt "Majorante" von $\sum\limits_{\nu=0}^{\infty} a_\nu$.)

Beweis. Nach Voraussetzung ist insbesondere $b_\nu \ge 0$ für alle $\nu \ge N$. Ohne Eischränkung können wir $b_\nu \ge 0$ für alle ν voraussetzen. Da $\sum\limits_{\nu=0}^{\infty} b_\nu$ konvergent ist, existiert ein $M > 0$ mit

$$\sum_{\nu=0}^{n} b_\nu \le M \qquad (n \in \mathbb{N}) .$$

Also gilt für alle $n \ge N$

$$\sum_{\nu=0}^{n} |a_\nu| = \sum_{\nu=0}^{N-1} |a_\nu| + \sum_{\nu=N}^{n} |a_\nu| \le \sum_{\nu=0}^{N-1} |a_\nu| + M .$$

Damit ist $\left(\sum\limits_{\nu=0}^{n} |a_\nu|\right)_n$ beschränkt, d. h. $\sum\limits_{\nu=0}^{\infty} a_\nu$ ist absolut konvergent nach S. 6.6. □

Beispiel 6.16 1. Wir betrachten die Reihe $\sum\limits_{\nu=0}^{\infty} a_\nu$ mit $a_\nu = (-1)^\nu \cdot \frac{\nu^2-3\nu}{4\nu^4+5}$.
Dann gilt (für $\nu > 1$)

$$\left|(-1)^\nu \frac{\nu^2 - 3\nu}{4\nu^4 + 5}\right| \nu^2 = \frac{1 - 3/\nu}{4 + 5/\nu^4} \to \frac{1}{4} \quad (\nu \to \infty) .$$

Also existiert ein $N \in \mathbb{N}$ mit

$$|a_\nu| \le \frac{1}{\nu^2} \qquad (\nu \ge N) .$$

Folglich ist die Reihe nach S. 6.15 und B. 6.7 konvergent .
2. Die Reihe $\sum\limits_{\nu=1}^{\infty} b_\nu$ mit $b_\nu = \frac{\nu^2-3\nu}{4\nu^3+5}$ ist divergent. Denn: Angenommen, die Reihe wäre konvergent. Da

$$\frac{\nu^2 - 3\nu}{4\nu^3 + 5} \cdot \nu \to \frac{1}{4} \quad (\nu \to \infty)$$

gilt, existiert ein $N \in \mathbb{N}$ mit

$$b_\nu \ge \frac{1}{5\nu} \qquad (\nu \ge N) .$$

Dann wäre auch $\sum\limits_{\nu=1}^{\infty} \frac{1}{5\nu}$ konvergent nach S. 6.15. Dies ist aber nicht der Fall. (Man nennt $\sum\limits_{\nu=1}^{\infty} \frac{1}{5\nu}$ eine "divergente Minorante" von $\sum\limits_{\nu=1}^{\infty} b_\nu$.)

Durch Anwendung von S. 6.15 auf die konvergente Majorante $\sum\limits_{\nu=0}^{\infty} q^\nu$ für $|q| < 1$ ergibt sich:

Satz 6.17 *(Wurzelkriterium)*
Es sei (a_ν) eine Folge in $\mathbb{K}$ und es sei

$$a := \overline{\lim_{\nu\to\infty}} \sqrt[\nu]{|a_\nu|} \; .$$

1. Ist $a < 1$, so ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ absolut konvergent.

2. Ist $a > 1$, so ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ divergent.

Beweis. 1. Ist $a < 1$, so existieren ein $N \in \mathbb{N}$ und ein $q < 1$ mit $\sqrt[\nu]{|a_\nu|} \leq q$ für alle $\nu \geq N$. Da $\sum\limits_{\nu=0}^{\infty} q^\nu$ konvergiert, ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ absolut konvergent nach S. 6.15.
2. Ist $a > 1$, so ist $\sqrt[\nu]{|a_\nu|} > 1$ (und damit auch $|a_\nu| > 1$) für ∞ viele $\nu \in \mathbb{N}_0$. Also ist (a_ν) keine Nullfolge. Nach S. 6.4 ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ divergent. □

Beispiel 6.18 1. Es sei $a_\nu = \nu^p q^\nu$ für ein festes $p \in \mathbb{N}$ und ein festes $q \in \mathbb{C}$ mit $|q| < 1$. Dann gilt, da $\sqrt[\nu]{\nu} \to 1$ für $\nu \to \infty$ ([Ü]),

$$\sqrt[\nu]{\nu^p |q|^\nu} = \left(\sqrt[\nu]{\nu}\right)^p |q| \to |q| \qquad (\nu \to \infty) \; .$$

Also gilt

$$\overline{\lim_{\nu\to\infty}} \sqrt[\nu]{|a_\nu|} = \lim_{\nu\to\infty} \sqrt[\nu]{|a_\nu|} = |q| < 1 \; .$$

Nach S. 6.17 ist $\sum\limits_{\nu=0}^{\infty} \nu^p q^\nu$ konvergent. (Insbesondere ergibt sich mit S. 6.4 auch $\nu^p q^\nu \to 0$ $(\nu \to \infty)$.)
2. Es sei $a_\nu = \frac{1}{\nu^p}$ für ein festes $p \in \mathbb{N}$. Dann gilt

$$\sqrt[\nu]{|a_\nu|} = \left(\frac{1}{\sqrt[\nu]{\nu}}\right)^p \to 1 \quad (\nu \to \infty) \; ,$$

d. h. es ist $a = 1$ in S. 6.17. Wie oben gesehen ist $\sum\limits_{\nu=1}^{\infty} 1/\nu$ divergent und $\sum\limits_{\nu=1}^{\infty} \frac{1}{\nu^p}$ für $p > 1$ (absolut) konvergent. Also lässt sich im Fall $a = 1$ i. a. keine Aussage machen.

Satz 6.19 *(Quotientenkriterium)*
Es sei (a_ν) eine Folge in $\mathbb{K}$ mit $a_\nu \neq 0$ für $\nu \geq \nu_0$. Wir setzen

$$\underline{a} := \underline{\lim_{\nu\to\infty}} \left|\frac{a_{\nu+1}}{a_\nu}\right| \; , \qquad \overline{a} := \overline{\lim_{\nu\to\infty}} \left|\frac{a_{\nu+1}}{a_\nu}\right| \; .$$

Dann gilt

1. *Ist* $\overline{a} < 1$, *so ist* $\sum\limits_{\nu=0}^{\infty} a_\nu$ *absolut konvergent.*

2. *Ist* $\underline{a} > 1$, *so ist* $\sum\limits_{\nu=0}^{\infty} a_\nu$ *divergent.*

Beweis. 1. Es sei $\overline{a} < 1$. Dann existiert zu jedem $q \in (\overline{a}, 1)$ ein $N = N_q \in \mathbb{N}$ mit

$$\left|\frac{a_{\nu+1}}{a_\nu}\right| \leq q \qquad (\nu \geq N) .$$

Für $n > N$ gilt dann

$$|a_n| = \left|\frac{a_n}{a_{n-1}}\right| \cdot \left|\frac{a_{n-1}}{a_{n-2}}\right| \cdots \left|\frac{a_{N+1}}{a_N}\right| |a_N| \leq q^{n-N}|a_N| = \frac{|a_N|}{q^N} \cdot q^n$$

also

$$\sqrt[n]{|a_n|} \leq \left(\frac{|a_N|}{q^N}\right)^{1/n} q \longrightarrow q \qquad (n \to \infty)$$

und damit

$$\overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|} \leq q .$$

Nach dem Wurzelkriterium (S. 6.17) ist $\sum a_\nu$ absolut konvergent.
2. Es sei $\underline{a} > 1$. Dann existiert ein $N \in \mathbb{N}$, mit

$$\left|\frac{a_{\nu+1}}{a_\nu}\right| \geq 1 \qquad (\nu \geq N) ,$$

d. h. $(|a_\nu|)_{\nu \geq N}$ ist monoton wachsend und damit keine Nullfolge. Also ist $\sum a_\nu$ sicher divergent. □

Bemerkung 6.20 Der Beweis zum Quotientenkriterium zeigt, dass (unter den Voraussetzungen des Quotientenkriteriums)

$$a = \overline{\lim_{\nu\to\infty}} \sqrt[\nu]{|a_\nu|} \leq \overline{a} = \overline{\lim_{\nu\to\infty}} \left|\frac{a_{\nu+1}}{a_\nu}\right|$$

gilt. Also liefert auch das Wurzelkriterium stets die Konvergenz, wenn diese durch das Quotientenkriterium nachgewiesen werden kann. Allerdings kann es u. U. wesentlich einfacher sein, $\overline{a} < 1$ als $a < 1$ nachzurechnen.

Beispiel 6.21 1. Wir betrachten $a_\nu := \frac{z^\nu}{\nu!}$ $(\nu \in \mathbb{N}_0)$, wobei $z \in \mathbb{C}$ fest ist. Dann gilt (für $z \neq 0$)

$$\left|\frac{a_{\nu+1}}{a_\nu}\right| = \frac{|z|^{\nu+1}}{(\nu+1)!} \frac{\nu!}{|z|^\nu} = \frac{|z|}{\nu+1} \to 0 \quad (\nu \to \infty)$$

also ist $\overline{a} = 0 < 1$. Damit liefert das Quotientenkriterium die absolute Konvergenz von $\sum\limits_{\nu=0}^{\infty} \frac{z^\nu}{\nu!}$ für alle $z \in \mathbb{C} \setminus \{0\}$ (und für $z = 0$ ist die Reihe trivialerweise konvergent).

2. Es sei $0 < q < 1$ und

$$a_\nu := \begin{cases} q^\nu & , \text{ falls } \nu \text{ gerade} \\ q^{\nu+2} & , \text{ falls } \nu \text{ ungerade} \end{cases}$$

Dann gilt

$$\left|\frac{a_{2k+2}}{a_{2k+1}}\right| = \frac{q^{2k+2}}{q^{2k+3}} = \frac{1}{q} \longrightarrow \frac{1}{q} \qquad (k \to \infty)$$

und

$$\left|\frac{a_{2k+1}}{a_{2k}}\right| = \frac{q^{2k+3}}{q^{2k}} = q^3 \longrightarrow q^3 \qquad (k \to \infty)$$

d. h.

$$\overline{a} = \frac{1}{q} > 1\,, \qquad \underline{a} = q^3 < 1\,.$$

Damit ist nach dem Quotientenkriterium keine Aussage möglich. Es gilt jedoch

$$\sqrt[\nu]{|a_\nu|} \le q < 1 \qquad (\nu \in \mathbb{N}_0)$$

d. h. $\sum\limits_{\nu=0}^{\infty} a_\nu$ ist nach dem Wurzelkriterium absolut konvergent. Das Beispiel und B. 6.20 zeigen, dass das Wurzelkriterium echt stärker ist als das Quotientenkriterium. Außerdem kann $\underline{a}$ in S. 6.19.2 nicht durch $\overline{a}$ ersetzt werden.

In S. 6.3 hatten wir gesehen, dass gewisse Rechenregeln, die wir von (endlichen) Summen kennen, auch für Reihen gelten. Wir werden nun sehen, dass bei Reihen i. a. die "Reihenfolge" der Glieder nicht abgeändert werden kann ohne die Konvergenz zu beeinflussen. Dazu müssen wir zunächst präzisieren, was wir unter "Änderung der Reihenfolge" verstehen.

Definition 6.22 Es sei $\sum\limits_{\nu=0}^{\infty} a_\nu$ eine Reihe in $\mathbb{K}$. Ist $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$ bijektiv, so heißt $\sum\limits_{\nu=0}^{\infty} a_{\varphi(\nu)} = \sum\limits_{\nu=0}^{\infty} a'_\nu$ eine *Umordnung* von $\sum\limits_{\nu=0}^{\infty} a_\nu$.

Beispiel 6.23 Wir betrachten die Reihe

$$\sum_{\nu=0}^{\infty} a_\nu = \sum_{\nu=0}^{\infty} (-1)^\nu \frac{1}{\nu+1} \left(= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} \pm \ldots \right)$$

die nach dem Leibniz-Kriterium konvergiert (vgl. B. 6.11). Es sei $s := \sum\limits_{\nu=0}^{\infty} a_\nu$. Hier ist

$$\sum_{\nu=0}^{\infty} a'_\nu = 1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \cdots + \frac{1}{2k-1} - \frac{1}{4k-2} - \frac{1}{4k} + \cdots$$

eine Umordnung von $\sum_{\nu=0}^{\infty} a_\nu$. (Es gilt $a'_\nu = a_{\varphi(\nu)}$ mit

$$\begin{array}{lcll} \varphi(3k) & = & 2k & \\ \varphi(3k+1) & = & 4k+1 & \qquad (k \in \mathbb{N}_0)\ .) \\ \varphi(3k+2) & = & 4k+3 & \end{array}$$

Ist $t_n = \sum_{\nu=0}^{n} a'_\nu$ und $s_n = \sum_{\nu=0}^{n} a_\nu$, so gilt für alle $k \in \mathbb{N}$

$$\begin{aligned} 2t_{3k-1} &= \underbrace{2-1}_{=1} - \frac{1}{2} + \underbrace{\frac{2}{3} - \frac{1}{3}}_{=1/3} - \frac{1}{4} + \cdots + \underbrace{\frac{2}{2k-1} - \frac{1}{2k-1}}_{=1/(2k-1)} - \frac{1}{2k} \\ &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2k-1} - \frac{1}{2k} = s_{2k-1}\ . \end{aligned}$$

Damit folgt

$$t_{3k-1} \to \frac{1}{2}s \qquad (k \to \infty)\ .$$

Aus $a'_\nu \to 0$ $(\nu \to \infty)$ ergibt sich auch

$$t_{3k} = t_{3k-1} + a'_{3k} \to \frac{1}{2}s \quad \text{und} \quad t_{3k+1} = t_{3k} + a'_{3k+1} \to \frac{1}{2}s$$

und deshalb ist

$$\sum_{\nu=0}^{\infty} a'_\nu = \lim_{n\to\infty} t_n = \frac{1}{2}s = \frac{1}{2}\sum_{\nu=0}^{\infty} a_\nu\ .$$

Da $s \neq 0$ ist (beachte: nach dem Beweis zu S. 6.10 ist $s \geq s_1 = 1 - 1/2 = 1/2$) ist $s/2 \neq s$, d. h. die Reihenwerte sind verschieden.

Für absolut konvergente Reihen gilt allerdings

Satz 6.24 *Die Reihe $\sum_{\nu=0}^{\infty} a_\nu$ in $\mathbb{K}$ sei absolut konvergent, und es sei $s := \sum_{\nu=0}^{\infty} a_\nu$. Ist $\sum_{\nu=0}^{\infty} a'_\nu$ eine beliebige Umordnung von $\sum_{\nu=0}^{\infty} a_\nu$, so ist auch $\sum_{\nu=0}^{\infty} a'_\nu$ absolut konvergent und es gilt $s = \sum_{\nu=0}^{\infty} a'_\nu$.*

Beweis. Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $N = N_\varepsilon \in \mathbb{N}$ mit

$$0 \leq \sum_{\nu=N+1}^{\infty} |a_\nu| = \sum_{\nu=0}^{\infty} |a_\nu| - \sum_{\nu=0}^{N} |a_\nu| < \varepsilon\ .$$

Wir wählen ein $M \in \mathbb{N}$ so groß, dass in der Teilsumme $\sum_{\nu=0}^{M} a'_\nu$ alle $a_0, \ldots, a_N$ als Summanden vorkommen (d. h. ist $a'_\nu = a_{\varphi(\nu)}$ so wählen wir M so groß, dass $\{0, \ldots, N\} \subset$

$\{\varphi(0), \ldots, \varphi(M)\}$ gilt).
Dann gilt für alle $n > m \geq M$

$$\sum_{\nu=m+1}^{n} |a'_\nu| = \sum_{\nu=m+1}^{n} |a_{\varphi(\nu)}| \leq \sum_{\nu=N+1}^{\infty} |a_\nu| < \varepsilon \ .$$

Nach dem Cauchy'schen Konvergenzkriterium ist $\sum\limits_{\nu=0}^{\infty} |a'_\nu|$ konvergent.
Weiter gilt für alle $n \geq M$

$$\begin{aligned}\left|\sum_{\nu=0}^{n} a'_\nu - s\right| &= \left|\sum_{\nu=0}^{n} a'_\nu - \sum_{\nu=0}^{\infty} a_\nu\right| = \left|\sum_{\nu=0}^{n} a'_\nu - \sum_{\nu=0}^{N} a_\nu - \sum_{\nu=N+1}^{\infty} a_\nu\right| \\ &\leq \left|\sum_{\nu=0}^{n} a'_\nu - \sum_{\nu=0}^{N} a_\nu\right| + \sum_{\nu=N+1}^{\infty} |a_\nu| \\ &\leq \sum_{\nu=N+1}^{\infty} |a_\nu| + \sum_{\nu=N+1}^{\infty} |a_\nu| < 2\varepsilon\end{aligned}$$

Also ist $s = \sum\limits_{\nu=0}^{\infty} a'_\nu$. □

Wir wollen uns jetzt noch (kurz) mit der Multiplikation von Reihen beschäftigen. Dazu orientieren wir uns zunächst an der Multiplikation von Summen: Es sei

$$A := \sum_{\nu=0}^{n} a_\nu \ , \qquad B = \sum_{\mu=0}^{m} b_\mu \ .$$

Dann ist

$$AB = \left(\sum_{\nu=0}^{n} a_\nu\right)\left(\sum_{\mu=0}^{m} b_\mu\right) = \sum_{\substack{0 \leq \nu \leq n \\ 0 \leq \mu \leq m}} a_\nu b_\mu$$

wobei die Reihenfolge der Summation beliebig ist. Entsprechend sollte beim Produkt $\left(\sum\limits_{\nu=0}^{\infty} a_\nu\right)\left(\sum\limits_{\mu=0}^{\infty} b_\mu\right)$ jeder der Faktoren $a_\nu b_\mu$ $(\nu \in \mathbb{N}_0, \mu \in \mathbb{N}_0)$ einmal auftauchen. Anders als bei endlichen Summen spielt dabei jedoch die "Reihenfolge" der Summation

i. a. eine Rolle. Eine mögliche Anordnung ist die folgende

$$\begin{array}{ccccccccc}
a_0b_0 & & a_0b_1 & & a_0b_2 & & a_0b_3 & & a_0b_4 \\
& \overset{+}{\swarrow} & & \overset{+}{\swarrow} & & \overset{+}{\swarrow} & & \overset{+}{\swarrow} & \\
a_1b_0 & & a_1b_1 & & a_1b_2 & & a_1b_3 & & \dots \\
& \overset{+}{\swarrow} & & \overset{+}{\swarrow} & & \overset{+}{\swarrow} & & & \\
a_2b_0 & & a_2b_1 & & a_2b_2 & & \dots & & \dots \\
& \overset{+}{\swarrow} & & \overset{+}{\swarrow} & & & & & \\
a_3b_0 & & a_3b_1 & & \dots & & \dots & & \dots \\
& \overset{+}{\swarrow} & & & & & & & \\
a_4b_0 & & \dots & & \dots & & \dots & & \dots
\end{array}$$

d. h. wir betrachten die Reihe $\sum\limits_{n=0}^{\infty} c_n$ mit

$$c_n = \sum_{\nu=0}^{n} a_\nu b_{n-\nu} \qquad (n \in \mathbb{N}_0) \ .$$

Die c_n stellen also gerade die Summe der Produkte in der n-ten Diagonale dar.

Definition 6.25 Es seien $\sum\limits_{\nu=0}^{\infty} a_\nu$ und $\sum\limits_{\nu=0}^{\infty} b_\nu$ Reihen in $\mathbb{K}$. Dann heißt die Reihe

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} \left(\sum_{\nu=0}^{n} a_\nu b_{n-\nu} \right) \left[= \sum_{n=0}^{\infty} \left(\sum_{\nu=0}^{n} a_{n-\nu} b_\nu \right) \right]$$

Cauchy'sche Produktreihe oder kurz *Cauchy-Produkt* von $\sum\limits_{\nu=0}^{\infty} a_\nu$ und $\sum\limits_{\nu=0}^{\infty} b_\nu$.

Es gilt damit

Satz 6.26 *Es seien* $\sum\limits_{\nu=0}^{\infty} a_\nu$ *und* $\sum\limits_{\nu=0}^{\infty} b_\nu$ *konvergente Reihen in* $\mathbb{K}$. *Ist eine der beiden Reihen absolut konvergent, so konvergiert die Cauchy'sche Produktreihe* $\sum\limits_{n=0}^{\infty} c_n$ *und es gilt*

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{\nu=0}^{\infty} a_\nu \right) \left(\sum_{\nu=0}^{\infty} b_\nu \right) \ .$$

Beweis. O. E. sei $\sum\limits_{\nu=0}^{\infty} |a_\nu|$ konvergent. Es sei $B_n := \sum\limits_{\nu=0}^{n} b_\nu$ und $A := \sum\limits_{\nu=0}^{\infty} a_\nu$, $B := \sum\limits_{\nu=0}^{\infty} b_\nu$. Dann gilt

$$\begin{aligned}\sum_{\nu=0}^{n} c_\nu &= a_0b_0 + (a_1b_0 + a_0b_1) + (a_2b_0 + a_1b_1 + a_0b_2) + \ldots + (a_nb_0 + \ldots + a_0b_n) \\ &= a_0B_n + a_1B_{n-1} + \ldots + a_nB_0 \\ &= a_0(B_n - B) + a_1(B_{n-1} - B) + \ldots + a_n(B_0 - B) + B \cdot \sum_{\nu=0}^{n} a_\nu \, .\end{aligned}$$

Wegen $B \cdot \sum\limits_{\nu=0}^{n} a_\nu \to A \cdot B \ (n \to \infty)$ reicht es also zu zeigen, dass gilt

$$\sum_{\nu=0}^{n} a_{n-\nu}(B_\nu - B) \to 0 \qquad (n \to \infty) \, .$$

Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $N = N_\varepsilon \in \mathbb{N}$ mit $|B_\nu - B| < \varepsilon \, (\nu \geq N)$. Für $n > N$ gilt dann

$$\begin{aligned}\left|\sum_{\nu=0}^{n} a_{n-\nu}(B_\nu - B)\right| &\leq \sum_{\nu=0}^{N} |a_{n-\nu}||B_\nu - B| + \varepsilon \sum_{\nu=N+1}^{n} |a_{n-\nu}| \\ &\leq \max_{0 \leq \nu \leq N} |B_\nu - B| \cdot \sum_{\nu=0}^{N} |a_{n-\nu}| + \varepsilon \cdot \sum_{\nu=0}^{\infty} |a_\nu| \, .\end{aligned}$$

Aus S. 6.4 folgt $a_{n-\nu} \to 0 \ (n \to \infty)$ für alle festen $\nu = 0, \ldots N$, also auch $\sum\limits_{\nu=0}^{N} |a_{n-\nu}| \to 0 \ (n \to \infty)$. Damit existiert ein $N'_\varepsilon \geq N$ so, dass

$$\max_{0 \leq \nu \leq N} |B_\nu - B| \cdot \sum_{\nu=0}^{N} |a_{n-\nu}| < \varepsilon \qquad (n \geq N'_\varepsilon) \, .$$

Also gilt insgesamt

$$\left|\sum_{\nu=0}^{n} a_{n-\nu}(B_\nu - B)\right| \leq \varepsilon \left(1 + \sum_{\nu=0}^{\infty} |a_\nu|\right) \qquad (n \geq N'_\varepsilon) \, .$$

Da $\varepsilon > 0$ beliebig war folgt $\sum\limits_{\nu=0}^{n} a_{n-\nu}(B_\nu - B) \to 0 \qquad (n \to \infty)$. □

Beispiel 6.27 Für $z \in \mathbb{C}, |z| < 1$ betrachten wir die (absolut) konvergenten geometrischen Reihen

$$\sum_{\nu=0}^{\infty} a_\nu = \sum_{\nu=0}^{\infty} b_\nu := \sum_{\nu=0}^{\infty} z^\nu = \frac{1}{1-z} \, .$$

Dann ist die Cauchy'sche Produktreihe gegeben durch $\sum\limits_{n=0}^{\infty} c_n$ mit

$$c_n = \sum_{\nu=0}^{n} a_\nu b_{n-\nu} = \sum_{\nu=0}^{n} z^\nu z^{n-\nu} = z^n \sum_{\nu=0}^{n} 1 = (n+1) z^n \ .$$

Nach S. 6.26 gilt

$$\sum_{n=0}^{\infty} (n+1) z^n = \sum_{n=0}^{\infty} c_n = \frac{1}{(1-z)^2} \ .$$

Beispiel 6.28 Wir betrachen $a_\nu = b_\nu = (-1)^\nu / \sqrt{\nu+1}$. Dann ist $\sum\limits_{\nu=0}^{\infty} a_\nu = \sum\limits_{\nu=0}^{\infty} b_\nu$ konvergent nach dem Leibnizkriterium. Für die Cauchy'sche Produktreihe $\sum\limits_{n=0}^{\infty} c_n$ gilt

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

also

$$|c_n| \geq \sum_{\nu=0}^{n} \frac{1}{\sqrt{n+1}\sqrt{n+1}} = 1 \ .$$

Also ist $\sum\limits_{n=0}^{\infty} c_n$ divergent nach S. 6.4.2. Dieses Beispiel zeigt insbesondere, dass auf die Voraussetzung der absoluten Konvergenz einer der beiden Reihen in S. 6.26 nicht verzichtet werden kann!

7 Die Exponentialfunktion (elementare Funktionen I)

Für $z \in \mathbb{C}$ betrachten wir die Reihe $\sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!}$, die nach B. 6.21.1 absolut konvergent ist. Mittels dieser Reihe können wir eine weitere elementare Funktion definieren.

Definition 7.1 Die Funktion $\exp : \mathbb{K} \to \mathbb{K}$, gegeben durch

$$\exp(z) := \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} \qquad (z \in \mathbb{K}) ,$$

heißt *Exponentialfunktion*.

Eine der zentralen Eigenschaften der Exponentialfunktion stellt die folgende Funktionalgleichung dar, die zeigt, dass exp aus der Addition eine Multiplikation macht.

Satz 7.2 *Es seien $z, w \in \mathbb{C}$. Dann gilt*

$$\exp(z + w) = \exp(z) \cdot \exp(w) .$$

Beweis. Die Reihen $\sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!}$ und $\sum_{\nu=0}^{\infty} \frac{w^\nu}{\nu!}$ konvergieren absolut nach B. 6.21.1. Also konvergiert nach S. 6.26 das Cauchy-Produkt der beiden Reihen, und es gilt

$$\begin{aligned} \exp(z) \cdot \exp(w) &= \left(\sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} \right) \left(\sum_{\nu=0}^{\infty} \frac{w^\nu}{\nu!} \right) = \sum_{n=0}^{\infty} \sum_{\nu=0}^{n} \frac{z^\nu}{\nu!} \frac{w^{n-\nu}}{(n-\nu)!} = \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\nu=0}^{n} \binom{n}{\nu} z^\nu w^{n-\nu} \overset{S. 2.12}{=} \sum_{n=0}^{\infty} \frac{1}{n!} (z+w)^n = \exp(z+w) . \end{aligned}$$

□

Als Folgerung erhalten wir

Satz 7.3 *Für alle $z \in \mathbb{C}$ gilt $\exp(-z) = \frac{1}{\exp(z)}$ (Insbesondere ist $e^z \neq 0$).*

Beweis. Nach D. 7.1 gilt $\exp(0) = 1$. Also folgt aus S. 7.2

$$1 = \exp(z - z) = \exp(z) \cdot \exp(-z)$$

d. h. $\exp(-z) = 1/\exp(z)$. □

Satz 7.4 *Für alle $z \in \mathbb{C}$ gilt*

$$\lim_{n\to\infty}\left(1+\frac{z}{n}\right)^n = \exp(z)\left(=\sum_{\nu=0}^{\infty}\frac{z^\nu}{\nu!}\right) .$$

Beweis. Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $k \in \mathbb{N}$ mit

$$\sum_{\nu=k+1}^{\infty}\frac{|z|^\nu}{\nu!} < \varepsilon/3 .$$

Es folgt (mit der binomischen Formel) für $n > k$

$$\begin{aligned}\left|\left(1+\frac{z}{n}\right)^n - \sum_{\nu=0}^{\infty}\frac{z^\nu}{\nu!}\right| &\le \sum_{\nu=0}^{k}\left|\binom{n}{\nu}\frac{z^\nu}{n^\nu}-\frac{z^\nu}{\nu!}\right| + \\ &+ \sum_{\nu=k+1}^{n}\binom{n}{\nu}\frac{|z|^\nu}{n^\nu} + \sum_{\nu=k+1}^{\infty}\frac{|z|^\nu}{\nu!} .\end{aligned}$$

Zur Abschätzung der mittleren Summe verwenden wir

$$\binom{n}{\nu}\frac{1}{n^\nu} = \frac{1}{\nu!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{\nu-1}{n}\right) \le \frac{1}{\nu!}$$

und damit

$$\sum_{\nu=k+1}^{n}\binom{n}{\nu}\frac{|z|^\nu}{n^\nu} \le \sum_{\nu=k+1}^{n}\frac{|z|^\nu}{\nu!} < \varepsilon/3 .$$

Wegen

$$\binom{n}{\nu}\frac{1}{n^\nu} \to \frac{1}{\nu!} \qquad (n\to\infty)$$

für festes $\nu \in \mathbb{N}_0$ existiert ein $N > k$ mit

$$\sum_{\nu=0}^{k}\left|\binom{n}{\nu}\frac{z^\nu}{n^\nu}-\frac{z^\nu}{\nu!}\right| = \sum_{\nu=0}^{k}|z|^\nu\left(\frac{1}{\nu!}-\binom{n}{\nu}\frac{1}{n^\nu}\right) < \varepsilon/3$$

für $n \ge N$. Für $n \ge N$ gilt dann

$$\left|\left(1+\frac{z}{n}\right)^n - \sum_{\nu=0}^{\infty}\frac{z^\nu}{\nu!}\right| < \varepsilon .$$

□

Bemerkung 7.5 Aus S. 7.4 ergibt sich insbesondere

$$\exp(1) = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e(=e^1) .$$

Hieraus ergibt sich durch Anwendung von S. 7.2

$$\exp(p/q) = e^{p/q}$$

für alle $p \in \mathbb{Z}, q \in \mathbb{N}$. Deshalb schreiben wir in Zukunft auch e^z statt $\exp(z)$ für allgemeines $z \in \mathbb{C}$.

Satz 7.6 *Die Funktionen* $\exp : \mathbb{R} \to (0, \infty)$ *und* $\ln : (0, \infty) \to \mathbb{R}$ *sind bijektiv und es gilt*

$$\exp = \ln^{-1} \ .$$

Für $x > 0, y \in \mathbb{R}$ *ist also*

$$y = \ln(x) \Longleftrightarrow e^y = x \ .$$

Beweis. Es sei $x \in \mathbb{R}$. Nach S. 5.16.1 gilt für n so groß, dass $1 + \frac{x}{n} > 0$ ist,

$$n\left(1 - \frac{1}{1 + x/n}\right) \le n \ln\left(1 + \frac{x}{n}\right) \le n\left(1 + \frac{x}{n} - 1\right) = x \ .$$

Aus

$$n\left(1 - \frac{1}{1 + x/n}\right) = \frac{x}{1 + x/n} \to x \qquad (n \to \infty)$$

ergibt sich

$$\ln\left(\left(1 + \frac{x}{n}\right)^n\right) = n \ln\left(1 + \frac{x}{n}\right) \to x \qquad (n \to \infty)$$

d.h. mit S. 5.17

$$x = \lim_{n \to \infty} \ln\left(\left(1 + \frac{x}{n}\right)^n\right) = \ln(e^x)$$

(man beachte: Aus $(1 + x/n)^n \ge 0$ für n genügend groß und $e^x \ne 0$ folgt $e^x > 0$).
Also ist $\ln \circ \exp = \mathrm{id}_{\mathbb{R}}$. Insbesondere ist damit $\ln : (0, \infty) \to \mathbb{R}$ surjektiv, also auch bijektiv nach S. 5.16.3. Hieraus folgt auch, dass exp bijektiv und $\exp = \ln^{-1}$ ist. □

Für die reelle Exponentialfunktion gilt darüber hinaus

Satz 7.7 *1. Für alle* $x \in \mathbb{R}$ *ist* $e^x \ge 1 + x$.

2. Für alle $x < 1$ *ist* $e^x \le \frac{1}{1-x}$.

3. Für alle $x, y \in \mathbb{R}$ *mit* $x < y$ *ist* $e^x < e^y$.

Beweis. 1. Nach S. 5.16.1 ist $1 + \ln y \le y$ für $y > 0$. Mit $y = e^x$ folgt die Behauptung.
2. Nach 1. gilt

$$e^{-x} \ge 1 - x \ ,$$

also für $x < 1$ auch

$$e^x = \frac{1}{e^{-x}} \le \frac{1}{1-x} \ .$$

3. Mit 1. ergibt sich $e^y/e^x = e^{y-x} \ge 1 + (y-x) > 1$, also $e^y > e^x$. □

Weiterhin definieren wir damit allgemeine Potenzen und Logarithmen.

Definition 7.8 Für $a > 0$ und $b \in \mathbb{C}$ setzen wir

$$a^b := \exp(b \cdot \ln a) = e^{b \cdot \ln a} \ .$$

Man beachte, dass aufgrund von S. 5.16.2 für $b = p/q$ mit $p \in \mathbb{Z}$, $q \in \mathbb{N}$ gilt:

$$\ln(a^{p/q}) = \frac{p}{q} \ln a \ .$$

Damit stimmt diese Definition für den Fall $b \in \mathbb{Q}$ mit der Definition aus D. 3.17 überein.

Aus den Rechenregeln für ln und exp erhält man weiterhin

Satz 7.9 *(allgemeine Potenzgesetze)*
1. Für $a > 0$ und $b_1, b_2 \in \mathbb{C}$ gilt $a^{b_1}a^{b_2} = a^{b_1+b_2}$ und $(a^{b_1})^{b_2} = a^{b_1 b_2}$

2. Für $a_1, a_2 > 0$ und $b \in \mathbb{C}$ gilt $a_1^b a_2^b = (a_1 a_2)^b$.

Beweis. 1. Es gilt

$$\begin{aligned} a^{b_1}a^{b_2} &= e^{b_1 \ln a} e^{b_2 \ln a} = e^{b_1 \ln a + b_2 \ln a} = \\ &= e^{(b_1+b_2)\ln a} = a^{b_1+b_2} \end{aligned}$$

und

$$(a^{b_1})^{b_2} = e^{b_2 \ln(a^{b_1})} = e^{b_2 \ln(e^{b_1 \ln a})} = e^{b_2 b_1 \ln a} = a^{b_1 b_2} \ .$$

2. [Ü]. □

Bemerkung und Definition 7.10 Wir betrachten ein festes $a > 0, a \neq 1$. Dann gilt für jedes $x > 0$ mit $y := \ln x / \ln a$:

$$a^y = e^{y \ln a} = e^{\ln x} = x$$

und y ist die einzige (reelle) Lösung dieser Gleichung (für $y_1 < y_2$ ist $a^{y_1} < a^{y_2}$ im Falle $a > 1$ und $a^{y_1} > a^{y_2}$ im Falle $0 < a < 1$). Wir definieren

$$\log_a x := \frac{\ln x}{\ln a} \qquad (x > 0) \ .$$

Es gilt also für $x > 0, y \in \mathbb{R}$

$$y = \log_a x \Longleftrightarrow a^y = x \ .$$

Wir schauen uns nun die Exponentialfunktion speziell für rein imaginäre Argumente an.

Satz 7.11 *Für alle $x \in \mathbb{R}$ ist $|e^{ix}| = 1$.*

Beweis. Zunächst gilt für beliebiges $z \in \mathbb{C}$

$$\overline{s_n(z)} = \overline{\sum_{\nu=0}^{n} \frac{z^\nu}{\nu!}} = \sum_{\nu=0}^{n} \frac{\overline{z}^\nu}{\nu!} = s_n(\overline{z}) \to e^{\overline{z}} \qquad (n \to \infty) \ .$$

Aus $s_n(z) \to e^z \quad (n \to \infty)$ folgt $\overline{s_n(z)} \to \overline{e^z} \quad (n \to \infty)$ ([Ü]). Also ist $\overline{e^z} = e^{\overline{z}}$. Speziell ergibt sich für $z = ix$

$$|e^{ix}|^2 = e^{ix}\overline{e^{ix}} = e^{ix}e^{\overline{ix}} = e^{ix}e^{-ix} \overset{S.7.3}{=} 1 \ .$$

□

Definition 7.12 Für $x \in \mathbb{R}$ setzen wir

$$\cos x := \ \mathrm{Re}\ (e^{ix}) \quad \text{und} \quad \sin x := \ \mathrm{Im}\ (e^{ix}) \ .$$

Die Funktion $\cos : \mathbb{R} \to \mathbb{R}$ heißt *Cosinus-Funktion* und die Funktion $\sin : \mathbb{R} \to \mathbb{R}$ heißt *Sinus-Funktion.*

Aus Eigenschaften der Exponentialfunktion erhält man damit leicht

Satz 7.13 *Für alle $x \in \mathbb{R}$ gilt*

1. $\sin^2 x + \cos^2 x = 1$,
2. $-1 \le \cos x \le 1$ *und* $-1 \le \sin x \le 1$,

3. $\cos(x) = \cos(-x)$ *und* $\sin(-x) = -\sin x$,

4. $\cos(0) = 1$ *und* $\sin(0) = 0$.

Beweis. 1. Nach 7.12 gilt die sog. EULER'sche Formel

$$\boxed{e^{ix} = \cos x + i \sin x} \qquad (x \in \mathbb{R}) \ .$$

Also gilt $1 = |e^{ix}|^2 = (\mathrm{Re}\,(e^{ix}))^2 + (\mathrm{Im}\,(e^{ix}))^2 = \cos^2 x + \sin^2 x$.

2. Folgt sofort aus 1.

3. Es gilt $\cos(-x) = \mathrm{Re}\,(e^{-ix}) = \mathrm{Re}\,(\overline{e^{ix}}) = \mathrm{Re}\,(e^{ix}) = \cos(x)$.
Entsprechend sieht man $\sin(-x) = -\sin x$.

4. E gilt $1 = e^{i0} = \cos(0) + i \sin(0)$. □

Weiter gilt

Satz 7.14 *(Additionstheoreme)*
Für $x, y \in \mathbb{R}$ gilt

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

2. $\sin(x + y) = \sin x \cos y + \cos x \sin y$.

Beweis. Für $x, y \in \mathbb{R}$ gilt

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

Durch Vergleich von Real- und Imaginärteil ergibt sich 1. und 2. □

Wir wollen zum Abschluss dieses Abschnitts auf verschiedene nützliche Ungleichungen eingehen, die später eine wichtige Rolle spielen werden. Die Basis ist folgendes Ergebnis

Satz 7.15 *(Hölder'sche Ungleichung)*
Es seien $z_1, \ldots, z_n$ und $w_1, \ldots, w_n$ komplexe Zahlen. Ferner seien p, q reelle Zahlen mit $p > 1$ und $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt

$$\sum_{\nu=1}^{n} |z_\nu w_\nu| \le \left(\sum_{\nu=1}^{n} |z_\nu|^p \right)^{1/p} \left(\sum_{\nu=1}^{n} |w_\nu|^q \right)^{1/q} \ .$$

Beweis. 1. Wir zeigen zunächst: Für $x \geq 0, y \geq 0$ gilt

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q} =: A \tag{$*$}$$

Da x^p und $y^p \geq 0$ sind, ist die Behauptung im Falle $x = 0$ oder $y = 0$ klar. Es gelte also $x > 0$ und $y > 0$. Aus S. 5.16.1/2 ergibt sich

$$\begin{aligned} p \cdot \ln x &= \ln \frac{x^p}{A} + \ln A \leq \frac{x^p}{A} - 1 + \ln A \\ q \cdot \ln y &= \ln \frac{y^q}{A} + \ln A \leq \frac{y^q}{A} - 1 + \ln A \end{aligned}$$

Division durch p (1. Ungleichung) in q (2. Ungleichung) und Addition ergibt

$$\ln(xy) \leq \frac{1}{A}\left(\frac{x^p}{p} + \frac{y^q}{q}\right) - \left(\frac{1}{p} + \frac{1}{q}\right) + \left(\frac{1}{p} + \frac{1}{q}\right) \ln A = \ln A$$

und damit auch $(*)$.
2. Wir setzen

$$z := \left(\sum_{\nu=1}^{n} |z_\nu|^p\right)^{1/p}, \qquad w := \left(\sum_{\nu=1}^{n} |w_\nu|^q\right)^{1/q}$$

Die Behauptung ist klar falls $z = 0$ oder $w = 0$ (dann ist $\sum\limits_{\nu=1}^{n} |z_\nu w_\nu| = 0$). Es sei also $z > 0$ und $w > 0$.
Aus 1. folgt für $\nu = 1, \ldots, n$

$$\frac{|z_\nu|}{z} \cdot \frac{|w_\nu|}{w} \leq \frac{1}{p}\frac{|z_\nu|^p}{z^p} + \frac{1}{q} \cdot \frac{|w_\nu|^q}{w^q}$$

und damit durch Aufsummieren

$$\frac{1}{z \cdot w} \sum_{\nu=1}^{n} |z_\nu w_\nu| \leq \frac{1}{p}\frac{1}{z^p} \sum_{\nu=1}^{n} |z_\nu|^p + \frac{1}{q}\frac{1}{w^q} \sum_{\nu=1}^{n} |w_\nu|^q = \frac{1}{p} + \frac{1}{q} = 1\,,$$

was äquivalent zur Behauptung ist. □

Bemerkung 7.16 Als wichtiger Spezialfall ergibt sich im Falle $p(= q) = 2$ für beliebige komplexe Zahlen $z_1, \ldots, z_n$ und $w_1, \ldots, w_n$ die *Cauchy-Schwarz'sche Ungleichung*

$$\sum_{\nu=1}^{n} |z_\nu w_\nu| \leq \sqrt{\sum_{\nu=1}^{n} |z_\nu|^2} \cdot \sqrt{\sum_{\nu=1}^{n} |w_\nu|^2}\,.$$

Als wichtige Konsequenz aus der Hölder-Ungleichung erhalten wir

Satz 7.17 *(Minkowski'sche Ungleichung)*
Es seien $z_1, \dots, z_n$ und $w_1, \dots, w_n$ komplexe Zahlen. Ferner sei $p \geq 1$. Dann gilt

$$\left(\sum_{\nu=1}^{n} |z_\nu + w_\nu|^p\right)^{1/p} \leq \left(\sum_{\nu=1}^{n} |z_\nu|^p\right)^{1/p} + \left(\sum_{\nu=1}^{n} |w_\nu|\right)^{1/p} .$$

Beweis. Wir setzen

$$s := \sum_{\nu=1}^{n} |z_\nu + w_\nu|^p .$$

Ist $s = 0$ oder $p = 1$, so ist die Ungleichung klar (Dreiecksungleichung). Es sei also $s > 0$ und $p > 1$.
Wir wählen $q > 0$ so, dass $\frac{1}{p} + \frac{1}{q} = 1$ gilt (d. h. $q = \frac{p}{p-1}$). Aus der Hölder-Ungleichung folgt

$$\begin{aligned}
s &= \sum_{\nu=1}^{n} |z_\nu + w_\nu|^p \leq \sum_{\nu=1}^{n} |z_\nu||z_\nu + w_\nu|^{p-1} + \sum_{\nu=1}^{n} |w_\nu||z_\nu + w_\nu|^{p-1} \\
&\leq \left(\sum_{\nu=1}^{n} |z_\nu|^p\right)^{1/p} \left(\sum_{\nu=1}^{n} |z_\nu + w_\nu|^{(p-1)q}\right)^{1/q} + \\
&+ \left(\sum_{\nu=1}^{n} |w_\nu|^p\right)^{1/p} \left(\sum_{\nu=1}^{n} |z_\nu + w_\nu|^{(p-1)q}\right)^{1/q} = \\
&= \left[\left(\sum_{\nu=1}^{n} |z_\nu|^p\right)^{1/p} + \left(\sum_{\nu=1}^{n} |w_\nu|^p\right)^{1/p}\right] s^{1/q} .
\end{aligned}$$

Division durch $s^{1/q}$ ergibt die Behauptung (da $1 - \frac{1}{q} = \frac{1}{p}$ gilt). $\square$

8 Metrische Räume

In diesem Abschnitt werden wir eine Klasse von Mengen mit einer gewissen Struktur kennen lernen, die sich als besonders geeignet für das "Betreiben von Analysis" herausstellt. Bevor wir dazu kommen, gehen wir jedoch zunächst kurz auf den Begriff der Mächtigkeit für ganz allgemeine Mengen ein.

Definition 8.1 Es seien M, M_1, M_2 beliebige Mengen.

1. M_1 und M_2 heißen *von gleicher Mächtigkeit* (oder kurz *gleichmächtig*), falls eine bijektive Abbildung $f : M_1 \to M_2$ existiert.

2. M heißt *endlich*, falls M gleichmächtig zu $\{1, \ldots, n\}$ für ein $n \in \mathbb{N}$ (oder auch $= \emptyset$) ist. Anderenfalls heißt M *unendlich*.

3. M heißt *abzählbar unendlich* falls M gleichmächtig zu $\mathbb{N}$ ist.

4. M heißt *abzählbar* falls M endlich oder abzählbar unendlich ist. Anderenfalls heißt M *überabzählbar*.

Bemerkung 8.2 Aus D. 8.1 ergibt sich sofort, dass $M \neq \emptyset$ genau dann abzählbar ist, wenn M in der Form $M = \{x_n : n \in \mathbb{N}\}$ geschrieben werden kann. Ist M abzählbar unendlich, so können die x_n dabei paarweise verschieden gewählt werden. Eine solche Darstellung heißt Abzählung von M. Weiter sieht man leicht, dass jede Teilmenge einer abzählbaren Menge wieder abzählbar ist. (Folglich ist auch jede Menge, die eine überabzählbare Menge enthält, selbst überabzählbar.)

Satz 8.3 *Es sei $I \neq \emptyset$ eine abzählbare Menge, und es seien A_n $(n \in I)$ abzählbare Mengen. Dann ist auch $\bigcup\limits_{n \in I} A_n$ abzählbar.*

Beweis. O. E. können wir $I = \mathbb{N}$ annehmen. (Ist I endlich, so können wir ohne Einschränkung $I = \{1, \ldots, n_0\}$ wählen und dann $A_n := A_1$ für $n > n_0$ setzen.) Es sei

$$A_1 = \left\{x_k^{(1)} : k \in \mathbb{N}\right\} \ , \quad A_2 = \left\{x_k^{(2)} : k \in \mathbb{N}\right\}, \ldots$$

Wir betrachten folgende Anordnung der Elemente $x_k^{(n)}; n \in \mathbb{N}, k \in \mathbb{N}$:

$$
\begin{array}{cccccccccccc}
x_1^{(1)} & \rightarrow & x_2^{(1)} & & x_3^{(1)} & \rightarrow & x_4^{(1)} & & x_5^{(1)} & \rightarrow & x_6^{(1)} & \dots \\
 & \swarrow & & \nearrow & & \swarrow & & \nearrow & & \swarrow & & \\
x_1^{(2)} & & x_2^{(2)} & & x_3^{(2)} & & x_4^{(2)} & & x_5^{(2)} & \dots & & \dots \\
\downarrow & \nearrow & & \swarrow & & \nearrow & & \swarrow & & & & \\
x_1^{(3)} & & x_2^{(3)} & & x_3^{(3)} & & x_4^{(3)} & \dots & & \dots & & \dots \\
 & \swarrow & & \nearrow & & \swarrow & & & & & & \\
x_1^{(4)} & & x_2^{(4)} & & x_3^{(4)} & \dots & & \dots & & \dots & & \dots \\
\downarrow & \nearrow & & \swarrow & & & & & & & & \\
x_1^{(5)} & & x_2^{(5)} & \dots & & \dots & & \dots & & \dots & & \dots \\
 & \swarrow & & & & & & & & & & \\
x_1^{(6)} & \dots & & \dots & & \dots & & \dots & & \dots & & \dots \\
\downarrow & & & & & & & & & & &
\end{array}
$$

Hierbei treten alle $x_k^{(n)}$ auf. Diese können durch "Verfolgen der Pfeile" zu einer Folge angeordnet werden:

$$x_1^{(1)}, x_2^{(1)}, x_1^{(2)}, x_1^{(3)}, x_2^{(2)}, x_3^{(1)}, x_4^{(1)}, x_3^{(2)}, \dots$$

Dies ergibt eine Abzählung von $\bigcup_{n \in \mathbb{N}} A_n$, d. h. $\bigcup_{n \in \mathbb{N}} A_n$ ist abzählbar. □

Insbesondere ergibt sich aus S. 8.3

Satz 8.4 *Die Menge $\mathbb{Q}$ der rationalen Zahlen ist abzählbar.*

Beweis. Zunächst sieht man leicht, dass $\mathbb{Z}$ abzählbar ist ([Ü]). Also ist $A_n := \{k/n : k \in \mathbb{Z}\}$ für alle $n \in \mathbb{N}$ abzählbar. Nach S. 8.3 ist folglich auch

$$\bigcup_{n \in \mathbb{N}} A_n = \left\{ \frac{k}{n} : k \in \mathbb{Z}, n \in \mathbb{N} \right\}$$

abzählbar, und damit auch

$$\mathbb{Q} = \left\{ \frac{k}{n} : k \in \mathbb{Z}, n \in \mathbb{N},\ k, n \text{ teilerfremd} \right\} .$$

□

Andererseits ergibt sich aus folgendem Ergebnis unmittelbar die Überabzählbarkeit von $\mathbb{R}$ oder $\mathbb{C}$.

Satz 8.5 *Jedes Intervall in $\mathbb{R}$ ist überabzählbar.*

Beweis.
1. Wir beweisen zunächst das sogenannte Intervallschachtlungsprinzip, das auch für sich genommen von Interesse ist:
Ist I_n eine Folge von abgeschlossenen Intervallen, d. h. $I_n = [a_n, b_n]$, mit $I_{n+1} \subset I_n$ $(n \in \mathbb{N})$ und $b_n - a_n \to 0\,(n \to \infty)$, so existiert ein $x \in \mathbb{R}$ mit

$$\{x\} = \bigcap_{n\in\mathbb{N}} I_n \,.$$

Denn: Nach Voraussetzung ist $(a_n) \uparrow$ und $(b_n) \downarrow$. Außerdem gilt $a_n \leq b_1$ und $b_n \geq a_1$ $(n \in \mathbb{N})$. Also existieren nach dem Hauptsatz über monotone Folgen

$$a := \lim_{n\to\infty} a_n \quad \text{und} \quad b := \lim_{n\to\infty} b_n \,.$$

Aus $a_n \leq b_n$ folgt $a \leq b$, d. h.

$$a_n \leq a \leq b \leq b_n \qquad (n \in \mathbb{N}) \,.$$

Folglich ist

$$[a, b] \subset \bigcap_{n\in\mathbb{N}} I_n \,.$$

Aus $b_n - a_n \to 0$ $(n \to \infty)$ folgt $a = b$, also ist $[a, b]$ einpunktig.

2. O. E. können wir das Intervall $[0, 1]$ betrachten. Angenommen, $[0, 1]$ ist abzählbar, d. h.

$$[0, 1] = \{x_n : n \in \mathbb{N}\} \,.$$

Wir teilen dann $[0, 1]$ in die drei gleich langen Intervalle $[0, 1/3], [1/3, 2/3]$ und $[2/3, 1]$ auf. Dann ist x_1 in einem dieser Intervalle (das wir I_1 nennen) nicht enthalten. Anschließend teilen wir I_1 in drei gleich lange Intervalle (also der Länge $1/9 = 1/3^2$) auf. Dann ist x_2 in einem dieser Intervalle (I_2 genannt) nicht enthalten. So fortfahrend erhalten wir induktiv eine Folge $I_n = [a_n, b_n]$ abgeschlossener Intervalle in $[0, 1]$ mit $I_{n+1} \subset I_n$ sowie $x_n \notin I_n$ und $b_n - a_n = 1/3^n \to 0$ $(n \to \infty)$. Also ist $\bigcap_{n\in\mathbb{N}} I_n = \{x\}$ für ein $x \in [0, 1]$. Nach Konstruktion gilt $x_n \notin \bigcap_{n\in\mathbb{N}} I_n$, d. h. $x_n \neq x$ für alle $n \in \mathbb{N}$. Widerspruch! □

Wie wir bereits in den vorhergehenden Abschnitten gesehen haben, besteht ein zentrales Anliegen der Analysis darin, "Grenzwerte" zu untersuchen. Grob gesagt bedeutet "$x_n \to x$", dass x_n für große n "nahe bei" x liegt. Es ist also wesentlich, "Abstände" zwischen Elementen einer Menge bestimmen zu können. Eine Klasse von Räumen mit dieser Eigenschaft wollen wir nun definieren, die sog. metrischen Räume. Es wird sich später zeigen, dass diese Räume für viele Fragen der Analysis den geeigneten Rahmen bilden.

Definition 8.6 Es sei $X \neq \emptyset$ eine Menge. Eine Abbildung $d : X \times X \to \mathbb{R}$ heißt *Metrik (auf X)*, falls folgende Bedingungen erfüllt sind:

(d.1) (Definitheit)
$d(x,x) = 0$ für alle $x \in X$ und $d(x,y) > 0$ für alle $x, y \in X$ mit $x \neq y$.

(d.2) (Symmetrie)
Für alle $x, y \in X$ ist $d(x,y) = d(y,x)$.

(d.3) (Dreiecksungleichung)
Für alle $x, y, z \in X$ gilt $d(x,y) \leq d(x,z) + d(z,y)$.

Das Paar (X, d) heißt dann *metrischer Raum.*

Beispiel 8.7 1. Es sei $X = \mathbb{R}$ und es sei

$$d(x,y) := d_{|\cdot|}(x,y) := |x - y| \qquad (x, y \in \mathbb{R})$$

(wobei $|\cdot|$ der Betrag in $\mathbb{R}$ ist).
Dann ist d eine Metrik auf $\mathbb{R}$, wie sich leicht aus den Eigenschaften von $|\cdot|$ ergibt ([Ü]). Entsprechend ist durch

$$d(z,w) := d_{|\cdot|}(z,w) := |z - w| \qquad (z, w \in \mathbb{C})$$

eine Metrik auf $\mathbb{C}$ gegeben.

2. Es sei $X \neq \emptyset$ eine beliebige Menge. Dann ist durch

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

eine Metrik auf X gegeben (die sog. diskrete Metrik). Auch dies ergibt sich leicht durch Überprüfen von (d.1)–(d.3).

Wir betrachten nun eine Klasse metrischer Räume (die $(\mathbb{R}, d_{|\cdot|})$ und $(\mathbb{C}, d_{|\cdot|})$ enthält), deren Metriken sich durch besonders schöne Eigenschaften auszeichnen. Wir stellen damit gleichzeitig erstmals eine Verbindung zur Linearen Algebra her.

Definition 8.8 Es sei $V = (V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$. Eine Abbildung $||\cdot|| : V \to \mathbb{R}$ heißt *Norm* (auf V), falls folgende Bedingungen erfüllt sind

(N.1) (Definitheit)
$||0|| = 0$ und $||x|| > 0$ für alle $x \neq 0$.

(N.2) (Homogenität)
Für alle $x \in V$ und alle $\lambda \in \mathbb{K}$ ist $||\lambda x|| = |\lambda| \cdot ||x||$.

(N.3) (Dreiecksungleichung)
Für alle $x, y \in V$ gilt $||x + y|| \leq ||x|| + ||y||$.

Wir nennen dann $(V, ||\cdot||)$ einen *normierten* Raum.

Satz 8.9 *Ist* $(V, ||\cdot||)$ *ein normierter Raum, so ist durch*

$$d(x,y) := d_{||\cdot||}(x,y) := ||x-y|| \qquad (x, y \in V)$$

eine Metrik auf V *gegeben.*

Beweis. (d.1) ergibt sich unmittelbar aus (N.1).
Zu (d.2): Sind $x, y \in V$, so gilt

$$d(x,y) = ||x-y|| = ||-(y-x)|| \overset{(N.2)}{=} ||y-x|| = d(y,x) \ .$$

Zu (d.3): Sind $x, y, z \in V$, so gilt

$$d(x,y) = ||x-y|| = ||x-z+z-y|| \overset{(N.3)}{\leq} ||x-z|| + ||z-y|| = d(x,z) + d(z,y)$$

□

Beispiel 8.10 1. Ist $|\cdot|$ der Betrag in $\mathbb{R}$ bzw. $\mathbb{C}$, so ist $|\cdot|$ eine Norm auf $\mathbb{R}$ bzw. $\mathbb{C}$ (als Vektorraum über $\mathbb{R}$ bzw. $\mathbb{C}$)
2. Es sei

$$\mathbb{R}^m := \{(x_1, \ldots, x_m) : x_j \in \mathbb{R} \ \text{ für } j = 1, \ldots, m\}$$

und

$$\mathbb{C}^m := \{(z_1, \ldots, z_m) : z_j \in \mathbb{C} \ \text{ für } j = 1, \ldots, m\}$$

(mit der üblichen Addition und Skalarmultiplikation; vgl. Lineare Algebra). Wir schreiben in Zukunft auch kurz $\mathbb{K}^m$ für $\mathbb{R}^m$ und $\mathbb{C}^m$. Für jedes $p \geq 1$ ist durch

$$||x|| := ||x||_p := \left(\sum_{j=1}^m |x_j|^p \right)^{1/p} \qquad (x = (x_1, \ldots, x_m) \in \mathbb{K}^m)$$

eine Norm auf $\mathbb{K}^m$ gegeben.
(Denn: (N.1) und (N.2) ergeben sich sofort aus Eigenschaften des Betrags $|\cdot|$ und der allgemeinen Potenzen; (N.3) ist die Minkowski-Ungleichung (S. 7.17)).
Damit ist für jedes $p \geq 1$ nach S. 8.9

$$d(x,y) := d_p(x,y) = ||x-y||_p \qquad (x, y \in \mathbb{K}^m)$$

eine Metrik auf $\mathbb{K}^m$.
Für $p = \infty$ setzen wir zudem

$$||x||_\infty := \max_{1 \leq j \leq m} |x_j| \qquad (x = (x_1, \ldots, x_m) \in \mathbb{K}^m) \ .$$

Dann ist auch $||\cdot||_\infty$ eine Norm auf $\mathbb{K}^m$ ([Ü]).
Im Falle $p = 2$ heißt die obige Norm $||\cdot||_2$ (für $\mathbb{K} = \mathbb{R}$) *euklidische Norm*, und die Metrik d_2 heißt *euklidische Metrik*. Wir schreiben für $||\cdot||_2$ auch kurz $|\cdot|$.
Falls nichts anderes angegeben ist, soll im weiteren Verlauf der Vorlesung $\mathbb{K}^m$ stets mit der Norm $|\cdot| = ||\cdot||_2$ und der Metrik $d_{|\cdot|} = d_2$ versehen sein.

Beispiel 8.11 Wir betrachten $x = (1, 3-2)$ und $y = (2, -1, 0)$ in $\mathbb{R}^3$. Es gilt

$$|x| = ||x||_2 = \sqrt{1+9+4} = \sqrt{14}\,, \qquad |y| = ||y||_2 = \sqrt{4+1} = \sqrt{5}$$

und

$$d_2(x,y) = |x-y| = ||x-y||_2 = |(-1,4,-2)| = \sqrt{1+16+4} = \sqrt{21}\,.$$

Bemerkung und Definition 8.12 Es sei $M \neq \emptyset$ eine Menge. Eine Funktion $f : M \to \mathbb{K}^m$ heißt *beschränkt*, falls ein $C > 0$ existiert mit

$$|f(t)| \leq C \qquad \text{für alle} \quad t \in M\,.$$

Wir setzen

$$B(M, \mathbb{K}^m) := \{f : M \to \mathbb{K}^m : f \text{ beschränkt}\}\,.$$

($B(M, \mathbb{K}^m)$ ist mit der üblichen Addition und Skalarmultiplikation für Funktionen, also

$$(f+g)(t) := f(t) + g(t), \quad (\lambda f)(t) := \lambda \cdot f(t) \quad (t \in M)$$

für $f, g \in B(M, \mathbb{K}^m)$ und $\lambda \in \mathbb{K}$, ein Vektorraum ([Ü]).)
Weiter setzen wir für $f, g \in B(M, \mathbb{K}^m)$

$$||f|| := ||f||_\infty := \sup\{|f(t)| : t \in M\}$$

($||f||$ heißt sup-Norm oder ∞-Norm von f) und

$$d(f,g) := ||f-g||_\infty\,.$$

Dann ist $||\cdot||$ eine Norm auf $B(M, \mathbb{K}^m)$ ([Ü]), also d eine Metrik auf $B(M, \mathbb{K}^m)$.

Bemerkung und Definition 8.13 Es sei für $p \geq 1$

$$\ell_p := \{a = (a_\nu)_{\nu=0}^\infty : a_\nu \in \mathbb{K}, \sum_{\nu=0}^\infty |a_\nu|^p < \infty\}$$

und

$$||a|| := ||a||_p = \left(\sum_{\nu=0}^\infty |a_\nu|^p\right)^{1/p} \qquad (a \in \ell_p)\,.$$

Dann ist ℓ_p ein Vektorraum und $||\cdot||_p$ eine Norm auf ℓ_p ([Ü]) und folglich ist durch

$$d(a,b) := d_p(a,b) := ||a-b||_p \qquad (a, b \in \ell_p)$$

eine Metrik auf $X = \ell_p$ gegeben.

Wir untersuchen nun Folgen in metrischen Räumen

Definition 8.14 Es sei (X,d) ein metrischer Raum, und es sei (x_n) eine Folge in X. Ferner sei $x \in X$. Die Folge (x_n) heißt *konvergent* (in X) gegen (den *Grenzwert*) x, falls gilt

$$d(x_n,x) \to 0 \qquad (n \to \infty)\ ,$$

d. h. falls zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ existiert mit

$$d(x_n,x) < \varepsilon \qquad (n \geq N_\varepsilon)\ .$$

Wir schreiben wieder

$$x = \lim_{n\to\infty} x_n \qquad \text{oder} \quad x_n \to x \quad (n \to \infty)\ .$$

Beispiel 8.15 Es sei $X = \mathbb{R}$ und $(x_n) = (1/n)$. Dann gilt $x_n \to 0$ $(n \to \infty)$ im metrischen Raum $(\mathbb{R}, d_{|\cdot|})$ aber **nicht** im metrischen Raum $(\mathbb{R}, \delta)$, wobei δ die diskrete Metrik ist. (Dort gilt: Für alle $x \in \mathbb{R}$ existiert ein $n_0 = n_0(x)$ so, dass $\delta(x_n,x) = 1$ für alle $n \geq n_0$.)
Dieses Beispiel zeigt insbesondere, dass die Konvergenz einer Folge von der zu Grunde liegenden Metrik abhängen kann!

Wie im Fall $(X,d) = (\mathbb{K}, d_{|\cdot|})$ gilt

Satz 8.16 *Jede Folge (x_n) in einem metrischen Raum (X,d) hat höchstens einen Grenzwert.*

Beweis. Es seien $x, \tilde{x}$ Grenzwerte von (x_n). Dann gilt mit (d.3)

$$d(x,\tilde{x}) \leq d(x,x_n) + d(x_n,\tilde{x}) \to 0 \qquad (n \to \infty)\ ,$$

also ist $d(x,\tilde{x}) = 0$. Aufgrund von (d.1) ist $x = \tilde{x}$. □

Bemerkung 8.17 Es sei $(X,d) = (\mathbb{K}^m, d_2)$, also

$$d(x,y) = d_2(x,y) = ||x-y||_2 = \left(\sum_{j=1}^{m} |x_j - y_j|^2\right)^{1/2} \qquad (x,y \in \mathbb{K}^m)$$

(vgl. B. 8.10.2). Es sei (x_n) eine Folge in $\mathbb{K}^m$ und

$$x_n = (x_{1n}, x_{2n}, \ldots, x_{mn}) \qquad (n \in \mathbb{N})$$

(d. h. $x_{1n}, \dots, x_{mn}$ sind die Koordinaten von x_n). Ist $x = (x_1, \dots, x_m) \in \mathbb{K}^m$, so gilt

$$\lim_{n\to\infty} x_n = x$$

(d. h. $||x_n - x||_2 \to 0$ $(n \to \infty)$) genau dann, wenn für die "Koordinatenfolgen" $(x_{jn})_n$ gilt

$$\lim_{n\to\infty} x_{jn} = x_j \qquad (j = 1, \dots, m)\,.$$

(Denn:
"$\Rightarrow$:" Gilt $x_n \to x$ $(n \to \infty)$, so ergibt sich für $j = 1, \dots, m$

$$|x_{jn} - x_j| \le \left(\sum_{j=1}^{m} |x_{jn} - x_j|^2 \right)^{1/2} = ||x_n - x||_2 \to 0 \qquad (n \to \infty)$$

also

$$x_{jn} \to x_j \qquad (n \to \infty)\,.$$

"$\Leftarrow$:" Es gelte nun $x_{jn} \to x_j$ $(n \to \infty)$ für $j = 1, \dots, m$. Ist $\varepsilon > 0$ gegeben, so existieren $N_\varepsilon^{(j)} \in \mathbb{N}$ für $j = 1, \dots, m$ mit

$$|x_{jn} - x_j| < \frac{\varepsilon}{\sqrt{m}} \qquad \left(n \ge N_\varepsilon^{(j)}\right)\,.$$

Also gilt für $n \ge \max\left(N_\varepsilon^{(1)}, \dots, N_\varepsilon^{(m)}\right)$:

$$||x_n - x||_2 = \left(\sum_{j=1}^{m} |x_{jn} - x_j|^2 \right)^{1/2} \le \left(m \cdot \frac{\varepsilon^2}{m} \right)^{1/2} = \varepsilon\,.$$

Da $\varepsilon > 0$ beliebig war, gilt $d(x_n, x) = ||x_n - x||_2 \to 0$ $(n \to \infty)$, d. h.
$x_n \to x$ $(n \to \infty)$.)

Beispiel 8.18 Es sei $m = 3$ und

$$x_n = \left(\frac{1}{n} \,,\; \left(1 + \frac{1}{n}\right)^n \,,\; 1 - \frac{1}{2n} \right) = (x_{1n} \,,\; x_{2n} \,,\; x_{3n})\,.$$

Dann gilt

$$x_{1n} \to 0\,, \quad x_{2n} \to e\,, \quad x_{3n} \to 1 \qquad (n \to \infty)$$

also folgt mit B. 8.17

$$||x_n - (0, e, 1)||_2 \to 0 \qquad (n \to \infty)\,,$$

d. h.

$$x_n \to (0, e, 1) \qquad (n \to \infty)\,.$$

Definition 8.19 Es sei $(V, ||\cdot||)$ ein normierter Raum. Eine Menge $M \subset V$ heißt *beschränkt*, falls ein $C > 0$ existiert mit $||x|| \leq C$ für alle $x \in M$. Eine Folge $(x_n)_{n\in\mathbb{N}}$ in V heißt *beschränkt*, falls $\{x_n : n \in \mathbb{N}\}$ beschränkt ist.

In Verallgemeinerung von S. 5.24 erhalten wir

Satz 8.20 *(Bolzano-Weierstraß für Folgen in $\mathbb{K}^m$)*
Es sei $m \in \mathbb{N}$. Dann gilt: Jede beschränkte Folge in $\mathbb{K}^m$ besitzt eine konvergente Teilfolge.

Beweis. (Per Induktion nach m)
1. Induktionsanfang $m = 1$: Ergibt sich aus S. 5.24.
2. Induktionsschritt $m \to m+1$: Es sei (x_n) eine Folge in $\mathbb{K}^{m+1}$ und

$$x_n = (x_{1,n}, \ldots, x_{m,n}, x_{m+1,n}) \qquad (n \in \mathbb{N}) .$$

Dann ist (y_n) mit

$$y_n := (x_{1,n}, \ldots, x_{m,n}) \qquad (n \in \mathbb{N})$$

eine Folge in $\mathbb{K}^m$ und es gilt

$$||y_n||_2 = \left(\sum_{j=1}^{m} |x_{j,n}|^2\right)^{1/2} \leq \left(\sum_{j=1}^{m+1} |x_{j,n}|^2\right)^{1/2} = ||x_n||_2$$

d. h. (y_n) ist beschränkt. Nach Induktionsveraussetzung existiert eine Teilfolge (y_{n_k}) mit

$$y_{n_k} \to y \qquad (k \to \infty)$$

für ein $y \in \mathbb{K}^m$.
Ist $z_k := x_{m+1,n_k}$, so ist $|z_k| \leq ||x_{n_k}||$ für alle $k \in \mathbb{N}$, also ist (z_k) beschränkt in $\mathbb{K}$. Nach S. 5.24 existieren eine Teilfolge $(z_{k_\ell})_\ell$ von $(z_k)_k$ und ein $z \in \mathbb{K}$ mit

$$x_{m+1,n_{k_\ell}} = z_{k_\ell} \to z \qquad (\ell \to \infty) .$$

Ist $y = (y_1, \ldots, y_m)$, so folgt aus $y_{n_{k_\ell}} \to y$ $(\ell \to \infty)$ und B. 8.17

$$x_{1,n_{k_\ell}} \to y_1 \ , \quad x_{2,n_{k_\ell}} \to y_2, \ldots, \quad x_{m,n_{k_\ell}} \to y_m \qquad (\ell \to \infty) .$$

Also gilt wieder mit B. 8.17

$$x_{n_{k_\ell}} = \left(x_{1,n_{k_\ell}}, \ldots, x_{m,n_{k_\ell}}, x_{m+1,n_{k_\ell}}\right) \to (y_1 \ldots, y_m, z)$$

für $\ell \to \infty$. □

Ähnlich wie wir die Definition der Folgenkonvergenz in allgemeinen metrischen Räumen auf die Definition in $\mathbb{R}$ zurückgeführt haben, gehen wir jetzt bei der Einführung von Cauchy-Folgen vor.

Definition 8.21 Es sei (X,d) ein metrischer Raum. Eine Folge (x_n) in X heißt *Cauchy-Folge*, falls zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ existiert mit

$$d(x_n, x_m) < \varepsilon \qquad (n, m \geq N_\varepsilon) .$$

Bemerkung 8.22 Ist (X,d) ein metrischer Raum, und ist (x_n) eine Folge in X, so gilt wie im Fall $(X,d) = (\mathbb{K}, d_{|\cdot|})$: Ist (x_n) konvergent, so ist (x_n) eine Cauchy-Folge. Die Umkehrung ist allerdings im Allgemeinen falsch!
Betrachtet man etwa $(X,d) = (\mathbb{Q}, d_{|\cdot|})$ (also die rationalen Zahlen mit der Betragsmetrik), so ist die Folge $(x_n)_{n=0}^\infty$ mit $x_0 = 2$ und

$$x_{n+1} := \frac{1}{2}\left(x_n + \frac{2}{x_n}\right) \qquad (n \in \mathbb{N}_0)$$

eine Folge in $X = \mathbb{Q}$. Betrachtet man (x_n) als Folge in $(\mathbb{R}, d_{|\cdot|})$, so gilt $x_n \to \sqrt{2}$ (vgl. B. 5.13.2), also ist (x_n) eine Cauchy-Folge in $(\mathbb{R}, d_{|\cdot|})$ und damit auch in $(\mathbb{Q}, d_{|\cdot|})$. Da jedoch $\sqrt{2} \notin \mathbb{Q}$ ist, kann die Folge in $(\mathbb{Q}, d_{|\cdot|})$ nicht konvergent sein. (Die Konvergenz in $(\mathbb{Q}, d_{|\cdot|})$ würde auch die Konvergenz in $(\mathbb{R}, d_{|\cdot|})$ implizieren, d. h., ein rationaler Grenzwert würde der Eindeutigkeit des Grenzwertes (B. 5.2) widersprechen.)

Definition 8.23 Es sei (X,d) ein metrischer Raum. (X,d) heißt *vollständig*, falls jede Cauchy-Folge in (X,d) konvergiert.

Beispiel 8.24 1. Nach S. 5.26 sind $\mathbb{R}$ und $\mathbb{C}$ (mit der Metrik $d_{|\cdot|}$) vollständig. Dasselbe gilt für $(\mathbb{K}^m, d_{|\cdot|})$ für beliebige $m \in \mathbb{N}$.
(Denn: Ist (x_n) eine Cauchy-Folge in $(\mathbb{K}^m, d_{|\cdot|})$, so gilt mit $x_n = (x_{1n}, \ldots, x_{mn})$:

$$|x_{jn} - x_{jm}| \leq ||x_n - x_m||_2 \qquad (n, m \in \mathbb{N}) ,$$

also sind auch die Folgen $(x_{jn})_n$ für $j = 1, \ldots, m$ Cauchy-Folgen in $\mathbb{K}$. Da $(\mathbb{K}, d_{|\cdot|})$ vollständig ist, sind diese konvergent und damit ist nach B. 8.17 auch (x_n) konvergent in $(\mathbb{K}^m, d_{|\cdot|})$.)

2. Nach B. 8.22 ist $(\mathbb{Q}, d_{|\cdot|})$ nicht vollständig.

3. Auf $\mathbb{R}$ ist durch

$$d(x,y) := \left| \frac{x}{1+|x|} - \frac{y}{1+|y|} \right| \qquad (x, y \in \mathbb{R})$$

eine Metrik definiert ([Ü]).
Die Folge (x_n) mit $x_n = n$ ist eine Cauchy-Folge in $(\mathbb{R}, d)$, die nicht konvergiert.
(Denn: Es gilt für $n, m \in \mathbb{N}$ mit $n \geq m$

$$d(x_n, x_m) = \left| \frac{n}{1+n} - \frac{m}{1+m} \right| = \frac{n-m}{(1+n)(1+m)} \leq \frac{n}{1+n} \cdot \frac{1}{1+m} \leq \frac{1}{1+m} .$$

Also existiert zu jedem $\varepsilon > 0$ ein N_ε mit $d(x_n, x_m) < \varepsilon$ für alle $n \geq m \geq N_\varepsilon$, d. h. (x_n) ist eine Cauchy-Folge in $(\mathbb{R}, d)$. Angenommen, (x_n) ist konvergent; $x_n \to x \in \mathbb{R}$. Ist $n_0 \in \mathbb{N}, n_0 > x$, so gilt für alle $n \geq n_0$

$$d(n, x) = \left| \frac{n}{1+n} - \frac{x}{1+|x|} \right| = \frac{n}{1+n} - \frac{x}{1+|x|} \geq \frac{n_0}{1+n_0} - \frac{x}{1+|x|} > 0$$

also $d(n, x) \not\to 0$ $(n \to \infty)$. Widerspruch!)

Damit ist $(\mathbb{R}, d)$ nicht vollständig. Es zeigt sich (mit 1.), dass die Völlständigkeit eine Eigenschaft ist, die nicht nur von der zugrunde liegenden Menge (hier $\mathbb{R}$), sondern auch von der Metrik abhängt!

9 Topologische Grundbegriffe

Wir untersuchen nun die Struktur von Teilmengen metrischer Räume. Dabei werden wir oft die Fälle $\mathbb{R}$ oder $\mathbb{R}^2$ zur Veranschaulichung heranziehen. Das Konzept ist jedoch wesentlich allgemeiner ohne das die Begriffe sich gegenüber diesen sehr speziellen Fällen ändern.

Definition 9.1 Es sei (X, d) ein metrischer Raum und es sei $x_0 \in X$.

1. Für $\varepsilon > 0$ heißt die Menge
$$U_\varepsilon(x_0) := \{x \in X : d(x, x_0) < \varepsilon\}$$
ε-*Umgebung* von x_0 (oder auch *Kugel* mit *Radius* ε um x_0).

2. Eine Menge $U \subset X$ heißt *Umgebung* von x_0, falls ein $\varepsilon > 0$ existiert mit $U_\varepsilon(x_0) \subset U$.

Beispiel 9.2 1. Ist $(X, d) = (\mathbb{R}, d_{|\cdot|})$, so ist für $x_0 \in \mathbb{R}$
$$U_\varepsilon(x_0) = \{x \in \mathbb{R} : |x - x_0| < \varepsilon\} = (x_0 - \varepsilon, x_0 + \varepsilon)\ .$$
2. Ist $(X, d) = (\mathbb{C}, d_{|\cdot|})$, so ist für $z_0 \in \mathbb{C}$
$$\begin{aligned} U_\varepsilon(z_0) &= \{z \in \mathbb{C} : |z - z_0| < \varepsilon\} = \\ &= \{z \in \mathbb{C} :\ \mathrm{Re}^2(z - z_0) +\ \mathrm{Im}^2(z - z_0) < \varepsilon^2\} \end{aligned}$$
der Kreis mit Radius ε um z_0.
3. Ist $(X, d) = \left(\mathbb{R}^3, d_{|\cdot|}\right)$ so ist für $x^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}\right) \in \mathbb{R}^3$
$$U_\varepsilon\left(x^{(0)}\right) = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : \left(x_1 - x_1^{(0)}\right)^2 + \left(x_2 - x_2^{(0)}\right)^2 + \left(x_3 - x_3^{(0)}\right)^2 < \varepsilon^2\right\}$$
die Kugel mit Radius ε um $x^{(0)}$.

Bemerkung 9.3 Aus D. 9.1 ergibt sich leicht folgende Charakterisierung der Konvergenz einer Folge (x_n) in einem metrischen Raum (X, d).
Es gilt $x_n \to x$ $(n \to \infty)$ genau dann, wenn zu jeder Umgebung U von x ein $N_U \in \mathbb{N}$ existiert mit
$$x_n \in U \qquad \text{für alle} \quad n \geq N_U\ .$$
([Ü]).

Definition 9.4 Es sei (X, d) ein metrischer Raum, und es sei $M \subset X$.

1. Ein Punkt $x_0 \in M$ heißt *innerer Punkt* von M, falls ein $\varepsilon > 0$ mit $U_\varepsilon(x_0) \subset M$ existiert (d. h. M ist Umgebung von x_0).

2. Die Menge
$$M^0 := \{x \in M : x \text{ innerer Punkt von } M\} \ (\subset M)$$
heißt *innerer Kern* (oder *Inneres*) von M.

3. M heißt *offen* falls $M = M^0$ gilt (d. h. jeder Punkt von M ist innerer Punkt).

Satz 9.5 *Es sei (X, d) ein metrischer Raum.*

1. *Für alle $x_0 \in X$ und $\varepsilon > 0$ ist $U_\varepsilon(x_0)$ offen.*

2. *Sind $M_1, M_2 \subset X$ mit $M_1 \subset M_2$, so ist $M_1^0 \subset M_2^0$.*

3. *Für $M \subset X$ ist M^0 offen und es gilt*
$$M^0 = \bigcup_{\substack{\mathcal{O}\subset M\\ \mathcal{O} \text{ offen}}} \mathcal{O} \ .$$

Beweis. 1. Es sei $x_1 \in U_\varepsilon(x_0)$. Für $\delta := \varepsilon - d(x_1, x_0)$ gilt $\delta > 0$ und für alle $x \in U_\delta(x_1)$ gilt
$$d(x, x_0) \leq d(x, x_1) + d(x_1, x_0) < \delta + d(x_1, x_0) = \varepsilon \ ,$$
d. h. $U_\delta(x_1) \subset U_\varepsilon(x_0)$.

2. Ist $x \in M_1^0$, so existiert ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subset M_1$. Also ist auch $U_\varepsilon(x) \subset M_2$ und damit $x \in (M_2)^0$.

3. a) Wir haben zu zeigen: $(M^0)^0 = M^0$. Da nach D. 9.4.2 jedenfalls $(M^0)^0 \subset M^0$ gilt, genügt es, $(M^0)^0 \supset M^0$ zu zeigen.
Es sei also $x \in M^0$. Dann existiert ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subset M$. Nach 1. ist $U_\varepsilon(x) = (U_\varepsilon(x))^0$ und nach 2. gilt damit
$$U_\varepsilon(x) = (U_\varepsilon(x))^0 \subset M^0 \ .$$
Folglich ist $x \in (M^0)^0$.
b) "$\subset$" ist klar, da $M^0 \subset M$ und M^0 offen nach a).
"$\supset$" Es sei $\mathcal{O} \subset M, \mathcal{O}$ offen. Dann gilt $\mathcal{O} = \mathcal{O}^0 \subset M^0$ nach 2. □

Beispiel 9.6 1. Wir betrachten $(X,d) = (\mathbb{R}, d_{|\cdot|})$ und $M = [a,b]$ für $a,b \in \mathbb{R}, a < b$. Dann ist $M^0 = (a,b) (\neq M$, also M nicht offen).
(Denn:
"$\subset$": Ist $x \in M^0$, so existiert ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subset M = [a,b]$, d. h. $(x-\varepsilon, x+\varepsilon) \subset [a,b]$. Also ist $x \neq a$ und $x \neq b$, d. h. $x \in (a,b)$.
"$\supset$": Ist $x \in (a,b)$, so gilt für $\varepsilon := \min(b-x, x-a)$

$$U = U_\varepsilon(x) = (x-\varepsilon, x+\varepsilon) \subset M$$

und damit ist $x \in M^0$.)

Nach S. 9.5.3 ist (a,b) offen. Die Intervalle $(a,b], [a,b)$ sind nicht offen. Hingegen sind die Intervalle $(-\infty, b)$ und (a, ∞) offen.

2. Für $a = (a_1, \dots, a_m), b = (b_1, \dots, b_m) \in \mathbb{R}^m$ setzen wir

$$\begin{aligned} (a,b) &:= \{x = (x_1, \dots, x_m) \in \mathbb{R}^m : a_j < x_j < b_j \quad (j = 1, \dots, m)\} \\ &= (a_1, b_1) \times \ldots \times (a_m, b_m) \end{aligned}$$

(und entsprechend $[a,b), (a,b]$ bzw. $[a,b]$ mit $[a_j, b_j), (a_j, b_j]$ bzw. $[a_j, b_j]$ anstelle von (a_j, b_j)).
Dann ist $[a,b]^0 = (a,b) \left(= [a,b)^0 = (a,b]^0\right)$ ([Ü]).

Satz 9.7 *Es sei (X,d) ein metrischer Raum.*

1. *Die Vereinigung beliebig vieler offener Mengen ist offen. (D. h. ist $I (\neq \emptyset)$ eine beliebige Indexmenge und sind $\mathcal{O}_\alpha$ offene Mengen für alle $\alpha \in I$, so ist $\bigcup_{\alpha \in I} \mathcal{O}_\alpha$ offen).*
2. *Der Durchschnitt endlich vieler offener Mengen ist offen. (D. h. sind $\mathcal{O}_1, \dots, \mathcal{O}_n$ offene Mengen, so ist $\bigcap_{j=1}^n \mathcal{O}_j$ offen.)*

Beweis. 1. Es sei $x \in \bigcup_{\alpha \in I} \mathcal{O}_\alpha$. Dann existiert ein $\alpha \in I$ mit $x \in \mathcal{O}_\alpha$. Da $\mathcal{O}_\alpha$ offen ist, existiert ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subset \mathcal{O}_\alpha$. Dann ist auch $U_\varepsilon(x) \subset \bigcup_{\alpha \in I} \mathcal{O}_\alpha$. Folglich ist $\bigcup_{\alpha \in I} \mathcal{O}_\alpha$ offen.

2. Es sei $x \in \bigcap_{j=1}^n \mathcal{O}_j$. Dann ist $x \in \mathcal{O}_j$ für alle $j \in \{1, \dots, n\}$. Da $\mathcal{O}_j$ offen ist, existiert ein $\varepsilon_j > 0$ mit $U_{\varepsilon_j}(x) \subset \mathcal{O}_j$. Für $\varepsilon := \min\{\varepsilon_1, \dots, \varepsilon_n\}$ gilt $\varepsilon > 0$ und $U_\varepsilon(x) \subset U_{\varepsilon_j}(x) \subset \mathcal{O}_j$ für $j = 1, \dots, n$ und damit $U_\varepsilon(x) \subset \bigcap_{j=1}^n \mathcal{O}_j$. Folglich ist $\bigcap_{j=1}^n \mathcal{O}_j$ offen. $\square$

Bemerkung 9.8 I. a. ist der Durchschnitt abzählbar vieler offener Mengen nicht mehr offen! Man betrachte etwa $\mathcal{O}_n = (-1/n, 1/n) \subset \mathbb{R}$ (mit $d = d_{|\cdot|}$). Dann ist $\mathcal{O}_n$ offen für alle $n \in \mathbb{N}$, aber

$$\bigcap_{n\in\mathbb{N}} \mathcal{O}_n = \bigcap_{n\in\mathbb{N}} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}$$

ist nicht offen.

Definition 9.9 Es sei (X, d) ein metrischer Raum, und es sei $M \subset X$.

1. M heißt *abgeschlossen*, falls $M^c = X \setminus M$ offen ist.
2. Die Menge

$$\overline{M} := \bigcap_{\substack{A\supset M \\ A \text{ abgeschlossen}}} A$$

heißt *Abschluss* (oder *abgeschlossene Hülle*) von M.

3. Die Menge

$$\partial M := \overline{M} \setminus M^0$$

heißt *Rand* von M.

Beispiel 9.10 1. Für jeden metrischen Raum (X, d) sind $\emptyset$ und X abgeschlossen (da $X = \emptyset^c$ und $\emptyset = X^c$ offen sind). Es gibt also Mengen, die sowohl offen als auch abgeschlossen sind.

2. Es sei $(X, d) = (\mathbb{R}, d_{|\cdot|})$. Dann ist $[a, b]$ abgeschlossen für $a < b$ (denn $\mathbb{R} \setminus [a, b] = (-\infty, a) \cup (b, \infty)$ ist nach B. 9.6.1 und S. 9.7.1 offen). Weiter sind die Intervalle $(-\infty, b]$ und $[a, \infty)$ abgeschlossen. Intervalle der Form $[a, b)$ und $(a, b]$ sind weder offen noch abgeschlossen.

Satz 9.11 *Es sei (X, d) ein metrischer Raum*

1. *Der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen.*
2. *Die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.*

Beweis. Sind A_α $(\alpha \in I)$ abgeschlossen, so gilt

$$\left(\bigcap_{\alpha\in I} A_\alpha\right)^c = \bigcup_{\alpha\in I} A_\alpha^c \qquad \text{und} \qquad \left(\bigcup_{\alpha\in I} A_\alpha\right)^c = \bigcap_{\alpha\in I} A_\alpha^c$$

(de MORGANsche Regeln). Nach Definition sind die A_α^c offen. Damit ergibt sich die Behauptung durch Anwendung von S. 9.7. □

Bemerkung 9.12 Die Vereinigung abzählbar vieler abgeschlossener Mengen ist i. a. nicht mehr abgeschlossen. So sind etwa $A_n = [-1+1/n, 1-1/n]\,(n \in \mathbb{N})$ in $(\mathbb{R}, d_{|\cdot|})$ abgeschlossen, aber

$$\bigcup_{n\in\mathbb{N}} A_n = (-1,1)$$

ist nicht abgeschlossen (da $(-1,1)^c$ nicht offen ist).

Satz 9.13 *Es seien (X,d) ein metrischer Raum und $M \subset X$. Dann gilt:*

1. *$\overline{M}$ ist abgeschlossen.*
2. *M ist genau dann abgeschlossen, wenn $M = \overline{M}$ gilt.*

Beweis. 1. Nach Definition ist $\overline{M}$ Durchschnitt abgeschlossener Mengen. Also ist auch $\overline{M}$ abgeschlossen nach S. 9.11.1.
2. Falls $M = \overline{M}$ gilt, so ist M abgeschlossen nach 1. Ist umgekehrt M abgeschlossen, so folgt $M = \overline{M}$ aus der Definition von $\overline{M}$. □

Damit gilt folgende wichtige Charakterisierung des Abschlusses:

Satz 9.14 *Ist (X,d) ein metrischer Raum und ist $M \subset X$, so gilt*

$$\overline{M} = M \cup H(M)\,.$$

Beweis. "$\subset$": Es sei $x \in \overline{M}$. Ist $x \in M$, so ist $x \in M \cup H(M)$. Es sei also $x \notin M$, und es sei $\varepsilon > 0$ gegeben. Angenommen $(U_\varepsilon(x) \setminus \{x\}) \cap M = \emptyset$. Dann ist $U_\varepsilon(x) \subset M^c$ (beachte: $x \in M^c$). Also ist $A := (U_\varepsilon(x))^c$ abgeschlossen (S. 9.5.1) und $A \supset M$. Folglich ist $A \supset \overline{M}$ und damit $x \in A = (U_\varepsilon(x))^c$. Widerspruch.

"$\supset$": Es sei $x \in M \cup H(M)$. Ist $x \in M$, so ist auch $x \in \overline{M}$. Ist $x \in H(M)$, und ist $A \supset M, A$ abgeschlossen, so ist $A^c \subset M^c$ offen. Angenommen $x \notin A$. Dann ist $x \in A^c$, also existiert ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subset A^c \subset M^c$. Dies widerspricht aber $x \in H(M)$. Folglich ist $x \in A$ und da $A \supset M, A$ abgeschlossen, beliebig war, ist $x \in \overline{M}$. □

Satz 9.15 *Ist (X,d) ein metrischer Raum, und ist $M \subset X$, so sind folgende Aussagen äquivalent:*

a) M ist abgeschlossen

b) $H(M) \subset M$

c) Für alle Folgen (x_n) *in* M *mit* $x_n \to x$ $(n \to \infty)$ *gilt* $x \in M$*.*

Beweis. 1. Die Äquivalenz von a) und b) folgt direkt aus S. 9.14 und S. 9.13.
2. b) $\Rightarrow$ c): Es sei (x_n) eine Folge in M mit $x_n \to x$. Angenommen, $x \notin M$. Dann ist $x_n \neq x$ für alle $n \in \mathbb{N}$, also ist $x \in H(M)$ nach B. **??**. Nach Voraussetzung ist dann auch $x \in M$. Widerspruch!
c) $\Rightarrow$ b): Ist $x \in H(M)$, so existiert nach B. **??** eine Folge (x_n) in M mit $x_n \to x$. Nach Voraussetzung (also c)) ist $x \in M$. □

Definition 9.16 Es sei (X, d) ein metrischer Raum.

1. Eine Teilmenge M von X heißt *dicht* (in X), falls $\overline{M} = X$ gilt.
2. X heißt *separabel*, falls eine abzählbare dichte Teilmenge existiert.

Beispiel 9.17 Es sei $(X, d) = (\mathbb{R}, d_{|\cdot|})$. Dann ist $\mathbb{Q}$ eine abzählbare dichte Teilmenge von $\mathbb{R}$. Also ist $\mathbb{R}$ separabel.

10 Stetige Funktionen zwischen metrischen Räumen

Wir untersuchen jetzt Funktionen $f : M \to Y$, wobei $X = (X, d), M \subset X$ und $Y = (Y, e)$ metrische Räume sind.
Wir wollen zunächst kurz darauf eingehen wie man in einfachen Fällen solche Funktionen veranschaulichen kann. Dazu setzen wir

$$S_f := \mathrm{graph}(f) := \{(x, y) : x \in M,\, y = f(x)\} = \{(x, f(x)) : x \in M\} \subset X \times Y \,.$$

S_f heißt *Graph* von f.

Beispiel 10.1 1. Es sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) = e^x (x \in \mathbb{R})$. Dann ist

$$S_f = \{(x, e^x) : x \in \mathbb{R}\} \subset \mathbb{R}^2$$

2. Es sei $f : [0, \infty) \to \mathbb{R}$ definiert durch

$$f(x) = \sqrt{x} \qquad (x \geq 0) \,.$$

Dann ist $S_f = \{(x, \sqrt{x}) : x \geq 0\} \subset \mathbb{R}^2$.

3. Es sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) := [x]$, wobei $[x] := \max\{n \in \mathbb{Z} : n \leq x\}$ (sog. Gaußklammer von x). Dann ist

$$S_f = \{(x, [x]) : x \in \mathbb{R}\} \subset \mathbb{R}^2 \,.$$

4. Es sei $f : \mathbb{R}^2 \to \mathbb{R}$
definiert durch

$$f(x_1, x_2) := \sin(x_1 x_2) \qquad ((x_1, x_2) \in \mathbb{R}^2) \,.$$

Dann ist

$$S_f = \{(x_1, x_2, \sin(x_1 x_2)) : (x_1, x_2) \in \mathbb{R}^2\} \subset \mathbb{R}^3 \,.$$

Eine der wichtigsten Struktureigenschaften von Funktionen (zwischen metrischen Räumen) ist die Stetigkeit:

Definition 10.2 Es seien (X, d) und (Y, e) metrische Räume, und es sei $f : M \to Y$, wobei $M \subset X$.
1. f heißt *stetig an der Stelle* $x_0 \in M$, falls zu jedem $\varepsilon > 0$ ein $\delta_\varepsilon (= \delta_{\varepsilon, x_0}) > 0$ existiert mit

$$e(f(x), f(x_0)) < \varepsilon \quad \text{für alle } x \in M \text{ mit } d(x, x_0) < \delta_\varepsilon \,.$$

2. f heißt *stetig auf der Menge* $M_0 \subset M$, falls f stetig an jeder Stelle $x_0 \in M_0$ ist. Ist $M_0 = M$, so heißt f kurz *stetig.*

Anschaulich bedeutet die Stetigkeit an einer Stelle x_0, dass die Funktionswerte $f(x)$ für x "nahe bei x_0" auch "nahe bei $f(x_0)$" liegen.

Beispiel 10.3 1. Es sei (X, d) ein metrischer Raum, und es sei $f : X \to X$ definiert durch $f(x) := x$ $(x \in X)$ (d. h. $f = \mathrm{id}_X$). Dann ist f stetig (auf X).
(Denn: Ist $x_0 \in X$ und $\varepsilon > 0$, so gilt für $\delta_\varepsilon := \varepsilon$:

$$d(f(x), f(x)) = d(x, x_0) < \varepsilon \quad \text{für alle} \quad x \in X \text{ mit } d(x, x_0) < \delta_\varepsilon = \varepsilon \text{ .)}$$

Ist $c \in X$ und ist $f : X \to X$ mit $g(x) \equiv c \, (x \in X)$, so ist auch g stetig (auf X).

2. Es sei $(X, d) = (\mathbb{R}, d_{|\cdot|})$ und es sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} 1, & \text{falls} \quad x = 0 \\ x^2, & \text{sonst} \end{cases}$$

Dann ist f stetig an allen Stellen $x_0 \neq 0$.
(Denn: Ist $x_0 \neq 0$ und $\varepsilon > 0$, so gilt für $\delta_\varepsilon := \min(\varepsilon/(3|x_0|), |x_0|)$

$$|f(x) - f(x_0)| = |x + x_0| \cdot |x - x_0| \leq (|x - x_0| + 2|x_0|) \cdot |x - x_0| \leq 3|x_0| \cdot \frac{\varepsilon}{3|x_0|} = \varepsilon$$

für alle $x \in \mathbb{R}$ mit $|x - x_0| < \delta_\varepsilon$).
Ferner ist f unstetig an $x_0 = 0$.
(Denn: Wir betrachten $\varepsilon = 1/2$. Ist $\delta > 0$ beliebig (wobei o. E. $\delta < 1$), so gilt für $x := \delta/2$ einerseits $|x| < \delta$ und andererseits

$$|f(x) - f(0)| = 1 - \delta^2/4 > 1/2 \text{).}$$

Wir wollen nun der obigen "ε-δ_ε-Definition" zwei Charakterisierungen zur Seite stellen, die oft besser zu handhaben sind. Dazu brauchen wir den Begriff des Grenzwertes einer Funktion:

Definition 10.4 Es seien (X, d) und (Y, e) metrische Räume, $M \subset X$, und es sei $f : M \to Y$. Ferner sei x_0 ein Häufungspunkt von M. Wir sagen, f habe an x_0 den *(Funktions-) Grenzwert* g, falls **für alle** Folgen (x_n) in M mit $x_n \to x_0$ $(n \to \infty)$ und $x_n \neq x_0$ $(n \in \mathbb{N})$ gilt

$$\lim_{n \to \infty} f(x_n) = g \text{ .}$$

Wir schreiben dann

$$\lim_{x \to x_0} f(x) = g \quad \text{oder} \quad \lim_{\substack{x \to x_0 \\ x \in M}} f(x) = g \quad \text{oder kurz} \quad f(x) \to g \quad (x \to x_0) \text{ .}$$

Ist speziell $(Y, e) = (\mathbb{R}, d_{|\cdot|})$, so gilt die entsprechende Definition auch für $g = \infty$ und $g = -\infty$.
Ist speziell $(X, d) = (\mathbb{R}, d_{|\cdot|})$ und $x_0 \in H(M^+)$, wobei $M^+ := \{x \in M : x \geq x_0\}$ (bzw. $x_0 \in H(M^-)$, wobei $M^- := \{x \in M : x \leq x_0\}$), so schreiben wir auch

$$f(x_0^+) := \lim_{x \to x_0^+} f(x) := \lim_{\substack{x \to x_0 \\ x \in M^+}} f(x) \quad \text{bzw.} \quad f(x_0^-) := \lim_{x \to x_0^-} f(x) := \lim_{\substack{x \to x_0 \\ x \in M^-}} f(x) .$$

($f(x_0^+)$ heißt *rechtsseitiger Grenzwert* und $f(x_0^-)$ *linksseitiger Grenzwert* an x_0).
Außerdem definieren wir für $X = \mathbb{R}$ die Grenzwerte $\lim_{x \to \infty} f(x)$ (bzw. $\lim_{x \to -\infty} f(x)$) wie oben mit $x_n \to \infty$ (bzw. $x_n \to -\infty$) anstelle von $x_n \to x_0$.

Bemerkung 10.5 1. Man beachte, dass im Falle von D. 10.4 der Punkt x_0 in M liegen kann oder auch nicht. Auch im Falle $x_0 \in M$ spielt der Funktionswert $f(x_0)$ bei der Grenzwertuntersuchung keine Rolle! So ist etwa in B. 10.3.2

$$\lim_{x \to 0} f(x) = 0 ,$$

da für alle Folgen (x_n) in $\mathbb{R}$ mit $x_n \to 0$ und $x_n \neq 0$ gilt $f(x_n) \to 0$.
2. Man sieht leicht ([Ü]), dass im Falle $X = \mathbb{R}$ und $x_0 \in H(M^+) \cap H(M^-)$ gilt: $g = \lim_{x \to x_0} f(x)$ existiert genau dann, wenn $f(x_0^+)$ und $f(x_0^-)$ existieren und $f(x_0^+) = f(x_0^-) = g$ erfüllt ist.

Beispiel 10.6 1. Es sei $(X, d) = (\mathbb{C}, d_{|\cdot|})$. Dann gilt

$$\frac{e^z - 1}{z} \to 1 \quad (z \to 0)$$

(Denn: Für $0 < |z| < 1$ ist

$$\left|\frac{e^z - 1}{z} - 1\right| = \left|\frac{1}{z} \sum_{\mu=1}^{\infty} \frac{z^\mu}{\mu!} - 1\right| = \left|\sum_{\nu=1}^{\infty} \frac{z^\nu}{(\nu+1)!}\right| \leq \sum_{\nu=1}^{\infty} \frac{|z|^\nu}{(\nu+1)!} \leq \sum_{\nu=1}^{\infty} |z|^\nu = \frac{|z|}{1 - |z|}.$$

Ist (z_n) eine Folge in $\mathbb{C} \setminus \{0\}$ mit $z_n \to 0$, so gilt

$$\frac{|z_n|}{1 - |z_n|} \to 0 \qquad (n \to \infty)$$

und folglich auch

$$\left|\frac{e^{z_n} - 1}{z_n} - 1\right| \to 0 \qquad (n \to \infty) .$$

Da (z_n) beliebig war, folgt die Behauptung.)

2. Es sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch

$$f(x) := \operatorname{sign}(x) := \begin{cases} 1, & \text{falls} \quad x > 0 \\ 0, & \text{falls} \quad x = 0 \\ -1, & \text{falls} \quad x < 0 \end{cases} .$$

Dann gilt

$$f(0^+) = \lim_{x \to 0^+} f(x) = 1, \qquad f(0^-) = \lim_{x \to 0^-} f(x) = -1$$

und $\lim_{x \to 0} f(x)$ existiert nicht.

3. Es ist

$$\lim_{x \to \infty} e^x = \infty \ , \qquad \lim_{x \to -\infty} e^x = 0$$

und

$$\lim_{x \to \infty} \ln x = \infty \ , \qquad \lim_{x \to 0^+} \ln x = -\infty$$

Wir beweisen folgende wichtige Charakterisierungen der Stetigkeit:

Satz 10.7 *Es seien (X, d) und (Y, e) metrische Räume, $M \subset X$ und $f : M \to Y$. Ferner sei $x_0 \in M$. Dann sind äquivalent:*

a) f ist stetig an der Stelle x_0.

b) Für alle Folgen (x_n) in M mit $x_n \to x_0$ gilt $f(x_n) \to f(x_0)$.

c) Entweder ist $x_0 \notin H(M)$ oder es ist $\lim_{x \to x_0} f(x) = f(x_0)$.

Beweis. a) $\Rightarrow$ b): Es sei (x_n) eine Folge in M mit $x_n \to x_0$ $(n \to \infty)$, und es sei $\varepsilon > 0$ gegeben. Da f stetig an x_0 ist existiert ein $\delta_\varepsilon > 0$ so, dass $e(f(x), f(x_0)) < \varepsilon$ für alle $x \in M$ mit $d(x, x_0) < \delta_\varepsilon$. Aus $x_n \to x_0$ folgt die Existenz eines $N_\varepsilon = N(\delta_\varepsilon) \in \mathbb{N}$ mit $d(x_n, x_0) < \delta_\varepsilon$ für alle $n \geq N_\varepsilon$. Also gilt auch $e(f(x_n), f(x_0)) < \varepsilon \, (n \geq N_\varepsilon)$.

b) $\Rightarrow$ c): Ist $x_0 \notin H(M)$, so ist nicht zu zeigen. Es sei also $x \in H(M)$ und es sei (x_n) eine Folge in M mit $x_n \to x_0 \, (n \to \infty)$ und $x_n \neq x_0 \, (n \in \mathbb{N})$. Dann gilt $f(x_n) \to f(x_0)$ nach Voraussetzung. Da (x_n) beliebig war, ist

$$\lim_{x \to x_0} f(x) = f(x_0) \ .$$

c) $\Rightarrow$ a): Angenommen, f ist nicht stetig an x_0. Dann existiert ein $\varepsilon > 0$ so, dass für alle $\delta > 0$ ein $x \in M$ existiert mit $d(x, x_0) < \delta$ und $e(f(x), f(x_0)) \geq \varepsilon$. Insbesondere existiert zu jedem $n \in \mathbb{N}$ ein $x_n \in M$ mit $d(x_n, x_0) < 1/n$ und $e(f(x_n), f(x_0)) \geq \varepsilon$ (also auch $x_n \neq x_0$). Damit ist $x_0 \in H(M)$ und für die Folge (x_n) in M gilt $x_n \to x_0 \, (n \to \infty), x_n \neq x_0$ und $f(x_n) \not\to f(x_0)$. Widerspruch! $\square$

Beispiel 10.8 1. Es sei $(X, d) = (\mathbb{R}, d_{|\cdot|})$ und es sei $f : (0, \infty) \to \mathbb{R}$ definiert durch $f(x) := \ln x \, (x > 0)$. Dann ist f stetig auf $(0, \infty)$.
(Denn: Nach S. 5.17 gilt für alle $x_0 \in (0, \infty)$ und alle Folgen (x_n) in $(0, \infty)$ mit $x_n \to x_0$ auch $f(x_n) \to f(x_0) \, (n \to \infty)$. Also ist f stetig an x_0 nach S. 10.7.)

2. Es sei $(X, d) = (\mathbb{R}, d_{|\cdot|})$ und es sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch

$$f(x) := \begin{cases} 1, & \text{falls} \quad x \in \mathbb{Q} \\ 0, & \text{falls} \quad x \in \mathbb{R} \setminus \mathbb{Q} \end{cases} .$$

Dann ist f unstetig an allen $x_0 \in \mathbb{R}$.
(Denn: Ist $x_0 \in \mathbb{R} \setminus \mathbb{Q}$, so existiert eine Folge (x_n) in $\mathbb{Q}$ mit $x_n \to x_0 \, (n \to \infty)$. Also gilt $f(x_n) = 1 \to 1 \neq f(x_0) = 0 \, (n \to \infty)$. Folglich ist f unstetig an x_0 nach S. 10.7.
Entsprechend existiert für $x_0 \in \mathbb{Q}$ eine Folge (x_n) in $\mathbb{R} \setminus \mathbb{Q}$ mit $x_n \to x_0$ (etwa $x_n = x_0 + \sqrt{2}/n$). Für diese gilt $f(x_n) = 0 \to 0 \neq f(x_0) = 1$. Wieder nach S. 10.7 ist f unstetig an x_0.)

Satz 10.9 *Es seien (X, d_X), (Y, d_Y) und (Z, d_Z) metrische Räume, und es seien $f : X \to Y$ und $g : Y \to Z$. Ist $x_0 \in X$ so, dass $y_0 = \lim\limits_{x \to x_0} f(x)$ existiert, und ist g stetig an y_0, so gilt $\lim\limits_{x \to x_0} (g \circ f)(x) = g(y_0)$. Ist zudem f stetig an x_0, so ist auch $g \circ f$ stetig an x_0.*

Beweis. Es sei (x_n) eine Folge in X mit $x_n \to x_0 \, (n \to \infty)$ und $x_n \neq x_0$. Dann gilt $f(x_n) \to y_0 \, (n \to \infty)$. Da g stetig an y_0 ist, gilt dann auch $g(f(x_n)) \to g(y_0) \, (n \to \infty)$ nach S. 10.7. Also ist $\lim\limits_{x \to x_0} (g \circ f)(x) = y_0$. Damit ergibt sich auch die zweite Aussage leicht durch Anwendung von S. 10.7. □

Bemerkung und Definition 10.10 Ist (X, d) ein metrischer Raum, $M \subset X$ und sind $f, g : M \to \mathbb{K}$ (also reell- oder komplexwertig), so definieren wir

$$\begin{array}{lclll} f + g & : & M \to \mathbb{K} & \text{durch} & (f+g)(x) := f(x) + g(x) & (x \in M) \\ f \cdot g & : & M \to \mathbb{K} & \text{durch} & (f \cdot g)(x) := f(x) \cdot g(x) & (x \in M) \\ f/g & : & M \to \mathbb{K} & \text{durch} & (f/g)(x) := f(x)/g(x) & (x \in M) \end{array}$$

(wobei im Falle f/g natürlich $g(x) \neq 0$ vorausgesetzt wird).
Gilt für ein $x_0 \in H(M)$

$$f(x) \to a \qquad \text{und} \qquad g(x) \to b \qquad (x \to x_0),$$

so folgt

$$(f+g)(x) \to a+b \qquad \text{und} \qquad (f\cdot g)(x) \to a\cdot b \quad (x \to x_0)$$

und im Falle $b \neq 0$ auch

$$(f/g)(x) \to a/b \qquad (x \to x_0)\ .$$

(Der Beweis ergibt sich leicht aus den Grenzwertsätzen für konvergente Folgen in $\mathbb{K}$ (S. 5.5).)
Hieraus ergibt sich durch Anwendung von S. 10.7: Ist $x_0 \in M$ und sind f und g stetig an x_0, so sind auch $f+g$, $f\cdot g$ und f/g stetig an x_0.

Beispiel 10.11 1. Es seien $a_0,\dots,a_n \in \mathbb{K}$ und es sei $P:\mathbb{K}\to\mathbb{K}$ definiert durch

$$P(x) := \sum_{\nu=0}^{n} a_\nu x^\nu \qquad (x \in K)\ .$$

(Funktionen dieser Form heißen *Polynome* (in $\mathbb{R}$ bzw. $\mathbb{C}$)).
Dann ist P stetig auf $\mathbb{K}$.
(Denn: Nach B. 10.3.1 sind $P_0, P_1 : \mathbb{K}\to\mathbb{K}$ mit $P_0(x) := c$, wobei $c\in\mathbb{K}$ fest, und $P_1(x) := x$, stetig auf $\mathbb{K}$. Durch sukzessive Anwendung von B/D 10.10 ergibt sich die Stetigkeit von P.)
Ist $Q:\mathbb{K}\to\mathbb{K}$ ein weiteres Polynom, so ist $R = P/Q$ definiert und stetig auf $\mathbb{K}\setminus N_Q$, wobei $N_Q := \{x\in\mathbb{K} : Q(x)=0\}$.

2. Die Funktion $\exp:\mathbb{C}\to\mathbb{C}$ ist stetig
(Denn: Zunächst gilt nach B. 10.6.1

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

Es sei nun $z_0\in\mathbb{C}$ beliebig. Dann gilt für $z\in\mathbb{C}$

$$e^z - e^{z_0} = e^{z_0}(e^{z-z_0}-1)\ .$$

Also folgt (da $z-z_0\to 0$ für $z\to z_0$)

$$\lim_{z\to z_0} e^z - e^{z_0} = e^{z_0}\cdot \lim_{z\to z_0}(e^{z-z_0}-1) = e^{z_0}0 = 0$$

und damit

$$\lim_{z\to z_0} e^z = e^{z_0}\ .$$

Mit S.10.7 ergibt sich die Behauptung.)
3. Aus 2. ergibt sich auch die Stetigkeit folgender Funktionen in $\mathbb{R}$:

(i) $f(x) = a^x \qquad (x \in \mathbb{R})$, wobei $a > 0$ fest,

(ii) $f(x) = \cos x \quad (x \in \mathbb{R})$,

(iii) $f(x) = \sin x \quad (x \in \mathbb{R})$,

(iv) $f(x) = x^\alpha \qquad (x > 0)$, wobei $\alpha \in \mathbb{R}$ fest.

(Denn: Für $x \in \mathbb{R}$ gilt

$$\begin{aligned} a^x &= e^{x \cdot \ln a} \\ \cos x &= \frac{1}{2}\left(e^{ix} + e^{-ix}\right) \\ \sin x &= \frac{1}{2i}\left(e^{ix} - e^{-ix}\right) \end{aligned}$$

und für $x > 0$ gilt $x^\alpha = e^{\alpha \cdot \ln x}$. Hieraus ergibt sich die Behauptung leicht mit B. 10.8.1 und 2. durch Anwendung von S. 10.9 und B./D. 10.10.)

Der folgende Satz macht deutlich, wie Stetigkeit und die topologischen Begriffe aus Abschnitt 9 zusammenhängen.

Satz 10.12 *Es seien (X, d) und (Y, e) metrische Räume, und es sei $f : X \to Y$. Dann sind äquivalent:*

a) f ist stetig (auf ganz X).

b) Für alle offenen Mengen $\mathcal{O} \subset Y$ ist $f^{-1}(\mathcal{O})$ offen in X.

c) Für alle abgschlossenen Mengen $A \subset Y$ ist $f^{-1}(A)$ abgeschlossen in X.

Beweis. a) $\Rightarrow$ b): Es sei $\mathcal{O} \subset Y$ offen, und es sei $x \in f^{-1}(\mathcal{O})$. Dann ist $y := f(x) \in \mathcal{O}$. Also existiert ein $\varepsilon > 0$ mit $U_\varepsilon(y) \subset \mathcal{O}$. Da f stetig an x ist, existiert ein $\delta > 0$ mit $f(U_\delta(x)) \subset U_\varepsilon(y)$, also $f(U_\delta(x)) \subset \mathcal{O}$. Dies impliziert wiederum $U_\delta(x) \subset f^{-1}(\mathcal{O})$ und damit ist $f^{-1}(\mathcal{O})$ offen.

b) $\Rightarrow$ c): Es sei $A \subset Y$ abgeschlossen. Dann ist A^c offen und damit ist nach Voraussetzung auch $f^{-1}(A^c)$ offen. Aus $f^{-1}(A^c) = (f^{-1}(A))^c$ folgt, dass $f^{-1}(A)$ abgeschlossen ist.

c) $\Rightarrow$ a): Es sei $x \in X$ und $V := U_\varepsilon(f(x))$, wobei $\varepsilon > 0$ beliebig. Dann ist V^c abgeschlossen, also auch $f^{-1}(V^c) = (f^{-1}(V))^c$. Folglich ist $f^{-1}(V)$ offen. Da $x \in f^{-1}(V)$ ist, existiert ein $\delta > 0$ mit $U_\delta(x) \subset f^{-1}(V)$, d. h. $f(U_\delta(x)) \subset f(f^{-1}(V)) \subset V$. Also ist f stetig an x. $\square$

Bemerkung 10.13 Für Bildmengen gilt eine S. 10.12 entsprechende Aussage i. a. nicht!
Ist etwa $(X, d) = (\mathbb{R}, d_{|\cdot|})$ und $f(x) = x^2$ $(x \in \mathbb{R})$, so ist für die offene Menge $\mathbb{R}$

$$f(\mathbb{R}) = [0, \infty)$$

was nicht offen ist. Ist $g(x) = \frac{x}{1+|x|}$ $(x \in \mathbb{R})$, so ist für die abgeschlossene Menge $\mathbb{R}$

$$g(\mathbb{R}) = (-1, 1) \ ,$$

was nicht abgeschlossen ist.

11 Stetige Funktionen auf kompakten Mengen

Wir untersuchen nun eine wichtige Klasse von Mengen, deren Struktur sich unter stetigen Funktionen auf die Bildmenge überträgt.

Definition 11.1 Es sei (X, d) ein metrischer Raum, und es sei $K \subset X$.

1. Eine Familie $\mathcal{U}$ offener Mengen in X heißt *offene Überdeckung* von K, falls
$$K \subset \bigcup_{U \in \mathcal{U}} U \, .$$

2. K heißt *kompakt*, falls jede offene Überdeckung $\mathcal{U}$ von K eine endliche Teilüberdeckung enthält, d. h. ist $\mathcal{U}$ eine offene Überdeckung von K, so existieren $U_1, \ldots, U_n \in \mathcal{U}$ mit
$$K \subset \bigcup_{j=1}^{n} U_j \, .$$

Beispiel 11.2 1. Es sei $(X, d) = (\mathbb{R}, d_{|\cdot|})$ und es sei $K = [a, b]$ für $a < b$. Dann ist für jedes $\varepsilon > 0$
$$\mathcal{U} := \{U_\varepsilon(x) : x \in [a, b]\} = \{(x - \varepsilon, x + \varepsilon) : x \in [a, b]\}$$
eine offene Überdeckung von K. Ist $n \in \mathbb{N}$ so, dass $(b - a)/n < \varepsilon$, so gilt für
$$x_j := a + \frac{b - a}{n} j \qquad (j = 0, \ldots, n)$$
hier
$$[a, b] \subset \bigcup_{j=0}^{n} U_\varepsilon(x_j) \, ,$$
d. h. $\mathcal{U}$ enthält eine offene Teilüberdeckung.
2. Es sei $(X, d) = (\mathbb{R}, d_{|\cdot|})$. Dann ist
$$\mathcal{U} := \{(-k, k) : k \in \mathbb{N}\}$$
eine offene Überdeckung von $\mathbb{R}$.
Sind $U_1, \ldots, U_n \in \mathcal{U}$, d. h. $U_1 = (-k_1, k_1), \ldots, U_n = (-k_n, k_n)$ so ist
$$\bigcup_{j=1}^{n} U_j = (-k, k)$$
wobei $k := \max\{k_1, \ldots, k_n\}$. Also existiert keine endliche Teilüberdeckung von $\mathcal{U}$ mit $\mathbb{R} \subset \bigcup\limits_{j=1}^{n} U_j$, d. h. $\mathbb{R}$ ist nicht kompakt. Dieses Beispiel zeigt, dass abgeschlossene Mengen i. a. nicht kompakt sind.
Es gilt jedoch umgekehrt

Satz 11.3 *Es seien (X, d) ein metrischer Raum und $K \subset X$ kompakt. Dann gilt*

1. *K ist abgeschlossen.*
2. *Ist $A \subset K$ abgeschlossen, so ist A kompakt.*

Beweis. 1. Es sei $x \in K^c$. Für alle $y \in K$ betrachten wir

$$U(y) := \{z \in X : d(z, y) < d(x, y)/2\} (= U_{d(x,y)/2}(y)) .$$

Dann ist $\mathcal{U} = \{U(y) : y \in K\}$ eine offene Überdeckung von K. Also existieren $y_1, \ldots, y_n \in K$ mit $K \subset \bigcup_{j=1}^{n} U(y_j)$. Für $V := U_\delta(x)$ mit $\delta := \min_{1 \le j \le n} d(x, y_j)/2$ gilt $V \cap U(y_j) = \emptyset$ für $j = 1, \ldots, n$, also auch $V \cap K = \emptyset$ und damit $V \subset K^c$. Folglich gilt $x \in (K^c)^0$ und da $x \in K^c$ beliebig war, ist K^c offen.
2. [Ü] □

Wesentlich tiefliegender ist folgender Satz, der die Kompaktheit in metrischen Räumen charakterisiert.

Satz 11.4 *Es sei (X, d) ein metrischer Raum, und es sei $K \subset X$. Dann sind äquivalent:*

a) *K ist kompakt.*

b) *Jede unendliche Teilmenge $K' \subset K$ hat einen Häufungspunkt, der in K liegt.*

c) *Jede Folge in K hat eine konvergente Teilfolge mit Grenzwert in K.*

Beim Beweis verwenden wir folgendes Hilfsresultat, das auch für sich genommen interessant ist.

Satz 11.5 *Ist (X, d) ein metrischer Raum mit der Eigenschaft, dass jede Folge in X eine konvergente Teilfolge besitzt, so ist X separabel.*

Beweis. 1. Es sei $\varepsilon > 0$ gegeben. Wir zeigen zunächst: Es existiert eine endliche Teilmenge A_ε von X mit

$$d(a, b) \ge \varepsilon \qquad \text{für alle } a, b \in A_\varepsilon,\ a \ne b \tag{$*$}$$

und

$$U_\varepsilon(x) \cap A_\varepsilon \ne \emptyset \qquad \text{für alle } x \in X . \tag{$**$}$$

Angenommen, eine solche Menge A_ε existiert nicht. Dann definieren wir induktiv eine Folge (x_n) in X so, dass $d(x_j, x_k) \geq \varepsilon$ für alle j, k mit $j \neq k$. Eine solche Folge hat keine konvergente Teilfolge (da keine Teilfolge eine Cauchy-Folge ist). Also ergibt sich ein Widerspruch zur Voraussetzung.
Wir setzen $x_1 \in X$ beliebig und nehmen an, dass wir $x_1, \ldots, x_n$ mit $d(x_j, x_k) \geq \varepsilon$ für $j, k = 1, \ldots, n, j \neq k$, bereits definiert haben. Die Menge $\{x_1, \ldots, x_n\}$ erfüllt dann $(*)$, so dass nach Annahme $(**)$ nicht erfüllt ist. Folglich existiert ein $x_{n+1} \in X$ mit

$$U_\varepsilon(x_{n+1}) \cap \{x_1, \ldots, x_n\} = \emptyset \, ,$$

d. h. $d(x_{n+1}, x_j) \geq \varepsilon$ für $j = 1, \ldots, n$. Damit ist (x_n) wie gewünscht.
2. Wir definieren

$$A := \bigcup_{n=1}^{\infty} A_{1/n} \, .$$

Dann ist A abzählbar (S. 8.3). Ist $x \in X$ und $\varepsilon > 0$ gegeben, so folgt für n mit $1/n < \varepsilon$ aus $(**)$:

$$U_\varepsilon(x) \cap A \supset U_{1/n}(x) \cap A_{1/n} \neq \emptyset \, .$$

Also ist $x \in A$ oder $x \in H(A)$, d. h. $x \in \overline{A}$ nach S. 9.14. Da $x \in X$ beliebig war, ist $\overline{A} = X$. □

Beweis zu S. 11.4. Wir können uns darauf beschränken, die Behauptung für den Spezialfall $K = X$ zu beweisen ([Ü]; man beachte dabei: $K \subset (X, d)$ ist kompakt genau dann, wenn $(K, d_{|K \times K})$ kompakt ist.)

c) $\Rightarrow$ a): Nach S. 11.5 existiert eine abzählbare dichte Teilmenge A von X. Es sei

$$\mathcal{B} := \{U_r(x) : x \in A, r \in \mathbb{Q}, r > 0\} \qquad \left(= \bigcup_{\substack{r>0 \\ r \in \mathbb{Q}}} \{U_r(x) : x \in A\} \right) \, .$$

Dann ist auch $\mathcal{B}$ abzählbar nach S. 8.3.
Es sei nun $\mathcal{U}$ eine offene Überdeckung von $X (= K)$ und

$$\mathcal{B}_0 := \{B \in \mathcal{B} : \exists U \in \mathcal{U} : B \subset U\} \, .$$

Für jedes $B \in \mathcal{B}_0$ wählen wir ein $V_B \in \mathcal{U}$ mit $B \subset V_B$ und setzen

$$\mathcal{U}_0 := \{V_B : B \in \mathcal{B}_0\} \quad (\subset \mathcal{U}) \, .$$

Dann ist $\mathcal{U}_0$ abzählbar (da $\mathcal{B}_0$ abzählbar ist). Wir zeigen, dass $\mathcal{U}_0$ eine offene Überdeckung von X ist. Dazu sei $z \in X$ beliebig. Wir wählen ein $U \in \mathcal{U}$ mit $z \in U$. Da U offen ist, existiert ein $\varepsilon > 0$ mit $U_\varepsilon(z) \subset U$. Nun wählen wir ein $x \in A$ mit $d(x, z) < \varepsilon/2$ und (nach S. 3.23) ein $r \in \mathbb{Q}$ mit $d(x, z) < r < \varepsilon/2$. Dann gilt

$$z \in U_r(x) \subset U_\varepsilon(z) \subset U$$

(beachte: für $y \in U_r(x)$ gilt $d(y,z) \leq d(y,x)+d(x,z) < r+\varepsilon/2 < \varepsilon$), also $B := U_r(x) \in \mathcal{B}_0$. Für V_B gilt $z \in B \subset V_B$ und $V_B \in \mathcal{U}_0$. Folglich ist $\mathcal{U}_0$ eine Überdeckung von X.
(Wir haben also bisher gezeigt: Jede offene Überdeckung besitzt eine abzählbare Teilüberdeckung.)
Es sei $\{V_n : n \in \mathbb{N}\}$ eine Abzählung von $\mathcal{U}_0$. Wir setzen

$$W_n := \bigcup_{j=1}^{n} V_j \,.$$

Dann ist $W_n \subset W_{n+1}\,(n \in \mathbb{N})$ und $\bigcup\limits_{n\in\mathbb{N}} W_n = \bigcup\limits_{j\in\mathbb{N}} V_j = X$. Es genügt zu zeigen: $W_n = X$ für ein $n \in \mathbb{N}$.
Angenommen dies ist nicht der Fall. Dann exisitert für alle $n \in \mathbb{N}$ ein $x_n \in X \setminus W_n$. Nach Voraussetzung hat (x_n) eine Teilfolge $(x_{n_k})_k$ mit Grenzwert $x \in X$. Wir wählen $N \in \mathbb{N}$ mit $x \in W_N$. Dann gilt einerseits $x_n \notin W_N$ für alle $n \geq N$ aber andererseits auch $x_{n_k} \in W_N$ für alle $k \geq k_0$ (da $x_{n_k} \to x\,(k \to \infty)$ und W_N offen). Widerspruch!

a) $\Rightarrow$ b): Angenommen, b) gilt nicht. Dann existiert eine unendliche Teilmenge D von X ohne Häufungspunkt. Also existiert zu jedem $x \in D$ eine offene Umgebung V_x von x mit $D \cap V_x = \{x\}$. Da D abgeschlossen ist (nach S. 9.15) ist $\mathcal{U} := \{U_x := V_x \cup D^c : x \in D\}$ eine offene Überdeckung von X. Ist $\{U_{x_1}, \ldots, U_{x_n}\}$ eine endliche Teilmenge von $\mathcal{U}$, so ist $X \not\subset \bigcup\limits_{j=1}^{n} U_{x_j}$, da für $x \in D \setminus \{x_1, \ldots, x_n\} (\neq \emptyset)$ gilt: $x \notin U_{x_j}$ $(j = 1, \ldots, n)$. Widerspruch zu a)!

b) $\Rightarrow$ c): Es sei (x_n) eine Folge in X. Ist $E := \{x_n : n \in \mathbb{N}\}$ endlich so existiert eine Teilfolge (x_{n_k}) von (x_n) mit $x_{n_k} \equiv x_{n_1}$ für alle $k \in \mathbb{N}$, also gilt $x_{n_k} \to x_{n_1} \in X \quad (n \to \infty)$. Ist E unendlich, so hat E nach Voraussetzung einen Häufungspunkt $x \in X$. Dann existiert auch eine Teilfolge (x_{n_k}) von (x_n) mit $x_{n_k} \to x(n \to \infty)$ (vgl. B. **??**). □

Wir ziehen eine erste wichtige Folgerung aus S. 11.4

Satz 11.6 *(Heine-Borel)*
Es sei $M \subset \mathbb{K}^m$ (mit $d = d_{|\cdot|}$). Dann sind äquivalent

a) M ist kompakt

b) M ist beschränkt und abgeschlossen.

Beweis. a) $\Rightarrow$ b): Nach S. 11.3 ist M abgeschlossen. Außerdem ist M auch beschränkt (Denn: Angenommen, nicht. Dann existiert eine Folge (x_n) in M mit $|x_n| \to \infty$ $(n \to \infty)$. Diese Folge besitzt keine konvergente Teilfolge; man beachte dabei: konvergente

Folgen sind notwendig beschränkt, auch in allgemeinen normierten Räumen. Widerspruch zu c) aus S. 11.4.)

b) $\Rightarrow$ a): Es sei (x_n) eine Folge in M. Da M beschränkt ist, existiert nach dem Satz von Bolzano und Weierstraß (S. 8.20) eine konvergente Teilfolge (x_{n_k}). Ist $x = \lim\limits_{k\to\infty} x_{n_k}$, so ergibt sich $x \in M$ aus der Abgeschlossenheit von M (S. 9.15). Nach S. 11.4 ist M kompakt. □

Beispiel 11.7 Es sei für $a = (a_1, \ldots, a_m),\ b = (b_1, \ldots, b_m) \in \mathbb{R}^m$ mit $a_j < b_j$

$$M = [a, b] = [a_1, b_1] \times \cdots \times [a_m, b_m] .$$

Dann ist M beschränkt und abgeschlossen, wie man sich leicht überlegen kann. Also ist M nach S. 11.6 auch kompakt.

Bemerkung 11.8 1. Eine S. 11.6 entsprechende Aussage gilt nicht mehr in allgemeineren normierten Räumen! Zwar folgt aus der Kompaktheit von M stets die Beschränktheit und die Abgschlossenheit (wie die Beweisrichtung a) $\Rightarrow$ b) von S. 11.6 zeigt). Jedoch sind umgekehrt Teilmengen M, die beschränkt und abgeschossen sind, i. a. nicht kompakt!
Ist etwa $(V, ||\cdot||) = (\ell_2, ||\cdot||_2)$, so ist

$$M := \{a \in \ell_2 : ||a||_2 = \left(\sum_{\nu=0}^{\infty} |a_\nu|^2\right)^{1/2} \le 1\}$$

beschränkt und abgeschlossen, aber nicht kompakt ([Ü]).

2. Ist $K \subset \mathbb{R}$ kompakt, so existieren aufgrund der Beschränktheit $\sup(K)$ und $\inf(K)$ und aus der Abgeschlossenheit folgt leicht ([Ü])

$$\sup(K) \in K \qquad \text{und} \qquad \inf(K) \in K .$$

Wir untersuchen jetzt stetige Funktionen auf kompakten Mengen.

Satz 11.9 *Es seien (X, d) und (Y, e) metrische Räume, und es sei $f : X \to Y$. Dann gilt: Ist $K \subset X$ kompakt und ist f stetig auf K, so ist auch $f(K) \subset Y$ kompakt.*

Beweis. Es sei (y_n) eine Folge in $f(K)$. Dann existieren $x_n \in K$ mit $y_n = f(x_n)$ $(n \in \mathbb{N})$. Da K kompakt ist, existieren nach S. 11.4 ein $x \in K$ und eine Teilfolge (x_{n_j}) von

(x_n) mit $x_{n_j} \to x$ $(j \to \infty)$.
Da f stetig ist, folgt

$$y_{n_j} = f(x_{n_j}) \to f(x) \in f(K) \qquad (j \to \infty) .$$

Wieder nach S. 11.4 ist $f(K)$ kompakt. □

Für reellwertige Funktionen hat der Satz eine wichtige Konsequenz. Um diese formulieren zu können, brauchen wir eine Definition.

Definition 11.10 Es sei (X, d) ein metrischer Raum, und es sei $f : M \to \mathbb{R}$ für eine Teilmenge M von X. Ferner sei $x_0 \in M$.

1. Man sagt, f habe an der Stelle x_0 ein *absolutes* (oder auch *globales*) *Maximum*, falls gilt
$$f(x) \leq f(x_0) \qquad \text{für alle} \quad x \in M ,$$
d. h.
$$f(x_0) = \max f(M) =: \max_M f(x) .$$
2. Man sagt, f habe an der Stelle x_0 ein *absolutes* (oder auch *globales*) *Minimum*, falls gilt
$$f(x) \geq f(x_0) \qquad \text{für alle} \quad x \in M ,$$
d. h.
$$f(x_0) = \min f(M) =: \min_M f(x) .$$

Beispiel 11.11 Es sei $X = \mathbb{R}$ und $f(x) := x^2$ $(x \in \mathbb{R})$. Dann hat f an $x_0 = 0$ ein absolutes Minimum, aber es existiert kein absolutes Maximum.

Satz 11.12 *Es seien (X, d) ein metrischer Raum, $K \subset X$ kompakt und $f : K \to \mathbb{R}$ stetig. Dann hat f ein absolutes Maximum und ein absolutes Minimum, d. h. es existieren $x_1, x_2 \in K$ mit*

$$f(x_1) \leq f(x) \leq f(x_2) \qquad \textit{für alle} \quad x \in K .$$

Beweis. Nach S. 11.9 ist $f(K) \subset \mathbb{R}$ kompakt. Nach B. 11.8.2 existieren $\sup f(K)$ und $\inf f(K)$ und es gilt

$$\sup f(K) \in f(K) , \qquad \inf f(K) \in f(K)$$

(d. h. $\sup f(K) = \max f(K)$ und $\inf f(K) = \min f(K)$). Dies ist lediglich eine Umformulierung der Behauptung. □

Beispiel 11.13 1. Es sei $[-a, a] \subset \mathbb{R}$ und $f(x) := x^2$ $(x \in [a, b])$. Dann ist

$$f(a) = f(-a) = \max_{[-a,a]} f(x)$$

d. h. f hat an a und $-a$ absolute Maxima. Außerdem ist

$$f(0) = \min_{[-a,a]} f(x) \; .$$

2. Es sei $f(x) = \frac{x}{1+|x|}$ $(x \in \mathbb{R})$. Dann hat $f : \mathbb{R} \to \mathbb{R}$ weder ein absolutes Maximum, noch ein absolutes Minimum. Für alle kompakten Intervalle $[a, b] \subset \mathbb{R}$ gilt jedoch

$$f(b) = \max_{[a,b]} f(x) \qquad \text{und} \quad f(a) = \min_{[a,b]} f(x) \; .$$

Eine sehr elegante Anwendung von S. 11.9 ist:

Satz 11.14 *Es seien (K, d) und (Y, e) metrische Räume, und es sei K kompakt. Ist $f : K \to Y$ bijektiv und stetig, so ist auch $f^{-1} : Y \to K$ stetig.*

Beweis. Es sei $A \subset K$ abgeschlossen. Nach S. 10.12 reicht es zu zeigen:
$(f^{-1})^{-1}(A) = f(A) \subset Y$ ist abgeschlossen.
Da K kompakt ist, ist A als abgeschlossene Teilmenge auch kompakt (S. 11.3). Also ist $f(A) \subset Y$ kompakt nach S. 11.9 und damit auch abgeschlossen nach S. 11.3. □

Eine weitere wichtige Eigenschaft stetiger Funktionen auf kompakten Mengen ist die gleichmäßige Stetigkeit:

Definition 11.15 Es seien (X, d) und (Y, e) metrische Räume, und es sei $f : X \to Y$. Ist $M \subset X$, so heißt f *gleichmäßig stetig* auf M, falls für alle $\varepsilon > 0$ ein $\delta_\varepsilon > 0$ so existiert, dass

$$e(f(x_1), f(x_2)) < \varepsilon \qquad \text{für alle } x_1, x_2 \in M \text{ mit } d(x_1, x_2) < \delta_\varepsilon \; .$$

Beispiel 11.16 Es sei $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$ $(x \in \mathbb{R})$. Dann ist f gleichmäßig stetig auf $[0, 1]$.
(Denn: Es sei $\varepsilon > 0$ gegeben. Dann gilt mit $\delta_\varepsilon := \varepsilon/2$ für alle $x_1, x_2 \in [0, 1]$ mit $|x_1 - x_2| < \delta_\varepsilon$:

$$|f(x_1) - f(x_2)| = |x_1 + x_2| \cdot |x_1 - x_2| \le 2|x_1 - x_2| < \varepsilon \; .)$$

Bemerkung 11.17 Aus der Definition ergibt sich sofort, dass jede auf einer Menge M gleichmäßig stetige Funktion auch stetig auf M (d. h. stetig in jedem Punkt $x \in M$) ist. Man beachte weiterhin, dass die gleichmäßige Stetigkeit eine sog. globale Eigenschaft, d. h. eine Eigenschaft der Funktion als Ganzes auf der Menge M, während die Stetigkeit eine sog. lokale Eigenschaft ist, d. h. eine Eigenschaft, die allein von der Funktion in der Nähe eines jeden Punktes abhängt.

Das folgende Beispiel zeigt, dass die Stetigkeit auf einer Menge i. a. nicht die gleichmäßige Stetigkeit impliziert.

Beispiel 11.18 Wir betrachten $f : (0, 1] \to \mathbb{R}$ mit

$$f(x) = \frac{1}{x} \qquad (x \in (0, 1]) \; .$$

Dann ist f stetig auf $(0, 1]$, aber nicht gleichmäßig stetig.
(Es sei $\varepsilon = 1$, und es sei $\delta > 0$ beliebig. Wir wählen $x_1 = 1/n, x_2 = 1/m$, wobei $n, m \in \mathbb{N}, m > n$ und so, dass $1/n < \delta$. Dann ist auch $|x_1 - x_2| = 1/n - 1/m < \delta$, aber

$$|f(x_1) - f(x_2)| = |n - m| \geq 1 = \varepsilon \; .$$

Folglich ist f nicht gleichmäßig stetig auf $(0, 1]$.)

Es gilt jedoch

Satz 11.19 *Es seien (X, d) und (Y, e) metrische Räume und $f : X \to Y$. Ist $K \subset X$ kompakt und ist f stetig auf K, so ist f auch gleichmäßig stetig auf K.*

Beweis. Angenommen nicht. Dann existiert ein $\varepsilon > 0$ so, dass für alle $n \in \mathbb{N}$ zwei Punkte $x_n, y_n \in K$ existieren mit

$$d(x_n, y_n) < 1/n \qquad \text{und} \quad e(f(x_n), f(y_n)) \geq \varepsilon \; .$$

Die Folge (x_n) besitzt nach S. 11.4 eine Teilfolge $(x_{n_j})_j$ mit $x_{n_j} \to x \in K$.
Dann gilt

$$d(x, y_{n_j}) \leq d(x, x_{n_j}) + d(x_{n_j}, y_{n_j}) \to 0 \qquad (n \to \infty) \; .$$

d. h. $y_{n_j} \to x$ $(j \to \infty)$. Also folgt aufgrund der Stetigkeit von f an der Stelle x

$$\varepsilon \leq e(f(x_{n_j}), f(y_{n_j})) \leq e(f(x_{n_j}), f(x)) + e(f(x), f(y_{n_j})) \to 0 \quad (j \to \infty) \; .$$

Widerspruch! □

12 Reellwertige Funktionen einer reellen Variablen

Wir untersuchen nun speziell Funktionen $f : M \to \mathbb{R}$, wobei $M \subset \mathbb{R}$, also reellwertige Funktionen einer reellen Variablen.
Zunächst beweisen wir

Satz 12.1 *(Zwischenwertsatz)*

Es sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall und $f : [a,b] \to \mathbb{R}$ sei stetig auf $[a,b]$. Ist $\underline{m} := \min\limits_{[a,b]} f(x)$, $\overline{m} := \max\limits_{[a,b]} f(x)$, so existiert zu jedem $\eta \in [\underline{m}, \overline{m}]$ ein $\xi \in [a,b]$ mit $f(\xi) = \eta$, d. h.

$$W(f) = [\underline{m}, \overline{m}]$$

ist wieder ein kompaktes Intervall (bzw. ein Punkt).

Beweis. Zunächst beachte man, dass nach S. 11.12 $\underline{m}$ und $\overline{m}$ existieren, d. h. es existieren $\alpha, \beta \in [a,b]$ mit $f(\alpha) = \underline{m}$, $f(\beta) = \overline{m}$. Ist $\underline{m} = \overline{m}$, so ist $f(x) \equiv \underline{m}(= \overline{m})$ auf $[a,b]$ und damit ist die Behauptung klar. Es sei also $\underline{m} < \overline{m}$ und o. E. $\alpha < \beta$. Weiter sei ein $\eta \in (\underline{m}, \overline{m})$ gegeben. Wir betrachten

$$M := \{x \in [\alpha, \beta) : f(x) \leq \eta\} .$$

Dann ist $M \neq \emptyset$ (da $\alpha \in M$) und beschränkt, also existiert

$$\xi := \sup M \in [\alpha, \beta] .$$

Behauptung: Es gilt $f(\xi) = \eta$.
Denn: Es existiert eine Folge (x_n) in M mit $x_n \to \xi$. Da f stetig an $\xi \in [\alpha, \beta] \subset [a,b]$ ist, gilt $f(x_n) \to f(\xi)$ also $f(\xi) \leq \eta$. Aus $f(\beta) = \overline{m} > \eta$ folgt $\xi < \beta$. Ist $(\tilde{x}_n)$ eine Folge in $(\xi, \beta]$ mit $\tilde{x}_n \to \xi$, so folgt $\eta < f(\tilde{x}_n) \to f(\xi)\,(n \to \infty)$, also $f(\xi) \geq \eta$ und damit $f(\xi) = \eta$. □

Bemerkung 12.2 1. Die Aussage von S. 12.1 ist i. a. falsch, falls f unstetig auf $[a,b]$ ist. Ist etwa $f : [-1,1] \to \mathbb{R}$

$$f(x) = \begin{cases} x+1, & x \geq 0 \\ x, & x < 0 \end{cases}$$

so ist $W(f) = [-1, 0) \cup [1, 2]$, also kein Intervall.

2. Ist $I \subset \mathbb{R}$ ein beliebiges (beschränktes oder unbeschränktes) Intervall, und ist $f : I \to \mathbb{R}$ stetig so ist stets $W(f)$ ein Intervall (u. U. ein Punkt).
(Denn: sind $y_1, y_2 \in W(f)$ mit $y_1 < y_2$, so existieren $x_1, x_2 \in I$ mit $f(x_j) = y_j (j = 1, 2)$. O. E. sei $x_1 < x_2$. Dann ist $[y_1, y_2] \subset W(f_{|[x_1,x_2]}) = f([x_1, x_2])$ nach S. 12.1, also auch $[y_1, y_2] \subset W(f)$. Da $y_1, y_2 \in W(f)$ beliebig waren, ist $W(f)$ ein Intervall.)

Damit können wir uns ein genaues Bild über die Umkehrung stetiger Funktionen in $\mathbb{R}$ machen.

Definition 12.3 Es sei $M \subset \mathbb{R}$ und $f : M \to \mathbb{R}$. Dann heißt f

1. *monoton wachsend*, falls $f(x_1) \leq f(x_2)$ für $x_1 < x_2$
2. *streng monoton wachsend*, falls $f(x_1) < f(x_2)$ für $x_1 < x_2$
3. *(streng) monoton-fallend*, falls $-f$ (streng) monoton wachsend ist

Es gilt

Satz 12.4 *Es sei $I \subset \mathbb{R}$ ein Intervall und es sei $f : I \to \mathbb{R}$ stetig und streng monoton wachsend. Dann existiert die Umkehrfunktion f^{-1} auf dem Intervall $J := W(f)$ und f^{-1} ist ebenfalls stetig (auf J) sowie streng monoton wachsend. Eine entsprechende Aussage gilt mit "streng monoton fallend" an Stelle von "streng monoton wachsend".*

Beweis. Aus D. 12.3 sieht man leicht, dass $f : I \to W(f)$ bijektiv ist, d. h. $f^{-1} : W(f) \to I$ existiert. Außerdem ist $J = W(f)$ nach B. 12.2.2 ein Intervall. Weiter ist $f^{-1} : J \to I$ streng monoton wachsend.
(Denn: Angenommen, es existieren $y_1, y_2 \in J$ mit $y_1 < y_2$ und $x_1 = f^{-1}(y_1) \geq f^{-1}(y_2) = x_2$. Dann gilt $y_1 = f(x_1) \geq f(x_2) = y_2$, da f (streng) monoton wachsend ist. Widerspruch!)
Schließlich ist $f^{-1} : J \to I$ stetig.
(Denn: Es sei $y_0 \in J$. Ist $y_0 \in J^0$, so existieren $y_1 < y_0 < y_2$ mit $[y_1, y_2] \subset J$. Ist $x_k := f^{-1}(y_j)\,(j = 1, 2)$, so ist $f : [x_1, x_2] \to [y_1, y_2]$ bijektiv, also f^{-1} stetig auf (y_1, y_2) nach S. 11.14. Insbesondere ist f^{-1} stetig an y_0. Ist $y_0 \in \partial J$, so argumentiert man ähnlich.) □

Wir kommen jetzt noch einmal auf die trigonometrischen Funktionen zu sprechen

Satz 12.5 *Für alle $x \in \mathbb{R}$ gilt*

$$\cos x = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu x^{2\nu}}{(2\nu)!} \qquad \textit{und} \qquad \sin x = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu x^{2\nu+1}}{(2\nu+1)!} \, .$$

Beweis. Nach Definition gilt

$$e^{ix} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} \qquad (x \in \mathbb{R}) ,$$

also erhalten wir

$$e^{ix} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{\nu=0}^{\infty} \frac{(ix)^{2\nu}}{(2\nu)!} + \sum_{\nu=0}^{\infty} \frac{(ix)^{2\nu+1}}{(2\nu+1)!} = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu x^{2\nu}}{(2\nu)!} + i \sum_{\nu=0}^{\infty} \frac{(-1)^\nu x^{2\nu+1}}{(2\nu+1)!} .$$

Da $e^{ix} = \cos(x) + i\sin(x)$ gilt, ergibt sich die Behauptung durch Vergleich von Real- und Imaginärteil. □

Unser Ziel ist es nun, die Zahl π analytisch zu definieren. Dazu zeigen wir

Satz 12.6 *Die Funktion* cos *hat im Intervall* $(0, \infty)$ *eine kleinste Nullstelle* x_0.

Beweis. Nach S. 7.13.4 ist $\cos(0) = 1$. Mit S. 12.5 gilt

$$\begin{aligned}
\cos x &= \sum_{\nu=0}^{\infty} \frac{(-1)^\nu x^{2\nu}}{(2\nu)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} \mp \dots \\
&= 1 - \frac{x^2}{2!}\left(1 - \frac{x^2}{3 \cdot 4}\right) - \frac{x^6}{6!}\left(1 - \frac{x^2}{7 \cdot 8}\right) - \dots \\
&= 1 - \frac{x^2}{2!}\left(1 - \frac{x^2}{3 \cdot 4}\right) - \sum_{k=1}^{\infty} \frac{x^{4k+2}}{(4k+2)!}\left(1 - \frac{x^2}{(4k+3)(4k+4)}\right)
\end{aligned}$$

Für $x = 2$ ergibt sich

$$1 - \frac{2^2}{(4k+3)(4k+4)} > 0 \qquad (k \in \mathbb{N}_0) ,$$

also

$$\cos 2 \leq 1 - \frac{2^2}{2!}\left(1 - \frac{2^2}{3 \cdot 4}\right) = -\frac{1}{3} < 0 .$$

Nach dem Zwischenwertsatz (S. 12.1) existiert ein $\xi \in (0, 2)$ mit $\cos(\xi) = 0$. Für $x_0 := \inf\{\xi : \cos(\xi) = 0,\ \xi > 0\}$ gilt auch $\cos(x_0) = 0$ (da cos stetig auf $\mathbb{R}$) und $x_0 > 0$. □

Definition 12.7 Ist x_0 wie in S. 12.6, so setzen wir

$$\pi := 2x_0 .$$

Bemerkung 12.8 Nach obiger Definition gilt also

$$\cos\left(\frac{\pi}{2}\right) = 0$$

und, da $\cos(x) = \cos(-x)$ ist (S. 7.13.3) ist $\cos(x) > 0$ für $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.
Weiter gilt

$$1 = |e^{i\pi/2}|^2 = \cos^2\left(\frac{\pi}{2}\right) + \sin^2\left(\frac{\pi}{2}\right) = \sin^2\left(\frac{\pi}{2}\right) .$$

Ähnlich wie im Beweis zu S. 12.6 zeigt man, dass $\sin x > 0$ für $x \in (0, 2]$ ist ([Ü]). Also folgt

$$\sin(\pi/2) = 1 .$$

Hieraus ergibt sich wiederum $e^{i\pi/2} = i$ und allgemeiner für $k \in \mathbb{Z}$

$$e^{i\pi k/2} = \left(e^{i\pi/2}\right)^k = i^k = \begin{cases} (-1)^{k/2}, & \text{falls} \quad k \text{ gerade} \\ i(-1)^{(k-1)/2}, & \text{falls} \quad k \text{ ungerade} \end{cases} .$$

Unter Ausnutzung der Additionstheoreme erhält man bekannte Eigenschaften der trigonometrischen Funktionen

Satz 12.9 *Für alle $x \in \mathbb{R}$ gilt*

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

   *(*cos *und* sin *sind "2π-periodisch")*
4. $\sin x = 0$ *genau dann, wenn* $x = k\pi$ *für ein* $k \in \mathbb{Z}$,
5. $\cos x = 0$ *genau dann, wenn* $x = \frac{\pi}{2} + k\pi$ *für ein* $k \in \mathbb{Z}$.

Beweis. Mit S. 7.14.1 und B. 12.8 erhalten wir

$$\cos(x + \pi/2) = \cos(x)\cos(\pi/2) - \sin(x)\sin(\pi/2) = -\sin x .$$

und

$$\sin(x + \pi/2) = \cos(x)\sin(\pi/2) + \sin(x)\cos(\pi/2) = \cos x .$$

Hieraus folgt wiederum

$$\cos(x + \pi) = \cos(x + \pi/2)\cos(\pi/2) - \sin(x + \pi/2)\sin(\pi/2) = -\cos x .$$

Die weiteren Behauptungen ergeben sich in ähnlicher Weise ([Ü]). □

Satz 12.10 *Es gilt*

1. $\sin$ *ist streng monoton wachsend in* $[-\pi/2, \pi/2]$,
2. $\cos$ *ist streng monoton fallend in* $[0, \pi]$.

Beweis. 1. Aus S. 7.14 folgt für $x_1, x_2 \in \mathbb{R}$

$$\sin x_1 - \sin x_2 = -2 \cdot \cos\left(\frac{x_1 + x_2}{2}\right) \cdot \sin\left(\frac{x_2 - x_1}{2}\right)$$

(beachte dabei: $x_1 = \frac{x_1+x_2}{2} + \frac{x_1-x_2}{2}$ und $x_2 = \frac{x_1+x_2}{2} + \frac{x_2-x_1}{2}$).
Ist $-\pi/2 \leq x_1 < x_2 \leq \pi/2$, so gilt $\cos\left(\frac{x_1+x_2}{2}\right) > 0$ und $\sin\left(\frac{x_2-x_1}{2}\right) > 0$, also $\sin x_1 < \sin x_2$.
2. Analog. □

Bemerkung und Definition 12.11 Die nach S. 12.4, S. 12.10 und B. 10.11.3 auf $[-1, 1] = \sin([-\pi/2, \pi/2])$ existierende (und dort streng monoton wachsende und stetige) Umkehrfunktion von sin heißt arcsin, d. h. $\arcsin : [-1, 1] \to [-\pi/2, \pi/2]$ erfüllt

$$y = \arcsin x \Leftrightarrow x = \sin y \qquad (x \in [-1, 1],\ y \in [-\pi/2, \pi/2])\ .$$

Entsprechend bezeichnet man die auf $[-1, 1]$ existierende (und dort streng monoton fallende und stetige) Umkehrfunktion von cos mit arccos, d. h. $\arccos : [-1, 1] \to [0, \pi]$ erfüllt

$$y = \arccos x \Leftrightarrow x = \cos y \quad (x \in [-1, 1],\ y \in [0, \pi])$$

Bemerkung und Definition 12.12 Wir definieren die Funktionen
$\tan : \mathbb{R} \setminus \{\pi/2 + k\pi : k \in \mathbb{Z}\} \to \mathbb{R}$ und $\cot : \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\} \to \mathbb{R}$ durch

$$\tan x = \frac{\sin x}{\cos x}\ , \qquad \cot x = \frac{\cos x}{\sin x}\ .$$

Dann gilt: tan ist streng monoton wachsend und stetig in $(-\pi/2, \pi/2)$ und cot ist streng monoton fallend und stetig in $(0, \pi)$.
Außerdem gilt $W(\tan_{|(-\pi/2,\pi/2)}) = W(\cot_{|(0,\pi)}) = \mathbb{R}$. Also existieren auf $\mathbb{R}$ die Umkehrfunktionen genannt arctan bzw. arccot.

Zum Schluss dieses Abschnitts wollen wir noch einmal kurz auf den Zusammenhang zwischen Monotonie und Stetigkeit eingehen. Es ist klar, dass i. a. monotone Funktionen nicht überall stetig sind (vgl. etwa B. 12.2.1). Wir werden jedoch sehen, dass die Unstetigkeitsstellen eine besondere Struktur haben und auch nicht “allzu häufig” sind.

Definition 12.13 Es sei $-\infty \leq a < b \leq \infty$ und $f : (a,b) \to \mathbb{R}$. Ferner sei f unstetig an der Stelle $x_0 \in (a,b)$.

1. Die Stelle x_0 heißt *Sprungstelle* (oder auch *Unstetigkeitsstelle 1. Art*), falls
$$f(x_0^+) = \lim_{x \to x_0^+} f(x) \quad \text{und} \quad f(x_0^-) = \lim_{x \to x_0^-} f(x)$$
beide existieren und nicht ∞ oder $-\infty$ sind.

2. Anderenfalls heißt x_0 *Unstetigkeitsstelle 2. Art.*

Beispiel 12.14 1. Es sei
$$f(x) = \operatorname{sign}(x) = \begin{cases} 1, & \text{falls} \quad x > 0 \\ 0, & \text{falls} \quad x = 0 \\ -1, & \text{falls} \quad x < 0 \end{cases} .$$
Dann ist $x_0 = 0$ eine Unstetigkeitsstelle 1. Art (es gilt $f(0^+) = 1$, $f(0^-) = -1$).

2. Es sei
$$f(x) = \begin{cases} 1/x & , \quad x \neq 0 \\ 0 & , \quad x = 0 \end{cases} .$$
Dann ist $x_0 = 0$ eine Unstetigkeitsstelle 2. Art (es gilt $f(0^+) = \infty$).

3. Es sei
$$f(x) := \begin{cases} \sin(1/x) & , \quad x \neq 0 \\ 0 & , \quad x = 0 \end{cases} .$$
Dann ist $x_0 = 0$ eine Unstetigkeitsstelle 2. Art (da $f(0^+)$ nicht existiert ([Ü])).

4. Es sei
$$f(x) = \begin{cases} 1 & , \quad x \in \mathbb{Q} \\ 0 & , \quad x \notin \mathbb{Q} \end{cases} .$$
Dann ist jedes $x_0 \in \mathbb{R}$ eine Unstetigkeitsstelle 2. Art ($f(x_0^+)$ und $f(x_0^-)$ existieren für kein $x_0 \in \mathbb{R}$).

Für monotone Funktionen können keine Unstetigkeitsstellen 2. Art auftreten:

Satz 12.15 *Es sei $-\infty \leq a < b \leq \infty$, und es sei $f : (a,b) \to \mathbb{R}$ monoton (wachsend oder fallend). Dann existieren für alle $x_0 \in (a,b)$ die Grenzwerte $f(x_0^+)$ und $f(x_0^-)$ und es gilt*
$$f(x_0^-) \leq f(x_0) \leq f(x_0^+) \qquad \textit{falls } f \textit{ monoton wächst}$$

und

$$f(x_0^-) \geq f(x_0) \geq f(x_0^+) \qquad \textit{falls } f \textit{ monoton fällt} .$$

Beweis. O. E. sei f monoton wachsend. Ist $x_0 \in (a,b)$, so ist $f(x) \leq f(x_0)$ für alle $x < x_0$, also existiert

$$g := \sup\{f(x) : x < x_0\}$$

und es gilt $g \leq f(x_0)$. Es sei (x_n) eine Folge in (a, x_0) mit $x_n \to x_0$, und es sei $\varepsilon > 0$ gegeben. Dann existiert ein $\xi \in (a, x_0)$ mit $f(\xi) > g - \varepsilon$ und aufgrund der Monotonie gilt

$$f(x) > g - \varepsilon \qquad \text{für alle} \quad x \in (\xi, x_0) .$$

Da $x_n \to x_0 \, (n \to \infty)$ gilt, existiert ein $N \in \mathbb{N}$ mit $x_n \in (\xi, x_0)$ für alle $n \geq N$ und damit auch

$$g \geq f(x_n) \geq g - \varepsilon \qquad \text{für alle} \quad n \geq N .$$

Da $\varepsilon > 0$ beliebig war, ist $\lim\limits_{n\to\infty} f(x_n) = g$ und da (x_n) beliebig (aus (a, x_0)) war, gilt

$$f(x_0^-) = \lim_{x \to x_0^-} f(x) = g(\leq f(x_0)) .$$

Entsprechend zeigt man die Existenz von $f(x_0^+)$ und $f(x_0^+) \geq f(x_0)$. □

Weiter folgt hieraus

Satz 12.16 *Es sei $f : (a,b) \to \mathbb{R}$ monoton. Dann hat f höchstens abzählbar viele Unstetigkeitsstellen (und diese sind 1. Art).*

Beweis. O. E. sei f monoton wachsend. Wir setzen

$$S(f) := \{x \in (a,b) : f \text{ unstetig an } x\} \; (= \{x \in (a,b) : x \text{ Sprungstelle von } f\})$$

Ist $S(f) = \emptyset$, so ist nichts zu zeigen. Es sei also $S(f) \neq \emptyset$. Für $x \in S(f)$ sei

$$I(x) := (f(x^-), f(x^+)) .$$

Dann folgt aus der Monotonie von f für $x_1, x_2 \in S(f)$ mit $x_1 < x_2$:

$$I(x_1) \cap I(x_2) = \emptyset$$

(da $f(x_1^+) \leq f\left(\frac{x_1+x_2}{2}\right) \leq f(x_2^-)$). Nach S. 3.23 ist $I(x) \cap \mathbb{Q} \neq \emptyset$, also können wir ein $\varphi(x) \in I(x) \cap \mathbb{Q}$ wählen. Es gilt dann $\varphi(x_1) \neq \varphi(x_2)$ für $x_1 \neq x_2$. Also ist φ eine bijektive Abbildung von $S(f)$ nach $W(\varphi) \subset \mathbb{Q}$. Da $\mathbb{Q}$ abzählbar ist, ist auch $W(\varphi)$ und damit auch $S(f)$ abzählbar. □

Beispiel 12.17 Es sei $f : (0,1) \to \mathbb{R}$ definiert durch

$$f(x) := \frac{1}{n} \qquad \text{für } x \in \left[\frac{1}{n+1}, \frac{1}{n}\right) \text{ und } n \in \mathbb{N} .$$

Dann ist f monoton wachsend auf $(0,1)$ und jede Stelle $x_n = 1/n \, (n \in \mathbb{N})$ ist eine Sprungstelle. Also hat f abzählbar unendlich viele Sprungstellen.

13 Differenzialrechnung von Funktionen einer reellen Variablen

Wir betrachten etwas allgemeiner als im vorigen Abschnitt Funktionen $f : M \to \mathbb{K}$, wobei $M \subset \mathbb{R}$ ist. Um die feinere Struktur solcher Funktionen untersuchen zu können, brauchen wir einen über die Stetigkeit hinausgehenden "Glattheitsbegriff". Grob gesagt wollen wir Funktionen definieren, deren Graph keine "Ecken" hat; d. h. Funktionen, die Tangenten an den Graph besitzen. Diese Tangenten spiegeln das Veränderungsverhalten der Funktion wider.
Die **Idee** ist folgende: Wir betrachten für zwei Punkte $x_0, x \in M$ Sekanten S_x durch die Punkte $(x_0, f(x_0))$ und $(x, f(x))$ in $\mathbb{R}^2$. Diese Sekanten sind festgelegt durch $(x_0, f(x_0))$ und ihre Steigung

$$\frac{f(x) - f(x_0)}{x - x_0}$$

(Verhältnis der Änderung von f zur Änderung von x). Für $x \to x_0$ wird die Grenzgerade T, falls existent, als Tangente an den Graph im Punkt $(x_0, f(x_0))$ angesehen. Die Steigung von T ist ein Maß für die Änderung der Funktionswerte in der Nähe von x_0. Dies fassen wir in eine exakte Definition.

Definition 13.1 Es seien $M \subset \mathbb{R}$ und $f : M \to \mathbb{K}$. Weiter sei $x_0 \in M \cap H(M)$.

1. f heißt *differenzierbar* an der Stelle x_0, falls der Grenzwert
$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$
existiert. Wir bezeichnen diesen Grenzwert als *Ableitung* von f an der Stelle x_0 und schreiben dafür $f'(x_0)$ (oder auch $\frac{df}{dx}(x_0)$).

2. f heißt *differenzierbar auf* $M_0 \subset M$ falls f in jedem Punkt $x_0 \in M_0$ differenzierbar ist. Die Funktion $f' : M_0 \to \mathbb{K}$ heißt dann *Ableitung* von f (*auf* M_0).

3. *Höhere Ableitungen* von f werden rekursiv definiert: Ist f' differenzierbar auf M_0, so schreiben wir $f^{(2)} := f'' := (f')'$. Ist f'' differenzierbar auf M_0, so ist $f^{(3)} := f''' = (f'')'$ u. s. w. Wir schreiben für die höheren Ableitungen stets $f^{(n)}$.

Beispiel 13.2 1. Ist $f(x) \equiv c$ $(x \in \mathbb{R})$ für eine Konstante $c \in \mathbb{K}$, so ist f differenzierbar auf $\mathbb{R}$ und es gilt
$$f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x} = 0 \qquad (x \in \mathbb{R}) .$$
2. Für $f(x) := cx$ $(x \in \mathbb{R})$, wobei $c \in \mathbb{K}$ eine Konstante ist, gilt
$$f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x} = \lim_{y \to x} \frac{c(y - x)}{y - x} = c \qquad (x \in \mathbb{R}) .$$

3. Für $f(x) = x^2$ $(x \in \mathbb{R})$ gilt

$$f'(x) = \lim_{y \to x} \frac{y^2 - x^2}{y - x} = \lim_{y \to x} \frac{(y+x)(y-x)}{y-x} = \lim_{y \to x} y + x = 2x \qquad (x \in \mathbb{R}) .$$

4. Die Funktion $f(x) := |x|$ $(x \in \mathbb{R})$ ist nicht differenzierbar an der Stelle $x_0 = 0$, denn ist (x_n) eine Folge mit $x_n \to 0$ und $x_n > 0$ $(n \in \mathbb{N})$, so gilt

$$\frac{f(x_n) - f(0)}{x_n - 0} = 1 \to 1 \qquad (n \to \infty)$$

und ist (x_n) eine Folge mit $x_n \to 0$ und $x_n < 0$, so gilt

$$\frac{f(x_n) - f(0)}{x_n - 0} = \frac{-x_n}{x_n} = -1 \to -1 \qquad (n \to \infty) .$$

Also existiert

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

nicht!
Dieses Beispiel zeigt, dass die Stetigkeit an einer Stelle i. a. **nicht** die Differenzierbarkeit an dieser Stelle impliziert. Es gilt jedoch umgekehrt:

Satz 13.3 *Ist $f : M \to \mathbb{K}$ differenzierbar auf der Menge $M_0 \subset M$, so ist f stetig auf M_0.*

Beweis. Es sei $x_0 \in M_0$ und es sei (x_n) eine Folge in M mit $x_n \to x_0$ $(n \to \infty)$ und $x_n \neq x_0$ $(n \in \mathbb{N})$. Dann gilt

$$f(x_n) - f(x_0) = \frac{f(x_n) - f(x_0)}{x_n - x_0}(x_n - x_0) \to f'(x_0)(x_0 - x_0) = 0 \qquad (n \to \infty) ,$$

d. h. $f(x_n) \to f(x_0)$ $(n \to \infty)$. Also ist $\lim_{x \to x_0} f(x) = f(x_0)$. □

Wir stellen wichtige Rechenregeln für die Differenziation zusammen.

Satz 13.4 *(Summen-, Produkt- und Quotientenregel)*
Es sei $M \subset \mathbb{R}$ und es seien $f, g : M \to \mathbb{K}$ differenzierbar an der Stelle $x_0 \in M$. Dann gilt

1. *Für beliebige $\alpha, \beta \in \mathbb{K}$ ist $\alpha f + \beta g$ differenzierbar an x_0, und es gilt*

$$(\alpha f + \beta g)'(x_0) = \alpha \cdot f'(x_0) + \beta \cdot g'(x_0) .$$

2. *$f \cdot g$ ist differenzierbar an x_0, und es gilt*

$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0) .$$

3. Ist $g(x_0) \neq 0$, so ist auch f/g differenzierbar an x_0, und es gilt

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)} \ .$$

Beweis. Wir beschränken uns auf den Beweis von 3. Der Beweis von 1. und 2. verläuft ähnlich.
Da g differenzierbar an x_0 ist, ist g auch stetig an x_0 (S. 13.3). Es sei (x_n) eine Folge in M mit $x_n \to x_0$ $(n \to \infty)$ und $x_n \neq x_0$ $(n \in \mathbb{N})$. Da $g(x_n) \to g(x_0) \neq 0$ gilt, ist auch $g(x_n) \neq 0$ für $n \geq N$. Dann gilt für $n \geq N$

$$\begin{aligned}
\frac{(f/g)(x_n) - (f/g)(x_0)}{x_n - x_0} &= \frac{1}{g(x_n)g(x_0)}\left[\frac{f(x_n)g(x_0) - f(x_0)g(x_n)}{x_n - x_0}\right] \\
&= \frac{1}{g(x_n)g(x_0)}\left[\frac{f(x_n) - f(x_0)}{x_n - x_0} \cdot g(x_0) - f(x_0) \cdot \frac{g(x_n) - g(x_0)}{x_n - x_0}\right] \\
&\to \frac{1}{g^2(x_0)}\left[f'(x_0)g(x_0) - f(x_0)g'(x_0)\right] \qquad (n \to \infty) \ .
\end{aligned}$$

Damit ist 3. bewiesen. □

Beispiel 13.5 1. Es sei $n \in \mathbb{N}$. Dann ist $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^n$ differenzierbar auf $\mathbb{R}$ mit

$$f'(x) = n \cdot x^{n-1} \qquad (x \in \mathbb{R})$$

(Denn: Für $n = 1$ gilt die Behauptung nach B. 13.2.2 (mit $0^0 := 1$).
Es gelte $(x^k)' = k \cdot x^{k-1}$ für ein $k \in \mathbb{N}$. Dann gilt mit der Produktregel (S. 13.4.2)

$$(x^{k+1})' = (xx^k)' = 1 \cdot x^k + xkx^{k-1} = (k+1)x^k \qquad (x \in \mathbb{R}) \ .)$$

2. Ist $P : \mathbb{R} \to \mathbb{R}$ ein Polynom, d. h. $P(x) = \sum_{\nu=0}^{n} a_\nu x^\nu$ für gewisse $a_0, \ldots, a_n \in \mathbb{R}$, so gilt

$$P'(x) = \sum_{\nu=1}^{n} \nu \cdot a_\nu x^{\nu-1} \qquad (x \in \mathbb{R})$$

(Folgt aus 1. und mehrfacher Anwendung von S. 13.4 1.)

Bevor wir uns mit der Ableitung von Hintereinanderausführungen beschäftigen, wollen wir eine Charakterisierung der Differenzierbarkeit formulieren.

Satz 13.6 *(Zerlegungsformel)*
Es seien $M \subset \mathbb{R}$ und $f : M \to \mathbb{K}$. Ferner sei $x_0 \in M \cap H(M)$. Dann sind folgende Aussagen äquivalent:

a) f ist differenzierbar an x_0.

b) Es existieren ein $c \in \mathbb{K}$ und eine in x_0 stetige Funktion $\varepsilon : M \to \mathbb{K}$ so, dass

$$f(x) = f(x_0) + c \cdot (x - x_0) + \varepsilon(x)|x - x_0| \qquad (x \in M)$$

und $\varepsilon(x_0) = 0$.

Außerdem ist in diesem Fall $f'(x_0) = c$.

Beweis. a) $\Rightarrow$ b): Wir setzen für $x \in M$

$$\varepsilon(x) := \begin{cases} \operatorname{sign}(x - x_0) \left\{ \dfrac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right\} & , \text{ falls } x \neq x_0 \\ 0 & , \text{ falls } x = x_0 \end{cases} .$$

Dann ist (da $x - x_0 = \operatorname{sign}(x - x_0) \cdot |x - x_0|$)

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \varepsilon(x)|x - x_0|$$

(also $c = f'(x_0)$) und es gilt

$$\varepsilon(x) \to 0 = \varepsilon(x_0) \qquad (x \to x_0)$$

auf Grund der Differenzierbarkeit von f an x_0.

b) $\Rightarrow$ a): Es gilt für $x \in M,\ x \neq x_0$

$$\varepsilon(x) \cdot \operatorname{sign}(x - x_0) = \frac{f(x) - f(x_0)}{x - x_0} - c\,.$$

Aus $\varepsilon(x) \to 0$ $(x \to x_0)$ folgt

$$\frac{f(x) - f(x_0)}{x - x_0} \to c \qquad (x \to x_0)$$

also ist f differenzierbar an x_0 mit $f'(x_0) = c$. $\square$

Satz 13.7 *(Kettenregel)*
Es sei $M \subset \mathbb{R}$ und es sei $h : M \to \mathbb{R}$. Ferner sei $g : L \to \mathbb{K}$ mit $L \supset W(h)$. Ist h differenzierbar an $x_0 \in M$ und ist g differenzierbar an $y_0 = h(x_0) \in L$, so ist $f = g \circ h$ differenzierbar an x_0 mit

$$f'(x_0) = (g \circ h)'(x_0) = g'(y_0) \cdot h'(x_0) = g'\big(h(x_0)\big)h'(x_0)\,.$$

Beweis. Nach S. 13.6 existieren Funktionen $\varepsilon_h : M \to \mathbb{R}$ und $\varepsilon_g : L \to \mathbb{K}$ mit

$$\begin{aligned} h(x) &= h(x_0) + h'(x_0)(x - x_0) + \varepsilon_h(x)|x - x_0| \\ g(y) &= g(y_0) + g'(y_0)(y - y_0) + \varepsilon_g(y)|y - y_0| \end{aligned}$$

und so, dass ε_h stetig an x_0 und ε_g stetig an $y_0 = h(x_0)$ mit $\varepsilon_h(x_0) = \varepsilon_g(y_0) = 0$ ist. Hieraus folgt für $x \in M$

$$\begin{aligned} f(x) &= g\Big(h(x)\Big) = \\ &= g\Big(h(x_0)\Big) + g'\Big(h(x_0)\Big)\Big(h(x) - h(x_0)\Big) + \varepsilon_g\Big(h(x)\Big)|h(x) - h(x_0)| \\ &= (g \circ h)(x_0) + g'\Big(h(x_0)\Big)\Big[h'(x_0)(x - x_0) + \varepsilon_h(x)|x - x_0|\Big] \\ &+ \varepsilon_g\Big(h(x)\Big)\Big|h'(x_0)(x - x_0) + \varepsilon_h(x)|x - x_0|\Big| \\ &= (g \circ h)(x_0) + g'\Big(h(x_0)\Big)h'(x_0)(x - x_0) \\ &+ |x - x_0|\Big[g'\Big(h(x_0)\Big)\varepsilon_h(x) + \varepsilon_g\Big(h(x)\Big)\Big|h'(x_0) + \varepsilon_h(x)\Big|\Big] \end{aligned}$$

Da h nach S. 13.3 stetig an x_0 ist, ist $\varepsilon_f : M \to \mathbb{K}$ mit

$$\varepsilon_f(x) := g'(y_0)\varepsilon_h(x) + \varepsilon_g\Big(h(x)\Big)|h'(x_0) + \varepsilon_h(x)|$$

stetig an x_0 mit $\varepsilon_f(x_0) = 0$. Aus S. 13.6 ergibt sich die Differenzierbarkeit von $f = g \circ h$ an x_0 und

$$f'(x_0) = g'\Big(h(x_0)\Big)h'(x_0) \; .$$

□

Beispiel 13.8 Es sei $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = (x^3 + 2x + 1)^5 \; .$$

Dann gilt $f = g \circ h$ mit $h(x) = x^3 + 2x + 1$ und $g(y) = y^5$. Also gilt

$$\begin{aligned} f'(x) = g'(h(x))h'(x) &= 5(x^3 + 2x + 1)^4 \cdot (x^3 + 2x + 1)' = \\ &= 5(x^3 + 2x + 1)^4(3x^2 + 2) \; . \end{aligned}$$

Satz 13.9 *(Umkehrregel)*
Es sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ sei stetig und streng monoton (wachsend oder fallend) auf I. Ist f differenzierbar an $x_0 \in I$ mit $f'(x_0) \neq 0$, so ist die Umkehrfunktion $f^{-1} : W(f) \to I$ differenzierbar an $y_0 = f(x_0)$ mit

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} \; .$$

Beweis. Es sei (y_n) eine Folge in $W(f) = D(f^{-1})$ mit $y_n \to y_0$ und $y_n \neq y_0$ $(n \in \mathbb{N})$ (man beachte: $W(f)$ ist ein Intervall nach S. 11.4). Für $x_n := f^{-1}(y_n)$ gilt dann $x_n \neq x_0$ $(n \in \mathbb{N})$ und $x_n \to x_0$ $(n \to \infty)$ nach S. 11.4. Also erhalten wir

$$\frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} = \frac{x_n - x_0}{f(x_n) - f(x_0)} \to \frac{1}{f'(x_0)} \qquad (n \to \infty) \; .$$

□

Nun wenden wir uns der Differentiation elementarer Funktionen zu. Die Basis dazu bildet der folgende Satz

Satz 13.10 *Es sei $c \in \mathbb{C}$ fest. Dann gilt*

$$(e^{cx})' = ce^{cx} \qquad (x \in \mathbb{R}) \; .$$

Beweis. Nach B. 10.6.1 gilt

$$\lim_{z \to 0} \frac{e^z - 1}{z} = 1 \; .$$

Hieraus folgt für $c \neq 0$

$$\lim_{x \to 0} \frac{e^{cx} - 1}{cx} = 1 \; ,$$

also

$$\lim_{x \to 0} \frac{e^{cx} - 1}{x} = c \; .$$

Für beliebiges $x_0 \in \mathbb{R}$ folgt

$$\lim_{x \to x_0} \frac{e^{cx} - e^{cx_0}}{x - x_0} = e^{cx_0} \lim_{x \to x_0} \frac{e^{c(x-x_0)} - 1}{x - x_0} = ce^{cx_0} \; .$$

d. h. die Behauptung gilt für $c \neq 0$. Ist $c = 0$, so ist $e^{cx} \equiv 1$, und die Behauptung folgt auch. □

Folgerung 13.11 1. Ist $a > 0$ fest, so gilt

$$(a^x)' = a^x \ln a \qquad (x \in \mathbb{R}) \; .$$

(Denn: $(a^x)' = (e^{x \cdot \ln a})' = \ln a \cdot e^{x \ln a} = \ln a \cdot a^x$.)

2. Es gilt

$$(\cos x)' = -\sin x \qquad \text{und} \qquad (\sin x)' = \cos x \qquad (x \in \mathbb{R}) \, .$$

(Denn:

$$(\cos x)' = \frac{1}{2}(e^{ix} + e^{-ix})' = \frac{1}{2}(ie^{ix} - ie^{-ix}) = -\frac{1}{2i}(e^{ix} - e^{-ix}) = -\sin x$$

und für $\sin x$ entsprechend.)

3. Für festes $a > 0, a \neq 1$, gilt

$$(\log_a x)' = \frac{1}{x} \cdot \frac{1}{\ln a} \qquad (x > 0)\ .$$

(Denn: Es sei $f(y) = a^y$ $(y \in \mathbb{R})$. Dann ist $f^{-1}(x) = \log_a x$ $(x > 0)$. Also gilt nach der Umkehrregel und 1. mit $y = \log_a x$

$$(f^{-1})'(x) = \frac{1}{f'(y)} = \frac{1}{a^y \ln a} = \frac{1}{x} \cdot \frac{1}{\ln a} \qquad (x > 0)\ .)$$

4. Für $\alpha \in \mathbb{R}$ fest gilt

$$(x^\alpha)' = \alpha x^{\alpha-1} \qquad (x > 0)$$

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

5. Es gilt

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}\ , \qquad (\arccos x)' = -\frac{1}{\sqrt{1-x^2}} \qquad (x \in (-1,1))$$

(Denn: Mit der Umkehrregel und 2. gilt für $x \in (-1,1)$ mit $y = \arcsin x$

$$(\arcsin x)' = \frac{1}{(\sin y)'} = \frac{1}{\cos y} = \frac{1}{\sqrt{1-\sin^2 y}} = \frac{1}{\sqrt{1-x^2}} \qquad (x \in (-1,1))\ .$$

Entsprechend für $\arccos x$.)

6. Es gilt

$$\begin{aligned}(\tan x)' &= \frac{1}{\cos^2 x} = 1 + \tan^2 x \qquad (x \neq \pi/2 + k\pi,\ k \in \mathbb{Z})\\ (\cot x)' &= -\frac{1}{\sin^2 x} = -1 - \cot^2 x \qquad (x \neq k\pi,\ k \in \mathbb{Z})\end{aligned}$$

und

$$\begin{aligned}(\arctan x)' &= \frac{1}{1+x^2} \qquad (x \in \mathbb{R})\\ (\operatorname{arccot} x)' &= -\frac{1}{1+x^2} \qquad (x \in \mathbb{R})\ .\end{aligned}$$

([Ü])

Wir beschäftigen uns nun mit der geometrischen Bedeutung der Ableitung. Dazu definieren wir zunächst, was wir unter lokalen Extrema verstehen.

Definition 13.12 Es sei (X, d) ein metrischer Raum,$M \subset X$, und es sei $f : M \to \mathbb{R}$. Ferner sei $x_0 \in M$.

1. Man sagt, f habe an x_0 ein *lokales* (oder auch *relatives*) *Maximum*, falls ein $\delta > 0$ existiert mit
$$f(x) \leq f(x_0) \qquad \text{für alle} \quad x \in M \cap U_\delta(x_0)$$
2. Man sagt, f habe an x_0 ein *lokales* (oder auch *relatives*) *Minimum*, falls ein $\delta > 0$ existiert mit
$$f(x) \geq f(x_0) \qquad \text{für alle} \quad x \in M \cap U_\delta(x_0)\ .$$

Die Stellen x_0, an denen lokale Maxima oder Minima vorliegen, bezeichnet man auch als *Extremstellen* von f.

Es gilt folgendes einfache notwendige Kriterium für Extrema differenzierbarer Funktionen.

Satz 13.13 *Es sei $I \subset \mathbb{R}$ ein* **offenes** *Intervall und die Funktion $f : I \to \mathbb{R}$ sei differenzierbar auf I. Dann gilt: Ist x_0 eine Extremstelle von f, so ist*
$$f'(x_0) = 0\ .$$
(Punkte x_0 mit $f'(x_0) = 0$ nennt man auch kritische Punkte von f.)

Beweis. O. E. liege an x_0 ein lokales Maximum vor. Dann existiert ein $\delta > 0$ mit $f(x) \leq f(x_0)$ für alle $x \in U_\delta(x_0)$. Da f differenzierbar an x_0 ist, gilt
$$f'(x_0) = \lim_{n\to\infty} \frac{f(x_0 + \frac{\delta}{2n}) - f(x_0)}{\delta/2n} \leq 0$$
und
$$f'(x_0) = \lim_{n\to\infty} \frac{f(x_0 - \frac{\delta}{2n}) - f(x_0)}{-\delta/2n} \geq 0\ ,$$
d. h. $f'(x_0) = 0$. □

Bemerkung 13.14 1. S. 13.13 liefert lediglich eine **notwendige** Bedingung für das Auftreten lokaler Maxima oder Minima. So hat etwa $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^3$ an $x_0 = 0$ einen kritischen Punkt, aber an $x_0 = 0$ liegt kein lokales Extremum vor.
2. Für nicht offene Intervalle I ist die Aussage von S. 13.13 i. a. falsch. So hat etwa $f : [-1, 1] \to \mathbb{R}$ mit $f(x) = x^2$ an ± 1 (sogar globale) Maxima, aber die Ableitung ist dort $\neq 0$.

Über das Auftreten kritischer Punkte gibt foldender Satz Auskunft

Satz 13.15 *(Rolle)*
Es sei $I = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall, und es sei $f : I \to \mathbb{R}$ stetig auf I und differenzierbar auf $I^0 = (a, b)$. Gilt $f(a) = f(b)$ so existiert ein $\xi \in (a, b)$ mit $f'(\xi) = 0$.

Beweis. Ist $f(x) \equiv f(a)$ auf $[a, b]$ so ist jedes $x \in (a, b)$ eine Extremstelle, d. h. $f'(x) \equiv 0$ auf (a, b) nach S. 13.13.
Es sei also $f \not\equiv f(a)(= f(b))$.
Nach S. 11.12 hat f absolute Maxima und Minima auf $[a, b]$. Mindestens eines davon liegt in (a, b). Nach S. 13.13 liegt dort ein kritischer Punkt vor. □

Als Folgerung erhalten wir

Satz 13.16 *(Mittelwertsatz)*
Es sei $I = [a, b]$ ein kompaktes Intervall, und $f : I \to \mathbb{R}$ sei stetig auf I und differenzierbar auf $I^0 = (a, b)$. Dann existiert ein $\xi \in (a, b)$ mit

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

Beweis. Man wende den Satz von Rolle auf die Funktion $\varphi : I \to \mathbb{R}$ mit

$$\varphi(x) := f(x) - \frac{f(b) - f(a)}{b - a}(x - a) \qquad (x \in I)$$

an! (Es gilt $\varphi(b) = \varphi(a) = f(a)$ und $\varphi'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$.) □

Wir ziehen verschiedene Folgerungen aus dem Mittelwertsatz

Satz 13.17 *Es sei $I = [a, b]$ und $f : I \to \mathbb{R}$ sei stetig auf I und differenzierbar auf (a, b). Dann ist f genau dann konstant, wenn $f'(x) \equiv 0$ auf (a, b) gilt.*

Beweis. "$\Rightarrow$" Ist $f(x) \equiv c$ auf I, so gilt natürlich $f'(x) \equiv 0$ auf I.

"$\Leftarrow$ Wir zeigen: $f(x) \equiv f(a)$ auf $(a, b]$. Dazu sei $x \in (a, b]$ gegeben. Dann erfüllt f (bzw. $f_{|[a,x]}$) die Voraussetzungen des Mittelwertsatzes auf $[a, x]$, d. h. es existiert ein $\xi \in (a, x)$ mit

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

d. h. $f(x) = f(a)$. □

Der folgende Satz gibt ein **hinreichendes** Kriterien für Extremstellen.

Satz 13.18 *Die Funktion $f : (a, b) \to \mathbb{R}$ sei differenzierbar auf (a, b), und es sei $x_0 \in (a, b)$ ein kritischer Punkt von f (d. h. $f'(x_0) = 0$)*

1. Existiert ein $\delta > 0$, so dass

$$f'(x) \begin{cases} \leq 0 & \text{für} \quad x_0 < x < x_0 + \delta \\ \geq 0 & \text{für} \quad x_0 - \delta < x < x_0 \end{cases} ,$$

so hat f an x_0 ein lokales Maximun.

2. Existiert ein $\delta > 0$, so dass

$$f'(x) \begin{cases} \geq 0 & \text{für} \quad x_0 < x < x_0 + \delta \\ \leq 0 & \text{für} \quad x_0 - \delta < x < x_0 \end{cases} ,$$

So hat f an x_0 ein lokales Minimum.

Beim Beweis verwenden wir:

Satz 13.19 *Es sei $I = [a, b]$ und $f : I \to \mathbb{R}$ sei stetig auf I und differenzierbar auf $I^0 = (a, b)$.*
1. Ist $f'(x) > 0$ für alle $x \in (a, b)$, so ist f streng monoton wachsend auf $[a, b]$.

2. Ist $f'(x) < 0$ für alle $x \in (a, b)$, so ist f streng monoton fallend auf $[a, b]$.

Ensprechende Aussagen gelten mit "$\geq$" bzw. "$\leq$" und ohne "streng".

Beweis. 1. Es seien $x_1, x_2 \in [a, b]$ mit $x_1 < x_2$. Durch Anwendung des Mittelwertsatzes für das Intervall $[x_1, x_2]$ ergibt sich mit einem $\xi \in (x_1, x_2)$

$$f(x_2) - f(x_1) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x_2 - x_1) = f'(\xi)(x_2 - x_1) > 0 ,$$

also $f(x_1) < f(x_2)$.
2. Entsprechend. $\square$

Beweis. (S. 13.18)
1. Nach S. 13.19 ist f monoton fallend auf $[x_0, x_0 + \delta)$ und monoton wachsend auf $(x_0 - \delta, x_0]$. Damit hat f an x_0 ein lokales Maximum.
2. Analog. $\square$

Beispiel 13.20 1. Es sei $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = 2x^3 + 3x^2 - 1 \ .$$

Dann gilt

$$f'(x) = 6x(x+1) \begin{cases} > 0 & , \quad x \in (-\infty, -1) \\ < 0 & , \quad x \in (-1, 0) \\ > 0 & , \quad x \in (0, \infty) \end{cases}$$

Also ist f streng monoton $\left\{ \begin{array}{l} \text{wachsend} \\ \text{fallend} \\ \text{wachsend} \end{array} \right\}$ auf $\left\{ \begin{array}{l} (-\infty, -1] \\ [-1, 0] \\ [0, \infty) \end{array} \right.$.

(Für $(-\infty, -1]$ und $[0, \infty)$ wende man S. 13.19 auf die kompakten Intervalle $[-n, -1]$ und $[0, n]$ $(n \in \mathbb{N})$ an.)
Weiter ist $f'(0) = f'(-1) = 0$. Anwendung von S. 13.18 zeigt, dass an 0 ein lokales Minimum und and -1 ein lokales Maximum vorliegt.

2. Es sei $f : (0, \infty) \to \mathbb{R}$ mit

$$f(x) := \left(1 + \frac{1}{x}\right)^x \qquad (x > 0) \ .$$

Dann ist f monoton wachsend.
(Denn: Es genügt zu zeigen: $g(x) := \ln f(x) = x \cdot \ln(1 + 1/x)$ ist monoton wachsend. Es gilt

$$g'(x) = \ln(1 + 1/x) + x \cdot \frac{1}{1 + 1/x} \cdot \frac{-1}{x^2} = \ln(1 + 1/x) - \frac{1}{1 + x} \geq 0$$

nach S. 5.16.1.)

Zum Abschluss dieses Abschnitts beschäftigen wir uns mit einer eleganten Methode um gewisse Grenzwerte auszurechnen. Dazu beweisen wir zunächst eine Verallgemeinerung des Mittelwertsatzes.

Satz 13.21 *(erweiterter Mittelwertsatz)*
Es sei $I = [a, b]$ und die Funktionen $f, g : I \to \mathbb{R}$ seien stetig auf I und differenzierbar auf $I^0 = (a, b)$. Ferner sei $g'(x) \neq 0$ für alle $x \in (a, b)$. Dann ist $g(a) \neq g(b)$ und es existiert ein $\xi \in (a, b)$ mit

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

Beweis. Nach S. 13.16 existiert ein $\eta \in (a, b)$ mit

$$g(b) - g(a) = \frac{g(b) - g(a)}{b - a}(b - a) = g'(\eta)(b - a) \neq 0$$

d. h. es ist $g(a) \neq g(b)$.
Wir betrachten die auf $[a, b]$ stetige und auf (a, b) differenzierbare Funktion φ mit

$$\varphi(x) := f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}[g(x) - g(a)] \, .$$

Dann ist $\varphi(a) = \varphi(b) = f(a)$ und

$$\varphi'(x) := f'(x) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(x) \qquad (x \in (a, b)) \, .$$

Nach dem Satz von Rolle existiert ein $\xi \in (a, b)$ mit $\varphi'(\xi) = 0$, d. h.

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

Da $g'(\xi) \neq 0$ ist, folgt die Behauptung. □

Damit beweisen wir

Satz 13.22 *(Regeln von de l'Hospital)*
Es sei $-\infty \leq a < b \leq \infty$ und die Funktion $f, g : (a, b) \to \mathbb{R}$ seien differenzierbar auf (a, b) mit $g'(x) \neq 0$ für alle $x \in (a, b)$. Ferner gelte

1. $\lim\limits_{x \to a^+} f(x) = \lim\limits_{x \to a^+} g(x) = 0$ *(Fall „$\frac{0}{0}$")*

oder

2. $\lim\limits_{x \to a^+} g(x) = \infty$ *(oder $-\infty$).*

Dann gilt: Existiert $\lim\limits_{x \to a^+} \frac{f'(x)}{g'(x)} \in \mathbb{R} \cup \{\pm\infty\}$, so existiert auch $\lim\limits_{x \to a^+} \frac{f(x)}{g(x)}$ und es ist

$$\lim_{x \to a^+} \frac{f(x)}{g(x)} = \lim_{x \to a^+} \frac{f'(x)}{g'(x)} \, .$$

Eine entsprechende Aussage gilt für $x \to b^-$.

Beweis. 1. Es gelte 1. und es sei $a > -\infty$. Durch $g(a) := f(a) := 0$ können f und g stetig nach $[a,b)$ fortgesetzt werden. Es sei (x_n) eine Folge in (a,b) mit $x_n \to a(n \to \infty)$. . Nach S. 13.21 existiert ein $\xi_n \in (a, x_n)$ mit

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

Da $\lim\limits_{x\to a^+} \frac{f'(x)}{g'(x)}$ existiert, ist

$$\lim_{n\to\infty} \frac{f(x_n)}{g(x_n)} = \lim_{n\to\infty} \frac{f'(\xi_n)}{g'(\xi_n)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)} \,.$$

Da (x_n) beliebig war, folgt die Behauptung.
Der Beweis für $x \to b^-$, $b < \infty$ verläuft analog. Die Fälle $a = -\infty$ und $b = \infty$ können leicht auf die obigen Fälle zurückgeführt werden ([Ü]).
2. Auf den technisch etwas aufwändigeren Beweis unter der Voraussetung 2. verzichten wir. □

Beispiel 13.23 1. Für $f(x) = \sqrt{x} - 1$ und $g(x) = \sqrt{x-1}$ $(x > 1)$ gilt

$$\lim_{x\to 1^+} f(x) = \lim_{x\to 1^+} g(x) = 0$$

sowie

$$\lim_{x\to 1^+} \frac{f'(x)}{g'(x)} = \lim_{x\to 1^+} \frac{\frac{1}{2\sqrt{x}}}{\frac{1}{2\sqrt{x-1}}} = \lim_{x\to 1^+} \frac{\sqrt{x-1}}{\sqrt{x}} = 0 \,.$$

Also ist auch

$$\lim_{x\to^+} \frac{f(x)}{g(x)} = \lim_{x\to 1} \frac{\sqrt{x}-1}{\sqrt{x-1}} = 0 \,.$$

2. Für $f(x) = 1 - \cos x$ und $g(x) = x^2$ gilt

$$1 - \cos x \to 0 \qquad \text{und} \qquad x^2 \to 0 \qquad (x \to 0)$$

sowie

$$\frac{(1-\cos x)'}{(x^2)'} = \frac{\sin x}{2x} \qquad (x \neq 0) \,.$$

Hier liegt wieder der Fall "$\frac{0}{0}$" vor, d. h.

$$\sin x \to 0 \qquad \text{und} \qquad 2x \to 0 \qquad (x \to 0) \,.$$

Weiter gilt

$$\frac{(\sin x)'}{(2x)'} = \frac{\cos x}{2} \to \frac{1}{2} \qquad (x \to 0) \,.$$

Also ist

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

3. Es gilt für jedes feste $\alpha > 0$:

$$\lim_{x\to\infty}\frac{\ln x}{x^\alpha}=0\ ,$$

d. h. die Logarithmusfunktion wächst langsamer als jede (noch so kleine) positive Potenz von x.

(Denn: Es gilt $\lim\limits_{x\to\infty} x^\alpha=\infty$ und

$$\lim_{x\to\infty}\frac{(\ln x)'}{(x^\alpha)'}=\lim_{x\to\infty}\frac{x^{-1}}{\alpha x^{\alpha-1}}=\lim_{x\to\infty}\frac{1}{\alpha x^\alpha}=0$$

also folgt die Behauptung mit S. 13.22.2.)

4. Für jedes $a > 1$ und jedes $n \in \mathbb{N}$ ist

$$\lim_{x\to\infty}\frac{a^x}{x^n}=\infty\ ,$$

d. h. geometrisches Wachstum ist schneller als das Wachstum jeder (noch so großen) Potenz von x.

(Denn: Es gilt $\lim\limits_{x\to\infty} x^n=\infty$ und damit ergibt sich durch wiederholte Anwendung von S. 13.22.2

$$\lim_{x\to\infty}\frac{a^x}{x^n}=\lim_{x\to\infty}\frac{a^x\ln a}{nx^{n-1}}=\lim_{x\to\infty}\frac{a^x(\ln a)^2}{n(n-1)x^{n-2}}=\ldots=\lim_{x\to\infty}\frac{a^x(\ln a)^n}{n!}=\infty\ .)$$

14 Der Taylorsche Satz und Anwendungen

Ein wesentliches Anliegen in der Analysis besteht darin, “komplizierte” Funktionen f durch andere, möglichst einfache Funktionen wie etwa Polynome P (die auf einem Computer relativ einfach auszuwerten sind) anzunähern, d. h. man will für f eine Darstellung der Form

$$f(x) = P(x) + R(x)$$

wobei R eine (möglichst kleine) Fehlerfunktion ist. Eine sehr einfache solche Darstellung kennen wir bereits:
Ist f differenzierbar an der Stelle x_0, so gilt nach der Zerlegungsformel

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + \varepsilon(x)|x - x_0| \\ &= T_1(x) + R(x) \end{aligned}$$

wobei $T_1(x) := f(x_0) + f'(x_0)(x - x_0)$ ein Polynom vom Grad ≤ 1 und $R(x) := \varepsilon(x)|x - x_0|$ für $x \to x_0$ schneller gegen 0 geht als $|x - x_0|$. Wie könnten sinnvolle Polynome höheren Grades aussehen? Betrachten wir zwei Beispiele:

Beispiel 14.1 1. Es sei $f : (-1, 1) \to \mathbb{R}$ mit $f(x) = \frac{1}{1-x}$ für $|x| < 1$. Dann gilt

$$f(x) = \sum_{\nu=0}^{\infty} x^\nu = \sum_{\nu=0}^{n} x^\nu + \sum_{\nu=n+1}^{\infty} x^\nu = \sum_{\nu=0}^{n} x^\nu + \frac{x^{n+1}}{1-x} .$$

für $|x| < 1$. Betrachten wir die Ableitungen von f, so gilt

$$\begin{aligned} f^{(0)}(x)\Big(:= f(x)\Big) &= \frac{1}{1-x} &&, \quad f^{(0)}(0) = 1 = 0\,! \\ f'(x) &= \frac{1}{(1-x)^2} &&, \quad f'(0) = 1 = 1\,! \\ f''(x) &= \frac{1 \cdot 2}{(1-x)^3} &&, \quad f''(0) = 1 \cdot 2 = 2\,! \\ &\vdots && \qquad \vdots \\ f^{(\nu)}(x) &= \frac{1 \cdot 2 \cdots \nu}{(1-x)^{\nu+1}} &&, \quad f^{(\nu)}(0) = 1 \cdot 2 \cdots \nu = \nu\,! \end{aligned}$$

Also ist

$$\begin{aligned} f(x) &= \sum_{\nu=0}^{n} x^\nu + \frac{x^{n+1}}{1-x} = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(0)}{\nu!} x^\nu + \frac{x^{n+1}}{1-x} \\ &= T_n(x) + R_n(x) \end{aligned}$$

mit

$$T_n(x) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(0)}{\nu!} x^\nu , \qquad R_n(x) = \frac{x^{n+1}}{1-x} .$$

Hierbei ist T_n ein Polynom vom Grad $\leq n$, dessen Koeffizienten durch die Ableitungen von f an x_0 ausgedrückt werden können.

2. Es sei $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^x$. Dann ist

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

und $f^{(\nu)}(x) = e^x$ ($\nu \in \mathbb{N}_0$), d . h., $f^{(\nu)}(0) = 1$ ($\nu \in \mathbb{N}_0$). Also ist auch hier

$$f(x) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(0)}{\nu!} x^\nu + R(x) = T_n(x) + R(x) ,$$

wobei jetzt $R(x) = \sum\limits_{\nu=n+1}^{\infty} \frac{x^\nu}{\nu!}$.

Definition 14.2 1. Es sei $M \subset \mathbb{R}$ und $f : M \to \mathbb{R}$. Existieren für ein $x_0 \in M$ und ein $n \in \mathbb{N}_0$ die Ableitungen $f^{(\nu)}(x_0)$ für $\nu = 0, \ldots, n$ (wobei $f^{(0)}(x_0) := f(x_0)$ gesetzt ist), so heißt das Polynom T_n vom Grad $\leq n$ mit

$$T_n(x) := T_{n,x_0}(x) := \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x_0)}{\nu!} (x - x_0)^\nu$$

n-tes Taylor-Polynom von f (*bzgl. der Entwicklungsmitte* x_0). Die Koeffizienten $f^{(\nu)}(x_0)/\nu!$ heißen *Taylor-Koeffizienten* von f (bzgl. x_0).
2. Ist I ein Intervall, so setzen wir

$$C^n(I) := \{f : I \to \mathbb{R} : f^{(n)} \text{ existiert und ist stetig auf } I\}$$

(d. h. $C^n(I)$ ist die Menge aller n-mal stetig differenzierbaren Funktionen auf I). Weiter setzen wir

$$C(I) := C^0(I) := \{f : I \to \mathbb{R} : f \text{ stetig auf } I\} .$$

Darüber hinaus schreiben wir im folgenden für zwei Punkte $x_1, x_2 \in \mathbb{R}, x_1 \neq x_2$

$$I(x_1, x_2) := \begin{cases} (x_1, x_2) & , \text{ falls } \ x_1 < x_2 \\ (x_2, x_1) & , \text{ falls } \ x_2 < x_1 \end{cases}$$

und $I[x_1, x_2]$ für die entsprechenden abgeschlossenen Intervalle.

Es gilt damit

Satz 14.3 *(Taylor)*
Es sei $I \subset \mathbb{R}$ ein beliebiges Intervall, und es sei $x_0 \in I$. Ferner sei für ein $n \in \mathbb{N}_0$ die

Funktion $f : I \to \mathbb{R}$ n-mal stetig differenzierbar auf I (d. h. $f \in C^n(I)$) und $(n+1)$ mal differenzierbar auf I^0 (d. h. $f^{(n+1)}$ existiert auf I^0). Dann gilt für alle $x \in I$ die "Taylor'sche Formel bzgl. der Entwicklungsmitte x_0":

$$f(x) = T_n(x) + R_n(x) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x_0)}{\nu!}(x - x_0)^\nu \ + R_n(x) \ ,$$

wobei für das "Restglied" $R_n(x)$ gilt: $R(x_0) = 0$ und

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1} \quad \textit{mit einem geeigneten } \xi = \xi(x) = \xi_{n,x_0}(x) \in I(x, x_0)$$

(sog. Lagrange-Form des Restgliedes).

Beweis. Ist $x = x_0$, so ist $f(x_0) = T_n(x_0)$, d h. $R_n(x_0) = 0$. Es sei also $x \neq x_0$. Wir betrachten die Hilfsfunktionen $\varphi, \psi : I[x, x_0] \to \mathbb{R}$ mit

$$\varphi(t) := f(x) - \sum_{\nu=0}^{n} \frac{f^{(\nu)}(t)}{\nu!}(x - t)^\nu \qquad \text{und} \qquad \psi(t) := (x - t)^{n+1} \ .$$

Dann sind nach Voraussetzung φ und ψ stetig auf $I[x, x_0]$ und differenzierbar auf $I(x, x_0)$. Ferner gilt für $t \in I(x, x_0)$

$$\begin{aligned}
\varphi'(t) &= -\sum_{\nu=0}^{n} \left\{ \frac{f^{(\nu)}(t)}{\nu!}(x - t)^\nu \right\}' = \\
&= -f'(t) - \sum_{\nu=1}^{n} \left\{ \frac{f^{(\nu+1)}(t)}{\nu!}(x - t)^\nu - \frac{f^{(\nu)}(t)}{(\nu - 1)!}(x - t)^{\nu-1} \right\} \\
&= -f'(t) - \left\{ \frac{f^{(n+1)}(t)}{n!}(x - t)^n - f'(t) \right\} \\
&= -\frac{f^{(n+1)}(t)}{n!}(x - t)^n \ ,
\end{aligned}$$

und

$$\psi'(t) = -(n+1)(x - t)^n \ .$$

Außerdem ist $\varphi(x) = 0 = \psi(x)$ und $\psi'(t) \neq 0$ in $I(x, x_0)$. Nach dem erweiterten Mittelwertsatz existiert ein $\xi = \xi(x) = \xi(x, x_0) \in I(x, x_0)$ mit

$$\frac{\varphi(x_0)}{\psi(x_0)} = \frac{\varphi(x_0) - \varphi(x)}{\psi(x_0) - \psi(x)} = \frac{\varphi'(\xi)}{\psi'(\xi)} \ ,$$

d. h.

$$\varphi(x_0) = \psi(x_0)\frac{\varphi'(\xi)}{\psi'(\xi)} \ .$$

Also folgt

$$\begin{aligned} R_n(x) &= f(x) - \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x_0)}{\nu!}(x-x_0)^\nu = \varphi(x_0) = \psi(x_0)\frac{\varphi'(\xi)}{\psi'(\xi)} \\ &= (x-x_0)^{n+1}\frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n \frac{1}{(n+1)(x-\xi)^n} = \\ &= \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1} . \end{aligned}$$

□

Bemerkung 14.4 1. Im Falle $n = 0$ (und unter den Voraussetzungen von S. 14.3) lautet die Taylor-Formel

$$f(x) = T_0(x) + R_0(x) = f(x_0) + f'(\xi)(x - x_0)$$

mit einem $\xi = \xi(x) \in I(x, x_0)$, d. h.

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

Dies ist die Aussage des Mittelwertsatzes. Der Taylor-Satz kann also als Verallgemeinerung des Mittelwertsatzes aufgefasst werden. Außerdem ist für $n = 1$, falls $|f''(\xi)| \leq M$ für alle $\xi \in I$ erfüllt ist,

$$\begin{aligned} f(x) &= T_1(x) + R_1(x) = \\ &= f(x_0) + f'(x_0)(x-x_0) + \frac{f''\big(\xi(x)\big)}{2}(x-x_0)^2 \end{aligned}$$

d. h.

$$\varepsilon(x) := \frac{f''\big(\xi(x)\big)}{2}|x - x_0|$$

erfüllt $|\varepsilon(x)| \leq \frac{M}{2}|x-x_0|$. Dies verbessert (unter den Zusatzvoraussetzungen) die Fehlerabschätzung in der Zerlegungsformel.

2. Weitere Darstellungen des Restgliedes R_n erhält man, indem man im Beweis zum Taylor-Satz andere Hilfsfunktionen ψ verwendet. So ergibt sich für jedes $p \in \{0, \dots, n\}$ mit

$$\psi(t) = \psi_p(t) := (x - t)^{n+1-p} \qquad (t \in I[x, x_0])$$

und dem gleichen Beweis wie zum Taylor-Satz

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^p \frac{(x-x_0)^{n+1-p}}{n+1-p}$$

mit einem geeigneten $\xi = \xi(x) = \xi_{p,n,x_0}(x) \in I(x, x_0)$. Insbesondere gilt für $p = n$

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n(x-x_0)$$

mit einem geeigneten $\xi \in I(x, x_0)$. Diese Darstellung des Restgliedes heißt *Cauchy-Form*.

Beispiel 14.5 Wir betrachten $f(x) = \cos x$ $(x \in \mathbb{R})$. Dann gilt

$$f^{(2k)}(x) = (-1)^k \cdot \cos x\ , \qquad f^{(2k+1)}(x) = (-1)^{k+1} \cdot \sin x$$

für $k \in \mathbb{N}_0$, d. h. für alle $n \in \mathbb{N}_0$ ist mit $x_0 = 0$

$$T_{2n}(x) = \sum_{\nu=0}^{2n} \frac{f^{(\nu)}(0)}{\nu!}x^\nu = \sum_{k=0}^{n} \frac{(-1)^k x^{2k}}{(2k)!} \ (= T_{2n+1}(x))$$

(vgl. S. 12.5). Nach S. 14.3 existiert für alle $x \in \mathbb{R}$ ein $\xi \in \mathbb{R}$ mit

$$\cos x - T_{2n}(x) = R_{2n+1}(x) = \frac{f^{(2n+2)}(\xi)}{(2n+2)!}x^{2n+2} = \frac{(-1)^{n+1}\cos(\xi)}{(2n+2)!}x^{2n+2}$$

also

$$|\cos x - T_{2n}(x)| \le \frac{x^{2n+2}}{(2n+2)!}\ .$$

Insbesondere haben wir also

$$|\cos x - T_2(x)| = \left|\cos(x) - 1 + \frac{x^2}{2}\right| \le \frac{x^4}{4!}$$

und

$$|\cos x - T_4(x)| = \left|\cos(x) - 1 + \frac{x^2}{2} - \frac{x^4}{4!}\right| \le \frac{x^6}{6!}\ .$$

Der folgende Satz liefert ein weiteres hinreichendes Kriterium für das Vorliegen von Extremstellen.

Satz 14.6 *Es sei $I = (a, b) \subset \mathbb{R}$ ein offenes Intervall, $n \in \mathbb{N}$, und es sei $f : I \to \mathbb{R}$ $(n+1)$-mal stetig differenzierbar auf I. Für einen Punkt $x_0 \in I$ gelte*

$$f'(x_0) = f''(x_0) = \ldots = f^{(n)}(x_0) = 0 \qquad \textit{und} \qquad f^{(n+1)}(x_0) \neq 0\ .$$

1. *Ist n ungerade, so ist x_0 eine Extremstelle von f. Dabei liegt im Falle $f^{(n+1)}(x_0) > 0$ ein lokales Minimum und im Falle $f^{(n+1)}(x_0) < 0$ ein lokales Maximum vor.*

2. *Ist n gerade, so ist x_0 keine Extremstelle von f.*

Beweis. Nach Voraussetzung gilt

$$T_n(x) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x_0)}{\nu!}(x-x_0)^\nu = f(x_0) \; .$$

1. Es gelte $f^{(n+1)}(x_0) > 0$. Wir wählen ein $\delta > 0$ so, dass $f^{(n+1)}(y) > 0$ für alle $y \in U_\delta(x_0)$ gilt. Nach S. 14.3 existiert zu jedem $x \in U_\delta(x_0)$ ein $\xi = \xi(x) \in U_\delta(x_0)$ mit

$$f(x) - f(x_0) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1} \; .$$

Ist n ungerade, so ist $(x-x_0)^{n+1} \geq 0$ für $x \in U_\delta(x_0)$, also

$$f(x) \geq f(x_0) \qquad (x \in U_\delta(x_0)) \; ,$$

d. h. an x_0 liegt ein lokales Minimum vor. Ist n gerade, so ist $(x-x_0)^{n+1} > 0$ für $x_0 < x < x_0+\delta$ und < 0 für $x_0-\delta < x < x_0$. Also ist $f(x) > f(x_0)$ für $x_0 < x < x_0+\delta$ und $f(x) < f(x_0)$ für $x_0 - \delta < x < x_0$. Damit liegt an x_0 kein Extremum vor.
2. Ist $f^{(n+1)}(x_0) < 0$, so verläuft der Beweis analog. □

Beispiel 14.7 1. Es sei $f(x) = 2x^3 + 3x^2 - 1$ $(x \in \mathbb{R})$ (vgl. B. 13.20.1).
Dann gilt

$$f'(x) = 6x^2 + 6x \; , \qquad f''(x) = 12x + 6 \qquad (x \in \mathbb{R}) \; .$$

Also gilt an den kritischen Stellen $x_1 = 0$ und $x_2 = -1$ (also an den Stellen x mit $f'(x) = 0$):

$$\begin{aligned} f''(x_1) &= f''(0) = 6 > 0 \\ f''(x_2) &= f''(-1) = -6 < 0 \; . \end{aligned}$$

Nach 14.6 hat f an 0 ein lokales Minimum und an -1 ein lokales Maximum.

2. Es seien $a_1, \ldots, a_N$ reelle Zahlen. Gesucht ist eine reelle Zahl a so, dass $\sum\limits_{j=1}^{N}(a-a_j)^2$ minimal wird. Wir betrachten also die Funktion $\varphi : \mathbb{R} \to \mathbb{R}$ mit

$$\varphi(x) := \sum_{j=1}^{N}(x-a_j)^2$$

Dann gilt

$$\varphi'(x) = 2\sum_{j=1}^{N}(x-a_j) = 0$$

genau dann, wenn $x = \frac{1}{N}\sum\limits_{j=1}^{N} a_j =: a$ ist (arithmetisches Mittel der a_j). Weiter ist

$$\varphi''(x) \equiv 2N > 0$$

und damit liegt an a ein lokales Minimum vor (dieses ist auch global).

Existieren für eine Funktion f die Ableitungen beliebiger Ordnung, so kann man die Folge der Taylor-Polynome $T_n(x)$ als die Teilsummefolge der Reihe $\sum\limits_{\nu=0}^{\infty} \frac{f^{(\nu)}(x_0)}{\nu!}(x-x_0)^\nu$ auffassen.

Definition 14.8 Es sei $M \subset \mathbb{R}$ und $x_0 \in M$. Ferner sei $f : M \to \mathbb{R}$ so, dass alle Ableitungen $f^{(\nu)}(x_0)$ für $\nu \in \mathbb{N}_0$ existieren. Dann heißt für $x \in \mathbb{R}$ die (formale) Reihe

$$\sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(x_0)}{\nu!}(x-x_0)^\nu$$

Taylor-Reihe von f an der Stelle x (mit *Entwicklungsmitte* x_0).

Bemerkung 14.9 Ist $I \subset \mathbb{R}$ ein Intervall, sind $x, x_0 \in I$ und ist

$$f \in C^\infty(I) := \bigcap_{n\in\mathbb{N}_0} C^n(I) = \{f : I \to \mathbb{R} : f \ \infty \text{ oft differenzierbar auf } I\})$$

so gilt nach dem Taylor-Satz

$$f(x) = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(x_0)}{\nu!}(x-x_0)^\nu$$

(d. h. $f(x)$ stimmt mit seiner Taylor-Reihe an der Stelle x überein) genau dann, wenn

$$R_n(x) = \frac{f^{(n+1)}(\xi_n(x))}{(n+1)!}(x-x_0)^{n+1} \to 0 \qquad (n \to \infty)$$

gilt. Insbesondere ist dies etwa dann erfüllt, wenn die Folge $(||f^{(n+1)}||_\infty)_n$ beschränkt ist, d. h. wenn ein $M > 0$ so existiert, dass

$$|f^{(n+1)}(t)| \le M \qquad (t \in I,\, n \in \mathbb{N}_0)\ .$$

(Denn: In diesem Fall gilt

$$|R_n(x)| \le \frac{M}{(n+1)!}|x-x_0|^n \to 0 \qquad (n \to \infty)\ .)$$

Beispiel 14.10 Es sei $f(x) = \cos x$ ($x \in \mathbb{R}$). Dann gilt nach B. 14.5

$$|f^{(n+1)}(t)| \leq 1 \qquad (t \in \mathbb{R},\ n \in \mathbb{N}_0)\ .$$

Also gilt (etwa für $x_0 = 0$) $R_n(x) \to 0$ für alle $x \in \mathbb{R}$ und damit

$$\cos x = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(0)}{\nu!} x^\nu = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} \qquad (x \in \mathbb{R})\ ,$$

(vgl. S. 12.5).

Bemerkung 14.11 Es gibt Funktionen $f \in C^\infty(\mathbb{R})$, für die die Taylor-Reihe (mit Entwicklungsmitte x_0) nur an dem Punkt x_0 den "richtigen" Wert $f(x)$ hat. Man kann etwa für $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

zeigen, dass f in $C^\infty(\mathbb{R})$ liegt und dass $f^{(\nu)}(0) = 0$ für alle $\nu \in \mathbb{N}_0$ gilt ([Ü]). Also ist hier $T_n(x) \equiv 0$ und damit ist auch die Taylor-Reihe (mit Entwicklungsmitte 0) $\equiv 0$ auf $\mathbb{R}$. Folglich stimmt $f(x)$ nur an $x_0 = 0$ mit "seiner" Taylor-Reihe überein.

Wir werden später noch eine ganze Klasse von Funktionen kennenlernen, die durch ihre Taylor-Reihe dargestellt werden können. Wir begnügen uns zunächst jedoch mit zwei weiteren Beispielen.

Beispiel 14.12 1. (Logarithmusreihe) Wir betrachten die Funktion $f : (-1, \infty) \to \mathbb{R}$ mit

$$f(x) = \ln(1 + x) \qquad (x > -1)\ .$$

Es gilt

$$f^{(\nu)}(x) = \frac{(-1)^{\nu-1}(\nu - 1)!}{(1 + x)^\nu} \qquad (\nu \in \mathbb{N})$$

also insbesondere

$$f^{(\nu)}(0) = (-1)^{\nu-1}(\nu - 1)! \qquad (\nu \in \mathbb{N}_0)\ .$$

Wir zeigen, daß für alle $x \in (-1, 1)$ gilt

$$\ln(1 + x) = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(0)}{\nu!} x^\nu = \sum_{\nu=1}^{\infty} \frac{(-1)^{\nu-1}}{\nu} x^\nu\ .$$

(Denn: O. E. sei $x \neq 0$. Nach B. 14.4.2 existiert ein $\theta = \theta_n(x) \in (0, 1)$ (nämlich $\theta = \xi/x$) mit

$$\begin{aligned} R_n(x) &= \frac{f^{(n+1)}(\theta x)}{n!}(x - \theta x)^n x = \frac{f^{(n+1)}(\theta x)}{n!}(1 - \theta)^n x^{n+1} \\ &= (-1)^n \frac{(1 - \theta)^n x^{n+1}}{(1 + \theta x)^{n+1}}\ . \end{aligned}$$

Also ist (da $1+\theta x > 0$)

$$|R_n(x)| \leq |x|^{n+1} \frac{(1-\theta)^n}{(1+\theta x)^{n+1}} \leq \frac{|x|^{n+1}}{1-|x|} \left(\frac{1-\theta}{1+\theta x}\right)^n .$$

Aus $0 < \frac{1-\theta}{1+\theta x} < 1$ folgt $|R_n(x)| \leq \frac{|x|^{n+1}}{1-|x|} \to 0\,(n \to \infty)$.)

2. (binomische Reihe) Es sei $f(x) = (1+x)^\alpha$ $(x > -1)$, wobei $\alpha \in \mathbb{R}$ fest ist. Hier gilt

$$f^{(\nu)}(x) = \alpha(\alpha-1)\cdots(\alpha-\nu+1)(1+x)^{\alpha-\nu} \qquad (\nu \in \mathbb{N})$$

also insbesondere

$$\frac{f^{(\nu)}(0)}{\nu!} = \frac{\alpha(\alpha-1)\cdots(\alpha-\nu+1)}{\nu!} = \binom{\alpha}{\nu} \qquad (\nu \in \mathbb{N}) .$$

Wir zeigen: Für alle $x \in (-1,1)$ ist

$$(1+x)^\alpha = \sum_{\nu=0}^{\infty} \binom{\alpha}{\nu} x^\nu$$

(Denn: O. E. sei $x \neq 0$. Nach B. 14.4.2 existiert ein $\theta = \theta_n(x) \in (0,1)$ mit

$$\begin{aligned} R_n(x) &= \frac{f^{(n+1)}(\theta x)}{n!}(1-\theta)^n x^{n+1} = \\ &= \alpha \cdot \frac{(\alpha-1)(\alpha-2)\cdots(\alpha-n)}{n!}(1+\theta x)^{\alpha-n-1}(1-\theta)^n x^{n+1} \\ &= \alpha \binom{\alpha-1}{n} x^{n+1}(1+\theta x)^{\alpha-1} \left(\frac{1-\theta}{1+\theta x}\right)^n \end{aligned}$$

also

$$\begin{aligned} |R_n(x)| &\leq \left|\binom{\alpha-1}{n} x^n\right| \cdot |\alpha x|\,(1+\theta x)^{\alpha-1} \left(\frac{1-\theta}{1+\theta x}\right)^n \\ &\leq M \cdot \left|\binom{\alpha-1}{n} x^n\right| \end{aligned}$$

mit einer von n unabhängigen Konstanten M. Da die Reihe $\sum\limits_{n=0}^{\infty} \binom{\alpha-1}{n} x^n$ nach den Quotientenkriterium konvergiert ([Ü]), gilt

$$\binom{\alpha-1}{n} x^n \to 0 \qquad (n \to \infty) ,$$

also auch $R_n(x) \to 0 \quad (n \to \infty)$.)

15 Funktionenfolgen und Funktionenreihen

Wir haben schon früher gesehen, dass wichtige elementare Funktionen wie die Exponentialfunktion oder die Logarithmusfunktion über gewisse Grenzwerte definiert sind. Ziel ist es nun, allgemeine Strukturaussagen über Funktionen zu machen, die sich als Grenzwerte von sog. Funktionenfolgen oder Funktionenreihen ergeben. Dabei betrachten wir $\mathbb{K}^m$-wertige Funktionen. Die meisten Ergebnisse gelten jedoch analog für Funktionen in metrische Räume.

Definition 15.1 1. Es sei X eine beliebige Menge, $X \neq \emptyset$.

1. Sind $f_n : X \to \mathbb{K}^m$ Funktionen, so nennt man die Folge $(f_n)_n$ eine *Funktionenfolge*. Die Funktionenfolge (f_n) heißt *punktweise konvergent* auf der Menge $M \subset X$, falls für alle $x \in M$ die Folge $(f_n(x))$ in $\mathbb{K}^m$ konvergiert. Die Funktion $f : M \to \mathbb{K}^m$ mit $f(x) := \lim\limits_{n\to\infty} f_n(x)$ heißt *Grenzfunktion* der Folge (f_n) (auf M).

2. Sind $a_\nu : X \to \mathbb{K}$ Funktionen, so heißt die Funktionenfolge (s_n) mit

$$s_n(x) := \sum_{\nu=0}^{n} a_\nu(x) \qquad (x \in X, n \in \mathbb{N}_0)$$

eine *Funktionenreihe*. Wir schreiben wieder $\sum\limits_{\nu=0}^{\infty} a_\nu$ oder $\sum\limits_{\nu=0}^{\infty} a_\nu(x)$. Die Funktionenreihe $\sum\limits_{\nu=0}^{\infty} a_\nu$ heißt *punktweise konvergent* auf $M \subset X$ falls (s_n) auf M punktweise konvergiert. Wir verwenden (wie bei Zahlenreihen) das Symbol $\sum\limits_{\nu=0}^{\infty} a_\nu$ bzw. $\sum\limits_{\nu=0}^{\infty} a_\nu(x)$ dann auch wieder für die Grenzfunktion, d. h.

$$\sum_{\nu=0}^{\infty} a_\nu(x) := \lim_{n\to\infty} \sum_{\nu=0}^{n} a_\nu(x) \qquad (x \in M)\,.$$

Natürlich definiert man entsprechend Funktionenreihen der Form $\sum\limits_{\nu=\nu_0}^{\infty} a_\nu(x)$ auch für allgemeine $\nu_0 \in \mathbb{Z}$.

Beispiel 15.2 1. Wir betrachten die Funktionen $f_n : \mathbb{R} \to \mathbb{R}$ mit

$$f_n(x) := x^n \qquad (x \in \mathbb{R},\, n \in \mathbb{N})\,.$$

Dann gilt

$$f_n(x) \to \begin{cases} 0 \;, & \text{falls} \quad x \in (-1,1) \\ 1 \;, & \text{falls} \quad x = 1 \end{cases} \quad ,$$

d. h. (f_n) konvergiert punktweise auf $(-1,1]$ und die Grenzfunktion $f:(-1,1]\to\mathbb{R}$ ist gegeben durch

$$f(x)=\begin{cases} 0 & , \quad x\in(-1,1) \\ 1 & , \quad x=1 \end{cases} \quad .$$

2. Es seien $a_\nu:\mathbb{R}\to\mathbb{R}$ definiert durch

$$a_\nu(x)=\frac{1}{\nu^x} \qquad (x\in\mathbb{R},\ \nu\in\mathbb{N}) .$$

Dann ist die Funktionenreihe

$$\sum_{\nu=1}^{\infty}\frac{1}{\nu^x}$$

auf $M=(1,\infty)$ punktweise konvergent (vgl. B. 6.9, wobei die dortige Behauptung auch für $\alpha\in\mathbb{R}$ gilt). Wir definieren $\zeta:(1,\infty)\to\mathbb{R}$ durch

$$\zeta(x):=\sum_{\nu=1}^{\infty}\frac{1}{\nu^x} \qquad (x>1) .$$

Die Funktion ζ heißt *(Riemann'sche) Zetafunktion.*

3. Es sei $a_\nu(z):=\frac{z^\nu}{\nu!}$ $(z\in\mathbb{C})$. Dann ist die Funktionenreihe

$$\sum_{\nu=0}^{\infty}a_\nu(z)=\sum_{\nu=0}^{\infty}\frac{z^\nu}{\nu!}$$

punktweise konvergent auf $\mathbb{C}$ (und es ist bekanntlich $e^z=\sum\limits_{\nu=0}^{\infty}z^\nu/\nu!$).

Das erste Beispiel zeigt insbesondere, dass die Grenzfunktion unstetig (an $x_0=1$) ist, obwohl alle Folgelieder f_n stetige Funktionen (und sogar ∞ oft differenzierbar) auf ganz $\mathbb{R}$ sind. Da wir an Aussagen der Form "f_n stetig $(n\in\mathbb{N})\Rightarrow f$ stetig" interessiert sind, führen wir einen "strengeren" Konvergenzbegriff ein, mit dessen Hilfe eine solche Aussage möglich wird.

Definition 15.3 1. Es sei $X\neq\emptyset$ eine beliebige Menge.

1. Sind $f_n:X\to\mathbb{K}^m$ Funktionen $(n\in\mathbb{N})$, so heißt die Funktionenfolge (f_n) *gleichmäßig konvergent* auf der Menge $M\subset X$ (gegen die Grenzfunktion $f:M\to\mathbb{K}^m$), falls zu jedem $\varepsilon>0$ ein $N_\varepsilon\in\mathbb{N}$ so existiert, das

$$|f(x)-f_n(x)|<\varepsilon \qquad \text{für alle } n\geq N_\varepsilon \text{ und alle } x\in M .$$

Wir schreiben dann

$$f_n\to f \qquad (n\to\infty) \quad \text{gleichmäßig auf } M .$$

2. Eine Funktionenreihe $\sum_{\nu=0}^{\infty} a_\nu(x)$ heißt *gleichmäßig konvergent* auf $M \subset X$, falls die Funktionenfolge (s_n) mit $s_n(x) = \sum_{\nu=0}^{n} a_\nu(x)$ gleichmäßig auf M konvergiert.

Bemerkung 15.4 1. Gilt $f_n \to f$ gleichmäßig auf M, so gilt insbesondere $f_n(x) \to f(x)\,(n \to \infty)$ für alle $x \in M$, d. h. (f_n) konvergiert auf M auch punktweise gegen f ("gleichmäßige Konvergenz impliziert punktweise Konvergenz").
Der Unterschied zur punktweisen Konvergenz liegt darin, dass das N_ε unabhängig von x gewählt werden kann. Punktweise Konvergenz bedeutet: Für alle $x \in M$ und alle $\varepsilon > 0$ existiert ein $N_{\varepsilon,x} \in \mathbb{N}$ mit $|f(x) - f_n(x)| < \varepsilon$ für alle $n \geq N_{\varepsilon,x}$.

2. Oft erweist sich folgende Charakterisierung der gleichmäßigen Konvergenz als äußerst nützlich: Es ist

$$f_n \to f \qquad \text{gleichmäßig auf } M$$

genau dann, wenn

$$||f - f_n||_\infty = \sup_M |f(x) - f_n(x)| \to 0 \qquad (n \to \infty)$$

gilt (wobei wir formal $\sup(X) := \infty$ für nach oben unbeschränkte Mengen $X \subset \mathbb{R}$ setzen).
(Denn: Gilt $f_n \to f$ gleichmäßig auf M, so existiert zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ mit

$$|f(x) - f_n(x)| < \varepsilon \qquad (x \in M,\ n \geq N_\varepsilon)\ .$$

Hieraus folgt

$$||f - f_n||_\infty = \sup_M |f(x) - f_n(x)| \leq \varepsilon \qquad (n \geq N_\varepsilon)\ .$$

Da $\varepsilon > 0$ beliebig war, ist $(||f - f_n||_\infty)$ eine Nullfolge in $\mathbb{R}$.
Es gelte umgekehrt

$$||f - f_n||_\infty \to 0 \qquad (n \to \infty)\ .$$

Ist $\varepsilon > 0$ gegeben, so existiert ein $N_\varepsilon \in \mathbb{N}$ mit

$$\sup_M |f(x) - f_n(x)| = ||f - f_n||_\infty < \varepsilon \qquad (n \geq N_\varepsilon)\ .$$

Dann ist auch

$$|f(x) - f_n(x)| < \varepsilon \qquad (n \geq N_\varepsilon,\ x \in M)\ .$$

Also gilt $f_n \to f$ gleichmäßig auf M.)

Beispiel 15.5 Wir betrachten noch einmal $f_n(x) = x^n$ $(n \in \mathbb{N},\ x \in \mathbb{R})$ (vgl. B. 15.2.1). Ist $M = [-1/2, 1/2]$ und $\varepsilon > 0$ gegeben, so gilt

$$|f_n(x) - f(x)| = |x|^n \leq 1/2^n \qquad (n \in \mathbb{N},\ x \in [-1/2, 1/2])\ .$$

Ist $N_\varepsilon \in \mathbb{N}$ so, dass $1/2^n < \varepsilon\,(n \geq N_\varepsilon)$, so ist also

$$|f_n(x) - f(x)| < \varepsilon \qquad (n \geq N_\varepsilon,\ x \in [-1/2, 1/2])\ .$$

Damit gilt

$$x^n \to 0 \qquad \text{gleichmäßig auf} \quad [-1/2, 1/2]\ .$$

Andererseits ist für $M = [0, 1)$

$$\sup_M |f_n(x) - f(x)| = \sup_{[0,1)} x^n = 1 \qquad (n \in \mathbb{N})\ .$$

Also ist (f_n) nicht gleichmäßig konvergent auf $[0, 1)$ nach B. 15.4.

In S. 6.15 hatten wir als wichtiges hinreichendes Kriterium für (absolute) Konvergenz von Zahlenreihen das Weierstraß'sche Majorantenkriterium kennen gelernt. Ein entsprechendes Ergebnis gibt es für Funktionenreihen.

Satz 15.6 *(Weierstraß'sches Majorantenkriterium)*
Es sei $M \neq \emptyset$ eine beliebige Menge, und es seien $a_\nu : M \to \mathbb{K}$ $(\nu \in \mathbb{N}_0)$. Ferner sei $\sum_{\nu=0}^{\infty} b_\nu$ eine konvergente Reihe mit $b_\nu \geq 0$ $(\nu \in \mathbb{N}_0)$. Existiert ein $N \in \mathbb{N}$ so, dass $|a_\nu(x)| \leq b_\nu$ für alle $\nu \geq N$ und $x \in M$, so konvergiert $\sum_{\nu=0}^{\infty} a_\nu$ gleichmäßig auf M.

Beweis. Nach S. 6.15 konvergiert $\sum_{\nu=0}^{\infty} a_\nu$ punktweise auf M, d. h. $\sum_{\nu=0}^{\infty} a_\nu(x)$ konvergiert für alle $x \in M$.
Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $N_\varepsilon \in \mathbb{N}$ (o. E. $N_\varepsilon \geq N$) mit

$$(0 \leq)\ \sum_{\nu=n+1}^{\infty} b_\nu = \sum_{\nu=0}^{\infty} b_\nu - \sum_{\nu=0}^{n} b_\nu < \varepsilon \qquad (n \geq N_\varepsilon)\ .$$

Also folgt

$$\begin{aligned} \left| \sum_{\nu=0}^{\infty} a_\nu(x) - \sum_{\nu=0}^{n} a_\nu(x) \right| &= \left| \sum_{\nu=n+1}^{\infty} a_\nu(x) \right| \leq \sum_{\nu=n+1}^{\infty} |a_\nu(x)| \\ &\leq \sum_{\nu=n+1}^{\infty} b_\nu < \varepsilon \qquad (n \geq N_\varepsilon, x \in M)\ . \end{aligned}$$

Damit ist $\sum_{\nu=0}^{\infty} a_\nu = \lim_{n\to\infty} \sum_{\nu=0}^{n} a_\nu$ gleichmäßig auf M. □

Beispiel 15.7 Es sei $a_\nu = \frac{1}{\nu^x}$ $(x > 1,\ \nu \in \mathbb{N})$ (vgl. B. 15.2.2). Ist $M = [\alpha, \infty)$, wobei $\alpha > 1$ ist, so ist $\sum\limits_{\nu=1}^{\infty} a_\nu(x) = \sum\limits_{\nu=1}^{\infty} \frac{1}{\nu^x}$ gleichmäßig konvergent auf M.
(Denn: Für alle $x \in M$ und alle $\nu \in \mathbb{N}$ ist

$$|a_\nu(x)| = \frac{1}{\nu^x} \le \frac{1}{\nu^\alpha}$$

und $\sum\limits_{\nu=1}^{\infty} \frac{1}{\nu^\alpha}$ ist konvergent. Also ergibt sich die Behauptung aus S. 15.6.)

Ein notwendiges und hinreichendes Kriterium für gleichmäßige Konvergenz liefert

Satz 15.8 *(Cauchy-Kriterium für gleichmäßige Konvergenz)*
Es sei $X \neq \emptyset$ eine Menge und es seien $f_n : X \to \mathbb{K}^m$ Funktionen $(n \in \mathbb{N})$. Ist $M \subset X$, so ist (f_n) gleichmäßig konvergent auf M genau dann, wenn zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ so existiert, dass

$$|f_n(x) - f_m(x)| < \varepsilon \qquad \text{für alle}\, n, m \ge N_\varepsilon \text{ und alle}\, x \in M$$

(sog. gleichmäßige Cauchy-Bedingung).

Beweis. 1. Es gelte $f_n \to f$ gleichmäßig auf M. Ist $\varepsilon > 0$ gegeben, so existiert ein $N_\varepsilon \in \mathbb{N}$ mit

$$|f(x) - f_n(x)| < \frac{\varepsilon}{2} \qquad (n \ge N_\varepsilon,\ x \in M)\ .$$

Für alle $n, m \ge N_\varepsilon$ und alle $x \in M$ folgt dann

$$|f_n(x) - f_m(x)| \le |f_n(x) - f(x)| + |f(x) - f_m(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\ .$$

Also ist die Cauchy-Bedingung erfüllt.

2. Es gelte die Cauchy-Bedingung. Dann ergibt sich insbesondere, dass für jedes feste $x \in M$ die Folge $(f_n(x))_n$ in $(\mathbb{K}^m, d_{|\cdot|})$ eine Cauchy-Folge ist, Da $(\mathbb{K}^m, d_{|\cdot|})$ vollständig ist (B. 8.24), ist $(f_n(x))_n$ konvergent, d. h. es existiert ein $y \in \mathbb{K}^m$ mit $f_n(x) \to y\ (n \to \infty)$. Wir definieren $f : M \to \mathbb{K}^m$ durch $f(x) := y\ (x \in M)$ und zeigen: $f_n \to f$ gleichmäßig auf M.
Dazu sei $\varepsilon > 0$ gegeben. Dann existiert ein $N_\varepsilon \in \mathbb{N}$ mit

$$|f_m(x) - f_n(x)| < \frac{\varepsilon}{2} \qquad (n, m \ge N_\varepsilon,\ x \in M)\ .$$

Hieraus folgt für festes $n \ge N_\varepsilon$ und $x \in M$

$$|f(x) - f_n(x)| = \lim_{m\to\infty} |f_m(x) - f_n(x)| \le \frac{\varepsilon}{2} < \varepsilon\ .$$

Da $\varepsilon > 0$ beliebig war, ergibt sich die Behauptung. □

Bemerkung 15.9 Aus S. 15.8 ergibt sich leicht, dass der Raum $B(M, \mathbb{K}^m)$ aller auf M beschränkten $\mathbb{K}^m$-wertigen Funktionen mit der "Metrik der gleichmäßigen Konvergenz"

$$d(f,g) := ||f-g||_\infty = \sup_M |f(x)-g(x)| \qquad (f,g \in B(M,\mathbb{K}^m))$$

aus B. 8.12 vollständig ist.
(Denn: Es sei (f_n) eine Cauchy-Folge in $B(M,\mathbb{K}^m)$. Dann existiert zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ mit

$$||f_n - f_m||_\infty = \sup_M |f_n(x) - f_m(x)| < \varepsilon \quad (n, m \geq N_\varepsilon) \ .$$

Also ist auch

$$|f_n(x) - f_m(x)| < \varepsilon \qquad (n, m \geq N_\varepsilon, \, x \in M) \ .$$

Nach S. 15.8 existiert eine Funktion $f : M \to \mathbb{K}^m$ mit $f_n \to f$ gleichmäßig auf M, d. h. $||f - f_n||_\infty \to 0$ nach B. 15.4.2. Es bleibt noch zu zeigen, dass $f \in B(M,\mathbb{K}^m)$ d. h. f beschränkt ist. Dazu wählen wir zu $\varepsilon = 1$ ein $N \in \mathbb{N}$ mit

$$\sup_M |f(x) - f_N(x)| < 1 \ .$$

Dann gilt für alle $x \in M$

$$\begin{aligned} |f(x)| &\leq |f(x) - f_N(x)| + |f_N(x)| \\ &\leq 1 + ||f_N||_\infty \end{aligned}$$

d. h. f ist beschränkt.)

Wir kommen nun zu dem bereits oben angedeuteten Ergebnis über die "Vererbung" der Stetigkeit auf die Grenzfunktion.

Satz 15.10 *Es seien (X,d) ein metrischer Raum und $f, f_n : X \to \mathbb{K}^m$ Funktionen. Dann gilt: Sind alle Funktionen f_n stetig an der Stelle $x_0 \in X$ und existiert eine Umgebung U von x_0 mit $f_n \to f$ gleichmäßig auf U, so ist auch f stetig an x_0.*

Beweis. Es sei $\varepsilon > 0$ gegeben. Auf Grund der gleichmäßigen Konvergenz von (f_n) gegen f auf U existiert ein $N = N_\varepsilon \in \mathbb{N}$ mit

$$|f(x) - f_n(x)| < \varepsilon/3 \qquad (n \geq N, \ x \in U) \ .$$

Da f_N stetig an x_0 ist, existiert ein $\delta_\varepsilon > 0$ so, dass

$$|f_N(x) - f_N(x_0)| < \varepsilon/3 \qquad \text{für alle } x \text{ mit } d(x,x_0) < \delta_\varepsilon$$

O. E. können wir δ_ε so wählen, dass $U_{\delta_\varepsilon}(x_0) \subset U$ gilt. Für alle x mit $d(x, x_0) < \delta_\varepsilon$ ergibt sich damit

$$\begin{aligned}|f(x) - f(x_0)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon \,.\end{aligned}$$

Folglich ist f stetig an x_0. □

Bemerkung und Definition 15.11 1. Natürlich läßt sich die Aussage von S. 15.10 sofort auf Funktionenreihen übertragen, indem man den Satz auf die Teilsummenfolge (s_n) anwendet.
2. Sind (X, d) ein metrischer Raum und (f_n) eine Funktionenfolge auf X, so nennt man (f_n) *lokal gleichmäßig konvergent* auf X, falls zu jedem $x_0 \in X$ eine Umgebung von x_0 existiert, auf der (f_n) gleichmäßig konvergiert. Insbesondere ist jede auf X gleichmäßig konvergente Folge auch lokal gleichmäßig konvergent.
S. 15.10 zeigt, dass die Grenzfunktion einer lokal gleichmäßig konvergenten Folge stetiger Funktionen ebenfalls stetig ist.
Weiter heißt eine Funktionenreihe $\sum\limits_{\nu=0}^{\infty} a_\nu$ *lokal gleichmäßig konvergent* auf X, falls die Teilsummenfolge (s_n) lokal gleichmäßig konvergent auf X ist. Man sieht also: Sind die a_ν stetig auf X und ist die Konvergenz lokal gleichmäßig auf X, so ist auch $\sum\limits_{\nu=0}^{\infty} a_\nu$ stetig auf X.

Beispiel 15.12 (vgl. B. 15.2.2 und B. 15.7)
Ist $\alpha > 1$, so gilt

$$\zeta(x) = \sum_{\nu=1}^{\infty} \frac{1}{\nu^x}$$

gleichmäßig auf $[\alpha, \infty)$. Insbesondere konvergiert deshalb die Reihe lokal gleichmäßig auf $(1, \infty)$. Da $1/\nu^x$ für alle $\nu \in \mathbb{N}$ stetig auf $\mathbb{R}$ ist, ist ζ stetig auf $(1, \infty)$ nach B/D. 15.11.

Wir wenden uns nun der Frage zu, unter welchen Voraussetzungen an ein Funktionenfolge oder -reihe man "lim" und Differenziation vertauschen kann, d. h. wann gilt

$$\lim_{n\to\infty} f_n'(x) = \Big(\lim_{n\to\infty} f_n(x)\Big)' \;?$$

Beispiel 15.13 Wir betrachten $f_n : \mathbb{R} \to \mathbb{R}$ mit

$$f_n(x) = \frac{\sin(nx)}{\sqrt{n}} \qquad (x \in \mathbb{R}\ ,\ n \in \mathbb{N})\ .$$

Dann gilt

$$\sup_{x\in\mathbb{R}} |f_n(x) - 0| \le \frac{1}{\sqrt{n}} \to 0 \qquad (n \to \infty)\ ,$$

also

$$f_n \to f \equiv 0 \qquad \text{gleichmäßig auf } \mathbb{R}\ .$$

Weiter erhalten wir

$$f_n'(x) = \sqrt{n}\cos(nx) \qquad (n \in \mathbb{N}\ ,\ x \in \mathbb{R})$$

also gilt etwa für $x = \pi$:

$$f_n'(\pi) = (-1)^n\sqrt{n}$$

d. h . $(f_n'(\pi))_n$ ist divergent. Insbesondere gilt **nicht** für alle $x \in \mathbb{R}$

$$\lim_{n\to\infty} f_n'(x) = (\lim_{n\to\infty} f_n)'(x) \equiv 0$$

(obwohl die Grenzfunktion $f \equiv 0$ überall differenzierbar ist).
Also: Aus $f_n \to f$ gleichmäßig auf M und f_n, f differenzierbar auf M folgt i. a. **nicht** $f_n' \to f'$ auf M.

Anders verhält sich die Sache, wenn die Folge der Ableitungen gleichmäßig konvergiert.

Satz 15.14 *Es sei $[a,b] \subset \mathbb{R}$ und die Funktionen $f_n : [a,b] \to \mathbb{R}$ seien differenzierbar auf $[a,b]$. Ferner existiere ein $x_0 \in [a,b]$ so, dass $(f_n(x_0))_n$ konvergiert und die Folge (f_n') sei gleichmäßig konvergent auf $[a,b]$. Dann gilt*

1. *(f_n) konvergiert gleichmäßig auf $[a,b]$.*
2. *Ist $f := \lim_{n\to\infty} f_n$ auf $[a,b]$, so ist f differenzierbar auf $[a,b]$, und es gilt*

$$f'(x) = \lim_{n\to\infty} f_n'(x) \qquad (x \in [a,b])\ .$$

Beweis. 1. Es sei $\varepsilon > 0$ gegeben. Nach Voraussetzung existiert ein $N_\varepsilon \in \mathbb{N}$ so, dass

$$|f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2}$$

und

$$|f_n'(t) - f_m'(t)| < \frac{\varepsilon}{2(b-a)}$$

für alle $n, m \geq N_\varepsilon$ und alle $t \in [a, b]$ gilt. (Man beachte: $(f_n(x_0))_n$ ist eine Cauchy-Folge in $\mathbb{R}$ und (f_n') ist nach S. 15.8 eine “gleichmäßige Cauchy-Folge” auf $[a, b]$.)
Es sei $x \in [a, b]$ beliebig. Dann existiert nach dem Mittelwertsatz für alle $n, m \in \mathbb{N}$ ein $\xi = \xi_{n,m,x} \in (a, b)$ mit

$$(f_n - f_m)(x) - (f_n - f_m)(x_0) = (f_n' - f_m')(\xi)(x - x_0) .$$

Also folgt für $n, m \geq N_\varepsilon$

$$\begin{array}{rcl} |f_n(x) - f_m(x)| & \leq & |(f_n - f_m)(x) - (f_n - f_m)(x_0)| + |f_n(x_0) - f_m(x_0)| \\ & < & \dfrac{\varepsilon}{2(b-a)}|x - x_0| + \dfrac{\varepsilon}{2} \leq \varepsilon . \end{array}$$

Damit ist (f_n) nach S. 15.8 gleichmäßig konvergent auf $[a, b]$. Wir setzen $f(x) := \lim\limits_{n\to\infty} f_n(x)$.

2. Es sei $x_1 \in [a, b]$ beliebig. Wir zeigen: f ist differenzierbar an x_1 und es gilt

$$f'(x_1) = \lim_{n\to\infty} f_n'(x_1) .$$

Dazu setzen wir für $n \in \mathbb{N}$ und $x \in [a, b]$

$$g_n(x) := \begin{cases} \dfrac{f_n(x) - f_n(x_1)}{x - x_1} & , \text{ falls } \quad x \neq x_1 \\ f_n'(x_1) & , \text{ falls } \quad x = x_1 \end{cases} .$$

Aufgrund der Differenzierbarkeit von f_n ist g_n stetig auf $[a, b]$.
Wir zeigen: (g_n) ist gleichmäßig konvergent auf $[a, b]$: Für alle $x \in [a, b], x \neq x_1$, ist nach dem Mittelwertsatz mit einem $\xi = \xi_{n,m,x}$

$$\begin{array}{rcl} |g_n(x) - g_m(x)| & = & \left| \dfrac{(f_n - f_m)(x) - (f_n - f_m)(x_1)}{x - x_1} \right| \\ & = & |(f_n' - f_m')(\xi)| \leq ||f_n' - f_m'||_\infty \end{array}$$

und für $x = x_1$ ist auch

$$|g_n(x_1) - g_m(x_1)| = |f_n'(x_1) - f_m'(x_1)| \leq ||f_n' - f_m'||_\infty .$$

Also gilt

$$||g_n - g_m||_\infty \leq ||f_n' - f_m'||_\infty .$$

Aus der gleichmäßigen Konvergenz von (f_n') ergibt sich (mit S. 15.8) die gleichmäßige Konvergenz von (g_n).
Es sei

$$g(x) := \lim_{n\to\infty} g_n(x) \qquad (n \in \mathbb{N}, \, x \in [a, b]) .$$

Dann ist g nach S. 15.10 stetig auf $[a,b]$.
Für $x \neq x_1$ gilt

$$g(x) = \lim_{n\to\infty} g_n(x) = \lim_{n\to\infty} \frac{f_n(x) - f_n(x_1)}{x - x_1} = \frac{f(x) - f(x_1)}{x - x_1}$$

und für $x = x_1$ gilt auf Grund der Stetigkeit von g

$$g(x_1) = \lim_{x\to x_1} g(x) = \lim_{x\to x_1} \frac{f(x) - f(x_1)}{x - x_1} .$$

Also existiert $f'(x_1)$ (und es ist $f'(x_1) = g(x_1)$).
Aus der Konvergenz von (g_n) gegen g ergibt sich schließlich noch

$$\lim_{n\to\infty} f_n'(x_1) = \lim_{n\to\infty} g_n(x_1) = g(x_1) = f'(x_1) .$$

Damit ist der Satz bewiesen. □

Bemerkung 15.15 1. Aus S. 15.14 ergibt sich leicht ([Ü]): Es sei $I \subset \mathbb{R}$ ein beliebiges Intervall, und es sei (f_n) eine Folge auf I differenzierbarer Funktionen so, dass (f_n') auf I lokal gleichmäßig und $(f_n(x_0))$ für ein x_0 konvergiert. Dann konvergiert (f_n) ebenfalls lokal gleichmäßig auf I und die Grenzfunktion f ist differenzierbar auf I mit

$$f'(x) = \lim_{n\to\infty} f_n'(x) \qquad (x \in I) .$$

2. Natürlich lassen sich die Aussagen von S. 15.14 und aus 1. sofort auf Funktionenreihen übertragen, indem man die entsprechenden Ergebnisse auf die Teilsummenfolge (s_n) anwendet. Ist etwa $I \subset \mathbb{R}$ ein Intervall und sind $a_\nu : I \to \mathbb{R}$ differenzierbar auf I und so, dass $\sum\limits_{\nu=0}^{\infty} a_\nu'$ lokal gleichmäßig auf I und $\sum\limits_{\nu=0}^{\infty} a_\nu(x_0)$ für ein x_0 konvergiert, so ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ lokal gleichmäßig konvergent und differenzierbar auf I mit

$$\left(\sum_{\nu=0}^{\infty} a_\nu(x)\right)' = \sum_{\nu=0}^{\infty} a_\nu'(x) \quad (x \in I) .$$

So gilt etwa für $a_\nu(x) = \frac{1}{\nu^x}$ $\quad (x > 1)$

$$a_\nu'(x) = \left(\frac{1}{\nu^x}\right)' = -\frac{\ln \nu}{\nu^x} \qquad (x > 1,\ \nu \in \mathbb{N}) .$$

Ist $\alpha > 1$, so gilt für $\beta := \frac{\alpha+1}{2}$ und $x \geq \alpha$ sowie $\nu \geq \nu_0(\alpha)$

$$\frac{\ln \nu}{\nu^x} \leq \frac{\ln \nu}{\nu^\alpha} \leq \frac{\nu^{\frac{\alpha-1}{2}}}{\nu^\alpha} = \frac{1}{\nu^\beta} .$$

Also konvergiert $\sum_{\nu=0}^{\infty} a'_\nu$ nach S. 15.6 gleichmäßig auf $[\alpha, \infty)$ und damit insbesondere lokal gleichmäßig auf $(1, \infty)$. Da $\sum_{\nu=0}^{\infty} a_\nu$ ebenfalls (lokal gleichmäßig) auf $(1, \infty)$ konvergiert, ist $\zeta(x) = \sum_{\nu=1}^{\infty} \frac{1}{\nu^x}$ differenzierbar auf $(1, \infty)$ und es gilt

$$\zeta'(x) = \left(\sum_{\nu=1}^{\infty} \frac{1}{\nu^x}\right)' = \sum_{\nu=1}^{\infty} \left(\frac{1}{\nu^x}\right)' = -\sum_{\nu=1}^{\infty} \frac{\ln \nu}{\nu^x} \qquad (x > 1) .$$

3. Ohne die Voraussetzung der Konvergenz von $(f_n(x_0))_n$ in mindestens einem Punkt x_0 ist die Aussage von S. 15.14 i. a. falsch. So gilt etwa für $f_n(x) \equiv n$ auf $\mathbb{R}$

$$f'_n(x) \equiv 0 \to 0 \qquad \text{gleichmäßig auf } \mathbb{R}$$

aber $(f_n(x))_n$ ist offenbar divergent in allen Punkten $x \in \mathbb{R}$.

Wir beweisen zum Abschluss noch folgendes interessante Ergebnis

Satz 15.16 *Es existiert eine auf $\mathbb{R}$ stetige Funktion, die nirgends differenzierbar ist.*

Beweis. Wir definieren $\varphi : \mathbb{R} \to \mathbb{R}$ durch $\varphi(x) := |x|$ für $x \in [-1, 1]$ und ansonsten durch 2-periodische Fortsetzung. Dann gilt offenbar

$$|\varphi(t) - \varphi(s)| \le |t - s| \qquad (t, s \in \mathbb{R}) .$$

Damit setzen wir

$$f(x) := \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \varphi(4^n x) \qquad (x \in \mathbb{R}) .$$

Aus

$$\left|\left(\frac{3}{4}\right)^n \varphi(4^n x)\right| \le \left(\frac{3}{4}\right)^n$$

ergibt sich die gleichmäßige Konvergenz der Reihe auf $\mathbb{R}$ (S. 15.6). Da die einzelnen Reihenglieder stetig auf $\mathbb{R}$ sind, ist auch f stetig auf $\mathbb{R}$ nach S. 15.10.
Nun sei $x \in \mathbb{R}$ fest. Für $m \in \mathbb{N}$ setzen wir

$$\delta_m := \pm \frac{1}{2} \frac{1}{4^m}$$

wobei das Vorzeichen so gewählt ist, dass zwischen $4^m x$ und $4^m(x + \delta_m)$ keine ganze Zahl liegt. Wir betrachten

$$\gamma_n := \frac{\varphi(4^n(x + \delta_m)) - \varphi(4^n x)}{\delta_m} \qquad (n \in \mathbb{N}_0) .$$

Für $n > m$ ist $\gamma = 0$, da $4^n \delta_m$ eine gerade ganze Zahl ist. Für $n \leq m$ ist $|\gamma_n| \leq 4^n$ und für $n = m$ ist $\gamma_m = 4^m$. Damit gilt

$$\left| \frac{f(x+\delta_m) - f(x)}{\delta_m} \right| = \left| \sum_{n=0}^{m} \left(\frac{3}{4}\right)^n \gamma_n \right| \geq 3^m - \sum_{n=0}^{m-1} 3^n = \frac{1}{2}(3^m + 1) \to \infty \quad (m \to \infty) .$$

Also ist f nicht differenzierbar an x. □

16 Potenzreihen

Wir untersuchen jetzt eine Klasse besonders wichtiger Funktionenreihen, sog. Potenzreihen. Hier sind die Teilsummen s_n Polynome vom Grad $\le n$, also von besonders einfacher Struktur. Einige Beispiele kennen wir schon sehr gut:
Für $s_n(z) = \sum\limits_{\nu=0}^{n} z^\nu$ ist

$$\frac{1}{1-z} = \sum_{\nu=0}^{\infty} z^\nu = \lim_{n\to\infty} s_n(z) \qquad (|z| < 1)$$

und für $s_n(z) = \sum\limits_{\nu=0}^{n} \frac{z^\nu}{\nu!}$ ist

$$e^z = \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} = \lim_{n\to\infty} s_n(z) \qquad (z \in \mathbb{K})\ .$$

Allgemeiner betrachten wir

Definition 16.1 Es sei $z_0 \in \mathbb{K}$ und es sei $(a_\nu)_{\nu=0}^{\infty}$ eine Folge in $\mathbb{K}$. Dann heißt die Funktionenreihe $\sum\limits_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu$ (also die Funktionenfolge (s_n) mit $s_n(x) = \sum\limits_{\nu=0}^{n} a_\nu (z-z_0)^\nu$) auf $\mathbb{K}$ eine *Potenzreihe* (mit der *Entwicklungsmitte* z_0 und der *Koeffizientenfolge* (a_ν)).

Man beachte, dass in der Definition keine Konvergenz der Potenzreihe gefordert wird. Über das Konvergenzverhalten von Potenzreihen gibt der folgende Satz Auskunft:

Satz 16.2 *(Cauchy-Hadamard)*
Es sei $\sum\limits_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu$ *eine Potenzreihe und es sei*

$$a := \overline{\lim_{\nu\to\infty}} \sqrt[\nu]{|a_\nu|}\ .$$

Dann gilt:

1. *Ist* $a = 0$ *so konvergiert die Potenzreihe absolut für alle* $z \in \mathbb{K}$.

2. *Ist* $0 < a < \infty$, *so konvergiert die Potenzreihe absolut für alle* $z \in \mathbb{K}$ *mit* $|z-z_0| < 1/a$ *und divergiert für alle* $z \in \mathbb{K}$ *mit* $|z-z_0| > 1/a$.

3. *Ist* $a = \infty$ *(d. h.* $(\sqrt[\nu]{|a_\nu|})_\nu$ *ist unbeschränkt), so divergiert die Potenzreihe für alle* $z \in \mathbb{K} \setminus \{z_0\}$.

Beweis. Es gilt für alle $\nu \in \mathbb{N}_0$

$$|a_\nu (z-z_0)^\nu|^{1/\nu} = |a_\nu|^{1/\nu} \cdot |z-z_0|\ ,$$

also ist im Falle $a < \infty$

$$\overline{\lim_{\nu\to\infty}} |a_\nu (z-z_0)^\nu|^{1/\nu} = a \cdot |z - z_0| \, .$$

Damit ergeben sich 1. und 2. sofort aus dem Wurzelkriterium (S. 6.17).
Ist $a = \infty$, so ist für $z \neq z_0$ auch $(|a_\nu(z-z_0)^\nu|^{1/\nu} = |a_\nu|^{1/\nu}|z-z_0|)_\nu$ unbeschränkt und folglich ist $(a_\nu(z-z_0)^\nu)_\nu$ insbesondere keine Nullfolge. Damit ergibt sich auch 3. $\square$

Definition 16.3 Ist $\sum\limits_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu$ eine Potenzreihe, so heißt

$$R := a^{-1} = \left(\overline{\lim_{\nu\to\infty}} \sqrt[\nu]{|a_\nu|} \right)^{-1} \in \mathbb{R}_0^+ \cup \{\infty\}$$

Konvergenzradius der Potenzreihe (wobei $\frac{1}{\infty} := 0$ und $\frac{1}{0} := \infty$ gesetzt ist). Im Falle $R > 0$ heißt weiterhin

$$U_R(z_0) = \{z \in \mathbb{K} : |z - z_0| < R\}$$

Konvergenzkreis (im Falle $\mathbb{K} = \mathbb{R}$ meist *Konvergenzintervall*) der Potenzreihe (wobei $U_\infty(z_0) = \{z \in \mathbb{K} : |z - z_0| < \infty\} = \mathbb{K}$ ist).

Beispiel 16.4 1. Für die Potenzreihe $\sum\limits_{\nu=0}^{\infty} z^\nu$ (also die geometrische Reihe) gilt $z_0 = 0$ und $a_\nu \equiv 1\,(\nu \in \mathbb{N}_0)$, also ist $R = 1$. Nach S. 16.2 oder den früheren Überlegungen konvergiert $\sum\limits_{\nu=0}^{\infty} z^\nu$ absolut im Konvergenzkreis $U_1(0) = \{z : |z| < 1\}$ und divergiert für alle z mit $|z| > 1$.

2. Für die Potenzreihe $\sum\limits_{\nu=0}^{\infty} \frac{z^\nu}{\nu!}$ ist $z_0 = 0$ und $a_\nu = \frac{1}{\nu!}\,(\nu \in \mathbb{N}_0)$. Hier ist $\overline{\lim\limits_{\nu=\infty}} |a_\nu|^{1/\nu} = \lim\limits_{\nu\to\infty} \frac{1}{\sqrt[\nu]{\nu!}} = 0$, d. h. $R = \infty$. Bekanntlich (oder jetzt auch nach S. 16.2) konvergiert die Potenzreihe absolut für alle $z \in \mathbb{C}$.

3. Für die Potenzreihe $\sum\limits_{\nu=1}^{\infty} \frac{(-1)^{\nu-1} z^\nu}{\nu}$ gilt $z_0 = 0$ und $a_\nu = \frac{(-1)^{\nu-1}}{\nu}\,(\nu \in \mathbb{N})$ (sowie $a_0 = 0$). Hier ist

$$\overline{\lim_{\nu\to\infty}} |a_\nu|^{1/\nu} = \lim_{\nu\to\infty} \frac{1}{\sqrt[\nu]{\nu}} = 1 \, ,$$

d. h. $R = 1$. Also konvergiert die Potenzreihe absolut in ihrem Konvergenzkreis $U_1(0) = \{z : |z| < 1\}$ und divergiert für alle $|z| > 1$. In B. 14.12.1 hatten wir gesehen, dass Konvergenz für alle $x \in \mathbb{R}$ mit $|x| < 1$ vorliegt (und dass

$$\ln(1+x) = \sum_{\nu=1}^{\infty} \frac{(-1)^{\nu-1} x^\nu}{\nu}$$

gilt).

4. Die Potenzreihe $\sum\limits_{\nu=0}^{\infty} \nu! z^\nu$ hat den Konvergenzradius $R = 0$. Also ist die Reihe nach S. 16.2.3 für alle $z \neq 0$ divergent.

Wir betrachten noch einmal die geometrische Reihe $\sum\limits_{\nu=0}^{\infty} z^\nu$. Es gilt

$$\sup_{|z|<1} |s_n(z) - s_{n-1}(z)| \geq \sup_{0<x<1} s_n(x) - s_{n-1}(x) = \sup_{0<x<1} x^n = \lim_{x\to 1^-} x^n = 1\ .$$

Also kann nach B. 15.4.2 und dem Cauchy-Kriterium die Konvergenz nicht gleichmäßig auf dem Konvergenzkreis $U_1(0) = \{z : |z| < 1\}$ sein.

Es gilt jedoch

Satz 16.5 *Es sei* $\sum\limits_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu$ *eine Potenzreihe mit Konvergenzradius* $R > 0$. *Dann konvergiert* $\sum\limits_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu$ *lokal gleichmäßig auf dem Konvergenzkreis* $U_R(z_0)$.

Beweis. Es sei $K \subset U_R(z_0)$ kompakt. Dann existiert nach S. 11.12 ein $z_1 \in K$ mit $|z_1 - z_0| = \max\limits_{z\in K} |z - z_0| =: r$. Also ist $r < R$ und nach S. 16.2 ist $\sum\limits_{\nu=0}^{\infty} |a_\nu| r^\nu = \sum\limits_{\nu=0}^{\infty} |a_\nu (z_1 - z_0)^\nu|$ konvergent. Weiter gilt für alle $z \in K$ und alle $\nu \in \mathbb{N}_0$

$$|a_\nu (z - z_0)^\nu| \leq |a_\nu| r^\nu\ .$$

Also folgt aus S. 15.6 die gleichmäßige Konvergenz der Potenzreihe auf K. Da jeder Punkt $z \in U_R(z_0)$ eine kompakte Umgebung in $U_R(z_0)$ besitzt (etwa $\overline{U_\delta(z)}$ mit $0 < \delta < R - |z - z_0|$), ergibt sich damit die Behauptung. □

Nach B./D. 15.11 stellt jede Potenzreihe eine in ihrem Konvergenzkreis stetige Funktion dar. Viel weitergehend erhält man unter der Ausnutzung von B. 15.15:

Satz 16.6 *Es sei* $\sum\limits_{\nu=0}^{\infty} a_\nu (x-x_0)^\nu$ *eine Potenzreihe in* $\mathbb{R}$ *mit Konvergenzradius* $R > 0$. *Ist* $I := (x_0 - R, x_0 + R)$ *und ist* $A(x) := \sum\limits_{\nu=0}^{\infty} a_\nu (x-x_0)^\nu$ *für* $x \in I$, *so ist* $A \in C^\infty(I)$ *und es gilt*

$$a_\nu = \frac{A^{(\nu)}(x_0)}{\nu!} \qquad (\nu \in \mathbb{N}_0)$$

d. h. die Potenzreihe $\sum\limits_{\nu=0}^{\infty} a_\nu (x-x_0)^\nu = \sum\limits_{\nu=0}^{\infty} \frac{A^{(\nu)}(x_0)}{\nu!} (x-x_0)^\nu$ *ist die Taylor-Reihe von* A.

Beweis. Ist $s_n(x) = \sum\limits_{\nu=0}^{n} a_\nu (x-x_0)^\nu$ die n-te Teilsumme von $\sum\limits_{\nu=0}^{\infty} a_\nu (x-x_0)^\nu$, so gilt

$$s_n'(x) = \sum_{\nu=1}^{n} \nu a_\nu (x-x_0)^{\nu-1} = \sum_{\nu=0}^{n-1} (\nu+1) a_{\nu+1} (x-x_0)^\nu \qquad (x \in \mathbb{R})\ .$$

Weiter ist

$$\overline{\lim_{\nu\to\infty}} |\nu a_\nu|^{1/\nu} = \lim_{\nu\to\infty} \nu^{1/\nu} \cdot \overline{\lim_{\nu\to\infty}} |a_\nu|^{1/\nu} = \overline{\lim_{\nu\to\infty}} |a_\nu|^{1/\nu} = 1/R\ .$$

Also hat auch die Potenzreihe $\sum\limits_{\nu=0}^{\infty} \nu a_\nu (x-x_0)^\nu$ den Konvergenzradius R und damit auch die Potenzreihe $\sum\limits_{\nu=0}^{\infty} (\nu+1) a_{\nu+1} (x-x_0)^\nu$ (warum?).
Nach S. 16.5 konvergiert die Folge der Teilsummen, also (s_n'), lokal gleichmäßig auf I und nach B. 15.15.2 ist folglich A differenzierbar auf I mit

$$A'(x) = \lim_{n\to\infty} s_n'(x) = \sum_{\nu=0}^{\infty} (\nu+1) a_{\nu+1} (x-x_0)^\nu\ .$$

Durch Anwendung der gleichen Argumentation auf $A'(x)$ sieht man: A'' existiert auf I und es gilt

$$A''(x) = \sum_{\nu=0}^{\infty} (\nu+1)(\nu+2) a_{\nu+2} (x-x_0)^\nu$$

auf I. Induktiv erhält man: Für alle $k \in \mathbb{N}_0$ existiert $A^{(k)}$ auf I und es ist

$$A^{(k)}(x) = \sum_{\nu=0}^{\infty} (\nu+1)(\nu+2)\cdots(\nu+k) a_{\nu+k} (x-x_0)^\nu$$

auf I. Also ist $A \in C^\infty(I)$ und für $x = x_0$ erhält man

$$A^{(k)}(x_0) = k! a_k \qquad (k \in \mathbb{N}_0)\ .$$

□

Bemerkung 16.7 S. 16.6 besagt insbesondere, dass Funktionen die durch Potenzreihen dargestellt werden, besonders "glatt", d. h. unendlich oft differenzierbar sind. Umgekehrt können C^∞-Funktionen nicht stets durch Potenzreihen dargestellt werden, wie etwa die Funktion

$$f(x) = \begin{cases} e^{-1/x^2} & , \quad x \neq 0 \\ 0 & , \quad x = 0 \end{cases}$$

aus B. 14.11 zeigt.

Am Rande ihres Konvergenzkreises können Potenzreihen ein sehr kompliziertes Verhalten zeigen. Wir geben zum Abschluss ein Ergebnis an, dass die Stetigkeit in einem Randpunkt sichert. Dabei nehmen wir der Einfachheit halber an, dass $z_0 = 0$ und $R = 1$ ist. Der allgemeine Fall ($z_0 \in \mathbb{K}$ und $0 < R < \infty$) kann hierauf leicht zurückgeführt werden.
Bevor wir den Satz formulieren bemerken wir: Ist $\sum\limits_{\nu=0}^{\infty} a_\nu = \sum\limits_{\nu=0}^{\infty} a_\nu 1^\nu$ konvergent, so gilt nach S. 16.2 für den Konvergenzradius R der Potenzreihe $\sum\limits_{\nu=0}^{\infty} a_\nu x^\nu$ jedenfalls $R \geq 1$.

Satz 16.8 *(Abelscher Grenzwertsatz)*
Die Reihe $\sum\limits_{\nu=0}^{\infty} a_\nu$ *in* $\mathbb{K}$ *sei konvergent. Ist*

$$A(x) := \sum_{\nu=0}^{\infty} a_\nu x^\nu \qquad (-1 < x < 1)\ ,$$

so gilt

$$\lim_{x \to 1^-} A(x) = \sum_{\nu=0}^{\infty} a_\nu\ ,$$

d. h. durch $A(1) := \sum\limits_{\nu=0}^{\infty} a_\nu$ *lässt sich* A *stetig auf* $(-1, 1]$ *fortsetzen.*

Beweis. Es sei

$$s := \sum_{\nu=0}^{\infty} a_\nu\ , \quad s_n := \sum_{\nu=0}^{n} a_\nu \qquad (n \in \mathbb{N}_0)\ .$$

Da (nach S. 16.2) für $|x| < 1$ die Potenzreihen

$$A(x) = \sum_{\nu=0}^{\infty} a_\nu x^\nu \qquad \text{und} \qquad \frac{1}{1-x} = \sum_{\nu=0}^{\infty} x^\nu$$

beide absolut konvergieren, gilt nach S. 6.26

$$\begin{aligned} A(x) \cdot \frac{1}{1-x} &= \left(\sum_{\nu=0}^{\infty} a_\nu x^\nu \right) \left(\sum_{\nu=0}^{\infty} x^\nu \right) \\ &= \sum_{\nu=0}^{\infty} x^\nu \left(\sum_{\mu=0}^{\nu} a_\mu \right) = \sum_{\nu=0}^{\infty} s_\nu x^\nu\ , \end{aligned}$$

d. h.

$$A(x) = (1-x) \sum_{\nu=0}^{\infty} s_\nu x^\nu\ .$$

Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $N = N_\varepsilon \in \mathbb{N}$ mit $|s_n - s| < \varepsilon/2$ für alle $n \geq N$. Wegen $1 = (1-x)\sum\limits_{\nu=0}^{\infty} x^\nu$ erhalten wir für $0 < x < 1$

$$\begin{aligned} |A(x) - s| &= |(1-x)\sum_{\nu=0}^{\infty}(s_\nu - s)x^\nu| \leq \\ &\leq (1-x)\sum_{\nu=0}^{N-1}|s_\nu - s|x^\nu + \frac{\varepsilon}{2}(1-x)\sum_{\nu=N}^{\infty}x^\nu \\ &\leq (1-x)\sum_{\nu=0}^{N-1}|s_\nu - s|x^\nu + \varepsilon/2 \end{aligned}$$

Weiter existiert ein $\delta = \delta(N_\varepsilon) > 0$ so, dass

$$(1-x)\sum_{\nu=0}^{N-1}|s_\nu - s|x^\nu < \varepsilon/2$$

für alle x mit $1 - \delta < x < 1$. Also gilt

$$|A(x) - s| < \varepsilon \qquad \text{für} \qquad 1 - \delta < x < 1\,.$$

□

Der Abelsche Grenzwertsatz hat verschiedene interessante Anwendungen. U. a. ist er dazu geeignet, Summen gewisser konvergenter Reihen zu berechnen.

Beispiel 16.9 1. Nach B. 14.12.1 gilt

$$\ln(1+x) = \sum_{\nu=1}^{\infty}\frac{(-1)^{\nu-1}x^\nu}{\nu} \qquad (-1 < x < 1)\,.$$

Außerdem konvergiert die Reihe $\sum\limits_{\nu=1}^{\infty}\frac{(-1)^{\nu-1}}{\nu}$ nach dem Leibnizkriterium . Also erhalten wir mit dem Abelschen Grenzwertsatz

$$\sum_{\nu=1}^{\infty}\frac{(-1)^{\nu-1}}{\nu} = \lim_{x\to 1^-}\sum_{\nu=0}^{\infty}\frac{(-1)^{\nu-1}x^\nu}{\nu} = \lim_{x\to 1^-}\ln(1+x) = \ln(2)\,.$$

2. Wir betrachten die Potenzreihe $\sum\limits_{\nu=0}^{\infty}(-1)^\nu\frac{x^{2\nu+1}}{2\nu+1}$ mit Konvergenzradius 1. Ist

$$A(x) = \sum_{\nu=0}^{\infty}(-1)^\nu\frac{x^{2\nu+1}}{2\nu+1} \qquad (-1 < x < 1)\,,$$

so ist nach S. 16.6

$$A'(x) = \sum_{\nu=0}^{\infty}\left((-1)^\nu\frac{x^{2\nu+1}}{2\nu+1}\right)' = \sum_{\nu=0}^{\infty}(-1)^\nu x^{2\nu} = \frac{1}{1+x^2}$$

für $-1 < x < 1$. Aus $(\arctan x)' = \frac{1}{1+x^2}$ folgt

$$A'(x) - (\arctan x)' \equiv 0 \qquad \text{auf} \quad (-1,1)\ ,$$

also ist nach S. 13.19 mit einer Konstanten $c \in \mathbb{R}$

$$A(x) = \arctan(x) + c \qquad (-1, < x, < 1)\ .$$

Mit $A(0) = 0 = \arctan 0$ ergibt sich $c = 0$, d. h.

$$\arctan x = A(x) = \sum_{\nu=0}^{\infty} (-1)^\nu \frac{x^{2\nu+1}}{2\nu+1} \qquad (-1 < x < 1)\ .$$

Da die Reihe $\sum\limits_{\nu=0}^{\infty} \frac{(-1)^\nu}{2\nu+1}$ nach dem Leibniz-Kriterium konvergiert, ergibt sich aus dem Abelschen Grenzwertsatz

$$\sum_{\nu=0}^{\infty} \frac{(-1)^\nu}{2\nu+1} = \lim_{x\to 1^-} \arctan(x) = \arctan(1) = \frac{\pi}{4}\ .$$

Eine weitere Anwendung ergibt sich im Zusammenhang mit Cauchy-Produkten.

Satz 16.10 *Sind* $\sum\limits_{\nu=0}^{\infty} a_\nu$ *und* $\sum\limits_{\nu=0}^{\infty} b_\nu$ *(nicht notwendig absolut) konvergente Reihen mit der Eigenschaft, das das Cauchy-Produkt* $\sum\limits_{n=0}^{\infty} c_n = \sum\limits_{n=0}^{\infty} \sum\limits_{\nu=0}^{n} a_\nu b_{n-\nu}$ *ebenfalls konvergiert, so gilt*

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{\nu=0}^{\infty} a_\nu\right)\left(\sum_{\nu=0}^{\infty} b_\nu\right)\ .$$

Beweis. Die Potenzreihen $\sum\limits_{\nu=0}^{\infty} a_\nu x^\nu$ $\sum\limits_{\nu=0}^{\infty} b_\nu x^\nu$ und $\sum\limits_{n=0}^{\infty} c_n x^n$ haben alle Konvergenzradius ≥ 1. Nach S. 16.2 konvergieren die Reihen absolut für jedes x mit $|x| < 1$. Also gilt nach S. 6.26

$$\begin{aligned}
\left(\sum_{\nu=0}^{\infty} a_\nu x^\nu\right)\left(\sum_{\nu=0}^{\infty} b_\nu x^\nu\right) &= \sum_{n=0}^{\infty} \sum_{\nu=0}^{n} a_\nu x^\nu b_{n-\nu} x^{n-\nu} \\
&= \sum_{n=0}^{\infty} x^n \sum_{\nu=0}^{n} a_\nu b_{n-\nu} = \sum_{n=0}^{\infty} c_n x^n\ .
\end{aligned}$$

Nach dem Abelschen Grenzwertsatz ist

$$\begin{aligned}
\left(\sum_{\nu=0}^{\infty} a_\nu\right)\left(\sum_{\nu=0}^{\infty} b_\nu\right) &= \left(\lim_{x\to 1^-} \sum_{\nu=0}^{\infty} a_\nu x^\nu\right)\left(\lim_{x\to 1^-} \sum_{\nu=0}^{\infty} b_\nu x^\nu\right) \\
&= \lim_{x\to 1^-} \sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} c_n\ .
\end{aligned}$$

□

17 Integralrechnung von Funktionen einer reellen Veränderlichen

Die Integralrechnung entstand ursprünglich aus der Frage nach Berechnung von Flächeninhalten. Ähnlich wie bei der Differenzialrechnung werden wir Integrale über einen gewissen Grenzprozess definieren. Dazu betrachten wir zunächst besonders einfache Funktionen, für die wir die "Fläche unter den Graphen" sehr leicht definieren können.

Definition 17.1 Es sei $I = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall.

1. Sind $a = x_0 < x_1 < x_2 < \ldots < x_n = b$, so heißt

$$Z = \{x_0, x_1, \ldots, x_n\}$$

eine Zerlegung von $[a, b]$.

2. Eine Funktion $\varphi : [a, b] \to \mathbb{K}$ heißt *Treppenfunktion*, wenn eine Zerlegung $Z = \{x_0, \ldots, x_n\}$ von $[a, b]$ und $c_1, \ldots, c_n \in \mathbb{K}$ existieren mit

$$\varphi(x) \equiv c_j \qquad (x \in (x_{j-1}, x_j),\ j = 1, \ldots, n)\ . \tag{17.1}$$

(Die Funktionswerte an den Stellen x_j spielen keine Rolle). Eine Zerlegung Z so, dass Konstanten c_j wie in (17.1) existieren, heißt *zulässig* für φ.

Definition 17.2 Ist $\varphi : [a, b] \to \mathbb{K}$ eine Treppenfunktion wie in D. 17.1.2, so setzen wir

$$\int_a^b \varphi := \int_a^b \varphi(x)\, dx := \sum_{j=1}^n c_j(x_j - x_{j-1})\ . \tag{17.2}$$

Die Zahl $\int_a^b \varphi(x)dx$ heißt *Integral* von φ (auf $[a, b]$).

!! Wichtig bei dieser Definition: Für eine Treppenfunktion φ gibt es verschiedene zulässige Zerlegungen. Damit $\int_a^b \varphi$ wohldefiniert ist, muss die Summe $\sum\limits_{j=1}^n c_j(x_j - x_{j-1})$ von der Zerlegung unabhängig sein. Man überlegt sich leicht, dass dies der Fall ist ([Ü]). Ist etwa

$$\varphi(x) = \begin{cases} 0 & , \quad 0 \le x \le 1/2 \\ 1 & , \quad 1/2 < x \le 1 \end{cases}$$

so ist $Z = \{0, 1/2, 1\}$ eine zulässige Zerlegung mit $c_1 = 0\ ,\ c_2 = 1$, und es gilt

$$\int_0^1 \varphi = 0 \cdot 1/2 + 1 \cdot 1/2 = 1/2\ .$$

Eine weitere zulässige Zerlegung ist $\{0, 1/2, 3/4, 1\}$ mit $c_1 = 0$, $c_2 = c_3 = 1$. Hierfür gilt auch

$$\int_0^1 \varphi = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} = \frac{1}{2} .$$

Satz 17.3 *Sind $\varphi, \psi : [a, b] \to \mathbb{K}$ Treppenfunktionen, und sind $\alpha, \beta \in \mathbb{K}$, so gilt*

1. $\int_a^b (\alpha\varphi + \beta\psi) = \alpha \int_a^b \varphi + \beta \int_a^b \psi$.
2. *Ist $\mathbb{K} = \mathbb{R}$ und $\varphi \leq \psi$ (d. h. $\varphi(x) \leq \psi(x)$ für alle $x \in [a, b]$), so ist*
$$\int_a^b \varphi \leq \int_a^b \psi .$$
3. *$|\varphi|$ ist eine Treppenfunkton und es gilt*
$$\left| \int_a^b \varphi \right| \leq \int_a^b |\varphi| \leq (b - a)||\varphi|| .$$
(Dabei ist $||\varphi|| := ||\varphi||_{\infty,[a,b]} := ||\varphi||_\infty = \sup \{|\varphi(x)| : x \in [a, b]\}$.)

Beweis. 1. Es seien Z_1 bzw. Z_2 zulässige Zerlegungen für φ bzw. ψ. Dann ist $Z_1 \cup Z_2$ sowohl für φ als auch für ψ zulässig. Ist $Z_1 \cup Z_2 = \{x_0, \dots, x_n\}$ und sind $c_1, \dots, c_n \in \mathbb{K}$ bzw. $d_1, \dots, d_n \in \mathbb{K}$ wie in (17.1) für φ bzw. ψ, so gilt

$$\begin{aligned} \int_a^b \alpha\varphi + \beta\psi &= \sum_{j=1}^n [\alpha c_j(x_j - x_{j-1}) + \beta d_j(x_j - x_{j-1})] = \\ &= \alpha \sum_{j=1}^n c_j(x_j - x_{j-1}) + \beta \sum_{j=1}^n d_j(x_j - x_{j-1}) = \\ &= \alpha \int_a^b \varphi + \beta \int_a^b \psi . \end{aligned}$$

Die Aussagen 2. und 3. ergeben sich ähnlich aus einfachen Eigenschaften von Summen. □

Wir werden nun allgemeinere Funktionen betrachten, die sich in geeigneter Weise durch Treppenfunktionen annähern lassen. Für diese Funktionen können wir dann das Integral (als “Fläche unter dem Graphen”) über die Integrale dieser Treppenfunktionen definieren.

Definition 17.4 Es sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall, und es sei $f : [a,b] \to \mathbb{K}$. Dann heißt f *Regelfunktion*, falls eine Folge (φ_n) von Treppenfunktionen existiert mit $\varphi_n \to f$ gleichmäßig auf $[a,b]$. Weiter setzen wir

$$R[a,b] := \{f : [a,b] \to \mathbb{K} : f \text{ Regelfunktion}\} .$$

Bemerkung 17.5 1. Es gilt: f ist genau dann eine Regelfunktion, wenn zu jedem $\varepsilon > 0$ eine Treppenfunktion φ existiert mit

$$||\varphi - f||_\infty < \varepsilon .$$

(vgl. B. 15.4.2). Insbesondere ist jedes $f \in R[a,b]$ beschränkt auf $[a,b]$.
2. Aus D. 17.4 ergibt sich leicht: Sind f, g Regelfunktionen und sind $\alpha, \beta \in \mathbb{K}$, so sind auch $\alpha f + \beta g$ sowie $f \cdot g$ Regelfunktionen (d. h. $R[a,b]$ ist eine sog. Funktionenalgebra). (Denn: Sind (φ_n) und (ψ_n) Folgen von Treppenfunktionen mit $\varphi_n \to f$, $\psi_n \to g$ gleichmäßig auf $[a,b]$, so gilt

$$||\alpha f + \beta g - (\alpha\varphi_n + \beta\psi_n)|| \leq |\alpha| \cdot ||f - \varphi_n|| + |\beta| \cdot ||g - \psi_n|| \to 0 \quad (n \to \infty) .$$

Da aus $||\psi_n - g|| \to 0$ die Beschränktheit von $(||\psi_n||)_n$ folgt, gilt weiter

$$||fg - \varphi_n\psi_n|| \leq ||\psi_n|| \cdot ||f - \varphi_n|| + ||f|| \cdot ||g - \psi_n|| \to 0 \quad (n \to \infty).)$$

Von besonderer Wichtigkeit ist folgende Charakterisierung der Regelfunktionen.

Satz 17.6 *Es sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall, und es sei $f : [a,b] \to \mathbb{K}$. Dann sind äquivalent*

a) f ist eine Regelfunktion.

b) Für jedes $x_0 \in (a,b)$ existieren die rechts- und die linksseitigen Grenzwerte $f(x_0^+)$ und $f(x_0^-)$ und es existieren $f(a^+)$ sowie $f(b^-)$.

Beweis. a) $\Rightarrow$ b): Es sei $x_0 \in [a,b)$. Wir zeigen: $f(x_0^+)$ existiert. Dazu sei $\varepsilon > 0$ gegeben. Dann existiert eine Treppenfunktion φ mit

$$||f - \varphi||_\infty < \varepsilon/2 .$$

Es existiert ein $t = t_\varepsilon > x_0$, so dass φ konstant auf (x_0, t) ist.
Für beliebige $x, y \in (x_0, t)$ gilt dann

$$|f(x) - f(y)| \leq |f(x) - \varphi(x)| + |\varphi(y) - f(y)| < \varepsilon .$$

Ist (x_n) eine Folge in (x_0, b) mit $x_n \to x_0$, so existiert ein $N_\varepsilon \in \mathbb{N}$ mit $x_n \in (x_0, t)$ für alle $n \geq N_\varepsilon$, also

$$|f(x_n) - f(x_m)| < \varepsilon \qquad (n, m \geq N_\varepsilon)\ .$$

Damit ist $(f(x_n))$ eine Cauchy-Folge, also konvergent in $\mathbb{K}$. Es sei $g := \lim\limits_{n\to\infty} f(x_n)$.
Ist $(\tilde{x}_n)$ eine weitere Folge in (x_0, b) mit $\tilde{x}_n \to x_0$, so existiert genauso $\tilde{g} := \lim\limits_{n\to\infty} f(\tilde{x}_n)$.
Es ist noch zu zeigen: $g = \tilde{g}$. Ist $\varepsilon > 0$ so ist wie oben $\tilde{x}_n \in (x_0, t)$ für n genügend groß. Also folgt

$$|g - \tilde{g}| \leq |g - f(x_n)| + |f(x_n) - f(\tilde{x}_n)| + |f(\tilde{x}_n) - \tilde{g}| < 3\varepsilon$$

für n genügend groß. Da $\varepsilon > 0$ dabei beliebig war, ist $g = \tilde{g}$. Deshalb ist $g = \lim\limits_{x\to x_0^+} f(x)$.
Entsprechend zeigt man die Existenz von $f(x_0^-)$ für alle $x_0 \in [a, b]$.

b) $\Rightarrow$ a): Angenommen, es existiert ein $\varepsilon > 0$, so dass für alle Treppenfunktionen φ gilt $||f - \varphi||_\infty > \varepsilon$.
Wir definieren induktiv eine Intervallschachtellung $([a_n, b_n])$ mit der Eigenschaft

$$(*) \qquad ||f - \varphi||_{\infty,[a_n,b_n]} > \varepsilon \qquad \text{für alle Treppenfunktionen } \varphi\ .$$

Dazu setzen wir $[a_0, b_0] := [a, b]$ und definieren die weiteren Intervalle durch sukzessive Intervall-Halbierung. Ist $[a_n, b_n]$ bereits konstruiert und ist M der Mittelpunkt von $[a_n, b_n]$, so ist

$$||f - \varphi||_{\infty,[a_n,M]} > \varepsilon \;\text{ für alle } \varphi \qquad \text{oder} \qquad ||f - \varphi||_{\infty,[M,b_n]} > \varepsilon \;\text{ für alle } \varphi\ .$$

(Denn: Angenommen, es existieren Treppenfunktionen φ_1 und φ_2 mit $||f - \varphi_1||_{\infty,[a_n,M]} \leq \varepsilon$ und $||f - \varphi_2||_{\infty,[M,b_n]} \leq \varepsilon$. Dann ist durch

$$\varphi(x) := \begin{cases} \varphi_1(x) & , \quad x \in [a_n, M] \\ \varphi_2(x) & , \quad x \in (M, b_n] \end{cases}$$

eine Treppenfunktion auf $[a_n, b_n]$ definiert mit $||f - \varphi||_{\infty,[a_n,b_n]} \leq \varepsilon$.)
Also können wir als $[a_{n+1}, b_{n+1}]$ eines der beiden Intervalle wählen.
Im Beweis zu S. 8.5 hatten wir gesehen, dass ein $x_0 \in [a, b]$ existiert mit $\{x_0\} = \bigcap\limits_{n\in\mathbb{N}} [a_n, b_n]$ (Intervallschachtelungsprinzip).

Wir betrachen den Fall $x_0 \in (a, b)$. Dann existiert ein $\delta = \delta_\varepsilon > 0$, so dass

$$|f(x) - f(x_0^+)| < \varepsilon \qquad (x \in (x_0, x_0 + \delta])$$
$$|f(x) - f(x_0^-)| < \varepsilon \qquad (x \in [x_0 - \delta, x_0))$$

Weiter existiert ein $N \in \mathbb{N}$ mit $[a_N, b_N] \subset [x_0 - \delta, x_0 + \delta]$. Die Treppenfunktion φ auf $[a_N, b_N]$ mit

$$\varphi(x) := \begin{cases} f(x_0^-) & , \quad x \in [a_N, x_0) \\ f(x_0) & , \quad x = x_0 \\ f(x_0^+) & , \quad x \in (x_0, b_N] \end{cases}$$

erfüllt dann

$$||f - \varphi||_{\infty,[a_N,b_N]} \leq \varepsilon$$

im Widerspruch zu $(*)$.
Ist $x_0 = a$ oder $x_0 = b$ argumentiert man analog. □

Satz 17.7 *Es sei* $[a, b] \subset \mathbb{R}$ *und* $f : [a, b] \to \mathbb{K}$.

1. *Ist* f *stetig auf* $[a, b]$*, so ist* f *eine Regelfunktion (d. h.* $C[a, b] \subset R[a, b]$*).*
2. *Ist* f *monoton auf* $[a, b]$*, so ist* f *eine Regelfunktion.*

Beweis.
1. Ergibt sich sofort aus S. 17.6
2. Nach S. 12.15 existieren alle rechts- und linksseitigen Grenzwerte. Also ist $f \in R[a, b]$ nach S. 17.6. □

Definition 17.8 Es sei $f : [a, b] \to \mathbb{K}$ eine Regelfunktion, und es sei (φ_n) eine Folge von Treppenfunktionen mit $\varphi_n \to f$ gleichmäßig auf $[a, b]$. Wir setzen

$$\int_a^b f := \int_a^b f(x)dx := \lim_{n\to\infty} \int_a^b \varphi_n(x)dx \ . \tag{17.3}$$

Dann heisst $\int_a^b f$ das *Integral* von f auf $[a, b]$.
!! Damit (17.3) eine sinnvolle Definition ist, müssen zwei Dinge sichergestellt werden:

1. Der Grenzwert muss existieren.
2. Er hängt nicht von der Wahl der approximierenden Fuktionen (φ_n) ab.

Beides ist erfüllt:
Zu 1.: Es gilt für $m, n \in \mathbb{N}$ nach S. 17.3.3

$$\left| \int_a^b \varphi_n - \int_a^b \varphi_m \right| \leq (b - a)||\varphi_n - \varphi_m|| \ .$$

Da (φ_n) eine gleichmäßige Cauchy-Folge auf $[a,b]$ ist, ist auch $(\int_a^b \varphi_n)$ eine Cauchy-Folge in $\mathbb{K}$, also konvergent.
Zu 2.: Ist (ψ_n) eine weitere Folge von Treppenfunktionen mit $\psi_n \to f$ gleichmäßig auf $[a,b]$, so gilt

$$\left| \int_a^b \varphi_n - \int_a^b \psi_n \right| \leq (b-a)||\varphi_n - \psi_n|| \leq (b-a)[||\varphi_n - f|| + ||f - \psi_n||] \to 0 \; .$$

also auch

$$\lim_{n\to\infty} \int_a^b \varphi_n = \lim_{n\to\infty} \int_a^b \psi_n \; .$$

Wir stellen einige Rechenregeln für das Integral zusammen

Satz 17.9 *Es seien $f, g \in R[a,b]$ und $\alpha, \beta \in \mathbb{K}$. Dann gilt*

1. $\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g$
2. *Ist $\mathbb{K} = \mathbb{R}$ und $f \leq g$ auf $[a,b]$, so ist*
$$\int_a^b f \leq \int_a^b g \; .$$
3. *$|f|$ ist eine Regelfunktion und*
$$\left| \int_a^b f \right| \leq \int_a^b |f| \leq (b-a)||f||$$
4. *Ist $c \in (a,b)$, so gilt*
$$\int_a^b f = \int_a^c f + \int_c^b f \; .$$

Beweis. Es seien (φ_n) und (ψ_n) Folgen von Treppenfunktionen mit $\varphi_n \to f$, $\psi_n \to g$ gleichmäßig auf $[a,b]$.

1. Dann gilt $\alpha\varphi_n + \beta\psi_n \to \alpha f + \beta g$ gleichmäßig auf $[a,b]$ und deshalb

$$\int_b^a \alpha f + \beta g = \lim_{n\to\infty} \int_a^b (\alpha\varphi_n + \beta\psi_n) = \alpha \int_a^b f + \beta \int_a^b g \; .$$

2. Hier sind φ_n und ψ_n reellwertig. Wir setzen

$$\begin{array}{rcl} \varphi_n^- & := & \varphi_n - ||f-\varphi_n|| \\ \psi_n^+ & := & \psi_n + ||g-\psi_n|| \end{array} .$$

Dann sind φ_n^-, ψ_n^+ Treppenfunktionen mit $\varphi_n^- \leq f \leq g \leq \psi_n^+$ sowie

$$||f-\varphi_n^-|| \to 0 \qquad \text{und} \qquad ||g-\psi_n^+|| \to 0 .$$

Also folgt mit S. 17.3.2

$$\int_a^b f = \lim_{n\to\infty} \int_a^b \varphi_n^- \leq \lim_{n\to\infty} \int_a^b \psi_n^+ = \int_a^b g .$$

3. Aus $||f(x)| - |\varphi_n(x)|| \leq |f(x) - \varphi_n(x)|$ folgt, dass auch $|f|$ eine Regelfunktion ist. Es sei $u \in \mathbb{C}$ mit $|u| = 1$ und so, dass $u \cdot \int_a^b f \geq 0$ gilt (ist $\int_a^b f = re^{i\varphi}$, so ist $u = e^{-i\varphi}$ geeignet). Mit 2. ergibt sich dann

$$\begin{aligned} \left| \int_a^b f(x)dx \right| &= u \cdot \int_a^b f(x)dx = \mathrm{Re} \int_a^b uf(x)dx + i\mathrm{Im} \int_a^b uf(x)dx = \\ &= \int_a^b \mathrm{Re}(uf(x))dx \leq \int_a^b |uf(x)|dx = \int_a^b |f(x)| \leq (b-a)||f|| . \end{aligned}$$

4. [Ü] □

Beispiel 17.10 Wir betrachen $f(x) = x$ auf $[0,1]$. Dann ist durch

$$\varphi_n(x) := \begin{cases} \frac{j-1}{n} & , \quad x \in [\frac{j-1}{n}, \frac{j}{n}),\ j = 1, \ldots, n \\ 1 & , \quad x = 1 \end{cases}$$

eine Folge von Treppenfunktionen auf $[0,1]$ gegeben mit $\varphi_n \to f$ gleichmäßig auf $[0,1]$. Es gilt

$$\int_0^1 \varphi_n = \sum_{j=1}^n \frac{j-1}{n}\frac{1}{n} = \frac{n(n-1)}{2n^2} \to \frac{1}{2} \quad (n \to \infty) ,$$

also ist

$$\int_0^1 f(x)dx = 1/2.$$

Bemerkung 17.11 Wir setzen noch für $a < b$ und $f \in R[a,b]$

$$\int_a^a f := 0 \qquad \text{und} \qquad \int_b^a f := -\int_a^b f$$

Wir wollen nun verschiedene Techniken zur Berechnung von Integralen herleiten. Wichtig ist dabei der Begriff der Stammfunktion.

Definition 17.12 Es sei $I \subset \mathbb{R}$ ein Intervall, und es sei $f : I \to \mathbb{K}$. Eine Funktion $F : I \to \mathbb{K}$ heißt *Stammfunktion* von f, falls $F' = f$ gilt.

Damit gilt folgender zentrale Satz der Analysis.

Satz 17.13 *(Hauptsatz der Differenzial- und Integralrechnung)*
Es sei $[a,b] \subset \mathbb{R}$ und $f \in R[a,b]$.

1. *Die Funktion $F : [a,b] \to \mathbb{K}$ mit*

$$F(x) := \int_a^x f(t)dt \qquad (x \in [a,b])$$

ist stetig auf $[a,b]$.

2. *Ist f stetig an der Stelle $x_0 \in [a,b]$, so ist F differenzierbar an x_0 mit*

$$F'(x_0) = f(x_0)\ .$$

Insbesondere gilt also: Ist f stetig auf $[a,b]$, so ist F eine Stammfunktion von f.

3. *Ist Φ eine beliebige Stammfunktion und ist f stetig auf $[a,b]$, so gilt*

$$\int_a^b f = \Phi(b) - \Phi(a) \quad (=: \Phi(x)|_a^b)\ . \tag{17.4}$$

Beweis. 1. Nach B. 17.5 ist f beschränkt, d. h. $||f|| < \infty$. Nach S. 17.9.3 und 4. gilt für $x, y \in [a,b]$ mit $x < y$

$$|F(y) - F(x)| = \left| \int_a^y f - \int_a^x f \right| = \left| \int_x^y f \right| \leq ||f||(y-x)\ .$$

Insbesondere ist F stetig auf $[a,b]$.

2. Es sei f stetig an x_0 und (x_n) eine Folge in $[a,b]$ mit $x_n \to x_0, x_n \neq x_0$. Ist $\varepsilon > 0$ gegeben, so existiert ein $\delta > 0$ mit $|f(x) - f(x_0)| < \varepsilon$ für alle $x \in [a,b]$ mit $|x - x_0| < \delta$. Für alle n mit $|x_n - x_0| < \delta$ gilt dann (beachte: $f(x_0) = \frac{1}{x-x_0}\int_{x_0}^{x} f(x_0)dt$)

$$\left|\frac{F(x_n) - F(x_0)}{x_n - x_0} - f(x_0)\right| = \left|\frac{1}{x_n - x_0}\int\limits_{x_0}^{x_n}\Big(f(t) - f(x_0)\Big)dt\right| \leq \frac{|x_n - x_0|}{|x_n - x_0|}\varepsilon = \varepsilon\ .$$

Da $\varepsilon > 0$ beliebig war, gilt $F'(x_0) = f(x_0)$.

3. Es sei Φ eine Stammfunktion zu f. Dann gilt

$$\Phi'(x) = f(x) = F'(x) \qquad (x \in [a,b])\ .$$

Also ist $(\Phi - F)'(x) \equiv 0$ und damit $(\Phi - F)(x) \equiv c$ für eine Konstante $c \in \mathbb{K}$ (S. 13.17 und [Ü] für $\mathbb{K} = \mathbb{C}$). Also gilt

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

□

Der Hauptsatz der Differenzial- und Integralrechnung (kurz HDI) zeigt, dass Integrale – unter geeigneten Voraussetzungen – durch Bestimmung einer Stammfunktion berechnet werden können!
Man beachte dabei: Sind F_1 und F_2 Stammfunktionen zu f, so exstiert ein $c \in \mathbb{K}$ mit $F_1(x) = F_2(x) + c$ für alle $x \in [a,b]$, d. h. zwei Stammfunktionen unterscheiden sich lediglich durch eine additive Konstante.
Wir schreiben auch

$$\int f(x)dx$$

(*unbestimmtes Integral*) für eine Stammfunktion (oder auch die Gesamtheit aller Stammfunktionen) von f.

Beispiel 17.14 1. Es sei $f(x) = \frac{1}{x}$ $(x > 0)$. Dann gilt

$$(\ln x)' = \frac{1}{x} = f(x) \qquad (x > 0)\ .$$

Also ist

$$\ln x = \int \frac{1}{x}\,dx \qquad (x > 0)\ .$$

Nach dem HDI gilt für $0 < a < b < \infty$:

$$\int\limits_a^b f = \int\limits_a^b \frac{1}{x}\,dx = \ln x|_a^b = \ln(b) - \ln(a) \quad \left(= \ln\left(\frac{b}{a}\right)\right)\ .$$

2. Es sei $f(x) = x^\alpha$ $(x > 0)$, wobei $\alpha \in \mathbb{R}$, $\alpha \neq -1$ fest ist. Dann gilt

$$\left(\frac{1}{\alpha+1}x^{\alpha+1}\right)' = x^\alpha = f(x) \qquad (x > 0) \ ,$$

also ist

$$\int x^\alpha dx = \frac{1}{\alpha+1}x^{\alpha+1} \quad (x > 0)$$

und folglich für $0 < a < b < \infty$

$$\int\limits_a^b x^\alpha \ dx = \frac{1}{\alpha+1}x^{\alpha+1}|_a^b = \frac{1}{\alpha+1}\left(b^{\alpha+1} - a^{\alpha+1}\right) \ .$$

Im Falle $\alpha \geq 0$ gilt dies auch für $a = 0$.

Durch Übertragung der Produkt- und der Kettenregel ergeben sich wichtige Techniken zur möglichen Berechnung von Integralen. Wir bemerken zunächst:
Sind f, g differenzierbar auf dem Intervall I, so gilt mit der Produktregel

$$(fg)'(x) = (f'g)(x) + (fg')(x) \qquad (x \in I) \ ,$$

also ist

$$fg = \int (fg)' = \int f'g + \int fg' \quad \text{auf } I. \tag{17.5}$$

Als Konsequenz erhalten wir

Satz 17.15 *(partielle Integration)*
Sind f und $g : [a, b] \to \mathbb{K}$ stetig differenzierbar (d. h. f' und g' sind stetig auf $[a, b]$), so gilt

$$\int\limits_a^b fg' = fg|_a^b - \int\limits_a^b f'g \ .$$

Beweis. Es gilt mit (17.5) und nach S. 17.13

$$\int\limits_a^b fg' + \int\limits_a^b f'g = fg|_a^b \ .$$

□

Beispiel 17.16 1. Für $\alpha \neq -1$ gilt

$$\begin{aligned}
\underset{\substack{\uparrow \\ g'}}{\int_a^b x^\alpha} \; \underset{\substack{\uparrow \\ f}}{\ln x} \, dx &= \frac{1}{\alpha+1} x^{\alpha+1} \ln x \Big|_a^b - \frac{1}{\alpha+1} \int_a^b x^\alpha \, dx = \\
&= \frac{1}{\alpha+1} \left(b^{\alpha+1} \ln b - a^{\alpha+1} \ln a - \frac{1}{\alpha+1} b^{\alpha+1} + \frac{1}{\alpha+1} a^{\alpha+1} \right)
\end{aligned}$$

oder auch kurz mit unbestimmter Integration

$$\begin{aligned}
\int x^\alpha \ln x \, dx &= \frac{1}{\alpha+1} x^{\alpha+1} \ln x - \frac{1}{\alpha+1} \int x^\alpha \, dx = \\
&= \frac{x^{\alpha+1}}{\alpha+1} \left(\ln x - \frac{1}{\alpha+1} \right)
\end{aligned}$$

2. Es gilt auf $(-1, 1)$

$$\begin{aligned}
\int \sqrt{1-x^2} \, dx &= \int 1 \cdot \sqrt{1-x^2} \, dx = x\sqrt{1-x^2} - \int \frac{x(-2x)}{2\sqrt{1-x^2}} \, dx \\
&= x\sqrt{1-x^2} + \int \frac{dx}{\sqrt{1-x^2}} - \int \frac{1-x^2}{\sqrt{1-x^2}} \, dx \\
&= x\sqrt{1-x^2} + \arcsin x - \int \sqrt{1-x^2} \, dx \, ,
\end{aligned}$$

also

$$\int \sqrt{1-x^2} \, dx = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x)$$

Mit Hilfe der Kettenregel sieht man: Ist F eine Stammfunktion zu f auf einem Intervall J und ist $t : I \to J$ differenzierbar, so gilt

$$(F \circ t)' = (F' \circ t)t' = (f \circ t)t'$$

also

$$\int f(t(x))t'(x)dx = F(t(x)) = \int f(t)dt_{|t=t(x)} \tag{17.6}$$

Für bestimmte Integration erhalten wir entsprechend

Satz 17.17 *(Substitutionsregel)*
Es sei $f : [\alpha, \beta] \to \mathbb{K}$ stetig. Weiter sei $t : [a, b] \to \mathbb{R}$ stetig differenzierbar mit $[\alpha, \beta] \supset W(t)$. Dann gilt

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

Beweis. Es sei F eine Stammfunktion zu f (existiert nach S. 17.13; etwa $F(t) = \int_\alpha^t f(s)ds$). Da $(f \circ t)t'$ stetig auf $[a, b]$ ist, gilt nach (17.6) und S. 17.13 (man beachte: (17.4) gilt auch, falls $a \geq b$ ist)

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

□

Beispiel 17.18 1. Es gilt auf $(-1, 1)$

$$\int \frac{2x}{\sqrt{1-x^2}}\,dx \overset{t=x^2}{=} \int \frac{dt}{\sqrt{1-t}}_{|t=x^2} = -2\sqrt{1-t}_{|t=x^2} = -2\sqrt{1-x^2}\,.$$

Weiter erhält man für $-1 < a < b < 1$ mit $t(x) = x^2$

$$\int\limits_a^b \frac{2x}{\sqrt{1-x^2}}\,dx = \int\limits_{a^2}^{b^2} \frac{dt}{\sqrt{1-t}} = -2\sqrt{1-t}\Big|_{a^2}^{b^2} = 2\left(\sqrt{1-a^2} - \sqrt{1-b^2}\right)$$

2. Es gilt auf $I = (0, \pi)$

$$\begin{aligned}
\int \frac{dx}{\sin x} &\overset{S.7.14}{=} \int \frac{dx}{2\sin(x/2)\cos(x/2)} = \\
&= \int \frac{\cos(x/2)}{\sin(x/2)} \frac{1}{2\cos^2(x/2)}\,dx \\
&= \int \frac{1}{\tan(x/2)} \frac{1}{2\cos^2(x/2)}\,dx \\
&\overset{t(x)=\tan(x/2)}{=} \int \frac{dt}{t}_{|t=\tan(x/2)} = \ln(\tan(x/2))
\end{aligned}$$

Bemerkung 17.19 1. Ist $f \in R[a, b]$, so besitzt f i. a. noch **keine** Stammfunktion!
Ist etwa $f(x) = \text{sign}(x)$ auf $[-1, 1]$, so ist $f \in R[a, b]$, aber f besitzt keine Stammfunktion auf $[-1, 1]$
(Denn angenommen, f hätte eine Stammfunktion F. Dann müssste (bis auf eine additive Konstante) $F(x) = |x|$ für alle $x \neq 0$ gelten und damit wäre F nicht differenzierbar an $x = 0$.)
2. Hat f eine Stammfunktion auf $[a, b]$, so gilt i. a. noch **nicht** $f \in R[a, b]$!
Ist etwa

$$f(x) := \begin{cases} \cos(1/x) + 2x\sin(1/x), & x > 0 \\ 0, & x = 0 \end{cases}$$

so ist

$$F(x) := \begin{cases} x^2 \sin(1/x), & x > 0 \\ 0, & x = 0 \end{cases}$$

eine Stammfunktion von f auf $[0, \infty)$, aber f ist keine Regelfunktion auf $[0, 1]$ (denn $f(0^+)$ existiert nicht).

Zum Abschluss befassen wir uns noch kurz mit der Vertauschung von "lim" und Integration für Funktionenfolgen und Funktionenreihen.

Satz 17.20 *Es sei* $[a, b] \subset \mathbb{R}$.

1. *Sind* $f_n \in R[a, b]$ $(n \in \mathbb{N})$ *und konvergiert die Folge* (f_n) *gleichmäßig auf* $[a, b]$ *gegen die Grenzfunktion* f*, so ist* $f \in R[a, b]$ *und es gilt*

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

2. *Sind* $g_\nu \in R[a, b]$ *und konvergiert die Reihe* $\sum_{\nu=0}^{\infty} g_\nu(x)$ *gleichmäßig auf* $[a, b]$*, so gilt*

$$\int_a^b \left\{ \sum_{\nu=0}^{\infty} g_\nu(x) \right\} dx = \sum_{\nu=0}^{\infty} \int_a^b g_\nu(x)\, dx\,.$$

Beweis. 1. Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $N = N_\varepsilon \in N$ mit

$$||f - f_n||_\infty < \frac{\varepsilon}{2} \qquad (n \geq N)\,.$$

Ferner existiert eine Treppenfunktion φ mit

$$||f_N - \varphi||_\infty < \frac{\varepsilon}{2}\,.$$

Also folgt

$$||f - \varphi||_\infty \leq ||f - f_N||_\infty + ||f_N - \varphi||_\infty < \varepsilon\,.$$

Damit ist $f \in R[a, b]$. Außerdem ergibt sich aus S. 17.9.3

$$\left| \int_a^b f - \int_a^b f_n \right| \leq (b - a)||f - f_n||_\infty \to 0 \qquad (n \to \infty)\,.$$

2. Die 2. Behauptung ergibt sich unmittelbar aus 1. durch Anwendung auf die Teilsummenfolge. □

Beispiel 17.21 Wir betrachten die Funktion $F : (0, \infty) \to \mathbb{R}$

$$F(x) := \int_0^x e^{-t^2}\, dt \qquad (x > 0) .$$

(Die Funktion e^{-t^2} besitzt keine "elementare" Stammfunktion.)
Es gilt für festes $x > 0$

$$e^{-t^2} = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu t^{2\nu}}{\nu!} \qquad \text{gleichmäßig auf } [0, x] .$$

Also gilt nach S. 17.20.2

$$\begin{aligned} F(x) &= \int_0^x \left(\sum_{\nu=0}^{\infty} \frac{(-1)^\nu t^{2\nu}}{\nu!} \right) dt = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu}{\nu!} \int_0^x t^{2\nu}\, dt = \\ &= \sum_{\nu=0}^{\infty} \frac{(-1)^\nu}{\nu!} \cdot \frac{x^{2\nu+1}}{2\nu+1} . \end{aligned}$$

Damit haben wir eine Potenzreihenentwicklung für die Funktion F gefunden.

18 Uneigentliche Integrale

Wir haben bisher nur Integrale auf kompakten Intervallen betrachtet. Wir wollen jetzt auch Integrale auf nichtkompakten Intervallen erklären. Dies werden wird dadurch tun, dass beliebige Intervalle durch kompakte "ausschöpfen".

Definition 18.1 Es sei $I = [a, b)$ wobei $-\infty < a < b \leq \infty$. Ferner sei $f : [a, b) \to \mathbb{K}$ so, dass $f \in R[a, B]$ für alle $a < B < b$ gilt. Existiert $\lim\limits_{B \to b^-} \int_a^B f(x)\,dx \in \mathbb{K}$, so heißt

$$\int\limits_a^b f(x)\,dx := \int\limits_a^{b^-} f(x)\,dx := \lim_{B \to b^-} \int\limits_a^B f(x)\,dx$$

uneigentliches Integral von f auf $[a, b)$. Ensprechend definiert man für $-\infty \leq a < b < \infty$ und $f : (a, b] \to \mathbb{K}$ das *uneigentliche Integral*

$$\int\limits_a^b f(x)\,dx := \int\limits_{a^+}^b f(x)\,dx := \lim_{A \to a^+} \int\limits_A^b f(x)\,dx$$

(im Falle der Existenz des Grenzwertes).
Ist $-\infty \leq a < b \leq \infty$ und $f : (a, b) \to \mathbb{K}$, so, dass $\int_{a+}^c f$ sowie $\int_c^{b-} f$ für ein $c \in (a, b)$ existieren, so setzen wir

$$\int\limits_a^b f(x)\,dx := \int\limits_{a^+}^{b^-} f(x)\,dx := \int\limits_{a^+}^c f(x)\,dx + \int\limits_c^{b^-} f(x)\,dx\,.$$

Im Falle der Existenz der entsprechenden Grenzwerte sprechen wir auch von der *Konvergenz* der Integrale.

Bemerkung 18.2 1. Man beachte, dass die letzte Definition unabhängig von der speziellen Wahl von $c \in (a, b)$ ist, d. h. ist $\tilde{c} \in (a, b)$, so existieren $\int_a^c f$ und $\int_c^b f$ genau dann, wenn $\int_a^{\tilde{c}}$ und $\int_{\tilde{c}}^b$ existieren, und es ist dann $\int_a^c f + \int_c^b f = \int_a^{\tilde{c}} f + \int_{\tilde{c}}^b f$.

2. Ist $-\infty < a < b < \infty$ und ist $f \in R[a, b]$, so gilt

$$\int\limits_a^b f = \int\limits_{a^+}^b f = \int\limits_a^{b^-} f = \int\limits_{a^+}^{b^-} f\ ,$$

d. h. die obige Definition stimmt im Falle der Existenz des (eigentlichen) Integrals mit diesem überein.

(Denn: Wir zeigen $\int_a^b f = \int_a^{b^-} f$. Die weiteren Identitäten ergeben sich entsprechend. Nach S. 17.13 (HDI) ist

$$F(x) := \int_a^x f(t)\,dt \qquad (x \in [a,b])$$

stetig auf $[a,b]$. Also ist

$$\int_a^b f = F(b) = \lim_{B\to b^-} F(B) = \lim_{B\to b^-} \int_a^B f = \int_a^{b^-} f\,.)$$

Beispiel 18.3 1. Es sei $\alpha \in \mathbb{R}$. Dann existiert das uneigentliche Integral $\displaystyle\int_{0^+}^1 \frac{dx}{x^\alpha}$ genau dann, wenn $\alpha < 1$ ist.

(Denn: Es gilt für $0 < A < 1$ und $\alpha \neq 1$

$$\int_A^1 \frac{dx}{x^\alpha} = \frac{x^{1-\alpha}}{1-\alpha}\bigg|_A^1 = \frac{1}{1-\alpha} - \frac{A^{1-\alpha}}{1-\alpha}\,.$$

Ist $\alpha < 1$, so existiert

$$\int_{0^+}^1 \frac{dx}{x^\alpha} = \lim_{A\to 0^+} \int_A^1 \frac{dx}{x^\alpha} = \frac{1}{1-\alpha} - \frac{1}{1-\alpha}\lim_{A\to 0^+} A^{1-\alpha} = \frac{1}{1-\alpha}\,.$$

Ist $\alpha > 1$, so existiert der Grenzwert nicht (als eigentlicher Grenzwert).
Ist $\alpha = 1$, so gilt

$$\int_A^1 \frac{dx}{x} = \ln x|_A^1 = -\ln A = \ln(1/A)\,.$$

Also existiert auch hier der Grenzwert nicht. Man beachte weiterhin: Ist $\alpha \le 0$, so existiert, da $x^{-\alpha}$ stetig auf $[0,1]$ (fortsetzbar) ist, auch das eigentliche Integral $\int_0^1 x^{-\alpha}\,dx$. Hier gilt

$$\int_0^1 x^{-\alpha}\,dx = \frac{x^{1-\alpha}}{1-\alpha}\bigg|_0^1 = \frac{1}{1-\alpha} = \int_{0^+}^1 x^{-\alpha}\,dx\,.$$

(vgl. B. 18.2.2).

2. Es sei $\alpha \in \mathbb{R}$. Dann existiert das uneigentliche $\displaystyle\int_1^\infty \frac{dx}{x^\alpha}$ genau dann, wenn $\alpha > 1$ ist.

(Denn: Für $\alpha \neq 1$ gilt

$$\int_1^B \frac{dx}{x^\alpha} = \frac{x^{1-\alpha}}{1-\alpha}\bigg|_1^B = \frac{B^{1-\alpha}}{1-\alpha} - \frac{1}{1-\alpha}\,.$$

Also existiert für $\alpha > 1$

$$\int_1^\infty \frac{dx}{x^\alpha} = \lim_{B\to\infty} \int_1^B \frac{dx}{x^\alpha} = \lim_{B\to\infty} \frac{B^{1-\alpha}}{1-\alpha} - \frac{1}{1-\alpha} = \frac{1}{\alpha-1} \,.$$

Für $\alpha < 1$ existiert der Grenzwert nicht. Für $\alpha = 1$ gilt schließlich

$$\int_1^B \frac{dx}{x} = \ln B \,,$$

also existiert auch hier der Grenzwert nicht.)

3. Aus 1. und 2. folgt, dass für kein $\alpha \in \mathbb{R}$ das uneigentliche Integral $\int_{0^+}^{\infty} \frac{dx}{x^\alpha}$ existiert (denn anderenfalls müssten beide uneigentlichen Integrale $\int_{0^+}^{1} \frac{dx}{x^\alpha}$ und $\int_1^\infty \frac{dx}{x^\alpha}$ existieren).

4. Wir betrachten $f(x) = \frac{1}{\sqrt{1-x^2}}$ $(x \in (-1,1))$. Dann gilt für $B > 0$

$$\int_0^B \frac{dx}{\sqrt{1-x^2}} = \arcsin(x)|_0^B = \arcsin(B)$$

und für $A < 0$

$$\int_A^0 \frac{dx}{\sqrt{1-x^2}} = -\arcsin(A) \,,$$

also ist

$$\int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} = \int_{-1^+}^{1^-} \frac{dx}{\sqrt{1-x^2}} = \lim_{B\to 1^-} \arcsin(B) - \lim_{A\to -1^+} \arcsin(A) = \pi \,.$$

Wir beschäftigen uns nun mit Konvergenzkriterien für uneigentliche Integrale. Dabei formulieren wir die Ergebnisse für Integrale der Form $\int_a^{b^-}$. Entsprechende Aussagen gelten natürlich für die anderen Typen uneigentlicher Integrale.

Satz 18.4 *(Vergleichskriterium)*
Es sei $-\infty < a < b \leq \infty$ und $f, g : [a,b) \to \mathbb{R}$ seien so, dass $f, g \in R[a,B]$ für alle $a < B < b$. Ferner gelte

$$0 \leq f(x) \leq g(x) \qquad (x \in [a,b)) \,.$$

Dann gilt: Existiert $\int_a^{b-} g$, so existiert auch $\int_a^{b-} f$ und es ist

$$\int_a^{b-} f \leq \int_a^{b-} g \, .$$

Beweis. Wir betrachten

$$F(x) := \int_a^x f \, , \qquad G(x) := \int_a^x g \qquad (x \in [a,b)) \, .$$

Dann sind F und G monoton wachsend auf $[a,b)$ und es gilt für $a < B < b$

$$F(B) \leq G(B) \leq \lim_{B \to b^-} \int_a^B g = \int_a^{b^-} g \, .$$

Also existiert auch $\lim\limits_{B\to b^-} F(B) = \int_a^{b^-} f$. (vgl. S. 12.15). □

Satz 18.5 *Es sei $-\infty < a < b \leq \infty$ und $f : [a,b) \to \mathbb{R}$ sei so, dass $f \in R[a,B]$ für alle $a < B < b$. Dann gilt: Existiert $\int_a^{b^-} |f|$, so existiert auch $\int_a^{b^-} f$.*

Beweis. Wir betrachten die Funktion $\varphi : [a,b) \to \mathbb{R}$ mit

$$\varphi(x) = f(x) + |f(x)| \, .$$

Dann ist $\varphi \in R[a,B]$ für alle $a < B < b$ (S. 17.9.3) und

$$0 \leq \varphi(x) \leq 2|f(x)| \qquad (x \in [a,b)) \, .$$

Mit S. 18.4 folgt die Existenz von $\int_a^{b^-} \varphi$ und aus

$$\int_a^B f = \int_a^B \varphi - \int_a^B |f| \qquad (a < B < b)$$

folgt die Existenz von $\int_a^{b^-} f = \lim\limits_{B\to b^-} \int_a^B f$. □

Beispiel 18.6 1. Wir betrachten für $\alpha > 1$ das uneigentliche Integral $\displaystyle\int_1^\infty \frac{\sin x}{x^\alpha}\, dx$.

Es gilt

$$\frac{|\sin x|}{x^\alpha} \leq \frac{1}{x^\alpha} \qquad (x \geq 1) \, .$$

Da $\int\limits_1^\infty \frac{dx}{x^\alpha}$ existiert, folgt die Existenz von $\int\limits_1^\infty \frac{|\sin x|}{x^\alpha}\,dx$ aus S. 18.4. Nach S. 18.5 existiert auch $\int\limits_1^\infty \frac{\sin x}{x^\alpha}\,dx$.

2. Wir betrachten das uneigentliche Integral $\int\limits_1^\infty \frac{\sin x}{x}\,dx$.

Es gilt

$$\begin{aligned}\int\limits_1^B \frac{\sin x}{x}\,dx &= -\frac{1}{x}\cos x\Big|_1^B - \int\limits_1^B \frac{\cos x}{x^2}\,dx\\ &= \cos 1 - \frac{\cos B}{B} - \int\limits_1^B \frac{\cos x}{x^2}\,dx\end{aligned}$$

Wie in 1. sieht man, dass $\int\limits_1^\infty \frac{\cos x}{x^2}\,dx$ existiert. Also existiert auch

$$\lim_{B\to\infty}\int\limits_1^B \frac{\sin x}{x}\,dx = \cos 1 - \int\limits_1^\infty \frac{\cos x}{x^2}\,dx\ ,$$

d. h. $\int\limits_1^\infty \frac{\sin x}{x}\,dx$ existiert.

Man kann zeigen ([Ü]): Das uneigentliche Integral $\int\limits_1^\infty \frac{|\sin x|}{x}\,dx$ existiert nicht! Also: Für uneigentliche Integrale folgt aus der Existenz von $\int_a^{b-} f$ i. a. noch nicht die Existenz von $\int_a^{b^-} |f|$.

Im folgenden Satz wird ein Zusammenhang zwischen der Konvergenz von Reihen und der Konvergenz uneigentlicher Integrale hergestellt:

Satz 18.7 *(Integralkriterium für Reihen)*
Es sei $f : [1,\infty) \to \mathbb{R}$ monoton fallend und es gelte $f(x) \geq 0$ $(x \geq 1)$. Dann existiert $g := \lim\limits_{n\to\infty}\left(\sum\limits_{\nu=1}^{n} f(\nu) - \int_1^{n+1} f\right)$ und es gilt

$$0 \leq g \leq f(1)\ .$$

Beweis. Zunächst folgt aus $f(\nu) \geq f(x) \geq f(\nu+1)$ in $[\nu, \nu+1]$

$$f(\nu) \geq \int\limits_{\nu}^{\nu+1} f(x)\,dx \geq f(\nu+1)\ .$$

Hieraus ergibt sich, dass die Folge (a_n) mit

$$a_n := \sum_{\nu=1}^{n} f(\nu) - \int\limits_{1}^{n+1} f$$

monoton wachsend ist mit $0 \leq a_n \leq f(1) - f(n+1)$.
(Denn: Für $n \geq 2$ ist

$$a_n - a_{n-1} = f(n) - \int\limits_{n}^{n+1} f \geq 0$$

also $a_n \uparrow$. Aus $a_1 = f(1) - \int_1^2 f \geq 0$ folgt $a_n \geq 0$ $(n \in \mathbb{N})$. Weiter ist $a_1 = f(1) - \int_1^2 f \leq f(1) - f(2)$ und aus $a_{n-1} \leq f(1) - f(n)$ folgt

$$a_n = a_n - a_{n-1} + a_{n-1} = f(n) - \int\limits_{n}^{n+1} f + a_{n-1} \leq f(1) - \int\limits_{n}^{n+1} f \leq f(1) - f(n+1)\ .)$$

Hiermit ergibt sich die Behauptung. □

Beispiel 18.8 Wir betrachten $f(x) = \dfrac{1}{x^\alpha}$ für $\alpha > 0$. Dann ist f monoton fallend auf $[1, \infty)$ und $f(x) \geq 0$. Also existiert

$$g = \lim_{n\to\infty} \left(\sum_{\nu=1}^{n} \frac{1}{\nu^\alpha} - \int\limits_{1}^{n+1} \frac{dx}{x^\alpha} \right)\ .$$

Ist $\alpha > 1$, so ist $\lim\limits_{n\to\infty} \int\limits_{1}^{n+1} \dfrac{dx}{x^\alpha} = \int\limits_{1}^{\infty} \dfrac{dx}{x^\alpha} = \dfrac{1}{\alpha - 1}$ und $\zeta(\alpha) = \sum\limits_{\nu=1}^{\infty} \dfrac{1}{\nu^\alpha}$. Nach S. 18.7 ist

$$0 \leq \zeta(\alpha) - \frac{1}{\alpha - 1} \leq 1 \qquad (\alpha > 1)\ .$$

Ist $\alpha = 1$, so ergibt sich die Konvergenz von

$$a_n = \sum_{\nu=1}^{n} \frac{1}{\nu} - \int\limits_{1}^{n+1} \frac{dx}{x} = \sum_{\nu=1}^{n} \frac{1}{\nu} - \ln(n+1)$$

(vgl. [Ü]).

Beispiel 18.9 Das uneigentliche Integral $\int_{0+}^{\infty} e^{-t}t^{x-1}dt$ existiert für jedes $x > 0$.
(Denn: Es gilt

$$\lim_{t\to\infty} \frac{t^{x+1}}{e^t} = 0 \, .$$

Also existiert eine Konstante $M > 0$ so, dass für alle $t \in [1, \infty)$ gilt

$$e^{-t}t^{x-1} \le \frac{M}{t^2}$$

(warum?). Aus der Konvergenz von $\int_1^\infty \frac{dt}{t^2}$ ergibt sich mit dem Vergleichskriterium auch die Konvergenz von $\int_1^\infty e^{-t}t^{x-1}dt$.
Weiter gilt

$$e^{-t}t^{x-1} \le t^{x-1} \qquad (t \in (0, 1]) \, .$$

Aus der Konvergenz von $\int_{0+}^1 t^{x-1}dt$ (siehe B. 18.3.1) folgt wieder mit dem Vergleichskriterium die Konvergenz von $\int_{0+}^1 e^{-t}t^{x-1}dt$. Insgesamt ergibt sich damit die Konvergenz von $\int_{0+}^\infty e^{-t}t^{x-1}dt$.)
Die Funktion $\Gamma : (0, \infty) \to \mathbb{R}$ mit

$$\Gamma(x) := \int_{0+}^{\infty} e^{-t}t^{x-1}dt \qquad (x > 0)$$

heißt (*Euler'sche*) *Gammafunktion*. Durch partielle Integration erhält man leicht die folgende "Funktionalgleichung" für Γ ([Ü]):

$$\Gamma(x + 1) = x \cdot \Gamma(x) \qquad (x > 0) \, . \tag{18.1}$$

Speziell gilt

$$\Gamma(1) = \int_0^\infty e^{-t}dt = \lim_{B\to\infty} -e^{-t}|_0^B = 1 \, ,$$

woraus sich wiederum mit (18.1) induktiv

$$\Gamma(n + 1) = n! \qquad (n \in \mathbb{N})$$

ergibt. Die Gammafunktion "interpoliert" also die Fakultäten; man kann $\Gamma(x)$ als "verallgemeinerte Fakultät" auffassen.

Zum Abschluss wollen wir noch eine Formel herleiten, die das Wachstum von $n!$ für $n \to \infty$ sehr genau beschreibt, die sog. Stirling'sche Formel. Dazu beweisen wir zunächst

Satz 18.10 *(Euler'sche Summenformel)*
Es sei $n \in \mathbb{N}$, und es sei $f : [0, n] \to \mathbb{R}$ stetig differenzierbar. Dann gilt

$$\sum_{\nu=0}^{n} f(\nu) = \int_0^n f + \frac{1}{2}\Big(f(0) + f(n)\Big) + \int_0^n B(x)f'(x)\, dx$$

mit

$$B(x) := x - [x] - 1/2 \,.$$

Beweis. Es gilt mit partieller Integration

$$\begin{aligned}
\int_0^n B(x)f'(x)\,dx &= \sum_{\nu=1}^{n} \int_{\nu-1}^{\nu} (x-\nu+1/2)f'(x)\,dx = \\
&= \sum_{\nu=1}^{n} \left[(x-\nu+1/2)f(x)\Big|_{\nu-1}^{\nu} - \int_{\nu-1}^{\nu} f(x)\,dx \right] = \\
&= \frac{1}{2}\sum_{\nu=1}^{n} \Big(f(\nu) + f(\nu-1) \Big) - \int_0^n f(x)\,dx \\
&= \sum_{\nu=0}^{n} f(\nu) - \frac{1}{2}\Big(f(0) + f(n) \Big) - \int_0^n f(x)\,dx \,.
\end{aligned}$$

□

Für zwei Folgen (a_n) und (b_n) mit $a_n, b_n \neq 0$ schreiben wir

$$a_n \sim b_n \qquad (n \to \infty)$$

falls $a_n/b_n \to 1\,(n \to \infty)$ gilt.

Satz 18.11 *(Stirling'sche Formel)*
Es ist

$$n! \sim n^n e^{-n} \sqrt{2\pi n} \qquad (n \to \infty) \,. \tag{18.2}$$

Beweis. 1. Wir wenden S. 18.10 auf $f(x) = \ln(1+x)$ an.
Dann gilt

$$\sum_{\nu=1}^{n} \log\nu = \sum_{\nu=0}^{n-1} f(\nu) = \int_0^{n-1} \log(1+x)\,dx + \frac{1}{2}\log n + \int_0^{n-1} \frac{B(x)}{1+x}\,dx \,.$$

wobei

$$\int_0^{n-1} \log(1+x)\,dx = (1+x)\Big(\log(1+x) - 1\Big)\Big|_0^{n-1} = n\log n - n + 1 \,.$$

2. Wir zeigen $\int\limits_0^\infty \frac{B(x)}{1+x}\,dx$ konvergiert.

Denn: Zunächst gilt für $R>0$

$$\left|\int\limits_0^R B(x)\,dx\right| \le \Big|\sum_{\nu=1}^{[R]} \underbrace{\int\limits_{\nu-1}^{\nu} B(x)dx}_{=0}\Big| + \int\limits_{[R]}^{R} \underbrace{|B(x)|}_{\le 1/2}\,dx \le 1/2\,.$$

Durch partielle Integration (ist möglich, da B "stückweise stetig") erhalten wir mit $C(y) := \int_0^y B(x)\,dx$

$$\int\limits_0^R \frac{B(x)}{1+x}\,dx = C(x)\frac{1}{1+x}\Big|_0^R + \int\limits_0^R \frac{C(x)}{(1+x)^2}\,dx\,.$$

Da $|C(x)| \le 1/2$ auf $[0,\infty)$ gilt, existiert $\int\limits_0^R \frac{C(x)}{(1+x)^2}\,dx$ nach S. 18.4 und S. 18.5.

Damit existiert auch

$$\int\limits_0^\infty \frac{B(x)}{1+x}\,dx = \lim_{R\to\infty}\int\limits_0^R \frac{B(x)}{1+x}\,dx \left(= \int\limits_0^\infty \frac{C(x)}{(1+x)^2}\,dx\right)\,.$$

3. Nun betrachten wir

$$a_n := \frac{n!}{n^n e^{-n}\sqrt{n}}$$

und zeigen: $a_n \to \sqrt{2\pi}$ $(n\to\infty)$. Dies ist äquivalent zur Behauptung. Es gilt nach 1.

$$\log a_n = \sum_{\nu=1}^{n}\log\nu - n\log n + n - \frac{1}{2}\log n = 1 + \int\limits_0^{n-1}\frac{B(x)}{1+x}\,dx\,,$$

also ist $(\log a_n)$ konvergent nach 2. und damit konvergiert auch (a_n). Wir setzen $a := \lim\limits_{n\to\infty} a_n$. Dann gilt unter Verwendung Wallis-Produktes für $\pi/2$ (siehe [Ü])

$$\begin{aligned}
\sqrt{\pi/2} &= \lim_{n\to\infty}\left[\frac{2\cdot 2}{1\cdot 3}\cdot\frac{4\cdot 4}{3\cdot 5}\cdot\dots\cdot\frac{(2n)(2n)}{(2n-1)(2n+1)}\right]^{1/2} \\
&= \lim_{n\to\infty}\frac{1}{\sqrt{2n+1}}\cdot\frac{2^n n!}{1\cdot 3\cdot 5\cdots(2n-1)} \\
&= \lim_{n\to\infty}\frac{1}{\sqrt{2n+1}}\cdot\frac{(2^n n!)^2}{(2n)!} \\
&= \lim_{n\to\infty}\frac{1}{\sqrt{2n+1}}\frac{2^{2n}(a_n n^n e^{-n}\sqrt{n})^2}{a_{2n}(2n)^{2n}e^{-2n}\sqrt{2n}} \\
&= \lim_{n\to\infty}\frac{1}{\sqrt{2n+1}}\frac{a_n^2}{a_{2n}}\frac{\sqrt{n}}{\sqrt{2}} = \frac{a}{2}
\end{aligned}$$

d. h. $a = \sqrt{2\pi}$.

□

Unter Ausnutzung des Wallis-Produktes erhält man auch

Satz 18.12 *Es ist*

$$\int_0^\infty e^{-u^2}\, du = \frac{\sqrt{\pi}}{2}$$

Beweis. Nach S. 7.7.2 ist $e^t \le \frac{1}{1-t}$ $(t < 1)$, also

$$e^{-x^2} \le \frac{1}{1+x^2} \qquad (x \in \mathbb{R})$$

und damit auch

$$e^{-nx^2} \le \left(\frac{1}{1+x^2}\right)^n \qquad (x \in \mathbb{R}, n \in \mathbb{N})\ .$$

Entsprechend folgt aus $1 + t \le e^t$

$$(1-x^2)^n \le e^{-nx^2} \qquad (0 \le x \le 1)\ .$$

Also erhalten wir

$$\begin{aligned}\int_0^{\pi/2} \sin^{2n+1} t\, dt &\overset{x=\cos t}{=} \int_0^1 (1-x^2)^n\, dx \le \int_0^1 e^{-nx^2}\, dx \\ &\le \int_0^\infty e^{-nx^2}\, dx = \frac{1}{\sqrt{n}} \int_0^\infty e^{-u^2}\, du \le \int_0^\infty \frac{dx}{(1+x^2)^n} \\ &\overset{x=\cot\tan t}{=} \int_0^{\pi/2} \sin^{2n-2} t\, dt\end{aligned}$$

d. h.

$$\sqrt{n} \int_0^{\pi/2} \sin^{2n+1} t\, dt \le \int_0^\infty e^{-u^2}\, du \le \sqrt{n} \int_0^{\pi/2} \sin^{2n-2} t\, dt\ .$$

Mit

$$\int_0^{\pi/2} \sin^{2n+1} t\, dt = \frac{2}{3} \cdot \frac{4}{5} \cdots \frac{2n}{2n+1}$$

und

$$\int_0^{\pi/2} \sin^{2n-2} t\, dt = \frac{\pi}{2} \cdot \frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-3}{2n-2}$$

(siehe [Ü]) ergibt sich durch Quadrieren

$$\frac{n}{2n+1}\left[\frac{2\cdot 2}{1\cdot 3}\cdot\frac{4\cdot 4}{3\cdot 5}\cdots\frac{(2n)(2n)}{(2n-1)(2n+1)}\right] \le \left(\int_0^\infty e^{-u^2}\,du\right)^2$$
$$\le \frac{n}{2n-1}\left(\frac{\pi}{2}\right)^2 \frac{1}{\left[\frac{2\cdot 2}{1\cdot 3}\cdot\frac{4\cdot 4}{3\cdot 5}\cdots\frac{(2n-2)(2n-2)}{(2n-3)(2n-1)}\right]}$$

Nun gilt (Wallis-Produkt; [Ü])

$$\lim_{n\to\infty}\frac{2\cdot 2}{1\cdot 3}\cdot\frac{4\cdot 4}{3\cdot 5}\cdots\frac{(2n)(2n)}{(2n-1)(2n+1)} = \frac{\pi}{2}$$

also erhalten wir für $n \to \infty$

$$\left(\int_0^\infty e^{-u^2}\,du\right)^2 = \frac{\pi}{4}\ .$$

□

19 Differenzialrechnung von Funktionen mehrerer Variablen

In Abschnitt 13 haben wir Ableitungen von Funktionen einer reellen Variablen untersucht. Wir studieren jetzt für $d, m \in \mathbb{N}$ Funktionen $f : M \to \mathbb{R}^m$, wobei $M \subset \mathbb{R}^d$ ist. Die Räume $\mathbb{R}^d$ bzw. $\mathbb{R}^m$ seien im Folgenden stets mit der euklidischen Metrik versehen. Damit stehen uns die Begriffe und Ergebnisse aus den Abschnitten 9 und 10 auch hier zur Verfügung.

Wir bemerken zunächst, dass jede Funktion $f : M \to \mathbb{R}^m$ in m "Koordinatenfunktionen" $f_1, \ldots, f_m : M \to \mathbb{R}$ zerlegt werden kann indem man $f_j(x)$ als die j-te Komponente des Vektors $f(x) \in \mathbb{R}^m$ definiert, d. h.

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{pmatrix} \in \mathbb{R}^m .$$

Bei vielen Untersuchungen kann man sich damit auf den Fall $m = 1$, d. h. auf den Fall reellwertiger Funktionen, beschränken. So ergibt sich etwa sofort aus B. 8.17, dass f genau dann stetig an einer Stelle $x \in M$ ist, wenn dies für alle Koordinatenfunktionen f_j gilt.

Betrachten wir also $f : M \to \mathbb{R}$. Bevor wir zum eigentlichen Begriff der Differenzierbarkeit von Funktionen mehrerer Veränderlicher kommen, untersuchen wir zunächst Ableitungen "in einer Richtung". Es handelt sich dabei tatsächlich um "eindimensionale" Ableitungen wie wir sie aus Abschnitt 13 schon kennen.

Definition 19.1 Ein Vektor $r \in \mathbb{R}^d$ mit $|r| = 1$ heißt *Richtung*.

Insbesondere sind

$$e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_d = (0, \ldots, 0, 1)$$

Richtungen. Unser Ziel ist nun die Untersuchung von Funktionen in einer gegebenen Richtung r, d. h. ist $f : M \to \mathbb{R}$, wobei $M \subset \mathbb{R}^d$, und ist $x^{(0)} \in M$, so betrachten wir die Gerade

$$G_{r,x^{(0)}} := \{x \in \mathbb{R}^d : x = x^{(0)} + t \cdot r \ , \ t \in \mathbb{R}\}$$

und die Funktion $f_{|G_{r,x^{(0)}} \cap M}$, d. h. wir betrachten $g = g_{r,x^{(0)}} : \tilde{M} \to \mathbb{R}$

$$g(t) := f(x^{(0)} + t \cdot r) \qquad (t \in \tilde{M}) \ ,$$

wobei

$$\tilde{M} := \{t \in \mathbb{R} : x^{(0)} + t \cdot r \in M\} \ .$$

Beispiel 19.2 Wir betrachten die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x_1, x_2) = x_1^2 + x_2 \qquad (x_1, x_2 \in \mathbb{R})$$

sowie $x^{(0)} = (0,0)$. Ist $r = (\cos\varphi, \sin\varphi)$, so ist

$$g(t) = f\Big((0,0) + t(\cos\varphi, \sin\varphi)\Big) = t^2\cos^2\varphi + t\sin\varphi \qquad (t \in \mathbb{R}) \ .$$

Ist speziell $\varphi = 0$, d. h. $r = e_1 = (1,0)$, so ist

$$g(t) = f\Big((0,0) + t \cdot (1,0)\Big) = t^2 \qquad (t \in \mathbb{R}) \ .$$

und ist speziell $\varphi = \pi/2$, d. h. $r = e_2 = (0,1)$, so ist

$$g(t) = t \qquad (t \in \mathbb{R}) \ .$$

Definition 19.3 Es sei $M \subset \mathbb{R}^d$ und es sei $f : M \to \mathbb{R}$. Ferner sei $x^{(0)} \in M$. Ist r eine Richtung in $\mathbb{R}^d$, so heißt f *differenzierbar in Richtung r an der Stelle $x^{(0)}$*, falls der Grenzwert

$$\lim_{t\to 0} \frac{f(x^{(0)} + t \cdot r) - f(x^{(0)})}{t} = \lim_{t\to 0} \frac{g(t) - g(0)}{t}$$

(wobei $g(t) = f(x^{(0)} + t \cdot r)$ wie oben) existiert. Wir nennen den Grenzwert *(Richtungs-) Ableitung* von f *in Richtung* r an $x^{(0)}$ und schreiben dafür

$$\partial_r f(x^{(0)}) \qquad \text{oder} \qquad \frac{\partial f}{\partial r}(x^{(0)}) \qquad \text{oder} \qquad f_r(x^{(0)}) \ .$$

Ist speziell $r = e_k$ der k-te Einheitsvektor, so sagen wir, f sei *partiell differenzierbar nach der Variablen* x_k an der Stelle $x^{(0)}$ und schreiben für $\partial_{e_k} f(x^{(0)})$ auch

$$\partial_k f(x^{(0)}) \qquad \text{oder} \qquad \frac{\partial f}{\partial x_k}(x^{(0)}) \qquad \text{oder} \qquad f_{x_k}(x^{(0)}) \ .$$

Dieser Wert heißt dann auch *partielle Ableitung von f nach der Variablen x_k* (an der Stelle $x^{(0)}$).

Bemerkung 19.4 Aus D. 19.3 ergibt sich sofort, dass die Ableitung in Richtung r an der Stelle $x^{(0)}$ genau dann existiert, wenn die Funktion $g = g_{r,x^{(0)}}$ an der Stelle 0 differenzierbar ist (und dann gilt $\partial_r f(x^{(0)}) = g'(0)$). Also handelt es sich bei Richtungs- und partiellen Ableitungen um Ableitungen von Funktionen einer reellen Variablen.

Folglich gelten auch alle Resultate, die wir hierfür in Abschnitt 13 hergeleitet haben.

Besonders einfach ist die Situation für partielle Ableitungen: Ist $x^{(0)} = (x_1^{(0)}, \dots, x_d^{(0)}) \in \mathbb{R}^d$, so ist für $k = 1, \dots, d$

$$\begin{aligned} \partial_k f(x^{(0)}) &= \lim_{t\to 0} \frac{f(x^{(0)} + t \cdot e_k) - f(x^{(0)})}{t} \\ &= \lim_{t\to 0} \frac{f(x_1^{(0)}, \dots, x_{k-1}^{(0)}, x_k^{(0)} + t, x_{k+1}^{(0)}, \dots, x_d^{(0)}) - f(x^{(0)})}{t} \ . \end{aligned}$$

Man sieht, dass $\partial_k f(x^{(0)})$ sich darstellt als die Ableitung der Funktion

$$x_k \mapsto f(x_1, \dots, x_k, \dots, x_d)$$

bei festgehaltenen Variablen $x_1, \dots, x_{k-1}$ und $x_{k+1}, \dots, x_d$ (diese werden sozusagen als Parameter aufgefasst).

Im Falle $d = 2$ schreibt man meist "$f(x, y)$" statt "$f(x_1, x_2)$". In diesem Falle sprechen wir auch von den partiellen Ableitungen nach x bzw. y und schreiben für $(x_0, y_0) \in \mathbb{R}^2$ auch

$$\frac{\partial f}{\partial x}(x_0, y_0) \qquad \text{bzw.} \qquad f_x(x_0, y_0)$$

sowie

$$\frac{\partial f}{\partial y}(x_0, y_0) \qquad \text{bzw.} \qquad f_y(x_0, y_0) \ .$$

Entsprechend schreiben wir im Falle $d = 3$ oft "$f(x, y, z)$" statt "$f(x_1, x_2, x_3)$" und

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \qquad \text{bzw.} \qquad f_x, f_y, f_z \ .$$

Beispiel 19.5 1. Es sei f wie in B. 19.2, d. h. $f(x, y) = x^2 + y$ $(x, y \in \mathbb{R})$. Dann gilt für $r = (\cos\varphi, \sin\varphi)$ und $x^{(0)} = (0, 0)$

$$\partial_r f(0,0) = \left(\frac{\partial f}{\partial r}(0,0) = f_r(0,0)\right) = \lim_{t\to 0} \frac{g(t) - g(0)}{t} = \lim_{t\to 0} \frac{t^2 \cos^2\varphi + t \sin\varphi}{t} = \sin\varphi \ .$$

Weiter gilt für allgemeines $(x_0, y_0) \in \mathbb{R}^2$

$$\partial_1 f(x_0, y_0) \left(= \frac{\partial f}{\partial x}(x_0, y_0) = f_x(x_0, y_0)\right) = 2x_0$$

und

$$\partial_2 f(x_0, y_0) \left(= \frac{\partial f}{\partial y}(x_0, y_0) = f_y(x_0, y_0)\right) = 1$$

2. Es sei $f : \mathbb{R}^3 \to \mathbb{R}$ definiert durch

$$f(x, y, z) = xy^2 z^3 \qquad (x, y, z \in \mathbb{R}) \ .$$

Dann gilt für alle $(x, y, z) \in \mathbb{R}^3$

$$\begin{aligned}
\partial_1 f(x,y,z) \quad & \Big(= \tfrac{\partial f}{\partial x}(x,y,z) \quad = \quad f_x(x,y,z) \Big) \quad = \quad y^2 z^3 \\
\partial_2 f(x,y,z) \quad & \Big(= \tfrac{\partial f}{\partial y}(x,y,z) \quad = \quad f_y(x,y,z) \Big) \quad = \quad 2xyz^3 \\
\partial_3 f(x,y,z) \quad & \Big(= \tfrac{\partial f}{\partial z}(x,y,z) \quad = \quad f_2(x,y,z) \Big) \quad = \quad 3xy^2 z^2 \ .
\end{aligned}$$

3. Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x,y) := \begin{cases} \dfrac{xy}{x^2+y^2} & , \quad (x,y) \neq (0,0) \\ 0 & , \quad (x,y) = (0,0) \end{cases}$$

Dann gilt für $(x_0, y_0) \neq (0,0)$:

$$\partial_1 f(x_0,y_0) = \frac{y_0}{x_0^2+y_0^2} - \frac{2x_0^2 y_0}{(x_0^2+y_0^2)^2}$$

und

$$\partial_2 f(x_0,y_0) = \frac{x_0}{x_0^2+y_0^2} - \frac{2x_0 y_0^2}{(x_0^2+y_0^2)^2} \ .$$

Für $(x_0, y_0) = (0,0)$ gilt

$$\partial_1 f(0,0) = \lim_{t\to 0} \frac{f(t,0) - f(0,0)}{t} = \lim_{t\to 0} 0 = 0$$

und

$$\partial_2 f(0,0) = \lim_{t\to 0} \frac{f(0,t) - f(0,0)}{t} = \lim_{t\to 0} 0 = 0$$

Also existieren die partiellen Ableitungen in allen Punkten (x_0, y_0). Die Funktion f ist allerdings nicht stetig an der Stelle $(0,0)$, denn für $(x_n, y_n) = \left(\frac{1}{n}, \frac{1}{n}\right)$ gilt

$$f(x_n,y_n) = \frac{1/n^2}{1/n^2 + 1/n^2} = \frac{1}{2} \to \frac{1}{2} \neq 0 = f(0,0) \qquad (n \to \infty) \ .$$

Das Beispiel zeigt, dass die Existenz der partiellen Ableitungen i. a. (anders als im Fall $d = 1$) noch nicht die Stetigkeit impliziert. Man beachte allerdings, dass die partiellen Ableitungen in der Nähe von $(0,0)$ unbeschränkt und damit insbesondere unstetig sind (es gilt für $x_0 \neq 0 y_0 \neq 0$

$$\partial_1 f(0,y_0) = 1/y_0 \ , \qquad \partial_2 f(x_0,0) = 1/x_0 \ ).$$

Wir kommen nun zum eigentlichen Begriff der Differenzierbarkeit von Funktionen mehrerer reeller Variablen. Wir werden sehen, dass – wie im eindimensionalen – die Differenzierbarkeit die Stetigkeit impliziert, und wir werden sehen, dass ein sehr enger Zusammenhang zu den partiellen Ableitungen besteht.

Die folgende Definition erweist sich als eine unmittelbare Übertragung der Zerlegungsformel (S. 13.6) in's Mehrdimensionale

Definition 19.6 Es sei $M \subset \mathbb{R}^d$ und es sei $f : M \to \mathbb{R}^m$ (wobei $d, m \in \mathbb{N}$ beliebig). Ferner sei $x^{(0)} \in M^0$. Dann heißt f *differenzierbar* an der Stelle $x^{(0)}$, falls eine Matrix $C \in \mathbb{R}^{m\times d}$ und eine in $x^{(0)}$ stetige Funktion $\varepsilon : M \to \mathbb{R}^m$ so existieren, dass gilt

$$f(x) = f(x^{(0)}) + C \cdot (x - x^{(0)}) + \varepsilon(x)|x - x^{(0)}| \quad (x \in M)$$

und $\varepsilon(x^{(0)}) = 0$. In diesem Fall heißt die Matrix C (*totale*) *Ableitung* von f an $x^{(0)}$.

Bemerkung 19.7 1. Die Differenzierbarkeit von f an der Stelle $x^{(0)}$ bedeutet anschaulich, dass f "in der Nähe" von $x^{(0)}$ gut angenähert werden kann durch die affinlineare Funktion

$$T(x) := f(x^{(0)}) + C \cdot (x - x^{(0)})$$

(nämlich so gut, dass der Fehler, den man dabei begeht, schneller als $|x - x^{(0)}|$ gegen 0 konvergiert für $x \to x^{(0)}$). Im Fall $m = 2$ und $d = 1$, also im Falle einer reellwertigen Funktion von zwei (reellen) Variablen ist der Graph von T eine Ebene im $\mathbb{R}^3$. Diese Ebene stellt die "Tangentialebene" an den Graphen von f im Punkt $x^{(0)}$ dar.

2. Ist $f : M \to \mathbb{R}^m$ mit den Koordinatenfunktionen $f_1, \ldots, f_m : M \to \mathbb{R}$, so gilt: f ist genau dann differenzierbar an $x^{(0)}$, wenn sämtliche Koordinatenfunktionen $f_1, \ldots, f_m$ differenzierbar an $x^{(0)}$ sind.
Denn: Ist f differenzierbar an $x^{(0)}$ und sind $C = (c_{j_k})$ sowie $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)^T$ wie in D. 19.6, so gilt für $j = 1, \ldots, m$ und $x \in M$ mit $c_j := (c_{j1}, \ldots, c_{jd})$

$$f_j(x) - f_j(x^{(0)}) - c_j(x - x^{(0)}) = \varepsilon_j(x)|x - x^{(0)}| \qquad (x \in M)$$

und aus $\varepsilon(x) \to 0\, (x \to x^{(0)})$ folgt auch $\varepsilon_j(x) \to 0$ $(x \to x^{(0)})$. Also ist f_j differenzierbar an $x^{(0)}$.
Sind umgekehrt $f_j\, (j = 1, \ldots, m)$ differenzierbar an $x^{(0)}$, so existieren Vektoren $c_j = (c_{j1}, \ldots, c_{jd}) \in \mathbb{R}^d$ und Funktionen $\varepsilon_j : M \to \mathbb{R}$ mit $\varepsilon_j(x) \to 0 = \varepsilon_j(0)$ $(x \to x^{(0)})$ und

$$f_j(x) = f_j(x^{(0)}) + c_j(x - x^{(0)}) + \varepsilon_j(x)|x - x^{(0)}| \; .$$

Für

$$C := (c_{jk})_{\substack{j=1,\ldots,m \\ k=1,\ldots,d}} \in \mathbb{R}^{m\times d}$$

und $\varepsilon := (\varepsilon_1, \ldots, \varepsilon_m)^T : M \to \mathbb{R}^m$ gilt dann

$$f(x) - f(x^{(0)}) - C(x - x^{(0)}) = \varepsilon(x)|x - x^{(0)}|$$

und aus B. 8.17 ergbit sich $\varepsilon(x) \to 0 = \varepsilon(x)$ für $x \to x^{(0)}$. Also ist f differenzierbar an $x^{(0)}$.

Definition 19.8 Ist $M \subset \mathbb{R}^d$ und ist $f : M \to \mathbb{R}^m$ so, dass für ein $x^{(0)} \in M$ die partiellen Ableitungen der Koordinatenfunktionen $f_1, \ldots, f_m$ nach allen Variablen existieren, so heißt die Matrix

$$J_f(x^{(0)}) := \left(\partial_k f_j(x^{(0)})\right)_{\substack{j=1,\ldots,m \\ k=1,\ldots,d}} \in \mathbb{R}^{m\times d}$$

Jacobi-Matrix von f an $x^{(0)}$. Ist $m = 1$, d. h. $f : M \to \mathbb{R}$, so schreibt man auch

$$\operatorname{grad} f(x^{(0)}) := \nabla f(x^{(0)}) := J_f(x^{(0)}) = \left(\partial_1 f(x^{(0)}), \ldots, \partial_d f(x^{(0)})\right) .$$

Dieser (Zeilen-) Vektor heißt *Gradient* von f an $x^{(0)}$.

Beispiel 19.9 Wir betrachten die Funktion $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit

$$f(x, y) = \begin{pmatrix} e^{xy^2} \\ y \sin x \end{pmatrix} = \begin{pmatrix} f_1(x, y) \\ f_2(x, y) \end{pmatrix}$$

Dann gilt

$$J_f(x, y) = \begin{pmatrix} y^2 e^{xy^2} & 2xye^{xy^2} \\ y \cos x & \sin x \end{pmatrix}$$

und

$$\operatorname{grad} f_1(x, y) = (y^2 e^{xy^2}, 2xye^{xy^2})$$

sowie

$$\operatorname{grad} f_2(x, y) = (y \cos x, \sin x)$$

Wir stellen nun eine Beziehung zwischen Ableitung und partiellen Ableitungen her.

Satz 19.10 *Es seien $M \subset \mathbb{R}^d$ und $f = (f_1, \ldots, f_m)^T : M \to \mathbb{R}^m$ differenzierbar an der Stelle $x^{(0)} \in M^0$. Dann gilt: Die Koordinatenfunktionen $f_1, \ldots, f_m$ sind partiell differenzierbar nach allen Variablen $x_1, \ldots, x_d$ und für die Ableitung C aus D. 19.6 gilt*

$$C = (c_{jk}))_{\substack{j=1,\ldots,m \\ k=1,\ldots,d}} = \left(\partial_k f_j(x^{(0)})\right)_{\substack{j=1,\ldots,m \\ k=1,\ldots,d}} = J_f(x^{(0)}) .$$

Wir schreiben dann auch $f'(x^{(0)}) := Df(x^{(0)}) := C$. (Man beachte dabei: Insbesondere ist C eindeutig bestimmt.)

Beweis. Es seien $C = (c_{jk})$ und ε wie in D. 19.6. Ist $c_j = (c_{j1}, \ldots, c_{jd})$ die j-te Zeile von C und ist $k \in \{1, \ldots, d\}$ fest, so gilt für $t \in \mathbb{R} \setminus \{0\}$ genügend klein

$$\begin{aligned} \frac{f_j(x^{(0)} + te_k) - f_j(x^{(0)})}{t} &= \frac{c_j \cdot te_k + \varepsilon_j(x^{(0)} + te_k)|te_k|}{t} \\ &= c_j \cdot e_k + \varepsilon_j(x^{(0)} + te_k) \operatorname{sign}(t) \\ &= c_{jk} + \varepsilon_j(x^{(0)} + te_k) \operatorname{sign}(t) . \end{aligned}$$

Aus $\varepsilon_j(x^{(0)} + te_k) \to 0$ $(t \to 0)$ folgt die Behauptung. □

Wir benötigen im folgenden einige Hilfsmittel aus der Linearen Algebra.

Satz 19.11 *Es seien $d, m \in \mathbb{N}$.*

1. *Für $A \in \mathbb{R}^{m \times d}$ setzen wir*
$$||A|| := \sup\{|Ax| : x \in \mathbb{R}^d\ ,\ |x| \leq 1\}\ .$$
Dann ist $||A|| < \infty$ und es gilt $|Ax| \leq ||A|| \cdot |x|$ für alle $x \in \mathbb{R}^d$.

2. *Durch $|| \cdot ||$ ist eine Norm auf $\mathbb{R}^{m \times d}$ gegeben (die sog.* Operatornorm*).*

3. *Sind $A \in \mathbb{R}^{m \times d}$ und $B \in \mathbb{R}^{p \times m}$, wobei $p \in \mathbb{N}$, so gilt*
$$||BA|| \leq ||B|| \cdot ||A||\ .$$

Beweis. 1. Es sei $x \in \mathbb{R}^d$ mit $|x| \leq 1$. Dann gilt $x = \sum\limits_{k=1}^{d} x_k e_k$ und es ist $|x_k| \leq 1$ für $k = 1, \ldots, d$. Also folgt

$$|Ax| = \left| \sum_{k=1}^{d} x_k A e_k \right| \leq \sum_{k=1}^{d} |x_k| \cdot |Ae_k| \leq \sum_{k=1}^{d} |Ae_k| =: M\ .$$

Da dies für alle x mit $|x| \leq 1$ gilt ist $||A|| \leq M$. (Wir bemerken für später: Es ist $M =$ Summe der euklidischen Normen der Spalten von A.)
Ist $x \in \mathbb{R}^d \setminus \{0\}$ beliebig, so gilt

$$|Ax| = |x| \cdot \left| A \left(\frac{x}{|x|} \right) \right| \leq ||A|| \cdot |x|$$

(und für $x = 0$ ist $A0 = 0$).

2. [Ü]

3. Es sei $x \in \mathbb{R}^d$ mit $|x| \leq 1$. Dann gilt mit 1.

$$|(BA)x| = |B(Ax)| \leq ||B|| \cdot |Ax| \leq ||B|| \cdot ||A||$$

also ist $||BA|| \leq ||B|| \cdot ||A||$ nach Definition. □

Damit erhalten wir wie im Falle einer Variablen

Satz 19.12 *Es sei $M \subset \mathbb{R}^d$ und es sei $f : M \to \mathbb{R}^m$. Ist f differenzierbar an der Stelle $x^{(0)} \in M$, so ist f stetig an der Stelle $x^{(0)}$.*

Beweis. Mit den Bezeichnungen aus D. 19.6 gilt

$$\begin{aligned} |f(x) - f(x^{(0)})| &\leq |C(x - x^{(0)})| + |\varepsilon(x)| \cdot |x - x^{(0)}| \\ &\leq ||C|| \cdot |x - x^{(0)}| + |\varepsilon(x)| \cdot |x - x^{(0)}| \to 0 \quad (x \to x^{(0)}) \,. \end{aligned}$$

Also ist

$$\lim_{x \to x^{(0)}} f(x) = f(x^{(0)})$$

und damit ist f stetig an $x^{(0)}$. □

In B. 19.5 hatten wir gesehen, dass die Existenz aller partiellen Ableitungen i. a. noch nicht die Stetigkeit, also nach S. 19.12 schon gar nicht die Differenzierbarkeit impliziert. Es gilt jedoch

Satz 19.13 *Es sei $U \subset \mathbb{R}^d$ offen, und es sei $f = (f_1, \ldots, f_m)^T : U \to \mathbb{R}^m$ so, dass die partiellen Ableitungen $\partial_k f_j$ für alle j, k auf U existieren. Ist $x^{(0)} \in U$ so, dass die $\partial_k f_j$ für alle j, k stetig an $x^{(0)}$ sind, so ist f differenzierbar an $x^{(0)}$. Außerdem ist in diesem Fall die Abbildung*

$$U \ni x \mapsto J_f(x) \in \mathbb{R}^{m \times d}$$

stetig an $x^{(0)}$ (wobei $\mathbb{R}^{m \times d}$ mit der von der Operatornorm herkommenden Metrik versehen ist).

Beweis. 1. Nach B. 19.7.2 können wir uns auf den Beweis für den Fall $m = 1$ (also f reellwertig) beschränken.
Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $\delta = \delta_\varepsilon > 0$ mit

$$|\partial_k f(y) - \partial_k f(x^{(0)})| < \frac{\varepsilon}{d} \quad (y \in U_\delta(x^{(0)}),\, k = 1, \ldots, d) \,.$$

Es sei $x \in U_\delta(x^{(0)})$. Wir setzen $h := x - x^{(0)}$. Ist $h = (h_1, \ldots, h_d) = \sum_{k=1}^{d} h_k e_k$, so betrachten wir $v_0 = 0$ und $v_m := \sum_{k=1}^{m} h_k e_k$ $(1 \leq m \leq d)$. Dann gilt (Teleskopsumme!)

$$\begin{aligned} f(x) - f(x^{(0)}) &= f(x^{(0)} + h) - f(x^{(0)}) \\ &= \sum_{k=1}^{d} \Big\{ f(x^{(0)} + v_k) - f(x^{(0)} + v_{k-1}) \Big\} \,. \end{aligned}$$

Wegen $|v_k| \leq |h| < \delta$ ist $x^{(0)} + v_k \in U_\delta(x^{(0)})$ für $k = 0, \ldots, d$. Weiter ist $v_k = v_{k-1} + h_k e_k$. Also existiert nach dem Mittelwertsatz (angewandt auf $t \mapsto f(x^{(0)} + v_{k-1} + t e_k)$) für jedes $k \in \{1, \ldots, d\}$ ein $\theta = \theta_k \in (0, 1)$ mit

$$f(x^{(0)} + v_k) - f(x^{(0)} + v_{k-1}) = h_k \partial_k f(x^{(0)} + v_{k-1} + \theta h_k e_k) \,.$$

Dabei ist $y^{(k)} := x^{(0)} + v_{k-1} + \theta h_k e_k \in U_\delta(x^{(0)})$. Also folgt

$$\begin{aligned}|f(x^{(0)}+h) - f(x^{(0)}) - \mathrm{grad} f(x^{(0)}) \cdot h| &= \\ &\leq \sum_{k=1}^{d} |f(x^{(0)}+v_k) - f(x^{(0)}+v_{k-1}) - \partial_k f(x^{(0)}) h_k| \\ &\leq \sum_{k=1}^{d} |h_k||\partial_k f(y^{(k)}) - \partial_k f(x^{(0)})| \leq \varepsilon |h| \,.\end{aligned}$$

Da $\varepsilon > 0$ beliebig war, ist

$$\lim_{x \to x^{(0)}} \frac{f(x) - f(x^{(0)}) - \ \mathrm{grad}\ f(x^{(0)}) \cdot (x - x^{(0)})}{|x - x^{(0)}|} = 0$$

was äquivalent zur Differenzierbarkeit von f an $x^{(0)}$ ist.

2. Für $k = 1, \ldots, d$ sind nach Voraussetzung (und B. 8.17) die Abbildungen

$$U \ni x \mapsto J_f(x) e_k \in \mathbb{R}^m$$

(also die Spalten von J_f) stetig an $x^{(0)}$. Damit erhalten wir für $x^{(0)} \in U$ mit der Abschätzung aus dem Beweis zu S. 19.11.1

$$||J_f(x) - J_f(x^{(0)})|| \leq \sum_{k=1}^{d} |J_f(x) e_k - J_f(x^{(0)}) e_k| \to 0 \qquad (x \to x^{(0)})$$

Dies impliziert die Stetigkeit von $x \mapsto J_f(x)$ an $x^{(0)}$. □

Beispiel 19.14 Es sei $f : \mathbb{R}^2 \to \mathbb{R}^2$ definiert durch

$$f(x, y) = \begin{pmatrix} e^{xy^2} \\ y \cdot \sin x \end{pmatrix} \qquad (x, y \in \mathbb{R})$$

(vgl. B. 19.9). Dann sind die partiellen Ableitungen

$$\begin{aligned}\partial_1 f_1(x, y) &= y^2 e^{xy^2} \quad , \quad \partial_2 f_1(x, y) = 2xy e^{xy^2} \\ \partial_1 f_2(x, y) &= y \cdot \cos x \ , \quad \partial_2 f_2(x, y) = \sin x\end{aligned}$$

alle stetig auf $\mathbb{R}^2$. Also ist f differenzierbar auf $\mathbb{R}^2$ nach S. 19.13.

Wir befassen uns zum Abschluss noch mit Rechenregeln für die Ableitung. Wie im Eindimensionalen gilt

Satz 19.15 *1. Es sei $M \subset \mathbb{R}^d$, und es seien $f, g : M \to \mathbb{R}^m$ differenzierbar an der Stelle $x^{(0)} \in M$. Dann ist für alle $\alpha, \beta \in \mathbb{R}$ auch $\alpha f + \beta g$ differenzierbar an $x^{(0)}$ und es gilt*

$$(\alpha f + \beta g)'(x^{(0)}) = \alpha f'(x^{(0)}) + \beta g'(x^{(0)})$$

(andere Schreibweise: $J_{\alpha f + \beta g}(x^{(0)}) = \alpha J_f(x^{(0)}) + \beta J_g(x^{(0)})$).

2. (Kettenregel) Es seien $M \subset \mathbb{R}^d$ und $f : M \to \mathbb{R}^m$. Ferner seien $L \subset \mathbb{R}^m$ und $g : L \to \mathbb{R}^p$ mit $L \supset W(f)$. Ist f differenzierbar an $x^{(0)} \in M$ und ist g differenzierbar an $y^{(0)} = f(x^{(0)})$, so ist auch $g \circ f : M \to \mathbb{R}^p$ differenzierbar an $x^{(0)}$ und es gilt

$$(g \circ f)'(x^{(0)}) = g'\big(f(x^{(0)})\big) \cdot f'(x^{(0)})$$

wobei "·" die übliche Matrixmultiplikation bezeichnet
(andere Schreibweise: $J_{g \circ f}(x^{(0)}) = J_g(f(x^{(0)})) \cdot J_f(x^{(0)})$).

Beweis. 1. ergibt sich unmittelbar aus D. 19.6 und der Beweis zu 2. verläuft wörtlich wie der Beweis zur Kettenregel im Eindimensionalen (S. 13.7). □

Beispiel 19.16 Es seien $f : \mathbb{R}^2 \to \mathbb{R}^2$ und $g : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x, y) = \begin{pmatrix} e^x \cos y \\ e^x \sin y \end{pmatrix}, \qquad g(u, v) := u^2 v\ .$$

Dann gilt

$$g(f(x, y)) = (e^x \cos y)^2 e^x \sin y = e^{3x} \cos^2 y \sin y\ ,$$

also

$$\begin{aligned} (g \circ f)'(x, y)\big(&= J_{g \circ f}(x, y) = \operatorname{grad}\,(g \circ f)(x, y)\big) \\ &= \ (3e^{3x} \cos^2 y \sin y\ ,\ e^{3x}(\cos^3 y - 2 \cos y \sin^2 y))\ . \end{aligned}$$

Andererseits gilt auch

$$f'(x, y) = J_f(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix},$$

und

$$g'(u, v) = (\operatorname{grad} g(u, v) = J_g(u, v)) = (2uv, u^2)\ ,$$

also

$$\begin{aligned} g'(f(x, y))f'(x, y) &= \operatorname{grad}\, g(f(x, y)) J_f(x, y) = \\ &= (2e^{2x} \sin y \cos y, e^{2x} \cos^2 y) \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} = \\ &= (3e^{3x} \sin y \cos^2 y, e^{3x}(\cos^3 y - 2 \cos y \sin^2 y))\ . \end{aligned}$$

20 Mittelwert- und Taylorsatz für Funktionen mehrerer Variablen

Ein besonders wichtiger Spezialfall der Kettenregel ergibt sich für differenzierbare Funktionen $f : M \to \mathbb{R}$, wobei $M \subset \mathbb{R}^d$, und $\varphi : I \to M$, wobei $I \subset \mathbb{R}$ ist. Für $f \circ \varphi : I \to \mathbb{R}$ erhalten wir dann

$$\begin{aligned}(f \circ \varphi)'(t) &= f'(\varphi(t))\varphi'(t) = \operatorname{grad} f(\varphi(t)) \cdot \begin{pmatrix} \varphi_1'(t) \\ \vdots \\ \varphi_d'(t) \end{pmatrix} \\ &= \sum_{j=1}^{d} \frac{\partial f}{\partial x_j}(\varphi(t)) \cdot \varphi_j'(t) \qquad (t \in I)\,. \end{aligned} \tag{20.1}$$

Als wichtige Konsequenz erhalten wir

Satz 20.1 *Es sei $M \subset \mathbb{R}^d$ und $f : M \to \mathbb{R}$ sei differenzierbar an $x^{(0)} \in M$. Für alle Richtungen $r \in \mathbb{R}^d$ existiert dann $\partial_r f(x^{(0)})$ und es gilt*

1. $\partial_r f(x^{(0)}) = \operatorname{grad} f(x^{(0)}) \cdot r$,

2. $|\partial_r f(x^{(0)})| \le |\operatorname{grad} f(x^{(0)})|$.

Beweis. 1. Da f differenzierbar an $x^{(0)}$ ist, existiert ein $\delta > 0$ mit $U_\delta(x^{(0)}) \subset M$. Wir betrachten $\varphi : (-\delta, \delta) \to \mathbb{R}^d$ mit

$$\varphi(t) := x^{(0)} + t \cdot r \qquad (t \in (-\delta, \delta))\,.$$

Dann gilt mit (20.1)

$$\partial_r f(x^{(0)}) = (f \circ \varphi)'(0) = \operatorname{grad} f(\varphi(0)) \cdot \varphi'(0) = \operatorname{grad} f(x^{(0)}) \cdot r$$

(beachte dabei: $\varphi_j'(0) = r_j$ für $j = 1, \ldots, d$).

2. Mit der Cauchy-Schwarz'schen Ungleichung ergibt sich aus 1. unmittelbar

$$|\partial_r f(x^{(0)})| = |\operatorname{grad} f(x^{(0)}) \cdot r| \le |\operatorname{grad} f(x^{(0)})| \cdot |r| = |\operatorname{grad} f(x^{(0)})|\,.$$

□

Bemerkung 20.2 Ist $\operatorname{grad} f(x^{(0)}) \neq 0$ und ist $r_0 = \operatorname{grad}^T f(x^{(0)})/|\operatorname{grad} f(x^{(0)})|$ (die sog. Gradientenrichtung), so gilt mit S. 20.1.1

$$\begin{aligned}\partial_{r_0} f(x^{(0)}) &= \operatorname{grad} f(x^{(0)}) \cdot r_0 = \frac{1}{|\operatorname{grad} f(x^{(0)})|} \operatorname{grad} f(x^{(0)}) \cdot \operatorname{grad}^T f(x^{(0)}) \\ &= |\operatorname{grad} f(x^{(0)})|\end{aligned}$$

und entsprechend

$$\partial_{-r_0} f(x^{(0)}) = -|\operatorname{grad} f(x^{(0)})| \; .$$

(Die Gradientenrichtung ist die "Richtung des steilsten Anstiegs" von f und die negative Gradientenrichtung ist die "Richtung des steilsten Abstiegs" von f.)
Außerdem gilt für $r_1 \perp \operatorname{grad} f(x^{(0)})$

$$\partial_{r_1} f(x^{(0)}) = \operatorname{grad} f(x^{(0)}) \cdot r_1 = 0 \, .$$

d. h. die Richtungsableitung senkrecht zur Gradientenrichtung verschwindet. Diese Erkenntnisse werden sich als grundlegend für die mehrdimensionale Optimierung erweisen.

Beispiel 20.3 Wir betrachten $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x, y) = x^2 + y^2 \qquad (x, y \in \mathbb{R}) \; .$$

Dann gilt

$$\operatorname{grad} f(x, y) = (2x, 2y) \; .$$

Ist etwa $(x_0, y_0) = (1, 1)$, so ist für $r = (r_1, r_2)^T$ mit $|r| = 1$

$$\operatorname{grad} f(1, 1) \cdot r = (2, 2) \cdot r = 2r_1 + 2r_2 \; .$$

Für $r_0 = \frac{1}{\sqrt{2}}(1, 1)^T$ ist $\operatorname{grad} f(1, 1) \cdot r_0 = 2\sqrt{2} = |\operatorname{grad} f(1, 1)|$, und für $r_1 = \frac{1}{\sqrt{2}}(1, -1)^T$ ist $\operatorname{grad} f(1, 1) \cdot r_1 = 0$.

Satz 20.4 *(Mittelwertsatz)*
Es sei $M \subset \mathbb{R}^d$ und es sei $f : M \to \mathbb{R}$ stetig auf M und differenzierbar auf M^0. Sind $x, y \in M$ so, dass

$$\overline{x, y} := \{x + t(y - x) : t \in (0, 1)\} \subset M^0$$

gilt, so existiert ein $\xi \in \overline{x, y}$ mit

$$f(y) - f(x) = \operatorname{grad} f(\xi) \cdot (y - x) \; .$$

Beweis. Wir betrachten $g : [0,1] \to \mathbb{R}$, definiert durch

$$g(t) := f(x + t(y-x)) \qquad (t \in [0,1]) \; .$$

Dann ist g stetig auf $[0,1]$ und differenzierbar auf $(0,1)$ mit

$$g'(t) = \operatorname{grad} f(x + t(y-x)) \cdot (y-x) \; .$$

Nach dem "eindimensionalen" Mittelwertsatz (S. 13.16) existieren $\tau \in (0,1)$ mit

$$g'(\tau) = g(1) - g(0) = f(y) - f(x) \; .$$

Also gilt mit $\xi := x + \tau(y-x)$ die Behauptung. □

Definition 20.5 Eine Menge $M \subset \mathbb{R}^d$ heißt *konvex*, falls $\overline{x,y} \subset M$ für alle $x, y \in M$.

Satz 20.6 *Es sei $M \subset \mathbb{R}^d$ konvex und offen, und es sei $f : M \to \mathbb{R}^m$ differenzierbar auf M mit*

$$c = \sup_{x \in M} ||J_f(x)|| < \infty \; .$$

Dann gilt für alle $x, y \in M$

$$|f(x) - f(y)| \leq c|x-y| \; .$$

Beweis. 1. Ist $m = 1$ und sind $x, y \in M$, so existiert nach S. 20.4 ein $\xi \in \overline{x,y}$ mit

$$\begin{aligned} |f(x) - f(y)| &= |\operatorname{grad} f(\xi) \cdot (x-y)| \\ &\leq |\operatorname{grad} f(\xi)| \cdot |x-y| \; . \end{aligned}$$

Aus $|\operatorname{grad} f(\xi)| = ||J_f(\xi)||$ (siehe [Ü]) folgt die Behauptung.
2. Es sei $m \in \mathbb{N}$ beliebig. Für $x, y \in M$ betrachtet man $g : M \to \mathbb{R}$, definiert durch

$$g(u) := (f(x) - f(y))^T f(u) \qquad (u \in M) \; .$$

Dann gilt

$$\operatorname{grad} g(u) = (f(x) - f(y))^T J_f(u) \qquad (u \in M) \; ,$$

also mit Beweisschritt 1. (angewandt auf g)

$$\begin{aligned} |f(x) - f(y)|^2 &= g(x) - g(y) \leq |\operatorname{grad} g(\xi)| \cdot |x-y| \\ &\leq |f(x) - f(y)| \cdot ||J_f(\xi)|| \cdot |x-y| \; . \end{aligned}$$

Ist $f(x) = f(y)$, so ist die Behauptung klar und ist $f(x) \neq f(y)$ ergibt sich die Behauptung nach Division durch $|f(x) - f(y)|$. □

Aus S. 20.6 ergibt sich insbesondere, dass eine auf einer offenen, konvexen Menge differenzierbare Funktion f mit verschwindender Ableitung (d. h. $J_f(x) \equiv 0$) eine Konstante ist. Wir wollen dies für allgemeinere offene Mengen beweisen.

Definition 20.7 Es sei $M \subset \mathbb{R}^d$
1. M heißt *Polygon-zusammenhängend*, falls für alle $x, y \in M$ endlich viele Punkte $x^{(0)}, \dots, x^{(k)} \in M$ existieren mit $x^{(0)} = x, x^{(k)} = y$ und $\overline{x^{(j-1)}, x^{(j)}} \subset M$ für alle $j \in \{1, \dots, k\}$.
2. M heißt *Gebiet*, falls M offen und Polygon-zusammenhängend ist.

Satz 20.8 *Ist $G \subset \mathbb{R}^d$ ein Gebiet und ist $f : G \to \mathbb{R}^m$ so, dass $J_f(x) \equiv 0$ gilt, so ist $f(x) \equiv$ const auf G.*

Beweis. Es genügt, zu zeigen: Für alle $x, y \in G$ gilt $f(x) = f(y)$. Sind $x, y \in G$, so existieren $x^{(0)} = x, x^{(1)}, \dots, x^{(k)} = y \in G$ mit $\overline{x^{(j-1)} x^{(j)}} \subset G\,(j = 1, \dots, k)$. Da $K_j := \overline{x^{(j-1)}, x^{(j)}} \cup \{x^{(j)}, x^{(j-1)}\} \subset G$ kompakt ist, existiert ein $\varepsilon > 0$ mit

$$U_j := \{x : |x - y| < \varepsilon \text{ für ein } y \in K_j\} = \bigcup_{y \in K_j} U_\varepsilon(y) \subset G\,.$$

Außerdem ist U_j offen und konvex. Also ist $f(x^{(j-1)}) = f(x^{(j)})$ nach S. 20.6 (man beachte: f ist differenzierbar auf G). Da dies für $j = 1, \dots, k$ gilt, ist $f(x) = f(y)$. □

Beispiel 20.9 Wir betrachten $f : U \to \mathbb{R}$ mit

$$f(x, y) = \arctan \frac{y}{x} + \arctan \frac{x}{y} \qquad ((x, y) \in U)\,,$$

wobei $U = \bigcup_{j=1}^{4} G_j$ und

$$\begin{aligned} G_1 &= \{(x, y) : x, y > 0\}\,, \qquad & G_2 &= \{(x, y) : x < 0, y > 0\}\,,\\ G_3 &= \{(x, y) : x, y < 0\}\,, \qquad & G_4 &= \{(x, y) : x > 0, y < 0\}\,. \end{aligned}$$

Dann sind G_1, G_2, G_3, G_4 Gebiete. Weiter gilt

$$\frac{\partial f}{\partial x}(x, y) \equiv 0\,, \qquad \frac{\partial f}{\partial y}(x, y) \equiv 0 \qquad \text{auf } U\,.$$

Nach. S. 20.8 ist $f(x, y) \equiv$ const auf jedem der Gebiete G_j. Allerdings ist f nicht $\equiv$ const auf U, denn

$$\begin{aligned} f(x, y) &\equiv f(1, 1) = \pi/2 && \text{auf } G_1\,,\\ f(x, y) &\equiv f(-1, 1) = -\pi/2 && \text{auf } G_2\,,\\ f(x, y) &\equiv f(-1, -1) = \pi/2 && \text{auf } G_3\,,\\ f(x, y) &\equiv f(1, -1) = -\pi/2 && \text{auf } G_4\,. \end{aligned}$$

Wir beschäftigen uns nun mit Ableitungen höherer Ordnung.

Definition 20.10 Es sei $M \subset \mathbb{R}^d$ und $f : M \to \mathbb{R}$. Ferner sei $x^{(0)} \in M$.
1. Sind r, s Richtungen in $\mathbb{R}^d$, und ist $\partial_r f$ definiert auf M und differenzierbar in Richtung s an $x^{(0)}$, so setzen wir

$$(\partial_s \partial_r f)(x^{(0)}) := \partial_s(\partial_r f)(x^{(0)}) \ .$$

Sind allgemeiner $r_1, \ldots, r_\ell$ Richtungen, so definieren wir induktiv

$$(\partial_{r_\ell} \ldots \partial_{r_1} f)(x^{(0)}) := \partial_{r_\ell}(\partial_{r_{\ell-1}} \ldots \partial_{r_1} f)(x^{(0)})$$

im Falle der Existenz der entsprechenden Grenzwerte (*Richtungsableitungen der Ordnung ℓ*). Für $r_1 = \ldots = r_\ell =: r$ schreiben wir kurz $\partial_r^\ell f(x^{(0)})$.

2. Sind Speziell $r_1 = e_{k_1}, \ldots, r_\ell = e_{k_\ell}$, so schreibt man wieder

$$(\partial_{k_\ell} \ldots \partial_{k_1} f)(x^{(0)}) \quad \text{oder} \quad \frac{\partial^\ell f}{\partial x_{k_\ell} \ldots \partial x_{k_1}}(x^{(0)})$$

an Stelle von $(\partial_{r_\ell} \ldots \partial_{r_1} f)(x^{(0)})$ (*partielle Ableitungen der Ordnung ℓ*). Für $k_1 = \ldots = k_\ell =: k$ schreibt man kurz $\partial_k^\ell f(x^{(0)})$ bzw. $\frac{\partial^\ell f}{\partial x_k^\ell}(x^{(0)})$.

(Bei zwei Variablen schreibt man meist (x, y) statt (x_1, x_2) und bei drei Variablen meist (x, y, z) statt (x_1, x_2, x_3).)

Beispiel 20.11 Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x, y) = e^{xy} + x^2 y \qquad (x, y \in \mathbb{R}) \ .$$

Dann gilt

$$\begin{aligned}
\partial_1 f(x, y) &= \frac{\partial f}{\partial x}(x, y) = y e^{xy} + 2xy \\
\partial_2 f(x, y) &= \frac{\partial f}{\partial y}(x, y) = x e^{xy} + x^2 \\
\partial_1^2 f(x, y) &= \frac{\partial^2 f}{\partial x^2}(x, y) = y^2 e^{xy} + 2y \\
\partial_2 \partial_1 f(x, y) &= \frac{\partial^2 f}{\partial y \partial x}(x, y) = e^{xy} + xy e^{xy} + 2x \\
\partial_1 \partial_2 f(x, y) &= \frac{\partial^2 f}{\partial x \partial y}(x, y) = e^{xy} + xy e^{xy} + 2x \\
\partial_2^2 f(x, y) &= \frac{\partial^2 f}{\partial y^2}(x, y) = x^2 e^{xy} \ .
\end{aligned}$$

Man beachte dabei, dass $\partial_2\partial_1 f = \partial_1\partial_2 f$ gilt.
Weiter erhalten wir etwa

$$\begin{aligned}
\partial_1\partial_2\partial_1 f(x,y) &= \frac{\partial^3 f}{\partial x\partial y\partial x}(x,y) = \frac{\partial}{\partial x}(e^{xy}(1+xy)+2x) = \\
&= ye^{xy}(1+xy)+ye^{xy}+2 \\
\partial_2\partial_1^2 f(x,y) &= \frac{\partial^3 f}{\partial y\partial x^2}(x,y) = \frac{\partial}{\partial y}(y^2e^{xy}+2y) = \\
&= 2ye^{xy}+xy^2e^{xy}+2
\end{aligned}$$

also ist $\partial_1\partial_2\partial_1 f = \partial_2\partial_1^2 f$. Man sieht also, dass die Reihenfolge, in der die partiellen Ableitungen gebildet werden, vertauscht werden kann.

Wir beweisen ganz allgemein für partielle Ableitungen der Ordnung 2:

Satz 20.12 *(Schwarz)*
Es sei $U \subset \mathbb{R}^d$ offen und es seien $p, q \in \{1, \dots, d\}$. Ferner sei $f : U \to \mathbb{R}$ so, dass $\partial_p f$, $\partial_q f$ und $\partial_q\partial_p f$ auf U existieren. Ist $\partial_q\partial_p f$ stetig an der Stelle $x^{(0)} \in U$, so existiert auch $\partial_p\partial_q f(x^{(0)})$ und es gilt

$$\partial_p\partial_q f(x^{(0)}) = \partial_q\partial_p f(x^{(0)}) \; .$$

Beweis. O. E. können wir $d = 2$ sowie $p = 1, q = 2$ annehmen. Wir wählen $h_0, k_0 > 0$ so, dass (mit $x^{(0)} = (x_0, y_0)$)

$$(x_0+h, y_0+k) \in U$$

für alle $|h| < h_0$ und $|k| < k_0$ gilt. Weiter betrachten wir die Funktion $g : I[x_0, x_0+h] \to \mathbb{R}$ mit

$$g(t) := f(t, y_0+k) - f(t, y_0) \qquad (t \in I[x_0, x_0+h])$$

wobei h, k wie oben fest gewählt sind. Zwei Anwendungen des Mittelwertsatzes (S. 13.16) zeigen, dass ein $\xi \in I(x_0, x_0+h)$ und ein $\eta \in I(y_0, y_0+k)$ existieren mit

$$\begin{aligned}
\Delta(h,k) &:= f(x_0+h, y_0+k) - f(x_0+h, y_0) - f(x_0, y_0+k) + f(x_0, y_0) = \\
&= g(x_0+h) - g(x_0) = hg'(\xi) = \\
&= h[\partial_1 f(\xi, y_0+k) - \partial_1 f(\xi, y_0)] = hk\partial_2\partial_1 f(\xi, \eta)
\end{aligned}$$

Nun setzen wir $A := \partial_2\partial_1 f(x_0, y_0)$ und geben $\varepsilon > 0$ vor. Sind $|h|$ und $|k|$ genügend klein, so ist

$$|A - \partial_2\partial_1 f(\xi, \eta)| < \varepsilon$$

($\partial_2\partial_1 f$ ist stetig an (x_0, y_0)). Also erhalten wir für $h, k \neq 0$

$$\left| \frac{\Delta(h,k)}{h \cdot k} - A \right| < \varepsilon$$

für $|h|, |k|$ genügend klein. Betrachtet man bei festem h den Grenzwert $\lim\limits_{k\to 0} \frac{\Delta(h,k)}{k}$, so erhält man damit auch

$$\left| \frac{\partial_2 f(x_0+h, y_0) - \partial_2 f(x_0, y_0)}{h} - A \right| < \varepsilon$$

für alle $h \neq 0$ mit $|h|$ genügend klein. Damit existiert $\partial_1 \partial_2 f(x_0, y_0)$ und es gilt

$$\partial_1 \partial_2 f(x_0, y_0) = A \, .$$

□

Bemerkung und Definition 20.13 1. Ist $U \subset \mathbb{R}^d$ offen und ist $n \in \mathbb{N}_0$, so bezeichnen wir mit $C^n(U)$ die Menge aller Funktionen $f : U \to \mathbb{R}$ mit der Eigenschaft, dass die partiellen Ableitungen der Ordnung $(\leq) \, n$ auf U existieren und dort stetig sind.

Durch sukzessive Anwendung von S. 20.12 sieht man: Ist $f \in C^n(U)$ und sind $k_1, \ldots, k_n \in \{1, \ldots, d\}$, so gilt für jede Permutation $p : \{1, \ldots, n\} \to \{1, \ldots, n\}$:

$$\partial_{k_n} \ldots \partial_{k_1} f(x) = \partial_{k_{p(n)}} \ldots \partial_{k_{p(1)}} f(x) \qquad (x \in U) \, ,$$

d. h. die Reihenfolge, in der die partiellen Ableitungen $\partial_{k_1}, \ldots, \partial_{k_n}$ gebildet werden, spielt keine Rolle.

2. Die Aussage von 1. wird i. a. falsch, wenn man auf die Voraussetzung der Stetigkeit der partiellen Ableitungen der Ordnung n verzichtet. So kann man etwa für die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x,y) := \begin{cases} xy \dfrac{x^2 - y^2}{x^2 + y^2} & , \quad (x,y) \neq (0,0) \\ 0 & , \quad (x,y) = (0,0) \end{cases}$$

zeigen: Alle partiellen Ableitungen der Ordnung 2 existieren auf $\mathbb{R}^2$ und sind stetig auf $\mathbb{R}^2 \setminus \{(0,0)\}$, aber es gilt

$$\partial_2 \partial_1 f(0,0) = -1 \, , \quad \partial_1 \partial_2 f(0,0) = 1 \, .$$

Wir wollen nun ein Analogon zum Taylor-Satz für Funktionen mehrerer Variablen herleiten. Dazu beweisen wir vorbereitend

Satz 20.14 *Es sei $U \subset \mathbb{R}^d$ offen und es sei $f \in C^n(U)$. Dann existiert $\partial_r^n f$ auf U für alle Richtungen $r \in \mathbb{R}^d$ und es gilt mit $r = (r_1, \ldots, r_d)$*

$$\partial_r^n f(x) = \sum_{k_1, \ldots, k_n = 1}^{d} (\partial_{k_n} \ldots \partial_{k_1} f)(x) \cdot r_{k_1} \cdots r_{k_n} \, .$$

Beweis. Wir beweisen die Behauptung per Induktion nach n.
$n = 1$: Ist $f \in C^1(U)$, so folgt aus S. 19.13, dass f differenzierbar auf U ist. Aus S. 20.1 erhalten wir

$$\partial_r^1 f(x) = \partial_r f(x) = \operatorname{grad} f(x) \cdot r = \sum_{k_1=1}^{d} \partial_{k_1} f(x) \cdot r_{k_1} \ .$$

$n \to n+1$: Da $f \in C^{n+1}(U)$ ist, folgt $\partial_r^n f \in C^1(U)$ mit der Induktionsvoraussetzung. Also ergibt sich wie oben beim Induktionsanfang

$$\begin{aligned} \partial_r^{n+1} f(x) &= \partial_r(\partial_r^n f)(x) = \operatorname{grad}(\partial_r^n f)(x) \cdot r \\ &= \sum_{k_{n+1}=1}^{d} \sum_{k_1,\ldots,k_n=1}^{d} \partial_{k_{n+1}}(\partial_{k_n} \ldots \partial_{k_1} f)(x) r_{k_1} \cdots r_{k_n} \cdot r_{k_{n+1}} \ , \end{aligned}$$

woraus die Behauptung für $n+1$ folgt. □

Damit erhalten wir

Satz 20.15 *(Taylor-Satz für Funktionen mehrerer Variablen)*
Es sei $U \subset \mathbb{R}^d$ offen und konvex, und es sei $f \in C^{n+1}(U)$ für ein $n \in \mathbb{N}_0$. Ferner sei $x^{(0)} \in U$. Dann existiert zu jedem $x \in U$, $x \neq x^{(0)}$, ein $\xi \in \overline{x, x^{(0)}}$ so, dass mit $r := (x - x^{(0)})/|x - x^{(0)}|$ gilt

$$f(x) = \sum_{\nu=0}^{n} \frac{\partial_r^\nu f(x^{(0)})}{\nu!} |x - x^{(0)}|^\nu + \frac{\partial_r^{n+1} f(\xi)}{(n+1)!} |x - x^{(0)}|^{n+1} \ .$$

Beweis. Wir setzen $h := |x - x^{(0)}|$ und betrachten die Funktion $g : [0, h] \to \mathbb{R}$ mit

$$g(t) := f(x^{(0)} + t \cdot r) \quad (t \in [0, h]) \ .$$

Dann gilt

$$g'(t) = \operatorname{grad} f(x^{(0)} + t \cdot r) \cdot r = (\partial_r f)(x^{(0)} + t \cdot r) \qquad (0 \leq t \leq h)$$

und damit auch allgemein (beachte: $\partial_r^\nu f$ existiert auf U für $\nu = 0, \ldots, n+1$)

$$g^{(\nu)}(t) = (\partial_r^\nu f)(x^{(0)} + t \cdot r) \qquad (0 \leq t \leq h \ , \ \nu = 0, \ldots, n+1) \ .$$

Nach dem eindimensionalen Taylor-Satz (S. 14.3) existiert ein $\tau \in (0, h)$ mit

$$g(h) = \sum_{\nu=0}^{n} \frac{g^{(\nu)}(0)}{\nu!} h^\nu + \frac{g^{(h+1)}(\tau)}{(n+1)!} h^{n+1}$$

d. h. mit $\xi := x^{(0)} + \tau \cdot r$ gilt $\xi \in \overline{x, x^{(0)}}$ und

$$f(x) = \sum_{\nu=0}^{n} \frac{\partial_r^\nu f(x^{(0)})}{\nu!} |x - x^{(0)}|^\nu + \frac{\partial_r^{n+1} f(\xi)}{\nu!} |x - x^{(0)}|^{n+1} \ .$$

□

Bemerkung 20.16 1. Ist $d = 1$, so ist für $x > x^{(0)}$, d. h. $r = 1$,

$$\partial_1^\nu f = f^{(\nu)}$$

und für $x < x^{(0)}$, d. h. $r = -1$, ist

$$\partial_{-1}^\nu f = (-1)^\nu f^{(\nu)} .$$

Damit gilt in beiden Fällen

$$f(x) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x^{(0)})}{\nu!}(x - x^{(0)})^\nu + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x^{(0)})^{n+1}$$

also die Aussage von S. 14.3.
2. Für $n = 0$ erhalten wir

$$f(x) = f(x^{(0)}) + \partial_r f(\xi)|x - x^{(0)}| = f(x^{(0)}) + \operatorname{grad} f(\xi)(x - x^{(0)})$$

also die Aussage von S. 20.4.
3. Für $n = 1$ ergibt sich

$$f(x) = f(x^{(0)}) + \partial_r f(x^{(0)})|x - x^{(0)}| + \frac{1}{2}\partial_r^2 f(\xi)|x - x^{(0)}|^2 .$$

Nach S. 20.14 ist

$$\partial_r^2 f(y) = \sum_{j,k=1}^{d} \partial_k \partial_j f(y) r_j r_k = r^T H_f(y) r ,$$

wobei

$$H_f(y) := \big(\partial_k \partial_j f(y)\big)_{j,k=1,\ldots,d} \in \mathbb{R}^{d\times d}$$

die sog. *Hesse-Matrix* von f an der Stelle y bezeichnet. Es gilt also hier

$$\begin{aligned} f(x) &= f(x^{(0)}) + \operatorname{grad} f(x^{(0)})(x - x^{(0)}) + \frac{1}{2} r^T H_f(\xi) r \cdot |x - x^{(0)}|^2 \\ &= f(x^{(0)}) + \operatorname{grad} f(x^{(0)})(x - x^{(0)}) + \frac{1}{2}(x - x^{(0)})^T H_f(\xi)(x - x^{(0)}) . \end{aligned}$$

Ähnlich wie im Eindimensionalen Fall werden wir als wesentliche Anwendung des Taylor-Satzes ein hinreichendes Kriterium für lokale Extrema herleiten können. Zunächst gilt folgendes wichtige **notwendige** Kriterium für Extremstellen.

Satz 20.17 *Es sei $U \subset \mathbb{R}^d$ offen und es sei $f : U \to \mathbb{R}$. Ferner sei $x^{(0)} \in U$ so, dass* $\operatorname{grad} f(x^{(0)})$ *existiert. Hat f an $x^{(0)}$ ein lokales Extremum, so gilt*

$$\operatorname{grad} f(x^{(0)}) = 0 .$$

Beweis. Hat f an $x^{(0)}$ ein lokales Extremum, so haben auch alle Funktionen $g_k : U_k \to \mathbb{R}$ mit

$$g_k(t) := f(x^{(0)} + te_k) \qquad (t \in U_k := \{s : x^{(0)} + se_k \in U\})$$

ein lokales Extremum an $t_0 = 0$. Also gilt nach S. 13.13 und D. 19.3

$$0 = g_k'(0) = \partial_k f(x^{(0)}) \qquad (k = 1, \ldots, d) \;.$$

□

Beispiel 20.18 1. Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x, y) = x^2 + y^2 \qquad (x, y \in \mathbb{R}) \;.$$

Dann hat f an $(0, 0)$ offenbar ein (sogar globales) Minimum. Es gilt $\operatorname{grad} f(x, y) = (2x, 2y)$, also tatsächlich $\operatorname{grad} f(0, 0) = 0$.
2. Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x, y) = x^2 - y^2 \qquad (x, y \in \mathbb{R}) \;.$$

Dann gilt $\operatorname{grad} f(x, y) = (2x, -2y)$, also $\operatorname{grad} f(0, 0) = (0, 0)$. Trotzdem hat f an $(x_0, y_0) = (0, 0)$ offenbar kein lokales Extremum, d. h. wie im Eindimensionalen ist die Bedingung "$\operatorname{grad} f(x, y) = (0, 0)$" i. a. **nicht hinreichend** für das Vorliegen einer Extremstelle.

Definition 20.19 Es sei $A \in \mathbb{R}^{d \times d}$ symmetrisch. Dann heißt A

1. *positiv definit*, falls $x^T Ax > 0$ für alle $x \in \mathbb{R}^d \setminus \{0\}$ gilt,
2. *positiv semidefinit*, falls $x^T Ax \geq 0$ für alle $x \in \mathbb{R}^d$ gilt,
3. *negativ (semi-) definit*, falls $-A$ positiv (semi-) definit ist,
4. *indefinit*, falls A weder positiv noch negativ semidefinit ist.

Bemerkung 20.20 1. Ist $f \in C^2(U)$, für eine offene Menge U in $\mathbb{R}^d$, so ist die Hesse-Matrix

$$H_f(x) = \big(\partial_k \partial_j f(x)\big)_{j,k=1,\ldots,d} \in \mathbb{R}^{d \times d}$$

nach dem Satz von Schwarz (S. 20.12) für alle $x \in U$ symmetrisch.

2. Für Charakterisierungen der Definitheitsbegriffe aus D. 20.19 verweisen wir auf die Lineare Algebra. Erwähnt werden soll hier nur folgendes: Sind $\lambda_1, \ldots, \lambda_d$ die Eigenwerte von A, so ist

$$A \left\{ \begin{array}{l} \text{positiv definit} \\ \text{positiv semidefinit} \\ \text{negativ definit} \\ \text{negativ semidefinit} \end{array} \right\} \iff \left\{ \begin{array}{l} \lambda_1, \ldots, \lambda_d > 0 \\ \lambda_1, \ldots, \lambda_d \geq 0 \\ \lambda_1, \ldots, \lambda_d < 0 \\ \lambda_1, \ldots, \lambda_d \leq 0 \end{array} \right\} .$$

Es gilt nun folgendes für die mehrdimensionale Optimierung zentrale Resultat

Satz 20.21 *Es sei $U \subset \mathbb{R}^d$ offen, und es sei $f \in C^2(U)$. Ferner sei $x^{(0)} \in U$ ein kritischer Punkt (d. h. $\operatorname{grad} f(x^{(0)}) = 0$). Dann gilt:*

1. *Ist $H_f(x^{(0)})$ positiv definit, so hat f an $x^{(0)}$ ein lokales Minimum.*
2. *Ist $H_f(x^{(0)})$ negativ definit, so hat f an $x^{(0)}$ ein lokales Maximum.*
3. *Hat f an $x^{(0)}$ ein lokales Minimum, so ist $H_f(x^{(0)})$ positiv semidefinit.*
4. *Hat f an $x^{(0)}$ ein lokales Maximum, so ist $H_f(x^{(0)})$ negativ semidefinit.*

Beweis. Es genügt, die Behauptungen 1. und 3. zu beweisen. Die Aussagen 2. und 4. ergeben sich dann durch Betrachtung von $-f$.
Dazu bemerken wir zunächst: Die Funktion $H_f : (U, |\cdot|) \to (\mathbb{R}^{d \times d}, ||\cdot||)$ ist stetig. (Denn: Für $x \in U$ ist

$$H_f(x) = J_{\operatorname{grad} f}(x)$$

und da nach Voraussetzung alle partiellen Ableitungen von $\operatorname{grad} f$ stetig auf U sind, ist H_f stetig auf U nach S. 19.13.)

1. Es sei $H_f(x^{(0)})$ positiv definit. Da die Funktion

$$\mathbb{R}^d \ni r \mapsto r^T H_f(x^{(0)}) r \in \mathbb{R}$$

stetig auf $\mathbb{R}^d$ und $S^{d-1} := \{r : |r| = 1\} \subset \mathbb{R}^d$ beschränkt und abgeschlossen, also kompakt ist, existiert ein $r_0 \in S^{d-1}$ mit

$$r^T H_f(x^{(0)}) r \geq r_0^T H_f(x^{(0)}) r_0 =: \varepsilon > 0 \qquad (r \in S^{d-1}) .$$

Weiter existiert nach der Vorbemerkung ein $\delta = \delta_\varepsilon > 0$ mit $||H_f(y) - H_f(x^{(0)})|| < \varepsilon$ für alle $y \in U_\delta(x^{(0)})$.
Es sei $x \in U_\delta(x^{(0)})$, $x \neq x^{(0)}$. Wir setzen

$$r := \frac{x - x^{(0)}}{|x - x^{(0)}|}.$$

Nach dem Taylor-Satz und B. 20.16.3 existiert ein $\xi \in U_\delta(x^{(0)})$ mit

$$\begin{aligned} \frac{f(x) - f(x^{(0)})}{|x - x^{(0)}|^2} &= \frac{1}{2} r^T H_f(\xi) r \\ &= \frac{1}{2} r^T (H_f(\xi) - H_f(x^{(0)})) r + \frac{1}{2} r^T H_f(x^{(0)}) r \ . \end{aligned}$$

Weiter folgt mit der Cauchy-Schwarz'schen Ungleichung und S. 19.11.1

$$\begin{aligned} |r^T (H_f(\xi) - H_f(x^{(0)})) r| &\leq |r^T| \cdot |(H_f(\xi) - H_f(x^{(0)}) r| \\ &\leq ||H_f(\xi) - H_f(x^{(0)})|| < \varepsilon \ . \end{aligned}$$

Also ist $f(x) > f(x^{(0)})$. Damit liegt an $x^{(0)}$ ein (sogar striktes) lokales Minimum vor.

2. Die Funktion f habe an $x^{(0)}$ ein lokales Minimum. Angenommen, $H_f(x^{(0)})$ ist nicht positiv semidefinit. Dann existiert ein $r_1 \in S^{d-1}$ mit $r_1^T H_f(x^{(0)}) r_1 < 0$. Ist $\rho > 0$ so, dass $U_\rho(x^{(0)}) \subset U$, so gilt für

$$x^{(n)} := x^{(0)} + \frac{\rho}{n} r_1$$

wie in 1. mit einer Folge $(\xi^{(n)})$

$$\frac{f(x^{(n)}) - f(x^{(0)})}{|x^{(n)} - x^{(0)}|^2} = \frac{1}{2} r_1^T (H_f(\xi^{(n)}) - H_f(x^{(0)})) r_1 + \frac{1}{2} r_1^T H_f(x^{(0)}) r_1 \ .$$

Aus $\xi^{(n)} \in \overline{x^{(n)}, x^{(0)}}$ folgt $\xi^{(n)} \to x^{(0)}$ $(n \to \infty)$ und damit $r_1^T (H_f(\xi^{(n)}) - H_f(x^{(0)})) r_1 \to 0 \, (n \to \infty)$. Also ist $f(x^{(n)}) < f(x^{(0)})$ für alle genügend großen n. Widerspruch! □

Beispiel 20.22 1. Ist $f(x, y) = x^2 + y^2$ für $x, y \in \mathbb{R}$ (vgl. B. 20.18.1), so gilt

$$H_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

also ist H_f stets positiv definit. Insbesondere liegt am kritischen Punkt $(0, 0)$ ein lokales Minimum vor.
2. Ist $f(x, y) = x^2 - y^2$ für $x, y \in \mathbb{R}$ (vgl. B. 20.18.2) so gilt

$$H_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \ .$$

Hier ist H_f stets indefinit. Also hat f nach S. 20.21.3./4. keine lokalen Extrema.
3. Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x, y) = 2x^3 - 3x^2 + 2y^3 + 3y^2 \qquad (x, y \in \mathbb{R}) \ .$$

Dann gilt

$$\operatorname{grad} f(x,y) = (6(x^2 - x)\ ,\ 6(y^2 + y)) = (0,0)$$

genau dann, wenn $x \in \{0,1\}$ und $y \in \{0,-1\}$. Also haben wir die kritischen Stellen

$$(0,0)\ ,\ (0,-1)\ ,\ (1,0)\ ,\ (1,-1)\ .$$

Weiter gilt

$$H_f(x,y) = \begin{pmatrix} 12x - 6 & 0 \\ 0 & 12y + 6 \end{pmatrix}$$

Also ist

$$\begin{aligned} H_f(0,0) &= \begin{pmatrix} -6 & 0 \\ 0 & 6 \end{pmatrix} && \text{indefinit} \\ H_f(0,-1) &= \begin{pmatrix} -6 & 0 \\ 0 & -6 \end{pmatrix} && \text{negativ definit} \\ H_f(1,0) &= \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix} && \text{positiv definit} \end{aligned}$$

und

$$H_f(1,-1) = \begin{pmatrix} 6 & 0 \\ 0 & -6 \end{pmatrix} \qquad \text{indefinit}$$

Damit ist f an $(0,-1)$ ein lokales Maximum, an $(1,0)$ ein lokales Minimum und ansonsten keine Extremstellen.

21 Hauptsätze der mehrdimensionalen Analysis

Wir werden uns in diesem Abschnitt mit zentralen Ergebnissen der mehrdimensionalen Analysis beschäftigen. Dazu starten wir mit einem sehr allgemeinen Ergebnis, das auch später noch an verschiedenen Stellen von Bedeutung sein wird.

Satz 21.1 *(Banach'scher Fixpunktsatz)*
Es sei (X, d) ein vollständiger metrischer Raum und es sei $\varphi : X \to X$. Ferner existiere ein $\alpha < 1$ so, dass

$$d(\varphi(x), \varphi(y)) \leq \alpha d(x, y) \qquad (x, y \in X)$$

*(eine Abbildung mit dieser Eigenschaft heißt α-*Kontraktion*). Dann existiert genau ein $x^* \in X$ mit $x^* = \varphi(x^*)$, also genau ein Fixpunkt von φ. Außerdem konvergiert für alle $x_0 \in X$ die Folge (x_n) mit*

$$x_{n+1} := \varphi(x_n) \qquad (n \in \mathbb{N}_0)$$

gegen x^* und es gilt für alle $n \in \mathbb{N}_0$

$$d(x_n, x^*) \leq \alpha^n d(x_0, x^*) \left(\leq \frac{\alpha^n}{1-\alpha} d(x_1, x_0) \right) .$$

Beweis. Für alle $n \in \mathbb{N}$ gilt

$$d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1})) \leq \alpha \cdot d(x_n, x_{n-1}) \leq \ldots \leq \alpha^n \cdot d(x_1, x_0) .$$

Also ist für $m, n \in \mathbb{N}$ mit $m > n$

$$\begin{aligned} d(x_m, x_n) &\leq \sum_{k=n}^{m-1} d(x_{k+1}, x_k) \leq \sum_{k=n}^{m-1} \alpha^k d(x_1, x_0) \\ &\leq d(x_1, x_0) \sum_{k=n}^{\infty} \alpha^k = \frac{\alpha^n}{1-\alpha} d(x_1, x_0) . \end{aligned}$$

Ist $\varepsilon > 0$, so existiert ein $N_\varepsilon \in \mathbb{N}$ mit

$$\frac{\alpha^n}{1-\alpha} d(x_1, x_0) < \varepsilon \qquad (n \geq N_\varepsilon) ,$$

also auch

$$d(x_m, x_n) < \varepsilon \qquad (m > n \geq N_\varepsilon) .$$

Folglich ist (x_n) eine Cauchy-Folge in X. Da X vollständig ist, existiert ein $x^* \in X$ mit $x_n \to x^*$ $(n \to \infty)$. Da φ nach Voraussetzung insbesondere stetig auf X ist, gilt damit

$$x^* \leftarrow x_{n+1} = \varphi(x_n) \to \varphi(x^*) \qquad (n \to \infty)$$

d. h. $x^* = \varphi(x^*)$. Außerdem erhalten wir

$$d(x_n, x^*) = d(\varphi(x_{n-1}), \varphi(x^*)) \leq \alpha d(x_{n-1}, x^*) \leq \ldots \leq \alpha^n d(x_0, x^*)$$

und zudem

$$d(x_0, x^*) \leq d(x_0, x_1) + d(x_1, x^*) \leq d(x_0, x_1) + \alpha d(x_0, x^*) \ ,$$

also

$$(1 - \alpha) d(x_0, x^*) \leq d(x_0, x_1) \ .$$

Ist schließlich $\tilde{x}$ ein weiterer Fixpunkt von φ, so ist

$$d(\tilde{x}, x^*) = d(\varphi(\tilde{x}), \varphi(x^*)) \leq \alpha d(\tilde{x}, x^*) \ ,$$

also $d(\tilde{x}, x^*) = 0$. □

Bemerkung 21.2 Eine typische Situation, unter der die Voraussetzungen von S. 21.1 erfüllt sind, ist die folgende: Es sei $U \subset \mathbb{R}^d$ offen und $\varphi : U \to \mathbb{R}^d$ differenzierbar. Ferner sei $M \subset U$ abgeschlossen mit $\varphi(M) \subset M$ und so, dass für $\alpha < 1$

$$||J_\varphi(x)|| \leq \alpha \qquad (x \in M) \ .$$

(Denn dann ist $(M, d_{|\cdot|})$ vollständig als abgeschlossene Teilmenge des vollständigen metrischen Raumes $(\mathbb{R}^d, d_{|\cdot|})$ und φ ist eine α-Kontraktion nach S. 20.6.)

Wir beschäftigen uns nun mit der "lokalen Umkehrbarkeit" von Funktionen $f : M \to \mathbb{R}^d$, wobei $M \subset \mathbb{R}^d$. Ist $f : M \to N$ bijektiv, so bedeutet dies, dass für jedes $y = (y_1, \ldots, y_d) \in N$ das Gleichungssystem

$$\begin{aligned} f_1(x_1, \ldots, x_d) &= y_1 \\ &\vdots \\ f_d(x_1, \ldots, x_d) &= y_d \end{aligned}$$

genau eine Lösung $x = (x_1, \ldots, x_d) \in M$ besitzt. Betrachten wir zunächst zwei bekannte Spezialfälle:
1. Ist $f : \mathbb{R}^d \to \mathbb{R}^d$ definiert durch

$$f(x) := Ax \qquad (x \in \mathbb{R}^d)$$

wobei $A \in \mathbb{R}^{d \times d}$ ist, so ist das lineare Gleichungssystem

$$Ax = f(x) = y$$

genau dann für alle $y \in \mathbb{R}^d$ eindeutig lösbar, wenn $\det A \neq 0$ ist. In diesem Falle ist also $f : \mathbb{R}^d \to \mathbb{R}^d$ bijektiv, und es gilt bekanntlich

$$x = f^{-1}(y) = A^{-1}y \qquad (y \in \mathbb{R}^d)\ .$$

2. Ist $\Omega \subset \mathbb{R}$ offen und $f : \Omega \to \mathbb{R}$ stetig differenzierbar mit $f'(x_0) \neq 0$ für ein $x_0 \in \Omega$, so existiert eine offene Umgebung U von x_0 so, dass $f'(x)$ entweder durchgehend > 0 oder < 0 auf U und damit f streng monoton auf U ist. Also existiert $f^{-1} := (f_{|U})^{-1}$ auf $V = f(U)$ und es gilt für $y = f(x) \in V$

$$(f^{-1})'(y) = \frac{1}{f'(x)}\ .$$

Schließlich ist auch V offen (da $V = (f^{-1})^{-1}(U)$ nach S. 10.12 offen ist) und f^{-1} stetig differenzierbar auf V.

Wir setzen für eine offene Menge $\Omega \in \mathbb{R}^d$ und $k, m \in \mathbb{N}$

$$C^k(\Omega, \mathbb{R}^m) := \{f = (f_1, \ldots, f_m)^T : \Omega \to \mathbb{R}^m : f_1, \ldots, f_m \in C^k(\Omega)\}\ .$$

Dann gilt allgemein

Satz 21.3 *(Hauptsatz über Umkehrfunktionen)*
Es sei $\Omega \subset \mathbb{R}^d$ offen, und es sei $f \in C^1(\Omega, \mathbb{R}^d)$. Ferner sei $x^{(0)} \in \Omega$ mit

$$\det J_f(x^{(0)}) \neq 0\ .$$

Dann gilt:

1. *Es existieren offene Umgebungen U von $x^{(0)}$ und V von $y^{(0)} := f(x^{(0)})$ so, dass $f_{|U} : U \to V$ bijektiv ist.*
2. *Ist $f^{-1} = (f_{|U})^{-1} : V \to U$, so ist $f^{-1} \in C^1(V, \mathbb{R}^d)$ und für $y = f(x) \in V$ gilt*

$$J_{f^{-1}}(y) = [J_f(x)]^{-1}\ .$$

Beweis. Wir setzen $A := J_f(x^{(0)})$. Nach Voraussetzung existiert $A^{-1}(\neq 0)$. Da $x \mapsto J_f(x)$ stetig auf Ω ist (S. 19.13), existiert ein $\delta > 0$ so, dass für $\varepsilon := (2||A^{-1}||)^{-1}$ gilt

$$||J_f(x) - A|| < \varepsilon \qquad \Big(x \in U_\delta(x^{(0)})\Big)\ .$$

Wir setzen

$$U := U_\delta(x^{(0)})\ , \qquad V := f(U)$$

und zeigen:

(i) $f_{|U} : U \to V$ ist injektiv (also bijektiv), d. h. $f^{-1} = (f_{|U})^{-1}$ existiert

(ii) $V \subset \mathbb{R}^d$ ist offen

(iii) $\det J_f(x) \neq 0 \quad (x \in U)$

(iv) $f^{-1} \in C^1(V, \mathbb{R}^d)$ und $J_{f^{-1}}(y) = [J_f(x)]^{-1}$ für alle $y \in V$, $y = f(x)$.

1. Es sei $y \in V$ fest. Wir betrachten $\varphi = \varphi_y : U \to \mathbb{R}^d$ mit

$$\varphi(x) := x + A^{-1}(y - f(x)) \qquad (x \in U)\ .$$

Dann ist $\varphi(x) = x$ genau dann, wenn $y = f(x)$ gilt. Also ist für (i) zu zeigen: φ hat höchstens (und damit genau) einen Fixpunkt. Es gilt für $x \in U$

$$J_\varphi(x) = E - A^{-1}J_f(x) = A^{-1}(A - J_f(x))$$

und damit

$$||J_\varphi(x)|| \leq ||A^{-1}|| \cdot ||A - J_f(x)|| < ||A^{-1}||\varepsilon = \frac{1}{2}\ .$$

Sind $x, \tilde{x}$ Fixpunkte von φ, so gilt nach S. 20.6

$$|x - \tilde{x}| = |\varphi(x) - \varphi(\tilde{x})| \leq \frac{1}{2}|x - \tilde{x}|$$

also $x = \tilde{x}$.

2. Es sei $y^{(1)} = f(x^{(1)}) \in V$. Da $U = U_\delta(x^{(0)})$ offen ist, existiert ein $\rho > 0$ mit $\overline{U_\rho(x^{(1)})} =: M \subset U$. Wir setzen $\eta := \varepsilon\rho$ und zeigen: $U_\eta(y^{(1)}) \subset V$.
Dazu sei $y \in U_\eta(y^{(1)})$ gegeben und $\varphi = \varphi_y$ wie in (i). Dann gilt $\varphi(M) \subset M$, denn für $x \in M$ ist

$$\begin{aligned}
|\varphi(x) - x^{(1)}| &\leq |\varphi(x) - \varphi(x^{(1)})| + |\varphi(x^{(1)}) - x^{(1)}| \\
&\leq \frac{1}{2}|x - x^{(1)}| + |A^{-1}(y - f(x^{(1)}))| \\
&\leq \frac{1}{2}|x - x^{(1)}| + ||A^{-1}|| \cdot |y - y^{(1)}| \leq \frac{\rho}{2} + \frac{\rho}{2} = \rho
\end{aligned}$$

also $\varphi(x) \in M$.
Außerdem ist $||J_\varphi(x)|| \leq 1/2$ für alle $x \in M$ nach 1. Nach S. 21.1 und B. 21.2 existiert (genau) ein $x \in M$ mit $\varphi(x) = x$ also $f(x) = y$.

3. Es gilt nach 1. (wobei $\varphi = \varphi_y$ mit einem beliebigen $y \in V$)

$$J_f(x) = A \cdot (E - J_\varphi(x)) \qquad (x \in U)$$

also

$$\det J_f(x) = \det(A) \cdot \det(E - J_\varphi(x)) \qquad (x \in U)\ .$$

Wieder nach 1. ist $||J_\varphi(x)|| \le \frac{1}{2}$ $(x \in U)$ und damit ist $E - J_\varphi(x)$ invertierbar. (Es gilt $(E-B)^{-1} = \sum\limits_{\nu=0}^{\infty} B^\nu$, falls $||B|| < 1$ ist; [Ü]). Also ist auch $\det J_f(x) \ne 0$.

4. Wir setzen $g := (f_{|U})^{-1}$ und betrachten ein beliebiges $y^{(1)} \in V$. Ist $\varphi = \varphi_{y^{(1)}}$ wie in 1. und $y^{(1)} = f(x^{(1)})$, so gilt

$$\varphi(x) = x + A^{-1}\left(f(x^{(1)}) - f(x)\right)$$

also für $x \in U$

$$\begin{aligned} \frac{1}{2}|x - x^{(1)}| &\ge |\varphi(x) - \varphi(x^{(1)})| = \\ &= |x - x^{(1)} + A^{-1}(f(x^{(1)}) - f(x))| \\ &\ge |x - x^{(1)}| - ||A^{-1}|| \cdot |f(x^{(1)}) - f(x)| \end{aligned}$$

und damit

$$|f(x) - f(x^{(1)})| \ge \frac{1}{2||A^{-1}||}|x - x^{(1)}| = \varepsilon |x - x^{(1)}| \,. \qquad (*)$$

Für $y \in V$, $y \ne y^{(1)}$, folgt (mit $y = f(x)$)

$$\begin{aligned} &\frac{|g(y) - g(y^{(1)}) - [J_f(x^{(1)})]^{-1}(y - y^{(1)})|}{|y - y^{(1)}|} \le \\ \le\ & ||[J_f(x^{(1)})]^{-1}|| \frac{|J_f(x^{(1)})(x - x^{(1)}) - (f(x) - f(x^{(1)}))|}{|f(x) - f(x^{(1)})|} \\ \le\ & \frac{||[J_f(x^{(1)})]^{-1}||}{\varepsilon} \frac{|f(x) - f(x^{(1)}) - J_f(x^{(1)})(x - x^{(1)})|}{|x - x^{(1)}|} \to 0 \quad (x \to x^{(1)}) \,. \end{aligned}$$

Für $y \to y^{(1)}$ folgt $x \to x^{(1)}$ aus $(*)$. Also ist g differenzierbar an $y^{(1)}$ und es gilt (Eindeutigkeit von C in D. 19.6)

$$J_g(y^{(1)}) = [J_f(x^{(1)})]^{-1} = [J_f(f^{-1}(y^{(1)}))]^{-1} \,.$$

Da $A \mapsto A^{-1}$ eine stetige Abbildung von $\mathbb{R}^{d\times d} \setminus \{B : \det(B) = 0\}$ in sich selbst ist (bzgl. der Operatornorm)([Ü]), ist $y \mapsto J_g(y)$ als Verknüpfung der stetigen Abbildungen f^{-1} sowie $x \mapsto J_f(x)$ und $A \mapsto A^{-1}$ stetig auf V. Dies ist äquivalent dazu, dass $g = f^{-1}$ in $C^1(V, \mathbb{R}^d)$ liegt. □

Beispiel 21.4 Wir betrachten $f : (0,\infty) \times \mathbb{R} \to \mathbb{R}^2$ mit

$$f(r,\varphi) := \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix} \qquad (r > 0, \varphi \in \mathbb{R}) \,.$$

Dann ist $f \in C^1(\Omega, \mathbb{R}^2)$, wobei $\Omega = (0, \infty) \times \mathbb{R}$, und es gilt

$$\det J_f(r, \varphi) = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r > 0 \,.$$

Also existiert nach S. 21.3 zu jedem $(r_0, \varphi_0) \in \Omega$ eine Umgebung U von (r_0, φ_0) so, dass $f : U \to f(U) = V$ bijektiv ist. Außerdem ist V offen und $f^{-1} = (f_{|U})^{-1} \in C^1(V, \mathbb{R}^2)$. Jedoch ist f nicht umkehrbar auf ganz Ω!
Wie sehen "maximale" offene Mengen aus, auf denen f umkehrbar ist?
Ist $f(r, \varphi) = f(\tilde{r}, \tilde{\varphi})$, so gilt $r\cos\varphi = \tilde{r}\cos\tilde{\varphi}$ und $r\sin\varphi = \tilde{r}\sin\tilde{\varphi}$, also

$$r^2(\cos^2\varphi + \sin^2\varphi) = \tilde{r}^2(\cos^2\tilde{\varphi} + \sin^2\tilde{\varphi})$$

und damit $r = \tilde{r}$. Folglich gilt $\cos\varphi = \cos\tilde{\varphi}$ und $\sin\varphi = \sin\tilde{\varphi}$, woraus wiederum $\tilde{\varphi} = \varphi + 2k\pi$ für ein $k \in \mathbb{Z}$ folgt. Also ist f etwa umkehrbar auf

$$U = U_\alpha = \{(r, \varphi) \in \Omega : \alpha < \varphi < \alpha + 2\pi\}$$

für beliebiges $\alpha \in \mathbb{R}$. Auf jeder echten offenen Obermenge von U_α ist f nicht mehr umkehrbar. Außerdem gilt $f(U_\alpha) = \mathbb{R}^2 \setminus \{(r\cos\alpha, r\sin\alpha) : r > 0\}$ für alle $\alpha \in \mathbb{R}$. Wie sieht eine (nicht **die**) Umkehrfunktion aus? Betrachten wir

$$U = U_{-\pi/2} = \left\{(r, \varphi) : r > 0 \,,\ \varphi \in \left(-\frac{\pi}{2}\,, \frac{3}{2}\pi\right)\right\} \,.$$

Ist $(x, y) \in f(U_{-\pi/2})$ mit

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}$$

so gilt $r = \sqrt{x^2 + y^2}$. Für $x = 0$ ist $\varphi = \pi/2$. Ist $x \neq 0$, so folgt

$$\frac{y}{x} = \tan\varphi$$

also

$$\varphi = \begin{cases} \arctan(y/x) & , \text{ falls } x > 0 \\ \arctan(y/x) + \pi & , \text{ falls } x < 0 \end{cases}$$

(beachte: $\cos\varphi > 0$, falls $\varphi \in (-\pi/2, \pi/2)$). Insgesamt erhalten wir

$$f^{-1}(x, y) = (f_{|U_{-\pi/2}})^{-1}(x, y) = \begin{cases} \left(\sqrt{x^2 + y^2}, \arctan(y/x)\right) & , \text{ falls } x > 0 \\ (y, \pi/2) & , \text{ falls } x = 0 \\ \left(\sqrt{x^2 + y^2}, \arctan(y/x) + \pi\right) & , \text{ falls } x < 0 \end{cases} \,.$$

In enger Beziehung zum Hauptsatz über Umkehrfunktionen steht ein weiterer Hauptsatz: der über implizite Funktionen. Worum geht es dabei?

Gegeben ist eine Funktion $F : M \to \mathbb{R}^m$ mit $M \subset \mathbb{R}^d \times \mathbb{R}^m$ und damit das Gleichungssystem

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

wobei $\mathbb{R}^d \times \mathbb{R}^m = \{(x, y) : x = (x_1, \ldots, x_d) \in \mathbb{R}^d\ ,\ y = (y_1, \ldots, y_m) \in \mathbb{R}^m\}$, d. h.

$$\begin{aligned} F_1(x_1, \ldots, x_d,\, y_1, \ldots, y_m) &= 0 \\ &\vdots \\ F_m(x_1, \ldots, x_d,\, y_1, \ldots, y_m) &= 0 \end{aligned}$$

(also $d+m$ "Unbekannte" und m Gleichungen). Ferner sei $(x^{(0)}, y^{(0)}) \in M$ eine Lösung, also $F(x^{(0)}, y^{(0)}) = 0$. Ziel ist, das Gleichungssystem "lokal nach y aufzulösen", d. h. wir suchen Umgebungen U von $x^{(0)}$ und V von $y^{(0)}$ sowie eine Funktion $f : U \to \mathbb{R}^m$ so, dass für alle $x \in U, y \in V$ gilt

$$F(x, y) = 0 \Leftrightarrow y = f(x)\ .$$

Wir orientieren uns wieder an zwei einfachen Beispielen

Beispiel 21.5 1. Wir betrachten die Gleichung

$$F(x, y) := x^2 - y^2 - 1 = 0 \qquad (x, y \in \mathbb{R})\ .$$

Dann ist offenbar $(x_0, y_0) = (\sqrt{2}, 1)$ eine Lösung der Gleichung. Hier gilt etwa für $U := (1, \infty)$ und $W := (1, \infty) \times (0, \infty)$

$$F(x, y) = 0 \Leftrightarrow y = \sqrt{x^2 - 1} =: f(x)$$

(d. h. $\{(x, y) \in W : F(x, y) = 0\} = \{(x, f(x)) : x \in U\}$; die Lösungsmenge ist lokal der Graph der Funktion f). Dies ist etwa falsch für $W = (1, \infty) \times \mathbb{R}$.
2. Es seien $A \in \mathbb{R}^{m \times d}$, $B \in \mathbb{R}^{m \times m}$ sowie $F : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^m$ mit

$$F(x, y) = Ax + By = (A\ \ B)\begin{pmatrix} x \\ y \end{pmatrix}$$

(also lineares GLS $Ax + By = 0$). Dann gilt im Falle $\det B \neq 0$

$$F(x, y) = 0 \Leftrightarrow By = -Ax \Leftrightarrow y = -B^{-1}Ax =: f(x)$$

also $L := \{(x, y) : F(x, y) = 0\} = \{(x, f(x)) : x \in \mathbb{R}^d\}$.

Satz 21.6 *(Hauptsatz über implizite Funktionen)*
Es sei $\Omega \subset \mathbb{R}^d \times \mathbb{R}^m$ offen und es sei $F \in C^1(\Omega, \mathbb{R}^m)$. Wir setzen zur Abkürzung

$$\frac{\partial F}{\partial y}(x, y) := \left(\frac{\partial F_\mu}{\partial y_\nu}(x, y)\right)_{\mu,\nu=1,\ldots,m} \qquad \textit{und} \quad \frac{\partial F}{\partial x}(x, y) := \left(\frac{\partial F_\mu}{\partial x_\nu}(x, y)\right)_{\substack{\mu=1,\ldots,m \\ \nu=1,\ldots,d}}\ .$$

Ferner sei $(x^{(0)}, y^{(0)}) \in \Omega$ *so, dass* $F(x^{(0)}, y^{(0)}) = 0$ *und*

$$\det \frac{\partial F}{\partial y}(x^{(0)}, y^{(0)}) \neq 0 \, .$$

Dann gilt

1. *Es existieren offene Umgebungen* U *von* $x^{(0)}$ *und* W *von* $(x^{(0)}, y^{(0)})$ *sowie eine Funktion* $f : U \to \mathbb{R}^m$ *so, dass für* $(x, y) \in W$ *gilt*

$$F(x, y) = 0 \Leftrightarrow y = f(x) \, .$$

2. *Es ist* $f \in C^1(U, \mathbb{R}^m)$ *und*

$$J_f(x) = -\left[\frac{\partial F}{\partial y}(x, f(x))\right]^{-1} \frac{\partial F}{\partial x}(x, f(x)) \qquad (x \in U) \, .$$

Beweis. 1. Wir betrachten $G : \Omega \to \mathbb{R}^d \times \mathbb{R}^m$ mit

$$G(x, y) = \Big(x, F(x, y)\Big) \qquad \Big((x, y) \in \Omega\Big) \, .$$

Dann gilt $G(x^{(0)}, y^{(0)}) = (x^{(0)}, 0)$ und $G \in C^1(\Omega, \mathbb{R}^d \times \mathbb{R}^m)$. Außerdem ist

$$J_G(x, y) = \left[\begin{array}{ccc} E_d & \vdots & 0 \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots\dots} \\ \frac{\partial F}{\partial x}(x, y) & \vdots & \frac{\partial F}{\partial y}(x, y) \end{array}\right]$$

und damit $\det J_G(x^{(0)}, y^{(0)}) = \det \frac{\partial F}{\partial y}(x^{(0)}, y^{(0)}) \neq 0$. Nach S. 21.3 existieren offene Umgebungen W von $(x^{(0)}, y^{(0)})$ und S von $(x^{(0)}, 0) = G(x^{(0)}, y^{(0)})$ so, dass $G_{|W} : W \to S$ bijektiv ist mit

$$G^{-1} = (G_{|W})^{-1} \in C^1(S, \mathbb{R}^d \times \mathbb{R}^m) \, .$$

Da S offen ist, existieren offene Umgebungen U von $x^{(0)}$ und V von 0 mit $U \times V \subset S$. Wir können also o. E. annehmen, dass $S = U \times V$ gilt. Aus der Definition von G ergibt sich, dass G^{-1} von der Form

$$G^{-1}(x, z) = (x, H(x, z)) \qquad ((x, z) \in S)$$

mit $H \in C^1(S, \mathbb{R}^m)$ ist.
Ist $\pi_2 : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^m$ definiert durch $\pi_2(x, y) := y$, so gilt

$$\pi_2(G(x, y)) = \pi_2(x, F(x, y)) = F(x, y) \qquad ((x, y) \in \Omega)$$

also für $(x, z) \in S$:

$$F(x, H(x, z)) = \pi_2(G(x, H(x, z))) = \pi_2(x, z) = z \, .$$

Definiert man $f : U \to \mathbb{R}^m$ durch

$$f(x) := H(x,0) \qquad (x \in U)$$

so gilt für $x \in U$

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

und aus $F(x,y) = 0$ für ein $(x,y) \in W$ folgt

$$G(x,y) = (x,0) = G(G^{-1}(x,0)) = G(x, f(x))$$

und damit, da $G_{|W}$ bijektiv ist, $y = f(x)$. Also ergibt sich 1.
2. Weiter folgt aus $H \in C^1(S, \mathbb{R}^m)$ auch $f \in C^1(U, \mathbb{R}^m)$ und mit der Kettenregel ergibt sich aus

$$\varphi(x) := F(x, f(x)) \equiv 0 \qquad (x \in U)$$

schließlich

$$0 = J_\varphi(x) = J_F(x, f(x)) \cdot \begin{pmatrix} E_d \\ J_f(x) \end{pmatrix} = \frac{\partial F}{\partial x}(x, f(x)) + \frac{\partial F}{\partial y}(x, f(x)) J_f(x)$$

und damit

$$J_f(x) = -\left[\frac{\partial F}{\partial y}(x, f(x))\right]^{-1} \frac{\partial F}{\partial x}(x, f(x)) \qquad (x \in U)\ .$$

(Man beachte dabei: Es gilt nach S. 21.3

$$\det \frac{\partial F}{\partial y}(x,y) = \det J_G(x,y) \neq 0$$

für alle $(x,y) \in W$.) □

Beispiel 21.7 (Lemniskate)
Es sei $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ definiert durch

$$F(x,y) = (x^2 + y^2)^2 - 2(x^2 - y^2) \qquad (x, y \in \mathbb{R})\ .$$

Dann gilt

$$0 = \frac{\partial F}{\partial y}(x,y) = 2(x^2 + y^2)2y + 4y = 4y(x^2 + y^2 + 1)$$

genau dann, wenn $y = 0$ ist. Ist $y = 0$ und $F(x,y) = 0$, so gilt $x^4 - 2x^2 = 0$, also $x = 0$ oder $x = \pm\sqrt{2}$.
Nach S. 21.6 ist für alle (x_0, y_0) mit $F(x_0, y_0) = 0$ und $(x_0, y_0) \notin \{(0,0), (\pm\sqrt{2}, 0)\}$ die Gleichung $F(x,y) = 0$ auf einer Umgebung W von (x_0, y_0) “auflösbar nach y”. “Implizites Differenzieren” ergibt für die Funktion $f = f_{(x_0,y_0)}$ aus S. 21.6

$$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))} = \frac{4x(x^2 + y^2 - 1)}{4y(x^2 + y^2 + 1)}_{|y = f(x)}$$

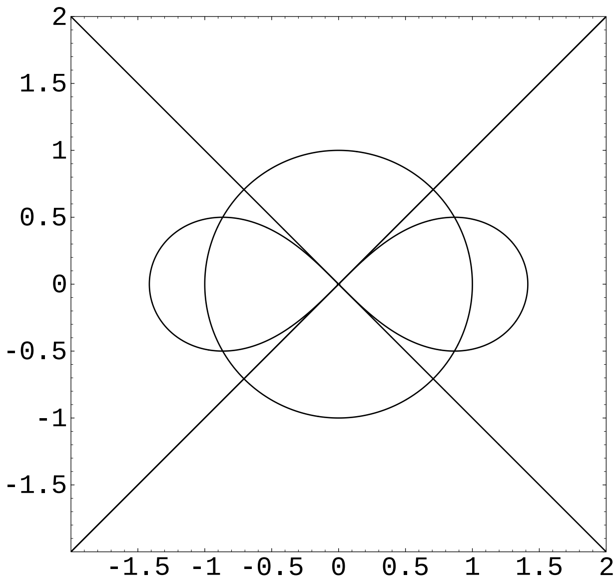
auf einer Umgebung U von x_0. Also hat f Extremstellen höchstens in Punkten x mit

$$x^2 + f^2(x) = x^2 + y^2 = 1 .$$

Um eine Vorstellung von der Lösungsmenge zu bekommen, betrachten wir Polarkoordinaten: Es gilt für $(x, y) \neq 0$

$$0 = F(x, y) = F(r\cos\varphi, r\sin\varphi) = r^4 - 2r^2(\cos^2\varphi - \sin^2\varphi) = r^2(r^2 - 2\cos 2\varphi)$$

genau dann, wenn $r = \sqrt{2\cos 2\varphi}$, wobei φ so, dass $\cos(2\varphi) > 0$ ist (beachte: $r > 0$).



Eine wichtige Anwendung des Hauptsatzes über implizite Funktionen ergibt sich im Bereich der Optimierung unter Nebenbedingungen. Wir untersuchen folgendes Problem:
Gegeben sind Funktionen $f : M \to \mathbb{R}$, $g : M \to \mathbb{R}^m$, wobei $M \subset \mathbb{R}^d$. Gesucht sind die "Extremstellen von f unter der Nebenbedingung $g(x) = 0$".

Wir wollen dies in einer Definition etwas präzisieren

Definition 21.8 Es sei $M \subset \mathbb{R}^d$, und es seien $f : M \to \mathbb{R}$ sowie $g : M \to \mathbb{R}^m$. Ist

$$L := \{x \in M : g(x) = 0\}$$

und ist $x^{(0)} \in L$, so heißt $x^{(0)}$ ein *lokales Maximum* (bzw. *Minimum*) *von f unter der Nebenbedingung* $g(x) = 0$, falls $x^{(0)}$ ein lokales Maximum (bzw. Minimum) von $f_{|L}$ ist.

Der folgende Satz liefert eine notwendige Bedingung für das Vorliegen einer solchen Extremstelle

Satz 21.9 *(Lagrange)*
Es sei $\Omega \subset \mathbb{R}^d$ offen und es seien $f \in C^1(\Omega)$ und $g \in C^1(\Omega, \mathbb{R}^m)$ mit $m < d$. Ferner sei $x^{(0)} \in \Omega$ so, dass

$$\operatorname{Rang}\big(J_g(x^{(0)})\big) = m$$

ist. Dann gilt: Hat f an $x^{(0)}$ ein lokales Extremum unter der Nebenbedingung $g(x) = 0$, so existiert ein $\lambda \in \mathbb{R}^m$ mit

$$\operatorname{grad} f(x^{(0)}) = \sum_{j=1}^{m} \lambda_j \operatorname{grad} g_j(x^{(0)}) = \lambda^T J_g(x^{(0)})$$

(die Komponenten $\lambda_1, \ldots, \lambda_m$ von λ heißen "Lagrange-Multiplikatoren").

Beweis. Wir schreiben zur Abkürzung $x = (y, z)$ mit

$$y = (x_1, \ldots, x_m)\,, \; z = (x_{m+1}, \ldots, x_d)$$

für $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ und insbesondere

$$(y^{(0)}, z^{(0)}) = x^{(0)}\ .$$

O. E. gelte

$$\det\left(\frac{\partial g}{\partial y}(x^{(0)})\right) = \det\left(\frac{\partial g_\mu}{\partial x_\nu}(x^{(0)})\right)_{\mu,\nu=1,\ldots,m} \neq 0\,,$$

d. h. $\frac{\partial g}{\partial y}(x^{(0)})$ habe vollen Rang m.

Nach dem Hauptsatz über implizite Funktionen ist $g(x) = g(y, z) = 0$ bei $x^{(0)}$ "lokal nach y auflösbar". Insbesondere existieren eine Umgebung U von $z^{(0)} = (x^{(0)}_{m+1}, \ldots, x^{(0)}_d)$ sowie eine Funktion $\varphi \in C^1(U, \mathbb{R}^m)$ so, dass $\varphi(z^{(0)}) = y^{(0)}$ und

$$g(\varphi(z), z) = 0 \qquad (z \in U) \tag{$*$}$$

gilt. Wir definieren $F \in C^1(U)$ durch

$$F(z) := f(\varphi(z), z) \quad (z \in U)\ .$$

Nach Voraussetzung hat F an $z^{(0)}$ ein lokales Extremum (ohne Nebenbedingung). Also gilt nach S. 20.17

$$\operatorname{grad} F(z^{(0)}) = 0$$

und mit der Kettenregel weiter

$$0 = \operatorname{grad} f(x^{(0)}) \cdot \begin{pmatrix} J_\varphi(z^{(0)}) \\ E_{d-m} \end{pmatrix} = \frac{\partial f}{\partial y}(x^{(0)}) J_\varphi(z^{(0)}) + \frac{\partial f}{\partial z}(x^{(0)}) \tag{$**$}$$

Wegen $\det\left(\frac{\partial g}{\partial y}(x^{(0)})\right) \neq 0$ hat das lineare Gleichungssystem

$$\lambda^T \cdot \frac{\partial g}{\partial y}(x^{(0)}) = \frac{\partial f}{\partial y}(x^{(0)}) \qquad (***)$$

genau eine Lösung $\lambda^T = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$. Aus $(*)$ ergibt sich durch "implizites Differenzieren"

$$0 = J_g(x^{(0)}) \begin{pmatrix} J_\varphi(z^{(0)}) \\ E_{d-m} \end{pmatrix} = \frac{\partial g}{\partial y}(x^{(0)}) J_\varphi(z^{(0)}) + \frac{\partial g}{\partial z}(x^{(0)})$$

und mit $(**)$ folgt

$$\frac{\partial f}{\partial z}(x^{(0)}) = -\frac{\partial f}{\partial y}(x^{(0)}) J_\varphi(z^{(0)}) = -\lambda^T \frac{\partial g}{\partial y}(x^{(0)}) J_\varphi(z^{(0)}) = \lambda^T \frac{\partial g}{\partial z}(x^{(0)}) \, .$$

Zusammen mit $(***)$ ergibt sich die Behauptung. □

Bemerkung 21.10 Ähnlich wie S. 20.17 liefert S. 21.9 lediglich eine notwendige Bedingung für das Vorliegen eines Extremums unter der Nebenbedingung $g(x) = 0$. Um die entsprechenden Punkte $x^{(0)}$ zu bestimmen, hat man die Gleichungen

$$g_j(x_1, \dots, x_d) = 0 \qquad (j = 1, \dots, m)$$

und

$$\frac{\partial f}{\partial x_k}(x_1, \dots, x_d) = \sum_{j=1}^m \lambda_j \frac{\partial g_j}{\partial x_k}(x_1, \dots, x_d) \qquad (k = 1, \dots, d)$$

zu lösen (also $(d+m)$ Gleichungen für die $(d+m)$ Unbekannten $x_1, \dots, x_d$ und $\lambda_1, \dots, \lambda_m$).

Beispiel 21.11 Die Produktion eines Unternehmens sei in Abhängigkeit der Produktionsfaktoren x, y beschrieben durch die (Cobb-Douglas-) Funktion

$$P(x, y) = x^\alpha y^{1-\alpha} \qquad (x, y > 0)$$

wobei $\alpha \in (0, 1)$ fest ist. Die Produktionskosten seien gegeben durch eine lineare Kostenfunktion K der Form

$$K(x, y) = px + qy \qquad (x, y > 0)$$

mit Konstanten $p, q > 0$. Gesucht ist eine kostenminimale Faktorkombination (x, y) zu einem vorgegebenen Produktionsniveau $c > 0$, d. h. wir wollen das Optimierungsproblem

$$K(x, y) \stackrel{!}{\to} \min$$

unter der Nebenbedingung

$$P(x, y) - c = 0$$

lösen. Nach S. 21.9 ist eine notwendige Bedingung gegeben durch

$$\operatorname{grad} K(x, y) = \lambda \operatorname{grad} P(x, y) ,$$

d. h. wir haben die 3 Gleichungen

$$\begin{aligned} x^\alpha y^{1-\alpha} &= c \\ p &= \lambda \alpha x^{\alpha-1} y^{1-\alpha} \\ q &= \lambda(1-\alpha) x^\alpha y^{-\alpha} \end{aligned}$$

für die Unbekannten x, y, λ. Division der 2. und 3. Gleichung ergibt

$$x = \frac{\alpha}{1-\alpha} \cdot \frac{q}{p} \cdot y$$

und mit der 1. Gleichung erhalten wir

$$y = c \left(\frac{p(1-\alpha)}{q\alpha} \right)^\alpha$$

und damit

$$x = c \left(\frac{q\alpha}{p(1-\alpha)} \right)^{1-\alpha} .$$

Also kann nur an dieser Stelle (x, y) ein Minimum unter der Nebenbedingung vorliegen. Man kann sich überlegen, dass dies tatsächlich der Fall ist.

22 Gewöhnliche Differenzialgleichungen: Beispiele und elementare Lösungemethoden

Differenzialgleichungen sind Gleichungen, in denen eine unabhängige Variable, Funktionen und Ableitungen der Funktionen auftauchen. Dabei bezeichnen wir die unabhängige Variable meist mit t (Zeit). Bevor wir uns mit der allgemeinen Theorie beschäftigen, betrachten wir einige einfache Spezialfälle mit Anwendungsbeispielen.

Beispiel 22.1 Wir betrachten ein sehr vereinfachtes Keynesianisches Modell des Wachstums einer Volkswirtschaft. Ist Y das (Volks)einkommen (etwa gemessen als Bruttosozialprodukt), so besagt der Keynesianische Ansatz, dass das die Veränderung Y' proportional zur Differenz von Nachfrage D und Einkommen ist, d.h.

$$Y'(t) = k(D(t) - Y(t)) \qquad \text{oder kurz } Y' = k(D - Y) \ ,$$

wobei k eine positive Konstante ist. Weiter nimmt man an, dass die Nachfrage D sich als Summe aus privatem Konsum C, Investitionen I und Staatsausgaben G ergibt (geschlossene Volkswirtschaft; ohne Ausland). Ein weiter vereinfachtes Modell geht davon aus, dass I und G konstant sind (man schreibt $I = \bar{I}, G = \bar{G}$), und dass $C(t)$ von der Form $C(t) = c_0 + cY(t)$ mit Konstanten $c_0 > 0$ und $c \in (0,1)$ ist. Damit erhalten wir

$$\begin{aligned} Y'(t) &= k(c_0 + cY(t) + \bar{I} + \bar{G} - Y(t)) = \\ &= k(c-1)Y(t) + k(c_0 + \bar{I} + \bar{G}) =: \alpha Y(t) + \beta \end{aligned}$$

mit Konstanten $\alpha < 0$ und $\beta > 0$. Gesucht ist nun eine Funktion Y, die die Gleichung unter der „Anfangsbedingung“ $Y(0) = Y_0$ löst.

Definition 22.2 Es sei $D \subset \mathbb{R} \times \mathbb{K}^d$ offen, und es sei $f : D \to \mathbb{K}^d$ stetig. Wir schreiben ein Element aus $\mathbb{R} \times \mathbb{K}^d$ meist in der Form (t, y), wobei $t \in \mathbb{R}$ und $y \in \mathbb{K}^d$ ist.

1. Eine *(gewöhnliche) Differenzialgleichung (1. Ordnung)* (bzw. *System gewöhnlicher Differenzialgleichungen (1. Ordnung)*) ist eine Gleichung der Form

$$y'(t) = f(t, y(t)) \tag{22.1}$$

oder kurz

$$y' = f(t, y) \ ,$$

wobei $y'(t) = (y_1'(t), ..., y_d'(t))$.

2. Eine Funktion $\varphi : I \to \mathbb{K}^d$ bzw. das Paar (φ, I) heißt *Lösung* der Differenzialgleichung (22.1) falls φ differenzierbar auf (dem Intervall) I ist, falls

$$\operatorname{graph}(\varphi) = \{(t, \varphi(t)) : t \in I\} \subset D$$

gilt, und falls (22.1) für alle $t \in I$ erfüllt ist, d.h.

$$\varphi'(t) = f(t, \varphi(t)) \qquad \text{für alle} \quad t \in I\,.$$

3. Ist $(t_0, y^{(0)}) \in D$, so heißt eine Beziehung der Form

$$y' = f(t, y) \qquad\qquad y(t_0) = y^{(0)} \tag{22.2}$$

ein *Anfangswertproblem* (für die Differenzialgleichung (22.1)) (kurz: AWP). Ist $t_0 \in I$ und ist (φ, I) eine Lösung von (22.1) mit $\varphi(t_0) = y^{(0)}$, so heißt φ bzw. (φ, I) *Lösung des AWP* (22.2).

Ist $d = 1$, so nennt man (22.1) auch *skalare* Differenzialgleichung. In diesem Fall haben wir eine Lösung $\varphi : I \to \mathbb{K}$ (reell- oder komplexwertig).
Ist weiterhin speziell $D = I \times J$ mit Intervallen $I, J \subset \mathbb{R}$ und f von der Form

$$f(t, y) = h(t)g(y)$$

mit $h : I \to \mathbb{R}$ und $g : J \to \mathbb{R}$, so spricht man von einer *Differenzialgleichung mit getrennten Veränderlichen* (oder von einer *separierbaren Differenzialgleichung*). In B.22.1 haben wir eine solche Gleichung mit $h(t) \equiv 1$ und $g(y) = \alpha y + \beta$ vorliegen. Wir beschäftigen uns zunächst mit Lösungen solcher Gleichungen.
Ist $g(y_0) = 0$ für ein $y_0 \in J$, so ist offenbar $\varphi(t) \equiv y_0$ eine Lösung von $y' = h(t)g(y)$ auf $\mathbb{R}$. Das folgende Beispiel zeigt, dass unter Umständen mehrere Lösungen existieren können.

Beispiel 22.3 Es sei $D = \mathbb{R} \times \mathbb{R}$ und

$$f(t, y) = g(y) = \begin{cases} \sqrt{y}\,, & \text{falls } y \geq 0 \\ 0\,, & \text{falls } y < 0 \end{cases}\,.$$

Dann ist offenbar $\varphi(t) \equiv 0$ eine Lösung des AWP

$$y' = g(y)\,, \qquad y(0) = 0\,.$$

Man rechnet nach, dass auch die Funktionen

$$\varphi_c(t) := \begin{cases} (t - c)^2/4\,, & \text{falls } t \geq c \\ 0\,, & \text{falls } t < c \end{cases}$$

für alle $c \geq 0$ Lösungen des AWPs darstellen. Also haben wir unendlich viele Lösungen.

Im Falle $g(y_0) \neq 0$ treten entsprechende Probleme nicht auf:

Satz 22.4 *Es seien $I, J \subset \mathbb{R}$ offene Intervalle, und es seien $h : I \to \mathbb{R}$ und $g : J \to \mathbb{R}$ stetig. Ferner gelte*

$$g(y) \neq 0 \qquad (y \in J)\,.$$

Ist $t_0 \in I$, $y_0 \in J$ und

$$H(t) := \int\limits_{t_0}^{t} h(s)ds \quad (t \in I), \qquad G(y) := \int\limits_{y_0}^{y} \frac{ds}{g(s)} \quad (y \in J),$$

so hat das AWP

$$y' = h(t)g(y)\,, \qquad y(t_0) = y_0$$

auf jedem Intervall $I_0 \subset I$ mit $t_0 \in I_0$ und $H(I_0) \subset G(J)$ genau eine Lösung (φ, I_0), und diese ergibt sich durch Auflösen der Gleichung

$$G(y) = H(t)$$

nach y (d.h. $\varphi(t) = G^{-1}(H(t))$ $(t \in I_0)$).

Beweis. 1. Da $G'(y) = 1/g(y) \neq 0$ für alle $y \in J$ gilt, ist G streng monoton (wachsend oder fallend) auf J und besitzt eine (ebenfalls stetig differenzierbare) Umkehrfunktion $G^{-1} : G(J) \to J$ (S. 12.4 und S. 13.9). Ist $I_0 \subset I$ mit $t_0 \in I_0$ und $H(I_0) \subset G(J)$, so betrachten wir $\varphi : I_0 \to \mathbb{R}$ mit

$$\varphi(t) = G^{-1}(H(t)) \qquad (t \in I_0)\,.$$

Dann gilt

$$\varphi'(t) \overset{S.13.9}{=} \frac{1}{G'(G^{-1}(H(t)))} \cdot H'(t) = g(\varphi(t))h(t) \qquad (t \in I)$$

und $\varphi(t_0) = G^{-1}(H(t_0)) = G^{-1}(0) = y_0$, d.h. (φ, I_0) löst das AWP.
2. Ist (ψ, I_0) eine weitere Lösung des AWP, so gilt

$$\frac{\psi'(t)}{g(\psi(t))} = h(t) \qquad (t \in I_0)$$

und damit für alle $t \in I_0$ (Substitution $u = \psi(s)$)

$$G(\psi(t)) = \int\limits_{\psi(t_0)}^{\psi(t)} \frac{du}{g(u)} = \int\limits_{t_0}^{t} \frac{\psi'(s)}{g(\psi(s))} ds = \int\limits_{t_0}^{t} h(s)ds = H(t)$$

d.h.

$$\psi(t) = G^{-1}(H(t)) = \varphi(t)\,.$$

□

Satz 22.4 erweist sich als äußerst nützlich, da die Aussage „konstruktiv“ ist, d.h. es wird ein Verfahren zur Berechnung der Lösung geliefert. Im Wesentlichen hat man zwei Stammfunktionen (F und G) zu berechnen und die Umkehrfunktion von G anzuwenden.

Beispiel 22.5 1. Wir betrachten das AWP $Y' = \alpha Y + \beta$; $Y(0) = Y_0$ aus B. 22.1, wobei wir zunächst $Y_0 < -\beta/\alpha$ annehmen.
Dann ist das AWP nach S. 22.4 (jedenfalls für $|t|$ genügend klein) eindeutig lösbar, und die Lösung ergibt sich aus

$$t = \int_0^t ds = \int_{Y_0}^{Y} \frac{ds}{\alpha s + \beta} = \frac{1}{\alpha} \ln(\alpha s + \beta)|_{Y_0}^{Y} = \frac{1}{\alpha}\Big(\ln(\alpha Y + \beta) - \ln(\alpha Y_0 + \beta)\Big)\,,$$

also

$$\ln(\alpha Y + \beta) = \alpha t + \ln(\alpha Y_0 + \beta)$$

d.h.

$$Y = Y(t) = e^{\alpha t}(Y_0 + \beta/\alpha) - \beta/\alpha = e^{k(c-1)t}\big(Y_0 - \frac{c_0 + \bar{I} + \bar{G}}{1-c}\big) + \frac{c_0 + \bar{I} + \bar{G}}{1-c}$$

Dies ist eine Lösung auf ganz $\mathbb{R}$, wie man sofort durch Nachrechnen sieht.
Eine entsprechende Rechnung gilt für $Y_0 > -\beta/\alpha$.
Ist $Y_0 = -\beta/\alpha$, so ist $g(Y_0) = \alpha Y_0 + \beta = 0$. Also ist in diesem Fall etwa $Y(t) \equiv -\beta/\alpha = (c_0 + \bar{I} + \bar{G})/(1-c)$ eine Lösung (sogenannte triviale Lösung).

2. Wir betrachten die in der Populationsdynamik zur Modellierung von Tier- oder Pflanzenpopulationen oft verwendete sogenannte logistische Gleichung

$$y' = g(y) = ay(b-y)\,,$$

wobei $a, b > 0$ Konstanten sind. Wir suchen die Lösung des AWP

$$y' = g(y)\,, \qquad y(0) = y_0$$

mit $y_0 > 0$. Ist $y_0 = b$, so ist $g(y_0) = 0$, und damit ist $y \equiv b$ eine Lösung des AWP (unter Umständen nicht die einzige).
Es sei nun $y_0 \neq b$. Dann erhalten wir nach S. 22.4 die Lösung aus

$$\begin{aligned} t = \int_0^t ds &= \int_{y_0}^{y} \frac{ds}{g(s)} = \frac{1}{a}\int_{y_0}^{y} \frac{ds}{s(b-s)} = \frac{1}{ab}\int_{y_0}^{y} \frac{ds}{s} + \frac{1}{ab}\int_{y_0}^{y} \frac{ds}{b-s} \\ &= \frac{1}{ab}\big[\ln|y| - \ln|b-y| + \ln|b-y_0| - \ln|y_0|\big] \end{aligned}$$

d.h.

$$\left|\frac{y}{b-y}\right| = e^{abt}\left|\frac{y_0}{b-y_0}\right| .$$

Für $y_0 < b$ erhalten wir (da dann auch $y(t) < b$ für $|t|$ klein)

$$\frac{y}{b-y} = \frac{b}{b-y} - 1 = e^{abt} \cdot \frac{y_0}{b-y_0}$$

bzw.

$$y = y(t) = b\left(\frac{e^{abt}\frac{y_0}{b-y_0}}{1+e^{abt}\frac{y_0}{b-y_0}}\right) = \frac{b}{1+(\frac{b}{y_0}-1)e^{-abt}} .$$

Eine entsprechende Rechnung gilt für $y_0 > b$.

3. Wir betrachten mit $I = J = \mathbb{R}$ das AWP

$$y' = h(t)g(y) = e^t(1+y^2), \qquad y(0) = 1 .$$

Die Lösung ergibt sich wieder nach S. 22.4 (auf einer Umgebung von $t_0 = 0$) aus

$$\underbrace{e^{t-1}}_{=H(t)} = \int_0^t e^s ds = \int_1^y \frac{ds}{1+s^2} = \underbrace{\arctan(y) - \frac{\pi}{4}}_{=G(y)} .$$

also

$$\arctan(y) = e^t + \frac{\pi}{4} - 1$$

bzw.

$$y = y(t) = \tan\left(e^t + \frac{\pi}{4} - 1\right)$$

(Es gilt dabei $G(\mathbb{R}) = (-3\pi/4, \pi/4)$, also ist $H(t) \in G(\mathbb{R})$ für $t \in (-\infty, \ln(1+\pi/4))$. Nach S. 22.4 ist y Lösung auf $(-\infty, \ln(1+\pi/4))$, und dieses Intervall ist offenbar maximal. Man sieht, dass die Lösung i.a. nur auf einem echten Teilintervall von I $(=\mathbb{R})$ existiert.

Eine weitere Klasse von Differenzialgleichungen, bei denen man die Lösungen i. W. mittels Integration bestimmen kann, sind (skalare) lineare Differenzialgleichungen 1. Ordnung:
Ist $I \subset \mathbb{R}$ ein offenes Intervall und sind $a : I \to \mathbb{K}$ und $b : I \to \mathbb{K}$ stetig, so heißt

$$y' = a(t)y + b(t)$$

eine *lineare Differenzialgleichung 1. Ordnung*. Ist $b = 0$, so spricht man von einer *homogenen Gleichung*. Es gilt dafür

Satz 22.6 *Es sei $I \subset \mathbb{R}$ ein offenes Intervall, und es sei $(t_0, y_0) \in I \times \mathbb{K}$. Ferner seien $a, b : I \to \mathbb{K}$ stetig.*

1. *Das AWP zur homogenen Gleichung*

$$y' = a(t)y\,, \qquad y(t_0) = y_0$$

   hat genau eine Lösung φ *auf* I*, gegeben durch*

$$\varphi(t) = y_0 e^{A(t)} \qquad (t \in I)\,,$$

   wobei $A(t) = \int_{t_0}^{t} a(s)ds\ (t \in I)$*.*

2. *Das AWP zur (inhomogenen) Gleichung*

$$y' = a(t)y + b(t)\,, \qquad y(t_0) = y_0$$

   hat genau eine Lösung φ *auf* I*, gegeben durch*

$$\varphi(t) = e^{A(t)}\Big[y_0 + \int\limits_{t_0}^{t} e^{-A(s)}b(s)ds\Big]\,.$$

Beweis. Offensichtlich ist 1. ein Spezialfall von 2. Es genügt also, 2. zu beweisen. Es gilt für $t \in I$

$$\varphi'(t) = e^{A(t)}a(t)\left[y_0 + \int\limits_{t_0}^{t} e^{-A(s)}b(s)ds\right] + \underbrace{e^{A(t)}e^{-A(t)}}_{=1}b(t) = a(t)\varphi(t) + b(t)$$

und

$$\varphi(t_0) = e^{A(t_0)}y_0 = y_0\,,$$

d.h. φ ist Lösung des AWP auf I.
Ist $(\tilde{\varphi}, I)$ eine weitere Lösung, so betrachten wir

$$\psi(t) := (\varphi(t) - \tilde{\varphi}(t))e^{-A(t)}\,.$$

Dann gilt

$$\psi'(t) = \underbrace{(\varphi'(t) - \tilde{\varphi}'(t))}_{=a(t)(\varphi(t)-\tilde{\varphi}(t))}e^{-A(t)} - (\varphi(t) - \tilde{\varphi}(t))e^{-A(t)}a(t) \equiv 0\,,$$

also ist $\psi \equiv 0$ auf I und damit (da $e^{-A(t)} \neq 0$) auch $\varphi \equiv \tilde{\varphi}$ auf I . □

Beispiel 22.7 Wir betrachten noch einmal das AWP aus B. 22.1. Nach S. 22.6 (mit $a(t) \equiv \alpha$ und $b(t) \equiv \beta$) ist durch

$$\varphi(t) = e^{\alpha t}\Big[Y_0 + \beta\int\limits_{0}^{t} e^{-\alpha s}ds\Big] = e^{\alpha t}\Big[Y_0 - \frac{\beta}{\alpha}e^{-\alpha s}|_0^t\Big] = e^{\alpha t}\Big[Y_0 + \frac{\beta}{\alpha}\Big] - \frac{\beta}{\alpha} \qquad (t \in \mathbb{R})$$

die (eindeutig bestimmte) Lösung auf $\mathbb{R}$ gegeben (vgl. B. 22.5.1).

Zum Abschluss wollen wir uns noch mit einer Klasse von Differenzialgleichungen beschäftigen, in denen höhere Ableitungen auftreten. Zunächst wieder ein Beispiel aus der Ökonomie.

Beispiel 22.8 Wir betrachten wieder ein dynamisches Modell einer Volkswirtschaft: Das Einkommen Y verändere sich proportional zur Differenz aus Nachfrage D und Einkommen, d.h.

$$Y'(t) = k(D(t) - Y(t)).$$

Weiter betrachten wir das Zinsniveau r, dessen Veränderung proportional zur Differenz aus Geldnachfrage und Geldangebot ist. Geht man weiter davon aus, dass die Geldnachfrage proportional zu Y und das Geldangebot konstant ($M = \bar{M}$) sind, so erhalten wir

$$r'(t) = m(dY(t) - \bar{M}).$$

Schließlich ergibt sich – nach dem Modell – die Nachfrage D als Summe von privatem Konstum C, der als proportional zu Y angenommen wird ($C = cY$), und Investitionen I, die ihrerseits als $I = -ar$ mit einer Konstanten $a > 0$ angesetzt sind. Insgesamt ergibt sich, wenn man noch vereinfachend $k = m = 1$ annimmt,

$$\begin{array}{rcl} Y'(t) & = & (c-1)Y(t) - ar(t) \\ r'(t) & = & dY(t) - \bar{M} \end{array} .$$

Differenziert man die erste Gleichung und setzt die zweite ein, so folgt

$$Y''(t) = (c-1)Y'(t) - ar'(t) == (c-1)Y'(t) - adY(t) - a\bar{M} \qquad .$$

Dies ist eine sogenannte Differenzialgleichung 2. Ordnung, in der erste und zweite Ableitungen auftauchen.

Allgemeiner betrachten wir nun Gleichungen n-ter Ordnung:

Definition 22.9 Es sei $D \subset \mathbb{R} \times \mathbb{K}^n$ offen, und es sei $f \in C(D, \mathbb{K})$. Eine Gleichung der Gestalt

$$y^{(n)}(t) = f\big(t, y(t), y'(t), ..., y^{(n-1)}(t)\big) \tag{22.3}$$

oder kurz

$$y^{(n)} = f\big(t, y, y', ..., y^{(n-1)}\big)$$

heißt *(gewöhnliche) Differenzialgleichung n-ter Ordnung.*
Eine Funktion $\varphi : I \to \mathbb{K}$ (bzw. das Paar (φ, I)) heißt *Lösung* von (22.3), falls φ n-mal differenzierbar auf I ist mit $(t, \varphi(t), ..., \varphi^{(n-1)}(t)) \in D$ und

$$\varphi^{(n)}(t) = f\big(t, \varphi(t), ..., \varphi^{(n-1)}(t)\big) \qquad \text{für alle} \quad t \in I.$$

Ist $(t_0, \eta_0, ..., \eta_{n-1}) \in D$, so heißt eine Beziehung der Form

$$y^{(n)} = f\big(t, y, ..., y^{(n-1)}\big), \quad y(t_0) = \eta_0, ..., y^{(n-1)}(t_0) = \eta_{n-1} \tag{22.4}$$

ein *Anfangswertproblem* (AWP) für (22.3). Schließlich heißt eine Lösung (φ, I) von (22.3) *Lösung des AWP* (22.4), falls

$$\varphi(t_0) = \eta_0, ..., \varphi^{(n-1)}(t_n) = \eta_{n-1}$$

gilt.

Bemerkung 22.10 Man kann eine DGL n-ter Ordnung stets auf ein System von Differenzialgleichungen 1. Ordnung wie aus D. 22.2 umschreiben:
Betrachten wir $F : D \to \mathbb{K}^n$ mit

$$\begin{array}{rcc} F_1(t, y_1, ..., y_n) & = & y_2 \\ \vdots & & \vdots \\ F_{n-1}(t, y_1, ..., y_n) & = & y_n \\ F_n(t, y_1, ..., y_n) & = & f(t, y_1, ..., y_n) \end{array},$$

so sieht man sofort: Ist $\varphi : I \to \mathbb{K}$ eine Lösung von (22.3) (bzw. (22.4)), so ist

$$\Phi(t) = \begin{pmatrix} \Phi_1(t) \\ \Phi_2(t) \\ \vdots \\ \Phi_n(t) \end{pmatrix} := \begin{pmatrix} \varphi(t) \\ \varphi'(t) \\ \vdots \\ \varphi^{(n-1)}(t) \end{pmatrix} \qquad (t \in I),$$

eine Lösung von

$$y' = F(t, y)$$

(bzw. $y' = F(t, y)$ und $y(t_0) = (\eta_0, ..., \eta_{n-1})^T$). Ist umgekehrt $\Phi : I \to \mathbb{K}^n$ eine Lösung von $y' = F(t, y)$ (bzw. $y' = F(t, y)$ und $y(t_0) = (\eta_0, ..., \eta_{n-1})^T$), so ist $\varphi := \Phi_1 : I \to \mathbb{K}$ (also die 1. Komponente von Φ) eine Lösung von (22.3) (bzw. (22.4)), auf I.
Dies zeigt, dass man sich bei einer allgemeinen Lösungstheorie auf Gleichungen 1. Ordnung beschränken kann. Einer solchen allgemeinen Lösungstheorie wenden wir uns als nächstes zu.

23 Lösungstheorie für allgemeine gewöhnliche Differentialgleichungen

Wir studieren nun allgemeine Anfangswertprobleme wie in (22.2), d.h. für eine gegebene Funktion $f \in C(D, \mathbb{K}^d)$, wobei $D \subset \mathbb{R} \times \mathbb{K}^d$ offen ist, und für $(t_0, y^0) \in D$ betrachten wir das AWP

$$y' = f(t, y) \qquad y(t_0) = y^0 .$$

In 22.3 hatten wir gesehen, dass nicht jedes solche AWP eine eindeutige Lösung besitzt. Im folgenden wollen wir zeigen, dass unter etwas stärkeren Voraussetzungen an f (außer der Stetigkeit) stets genau eine sog. maximale Lösung des AWPs existiert.

Definition 23.1 Es sei $D \subset \mathbb{R} \times \mathbb{K}^d$ offen, und es sei $f \in C(D, \mathbb{K}^d)$. Man sagt, f *genügt* auf D *einer lokalen Lipschitz-Bedingung bezüglich* y, falls für alle $K \subset D$, K kompakt, eine Konstante $L = L(K)$ existiert mit

$$|f(t, y) - f(t, \tilde{y})| \le L|y - \tilde{y}|$$

für alle (t, y) und $(t, \tilde{y}) \in K$ (wobei $|\cdot|$ stets euklidische Norm ist).

Bemerkung 23.2 1. Es gilt: f genügt auf D einer lokalen Lipschitz-Bedingung bezüglich y genau dann, wenn zu jedem $(t_0, y^0) \in D$ eine Umgebung $U = U(t_0, y^0)$ von (t_0, y^0) und eine Konstante $L = L(U)$ existieren mit

$$|f(t, y) - f(t, \tilde{y})| < L|y - \tilde{y}| \qquad \text{für alle} \quad (t, y), (t, \tilde{y}) \in U .$$

(Denn: "$\Rightarrow$" ist klar (wähle $U = U_\delta(t_0, y^0)$ so, dass $\bar{U} \subset D$).
"$\Leftarrow$" Es sei $K \subset D$ kompakt. Angenommen, es existiert keine Konstante L wie gewünscht. Dann existieren Folgen $(t_n, y^{(n)})$ und $(t_n, \tilde{y}^{(n)})$ in K und eine Folge $L_n \to \infty$ mit

$$|f(t_n, y^{(n)}) - f(t_n, \tilde{y}^{(n)})| > L_n |y^{(n)} - \tilde{y}^{(n)}| \qquad (n \in \mathbb{N}) .$$

Da K kompakt ist, besitzt $(t_n, y^{(n)})$ eine konvergente Teilfolge mit Grenzwert (t_0, y^0) in K. O. E. können wir annehmen, dass die Folge $(t_n, y^{(n)})$ selbst konvergiert. Aus

$$L_n |y^{(n)} - \tilde{y}^{(n)}| \le 2 \max_{(t,y) \in K} |f(t, y)| \qquad (n \in \mathbb{N})$$

folgt, dass auch $\tilde{y}^{(n)} \to y^0$ gilt. Nach Voraussetzung existieren eine Umgebung U von (t_0, y^0) und ein $L > 0$ mit

$$|f(t, y) - f(t, \tilde{y})| \le L|y - \tilde{y}| \qquad \text{für alle } (t, y), (t, \tilde{y}) \in U .$$

Da $(t_n, y^{(n)})$ und $(t_n, \tilde{y}^{(n)})$ für n genügend groß in U liegen, ergibt sich ein Widerspruch.)

2. Ist $\mathbb{K} = \mathbb{R}$, existiert für jedes $(t, y) \in D$

$$\frac{\partial f}{\partial y}(t,y) := \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(t,y) & \cdots & \frac{\partial f_1}{\partial y_d}(t,y) \\ \cdots & & \cdots \\ \frac{\partial f_d}{\partial y_1}(t,y) & \cdots & \frac{\partial f_d}{\partial y_d}(t,y) \end{pmatrix}$$

und sind sämtliche partiellen Ableitungen $\partial f_j/\partial y_k : D \to \mathbb{R}$ stetig auf D, so genügt f auf D einer lokalen Lipschitz-Bedingung bezüglich y.
(Denn: Wie im Beweis zu Satz 19.13 sieht man, dass $\partial f/\partial y : D \to \mathbb{R}^{d\times d}$ stetig ist (wobei $\mathbb{R}^{d\times d}$ mit der von der Operatornorm herkommenden Metrik versehen ist). Ist $(t_0, y^0) \in D$, so existiert eine Kugel $U_\delta(t_0, y^0)$ mit $\bar{U} \subset D$. Da $\bar{U}$ kompakt ist, existiert

$$L := L(U) := \max_{(t,y)\in\bar{U}} \| \frac{\partial f}{\partial y}(t,y) \| \ .$$

Nach S.20.6 (angewandt auf die Funktionen $y \mapsto f(t,y)$ bei festem t; wichtig dabei $\{y \in \mathbb{R}^d : (t,y) \in U\}$ ist konvex) gilt

$$|f(t,y) - f(t,\tilde{y})| \le L|y - \tilde{y}| \qquad ((t,y),(t,\tilde{y}) \in U)\,.$$

Also folgt aus 1. die Behauptung.)

Beispiel 23.3 1. Es sei

$$f(t,y) = e^t(1+y^2) \qquad (t, y \in \mathbb{R})$$

(vgl. B. 22.5.3). Dann gilt

$$\frac{\partial f}{\partial y}(t,y) = e^t 2y \qquad (t, y \in \mathbb{R})\,.$$

Insbesondere ist $\partial f/\partial y$ stetig auf $\mathbb{R}^2$. Nach B. 23.2.2 genügt f auf $\mathbb{R}^2$ einer lokalen Lipschitz-Bedingung bezüglich y.
Man beachte jedoch: Für $t_0 \in \mathbb{R}$ fest existiert <u>kein</u> $L > 0$ so, dass

$$|f(t_0,y) - f(t_0,\tilde{y})| = e^{t_0}|y+\tilde{y}||y-\tilde{y}| \le L|y-\tilde{y}|$$

für alle $y, \tilde{y} \in \mathbb{R}$ erfüllt ist (f genügt keiner globalen Lipschitz-Bedingung bzgl. y).

2. Es sei

$$f(t,y) := \begin{cases} \sqrt{y}, & y \ge 0 \\ 0, & y < 0 \end{cases}, \qquad t \in \mathbb{R}$$

(vgl B. 22.3). Dann gilt für alle t und $y, \tilde{y} > 0$

$$|f(t,y) - f(t,\tilde{y})| = |\sqrt{y} - \sqrt{\tilde{y}}| = \frac{|y-\tilde{y}|}{\sqrt{y}+\sqrt{\tilde{y}}}\,.$$

Ist $L > 0$ beliebig, so gilt für $y, \tilde{y}$ genügend klein

$$\frac{1}{\sqrt{y}+\sqrt{\tilde{y}}} > L\,.$$

Also existieren <u>keine</u> Umgebung U von $(0,0)$ und $L \geq 0$ so, dass

$$|f(t,y) - f(t,\tilde{y})| \leq L|y-\tilde{y}| \qquad \text{für alle} \quad (t,y),(t,\tilde{y}) \in U$$

erfüllt ist. Folglich genügt f auf keiner offenen Menge D, die $(0,0)$ enthält, einer lokalen Lipschitz-Bedingung bezüglich y.

Wir zeigen nun, dass das AWP (22.2) äquivalent ist zu einer gewissen Integralgleichung. Um dies für $\mathbb{K}^d$-wertige Funktionen formulieren zu können, setzen wir für $f = (f_1, ..., f_d) = [a,b] \to \mathbb{K}^d$, mit $f_j \in R[a,b]$ für $j = 1, ..., d$

$$\int\limits_a^b f(s)ds := \left(\int\limits_a^b f_1(s)ds, ..., \int\limits_a^b f_d(s)ds \right)\,.$$

Dann hat $\int$ die üblichen Eigenschaften skalarer Integrale; einzig nichttrivial ist dabei $|\int_a^b f| \leq \int_a^b |f|$:
Es gilt für alle $f = (f_1, ..., f_d)$ mit $f_j \in R[a,b]$

$$\Big|\int\limits_a^b f(s)ds\Big| \leq \int\limits_a^b |f(s)|ds\,.$$

(Denn: Es sei

$$u := \int\limits_a^b f(s)ds \in \mathbb{K}^d\,.$$

Dann gilt mit der Cauchy-Schwarzschen Ungleichung

$$\begin{aligned}
|u|^2 &= \bar{u}^T u = \sum_{j=1}^d \bar{u}_j \int\limits_a^b f_j(s)ds = \int\limits_a^b \left(\sum_{j=1}^d \bar{u}_j f_j(s) \right) ds \\
&= \int\limits_a^b \bar{u}^T f(s)ds = \operatorname{Re} \int\limits_a^b \bar{u}^T f(s)ds = \int\limits_a^b \operatorname{Re}(\bar{u}^T f(s))ds \\
&\leq \int\limits_a^b |\bar{u}^T f(s)|ds \leq \int\limits_a^b |u|\,|f(s)|ds\,,
\end{aligned}$$

also

$$|u| \leq \int_a^b |f(s)| ds \,.)$$

Satz 23.4 *Es sei $D \subset \mathbb{R} \times \mathbb{K}^d$ offen, und es sei $f \in C(D, \mathbb{K}^d)$. Ferner sei $(t_0, y^0) \in D$. Ist $I \subset \mathbb{R}$ ein Intervall mit $t_0 \in I$, so sind für $\varphi \in C(I, \mathbb{K}^d)$ äquivalent:*

a) (φ, I) ist eine Lösung von (22.2), d.h. $\varphi(t_0) = y^0$ und

$$\varphi'(t) = f(t, \varphi(t)) \qquad (t \in I)\,.$$

b) Für alle $t \in I$ gilt

$$\varphi(t) = y^0 + \int_{t_0}^t f(s, \varphi(s)) ds \,.$$

Beweis. a) $\Rightarrow$ b): Es sei $\varphi = (\varphi_1, ..., \varphi_d)$, $y^0 = (y_1^0, ..., y_d^0)$. Aus $\varphi_j'(s) = f_j(s, \varphi(s))$ und $\varphi_j(t_0) = y_j^0$ ergibt sich durch Aufintegrieren und Anwendung des HDI (Teil 2) für alle $t \in I$ und $j = 1, ..., d$

$$\varphi_j(t) - y_j^0 = \varphi_j(t) - \varphi_j(t_0) = \int_{t_0}^t \varphi_j'(s) ds = \int_{t_0}^t f_j(s, \varphi(s)) ds \,.$$

(Man beachte dabei φ_j' ist stetig auf I, da $s \mapsto f_j(s, \varphi(s))$ stetig ist.)
b) $\Rightarrow$ a): Da φ stetig auf I und f stetig auf D sind, ist $s \mapsto f(s, \varphi(s))$ stetig auf I. Also ergibt sich a) wieder durch Anwendung des HDI (diesmal Teil 1). □

Damit können wir folgende (zunächst „lokale“) Version eines Existenz- und Eindeutigkeitssatzes für Anfangswertprobleme beweisen.

Satz 23.5 *(Picard-Lindelöf; lokale Version)*
Es sei $D \subset \mathbb{R} \times \mathbb{K}^d$ offen, und es sei $f \in C(D, \mathbb{K}^d)$. Ferner genüge f auf D einer lokalen Lipschitz-Bedingung bezüglich y. Dann existiert für jedes $K \subset D$, K kompakt, ein $\alpha_0 = \alpha_0(K) > 0$ so, dass das AWP

$$y' = f(t, y)\,, \qquad y(\xi) = \eta$$

für jedes $(\xi, \eta) \in K$ und jedes Intervall $I \subset [\xi - \alpha_0, \xi + \alpha_0]$ mit $\xi \in I$ genau eine Lösung auf I besitzt.

Beweis. Unser Beweis beruht auf einer geeigneten Anwendung des Banachschen Fixpunktsatzes.

1. Es sei $K \subset D$ kompakt. Wir wählen eine offene und beschränkte Menge V mit

$$K \subset V \subset \bar{V} \subset D\,.$$

(Eine solche existiert: Ist $D = \mathbb{R} \times \mathbb{K}^d$, so ist die Existenz klar. Ist $D \neq \mathbb{R} \times \mathbb{K}^d$ (also $D^c \neq \emptyset$), so ist

$$\operatorname{dist}(K, D^c) > 0$$

([Ü]). Für $0 < \gamma < \operatorname{dist}(K, \partial D)$ hat dann etwa

$$V = \bigcup_{(\xi,\eta)\in K} U_\gamma((\xi,\eta))$$

die gewünschten Eigenschaften.)
Da $\operatorname{dist}(K, V^c) > 0$ ist, existieren $\alpha, \beta > 0$ so, dass für alle $(\xi, \eta) \in K$ gilt

$$R(\xi,\eta) := \{(t,y) : |t-\xi| \le \alpha,\ |y-\eta| \le \beta\} \subset V\,.$$

Wir setzen weiter (beachte: $\bar{V}$ kompakt)

$$M := \max_{(t,y)\in\bar{V}} |f(t,y)|$$

und (mit $L = L(\bar{V})$ wie in D. 23.1 und $\varrho/0 := \infty$ für $\varrho > 0$)

$$\alpha_0 := \min\left(\alpha, \frac{\beta}{M}, \frac{1}{2L}\right)\,.$$

2. Es sei $(\xi, \eta) \in K$ fest, und es sei I ein kompaktes Intervall mit $\xi \in I$ und

$$I \subset [\xi - \alpha_0, \xi + \alpha_0]\,.$$

Wir setzen

$$A := \{\varphi \in C(I, \mathbb{K}^d) : |\varphi(t) - \eta| \le \beta \text{ für alle } t \in I\}\,.$$

Nach B. 15.9 ist $(B(I, \mathbb{K}^d), d)$ mit

$$d(f,g) := \| f - g \|_\infty$$

ein vollständiger metrischer Raum. Weiter ist $A \subset B(I, \mathbb{K}^d)$ abgeschlossen.
(Denn: Es sei $(\varphi_n)_n$ eine Folge in A mit $\varphi_n \to \varphi$ für ein $\varphi \in B(I, \mathbb{K}^d)$. Dann ist nach S. 15.10 φ stetig auf I, d.h. $\varphi \in C(I, \mathbb{K}^d)$. Außerdem gilt für alle $t \in I$, $n \in \mathbb{N}$

$$|\varphi(t) - \eta| \le \underbrace{|\varphi_n(t) - \eta|}_{\le\beta} + \underbrace{|\varphi_n(t) - \varphi(t)|}_{\to 0 \quad (n\to\infty)}\,,$$

also auch $|\varphi(t) - \eta| \le \beta$ und damit $\varphi \in A$. Nach S. 9.15 ist A abgeschlossen.)
Als abgeschlossene Teilmenge des vollständigen Raumes $(B(I, \mathbb{K}^d), d)$ ist (A, d) ebenfalls vollständig ([Ü]).

3. Wir definieren für $\varphi \in A$

$$T\varphi(t) := \eta + \int_\xi^t f(s,\varphi(s))ds \qquad (t \in I)\,.$$

(Man beachte: aus $\varphi \in A$ folgt $(s,\varphi(s)) \in R(\xi,\eta) \subset V \subset D$ für alle $s \in I$).
Da $s \mapsto f(s,\varphi(s))$ stetig auf I ist, ist auch $T\varphi$ stetig auf I nach dem HDI (sogar differenzierbar). Weiter gilt für $t \in I$

$$\begin{aligned} |T\varphi(t) - \eta| &= \Big|\int_\xi^t f(s,\varphi(s))ds\Big| \le \Big|\int_\xi^t \underbrace{|f(s,\varphi(s))|}_{\le M} ds\Big| \\ &\le M \cdot |t-\xi| \le M \cdot \alpha_0 \le \beta\,, \end{aligned}$$

also ist $T\varphi \in A$. Damit gilt $T : A \to A$.
4. Behauptung: $T : A \to A$ ist eine 1/2-Kontraktion.
Denn: Für $\varphi, \tilde\varphi \in A$ und $t \in I$ gilt

$$\begin{aligned} |T\varphi(t) - T\tilde\varphi(t)| &\le \Big|\int_\xi^t |f(s,\varphi(s)) - f(s,\tilde\varphi(s))|ds\Big| \le \\ &\le L\Big|\int_\xi^t |\varphi(s) - \tilde\varphi(s)|ds\Big| \le L \parallel \varphi - \tilde\varphi \parallel_\infty |t-\xi| \\ &\le L \parallel \varphi - \tilde\varphi \parallel_\infty \cdot \alpha_0 \le \frac{1}{2} \parallel \varphi - \tilde\varphi \parallel_\infty \,. \end{aligned}$$

5. Nach dem Banachschen Fixpunktsatz hat T genau einen Fixpunkt $\varphi \in A$, d.h.

$$\varphi(t) = \eta + \int_\xi^t f(s,\varphi(s))ds \qquad (t \in I)\,.$$

Also ist φ nach S. 23.4 die eindeutig bestimmte Lösung des AWPs $y' = f(t,y)$, $y(\xi) = \eta$ auf I, jedenfalls mit $|\varphi(t) - \eta| \le \beta$. Dies gilt aber notwendigerweise für jede Lösung ([Ü]). □

Bemerkung 23.6 1. Man kann zeigen, dass auch ohne die Voraussetzung einer lokalen Lipschitz-Bedingung die Existenz einer Lösung des AWP (22.2) auf einer Umgebung eines jeden Anfangswertepaares $(t_0, y^0) \in D$ gesichert ist. (Existenzsatz von Peano). Wie etwa B. 22.3 zeigt, sind die Lösungen in diesem Fall allerdings nicht mehr eindeutig.

Auf den Beweis des Satzes von Peano, der weitergehende Hilfsmittel der Analysis erfordert, wollen wir nicht eingehen.

2. Mit den Bezeichnungen aus dem Beweis zu S. 23.5 ergibt sich mit dem Banachschen Fixpunktsatz folgendes interessante Verfahren zur näherungsweisen Berechnung der Lösung des AWP (22.2) auf einer Umgebung von t_0:
Ist $\varphi_0 \in A$ (etwa $\varphi_0(t) \equiv y^0$), so konvergiert die Folge (φ_0) in A mit

$$\varphi_{n+1}(t) := T\varphi_n(t) = y^0 + \int_{t_0}^{t} f(s, \varphi_n(s))ds$$

für $|t - t_0|$ genügend klein gegen die Lösung φ. Außerdem ergibt sich aus dem Banachschen Fixpunktsatz auch eine Abschätzung für den Fehler $\| \varphi - \varphi_n \|_\infty$.
Dieses iterative Näherungsverfahren zur Bestimmung der Lösung heißt „Picard-Lindelöfsches Iterationsverfahren". Für das einfache Beispiel $y' = \alpha y$, $y(0) = 1$ erhalten wir etwa mit $\varphi_0 \equiv 1$

$$\varphi_n(t) = \sum_{\nu=0}^{n} \frac{\alpha^\nu t^\nu}{\nu!} \, ,$$

also $\varphi(t) = \lim\limits_{n\to\infty} \varphi_n(t) = e^{\alpha t}$ (hier sogar für alle $t \in \mathbb{R}$).

Definition 23.7 1. Es seien (φ, I) und $(\tilde{\varphi}, \tilde{I})$ Lösungen des AWP (22.2). Dann heißt $(\tilde{\varphi}, \tilde{I})$ (*echte*) *Fortsetzung* von (φ, I), falls $\tilde{I} \supset I$ $(\tilde{I} \neq I)$ und $\tilde{\varphi}_{|I} = \varphi$ gilt.
2. Eine Lösung (φ_0, I_0) von (22.2) heißt *maximal*, falls (φ_0, I_0) Fortsetzung jeder Lösung (φ, I) von (22.2) ist. Das Intervall I_0 heißt dann *maximales Existenzintervall* der Lösung des AWP. (Man beachte: (φ_0, I_0) ist eindeutig.)

Es gilt damit

Satz 23.8 *(Picard-Lindelöf, globale Version)*
Es sei $D \subset \mathbb{R} \times \mathbb{K}^d$ offen, und es sei $f \in C(D, \mathbb{K}^d)$. Ferner genüge f auf D einer lokalen Lipschitz-Bedingung bezüglich y. Dann existiert zu jedem $(t_0, y^0) \in D$ (genau) eine maximale Lösung (φ_0, I_0) des AWPs (22.2). Weiter gilt: I_0 ist offen, und die maximale Lösung „verlässt jede kompakte Teilmenge von D", d.h., zu jedem $K \subset D$, K kompakt, existieren $T_1, T_2 \in I_0$, $T_1 < T_2$ mit

$$(t, \varphi_0(t)) \notin K$$

für alle $t < T_1$ und $t > T_2$.

Beweis.
1. Wir zeigen zunächst: Sind (φ_1, I_1) und (φ_2, I_2) Lösungen des AWPs (22.2), so gilt

$$\varphi_1(t) = \varphi_2(t) \qquad (t \in I_1 \cap I_2)\,.$$

Denn: Angenommen, es existiert ein $\bar{t} \in I_1 \cap I_2$ mit $\varphi_1(\bar{t}) \neq \varphi_2(\bar{t})$. O.E. betrachten wir den Fall $\bar{t} > t_0$. Dann setzen wir

$$\bar{s} := \inf\{t \in I_1 \cap I_2 : t > t_0,\ \varphi_1(t) \neq \varphi_2(t)\}\,.$$

Nach S. 23.5 ist $t_0 < \bar{s}(\leq \bar{t})$ und nach Definition gilt

$$\varphi_1(t) = \varphi_2(t) \qquad \text{für alle} \quad t \in [t_0, \bar{s})\,,$$

und da φ_1 und φ_2 stetig auf $[t_0, \bar{s}] \subset I_1 \cap I_2$ sind, gilt auch

$$\varphi_1(\bar{s}) = \varphi_2(\bar{s})\,.$$

Damit ist $\bar{s}$ kein Randpunkt von $I_1 \cap I_2$ und φ_1 und φ_2 sind Lösungen von $y' = f(t, y)$, $y(\bar{s}) = \varphi_1(\bar{s})(= \varphi_2(\bar{s}))$ auf einer Umgebung von $\bar{s}$. Nach S. 23.5 muss dann aber $\varphi_1(t) = \varphi_2(t)$ auf einer Umgebung von $\bar{s}$ gelten, im Widerspruch zur Definition von $\bar{s}$.
2. Es sei I_0 die Vereinigung aller Intervalle I mit $t_0 \in I$ und so, dass auf I eine Lösung $\varphi = \varphi_I$ existiert (solche Intervalle existieren nach S. 23.5). Für $t \in I_0$ setzen wir

$$\varphi_0(t) := \varphi_I(t) \qquad \text{falls} \quad t \in I\,.$$

Dann ist φ_0 nach 1. wohldefiniert, denn sind $\varphi = \varphi_I$ und $\tilde{\varphi} = \tilde{\varphi}_{\tilde{I}}$ zwei Lösungen mit $t \in I \cap \tilde{I}$, so gilt $\tilde{\varphi}(t) = \varphi(t)$.
Außerdem ergibt sich aus der Definition, dass φ_0 maximale Lösung von (22.2) auf I_0 ist. Schließlich ist I_0 auch offen, denn angenommen, es existiert ein Randpunkt b von I_0 mit $b \in I_0$ (o.E. b rechter Randpunkt). Nach S. 23.5 hat das AWP

$$y' = f(t, y), \qquad y(b) = \varphi_0(b)$$

eine Lösung $\tilde{\varphi}$ auf $[b - \alpha_0, b + \alpha_0]$ für ein $\alpha_0 > 0$. Dann ist durch

$$\bar{\varphi}(t) := \begin{cases} \varphi_0(t)\,, & t \in [t_0, b] \\ \tilde{\varphi}(t)\,, & t \in (b, b + \alpha_0] \end{cases}$$

eine Lösung von (22.2) auf $[t_0, b+\alpha_0]$ gegeben (wichtig dabei: $\bar{\varphi}'(b) = f(b, \varphi_0(b))$, auch „rechtsseitig“). Dies widerspricht der Maximalität von I_0.
3. Es sei $K \subset D$ kompakt. Angenommen, es existieren keine T_1, T_2 wie gefordert. Wir setzen $I_0 := (a, b)$.
Dann existiert eine Folge (t_n) in I_0 mit $t_n \uparrow b$ oder $t_n \downarrow a$ und

$$(t_n, \varphi_0(t_n)) \in K \qquad (n \in \mathbb{N})\,.$$

O.E. gelte $t_n \uparrow b$. Dann ist insbesondere $b < \infty$.
Es sei nun $\alpha_0 = \alpha_0(K)$ wie in S. 23.5. Wir wählen ein $N \in \mathbb{N}$ mit $b - t_N < \alpha_0$. Wie in Beweisschritt 2. (Anwendung von S. 23.5 auf das AWP

$$y' = f(t, y), \qquad y(t_N) = \varphi_0(t_N))$$

sieht man, dass φ_0 zu einer Lösung auf $[t_0, t_N + \alpha_0]$ fortgesetzt werden kann.
Da $t_N + \alpha_0 > b$ ist, ergibt sich ein Widerspruch zur Maximalität von I_0. □

Bemerkung 23.9 Unter den Voraussetzungen von S. 23.8 existiert also zu jedem Punkt $(t_0, y^0) \in D$ eine eindeutig bestimmte maximale Lösung (φ_0, I_0). Wir schreiben statt φ_0 auch $\varphi(\cdot, t_0, y^0)$ und statt I_0 auch $I_{(t_0,y^0)}$. Die dadurch definierte Funktion $\varphi : \Omega \to \mathbb{K}^d$ mit

$$\Omega := \bigcup_{(t_0,y^0)\in D} \{(t, t_0, y^0) : t \in I(t_0, y^0)\} \subset \mathbb{R} \times \mathbb{R} \times \mathbb{K}^d$$

betrachten wir als die „allgemeine Lösung“ der Differentialgleichung $y' = f(t, y)$.
Man kann dafür zeigen: Ω ist offen und φ ist stetig auf Ω, d.h. die Lösungen der AWPe $y' = f(t, y)$, $y(t_0) = y^0$ hängen insbesondere „stetig von den Anfangswerten ab“.
Auf den Beweis verzichten wir.

Beispiel 23.10 1. Es sei $D = \mathbb{R} \times \mathbb{R}$. Für festes $\lambda \in \mathbb{R}$ hat das AWP

$$y' = \lambda y\,, \qquad y(t_0) = y_0$$

die maximale Lösung

$$\varphi(t, t_0, y_0) = y_0 e^{\lambda(t-t_0)} \quad \text{auf} \quad I_{(t_0,y_0)} = \mathbb{R}\,.$$

2. Es sei $D = \mathbb{R} \times \mathbb{R}$. Wir betrachten das AWP

$$y' = y^2\,, \qquad y(t_0) = y_0$$

für $(t_0, y^0) \in \mathbb{R} \times \mathbb{R}$. Nach S. 22.4 ergibt sich die Lösung (jedenfalls lokal) für $y_0 \neq 0$ durch Auflösen von

$$t - t_0 = \int_{t_0}^{t} ds = \int_{y_0}^{y} \frac{ds}{s^2} = -\frac{1}{y} + \frac{1}{y_0}\,,$$

also

$$y = y(t) = \frac{y_0}{1 - y_0(t - t_0)} \qquad \text{für } |t - t_0| \text{ klein}\,.$$

Man sieht (durch Differenzieren), dass dabei gilt

$$\varphi(t,t_0,y_0) = \frac{y_0}{1-y_0(t-t_0)} \quad \text{für} \quad \begin{cases} t > t_0 + 1/y_0\,, & \text{falls } y_0 < 0 \\ t \in \mathbb{R}\,, & \text{falls } y_0 = 0 \\ t < t_0 + 1/y_0\,, & \text{falls } y_0 > 0 \end{cases},$$

also

$$I_{(t_0,y_0)} = \begin{cases} (t_0 + 1/y_0, \infty)\,, & \text{falls } y_0 < 0 \\ (-\infty,\infty)\,, & \text{falls } y_0 = 0 \\ (-\infty, t_0 + 1/y_0) & \text{falls } y_0 > 0 \end{cases}.$$

Obwohl $D = \mathbb{R} \times \mathbb{R}$ ist, sind die maximalen Lösungsintervalle dabei i.a. nicht ganz $\mathbb{R}$; die Lösungen haben eine „endliche Entweichzeit".
Man beachte auch: Alle Lösungen verlassen jede kompakte Teilmenge von $\mathbb{R}^2$, sowohl, wenn t sich dem rechten Randpunkt, als auch, wenn t sich dem linken Randpunkt von $I_{(t_0,y_0)}$ annähert. Genauer gilt hier

$$\varphi(t,t_0,y_0) \to \begin{cases} -\infty & \text{für } t \to (t_0 + 1/y_0)^+\,, & \text{falls } y_0 < 0 \\ \infty & \text{für } t \to (t_0 + 1/y_0)^-\,, & \text{falls } y_0 > 0 \end{cases}.$$

Komplizierter ist das Randverhalten bei folgendem (ersten höherdimensionalen) Beispiel.
3. Es sei $D = (0,\infty) \times \mathbb{R}^2$ und

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = f(t,y_1,y_2) = \begin{pmatrix} -y_2/t^2 \\ y_1/t^2 \end{pmatrix}, \qquad y\left(\frac{1}{\pi}\right) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Dann ist

$$\varphi\left(t, \frac{1}{\pi}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}\right) = \begin{pmatrix} \sin(1/t) \\ \cos(1/t) \end{pmatrix} \qquad (t \in (0,\infty))$$

die maximale Lösung des AWP. Hier existiert $\lim\limits_{n\to 0^+} \varphi(t)$ nicht!

Unter starken Voraussetzungen an f kann man eine Aussage über die Größe der maximalen Lösungsintervalle machen.

Satz 23.11 *Es sei $D = I \times \mathbb{K}^d$, wobei $I \subset \mathbb{R}$ ein offenes Intervall ist. Ferner genüge $f \in C(D,\mathbb{K}^d)$ einer lokalen Lipschitz-Bedingung bezüglich y, und es gelte*

$$|f(t,y)| \le \varrho(t)|y| + \sigma(t) \qquad (t \in I)$$

mit stetigen Fuktionen $\varrho, \sigma : I \to (0,\infty)$. Dann gilt für alle $(t_0, y^0) \in D$:

$$I_{(t_0,y^0)} = I\,.$$

Beim Beweis verwenden wir das äußerst nützliche sogenannte „Lemma von Gronwall“:

Satz 23.12 *Es sei $\psi \in C([0,T))$ für ein $T > 0$, und es gelte*

$$\psi(t) \leq A + B \int_0^t \psi(s)ds \qquad (t \in [0,T))$$

für gewisse Konstanten $A \in \mathbb{R}$ und $B > 0$. Dann ist

$$\psi(t) \leq Ae^{Bt} \qquad (t \in [0,T))\,.$$

Beweis. Es sei $\delta > 0$ und

$$g(t) = g_\delta(t) = (A+\delta)e^{Bt} \qquad (t \in [0,T))\,.$$

Dann gilt

$$g(t) = A + \delta + B \int_0^t g(s)ds \qquad (t \in [0,T))\,.$$

Wir zeigen: $\psi(t) < g(t)$ auf $[0,T)$. (Da $\delta > 0$ beliebig war, ergibt sich hieraus die Behauptung.)
Für $t = 0$ ist jedenfalls $\psi(0) < g(0)$. Angenommen, es existiert ein $t_0 > 0$ mit $\psi(t_0) \geq g(t_0)$. Für

$$t_1 := \inf\{t_0 > 0,\ \psi(t_0) \geq g(t_0)\} \quad (> 0)$$

gilt dann $\psi(t_1) = g(t_1)$ und $\psi(t) \leq g(t)$ für $t \in [0,t_1]$ und deshalb

$$\psi(t_1) \leq A + B \int_0^{t_1} \psi(s)ds < A + \delta + B \int_0^{t_1} g(s)ds = g(t_1)\,.$$

Widerspruch! □

Beweis. (**zu S. 23.11**) Angenommen, $(\alpha,\beta) := I_{(t_0,y^0)} \neq I =: (a,b)$. O.E. sei dann $\beta < b$. Wir setzen

$$R := \max_{t\in[0,\beta]} \varrho(t), \qquad S := \int_{t_0}^{\beta} \sigma(t)dt\,.$$

Dann gilt für $t_0 \leq t < \beta$

$$\varphi(t) = \varphi(t,t_0,y^0) = y^0 + \int_{t_0}^{t} f(s,\varphi(s))ds$$

und damit

$$\begin{aligned} |\varphi(t)| &\leq |y^0| + \int_{t_0}^{t} \varrho(s)|\varphi(s)|ds + \int_{t_0}^{t} \sigma(s)ds \\ &\leq |y^0| + S + R\int_{t_0}^{t} |\varphi(s)|ds \qquad (t \in [t_0, \beta))\,. \end{aligned}$$

Nach dem Lemma von Gronwall (angewandt auf $\psi(t) := |\varphi(t + t_0)|$) gilt

$$\begin{aligned} |\varphi(t)| &\leq (|y^0| + S)e^{R(t-t_0)} \\ &\leq (|y^0| + S)e^{R(\beta-t_0)} \qquad (t \in [t_0, \beta))\,, \end{aligned}$$

also ist φ beschränkt auf $[t_0, \beta)$. Das widerspricht aber der Tatsache, dass $\varphi_{|[t_0,\beta)}$ nach S. 23.8 „jede kompakte Teilmenge von $D = I \times \mathbb{R}^d$ verlässt“. $\square$

24 Allgemeine lineare Differenzialgleichungen

Bereits in Abschnitt 22 hatten wir uns kurz mit (skalaren) linearen Differenzialgleichungen beschäftigt. Wir untersuchen jetzt den wesentlich allgemeineren Fall von Systemen linearer Differenzialgleichungen.
Es seien im Folgenden stets $I \subset \mathbb{R}$ ein offenes Intervall und

$$A = (a_{jk}) : I \to \mathbb{K}^{d\times d}$$

sowie

$$b : I \to \mathbb{K}^d$$

stetig. Eine Differenzialgleichung der Form

$$y' = A(t)y + b(t) \tag{24.1}$$

nennen wir ein *lineares System* (von Differenzialgleichungen) oder kurz *lineare Differenzialgleichung*. Die Gleichung

$$y' = A(t)y \tag{24.2}$$

heißt *zugehörige homogene Gleichung*. Ist $b = 0$, so heißt (24.1) *homogen*. Meist betrachten wir auch jetzt wieder zugehörige AWPe der Form

$$y' = A(t)y + b(t)\,, \qquad y(\xi) = \eta \tag{24.3}$$

für $\xi \in I, \eta \in \mathbb{K}^d$.
Aus den Ergebnissen des vorigen Abschnittes erhalten wir unmittelbar

Satz 24.1 *Es seien $I \subset \mathbb{R}$ ein offenes Intervall und $A : I \to \mathbb{K}^{d\times d}, b : I \to \mathbb{K}^d$ stetig. Dann hat für jedes $(\xi, \eta) \in I \times \mathbb{K}^d$ das AWP (24.3), genau eine maximale Lösung $\varphi = \varphi(\cdot, \xi, \eta)$ mit $I_{(\xi,\eta)} = I$.*

Beweis. Wir betrachten $f : I \times \mathbb{K}^d \to \mathbb{K}^d$

$$f(t, y) := A(t)y + b(t) \qquad (t \in I, y \in \mathbb{K}^d)\,.$$

Dann ist f stetig (warum ?), und es gilt

$$|f(t, y) - f(t, \tilde{y})| = |A(t)(y - \tilde{y})| \leq ||A(t)||\, |y - \tilde{y}| \qquad (t \in I, y \in \mathbb{K}^d)\,.$$

Ist $I_0 \subset I$ kompakt, so existiert

$$L := \max_{t\in I_0} ||A(t)||\,.$$

Also gilt
$$|f(t,y)-f(t,\tilde{y})| \le L|y-\tilde{y}| \qquad (t \in I_0, y \in \mathbb{K}^d)\,.$$
Insbesondere impliziert dies, dass f auf $I \times \mathbb{K}^d$ einer lokalen Lipschitz-Bedingung bezüglich y genügt. Ferner gilt
$$|f(t,y)| \le ||A(t)||\,|y| + |b(t)| \qquad (t \in I, y \in \mathbb{K}^d)\,.$$
Aus S. 23.11 (angewandt mit $\rho(t) = ||A(t)||$ und $\sigma(t) = |b(t)|$) ergibt sich die Behauptung. □

Wir wollen uns nun die Struktur der Lösungsgesamtheit von (24.1) genauer anschauen. Dann betonen wir die Abhängigkeit von der Inhomogenität b und schreiben $\varphi = \varphi_b(\cdot,\xi,\eta)$ für die maximale Lösung von (24.3). Außerdem setzen wir
$$L_b := \{\varphi_b(\cdot,\xi,\eta) : \xi \in I, \eta \in \mathbb{K}^d\} = \{\varphi : \varphi \,\text{löst (24.1) auf}\, I\}\,,$$
also insbesondere
$$L_0 := \{\varphi_0(\cdot,\xi,\eta) : \xi \in I, \eta \in \mathbb{K}^d\} = \{\varphi : \varphi \,\text{löst (24.2) auf}\, I\}\,.$$
Es gilt

Satz 24.2 *1. L_0 ist ein d-dimensionaler Unterraum von $C(I,\mathbb{K}^d)$, und für jedes $\xi \in I$ ist die Abbildung $\eta \mapsto \varphi_0(\cdot,\xi,\eta)$ von $\mathbb{K}^d$ auf L_0 ein Isomorphismus (linear und bijektiv).*

2. Weiter gilt: Ist $\varphi \in L_b$ fest, so gilt
$$L_b = \varphi + L_0 \ (:= \{\varphi + \psi : \psi \in L_0\})\,,$$
d.h. L_b ist ein (d-dimensionaler) affiner Unterraum von $C(I,\mathbb{K}^d)$.

Beweis. 1. Wir zeigen: Die Abbildung $T = T_\xi : \mathbb{K}^d \to C(I,\mathbb{K}^d)$ mit
$$T(\eta) = \varphi_0(\cdot,\xi,\eta) \qquad (\eta \in \mathbb{K}^d)$$
ist linear und injektiv. (Hieraus folgt, dass $L_0 = T(\mathbb{K}^d)$ ein d-dimensionaler Unterraum von $C(I,\mathbb{K}^d)$ ist.)
Es gilt für $\eta_1, \eta_2 \in \mathbb{K}^d$ und $\lambda_1, \lambda_2 \in \mathbb{K}$ mit $\varphi_1 := \varphi_0(\cdot,\xi,\eta_1)$ und $\varphi_2 := \varphi_0(\cdot,\xi,\eta_2)$
$$\begin{aligned}(\lambda_1\varphi_1 + \lambda_2\varphi_2)' &= \lambda_1\varphi_1' + \lambda_2\varphi_2' = \\ &= \lambda_1 A(t)\varphi_1 + \lambda_2 A(t)\varphi_2 = \\ &= A(t)[\lambda_1\varphi_1 + \lambda_2\varphi_2]\end{aligned}$$

auf I, sowie

$$(\lambda_1\varphi_1 + \lambda_2\varphi_2)(\xi) = \lambda_1\varphi_0(\xi, \xi, \eta_1) + \lambda_2\varphi_0(\xi, \xi, \eta_2) = \lambda_1\eta_1 + \lambda_2\eta_2 \, .$$

Also ist (Eindeutigkeit der Lösung von (24.3))

$$\lambda_1\varphi_0(\cdot, \xi, \eta_1) + \lambda_2\varphi_0(\cdot, \xi, \eta_2) = \varphi_0(\cdot, \xi, \lambda_1\eta_1 + \lambda_2\eta_2) \, ,$$

mit anderen Worten

$$\lambda_1 T(\eta_1) + \lambda_2 T(\eta_2) = T(\lambda_1\eta_1 + \lambda_2\eta_2) \, .$$

Folglich ist T linear.
Ist $T(\eta_1) = T(\eta_2)$, so gilt $\varphi(\cdot, \xi, \eta_1) = \varphi(\cdot, \xi, \eta_2)$, also insbesondere

$$\eta_1 = \varphi(\xi, \xi, \eta_1) = \varphi(\xi, \xi, \eta_2) = \eta_2 \, .$$

Also ist T injektiv.
2. „ $\supset$ “ Es sei $\psi \in \varphi + L_0$, d.h. $\psi = \varphi + \varphi_0(\cdot, \xi, \eta)$ für ein $(\xi, \eta) \in I \times \mathbb{K}^d$. Dann gilt

$$\begin{aligned} \psi' &= \varphi' + \varphi_0'(\cdot, \xi, \eta) = A(t)\varphi' + b(t) + A(t)\varphi_0(\cdot, \xi, \eta) \\ &= A(t)\psi + b(t) \end{aligned}$$

auf I, d.h. ψ löst (24.1). Damit ist $\psi \in L_b$.
„ $\subset$ “ Es sei $\psi \in L_b$, d.h. $\psi' = A(t)\psi + b(t)$ auf I. Dann gilt $(\psi - \varphi)' = A(t)(\psi - \varphi)$, d.h. $\psi - \varphi \in L_0$. Also ist $\psi = \varphi + (\psi - \varphi) \in \varphi + L_0$. □

Bemerkung und Definition 24.3 Da nach S. 24.2 der Lösungsraum L_0 der homogenen Gleichung $y' = A(t)y$ ein d-dimensionaler linearer Raum ist, reicht es, zur Bestimmung einer beliebigen Lösung eine Basis von L_0 zu kennen (jede Lösung ist dann Linearkombination der Basiselemente). Eine solche Basis heißt *Fundamentalsystem*.
Außerdem zeigt der zweite Teil des Satzes, dass sich die Bestimmung einer beliebigen Lösung des inhomogenen Systems $y' = A(t)y + b(t)$ auf die Bestimmung einer speziellen Lösung des inhomogenen Systems und einer Basis des Lösungsraumes L_0 der zugehörigen inhomogenen Gleichung reduziert.

Wir werden uns zunächst mit homogenen Gleichungen befassen. Der erste Satz zeigt, dass die lineare Unabhängigkeit von Lösungen äquivalent ist zur linearen Unabhängigkeit der Anfangswerte.

Satz 24.4 *Es seien $\psi^{(1)}, ..., \psi^{(m)} \in L_0$, also Lösungen des homogenen Systems $y' = A(t)y$. Dann sind äquivalent:*

a) $\psi^{(1)}, ..., \psi^{(m)}$ *sind linear unabhängig.*

b) Für alle $\xi \in I$ *sind* $\psi^{(1)}(\xi), ..., \psi^{(m)}(\xi)$ *linear unabhängig.*

c) Es existiert ein $t_0 \in I$ *so, dass* $\psi^{(1)}(t_0), ..., \psi^{(m)}(t_0)$ *linear unabhängig sind.*

Beweis. Der Beweis ergibt sich unmittelbar aus der Beziehung

$$\psi^{(j)} = \varphi_0(\cdot, \xi, \psi^{(j)}(\xi)),$$

die für alle $\xi \in I$ gilt (auf beiden Seiten steht die Lösung von $y' = A(t)y$, $y(\xi) = \psi^{(j)}(\xi)$ auf I):
Gilt a), so sind, da nach S. 24.2 $\eta \mapsto \varphi_0(\cdot, \xi, \eta)$ für alle $\xi \in I$ ein Isomorphismus ist, auch $\eta^{(1)} = \psi^{(1)}(\xi), ..., \eta^{(m)} = \psi^{(m)}(\xi)$ linear unabhängig, d.h. b) gilt.
b) $\Rightarrow$ c) ist klar.
Gilt schließlich c), so folgt wiederum a) aus der Tatsache, dass $\eta \mapsto \varphi_0(\cdot, t_0, \eta)$ ein Isomorphismus von $\mathbb{K}^d$ auf L_0 ist. □

Beispiel 24.5 (vgl. B. 23.10.3) Wir betrachten $I = (0, \infty)$ und

$$A(t) := \begin{pmatrix} 0 & -1/t^2 \\ 1/t^2 & 0 \end{pmatrix} \qquad (t \in (0, \infty)).$$

Dann ist nach B. 23.10.3

$$\psi^{(1)}(t) = \begin{pmatrix} \sin(1/t) \\ \cos(1/t) \end{pmatrix} \qquad (t \in (0, \infty))$$

eine Lösung des homogenen Systems

$$y' = \begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & -1/t^2 \\ 1/t^2 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = A(t)\, y$$

(genauer gilt $\psi^{(1)}(t) = \varphi_0(t, 1/\pi, \binom{0}{-1})$ für $t > 0$). Man rechnet leicht nach, dass auch

$$\psi^{(2)}(t) = \begin{pmatrix} -\cos(1/t) \\ \sin(1/t) \end{pmatrix} \qquad (t > 0)$$

eine Lösung von $y' = A(t)y$ ist (genauer $\psi^{(2)}(t) = \varphi_0(t, 1/\pi, \binom{1}{0})$).
Da $\psi^{(1)}(1/\pi) = \binom{0}{-1}$ und $\psi^{(2)}(1/\pi) = \binom{1}{0}$ in $\mathbb{R}^2$ linear unabhängig sind, sind auch $\psi^{(1)}, \psi^{(2)}$ linear unabhängig, also eine Basis des Lösungsraumes L_0. Folglich ist jede Lösung von $y' = A(t)y$ von der Form

$$\varphi = \lambda_1 \psi^{(1)} + \lambda_2 \psi^{(2)}$$

für gewisse $\lambda_1, \lambda_2 \in \mathbb{R}$.

Bemerkung 24.6 1. Sind $\psi^{(1)}, ..., \psi^{(d)}$ beliebige Lösungen von $y' = A(t)y$ auf I, so bezeichnet man die Determinante

$$W(t) := W(\psi^{(1)}, ..., \psi^{(d)}; t) := \begin{vmatrix} \psi_1^{(1)}(t) & \dots & \psi_1^{(d)}(t) \\ \vdots & & \vdots \\ \psi_d^{(1)}(t) & \dots & \psi_d^{(d)}(t) \end{vmatrix} \qquad (t \in I)$$

als *Wronski-Determinante* von $\psi^{(1)}, ..., \psi^{(d)}$. Es gilt nach S. 24.4: $\psi^{(1)}, ..., \psi^{(d)}$ ist ein Fundamentalsystem (also eine Basis des Lösungsraumes) genau dann, wenn $W(t_0) \neq 0$ für ein $t_0 \in I$ ist. Außerdem ist in diesem Falle schon $W(t) \neq 0$ für alle $t \in I$! (Also: Entweder ist $W(t) \neq 0$ für alle $t \in I$ oder $W(t) \equiv 0$ auf I.)
2. Bilden $\psi^{(1)}, ..., \psi^{(d)}$ ein Fundamentalsystem, so heißt die Matrix

$$\Phi(t) := \begin{pmatrix} \psi_1^{(1)}(t) & \dots & \psi_1^{(d)}(t) \\ \vdots & & \vdots \\ \psi_d^{(1)}(t) & \dots & \psi_d^{(d)}(t) \end{pmatrix} \qquad (t \in I)$$

eine *Fundamentalmatrix*. Nach 1. ist $\Phi(t)$ für alle $t \in I$ invertierbar ($W(t) = \det \Phi(t) \neq 0$). Es gilt damit

$$\varphi_0(t, \xi, \eta) = \Phi(t) \cdot \Phi^{-1}(\xi) \cdot \eta \qquad (t \in I) \tag{24.4}$$

d.h. die allgemeine Lösung ergibt sich als Produkt der matrixwertigen Funktion Φ mit dem Vektor $\Phi^{-1}(\xi)\eta$.
(Denn: Für jedes (ξ, η) ist

$$\psi = \Phi(\cdot)\Phi^{-1}(\xi) \cdot \eta$$

eine Linearkombination der Funktionen $\psi^{(1)}, ..., \psi^{(d)}$, also eine Lösung von $y' = A(t)y$. Außerdem gilt

$$\psi(\xi) = \Phi(\xi)\Phi^{-1}(\xi)\eta = \eta\,,$$

also ist auch die Anfangsbedingung erfüllt. Folglich ist $\psi = \varphi_0(\cdot, \xi, \eta)$.)

Beispiel 24.7 Es sei A wie in B. 24.5. Dann ist

$$\Phi(t) = \begin{pmatrix} \sin(1/t) & -\cos(1/t) \\ \cos(1/t) & \sin(1/t) \end{pmatrix} = \begin{pmatrix} \psi_1^{(1)} & \psi_1^{(2)} \\ \psi_2^{(1)} & \psi_2^{(2)} \end{pmatrix}$$

eine Fundamentalmatrix. Speziell gilt

$$\Phi\Big(\frac{1}{\pi}\Big) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

und damit

$$\Phi^{-1}\Big(\frac{1}{\pi}\Big) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Also ist die Lösung von $y' = A(t)y$, $y(1/\pi) = \eta$ gegeben durch

$$\varphi_0\Big(t, \frac{1}{\pi}, \eta\Big) = \Phi(t)\begin{pmatrix} -\eta_2 \\ \eta_1 \end{pmatrix} = = \begin{pmatrix} -\eta_2 \sin(1/t) - \eta_1 \cos(1/t) \\ -\eta_2 \cos(1/t) + \eta_1 \sin(1/t) \end{pmatrix}.$$

Wir kommen nun zurück zum inhomogenen System (24.1). Es gilt hierfür

Satz 24.8 *(Variation der Konstanten)*
Es sei $\Phi(\cdot)$ eine Fundamentalmatrix von (24.2). Dann gilt für $\xi \in I$: Die Funktion

$$t \mapsto \Phi(t) \int_\xi^t \Phi^{-1}(s) b(s) ds$$

ist eine spezielle Lösung von (24.1), nämlich $\varphi_b(t, \xi, 0)$, und es gilt

$$\begin{aligned} \varphi_b(t, \xi, \eta) &= \Phi(t)\Big[\Phi^{-1}(\xi)\eta + \int_\xi^t \Phi^{-1}(s) b(s) ds\Big] \qquad (t \in I) \qquad (24.5) \\ (&= \varphi_0(t, \xi, \eta) + \varphi_b(t, \xi, 0)\,). \end{aligned}$$

Beweis. Wir schreiben für eine Funktion $C : I \to \mathbb{K}^{d\times d}$ mit $C = (c_{jk})_{j=1,\dots,d}$ und differenzierbaren Funktionen $c_{jk} : I \to \mathbb{K}$ kurz

$$C'(t) := \begin{pmatrix} c'_{11}(t) & \dots & c'_{1d}(t) \\ \vdots & & \\ c'_{d1}(t) & \dots & c'_{dd}(t) \end{pmatrix} \qquad (t \in I).$$

Dann gilt folgende Produktregel für differenzierbares $g : I \to \mathbb{K}^d$

$$(Cg)'(t) = C'(t)g(t) + C(t)g'(t) \qquad (t \in I).$$

(Denn: für $t \in I$ ist

$$\begin{aligned} (Cg)'_j(t) &= \Big(\sum_{k=1}^d c_{jk}(t) g_k(t)\Big)' = \sum_{k=1}^d c'_{jk}(t) g_k(t) \\ &+ \sum_{k=1}^d c_{jk}(t) g'_k(t) = (C'g)_j(t) + (Cg')_j(t)\Big). \end{aligned}$$

Folglich gilt (HDI komponentenweise angewandt)

$$\begin{aligned} \frac{d}{dt}\Big(\Phi(t) \int_\xi^t \Phi^{-1}(s) b(s) ds\Big) &= \Phi'(t) \int_\xi^t \Phi^{-1}(s) b(s) ds + \Phi(t)\Phi^{-1}(t) b(t) \\ &\overset{[\ddot{U}]}{=} A(t)\Phi(t) \int_\xi^t \Phi^{-1}(s) b(s) ds + b(t) \end{aligned}$$

d.h.

$$t \mapsto \Phi(t) \int_{\xi}^{t} \Phi^{-1}(s) b(s) ds$$

ist Lösung von (24.1) mit Funktionswert 0 an $t = \xi$. Weiter ist $L_b = \varphi_b(\cdot, \xi, 0) + L_0$ nach S. 24.2, und nach (24.4) ist $L_0 = \{\Phi(\cdot)\Phi^{-1}(\xi)\eta : (\xi, \eta) \in I \times \mathbb{K}^n\}$. Da

$$\Phi(\xi)\Big[\Phi^{-1}(\xi)\eta + \int_{\xi}^{t} \Phi^{-1}(s) b(s) ds\Big] = \eta$$

gilt, folgt die Behauptung. □

Beispiel 24.9 Wir betrachten wieder A aus B. 24.7. Ferner sei

$$b(t) = \begin{pmatrix} 1/t^2 \\ 0 \end{pmatrix} \qquad (t > 0)\,.$$

Dann ist für $\xi = 1/\pi$ mit Φ aus B. 24.7 nach S. 24.8

$$\varphi_b\Big(t, \frac{1}{\pi}, 0\Big) = \Phi(t) \int_{1/\pi}^{t} \Phi^{-1}(s) b(s) ds\,.$$

Weiter gilt

$$\Phi^{-1}(t) = \begin{pmatrix} \sin(1/t) & \cos(1/t) \\ -\cos(1/t) & \sin(1/t) \end{pmatrix},$$

also

$$\int_{1/\pi}^{t} \Phi^{-1}(s) b(s) ds = \int_{1/\pi}^{t} \begin{pmatrix} \sin(1/s)/s^2 \\ -\cos(1/s)/s^2 \end{pmatrix} ds \overset{1/s=u}{=} \begin{pmatrix} \cos(1/t) + 1 \\ \sin(1/t) \end{pmatrix}$$

und damit ist

$$\begin{aligned} \Phi(t) \int_{1/\pi}^{t} \Phi^{-1}(s) b(s) ds &= \begin{pmatrix} \sin(1/t) & -\cos(1/t) \\ \cos(1/t) & \sin(1/t) \end{pmatrix} \begin{pmatrix} \cos(1/t) + 1 \\ \sin(1/t) \end{pmatrix} \\ &= \begin{pmatrix} \sin(1/t) \\ 1 + \cos(1/t) \end{pmatrix} \end{aligned}$$

eine spezielle Lösung der inhomogenen Gleichung. Insgesamt ist nach S. 24.8

$$\varphi_b\Big(t, \frac{1}{\pi}, \eta\Big) = \varphi_0\Big(t, \frac{1}{\pi}, \eta\Big) + \begin{pmatrix} \sin(1/t) \\ 1 + \cos(1/t) \end{pmatrix}$$

mit $\varphi_0(\cdot, 1/\pi, \eta)$ wie in B. 24.7.

Wir wollen nun die obigen Ergebnisse auf lineare Differenzialgleichungen n-ter Ordnung anwenden: Sind $a_0, a_1, ..., a_{n-1} : I \to \mathbb{K}$ und $b : I \to \mathbb{K}$ stetig, so heißt eine Gleichung der Form

$$y^{(n)} = a_{n-1}(t)y^{(n-1)} + \ldots + a_0(t)y + b(t) = \sum_{\nu=0}^{n-1} a_j(t)y^{(\nu)} + b(t) \tag{24.6}$$

eine *lineare Differenzialgleichung n-ter Ordnung*. Die Gleichung

$$y^{(n)} = \sum_{\nu=0}^{n-1} a_j(t)y^{(\nu)} \tag{24.7}$$

heißt *zugehörige homogene* Gleichung. Entsprechende Anfangswertprobleme sind von der Form

$$y^{(n)} = \sum_{\nu=0}^{n-1} a_\nu(t)y^{(\nu)} + b(t), \qquad y^{(\nu)}(\xi) = \eta_\nu \quad (\nu = 0, ..., n-1) \tag{24.8}$$

mit $\xi \in I, \eta_0, ..., \eta_{n-1} \in \mathbb{K}$.
In B. 22.10 haten wir gesehen, dass man solche Gleichungen bzw. Anfangswertprobleme in Systeme 1. Ordnung umschreiben kann. Hier lautet das entsprechende System

$$\begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_{n-1}' \\ y_n' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \ldots & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & & & & & \vdots \\ 0 & \ldots & \ldots & \ldots & 0 & 1 \\ a_0(t) & a_1(t) & \ldots & \ldots & \ldots & a_{n-1}(t) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ b(t) \end{pmatrix}, \tag{24.9}$$

also ein lineares System. Nach B. 22.10 lassen sich sämliche Ergebnisse über Lösungen dieses Systems in Ergebnisse über die Lösungen von (24.6) bzw. (24.8) übertragen. Insbesondere erhalten wir aus S. 24.1

Satz 24.10 *Ist $I \subset \mathbb{R}$ ein offenes Intervall und sind $a_0, ..., a_{n-1}, b : I \to \mathbb{K}$ stetig, so hat für jedes $(\xi, \eta) \in I \times \mathbb{K}^d$ das AWP (24.8) (mit $\eta = (\eta_0, ..., \eta_{n-1})$) genau eine Lösung φ auf I. Wir schreiben dafür*

$$u_b(\cdot; \xi; \eta_0, ..., \eta_{n-1}) = u_b(\cdot; \xi; \eta)\,.$$

Um eine S. 24.2 entsprechende Aussage über die Lösungsgesamtheit machen zu können, unterscheiden wir auch wieder

$$M_0 := \{u_0(\cdot, \xi, \eta) : (\xi, \eta) \in I \times \mathbb{K}^n\}$$

und

$$M_b := \{u_b(\cdot,\xi,\eta) : (\xi,\eta) \in I \times \mathbb{K}^n\}.$$

Die Abbildung $j : L_b \to M_b$ mit

$$j(\varphi_1,...,\varphi_n) := \varphi_1 \qquad ((\varphi_1,...,\varphi_n) \in \mathrm{L}_b),$$

wobei L_b die Lösungsmenge von (24.9) ist, ist nach B. 22.10 bijektiv und im Falle $b = 0$ auch linear. Damit lassen sich auch die weiteren Ergebnisse über lineare Systeme übertragen. Wir fassen die wesentlichen im folgenden Satz zusammen:

Satz 24.11 *1. M_0 ist ein n-dimensionaler Unterraum (von $C(I,\mathbb{K})$).*

2. $M_b = \varphi + M_0$ für jedes $\varphi \in M_b$.

3. Sind $v_1,...,v_n$ Lösungen der homogenen Gleichung (24.7), d.h. $v_1,...,v_n \in M_0$, so sind $v_1,...,v_n$ linear unabhängig (ein „Fundamentalsystem"), genau dann, wenn die „Wronski-Determinante" $W(t) = \det \Phi(t)$, wobei

$$\Phi(t) = \begin{pmatrix} v_1(t) & \dots & \dots & v_n(t) \\ v_1'(t) & \dots & \dots & v_n'(t) \\ \vdots & & & \vdots \\ v_1^{(n-1)}(t) & \dots & \dots & v_n^{(n-1)}(t) \end{pmatrix} \in \mathbb{K}^{n\times n},$$

für ein $t \in I$ nicht verschwindet. In diesem Fall ist schon $W(t) \neq 0$ für alle $t \in I$.

4. (Variation der Konstanten) Ist $v_1,...,v_n$ ein Fundamentalsystem der homogenen Gleichung 24.7, so ist

$$t \mapsto \big(v_1(t),...,v_n(t)\big) \cdot \int_\xi^t \Phi^{-1}(s) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(s) \end{pmatrix} ds \qquad (t \in I)$$

eine Lösung der inhomogenen Gleichung (24.6), nämlich $u_b(t;\xi;0)$, und für $(\xi,\eta) \in I \times \mathbb{K}^n, \eta = (\eta_0,...,\eta_{n-1})^T$ gilt

$$\begin{aligned} u_b(t;\xi;\eta) &= (v_1(t),...,v_n(t)) \left[\Phi^{-1}(\xi) \begin{pmatrix} \eta_0 \\ \vdots \\ \eta_{n-1} \end{pmatrix} + \int_\xi^t \Phi^{-1}(s) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(s) \end{pmatrix} ds \right] \\ (&= u_0(t;\xi;\eta) + u_b(t;\xi;0)\ ) \end{aligned} \tag{24.10}$$

Beweis. Die Ergebnisse ergeben sich alle aus den entsprechenden Ergebnissen für das lineare System (24.9) durch Anwendung von $j : L_b \to M_b$ bzw. $j^{-1} : M_b \to L_b$, gegeben durch

$$j^{-1}(u) = (u, u', ..., u^{(n-1)}) \qquad (u \in M_b)\,.$$

Man beachte dabei insbesondere, dass die rechte Seite in (24.10) die erste Komponente von (24.5) für das entsprechende System ist. $\square$

Bemerkung 24.12 Will man (24.10) anwenden, so hat man insbesondere $\Phi^{-1}(s)\begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(s) \end{pmatrix}$ zu bestimmen, d.h. man hat das lineare Gleichungssystem

$$\Phi(s)\begin{pmatrix} c_1(s) \\ \vdots \\ c_n(s) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(s) \end{pmatrix}$$

zu lösen. Auf Grund der speziellen Struktur der rechten Seite ist dies mit Hilfe der Cramerschen Regel relativ einfach durchzuführen. Man erhält hier

$$\begin{aligned} c_j(s) &= \frac{1}{\det \Phi(s)} \det \begin{pmatrix} v_1 & \dots & v_{j-1} & 0 & v_{j+1} & \dots & v_n \\ v_1' & & v_{j-1}' & 0 & v_{j+1}' & & v_n' \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ v_1^{(n-1)} & \dots & v_{j-1}^{(n-1)} & b(s) & v_{j+1}^{(n-1)} & \dots & v_n^{(n-1)} \end{pmatrix} \\ &= \frac{1}{W(s)}(-1)^{n+j} b(s) W_j(s)\,, \end{aligned}$$

wobei

$$W_j(s) = \det \begin{pmatrix} v_1 & \dots & v_{j-1} & v_{j+1} & \dots & v_n \\ v_1' & & v_{j-1}' & v_{j+1}' & & v_n' \\ \vdots & & \vdots & \vdots & & \vdots \\ v_1^{(n-2)} & \dots & v_{j-1}^{(n-2)} & v_{j+1}^{(n-2)} & \dots & v_n^{(n-2)} \end{pmatrix}$$

ist. Also erhalten wir

$$u_b(t;\xi;0) = \sum_{j=1}^{n} v_j(t)(-1)^{n+j} \int_{\xi}^{t} \frac{b(s)W_j(s)}{W(s)} ds \qquad (t \in I)\,.$$

Diese Darstellung erklärt den Namen „Variation der Konstanten“: Während jede Lösung der homogenen Gleichung von der Form $\sum_{j=1}^{n} \lambda_j v_j(t)$ ist (also Linearkombination der $v_1, ..., v_n$), ist hier

$$u_b(t;\xi;0) = \sum_{j=1}^{n} \lambda_j(t) v_j(t)\,,$$

also Linearkombination der $v_1(t), ..., v_n(t)$ mit bezüglich t variierenden Koeffizienten $\lambda_1(t), ..., \lambda_n(t)$.

Beispiel 24.13 Wir betrachten die lineare Differenzialgleichung 2. Ordnung

$$y'' = y + e^t$$

und suchen die Lösung mit der Anfangsbedingung $y(0) = 0$, $y'(0) = 1$. Man rechnet sofort nach, dass

$$v_1(t) = e^t\,, \qquad v_2(t) = e^{-t} \qquad\qquad (t \in \mathbb{R})$$

ein Fundamentalsystem der homogenen Gleichung $y'' = y$ ist. Also erhalten wir nach B. 24.12 mit

$$\begin{aligned}
&b(t) = e^t\,,\\
&W(t) = \det\begin{pmatrix} e^t & e^{-t} \\ e^t & -e^{-t} \end{pmatrix} = -2\\
&W_1(t) = \det(e^{-t}) = e^{-t}, \qquad W_2(t) = \det(e^t) = e^t
\end{aligned}$$

die spezielle Lösung

$$u_b(t;0;0,0) = -v_1(t)\int_0^t \frac{b(s)W_1(s)}{W(s)}ds + v_2(t)\int_0^t \frac{b(s)W_2(s)}{W(s)}ds = \frac{t}{2}e^t - \frac{1}{4}e^t + \frac{1}{4}e^{-t}\,.$$

Weiter ergibt sich mit (24.10), da

$$\Phi(t) = \begin{pmatrix} e^t & e^{-t} \\ e^t & -e^{-t} \end{pmatrix} \qquad \text{und} \qquad \Phi^{-1}(0)\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}\frac{1}{2}\\ -\frac{1}{2}\end{pmatrix}$$

damit

$$\begin{aligned}
u_b(t;0;0,1) &= \big(v_1(t), v_2(t)\big)\Phi^{-1}(0)\begin{pmatrix}0\\1\end{pmatrix} + u_b(t;0;0,0)\\
&= \frac{1}{2}e^t - \frac{1}{2}e^{-t} + \frac{t}{2}e^t - \frac{1}{4}e^t + \frac{1}{4}e^{-t} = \frac{1}{4}e^t - \frac{1}{4}e^{-t} + \frac{t}{2}e^t
\end{aligned}$$

Also: Die obigen Ergebnisse zeigen, dass wir zur Lösung linearer Systeme oder linearer Differenzialgleichungen n-ter Ordnung Fundamentalsysteme finden müssen. Leider ist dies ein i A. sehr schwieriges Unterfangen, und es gibt keineswegs eine geschlossene

Theorie dazu. Wir werden uns im nächsten Abschnitt intensiv mit dem Fall linearer Gleichungen mit konstanten Koeffzienten (d.h. $A(t) \equiv A$ auf $\mathbb{R}$ bzw. $a_j(t) \equiv a_j$ auf $\mathbb{R}$ für $j = 1, ..., n$) beschäftigen. Zum Abschluss dieses Abschnittes stellen wir noch kurz zwei mögliche Ansätze für allgemeine lineare Gleichungen vor: Kennt man (woher auch immer) bereits eine nichtverschwindende Lösung einer linearen Differenzialgleichung der Ordnung n, so lässt sich dies nutzen, um weitere Lösungen aus einer linearen Differenzialgleichung $(n-1)$-ter Ordnung zu bestimmen („Reduktion der Ordnung“). Wir beschränken uns bei der Darstellung dieses Verfahrens auf den Fall $n = 2$.

Satz 24.14 *Es sei $I \subset \mathbb{R}$ ein offenes Intervall, und es seien $a_0, a_1 : I \to \mathbb{K}$ stetig. Ist (u, I) eine Lösung der Differenzialgleichung*

$$y'' = a_1(t)y' + a_0(t)y$$

mit $u(t) \neq 0$ für alle $t \in J$, wobei $J \subset I$ ein offenes Intervall ist, so erhält man eine zweite, von u linear unabhängige Lösung $v : J \to \mathbb{K}$ durch den Ansatz

$$v(t) = z(t)u(t) \qquad (t \in J)\,,$$

wobei $z' = w \not\equiv 0$ eine Lösung der linearen Differenzialgleichung 1. Ordnung

$$w' = \Big(a_1(t) - 2\frac{u'(t)}{u(t)}\Big)w$$

ist.

Beweis. Aus $v = zu$ folgt

$$v' = z'u + zu'\,, \qquad v'' = z''u + 2z'u' + u''z\,,$$

und mit $u'' = a_1u' + a_0u$ erhalten wir, falls $z'' = (a_1 - 2u'/u)z'$ gilt,

$$v'' - a_1v' - a_0v = z''u + 2z'u' + u''z - a_1z'u - a_1zu' - a_0zu = 0\,,$$

d.h. v ist ebenfalls Lösung (auf J).
Ist $z' \not\equiv 0$, so ist z nicht konstant auf J, und damit sind u und v linear unabhängig. □

Beispiel 24.15 Wir betrachten die lineare Differenzialgleichung 2. Ordnung

$$y'' = \frac{2t}{1-t^2}y' - \frac{2}{1-t^2}y$$

auf $I = (-1, 1)$. Man sieht sofort, dass

$$u(t) = t$$

eine Lösung auf $(-1,1)$ ist. Also erhalten wir nach S. 24.14 eine zweite, linear unabhängige Lösung v, etwa auf $(0,1)$ durch

$$v = zu\,,$$

wobei $w = z'$ Lösung von

$$w' = \Big(a_1 - 2\frac{u'}{u}\Big)w = \Big(\frac{2t}{1-t^2} - \frac{2}{t}\Big)w$$

ist. Nach S. 22.6 ist mit

$$\begin{aligned} A(t) &= \int \frac{2tdt}{1-t^2} - \int \frac{2dt}{t} = \ln\Big(\frac{1}{1-t^2}\Big) - 2\ln(t) \quad [+c] \\ &= \ln\Big(\frac{1}{1-t^2}\Big) + \ln\Big(\frac{1}{t^2}\Big) \quad [+c] \end{aligned}$$

die Funktion

$$w(t) = e^{A(t)} = \frac{1}{(1-t^2)t^2}$$

Lösung, also

$$z(t) = \int \frac{dt}{(1-t^2)t^2} = \int \frac{1}{t^2} + \frac{1}{2}\Big(\frac{1}{1-t} + \frac{1}{1+t}\Big)dt = -\frac{1}{t} + \frac{1}{2}\ln\Big(\frac{1+t}{1-t}\Big)$$

und damit

$$v(t) = tz(t) = \frac{t}{2}\ln\Big(\frac{1+t}{1-t}\Big) - 1\,.$$

Folglich bilden (u,v) ein Fundamentalsystem der Ausgangsgleichung (zunächst auf $(0,1)$, aber tatsächlich auch auf $(-1,1)$).

Im manchen Fällen ist es möglich, Lösungen einer Differenzialgleichung durch einen sogenannten Potenzreihenansatz zu gewinnen. Wir erläutern die zu Grunde liegende Idee auch nur für den Fall linearer Differenzialgleichungen 2. Ordnung.

Satz 24.16 *Es sei $I = (-r,r)$ für ein $r > 0$, und es seien $p,q : I \to \mathbb{K}$ gegeben durch*

$$p(t) = \sum_{\nu=0}^{\infty} p_\nu t^\nu\,, \qquad q(t) = \sum_{\nu=0}^{\infty} q_\nu t^\nu\,.$$

Ferner sei $p_0 \notin \mathbb{N}$. Sind $\eta_0, \eta_1 \in \mathbb{K}$ beliebig, so definieren wir die Folge $(a_\nu)_{\nu\in\mathbb{N}_0}$ rekursiv durch

$$\begin{array}{lll} a_0 = \eta_0\,, & a_1 = \eta_1\,, & \textit{falls} \quad p_0 = q_0 = 0 \\ a_0 = \eta_0\,, & a_1 = -\frac{q_0}{p_0}a_0\,, & \textit{falls} \quad p_0 \neq 0 \\ a_0 = 0\,, & a_1 = \eta_1\,, & \textit{falls} \quad p_0 = 0, q_0 \neq 0 \end{array}$$

sowie

$$a_{\nu+1} := \frac{1}{(\nu+1)(\nu-p_0)} \sum_{\mu=0}^{\nu} (\mu\, p_{\nu+1-\mu} + q_{\nu-\mu})a_\mu \qquad (\nu \in \mathbb{N}).$$

Dann gilt: Ist die Potenzreihe $\sum\limits_{\nu=0}^{\infty} a_\nu t^\nu$ *konvergent auf* I*, so ist*

$$\varphi(t) = \sum_{\nu=0}^{\infty} a_\nu t^\nu$$

eine Lösung der Differenzialgleichung

$$y'' = \frac{p(t)}{t} y' + \frac{q(t)}{t} y$$

auf $I \cap (0, \infty)$ *und auf* $I \cap (-\infty, 0)$*.*

Beweis. Nach dem Beweis zu S. 16.6 können wir φ gliedweise differenzieren (beachte: $I \subset$ Konvergenzkreis von φ) und erhalten

$$\begin{aligned} \varphi'(t) &= \sum_{\nu=0}^{\infty} (\nu+1) a_{\nu+1} t^\nu \\ \varphi''(t) &= \sum_{\nu=0}^{\infty} (\nu+1)(\nu+2) a_{\nu+2} t^\nu \end{aligned} \qquad (t \in I)\,.$$

Also erhalten wir mit Hilfe des Cauchy-Produktes

$$\begin{aligned} & t\varphi''(t) - p(t)\varphi'(t) - q(t)\varphi(t) = \\ &= \sum_{\nu=0}^{\infty} (\nu+1)\nu\, a_{\nu+1} t^\nu - \left(\sum_{\nu=0}^{\infty} p_\nu t^\nu\right)\left(\sum_{\nu=0}^{\infty} (\nu+1) a_{\nu+1} t^\nu\right) - \left(\sum_{0}^{\infty} q_\nu t^\nu\right)\left(\sum_{0}^{\infty} a_\nu t^\nu\right) \\ &= \sum_{\nu=0}^{\infty} (\nu+1)\nu\, a_{\nu+1} t^\nu - \sum_{\nu=0}^{\infty} t^\nu \sum_{\mu=0}^{\nu} (\mu+1) a_{\mu+1} p_{\nu-\mu} - \sum_{\nu=0}^{\infty} t^\nu \sum_{\mu=0}^{\nu} q_{\nu-\mu} a_\mu \\ &= \sum_{\nu=0}^{\infty} t^\nu \underbrace{\left[(\nu+1)\nu\, a_{\nu+1} - (\nu+1) p_0 a_{\nu+1} - \sum_{\mu=0}^{\nu} (\mu\, p_{\nu+1-\mu} + q_{\nu-\mu}) a_\mu\right]}_{=: b_\nu} \end{aligned} .$$

Nach Voraussetzung gilt $b_\nu = 0$ $(\nu \in \mathbb{N})$, und für $\nu = 0$ erhalten wir

$$b_0 = -p_0 a_1 - q_0 a_0\,.$$

Die Bedingungen an a_0, a_1 sind gerade so eingerichtet, dass stets $b_0 = 0$ ist. Also löst φ die Differenzialgleichung. □

Bemerkung 24.17 1. Man kann beweisen, dass unter den Bedingungen von S. 24.16 gilt: Haben die Potenzreihen p und q Konvergenzradius $r > 0$, so hat auch die Potenzreihe φ (mindestens) Konvergenzradius r, d.h. $\sum_{\nu=0}^{\infty} \varphi_\nu t^\nu$ konvergiert stets auf I.

Wir verzichten auf den (mit den uns derzeit zur Verfügung stehenden Mitteln etwas aufwändigen) Beweis. In den Beispielen ist die Konvergenz meist leicht zu sehen.
2. Gilt $p_0 = q_0 = 0$, so folgt

$$\tilde{p}(t) = \frac{p(t)}{t} = \sum_{\nu=1}^{\infty} p_\nu t^{\nu-1}\,, \tilde{q}(t) = \frac{q(t)}{t} = \sum_{\nu=1}^{\infty} q_\nu t^{\nu-1}\,,$$

d.h. wir können die Differenzialgleichung

$$y'' = \tilde{p}(t)y' + \tilde{q}(t)y \quad (= \frac{p(t)}{t}y' + \frac{q(t)}{t}y \text{ für } t \neq 0)$$

auf ganz $I = (-r, r)$ betrachten. In diesem Fall erhalten wir, indem wir speziell etwa die beiden Vektoren

$$\begin{pmatrix} \eta_0 \\ \eta_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{und} \qquad \begin{pmatrix} \eta_0 \\ \eta_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

als Anfangsbedingungen wählen (man beachte: dann gilt $\varphi(0) = \eta_0, \varphi'(0) = \eta_1$), ein Fundamentalsystem für diese Gleichung.

Beispiel 24.18 1. Wir betrachten für festes $\lambda \in \mathbb{R}$ die Gleichung

$$y'' = 2ty' - \lambda y \qquad \text{auf} \quad I = \mathbb{R}$$

(Hermitesche Differenzialgleichung). Hier ist

$$p(t) = 2t^2 \qquad q(t) = -\lambda t\,,$$

d.h.

$$\begin{array}{ll} p_2 = 2\,, & p_\nu = 0 \quad \text{sonst} \\ q_1 = -\lambda\,, & q_\nu = 0 \quad \text{sonst} \end{array}\,.$$

Wir sind also in der Situation von B. 24.17.2 (d. h. $p_0 = q_0 = 0$).
Die Rekursionsformel aus S. 24.16 ergibt hier

$$a_{\nu+1} = \frac{2(\nu-1)-\lambda}{(\nu+1)\nu} a_{\nu-1} \qquad (\nu \in \mathbb{N})\,.$$

Wählen wir speziell $a_0 = \eta_0 = 1, a_1 = \eta_1 = 0$, so erhalten wir

$$u_0(t;0;1,0) = 1 - \frac{\lambda}{z!}t^2 - \frac{(4-\lambda)\lambda}{4!}t^4 - \frac{(8-\lambda)(4-\lambda)\lambda}{6!}t^6 \ldots \left(= \sum_{\mu=0}^{\infty} \frac{1}{(2\mu)!} \prod_{k=0}^{\mu-1} (4k-\lambda) \cdot t^{2\mu} \right)\,.$$

Mit $a_0 = \eta_0 = 0$ und $a_1 = \eta_1 = 1$ ergibt sich entsprechend

$$u_0(t;0;0,1) = t + \frac{2-\lambda}{3!}t^3 + \frac{(6-\lambda)(2-\lambda)}{5!}t^5 \ldots \left(= \sum_{\mu=0}^{\infty} \prod_{k=0}^{\mu-1} (4k+2-\lambda) \frac{t^{2\mu+1}}{(2\mu+1)!} \right)\,.$$

Beide Potenzreihen konvergieren auf $\mathbb{R}$ und bilden ein Fundamentalsystem der Gleichung.

2. Wir betrachten für feste $a, b \in \mathbb{C}, -b \notin \mathbb{N}_0$, die Gleichung

$$y'' = \frac{t-b}{t} y' + \frac{a}{t} y$$

(Kummersche Differenzialgleichung). Hier ist

$$p(t) = t - b, \qquad q(t) = a\,,$$

d.h.

$$\begin{array}{llll} p_0 = -b, & p_1 = 1, & p_\nu = 0 & \text{sonst} \\ q_0 = a, & q_\nu = 0 & \text{sonst.} & \end{array}$$

Die Rekursionsformel aus S. 24.16 ergibt hier $a_1 = aa_0/b$ und

$$a_{\nu+1} = -\frac{\nu + a}{(\nu+1)(\nu+b)} \cdot a_\nu \qquad (\nu \in \mathbb{N})\,.$$

Wählen wir etwa $\eta_0 = a_0 = 1$, so ergibt sich

$$\varphi(t) = 1 + \frac{a}{b} t + \frac{(1+a)a}{(1+b)b} \frac{t^2}{2} + \frac{(2+a)(1+a)a}{(2+b)(1+b)b} \frac{t^3}{3!} + \dots\,,$$

d.h.

$$\varphi(t) = \sum_{\nu=0}^{\infty} \frac{(a)_\nu}{(b)_\nu \nu!} t^\nu \qquad \text{mit} \qquad \begin{array}{l} (c)_\nu := c(c+1)\dots(c+\nu-1) \\ (c)_0 := 1 \end{array} .$$

Die Potenzreihe konvergiert auf $\mathbb{R}$. Die Funktion φ heißt Kummer-Funktion (mit Parametern a, b) oder auch konfluente hypergeometrische Funktion.

25 Lineare Differenzialgleichungen mit konstanten Koeffizienten

Nach den Überlegungen der letzten Abschnitte stellt sich weiter die Frage, wie man Fundamentalsysteme homogener Systeme bestimmen kann. Wir wollen diese Frage (jedenfalls im Prinzip) klären für Systeme der Gestalt

$$y' = Ay\,, \tag{25.1}$$

wobei $A \in \mathbb{K}^{d\times d}$ eine feste Matrix ist (also unabhängig von t). Eine solche Gleichung heißt *lineares System mit konstanten Koeffizienten*.
Eine sehr kompakte Darstellung für Fundamentalmatrizen erhält man durch Einführung der Matrix-Exponentialfunktion. Dazu werden wir zunächst allgemeiner das Konzept sogenannter Matrix-Potenzreihen kurz vorstellen:
Es sei $f(z) = \sum\limits_{\nu=0}^{\infty} c_\nu z^\nu$ eine Potenzreihe mit Konvergenzradius $r > 0$ (und o. E. mit Entwicklungsmitte 0). Ferner sei $(B_\nu)_{\nu\in\mathbb{N}_0}$ eine Folge in $\mathbb{K}^{d\times d}$ mit $\overline{\lim\limits_{\nu\to\infty}}||B_\nu||^{1/\nu} < r$. Ist $B_\nu = (b_{jk}^{(\nu)})_{j,k=1,\ldots,d}$, so definieren wir

$$\sum_{\nu=0}^{\infty} c_\nu B_\nu := \left(\sum_{\nu=0}^{\infty} c_\nu b_{jk}^{(\nu)}\right)_{j,k=1,\ldots,d} \in \mathbb{K}^{d\times d}\,.$$

(Man beachte dabei: die Reihen $\sum\limits_{\nu=0}^{\infty} c_\nu b_{jk}^{(\nu)}$ sind absolut konvergent in $\mathbb{K}$, denn nach Voraussetzung existiert ein $s < r$ so, dass $||B_\nu|| \le s^\nu$ für alle $\nu \ge \nu_0$. Also gilt auch

$$|b_{jk}^{(\nu)}| \le ||B_\nu|| \le s^\nu \qquad (\nu \ge \nu_0;\ j,k = 1,...,d)\,.$$

Da $\sum\limits_{\nu=0}^{\infty} |c_\nu| s^\nu$ konvergiert, konvergiert nach dem Weierstraßschen Majorantenkriterium auch $\sum\limits_{\nu=0}^{\infty} c_\nu b_{jk}^{(\nu)}$ absolut.)
Man kann dafür zeigen: Ist $S_n := \sum\limits_{\nu=0}^{n} c_\nu B_\nu$, so ist

$$\sum_{\nu=0}^{\infty} c_\nu B_\nu = \lim_{n\to\infty} S_n \quad \text{in} \quad (\mathbb{K}^{d\times d}, ||\cdot||)\,.$$

Außerdem gilt für $A \in \mathbb{K}^{d\times d}$ und $C \in \mathbb{K}^{d\times d}$

$$(*) \qquad A\left(\sum_{\nu=0}^{\infty} c_\nu B_\nu\right) = \sum_{\nu=0}^{\infty} c_\nu A B_\nu \qquad \text{und} \qquad \left(\sum_{\nu=0}^{\infty} c_\nu B_\nu\right) C = \sum_{\nu=0}^{\infty} c_\nu B_\nu C\,.$$

(Denn: Zunächst gilt $||AB_\nu||^{1/\nu} \le ||A||^{1/\nu}||B_\nu||^{1/\nu}$ und aus $||A||^{1/\nu} \to 1$ (bzw. 0 für $||A|| = 0$) folgt, dass auch $\sum\limits_{\nu=0}^{\infty} c_\nu A B_\nu$ existiert. Weiter gilt, falls $A = (a_{jk})$, für den

(j,k)-ten Eintrag von $A \cdot \sum\limits_{\nu=0}^{\infty} c_\nu B_\nu$

$$\sum_{\ell=1}^{d} a_{j\ell} \cdot \sum_{\nu=0}^{\infty} c_\nu b_{\ell k}^{(\nu)} = \sum_{\nu=0}^{\infty} c_\nu \sum_{\ell=1}^{d} a_{j\ell} b_{\ell k}^{(\nu)} ,$$

was auch der (j,k)-te Eintrag von $\sum\limits_{\nu=0}^{\infty} c_\nu A B_\nu$ ist. Eine entsprechende Überlegung zeigt die zweite Behauptung.)

Ist speziell $B_\nu = B^\nu$ für eine feste Matrix $B \in \mathbb{K}^{d\times d}$ mit $||B|| < r$ (wobei $B^0 := E_d$), so ist $||B_\nu|| = ||B^\nu|| \leq ||B||^\nu$, also existiert $\sum\limits_{\nu=0}^{\infty} c_\nu B^\nu$. Wir schreiben dann kurz

$$f(B) := \sum_{\nu=0}^{\infty} c_\nu B^\nu .$$

Mit dieser Notation gilt

Satz 25.1 *Es sei* $f(z) = \sum\limits_{\nu=0}^{\infty} c_\nu z^\nu$ *eine Potenzreihe mit Konvergenzradius* $r > 0$. *Ist* $A \in \mathbb{K}^{d\times d}$, *so definieren wir* $\psi : (-r/||A||, r/||A||) \to \mathbb{K}^{d\times d}$ *durch*

$$\psi(t) := f(tA) \qquad \big(|t| < \frac{r}{||A||}\big) .$$

Dann gilt: ψ *ist differenzierbar auf* $(-r/||A||, r/||A||)$ *und es ist*

$$\psi'(t) = A \cdot f'(tA) \qquad \big(|t| < \frac{r}{||A||}\big) .$$

Beweis. Zunächst beachte man, dass nach der Vorbemerkung $f(tA)$ existiert, da $||tA|| = |t| \cdot ||A|| < r$ gilt. Da auch die Potenzreihe

$$f'(z) = \sum_{\nu=1}^{\infty} \nu c_\nu z^{\nu-1} = \sum_{\nu=0}^{\infty} (\nu+1) c_{\nu+1} z^\nu$$

Konvergenzradius r hat, existiert auch $f'(tA)$ für alle t mit $|t| < r/||A||$. Außerdem ist (mit $A^\nu = (a_{jk}^{(\nu)})_{j,k=1,\ldots,d}$)

$$f(tA) = \left(\sum_{\nu=0}^{\infty} c_\nu a_{jk}^{(\nu)} t^\nu \right)_{j,k=1,\ldots,d} ,$$

d.h. alle Einträge der Matrix sind Potenzreihen in der Variablen t. Da diese gliedweise differenziert werden dürfen (vgl. Beweis zu S. 16.6), erhalten wir für $|t| < r/||A||$

$$\begin{aligned}
\psi'(t) &= \left(\Big(\sum_{\nu=0}^{\infty} c_\nu a_{jk}^{(\nu)} t^\nu\Big)'\right)_{j,k=1,\dots,d} = \\
&= \left(\sum_{\nu=0}^{\infty} (\nu+1) c_{\nu+1} a_{jk}^{(\nu+1)} t^\nu\right)_{j,k=1,\dots,d} = \\
&= \sum_{\nu=0}^{\infty} (\nu+1) c_{\nu+1} t^\nu A^{\nu+1} \overset{(*)}{=} A f'(tA)\,.
\end{aligned}$$

□

Betrachten wir speziell die Potenzreihe

$$e^z = \exp(z) = \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!}$$

mit Konvergenzradius $r = \infty$, so ergibt sich für $\psi(t) = \exp(tA)$ die fürs Weitere zentrale Folgerung

$$\psi'(t) = A \cdot \exp(tA) \qquad (t \in \mathbb{R})\,.$$

Für $\exp(B)$ schreiben wir auch wieder kurz e^B. Damit ergibt sich

$$\boxed{(e^{tA})' = Ae^{tA}.} \tag{25.2}$$

Die Eleganz dieser Betrachtungen belegt folgender

Satz 25.2 *Es sei $A \in \mathbb{K}^{d\times d}$. Dann ist für alle $\eta \in \mathbb{K}^d$ die (maximale) Lösung des AWP*

$$y' = Ay\,, \qquad y(0) = \eta$$

gegeben durch

$$\varphi(t) = \varphi(t, 0, \eta) = e^{tA} \cdot \eta \qquad (t \in \mathbb{R})\,.$$

Außerdem ist e^{tA} eine Fundamentalmatrix für (25.1).

Beweis. Mit der Produktregel für matrixwertige Funktionen aus dem Beweis zu S. 24.8 ergibt sich

$$\frac{d}{dt}(e^{tA} \cdot \eta) = (e^{tA})' \cdot \eta \overset{(25.2)}{=} Ae^{tA} \cdot \eta \qquad (t \in \mathbb{R})\,,$$

d.h. $t \mapsto e^{tA} \cdot \eta$ ist Lösung von (25.1). Außerdem ist

$$e^{tA} \cdot \eta|_{t=0} = e^{0A} \cdot \eta = E_d \eta = \eta\,.$$

Also ist $\varphi(t,0,\eta) = e^{tA}\cdot\eta$.
Wählt man speziell $\eta = e_k$, wobei e_k den k-ten Einheitsvektor bezeichnet, so ist $e^{tA}\cdot e_k$ die k-te Spalte von e^{tA}, d.h. jede Spalte ist Lösung von (25.1). Da $e^{0A} = E_d$ gilt, sind die Spalten (nach S. 24.4 mit $\xi = 0$) linear unabhängig. Folglich ist $e^{tA}e_1, ..., e^{tA}e_d$ ein Fundamentalsystem, d.h. e^{tA} eine Fundamentralmatrix. □

Im Prinzip haben wir also ein Fundamentalsystem für (25.1) gefunden (nämlich das System $e^{tA}e_1, ..., e^{tA}e_d$). Es stellt sich allerdings dabei die Frage, wie man e^{tA} „konkret" berechnen kann. Dazu versucht man, die Berechnung von e^{tA} für allgemeines A auf die Berechnung von $e^{t\tilde{A}}$ für gewisse einfache Matrizen $\tilde{A}$ zurück zu führen. Entsprechende Normalformen $\tilde{A}$ kennt man aus der linearen Algebra.
Um in diese Richtung weitergehen zu können, brauchen wir zunächst einige Rechenregeln für die Matrixexponentialfunktion.

Satz 25.3 *Es seien $A, B, C \in \mathbb{K}^{d\times d}$.*

1. *Gilt $AB = BA$, so folgt $e^Ae^B = e^{A+B}$.*
2. *Ist C invertierbar, so gilt*
$$e^{C^{-1}AC} = C^{-1}e^AC\,.$$
3. *Hat A Blockdiagonalform*
$$A = \begin{pmatrix} \boxed{A_1} & & \\ & \ddots & \\ & & \boxed{A_m} \end{pmatrix},$$
d.h. $A = \operatorname{diag}(A_1, ..., A_m)$, so gilt
$$e^A = \operatorname{diag}(e^{A_1}, ..., e^{A_m})\,.$$

Beweis. 1. Aus $AB = BA$ folgt
$$(A+B)^n = \sum_{\nu=0}^{n}\binom{n}{\nu}A^\nu B^{n-\nu} \qquad (n\in\mathbb{N}_0)\,.$$
Weiter erhält man für den (j,k)-ten Eintrag von e^Ae^B (Cauchy-Produkt)
$$\begin{aligned}\sum_{\ell=1}^{d}\left(\sum_{\nu=0}^{\infty}\frac{1}{\nu!}a_{j\ell}^{(\nu)}\right)\left(\sum_{\nu=0}^{\infty}\frac{1}{\nu!}b_{\ell k}^{(\nu)}\right) &= \\ = \sum_{\ell=1}^{d}\sum_{n=0}^{\infty}\sum_{\nu=0}^{n}\underbrace{\frac{1}{\nu!}\,\frac{1}{(n-\nu)!}}_{=\frac{1}{n!}\binom{n}{\nu}}\,a_{j\ell}^{(\nu)}\,b_{\ell k}^{(n-\nu)} &= \\ = \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{\nu=0}^{n}\binom{n}{\nu}\sum_{\ell=1}^{d}a_{j\ell}^{(\nu)}\,b_{\ell k}^{(n-\nu)}\,, &\end{aligned}$$

was auch der (j,k)-te Eintrag von e^{A+B} ist.
2. Es gilt

$$(C^{-1}AC)^{\nu} = C^{-1}A^{\nu}C \qquad (\nu \in \mathbb{N}_0)$$

(Beweis per Induktion). Also erhalten wir

$$C^{-1}e^{A}C = C^{-1}\left(\sum_{\nu=0}^{\infty}\frac{1}{\nu!}A^{\nu}\right)C \overset{(*)}{=} \sum_{\nu=0}^{\infty}\frac{1}{\nu!}C^{-1}A^{\nu}C = e^{C^{-1}AC}\,.$$

3. Ist

$$A = \begin{pmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_m} \end{pmatrix} \qquad (\nu \in \mathbb{N}_0)\,,$$

so ist

$$A^{\nu} = \begin{pmatrix} \boxed{A_1^{\nu}} & & & \\ & \boxed{A_2^{\nu}} & & \\ & & \ddots & \\ & & & \boxed{A_m^{\nu}} \end{pmatrix} \qquad (\nu \in \mathbb{N}_0)\,,$$

also ergibt sich 3. nach der (elementweisen) Definition von e^{A}. □

Bemerkung 25.4 Es sei $A \in \mathbb{K}^{d\times d}$ diagonalisierbar, d.h. es existieren eine Diagonalmatrix $\Lambda = \operatorname{diag}(\lambda_1, ..., \lambda_d)$ sowie eine invertierbare Matrix $C \in \mathbb{K}^{d\times d}$ mit

$$A = C\Lambda C^{-1}$$

(dabei sind $\lambda_1, ..., \lambda_d$ die Eigenwerte von A und die Spalten von C, also $Ce_1, ..., Ce_d$, bilden eine Basis aus Eigenvektoren von A).
Dann gilt mit S. 25.3

$$e^{tA} = Ce^{t\Lambda}C^{-1} = C\cdot\operatorname{diag}(e^{\lambda_1 t}, ..., e^{\lambda_d t})C^{-1}\,.$$

Die Spalten bilden ein Fundamentalsystem. Außerdem ist mit $c_k := Ce_k$ auch

$$e^{\lambda_1 t}c_1, \ldots, e^{\lambda_d t}c_d$$

ein Fundamentalsystem (folgt etwa aus $e^{\lambda_k t}c_k = e^{tA}c_k$ und S. 25.2).

Beispiel 25.5 1. Wir betrachten das lineare System

$$y' = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} y = Ay\,.$$

Es gilt

$$\det(A - \lambda E) = -\lambda^3 + 3\lambda - 2 = 0 \quad \Leftrightarrow \quad \lambda = 1 \text{ oder } \lambda = -2\,,$$

also haben wir die Eigenwerte 1 und -2. Weiter rechnet man nach, dass

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \qquad \text{und} \qquad \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

(linear unabhängige) Eigenvektoren zu $\lambda_1 = 1$ sind, und dass

$$\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

ein Eigenvektor zu $\lambda_2 = 2$ ist. Also gilt hier

$$\begin{aligned} e^{tA} &= \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} e^t & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^{-2t} \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}^{-1}}_{=\frac{1}{3}\begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ -1 & -1 & 1 \end{pmatrix}} \\ &= \frac{1}{3} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2e^t & -e^t & e^t \\ -e^t & 2e^t & e^t \\ -e^{-2t} & -e^{-2t} & e^{-2t} \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 2e^t + e^{-2t} & -e^t + e^{-2t} & e^t - e^{-2t} \\ -e^t + e^{-2t} & 2e^t + e^{-2t} & e^t - e^{-2t} \\ e^t - e^{-2t} & e^t - e^{-2t} & 2e^t + e^{-2t} \end{pmatrix}. \end{aligned}$$

Die Spalten bilden bilden nach S. 25.2 ein Fundamentalsystem.
Direkter erhält man allerdings (vgl. B. 25.4) das Fundamentalsystem

$$e^t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad e^t \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \qquad e^{-2t} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

2. Wir betrachten das lineare System

$$y' = \begin{pmatrix} 1 & 3 & 4 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} y = Ay\,.$$

Es gilt

$$\det(A - AE) = (1-\lambda)(\lambda^2 - 2\lambda + 2)\,,$$

also haben wir (in $\mathbb{C}$) die Eigenwerte

$$\lambda_1 = 1\,, \quad \lambda_2 = 1 + i\,, \quad \lambda_3 = 1 - i \quad (= \overline{\lambda_2})\,.$$

Zugehörige Eigenvektoren sind

$$c_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad c_2 = \begin{pmatrix} 3-4i \\ i \\ 1 \end{pmatrix}, \qquad c_3 = \begin{pmatrix} 3+4i \\ -i \\ 1 \end{pmatrix} \qquad (= \overline{c_2})\,.$$

Damit bilden etwa

$$e^t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad e^{(1+i)t} \begin{pmatrix} 3-4i \\ i \\ 1 \end{pmatrix}, \qquad e^{(1-i)t} \begin{pmatrix} 3+4i \\ -i \\ 1 \end{pmatrix}$$

ein Fundamentalsystem von $y' = Ay$ (als Gleichung in $\mathbb{C}$ betrachtet).
Ein reelles Fundamentalsystem erhält man, indem man

$$e^{(1\pm i)t} \begin{pmatrix} 3 \mp 4i \\ \pm i \\ i \end{pmatrix}$$

durch

$$\mathrm{Re}\left(e^{(1+i)t} \begin{pmatrix} 3-4i \\ i \\ 1 \end{pmatrix} \right) \qquad \text{und} \qquad \mathrm{Im}\left(e^{(1+i)t} \begin{pmatrix} 3-4i \\ i \\ 1 \end{pmatrix} \right)$$

ersetzt. ([Ü])

Außerdem gilt

$$e^{tA} = \begin{pmatrix} 1 & 3-4i & 3+4i \\ 0 & i & -i \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} e^t & 0 & 0 \\ 0 & e^{(1+i)t} & 0 \\ 0 & 0 & e^{(1-i)t} \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 3-4i & 3+4i \\ 0 & i & -1 \\ 0 & 1 & 1 \end{pmatrix}^{-1}}_{\begin{pmatrix} 1 & 4 & -3 \\ 0 & -i/2 & 1/2 \\ 0 & i/2 & 1/2 \end{pmatrix}}$$

$$= \begin{pmatrix} e^t & 4e^t - 4e^t\cos t + 3e^t \sin t & -3e^t + 3e^t \cos t + 4e^t \sin t \\ 0 & e^t \cos t & -e^t \sin t \\ 0 & e^t \sin t & e^t \cos t \end{pmatrix}.$$

Schwieriger wird die Berechnung eines Fundamentalsystems natürlich dann, wenn A nicht diagonalisierbar ist. Aus der Linearen Algebra sollte bekannt sein, dass jede Matrix $A \in \mathbb{C}^{d\times d}$ (die natürlich auch rein reelle Einträge haben kann) ähnlich zu einer Blockdiagonalmatrix der Form

$$B = \begin{pmatrix} \boxed{J_1} & & & \\ & \boxed{J_2} & & 0 \\ & & \ddots & \\ & 0 & & \boxed{J_m} \end{pmatrix}$$

ist (d.h. $A = CBC^{-1}$ für eine Matrix C mit $\det(C) \neq 0$), wobei der Jordan-Block J_k die Form

$$J_k = \begin{pmatrix} \lambda_k & 1 & 0 & 0 & \dots & 0 \\ & \lambda_k & 1 & 0 & \dots & 0 \\ & & \ddots & & & \vdots \\ O & & & \ddots & & 0 \\ & & & & & 1 \\ & & & & & \lambda_k \end{pmatrix} \quad \text{bzw.} \quad J_k = (\lambda_k)$$

hat.
Nach S. 25.3 reduziert sich die Berechnung der Matrix e^{tA} auf die Berechnung der Matrizen C, C^{-1} sowie e^{tJ_k} $(k = 1, ..., m)$. Aus Sicht der Analysis stellt sich dabei insbesondere die Frage nach der Berechnung von e^{tJ}, wobei J eine Jordan-Matrix obiger Gestalt ist.

Satz 25.6 *Es sei*

$$J = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & O & & \ddots & 1 \\ & & & & \lambda \end{pmatrix} \in \mathbb{K}^{r\times r}$$

eine Jordan-Matrix. Dann gilt

$$e^{tJ} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \dots & \frac{t^{r-1}}{(r-1)!} \\ & 1 & t & \dots & \frac{t^{r-2}}{(r-2)!} \\ & & \ddots & \ddots & \vdots \\ & O & & \ddots & t \\ & & & & 1 \end{pmatrix} .$$

Beweis. Zunächst ist

$$J = \lambda E_r + F \qquad \text{mit} \qquad F = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & O & & \ddots & 1 \\ & & & & 0 \end{pmatrix} \in \mathbb{K}^{r\times r} ,$$

also (beachte: λE_r und F vertauschen)

$$e^{tJ} = e^{t\lambda E_r} e^{tF} = e^{t\lambda} e^{tF} .$$

Weiter rechnet man nach, dass

$$F^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \dots & 0 \\ & \ddots & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \ddots & 0 \\ & & & \ddots & \ddots & 1 \\ & O & & & \ddots & 0 \\ & & & & & 0 \end{pmatrix}, \dots, F^{r-1} = \begin{pmatrix} 0 & \dots & 0 & 1 \\ & \ddots & & 0 \\ & O & \ddots & \vdots \\ & & & 0 \end{pmatrix} , \quad F^\nu = 0 \quad (\nu \geq r)$$

gilt (d.h. die „1-Diagonale rückt jeweils um einen Schritt nach rechts“). Hieraus folgt

$$e^{tF} = \sum_{\nu=0}^{r-1} \frac{1}{\nu!} t^\nu F^\nu = \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{r-1}}{(r-1)!} \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & t \\ & & & & 1 \end{pmatrix} .$$

□

Beispiel 25.7 Es sei

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} \boxed{J_1} & O \\ O & \boxed{J_2} \end{pmatrix}$$

mit

$$J_1 = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}, \qquad J_2 = (2) .$$

Dann gilt

$$e^{tA} = \begin{pmatrix} e^{-t} & te^{-t} & 0 \\ 0 & e^{-t} & 0 \\ 0 & 0 & e^{2t} \end{pmatrix} .$$

Zusammenfassend erhalten wir mit S. 25.6 und den vorangegangenen Überlegungen folgendes Ergebnis über die Struktur des Lösungsraumes.

Satz 25.8 *Es sei $A \in \mathbb{C}^{d\times d}$, und es seien $J_1, ..., J_m$ die zugehörigen Jordanblöcke. Dann existiert ein Fundamentalsystem der Form*

$$e^{\lambda_k t} P^{(k,\ell)}(t) \qquad k = 1, ..., m;\ \ell = 0, ..., r_k - 1 ,$$

wobei $J_k = \lambda_k E_{r_k} + F_k$ wie oben, und wobei die Komponenten $P_1^{(k,\ell)}, ..., P_d^{(k,\ell)}$ von $P^{(k,\ell)}$ Polynome vom Grad $\leq \ell$ sind.

Beweis. Es sei $C \in \mathbb{C}^{d\times d}$ wie oben, d.h.

$$e^{tA} C = C \cdot \mathrm{diag}(e^{tJ_1}, ..., e^{tJ_m}) = C \cdot \mathrm{diag}(e^{\lambda_1 t} e^{tF_1}, ..., e^{\lambda_m t} e^{tF_m}) .$$

Da e^{tA} eine Fundamentalmatrix ist, ist

$$e^{tA}Ce_1, ..., e^{tA}Ce_d$$

ein Fundamentalsystem. Es gilt mit $C = (c_{jk})$

$$C \cdot \operatorname{diag}(e^{\lambda_1 t}e^{tF_1}, ..., e^{\lambda_m t}e^{tF_m}) =$$

$$\begin{pmatrix} c_{11} & \dots & c_{1,r_1} & \dots & c_{1,r_2} & \dots & c_{1,d} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots & & \vdots \\ c_{d1} & \dots & c_{d,r_1} & \dots & c_{d,r_2} & \dots & c_{d,d} \end{pmatrix} \begin{pmatrix} \boxed{e^{\lambda_1 t}e^{tF_1}} & & & O \\ & \boxed{e^{\lambda_2 t}e^{tF_2}} & & \\ & & \ddots & \\ O & & & \boxed{e^{\lambda_m t}e^{tF_m}} \end{pmatrix}.$$

Die ersten r_1-Spalten haben die Form

$$\begin{matrix} e^{\lambda_1 t}c_{11} & e^{\lambda_1 t}(c_{11}t + c_{12}) & \dots & e^{\lambda_1 t}\Big(c_{11}\frac{t^{r_1-1}}{(r_1-1)!} + c_{12}\frac{t^{r_1-2}}{(r_2-1)!} + \dots + c_{1r_1}\Big) \\ \vdots & \vdots & & \vdots \\ e^{\lambda_1 t}c_{d1} & e^{\lambda_1 t}(c_{d1}t + c_{d2}) & \dots & e^{\lambda_1 t}\Big(c_{d1}\frac{t^{r_1-1}}{(r_1-1)!} + c_{d2}\frac{t^{r_1-2}}{(r_2-1)!} + \dots + c_{dr_1}\Big) \end{matrix}$$

und entsprechend in den

$$\begin{matrix} \text{Spalten}\, r_1 + 1, ..., r_2 & \text{mit} & e^{\lambda_2 t}(\dots) \\ \vdots & & \\ \text{Spalten}\, r_{m-1} + 1, ..., r_m & \text{mit} & e^{\lambda_m t}(\dots) \end{matrix}.$$

Damit ergibt sich die Behauptung. □

Ein wesentlicher Untersuchungsgegenstand bei Differentialgleichungen ist die Frage nach dem Verhalten von Lösungen, wenn t sich einem der Randpunkte des Lösungsintervalls nähert. Bei linearen Systemen mit konstanten Koeffizienten interessiert also das Verhalten für $t \to \infty$ (oder $t \to -\infty$; wir beschränken uns auf den Fall $t \to \infty$). Als Anwendung von S. 25.8 erhalten wir

Satz 25.9 *(Stabilitätskriterium)*
Es sei $A \in \mathbb{C}^{d\times d}$ und $\sigma(A) := \{\lambda : \lambda$ Eigenwert von $A\}$. Dann sind äquivalent:

a) Es existieren $M, \alpha > 0$ mit $||e^{tA}|| \le Me^{-\alpha t}$ für $t \ge 0$.

b) Für alle Lösungen φ von (25.1) gilt $\varphi(t) \to 0$ $(t \to \infty)$.

c) Für alle $\lambda \in \sigma(A)$ ist

$$\operatorname{Re}(\lambda) < 0\,.$$

Beweis. c) $\Rightarrow$ a): Es sei $\alpha > 0$ so, dass $\mathrm{Re}(\lambda) < -\alpha (< 0)$ für alle $\lambda \in \sigma(A)$. Für jedes $k \in \{1, ..., d\}$ ist $e^{tA}e_k$ Linearkombination von Funktionen der Form

$$e^{\lambda t} t^\ell y\,,$$

wobei $\ell \in \mathbb{N}_0, \lambda \in \sigma(A)$ und $y \in \mathbb{K}^d$. Für jedes solche Tripel (ℓ, λ, y) existiert ein $M_{\ell,\lambda,y} > 0$ mit

$$|e^{\lambda t} t^\ell y| \leq e^{t\mathrm{Re}\lambda} t^\ell |y| \leq M_{\ell,\lambda,y} e^{-\alpha t} \qquad (t \geq 0)\,.$$

Also existieren $M_k > 0$ mit $|e^{tA}e_k| \leq M_k e^{-\alpha t}$ für $t \geq 0$. Da

$$||e^{tA}|| \leq \sum_{k=1}^{d} |e^{tA}e_k| \leq e^{-\alpha t} \sum_{k=1}^{d} M_k$$

gilt, folgt a).

a) $\Rightarrow$ b): Ist $\varphi(0) = \eta$, so ist $\varphi(t) = e^{tA}\eta$, also $|\varphi(t)| \leq ||e^{tA}||\,|\eta| \to 0\,(t \to \infty)$.

b) $\Rightarrow$ c): Angenommen, es existiert ein Eigenwert λ von A mit $\mathrm{Re}(\lambda) \geq 0$. Ist $P^{(0)}$ das zugehörige Polynom vom Grad 0 aus S. 25.8 (dann ist $P^{(0)} \neq 0$ ein Eigenvektor von A), so ist

$$\varphi(t) = e^{\lambda t} P^{(0)}$$

eine Lösung von (25.1) mit $|\varphi(t)| = e^{t\mathrm{Re}\lambda}|P^{(0)}| \geq |P^{(0)}|$ für alle $t \geq 0$. Widerspruch zu $\varphi(t) \to 0$ for $t \to \infty$! $\square$

Bemerkung 25.10 Hat man ein (inhomogenes) lineares System mit konstanter Koeffizientenmatrix A, also

$$y' = Ay + b(t)$$

mit $A \in \mathbb{K}^{d\times d}$ und $b : I \to \mathbb{K}^d$ stetig, so erhält man eine spezielle Lösung der inhomogenen Gleichung nach S. 24.8 durch

$$\varphi(t) = \varphi_b(t, \xi, 0) = e^{tA} \int_\xi^t e^{-sA} b(s)ds$$

(Man beachte: Nach S. 25.3.1 ist für beliebige Matrizen $B \in \mathbb{K}^{d\times d}$

$$E_d = e^O = e^{B-B} = e^B e^{-B}\,,$$

also ist e^B invertierbar mir $(e^B)^{-1} = e^{-B}$).

Außerdem ist dann

$$\varphi_b(t, \xi, \eta) = \varphi_0(t, \xi, \eta) + \varphi(t) = e^{tA}\Big[e^{-\xi A}\eta + \int_\xi^t e^{-sA} b(s)ds\Big]\,.$$

Ist auch die rechte Seite b unabhängig von t (also $b(t) \equiv b$ auf $\mathbb{R}$), so erhält man dabei, falls A invertierbar ist,

$$\begin{aligned}\varphi_b(t,\xi,\eta) &= e^{tA}(e^{-\xi A}\eta - A^{-1}e^{-tA}b + A^{-1}e^{-\xi A}b)\\ &= e^{(t-\xi)A}(\eta + A^{-1}b) - A^{-1}b\end{aligned}$$

Beispiel 25.11 Wir betrachten das lineare System aus B. 22.8, d.h.

$$\begin{pmatrix} Y' \\ r' \end{pmatrix} = \begin{pmatrix} -s & -a \\ d & 0 \end{pmatrix} \begin{pmatrix} Y \\ r \end{pmatrix} + \begin{pmatrix} 0 \\ -\overline{M} \end{pmatrix}$$

(mit Konstanten $s := 1 - c, a, d, \overline{M} > 0$). Es gilt

$$\begin{aligned}&\det(A - \lambda E) := (-s-\lambda)(-\lambda) + ad = 0\\ \Leftrightarrow &\begin{cases} \lambda_{1,2} = -\frac{s}{2} \pm \sqrt{\frac{s^2}{4} - ad}\,, & \text{falls} \quad \Delta := \frac{s^2}{4} - ad \geq 0 \\ \lambda_{n,2} = -\frac{s}{2} \pm i\sqrt{ad - \frac{s^2}{4}} & \text{falls} \quad \Delta < 0 \end{cases}.\end{aligned}$$

Dabei gilt in beiden Fällen

$$\operatorname{Re}(\lambda_{1,2}) < 0\,,$$

also konvergieren nach S. 25.9 alle Lösungen der homogenen Gleichung gegen 0 für $t \to \infty$ (mit exponentieller Geschwindigkeit).
Folglich erhalten wir mit B. 25.10 für alle Lösungen φ der inhomogenen Gleichung

$$\begin{pmatrix} Y(t) \\ r(t) \end{pmatrix} = \varphi(t) \to -A^{-1}b = \begin{pmatrix} 0 & -1/d \\ 1/a & s/(da) \end{pmatrix} \begin{pmatrix} 0 \\ -\overline{M} \end{pmatrix} = \begin{pmatrix} \overline{M}/d \\ -s\overline{M}/(da) \end{pmatrix} \qquad (t \to \infty)\,.$$

Insbesondere würde nach diesem Modell das Einkommen $Y(t)$ sich für $t \to \infty$ der Konstante $\overline{M}/d$ annähern, d.h. ist $\overline{M}$ (Geldangebot) „groß“ und d (Verhältnis von Geldnachfrage zu Einkommen) „klein“, so wird das Einkommen im Zeitverlauf „groß“ werden. [„It's money, that matters“]

Wie im Abschnitt vorher wollen wir uns auch hier gesondert mit (skalaren) linearen Differenzialgleichungen höherer Ordnung beschäftigen. Es seien $a_0, a_1, ..., a_{n-1} \in \mathbb{K}$. Dann heißt eine Gleichung der Form

$$y^{(n)} = \sum_{\nu=0}^{n-1} a_\nu y^{(\nu)}$$

bzw.

$$L(y) := y^{(n)} - \sum_{\nu=0}^{n-1} a_\nu y^{(\nu)} = 0 \qquad (25.3)$$

eine (homogene) *lineare Differenzialgleichung $n-$ter Ordnung mit konstanten Koeffizienten.*

Natürlich kann man zur Berechnung eines Fundamentalsystems, also n linear unabhängiger Lösungen von (25.3) auf $\mathbb{R}$, so vorgehen, dass man die Matrix e^{tA} für das entsprechende lineare System (24.9) mit

$$A = \begin{pmatrix} 0 & 1 & \dots & \dots & \dots & 0 \\ & \ddots & & & & \vdots \\ & & \ddots & & & \vdots \\ & & & \ddots & & \vdots \\ & & & & 0 & 1 \\ a_0 & a_1 & \dots & \dots & a_{n-2} & a_{n-1} \end{pmatrix} \in \mathbb{K}^{n \times n}$$

bestimmt. Wir wollen hier jedoch eine direktere Methode herleiten, die lediglich die Eigenwerte von A (inkl. algebraischen Vielfachheiten), also die Nullstellen des charakteristischen Polynoms, benötigt.
Durch Entwicklung nach der letzten Zeile sieht man, dass gilt

$$P(\lambda) := \det(A - \lambda E) = (-1)^n(\lambda^n - a_{n-1}\lambda^{n-1} - \dots - a_1\lambda - a_0)\,,$$

d.h. die Koeffizienten sind bis auf Vorzeichen die $a_0, ..., a_{n-1}$. Es gilt nun

Satz 25.12 *Es seien $\lambda_1, ..., \lambda_m$ die paarweise verschiedenen Nullstellen des charakteritischen Polynoms P mit algebraischen Vielfachheiten $r_1, ..., r_m$. Dann bilden die Funktionen*

$$e^{\lambda_k t} t^{\ell} \qquad (\ell = 0, ..., r_k - 1; k = 1, ..., m)$$

ein Fundamentalsystem von (25.3) (im Falle $\mathbb{K} = \mathbb{C}$).

Beweis. Ist

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & \dots & \dots & 0 \\ & & & & & \vdots \\ & & & & 0 & 1 \\ a_0 & a_1 & a_2 & \dots & a_{n-2} & a_{n-1} \end{pmatrix}$$

die Koeffizientenmatrix des entsprechenden Systems (24.9), so existiert nach S. 25.8 ein Fundamentalsystem für dieses System, bestehend aus Funktionen der Form

$$e^{\lambda_{\mu_k} t} P^{(k,\ell)} \qquad (k = 1, ..., \tilde{m};\ \ell = 0, ..., \tilde{r}_k - 1)\,,$$

wobei $\tilde{m}$ die Anzahl der entsprechenden Jordanblöcke und $\tilde{r}_k$ deren Dimension sowie $\mu_k \in \{1, ..., m\}$ ist (dabei gilt $\tilde{m} \geq m$). Da die Abbildung $j : L_0 - M_0$ aus dem Beweis zu S. 24.11 ein Isomorphismus ist, ergeben die ersten Komponenten

$$e^{\lambda_{\mu_k} t} Q^{(k,\ell)}(t) := e^{\lambda_k t} P_1^{(k,\ell)}(t) \qquad (k = 1, ..., \tilde{m};\ \ell = 0, ..., \tilde{r}_k - 1)$$

eine Basis von M_0, also ein Fundamentalsystem zu (25.3). Dabei sind die Funktionen $Q^{(k,\ell)}$ Polynome vom Grad $\leq \ell$. Hieraus folgt, dass $\tilde{m} = m$ sein muss, d.h. zu jedem Eigenwert gehört nur ein Jordanblock (was i.a. nicht der Fall ist). (Denn sonst wäre $\lambda_{\mu_k} = \lambda_{\mu_{k'}}$ für $k, k' \in \{1, ..., m\}, k \neq k'$. Dann wären aber $e^{\lambda_{\mu_k} t} Q^{(k,0)}$ und $e^{\lambda_{\mu_{k'}} t} Q^{(k',0)}$ linear abhängig.)
O.E. sei nun $\mu_k = k$ für $k = 1, ..., m$ (ansonsten: Umnummerieren).
Damit ist, da $n = \sum\limits_{k=1}^{\tilde{m}} \tilde{r}_k = \sum\limits_{k=1}^{m} r_k$, auch $r_k = \tilde{r}_k$. Also haben wir ein Fundamentalsystem der Form

$$e^{\lambda_k t} \cdot Q^{(k,\ell)}(t) \qquad (k = 1, ..., m;\ \ell = 0, ..., r_k - 1)\,.$$

Weiter bilden die Polynome $Q^{(k,0)}, ..., Q^{(k,r_k-1)}$ eine Basis des Raumes der Polynome vom Grad $\leq r_k - 1$ (denn aus der linearen Unabhängigkeit von

$$e^{\lambda_k t} Q^{(k,0)}, ..., e^{\lambda_k t} Q^{(k,r_k-1)}$$

folgt auch die lineare Unabhängigkeit der Polynome $Q^{(k,0)}, ..., Q^{(k,r_k-1)}$). Also lässt sich jedes Monom $t^\ell, \ell = 0, ..., r_k - 1$, als Linearkombination der $Q^{(k,0)}, ..., Q^{(k,r_k-1)}$ schreiben. Folglich sind

$$e^{\lambda_k t} t^\ell \qquad (k = 1, ..., m;\ \ell = 0, ..., r_k - 1)$$

Lösungen von (25.3). Nach obigen Überlegungen ist jedes $\varphi \in M_0$ Linearkombination dieser (insgesamt n) Funktionen. Also bilden diese eine Basis des n-dimensionalen Raumes M_0. □

Bemerkung 25.13 Ist $\lambda = \mu + i\sigma$ eine nichtreelle r-fache Nullsteelle und ist $A \in \mathbb{R}^{d \times d}$, so ist auch $\overline{\lambda} = \mu - i\sigma$ eine r-fache Nullstelle charakteristischen Polynoms. In diesem Fall haben wir also die $2r$ Lösungen

$$\begin{aligned} e^{\lambda t} t^\ell &= e^{\mu t} e^{i\sigma t} t^\ell \\ e^{\overline{\lambda} t} t^\ell &= e^{\mu t} e^{-i\sigma t} t^\ell \end{aligned} \qquad \ell = 0, ..., r - 1\,.$$

Ein reelles Fundamentalsystem erhält man dann, indem man für alle solchen Eigenwerte diese Lösungen ersetzt durch ([Ü])

$$\begin{aligned} \mathrm{Re}(e^{\lambda t} t^\ell) &= e^{\mu t} \cos(\sigma t) t^\ell \\ \mathrm{Im}(e^{\lambda t} t^\ell) &= e^{\mu t} \sin(\sigma t) t^\ell \end{aligned} \qquad \ell = 0, ..., r - 1\,.$$

Beispiel 25.14 Auch hier betrachten wir noch einmal das B. 22.8. Dort hatten wir für Y auch die lineare Differenzialgleichung 2. Ordnung

$$Y - (c-1)Y' - adY + a\overline{M}$$

hergeleitet. Für das charakteristische Polynom gilt hier (mit $s = 1 - c$)

$$P(\lambda) = \lambda^2 - (c-1)\lambda + ad = \lambda^2 + s\lambda + ad\,,$$

also

$$P(\lambda) = 0 \Leftrightarrow \lambda_{1,2} = \begin{cases} -\frac{s}{2} \pm \sqrt{\frac{s^2}{4} - ad}\,, & \text{falls } \frac{s^2}{4} - ad \geq 0 \\ -\frac{s}{2} \pm i\sqrt{ad - \frac{s^2}{4}}\,, & \text{sonst} \end{cases}$$

(vgl. B. 25.11). Offensichtlich ist

$$u_t(t) = \frac{\overline{M}}{d}$$

eine spezielle Lösung der inhomogenen Gleichung.
Nach S. 24.11, S. 25.12 und B. 25.13 sind alle Lösungen von der Form

$$U(t) = \begin{cases} \overline{M}/d + c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}\,, & \text{falls } s^2/4 - ad > 0 \\ \overline{M}/d + c_1 e^{st/2} + c_2 t e^{-st/2}\,, & \text{falls } s^2/4 - ad = 0 \\ \overline{M}/d + c_1 e^{-st/2}\cos(\sigma t) + c_2 e^{-st/2}\sin(\sigma t)\,, & \text{falls } \sigma^2 := ad - s^2/4 > 0 \end{cases} ,$$

wobei $c_1, c_2 \in \mathbb{R}$ beliebig sind.

26 Wege und Kurven

Ist $\varphi : I \to \mathbb{K}^d$ Lösung einer Differentialgleichung, so beschreibt $\{\varphi(t) : t \in I_0\}$, wobei $I_0 \subset I$ kompakt ist, eine sogenannte Kurve in $\mathbb{K}^d$. Bevor wir uns etwas systematischer mit Kurven beschäftigen, betrachten wir zunächst eine weitere Klasse skalarwertiger Funktionen.

Definition 26.1 Es sei $f : [a,b] \to \mathbb{K}$. Ist $Z = \{x_0, x_1, ..., x_n\}$ eine Zerlegung von $[a,b]$, so setzen wir

$$V_Z(f) = \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| \,.$$

Ist

$$V(f) := \bigvee_a^b (f) := \sup\{V_z(f) : Z \text{ Zerlegung von } [a,b]\} < \infty \,,$$

so heißt f *von beschränkter Variation* (oder *von beschränkter Schwankung*) auf $[a,b]$. Wir setzen

$$BV[a,b] := BV([a,b],\mathbb{K}) := \{f : [a,b] \to \mathbb{K} : f \text{ von beschränkter Variation auf } [a,b]\} \,.$$

Die Zahl $\bigvee_a^b(f) \in [0,\infty]$ heißt *(Total-)Variation* von f auf $[a,b]$.

Beispiel 26.2 1. Es sei $f(x) = x$ $(x \in [a,b])$. Dann gilt für alle Zerlegungen $Z = \{x_1, ..., x_n\}$

$$V_Z(f) = \sum_{j=1}^{k} (x_j - x_{j-1}) = b - a \,.$$

Also ist $f \in BV[a,b]$ mit

$$\bigvee_a^b (f) = b - a \,.$$

2. Es sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} x \cdot \cos\left(\frac{\pi}{x}\right), & x \neq 0 \\ 0\,, & x = 0 \end{cases} \,.$$

Dann ist f stetig auf $\mathbb{R}$ ([Ü]). Wir betrachten $[a,b] = [0,1]$. Ist $Z = Z_n = \{0, \frac{1}{n}, \frac{1}{n-1}, ..., \frac{1}{2}, 1\} = \{x_0, ..., x_n\}$, so gilt

$$\begin{aligned} V_Z(f) &\geq \sum_{j=2}^{n} |f(x_j) - f(x_{j-1})| = \sum_{j=2}^{n} \left| f\left(\frac{1}{n+1-j}\right) - f\left(\frac{1}{n+2-j}\right) \right| \\ &= \sum_{j=2}^{n} \left| \frac{1}{n+1-j} \underbrace{\cos(\pi(n+1-j))}_{=(-1)^{n+1-j}} - \frac{1}{n+2-j} \underbrace{\cos(\pi(n+2-j))}_{=(-1)^{n+2-j}} \right| \\ &= \sum_{j=2}^{n} \left(\frac{1}{n+1-j} + \frac{1}{n+2-j} \right) = \sum_{k=1}^{n-1} \frac{1}{k} + \sum_{k=2}^{n} \frac{1}{k} \to \infty \qquad (n \to \infty)\,, \end{aligned}$$

also ist f nicht von beschränkter Variation auf $[0,1]$. (Dasselbe gilt für jedes Intervall $[a,b]$ mit $0 \in [a,b]$.)

Im folgenden Satz werden zwei einfache hinreichende Bedingungen für die Beschränktheit der Variation gegeben.

Satz 26.3 *Es sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall und $f : [a,b] \to \mathbb{K}$.*

1. *Ist f stetig auf $[a,b]$ und differenzierbar auf (a,b) mit beschränkter Ableitung, so ist $f \in BV[a,b]$.*
2. *Ist f (reellwertig und) monoton, so ist $f \in BV[a,b]$, und es ist*

$$\bigvee_a^b(f) = |f(b) - f(a)|.$$

Beweis. 1. Es sei $M > 0$ so, dass $|f'(\xi)| \leq M$ für alle $\xi \in (a,b)$ gilt. Ist $Z = \{x_0, ..., x_n\}$ eine Zerlegung von $[a,b]$, so existieren, falls f reellwertig ist, nach dem Mittelwertsatz $\xi_1, ..., \xi_n \in (a,b)$ mit

$$f(x_j) - f(x_{j-1}) = f'(\xi_j)(x_j - x_{j-1}) \qquad (j = 1, ..., n)\,.$$

Also gilt

$$V_Z(f) = \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| \leq M \sum_{j=1}^{n}(x_j - x_{j-1}) = M(b-a)$$

und damit auch

$$\bigvee_a^b(f) = \sup_Z V_Z(f) \leq M(b-a)\,.$$

Ist f komplexwertig, $f = u + iv$ wobei u, v reellwertig, so wendet man die gleiche Argumentation auf u und v an.
2. Es sei o.E. $f \uparrow$ und es sei $Z := \{x_0, ..., x_n\}$ eine Partition von $[a,b]$. Dann gilt

$$V_Z(f) = \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| = \sum_{j=1}^{n}(f(x_j) - f(x_{j-1})) = f(b) - f(a)\,.$$

Da Z beliebig war, folgt die Behauptung. □

Wir stellen einige einfache Eigenschaften der Klasse $BV[a,b]$ zusammen.

Satz 26.4 *Es sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall. Dann gilt*

1. *$BV[a, b]$ ist ein linearer Teilraum von $B([a, b], \mathbb{K})$.*

2. *Mit $f, g \in BV[a, b]$ gilt auch $f \cdot g \in BV[a, b]$.*

3. *Ist $c \in (a, b)$, so gilt $f \in BV[a, b]$ genau dann, wenn $f \in BV[a, c]$ und $f \in BV[c, b]$, und in diesem Falle ist*

$$\bigvee_a^b(f) = \bigvee_a^c(f) + \bigvee_c^b(f)\,.$$

Beweis. 1. und 2. ergeben sich leicht aus der Definition 26.1 ([Ü]).
Wir zeigen die Aussage 3.
„$\Rightarrow$ “ Es sei $f \in BV[a, b]$. Sind $Z_{a,c}$ bzw. $Z_{c,b}$ Zerlegungen von $[a, c]$ bzw. $[c, b]$, so ist $Z = Z_{a,c} \cup Z_{c,b}$ eine Zerlegung von $[a, b]$, und es gilt

$$V_{Z_{a,c}}(f) + V_{Z_{c,b}}(f) = V_Z(f) \le \bigvee_a^b(f)\,.$$

Also gilt: $\bigvee_a^c(f) < \infty$, $\bigvee_c^b(f) < \infty$ und

$$\bigvee_a^c(f) + \bigvee_c^b(f) \le \bigvee_a^b(f)\,.$$

„$\Leftarrow$“ Es sei $Z = \{x_0, ..., x_n\}$ eine Zerlegung von $[a, b]$. Dann existiert ein $k \in \{0, \dots, n-1\}$ mit $x_k \le c < x_{k+1}$, und damit ist $Z_{a,c} = \{x_0, ..., x_k, c\}$ (bzw. $Z_{a,c} = \{x_0, ..., x_{k-1}, c\}$ im Falle $x_k = c$) eine Zerlegung von $[a, c]$ und $Z_{c,b} = \{c, x_{k+1}, ..., x_n\}$ eine Zerlegung von $[a, b]$. Es gilt

$$\begin{aligned} V_Z(f) &= \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| \le \sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| + |f(c) - f(x_k)| \\ &+ |f(x_{k+1}) - f(c)| + \sum_{j=k+1}^{n} |f(x_j) - f(x_{j-1})| \\ &= V_{Z_{a,c}}(f) + V_{Z_{c,b}}(f) \le \bigvee_a^c(f) + \bigvee_c^b(f)\,. \end{aligned}$$

Also ist $\bigvee_a^b(f) < \infty$ (d.h. $f \in BV[a, b]$) und

$$\bigvee_a^b(f) \le \bigvee_a^c(f) + \bigvee_c^b(f)\,.$$

□

Als Konsequenz erhalten wir

Satz 26.5 *Es sei $f \in BV[a,b]$. Wir definieren $V : [a,b] \to \mathbb{R}$ durch*

$$V(x) := \bigvee_a^x (f) \qquad (x \in [a,b])$$

(wobei $\bigvee_a^a (f) := 0$). Dann gilt:

1. *V ist monoton wachsend.*
2. *Ist $\mathbb{K} = \mathbb{R}$ so, sind auch $V \pm f$ monoton wachsend.*

Beweis. Die Monotonie von V ergibt sich direkt aus S. 26.4.3
(Denn: Ist $x_1 < x_2$, so gilt

$$V(x_2) - V(x_1) = \bigvee_a^{x_2}(f) - \bigvee_a^{x_1}(f) = \bigvee_{x_1}^{x_2}(f) + \bigvee_a^{x_1}(f) - \bigvee_a^{x_1}(f) = \bigvee_{x_1}^{x_2}(f) \geq 0.)$$

Weiter gilt für $x_1 < x_2$

$$|f(x_2) - f(x_1)| \leq \bigvee_{x_1}^{x_2}(f)\,,$$

also

$$(V(x_2) \pm f(x_2)) - (V(x_1) \pm f(x_1)) = \bigvee_{x_1}^{x_2}(f) \pm (f(x_2) - f(x_1)) \geq 0\,,$$

d.h. $V + f, V - f$ sind monoton wachsend. □

Damit ergibt sich unmittelbar folgende interessante Charakterisierung der Funktionen von beschränkter Variation.

Satz 26.6 *Ist $\mathbb{K} = \mathbb{R}$, so ist $f \in BV[a,b]$ genau dann, wenn monoton wachsende Funktionen g und h auf $[a,b]$ existieren mit*

$$f = g - h\,.$$

Beweis. Ist $f = g - h$ mit monoton wachsenden Funktionen g, h, so ist $f \in BV[a,b]$ nach Satz 26.3.2 und Satz 26.4.1.
Es sei $f \in BV[a,b]$. Dann gilt für V aus S. 26.5

$$f = \frac{1}{2}(V + f) - \frac{1}{2}(V - f)\,,$$

d.h. $g = \frac{1}{2}(V + f)$ und $h = \frac{1}{2}(V - f)$ sind wie gewünscht. □

Bemerkung 26.7 1. Ist $M[a,b] := \{f : [a,b] \to \mathbb{R} : f \uparrow\}$, so besagt S. 26.6

$$BV[a,b] = M[a,b] - M[a,b]$$

und

$$BV[a,b] = \operatorname{span}(M[a,b])$$

(in $B([a,b],\mathbb{R})$).
2. Mit Hilfe von S. 26.6 lassen sich viele Ergebnisse über monotone Funktionen auf Funktionen von beschränkter Variation übertragen. So ergibt sich etwa aus S. 12.16, dass jedes $f \in BV[a,b]$ höchstens abzählbar viele Unstetigkeitsstellen besitzt und dass diese alle von 1. Art (also Sprungstellen) sind.
Da $R[a,b]$ ein linearer Raum ist, folgt aus S. 17.7, dass jede Funktion $f \in BV[a,b]$ eine Regelfunktion ist (also $BV[a,b] \subset R[a,b]$).

Wir kommen nun zum eigentlichen Thema dieses Abschnittes.

Definition 26.8 Es sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall, und es sei (X,d) ein metrischer Raum.

1. Ist $\gamma : [a,b] \to X$ stetig, so heißt γ ein *Weg* (in X). Der Punkt $\gamma(a)$ heißt *Anfangspunkt* des Weges und $\gamma(b)$ heißt *Endpunkt* des Weges. Gilt $\gamma(a) = \gamma(b)$, so heißt der Weg *geschlossen*. Ferner heißt γ ein *Jordanweg*, falls $\gamma_{[a,b)}$ injektiv ist.

2. Es sei $\Gamma \subset X$. Dann heißt Γ eine *Kurve* (oder ein *Bogen*) in X, falls ein Weg γ existiert mit
$$\Gamma = \gamma([a,b]) = \{\gamma(t) : t \in [a,b]\}.$$
In diesem Fall heißt γ eine *Parameterdarstellung* (oder *Parametrisierung*) von Γ. Eine Kurve Γ heißt *Jordankurve*, falls eine Parameterdarstellung γ so existiert, dass der Weg γ ein Jordanweg ist. Eine solche Parameterdarstellung heißt dann auch *Jordan-Darstellung* von Γ. Schließlich nennt man eine Jordankurve *geschlossen*, falls eine Jordan-Darstellung existiert, die geschlossen ist.

Beispiel 26.9 1. Es sei $\varphi : \mathbb{R} \to \mathbb{R}^2$ definiert durch

$$\varphi(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \qquad (t \in \mathbb{R}).$$

Für $[a,b] \subset \mathbb{R}$ ist dann $\gamma = \gamma_{a,b} = \varphi_{|[a,b]}$ ein Weg. Also ist

$$\Gamma = \Gamma_{a,b} = \left\{ \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} : t \in [a,b] \right\}$$

eine Kurve in $\mathbb{R}^2$. Ist dabei $b - a \geq 2\pi$, so ist

$$\Gamma = \Gamma_{a,b} = \{x \in \mathbb{R}^2 : |x| = 1\}$$

der Einheitskreis.
Für $a = 0, b = 2\pi$ ist

$$\gamma_{a,b}(b) = \varphi(2\pi) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \varphi(0) = \gamma_{a,b}(a)\,,$$

also ist der Weg $\gamma_{0,2\pi}$ geschlossen.
Entsprechendes gilt für die Wege $\gamma_{a,a+2k\pi}$ für $k \in \mathbb{N}, a \in \mathbb{R}$. Ist $b - a \neq 2k\pi$ für alle $k \in \mathbb{Z}$, so ist $\gamma_{a,b}$ nicht geschlossen. Weiter ist für alle a, b die Kurve $\Gamma = \Gamma_{a,b}$ eine Jordankurve. (Denn: Man kann stets eine Parameterdarstellung so wählen, dass $b - a \leq 2\pi$ gilt. Dann ist $\varphi_{|[a,b)}$ injektiv, denn gilt

$$\begin{pmatrix} \cos t \\ \sin t \end{pmatrix} = \begin{pmatrix} \cos \tilde{t} \\ \sin \tilde{t} \end{pmatrix}$$

für $t, \tilde{t} \in \mathbb{R}$, so ist $\tilde{t} = t + 2k\pi$ für ein $k \in \mathbb{Z}$.)
Man beachte aber dabei: Nicht jede Parameterdarstellung γ für Γ ist eine Jordan-Darstellung! (Ist $b - a > 2\pi$, so ist $\varphi_{|[a,b)}$ nicht mehr injektiv.)
Außerdem beachte man, dass Γ viele weitere Parameterdarstellungen hat. So gilt etwa

$$\begin{aligned} \Gamma_{0,\pi} &= \left\{(\cos t, \sin t)^T : t \in [0,\pi]\right\} = \\ &= \left\{(t, \sqrt{1-t^2})^T : t \in [-1,1]\right\} \\ &= \ldots \end{aligned}$$

2. Wir betrachten die Funktion $\gamma : [0, 2\pi] \to \mathbb{R}^2$ mit

$$\gamma(t) = (1 + \cos t)\begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \qquad (t \in [0, 2\pi])\,.$$

Dann ist $\Gamma = \gamma([0, 2\pi])$, die sogenannte Kordioide (vgl. [Ü]), eine geschlossene Jordankurve (denn $\gamma_{|[0,2\pi)}$, ist injektiv und $\gamma(0) = \gamma(2\pi)$).
3. Es sei $\varphi : \mathbb{R} \to \mathbb{C}$ definiert durch $\varphi(t) = e^{(1+i)t} = e^t e^{it}$. Dann ist für jedes $[a, b] \subset \mathbb{R}$

$$\Gamma = \varphi([a, b])$$

eine Jordankurve („logarithmische Spirale“).

Definition 26.10 Sind $a_0 < a_1 < a_2 < \ldots < a_n$ reelle Zahlen und sind $\gamma^{(j)} : [a_{j-1}, a_j] \to X$ Wege mit $\gamma^{(j+1)}(a_j) = \gamma^{(j)}(a_j) (j = 1, ..., n-1)$, so heißt der Weg $\gamma : [a_0, a_n] \to X$ mit

$$\gamma(t) := \gamma^{(j)}(t), \quad \text{falls } t \in [a_{j-1}, a_j]$$

(wohldefiniert!) Summe der Wege $\gamma^{(1)}, ..., \gamma^{(n)}$. Wir schreiben dafür auch

$$\gamma = \gamma_1 \oplus \ldots \oplus \gamma_n = \bigoplus_{j=1}^{n} \gamma_j \,.$$

Ist $\Gamma_j = \gamma^{(j)}([a_{j-1}, a_j])$ (also $\gamma^{(j)}$ eine Parameterdarstellung von Γ_j) für $j = 1, ..., n$, so schreiben wir für $\Gamma = \gamma([a_0, a_n])$ auch

$$\Gamma = \Gamma_1 \oplus \ldots \oplus \Gamma_n = \bigoplus_{j=1}^{n} \Gamma_j \,.$$

Beispiel 26.11 Es seien $x^{(0)}, ..., x^{(n)} \in \mathbb{K}^d$ und $\gamma^{(j)} : [j-1, j] \to \mathbb{K}^d$ definiert durch

$$\gamma^{(j)}(t) = x^{(j-1)} + (t - j + 1)(x^{(j)} - x^{(j-1)}) \,.$$

Dann gilt $\gamma^{(j+1)}(j) = x^{(j)} = \gamma^{(j)}(j)$ und

$$\begin{aligned} \Gamma_j &= \{\gamma^{(j)}(t) : t \in [j-1, j]\} \\ &= \overline{x^{(j-1)} x^{(j)}} \cup \{x^{(j-1)}, x^{(j)}\} \end{aligned}$$

(Strecke von $x^{(j-1)}$ nach $x^{(j)}$ mit Endpunkten). Die Kurve

$$\Gamma = \Gamma_1 \oplus \ldots \oplus \Gamma_n$$

nennt man den *Polygonzug* durch $x^{(0)}, x^{(1)}, ..., x^{(n)}$.

Wir werden nun die Frage untersuchen, wie man einem Weg in sinnvoller Weise eine Länge zuordnen kann. Dazu nähern wir einen gegebenen Weg $\gamma : [a, b] \to \mathbb{K}^d$ durch Polygonzüge an: Ist $Z = \{t_0, t_1, ..., t_n\}$ eine Zerlegung von $[a, b]$, so betrachten wir den Polygonzug durch $x^{(0)} = \gamma(t_0), ..., x^{(n)} = \gamma(t_n)$.
Dessen „Länge“ ist $\sum_{j=1}^{n} |x^{(j)} - x^{(j-1)}|$. Wählt man nun immer feinere Zerlegungen von $[a, b]$, so sollten die Längen der entsprechenden Polygonzüge das annähern, was wir als Länge von γ betrachten wollen. Dies führt zu folgender Definition.

Definition 26.12 Es sei $\gamma : [a, b] \to \mathbb{K}^d$ ein Weg. Ist $Z = \{t_0, ..., t_n\}$ eine Zerlegung von $[a, b]$, so setzen wir

$$L_Z(\gamma) := \sum_{j=1}^{n} |\gamma(t_j) - \gamma(t_{j-1})| \,.$$

Ist

$$L(\gamma) := \sup\{L_Z(\gamma) : Z \text{ Zerlegung von } [a, b]\} < \infty \,,$$

so heißt γ *rektifizierbar* und $L(\gamma)$ heißt *Länge* von Γ.

Es gilt dafür

Satz 26.13 *Ist $\gamma = (\gamma_1, ..., \gamma_d) : [a,b] \to \mathbb{K}^d$ ein Weg, so ist γ genau dann rektifizierbar, wenn alle Komponentenfunktionen $\gamma_1, ..., \gamma_d$ von beschränkter Variation auf $[a,b]$ sind. Außerdem gilt dann*

$$\max_{1\le k\le d} \bigvee_a^b(\gamma_k) \le L(\gamma) \le \sum_{k=1}^d \bigvee_a^b(\gamma_k)\,.$$

Beweis. „$\Rightarrow$" Ist γ rektifizierbar, so gilt für alle $k \in \{1, ..., d\}$ und alle Zerlegungen $Z = \{t_0, ..., t_n\}$ von $[a,b]$:

$$V_Z(\gamma_k) = \sum_{j=1}^n |\gamma_k(t_j) - \gamma_k(t_{j-1})| \le \sum_{j=1}^n |\gamma(t_j) - \gamma(t_{j-1})| \le L(\gamma)\,,$$

also ist γ_k in $BV[a,b]$, und es gilt

$$\bigvee_a^b(\gamma_k) \le L(\gamma)\,.$$

„$\Leftarrow$" Ist $\gamma_k \in BV[a,b]$ für alle $k \in \{1, ..., d\}$, so gilt für jede Zerlegung $Z = \{t_0, ..., t_n\}$ von $[a,b]$ (beachte: $|x| \le ||x||_1 = \sum_{j=1}^d |x_j|$ für $x \in \mathbb{K}^d$)

$$L_Z(\gamma) = \sum_{j=1}^n |\gamma(t_j) - \gamma(t_{j-1})| \le \sum_{j=1}^n \sum_{k=1}^d |\gamma_k(t_j) - \gamma_k(t_{j-1})| \le \sum_{k=1}^d \bigvee_a^b(\gamma_k)\,.$$

Also ist γ rektifizierbar, und es gilt

$$L(\gamma) \le \sum_{k=1}^d \bigvee_a^b(\gamma_k)\,.$$

□

Bemerkung 26.14 Ist $\gamma = \oplus_{j=1}^n \gamma^{(j)}$ die Summe von Wegen $\gamma_1, ..., \gamma_n$, so ist γ genau dann rektifizierbar, wenn alle $\gamma^{(1)}, ..., \gamma^{(n)}$ rektifizierbar sind, und in diesem Falle gilt

$$L(\gamma) = L\Big(\bigoplus_{j=1}^n \gamma^{(j)}\Big) = \sum_{j=1}^n L(\gamma^{(j)})\,.$$

(Der Beweis verläuft (für $n = 2$) wie der Beweis zu S. 26.4.3, wobei lediglich „$V_Z(f)$" durch „$L_Z(\gamma)$" ersetzt werden muss.)

Es drängt sich natürlich die Frage auf, wie man Längen von Kurven konkret berechnen kann. Die Definition erweist sich schon in einfachsten Fällen als sehr ungeeignet dafür. Als wesentlich günstiger erweist sich

Satz 26.15 *Es sei $\gamma : [a, b] \to \mathbb{K}^d$ ein Weg. Ferner sei $\gamma = (\gamma_1, ..., \gamma_d)$ stetig differenzierbar auf $[a, b]$ (d.h. γ_k' existiert auf $[a, b]$ und ist stetig auf $[a, b]$ für $k = 1, ..., d$). Dann ist γ rektifizierbar, und es gilt*

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

Beweis. 1. Es sei $Z = \{t_0, ..., t_n\}$ eine Zerlegung von $[a, b]$. Dann gilt nach dem HDI (komponentenweise angewandt; vgl. Beweis zu S. 23.4)

$$|\gamma(t_j) - \gamma(t_{j-1})| = \Big| \int_{t_{j-1}}^{t_j} \gamma'(t) dt \Big| \le \int_{t_{j-1}}^{t_j} |\gamma'(t)| dt \qquad (j = 1, ..., n) \,,$$

also

$$L_Z(\gamma) = \sum_{j=1}^n |\gamma(t_j) - \gamma(t_{j-1})| \le \sum_{j=1}^n \int_{t_{j-1}}^{t_j} |\gamma'(t)| dt = \int_a^b |\gamma'(t)| dt \,.$$

Da Z beliebig war, ist γ rektifizierbar, und es gilt jedenfalls

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

2. Es bleibt noch die Gleichheit zu zeigen.
Dazu definieren wir $s : [a, b] \to \mathbb{R}$ durch

$$s(t) := L(\gamma_{|[a,t]}) \qquad (t \in [a, b])$$

mit $L(\gamma_{|[a,a]}) := 0$ (sog. Weglängenfunktion).
Dann gilt mit B. 26.14 für $t \in [a, b)$, $h > 0$ (so, dass $t + h \le b$)

$$s(t + h) - s(t) = L(\gamma_{|[0,t+h]}) - L(\gamma_{|[0,t]}) = L(\gamma_{|[t,t+h]})$$

(da $\gamma_{|[a,t+h]} = \gamma_{|[a,t]} \oplus \gamma_{|[t,t+h]}$). Folglich ist mit 1.

$$\frac{|\gamma(t + h) - \gamma(t)|}{h} \le \frac{1}{h} L(\gamma_{|[t,t+h]}) = \frac{s(t + h) - s(t)}{h} \le \frac{1}{h} \int_t^{t+h} |\gamma'| \,.$$

Für $h \to 0$ konvergieren die rechte und die linke Seite diese Ungleichung gegen $|\gamma'(t)|$ (beachte: Ist $F(t) := \int_a^t |\gamma'(s)|ds$, so ist nach dem HDI

$$|\gamma'(t)| = F'(t) = \lim_{h\to 0} \frac{F(t+h) - F(t)}{h} = \lim_{h\to 0} \frac{1}{h} \int\limits_t^{t+h} |\gamma'|\,.)$$

Also ist s rechtsseitig differenzierbar an der Stelle $t \in [a, b)$ mit

$$\lim_{h\to 0^+} \frac{s(t+h) - s(t)}{h} = |\gamma'(t)|\,.$$

Entsprechend sieht man, dass s linksseitig differenzierbar an allen Stellen $t \in (a, b]$ ist mit

$$\lim_{h\to 0^-} \frac{s(t+h) - s(t)}{h} = |\gamma'(t)|\,.$$

Also ist s differenzierbar auf $[a, b]$ mit $s'(t) = |\gamma'(t)|\,(t \in [a, b])$. Wieder mit dem HDI ergibt sich (beachte: $s' = |\gamma'|$ ist stetig)

$$L(\gamma) = s(b) - s(a) = \int\limits_a^b s'(t)\,dt = \int\limits_a^b |\gamma'(t)|dt\,.$$

□

Beispiel 26.16 Für $0 < \alpha \le \beta$ sei $\gamma : [a, b] \to \mathbb{R}^2$ definiert durch

$$\gamma(t) = \begin{pmatrix} \alpha \cos t \\ \beta \sin t \end{pmatrix} \qquad (t \in [a, b])\,.$$

Dann beschreibt γ eine Ellipse (für $b - a > 2\pi$). Es gilt

$$|\gamma'(t)| = (\alpha^2 \sin^2 t + \beta^2 \cos^2 t)^{1/2} = \beta\Big(1 - \Big(1 - \frac{\alpha^2}{\beta^2}\Big) \sin^2 t\Big)^{1/2},$$

also

$$L(\gamma) = \beta \int\limits_a^b \sqrt{1 - k^2 \sin^2 t}\,dt$$

mit $k := \sqrt{1 - \alpha^2/\beta^2}$ (sog. elliptisches Integral). Ist $\alpha = \beta$ (Kreis mit Radius β), so ist

$$L(\gamma) = \beta(b - a)\,,$$

also für $a = 0$, $b = 2\pi$ insbesondere $L(\gamma) = 2\pi\beta$.

Bemerkung und Definition 26.17 Es sei $\gamma : [a, b] \to \mathbb{K}^d$ ein Weg. Existieren Wege $\gamma^{(j)} : [a_{j-1}, a_j] \to \mathbb{K}^d\,(j = 1, ..., n)$ mit

$$\gamma = \bigoplus_{j=1}^{n} \gamma^{(j)}$$

und so, dass jedes $\gamma^{(j)}$ stetig differenzierbar ist, so heißt γ ein *Pfad*. Nach B. 26.14 und S. 26.15 ist jeder Pfad rektifizierbar, und es gilt

$$L(\gamma) = \sum_{j=1}^{d} L(\gamma^{(j)}) = \sum_{j=1}^{d} \int_{a_{j-1}}^{a_j} |\gamma^{(j)\prime}(t)|dt\,.$$

Man würde natürlich gerne einer Kurve eine Länge zuordnen, die in gewisser Weise unabhängig von der Parameterdarstellung ist. Dass dieses Vorhaben nicht unproblematisch ist, zeigt schon der einfache Fall des Kreises $\Gamma = \{x \in \mathbb{R}^2 : |x| = 1\}$. Wir hatten in B. 26.9 gesehen, dass gilt

$$\Gamma = \varphi([a, b])$$

für alle a, b mit $b - a \geq 2\pi$. Es gilt nach B. 26.16 dabei

$$L(\varphi_{|[a,b]}) = b - a\,,$$

was jeden Wert $\geq 2\pi$ annehmen kann.
Wenn man sich jedoch auf Jordankurven und Jordandarstellungen beschränkt, so kann man zeigen, dass jeder Jordankurve Γ in eindeutiger Weise eine Länge $L(\Gamma)$ (u.U. $= \infty$) zugeordnet werden kann, nämlich die Länge einer (beliebigen) Jordandarstellung (wichtig dabei: Man kann zeigen, dass für alle Jordandarstellungen die Längen gleich sind). Wir wollen auf den Beweis dazu hier nicht eingehen. Eine ausführliche Darstellung findet man etwa in Heuser, H., Lehrbuch der Analysis, Teil 2.
Für den Einheitskreis Γ von oben ergibt sich dabei natürlich

$$L(\Gamma) = L(\varphi_{|[0,2\pi]}) = 2\pi\,.$$

Wir wenden uns zum Abschluss dieses Abschnittes noch einem Begriff zu, der zu den topologischen Grundbegriffen gezählt werden kann.

Definition 26.18 Es sei (X, d) ein metrischer Raum.

1. X heißt *unzusammenhängend*, falls offene Mengen $U, V \subset X$ existieren mit

$$X = U \cup V,\ U \neq \emptyset,\ V \neq \emptyset,\ U \cap V = \emptyset\,.$$

   Andernfalls heist X *zusammenhängend*.

2. $M \subset X$ heißt *unzusammenhängend*, falls $(M, d_{|M\times M})$ unzusammenhängend ist (d. h. falls offene Mengen $U, V \subset X$ existieren mit $M \subset U \cup V, U \cap M \neq \emptyset$, $V \cap M \neq \emptyset$, $U \cap V \cap M = \emptyset$). Anderenfalls heißt M *zusammenhängend.*

Beispiel 26.19 1. Es sei $(X, d) = (\mathbb{R}, d_{|\cdot|})$ und $M = [0,1] \cup [2,3]$. Dann gilt für die offenen Mengen $U = (-\frac{1}{2}, \frac{3}{2}), V = (\frac{3}{2}, \frac{7}{2})$:

$$M \subset U \cup V, \quad U \cap M \neq \emptyset, \quad V \cap M \neq \emptyset \qquad \text{und} \qquad U \cap V \cap M = \emptyset\,,$$

also ist M unzusammenhängend.
2. Es sei $(X, d) = (\mathbb{R}, d_{|\cdot|})$ und $M = \mathbb{Q}$. Dann gilt für die offenen Mengen $U = (-\infty, \sqrt{2}), V = (\sqrt{2}, \infty)$:

$$\mathbb{Q} \subset U \cup V, \quad U \cap \mathbb{Q} \neq \emptyset, \quad V \cap \mathbb{Q} \neq \emptyset \qquad \text{und} \qquad U \cap V \cap \mathbb{Q} = \emptyset\,,$$

also ist $\mathbb{Q}$ unzusammenhängend.

3. In jedem metrischen Raum (X, d) sind die einpunktigen Mengen zusammenhängend ([Ü]).

In $(\mathbb{R}, d_{|\cdot|})$ gilt

Satz 26.20 *Eine Teilmenge M von $\mathbb{R}$ (mit üblicher Betragsmetrik $d_{|\cdot|}$) ist genau dann zusammenhängend, wenn sie ein Intervall oder einpunktig ist.*

Beweis. „$\Rightarrow$“ Es sei $M \subset \mathbb{R}$ mit $|M| \geq 2$, und M sei kein Intervall. Dann existieren Punkte a, b, c mit $a < b < c$ und $a, c \in M, b \notin M$. Folglich gilt für $U := (-\infty, b), V := (b, \infty)$

$$M \subset U \cup V, \quad U \cap M \neq \emptyset, \quad V \cap M \neq \emptyset \qquad \text{und} \qquad U \cap V \cap M = \emptyset\,,$$

also ist M unzusammenhängend.
„$\Leftarrow$“ Es sei M ein Intervall. Angenommen, M ist unzusammenhängend. Dann existieren offene Mengen U und V mit

$$M \subset U \cup V, \quad U \cap M \neq \emptyset, \quad V \cap M \neq \emptyset, \quad U \cap V \cap M = \emptyset\,.$$

Es sei $a \in U \cap M, b \in V \cap M$ und o.E. $a < b$. Wir zeigen: $(a, b) \not\subset M$, und damit ist M kein Intervall (also Widerspruch!).
Wäre $(a, b) \subset M$, also auch $[a, b] \subset M$, so würde für $U \cap [a, b]$ und $V \cap [a, b]$ gelten

$$U \cap [a, b] = V^c \cap [a, b] \qquad \text{und} \qquad V \cap [a, b] = U^c \cap [a, b]\,,$$

also sind beide Mengen abgeschlossen. Insbesondere ist

$$\xi := \sup(U \cap [a,b]) \in U \cap [a,b]\,.$$

Dabei muss $\xi \neq b$ gelten (sonst wäre $b \in U \cap V \cap M$). Nach Definition von ξ ist $(\xi, b] \subset U^c$, also $(\xi, b] \subset U^c \cap [a,b] = V \cap [a,b]$. Da $V \cap [a,b]$ abgeschlossen ist, gilt $\xi \in V \cap [a,b]$, also

$$\xi \in V \cap U \cap [a,b] \subset V \cap U \cap M\,.$$

Wieder Widerspruch!
Mit B. 26.19.3 folgt „$\Leftarrow$“. □

Der folgende Satz zeigt, dass der Zusammenhang einer Menge sich unter stetigen Abbildungen auf die Bildmenge überträgt.

Satz 26.21 *Es seien (X,d) und (Y,e) metrische Räume. Ist $M \subset X$ zusammenhängend und $f : M \to Y$ stetig, so ist auch $f(M) \subset Y$ zusammenhängend.*

Beweis. Wir können uns beim Beweis auf den Fall $M = X$ und $Y = f(M)$ beschränken.
Angenommen, Y ist unzusammenhängend. Dann existieren offene Mengen V_1, $V_2 \subset M$ mit

$$Y = V_1 \cup V_2 \quad V_1 \neq \emptyset,\ V_2 \neq \emptyset, \quad V_1 \cap V_2 = \emptyset\,.$$

Da $f : X \to Y$ stetig ist, sind

$$U_1 = f^{-1}(V_1) \qquad U_2 = f^{-1}(V_2)$$

offen (in (X,d)). Es gilt dafür

$$U_1 \neq \emptyset, \quad U_2 \neq \emptyset, \quad U_1 \cap U_2 = f^{-1}(V_1 \cap V_2) = f^{-1}(\emptyset) = \emptyset$$

und

$$U_1 \cup U_2 = f^{-1}(U_1 \cup U_2) = f^{-1}(Y) = X\,,$$

also ist X unzusammenhängend. Widerspruch! □

Als Konsequenz aus S. 26.20 und S. 26.21 erhalten wir weitreichende Verallgemeinerungen des Zwischenwertsatzes und seiner Folgerungen (S. 12.1, B. 12.2):

Satz 26.22 *Es sei (X,d) ein metrischer Raum, und es sei $M \subset X$ zusammenhängend. Ist $f : M \to \mathbb{R}$ stetig und nicht konstant, so ist $f(M) \subset \mathbb{R}$ ein Intervall. Ist M überdies kompakt, so ist $f(M)$ ein kompaktes Intervall.*

Beweis. Nach S. 26.21 ist $f(M) \subset \mathbb{R}$ zusammenhängend, und da $f \not\equiv$ const ist, ist $f(M)$ nicht einpunktig, also ein Intervall nach S. 26.20. Ist M kompakt, so ist auch $f(M)$ kompakt nach S. 11.9 (angewandt auf den kompakten metrischen Raum $(M, d_{|M\times M})$). □

Definition 26.23 Es sei (X, d) ein metrischer Raum. Eine Menge $M \subset X$ heißt *Bogen-zusammenhängend*, falls zu allen Punkten $x, y \in X$ eine Kurve Γ existiert mit $x, y \in \Gamma$ und $\Gamma \subset M$.

Wir zeigen, dass dieser Zusammenhangsbegriff strenger ist als der oben definierte.

Satz 26.24 *Es sei (X, d) ein metrischer Raum. Ist $M \subset X$ Bogen-zusammenhängend, so ist M auch zusammenhängend.*

Beweis. Wir können uns wieder auf den Spezialfall $X = M$ beschränken.
Angenommen, X ist nicht zusammenhängend, d.h. es existieren $U, V \subset X$ offen mit

$$U \cup V = X, \qquad U \neq \emptyset, \qquad V \neq \emptyset, \qquad U \cap V = \emptyset\,.$$

Es sei $x \in U, y \in V$. Da X Bogen-zusammenhängend ist, existiert eine Kurve Γ in X mit $x, y \in \Gamma$. Es sei $\gamma : [a, b] \to X$ eine Parameterdarstellung von Γ. Wir betrachten

$$\begin{aligned} A &:= \{t \in [a, b] : \gamma(t) \in U\} = \gamma^{-1}(U) \\ B &:= \{t \in [a, b] : \gamma(t) \in V\} = \gamma^{-1}(V) \end{aligned}\,.$$

Dann sind A und B abgeschlossen (und nicht leer).
(Denn: Ist (t_k) eine Folge in A mit $t_k \to s\,(k \to \infty)$, so ist jedenfalls $s \in [a, b]$. Also gilt auf Grund der Stetigkeit von γ dann auch $\gamma(t_k) \to \gamma(s) \in \Gamma\,(k \to \infty)$. Wäre $\gamma(s) \notin U$, so wäre $\gamma(s) \in V$ und damit $\gamma(t_k) \in V$ für k genügend groß (da V offen), also $\gamma(t_k) \in U \cap V = \emptyset$, was beim besten Willen nicht geht. Damit ist $\gamma(s) \in U$, also $s \in A$. Entsprechendes gilt für B.)
Da U, V offen sind, sind nach S. 10.12 auch A und B offen (im metrischen Raun $([a, b], d_{|\cdot|})$!). Da $([a, b], d_{|\cdot|})$ zusammenhängend ist, muss damit $A = B = [a, b]$ sein. (Ist (Y, e) ein zusammenhängender metrischer Raum, so sind $\emptyset$ und Y die einzigen Mengen, die sowohl offen als auch abgeschlossen sind ([Ü]).)
Dann ist aber $\Gamma \subset U \cap V$. Widerspruch zu $U \cap V = \emptyset$! □

Beispiel 26.25 I.a. folgt aus „zusammenhängend“ nicht „Bogen-zusammenhängend“. So kann man etwa für die Menge $M \subset \mathbb{R}^2$ mit

$$M = \left\{ \left(t, \sin \frac{1}{t} \right) : t \in (0,1) \right\} \cup (\{0\} \times [-1,1])$$

zeigen: M ist zusammenhängend ([Ü]) und M ist nicht Bogen-zusammenhängend (ohne Beweis).

In D. 20.7 hatten wir (für Teilmengen von $\mathbb{R}^d$, dieselbe Definition verwenden wir für $\mathbb{C}^d$) festgelegt, was „Polygon-zusammenhängend“ und „Gebiet“ heißt. Offensichtlich ist jede Polygon-zusammenhängende Menge auch Bogen-zusammenhängend. Also haben wir (in $\mathbb{K}^d$):

Polygon-zusammenhängend $\Rightarrow$ Bogen-zusammenhängend $\Rightarrow$ zusammenhängend.

Ist $G \subset \mathbb{K}^d$ *offen*, so sind alle drei Begriffe gleichbedeutend, denn es gilt

Satz 26.26 *Es sei $G \subset \mathbb{K}^d$ offen und zusammenhängend. Dann ist G auch Polygon-zusammenhängend, also ein Gebiet.*

Beweis. Es sei (o.E. $G \neq \emptyset$ und) $x^{(0)} \in G$ fest. Wir setzen

$$M := \{x \in G : \exists \, \text{Polygonzug}\, \Gamma \,\text{in}\, G, \text{ der } x, x^{(0)} \text{ verbindet}\}$$

und zeigen: M ist offen und abgeschlossen im metrischen Raum $(G, d_{|\cdot|})$. Da $M \neq \emptyset$ ist, ist dann $M = G$ ([Ü]), und damit folgt, dass G Polygon-zusammenhängend ist.
Also: M ist offen, dann ist $x \in M$, so existiert ein $\delta > 0$ mit $U_\delta(x) \subset G$. Da man x und $x^{(0)}$ durch einen Polygonzug in G verbinden kann, gilt dies auch für $x^{(0)}$ und alle $y \in U_\delta(x)$. Also ist $U_\delta(x) \subset M$.
M ist auch abgeschlossen (in $(G, d_{|\cdot|})$), denn es sei $(y^{(n)})$ eine Folge in M mit $y^{(n)} \to y \in G$. Es sei wieder $\delta > 0$ so ,dass $U_\delta(y) \subset G$ gilt. Dann existiert ein n mit $y^{(n)} \in U_\delta(y)$. Da $x^{(0)}$ und $y^{(n)}$ durch einen Polygonzug in G verbunden werden können, gilt dies auch für $x^{(0)}$ und y. Also ist $y \in M$. $\square$

Als eine Anwendung obiger Resultate ergibt sich etwa folgendes Ergebnis über die „Gebietstreue“ gewisser C^1-Abbildungen.

Satz 26.27 *Es sei $G \subset \mathbb{R}^d$ ein Gebiet, und es sei $f \in C^1(G, \mathbb{R}^d)$. Ferner gelte*

$$\det J_f(x) \neq 0 \qquad (x \in G)\,.$$

Dann ist $f(G) \subset \mathbb{R}^d$ ein Gebiet.

Beweis. Da G ein Gebiet, also insbesondere (Polygon-)zusammenhängend, ist, ist auch $f(G)$ nach S. 26.21 zusammenhängend.
Ist $y^{(0)} \in f(G)$, also $y^{(0)} = f(x^{(0)})$ für ein $x^{(0)} \in G$, so existiert nach S. 21.3 eine Umgebung V von $y^{(0)}$ mit $V \subset f(G)$. Damit ist $f(G)$ offen.
Nach S. 26.26 ist folglich $f(G)$ ein Gebiet. □

27 Lebesgue-Integral für Funktionen mehrerer Veränderlicher

Integrale haben Sie schon an verschiedenen Stellen kennen gelernt. Zum einen haben wir uns in Abschnitt 17 mit Integralen für Regelfunktionen, also gleichmäßigen Grenzwerten von Treppenfunktionen auf kompakten Intervallen $[a, b] \subset \mathbb{R}$, befasst. Weiter sind in Abschnitt 18 auch uneigentliche Integrale auf offenen Intervallen studiert worden. Schließlich haben Sie in der Wahrscheinlichkeitstheorie den Begriff des Integrals bezüglich eines allgemeinen Maßes kennen gelernt. Insbesondere wurde dabei das Integral bezüglich des Lebesgue-Maßes auf $\mathbb{R}^d$ untersucht.

Wir werden uns im Folgenden der dort erarbeiteten Ergebnisse bedienen, um insbesondere Techniken zur Berechnung von (Lebesgue-)Integralen für Funktionen mehrerer (reeller) Variablen zu entwickeln.

Zunächst werden die wichtigsten Definitionen und Ergebnisse, die in der Wahrscheinlichkeitstheorie ausführlich dargestellt wurden, noch einmal in Erinnerung gerufen.

Definition 27.1 1. Es sei Ω eine Menge. Ein Mengensystem $\mathcal{S} \subset \text{Pot}(\Omega)$ heißt *σ-Algebra* (auf Ω), falls gilt

(σ1): $\emptyset \in \mathcal{S}$

(σ1): Aus $A \in \mathcal{S}$ folgt $A^c = \Omega \backslash A \in \mathcal{S}$

(σ1): Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{S}$, so ist $\bigcup\limits_{n\in\mathbb{N}} A_n \in \mathcal{S}$.

Das Paar $(\Omega, \mathcal{S})$ heißt dann *Messraum*.

2. Ist $(\Omega, \mathcal{S})$ ein Messraum, so heißt eine Abbildung $\mu : \mathcal{S} \to [0, \infty]$ *Maß* auf $(\Omega, \mathcal{S})$, falls

(M1): $\mu(\emptyset) = 0$

(M2): Ist (A_n) eine Folge in $\mathcal{S}$ mit $A_j \cap A_k = \emptyset$, für alle $k \neq j$ („A_n paarweise disjunkt"), so gilt

$$\mu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n=1}^{\infty} \mu(A_n)\,.$$

Das Tripel $(\Omega, \mathcal{S}, \mu)$ heißt dann *Maßraum*.

Bemerkung 27.2 Ist $\mathcal{S}$ eine σ-Algebra auf Ω, so gilt auch

- $\Omega \in \mathcal{S}$

- $A, B \in \mathcal{S} \Rightarrow A \backslash B \in \mathcal{S}$
- (A_n) Folge in $\mathcal{S} \Rightarrow \bigcap\limits_{n\in\mathbb{N}} A_n \in \mathcal{S}$

Beispiel 27.3 (Unser „Held") Ist $\Omega = \mathbb{R}^d$, so heißt die σ-Algebra

$$\mathbb{B}^d := \cap\{\mathcal{S} : \mathcal{S}\ \sigma\text{-Algebra, die alle offenen Mengen enthält}\}$$

Borel σ-Algebra (auf $\mathbb{R}^d$) (die „kleinste"σ-Algebra, die alle offenen Mengen enthält). Dafür kann man zeigen: Es gibt genau ein Maß $\lambda = \lambda^d$ auf $\mathbb{B}^d$ mit

$$\lambda\Big(\prod_{j=1}^d (a_j, b_j]\Big) = \prod_{j=1}^d (b_j - a_j)\,.$$

Dieses Maß heißt *Lebesgue-Maß* auf $(\mathbb{R}^d, \mathbb{B}^d)$. Es gilt dabei mit

$$\mathcal{E}_d := \{\prod_{j=1}^d (a_j, b_j] = (a, b] : a_j \leq b_j\ (j = 1, ..., d)\}$$

und $V(I) := \prod_{j=1}^d (b_j - a_j)$ für $I \in \mathcal{E}_d$

$$\lambda(A) = \inf\Big\{\sum_{k\in\mathbb{N}} V(I_k) : (I_k) \text{ in } \mathcal{E}_d \text{ mit } A \subset \bigcup_{k\in\mathbb{N}} I_k\Big\} \qquad (A \in \mathbb{B}^d)\,.$$

Wichtig für uns ist insbesondere das Verhalten des Lebesgue-Maßes unter affin-linearen Transformationen: Ist $A \in \mathbb{R}^{d\times d}, b \in \mathbb{R}^d$ und $T : \mathbb{R}^d \to \mathbb{R}^d$ mit $T(x) = Ax + b$ für $x \in \mathbb{R}^d$, so gilt

$$\lambda(T(M)) = |\det(A)|\lambda(M) \qquad (M \in \mathbb{B}^d)\,.$$

Definition 27.4 Es sei $f : \mathbb{R}^d \to \mathbb{R}^m$. Dann heißt f *(Borel-)messbar*, falls

$$f^{-1}(B) \in \mathbb{B}^d \qquad (B \in \mathbb{B}^m)\,.$$

Für $m = 1$ ist es sinnvoll, allgemeiner Funktionen $f : \mathbb{R}^d \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ und die σ-Algebra $\overline{\mathbb{B}^1} := \{A \subset \overline{\mathbb{R}} : A \cap \mathbb{R} \in \mathbb{B}^1\}$ zu betrachten. Eine solche Funktion heißt *(Borel-)messbar*, falls

$$f^{-1}(B) \in \mathbb{B}^d$$

für alle $B \in \overline{\mathbb{B}^1}$ gilt.

Damit kann man wie folgt das Lebesgue-Integral in drei Schritten definieren.

1. Es sei $f : \mathbb{R}^d \to [0, \infty)$ messbar mit endlich vielen Funktionewerten $\alpha_1, ..., \alpha_n \in [0, \infty)$, d.h. $f = \sum_{j=1}^n \alpha_j 1_{\{f=\alpha_j\}}$, wobei $\{f = \alpha\} = f^{-1}(\{\alpha\})$. (Eine solche Funktion nennen wir *Elementarfunktion.*) Dann setzt man

$$\int f d\lambda^d := \sum_{j=1}^n \alpha_j \lambda^d(\{f = \alpha_j\}) \in [0, \infty]\,,$$

wobei $0 \cdot \infty := 0$ und $\alpha \cdot \infty := \infty$ für $\alpha > 0$.

2. Es sei $f : \mathbb{R}^d \to [0, \infty]$ messbar. Dann setzt man

$$\int f d\lambda^d := \sup \left\{ \int g d\lambda^d : g \text{ Elementarfunktion mit } g \leq f \right\}.$$

3. Es sei $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ messbar und

$$f^+ := \max(f, 0), \qquad f^- := \max(-f, 0).$$

Ist $\int f^+ d\lambda^d < \infty$ oder $\int f^- d\lambda^d < \infty$, so heißt f *(Lebesgue-) integrierbar* (auf $\mathbb{R}^d$), und man definiert

$$\int f d\lambda^d := \int f^+ d\lambda^d - \int f^- d\lambda^d$$

(wobei $\infty - \alpha := \infty, \alpha - \infty := -\infty$ für $\alpha \in \mathbb{R}$). Ferner heißt f *absolut (Lebesgue-) integrierbar* (auf $\mathbb{R}^d$), falls $\int f^+ d\lambda^d < \infty$ und $\int f^- d\lambda^d < \infty$.
Ist allgemeiner $A \in \mathbb{B}^d$ und ist $f 1_A$ (absolut) integrierbar auf $\mathbb{R}^d$, so sagt man, f sei *(absolut) integrierbar* auf A, und setzt

$$\int_A f d\lambda^d := \int f 1_A d\lambda^d.$$

Bemerkung 27.5 Man kann leicht zeigen: Ist $f \in R[a, b]$, so ist $\tilde{f} : \mathbb{R} \to \mathbb{R}$ mit

$$\tilde{f}(x) = \begin{cases} f(x) & , x \in [a, b] \\ 0 & , x \notin [a, b] \end{cases}$$

absolut Lebesgue-integrierbar, und es gilt

$$\int \tilde{f} d\lambda = \int_{[a,b]} \tilde{f} d\lambda = \int_a^b f(x) dx.$$

Außerdem gilt: Existiert das uneigentliche Integral $\int_a^{b-} f$ (bzw. $\int_{a+}^b f$ bzw. $\int_{a+}^{b-} f$), so ist $\tilde{f}$ (definiert wie oben mit $[a, b)$ bzw. $(a, b]$ bzw. (a, b) statt $[a, b]$) messbar und im Falle, dass $\tilde{f}$ Lebesgue-integrierbar ist, erhält man

$$\int \tilde{f} d\lambda = \int_{[a,b)} \tilde{f} d\lambda = \int_a^{b-} f(x) dx$$

(entsprechend für $\int_{a+}^{f} f, \int_{a+}^{b-} f$). Also: Immer, wenn Regel- und Lebesgue-Integral beide existieren, stimmen sie überein. Wir schreiben deshalb auch $\int_a^b f = \int_a^b f(x) dx$ für Lebesgue-Integrale.

Im folgenden „Mega-Satz“ fassen wir die für uns wichtigen Ergebnisse über Eigenschaften des Lebesgue-Integrals zusammen. Für die Beweise verweisen wir auch hier auf die Wahrscheinlichkeitstheorie.

Satz 27.6 • *(Satz von der monotonen Konvergenz; Levi)*
Es sei (f_n) *eine Folge messbarer Funktionen* $f_n : \mathbb{R}^d \to [0, \infty]$ *mit* $f_n \le f_{n+1}$ $(n \in \mathbb{N})$. *Dann gilt* $f : \mathbb{R}^d \to [0, \infty]$ *mit*

$$f(x) := \lim_{n\to\infty} f_n(x) \qquad (x \in \mathbb{R}^d)$$

ist messbar und

$$\int f\, d\lambda^d = \int (\lim_{n\to\infty} f_n)\, d\lambda^d = \lim_{n\to\infty} \int f_n\, d\lambda^d .$$

- *(Satz von der dominierten Konvergenz; Lebesgue)*
Es sei (f_n) *eine Folge messbarer Funktionen* $f : \mathbb{R}^d \to \overline{\mathbb{R}}$, *und es sei* $g : \mathbb{R}^d \to \overline{\mathbb{R}}$ *absolut integrierbar. Ferner gelte*

  (i) $f_n(x) \to f(x)$ *für alle* $x \in \mathbb{R}^d$,

  (ii) $|f_n| \le g$ *auf* $\mathbb{R}^d$.

  Dann ist $f : \mathbb{R}^d \to \mathbb{R}$ *absolut integrierbar, und es gilt*

$$\int f d\lambda^d = \lim_{n\to\infty} \int f_n d\lambda^d .$$

- *(Satz von Fubini für Lebesgue-Integral)*
Es sei $f : \mathbb{R}^{d+m} \to \overline{\mathbb{R}}$ *messbar. Ist* f *absolut* λ^{d+m} *integrierbar, so existiert eine Menge* $N \in \mathbb{B}^m$ *mit* $\lambda^m(N) = 0$ *und so, dass*

$$x \longmapsto f(x, y)$$

  für alle $y \in \mathbb{R}^m \setminus N$ *absolut* λ^d*-integrierbar ist. Außerdem gilt dann*

$$\int f d\lambda^{d+m} = \int \int f(x, y) d\lambda^d(x) d\lambda^m(y)$$

  (wobei wir für $y \in N$ *etwa* $\int f(x, y) d\lambda^d(x) := 0$ *setzen). Entsprechend existiert eine Menge* $M \in \mathbb{B}^d$ *mit* $\lambda^d(M) = 0$ *und so, dass*

$$y \longmapsto f(x, y)$$

  für alle $x \in \mathbb{R}^d \setminus M$ *absolut* λ^d*-integrierbar ist, und dass gilt*

$$\int f d\lambda^{d+m} = \int \int f(x, y) d\lambda^m(y) d\lambda^d(x)$$

  (wobei $\int f(x, y) d\lambda^m(y) := 0$ *für* $x \in M$*).*

Oft beweist man Ergebnisse für integrierbare Funktionen (wie etwa den Satz von Fubini) nach folgendem Schema: Man zeigt die Behauptung

1. für Indikatorfunktionen f, also $f = 1_A$, wobei $A \in \mathbb{B}^d$ (hier ist $\int f d\lambda = \lambda(A)$),
2. für Elementarfunktionen durch Summieren,
3. für nichtnegative messbare Funktionen durch Supremumbildung und
4. für integrierbare Funktionen durch Anwendung auf f^+, f^-.

Dieses Vorgehen heißt Standardschluss (der Integrationstheorie). Wir beweisen so auch folgenden zentralen Satz.

Satz 27.7 *(Substitutionsregel)*
Es seien $U, V \subset \mathbb{R}^d$ offen, und es sei $\varphi : U \to V$ bijektiv mit $\varphi \in C^1(U, \mathbb{R}^d)$ und $\det J_\varphi(x) \neq 0$ $(x \in U)$. Ist $f : \mathbb{R}^d \to \mathbb{R}$ messbar, so ist f genau dann Lebesgue-integrierbar auf V, wenn $(f \circ \varphi)\,|\det J_\varphi|$ auf U integrierbar ist, und in diesem Falle gilt

$$\int_V f d\lambda = \int_U (f \circ \varphi)\,|\det J_\varphi|\, d\lambda\,.$$

Bevor wir zum eigentlichen Beweis der Substitutionsregel kommen, beweisen wir ein Hilfsresultat, das auch für sich genommen interessant ist.

Satz 27.8 *Es sei $A \subset \mathbb{R}^d$ eine Borelmenge. Dann gilt*

1. $\lambda^d(A) = \inf\{\lambda^d(U) : U \supset A \text{ offen}\}$
2. $\lambda^d(A) = \sup\{\lambda^d(K) : K \subset A \text{ kompakt}\}$.

Beweis. 1. Es sei $\varepsilon > 0$ gegeben. Aus der Definition von λ (vgl. B. 27.3) folgt die Existenz *offener* Intervalle I_k $(k \in \mathbb{N})$ mit $\bigcup_{k \in N} I_k \supset A$ und

$$\sum_{k=1}^{\infty} \lambda(I_k) < \lambda(A) + \varepsilon$$

([Ü]). Für die offene Menge $U := \bigcup_{k \in N} I_k$ gilt dann: $A \subset U$ und

$$\lambda(A) \leq \lambda(U) \leq \sum_{I}^{\infty} \lambda(I_k) \leq \lambda(A) + \varepsilon\,.$$

Da $\varepsilon > 0$ beliebig war, folgt 1.

2. Zunächst betrachten wir den Spezialfall einer beschränkten Borelmenge A. Wir wählen ein offenes Intervall $I \subset \mathbb{R}^d$ mit $\overline{A} \subset I$. Dann existiert nach 1. zu jedem $\varepsilon > 0$ eine offene Menge $V = V_\varepsilon$ mit $I \setminus A \subset V \subset I$ und

$$\lambda(V) < \lambda(I \setminus A) + \varepsilon\,.$$

Da A eine Borelmenge ist, gilt

$$\lambda(I) = \lambda(I \cap A) + \lambda(I \setminus A) = \lambda(A) + \lambda(I \setminus A)\,,$$

also

$$\lambda(I \setminus V) = \lambda(I) - \lambda(V) \geq \lambda(I) - \lambda(I \setminus A) - \varepsilon = \lambda(A) - \varepsilon\,.$$

Es sei $K := I \setminus V$. Dann ist nach obiger Konstruktion K abgeschlossen und beschränkt, also kompakt in $\mathbb{R}^d$. Da $\varepsilon > 0$ beliebig war, folgt 2. (für A beschränkt).
Es sei nun $A \in \mathbb{B}^d$ beliebig. Dann ist mit

$$A_n := A \cap \{x : n-1 \leq |x| < n\} \qquad (n \in \mathbb{N})$$

offenbar $A_m \cap A_n = \emptyset$ $(m \neq n)$ und

$$A = \bigcup_{n \in \mathbb{N}} A_n\,.$$

Also ist

$$\lambda(A) = \sum_{n \in \mathbb{N}} \lambda(A_n)\,.$$

Es sei $\varepsilon > 0$ gegeben. Dann existiert in $N \in \mathbb{N}$ mit

$$\sum_{n=1}^{N} \lambda(A_n) > \begin{cases} \lambda(A) - \varepsilon & , \text{ falls } \lambda(A) < \infty \\ 1/\varepsilon & , \text{ falls } \lambda(A) = \infty \end{cases}.$$

Da A_n beschränkt ist, existiert eine kompakte Menge $K \subset A_n$ mit $\lambda(K_n) \geq \lambda(A_n) - \varepsilon/2^n$. Also folgt für $K := \bigcup_{n=1}^{N} K_n$ (da die K_n paarweise disjunkt sind)

$$\lambda(K) = \sum_{n=1}^{N} \lambda(K_n) \geq \sum_{n=1}^{N} \lambda(A_n) - \varepsilon \sum_{n=1}^{N} 1/2^n \geq \begin{cases} \lambda(A) - 2\varepsilon & , \text{ falls } \lambda(A) < \infty \\ 1/\varepsilon - \varepsilon & , \text{ falls } \lambda(A) = \infty \end{cases}.$$

Da $\varepsilon > 0$ beliebig war, folgt die Behauptung. □

Beweis zu S. 24.7.
1. Es sei für $\xi = (\xi_1, ..., \xi_d) \in \mathbb{R}^d$ und $h > 0$

$$Q := Q_{\xi,h} := (\xi_1 - h, \xi_1 + h] \times \ldots \times (\xi_d - h, \xi_d]$$

der „halboffene Würfel mit Zentrum ξ" (und Kantenlänge $2h$). Ferner setzen wir für $\xi \in U$

$$T_\xi(x) := \varphi(\xi) + J_\varphi(\xi)(x-\xi) \qquad (x \in \mathbb{R}^d)\,.$$

Wir zeigen: Für alle kompakte Teilmengen K von U und alle $\varepsilon > 0$ existiert ein $\delta = \delta(K,\varepsilon) > 0$ so, dass für jeden Würfel $Q = Q_{\xi,h} \subset U$ mit Zentrum $\xi \in K$ und Durchmesser $< 2\delta$ (d.h. $h\sqrt{d} < \delta$) gilt

$$\varphi(Q) \subset T_\xi(Q_{\xi,(1+\varepsilon)h})\,.$$

Beweis dazu: Es seien $K \subset U$ kompakt und $\varepsilon > 0$. Nach S. 21.3 (Hauptsatz über Umkehrfunktionen) ist auch φ^{-1} stetig differenzierbar, und es gilt

$$J_{\varphi^{-1}}(\varphi(x)) = (J_\varphi(x))^{-1} \qquad (x \in U)\,.$$

Da K kompakt ist, existiert

$$M := \max_{\xi \in K} \|\, (J_\varphi(\xi))^{-1} \,\| \;.$$

Wir zeigen: Zu $\eta := \frac{\varepsilon}{M\sqrt{d}}$ existiert ein $\delta > 0$ so, dass für alle (x,ξ) mit $\xi \in K$ und $x \in U_\delta(\xi)$ gilt

$$|\varphi(x) - T_\xi(x)| < \eta|x-\xi|\,.$$

(Denn: Zunächst existiert eine beschränkte offene Menge V mit

$$K \subset V \subset \overline{V} \subset U$$

(vgl. Beweis zu S. 23.5). Für $\rho > 0$ genügend klein gilt dann

$$U_\rho(\xi) \subset V$$

für alle $\xi \in K$ (da $\operatorname{dist}(K, V^c) > 0$). Nach S. 19.13 ist für jedes $j \in \{1,...,d\}$ die Funktion $\operatorname{grad}\varphi_j$ stetig auf U, also gleichmäßig stetig auf der kompakten Menge $\overline{V} \subset U$. Folglich existiert ein $\delta \in (0,\rho)$ so, dass

$$|\operatorname{grad}\varphi_j(\tau) - \operatorname{grad}\varphi_j(\xi)| < \frac{\eta}{d}$$

für alle $\xi \in K, \tau \in U_\delta(\xi)$ und alle $j = 1,...,d$.
Aus dem Mittelwertsatz (S. 20.4) folgt für alle $\xi \in K$ und $x \in U_\delta(\xi)$ mit gewissen $\tau_j \in U_\delta(\xi)$

$$\begin{aligned}
|\varphi(x) - T_\xi(x)| &\le \sum_{j=1}^d |\varphi_j(x) - (T_\xi)_j(x)| \\
&= \sum_{j=1}^d |\varphi_j(x) - \varphi_j(\xi) - \operatorname{grad}\varphi_j(\xi)(x-\xi)| \\
&= \sum_{j=1}^d |(\operatorname{grad}\varphi_j(\tau_j) - \operatorname{grad}\varphi_j(\xi))(x-\xi)| \\
&\le \sum_{j=1}^d \frac{\eta}{d}\cdot|x-\xi| = \eta|x-\xi|\,.)
\end{aligned}$$

Da $T_\xi^{-1}y - T_\xi^{-1}z = (J_\varphi(\xi))^{-1}(y-z)$ für alle $\xi \in K$ und $y, z \in \mathbb{R}^d$ gilt, folgt hieraus für alle $\xi \in K$ und $x \in U_\delta(\xi)$

$$|T_\xi^{-1}(\varphi(x)) - x| = |T_\xi^{-1}(\varphi(x)) - T_\xi^{-1}(T_\xi(x))| \le M|\varphi(x) - T_\xi(x)| \le M\eta|x - \xi|\,.$$

Ist also $x \in Q = Q_{\xi,h}$ mit $h\sqrt{d} < \delta$, so ist $|x - \xi| < h\sqrt{d} < \delta$ und damit

$$|T_\xi^{-1}(\varphi(x)) - x| < \varepsilon h\,.$$

Folglich ist

$$T_\xi^{-1}\varphi(x) \in Q_{\xi,(1+\varepsilon)h}$$

bzw.

$$\varphi(x) \in T_\xi(Q_{\xi,(1+\varepsilon)h})\,.$$

2. Wir zeigen: Ist $A \in B^d, A \subset U$, so gilt

$$\lambda(\varphi(A)) \le \int_A |\det J_\varphi| d\lambda\,.$$

Beweis dazu: Es sei B eine kompakte Teilmenge von $\varphi(A)$. Nach S. 27.8 genügt es, zu zeigen:

$$\lambda(B) \le \int_A |\det J_\varphi| d\lambda\,.$$

Wir setzen $K := \varphi^{-1}(B)$. Dann ist $K \subset A \subset U$ kompakt (S. 11.9). Für $j \in \mathbb{N}$ betrachten wir

$$K_j := \Big\{x : \operatorname{dist}(x, K) \le \frac{1}{j}\Big\} = \bigcup_{\xi \in K} \overline{U_{1/j}(\xi)}\,.$$

Dann ist K_j kompakt für alle j (warum ?) und $K_j \subset U$ für $j \ge j_0$. Außerdem ist $K_j \searrow$ und $\bigcap\limits_{j\in\mathbb{N}} K_j = K$.
Weiter ist $\det J_\varphi$ stetig auf U ([Ü]). Also ist für $j \ge j_0$

$$\int_{K_j} |\det J_\varphi| d\lambda = \max_{K_j} |\det J_\varphi| \cdot \lambda(K_j) < \infty\,.$$

Aus dem Satz von der monotonen Konvergenz (gilt auch für fallende Folgen g_n mit $\int g_n d\lambda < \infty$) ergibt sich

$$\int_K |\det J_\varphi| d\lambda = \lim_{j\to\infty} \int_{K_j} |\det J_\varphi| d\lambda\,.$$

Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $j = j(\varepsilon) \ge j_0$ mit

$$\int_{K_j} |\det J_\varphi| d\lambda < \int_K |\det J_\varphi| d\lambda + \varepsilon\,.$$

Da $\det J_\varphi$ stetig auf der kompakten Menge K_j ist, ist $\det J_\varphi$ gleichmäßig stetig auf K_j. Also existiert ein $\delta = \delta_\varepsilon > 0$ mit

$$|\det J_\varphi(x) - \det J_\varphi(y)| < \varepsilon \qquad \text{für alle} \quad x, y \in K_j,\ |x-y| < \delta\,.$$

O.E. sei $\delta < \delta(K_j, \varepsilon)$, wobei $\delta(K_j, \varepsilon)$ wie in 1., und $\delta < (2j)^{-1}$ (dann ist jeder Würfel mit Durchmesser $< 2\delta$, der K schneidet, in K_j enthalten).
Da K beschränkt ist, existieren endlich viele (halboffene) Würfel $Q_1, ..., Q_N$ mit folgenden Eigenschaften

(i) $K \subset \bigcup\limits_{m=1}^{N} Q_m$,

(ii) $Q_m \cap K \neq \emptyset$ für alle $m \in \{1, ..., N\}$,

(iii) Durchmesser $Q_m < 2\delta$ für alle $m \in \{1, ..., N\}$,

(iv) $Q_m \cap Q_{m'} = \emptyset$ für alle $m, m' \in \{1, ..., N\}, m \neq m'$ (hierfür haben wir „halboffene“ Würfel gewählt).

Dann gilt $B = \varphi(K) \overset{(i)}{\subset} \varphi(\bigcup_{m=1}^{N} Q_m)$. Ist $Q_m = Q_{\xi_m, h_m}$, so gilt $h_m\sqrt{d} < \delta$, also $Q_m \subset K_j$, und

$$\begin{aligned}
\lambda(B) &\leq \lambda\Big(\varphi\big(\bigcup_{m=1}^{N} Q_m\big)\Big) = \sum_{m=1}^{N} \lambda(\varphi(Q_m)) \\
&\overset{1.}{\leq} \sum_{m=1}^{N} \lambda(T_{\xi_m}(Q_{\xi_m,(1+\varepsilon)h_m})) \\
&\overset{B.27.3}{\leq} \sum_{m=1}^{N} |\det J_\varphi(\xi_m)| \lambda(Q_{\xi_m,(1+\varepsilon)h_m}) \\
&= \sum_{m=1}^{N} |\det J_\varphi(\xi_m)| (1+\varepsilon)^d \lambda(Q_m) \\
&\leq (1+\varepsilon)^d \sum_{m=1}^{N} \int\limits_{Q_m} (|\det J_\varphi(x)| + \varepsilon) d\lambda(x) \\
&\overset{(iv)}{=} (1+\varepsilon)^d \int\limits_{\bigcup_1^N Q_m} (|\det J_\varphi(x)| + \varepsilon) d\lambda(x) \\
&\leq (1+\varepsilon)^d \Big(\int\limits_{K_j} |\det J_\varphi(x)| d\lambda(x) + \varepsilon \lambda(K_j) \Big)\,.
\end{aligned}$$

Da $\int_{K_j} |\det J_\varphi(x)| d\lambda(x) \leq \int_K |\det J_\varphi(x)| d\lambda(x) + \varepsilon$ gilt, erhalten wir

$$\lambda(B) \leq (1+\varepsilon)^d \Big[\int\limits_K |\det J_\varphi(x)| d\lambda(x) + \varepsilon(\lambda(K_{j_0}) + 1) \Big]\,.$$

Da $\varepsilon > 0$ beliebig war, folgt

$$\lambda(B) \leq \int_K |\det J_\varphi| d\lambda \leq \int_A |\det J_\varphi| d\lambda \,.$$

3. Mit Hilfe des „Standardschlusses der Integrationstheorie“ (s.o.) erhalten wir aus 2. jedenfalls für alle messbaren nichtnegativen Funktionen f

$$\int_V f d\lambda \leq \int_U (f \circ \varphi) |\det J_\varphi| d\lambda$$

(man beachte dabei: die Abschätzung aus 2. ist der Spezialfall $f \circ \varphi = 1_A$ bzw. $f = 1_{\varphi(A)}$).
4. Wir wenden die Abschätzung aus 3. mit φ^{-1} an Stelle von φ und mit $(f \circ \varphi)|\det J_\varphi|$ an Stelle von f an. Dann gilt

$$\begin{aligned} \int_U (f \circ \varphi) |\det J_\varphi| d\lambda &\leq \int_V (f \circ \varphi \circ \varphi^{-1})(|\det J_\varphi| \circ \varphi^{-1}) |\det J_{\varphi^{-1}}| d\lambda \\ &= \int_V f |\det J_{\varphi \circ \varphi^{-1}}| d\lambda = \int_V f d\lambda \end{aligned}$$

(man beachte: mit der Kettenregel gilt

$$J_{\varphi \circ \varphi^{-1}} = (J_\varphi \circ \varphi^{-1}) J_{\varphi^{-1}} \,,$$

also

$$\det J_{\varphi \circ \varphi^{-1}} = \det(J_\varphi \circ \varphi^{-1}) \cdot \det J_{\varphi^{-1}} \,.)$$

Folglich ist

$$\int_V f d\lambda = \int_U (f \circ \varphi) |\det J_\varphi| d\lambda$$

für alle messbaren $f : V \to [0, \infty)$. Durch Anwendung dieser Identität auf f^+ und f^- erhalten wir schließlich die Aussage des Satzes (beachte: $(f \circ \varphi)^+ = f^+ \circ \varphi$, $(f \circ \varphi)^- = f^- \circ \varphi$). □

In Anwendungen der Substitutionsregel treten Transformationen mit Polarkoordinaten besonders häufig auf.

Beispiel 27.9 (ebene Polarkoordinaten, vgl. B. 21.4).
Wir betrachten die Funktion $f : U \to V$, definiert durch

$$g(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} \qquad ((r, \varphi) \in U) \,,$$

wobei $U = (0, \infty) \times (0, 2\pi)$ und $V = \mathbb{R}^2 \setminus ([0, \infty) \times \{0\})$.
Dann ist $g : U \to V$ bijektiv und $g \in C^1(U)$ mit

$$\det J_g(r, \varphi) = r > 0$$

(siehe B. 21.4). Also erhalten wir mit S. 27.7 für alle messbaren $f : \mathbb{R}^2 \to \mathbb{R}$:
f ist genau dann integrierbar auf V, wenn $(r, \varphi) \to r f(g(r, \varphi))$ integrierbar auf U ist, und dann gilt

$$\int_V f(x, y) d\lambda^2(x, y) = \int_U f(r\cos\varphi, r\sin\varphi)\, r\, d\lambda^2(r, \varphi)\,.$$

Da $\lambda^2(\mathbb{R}^2 \setminus V) = \lambda^2([0, \infty) \times \{0\}) = 0$ und $\lambda^2(([0, \infty) \times [0, 2\pi]) \setminus U) = \lambda^1(\partial U) = 0$ gilt, folgt dann auch (mit dem Satz von Fubini)

$$\int f d\lambda^2 = \int_{[0,\infty]\times[0,2\pi]} f(r\cos\varphi, r\sin\varphi)\, r\, d\lambda^2(r, \varphi) = \int_0^{2\pi}\int_0^{\infty} f(r\cos\varphi, r\sin\varphi)\, r\, dr d\varphi\,.$$

Ist speziell $f = 1_{U_R(0)}$, so gilt

$$f(g(r, \varphi)) = \begin{cases} 1, & \text{falls} \quad 0 \le r < R \\ 0, & \text{sonst.} \end{cases} ,$$

also

$$\lambda^2(U_R(0)) = \int f d\lambda^2 = \int_0^{2\pi}\int_0^{R} r dr d\varphi = \int_0^{2\pi} \frac{r^2}{2}\Big|_0^R d\varphi = 2\pi \frac{R^2}{2} = \pi R^2\,.$$

Für Mengen $M \subset \mathbb{R}^2$ der Form

$$M = \{r\cos\varphi, r\sin\varphi) : 0 \le r \le \rho(\varphi),\ \varphi \in [0, 2\pi]\}\,,$$

mit einer (messbaren) Funktion $\rho : [0, 2\pi] \to [0, \infty)$ ergibt sich allgemeiner

$$\lambda^2(M) = \int 1_M d\lambda^2 = \int_0^{2\pi}\int_0^{\rho(\varphi)} r dr d\varphi = \frac{1}{2}\int_0^{2\pi} \rho^2(\varphi) d\varphi$$

(ist etwa $M = S_{\alpha,\beta} = \{(\cos\varphi, \sin\varphi) : \alpha \le \varphi \le \beta\}$ ein Kreissektor, so ist

$$\rho(\varphi) = \begin{cases} 1, & \alpha \le \varphi \le \beta \\ 0, & \text{sonst.} \end{cases}$$

also

$$\lambda^2(S_{\alpha,\beta}) = \frac{1}{2}\int_\alpha^\beta d\varphi = \frac{1}{2}(\beta - \alpha)\,.)$$

Beispiel 27.10 (Sphärische Polarkoordinaten)
Wir betrachten die Abbildung $g : U \to V$ mit

$$g(r,\varphi,\vartheta) = \begin{pmatrix} r\cos\varphi\cos\vartheta \\ r\sin\varphi\cos\vartheta \\ r\sin\vartheta \end{pmatrix} \qquad ((r,\varphi,\vartheta) \in U)\,,$$

wobei

$$U = (0,\infty) \times (0,2\pi) \times (-\pi/2, \pi/2)$$

und

$$V = g(U) = \mathbb{R}^3 \setminus ([0,\infty) \times \{0\} \times \mathbb{R})\,.$$

Man kann dafür zeigen ([Ü]): Es ist $g \in C^1(U,\mathbb{R}^3)$ und $g : U \to V$ bijektiv mit

$$|\det J_g(r,\varphi,\vartheta)| = r^2\cos\vartheta \qquad ((r,\varphi,\vartheta) \in U)$$

(Entwicklung nach der letzten Zeile). Also erhalten wir mit S. 27.7: Ist $f : \mathbb{R}^3 \to \mathbb{R}$ messbar, so ist f genau dann integrierbar auf V, wenn

$$(r,\varphi,\vartheta) \longmapsto f(g(r,\varphi,\vartheta))r^2\cos\vartheta$$

auf U integrierbar ist, und dann gilt

$$\int_V f d\lambda^2 = \int_U f(g(r,\varphi,\vartheta))\, r^2\cos\vartheta\, d\lambda^3(r,\varphi,\vartheta)\,.$$

Da wieder $\lambda^3(\mathbb{R}^3 \setminus V) = \lambda^3(\partial U) = 0$ ist, ergibt sich dann auch

$$\begin{aligned} \int f d\lambda^3 &= \int_{[0,\infty)\times[0,2\pi]\times[-\frac{-\pi}{2},\frac{\pi}{2}]} f(g(r,\varphi,\vartheta)\, r^2\cos\vartheta\, d\lambda^3(r,\varphi,\vartheta) \\ &\overset{\text{Fubini}}{=} \int_0^{2\pi}\int_{-\pi/2}^{\pi/2}\int_0^{\infty} f(g(r,\varphi,\vartheta))\, r^2\cos\vartheta\, dr d\vartheta d\varphi \end{aligned}$$

Ist etwa $f = 1_{U_R(0)}$ (in $\mathbb{R}^3$), so gilt

$$f(g(r,\varphi,\vartheta)) = \begin{cases} 1, & \text{falls} \quad 0 \le r \le R \\ 0, & \text{sonst.} \end{cases}\,,$$

also

$$\begin{aligned} \lambda^3(U_R(0)) &= \int f d\lambda^3 = \int_0^{2\pi}\int_{-\pi/2}^{\pi/2}\int_0^{R} r^2\cos\vartheta\, dr d\vartheta d\varphi = \\ &= \frac{R^3}{3}\int_0^{2\pi}\underbrace{\int_{-\pi/2}^{\pi/2}\cos\vartheta\, d\vartheta}_{=2}\, d\varphi = \frac{4\pi R^3}{3}\,. \end{aligned}$$

Manchmal lassen sich über den Umweg mehrdimensionaler Integration auch eindimensionale Integrale elegant berechnen. Wir betrachten auch hierzu ein Beispiel.

Beispiel 27.11 Die Beta-Funktion $B : (0,\infty) \times (0,\infty) \to \mathbb{R}$ ist definiert durch

$$B(p,q) := \int_{0+}^{1-} t^{p-1}(1-t)^{q-1}dt \qquad (p,q>0)$$

(man beachte: das uneigentliche Integral existiert). Substituiert man $t = \cos^2\varphi$, so folgt

$$B(p,q) = 2\int_0^{\pi/2} \cos^{2p-1}(\varphi)\sin^{2q-1}(\varphi)d\varphi\,.$$

Weiter gilt für $m > 0$

$$\int_0^\infty r^{2m-1}e^{-r^2}dr = \frac{1}{2}\int_0^\infty t^{m-1}e^{-t}dt = \frac{1}{2}\Gamma(m)\,,$$

wobei Γ die Eulersche Gammafunktion bezeichnet. Also ist

$$\begin{aligned}
\frac{1}{2}B(p,q)\Gamma(p+q) &= \int_0^\infty B(p,q)r^{2p+2q-1}e^{-r^2}dr \\
&= 2\int_0^\infty\int_0^{\pi/2}(r\cos\varphi)^{2p-1}(r\sin\varphi)^{2q-1}e^{-r^2}r\,d\varphi dr \\
&\overset{B.27.9}{=} 2\int_{(0,\infty)\times(0,\infty)} x^{2p-1}y^{2p-1}e^{-x^2-y^2}d\lambda^2(x,y) \\
&\overset{\text{Fubini}}{=} 2\int_0^\infty x^{2p-1}e^{-x^2}d\lambda(x)\int_0^\infty y^{2p-1}e^{-y^2}d\lambda(y) \\
&= \frac{1}{2}\Gamma(p)\Gamma(q)\,,
\end{aligned}$$

d.h.

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}\,.$$

28 Fourier-Reihen

In den Abschnitten 14 und 16 hatten wir uns mit Taylor-Reihen und Potenzreihen beschäftigt. Wir hatten gesehen, dass unter geeigneten Voraussetzungen an die Ableitungen $f^{(n)}$ eine Funktion f einer reellen Variablen durch ihre Taylor-Reihe dargestellt werden kann, genauer

$$f(x) = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(x_0)}{\nu!} (x - x_0)^\nu$$

auf einem gewissen Intervall (I) (siehe etwa S. 16.6). Wir wollen nun eine andere Art einer Reihenentwicklung untersuchen, die den wesentlichen Vorteil hat, dass keine Ableitungen von f benötigt werden. Dazu stellen wir zunächst einige Vorüberlegungen an.

Bemerkung 28.1 Wir betrachten den linearen Raum

$$C_{2\pi} := \{f \in C([-\pi, \pi], \mathbb{C}) : f(-\pi) = f(\pi)\} ,$$

versehen mit dem Skalarprodukt

$$< f, g > := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\overline{g}(t)dt \qquad (f, g \in C_{2\pi}) .$$

(Dann ist bekanntlich ($\rightarrow$ lineare Algebra) durch

$$||f||_2 := \sqrt{< f, f >} = \Big(\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|^2 dt\Big)^{1/2} \qquad (f \in C_{2\pi})$$

eine Norm auf $C_{2\pi}$ definiert.)
Ist

$$f_n(t) := e^{int} \qquad (t \in [-\pi, \pi], n \in \mathbb{Z}) ,$$

so gilt damit

$$< f_n, f_m > = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)t} dt = \begin{cases} \frac{1}{2\pi} \frac{1}{i(n-m)} e^{i(n-m)t}\Big|_{-\pi}^{\pi} = 0 & , \quad n \neq m \\ 1 & , \quad n = m \end{cases} ,$$

d.h. $(f_n)_{n\in\mathbb{Z}}$ ist ein Orthogonalsystem in $(C_{2\pi}, < ., . >)$.
Ist J_n die Menge der „trigonometrischen Polynome vom Grad $\leq n$“, d.h.

$$J_n := \operatorname{lin\,span}\{f_\nu : \nu = -n, ..., 0, ..., n\} \subset C_{2\pi} ,$$

so gilt nach dem Projektionssatz ($\rightarrow$ Lineare Algebra) für alle $f \in C_{2\pi}$

$$||f - \sum_{\nu=-n}^{n} < f, f_\nu > \; f_\nu||_2 = \operatorname{dist}(f, J_n) (= \min_{g\in J_n} ||f - g||_2) .$$

Definiert man

$$a_\nu := < f, f_\nu > = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-i\nu t}dt \qquad (\nu \in \mathbb{Z})$$

und

$$S_n(f)(t) := \sum_{\nu=-n}^{n} a_\nu e^{i\nu t} \qquad (t \in [-\pi, \pi],\, n \in \mathbb{N}_0)\,,$$

so ist $S_n(f) \in C_{2\pi}$ die beste Näherung an f aus J_n bzgl. der 2-Norm. Außerdem impliziert die Besselsche Ungleichung (→ Lineare Algebra) hier: $\sum\limits_{\nu=-n}^{n} |c_\nu|^2 \leq ||f||_2^2$ für alle $n \in \mathbb{N}_0$.

Definition 28.2 Es sei $f : [-\pi, \pi] \to \mathbb{C}$ eine Regelfunktion. Dann heißt für $\nu \in \mathbb{Z}$

$$a_\nu := \hat{f}(\nu) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-i\nu t}dt$$

ν-ter Fourier-Koeffizient von f. Weiter heißt für $n \in \mathbb{N}_0$

$$S_n(f)(t) := \sum_{\nu=-n}^{n} a_\nu e^{i\nu t} = \sum_{\nu=-n}^{n} \hat{f}(\nu)e^{i\nu t}$$

n-te Fourier-Teilsumme von f und die (formale) Reihe

$$\sum_{\nu=-\infty}^{\infty} a_\nu e^{i\nu t} = \sum_{\nu=-\infty}^{\infty} \hat{f}(\nu)e^{i\nu t}\,,$$

also die Folge der Teilsummen $(S_n(f))_{n\in\mathbb{N}_0})$, die *Fourier-Reihe* von f.

Bemerkung 28.3 Man beachte, dass für $f \in R[-\pi, \pi]$ und $\nu \in \mathbb{Z}$ auch die Funktion $t \mapsto f(t)e^{-i\nu t}$ in $R[-\pi, \pi]$ liegt, also existieren die Fourier-Koeffizienten für alle Regelfunktionen.

Beispiel 28.4 Wir betrachten

$$f(t) := |t| \qquad (t \in [-\pi, \pi])\,.$$

Dann gilt für $\nu \neq 0$

$$2\pi a_\nu = \int_{-\pi}^{\pi} |t|\cos(-\nu t)\,dt + \underbrace{i\int_{-\pi}^{\pi} |t|\sin(-\nu t)\,dt}_{=0} = 2\int_{0}^{\pi} t\cos(\nu t)\,dt = \frac{2}{\nu^2}((-1)^\nu - 1)$$

also

$$a_0 = \frac{\pi}{2}, \qquad a_{2k+1} = -\frac{2}{\pi(2k+1)^2} \quad (k \in \mathbb{Z}), \qquad a_{2k} = 0 \quad (k \in \mathbb{Z},\, k \neq 0).$$

Folglich ist

$$\begin{aligned}
\sum_{\nu=-\infty}^{\infty} a_\nu e^{i\nu t} &= \frac{\pi}{2} - \frac{2}{\pi} \sum_{k=-\infty}^{\infty} \frac{e^{i(2k+1)t}}{(2k+1)^2} = \\
&= \frac{\pi}{2} - \frac{2}{\pi} \left[\sum_{k=0}^{\infty} \frac{\cos((2k+1)t)}{(2k+1)^2} + \sum_{k=1}^{\infty} \frac{\cos((-2k+1)t)}{(-2k+1)} \right] \\
&\quad -\frac{2}{\pi}\, i \underbrace{\left(\sum_{k=0}^{\infty} \frac{\sin((2k+1)t)}{(2k+1)^2} + \sum_{k=1}^{\infty} \frac{\sin((-2k+1)t)}{(-2k+1)^2} \right)}_{=0} \\
&= \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\cos((2k-1)t)}{(2k-1)^2}.
\end{aligned}$$

(Da $\sum\limits_{k=1}^{\infty} 1/(2k-1)^2$ konvergiert, ist die Fourier-Reihe (absolut und) gleichmäßig konvergent auf $[-\pi, \pi]$.)

Bisher wissen wir nichts darüber, unter welchen Bedingungen die Fourierreihe die Funktion f darstellt, d.h. wann (und in welchem Sinne)

$$f = \lim_{n\to\infty} S_n(f)$$

gilt. Wenn wir etwa nach Konvergenz bzgl. $||\cdot||_2$ fragen, so zeigen die Überlegungen aus B. 28.1, dass

$$||f - S_n(f)||_2 \to 0 \qquad (n \to \infty)$$

schon dann gilt, wenn nur $\operatorname{dist}(f, J_n) \to 0$ erfüllt ist, m.a.W., wenn $\bigcup\limits_{n\in\mathbb{N}} J_n$ (die Menge der trigonometrischen Polynome) dicht in $(C_{2\pi}, ||\cdot||_2)$ ist. Da für alle $f \in C_{2\pi}$

$$||f||_2 = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |f|^2 \right)^{1/2} \leq ||f||_\infty \Big(= \max_{[-\pi,\pi]} |f(t)| \Big)$$

gilt, reicht es dafür zu zeigen, dass $\bigcup\limits_{n\in\mathbb{N}} J_n$ dicht in $(C_{2\pi}, ||\cdot||_\infty)$ ist.

Satz 28.5 *(Weierstraßscher Approximationssatz für trigonometrische Approximation)*
Es sei $f \in C_{2\pi}$. Dann existiert zu jedem $\varepsilon > 0$ ein trigonometrisches Polynom P (also $P \in \bigcup\limits_{n\in\mathbb{N}} J_n$) mit

$$||f - P||_\infty = \max_{t\in[-\pi,\pi]} |f(t) - P(t)| < \varepsilon$$

m.a.W. $J := \bigcup\limits_{n\in\mathbb{N}} J_n$ ist dicht in $(C_{2\pi}, ||\cdot||_\infty)$.

Beweis.
1. Wir zeigen zunächst: Existiert eine Folge (Q_k) in J, so dass

(i) $Q_k(t) \geq 0 \qquad (t \in [-\pi, \pi], k \in N)$,

(ii) $\frac{1}{2\pi} \int\limits_{-\pi}^{\pi} Q_k(t)dt = 1 \qquad (k \in \mathbb{N})$,

(iii) Für alle $\delta > 0$ gilt $Q_k \to 0$ gleichmäßig auf $[-\pi, -\delta] \cup [\delta, \pi]$,

so gilt die Behauptung.
Denn: Es sei $f \in C_{2\pi}$. Wir definieren eine Folge (P_k) durch

$$P_k(t) := \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} f(t-s)Q_k(s)ds \qquad (k \in \mathbb{N}, t \in [-\pi, \pi])$$

Mit der Substitutionsregel sieht man leicht, dass dann auch

$$P_k(t) = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} f(u)Q_k(t-u)du \qquad (k \in \mathbb{N}, t \in [-\pi, \pi])$$

gilt (wichtig dabei: f, Q_k können 2π-periodisch stetig auf $\mathbb{R}$ fortsetzt werden). Da Q_k ein trigonometrisches Polynom ist, hat Q_k die Form

$$Q_k(t) = \sum_{\nu=-n_k}^{n_k} c_{k\nu} e^{i\nu t}$$

mit gewissen $c_{k\nu} \in \mathbb{C}, n_k \in \mathbb{N}$. Also folgt

$$\begin{aligned} P_k(t) &= \sum_{\nu=-n_k}^{n_k} a_{k\nu} \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} f(s) e^{i\nu(t-s)} ds = \\ &= \sum_{\nu=-n_k}^{n_k} a_{k\nu} \hat{f}(\nu) e^{i\nu t}, \end{aligned}$$

d.h. P_k ist ebenfalls ein trigonometrisches Polynom.
Da f stetig auf dem kompakten Intervall $[-\pi, \pi]$, also gleichmäßig stetig ist, existiert zu jedem $\varepsilon > 0$ ein $\delta = \delta_\varepsilon > 0$

$$|f(t) - f(s)| < \varepsilon \qquad \text{für alle } t, s \text{ mit } |t-s| < \delta\,.$$

Weiter ergibt sich mit (ii)

$$P_k(t) - f(t) = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} \{f(t-s) - f(t)\} Q_k(s) ds \qquad (t \in [-\pi, \pi], k \in \mathbb{N})\,,$$

also mit (i)

$$\begin{aligned} |P_k(t) - f(t)| &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t-s) - f(t)| Q_k(s) ds = \\ &= \frac{1}{2\pi} \Big(\int_{-\delta}^{\delta} + \int_{[-\pi,-\delta]\cup[\delta,\pi]} \Big) \ldots =: I_1(t) + I_2(t) \,. \end{aligned}$$

Dabei gilt

$$I_1(t) \leq \frac{\varepsilon}{2\pi} \int_{-\delta}^{\delta} Q_k(t) dt \leq \varepsilon$$

und

$$I_2(t) \leq 2||f||_\infty \cdot \sup\{Q_k(t) : t \in [-\pi, -\delta] \cup [\delta, \pi]\} \to 0 \qquad (k \to \infty)\,,$$

d.h. es existiert ein $K_\varepsilon \in \mathbb{N}$ mit $I_2(t) < \varepsilon$ für $k \geq K_\varepsilon, t \in [-\pi, \pi]$. Also folgt

$$||P_k - f||_\infty \leq 2\varepsilon \qquad (k \geq K_\varepsilon)\,.$$

Da $\varepsilon > 0$ beliebig war, ergibt sich die Behauptung.
2. Es bleibt zu zeigen, dass eine Folge (Q_k) mit den Eigenschaften (i), (ii), (iii) existiert. Eine mögliche Wahl ist die folgende:

$$Q_k(t) := c_k \Big(\frac{1 + \cos t}{2}\Big)^k,$$

wobei c_k so ist, dass (ii) gilt, d.h.

$$\frac{1}{c_k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\frac{1 + \cos t}{2})^k dt.$$

Dann sind (i) und (ii) erfüllt. Es bleibt noch (iii) zu zeigen. Dazu beachten wir, dass Q_k gerade ist, also impliziert (ii)

$$1 = \frac{c_k}{\pi} \int_0^{\pi} \Big(\frac{1 + \cos t}{2}\Big)^k dt \geq \frac{c_k}{\pi} \int_0^{\pi} \Big(\frac{1 + \cos t}{2}\Big)^k \sin t \, dt = \frac{2c_k}{\pi} \int_0^1 x^k dx = \frac{2c_k}{\pi(k+1)}\,.$$

Da Q_k monoton fallend auf $[0, \pi]$ ist, folgt für $0 < \delta \leq |t| \leq \pi$ (da $c_k \leq \pi(k+1)/2$)

$$Q_k(t) \leq Q_k(\delta) \leq \frac{\pi(k+1)}{2} \Big(\underbrace{\frac{1 + \cos(\delta)}{2}}_{<1}\Big)^k \to 0 \qquad (k \to \infty)\,,$$

also (iii). $\square$

Bemerkung 28.6 Wie oben bereits erläutert, impliziert S. 28.5 insbesondere, dass für alle $f \in C_{2\pi}$ gilt

$$||f - S_n(f)||_2 \to 0 \qquad (n \to \infty)$$

d.h. $S_n(f) \to f$ „im Mittel". Dies bedeutet jedoch noch nicht, dass stets auch

$$S_n(f)(t) \to f(t)$$

für alle $t \in [-\pi, \pi]$ gilt. Wir werden später sehen, dass dies tatsächlich nicht für alle $f \in C_{2\pi}$ erfüllt ist (wir werden sogar sehen, dass $S_n(f)(t)$ noch nicht einmal für alle t konvergiert).

Als weitere Folgerung aus S. 28.5 erhalten wir

Satz 28.7 *Es sei $f \in C_{2\pi}$. Dann gilt*

1. $||f||_2^2 = \sum_{\nu=-\infty}^{\infty} |\hat{f}(\nu)|^2$ *(Parsevalsche Gleichung).*

2. *Gilt $\hat{f}(\nu) = 0$ für alle $\nu \in \mathbb{Z}$, so ist $f \equiv 0$.*

3. *Konvergiert $S_n(f)$ gleichmäßig auf $[-\pi, \pi]$, so gilt*

$$S_n(f) \to f\,.$$

Beweis.
1. Es gilt $S_n(f) \in J_n$ und $f - S_n(f) \in J_n^{\perp}$ ($\to$ Lineare Algebra). Also ergibt sich aus dem Satz von Pythagoras

$$||f||_2^2 \;=\; ||f - S_n(f)||_2^2 + ||S_n(f)||_2^2\,.$$

Nach B. 28.6 gilt $||f - S_n(f)||_2^2 \to 0\,(n \to \infty)$, d.h. $||S_n(f)||_2^2 \to ||f||_2^2\,(n \to \infty)$. Außerdem ist (wieder mit Pythagoras)

$$||S_n(f)||_2^2 \;=\; ||\sum_{\nu=-n}^{n} < f, f_\nu > f_\nu||_2^2 \;=\; \sum_{\nu=-n}^{n} |< f, f_\nu >|^2 \underbrace{||f_\nu||_2^2}_{=1} = \sum_{\nu=-n}^{n} |\hat{f}(\nu)|^2\,.$$

Damit ergibt sich 1.
2. Ist $\hat{f}(\nu) = 0\,(\nu \in \mathbb{Z})$, so ist $||f||_2^2 = 0$ nach 1., also $f \equiv 0$.
3. Nach Voraussetzung ist $g : [-\pi, \pi] \to \mathbb{C}$ mit

$$g(t) := \lim_{n\to\infty} S_n(f)(t) = \sum_{\nu=-\infty}^{\infty} \hat{f}(\nu) e^{i\nu t}$$

stetig auf $[-\pi, \pi]$ (S. 15.10) mit $g(\pi) = g(-\pi)$, also $g \in C_{2\pi}$. Außerdem gilt

$$\hat{g}(k) = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} g(t) e^{-ikt} dt = \sum_{\nu=-\infty}^{\infty} \hat{f}(\nu) \underbrace{\frac{1}{2\pi} \int\limits_{-\pi}^{\pi} e^{i(\nu-k)t} dt}_{=\delta_{\nu k}} = \hat{f}(k)\,,$$

d.h. $\hat{f}(k) - \hat{g}(k) = 0$. Nach 2. ist $f - g \equiv 0$. □

Der folgende Satz liefert ein einfaches hinreichendes Kriterium für die gleichmäßige Konvergenz.

Satz 28.8 *Es sei $f \in C_{2\pi}$ stetig differenzierbar auf $[-\pi, \pi]$. Dann gilt*

$$S_n(f) \to f \qquad \textit{gleichmäßig auf } [-\pi, \pi].$$

Beweis. Für $\nu \in \mathbb{Z}, \nu \neq 0$ gilt (partielle Integration, 2π-Periodizität)

$$\int\limits_{-\pi}^{\pi} f(t) e^{-i\nu t} dt = \frac{1}{i\nu} \int\limits_{-\pi}^{\pi} f'(t) e^{-i\nu t} dt\,,$$

d.h. $\hat{f}(\nu) = \frac{1}{i\nu} \widehat{f'}(\nu)$. Also folgt (mit Cauchy-Schwarzscher Ungleichung)

$$\begin{aligned} \sum_{\nu=-n}^{n} |\hat{f}(\nu)| \; &= \; |\hat{f}(0)| \; + \sum_{\substack{\nu=-n \\ \nu \neq 0}}^{n} \frac{|\widehat{f'}(\nu)|}{\nu} \leq \\ & |\hat{f}(0)| \; + \Big(2 \sum_{\nu=1}^{n} \frac{1}{\nu^2}\Big)^{1/2} \Big(\sum_{\substack{\nu=-n \\ \nu \neq 0}}^{n} |\widehat{f'}(\nu)|^2 \Big)^{1/2}. \end{aligned}$$

Aus der Parsevalschen Gleichung folgt $\sum\limits_{\nu=-\infty}^{\infty} |\widehat{f'}(\nu)|^2 = ||f'||_2^2$, also ist $\sum\limits_{-\infty}^{\infty} |\hat{f}(\nu)|$ konvergent. Nach dem Weierstraßschen Majorantenkriterium ist $\sum\limits_{-\infty}^{\infty} \hat{f}(\nu) e^{i\nu t}$ gleichmäßig konvergent auf $[-\pi, \pi]$ (und nach S. 28.7.3 gegen f). □

Die Frage nach der (punktweisen oder gleichmäßigen) Konvergenz von Fourier-Reihen ist nicht leicht zu beantworten. Es gibt weitere hinreichende Bedingungen, die z. T. wesentlich feiner sind als die aus S. 28.8. Wir wollen hierauf nicht weiter eingehen.
Unser Ziel im weiteren Verlauf ist es, zu zeigen, dass Fourier-Reihen stetiger Funktionen nicht stets (punktweise) konvergieren. Dazu beschäftigen wir uns zunächst mit weiteren allgemeinen Ergebnissen der Analysis, die auch für sich genommen von weitreichendem Interesse sind.

Satz 28.9 *(Baire)*
Es sei (X,d) ein vollständiger metrischer Raum. Ist $(\mathcal{O}_n)_{n\in\mathbb{N}}$ eine Folge offener und dichter Teilmengen von X, so ist auch $\bigcap_{n\in\mathbb{N}} \mathcal{O}_n$ dicht in X (also insbesondere $\neq \emptyset$).

Beweis. Es sei $x_0 \in X$ und $r_0 > 0$ gegeben. Wir haben zu zeigen: Für $U := U_{r_0}(x_0)$ gilt

$$U \cap \bigcap_{n\in\mathbb{N}} \mathcal{O}_n \neq \emptyset .$$

Da $\overline{\mathcal{O}}_1 = X$ gilt, ist $U \cap \mathcal{O}_1 \neq \emptyset$ und offen. Also existieren ein $x_1 \in X$ und ein $r_1 \in (0,1)$ mit

$$\overline{U_{r_1}(x_1)} \subset U \cap \mathcal{O}_1. \tag{$*$}$$

Nun können wir genauso mit $U_{r_1}(x_1)$ statt U und $\mathcal{O}_2$ statt $\mathcal{O}_1$ argumentieren. Induktiv erhält man so eine Folge (x_n) in X und eine Folge (r_n) in $(0,1)$ (o.E. so, dass $r_n < 1/n$) mit

$$\overline{U_{r_n}(x_n)} \subset U_{r_{n-1}}(x_{n-1}) \cap \mathcal{O}_n \qquad (n \in \mathbb{N}) \; . \tag{$**$}$$

Ist $n \in \mathbb{N}$ gegeben, so gilt dabei $x_j \in U_{r_n}(x_n) \subset U_{1/n}(x_n)$ für alle $j > n$, also gilt für $j, k > n$

$$d(x_j, x_k) < \frac{2}{n} .$$

Damit ist (x_n) eine Cauchy-Folge in X, also konvergent. Es sei

$$x := \lim_{n\to\infty} x_n .$$

Ist wieder $n \in \mathbb{N}$ gegeben, so ist $x_j \in \overline{U_{r_n}(x_n)}$ für alle $j > n$, also auch $x \in \overline{U_{r_n}(x_n)}$. Folglich ist $x \in \mathcal{O}_n$ nach $(**)$. Außerdem ist auch $x \in U$ nach $(*)$. □

Wir werden den Satz von Baire verwenden, um einen der grundlegenden Sätze der sog. Funktionalanalysis zu beweisen: den Satz von der gleichmäßigen Beschränktheit. Dabei geht es um lineare Abbildungen, auf die wir erst kurz zu sprechen kommen.

Definition 28.10 Es seien $(E, ||\cdot||)_E)$ und $(F, ||\cdot||_F)$ normierte Räume über $\mathbb{K}$, und es sei $T : E \to F$ linear. Wir setzen

$$||T|| := \sup\{||Tx||_F : x \in E, ||x||_E \leq 1\} .$$

Ist $||T|| < \infty$, so heißt T *beschränkt.* Weiter setzen wir

$$L(E,F) := \{T : E \to F : T \text{ linear und beschränkt}\} .$$

(Man kann zeigen ([Ü]): $||\cdot||$ ist eine Norm auf $L(E,F)$, die sogenannte Operatornorm; vgl. S. 19.11 für den Fall $E = \mathbb{K}^d, F = \mathbb{K}^n$.)

Bemerkung 28.11 Anders als im Fall endlich-dimensionaler Räume E, F sind im unendlich-dimensionalen Fall lineare Abbildungen nicht immer stetig. Man kann zeigen ([Ü]): Sind $(E, ||\cdot||_E)$ und $(F, ||\cdot||_F)$ normierte Räume und ist $T : E \to F$ linear, so sind äquivalent

a) T ist beschänkt,

b) T ist stetig (auf E),

c) T ist stetig an $x_0 = 0$.

Beispiel 28.12 Es sei $E = C[a, b]$ und $F = (\mathbb{K}, |\cdot|)$. Weiter sei für ein $t_0 \in [a, b]$ die Abbildung $T = T_{t_0} : E \to \mathbb{K}$ definiert durch

$$Tf := f(t_0) \qquad (f \in C[a, b])\,.$$

Dann ist T offenbar linear.
Ist $||\cdot||_E = ||\cdot||_\infty$, so gilt für alle f mit $||f||_\infty \leq 1$

$$|Tf| \;=\; |f(t_0)| \leq ||f||_\infty\,,$$

also ist $||T|| \leq 1$. (Ist $f \equiv 1$, so gilt $|Tf| = 1$, also ist genauer $||T|| = 1$.) Damit ist T beschränkt, also auch stetig nach B. 28.11.
Es sei nun $t_0 \in (a, b)$ und $||\cdot||_E = ||\cdot||_1$, wobei

$$||f||_1 := \int_a^b |f(t)|\,dt.$$

Dann gilt für f_n mit

$$f_n(t) = \begin{cases} n - n^2|t - t_0| & , \quad |t - t_0| < 1/n \\ 0 & , \quad |t - t_0| \geq 1/n \end{cases}$$

$f_n \in C[a, b]$ und $\int\limits_a^b f_n = \int\limits_{t_0-1/n}^{t_0+1/n} f_n = 1$ $(n \geq n_0)$, also $||f_n||_1 = 1$ $(n \geq n_0)$. Außerdem ist $Tf_n = n\,(n \in \mathbb{N})$, d.h. $||T|| = \infty$. Also ist T unbeschränkt.

Definition 28.13 Es sei $(E, ||\cdot||_E)$ ein normierter Raum. Ist $(E, d_{||\cdot||})$ vollständig, so heißt $(E, ||\cdot||_E)$ ein *Banachraum*.

Der folgende Satz zeigt, dass eine Familie beschränkter linearer Abbildungen entweder „gleichmäßig beschränkt“ ist oder dass für viele Punkte die Familie „punktweise unbeschränkt“ ist. Also:

Satz 28.14 *(Banach-Steinhaus; Satz von der gleichmäßigen Beschränktheit)*
Es seien $(E, ||\cdot||_E)$ ein Banachraum, $(F, ||\cdot||_F)$ ein normierter Raum und $(T_\iota)_{\iota\in I}$ eine Familie beschränkter linearer Abbildungen $T_\iota : E \to F$. Dann ist entweder

$$\sup_{\iota\in I} ||T_\iota||_F < \infty$$

oder es existiert eine dichte G_δ-Menge $A \subset E$ (G_δ-Mengen sind abzählbare Durchschnitte offener Mengen) mit

$$\sup_{\iota\in I} ||T_\iota x||_F = \infty \qquad (x \in A)\,.$$

Beweis. Wir betrachten $\varphi : E \to [0,\infty]$ mit

$$\varphi(x) := \sup_{\iota\in I} ||T_\iota x||_F \qquad (x \in E)$$

und

$$V_n := \{x \in E : \varphi(x) > n\} \qquad (n \in \mathbb{N})\,.$$

Dann it V_n offen in E.
(Denn: Es sei $x_0 \in V_n$, d.h. $\varphi(x_0) > n$. Dann existiert ein $\iota \in I$ mit $||T_\iota x|| > n$. Da $E \ni x \mapsto ||T_\iota x||_F \in \mathbb{R}$ stetig ist, existiert ein $\delta > 0$ mit $||T_\iota x||_F > n$ für alle $x \in U_\delta(x_0)$. Also ist auch $\varphi(x) > n$ für alle $x \in U_\delta(x_0)$, d.h. $U_\delta(x) \subset V_n$.)

<u>1. Fall:</u> Es existiert ein $N \in \mathbb{N}$ so, dass $\overline{V_N} \neq E$ ist (d.h. V_N ist nicht dicht in E). Dann existieren ein $x_0 \in E$ und ein $r > 0$ mit $\overline{U_r(x_0)} \cap V_n = \emptyset$, d.h. ist $||x||_E \le r$, so ist $x_0 + x \notin V_n$. Folglich ist $\varphi(x_0 + x) \le N$, also

$$||T_\iota(x_0 + x)||_F \le N \qquad (\iota \in I, ||x||_E \le r)\,.$$

Aus $x = (x_0 + x) - x_0$ ergibt sich

$$||T_\iota x||_F \le ||T_\iota(x_0 + x)||_F + ||T_\iota(x_0)||_F \le 2N \qquad (||x||_E \le r, \iota \in I)\,.$$

Hieraus folgt $||T_\iota x||_F \le 2N/r$ für $||x||_E \le 1$ und $\iota \in I$, d.h. $||T_\iota|| \le 2N/r$ für alle $\iota \in I$. Also ist $\sup\limits_{\iota\in I} ||T_\iota|| \le 2N/r < \infty$.
<u>2. Fall:</u> Für alle $n \in \mathbb{N}$ ist V_n dicht in E. Dann ist $\bigcap\limits_{n\in\mathbb{N}} V_n$ nach S. 28.9 eine dichte G_δ-Menge. Dabei ist $\varphi(x) = \infty$ für alle $x \in \bigcap\limits_{n\in\mathbb{N}} V_n$. □

Nun wenden wir uns wieder weniger abstrakten Dingen zu: Es sei $f \in C_{2\pi}$. Dann ist

$$S_n(f)(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)D_n(t-s)\,ds \qquad (t \in [-\pi,\pi], n \in \mathbb{N})\,, \tag{28.1}$$

wobei

$$D_n(t) = \sum_{\nu=-n}^{n} e^{i\nu t}$$

(D_n heißt *Dirichlet-Kern*). Wir verwenden diese Darstellung von $S_n(f)$, um folgendes Ergebnis zu beweisen.

Satz 28.15 *Es sei $t_0 \in [-\pi, \pi]$. Dann existiert eine dichte G_δ-Menge E_{t_0} in $C_{2\pi}$ so, dass*

$$\sup_{n\in\mathbb{N}} |S_n(f)(t_0)| = \infty \qquad (f \in E_{t_0})$$

Insbesondere ist also $(S_n(f)(t_0))_n$ divergent für alle $f \in E_{t_0}$.

Beweis.
1. Wir betrachten o. E. $t_0 = 0$ und definieren $T_n : C_{2\pi} \to \mathbb{C}$ durch

$$T_n(f) := S_n(f)(0) \qquad (f \in C_{2\pi}).$$

Dann ist $T_n : (C_{2\pi}, ||\cdot||_\infty) \to (\mathbb{C}, |\cdot|)$ linear, und es gilt

$$|T_n(f)| = \Big|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)D(-s)ds\Big| \leq ||f||_\infty \cdot \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(s)|\, ds = ||f||_\infty \cdot ||D_n||_1\,,$$

also

$$||T_n|| \leq ||D_n||_1$$

(insbesondere ist T_n beschränkt).
Wir zeigen: $||T_n|| \to \infty$ $(n \to \infty)$, und dazu zeigen wir $||T_n|| = |D_n||_1$ und $||D_n||_1 \to \infty$ $(n \to \infty)$.
Zunächst gilt für $t \in [-\pi, \pi], t \neq 0$

$$\begin{aligned}
2i\sin\left(\frac{t}{2}\right) D_n(t) &= e^{it/2}D_n(t) - e^{-it/2}D_n(t) = \\
&= \sum_{\nu=-n}^{n} e^{i(\nu+1-1/2)t} - \sum_{\nu=-n}^{n} e^{i(\nu-1/2)t} = \\
&= e^{i(n+1/2)t} - e^{i(-n-1/2)t} = 2i\sin\Big((n+\frac{1}{2})t\Big)\,,
\end{aligned}$$

d.h.

$$D_n(t) = \begin{cases} \dfrac{\sin((n+1/2)t)}{\sin(t/2)} &, \quad t \in [-\pi,\pi]\setminus\{0\} \\ 2n+1 &, \quad t = 0 \end{cases}.$$

Aus $|\sin x| \le |x|$ folgt

$$\begin{aligned} ||D_n||_1 &\ge \frac{2}{\pi}\int\limits_{0+}^{\pi} |\sin\Big((n+\frac{1}{2})t\Big)|\frac{dt}{t} = \frac{2}{\pi}\int\limits_{0}^{(n+1/2)\pi} |\sin(u)|\frac{du}{u} \\ &\ge \frac{2}{\pi}\sum_{k=1}^{n}\frac{1}{k\pi}\int\limits_{(k-1)\pi}^{k\pi} |\sin u|du = \frac{4}{\pi^2}\sum_{k=1}^{n}\frac{1}{k} \to \infty \qquad (n\to\infty)\,. \end{aligned}$$

Es bleibt noch $||D_n||_1 \le ||T_n||$ zu zeigen.
Es gilt

$$||D_n||_1 = \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} |D_n(t)|dt = \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} \underbrace{\operatorname{sign}(D_n)(t)}_{=g(t)} D_n(t)\,dt\,.$$

Dabei ist g eine Treppenfunktion, genauer existieren (äquidistante) Punkte $-\pi = t_0 < t_1 < \ldots < t_m = \pi$ mit

$$g(t) \equiv 1 \qquad \text{oder} \qquad g(t) \equiv -1 \qquad \text{auf } (t_{j-1}, t_j)\,.$$

Ist $\varepsilon > 0$ gegeben, so existiert eine Funktion $g_\varepsilon \in C_{2\pi}$ mit $|g_\varepsilon|_\infty \le 1$ und

$$\frac{1}{2\pi}\int\limits_{-\pi}^{\pi} |g(t) - g_\varepsilon(t)|\,dt < \varepsilon\,,$$

(„lineare Abänderung von g in der Nähe der t_j “). Also folgt (man beachte: D_n ist gerade)

$$\begin{aligned} ||T_n|| \ge |T_n(g_\varepsilon)| &= \Big|\frac{1}{2\pi}\int\limits_{-\pi}^{\pi} g_\varepsilon(s)D_n(-s)ds\Big| \\ &= \Big|\frac{1}{2\pi}\int\limits_{-\pi}^{\pi} g(s)D_n(-s)ds - \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} (g_\varepsilon(s) - g(s))D_n(-s)ds\Big| \\ &\ge ||D_n||_1 - \varepsilon||D_n||_\infty\,. \end{aligned}$$

Da $\varepsilon > 0$ beliebig war, ist $||T_n|| \ge ||D_n||_1$.
2. Nach dem Satz von Banach-Steinhaus und 1. existiert eine dichte G_δ-Menge $E_0 \subset C_{2\pi}$ mit

$$\sup_{n\in\mathbb{N}} |T_n f| = \sup_{n\in\mathbb{N}} |S_n(f)(0)| = \infty \qquad (f \in E_0)\,.$$

Offensichtlich kann eine analoge Argumentation für beliebiges $t_0[-\pi,\pi]$ an Stelle von $t_0 = 0$ geführt werden. $\square$

Bemerkung 28.16 1. Man kann S. 28.15 noch dahingehend verschärfen, dass sogar eine dichte G_δ-Menge $E \subset C_{2\pi}$ existiert mit der Eigenschaft, dass für alle $f \in E$ die Menge

$$A_f := \{t \in [-\pi, \pi] : \sup_{n\in\mathbb{N}} |S_n(f)(t)| = \infty\}$$

eine dichte G_δ-Menge in $[-\pi, \pi]$ ist. Insbesondere ist A_f dann auch überabzählbar ([Ü]). Es gibt also „sehr viele" stetige Funktionen, deren Fourier-Reihe auf einer überabzählbaren Menge divergiert.
2. Auf der anderen Seite gibt es allerdings auch ein (sehr tiefliegendes) Resultat, das besagt, dass für jede Funktion $f \in C_{2\pi}$ die Fourier-Reihe $\sum\limits_{\nu=-\infty}^{\infty} \hat{f}(\nu)e^{i\nu t}$ von f für alle t bis auf eine Ausnahmemenge mit Lebesgue-Maß 0 gegen $f(t)$ konvergiert. Dies ist der berühmte Satz von Carleson (1966).

Wir zeigen noch, dass die arithmetischen Mittel der Fourier-Teilsummen für alle $f \in C_{2\pi}$ gleichmäßig auf $[-\pi, \pi]$ gegen f konvergieren. Betrachten wir also

$$\sigma_n(f)(t) := \frac{1}{n+1}\sum_{\nu=0}^{n} S_\nu(f)(t) \qquad (t \in [-\pi, \pi], n \in \mathbb{N}_0)\,.$$

Dann gilt offenbar $\sigma_n(f) \in T_n$ für alle $n \in \mathbb{N}$. Es gilt dafür

Satz 28.17 *(Fejér)*
Es sei $f \in C_{2\pi}$. Dann gilt

$$\sigma_n(f) \to f \qquad \textit{gleichmäßig auf } [-\pi, \pi].$$

Beweis. Es gilt

$$\begin{aligned}\sigma_n(f)(t) &= \frac{1}{n+1}\sum_{\nu=0}^{n}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)D_\nu(t-s)ds \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)\Big(\frac{1}{n+1}\sum_{\nu=0}^{n} D_\nu(t-s)\Big)ds \qquad (t \in [-\pi, \pi])\,.\end{aligned}$$

Wir setzen

$$F_n(t) := \frac{1}{n+1}\sum_{\nu=0}^{n} D_\nu(t) \qquad (t \in [-\pi, \pi], n \in \mathbb{N}_0)$$

(sog. Fejér-Kern). Dann gilt

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(t)dt = \frac{1}{n+1}\sum_{\nu=0}^{n}\underbrace{\frac{1}{2\pi}\int_{-\pi}^{\pi} D_\nu(t)dt}_{=1} = 1$$

und weiter

$$F_n(t) = \begin{cases} \dfrac{1}{n+1} \cdot \dfrac{\sin^2((n+1)t/2)}{\sin^2(t/2)} & , \quad t \neq 0 \\ n+1 & , \quad t = 0 \end{cases} .$$

(Denn: Es gilt für $t \neq 0$

$$D_\nu(t) = \frac{\sin(\nu + 1/2)t)}{\sin(t/2)} ,$$

also

$$\begin{aligned} F_n(t) &= \frac{1}{n+1} \sum_{\nu=0}^{n} D_\nu(t) = \frac{1}{n+1} \cdot \frac{1}{\sin^2(t/2)} \sum_{\nu=0}^{n} \underbrace{\sin((\nu + 1/2)t) \sin(t/2)}_{=\frac{1}{2}[\cos(\nu t) - \cos((\nu+1)t)]} \\ &= \frac{1}{n+1} \cdot \frac{1}{\sin^2(t/2)} \cdot \frac{1}{2}(1 - \cos((n+1)t)) = \frac{1}{n+1} \frac{\sin^2((n+1)t/2)}{\sin^2(t/2)} \end{aligned}$$

und für $t = 0$

$$F_n(0) = \frac{1}{n+1} \sum_{\nu=0}^{n} (2\nu + 1) = n + 1 \, .)$$

Hieraus folgt für jedes $\delta > 0$:

$$\max_{[-\pi,\delta] \cup [\delta,\pi]} |F_n(t)| \leq \frac{1}{n+1} \frac{1}{\sin^2(\delta/2)} \to \infty \qquad (n \to \infty)$$

und $F_n(t) \geq 0$ auf $[-\pi, \pi]$. Der Beweisteil 1. zu S. 28.5 zeigt nun, dass

$$\sigma_n(f)(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) F_n(t - s) ds = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t - s) F_n(s) ds$$

gleichmäßig auf $[-\pi, \pi]$ gegen f konvergiert. □

Wir wollen zum Abschluss noch ein dem Fejérschen Satz entsprechendes Ergebnis für die Approximation mit algebraischen Polynomen herleiten.

Bemerkung und Definition 28.18 Für $n \in \mathbb{N}_0$ betrachten wir die Funktionen $T_n : [-1, 1] \to \mathbb{R}$ mit

$$T_n(x) := \cos(n \arccos x) \qquad (x \in [-1, 1]) \, .$$

Dann gilt für $n \in \mathbb{N}$

$$T_{n+1}(x) = 2x T_n(x) - T_{n-1}(x) \qquad (x \in [-1, 1]) \, .$$

(Denn: Für $n \in \mathbb{N}$ folgt aus den Additionstheoremen

$$\cos((n+1)\varphi) + \cos(((n-1)\varphi) = 2 \cos \varphi \cos n\varphi \qquad (\varphi \in \mathbb{R}) \, .$$

Für $x \in [-1,1]$ ergibt sich mit $\varphi = \arccos x$ daraus

$$\cos((n+1)\arccos x) = 2\cos(\arccos x)\cos(n\arccos x) - \cos((n-1)\arccos x)\,,$$

also die Behauptung.)
Mit $T_0(x) \equiv 1$ und $T_1(x) = x$ auf $[-1,1]$ sieht man damit induktiv, dass T_n ein Polynom vom Grad n (mit führendem Koeffizienten 2^{n-1} für $n \neq 0$) ist. Das Polynom T_n (fortgesetzt auf $\mathbb{R}$) heißt n-tes *Tschebyscheff-Polynom.*
Wir zeigen damit

Satz 28.19 *(Weierstraßscher Approximationssatz für polynomiale Approximation) Es sei $f \in C[a,b]$, und es sei $\varepsilon > 0$. Dann existiert ein Polynom P mit*

$$||f - P||_\infty = \max_{x\in[a,b]} |f(x) - P(x)| < \varepsilon\,.$$

Beweis.
1. Es sei zunächst $[a,b] = [-1,1]$, und es sei $f \in C[-1,1]$. Wir definieren $g : [-\pi,\pi] \to \mathbb{C}$ durch

$$g(t) := f(\cos t) \qquad (t \in [-\pi,\pi])\,.$$

Dann ist $g \in C_{2\pi}$ (beachte: $\cos(\pi) = \cos(-\pi)$). Da g eine gerade Funktion ist, gilt (mit $a_\nu = \hat{g}(\nu)$)

$$S_n(g)(t) = a_0 + 2\sum_{\nu=1}^{n} a_\nu \cos(\nu t)$$

([Ü]) und damit

$$\begin{aligned}
\sigma_n(g)(t) &= \frac{1}{n+1}\sum_{\nu=0}^{n} S_\nu(g)(t) = \\
&= a_0 + \frac{1}{n+1}\sum_{\nu=1}^{n}\sum_{\mu=1}^{\nu} 2a_\mu\cos(\mu t) = \\
&= a_0 + \frac{1}{n+1}\sum_{\mu=1}^{n} 2a_\mu\cos(\mu t)\cdot \underbrace{\sum_{\nu=\mu}^{n} 1}_{=n-\mu+1} \\
&= a_0 + 2\sum_{\mu=1}^{n}\Big(1 - \frac{\mu}{n+1}\Big)a_\mu\cos(\mu t)\,.
\end{aligned}$$

Für

$$P_n := a_0 + 2\sum_{\mu=1}^{n}\Big(1 - \frac{\mu}{n+1}\Big)a_\mu T_\mu$$

gilt dann: P_n ist ein Polynom vom Grad $\leq n$, und für alle $x \in [-1,1]$ ist mit $x = \cos t$

$$|f(x) - P_n(x)| = |g(t) - \sigma_n(g)(t)| \leq ||g - \sigma_n(g)||_\infty \to 0 \qquad (n \to \infty)\,.$$

Da die rechte Seite unabhängig von x ist, folgt

$$||f - P_n||_\infty \to 0 \qquad (n \to \infty)\,.$$

2. Durch eine geeignete affin-lineare Transformation lässt sich der Fall eines beliebigen Intervalls $[a,b]$ leicht auf den Fall $[a,b] = [-1,1]$ zurückführen ([Ü]). □

Bemerkung 28.20 Ist f reellwertig, so ergibt sich aus dem Beweis, dass auch P reell (also mit reellen Koeffizienten) gewählt werden kann.

29 Holomorphe Funktionen

Bereits in Abschnitt 13 haben wir uns mit der Differenzialrechnung von komplexwertigen Funktionen beschäftigt. Dort haben wir uns allerdings auf den Fall von Funktionen f einer reellen Variablen beschränkt, d.h. $f : M \to \mathbb{C}$, wobei $M \subset \mathbb{R}$ war. Unser Thema wird im Folgenden die Untersuchung von Funktionen einer komplexen Variablen, d.h. $f : M \to \mathbb{C}$ mit $M \subset \mathbb{C}$ sein. Natürlich sind die Ergebnisse, die wir in den Abschnitten 10 und 11 für Funktionen auf auf metrischen Räumen hergeleitet haben, insbesondere in diesem Spezialfall anwendbar.

Auch die Definition der Differenzierbarkeit (D. 13.1) überträgt sich sofort, in dem man „$M \subset \mathbb{R}$“ durch „$M \subset \mathbb{C}$“ (oder „$M \subset \mathbb{K}$“) ersetzt. Schließlich gelten die Sätze 13.3 (Differenzierbarkeit impliziert Stetigkeit), 13.4 (Summen, Produkt-und Quotientenregel), 13.6 (Zerlegungsformel) und 13.7 (Kettenregel) in analoger Wiese (beim Beweis der Zerlegungsformel hat man $\mathrm{sign}(x - x_0)$ durch $(x - x_0)/|x - x_0|$ zu ersetzen). Auch die Umkehrregel gilt unter geeigneten Voraussetzungen. Wir werden später noch genauer darauf zu sprechen kommen. Soweit also alles analog. Wo liegen aber die Unterschiede zur reellen Differenzierbarkeit? Bekanntlich ist $\mathbb{C}$ nichts anderes als der lineare Raum $\mathbb{R}^2$, versehen mit einer geeigneten Multiplikation als zusätzlicher Struktur: In Abschnitt 19 hatten wir uns mit Differenzialrechnung von Funktionen $g : M \to \mathbb{R}^m$, wobei $M \subset \mathbb{R}^d$ für $m, d \in \mathbb{N}$ ist, beschäftigt. Also haben wir insbesondere ein Konzept der Differenzierbarkeit von Funktionen $g : M \to \mathbb{R}^2$, wobei $M \subset \mathbb{R}^2$ ist, erarbeitet. Man könnte nun auf den Gedanken kommen, dass für eine Funktion $f : M \to \mathbb{C}$ mit $M \subset \mathbb{C}$ gilt: Ist $f(z) = u(x, y) + iv(x, y)$ mit $u : M \to \mathbb{R}$, $v : M \to \mathbb{R}$ (und M als Teilmenge von $\mathbb{R}^2$ aufgefasst), also f zerlegt in Real- und Imaginärteil, so ist f differenzierbar an $z_0 \in M$, falls $g = (u, v) : M \to \mathbb{R}^2$ differenzierbar an (x_0, y_0) (mit $z_0 = x_0 + iy_0$) ist. Dies ist jedoch i.a. *nicht* der Fall:

Beispiel 29.1 Es sei $f : \mathbb{C} \to \mathbb{C}$ definiert durch

$$f(z) = \mathrm{Re}(z) \qquad (z \in \mathbb{C}) .$$

Hier gilt

$$\begin{aligned} u(x, y) &= x \\ v(x, y) &= 0 \end{aligned} \qquad ((x, y) \in \mathbb{R}^2) .$$

Dann ist natürlich $g = (u, v) : \mathbb{R}^2 \to \mathbb{R}^2$ differenzierbar (da $\partial u/\partial x \equiv 1$, $\partial u/\partial y = \partial v/\partial x = \partial v/\partial y \equiv 0$ stetig auf $\mathbb{R}^2$ sind; vgl. 19.13). Aber:

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

existiert für *kein* $z_0 \in \mathbb{C}$.

(Denn: Ist $z_0 \in \mathbb{C}$, so gilt für $h \in \mathbb{R}, h > 0$

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{x_0 + h - x_0}{h} \equiv 1 \to 1 \qquad (h \to 0^+)$$

und für $z = z_0 + ih = x_0 + i(y_0 + h)$

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{x_0 - x_0}{ih} \equiv 0 \to 0 \qquad (h \to 0^+)\,,$$

also existiert der Grenzwert $\lim\limits_{z \to z_0} (f(z) - f(z_0))/(z - z_0)$ nicht. Damit ist f an keiner Stelle $z_0 \in \mathbb{C}$ (komplex) differenzierbar!

Anders als im 2-dimensionalen reellen Fall impliziert die (komplexe) Differenzierbarkeit eine starke „Kopplung“ zwischen Real- und Imaginärteil der Funktion.
Wir beschränken uns ab nun auf Funktionen, die auf einer offenen Menge definiert sind.

Definition 29.2 Es sei $\Omega \subset \mathbb{C}$ offen und es sei $f : \Omega \to \mathbb{C}$. Dann heißt f *holomorph* (oder *regulär*) (in Ω), falls f differenzierbar in jedem Punkt $z_0 \in \Omega$ ist. Wir setzen

$$H(\Omega) := \{f : \Omega \to \mathbb{C} : f \text{ holomorph in } \Omega\}.$$

Beispiel 29.3 1. Wie in B. 13.5 sieht man: Ist $P : \mathbb{C} \to \mathbb{C}$ ein Polynom, $P(z) = \sum\limits_{\nu=0}^{n} a_\nu z^\nu$, so ist P holomorph in $\mathbb{C}$, und es gilt:

$$P'(z) = \sum_{\nu=1}^{n} \nu a_\nu z^{\nu-1} \qquad \Big(= \sum_{\nu=0}^{n-1} (\nu+1) a_{\nu+1} z^\nu \Big)$$

2. Es sei $c \in \mathbb{C}$ fest und $f : \mathbb{C} \to \mathbb{C}$ definiert durch

$$f(z) = e^{cz} \qquad (z \in \mathbb{C})\,.$$

Dann ist f holomorph in $\mathbb{C}$ mit

$$f'(z) = ce^{cz} \qquad (z \in \mathbb{C})\,.$$

(Denn: O.E. sei $c \neq 0$. Nach B. 10.6.1 ist

$$\lim_{z \to 0} \frac{e^{cz} - 1}{cz} = 1$$

bzw.

$$\lim_{z \to 0} \frac{e^{cz} - 1}{z} = c \qquad (= f'(0))\,.$$

Für beliebiges $z_0 \in \mathbb{C}$ ist

$$\lim_{z \to z_0} \frac{e^{cz} - e^{cz_0}}{z - z_0} = e^{cz_0} \lim_{z \to z_0} \frac{e^{c(z-z_0)} - 1}{z - z_0} = ce^{cz_0} ,$$

also ist f differenzierbar an z_0 mit $f'(z_0) = ce^{cz_0}$.)
3. Die Funktionen $\cos : \mathbb{C} \to \mathbb{C}$ und $\sin : \mathbb{C} \to \mathbb{C}$ sind definiert durch

$$\left.\begin{aligned} \cos z &:= \frac{1}{2}(e^{iz} + e^{-iz}) \\ \sin z &:= \frac{1}{2i}(e^{iz} - e^{-iz}) \end{aligned}\right. \qquad (z \in \mathbb{C}) .$$

(Nach D. 7.12 gilt für $x \in \mathbb{R}$

$$\begin{aligned} \cos x &= \mathrm{Re}(e^{ix}) = \frac{1}{2}(e^{ix} + e^{-ix}) \\ \sin x &= \mathrm{Im}(e^{ix}) = \frac{1}{2i}(e^{ix} - e^{-ix}) \end{aligned} ,$$

also stimmt die neue Definition für reelle Argumente mit der bestehenden überein).
Es gilt: sin und cos sind holomorph in $\mathbb{C}$ mit

$$\begin{aligned} (\sin z)' &= \cos z \\ (\cos z)' &= -\sin z \end{aligned} \qquad (z \in \mathbb{C})$$

(ergibt sich aus 2. mit Summen- und Kettenregel).
4. Es sei $a \in \mathbb{C}$, $\Omega := \mathbb{C} \setminus \{a\}$ und $k \in \mathbb{N}$. Dann ist $f : \Omega \to \mathbb{C}$ mit $f(z) = 1/(a - z)^k$ für $z \neq a$ holomorph in Ω mit

$$f'(z) = \frac{k}{(a - z)^{k+1}} \qquad (z \in \mathbb{C} \setminus \{a\})$$

(ergibt sich aus der Quotientenregel und 1.).

Es gilt folgender wichtige Zusammenhang zwischen 2-dimensionaler reeller und komplexer Differenzierbarkeit.

Satz 29.4 *Es sei $\Omega \subset \mathbb{C}$ offen, und es sei $f : \Omega \to \mathbb{C}$ mit*

$$f(z) = u(x, y) + iv(x, y) \qquad (z = x + iy \in \Omega)$$

und $u, v : \Omega \to \mathbb{R}$ (d.h. $u = \mathrm{Re} f, v = \mathrm{Im} f$). Ist $z_0 \in \Omega$, so sind folgende Aussagen äquivalent:

a) f ist (komplex) differenzierbar an z_0.

b) u und v sind (reell) differenzierbar an (x_0, y_0) mit

$$\begin{aligned} \frac{\partial u}{\partial x}(x_0,y_0) &= \frac{\partial v}{\partial y}(x_0,y_0) \\ \frac{\partial u}{\partial y}(x_0,y_0) &= -\frac{\partial v}{\partial x}(x_0,y_0) \end{aligned} \tag{29.1}$$

(sog. Cauchy-Riemannsche Differenzialgleichungen).

Außerdem gilt in diesem Fall

$$\begin{aligned} f'(z_0) &= \frac{\partial u}{\partial x}(x_0,y_0) + i\,\frac{\partial v}{\partial x}(x_0,y_0) \\ \Big(&= \frac{\partial v}{\partial y}(x_0,y_0) - i\,\frac{\partial u}{\partial y}(x_0,y_0)\Big)\,. \end{aligned}$$

Beweis. a) $\Rightarrow$ b): Ist f differenzierbar an z_0, so gilt einerseits für $z = z_0 + h$, wobei h reell ist,

$$\begin{aligned} f'(z_0) &= \lim_{h\to 0} \frac{f(z_0+h)-f(z_0)}{h} \\ &= \lim_{h\to 0} \frac{u(x_0+h,y_0)-u(x_0,y_0)}{h} + i\,\frac{v(x_0+h,y_0)-v(x_0,y_0)}{h}\,. \end{aligned}$$

Also existieren die beiden (reellen) Grenzwerte

$$\lim_{h\to 0} \frac{u(x_0+h,y_0)-u(x_0,y_0)}{h} \qquad \text{und} \qquad \lim_{h\to 0} \frac{v(x_0+h,y_0)-v(x_0,y_0)}{h}\,.$$

Nach D. 19.3 existieren die partiellen Ableitungen $\frac{\partial u}{\partial x}(x_0,y_0)$ und $\frac{\partial v}{\partial x}(x_0,y_0)$, und es gilt

$$f'(z_0) = \frac{\partial u}{\partial x}(x_0,y_0) + i\,\frac{\partial v}{\partial x}(x_0,y_0)\,.$$

Andererseits gilt mit $z = z_0 + ih$ und h reell auch

$$\begin{aligned} f'(z_0) &= \lim_{h\to 0} \frac{f(z_0+ih)-f(z_0)}{ih} \\ &= \lim_{h\to 0} \frac{u(x_0,y_0+h)-u(x_0,y_0)}{ih} + \frac{v(x_0,y_0+h)-v(x_0,y_0)}{h} \\ &= -i\,\frac{\partial u}{\partial y}(x_0,y_0) + \frac{\partial v}{\partial y}(x_0,y_0)\,. \end{aligned}$$

Damit ergibt sich die Zusatzbehauptung und durch Vergleich von Real- und Imaginärteil auch die Gültigkeit von (29.1).
Nach der Zerlegungsformel (S. 13.6 mit $\mathbb{C}$ statt $\mathbb{R}$) existiert eine an z_0 stetige Funktion $\varepsilon : \Omega \to \mathbb{C}$ mit

$$f(z) = f(z_0) + f'(z_0)(z-z_0) + \varepsilon(z)|z-z_0|$$

und $\varepsilon(z_0) = 0$. Ist $\varepsilon(z) = \alpha(x,y) + i\beta(x,y)$ mit $\alpha, \beta : \Omega \to \mathbb{R}$, so bedeutet dies für den Realteil (da $f'(z_0) = \frac{\partial u}{\partial x}(x_0,y_0) - i\frac{\partial u}{\partial y}(x_0,y_0)$)

$$u(x,y) = u(x_0,y_0) + \frac{\partial u}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial u}{\partial y}(x_0,y_0)(y-y_0) + \alpha(x,y)|(x-x_0,y-y_0)|$$

und für den Imaginärteil (da $f'(z_0) = \frac{\partial v}{\partial y}(x_0,y_0) + i\frac{\partial v}{\partial x}(x_0,y_0)$)

$$v(x,y) = v(x_0,y_0) + \frac{\partial v}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial v}{\partial y}(x_0,y_0)(y-y_0) + \beta(x,y)|(x-x_0,y-y_0)|\,.$$

Da $\alpha(x,y) \to 0$ und $\beta(x,y) \to 0$ für $(x,y) \to (x_0,y_0)$ gilt, sind u und v differenzierbar an (x_0,y_0) nach D. 19.6.

b) $\Rightarrow$ a): Sind u, v differenzierbar an (x_0,y_0) und gelten die Gleichungen (29.1), so ergibt sich aus D. 19.6, B. 19.7.2 und S. 19.10 die Existenz einer Funktion $(\alpha,\beta) : \Omega \to \mathbb{R}^2$ mit

$$\begin{pmatrix} u(x,y) - u(x_0,y_0) \\ v(x,y) - v(x_0,y_0) \end{pmatrix} = J_{(u,v)}(x_0,y_0)\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} + \begin{pmatrix} \alpha(x,y) \\ \beta(x,y) \end{pmatrix}|(x-x_0,y-y_0)|$$

mit $\alpha(x,y) \to 0$, $\beta(x,y) \to 0$ für $(x,y) \to (x_0,y_0)$. Nach (29.1) ist

$$J_{(u,v)}(x_0,y_0) = \begin{pmatrix} \frac{\partial u}{\partial x}(x_0,y_0) & -\frac{\partial v}{\partial x}(x_0,y_0) \\ & \\ \frac{\partial v}{\partial x}(x_0,y_0) & \frac{\partial u}{\partial x}(x_0,y_0) \end{pmatrix},$$

also in komplexer Schreibweise mit $\varepsilon(z) := \alpha(x,y) + i\beta(x,y)$

$$f(z) - f(z_0) = \Big(\frac{\partial u}{\partial x}(x_0,y_0) + i\,\frac{\partial v}{\partial x}(x_0,y_0)\Big)(z-z_0) + \varepsilon(z)|z-z_0|\,.$$

Da $\varepsilon(z) \to 0\,(z \to z_0)$ gilt, ist f differenzierbar an z_0 nach der Zerlegungsformel (S.13.6 mit $\mathbb{C}$ statt $\mathbb{R}$), und es gilt

$$f'(z_0) = \frac{\partial u}{\partial x}(x_0,y_0) + i\,\frac{\partial v}{\partial x}(x_0,y_0) \overset{(29.1)}{=} \frac{\partial v}{\partial y}(x_0,y_0) - i\,\frac{\partial u}{\partial y}(x_0,y_0)\,.$$

□

Bemerkung 29.5 1. Insbesondere ergibt sich aus S. 29.4, dass $f : \Omega \to \mathbb{C}$ genau dann holomorph in Ω ist, wenn u und v differenzierbar auf Ω sind mit

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \qquad \text{und} \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \qquad \text{auf } \Omega\,.$$

Hinreichend für die Differenzierbarkeit von u und v auf Ω ist nach S. 19.13 die Existenz und Stetigkeit der partiellen Ableitungen $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}$ auf Ω.

2. Die Cauchy-Riemannschen Differenzialgleichungen (29.1) lassen sich auch durch eine (komplexe) Differenzialgleichung ausdrücken: Definiert man für $f : \Omega \to \mathbb{C}$, $f = u + iv$, und $z_0 = x_0 + iy_0 \in \Omega$ (im Falle der Existenz)

$$\begin{aligned} \overline{\partial} f(z_0) &:= \frac{1}{2}\Big(\frac{\partial f}{\partial x}(z_0) + i\,\frac{\partial f}{\partial y}(z_0)\Big) \\ &= \frac{1}{2}\Big(\frac{\partial u}{\partial x} + i\,\frac{\partial v}{\partial x} + i\,\Big(\frac{\partial u}{\partial y} + i\,\frac{\partial v}{\partial y}\Big)\Big)(x_0, y_0)\,, \end{aligned}$$

so gilt (29.1) genau dann, wenn

$$\overline{\partial} f(z_0) = 0$$

erfüllt ist.

In Abschnitt 16 haben wir uns bereits recht intensiv mit Potenzreihen beschäftigt. U. a. haben wir gezeigt, dass Potenzreihen stets stetig in ihrem Konvergenzkreis sind und dass sie im reellen Fall beliebig oft differenzierbar sind. Wir zeigen nun eine entsprechende Aussage für den komplexen Fall.

Satz 29.6 *Es sei $\sum\limits_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu$ eine Potenzreihe mit Konvergenzradius $R > 0$, und es sei $A(z) := \sum\limits_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu$. Dann existieren die Ableitungen $A^{(k)}$ für alle $k \in \mathbb{N}$ im Konvergenzkreis $U_R(z_0)$ (also $A^{(k)} \in H(U_R(z_0))$ für alle $k \in \mathbb{N}_0$), und es gilt*

$$A^{(k)}(z) = \sum_{\nu=0}^{\infty} (\nu+1)(\nu+2)\dots(\nu+k) a_{\nu+k} (z-z_0)^\nu$$

in $U_R(z_0)$. Insbesondere ist

$$k! a_k = A^{(k)}(z_0) \qquad (k \in \mathbb{N}_0)\,,$$

also

$$A(z) = \sum_{\nu=0}^{\infty} \frac{A^{(\nu)}(z_0)}{\nu!} (z-z_0)^\nu \qquad (z \in U_R(z_0))\,.$$

Beweis. 1. Wir zeigen $A \in H(U_R(z_0))$ und

$$A'(z) = \sum_{\nu=1}^{\infty} \nu a_\nu (z-z_0)^{\nu-1} = \sum_{\nu=0}^{\infty} (\nu+1) a_{\nu+1} (z-z_0)^\nu \qquad (z \in U_R(z_0))\,.$$

O. E. können wir $z_0 = 0$ annehmen.
Zunächst hat auch die Potenzreihe $\sum\limits_{\nu=1}^{\infty} \nu a_\nu z^\nu$ den Konvergenzradius R. Also konvergiert $\sum\limits_{\nu=1}^{\infty} \nu |a_\nu| r^{\nu-1}$ für alle $r < R$ (S. 16.2).

Es sei $z_1 \in U_R(0)$ fest. Wir wählen ein $r \in (|z_1|, R)$ und setzen $\delta := r - |z_1|$. Für $z \in U_R(0)$ gilt

$$A(z) - A(z_1) = \sum_{\nu=1}^{\infty} a_\nu (z^\nu - z_1^\nu) = (z - z_1) \sum_{\nu=1}^{\infty} a_\nu \underbrace{\sum_{k=0}^{\nu-1} z_1^k z^{\nu-1-k}}_{=: \phi_\nu(z)} .$$

Ist $|z - z_1| < \delta$, so ist $|z| < r$ (und $|z_1| < r$), also $|\phi_\nu(z)| \le \nu r^{\nu-1}$. Nach dem Weierstraßschen Majorantenkriterium (S. 15.6) konvergiert $\sum\limits_{\nu=1}^{\infty} a_\nu \phi_\nu(z)$ gleichmäßig auf $U_\delta(z_1)$. Folglich ist $\phi(z) := \sum\limits_{\nu=1}^{\infty} a_\nu \phi_\nu(z)$ stetig auf $U_\delta(z_1)$ (S. 15.10), und es gilt

$$\lim_{z \to z_1} \frac{A(z) - A(z_1)}{z - z_1} = \lim_{z \to z_1} \phi(z) = \phi(z_1) = \sum_{\nu=1}^{\infty} \nu a_\nu z_1^{\nu-1} .$$

2. Durch Anwendung der gleichen Argumentation auf A' sieht man: $A' \in H(U_R(z_0))$ und

$$A''(z) = \sum_{\nu=0}^{\infty} (\nu+1)(\nu+2) a_{\nu+2} (z - z_0)^\nu \qquad (z \in U_R(z_0)) .$$

Induktiv erhält man: Für alle $k \in \mathbb{N}$ ist $A^{(k-1)} \in H(U_R(z_0))$ und

$$A^{(k)}(z) = \sum_{\nu=0}^{\infty} (\nu+1)(\nu+2) \ldots (\nu+k) a_{\nu+k} (z - z_0)^\nu .$$

Die Zusatzbehauptung $k! a_k = A^{(k)}(z_0)$ ergibt sich für $z = z_0$. □

Definition 29.7 Es sei $\Omega \subset \mathbb{K}$ offen, und es sei $f : \Omega \to \mathbb{K}$. Dann heißt f *analytisch* (in Ω), falls zu jedem $z_0 \in \Omega$ ein $R = R(z_0) > 0$ so existiert, dass

$$f(z) = \sum_{\nu=0}^{\infty} a_\nu (z - z_0)^\nu \qquad (z \in U_R(z_0))$$

für eine gewisse Folge (a_ν) (abhängig von z_0) in $\mathbb{K}$ gilt (d.h. f ist in jedem Punkt „lokal“ durch eine Potenzreihe darstellbar). [Dann ist natürlich notwendigerweise $a_\nu = f^{(\nu)}(z_0)/\nu!$.]

Bemerkung 29.8 Aus S.29.6 ergibt sich mit dieser Sprechweise: Ist $\Omega \subset \mathbb{C}$ offen und $f : \Omega \to \mathbb{C}$ analytisch, so ist insbesondere f holomorph in Ω.
Von fundamentaler Bedeutung für die gesamte Funktionentheorie ist die Tatsache, dass auch die Umkehrung dieser Aussage wahr ist. Dies werden wir im nächsten Abschnitt beweisen.

Eine große Klasse analytischer Funktionen liefert

Satz 29.9 *Es sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall, und es seien $\varphi, \psi : [a,b] \to \mathbb{K}$ Regelfunktionen. Ferner sei $\Omega \subset \mathbb{K}$ offen mit $\Omega \cap \psi([a,b]) = \emptyset$. Wir definieren $f : \Omega \to \mathbb{K}$ durch*

$$f(z) := \int_a^b \frac{\varphi(t)}{\psi(t) - z}\, dt \qquad (z \in \Omega)\,.$$

Dann ist f analytisch in Ω, und es gilt

$$f^{(k)}(z) = k! \int_a^b \frac{\varphi(t)}{(\psi(t) - z)^{k+1}}\, dt \qquad (z \in \Omega,\ k \in \mathbb{N})\,.$$

Beweis. Es sei $z_0 \in \Omega$ und $U_R(z_0) \subset \Omega$. Aus

$$\left| \frac{z - z_0}{\psi(t) - z_0} \right| \le \frac{|z - z_0|}{R} < 1$$

für alle $z \in U_R(z_0)$ und alle $t \in [a,b]$ folgt, dass die Reihe

$$\sum_{\nu=0}^{\infty} \frac{(z - z_0)^\nu}{(\psi(t) - z_0)^{\nu+1}} = \frac{1}{\psi(t) - z_0} \cdot \frac{1}{1 - \frac{z - z_0}{\psi(t) - z_0}} = \frac{1}{\psi(t) - z}$$

für jedes feste $z \in U_R(z_0)$ gleichmäßig auf $[a,b]$ konvergiert. (Weierstraßsches Majorantenkriterium; S. 15.6). Also erhalten wir mit S. 17.20

$$f(z) = \int_a^b \varphi(t) \sum_{\nu=0}^{\infty} \frac{(z - z_0)^\nu}{(\psi(t) - z_0)^{\nu+1}}\, dt = \sum_{\nu=0}^{\infty} (z - z_0)^\nu \cdot \int_a^b \frac{\varphi(t)}{(\psi(t) - z_0)^{\nu+1}}\, dt\,,$$

d.h. mit

$$a_\nu := \int_a^b \frac{\varphi(t)}{(\psi(t) - z_0)^{\nu+1}}\, dt \qquad (\nu \in \mathbb{N}_0)$$

gilt für alle $z \in U_R(z_0)$

$$f(z) = \sum_{\nu=0}^{\infty} a_\nu (z - z_0)^\nu\,.$$

Folglich ist f analytisch in Ω. Außerdem erhalten wir für $z = z_0$

$$f^{(k)}(z_0) = k! a_k = k! \int_a^b \frac{\varphi(t)}{(\psi(t) - z_0)^{k+1}}\, dt \qquad (k \in \mathbb{N}_0)$$

nach S. 29.6. □

Zum Abschluss beweisen wir eine Eigenschaft, die analytische Funktionen wesentlich gegenüber „gewöhnlichen“ C^∞-Funktionen auszeichnet.

Satz 29.10 *Es sei $G \subset \mathbb{K}$ ein Gebiet, und es sei $f : G \to \mathbb{K}$ analytisch. Wir setzen*

$$Z(f) := \{z_0 \in G : f(z_0) = 0\}\,.$$

Dann gilt: Entweder ist $Z(f) = G$ (d.h. $f \equiv 0$) oder $Z(f)$ hat keinen Häufungspunkt in G. Im zweiten Fall existieren zu jedem $z_0 \in Z(f)$ eine eindeutig bestimmte Zahl $m = m(z_0) \in \mathbb{N}$ und eine auf einer Umgebung U von z_0 stetige Funktion g mit $g(z_0) \neq 0$ und

$$f(z) = (z - z_0)^m g(z) \qquad (z \in U)\,.$$

Beweis. Es sei $z_0 \in Z(f)$ fest, und es sei $R = R(z_0) > 0$ so, dass

$$f(z) = \sum_{\nu=1}^{\infty} a_\nu (z - z_0)^\nu \qquad (z \in U_R(z_0))$$

(mit $a_\nu = f^{(\nu)}(z_0)/\nu!$) gilt. Nun sind zwei Fälle möglich: Entweder ist $a_\nu = 0$ für alle $\nu \in \mathbb{N}_0$, also $f(z) \equiv 0$ auf $U_R(z_0)$ oder es existiert eine kleinste Zahl $m \in \mathbb{N}$ mit $a_m \neq 0$. In diesem Fall setzen wir

$$g(z) := \begin{cases} (z - z_0)^{-m} f(z) & \text{für } z \in U_R(z_0) \setminus \{z_0\} \\ a_m & \text{für } z = z_0 \end{cases}$$

Dann gilt $f(z) = (z - z_0)^m g(z)$ für alle $z \in U_R(z_0)$, und außerdem ist

$$g(z) = \sum_{\nu=0}^{\infty} a_{m+\nu} (z - z_0)^\nu \qquad (z \in U_R(z_0))\,.$$

Insbesondere ist g stetig auf $U_R(z_0)$. Aus $g(z_0) = a_m \neq 0$ und der Stetigkeit von g folgt, dass $g(z) \neq 0$ auf einer Umgebung von z_0 gilt. Damit gilt im Falle $a_m \neq 0$ für ein $m \in \mathbb{N}$ die Zusatzbehauptung und z_0 ist ein isolierter Punkt von $Z(f)$.
Es sei $A := H(Z(f)) \cap G$ die Menge der Häufungspunkte von $Z(f)$ in G. Da f stetig auf G ist, gilt $A \subset Z(f)$.
Also: Ist $A \neq \emptyset$ und $z_0 \in A$, so tritt notwendigerweise der erste Fall auf, d.h. $f(z) \equiv 0$ auf einer Umgebung von z_0. Damit ist $z_0 \in A^0$, also A offen. Außerdem ist A auch abgeschlossen (in $(G, d_{|\cdot|})$) als Menge von Häufungspunkten ([Ü]). Da $(G, d_{|\cdot|})$ zusammenhängend ist, gilt schon $A = \Omega$. Dies war zu zeigen. □

Als Folgerung erhalten wir unmittelbar folgendes fundamentale Ergebnis.

Satz 29.11 *(Identitätssatz für analytische Funktionen)*
Es sei $G \subset \mathbb{K}$ ein Gebiet, und es seien $f, g : G \to \mathbb{K}$ analytisch. Existiert eine Folge (z_n) in G mit Häufungspunkt in G und so, dass

$$f(z_n) = g(z_n)$$

für alle $n \in \mathbb{N}$ gilt, so ist $f \equiv g$ in G.

Beweis. Mit f und g ist offenbar auch $f - g$ analytisch in G. Aus $z_n \in Z(f - g)$ für alle $n \in \mathbb{N}$ ergibt sich die Behauptung sofort aus S. 29.10. $\square$

30 Cauchyscher Integralsatz und Cauchysche Integralformel

Wir kommen nun zu den zentralen Ergebnissen der sogenannten Cauchyschen Funktionentheorie. Dabei werden wir, wie bereits oben angedeutet, insbesondere sehen, dass jede holomorphe Funktion analytisch und damit beliebig oft differenzierbar ist – eine Art „mathematisches Wunder"!
Vorbereitend beschäftigen wir uns kurz mit dem Konzept komplexer Wegintegrale. Wir werden uns dabei auf Integrale längs Pfaden beschränken (vgl. B/D. 26.17).

Definition 30.1 Es sei $\gamma : [\alpha, \beta] \to \mathbb{C}$ ein Pfad, d.h. γ ist stetig, und es existieren $\alpha = t_0 < t_1 < \ldots < t_n = \beta$ so, dass $\gamma_{|[t_{j-1}, t_j]}$ stetig differenzierbar auf $[t_{j-1}, t_j]$ ist $(j = 1, \ldots, n)$ [Man beachte dabei: an den Punkten $t_1, \ldots, t_{n-1}$ können rechts- und linksseitige Ableitung unterschiedlich sein.]. Ferner sei $\Gamma = \gamma([\alpha, \beta])$ die Kurve mit Parameterdarstellung γ. Ist $f : \Gamma \to \mathbb{C}$ stetig auf Γ, so ist $f \circ \gamma \cdot \gamma' \in R[\alpha, \beta]$ nach S.17.6. Wir definieren das *Integral von f längs γ* durch

$$\int_\gamma f := \int_\gamma f(z)dz := \int_\alpha^\beta f \circ \gamma \cdot \gamma' = \int_\alpha^\beta f(\gamma(t))\gamma'(t)dt$$

(wobei γ' beliebig definiert ist an den Stellen t_j).

Bemerkung 30.2 1. Es seien $[\alpha_1, \beta_1]$ und $[\alpha, \beta]$ Intervalle, und es sei $\varphi : [\alpha_1, \beta_1] \to [\alpha, \beta]$ stetig differenzierbar mit $\varphi(\alpha_1) = \alpha$, $\varphi(\beta_1) = \beta$. Ist $\gamma : [\alpha, \beta] \to \mathbb{C}$ ein Pfad, so ist auch $\gamma_1 := \gamma \circ \varphi$ ein Pfad (mit $\Gamma = \gamma([\alpha, \beta]) = \gamma_1([\alpha_1, \beta_1])$, d.h. beide parametrisieren die selbe Kurve). Mit der Substitutionsregel („stückweise" angewandt) gilt dann

$$\int_{\gamma_1} f = \int_{\alpha_1}^{\beta_1} f \circ \gamma_1 \cdot \gamma_1' = \int_{\alpha_1}^{\beta_1} f \circ \gamma \circ \varphi \cdot \gamma' \circ \varphi \cdot \varphi' = \int_\alpha^\beta f \circ \gamma \cdot \gamma' = \int_\gamma f\,.$$

Sind γ_1 und γ_2 Pfade, die beide die selbe Kurve Γ parametrisieren, so sagen wir, γ_1 und γ_2 seien *$C(\Gamma)$-äquivalent*, falls

$$\int_{\gamma_1} f = \int_{\gamma_2} f$$

für alle $f \in C(\Gamma)$ gilt. Insbesondere zeigt obige Überlegung, dass zu jedem Pfad γ und zu jedem kompakten Intervall $[\sigma, \tau]$ ein zu γ $C(\Gamma)$-äquivalenter Pfad $\tilde{\gamma} : [\sigma, \tau] \to \mathbb{C}$ existiert. Wir können also das Parameterintervall stets nach Wunsch wählen.

Man beachte jedoch: Ist γ wie oben und $\varphi(t) = \alpha + \beta - t\,(t \in [\alpha, \beta])$, so gilt für $\gamma_1 = \gamma \circ \varphi$ (also $\gamma_1(t) = \gamma(\alpha + \beta - t)$)

$$\int_{\gamma_1} f = \int_\alpha^\beta f \circ \gamma_1 \cdot \gamma_1' = \int_\alpha^\beta f \circ \gamma \circ \varphi \cdot \gamma' \circ \varphi \cdot \varphi' = \int_\beta^\alpha f \circ \gamma \cdot \gamma' = -\int_\gamma f\,.$$

Wir schreiben für den Weg γ_1 deshalb auch $-\gamma$. Es gilt damit sehr suggestiv

$$\int_{-\gamma} f = -\int_\gamma f \qquad (f \in C(\Gamma))\,.$$

2. Eine einfache, aber oft sehr nützliche Abschätzung für das Integral von f längs γ ergibt sich (mit $\|f\|_\infty = \sup_{\zeta \in \Gamma} |f(\zeta)|$) als

$$\begin{aligned} \Big|\int_\gamma f\Big| &= \Big|\int_\alpha^\beta f \circ \gamma \cdot \gamma'\Big| \leq \int_\alpha^\beta |f \circ \gamma||\gamma'| \\ &\leq \|f\|_\infty \cdot \int_\alpha^\beta |\gamma'| \overset{26.17}{=} \|f\|_\infty \cdot L(\gamma) \end{aligned}$$

3. Ist $\gamma = \bigoplus_{j=1}^n \gamma^{(j)}$ die Summe von Pfaden $\gamma^{(1)}, \ldots, \gamma^{(n)}$, so folgt aus D. 30.1 unmittelbar $\int_\gamma f = \sum_{j=1}^n \int_{\gamma^{(j)}} f$ für alle stetigen $f : \Gamma \to \mathbb{C}$.

Beispiel 30.3 1. Es sei $\gamma : [\alpha, \beta] \to \mathbb{C}$ ein beliebiger Pfad, $m \in \mathbb{N}_0$ und $f(z) = z^m$ $(z \in \mathbb{C})$. Dann gilt nach dem HDI (stückweise angewandt)

$$\begin{aligned} \int_\gamma z^m dz &= \int_\alpha^\beta \gamma^m(t)\gamma'(t)dt = \frac{1}{m+1}\int_\alpha^\beta (\gamma^{m+1})'(t)dt \\ &= \frac{1}{m+1}(b^{m+1} - a^{m+1})\,, \end{aligned}$$

wobei $\gamma(\alpha) = a, \gamma(\beta) = b$. Also: Der Wert des Integrals hängt nur von den Anfangs- und Endpunkten von γ ab, nicht aber vom Weg γ! Insbesondere gilt für jeden geschlossenen Pfad γ

$$\int_\gamma z^m dz = 0\,.$$

2. Es sei $z_0 \in \mathbb{C}, R > 0$ und $\gamma : [0, 2\pi] \to \mathbb{C}$ mit $\gamma(t) = z_0 + Re^{it}$ (dann ist $\Gamma = \gamma([0, 2\pi])$ der Kreis mit Radius R um z_0). Hier gilt für $m \in \mathbb{Z}$

$$\int_\gamma (z - z_0)^m dz = \begin{cases} 0 & \text{, falls } m \neq -1 \\ 2\pi i & \text{, falls } m = -1 \end{cases}\,.$$

(Denn: Für $m \neq -1$ ist

$$\begin{aligned}\int_\gamma (z-z_0)^m dz &= \int_0^{2\pi} (Re^{it})^m\, iRe^{it}\, dt = iR^{m+1} \int_0^{2\pi} e^{it(m+1)} dt \\ &= iR^{m+1} \frac{1}{i(m+1)} e^{it(m+1)}\Big|_0^{2\pi} = 0\end{aligned}$$

und für $m = -1$ ist

$$\int_\gamma \frac{dz}{z-z_0} = \int_0^{2\pi} (Re^{it})^{-1}\, iRe^{it}\, dt = i\int_0^{2\pi} dt = 2\pi i.)$$

Man beachte auch hier: der Wert des Integrals hängt nicht vom Radius R ab.

Bemerkung 30.4 Oft werden wir Integrale längs Kreisen oder Polygonzügen betrachten. Wir schreiben daher für $\gamma : [0, 2\pi]$ mit $\gamma(t) = z_0 + Re^{it}$ wie in B.30.3.2 auch

$$\int_{|z-z_0|=R} f = \int_{K_R(z_0)} f := \int_\gamma f \qquad (f \in C(\Gamma))\,.$$

Dabei ist

$$K_R(z_0) := \{z : |z-z_0| = R\} \quad (= \partial U_R(z_0))$$

der Kreis mit Radius R um z_0).

Ist ferner Γ der Polygonzug durch $x^{(0)}, \ldots, x^{(n)}$ wie in B.26.11, und ist $\gamma = \bigoplus_{j=1}^{n} \gamma^{(j)}$ wie dort, so setzen wir für $f \in C(\Gamma)$

$$\int_\Gamma f := \int_\gamma f \qquad \Big(= \sum_{j=1}^{n} \int_{\gamma^{(j)}} f\Big)\,.$$

Ist insbesondere Γ ein Polygonzug durch $x^{(0)}, x^{(1)}$, also die Strecke von $x^{(0)}$ nach $x^{(1)}$ (inklusive Endpunkten), so schreiben wir auch

$$\int_{[x^{(0)},x^{(1)}]} f = \int_\Gamma f\,.$$

Da der Übergang zu äquivalenten Parametrisierungen den Wert der Integrale nicht ändert, ergibt sich insbesondere für den Polygonzug Γ durch $x^{(0)}, \ldots, x^{(n)}$

$$\int_\Gamma f = \sum_{j=1}^{n} \int_{[x^{(j-1)},x^{(j)}]} f \qquad (f \in C(\Gamma))\,.$$

Wir werden nun zeigen, dass unter geeigneten Bedingungen Integrale längs geschlossener Pfade – so wie in B.30.3.1 – verschwinden.

Satz 30.5 *Es sei $G \subset \mathbb{C}$ ein Gebiet, und es sei $f : G \to \mathbb{C}$ stetig. Dann sind folgende Aussagen äquivalent:*

a) Es existiert eine Funktion $F \in H(G)$ mit $F' = f$ in G (d.h. F ist eine Stammfunktion zu f in G).

b) Für alle geschlossenen Pfade γ in G gilt

$$\int_\gamma f = 0\,.$$

Außerdem gilt dann: Ist F eine Stammfunktion zu f in G und ist γ ein Pfad in G mit Anfangspunkt a und Endpunkt b, so ist

$$\int_\gamma f = F(b) - F(a)\,.$$

Beweis. a) $\Rightarrow$ b): Es sei $\gamma : [\alpha, \beta] \to G$ mit $\gamma(\alpha) = a$ und $\gamma(\beta) = b$. Dann gilt

$$\begin{aligned} \int_\gamma f &= \int_\alpha^\beta f \circ \gamma \cdot \gamma' = \int_\alpha^\beta (F \circ \gamma)' = F(\gamma(\beta)) - F(\gamma(\alpha)) = \\ &= F(b) - F(a) \end{aligned}$$

nach dem HDI (stückweise angewandt).

b) $\Rightarrow$ a): Es sei $z_* \in G$ fest. Wir definieren $F : G \to \mathbb{C}$ durch

$$F(z) = \int_{z_*}^{z} f(\zeta) d\zeta := \int_{\gamma_z} f(\zeta) d\zeta \qquad (z \in G)\,,$$

wobei γ_z ein beliebiger Pfad in G mit Anfangspunkt z_* und Endpunkt z ist (ein solcher existiert stets, da G Polygon-zusammenhängend ist).
Wichtig ist dabei, dass die Definition unabhängig von der speziellen Wahl von γ_z ist! (Denn: Sind γ_1 und γ_2 Pfade in G mit Anfangspunkt z_* und Endpunkt z, so können wir o.E. annehmen, dass γ_1 etwa auf $[-1, 0]$ und γ_2 auf $[0, 1]$ definiert sind. Dann ist der Pfad $\gamma_1 \oplus (-\gamma_2)$ geschlossen, und deshalb gilt

$$0 = \int_{\gamma_1 \oplus (-\gamma_2)} f = \int_{\gamma_1} f + \int_{-\gamma_2} f = \int_{\gamma_1} f - \int_{\gamma_2} f\,,$$

also $\int_{\gamma_1} f = \int_{\gamma_2} f$.)
Nun sei $z_0 \in G$ fest. Wir zeigen: F ist differenzierbar an z_0 mit $F'(z_0) = f(z_0)$. Für $z_1, z_2 \in \mathbb{C}$ setzen wir $\gamma_{z_1,z_2}(t) := z_1 + t(z_2 - z_1)$, wobei $t \in [0, 1]$. Ist γ_{z_0} ein Pfad in G,

der z_* und z_0 verbindet (o. E. auf $[-1,0]$ definiert), und ist $r > 0$ so, dass $U_r(z_0) \subset G$ ist, so ist $\gamma_z = \gamma_{z_0} \oplus \gamma_{z_0,z}$ für $z \in U_r(z_0)$ ein Pfad in G, der z_* und z verbindet. Es gilt

$$F(z) - F(z_0) = \int\limits_{\gamma_z} f - \int\limits_{\gamma_{z_0}} f = \int\limits_{\gamma_{z_0,z}} f \left(= \int\limits_{[z_0,z]} f \right) .$$

Also ergibt sich weiter für $z \neq z_0$ (beachte: $\int\limits_{[z_0,z]} c dz = c(z - z_0)$),

$$\frac{F(z) - F(z_0)}{z - z_0} - f(z_0) = \frac{1}{z - z_0} \int\limits_{[z_0,z]} [f(\zeta) - f(z_0)] d\zeta .$$

Ist $\varepsilon > 0$ gegeben, so folgt aus der Stetigkeit von f an z_0 die Existenz eines $\delta \in (0, r)$ mit $|f(\zeta) - f(z_0)| < \varepsilon$ für $|\zeta - z_0| < \delta$. Damit ist

$$\Big|\frac{1}{z - z_0} \int\limits_{[z_0,z]} [f(\zeta) - f(z_0)] d\zeta\Big| \leq \frac{1}{|z - z_0|} \varepsilon \cdot |z - z_0| = \varepsilon .$$

Dies zeigt, dass F differenzierbar an z_0 ist mit $F'(z_0) = f(z_0)$. □

Definition 30.6 Es sei $G \subset \mathbb{C}$ ein Gebiet. Dann heißt G *sternförmig* (bezüglich z_*), falls ein $z_* \in G$ so existiert, dass die Strecke $[z_*, z] = \{z_* + t(z - z_*) : t \in [0, 1]\}$ in G liegt für alle $z \in G$.

Bemerkung 30.7 Offensichtlich ist jedes konvexe Gebiet auch sternförmig. Ein Beispiel eines nicht-konvexen, sternförmigen Gebietes ist etwa $G = \mathbb{C} \setminus [1, \infty)$. Damit beweisen wir folgendes Ergebnis, das von fundamentaler Bedeutung für die komplexe Analysis ist.

Satz 30.8 *(Cauchyscher Integralsatz für sternförmige Gebiete)*
Es sei $G \subset \mathbb{C}$ ein sternförmiges Gebiet, und es sei f stetig in G und holomorph in $G \setminus \{p\}$ für einen Punkt $p \in G$. Dann ist

$$\int\limits_{\gamma} f = 0$$

für alle geschlossenen Pfade in G.

Beweis. 1. Es seien $a, b, c \in \mathbb{C}$. Wir bezeichnen mit $\Delta = \Delta_{a,b,c}$ das abgeschlossene Dreieck mit den Eckpunkten a, b, c (also die konvexe Hülle von a, b, c) und fassen

$\partial\Delta = \partial\Delta_{a,b,c}$ als Polygonzug durch a, b, c auf.
Wir zeigen nun: Ist $\Delta = \Delta_{a,b,c}$ ein Dreieck in G, so gilt

$$I := \int\limits_{\partial\Delta} f = 0\,. \tag{30.1}$$

Wir betrachten zunächst den Fall, dass der Ausnahmepunkt p nicht in Δ liegt.
Sind a', b', c' die Mittelpunkte von $[b, c], [c, a]$ bzw. $[a, b]$, so betrachten wir die 4 Dreiecke

$$\Delta_{a,c',b'},\ \Delta_{b,a',c'},\ \Delta_{c,b',a'} \quad \text{und} \quad \Delta_{a',b',c'}\,.$$

Dann gilt (beachte $\int\limits_{[z_1,z_2]} f = -\int\limits_{[z_2,z_1]} f$ nach B.30.2.1)

$$I = \int\limits_{\partial\Delta_{a,c',b'}} f + \int\limits_{\partial\Delta_{b,a',c'}} f + \int\limits_{\partial\Delta_{c,b',a'}} f + \int\limits_{\partial\Delta_{a',b',c'}} f\,.$$

Also ist der Betrag zumindest eines der Integrale auf der rechten Seite $\geq |I|/4$. Das entsprechende Dreieck heiße Δ_1.
Nun gehen wir analog mit Δ_1 an Stelle von Δ vor. Induktiv erhalten wir so eine Folge von Dreiecken Δ_n mit $\Delta \supset \Delta_1 \supset \Delta_2 \supset \ldots$ und Länge $(\partial\Delta_n) = L/2^n$, wobei L die Länge von $\partial\Delta$ bezeichnet, und so, dass gilt

$$|I| \leq 4^n \Big|\int\limits_{\partial\Delta_n} f\Big| \qquad (n \in \mathbb{N})\,.$$

Es gilt

$$\bigcap_{n\in\mathbb{N}} \Delta_n = \{z_0\}$$

für ein $z_0 \in \Delta \subset G$. Da $z_0 \neq p$ ist, ist f differenzierbar an z_0. Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $\delta > 0$ mit

$$|f(z) - f(z_0) - f'(z_0)(z - z_0)| < \varepsilon |z - z_0|$$

für alle $z \in U_\delta(z_0)$ (Zerlegungsformel). Außerdem ist $\Delta_n \subset U_\delta(z_0)$ für alle $n \geq n_0$. Für diese n gilt

$$\begin{aligned} \Big|\int\limits_{\partial\Delta_n} f\Big| &\overset{B.30.3.1}{=} \Big|\int\limits_{\partial\Delta_n} [f(z) - f(z_0) - f'(z_0)(z - z_0)]\,dz\Big| \\ &\leq \varepsilon \int\limits_{\partial\Delta_n} \underbrace{|z - z_0|}_{\leq 2^{-n}L} \leq \varepsilon \cdot (2^{-n}L)^2\,. \end{aligned}$$

Also erhalten wir

$$|I| \leq \varepsilon L^2\,.$$

Da $\varepsilon > 0$ beliebig war, ist $I = 0$.
Nun betrachten wir den Fall $p \in \Delta$. Ist p ein Eckpunkt von Δ, etwa $p = a$, so wählen wir Punkte $x_n \in \overline{a,b}$ und $y_n \in \overline{a,c}$ mit $x_n \to a, y_n \to a\,(n \to \infty)$. Dann gilt mit dem bereits Bewiesenen

$$|\int\limits_{\partial\Delta} f| = |\int\limits_{\partial\Delta_{a,x_n,y_n}} f + \underbrace{\int\limits_{\partial\Delta_{x_n,b,y_n}} f}_{=0} + \underbrace{\int\limits_{\partial\Delta_{y_n,b,c}} f}_{=0} | \leq ||f||_{\infty,\Delta} \cdot \text{Länge}(\partial\Delta_{a,x_n,y_n}) \to 0\,,$$

also $\int\limits_{\partial\Delta} f = 0$ auch hier.
Ist $p \in \Delta^0$, so wenden wir das eben Bewiesene auf die Dreiecke $\Delta_{a,b,p}$, $\Delta_{c,a,p}$ und $\Delta_{b,c,p}$ an. Ist $p \in \partial\Delta$ kein Eckpunkt, etwa $p \in \overline{a,b}$, so betrachtet man $\Delta_{a,p,c}$ und $\Delta_{c,p,b}$.
2. Es sei G sternförmig bezüglich z_*. Wir definieren $F : G \to \mathbb{C}$ durch

$$F(z) := \int\limits_{[z_*,z]} f \qquad (z \in G)\,.$$

Es sei $z_0 \in \Omega$ gegeben. Da G offen ist, existiert ein $r > 0$ mit $U_r(z_0) \subset \Omega$. Dann gilt auch: Ist $z \in U_r(z_0)$, so liegt das Dreieck Δ_{z_*,z,z_0} in G (da G sternförmig ist). Also erhalten wir mit

$$\begin{aligned} 0 &= \int\limits_{\partial\Delta_{z_*,z,z_0}} f = \int\limits_{[z_*,z]} f + \int\limits_{[z,z_0]} f - \int\limits_{[z_*,z_0]} \\ &= F(z) - F(z_0) + \int\limits_{[z,z_0]} f \end{aligned}$$

d.h.

$$F(z) - F(z_0) = -\int\limits_{[z,z_0]} f = \int\limits_{[z_0,z]} f\,.$$

Genau wie im Beweis zu S. 30.5 sieht man nun: F ist differenzierbar auf G mit $F' = f$ auf G. Aus S.30.5 folgt die Behauptung. □

Bemerkung und Definition 30.9 Man sieht leicht, dass die Aussage von S. 30.8 nicht für jedes Gebiet $G \subset \mathbb{C}$ gilt: Ist etwa $G = \mathbb{C} \setminus \{0\}$ und $f(z) = 1/z\,(z \in G)$, so ist f holomorph in G, aber es gilt

$$\int\limits_{|z|=1} \frac{dz}{z} = 2\pi i \neq 0\,.$$

Wir sagen, ein Gebiet $G \subset C$ sei *(homologisch) einfach zusammenhängend*, falls für alle f wie in S.30.8 und alle geschlossenen Pfade γ in G gilt $\int\limits_{\gamma} f = 0$ (also die Aussage von

S.30.8). S. 30.8 zeigt, dass sternförmige Gebiete jedenfalls einfach zusammenhängend sind.

Als erste Anwendung des Cauchyschen Integralsatzes beweisen wir eine wichtige Integraldarstellung für holomorphe Funktionen. Dazu gibt es vorbereitend

Bemerkung und Definition 30.10 Es sei (X, d) ein metrischer Raum, und es sei $M \subset X$. Für $x \in M$ heißt

$$Z_M(x) := \bigcup \{A \subset M : x \in A, A \text{ zusammenhängend}\}$$

(Zusammenhangs-)Komponente von M (bezüglich x). Man sieht leicht ([Ü]): $Z_M(x)$ ist zusammenhängend, und für $x, y \in M$ gilt entweder $Z_M(x) = Z_M(y)$ oder $Z_M(x) \cap Z_M(y) = \emptyset$.
Ist speziell $\Omega \subset \mathbb{C}$ offen, so existieren höchstens abzählbar viele paarweise disjunkte Komponenten $(G_\alpha)_{\alpha \in I}$ von Ω mit

$$\Omega = \bigcup_{\alpha \in I} G_\alpha$$

([Ü]). Die G_α sind dann jeweils Gebiete.
Außerdem gilt: Ist $K \subset \mathbb{C}$ kompakt, so hat die offene Menge $\mathbb{C} \setminus K$ genau eine unbeschränkte Komponente (nämlich die, die $\{z : |z| > \sup\{|\zeta| : \zeta \in K\}\}$ enthält).

Satz 30.11 *Es sei $\gamma : [\alpha, \beta] \to \mathbb{C}$ ein geschlossener Pfad, und es sei $\Gamma := \gamma([\alpha, \beta])$. Dann ist $\operatorname{ind}_\gamma : \mathbb{C} \setminus \Gamma \to \mathbb{C}$, definiert durch*

$$\operatorname{ind}_\gamma(z) := \frac{1}{2\pi i} \int_\gamma \frac{d\zeta}{\zeta - z} \qquad (z \in \mathbb{C} \setminus \Gamma),$$

eine analytische Funktion mit ganzzahligen Werten (d.h. $\operatorname{ind}_\gamma(\mathbb{C} \setminus \Gamma) \subset \mathbb{Z}$). Außerdem gilt $\operatorname{ind}_\gamma(z) = \text{const}$ auf jeder Komponente von $\mathbb{C} \setminus \Gamma$ und $\operatorname{ind}_\gamma(z) \equiv 0$ auf der unbeschränkten Komponente von $\mathbb{C} \setminus \Gamma$.

Beweis. 1. Es gilt

$$\operatorname{ind}_\gamma(z) = \frac{1}{2\pi i} \int_\alpha^\beta \frac{\gamma'(s)}{\gamma(s) - z}\, ds \qquad (z \in \mathbb{C} \setminus \Gamma).$$

Also ist ind_γ nach S.29.9 analytisch in $\mathbb{C} \setminus \Gamma$. Für $w \in \mathbb{C}$ sieht man leicht ([Ü]), dass $w/2\pi i \in \mathbb{Z}$ äquivalent ist zu $e^w = 1$. Also ist die Aussage, dass ind_γ ganzzahlige Werte hat, äquivalent dazu $\varphi = \varphi_z : [\alpha, \beta] \to \mathbb{C}$ mit

$$\varphi(t) := \exp\Big(\int\limits_\alpha^t \frac{\gamma'(s)}{\gamma(s) - z} ds\Big) \qquad (t \in [\alpha, \beta])$$

für alle $z \in \mathbb{C} \setminus \Gamma$ die Bedingung $\varphi(\beta) = 1$ erfüllt.
Es gilt

$$\frac{\varphi'(t)}{\varphi(t)} = \frac{\gamma'(t)}{\gamma(t) - z}$$

für alle $t \in [\alpha, \beta]$ bis auf eine endliche Ausnahmemenge S. Für die stetige Funktion $\varphi/(\gamma - z)$ gilt also

$$\Big(\frac{\varphi}{\gamma - z}\Big)' = \frac{\varphi'(\gamma - z) - \varphi\gamma'}{(\gamma - z)^2} \equiv 0 \quad \text{auf } [\alpha, \beta] \setminus S$$

Hieraus folgt, dass $\varphi(t)/(\gamma(t) - z) = \mathrm{const}$ auf $[\alpha, \beta]$ ist. Aus $\varphi(\alpha) = 1$ ergibt sich

$$\varphi(t) = \frac{\gamma(t) - z}{\gamma(\alpha) - z} \qquad (t \in [\alpha, \beta])$$

und, da γ geschlossen ist, erhält man $\varphi(\beta) = 1$.
2. Es sei G eine Komponente von $\mathbb{C} \setminus \Gamma$. Da G zusammenhängend und ind_γ stetig und ganzzahlig auf G ist, ist $\mathrm{ind}_\gamma(z) \equiv \mathrm{const}$ in G nach S.26.22.
Da Γ kompakt ist, existiert weiter $R := \max\limits_{\zeta \in \Gamma} |\zeta|$. Es gilt für $|z| > R$

$$|\mathrm{ind}_\gamma(z)| \leq \frac{1}{|z| - R} \frac{L(\gamma)}{2\pi} .$$

Also ist $|\mathrm{ind}_\gamma(z)| < 1$ für $|z|$ genügend groß. Nach Beweisschritt 1 muss dann $\mathrm{ind}_\gamma(z) \equiv 0$ für $|z|$ genügend groß sein. Da $\mathrm{ind}_\gamma(z) \equiv \mathrm{const}$ auf der unbeschränkten Komponente von $\mathbb{C} \setminus \Gamma$ ist, ist $\mathrm{ind}_\gamma(z) \equiv 0$ dort. □

Damit können wir die angekündigte Integraldarstellung für holomorphe Funktionen, jedenfalls in sternförmigen Gebieten, formulieren.

Satz 30.12 *(Cauchysche Integralformel)*
Es seien $G \subset \mathbb{C}$ ein einfach zusammenhängendes Gebiet, $f \in H(G)$ und γ ein geschlossener Pfad in G. Dann gilt

$$f(z) \cdot \mathrm{ind}_\gamma(z) = \frac{1}{2\pi i} \int\limits_\gamma \frac{f(\zeta)}{\zeta - z} d\zeta \qquad (z \in G \setminus \Gamma)$$

(wobei Γ das Bild von γ ist).

Beweis. Es sei $z \in \mathbb{C} \setminus \Gamma$ fest. Dann gilt für die Funktion $g : G \to \mathbb{C}$ mit

$$g(\zeta) := \begin{cases} \dfrac{f(\zeta) - f(z)}{\zeta - z} & , \quad \zeta \neq z \\ \\ f'(z) & , \quad \zeta = z \end{cases}$$

nach Voraussetzung: g ist stetig auf G und holomorph in $G \setminus \{z\}$.
Da G einfach zusammenhängend ist, folgt

$$\begin{aligned} 0 &= \int\limits_\gamma \frac{f(\zeta) - f(z)}{\zeta - z}\, d\zeta = \int\limits_\gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta - f(z) \int\limits_\gamma \frac{d\zeta}{\zeta - z} = \\ &= \int\limits_\gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta - 2\pi i\, f(z) \cdot \operatorname{ind}_\gamma(z)\,. \end{aligned}$$

und damit die Behauptung. □

Bemerkung 30.13 Ist speziell $\gamma(t) = z_0 + Re^{it}$ für $t \in [0, 2\pi]$, also $\Gamma = K_R(z_0) = \{z : |z - z_0| = R\}$, so ist

$$\operatorname{ind}_\gamma(z_0) = \frac{1}{2\pi i} \int\limits_{|\zeta - z_0| = R} \frac{d\zeta}{\zeta - z_0} = 1\,,$$

also ist nach S. 30.11

$$\operatorname{ind}_\gamma(z) = \begin{cases} 1 & , \quad \text{falls } |z - z_0| < R \\ 0 & , \quad \text{falls } |z - z_0| > R \end{cases}\,.$$

Ist Ω eine beliebige offene Menge mit $\overline{U_R(z_0)} \subset \Omega$ und ist $f \in H(\Omega)$ so folgt aus S. 30.12 (angewandt auf ein sternförmiges Gebiet G mit $\overline{U_R(z_0)} \subset G \subset \Omega$)

$$f(z) = \frac{1}{2\pi i} \int\limits_{|\zeta - z_0| = R} \frac{f(\zeta)}{\zeta - z} d\zeta \qquad (z \in U_R(z_0)) \tag{30.2}$$

Man beachte: (30.2) zeigt insbesondere, dass die Funktionswerte in $U_R(z_0)$ bereits vollständig durch die Werte am Rand $K_R(z_0)$ festgelegt sind!
Wählt man speziell $z = z_0$, so ergibt sich

$$f(z_0) = \frac{1}{2\pi i} \int\limits_0^{2\pi} \frac{f(z_0 + Re^{it})}{Re^{it}}\, i\, Re^{it} dt = \frac{1}{2\pi} \int\limits_0^{2\pi} f(z_0 + Re^{it}) dt\,. \tag{30.3}$$

Also: Der Funktionswert im Kreismittelpunkt ergibt sich als „Integralmittel“ der Funktionswerte am Rand des Kreises.

Mit der Cauchyschen Integralformel zur Hand können wir reichlich ernten. Als erstes formulieren wir den schon eingangs des Abschnitts angekündigten

Satz 30.14 *Es sei $\Omega \subset \mathbb{C}$ offen, und es sei $f : \Omega \to \mathbb{C}$. Dann gilt*

$$f \text{ holomorph in } \Omega \Leftrightarrow f \text{ analytisch in } \Omega.$$

Außerdem gilt dann: Für alle $z_0 \in \Omega$ ist (mit $\operatorname{dist}(z_0, \emptyset) := \infty$)

$$f(z) = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(z_0)}{\nu!}(z - z_0)^\nu \qquad \text{in } |z - z_0| < \operatorname{dist}(z_0, \partial\Omega)\,.$$

Beweis. 1. Es sei f holomorph in Ω, und es sei $z_0 \in \Omega$. Ist $R < \operatorname{dist}(z_0, \partial\Omega)$, so gilt $\overline{U_R(z_0)} \subset \Omega$. Also besagt (30.2), dass für alle $z \in U_R(z_0)$ gilt

$$f(z) = \frac{1}{2\pi i} \int\limits_{|\zeta - z_0| = R} \frac{f(\zeta)}{\zeta - z} d\zeta = \frac{1}{2\pi} \int\limits_0^{2\pi} \frac{f(z_0 + Re^{it})}{z_0 + \mathrm{Re}^{it} - z} Re^{it} dt\,.$$

Nach S. 29.9 (angewandt mit $[a, b] = [0, 2\pi], \psi(t) = z_0 + Re^{it}$ und $\varphi(t) = f(z_0 + Re^{it})Re^{it}/2\pi$) gilt

$$f(z) = \sum_{\nu=0}^{\infty} a_\nu (z - z_0)^\nu = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(z_0)}{\nu!}(z - z_0)^\nu$$

für alle $z \in U_R(z_0)$. Da $R < \operatorname{dist}(z_0, \partial\Omega)$ beliebig war, gilt die Darstellung in $|z - z_0| < \operatorname{dist}(z_0, \partial\Omega)$. Insbesondere ist f auch analytisch in Ω.
2. Ist f analytisch in Ω, so ist f holomorph in Ω nach B.29.8. □

Bemerkung 30.15 Kombiniert man S. 30.14 und S. 29.6, so ergibt sich insbesondere, dass jede Funktion $f \in H(\Omega)$ beliebig oft differenzierbar auf Ω ist. Außerdem gilt dann auch folgende verallgemeinerte Cauchysche Integralformel: Ist $G \subset \mathbb{C}$ einfach zusammenhängend und γ ein geschlossener Pfad in G, so gilt für alle $k \in \mathbb{N}_0$

$$f^{(k)}(z) \cdot \operatorname{ind}_\gamma(z) = \frac{k!}{2\pi i} \int\limits_\gamma \frac{f(\zeta)}{(\zeta - z)^{k+1}} d\zeta \qquad (z \in G \setminus \Gamma)\,,$$

wobei Γ das Bild von γ ist. Insbesondere ergibt sich wieder für alle $z_0 \in \Omega$ und $R < \operatorname{dist}(z_0, \partial\Omega)$

$$f^{(k)}(z) = \frac{1}{2\pi i} \int\limits_{|\zeta - z_0| = R} \frac{f(\zeta)}{(\zeta - z)^{k+1}} d\zeta \qquad (z \in U_R(z_0))\,.$$

(Denn: Ist $\gamma : [\alpha, \beta] \to G$ ein geschlossener Pfad in G, so gilt nach der Cauchyschen Integralformel

$$f(z) \cdot \operatorname{ind}_\gamma(z) = \frac{1}{2\pi i} \int\limits_\gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta = \frac{1}{2\pi i} \int\limits_\alpha^\beta \frac{f(\gamma(t))\gamma'(t)}{\gamma(t) - z}\, dt$$

für alle $z \in G \setminus \Gamma$. Ist $z \in G \setminus \Gamma$, so ist $\operatorname{ind}_\gamma(w) = \text{const} = c$ auf einer Umgebung von z (S. 30.11). Also folgt aus S. 29.9

$$f^{(k)}(z) \cdot c = \frac{k!}{2\pi i} \int\limits_\alpha^\beta \frac{f(\gamma(t))\gamma'(t)}{(\gamma(t) - z)^{k+1}}\, dt = \frac{k!}{2\pi i} \int\limits_\gamma \frac{f(\zeta)}{(\zeta - z)^{k+1}}\, d\zeta\,.$$

Da $z \in G \setminus \Gamma$ beliebig war, ergibt sich die Behauptung.)

Wir haben oben gesehen, dass das Verschwinden der Integrale über den Rand von Dreiecken für holomorphe Funktionen, also die Gleichung (30.1), der Schlüssel zum Beweis des Cauchyschen Integralsatzes ist.
Interessanterweise gilt auch eine gewisse Umkehrung dieses Ergebnisses:

Satz 30.16 *Es sei $\Omega \subset \mathbb{C}$ offen, und es sei $f : \Omega \to \mathbb{C}$ stetig. Existiert zu jedem $z_0 \in \Omega$ ein $r > 0$ so, dass*

$$\int\limits_{\partial\Delta} f = 0$$

für alle Dreiecke Δ in $U_r(z_0)$ gilt, so ist f holomorph in Ω.

Beweis. Es sei $z_0 \in \Omega$ und $G := U_r(z_0)$. Dann ist insbesondere G sternförmig. Im Beweisschritt 2 zu S. 30.8 haben wir lediglich die Stetigkeit von f und die Bedingung (30.1) verwandt. Da beides jetzt nach Voraussetzung gilt, existiert nach erwähntem Beweisschritt 2 ein $F \in H(G)$ mit $F' = f$ auf G. Also ist $f \in H(G)$ nach B. 30.15. Da $z_0 \in \Omega$ beliebig war, ist $f \in H(\Omega)$. □

Bemerkung 30.17 Insbesondere ergibt sich aus S. 30.16 auch, dass jede Funktion $f : \Omega \to \mathbb{C}$, die stetig auf Ω und holomorph in $\Omega \setminus \{p\}$ für ein $p \in \Omega$ ist, bereits „automatisch“ holomorph in Ω ist.
(Denn: Ist $z_0 \in \Omega$ und $G = U_r(z_0) \subset \Omega$, so gilt nach dem Cauchyschen Integralsatz

$$\int\limits_{\partial\Delta} f = 0$$

für alle Dreiecke Δ in G. Also ist f holomorph in G nach S. 30.16.)

Zum Abschluss dieses Abschnitts wollen wir uns mit Logarithmen und allgemeinen Potenzen in $\mathbb{C}$ beschäftigen.
In Abschnitt 7 hatten wir die reelle Logarithmusfunktion als Umkehrung der (reellen) Exponentialfunktion erkannt. Schon die Tatsache, dass die Exponentialfunktion im Komplexen nicht mehr injektiv ist, deutet an, dass die Situation hier komplizierter wird. Es gilt jedenfalls

Satz 30.18 *Es sei $G \subset \mathbb{C}$ ein einfach zusammenhängendes Gebiet, und es sei $g \in H(G)$ mit $g(z) \neq 0$ für alle $z \in G$. Ferner sei $z_0 \in G$ und $g(z_0) = r_0 e^{i\varphi_0}$. Dann gilt*

1. *Es existiert eine Funktion $f \in H(G)$ mit $f(z_0) = \ln r_0 + i\varphi_0$ und*

$$e^{f(z)} = g(z) \qquad \text{für alle } z \in G\,.$$

2. *Ist $\tilde{f} : G \to \mathbb{C}$ stetig mit $e^{\tilde{f}(z)} = g(z)$ für alle $z \in G$, so existiert ein $k \in \mathbb{Z}$ mit*

$$\tilde{f}(z) = f(z) + 2k\pi i \qquad (z \in G)\,.$$

Beweis. 1. Da $g(z) \neq 0$ ist, ist g'/g holomorph in G. Da G einfach zusammenhängend ist, existiert nach S. 30.5 eine Funktion $f \in H(G)$ mit $f' = g'/g$. Dann gilt

$$(e^f)' = e^f \cdot f' = e^f g'/g \quad \text{in } G,$$

d.h.

$$(e^f)'g - e^f g' \equiv 0 \quad \text{in } G.$$

Also ist

$$\frac{e^{f(z)}}{g(z)} \equiv \text{const in } G.$$

Ist $z_0 \in G$ und $g(z_0) = r_0 e^{i\varphi_0}$, so können wir f so wählen, dass $f(z_0) = \ln r_0 + i\varphi_0$ gilt (Addition einer geeigneten Konstante). Dann erhalten wir

$$e^{f(z_0)} = r_0 e^{i\varphi_0} = g(z_0)$$

und damit

$$e^{f(z)} = g(z) \qquad \text{für alle } z \in G.$$

2. Ist $\tilde{f} \in C(G)$ mit $e^{\tilde{f}} = g$, so gilt

$$e^{\tilde{f}(z) - f(z)} = e^{\tilde{f}(z)}/e^{f(z)} \equiv 1 \text{ in } G.$$

Damit ist
$$\varphi(z) = \frac{\tilde{f}(z) - f(z)}{2\pi i} \in \mathbb{Z}$$
für alle $z \in G$. Da G zusammenhängend und φ stetig auf G ist, ist $\varphi(z) \equiv \text{const}$ auf G, d.h. es existiert ein $k \in \mathbb{Z}$ mit
$$\tilde{f}(z) = f(z) + 2k\pi i \qquad (z \in G)\,. \qquad \square$$

Bemerkung und Definition 30.19 Jede Funktion $f \in H(G)$ mit $e^f = g$ in G heißt *Zweig des Logarithmus* von g in G. Ist f ein solcher Zweig, so ist auch $\tilde{f}$ mit $\tilde{f}(z) = f(z) + 2k\pi i$ für ein $k \in \mathbb{Z}$ ein Zweig. Nach S. 30.18.2 sind durch diese (abzählbar unendlich vielen) Funktionen alle Zweige gegeben. Außerdem zeigt S. 30.18.2 auch, dass jede stetige Funktion f mit $e^f = g$ „automatisch“ holomorph in G ist.

Beispiel 30.20 Es sei
$$G_- := \mathbb{C} \setminus (-\infty, 0]\,.$$
Dann ist G_- sternförmig, also einfach zusammenhängend nach dem Cauchyschen Integralsatz. Ist $g(z) = z\,(z \in G_-)$, so existiert nach S. 30.18.1 eine in Funktion $f \in H(G_-)$ mit $f(1) = 0$ und
$$e^{f(z)} = z \qquad \text{für alle } z \in G_-.$$
Ist $z \in G_-$, so existieren eindeutig bestimmte $r > 0$ und $\varphi \in (-\pi, \pi)$ mit $z = re^{i\varphi}$. Die Abbildung $p : G_- \to (0, \infty) \times (-\pi, \pi)$ mit
$$p(z) = (r, \varphi) \qquad (z = re^{i\varphi} \in G_-)$$
ist stetig auf G_- (vgl. B.21.4). Setzt man $\arg z := \varphi$, so ist auch $\tilde{f} : G_- \to \mathbb{C}$ mit
$$\tilde{f}(z) = \ln r + i\varphi = \ln|z| + i \arg z \qquad (z \in G_-)$$
stetig. Weiter gilt natürlich auch
$$e^{\tilde{f}(z)} = e^{\ln r + i\varphi} = re^{i\varphi} = z \qquad (z \in G_-)\,.$$
Also ist $\tilde{f}$ ein Zweig des Logarithmus in G_-, und da $\tilde{f}(1) = 0 = f(1)$ gilt, ist $f(z) \equiv \tilde{f}(z)$ in G_-. Für $z = r > 0$ haben wir insbesondere $f(r) = \ln r$, d.h. dieser Zweig setzt den „reellen Logarithmus“ ln holomorph auf G_- fort.
Wir nennen f den Hauptzweig des Logarithmus (von z) in G_- und schreiben dafür auch
$$f(z) =: \log z \qquad (z \in G_-)\,.$$

Nach S. 30.18.2 sind alle weiteren Zweige von der Form

$$z \mapsto \log z + 2k\pi i = \ln r + i(\varphi + 2k\pi)$$

für ein $k \in \mathbb{Z}$.
In B.14.12.1 hatten wir gesehen, dass für $x \in (-1, 1)$ gilt

$$\ln(1 + x) = \sum_{\nu=1}^{\infty} \frac{(-1)^{\nu-1} x^{\nu}}{\nu}\,.$$

Da $z \mapsto \log(1+z)$ holomorph in $\mathbb{C}\backslash(-\infty, -1]$ ist, gilt nach dem Identitätssatz (S.29.11) allgemeiner

$$\log(1 + z) = \sum_{\nu=1}^{\infty} \frac{(-1)^{\nu-1} z^{\nu}}{\nu}$$

für alle $z \in \mathbb{C}$ mit $|z| < 1$.
Wie steht es mit der Gültigkeit der Funktionalgleichung S.5.16.2 für komplexe Zahlen z und w?
Ist $z = re^{i\varphi}$, $w = \varrho e^{i\vartheta}$ mit $\varphi, \vartheta \in (-\pi, \pi)$ und $\varphi + \vartheta \in (-\pi, \pi)$, so ist $zw = r\varrho e^{i(\varphi+\vartheta)}$. Es gilt also

$$\begin{aligned} \log(zw) &= \ln(r\varrho) + i(\varphi + \vartheta) = \ln r + i\varphi + \ln \varrho + i\vartheta = \\ &= \log z + \log w\,. \end{aligned}$$

Ist jedoch etwa $\varphi + \vartheta > \pi$, so ist $zw = r\varrho e^{i(\varphi+\vartheta-2\pi)}$, also

$$\begin{aligned} \log(zw) &= \ln(r\varrho) + i(\varphi + \vartheta - 2\pi) = \\ &= \log z + \log w - 2\pi i\,. \end{aligned}$$

Es kommt also ein „Korrekturterm“ $2\pi i$ hinzu. Das Beispiel zeigt, dass Vorsicht im Umgang mit komplexen Logarithmen angebracht ist!

Wie im Reellen definieren wir allgemein Potenzen mit Hilfe von Logarithmen. Wir beschränken uns dabei auf Potenzen, die unter Verwendung des Hauptzweiges definiert sind.

Definition 30.21 Es sei $G_- = \mathbb{C} \setminus (-\infty, 0]$, und es sei $b \in \mathbb{C}$. Wir setzen

$$z^b := e^{b \cdot \log z} \qquad (z \in G_-)\,.$$

Ist speziell $b = 1/k$ für ein $k \in \mathbb{N}, k \geq 2$, so schreiben wir auch $\sqrt[k]{z}$ an Stelle von $z^{1/k}$, und ist $k = 2$, so schreiben wir kurz $\sqrt{z}$. Die Funktion $z \mapsto \sqrt[k]{z}$ heißt *Hauptzweig der k-ten Wurzel* (für $k = 2$ *Hauptzweig der Wurzel*) (von z) in G_-.
Für $z = r > 0$ stimmt diese Definition nach B.30.20 mit der aus D. 7.8 überein.

Bemerkung 30.22 1. Für $z \in G_-$ und $b_1, b_2 \in \mathbb{C}$ gilt

$$z^{b_1+b_2} = z^{b_1} z^{b_2} .$$

Weiter ist $z \mapsto z^b$ holomorph in G_- mit

$$(z^b)' = b \cdot z^{b-1} .$$

2. Ist $z = re^{i\varphi} \in G_-$ mit $\varphi \in (-\pi, \pi)$, so gilt $\sqrt[k]{z} = \sqrt[k]{r} e^{i\varphi/k}$. Andere Zweige der k-ten Wurzel erhält man durch Verwendung anderer Zweige des Logarithmus. Außerdem kann man allgemeine Potenzen auch in anderen einfach zusammenhängenden Gebieten betrachten.

31 Anwendungen der Cauchyschen Integralformel

Wir werden nun einige zentrale funktionentheoretische Konsequenzen aus der Cauchyschen Integralformel herleiten.

Eine erste, einfache Folgerung ist

Satz 31.1 *(Cauchysche Ungleichung)*
Es sei $\Omega \subset \mathbb{C}$ offen, $f \in H(\Omega)$, und es seien $z_0 \in \Omega$ und $R > 0$ so, dass $\overline{U_R(z_0)} \subset \Omega$ gilt. Dann ist für $0 \le r < R$

$$|f^{(k)}(z)| \le \frac{k!R}{(R-r)^{k+1}} \max_{\zeta \in K_R(z_0)} |f(\zeta)| \qquad (k \in \mathbb{N}_0,\, |z - z_0| \le r)\,.$$

Beweis. Nach B. 30.15 gilt für $|z - z_0| \le r$

$$\begin{aligned} |f^{(k)}(z)| &= \Big|\frac{k!}{2\pi i} \int\limits_{K_R(z_0)} \frac{f(\zeta)}{(\zeta - z)^{k+1}}\, d\zeta\Big| \\ &\le \max_{\zeta \in K_R(z_0)} |f(\zeta)| \frac{k!}{2\pi(R-r)^{k+1}}\, 2\pi R\,, \end{aligned}$$

also die Behauptung. □

Bemerkung und Definition 31.2 Eine in $\mathbb{C}$ holomorphe Funktion f heißt *ganze Funktion*. Ist f ganz, so gilt nach S. 30.14 für alle $z_0 \in \mathbb{C}$ (da $\operatorname{dist}(z_0, \emptyset) = \infty$)

$$f(z) = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(z_0)}{\nu!} (z - z_0)^{\nu} \qquad (z \in \mathbb{C})\,.$$

Satz 31.3 *(Liouville)*
Ist f ganz und nicht konstant, so existiert eine Folge (z_n) in $\mathbb{C}$ mit $f(z_n) \to \infty$ für $n \to \infty$, d.h. $f(\mathbb{C})$ ist unbeschränkt.

Beweis. Angenommen, es existiert ein $M > 0$ mit $|f(z)| \le M$ für alle $z \in \mathbb{C}$. Wir zeigen: Dann ist $f \equiv \text{const}$.
Nach der Cauchyschen Ungleichung gilt für alle $k \in \mathbb{N}$, $R > 0$

$$|f^{(k)}(0)| \le \frac{k!}{R^k} M \to 0 \qquad (R \to \infty)\,.$$

Also ist $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}$, d.h.

$$f(z) = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(0)}{\nu!} z^\nu = f(0) \qquad (z \in \mathbb{C})\,.$$

□

Bemerkung 31.4 Aus dem Satz von Bolzano-Weierstraß ergibt sich sofort, dass für jede Folge (z_n) mit $f(z_n) \to \infty$ auch $z_n \to \infty$ gelten muss.

Beispiel 31.5 1. Es sei $P(z) = \sum\limits_{\nu=0}^{n} a_\nu z^\nu$ ein Polynom vom Grad $n \geq 1$. Dann ist P ganz und $P \not\equiv \text{const}$. Also existiert eine Folge (z_n) mit $P(z_n) \to \infty$ $(n \to \infty)$. Tatsächlich gilt bei Polynomen viel mehr; denn mit $b_\nu := a_\nu / a_n$ ist

$$\frac{P(z)}{a_n z^n} = 1 + \underbrace{\sum_{\nu=0}^{n-1} \frac{b_\nu}{z^{n-\nu}}}_{\to 0\,(z \to \infty)} \to 1 \qquad (z \to \infty)\,,$$

d.h. zu jedem $\varepsilon > 0$ existiert ein $R > 0$ mit

$$|a_n||z|^n(1+\varepsilon) > |P(z)| > |a_n||z|^n(1-\varepsilon) \qquad (|z| > R)\,.$$

Insbesondere gilt also $P(z) \to \infty\,(z \to \infty)$.
2. Ist $f(z) = \sin z$, so gilt etwa für $z_n = in$

$$f(z_n) = \frac{1}{2\pi}(e^{-n} - e^n) \to \infty \qquad (n \to \infty)\,.$$

Als interessante Folgerung aus S. 31.3 erhalten wir

Satz 31.6 *(Fundamentalsatz der Algebra)*
Es sei $P(z) = \sum\limits_{\nu=0}^{n} a_\nu z^\nu$ *ein Polynom vom Grad* $n \in \mathbb{N}$. *Dann existieren* $z_1, ..., z_n \in \mathbb{C}$ *mit*

$$P(z) = a_n \prod_{k=1}^{n} (z - z_k)\,.$$

Beweis. 1. Wir zeigen: P hat eine Nullstelle $z_1 \in \mathbb{C}$.
Angenommen, nicht, d.h. $1/P$ ist eine ganze Funktion. Dann existiert nach S. 31.3 eine Folge (z_n) in $\mathbb{C}$ mit $1/P(z_n) \to \infty\,(n \to \infty)$, also $P(z_n) \to 0\,(n \to \infty)$. Außerdem

gilt $z_n \to \infty$ nach B. 31.4. Dies widerspricht aber $P(z) \to \infty\,(n \to \infty)$.
2. Ist z_1 eine Nullstelle von P, so gilt

$$\begin{aligned} P(z) &= P(z) - P(z_1) = \sum_{\nu=1}^{n} a_\nu (z^\nu - z_1^\nu) = \\ &= (z - z_1) \sum_{\nu=1}^{n} a_\nu \sum_{k=0}^{\nu-1} z_1^k z^{\nu-1-k} = (z - z_1) Q(z)\,, \end{aligned}$$

wobei Q ein Polynom vom Grad $n-1$ mit führendem Koeffizienten a_n ist. Induktiv erhält man damit die Behauptung. □

Wir untersuchen nun Folgen holomorpher Funktionen. Es gilt folgendes „angenehme“ Resultat:

Satz 31.7 *Es sei $\Omega \subset \mathbb{C}$ offen, und es seien $f_n : \Omega \to \mathbb{C}$ stetige Funktionen. Ferner gelte $f_n \to f$ lokal gleichmäßig auf Ω. Dann gilt*

1. *Ist γ ein Pfad in Ω, so ist*

$$\int_\gamma f = \lim_{n\to\infty} \int_\gamma f_n\,.$$

2. *Sind die f_n holomorph in Ω, so ist auch f holomorph in Ω, und es gilt für alle $k \in \mathbb{N}$*

$$f_n^{(k)} \to f^{(k)} \qquad (n \to \infty)$$

lokal gleichmäßig auf Ω.

Beweis. 1. Zunächst ist f als lokal gleichmäßiger Grenzwert stetiger Funktionen ebenfalls stetig (B. 15.11).
Außerdem gilt: Ist $K \subset \Omega$ kompakt, so konvergiert (f_n) gleichmäßig auf K.
(Denn: Zu jedem $z \in \Omega$ existiert ein $\delta = \delta(z) > 0$ mit $f_n \to f$ gleichmäßig auf $U_\delta(z)$. Da

$$K \subset \bigcup_{z\in K} U_\delta(z)$$

gilt, existieren $z_1, \ldots, z_m \in K$ mit

$$K \subset \bigcup_{j=1}^{m} U_\delta(z_j)\,.$$

Ist $\varepsilon > 0$ gegeben, so existieren $N_j(\varepsilon) \in \mathbb{N}$ mit

$$|f_n(z) - f(z)| < \varepsilon \qquad (z \in U_\delta(z_j),\ n \geq N_j(\varepsilon))$$

für $j = 1, \dots, m$. Also gilt für $n \geq N_\varepsilon := \max\limits_{j=1,\dots,m} N_j(\varepsilon)$

$$|f_n(z) - f(z)| < \varepsilon \qquad (z \in K, n \geq N(\varepsilon)\,.)$$

Ist $\gamma : [\alpha, \beta] \to \Omega$ ein Pfad, so ist $\Gamma := \gamma([\alpha, \beta])$ kompakt in Ω. Aus S. 17.20 und der Definition von Pfadintegralen folgt

$$\int\limits_\gamma f = \lim_{n\to\infty} \int\limits_\gamma f_n\,.$$

2. Nach 1. ist f stetig auf Ω, und es gilt für alle Dreiecke Δ in Ω

$$\int\limits_{\partial\Delta} f = \lim_{n\to\infty} \int\limits_{\partial\Delta} f_n\,.$$

Aus dem Cauchyschen Integralsatz (angewandt etwa auf ein konvexes Gebiet G mit $\Delta \subset G \subset \Omega$) folgt $\int\limits_{\partial\Delta} f_n = 0$ für alle $n \in \mathbb{N}$, also auch $\int\limits_{\partial\Delta} f = 0$. Nach S. 30.16 ist f holomorph in Ω.

Ist $z_0 \in \Omega$, so existiert ein $R > 0$ mit $f_n \to f$ gleichmäßig auf $\overline{U_R(z_0)}$. Aus der Cauchyschen Ungleichung folgt für alle festen $k \in \mathbb{N}$

$$\max_{|z-z_0|\leq R/2} |f^{(k)}(z) - f_n^{(k)}(z)| \leq \frac{k!2^{k+1}}{R^k} \max_{\zeta \in K_R(z_0)} |f(\zeta) - f_n(\zeta)| \to 0 \qquad (n \to \infty)\,.$$

Also gilt $f_n^{(k)} \to f^{(k)}$ lokal gleichmäßig auf Ω. □

Beispiel 31.8 1. Ist $\sum\limits_{\nu=0}^{\infty} a_\nu (z - z_0)^\nu$ eine Potenzreihe mit Konvergenzradius $R > 0$, so konvergiert die Folge der Teilsummen $f_n(z) = \sum\limits_{\nu=0}^{n} a_\nu (z - z_0)^\nu$ lokal gleichmäßig auf $U_R(z_0)$ nach S. 16.5. Da die f_n hier Polynome sind, ist $f(z) = \sum\limits_{\nu=0}^{\infty} a_\nu (z - z_0)^\nu$ nach S. 31.7 holomorph in $U_R(z_0)$, und für die Ableitungen gilt

$$f^{(k)}(z) = \lim_{n\to\infty} f_n^{(k)}(z) = \sum_{\nu=0}^{n-k} (\nu+1)\dots(\nu+k)\, a_{\nu+k}\, (z - z_0)^\nu$$

lokal gleichmäßig auf $U_R(z_0)$. Dies ist wieder S. 29.6.

2. In B. 15.2.2 hatten wir die Riemannsche Zeta-Funktion $\zeta : (1, \infty) \to \mathbb{R}$ definiert. Wir setzen nun allgemein für $z \in \mathbb{C}$ mit $\operatorname{Re} z > 1$

$$\zeta(z) := \sum_{\nu=1}^{\infty} \frac{1}{\nu^z} \qquad \Big(= \sum_{\nu=1}^{\infty} \frac{1}{e^{z \cdot \ln \nu}} \Big)\,.$$

Dann konvergieren die Teilsummen $s_n(z) = \sum\limits_{\nu=1}^{n} \frac{1}{\nu^z}$ lokal gleichmäßig auf $\Omega := \{z : \operatorname{Re} z > 1\}$ ([Ü]). Da die Teilsummen ganze Funktionen sind, ist ζ holomorph in Ω nach S. 31.7.

Bemerkung 31.9 Ist für $\Omega \subset \mathbb{K}$

$$A(\Omega) := \{f : \Omega \to \mathbb{K} : f \text{ analytisch in } \Omega\},$$

so gilt im Falle $\mathbb{K} = \mathbb{C}$ nach S. 30.14

$$A(\Omega) = H(\Omega).$$

Also gilt nach S. 31.7 insbesondere: Aus $f_n \in A(\Omega)$ und $f_n \to f$ lokal gleichmäßig auf Ω folgt $f \in A(\Omega)$. Eine entsprechende Aussage gilt *nicht* im Falle $K = \mathbb{R}$.
Ist etwa $\Omega = \mathbb{R}$ und ist $f \in C(\mathbb{R})$ beliebig, so existiert nach dem Weierstraßschen Approximationssatz zu jedem $n \in \mathbb{N}$ ein Polynom P_n mit

$$\max_{[-n,n]} |f(x) - P_n(x)| < \frac{1}{n}.$$

Also konvergiert die Folge (P_n) lokal gleichmäßig auf $\mathbb{R}$ gegen f, d.h. jede auf $\mathbb{R}$ stetige Funktion ist lokal gleichmäßiger Grenzwert in $\mathbb{R}$ analytischer Funktionen.

Eines der zentralen Themen der reellen Analysis ist die Frage nach Extremstellen von Funktionen (mit Werten in $\mathbb{R}$). Da wir keine Ordnung in $\mathbb{C}$ haben, macht eine solche Fragestellung für komplexwertige Funktionen keinen Sinn. Wir können jedoch nach Extremstellen von $|f|$ suchen.
Bei holomorphen Funktionen bleibt diese meist erfolglos. Es gilt nämlich

Satz 31.10 *(Maximumprinzip; negative Formulierung)*
Es seien $G \subset \mathbb{C}$ ein Gebiet und $f \in H(G)$, $f \not\equiv \text{const}$. Dann hat $|f|$ kein lokales Maximum.

Beweis. 1. Wir zeigen zunächst: Existieren ein $z_0 \in G$ und ein $r > 0$ so, dass

$$|f|(z) \equiv \text{const}$$

auf $U_r(z_0)$ ist, so ist $f(z) \equiv \text{const}$ auf G.
Nach dem Identitätssatz (S. 29.11) genügt es dazu zu zeigen, dass $f(z) \equiv \text{const}$ auf $U_r(z_0)$ ist.
Ist $|f| \equiv 0$ auf $U_r(z_0)$, so ist dies klar.
Es sei also $|f| \equiv \text{const} \neq 0$, o. E. $|f| \equiv 1$ auf $U_r(z_0)$.
Dann existiert nach S. 30.18 eine Funktion $g \in H(U_r(z_0))$ mit $e^g = f$. Hierfür gilt

$$e^{\operatorname{Re} g} = |e^g| = |f| = 1,$$

also $\operatorname{Re} g \equiv 0$ in $U_R(z_0)$. Aus den Cauchy-Riemannschen Differenzialgleichungen (29.1) folgt, dass $g \equiv \text{const}$ und damit auch $e^g \equiv \text{const}$ auf $U_r(z_0)$ ist.
2. Es sei z_0 ein lokales Maximum von $|f|$, d.h. es existiert ein $r > 0$ mit

$$|f(z_0)| \geq |f(z)| \qquad \text{für alle } z \in U_r(z_0).$$

Angenommen, es existiert ein $z_1 \in U_r(z_0)$ mit $|f(z_1)| < |f(z_0)|$. Ist $\varrho = |z_1 - z_0|$, so gilt auf Grund der Stetigkeit von $t \mapsto f(z_0 + \varrho e^{it})$ auf $[0, 2\pi]$ und $|f(z_0 + \varrho e^{it})| \leq |f(z_0)|$

$$\frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + \varrho e^{it})| dt < |f(z_0)| \frac{1}{2\pi} \int_0^{2\pi} dt = |f(z_0)| \,,$$

also mit der Mittelwertformel (30.3)

$$|f(z_0)| \leq \frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + \varrho e^{it})| dt < |f(z_0)| \,.$$

Widerspruch! Also ist $|f| \equiv \text{const}$ auf $U_r(z_0)$. □

Als unmittelbare Folgerung erhalten wir

Satz 31.11 *(Maximumprinzip; positive Formulierung)*
Es sei $G \subset \mathbb{C}$ ein beschränktes Gebiet, und es sei $f : \overline{G} \to \mathbb{C}$ stetig auf $\overline{G}$ und holomorph in G. Dann existiert ein $z_0 \in \partial G$ mit

$$|f(z_0)| = \max_{z \in \overline{G}} |f(z)| \,.$$

Beweis. Da G beschränkt ist, ist $\overline{G} = G \cup \partial G$ kompakt. Also existiert ein $z_0 \in \overline{G}$ mit $|f(z_0)| = \max_{z \in \overline{G}} |f(z)|$ (beachte: $|f|$ stetig auf $\overline{G}$). Ist $f \equiv \text{const}$, so kann natürlich $z_0 \in \partial G$ gewählt werden. Ist $f \not\equiv \text{const}$, so gilt $z_0 \notin G$ nach S. 31.10, also $z_0 \in \partial G$. □

Bemerkung und Definition 31.12 Es sei $f : \mathbb{C} \to \mathbb{C}$ ganz. Wir setzen

$$M(r, f) := \max_{|z|=r} |f(z)| \qquad (r \geq 0) \,.$$

Dann ist

$$M(r, f) = \max_{|z| \leq r} |f(z)|$$

nach S. 31.11, also ist $M(\cdot, f)$ monoton wachsend.

Beispiel 31.13 Wir betrachten

$$f(z) = \cos z \,.$$

Hier gilt: Ist $z = x + iy$, so folgt

$$|\cos z| \leq \frac{1}{2} (|e^{iz}| + |e^{-iz}|) = \frac{1}{2}(e^{-y} + e^{y}) = \cosh(\pm y) = \cos(\pm iy) \,.$$

Da cosh auf $[0, \infty)$ monoton wächst, folgt für $|z| \le r$

$$|\cos z| \le \cosh(r) = \cos(ir)$$

d.h.

$$M(r, f) = \cosh(r) = \cos(ir)\,.$$

Bemerkung 31.14 Ist $G \subset \mathbb{C}$ ein Gebiet und ist $f \in H(G)$, so gilt natürlich für alle Nullstellen z_0 von f

$$|f(z_0)| = 0 \le |f(z)| \qquad (z \in G)\,,$$

d.h. Nullstellen sind Minima von $|f|$. Ist aber $f(z) \ne 0$ für alle $z \in G$ (d.h. $Z(f) = \emptyset$), so hat f im Falle $f \not\equiv$ const auch kein lokales Minimum in G (Minimumprinzip; negative Formulierung).
Außerdem existiert dann im Falle, dass G beschränkt ist, stets ein $z_0 \in \partial G$ mit

$$|f(z_0)| = \min_{z \in \overline{G}} |f(z)|$$

(Minimumprinzip; positive Formulierung).
Beides ergibt sich unmittelbar durch Anwendung obiger Maximumprinzipien auf $1/f$.

Wir wollen nun das lokale Abbildungsverhalten einer holomorphen Funktion etwas genauer beleuchten. Das Maximumprinzip wird sich dabei auch noch einmal als Konsequenz eines allgemeineren Resultats ergeben.

Satz 31.15 *Es sei $\Omega \subset \mathbb{C}$ offen, und es sei $f \in H(\Omega)$. Ist $z_0 \in \Omega$ mit $f'(z_0) \ne 0$, so existieren offene Umgebungen U von z_0 in Ω und V von $w_0 = f(z_0)$ in $f(\Omega)$ so, dass $f_{|U} : U \to V$ bijektiv ist mit $f'(z) \ne 0$ in U. Außerdem ist dann $f^{-1} := (f_{|U})^{-1} : V \to U$ holomorph mit*

$$(f^{-1})'(w) = \frac{1}{f'(f^{-1}(w))} \qquad (w \in V)\,.$$

Beweis. Es seien $u = \mathrm{Re} f, v = \mathrm{Im} f$. Dann folgt aus (29.1) unmittelbar ([Ü])

$$|f'(z)|^2 = \det J_{(u,v)}(x, y)\,,$$

wobei $z = x + iy \in \Omega$. Also ergeben sich der erste Teil der Aussage und die Stetigkeit von f^{-1} durch Anwendung von S. 21.3 auf $g = (u, v) : \Omega \to \mathbb{R}^2$.
(Man beachte: Da f holomorph ist, ist f' stetig auf Ω und damit ist $g \in C^1(\Omega, \mathbb{R}^2)$ nach S. 29.4).
Ist $w \in V$ fest und ist w_n eine Folge in V mit $w_n \to w$, $w_n \ne w$, gilt für $z_n = f^{-1}(w_n)$ und $z = f^{-1}(w)$ (da $z_n \ne z$ für alle n und $z_n \to z$)

$$\frac{f^{-1}(w_n) - f^{-1}(w)}{w_n - w} = \frac{z_n - z}{f(z_n) - f(z)} \to \frac{1}{f'(z)} \qquad (n \to \infty)\,.$$

Folglich ist f^{-1} differenzierbar an w mit

$$(f^{-1})'(w) = \frac{1}{f'(z)}\,.$$

Da $w \in V$ beliebig war, ist f^{-1} holomorph in V. □

Beispiel 31.16 Wir betrachten $f(z) = z^2$ ($z \in \mathbb{C}$). Dann gilt $f'(z) = 2z \neq 0$ für alle $z \neq 0$. Ist also $z_0 \neq 0$, so existieren offene Umgebungen U von z_0 und V von z_0^2 so, dass $f_{|U} : U \to V$ bijektiv ist.
Man beachte jedoch: f ist nicht injektiv auf $\mathbb{C} \setminus \{0\}$ (da $f(z) = f(-z)$ für alle $z \in \mathbb{C}$ gilt).

Bemerkung und Definition 31.17 Es sei $\Omega \subset \mathbb{C}$ offen, und es sei $f \in H(\Omega)$. Ein Punkt $z \in \Omega$ heißt *Nullstelle der Ordnung* $m \in \mathbb{N}$ *von* f (oder m*-fache Nullstelle*), falls m wie in S. 29.10 ist, d.h. es existiert eine auf einer Umgebung U von z_0 stetige Funktion g mit $g(z_0) \neq 0$ und

$$f(z) = (z - z_0)^m g(z) \qquad (z \in U)\,.$$

Der Beweis zu S. 29.10 zeigt: g kann stets in $H(\Omega)$ gewählt werden, und außerdem ist z_0 genau dann m-fache Nullstelle, wenn $f^{(j)}(z_0) = 0$ für $j = 0, \ldots, m-1$ und $f^{(m)}(z_0) \neq 0$ gilt.

Satz 31.18 *Es sei $\Omega \subset \mathbb{C}$ offen, und es sei $f \in H(\Omega)$. Ferner sei $z_0 \in \Omega$ und $w_0 = f(z_0)$, wobei z_0 eine Nullstelle der Ordnung m von $f - w_0$ ist. Dann existieren eine offene Umgebung U von z_0 und eine in U holomorphe Funktion φ mit folgenden Eigenschaften*

1. *Für alle $z \in U$ ist*
$$f(z) = w_0 + \varphi^m(z)\,.$$
2. *φ' hat keine Nullstelle in U und φ ist eine bijektive Abbildung von U auf $U_r(0)$ für ein $r > 0$.*

Beweis. 1. Es sei $G \subset \Omega$ eine offene, sternförmige Umgebung von z_0 so, dass $f(z) \neq w_0$ für alle $z \in G \setminus \{z_0\}$ (existiert nach S. 29.10). Ist $g : G \to \mathbb{C}$ definiert durch

$$g(z) := \begin{cases} (z - z_0)^{-m}(f(z_0) - w_0) & \text{für } z \in G \setminus \{z_0\} \\ f^{(m)}(z_0)/m! & \text{für } z = z_0 \end{cases},$$

so ist $g \in H(G)$ (vgl. Beweis zu S. 29.10) mit $g(z) \neq 0$ in G. Also existiert nach S. 30.18 eine Funktion $h \in H(G)$ mit $e^h = g$. Ist $\varphi : G \to \mathbb{C}$ definiert durch

$$\varphi(z) = (z - z_0)e^{h(z)/m} \qquad (z \in G)\,,$$

so gilt

$$\varphi^m(z) = (z - z_0)^m e^{h(z)} = (z - z_0)^m g(z) = f(z) - w_0 \qquad (z \in G)\,.$$

2. Nach 1. gilt insbesondere $\varphi(z_0) = 0$ und $\varphi'(z_0) \neq 0$. Da φ' stetig ist, erhält man die Aussage 2. durch Anwendung von S. 31.15, wobei man V dort gegebenenfalls zum Kreis $U_r(0)$ „verkleinert“. □

Als unmittelbare Konsequenz erhalten wir

Satz 31.19 *(Gebietstreue holomorpher Funktionen)*
Ist $G \subset \mathbb{C}$ ein Gebiet und ist $f \in H(G)$, $f \neq \mathrm{const}$, so ist auch $f(G)$ ein Gebiet.

Beweis. Es sei $w_0 = f(z_0) \in f(G)$, und es seien U und φ wie in S. 31.18 (man beachte: jede Nullstelle von $f - w_0$ hat endliche Ordnung nach S. 29.10). Dann ist insbesondere $U_{r^m}(w_0) = f(U) \subset f(G)$, also ist $f(G)$ offen. Da f insbesondere stetig auf G ist, ist $f(G)$ auch zusammenhängend nach S. 26.21. □

Bemerkung 31.20 Ist $G \subset \mathbb{C}$ ein Gebiet und ist $f \in H(G)$, $f \not\equiv \mathrm{const}$, so ist für alle $w_0 \in f(G)$ und alle $r > 0$ mit $U_r(z_0) \subset G$ die Menge $f(U_r(z_0))$, wobei z_0 so, dass $f(z_0) = w_0$, offen. Also existiert insbesondere stets ein $w_1 \in f(U_r(z_0))$ mit $|w_1| > |w_0|$. Damit hat $|f|$ keine lokalen Maxima in G. Dies zeigt, dass S. 31.19 das Maximumprinzip umfasst.

32 Isolierte Singularitäten und Laurent-Reihen

Bisher haben wir uns mit dem Verhalten holomorpher Funktionen $f : \Omega \to \mathbb{C}$ in der Menge Ω beschäftigt. Oft ist aber das Verhalten holomorpher Funktionen bei Annäherung an Randpunkte von Ω besonders interessant. Der einfachste Fall eines solchen Randpunktes ist der eines isolierten Punktes, mit dem wir uns jetzt genauer befassen.

Definition 32.1 Es sei $\Omega \subset \mathbb{C}$ offen und $a \in \Omega$. Ist $f \in H(\Omega \setminus \{a\})$, so heißt a eine *isolierte Singuarlität* von f.

Beispiel 32.2 1. Es sei $f : \mathbb{C} \setminus \{0\} \to \mathbb{C}$ definiert durch

$$f(z) = \frac{e^z - 1}{z} \qquad (z \neq 0)\,.$$

Dann ist $a = 0$ eine isolierte Singularität von f, und es gilt

$$\lim_{z \to 0} f(z) = 1\,.$$

(Nach B. 30.17 ist dann f zu einer in $\mathbb{C}$ holomorphen Funktion fortsetzbar, indem man $f(0) := 1$ setzt.)
2. Es sei $a \in \mathbb{C}$, $p \in \mathbb{N}$ und $f : \mathbb{C} \setminus \{a\} \to \mathbb{C}$ definiert durch

$$f(z) = \frac{1}{(z-a)^p} \qquad (z \neq a).$$

Dann ist a eine isolierte Singularität von f und es gilt

$$\lim_{z \to a} f(z) = \infty.$$

3. Es sei $f : \mathbb{C} \setminus \{0\} \to \mathbb{C}$ definiert durch

$$f(z) = e^{1/z} \qquad (z \neq 0)\,.$$

Dann ist $a = 0$ eine isolierte Singularität von f. Hier existiert $\lim\limits_{z \to a} f(z)$ nicht.

Wir zeigen im Folgenden, dass sich isolierte Singularitäten in natürlicher Weise in drei Typen einteilen lassen. Dabei gehören obige Beispiele zu jeweils unterschiedlichen Typen.

Definition 32.3 Es sei $\Omega \subset \mathbb{C}$ offen, und es sei $f \in H(\Omega \setminus \{a\})$ für ein $a \in \Omega$. Dann heißt a

1. *hebbare Singularität* von f, falls f zu einer in Ω holomorphen Funktion fortgesetzt werden kann.

2. *Pol (der Ordnung p)*, falls eine in einer offenen Umgebung U von a holomorphe Funktion g existiert mit $g(a) \neq 0$ und

$$f(z) = \frac{g(z)}{(z-a)^p} \qquad (z \in U \setminus \{a\}).$$

3. *wesentliche Singularität*, falls a weder hebbare Singularität noch Pol von f ist.

Zunächst gilt folgende Charakterisierung hebbarer Singularitäten.

Satz 32.4 *(Riemannscher Hebbarkeitssatz)*
Es sei $\Omega \subset \mathbb{C}$ offen und $f \in H(\Omega \setminus \{a\})$ für ein $a \in \Omega$. Dann sind äquivalent

a) f hat an a eine hebbare Singularität.

b) Es existiert ein $r > 0$ so, dass f in $U_r(a) \setminus \{a\}$ beschränkt ist.

Beweis. a) $\Rightarrow$ b): Ist a eine hebbare Singularität von f, so existiert eine Funktion f_0 in $H(\Omega)$ mit $f(z) = f_0(z)$ für alle $z \in \Omega \setminus \{a\}$. Insbesondere gilt dann

$$\lim_{z\to a} f(z) = \lim_{z\to a} f_0(z) = f_0(a).$$

Also existiert ein $r > 0$ so, dass $|f(z)| < |f_0(z)| + 1$ für alle $z \in U_r(a) \setminus \{a\}$.
b) $\Rightarrow$ a): Es seien $r > 0$ und $M > 0$ so, dass $|f(z)| \leq M$ für $0 < |z-a| < r$. Wir definieren $h : \Omega \to \mathbb{C}$ durch $h(a) := 0$ und

$$h(z) := (z-a)^2 f(z) \qquad (z \in \Omega \setminus \{a\}).$$

Dann gilt für $0 < |z-a| < r$

$$\left|\frac{h(z)-h(a)}{z-a}\right| = |(z-a)f(z)| \leq M|z-a| \to 0 \qquad (z \to 0).$$

Also ist h differenzierbar an der Stelle a mit $h'(a) = 0$.
Damit ist $h \in H(\Omega)$, und es gilt nach S. 30.14

$$h(z) = \sum_{\nu=0}^{\infty} \frac{h^{(\nu)}(a)}{\nu!}(z-a)^\nu = \sum_{\nu=2}^{\infty} a_\nu (z-a)^\nu$$

in $U_r(a)$ (wobei $a_\nu := h^{(\nu)}(a)/\nu!$ für $\nu \geq 2$). Folglich ist

$$f(z) = \sum_{\nu=2}^{\infty} a_\nu (z-a)^{\nu-2} = \sum_{\nu=0}^{\infty} a_{\nu+2}(z-a)^\nu$$

in $U_r(a) \setminus \{a\}$. Da die rechte Seite holomorph in $U_r(a)$ ist, ist f (durch $f(a) := a_2$) holomorph nach Ω fortsetzbar. □

Für Pole der Ordnung p gilt folgende Charakterisierung.

Satz 32.5 *Es sei $\Omega \subset \mathbb{C}$ offen, $a \in \Omega$ und $f \in H(\Omega \setminus \{a\})$. Dann sind äquivalent:*

a) f hat an a einen Pol der Ordnung p.

b) $1/f$ ist holomorph (fortsetzbar) auf einer offenen Umgebung U von a mit Nullstelle der Ordnung p an der Stelle a.

Beweis. a) $\Rightarrow$ b): Es sei $g \in H(U)$ wie in D. 32.3.2. O. E. gelte $g(z) \neq 0$ in U. Dann ist

$$\frac{1}{f(z)} = (z-a)^p \frac{1}{g(z)} \qquad (z \in U \setminus \{a\}).$$

Also ist $1/f$ holomorph fortsetzbar an der Stelle a, und die Fortsetzung hat nach B./D. 31.17 eine Nullstelle der Ordnung p an a.
b) $\Rightarrow$ a): Nach Voraussetzung gilt

$$\frac{1}{f(z)} = (z-a)^p \cdot h(z) \qquad (z \in \Omega \setminus \{a\})$$

mit einer Funktion $h \in H(\Omega)$ mt $h(a) \neq 0$. Dann ist auch $h(z) \neq 0$ auf einer offenen Umgebung U von a. Also ist

$$f(z) = \frac{1/h(z)}{(z-a)^p} \qquad (z \in U \setminus \{a\}),$$

d.h. f hat einen Pol der Ordnung p an a. □

Als Folgerung erhalten wir

Satz 32.6 *Es sei $\Omega \subset \mathbb{C}$ offen, $a \in \Omega$ und $f \in H(\Omega \setminus \{a\})$. Dann hat f an a genau dann einen Pol, wenn gilt*

$$\lim_{z \to a} f(z) = \infty$$

(d.h. $\lim_{z \to a} |f(z)| = \infty$).

Beweis. Hat f an a einen Pol, etwa der Ordnung p, so gilt mit $g \in H(U)$ wie in D. 32.3.2 (da $g(a) \neq 0$)

$$|f(z)| = \frac{|g(z)|}{|z-a|^p} \to \infty \qquad (z \to a).$$

Gilt umgekehrt

$$|f(z)| \to \infty \qquad (z \to \infty).$$

so existiert eine offene Umgebung U von a mit $f(z) \neq 0$ in U und

$$\lim_{z \to a} \frac{1}{f(z)} = 0.$$

also ist $1/f$ nach S. 32.4 (oder B. 30.17) holomorph fortsetzbar an der Stelle a mit Nullstellen, etwa der Ordnung p. Dann hat f nach S. 32.5 einen Pol der Ordnung p. □

Beispiel 32.7 Die Funktionen $\cot : \mathbb{C} \setminus Z(\sin)$ und $\tan : \mathbb{C} \setminus Z(\cos)$ sind definiert durch

$$\cot z := \frac{\cos z}{\sin z} \qquad (z \in \mathbb{C} \setminus Z(\sin))$$

und

$$\tan z := \frac{\sin z}{\cos z} \qquad (z \in \mathbb{C} \setminus Z(\cos)) .$$

Es gilt dabei ([Ü])

$$Z(\sin) = \{k\pi : k \in \mathbb{Z}\}, \quad Z(\cos) = \left\{\left(k + \frac{1}{2}\right)\pi : k \in \mathbb{Z}\right\},$$

und alle Nullstellen sind von der Ordnung 1. Damit hat auch tan an den Stellen $k\pi$ Nullstellen der Ordnung 1 und cot an den Stellen $(k + 1/2)\pi$. Nach S. 32.5 hat $\cot = 1/\tan$ Pole der Ordnung 1 an den Stellen $k\pi$ und $\tan = 1/\cot$ Pole der Ordnung 1 an den Stellen $(k + 1/2)\pi$.
Nach S. 32.6 gilt für alle $k \in \mathbb{Z}$

$$\lim_{z \to k\pi} \cot z = \lim_{z \to (k+\frac{1}{2})\pi} \tan z = \infty .$$

Bleibt noch, das Verhalten in der Nähe von wesentlichen Singularitäten zu charakterisieren. Bitte schön:

Satz 32.8 *(Casorati-Weierstraß)*
Es sei $\Omega \subset \mathbb{C}$ offen, $a \in \Omega$ und $f \in H(\Omega \setminus \{a\})$. Dann sind äquivalent:

a) f hat an a eine wesentliche Singularität.

b) Für alle offenen Umgebungen U von a in Ω gilt

$$\overline{f(U \setminus \{a\})} = \mathbb{C},$$

m.a.W. zu jedem $w \in \mathbb{C} \cup \{\infty\}$ existiert eine Folge z_n in $\Omega \setminus \{a\}$ mit $z_n \to a$ und $f(z_n) \to w$ für $n \to \infty$.

Beweis. b) ⇒ a): Gilt die Bedingung b), so ist f unbeschränkt in jeder Umgebung von a, und es gilt sicher nicht $f(z) \to \infty$ für $z \to a$. Folglich hat f an a weder eine hebbare Singularität noch einen Pol (S. 32.4 bzw. S. 32.6). Also hat f an a eine wesentliche

Singuarlität.
a) $\Rightarrow$ b): Angenommen, es existiert eine offene Umgebung U von a mit

$$\overline{f(U \setminus \{a\})} \neq \mathbb{C}\,.$$

Dann existieren ein $w \in \mathbb{C}$ und ein $\delta > 0$ so, dass $|f(z) - w| \geq \delta$ für alle $z \in U \setminus \{a\}$ gilt. Wir definieren $g : U \setminus \{a\} \to \mathbb{C}$ durch

$$g(z) := \frac{1}{f(z) - w} \qquad (z \in U \setminus \{a\})\,.$$

Dann ist $g \in H(U \setminus \{a\})$ mit $|g(z)| < 1/\delta$ für alle $z \in U, z \neq a$. Also hat g an a eine hebbare Singuarlität nach S. 32.4 (wir schreiben auch für die Fortsetzung wieder g). Ist $g(a) \neq 0$, so ist $f(z) = w + 1/g(z)$ beschränkt in einer Umgebung von a, also hat f wieder nach S. 32.4 eine hebbare Singuarlität. Widerspruch!
Ist $g(a) = 0$, so gilt

$$f(z) - w = \frac{1}{g(z)} \to \infty \qquad (z \to a)\,,$$

also auch

$$f(z) \to \infty \qquad (z \to a)\,.$$

Damit hat f an a einen Pol nach S. 32.6. Widerspruch! □

Beispiel 32.9 Wir betrachten die Funktion $f : \mathbb{C} \setminus \{0\} \to \mathbb{C}$

$$f(z) = e^{1/z} \qquad (z \neq 0)\,.$$

Für die Folge $(1/n)$ gilt $f(1/n) \to \infty$ $(n \to \infty)$ und für die Folge $(-1/n)$ gilt $f(-1/n) \to \infty$ $(n \to \infty)$. Ist $w \in \mathbb{C}, w \neq 0$ und $w = re^{i\varphi}$ mit $\varphi \in (-\pi, \pi]$, so gilt für die Folge

$$z_n = [\ln r + i(\varphi + 2n\pi)]^{-1}$$

$z_n \to 0$ $(n \to \infty)$ und

$$f(z_n) = e^{\ln r + i(\varphi + 2n\pi)} = re^{i\varphi} = w\,.$$

Also gilt hier sogar $f(z_n) = w$ für alle $n \in \mathbb{N}$ (und damit natürlich insbesondere $f(z_n) \to w$ $(n \to \infty)$). Es ist also hier tatsächlich

$$f(U \setminus \{0\}) = \mathbb{C} \setminus \{0\}$$

für alle Umgebungen U von 0, d.h. in jeder (noch so kleinen) Umgebung von 0 wird jeder Wert $w \in \mathbb{C} \setminus \{0\}$ (unendlich oft) als Funktionswert angenommen!

Wir wollen nun die Typen von isolierten Singularitäten noch in einer weiteren Form charakterisieren. Dazu betrachten wir zunächst holomorphe Funktionen in Kreisringen

$$V_{r,R}(z_0) := \{z \in \mathbb{C} : r < |z - z_0| < R\},$$

wobei $z_0 \in \mathbb{C}$ und $0 \le r < R \le \infty$.

Definition 32.10 Es sei $(a_\nu)_{\nu\in\mathbb{Z}}$ eine Folge in $\mathbb{C}$. Dann heißt die (formale) Reihe $\sum\limits_{\nu=-\infty}^{\infty} a_\nu(z - z_0)^\nu$, d.h. die Folge der n-ten Teilsummen $\big(\sum\limits_{\nu=-n}^{n} a_\nu(z - z_0)^\nu\big)_{n\in\mathbb{N}_0}$, eine *Laurent-Reihe* (mit Entwicklungsmitte $z_0 \in \mathbb{C}$ und Koeffizientenfolge (a_ν)).
Die (Potenz-)Reihe $\sum\limits_{\nu=0}^{\infty} a_\nu(z - z_0)^\nu$ heißt *Regulärteil* (oder auch *Nebenteil*) und die Reihe $\sum\limits_{\nu=-\infty}^{-1} a_\nu(z - z_0)^\nu = \sum\limits_{\nu=1}^{\infty} a_{-\nu}(z - z_0)^{-\nu}$ heißt *Hauptteil* der Laurent-Reihe.

Bemerkung 32.11 Man sieht leicht: Ist

$$\overline{\lim_{\nu\to\infty}} \sqrt[\nu]{|a_\nu|} = \frac{1}{R} \qquad \text{und} \qquad \overline{\lim_{\nu\to\infty}} \sqrt[\nu]{|a_{-\nu}|} = r\,,$$

so konvergiert $\sum\limits_{\nu=0}^{\infty} a_\nu(z-z_0)^\nu$ lokal gleichmäßig in $|z-z_0| < R$ (S. 16.2) und $\sum\limits_{\nu=1}^{\infty} a_{-\nu}(z-z_0)^{-\nu}$ lokal gleichmäßig in $|z - z_0| > r$ (Anwendung von S. 16.2 auf $\sum\limits_{\nu=1}^{\infty} a_{-\nu}\zeta^\nu$ mit $\zeta = (z-z_0)^{-1}$). Also ist im Falle $r < R$ die Laurent-Reihe lokal gleichmäßig konvergent im Kreisring $V_{r,R}(z_0)$. Nach S. 31.7 ist

$$\begin{aligned} f(z) \quad &:= \quad \sum_{\nu=-\infty}^{\infty} a_\nu(z - z_0)^\nu := \lim_{n\to\infty} \sum_{\nu=-n}^{n} a_\nu(z - z_0)^n \\ &\Big(= \quad \sum_{\nu=0}^{\infty} a_\nu(z - z_0)^\nu + \sum_{\nu=1}^{\infty} a_{-\nu}(z - z_0)^{-\nu}\Big) \end{aligned}$$

holomorph in $V_{r,R}(z_0)$.

Beispiel 32.12 Für $a_\nu := 1/|\nu|!$ ($\nu \in \mathbb{Z}$) gilt

$$\lim_{\nu\to\infty} \sqrt[\nu]{|a_\nu|} = 0 \qquad \text{und} \qquad \lim_{\nu\to\infty} \sqrt[\nu]{|a_{-\nu}|} = 0\,,$$

also $r = 0$ und $R = \infty$. Hier ist

$$f(z) = \sum_{\nu=-\infty}^{\infty} a_\nu z^\nu = \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} + \sum_{\nu=1}^{\infty} \frac{1}{\nu! z^\nu} = e^z + e^{1/z} - 1$$

holomorph in $V_{0,\infty}(0) = \mathbb{C} \setminus \{0\}$.

Wir zeigen nun, dass umgekehrt jede in einem Kreisring holomorphe Funktion durch eine Laurent-Reihe dargestellt wird. Vorbereitend zeigen wir

Satz 32.13 *Es seien $0 \le r < R \le \infty, z_0 \in \mathbb{C}$, und es sei $f \in H(V_{r,R}(z_0))$. Dann gilt für $r < \varrho \le \sigma < R$*
$$\int\limits_{K_\varrho(z_0)} f = \int\limits_{K_\sigma(z_0)} f\,.$$

Beweis. O. E. können wir $z_0 = 0$ annehmen.
Wir wählen $\alpha, \beta \in \mathbb{R}$ mit $e^\alpha = \varrho$, $e^\beta = \sigma$ und betrachten den Polygonzug Γ durch $\alpha - i\pi, \beta - i\pi, \beta + i\pi, \alpha + i\pi, \alpha - i\pi$. Ist γ die Parametrisierung aus B. 26.11, so gilt für $\gamma_1 := \exp \circ \gamma : [0,4] \to \mathbb{C}$
$$\begin{aligned}\int\limits_{[-\varrho,-\sigma]} f + \int\limits_{K_\sigma(0)} f - \int\limits_{[-\varrho,-\sigma]} f - \int\limits_{K_\varrho(0)} f &= \int\limits_{\gamma_1} f = \int\limits_0^4 f \circ \gamma_1 \cdot \gamma_1' \\ &= \int\limits_0^4 f \circ \exp \circ \gamma \cdot \exp \circ \gamma \cdot \gamma' = \int\limits_\gamma f \circ \exp \cdot \exp\,.\end{aligned}$$
Da $f \circ \exp \cdot \exp$ holomorph ist in einem konvexen Gebiet, das das Rechteck $\{w : \alpha \le \operatorname{Re} w \le \beta, -\pi \le \operatorname{Im} w \le \pi\}$ enthält, gilt $\int\limits_\gamma f \circ \exp \cdot \exp = 0$ nach dem Cauchyschen Integralsatz. Also folgt
$$\int\limits_{K_\sigma(0)} f = \int\limits_{K_\varrho(0)} f\,.$$
□

Satz 32.14 *Es seien $0 \le r < R \le \infty, z_0 \in \mathbb{C}$, und es sei $f \in H(V_{r,R}(z_0))$. Ferner sei*
$$a_\nu = \frac{1}{2\pi i} \int\limits_{|\zeta - z_0| = \varrho} \frac{f(\zeta)}{(\zeta - z_0)^{\nu+1}} d\zeta \qquad (\nu \in \mathbb{Z})$$
für ein $\varrho \in (r, R)$. Dann gilt
$$f(z) = \sum_{\nu=-\infty}^{\infty} a_\nu (z - z_0)^\nu$$
mit lokal gleichmäßiger Konvergenz von Regulär- und Hauptteil in $V_{r,R}(z_0)$, Außerdem ist die Darstellung eindeutig, d.h. ist $\sum\limits_{\nu=-\infty}^{\infty} b_\nu (z - z_0)^\nu = f(z)$ lokal gleichmäßig in $V_{r,R}(z_0)$, so ist $b_\nu = a_\nu$ für alle $\nu \in \mathbb{Z}$.

Beweis. 1. Es sei $\varrho \in (r, R)$ fest. Wir betrachten die Funktion $g : [0, 2\pi] \to \mathbb{C}$ mit

$$g(t) := f(z_0 + \varrho e^{it}) \qquad (t \in [0, 2\pi])\,.$$

Dann ist $g \in C_{2\pi}$ (d.h. stetig mit $g(0) = g(2\pi)$) und $g \in C^1[0, 2\pi]$ (es gilt $g'(t) = f'(z_0 + \varrho e^{it}) \cdot \varrho i e^{it}$ auf $[0, 2\pi]$). Also hat g nach S. 28.8 eine gleichmäßig konvergente Fourier-Reihen-Darstellung

$$g(t) = \sum_{\nu=-\infty}^{\infty} c_\nu e^{i\nu t} \qquad (t \in [0, 2\pi])$$

mit

$$\begin{aligned} c_\nu &= \frac{1}{2\pi} \int_0^{2\pi} g(s) e^{-i\nu s} ds = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + \varrho e^{is}) e^{-i\nu s} ds \\ &= \frac{\varrho^\nu}{2\pi i} \int_{|\zeta - z_0| = \varrho} \frac{f(\zeta)}{(\zeta - z_0)^{\nu+1}} d\zeta = \varrho^\nu a_\nu \qquad (\nu \in \mathbb{Z})\,. \end{aligned}$$

Wir zeigen: Die Laurent-Reihe $\sum\limits_{\nu=-\infty}^{\infty} a_\nu (z - z_0)^\nu$ konvergiert lokal gleichmäßig in $V_{r,R}(z_0)$. Denn: Es gilt für $r < \sigma < R$ nach S. 32.13 (beachte: $\zeta \mapsto f(\zeta)/(\zeta - z_0)^{\nu+1}$ ist holomorph in $V_{r,R}(z_0)$)

$$|a_\nu| = \left| \frac{1}{2\pi i} \int_{|\zeta - z_0| = \sigma} \frac{f(\zeta)}{(\zeta - z_0)^{\nu+1}} d\zeta \right| \le \frac{1}{2\pi} \frac{2\pi\sigma}{\sigma^{\nu+1}} \max_{|\zeta - z_0| = \sigma} |f(\zeta)| = \frac{M}{\sigma^\nu} \qquad (\nu \in \mathbb{Z})$$

mit $M := \max\limits_{|\zeta - z_0| = \sigma} |f(\zeta)|$. Also folgt

$$\overline{\lim_{\nu \to \infty}} \sqrt[\nu]{|a_\nu|} \le \frac{1}{\sigma} \qquad \text{und} \qquad \overline{\lim_{\nu \to \infty}} \sqrt[\nu]{|a_{-\nu}|} \le \sigma\,.$$

Da $\sigma \in (r, R)$ beliebig war, ist

$$\overline{\lim_{\nu \to \infty}} \sqrt[\nu]{|a_\nu|} \le \frac{1}{R} \qquad \text{und} \qquad \overline{\lim_{\nu \to \infty}} \sqrt[\nu]{|a_{-\nu}|} \le r\,.$$

Aus B. 32.11 folgt die Behauptung.
Es sei $\tilde{f}(z) := \sum\limits_{\nu=-\infty}^{\infty} a_\nu (z - z_0)^\nu$. Dann ist $f \in H(V_{r,R}(z_0))$, und es gilt für $z = z_0 + \varrho e^{it}$

$$\tilde{f}(z) = \tilde{f}(z_0 + \varrho e^{it}) = \sum_{\nu=-\infty}^{\infty} \underbrace{a_\nu \varrho^\nu}_{=c_\nu} e^{i\nu t} = g(t) = f(z)\,.$$

Nach dem Identitätssatz (S. 29.11) ist $f = \tilde{f}$ in $V_{r,R}(z_0)$.
2. Ist $f(z) = \sum\limits_{\nu=-\infty}^{\infty} b_\nu (z - z_0)^\nu$ lokal gleichmäßig in $V_{r,R}(z_0)$, so gilt für $\varrho \in (r, R)$ und $k \in \mathbb{Z}$ nach B. 30.3.2

$$a_k = \frac{1}{2\pi i} \int_{|\zeta - z_0| = \varrho} \frac{f(\zeta)}{(\zeta - z_0)^{k+1}} d\zeta = \sum_{\nu=-\infty}^{\infty} b_\nu \frac{1}{2\pi i} \int_{|\zeta - z_0| = \varrho} (\zeta - z_0)^{\nu-k-1} d\zeta = b_k\,.$$

□

Nach S. 32.14 hat eine holomorphe Funktion um jede isolierte Singularität eine Laurent-Entwicklung. Wir zeigen nun, dass sich die drei Typen isolierter Singularitäten auch in natürlicher Weise durch die Hauptteile dieser Laurent-Entwicklungen charakterisieren lassen.

Satz 32.15 *Es sei $\Omega \subset \mathbb{C}$ offen, $a \in \Omega$ und $f \in H(\Omega \setminus \{a\})$. Ist $R := \operatorname{dist}(a, \partial\Omega)$, so hat f in $V_{0,R}(a) = U_R(a) \setminus \{a\}$ eine Laurent-Entwicklung*

$$f(z) = \sum_{\nu=-\infty}^{\infty} a_\nu (z-a)^\nu = \sum_{\nu=0}^{\infty} a_\nu (z-a)^\nu + \sum_{\nu=1}^{\infty} a_{-\nu}(z-a)^{-\nu}$$

gemäß S. 32.14*. Dabei gilt für $h(z) = \sum\limits_{\nu=1}^{\infty} a_{-\nu}(z-a)^{-\nu}$*

1. *a ist hebbare Singularität genau dann, wenn $a_{-\nu} = 0$ gilt für alle $\nu \in \mathbb{N}$ (d.h. $h \equiv 0$).*
2. *a ist Pol der Ordnung p genau dann, wenn $a_{-p} \neq 0$ und $a_{-\nu} = 0$ für alle $\nu > p$ gilt (d.h. $h(z) = \sum\limits_{\nu=1}^{p} a_{-\nu}(z-a)^{-\nu}$).*
3. *a ist wesentliche Singularität genau dann, wenn $a_{-\nu} \neq 0$ für ∞ viele $\nu \in \mathbb{N}$ gilt.*

Beweis.
1. f hat genau dann eine hebbare Singularität an a, wenn f eine Potenzreihenentwicklung $f(z) = \sum\limits_{\nu=0}^{\infty} a_\nu(z-a)^\nu$ in $U_R(a) \setminus \{a\}$ hat. Nach der Eindeutigkeitsaussage von S. 32.14 ist dies genau dann der Fall, wenn $h(z) \equiv 0$ auf $U_R(a) \setminus \{a\}$ ist.
2. f hat an a genau dann einen Pol der Ordnung p, wenn eine Umgebung U von a sowie eine Funktion $g \in H(U)$ mit $g(a) \neq 0$ und

$$f(z) = \frac{g(z)}{(z-a)^p} \quad \text{in} \quad U \setminus \{a\}$$

existieren. Dabei kann stets $g \in H(U_R(a))$ angenommen werden (man setze $g(z) := (z-a)^p f(z)$ in $U_R(a) \setminus \{a\}$). Somit hat g eine Potenzreihendarstellung

$$g(z) = \sum_{\nu=0}^{\infty} b_\nu (z-a)^\nu \quad \text{in} \quad U_R(a)\,.$$

Also: Hat f einen Pol der Ordnung p, so gilt in $V_{0,R}(a)$

$$f(z) = \sum_{\nu=0}^{\infty} b_\nu (z-a)^{\nu-p} = \sum_{\nu=-p}^{\infty} b_{\nu+p}(z-a)^\nu\,,$$

also ist $h(z) = \sum\limits_{\nu=-p}^{-1} b_{\nu+p}(z-a)^{\nu} = \sum\limits_{\nu=1}^{p} b_{p-\nu}(z-a)^{-\nu}$ der Haupttteil der Laurent-Reihe, und es gilt $a_{-p} = b_0 = g(a) \neq 0$.
Ist umgekehrt

$$f(z) = \sum_{\nu=-p}^{\infty} a_{\nu}(z-a)^{\nu} \quad \text{in} \quad V_{0,R}(a)$$

mit $a_{-p} \neq 0$, so ist

$$g(z) := \sum_{\nu=-p}^{\infty} a_{\nu}(z-a)^{\nu+p} = \sum_{\nu=0}^{\infty} a_{\nu-p}(z-a)^{\nu}$$

holomorph in $U_R(0)$ mit $g(a) = a_{-p} \neq 0$ und

3. Ergibt sich aus 1. und 2. $f(z) = \dfrac{g(z)}{(z-a)^p} \quad \text{in} \quad V_{0,R}(a)\,.$ □

Beispiel 32.16 Wir betrachten noch einmal die Beispiele aus B. 32.2.
1. Für $f(z) = (e^z - 1)/z$ gilt

$$f(z) = \sum_{\nu=0}^{\infty} \frac{z^{\nu}}{(\nu+1)!} \quad \text{in} \quad V_{0,\infty}(0)\,,$$

also ist hier $h(z) \equiv 0$. d.h. $a_{-\nu} = 0$ für alle $\nu \in \mathbb{N}$ (vgl. S. 32.15.1). 2. Für $f(z) = (z-a)^{-p}$ gilt $h(z) = (z-a)^{-p} (= f(z))$ in $V_{0,\infty}(a)$, d.h. $a_{-p} = 1$ und $a_{-\nu} = 0$ für $\nu \in \mathbb{N}, \nu \neq p$ (vgl. B. 32.15.2). 3. Für $f(z) = e^{1/z}$ gilt

$$f(z) = 1 + \sum_{\nu=1}^{\infty} \frac{1}{\nu!} z^{-\nu} \quad \text{in} \quad V_{0,R}(0)\,,$$

also ist hier $h(z) = \sum\limits_{\nu=1}^{\infty} z^{-\nu}/\nu! = \sum\limits_{\nu=1}^{\infty} a_{-\nu} z^{-\nu}$, d.h. $a_{-\nu} = 1/|\nu|! \neq 0$ für alle $\nu \in \mathbb{N}$ (vgl. S. 32.15.3).

Bemerkung 32.17 Eine ganze Funktion f heißt *transzendent*, falls f kein Polynom ist. Durch Übertragung des Satzes von Casorati-Weierstraß sieht man: Ist f transzendent, so existiert zu jedem $w \in \mathbb{C} \cup \{\infty\}$ eine Folge (z_n) in $\mathbb{C}$ mit $z_n \to \infty$ und $f(z_n) \to w\ (n \to \infty)$.
(Denn: Es sei $f(z) = \sum\limits_{\nu=0}^{\infty} a_{\nu} z^{\nu}\ (z \in \mathbb{C})$. Dann hat $g : \mathbb{C} \setminus \{0\} \to \mathbb{C}$,

$$g(z) = f\Big(\frac{1}{z}\Big) \qquad (z \neq 0)\,,$$

die Laurent-Entwicklung

$$g(z) = a_0 + \sum_{\nu=1}^{\infty} \frac{a_\nu}{z^\nu} \qquad (z \neq 0)\,.$$

Da $a_\nu \neq 0$ für ∞ viele ν gilt (beachte: f ist kein Polynom), hat g an 0 eine wesentliche Singularität nach S. 32.15. Also existiert nach S. 32.8 zu jedem $w \in \mathbb{C} \cup \{\infty\}$ eine Folge (ζ_n) in $\mathbb{C} \setminus \{0\}$ mit $\zeta_n \to 0$ und $g(\zeta_n) \to w\,(n \to \infty)$. Die Folge (z_n) mit $z_n = 1/\zeta_n$ erfüllt dann $z_n \to \infty$ und $f(z_n) \to w\,(n \to \infty)$.)

Man sagt auch, eine ganze Funktion habe eine „isolierte Singularität an ∞“. Konstante Funktionen haben dann eine „hebbare Singularität an ∞“, Polynome vom Grad ≥ 1 haben einen „Pol an ∞“, und transzendente Funktionen haben eine „wesentliche Singularität an ∞“.

33 Der Residuenkalkül

Wir wollen nun den sogenannten Residuensatz beweisen, ein Ergebnis, das man als Verallgemeinerung der Cauchyschen Integralformel auffassen kann. Im weiteren werden wir den Satz nutzen, um gewisse (z. T. reelle) Integrale bequem zu berechnen. Zunächst zum Begriff des Residuums.

Definition 33.1 Es sei $\Omega \subset \mathbb{C}$ offen, $a \in \Omega$ und $f \in H(\Omega \setminus \{a\})$. Ferner sei $f(z) = \sum_{\nu=-\infty}^{\infty} a_\nu (z-a)^\nu$ die Laurent-Entwicklung von f mit Entwicklungsmitte a gemäß S. 32.14. Dann heißt

$$\operatorname{Res}(f,a) := a_{-1} \left(= \frac{1}{2\pi i} \int_{|\zeta - a| = r} f(\zeta) d\zeta \text{ für } 0 < r < \operatorname{dist}(a, \partial\Omega) \right)$$

Residuum von f an der Stelle a.

Beispiel 33.2 (vgl. B.32.16)
1. Hat f an a eine hebbare Singularität, so gilt

$$\operatorname{Res}(f,a) = a_{-1} = \frac{1}{2\pi i} \int_{|\zeta - a| = r} f(\zeta) d\zeta = 0$$

nach S. 31.15 (oder dem CIS).
2. Es sei $f(z) = 1/(z-a)^p$ für ein $a \in \mathbb{C}$ und ein $p \in \mathbb{N}$. Dann gilt

$$\operatorname{Res}(f,a) = a_{-1} = \begin{cases} 0, & \text{falls } p > 1 \\ 1, & \text{falls } p = 1 \end{cases} .$$

3. Es sei

$$f(z) = e^{1/z} = 1 + \sum_{\nu=1}^{\infty} \frac{1}{\nu!} \frac{1}{z^\nu}$$

für $z \in \mathbb{C} \setminus \{0\}$. Dann gilt

$$\operatorname{Res}(f,0) = a_{-1} = 1 .$$

Es gilt

Satz 33.3 *(Residuensatz)*
Es sei $G \subset \mathbb{C}$ ein einfach zusammenhängendes Gebiet, und es sei $A \subset G$ endlich. Ferner sei f holomorph in $G \setminus A$. Dann gilt für alle geschlossenen Pfade γ in $G \setminus A$

$$\int_\gamma f = 2\pi i \sum_{w \in A} \operatorname{ind}_\gamma(w) \cdot \operatorname{Res}(f,w) .$$

Beweis. O. E. können wir $A \neq \emptyset$ annehmen (vgl. D.30.9).
Wir wählen $\delta > 0$ so, dass $U_\delta(w) \subset G$ für alle $w \in A$ und

$$|w - \tilde{w}| > \delta$$

für alle $w, \tilde{w} \in A, w \neq \tilde{w}$ gilt. Dann hat f für alle $w \in A$ nach S.32.14 in $V_{0,\delta}(w)$ eine Laurent-Entwicklung, d.h. es existieren $a_\nu = a_\nu(w) \in \mathbb{C}$ mit

$$f(z) = \sum_{\nu=-\infty}^{\infty} a_\nu(w)(z-w)^\nu$$

lokal gleichmäßig in $V_{0,\delta}(w)$. Der Hauptteil

$$h_w(z) := \sum_{\nu=1}^{\infty} a_{-\nu}(w)(z-w)^{-\nu}$$

konvergiert dann lokal gleichmäßig in $\mathbb{C} \setminus \{w\}$ (vgl. B. 32.11 und Beweis zu S. 30.14). Also folgt für $w \in A$

$$\int_\gamma h_w(z)dz = \sum_{\nu=1}^{\infty} a_{-\nu}(w) \int_\gamma \frac{dz}{(z-w)^\nu} = a_{-1}(w) \cdot 2\pi i \operatorname{ind}_\gamma(w) = 2\pi i \operatorname{ind}_\gamma(w) \cdot \operatorname{Res}(f, w)\,.$$

(Man beachte dabei: Für $\nu > 1$ ist $z \mapsto (z-w)^{1-\nu}/(1-\nu)$ eine Stammfunktion zu $z \mapsto (z-w)^{-\nu}$ in $\mathbb{C} \setminus \{w\}$, also ist $\int_\gamma (z-w)^{-\nu} dz = 0$ nach S.30.5.)
Die Funktion $g : G \setminus A \to \mathbb{C}$

$$g(z) := f(z) - \sum_{w \in A} h_w(z) \qquad (z \in G \setminus A)$$

ist holomorph in $G \setminus A$, und für $w \in A$ gilt in $V_{0,\delta}(w)$

$$g(z) = \sum_{\nu=0}^{\infty} a_\nu(w)(z-w)^\nu - \sum_{\substack{\tilde{w} \in A \\ w \neq \tilde{w}}} h_{\tilde{w}}(z)\,.$$

Da die rechte Seite holomorph in $U_\delta(w)$ ist, hat g an w eine hebbare Singularität. Also ist g holomorph fortsetzbar nach G. Damit gilt, da G einfach zusammenhängend ist,

$$\int_\gamma g = 0\,.$$

Folglich ist

$$0 = \int_\gamma f - \sum_{w \in A} \int_\gamma h_w = \int_\gamma f - 2\pi i \sum_{w \in A} \operatorname{ind}_\gamma(w) \cdot \operatorname{Res}(f, w)\,.$$

$\square$

Bemerkung 33.4 Der Residuensatz kann als Verallgemeinerung des CIS und der CIF aufgefasst werden:
Ist f holomorph in G und $A = \emptyset$, so gilt (mit $\sum\limits_{\emptyset} = 0$)

$$\int\limits_{\gamma} f = 0$$

für alle geschlossenen Pfade γ in G, also die Aussage des CIS (vgl. D.30.9).
Ist $z \in G$ und $A = \{z\}$, so hat die Funktion $g : G \setminus \{z\} \to \mathbb{C}$ mit

$$g(\zeta) := \frac{f(\zeta)}{\zeta - z} \qquad (\zeta \in G \setminus \{z\})$$

an der Stelle z eine isolierte Singularität. Es gilt dabei

$$g(\zeta) = \sum_{\nu=-1}^{\infty} \frac{f^{(\nu+1)}(z)}{(\nu+1)!} (\zeta - z)^{\nu}$$

für $|\zeta - z|$ genügend klein, also $\text{Res}(g, z) = f(z)$. Damit erhalten wir aus S.33.3 für alle geschlossenen Pfade γ in G

$$\frac{1}{2\pi i} \int\limits_{\gamma} \frac{f(\zeta}{\zeta - z} d\zeta = \frac{1}{2\pi i} \int\limits_{\gamma} g(\zeta) d\zeta \;=\; \text{ind}_{\gamma}(z) \cdot \text{Res}(g, z) \;=\; \text{ind}_{\gamma}(z) \cdot f(z)\,,$$

d.h. die CIF.

Um den Residuensatz anwenden zu können, ist es wichtig, Techniken zur Berechnung von Residuen zur Verfügung zu haben. Für Pole gilt:

Satz 33.5 *Es sei $\Omega \subset \mathbb{C}$ offen, $a \in \Omega$ und $f \in H(\Omega \setminus \{a\})$.*

1. *Hat f an a einen Pol der Ordnung p, so ist $(z-a)^p f(z)$ holomorph fortsetzbar nach Ω, und es gilt*

$$\text{Res}(f, a) = \frac{1}{(p-1)!} \lim_{z \to a} \frac{d^{p-1}}{dz^{p-1}} \big((z-a)^p f(z)\big)\,.$$

2. *Existieren eine offene Umgebung U von a und Funktionen $g, h \in H(U)$ mit $g(a) \neq 0, h(a) = 0, h'(a) \neq 0$ und $f = g/h$ in $U \setminus \{a\}$, so gilt*

$$\text{Res}(f, a) = \frac{g(a)}{h'(a)}\,.$$

Beweis. 1. Es gilt nach S.32.15 für $0 < |z-a| < r$, wobei $U_R(a) \subset \Omega$,

$$(z-a)^p f(z) = \sum_{\nu=-p}^{\infty} a_\nu (z-a)^{\nu+p} = \sum_{\nu=0}^{\infty} a_{\nu-p}(z-a)^\nu .$$

Also ist (da die rechte Seite holomorph in $U_r(a)$ ist)

$$\operatorname{Res}(f,a) = a_{-1} = \frac{1}{(p-1)!} \lim_{z\to a} \frac{d^{p-1}}{dz^{p-1}}\big((z-a)^p f(z)\big) .$$

2. Nach Voraussetzung hat h/g eine Nullstelle der Ordnung 1 an a, also hat f einen Pol der Ordnung 1 an a. Nach 1. ist

$$\operatorname{Res}(f,a) = \lim_{z\to a}(z-a)\frac{g(z)}{h(z)} = \lim_{z\to a} g(z) \lim_{z\to a} \frac{1}{\frac{h(z)-h(a)}{z-a}} = \frac{g(a)}{h'(a)} .$$ □

Beispiel 33.6 1. Es sei $f(z) = \cot z = \cos z/\sin z$ $(z \in \mathbb{C}\setminus\{k\pi : k \in \mathbb{Z}\})$. Dann gilt mit $g(z) = \cos z, h(z) = \sin z$ nach S.33.5.2

$$g(k\pi) = (-1)^k, \qquad h(k\pi) = 0, \qquad h'(k\pi) = \cos(k\pi) = (-1)^k$$

und damit

$$\operatorname{Res}(f,k\pi) = \frac{g(k\pi)}{h'(k\pi)} = 1 \qquad (k \in \mathbb{Z}) .$$

2. Es sei $f(z) = 1/(1+z^2)$ $(z \in \mathbb{C}\setminus\{\pm i\})$. Dann gilt mit $g(z) = 1$, $h(z) = 1+z^2$ nach S.33.5.2

$$\operatorname{Res}(f,\pm i) = \pm\frac{1}{2i} = \mp\frac{i}{2} .$$

Dies sieht man auch leicht direkt mittels Partialbruchzerlebung: Es gilt

$$f(z) = \frac{1}{(z+i)(z-i)} = \frac{i}{2}\frac{1}{z+i} - \frac{i}{2}\frac{1}{z-i} ,$$

also

$$\operatorname{Res}(f,i) = \frac{1}{2\pi i}\int\limits_{|z-i|=1} f(z)dz = -\frac{i}{2}\frac{1}{2\pi i}\int\limits_{|z-i|=1}\frac{dz}{z-i} = -\frac{i}{2}$$

(und entsprechend für $-i$).

3. Es sei $f(z) = 1/(1+z^2)^2$ $(z \neq 0)$. Dann hat f an $\pm i$ Pole der Ordnung 2. Es gilt nach S.33.5.1

$$\operatorname{Res}(f,i) = \lim_{z\to i}\frac{d}{dz}\big((z-i)^2 f(z)\big) = \lim_{z\to i}\frac{d}{dz}\left(\frac{1}{(z+i)^2}\right) = \lim_{z\to i}\frac{2}{(z+i)^3} = \frac{1}{4i} .$$

Wir werden nun zeigen, dass man den Residuensatz insbesondere dafür nutzen kann, uneigentliche Integrale der Form

$$\int\limits_{-\infty}^{\infty} f(x)dx$$

zu berechnen.

Satz 33.7 *Es sei* $E := \{z \in \mathbb{C} : \mathrm{Im}(z) \geq 0\}$*, und es sei* $A \subset E^0$ *endlich. Ferner sei* $f \in H(G \setminus A)$*, wobei* G *ein konvexes Gebiet mit* $E \subset G$ *ist, und so, dass Konstanten* $R_0, M > 0$ *und* $\alpha > 1$ *existieren mit*

$$|f(z)| \leq \frac{M}{|z|^\alpha} \qquad (z \in E, |z| \geq R_0)\,.$$

Dann existiert $\int\limits_{-\infty}^{\infty} f(x)dx$*, und es gilt*

$$\int\limits_{-\infty}^{\infty} f(x)dx \;=\; 2\pi i \sum_{w \in A} \mathrm{Res}(f, w)\,.$$

Beweis. Zunächst folgt aus $|f(x)| \leq M/|x|^\alpha$ für $x \in \mathbb{R}, |x| \geq R_0$ und der Existenz der Integrale $\int\limits_1^\infty dx/x^\alpha$ und $\int\limits_{-\infty}^{-1} dx/|x|^\alpha$ auch die Existenz der Integrale $\int\limits_0^\infty f(x)dx$ und $\int\limits_{-\infty}^{0} f(x)dx$ nach S.18.4 (man beachte dabei: f ist stetig auf $\mathbb{R}$, also existiert $\int\limits_{-R_0}^{R_0} f(x)dx$).

O.E. können wir davon ausgehen, dass $|w| < R_0$ für alle $w \in A$ gilt. Für $R \geq R_0$ betrachten wir den Pfad $\gamma_R = \gamma_R^{(1)} \oplus \gamma_R^{(2)}$, wobei

$$\gamma_R^{(1)}(t) = t \qquad (t \in [-R, R])$$

und

$$\gamma_R^{(2)}(t) = Re^{i(t-R)} \qquad (t \in [R, R+\pi])\,.$$

Dann gilt ([Ü]) $\mathrm{ind}_{\gamma_R}(w) = 1$ für alle $w \in E^0, |w| < R$. Also erhalten wir nach dem Residuensatz

$$\int\limits_{\gamma_R} f = 2\pi i \sum_{w \in A} \mathrm{Res}(f, w) \qquad \text{für alle } R \geq R_0.$$

Weiter gilt

$$\int\limits_{\gamma_R^{(1)}} f = \int\limits_{-R}^{R} f(x)dx$$

und

$$\Big|\int\limits_{\gamma_R^{(2)}} f(z)dz\Big| \;=\; \frac{M}{R^\alpha} \cdot \pi R \;=\; \frac{M\pi}{R^{\alpha-1}} \to 0 \qquad (R \to \infty)\,.$$

Also erhalten wir

$$\begin{aligned}\int\limits_{-R}^{R} f(x)dx \;&=\; \int\limits_{\gamma_R} f - \int\limits_{\gamma_R^{(2)}} = 2\pi i \sum_{w \in A} \mathrm{Res}(f, w) - \int\limits_{\gamma_R^{(2)}} f \\ &\to\; 2\pi i \sum_{w \in A} \mathrm{Res}(f, w) \qquad (R \to \infty)\,,\end{aligned}$$

woraus die Behauptung folgt. □

Bemerkung 33.8 Insbesondere lässt sich S.33.7 bei Integranden der Form

$$f(x) = e^{iax}\frac{P(x)}{Q(x)} = \cos(ax)\frac{P(x)}{Q(x)} + i\,\sin(ax)\frac{P(x)}{Q(x)}$$

anwenden, wobei $a \geq 0$ ist und wobei P und Q Polynome sind mit $\deg(Q) \geq \deg(P)+2$ und $Q(x) \neq 0\,(x \in \mathbb{R})$.
(Denn: Es sei

$$A := Z(Q) \cap E^0$$

die Menge der Nullstellen von Q in der oberen Halbebene E^0. Dann ist

$$f(z) := e^{iaz}\frac{P(z)}{Q(z)}$$

holomorph in $G_\delta \setminus A$, für $G_\delta := \{z : \mathrm{Im}(z) > -\delta\}$ mit einem $\delta > 0$. Außerdem gilt für $\mathrm{Im} z \geq 0, z = x + iy$ mit einem $M > 0$

$$|f(z)| = \underbrace{e^{-ay}}_{\leq 1}\frac{|P(z)|}{|Q(z)|} \leq \frac{|P(z)|}{|Q(z)|} \leq \frac{M}{|z|^2}$$

für $|z|$ genügend groß. Damit sind alle Voraussetzungen von S.33.7 erfüllt.)

Beispiel 33.9 Für $a > 0$ sei $f : \mathbb{C} \setminus \{\pm i\}$ definiert durch

$$f(z) = \frac{e^{iaz}}{1+z^2} \qquad (z \in \mathbb{C} \setminus \{\pm i\})\,.$$

Dann ist f wie in S.33.8. Also gilt nach S.33.7

$$\int\limits_{-\infty}^{\infty} \frac{\cos(ax)}{1+x^2}\,dx = \mathrm{Re}\int\limits_{-\infty}^{\infty} \frac{e^{iax}}{1+x^2}\,dx = \mathrm{Re}(2\pi i \mathrm{Res}(f,i))\,.$$

Weiter ist (etwa nach S.33.5.2)

$$\mathrm{Res}(f,i) = \frac{e^{iaz}}{2z}\Big|_{z=i} = \frac{e^{-a}}{2i}\,,$$

also

$$\int\limits_{-\infty}^{\infty} \frac{\cos(ax)}{1+x^2}\,dx = \pi e^{-a}$$

(klar ist, dass $\int\limits_{-\infty}^{\infty} \sin(ax)/(1+x^2) = 0$ gilt, da der Integrand ungerade ist).

Eine weitere interessante Klasse von Integralen, die mittels des Residuensatzes oft leicht berechnet werden können, sind Integrale der Form

$$\int_0^{2\pi} R(\cos t)dt \qquad \text{bzw.} \qquad \int_0^{2\pi} R(\sin t)dt\,.$$

wobei R eine rationale Funktion ist. Beachtet man, dass

$$\cos t = \frac{1}{2}\big(z + \frac{1}{z}\big)\,, \qquad \sin t = \frac{1}{2i}\big(z - \frac{1}{z}\big)$$

für $z = e^{it}$ gilt, so ergibt sich mit

$$R^*(z) = \frac{1}{z}\,R\Big(\frac{1}{2}\big(z + \frac{1}{z}\big)\Big) \qquad \text{bzw.} \qquad R^{**}(z) = \frac{1}{z}\,R\Big(\frac{1}{2i}\big(z - \frac{1}{z}\big)\Big)$$

dabei (falls R^* bzw. R^{**} stetig auf $|z| = 1$ sind)

$$\int_{|z|=1} R^*(z)dz \;=\; i\int_{|z|=1} R\Big(\frac{1}{2}\big(z + \frac{1}{z}\big)\Big)\frac{dz}{iz} \;=\; i\int_0^{2\pi} R(\cos t)dt \tag{33.1}$$

bzw.

$$\int_{|z|=1} R^{**}(z)dz \;=\; i\int_{|z|=1} R\Big(\frac{1}{2i}\big(z - \frac{1}{z}\big)\Big)\frac{dz}{iz} \;=\; i\int_0^{2\pi} R(\sin t)dt \tag{33.2}$$

Dies beweist schon im Wesentlichen folgenden

Satz 33.10 *Es sei R eine rationale Funktion.*

1. *Ist*
$$R^*(z) := \frac{1}{z}R\Big(\frac{1}{2}\big(z + \frac{1}{z}\big)\Big)$$
holomorph in $U_\delta(0) \setminus A^*$ für eine endliche Menge $A^* \subset \mathbb{D} := \{|z| < 1\}$ und ein $\delta > 1$, so gilt
$$\int_0^{2\pi} R(\cos t)dt = 2\pi \sum_{w\in A^*} \operatorname{Res}(R^*, w)\,.$$

2. *Ist*
$$R^{**}(z) := \frac{1}{z}R\Big(\frac{1}{2i}\big(z - \frac{1}{z}\big)\Big)$$
holomorph in $U_\delta(0) \setminus A^{**}$ für eine endliche Menge $A^{**} \subset \mathbb{D}$ und ein $\delta > 1$, so folgt
$$\int_0^{2\pi} R(\sin t)dt = 2\pi \sum_{w\in A^{**}} \operatorname{Res}(R^{**}, w)\,.$$

Beweis. Die Behauptungen ergeben sich unmittelbar aus (33.1) bzw. (33.2) und dem Residuensatz, angewandt auf R^* bzw. R^{**} und $G = U_\delta(0)$ sowie $\gamma(t) = e^{it}$ $(t \in [0, 2\pi])$. □

Beispiel 33.11 1. Für $0 < r < 1$ betrachten wir das Integral

$$\int_0^{2\pi} \frac{dt}{1 - 2r\cos t + r^2}\,.$$

Ist

$$R(\cos t) = \frac{1}{1 - 2r\cos t + r^2}$$

und

$$R^*(z) = \frac{1}{z} R\Big(\frac{1}{2}\big(z + \frac{1}{z}\big)\Big) = \frac{1}{z} \cdot \frac{1}{1 - r(z + 1/z) + r^2} = \frac{1}{(z - r)(1 - rz)}\,,$$

so hat R^* die beiden einfachen Pole $r < 1$ und $1/r > 1$. Also gilt nach S.33.10 und S.33.5.1

$$\int_0^{2\pi} \frac{dt}{1 - 2r\cos t + r^2} = 2\pi \cdot \operatorname{Res}(R^*, r) \cdot = 2\pi \cdot \frac{1}{1 - rz}_{|z=r} = \frac{2\pi}{1 - r^2}\,.$$

[Für die Funktion

$$P(r, t) := \frac{1 - r^2}{1 - 2r\cos t + r^2} \qquad (0 < r < 1, 0 \le t \le 2\pi)$$

folgt also

$$\frac{1}{2\pi} \int_0^{2\pi} P(r, t)dt = 1$$

für alle r. Diese Funktion, der sogenannte Poisson-Kern, spielt eine wichtige Rolle in der Theorie harmonischer Funktionen, wie wir später noch sehen werden.]
2. Für $p \in N$ und $a > 1$ betrachten wir das Integral

$$\int_0^{2\pi} \frac{dt}{(a + \cos t)^p}\,.$$

Hier ist

$$R(\cos t) = \frac{1}{(a + \cos t)^p}\,,$$

also

$$R^*(z) = \frac{1}{z} \cdot \frac{1}{(a + (z + 1/z)/2)^p} = \frac{2^p z^{p-1}}{(z^2 + 2az + 1)^p} = \frac{2^p z^{p-1}}{(z - z_1^*)^p (z - z_2^*)^p}$$

mit $w_1 = -a + \sqrt{a^2-1} \in (-1,0)$ und $w_2 = -a - \sqrt{a^2-1} < -1$.
Also ergibt sich aus S.33.10 und S.33.5.1

$$\int_0^{2\pi} \frac{dt}{(a+\cos t)^p} = 2\pi \cdot \operatorname{Res}(R^*, w_1) = \frac{2\pi}{(p-1)!} \frac{d^{p-1}}{dz^{p-1}} \Big(\frac{2^p z^{p-1}}{(z-w_2)^p}\Big)_{|z=w_1} .$$

Für $p = 1$ erhalten wir

$$\int_0^{2\pi} \frac{dt}{a+\cos t} = 2\pi \cdot \frac{2}{2\sqrt{a^2-1}} = \frac{2\pi}{\sqrt{a^2-1}} ,$$

und für $p = 2$ folgt

$$\begin{aligned}\int_0^{2\pi} \frac{dt}{(a+\cos t)^2} &= 2\pi \cdot \frac{d}{dt} \Big(\frac{4z}{(z-w_2)^2}\Big)_{|z=w_1} \\ &= 2\pi \cdot 4 \cdot \frac{-w_1 - w_2}{(w_1-w_2)^3} = \frac{2\pi a}{\sqrt{a^2-1}^3} .\end{aligned}$$

Wir kommen zum Schluss dieses Abschnittes zu zwei interessanten funktionentheoretischen Anwendungen des Residuensatzes.

Definition 33.12 Es sei $\Omega \subset \mathbb{C}$ offen. Eine Funktion f heißt *meromorph* in Ω, falls eine Menge $A \subset \Omega$ ohne Häufungspunkt in Ω existiert mit $f \in H(\Omega \setminus A)$ und so, dass f an den Stellen $a \in A$ (falls $A \neq \emptyset$) Pole hat.

Bemerkung 33.13 1. Ist $f \in H(\Omega)$, so ist f auch meromorph in Ω (da $A = \emptyset$ zulässig ist).
2. Indem wir formal $f(a) := \infty$ für alle $a \in A$ (im Falle $A \neq \emptyset$) setzen, können wir f auch als auf ganz Ω definiert betrachten. Man beachte dabei: Es gilt stets $\lim\limits_{z \to a} f(z) = \infty$ für $a \in A$ (S.32.6).

Beispiel 33.14 1. Die Funktionen cot und tan sind meromorph in $\mathbb{C}$, denn mit $A = \{k\pi : k \in \mathbb{Z}\}$ gilt: $\cot \in H(\mathbb{C} \setminus A)$ und cot hat an den Stellen $a \in A$ Pole (1. Ordnung). Entsprechendes gilt für tan mit $\tilde{A} = \{(k + \frac{1}{2})\pi : k \in \mathbb{Z}\}$ (vgl. B. 32.7). Man beachte dabei: A bzw. $\tilde{A}$ haben keine Häufungspunkte in $\mathbb{C}$.
2. Es sei $f : \mathbb{C} \setminus (\{1/(k\pi) : k \in \mathbb{Z}\} \cup \{0\}) \to \mathbb{C}$, definiert durch

$$f(z) = \frac{1}{\sin(1/z)} \qquad \big(z \neq \frac{1}{k\pi}, z \neq 0\big) .$$

Dann ist f meromorph in $\Omega = \mathbb{C} \setminus \{0\}$ (mit Polstellenmenge $A = \{1/(k\pi) : k \in \mathbb{Z}\}$), jedoch nicht in $\mathbb{C}$ (da 0 ein Häufungspunkt von A ist).

Als Folgerung aus dem Residuensatz erhalten wir

Satz 33.15 *(Argumentprinzip)*
Es sei $G \subset \mathbb{C}$ ein einfach zusammenhängendes Gebiet, und es sei f meromorph in G. Ferner seien $Z(f) := \{\text{Nullstellen von } f \text{ in } G\}$ und $P(f) := \{\text{Polstellen von } f \text{ in } G\}$ endlich. Ist γ ein geschlossener Pfad in $G \setminus (Z(f) \cup P(f))$, so gilt

$$\frac{1}{2\pi i} \int\limits_\gamma \frac{f'}{f} = \sum_{w \in Z(f)} \mathrm{ind}_\gamma(w) \cdot n(w) - \sum_{w \in P(f)} \mathrm{ind}_\gamma(w) \cdot p(w)\,,$$

wobei $n(w) = n_f(w)$ die Ordnung der Nullstelle w von f und $p(w) = p_f(w)$ die Ordnung der Polstelle w von f bezeichnet.

Beweis. 1. Ist a eine Nullstelle von f der Ordnung $n(a)$, so existiert eine in einer offenen Umgebung U von a holomorphe Funktion g mit $g(z) \neq 0$ in U und

$$f(z) = (z-a)^{n(a)} g(z) \qquad (z \in U)\,.$$

Also folgt

$$\frac{f'(z)}{f(z)} = \frac{n(a)}{z-a} + \frac{g'(z)}{g(z)} \qquad \text{in } U \setminus \{a\}\,.$$

d.h. f'/f hat an a einen Pol der Ordnung 1, und es gilt

$$\mathrm{Res}\big(\frac{f'}{f}, a\big) = n(a)\,.$$

2. Ist a ein Pol der Ordnung $p(a)$ von f, so existiert ein in einer Umgebung U von a holomorphe Funktion g mit $g(z) \neq 0$ in U und

$$f(z) = \frac{g(z)}{(z-a)^{p(a)}} \qquad (z \in U \setminus \{a\})\,.$$

Aus

$$f'(z) = g'(z) \cdot \frac{1}{(z-a)^{p(a)}} + g(z) \cdot \frac{-p(a)}{(z-a)^{p(a)+1}} \qquad (z \in U \setminus \{a\})$$

folgt

$$\frac{f'(z)}{f(z)} = \frac{g'(z)}{g(z)} - \frac{p(a)}{z-a} \qquad (z \in U\{a\})\,,$$

d.h. f'/f hat wieder einen Pol 1. Ordnung an a mit

$$\mathrm{Res}\Big(\frac{f'}{f}, a\Big) = -p(a)\,.$$

3. Ist γ ein geschlossener Pfad in $G \setminus (Z(f) \cup P(f))$, so gilt nach S. 33.3 und 1. und 2.

$$\frac{1}{2\pi i} \int\limits_\gamma \frac{f'}{f} = \sum_{w \in Z(f)} \mathrm{ind}_\gamma(w) \cdot n(w) - \sum_{w \in P(f)} \mathrm{ind}_\gamma(w) \cdot p(w)\,. \qquad \square$$

Bemerkung und Definition 33.16 Es sei $\gamma : [\alpha, \beta] \to \mathbb{C}$ ein geschlossener Pfad. Dann heißt γ *einfach geschlossen*, falls $\mathbb{C} \setminus \Gamma$, wobei $\Gamma := \gamma([\alpha, \beta])$ nur zwei Komponenten hat (d.h. nur eine beschränkte Komponente; vgl. B/D.30.10), und falls $|\mathrm{ind}_\gamma(z)| = 1$ in der beschränkten Komponente gilt. Dann nennen wir die unbeschränkte Komponente das *Außengebiet* $\mathrm{Ext}(\gamma)$ und die beschränkte Komponente das *Innengebiet* $\mathrm{Int}(\gamma)$. Außerdem sagen wir, γ sei *positiv orientiert*, falls $\mathrm{ind}_\gamma(z) = 1$ für alle $z \in \mathrm{Int}(\gamma)$ gilt. Für positiv orientierte, einfach geschlossene Pfade γ in $G \setminus (Z(f) \cup P(f))$ ergibt sich unter den Voraussetzungen von S. 33.15

$$\frac{1}{2\pi i} \int_\gamma \frac{f'}{f} = \sum_{w \in Z(f) \cap \mathrm{Int}(\gamma)} n(w) - \sum_{w \in P(f) \cap \mathrm{Int}(\gamma)} p(w) \quad , \tag{33.3}$$

d.h. das Integral auf der linken Seite gibt die Differenz zwischen der Anzahl der Nullstellen und der Polstellen von f in $\mathrm{Int}(\gamma)$ (inkl. Vielfachheiten) an.

Bemerkung 33.17 Ist unter den Voraussetzungen von S. 33.15 zusätzlich f holomorph in G (m. a. W. $P(f) = \emptyset$) und ist γ ein positiv orientierter, einfach geschlossener Pfad in $G \setminus Z(f)$, so erhalten wir aus (33.3)

$$\frac{1}{2\pi i} \int_\gamma \frac{f'}{f} = \sum_{w \in Z(f) \cap \mathrm{Int}(\gamma)} n(w) \quad , \tag{33.4}$$

d.h. das Integral auf der linken Seite „zählt“ die Nullstellen von f in $\mathrm{Int}(\gamma)$ (inkl. Vielfachheiten).

Als Folgerung aus dem Argument-Prinzip (bzw. (33.4)) erhalten wir

Satz 33.18 *(Rouché)*
Es sei $G \subset \mathbb{C}$ ein einfach zusammenhängendes Gebiet, und es seien $f, g \in H(G)$ mit nur endlich vielen Nullstellen in G. Ferner sei γ ein einfach geschlossener Pfad in G so, dass

$$|f(z) - g(z)| < |f(z)| \quad \text{für alle } z \in \Gamma\,.$$

Dann haben f und g die gleiche Anzahl von Nullstellen in $\mathrm{Int}(\gamma)$ (incl. Vielfachheiten), d. h.

$$\sum_{w \in Z(f) \cap \mathrm{Int}(\gamma)} n_f(w) = \sum_{w \in Z(g) \cap \mathrm{Int}(\gamma)} n_g(w)\,.$$

Beweis. Für $t \in [0, 1]$ betrachten wir die Funktionen $\varphi_t \in H(G)$ mit

$$\varphi_t(z) := f(z) - t\big(f(z) - g(z)\big) \qquad (z \in G)\,.$$

Für $z \in \Gamma$ gilt

$$|\varphi_t(z)| \geq |f(z)| - t|f(z) - g(z)| \geq |f(z)| - |f(z) - g(z)| > 0\,,$$

so dass φ_t auf Γ keine Nullstellen hat. Nach (33.4) gilt für $t \in [0,1]$

$$\frac{1}{2\pi i}\int\limits_\gamma \frac{\varphi_t'(z)}{\varphi_t(z)}dz \;=\; \sum_{w \in Z(\varphi_t)\cap \mathrm{Int}(\gamma)} n(w) \;=:\; N(t)\,.$$

Die Funktion $N : [0,1] \to \mathbb{N}_0$ ist stetig.
(Denn: Sind $t_1, t_2 \in [0,1]$, so gilt mit $h(z) := (f-g)(z)$

$$\left|2\pi i\big(N(t_2) - N(t_1)\big)\right| = \left|\int\limits_\gamma \frac{(t_2 - t_1)(f'h - fh')}{(f - t_1 h)(f - t_2 h)}\right| \leq |t_2 - t_1| \left\|\frac{fh' - f'h}{(|f| - |h|)^2}\right\|_{\infty,\Gamma} L(\gamma)\,.$$

Dies zeigt die (Lipschitz-)Stetigkeit von N.)
Da $[0,1]$ zusammenhängend ist, ist $N(t) \equiv \mathrm{const}$ auf $[0,1]$, also insbesondere

$$\frac{1}{2\pi i}\int\limits_\gamma \frac{f'(z)}{g(z)}dz = N(1) = N(0) = \frac{1}{2\pi i}\int\limits_\gamma \frac{f'(z)}{f(z)}dz\,,$$

d.h.

$$\sum_{w \in Z(g)\cap \mathrm{Int}(\gamma)} n_g(w) \;=\; \sum_{w \in Z(f)\cap \mathrm{Int}(\gamma)} n_f(w)\,.$$

□

Wir geben einige typische Anwendungsbeispiele zum Satz von Rouché.

Beispiel 33.19 1. Wir beweisen noch einmal den Fundamentalsatz der Algebra. Also: Es sei $P(z) = \sum\limits_{\nu=0}^{n} a_\nu z^\nu$ ein Polynom vom Grad $n \in \mathbb{N}$. Dann ist $\sum\limits_{w \in Z(P)} n(w) = n$, d.h. P hat n Nullstellen inkl. Vielfachheiten.
(Denn: Ist $Q(z) := a_n z^n$, so gilt für R genügend groß

$$|P(z) - Q(z)| \;=\; \Big|\sum_{\nu=0}^{n-1} a_\nu z^\nu\Big| \;<\; |a_n z^n| \;=\; |Q(z)| \qquad (|z| = R)\,.$$

Also ergibt sich aus S. 33.18 (mit $\gamma(t) = \mathrm{R}e^{it}$)

$$\sum_{w \in Z(P)\cap U_R(0)} n_P(w) \;=\; \sum_{w \in Z(Q)\cap U_R(0)} n_Q(w) \;=\; n_Q(0) = n\,.)$$

2. Wir betrachten die (transzendente) Gleichung

$$e^z \;=\; 1 + 2z\,.$$

Wir suchen alle Lösungen in $\mathbb{D}$. Offensichtlich ist $z = 0$ eine Lösung. Mit $f(z) = 2z$ und $g(z) = 1 + 2z - e^z$ gilt für $|z| = 1$

$$|f(z) - g(z)| \;=\; |e^z - 1| \;<\; 2 = |f(z)|$$

(beachte: für $|z| \leq 1$ gilt

$$|e^z - 1| \;\leq\; \sum_{\nu=1}^{\infty} \frac{|z|^\nu}{\nu!} \;=\; e^{|z|} - 1 \;\leq\; e - 1 < 2\,).$$

Also haben f und g die gleiche Anzahl von Nullstellen, nämlich eine in $\mathbb{D}$. Folglich ist $z = 0$ die einzige Lösung der Gleichung in $\mathbb{D}$.

Wir studieren zum Abschluss einige Auswirkungen des Satzes von Rouché auf Funktionenfolgen.

Satz 33.20 *Es sei $G \subset \mathbb{C}$ ein einfach zusammenhängendes Gebiet, und es seien $f_n \in H(G)$ mit $f_n \to f$ lokal gleichmäßig auf G.*
Ist $\gamma : [a, b] \to \mathbb{C}$ ein einfach geschlossener Pfad in G und ist $f(z) \neq 0$ für alle $z \in \Gamma := \gamma([a, b])$, so haben f und f_n für n genügend groß die gleiche Anzahl von Nullstellen (inkl. Vielfachheiten) in $\mathrm{Int}(\gamma)$.

Beweis. Da f stetig auf Γ ist, gilt

$$\delta := \min_{z \in \Gamma} |f(z)| > 0\,.$$

Da $\Gamma \subset G$ kompakt ist, existiert ein $N = N(\delta) > 0$ so, dass

$$\max_{z \in \Gamma} |f(z) - f_n(z)| < \delta$$

für alle $n \geq N$ gilt. Damit folgt die Behauptung aus dem Satz von Rouché (S. 33.18).
□

Beispiel 33.21 Es sei

$$f(z) = e^z = \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} \qquad (z \in \mathbb{C})\,.$$

Dann gilt für die n-ten Teilsummen $s_n(z) = \sum\limits_{\nu=0}^{n} z^\nu/\nu!$ nach S. 33.20 (angewandt mit $\gamma(t) = \mathrm{R}e^{it} (t \in [0, 2\pi])$: Es existiert ein $N = N(R)$ so, dass s_n für alle $n \geq N$ in $|z| < R$ keine Nullstelle hat (d.h. die Nullstellen, die nach dem Fundamentalsatz der Algebra ja existieren, rücken mit wachsendem n immer weiter „Richtung ∞“).

Als Folgerung aus S. 33.20 erhalten wir insbesondere

Satz 33.22 *(Hurwitz)*
Ist $G \subset \mathbb{C}$ ein Gebiet und ist (f_n) eine Folge in G holomorpher Funktionen mit $f_n \to f$ lokal gleichmäßig auf G und $f_n(z) \neq 0$ in G, so ist entweder $f \equiv 0$ in G oder es ist $f(z) \neq 0$ in G.

Beweis. Es sei $f \not\equiv 0$ in G. Angenommen, f habe eine Nullstelle $z_0 \in G$, etwa der Ordnung $m \in \mathbb{N}$. Ist $r > 0$ so, dass $f(z) \neq 0$ in $\overline{U_r(z_0)} \setminus \{z_0\}$ gilt, so folgt aus S. 33.20 (angewandt auf $\gamma(t) = z_0 + re^{it}$), dass f_n für n genügend groß in $U_r(z_0)$ ebenfalls m Nullstellen hat. Widerspruch! □

Definition 33.23 Es sei $G \subset \mathbb{C}$ ein Gebiet, und es sei $f \in H(G)$. Dann heißt f *schlicht* in G, falls f injektiv ist.

Satz 33.24 *Es sei $G \subset \mathbb{C}$ ein Gebiet und sei (f_n) eine Folge in G schlichter Funktionen mit $f_n \to f$ lokal gleichmäßig auf G. Dann ist entweder $f \equiv$* const *oder f ist schlicht in G.*

Beweis. Ist $z_0 \in G$ fest, so sind nach Voraussetzung die Funktionen $f_n - f_n(z_0)$ nullstellenfrei in $G \setminus \{z_0\}$. Also ist nach S. 33.22 entweder $f - f(z_0) \equiv 0$, d.h. $f \equiv f(z_0)$ in G, oder $f - f(z_0)$ ist nullstellenfrei in $G \setminus \{z_0\}$, d.h. $f(z) \neq f(z_0)$ für alle $z \neq z_0$. Ist also $f \not\equiv$ const in G, so gilt für alle $z_0 \in G$ die zweite Alternative. Folglich ist f injektiv (und damit schlicht) in G. □

34 Der Rungesche Approximationssatz mit Anwendungen

In diesem Abschnitt befassen wir uns mit einem Ergebnis, mit dessen Hilfe viele Aussagen der sogenannten konstruktiven Funktionentheorie einfach bewiesen werden können: der Rungesche Approximationssatz. Es geht um die Frage, unter welchen Bedingungen holomorphe Funktionen durch Polynome oder rationale Funktionen angenähert werden können.
Ist etwa $G = U_r(z_0)$ und $f \in H(G)$, so gilt bekanntlich

$$f(z) = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(z_0)}{\nu!}\,(z - z_0)^{\nu}$$

lokal gleichmäßig in G. Ist also

$$P_n(z) := \sum_{\nu=0}^{n} \frac{f^{(\nu)}(z_0)}{\nu!}(z - z_0)^{\nu}\ ,$$

so konvergiert die Polynomfolge (P_n) gleichmäßig gegen f auf allen kompakten Teilmengen K von G. Auf der anderen Seite haben wir

Beispiel 34.1 Es sei $G = \mathbb{C} \setminus \{0\}$ und $f(z) = 1/z\,(z \in G)$. Dann existiert keine Folge (P_n) von Polynomen mit $P_n \to f$ gleichmäßig auf $K = \{z : |z| = 1\}$, also insbesondere keine Polynomfolge mit $P_n \to f$ lokal gleichmäßig auf G.
(Denn: Angenommen, doch. Dann gilt nach S. 31.7

$$0 = \int\limits_{|z|=1} P_n(z)dz \;\to\; \int\limits_{|z|=1} \frac{dz}{z} \;=\; 2\pi i\,.$$

Widerspruch!)

Entsprechende Probleme treten nicht mehr auf, wenn man mit rationalen Funktionen statt mit Polynomen approximiert.

Satz 34.2 *(Rungescher Approximationssatz für rationale Approximation)*
Es sei $K \subset \mathbb{C}$ kompakt, und es sei $f \in H(\Omega)$ für eine offene Menge $\Omega \supset K$. Dann existiert zu jedem $\varepsilon > 0$ eine rationale Funktion R der Form $R(z) = \sum\limits_{k=1}^{M} \alpha_k/(z - \zeta_k)$, wobei $\alpha_1, \ldots, \alpha_M \in \mathbb{C}$ und $\zeta_1, \ldots, \zeta_M \in \mathbb{C} \setminus K$, mit

$$\max_{z\in K} |f(z) - R(z)| < \varepsilon\,.$$

Als Vorbereitung für den Beweis zu S. 34.2 brauchen wir eine allgemeine Version der Cauchyschen Integralformel:

Satz 34.3 *(Cauchysche Integralformel für allgemeine Kompakta)*
Es sei $K \subset \mathbb{C}$ kompakt, und es sei $\Omega \subset \mathbb{C}$ offen mit $K \subset \Omega$. Dann existieren (orientierte) Strecken $\Gamma_1, \ldots, \Gamma_m$ in $\Omega \setminus K$ so, dass für alle $f \in H(\Omega)$ und alle $z \in K$ gilt

$$f(z) = \frac{1}{2\pi i} \sum_{j=1}^{m} \int_{\Gamma_j} \frac{f(\zeta)}{\zeta - z}\, d\zeta\,.$$

Beweis. Da $K \subset \Omega$ kompakt ist, gilt

$$\delta := \operatorname{dist}(K, \partial\Omega) > 0\,.$$

Legt man über die Ebene $\mathbb{C}$ ein achsenparalleles Gitter aus kompakten Quadraten mit Maschenweite d so, dass $d\sqrt{2} < \delta$ gilt, so wird K von nur endlich vielen dieser Quadrate getroffen. Wir bezeichnen diese mit $Q_1, \ldots, Q_N$.
Behauptung: Es gilt

$$K \subset \bigcup_{j=1}^{N} Q_j \subset \Omega\,.$$

(Nach Definition der $Q_1, \ldots, Q_N$ gilt $K \subset \bigcup\limits_{j=1}^{N} Q_j$.
Ist $j \in \{1, \ldots, m\}$ und $z_j \in Q_j \cap K$, so gilt $U_\delta(z_j) \subset \Omega$ nach Definition von δ. Da Q_j Durchmesser $d\sqrt{2} < \delta$ hat, gilt außerdem $Q_j \subset U_\delta(z_j)$. Also ist $Q_j \subset \Omega$.)
Wir betrachten nun die Strecken, die zum Rand *genau* eines der Quadrate $Q_1, \ldots, Q_N$ gehören (als nicht gemeinsame Seiten zweier der Quadrate). Die Ränder der $Q_1, \ldots, Q_N$ seien positiv orientiert. Damit ergibt sich eine Orientierung für die entsprechenden Strecken. Wir nennen die (so orientierten) Strecken $\Gamma_1, \ldots, \Gamma_m$. Es gilt dann

$$\bigcup_{k=1}^{m} \Gamma_k \subset \Omega \setminus K\,.$$

(Denn: Würde eine der Strecken aus $\bigcup\limits_{k=1}^{m} \Gamma_k$ die Menge K treffen, so hätten die beiden an diese Strecke angrenzenden Quadrate gemeinsame Punkte mit K, was der Auswahl der $Q_1, \ldots, Q_N$ widerspricht. Außerdem gilt $\bigcup\limits_{k=1}^{m} \Gamma_k \subset \Omega$, da $\bigcup\limits_{j=1}^{N} Q_j \subset \Omega$ ist.
Da jede Seite, die zum Rand zweier der Quadrate $Q_1, \ldots, Q_N$ gehört, entgegengesetzt durchlaufen wird (wenn ∂Q_j positiv orientiert ist), folgt

$$\sum_{j=1}^{N} \int_{\partial Q_j} \frac{f(\zeta)}{\zeta - z} d\zeta = \sum_{k=1}^{m} \int_{\Gamma_k} \frac{f(\zeta)}{\zeta - z} d\zeta \qquad \text{für alle } z \in \Omega \setminus \bigcup_{j=1}^{N} Q_j.$$

Ist $z \in Q_j^0$ für ein $j \in \{1, \ldots, N\}$, so gilt nach der CIF (S. 30.12)

$$\frac{1}{2\pi i} \int\limits_{\partial Q_j} \frac{f(\zeta)}{\zeta - z} d\zeta = f(z)$$

und nach dem CIS (S. 30.8)

$$\frac{1}{2\pi i} \int\limits_{\partial Q_\ell} \frac{f(\zeta)}{\zeta - z}\, d\zeta = 0 \qquad (\ell = 1, \ldots, N; \ell \neq j)\,.$$

Also gilt die Behauptung für alle $z \in \bigcup\limits_{j=1}^{N} Q_j^0 (\cap K)$.

Ist $z \in K \cap \partial Q_j$ für ein $j \in \{1, \ldots, N\}$, so ist $z \notin \bigcup\limits_{k=1}^{m} \Gamma_k$. Nach S. 29.9 ist

$$z \mapsto \sum_{k=1}^{m} \int\limits_{\Gamma_k} \frac{f(\zeta)}{\zeta - z}\, d\zeta$$

analytisch in $\mathbb{C} \setminus \bigcup\limits_{k=1}^{m} \Gamma_k$. Ist (z_n) eine Folge in Q_j^0 mit $z_n \to z$, so gilt für $n \to \infty$

$$f(z) \leftarrow f(z_n) = \frac{1}{2\pi i} \sum_{k=1}^{m} \int\limits_{\Gamma_k} \frac{f(\zeta)}{\zeta - z_n}\, d\zeta \to \frac{1}{2\pi i} \sum_{k=1}^{m} \int\limits_{\Gamma_k} \frac{f(\zeta)}{\zeta - z}\, d\zeta\,.$$

Damit gilt die Behauptung für alle $z \in K$. □

Nun zum

Beweis zu S. 34.2. Es seien $\Gamma_1, \ldots, \Gamma_m$ (abhängig von K und Ω) wie in S. 34.3.
Dann betrachten wir (mit: $\Gamma := \bigcup\limits_{k=1}^{m} \Gamma_k$) die Funktion $v : \Gamma \times K \to \mathbb{C}$

$$v(\zeta, z) := \frac{f(\zeta)}{\zeta - z} \quad \big((\zeta, z) \in \Gamma \times K\big)\,.$$

Da v stetig auf $\Gamma \times K$ und $\Gamma \times K$ kompakt ist, ist v gleichmäßig stetig auf $\Gamma \times K$ (S. 11.19).
Ist $\varepsilon > 0$ gegeben, so existiert ein $\delta_\varepsilon > 0$ mit

$$|v(\zeta, z) - v(\zeta', z)| < \varepsilon \text{ für alle } \zeta, \zeta' \in \Gamma,\ |\zeta - \zeta'| < \delta_\varepsilon \text{ und für alle } z \in K.$$

Wir unterteilen Γ in Teilstrecken $S_1, \ldots, S_M$ der Länge $< \delta_\varepsilon$ (und jeweils gleicher Orientierung wie vorher). Wählt man $\zeta_k \in S_k$, so gilt

$$\Bigg| \int\limits_{S_k} \frac{f(\zeta)}{\zeta - z}\, d\zeta - \int\limits_{S_k} \frac{f(\zeta_k)}{\zeta_k - z}\, d\zeta \Bigg| < \varepsilon \cdot L(S_k)\,,$$

wobei $L(S_k)$ die Länge von S_k bezeichnet. Also erhalten wir insgesamt mit

$$\alpha_k := -f(\zeta_k) \cdot \frac{1}{2\pi i} \int_{S_k} d\zeta$$

aus S. 34.3

$$\begin{aligned}\Big|f(z) - \sum_{k=1}^{M} \frac{\alpha_k}{z-\zeta_k}\Big| &= \Big|\frac{1}{2\pi i} \sum_{k=1}^{M} \int_{S_k} \frac{f(\zeta)}{\zeta - z}\, d\zeta - \frac{1}{2\pi i} \sum_{k=1}^{M} \int_{S_k} \frac{f(\zeta_k)}{\zeta_k - z}\, d\zeta\Big| \\ &\leq \frac{1}{2\pi i} \sum_{k=1}^{M} \Big| \int_{S_k} \frac{f(\zeta)}{\zeta - z}\, d\zeta - \int_{S_k} \frac{f(\zeta_k)}{\zeta_k - k} \Big| \leq \frac{\varepsilon}{2\pi} \cdot L(\Gamma)\end{aligned}$$

(beachte: $L(\Gamma) = \sum_{k=1}^{M} L(S_k)$ ist unabhängig von ε). Da $\varepsilon > 0$ beliebig war, folgt die Behauptung. □

Wir werden nun sehen, dass wir die Polstellen in einem gewissen Rahmen verschieben können.

Satz 34.4 *(Rungescher Approximationssatz für rationale Approximation mit vorgeschriebenen Polstellen)*
Es sei $K \subset \mathbb{C}$ kompakt, und es sei $A \subset \mathbb{C} \setminus K$ eine Menge mit der Eigenschaft, dass $A \cap G \neq \emptyset$ gilt für jede beschränkte Komponente G von $\mathbb{C} \setminus K$. Ist $f \in H(\Omega)$ für eine offene Menge $\Omega \supset K$, so existiert zu jedem $\varepsilon > 0$ eine rationale Funktion R, deren Pole alle in A (bzw. ∞) liegen, und für die gilt

$$\max_{z \in K} |f(z) - R(z)| < \varepsilon\,.$$

Bevor wir zum Beweis des Satzes kommen, zeigen wir, dass man als einfache Folgerung auch eine Aussage über polynomiale Approximation erhält:

Satz 34.5 *(Rungescher Approximationssatz für polynomiale Approximation)*
Es sei $K \subset \mathbb{C}$ kompakt und so, dass $\mathbb{C} \setminus K$ zusammenhängend ist. Ist $f \in H(\Omega)$ für eine offene Menge $\Omega \supset K$, so existiert zu jedem $\varepsilon > 0$ ein Polynom P mit

$$\max_{z \in K} |f(z) - P(z)| < \varepsilon\,.$$

Beweis. Da $\mathbb{C} \setminus K$ zusammenhängend ist, kann man $A = \emptyset$ in S. 34.4 wählen. Rationale Funktionen, die keine Pole in $\mathbb{C}$ haben, sind Polynome. Also ergibt sich die Behauptung aus S. 34.4 □

Doch nun zum **Beweis zu S. 34.4** selbst. Im Wesentlichen wird es darum gehen, bei rationalen Funktionen wie in S. 34.2 Pole zu „verschieben".

1. („Polverschiebung in $\mathbb{C}$")

Wir zeigen: Ist G eine Komponente von $\mathbb{C} \setminus K$ und ist $a \in G$, so ist für alle $\zeta \in G$ die Funktion $z \mapsto \frac{1}{z-\zeta}$ auf K „gleichmäßig durch Polynome in $\frac{1}{z-a}$ approximierbar" (d.h. für jedes $\varepsilon > 0$ existiert ein Polynom P mit

$$\max_{z \in K} \left| \frac{1}{z-\zeta} - P\Big(\frac{1}{z-a}\Big) \right| < \varepsilon \,).$$

Dazu definieren wir

$$V := \{\zeta \in G : \zeta \text{ hat diese Eigenschaft}\}$$

und zeigen: $V = G$.

Da G zusammenhängend ist, genügt es wiederum zu zeigen, dass V offen und abgeschlossen in G ist (beachte: $V \neq \emptyset$, da $a \in V$).

Also: V ist abgeschlossen, denn ist (ζ_n) eine Folge in V mit $\zeta_n \to \zeta \in G$, so gilt

$$\frac{1}{z-\zeta_n} \to \frac{1}{z-\zeta} \qquad \text{gleichmäßig in } z \in K\,.$$

(Denn: Ist $\delta := \operatorname{dist}(\zeta, K) (> 0)$, so gilt $\operatorname{dist}(\zeta_n, K) \geq \delta/2$ für $n \geq n_0$, also folgt

$$\max_{z \in K} \left| \frac{1}{z-\zeta_n} - \frac{1}{z-\zeta} \right| = \max_{z \in K} \left| \frac{\zeta_n - \zeta}{(z-\zeta_n)(z-\zeta)} \right| \leq \frac{2}{\delta^2} |\zeta_n - \zeta| \to 0 \qquad (n \to \infty).)$$

Für alle $n \in \mathbb{N}$ existiert ein Polynom P_n mit

$$\max_{z \in K} \left| \frac{1}{z-\zeta_n} - P_n\Big(\frac{1}{z-a}\Big) \right| < \frac{1}{n}$$

und damit

$$\max_{z \in K} \left| \frac{1}{z-\zeta} - P_n\Big(\frac{1}{z-a}\Big) \right| \leq \max_{z \in K} \left| \frac{1}{z-\zeta} - \frac{1}{z-\zeta_n} \right| + \frac{1}{n} \to 0 \qquad (n \to \infty)\,.$$

Folglich ist $\zeta \in A$.

Um zu zeigen, dass V offen ist, sei $\zeta_0 \in V$ gegeben. Dann ist $\delta := \operatorname{dist}(\zeta_0, K) > 0$. Ist $\zeta \in U_{\delta/2}(\zeta_0)$, so gilt für $z \in K$

$$\frac{1}{z-\zeta} = \frac{1}{z-\zeta_0+\zeta_0-\zeta} = \frac{1}{z-\zeta_0} \cdot \sum_{\nu=0}^{\infty} \Big(\frac{\zeta-\zeta_0}{z-\zeta_0}\Big)^{\nu}.$$

Dabei ist

$$\left| \frac{\zeta-\zeta_0}{z-\zeta_0} \right| \leq \frac{1}{2} \qquad \text{für alle } z \in K\,,$$

also ist die Reihe gleichmäßig konvergent auf K. Da die Teilsummen Polynome in $(z-\zeta_0)^{-1}$ sind, ist auch $\zeta \in V$ (Man beachte: Mit $(z-\zeta_0)^{-1}$ ist auch jedes Polynom

in $(z-\zeta_0)^{-1}$ auf K gleichmäßig durch Polynome in $(z-a)^{-1}$ approximierbar.). Damit ist $U_{\delta/2}(\zeta_0) \subset V$, also V offen.
2. („Polverschiebung nach ∞")
Nächste Behauptung: Ist G die unbeschränkte Komponente von $\mathbb{C} \setminus K$, so ist für alle $\zeta \in G$ die Funktion $z \mapsto \frac{1}{z-\zeta}$ auf K gleichmäßig durch Polynome approximierbar.
Denn: Es sei $\varepsilon > 0$ gegeben. Ist $a > 0$ so, dass $\{|z| \geq a\} \subset G$ gilt, so existiert nach 1. ein Polynom P mit

$$\max_{z\in K} \left|\frac{1}{z-\zeta} - P\Big(\frac{1}{z-a}\Big)\right| < \frac{\varepsilon}{2}.$$

Da $\varphi(z) := P(\frac{1}{z-a})$ holomorph in $|z| < a$ ist und da $K \subset \{|z| < a\}$ kompakt ist, existiert ein $n \in \mathbb{N}$ mit

$$\max_{z\in K} \left|P\Big(\frac{1}{z-a}\Big) - \sum_{\nu=0}^{n} \frac{\varphi^{(\nu)}(0)}{\nu!}\, z^{\nu}\right| < \frac{\varepsilon}{2}.$$

Also ist

$$\max_{z\in K} \left|\frac{1}{z-\zeta} - \sum_{\nu=0}^{n} \frac{\varphi^{(\nu)}(0)}{\nu!}\, z^{\nu}\right| < \varepsilon.$$

3. Es sei $\varepsilon > 0$ gegeben. Dann existieren nach S. 34.2 Punkte $\zeta_1, \ldots, \zeta_M \in \mathbb{C} \setminus K$ und $\alpha_1, \ldots, \alpha_M \in \mathbb{C}$ mit

$$\max_{z\in K} \left|f(z) - \sum_{k=1}^{M} \frac{\alpha_k}{z-\zeta_k}\right| < \frac{\varepsilon}{2}.$$

Ist G_k die Komponente von $\mathbb{C} \setminus K$, die ζ_k enthält, und ist G_k beschränkt, so existiert ein $a_k \in A \cap G_k$. Nach 1. existiert also ein Polynom g_k in $(z-a_k)^{-1}$ mit

$$\max_{z\in K} \left|\frac{\alpha_k}{z-\zeta_k} - g_k(z)\right| < \frac{\varepsilon}{2M}.$$

Ist G_k unbeschränkt, so existiert nach 2. ein Polynom g_k mit

$$\max_{z\in K} \left|\frac{\alpha_k}{z-\zeta_k} - g_k(z)\right| < \frac{\varepsilon}{2M}.$$

Die Funktion $R := \sum_{k=1}^{M} g_k$ ist eine rationale Funktion, deren Pole alle in A liegen, und für die gilt

$$\max_{z\in K} |f(z) - R(z)| \leq \frac{\varepsilon}{2} + \sum_{k=1}^{M} \max_{z\in K} \left|\frac{\alpha_k}{z-\zeta_k} - g_k(z)\right| < \varepsilon.$$

Also ist R wie gewünscht. □

Im Folgenden werden wir uns mit diversen Anwendungsbeispielen zum Runge-Satz beschäftigen. Dabei geht es meist um die Konstruktion gewisser Funktionen oder Funktionenfolgen mit vorgegebenen Eigenschaften.

Beispiel 34.6 1. Es existiert eine Folge (P_n) von Polynomen mit

$$P_n(0) \to \infty \qquad \text{und} \qquad P_n(z) \to 0 \qquad \text{für alle } z \neq 0\,.$$

Denn: Für $n \in \mathbb{N}$ sei

$$L_n := \{z = re^{i\varphi} : \frac{1}{n} \leq r \leq n\,,\ \frac{2\pi}{n} \leq \varphi \leq 2\pi\}$$

und

$$K_n := \overline{U_{\frac{1}{2n}}(0)} \cup L_n\,.$$

Dann ist $K_n \subset \mathbb{C}$ kompakt und $\mathbb{C} \setminus K$ zusammenhängend. Also existiert nach S. 34.5, angewandt auf die Funktion $f_n : K_n \to \mathbb{C}$ mit

$$f_n(z) = \begin{cases} 0, & z \in L_n \\ n, & z \in \overline{U_{\frac{1}{2n}}(0)} \end{cases}$$

(man beachte: $f_n \in H(K_n)$ für alle $n \in \mathbb{N}$), ein Polynom P_n mit

$$\max_{z \in K_n} |f_n(z) - P_n(z)| \,<\, \frac{1}{n}\,.$$

Ist $z \neq 0$, so ist $z \in L_n$ für n genügend groß. Also gilt für n genügend groß

$$|P_n(z)| \;=\; |P_n(z) \;-\; f_n(z)| \;<\; \frac{1}{n} \;\to\; 0 \qquad (n \to \infty)\,.$$

Ist $z = 0$, so gilt für alle $n \in \mathbb{N}$

$$|P_n(0)| \geq |f_n(0)| - |f_n(0) - P_n(0)| > n - 1/n \to \infty \quad (n \to \infty)\,.$$

2. Es existiert eine Folge (P_n) von Polynomen mit

$$P_n(z) \to 0 \qquad \text{für alle } z \in \mathbb{C}$$

und

$$\max_{|z| \leq q} |P_n(z)| \;\to\; \infty \qquad \text{für alle } q > 0\,.$$

Denn: Es sei

$$L_n := \big\{z : |z| \leq n\,; \quad \mathrm{Im}(z) \leq 0 \ \text{ oder } \ \mathrm{Im}(z) \geq \frac{1}{n}\big\}\,,$$

$$M_n := \big\{z : |z| \leq n\,; \quad \frac{1}{3n} \leq \mathrm{Im}\, z \leq \frac{2}{3n}\big\}$$

und $K_n := L_n \cup M_n$. Dann ist $K_n \subset \mathbb{C}$ kompakt und $\mathbb{C} \setminus K_n$ zusammenhängend. Also existiert nach S. 34.5, angewandt auf $f_n : K \to \mathbb{C}$ mit

$$f_n(z) := \begin{cases} 0, & z \in L_n \\ n, & z \in M_n \end{cases}\,,$$

zu jedem $n \in \mathbb{N}$ ein Polynom P_n mit

$$\max_{z \in K_n} |f_n(z) - P_n(z)| \,<\, \frac{1}{n}\,.$$

Ist $z_0 \in \mathbb{C}$, so existiert ein $n_0 \in \mathbb{N}$ mit $z_0 \in L_n$ für alle $n \geq n_0$. Also gilt $P_n(z_0) \to 0\,(n \to \infty)$. Andererseits gilt: Ist $q > 0$, so existiert ein $n_1 \in \mathbb{N}$ mit

$$M_n \cap \overline{U_q(0)} \neq \emptyset \qquad (n \geq n_1)\,.$$

Hieraus folgt mit $z_n \in M_n \cap \overline{U_q(0)}$ für $n \geq n_1$

$$\max_{|z| \leq q} |P_n(z)| \;\geq\; |P_n(z_n)| \;\geq\; n - 1/n \;\to\; \infty \qquad (n \to \infty)\,.$$

Wir zeigen im folgenden Satz, dass es (in $\mathbb{D}$) holomorphe Funktionen gibt, die ein sehr kompliziertes Randverhalten aufweisen.

Satz 34.7 *Es existiert eine Funktion $f \in H(\mathbb{D})$ so, dass für alle $\varphi \in (0, 2\pi]$ und alle $w \in \mathbb{C} \cup \{\infty\}$ eine Folge (r_j) existiert mit $0 < r_j \uparrow 1$ und*

$$f(r_j e^{i\varphi}) \;\to\; w \qquad (j \to \infty)\,.$$

Beweis. Es sei (s_m) eine Folge mit $0 < s_m \uparrow 1$, und es sei

$$\Gamma_m := \left\{ s_m e^{i\vartheta} :\; 0 \leq \vartheta \leq 2\pi(1 - 1/m) \right\}.$$

Ferner sei $\{w_k : k \in \mathbb{N}\}$ eine Abzählung der Menge $\mathbb{Q} + i\mathbb{Q}(:= \{u + iv : u, v \in \mathbb{Q}\})$. Da $\mathbb{Q}+i\mathbb{Q}$ dicht in $\mathbb{C}$ ist, erhalten wir damit: Für alle $\varphi \in (0, 2\pi]$ und alle $w \in \mathbb{C} \cup \{\infty\}$ existiert eine Folge $(k_j)_{j \in \mathbb{N}}$ in $\mathbb{N}$ mit $s_{k_j} e^{i\varphi} \in \Gamma_{k_j}$ und $w_{k_j} \to w$ $(j \to \infty)$.
Wir definieren zunächst induktiv eine Folge (P_m) von Polynomen. Dazu setzen wir $P_0(z) \equiv 0$ und gehen davon aus, dass $P_0, \ldots, P_{m-1}$ bereits konstruiert sind. Da $K_m := \Gamma_m \cup \overline{U_{s_{m-1}}(0)}$ kompakt und $\mathbb{C} \setminus K_m$ zusammenhängend ist, existiert nach S. 34.5 ein Polynom P_m mit

$$\max_{|z| \leq s_{m-1}} |P_m(z) - P_{m-1}(z)| \,<\, \frac{1}{2^m}$$

und

$$\max_{z \in \Gamma_m} |P_m(z) - w_m| \,<\, \frac{1}{2^m}\,.$$

Für die so definierte Folge (P_m) erhalten wir: Die Reihe

$$f(z) := \sum_{m=1}^{\infty} \big(P_m(z) - P_{m-1}(z)\big)$$

konvergiert lokal gleichmäßig in $\mathbb{D}$.
(Denn: Ist $q < 1$, so gilt $\overline{U_q(0)} \subset \overline{U_{s_{m-1}}(0)}$ für $m \geq m_q$, alo nach Konstruktion

$$|P_m(z) - P_{m-1}(z)| \leq \frac{1}{2^m}$$

für alle z mit $|z| \leq q$ und alle $m \geq m_q$. Somit ist die Reihe nach S. 15.6 gleichmäßig konvergent in $|z| \leq q$.)
Nach S. 31.7 ist insbesondere $f \in H(\mathbb{D})$. Außerdem gilt

$$P_m = \sum_{\nu=1}^{m} (P_\nu - P_{\nu-1}),$$

also

$$f = P_m + \sum_{\nu=m+1}^{\infty} (P_\nu - P_{\nu-1}).$$

Damit erhalten wir für $z \in \Gamma_m$

$$|f(z) - w_m| \leq |P_m(z) - w_m| + \sum_{\nu=m+1}^{\infty} |P_\nu(z) - P_{\nu-1}(z)| \leq \frac{1}{2^m} + \sum_{\nu=m+1}^{\infty} \frac{1}{2^\nu} = \frac{1}{2^{m-1}}.$$

Sind φ und w gegeben, so folgt mit $r_j := s_{k_j}$ von oben

$$|f(r_j e^{i\varphi}) - w| \leq |f(r_j e^{i\varphi}) - w_{k_j}| + |w_{k_j} - w| \to 0 \qquad (j \to \infty).$$

□

In den bisherigen Anwendungsbeispielen hatten wir es mit Funktionen bzw. Funktionenfolgen auf $\mathbb{D}$ bzw. $\mathbb{C}$ zu tun. Wir betrachten nun allgemeinere offene Mengen Ω. Wichtig ist dabei, geeignete „Ausschöpfungen“, bestehend aus kompakten Mengen, zur Verfügung zu haben.

Satz 34.8 *Es sei $\Omega \subset \mathbb{C}$ offen. Dann existiert eine Folge (K_n) kompakter Mengen mit $\Omega = \bigcup_{n\in\mathbb{N}} K_n$ und folgenden Eigenschaften:*

(i) $K_n \subset K_{n+1}^0$ für alle $n \in \mathbb{N}$.

(ii) Ist $K \subset \Omega$ kompakt, so gilt $K \subset K_n$ für n genügend groß.

(iii) Ist G eine beschränkte Komponente von K_n^c, so existiert eine (beschränkte) Komponente E von $\mathbb{C} \setminus \Omega$ mit $E \subset G$.

Beweis. Wir setzen (mit $\operatorname{dist}(z, \emptyset) := \infty$)

$$K_n := \left\{ z \in \Omega : |z| \leq n\,,\ \operatorname{dist}(z, \Omega^c) \geq \frac{1}{n} \right\} \qquad (n \in \mathbb{N}).$$

Dann ist $K_n \subset \Omega$ kompakt ([Ü]) und $K_n \subset K_{n+1}^0$ für alle $n \in \mathbb{N}$ (ist $z \in K_n$, so ist $U_{\frac{1}{n}-\frac{1}{n+1}}(z) = U_{\frac{1}{n(n+1)}}(z) \subset K_{n+1}$). Damit gilt (i).
Ist ferner $K \subset \Omega$ kompakt, so ist $\operatorname{dist}(K, \Omega^c) > 0$ und K beschränkt. Also ist $K \subset K_n$ für $n \geq n(K)$, d.h. (ii) ist erfüllt (und es gilt natürlich auch $\bigcup_{n\in\mathbb{N}} K_n = \Omega$).
Schließlich gilt

$$K_n^c = V_{n,\infty}(0) \cup \bigcup_{\zeta\in\Omega^c} U_{\frac{1}{n}}(\zeta) \qquad (n \in \mathbb{N})\,.$$

Insbesondere enthält die unbeschränkte Komponente von K_n^c die Menge $V_{n,\infty}(0)$).
Ist also G eine beschränkte Komponente von K_n^c (falls existent), so folgt $G \subset \{z : |z| < n\}$, und es existiert ein $\zeta \in \Omega^c$ mit $U_{\frac{1}{n}}(\zeta) \subset G$, also insbesondere $\zeta \in G$. Also ist $G = Z_{K_n^c}(\zeta)$. Da $\Omega^c \subset K_n^c$ gilt, ist weiterhin $Z_{\Omega^c}(\zeta) \subset Z_{K_n^c}(\zeta)$ nach Definition (vgl. B/D. 30.10). Also ist $E := Z_{\Omega^c}(\zeta) \subset G$. □

Damit erhalten wir

Satz 34.9 *(Rungescher Approximationssatz für offene Mengen)*
Es sei $\Omega \subset \mathbb{C}$ offen, und es sei $f \in H(\Omega)$. Dann gilt

1. *Ist $A \subset \Omega^c$ so, dass A jede beschränkte Komponente von Ω^c trifft, so existiert eine Folge rationaler Funktionen (R_n) mit $R_n \in H(\mathbb{C} \setminus A)$ (d.h. sämtliche Pole liegen in A) und*
$$R_n \to f \qquad \text{lokal gleichmäßig auf } \Omega.$$
2. *Hat Ω^c keine beschränkte Komponente, so existiert eine Folge (P_n) von Polynomen mit*
$$P_n \to f \qquad \text{lokal gleichmäßig auf } \Omega.$$

Beweis. 1. Ist (K_n) die Folge aus S. 34.8, so trifft nach Voraussetzung und Eigenschaft (iii) aus S. 34.8 die Menge A jede beschränkte Komponente von K_n^c $(n \in \mathbb{N})$. Also existiert nach S. 34.4 zu jedem $n \in \mathbb{N}$ eine rationale Funktion R_n mit $R_n \in H(\mathbb{C} \setminus A)$ (also alle Pole in A) und

$$\max_{z\in K_n} |f(z) - R_n(z)| < \frac{1}{n}\,.$$

Ist $z_0 \in \Omega$, so existiert ein $\delta > 0$ mit $K := \overline{U_\delta(z_0)} \subset \Omega$. Also ist $K \subset K_n$ für alle n genügend groß (S. 34.8 (ii)). Damit gilt

$$\max_{z\in K} |f(z) - R_n(z)| < \frac{1}{n}$$

für n genügend groß, d.h. $R_n \to f$ gleichmäßig auf K. Da $z_0 \in \Omega$ beliebig war, gilt die erste Behauptung.
2. Ergibt sich als Spezialfall mit $A = \emptyset$ (vgl. S. 34.5). □

Bemerkung und Definition 34.10 Ist $\Omega \subset \mathbb{C}$ offen, so nennen wir eine beschränkte Komponente von Ω^c (falls existiert) ein *Loch* von Ω. Der zweite Teil von S. 34.9 besagt dann, dass für jede offene Menge Ω ohne Löcher jede Funktion $f \in H(\Omega)$ „lokal gleichmäßig auf Ω durch Polynome approximierbar ist“. Eine für uns wichtige Folgerung hieraus:

Ist $G \subset \mathbb{C}$ ein Gebiet ohne Löcher, so ist G einfach zusammenhängend.
(Denn: Ist $f \in H(G)$, so existiert eine Folge (P_n) von Polynomen mit $P_n \to f$ lokal gleichmäßig auf G. Ist γ ein geschlossener Pfad in G, so folgt aus S. 31.7

$$0 = \int_\gamma P_n \to \int_\gamma f \qquad (n \to \infty)\,,$$

also $\int_\gamma f = 0$. Damit ist G einfach zusammenhängend (man beachte, dass der „Ausnahmepunkt“ in S. 30.8 nach B. 30.17 keine Rolle spielt).)
Damit haben wir eine angenehme (topologische) hinreichende Bedingung für den einfachen Zusammenhang eines Gebietes. Man kann zeigen (ohne Beweis!): Ein Gebiet G ist genau dann einfach zusammenhängend, wenn G keine Löcher hat.

Als weitere Anwendung des Runge-Satzes beweisen wir, dass meromorphe Funktionen zu beliebig vorgeschriebenen Hauptteilen existieren.

Satz 34.11 *(Mittag-Leffler)*
Es sei $\Omega \subset \mathbb{C}$ offen, und es sei $A \subset \Omega$ ohne Häufungspunkt in Ω. Ferner seien P_w für $w \in A$ Polynome. Dann existiert eine in $\Omega \setminus A$ holomorphe Funktion f so, dass für alle $w \in A$ der Hauptteil der Laurent-Entwicklung um w gerade $P_w((1/(z-w))$ ist.

Beweis. Es sei (K_n) eine Ausschöpfung von Ω wie in S. 34.8. Für $n \in \mathbb{N}$ setzen wir $B_n := A \cap (K_{n+1} \setminus K_n)$ und

$$f_n(z) := \sum_{w \in B_n} P_w\Big(\frac{1}{z-w}\Big) \qquad (z \in \mathbb{C} \setminus B_n)$$

(beachte: B_n ist endlich, da A keinen Häufungspunkt in Ω hat).
Dann ist $f_n \in H(\mathbb{C} \setminus B_n)$, und es gilt $K_n \subset \mathbb{C} \setminus B_n$.
Da jede beschränkte Komponente von K_n^c einen Punkt aus Ω^c enthält, existiert nach S. 34.4 zu jedem $n \in \mathbb{N}$ eine rationale Funktion R_n mit $R_n \in H(\Omega)$ (also alle Pole außerhalb von Ω) und

$$\max_{z \in K_n} |f_n(z) - R_n(z)| < \frac{1}{2^n}\,.$$

Mit $f_0(z) := \sum\limits_{w \in A \cap K_1} P_w(\frac{1}{z-w})$ gilt:

$$f(z) := f_0(z) + \sum_{\nu=1}^{\infty} (f_\nu(z) - R_\nu(z)) \qquad (z \in \Omega \setminus A)$$

konvergiert gleichmäßig auf jeder kompakten Teilmenge K von $\Omega \setminus A$ (S. 15.6, und $K \subset K_n$ für n genügend groß). Damit ist $f \in H(\Omega \setminus A)$. Da für alle $n \in \mathbb{N}$

$$f = \sum_{\nu=0}^{n} f_\nu + \underbrace{\sum_{\nu=n+1}^{\infty} (f_\nu - R_\nu) - \sum_{\nu=1}^{n} R_\nu}_{=:\, g_n}$$

mit $g_n \in H(\Omega \setminus \bigcup_{\nu=n+1}^{\infty} B_\nu)$ gilt, und da $K_{n+1} \subset \Omega \setminus \bigcup_{\nu=n+1}^{\infty} B_\nu$ ist, hat f in K_{n+1} genau die vorgeschriebenen Pole mit entsprechenden Hauptteilen. Da $n \in \mathbb{N}$ beliebig war, folgt die Behauptung. $\square$

35 Harmonische Funktionen

Im letzten Abschnitt beschäftigen wir uns mit einer Klasse von Funktionen, die in gewisser Weise an der Schnittstelle zwischen reeller und komplexer Analysis zu finden sind.

Definition 35.1 Es sei $\Omega \subset \mathbb{R}^d$ offen. Eine Funktion $u : \Omega \to \mathbb{R}$ heißt *harmonisch* (in Ω), falls $u \in C^2(\Omega)$ ist, und falls

$$\triangle u := \sum_{j=1}^{d} \frac{\partial^2 u}{\partial x_j^2} \Big(= \sum_{j=1}^{d} \partial_j^2 u \Big) \equiv 0$$

auf Ω gilt. Wir setzen

$$\mathrm{Har}(\Omega) := \{ f : \Omega \to \mathbb{R} : f \text{ harmonisch in } \Omega \} .$$

Wir werden uns im Folgenden auf die Untersuchung harmonischer Funktionen von zwei Variablen beschränken (also $\Omega \subset \mathbb{R}^2$). In diesem Fall ergibt sich ein sehr direkter Zusammenhang zwischen harmonischen und holomorphen Funktionen. Wir fassen dazu Ω wieder als Teilmenge von $\mathbb{C}$ auf.

Satz 35.2 *Es sei $G \subset \mathbb{C}$ ein Gebiet. Dann gilt:*

1. *Ist $f \in H(G)$, so ist $u = \mathrm{Re}(f) \in \mathrm{Har}(G)$.*
2. *Ist $u \in \mathrm{Har}(G)$ und ist G einfach zusammenhängend, so existiert eine Funktion $f \in H(G)$ mit $u = \mathrm{Re}(f)$. Dabei ist f bis auf eine additive Konstante (der Form ic mit $c \in \mathbb{R}$) eindeutig bestimmt.*

Beweis. 1. Es sei $f = u + iv$ mit $u, v =: \Omega \to \mathbb{R}$. Dann gilt nach S. 29.4

$$f' = \frac{\partial u}{\partial x} + i\,\frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i\,\frac{\partial u}{\partial y} \qquad \text{auf } G .$$

Da f analytisch in Ω ist, zeigt die neuerliche Anwendung von S. 29.4 auf f', dass alle partiellen Ableitungen 2. Ordnung von u und v auf G existieren und stetig sind. Also gilt $u, v \in C^2(G)$. Außerdem ergibt sich aus den Cauchy-Riemannschen Differenzialgleichungen (29.1) und S. 20.12

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial}{\partial x}\Big(\frac{\partial v}{\partial y}\Big) - \frac{\partial}{\partial y}\Big(\frac{\partial v}{\partial x}\Big) \equiv 0 \text{ auf } G .$$

2. Wir betrachten

$$g := \frac{\partial u}{\partial x} - i\,\frac{\partial u}{\partial y} .$$

Dabei sind $\partial u/\partial x,\ \partial u/\partial y \in C^1(G)$, und es gilt

$$\frac{\partial}{\partial x}\frac{\partial u}{\partial x} \overset{\Delta u=0}{=} \frac{\partial}{\partial y}\Big(-\frac{\partial u}{\partial y}\Big) \qquad \text{sowie} \qquad \frac{\partial}{\partial y}\frac{\partial u}{\partial x} \overset{\text{S.20.12}}{=} -\frac{\partial}{\partial x}\Big(-\frac{\partial u}{\partial y}\Big).$$

Also erfüllt g die Cauchy-Riemannschen Gleichungen (29.1), und damit ist $g \in H(G)$. Da G einfach zusammenhängend ist, existiert nach S. 30.5 eine Stammfunktion f zu g, d.h.

$$f' = g = \frac{\partial u}{\partial x} - i\,\frac{\partial u}{\partial y} \ \text{ auf } G.$$

Für $\tilde{u} := \mathrm{Re} f$ gilt

$$\frac{\partial \tilde{u}}{\partial x} - i\,\frac{\partial \tilde{u}}{\partial y} \overset{\text{S.29.4}}{=} f' = g = \frac{\partial u}{\partial x} - i\,\frac{\partial u}{\partial y}.$$

also $\partial u/\partial x - \partial \tilde{u}/\partial x \equiv 0$ und $\partial \tilde{u}/\partial y - \partial u/\partial y \equiv 0$ auf G. Damit ist $u - \tilde{u} \equiv \text{const}$ auf G. Durch Addition einer geeigneten Konstante können wir weiter $f(z_0) = u(z_0)$ für ein $z_0 \in G$ erreichen. Dann ist aber schon $u \equiv \tilde{u}$ auf G, d.h. $u = \mathrm{Re} f$. Ist $\tilde{f} \in H(G)$ eine weitere Funktion mit $u = \mathrm{Re}\tilde{f}$, so gilt nach S. 29.4

$$\tilde{f}' = \frac{\partial u}{\partial x} - i\,\frac{\partial u}{\partial y} = g \quad \text{auf } G\,,$$

also ist $\tilde{f}$ ebenfalls eine Stammfunktion zu g auf G. Damit unterscheiden sich f und $\tilde{f}$ nur um eine additive Konstante. □

Bemerkung 35.3 Der Beweis zu S. 35.2.2 zeigt genauer: Ist G ein beliebiges Gebiet in $\mathbb{C}$ und ist $u \in \mathrm{Har}(G)$, so ist

$$g := \frac{\partial u}{\partial x} - i\,\frac{\partial u}{\partial y}$$

holomorph in G. Eine Funktion $f \in H(G)$ erfüllt genau dann $u = \mathrm{Re} f$ in G, wenn f eine Stammfunktion von g in G ist mit $\mathrm{Re} f(z_0) = u(z_0)$ für ein $z_0 \in G$.
Insbesondere zeigt dies auch, dass die Aussage von S. 35.2.2 im Allgemeinen falsch wird für nicht einfach zusammenhängende Gebiete. Ist etwa $G = \mathbb{C} \setminus \{0\}$ und $u(z) = \ln|z|$ für $z \neq 0$, so ist $u \in \mathrm{Har}(G)$, und es gilt ([Ü])

$$g(z) = \frac{\partial u}{\partial x}(z) - i\,\frac{\partial u}{\partial y}(z) = \frac{1}{z}.$$

Jedoch hat bekanntlich g keine Stammfunktion in $\mathbb{C} \setminus \{0\}$ (S. 30.5 und $\int\limits_{|z|=1} dz/z = 2\pi i \neq 0$). Also existiert keine Funktion $f \in H(G)$ mit $u = \mathrm{Re} f$.

Aus der Tatsache, dass harmonische Funktionen sich (lokal) als Realteile holomorpher Funktionen ergeben, lassen sich viele wichtige Eigenschaften von holomorphen auf harmonische Funktionen übertragen. Wir sammeln einige im folgenden Satz.

Satz 35.4 *Es sei $G \subset \mathbb{C}$ ein Gebiet, und es sei u harmonisch in G. Dann gilt*

1. *(Mittelwert-Eigenschaft)*
   Ist $z_0 \in G$ und $U_\rho(z_0) \subset G$, so gilt für $0 < r < \rho$

$$u(z_0) = \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{it})dt \tag{35.1}$$

2. *(Identitätssatz)*
   Ist v harmonisch in G und gilt $u(z) = v(z)$ auf einer offenen Menge $U \subset G, U \neq \emptyset$, so ist $u \equiv v$ in G.

3. *(Maximumprinzip / Minimumprinzip)*
   Ist $u \not\equiv \mathrm{const}$ in G, so hat u kein lokales Maximum und kein lokales Minimum. Ist G beschränkt und u stetig auf $\overline{G}$, so existieren $z_0, z_1 \in \partial G$ mit

$$u(z_0) = \max_{z \in \overline{G}} u(z) \qquad \textit{und} \qquad u(z_1) = \min_{z \in \overline{G}} u(z)\,.$$

Beweis. 1. Da $U_\rho(z_0)$ einfach zusammenhängend ist, existiert eine Funktion $f \in H(U_\rho(z_0))$ mit $u = \mathrm{Re}f$. Nach der Mittelwertformel (30.3) gilt

$$u(z_0) = \mathrm{Re}(f(z_0)) = \mathrm{Re}\Big(\frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{it})dt\Big) = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{Re}f(z_0 + re^{it})dt\,.$$

2. O. E. sei $v \equiv 0$ (sonst betrachten wir $\tilde{u} = u - v$). Es sei

$$g = \frac{\partial u}{\partial x} - i\,\frac{\partial u}{\partial y}\,.$$

Dann ist $g \in H(G)$ (siehe B. 35.2), und es gilt $g(z) = 0$ für alle $z \in U$ (da $u \equiv 0$ auf U nach Voraussetzung). Nach dem Identitätssatz für holomorphe Funktionen ist $g \equiv 0$ auf G, also $\partial u/\partial x = \partial u/\partial y \equiv 0$ in G. Folglich ist $u \equiv \mathrm{const}$ auf G, und aus $u(z) = 0$ auf U folgt $u \equiv 0$ in G.
3. Genau wie im Beweisteil 2. zu S. 31.10 sieht man (unter Verwendung der Mittelwertformel (35.1)): Hat U an $z_0 \in G$ ein lokales Maximum, so ist $u(z) \equiv u(z_0)$ auf $U_r(z_0)$ für ein $r > 0$. Nach 2. ist dann $u \equiv u(z_0)$ in G.
Die zweite Aussage ergibt sich genau so wie S. 31.11 sich aus S. 30.10 ergibt.
Da mit u uach $-u$ harmonisch in G ist, gelten entsprechende Aussagen auch für Minima. □

Bemerkung 35.5 Für analytische – und damit für holomorphe – Funktionen gilt bekanntlich eine stärkere Version des Identitätssatzes (S. 29.11). Ein entsprechendes Resultat ist i. A. nicht gültig für harmonische Funktionen.
So gilt etwa für

$$u(z) = \mathrm{Re}\, z\,, \qquad v(z) \equiv 0 \qquad (z \in \mathbb{C})\,.$$

offenbar $u = v$ auf $i\mathbb{R}$, aber es ist $u \not\equiv v$ in $\mathbb{C}$.

Eine der zentralen Fragestellungen im Zusammenhang mit harmonischen Funktionen ist die nach der „harmonischen Fortsetzbarkeit“ stetiger Funktionen, die am Rande eines Gebietes vorgegeben sind. Genauer:

Definition 35.6 Es sei $G \subset \mathbb{C}$ ein Gebiet, und es sei $\varphi : \partial G \to \mathbb{R}$ stetig. Unter dem *Dirichlet-Problem* $D(G, \varphi)$ versteht man die Frage nach der Existenz und Eindeutigkeit in G harmonischer Funktionen u mit

$$\lim_{z \to \zeta} u(z) \Big(= \lim_{\substack{z \to \zeta \\ z \in G}} u(z) \Big) = \varphi(\zeta) \qquad \text{für alle } \zeta \in \partial G\,,$$

m. a. W. man sucht Funktionen $u \in C(\overline{G})$ mit $\Delta u \equiv 0$ in G und $u_{|\partial G} = \varphi$.

Die Frage nach der Eindeutigkeit ist für beschränkte Gebiete leicht zu beantworten:

Satz 35.7 *Ist $G \subset \mathbb{C}$ ein beschränktes Gebiet, so existiert zu jedem $\varphi \in C(\partial G)$ höchstens eine Lösung des Dirichlet-Problems $D(G, \varphi)$.*

Beweis. Sind u, v Lösungen von $D(G, \varphi)$, so ist $u - v \in C(\overline{G})$ harmonisch in G, und es gilt $u - v = 0$ auf ∂G. Nach S. 35.4.3 ist damit $u - v \equiv 0$ in $\overline{G}$. □

Bemerkung 35.8 Mit einem ähnlichen Argument sieht man auch, dass nicht jedes Dirichlet-Problem $D(G, \varphi)$ lösbar ist. Ist etwa $G = \mathbb{D} \setminus \{0\}$ und

$$\varphi(\zeta) := \begin{cases} 0\,, & \zeta \in \partial\mathbb{D} \\ 1\,, & \zeta = 0 \end{cases}\,,$$

so existiert keine Lösung von $D(G, \varphi)$.
(Man kann zeigen (ohne Beweis): Ist $u \in \mathrm{Har}(G)$ und stetig auf $\mathbb{D}$, so ist $u \in \mathrm{Har}(\mathbb{D})$. Aus $u = 0$ auf $\partial\mathbb{D}$ folgt dann nach S. 34.5.3 schon $u \equiv 0$ in $\overline{\mathbb{D}}$, also $u(0) = 0 \neq 1$.)

Die Frage nach der Existenz von Lösungen erweist sich als recht diffizil. Die Antwort hängt wesentlich von der Struktur von G ab. Wir untersuchen den einfachen Fall eines Kreises $G = U_\rho(z_0)$ für $z_0 \in \mathbb{C}$, $\rho > 0$. Dabei reicht es i. W., den Fall $z_0 = 0$ und $\rho = 1$, also $G = \mathbb{D}$ zu klären.
Dazu stellen wir eine kleine Vorüberlegung an: Angenommen, die Randfunktion φ sei darstellbar durch ihre Fourier-Reihe, d.h.

$$\varphi(e^{i\theta}) = \sum_{\nu=-\infty}^{\infty} a_\nu e^{i\nu\theta} \qquad \text{mit} \qquad a_\nu = \frac{1}{2\pi}\int_0^{2\pi} \varphi(e^{it})e^{-i\nu t}dt$$

(vgl. Abschnitt 28). Dann wäre eine „sinnvolle“ Erweiterung u von φ für $z = re^{i\theta} \in \mathbb{D}$ gegeben durch

$$\begin{aligned} u(z) = u(re^{i\theta}) \;&:=\; \sum_{\nu=-\infty}^{\infty} a_\nu r^{|\nu|}e^{i\nu\theta} \\ &= \sum_{\nu=-\infty}^{\infty} \Big(\frac{1}{2\pi}\int_0^{2\pi} \varphi(e^{it})e^{-i\nu t}dt\Big) r^{|\nu|}e^{i\nu\theta} \\ &= \frac{1}{2\pi}\int_0^{2\pi} \sum_{\nu=-\infty}^{\infty} r^{|\nu|}e^{i\nu(\theta-t)}\varphi(e^{it})dt \end{aligned}$$

(aus $|r^{|\nu|}e^{i\nu(\theta-t)}\varphi(e^{it})| \le r^{|\nu|}||\varphi||_\infty$ folgt mit S. 15.6 die gleichmäßige Konvergenz der Reihe auf $[0, 2\pi]$ bei festem $z = re^{i\theta}$). Mit

$$P_r(s) := \sum_{\nu=-\infty}^{\infty} r^{|\nu|}e^{i\nu s} \qquad (r < 1, s \in [0, 2\pi])$$

ist also

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi} P_r(\theta - t)\varphi(e^{it})dt \qquad (z = re^{i\theta} \in \mathbb{D})\,.$$

Nun wäre es schön, wenn u harmonisch in $\mathbb{D}$ wäre. Bevor wir zeigen, dass dies tatsächlich der Fall ist, stellen wir $P_r(s)$ geschlossen dar: Es gilt

$$\begin{aligned} P_r(s) \;&=\; \sum_{\nu=0}^{\infty} r^\nu e^{i\nu s} + \sum_{\nu=1}^{\infty} r^\nu e^{-i\nu s} \\ &= \frac{1}{1 - re^{is}} + \frac{re^{-is}}{1 - re^{-is}} = \frac{1 - r^2}{1 - 2r\cos(s) + r^2} \end{aligned}$$

oder auch in komplexer Schreibweise mit $z = re^{i\theta}$

$$\begin{aligned} \operatorname{Re}\left(\frac{e^{it} + z}{e^{it} - z}\right) \;&=\; \operatorname{Re}\left(\frac{(e^{it} + z)(e^{-it} - \overline{z})}{|e^{it} - z|^2}\right) \\ &= \frac{1 - |z|}{|e^{it} - z|^2} = \frac{1 - r^2}{|e^{it} - re^{i\theta}|^2} = \frac{1 - r^2}{1 - 2r\cos(\theta - t) + r^2} = P_r(\theta - t) \end{aligned} \tag{35.2}$$

Diese Dinge gießen wir in eine Definition

Definition 35.9 1. Die Funktion $P : \mathbb{D} \times \partial\mathbb{D} \to \mathbb{R}$, definiert durch

$$P(z, e^{it}) := \frac{1 - |z|^2}{|e^{it} - z|^2} = \mathrm{Re}\Big(\frac{e^{it} + z}{e^{it} - z}\Big) \qquad (z \in \mathbb{D}, t \in [0, 2\pi])$$

heißt *Poisson-Kern.*
2. Ist $t \mapsto \varphi(e^{it})$ eine Regelfunktion auf $[0, 2\pi]$, so heißt $T\varphi : \mathbb{D} \to \mathbb{R}$ mit

$$\begin{aligned} T\varphi(z) &:= \frac{1}{2\pi} \int_0^{2\pi} P(z, e^{it})\varphi(e^{it})dt \\ \Big(&= \frac{1}{2\pi} \int_0^{2\pi} P_r(\theta - t)\varphi(e^{it})dt\,, \quad z = re^{i\theta}\Big) \end{aligned}$$

Poisson-Integral von φ.

Wir notieren zunächst einige wichtige Eigenschaften des Poisson-Kerns.

Satz 35.10 *Es gilt*

1. $P(z, e^{it}) > 0$ *für* $z \in \mathbb{D}$, $t \in [0, 2\pi]$,
2. $\displaystyle \frac{1}{2\pi} \int_0^{2\pi} P(z, e^{it})dt = \frac{1}{2\pi} \int_0^{2\pi} P_r(s)ds = 1$ *für* $z \in \mathbb{D}$.
3. *Für* $\zeta_0 \in \partial\mathbb{D}$ *und* $\delta > 0$ *ist*

$$\lim_{z \to \zeta_0} \sup_{|e^{it} - \zeta_0| \geq \delta} P(z, e^{it}) = 0\,.$$

Beweis. 1. Ergibt sich direkt aus (35.2).
2. Siehe B. 33.11.
3. Ist $|z - \zeta_0| < \delta/2$, so gilt mit (35.2) für alle t mit $|e^{it} - \zeta_0| \geq \delta$

$$P(z, e^{it}) \leq \frac{1 - |z|^2}{(|e^{it} - \zeta_0| - |\zeta_0 - z|)^2} \leq \frac{1 - |z|^2}{\delta^2/4} \to 0 \qquad (z \to \zeta_0)\,. \qquad \square$$

Damit können wir das Dirichlet-Problem $(\mathbb{D}, \varphi)$ lösen:

Satz 35.11 *Es sei $t \mapsto \varphi(e^{it})$ eine Regelfunktion. Dann gilt:*
1. $T\varphi$ ist harmonisch in $\mathbb{D}$.
2. Ist φ stetig an $\zeta_0 \in \partial\mathbb{D}$, so ist $\lim\limits_{z\to\zeta_0} T\varphi(z) = \varphi(\zeta_0)$.
Insbesondere ist im Falle $\varphi \in C(\partial\mathbb{D})$ die Funktion $T\varphi$ (die) Lösung von $(\mathbb{D}, \varphi)$.

Beweis. 1. Für $z \in \mathbb{D}$ gilt mit (35.2)

$$T\varphi(z) = \mathrm{Re}\left(\frac{1}{2\pi}\int_0^{2\pi} \frac{e^{it}\varphi(e^{it})}{e^{it}-z}\,dt + \frac{z}{2\pi}\int_0^{2\pi}\frac{\varphi(e^{it})}{e^{it}-z}\,dt\right)$$

Nach S. 29.9 ist

$$f(z) := \frac{1}{2\pi}\int_0^{2\pi} \frac{e^{it}\varphi(e^{it})}{e^{it}-z}\,dt + \frac{z}{2\pi}\int_0^{2\pi}\frac{\varphi(e^{it})}{e^{it}-z}\,dt$$

analytisch (also holomorph) in $\mathbb{D}$. Damit ist $T\varphi \in \mathrm{Har}(\mathbb{D})$ nach S. 35.2.
2. Es sei φ stetig an $\zeta_0 \in \partial\mathbb{D}$. Ist $\varepsilon > 0$ gegeben, so existiert ein $\delta > 0$ mit

$$|\varphi(\zeta) - \varphi(\zeta_0)| < \varepsilon \qquad \text{für alle } \zeta \in \partial\mathbb{D}, |\zeta - \zeta_0| < \delta\,.$$

Für $z \in \mathbb{D}$ erhalten wir

$$\begin{aligned}|T\varphi(z) - \varphi(\zeta_0)| &\overset{\text{S.35.10.2}}{=} \left|\frac{1}{2\pi}\int_0^{2\pi} P(z,e^{it})(\varphi(e^{it} - \varphi(\zeta_0))\,dt\right| \\ &\overset{\text{S.35.10.1}}{\leq} \frac{1}{2\pi}\int_0^{2\pi} P(z,e^{it})|\varphi(e^{it}) - \varphi(\zeta_0)|\,dt\,.\end{aligned}$$

Wieder mit S. 35.10.2 folgt

$$\frac{1}{2\pi}\int\limits_{|e^{it}-\zeta_0|<\delta} P(z,e^{it})|\varphi(e^{it}) - \varphi(\zeta_0)|dt < \varepsilon\,.$$

Außerdem existiert nach S. 35.10.3 ein $\delta' > 0$ mit

$$\sup_{|e^{it}-\zeta_0|\geq\delta} P(z,e^{it}) < \varepsilon \qquad \text{für alle } z \in \mathbb{D} \text{ mit } |z-\zeta_0| < \delta.$$

Also gilt für $|z - \zeta_0| < \delta'$:

$$\frac{1}{2\pi}\int\limits_{|e^{it}-\zeta_0|\geq\delta} P(z,e^{it})|\varphi(e^{it}) - \varphi(\zeta_0)|dt \leq \frac{\varepsilon}{2\pi}\int_0^{2\pi} |\varphi(e^{it}) - \varphi(\zeta_0)|dt$$

$$\leq \varepsilon\Big(\frac{1}{2\pi}\int_0^{2\pi} |\varphi(e^{it})|dt + |\varphi(\zeta_0)|\Big)\,,$$

und damit insgesamt

$$|T\varphi(z) - \varphi(\zeta_0)| \le \varepsilon\Big(1 + \frac{1}{2\pi}\int_0^{2\pi} |\varphi(e^{it})|dt + |\varphi(\zeta_0)|\Big).$$

Da $\varepsilon > 0$ beliebig war, gilt $\lim\limits_{z\to\zeta_0} T\varphi(z) = \varphi(\zeta_0)$. □

Bemerkung 35.12 Es ist klar, dass sich S. 35.11 leicht auf beliebige Kreise $G = U_\rho(z_0)$ für $\rho > 0, z_0 \in \mathbb{C}$ übertragen läßt: Ist φ stetig auf $K_\rho(z_0)$, so ist $T_G\varphi : G \to \mathbb{R}$, definiert durch

$$T_G\varphi(z) := \frac{1}{2\pi}\int_0^{2\pi} P\Big(\frac{z-z_0}{\rho}, e^{it}\Big)\varphi(z_0 + \rho e^{it})dt \qquad (z \in G)$$

$$\Big(= \frac{1}{2\pi}\int_0^{2\pi} \frac{\rho^2 - r^2}{\rho^2 - 2r\rho\cos(\theta - t) + r^2}\,\varphi(z_0 + \rho e^{it})dt \qquad (z = z_0 + re^{i\theta})\Big)$$

Lösung des Dirichlet-Problems (G, φ). Als Konsequenz erhält man die sogenannte Poisson-Integralformel: Ist u harmonisch in einer offenen Umgebung von $\overline{U_\rho(z_0)}$, so ist

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{\rho^2 - r^2}{\rho^2 - 2r\rho\cos(\theta - t) + r^2}\,u(z_0 + \rho e^{it})\,dt$$

für alle $z = z_0 + re^{i\theta} \in U_\rho(z_0)$. (Dies ergibt sich aus S. 35.7, da sowohl u als auch $T_G\varphi$ mit $\varphi = u_{|\partial G}$ das Dirichlet-Problem (G, φ) lösen.)
Speziell folgt für $z = z_0$

$$u(z_0) = \frac{1}{2\pi}\int_0^{2\pi} u(z_0 + \rho e^{it})\,dt\,,$$

also wieder die Mittelwertformel (35.1).

Wir zeigen, dass die Mittelwertformel harmonische Funktionen charakterisiert:

Satz 35.13 *Es sei $\Omega \subset \mathbb{C}$ offen, und es sei $h \in C(\Omega, \mathbb{R})$. Existiert zu jedem $w \in \Omega$ ein $\rho(w) > 0$ so, dass*

$$h(w) = \frac{1}{2\pi}\int_0^{2\pi} h(w + \rho e^{it})\,dt$$

für $0 < \rho < \rho(w)$ gilt, so ist h harmonisch in Ω.

Beweis. Es sei $z_0 \in \Omega$ und $R > 0$ mit $\overline{U_R(z_0)} \subset \Omega$. Wir definieren mit $G := U_R(z_0)$ die Funktion $k : \overline{G} \to \mathbb{R}$ durch

$$k(z) := \begin{cases} h(z) - T_G h(z) & , z \in G \\ 0 & , z \in \partial G \end{cases} .$$

Dann ist k stetig auf $\overline{G}$ nach B. 35.12. Außerdem ist $T_G h$ harmonisch in G (vgl. S. 35.11.1). Ist also $w \in G$, so gilt nach Voraussetzung und (35.1) für ρ genügend klein:

$$k(w) = \frac{1}{2\pi} \int_0^{2\pi} k(w + \rho e^{it})\, dt \, ,$$

d.h. k erfüllt ebenfalls die Mittelwert-Eigenschaft (lokal). Da der Beweis des Maximum-/Minimumprinzips lediglich die Mittelwert-Eigenschaft verwendet (vgl. Beweis zu S. 35.4), ergibt sich genau wie dort: Es existieren $z_0, z_1 \in \partial G$ mit

$$k(z_0) = \max_{z \in \overline{G}} k(z) \qquad \text{und} \qquad k(z_1) = \max_{z \in \overline{G}} k(z)$$

aus $k(z_0) = k(z_1) = 0$ folgt $k \equiv 0$ auf $\overline{G}$, d.h. h stimmt auf $U_R(z_0)$ mit der (harmonischen) Funktion $T_G h$ überein. Damit ist h harmonisch in $U_R(z_0)$, und da $z_0 \in \Omega$ beliebig war, folgt die Behauptung. □

Bemerkung 35.14 Als wichtige Konsequenz erhalten wir ([Ü]): Ist $\Omega \subset \mathbb{C}$ offen und sind $u_n \in \mathrm{Har}(\Omega)$ mit $u_n \to u$ lokal gleichmäßig auf Ω, so ist auch $u \in \mathrm{Har}(\Omega)$.