Jürgen Müller

Einführung in die Mathematik/Analysis einer und mehrerer Veränderlicher

Skriptum zur den Vorlesung
Wintersemester 2012/2013 und Sommersemster 2103
Universität Trier
Fachbereich IV
Mathematik/Analysis

Dank an Elke Gawronski für die Mithilfe bei der Erstellung

Inhaltsverzeichnis

1 Mengen und Abbildungen

Wir starten mit einigen einführenden Definitionen und Ergebnissen aus der Theorie der Mengen und Abbildungen, die Grundlage 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 mehrere Möglichkeiten der Darstellung von Mengen: etwa die aufzählende Schreibweise (etwa $M = \{1, 3, 5, 7\}$) oder auch 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

$$M = \{x : x \text{ ungerade natürliche Zahl kleiner als } 9\}.$$

Die Menge ohne Elemente heißt die *leere Menge* (Schreibweise: $\emptyset$)
Wir werden im Weiteren davon ausgehen, dass natürliche, ganze, rationale und reelle Zahlen bekannt sind, werden aber in einem eher informellen eigenen Teil der Vorlesung auf eine axiomatische bzw. konstruktive Einführung zu sprechen kommen.

$$\begin{aligned}
\mathbb{N} &:= \{x : x \text{ natürliche Zahl}\} \\
\mathbb{N}_0 &:= \{x : x \text{ natürliche Zahl oder } x = 0\} \\
\mathbb{Z} &:= \{x : x \text{ ganze Zahl}\} \\
\mathbb{Q} &:= \{x : x \text{ rationale Zahl}\} \\
\mathbb{R} &:= \{x : x \text{ reelle Zahl}\}
\end{aligned}$$

Außerdem setzen wir die üblichen arithmetischen Operationen ($+$ und $\cdot$) und die Ordnungsrelationen ($<$, $>$, $\leq$, $\geq$) als bekannt voraus. Auch diese Dinge werden wir im informellen Teil der Vorlesung thematisieren.
Wir setzen noch für $a, b \in \mathbb{R}$ (mit $a \leq b$)

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

Diese Mengen heißen *Intervalle*.

Definition 1.1 Es seien A, B Mengen.

1. A heißt *Teilmenge* von B (Schreibweie: $A \subset B$), falls aus $x \in A$ auch $x \in B$ folgt.

2. A und B heißen *gleich*, 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 bezüglich B.

Definition 1.2 Es sei $I \neq \emptyset$ eine Menge, und es seien A_α Mengen für alle $\alpha \in I$ (d.h. jedem $\alpha \in I$ ist eine Menge A_α „zugeordnet“). Dann heißen

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

Vereinigung der Mengen A_α $(\alpha \in I)$ und

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

Durchschnitt der Mengen A_α $(\alpha \in I)$. Unmittelbar aus den Definitionen ergeben sich folgende praktische Äquivalenzen: Ist B eine weitere Menge, so ist

$$A_\alpha \subset B \ (\alpha \in I) \text{ genau dann, wenn } \bigcup_{\alpha \in I} A_\alpha \subset B$$

und

$$B \subset A_\alpha \ (\alpha \in I) \text{ genau dann, wenn } B \subset \bigcap_{\alpha \in I} A_\alpha \, .$$

Ist I in aufzählender Form gegeben, so setzen wir $\bigcup$ bzw. $\bigcap$ auch zwischen die einzelnen Mengen, also etwa im Falle $I = \{1, 2, 3\}$

$$A_1 \cup A_2 \cup A_3 := \bigcup_{\alpha \in I} A_\alpha, \qquad A_1 \cap A_2 \cap A_3 := \bigcap_{\alpha \in I} A_\alpha.$$

Insbesondere sind damit für eine Menge von Mengen (einem so genannten Mengensystem) $\mathcal{F}$ auch

$$\bigcup_{M \in \mathcal{F}} M \qquad \text{und} \qquad \bigcap_{M \in \mathcal{F}} M$$

definiert (hier ist speziell $I = \mathcal{F}$ und $A_M = M$). Auch in diesem Fall schreibt man alternativ $\bigcup$ und $\bigcap$ auch zwischen die einzelnen Mengen, falls $\mathcal{F}$ aufzählend gegeben ist, also etwa im Falle $\mathcal{F} = \{A, B, C\}$

$$A \cup B \cup C \quad \text{und} \quad A \cap B \cap C \, .$$

Beispiel 1.3 1. Ist $I = \mathbb{N}$ und ist $A_m := (0, m]$ für $m \in \mathbb{N}$, so ist

$$\bigcup_{m\in\mathbb{N}} A_m = \mathbb{R}_+, \quad \bigcap_{m\in\mathbb{N}} A_m = (0,1].$$

Ist $A_m := (m, \infty)$ für $m \in \mathbb{N}$, so ist

$$\bigcap_{m\in\mathbb{N}} A_m = \emptyset\ .$$

2. Sind $A = \{2k : k \in \mathbb{Z}\}$ und $B = \{3k : k \in \mathbb{Z}\}$, so gilt

$$A \cap B = \{6k : k \in \mathbb{Z}\}.$$

Satz 1.4 *Ist A eine Menge und sind B_α Mengen $(\alpha \in I)$, so gilt*

1. $\qquad A \cap (\bigcup_{\alpha\in I} B_\alpha) = \bigcup_{\alpha\in I} (A \cap B_\alpha),$

2. $\qquad A \cup (\bigcap_{\alpha\in I} B_\alpha) = \bigcap_{\alpha\in I} (A \cup B_\alpha).$

Beweis. 1. „$\subset$“ Es sei $x \in A \cap (\bigcup_{\alpha\in I} B_\alpha)$. Dann ist $x \in A$ und $x \in \bigcup_{\alpha\in I} B_\alpha$, also $x \in A$ und $x \in B_\beta$ für ein $\beta \in I$. Damit ist $x \in A \cap B_\beta$, also auch $x \in \bigcup_{\alpha\in I} (A \cap B_\alpha)$.
„$\supset$“ Es sei $x \in \bigcup_{\alpha\in I} (A\cap B_\alpha)$. Dann existiert ein $\beta \in I$ mit $x \in A\cap B_\beta$. Damit ist $x \in A$ und $x \in B_\beta$, also auch $x \in A$ und $x \in \bigcup_{\alpha\in I} B_\alpha$, d. h. $x \in A \cap (\bigcup_{\alpha\in I} B_\alpha)$.
2. [Ü] $\square$

Satz 1.5 *(De Morgansche Regeln)*
Es sei B eine Menge, und es seien $A_\alpha \subset B$ für alle $\alpha \in I$. Dann gilt

1. $C_B(\bigcup_{\alpha\in I} A_\alpha) = \bigcap_{\alpha\in I} C_B(A_\alpha),$

2. $C_B(\bigcap_{\alpha\in I} A_\alpha) = \bigcup_{\alpha\in I} C_B(A_\alpha).$

Beweis. 1. „$\subset$“ Es sei $x \in C_B(\bigcup_{\alpha\in I} A_\alpha)$. Dann ist $x \in B$ und $x \notin \bigcup_{\alpha\in I} A_\alpha$, also $x \in B$ und $x \notin A_\alpha$ für alle $\alpha \in I$. Damit ist $x \in B \setminus A_\alpha$ für alle $\alpha \in I$, also $x \in \bigcap_{\alpha\in I} C_B(A_\alpha)$.

„$\supset$“ Es sei $x \in \bigcap_{\alpha\in I} C_B(A_\alpha)$. Dann ist $x \in C_B(A_\alpha)$ für alle $\alpha \in I$, also $x \in B$ und $x \notin A_\alpha$ für alle $\alpha \in I$. Damit ist $x \in B$ und $x \notin \bigcup_{\alpha\in I} A_\alpha$, d. h. $x \in C_B(\bigcup_{\alpha\in I} A_\alpha)$.

2. [Ü] $\square$

Definition 1.6 Ähnlich wie bei der Einführung von Mengen wollen wir auf eine eher intuitive Definition des zweiten grundlegenden Begriffes der Mathematik zurückgreifen: Es seinen X und Y (nichtleere) Mengen. Eine *Funktion* (oder *Abbildung*)

$$f : X \to Y$$

ist eine Vorschrift, die jedem $x \in X$ **genau ein** Element $f_x = f(x) \in Y$ zuordnet. Wir schreiben auch $X \ni x \mapsto f(x) \in Y$ oder auch $(f_x)_{x\in X}$.
Dabei heißen X der *Definitionsbereich* und Y der *Zielbereich* von f. Wir setzen

$$Y^X := \{f : X \to Y\} .$$

Sind $f, g \in Y^X$, so heißen f und g *gleich*, falls $f(x) = g(x)$ für alle $x \in X$ gilt.
Ist $X_0 \subset X$, so heißt $f_{|X_0} : X_0 \to Y$, definiert durch $f_{|X_0}(x) := f(x)$ für alle $x \in X_0$, die *Einschränkung* von f auf X_0.
Hat man speziell $X = \{1, \ldots, n\}$ für ein $n \in \mathbb{N}$, so schreibt man meist $(y_1, \ldots, y_n)$ statt $(y_j)_{j\in\{1,\ldots,n\}}$ und kurz

$$Y^n := Y \times \cdots \times Y := Y^{\{1,\ldots,n\}} = \{y = (y_1, \ldots, y_n)\} .$$

Schließlich definieren wir für beliebgie Mengen $A_1, \ldots, A_n$ (mit $Y := \bigcup_{j\in\{1,\ldots,n\}} A_j$)

$$A_1 \times \cdots \times A_n := \{(y_1, \ldots, y_n) \in Y^n : y_j \in A_j \, (j = 1, \ldots, n)\} .$$

Beispiel 1.7 Es seien $X = Y = \mathbb{N}$, und es sei $f : \mathbb{N} \to \mathbb{N}$ definiert durch

$$f(x) := \begin{cases} x \ , & \text{falls } x \text{ gerade} \\ 2x, & \text{falls } x \text{ ungerade} \end{cases} .$$

Ist $X_0 := \{x \in \mathbb{N} : x \text{ ungerade}\}$, so ist

$$f_{|X_0}(x) = 2x \qquad (x \in X_0).$$

Definition 1.8 Sind X, Y Mengen und ist $f : X \to Y$, so heißt für $B \subset Y$

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

Urbildmenge von B unter f und für $A \subset X$

$$f(A) := \{f(x) : x \in A\} = \{y \in Y : y = f(x) \text{ für ein } x \in A\}$$

Bildmenge von A unter f. Speziell heißt

$$W(f) := f(X)$$

Wertebereich von f.

Beispiel 1.9 In der Situation von B. 1.7 ist etwa

$$f^{-1}\big(\{2,4,6\}\big) = f^{-1}\big(\{1,2,3,4,5,6\}\big) = \{1,2,3,4,6\}$$

und

$$f\big(\{1,2,3\}\big) = \{2,6\}.$$

Außerdem ist

$$W(f) = \{y \in \mathbb{N} : y \text{ gerade}\}.$$

Satz 1.10 *Es seien X, Y Mengen und $f : X \to Y$.*

1. *Sind $B_\alpha \subset Y$ für $\alpha \in I$, so gilt*

   (i) $f^{-1}\big(\bigcup_{\alpha \in I} B_\alpha\big) = \bigcup_{\alpha \in I} f^{-1}(B_\alpha)$,

   (ii) $f^{-1}\big(\bigcap_{\alpha \in I} B_\alpha\big) = \bigcap_{\alpha \in I} f^{-1}(B_\alpha)$.

2. *Sind $A_\alpha \subset X$ für $\alpha \in I$, so gilt*

   (i) $f\big(\bigcup_{\alpha \in I} A_\alpha\big) = \bigcup_{\alpha \in I} f(A_\alpha)$,

   (ii) $f\big(\bigcap_{\alpha \in I} A_\alpha\big) \subset \bigcap_{\alpha \in I} f(A_\alpha)$.

Beweis. 1. (i) „$\subset$“: Es sei $x \in f^{-1}\big(\bigcup_{\alpha \in I} B_\alpha\big)$. Dann ist $f(x) \in \bigcup_{\alpha \in I} B_\alpha$, d.h. es existiert ein $\beta \in I$ mit $f(x) \in B_\beta$. Also ist $x \in f^{-1}(B_\beta)$ und damit auch $x \in \bigcup_{\alpha \in I} f^{-1}(B_\alpha)$.

„$\supset$“: Ist $\beta \in I$, so ist $B_\beta \subset \bigcup\limits_{\alpha\in I} B_\alpha$, also auch $f^{-1}(B_\beta) \subset f^{-1}\big(\bigcup\limits_{\alpha\in I} B_\alpha\big)$. Da $\beta \in I$ beliebig war, gilt $\bigcup\limits_{\alpha\in I} f^{-1}(B_\alpha) \subset f^{-1}\big(\bigcup\limits_{\alpha\in I} B_\alpha\big)$.
(ii) „$\subset$“: Für alle $\beta \in I$ ist $\bigcap\limits_{\alpha\in I} B_\alpha \subset B_\beta$, also auch $f^{-1}\big(\bigcap\limits_{\alpha\in I} B_\alpha\big) \subset f^{-1}(B_\beta)$. Da β beliebig war, ist $f^{-1}\big(\bigcap\limits_{\alpha\in I} B_\alpha\big) \subset \bigcap\limits_{\beta\in I} f^{-1}(B_\beta)$.
„$\supset$“: Es sei $x \in \bigcap\limits_{\alpha\in I} f^{-1}(B_\alpha)$. Dann ist $x \in f^{-1}(B_\alpha)\,(\alpha \in I)$, also $f(x) \in B_\alpha\,(\alpha \in I)$ und damit auch $f(x) \in \bigcap\limits_{\alpha\in I} B_\alpha$. Folglich ist $x \in f^{-1}\big(\bigcap\limits_{\alpha\in I} B_\alpha\big)$.
2. (i) „$\subset$“: Es sei $y \in f\big(\bigcup\limits_{\alpha\in I} A_\alpha\big)$. Dann existiert ein $x \in \bigcup\limits_{\alpha\in I} A_\alpha$ mit $f(x) = y$. Ist $\beta \in I$ mit $x \in A_\beta$, so ist also $y = f(x) \in f(A_\beta)$. Damit ist $y \in \bigcup\limits_{\alpha\in I} f(A_\alpha)$.
„$\supset$“: Ist $\beta \in I$, so ist $A_\beta \subset \bigcup\limits_{\alpha\in I} A_\alpha$, also auch $f(A_\beta) \subset f\big(\bigcup\limits_{\alpha\in I} A_\alpha\big)$. Da $\beta \in I$ beliebig war, gilt „$\supset$“.
(ii) [Ü] $\square$

Definition 1.11 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 gilt: sind $x_1, x_2 \in X$ mit $x_1 \neq x_2$, so ist $f(x_1) \neq f(x_2)$,
3. *bijektiv*, falls f injektiv und surjektiv ist.

Beispiel 1.12 Es sei f wie im B 1.7. Dann ist f weder surjektiv noch injektiv (es gilt etwa $1 \notin W(f)$ und $f(2) = f(1)$), dagegen ist $f_{|X_0}$ injektiv.

Definition 1.13 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

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

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

Beispiel 1.14 Sind $X = Y = Z = \mathbb{N}$ und $f, g : \mathbb{N} \to \mathbb{N}$ definiert durch

$$f(x) = x^2, \quad g(x) = x + 1 \quad (x \in \mathbb{N}) ,$$

so ist $g \circ f : \mathbb{N} \to \mathbb{N}$ gegeben durch

$$(g \circ f)(x) = x^2 + 1 \quad (x \in \mathbb{N}) .$$

Man beachte: Hier ist auch $f \circ g : \mathbb{N} \to \mathbb{N}$ definiert und es gilt

$$(f \circ g)(x) = (x + 1)^2 \quad (x \in \mathbb{N})$$

Dabei ist $g \circ f \neq f \circ g$.

Satz 1.15 *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

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

□

Bemerkung und Definition 1.16 Es seien X, Y Mengen und es sei $f : X \to Y$ bijektiv. Dann existiert zu jedem $y \in Y$ **genau** ein $x \in X$ mit $f(x) = y$. 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.

2 Körper

Definition 2.1 Es seien $G \neq \emptyset$ eine Menge, und es sei $* : G \times G \to G$ eine Abbildung. Dann heißt $(G, *)$ *Gruppe*, falls gilt

(G.1) (Assoziativgesetz)
Für alle $x, y, z \in G$ ist
$$(x * y) * z = x * (y * z).$$

(G.2) (Existenz eines neutralen Elementes)
Es existiert ein $e \in G$ mit $x * e = e * x = x$ für alle $x \in G$.

(G.3) (Existenz linksinverser Elemente)
Für alle $x \in G$ existiert ein $y \in G$ mit $y * x = e$.

Gilt zudem
(G.4) (Kommutativgesetz)
Für alle $x, y \in G$ ist
$$x * y = y * x,$$
so heißt $(G, *)$ eine *abelsche* (oder *kommutative*) Gruppe.

Bemerkung 2.2 1. Für $(G.4)$ ist es wichtig, dass das neutrale Element eindeutig bestimmt ist, d. h. es existiert nur ein $e = e_G \in G$ mit $x * e = e * x = x$ $(x \in G)$.
(Denn: Sind $e, e' \in G$ neutrale Elemente, so gilt
$$e' = e * e' = e.)$$

2. Ist y linksinvers zu x, so ist y auch rechtsinvers, d.h. es ist auch $x * y = e$.
(Denn: Ist z so, dass $z * y = e$, so folgt
$$x*y = e*(x*y) = (z*y)*(x*y) = z*(y*(x*y)) = z*((y*x)*y) = z*(e*y) = z*y = e).$$

Außerdem existiert zu jedem $x \in G$ nur ein $y \in G$ mit $y * x = e$.
(Denn: Sind y und $y' \in G$ mit $y * x = y' * x = e$, so folgt
$$y' = y' * e = y' * (x * y) = (y' * x) * y = e * y = y.)$$

Wir schreiben x^{-1} für das inverse Element von x.
Es gilt damit für $x, y \in G$ ([Ü])
$$(x * y)^{-1} = y^{-1} * x^{-1} \quad \text{und} \quad (x^{-1})^{-1} = x.$$

Beispiel 2.3 1. $(\mathbb{Z}, +)$ ist eine abelsche Gruppe.
2. $(\mathbb{N}_0, +)$ ist keine Gruppe (es existiert kein $y \in \mathbb{N}_0$ mit $1 + y = 0$). Genauso ist $(\mathbb{Z} \setminus \{0\}, \cdot)$ keine Gruppe (es existiert kein $y \in \mathbb{Z} \setminus \{0\}$ mit $2 \cdot y = 1$).
3. Ist X eine (nichtleere) Menge, so setzen wir

$$S(X) := \{f : X \to X : f \text{ bijektiv}\} .$$

Dann ist $(S(X), \circ)$ ein Gruppe. Hat X mindestens 3 Elemente, so ist $(S(X), \circ)$ nicht abelsch ([Ü]).

Satz 2.4 *Es sei $(G, *)$ eine Gruppe, und es seien $a, b \in G$. Dann sind die Gleichungen $a * x = b$ und $y * a = b$ eindeutig lösbar mit den Lösungen $x = a^{-1} * b$ bzw. $y = b * a^{-1}$.*

Beweis. Wir zeigen die Behauptung nur für die erste Gleichung.
Zunächst gilt

$$a * (a^{-1} * b) = (a * a^{-1}) * b = e * b = b,$$

also ist $a^{-1} * b$ eine Lösung.
Ist andererseits x irgendeine Lösung, so folgt aus $a * x = b$

$$a^{-1} * b = a^{-1} * (a * x) = (a^{-1} * a) * x = e * x = x .$$

□

Definition 2.5 Es sei K eine Menge mit mindestens zwei Elementen. Weiter seien $+ : K \times K \to K$ und $\cdot : K \times K \to K$ Abbildungen. Dann heißt $(K, +, \cdot)$ *Körper*, falls gilt

(K.1) $(K, +)$ ist eine abelsche Gruppe (mit neutralem Element genannt $0 = 0_K$).

(K.2) $(K \setminus \{0_K\}, \cdot)$ ist eine abelsche Gruppe (mit neutralem Element genannt $1 = 1_K$).

(K.3) (Distributivgesetze)
Für alle $x, y, z \in K$ gilt

$$(x + y) \cdot z = (x \cdot z) + (y \cdot z) \quad \text{und} \quad x \cdot (y + z) = (x \cdot y) + (x \cdot z).$$

Im Folgenden schreiben wir auch kurz xy statt $x \cdot y$ und $x + yz$ statt $x + (yz)$ („Punktrechnung vor Strichrechnung"). Weiter schreiben wir für $x \in X$ wie üblich $-x$ für das inverse Element bezüglich „+". Schließlich schreiben wir noch $x - y$ statt $x + (-y)$ und x/y statt xy^{-1}.

Satz 2.6 *Es sei* $(K, +, \cdot)$ *ein Körper. Ferner seien* $x, y \in K$*. Dann gilt*

1. $0_K \cdot x = x \cdot 0_K = 0_K$,
2. $-(xy) = (-x)y = x(-y)$,
3. $xy = 0_K$ *genau dann, wenn* $x = 0_K$ *oder* $y = 0_K$ *(Nullteilerfreiheit).*

Beweis.
1. Aus $0_K \cdot x = (0_K + 0_K)x \overset{(K.3)}{=} 0_Kx + 0_Kx$ folgt

$$0_Kx = 0_Kx - 0_Kx = 0_K$$

nach S. 2.4 (mit $(G, *) = (K, +)$). Entsprechend sieht man $x \cdot 0_K = 0_K$.
2. und 3. [Ü] □

Beispiel 2.7 1. $(\mathbb{Q}, +, \cdot)$ ist ein Körper.
2. (Binärkörper) Es sei $K = \{n, e\}$ mit den Rechenoperationen

$+$	n	e
n	n	e
e	e	n

$\cdot$	n	e
n	n	n
e	n	e

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

3. $(\mathbb{N}, +, \cdot)$ und $(\mathbb{Z}, +, \cdot)$ bilden keine Körper (vgl. B. 2.3).

Definition 2.8 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 $x_1 = \ldots = x_n =: x$, 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 für $n \in \mathbb{N}$

$$(-n)x := n(-x)\,, \qquad x^{-n} := (x^{-1})^n$$

(Letzteres natürlich nur im Falle $x \neq 0$) und

$$0 \cdot x := 0_K\,,\ x^0 := 1_K \quad \text{wobei 0 die Null in } \mathbb{Z} \text{ bezeichnet}\,.$$

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 beliebiges $k \in \mathbb{N}$ 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.
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 für $n = N$ und den Induktionsschritt von k auf $k+1$ für beliebiges $k \geq N$.

Ein typischer Induktionsbeweis ist der Beweis zu

Satz 2.9 *Für alle $n \in \mathbb{N}$ gilt*

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

Beweis.

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

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

   b) Wir zeigen: aus a) folgt $\sum_{\nu=1}^{k+1} \nu = \frac{(k+1)(k+2)}{2}$ (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.

Bemerkung und Definition 2.10 Es seien $(K, +, \cdot)$ ein Körper und $\varphi : \{1, \ldots, n\} \to \{1, \ldots, n\}$ bijektiv. Dann gilt für $x_1, \ldots, x_n \in K$

$$\sum_{\nu=1}^{n} x_{\varphi(\nu)} = \sum_{\nu=1}^{n} x_\nu \qquad \text{und} \qquad \prod_{\nu=1}^{n} x_{\varphi(\nu)} = \prod_{\nu=1}^{n} x_\nu .$$

Damit wird folgende Schreibweise sinnvoll: Ist $n \in \mathbb{N}$ und ist I eine beliebige n-elementige Menge, so setzen wir für $x_j \in K$ $(j \in I)$

$$\sum_{j \in I} x_j := \sum_{k=1}^{n} x_{j_k} \quad \text{und} \quad \prod_{j \in I} x_j := \prod_{k=1}^{n} x_{j_k} ,$$

wobei $\{j_1, \ldots, j_n\}$ eine beliebige Aufzählung von I ist. Sind weiter $y_j \in K$ $(j \in I)$ und $x \in K$, so gilt auch etwa

$$\sum_{j \in I} x x_j = x \sum_{j \in I} x_j ,$$

und
$$\sum_{j\in I}(x_j+y_j)=\sum_{j\in I}x_j+\sum_{j\in I}y_j, \qquad \prod_{j\in I}(x_jy_j)=\prod_{j\in I}x_j\prod_{j\in I}y_j.$$
Die Beweise ergeben sich (nicht ganz leicht) per Induktion.
Weiter kann man hiermit (leicht) zeigen, dass für $m, m_1, m_2 \in \mathbb{Z}$ und $x_1, x_2, x \in K$ folgende Vervielfachungs- und Potenzgesetze gelten:

$$\begin{aligned} m_1x + m_2x &= (m_1+m_2)x\ , \\ mx_1 + mx_2 &= m(x_1+x_2)\ , \\ (m_1m_2)x &= m_1(m_2x)\ . \end{aligned}$$

und (für $x_1, x_2, x \neq 0$ falls negative Potenzen auftreten)

$$\begin{aligned} x^{m_1}x^{m_2} &= x^{m_1+m_2}\ , \\ x_1^m x_2^m &= (x_1x_2)^m\ , \\ (x^{m_1})^{m_2} &= x^{m_1m_2}\ . \end{aligned}$$

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

Definition 2.11 1. Es sei $X \neq \emptyset$ eine Menge. Eine Teilmenge R von $X \times X$ (genauer das Paar (R, X)) nennt man auch eine *Relation auf* (oder auch *in*) X. Ist $(x, y) \in R$, so schreibt man oft xRy.
Eine Relation $<$ auf X heißt eine *Ordnung* (auf X), falls gilt

(O.1) Für alle $x, y \in X$ 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).

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 $y \geq x$.

2. Es sei $K = (K, +, \cdot)$ ein Körper. Ist $<$ eine Ordnung auf K, so heißt $K = (K, +, \cdot, <)$ *geordnet*, wenn $<$ (neben den Bedingungen aus 1.) noch folgende Verträglichkeiten mit der Addition und Multiplikation erfüllt

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

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

Wir nennen $x \in K$ *positiv*, falls $x > 0_K$ gilt und *negativ*, falls $x < 0_K$ gilt.

Satz 2.12 *Es seien $K = (K, +, \cdot, <)$ ein geordneter Körper und $x, y \in K$. Dann gilt*

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

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. Sind $x, y > 0$ so folgt mit (O.4) sofort $0 = 0y < xy$.
Es seien $x, y < 0$. 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. Ergibt sich unmittelbar aus 2. und (O.1).

4. Wir zeigen zunächst: $x^{-1} > 0$. (Denn: 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.) Genauso ist $y^{-1} > 0$. Damit ergibt sich aus $x < y$ mit (O.4) $xy^{-1} < yy^{-1} = 1$ und wieder mit (O.4) $x^{-1}xy^{-1} < x^{-1}1 = x^{-1}$, also $y^{-1} < x^{-1}$.

□

Bemerkung 2.13 Es sei K ein geordneter Körper. Per Induktion sieht man leicht:

1. Ist $n \in \mathbb{N}$ und ist $x < y$, so gilt $nx < ny$ und im Falle $x > 0_K$ auch $0_K < x^n < y^n$.

2. Ist $x > 0_K$, so ist auch $nx > mx > 0_K$ für alle $n, m \in \mathbb{N}$ mit $n > m$.

Insbesondere folgt aus 2., dass K unendlich ist. Genauer ergibt sich auch:
Sind $x, y \in K$ mit $x < y$, so liegen zwischen x und y unendlich viele Elemente aus K (Denn: Für alle $n\mathbb{N}$ ist $(n+1)1_K > n1_K > 0_K$, also $(n1_K)^{-1} > ((n+1)1_K)^{-1} > 0_K$ und folglich

$$y = x + (y-x)(1_K)^{-1} > x + (y-x)(2 \cdot 1_K)^{-1} > x + (y-x)(3 \cdot 1_K)^{-1} \cdots > x\ .)$$

Beispiel 2.14 1. $(\mathbb{Q}, +, \cdot, <)$ ist ein geordneter Körper.
2. Im Binärkörper $(K, +, \cdot)$ aus B. 2.7.2 existiert keine Ordnungsrelation, da jeder geordnete Körper unendlich viele Elemente enthält (vgl. B 2.13).

3 Geometrische Summenformel und binomische Formel

Eine wichtige Formel für Summen von Potenzen in Körpern ist die

Satz 3.1 *(geometrische Summenformel)*
Es sei $(K, +, \cdot)$ ein Körper. Dann gilt für alle $x \in K$ und alle $n \in \mathbb{N}$

$$(x - 1_K) \cdot \sum_{\nu=0}^{n-1} x^\nu = x^n - 1_K \tag{3.1}$$

und für $x \neq 1_K$ damit

$$\sum_{\nu=0}^{n-1} x^\nu = \frac{x^n - 1_K}{x - 1_K} \tag{3.2}$$

Beweis. Es gilt

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

□

Bemerkung 3.2 Allgemeiner kann man zeigen ([Ü]): Ist K ein Körper und sind $a, b \in K$, so gilt für alle $n \in \mathbb{N}$

$$a^n - b^n = (a-b)\sum_{\nu=0}^{n-1} a^\nu b^{n-1-\nu} \,.$$

Neben der geometrischen Summenformel gibt es eine weitere Formel in Körpern, die binomische Formel. Es handelt sich dabei um eine Summenformel für die Ausdrücke $(a+b)^n$, wobei $a, b \in K$ und $n \in \mathbb{N}$ ist. Um die allgemeine Formel angeben zu können, brauchen wir

Definition 3.3 1. Wir definieren $n!$ (“*n-Fakultät*”) für $n \in \mathbb{N}_0$ durch

$$n! := \prod_{\nu=1}^{n} \nu \ ,$$

wobei $\prod_{\nu=m}^{n} x_\nu := 1$ im Falle $n < m$ gesetzt ist (also $0! = 1$).

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

$$\binom{n}{\nu} := \frac{1}{\nu!} \prod_{k=n-\nu+1}^{n} k \qquad \left(= \frac{1}{\nu!} \prod_{k=1}^{\nu} (n+1-k) \right)$$

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

Es gilt also etwa

$$6! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 = 720 \ , \qquad 10! = 3.628.800 \ ,$$

$$\binom{7}{5} = \frac{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3}{5!} = 21$$

Wir stellen einige Eigenschaften der Binomialkoeffizienten zusammen.

Satz 3.4 *Es seien $n, \nu \in \mathbb{N}_0$. Dann gilt*

1. $\binom{n}{\nu} = \frac{n!}{\nu!(n-\nu)!} = \binom{n}{n-\nu}$ *falls* $\nu \leq n$.
2. $\binom{n}{\nu} = 0$ *falls* $\nu > n$.

Beweis.

1. Es gilt für $\nu \leq n$

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

Damit ist auch

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

2. Für $\nu > n$ ist $n-\nu+1 \leq 0$ und damit $\prod\limits_{k=n-\nu+1}^{n} k = 0$, also auch $\binom{n}{\nu} = 0$.

□

Besonders wichtig ist folgende Rekursionsformel:

Satz 3.5 *Für $n, \nu \in \mathbb{N}$ gilt*

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

Beweis. Nach S. 3.4.1 gilt für $\nu \in \{1, \ldots, n\}$

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

Für $\nu = n+1$ ist nach S. 3.4.2

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

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 $\binom{n}{0}, \ldots, \binom{n}{n}$ stehen, so entsteht das sog. *Pascal'sche Dreieck*:

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

Die ersten Zeilen berechnen sich etwa unter Ausnutzung von S. 3.5 zu

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

Damit gilt

Satz 3.6 *Es sei K ein Körper. Dann gilt für alle $n \in \mathbb{N}_0$ und $x \in K$*

$$(1_K + x)^n = \sum_{\nu=0}^{n} \binom{n}{\nu} x^\nu$$

Beweis.
1. Für $n = 0$ gilt $(1+x)^0 = 1 = \sum\limits_{\nu=0}^{0} \binom{0}{\nu} x^\nu$.
2. Für ein $k \in \mathbb{N}_0$ gelte

$$(1+x)^k = \sum_{\nu=0}^{k} \binom{k}{\nu} x^\nu \ .$$

Dann gilt mit S. 3.5

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

□

Bemerkung 3.7 Aus S. 3.6 ergibt sich unmittelbar die allgemeine *binomische Formel*: Sind $a, b \in K$, so ist

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

(Denn: Im Falle $b \neq 0$ ist

$$(a+b)^n = b^n \big(1 + (a/b)\big)^n = b^n \sum_{\nu=0}^{n} \binom{n}{\nu} (a/b)^\nu = \sum_{\nu=0}^{n} \binom{n}{\nu} a^\nu b^{n-\nu}$$

und im Falle $b = 0$ steht auf beiden Seiten a^n.)

Beispiel 3.8 Es gilt etwa

$$\begin{aligned} (a+b)^6 &= \sum_{\nu=0}^{6} \binom{6}{\nu} a^\nu b^{6-\nu} \\ &= 1 \cdot b^6 + 6 \cdot ab^5 + 15a^2b^4 + 20a^3b^3 + 15a^4b^2 + 6a^5b + 1 \cdot a^6 \, . \end{aligned}$$

Bemerkung 3.9 Als Spezialfälle aus S. 3.6 ergeben sich interessante Beziehungen für das Pascal'sche Dreieck:
Für ($K = \mathbb{Q}$ und) $a = 1, b = 1$ ergibt sich

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

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

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

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

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

Bemerkung 3.10 (Bernoullische Ungleichung)
Ist $(K, +, \cdot, <)$ ein geordneter Körper, so gilt nach S. 3.6 für alle $n \in \mathbb{N}, x \geq 0_K$

$$(1+x)^n = \sum_{\nu=0}^{n} \binom{n}{\nu} x^\nu \geq \sum_{\nu=0}^{1} \binom{n}{\nu} x^\nu = 1 + nx.$$

Tatsächlich gilt diese Abschätzung auch für $x \geq -1_K$.
(Denn: Für $k = 1$ ist $(1+x)^1 = 1 + 1 \cdot x$.
Gilt die Abschätzung für ein $k \in \mathbb{N}$ und ist $x \geq -1_K$, so folgt (da $1 + x \geq 0_K$)

$$(1+x)^{k+1} = (1+x)^k(1+x) \geq (1+kx)(1+x) = 1 + (k+1)x + kx^2 \text{ .)}$$

Bemerkung 3.11 Zum Abschluss beschäftigen wir uns kurz mit der Bedeutung der Fakultäten und Binomialkoeffizienten im Bereich der „Kombinatorik“.
Für eine endliche Menge M setzen wir

$$|M| := \text{Anzahl der Elemente von } M.$$

Dann gilt: Sind M, N endliche Mengen, so existiert eine bijektive Abbildung $f : M \to N$ genau dann, wenn $|M| = |N|$ ist (d. h., wenn M und N gleich viele Elemente haben).

1. Es sei $n \in \mathbb{N}$. Dann gilt: Sind I, J n-elementige Mengen und ist

$$S(I, J) := \{\varphi : I \to J : \varphi \text{ bijektiv}\}$$

so ist

$$\big|S(I, J)\big| = n! \text{ .}$$

(Denn: Wir führen den Beweis per Induktion nach n.
1. Induktionsanfang: Für $n = 1$ ist die Behauptung klar.
2. Induktionsschritt: Die Behauptung gelte für ein $k \in \mathbb{N}$.
Es seien I, J $(k+1)$-elementige Mengen. O.E. können wir $J = \{1, \ldots, k+1\}$ annehmen.
(Denn: Wie oben bemerkt, existiert eine bijektive Abbildung $f : \{1, \ldots, k+1\} \to J$.
Dann ist die Abbildung

$$S(I, \{1, \ldots, k+1\}) \ni \varphi \mapsto f \circ \varphi \in S(I, J)$$

bijektiv. Also ist $|S(I, \{1, \ldots, k+1\})| = |S(I, J)|$.)
Für $i \in I$ definieren wir

$$T_i := \big\{\varphi \in S(I, \{1, \ldots, k+1\}) : \varphi(i) = k+1\big\} \text{ .}$$

Dann ist
$$\bigcup_{i\in I} T_i = S(I,\{1,\ldots,k+1\}) \quad \text{und} \quad T_i \cap T_{i'} = \emptyset \quad (i \neq i').$$
Also ist $\big|S(I,\{1,\ldots,k+1\})\big| = \sum_{i\in I} |T_i|$.
Definiert man für $\varphi \in T_i$ die Funktion $\psi : I \setminus \{i\} \to \{1,\ldots,k\}$ durch
$$\psi(j) := \varphi(j) \quad (j \in I,\, j \neq i)$$
(also Definitionsbereich und Zielbereich jeweils um ein Element „verkleinert“), so ist
$$T_i \ni \varphi \mapsto \psi \in S(I \setminus \{i\}, \{1,\ldots,k\})$$
eine bijektive Abbildung.
Nach Induktionsannahme gilt $\big|S\big(I\setminus\{i\},\{1,\ldots,k\}\big)\big| = k!$ und damit auch $|T_i| = k!$.
Also ist $\big|S(I,\{1,\ldots,k+1\})\big| = \sum_{i\in I} k! = (k+1)k! = (k+1)!$.)

2. Ist M eine n-elementige Menge und ist $\mathcal{M}_\nu \subset \mathrm{Pot}(M)$ die Menge der ν-elementigen Teilmengen von M (wobei $\nu \in \{0,\ldots,n\}$), so ist ([Ü])
$$|\mathcal{M}_\nu| = \binom{n}{\nu}.$$
Nach B. 3.9 ist damit auch
$$|\mathrm{Pot}(M)| = 2^n\,.$$

4 Reelle und komplexe Zahlen

Ist $(K, +, \cdot, <)$ ein geordneter Körper (etwa $K = \mathbb{Q}$), so sind Gleichungen der Form

$$nx = c$$

für alle $n \in \mathbb{N}, c \in K$ lösbar ($x = c/n := c/(n1_K)$ ist die eindeutige Lösung). Im Allgemeinen gilt dies nicht für Gleichungen der Form

$$x^n = c,$$

wobei $c \in K, n \in \mathbb{N}, n > 1$. Sind $c < 0$ und n gerade, so ist dies nach S. 2.12.3 ohnehin ausgeschlossen. Aber auch im Falle $c > 0$ existiert im Allgemeinen keine Lösung (wie schon seit der Antike bekannt ist).

Satz 4.1 *Für alle $x \in \mathbb{Q}$ ist $x^2 \neq 2$.*

Beweis. 1. Allgemein gilt: Ist $m \in \mathbb{Z}$ ungerade, so ist auch m^2 ungerade (denn: ist $m = 2\ell + 1$ mit $\ell \in \mathbb{Z}$, so ist $m^2 = 4(\ell^2 + \ell) + 1$, also ebenfalls ungerade).
2. Angenommen, es existiert $p/q \in \mathbb{Q}$ mit $(p/q)^2 = 2$. Wir können ohne Einschränkung annehmen, dass $p \in \mathbb{Z}$, $q \in \mathbb{N}$ teilerfremd und damit insbesondere nicht beide gerade sind.
Dann 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 ergibt sich ein Widerspruch. Damit ist die Annahme falsch d. h. es existiert kein $x \in \mathbb{Q}$ mit $x^2 = 2$. $\square$

Unsere Ziele im Weiteren sind:
1. Erweitern von $(\mathbb{Q}, +, \cdot, <)$ zu einem geordneten Körper $(\mathbb{R}, +, \cdot, <)$ so, dass $x^n = c$ für alle $n \in \mathbb{N}$ und $c \geq 0$ lösbar ist.
2. Erweitern von $(\mathbb{R}, +, \cdot)$ zu einem Körper $(\mathbb{C}, +, \cdot)$ so, dass $x^n = c$ für alle $n \in \mathbb{N}, c \in \mathbb{C}$ lösbar ist.

Definition 4.2 Es seien X eine (nichtleere) Menge und $<$ eine Ordnung auf X (wir sagen dann, $(X, <)$ ist geordnet). Weiter sei $M \subset X$.

1. M heißt *nach oben beschränkt*, wenn ein $\overline{s} \in X$ 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 $\underline{s} \in X$ 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 4.3 Es sei $X = \mathbb{Q}$ und
$$M := \{x \in \mathbb{Q} : x \geq 0,\ x^2 \leq 2\} \ \big(= \{x \in \mathbb{Q} : x \geq 0,\ x^2 < 2\}\big) \ .$$
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 $\overline{\overline{s}} \in X$ 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 4.4 Es sei $(X, <)$ geordnet, und es sei $M \subset X$.

1. Ein obere Schranke $\overline{\xi} \in X$ von M heißt *kleinste obere Schranke* (oder *Supremum*) von M, falls für jede obere Schranke $\overline{s}$ von M gilt
$$\overline{s} \geq \overline{\xi} \ .$$

2. Eine untere Schranke $\underline{\xi} \in X$ von M heißt *größte untere Schranke* (oder *Infimum*) von M, falls für jede untere Schranke $\underline{s}$ von M gilt
$$\underline{s} \leq \underline{\xi} \ .$$

Bemerkung und Definition 4.5 Aus der Definition ergibt sich sofort, dass für jedes M höchstens ein Supremum $\overline{\xi}$ und ein Infimum $\underline{\xi}$ existieren. Wir schreiben (im Falle der Existenz)
$$\overline{\xi} = \sup\ M \ .$$

Zudem nennen wir $\overline{\xi}$ *Maximum* von M (und schreiben $\overline{\xi} = \max M$), falls zusätzlich $\overline{\xi} \in M$ gilt.
Weiter schreiben wir (im Falle der Existenz)

$$\underline{\xi} := \inf M.$$

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

Beispiel 4.6 Es sei $X = \mathbb{Q}$.

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

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

   (Denn: Offensichtlich ist 1 obere Schranke. Ist andererseits $s < 1$, so ist s keine obere Schranke, da etwa $x := \max\{1/2, (s+1)/2\} \in M$ und $s < x$. Also ist $1 = \sup M$. Entsprechend zeigt man, dass $0 = \inf M$.)

   Man sieht, dass $\inf M$ und $\sup M$ nicht in M liegen, d. h. $\max M$ und $\min M$ existieren hier nicht!

3. Es sei $M = \{x \in \mathbb{Q} : x \geq 0,\ x^2 \leq 2\}$. Nach B. 4.3 ist M beschränkt. Hier ist $\inf M = 0$, es existiert aber **kein Supremum** von M. Dies ergibt sich aus S. 4.1 und dem folgenden Resultat.

Satz 4.7 *Es sei K ein geordneter Körper, und es seien $n \in \mathbb{N}$ sowie $c \in K$ mit $c > 0_K$. Wir setzen $M := \{x \in K : x \geq 0, x^n \leq c\}$. Dann gilt*

1. *M ist nichtleer und nach oben beschränkt.*

2. *Existiert $s := \sup M$, so gilt $s^n = c$.*

Beweis.

1. Stets ist $0 \in M$. Weiter ist $1+c$ obere Schranke vom M. Ist $x \in K$ mit $x > 1+c$, so gilt nach der Bernoullischen Ungleichung

$$x^n > (1+c)^n \geq 1 + nc > nc \geq c$$

und damit ist $x \notin M$.)

2. Zunächst gilt für $0 \leq b \leq a$

$$a^n - b^n \leq n(a-b)a^{n-1}\,. \tag{$*$}$$

Denn: Nach der verallgemeinerten geometrischen Summenformel (B. 3.2) ist

$$a^n - b^n = (a-b)\sum_{\nu=0}^{n-1} a^\nu b^{n-\nu-1} \leq n(a-b)a^{n-1}.$$

a) Angenommen, es ist $s^n > c$. Dann existiert ein $\delta > 0$ mit $(s-\delta)^n \geq c$ (nach $(*)$ ist $\delta := \dfrac{s^n - c}{ns^{n-1}}$ geeignet).
Ist $x \in M$, so folgt $x^n \leq c \leq (s-\delta)^n$ und damit auch $x \leq s - \delta$. Also ist $s - h$ obere Schranke von M im Widerspruch dazu, dass s kleinste obere Schranke ist.

b) Angenommen, es ist $s^n < c$. Dann existiert ein $\delta > 0$ mit $(s+\delta)^n \leq c$ (nach $(*)$ ist

$$\delta := \min\left\{1, \frac{c - s^n}{n(s+1)^{n-1}}\right\}$$

geeignet). Dann ist aber $s + \delta \in M$ und damit s keine obere Schranke von M. Widerspruch.

□

Definition 4.8 Ein geordnete Menge $(X, <)$ heißt *(ordnungs-)vollständig*, falls jede nichtleere, nach oben beschränkte Teilmenge M von X ein Supremum hat.

Bemerkung und Definition 4.9 Es sei K ein vollständiger geordneter Körper.
1.Für jedes $c \in K, c \geq 0$ und jedes $n \in \mathbb{N}$ hat die Gleichung

$$x^n = c$$

genau eine Lösung $s \in K$ mit $s \geq 0$.
(Denn: Die Existenz einer Lösung x ergibt sich aus S. 4.7. Sind $x_1, x_2 \in K$ mit $0 \leq x_1 < x_2$ so ergibt sich auch $x_1^n < x_2^n$. Also hat die Gleichung $x^n = c$ höchstens eine Lösung.)
Wir setzen

$$\sqrt[n]{c} := s\,.$$

Damit ist für $0 \leq c < d$ auch $\sqrt[n]{c} < \sqrt[n]{d}$.
Schließlich ergibt sich für $c, d \in K$ $c, d \geq 0$ aus den entsprechenden Potenzgesetzen leicht ([Ü])

$$\sqrt[n]{cd} = \sqrt[n]{c}\,\sqrt[n]{d} \qquad \text{und} \qquad \sqrt[m]{\sqrt[n]{c}} = \sqrt[nm]{c}\,.$$

2. Aus der Vollständigkeit von K folgt auch die sogenannte archimedische Eigenschaft: Ist $x \in K$ positiv, so existiert ein $n \in \mathbb{N}$ mit $n1_K > x$, d. h. man kann jedes vorgegebene $x > 0$ durch „Aneinanderreihen“ genügend vieler Einsen übertreffen.

Von zentraler Bedendeutung für die Analysis ist die folgende – alles andere als leicht zu beweisende – Tatsache

Satz 4.10 *Es existiert ein vollständiger geordneter Körper* $(\mathbb{R}, +, \cdot, <)$*, der eine Erweiterung von* $(\mathbb{Q}, +, \cdot, <)$ *darstellt (d. h.* $\mathbb{Q} \subset \mathbb{R}$ *und die Einschränkungnen von* $+$*,* $\cdot$ *und* $<$ *auf* $\mathbb{Q}$ *stimmen mit den entsprechenden Funktionen bzw. Relationen in* $\mathbb{Q}$ *überein).*

Bemerkung und Definition 4.11 Die Elemente von $\mathbb{R}$ heißen *reelle Zahlen*. Auf eine mögliche Konstruktion der reellen Zahlen werden wir im Anhang eingehen.
Aus der archimedischen Eigenschaft von $\mathbb{R}$ ergibt sich folgende wichtige Konsequenz: Sind $x, y \in \mathbb{R}$, so existiert ein $r \in \mathbb{Q}$ mit $x < r < y$.

Wie wir oben gesehen haben, hat damit 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 $x \in \mathbb{R}$ nach S. 2.12.3). Unser Ziel ist es nun, den Körper der reellen Zahlen so zu erweitern, dass $x^2 = c$ auch für $c < 0$ (also etwa $x^2 = -1$) lösbar ist. (Wir werden später sehen, dass tatsächlich dann auch $x^n = c$ für beliebiges c lösbar ist.)

Bemerkung und Definition 4.12 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 $0 = 0_{\mathbb{C}} = (0_{\mathbb{R}}, 0_{\mathbb{R}})$ und die Eins in $\mathbb{C}$ ist gegeben durch $1 = 1_{\mathbb{C}} = (1_{\mathbb{R}}, 0_{\mathbb{R}})$. 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 $z = (x, y) \in \mathbb{C}$

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

und

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

Beispiel 4.13 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 4.14 Indem wir die komplexe Zahl $(x, 0)$ mit der reellen x identifizieren, können wir $\mathbb{C}$ als Erweiterung von $\mathbb{R}$ auffassen. Die reellen Zahlen entsprechen also den komplexen Zahlen mit Imaginärteil 0. *Wir schreiben dann auch kurz x statt* $(x, 0)$. Die Addition und die Multiplikation in $\mathbb{R}$ ergeben sich dabei als Einschränkungen der Addition und der Multiplikation in $\mathbb{C}$.
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 4.15 In $(\mathbb{C}, +, \cdot)$ ist es nicht möglich, eine Ordnungsrelation $<$ (mit den Eigenschaften aus D. 2.11) zu definieren!
(Denn: Angenommen, doch. Dann gilt $1_\mathbb{C} > 0_\mathbb{C}$ nach S. 2.12.3, also $-1_\mathbb{C} < 0_\mathbb{C}$ nach S. 2.12.1. Für $z = i$ gilt mit S. 2.12.3 aber andererseits $0 < i^2 = -1_\mathbb{C}$ also Widerspruch zu (O.1).)

Definition 4.16 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 4.17 *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. $\operatorname{Re}(z) = \dfrac{z + \overline{z}}{2}$ *und* $\operatorname{Im}(z) = \dfrac{z - \overline{z}}{2i}$.

Beweis. [Ü] □

Für das Rechnen mit Beträgen gelten folgende Regeln

Satz 4.18 *Für $z, z_1, z_2 \in \mathbb{C}$ gilt*

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

Beweis. 1., 2. und 3. als [Ü].
4. 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.
5. Es gilt

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

Durch Wurzelziehen folgt die Behauptung für $z_1 + z_2$. Damit erhält man dann auch

$$|z_1 - z_2| \leq |z_1| + |-z_2| = |z_1| + |z_2| \ .$$

$\square$

Beispiel 4.19 Es gilt für $z = (3, -1) = 3 - i$

$$\begin{aligned} |z| &= \sqrt{9+1} = \sqrt{10} \\ \overline{z} &= 3 - (-i) = 3 + i \\ z\overline{z} &= (3-i)(3+i) = 9 - 3i + 3i - i^2 = 9 + 1 \left(= |z|^2\right) \end{aligned}$$

Definition 4.20 In Verallgemeinerung von D. 3.3 setzen wir noch für $z \in \mathbb{C}$ und $\nu \in \mathbb{N}_0$

$$\binom{z}{\nu} := \frac{1}{\nu!}\prod_{k=1}^{\nu}(z+1-k) = \begin{cases} \dfrac{z(z-1)\cdots(z-\nu+1)}{\nu!}, & \text{falls} \quad \nu > 0 \\ 1, & \text{falls} \quad \nu = 0 \end{cases}$$

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

5 Folgen reeller und komplexer Zahlen

Es seien I, Y nichtleere Mengen und $a : I \to Y$. Man schreibt oft (wie im ersten Abschnitt angedeutet) auch $(a_j)_{j\in I}$ oder kurz (a_j).
Ist $I \subset \mathbb{N}$ oder $I \subset \mathbb{Z}$ (unendlich), so spricht man von einer *Folge* (*in* Y). Im Falle $I = \{n \in \mathbb{Z} : n \geq N\}$ schreibt man dabei auch $(a_n)_{n=N}^{\infty}$. Da die meisten der Begriffe im Weiteren nicht davon abhängen, ob eine Folge um endlich viele Folgeglieder a_n „erweitert" oder „verringert" wird, werden wir meist ohne Einschränkung $I = \mathbb{N}$ betrachten, also $(a_n) = (a_1, a_2, a_3, \ldots)$.
In diesem Abschnitt wird $Y = \mathbb{R}$ (Folgen reeller Zahlen) oder $Y = \mathbb{C}$ (Folgen komplexer Zahlen) sein. Um die beiden Fälle $\mathbb{R}$ und $\mathbb{C}$ einheitlich bezeichnen zu können, schreiben wir $\mathbb{K} := Y$, falls $Y \in \{\mathbb{R}, \mathbb{C}\}$.

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

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

2. *Cauchy-Folge*, falls zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ existiert mit

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

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

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

(für Folgen in $\mathbb{R}$ kann man die Bedingung ersetzen durch

$$a - \varepsilon < a_n < a + \varepsilon \quad \text{für alle } n \geq N_\varepsilon \text{).}$$

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

$$a_n \to a \quad (n \to \infty) \,.$$

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

4. eine *Nullfolge*, falls $a_n \to 0$ $(n \to \infty)$ gilt.

Bemerkung 5.2 1. Man sieht leicht, dass jede Folge (a_n) höchstens einen Grenzwert a hat ([Ü]). Man setzt im Falle der Konvergenz

$$a =: \lim_{n\to\infty} a_n \,.$$

Aus der Definition der Konvergenz ergibt sich außerdem sofort: Es gilt $a_n \to a$ für $n \to \infty$ genau dann, wenn $(|a_n - a|)$ eine Nullfolge ist.
2. Es sei $A(n)$ eine Aussage ($n \in \mathbb{N}$). Man sagt, $A(n)$ gilt *für alle n genügend groß*, falls die Menge $\{n \in \mathbb{N} : A(n) \text{ ist nicht wahr}\}$ endlich ist. So lässt sich etwa die Konvergenz einer Folge auch folgendermaßen formulieren: (a_n) ist genau dann konvergent gegen a, falls für alle $\varepsilon > 0$ gilt

$$|a_n - a| < \varepsilon$$

für alle n genügend groß.
3. Ein einfaches, aber sehr nützliches Nullfolgenkriterium ist das folgende: Sind (a_n) eine Folge in $\mathbb{K}$, $c > 0$ eine Konstante und (b_n) eine Nullfolge mit $|a_n| \leq c\,b_n$ für n genügend groß, so ist auch (a_n) eine Nullfolge.
(Denn: Es sei $N \in \mathbb{N}$ so, dass $|a_n| \leq c\,b_n$ für alle $n \geq N$. Ist $\varepsilon > 0$ gegeben, so existiert ein $N'_\varepsilon \in \mathbb{N}$ mit $b_n < \varepsilon/c$ ($n \geq N'_\varepsilon$). Dann ist für $n \geq N_\varepsilon := \max\{N'_\varepsilon, N\}$ auch $|a_n| \leq cb_n < \varepsilon$ ($n \geq N_\varepsilon$).)
Entsprechend gilt: Ist (b_n) unbeschränkt und $|a_n| \geq c\,b_n$ für n genügend groß, so ist auch (a_n) unbeschränkt.

Beispiel 5.3 1. Ist $a \in \mathbb{K}$ fest und ist $a_n = a$ ($n \in \mathbb{N}$), so ist (a_n) konvergent mit $a = \lim_{n\to\infty} a_n$.
2. Die Folge (a_n) in $\mathbb{R}$ mit $a_n = n$ ist unbeschränkt (nach der archimedischen Eigenschaft von $\mathbb{R}$).
3. Die Folge (a_n) in $\mathbb{R}$ mit $a_n = 1/n$ (also $(a_n) = (1, \frac{1}{2}, \frac{1}{3}, \ldots)$) ist eine Nullfolge, 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 $N_\varepsilon > 1/\varepsilon$ (wieder archimedische Eigenschaft von $\mathbb{R}$). Also gilt für alle $n \geq N_\varepsilon$

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

4. (*geometrische Folge*)
Für festes $q \in \mathbb{K}$ sei $a_n = q^n$, also $(a_n) = (q, q^2, q^3, \ldots)$. Dann gilt

a) Für $|q| < 1$ ist (q^n) eine Nullfolge, also $\lim_{n\to\infty} q^n = 0$.

b) Für $|q| > 1$ ist (q^n) unbeschränkt.

(Denn:
Für $q = 0$ ist die Behauptung klar.

Es sei $0 < |q| < 1$. Dann ist mit einem $a > 0$

$$1/|q| = 1 + a\ .$$

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

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

Aus $1/n \to 0$ folgt mit B. 5.2.3 die Behauptung.
Nun sei $|q| > 1$, d. h. $|q| = 1 + b$ mit einem $b > 0$. Mit der Bernoullischen Ungleichung gilt

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

d. h. (q^n) ist unbeschränkt nach 5.2.3.
5. Ist $q = -1$ in 4., also $a_n = (-1)^n$, so ist (a_n) beschränkt (da $|a_n| = 1$ für alle $n \in \mathbb{N}$), aber keine Cauchy-Folge.
(Denn: Für alle $k \in \mathbb{N}$ ist

$$|a_{2k} - a_{2k+1}| = 2\ ,$$

d. h. etwa zu $\varepsilon = 2$ existiert kein $N_\varepsilon \in \mathbb{N}$ mit $|a_n - a_m| < \varepsilon$ für alle $n, m \geq N_\varepsilon$.)
Aus dem nächsten Satz ergibt sich, dass (a_n) auch divergiert.

Satz 5.4 *1. Jede Cauchy-Folge (a_n) in $\mathbb{K}$ ist beschränkt.*
2. Jede konvergente Folge (a_n) in $\mathbb{K}$ ist eine Cauchy-Folge.

Beweis.
1. Zu $\varepsilon = 1$ existiert ein $N \in \mathbb{N}$ mit

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

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

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

gilt dann

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

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

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

□

Der folgende Satz zeigt, dass Grenzwertbildung mit den algebraischen Operationen in $\mathbb{K}$ verträglich ist.

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 $(a_n \pm b_n)$ konvergiert gegen $a \pm b$, d. h.*
$$\lim_{n\to\infty}(a_n \pm b_n) = a \pm b \left(= \lim_{n\to\infty} a_n \pm \lim_{n\to\infty} b_n\right)\,.$$
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 $n \geq N_\varepsilon := \max(N_\varepsilon^{(1)}, N_\varepsilon^{(2)})$:
$$|a_n \pm b_n - (a \pm b)| = |a_n - a \pm (b_n - b)| \leq |a_n - a| + |b_n - b| < \varepsilon/2 + \varepsilon/2 = \varepsilon\,.$$
2. Nach S. 5.4 ist (b_n) beschränkt, d. h. es existiert ein $M > 0$ mit $|b_n| \leq M$ für alle $n \in \mathbb{N}$. Damit ist
$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - a b_n + a b_n - ab| \leq \\ &\leq |b_n||a_n - a| + |a||b_n - b| \leq M|a_n - a| + |a||b_n - b|\,. \end{aligned}$$
Nach B. 5.2 und 1. konvergiert die rechte Seite gegen 0. Also konvergiert wieder nach B. 5.2 $(a_n b_n)$ gegen ab.

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!)
$$|b_n| = |b + b_n - b| \geq |b| - |b_n - b| > |b| - |b|/2 = |b|/2 > 0 \qquad (n \geq N)\ .$$
Damit ergibt sich
$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b_n - b|}{|b_n b|} < \frac{2}{|b|^2}|b_n - b|.$$
Aus $|b_n - b| \to 0$ folgt $1/b_n \to 1/b$ $(n \to \infty)$ mit B. 5.2.

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 $a_n = \frac{3n^2-4n}{2n^2+5}$. Dann folgt mit S. 5.5 und B. 5.3.1./3.
$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{3 - 4/n}{2 + 5/n^2} = \frac{3 - 4\lim\limits_{n\to\infty} 1/n}{2 + 5\lim\limits_{n\to\infty} 1/n^2} = \frac{3 - 4\cdot 0}{2 + 0} = \frac{3}{2}\ .$$

Bemerkung und Definition 5.7 Manchmal erweist es sich als nützlich, für reelle Folgen (a_n) erweiterte Grenzwerte $\pm\infty$ zu betrachten. Wir sagen deshalb

1. (a_n) heißt *bestimmt divergent* 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* 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 $a_n \to -\infty$ $(n \to \infty)$

Wie in B. 5.2.3 sieht man leicht: Gilt $b_n \to \infty$ und ist $a_n \geq c\, b_n$ für n genügend groß und eine Konstante $c > 0$, so gilt auch $a_n \to \infty$. Aus der Überlegung in B. 5.3.4 folgt damit etwa

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

im Falle $q > 1$.

Für $q < -1$ ist (q^n) weder bestimmt divergent gegen ∞ noch bestimmt divergent gegen $-\infty$.

6 Hauptsatz monotone Folgen und Konsequenzen

Wir betrachten zunächst speziell Folgen reeller Zahlen, wobei wir entscheidend von der Ordnungsstruktur in $\mathbb{R}$ Gebrauch machen. Insbesondere verwenden wir häufig folgende Aussage über die Verträglichkeit von Grenzwerten und $\leq$-Relation: Sind (a_n) und (b_n) konvergente Folgen in $\mathbb{R}$ mit $a_n \leq b_n$ für n genügend groß, so gilt auch ([Ü])

$$\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n\ .$$

Definition 6.1 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 6.2 1. Für festes $k \in \mathbb{N}$ sind die Folgen $\left(\frac{1}{n^k}\right)_{n=1}^\infty$ und $\left(\frac{1}{\sqrt[k]{n}}\right)_{n=1}^\infty$ streng monoton fallend.
2. $(q^n)_n$ ist streng monoton wachsend falls $q > 1$, streng monoton fallend falls $0 < q < 1$ und weder monoton wachsend noch monoton fallend falls $q < 0$. Außerdem ist $(1)_n$ monoton wachsend und fallend, aber nicht streng monoton wachsend oder fallend.
3. Die Folge (a_n) mit

$$a_n = (1 + 1/n)^n = \big(2, (3/2)^2, (4/3)^3, \dots\big)$$

ist (streng) monoton wachsend.
(Denn: Für alle $n \geq 2$ gilt

$$\begin{aligned}\frac{a_n}{a_{n-1}} &= \frac{\left(\frac{n+1}{n}\right)^n}{\left(\frac{n}{n-1}\right)^{n-1}} = \frac{n+1}{n} \cdot \left(\frac{(n+1)}{n} \cdot \frac{(n-1)}{n}\right)^{n-1} = \\ &= \frac{n+1}{n}\left(1 - \frac{1}{n^2}\right)^{n-1} \overset{B.3.10}{\geq} \frac{n+1}{n}\left(1 - \frac{n-1}{n^2}\right) = \frac{n^3+1}{n^3} > 1\ .)\end{aligned}$$

Ählich kann man zeigen ([Ü]): Die Folge (b_n) mit

$$b_n = (1 + 1/n)^{n+1} = \big(4, (3/2)^3, (4/3)^4, \dots\big) \quad (n \in \mathbb{N})$$

ist streng monoton fallend.

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

Satz 6.3 *(Hauptsatz über monotone Folgen)*
Es sei (a_n) eine Folge in $\mathbb{R}$. Ist (a_n) monoton (wachsend oder fallend) und beschränkt, so ist (a_n) konvergent.

Beweis. Wir betrachten die Menge

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

Dann ist $A \neq \emptyset$ und beschränkt. Da $\mathbb{R}$ vollständig ist, existieren

$$a := \sup A \qquad \text{und} \qquad b := \inf A .$$

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

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

Da $\varepsilon > 0$ beliebig war, folgt $a_n \to a$ $(n \to \infty)$.
2. Ist (a_n) monoton fallend, so sieht man entsprechend, dass $a_n \to b$ gilt. □

Bemerkung 6.4 Der Hauptsatz über monotone Folgen ist eines der wichtigsten Kriterien für die Konvergenz von Folgen. Man beachte, dass die Kenntnis des Grenzwerts bei der Konvergenzuntersuchung nicht vonnöten ist.
Der Beweis zeigt: Ist $A = \{a_n : n \in \mathbb{N}\}$ die Menge der Folgeglieder, so gilt

$$a_n \to \begin{cases} \sup A & \text{falls } a_n \nearrow \\ \inf A & \text{falls } a_n \searrow \end{cases} .$$

Eine gewisse Umkehrung dieser Aussage gilt ebenfalls: Ist $A \subset \mathbb{R}$ nichtleer und nach oben (bzw. nach unten) beschränkt, so existiert eine monoton wachsende (bzw. monton fallende) Folge (a_n) in A mit $a_n \to \sup A$ (bzw. $a_n \to \inf A$) für $n \to \infty$ ([Ü]).

Beispiel 6.5 (Eulersche Zahl e)
Wir betrachten noch einmal die Folgen (a_n) und (b_n) aus B. 6.2.3. d. h.

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

Aufgrund der Monotonieaussagen aus B. 6.2.3. gilt $2 = a_1 \le a_n \le b_n \le b_1 = 4$. Nach dem Hauptsatz über monotone Folgen sind (a_n) und (b_n) konvergent und mit S. 5.5.2 folgt

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} (1 + 1/n)a_n = \lim_{n\to\infty} a_n$$

Wir setzen

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

(e heißt *Eulersche Zahl*).

Viele interessante Folgen sind rekursiv definiert: Ist $\varphi : X \to X$ eine Abbildung, so ist mit einem Startwert $a_0 \in X$ durch

$$a_{n+1} := \varphi(a_n) \quad (n \in \mathbb{N}_0)$$

eine Folge $(a_n)_n$ in X gegeben. Ein typisches – und schon seit mehr als zwei Jahrtausenden bekanntes – Beispiel ist das folgende.

Beispiel 6.6 *Heron-Verfahren; Babylonisches Wurzelziehen*
Es sei $c > 0$ gegeben und $\varphi : (0, \infty) \to (0, \infty)$ definiert durch

$$\varphi(x) := \frac{1}{2}\left(x + \frac{c}{x}\right) \quad (x > 0) .$$

Wir betrachten mit einem beliebigen Startwert $a_0 > 0$ (der Startwert a_0 sollte eine grobe Näherung an $\sqrt{c}$ sein) die Folge (a_n) in $(0, \infty)$ mit

$$a_{n+1} := \varphi(a_n) = \frac{1}{2}\left(a_n + \frac{c}{a_n}\right) = \frac{1}{2a_n}\left(a_n^2 + c\right) .$$

Behauptung: (a_n) ist konvergent mit $\lim_{n\to\infty} a_n = \sqrt{c}$.

Denn: 1. Wir zeigen: $a_n \ge \sqrt{c}$ $(n \in \mathbb{N})$. Denn: Für $n \in \mathbb{N}_0$ gilt

$$a_{n+1} - \sqrt{c} = \frac{1}{2a_n}(a_n^2 - 2a_n\sqrt{c} + c) = \frac{1}{2a_n}(a_n - \sqrt{c})^2 \ge 0 ,$$

also ist $a_{n+1} \geq \sqrt{c}$.
2. Die Folge $(a_n)_{n=1}^{\infty}$ ist monoton fallend, denn für $n \in \mathbb{N}$ ist

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

Aus $a_n \geq \sqrt{c}$ folgt $a_n^2 \geq c$, also $a_n - a_{n+1} \geq 0$.
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 $a \geq \sqrt{c} > 0$)

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

Aus der Eindeutigkeit des Grenzwertes folgt

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

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

Man kann also die Folgeglieder a_n als Näherungen für $\sqrt{c}$ verwenden (dabei sind im Falle $c \in \mathbb{Q}$ und $a_0 \in \mathbb{Q}$ die Näherungen a_n stets rationale Zahlen).
Wie sieht es dabei mit dem Fehler aus, wenn man a_n statt $\sqrt{c}$ verwendet? Wir schätzen den Fehler nach oben ab. Dazu sei

$$a_n = \sqrt{c}(1 + f_n) \quad (n \in \mathbb{N})$$

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

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

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} im nächsten Schritt $\leq \frac{1}{2}(10^{-m})^2 = \frac{1}{2}10^{-2m}$; die Anzahl der “exakten Stellen” verdoppelt sich im Wesentlichen!

Bemerkung und Definition 6.7 Es sei (a_n) eine beschränkte Folge in $\mathbb{R}$. Wir betrachten die Mengen

$$B_n := \{a_k : k \geq n\}$$

Es gilt dabei $B_{n+1} \subset B_n$ (und die B_n sind beschränkt). Also existieren

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

und nach Definition von sup und inf sind die Folgen $(\overline{b_n})$ bzw. $(\underline{b_n})$ monoton fallend bzw. monoton wachsend (und beschränkt, da $\underline{b_1} \leq \underline{b_n} \leq \overline{b_n} \leq \overline{b_1}$). Also sind beide Folgen konvergent nach dem Hauptsatz über monotone Folgen. Damit existieren

$$\overline{a} := \lim_{n\to\infty} \overline{b_n} \qquad \text{und} \qquad \underline{a} := \lim_{n\to\infty} \underline{b_n}$$

Die reelle Zahl $\overline{a}$ heißt *Limes superior* von (a_n) (Schreibweise: $\overline{a} =: \overline{\lim\limits_{n\to\infty}} a_n$ oder auch $\limsup\limits_{n\to\infty} a_n$), und $\underline{a}$ heißt *Limes inferior* von (a_n) (Schreibweise: $\underline{a} = \underline{\lim\limits_{n\to\infty}} a_n$ oder auch $\liminf\limits_{n\to\infty} a_n$).

Satz 6.8 *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 $\varepsilon > 0$ gilt $a_n < a + \varepsilon$ für alle n genügend groß.*

   b) *Für alle $\varepsilon > 0$ gilt $a_n > a - \varepsilon$ für unendlich viele $n \in \mathbb{N}$.*

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

   a) *Für alle $\varepsilon > 0$ gilt $a_n > a - \varepsilon$ für alle n genügend groß.*

   b) *Für alle $\varepsilon > 0$ gilt $a_n < a + \varepsilon$ für unendlich viele $n \in \mathbb{N}$.*

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

$$\overline{b_n} = \sup\{a_k : k \geq n\} \leq a + \varepsilon$$

für alle n genügend groß. Also ist

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

Aus b) folgt $\overline{b_n} > a - \varepsilon$ für alle $n \in \mathbb{N}$. Also ergibt sich

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

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

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

Beispiel 6.9 Es sei $a_n = (-1)^n (1 + 1/n)$. Dann ist (a_n) beschränkt und mit S. 6.8 ergibt sich

$$\overline{\lim}\, a_n = 1 \quad \text{und} \quad \underline{\lim}\, a_n = -1.$$

Satz 6.10 *Es sei (a_n) eine beschränkte Folge in $\mathbb{R}$. Dann gilt*

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

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_ε mit

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

Also sind a) und b) aus S. 6.8.1 und 2. erfüllt und folglich gilt mit S. 6.8

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

"$\Longleftarrow$" Gilt $\overline{\lim} a_n = \underline{\lim} a_n =: a$, so folgt aus S. 6.8 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 $\lim a_n = a$. □

Bemerkung und Definition 6.11 Es seien $X \neq \emptyset$ eine Menge und (a_n) eine beliebige Folge in X. Ist (n_k) eine streng monoton wachsende Folge in $\mathbb{N}$, so heißt $(a_{n_k})_{k\in\mathbb{N}}$ eine *Teilfolge* von $(a_n)_{n\in\mathbb{N}}$.
Aus den Definitionen ergibt sich unmittelbar für Folgen (a_n) in $\mathbb{K}$: Ist (a_n) konvergent, so ist auch jede Teilfolge konvergent (mit gleichem Grenzwert) und ist (a_n) eine Cauchy-Folge, so ist auch jede Teilfolge eine Cauchy-Folge.
Weiter gilt: Ist (a_n) eine Cauchy-Folge und exsitiert eine konvergente Teilfoge (a_{n_k}), so ist auch die Gesamtfolge (a_n) konvergent.
(Denn: Es sei $a := \lim\limits_{k\to\infty} a_{n_k}$ und es sei $\varepsilon > 0$ gegeben. Dann esistiert ein $N_\varepsilon \in \mathbb{N}$ mit

$$|a_n - a_m| < \varepsilon/2 \qquad (n, m \geq N_\varepsilon)\ .$$

Weiter existiert ein $k = k_\varepsilon$ so, dass $n_k \geq N_\varepsilon$ und $|a_{n_k} - a| < \varepsilon/2$. Damit ist

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \varepsilon \qquad (n \geq N_\varepsilon)\ .)$$

Beispiel 6.12 1. Es sei $a_n = (-1)^n(1 + 1/n)$ (vgl. B 6.9). Dann sind

$$(a_{2k})_{k\in\mathbb{N}} = \Big(1 + \frac{1}{2k}\Big)_{k\in\mathbb{N}} \qquad \text{und} \qquad (a_{2k-1})_{k\in\mathbb{N}} = \Big(-1 - \frac{1}{2k-1}\Big)_{k\in\mathbb{N}}$$

Teilfolgen von (a_n). Dabei gilt $a_{2k} \to 1 = \overline{\lim}\, a_n$ und $a_{2k-1} \to -1 = \underline{\lim}\, a_n$ für $k \to \infty$. Die Folge (a_n) selbst ist divergent.

Allgemein erhalten wir

Satz 6.13 *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*

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

Beweis. Es sei $a := \overline{\lim\limits_{n\to\infty}}\, a_n$.
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 $|a_{n_{k+1}} - a| < 1/k$ (existiert nach S. 6.8.1 mit $\varepsilon = 1/k$). Dann gilt $a_{n_k} \to a$ $(k \to \infty)$.

2. Analog
3. Es sei (a_{ℓ_k}) eine konvergente Teilfolge von (a_n). Aus $a_{\ell_k} \le \overline{b_{\ell_k}}$ folgt $\lim\limits_{k\to\infty} a_{\ell_k} \le \lim\limits_{k\to\infty} \overline{b_{\ell_k}} = a$. Entsprechend sieht man, dass $\underline{\lim}\, a_n \le \lim\limits_{k\to\infty} a_{\ell_k}$ gilt. □

Als Konsequenz erhalten wir zwei der zentralen Ergebnisse der Analysis, den Satz von Bolzano und Weierstraß sowie das Cauchysche Konvergenzkriterium.

Satz 6.14 *(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. 6.13.1.
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 $|\alpha_n| \le |a_n|$ und $|\beta_n| \le |a_n|$). 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 $\alpha_{n_{k_\ell}} + i\beta_{n_{k_\ell}} \to \alpha + i\beta$ $(\ell \to \infty)$. □

Beispiel 6.15 Es sei $a_n = q^n$ mit $q \in \mathbb{C}$, $|q| = 1$. Dann ist auch $|a_n| = 1$ für alle n. Also hat nach S. 6.14 die geometrische Folge (a_n) eine konvergente Teilfolge. Ist etwa $q = i$, so gilt

$$a_{4k} = 1 \to 1, \quad a_{4k+1} = i \to i, \quad a_{4k+2} = -1 \to -1, \quad a_{4k+3} = -i \to -i\,.$$

Satz 6.16 *(Cauchysches Konvergenzkriterium)*
Es sei (a_n) eine Folge in $\mathbb{K}$. Genau dann ist (a_n) konvergent wenn (a_n) eine Cauchy-Folge ist.

Beweis. Ist (a_n) konvergent, so ist (a_n) nach S. 5.4.2 eine Cauchy-Folge.
Es sei umgekehrt (a_n) eine Cauchy-Folge. Dann ist (a_n) jedenfalls beschränkt nach S. 5.4.1. Also hat (a_n) nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge (a_{n_k}). Nach B./D. 6.11 ist (a_n) konvergent. □

7 Reihen

Wir betrachten wieder die Folge $(q^\nu)_{\nu=0}^\infty$ für ein $q \in \mathbb{K}$ mit $|q| < 1$. Für die Summe der ersten $(n+1)$ Folgelieder 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 $\frac{1}{1-q}$.

Bemerkung und Definition 7.1 Es seien $n_0 \in \mathbb{Z}$ und $(a_\nu)_{\nu=n_0}^\infty$ eine Folge in $\mathbb{K}$.
1. Die der Folge (a_ν) zugeordnete Folge $(s_n)_{n=n_0}^\infty$ der *Partial-* oder *Teilsummen*

$$s_n := \sum_{\nu=n_0}^{n} a_\nu \qquad (n \in \mathbb{Z},\, n \geq n_0)$$

heißt (*die mit* (a_ν) *gebildete*) *Reihe* und wird mit $\sum\limits_{\nu=n_0}^{\infty} a_\nu$ bezeichnet. Die a_ν heißen dann *Reihenglieder*.
2. Ist die Reihe $\sum\limits_{\nu=n_0}^{\infty} a_\nu$ (also die Folge (s_n)) konvergent gegen s, so schreiben wir $s =: \sum\limits_{\nu=n_0}^{\infty} a_\nu$. Die Zahl s heißt dann der *Reihenwert*.

Man beachte, dass das Symbol $\sum\limits_{\nu=n_0}^{\infty} a_\nu$ also zwei Bedeutungen hat: Erstens steht es für die Folge (s_n) der Teilsummen und zweitens (im Falle der Konvergenz!) für deren Grenzwert.

Beispiel 7.2 1. (geometrische Reihe) 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

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

Speziell ergibt sich etwa für $q = 1/2$

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

2. 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

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

Für das "Rechnen" mit konvergenten Reihen gilt

Satz 7.3 *Es seien* $\sum\limits_{\nu=n_0}^{\infty} a_\nu$ *und* $\sum\limits_{\nu=n_0}^{\infty} b_\nu$ *konvergente Reihen in* $\mathbb{K}$*, und es seien* $\alpha, \beta \in \mathbb{K}$*. Dann ist auch* $\sum\limits_{\nu=n_0}^{\infty} (\alpha a_\nu + \beta b_\nu)$ *konvergent mit*

$$\sum_{\nu=n_0}^{\infty} (\alpha a_\nu + \beta b_\nu) = \alpha \sum_{\nu=n_0}^{\infty} a_\nu + \beta \sum_{\nu=n_0}^{\infty} b_\nu \, .$$

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

Damit gilt etwa

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

Insbesondere erhält man aus S. 7.3 auch: Ist $n > n_0$, so ist $\sum\limits_{\nu=n_0}^{\infty} a_\nu$ genau dann konvergent, wenn $\sum\limits_{\nu=n}^{\infty} a_\nu$ konvergiert, und in diesem Fall ist

$$\sum_{\nu=n_0}^{\infty} a_\nu = \sum_{\nu=n_0}^{n-1} a_\nu + \sum_{\nu=n}^{\infty} a_\nu \, .$$

Für Konvergenzuntersuchungen ist es also unwichtig, wie die untere Summationsgrenze aussieht. Wir werden daher im Weiteren meist o. E. $n_0 = 0$ (oder $n_0 = 1$) betrachten.

Bemerkung 7.4 Eine **notwendige Bedingung** für die Konvergenz einer Reihe ist, dass die Reihenglieder eine Nullfolge bilden, d. h. ist $\sum_{\nu=0}^{\infty} a_\nu$ konvergent, so gilt

$$a_n \to 0 \qquad (n \to \infty).$$

(Denn: Ist $s_n = \sum_{\nu=0}^{n} a_\nu$, so gilt

$$a_n = s_n - s_{n-1} \to \sum_{\nu=0}^{\infty} a_\nu - \sum_{\nu=0}^{\infty} a_\nu = 0 \quad (n \to \infty).)$$

Ist etwa $a_n = q^n$ mit $|q| \geq 1$, so ist $|a_n| \geq 1$ $(n \in \mathbb{N})$, also ist $\sum_{\nu=0}^{\infty} q^\nu$ sicher divergent. Damit ergibt sich für geometrische Reihen insgesamt:
$\sum_{\nu=0}^{\infty} q^\nu$ **ist genau dann konvergent, wenn** $|q| < 1$ **ist**.

Wir betrachten jetzt speziell Reihen mit nichtnegativen Gliedern.

Bemerkung 7.5 Ist $\sum_{\nu=0}^{\infty} a_\nu$ eine Reihe mit $a_\nu \geq 0$ $(n \in \mathbb{N})$, so ist $s_n = \sum_{\nu=0}^{n} a_\nu \uparrow$. Also gilt

- entweder ist (s_n) beschränkt und damit $\sum_{\nu=0}^{\infty} a_\nu$ konvergent (Hauptsatz über monotone Folgen)
- oder (s_n) ist unbeschränkt und damit $s_n \to \infty$ $(n \to \infty)$.

Wir schreiben im ersten Fall dann auch $\sum_{\nu=0}^{\infty} a_\nu < \infty$ und im zweiten Fall $\sum_{\nu=0}^{\infty} a_\nu = \infty$.

Satz 7.6 *(Cauchyscher Verdichtungssatz)*
Es sei $a_\nu \geq 0$ $(n \in \mathbb{N})$ *und* (a_ν) *monoton fallend.*

1. *Ist* $s_n = \sum_{\nu=1}^{n} a_\nu$ *und* $\sigma_n = \sum_{k=0}^{n} 2^k a_{2^k}$, *so gilt*

$$s_{2^{n+1}-1} \leq \sigma_n \leq 2 s_{2^n} \qquad (n \in \mathbb{N}).$$

2. $\sum_{\nu=1}^{\infty} a_\nu$ *ist genau dann konvergent, wenn* $\sum_{k=0}^{\infty} 2^k a_{2^k}$ *konvergiert.*

Beweis. 1. Für $n \in \mathbb{N}$ ist

$$\begin{aligned} s_{2^{n+1}-1} &= a_1 + (a_2 + a_3) + (a_4 + \cdots + a_7) + \cdots + (a_{2^n} + \cdots + a_{2^{n+1}-1}) \\ &\leq 2^0 a_1 + 2 \cdot a_2 + 4 \cdot a_4 + \cdots + 2^n a_{2^n} (= \sigma_n) \\ &= 2\left(\frac{1}{2}a_1 + a_2 + 2 \cdot a_4 + 4 \cdot a_8 + \cdots + 2^{n-1} a_{2^n}\right) \\ &\leq 2\big(a_1 + a_2 + (a_3 + a_4) + (a_5 + \cdots a_8) + (a_{2^{n-1}+1} + \cdots + a_{2^n})\big) \\ &= 2\, s_{2^n} \end{aligned}$$

2. „$\Longrightarrow$“ Ist $\sum\limits_{\nu=0}^{\infty} a_\nu < \infty$, also insbesondere (s_{2^n}) beschränkt, so ist auch (σ_n) beschränkt nach 1. Damit ist $\sum\limits_{k=0}^{\infty} 2^k a_{2^k} < \infty$.

„$\Longleftarrow$ “ Ist $\sum\limits_{k=0}^{\infty} 2^k a_{2^k} < \infty$, so ist (σ_m) beschränkt, also nach 1. auch $(s_{2^{m+1}-1})_m$. Da (s_n) monoton wachsend ist, ist (s_n) beschränkt, also $\sum\limits_{\nu=0}^{\infty} a_\nu < \infty$. □

Beispiel 7.7 (Harmonische Reihen)
Es sei $p \in \mathbb{N}$. Dann gilt

$$\sum_{\nu=1}^{\infty} \frac{1}{\nu^p} \begin{cases} = \infty & \text{, falls } p = 1 \\ < \infty & \text{, falls } p > 1 \end{cases}.$$

(Denn: Es ist $0 \leq 1/\nu^p \downarrow$ und

$$\sum_{k=0}^{\infty} 2^k a_{2^k} = \sum_{k=0}^{\infty} \frac{2^k}{(2^k)^p} = \sum_{k=0}^{\infty} \left(\frac{1}{2^{p-1}}\right)^k.$$

Mit S. 7.6 und B. 7.4 ergibt sich die Behauptung.)
Insbesondere zeigt der Fall $p = 1$, dass aus $a_n \to 0$ $(n \to \infty)$ im Allgemeinen noch nicht die Konvergenz von $\sum\limits_{\nu=0}^{\infty} a_\nu$ folgt.

Satz 7.8 *(Leibniz-Kriterium für alternierenede Reihen)*
Es sei $a_\nu \geq 0$ $(\nu \in \mathbb{N}_0)$ und (a_ν) monoton fallend.

1. *Ist $s_n = \sum\limits_{\nu=0}^{n} (-1)^\nu a_\nu$, so ist (s_{2m}) monoton fallend und (s_{2m+1}) monoton wachsend. Außerdem gilt $s_{2m} \geq 0$ für alle $m \in \mathbb{N}$.*

2. Gilt $a_n \to 0$ für $n \to \infty$, so ist $\sum\limits_{\nu=0}^{\infty} (-1)^\nu a_\nu$ konvergent.

Beweis. 1. Für $n \geq 2$ ist

$$s_n - s_{n-2} = \sum_{\nu=0}^{n} (-1)^\nu a_\nu - \sum_{\nu=0}^{n-2} (-1)^\nu a_\nu = (-1)^n \underbrace{(a_n - a_{n-1})}_{\leq 0}.$$

Also ist $s_{2m} - s_{2m-2} \leq 0$ und $s_{2m+1} - 2_{2m-1} \geq 0$ ($m \in \mathbb{N}$). Weiter gilt

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

2. Nach dem Hauptsatz über monotone Folgen ist (s_{2m}) konvergent. Ist $s := \lim\limits_{m \to \infty} s_{2m}$, so gilt auch

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

Hieraus folgt $s_n \to s$ für $n \to \infty$ ([Ü]). □

Beispiel 7.9 (alternierende harmonische Reihe)
Die Reihe $\sum\limits_{\nu=1}^{\infty} (-1)^\nu / \nu$ konvergiert nach S. 7.8 (denn: $a_n \downarrow 0$ ($n \to \infty$)). Während also $\sum\limits_{\nu=1}^{\infty} 1/\nu$ (nach B. 7.7) divergiert, führt das "Anbringen" abwechselnder Vorzeichen zur Konvergenz.

Satz 7.10 *(Cauchy-Kriterium für Reihen)*
Es sei (a_ν) eine Folge in $\mathbb{K}$. Dann ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ genau dann konvergent, wenn für alle $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ so existiert, dass

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

Beweis. Ist $s_n = \sum\limits_{\nu=0}^{n} a_\nu$, so ist für $n > m \geq 0$

$$|s_n - s_m| = |s_m - s_n| = \Big| \sum_{\nu=m+1}^{n} a_\nu \Big|.$$

Damit ergibt sich die Behauptung unmittelbar aus dem Cauchyschen Konvergenzkriterium für Folgen. □

Eines der wichtigsten Kriterien für Konvergenz ist

Satz 7.11 *(Weierstraßsches Majorantenkriterium)*
Es seien (a_ν) eine Folge in $\mathbb{K}$ und $b_\nu \geq 0$ mit $|a_\nu| \leq b_\nu$ für ν genügend groß. Ist $\sum\limits_{\nu=0}^{\infty} b_\nu < \infty$, so ist auch $\sum\limits_{\nu=0}^{\infty} a_\nu$ konvergent.

Beweis. Es sei n_0 so, dass $|a_\nu| \leq b_\nu$ für $\nu \geq n_0$.
Ist $\varepsilon > 0$ gegeben, so existiert nach S. 7.10 ein $N_\varepsilon\,(> n_0)$ so, dass

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

Aus der Dreiecksungleichung folgt

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

Wieder nach S. 7.10 ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ konvergent. □

Beispiel 7.12

1. Es sei $a_\nu = (-1)^\nu \dfrac{\nu^2 + 3\nu}{2\nu^4 + 4}$. Dann gilt

$$\nu^2 |a_\nu| = \frac{1 + 3/\nu}{2 + 4/\nu^4} \to \frac{1}{2} \qquad (\nu \to \infty).$$

   Also ist $|a_\nu| \leq 1/\nu^2$ für ν genügend groß. Aus $\sum\limits_{\nu=1}^{\infty} 1/\nu^2 < \infty$ ergibt sich die Konvergenz von $\sum\limits_{\nu=1}^{\infty} a_\nu$ mit S. 7.11.

2. Es sei $p \in \mathbb{N}$ und $b_\nu = 1/\sqrt[p]{\nu}$. Dann ist

$$b_\nu = \frac{1}{\sqrt[p]{\nu}} \geq \frac{1}{\nu} \qquad (\nu \in \mathbb{N}).$$

   Also ist $\sum\limits_{\nu=0}^{\infty} 1/\sqrt[p]{\nu} = \infty$ (denn wäre $\sum\limits_{\nu=1}^{\infty} 1/\sqrt[p]{\nu} < \infty$, so wäre auch $\sum\limits_{\nu=1}^{\infty} 1/\nu < \infty$ nach S. 7.11).

Bemerkung und Definition 7.13 Es sei (a_ν) eine Folge in $\mathbb{K}$. Gilt

$$\sum_{\nu=0}^{\infty} |a_\nu| < \infty$$

so ist nach dem Majorantenkriterium S. 7.11 (angewandt mit $b_n = |a_n|$) auch $\sum\limits_{\nu=0}^{\infty} a_\nu$ konvergent. Außerdem gilt dann ([Ü])

$$\left| \sum_{\nu=0}^{\infty} a_\nu \right| \leq \sum_{\nu=0}^{\infty} |a_\nu| \, .$$

Die Reihe $\sum\limits_{\nu=0}^{\infty} a_\nu$ heißt *absolut konvergent*, falls $\sum\limits_{\nu=0}^{\infty} |a_\nu| < \infty$ ist.
Ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ konvergent und $\sum\limits_{\nu=0}^{\infty} |a_\nu| = \infty$, so heißt $\sum\limits_{\nu=0}^{\infty} a_\nu$ *bedingt konvergent*.

Beispiel 7.14 1. Die Reihe aus B. 7.12.1 ist absolut konvergent.
2. Es sei $p \in \mathbb{N}$. Die Reihe $\sum\limits_{\nu=1}^{\infty} (-1)^\nu/\nu^p$ ist für $p = 1$ bedingt konvergent und für $p \geq 2$ absolut konvergent (B. 7.9 und B. 7.7).

Wir schreiben für eine Folge (c_n) in $\mathbb{R}$

$$\overline{\lim_{n\to\infty}}\, c_n = \infty$$

falls für jedes $R > 0$ unendlich viele n existieren mit $c_n > R$ und

$$\underline{\lim_{n\to\infty}}\, c_n = \infty$$

falls $c_n \to \infty$ gilt. Außerdem gelte die Konvention $\infty > x$ für alle $x \in \mathbb{R}$.

Durch Anwendung von S. 7.12 auf die konvergente Majorante $\sum\limits_{\nu=0}^{\infty} q^\nu$ für geeignetes $0 < q < 1$ ergibt sich:

Satz 7.15 *(Wurzelkriterium)*
Es sei (a_ν) eine Folge in $\mathbb{K}$ und es sei

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

Dann gilt
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 existiert zu $q := (1+a)/2 (< 1)$ ein $N \in \mathbb{N}$ mit $\sqrt[\nu]{|a_\nu|} \le q$ (und damit auch $|a_\nu| \le q^\nu$) für alle $\nu \ge N$. Da $\sum\limits_{\nu=0}^{\infty} q^\nu$ konvergiert, ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ absolut konvergent nach S. 7.11.
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) sicher nicht konvergent gegen 0. Nach B. 7.4 ist $\sum\limits_{\nu=0}^{\infty} a_\nu$ divergent. □

Beispiel 7.16 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]),

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

Nach S. 7.15 ist $\sum\limits_{\nu=0}^{\infty} \nu^p q^\nu$ absolut konvergent. (Insbesondere ergibt sich mit B. 7.4 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 der Situation von S. 7.15. Wie oben gesehen ist $\sum\limits_{\nu=1}^{\infty} 1/\nu$ divergent und $\sum\limits_{\nu=1}^{\infty} 1/\nu^p$ für $p > 1$ (absolut) konvergent. Also lässt sich im Fall $a = 1$ im Allgemeinen keine Aussage machen.

Satz 7.17 *(Quotientenkriterium)*
Es sei (a_ν) eine Folge in $\mathbb{K}$ mit $a_\nu \neq 0$ für $\nu \in \mathbb{N}_0$. Dann gilt

1. $\varliminf\limits_{\nu\to\infty} \left|\frac{a_{\nu+1}}{a_\nu}\right| \le \varliminf\limits_{\nu\to\infty} \sqrt[\nu]{|a_\nu|} \le \varlimsup\limits_{\nu\to\infty} \sqrt[\nu]{|a_\nu|} \le \varlimsup\limits_{\nu\to\infty} \left|\frac{a_{\nu+1}}{a_\nu}\right|$
2. *Ist* $\varlimsup\limits_{\nu\to\infty} \left|\frac{a_{\nu+1}}{a_\nu}\right| < 1$, *so ist* $\sum\limits_{\nu=0}^{\infty} a_\nu$ *absolut konvergent.*

Beweis. 1. Es sei $a := \varlimsup\limits_{\nu\to\infty} \left|\frac{a_{\nu+1}}{a_\nu}\right|$ (o. E. $< \infty$). Dann existiert zu jedem $\varepsilon > 0$ ein $N = N_\varepsilon \in \mathbb{N}$ mit

$$\left|\frac{a_{\nu+1}}{a_\nu}\right| \le a + \varepsilon \qquad (\nu \ge N) .$$

Für $\nu > N$ gilt dann

$$|a_\nu| = \left|\frac{a_\nu}{a_{\nu-1}}\right| \cdot \left|\frac{a_{\nu-1}}{a_{\nu-2}}\right| \cdots \left|\frac{a_{N+1}}{a_N}\right| |a_N| \le (a+\varepsilon)^{\nu-N}|a_N| = \frac{|a_N|}{(a+\varepsilon)^N} \cdot (a+\varepsilon)^\nu$$

also

$$\sqrt[\nu]{|a_\nu|} \le (a+\varepsilon)\sqrt[\nu]{\frac{|a_N|}{(a+\varepsilon)^N}} \longrightarrow a+\varepsilon \qquad (\nu \to \infty)$$

und damit

$$\overline{\lim_{\nu\to\infty}} \sqrt[\nu]{|a_\nu|} \le a+\varepsilon\ .$$

Da $\varepsilon > 0$ beliebig war, folgt die Behauptung.
Der Beweis für $\underline{\lim}$ verläuft ähnlich.
2. Nach dem Wurzelkriterium (S. 7.15) und 1. ist $\sum a_\nu$ absolut konvergent. □

Beispiel 7.18 Wir betrachten $a_\nu := \frac{z^\nu}{\nu!}$ $(\nu \in \mathbb{N}_0)$, wobei $z \in \mathbb{C}$ fest ist. Dann gilt (für $z \ne 0$)

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

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).

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

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

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 Summanden $a_\nu b_\mu\,(\nu \in \mathbb{N}_0, \mu \in \mathbb{N}_0)$ genau einmal auftauchen. Anders als bei endlichen Summen spielt dabei jedoch die “Reihenfolge” der

Summation im Allgemeinen eine Rolle. Eine mögliche Anordnung ist die folgende

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

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 7.19 Es seien $a = (a_\nu)_{\nu=0}^{\infty}$ und $b = (b_\nu)_{\nu=0}^{\infty}$ Folgen in $\mathbb{K}$. Wir setzen

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

Dann heißt die Reihe

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

Cauchysche 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 7.20 *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 Cauchysche Produktreihe $\sum\limits_{n=0}^{\infty} (a * b)_n$ und es gilt*

$$\sum_{n=0}^{\infty} (a * b)_n = \left(\sum_{\nu=0}^{\infty} a_\nu \right) \left(\sum_{\nu=0}^{\infty} b_\nu \right) \ .$$

Beweis. 1. Wir zeigen allgemein: Es seien $c_{n\nu} \in \mathbb{K}$ ($\nu = 0, \ldots, n$; $n \in \mathbb{N}_0$) mit folgenden Eigenschaften:

(i) Für alle festen $\nu \in \mathbb{N}_0$ ist $(c_{n\nu})_{n=\nu}^{\infty}$ eine Nullfolge.

(ii) Es existiert ein $M > 0$ so, dass $\sum\limits_{\nu=0}^{n} |c_{n\nu}| \leq M$ für alle $n \in \mathbb{N}_0$.

Dann gilt: Ist $(d_n)_n$ eine Nullfolge in $\mathbb{K}$, so ist auch $\left(\sum_{\nu=0}^{n} c_{n\nu} d_\nu\right)_n$ eine Nullfolge. Denn: Es sei $d > 0$ so, dass $|d_n| \leq d$ für alle $n \in \mathbb{N}_0$. Ist $\varepsilon > 0$ gegeben, so existiert ein $N = N_\varepsilon \in \mathbb{N}$ mit $|d_n| < \varepsilon$ ($n \geq N$). Damit ergibt sich für $n \geq N$

$$\left|\sum_{\nu=0}^{n} c_{n\nu} d_\nu\right| \leq \sum_{\nu=0}^{N-1} |c_{n\nu}| \underbrace{|d_\nu|}_{\leq d} + \sum_{\nu=N}^{n} |c_{n\nu}| \underbrace{|d_\nu|}_{\leq \varepsilon} \leq d \cdot \sum_{\nu=0}^{N-1} |c_{n\nu}| + \varepsilon \cdot M \ .$$

Nach Voraussetzung gilt $c_{n\nu} \to 0$ ($n \to \infty$) für alle festen $\nu = 0, \ldots, N-1$, also auch $\sum\limits_{\nu=0}^{N-1} |c_{n\nu}| \to 0$ ($n \to \infty$). Folglich existiert ein $N'_\varepsilon \geq N$ so, dass

$$\sum_{\nu=0}^{N-1} |c_{n\nu}| < \varepsilon \qquad (n \geq N'_\varepsilon) \ .$$

Also gilt insgesamt

$$\left|\sum_{\nu=0}^{n} c_{n\nu} d_\nu\right| < \varepsilon(d + M) \qquad (n \geq N'_\varepsilon) \ .$$

Da $\varepsilon > 0$ beliebig war folgt die Behauptung.

2. 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

$$\begin{aligned}
\sum_{\nu=0}^{n} (a * b)_\nu &= a_0 b_0 + (a_0 b_1 + a_1 b_0) + (a_0 b_2 + a_1 b_1 + a_2 b_0) + \ldots + (a_0 b_n + \ldots + a_n b_0) \\
&= a_0 B_n + a_1 B_{n-1} + \ldots + a_n B_0 \\
&= a_0(B_n - B) + a_1(B_{n-1} - B) + \ldots + a_n(B_0 - B) + B \cdot \sum_{\nu=0}^{n} a_\nu \ .
\end{aligned}$$

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

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

Dazu sei $c_{n\nu} := a_{n-\nu}$ für $\nu = 0, \ldots, n$, $n \in \mathbb{N}_0$. Dann gilt

$$\sum_{\nu=0}^{n} |c_{n\nu}| = \sum_{\nu=0}^{n} |a_{n-\nu}| = \sum_{\mu=0}^{n} |a_\mu| \leq \sum_{\mu=0}^{\infty} |a_\mu| \quad (n \in \mathbb{N}_0),$$

also ist (ii) aus 1. erfüllt. Außerdem gilt $a_n \to 0$ $(n \to \infty)$ nach B. 7.4 und damit auch $c_{n\nu} = a_{n-\nu} \to 0$ $(n \to \infty)$ für alle ν, d. h. (i) aus 1. ist ebenfalls erfüllt. Da $d_\nu := B_\nu - B \to 0$ gilt, ergibt sich die Behauptung nach 1. □

Beispiel 7.21 1. 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 Cauchysche Produktreihe gegeben durch $\sum\limits_{n=0}^{\infty} (a * b)_n$ mit

$$(a * b)_n = \sum_{\nu=0}^{n} a_\nu b_{n-\nu} = \sum_{\nu=0}^{n} z^\nu z^{n-\nu} = z^n \sum_{\nu=0}^{n} 1 = (n+1) z^n .$$

Nach S. 7.20 gilt

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

2. 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 Cauchysche Produktreihe $\sum\limits_{n=0}^{\infty} (a * b)_n$ gilt

$$(a * b)_n = (-1)^n \sum_{\nu=0}^{n} \frac{1}{\sqrt{\nu+1}\sqrt{n-\nu+1}} ,$$

also

$$|(a * b)_n| \geq \sum_{\nu=0}^{n} \frac{1}{\sqrt{n+1}\sqrt{n+1}} = 1 .$$

Also ist $\sum\limits_{n=0}^{\infty} (a * b)_n$ divergent nach B. 7.4. Dieses Beispiel zeigt insbesondere, dass auf die Voraussetzung der absoluten Konvergenz einer der beiden Reihen in S. 7.20 nicht verzichtet werden kann!

8 Elementare Funktionen und Zwischenwertsatz

Definition 8.1 Es seien $M \subset \mathbb{K}$ und $f : M \to K$. Ist $x_0 \in M$, so heißt f *(folgen-) stetig* an der Stelle x_0, falls für alle Folgen (x_n) in M mit $x_n \to x_0$ auch

$$f(x_n) \to f(x_0) \qquad (n \to \infty)$$

gilt. Ferner heißt f *stetig* (auf M), falls f stetig an allen Stellen $x_0 \in M$ ist.

Bemerkung 8.2 Es seien M eine beliebige Menge und $f, g : M \to \mathbb{K}$. Wir definieren

$$\begin{aligned} f \pm g : M \to \mathbb{K},\ (f \pm g)(x) := f(x) + g(x) \\ f \cdot g : M \to \mathbb{K},\ (f \cdot g)(x) := f(x) \cdot g(x) \end{aligned}$$

und im Falle $g(x) \neq 0$ für alle $x \in M$

$$f/g : M \to \mathbb{K},\ (f/g)(x) := f(x)/g(x).$$

Aus S. 5.5 ergibt sich unmittelbar: Ist $M \subset \mathbb{K}$ und sind f, g stetig an $x_0 \in M$, so sind auch $f \pm g, f \cdot g$ und (im Falle der Existenz) f/g stetig an x_0.

Beispiel 8.3 Ein *Polynom* ist eine Funktion $P : \mathbb{K} \to \mathbb{K}$ der Form

$$P(x) = \sum_{\nu=0}^{d} a_\nu x^\nu$$

mit $a_0, \ldots, a_d \in \mathbb{K}$. Ist $a_d \neq 0$, so heißt d der Grad von P.
Jedes Polynom ist stetig.
(Denn: Offensichtlich sind $\mathbb{K} \ni x \mapsto 1 \in \mathbb{K}$ und $\mathrm{id}_{\mathbb{K}}$ stetig. Durch wiederholte Anwendung von B. 8.2 ergibt sich die Stetigkeit von P.)
Sind P, Q Polynome und ist $Z(Q) := \{x \in \mathbb{K} : Q(x) = 0\}$ die Nullstellenmenge von Q, so ist auch $P/Q : \mathbb{K} \setminus Z(Q) \to \mathbb{K}$ stetig (wieder nach B. 8.2).

Für $z \in \mathbb{C}$ betrachten wir die Reihe $\sum\limits_{\nu=0}^{\infty} \frac{z^\nu}{\nu!}$, die nach B. 7.18 absolut konvergent ist.

Definition 8.4 Die Funktion $\exp : \mathbb{C} \to \mathbb{C}$, gegeben durch

$$\exp(z) := \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} \qquad (z \in \mathbb{C})\ ,$$

heißt *(komplexe) Exponentialfunktion.*

Eine der zentralen Eigenschaften der Exponentialfunktion stellt die folgende Funktionalgleichung dar, die zeigt, dass exp aus der Addition eine Multiplikation macht.

Satz 8.5 *Es seien $z, w \in \mathbb{C}$. Dann gilt*

$$\exp(z+w) = \exp(z) \cdot \exp(w) \; .$$

Beweis. Die Reihen $\sum\limits_{\nu=0}^{\infty} \frac{z^\nu}{\nu!}$ und $\sum\limits_{\nu=0}^{\infty} \frac{w^\nu}{\nu!}$ konvergieren absolut nach B. 7.18. Also konvergiert nach S. 7.20 das Cauchy-Produkt der beiden Reihen, und es gilt mit der binomischen Formel

$$\begin{aligned}\exp(z) \cdot \exp(w) &= \left(\sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!}\right)\left(\sum_{\nu=0}^{\infty} \frac{w^\nu}{\nu!}\right) = \sum_{n=0}^{\infty}\sum_{\nu=0}^{n} \frac{z^\nu}{\nu!}\frac{w^{n-\nu}}{(n-\nu)!} = \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\nu=0}^{n} \binom{n}{\nu} z^\nu w^{n-\nu} = \sum_{n=0}^{\infty} \frac{1}{n!}(z+w)^n = \exp(z+w) \; .\end{aligned}$$

□

Als Folgerung erhalten wir

Satz 8.6 *1. Für alle $z \in \mathbb{C}$ ist $\exp(z) \neq 0$ und es gilt $\exp(-z) = 1/\exp(z)$.*

2. exp *ist stetig.*

Beweis. 1. Nach D. 8.4 gilt $\exp(0) = 1$. Also folgt aus S. 8.5 auch

$$1 = \exp(z-z) = \exp(z) \cdot \exp(-z)$$

d. h. $\exp(-z) = 1/\exp(z)$.
2. Zunächst gilt für $|z| < 1$ mit B. 7.13

$$|\exp(z) - 1| = \Big|\sum_{\nu=1}^{\infty} \frac{z^\nu}{\nu!}\Big| \leq \sum_{\nu=1}^{\infty} \frac{|z|^\nu}{\nu!} = |z| \sum_{\nu=1}^{\infty} \frac{|z|^{\nu-1}}{\nu!} \leq |z| \sum_{\nu=1}^{\infty} \frac{1}{\nu!} = |z|(\exp(1)-1) \; .$$

Ist also (z_n) eine Folge in $\mathbb{C}$ mit $z_n \to 0$, so folgt für n genügend groß

$$|\exp(z_n) - 1| \leq |z_n|(\exp(1)-1) \to 0 \qquad (n \to \infty),$$

also $\exp(z_n) \to 1 (= \exp(0))$ $(n \to \infty)$.
Ist $z_0 \in \mathbb{C}$ beliebig, so ergibt sich mit S. 8.5 für $(z_n) \in \mathbb{C}$ mit $z_n \to z_0$

$$\exp(z_n) - \exp(z_0) = \exp(z_0)\Big(\exp\underbrace{(z_n - z_0)}_{\to 0 (n\to\infty)} - 1\Big) \to 0 \qquad (n \to \infty),$$

also $\exp(z_0) \to \exp(z_0)$. Damit ist exp stetig. □

Satz 8.7 *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. Wir setzen

$$s_n := \sum_{\nu=0}^{n} \frac{z^\nu}{\nu!} \qquad \text{und} \qquad b_n := (1 + z/n)^n .$$

Damit genügt es, zu zeigen $s_n - b_n \to 0$ (denn dann folgt $b_n = s_n - (s_n - b_n) \to \exp(z)$).
Es gilt mit der binomischen Formel

$$b_n = \sum_{\nu=0}^{n} \binom{n}{\nu} \frac{z^\nu}{n^\nu}$$

also

$$s_n - b_n = \sum_{\nu=0}^{n} \left(1 - \frac{n(n-1)\cdots(n-\nu+1)}{n^\nu}\right) \frac{z^\nu}{\nu!} = \sum_{\nu=0}^{n} c_{n\nu} \frac{(2z)^\nu}{\nu!}$$

mit

$$c_{n\nu} := \underbrace{\left(1 - \frac{n(n-1)\cdots(n-\nu+1)}{n^\nu}\right)}_{\le 1} \frac{1}{2^\nu}.$$

Dabei ist

$$\sum_{\nu=0}^{n} |c_{n\nu}| \le \sum_{\nu=0}^{\infty} \left(\frac{1}{2}\right)^\nu = 2 \qquad (n \in \mathbb{N}_0)$$

und für alle $\nu \in \mathbb{N}_0$ fest gilt

$$c_{n\nu} \to 0 \qquad (n \to \infty).$$

Da $((2z)^n/n!)$ eine Nullfolge ist, gilt $s_n - b_n \to 0$ $(n \to \infty)$ nach Beweisschritt 1. zu S. 7.20. □

Bemerkung 8.8 Aus S. 8.7 ergibt sich insbesondere

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

Hieraus folgt aus S. 8.5 (induktiv angewandt) und S. 8.6 auch

$$\exp(n) = e^n$$

für alle $n \in \mathbb{Z}$. Deshalb ist die Schreibweise e^z statt $\exp(z)$ für allgemeines $z \in \mathbb{C}$ konsistent mit der (alten) Definition ganzzahliger Potenzen. Wir werden diese im Weiteren meist verwenden.

Bemerkung und Definition 8.9 1. Die Funktion $\cos : \mathbb{C} \to \mathbb{C}$, gegeben durch

$$\cos(z) := \frac{1}{2}(e^{iz} + e^{-iz}) \qquad (z \in \mathbb{C})\ ,$$

heißt *(komplexe) Cosinusfunktion.*
2. Die Funktion $\sin : \mathbb{C} \to \mathbb{C}$, gegeben durch

$$\sin(z) := \frac{1}{2i}(e^{iz} - e^{-iz}) \qquad (z \in \mathbb{C})\ ,$$

heißt *(komplexe) Sinusfunktion.*
Mit B. 8.2 und der Stetigkeit von exp ergibt sich leicht die Stetigkeit von cos und sin. Weiter ergibt sich aus der Definition sofort $\cos(0) = 1$, $\sin(0) = 0$ und

$$\cos(-z) = \cos z\ , \qquad \sin(-z) = -\sin z$$

sowie ([Ü])

$$\cos(z) = \sum_{\substack{\nu=0 \\ \nu \text{ gerade}}}^{\infty} \frac{(-1)^{\nu/2} z^\nu}{\nu!} = \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k}}{(2k)!}$$

und

$$\sin(z) = \sum_{\substack{\nu=0 \\ \nu \text{ ungerade}}}^{\infty} \frac{(-1)^{(\nu-1)/2} z^\nu}{\nu!} = \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k+1}}{(2k+1)!}$$

(mit absoluter Konvergenz der Reihen). Hieraus folgt auch, dass $\sin(x)$ und $\cos(x)$ reell sind für reelle x.
Schließlich sieht man leicht, dass für alle $z \in \mathbb{C}$ die *Eulersche Formel*

$$e^{iz} = \cos z + i \sin z \qquad (z \in \mathbb{C}) \tag{8.3}$$

erfüllt ist. Für reelle x ist damit

$$\cos x = \operatorname{Re}(e^{ix}) \quad \text{und} \quad \sin x = \operatorname{Im}(e^{ix})\ .$$

Satz 8.10 *(Additionstheoreme)*
Für $z, w \in \mathbb{C}$ gilt

1. $\cos(z+w) = \cos z \cos w - \sin z \sin w$,
2. $\sin(z+w) = \sin z \cos w + \cos z \sin w$.

Beweis. 1. Für $z, w \in \mathbb{C}$ gilt

$$\begin{aligned}\cos(z)\cos(w) - \sin(z)\sin(w) &= \frac{1}{4}\big((e^{iz}+e^{-iz})(e^{iw}+e^{-iw}) + (e^{iz}-e^{-iz})(e^{iw}-e^{-iw})\big) \\ &= \frac{1}{2}\big(e^{iz}e^{iw}+e^{-iz}e^{-iw}\big) = \frac{1}{2}\big(e^{i(z+w)}+e^{-i(z+w)}\big) \\ &= \cos(z+w)\end{aligned}$$

2. ergibt sich durch eine entsprechende Rechnung. □

Wir schauen uns nun trigonometrischen Funktionen für reelle Argumente an. In diesem Zusammenhang definieren wird eine der wichtigsten Konstanten der Mathematik, die Kreiszahl π. Die Existenz wird eine Konsequenz aus dem folgenden zentralen Ergebnis über stetige Funktionen auf Intervallen sein, das wiederum eine direkte Folgerung aus der Vollständigkeit von $\mathbb{R}$ ist.

Satz 8.11 *(Zwischenwertsatz)*
Es sei $I \neq \emptyset$ ein Intervall in $\mathbb{R}$ und $f : I \to \mathbb{R}$ sei stetig auf I. Dann gilt

1. *Für alle $\underline{y}, \overline{y} \in W(f)(= f(I))$ mit $\underline{y} \leq \overline{y}$ ist $[\underline{y}, \overline{y}] \subset W(f)$.*
2. *$W(f)$ ist ein Intervall.*

Beweis. 1. Es seien $\underline{y}, \overline{y} \in W(f)$, wobei o. E. $\underline{y} < \overline{y}$. Dann ist zu zeigen: Ist $\eta \in (\underline{y}, \overline{y})$ so existiert ein $\xi \in I$ mit $f(\xi) = \eta$.
Zunächst existieren $\underline{x}, \overline{x} \in I$ mit $f(\underline{x}) = \underline{y}$ und $f(\overline{x}) = \overline{y}$. O. E. sei $\underline{x} < \overline{x}$. Wir setzen

$$M := \{x \in [\underline{x}, \overline{x}] : f(x) \leq \eta\} .$$

Dann ist $M \neq \emptyset$ (da $\underline{x} \in M$) und beschränkt, also existiert

$$\xi := \sup M \in [\underline{x}, \overline{x}] \subset I .$$

Behauptung: Es gilt $f(\xi) = \eta$.

Denn: Es existiert eine Folge (x_n) in M mit $x_n \to \xi$. Da f stetig an $\xi \in I$ ist, gilt $f(x_n) \to f(\xi)$ also $f(\xi) \leq \eta$. Aus $f(\overline{x}) = \overline{y} > \eta$ folgt $\xi < \overline{x}$. Ist $(\tilde{x}_n)$ eine Folge in $(\xi, \overline{x}]$ 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 $f(\xi) = \eta$.
2. Man sieht leicht: $J \subset \mathbb{R}$ ist genau dann ein Intervall, wenn für alle $\underline{y}, \overline{y} \in J$ mit $\underline{y} \leq \overline{y}$ auch $[\underline{y}, \overline{y}] \subset J$ gilt. □

Bemerkung 8.12 Ist $d \in \mathbb{N}$ und $f(x) = x^d$ für $x \geq 0$, so folgt aus $f(0) = 0$ und $f(n) = n^d \to \infty$ $(n \to \infty)$ nach dem Zwischenwertsatz

$$f([0, \infty)) = [0, \infty)\ .$$

Damit hat für jedes $c \geq 0$ die Gleichung $x^d = c$ eine Lösung. Wir erhalten also aus dem Zwischenwertsatz noch einmal (jetzt völlig schmerzlos) die Existenz von Wurzeln positiver Zahlen.

Satz 8.13 *1. Für $x \in \mathbb{R}$ ist*

$$|e^{ix}| = 1$$

also auch $\cos^2 x + \sin^2 x = |e^{ix}|^2 = 1$ und damit insbesondere $-1 \leq \cos x \leq 1$ und $-1 \leq \sin x \leq 1$.

2. Es existiert ein $x \in (0, 2)$ mit $e^{ix} = i$.

Beweis. Zunächst gilt für beliebiges $z \in \mathbb{C}$ und $s_n(z) = \sum\limits_{\nu=0}^{n} \frac{z^\nu}{\nu!}$

$$\overline{s_n(z)} = \overline{\sum_{\nu=0}^{n} \frac{z^\nu}{\nu!}} = \sum_{\nu=0}^{n} \frac{\overline{z}^\nu}{\nu!} = s_n(\overline{z}) \to e^{\overline{z}} \qquad (n \to \infty)\ .$$

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 $z = ix$

$$\cos^2 x + \sin^2 x = |e^{ix}|^2 = e^{ix}\overline{e^{ix}} = e^{ix}e^{\overline{ix}} = e^{ix}e^{-ix} \overset{S.8.6}{=} 1\ .$$

2. (i) Zunächst gilt $\sin x \geq 0$ für alle $x \in [0, 2]$.
Denn: Ist

$$a_\nu = \frac{x^{2\nu+1}}{(2\nu+1)!}\ ,$$

so gilt für $\nu \in \mathbb{N}$
$$\frac{a_\nu}{a_{\nu-1}} = \frac{x^2}{(2\nu+1)(2\nu)} \leq \frac{4}{6} < 1$$
also ist $0 \leq a_\nu \downarrow$. Aus dem Leibniz-Kriterium (S. 7.8.1) ergibt sich
$$\sin x = \sum_{\nu=0}^{\infty} (-1)^\nu \frac{x^{2\nu+1}}{(2\nu+1)!} = \lim_{m\to\infty} s_{2m} \geq 0\ .$$

(ii) Setzen wir nun
$$a_\nu = \frac{4^{\nu+2}}{(2\nu+4)!} \qquad (\nu \in \mathbb{N}_0),$$
so gilt
$$\frac{a_\nu}{a_{\nu-1}} = \frac{4}{(2\nu+3)(2\nu+4)} < 1,$$
also $0 \leq a_\nu \downarrow$. Damit ist wieder nach dem Leibniz-Kriterium (S. 7.8.1)
$$\sum_{k=2}^{\infty} \frac{(-1)^k}{(2k)!} 2^{2k} = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu 4^{\nu+2}}{(2\nu+4)!} = \lim_{m\to\infty} s_{2m} \leq s_0 = a_0 = \frac{4^2}{4!} = \frac{2}{3}$$
und folglich
$$\cos(2) = 1 - \frac{2^2}{2!} + \sum_{k=2}^{\infty} \frac{(-1)^k 2^{2k}}{(2k)!} \leq -1 + \frac{2}{3} < 0.$$
Da cos stetig ist, existiert nach dem Zwischenwertsatz ein $x \in (0,2)$ mit $\cos x = 0$. Weiter ist $\sin x \geq 0$ nach (i) und $\sin^2 x = 1$ nach 1. Folglich ist $\sin x = 1$ und damit auch $e^{ix} = \cos x + i \sin x = i$.

□

Bemerkung und Definition 8.14 Wir setzen $M := \{x > 0 : e^{ix} = i\}$. Dann ist $M \neq \emptyset$ wie eben gesehen. Ist $x_0 := \inf M$, so ist aufgrund der Stetigkeit der Exponentialfunktion auch $e^{ix_0} = i$ (d. h. $x_0 = \min M$). Damit schreiben wir
$$\pi := 2x_0\ .$$
Nach dieser Definition gilt also
$$\cos\left(\frac{\pi}{2}\right) = 0, \qquad \sin\left(\frac{\pi}{2}\right) = 1$$
und mit $\cos(0) = 1$, dem Zwischenwertsatz und $\cos(x) = \cos(-x)$ folgt $\cos(x) > 0$ für $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.

Unter Ausnutzung der Additionstheoreme erhält man Periodizitätseigenschaften der trigonometrischen Funktionen

Satz 8.15 *Für alle $z \in \mathbb{C}$ gilt*

1. $\cos\left(z + \frac{\pi}{2}\right) = -\sin z$, $\qquad \sin\left(z + \frac{\pi}{2}\right) = \cos z$,
2. $\cos(z + \pi) = -\cos z$, $\qquad \sin(z + \pi) = -\sin z$,
3. $\cos(z + 2\pi) = \cos z$, $\qquad \sin(z + 2\pi) = \sin z$,

   ($\cos$ *und* $\sin$ *sind* 2π*-periodisch)*

Beweis. Mit S. 8.10.1 erhalten wir

$$\cos(z + \pi/2) = \cos(z)\cos(\pi/2) - \sin(z)\sin(\pi/2) = -\sin z \ .$$

und

$$\sin(z + \pi/2) = \cos(z)\sin(\pi/2) + \sin(z)\cos(\pi/2) = \cos z \ .$$

Hieraus folgt wiederum

$$\cos(z + \pi) = \cos(z + \pi/2 + \pi/2) = -\sin(z + \pi/2) = -\cos z \ .$$

Die weiteren Behauptungen ergeben sich in ähnlicher Weise ([Ü]). □

9 Monotonie und Umkehrfunktionen

Definition 9.1 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.

Satz 9.2 *1. Für alle $x \in \mathbb{R}$ ist $e^x \geq \max(1+x, 0)$.*

2. Für alle $x < 1$ ist $e^x \leq \frac{1}{1-x}$.

3. $\exp : \mathbb{R} \to \mathbb{R}$ ist streng monoton wachsend mit $\exp(\mathbb{R}) = W(\exp_{|\mathbb{R}}) = (0, \infty)$.

Beweis. 1. Aus $(1 + x/n)^n \geq 0$ für n genügend groß folgt $e^x \geq 0$. Ist $x \geq -1$ so ergibt sich mit der Bernoullischen Ungleichung $(1+x/n)^n \geq 1+x$ für alle n, also auch $e^x \geq 1 + x$.
2. Nach 1. gilt für $x < 1$
$$e^{-x} \geq 1 - x \ ,$$
also auch
$$e^x = \frac{1}{e^{-x}} \leq \frac{1}{1-x} \ .$$
3. Für $x_1 < x_2$ ergibt sich $e^{x_2}/e^{x_1} = e^{x_2 - x_1} \geq 1 + (x_2 - x_1) > 1$ nach 1. Also ist $\exp : \mathbb{R} \to \mathbb{R}$ streng monoton wachsend. Weiter folgt aus 1. und 2. auch
$$e^n \to \infty \qquad \text{und} \qquad e^{-n} \to 0 \quad (n \to \infty)$$
und damit ist $\exp(\mathbb{R}) = (0, \infty)$ nach dem Zwischenwertsatz. □

Satz 9.3 *1. $f := \sin_{|[-\pi/2, \pi/2]}$ ist streng monoton wachsend mit $W(f) = [-1, 1]$.*

2. $g := \cos_{|[0,\pi]}$ ist streng monoton fallend mit $W(g) = [-1, 1]$.

Beweis. 1. Aus den Additionstheoremen ergibt sich für $z, w \in \mathbb{C}$
$$\sin(z + w) - \sin(z - w) = 2\cos(z)\sin(w) \ .$$

Also folgt für $x_1, x_2 \in \mathbb{R}$

$$\begin{aligned}
\sin x_2 - \sin x_1 &= \sin\left(\frac{x_2+x_1}{2} + \frac{x_2-x_1}{2}\right) - \sin\left(\frac{x_2+x_1}{2} - \frac{x_2-x_1}{2}\right) \\
&= 2 \cdot \cos\left(\frac{x_2+x_1}{2}\right) \cdot \sin\left(\frac{x_2-x_1}{2}\right) .
\end{aligned}$$

Ist $-\pi/2 \le x_1 < x_2 \le \pi/2$, so gilt

$$\frac{x_2+x_1}{2} \in (-\pi/2, \pi/2) \quad \text{und} \quad \frac{x_2-x_1}{2} \in (0, \pi/2] \subset (0, \pi)$$

und damit $\cos\left(\frac{x_2+x_1}{2}\right) > 0$ und $\sin\left(\frac{x_2-x_1}{2}\right) > 0$, also $\sin x_1 < \sin x_2$.
Schließlich folgt aus $\sin(0) = 0$ und $\sin(\pi/2) = 1$ sowie $\sin(-x) = -\sin x$ mit dem Zwischenwertsatz $\sin([-\pi/2, \pi/2]) = [-1, 1]$.
2. Ergibt sich aus 1. und S. 8.15. □

Also Folgerung erhalten wir

Satz 9.4 *Für $z \in \mathbb{C}$ gilt*

1. $e^z = 1$ *genau dann, wenn* $z = 2k\pi i$ *für ein* $k \in \mathbb{Z}$,
2. $\sin z = 0$ *genau dann, wenn* $z = k\pi$ *für ein* $k \in \mathbb{Z}$,
3. $\cos z = 0$ *genau dann, wenn* $z = k\pi + \pi/2$ *für ein* $k \in \mathbb{Z}$,

Beweis. 1. [Ü]
2. Es gilt $0 = 2i \sin z = e^{iz} - e^{-iz}$ genau dann, wenn $e^{2iz} - 1 = 0$ ist. Aus 1. ergibt sich damit 2.
3. Mit 2. und S. 8.15. □

Bemerkung und Definition 9.5 Wir definieren die Funktionen

$$\tan : \mathbb{C} \setminus \{\pi/2 + k\pi : k \in \mathbb{Z}\} \to \mathbb{C}$$

und

$$\cot : \mathbb{C} \setminus \{k\pi : k \in \mathbb{Z}\} \to \mathbb{C}$$

durch

$$\tan z = \frac{\sin z}{\cos z} , \qquad \cot z = \frac{\cos z}{\sin z} .$$

Dann sind tan und cot stetig auf ihren jeweiligen Definitionsbereichen.

Wir befassen uns nun mit der Umkehrbarkeit der elementaren Funktionen. Dazu beweisen wir zunächst folgendes allgemeine Ergebnis.

Satz 9.6 *Es sei $I \subset \mathbb{R}$ ein Intervall und es sei $f : I \to W(f) \subset \mathbb{R}$ streng monoton wachsend (bzw. fallend). Dann existiert die Umkehrfunktion $f^{-1} : W(f) \to I$ und f^{-1} ist ebenfalls streng monoton wachsend (bzw. fallend). Außerdem ist f^{-1} stetig.*

Beweis. Es sei $J := W(f)$. Aus D. 9.1 sieht man leicht, dass $f : I \to J$ bijektiv ist, d. h. $f^{-1} : J \to I$ existiert. 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 seien $y_0 \in J$ und $x_0 := f^{-1}(y_0)$ sowie (y_n) eine Folge in J mit $y_n \to y_0$. Weiter sei $\varepsilon > 0$.
Ist x_0 nicht der rechte Randpunkt von I (d. h. $x_0 \neq \sup I$), so existiert ein $x_\varepsilon \in I$ mit $x_0 < x_\varepsilon \leq x_0 + \varepsilon$. Wir setzen

$$\delta_\varepsilon^+ := f(x_\varepsilon) - f(x_0) \ .$$

Dann ist $\delta_\varepsilon^+ > 0$ und für alle $y \in J$ mit $y_0 \leq y < y_0 + \delta_\varepsilon^+ = f(x_\varepsilon)$ (falls existent) folgt

$$0 \leq f^{-1}(y) - f^{-1}(y_0) < f^{-1}(y_0 + \delta_\varepsilon^+) - f^{-1}(y_0) = x_\varepsilon - x_0 \leq \varepsilon \ .$$

Ist x_0 nicht der linke Randpunkt von I, so sieht man entsprechend: Es existiert ein $\delta_\varepsilon^- > 0$ so, dass

$$0 \leq f^{-1}(y_0) - f^{-1}(y) < \varepsilon \quad \text{für alle } y \in J \text{ mit } y_0 - \delta_\varepsilon^- < y \leq y_0 \ .$$

Damit ergibt sich $|f(y) - f(y_0)| < \varepsilon$ für alle $y \in J$ mit $|y - y_0| < \delta_\varepsilon := \min\{\delta_\varepsilon^+, \delta_\varepsilon^-\}$. Aus $|y_n - y_0| < \delta_\varepsilon$ für n genügend groß folgt $|f(y_n) - f(y_0)| < \varepsilon$ für n genügend groß. □

Bemerkung und Definition 9.7 Nach S. 9.2.3 und S. 9.6 existiert die Umkehrfunktion von exp auf dem Intervall $(0, \infty)$ und ist dort stetig und streng monoton wachsend. Diese Funktion nennnen wir *(natürliche) Logarithmusfunktion* und schreiben dafür ln oder auch log. Es gilt also für $x \in (0, \infty)$ und $y \in \mathbb{R}$

$$y = \ln(x) \iff e^y = x \ .$$

Aus den entsprechenden Eigenschaften der Exponentialfunktion ergibt sich leicht ([Ü]):

1. Für alle $x_1, x_2 > 0$ ist $\ln(x_1 x_2) = \ln(x_1) + \ln(x_2)$.
2. Für alle $x > 0$ und alle $m \in \mathbb{Z}$ ist $\ln(x^m) = m \ln(x)$.

Damit ist es möglich, allgemeine Potenzen und Logarithmen zu definieren.

Definition 9.8 Für $a > 0$ und $m \in \mathbb{Z}$ ist

$$a^m = e^{\ln(a^m)} = e^{m \ln a}$$

nach B./D. 9.7.2. Wir setzen für allgemeines $b \in \mathbb{C}$

$$a^b := \exp(b \cdot \ln a) = e^{b \cdot \ln a} \ .$$

Aus den Rechenregeln für ln und exp erhält man

Satz 9.9 *(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 im Falle $b_1 \in \mathbb{R}$ auch $(a^{b_1})^{b_2} = a^{b_1 b_2}$.*

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

$$a^{b_1} a^{b_2} = e^{b_1 \ln a} e^{b_2 \ln a} = e^{b_1 \ln a + b_2 \ln a} = e^{(b_1 + b_2) \ln a} = a^{b_1 + b_2} \ .$$

Im Falle $b_1 \in \mathbb{R}$ ist $a^{b_1} > 0$ und damit gilt dann auch

$$(a^{b_1})^{b_2} = e^{b_2 \ln(a^{b_1})} = e^{b_2 \ln(e^{b_1 \ln a})} = e^{b_2 b_1 \ln a} = a^{b_1 b_2} \ .$$

2. [Ü]. □

Bemerkung und Definition 9.10 Wir betrachten ein festes $a > 0, a \neq 1$. Dann gilt für $x > 0$ und $y \in \mathbb{R}$

$$x = a^y = e^{y \ln a} \Longleftrightarrow y \ln a = \ln x \Longleftrightarrow y = \frac{\ln x}{\ln a} \ .$$

Wir definieren

$$\log_a x := \frac{\ln x}{\ln a} \qquad (x > 0) .$$

Damit gilt also für $x > 0, y \in \mathbb{R}$

$$y = \log_a x \Longleftrightarrow a^y = x .$$

Bemerkung und Definition 9.11 Die nach S. 9.6 und S. 9.3 auf $\sin([-\pi/2, \pi/2]) = [-1, 1]$ existierende (und dort streng monoton wachsende und stetige) Umkehrfunktion von sin heißt arcsin, d. h. $\arcsin : [-1, 1] \to [-\pi/2, \pi/2]$ erfüllt für $x \in [-1, 1]$, $y \in [-\pi/2, \pi/2]$

$$y = \arcsin x \Leftrightarrow x = \sin y .$$

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 für $x \in [-1, 1]$, $y \in [0, \pi]$

$$y = \arccos x \Leftrightarrow x = \cos y .$$

Außerdem gilt: tan ist streng monoton wachsend in $(-\pi/2, \pi/2)$ und cot ist streng monoton fallend in $(0, \pi)$ mit $W(\tan_{|(-\pi/2,\pi/2)}) = W(\cot_{|(0,\pi)}) = \mathbb{R}$. Also existieren auf $\mathbb{R}$ die (stetigen) Umkehrfunktionen, genannt arctan bzw. arccot, mit entsprechenden Monotonieeigenschaften.

Bemerkung 9.12 (Polarkoordinaten)
1. Für alle $\varphi \in \mathbb{R}$ ist $|e^{i\varphi}| = 1$. Umgekehrt gilt auch: Ist $z = x + iy \in \mathbb{C}$ mit $|z| = 1$, so existiert ein $\varphi \in [-\pi, \pi)$ mit

$$z = e^{i\varphi}$$

oder, anders ausgedrückt,

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

(Denn:
1. Fall: $y > 0$. Dann betrachten wir $\varphi := \arccos(x) \in (0, \pi)$. Dafür gilt $x = \cos\varphi$ und

$$y^2 = 1 - x^2 = 1 - \cos^2\varphi = \sin^2\varphi .$$

Da $\varphi \in (0, \pi)$ ist, ist $\sin\varphi > 0$. Damit ist $y = \sin\varphi$.
2. Fall: $y \le 0$. Dann gilt für $\varphi := -\arccos(x) \in [-\pi, 0]$ genauso $x = \cos(-\varphi) = \cos\varphi$ und $y^2 = \sin^2\varphi$. Da jetzt $\sin\varphi \le 0$ ist, folgt wieder $y = \sin\varphi$.)

2. Wir betrachten $f : (0,\infty) \times \mathbb{R} \to \mathbb{C}$ mit

$$f(r,\varphi) := re^{i\varphi} = r\cos(\varphi) + ir\sin(\varphi) \qquad (r > 0, \varphi \in \mathbb{R}).$$

Dann ergibt sich aus S. 9.4.1

$$f(r,\varphi) = e^{\ln r + i\varphi} = e^{\ln \tilde{r} + i\tilde{\varphi}} = f(\tilde{r},\tilde{\varphi})$$

genau dann, wenn $r = \tilde{r}$ und $\varphi = \tilde{\varphi} + 2k\pi$ für ein $k \in \mathbb{Z}$.
Damit erhalten wir:
Ist $\alpha \in \mathbb{R}$, so ist $f_\alpha := f_{|(0,\infty)\times[\alpha,\alpha+2\pi)}$ injektiv. Außerdem ist nach 1. und der 2π-Periodizität von $\varphi \mapsto e^{i\varphi}$ für alle $r > 0$

$$f_\alpha(\{r\} \times [\alpha, \alpha + 2\pi)) = \{z \in \mathbb{C} : |z| = r\}$$

und damit auch

$$W(f_\alpha) = \mathbb{C} \setminus \{0\}.$$

Bemerkung und Definition 9.13 Eine wichtige Folgerung aus der Polarkoordinatendarstellung komplexer Zahlen ist die Existenz von Wurzeln komplexer Zahlen: Es sei $c \in \mathbb{C} \setminus \{0\}$ beliebig. Dann existieren $r > 0$ und $\varphi \in \mathbb{R}$ mit $c = re^{i\varphi}$. Für $n \in \mathbb{N}$ gilt

$$z^n = c$$

genau dann, wenn

$$z = z_k = \sqrt[n]{r}\, e^{i(\varphi + 2k\pi)/n} \qquad \text{für ein } k \in \mathbb{Z}.$$

(Denn: Da $\varphi \mapsto e^{i\varphi}$ 2π-perodisch ist, gilt einerseits

$$z_k^n = r \cdot e^{i(\varphi+2k\pi)} = re^{i\varphi} = c \qquad (k \in \mathbb{Z}).$$

Andererseits folgt für $z = \rho e^{i\psi}$ mit $z^n = c$

$$\rho^n e^{in\psi} = z^n = re^{i\varphi}$$

und damit $\rho = \sqrt[n]{r}$ und $\psi = \psi_k = (\varphi + 2k\pi)/n$ für ein $k \in \mathbb{Z}$.)
Die Zahlen z_k heißen *n-te (komplexe) Wurzeln* aus c. Man beachte, dass nur n davon paarweise verschieden sind (etwa $z_0, \ldots, z_{n-1}$).
Ist speziell $c = 1$, so heißen die n Zahlen

$$z_k = e^{2k\pi i/n} \qquad (k = 0, \ldots, n-1)$$

n-te Einheitswurzeln. So sind etwa ± 1 die zweiten Einheitswurzeln und $\pm i, \pm 1$ die vierten Einheitswurzeln.

10 Normierte und metrische Räume

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 in diesem Abschnitt definieren, die sog. metrischen Räume. Es wird sich später zeigen, dass diese Räume für viele Fragen der Analysis den geeigneten Rahmen bilden.
Wir betrachten zunächst jedoch eine speziellere Klasse von Räumen, wobei wir damit gleichzeitig erstmals eine Verbindung zur Linearen Algebra herstellen.

Definition 10.1 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.

Beispiel 10.2 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 $\mathbb{C}^m$. Dann sind durch

$$||x||_1 := \sum_{j=1}^{m} |x_j|$$

$$||x||_\infty := \max\{|x_j| : j = 1, \ldots, m\} ,$$

wobei $x = (x_1, \ldots, x_m) \in \mathbb{K}^m$, Normen auf $\mathbb{K}^m$ gegeben.

Wir wollen weitere Normen auf $\mathbb{K}^m$ einführen. Dazu setzen wir noch $0^p := 0$ für $p > 0$.

Satz 10.3 *Es sei $p \in (1, \infty)$. Dann ist durch*

$$\|x\|_p := \Big(\sum_{j=1}^m |x_j|^p \Big)^{1/p} \qquad \big(x = (x_1, \dots, x_m) \in \mathbb{K}^m\big)$$

eine Norm auf $\mathbb{K}^m$ gegeben.

Beweis. (N.1) und (N.2) sind klar. (N.3) ergibt sich aus der Minkowski-Ungleichung (S. C.2). □

Bemerkung 10.4 Besonders wichtig ist der Fall $p = 2$, also

$$||x||_2 = \sqrt{\sum_{j=1}^m |x_j|^2} \quad (x \in \mathbb{K}^m)$$

(für $\mathbb{K} = \mathbb{R}$ spricht man von der *euklidschen Länge* von x). Aus

$$|x_{j_0}|^2 \le \sum_{j=1}^m |x_j|^2 \le \sum_{j,k=1}^m |x_j x_k| = \Big(\sum_{j=1}^m |x_j| \Big)^2$$

für alle $x \in \mathbb{K}^m$ und $j_0 \in \{1, \dots, m\}$ folgt

$$||x||_\infty \le ||x||_2 \le ||x||_1 \ .$$

Falls nichts anderes angegeben ist, soll im weiteren Verlauf der Vorlesung $\mathbb{K}^m$ stets mit der Norm $|| \cdot ||_2$ versehen sein. Wir schreiben hierfür auch kurz

$$| \cdot | := || \cdot ||_2 \ .$$

Definition 10.5 Es sei $(V, || \cdot ||)$ ein normierter Raum. Eine Menge $W \subset V$ heißt *beschränkt*, falls ein $C > 0$ existiert mit $||x|| \le C$ für alle $x \in W$. Sind X eine Menge, $f : X \to V$ und $M \subset X$, so heißt *beschränkt auf* M, falls $f(M) \subset V$ beschränkt ist. Im Falle $M = X$ sagen wir kurz, f sei *beschränkt*.

Beispiel 10.6 Es sei $(V, ||\cdot||)$ ein normierter Raum. Dann sind

$$S := S_V := \{x \in V : ||x|| = 1\} \qquad \text{und} \qquad B := B_V := \{x \in V : ||x|| \leq 1\}$$

(die sog. Einheitssphäre bzw. Einheitskugel in V) beschränkt.

Bemerkung und Definition 10.7 1. Es sei E ein Vektorraum (über $\mathbb{K}$). Ist M eine nichtleere Menge, so sind für $f, g : M \to E$ und $\lambda \in \mathbb{K}$ die Funktionen $f + g : M \to E$ und $\lambda f : M \to E$ definert durch

$$(f+g)(t) := f(t) + g(t), \quad (\lambda f)(t) := \lambda \cdot f(t) \quad (t \in M)\ .$$

Damit ist auch $E^M = \{f : M \to E\}$ ein Vektorraum über $\mathbb{K}$ (siehe Lineare Algebra).
2. Wir setzen für $m \in \mathbb{N}$

$$B(M, \mathbb{K}^m) := \{f : M \to \mathbb{K}^m : f \text{ beschränkt}\}\ .$$

Wie man leicht sieht, ist $B(M, \mathbb{K}^m)$ ein Unterraum von $\{f : M \to \mathbb{K}^m\}$ und damit selbst ein Vektorraum. Weiter definieren wir für $f, g \in B(M, \mathbb{K}^m)$

$$||f||_\infty := \sup\{|f(t)| : t \in M\}\ .$$

Dann ist $||\cdot||_\infty$ eine Norm auf $B(M, \mathbb{K}^m)$ ([Ü]).

Definition 10.8 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.*

Bemerkung 10.9 Es sei $X \neq \emptyset$ eine beliebige Menge. Dann ist durch

$$\delta(x,y) := \begin{cases} 0, & \text{falls} \quad x = y \\ 1, & \text{falls} \quad x \neq y \end{cases}$$

eine Metrik auf X gegeben (die sog. *diskrete Metrik*). Dies ergibt sich leicht durch Überprüfen von (d.1)–(d.3).

Satz 10.10 *Ist* $(V, ||\cdot||)$ *ein normierter Raum, so ist durch*

$$d_{||\cdot||}(x,y) := ||x-y|| \qquad (x,y \in V)$$

eine Metrik auf V *gegeben (die sog. induzierte Metrik).*

Beweis. (d.1) ergibt sich unmittelbar aus (N.1).
Zu (d.2): Sind $x, y \in V$, so gilt mit (N.2)

$$d(x,y) = ||x-y|| = ||-(y-x)|| = ||y-x|| = d(y,x) \; .$$

Zu (d.3): Sind $x, y, z \in V$, so gilt mit (N.3)

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

□

Beispiel 10.11 1. Ist (X,d) ein metrischer Raum und ist $M \subset X$, so ist auch (M,d) (d.h. genauer $(M, d_{|M\times M})$) ein metrischer Räum.
2. Es sei $M \subset \mathbb{K}^m$. Dann ist nach S. 10.10 und 1. durch

$$d_{|\cdot|}(x,y) := d_{||\cdot||_2}(x,y) := |x-y| \qquad (x,y \in M)$$

eine Metrik auf M gegeben (die sog. *euklidsche Metrik*). **Im Weiteren soll** $M \subset \mathbb{K}^m$ **stets mit dieser Metrik versehen sein, wenn nichts anderes explizit gesagt ist**.

Wir untersuchen nun Folgen in metrischen Räumen

Definition 10.12 Es sei (X,d) ein metrischer Raum, und es sei (x_n) eine Folge in X.

1. (x_n) heißt *konvergent* (in (X,d)) falls ein $x \in X$ so existiert, dass

   $$d(x_n, x) \to 0 \qquad (n \to \infty) \; .$$

   Dann heißt wieder x *Grenzwert* von (x_n) und wir schreiben

   $$x_n \to x \quad (n \to \infty)$$

   (in (X,d)).

2. (x_n) heißt *Cauchy-Folge* (in (X, d)), falls zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ existiert mit
$$d(x_n, x_m) < \varepsilon \qquad (n, m \geq N_\varepsilon)\ .$$

Ist $(X, d) = (V, d_{||\cdot||})$ so sprechen wir auch von Konvergenz bzw. von einer Cauchy-Folge in V (oder genauer in $(V, ||\cdot||)$).

Bemerkung 10.13 Es sei (X, d) ein metrischer Raum. Wie im Fall $(X, d) = (\mathbb{K}, d_{|\cdot|})$ gilt

1. Jede Folge in X hat höchstens einen Grenzwert x. Wir schreiben wieder
$$x = \lim_{n\to\infty} x_n.$$
2. Jede konvergente Folge ist eine Cauchy-Folge.
3. Jede Cauchy-Folge, die eine konvergente Teilfolge besitzt, ist konvergent.

(Denn: 1. Es seien $x, \tilde{x}$ Grenzwerte von (x_n). Dann gilt mit (d.2) und (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 $x = \tilde{x}$.
2. Wie im Beweis zu S. 5.4 mit der Dreiecksungleichung (d.3).
3. Wie in B./D. 6.11.)

Beispiel 10.14 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 (x_n) konvergiert **nicht** im metrischen Raum $(\mathbb{R}, \delta)$, wobei δ die diskrete Metrik ist. (Im metrischen Raum (X, δ) gilt: $x_n \to x$ genau dann, wenn $x_n = x$ für n genügend groß; [Ü].)
Das Beispiel zeigt insbesondere, dass die Konvergenz einer Folge von der zu Grunde liegenden Metrik abhängen kann!

Bemerkung 10.15 Ist (X, d) ein metrischer Raum, so ist i. A. nicht jede Cauchy-Folge konvergent!
Betrachtet man etwa $X = \mathbb{R}_+$ (also die positiven reellen Zahlen) mit der Betragsmetrik $d_{|\cdot|}$, so ist $(x_n) = (1/n)$ eine Cauchy-Folge in $\mathbb{R}_+$, die nicht konvergiert.
(Denn: Offensichtlich ist $(1/n)$ eine Cauchy-Folge in $(\mathbb{R}_+, d_{|\cdot|})$. Da jedoch $1/n \to 0$ in $\mathbb{R}$ gilt, kann die Folge in $(\mathbb{R}_+, d_{|\cdot|})$ nicht konvergent sein. (Die Konvergenz würde der Eindeutigkeit des Grenzwertes in $\mathbb{R}$ widersprechen.))

Bemerkung 10.16 1. Es sei $\big(x^{(n)}\big)_n = \big((x_1^{(n)}, \ldots, x_m^{(n)})\big)_n$ eine Folge in $\mathbb{K}^m$.
Dann gilt: $(x^{(n)})$ ist konvergent (bzw. eine Cauchy-Folge) in $\mathbb{K}^m$ genau dann, wenn für alle $j \in \{1, \ldots, m\}$ die sogenanten *Komponentenfolgen* $(x_j^{(n)})_n$ konvergent (bzw. Cauchy-Folgen) in $\mathbb{K}$ sind. Außerdem gilt dann

$$\lim_{n\to\infty} x^{(n)} = \big(\lim_{n\to\infty} x_1^{(n)}, \ldots, \lim_{n\to\infty} x_m^{(n)}\big).$$

(Denn: Zur Konvergenz:
„$\Rightarrow$“ Ist $x = \lim x^{(n)}$, so folgt für $j = 1, \ldots, m$ aus B. 10.4 mit $x := (x_1 \ldots, x_m)$

$$|x_j^{(n)} - x_j| \leq \|x^{(n)} - x\|_\infty \leq \|x^{(n)} - x\|_2 \to 0 \qquad (n \to \infty).$$

„$\Leftarrow$“ Ist $x_j := \lim\limits_{n\to\infty} x_j^{(n)}$ für $j = 1, \ldots, m$, so folgt für $x = (x_1, \ldots, x_m)$ wieder mit B. 10.4

$$\|x^{(n)} - x\|_2 \leq \|x^{(n)} - x\|_1 = \sum_{j=1}^m |x_j^{(n)} - x_j| \to 0 \qquad (n \to \infty).$$

Entsprechend ergibt sich die Äquivalenz für Cauchy-Folgen.)
2. Ist $(V, \|\cdot\|)$ ein normierter Raum, so gilt (wie in $V = \mathbb{K}$): Ist (x_n) konvergent, so ist (x_n) auch beschränkt.
(Denn: Für $x = \lim x_n$ ist

$$\|x_n\| \leq \|x\| + \|x_n - x\| \leq \|x\| + 1$$

für n genügend groß.)

Wir zeigen folgende Verallgemeinerung des Satzes von Bolzano-Weierstraß.

Satz 10.17 *(Bolzano-Weierstraß in $\mathbb{K}^m$)*
Es sei $m \in \mathbb{N}$. Dann hat jede beschränkte Folge in $\mathbb{K}^m$ eine konvergente Teilfolge.

Beweis. Wir beweisen die Behauptung per Induktion nach m.
$m = 1$ ist der Satz von Bolzano-Weierstraß in $\mathbb{K}$ (S. 6.14).
$m \to m+1$: Es sei $(x^{(n)})_n$ eine bechränkte Folge in $\mathbb{K}^{m+1}$. Wir schreiben

$$x^{(n)} = \left(x_1^{(n)}, \ldots, x_m^{(n)}, x_{m+1}^{(n)}\right) =: (y^{(n)}, z^{(n)})\,.$$

wobei $y^{(n)} \in \mathbb{K}^m$. Dann gilt $|y^{(n)}| \leq |x^{(n)}|$ und $|z^{(n)}| \leq |x^{(n)}|$. Insbesondere ist $\{y^{(n)} : n \in \mathbb{N}\}$ beschränkt. Also besitzt nach Induktionsvoraussetzung $(y^{(n)})$ eine konvergente Teilfolge $(y^{(n_j)})$. Da auch $(z^{(n_j)})$ beschränkt in $\mathbb{K}$ ist, existiert eine konvergente Teilfolge $(z^{(n_{j_\ell})})$ (wieder S. 6.14). Durch zweimalige Anwendung von B. 10.16 ergibt sich die Konvergenz von $(x^{(n_{j_\ell})}) = (y^{(n_{j_\ell})}, z^{(n_{j_\ell})})$ in $\mathbb{K}^{n+1}$ (beachte: alle Komponentenfolgen von $(y^{(n_{j_\ell})})$ sind konvergent.) □

Definition 10.18 Es sei (X, d) ein metrischer Raum. Eine Menge $M \subset X$ heißt *relativ kompakt*, falls jede Folge in M eine konvergente Teilfolge besitzt. M heißt *kompakt*, falls jede Folge in M eine konvergente Teilfolge mit Grenzwert in M besitzt.

Bemerkung 10.19 Ist $\big(V, \|\cdot\|\big)$ ein beliebiger normierter Raum, so ist jede relativ kompakte Teilmenge M beschränkt.
(Denn: Angenommen, M ist unbeschränkt. Dann existiert zu jedem $n \in \mathbb{N}$ ein $x_n \in M$ mit $\|x_n\| > n$. Da jede Teilfolge von (x_n) unbeschränkt ist, hat (x_n) keine konvergente Teilfolge. Widerspruch zu B. 10.16.2.)
Im Allgemeinen ist **nicht** jede beschränkte Menge in einem normierten Raum relativ kompakt ([Ü]). Nach dem Satz von Bolzano-Weierstraß gilt in Falle $V = \mathbb{K}^m$ jedoch:

Jede beschränkte Menge M in $\mathbb{K}^m$ ist relativ kompakt.

(Denn: Ist (x_n) eine Folge in M, so ist (x_n) beschränkt. Damit existiert nach S. 10.17 eine konvergente Teilfolge.)

Definition 10.20 Es sei (X, d) ein metrischer Raum, und es sei $x_0 \in X$. Für $\varepsilon > 0$ heißt die Menge

$$U_\varepsilon(x_0) := \{x \in X : d(x, x_0) < \varepsilon\}$$

ε-Umgebung von x_0.

Beispiel 10.21 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 $z_0 = x_0 + iy_0 \in \mathbb{C}$

$$\begin{aligned} U_\varepsilon(z_0) &= \{z \in \mathbb{C} : |z - z_0| < \varepsilon\} = \\ &= \{z = x + iy \in \mathbb{C} :\ (x - x_0)^2 + (y - y_0)^2 < \varepsilon^2\} \end{aligned}$$

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 $x^{(0)}$.

Definition 10.22 Es seien (X,d) ein metrischer Raum und $M \subset X$.

(i) *offen*, falls für jedes $x \in M$ ein $\varepsilon > 0$ existiert mit $U_\varepsilon(x) \subset M$,

(ii) *abgeschlossen*, falls $M^c = X \setminus M$ offen ist,

Beispiel 10.23 1. In jedem metrischen Raum (X,d) sind X und $\emptyset$ offen und abgeschlossen.
2. Es seien $a, b \in \mathbb{R}$ mit $a \leq b$. Dann sind die Intervalle (a,b), $(-\infty, b)$ und (a,∞) offen in $\mathbb{R}$ und die Intervalle $[a,b]$, $(-\infty, b]$ sowie $[a,\infty)$ abgeschlossen in $\mathbb{R}$.

Der folgende Satz liefert eine Charakterisierung der Abgeschlossenheit mittels Folgen.

Satz 10.24 *Sind (X,d) ein metrischer Raum und $M \subset X$, so ist M abgeschlossen genau dann, wenn für alle konvergenten Folgen in M auch der Grenzwert in M liegt.*

Beweis. „$\Rightarrow$“ Ist M abgeschlossen, so ist M^c offen. Es sei (x_n) eine Folge in M mit $x_n \to x\ (n \to \infty)$. Dann gilt für alle $\varepsilon > 0$

$$M \cap U_\varepsilon(x) \neq \emptyset$$

(da $x_n \in U_\varepsilon(x)$ für n genügend groß). Damit ist $x \notin M^c$, d.h. $x \in M$.
„$\Leftarrow$“ Angenommen, M^c ist nicht offen. Dann existiert ein $x \in M^c$ mit $U_{1/n}(x) \cap M \neq \emptyset$ für alle $n \in \mathbb{N}$. Ist $x_n \in U_{1/n}(x) \cap M$, so gilt $x_n \to x\ (n \to \infty)$. Widerspruch. $\square$

Bemerkung 10.25 Ist (X, d) ein metrischer Raum, so ist $M \subset X$ kompakt genau dann, wenn M relativ kompakt und abgeschlossen ist.
(Denn: „$\Leftarrow$“ Da M relativ kompakt ist, hat jede Folge in M eine konvergente Teilfolge. Nach S. 10.24 liegt der Grenzwert jeder solchen Teilfolge in M.
„$\Rightarrow$“ Offenbar ist M relativ kompakt. Ist (x_n) eine Folge in M mit $x_n \to x$, so gilt $x \in M$, da auch jede Teilfolge gegen x konvergiert.)

Ein weiterer zentraler Baustein der Analysis ist

Satz 10.26 *(Heine-Borel)*
Für $M \subset \mathbb{K}^m$ sind äquivalent:

a) M ist kompakt,

b) M ist beschränkt und abgeschlossen.

Beweis. Ist M kompakt, so ist M nach B. 10.19 beschränkt und nach B. 10.25 auch abgeschossen.
Ist umgekehrt M beschränkt und abgeschlossen, so ist M nach B. 10.19 relativ kompakt (und abgeschlossen), also kompakt nach B. 10.25. $\square$

Beispiel 10.27 Sind $a = (a_1, \dots, a_m)$ und $b = (b_1, \dots, b_m) \in \mathbb{R}^m$ mit $a_j \leq b_j$ für $1 = 1, \dots, m$, so ist

$$[a, b] := \big\{x = (x_1, \dots, x_n) \in \mathbb{R}^m : x_j \in [a_j, b_j] \ (j = 1, \dots, m)\big\}$$

kompakt in $\mathbb{R}^m$.
(Denn: Es sei $(x^{(n)})$ eine Folge in $[a, b]$ mit $x^{(n)} \to x$. Dann gilt $x_j^{(n)} \to x_j$ $(n \to \infty)$ für $j = 1, \dots, m$. Aus $a_j \leq x_j^{(n)} \leq b_j$ folgt $a_j \leq x_j \leq b_n$ für $j = 1, \dots, m$, also $x \in [a, b]$. Damit ist $[a, b]$ abgeschlossen.
Außerdem ist $[a, b]$ auch beschränkt, da für alle $x \in [a, b]$

$$\|x\|_2 \leq \|x\|_1 = \sum_{j=1}^{m} |x_j| \leq \sum_{j=1}^{m} \max\big(|a_j|, |b_j|\big)$$

gilt. Nach dem Satz von Heine-Borel ist $[a, b]$ kompakt.)

11 Stetige Funktionen zwischen metrischen Räumen

Eine der wichtigsten Struktureigenschaften von Funktionen zwischen metrischen Räumen ist die Stetigkeit:

Definition 11.1 Es seien (X, d_X), (Y, d_Y) metrische Räume und $f : X \to Y$.
1. f heißt *stetig an der Stelle* $x_0 \in X$ (bzgl. d_X, d_Y), falls zu jedem $\varepsilon > 0$ ein $\delta_\varepsilon > 0$ existiert mit

$$d_Y(f(x), f(x_0)) < \varepsilon \quad \text{für alle } x \in X \text{ mit } d_X(x, x_0) < \delta_\varepsilon \ ,$$

d. h.

$$f(U_\delta(x_0)) \subset U_\varepsilon(f(x_0)) \ .$$

2. f heißt *stetig auf der Menge* $M \subset X$, falls f stetig an jeder Stelle $x_0 \in M$ ist. Ist $M = X$, 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.
In Abschnitt 8 hatten wir bereits (Folgen-) Stetigkeit für Funktionen $f : X \to \mathbb{K}$, wobei $X \subset \mathbb{K}$ ist, definiert. Der folgende Satz zeigt, dass die beiden Definition harmonieren.

Satz 11.2 *Es seien (X, d_X) und (Y, d_Y) metrische Räume und $f : X \to Y$. Ist $x_0 \in X$, so sind äquivalent:*

a) f ist stetig an der Stelle x_0.

b) Für alle Folgen (x_n) in X mit $x_n \to x_0$ gilt $f(x_n) \to f(x_0)$.

Beweis.
a) $\Rightarrow$ *b)*: Es sei (x_n) eine Folge in X 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 $d_Y(f(x), f(x_0)) < \varepsilon$ für alle x mit $d_X(x, x_0) < \delta_\varepsilon$. Aus $x_n \to x_0$ folgt die Existenz eines $N_\varepsilon = N(\delta_\varepsilon) \in \mathbb{N}$ mit $d_X(x_n, x_0) < \delta_\varepsilon$ für alle $n \geq N_\varepsilon$. Also gilt auch $d_Y(f(x_n), f(x_0)) < \varepsilon \, (n \geq N_\varepsilon)$.
b) $\Rightarrow$ *a)*: Angenommen, f ist unstetig an x_0. Dann existiert ein $\varepsilon > 0$ so, dass für alle $\delta > 0$ ein x existiert mit $d_X(x, x_0) < \delta$ und $d_Y(f(x), y_0) \geq \varepsilon$. Insbesondere existiert zu jedem $n \in \mathbb{N}$ ein x_n mit $d_X(x_n, x_0) < 1/n$ und $d_Y(f(x_n), y_0) \geq \varepsilon$. Für die Folge (x_n) gilt damit $x_n \to x_0 \, (n \to \infty)$ und $f(x_n) \not\to f(x_0)$. Widerspruch! $\square$

Bemerkung 11.3 Es sei (X, d) ein metrischer Raum.
1. Ist $f : X \to X$ definiert durch $f(x) := x$ $(x \in X)$ (d. h. $f = \mathrm{id}_X$), so ist ist f stetig (bzgl. der Metriken $d = d_X = d_Y$).
2. Wie in B. 8.2 ergibt sich aus den entsprechenden Grenzwertsätzen für Folgen und S. 11.2: Sind $f, g : X \to \mathbb{K}$ stetig an x_0, so sind auch $f \pm g$, $f \cdot g$ und (falls definiert) f/g stetig an x_0.
Aus S. 11.2 ergibt sich auch leicht: Sind (Y, d_Y) und (Z, d_Z) weitere metrische Räume, ist $f : X \to Y$ stetig an x_0 und ist $g : Y \to Z$ stetig an $f(x_0)$, so ist auch $g \circ f$ stetig an x_0.

Wir wollen nun eine weitere Charakterisierung der Stetigkeit herleiten.

Definition 11.4 Es seien (X, d) ein metrischer Raum und $M \subset X$. Ein Punkt $x_0 \in X$ heißt *Häufungspunkt* von M, falls zu jedem $\varepsilon > 0$ ein $x \in M$ mit $0 < d(x, x_0) < \varepsilon$ existiert. Wir schreiben M' für die Menge aller Häufungspunkte von M. Ist $x \in M$ und kein Häufungspunkt, so heißt x_0 ein *isolierter Punkt* von M.

Beispiel 11.5 Es seien $X = \mathbb{R}$ und $M = \{1/k,\, k \in \mathbb{N}\}$. Dann ist $0 \in M'$.

Bemerkung und Definition 11.6 Es seien (X, d_X) und (Y, d_Y) metrische Räume. Ferner seien $x_0 \in X'$ und $f : X \to Y$ oder $f : X \setminus \{x_0\} \to Y$.
1. Wir sagen, f hat den *Grenzwert* $y_0 \in Y$ an x_0, falls zu jedem $\varepsilon > 0$ ein $\delta_\varepsilon > 0$ existiert mit

$$d_Y(f(x), y_0) < \varepsilon \qquad \text{für alle } x \in X \text{ mit } 0 < d_X(x, x_0) < \delta_\varepsilon. \tag{11.1}$$

In diesem Fall schreiben wir

$$f(x) \to y_0 \qquad (x \to x_0)\ .$$

Man beachte, dass auch im Falle $f : X \to Y$, also $x_0 \in D(f) = X$, der Wert $f(x_0)$ keine Rolle spielt.
Aus den Definitionen ergibt sich unmittelbar: Es gilt $f(x) \to y_0$ $(x \to x_0)$ genau dann, wenn die Funktion $f_0 : X \to Y$, definiert durch

$$f_0(x) = \begin{cases} y_0, & \text{falls} \quad x = x_0 \\ f(x), & \text{falls} \quad x \neq x_0, \end{cases}$$

stetig an x_0 ist. Ist $f : X \to Y$, so gilt insbesondere für $x_0 \in X'$ damit:

$$f \text{ stetig an } x_0 \quad \Leftrightarrow \quad f(x) \to f(x_0) \quad (x \to x_0).$$

(Aus der Definition der Stetigkeit folgt zudem sofort: Ist $x_0 \notin X'$, so ist f stets stetig an x_0.)

2. Manchmal ist es sinnvoll, allgemeiner Grenzwerte bezüglich geeigneter Teilmengen von X zu betrachten. Es sei also $M \subset X$ mit $x_0 \in M'$. Gilt dann (11.1) für alle $x \in M$ (mit $0 < d_X(x, x_0) < \delta_\varepsilon$), so heißt $y_0 \in Y$ *Grenzwert* von f an x_0 *bezüglich* M. Wir schreiben dann

$$f(x) \to y_0 \quad (x \to x_0,\, x \in M)\ .$$

Aus $x_0 \in M'$ folgt, dass höchstens ein Grenzwert y_0 von f an x_0 bzgl. M existiert. Wir schreiben im Falle der Existenz auch

$$\lim_{x \to x_0, x \in M} f(x) := y_0$$

und in Fall 1. kurz

$$\lim_{x \to x_0} f(x) := y_0\ .$$

Ganz praktisch ist schließlich die Tatsache, dass $f(x) \to y_0$ $(x \to x_0, x \in M)$ genau dann gilt, wenn $d_Y(f(x), y_0) \to 0$ $(x \to x_0, x \in M)$.

Beispiel 11.7 Es sei $(X, d) = (\mathbb{C}, d_{|\cdot|})$. Dann gilt

$$\frac{e^z - 1}{z} \to 1 \quad (z \to 0)$$

(Denn: Für $z \neq 0$ ist

$$|\frac{e^z - 1}{z} - 1| = \left| \frac{1}{z} \sum_{\mu=1}^{\infty} \frac{z^\mu}{\mu!} - 1 \right| = \left| \sum_{\nu=1}^{\infty} \frac{z^\nu}{(\nu+1)!} \right| \leq \sum_{\nu=1}^{\infty} \frac{|z|^\nu}{(\nu-1)!} \leq |z| e^{|z|}.$$

Da $z \mapsto |z| e^{|z|}$ stetig (an 0) ist, ergibt sich zunächst

$$|z| e^{|z|} \to 0 \qquad (z \to 0)$$

und damit auch

$$\left| \frac{e^z - 1}{z} - 1 \right| \to 0 \qquad (z \to 0)\ .)$$

Bemerkung 11.8 Aus den jeweiligen Definitonen ergibt sich leicht eine Charakterisierung des Grenzwertes mittels Folgenkonvergenz ([Ü]): Es gilt

$$f(x) \to y_0 \quad (x \to x_0,\, x \in M)$$

genau dann, wenn **für alle** Folgen (x_n) in M mit $x_n \to x_0$ und $x_n \neq x_0$ (wichtig!) auch $f(x_n) \to y_0$ $(n \to \infty)$ gilt. Aus den Grenzwertsätzen für Folgen erhalten wir damit auch: Sind $f, g : X \setminus \{x_0\} \to \mathbb{K}$ (oder $f, g : X \to \mathbb{K}$) so, dass

$$f(x) \to y_0, \quad g(x) \to z_0 \quad (x \to x_0,\, x \in M)\ ,$$

so gilt auch

$$(f \pm g)(x) \to y_0 \pm z_0, \quad (fg)(x) \to y_0 z_0 \quad (x \to x_0,\, x \in M)$$

und im Falle $g(x) \neq 0$ und $z_0 \neq 0$ auch

$$(f/g)(x) \to y_0/z_0 \quad (x \to x_0,\, x \in M)\ .$$

Bemerkung und Definition 11.9 Ist in B./D. 11.6.2 speziell $X \subset \mathbb{R}$ ein Intervall und $M = X \cap (-\infty, x_0)$, so sagt man, dass f an x_0 den *linksseitigen Grenzwert* y_0 hat. Man schreibt dann $f(x) \to y_0$ $(x \to x_0^-)$ sowie

$$f(x_0^-) := \lim_{x \to x_0^-} f(x) := y_0.$$

Entsprechend spricht man im Falle $M = X \cap (x_0, \infty)$ vom *rechtsseitigen Grenzwert* y_0 und schreibt dann $f(x) \to y_0$ $(x \to x_0^+)$ sowie

$$f(x_0^+) := \lim_{x \to x_0^+} f(x) := y_0.$$

Es gilt damit ([Ü]): Ist x_0 kein Randpunkt von X, so existiert $y_0 = \lim\limits_{x \to x_0} f(x)$ genau dann, wenn $f(x_0^+)$ und $f(x_0^-)$ existieren und

$$f(x_0^+) = f(x_0^-) = y_0$$

erfüllt ist. Existieren $f(x_0^+)$ und $f(x_0^-)$ und ist $f(x_0^+) \neq f(x_0^-)$, so heißt x_0 *Sprungstelle* von f.

Ist X ein nach oben unbeschränktes Intervall, so schreiben wir ferner

$$f(x) \to y_0 \quad (x \to \infty) \qquad \text{oder auch} \qquad \lim_{x \to \infty} f(x) := y_0$$

falls für alle $\varepsilon > 0$ ein $R_\varepsilon > 0$ so existiert, dass

$$d_Y(f(x), y_0) < \varepsilon \quad \text{für alle } x \in X,\, x > R_\varepsilon\ .$$

Entsprechend definiert man $\lim\limits_{x \to -\infty} f(x)$, falls X nach unten unbeschränkt ist.

Beispiel 11.10 1. 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

$$f(0^+) = \lim_{x \to 0^+} f(x) = 1, \qquad f(0^-) = \lim_{x \to 0^-} f(x) = -1$$

Damit ist x_0 eine Sprungstelle von f und $\lim\limits_{x \to 0} f(x)$ existiert nicht.
2. Es sei $X = [0, \infty)$ und $f : (0, \infty) \to \mathbb{R}$ definiert durch

$$f(x) := \sin(\pi/x) \quad (x > 0) .$$

Dann existiert kein rechtsseitiger Grenzwert von f an $x_0 = 0$.
(Denn: Für $x_n = 1/(n + 1/2)$ gilt $0 < x_n \to 0$, aber $(f(x_n)) = ((-1)^n)$ ist divergent. Aus B. 11.8 folgt die Behauptung.)
3. 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 existiert für **kein** x_0 in $\mathbb{R}$ der (rechts- oder linksseitige) Grenzwert!
(Denn: Ist $x_0 \in \mathbb{R}$ und ist $\delta > 0$, so existieren $x \in \mathbb{Q}$ und $y \in \mathbb{R} \setminus \mathbb{Q}$ (also $f(x) = 1$ und $f(y) = 0$) mit $x_0 < x, y < x_0 + \delta$. Hieraus folgt, dass kein rechtsseitiger Grenzwert an x_0 existiert. Entsprechend sieht man, dass kein linksseitiger Grenzwert existiert.)
Insbesondere ist damit f unstetig an allen Stellen $x_0 \in \mathbb{R}$.
4. Es gilt

$$\lim_{x \to -\infty} e^x = 0 \qquad \text{und} \qquad \lim_{x \to \infty} 1/x = 0$$

Definition 11.11 Unter den Bedingungen von D. 11.6 sei speziell $Y = \mathbb{R}$. Dann schreiben wir

$$f(x) \to +\infty \qquad (\text{bzw. } f(x) \to -\infty) \qquad (x \to x_0, x \in M),$$

falls für alle $R > 0$ ein $\delta_R > 0$ so existiert, dass

$$f(x) > R \qquad (\text{bzw. } f(x) < -R) \quad \text{für alle } x \in M \text{ mit } 0 < d(x, x_0) < \delta_R . \tag{11.2}$$

Entsprechend definieren wir wie in D. 11.9

$$f(x) \to \pm\infty \qquad (x \to x_0^{\pm}),$$

sowie

$$f(x) \to \pm\infty \quad (x \to \infty) \qquad \text{bzw.} \qquad f(x) \to \pm\infty \quad (x \to -\infty)\ .$$

So gilt etwa

$$\ln x \to -\infty \quad (x \to 0^+) \qquad \text{und} \qquad e^x \to \infty \quad (x \to \infty)$$

(statt $+\infty$ schreibt man auch kurz ∞).

Wir befassen uns nun mit Eigenschaften stetiger Funktionen.

Satz 11.12 *Es seien (X, d_X) und (Y, d_Y) metrische Räume, und es sei $f : X \to Y$. Dann sind äquivalent:*

a) f ist stetig.

b) Für alle abgschlossenen Mengen $A \subset Y$ ist $f^{-1}(A)$ abgeschlossen in X.

c) Für alle offenen Mengen $V \subset Y$ ist $f^{-1}(V)$ offen in X.

Beweis. a) $\Rightarrow$ b): Es sei (x_n) eine Folge in $f^{-1}(A)$ mit $x_n \to x$ $(n \to \infty)$. Aus der Stetigkeit von f folgt $A \ni f(x_n) \to f(x)$ $(n \to \infty)$. Da A abgeschlossen ist, gilt $f(x) \in A$, also $x \in f^{-1}(A)$. Damit ist auch $f^{-1}(A)$ abgeschlossen.
b) $\Rightarrow$ c) Ist V offen, so ist V^c abgeschlossen. Also ist auch $(f^{-1}(V))^c = f^{-1}(V^c)$ abgeschlossen und damit $f^{-1}(V)$ offen.
c) $\Rightarrow$ a): Es sei $x \in X$ und $\varepsilon > 0$. Dann ist $V := U_\varepsilon(f(x))$ offen in Y, und damit nach Voraussetzung auch $f^{-1}(V)$ offen in X. 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. □

Bemerkung 11.13 Für Bildmengen gilt eine S. 11.12 entsprechende Aussage i. A. nicht!
Ist etwa $(X, d_X) = (Y, d_Y) = (\mathbb{R}, d_{|\cdot|})$, so ist $\mathbb{R}$ offen und abgeschlossen. Ist $f(x) = x^2$ $(x \in \mathbb{R})$, so ist $f(\mathbb{R}) = [0, \infty)$ nicht offen. Ist $g(x) = \arctan x$ $(x \in \mathbb{R})$, so ist $g(\mathbb{R}) = (-\pi/2, \pi/2)$ nicht abgeschlossen.

Anders als Offenheit oder Abgeschlosssenheit überträgt sich Kompaktheit unter stetigen Funktionen auf Bildmengen.

Satz 11.14 *Es seien (X, d_X), (Y, d_Y) metrische Räume und $f : X \to Y$ stetig. Dann gilt: Ist $M \subset X$ (relativ) kompakt , so ist auch $f(M) \subset Y$ (relativ) kompakt.*

Beweis. Es sei (y_n) eine Folge in $f(M)$. Dann existieren $x_n \in M$ mit $y_n = f(x_n)$ ($n \in \mathbb{N}$). Ist M relativ kompakt, so existieren ein $x \in X$ und eine Teilfolge (x_{n_j}) von (x_n) mit $x_{n_j} \to x$ ($j \to \infty$). Da f stetig ist, folgt

$$y_{n_j} = f(x_{n_j}) \to f(x) \in f(M) \qquad (j \to \infty) .$$

Damit ist $f(M)$ relativ kompakt. Ist M sogar kompakt, so ist $x \in M$ (da M abgeschlossen) und damit auch $f(x) \in f(M)$. Damit ist auch $f(M)$ kompakt. □

Für reellwertige Funktionen hat der Satz eine wichtige Konsequenz. Um diese formulieren zu können, brauchen wir eine Definition.

Definition 11.15 Es sei $X \neq \emptyset$ eine Menge, und es sei $f : X \to \mathbb{R}$. Ferner sei $M \subset X$.

1. Man sagt, f hat ein *Maximum* bzgl. M, falls $\max_M f(x) := \max_{x \in M} f(x) := \max f(M)$ existiert. Ist $x_0 \in M$ so, dass $f(x_0) = \max_M f(x)$ gilt, d. h. ist

   $$f(x) \leq f(x_0) \qquad \text{für alle} \quad x \in M ,$$

   so sagt man, f hat ein Maximum bzgl. M *an* x_0.

2. Man sagt, f hat ein *Minimum* bzgl. M, falls $\min_M f(x) := \min_{x \in M} f(x) := \min f(M)$ existiert. Ist $x_0 \in M$ so, dass $f(x_0) = \min_M f(x)$ gilt, d. h. ist

   $$f(x) \geq f(x_0) \qquad \text{für alle} \quad x \in M ,$$

   so sagt man, f hat ein Minimum bzgl. M *an* x_0.

Ist $X = M$ so spricht man auch von *absolutem* (oder *globalem*) Maximum bzw. Minimum.

Beispiel 11.16 Es sei $X = \mathbb{R}$. Ist $f(x) := x^2$ ($x \in \mathbb{R}$), so hat f an $x_0 = 0$ ein absolutes Minimum, aber f hat kein absolutes Maximum. Ist $g(x) = \arctan x$ ($x \in \mathbb{R}$), so ist g beschränkt, aber g hat weder ein absolutes Maximum noch ein absolutes Minimum.

Satz 11.17 *Es seien (X, d_X) ein metrischer Raum und $f : X \to \mathbb{R}$ stetig. Ist $K \subset X$* ***kompakt****, so hat f ein Maximum und ein Minimum bzgl. K, d. h. es existieren $x_1, x_2 \in K$ mit*

$$f(x_1) \leq f(x) \leq f(x_2) \qquad \text{für alle} \quad x \in K\ .$$

Beweis. Nach S. 11.14 ist $f(K) \subset \mathbb{R}$ kompakt, also auch beschränkt und abgeschlossen. Damit existieren $\max f(K)$ und $\min f(K)$ ([Ü]). □

Beispiel 11.18 Es sei $X = \mathbb{R}$ und $f(x) := x^2$ $(x \in \mathbb{R})$. Dann ist

$$f(a) = f(-a) = \max_{[-a,a]} f(x)$$

d. h. f hat an a und $-a$ Maxima bzgl. $[-a, a]$.

Im Allgemeinen ist die Umkehrfunktion einer stetigen, bijektiven Funktion nicht stetig ([Ü]). Als weitere sehr elegante Anwendung von S. 11.14 erhält man jedoch:

Satz 11.19 *Es seien (X, d_X) und (Y, d_Y) metrische Räume, und es sei X* ***kompakt****. Ist $f : X \to Y$ bijektiv und stetig, so ist auch Y kompakt und $f^{-1} : Y \to X$ stetig.*

Beweis. Zunächst ist $Y = f(X)$ nach S. 11.14 kompakt.
Wir beweisen: f^{-1} ist stetig. Dazu sei $A \subset X$ abgeschlossen. Nach S. 11.12 reicht es zu zeigen, dass $(f^{-1})^{-1}(A) = f(A) \subset Y$ abgeschlossen ist.
Da Teilmengen relativ kompakter Mengen wieder relativ kompakt sind, und da A abgeschlossen ist, ist A kompakt nach B. 10.25. Also ist $f(A) \subset Y$ kompakt nach S. 11.14 und damit insbesondere abgeschlossen (wieder nach B. 10.25). □

Eine weitere wichtige Eigenschaft stetiger Funktionen auf kompakten Mengen ist die gleichmäßige Stetigkeit:

Definition 11.20 Es seien (X, d_X) und (Y, d_Y) metrische Räume, und es sei $f : X \to Y$. Dann heißt f *gleichmäßig stetig*, falls für alle $\varepsilon > 0$ ein $\delta_\varepsilon > 0$ so existiert, dass

$$d_Y(f(x_1), f(x_2)) < \varepsilon \qquad \text{für alle } x_1, x_2 \in X \text{ mit } d_X(x_1, x_2) < \delta_\varepsilon\ .$$

Aus der Definition ergibt sich sofort, dass jede gleichmäßig stetige Funktion auch stetig ist. Das folgende Beispiel zeigt, dass die Stetigkeit i. A. nicht die gleichmäßige Stetigkeit impliziert.

Beispiel 11.21 Es sei $f : R \to \mathbb{R}$ mit $f(x) = x^2\,(x \in \mathbb{R})$. Dann ist f stetig auf $\mathbb{R}$, aber nicht gleichmäßig stetig.
(Es sei $\varepsilon = 1$, und es sei $\delta > 0$ beliebig. Wir wählen $x_1 = 1/\delta, x_2 = 1/\delta + \delta/2$. Dann ist $|x_1 - x_2| = \delta/2 < \delta$, aber

$$|f(x_1) - f(x_2)| = |x_1 + x_2| \cdot |x_1 - x_2| > 2/\delta \cdot \delta/2 = 1 = \varepsilon\ .$$

Folglich ist f nicht gleichmäßig stetig auf $\mathbb{R}$.)

Es gilt allgemein

Satz 11.22 *Es seien (X, d_X) und (Y, d_Y) metrische Räume. Ist X* ***kompakt*** *und $f : X \to Y$ stetig, so ist f gleichmäßig stetig.*

Beweis. Angenommen nicht. Dann existiert ein $\varepsilon > 0$ so, dass für alle $n \in \mathbb{N}$ zwei Punkte $x_n, y_n \in X$ existieren mit

$$d_X(x_n, y_n) < 1/n \qquad \text{und} \quad d_Y(f(x_n), f(y_n)) \geq \varepsilon\ .$$

Da X kompakt ist, besitzt die Folge (x_n) eine Teilfolge $(x_{n_j})_j$ mit $x_{n_j} \to x \in X$. Damit gilt auch

$$d_X(x, y_{n_j}) \leq d_X(x, x_{n_j}) + d_X(x_{n_j}, y_{n_j}) \to 0 \qquad (j \to \infty)\ .$$

d. h. $y_{n_j} \to x$ $(j \to \infty)$. Also folgt auf Grund der Stetigkeit von f an der Stelle x

$$\varepsilon \leq d_Y(f(x_{n_j}), f(y_{n_j})) \leq d_Y(f(x_{n_j}), f(x)) + d_Y(f(x), f(y_{n_j})) \to 0 \quad (j \to \infty)\ .$$

Widerspruch! □

12 Differenzialrechnung von Funktionen einer Variablen

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 eine Funktion $f : \mathbb{R} \to \mathbb{R}$ und für $h \neq 0$ die Sekanten S_h durch die Punkte $(x_0, f(x_0))$ und $(x_0+h, f(x_0+h))$ in $\mathbb{R}^2$. Diese Sekanten sind festgelegt durch $(x_0, f(x_0))$ und ihre Steigung

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

Für $h \to 0$ 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, wobei wir allgemeinere Definitionsbereiche in $\mathbb{C}$ und komplexwertige Funktionen zulassen.

Definition 12.1 Es seien $X \subset \mathbb{K}$ mit $X \subset X'$ und $f : X \to \mathbb{K}$.

1. f heißt *differenzierbar* an der Stelle $x_0 \in X$, falls der Grenzwert
$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} \left(= \lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} \right)$$
(genauer eigentlich $h \to 0$, $h \in X - \{x_0\}$) 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* $M \subset X$ falls f in jedem Punkt $x \in M$ differenzierbar ist. Ist $M = X$, so heißt die Funktion $f' : X \to \mathbb{K}$ *Ableitung* von f.

3. *Höhere Ableitungen* von f werden rekursiv definiert: Ist $f^{(0)} := f$ und $f^{(k-1)}$ differenzierbar auf X, wobei $k \in \mathbb{N}$, so heißt $f^{(k)} := (f^{(k-1)})' : X \to \mathbb{K}$ k-te Ableitung von f. Wir schreiben dabei meist $f'' := f^{(2)}$ und $f''' := f^{(3)}$.

Beispiel 12.2 1. Ist $f(z) = az + b$ $(z \in \mathbb{C})$ mit Konstanten $a, b \in \mathbb{C}$, so ist f differenzierbar auf $\mathbb{C}$ und es gilt

$$f'(z) = \lim_{h\to 0} \frac{f(z+h)-f(z)}{h} = \lim_{h\to 0} \frac{ah}{h} = a \qquad (z \in \mathbb{C}) \ .$$

2. Für $f(z) = z^2$ $(z \in \mathbb{C})$ gilt

$$f'(z) = \lim_{h\to 0} \frac{(z+h)^2 - z^2}{h} = \lim_{h\to 0} \frac{2zh + h^2}{h} = \lim_{h\to 0}(2z+h) = 2z \qquad (z \in \mathbb{C}) .$$

Allgemeiner kann man (mit Hilfe der binomischen Formel; [Ü]) zeigen: Ist $n \in \mathbb{N}$ und $f(z) = z^n$ $(z \in \mathbb{C})$, so ist

$$f'(z) = nz^{n-1} \qquad (z \in \mathbb{C}) .$$

3. (Wichtig !!) Die Funktion $\exp : \mathbb{C} \to \mathbb{C}$ ist differenzierbar auf $\mathbb{C}$ mit

$$\exp' = \exp \ .$$

(Denn. Nach B. 11.7 gilt

$$\lim_{h\to 0} \frac{e^h - 1}{h} = 1$$

und damit für alle $z \in \mathbb{C}$

$$\frac{e^{z+h} - e^z}{h} = e^z \cdot \frac{e^h - 1}{h} \to e^z \qquad (h \to 0) .)$$

4. Die Funktion $f(x) := |x|$ $(x \in \mathbb{R})$ ist nicht differenzierbar an der Stelle $x_0 = 0$, denn ist $h > 0$, so gilt

$$\frac{f(0+h) - f(0)}{h} = 1 \to 1 \qquad (h \to 0^+)$$

und ist $h < 0$, so gilt

$$\frac{f(0+h) - f(0)}{h} = \frac{-h}{h} = -1 \to -1 \qquad (h \to 0^-) .$$

Also existiert

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

nicht!

Das letzte Beispiel zeigt, dass die Stetigkeit an einer Stelle i. A. **nicht** die Differenzierbarkeit an dieser Stelle impliziert. Umgekehrt implizert Differenzierbarkeit jedoch Stetigkeit, wie der zweite Teil des folgenden Satzes zeigt. Der erste Teil liefert eine wichtige Charakterisierung der Differenzierbarkeit.

Satz 12.3 *Es seien $X \subset \mathbb{K}$ mit $X \subset X'$, $f : X \to \mathbb{K}$ und $x_0 \in X$.*
1. Für $X_0 := X - \{x_0\}$ sind folgende Aussagen äquivalent.

a) f ist differenzierbar an x_0.

b) (Zerlegungsformel) Es existieren ein $c \in \mathbb{K}$ *und eine an* 0 *stetige Funktion* $\varepsilon : X_0 \to \mathbb{K}$ *mit* $\varepsilon(0) = 0$ *und*

$$f(x_0 + h) = f(x_0) + c \cdot h + \varepsilon(h) \cdot h \qquad (h \in X_0) \; .$$

c) Es existiert eine an 0 *stetige Funktion* $\varphi : X_0 \to \mathbb{K}$ *mit*

$$f(x_0 + h) - f(x_0) = \varphi(h) \cdot h \qquad (h \in X_0)$$

Außerdem ist in diesem Fall $f'(x_0) = c = \varphi(0)$*.*
2. Ist f *differenzierbar an* x_0*, so ist* f *auch stetig an* x_0*.*

Beweis. 1. a) $\Rightarrow$ b): Wir setzen $\varepsilon(0) := 0$ und für $h \in X_0$, $h \neq 0$,

$$\varepsilon(h) := \frac{f(x_0 + h) - f(x_0)}{h} - f'(x_0).$$

Dann ist

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \varepsilon(h) \cdot h$$

(also $c = f'(x_0)$) und es gilt $\varepsilon(h) \to 0 = \varepsilon(0)$ $(h \to 0)$ aufgrund der Differenzierbarkeit von f an x_0.
b) $\Rightarrow$ c): Klar mit $\varphi(h) := c + \varepsilon(h)$.
c) $\Rightarrow$ a): Aus

$$\frac{f(x_0 + h) - f(x_0)}{h} = \varphi(h) \to \varphi(0) \qquad (h \to 0)$$

ergibt sich die Differenzierbarkeit von f an x_0 und $f'(x_0) = \varphi(0)$.

2. Folgt aus 1. und $\varphi(h) \to \varphi(0)$ für $h \to 0$. □

Wir stellen wichtige Rechenregeln für Ableitungen zusammen.

Satz 12.4 *(Summen-, Produkt- und Quotientenregel)*
Es sei $X \subset \mathbb{K}$ *und es seien* $f, g : X \to \mathbb{K}$ *differenzierbar an der Stelle* $x_0 \in X$*. 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. 1. Klar nach Definition der Ableitung (oder nach der Zerlegungsformel).
2. Da g differenzierbar an x_0 ist, ist g auch stetig an x_0 (S. 12.3). Damit gilt für $h \in X_0$

$$\begin{aligned}\frac{1}{h}\left((fg)(x_0+h) - (fg)(x_0)\right) &= \frac{f(x_0+h) - f(x_0)}{h} g(x_0+h) + f(x_0)\frac{g(x_0+h) - g(x_0)}{h} \\ &\to f'(x_0)g(x_0) + f(x_0)g'(x_0) \, .\end{aligned}$$

3. Da g stetig an 0 ist mit $g(x_0) \neq 0$, existiert ein $\delta > 0$ mit $g(x) \neq 0$ für alle $x \in X$ mit $|x - x_0| < \delta$. Es sei $h \in X_0$ (o. E. mit $|h| < \delta$). Dann gilt

$$\begin{aligned}\frac{(1/g)\,(x_0+h) - (1/g)\,(x_0)}{h} &= \frac{1}{g(x_0+h)g(x_0)}\left(\frac{g(x_0) - g(x_0+h)}{h}\right) \\ &\to \frac{1}{g^2(x_0)}(-g'(x_0)) \qquad (h \to 0) \, .\end{aligned}$$

Damit ist 3. für $f \equiv 1$ bewiesen. Die Aussage für allgemeines f folgt hieraus mit der Produktregel. □

Beispiel 12.5 Ist $P(z) = \sum\limits_{\nu=0}^{n} a_\nu z^\nu$ $(z \in \mathbb{C})$ mit $a_0, \ldots, a_n \in \mathbb{C}$, so gilt

$$P'(z) = \sum_{\nu=1}^{n} \nu \cdot a_\nu z^{\nu-1} \qquad (z \in \mathbb{C})$$

(Folgt aus B. 12.2.2 und mehrfacher Anwendung von S. 12.4 1.)

Satz 12.6 *(Kettenregel)*
Es seien $X, Y \subset \mathbb{K}$ und es sei $f : X \to Y$. Ferner sei $g : Y \to \mathbb{K}$. Ist f differenzierbar an $x_0 \in X$ und ist g differenzierbar an $y_0 := f(x_0) \in Y$, so ist $g \circ f$ differenzierbar an x_0 mit

$$(g \circ f)'(x_0) = g'(y_0) \cdot f'(x_0) = g'\big(f(x_0)\big) f'(x_0) \, .$$

Beweis. Es sei $Y_0 := Y - \{y_0\}$. Nach S. 12.3 existieren an 0 stetige Funktionen $\varphi : X_0 \to \mathbb{K}$ und $\psi : Y_0 \to \mathbb{K}$ mit

$$\begin{aligned} f(x_0+h) - f(x_0) &= \varphi(h) \cdot h \\ g(y_0+u) - g(y_0) &= \psi(u) \cdot u \end{aligned}$$

Für $h \in X_0$ ist damit

$$g\Big(f(x_0+h)\Big) - g\Big(f(x_0)\Big) = g(y_0 + \varphi(h) \cdot h) - g(y_0) = \psi(\varphi(h)h)\varphi(h) \cdot h\ .$$

Da $(\psi \circ (\varphi \cdot \mathrm{id}_{X_0})) \cdot \varphi$ stetig an 0 ist, folgt aus S. 12.3 die Differenzierbarkeit von $g \circ f$ an x_0 sowie

$$(g \circ f)'(x_0) = \psi(0)\varphi(0) = g'(y_0)f'(x_0)\ .$$

□

Beispiel 12.7 1. Es sei $P : \mathbb{C} \to \mathbb{C}$ mit

$$P(z) = (z^3 + 2z + 1)^5\ .$$

Dann gilt $P = g \circ f$ mit $f(z) = z^3 + 2z + 1$ und $g(w) = w^5$. Also ergibt sich aus der Kettenregel

$$P'(z) = g'(f(z))f'(z) = 5(z^3 + 2z + 1)^4(3z^2 + 2)\ .$$

2. Es gilt auf $\mathbb{C}$

$$\cos' = -\sin \qquad \text{und} \qquad \sin' = \cos\,.$$

(Denn: Für $z \in \mathbb{C}$ ist

$$\cos'(z) = \frac{1}{2}(e^{iz} + e^{-iz})' = \frac{1}{2}(ie^{iz} - ie^{-iz}) = -\frac{1}{2i}(e^{iz} - e^{-iz}) = -\sin z$$

und für sin entsprechend.)

3. Aus 1. und der Quotientenregel folgt ([Ü])

$$\begin{aligned} \tan'(z) &= \frac{1}{\cos^2 z} = 1 + \tan^2 z \qquad (z \neq \pi/2 + k\pi,\ k \in \mathbb{Z}) \\ \cot'(z) &= -\frac{1}{\sin^2 z} = -1 - \cot^2 z \qquad (z \neq k\pi,\ k \in \mathbb{Z}) \end{aligned}$$

4. Ist $a > 0$ fest, so gilt

$$(a^z)' = a^z \ln a \qquad (z \in \mathbb{C}) .$$

(Denn: $(a^z)' = (e^{z \cdot \ln a})' = \ln a \cdot e^{z \ln a} = \ln a \cdot a^z$.)

Satz 12.8 *(Umkehrregel)*
Es sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ sei 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. Zunächst ist y_0 ein Häufungspunkt von $W(f) = D(f^{-1})$ (folgt aus der Stetigkeit von f an x_0). Für $J_0 := W(f) - \{0\}$ und

$$h(u) := f^{-1}(y_0 + u) - f^{-1}(y_0) \qquad (u \in J_0)$$

gilt $0 \neq h(u) \to 0$ $(u \to 0)$, da f^{-1} stetig (an y_0) nach S. 9.6. Also erhalten wir

$$\frac{f^{-1}(y_0 + u) - f^{-1}(y_0)}{u} = \frac{h(u)}{f(x_0 + h(u)) - f(x_0)} \to \frac{1}{f'(x_0)} \qquad (u \to 0) .$$

□

Beispiel 12.9 1. 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 $f(x) = a^x$ $(x \in \mathbb{R})$. Dann ist $f^{-1}(y) = \log_a y$ $(y > 0)$. Also gilt nach der Umkehrregel und 1. mit $x = \log_a y$

$$(f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{a^x \ln a} = \frac{1}{y} \cdot \frac{1}{\ln a} \qquad (y > 0) .)$$

2. 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 1. gilt für $y \in (-1,1)$ mit $x = \arcsin y$

$$\arcsin'(y) = \frac{1}{\sin'(x)} = \frac{1}{\cos x} = \frac{1}{\sqrt{1-\sin^2 x}} = \frac{1}{\sqrt{1-y^2}} \qquad (y \in (-1,1))\ .$$

Entsprechend für $\arccos x$.)
3. Es gilt ([Ü])

$$\begin{aligned}\arctan'(x) &= \frac{1}{1+x^2} \qquad (x \in \mathbb{R})\\ \operatorname{arccot}'(x) &= -\frac{1}{1+x^2} \qquad (x \in \mathbb{R})\ .\end{aligned}$$

4. Für $\alpha \in \mathbb{C}$ fest gilt

$$(x^\alpha)' = \alpha x^{\alpha-1} \qquad (x > 0)$$

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

13 Der Mittelwertsatz und Anwendungen

Wir beschäftigen uns nun mit der geometrischen Interpretation der Ableitung. Dazu definieren wir zunächst, was wir unter lokalen Extrema verstehen.

Definition 13.1 Es seien (X, d) ein metrischer Raum und $f : X \to \mathbb{R}$. Ferner sei $x_0 \in X$.

1. Man sagt, f habe an x_0 ein *lokales* (oder auch *relatives*) *Maximum*, falls ein $\delta > 0$ existiert mit
$$f(x) \leq f(x_0) \qquad \text{für alle} \quad x \in U_\delta(x_0)$$
2. Man sagt, f habe an x_0 ein *lokales* (oder auch *relatives*) *Minimum*, falls ein $\delta > 0$ existiert mit
$$f(x) \geq f(x_0) \qquad \text{für alle} \quad x \in U_\delta(x_0)\ .$$

Eine Stelle x_0, an der ein lokales Maximum oder ein lokales Minimum vorliegt, bezeichnet man als *Extremstelle* (oder *lokales Extremum*) von f.
Gilt $<$ statt $\leq$ für $x \neq x_0$ in 1. (bzw. $>$ statt $\geq$ für $x \neq x_0$ in 2.), so spricht man von einem *strikten* lokalen Maximum (bzw. *strikten* lokalen Minimum).

Es gilt folgendes einfache **notwendige** Kriterium für Extrema differenzierbarer Funktionen.

Satz 13.2 *Es sei $I \subset \mathbb{R}$ ein* **offenes** *Intervall und die Funktion $f : I \to \mathbb{R}$ sei differenzierbar an der Stelle $x_0 \in I$. Ist x_0 eine Extremstelle von f, so gilt*
$$f'(x_0) = 0\ .$$

Beweis. O. E. liege an x_0 ein lokales Minimum vor. Dann existiert ein $\delta > 0$ mit $f(x) \geq f(x_0)$ für alle $x \in U_\delta(x_0)$. Da f differenzierbar an x_0 ist, gilt
$$f'(x_0) = \lim_{h \to 0^+} \frac{f(x_0 + h) - f(x_0)}{h} \geq 0$$
und
$$f'(x_0) = \lim_{h \to 0^-} \frac{f(x_0 + h) - f(x_0)}{h} \leq 0\ ,$$
d. h. $f'(x_0) = 0$. □

Bemerkung und Definition 13.3 1. Punkte x_0, an denen $f'(x_0) = 0$ ist, nennt man auch *kritische Punkte* von f.
2. Wie bereits erwähnt, liefert S. 13.2 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.
3. Für nicht offene Intervalle I bzw. nicht differenzierbare Funktionen ist die Aussage von S. 13.2 i. A. falsch. So hat etwa $f : [0,1] \to \mathbb{R}$ mit $f(x) = x$ an 0 ein (sogar globales) Minimum, aber 0 ist kein kritischer Punkt (sondern ein sogenanntes *Randextremum*). Außerdem hat die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = |x|$ an der Stelle 0 ein (ebenfalls sogar globales) Minimum, aber x_0 ist kein kritischer Punkt (da $f'(0)$ überhaupt nicht existiert).

Über das Auftreten kritischer Punkte gibt folgender Satz Auskunft

Satz 13.4 *(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 (a,b). Gilt $f(a) = f(b)$ so existiert ein $\xi \in (a,b)$ mit $f'(\xi) = 0$.

Beweis. Ist $f(x) \equiv f(a)$ so ist $f'(x) \equiv 0$ auf $[a,b]$.
Es sei also $f \not\equiv f(a)(= f(b))$. Nach S. 11.17 existieren $x_1, x_2 \in [a,b]$ mit $f(x_1) \leq f(x) \leq f(x_2)$ für alle $x \in [a,b]$. Dabei liegt (mindestens) einer der beiden Punkte x_1, x_2 in (a,b). Nach S. 13.2 liegt dort ein kritischer Punkt vor. □

Als Folgerung erhalten wir

Satz 13.5 *(erweiterterter Mittelwertsatz)*
Es sei $I = [a,b]$ ein kompaktes Intervall, und $f, g : I \to \mathbb{R}$ seien stetig auf I und differenzierbar auf (a,b). Dann existiert ein $\xi \in (a,b)$ mit

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

Beweis. Man wende den Satz von Rolle auf die Funktion $\varphi : I \to \mathbb{R}$ mit

$$\varphi(x) := f(x)\big(g(b) - g(a)\big) - \big(f(b) - f(a)\big)g(x) \qquad (x \in I)$$

an! (Beachte: es gilt $\varphi(b) = \varphi(a) = f(a)g(b) - f(b)g(a)$.) □

Wir ziehen nun wichtige Folgerungen aus dem erweiterten Mittelwertsatz.

Satz 13.6 *Es sei $I = [a, b]$ ein kompaktes Intervall, und $f : I \to \mathbb{K}$ sei stetig auf I und differenzierbar auf (a, b).*

1. *(klassischer Mittelwertsatz) Ist $\mathbb{K} = \mathbb{R}$, so existiert ein $\xi \in (a, b)$ mit*
$$f'(\xi) = \frac{f(b) - f(a)}{b - a} .$$
2. *(Schrankensatz) Es existiert ein $\xi \in (a, b)$ mit*
$$|f(b) - f(a)| \leq |f'(\xi)| \cdot (b - a) .$$

Beweis. 1. Ergibt sich sofort aus dem erweiterten Mittelwertsatz mit $g(x) = x$.
2. Im Falle $\mathbb{K} = \mathbb{R}$ folgt dies sofort aus 1. Der Fall $\mathbb{K} = \mathbb{C}$ kann darauf zurückgeführt werden ([Ü]). □

Im Weiteren sei für ein Intervall I stets I^0 das Intervall I ohne die Randpunkte.

Bemerkung 13.7 Es sei I ein Intervall und $f : I \to \mathbb{K}$ sei stetig auf I und differenzierbar auf I^0. Dann ist f genau dann konstant, wenn $f'(x) \equiv 0$ auf I^0 gilt.
(Denn:
"$\Rightarrow$" Ist $f(x) \equiv c$ auf I, so gilt natürlich $f'(x) \equiv 0$ auf I^0.
"$\Leftarrow$" Sind $x_1, x_2 \in I$ mit $x_1 < x_2$, so folgt aus dem Schrankensatz mit einem $\xi \in (x_1, x_2)$
$$|f(x_1) - f(x_2)| = |f'(\xi)| \cdot (x_2 - x_1) = 0$$
und damit $f(x_1) = f(x_2)$).

Satz 13.8 *Es sei I ein Intervall und $f : I \to \mathbb{R}$ sei stetig auf I und differenzierbar auf I^0.*
1. Ist $f'(x) \geq 0$ (bzw. > 0) für alle $x \in I^0$, so ist f monoton wachsend (bzw. streng monoton wachsend) auf I.

2. Ist $f'(x) \leq 0$ (bzw. < 0) für alle $x \in I^0$, so ist f monoton fallend (bzw. streng monoton fallend) auf I.

Beweis. 1. Sind $x_1, x_2 \in I$ mit $x_1 < x_2$, so ergibt sich aus dem Mittelwertsatz mit einem $\xi \in (x_1, x_2)$
$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1) \geq 0 ,$$

also $f(x_1) \le f(x_2)$. Ist dabei $f'(\xi) > 0$, so ist $f(x_1) < f(x_2)$.
2. Entsprechend. $\square$

Beispiel 13.9 Es sei $f : (0, \infty) \to \mathbb{R}$ mit

$$f(x) := \left(1 + \frac{1}{x}\right)^x \qquad (x > 0)\ .$$

Dann ist f streng monoton wachsend.
(Denn: Es genügt zu zeigen: $g(x) := \ln f(x) = x \cdot \ln(1 + 1/x)$ ist streng monoton wachsend. Es gilt für $x > 0$

$$g'(x) = \ln(1 + 1/x) + x \cdot \frac{1}{1 + 1/x} \cdot \frac{-1}{x^2} = \ln(1 + 1/x) - \frac{1}{1 + x}\ .$$

Aus $e^t < 1/(1-t)$ für $0 < t < 1$ (vgl. S. 9.2.2) folgt

$$e^{\frac{1}{1+x}} < \left(1 - \frac{1}{1+x}\right)^{-1} = 1 + 1/x$$

und damit $g'(x) > 0$ für $x > 0$. Mit S. 13.8 folgt die Behauptung.)

Der folgende Satz gibt ein **hinreichendes** Kriterien für Extremstellen.

Satz 13.10 *(Vorzeichenwechsel-Kriterium)*
Die Funktion $f : I \to \mathbb{R}$ sei stetig auf dem Intervall I und differenzierbar auf I^0. Ferner sei $x_0 \in I$.

1. Existiert ein $\delta > 0$, so dass

$$f'(x) \begin{cases} \le 0 \ (\textit{bzw.} <) & \textit{für} \quad x \in I,\ x_0 < x < x_0 + \delta \\ \ge 0 \ (\textit{bzw.} >) & \textit{für} \quad x \in I,\ x_0 - \delta < x < x_0 \end{cases},$$

so hat f an x_0 ein lokales Maximun (bzw. stiktes lokales Maximun).

2. Existiert ein $\delta > 0$, so dass

$$f'(x) \begin{cases} \ge 0 \ (\textit{bzw.} >) & \textit{für} \quad x \in I,\ x_0 < x < x_0 + \delta \\ \le 0 \ (\textit{bzw.} <) & \textit{für} \quad x \in I,\ x_0 - \delta < x < x_0 \end{cases},$$

So hat f an x_0 ein lokales Minimum (bzw. stiktes lokales Minimun).

Beweis.
1. Nach S. 13.8 ist f (streng) monoton fallend auf $I \cap [x_0, x_0+\delta)$ und (streng) monoton wachsend auf $I \cap (x_0-\delta, x_0]$. Damit hat f an x_0 ein (stiktes) lokales Maximum.
2. Analog. □

Beispiel 13.11 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 $\left\{ \begin{array}{l} (-\infty, -1] \\ [-1, 0] \\ [0, \infty) \end{array} \right.$.

Nach S. 13.10 liegt an 0 ein striktes lokales Minimum und and -1 ein striktes lokales Maximum vor.

Bemerkung 13.12 Es seien $I \subset \mathbb{R}$ ein offenes Intervall und $f : I \to \mathbb{R}$ so, dass f'' auf I existiert. Ferner sei $x_0 \in I$ so, dass f'' stetig an x_0 ist mit $f'(x_0) = 0$. Dann gilt: Ist $f''(x_0) > 0$, so hat f an x_0 ein lokales Minimum, und ist $f''(x_0) < 0$, so hat f an x_0 ein striktes lokales Maximum.
(Denn: Es sei $f''(x_0) > 0$. Da f'' stetig an x_0 ist, existiert ein $\delta > 0$ mit $f''(x) > 0$ für alle $x \in U_\delta(x_0)$. Damit ist f' nach S. 13.8 streng monoton wachsend und $U_\delta(x_0)$. Mit $f'(x_0) = 0$ folgt

$$f'(x) \begin{cases} > 0 & \text{für} \quad x_0 < x < x_0 + \delta \\ < 0 & \text{für} \quad x_0 - \delta < x < x_0 \end{cases} .$$

Nach dem Vorzeichenwechsel-Kriterium hat f an x_0 ein stiktes lokales Minimum. Entsprechendes gilt, falls $f''(x_0) < 0$ ist.)

Zum Abschluss dieses Abschnitts beschäftigen wir uns mit einer eleganten Methode um gewisse Grenzwerte auszurechnen. Damit beweisen wir

Satz 13.13 *(Regeln von de l'Hospital)*
Es seien $I \subset \mathbb{R}$ ein Intervall, $x_0 \in I$ und die Funktionen $f, g : I \setminus \{x_0\} \to \mathbb{R}$ seien differenzierbar mit $g'(x) \neq 0$ für alle $x \in I$, $x \neq x_0$. Ferner gelte

1. $\lim\limits_{x \to x_0} f(x) = \lim\limits_{x \to x_0} g(x) = 0$ *(Fall "$\frac{0}{0}$")*

oder

2. $\lim\limits_{x \to x_0} g(x) = \infty$ *(oder $-\infty$).*

Dann gilt: Aus

$$\frac{f'(x)}{g'(x)} \to c \in \mathbb{R} \cup \{\pm\infty\} \qquad (x \to x_0)$$

folgt

$$\frac{f(x)}{g(x)} \to c \qquad (x \to x_0)\ .$$

Entsprechende Aussagen gelten für $x \to \pm\infty$.

Beweis. Wir betrachten nur den Fall $c \in \mathbb{R}$ und $x \to x_0^+$. Die Beweise für die anderen Fälle verlaufen ähnlich.
1. Es gelte die Voraussetzung 1. Durch $g(x_0) := f(x_0) := 0$ können f und g stetig nach I fortgesetzt werden.
Ist $x \in I$, $x > x_0$, so existiert nach S. 13.5 ein $\xi(x) \in (x_0, x)$ mit

$$\frac{f'(\xi(x))}{g'(\xi(x))} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f(x)}{g(x)}\ .$$

Da $\lim\limits_{x \to x_0^+} \xi(x) = x_0$ und $\xi(x) > x_0$ gilt, ist

$$\lim_{x \to x_0^+} \frac{f(x)}{g(x)} = \lim_{x \to x_0^+} \frac{f'(\xi(x))}{g'(\xi(x))} = c\ .$$

2. Es gelte o.E. $g(x) \to \infty$.
Weiter sei $\varepsilon > 0$ gegeben. Dann existiert ein $\delta_0 > 0$ mit

$$|f'(t)/g'(t) - c| < \varepsilon \qquad (x_0 < t < x_0 + \delta)\ .$$

Es sei $x_0 < s < x_0 + \delta_0$. Dann existiert ein $\delta \in (0, s)$ so, dass $g(x) > 0$ und

$$|f(s)|/g(x) < \varepsilon \quad \text{sowie} \quad |g(s)|/g(x) < \varepsilon \quad (x_0 < x < x_0 + \delta)\ .$$

Nach dem erweiterten Mittelwertsatz existiert zu jedem solchen x ein $\xi \in (x,s)$ mit

$$f'(\xi)(g(s)-g(x)) = (f(s)-f(x))g'(\xi) \ ,$$

also auch

$$\frac{f'(\xi)}{g'(\xi)}\left(\frac{g(s)}{g(x)}-1\right) = \frac{f(s)-f(x)}{g(x)},$$

und folglich

$$\frac{f(x)}{g(x)} = \frac{f(s)}{g(x)} + \frac{f'(\xi)}{g'(\xi)}\left(1-\frac{g(s)}{g(x)}\right) \ .$$

Damit erhalten wir

$$\left|\frac{f(x)}{g(x)}-c\right| \le \left|\frac{f(s)}{g(x)}\right| + \left|\frac{f'(\xi)}{g'(\xi)}-c\right| + \left|\frac{f'(\xi)}{g'(\xi)}\frac{g(s)}{g(x)}\right| \le \varepsilon+\varepsilon+(|c|+\varepsilon)\varepsilon \ .$$

Da $\varepsilon > 0$ beliebig war, folgt die Behauptung. □

Beispiel 13.14 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 nach S. 13.13 auch

$$\lim_{x\to 1^+}\frac{f(x)}{g(x)} = \lim_{x\to 1^+}\frac{\sqrt{x}-1}{\sqrt{x-1}} = 0 \ .$$

2. 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. 13.13)

14 Funktionenfolgen und Funktionenreihen

Wir haben schon früher gesehen, dass wichtige elementare Funktionen wie die Exponentialfunktion 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.

Definition 14.1 Es sei X eine beliebige Menge, $X \neq \emptyset$, und es sei (Y, d) ein metrischer Raum. Sind $f_n : X \to Y$ 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 Y konvergiert. Die Funktion $f : M \to Y$ mit $f(x) := \lim_{n\to\infty} f_n(x)$ heißt *Grenzfunktion* der Folge (f_n) (auf M). Wir schreiben dann auch

$$f_n \to f \qquad (n \to \infty) \quad \text{punktweise auf } M$$

Beispiel 14.2 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

$$f_n(x) \to \begin{cases} 0 & , \quad \text{falls} \quad x \in (-1, 1) \\ 1 & , \quad \text{falls} \quad x = 1 \end{cases} \quad ,$$

d. h. (f_n) konvergiert punktweise auf $(-1, 1]$ und die Grenzfunktion $f : (-1, 1] \to \mathbb{R}$ ist gegeben durch

$$f(x) = \begin{cases} 0 & , \quad x \in (-1, 1) \\ 1 & , \quad x = 1 \end{cases} \quad .$$

Das Beispiel zeigt insbesondere, dass die Grenzfunktion unstetig (an $x_0 = 1$) ist, obwohl alle Folgelieder f_n stetige Funktionen 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 14.3 Es seien $X \neq \emptyset$ eine beliebige Menge und (Y, d) ein metrischer Raum.
1. Für $f, g : X \to Y$ setzen wir

$$d_M(f, g) := \sup_{x\in M} d(f(x), g(x))$$

(wobei $\sup A := \infty$, falls A nach oben unbeschränkt ist).
2.Sind $f_n : X \to Y$ Funktionen ($n \in \mathbb{N}$), so heißt die Funktionenfolge (f_n) *gleichmäßig konvergent* auf der Menge $M \subset X$ (gegen die Grenzfunktion $f : M \to Y$), falls

$$d_M(f, f_n) \to 0 \qquad (n \to \infty).$$

Wir schreiben dann

$$f_n \to f \qquad (n \to \infty) \quad \text{gleichmäßig auf } M$$

oder auch

$$f_n(x) \to f(x) \qquad (n \to \infty) \quad \text{gleichmäßig auf } M\ .$$

Bemerkung 14.4 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. **gleichmäßige Konvergenz impliziert punktweise Konvergenz**.
Gleichmäßige Konvergenz bedeutet, dass zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ existiert mit

$$d(f(x), f_n(x)) < \varepsilon \qquad \text{für alle } n \geq N_\varepsilon \text{ und alle } x \in M\ .$$

Der Unterschied zur punktweisen Konvergenz liegt darin, dass 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 $d(f(x), f_n(x)) < \varepsilon$ für alle $n \geq N_{\varepsilon,x}$.

Beispiel 14.5 Wir betrachten noch einmal $f_n(x) = x^n\ (n \in \mathbb{N},\ x \in \mathbb{R})$ (vgl. B. 14.2). Ist $M = [-1/2, 1/2]$, so gilt

$$\sup_M |f_n(x) - f(x)| = 1/2^n \to 0 \qquad (n \to \infty)\ .$$

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 $[0, 1)$.

Wir kommen nun zu dem bereits angedeuteten Ergebnis über die „Vererbung“ der Stetigkeit auf die Grenzfunktion.

Definition 14.6 Sind (X, d) ein metrischer Raum und $x_0 \in X$, so heißt $M \subset X$ eine *Umgebung* von x_0, falls ein $\delta > 0$ existiert mit $U_\delta(x_0) \subset M$.

Satz 14.7 *Es seien $(X, d = d_X)$, $(Y, d = d_Y)$ metrische Räume und $f, f_n : X \to Y$ Funktionen. Ferner sei $x_0 \in X$. Sind die Funktion f_n stetig an der Stelle x_0 und gilt $f_n \to f$ gleichmäßig auf einer Umgebung M von x_0, so ist auch die Grenzfunktion f stetig an x_0.*

Beweis. Es sei $\varepsilon > 0$ gegeben. Auf Grund der gleichmäßigen Konvergenz von (f_n) gegen f auf M existiert ein $N = N_\varepsilon \in \mathbb{N}$ mit

$$d(f(x), f_n(x)) < \varepsilon/3 \qquad (n \geq N,\ x \in M)\ .$$

Da f_N stetig an x_0 ist, existiert ein $\delta = \delta_\varepsilon > 0$ so, dass ($U_\delta(x_0) \subset M$ und)

$$d(f_N(x), f_N(x_0)) < \varepsilon/3 \qquad \text{für alle } x \in U_\delta(x_0)\ .$$

Damit ist für $x \in U_\delta(x_0)$

$$\begin{aligned} d(f(x), f(x_0)) &\leq d(f(x), f_N(x)) + d(f_N(x), f_N(x_0)) + d(f_N(x_0), f(x_0)) \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon\ . \end{aligned}$$

□

Ein notwendiges und hinreichendes Kriterium für gleichmäßige Konvergenz liefert

Satz 14.8 *(Cauchy-Kriterium für gleichmäßige Konvergenz)*
Es seien $X \neq \emptyset$ eine Menge, (Y, d) ein ***vollständiger*** *metrischer Raum und $f_n : X \to Y$ Funktionen ($n \in \mathbb{N}$). Ist $M \subset X$, so ist (f_n) gleichmäßig konvergent auf M genau dann, wenn (f_n) eine gleichmäßige Cauchy-Folge auf M ist, d. h. wenn zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ so existiert, dass*

$$d_M(f_n, f_m) < \varepsilon \qquad \textit{für alle } n, m \geq N_\varepsilon.$$

Beweis. 1. Es gelte $f_n \to f$ gleichmäßig auf M. Ist $\varepsilon > 0$ gegeben, so existiert ein $N_\varepsilon \in \mathbb{N}$ mit

$$d_M(f, f_n) < \frac{\varepsilon}{2} \qquad (n \geq N_\varepsilon)\ .$$

Für alle $n, m \geq N_\varepsilon$ und $x \in M$ folgt dann

$$d(f_n(x), f_m(x)) \leq d(f_n(x), f(x)) + d(f(x), f_m(x)) \leq d_M(f_n, f) + d(f, f_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\ ,$$

also auch $d_M(f_n, f_m) \leq \varepsilon$.

2. Es sei (f_n) eine gleichmäßige Cauchy-Folge auf M. Dann ist insbesondere für jedes feste $x \in M$ die Folge $(f_n(x))_n$ eine Cauchy-Folge in (Y, d). Da (Y, d) vollständig ist, ist $(f_n(x))_n$ konvergent, d. h. es existiert ein $y \in Y$ mit $f_n(x) \to y$ $(n \to \infty)$. Wir definieren $f : M \to Y$ durch $f(x) := y\,(x \in M)$ und zeigen: $f_n \to f$ gleichmäßig auf M.
Zunächst gilt für $z \in Y$ und (y_m) in Y mit $y_m \to y$ nach der umgekehrten Dreiecksungleichung ([Ü]) $d(y_m, z) \to d(y, z)$ für $m \to \infty$. Damit ergibt für festes $x \in M$ und $n \in \mathbb{N}$

$$d_M(f_m, f_n) \geq d(f_m(x), f_n(x)) \to d(f(x), f_n(x)) \qquad (m \to \infty)\ .$$

Es sei $\varepsilon > 0$ gegeben. Dann existert ein $N = N_\varepsilon \in \mathbb{N}$ mit $d_M(f_m, f_n) < \varepsilon$ für $n, m \geq N$. Also ist $d(f(x), f_n(x)) \leq \varepsilon$ für alle $x \in M$ und $n \geq N$ und damit auch

$$d_M(f, f_n) \leq \varepsilon \qquad (n \geq N).$$

□

Bemerkung 14.9 Ist $M \neq \emptyset$ eine Menge so gilt für Funktionen $f, g \in B(M, \mathbb{K}^m)$ (vgl. B. 10.7)

$$||f - g||_\infty = \sup_{x \in M} |f(x) - g(x)| = \sup_{x \in M} d(f(x), g(x)) = d_M(f, g)\ .$$

Also bedeutet gleichmäßige Konvergenz in diesem Fall Konvergenz bezüglich der sup-Norm $||\cdot||_\infty$, d. h. der Metrik $d_{||\cdot||_\infty}$).
Aus S. 14.8 folgt, dass der Raum $(B(M, \mathbb{K}^m), ||\cdot||_\infty)$ ein vollständiger normierter Raum (ein sog. Banachraum) ist.
(Denn: Es sei (f_n) eine Cauchy-Folge in $B(M, \mathbb{K}^m)$. Dann ist (f_n) auch gleichmäßige Cauchy-Folge auf M. Nach S. 14.8 existiert eine Funktion $f : M \to \mathbb{K}^m$ mit $f_n \to f$ gleichmäßig auf M, d. h. $||f - f_n||_\infty \to 0$. 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

$$||f - f_N||_\infty < 1\ .$$

Dann gilt für alle $x \in M$

$$|f(x)| \leq |f(x) - f_N(x)| + |f_N(x)| \leq 1 + ||f_N||_\infty$$

und damit ist f beschränkt.)

Definition 14.10 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 die Funktionenfolge (s_n) auf M punktweise konvergiert. Entsprechend heißt die Funktionenreihe *gleichmäßig konvergent* auf M, falls (s_n) gleichmäßig auf M konvergiert. Wir verwenden (wie bei Zahlenreihen) das Symbol $\sum\limits_{\nu=0}^{\infty} a_\nu$ dann auch wieder für die Grenzfunktion. Natürlich definiert man entsprechend Funktionenreihen der Form $\sum\limits_{\nu=\nu_0}^{\infty} a_\nu(x)$ auch für allgemeine $\nu_0 \in \mathbb{Z}$.

Es stellt sich die Frage, wie man gleichmäßige Konvergenz ggfs. nachweisen kann. Wie beim Majorantenkriterium für Zahlenreihen (S. 7.11) ergibt sich

Satz 14.11 *(Weierstraßsches Majorantenkriterium für Funktionenreihen)*
Es seien $M \neq \emptyset$ eine Menge und $a_\nu : M \to \mathbb{K}$ für $\nu \in \mathbb{N}_0$. Ferner seien $b_\nu \geq 0$ so, dass $||a_\nu||_\infty = \sup\limits_{x \in M} |a_\nu(x)| \leq b_\nu$ für ν genügend groß. Dann gilt: Ist $\sum\limits_{\nu=0}^{\infty} b_\nu < \infty$, so konvergiert $\sum\limits_{\nu=0}^{\infty} a_\nu$ gleichmäßig auf M.

Beweis. Es sei $\varepsilon > 0$ gegeben. Dann existiert nach S. 7.10 ein $N = N_\varepsilon \in \mathbb{N}$ mit ($||a_\nu||_\infty \leq b_\nu$ für $\nu \geq N$ und)

$$\sum_{\nu=m+1}^{n} b_\nu < \varepsilon \qquad (n > m \geq N).$$

Damit ergibt sich für $n > m \geq N$ auch

$$d_M(s_n, s_m) = ||s_n - s_m||_\infty = ||\sum_{\nu=m+1}^{n} a_\nu||_\infty \leq \sum_{\nu=m+1}^{n} ||a_\nu||_\infty \leq \sum_{\nu=m+1}^{n} b_\nu < \varepsilon.$$

Also ist (s_n) eine gleichmäßige Cauchy-Folge auf M. Mit S. 14.8 folgt die Behauptung.
□

Beispiel 14.12 Es seien $a_\nu : \mathbb{C} \to \mathbb{C}$ definiert durch

$$a_\nu(z) = \frac{1}{\nu^z} \qquad (z \in \mathbb{C},\, \nu \in \mathbb{N})\ .$$

Für $\alpha > 1$ sei weiter

$$M_\alpha := \{z \in \mathbb{C} : \operatorname{Re} z \geq \alpha\}\ .$$

Dann ist die Funktionenreihe $\sum\limits_{\nu=1}^{\infty} a_\nu$ gleichmäßig konvergent auf M_α.
(Denn: Für alle $z \in M_\alpha$ und alle $\nu \in \mathbb{N}$ ist

$$|a_\nu(z)| = \frac{1}{\nu^{\operatorname{Re} z}} \leq \frac{1}{\nu^\alpha}$$

und $\sum\limits_{\nu=1}^{\infty} 1/\nu^\alpha$ ist konvergent. Also ergibt sich die Behauptung aus S. 14.11.)
Insbesondere ist die Funktionenreihe damit auf $M := \{z \in \mathbb{C} : \operatorname{Re} z > 1\} = \bigcup_{\alpha > 1} M_\alpha$ punktweise konvergent. Wir definieren $\zeta : M \to \mathbb{C}$ durch

$$\zeta(z) := \sum_{\nu=1}^{\infty} \frac{1}{\nu^z} \qquad (z \in M)\ .$$

Die Funktion ζ heißt *(Riemannsche) Zetafunktion*.
Da $z \to 1/\nu^z$ für alle $\nu \in \mathbb{N}$ stetig auf $\mathbb{C}$ ist, folgt aus S. 14.7 die Stetigkeit der Zetafunktion auf M (Man beachte: Ist $\operatorname{Re}(z_0) > 1$, so ist M_α für $1 < \alpha < \operatorname{Re}(z_0)$ eine Umgebung von z_0).

15 Potenzreihen

Wir untersuchen jetzt eine Klasse besonders wichtiger Funktionenreihen, sog. Potenzreihen. Hier sind die Teilsummen s_n Polynome vom Grad $\leq n$, also von besonders einfacher Struktur.

Definition 15.1 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) gegeben durch $s_n(z) = \sum\limits_{\nu=0}^{n} a_\nu (z-z_0)^\nu$) auf $\mathbb{K}$ eine *Potenzreihe* (mit der *Entwicklungsmitte* z_0 und der *Koeffizientenfolge* (a_ν)).

2. Ist $\sum\limits_{\nu=0}^{\infty} a_\nu (z - z_0)^\nu$ eine Potenzreihe, so heißt

$$R := \frac{1}{\overline{\lim\limits_{\nu\to\infty}} \sqrt[\nu]{|a_\nu|}} \in [0, \infty]$$

Konvergenzradius der Potenzreihe (wobei $1/\infty := 0$ und $1/0 := \infty$ 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 15.2 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 $R = 1$. Hier konvergiert $\sum\limits_{\nu=0}^{\infty} z^\nu$ absolut im Konvergenzkreis $U_1(0) = \{z : |z| < 1\}$ und divergiert für alle z mit $|z| > 1$. Ferner gilt

$$\sum_{\nu=0}^{\infty} z^\nu = \frac{1}{1-z} \quad (|z| < 1) .$$

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 $\overline{\lim\limits_{\nu=\infty}} |a_\nu|^{1/\nu} = \lim\limits_{\nu\to\infty} \frac{1}{\sqrt[\nu]{\nu!}} = 0$ (warum?), d. h. $R = \infty$. Bekanntlich konvergiert die Potenzreihe absolut für alle $z \in \mathbb{C}$ und es gilt

$$\sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} = e^z .$$

Über das Konvergenzverhalten allgemeiner Potenzreihen gibt der folgende Satz Auskunft.

Satz 15.3 *(Cauchy-Hadamard)*
Es sei $\sum_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu$ *eine Potenzreihe mit Konvergenzradius* R*. Dann gilt:*

1. *Ist* $R = \infty$*, so konvergiert die Potenzreihe absolut für alle* $z \in \mathbb{K} = U_\infty(z_0)$*.*

2. *Ist* $0 < R < \infty$*, so konvergiert die Potenzreihe absolut für alle* $z \in U_R(z_0)$ *und divergiert für alle* $z \in \mathbb{K}$ *mit* $|z-z_0| > R$*.*

3. *Ist* $R = 0$*, so divergiert die Potenzreihe für alle* $z \in \mathbb{K} \setminus \{z_0\}$*.*

Beweis. Es gilt für alle $\nu \in \mathbb{N}_0$ und $z \neq z_0$

$$|a_\nu (z-z_0)^\nu|^{1/\nu} = |a_\nu|^{1/\nu} \cdot |z-z_0| ,$$

also ist im Falle $R > 0$

$$\overline{\lim_{\nu\to\infty}} |a_\nu (z-z_0)^\nu|^{1/\nu} = \frac{1}{R} |z-z_0| .$$

Damit ergeben sich 1. und 2. sofort aus dem Wurzelkriterium (S. 7.15).
Ist $R = 0$ (d.h. $(|a_\nu|^{1/\nu})$ ist unbeschränkt), so ist auch $(|a_\nu (z-z_0)^\nu|^{1/\nu})$ unbeschränkt und folglich ist $(a_\nu (z-z_0)^\nu)_\nu$ insbesondere keine Nullfolge. Damit ergibt sich auch 3.
□

Bemerkung 15.4 Ist $\sum_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu$ eine Potenzreihe mit Konvergenzradius $R > 0$, so konvergiert $\sum_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu$ gleichmäßig auf jeder kompakten Teilmenge des Konvergenzkreises $U_R(z_0)$.
(Denn: Ist $K \subset U_R(z_0)$ kompakt, so hat die stetige Funktion $z \mapsto |z-z_0|$ nach S. 11.17 ein Maximum auf K, d. h. es gibt ein $z_1 \in K$ mit $|z_1 - z_0| = \sup_{z\in K} |z-z_0|$. Damit gilt für alle ν

$$\sup_{z\in K} |a_\nu (z-z_0)^\nu| = |a_\nu| \cdot |z_1 - z_0|^\nu =: b_\nu .$$

Da nach S. 15.3 die Reihe $\sum_{\nu=0}^{\infty} b_\nu$ konvergiert, folgt die gleichmäßige Konvergenz auf K aus dem Weierstraßschen Majorantenkriterium.)
Aus S. 14.7 folgt, dass die Funktion $f = f_{(a_\nu)} : U_R(z_0) \to \mathbb{K}$ mit

$$f(z) := \sum_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu \qquad (z \in U_R(z_0))$$

stetig ist. Wir zeigen nun, dass diese Funktion tatsächlich beliebig oft differenzierbar ist.

Satz 15.5 *Es sei* $\sum_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu$ *eine Potenzreihe mit Konvergenzradius* $R > 0$, *und es sei* $f(z) := \sum_{\nu=0}^{\infty} a_\nu (z-z_0)^\nu$ *für* $z \in U_R(z_0)$. *Dann existieren die Ableitungen* $f^{(k)}$ *für alle* $k \in \mathbb{N}$ *und es gilt*

$$f^{(k)}(z) = \sum_{\nu=0}^{\infty} (\nu+1)(\nu+2)\dots(\nu+k) a_{\nu+k}(z-z_0)^\nu \qquad (z \in U_R(z_0)) .$$

Insbesondere ist

$$k! a_k = f^{(k)}(z_0) \qquad (k \in \mathbb{N}_0) ,$$

also

$$f(z) = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(z_0)}{\nu!}(z-z_0)^\nu \qquad (z \in U_R(z_0)) .$$

Beweis. 1. Wir zeigen: f ist differenzierbar auf $U_R(z_0)$ und

$$f'(z) = \sum_{\nu=1}^{\infty} \nu a_\nu (z-z_0)^{\nu-1} = \sum_{\nu=0}^{\infty} (\nu+1) a_{\nu+1}(z-z_0)^\nu \qquad (z \in U_R(z_0)) .$$

O. E. können wir $z_0 = 0$ annehmen.
Es sei $z_1 \in U_R(0)$ fest. Für $z \in U_R(0)$ gilt

$$f(z) - f(z_1) = \sum_{\nu=1}^{\infty} a_\nu (z^\nu - z_1^\nu) = (z - z_1) \sum_{\nu=1}^{\infty} a_\nu \underbrace{\sum_{k=0}^{\nu-1} z_1^k z^{\nu-1-k}}_{=: \phi_\nu(z)} .$$

Wir wählen ein $r \in (|z_1|, R)$ und setzen $\delta := r - |z_1|$. Ist $|z - z_1| < \delta$, so ist $|z| < r$ (und $|z_1| < r$), also $|\phi_\nu(z)| \le \nu r^{\nu-1}$. Aus $\nu^{1/\nu} \to 1$ $(\nu \to \infty)$ folgt, dass auch die Potenzreihe $\sum_{\nu=1}^{\infty} \nu a_\nu z^\nu$ den Konvergenzradius R hat. Insbesondere konvergiert damit $\sum_{\nu=1}^{\infty} \nu |a_\nu| r^{\nu-1}$ nach dem Satz von Cauchy-Hadamard. Nach dem Weierstraßschen Majorantenkriterium (S. 14.11) konvergiert $\sum_{\nu=1}^{\infty} a_\nu \phi_\nu(z)$ gleichmäßig auf $U_\delta(z_1)$. Folglich ist $\phi(z) := \sum_{\nu=1}^{\infty} a_\nu \phi_\nu(z)$ stetig an z_1 (S. 14.7), d. h. es gilt

$$\lim_{z \to z_1} \frac{f(z) - f(z_1)}{z - z_1} = \lim_{z \to z_1} \phi(z) = \phi(z_1) = \sum_{\nu=1}^{\infty} \nu a_\nu z_1^{\nu-1} .$$

2. Durch Anwendung der gleichen Argumentation auf f' sieht man: f' ist differenzierbar auf $U_R(z_0)$ und

$$f''(z) = \sum_{\nu=0}^{\infty} (\nu+1)(\nu+2)a_{\nu+2}(z-z_0)^\nu \qquad (z \in U_R(z_0))\,.$$

Induktiv erhält man: Für alle $k \in \mathbb{N}$ ist $f^{(k-1)}$ differenzierbar auf $U_R(z_0)$ und

$$f^{(k)}(z) = \sum_{\nu=0}^{\infty} (\nu+1)(\nu+2)\dots(\nu+k)a_{\nu+k}(z-z_0)^\nu\,.$$

Die Zusatzbehauptung $k!a_k = f^{(k)}(z_0)$ ergibt sich für $z = z_0$. □

Beispiel 15.6 Für die Funktion $f : \mathbb{C} \to \mathbb{C}$ mit

$$f(z) = \begin{cases} \dfrac{\sin z}{z} & , \; z \neq 0 \\ 1 & , \; z = 0 \end{cases}$$

gilt

$$f(z) = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu z^{2\nu}}{(2\nu+1)!} = \sum_{\substack{\mu=0 \\ \mu \text{ gerade}}}^{\infty} \frac{(-1)^{\mu/2}}{(\mu+1)!} z^\mu\,.$$

(Denn: Es gilt für $z \neq 0$

$$f(z) = \frac{1}{z}\sum_{\nu=0}^{\infty} \frac{(-1)^\nu z^{2\nu+1}}{(2\nu+1)!} = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu z^{2\nu}}{(2\nu+1)!}$$

und für $z = 0$

$$f(0) = 1 = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu 0^{2\nu}}{(2\nu+1)!}\,.)$$

Also ist f nach S. 15.5 beliebig oft differenzierbar auf $\mathbb{C}$ mit

$$f^{(k)}(0) = \begin{cases} \frac{(-1)^{k/2}}{k+1} & \text{falls } k \text{ gerade} \\ 0 & \text{falls } k \text{ ungerade} \end{cases}\,.$$

Am Rande ihres Konvergenzkreises können Potenzreihen ein sehr kompliziertes Verhalten zeigen. Da ist es schon sehr beruhigend, ggfs. auf folgendes Resultat zurückgreifen zu können. (Die Beschränkung auf die Entwicklungsmitte $z_0 = 0$ und den Konvergenzradius 1 ist dabei unwesentlich.)

Satz 15.7 *(Abelscher Grenzwertsatz)*
Es sei $\sum\limits_{\nu=0}^{\infty} a_\nu z^\nu$ *eine Potenzreihe mit Konvergenzradius* 1*. Ferner sei die Reihe* $\sum\limits_{\nu=0}^{\infty} a_\nu$ *konvergent. Ist*

$$f(z) := \sum_{\nu=0}^{\infty} a_\nu z^\nu \qquad (|z| < 1) \ ,$$

so gilt

$$\lim_{x \to 1^-} f(x) = \sum_{\nu=0}^{\infty} a_\nu \ .$$

Beweis. Wir setzen $s := \sum\limits_{\nu=0}^{\infty} a_\nu$ und $s_n := \sum\limits_{\nu=0}^{n} a_\nu$ für $n \in \mathbb{N}_0$. Da (s_n) konvergent und damit auch beschränkt ist, hat die Potenzreihe $\sum\limits_{\nu=0}^{\infty} s_\nu z^\nu$ Konvergenzradius ≥ 1. Also gilt mit $s_{-1} := 0$ für $|z| < 1$

$$(1-z)\sum_{\nu=0}^{\infty} s_\nu z^\nu = \sum_{\nu=0}^{\infty} s_\nu z^\nu - \sum_{\nu=0}^{\infty} s_\nu z^{\nu+1} = \sum_{\nu=0}^{\infty}(s_\nu - s_{\nu-1})z^\nu = \sum_{\nu=0}^{\infty} a_\nu z^\nu = f(z) \ .$$

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 $0 < x < 1$

$$\begin{aligned} |f(x) - s| &= |(1-x)\sum_{\nu=0}^{\infty}(s_\nu - s)x^\nu| \leq \\ &\leq (1-x)\sum_{\nu=0}^{N-1} |s_\nu - s| x^\nu + \frac{\varepsilon}{2}(1-x)\sum_{\nu=N}^{\infty} x^\nu \\ &\leq (1-x)\sum_{\nu=0}^{N-1} |s_\nu - s| x^\nu + \varepsilon/2 \end{aligned}$$

Weiter existiert ein $\delta = \delta(N_\varepsilon) > 0$ so, dass

$$(1-x)\sum_{\nu=0}^{N-1} |s_\nu - s| < \varepsilon/2$$

für alle x mit $1 - \delta < x < 1$. Also gilt

$$|f(x) - s| < \varepsilon \qquad \text{für} \qquad 1 - \delta < x < 1 \ .$$

□

Beispiel 15.8 Es gilt ([Ü)])

$$\ln(1+x) = \sum_{\nu=1}^{\infty} \frac{(-1)^{\nu-1}}{\nu} \, x^{\nu} \qquad \big(x \in (-1,1)\big).$$

Nach dem Leibniz-Kriterium und dem Abelschen Grenzwertsatz ergibt sich damit

$$\ln(2) = \lim_{x \to 1^-} \ln(1+x) = \sum_{\nu=1}^{\infty} \frac{(-1)^{\nu-1}}{\nu}$$

16 Integrale und Stammfunktionen

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 (orientierte) „Fläche unter den Graphen“ sehr leicht definieren können.

Definition 16.1 Es sei $I = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall.

1. Sind $I_1, \ldots, I_m \subset [a, b]$ (nichtleere) Intervalle, so heißt
$$Z = \{I_1, \ldots, I_m\}$$
eine Zerlegung von $[a, b]$, falls die I_j paarweise disjunkt sind und $\bigcup\limits_{j=1}^{m} I_j = [a, b]$ gilt.

2. Eine Funktion $\varphi : [a, b] \to \mathbb{K}$ heißt *Treppenfunktion*, wenn eine Zerlegung Z von $[a, b]$ und $c_I \in \mathbb{K}$ existieren mit
$$\varphi_{|I} = c_I \quad (I \in Z) \,. \tag{16.1}$$
Eine Zerlegung Z so, dass Konstanten c_I wie in (16.1) existieren, heißt *zulässig* für φ.

Definition 16.2 Ist $\varphi : [a, b] \to \mathbb{K}$ eine Treppenfunktion wie in D. 16.1.2, so setzen wir
$$\int\limits_a^b \varphi := \int\limits_a^b \varphi(x)\; dx := \sum_{I \in Z} c_I \lambda(I) \,,$$
wobei $\lambda(I)$ die Länge von I bezeichnet. 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 $\sum\limits_{I \in Z} c_I \lambda(I)$ von der Wahl der Zerlegung unabhängig sein. Man kann sich überlegen, dass dies der Fall ist ([Ü]). Ist etwa
$$\varphi(x) = \begin{cases} 0 \quad , \quad 0 \leq x \leq 1/2 \\ 1 \quad , \quad 1/2 < x \leq 1 \end{cases}$$

so ist $Z = \{I_1, I_2\} = \{[0, 1/2], (1/2, 1]\}$ eine zulässige Zerlegung mit $c_{I_1} = 0$, $c_{I_2} = 1$, und es gilt

$$\int_0^1 \varphi = 0 \cdot 1/2 + 1 \cdot 1/2 = 1/2 \, .$$

Eine weitere zulässige Zerlegung ist $Z = \{I_1, I_2, I_3\} = \{[0, 1/2], (1/2, 3/4], (3/4, 1]\}$ mit $c_{I_1} = 0$, $c_{I_2} = c_{I_3} = 1$. Hierfür gilt auch

$$\int_0^1 \varphi = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} = \frac{1}{2} \, .$$

Bemerkung 16.3 Es seien $\varphi, \psi : [a, b] \to \mathbb{K}$ Treppenfunktionen. Dann gilt:

1. Sind $\alpha, \beta \in \mathbb{K}$, so ist $\alpha\varphi + \beta\psi$ eine Treppenfunktion und es gilt
$$\int_a^b (\alpha\varphi + \beta\psi) = \alpha \int_a^b \varphi + \beta \int_a^b \psi.$$
2. Ist $\mathbb{K} = \mathbb{R}$ und $\varphi \le \psi$ (d. h. $\varphi(x) \le \psi(x)$ für alle $x \in [a, b]$), so ist
$$\int_a^b \varphi \le \int_a^b \psi \, .$$
3. $|\varphi|$ ist eine Treppenfunkton und es gilt
$$\left| \int_a^b \varphi \right| \le \int_a^b |\varphi| \le (b - a)||\varphi|| \, .$$
(Dabei ist $||\varphi|| := ||\varphi||_\infty = \sup\{|\varphi(x)| : x \in [a, b]\}$.)

(Denn: 1. Es seien Z_1 bzw. Z_2 zulässige Zerlegungen für φ bzw. ψ. Dann ist $Z := \{I_1 \cap I_2 : I_1 \in Z_1, I_2 \in Z_2, I_1 \cap I_2 \neq \emptyset\}$ sowohl für φ als auch für ψ zulässig. Sind $c_I \in \mathbb{K}$ bzw. $d_I \in \mathbb{K}$ wie in (16.1) für φ bzw. ψ, so gilt $(\alpha\varphi + \beta\psi)_{|I} = \alpha c_I + \beta d_I$ für $I \in Z$ und damit

$$\begin{aligned} \int_a^b \alpha\varphi + \beta\psi &= \sum_{I \in Z} (\alpha c_I + \beta d_I)\lambda(I) = \\ &= \alpha \sum_{I \in Z} c_I \lambda(I) + \beta \sum_{I \in Z} d_I \lambda(I) = \alpha \int_a^b \varphi + \beta \int_a^b \psi \, . \end{aligned}$$

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 über die Integrale der entsprechenden Treppenfunktionen definieren.

Definition 16.4 Es sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall, und es sei $f \in B([a,b],\mathbb{K})$. Dann heißt f *Regelfunktion*, falls eine Folge (φ_n) von Treppenfunktionen $\varphi_n : [a,b] \to \mathbb{K}$ existiert mit

$$\varphi_n \to f \quad \text{gleichmäßig auf } [a,b]$$

(oder, anders ausgedrückt, $||f - \varphi_n||_\infty \to 0$ für $n \to \infty$; vgl. B. 14.9). Weiter setzen wir

$$R[a,b] := R([a,b],\mathbb{K}) := \{f : [a,b] \to \mathbb{K} : f \text{ Regelfunktion}\} .$$

Wir zeigen, dass insbesondere stetige Funktionen Regelfunktionen sind.

Satz 16.5 *Es sei $f : [a,b] \to \mathbb{K}$ stetig. Mit $x_j^{(n)} := a + j(b-a)/n$ $(j = 0,\ldots,n)$, $I_j^{(n)} := (x_{j-1}^{(n)}, x_j^{(n)}]$ $(j = 1,\ldots,n)$ und $I_0^{(n)} := \{a\}$ für $n \in \mathbb{N}$ ist durch*

$$\varphi_n(x) := f(x_j^{(n)}) \qquad (x \in I_j^{(n)},\ j = 0,\ldots,n)$$

eine Folge (φ_n)von Treppenfunktionen definiert mit $\varphi_n \to f$ gleichmäßig auf $[a,b]$. Insbesondere ist also $f \in R[a,b]$.

Beweis. Es sei $\varepsilon > 0$ gegeben. Da f gleichmäßig stetig auf $[a,b]$ ist, existiert ein $\delta_\varepsilon > 0$ mit $|f(x) - f(\tilde{x}| < \varepsilon$ für alle $x, \tilde{x}$ mit $|x - \tilde{x}| < \delta_\varepsilon$. Weiter existiert ein N_ε so, dass $(b-a)/n < \delta_\varepsilon$ für alle $n \geq N_\varepsilon$.
Ist also $n \geq N_\varepsilon$ und $x \in [a,b]$, so gilt für $j = j_n(x)$ so dass $x \in I_j^{(n)}$

$$|f(x) - \varphi_n(x)| = |f(x) - f(x_j^{(n)})| < \varepsilon$$

(beachte: es ist $0 \leq x_j^{(n)} - x < \delta_\varepsilon$). Da x beliebig war, folgt

$$||f - \varphi_n||_\infty \leq \varepsilon \quad (n \geq N_\varepsilon) .$$

□

Bemerkung 16.6 Man kann zeigen ([Ü]), dass auch monotone Funktionen $f : [a,b] \to \mathbb{R}$ stets Regelfunktionen sind.

Bemerkung und Definition 16.7 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

$$\int_a^b f := \int_a^b f(x)dx := \lim_{n\to\infty} \int_a^b \varphi_n(x)dx \, .$$

Dann heisst $\int_a^b f$ das *(Regel-) Integral* (oder auch *Cauchy-Integral*) von f auf $[a,b]$.
!! Damit dies 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. 16.3.1 und 3

$$\left| \int_a^b \varphi_n - \int_a^b \varphi_m \right| = \left| \int_a^b (\varphi_n - \varphi_m) \right| \leq (b-a)||\varphi_n - \varphi_m|| \, .$$

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

$$\left| \int_a^b \varphi_n - \int_a^b \psi_n \right| \leq (b-a)||\varphi_n - \psi_n|| \leq (b-a)\big(||\varphi_n - f|| + ||f - \psi_n||\big) \to 0 \, .$$

also auch

$$\lim_{n\to\infty} \int_a^b \varphi_n = \lim_{n\to\infty} \int_a^b \psi_n \, .$$

Beispiel 16.8 Wir betrachen $f(x) = x$ auf $[0,1]$. Dann ist nach S. 16.5 durch

$$\varphi_n(x) := \begin{cases} \frac{j}{n} & , \quad x \in (\frac{j-1}{n}, \frac{j}{n}],\ j = 1, \ldots, n \\ 0 & , \quad x = 0 \end{cases}$$

eine Folge von Treppenfunktionen auf $[0,1]$ gegeben mit $\varphi_n \to f$ gleichmäßig auf $[0,1]$. Es gilt

$$\int_0^1 \varphi_n = \frac{1}{n}\sum_{j=1}^n \frac{j}{n} = \frac{n(n+1)}{2n^2} \to \frac{1}{2} \quad (n \to \infty)\ ,$$

also ist

$$\int_0^1 f(x)dx = 1/2.$$

Wir stellen einige Rechenregeln zusammen, die sich mehr oder weniger unmittelbar aus den entsprechenden Eigenschaften des Integrals für Treppenfunktionen ergeben.

Satz 16.9 *Es seien $f, g \in R[a,b]$. Dann gilt*

1. *Sind $\alpha, \beta \in \mathbb{K}$, so ist auch $\alpha f + \beta g \in R[a,b]$ und es gilt*

$$\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g$$

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*

$$\left|\int_a^b f\right| \leq \int_a^b |f| \leq (b-a)||f||$$

Beweis. Es seien (φ_n) und (ψ_n) Folgen von Treppenfunktionen mit $\varphi_n \to f$, $\psi_n \to g$ gleichmäßig auf $[a,b]$. Dann gilt

$$||\alpha f + \beta g - (\alpha\varphi_n + \beta\psi_n)|| \leq |\alpha| \cdot ||f - \varphi_n|| + |\beta| \cdot ||g - \psi_n|| \to 0 \quad (n \to \infty)\ ,$$

also $\alpha\varphi_n + \beta\psi_n \to \alpha f + \beta g$ gleichmäßig auf $[a,b]$ und deshalb mit B. 16.3.1

$$\int_b^a \alpha f + \beta g = \lim_{n\to\infty} \int_a^b (\alpha\varphi_n + \beta\psi_n) = \alpha \int_a^b f + \beta \int_a^b g\ .$$

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 B. 16.3.2

$$\int_a^b f = \lim_{n\to\infty} \int_a^b \varphi_n^- \leq \lim_{n\to\infty} \int_a^b \psi_n^+ = \int_a^b g \,.$$

3. Aus $\big| |f(x)| - |\varphi_n(x)| \big| \leq |f(x) - \varphi_n(x)|$ folgt, dass auch $|f|$ eine Regelfunktion ist und dass $|\varphi_n| \to |f|$ gleichmäßig auf $[a,b]$ gilt. Damit ergibt sich mit B. 16.3.3 und mit 2. (beachte: $|f(x)| \leq ||f||$ für $x \in [a,b]$)

$$\left| \int_a^b f(x)dx \right| = \lim_{n\to\infty} \left| \int_a^b \varphi_n(x)dx \right| \leq \lim_{n\to\infty} \int_a^b |\varphi_n(x)|dx = \int_a^b |f(x)|dx \leq (b-a)||f|| \,.$$

□

Satz 16.10 *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* $\sum\limits_{\nu=0}^{\infty} g_\nu(x)$ *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. Aus der Definition folgt unmittelbar $f \in R[a,b]$ ([Ü]). Außerdem ergibt sich aus S. 16.9.3

$$\left|\int_a^b f - \int_a^b f_n\right| \leq (b-a)||f-f_n||_\infty \to 0 \qquad (n \to \infty) .$$

2. Die 2. Behauptung ergibt sich unmittelbar aus 1. durch Anwendung auf die Teilsummenfolge. □

Bemerkung und Definition 16.11 Es sei I ein (beliebiges!) Intervall. Wir setzen

$$R(I) := R(I,\mathbb{K}) := \{f : I \to \mathbb{K} : f_{|J} \in R(J,\mathbb{K}) \text{ für alle kompakten Intervalle } J \subset I\} .$$

Ist $f \in R(I)$, so definieren wir noch für $x, y \in I$ mit $x < y$

$$\int_x^x f := 0 \qquad \text{und} \qquad \int_y^x f := -\int_x^y f$$

Damit gilt für $x, y, z \in I$ ([Ü])

$$\int_x^y f + \int_y^z f = \int_x^z f$$

und

$$\left|\int_x^y f\right| \leq ||f||_{\infty, I[x,y]}|y-x|$$

wobei $||\cdot||_{\infty,M}$ die Supremumsnorm auf M und $I[x,y] := \{x + t(y-x) : t \in [0,1]\}$ die Stecke zwischen x und y bezeichnet.

Wir kommen zu einem der zentralen Sätze der Analysis, der die Beziehung zwischen der Differenzial- und der Integralrechnung herstellt.

Satz 16.12 *(Hauptsatz der Differenzial- und Integralrechnung; HDI)*
Es seien $I \subset \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f \in R(I)$. Die Funktion $F : I \to \mathbb{K}$ sei definiert durch

$$F(x) := \int_{x_0}^x f(t)dt \qquad (x \in I)$$

(sog. Integralfunktion von f bzgl. x_0). Dann gilt:

1. *F ist stetig auf I.*

2. *Ist f stetig an der Stelle $x \in I$, so ist F differenzierbar an x mit*

$$F'(x) = f(x) \ .$$

Beweis.
1. Es sei $x \in I$ beliebig. Dann existiert ein $r > 0$ so, dass

$$J := [x - r, x + r] \cap I$$

kompakt ist. Also ist f beschränkt auf J (da $f \in R(J)$). Damit gilt für h mit $x+h \in J$

$$|F(x+h) - F(x)| = \Big| \int\limits_{x_0}^{x+h} f - \int\limits_{x_0}^{x} f \Big| = \Big| \int\limits_{x}^{x+h} f \Big| \leq \|f\|_{\infty,J} \cdot |h| \to 0 \quad (h \to 0).$$

2. Es sei f stetig an x. Dann gilt für $h \in I - \{x\}$

$$\big\|f - f(x)\big\|_{\infty,I[x,x+h]} \to 0 \qquad (h \to 0)$$

und folglich

$$\Big|\frac{F(x+h) - F(x)}{h} - f(x)\Big| = \Big|\frac{1}{h} \int\limits_{x}^{x+h} \big(f(t) - f(x)\big)\, dt\Big| \leq \frac{|h|}{|h|} \big\|f - f(x)\big\|_{\infty,I[x,x+h]} \to 0$$

für $h \to 0$. $\square$

Wir wollen nun verschiedene Techniken zur Berechnung von Integralen herleiten. Von zentraler Bedeutung ist dabei der Begriff der Stammfunktion.

Bemerkung und Definition 16.13 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 (auf I), falls F differenzierbar ist mit $F' = f$ auf I.
Nach S. 16.12 sind die dort definierten Integralfunktionen für **stetige** Funktionen f auch Stammfunktionen. Für $f \in R(I)$ sind im Allgemeinen Integralfunktionen keine Stammfunktionen! ([Ü])
Wichtig für das Weitere: Sind F und G Stammfunktionen zu f, so existiert ein $c \in \mathbb{K}$ mit $F(x) = G(x) + c$ für alle $x \in [a, b]$, d. h. zwei Stammfunktionen unterscheiden sich lediglich durch eine additive Konstante ([Ü]). Durch

$$F \sim G :\Leftrightarrow F - G \text{ konstant}$$

ist eine Äquivalenzrelation auf der Menge der auf I differenzierbaren Funktionen gegeben. Wir schreiben für die zu einer Stammfunktion F von f gehörende Äquivalenzklasse $[F]$ auch

$$x \mapsto \int f(x)\, dx$$

(sog. *unbestimmtes Integral* von f). Addition und Skalarmultiplikation sind repräsentantenweise (wohl-)definiert.

Der nun folgende zweite Teil des Hauptsatzes der Differenzial- und Integralrechnung zeigt, dass Integrale – jedenfalls für stetige Funktionen – durch Bestimmung einer Stammfunktion berechnet werden können:

Satz 16.14 *(HDI, Teil 2)*
Es seien I ein Intervall und $f : I \to \mathbb{K}$ stetig. Ist Φ eine beliebige Stammfunktion zu f auf I, so gilt für alle $u, v \in I$

$$\int_u^v f = \Phi(v) - \Phi(u) =: \Phi|_u^v \; . \tag{16.2}$$

Beweis. Nach S. 16.12 ist F (wie dort definiert) eine Stammfunktion zu f auf I. Ist Φ eine beliebige Stammfunktion zu f auf I, so gilt $\Phi(x) = F(x) + c$ für eine Konstante $c \in \mathbb{K}$. Damit ist

$$\int_u^v f = F(v) - F(u) = \Phi(v) - \Phi(u) \; .$$

□

Beispiel 16.15 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 $u, v > 0$:

$$\int_u^v f = \int_u^v \frac{1}{x}\, dx = \ln x|_u^v = \ln(v) - \ln(u) \quad \left(= \ln\left(\frac{v}{u}\right)\right) \; .$$

Ist $f(x) = x^\alpha$ $(x > 0)$, wobei $\alpha \in \mathbb{R}$, $\alpha \neq -1$ fest ist, so 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 $u, v > 0$

$$\int\limits_u^v x^\alpha \, dx = \frac{1}{\alpha+1}x^{\alpha+1}|_u^v = \frac{1}{\alpha+1}\left(v^{\alpha+1} - u^{\alpha+1}\right) \ .$$

Im Falle $\alpha \geq 0$ gilt dies auch für $u = 0$.

Bemerkung 16.16 Durch Übertragung der Produkt- und der Kettenregel ergeben sich wichtige Techniken zur möglichen Berechnung von Stammfunktionen.
1. Sind $f, g : I \to \mathbb{K}$ und sind F bzw. G Stammfunktionen zu f bzw. g auf I, so folgt aus der Produktregel, dass FG eine Stammfunktion zu $fG + Fg$ auf I ist. Als Konsequenz erhalten wir

$$\int f(x)G(x)dx = F(x)G(x) - \int F(x)g(x)\, dx \quad \text{auf } I$$

(unbestimmte partielle Integration).
2. Mit Hilfe der Kettenregel sieht man: Ist $t : I \to J$ differenzierbar (wobei I, J Intervalle) und ist F eine Stammfunktion zu f auf J, so ist nach der Kettenregel $F \circ t$ eine Stammfunktion zu $(f \circ t) \cdot t'$ auf I, also

$$\int f(t(x))t'(x)dx = F(t(x)) = \int f(t)dt_{|t=t(x)} \quad \text{auf } I$$

(Substitutionsregel; unbestimmt).

Beispiel 16.17 1. Für $\alpha \neq -1$ und $x > 0$ gilt

$$\begin{aligned} \int x^\alpha \ln x \, dx &= \frac{1}{\alpha+1}x^{\alpha+1}\ln x - \frac{1}{\alpha+1}\int x^\alpha \, dx = \\ &= \frac{x^{\alpha+1}}{\alpha+1}\left(\ln x - \frac{1}{\alpha+1}\right) \end{aligned}$$

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 (übrigens auch auf $[-1, 1]$)

$$\int \sqrt{1-x^2}\,dx = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x)$$

Damit folgt insbesondere mit dem HDI, Teil 2,

$$\int\limits_{-1}^{1} \sqrt{1-x^2}\,dx = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x)|_{-1}^{1} = \frac{\pi}{2}$$

(Fläche der oberen Hälfte des Einheitskreises).

3. Es gilt auf $(-1, 1)$

$$\int \frac{2x}{\sqrt{1-x^2}}\,dx \overset{t=x^2}{=} \int \frac{dt}{\sqrt{1-t}}_{|t=x^2} = -2\sqrt{1-t}_{|t=x^2} = -2\sqrt{1-x^2}\,.$$

4. Es gilt auf $I = (0, \pi)$

$$\begin{aligned}\int \frac{dx}{\sin x} &\overset{S.8.10}{=} \int \frac{dx}{2\sin(x/2)\cos(x/2)} = \\ &= \int \frac{\cos(x/2)}{\sin(x/2)} \frac{1}{2\cos^2(x/2)}\,dx \\ &= \int \frac{1}{\tan(x/2)} \frac{1}{2\cos^2(x/2)}\,dx \\ &\overset{t(x)=\tan(x/2)}{=} \int \frac{dt}{t}_{|t=\tan(x/2)} = \ln(\tan(x/2))\end{aligned}$$

Bemerkung 16.18 Ist $\sum\limits_{\nu=0}^{\infty} a_\nu (x-x_0)^\nu$ eine Potenzreihe mit Konvergenzradius $R > 0$ und ist

$$f(x) = \sum_{\nu=0}^{\infty} a_\nu (x-x_0)^\nu \qquad \big(x \in I := (x_0 - R, x_0 + R)\big),$$

so ist nach S. 15.5 durch

$$F(x) := \sum_{\nu=0}^{\infty} \frac{a_\nu}{\nu+1} (x-x_0)^{\nu+1} \ \Big(= \sum_{\nu=1}^{\infty} \frac{a_{\nu-1}}{\nu} (x-x_0)^\nu \Big) \quad (x \in I)$$

eine Stammfunktion von f auf I definiert. Dabei gilt nach dem HDI, Teil 2 (da f stetig auf I ist)

$$F(v) - F(u) = \int\limits_u^v f \qquad (u, v \in I) \ .$$

Beispiel 16.19 Ist

$$f(x) := e^{-x^2} = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu x^{2\nu}}{\nu!} \qquad (x \in \mathbb{R})$$

so ist durch

$$F(x) = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu}{\nu!(2\nu+1)} \, x^{2\nu+1} \qquad (x \in \mathbb{R})$$

eine Stammfunktion von f auf $\mathbb{R}$ gegeben. Die Funktion

$$\mathrm{erf} := \frac{2}{\sqrt{\pi}} F$$

heißt *Fehlerfunktion* (error function).

17 Uneigentliche Integrale

Wir haben bisher nur Integrale auf kompakten Intervallen betrachtet. Wir wollen jetzt auch Integrale auf nichtkompakten Intervallen erklären.

Definition 17.1 Es sei I ein Intervall, wobei $I^0 = (a, b)$ mit $-\infty \leq a < b \leq \infty$. Ferner seien $f \in R(I)$ und $x_0 \in I$. Ist $F : I \to \mathbb{K}$ die Integralfunktion von f bzgl. x_0, d. h. $F(x) = \int\limits_{x_0}^{x} f$ für $x \in I$, so heißt f (*uneigentlich*) *integrierbar* auf I, falls $F(a^+)$ und $F(b^-)$ existieren. In diesem Fall sagt man auch, dass das uneigentliche Integral $\int\limits_{a^+}^{b^-} f$ existiert und die Zahl

$$\int\limits_{a^+}^{b^-} f := \int\limits_{a^+}^{b^-} f(x)dx := F(b^-) - F(a^+)$$

heißt dann (*uneigentliches*) *Integral* von f auf I.

Bemerkung 17.2 1. Sind $x_1 \in I$ und $G(x) = \int\limits_{x_1}^{x} f$, so unterscheiden sich F und G nur durch eine additive Konstante. Daher ist die Definition unabängig von der Wahl von x_0.

2. Ist $a \in I$, so gilt $F(a) = F(a^+)$ nach S. 16.12.1. Entsprechend ist $F(b) = F(b^-)$ im Falle $b \in I$. Also ist im Falle $I = [a, b]$

$$\int\limits_{a^+}^{b^-} f = F(b^-) - F(a^+) = F(b) - F(a) = \int\limits_{a}^{b} f,$$

d.h. ,,eigentliches" und uneigentliches Integral stimmen überein. Wir schreiben daher auch manchmal kurz $\int\limits_{a}^{b} f$ statt $\int\limits_{a^+}^{b^-} f$.

3. (erweiterter HDI, Teil 2) Es sei $f : I \to \mathbb{K}$ stetig. Nach dem HDI, Teil 1, ist F aus D. 17.1 eine Stammfunktion von f auf I. Ist Φ eine (beliebige) Stammfunktion von f auf I, so ist f genau dann integrierbar, wenn $\Phi(b^-)$ und $\Phi(a^+)$ existieren. Außerdem ist dann

$$\int\limits_{a^+}^{b^-} f = \Phi(b^-) - \Phi(a^+) =: \Phi|_a^b.$$

(Denn: Die Behauptung ergibt sich aus dem HDI, Teil 2 durch Grenzübergang $u \to a^+$ und $v \to b^-$.)

Beispiel 17.3 1. Für $\alpha \in \mathbb{R}$ sei $f_\alpha(x) := 1/x^\alpha$ auf $I = (0, \infty)$. Dann ist $\Phi_\alpha : (0, \infty) \to \mathbb{R}$ definiert durch

$$\Phi_\alpha(x) = \begin{cases} \dfrac{x^{1-\alpha}}{1-\alpha} & , \quad \alpha \neq 1 \\ \ln(x) & , \quad \alpha = 1 \end{cases}$$

eine Stammfunktion zu f_α auf I. Außerdem erhalten wir für $x \to \infty$

$$\Phi_\alpha(x) \to \begin{cases} 0 & , \quad \alpha > 1 \\ \infty & , \quad \alpha \leq 1 \end{cases}$$

und für $x \to 0^+$

$$\Phi_\alpha(x) \to \begin{cases} 0 & , \quad \alpha < 1 \\ \infty & , \quad \alpha \geq 1 \end{cases} .$$

Also ist f_α nach B. 17.2.3 genau dann (uneigentlich) integrierbar auf $[1, \infty)$, wenn $\alpha > 1$ ist und in diesem Falle ist

$$\int_1^\infty \frac{dx}{x^\alpha} = \Phi_\alpha(x)|_1^\infty = 0 - \frac{1}{1-\alpha} = \frac{1}{\alpha - 1} .$$

Entsprechend ist f_α genau dann integrierbar auf $(0, 1]$, wenn $\alpha < 1$ ist, und dann gilt

$$\int_{0^+}^1 \frac{dx}{x^\alpha} = \Phi_\alpha(x)|_0^1 = \frac{1}{1-\alpha} .$$

Hieraus folgt auch, dass f_α für kein $\alpha \in \mathbb{R}$ auf $(0, \infty)$ integrierbar ist.

2. Wir betrachten $f(x) = \dfrac{1}{\sqrt{1-x^2}}$ auf $I = (-1, 1)$. Es gilt $\arcsin' = f$ auf $(-1, 1)$ und arcsin ist stetig auf $[-1, 1]$. Nach B. 17.2.3 existiert $\displaystyle\int_{-1}^1 \frac{dx}{\sqrt{1-x^2}}$ und es ist

$$\int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} = \arcsin(x)|_{-1}^1 = \pi .$$

Satz 17.4 *Es sei I ein Intervall, $I^0 = (a, b)$, und es seien $f, g \in R(I)$. Dann gilt:*

1. *Sind f und g integrierbar auf I, so ist für alle $\alpha, \beta \in \mathbb{K}$ auch $\alpha f + \beta g$ integrierbar mit*

$$\int_{a^+}^{b^-} \alpha f + \beta g = \alpha \int_{a^+}^{b^-} f + \beta \int_{a^+}^{b^-} g .$$

2. *Sind f, g (reellwertig und) integrierbar mit $f \leq g$ auf I, so ist*

$$\int_{a^+}^{b^-} f \leq \int_{a^+}^{b^-} g\,.$$

3. *(Majorantenkriterium für Integrale) Ist g integrierbar auf I und gilt $|f| \leq g$, so ist auch f integrierbar auf I mit*

$$\Big|\int_{a^+}^{b^-} f\Big| \leq \int_{a^+}^{b^-} g\,.$$

Beweis. Es seien $x_0 \in I$ fest und $F(x) := \int\limits_{x_0}^{x} f$ und $G(x) := \int\limits_{x_0}^{x} g$ für $x \in I$.

1. Folgt sofort aus S. 16.9.1 und D. 17.1.

2. Für $u, v \in I$ mit $u \leq v$ gilt mit S. 16.9.2

$$F(v) - F(u) = \int_u^v f \leq \int_u^v g = G(v) - G(u)$$

und damit $F(b^-) - F(a^+) \leq G(b^-) - G(a^+)$ durch Grenzübergang $v \to b^-$ und $u \to a^+$.

3. Wir zeigen: $F(b^-)$ existiert. Dazu reicht es, zu zeigen ([Ü]): Für alle Folgen (x_n) in I^0 mit $x_n \to b$ ist $\big(F(x_n)\big)$ eine Cauchy-Folge.

Es sei also (x_n) eine solche Folge. Mit S. 16.9.2. und 3. folgt

$$\big|F(x_n) - F(x_m)\big| = \Big|\int_{x_m}^{x_n} f\Big| \leq \int_{x_m}^{x_n} g = G(x_n) - G(x_m)$$

im Falle $x_n \geq x_m$ und damit stets

$$|F(x_n) - F(x_m)| \leq |G(x_n) - G(x_m)|\,.$$

Da $(G(x_n))$ eine Cauchy-Folge ist, ist folglich auch $(F(x_n))$ eine Cauchy-Folge.
Genauso sieht man, dass $F(a^+)$ existiert.
Schließlich folgt aus

$$|F(v) - F(u)| = \Big|\int_u^v f\Big| \leq \int_u^v g = G(v) - G(u)$$

für $u \leq v$ auch $|F(b^-) - F(a^+)| \leq G(b^-) - G(a^+)$, wieder durch Grenzübergang $v \to b^-$ und $u \to a^+$. $\square$

Bemerkung 17.5 Insbesondere ergibt sich aus S. 17.4.3 (mit $g := |f|$): Ist $f \in R(I)$ (und damit auch $|f| \in R(I)$), so folgt aus der Integrierbarkeit von $|f|$ auch die von f. Wir sprechen dann auch von *absoluter Integrierbarkeit* von f. Wie bei Reihen gilt also: Ist f absolut integrierbar, so ist f integrierbar. Außerdem gilt dann:

$$\Big|\int_{a^+}^{b^-} f\Big| \leq \int_{a^+}^{b^-} |f| \,.$$

Beispiel 17.6 1. Wir betrachten für $\alpha > 1$ das uneigentliche Integral $\displaystyle\int_1^\infty \frac{\cos x}{x^\alpha}\, dx$.

Es gilt

$$\frac{|\cos x|}{x^\alpha} \leq \frac{1}{x^\alpha} \qquad (x \geq 1) \,.$$

Da $\displaystyle\int_1^\infty \frac{dx}{x^\alpha}$ existiert, folgt die Existenz von $\displaystyle\int_1^\infty \frac{|\cos x|}{x^\alpha}\, dx$ aus S. 17.4.3. Nach B. 17.5 existiert auch $\displaystyle\int_1^\infty \frac{\cos x}{x^\alpha}\, dx$. Entsprechendes gilt für das uneigentliche Integral $\displaystyle\int_1^\infty \frac{\sin x}{x^\alpha}\, dx$.

Satz 17.7 *(Uneigentliche partielle Integration)*
Es seien $I \subset \mathbb{R}$ mit $I^0 = (a,b)$ ein Intervall und $f, g : I \to \mathbb{K}$ stetig. Sind F, G zugehörige Stammfunktionen auf I und existieren $(FG)(b^-)$ sowie $(FG)(a^+)$, so gilt: fG ist genau dann integrierbar auf I, wenn Fg integrierbar auf I ist, und in diesem Fall ist

$$\int_{a^+}^{b^-} fG = FG\Big|_a^b - \int_{a^+}^{b^-} Fg.$$

Beweis. Da F, G stetig auf I sind, ist $fG + Fg$ stetig. Außerdem ist FG eine Stammfunktion zu $fG + Fg$. Nach B. 17.2.3 ist $fG + Fg$ integrierbar mit

$$\int_{a^+}^{b^-} (fG + Fg) = FG\Big|_a^b.$$

Ist nun etwa fG integrierbar, so ist auch $Fg = (Fg + fG) - fG$ integrierbar nach S. 17.4.1, und es gilt

$$\int_{a^+}^{b^-} fG = FG\Big|_a^b - \int_{a^+}^{b^-} Fg.$$

Entsprechendes gilt, falls Fg integrierbar ist. $\square$

Beispiel 17.8 Es sei $f : [1, \infty) \to \mathbb{K}$ stetig. Ist F eine Stammfunktion von f und ist F beschränkt, so existiert $\int\limits_1^\infty f(x)/x\, dx$.

(Denn: Da F beschränkt (und stetig) ist, existiert das Integral $\int\limits_1^\infty F(x)/x^2\, dx$ nach dem Majorantenkriterium. Mit $G(x) = 1/x$ und $g(x) = -1/x^2$ ergibt sich die Behauptung aus S. 17.7 (beachte: es gilt $F(x)G(x) = F(x)/x \to 0$ für $x \to \infty$).)

Wir betrachten $f(x) = \sin x$. Hier ist $F(x) = -\cos x$ beschränkt auf $[1, \infty)$. Also existiert das uneigentliche Integral $\displaystyle\int\limits_1^\infty \frac{\sin x}{x}\, dx$. Man kann zeigen ([Ü]): Das uneigentliche Integral $\displaystyle\int\limits_1^\infty \frac{|\sin x|}{x}\, dx$ 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 17.9 *(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 mit $a_\nu := f(\nu)$

$$0 \leq a_\nu - \int\limits_\nu^{\nu+1} f \leq a_\nu - a_{\nu+1}$$

und damit ist die Folge (s_n) mit

$$s_n := \sum_{\nu=1}^{n} a_\nu - \int_1^{n+1} f = \sum_{\nu=1}^{n} \left(a_\nu - \int_\nu^{\nu+1} f \right)$$

monoton wachsend mit $0 \le s_n \le f(1) - f(n+1) \le f(1)$. Also ergibt sich die Behauptung mit dem Hauptsatz über monotone Folgen. $\square$

Beispiel 17.10 Wir betrachten $f(x) = \dfrac{1}{x^\alpha}$ für $\alpha > 0$. Dann ist f monoton fallend auf $[1, \infty)$ und $f(x) \ge 0$. Also existiert

$$g = \lim_{n\to\infty} \left(\sum_{\nu=1}^{n} \frac{1}{\nu^\alpha} - \int_1^{n+1} \frac{dx}{x^\alpha} \right) .$$

Ist $\alpha > 1$, so ist $\lim\limits_{n\to\infty} \displaystyle\int_1^{n+1} \frac{dx}{x^\alpha} = \int_1^{\infty} \frac{dx}{x^\alpha} = \frac{1}{\alpha - 1}$ und $\zeta(\alpha) = \displaystyle\sum_{\nu=1}^{\infty} \frac{1}{\nu^\alpha}$. Nach S. 17.9 ist

$$0 \le \zeta(\alpha) - \frac{1}{\alpha - 1} \le 1 \qquad (\alpha > 1) .$$

Ist $\alpha = 1$, so ergibt sich die Konvergenz von

$$a_n = \sum_{\nu=1}^{n} \frac{1}{\nu} - \int_1^{n+1} \frac{dx}{x} = \sum_{\nu=1}^{n} \frac{1}{\nu} - \ln(n+1) .$$

Ist schließlich $0 < \alpha < 1$, so ergibt sich die Konvergenz von

$$a_n = \sum_{\nu=1}^{n} \frac{1}{\nu^\alpha} - \int_1^{n+1} \frac{dx}{x^\alpha} = \sum_{\nu=1}^{n} \frac{1}{\nu^\alpha} - \frac{(n+1)^{1-\alpha} - 1}{1 - \alpha} .$$

Beispiel 17.11 Das uneigentliche Integral $\int_0^\infty e^{-t} t^{z-1} dt$ existiert für jedes $z \in \mathbb{C}$ mit $\mathrm{Re}(z) > 0$.
(Denn: Es gilt

$$\lim_{t\to\infty} \frac{t^{z+1}}{e^t} = 0 .$$

Also existiert eine Konstante $M = M_x > 0$ so, dass für alle $t \in [1, \infty)$ gilt

$$|e^{-t} t^{z-1}| \le \frac{M}{t^2}$$

(warum?). Aus der Existenz von $\int_1^\infty 1/t^2\,dt$ ergibt sich mit dem Majorantenkriterium auch die Existenz von $\int_1^\infty e^{-t}t^{z-1}dt$.
Weiter gilt

$$|e^{-t}t^{z-1}| \le t^{\mathrm{Re}(z)-1} \qquad (t \in (0,1]) \ .$$

Aus der Existenz von $\int_0^1 t^{\mathrm{Re}(z)-1}dt$ (siehe B. 17.3.1) folgt wieder mit dem Majorantenkriterium die Existenz von $\int_0^1 e^{-t}t^{z-1}dt$. Insgesamt ergibt sich damit die Existenz von $\int_0^\infty e^{-t}t^{z-1}dt$.)
Die Funktion $\Gamma : \{z : \mathrm{Re}(z) > 0\} \to \mathbb{C}$ mit

$$\Gamma(z) := \int_{0+}^{\infty} e^{-t}t^{z-1}dt \qquad (\mathrm{Re}(z) > 0)$$

heißt (*Eulersche*) *Gammafunktion*. Durch uneigentliche partielle Integration erhält man unmittelbar die folgende Funktionalgleichung für Γ:

$$\Gamma(z+1) = z \cdot \Gamma(z) \qquad (\mathrm{Re}(z) > 0) \ . \tag{17.1}$$

Speziell gilt

$$\Gamma(1) = \int_0^\infty e^{-t}dt = -e^{-t}|_0^\infty = 1 \ ,$$

woraus sich wiederum mit (17.1) 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.

Wir wollen nun 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 17.12 *(Eulersche Summenformel)*
Es sei $n \in \mathbb{N}$, $n \ge 2$, und es sei $f : [1,n] \to \mathbb{K}$ stetig differenzierbar (d. h. f ist differenzerbar auf $[1,n]$ mit stetiger Ableitung). Dann gilt

$$\sum_{\nu=1}^{n} f(\nu) = \int_1^n f + \frac{1}{2}\Big(f(1) + f(n)\Big) + \int_1^n b(x)f'(x)\,dx$$

mit

$$b(x) := x - [x] - 1/2 \quad \textit{und} \quad [x] := \max\{\nu \in \mathbb{Z} : \nu \le x\}$$

Beweis. Es gilt mit partieller Integration

$$\begin{aligned}\int_1^n b(x)f'(x)\,dx &= \sum_{\nu=1}^{n-1}\int_\nu^{\nu+1}(x-\nu-1/2)f'(x)\,dx = \\ &= \sum_{\nu=1}^{n-1}\left[(x-\nu-1/2)f(x)\Big|_\nu^{\nu+1} - \int_\nu^{\nu+1} f(x)\,dx\right] = \\ &= \frac{1}{2}\sum_{\nu=1}^{n-1}\Big(f(\nu+1)+f(\nu)\Big) - \int_1^n f(x)\,dx \\ &= \sum_{\nu=1}^{n} f(\nu) - \frac{1}{2}\Big(f(1)+f(n)\Big) - \int_1^n f(x)\,dx\,.\end{aligned}$$

□

Für zwei Folgen (a_n) und (b_n) in $\mathbb{K}\setminus\{0\}$ schreiben wir

$$a_n \sim b_n \qquad (n\to\infty)$$

falls $a_n/b_n \to 1\,(n\to\infty)$ gilt.

Satz 17.13 *(Stirlingsche Formel)*
Es ist

$$n! \sim n^n e^{-n}\sqrt{2\pi n} \qquad (n\to\infty)\,. \tag{17.2}$$

Beweis. 1. Für $f(x)=\ln(x)$ $(x\geq 1)$ gilt

$$\int_1^n \ln(x)\,dx = x\ln(x) - x\Big|_1^n = n\ln n - n + 1$$

und damit mit S. 17.12

$$\sum_{\nu=1}^{n}\ln\nu = n\ln n - n + 1 + \frac{1}{2}\ln n + \int_1^n \frac{b(x)}{x}\,dx\,.$$

Weiter gilt für die Integralfunktion B von b bzgl. $x_0=1$

$$B(x) = \int_1^x b(t)dt = \frac{1}{2}\left(x-[x]-\frac{1}{2}\right)^2 - \frac{1}{8} \qquad (x\geq 1)$$

und damit mit partieller Integration (man beachte: B ist Stammfunktion zu b auf $(\nu, \nu+1)$ und $B(\mathbb{N}) = \{0\}$)

$$\int_1^n \frac{b(x)}{x}\,dx = \sum_{\nu=1}^{n-1} \int_\nu^{\nu+1} \frac{b(x)}{x}\,dx = \sum_{\nu=1}^{n-1} \left[\frac{B(x)}{x}\Big|_\nu^{\nu+1} + \int_\nu^{\nu+1} \frac{B(x)}{x^2}\,dx \right] = \int_1^n \frac{B(x)}{x^2}\,dx \,.$$

Da B beschränkt ist (konkret durch 1/8), ist nach dem Majorantenkriterium $x \mapsto B(x)/x^2$ integrierbar auf $[1, \infty)$. Damit ist die Folge $\left(\int_1^n b(x)/x\,dx \right)_n$ konvergent.

2. 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.

$$\ln a_n = \sum_{\nu=1}^n \ln \nu - n \ln n + n - \frac{1}{2} \ln n = 1 + \int_1^n \frac{b(x)}{x}\,dx \,,$$

also ist $(\ln a_n)$ konvergent nach 1. und damit konvergiert auch (a_n). Wir setzen $a := \lim\limits_{n\to\infty} a_n$. Dann ist $a > 0$ und es gilt unter Verwendung Wallis-Produktes für $\pi/2$ (siehe [Ü])

$$\begin{aligned} \sqrt{\pi/2} &= \lim_{n\to\infty} \frac{1}{\sqrt{2n+1}} \cdot \frac{2^n n!}{1 \cdot 3 \cdot 5 \cdots (2n-1)} \\ &= \lim_{n\to\infty} \frac{1}{\sqrt{2n+1}} \cdot \frac{(2^n n!)^2}{(2n)!} \\ &= \lim_{n\to\infty} \frac{1}{\sqrt{2n+1}} \frac{2^{2n} (a_n n^n e^{-n} \sqrt{n})^2}{a_{2n} (2n)^{2n} e^{-2n} \sqrt{2n}} \\ &= \lim_{n\to\infty} \frac{1}{\sqrt{2n+1}} \frac{a_n^2}{a_{2n}} \frac{\sqrt{n}}{\sqrt{2}} = \frac{a}{2} \end{aligned}$$

d. h. $a = \sqrt{2\pi}$.

□

Unter Ausnutzung des Wallis-Produktes erhält man auch

Satz 17.14 *Es ist*

$$\int_0^\infty e^{-u^2}\,du = \frac{\sqrt{\pi}}{2} \,.$$

Beweis. Nach S. 9.2.2 ist $e^x \leq \frac{1}{1-x}$ $(x < 1)$, also

$$e^{-t^2} \leq \frac{1}{1+t^2} \qquad (t \in \mathbb{R})$$

und damit auch

$$e^{-nt^2} \leq \left(\frac{1}{1+t^2}\right)^n \qquad (t \in \mathbb{R}, n \in \mathbb{N}) \; .$$

Entsprechend folgt aus $1 + x \leq e^x$

$$(1-t^2)^n \leq e^{-nt^2} \qquad (0 \leq t \leq 1) \; .$$

Also erhalten wir

$$\int_0^{\pi/2} \sin^{2n+1} x \; dx \overset{t(x)=\cos x}{=} \int_0^1 (1-t^2)^n \; dt \leq \int_0^1 e^{-nt^2} \; dt$$

$$\leq \int_0^\infty e^{-nt^2} \; dt \Big(= \frac{1}{\sqrt{n}} \int_0^\infty e^{-u^2} \; du \Big) \leq \int_0^\infty \frac{dt}{(1+t^2)^n} \overset{t(x)=\cot x}{=} \int_0^{\pi/2} \sin^{2n-2} x \; dx$$

d. h.

$$\sqrt{n} \int_0^{\pi/2} \sin^{2n+1} x \; dx \leq \int_0^\infty e^{-u^2} \; du \leq \sqrt{n} \int_0^{\pi/2} \sin^{2n-2} x \; dx \; .$$

Mit $w_n := \dfrac{2^n n!}{1 \cdot 3 \cdot 5 \cdots (2n-1)}$ gilt weiter ([Ü])

$$\int_0^{\pi/2} \sin^{2n+1} x \; dx = \frac{w_n}{2n+1} \qquad \text{und} \qquad \int_0^{\pi/2} \sin^{2n-2} x \; dx = \frac{\pi}{2} \cdot \frac{1}{w_{n-1}}$$

und damit

$$\frac{\sqrt{n}}{\sqrt{2n+1}} \frac{w_n}{\sqrt{2n+1}} \leq \int_0^\infty e^{-u^2} \; du \leq \frac{\pi}{2} \frac{\sqrt{n}}{\sqrt{2n-1}} \frac{\sqrt{2n-1}}{w_{n-1}} \; .$$

Da die rechte und die linke Seite gegen $\sqrt{\pi}/2$ konvergieren (Wallis-Produkt; siehe Beweis zu S. 17.13), ergibt sich die Behauptung. □

18 Differenzialrechnung von Funktionen mehrerer Variablen

In Abschnitt 12 haben wir Ableitungen von Funktionen einer (reellen oder komplexen) Variable untersucht. Wir studieren jetzt für $d, m \in \mathbb{N}$ Funktionen $f : M \to \mathbb{K}^m$, wobei $M \subset \mathbb{R}^d$ ist. Die Räume $\mathbb{R}^d$ bzw. $\mathbb{K}^m$ seien im Folgenden stets mit der euklidschen Metrik versehen. Damit stehen uns die Begriffe und Ergebnisse aus den Abschnitten 10 und 11 auch hier zur Verfügung.
Bei vielen Untersuchungen kann man sich auf den Fall $m = 1$, d. h. auf den Fall reell- oder komplexwertiger Funktionen, beschränken. Betrachten wir also zunächst Funktionen $f : M \to \mathbb{K}$.
Bevor wir zum eigentlichen Begriff der Differenzierbarkeit von Funktionen mehrerer Veränderlicher kommen, werden wir noch kurz auf weitere Grundbegriffe und Ergebnisse aus der Topologie eingehen.

Satz 18.1 *Es seien (X, d) ein metrischer Raum und $M_\alpha \subset X$ für $\alpha \in I$.*

1. Ist $M_\alpha \left\{ \begin{array}{l} \textit{offen} \\ \textit{abgeschlossen} \end{array} \right\}$ *für alle* $\alpha \in I$, *so ist* $\begin{cases} \bigcup_{\alpha \in I} M_\alpha \textit{ offen} \\ \bigcap_{\alpha \in I} M_\alpha \textit{ abgeschlossen} \end{cases}$

2. Ist I endlich und $M_\alpha \left\{ \begin{array}{l} \textit{offen} \\ \textit{abgeschlossen} \end{array} \right\}$ *für alle* $\alpha \in I$, *so ist* $\begin{cases} \bigcap_{\alpha \in I} M_\alpha \textit{ offen} \\ \bigcup_{\alpha \in I} M_\alpha \textit{ abgeschlossen} \end{cases}$.

Beweis. 1. Es seien M_α offen ($\alpha \in I$). Ist $x \in \bigcup_{\alpha \in I} M_\alpha$, so existiert ein $\alpha \in I$ mit $x \in M_\alpha$. Da M_α offen ist, existiert ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subset M_\alpha$. Damit ist $U_\varepsilon(x) \subset \bigcup_{\alpha \in I} M_\alpha$.
Ist M_α abgeschlossen ($\alpha \in I$), so ist M_α^c offen ($\alpha \in I$), also auch

$$\Big(\bigcap_{\alpha \in I} M_\alpha\Big)^c = \bigcup_{\alpha \in I} M_\alpha^c \quad \text{offen.}$$

2. Es seien $M_\alpha \,(\alpha \in I)$ offen und $x \in \bigcap_{\alpha \in I} M_\alpha$. Da M_α offen ist, existiert ein $\varepsilon_\alpha > 0$ mit $U_{\varepsilon_\alpha}(x) \subset M_\alpha$. Für $\varepsilon := \min\{\varepsilon_\alpha : \alpha \in I\}$ gilt dann $U_\varepsilon(x) \subset M_\alpha$ für alle $\alpha \in I$, also $U_\varepsilon(x) \subset \bigcap_{\alpha \in I} M_\alpha$.
Die Behauptung für abgeschlossene Mengen ergibt sich wie oben durch Komplementbildung. □

Bemerkung 18.2 Die Aussagen von 18.1.2 sind im Allgemeinen falsch für unendliche Indexmengen I. Ist etwa $(X, d) = (R, d_{|\cdot|})$, so ist für die offene Mengen

$$M_n := \big(-1/n, 1/n\big) \qquad (n \in \mathbb{N})$$

der abzählbare Schnitt

$$\bigcap_{n\in\mathbb{N}} M_n = \{0\}$$

nicht mehr offen. Durch Komplementbildung sieht man, dass auch unendliche Vereinigungen abgeschlossener Mengen im Allgemeinen nicht mehr abgeschlossen sind.

Definition 18.3 Es seien (X, d) ein metrischer Raum und $M \subset X$.

1. Die Menge

   $$M^0 := \bigcup_{\substack{U\subset M\\ U \text{ offen}}} U$$

   heißt *innerer Kern* (oder *Inneres*) von M.

   Jedes $x \in M^0$ heißt *innerer Punkt* von M.

2. Die Menge

   $$\overline{M} := \bigcap_{\substack{A\supset M\\ A \text{ abgeschlossen}}} A$$

   heißt *abgeschlossene Hülle* (oder *Abschluss*) von M.

3. Die Menge $\partial M := \overline{M} \setminus M^0$ heißt *Rand* von M.

4. M heißt *dicht* (in X), falls $\overline{M} = X$ gilt.

Bemerkung 18.4 Es sei (X, d) ein metrischer Raum. Aus S. 18.1 ergibt sich sofort, dass für jede Menge $M \subset X$ das Innere M^0 offen und der Abschluss $\overline{M}$ abgeschlossen sind. Außerdem folgt aus $\partial M = \overline{M} \cap (M^0)^c$ auch, dass ∂M abgeschlossen ist. Weiter ist M genau dann offen, wenn $M = M^0$ gilt, und genau dann abgeschlossen, wenn $M = \overline{M}$ gilt ([Ü]). Schließlich gilt (auch [Ü])

$$M^0 = \{x \in M : \exists\, \varepsilon > 0 : U_\varepsilon(x) \subset M\}$$

und

$$\overline{M} = \{x \in X : \exists\, (x_n) \text{ in } M : x_n \to x\, (n \to \infty)\}$$

Beispiel 18.5 1. Es sei $(X,d) = (\mathbb{R}^2, d_{|\cdot|}) = (\mathbb{R}^d, d_{\|\cdot\|_2})$. Ist

$$M := \{(x_1,x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = |x|^2 < 1, x_1 \geq 0\},$$

so ist

$$\begin{aligned} \overline{M} &= \{(x_1,x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1, x_1 \geq 0\} \\ M^0 &= \{(x_1,x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 < 1, x_1 > 0\} \end{aligned}$$

und

$$\partial M = \{(x_1,x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1, x_1 > 0\} \cup (\{0\} \times [-1,1]).$$

2. Ist $(X,d) = (\mathbb{R}, d_{|\cdot|})$, so gilt

$$\mathbb{Q}^0 = \emptyset, \qquad \overline{\mathbb{Q}} = \mathbb{R}, \qquad \partial\mathbb{Q} = \mathbb{R}.$$

Insbesondere ist $\mathbb{Q}$ dicht in $\mathbb{R}$.

Definition 18.6 Ein Vektor $\mathbf{v} \in \mathbb{R}^d \setminus \{0\}$ heißt *Richtung* in $\mathbb{R}^d$. Weiter setzen wir für $A \subset \mathbb{K}^d$, $x \in \mathbb{K}^d$ und $B \subset \mathbb{K}$

$$A \pm x := \{a \pm x : a \in A\}, \quad x \pm A := \{x \pm a : a \in A\}, \quad B \cdot x := \{tx : t \in B\}\,.$$

Damit ist für eine Richtung $\mathbf{v}$ in $\mathbb{R}^d$ und $x \in \mathbb{R}^d$ durch $x + \mathbb{R} \cdot \mathbf{v}$ die Gerade durch x in Richtung $\mathbf{v}$ gegeben. Ist speziell $\mathbf{v} = \mathbf{e}^{(k)}$ der k-te Einheitsvektor, so ist $x + \mathbb{R} \cdot \mathbf{e}^{(k)}$ die Gerade durch x parallel zur k-ten Koordinatenachse.
Zudem sei im Weiteren $c^T h := \sum_{j=1}^{d} c_j h_j$ für $c = (c_1, \dots, c_d)$, $h = (h_1, \dots, h_d) \in \mathbb{K}^d$.

Bemerkung und Definition 18.7 Es seien $M \subset \mathbb{R}^d$ und $f : M \to \mathbb{K}$.
1. Ist $x^{(0)} \in M^0$ (also innerer Punkt von M), so heißt f *differenzierbar* an der Stelle $x^{(0)}$, falls ein Vektor $c \in \mathbb{K}^d$ und eine an 0 stetige Funktion $\varepsilon : M_0 := (M - x^{(0)}) \to \mathbb{K}$ existieren mit $\varepsilon(0) = 0$ und

$$f(x^{(0)} + h) = f(x^{(0)}) + c^T h + |h| \cdot \varepsilon(h) \quad (h \in M_0)$$

(vgl. S. 12.3; Zerlegungsformel).
2. Ist $\mathbf{v}$ eine Richtung und ist f differenzierbar an $x^{(0)}$, so gilt für $t \in \mathbb{R}$ genügend klein

$$f(x^{(0)} + t\mathbf{v}) = f(x^{(0)}) + tc^T \mathbf{v} + |t| \cdot |\mathbf{v}| \cdot \varepsilon(t\mathbf{v})$$

und damit für $t \neq 0$

$$\frac{f(x^{(0)} + t\mathbf{v}) - f(x^{(0)})}{t} = c^T \mathbf{v} + \frac{|t|}{t} |\mathbf{v}| \varepsilon(t\mathbf{v}) \to c^T \mathbf{v} \quad (t \to 0)\,.$$

Allgemein heißt f *differenzierbar* an der Stelle $x^{(0)} \in M$ *in Richtung* $\mathbf{v}$, falls

$$\partial_{\mathbf{v}} f(x^{(0)}) := \frac{\partial f}{\partial \mathbf{v}}(x^{(0)}) := D_{\mathbf{v}} f(x^{(0)}) := \lim_{t \to 0} \frac{f(x^{(0)} + t\mathbf{v}) - f(x^{(0)})}{t}$$

existiert. In diesem Fall heißt $\partial_{\mathbf{v}} f(x^{(0)})$ die *Richtungsableitung* von f an $x^{(0)}$ *in Richtung* $\mathbf{v}$. Ist speziell $\mathbf{v} = \mathbf{e}^{(k)}$, so sagt man auch, f sei *partiell differenzierbar* an $x^{(0)}$ *nach der* k*-ten Variablen.* Dann schreiben wir auch $\partial_k f(x^{(0)})$ statt $\partial_{\mathbf{e}^{(k)}} f(x^{(0)})$ und sprechen von der *partiellen Ableitung* von f an $x^{(0)}$ *nach der* k*-ten Variablen.*
Ist f partiell differenzierbar nach allen Variablen, so heißt der Vektor

$$\mathrm{grad} f(x^{(0)}) := \nabla f(x^{(0)}) := (\partial_1 f(x^{(0)}), \ldots, \partial_d f(x^{(0)}))$$

Gradient von f an $x^{(0)}$. Weiter ergibt sich aus der Definition sofort, dass für $\lambda \in \mathbb{R} \setminus \{0\}$

$$\partial_{\lambda \mathbf{v}} f(x^{(0)}) = \partial_{\mathbf{v}} f(x^{(0)}) \cdot \lambda$$

gilt. Man kann sich also i. W. auf Richtungen $\mathbf{v}$ der Länge $|\mathbf{v}| = 1$ beschränken.
3. Die Überlegung am Anfang von 2. zeigt, dass aus der Differenzierbarkeit von f an $x^{(0)}$ insbesondere die Differenzierbarkeit in alle Richtungen $\mathbf{v}$ folgt, und dass $c^T \mathbf{v} = \partial_{\mathbf{v}} f(x^{(0)})$ für jedes c wie in der Definition gilt. Insbesondere ergibt sich

$$c = (c_1, \ldots, c_d) = (c^T \mathbf{e}^{(1)}, \ldots, c^T \mathbf{e}^{(d)}) = \big(\partial_1 f(x^{(0)}), \ldots, \partial_d f(x^{(0)})\big) = \mathrm{grad} f(x^{(0)})$$

(also ist c aus 1. insbesondere eindeutig bestimmt!).
Es gilt damit auch

$$\partial_{\mathbf{v}} f(x^{(0)}) = \mathrm{grad}^T f(x^{(0)}) \mathbf{v}\ . \tag{18.1}$$

Schließlich nennt man die ($\mathbb{R}$-)lineare Abbildung $\mathbb{R}^d \ni h \mapsto c^T h \in \mathbb{K}$ *Ableitung* von f an $x^{(0)}$ und schreibt dafür auch $df(x^{(0)})$ oder $Df(x^{(0)})$ oder kurz $f'(x^{(0)})$.

Bemerkung und Definition 18.8 Die Definition zeigt, dass Richtungs- und partielle Ableitungen nicht anderes als Ableitungen von Funktionen einer (reellen) Variablen sind. Folglich stehen auch alle Rechenregeln und Ergebnisse aus Abschnitt 12 hier zur Verfügung.
Besonders einfach ist die Situation für partielle Ableitungen: Ist $x^{(0)} = (x_1^{(0)}, \ldots, x_d^{(0)}) \in \mathbb{R}^d$, so ist für $k = 1, \ldots, d$

$$\begin{aligned}
\partial_k f(x^{(0)}) &= \lim_{t \to 0} \frac{f(x^{(0)} + t \cdot \mathbf{e}^{(k)}) - f(x^{(0)})}{t} \\
&= \lim_{t \to 0} \frac{f(x_1^{(0)}, \ldots, x_{k-1}^{(0)}, x_k^{(0)} + t, x_{k+1}^{(0)}, \ldots, x_d^{(0)}) - f(x^{(0)})}{t}\ .
\end{aligned}$$

Man sieht, dass $\partial_k f(x^{(0)})$ sich darstellt als die Ableitung der Funktion

$$x_k \mapsto f(x_1, \ldots, x_k, \ldots, x_d)$$

bei festgehaltenen Variablen $x_1, \ldots, x_{k-1}$ und $x_{k+1}, \ldots, x_d$ (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 18.9 1. Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x, y) = x^2 + y$ ($x, y \in \mathbb{R}$). Dann gilt für $\mathbf{v} = (\cos\varphi, \sin\varphi)$ und $x^{(0)} = (0, 0)$

$$f\Big((0,0) + t(\cos\varphi, \sin\varphi)\Big) = t^2\cos^2\varphi + t\sin\varphi \qquad (t \in \mathbb{R})\ .$$

Damit erhalten wir

$$\partial_{\mathbf{v}} f(0,0) = \left(\frac{\partial f}{\partial \mathbf{v}}(0,0) = D_{\mathbf{v}} f(0,0)\right) = \lim_{t\to 0} \frac{t^2\cos^2\varphi + t\sin\varphi}{t} = \sin\varphi\ .$$

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

$$f(x, y, z) = xy^2z^3 \qquad (x, y, z \in \mathbb{R})\ .$$

Dann gilt für alle $(x, y, z) \in \mathbb{R}^3$

$$\begin{aligned}
\partial_1 f(x,y,z) \quad &\left(= \tfrac{\partial f}{\partial x}(x,y,z) \quad = \quad f_x(x,y,z) \right) \quad = \quad y^2 z^3 \\
\partial_2 f(x,y,z) \quad &\left(= \tfrac{\partial f}{\partial y}(x,y,z) \quad = \quad f_y(x,y,z) \right) \quad = \quad 2xyz^3 \\
\partial_3 f(x,y,z) \quad &\left(= \tfrac{\partial f}{\partial z}(x,y,z) \quad = \quad f_2(x,y,z) \right) \quad = \quad 3xy^2 z^2 \; .
\end{aligned}$$

Bemerkung 18.10 Ist $f : M \to \mathbb{K}$ differenzierbar an $x^{(0)}$, so folgt aus der Definition leicht, dass f auch stetig an $x^{(0)}$ ist. Aus der Existenz der partiellen Ableitungen an $x^{(0)}$ folgt im Allgemeinen noch nicht die Stetigkeit (und damit schon gar nicht die Differenzierbarkeit) an $x^{(0)}$:
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

$$f(x_n,y_n) = \frac{1/n^2}{1/n^2+1/n^2} = \frac{1}{2} \to \frac{1}{2} \neq 0 = f(0,0) \qquad (n \to \infty) \; .$$

Bemerkung 18.11 Es seien $M \subset \mathbb{R}^d$, $a \in \mathbb{R}^d$ und $\mathbf{v}$ eine Richtung. Ist I ein offenes Intervall mit $a + I\mathbf{v} \subset M$, so ist $f : M \to \mathbb{K}$ genau dann differenzierbar auf $a + I\mathbf{v}$ in Richtung $\mathbf{v}$, wenn die Funktion $g = g_{a,\mathbf{v}} : I \to \mathbb{K}$, definiert durch

$$g(t) := f(a + t\mathbf{v}) \qquad (t \in I),$$

differenzierbar auf I (im Sinne von D. 12.1) ist. Außerdem gilt dann

$$g'(t) = \partial_{\mathbf{v}} f(a + t\mathbf{v}) \qquad (t \in I)\,.$$

(Denn: Ist $x^{(0)} = a + t_0\mathbf{v}$, so gilt für $t \in I - t_0$

$$\frac{g(t_0 + t) - g(t_0)}{t} = \frac{f(x^{(0)} + t\mathbf{v}) - f(x^{(0)})}{t}\,.)$$

Wir setzen für $x, y \in \mathbb{K}^d$

$$I[x, y] := \big\{x + t(y - x) : t \in [0, 1]\big\}$$

und

$$I(x, y) := I[x, y] \setminus \{x, y\}.$$

Unter Verwendung von (18.1) erhalten wir damit

Satz 18.12 *(Mittelwertsatz)*
Es seien $M \subset \mathbb{R}^d$ und $f : M \to \mathbb{R}$ stetig. Ist $\mathbf{v}$ eine Richtung mit $I[x, x + \mathbf{v}] \subset M$ und ist f differenzierbar auf $I(x, x + \mathbf{v})$ in Richtung $\mathbf{v}$, so existiert ein $\xi \in I(x, x + \mathbf{v})$ mit

$$f(x + \mathbf{v}) - f(x) = \partial_{\mathbf{v}} f(\xi).$$

Beweis. Wir definieren $g : [0, 1] \to \mathbb{R}$ durch

$$g(t) := f(x + t\mathbf{v}) \qquad (t \in [0, 1])\,.$$

Dann ist g stetig auf $[0, 1]$ und (nach B. 18.11) differenzierbar auf $(0, 1)$. Nach dem Mittelwertsatz für Funktionen einer Variable existiert ein $\tau \in (0, 1)$ mit $g(1) - g(0) = g'(\tau)$. Also folgt wieder mit B. 18.11 und $\xi := x + \tau\mathbf{v}$

$$f(x + \mathbf{v}) - f(x) = g(1) - g(0) = \partial_{\mathbf{v}} f(\xi)\,.$$

$\square$

Bemerkung 18.13 Es seien $M \subset \mathbb{R}^d$ und $f : M \to \mathbb{R}$ stetig auf M sowie differenzierbar auf M^0. Dann existiert für alle $x, y \in M$ mit $I(x,y) \subset M^0$ ein $\xi \in I(x,y)$ so, dass

$$f(y) - f(x) = \operatorname{grad}^T f(\xi)(y-x)\,.$$

(Ergibt sich mit $\mathbf{v} := y - x$ (wobei o. E. $x \neq y$) sofort aus dem Mittelwertsatz und (18.1).)

B. 18.10 zeigt, dass die Existenz der partiellen Ableitungen i. A. (anders als im Fall $d = 1$) noch nicht die Stetigkeit impliziert. Man beachte allerdings, dass dort die partiellen Ableitungen in der Nähe von $(0,0)$ unbeschränkt und damit insbesondere unstetig sind (für $x_0 \neq 0$, $y_0 \neq 0$ ist

$$\partial_1 f(0, y_0) = 1/y_0\ , \qquad \partial_2 f(x_0, 0) = 1/x_0\ ).$$

Es gilt jedoch

Satz 18.14 *Es seien $U \subset \mathbb{R}^d$ offen und $f : U \to \mathbb{K}$ so, dass die partiellen Ableitungen $\partial_1 f, \ldots, \partial_d f$ auf U existieren. Sind $\partial_1 f, \ldots, \partial_d f$ stetig an $x^{(0)} \in U$, so ist f differenzierbar an $x^{(0)}$.*

Beweis. Wir können uns auf den Fall $\mathbb{K} = \mathbb{R}$ beschränken (ansonsten: $\operatorname{Re} f$ und $\operatorname{Im} f$ betrachten). Außerdem sei ohne Einschränkung $x^{(0)} = 0$.
Wir zeigen

$$\frac{1}{|h|}\big(f(h) - f(0) - \operatorname{grad}^T f(0) \cdot h\big) \to 0 \qquad (h \to 0),$$

was äquivalent zur Differenzierbarkeit an 0 ist.
Dazu sei $\varepsilon > 0$ gegeben. Dann existiert ein $\delta > 0$ so, dass

$$\big|\partial_k f(x) - \partial_k f(0)\big| < \frac{\varepsilon}{d}$$

für alle $x \in U_\delta(0)$ und $k = 1, \ldots, d$ gilt.
Ist $h \in U_\delta(0), h \neq 0$, so setzen wir

$$v^{(m)} := \sum_{k=1}^{m} h_k \mathbf{e}^{(k)} \qquad (m = 0, \ldots, d)$$

(dann ist $v^{(0)} = 0$ und $v^{(d)} = \sum\limits_{k=1}^{d} h_k \mathbf{e}^{(k)} = (h_1, \ldots, h_d) = h$). Also gilt (Teleskopsumme)

$$f(h) - f(0) = \sum_{k=1}^{d} \big(f(v^{(k)}) - f(v^{(k-1)})\big).$$

Wegen $|v^{(k)}| \leq |h| < \delta$ ist $v^{(k)} \in U_\delta(0)$ für $k = 0, \ldots, d$.
Außerdem ist $v^{(k)} = v^{(k-1)} + h_k \mathbf{e}^{(k)}$, also folgt aus dem Mittelwertsatz die Existenz eines $\xi^{(k)} \in U_\delta(0)$ mit

$$f(v^{(k)}) - f(v^{(k-1)}) = \partial_{h_k \mathbf{e}^{(k)}} f(\xi^{(k)}) = \partial_k f(\xi^{(k)}) \cdot h_k \,.$$

Folglich erhalten wir

$$\frac{1}{|h|} \big| f(h) - f(0) - \operatorname{grad}^T f(0) \cdot h \big| \leq$$

$$\leq \frac{1}{|h|} \sum_{k=1}^{d} \big| f(v^{(k)}) - f(v^{(k-1)}) - \partial_k f(0) \cdot h_k \big|$$

$$= \frac{1}{|h|} \sum_{k=1}^{d} \big| \big(\partial_k f(\xi^{(k)}) - \partial_k f(0)\big) h_k \big| \leq \frac{1}{|h|} \frac{\varepsilon}{d} \sum_{k=1}^{d} |h_k| \leq \varepsilon \,.$$

□

Beispiel 18.15 Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x, y) = e^{xy^2} \qquad (x, y \in \mathbb{R})$$

Dann sind die partiellen Ableitungen

$$\partial_1 f(x, y) = y^2 e^{xy^2} \quad \text{und} \quad \partial_2 f(x, y) = 2xye^{xy^2}$$

stetig auf $\mathbb{R}^2$. Also ist f differenzierbar auf $\mathbb{R}^2$ nach S. 18.14.

Wir wollen zum Schluss des Abschnitts noch einmal kurz auf die geometrische Bedeutung der Ableitung im Mehrdimensionalen eingehen.

Bemerkung 18.16 Ist $f : M \to \mathbb{R}$ differenzierbar an der Stelle $x \in M^0$ mit

$$\mathbf{v}^* := \operatorname{grad} f(x) \neq 0,$$

so gilt für jede Richtung $\mathbf{v}$ nach der Cauchy-Schwarzschen Ungleichung $|(\mathbf{v}^*)^T \cdot \mathbf{v}| \leq |\mathbf{v}^*| \cdot |\mathbf{v}|$, also

$$-|\mathbf{v}^*| \cdot |\mathbf{v}| \leq \partial_{\mathbf{v}} f(x) = (\mathbf{v}^*)^T \cdot \mathbf{v} \leq |\mathbf{v}^*| \cdot |\mathbf{v}|$$

Weiter ist für $\mathbf{v} = \mathbf{v}^*$ (die sog. Gradientenrichtung)

$$\partial_{\mathbf{v}^*} f(x) = (\mathbf{v}^*)^T \cdot \mathbf{v}^* = |\mathbf{v}^*|^2$$

und entsprechend

$$\partial_{-\mathbf{v}^*} f(x) = -|\mathbf{v}^*|^2 \ .$$

Angesichts des Mittelwertsatzes ($f(x+\mathbf{v}) = f(x) + \partial_{\mathbf{v}} f(\xi)$ für ein $\xi \in I(x, x+\mathbf{v})$) kann man also – jedenfalls im Falle stetiger Richtungsableitungen – die Gradientenrichtung als die „Richtung des steilsten Anstiegs“ von f und die negative Gradientenrichtung als die „Richtung des steilsten Abstiegs“ von f ansehen.
Außerdem gilt für $\mathbf{v}^{(0)} \perp \mathbf{v}^*$

$$\partial_{\mathbf{v}^{(0)}} f(x) = (\mathbf{v}^*)^T \cdot \mathbf{v}^{(0)} = 0 \, .$$

d. h. die Richtungsableitungen der zur Gradientenrichtung senkrechten Richtungen verschwinden.
Diese Erkenntnisse werden sich als grundlegend für die mehrdimensionale Optimierung erweisen.

Beispiel 18.17 Es sei $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = x^2 + y^2 \qquad (x, y \in \mathbb{R}) \ .$$

Dann gilt

$$\operatorname{grad} f(x,y) = (2x, 2y) \ .$$

Betrachtet man etwa $(x,y) = (1,1)$, so ist $\mathbf{v}^* = (2,2)$ die Richtung des steilsten Anstiegs. Für $\mathbf{v}^{(0)} = (1,-1)$ gilt dabei $\partial_{\mathbf{v}^{(0)}} f(1,1) = 0$.

19 Taylorsatz und Extremstellen von Funktionen mehrerer Variablen

Wir beschäftigen uns zunächst mit Ableitungen höherer Ordnung.

Definition 19.1 Es seien $U \subset \mathbb{R}^d$ offen und $f : U \to \mathbb{K}$.
1. Sind $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \ldots$ Richtungen in $\mathbb{R}^d$, so definieren wir induktiv für $\ell \geq 2$ (soweit existent!)

$$\partial_{\mathbf{v}^{(\ell)}} \ldots \partial_{\mathbf{v}^{(1)}} f := \partial_{\mathbf{v}^{(\ell)}} (\partial_{\mathbf{v}^{(\ell-1)}} \ldots \partial_{\mathbf{v}^{(1)}} f) : U \to \mathbb{K}$$

(*Richtungsableitungen der Ordnung* ℓ auf U). Für $\mathbf{v}^{(1)} = \ldots = \mathbf{v}^{(\ell)} =: \mathbf{v}$ schreiben wir kurz $\partial_{\mathbf{v}}^{\ell} f$. Weiter setzen wir

$$C_{\mathbf{v}}^{n}(U) := C_{\mathbf{v}}^{n}(U, \mathbb{K}) := \{ f : U \to \mathbb{K} : \partial_{\mathbf{v}}^{\ell} f \text{ existiert für } \ell \leq n \text{ und ist stetig} \} .$$

2. Sind Speziell $\mathbf{v}^{(1)} = \mathbf{e}^{(k_1)}, \ldots, \mathbf{v}^{(\ell)} = \mathbf{e}^{(k_\ell)}$, so schreibt man wieder

$$\partial_{k_\ell} \ldots \partial_{k_1} f \quad \text{oder} \quad \frac{\partial^\ell f}{\partial x_{k_\ell} \ldots \partial x_{k_1}}$$

an Stelle von $\partial_{\mathbf{v}^{(\ell)}} \ldots \partial_{\mathbf{v}^{(1)}} f$ (*partielle Ableitungen der Ordnung* ℓ). Für $k_1 = \ldots = k_\ell =: k$ schreibt man kurz $\partial_k^\ell f$ bzw. $\frac{\partial^\ell f}{\partial x_k^\ell}$.

3. Schließlich setzen wir noch $\partial_{\mathbf{v}}^0 f := f$ und $\partial_k^0 f := f$ (d.h. die Richtungs- bzw. partiellen Ableitungen der Ordung 0 sind f selbst).

Beispiel 19.2 Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x, y) = x^2 y \qquad (x, y \in \mathbb{R}) .$$

Dann gilt

$$\begin{aligned}
\partial_1 f(x,y) &= \frac{\partial f}{\partial x}(x,y) = 2xy \\
\partial_2 f(x,y) &= \frac{\partial f}{\partial y}(x,y) = x^2 \\
\partial_1^2 f(x,y) &= \frac{\partial^2 f}{\partial x^2}(x,y) = 2y \\
\partial_2 \partial_1 f(x,y) &= \frac{\partial^2 f}{\partial y \partial x}(x,y) = 2x \\
\partial_1 \partial_2 f(x,y) &= \frac{\partial^2 f}{\partial x \partial y}(x,y) = 2x \\
\partial_2^2 f(x,y) &= \frac{\partial^2 f}{\partial y^2}(x,y) = 0 .
\end{aligned}$$

Man beachte dabei, dass $\partial_2\partial_1 f = \partial_1\partial_2 f$ gilt.
Weiter erhalten wir etwa

$$\begin{aligned} \partial_1\partial_2\partial_1 f(x,y) &= \frac{\partial^3 f}{\partial x\partial y\partial x}(x,y) = 2 \\ \partial_2\partial_1^2 f(x,y) &= \frac{\partial^3 f}{\partial y\partial x^2}(x,y) = 2 \end{aligned}$$

also ist $\partial_1\partial_2\partial_1 f = \partial_2\partial_1^2 f$. Man sieht, dass die Reihenfolge, in der die partiellen Ableitungen gebildet werden, vertauscht werden kann.

Wir beweisen ganz allgemein für partielle Ableitungen der Ordnung 2:

Satz 19.3 *(Schwarz)*
Es seien $U \subset \mathbb{R}^d$ offen und es seien $p, q \in \{1, \ldots, d\}$. Ferner sei $f : U \to \mathbb{K}$ 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 $p = 1, q = 2$ und $\mathbb{K} = \mathbb{R}$ annehmen. Außerdem sei o. E. $x^{(0)} = (0,0)$. Da U offen ist, existiert ein $R > 0$ so, dass $(x,y) \in U$ für $x, y \in \mathbb{R}$ mit $|x|, |y| < R$. Für solche (x,y) sei

$$\Delta(x,y) := f(x,y) - f(x,0) - f(0,y) + f(0,0) = g(x,y) - g(0,y)$$

mit $g(x,y) := f(x,y) - f(x,0)$.
Zwei Anwendungen des Mittelwertsatzes zeigen, dass ein $\xi = \xi(x,y) \in I[0,x]$ und ein $\eta = \eta(x,y) \in I[0,y]$ existieren mit

$$\Delta(x,y) = \partial_1 g(\xi,y)\cdot x = [\partial_1 f(\xi,y) - \partial_1 f(\xi,0)]\cdot x = \partial_2\partial_1 f(\xi,\eta)\cdot x\cdot y$$

Nun geben wir $\varepsilon > 0$ vor. Da $\partial_2\partial_1 f$ stetig an $(0,0)$ ist, existiert ein $0 < \delta(\leq R)$ so, dass für alle $u, v \in \mathbb{R}$ mit $|u| < \delta$ und $|v| < \delta$

$$|\partial_2\partial_1 f(\xi,\eta) - \partial_2\partial_1 f(0,0)| < \varepsilon$$

gilt. Also erhalten wir für $0 < |x|, |y| < \delta$

$$\left|\frac{\Delta(x,y)}{x\cdot y} - \partial_2\partial_1 f(0,0)\right| < \varepsilon \ .$$

Beachtet man, dass bei festem x

$$\lim_{y\to 0} \frac{\Delta(x,y)}{xy} = \frac{\partial_2 f(x,0) - \partial_2 f(0,0)}{x}$$

ist, so erhält man auch

$$\left| \frac{\partial_2 f(x,0) - \partial_2 f(0,0)}{x} - \partial_2\partial_1 f(0,0) \right| \leq \varepsilon$$

für alle x mit $0 < |x| < \delta$. Da $\varepsilon > 0$ beliebig war, existiert $\partial_1\partial_2 f(0,0)$ und es gilt

$$\partial_1\partial_2 f(0,0) = \partial_2\partial_1 f(0,0) \ .$$

$\square$

Bemerkung und Definition 19.4 1. Ist $U \subset \mathbb{R}^d$ offen und ist $n \in \mathbb{N}_0$, so bezeichnen wir mit $C^n(U) := C^n(U,\mathbb{K})$ die Menge aller Funktionen $f : U \to \mathbb{K}$ mit der Eigenschaft, dass die partiellen Ableitungen der Ordnung $(\leq)\, n$ auf U existieren und dort stetig sind. Außerdem setzen wir

$$C^\infty(U) := C^\infty(U,\mathbb{K}) := \bigcap_{n\in\mathbb{N}_0} C^n(U,\mathbb{K}) \ .$$

Durch mehrfache Anwendung von S. 19.3 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 S. 19.3 wird i. A. falsch, wenn man auf die Voraussetzung der Stetigkeit der partiellen Ableitung $\partial_q\partial_p f$ an $x^{(0)}$ verzichtet. So kann man etwa für die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x,y) := \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2} & , \quad (x,y) \neq (0,0) \\ \quad 0 & , \quad (x,y) = (0,0) \end{cases}$$

zeigen ([Ü]): 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 uns nun einem weiteren zentalen Ergebnis der Analysis zuwenden, dem Taylor-Satz. Wir beweisen zunächst eine Richtungsversion.

Satz 19.5 *(Taylor für Richtungen)*
Es sei $U \subset \mathbb{R}^d$ offen, und es seien $n \in \mathbb{N}_0$ sowie $\mathbf{v}$ eine Richtung. Ferner seien $f \in C^{n+1}_{\mathbf{v}}(U)$ und $x \in U$ mit $I[x, x+\mathbf{v}] \subset U$. Dann gilt

$$f(x+\mathbf{v}) = \sum_{\ell=0}^{n} \frac{\partial^\ell_{\mathbf{v}} f(x)}{\ell!} + \frac{1}{n!} \int\limits_0^1 (1-s)^n \partial^{n+1}_{\mathbf{v}} f(x+s\mathbf{v})\, ds\ .$$

Beweis. Wir beweisen die Behauptung per Induktion nach n.
1. Induktionsanfang $n = 0$: Aus $f \in C^1_{\mathbf{v}}(U)$ folgt nach dem HDI, Teil 2, angewandt auf $g(s) := f(x+s\mathbf{v})$ für $s \in [0,1]$ (vgl. B. 18.11),

$$f(x+\mathbf{v}) - f(x) = \int\limits_0^1 g'(s)\, ds = \int\limits_0^1 \partial_{\mathbf{v}} f(x+s\mathbf{v})\, ds\ .$$

2. Induktionsschritt $n-1 \to n$: Es sei $f \in C^{n+1}_{\mathbf{v}}(U)$. Ist $g(s) := \partial^n_{\mathbf{v}} f(x+s\mathbf{v})$ für $s \in [0,1]$, so ist $g'(s) = \partial^{n+1}_{\mathbf{v}} f(x+s\mathbf{v})$. Also gilt nach Induktionsannahme und mit partieller Integration

$$\begin{aligned} f(x+\mathbf{v}) \quad - \quad \sum_{\ell=0}^{n-1} \frac{\partial^\ell_{\mathbf{v}} f(x)}{\ell!} &= \frac{1}{(n-1)!} \int\limits_0^1 (1-s)^{n-1} g(s)\, ds \\ &= \frac{1}{(n-1)!} \left[-\frac{(1-s)^n}{n} g(s)\Big|_0^1 + \int\limits_0^1 \frac{(1-s)^n}{n} g'(s)\, ds \right] \\ &= \frac{1}{n!} \partial^n_{\mathbf{v}} f(x) + \frac{1}{n!} \int\limits_0^1 (1-s)^n \partial^{n+1}_{\mathbf{v}} f(x+s\mathbf{v})\, ds\ . \end{aligned}$$

□

Bemerkung 19.6 Ist $\mathbb{K} = \mathbb{R}$, so existiert (unter den Bedingungen des Taylor-Satzes) ein $\xi \in I[x, x+\mathbf{v}]$ so, dass

$$\frac{1}{n!} \int\limits_0^1 (1-s)^n \partial^{n+1}_{\mathbf{v}} f(x+s\mathbf{v})\, ds = \frac{1}{(n+1)!} \partial^{n+1}_{\mathbf{v}} f(\xi)\ .$$

(Denn: Man kann zeigen ([Ü]), dass folgende Variante eines (eindimensionalen) Mittelwertsatzes gilt: Ist $[a,b] \subset \mathbb{R}$ und sind $h : [a,b] \to \mathbb{R}$ stetig sowie $g \in R[a,b]$ mit $g \geq 0$ auf $[a,b]$ (oder $g \leq 0$ auf $[a,b]$), so existiert ein $\tau \in [a,b]$ mit

$$\int_a^b hg = h(\tau) \int_a^b g$$

Wendet man dies auf $g(s) := (1-s)^n$ und $h(s) := \partial_{\mathbf{v}}^{n+1} f(x+s\mathbf{v})$ mit $[a,b] = [0,1]$ an, und beachtet man, dass

$$\int_0^1 (1-s)^n \, ds = \frac{1}{(n+1)}$$

gilt, so ergibt sich die Existenz eines $\tau \in [0,1]$ mit

$$\frac{1}{n!} \int_0^1 (1-s)^n \partial_{\mathbf{v}}^{n+1} f(x+s\mathbf{v}) \, ds = \frac{1}{(n+1)!} \partial_{\mathbf{v}}^{n+1} f(x+\tau\mathbf{v}) \; .$$

Für $\xi := x + \tau\mathbf{v}$ ergibt sich die Behauptung.)

Wir wollen uns in einer zweiten Version von Richtungsableitungen befreien. Dazu setzen wir zur Abkürzung für $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$

$$\alpha! = \prod_{j=1}^d (\alpha_j)!, \qquad \partial^\alpha := \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d}$$

und für $h = (h_1, \ldots, h_d) \in \mathbb{R}^d$

$$h^\alpha := \prod_{j=1}^d h_j^{\alpha_j}.$$

In Verallgemeinerung von (18.1) beweisen wir zunächst

Satz 19.7 *Es sei $U \subset \mathbb{R}^d$ offen und es sei $f \in C^n(U)$. Dann ist $f \in C_{\mathbf{v}}^n(U)$ für alle Richtungen $\mathbf{v}$ und es gilt mit $\mathbf{v} = (v_1, \ldots, v_d)$*

$$\partial_{\mathbf{v}}^n f(x) = \sum_{(k_1,\ldots,k_n) \in \{1,\ldots,d\}^n} (\partial_{k_n} \ldots \partial_{k_1} f)(x) \cdot v_{k_1} \cdots v_{k_n} = \sum_{\alpha \in \mathbb{N}_0^d,\, \|\alpha\|_1 = n} \frac{n!}{\alpha!} (\partial^\alpha f)(x) \mathbf{v}^\alpha.$$

Beweis. 1. Wir beweisen die erste Gleichung per Induktion nach n.
$n = 1$: Ist $f \in C^1(U)$, so folgt aus S. 18.14, dass f differenzierbar auf U ist. Aus (18.1) erhalten wir

$$\partial_{\mathbf{v}}^1 f(x) = \partial_{\mathbf{v}} f(x) = \operatorname{grad}^T f(x) \cdot \mathbf{v} = \sum_{k_1=1}^{d} \partial_{k_1} f(x) \cdot v_{k_1} \qquad (x \in U) .$$

$n \to n+1$: Da $f \in C^{n+1}(U)$ ist, folgt $\partial_{\mathbf{v}}^n f \in C^1(U)$ mit der Induktionsvoraussetzung (man beachte: $x \mapsto \partial_{\mathbf{v}}^n f(x)$ ist eine Linearkombination von partiellen Ableitungen der Ordnung n). Also ergibt sich wie oben beim Induktionsanfang

$$\begin{aligned}\partial_{\mathbf{v}}^{n+1} f(x) &= \partial_{\mathbf{v}}(\partial_{\mathbf{v}}^n f)(x) = \operatorname{grad}^T (\partial_{\mathbf{v}}^n f)(x) \cdot \mathbf{v} \\ &= \sum_{k_{n+1}=1}^{d} \sum_{k_1,\dots,k_n=1}^{d} \partial_{k_{n+1}}(\partial_{k_n} \dots \partial_{k_1} f)(x) v_{k_1} \cdots v_{k_n} \cdot v_{k_{n+1}} \qquad (x \in U) ,\end{aligned}$$

woraus die Behauptung für $n+1$ folgt.
2. Nach B. 19.4 kann man die Reihenfolgen der partiellen Ableitungen in der ersten Summe beliebig permutieren. Da $\frac{n!}{\alpha_1! \dots \alpha_d!}$ Tupel $(k_1, \dots, k_n) \in \{1, \dots, d\}^n$ existieren, bei denen die Zahl $j \in \{1, \dots, d\}$ genau α_j-mal vorkommt (siehe B. 3.11 für den Fall $d = 2$), ergibt sich auch die zweite Gleichung, also

$$\partial_{\mathbf{v}}^n f(x) = \sum_{(\alpha_1,\dots,\alpha_d) \in \mathbb{N}_0^d,\ \sum_{j=1}^{d} \alpha_j = n} \frac{n!}{\alpha_1! \cdots \alpha_d!} (\partial_1^{\alpha_1} \dots \partial_d^{\alpha_d} f)(x) v_1^{\alpha_1} \cdots v_d^{\alpha_d} .$$

□

Beispiel 19.8 Wir werden nun sehr bescheiden und betrachten speziell die Fälle $n = 0$ und $n = 1$ im Taylor-Satz und in B. 19.6.
Ist $f \in C^1(U, \mathbb{R})$, so ergibt sich für $x, h \in U$ mit $I[x, x+h] \subset U$ die Existenz eines $\xi \in I[x, x+h]$ so, dass

$$f(x+h) = f(x) + \partial_h f(\xi) = f(x) + \operatorname{grad}^T f(\xi) h$$

also wieder die Aussage des Mittelwertsatzes.
Ist sogar $f \in C^2(U, \mathbb{R})$, so gilt nach S. 19.7

$$\partial_h^2 f = \sum_{j,k=1}^{d} (\partial_k \partial_j f) h_j h_k = h^T (Hf) h ,$$

wobei für $y \in U$

$$(Hf)(y) := \big(\partial_k \partial_j f(y)\big)_{j,k=1,\ldots,d} \in \mathbb{R}^{d\times d}$$

die sog. *Hesse-Matrix* von f an der Stelle y bezeichnet. Damit erhalten wir mit einem $\xi \in I[x, x+h]$

$$f(x+h) = f(x) + \partial_h f(x) + \frac{1}{2}\partial_h^2 f(\xi) = f(x) + \operatorname{grad}^T f(x) \cdot h + \frac{1}{2} h^T Hf(\xi) h\ .$$

Satz 19.9 *(Taylor)*
Es seien $U \subset \mathbb{R}^d$ offen und $f \in C^{n+1}(U,\mathbb{R})$ für ein $n \in \mathbb{N}_0$. Sind $x, h \in U$ mit $I[x, x+h] \subset U$, so existiert ein $\xi \in I[x, x+h]$ so, dass

$$f(x+h) = \sum_{\|\alpha\|_1 \le n} \frac{\partial^\alpha f(x)}{\alpha!}\, h^\alpha + \sum_{\|\alpha\|_1 = n+1} \frac{\partial^\alpha f(\xi)}{\alpha!}\, h^\alpha.$$

Beweis. O. E. sei $h \neq 0$. Dann ergibt sich aus dem Taylor-Satz für Richtungen und S. 19.7 mit ξ wie in B.19.6

$$f(x+h) - \sum_{\|\alpha\|_1 = n+1} \frac{\partial^\alpha f(\xi)}{\alpha!}\, h^\alpha = \sum_{\ell=0}^{n} \frac{1}{\ell!} \sum_{\|\alpha\|_1 = \ell} \frac{\ell!}{\alpha!} \partial^\alpha f(x)\, h^\alpha = \sum_{\|\alpha\|_1 \le n} \frac{\partial^\alpha f(x)}{\alpha!}\, h^\alpha.$$

□

Bemerkung 19.10 Der erste Summand

$$T_{n,x}(h) := \sum_{\|\alpha\|_1 \le n} \frac{\partial^\alpha f(x)}{\alpha!}\, h^\alpha$$

ist bei festem x ein Polynom der d Variablen $h_1, \ldots, h_d$ (vom Grad $\le n$). Dieses Polynom heißt n-tes *Taylor-Polynom* von f bezüglich der Entwicklungsmitte x. Ist speziell $d = 1$, so ist

$$T_{n,x}(h) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x)}{\nu!}\, h^\nu \qquad (h \in \mathbb{R})$$

ein Polynom einer Variablen.
Der zweite Summand $\sum\limits_{\|\alpha\|_1 = n+1} \frac{\partial^\alpha f(\xi)}{\alpha!}\, h^\alpha$ heißt Restglied in *Lagrange-Form.*

Als wesentliche Anwendung des Taylor-Satzes werden wir nun ein hinreichendes Kriterium für lokale Extrema herleiten. Zunächst gilt folgendes wichtige **notwendige** Kriterium für Extremstellen.

Satz 19.11 *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 Maximum oder Minimum, so ist $x^{(0)}$ ein kritischer Punkt, d. h.*

$$\operatorname{grad} f(x^{(0)}) = 0 .$$

Beweis. Es sei I ein offenes Intervall mit $0 \in I$ und so, dass $x^{(0)} + I\mathbf{e}^{(k)} \subset U$ für alle k gilt. Hat f an $x^{(0)}$ ein lokales Extremum, so haben auch alle Funktionen $g_k : I \to \mathbb{R}$ mit

$$g_k(t) := f(x^{(0)} + t\mathbf{e}^{(k)}) \qquad (t \in I)$$

ein lokales Extremum an $t_0 = 0$. Also gilt nach S. 13.2

$$0 = g_k'(0) = \partial_k f(x^{(0)}) \qquad (k = 1, \ldots, d) .$$

□

Beispiel 19.12 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 $(0, 0)$ ein (offenbar 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 $(x_0, y_0) = (0, 0)$ 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.

Satz 19.13 *Für $A \in \mathbb{K}^{m\times d}$ sei*

$$\|A\| := \sup\left\{|Ax| : x \in \mathbb{K}^d, |x| \le 1\right\}.$$

Dann gilt:

1. *$\|\cdot\|$ ist eine Norm auf $\mathbb{K}^{m\times d}$ (die sog. Operatornorm).*

2. *Für alle $x \in \mathbb{K}^d$ ist $|Ax| \le \|A\| \cdot |x|$.*

3. *Ist $B \in \mathbb{K}^{p\times m}$, so ist $\|BA\| \le \|B\| \cdot \|A\|$.*

4. *Ist $A = (a_{jk})_{\substack{j=1,\dots,m\\k=1,\dots,d}}$, so ist*

$$\max_{j,k} |a_{jk}| \le \|A\| \le d\sqrt{m} \max_{j,k} |a_{jk}|.$$

Beweis.
1. Axiome nachrechnen.
2. Ohne Einschränkung sei $x \neq 0$. Dann gilt

$$|Ax| = \left|A\Big(|x|\,\frac{x}{|x|}\Big)\right| = |x|\left|A\Big(\frac{x}{|x|}\Big)\right| \le \|A\| \cdot |x|.$$

3. Für $x \in \mathbb{K}^d, |x| \le 1$ gilt mit 2.

$$|BAx| \le \|B\| \cdot |Ax| \le \|B\| \cdot \|A\|.$$

4. ([Ü]) □

Ist $f \in C^2(U)$ für eine offene Menge U in $\mathbb{R}^d$, so ist die Hesse-Matrix

$$Hf(x) = \big(\partial_k\partial_j f(x)\big)_{j,k=1,\dots,d} \in \mathbb{R}^{d\times d}$$

nach dem Satz von Schwarz (S. 19.3) für alle $x \in U$ symmetrisch. Für das Weitere müssen wir kurz auf symmetrische Matrizen eingehen.

Definition 19.14 Ist $A \in \mathbb{R}^{d\times d}$ symmetrisch, so heißt A

1. *positiv definit*, falls $\mathbf{v}^T A\mathbf{v} > 0$ für alle Richtungen $\mathbf{v}$ gilt,

2. *positiv semidefinit*, falls $\mathbf{v}^T A\mathbf{v} \ge 0$ für alle Richtungen $\mathbf{v}$ gilt,

3. *negativ (semi-) definit*, falls $-A$ positiv (semi-) definit ist,

4. *indefinit*, falls A weder positiv noch negativ semidefinit ist.

Bemerkung 19.15 Für Charakterisierungen der Definitheitsbegriffe aus D. 19.14 verweisen wir auf die Lineare Algebra.

1. Im Falle $d = 2$ ergibt sich (leicht): $A = (a_{jk}) \in \mathbb{R}^{2\times 2}$ ist genau dann positiv definit (bzw. semidefinit), wenn $\det(A) > 0$ (bzw. ≥ 0) und $a_{11} > 0$ (bzw. $a_{11}, a_{22} \geq 0$) gilt. Entsprechend ist A ist genau dann negativ definit (bzw. semidefinit), wenn $\det(A) > 0$ (bzw. ≥ 0) und $a_{11} < 0$ (bzw. $a_{11}, a_{22} \leq 0$) gilt.
2. Wir setzen

$$S^{d-1} := \{\mathbf{v} \in \mathbb{R}^d : |\mathbf{v}| = 1\}$$

(Einheitssphäre in $\mathbb{R}^d$). Dann ist S^{d-1} beschränkt und abgeschlossen, also kompakt (Satz von Heine-Borel). Weiter ist die Funktion $\mathbf{v} \mapsto \mathbf{v}^T A\mathbf{v}$ stetig auf $\mathbb{R}^d$. Damit existieren $\min\limits_{\mathbf{v}\in S^{d-1}} \mathbf{v}^T A\mathbf{v}$ und $\max\limits_{\mathbf{v}\in S^{d-1}} \mathbf{v}^T A\mathbf{v}$. Genauer gilt (nicht leicht): Ist $\sigma(A) = \{\lambda_1, \ldots, \lambda_d\}$ die Menge der Eigenwerte von A (das sog. Spektrum von A), so ist

$$\min_{\mathbf{v}\in S^{d-1}} \mathbf{v}^T A\mathbf{v} = \min \sigma(A) \quad \text{und} \max_{\mathbf{v}\in S^{d-1}} \mathbf{v}^T A\mathbf{v} = \max \sigma(A) \ .$$

Insbesodere folgt daraus

$$A \left\{ \begin{array}{l} \text{positiv definit} \\ \text{positiv semidefinit} \\ \text{negativ definit} \\ \text{negativ semidefinit} \end{array} \right\} \Longleftrightarrow \left\{ \begin{array}{l} \lambda_1, \ldots, \lambda_d > 0 \\ \lambda_1, \ldots, \lambda_d \geq 0 \\ \lambda_1, \ldots, \lambda_d < 0 \\ \lambda_1, \ldots, \lambda_d \leq 0 \end{array} \right\} .$$

Wir betrachten jetzt Funktionen $f \in C^2(U, \mathbb{R})$. Da $\partial_k \partial_j f$ für alle j, k stetig auf U ist, ist auch $Hf : (U, |\cdot|) \to (\mathbb{R}^{d\times d}, ||\cdot||)$ nach S. 19.13.4 stetig.
Es gilt nun folgendes für die mehrdimensionale Optimierung zentrale Resultat

Satz 19.16 *Es seien $U \subset \mathbb{R}^d$ offen und $f \in C^2(U, \mathbb{R})$. Ferner sei $x^{(0)} \in U$ ein kritischer Punkt (d. h.* $\operatorname{grad} f(x^{(0)}) = 0$*). Dann gilt:*

1. *Ist $Hf(x^{(0)})$ positiv definit, so hat f an $x^{(0)}$ ein (striktes) lokales Minimum.*
2. *Ist $Hf(x^{(0)})$ negativ definit, so hat f an $x^{(0)}$ ein (striktes) lokales Maximum.*
3. *Hat f an $x^{(0)}$ ein lokales Minimum, so ist $Hf(x^{(0)})$ positiv semidefinit.*
4. *Hat f an $x^{(0)}$ ein lokales Maximum, so ist $Hf(x^{(0)})$ negativ semidefinit.*

Beweis. Es genügt, die Behauptungen 1. und 3. zu beweisen. Die Aussagen 2. und 4. ergeben sich dann durch Betrachtung von $-f$.

1. Es sei $A := Hf(x^{(0)})$ positiv definit. Wir setzen

$$\varepsilon := \min_{\mathbf{v} \in S^{d-1}} \mathbf{v}^T A \mathbf{v} (> 0) .$$

Nach der Vorbemerkung existiert ein $\delta = \delta_\varepsilon > 0$ mit $||Hf(x) - A|| < \varepsilon$ für alle $x \in U_\delta(x^{(0)})$.
Es sei $h \in U_\delta(0)$, $h \neq 0$. Nach B. 19.8 existiert ein $\xi \in I[x^{(0)}, x^{(0)} + h] \subset U_\delta(x^{(0)})$ mit

$$\begin{aligned} f(x^{(0)} + h) - f(x^{(0)}) &= \frac{1}{2} h^T Hf(\xi) h = \frac{1}{2} h^T (Hf(\xi) - A) h + \frac{1}{2} h^T A h \\ &\geq \frac{1}{2} h^T A h - \frac{1}{2} \left| h^T (Hf(\xi) - A) h \right| \end{aligned}$$

(man beachte: $x^{(0)}$ ist kritischer Punkt!). Dabei gilt

$$h^T A h = |h|^2 \left(\frac{h}{|h|} \right)^T A \left(\frac{h}{|h|} \right) \geq \varepsilon |h|^2$$

Mit der Cauchy-Schwarz'schen Ungleichung ergibt sich für für beliebige Matrizen $B \in \mathbb{R}^{d \times d}$ und $x \in \mathbb{R}^d$

$$|x^T B x| \leq |x| \cdot |Bx| \leq ||B|| \cdot |x|^2 .$$

Also folgt

$$|h^T (Hf(\xi) - A) h| \leq ||Hf(\xi) - A|| \cdot |h|^2 < \varepsilon |h|^2$$

und damit ist $f(x^{(0)} + h) > f(x^{(0)})$.
2. Die Funktion f habe an $x^{(0)}$ ein lokales Minimum. Ferner sei $\mathbf{v}$ eine Richtung. Dann gilt für $|t|$ genügend klein

$$f(x^{(0)} + t\mathbf{v}) - f(x^{(0)}) \geq 0 .$$

Wieder B. 19.8 existiert (für $t \neq 0$) ein $\xi_t \in I[x^{(0)}, x^{(0)} + t\mathbf{v}]$ mit

$$0 \leq \frac{f(x^{(0)} + t\mathbf{v}) - f(x^{(0)})}{t^2} = \frac{1}{2} \mathbf{v}^T Hf(\xi_t) \mathbf{v}$$

Aus $\xi_t \to x^{(0)}$ für $t \to 0$ folgt $||Hf(\xi_t) - A|| \to 0$ und damit (siehe 1.) auch $\mathbf{v}^T Hf(\xi_t) \mathbf{v} \to \mathbf{v}^T A \mathbf{v}$ für $t \to 0$. Also ist auch $\mathbf{v}^T A \mathbf{v} \geq 0$. □

Beispiel 19.17 1. Ist $f(x, y) = x^2 + y^2$ für $x, y \in \mathbb{R}$ (vgl. B. 19.12.1), so gilt

$$Hf(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

also ist Hf 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. 19.12.2) so gilt

$$Hf(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} .$$

Hier ist Hf stets indefinit. Also hat f nach S. 19.16.3./4. keine lokalen Extrema.
3. Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x,y) = 2x^3 - 3x^2 + 2y^3 + 3y^2$ $(x, y \in \mathbb{R})$. 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

$$Hf(x,y) = \begin{pmatrix} 12x - 6 & 0 \\ 0 & 12y + 6 \end{pmatrix}$$

Also ist

$$\begin{aligned} Hf(0,0) &= \begin{pmatrix} -6 & 0 \\ 0 & 6 \end{pmatrix} & \text{indefinit} \\ Hf(0,-1) &= \begin{pmatrix} -6 & 0 \\ 0 & -6 \end{pmatrix} & \text{negativ definit} \\ Hf(1,0) &= \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix} & \text{positiv definit} \end{aligned}$$

und

$$Hf(1,-1) = \begin{pmatrix} 6 & 0 \\ 0 & -6 \end{pmatrix} \quad \text{indefinit}$$

Damit ist f an $(0,-1)$ ein lokales Maximum, an $(1,0)$ ein lokales Minimum und ansonsten keine Extremstellen.

20 Hauptsätze der mehrdimensionalen Analysis

Wir werden uns in diesem Abschnitt mit zentralen Ergebnissen der mehrdimensionalen Analysis beschäftigen. Dazu starten wir mit einem sehr allgemeinen Hilfsresultat.

Bemerkung und Definition 20.1 Es seien (X, d) ein metrischer Raum, $M \subset X$ und $\alpha \in [0, 1)$. Eine Abbildung $\varphi : M \to X$ heißt *α-Kontraktion*, falls

$$d(\varphi(x), \varphi(y)) \leq \alpha d(x, y) \qquad (x, y \in M) .$$

Offensichtlich ist jede α-Kontraktion stetig. Außerdem existiert höchstens ein $x^* \in M$ mit $x^* = \varphi(x^*)$ d. h. φ hat höchstens einen Fixpunkt.
(Denn: Ist $\tilde{x}$ ein weiterer Fixpunkt von φ, so ist

$$d(\tilde{x}, x^*) = d(\varphi(\tilde{x}), \varphi(x^*)) \leq \alpha d(\tilde{x}, x^*) ,$$

also $d(\tilde{x}, x^*) = 0$.)

Satz 20.2 *(Banachscher Fixpunktsatz)*
Es seien (X, d) ein vollständiger metrischer Raum, $M \subset X$ abgeschlossen und $\varphi : M \to X$ mit $\varphi(M) \subset M$ (eine Selbstabbildung auf M). Ist φ eine α-Kontraktion, so existiert genau ein Fixpunkt $x^* \in M$ und zudem konvergiert für alle $x_0 \in M$ die Folge (x_n) mit

$$x_{n+1} := \varphi(x_n) \qquad (n \in \mathbb{N}_0)$$

gegen x^*. Genauer 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 $x_n \to x^*$ $(n \to \infty)$ und aufgrund der Abgeschlossenheit von M ist auch $x^* \in M$. Da φ insbesondere stetig 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

$$(1-\alpha) d(x_0, x^*) \leq d(x_0, x_1) .$$

□

Bemerkung und Definition 20.3 Es seien $M \subset \mathbb{R}^d$ und $f : M \to \mathbb{K}^m$ für ein $m \in \mathbb{N}$. Ist $f = (f_1, \ldots, f_m)$ mit $f_j : M \to \mathbb{K}$, so heißt f *differenzierbar* an der Stelle $x^{(0)} \in M^0$, falls $f_1, \ldots, f_m$ differenzierbar an $x^{(0)}$ sind.
Nach B./D.18.7 gilt: $f = (f_1, \ldots, f_m)$ ist genau dann differenzierbar an $x^{(0)}$, wenn eine Matrix $C \in \mathbb{K}^{m \times d}$ und eine Funktion an 0 stetige Funktion $\varepsilon : M - x^{(0)} \to \mathbb{K}^m$ existieren mit $\varepsilon(h) = 0$ und

$$f(x^{(0)} + h) = f(x^{(0)}) + Ch + |h| \varepsilon(h) .$$

Nach B./D. 18.7.3 ist C eindeutig bestimmt und es gilt

$$C = \begin{pmatrix} \operatorname{grad}^T f_1(x^{(0)}) \\ \vdots \\ \operatorname{grad}^T f_m(x^{(0)}) \end{pmatrix} =: Jf(x^{(0)}).$$

Die Matrix $Jf(x^{(0)})$ heißt *Jacobi-Matrix* von f an der Stelle $x^{(0)}$. Schließlich heißt wieder die ($\mathbb{R}$-)lineare Abbildung $\mathbb{R}^d \ni h \mapsto Ch \in \mathbb{K}^m$ *Ableitung* von f an $x^{(0)}$.

Satz 20.4 *(Kettenregel)*
Es seien $M \subset \mathbb{R}^d$, $L \subset \mathbb{R}^m$ und $f : M \to L$ differenzierbar an $x^{(0)}$ sowie $g : L \to \mathbb{K}^p$ differenzierbar an $f(x^{(0)})$. Dann ist auch $g \circ f$ differenzierbar an $x^{(0)}$ mit

$$J(g \circ f)(x^{(0)}) = (Jg)\big(f(x^{(0)})\big) \cdot (Jf)(x^{(0)}).$$

Beweis. Wir setzen $C := Jf(x^{(0)})$ und $D := Jg(y^{(0)})$, wobei $y^{(0)} := f(x^{(0)})$. Nach Voraussetzung existieren an 0 stetige Funktionen $\varepsilon : M - x^{(0)} \to \mathbb{R}^m$ und $\tau : L - y^{(0)} \to \mathbb{K}^p$ mit $\varepsilon(0) = 0$, $\tau(0) = 0$ und

$$\begin{aligned} f(x^{(0)} + h) &= f(x^{(0)}) + C \cdot h + |h|\varepsilon(h) \\ g(y^{(0)} + u) &= g(y^{(0)}) + D \cdot u + |u|\tau(u)\ . \end{aligned}$$

Setzt man $\varphi(h) := C \cdot h + |h|\varepsilon(h)$, so gilt für $|h| \neq 0$ genügend klein

$$\frac{|\varphi(h)|}{|h|} \leq ||C|| + |\varepsilon(h)| \leq ||C|| + 1.$$

Also erhält man

$$\begin{aligned} \frac{1}{|h|}|g(f(x^{(0)} + h)) - g(f(x^{(0)})) - DCh| &= \frac{1}{|h|}\big|D(|h|\varepsilon(h)) + |\varphi(h)|\tau(\varphi(h))\big| \\ &\leq ||D|| \cdot |\varepsilon(h)| + (||C|| + 1)|\tau(\varphi(h))| \to 0 \end{aligned}$$

für $h \to 0$. Hieraus ergibt sich die Behauptung (vgl. Beweis zu S. 18.14). □

Beispiel 20.5 Es seien $f : \mathbb{R}^2 \to \mathbb{R}^2$ und $g : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x, y) = \begin{pmatrix} e^x \cos y \\ e^x \sin y \end{pmatrix}, \qquad g(u, v) := uv\ .$$

Dann gilt

$$(g \circ f)(x, y) = e^{2x} \cos y \sin y\ ,$$

also

$$\operatorname{grad}^T(g \circ f)(x, y)\Big(= J(g \circ f)(x, y)\Big) = (2e^{2x} \cos y \sin y\ ,\ e^{2x}(\cos^2 y - \sin^2 y))\ .$$

Andererseits gilt auch

$$Jf(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}\ ,$$

und

$$\mathrm{grad}^T g(u,v) = Jg(u,v) = (v,u) \ ,$$

also

$$\begin{aligned} \mathrm{grad}^T\, g(f(x,y)) \cdot Jf(x,y) &= (e^x \sin y, e^x \cos y) \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} \\ &= (2e^{2x} \sin y \cos y, e^{2x}(\cos^2 y - \sin^2 y)) \ . \end{aligned}$$

Eine wichtige Folgerung aus dem Mittelwertsatz ist der Schrankensatz (vgl. S.13.6). Wir zeigen nun, dass dieser auch für vektorwertige Funktionen gilt.

Satz 20.6 *(Schrankensatz)*
Es sei $M \subset \mathbb{R}^d$ und $f : M \to \mathbb{K}^m$ sei stetig auf M sowie differenzierbar auf M^0. Sind $x, y \in M$ mit $I(x,y) \subset M^0$, so existiert ein $\xi \in I(x,y)$ mit

$$|f(y) - f(x)| \leq ||Jf(\xi)|| \cdot |y - x| \ .$$

Beweis. O. E. können wir $\mathbb{K} = \mathbb{R}$ annehmen (ansonsten: f als Funktion nach $\mathbb{R}^{2m}$ auffassen).
Es sei $g : M \to \mathbb{R}$ definiert durch

$$g(u) := (f(y) - f(x))^T f(u) \qquad (u \in M) \ .$$

Dann ist g stetig auf M und es gilt mit der Kettenregel

$$\mathrm{grad}^T g(u) = ((f(y) - f(x))^T (Jf)(u) \qquad (u \in M^0) \ .$$

Nach dem Mittelwertsatz existiert ein $\xi \in I(x,y)$ so, dass

$$\begin{aligned} |f(y) - f(x)|^2 &= g(y) - g(x) = \mathrm{grad}^T g(\xi) \cdot (y - x) \\ &= (f(y) - f(x))^T Jf(\xi)(y - x) \leq |f(y) - f(x)| \cdot |Jf(\xi)(y - x)| \\ &\leq |f(y) - f(x)| \cdot ||Jf(\xi)|| \cdot |y - x| \ . \end{aligned}$$

Ist $f(x) = f(y)$, so ist die Behauptung klar und ist $f(x) \neq f(y)$ ergibt sich die Behauptung nach Division durch $|f(y) - f(x)|$. $\square$

Wir beschäftigen uns nun mit der „lokalen Umkehrbarkeit“ von Funktionen $f : \Omega \to \mathbb{R}^d$, wobei $\Omega \subset \mathbb{R}^d$. Ist $f_{|U}$ für eine Menge $U \subset \Omega$ injektiv, so bedeutet dies, dass für jedes $y = (y_1, \ldots, y_d) \in f(U)$ 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 $x = (x_1, \ldots, x_d) \in U$ besitzt, nämlich $x = f_U^{-1}(y)$ wobei

$$f_U^{-1} := \Big(f_{|U} : U \to f(U)\Big)^{-1}.$$

Betrachten wir zunächst zwei bekannte Spezialfälle:
1. (linearer Fall) Ist $f : \mathbb{R}^d \to \mathbb{R}^d$ definiert durch

$$f(x) := Ax \qquad (x \in \mathbb{R}^d)$$

wobei $A \in \mathbb{R}^{d\times d}$ mit $\det(A) \neq 0$, so f bijektiv, d. h. das lineare Gleichungssystem

$$Ax = f(x) = y$$

für alle $y \in \mathbb{R}^d$ eindeutig lösbar. Außerdem gilt dann

$$x = f^{-1}(y)(= A^{-1}y) \qquad (y \in \mathbb{R}^d)\ .$$

2. (skalarer Fall) Ist $\Omega \subset \mathbb{R}$ offen und $f \in C^1(\Omega)$ mit $f'(x_0) \neq 0$ für ein $x_0 \in \Omega$, so existiert ein $\delta > 0$ so, dass $f'(x)$ entweder durchgehend > 0 oder < 0 auf $U := U_\delta(x_0)$ und damit f streng monoton auf U ist. Also existiert f_U^{-1} und es gilt für $y = f(x) \in f(U)$

$$(f_U^{-1})'(y) = \frac{1}{f'(x)}\ .$$

Wir setzen für eine offene Menge $\Omega \subset \mathbb{R}^d$ und $k, m \in \mathbb{N}$

$$C^k(\Omega, \mathbb{K}^m) := \{f = (f_1, \ldots, f_m)^T : \Omega \to \mathbb{K}^m : f_1, \ldots, f_m \in C^k(\Omega, \mathbb{K})\}\ .$$

Insbesondere gilt dabei $f \in C^1(\Omega, \mathbb{K}^m)$ genau dann, wenn $f : \Omega \to \mathbb{K}^m$ differenzierbar und $Jf : \Omega \to \mathbb{K}^{m\times d}$ (versehen mit der Operatornorm $||\cdot||$) stetig ist (ergibt sich leicht aus S. 19.13.4.)

Satz 20.7 *(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

$$\det Jf(x^{(0)}) \neq 0 \, .$$

Dann existiert eine offene Umgebung U von $x^{(0)}$ so, dass $f_{|U}$ injektiv und $V := f(U)$ offen ist. Außerdem ist $f_U^{-1} \in C^1(V, \mathbb{R}^d)$ und es gilt

$$Jf_U^{-1}(y) = [Jf(x)]^{-1} \qquad (y = f(x) \in V) \, .$$

Beweis. Wir setzen $A := Jf(x^{(0)})$. Nach Voraussetzung existiert $A^{-1}(\neq 0)$. Da $x \mapsto Jf(x)$ stetig auf Ω ist, existiert ein $\delta > 0$ so, dass für $a := (2||A^{-1}||)^{-1}$ gilt

$$||Jf(x) - A|| < a \qquad \Big(x \in U_\delta(x^{(0)})\Big) \, .$$

Damit definieren wir

$$U := U_\delta(x^{(0)}) \, , \qquad V := f(U)$$

und zeigen:

1. $f_{|U} : U \to \mathbb{R}^d$ ist injektiv, d. h. $(f_{|U})^{-1} : V \to U$ existiert
2. $V \subset \mathbb{R}^d$ ist offen
3. $\det Jf(x) \neq 0 \quad (x \in U)$
4. $f_U^{-1} \in C^1(V, \mathbb{R}^d)$ und $Jf_U^{-1}(y) = [Jf(x)]^{-1}$ für alle $y \in V$, $y = f(x)$.

An verschiedenen Stellen wird folgende Hilfsfunktion von Nutzen sein: Es sei $y \in \mathbb{R}^d$ und $\varphi = \varphi_y : U \to \mathbb{R}^d$ definiert durch

$$\varphi(x) := x + A^{-1}(y - f(x)) \qquad (x \in U) \, .$$

Dann ist $\varphi(x) = x$ genau dann, wenn $y = f(x)$ gilt. Ferner gilt für $x \in U$

$$J\varphi(x) = E - A^{-1}Jf(x) = A^{-1}(A - Jf(x))$$

und damit

$$||J\varphi(x)|| \leq ||A^{-1}|| \cdot ||A - Jf(x)|| < ||A^{-1}||a = \frac{1}{2} \, .$$

Also ist φ nach dem Schrankensatz eine 1/2-Kontraktion.
1. Es sei $y \in \mathbb{R}^d$ und $\varphi = \varphi_y$ wie in 1. Dann hat φ höchstens einen Fixpunkt nach B./D. 20.1. Also ist $f_{|U}$ injektiv.

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 zeigen: $U_{a\rho}(y^{(1)}) \subset V$.
Dazu sei $y \in U_{a\rho}(y^{(1)})$ gegeben und $\varphi = \varphi_y$ wie in 1. 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 $\varphi(x) \in M$. Da $M \subset \mathbb{R}^d$ abgeschlossen ist, folgt aus dem Banachschen Fixpunktsatz, dass (genau) ein $x \in M$ existiert mit $\varphi(x) = x$, also $f(x) = y$.

3. Es gilt nach 1. (wobei $\varphi = \varphi_y$ mit einem beliebigen $y \in V$)

$$Jf(x) = A \cdot (E - J\varphi(x)) \qquad (x \in U)$$

also

$$\det Jf(x) = \det(A) \cdot \det(E - J\varphi(x)) \qquad (x \in U)\ .$$

Wieder nach 1. ist $||J\varphi(x)|| \leq \frac{1}{2}$ $(x \in U)$ 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 $\det Jf(x) \neq 0$.
4. Es sei $y \in V, y = f(x)$. Dann gilt für alle $h \in U - x$

$$\begin{aligned} \frac{1}{2}\,|h| &\geq \big|\varphi(x+h) - \varphi(x)\big| \\ &= \big|x + h - x - A^{-1}\big(f(x+h) - f(x)\big)\big| \\ &\geq |h| - \|A^{-1}\| \cdot \big|f(x+h) - f(x)\big|, \end{aligned}$$

also

$$\big|f(x+h) - f(x)\big| \geq \frac{1}{2}\,\frac{1}{\|A^{-1}\|}\,|h| = a|h|.$$

Für $u \in V - y, u \neq 0$ setzen wir

$$h(u) := f_U^{-1}(y+u) - f_U^{-1}(y).$$

Dann ist $u = f(h(u) + x) - f(x)$ und damit

$$|u| \geq a\big|h(u)\big|$$

(insbesondere zeigt dies, dass f_U^{-1} stetig ist). Außerdem ergibt sich

$$\begin{aligned}
&\frac{1}{|u|}\left|f_U^{-1}(y+u) - f_U^{-1}(u) - \left[Jf(x)\right]^{-1}u\right| \\
&\leq \left\|\left[Jf(x)\right]^{-1}\right\| \cdot \frac{1}{a} \cdot \frac{\left|(Jf)(x)h(u) - u\right|}{\left|h(u)\right|} \\
&= \left\|\left[Jf(x)\right]^{-1}\right\| \cdot \frac{1}{a} \cdot \frac{\left|f\big(x+h(u)\big) - f(x) - (Jf)(x)h(u)\right|}{\left|h(u)\right|} \\
&\qquad \longrightarrow 0 \qquad (u \to 0),
\end{aligned}$$

da f differenzierbar an x ist und $h(u) \to 0$ für $u \to 0$ gilt.
Also ist f_U^{-1} differenzierbar an y und es gilt

$$Jf_U^{-1}(y) = [Jf(x)]^{-1} = [Jf(f_U^{-1}(y))]^{-1} .$$

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 von der Operatornorm induzierten Metrik)([Ü]), ist $y \mapsto Jf_U^{-1}(y)$ als Verknüpfung der stetigen Abbildungen f_U^{-1} sowie $x \mapsto Jf(x)$ und $A \mapsto A^{-1}$ stetig auf V. Dies ist äquivalent dazu, dass f_U^{-1} in $C^1(V, \mathbb{R}^d)$ liegt. □

Beispiel 20.8 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

$$\det Jf(r,\varphi) = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r > 0 .$$

Also existieren nach S. 20.7 zu jedem $(r_0, \varphi_0) \in \Omega$ eine offene Umgebung U von (r_0, φ_0) und eine offene Umgebung V von $f(r_0, \varphi_0)$ so, dass $f_{|U} : U \to V$ bijektiv ist. Außerdem ist $f^{-1} = (f_{|U})^{-1} \in C^1(V, \mathbb{R}^2)$.
Da $f(1, 2k\pi) = (1, 0)$ für alle $k \in \mathbb{Z}$ gilt, ist f jedoch 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

$$U = U_\alpha = \{(r,\varphi) \in \Omega : \alpha - \pi < \varphi < \alpha + \pi\}$$

für beliebiges $\alpha \in \mathbb{R}$. Auf jeder echten offenen Obermenge von U_α ist f nicht mehr umkehrbar. Außerdem gilt $f(U_\alpha) = \mathbb{R}^2 \setminus \{(r\cos(\alpha+\pi), r\sin(\alpha+\pi)) : r > 0\}$ für alle $\alpha \in \mathbb{R}$. Wie sieht eine (nicht **die**) Umkehrfunktion aus? Betrachten wir

$$U = U_0 = \{(r,\varphi) : r > 0\ ,\ \varphi \in (-\pi,\pi)\}\ .$$

Ist $(x,y) \in f(U_0)$ 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 $\varphi = \operatorname{sign}(y)\pi/2$. Ist $x > 0$, so folgt

$$\frac{y}{x} = \tan\varphi$$

also $\varphi = \arctan(y/x)$. Für $x < 0$ ergibt sich $\varphi = \operatorname{sign}(y)\pi + \arctan(y/x)$. Insgesamt erhalten wir

$$f_{U_0}^{-1}(x,y) = (x,y) = \begin{cases} \left(\sqrt{x^2+y^2}, \arctan(y/x)\right) & , \text{ falls } x > 0 \\ (|y|, \operatorname{sign}(y)\pi/2) & , \text{ falls } x = 0 \\ \left(\sqrt{x^2+y^2}, \arctan(y/x) + \operatorname{sign}(y)\pi\right) & , \text{ falls } x < 0 \end{cases}\ .$$

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 W von $(x^{(0)}, y^{(0)})$ sowie eine Funktion $f : U \to \mathbb{R}^m$ so, dass für alle $(x, y) \in W$ gilt

$$F(x, y) = 0 \Leftrightarrow y = f(x) \, .$$

Wir orientieren uns wieder an zwei einfachen Beispielen

Beispiel 20.9 1. Wir betrachten die Gleichung

$$F(x, y) := x^2 + y^2 - 1 = 0 \qquad (x, y \in \mathbb{R}) \, .$$

Dann ist offenbar $(x_0, y_0) = (1/\sqrt{2}, 1/\sqrt{2})$ eine Lösung der Gleichung. Hier gilt etwa für $U := (-1, 1)$ und $W := (-1, 1) \times (0, \infty)$

$$F(x, y) = 0 \Leftrightarrow y = \sqrt{1 - x^2} =: f(x)$$

(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 $W = (-1, 1) \times \mathbb{R}$.
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 20.10 *(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

$$\frac{\partial F}{\partial y}(x, y) := \left(\frac{\partial F_\mu}{\partial y_\nu}(x, y) \right)_{\mu, \nu = 1, \ldots, m} \quad \text{und} \quad \frac{\partial F}{\partial x}(x, y) := \left(\frac{\partial F_\mu}{\partial x_\nu}(x, y) \right)_{\substack{\mu = 1, \ldots, m \\ \nu = 1, \ldots, d}} \, .$$

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 (genau) 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*

$$Jf(x) = -\left[\frac{\partial F}{\partial y}(x, f(x))\right]^{-1} \frac{\partial F}{\partial x}(x, f(x)) \qquad (x \in U)\ .$$

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

$$JG(x, y) = \left[\begin{array}{ccc} E_d & \vdots & 0 \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots\dots} \\ \frac{\partial F}{\partial x}(x, y) & \vdots & \frac{\partial F}{\partial y}(x, y) \end{array}\right]$$

und damit $\det JG(x^{(0)}, y^{(0)}) = \det \frac{\partial F}{\partial y}(x^{(0)}, y^{(0)}) \neq 0$. Nach S. 20.7 existieren offene Umgebungen R von $(x^{(0)}, y^{(0)})$ und S von $(x^{(0)}, 0) = G(x^{(0)}, y^{(0)})$ so, dass $G_{|R} : R \to S$ bijektiv ist mit

$$G_R^{-1} \in C^1(S, \mathbb{R}^d \times \mathbb{R}^m)\ .$$

Da S offen ist, existieren offene Umgebungen U von $x^{(0)}$ und V von 0 mit $U \times V \subset S$. Dann ist $W := G_R^{-1}(U \times V)$ offen in R (nach S. 11.12) also auch in $\mathbb{R}^d \times \mathbb{R}^m$. Wir setzen $G^{-1} := (G_R^{-1})_{|U \times V}$. Aus der Definition von G ergibt sich, dass G^{-1} von der Form

$$G^{-1}(x, z) = (x, H(x, z)) \qquad ((x, z) \in U \times V)$$

mit $H \in C^1(U \times V, \mathbb{R}^m)$ 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 $(x, z) \in U \times V$:

$$F(x, H(x, z)) = \pi_2(G(x, H(x, z))) = \pi_2(x, z) = z\ .$$

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 $G_{|W}$ injektiv ist, $y = f(x)$. Also ergibt sich 1.
2. Weiter folgt aus $H \in C^1(U \times V, \mathbb{R}^m)$ 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
$$0 = J\varphi(x) = JF(x, f(x)) \cdot \begin{pmatrix} E_d \\ Jf(x) \end{pmatrix} = \frac{\partial F}{\partial x}(x, f(x)) + \frac{\partial F}{\partial y}(x, f(x)) Jf(x)$$
und damit
$$Jf(x) = -\left[\frac{\partial F}{\partial y}(x, f(x))\right]^{-1} \frac{\partial F}{\partial x}(x, f(x)) \qquad (x \in U)\ .$$
(Man beachte dabei: Es gilt nach S. 20.7
$$\det \frac{\partial F}{\partial y}(x, y) = \det JG(x, y) \neq 0$$
für alle $(x, y) \in W$.) □

Beispiel 20.11 (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. 20.10 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“. Sogenanntes implizites Differenzieren ergibt für die Funktion $f = f_{(x_0, y_0)}$ aus S. 20.10
$$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))} = -\frac{4x(x^2 + y^2 - 1)}{4y(x^2 + y^2 + 1)}_{|y = f(x)}$$

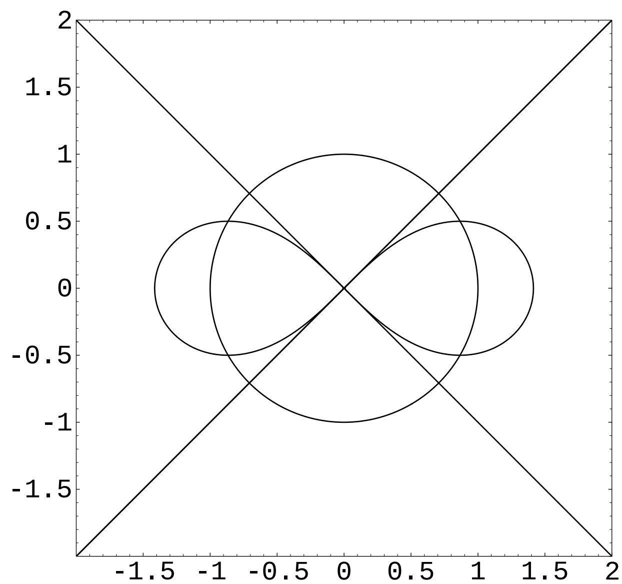
auf einer Umgebung U von x_0. Also hat f Extremstellen höchstens in Punkten x mit

$$x^2 + f^2(x) = x^2 + y^2 = 1 \, .$$

Man rechnet leicht nach, dass $F(x, y) = 0$ genau dann gilt, wenn $|z^2 - 1| = 1$ ist, wobei $z = x + iy$.
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 20.12 Es seien (X, d) ein metrischer Raum, und es seien $f : X \to \mathbb{R}$ sowie $g : X \to \mathbb{R}^m$. Ist

$$L := \{x \in X : g(x) = 0\}$$

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 20.13 *(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

$$\operatorname{Rang}\big(Jg(x^{(0)})\big) = m$$

ist. Dann gilt: Hat f an $x^{(0)}$ ein lokales Maximum oder Minimum unter der Nebenbedingung $g(x) = 0$, so existiert ein $\lambda \in \mathbb{R}^m$ mit

$$\operatorname{grad} f(x^{(0)}) = \sum_{j=1}^{m} \lambda_j \operatorname{grad} g_j(x^{(0)}) = (Jg)^T(x^{(0)})\lambda$$

(die Komponenten $\lambda_1, \dots, \lambda_m$ von λ heißen Lagrange-Multiplikatoren).

Beweis. Wir schreiben zur Abkürzung $x = (y, z)$ mit

$$y = (x_1, \dots, x_m)\,,\ z = (x_{m+1}, \dots, x_d)$$

für $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ und insbesondere

$$(y^{(0)}, z^{(0)}) = x^{(0)}\ .$$

O. E. gelte

$$\det\left(\frac{\partial g}{\partial y}(x^{(0)})\right) = \det\left(\frac{\partial g_\mu}{\partial x_\nu}(x^{(0)})\right)_{\mu,\nu=1,\dots,m} \neq 0\ ,$$

d. h. $\frac{\partial g}{\partial y}(x^{(0)})$ habe vollen Rang m. Nach dem Hauptsatz über implizite Funktionen ist die Gleichung $g(x) = g(y, z) = 0$ an $x^{(0)}$ „lokal nach y auflösbar“. Insbesondere existieren eine Umgebung U von $z^{(0)} = (x^{(0)}_{m+1}, \dots, x^{(0)}_d)$ sowie eine Funktion $\varphi \in C^1(U, \mathbb{R}^m)$ so, dass $\varphi(z^{(0)}) = y^{(0)}$ und

$$g(\varphi(z), z) \equiv 0 \qquad (z \in U)$$

gilt. Hieraus folgt

$$0 = Jg(x^{(0)}) \begin{pmatrix} J\varphi(z^{(0)}) \\ E_{d-m} \end{pmatrix} = \frac{\partial g}{\partial y}(x^{(0)}) J\varphi(z^{(0)}) + \frac{\partial g}{\partial z}(x^{(0)})$$

Weiter definieren wir $F \in C^1(U)$ durch

$$F(z) := f(\varphi(z), z) \quad (z \in U) .$$

Nach Voraussetzung ist $z^{(0)}$ eine Extremstelle von F (ohne Nebenbedingung). Also gilt nach S. 19.11

$$\operatorname{grad} F(z^{(0)}) = 0$$

und mit der Kettenregel daher

$$0 = \operatorname{grad}^T f(x^{(0)}) \cdot \begin{pmatrix} J\varphi(z^{(0)}) \\ E_{d-m} \end{pmatrix} = \left(\frac{\partial f}{\partial y}\right)^T (x^{(0)}) \cdot J\varphi(z^{(0)}) + \left(\frac{\partial f}{\partial z}\right)^T (x^{(0)}) .$$

Schließlich hat das lineare Gleichungssystem

$$\left(\frac{\partial g}{\partial y}\right)^T (x^{(0)}) \cdot \lambda = \frac{\partial f}{\partial y}(x^{(0)})$$

genau eine Lösung $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$. Hieraus ergibt sich wiederum

$$\left(\frac{\partial f}{\partial z}\right)^T (x^{(0)}) = -\left(\frac{\partial f}{\partial y}\right)^T (x^{(0)}) \cdot J\varphi(z^{(0)}) = -\lambda^T \frac{\partial g}{\partial y}(x^{(0)}) \cdot J\varphi(z^{(0)}) = \lambda^T \frac{\partial g}{\partial z}(x^{(0)})$$

und damit insgesamt $\operatorname{grad} f(x^{(0)}) = (Jg)^T(x^{(0)}) \cdot \lambda$. □

Bemerkung 20.14 Ähnlich wie S. 19.11 liefert S. 20.13 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

$$g_j(x_1, \ldots, x_d) = 0 \qquad (j = 1, \ldots, m)$$

und

$$\frac{\partial f}{\partial x_k}(x_1, \ldots, x_d) = \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_k}(x_1, \ldots, x_d) \qquad (k = 1, \ldots, d)$$

zu lösen (also $(d + m)$ Gleichungen für die $(d + m)$ Unbekannten $x_1, \ldots, x_d$ und $\lambda_1, \ldots, \lambda_m$).

Beispiel 20.15 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. 20.13 ist eine notwendige Bedingung gegeben durch

$$\operatorname{grad} K(x,y) = \lambda \operatorname{grad} P(x,y) ,$$

d. h. wir haben die 3 Gleichungen

$$\begin{array}{rcl} x^\alpha y^{1-\alpha} & = & c \\ p & = & \lambda \alpha x^{\alpha-1} y^{1-\alpha} \\ q & = & \lambda(1-\alpha) x^\alpha y^{-\alpha} \end{array}$$

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.

A Von den natürlichen zu den reellen Zahlen

Bemerkung und Definition A.1 Die Menge $\mathbb{N}$ der natürlichen Zahlen kann axiomatisch beschrieben werden durch die folgenden Bedingungen (Peano-Axiome):

(N1) $\mathbb{N}$ enthält ein Element, genannt 1.

(N2) (Nachfolgerfunktion) Es gibt eine injektive Abbildung $\varphi : \mathbb{N} \to \mathbb{N}$ mit $1 \notin W(\varphi)$.

(N3) (Prinzip der vollständigen Induktion) Ist $M \subset \mathbb{N}$ mit $1 \in M$ und so, dass $\varphi(n) \in M$ für alle $n \in M$, so ist $M = \mathbb{N}$.

Damit kann man zeigen: Es existieren eindeutig bestimmte Abbildungen $+$ und $\cdot$: $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ so, dass

$$n + 1 = \varphi(n), \qquad n + \varphi(m) = \varphi(n + m)$$

und

$$\begin{aligned} n \cdot 1 &= n, \\ n \cdot \varphi(m) &= n \cdot m + n. \end{aligned}$$

Außerdem sind damit die üblichen Relationen $<$ und $\leq$ gegeben durch

$$n < m :\Leftrightarrow \exists\, k \in \mathbb{N} : m = n + k$$

und

$$n \leq m :\Leftrightarrow n < m \text{ oder } n = m.$$

Schließlich gilt das Wohlordnungsprinzip: Jede Menge $\emptyset \neq M \subset \mathbb{N}$ hat ein Minimum, d.h. es existiert ein $m \in M$ mit $m \leq n$ für alle $n \in M$.
Damit kann man zeigen, dass zu jedem Paar $(n, a) \in \mathbb{N} \times \mathbb{N}$ ein $m \in \mathbb{N}_0$ und ein $r \in \{0, 1, \ldots, a - 1\}$ existieren mit

$$n = ma + r$$

(Division mit Rest). Hier hat man im Übrigen schonmal stillschweigend $\mathbb{N}$ um die Null zu $\mathbb{N}_0$ (mit passenden Rechenregeln) erweitert.
Definiert man Potenzen in $\mathbb{N}_0$ wie in B./D. 2.8, so gilt folgende wichtige Aussage über die **Darstellung** natürlicher Zahlen:

Es sei $q \in \mathbb{N}, q \geq 2$ fest. Dann existeren für alle $n \in \mathbb{N}$ ein $d \in \mathbb{N}_0$ und $a_0, \ldots, a_d \in \{0, \ldots, q-1\}$ mit $a_d \neq 0$ und so, dass

$$n = \sum_{j=0}^{d} a_j q^j .$$

(Denn: $n = 1$: $d = 0$ und $a_0 = 1$ ist geeignet.
$n - 1 \to n$: Wir wählen $d \in \mathbb{N}_0$ so, dass $q^d \leq n < q^{d+1}$. Division mit Rest ergibt

$$n = mq^d + n'$$

mit $0 < m < q$ und $0 \leq n' < q^d$, also insbesondere $n' < n$.

1. Fall: $n' = 0$. Dann setzen wir $a_d := m$ und $a_0 := \cdots := a_{d-1} := 0$.
2. Fall: $n' \neq 0$.
Nach Induktionsvoraussetzung (Behauptung gilt für alle $n' < n$) existieren ein $d' \in \mathbb{N}_0$ und $a'_0, \ldots, a'_{d'} \in \{0, 1, \ldots, q-1\}$ so, dass $a'_{d'} \neq 0$ und

$$n' = \sum_{j=0}^{d'} a'_j q^j .$$

Aus $n' < q^d$ ergibt sich $d' < d$. Hiermit ist

$$n = mq^d + n' = \sum_{j=0}^{d} a_j q^j \; ,$$

wobei $a_j := a'_j$ für $j \in \{0, \ldots, d'\}$, $a_d := m$ und $a_j = 0$ für $j \in \{d'+1, \ldots, d-1\}$. Also hat man die gewünschte Darstellung für n.)

Damit heißt

$$(a_d a_{d-1} \ldots a_0)_q$$

die *q-adische Darstellung* von n (man kann zeigen, dass die Darstellung eindeutig ist; [Ü]). Im Falle $q = 10$ (besser: $q = \mathrm{X}$) spricht man auch von der Dezimal-, im Falle $q = 2$ von der Binär- und im Falle $q = 16$ (oder besser $q = \mathrm{XVI}$) von der Hexadezimaldarstellung. Schließlich schreibt man im Dezimalfall auch kurz $a_r \ldots a_0$ statt $(a_r \ldots a_0)_X$.
So ist etwa für $n = \mathrm{XXIII}$

$$n = 1 \cdot 2^4 + 0 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = (10111)_2$$

und

$$n = 2 \cdot \mathrm{X}^1 + 3 \cdot \mathrm{X}^0 = (23)_\mathrm{X} = 23 \, .$$

Definition A.2 Eine Relation $\sim$ in X heißt *Äquivalenzrelation* (auf X), falls für alle $x, y, z \in X$ gilt

a) $x \sim x$ (Reflexivität),

b) aus $x \sim y$ folgt $y \sim x$ (Symmetrie),

c) aus $x \sim y$ und $y \sim z$ folgt $x \sim z$ (Transitivität).

2. Ist $\sim$ eine Äquivalenzrelation, so heißt $[x] := \{x' \in X : x \sim x'\}$ *die von x erzeugte Äquivalenzklasse*. Außerdem heißt $X/_\sim := \big\{[x] : x \in X\big\}$ der *Quotient* von X *modulo* $\sim$.

Beispiel A.3 1. Es sei $X = \mathbb{N}_0 \times \mathbb{N}_0$. Wir definieren

$$(a, b) \sim (c, d) \quad :\Leftrightarrow \quad a + d = b + c.$$

Dann ist $\sim$ eine Äquivalenzrelation auf X (ergibt sich aus Rechenregeln für die Addition in $\mathbb{N}_0$).
Formal ist damit die Menge $\mathbb{Z}$ der ganzen Zahlen definiert als

$$\mathbb{Z} := X/_\sim = (\mathbb{N}_0 \times \mathbb{N}_0)/_\sim.$$

Man schreibt dann $a - b$ statt $[(a, b)]$ und im Falle $b = 0$ kurz a (so findet man $\mathbb{N}_0$ in $\mathbb{Z}$ wieder). Außerdem schreibt man im Falle $a = 0$ kurz $-b$. Die Rechenoperationen $+$ und $\cdot$ sowie die Relation $<$ lassen sich auf $\mathbb{Z}$ übertragen (konkret durch

$$(a - b) + (c - d) := (a + c) - (b + d) \quad \text{und} \quad (a - b) \cdot (c - d) := (ac + bc) - (ad + bc)$$

sowie

$$a - b < c - d :\Leftrightarrow a + d < b + c \text{).}$$

2. Es sei $X = \mathbb{Z} \times \big(\mathbb{Z} \setminus \{0\}\big)$. Wir definieren

$$(a, b) \sim (c, d) \quad :\Leftrightarrow \quad ad = bc.$$

Dann ist $\sim$ eine Äquivalenzrelation auf X (ergibt sich aus Rechenregeln für die Multiplikation in $\mathbb{Z}$).
Formal ist damit die Menge $\mathbb{Q}$ der rationalen Zahlen definiert als

$$\mathbb{Q} := \big(\mathbb{Z} \times (Z \setminus \{0\})\big)/_\sim \,.$$

Für die rationale Zahl $\big[(a,b)\big]$ schreibt man a/b. Wieder lassen sich die Rechenoperationen $+$ und $\cdot$ sowie die Relation $<$ auf $\mathbb{Q}$ übertragen, hier etwa für $+$ und $\cdot$ durch

$$\frac{a}{b}+\frac{c}{d}:=\frac{ad+cb}{bd} \quad \text{und} \quad \frac{a}{b}\cdot\frac{c}{d}:=\frac{ac}{bd}\,.$$

Außerdem ergibt sich $\mathbb{Z}\subset\mathbb{Q}$, indem man $a\in\mathbb{Z}$ mit $a/1$ identifiziert.

Bemerkung und Definition A.4 Es sei $q\in\mathbb{N}$, $q\geq 2$. Eine Folge

$$x=(x_j)_{j\in\mathbb{Z}}=(\ldots x_2,x_1,x_0,x_{-1},x_{-2},\ldots)$$

mit $x_j\in\{0,1,\ldots,q-1\}$ und der Eigenschaft, dass ein $d=d(x)\in\mathbb{Z}$ existiert mit

$$x_d\neq 0 \qquad \text{und} \qquad x_j=0 \qquad (j>d)$$

heißt eine *q-adische Entwicklung*. Im Falle $q=2$ spricht man von *Binärentwicklung* und im Falle $q=X$ von *Dezimalentwicklung*. Wir schreiben dann kurz

$$x=\begin{cases} x_dx_{d-1}\ldots x_0,x_{-1}\ldots, & \text{falls } d\geq 0\\ 0,0\ldots 0x_dx_{d-1}\ldots, & \text{falls } d<0\end{cases}.$$

Im Weiteren beschränken wir uns auf den Fall $q=2$ und setzen

$$\mathbb{R}_+:=\{x: x \text{ Binärentwicklung}\},$$

wobei wir die beiden Binärentwicklungen

$$x=(x_j)_{j\in\mathbb{Z}} \quad \text{und} \quad y=(y_j)_{j\in\mathbb{Z}}$$

im Falle, dass für ein $m\in\mathbb{Z}$

$$x_j=y_j\ (j>m), \quad x_m=1,\ y_m=0 \quad \text{und} \quad x_j=0,\ y_j=1\ (j<m)$$

gilt, identifizieren (wir werden sehen, dass dies sinnvoll bzw. sogar zwingend ist). Im Zweifelsfall wählen wir die erste Darstellung, wenn wir eine Eindeutigkeit erzwingen möchten.
Weiter heißt eine Binärentwicklung *abbrechend*, falls ein $m\in\mathbb{Z}$ existiert mit $x_j=0$ $(j<m)$, d. h.

$$x=x_d\ldots x_m000\ldots \quad \text{bzw.} \quad x=0,0\ldots 0x_d\ldots x_m000\ldots\,.$$

Für abbrechende Binärentwicklungen identifizieren wir x und

$$\sum_{j=m}^{d} x_j 2^j \in \mathbb{Q}_+ \,.$$

Dann ist x abbrechend genau dann, wenn

$$2^m x \in \mathbb{N}$$

für ein $m \in \mathbb{N}$. Hieraus folgt sofort, dass für abbrechende x und y auch $x+y$ und $x \cdot y$ (in $\mathbb{Q}$ definiert!) abbrechend sind ([Ü]). Man beachte, dass $1/x$ im Allgemeinen nicht abbrechend ist.

Wir definieren nun für $x, y \in \mathbb{R}_+$ mit $x \neq y$

$$x < y,$$

falls entweder

$$d(x) < d(y)$$

oder

$$d(x) = d(y) \quad \text{und} \quad x_m = 0, y_m = 1 \text{ für } m := \max\{j : x_j \neq y_j\}.$$

Damit kann man zeigen, dass $(\mathbb{R}_+, <)$ geordnet ist. Der entscheidende Vorteil gegenüber $\mathbb{Q}_+$ liegt darin, dass

$$(\mathbb{R}_+, <) \text{ vollständig}$$

ist.

Wir wollen den Beweis andeuten, indem wir eine ,,Konstruktionsvorschrift" für $\sup M$ für nach oben beschränkte und nichtleere M angeben.

Dazu ist es hilfreich, *Abschneidungen* zu definieren. Ist $x = (x_j)_{j \in \mathbb{Z}}$ eine Binärentwicklung, so setzen wir für $m \in \mathbb{Z}$

$$[x]_m := \sum_{j=m}^{d} x_j 2^j \in \mathbb{Q}_+ \cup \{0\}$$

mit $[x]_m := 0 \in \mathbb{Q}$ falls $m > d$. .

Es sei also $M \subset \mathbb{R}_+$ nach oben beschränkt (und nichtleer). Wir definieren ein $\xi \in \mathbb{R}_+$ rekursiv:

Zunächst sei

$$d := \max\{m \in \mathbb{Z} : \exists\, x \in M : x \geq 2^m\}.$$

Damit setzen wir $\xi_d := 1$ und $\xi_j := 0$ $(j > d)$. Weiter definieren wir

$$\xi_{d-1} := \begin{cases} 1, & \text{falls } [x]_{d-1} = 2^d + 2^{d-1} \text{ für ein } x \in M \\ 0, & \text{sonst} \end{cases}$$

und entsprechend für $n \in \mathbb{N}$

$$\xi_{d-n-1} := \begin{cases} 1, & \text{falls } [x]_{d-n-1} = 2^d + \xi_{d-1}2^{d-1} + \ldots + \xi_{d-n}2^{d-n} + 2^{d-n-1} \text{ für ein } x \in M \\ 0, & \text{sonst.} \end{cases}$$

Hierdurch ist induktiv ein $\xi = (\xi_j)_{j\in\mathbb{Z}} \in \mathbb{R}_+$ definiert. Es gilt dabei ([Ü])

$$\xi = \sup M.$$

Damit definieren wir für $x, y \in \mathbb{R}_+ \cup \{0\}$ (mit $[0]_m := 0$; man identifiziert also 0 mit $(0)_{j\in\mathbb{Z}}$)

$$\begin{aligned} x + y &:= \sup\big\{[x]_m + [y]_m : m \in \mathbb{Z}\big\}, \\ x \cdot y &:= \sup\big\{[x]_m \cdot [y]_m : m \in \mathbb{Z}\big\}. \end{aligned}$$

Schließlich definieren wir wie bei der Erweiterung von $\mathbb{N}_0$ zu $\mathbb{Z}$ für $(a, b), (c, d) \in (\mathbb{R}_+ \cup \{0\}) \times (\mathbb{R}_+ \cup \{0\})$

$$(a, b) \sim (c, d) \quad :\Leftrightarrow \quad a + d = b + c.$$

Dann ist $\sim$ eine Äquivalenzrelation und

$$\mathbb{R} := \big((\mathbb{R}_+ \cup \{0\}) \times (\mathbb{R}_+ \cup \{0\})\big)/_\sim .$$

Man schreibt wieder $a - b$ statt $[(a, b)]$ und im Falle $b = 0$ kurz a sowie im Falle $a = 0$ kurz $-b$. Die Rechenoperationen $+$ und $\cdot$ sowie die Relation $<$ lassen sich auf $\mathbb{R}$ übertragen und zwar so, dass damit $(R, +, \cdot, <)$ ein vollständiger geordneter Körper wird.

Statt über Binärentwicklungen kann man die (positiven) reellen Zahlen auch über andere q-adische Entwicklungen einführen, also insbesondere über Dezimalentwicklungen. Dabei ist zu zeigen, dass am Ende derselbe geordnete Körper entsteht (also unabhängig von der Wahl von q).

B Mächtigkeit von Mengen

Wir werden im Folgenden sehen, dass in gewisser Weise „sehr viele“ reelle Zahlen irrational sind. Dazu beschäftigen wir uns zunächst mit dem Begriff der Mächtigkeit einer Menge. Zunächst präzisieren wir, dass wir unter einer endlichen Mengen eine Menge verstehen, die bijektiv auf eine der Mengen $\{1, \dots, n\}$ abgebildet werden kann. Außerdem sehen wir auch die leere Menge als endlich an.
Sind M_1, M_2 endliche Mengen, so haben M_1 und M_2 gleich viele Elemente (also $|M_1| = |M_2|$) genau dann, wenn eine bijektive Abbildung $\varphi : M_1 \to M_2$ existiert. Wir werden (G. Cantor folgend) diesen Gedanken nun auf unendliche Mengen übertragen.

Definition B.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 $\varphi : M_1 \to M_2$ existiert.

2. M heißt *abzählbar unendlich* falls M gleichmächtig zu $\mathbb{N}$ ist.

3. M heißt *abzählbar* falls M endlich oder abzählbar unendlich ist. Anderenfalls heißt M *überabzählbar*.

Bemerkung B.2 1. Aus D. B.1 ergibt sich sofort, dass eine Menge M genau dann abzählbar unendlich ist, wenn eine Folge $(x_n)_n$ in M existiert mit $x_n \neq x_m$ für $n \neq m$ und $M = \{x_n : n \in \mathbb{N}\}$.
Weiter folgt aus D. B.1 auch, dass $M \neq \emptyset$ genau dann abzählbar ist, wenn eine Folge $(x_n)_n$ in M existiert mit $M = \{x_n : n \in \mathbb{N}\}$. Eine solche Darstellung nennt man auch eine *Abzählung* von M.
(Denn: „$\Rightarrow$ “: Ist M abzählbar unendlich, so existiert eine solche Folge nach Definition. Ist M endlich, etwa $M = \{x_1, \dots, x_n\}$, so setze man $x_j := x_n$ für $j > n$.
„$\Leftarrow$ “: Ist M endlich, so sind wir fertig. Es sei also M unendlich. Wir definieren $n_1 := 1$. Sind $n_1, \dots, n_k$ bereits definiert, so setzen wir $n_{k+1} := \min\{n : x_n \neq x_{n_1}, \dots, x_{n_k}\}$. (Man beachte: Nach dem Wohlordnungssatz (siehe (Ü)) hat jede nichtleere Teilmenge von $\mathbb{N}$ ein Minimum.) Dann gilt nach Konstruktion $M = \{x_{n_k} : k \in \mathbb{N}\}$ und $x_{n_j} \neq x_{n_k}$ für $j \neq k$.)
2. Man sieht mit einer ähnlichen Überlegung wie in 1. auch, dass jede unendliche Menge eine abzählbar unendliche Teilmenge besitzt.

Satz B.3 *1. Jede Teilmenge B einer abzählbaren Menge A ist wieder abzählbar.*

2. *Es sei $I \neq \emptyset$ eine abzählbare Menge, und es seien A_n $(n \in I)$ abzählbare Mengen. Dann ist auch $\bigcup_{n\in I} A_n$ abzählbar.*

Beweis. 1. O. E. sei $B \neq \emptyset$. Nach B. B.2 existiert eine Folge (x_n) mit $A = \{x_n : n \in \mathbb{N}\}$. Ist n_0 so, dass $x_{n_0} \in B$, so definieren wir

$$y_n := \begin{cases} x_n, & \text{falls } x_n \in B \\ x_{n_0}, & \text{falls } x_n \notin B \end{cases} .$$

Dann ist $B = \{y_n : n \in \mathbb{N}\}$, also B abzählbar nach B. B.2.
2. Ohne Einschränkung können wir $A_n \neq \emptyset$ für alle $n \in I$ annehmen. Zudem koennen wir uns auch auf den Fall $I = \mathbb{N}$ beschränken. (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

$$A_1 = \left\{x_k^{(1)} : k \in \mathbb{N}\right\} \ , \quad A_2 = \left\{x_k^{(2)} : k \in \mathbb{N}\right\}, \ldots$$

Dann ist

$$\bigcup_{n\in\mathbb{N}} A_n = \left\{x_k^{(n)} : k, n \in \mathbb{N}\right\} \ .$$

Wir betrachten folgende Anordnung der Elemente $x_k^{(n)}; n \in \mathbb{N}, k \in \mathbb{N}$:

$$\begin{array}{ccccccccccc}
x_1^{(1)} & & x_2^{(1)} & & x_3^{(1)} & & x_4^{(1)} & & x_5^{(1)} & & x_6^{(1)} \ \ldots \\
& \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & \\
x_1^{(2)} & & x_2^{(2)} & & x_3^{(2)} & & x_4^{(2)} & & x_5^{(2)} & \ldots & \ \ \ \ldots \\
& \swarrow & & \swarrow & & \swarrow & & \swarrow & & & \\
x_1^{(3)} & & x_2^{(3)} & & x_3^{(3)} & & x_4^{(3)} & \ldots & & \ldots & \ \ \ \ldots \\
& \swarrow & & \swarrow & & \swarrow & & & & & \\
x_1^{(4)} & & x_2^{(4)} & & x_3^{(4)} & \ldots & & \ldots & & \ldots & \ \ \ \ldots \\
& \swarrow & & \swarrow & & & & & & & \\
x_1^{(5)} & & x_2^{(5)} & \ldots & & \ldots & & \ldots & & \ldots & \ \ \ \ldots \\
& \swarrow & & & & & & & & & \\
x_1^{(6)} & \ldots & & \ldots & & \ldots & & \ldots & & \ldots & \ \ \ \ldots
\end{array}$$

Hierbei treten alle $x_k^{(n)}$ auf. Diese können durch sukzessives Aneinanderreihen der Diagonalen zu einer Folge $(y_n)_{n\in\mathbb{N}}$ angeordnet werden; genauer ist $(y_n)_{n\in\mathbb{N}}$ für $m \in \mathbb{N}$ und $j = 1, \ldots, m$ definiert durch

$$y_{\frac{m(m-1)}{2}+j} := x_{m+1-j}^{(j)}.$$

(Wer's ganz genau wissen möchte beachte dabei: Für $M := \bigcup_{m\in\mathbb{N}} \big(\{m\} \times \{1, \dots, m\}\big)$ sind $\varphi : M \to \mathbb{N}$ bzw. $\psi : M \to \mathbb{N} \times \mathbb{N}$ mit

$$\varphi(m, j) := \frac{m(m-1)}{2} + j \quad \text{bzw.} \quad \psi(m, j) = (m + 1 - j, j)$$

bijektiv.) □

Insbesondere ergibt sich aus S. B.3

Satz B.4 *Die Menge $\mathbb{Q}$ der rationalen Zahlen ist abzählbar.*

Beweis. Zunächst sieht man leicht, dass $\mathbb{Z}$ abzählbar ist (scheibe $\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$). Damit sind nach S. B.3 auch $\mathbb{Z} \setminus \{0\}$ und

$$\mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) = \bigcup_{m\in\mathbb{Z}\setminus\{0\}} (\mathbb{Z} \times \{m\})$$

abzählbar. Ist $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) = \{x_j : j \in \mathbb{N}\}$ so ist

$$\mathbb{Q} = \Big(\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})\Big)/_\sim = \{[x_j] : j \in \mathbb{N}\} ,$$

also $\mathbb{Q}$ abzählbar. □

Wir wollen nun zeigen, dass jedes (nichttriviale) Intervall in $\mathbb{R}$ überabzählbar ist. Daraus ergibt sich auch unmittelbar die Überabzählbarkeit von $\mathbb{R}$ oder $\mathbb{C}$. Vorbereitend zeigen wir:

Satz B.5 *(Intervallschachtelungsprinzip)*

1. *Ist I_n eine Folge von Intervallen der Form $I_n = [a_n, b_n]$, mit $I_{n+1} \subset I_n$ $(n \in \mathbb{N})$, so existieren $a, b \in \mathbb{R}$ mit $a \le b$ und*

$$\bigcap_{n\in\mathbb{N}} I_n = [a, b] .$$

2. *Gilt zusätzlich $b_n - a_n \to 0\,(n \to \infty)$, so ist $a = b$, d. h. $\bigcap_{n\in\mathbb{N}} I_n$ ist einpunktig.*

Beweis. 1. Nach Voraussetzung ist $(a_n) \uparrow$ und $(b_n) \downarrow$. Außerdem gilt $a_1 \leq a_n \leq b_n \leq b_1$. Also gilt nach dem Hauptsatz über monotone Folgen

$$a_n \to \sup\{a_k : k \in \mathbb{N}\} =: a \quad \text{und} \quad b_n \to \inf\{b_k : k \in \mathbb{N}\} =: b \, .$$

Aus $a_n \leq b_n$ folgt $a \leq b$, also insgesamt

$$a_n \leq a \leq b \leq b_n \qquad (n \in \mathbb{N}) \, .$$

Folglich ist

$$[a, b] \subset \bigcap_{n \in \mathbb{N}} I_n \, .$$

Andererseits folgt für $x \in \bigcap\limits_{n \in \mathbb{N}} I_n$ aus $a_n \leq x \leq b_n$ für alle n auch $a \leq x \leq b$, also $x \in [a, b]$.

2. Aus $0 \leq b - a \leq b_n - a_n \to 0$ $(n \to \infty)$ folgt $b - a = 0$. □

Satz B.6 *Jedes Intervall $I \subset \mathbb{R}$, das mehr als einem Punkt enthält, ist überabzählbar.*

Beweis. Es reicht ([Ü]), das Intervall $[0, 1]$ zu 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 $I_n = [a_n, b_n]$ von Intervallen 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 $\bigcap\limits_{n \in \mathbb{N}} I_n = \{x\}$ für ein $x \in [0, 1]$. Ist $k \in \mathbb{N}$ so gilt nach Konstruktion $x_k \notin \bigcap\limits_{n \in \mathbb{N}} I_n$, d. h. $x_k \neq x$ für alle $k \in \mathbb{N}$. Widerspruch! □

Bemerkung B.7 Allgemein gilt: Ist M überabzählbar und ist A abzählbar, so ist auch $M \setminus A$ überabzählbar (denn sonst wäre nach S. B.3 auch $M = (M \cap A) \cup (M \setminus A)$ abzählbar). Also ist insbesondere nach S. B.4 und S. B.6 die Menge der irrationalen Zahlen in jedem Intervall I (mit $|I| > 1$) überabzählbar.

Bemerkung B.8 Da mit $\mathbb{Q}$ auch $\mathbb{Q} \cap (0,1)$ abzählbar ist, können wir $\mathbb{Q} \cap (0,1) = \{r_n : n \in \mathbb{N}\}$ schreiben. Für die (beschränkte) Folge (r_n) gilt dann:
Zu jedem $x \in [0,1]$ exstiert eine Teilfolge (r_{n_k}) mit

$$r_{n_k} \to x \quad (k \to \infty) \, ,$$

d.h. jede der überabzählbar vielen reellen Zahlen $x \in [0,1]$ ist Grenzwert einer Teilfolge von (r_n).
(Denn: Da jedes Intervall (a,b) mit $a < b$ eine rationale Zahl enthält, ist $(a,b) \cap \mathbb{Q}$ sogar unendlich (warum?).
Es sei $x \in [0,1]$. Dann ist

$$I_k := (x - 1/k, x + 1/k) \cap (0,1) \cap \mathbb{Q}$$

für jedes $k \in \mathbb{N}$ unendlich. Also existieren unendlich viele $n \in \mathbb{N}$ mit $r_n \in I_k$. Damit definieren wir (n_k) induktiv: Wir setzen $n_1 := 1$. Sind $n_1, \ldots, n_k$ bereits definiert, so wählen wir $n_{k+1} > n_k$ so, dass

$$r_{n_{k+1}} \in I_k \, .$$

Dann gilt nach Definition $r_{n_k} \to x$ $(k \to \infty)$.)

C Hölder- und Minkowski-Ungleichung

Satz C.1 *(Hölder-Ungleichung)*
Es seien $p, q > 1$ *mit* $\frac{1}{p} + \frac{1}{q} = 1$. *Dann gilt für* $u_1, \ldots, u_m, v_1, \ldots, v_m \geq 0$

$$\sum_{j=1}^{m} u_j v_j \leq \Big(\sum_{j=1}^{m} u_j^p\Big)^{1/p} \Big(\sum_{j=1}^{m} v_j^q\Big)^{1/q}.$$

Beweis. 1. Wir zeigen zunächst: Sind $u, v \geq 0$, so gilt

$$uv \leq \frac{1}{p}\, u^p + \frac{1}{q}\, v^q.$$

Denn: O. E. seien $u, v > 0$. Aus S. 9.2.1 folgt

$$\ln t \leq t - 1 \qquad (t > 0).$$

Hieraus ergibt sich für $A := \frac{1}{p}\, u^p + \frac{1}{q}\, v^q$

$$\begin{aligned} \ln(uv) &= \frac{1}{p}\,\ln(u^p) + \frac{1}{q}\,\ln(v^q) \\ &\leq \frac{1}{p} \cdot \ln\Big(\frac{u^p}{A}\Big) + \frac{1}{q} \cdot \ln\Big(\frac{v^q}{A}\Big) + \underbrace{\Big(\frac{1}{p} + \frac{1}{q}\Big)}_{=1} \ln A \\ &\leq \frac{1}{p}\Big(\frac{u^p}{A} - 1\Big) + \frac{1}{q}\Big(\frac{v^q}{A} - 1\Big) + \ln A = \ln A, \end{aligned}$$

also auch $uv \leq A$.
2. Wir setzen $U := \Big(\sum\limits_{j=1}^{m} u_j^p\Big)^{1/p}, V := \Big(\sum\limits_{j=1}^{m} v_j^q\Big)^{1/p}$.
Ohne Einschränkung können wir $U, V > 0$ annehmen. Aus 1. folgt

$$\begin{aligned} \frac{1}{U \cdot V} \sum_{j=1}^{m} u_j v_j &= \sum_{j=1}^{m} \frac{u_j}{U}\,\frac{v_j}{V} \leq \sum_{j=1}^{m} \Big(\frac{1}{p}\,\Big(\frac{u_j}{U}\Big)^p + \frac{1}{q}\,\Big(\frac{v_j}{V}\Big)^q\Big) \\ &= \frac{1}{pU^p} \sum_{j=1}^{m} u_j^p + \frac{1}{qV^q} \sum_{j=1}^{m} v_j^q = \frac{1}{p} + \frac{1}{q} = 1. \end{aligned}$$

□

Satz C.2 *(Minkowski-Ungleichung) Für alle $p > 1$ und alle $x, y \in \mathbb{K}^m$ gilt*

$$\left(\sum_{j=1}^{m} |x_j + y_j|^p\right)^{1/p} \le \left(\sum_{j=1}^{m} |x_j|^p\right)^{1/p} + \left(\sum_{j=1}^{m} |y_j|^p\right)^{1/p} .$$

Beweis. Wir setzen $S := \sum\limits_{j=1}^{m} |x_j + y_j|^p$. Mit S. C.1 gilt für $q := p/(p-1)$ (und damit q so, dass $\frac{1}{p} + \frac{1}{q} = 1$)

$$\begin{aligned} S &\le \sum_{j=1}^{m} \big(|x_j| + |y_j|\big)|x_j + y_j|^{p-1} = \sum_{j=1}^{m} |x_j| \cdot |x_j + y_j|^{p-1} + \sum_{j=1}^{m} |y_j| \cdot |x_j + y_j|^{p-1} \\ &\le \left[\Big(\sum_{j=1}^{m} |x_j|^p\Big)^{1/p} + \Big(\sum_{j=1}^{m} |y_j|^p\Big)^{1/p}\right] \Big(\underbrace{\sum_{j=1}^{m} |x_j + y_j|^{(p-1)q}}_{=S}\Big)^{1/q} \end{aligned}$$

Nach Division durch $S^{1/q}$ (ohne Einschränkung $S > 0$) ergibt sich die Behauptung (man beachte $S^{1/p} = S^{1-1/q}$). □

D Stetigkeitseigenschaften monotoner Funktionen

Es ist klar, dass i. A. monotone Funktionen nicht überall stetig sind (etwa sign). Tatsächlich können unendlich viele Sprungstellen auftreten (auch auf endlichen Intervallen), wie folgendes Beispiel zeigt.

Beispiel D.1 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.

Der folgende Satz zeigt, dass monotone Funktionen an vielen Stellen stetig sind, und dass als Unstetigkeitsstellen nur Sprungstellen in Frage kommen.

Satz D.2 *Es sei $I \neq \emptyset$ ein offenes Intervall, und es sei $f : I \to \mathbb{R}$ monoton (wachsend oder fallend). Dann ist f stetig bis auf höchstens abzählbar viele Sprungstellen. Außerdem gilt für alle $x_0 \in I$*

$$f(x_0^-) \leq f(x_0) \leq f(x_0^+) \qquad \textit{falls } f \textit{ monoton wächst}$$

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 (ansonsten betrachte man $-f$).
1. Wir zeigen: Ist $x_0 \in I$, so existiert $f(x_0^+) \in \mathbb{R}$ und es gilt $f(x_0^+) \geq f(x_0)$:
Da $f(x) \geq f(x_0)$ für alle $x > x_0$ gilt, existiert

$$y_0 := \inf_{x \in I \cap (x_0, \infty)} f(x)$$

und es gilt $y_0 \geq f(x_0)$. Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $x_\varepsilon > x_0$ mit $f(x_\varepsilon) < y_0 + \varepsilon$. Mit $\delta_\varepsilon := x_\varepsilon - x_0$ gilt dann für alle x mit $x_0 < x < x_0 + \delta_\varepsilon = x_\varepsilon$

$$y_0 \leq f(x) \leq f(x_\varepsilon) < y_0 + \varepsilon$$

Damit ist $f(x_0^+) = y_0$.
Entsprechend zeigt man die Existenz von $f(x_0^-) \in \mathbb{R}$ mit $f(x_0^-) \leq f(x_0)$.

2. Nach 1. können als Unstetigkeitsstellen nur Sprungstellen auftreten. Wir setzen

$$S(f) := \{x \in I : x \text{ Sprungstelle von } f\}$$

und zeigen, dass $S(f)$ abzählbar ist. O. E. sei $S(f)$ mindestens zweipunktig. Für $x \in S(f)$ setzen wir

$$I(x) := (f(x^-), f(x^+)) \, .$$

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 $f(x_1^+) \leq f\left(\frac{x_1+x_2}{2}\right) \leq f(x_2^-)$). Weiter 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. $\square$

Bemerkung D.3 Ist I ein beliebiges Intervall und ist $f : I \to \mathbb{R}$ monoton wachsend (bzw. fallend), so gilt zusätzlich: Ist der rechte Randpunkt b in I, so existiert $f(b^-)$, und es gilt $f(b^-) \leq f(b)$ (bzw. $\geq$); Ist der linke Randpunkt a in I, so existiert $f(a^+)$ und es gilt $f(a^+) \geq f(a)$ (bzw. $\leq$).

E Fundamentalsatz der Algebra

In B./D. 9.13 hatten wir die Existenz komplexer Wurzeln nachgewiesen. Die Existenz von Wurzeln bedeutet, dass Polynome $P : \mathbb{C} \to \mathbb{C}$ der Form $P(z) = z^n - c$ stets Nullstellen besitzen. Wir wollen zeigen, dass jedes Polynom $P : \mathbb{C} \to \mathbb{C}$ eine Nullstelle hat.

Bemerkung und Definition E.1 Sind $(V, \|\cdot\|)$ ein normierter Raum, (Y, d_Y) ein metrischer Raum, $X \subset V$ unbeschränkt und $f : X \to Y$, so schreiben wir

$$f(x) \to y_0 \qquad (\|x\| \to \infty)$$

falls zu jedem $\varepsilon > 0$ ein $R = R_\varepsilon > 0$ so existiert, dass

$$d_Y(f(x), y_0) < \varepsilon \qquad \text{für alle } x \in X \text{ mit } \|x\| > R\,.$$

Es gilt damit für alle $k \in \mathbb{N}$:

$$1/z^k \to 0 \qquad (|z| \to \infty)\,.$$

Ist also $P(z) = \sum\limits_{\nu=0}^{n} a_\nu z^\nu$ ein Polynom vom Grad $n \in \mathbb{N}$, so folgt

$$\frac{P(z)}{a_n z^n} = 1 + \sum_{\nu=0}^{n-1} \frac{a_\nu}{a_n} \frac{1}{z^{n-\nu}} \to 1 \qquad (|z| \to \infty)\,.$$

Satz E.2 *(Fundamentalsatz der Algebra)*
Es sei $n \in \mathbb{N}$ und $P : \mathbb{C} \to \mathbb{C}$ ein Polynom vom Grad n. Dann hat P eine Nullstelle $z_0 \in \mathbb{C}$.

Beweis. Nach B./D. E.1 existiert ein $R > 0$ so, dass

$$|P(z)| \geq \frac{1}{2}\, |a_n|\, |z_n|^n \qquad (|z| > R)\,.$$

Wählt man noch R so groß, dass $\frac{1}{2}|a_n| R^n > |a_0|$ ist, so folgt

$$|P(z)| > |a_0| = |P(0)| \qquad (|z| > R)\,.$$

Da $|P| : \mathbb{C} \to \mathbb{R}$ stetig ist, hat $|P|$ nach S. 11.17 ein Minimum bezüglich der kompakten Menge $\{z \in \mathbb{C} : |z| \leq R\}$, d.h. es existiert ein z_0 mit

$$|P(z_0)| \leq |P(z)| \qquad (|z| \leq R)\,.$$

Aus $|P(z_0)| \leq |P(0)|$ ergibt sich damit, dass $|P|$ ein absolutes Minimum an z_0 hat, also $|P(z_0)| \leq |P(z)|$ für alle $z \in \mathbb{C}$.

Wir zeigen: $P(z_0) = 0$.
Angenommen, nicht. Dann ist $Q : \mathbb{C} \to \mathbb{C}$ mit

$$Q(z) := \frac{P(z + z_0)}{P(z_0)} \qquad (z \in \mathbb{C})$$

ein Polynom vom Grad n mit $Q(0) = 1$ und $|Q(z)| \geq 1$ $(z \in \mathbb{C})$. Damit ist durch $R(z) := 1 - Q(z)$ ein Polynom vom Grad n definiert mit $R(0) = 0$. Weiter existieren ein $m \in \mathbb{N}$ und $b_m, \ldots, b_n \in \mathbb{C}$ mit $b_m \neq 0$ und

$$R(z) = b_m z^m + b_{m+1} z^{m+1} + \ldots + b_n z^n \, .$$

Ist $b_m = |b_m| e^{i\varphi}$, so gilt für $\psi = -\varphi/m$ und $t > 0$

$$R(te^{i\psi}) = t^m \left(|b_m| + \sum_{\nu=m+1}^{n} b_\nu t^{\nu-m} e^{i\nu\psi} \right)$$

mit

$$\delta(t) := \sum_{\nu=m+1}^{n} b_\nu e^{i\nu\psi} t^{\nu-m} \to 0 \;\; (t \to 0^+) \, .$$

Wählt man $0 < t < 1/\sqrt[m]{|b_m|}$ mit $|\delta(t)| < |b_m|$, so folgt (da $t^m |b_m| < 1$)

$$\begin{aligned} |Q(te^{i\psi})| &= |1 - R(te^{i\psi})| \leq \big|1 - t^m |b_m|\big| + t^m |\delta(t)| \\ &= 1 - t^m |b_m| + t^m |\delta(t)| < 1 \, . \end{aligned}$$

Widerspruch zu $|Q(z)| \geq 1$. □