Jürgen Müller

Einführung in die Mathematik/Analysis einer und mehrerer Veränderlicher

Skriptum zur den Vorlesung
Wintersemester 2016/2017 und Sommersemster 2017
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 $x \notin M$). Die Menge ohne Elemente heißt die **leere Menge** (Schreibweise: $\emptyset$ oder $\{\}$)
Es gibt verschiedene Möglichkeiten der Darstellung von Mengen, etwa die aufzählende Schreibweise oder auch die beschreibende Schreibweise, also eine Charakterisierung der Elemente. Die beschreibende Schreibweise hat allgemein die Form

$$M := \{x : x \text{ hat die Eigenschaft } E\},$$

wobei E irgendeine „Eigenschaft“ ist. Alternativ schreibt man statt $x :$ auch $x|$. In Weiteren werden wir das Symbol $:=$ als definierendes Gleichheitszeichen verwenden, d. h. die linke Seite wird durch die rechte definiert. Wir werden außerdem davon ausgehen, dass natürliche, ganze und rationale Zahlen samt arithmetischer Eigenschaften bekannt sind, werden aber im Anhang A kurz auf eine axiomatische Einführung eingehen. Man definiert

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

Definition 1.1 Es seien A, B Mengen.

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

2. A und B heißen **gleich** (Schreibweise $A = B$), falls $A \subset B$ und $B \subset A$ gilt. Sind dabei speziell $A := \{x\}$ und $B := \{y\}$ einpunktig, so nennen wir x und y **gleich** (Schreibweise: $x = y$; sind x und y ungleich, so schreibt man $x \neq y$).

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.

Ähnlich wie bei der obigen der Einführung von Mengen wollen wir auf eine eher intuitive Definition des zweiten grundlegenden Begriffes der Mathematik zurückgreifen, nämlich den einer Funktion.

Es seien dazu X und Y nichtleere Mengen. Eine **Funktion** (oder **Abbildung**) f von X nach Y ist eine „Vorschrift", die jedem $x \in X$ *genau ein* Element $f_x = f(x) \in Y$ zuordnet. Dabei heißen X der **Definitionsbereich** und Y der **Zielbereich** von f. Man schreibt $f : X \to Y$ und alternativ auch $X \ni x \mapsto f(x) \in Y$ oder $(f_x)_{x \in X}$. Im letzten Fall spricht man auch von einer **Familie** in Y und nennt dann X **Indexmenge**. Weiter setzt man
$$Y^X := \operatorname{Abb}(X, Y) := \{f : f \text{ Abbildung von } X \text{ nach } Y\}\ .$$
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.

Im Weiteren verwenden wir oft die Schreibweise $(x \in X)$ anstelle von „für alle $x \in X$".

Beispiel 1.2 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 1.3 Sind $n \in \mathbb{N}$ und X eine nichtleere Menge und ist $x : \{1, \ldots, n\} \to X$ eine Funktion, so schreibt man meist $(x_1, \ldots, x_n)$ statt $(x_j)_{j\in\{1,\ldots,n\}}$ und spricht dann von einem n-**Tupel** in X. Im Fall $n = 2$ spricht man auch von (geordneten) Paaren und im Fall $n = 3$ von Tripeln und verwendet dann oft eine indexfreie Schreibweise wie etwa (u, v) statt (x_1, x_2) beziehungsweise (u, v, w) statt (x_1, x_2, x_3).
Weiter setzt man

$$X^n := X^{\{1,\ldots,n\}}$$

und für beliebige Mengen $A_1, \ldots, A_n \subset X$

$$A_1 \times \cdots \times A_n := \{(x_1, \ldots, x_n) \in X^n : x_j \in A_j \text{ für } j \in \{1, \ldots, n\}\} .$$

Eine Teilmenge R von $X \times X$ heißt eine **Relation** auf (oder in) X. Man schreibt dann auch uRv, falls $(u, v) \in R$ gilt.

Definition 1.4 Es sei M eine nichtleere Menge. Eine Funktion $f : M \times M \to M$ heißt **(innere) Verknüpfung** auf M. Man wählt dann oft ein nichtalphabetisches Zeichen wie $\cdot$, $\circ$, $*$, $\times$, $+$ für f und schreibt wieder xfy statt $f(x, y) := f\big((x, y)\big)$ für $x, y \in M$, also etwa

$$x \cdot y\,,\ x \circ y\,,\ x * y\,,\ x \times y\,,\ x + y\,.$$

Im Fall des **Multiplikationszeichens** $\cdot$ schreibt man meist kurz xy statt $x \cdot y$.

Wir werden im Weiteren, wie bereits angedeutet, die Kenntnis der „üblichen“ Verknüpfungen $+$ und $\cdot$ auf $\mathbb{N}, \mathbb{N}_0, \mathbb{Z}$ und $\mathbb{Q}$ sowie der Relationen $<$ oder auch $\leq$ auf $\mathbb{Q}$ (samt Eigenschaften) voraussetzen. Genaueres findet sich im Anhang A.

Definition 1.5 Sind X, Y Mengen und ist $f : X \to Y$, so heißt für $B \subset Y$

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

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

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

Bildmenge von A unter f. Speziell heißt

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

Wertebereich von f. Ist $W(f)$ einpunktig, so heißt f **konstant**.

Beispiel 1.6 In der Situation von B. 1.2 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 $W(f) = \{y \in \mathbb{N} : y \text{ gerade}\}$.

Definition 1.7 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**, 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 1.8 Es sei f wie im B 1.2. 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 1.9 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

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

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

Beispiel 1.10 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 $g \circ f \neq f \circ g$ (da etwa $(g \circ f)(1) = 2 \neq 4 = (f \circ g)(1)$).

Satz 1.11 *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

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

Damit sind die beiden Funktionen gleich. □

Bemerkung und Definition 1.12 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 $f^{-1} \circ f : X \to X$ und

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

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 **identische Abbildung** auf X bezeichnet. Genauso gilt $f \circ f^{-1} = \mathrm{id}_Y$ und außerdem ist auch $f^{-1} : Y \to X$ bijektiv.

Definition 1.13 Es sei $I \neq \emptyset$ eine Menge, und es seien $(A_\alpha) := (A_\alpha)_{\alpha \in I}$ eine Familie von Mengen. Dann heißen

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

Vereinigung von (A_α) und

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

Durchschnitt von (A_α).
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 $A_M = M$).
Ist I in aufzählender Form gegeben, so setzt man $\bigcup$ bzw. $\bigcap$ meistens zwischen die einzelnen Mengen, also etwa im Falle $I = \{1, 2, 3\}$

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

Auch im Fall eines Mengensystems schreibt man alternativ $\bigcup$ und $\bigcap$ 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 1.14 Sind $A := \{2k : k \in \mathbb{Z}\}$ und $B := \{3k : k \in \mathbb{Z}\}$, so gilt

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

Nach Definition sind zwei Mengen gleich, wenn die erste Teilmenge der zweiten und die zweite Teilmenge der ersten ist. Daher beweist man üblicherweise die Gleichheit, indem man die beiden Inklusionen getrennt nachweist. Wir deuten dies im Weiteren durch die Schreibweise $\subset$: und $\supset$: in den entsprechenden Beweisen an.

Satz 1.15 *Es seien $(A_\alpha)_{\alpha \in I}$ eine Familie von Mengen und B eine Menge. Dann gilt*

1.
$$B \cap \Big(\bigcup_{\alpha \in I} A_\alpha\Big) = \bigcup_{\alpha \in I} (B \cap A_\alpha) \quad \textit{und} \quad B \cup \Big(\bigcap_{\alpha \in I} A_\alpha\Big) = \bigcap_{\alpha \in I} (B \cup A_\alpha).$$

2. *(**De Morgansche Regeln**):*
$$B \setminus (\bigcup_{\alpha \in I} A_\alpha) = \bigcap_{\alpha \in I} (B \setminus A_\alpha) \quad \textit{und} \quad B \setminus (\bigcap_{\alpha \in I} A_\alpha) = \bigcup_{\alpha \in I} (B \setminus A_\alpha).$$

Beweis. Wir werden exemplarisch die Beweise der ersten Aussage in 1. und 2. führen. Die zweiten ergeben sich in ähnlicher Weise. Wir schreiben kurz $\bigcup$ statt $\bigcup\limits_{\alpha \in I}$.
1. $\subset$: Es sei $x \in B \cap (\bigcup A_\alpha)$. Dann ist $x \in B$ und $x \in \bigcup A_\alpha$, also $x \in B$ und $x \in A_\beta$ für ein $\beta \in I$. Damit ist $x \in B \cap A_\beta$, also auch $x \in \bigcup (B \cap A_\alpha)$.
$\supset$: Es sei $x \in \bigcup (B \cap A_\alpha)$. Dann existiert ein $\beta \in I$ mit $x \in B \cap A_\beta$. Damit ist $x \in B$ und $x \in A_\beta$, also auch $x \in B$ und $x \in \bigcup A_\alpha$, d. h. $x \in B \cap (\bigcup A_\alpha)$.
2. $\subset$: Es sei $x \in B \setminus (\bigcup A_\alpha)$. Dann ist $x \in B$ und $x \notin \bigcup A_\alpha$, 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 $x \in \bigcap (B \setminus A_\alpha)$.

$\supset$: Es sei $x \in \bigcap(B \setminus A_\alpha)$. Dann ist $x \in B \setminus A_\alpha$ 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 $x \notin \bigcup A_\alpha$, d. h. $x \in B \setminus (\bigcup A_\alpha)$. $\square$

Satz 1.16 *Es seien X, Y Mengen und $f : X \to Y$.*

1. *Ist $(B_\alpha)_{\alpha \in I}$ eine Familie von Mengen in Y, so gilt*
$$f^{-1}\big(\bigcup_{\alpha \in I} B_\alpha\big) = \bigcup_{\alpha \in I} f^{-1}(B_\alpha) \quad \textit{und} \quad f^{-1}\big(\bigcap_{\alpha \in I} B_\alpha\big) = \bigcap_{\alpha \in I} f^{-1}(B_\alpha).$$

2. *Ist $(A_\alpha)_{\alpha \in I}$ eine Familie von Mengen in X, so gilt*
$$f\big(\bigcup_{\alpha \in I} A_\alpha\big) = \bigcup_{\alpha \in I} f(A_\alpha) \quad \textit{und} \quad f\big(\bigcap_{\alpha \in I} A_\alpha\big) \subset \bigcap_{\alpha \in I} f(A_\alpha).$$

Beweis. Wir beschränken uns wieder auf die jeweils ersten Aussagen.
1. $\subset$: Es sei $x \in f^{-1}\big(\bigcup B_\alpha\big)$. Dann ist $f(x) \in \bigcup B_\alpha$, 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 $x \in \bigcup f^{-1}(B_\alpha)$.
$\supset$: Ist $\beta \in I$, so ist $B_\beta \subset \bigcup B_\alpha$, also auch $f^{-1}(B_\beta) \subset f^{-1}\big(\bigcup B_\alpha\big)$. Da $\beta \in I$ beliebig war, gilt $\bigcup f^{-1}(B_\alpha) \subset f^{-1}\big(\bigcup B_\alpha\big)$.
2. $\subset$: Es sei $y \in f\big(\bigcup A_\alpha\big)$. Dann existiert ein $x \in \bigcup A_\alpha$ 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 $y \in \bigcup f(A_\alpha)$.
$\supset$: Ist $\beta \in I$, so ist $A_\beta \subset \bigcup A_\alpha$, also auch $f(A_\beta) \subset f\big(\bigcup A_\alpha\big)$. Da $\beta \in I$ beliebig war, gilt $\supset$. $\square$

Bemerkung 1.17 Man beachte, dass in der letzen Aussage in S. 1.16 kein Gleichheitszeichen steht. Tatsächlich liegt Gleichheit für alle Familien (A_α) genau dann vor, wenn f injektiv ist ([Ü]).

2 Monoide, Gruppen und Ringe

Ziel dieses Abschnittes ist es, algebraische Strukturen zu formalisieren. Dazu betrachten wir allgemein Mengen, die mit gewissen Verknüpfungen versehen sind.

Definition 2.1 Es seien M eine nichtleere Menge und $\cdot$ eine Verknüpfung auf M.

1. Die Verknüpfung heißt **assoziativ**, falls $x\,(yz) = (xy)\,z$ für $x, y, z \in M$ gilt, und **kommutativ**, falls $xy = yx$ für $x, y \in M$ gilt. Ist $\cdot$ assoziativ, so heißt $(M, \cdot)$ eine **Halbgruppe**. Ist $\cdot$ zudem kommutativ, so heißt die Halbgruppe **abelsch** (oder auch kommutativ).

2. Ein $e \in M$ heißt **neutral** (bezüglich $\cdot$) falls $ex = xe = x$ für alle $x \in M$ gilt. Existiert in einer Halbgruppe $(M, \cdot)$ ein neutrales Element e, so heißt $(M, \cdot, e)$. ein **Monoid**.

Bei assoziativen Verknüpfungen lässt man die Klammern meist weg, setzt also zum Beispiel $xyz := (xy)z = x(yz)$. Das **Pluszeichen** $+$ wird üblicherweise nur für kommutative Verknüpfungen benutzt. Schließlich schreibt man auch kurz M statt $(M, \cdot)$ oder $(M, \cdot, e)$.

Beispiel 2.2 1. Das Paar $(\mathbb{N}, +)$ ist eine abelsche Halbgruppe, $(\mathbb{N}_0, +, 0)$, $(\mathbb{N}, \cdot, 1)$ und $(\mathbb{Z}, \cdot, 1)$ sind abelsche Monoide.
2. Es sei X eine Menge. Dann heißt $\mathscr{P}(X) := \{A : A \subset X\}$ die **Potenzmenge** von X. Ist $(X, \cdot)$ eine Halbguppe, so definiert das **Komplexprodukt**

$$A \cdot B := \{xy : x \in A,\, y \in B\} \qquad (A, B \subset X)$$

eine assoziative Verknüpfung $\cdot$ auf $\mathscr{P}(X)$, also ist $(\mathscr{P}(X), \cdot)$ eine Halbgruppe. Ist $(X, \cdot, e)$ ein Monoid, so ist auch $(\mathscr{P}(X), \cdot, \{e\})$ ein Monoid. Im Falle einer einpunktigen Menge $A = \{x\}$ schreibt man meist kurz xB statt $\{x\} \cdot B$ und im Falle des Pluszeichens als Verknüpfung auf X natürlich auch $A + B$ statt $A \cdot B$. Die Menge $A + B$ heißt dann auch **Minkowski-Summe** von A und B.

Definition 2.3 Es sei $(M, \cdot, e)$ ein Monoid. Ist $x \in M$, so heißt ein $y \in M$ **linksinvers** zu x, falls $yx = e$, **rechtsinvers** zu x, falls $xy = e$, und **invers** zu x, falls $yx = xy = e$ gilt. Entsprechend heißt dann jeweils x **(links-, rechts-)invertierbar**. Ist jedes $x \in M$ invertierbar, so heißt M eine **Gruppe**.

Bemerkung 2.4 Es sei $(M, \cdot, e)$ ein Monoid.
1. Neutrale Elemente sind eindeutig, denn sind e und e' neutral, so ist

$$e' = ee' = e.$$

Auch inverse Elemente sind im Falle der Existenz eindeutig. Genauer gilt: Sind $x, y_1, y_2 \in M$ mit y_1 links- und y_2 rechtsinvers zu x, so ist

$$y_1 \;=\; y_1 e \;=\; y_1(xy_2) \;=\; (y_1 x)y_2 \;=\; ey_2 \;=\; y_2.$$

Man bezeichnet das inverse Element zu x mit x^{-1}. Bei Verwendung des Verknüpfungszeichens $+$ schreibt man meist $-x$ (und dann auch kurz $x - y$ satt $x + (-y)$).
2. Es seien $x, y \in M$ invertierbar. Dann sind auch x^{-1} und xy invertierbar mit

$$\left(x^{-1}\right)^{-1} = x \quad \text{und} \quad (xy)^{-1} \;=\; y^{-1}x^{-1}$$

(da $x^{-1}x = xx^{-1} = e$ und $xyy^{-1}x^{-1} = xx^{-1} = e = y^{-1}y = y^{-1}x^{-1}xy$).
Ist U die Menge der invertierbaren Elemente in M, so ist damit $(U, \cdot, e)$ eine Gruppe (mit $\cdot$ eingeschränkt auf $U \times U$ und Zielbereich U).
3. Ist jedes $x \in M$ linksinvertierbar, so ist M schon eine Gruppe ([Ü]). Entsprechendes gilt mit rechts statt links.
4. Sind $a, b \in M$ und ist a invertierbar, so sind die Gleichungen $ax = b$ und $ya = b$ eindeutig lösbar, nämlich durch $x = a^{-1}b$ beziehungsweise $y = ba^{-1}$ ([Ü]). Ist M eine Gruppe, so sind die Gleichungen damit für alle a, b eindeutig lösbar.

Beispiel 2.5 1. Die Tripel $(\mathbb{Z}, +, 0)$, $(\mathbb{Q}, +, 0)$ und $(\mathbb{Q}\backslash\{0\}, \cdot, 1)$ sind abelsche Gruppen. Im Monoid $(\mathbb{Z}, \cdot, 1)$ sind nur ± 1 invertierbar.
2. Es sei $X \neq \emptyset$ ein Menge. Dann ist $\mathrm{Abb}(X) := \mathrm{Abb}(X, X)$ mit der Komposition $\circ$ von Funktionen als Verknüpfung ein Monoid mit neutralem Element id_X. Dabei ist $f : X \to X$ genau dann invertierbar, wenn f bijektiv ist ([Ü]). Nach B. 2.4.2 ist damit

$$S(X) \;:=\; \{f \in \mathrm{Abb}(X) : f \text{ bijektiv}\}$$

eine Gruppe. Zu $f \in S(X)$ invers ist die Umkehrfunktion, die glücklicherweise sowieso schon mit f^{-1} bezeichnet wird. $S(X)$ heißt **symmetrische Gruppe** von X, und ein Element $f \in S(X)$ heißt **Permutation** von X.
Für $n \in \mathbb{N}$ heißt speziell $S_n := S(\{1, \ldots, n\})$ die n-te symmetrische Gruppe. Für $n \geq 3$ ist S_n nicht abelsch ([Ü]).

Wir wollen nun Produkte und Summen von mehr als zwei Faktoren beziehungsweise Summanden definieren.

Bemerkung und Definition 2.6 Es seien $(M, \cdot, e)$ ein Monoid und $N \in \mathbb{N}$. Sind $x_1, \ldots, x_N \in M$, so setzen wir $\prod_{k=1}^{0} x_k := e$ und

$$\prod_{k=1}^{n} x_k := \Big(\prod_{k=1}^{n-1} x_k\Big) \cdot x_n$$

für $n = 1, \ldots, N$. Außerdem schreiben wir $x^n := \prod_{k=1}^{n} x$, also im Falle $x_1 = \ldots = x_n = x$. Insbesondere ist $x^0 = e$. Ist x invertierbar, so setzen wir auch $x^{-n} := (x^{-1})^n$ für $n \in \mathbb{N}$. Sind allgemeiner $x_{m+1}, \ldots, x_n \in M$ für $m, n \in \mathbb{Z}$ und $n \geq m$, so schreiben wir auch

$$\prod_{j=m+1}^{n} x_j := \prod_{k=1}^{n-m} x_{k+m}.$$

Im Falle des Pluszeichens als Verknüpfung schreiben wir statt $\prod$ jeweils $\sum$. Außerdem schreiben wir dann nx statt x^n. Man beachte, dass die Abbildung $(n, x) \mapsto nx$ im Allgemeinen keine Verknüpfung auf M ist.

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. Man nimmt an, dass $A(n)$ oder auch $A(1), \ldots, A(n)$ für ein beliebiges $n \in \mathbb{N}$ richtig ist (Induktionsannahme) und zeigt, dass aus der Induktionsannahme die Richtigkeit der Aussage $A(n+1)$ folgt (Induktionsschritt).

Aus 1. und 2. ergibt sich, dass $A(n)$ für alle $n \in \mathbb{N}$ richtig ist.
Manchmal möchte man statt für $n \geq 1$ die Behauptung $A(n)$ für alle $n \in \mathbb{N}_0$, $n \geq n_0$ zeigen. Dann macht man den Induktionsanfang nicht für $n = 1$, sondern für $n = n_0$ und den Induktionsschritt von n auf $n+1$ für beliebiges $n \geq n_0$.

Bemerkung 2.7 Es sei $q \in \mathbb{N}$. Zur Illustation eines Induktionsbeweises zeigen wir folgende Aussage, die die Eindeutigkeit der q-adischen Darstellung natürlicher Zahlen (vgl. Anhang A) impliziert.

Für alle $n \in \mathbb{N}$ gilt: Sind $a_j \in \{0, \dots, q-1\}$ für $j = 0, \dots, n-1$, so ist

$$\sum_{j=0}^{n-1} a_j q^j < q^n \ .$$

1. Induktionsanfang: Für $n = 1$ und $a_0 \in \{0, \dots, q-1\}$ gilt $a_0 q^0 = a_0 < q$.
2. Induktionsannahme:

Für ein $n \in \mathbb{N}$ gelte: Sind $a_0, \dots, a_{n-1} \in \{0, \dots, q-1\}$, so ist $\sum\limits_{j=0}^{n-1} a_j q^j < q^n$.

Induktionsschritt:

Sind $a_0, \dots, a_n \in \{0, \dots, q-1\}$, so ist nach Induktionsannahme

$$\sum_{j=0}^{n} a_j q^j = \big(\sum_{j=0}^{n-1} a_j q^j\big) + a_n q^n < q^n + a_n q^n \leq q^n + (q-1)q^n = q^{n+1}.$$

Also gilt die Behauptung für $n+1$.

Bemerkung und Definition 2.8 Es sei $(M, \cdot, e)$ ein abelsches Monoid. Dann kann man (induktiv) zeigen, dass für $x_1, x_2, x \in M$ und $m, m_1, m_2 \in \mathbb{N}_0$ folgende Potenzgesetze gelten:

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

Ist M eine ablesche Gruppe, so gelten die Potenzgesetze auch für $m, m_1, m_2 \in \mathbb{Z}$.

Weiterhin kann man (induktiv und nicht ganz leicht) zeigen, dass für $\varphi \in S_n$ und $x_1, \dots, x_n \in M$

$$\prod_{k=1}^{n} x_{\varphi(k)} = \prod_{k=1}^{n} x_k$$

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

$$\prod_{j \in I} x_j := \prod_{k=1}^{n} x_{\psi(k)} \ ,$$

wobei $\psi : \{1, \ldots, n\} \to I$ eine beliebige bijektive Abbildung ist. Ist $(y_j)_{j\in I}$ eine weitere Familie in M, so gilt damit

$$\prod_{j\in I}(x_j y_j) = \prod_{j\in I} x_j \prod_{j\in I} y_j.$$

Wir kommen jetzt zu algebraischen Strukturen mit zwei Verknüpfungen.

Definition 2.9 Es sei R eine Menge und es seien $+$ und $\cdot$ Verknüpfungen auf R mit:

(R1) $(R, +, 0)$ ist eine abelsche Gruppe.

(R2) $(R, \cdot, 1)$ ist ein Monoid.

(R3) Die Verknüpfung $\cdot$ ist **distributiv über** $+$, d.h. für $x, y, z \in R$ gilt

$$x(y + z) \;=\; (xy) + (xz) \qquad \text{und} \qquad (x + y)z \;=\; (xz) + (yz).$$

Dann heißen $(R, +, \cdot)$ **Ring**, das neutrale Element 0 zu $+$ **Null(element)** und das neutrale Element 1 zu $\cdot$ **Eins(element)**. Ist $(R, \cdot, 1)$ abelsch, so heißt der Ring **kommutativ**. Wir schreiben manchmal deutlicher 0_R und 1_R für die neutralen Elemente eines Ringes. Andererseits schreiben wir oft kurz R statt $(R, +, \cdot)$.

Bemerkung 2.10 Standardbeispiele kommutativer Ringe sind $(\mathbb{Z}, +, \cdot)$ und $(\mathbb{Q}, +, \cdot)$. Man verwendet wie in $(\mathbb{Q}, +, \cdot)$ auch in allgemeinen Ringen R Punkt-vor-Strich-Schreibweisen, also zum Beispiel $x + yz := x + (yz)$. Induktiv ergeben sich für $x \in R$ und endliche Familien $(x_j)_{j\in I}$ in R die allgemeinen Distributivgesetze

$$x \sum_{j\in I} x_j = \sum_{j\in I} x x_j \quad \text{und} \quad \Big(\sum_{j\in I} x_j\Big)x = \sum_{j\in I} x_j x.$$

Bemerkung und Definition 2.11 Es sei R ein Ring. Dann gilt

1. 0 ist **absorbierend** für R, d. h. $0 \cdot x = x \cdot 0 = 0$ für $x \in R$.
2. $(-x)y = x(-y) = -xy$ und $(-x)(-y) = xy$ für $x, y \in R$.
3. $x(y - z) = xy - xz$ und $(x - y)z = xz - yz$ für $x, y, z \in R$.

Bemerkung 2.12 Es seien R ein Ring und X eine nichtleere Menge. Wir definieren für $f, g \in R^X$ die Funktionen $f + g \in R^X$ und $f \cdot g \in R^X$ argumentweise durch

$$(f+g)(x) := f(x) + g(x) \quad \text{und} \quad (f \cdot g)(x) := f(x) \cdot g(x) \quad (x \in X).$$

Damit ist $R^X = (R^X, +, \cdot)$ ein Ring mit Nullelement 0_{R^X} und Einselement 1_R, definiert durch $0_{R^X}(x) := 0_R$ und $1_{R^X}(x) := 1_R$ für $x \in X$. Ist R kommutativ, so ist auch R^X kommutativ.

Definition 2.13 Ein Ring $(R, +, \cdot)$ mit Nullelement 0_R und Einselemnt 1_R heißt **Körper**, falls $(R^*, \cdot, 1_R)$ mit

$$R^* := R \setminus \{0\}$$

eine abelsche Gruppe ist. Wir schreiben im Weiteren auch $1/x$ statt x^{-1} für das inverse Element von $x \neq 0_R$ bezüglich der Multiplikation und x/y statt $xy^{-1} (= y^{-1}x)$.

Bemerkung 2.14 Eine wichtige Eigenschaft von Körpern ist die **Nullteilerfreiheit**: Sind $x, y \in R$ mit $xy = 0$, so ist $x = 0$ oder $y = 0$ (da $\cdot$ eine Verknüpfung auf $\mathbb{R}^*$ ist!).

Beispiel 2.15 1. Der Ring $(\mathbb{Q}, +, \cdot)$ ist ein Körper, der Ring $(\mathbb{Z}, +, \cdot)$ nicht.
2. (Binärkörper) Es sei $\mathbb{F}_2 := \{\heartsuit, \clubsuit\}$ mit den Verknüpfungen

$+$	$\heartsuit$	$\clubsuit$
$\heartsuit$	$\heartsuit$	$\clubsuit$
$\clubsuit$	$\clubsuit$	$\heartsuit$

$\cdot$	$\heartsuit$	$\clubsuit$
$\heartsuit$	$\heartsuit$	$\heartsuit$
$\clubsuit$	$\heartsuit$	$\clubsuit$

Dann ist $(\mathbb{F}_2, +, \cdot)$ ein Körper mit $\heartsuit = 0 = 0_{\mathbb{F}_2}$ und $\clubsuit = 1 = 1_{\mathbb{F}_2}$ ([Ü]). In der Binärarithmetik ist also $1 + 1 = 0$.

3 Geometrische Summenformel und binomische Formel

Wir kommen nun zu verschiendenen Formeln, die in kommutativen Ringen gelten.

Satz 3.1 *Es sei* $(R,+,\cdot)$ *ein kommutativer Ring. Dann gilt für alle* $a,b \in R$ *und alle* $n \in \mathbb{N}$

$$a^n - b^n = (a-b)\sum_{\nu=0}^{n-1} a^\nu b^{n-1-\nu}\ . \tag{3.1}$$

Beweis. Es gilt

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

□

Bemerkung 3.2 (geometrische Summenformel) Ist $(R,+,\cdot)$ ein Körper, so ist für $x \neq 1$ nach S. 3.1

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

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

Definition 3.3 Für $n \in \mathbb{N}_0$ definiert man n-**Fakultät** durch

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

und für $n,\nu \in \mathbb{N}_0$ den **Binomialkoeffizient** n **über** ν durch

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

Es gilt also etwa $6! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 = 720$ und

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

Wir stellen einige Eigenschaften der Binomialkoeffizienten zusammen.

Satz 3.4 *Für $n, \nu \in \mathbb{N}_0$ gilt*

$$\binom{n}{\nu} = \begin{cases} \dfrac{n!}{\nu!(n-\nu)!} = \dbinom{n}{n-\nu}, & \text{falls } \nu \le n \\ = 0, & \text{falls } \nu > n \end{cases} .$$

Beweis. Es gilt für $\nu \le n$

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

Damit ist auch

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

Für $\nu > n$ ist $n - \nu + 1 \le 0$ und damit $\prod_{k=1}^{\nu}(n-k+1) = 0$, 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 gilt für $\nu \in \{1, \dots, n\}$

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

Für $\nu = n+1$ ist nach S. 3.4

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

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

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 sogenannte **Pascalsche 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

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

Satz 3.6 *(binomische Formel)*
Es sei $(R, +, \cdot)$ ein kommutativer Ring. Dann gilt für alle $a, b \in R$ und alle $n \in \mathbb{N}_0$

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

Beweis.

1. Für $n = 0$ gilt $(a+b)^0 = 1 = \sum\limits_{\nu=0}^{0} \binom{0}{\nu} a^\nu b^{0-\nu}$.

2. Für ein $n \in \mathbb{N}_0$ gelte $(a+b)^n = \sum\limits_{\nu=0}^{n} \binom{n}{\nu} a^\nu b^{n-\nu}$. Dann folgt mit S. 3.5

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

Also gilt die Behauptung dann auch für $n + 1$. □

Beispiel 3.7 Für $n = 6$ gilt

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

Bemerkung 3.8 Als Spezialfälle aus S. 3.6 ergeben sich interessante Beziehungen für das Pascalsche Dreieck: Für ($R = \mathbb{Z}$ und) $a = 1, b = 1$ ergibt sich

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

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

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

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

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

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, also genau dann, wenn M und N gleich viele Elemente haben. Außerdem gilt: Ist $(M_j)_{j\in J}$ eine endliche Familie endlicher Mengen, und sind die Mengen **paarweise disjunkt**, d. h. $M_j \cap M_k = \emptyset$ für $j, k \in J$, $j \neq k$, so ist

$$\# \bigcup_{j\in J} M_j = \sum_{j\in J} \#M_j.$$

Satz 3.9 *Für $n \in \mathbb{N}$ gilt: Ist X eine n-elementige Menge, so ist*

$$\#(S(X)) = n!\ .$$

Beweis. Wir führen den Beweis per Induktion nach n.
1. Induktionsanfang: Für $n = 1$ ist die Behauptung klar.
2. Induktionsschritt: Es sei X eine $(n + 1)$-elementige Menge. Ohne Einschränkung können wir $X = \{1, \ldots, n + 1\}$ (also $S(X) = S_{n+1}$) annehmen. Wir definieren

$$T_j := \big\{\sigma \in S_{n+1} : \sigma(j) = n + 1\big\} \qquad (j = 1, \ldots, n + 1)\ .$$

Dann ist

$$\bigcup_{j=1}^{n+1} T_j = S_{n+1} \quad \text{und} \quad T_j \cap T_k = \emptyset \quad (j \neq k).$$

Also ist $\#(S_{n+1}) = \sum\limits_{j=1}^{n+1} \#(T_j)$.

Definiert man für $\sigma \in T_j$ die Funktion $\tau_\sigma \in S_n$ durch

$$\tau_\sigma(k) := \begin{cases} \sigma(k), & (k = 1, \ldots, j-1) \\ \sigma(k+1), & (k = j, \ldots, n) \end{cases},$$

so ist $T_j \ni \sigma \mapsto \tau_\sigma \in S_n$ eine bijektive Abbildung. Nach Induktionsannahme gilt $\#(S_n) = n!$ und damit auch $\#(T_j) = n!$. Also ist

$$\#(S_{n+1}) = \sum_{j=1}^{n+1} n! = (n+1)n! = (n+1)!\ .$$

$\square$

Bemerkung 3.10 Ist X eine n-elementige Menge und ist $\mathscr{A}_\nu \subset \mathscr{P}(X)$ für $\nu \in \{0, \ldots, n\}$ die Menge der ν-elementigen Teilmengen von X, so ist ([Ü])

$$\#(\mathscr{A}_\nu) = \binom{n}{\nu}.$$

Nach B. 3.8 ist damit auch

$$\#(\mathscr{P}(X)) = \sum_{\nu=0}^{n} \#(\mathscr{A}_\nu) = 2^n\ .$$

4 Geordnete Körper, reelle und komplexe Zahlen

Definition 4.1 Es sei $X \neq \emptyset$ eine Menge. Eine Relation $<$ auf X heißt **Ordnung** (auf X), falls gilt

($O1$) Für alle $x, y \in X$ gilt entweder $x = y$ oder $x < y$ oder $y < x$ (Trichotomie).

($O2$) Für $x, y, z \in X$ gilt: aus $x < y$ und $y < z$ folgt $x < z$ (Transitivität).

Das Paar $(X, <)$ heißt dann eine **geordnete Menge**. Außerdem bedeutet $x \leq y$, dass entweder $x < y$ oder $x = y$ gilt. Schießlich scheiben wir auch $y > x$ statt $x < y$ und $y \geq x$ statt $x \leq y$.

Definition 4.2 Es seien $(X, <)$ geordnet und $M \subset X$.
1. M heißt **nach oben beschränkt**, wenn ein $s \in X$ existiert mit

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

Ein solches s heißt dann eine **obere Schranke** von M. Ist dabei $s \in M$, so heißt s **Maximum** von M (Schreibweise: $\max M := s$)
2. M heißt **nach unten beschränkt**, wenn ein $s \in X$ existiert mit

$$x \geq s \quad \text{für alle} \quad x \in M\ .$$

Ein solches s heißt dann **untere Schranke** von M. Ist dabei $s \in M$, so heißt s **Minimum** von M (Schreibweise: $\min M := s$)
3. M heißt **beschränkt** wenn M nach oben und nach unten beschränkt ist.

Beispiel 4.3 Es sei $(X, <) = (\mathbb{Q}, <)$. Dann ist die Menge $M = \{1/n : n \in \mathbb{N}\}$ beschränkt mit $\max M = 1$, aber M hat kein Minimum!

Bemerkung 4.4 Per Induktion kann man zeigen ([Ü]): Ist $(X, <)$ geordnet, so hat jede nichtleere, endliche Menge $M \subset X$ ein Maximum und ein Minimum. Per Induktion (oder aus dem Wohlordnungsprinzip; siehe Anhang A) folgt auch, dass jede nichtleere, nach oben beschränkte Menge $M \subset \mathbb{Z}$ ein Maximum und jede nichtleere, nach unten beschränkte Menge $M \subset \mathbb{Z}$ ein Minimum hat.

Definition 4.5 Es sei $K = (K, +, \cdot)$ ein Körper. Ist $<$ eine Ordnung auf K, so heißt $K = (K, +, \cdot, <)$ **geordnet**, wenn für $x, y \in K$ folgende Verträglichkeiten mit der Addition und Multiplikation erfüllt sind:

($O3$) Aus $x < y$ folgt $x + z < y + z$ für alle $z \in K$ (1. Monotoniegesetz).
($O4$) Aus $x < y$ und $z > 0$ folgt $xz < yz$ (2. Monotoniegesetz).

Wir nennen $x \in K$ **positiv**, falls $x > 0$ gilt und **negativ**, falls $x < 0$ gilt. Außerdem setzen wir $K_+ := \{x \in K : x > 0\}$ und $K_- := \{x \in K : x < 0\}$.

Satz 4.6 *Es seien $K = (K, +, \cdot, <)$ ein geordneter Körper und $x, y \in K$. Dann gilt*

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

Beweis. 1. Aus $0 < x$ folgt $-x = 0 + (-x) < x + (-x) = 0$, mit ($O3$), d. h. $-x < 0$. Entsprechend folgt aus $-x < 0$ auch $0 = x + (-x) < x + 0 = x$.
2. Sind $x, y > 0$, so folgt mit ($O4$) sofort $0 = 0y < xy$.
Es seien $x, y < 0$. Dann ist $-y > 0$ nach 1. Wegen $x < 0$ ergibt sich mit ($O4$)

$$-(xy) = x(-y) < 0(-y) = 0\ ,$$

also $xy > 0$ mit 1.
3. Ergibt sich unmittelbar aus 2. und ($O1$).
4. Zunächst ist $1/x > 0$, denn angenommen, es gilt $1/x \leq 0$ und damit $1/x < 0$. Dann folgt $1 = x/x < x \cdot 0 = 0$ mit ($O4$), im Widerspruch zu 3. Genauso ist $1/y > 0$. Aus $x < y$ ergibt sich also $x/y < y/y = 1$ mit ($O4$) und wieder mit ($O4$)

$$\frac{1}{y} = \frac{1}{y} \cdot \frac{x}{x} = \frac{x}{y} \cdot \frac{1}{x} < 1 \cdot \frac{1}{x} = \frac{1}{x}.$$

□

Beispiel 4.7 1. $(\mathbb{Q}, +, \cdot, <)$ ist ein geordneter Körper.
2. Im Binärkörper $(\mathbb{F}_2, +, \cdot)$ existiert keine Ordnungsrelation mit den Eigenschaften aus D. 4.5. (Angenommen doch. Dann ist $1 > 0$ nach S. 4.6.3, also auch $0 = 1 + 1 > 1 + 0 = 1$, im Widerspruch zu ($O1$).)

Bemerkung 4.8 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 $x > 0$ auch $0 < x^n < y^n$.
2. Ist $x > 0$, so ist auch $nx > mx > 0$ 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 ist die Menge $\{z \in K : x < z < y\}$ unendlich.
(Denn: Für alle $n, m \in \mathbb{N}$ mit $n > m$ ist $n1 > m1 > 0$, also $1/(m1) > 1/(n1) > 0$ und folglich

$$x < x + (y - x)/(n1) < x + (y - x)/(m1) \leq y.)$$

Im Allgemeinen sind in geordneten Körpern Gleichungen der Form

$$x^n = c,$$

wobei $c \in K, n \in \mathbb{N}, n > 1$ nicht lösbar. Ist $c < 0$ und ist n gerade, so ist dies nach S. 4.6.3 ohnehin ausgeschlossen. Aber auch im Falle $c > 0$ existiert im Allgemeinen keine Lösung.

Satz 4.9 *Für alle $x \in \mathbb{Q}$ ist $x^2 \neq 2$.*

Beweis. 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. Damit 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:

- 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.
- Erweitern von $(\mathbb{R}, +, \cdot)$ zu einem Körper $(\mathbb{C}, +, \cdot)$ so, dass $x^n = c$ für alle $n \in \mathbb{N}$ und $c \in \mathbb{C}$ lösbar ist.

Bemerkung und Definition 4.10 Es sei $(X, <)$ geordnet, und es sei $M \subset X$. Mit einer oberen Schranke s von M ist natürlich jedes $t \in X$ mit $t > s$ ebenfalls eine obere Schranke für M. Es stellt sich in natürlicher Weise die Frage nach kleinsten oberen (und größten unteren) Schranken.
Eine obere Schranke $s^* \in X$ von M heißt **kleinste obere Schranke** (oder **Supremum**) von M, falls $s \geq s^*$ für jede obere Schranke s von M gilt. Eine untere Schranke $s_* \in X$ von M heißt **größte untere Schranke** (oder **Infimum**) von M, falls $s \leq s_*$ für jede untere Schranke s von M gilt.
Aus der Definition ergibt sich sofort, dass für jedes M höchstens ein Supremum und ein Infimum existieren. Wir schreiben im Falle der Existenz

$$\sup M := s^*$$

beziehungsweise

$$\inf M := s_* .$$

Existiert $\max M$, so gilt $\sup M = \max M$. Im Falle der Existenz von $\min M$ ist $\inf M = \min M$.

Beispiel 4.11 Es sei $(X, <) = (\mathbb{Q}, <)$.
1. Ist $M = \{1/n : n \in \mathbb{N}\}$, so gilt $1 = \max M = \sup M$. Obwohl M kein Minimum hat, existiert $\inf M$ und es gilt

$$\inf M = 0.$$

(Denn: Zunächst ist 0 eine untere Schranke von M. Ist $s > 0$, also $s = p/q$ mit $p, q \in \mathbb{N}$, so ist $1/(q+1) < s$ und $1/(q+1) \in M$. Also ist s keine untere Schranke von M. Damit ist jede untere Schranke $s \leq 0$.)
2. Ist

$$M := \{x \in \mathbb{Q} : x \geq 0,\ x^2 \leq 2\} \ \bigl(= \{x \in \mathbb{Q} : x \geq 0,\ x^2 < 2\}\bigr) .$$

Dann ist M beschränkt, denn 0 ist eine untere Schranke und $3/2$ ist eine obere Schranke von M. (Ist $x > 3/2$, so folgt $x^2 > (3/2)^2 = 9/4 > 2$, also $x \notin M$.)
Hier ist $\inf M = \min M = 0$, es existiert aber *kein Supremum* von M, wie wir jetzt zeigen werden.

Bemerkung 4.12 Es sei K ein geordneter Körper. Dann gilt für alle $x > -1$ die **Bernoulli-Ungleichung**

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

([Ü]). Außerdem ist für $0 \le a \le b$ nach S. 3.1

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

Satz 4.13 *Es seien K ein geordneter Körper und $n \in \mathbb{N}$ sowie $0 \le c \in K$. Für*

$$M := \{x \in K : x \ge 0, x^n \le c\}.$$

gilt dann

1. *M ist nichtleer und nach oben beschränkt.*
2. *Existiert $s := \sup M$, so gilt $s^n = c$.*

Beweis. 1. Es gilt $0 \in M$, also $M \neq \emptyset$. Außerdem ist $1+c$ obere Schranke vom M, denn ist $x \in K$ mit $x > 1+c$, so gilt nach der Bernoullischen Ungleichung

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

und damit ist $x \notin M$.
2. Wir zeigen, dass weder $s^n > c$ noch $s^n < c$ gelten kann (damit ist $s^n = c$).
Angenommen, es ist $s^n > c$. Dann gilt für $\delta := \frac{s^n - c}{ns^{n-1}}$ nach B. 4.12

$$s^n - (s-\delta)^n \le n\delta s^{n-1} \le s^n - c$$

also $(s-\delta)^n \ge c$.
Ist $x \in M$, so folgt $x^n \le c \le (s-\delta)^n$ und damit auch $x \le s - \delta$. Also ist $s - \delta$ obere Schranke von M im Widerspruch dazu, dass s kleinste obere Schranke ist.
Angenommen, es ist $s^n < c$. Dann existiert ein $\delta > 0$ mit $(s+\delta)^n \le c$ (nach B. 4.12 mit $a = s + \delta$ und $b = s$ 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 4.14 Ein geordnete Menge $(X, <)$ heißt **(ordnungs-)vollständig**, falls jede nichtleere, nach oben beschränkte Teilmenge M von X ein Supremum hat.

Bemerkung und Definition 4.15 Es sei K ein vollständiger geordneter Körper. 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 s ergibt sich aus S. 4.13. Sind $s_1, s_2 \in K$ mit $0 \leq s_1 < s_2$ so ergibt sich auch $s_1^n < s_2^n$. Also hat die Gleichung $x^n = c$ höchstens eine Lösung.)
Wir setzen

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

Damit ergibt sich für $c, d \in K$ mit $c, d \geq 0$ sowie $n, m \in \mathbb{N}$ aus den entsprechenden Potenzgesetzen leicht ([Ü])

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

und für $0 \leq c < d$ auch $\sqrt[n]{c} < \sqrt[n]{d}$.

Von fundamentaler Bedendeutung für die Analysis ist das folgende Ergebnis.

Satz 4.16 *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 Einschränkungen von* $+, \cdot$ *und* $<$ *auf* $\mathbb{Q}$ *stimmen mit den entsprechenden Funktionen bzw. Relationen in* $\mathbb{Q}$ *überein).*

Bemerkung 4.17 Die Elemente von $\mathbb{R}$ heißen **reelle Zahlen**. Auf eine mögliche Konstruktion der reellen Zahlen und den Beweis zum obigen Satz werden wir im Anhang A eingehen. Als eine wichtige Folgerung aus der Vollständigkeit ergibt sich:

1. $\mathbb{N}$ ist unbeschränkt in $\mathbb{R}$ (archimedische Eigenschaft von $\mathbb{R}$).

2. Sind $x, y \in \mathbb{R}$ mit $x < y$, so existiert ein $r \in \mathbb{Q}$ mit $x < r < y$ (Dichtheit von $\mathbb{Q}$ in $\mathbb{R}$).

Bemerkung und Definition 4.18 Manchmal ist es praktisch und sinnvoll, die geordnete Menge $(\mathbb{R}, <)$ um zwei Punkte $+\infty$ (oder kurz ∞) und $-\infty$ so zu erweitern, dass definitionsgemäß $-\infty < x < \infty$ für alle $x \in \mathbb{R}$ ist. Für $M \subset \mathbb{R}$ ist damit $\sup M = \infty$, falls M nach oben unbeschränkt ist, und $\inf M = -\infty$, falls M nach unten unbeschänkt ist.
Eine nichtleere Menge $I \subset \mathbb{R}$ heißt **Intervall**, falls $x \in I$ für alle $x \in \mathbb{R}$ mit $\inf I < x < \sup I$ gilt. Für $a, b \in \mathbb{R} \cup \{\pm\infty\}$ setzen wir

$$\begin{aligned}
[a,b] &:= \{x \in \mathbb{R} : a \le x \le b\}, \text{ falls } -\infty < a \le b < \infty \\
(a,b) &:= \{x \in \mathbb{R} : a < x < b\}, \text{ falls } -\infty \le a < b \le \infty \\
[a,b) &:= \{x \in \mathbb{R} : a \le x < b\}, \text{ falls } -\infty < a < b \le \infty \\
(a,b] &:= \{x \in \mathbb{R} : a < x \le b\}, \text{ falls } -\infty \le a < b < \infty.
\end{aligned}$$

Jedes Intervall hat genau eine solche Form, wobei stets $a = \inf I$ und $b = \sup I$.

Wie wir oben gesehen haben, hat damit in $\mathbb{R}$ jede Gleichung $x^n = c$ für $n \in \mathbb{N}$ und $c \ge 0$ eine Lösung. Leider gilt dies nicht mehr im Falle $c < 0$ und n gerade (da $x^n \ge 0$ für gerades n und beliebiges $x \in \mathbb{R}$ nach S. 4.6.3). 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 4.19 Wir betrachten die abelsche Gruppe

$$(\mathbb{R}^2, +, (0,0)) = (\mathbb{R}^{\{1,2\}}, +, 0)$$

aus B. 2.12. Mit der dort allgemein definierten argumentweisen Multiplikation ist $\mathbb{R}^2$ zwar ein kommutativer Ring, aber nicht nullteilerfrei und damit insbesondere kein Körper. Wir definieren alternativ für $x = (s,t)$ und $y = (u,v)$ in $\mathbb{R}^2$

$$x \cdot y = (s,t) \cdot (u,v) := (su - tv, sv + tu).$$

Man rechnet leicht nach, dass damit $(\mathbb{R}^2, +, \cdot)$ ein Körper ist mit $1 = (1_\mathbb{R}, 0_\mathbb{R})$. Man schreibt dann meist $\mathbb{C}$ statt $\mathbb{R}^2$ und nennt die Elemente von $\mathbb{C}$ **komplexe Zahlen**. Traditionell verwendet man z oder w als Bezeichung für eine komplexe Zahl. Sind etwa $z = (3, -1)$ und $w = (1, 2)$, so ist

$$z \cdot w = (3, -1) \cdot (1, 2) = (3 - (-2), 6 - 1) = (5, 5)\ .$$

Für $z = (s,t) \neq 0$ gilt

$$\frac{1}{z} = \left(\frac{s}{s^2 + t^2}, \frac{-t}{s^2 + t^2} \right).$$

Bemerkung und Definition 4.20 Aus der Definition der Addition und der Multiplikation ergibt sich $(s,0)+(u,0) = (s+u,0)$ und $(s,0)(u,0) = (su,0)$, d. h. Addition und Multiplikation der komplexen Zahlen $(s,0)$ und $(u,0)$ entprechen der Addition und der Multiplikation von s und u in $\mathbb{R}$. Indem wir die komplexe Zahl $(s,0)$ mit der reellen s identifizieren, können wir den Körper $\mathbb{C}$ damit als Erweiterung des Körpers $\mathbb{R}$ auffassen. Wir schreiben dann auch kurz s statt $(s,0)$. Man nennt weiterhin

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

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

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

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

$$z = (s,t) = (s,0) + (0,1)(t,0) = s + it$$

schreiben. Diese Darstellung heißt **Normalform** (oder **kartesische Form**) von z. So gilt etwa

$$z = (3,-1) = 3 + i(-1)(= 3 - i).$$

Weiter nennen wir $\mathrm{Re}\, z := s$ **Realteil** von z und $\mathrm{Im}\, z := t$ **Imaginärteil** z.

Bemerkung 4.21 In $(\mathbb{C},+,\cdot)$ ist es nicht möglich, eine Ordnungsrelation $<$ (mit den Eigenschaften aus D. 4.5) zu definieren.
(Denn: Angenommen, doch. Dann wäre $1 > 0$ nach S. 4.6.3, also $-1 < 0$ nach S. 4.6.1. Für $z = i$ wäre mit S. 4.6.3 aber auch $0 < i^2 = -1$, also Widerspruch zu $(O1)$.)
Der Beweis zeigt, dass kein Körper, in dem die Gleichung $x^2 = -1$ eine Lösung hat, zu einem geordneten Körper gemacht werden kann.

Bemerkung und Definition 4.22 Es sei $z = s + it$ eine komplexe Zahl.

1. Die komplexe Zahl $\overline{z} := s - it$ heißt zu z **konjugiert komplex**.
2. Die Zahl $|z| := \sqrt{s^2 + t^2} \in [0,\infty)$ heißt **Betrag** von z.

Geometrisch entsteht $\overline{z}$ durch Spiegelung von z an der reellen Achse. Der Betrag $|z|$ beschreibt – nach dem Satz des Pythagoras – anschaulich die Länge der Strecke von 0 nach z in der euklidschen Ebene.
Für $z, w \in \mathbb{C}$ ergibt sich leicht

$$\overline{z+w} = \overline{z} + \overline{w}, \qquad \overline{zw} = \overline{z} \cdot \overline{w}, \qquad \overline{(\overline{z})} = z$$

sowie

$$\mathrm{Re}(z) = \frac{1}{2}(z + \overline{z}) \quad \text{und} \quad \mathrm{Im}(z) = \frac{1}{2i}(z - \overline{z}).$$

Satz 4.23 *Für $z, w \in \mathbb{C}$ 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* $1/z = \overline{z}/|z|^2$, *falls* $z \neq 0$,
4. $|zw| = |z|\,|w|$ *und* $|z+w|^2 = |z|^2 + 2\mathrm{Re}(z\overline{w}) + |w|^2$,
5. *(Dreiecksungleichung)* $|z \pm w| \leq |z| + |w|$.

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

$$|zw|^2 = (zw)(\overline{zw}) = (z\overline{z})(w\overline{w}) = |z|^2|w|^2 = (|z||w|)^2\,.$$

Durch Wurzelziehen folgt die erste Behauptung. Weiter gilt

$$|z+w|^2 = (z+w)(\overline{z}+\overline{w}) = z\overline{z} + z\overline{w} + w\overline{z} + w\overline{w} = |z|^2 + 2\,\mathrm{Re}\,(z\overline{w}) + |w|^2\,.$$

5. Nach 2. und 4. ist

$$|z+w|^2 \leq |z|^2 + 2|z\overline{w}| + |w|^2 = |z|^2 + 2|z||w| + |w|^2 = (|z|+|w|)^2$$

Durch Wurzelziehen folgt die Behauptung für $z+w$. Damit erhält man dann auch

$$|z-w| \leq |z| + |-w| = |z| + |w|\,.$$

□

Beispiel 4.24 Für $z = 3 - i$ gilt

$$|z| = \sqrt{9+1} = \sqrt{10}, \qquad \overline{z} = 3 - i(-1) = 3 + i$$

und

$$z\overline{z} = (3-i)(3+i) = 9 + 1 \left(= |z|^2\right).$$

Definition 4.25 In Verallgemeinerung von D. 3.3 setzen wir noch für $z \in \mathbb{C}$ und $\nu \in \mathbb{N}_0$

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

Die komplexe Zahl $\binom{z}{\nu}$ heißt **Binomialkoeffizient** z **über** ν.

5 Stetigkeit und Grenzwerte

Im Weiteren sei stets $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, also $\mathbb{K}$ der Körper der reellen oder der komplexen Zahlen.

Definition 5.1 Es sei $X \subset \mathbb{C}$.
1. Ist $a \in X$ und ist $f : X \to \mathbb{K}$, so heißt f **stetig an der Stelle** a, falls zu jedem $\varepsilon > 0$ ein $\delta = \delta_\varepsilon > 0$ existiert mit

$$|f(x) - f(a)| < \varepsilon$$

für alle ($x \in X$ mit) $|x - a| < \delta$. Weiter heißt f **stetig auf der Menge** $M \subset X$, falls f stetig an jeder Stelle $a \in M$ ist. Ist $M = X$, so heißt f kurz **stetig**.
2. Mit $C(X, \mathbb{K})$ bezeichen wir die Menge aller stetigen Funktionen $f : X \to \mathbb{K}$. Außerdem setzen wir $C(X) := C(X, \mathbb{C})$.

Bemerkung und Definition 5.2 Aus der Definition folgt sofort:
1. Die identische Abbildung $f = \mathrm{id}_{\mathbb{K}}$ ist stetig.
2. Konstante Funktionen $f : \mathbb{C} \to \mathbb{K}$ sind stetig.
3. Ist f stetig, so ist für jede Menge $M \subset X$ auch $f|_M$ stetig.
4. Ist $\rho > 0$ und

$$U_\rho(a) := U_{\rho,X}(a) := \{x \in X : |x - a| < \rho\},$$

so ist $f : X \to \mathbb{K}$ genau dann stetig an der Stelle a, wenn $f|_{U\rho(a)}$ stetig an a ist. Die Menge $U_\rho(a)$ heißt ρ-**Umgebung** von a (bezüglich X). Im Falle $X = \mathbb{R}$ ist $U_\rho(a)$ das Intervall $(a-\rho, a+\rho)$ und im Falle $X = \mathbb{C}$ ist $U_\rho(a)$ eine Kreisscheibe mit Mittelpunkt a und Radius ρ ([Ü]).

Wir wollen eine Charakterisierung der Stetigkeit herleiten, die auf dem zentralen Begriff des Grenzwertes beruht.

Definition 5.3 Es sei $X \subset \mathbb{C}$. Ein Punkt $a \in \mathbb{C}$ heißt **Häufungspunkt** von X, falls zu jedem $\varepsilon > 0$ ein $x \in X$ mit $0 < |x - a| < \varepsilon$ existiert (also $U_{\varepsilon,X}(a) \setminus \{a\}$ nichtleer ist). Wir schreiben X' für die Menge aller Häufungspunkte von X. Ist $a \in X$ und kein Häufungspunkt, so heißt a ein **isolierter Punkt** von X.

Beispiel 5.4 Für $X = \{1/k,\ k \in \mathbb{N}\}$ ist $0 \in X'$ (aber $0 \notin X$). Zudem ist jedes $a \in X$ ein isolierter Punkt von X.

Bemerkung und Definition 5.5 Es seien X eine Menge und $f : X \to \mathbb{K}$.
1. Ist $M \subset X$, so heißt f **beschränkt auf** M, falls die Menge $\{|f(x)| : x \in M\}$ (nach oben) beschränkt in $\mathbb{R}$ ist, also falls ein $M > 0$ existiert mit $|f(x)| \leq M$ für alle $x \in M$. Im Falle $M = X$ sagen wir kurz, f sei **beschränkt**.
2. Ist $X \subset \mathbb{C}$ und ist $a \in X'$, so heißt f **abklingend** an a, falls zu jedem $\varepsilon > 0$ ein $\delta = \delta_\varepsilon > 0$ existiert mit $|f(x)| < \varepsilon$ für alle ($x \in X$ mit) $0 < |x - a| < \delta$. Existiert eine Konstante $c \in \mathbb{K}$ so, dass $f - c$ abklingend an a ist, so heißt f **konvergent** an der Stelle a und c **Grenzwert** von f an der Stelle a. Wir schreiben dann

$$f(x) \to c \qquad (x \to a).$$

Man beachte dabei, dass auch im Falle $a \in X$, also im Falle, dass $f(a)$ existiert, der Funktionswert $f(a)$ beim Grenzwert keine Rolle spielt!
Aus $a \in X'$ folgt, dass höchstens ein Grenzwert c von f an a existiert ([Ü]). Wir schreiben im Falle der Existenz auch

$$\lim_{x \to a} f(x) := c\ .$$

Bemerkung 5.6 Es seien $X \subset \mathbb{C}$, $f, g : X \to \mathbb{K}$ und $a \in X'$.
1. Sind f und g abklingend an a, so sind auch $f \pm g$ abklingend an a.
2. Existiert ein $\rho > 0$ so, dass f auf $U_\rho(a)$ beschränkt ist, und ist g abklingend an a, so ist auch $f \cdot g$ abklingend an a.
(Denn:
1. Es sei $\varepsilon > 0$ gegeben. Dann existieren ein $\eta > 0$ und ein $0 < \delta \leq \eta$ mit $|f(x)| < \varepsilon/2$ für $0 < |x - a| < \eta$ und $|g(x)| < \varepsilon/2$ für $0 < |x - a| < \delta$. Also gilt für $0 < |x - a| < \delta$ nach der Dreiecksungleichung

$$|f(x) \pm g(x)| \leq |f(x)| + |g(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon\ .$$

2. Es seien $M, \rho > 0$ so, dass $|f(x)| \leq M$ für alle $x \in U_\rho(a)$. Ist $\varepsilon > 0$ gegeben, so existiert ein $0 < \delta \leq \rho$ mit $|g(x)| < \varepsilon/M$ für $0 < |x - a| < \delta$. Dann ist

$$|f(x)g(x)| \leq M|g(x)| < \varepsilon \quad (0 < |x - a| < \delta).)$$

Definition 5.7 Ist X eine Menge und ist $M \subset X$, so definieren wir die **Indikatorfunktion** $1_M = 1_{M,X} : X \to \mathbb{R}$ von M (bezüglich X) durch

$$1_M(x) := \begin{cases} 1, & \text{falls } x \in M \\ 0, & \text{falls } x \in X \setminus M \end{cases}.$$

Beispiel 5.8 Es seien $X = \mathbb{R}$ und $M := \{1/k : k \in \mathbb{N}\}$.
1. Für $f = 1_M$ gilt

$$f(x) \to 0 \quad (x \to a)$$

für alle $a \in \mathbb{R}^*$ (zu jedem $a \neq 0$ existiert ein $\rho > 0$ mit $f(x) = 0$ für alle $0 < |x - a| < \rho$). An der Stelle $a = 0$ existiert *kein* Grenzwert ([Ü]).
2. Für $f = 1_M \cdot \mathrm{id}_{\mathbb{R}}$, also

$$f(x) = \begin{cases} x, & \text{falls } x \in M \\ 0, & \text{falls } x \notin M \end{cases},$$

gilt

$$f(x) \to 0 \quad (x \to a)$$

für alle $a \in \mathbb{R}$ (für $a \neq 0$ kann man wie in 1. argumentieren und für $a = 0$ folgt die Behauptung etwa aus B. 5.6, da 1_M beschränkt ist).

Der folgende Satz zeigt, dass Grenzwertbildung mit den algebraischen Operationen in $\mathbb{C}$ verträglich ist. Wir definieren dazu in Ergänzung zu B. 2.12 für eine beliebige Menge X und eine nullstellenfreie Funktion $g : X \to \mathbb{K}$ (also $g(x) \neq 0$ für alle $x \in X$) die Funktion $1/g : X \to \mathbb{K}$ argumentweise durch

$$(1/g)(x) := 1/g(x) \qquad (x \in X)$$

und damit auch $f/g := f \cdot (1/g)$ für $f : X \to \mathbb{K}$.

Satz 5.9 *Es seien $X \subset \mathbb{C}$ und $f, g : X \to \mathbb{K}$. Weiter sei $a \in X'$ mit*

$$f(x) \to b \quad \text{und} \quad g(x) \to c \quad (x \to a).$$

Dann gilt

1. *$(f \pm g)(x) \to b \pm c$ $(x \to a)$.*
2. *$(f \cdot g)(x) \to b \cdot c$ $(x \to a)$.*
3. *Ist g nullstellenfrei und ist $c \neq 0$, so folgt $(f/g)(x) \to b/c$ $(x \to a)$.*

Beweis. 1. Die erste Aussage ergibt sich sofort aus B. 5.6.1.

2. Zunächst existiert ein $\rho > 0$ mit $|f(x) - b| < 1$ für $0 < |x - a| < \rho$ und somit auch

$$|f(x)| = |f(x) - b + b| \leq 1 + |b|.$$

Damit ist f beschränkt auf $U_\rho(a)$. Weiter ist

$$f(x)g(x) - bc = f(x)g(x) - cf(x) + cf(x) - bc = f(x)(g(x) - c) + c(f(x) - b).$$

Nach B. 5.6 ist die rechte Seite abklingend an a. Also folgt $(fg)(x) \to bc$ $(x \to a)$.
3. Nach 2. reicht es, zu zeigen:

$$(1/g)(x) \to 1/c \qquad (x \to a).$$

Da $c \neq 0$ ist, existiert ein $\rho > 0$ mit

$$|g(x) - c| < |c|/2 \qquad (0 < |x - a| < \rho)\ .$$

Also gilt mit der umgekehrten Dreiecksungleichung ([Ü])

$$|g(x)| = |c + g(x) - c| \geq |c| - |g(x) - c| > |c| - |c|/2 = |c|/2 > 0$$

und damit $|1/g(x)| \leq 2/|c|$ für $0 < |x - a| < \rho$. Folglich ist $1/g$ beschränkt auf $U_\rho(a)$. Nach B. 5.6 ist

$$\frac{1}{g(x)} - \frac{1}{c} = \frac{1}{g(x)}(1 - g(x)/c)$$

abklingend an a, also gilt $1/g(x) \to 1/c$ $(x \to a)$. □

Bemerkung 5.10 Ist unter den Voraussetzungen des vorherigen Satzes $\mathbb{K} = \mathbb{R}$ und $f \leq g$ (also $f(x) \leq g(x)$ für alle $x \in X$), so gilt $b \leq c$ ([Ü]). Im Allgemeinen folgt aus $f < g$ jedoch nicht $b < c$ (sondern eben nur $b \leq c$).

Bemerkung 5.11 Es seien $X \subset \mathbb{C}$, $a \in X'$ und $f : X \to \mathbb{K}$. Aus den obigen Definitionen ergibt sich unmittelbar: Es gilt $f(x) \to c$ $(x \to a)$ genau dann, wenn die Funktion $f_{a,c} : X \cup \{a\} \to \mathbb{C}$, definiert durch

$$f_{a,c}(x) := \begin{cases} c, & \text{falls } x = a \\ f(x), & \text{falls } x \neq a \end{cases},$$

stetig an a ist. Ist $a \in X$, so gilt damit auch:

f ist genau dann stetig an a, wenn $f(x) \to f(a)$ $(x \to a)$ gilt.

Aus der Definition der Stetigkeit ergibt zudem sofort: Ist a ein isolierter Punkt von X, so ist f stets stetig an a.

Mit S. 5.9 erhält man damit: Ist $X \subset \mathbb{C}$ und sind f, g stetig an $a \in X$, so sind $f \pm g$, $f \cdot g$ und für nullsetllenfreies g auch f/g stetig an a.

Beispiel 5.12 Ist $M = \{1/k : k \in \mathbb{N}\}$ und $f = 1_M$ wie in B. 5.8, so gilt für $k \in \mathbb{N}$

$$f(x) \to 0 \neq 1 = f(1/k) \quad (x \to 1/k)$$

und damit ist f nicht stetig an $1/k$. Da an 0 kein Grenzwert existiert, ist f auch nicht stetig an 0. Die Funktion $f = 1_M \mathrm{id}_\mathbb{R}$ aus B. 5.8 ist ebenfalls unstetig an allen Stellen $1/k$, aber stetig an der Stelle 0.

Definition 5.13 Es seien X eine Menge und $f : X \to \mathbb{K}$. Ist $f(x) = 0$, so heißt x eine **Nullstelle** von f. Mit $Z(f)$ bezeichnen wir die Menge der Nullstellen, also

$$Z(f) := \{x \in X : f(x) = 0\}.$$

Beispiel 5.14 Eine **Polynomfunktion** (oder kurz **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 $\deg(P) := d$ der **Grad** von P und $a_0, \ldots, a_d$ die **Koeffizienten** von P. Sind P, Q Polynome, so ist auch $P \cdot Q$ ein Polynom, und zwar von Grad $\deg(P) + \deg(Q)$. Aus B./D. 5.2.1 ergibt sich durch wiederholte Anwendung von B. 5.11, dass jedes Polynom stetig ist.

Sind P, Q Polynome, so ist zudem $P/Q : \mathbb{K} \setminus Z(Q) \to \mathbb{K}$ stetig, wieder nach B. 5.11. Funktionen der Form P/Q heißen **rational**.

Bemerkung 5.15 Es seien $U \subset \mathbb{C}$, $\alpha \in U'$ und $\varphi : U \to \mathbb{C}$ mit $\varphi(u) \to a$ $(u \to \alpha)$. Weiter seien $X \subset \mathbb{C}$ mit $\varphi(U) \subset X$ und $f : X \to \mathbb{K}$.

1. Ist $a \in X$ und ist f stetig an a, so gilt $(f \circ \varphi)(u) \to f(a)$ $(u \to \alpha)$. Ist zusätzlich $\alpha \in U$ und φ stetig an α, so ist auch $f \circ \varphi$ stetig an α.

2. Ist $a \in X' \setminus \varphi(U \setminus \{\alpha\})$ mit $f(x) \to c$ $(x \to a)$, so gilt $(f \circ \varphi)(u) \to c$ $(u \to \alpha)$.

(Denn: 1. Ist $\varepsilon > 0$ gegeben, so existiert ein $\eta > 0$ mit

$$|f(x) - f(a)| < \varepsilon \qquad (|x - a| < \eta).$$

Weiter existiert ein $\delta > 0$ so, dass

$$|\varphi(u) - a| < \eta \qquad (0 < |u - \alpha| < \delta).$$

Damit ist $|f(\varphi(u)) - f(a)| < \varepsilon$ für $0 < |u - \alpha| < \delta$.
2. Die 2. Behauptung ergibt sich aus 1. durch Anwendung auf $f_{a,c}$.)

Bemerkung und Definition 5.16 Es seien $X \subset \mathbb{C}$ und $f : X \to \mathbb{K}$. Ist $a \in \mathbb{C}$ und ist $M \subset X$ mit $a \in M'$, so schreiben wir

$$f(x) \to c \quad (x \to a,\, x \in M) \qquad \text{und} \qquad \lim_{M \ni x \to a} f(x) := c,$$

falls $f|_M(x) \to c$ $(x \to a)$ gilt. Sind dabei speziell $a \in \mathbb{R}$ und $M := X \cap (a, \infty)$, so sagt man, dass f an a den **rechtsseitigen Grenzwert** c hat und schreibt dann $f(x) \to c$ $(x \to a^+)$ sowie

$$f(a^+) := \lim_{x \to a^+} f(x) := c.$$

Entsprechend spricht man im Falle $M = X \cap (-\infty, a)$ vom **linksseitigen Grenzwert** c und schreibt dann $f(x) \to c$ $(x \to a^-)$ sowie

$$f(a^-) := \lim_{x \to a^-} f(x) := c.$$

Man sieht damit leicht: Ist $a \in \mathbb{R}$ Häufungspunkt von $X \cap (-\infty, a)$ und von $X \cap (a, \infty)$, so gilt $f(x) \to c$ $(x \to a)$ genau dann, wenn $f(a^+)$ und $f(a^-)$ existieren und

$$f(a^+) = f(a^-) = c$$

erfüllt ist. Existieren $f(a^+)$ und $f(a^-)$ mit

$$f(a^+) \neq f(a^-),$$

so heißt a **Sprungstelle** von f. An Sprungstellen hat f keinen (beidseitigen) Grenzwert.

Beispiel 5.17 1. Die **Vorzeichenfunktion** $\operatorname{sgn} : \mathbb{R} \to \mathbb{R}$ ist definiert durch

$$\operatorname{sgn}(x) := \begin{cases} 1, & \text{falls} \quad x > 0 \\ 0, & \text{falls} \quad x = 0 \\ -1, & \text{falls} \quad x < 0 \end{cases}.$$

Hier gilt

$$\operatorname{sgn}(0^+) = \lim_{x\to 0^+} \operatorname{sgn}(x) = 1, \qquad \operatorname{sgn}(0^-) = \lim_{x\to 0^-} \operatorname{sgn}(x) = -1.$$

Damit ist $a = 0$ eine Sprungstelle.
2. Die Indikatorfunktion $1_{\mathbb{Q}}$ von $\mathbb{Q}$ (bezüglich $\mathbb{R}$) heißt **Dirichlet-Funktion**. Die Dirichlet-Funktion hat für *kein* a in $\mathbb{R}$ einen rechts- oder linksseitigen Grenzwert!
(Denn: Ist $a \in \mathbb{R}$ 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 $a < x, y < a + \delta$. Hieraus folgt, dass kein rechtsseitiger Grenzwert an a existiert. Entsprechend sieht man, dass kein linksseitiger Grenzwert existiert.)
Insbesondere ist damit f unstetig an allen Stellen $a \in \mathbb{R}$.

Wie etwa die Funktion sgn zeigt, sind Bildmengen von Intervallen unter Funktionen mit Sprungstellen im Allgemeinen keine Intervalle. Anders ist die Situation bei stetigen Funktionen, wie der folgende, für die Analysis zentrale Satz zeigt. Der Beweis beruht ganz wesentlich auf der Vollständigkeit von $\mathbb{R}$.

Satz 5.18 *(Zwischenwertsatz)*
Es seien $X \subset \mathbb{R}$ und $f : X \to \mathbb{R}$ stetig. Ist $I \subset X$ ein Intervall, so ist auch $f(I)$ ein Intervall.

Beweis. Wir (müssen) zeigen: Ist $\eta \in \mathbb{R}$ mit $\inf f(I) < \eta < \sup f(I)$, so existiert ein $\xi \in I$ mit $f(\xi) = \eta$, also $\eta \in f(I)$.
Zunächst existieren nach Definition des Supremums und des Infimums $u, v \in f(I)$ mit $u < \eta < v$. Also existieren $\alpha, \beta \in I$ mit $f(\alpha) = u$ und $f(\beta) = v$. Ohne Einschränkung können wir $\alpha < \beta$ annehmen. Da I ein Intervall ist, folgt $[\alpha, \beta] \subset I$. Wir setzen

$$M := \{x \in [\alpha, \beta] : f(x) \le \eta\} .$$

Dann ist $M \ne \emptyset$ (da $\alpha \in M$) und beschränkt, also existiert aufgrund der Vollständigkeit von $\mathbb{R}$

$$\xi := \sup M \in \mathbb{R}.$$

Dabei gilt $\xi \in [\alpha, \beta]$, da β eine obere Schranke von M ist, also $\xi \in I$. Geanuer gilt $\xi \in M$, also $f(\xi) \le \eta$. (Denn: Es gilt jedenfalls $\xi \in M' \cup M$ ([Ü]). Ist $\xi \in M'$, so folgt $f(x) \to f(\xi)$ $(x \to \xi,\ x \in M)$, da f stetig an $\xi \in I$ ist. Aus $f(x) \le \eta$ für alle $x \in M$ folgt $f(\xi) \le \eta$ nach B. 5.10.)
Damit ergibt sich $\xi < \beta$ und mit $\eta < f(x) \to f(\xi)$ $(x \to \xi^+)$ auch $\eta \le f(\xi)$ nach B. 5.10, also insgesamt $f(\xi) = \eta$. □

Bemerkung 5.19 Für $n \in \mathbb{N}$ betrachten wir $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) := x^n$.
Ist $I = [0, \infty)$, so gilt $0 = f(0) \in f(I)$ und $\sup f(I) = \infty$ (ist $M \geq 1$ und $x > M$, so gilt $f(x) = x^n \geq x > M$). Da f stetig ist, folgt $f(I) \supset I$ aus dem Zwischenwertsatz (und damit $f(I) = I$). Also hat für jedes $c \geq 0$ (also $c \in I$) die Gleichung $x^n = c$ eine Lösung in I. Wir erhalten also aus dem Zwischenwertsatz noch einmal – und jetzt völlig schmerzlos – die Existenz n-ter Wurzeln.

Bemerkung und Definition 5.20 Es seien $X \subset \mathbb{R}$ und $f : X \to \mathbb{K}$.
1. Wir schreiben $\infty \in X'$, falls X nach oben unbeschränkt ist und $-\infty \in X'$ falls X nach unten unbeschränkt ist. Außerdem setzen wir

$$U_\rho(\pm\infty) := U_{\rho,X}(\pm\infty) := \{x \in X : \pm x > 1/\rho\}.$$

2. Ist $\infty \in X'$, so schreiben wir

$$f(x) \to c \quad (x \to \infty) \qquad \text{oder auch} \qquad f(\infty^-) := \lim_{x\to\infty} f(x) := c,$$

wenn zu jedem $\varepsilon > 0$ ein $\delta = \delta_\varepsilon > 0$ so existiert, dass $|f(x) - c| < \varepsilon$ für alle $x \in U_\delta(\infty)$. Dies ist genau dann der Fall, wenn $f(1/u) \to c$ $(u \to 0^+)$ gilt. Damit gelten die Aussagen von S. 5.9 auch hier. Entsprechend schreiben wir im Fall $-\infty \in X'$

$$f(x) \to c \quad (x \to -\infty) \qquad \text{oder auch} \qquad f(-\infty^+) := \lim_{x\to-\infty} f(x) := c,$$

wenn zu jedem $\varepsilon > 0$ ein $\delta = \delta_\varepsilon > 0$ so existiert, dass $|f(x)-c| < \varepsilon$ für alle $x \in U_\delta(-\infty)$. Dies ist genau dann der Fall, wenn $f(1/u) \to c$ $(u \to 0^-)$ gilt.

Beispiel 5.21 1. Für $n \in \mathbb{N}$ sei $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ definiert durch $f(x) = 1/x^n$. Aus $f(1/u) = u^n \to 0$ $(u \to 0)$ folgt

$$1/x^n = f(x) \to 0 \qquad (x \to \pm\infty).$$

2. Es seien $P(x) = \sum\limits_{\nu=0}^{d} a_\nu x^\nu$ und $Q(x) = \sum\limits_{\mu=0}^{d} b_\mu x^\mu$ Polynome mit $d = \deg(Q) > 0$ (und $\deg(P) \leq d$). Ist $f := P/Q : \mathbb{R} \setminus Z(Q) \to \mathbb{K}$, so gilt

$$f(x) \to a_d/b_d \qquad (x \to \pm\infty).$$

(Denn: Zunächst ist $Z(Q)$ endlich ([Ü]) und damit ist $\mathbb{R} \setminus Z(Q)$ jedenfalls nach oben und unten unbeschränkt. Da $b_d \neq 0$ ist, gilt für $x \notin Z(Q)$

$$\frac{P(x)}{Q(x)} = \frac{a_d + a_{d-1}x^{-1} + \cdots + a_0 x^{-d}}{b_d + b_{d-1}x^{-1} + \cdots + b_0 x^{-d}} \to \frac{a_d}{b_d} \qquad (x \to \pm\infty)$$

nach 1. und S. 5.9.)

Definition 5.22 Sind $X \subset \mathbb{R}$ und $f : X \to \mathbb{K}$ mit $f(X) \subset \mathbb{R}$, so 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. **(monoton) fallend** beziehungsweise **streng (monoton) fallend**, falls $-f$ wachsend beziehungsweise streng wachsend ist.

Ist f wachsend oder fallend, so sagen wir kurz, f sei **monoton**.

Die Dirichlet-Funktion zeigt, dass beschränkte Funktionen im Allgemeinen keine rechts- oder linksseitigen Grenzwerte haben. Der folgende Satz zeigt, dass *monotone* Funktionen stets rechts- und linksseitige Grenzwerte besitzen. Wieder basiert der (einfache) Beweis wesentlich auf der Existenz des Supremums und des Infimums beschränkter Mengen, also der Vollständigkeit von $\mathbb{R}$.

Satz 5.23 *Es seien $X \subset \mathbb{R}$ und $f : X \to \mathbb{R}$ beschränkt und monoton.*

1. *Ist $a \in (X \cap (a, \infty))'$, so existiert $f(a^+)$ und es gilt*
$$f(a^+) = \begin{cases} \inf f(X \cap (a, \infty)), & \text{falls } f \text{ wachsend ist} \\ \sup f(X \cap (a, \infty)), & \text{falls } f \text{ fallend ist} \end{cases}.$$
2. *Ist $b \in (X \cap (-\infty, b))'$, so existiert $f(b^-)$ und es gilt*
$$f(b^-) = \begin{cases} \sup f(X \cap (-\infty, b)), & \text{falls } f \text{ wachsend ist} \\ \inf f(X \cap (-\infty, b)), & \text{falls } f \text{ fallend ist} \end{cases}.$$

Beweis. Wir zeigen nur 1. Die Aussagen in 2. ergeben sich in analoger Weise. Weiter sei ohne Einschränkung f wachsend (ansonsten betrachte man $-f$).
Da f beschränkt ist, existiert
$$c := \inf f(X \cap (a, \infty)) \in \mathbb{R}.$$
Ist $\varepsilon > 0$ gegeben, so existiert ein $x_\varepsilon > a$ mit $f(x_\varepsilon) < c + \varepsilon$. Da f wachsend ist, gilt auch
$$c \leq f(x) \leq f(x_\varepsilon) < c + \varepsilon$$
für alle x mit $a < x < x_\varepsilon$. Damit ist $f(a^+) = c$. $\square$

Bemerkung und Definition 5.24 Es seien $X \subset \mathbb{C}$ und $a \in X'$ (im Fall $X \subset \mathbb{R}$ gegebenenfalls auch $a = \pm\infty$). Ist $f : X \to \mathbb{R}$ beschränkt, so sind Funktionen $g^* : (0,\infty) \to \mathbb{R}$ und $g_* : (0,\infty) \to \mathbb{R}$, definiert durch

$$g^*(r) := \sup f(U_r(a) \setminus \{a\}) \qquad \text{und} \qquad g_*(r) := \inf f(U_r(a) \setminus \{a\}),$$

beschränkt und monoton. Der nach S. 5.23 existierende Grenzwert

$$\limsup_{x\to a} f(x) := \overline{\lim_{x\to a}} f(x) := \lim_{r\to 0^+} g^*(r)$$

heißt **Limes superior** von f an a. Entprechend definiert man den **Limes inferior** von f an a mit g_* statt g^*. Man schreibt dafür $\liminf\limits_{x\to a} f(x)$ oder $\underline{\lim}_{x\to a} f(x)$.

Beispiel 5.25 Die Dirichlet-Funktion $f = 1_{\mathbb{Q}}$ ist beschränkt und damit existieren Limes superior und Limes inferior an allen Stellen $a \in \mathbb{R} \cup \{\pm\infty\}$. Genauer gilt hier

$$\limsup_{x\to a} f(x) = 1 \quad \text{und} \quad \liminf_{x\to a} f(x) = 0$$

für alle a (da g^* konstant $= 1$ und g_* konstant $= 0$ ist).

Definition 5.26 Es seien $X \subset \mathbb{C}$ und $a \in X'$ (wieder ist $a = \pm\infty$ für $X \subset \mathbb{R}$ zugelassen). Ist $f : X \to \mathbb{R}$, so schreiben wir

$$f(x) \to \pm\infty \qquad (x \to a),$$

falls für alle $\varepsilon > 0$ ein $\delta = \delta_\varepsilon > 0$ so existiert, dass $f(x) \in U_\varepsilon(\pm\infty)$ für alle $x \in U_\delta(a) \setminus \{a\}$ gilt. Unter den Bedingungen von B./D. 5.16 definiert man wie dort auch

$$f(x) \to \pm\infty \qquad (x \to a, x \in M).$$

6 Folgen und Reihen in $\mathbb{K}$

Viele Verfahren in der Mathematik beruhen auf der iterativen Anwendung einer Selbstabbildung $\varphi : X \to X$, wobei X eine nichtleere Menge bezeichnet. Ist $a_0 \in X$ (der **Startpunkt** der Iteration), so setzt man

$$a_{n+1} := \varphi(a_n) \quad (n \in \mathbb{N}_0). \tag{6.1}$$

Damit ist eine Funktion $(a_n)_{n\in\mathbb{N}_0} : \mathbb{N}_0 \to X$ rekursiv definiert.
Man interessiert sich typischerweise für das Verhalten von a_n für große n, wobei wir allgemein für eine nach oben unbeschränkte Menge $M \subset \mathbb{R}$ sagen, dass eine Eigenschaft für $x \in M$ **(genügend) groß** gilt, falls ein $\rho > 0$ so existiert, dass die Eigenschaft auf $U_{\rho,M}(\infty)$ (also für alle $x \in M$ mit $x > 1/\rho$) gilt.

Definition 6.1 Sind $N \subset \mathbb{N}_0$ unendlich (und damit nach oben unbeschränkt in $\mathbb{R}$) und X eine nichtleere Menge, so nennen wir eine Funktion $(a_n)_{n\in N} : N \to X$ eine **Folge** (in X). Die a_n heißen dann **Folgeglieder**. Im Falle $N = \{n \in \mathbb{N}_0 : n \geq n_0\}$ schreiben wir dabei auch $(a_n)_{n=n_0}^{\infty}$ oder $(a_n)_{n\geq n_0}$ oder kurz (a_n), wenn n_0 klar oder irrelevant ist. [1] Ist $J \subset N$ unendlich, so nennen wir $(a_n)_{n\in J}$ eine **Teilfolge** von $(a_n)_{n\in N}$.

In diesem Abschnitt betrachten wir ausschließlich $X = \mathbb{K}$, also Folgen reeller oder komplexer Zahlen. Da Folgen in $\mathbb{K}$ spezielle $\mathbb{K}$-wertige Funktionen (mit Definitionsbereich $N \subset \mathbb{R}$) sind, stehen sämtliche Begriffe und Ergebnisse des vorherigen Abschnitts zur Verfügung.

Bemerkung und Definition 6.2 Eine Folge $(a_n)_{n\in N}$ in $\mathbb{K}$ heißt **konvergent**, falls ein $c \in \mathbb{K}$ existiert mit

$$a_n \to c \quad (n \to \infty).$$

Wir schreiben dann auch $\lim\limits_{n\to\infty} a_n = c$. Ist dabei $c = 0$, so spricht man auch von einer **Nullfolge**. Eine Folge, die nicht konvergent ist, heißt **divergent**. Nach B./D. 5.20 ist $(a_n)_{n\in N}$ genau dann konvergent gegen c, wenn gilt: Für alle $\varepsilon > 0$ ist $|a_n - c| < \varepsilon$ für $n \in N$ groß. Außerdem folgt aus

$$|a_n| \leq |a_n - c| + |c| < 1 + |c|$$

[1] Wenn man möchte, kann man sich bei Fragen nach dem Verhalten von Folgen für große n stets auf die Indexmenge $\mathbb{N}_0$ zurückziehen. Ist nämlich $N \subset \mathbb{N}_0$ unendlich, so existiert genau eine streng wachsende Folge $(n_k)_{k=0}^{\infty}$ mit $N = \{n_k : k \in \mathbb{N}_0\}$, induktiv definiert durch $n_0 := \min N$ und $n_{k+1} := \min(N \setminus \{n_0, \ldots, n_k\})$. Das Verhalten von $(a_n)_{n\in N}$ für große n entspricht dann dem von $(a_{n_k})_{k=0}^{\infty}$ für große k.

für n groß, dass konvergente Folgen notwendig beschränkt sind.
Konvergiert eine Teilfolge $(a_n)_{n\in J}$ von $(a_n)_{n\in N}$ gegen c, so schreiben wir $a_n \to c$ $(n \to \infty,\ n \in J)$. Aus der Definition ergibt sich sofort: Ist eine Folge konvergent, so ist auch jede Teilfolge konvergent, und zwar mit gleichem Grenzwert.

Beispiel 6.3 Ein wichtige Familie von Folgen sind die **geometrischen Folgen** $(a_n) = (q^n)$ für $q \in \mathbb{K}$, die sich rekursiv durch (6.1) für $\varphi : \mathbb{K} \to \mathbb{K}$ mit $\varphi(x) = qx$ und $a_0 = 1$ ergeben. Es gilt

1. Für $|q| > 1$ ist (q^n) unbeschränkt.
2. Für $|q| < 1$ ist (q^n) eine Nullfolge, also $q^n \to 0$ $(n \to \infty)$.

(Denn: 1. Hier ist $|q| = 1 + \delta$ für ein $\delta > 0$. Mit der Bernoullischen Ungleichung gilt

$$|q^n| = |q|^n = (1+\delta)^n \geq 1 + n\delta > n\delta \qquad (n \in \mathbb{N}).$$

Also ist (q^n) unbeschränkt.
2. Für $q = 0$ ist die Behauptung klar. Es sei also $0 < |q| < 1$. Dann ist $1/|q| = 1 + \delta$ mit einem $\delta > 0$ und wie in 1.

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

Aus $1/n \to 0$ folgt $q^n \to 0$ $(n \to \infty)$.)

Bemerkung 6.4 (*Hauptsatz über monotone Folgen*)
Als Spezialfall von S. 5.23 erhält man ein einfache und wichtige hinreichende Bedingung für die Konvergenz von Folgen in $\mathbb{R}$:

Ist (a_n) monoton und beschränkt, so ist (a_n) konvergent.

Beispiel 6.5 (*Heron-Verfahren; Babylonisches Wurzelziehen*)
Es sei $c > 0$ gegeben und $\varphi : (0, \infty) \to (0, \infty)$ definiert durch

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

Dann gilt $\sqrt{c} = \sqrt{x \cdot c/x} \leq \varphi(x)$ (geometrisches Mittel $\leq$ arithmetisches Mittel; [Ü]).
Wir betrachten mit einem beliebigen Startwert $a_0 > 0$ die Folge (a_n) in $(0, \infty)$ mit

$$a_{n+1} := \varphi(a_n) = \frac{1}{2}\left(a_n + \frac{c}{a_n}\right) \qquad (n \in \mathbb{N}_0)\ .$$

Mithilfe des Hauptsatzes über monotone Folgen kann man zeigen, dass (a_n) konvergent ist mit

$$\lim_{n\to\infty} a_n = \sqrt{c}$$

([Ü]), d. h. die Folgeglieder a_n sind Approximationen (also Näherungen) für $\sqrt{c}$. Dabei sind im Falle $c \in \mathbb{Q}$ und $a_0 \in \mathbb{Q}$ die 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

$$\varepsilon_n = \frac{a_n - \sqrt{c}}{\sqrt{c}} = \frac{a_n}{\sqrt{c}} - 1 \geq 0 \quad (n \in \mathbb{N})$$

der relative Fehler. Dann gilt

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

also

$$\varepsilon_{n+1} = \frac{1}{2}\frac{\varepsilon_n^2}{1+\varepsilon_n} \leq \frac{1}{2}\varepsilon_n^2 .$$

Hat man nach n Schritten für a_n einen Fehler $\varepsilon_n \leq 10^{-m}$, so ist der Fehler ε_{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!

Wir betrachen nun allgemeine beschränkte Folgen.

Beispiel 6.6 Für $|q| = 1$ ist die geometrische Folge (q^n) beschränkt. Ist speziell $q = -1$, also $q^n = (-1)^n$, so ist die Teilfolge $((-1)^n)_{n\in 2\mathbb{N}_0}$ konstant mit Wert 1 (und damit konvergent gegen 1) und die Teilfolge $((-1)^n)_{n\in 2\mathbb{N}_0+1}$ konstant mit Wert -1 (und damit konvergent gegen -1). Es existieren also zwei Teilfolgen mit unterschiedlichen Grenzwerten. Insbesondere ist die Folge $((-1)^n)$ divergent.

Bemerkung 6.7 Ist $(a_n)_{n\in N}$ eine beschränkte Folge in $\mathbb{R}$, so existiert eine Teilfolge, die gegen $\limsup\limits_{n\to\infty} a_n$ konvergiert und eine Teilfolge, die gegen $\liminf\limits_{n\to\infty} a_n$ konvergiert.
(Denn: Nach B. 5.24 existiert $c := \limsup\limits_{n\to\infty} a_n$. Wir setzen $n_0 := \min N$ und definieren eine Folge $(n_k)_{k=0}^{\infty}$ in N induktiv: Sind $n_0, \ldots, n_k$ definiert, so ist nach Definition des Limes superior die Menge

$$N_k := \{n \in N : n > n_k,\ c - 1/k < a_n < c + 1/k\}$$

nichtleer ([Ü]). Damit setzen wir $n_{k+1} := \min N_k$. Für die so definierte Folge (n_k) gilt dann $|a_{n_k} - c| < 1/k \to 0$ $(k \to \infty)$. Ist $J := \{n_k : k \in \mathbb{N}_0\}$, so folgt $a_n \to c$ $(n \to \infty,\ n \in J)$. Die Behauptung für den Limes inferior ergibt sich analog.)

Als Konsequenz erhalten wir ein weiteres zentrales Ergebnis der Analysis:

Satz 6.8 *(Bolzano-Weierstraß)*
Jede beschränkte Folge $(a_n)_{n\in N}$ in $\mathbb{K}$ besitzt eine konvergente Teilfolge.

Beweis. 1. Es sei $\mathbb{K} = \mathbb{R}$. Dann wähle man eine Teilfolge wie in B. 6.7.
2. Es sei $\mathbb{K} = \mathbb{C}$. Ist

$$a_n = \alpha_n + i\beta_n$$

die Normalform von a_n, so sind die Folgen $(\alpha_n)_{n\in N}$ und $(\beta_n)_{n\in N}$ in $\mathbb{R}$ beschränkt (es gilt $|\alpha_n| \leq |a_n|$ und $|\beta_n| \leq |a_n|$). Nach 1. existieren eine Teilfolge $(\alpha_n)_{n\in I}$ von $(\alpha_n)_{n\in N}$ und ein $\alpha \in \mathbb{R}$ mit $\alpha_n \to \alpha$ $(n \to \infty, n \in I)$. Wieder nach 1. existieren auch eine Teilfolge $(\beta_n)_{n\in J}$ von $(\beta_n)_{n\in I}$ und ein $\beta \in \mathbb{R}$ mit $\beta_n \to \beta$ $(n \to \infty, n \in J)$. Mit S. 5.9 folgt $a_n = \alpha_n + i\beta_n \to \alpha + i\beta$ $(n \to \infty, n \in J)$. □

Bemerkung und Definition 6.9 Es sei $(a_n)_{n\geq m}$ eine Folge in $\mathbb{K}$. Die Folge $(s_n)_{n\geq m}$ der **Partial-** oder **Teilsummen**

$$s_n := \sum_{\nu=m}^{n} a_\nu =: a_m + \cdots + a_n \qquad (n \geq m)$$

heißt die (**mit** (a_n) **gebildete**) **Reihe**. Die a_ν heißen dann **Reihenglieder**.
Ist die Folge (s_n) konvergent, so heißt $\lim\limits_{n\to\infty} s_n$ der **Reihenwert** und man schreibt

$$\sum_{\nu=m}^{\infty} a_\nu := \lim_{n\to\infty} s_n.$$

Traditionell wird neben dem Reihenwert auch die Teilsummenfolge (s_n) mit $\sum\limits_{\nu=m}^{\infty} a_\nu$ bezeichnet. Das ist ganz praktisch, weil man dann kurz von Konvergenz oder Divergenz von $\sum\limits_{\nu=m}^{\infty} a_\nu$ sprechen kann. Man beachte aber, dass das Symbol $\sum\limits_{\nu=m}^{\infty} a_\nu$ damit zwei Bedeutungen hat: Erstens steht es für die Folge (s_n) der Teilsummen und zweitens (im Falle der Konvergenz!) für deren Grenzwert.

Beispiel 6.10 (**geometrische Reihen**) Es sei $a_n = q^n$ für ein $q \in \mathbb{K}$, $|q| < 1$. Dann ist $\sum_{\nu=0}^{\infty} q^\nu$ konvergent mit

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

Für $q = 1/2$ ergibt sich etwa $\sum_{\nu=0}^{\infty} 1/2^\nu = 2$.

Bemerkung 6.11 Durch Anwendung von S. 5.9 ergibt sich leicht: Sind $\sum_{\nu=m}^{\infty} a_\nu$ und $\sum_{\nu=m}^{\infty} b_\nu$ konvergente Reihen in $\mathbb{K}$ und ist $\lambda \in \mathbb{K}$, so sind auch $\sum_{\nu=m}^{\infty} (a_\nu + b_\nu)$ und $\sum_{\nu=m}^{\infty} \lambda a_\nu$ konvergent mit

$$\sum_{\nu=m}^{\infty} (a_\nu + b_\nu) = \sum_{\nu=m}^{\infty} a_\nu + \sum_{\nu=m}^{\infty} b_\nu \quad \text{und} \quad \sum_{\nu=m}^{\infty} \lambda a_\nu = \lambda \sum_{\nu=m}^{\infty} a_\nu .$$

Beispiel 6.12 Mit B. 6.10 und B. 6.11 ist

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

Insbesondere erhält man aus B. 6.11 auch: Ist $k > m$, so ist $\sum_{\nu=k}^{\infty} a_\nu$ genau dann konvergent, wenn $\sum_{\nu=m}^{\infty} a_\nu$ konvergiert, und in diesem Fall ist

$$\sum_{\nu=m}^{\infty} a_\nu = \sum_{\nu=m}^{k-1} a_\nu + \sum_{\nu=k}^{\infty} a_\nu \,.$$

Für Konvergenzuntersuchungen ist es also unwichtig, wie die untere Summationsgrenze aussieht.

Bemerkung 6.13 Eine *notwendige Bedingung* für die Konvergenz einer Reihe ist, dass die Reihenglieder eine Nullfolge bilden, d. h. ist $\sum_{\nu=m}^{\infty} a_\nu$ konvergent, so gilt

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

(Denn: Ist $s_n = \sum_{\nu=m}^{n} a_\nu$, so gilt $a_n = s_n - s_{n-1} \to 0$ $(n \to \infty)$.)

Beispiel 6.14 1. Ist $a_n = q^n$ mit $|q| \geq 1$, so ist $|a_n| \geq 1$ $(n \in \mathbb{N})$, also ist $\sum\limits_{\nu=0}^{\infty} q^\nu$ sicher divergent. Damit ergibt sich für geometrische Reihen insgesamt: $\sum\limits_{\nu=0}^{\infty} q^\nu$ ist *genau dann* konvergent, wenn $|q| < 1$ ist.
2. Wir betrachten die **harmonische Reihe** $\sum\limits_{\nu=1}^{\infty} 1/\nu$. Hier gilt $a_n = 1/n \to 0$ $(n \to \infty)$, aber

$$s_{2^k} = \sum_{\nu=1}^{2^k} \frac{1}{\nu} = 1 + \sum_{\ell=1}^{k} \sum_{\nu=2^{\ell-1}+1}^{2^\ell} \frac{1}{\nu} \geq 1 + \sum_{\ell=1}^{k} 2^{\ell-1}\frac{1}{2^\ell} = 1 + \frac{k}{2} \to \infty \quad (k \to \infty).$$

Also ist (s_n) unbeschränkt und damit ist die harmonische Reihe divergent. Das Beispiel zeigt, dass die notwendige Bedingung $a_n \to 0$ $(n \to \infty)$ aus B. 6.13 im Allgemeinen nicht hinreichend für die Konvergenz der mit a_n gebildeten Reihe ist.

Im Falle von Reihen mit nichtnegativen Gliedern, wie etwa der harmonischen Reihe, können nur zwei wesentlich unterschiedliche Situationen auftreten.

Bemerkung 6.15 Ist (a_n) eine Folge in $[0, \infty)$, also $a_n \geq 0$ für alle n, so ist die Teilsummenfolge (s_n) wachsend. Damit ist

- entweder (s_n) beschränkt und dann konvergent nach S. 5.23 mit

$$\sum_{\nu=m}^{\infty} a_\nu = \sup\{s_n : n \in \mathbb{N}\} < \infty$$

- oder (s_n) unbeschränkt mit $s_n \to \infty$ $(n \to \infty)$.

Wir schreiben im zweiten Fall auch $\sum\limits_{\nu=m}^{\infty} a_\nu = \infty$.

Bemerkung und Definition 6.16 (*Majorantenkriterium*)
Es seien $(a_n)_{n \geq m}$ und $(b_n)_{n \geq k}$ Folgen mit

$$0 \leq a_n \leq b_n$$

für n groß. Man nennt dann $\sum\limits_{\nu=k}^{\infty} b_\nu$ eine **Majorante** (von $\sum\limits_{\nu=m}^{\infty} a_\nu$). Ist $\sum\limits_{\nu=k}^{\infty} b_\nu$ konvergent, so ist auch $\sum\limits_{\nu=m}^{\infty} a_\nu$ konvergent.

(Denn: Es sei $n_0 \in \mathbb{N}$ mit $0 \le a_n \le b_n$ für $n \ge n_0$. Aus

$$\sum_{\nu=n_0}^{n} a_\nu \le \sum_{\nu=n_0}^{n} b_\nu \le \sum_{\nu=n_0}^{\infty} b_\nu \qquad (n \ge n_0)$$

folgt die Beschränktheit der Teilsummen $s_n = \sum_{\nu=n_0}^{n} a_\nu$. Nach B. 6.15 ist $\sum_{\nu=n_0}^{\infty} a_\nu$ konvergent und damit auch $\sum_{\nu=m}^{\infty} a_\nu$.)

Beispiel 6.17 (allgemeine harmonische Reihen) Es sei $d \in \mathbb{N}$. Dann gilt

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

(Denn: Für $d = 1$ ergibt sich die Behauptung aus B. 6.14. Für $d > 1$ ist

$$\frac{1}{\nu^d} \le \frac{1}{\nu^2} \le \frac{1}{(\nu-1)\nu} =: b_\nu .$$

Weiter ist $\sum_{\nu=2}^{n} b_\nu = 1 - 1/n$ ([Ü]), also $\sum_{\nu=2}^{\infty} b_\nu = 1$. Damit ist $\sum_{\nu=2}^{\infty} b_\nu$ eine konvergente Majorante. Mit B./D. 6.16 folgt die Behauptung.)

Wählt man als spezielle Majorante eine geometrische Reihe, so erhält man weitere Konvergenzkriterien:

Satz 6.18 *(Wurzelkriterium/Quotientenkriterium)*
Es sei $(a_n)_{n \ge m}$ eine Folge in $[0, \infty)$. Dann ist $\sum_{\nu=m}^{\infty} a_\nu$ konvergent, falls eine der folgenden Bedingungen erfüllt ist

1. *Es existiert ein $q < 1$ mit $\sqrt[n]{a_n} \le q$ für n groß.*

2. *Es existiert ein $q < 1$ mit $a_n > 0$ und $a_{n+1}/a_n \le q$ für n groß.*

Beweis. 1. Nach Voraussetzung ist $0 \le a_n \le q^n$ für n groß. Aus der Konvergenz der geometrische Reihe $\sum_{\nu=0}^{\infty} q^\nu$ folgt die Konvergenz von $\sum_{\nu=m}^{\infty} a_\nu$ mit B./D. 6.16.
2. Es sei $n_0 \in \mathbb{N}$ so, dass $a_n > 0$ und $a_{n+1}/a_n \le q$ für $n \ge n_0$. Induktiv ergibt sich mit $\lambda := a_{n_0} q^{-n_0}$

$$a_n \le q^{n-n_0} a_{n_0} = \lambda q^n \qquad (n \ge n_0).$$

Wieder folgt die Behauptung aus der Konvergenz von $\sum_{\nu=0}^{\infty} q^\nu$ und B./D. 6.16. □

Beispiel 6.19 Es seien $d \in \mathbb{N}$ und $0 < r < 1$. Dann ist $\sum_{\nu=1}^{\infty} \nu^d r^\nu$ konvergent.
(Denn: Aus

$$\frac{(n+1)^d r^{n+1}}{n^d r^n} = r\left(\frac{n+1}{n}\right)^d \to r < 1 \quad (n \to \infty)$$

folgt, dass die Bedingung 2. aus S. 6.18 für jedes $q \in (r, 1)$ erfüllt ist.)
Insbesondere folgt hieraus mit B. 6.13, dass $(n^d r^n)$ eine Nullfolge ist. Die Konvergenz der Reihe zeigt, dass die Folge (r^n) tatsächlich viel schneller abklingt, als die Folge (n^d) wächst. Eine Erkenntnis, die von fundamentaler Bedeutung für die gesamte Mathematik ist.

Wir betrachten zum Abschluss Reihen der Form $\sum_{\nu=m}^{\infty} (-1)^\nu a_\nu$ mit $a_n \geq 0$. Man nennt solche Reihen **alternierend**.

Satz 6.20 *Es sei $(a_n)_{n\geq 0}$ eine fallende Folge in $[0, \infty)$.*

1. *Für $s_n := \sum_{\nu=0}^{n} (-1)^\nu a_\nu$ gilt $s_n \geq 0$ mit fallender Teilfolge $(s_n)_{n\in 2\mathbb{N}_0}$ und wachsender Teilfolge $(s_n)_{n\in 2\mathbb{N}_0+1}$.*
2. *(Leibniz-Kriterium) Ist (a_n) eine Nullfolge, so ist $\sum_{\nu=0}^{\infty} (-1)^\nu a_\nu$ konvergent.*

Beweis. 1. Für ungerades $n \in \mathbb{N}$ ist

$$s_n = \sum_{\nu=0}^{n} (-1)^\nu a_\nu = (a_0 - a_1) + (a_2 - a_3) + \cdots + (a_{n-1} - a_n) \geq 0$$

und damit auch $s_{n+1} = s_n + a_{n+1} \geq 0$. Weiter gilt für $n \geq 2$

$$s_n - s_{n-2} = (-1)^{n-1}(a_{n-1} - a_n) = \begin{cases} \leq 0, & \text{falls } n \text{ gerade} \\ \geq 0, & \text{falls } n \text{ ungerade} \end{cases}.$$

Damit ist $(s_n)_{n\in 2\mathbb{N}_0}$ fallend und $(s_n)_{n\in 2\mathbb{N}_0+1}$ wachsend.
2. Nach 1. und dem Hauptsatz über monotone Folgen existiert ein $s \in [0, \infty)$ mit $s_n \to s$ $(n \to \infty,\, n \in 2\mathbb{N}_0)$. Dann gilt auch $s_{n+1} = s_n - a_{n+1} \to s$ $(n \to \infty,\, n \in 2\mathbb{N}_0)$. Zusammen ergibt sich $s_n \to s$ für $n \to \infty$. □

Beispiel 6.21 (alternierende harmonische Reihe)
Während die harmonische Reihe $\sum\limits_{\nu=1}^{\infty} 1/\nu$ (nach B. 6.14) divergiert, konvergiert $\sum\limits_{\nu=1}^{\infty} (-1)^\nu/\nu$ nach S. 6.20. Die abwechselnden Vorzeichen bewirken eine Art Stabilisierung der Teilsummen.

7 Cauchy-Kriterium und elementare Funktionen

Definition 7.1 Eine Folge $(a_n)_{n\in N}$ in $\mathbb{K}$ heißt **Cauchyfolge**, falls zu jedem $\varepsilon > 0$ ein $R = R_\varepsilon > 0$ so existiert, dass

$$|a_n - a_{n'}| < \varepsilon$$

für alle ($n \in N$ mit) $n, n' > R$, also für $n, n' \in N$ genügend groß.

Bemerkung 7.2 Aus der Definition ergibt sich leicht:

1. Jede Cauchyfolge ist beschränkt.

2. Eine Cauchyfolge, die eine konvergente Teilfolge hat, ist konvergent.

(Denn: Es sei $(a_n)_{n\in N}$ eine Cauchyfolge.
1. Zu $\varepsilon = 1$ existiert ein $n_0 \in N$ so, dass $|a_n - a_{n'}| < 1$ für alle $n, n' \in N$ mit $n, n' \in N$, also auch

$$|a_n| = |a_n - a_{n_0} + a_{n_0}| \leq |a_n - a_{n_0}| + |a_{n_0}| < |a_{n_0}| + 1$$

für alle $n \in N$ mit $n \geq n_0$. Damit ist

$$|a_n| \leq \max\{|a_n| + 1 : n \in N,\, n \leq n_0\} \quad (n \in N).$$

2. Es sei $(a_n)_{n\in J}$ eine Teilfolge mit $a_n \to c$ ($n \to \infty, n \in J$). Wir zeigen: $a_n \to c$ ($n \to \infty$). Dazu sei $\varepsilon > 0$ gegeben. Dann existiert ein $R > 0$ mit

$$|a_n - a_{n'}| < \varepsilon/2 \qquad (n, n' > R)\ .$$

Weiter existiert ein $j \in J$ so, dass $j > R$ und $|a_j - c| < \varepsilon/2$. Damit ist

$$|a_n - c| \leq |a_n - a_j| + |a_j - c| < \varepsilon \qquad (n > R)\ .)$$

Ein weiterer zentraler Baustein der Analysis ist

Satz 7.3 *(Cauchy-Kriterium für Folgen)*
Eine Folge $(a_n)_{n\in N}$ in $\mathbb{K}$ ist genau dann konvergent, wenn sie eine Cauchyfolge ist.

Beweis. $\Rightarrow$: Es sei $c \in \mathbb{K}$ mit $a_n \to c$ ($n \to \infty$). Dann existiert zu jedem $\varepsilon > 0$ ein $R > 0$ mit $|a_n - c| < \varepsilon/2$ ($n > R$). Also gilt

$$|a_n - a_{n'}| \leq |a_n - a| + |a - a_{n'}| < \varepsilon \qquad (n, n' > R).$$

$\Leftarrow$: Es sei $(a_n)_{n\in N}$ eine Cauchyfolge. Dann ist $(a_n)_{n\in N}$ beschränkt nach B. 7.2.1. Also hat $(a_n)_{n\in N}$ nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge. Nach B. 7.2.2 ist $(a_n)_{n\in N}$ konvergent. □

Wir ziehen erste Folgerungen aus dem Cauchy-Kriterium.

Bemerkung 7.4 Es seien $X \subset \mathbb{C}$ und $a \in X'$. Ist $f : X \to \mathbb{K}$, so sind folgende Aussagen äquivalent:

a) f hat einen Grenzwert an der Stelle a.

b) Für alle Folgen (x_n) in $X \setminus \{a\}$ mit $x_n \to a$ ist die Bildfolge $(f(x_n))$ konvergent.

c) Für alle Folgen (x_n) in $X \setminus \{a\}$ mit $x_n \to a$ ist die Bildfolge $(f(x_n))$ eine Cauchyfolge.

(Denn: Die Äquivalenz von b) und c) ergibt sich aus dem Cauchy-Kriterium. Die Implikation a) $\Rightarrow$ b) folgt aus B. 5.15 mit (x_n) anstelle von φ.
b) $\Rightarrow$ a): Es sei (x_n) eine Folge in $X \setminus \{a\}$ mit $x_n \to a$. Dann existiert ein $c \in \mathbb{K}$ mit $f(x_n) \to c$ $(n \to \infty)$. Angenommen, es gilt nicht $f(x) \to c$ $(x \to a)$. Dann existieren ein $\varepsilon > 0$ und eine Folge (y_n) in $X \setminus \{a\}$ mit $y_n \to a$ $(n \to \infty)$ und $|f(y_n) - c| \geq \varepsilon$ für alle n. Für die Folge (z_n) in $X \setminus \{a\}$ mit

$$z_n := \begin{cases} x_n, & \text{falls } n \text{ gerade} \\ y_n, & \text{falls } n \text{ ungerade} \end{cases}$$

gilt $z_n \to a$ $(n \to \infty)$, aber die Bildfolge $(f(z_n))$ ist nicht konvergent, da die Teilfolge $(f(z_n))_{n\in 2\mathbb{N}_0}$ gegen c konvergiert, die Teilfolge $(f(z_n))_{n\in 2\mathbb{N}_0+1}$ aber nicht. Dies widerspricht der Voraussetzung b). Also gilt $f(x) \to c$ $(x \to a)$.)

Satz 7.5 *(Cauchy-Kriterium für Reihen)*
Es sei $(a_n)_{n\geq m}$ eine Folge in $\mathbb{K}$. Dann ist $\sum\limits_{\nu=m}^{\infty} a_\nu$ genau dann konvergent, wenn für alle $\varepsilon > 0$ ein $R > 0$ so existiert, dass

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

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

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

Damit ergibt sich die Behauptung sofort aus dem Cauchy-Kriterium für Folgen. □

Satz 7.6 *Es sei $(a_n)_{n\geq m}$ eine Folge in $\mathbb{K}$. Ist $\sum_{\nu=m}^{\infty} |a_\nu|$ konvergent, so ist auch $\sum_{\nu=m}^{\infty} a_\nu$ konvergent.*

Beweis. Ist $\varepsilon > 0$ gegeben, so existiert nach S. 7.5 ein $R > 0$ so, dass

$$\sum_{\nu=n'+1}^{n} |a_\nu| < \varepsilon \qquad (n > n' > R).$$

Aus der Dreiecksungleichung folgt

$$\Big| \sum_{\nu=n'+1}^{n} a_\nu \Big| \leq \sum_{\nu=n'+1}^{n} |a_\nu| < \varepsilon \qquad (n > n' > R).$$

Wieder nach S. 7.5 ist $\sum_{\nu=m}^{\infty} a_\nu$ konvergent. □

Bemerkung und Definition 7.7 Es sei $(a_n)_{n\geq m}$ eine Folge in $\mathbb{K}$. Die Reihe $\sum_{\nu=m}^{\infty} a_\nu$ heißt **absolut konvergent**, falls $\sum_{\nu=m}^{\infty} |a_\nu|$ konvergiert. Nach S. 7.6 ist jede absolut konvergente Reihe auch konvergent. Außerdem gilt dann ([Ü])

$$\Big| \sum_{\nu=m}^{\infty} a_\nu \Big| \leq \sum_{\nu=m}^{\infty} |a_\nu|.$$

Ist $\sum_{\nu=m}^{\infty} a_\nu$ konvergent und ist $\sum_{\nu=m}^{\infty} |a_\nu|$ divergent, so heißt $\sum_{\nu=m}^{\infty} a_\nu$ **bedingt konvergent**.

Beispiel 7.8 1. Für $|q| < 1$ ist die geometrische Reihe $\sum_{\nu=0}^{\infty} q^\nu$ absolut konvergent (da $\sum_{\nu=0}^{\infty} |q|^\nu$ konvergiert).

2. Es sei $d \in \mathbb{N}$. Die Reihe $\sum_{\nu=1}^{\infty} (-1)^\nu/\nu^d$ ist für $d = 1$ bedingt konvergent und für $d \geq 2$ absolut konvergent (B. 6.21 und B. 6.17).

Bemerkung und Definition 7.9 Es seien $z \in \mathbb{C}$ und $a_n := z^n/n!$ für $n \in \mathbb{N}_0$. Dann gilt für $z \neq 0$

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

Damit ist die Reihe $\sum\limits_{\nu=0}^{\infty} z^\nu/\nu!$ nach dem Quotientenkriterium für $z \neq 0$ absolut konvergent. Für $z = 0$ ist die Reihe auch konvergent (alle Teilsummen sind $= 1$).
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**. Nach Definition ist $\exp(0) = 1$ und zudem $\exp(\mathbb{R}) \subset \mathbb{R}$. Allgemeiner gilt für alle $z \in \mathbb{C}$ ([Ü])

$$\exp(\overline{z}) = \overline{\exp(z)}.$$

Wir wollen Eigenschaften der Exponentialfunktion herleiten, die von fundamentaler Bedeutung für die Mathematik sind.

Satz 7.10 *Für $z, w \in \mathbb{C}$ gilt*

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

und

$$\exp(-z) = 1/\exp(z).$$

Beweis. 1. Für $n \in \mathbb{N}$ setzen wir $L_n := \{(\mu, \nu) \in \{0, \ldots, n\}^2 : \mu + \nu \leq n\}$ und

$$J_n := \{(\mu, \nu) \in \{0, \ldots, n\}^2 : \mu + \nu > n\}.$$

Dann gilt für $z, w \in \mathbb{C}$

$$\left(\sum_{\mu=0}^{n} \frac{z^\mu}{\mu!}\right)\left(\sum_{\nu=0}^{n} \frac{w^\nu}{\nu!}\right) = \sum_{(\mu,\nu)\in L_n} \frac{z^\mu w^\nu}{\mu!\nu!} + \sum_{(\mu,\nu)\in J_n} \frac{z^\mu w^\nu}{\mu!\nu!} =: s_n + \varepsilon_n\,.$$

Dabei konvergiert die linke Seite für $n \to \infty$ gegen $\exp(z)\exp(w)$ und außerdem gilt mit der binomischen Formel

$$s_n = \sum_{k=0}^{n}\sum_{\mu=0}^{k} \frac{z^\mu w^{k-\mu}}{\mu!(k-\mu)!} = \sum_{k=0}^{n} \frac{1}{k!}\,(z+w)^k \to \exp(z+w) \quad (n \to \infty).$$

Daher reicht es zu zeigen, dass eine Teilfolge von (ε_n) eine Nullfolge ist. Wir setzen dazu $r := \max\{|z|, |w|, 1\}$. Aus $\#(J_n) = n(n-1)/2$ folgt für $n = 2m \in 2\mathbb{N}$ mit $\max\{\mu, \nu\} \geq m + 1$ für $(\mu, \nu) \in J_{2m}$

$$\begin{aligned} |\varepsilon_{2m}| &\leq \frac{1}{(m+1)!} \sum_{(\mu,\nu)\in J_{2m}} r^{\mu+\nu} \leq \frac{r^{4m}}{(m+1)!} m(2m-1) \\ &\leq 2r^4 \frac{(r^4)^{m-1}}{(m-1)!} \to 0 \qquad (m \to \infty). \end{aligned}$$

Damit gilt $\varepsilon_n \to 0$ $(n \to \infty,\ n \in 2\mathbb{N})$.
2. Aus 1. folgt $1 = \exp(0) = \exp(z - z) = \exp(z) \cdot \exp(-z)$. □

Bemerkung und Definition 7.11 Aus S. 7.10 folgt

$$\exp(mz) = (\exp(z))^m \qquad (z \in \mathbb{C},\ m \in \mathbb{Z}).$$

Die Zahl $e := \exp(1)$ heißt **Eulersche Zahl**. Es gilt damit $\exp(m) = e^m$ für alle $m \in \mathbb{Z}$ und deshalb ist die Schreibweise e^z statt $\exp(z)$ für allgemeines $z \in \mathbb{C}$ konsistent mit der Definition ganzzahliger Potenzen. Wir werden diese im Weiteren meist verwenden.

Satz 7.12 exp *ist stetig.*

Beweis. Für $|h| \leq 1$ gilt mit B. 7.7

$$|e^h - 1| = \Big|\sum_{\nu=1}^{\infty} \frac{h^\nu}{\nu!}\Big| \leq \sum_{\nu=1}^{\infty} \frac{|h|^\nu}{\nu!} = |h| \sum_{\nu=1}^{\infty} \frac{|h|^{\nu-1}}{\nu!} \leq |h| \sum_{\nu=1}^{\infty} \frac{1}{\nu!} = |h|(e-1).$$

Hieraus ergibt sich $e^h - 1 \to 0$ $(h \to 0)$. Für $a, h \in \mathbb{C}$ folgt

$$e^{h+a} - e^a = e^a \cdot (e^h - 1) \to 0 \quad (h \to 0),$$

also $e^{h+a} \to e^a$ $(h \to 0)$ und damit $e^z \to e^a$ $(z \to a)$. □

Definition 7.13 Die Funktion $\cos : \mathbb{C} \to \mathbb{C}$, gegeben durch

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

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

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

heißt **(komplexe) Sinusfunktion**. Damit gilt für alle $z \in \mathbb{C}$ die **Eulersche Formel**

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

Bemerkung 7.14 Mit B. 5.11 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])

$$\cos(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} z^{2k}$$

und

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

mit absoluter Konvergenz der Reihen. Hieraus folgt insbesondere $\sin(\mathbb{R}) \subset \mathbb{R}$ und $\cos(\mathbb{R}) \subset \mathbb{R}$. Genauer ergibt sich für $t \in \mathbb{R}$ mit $e^{-it} = \overline{e^{it}}$

$$\cos t = \frac{1}{2}(e^{it} + \overline{e^{it}}) = \mathrm{Re}(e^{it}), \qquad \sin t = \frac{1}{2i}(e^{it} - \overline{e^{it}}) = \mathrm{Im}(e^{it})$$

und

$$\cos^2 t + \sin^2 t = |e^{it}|^2 = e^{it}e^{-it} = 1.$$

Satz 7.15 *(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. Für beliebige $u, v \in \mathbb{C}^*$ gilt

$$(u \pm u^{-1})(v + v^{-1}) + (u \mp u^{-1})(v - v^{-1}) = 2(uv \pm u^{-1}v^{-1}) .$$

Mit $u = e^{iz}$ und $v = e^{iw}$ ergibt sich

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

Nach Division durch 4 ergibt sich 1. Entsprechend sieht man 2. □

Wir nutzen den Zwischenwertsatz (und damit einmal mehr die Vollständigkeit von $\mathbb{R}$) um die Kreiszahl π zu definieren. Dazu beweisen wir zunächst

Satz 7.16 *Es existiert ein* $t \in (0,2)$ *mit* $\cos t = 0$ *und es gilt* $\sin|_{[0,2]} \geq 0$.

Beweis. 1. Es gilt

$$\cos(2) = 1 - \frac{2^2}{2!} + \sum_{k=2}^{\infty} \frac{(-1)^k 2^{2k}}{(2k)!} = -1 + \sum_{\nu=0}^{\infty} (-1)^\nu \frac{4^{\nu+2}}{(2\nu+4)!}.$$

Setzen wir

$$a_n := \frac{4^{n+2}}{(2n+4)!} \qquad (n \in \mathbb{N}_0),$$

so gilt

$$\frac{a_n}{a_{n-1}} = \frac{4}{(2n+3)(2n+4)} < 1 \qquad (n \in \mathbb{N}).$$

Also ist (a_n) fallend und damit nach S. 6.20.1

$$s_n \leq s_0 = a_0 = \frac{4^2}{4!} = \frac{2}{3} \qquad (n \in 2\mathbb{N}_0).$$

Folglich ist auch $\sum_{\nu=0}^{\infty} (-1)^\nu a_\nu \leq 2/3$ und damit $\cos(2) \leq -1 + 2/3 < 0$. Da cos stetig ist, existiert nach dem Zwischenwertsatz ein $t \in (0,2)$ mit $\cos t = 0$.
2. Wir setzen für $t \in (0,2]$

$$a_n := \frac{t^{2n+1}}{(2n+1)!} \qquad (n \in \mathbb{N}_0).$$

Dann gilt

$$\frac{a_n}{a_{n-1}} = \frac{t^2}{(2n+1)(2n)} \leq \frac{4}{6} < 1 \qquad (n \in \mathbb{N}),$$

also ist (a_n) fallend. Aus S. 6.20.1 ergibt sich

$$\sin t = \sum_{\nu=0}^{\infty} (-1)^\nu \frac{t^{2\nu+1}}{(2\nu+1)!} = \lim_{n\to\infty} s_n \geq 0\,.$$

□

Bemerkung und Definition 7.17 Nach S. 7.16 ist $M := \{t > 0 : \cos(t) = 0\}$ nichtleer, also existiert $s := \inf M$. Aus der Stetigkeit von cos folgt $\cos(s) = 0$ und mit $\cos(0) = 1$ ist $s > 0$. Nach dem Zwischenwertsatz sowie $\cos(t) = \cos(-t)$ ist zudem $\cos(t) > 0$ für $t \in (-s, s)$.
Die **Kreiszahl** π definieren wir nun als $\pi := 2s$. Dann gilt mit B. 7.14 und S. 7.16

$$\cos(\pi/2) = 0, \qquad \sin(\pi/2) = 1$$

und damit auch

$$e^{i\pi/2} = i.$$

Hieraus ergibt sich wiederum $e^{\pi i} = i^2 = -1$ und $e^{2\pi i} = 1$, also

$$e^{z+2k\pi i} = e^z$$

für alle $k \in \mathbb{Z}$, d. h. exp ist $2\pi i$-periodisch[2].

Unter Ausnutzung der Additionstheoreme erhält man Periodizitätseigenschaften der trigonometrischen Funktionen

Satz 7.18 *Für alle $z \in \mathbb{C}$ gilt*

1. $\cos(z + \pi/2) = -\sin z$, $\qquad \sin(z + \pi/2) = \cos z$,
2. $\cos(z + \pi) = -\cos z$, $\qquad \sin(z + \pi) = -\sin z$,
3. $\cos(z + 2\pi) = \cos z$, $\qquad \sin(z + 2\pi) = \sin z$.

Beweis. 1. Mit S. 7.15 erhalten wir

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

und

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

Die Aussagen in 2. ergeben sich durch zweimalige Anwendung der ersten und die in 3. durch zweimalige Anwendung von 2. □

[2]Eine Funktion $f : \mathbb{K} \to \mathbb{K}$ heißt a-periodisch, falls $f(x + a) = f(x)$ für alle $x \in \mathbb{K}$ gilt.

Satz 7.19 $\exp|_{\mathbb{R}}$ *ist streng wachsend mit* $\exp(\mathbb{R}) = (0, \infty)$.

Beweis. Aus

$$e^x = 1 + x + \sum_{\nu=2}^{\infty} \frac{x^\nu}{\nu!} \geq 1 + x$$

für $x \geq 0$ und $e^{-x} = 1/e^x$ folgt $\exp(\mathbb{R}) \subset (0, \infty)$ und zudem $e^t \to \infty$ $(t \to \infty)$ sowie $e^t = 1/e^{-t} \to 0$ $(t \to -\infty)$. Nach dem Zwischenwertsatz ist $\exp(\mathbb{R}) = (0, \infty)$. Für $s < t$ ergibt sich $e^t/e^s = e^{t-s} \geq 1 + (t - s) > 1$ und damit $e^t > e^s$. Also ist $\exp|_{\mathbb{R}}$ streng wachsend. □

Satz 7.20 *Es gilt*

1. $\sin|_{[-\pi/2, \pi/2]}$ *ist streng wachsend mit* $\sin([-\pi/2, \pi/2]) = [-1, 1]$.
2. $\cos|_{[0,\pi]}$ *ist streng fallend mit* $\cos([0, \pi]) = [-1, 1]$.

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 $s, t \in \mathbb{R}$

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

Ist $-\pi/2 \leq s < t \leq \pi/2$, so gilt

$$(t + s)/2 \in (-\pi/2, \pi/2) \quad \text{und} \quad (t - s)/2 \in (0, \pi/2].$$

Nach B. 7.17 ist $\cos((t + s)/2) > 0$ und mit S. 7.18 folgt $\sin((t - s)/2) > 0$. Also ist $\sin s < \sin t$. Aus $\sin(\pi/2) = 1$ sowie $\sin(-\pi/2) = -\sin(\pi/2) = -1$ ergibt sich daher mit dem Zwischenwertsatz auch $\sin([-\pi/2, \pi/2]) = [-1, 1]$.
2. Die zweite Aussage folgt aus 1. und S. 7.18. □

Satz 7.21 *Es gilt*

$$Z(\exp -1) = 2\pi i\mathbb{Z}, \quad Z(\sin) = \pi\mathbb{Z} \quad \text{und} \quad Z(\cos) = \pi(\mathbb{Z} + 1/2).$$

Beweis. 1. Aus der $2\pi i$-Periodizität von exp und $e^0 = 1$ folgt $Z(\exp -1) \supset 2\pi i\mathbb{Z}$.
$\subset$: Da cos streng fallend auf $[0, \pi]$ ist und $\cos(2\pi - t) = \cos(-t) = \cos t$ für alle $t \in \mathbb{R}$ gilt, ist $\cos(t) < 1$ für alle $t \in \mathbb{R} \setminus 2\pi\mathbb{Z}$, also $e^{it} \neq 1$ für $t \in \mathbb{R} \setminus 2\pi\mathbb{Z}$.
Ist nun $z = s + it$ mit $e^z = 1$, so gilt $1 = |e^z| = e^s|e^{it}| = e^s$ und folglich $s = 0$. Damit ist $e^{it} = 1$, also $t \in 2\pi\mathbb{Z}$ d. h. $z = it \in 2\pi i\mathbb{Z}$.
2. Es gilt $0 = 2i \sin z = e^{iz} - e^{-iz}$ genau dann, wenn $e^{2iz} - 1 = 0$ ist. Aus 1. ergibt sich damit $Z(\sin) = \pi\mathbb{Z}$ und mit S. 7.18 dann auch $Z(\cos) = \pi(\mathbb{Z} + 1/2)$. □

Bemerkung und Definition 7.22 Die **Tangensfunktion** $\tan : \mathbb{C} \setminus \pi(\mathbb{Z}+1/2) \to \mathbb{C}$ und die **Kotangensfunktion** $\cot : \mathbb{C} \setminus \pi\mathbb{Z} \to \mathbb{C}$ sind definiert durch

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

Nach B. 5.11 und B. 7.14 sind tan und cot stetig auf ihren jeweiligen Definitionsbereichen.

Bemerkung und Definition 7.23 Wir schreiben $\mathbb{S} := \{z \in \mathbb{C} : |z| = 1\}$ für den **Einheitskreis** in $\mathbb{C}$. Ist $z \in \mathbb{C}^*$, so gilt $z = r\zeta$ mit $r = |z| > 0$ und $\zeta = z/|z| \in \mathbb{S}$. Diese Darstellung von z nennt man **Polarform**.

Satz 7.24 *Es gilt* $\exp(i\mathbb{R}) = \mathbb{S}$ *und* $\exp(\mathbb{C}) = \mathbb{C}^*$.

Beweis. 1. Nach B. 7.14 ist $\exp(i\mathbb{R}) \subset \mathbb{S}$. Wir zeigen $\supset$. Dazu sei $w \in \mathbb{S}$ mit Normalform $u + iv$. Ohne Einschränkung können wir $v \geq 0$ annehmen (ist $e^{it} = u + iv$, so ist $e^{-it} = u - iv$). Nach S. 7.20 existiert ein $t \in [0, \pi]$ mit $u = \cos t$. Dann ist $\sin t \geq 0$ und

$$v^2 = 1 - u^2 = 1 - \cos^2 t = \sin^2 t$$

also $v = \sin t$ und damit $e^{it} = \cos t + i \sin t = w$.
2. Mit 1. ergibt sich nach B./D. 7.23 und S. 7.19

$$\mathbb{C}^* = (0, \infty) \cdot \mathbb{S} = \exp(\mathbb{R}) \cdot \exp(i\mathbb{R}) = \exp(\mathbb{R} + i\mathbb{R}).$$

□

Bemerkung und Definition 7.25 Wie bereits angedeutet, sind im Körper $\mathbb{C}$ Gleichungen der Form $z^n = c$ stets, also für alle $c \in \mathbb{C}$ und $n \in \mathbb{N}$, lösbar. Einen Beweis haben wir bisher noch nicht erbracht. Mithilfe von S. 7.24 ergibt sich die Behauptung sehr einfach: Ohne Einschränkung sei $c \neq 0$. Dann ist $c = e^w$ für ein $w \in \mathbb{C}$. Für $z := e^{w/n}$ ergibt sich

$$z^n = (e^{w/n})^n = c.$$

Aufgrund der $2\pi i$-Periodizität von exp ist dann auch $(ze^{2k\pi i/n})^n = c$ für $k \in \mathbb{Z}$, wobei die n Zahlen

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

paarweise verschieden sind und damit alle Lösungen der Gleichung darstellen. Man nennt $z_0, \dots, z_{n-1}$ die n-ten **Wurzeln** aus c. Im Fall $c = 1$ spricht man auch von den n-ten **Einheitswurzeln**. So sind etwa ± 1 die zweiten Einheitswurzeln und $\pm i, \pm 1$ die vierten Einheitswurzeln.

Bemerkung und Definition 7.26 (Polarkoordinaten) Aus S. 7.24 ergibt sich eine weitere Möglichkeit der Darstellung komplexer Zahlen, die sich für viele Zwecke als angemessen erweist. Ist $z \in \mathbb{C}^*$ mit Polarform $z = r\zeta$, so existiert ein $\theta \in \mathbb{R}$ mit $\zeta = e^{i\theta}$, also

$$z = re^{i\theta}.$$

Fixiert man $\alpha \in \mathbb{R}$ und beschränkt man θ auf das Intervall $(\alpha - \pi, \alpha + \pi]$, so ist die Darstellung nach S. 7.21 eindeutig. Ist $z = (s, t) \in \mathbb{R}^2$, so ist damit auch

$$(s, t) = (r \cos\theta, r \sin\theta).$$

Man nennt dann $r > 0$ und $\theta \in (\alpha - \pi, \alpha + \pi]$ die **Polarkoordinaten** von (s, t) bezüglich α. Meist wählt man $\alpha = 0$ oder $\alpha = \pi$.

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

Satz 7.27 *Es sei $I \subset \mathbb{R}$ ein Intervall und es sei $f : I \to W(f) \subset \mathbb{R}$ streng wachsend (bzw. fallend). Dann ist f bijektiv und es gilt*

1. *f^{-1} ist streng wachsend (bzw. fallend).*
2. *f^{-1} ist stetig.*

Beweis. Wir setzen $J := W(f)$. Aus der stengen Monotonie folgt, dass $f : I \to J$ injektiv (also auch bijektiv) ist, d. h. $f^{-1} : J \to I$ existiert.
1. Ohne Einschränkung sei f steng wachsend. Angenommen, es existieren $u, v \in J$ mit $u < v$ und $s := f^{-1}(u) \geq f^{-1}(v) =: t$. Dann gilt $u = f(s) \geq f(t) = v$, da f (streng) wachsend ist. Widerspruch! Also ist $f^{-1} : J \to I$ streng wachsend.
2. Es seien $u \in J$ und $\varepsilon > 0$ gegeben. Wir setzen $t := f^{-1}(u)$. Ist $t \neq \sup I$, so existiert ein $h = h_\varepsilon \in (0, \varepsilon)$ mit $t + h \in I$. Wir setzen

$$\delta^+ := \delta^+_\varepsilon := f(t + h) - f(t) .$$

Dann ist $\delta^+ > 0$ und für alle $v \in J$ mit $u \leq v < u + \delta^+ = f(t + h)$ folgt

$$0 \leq f^{-1}(v) - f^{-1}(u) < f^{-1}(u + \delta^+) - f^{-1}(u) = h < \varepsilon .$$

Ist $t \neq \inf I$, so sieht man entsprechend: Es existiert ein $\delta^- > 0$ so, dass

$$0 \leq f^{-1}(u) - f^{-1}(v) < \varepsilon$$

für alle $v \in J$ mit $u - \delta^- < v \leq u$. Damit ergibt sich $|f^{-1}(v) - f^{-1}(u)| < \varepsilon$ für alle $v \in J$ mit $|v - u| < \delta := \min\{\delta^+, \delta^-\}$. □

Bemerkung und Definition 7.28 Nach S. 7.19 und S. 7.27 existiert die Umkehrfunktion von exp auf dem Intervall $(0, \infty)$ und ist dort stetig und streng wachsend. Diese Funktion nennnen wir(natürliche) **Logarithmusfunktion** und schreiben dafür ln oder auch log. Aus den entsprechenden Eigenschaften der Exponentialfunktion ergibt sich leicht ([Ü]):

1. Für alle $s, t > 0$ ist $\ln(st) = \ln(s) + \ln(t)$.

2. Für alle $t > 0$ und alle $m \in \mathbb{Z}$ ist $\ln(t^m) = m \ln(t)$.

Definition 7.29 Für $a > 0$ und $m \in \mathbb{Z}$ ist

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

nach B./D. 7.28.2. Wir setzen für allgemeines $z \in \mathbb{C}$

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

Aus den Rechenregeln für ln und exp erhält man $a^{1/n} = \sqrt[n]{a}$ für $a > 0$ und die folgenden Potenzgesetze.

Satz 7.30 *Es seien $a, b > 0$, $c \in \mathbb{R}$ und $z, w \in \mathbb{C}$. Dann gilt*

$$a^z a^w = a^{z+w}, \quad a^z b^z = (ab)^z \quad \textit{und} \quad (a^c)^z = a^{cz}.$$

Beweis. Es gilt

$$a^z a^w = e^{z \ln a} e^{w \ln a} = e^{z \ln a + w \ln a} = e^{(z+w) \ln a} = a^{z+w}$$

und

$$a^z b^z = e^{z \ln a} e^{z \ln b} = e^{z(\ln a + \ln b)} = e^{z \ln(ab)} = (ab)^z.$$

Für $c \in \mathbb{R}$ ist $a^c = e^{c \ln a} > 0$ und damit

$$(a^c)^z = e^{z \ln(e^{c \ln a})} = e^{cz \ln a} = a^{cz}\ .$$

□

Bemerkung und Definition 7.31 Die nach S. 7.27 und S. 7.20 auf $[-1, 1]$ existierende (und dort streng wachsende und stetige) Umkehrfunktion von sin heißt arcsin. Entsprechend bezeichnet man die auf $[-1, 1]$ existierende (und dort streng fallende und stetige) Umkehrfunktion von cos mit arccos.
Außerdem gilt: tan ist streng wachsend in $(-\pi/2, \pi/2)$ und cot ist streng 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.

8 Metrische Räume

Wie wir bereits in den vorhergehenden Abschnitten gesehen haben, spielt in der Analysis das Konzept der Grenzwerte eine zentrale Rolle. Dabei ist es wesentlich, von Abständen zwischen zwei Elementen in einer Menge sprechen zu können.

Definition 8.1 Es sei $X \neq \emptyset$ eine Menge. Eine Abbildung $d : X \times X \to \mathbb{R}$ heißt **Metrik** (oder **Abstand**) auf X, falls folgende Bedingungen erfüllt sind:

(d1) (Definitheit) Für alle $x, y \in X$ ist $d(x, x) = 0$ und $d(x, y) > 0$ falls $x \neq y$.

(d2) (Symmetrie) Für alle $x, y \in X$ ist $d(x, y) = d(y, x)$.

(d3) (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 und Definition 8.2 1. Ist $X \neq \emptyset$ eine beliebige Menge, so definiert

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

eine Metrik auf X, die sogenannte **diskrete Metrik**. Insbesondere kann also jede nichtleere Menge mit einer Metrik versehen werden.
2. Ist (X, d) ein metrischer Raum und ist $M \subset X$ nichtleer, so ist durch $d_M := d|_{M \times M}$ eine Metrik auf M gegeben.
3. Sind $(X_1, d_1), \ldots, (X_m, d_m)$ metrische Räume, so sind durch

$$d_\infty(x, y) := \max\{d_j(x_j, y_j) : j = 1, \ldots, m\} \quad (x = (x_1, \ldots, x_m), y = (y_1, \ldots, y_m))$$

und

$$d_{\text{sum}}(x, y) := \sum_{j=1}^{m} d_j(x_j, y_j) \qquad (x = (x_1, \ldots, x_m), y = (y_1, \ldots, y_m))$$

Metriken auf $X_1 \times \ldots \times X_m$ definiert mit ([Ü])

$$d_\infty \leq d_{\text{sum}} \leq m \cdot d_\infty \, . \tag{8.1}$$

Bemerkung und Definition 8.3 Ist $(V, \|\cdot\|)$ ein normierter Raum, so ist durch

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

eine Metrik auf V gegeben, die sogenannte **induzierte Metrik**. Insbesondere ist durch

$$d(x,y) := |x-y| = ||x-y||_2 \qquad (x,y \in \mathbb{K}^m)$$

eine Metrik auf $\mathbb{K}^m$ definiert (vgl. B./D. C.5). *Wenn wir im Weiteren von $\mathbb{K}^m$ als metrischem Raum sprechen soll stets diese Metrik gemeint sein* (falls nichts Anderes gesagt wird). Weiter gilt hier $d_{||\cdot||_\infty} = d_\infty$ sowie $d_{||\cdot||_1} = d_{\text{sum}}$ und mit B./D. C.5

$$d_\infty(x,y) \le d(x,y) \le d_{\text{sum}}(x,y) \qquad (x,y \in \mathbb{K}^d). \tag{8.2}$$

Wir untersuchen nun Folgen in metrischen Räumen

Bemerkung und Definition 8.4 Es sei (X,d) ein metrischer Raum. Eine Folge $(x_n)_{n\in N}$ in X heißt (d-)**konvergent** falls $(d(x_n,c))_{n\in N}$ für ein $c \in X$ eine Nullfolge (in $\mathbb{R}$) ist. Dann ist c wieder eindeutig bestimmt. Wir nennen c den (d)-**Grenzwert** und wir schreiben dann auch $x_n \to c$ $(n \to \infty)$ und

$$\lim x_n := \lim_{n\to\infty} x_n := c.$$

Im Fall $d = d_{||\cdot||}$ schreiben wir auch kurz $||\cdot||$ statt $d_{||\cdot||}$, also etwa $||\cdot||$-konvergent.

Bemerkung 8.5 Es seien X eine Menge und δ die diskrete Metrik auf X. Eine Folge (x_n) in X ist genau dann δ-konvergent mit Grenzwert c, wenn $x_n = c$ für n groß ist ([Ü]). Ist $X = \mathbb{R}$, so ist $(1/n)$ also *nicht* δ-konvergent (obwohl $(1/n)$ natürlich $|\cdot|$-konvergent ist).

Bemerkung und Definition 8.6 Ist (X,d) ein metrischer Raum, so heißt eine Folge $(x_n)_{n\in N}$ in X eine (d-)**Cauchyfolge**, falls zu jedem $\varepsilon > 0$ ein $R > 0$ existiert mit

$$d(x_n, x_{n'}) < \varepsilon \qquad (n, n' > R)\ .$$

Ganz allgemein gilt (mit gleichem Beweis wie im Fall $(\mathbb{K}, d_{|\cdot|})$):

1. Jede konvergente Folge ist eine Cauchyfolge.

2. Jede Cauchyfolge, die eine konvergente Teilfolge besitzt, ist konvergent.

Bemerkung 8.7 Ist (X,d) ein metrischer Raum und (x_n) eine Folge in X mit $x_n \ne c$ für alle n und $x_n \to c$ $(n \to \infty)$, so ist mit $M := X \setminus \{c\}$ die Folge (x_n) eine d_M-Cauchyfolge, aber nicht d_M-konvergent.
Im Allgemeinen ist also *nicht jede Cauchyfolge konvergent*!

Definition 8.8 Eine Metrik d auf X beziehungsweise der metrische Raum (X, d) heißt **vollständig**, falls jede Cauchyfolge konvergiert. Ist V ein normierter Raum so, dass $d_{||\cdot||}$ vollständig ist, so nennt man $(V, ||\cdot||)$ einen **Banachraum**.

Bemerkung 8.9 1. Es seien $(X_1, d_1), \ldots, (X_m, d_m)$ metrische Räume und es sei

$$(x_n)_{n\in N} = (x_{1,n}, \ldots, x_{m,n})_{n\in N}$$

eine Folge in $X_1 \times \cdots \times X_m$. Aus den Abschätzungen (8.1) folgt unmittelbar, dass folgende Aussagen äquivalent sind:

a) $(x_n)_{n\in N}$ ist d_∞-konvergent (bzw. eine d_∞-Cauchyfolge)

b) $(x_n)_{n\in N}$ ist d_{sum}-konvergent (bzw. eine d_{sum}-Cauchyfolge)

c) Jede **Komponentenfolge** $(x_{j,n})_{n\in N}$ ist d_j-konvergent (bzw. eine d_j-Cauchyfolge).

Außerdem gilt im Falle der Konvergenz

$$\lim_{n\to\infty} x_n = \big(\lim_{n\to\infty} x_{1,n}, \ldots, \lim_{n\to\infty} x_{m,n}\big).$$

Ist speziell $X_1 \times \cdots \times X_m = \mathbb{K}^m$, so ergibt sich mit (8.2), dass a)-c) auch äquivalent dazu sind, dass die Folge $|\cdot|$-konvergent (bzw. eine $|\cdot|$-Cauchyfolge) ist.

Bemerkung 8.10 Unter Verwendung des Cauchy-Kriteriums für Folgen in $\mathbb{K}$ ergibt sich mit B. 8.9 die Vollständigkeit von $(\mathbb{K}^m, d_{|\cdot|})$. Also ist $(\mathbb{K}^m, |\cdot|)$ ein Banachraum.

Definition 8.11 Es sei (X, d) ein metrischer Raum.
1. Für $a \in X$ und $\rho > 0$ heißt die Menge

$$U_\rho(a) := U_{\rho,d}(a) := \{x \in X : d(x, a) < \rho\}$$

ρ-Umgebung (bzw. (ρ, d)-Umgebung) von a. Weiter setzen wir

$$B_\rho(a) := B_{\rho,d}(a) := \{x \in X : d(x, a) \le \rho\}$$

und

$$K_\rho(a) := K_{\rho,d}(a) := \{x \in X : d(x, a) = \rho\}.$$

2. Ist $M \subset X$, so heißt $a \in M$ ein **innerer Punkt** von M (in (X, d)), falls ein $\rho > 0$ existiert mit $U_\rho(a) \subset M$. In diesem Fall heißt zudem M eine **Umgebung** von a (in (X, d)).

Definition 8.12 Es seien (X, d) ein metrischer Raum und $M \subset X$. Dann heißt M

1. **offen** (in (X, d)), falls jeder Punkt $x \in M$ ein innerer Punkt von M ist,

2. **abgeschlossen** (in (X, d)), falls $M^c = X \setminus M$ offen ist.

Beispiel 8.13 1. In jedem metrischen Raum (X, d) sind X und $\emptyset$ offen und abgeschlossen. Außerdem folgt aus der Dreiecksungleichung leicht, dass $U_\rho(a)$ offen ist für beliebige $a \in X$ und $\rho > 0$.
2. Für $a, b \in \mathbb{R} \cup \{\pm\infty\}$ mit $a < b$ ist das Intervall (a, b) offen in $\mathbb{R}$. Sind $a, b \in \mathbb{R}$, so sind die Intervalle $[a, b]$, $(-\infty, b]$ sowie $[a, \infty)$ abgeschlossen in $\mathbb{R}$. Intervalle der Form $(a, b]$ oder $[a, b)$ sind weder offen noch abgeschlossen in $\mathbb{R}$.

Der folgende Satz liefert eine Charakterisierung der Abgeschlossenheit mittels Folgen.

Satz 8.14 *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. $\Rightarrow$: Es sei (x_n) eine Folge in M mit $x_n \to a$ $(n \to \infty)$. Für beliebiges $\varepsilon > 0$ ist dann $M \cap U_\varepsilon(a) \neq \emptyset$ (da $x_n \in U_\varepsilon(a)$ für n groß). Nach Voraussetzung ist M^c offen. Also gilt $a \notin M^c$, d. h. $a \in M$.
$\Leftarrow$: Angenommen, M^c ist nicht offen. Dann existiert ein $a \in M^c$ mit $U_{1/n}(a) \cap M \neq \emptyset$ für alle $n \in \mathbb{N}$. Ist $x_n \in U_{1/n}(a) \cap M$, so gilt $x_n \to a$ $(n \to \infty)$. Widerspruch. $\square$

Satz 8.15 *Es sei (X, d) ein metrischer Raum. Ist $(M_\alpha)_{\alpha \in I}$ eine Familie von Mengen in X, so gilt*

1. *Sind alle M_α offen, so ist $\bigcup_{\alpha \in I} M_\alpha$ und für endliches I auch $\bigcap_{\alpha \in I} M_\alpha$ offen.*

2. *Sind alle M_α abgeschlossen, so ist $\bigcap_{\alpha \in I} M_\alpha$ und für endliches I auch $\bigcup_{\alpha \in I} M_\alpha$ abgeschlossen.*

Beweis. 1. Ist $x \in M_\alpha$, so existiert nach Vorausssetzung ein $\varepsilon_\alpha = \varepsilon_{\alpha,x} > 0$ mit $U_{\varepsilon_\alpha}(x) \subset M_\alpha$. Ist also $x \in \bigcup_{\alpha \in I} M_\alpha$, so existiert ein $\beta \in I$ mit $x \in M_\beta$ und damit

$U_{\varepsilon_\beta}(x) \subset M_\beta$. Ist I endlich, $x \in \bigcap_{\alpha \in I} M_\alpha$ und $\varepsilon := \min\{\varepsilon_\alpha : \alpha \in I\}$, so gilt $U_\varepsilon(x) \subset M_\alpha$ für alle $\alpha \in I$.
2. Sind alle M_α abgeschlossen, so sind alle M_α^c offen und damit nach den De Morganschen Regeln und 1. auch

$$\Big(\bigcap_{\alpha \in I} M_\alpha\Big)^c = \bigcup_{\alpha \in I} M_\alpha^c.$$

Entsprechend ergibt sich die Behauptung für $\bigcup_{\alpha \in I} M_\alpha$ im Falle einer endlichen Indexmenge I. □

Bemerkung 8.16 Im Allgemeinen sind unendliche Schnitte offener Mengen nicht mehr offen: Ist etwa $X = \mathbb{R}$, so ist für die offenen 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 im Allgemeinen auch unendliche Vereinigungen abgeschlossener Mengen nicht mehr abgeschlossen sind.

Bemerkung und Definition 8.17 Es seien (X, d_X), (Y, d_Y) metrische Räume und $f : X \to Y$. Man kann die Stetigkeit von f and einer Stelle $a \in X$ genau wie in im Falle $X \subset \mathbb{C}$ und $Y = \mathbb{K}$ definieren, wobei man lediglich $|x - a|$ durch $d_X(x, a)$ und $|f(x) - f(a)|$ durch $d_Y(f(x), f(a))$ ersetzt, also: f heißt **stetig** an a, falls zu jedem $\varepsilon > 0$ ein $\delta = \delta_\varepsilon > 0$ existiert mit $d_Y(f(x), f(a)) < \varepsilon$ für alle $x \in X$ mit $d_X(x, a) < \delta$. Dies ist genau dann der Fall, wenn zu jeder offenen Umgebung V von $f(a)$ eine (offene) Umgebung U von a existiert mit

$$f(U) \subset V.$$

Außerdem ist f genau dann stetig an a, wenn f **folgenstetig** an a ist, d. h. wenn für alle Folgen (x_n) in X mit $x_n \to a$ auch $f(x_n) \to f(a)$ gilt ([Ü]). Man beachte, dass die Stetigkeit wesentlich von den Metriken auf X und Y abhängt. Wir schreiben daher auch $f : (X, d_X) \to (Y, d_Y)$. Wie üblich heißt f stetig auf $M \subset X$, wenn f stetig an jedem Punkt aus M ist und im Fall $M = X$ kurz stetig. Wir setzen

$$C(X, Y) := \{f : X \to Y : f \text{ stetig}\}$$

und $C(X) := C(X, \mathbb{C})$.

Bemerkung und Definition 8.18 Es seien (X, d) ein metrischer Raum, Y eine Menge und $f : X \to Y$.
1. Wir sagen, dass f eine Eigenschaft **lokal** an der Stelle $a \in X$ erfüllt, wenn eine Umgebung U von a so existiert, dass die Eigenschaft für $f|_U$ gilt. So heißt etwa f **lokal konstant** an der Stelle $a \in X$, falls $f|_U$ für eine Umgebung U von a konstant ist. Ist d_Y eine beliebige Metrik auf Y, so ergibt sich aus der Definition der Setigkeit sofort, dass jede an a lokal konstante Funktion auch stetig an a ist. Ist f lokal konstant an jeder Stelle $a \in X$, so heißt f lokal konstant.
2. Mit der Äquivalenz von Stetigkeit und Folgensetigkeit ergibt sich aus B. 8.9: Ist $f = (f_1, \dots, f_m) : X \to \mathbb{K}^m$, so ist f stetig an der Stelle $a \in X$ genau dann, wenn jede **Komponentenfunktion** $f_j : X \to \mathbb{K}$ stetig an a ist.

Bemerkung 8.19 Es seien (X, d_X), (Y, d_Y) und (Z, d_Z) metrische Räume.
1. Wie in B. 5.15 ergibt sich leicht: Ist $f : X \to Y$ stetig an a und ist $g : Y \to Z$ stetig an $f(a)$, so ist auch $g \circ f$ stetig an a.
2. Die Grenzwertaussagen aus B. 5.11 zeigen, dass die Abbildungen $(x, y) \mapsto x \pm y$ und $(x, y) \mapsto xy$ (folgen-)stetig auf $\mathbb{K}^2$ sind. Das Gleiche gilt für $(x, y) \mapsto x/y$ auf $\mathbb{K} \times \mathbb{K}^*$. Mit 1. und B. 8.18.2 ergibt sich: Sind $f, g : X \to \mathbb{K}$ stetig an a, so sind auch $f \pm g$, $f \cdot g$ und (falls definiert) f/g stetig an a. Insbesondere ist $C(X, \mathbb{K})$ eine $\mathbb{K}$-Algebra (siehe B. C.2).

Satz 8.20 *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 offenen Mengen $V \subset Y$ ist $f^{-1}(V) \subset X$ offen.

c) Für alle abgschlossenen Mengen $B \subset Y$ ist $f^{-1}(B) \subset X$ abgeschlossen.

Beweis. a) $\Rightarrow$ b): Es sei $V \subset Y$ offen. Ist $a \in f^{-1}(V)$, so ist V eine offene Umgebung von $f(a)$. Da f stetig an a ist, existiert eine Umgebung U von a mit $f(U) \subset V$, also $U \subset f^{-1}(f(U)) \subset f^{-1}(V)$. Damit ist $f^{-1}(V)$ offen.
b) $\Rightarrow$ a): Es seien $a \in X$ und V eine offene Umgebung von $f(a)$. Nach Voraussetzung ist $U := f^{-1}(V)$ offen in X. Da $a \in U$ gilt, ist U eine Umgebung von a mit $f(U) = f(f^{-1}(V)) \subset V$. Also ist f stetig an a.
Die Äquivalenz von b) und c) ergibt sich durch Komplementbildung (man beachte, dass $(f^{-1}(M))^c = f^{-1}(M^c)$ für beliebige Mengen $M \subset Y$ gilt). $\square$

Definition 8.21 Es seien (X, d) ein metrischer Raum und $M \subset X$.
1. Ein Punkt $a \in X$ heißt **Häufungspunkt** von M (in (X, d)), falls für jede Umgebung U von a die Menge $M \cap (U \setminus \{a\})$ nichtleer ist (oder äquivalent, wenn eine Folge in M existiert mit $a \neq x_n \to a$). Wir schreiben $M' := M'_{(X,d)}$ für die Menge der Häufungspunkte von M. Ist $M' = \emptyset$, so heißt M **diskret** (in (X, d)).
2. Die Menge $M^\circ = M^\circ_{(X,d)}$ der inneren Punkte von M heißt **Inneres** von M, die Menge $\overline{M} = \overline{M}_{(X,d)} := M \cup M'$ **Abschluss** von M und $\partial M := \partial_{(X,d)} M := \overline{M} \setminus M^o$ **Rand** von M.
3. M heißt **dicht** (in (X, d)), falls $\overline{M} = X$ gilt. Außerdem heißt (X, d) **separabel**, falls eine abzählbare dichte Teilmenge existiert.

Bemerkung 8.22 Es sei (X, d) ein metrischer Raum. Aus der Definition ergibt sich sofort, dass für jede Menge $M \subset X$ das Innere M° offen ist und dass M genau dann offen ist, wenn $M = M^\circ$ gilt. Entsprechend ist der Abschluss $\overline{M}$ abgeschlossen und M genau dann abgeschlossen, wenn $M = \overline{M}$ gilt ([Ü]). Aus $\partial M = \overline{M} \cap (M^\circ)^c$ folgt mit S. 8.15, dass auch ∂M abgeschlossen ist.

Beispiel 8.23 1. Es sei $X = \mathbb{C}$. Wir setzen

$$\mathbb{D} := U_1(0) = \{z \in \mathbb{C} : |z| < 1\}.$$

Hier ist $\overline{\mathbb{D}} = \{z : |z| \leq 1\} = B_1(0)$ und $\partial\mathbb{D} = \{z : |z| = 1\} = K_1(0) = \mathbb{S}$.
2. In $X = \mathbb{R}$ gilt $\mathbb{Q}^\circ = \emptyset$ und $\overline{\mathbb{Q}} = \mathbb{Q}' = \partial\mathbb{Q} = \mathbb{R}$. Insbesondere ist $\mathbb{R}$ separabel.

Definition 8.24 Ein metrischer Raum (X, d) heißt (X, d) **(folgen-)kompakt**, falls jede Folge in X eine konvergente Teilfolge besitzt. Eine Menge $M \subset X$ heißt **kompakt**, falls (M, d_M) kompakt ist (also jede Folge in M eine konvergente Teilfolge mit Grenzwert in M hat). Weiter heißt M **relativ kompakt** (in (X, d)), falls jede Folge in M eine (in X) konvergente Teilfolge hat.

Bemerkung 8.25 Ist (X, d) ein metrischer Raum, so ist $M \subset X$ kompakt genau dann, wenn M relativ kompakt und abgeschlossen ist.
(Denn: Es sei M kompakt. Dann ist M auch relativ kompakt. Ist (x_n) eine Folge in M mit $x_n \to a$, so gilt $a \in M$, da auch jede Teilfolge gegen a konvergiert. Nach S. 8.14 ist M abgeschlossen. Ist umgekehrt M relativ kompakt und abgschlossen, so hat jede Folge in M eine konvergente Teilfolge und nach S. 8.14 liegt der Grenzwert in M.)

Beispiel 8.26 Nach dem Satz von Bolzano-Weierstraß sind beschränkte Mengen in $\mathbb{K}$ relativ kompakt. Mit B. 8.25 folgt, dass Intervalle der Form $[a, b]$ kompakt sind.

Satz 8.27 *Es seien $(X_1, d_1), \ldots, (X_m, d_m)$ kompakte metrische Räume. Dann ist auch $(X_1 \times \cdots \times X_m, d_\infty)$ kompakt.*

Beweis. Wir beweisen die Behauptung per Induktion nach m.
Für $m = 1$ ist nichts zu zeigen.
$m \to m + 1$: Wir setzen $X := X_1 \times \cdots \times X_m$ und definieren $p : X \times X_{m+1} \to X$, $q : X \times X_{m+1} \to X_{m+1}$ durch

$$p(x) := (x_1, \ldots, x_m), \quad q(x) := x_{m+1} \quad (x = (x_1, \ldots, x_m, x_{m+1}) \in X \times X_{m+1}).$$

Ist $(x_n)_{n\in N}$ eine Folge in $X_1 \times \cdots \times X_{m+1}$, so ist $(p(x_n))_{n\in N}$ eine Folge in X. Also existiert nach Induktionsvoraussetzung eine d_∞-konvergente Teilfolge $(p(x_n))_{n\in I}$. Weiter ist $(q(x_n))_{n\in I}$ eine Folge in X_{m+1}. Also existiert eine konvergente Teilfolge $(q(x_n))_{n\in J}$. Aus

$$d_\infty(x_n, c) = \max\{d_\infty(p(x_n), p(c)), d_{m+1}(q(x_n), q(c))\}$$

für $c \in X \times X_{m+1}$ ergibt sich durch Anwendung von B. 8.9 die d_∞-Konvergenz von $(x_n)_{n\in J} = (p(x_n), q(x_n))_{n\in J}$. □

Beispiel 8.28 Nach B. 8.26 uns S. 8.27 sind Mengen der Form $[a_1, b_1] \times \ldots \times [a_m, b_m]$, wobei $a_j, b_j \in \mathbb{R}$ mit $a_j \leq b_j$ für $j = 1, \ldots, m$, kompakt.

Bemerkung und Definition 8.29 Es sei $(V, ||\cdot||)$ ein normierter Raum. Eine Menge $M \subset V$ heißt **beschränkt**, falls ein $R > 0$ existiert mit $||x|| \leq R$ für alle $x \in M$ (also $M \subset B_R(0)$ gilt). Ist $M \subset V$ relativ kompakt (also relativ kompakt im metrischen Raum $(V, d_{||\cdot||})$), so ist M beschränkt (anderenfalls würde eine Folge (x_n) in M existieren mit $\|x_n\| \to \infty$. Diese Folge hätte keine konvergente Teilfolge).
Für *Teilmengen M von $\mathbb{K}^m$* gilt umgekehrt auch: Ist M beschränkt, so ist M schon relativ kompakt. (Denn: Es existiert ein $R > 0$ mit $M \subset B_R(0) \subset [-R, R]^m$. Also ist M nach B. 8.28 Teilmenge einer kompakten Menge und damit relativ kompakt.)

Ein weiteres zentrales Resultat der Analysis ist mit den obigen Überlegungen bereits so gut wie bewiesen:

Satz 8.30 *(Heine-Borel)*
Teilmengen von $\mathbb{K}^m$ sind genau dann kompakt, wenn sie beschränkt und abgeschlossen sind.

Beweis. Nach B. 8.29 ist für Mengen $M \subset K^m$ Beschränktheit gleichbedeutend mit relativer Kompaktheit. Mit B. 8.25 ergibt sich damit die Behauptung $\square$

Bemerkung 8.31 Ist $M \subset \mathbb{R}$ beschränkt, so gilt $\sup M \in \overline{M}$ und $\inf M \in \overline{M}$. Ist also M kompakt (und damit zusätzlich abgeschlossen), so hat M Maximum und Minimum.

In vielen mathematischen Fragestellungen geht es darum, reellwertige Funktionen zu maximieren bzw. zu minimieren. Wir wollen zunächst die passenden Begriffe einführen.

Definition 8.32 Es seien $X \neq \emptyset$ eine Menge und $f : X \to \mathbb{R}$.

1. Man sagt, f hat ein **Maximum**, falls $\max\limits_X f := \max\limits_{x \in X} f(x) := \max f(X)$ existiert. Ist $x_0 \in X$ so, dass $f(x_0) = \max\limits_X f$ gilt, d. h. ist $f(x) \leq f(x_0)$ für alle $x \in X$, so sagt man, dass f das Maximum an x_0 annnimmt.
2. Man sagt, f hat ein **Minimum**, falls $\max\limits_X f := \min\limits_{x \in X} f(x) := \min f(X)$ existiert. Ist $x_0 \in X$ so, dass $f(x_0) = \min\limits_X f$ gilt, d. h. ist $f(x) \geq f(x_0)$ für alle $x \in X$, so sagt man, dass f das Minumim an x_0 annimmt.

Ist $M \subset X$, so sagt man auch f habe ein **Maximum auf** M (oder Maximim bezüglich M), falls $f|_M$ ein Maximum hat. Entsprechendes gilt für Minimum statt Maximum.

Beschränkte und stetige reellwertige Funktionen haben im Allgemeinen weder ein Maximum noch ein Minimum, wie etwa $\arctan : \mathbb{R} \to \mathbb{R}$ zeigt (der Wertebereich

$$\arctan(\mathbb{R}) = (-\pi/2, \pi/2)$$

ist offen). Günstiger ist die Situation bei stetigen Funktionen auf kompakten Mengen.

Satz 8.33 *Es seien* (X, d_X), (Y, d_Y) *metrische Räume und* $f : X \to Y$ *stetig. Dann gilt:*

1. *Ist* $M \subset X$ *relativ kompakt, so ist auch* $f(M) \subset Y$ *relativ kompakt.*
2. *Ist* X *kompakt, so ist auch* $f(X)$ *kompakt.*
3. *Ist* X *kompakt und* $Y = \mathbb{R}$*, so hat* f *Maximum und Minimum.*

Beweis. 1. Es sei $(y_n)_{n\in N}$ eine Folge in $f(M)$. Wir wählen $x_n \in M$ mit $y_n = f(x_n)$. Nach Voraussetzung existieren ein $a \in X$ und eine Teilfolge $(x_n)_{n\in I}$ von $(x_n)_{n\in N}$ mit $x_n \to a$ $(n \to \infty,\, n \in I)$. Da f stetig ist, folgt

$$y_n = f(x_n) \to f(a) \qquad (n \to \infty,\, n \in I)\ .$$

Damit ist $f(M)$ relativ kompakt.
2. Ist $X = M$ in 1., so ist natürlich $f(a) \in f(X)$.
3. Nach 2. und B. 8.31 existieren $\max\limits_X f$ und $\min\limits_X f$. □

Bemerkung 8.34 Es seien $I = (a, b)$ ein offenes Intervall und $f : I \to \mathbb{R}$ stetig. Existieren die Grenzwerte $f(a^+)$ und $f(b^-)$, so ist f beschränkt auf I, jedoch existieren im Allgemeinen weder Maximum noch Minimum, wie etwa wieder $f = \arctan$ auf $\mathbb{R}$ zeigt. Gilt jedoch $f(a^+) = f(b^-)$, so hat f jedenfalls Maximum *oder* Minimum.
(Denn: Ist f konstant, so ist die Behauptung klar. Ist f nicht konstant, so existiert ein $x_0 \in I$ mit $f(x_0) \neq f(a^+)(= f(b^-))$. Ist $f(x_0) > f(a^+)$, so existiert ein kompaktes Intervall $J \subset I$ mit $f(x_0) > f(x)$ für alle $x \in I\backslash J$. Nach S. 8.33 hat f ein Maximum auf J. Also ist $\max\limits_J f = \max\limits_I f$. Ist $f(x_0) < f(a^+)$, so hat entsprechend f ein Minimum.)

Bemerkung und Definition 8.35 Es seien $X \neq \emptyset$ eine Menge und $g : X \to \mathbb{R}$ beschränkt. Dann schreiben wir

$$\sup_X g := \sup_{x\in X} g(x) := \sup g(X).$$

Ist $(E, |\cdot|_E)$ ein normierter Raum, so setzt man

$$B(X, E) := \{f : X \to E : |f|_E \text{ beschränkt}\}$$

und $B(X) := B(X, \mathbb{C})$. Wie man leicht sieht, ist $B(X, E)$ ein Unterraum von E^X und damit ein Vektorraum. Aus $\sup(\alpha A) = \alpha \sup(A)$ und $\sup(A + B) \leq \sup A + \sup B$ für $\alpha \geq 0$ und beschränkte Mengen $A, B \subset [0, \infty)$ (tatsächlich gilt auch im zweiten Fall Gleichhheit; [Ü]) folgt, dass

$$||f||_\infty := ||f||_{\infty,X} := \sup_X |f|_E \qquad (f \in B(X, E))$$

mit eine Norm auf $B(X, E)$ definiert. Man nennt $||\cdot||_\infty$ die **Supremumsnorm** auf X (bezüglich $|\cdot|_E$). Im Falle eines kompakten metrischen Raumes (X, d) gilt $C(X, E) \subset B(X, E)$ nach S. 8.33 und $||f||_\infty = \max_X |f|_E$.

Eine stetige, bijektive Funktion zwischen metrischen Räumen, deren Umkehrabbildung ebenfalls stetig ist, nennt man einen **Homöomorphismus**. Im Allgemeinen ist die Umkehrfunktion einer stetigen, bijektiven Funktion nicht stetig ([Ü]). Als Anwendung von S. 8.33 erhält man jedoch:

Satz 8.36 *Es seien (X, d_X) und (Y, d_Y) metrische Räume. Ist (X, d) kompakt und ist $f : X \to Y$ bijektiv und stetig, so ist auch Y kompakt und f ein Homöomorphismus.*

Beweis. Zunächst ist $Y = f(X)$ nach S. 8.33 kompakt. Es sei $A \subset X$ abgeschlossen. Zm Beweis der Stetigkeit von f^{-1} reicht es nach S. 8.20 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. 8.25. Also ist $f(A) \subset Y$ kompakt nach S. 8.33 und damit insbesondere abgeschlossen (wieder nach B. 8.25). □

Eine weitere wichtige Eigenschaft stetiger Funktionen auf kompakten Mengen ist die gleichmäßige Stetigkeit:

Definition 8.37 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 $\delta = \delta_\varepsilon > 0$ so existiert, dass

$$d_Y(f(x), f(y)) < \varepsilon$$

für alle $x, y \in X$ mit $d_X(x, y) < \delta$.

Aus der Definition ergibt sich sofort, dass jede gleichmäßig stetige Funktion auch stetig ist. Das folgende Beispiel zeigt, dass die Stetigkeit im Allgemeinen nicht die gleichmäßige Stetigkeit impliziert.

Beispiel 8.38 Es sei $f : \mathbb{R} \to \mathbb{R}$ 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 $x = 1/\delta, y = 1/\delta + \delta/2$. Dann ist $|x - y| = \delta/2 < \delta$, aber

$$|f(x) - f(y)| = |x + y| \cdot |x - y| > 2/\delta \cdot \delta/2 = 1 = \varepsilon \; .$$

Folglich ist f nicht gleichmäßig stetig auf $\mathbb{R}$.)

Satz 8.39 *Es seien (X, d_X) und (Y, d_Y) metrische Räume. Ist (X, d) 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 $(x_n)_{n\in I}$ mit $x_n \to a \in X$ $(n \to \infty,\, n \in I)$. Damit gilt auch

$$d_X(a, y_n) \leq d_X(a, x_n) + d_X(x_n, y_n) \to 0 \qquad (n \to \infty,\, n \in I),$$

also $y_n \to a$ $(n \to \infty,\, n \in I)$. Aufgrund der (Folgen-)Stetigkeit von f an der Stelle a ergibt sich

$$\varepsilon \leq d_Y(f(x_n), f(y_n)) \leq d_Y(f(x_n), f(a)) + d_Y(f(a), f(y_n)) \to 0 \quad (n \to \infty, n \in I) \; .$$

Widerspruch! $\square$

9 Funktionenfolgen und Funktionenreihen

Wir haben schon früher gesehen, dass wichtige 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 9.1 Es seien $X \neq \emptyset$ eine Menge und (Y, d) ein metrischer Raum. Eine Folge $(f_n)_{n\in N}$ in $\mathrm{Abb}(X, Y) = Y^X$ nennt man eine **Funktionenfolge**. Die Funktionenfolge $(f_n)_{n\in N}$ heißt **punktweise konvergent** auf der Menge $M \subset X$, falls für alle $x \in M$ die Folge $(f_n(x))_{n\in N}$ in Y konvergiert. Die Funktion $f : M \to Y$ mit $f(x) := \lim_{n\to\infty} f_n(x)$ heißt **Grenzfunktion** der Folge $(f_n)_{n\in N}$ (auf M). Wir schreiben dann auch

$$f_n \to f\ (n \to \infty) \quad \text{punktweise auf } M.$$

Beispiel 9.2 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

$$f_n(x) \to \begin{cases} 0 & , \text{ falls } \ x \in (-1, 1) \\ 1 & , \text{ falls } \ x = 1 \end{cases} ,$$

d. h. (f_n) konvergiert punktweise auf $(-1, 1]$ mit der Grenzfunktion $f = 1_{\{1\},(-1,1]}$. Das Beispiel zeigt insbesondere, dass die Grenzfunktion unstetig (an der Stelle 1) ist, obwohl alle Folgelieder f_n stetige Funktionen auf $\mathbb{R}$ sind.

Wir führen nun einen strengeren Konvergenzbegriff für Funktionenfolgen ein, der den entscheidenden Vorteil hat, dass sich Stetigkeit auf die Grenzfunktion überträgt.

Bemerkung und Definition 9.3 Es seien $X \neq \emptyset$ eine Menge und (Y, d) ein metrischer Raum.

1. Für $f, g : X \to Y$ so, dass $X \ni x \mapsto d(f(x), g(x)) \in \mathbb{R}$ beschränkt ist, setzen wir

$$d_\infty(f, g) := d_{\infty,X}(f, g) := \sup_{x\in X} d(f(x), g(x)).$$

Ist $h : X \to Y$ so, dass $x \mapsto d(g(x), h(x))$ beschränkt ist, so gilt die Dreiecksungleichung

$$d_\infty(f, h) \leq d_\infty(f, g) + d_\infty(g, h).$$

Ist $(E, |\cdot|_E)$ ein normierter Raum und $(Y, d) = (E, d_{|\cdot|_E})$, so ist zudem

$$d_\infty(f, g) = \|f - g\|_\infty \qquad \big(f, g \in B(X, E)\big),$$

also d_∞ die von der Supremumsnorm induzierte Metrik auf $B(X, E)$.
2. Eine Folge $(f_n)_{n\in N}$ in $\mathrm{Abb}(X, Y)$ heißt **gleichmäßig konvergent** auf der Menge $M \subset X$ gegen die Grenzfunktion $f : M \to Y$, falls $(d_{\infty,M}(f, f_n))_{n\in N}$ (existiert und) eine Nullfolge ist. Wir schreiben dann

$$f_n \to f \ (n \to \infty) \quad \text{gleichmäßig auf } M$$

oder auch

$$f_n(x) \to f(x) \ (n \to \infty) \quad \text{gleichmäßig auf } M \ .$$

Ist $(E, |\cdot|_E)$ ein normierter Raum, so ist also für $f_n, f \in B(X, E)$ *gleichmäßige Konvergenz auf M dasselbe wie $||\cdot||_{\infty,M}$-Konvergenz.*
3. Gilt $f_n \to f$ gleichmäßig auf M, so folgt $f_n(x) \to f(x)$ $(n \to \infty)$ für alle $x \in M$ (da $d(f_n(x), f(x)) \le d_{\infty,M}(f_n, f)$); mit anderen Worten: gleichmäßige Konvergenz impliziert punktweise Konvergenz.

Beispiel 9.4 Wir betrachten noch einmal f_n und f aus B. 9.2. Ist $M = [-1/2, 1/2]$, so gilt

$$d_{\infty,M}(f_n, f) = ||f_n - f||_{\infty,M} = \max_{x\in M} |x^n| = 1/2^n \to 0 \qquad (n \to \infty) \ ,$$

also $f_n \to 0\,(= f|_M)$ $(n \to \infty)$ gleichmäßig auf $[-1/2, 1/2]$. Für $M = [0, 1)$ ist

$$d_{\infty,M}(f_n, f) = \|f_n - f\|_{\infty,M} = \sup_{x\in[0,1)} x^n = 1 \quad (n \in \mathbb{N}) \ .$$

Also ist (f_n) nicht gleichmäßig konvergent auf $[0, 1)$.

Satz 9.5 *Es seien (Y, d) ein metrischer und (X, d_X) ein kompakter metrischer Raum. Dann ist*

$$d_\infty(f, g) = \max_{x\in X} d\big(f(x), g(x)\big) \qquad (f, g \in C(X, Y))$$

und d_∞ eine Metrik auf $C(X, Y)$.

Beweis. Aus Dreiecksungleichung und umgekehrter Dreiecksungleichung ([Ü]) ergibt sich für $u, v, u', v' \in Y$

$$\big|d(u, v) - d(u', v')\big| \le \big|d(u, v) - d(u', v)\big| + \big|d(u', v) - d(u', v')\big| \le d(u, u') + d(v, v')$$

und damit

$$\big|d(f(x), g(x)) - d\big(f(y), g(y)\big)\big| \le d(f(x), f\big(y)\big) + d\big(g(x), g(y)\big)$$

für alle $f, g \in C(X, Y)$ und $x, y \in X$. Da X kompakt ist, sind f und g nach S. 8.39 gleichmäßig stetig. Also ist auch $x \mapsto d\big(f(x), g(x)\big)$ (gleichmäßig) stetig und hat folglich nach S. 8.33 ein Maximum. Wie bereits in B. 9.3 erwähnt gilt die Dreiecksungleichung, also (d3). Definitheit und Symmetrie von d_∞ folgen sofort aus den entsprechenden Eigenschaften von d. □

Wir kommen nun zu dem bereits angedeuteten Ergebnis über die Vererbung der Stetigkeit auf die Grenzfunktion.

Satz 9.6 *Es seien (X, d_X), (Y, d) metrische Räume und $a \in X$. Weiter sei $(f_n)_{n\in N}$ eine Folge in $\operatorname{Abb}(X, Y)$, die gleichmäßig auf einer Umgebung U von a gegen f konvergiert. Sind die Funktionen f_n stetig an a, so ist auch f stetig an a.*

Beweis. Es sei $\varepsilon > 0$ gegeben. Aufgrund der gleichmäßigen Konvergenz von (f_n) gegen f auf U existiert ein $n \in N$ mit

$$\sup_{x\in U} d(f(x), f_n(x)) < \varepsilon/3 \,.$$

Da f_n stetig an a ist, existiert ein $\delta = \delta_\varepsilon > 0$ so, dass $U_\delta(a) \subset U$ und

$$d(f_n(x), f_n(a)) < \varepsilon/3 \qquad (x \in U_\delta(a)) \,.$$

Damit ist für $x \in U_\delta(a)$

$$d(f(x), f(a)) \le d(f(x), f_n(x)) + d(f_n(x), f_n(a)) + d(f_n(a), f(a)) < \varepsilon \,.$$

□

Ein notwendiges und hinreichendes Kriterium für gleichmäßige Konvergenz liefert

Satz 9.7 *(Cauchy-Kriterium für gleichmäßige Konvergenz)*
Es seien $X \neq \emptyset$ eine Menge, (Y, d) ein vollständiger metrischer Raum und $(f_n)_{n\in N}$ eine Folge in $\operatorname{Abb}(X, Y)$. Ist $M \subset X$, so ist $(f_n)_{n\in N}$ gleichmäßig konvergent auf M genau dann, wenn zu jedem $\varepsilon > 0$ ein $R > 0$ existiert mit

$$d_{\infty,M}(f_n, f_{n'}) < \varepsilon \qquad (n, n' > R).$$

Beweis. $\Rightarrow$: Es gelte $f_n \to f$ gleichmäßig auf M. Ist $\varepsilon > 0$ gegeben, so existiert ein $R > 0$ mit

$$d_\infty(f, f_n) < \varepsilon/2 \qquad (n > R) .$$

Für alle $n, n' > R$ gilt dann mit der Dreiecksungleichung (siehe B./D. 9.3)

$$d_\infty(f_n, f_{n'}) \leq d_\infty(f_n, f) + d_\infty(f, f_{n'}) < \varepsilon.$$

$\Leftarrow$: Nach Voraussetzung ist insbesondere für jedes feste $x \in M$ die Folge $(f_n(x))_{n \in N}$ eine d-Cauchyfolge. Da (Y, d) vollständig ist, ist $(f_n(x))_{n \in N}$ konvergent. Wir definieren $f : M \to Y$ durch $f(x) := \lim_{n\to\infty} f_n(x)$ $(x \in M)$ und zeigen, dass (f_n) gleichmäßig auf M gegen f konvergiert.
Dazu sei $\varepsilon > 0$ gegeben. Dann existert ein $R = R_\varepsilon > 0$ mit $d_\infty(f_n, f_{n'}) < \varepsilon$ für $n, n' > R$. Es sei nun $n > R$ fest. Ist $x \in M$, so folgt aus der Stetigkeit der Abbildung $y \mapsto d(f_n(x), y)$ ([Ü]) für $m > R$

$$\varepsilon > d(f_n(x), f_m(x)) \to d(f_n(x), f(x)) \qquad (m \to \infty) .$$

Also ist $d(f_n(x), f(x)) \leq \varepsilon$. Da $x \in M$ beliebig war, folgt $d_\infty(f, f_n) \leq \varepsilon$. □

Bemerkung 9.8 Sind $X \neq \emptyset$ eine Menge und $(E, |\cdot|_E)$ ein Banachraum, so ist auch $(B(X, E), ||\cdot||_\infty)$ ein Banachraum.
(Denn: Es sei $(f_n)_{n\in N}$ eine $||\cdot||_\infty$-Cauchyfolge. Nach S. 9.7 existiert eine Funktion $f : X \to E$ mit $f_n \to f$ gleichmäßig auf X. Nach B. 9.3 reicht es zu zeigen, dass $|f|_E$ beschränkt ist. Dazu wählen wir ein $n \in N$ mit $||f - f_n||_\infty < 1$. Dann gilt für alle $x \in X$

$$|f(x)|_E \leq |f(x) - f_n(x)|_E + |f_n(x)|_E \leq 1 + ||f_n||_\infty.$$

Also ist $|f|_E$ beschränkt.)

Bemerkung 9.9 Es seien (X, d_X) *kompakt* und (Y, d) *vollständig*. Dann ist nach S. 9.7 und S. 9.6 der metrische Raum $(C(X, Y), d_\infty)$ vollständig.
Es sei nun (Y, d) *kompakt*. Ist X *endlich*, so ist $C(X, Y) = \text{Abb}(X, Y)$ und mit S. 8.27 sieht man, dass dann auch $(\text{Abb}(X, Y), d_\infty)$ kompakt ist. Im Allgemeinen ist jedoch der Raum $(C(X, Y), d_\infty)$ nicht kompakt: Sind etwa $X = Y = [0, 1]$ (jeweils mit der Betragsmetrik), so ergibt sich aus B. 9.4, dass die Folge (f_n) in $C(X, Y)$ mit

$$f_n(x) = x^n \qquad \big(x \in [0, 1]\big)$$

keine d_∞-konvergente (also auf $[0, 1]$ gleichmäßig konvergente) Teilfolge hat.

Definition 9.10 Sind $X \neq \emptyset$ eine Menge, $(E, |\cdot|_E)$ ein normierter Raum und $(f_n)_{n\geq m}$ eine Folge in $\mathrm{Abb}(X, E)$, so heißt die Funktionenfolge (s_n) mit

$$s_n(x) := \sum_{\nu=m}^{n} f_\nu(x) \qquad (x \in X,\, n \geq m)$$

eine **Funktionenreihe**. Man schreibt wieder $\sum\limits_{\nu=m}^{\infty} f_\nu$ statt $(s_n)_{n\geq m}$. Die Funktionenreihe $\sum\limits_{\nu=m}^{\infty} f_\nu$ 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. Man verwendet das Symbol $\sum\limits_{\nu=m}^{\infty} f_\nu$ dann auch wieder für die Grenzfunktion.

Es stellt sich die Frage nach hinreichenden Bedingungen für die gleichmäßige Konvergenz. Analog zu S. 7.6 ergibt sich

Satz 9.11 *(Weierstraß-Kriterium)*
Es seien $X \neq \emptyset$ eine Menge, $(E, |\cdot|_E)$ ein Banachraum und $(f_n)_{n\geq m}$ eine Folge in $B(X, E)$. Konvergiert die Reihe $\sum\limits_{\nu=m}^{\infty} ||f_\nu||_\infty$, so konvergiert $\sum\limits_{\nu=m}^{\infty} f_\nu$ gleichmäßig auf X.

Beweis. Zunächst gilt $s_n \in B(X, E)$, da $B(X, E)$ ein linearer Raum ist. Ist $\varepsilon > 0$ gegeben, so existiert nach S. 7.5 ein $R > 0$ mit

$$\sum_{\nu=n'+1}^{n} ||f_\nu||_\infty < \varepsilon \qquad (n > n' > R).$$

Damit ergibt sich für $n > n' > R$

$$||s_n - s_{n'}||_\infty = \Big\| \sum_{\nu=n'+1}^{n} f_\nu \Big\|_\infty \leq \sum_{\nu=n'+1}^{n} ||f_\nu||_\infty \leq \sum_{\nu=n'+1}^{n} ||f_\nu||_\infty < \varepsilon.$$

Also ist (s_n) eine $||\cdot||_\infty$-Cauchyfolge. Nach B. 9.8 ist (s_n) auch $||\cdot||_\infty$-konvergent. □

Beispiel 9.12 Für $\alpha > 1$ sei $S_\alpha := \{z \in \mathbb{C} : \mathrm{Re}\, z \geq \alpha\}$. Dann ist die Funktionenreihe $\sum\limits_{\nu=1}^{\infty} \nu^{-z}$ gleichmäßig konvergent auf S_α. (Denn: Für alle $z \in S_\alpha$ und alle $\nu \in \mathbb{N}$ ist

$$|n^{-z}| = |e^{-z \ln n}| = e^{-\mathrm{Re}(z) \ln n} \leq e^{-\alpha \ln n} = n^{-\alpha}.$$

Da die Reihe $\sum_{\nu=1}^{\infty} \nu^{-\alpha}$ konvergiert ([Ü]), ergibt sich die Behauptung aus S. 9.11.)
Mit $S := \{z \in \mathbb{C} : \operatorname{Re} z > 1\}$ heißt die Funktion $\zeta : S \to \mathbb{C}$, definiert durch

$$\zeta(z) := \sum_{\nu=1}^{\infty} \nu^{-z} \qquad (z \in S) ,$$

(Riemannsche) Zetafunktion. Da $z \mapsto n^{-z}$ für alle $n \in \mathbb{N}$ stetig auf $\mathbb{C}$ ist, folgt aus S. 9.6 die Stetigkeit der Zetafunktion auf S (Man beachte: Ist $\operatorname{Re}(a) > 1$ und $1 < \alpha < \operatorname{Re}(a)$, so ist S_α eine Umgebung von a).

Bemerkung und Definition 9.13 Es seien $a \in \mathbb{K}$ und $(c_n)_{n=0}^{\infty}$ eine Folge in $\mathbb{K}$. Dann heißt die Funktionenreihe $\sum_{\nu=0}^{\infty} f_\nu$ mit $f_n : \mathbb{K} \to \mathbb{K}$, definiert durch

$$f_n(x) := c_n(x-a)^n \quad (x \in \mathbb{K},\ n \in \mathbb{N}_0),$$

Potenzreihe mit **Entwicklungsmitte** a und **Koeffizientenfolge** (c_n). Man kann zeigen ([Ü]): Ist

$$R := \sup\{|h| : \sum_{\nu=0}^{\infty} c_\nu h^\nu \text{ konvergent}\} > 0,$$

so ist die Potenzreihe gleichmäßig konvergent auf $B_r(a)$ für alle $r < R$. Aus S. 9.6 ergibt sich damit die Stetigkeit von $x \mapsto \sum_{\nu=0}^{\infty} c_\nu (x-a)^\nu$ auf $U_R(a)$ (mit $U_\infty(a) := \mathbb{K}$). Man nennt R den **Konvergenzradius** und $U_R(a)$ den **Konvergenzkreis** (im Falle $\mathbb{K} = \mathbb{R}$ meist das **Konvergenzintervall**) der Potenzreihe.

10 Differenzialrechnung

Wir untersuchen nun Funktionen $f : X \to \mathbb{C}$, wobei $X \subset \mathbb{K}$, also Funktionen einer reellen oder komplexen Variable. Um die feinere Struktur des Veränderungsverhaltens solcher Funktionen untersuchen zu können, brauchen wir einen über die Stetigkeit hinausgehenden Glattheitsbegriff. Grob gesagt wollen wir Funktionen definieren, die lokal sehr gut durch affin-lineare Abbildungen approximiert werden können.

Definition 10.1 Es seien $X \subset \mathbb{K}$ und $f : X \to \mathbb{C}$.

1. f heißt **differenzierbar an der Stelle** $a \in X$, falls a Häufungspunkt von X ist und der Grenzwert
$$f'(a) := \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$
existiert. Man bezeichnet $f'(a)$ als **(erste) Ableitung** von f an der Stelle a.

2. f heißt **differenzierbar** (auf X), falls f in jedem Punkt $x \in X$ differenzierbar ist. In diesem Fall heißt die Funktion $f' : X \to \mathbb{C}$ **(erste) Ableitung** von f. Weitere Schreibweisen sind etwa Df oder df oder auch (df/dx)[3]. Ist f' stetig, so heißt f **stetig differenzierbar** (auf X).

Bemerkung 10.2 Es seien wieder $X \subset \mathbb{K}$, $a \in X \cap X'$ und $f : X \to \mathbb{C}$. Definiert man die Funktion $\tau_a f : (X - a) \to \mathbb{C}$ durch

$$(\tau_a f)(h) := f(a + h) - f(a) \qquad (h \in X - a),$$

so ist $\tau_a f(0) = 0$ und f nach Bemerkung 5.15 genau dann differenzierbar an a, wenn

$$\lim_{h \to 0} \frac{1}{h}\, \tau_a f(h) = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h}$$

existiert, also $\tau_a f$ differenzierbar an 0 ist, und in diesem Fall ist

$$f'(a) = (\tau_a f)'(0).$$

Man kann sich also bei Bedarf bei der Untersuchung von Ableitungen auf den Fall $a = f(a) = 0$ zurückziehen.

[3]Diese Schreiweise ist zwar suggestiv und praktisch in manchen Situationen, aber problematisch, da man dabei stillschweigend davon ausgeht, dass die Variable x heißt. Außerdem handelt man sich erhebliche formale Probleme ein, wenn man die Ableitung an der Stelle x betrachten will, also so etwas wie $(df/dx)(x)$. Wir werden daher im Weiteren auf diese Notation verzichten.

Beispiel 10.3 1. Ist $f(x) := cx + b$ $(x \in \mathbb{K})$ mit Konstanten $b, c \in \mathbb{C}$, so ist f differenzierbar auf $\mathbb{K}$ und es gilt

$$f'(x) = \lim_{h\to 0} \frac{1}{h}\,\tau_x f(h) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \lim_{h\to 0} \frac{ch}{h} = c \qquad (x \in \mathbb{K})\ .$$

2. Für $f(x) := x^2$ $(x \in \mathbb{K})$ gilt

$$f'(x) = \lim_{h\to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h\to 0} \frac{2xh+h^2}{h} = \lim_{h\to 0}(2x+h) = 2x \qquad (x \in \mathbb{K})\ .$$

Allgemeiner kann man mithilfe der binomischen Formel ([Ü]) zeigen: Ist $n \in \mathbb{N}$ und $f(x) = x^n$ $(x \in \mathbb{K})$, so ist

$$f'(x) = nx^{n-1} \qquad (x \in \mathbb{K})\ .$$

3. Ist $f(x) := |x|$ $(x \in \mathbb{R})$, so ist stetig auf $\mathbb{R}$, aber nicht differenzierbar an der Stelle $a = 0$, da $f(h)/h = 1$ für $h > 0$ und $f(h)/h = -h/h = -1$ für $h < 0$ gilt. Also hat $h \mapsto f(h)/h$ keinen (beidseitigen) Grenzwert an 0. Das Beispiel zeigt, dass Stetigkeit an einer Stelle im Allgemeinen *nicht* die Differenzierbarkeit an dieser Stelle impliziert.

Satz 10.4 *Die Funktionen* exp *ist differenzierbar auf* $\mathbb{C}$ *mit* $\exp' = \exp$.

Beweis. Nach Bemerkung 9.13 ist $\varepsilon : \mathbb{C} \to \mathbb{C}$ mit

$$\varepsilon(h) := \sum_{\mu=0}^{\infty} \frac{h^\mu}{(\mu+1)!}$$

stetig auf $\mathbb{C}$ mit $\varepsilon(0) = 1$. Damit gilt für $h \in \mathbb{C}^*$

$$\frac{1}{h}(e^h - 1) = \sum_{\nu=1}^{\infty} \frac{h^{\nu-1}}{\nu!} = \varepsilon(h) \to 1 \qquad (h \to 0).$$

Für beliebiges $a \in \mathbb{C}$ folgt

$$(e^{a+h} - e^a)/h = e^a \cdot (e^h - 1)/h \to e^a \cdot 1 \quad (h \to 0).$$

□

Der erste Teil des folgenden Satzes zeigt, dass Differenzierbarkeit einer Funktion f an einer Stelle a bedeutet, dass f lokal an a mit einer gewisssen Güte durch eine affin-lineare Funktion der Form

$$h \mapsto f(a) + c \cdot h$$

approximiert werden kann. Der zweite zeigt, dass Differenzierbarkeit Stetigkeit impliziert. Wir schreiben dabei $X_a := (X - a) \setminus \{0\}$.

Satz 10.5 *Es seien $X \subset \mathbb{K}$, $a \in X \cap X'$ und $f : X \to \mathbb{C}$. Dann gilt*

1. *(Zerlegungsformel, affin-lineare Approximation) f ist genau dann differenzierbar an a, wenn ein $c \in \mathbb{C}$ und eine an 0 abklingende Funktion $\varepsilon = \varepsilon_{f,a} : X_a \to \mathbb{C}$ existieren mit*

$$f(a+h) = f(a) + c \cdot h + \varepsilon(h) \cdot h \qquad (h \in X_a) \ .$$

   Außerdem ist in diesem Fall $f'(a) = c$.

2. *Ist f differenzierbar an a, so ist f auch stetig an a.*

Beweis. 1. $\Rightarrow$: Wir setzen für $h \in X_a$

$$\varepsilon(h) := (\tau_a f)(h)/h - f'(a).$$

Dann ist ε abklingend an 0 aufgrund der Differenzierbarkeit von f an a und es gilt

$$f(a+h) = f(a) + f'(a) \cdot h + \varepsilon(h) \cdot h \qquad (h \in X_a),$$

also $c = f'(a)$
$\Leftarrow$: Klar mit $(\tau_a f)(h)/h = c + \varepsilon(h)$ für $h \in X_a$.
2. Folgt aus der Zerlegungsformel für $h \to 0$. □

Satz 10.6 *(Summen-, Produkt- und Quotientenregel)*
Es seien $X \subset \mathbb{K}$ und $f, g : X \to \mathbb{C}$ differenzierbar an der Stelle $a \in X$. Dann gilt

1. *$f + g$ ist differenzierbar an a mit*

$$(f+g)'(a) = f'(a) + g'(a) \ .$$

2. *$f \cdot g$ ist differenzierbar an a mit*

$$(f \cdot g)'(a) = f'(a) \cdot g(a) + f(a) \cdot g'(a) \ .$$

3. *Ist g nullstellenfrei, so ist f/g differenzierbar an a mit*

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

Beweis. Wie man leicht nachrechnet, gilt für $h \in X - a$

$$\tau_a(f+g)(h) = \tau_a f(h) + \tau_a g(h)$$

und

$$\tau_a(f \cdot g)(h) = g(a+h) \cdot \tau_a f(h) + f(a) \cdot \tau_a g(h).$$

Nach Satz 10.5 ist g stetig an a. Damit ergeben sich die Summenregel und die Produktregel jeweils nach Divsion durch h und Grenzwertbildung für $h \to 0$. Ist zusätzlich $Z(g) = \emptyset$, so gilt auch

$$\tau_a(1/g)(h) = \frac{-\tau_a g(h)}{g(a+h)g(a)}$$

Wieder nach Divsion durch h und Grenzwertbildung für $h \to 0$ ergibt sich die Quotientenregel für $f = 1$. Die allgemeine Aussage folgt mit der Produktregel. □

Beispiel 10.7 Ist $p(x) = \sum\limits_{\nu=0}^{d} c_\nu x^\nu$ $(x \in \mathbb{K})$ ein Polynom, so folgt aus Satz 10.6 und Beispiel 10.3

$$p'(x) = \sum_{\nu=1}^{d} \nu \cdot c_\nu x^{\nu-1} \qquad (x \in \mathbb{K}).$$

Satz 10.8 *(Kettenregel)*
Es seien $X, Y \subset \mathbb{K}$ und es sei $f : X \to Y$. Ferner sei $g : Y \to \mathbb{C}$. Ist f differenzierbar an $a \in X$ und ist g differenzierbar an $f(a)$, so ist $g \circ f$ differenzierbar an a mit

$$(g \circ f)'(a) = g'\big(f(a)\big) f'(a)\ .$$

Beweis. Wir betrachten zunächst den Spezialfall $a = 0$ und $f(0) = g(0) = 0$. Ist $\varepsilon = \varepsilon_{g,0}$ wie in der Zerlegungsformel und $\varepsilon(0) := 0$, so ist $\varepsilon \circ f$ abklingend an 0 und damit gilt für $h \in X$, $h \neq 0$

$$\begin{aligned} \frac{1}{h}(g \circ f)(h) &= \frac{1}{h}\big(g(f(h)) - g'(0)f(h)\big) + g'(0)\frac{f(h)}{h} \\ &= \varepsilon(f(h))\frac{f(h)}{h} + g'(0)\frac{f(h)}{h} \to g'(0)f'(0) \quad (h \to 0). \end{aligned}$$

Also ist $g \circ f$ differenzierbar an 0 mit $(g \circ f)'(0) = g'(0)f'(0)$.
Sind nun a sowie $f(a)$ und $g(f(a))$ beliebig, so gilt

$$\tau_a(g \circ f) = \tau_{f(a)} g \circ \tau_a f$$

und damit ergibt sich die Behauptung durch Anwendung des Spezialfalls auf die rechte Seite. □

Beispiel 10.9 Für $p : \mathbb{C} \to \mathbb{C}$ mit

$$p(z) = (z^3 + 2z + 1)^5 \quad (z \in \mathbb{C})$$

gilt $p = g \circ f$ mit $f(z) = z^3+2z+1$ und $g(w) = w^5$. Also ergibt sich aus der Kettenregel

$$p'(z) = g'(f(z))f'(z) = 5(z^3 + 2z + 1)^4(3z^2 + 2) \quad (z \in \mathbb{C}).$$

Satz 10.10 *Die Funktionen* sin *und* cos *sind differenzierbar auf* $\mathbb{C}$ *mit*

$$\sin' = \cos \quad \textit{und} \quad \cos' = -\sin .$$

Beweis. Für $z \in \mathbb{C}$ ist nach der Kettenregel

$$\sin'(z) = \frac{1}{2i}(ie^{iz} + ie^{-iz}) = \frac{1}{2}(e^{iz} + e^{-iz}) = \cos z .$$

Entsprechend ergibt sich $\cos' = -\sin$. □

Satz 10.11 *(Umkehrregel)*
Es seien $X \subset \mathbb{K}$ und $f : X \to Y \subset \mathbb{C}$ bijektiv. Ist f differenzierbar an der Stelle $a \in X$ mit $f'(a) \neq 0$ und ist die Umkehrfunktion f^{-1} stetig an $c := f(a)$, so ist f^{-1} differenzierbar an c mit

$$(f^{-1})'(c) = 1/f'(a) = 1/(f'(f^{-1}(c)) .$$

Beweis. Ist (x_n) eine Folge in X mit $a \neq x_n \to a$ $(n \to \infty)$, so gilt aufgrund der Stetigkeit von f an a und der Injektivität auch

$$f(a) \neq f(x_n) \to c \qquad (n \to \infty) .$$

Damit in $c \in Y'$. Es sei zunächst $a = c = 0$. Dann ist f^{-1} stetig an 0. Also gilt $f^{-1}(u) \to 0$ $(u \to 0)$ und folglich mit Bemerkung 5.15 und $f^{-1}(u) \neq 0$ für $u \neq 0$

$$\frac{f^{-1}(u)}{u} = \frac{f^{-1}(u)}{f(f^{-1}(u))} \to \frac{1}{f'(0)} \qquad (u \to 0) .$$

Der allgemeine Fall ergibt sich aus Bemerkung 10.2 und $\tau_c f^{-1} = (\tau_a f)^{-1}$. □

Bemerkung 10.12 Es seien $X \subset \mathbb{K}$ und $f : X \to Y \subset \mathbb{C}$ bijektiv mit stetiger Umkehrfunktion. Ist $g : Y \to \mathbb{C}$ nullstellenfrei und so, dass $f' = g \circ f$ auf X gilt, so ist nach der Umkehrregel

$$(f^{-1})' = 1/(f' \circ f^{-1}) = 1/g\ .$$

Beispiel 10.13 1. Ist $a > 0$ fest, so gilt mit $a^z = e^{z \ln a}$ und der Kettenregel

$$(z \mapsto a^z)' = (z \mapsto a^z \ln a).$$

2. Es gilt

$$\ln'(t) = 1/t \qquad (t > 0)\ .$$

(Denn: Nach Bemerkung 10.12, angewandt auf $g(t) := t$ für $t > 0$ und $f = \exp|_{\mathbb{R}}$, ist $\ln' = (f^{-1})' = 1/g$.)
Damit ergibt sich für festes $\alpha \in \mathbb{C}$ mit der Kettenregel

$$(t \mapsto t^\alpha)' = (t \mapsto e^{\alpha \ln t})' = (t \mapsto \alpha t^{\alpha - 1}).$$

3. Aus Satz 10.10 und der Quotientenregel folgt

$$\tan' = \frac{1}{\cos^2}\big(\cos^2 + \sin^2\big) = \frac{1}{\cos^2} = 1 + \tan^2$$

auf $\mathbb{C} \setminus \pi(\mathbb{Z} + 1/2)$ und entsprechend auf $\mathbb{C} \setminus \pi\mathbb{Z}$

$$\cot' = -\sin^{-2} = -1 - \cot^2.$$

4. Es gilt

$$\arctan'(t) \ = \ \frac{1}{1+t^2}, \qquad \operatorname{arccot}'(t) = -\frac{1}{1+t^2} \qquad (t \in \mathbb{R})\ .$$

(Denn: Nach 3. und Bemerkung 10.12, angewandt auf $f = \tan|_{(-\pi/2,\pi/2)}$, ist $\arctan' = (f^{-1})' = 1/g$ mit $g(t) := 1 + t^2$ für $t \in \mathbb{R}$. Entsprechendes gilt für arccot.)
5. Es gilt ([Ü])

$$\arcsin'(t) = \frac{1}{\sqrt{1-t^2}}\ , \qquad \arccos'(t) = -\frac{1}{\sqrt{1-t^2}} \qquad (t \in (-1,1)).$$

Wir beschäftigen uns nun damit, wie man die Differenzialrechnung nutzen kann, um insbesondere das lokale Verhalten differenzierbarer Funktionen genauer zu untersuchen. Dazu definieren wir zunächst, was wir unter Extremstellen verstehen.

Definition 10.14 Es seien (X,d) ein metrischer Raum und $f: X \to \mathbb{R}$.

1. Ein Punkt $x_{\max} \in X$ heißt **Maximalstelle** (von f) und $f(x_{\max})$ dann **lokales Maximum**, falls eine Umgebung U von $x_{\max}$ existiert mit
$$f(x) \le f(x_{\max}) \qquad (x \in U)\,,$$
falls also $f|_U$ an der Stelle $x_{\max}$ maximal wird. Gilt $<$ statt $\le$ für $x \neq x_{\max}$, so spricht man von einem **strikten** lokalen Maximum.

2. Ein Punkt $x_{\min} \in X$ heißt **Minimalstelle** (von f) und $f(x_{\min})$ dann ein **lokales Minimum**, falls eine Umgebung U von $x_{\min}$ existiert mit
$$f(x) \ge f(x_{\min}) \qquad (x \in U)\,,$$
falls also $f|_U$ an der Stelle $x_{\min}$ minimal wird. Gilt $>$ statt $\ge$ für $x \neq x_{\min}$, so spricht man von einem **strikten** lokalen Minimum.

Ist a eine Maximal- oder eine Minimalstelle von f, so nennt man a auch eine **Extremstelle** von f.

Wir werden im Weiteren verdeutlichen, dass die Differenzialrechung ein effizientes Instrumentarium zur Bestimmung von Extremstellen zur Verfügung stellt. Ist $f: X \to \mathbb{C}$ differenzierbar, so nennt man eine Nullstelle von f' auch **kritische Stelle** von f.

Satz 10.15 *Es seien $X \subset \mathbb{R}$ und a ein innerer Punkt von X. Ist $f: X \to \mathbb{R}$ differenzierbar an a und ist a eine Extremstelle von f, so ist a eine kritische Stelle, also $f'(a) = 0$.*

Beweis. Ohne Einschränkung sei a eine Minimalstelle. Dann existiert ein $\delta > 0$ mit $\tau_a f(h) \ge 0$ für $|h| < \delta$. Da f differenzierbar an a ist, gilt
$$0 \le \lim_{h \to 0^+} (\tau_a f)(h)/h = f'(a) = \lim_{h \to 0^-} (\tau_a f)(h)/h \le 0.$$
□

Bemerkung und Definition 10.16 1. Das Verschwinden von f' an einer Stelle a ist lediglich eine *notwendige* Bedingung dafür, dass a eine Maximal- oder Minimalstelle ist. So hat etwa $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^3$ an 0 einen kritischen Punkt, aber 0 ist keine Extremstelle. Daher nutzt man Satz 10.15 typischerweise dafür, die Extremalität von f an den inneren Punkten a, die nicht kritisch sind, auszuschließen. Kritische Stellen sind die einzigen „Kandidatinnen" für Extremstellen im Inneren von X.
2. Ist a eine Extremstelle, aber kein innerer Punkt von X, so ist a nicht notwendig eine kritische Stelle. So sind ± 1 für $f : [-1, 1] \to \mathbb{R}$ mit $f(x) = |x|$ Maximalstellen, aber keine kritischen Stellen. Außerdem ist 0 eine Minimalstelle, aber auch kein kritischer Punkt (da f an 0 gar nicht differenzierbar ist).

Satz 10.17 *(Rolle)*
Es sei $I = (\alpha, \beta)$ ein offenes Intervall, und es sei $f : I \to \mathbb{R}$ differenzierbar. Existieren $f(\alpha^+)$ und $f(\beta^-)$ und ist $f(\alpha^+) = f(\beta^-)$, so hat f einen kritischen Punkt.

Beweis. Nach Bemerkung 8.34 hat f ein Maximum oder ein Minimum. Nach Satz 10.15 hat f damit einen kritischen Punkt. □

Als Folgerung erhalten wir

Satz 10.18 *Es seien $I = (\alpha, \beta)$ ein offenes Intervall und $f, g : I \to \mathbb{R}$ differenzierbar. Existieren $f(\alpha^+)$ und $f(\beta^-)$ sowie $g(\alpha^+)$ und $g(\beta^-)$, so existiert ein $\xi \in (\alpha, \beta)$ mit*

$$f'(\xi)(g(\beta^-) - g(\alpha^+)) = (f(\beta^-) - f(\alpha^+))g'(\xi) .$$

Beweis. Wir betrachten die Funktion $\varphi : I \to \mathbb{R}$ mit

$$\varphi(x) := f(x)\big(g(\beta^-) - g(\alpha^+)\big) - \big(f(\beta^-) - f(\alpha^+)\big)g(x) \qquad (x \in I).$$

Da $\varphi(\beta^-) = \varphi(\alpha^+) = f(\alpha^+)g(\beta^-) - f(\beta^-)g(\alpha^+)$ gilt, ergibt sich die Aussage durch Anwendung des Satzes von Rolle mit φ statt f. □

Definition 10.19 Es seien V ein $\mathbb{K}$-Vektorraum und $u, v \in V$. Wir definieren $s_u^v : [0, 1] \to V$ durch

$$s_u^v(t) := u + t(v - u) \quad (t \in [0, 1])$$

und nennen s_u^v **orientierte Strecke** von u nach v. Außerdem setzen wir

$$[u,v] := s_u^v([0,1]) = \{u + t(v-u) : t \in [0,1]\}$$

und

$$(u,v) := s_u^v((0,1)) = \{u + t(v-u) : t \in (0,1)\}.$$

Damit kommen wir zu einem Satz, der gleich zwei weitere zentrale Ergebnisse der Analysis enthält.

Satz 10.20 *Es seien $a \in \mathbb{K}$ und $h \in \mathbb{K}^*$. Weiter sei $f : [a, a+h] \to \mathbb{C}$ stetig auf $[a, a+h]$ und differenzierbar auf $(a, a+h)$. Dann gilt*

1. *(Mittelwertsatz) Ist f reellwertig, so existiert ein $\xi \in (a, a+h)$ mit*

$$f(a+h) - f(a) = f'(\xi) \cdot h\ .$$

2. *(Schrankensatz) Es existiert ein $\xi \in (a, a+h)$ mit*

$$|f(a+h) - f(a)| \leq |f'(\xi)| \cdot |h|\ .$$

Beweis. 1. Ergibt sich durch Anwendung von Satz 10.18 mit $I = (0,1)$ sowie $f \circ s_a^{a+h}$ statt f und $g(x) = x$.
2. Ohne Einschränkung können wir $w := f(a+h) - f(a) \neq 0$ annehmen. Wir setzen $\zeta := |w|/w$ und betrachten die Funktion $\varphi := \mathrm{Re}(\zeta f \circ s_a^{a+h}) : [0,1] \to \mathbb{R}$. Nach 1. existiert ein $\tau \in (0,1)$ mit

$$|w| = \mathrm{Re}(\zeta w) = \varphi(1) - \varphi(0) = \varphi'(\tau) = \mathrm{Re}\big(\zeta f'(s_a^{a+h}(\tau)) \cdot h\big) \leq |f'(s_a^{a+h}(\tau))| \cdot |h|.$$

Mit $\xi := s_a^{a+h}(\tau)$ ergibt sich die Behauptung. □

Definition 10.21 Es sei V ein $\mathbb{K}$-Vektorraum. Eine Menge $X \subset V$ heißt **sternförmig** bezüglich $a \in X$, falls

$$X = \bigcup_{x \in X} [a, x]$$

gilt und kurz sternförmig, falls sie sternförmig bezüglich eines Punktes a ist. Eine nichtleere Menge in $\mathbb{R}$ ist genau dann sternförmig, wenn sie ein Intervall ist.

Satz 10.22 *Es seien $X \subset \mathbb{K}$ sternförmig und $f : X \to \mathbb{C}$. Ist f differenzierbar mit $f' = 0$, so ist f konstant.*

Beweis. Es sei X sternförmig bezüglich a. Ist $x \in X$ mit $x \neq a$, so folgt aus dem Schrankensatz mit einem $\xi \in (x, a)$

$$|f(x) - f(a)| \leq |f'(\xi)| \cdot |x - a| = 0$$

und damit $f(x) = f(a)$. Also ist f konstant ($= f(a)$) auf X. □

Satz 10.23 *Es sei I ein Intervall und $f : I \to \mathbb{R}$ sei stetig auf I und differenzierbar auf I°. Dann gilt*

1. *Ist $f' \geq 0$ auf I°, so ist f wachsend auf I.*
2. *Ist $f' \leq 0$ auf I°, so ist f fallend auf I.*

Ist dabei $f' > 0$ beziehungsweise $f' < 0$, so ist die Monotonie streng.

Beweis. 1. Sind $s, t \in I$ mit $s < t$, so ergibt sich aus dem Mittelwertsatz mit einem $\xi \in (s, t)$

$$f(t) - f(s) = f'(\xi)(t - s) \geq 0 \ ,$$

also $f(s) \leq f(t)$. Ist dabei $f'(\xi) > 0$, so ist $f(s) < f(t)$.
2. Ergibt sich durch Anwendung von 1. auf $-f$. □

Beispiel 10.24 Die Funktion $\cosh|_{\mathbb{R}}$ ist streng wachsend auf $[0, \infty)$ und streng fallend auf $(-\infty, 0]$, da $\cosh' = \sinh$ mit $\sinh > 0$ auf $(0, \infty)$ und $\sinh < 0$ auf $(-\infty, 0)$.

Der folgende Satz gibt ein *hinreichendes* Kriterien für Extremstellen.

Satz 10.25 *(Vorzeichenwechsel-Kriterium)*
Die Funktion $f : I \to \mathbb{R}$ sei stetig auf dem Intervall I und differenzierbar auf I°. Ferner sei $a \in I$. Dann gilt

1. *Existiert ein $\delta > 0$ mit $f' \leq 0$ auf $I \cap (a, a+\delta)$ und $f' \geq 0$ auf $I \cap (a-\delta, a)$, so ist a eine Maximalstelle von f.*

2. *Existiert ein $\delta > 0$ mit $f' \geq 0$ auf $I \cap (a, a+\delta)$ und $f' \leq 0$ auf $I \cap (a-\delta, a)$ so ist a eine Minimalstelle von f.*

Gilt jeweils $f' < 0$ beziehungweise $f' > 0$, so ist das Extremum strikt.

Beweis.
1. Nach Satz 10.23 ist f fallend auf $I \cap [a, a+\delta)$ und wachsend auf $I \cap (a-\delta, a]$. Damit ist a eine Maximalstelle. Gilt dabei jeweils die strikte Ungleichung, so ist das Maximum strikt.
2. Wieder durch Anwendung von 1. auf $-f$. □

Beispiel 10.26 Es sei $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = 2x^3 + 3x^2 - 1 \; .$$

Dann gilt

$$f'(x) = 6x(x+1) \begin{cases} > 0, & \text{für } x \in (-\infty, -1) \\ < 0, & \text{für } x \in (-1, 0) \\ > 0, & \text{für } x \in (0, \infty) \end{cases} .$$

Nach Satz 10.23 ist f streng wachsend auf $(-\infty, -1]$, streng fallend auf $[-1, 0]$ und streng wachsend auf $[0, \infty)$.
Zudem liegt nach Satz 10.25 an 0 ein striktes lokales Minimum und and -1 ein striktes lokales Maximum vor.

Zum Abschluss dieses Abschnitts beschäftigen wir uns mit einer eleganten Methode um gewisse Grenzwerte auszurechnen.

Satz 10.27 *(Regeln von de l'Hospital)*
Es seien $I = (a, b) \subset \mathbb{R}$ ein Intervall und $f, g : I \to \mathbb{R}$ differenzierbar mit

$$f(a^+) = g(a^+) = 0 \quad \text{oder} \quad g(x) \to \pm\infty \quad (x \to a).$$

Ist g' nullstellenfrei und gilt

$$f'(t)/g'(t) \to c \in \mathbb{R} \cup \{\pm\infty\} \qquad (t \to a),$$

so folgt

$$f(x)/g(x) \to c \qquad (x \to a) \; .$$

Eine entsprechende Aussage gilt für Grenzwerte $x \to b$.

Beweis. 1. Es gelte $f(a^+) = g(a^+) = 0$. Ist $x \in I$, so existiert nach Satz 10.18 ein $\xi(x) \in (a, x)$ mit

$$\frac{f'(\xi(x))}{g'(\xi(x))} = \frac{f(x) - f(a^+)}{g(x) - g(a^+)} = \frac{f(x)}{g(x)} \; .$$

Dabei gilt $a < \xi(x) \to a$ $(x \to a)$. Also folgt mit Bemerkung 5.15

$$\frac{f'(\xi(x))}{g'(\xi(x))} \to c \quad (x \to a).$$

2. Es gelte ohne Einschränkung $g(x) \to \infty$. Ist $\varepsilon > 0$ gegeben, so existiert ein $\delta > 0$ mit $g(t) > 0$ und $f'(t)/g'(t) \in U_\varepsilon(c)$ für $t \in U_\delta(a)$.
Es sei $s \in U_\delta(a)$ fixiert. Ist $a < x < s$, so existiert nach Satz 10.18 ein $\xi(x) \in (x, s)$ mit

$$f'(\xi(x))(g(s) - g(x)) = (f(s) - f(x))g'(\xi(x)) \; ,$$

also auch

$$\frac{f'(\xi(x))}{g'(\xi(x))} \left(\frac{g(s)}{g(x)} - 1 \right) = \frac{f(s) - f(x)}{g(x)}$$

und folglich

$$\frac{f(x)}{g(x)} = \frac{f(s)}{g(x)} + \frac{f'(\xi(x))}{g'(\xi(x))} \left(1 - \frac{g(s)}{g(x)} \right) \; .$$

Nach Voraussetzung gilt $f(s)/g(x) \to 0$ und $g(s)/g(x) \to 0$ $(x \to a^+)$. Da

$$f'(\xi(x))/g'(\xi(x)) \in U_\varepsilon(c)$$

gilt, existiert ein $\eta > 0$ mit $f(x)/g(x) \in U_{2\varepsilon}(c)$ für $x \in U_\eta(a)$. Da $\varepsilon > 0$ beliebig war, folgt die Behauptung. □

Beispiel 10.28 1. Für alle $\alpha > 0$ gilt

$$\lim_{x \to \infty} \frac{\ln x}{x^\alpha} = 0 \; ,$$

d. h. die Logarithmusfunktion wächst langsamer als jede positive Potenz von x. (Denn: Es gilt $\lim\limits_{x\to\infty} x^\alpha = \infty$ und

$$\lim_{x\to\infty} \frac{x^{-1}}{\alpha x^{\alpha-1}} = \lim_{x\to\infty} \frac{1}{\alpha x^\alpha} = 0$$

also folgt die Behauptung mit Satz 10.27)

2. Für alle $c \in \mathbb{R}$ gilt ([Ü])

$$x\ln(1 + c/x) \to c \qquad (x \to \infty)$$

und damit auch

$$\left(1 + \frac{c}{x}\right)^x \to e^c \qquad (x \to \infty).$$

11 Höhere Ableitungen

Definition 11.1 Es seien $X \subset \mathbb{K}$ mit $X \subset X'$ und $f : X \to \mathbb{C}$. Mit $f^{(0)} := f$ kann man höhere Ableitungen rekursiv definieren: Ist $n \in \mathbb{N}$ und $f^{(n-1)}$ differenzierbar, so heißt f **n-mal differenzierbar** und die Funktion

$$f^{(n)} := (f^{(n-1)})' : X \to \mathbb{C}$$

die **n-te Ableitung** von f. Dabei schreibt man meist $f'' := f^{(2)}$ und $f''' := f^{(3)}$. Ist $f^{(n)}$ stetig, so sagt man, f sei **n-mal stetig differenzierbar** und existiert $f^{(n)}$ für alle $n \in \mathbb{N}_0$, so sagt man f sei **beliebig oft differenzierbar**.

Satz 11.2 *Es seien $I \subset \mathbb{R}$ ein offenes Intervall, $f : I \to \mathbb{R}$ zweimal stetig differenzierbar und $a \in I$ eine kritische Stelle. Dann gilt:*

1. *Ist $f''(a) > 0$, so hat f an a ein striktes lokales Minimum.*
2. *Ist $f''(a) < 0$, so hat f an a ein striktes lokales Maximum.*

Beweis. 1. Da f'' stetig an a ist, existiert ein $\delta > 0$ mit $f''(x) > 0$ für alle $x \in U_\delta(a)$. Damit ist f' nach Satz 10.23 streng monoton wachsend und $U_\delta(a)$. Mit $f'(a) = 0$ folgt

$$f'(x) \begin{cases} > 0 & \text{für} \quad a < x < a + \delta \\ < 0 & \text{für} \quad a - \delta < x < a \end{cases} .$$

Nach dem Vorzeichenwechsel-Kriterium hat f an a ein stiktes lokales Minimum.
2. Ergibt sich wieder aus 1. durch Anwendung auf $-f$. $\square$

Beispiel 11.3 Wir betrachten noch einmal das Polynom f aus Beispiel 10.26, also

$$f(x) = 2x^3 + 3x^2 - 1 .$$

Hier ist

$$f''(x) = 12x + 6$$

und damit gilt $f''(0) = 6 > 0$ an der kritischen Stelle 0 und $f''(-1) = -6 < 0$ an der kritischen Stelle -1. Also ist 0 eine Minimalstelle und -1 eine Maximalstelle.

Wir haben bereits gesehen, dass Funktionen, die durch Potenzreihen darstellbar sind, stetig auf dem Konvergenzkreis sind. Wir zeigen nun viel schärfer, dass solche Funktionen tatsächlich beliebig oft differenzierbar ist. Dabei werden wir uns auf Potenzreihen mit Entwicklungsmitte $a = 0$ beschränken. Der allgemeine Fall kann problemlos darauf zurückgeführt werden.

Satz 11.4 *Es sei* $\sum_{\nu=0}^{\infty} c_\nu x^\nu$ *eine Potenzreihe mit Konvergenzradius* $R > 0$*, und es sei* $f(x) := \sum_{\nu=0}^{\infty} c_\nu x^\nu$ *für* $x \in U_R(0)$*. Dann ist* f *beliebig oft differenzierbar auf* $U_R(0)$ *und es gilt*

$$f^{(k)}(x) = \sum_{\nu=0}^{\infty} (\nu+1)(\nu+2)\dots(\nu+k)c_{\nu+k}\, x^\nu \qquad (x \in U_R(0))\,.$$

Insbesondere ist

$$f^{(k)}(0) = k!c_k \qquad (k \in \mathbb{N}_0)\,,$$

also

$$f(x) = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(0)}{\nu!} x^\nu \qquad (x \in U_R(0))\,.$$

Beweis. 1. Wir zeigen: f ist differenzierbar auf $U_R(0)$ und

$$f'(x) = \sum_{\nu=1}^{\infty} \nu c_\nu x^{\nu-1} = \sum_{\nu=0}^{\infty} (\nu+1)c_{\nu+1}x^\nu \qquad (x \in U_R(0))\,.$$

Es sei $a \in U_R(0)$. Wir wählen ein r mit $|a| < r < R$. Mit Satz 3.1 gilt

$$f(a+h) - f(a) = \sum_{\nu=1}^{\infty} c_\nu((a+h)^\nu - a^\nu) = h\sum_{\nu=1}^{\infty} \phi_\nu(h)$$

für $|h| \le \delta := r - |a|$ mit

$$\phi_n(h) := c_n \sum_{k=0}^{n-1} (a+h)^k a^{n-1-k}.$$

Aus $|(a+h)^k a^{n-1-k}| \le r^{n-1}$ für $k = 0, \dots, n$ folgt $|\phi_n(h)| \le n|c_n|r^{n-1}$.
In Bemerkung 9.13 haben wir gesehen, dass ein $q < 1$ existiert mit $|c_n|r^n \le q^n$ für n genügend groß. Nach dem Wurzelkriterium konvergiert damit die Reihe $\sum_{\nu=1}^{\infty} \nu|c_\nu|r^{\nu-1}$ und nach dem Weierstraß-Kriterium (Satz 9.11) folglich die Funktionenreihe $\sum_{\nu=1}^{\infty} \phi_\nu$

gleichmäßig auf $B_\delta(0)$. Mit Satz 9.6 ergibt sich, dass die Grenzfunktion $\phi := \sum\limits_{\nu=1}^{\infty} \phi_\nu$ stetig an 0 ist. Also gilt

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = \lim_{h\to 0} \phi(h) = \phi(0) = \sum_{\nu=1}^{\infty} \phi_\nu(0) = \sum_{\nu=1}^{\infty} \nu c_\nu a^{\nu-1}\,.$$

2. Induktiv erhält man mit Beweisschritt 1, dass $f^{(k-1)}$ für alle $k \in \mathbb{N}$ differenzierbar auf $U_R(0)$ ist mit

$$f^{(k)}(x) = \sum_{\nu=0}^{\infty} (\nu+1)(\nu+2)\ldots(\nu+k)c_{\nu+k}x^\nu\,.$$

Die Zusatzbehauptung $k!c_k = f^{(k)}(0)$ ergibt sich für $x = 0$. □

Beispiel 11.5 Für die Funktion $f : \mathbb{C} \to \mathbb{C}$ mit

$$f(z) = \begin{cases} \sin(z)/z\,, & z \neq 0 \\ 1\,, & z = 0 \end{cases}$$

gilt

$$f(z) = \sum_{\mu=0}^{\infty} \frac{(-1)^\mu z^{2\mu}}{(2\mu+1)!} = \sum_{\nu=0}^{\infty} c_\nu z^\nu$$

mit $c_k = (-1)^{k/2}/(k+1)!$ für gerade k und $c_k = 0$ für ungerade k. Insbesondere ist f nach Satz 11.4 beliebig oft differenzierbar auf $\mathbb{C}$ mit

$$f^{(k)}(0) = \begin{cases} \frac{(-1)^{k/2}}{k+1} & \text{falls } k \text{ gerade} \\ 0 & \text{falls } k \text{ ungerade} \end{cases}\,.$$

Bemerkung und Definition 11.6 Es seien $X \subset \mathbb{K}$ und $f : X \to \mathbb{C}$. Eine Funktion $F : X \to \mathbb{C}$ heißt **Stammfunktion** zu f (auf X), falls F differenzierbar ist mit $F' = f$ auf X. Ist $X \subset \mathbb{K}$ sternförmig und sind F und G Stammfunktionen zu f auf X, so ist $(F-G)' = 0$ auf X und damit ist $F - G$ nach Satz 10.22 konstant auf X, mit anderen Worten, F und G unterscheiden sich lediglich durch eine additive Konstante.

Bemerkung 11.7 Ist $\sum\limits_{\nu=0}^{\infty} c_\nu x^\nu$ eine Potenzreihe mit Konvergenzradius $R > 0$ und ist

$$f(x) = \sum_{\nu=0}^{\infty} c_\nu x^\nu \qquad (x \in U_R(0)),$$

so ist nach Satz 11.4 durch

$$F(x) := \sum_{\nu=0}^{\infty} \frac{c_\nu}{\nu+1}\, x^{\nu+1} \left(= \sum_{\nu=1}^{\infty} \frac{c_{\nu-1}}{\nu}\, x^\nu \right) \quad (x \in U_R(0))$$

eine Stammfunktion zu f auf $U_R(0)$ definiert (man beachte: der Konvergenzradius ist auch R).

Beispiel 11.8 Ist $f(z) := 1/(1-z)$ für $z \in \mathbb{C} \setminus \{1\}$, so ist durch

$$F(z) = \sum_{\nu=1}^{\infty} \frac{z^\nu}{\nu} \qquad (|z| < 1)$$

nach Bemerkung 11.7 eine Stammfunktion zu f auf $\mathbb{D}$ gegeben. Mit der Kettenregel sieht man sofort, dass auf $(-\infty, 1)$ durch

$$G(x) = \ln(1/(1-x)) \qquad (x < 1)$$

ebenfalls eine Stammfunktion zu f gegeben ist. Nach Bemerkung 11.6 sind F und G auf $I = (-1, 1)$ bis auf eine additive Konstante gleich. Da $F(0) = 0 = G(0)$ gilt, ist die additive Konstante $= 0$ und damit stimmen F und G auf I überein. Also folgt

$$\ln\left(\frac{1}{1-x}\right) = \sum_{\nu=1}^{\infty} \frac{x^\nu}{\nu} \qquad (-1 < x < 1).$$

Am Rande ihres Konvergenzkreises können Funktionen, die durch Potenzreihen definiert sind, ein sehr kompliziertes Verhalten zeigen. Konvergiert die Potenzreihe an einem Randpunkt ζ des Konvergenzkreises, so existiert jedenfalls der sogenannte radiale Randwert der Grenzfunktion an der Stelle ζ. Geanuer gilt

Satz 11.9 *(Abelscher Grenzwertsatz)*
Es sei $\sum\limits_{\nu=0}^{\infty} c_\nu x^\nu$ *eine Potenzreihe mit Konvergenzradius* $0 < R < \infty$ *und*

$$f(x) := \sum_{\nu=0}^{\infty} c_\nu x^\nu \qquad (x \in U_R(0)).$$

Ist für ein $\zeta \in K_R(0)$ die Reihe $\sum\limits_{\nu=0}^{\infty} c_\nu \zeta^\nu$ konvergent, so gilt

$$\lim_{r \to 1^-} f(r\zeta) = \sum_{\nu=0}^{\infty} c_\nu \zeta^\nu \ .$$

Beweis. Ohne Einschränkung können wir $R = 1$ und $\zeta = 1$ annehmen (ansonsten betrachte man $g(x) := f(\zeta x)$).
Wir setzen $s_n := \sum\limits_{\nu=0}^{n} c_\nu$ für $n \in \mathbb{N}_0$ und $s := \sum\limits_{\nu=0}^{\infty} c_\nu = \lim s_n$. Da (s_n) beschränkt ist, hat die Potenzreihe $\sum\limits_{\nu=0}^{\infty} s_\nu x^\nu$ Konvergenzradius ≥ 1. Also gilt mit $s_{-1} := 0$ für $|x| < 1$

$$(1-x)\sum_{\nu=0}^{\infty} s_\nu x^\nu = \sum_{\nu=0}^{\infty} s_\nu x^\nu - \sum_{\nu=0}^{\infty} s_\nu x^{\nu+1} = \sum_{\nu=0}^{\infty}(s_\nu - s_{\nu-1})x^\nu = \sum_{\nu=0}^{\infty} c_\nu x^\nu = f(x) \ .$$

Es sei $\varepsilon > 0$ gegeben. Dann existiert ein $n \in \mathbb{N}$ mit $|s_\nu - s| < \varepsilon$ für alle $\nu > n$. Für $0 < r < 1$ ergibt sich mit $1 = (1-r)\sum\limits_{\nu=0}^{\infty} r^\nu$

$$\begin{aligned} |f(r) - s| &= |(1-r)\sum_{\nu=0}^{\infty}(s_\nu - s)r^\nu| \\ &\leq (1-r)\sum_{\nu=0}^{n}|s_\nu - s| + \varepsilon(1-r)\sum_{\nu=n+1}^{\infty} \leq (1-r)\sum_{\nu=0}^{n}|s_\nu - s|r^\nu + \varepsilon. \end{aligned}$$

Aus $(1-r)\sum\limits_{\nu=0}^{n}|s_\nu - s| \to 0$ für $r \to 1^-$ folgt die Existenz eines $\delta > 0$ mit $|f(r)-s| < 2\varepsilon$ für $1 - \delta < r < 1$. □

Beispiel 11.10 Nach Beispiel 11.8 ist

$$\ln\left(\frac{1}{1-x}\right) = \sum_{\nu=1}^{\infty} \frac{x^\nu}{\nu} \qquad \big(x \in (-1,1)\big).$$

Da die alternierende Reihe $\sum\limits_{\nu=1}^{\infty}(-1)^\nu/\nu$ nach dem Leibniz-Kriterium konvergiert, ergibt sich mit dem Abelschen Grenzwertsatz für $\zeta = -1$

$$\ln(1/2) = \lim_{r \to 1^-} \ln\left(\frac{1}{1+r}\right) = \sum_{\nu=1}^{\infty} \frac{(-1)^\nu}{\nu}.$$

12 Integralrechnung

Die Integralrechnung entstand ursprünglich aus der Frage nach der Definition und der Berechnung von Flächeninhalten. Ähnlich wie bei der Differenzialrechnung werden wir Integrale über einen gewissen Grenzprozess einführen. Dazu betrachten wir zunächst besonders einfache Funktionen, für die wir die „orientierte Fläche unter den Graphen“ in sehr natürlicher Weise über die Flächen von Rechtecken definieren können.

Definition 12.1 Sind (X, d) ein metrischer Raum und $M \subset X$, so heißt

$$\operatorname{diam}(M) = \sup\{d(x,y) : x, y \in M\}$$

(mit $\operatorname{diam}(\emptyset) := 0$) der **Durchmesser** von M. Ist $I \subset \mathbb{R}$ ein Intervall, so schreiben wir auch

$$|I| := \operatorname{diam}(I)$$

und nennen $|I|$ die **Länge** von I.

Bemerkung und Definition 12.2 Es sei $[\alpha, \beta] \subset \mathbb{R}$ ein kompaktes Intervall.
1. Eine endliche Familie $(I_j)_{j\in J}$ *paarweise disjunkter* Intervalle mit

$$\bigcup_{j\in J} I_j = [\alpha, \beta]$$

nennt man eine **Zerlegung** des Intervalls $[\alpha, \beta]$. Ist $(I_j)_{j\in J}$ eine Zerlegung von $[\alpha, \beta]$ und ist $I \subset [\alpha, \beta]$ ein weiteres Intervall, so gilt (mit $|\emptyset| := 0$)

$$|I| = \sum_{j\in J} |I \cap I_j|.$$

2. Eine Funktion $\varphi \in B[\alpha, \beta] := B([\alpha, \beta], \mathbb{C})$ heißt **Treppenfunktion** (auf $[\alpha, \beta]$), falls eine Zerlegung $(I)_{I\in E}$ von $[\alpha, \beta]$ und Konstanten $c(I) = c_\varphi(I) \in \mathbb{C}$ existieren mit

$$\varphi = \sum_{I\in E} c(I) \cdot 1_I = \sum_{I\in E} c(I) \cdot 1_{I,[\alpha,\beta]},$$

also so, dass φ konstant $= c(I)$ auf I ist. Eine Zerlegung, zu der entsprechende Konstanten $c(I)$ existieren, nennen wir **zulässig** für die Funktion φ. Wir schreiben $T[\alpha, \beta]$ für die Menge der Treppenfunktionen auf $[\alpha, \beta]$.

Bemerkung und Definition 12.3 Es seien $(I)_{I\in E}$ bzw. $(J)_{J\in F}$ Zerlegungen von $[\alpha,\beta]$. Ist

$$G := \{(I,J) \in E \times F : I \cap J \neq \emptyset\},$$

so heißt $(I\cap J)_{(I,J)\in G}$ die **gemeinsame Verfeinerung** von $(I)_{I\in E}$ und $(J)_{J\in F}$. Aus Satz 1.15 ergibt sich, dass die gemeinsame Verfeinerung ebenfalls eine Zerlegung von $[\alpha,\beta]$ ist.
Sind $(I)_{I\in E}$ und $(J)_{J\in F}$ zulässig für φ, so ist auch die gemeinsame Verfeinerung zulässig für φ und es gilt $\varphi|_{I\cap J} = c(I) = c(J)$ für $(I,J) \in G$. Mit Bemerkung 12.2 ergibt sich

$$\sum_{I\in E} c(I)|I| = \sum_{(I,J)\in G} c(I)|I\cap J| = \sum_{(I,J)\in G} c(J)|I\cap J| = \sum_{J\in F} c(J)|J|\ .$$

Definition 12.4 1. Ist $\varphi : [\alpha,\beta] \to \mathbb{C}$ eine Treppenfunktion wie in D. 12.2, so heißt

$$\int_\alpha^\beta \varphi := \int_\alpha^\beta \varphi(t)\,dt := \sum_{I\in E} c_I|I|\ ,$$

Integral von φ (auf $[\alpha,\beta]$). (Man beachte dabei: Die Summe auf der rechten Seite ist nach Bemerkung 12.3 unabhängig von der Wahl der Zerlegung).
2. Ist $U \subset B[\alpha,\beta]$ ein Unterraum und ist ℓ ein lineares Funktional auf U (also eine lineare Abbildung von U nach $\mathbb{C}$), so sagen wir, ℓ sei **nichtnegativ**, falls $\ell(f) \geq 0$ für alle f mit $f \geq 0$ gilt. In diesem Fall ist allgemeiner $\ell(f) \leq \ell(g)$ für alle reellwertigen f, g mit $f \leq g$.

Satz 12.5 *Die Abbildung* $\int_\alpha^\beta : T[\alpha,\beta] \to \mathbb{C}$ *ist linear und nichtnegativ. Außerdem gilt für jede Treppenfunktion* φ*:*

1. $|\varphi|$ *ist eine Treppenfunkton mit*

$$\Big|\int_\alpha^\beta \varphi\Big| \leq \int_\alpha^\beta |\varphi| \leq ||\varphi||_\infty(\beta-\alpha)\ .$$

2. *Für* $\tau \in [\alpha,\beta]$ *ist*

$$\int_\alpha^\beta \varphi = \int_\alpha^\tau \varphi + \int_\tau^\beta \varphi\ .$$

Beweis. Es seien φ, ψ Treppenfunktionen und $\lambda \in \mathbb{C}$. Sind $(I)_{I\in E}$ bzw. $(J)_{J\in J}$ zulässige Zerlegungen für φ beziehungsweise ψ, so ist die gemeinsame Verfeinerung $(I\cap J)_{(I,J)\in G}$ sowohl für φ als auch für ψ zulässig. Sind $c_\varphi(I) \in \mathbb{C}$ bzw. $c_\psi(J) \in \mathbb{C}$ wie in Bemerkung 12.2 für φ bzw. ψ, so ist $\lambda\varphi + \psi$ konstant $= \lambda c_\varphi(I) + c_\psi(J)$ auf $I \cap J$ für $(I, J) \in G$. Also ist $\lambda\varphi + \psi$ eine Treppenfunktion (und damit $T[\alpha, \beta]$ ein Unterraum von $B[\alpha, \beta]$) und es gilt

$$\begin{aligned}
\int\limits_\alpha^\beta (\lambda\varphi + \psi) &= \sum_{(I,J)\in G} (\lambda c_\varphi(I) + c_\psi(J))|I \cap J| \\
&= \lambda \sum_{(I,J)\in G} c_\varphi(I)|I \cap J| + \sum_{(I,J)\in G} c_\psi(J)|I \cap J| = \lambda \int\limits_\alpha^\beta \varphi + \int\limits_\alpha^\beta \psi\,.
\end{aligned}$$

Die Nichtnegativität und 1. ergeben sich unmittelbar aus der Definition und entsprechenden Eigenschaften von Summen. Die Aussage 2. folgt mit

$$\varphi = \varphi \cdot 1_{[\alpha,\tau]} + \varphi \cdot 1_{(\tau,\beta]}$$

und $\int_\alpha^\tau \varphi = \int_\alpha^\beta \varphi \cdot 1_{[\alpha,\tau]}$ sowie $\int_\tau^\beta \varphi = \int_\alpha^\beta \varphi \cdot 1_{(\tau,\beta]}$ aus der Linearität. □

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.

Bemerkung und Definition 12.6 Eine Funktion $f \in B[\alpha, \beta]$ heißt **Regelfunktion** (auf $[\alpha, \beta]$), falls eine Folge (φ_n) von Treppenfunktionen existiert mit $||f - \varphi_n||_\infty \to 0$ für $n \to \infty$, also $\varphi_n \to f$ gleichmäßig auf $[\alpha, \beta]$. Wir schreiben $R[\alpha, \beta]$ für die Menge der Regelfunktionen. Damit ist $R[\alpha, \beta]$ der Abschluss von $T[\alpha, \beta]$ in $B[\alpha, \beta]$.

Satz 12.7 *Es sei $f : [0, 1] \to \mathbb{C}$ stetig. Dann ist mit $I_0^{(n)} := \{0\}$ und*

$$I_j^{(n)} := ((j-1)/n, j/n] \quad (j = 1, \ldots, n)$$

für $n \in \mathbb{N}$ durch

$$\varphi_n(t) := f(j/n) \qquad (t \in I_j^{(n)},\ j = 0, \ldots, n)$$

eine Folge (φ_n) von Treppenfunktionen definiert mit $||\varphi_n - f||_\infty \to 0$. Insbesondere ist f eine Regelfunktion.

Beweis. Es sei $\varepsilon > 0$ gegeben. Da f gleichmäßig stetig auf $[0,1]$ ist, existiert ein $\delta > 0$ so, dass $|f(t) - f(s)| < \varepsilon$ für alle t, s mit $|t - s| < \delta$. Ist $n > 1/\delta$ und ist $t \in [0,1]$, so existiert ein $j = j_{n,t}$ so, dass $t \in I_j^{(n)}$ und damit (da $0 \leq j/n - t < 1/n < \delta$)

$$|f(t) - \varphi_n(t)| = |f(t) - f(j/n)| < \varepsilon\,.$$

Da t beliebig war, folgt $||f - \varphi_n||_\infty \leq \varepsilon$ für $n > 1/\delta$. □

Bemerkung 12.8 Aus Satz 12.7 folgt leicht, dass $C[\alpha,\beta] \subset R[\alpha,\beta]$ für beliebige kompakte Intervalle $[\alpha,\beta]$ gilt. Man kann zeigen ([Ü]), dass auch monotone Funktionen $f : [\alpha,\beta] \to \mathbb{R}$ stets Regelfunktionen sind.

Bemerkung und Definition 12.9 Es seien f eine Regelfunktion und (φ_n) eine Folge von Treppenfunktionen mit $\varphi_n \to f$ gleichmäßig auf $[\alpha,\beta]$. Dann gilt:

1. Die Folge $(\int_\alpha^\beta \varphi_n)_n$ konvergiert in $\mathbb{C}$, denn für $n, n' \in \mathbb{N}$ gilt nach Satz 12.5

$$\Big|\int_\alpha^\beta \varphi_n - \int_\alpha^\beta \varphi_{n'}\Big| = \Big|\int_\alpha^\beta (\varphi_n - \varphi_{n'})\Big| \leq ||\varphi_n - \varphi_{n'}||_\infty(\beta - \alpha)$$

   und da (φ_n) eine Cauchy-Folge in $B[\alpha,\beta]$ ist, ist auch $(\int_\alpha^\beta \varphi_n)$ eine Cauchy-Folge in $\mathbb{C}$, also konvergent.

2. Ist (ψ_n) eine weitere Folge von Treppenfunktionen mit $\psi_n \to f$ gleichmäßig auf $[\alpha,\beta]$, so gilt

$$\Big|\int_\alpha^\beta \varphi_n - \int_\alpha^\beta \psi_n\Big| \leq ||\varphi_n - \psi_n||_\infty(\beta-\alpha) \leq \big(||\varphi_n - f||_\infty + ||f - \psi_n||_\infty\big)(\beta-\alpha) \to 0\,,$$

   also $\lim\limits_{n\to\infty} \int_\alpha^\beta \varphi_n = \lim\limits_{n\to\infty} \int_\alpha^\beta \psi_n$.

Damit setzen wir

$$\int_\alpha^\beta f := \int_\alpha^\beta f(t)\,dt := \lim_{n\to\infty} \int_\alpha^\beta \varphi_n$$

und nennen $\int_\alpha^\beta f$ das **(Regel-)Integral** (oder auch **Cauchyintegral**) von f auf $[\alpha,\beta]$. Nach 2. ist dabei der Wert unabhängig von der speziellen Wahl der Treppenfunktionenfolge.

Beispiel 12.10 Wir betrachen $f(t) = t$ auf $[0,1]$. Dann ist nach Satz 12.7 durch

$$\varphi_n(t) := \begin{cases} j/n & , \quad t \in ((j-1)/n, j/n],\ j = 1, \ldots, n \\ 0 & , \quad t = 0 \end{cases}$$

eine Folge von Treppenfunktionen auf $[0,1]$ gegeben mit $\varphi_n \to f$ gleichmäßig auf $[0,1]$. Es gilt

$$\int_0^1 \varphi_n = \sum_{j=1}^n \frac{j}{n} \cdot |I_j^{(n)}| = \frac{1}{n} \sum_{j=1}^n \frac{j}{n} = \frac{n(n+1)}{2n^2} = \frac{1}{2} + \frac{1}{2n} \to \frac{1}{2} \quad (n \to \infty)\ ,$$

also ist

$$\int_0^1 f = \int_0^1 t\,dt = 1/2.$$

Wir stellen einige Rechenregeln für Regelintegrale zusammen, die sich aus der Approximation durch Treppenfunktionen ergeben.

Satz 12.11 *Die Abbildung $\int_\alpha^\beta : R[\alpha,\beta] \to \mathbb{C}$ ist linear und nichtnegativ. Außerdem gilt*

1. *$|f|$ ist eine Regelfunktion und*

$$\Big| \int_\alpha^\beta f \Big| \leq \int_\alpha^\beta |f| \leq ||f||_\infty (\beta - \alpha).$$

2. *Für $\tau \in [\alpha,\beta]$ sind $f|_{[\alpha,\tau]} \in R[\alpha,\tau]$ und $f|_{[\tau,\beta]} \in R[\tau,\beta]$, und es gilt*

$$\int_\alpha^\beta f = \int_\alpha^\tau f + \int_\tau^\beta f\ .$$

Beweis. Es seien $f, g \in R[\alpha,\beta]$ und $\lambda \in \mathbb{C}$. Nach Voraussetzung existieren Folgen von Treppenfunktionen (φ_n) und (ψ_n) mit $\varphi_n \to f$ und $\psi_n \to g$ gleichmäßig auf $[\alpha,\beta]$ und damit

$$||\lambda f + g - (\lambda\varphi_n + \psi_n)||_\infty \leq |\lambda| \cdot ||f - \varphi_n||_\infty + ||g - \psi_n||_\infty \to 0$$

für $n \to \infty$. Also ist $\lambda f + g \in R[\alpha, \beta]$ und mit Satz 12.5 gilt

$$\int_\alpha^\beta \lambda f + g = \lim_{n\to\infty} \int_\alpha^\beta (\lambda\varphi_n + \psi_n) = \lambda \int_\alpha^\beta f + \int_\alpha^\beta g \ .$$

Es sei nun $f \geq 0$. Wir setzen

$$\varphi_n^+ := \varphi_n + ||f - \varphi_n||_\infty \ .$$

Dann sind φ_n^+ Treppenfunktionen mit $\varphi_n^+ \geq f \geq 0$ sowie $||f - \varphi_n^+||_\infty \to 0$. Also folgt mit Satz 12.5

$$\int_\alpha^\beta f = \lim_{n\to\infty} \int_\alpha^\beta \varphi_n^+ \geq 0 \ .$$

1. Aus $\big| |f| - |\varphi_n| \big| \leq |f - \varphi_n|$ folgt, dass auch $|f|$ eine Regelfunktion ist und dass $|\varphi_n| \to |f|$ gleichmäßig auf $[\alpha, \beta]$ gilt. Damit ergibt sich mit Satz 12.5.1 und $|f| \leq ||f||_\infty$

$$\Big|\int_\alpha^\beta f\Big| = \lim_{n\to\infty} \Big|\int_\alpha^\beta \varphi_n\Big| \leq \lim_{n\to\infty} \int_\alpha^\beta |\varphi_n| = \int_\alpha^\beta |f| \leq ||f||_\infty(\beta - \alpha) \ .$$

2. ergibt sich wie oben aus Satz 12.5.2 durch Grenzübergang $\varphi_n \to f$. □

Wir kommen zu zentralen Sätzen der eindimensionalen Analysis, die die Beziehung zwischen der Differenzial- und der Integralrechnung herstellen. Wir setzen dabei für $f \in R[\alpha, \beta]$

$$\int_\beta^\alpha f := -\int_\alpha^\beta f$$

und für allgemeine Intervalle I

$$R(I) := \{f : I \to \mathbb{C} : f|_{[\alpha,\beta]} \in R[\alpha, \beta] \text{ für alle } [\alpha, \beta] \subset I\}.$$

Damit gilt

$$\int_u^w f = \int_u^v f + \int_v^w f$$

für beliebige $u, v, w \in I$ und $f \in R(I)$.

Satz 12.12 *(Hauptsatz über Integralfunktionen)*
Es seien I ein Intervall, $f \in R(I)$ und $u \in I$. Dann ist die Funktion $Vf = V_u f : I \to \mathbb{C}$, definiert durch

$$(Vf)(x) := \int_u^x f \qquad (x \in I),$$

stetig auf I und differenzierbar an allen Stetigkeitstellen x von f mit

$$(Vf)'(x) = f(x)\ .$$

Beweis. Es sei $x \in I$ beliebig. Dann existiert ein $\delta > 0$ so, dass $J := I \cap [x-\delta, x+\delta]$ ein kompaktes Intervall ist. Also ist f beschränkt auf J. Für $h \in J - x$ gilt dann

$$|(Vf)(x+h) - (Vf)(x)| = \Big| \int_u^{x+h} f - \int_u^x f \Big| = \Big| \int_x^{x+h} f \Big| \leq \|f\|_{\infty,J} \cdot |h| \to 0 \quad (h \to 0).$$

Es sei nun f stetig an der Stelle x. Dann gilt für $h \neq 0$

$$\Big| \frac{(Vf)(x+h) - (Vf)(x)}{h} - f(x) \Big| = \Big| \frac{1}{h} \int_x^{x+h} (f - f(x)) \Big| \leq \big\| f - f(x) \big\|_{\infty,[x,x+h]} \to 0$$

für $h \to 0$. □

Bemerkung und Definition 12.13 Wir nennen die Funktion $Vf = V_u f$ aus Satz 12.12 die **Integralfunktion**[4] von f bezüglich u. Ist $w \in I$, so unterscheiden sich die Funktionen $V_u f$ und $V_w f$ lediglich durch eine additive Konstante (genauer ist $V_u f = V_w f + \int_u^w f$). Nach Satz 12.12 ist Vf im Falle einer *stetigen* Funktion f auch eine Stammfunktion zu f auf I. Für nichtstetige f sind Integralfunktionen nicht stets Stammfunktionen: Ist etwa $f = 1_{[0,\infty)}$, so ist $V_0 f = \mathrm{id}_{\mathbb{R}} f$ nicht differenzierbar an der Sprungstelle 0 von f.

Der folgende Satz beinhaltet *das* zentrale Ergebnis zur Berechnung von Integralen.

[4] Die lineare Abbildung $V : R(I) \to C(I)$ nennt man Volterra-Operator auf $R(I)$; daher das V.

Satz 12.14 *(Hauptsatz der Differenzial- und Integralrechnung)*
Es seien I ein Intervall $f : I \to \mathbb{C}$ stetig. Ist F eine Stammfunktion zu f auf I, so ist

$$\int_u^v f = F(v) - F(u) =: F(t)\Big|_u^v =: F\Big|_u^v$$

für alle $u, v \in I$.

Beweis. Da f stetig ist, ist auch $V_u f$ eine Stammfunktion zu f auf I. Nach Bemerkung 11.6 ist die Differenz $F - V_u f$ konstant auf I. Damit ergibt sich

$$\int_v^u f = (V_u f)(v) = (V_u f)(v) - (V_u f)(u) = F(v) - F(u)$$

für alle $u, v \in I$. □

Beispiel 12.15 1. Es sei $f(t) = 1/t$ $(t > 0)$. Dann ist $t \mapsto \ln(t)$ eine Stammfunktion zu f auf $(0, \infty)$. Nach dem Hauptsatz der Differenzial- und Integralrechnung gilt für $0 < u, v < \infty$

$$\int_u^v f = \int_u^v \frac{1}{t}\, dt = \ln t\Big|_u^v = \ln(v) - \ln(u) \quad \left(= \ln\left(\frac{v}{u}\right)\right) .$$

2. Für $\alpha \in \mathbb{C} \setminus \{-1\}$ sei $f(t) = t^\alpha$ $(t > 0)$. Dann ist $t \mapsto \frac{1}{\alpha+1} t^{\alpha+1}$ eine Stammfunktion zu f auf $(0, \infty)$ und folglich ist für $0 < u, v < \infty$

$$\int_u^v t^\alpha\, dt = \frac{1}{\alpha+1} t^{\alpha+1}\Big|_u^v = \frac{1}{\alpha+1}\left(v^{\alpha+1} - u^{\alpha+1}\right) .$$

Im Falle $\mathrm{Re}(\alpha) \geq 0$ gilt dies auch für $u = 0$.

Bemerkung 12.16 Es sei I ein Intervall.
1. (Substitutionsregel) Es seien $\gamma : I \to \mathbb{K}$ stetig differenzierbar und ist $f : \gamma(I) \to \mathbb{C}$. Ist F eine Stammfuktion zu f auf $\gamma(I)$, so ist $F \circ \gamma$ nach der Kettenregel eine Stammfunktion zu $(f \circ \gamma)\gamma'$ auf I. Ist dabei $\mathbb{K} = \mathbb{R}$ und f stetig, so gilt für $u, v \in I$

$$\int_u^v (f \circ \gamma)\gamma' = \int_{\gamma(u)}^{\gamma(v)} f .$$

(Denn: Nach dem Hauptsatz über Integralfunktionen existiert eine Stammfunktion F zu f auf (dem Intervall) $\gamma(I)$ und damit haben beide Integrale nach dem Hauptsatz der Differenzial- und Integralrechnung den Wert $F(\gamma(v)) - F(\gamma(u))$.)
2. (partielle Integration) Sind $f, g : I \to \mathbb{C}$ 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. Sind f, g stetig, so ergibt sich mit dem Hauptsatz der Differenzial- und Integralrechnung für $u, v \in I$

$$\int_u^v fG = FG\big|_u^v - \int_u^v Fg \; .$$

Man kann also unter Umständen die Berechnug des Integrals $\int_u^v fG$ auf die von $\int_u^v Fg$ zurückführen.

Beispiel 12.17 1. Mit $\gamma(t) := t^2$ auf $(-1, 1)$ gilt für $u, v \in (-1, 1)$

$$\int_u^v \frac{2t}{\sqrt{1-t^2}}\, dt = \int_{u^2}^{v^2} \frac{1}{\sqrt{1-s}}\, ds = -2\sqrt{1-s}\,\Big|_{u^2}^{v^2} \; .$$

2. Für $\alpha \neq -1$ und $u, v > 0$ gilt

$$\int_u^v t^\alpha \ln t \, dt = \frac{t^{\alpha+1}}{\alpha+1} \ln t\Big|_u^v - \frac{1}{\alpha+1} \int t^\alpha \, dt = \frac{t^{\alpha+1}}{\alpha+1}\Big(\ln t - \frac{1}{\alpha+1}\Big)\Big|_u^v \; .$$

Bemerkung 12.18 Sind $a, b \in \mathbb{R}$, so ergibt sich mit der Substitutionsregel und $\gamma = s_a^b$ für $f \in C[a, b]$

$$\int_a^b f = \int_0^1 (f(s_a^b(t))(b-a)\, dt = (b-a) \int_0^1 f(a + t(b-a))\, dt.$$

Wir setzen für beliebige $a, b \in \mathbb{K}$ und $f \in C[a, b]$

$$\int_a^b f := (b-a) \int_0^1 f(a + t(b-a))\, dt.$$

Ist dabei F eine Stammfunktion zu f auf $[a, b]$, so ist $F \circ s_a^b$ eine Stammfunktion zu $(b-a)(f \circ s_a^b)$ auf $[0, 1]$ und damit nach dem Hauptsatz der Differenzial- und Integralrechnung

$$\int_a^b f = F(b) - F(a).$$

Satz 12.19 *(Taylor)*
Es seien $X \subset \mathbb{K}$ sternförmig bezüglich a und $n \in \mathbb{N}_0$. Ist f $(n+1)$-mal stetig differenzierbar auf X, so gilt für $h \in X - a$

$$f(a+h) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(a)}{\nu!} h^\nu + \frac{h^{n+1}}{n!} \int_0^1 (1-t)^n f^{(n+1)}(a+th)\,dt\,.$$

Beweis. Wir beweisen die Behauptung per Induktion nach n. Für $n = 0$ und f stetig differenzierbar auf X ergibt sich mit Bemerkung 12.18

$$f(a+h) - f(a) = h \int_0^1 f'(a+th)\,dt\,.$$

Gilt die Behauptung für $n-1$ und ist f $(n+1)$-mal stetig differenzierbar, so folgt mit partieller Integration

$$\begin{aligned} f(a+h) \quad - \quad \sum_{\nu=0}^{n-1} \frac{f^{(\nu)}(a)}{\nu!} h^\nu &= \frac{h^n}{(n-1)!} \int_0^1 (1-t)^{n-1} f^{(n)}(a+th)\,dt \\ &= \frac{h^n}{(n-1)!} \Big(- \frac{(1-t)^n}{n} f^{(n)}(a+th)\Big|_0^1 + h \int_0^1 \frac{(1-t)^n}{n} f^{(n+1)}(a+th)\,dt \Big) \\ &= \frac{h^n}{n!} f^{(n)}(a) + \frac{h^{n+1}}{n!} \int_0^1 (1-t)^n f^{(n+1)}(a+th)\,dt\,. \end{aligned}$$

□

Bemerkung und Definition 12.20 In der Situation von Satz 12.19 heißen

$$c_{k,a}(f) := f^{(k)}(a)/k!$$

für $k = 0, \ldots, n$ der k-te **Taylor-Koeffizient** von f bezüglich a, das Polynom $T_{n,a}f$, definiert durch

$$(T_{n,a}f)(h) := \sum_{\nu=0}^{n} c_{\nu,a}(f) h^\nu \qquad (h \in \mathbb{C}),$$

das n-te **Taylor-Polynom** von f bezüglich a und

$$(R_{n,a}f)(h) := f(a+h) - (T_{n,a}f)(h) \qquad (h \in X - a)$$

das n-te **Restglied**. Ist f beliebig oft differenzierbar, so heißt die Funktionenreihe $\sum\limits_{\nu=0}^{\infty} c_{\nu,a}(f)h^{\nu}$ die **Taylor-Reihe** von f bezüglich a.

Aus Satz 11.4 ergibt sich: Ist $\sum\limits_{\nu=0}^{\infty} c_{\nu}(x-a)^{\nu}$ eine Potenzreihe mit positivem Konvergenzradius und f die Grenzfunktion auf dem Konvergenzkreis, so ist $c_k = c_{k,a}(f)$ für alle $k \in \mathbb{N}_0$ und damit $\sum\limits_{\nu=0}^{\infty} c_{\nu}h^{\nu}$ die Taylor-Reihe von f bezüglich a.

Bemerkung 12.21 (Lagrangeform des Restgliedes) Man kann zeigen ([Ü]), dass folgende Variante eines Mittelwertsatzes für Integrale gilt: Ist $[\alpha,\beta] \subset \mathbb{R}$ und sind $\varphi : [\alpha,\beta] \to \mathbb{R}$ stetig sowie $\psi \in R[\alpha,\beta]$ mit $\psi \geq 0$ auf $[\alpha,\beta]$, so existiert ein $\tau \in [\alpha,\beta]$ mit

$$\int\limits_{\alpha}^{\beta} \varphi\psi = \varphi(\tau) \int\limits_{\alpha}^{\beta} \psi \ .$$

Wendet man dies auf $\varphi(t) := f^{(n+1)}(a+th)$ und $\psi(t) := (1-t)^n$ mit $[\alpha,\beta] = [0,1]$ an und beachtet man dabei, dass

$$\int\limits_{0}^{1} \psi = \int\limits_{0}^{1} (1-t)^n \, dt = \frac{1}{n+1}$$

gilt, so ergibt sich unter den Bedingungen des Taylor-Satzes für reellwertige $f^{(n+1)}$ und $h \in X - a$ die Existenz eines $\tau = \tau_{n,a,h}(f) \in [0,1]$ mit

$$(R_{n,a}f)(h) = \frac{f^{(n+1)}(a+\tau h)}{(n+1)!} \, h^{n+1} \ .$$

13 Uneigentliche Integrale

Wir haben bisher nur Integrale auf kompakten Intervallen definiert. Wir wollen jetzt auch nichtkompakte Intervalle betrachten.

Bemerkung und Definition 13.1 Es sei I ein Intervall und $a := \inf I$, $b := \sup I$. Eine Funktion $f \in R(I)$ heißt **integrierbar** auf I, falls $(V_u f)(a^+)$ und $(V_u f)(b^-)$ für ein $u \in I$ existieren. In diesem Fall existieren die beiden Grenzwerte für jedes $u \in I$ und die Differenz $(V_u f)(b^-) - (V_u f)(a^+)$ ist nach Bemerkung 12.13 unabhängig von u. Man sagt dann auch, dass das uneigentliche Integral $\int_{a^+}^{b^-} f$ existiert (beziehungeweise konvergiert) und die Zahl

$$\int_{a^+}^{b^-} f := \int_{a^+}^{b^-} f(t)\,dt := (V_u f)(b^-) - (V_u f)(a^+)$$

heißt (**uneigentliches**) **Integral** von f auf I. Ist $b = \infty$, so schreibt man meist ∞ statt ∞^- und entsprechend für $a = -\infty$. Aus Satz 5.9, Bemerkung 5.10 und Satz 12.11 folgt leicht, dass durch $f \mapsto \int_{a^+}^{b^-} f$ ein lineares und nichtnegatives Funktional auf der Menge der auf I integrierbaren Funktionen definiert ist.

Bemerkung 13.2 1. Ist $b \in I$, so gilt $(V_u f)(b) = (V_u f)(b^-)$ nach dem Hauptsatz über Integralfunktionen. Entsprechend ist $(V_u f)(a) = (V_u f)(a^+)$ im Falle $a \in I$. Also ist im Falle $I = [a, b]$

$$\int_{a^+}^{b^-} f = (V_u f)(b^-) - (V_u f)(a^+) = \int_a^b f,$$

d. h. „eigentliches“ und uneigentliches Integral stimmen überein. Man schreibt daher manchmal auch (und im Falle $b \in I$ meist) kurz b statt b^-. Entsprechendes gilt für a.
2. (erweiterter Hauptsatz der Differenzial- und Integralrechnung) Es sei $f : I \to \mathbb{C}$ stetig. Ist F eine beliebige Stammfunktion zu f auf I, so ist $F - V_u f$ konstant auf I. Also ist f genau dann integrierbar, wenn $F(b^-)$ und $F(a^+)$ existieren. Außerdem gilt dann

$$\int_{a^+}^{b^-} f = F(b^-) - F(a^+) =: F(t)|_a^b.$$

Beispiel 13.3 1. Für $\alpha \in \mathbb{R}$ sei $f_\alpha(t) := t^{-\alpha}$ auf $I = (0, \infty)$. Dann ist durch

$$F_\alpha(t) = \begin{cases} \frac{t^{1-\alpha}}{1-\alpha}, & \alpha \neq 1 \\ \ln(t), & \alpha = 1 \end{cases}$$

eine Stammfunktion F_α zu f_α auf I definiert. Außerdem erhalten wir für $t \to \infty$

$$F_\alpha(t) \to \begin{cases} 0, & \alpha > 1 \\ \infty, & \alpha \leq 1 \end{cases}$$

und für $t \to 0^+$

$$F_\alpha(t) \to \begin{cases} 0, & \alpha < 1 \\ -\infty, & \alpha \geq 1 \end{cases}.$$

Also ist f_α nach Bemerkung 13.2.2 genau dann integrierbar auf $[1, \infty)$, wenn $\alpha > 1$ ist und in diesem Falle ist

$$\int\limits_1^\infty t^{-\alpha}\, dt = F_\alpha(t)\big|_1^\infty = 0 - \frac{1}{1-\alpha} = \frac{1}{\alpha - 1}\ .$$

Entsprechend ist f_α genau dann integrierbar auf $(0, 1]$, wenn $\alpha < 1$ ist, und dann gilt

$$\int\limits_{0^+}^1 t^{-\alpha}\, dt = F_\alpha(t)\big|_0^1 = \frac{1}{1-\alpha}\ .$$

Hieraus folgt auch, dass f_α für kein $\alpha \in \mathbb{R}$ auf $(0, \infty)$ integrierbar ist.
2. Wir betrachten $f(t) = (1 - t^2)^{-1/2}$ auf $I = (-1, 1)$. Es gilt $\arcsin' = f$ auf $(-1, 1)$ und arcsin ist stetig auf $[-1, 1]$. Nach Bemerkung 13.2.2 ist f integrierbar und es gilt

$$\int\limits_{-1^+}^{1^-} \frac{dt}{\sqrt{1-t^2}} = \arcsin(t)\big|_{-1}^1 = \pi\ .$$

Satz 13.4 *(Uneigentliche partielle Integration)*
Es seien $I \subset \mathbb{R}$ ein Intervall, $a := \inf I$, $b := \sup I$ und $f, g : I \to \mathbb{C}$ 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

$$\int\limits_{a^+}^{b^-} fG = FG\big|_a^b - \int\limits_{a^+}^{b^-} Fg.$$

Beweis. Da F, G stetig auf I sind, ist $fG + Fg$ stetig. Außerdem ist FG eine Stammfunktion zu $fG + Fg$. Nach Bemerkung 13.2.2 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 und es gilt

$$\int_{a^+}^{b^-} fG = FG\Big|_a^b - \int_{a^+}^{b^-} Fg.$$

Entsprechendes gilt, falls Fg integrierbar ist. $\square$

Beispiel 13.5 (Fläche des Einheitskreises) Wir betrachten die auf $[-1, 1]$ stetigen Funktionen $f(t) := 1$ und $G(t) := \sqrt{1 - t^2}$. Dann existiert das (eigentliche) Integral $\int_{-1}^{1} fG$. Mit $F(t) = t$ auf $[-1, 1]$ ergibt sich nach Satz 13.4 (da $t^2 = 1 - (1 - t^2)$)

$$\int_{-1}^{1} \sqrt{1 - t^2}\, dt = t\sqrt{1 - t^2}\Big|_{-1}^{1} + \int_{-1^+}^{1^-} \frac{t^2}{\sqrt{1 - t^2}}\, dt = \int_{-1^+}^{1^-} \frac{dt}{\sqrt{1 - t^2}} - \int_{-1}^{1} \sqrt{1 - t^2}\, dt$$

und damit

$$2\int_{-1}^{1} \sqrt{1 - t^2}\, dt = \int_{-1^+}^{1^-} \frac{dt}{\sqrt{1 - t^2}} = \pi.$$

Satz 13.6 *(Majorantenkriterium für Integrale)*
Es sei I ein Intervall und $a := \inf I$, $b := \sup I$. Ist $f \in R(I)$ und ist g integrierbar auf I mit $|f| \le g$, so ist auch f integrierbar auf I mit

$$\Big|\int_{a^+}^{b^-} f\Big| \le \int_{a^+}^{b^-} g\,.$$

Beweis. Es seien $F := V_u f$ und $G := V_u g$ für ein $u \in I$. Sind $s, t \in I$ mit $s < t$, so gilt

$$|F(t) - F(s)| = \Big|\int_s^t f\Big| \le \int_s^t |f| \le \int_s^t g = G(t) - G(s) \tag{13.1}$$

und damit $|F(t) - F(s)| \leq |G(t) - G(s)|$ für beliebige $s, t \in I$.
Wir zeigen: $F(b^-)$ existiert. Dazu sei (t_n) eine Folge in I mit $b > t_n \to b$. Dann ist $|F(t_n) - F(t_{n'})| \leq |G(t_n) - G(t_{n'})|$ für $n, n' \in \mathbb{N}$. Da $(G(t_n))$ nach Bemerkung 7.4 eine Cauchy-Folge ist, ist auch $(F(t_n))$ eine Cauchy-Folge und damit existiert $F(b^-)$, wieder nach Bemerkung 7.4. Genauso sieht man, dass $F(a^+)$ existiert. Schließlich folgt aus (13.1) auch $|F(b^-) - F(a^+)| \leq G(b^-) - G(a^+)$. □

Bemerkung und Definition 13.7 Insbesondere ergibt sich aus Satz 13.6 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 13.8 Für $\alpha > 1$ betrachten wir die Funktion $f : [1, \infty) \to \mathbb{R}$ mit

$$f(t) := t^{-\alpha} \cos(t) \qquad (t \geq 1).$$

Es gilt $|\cos t| t^{-\alpha} \leq t^{-\alpha}$ für $t \geq 1$. Da $\int_1^\infty t^{-\alpha}\, dt$ nach Beispiel 13.3.1 existiert, folgt die absolute Integrierbarkeit von f aus Satz 13.6. Entsprechendes gilt für die Funktion $t \mapsto t^{-\alpha} \sin(t)$ auf $[1, \infty)$.

Beispiel 13.9 Es sei $f : [1, \infty) \to \mathbb{C}$ stetig. Ist F eine Stammfunktion zu f und ist F beschränkt, so ist $t \mapsto t^{-1} f(t)$ integrierbar auf $[1, \infty)$.
(Denn: Da F beschränkt (und stetig) ist, existiert das Integral $\int_1^\infty F(t) t^{-2}\, dt$ nach dem Majorantenkriterium. Mit $G(t) = t^{-1}$ und $g(t) = -t^{-2}$ ergibt sich die Behauptung aus Satz 13.4 und $(FG)(t) = F(t) t^{-1} \to 0$ für $t \to \infty$.)
Wir betrachten $f(t) = \sin t$. Hier ist $F(t) = -\cos t$ beschränkt auf $[1, \infty)$. Also ist $t \mapsto t^{-1} \sin(t)$ integrierbar auf $[1, \infty)$. Man kann zeigen ([Ü]): $t \mapsto t^{-1} |\sin t|$ ist nicht integrierbar auf $[1, \infty)$. Also: Absolute Integrierbarkeit ist eine echt stärkere Eigenschaft als Integrierbarkeit.

Im folgenden Satz wird ein Zusammenhang zwischen der Konvergenz von Reihen und der Existenz uneigentlicher Integrale hergestellt:

Satz 13.10 *(Integralkriterium für Reihen)*
Es sei $f : [1, \infty) \to [0, \infty)$ *fallend. Dann existiert* $c := \lim\limits_{n\to\infty} \left(\sum\limits_{\nu=1}^{n} f(\nu) - \int_1^{n+1} f \right)$ *und es gilt* $0 \le c \le f(1)$.

Beweis. Wir setzen $a_n := f(n)$ für $n \in \mathbb{N}$. Aus $a_n \ge f(t) \ge a_{n+1}$ für $t \in [n, n+1]$ folgt $a_n \ge \int_n^{n+1} f \ge a_{n+1}$ und damit

$$0 \le a_n - \int\limits_n^{n+1} f \le a_n - a_{n+1}.$$

Also ist die Folge (s_n) mit

$$s_n := \sum_{\nu=1}^{n} a_\nu - \int\limits_1^{n+1} f = \sum_{\nu=1}^{n} \Big(a_\nu - \int\limits_\nu^{\nu+1} f \Big)$$

wachsend mit $0 \le s_n \le a_1 - a_{n+1} \le a_1$. Nach dem Hauptsatz über monotone Folgen ist s_n konvergent mit $0 \le \lim s_n \le a_1$. □

Beispiel 13.11 Es seien $\alpha > 0$ und $f(t) := t^{-\alpha}$ für $t \ge 1$. Dann ist f fallend auf $[1, \infty)$ und $f \ge 0$. Also existiert nach Satz 13.10

$$c = \lim_{n\to\infty} \Big(\sum_{\nu=1}^{n} \nu^{-\alpha} - \int\limits_1^{n+1} t^{-\alpha}\, dt \Big) .$$

Ist $\alpha > 1$, so ist $\lim_{n\to\infty} \int_1^{n+1} t^{-\alpha}\, dt = \int_1^\infty t^{-\alpha}\, dt = (\alpha - 1)^{-1}$ und $\zeta(\alpha) = \sum\limits_{\nu=1}^{\infty} \nu^{-\alpha}$. Nach Satz 13.10 ist

$$0 \le \zeta(\alpha) - \frac{1}{\alpha - 1} \le 1 \qquad (\alpha > 1) .$$

Ist $\alpha = 1$, so ergibt sich die Konvergenz von

$$s_n := \sum_{\nu=1}^{n} \frac{1}{\nu} - \int\limits_1^{n+1} \frac{dt}{t} = \sum_{\nu=1}^{n} \frac{1}{\nu} - \ln(n + 1) .$$

Der Grenzwert $c = \lim s_n$ heißt **Euler-Mascheroni Konstante**. Ist $0 < \alpha < 1$, so ergibt sich aus die Konvergenz von

$$s_n := \sum_{\nu=1}^{n} \nu^{-\alpha} - \int\limits_1^{n+1} t^{-\alpha}\, dt = \sum_{\nu=1}^{n} \nu^{-\alpha} - \frac{(n + 1)^{1-\alpha} - 1}{1 - \alpha} .$$

Satz 13.12 *Die Funktion* $t \mapsto e^{-t}t^{z-1}$ *ist für alle* $z \in \mathbb{C}$ *absolut integrierbar auf* $[1, \infty)$ *und für* $\mathrm{Re}(z) > 0$ *absolut integrierbar auf* $(0, \infty)$*.*

Beweis. Wir setzen $f(t) := e^{-t}t^{z-1}$ für $t > 0$. Aus $t^{z+1}e^{-t} \to 0$ für $t \to \infty$ folgt, dass $t \mapsto |e^{-t}t^{z+1}|$ ein Maximum auf $[1, \infty)$ hat. Also existiert eine Konstante $M > 0$ so, dass $|f(t)| \le Mt^{-2}$ für alle $t \in [1, \infty)$ gilt. Aus der Existenz von $\int_1^\infty t^{-2}\,dt$ ergibt sich mit dem Majorantenkriterium die absolute Integrierbarkeit von f auf $[1, \infty)$. Weiter gilt $|f(t)| \le t^{\mathrm{Re}(z)-1}$ für $t \in (0, 1]$. Ist $\mathrm{Re}(z) > 0$, so folgt aus der Existenz von $\int_0^1 t^{\mathrm{Re}(z)-1}dt$ (nach Beispiel 13.3.1) wieder mit dem Majorantenkriterium die absolute Integrierbarkeit von f auf $(0, 1]$ und damit auch auf $(0, \infty)$. $\square$

Bemerkung und Definition 13.13 Es sei $\Omega := \{z \in \mathbb{C} : \mathrm{Re}(z) > 0\}$ die offene rechte Halbebene. Die Funktion $\Gamma : \Omega \to \mathbb{C}$ 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

$$\Gamma(z+1) = z \cdot \Gamma(z) \qquad (\mathrm{Re}(z) > 0)\ . \tag{13.2}$$

Speziell gilt $\Gamma(1) = \int_0^\infty e^{-t}dt = -e^{-t}\big|_0^\infty = 1$, woraus sich wiederum mit (13.2) induktiv

$$\Gamma(n+1) = n!$$

für alle $n \in \mathbb{N}$ ergibt. Die Gammafunktion „interpoliert“ also die Fakultäten; man kann die Werte $\Gamma(z)$ als verallgemeinerte Fakultäten auffassen.

14 Mehrdimensionale Differenzialrechnung

Wir wollen jetzt Ableitungen für Funktionen mehrerer (meist reeller) Variablen definieren und untersuchen. Dazu betrachten wir zunächst allgemeiner Banachräume V, E über $\mathbb{K}$ und Abbildungen $f : X \to E$, wobei $X \subset V$. Im Weiteren seien also V, W und E stets Banachräume über $\mathbb{K}$ mit Normen $|\cdot|_V, |\cdot|_W$ und $|\cdot|_E$. Dabei lassen wir die Indizes meist weg, wenn der Bezug aus dem Zusammenhang klar ist.

Bemerkung und Definition 14.1 Eine lineare Abbildung $A : V \to E$ heißt **beschränkt**, falls $A|_{B_1(0)}$ beschränkt ist. Wir schreiben $L(V, E)$ für den Raum der beschränkten linearen Abbildungen von V nach E. Indem man A mit $A|_{B_1(0)}$ identifiziert, kann man $L(V, E)$ als Unterraum von $B(B_1(0), E)$ auffassen und dann ist durch

$$\|A\| := \sup_{x \in B_1(0)} |A(x)| = \|A|_{B_1(0)}\|_\infty$$

eine Norm auf $L(V, E)$ gegeben. Man nennt $\|\cdot\|$ die **Operatornorm** auf $L(V, E)$. Es gilt dabei

1. Für alle $x \in V$ ist $|A(x)| \leq \|A\| \cdot |x|$.
2. Ist $B \in L(E, W)$, so ist (mit $BA := B \circ A$)

$$\|BA\| \leq \|B\| \cdot \|A\|.$$

(Denn: 1. Ohne Einschränkung sei $x \neq 0$. Dann gilt

$$|A(x)| = \left|A\Big(|x|\,\frac{x}{|x|}\Big)\right| = |x|\left|A\Big(\frac{x}{|x|}\Big)\right| \leq \|A\| \cdot |x|.$$

2. Für $|x| \leq 1$ gilt mit 1.

$$|BA(x)| \leq \|B\| \cdot |A(x)| \leq \|B\| \cdot \|A\|.)$$

Sind $X \subset V$, $a \in X'$ und $f : X \to E$, so sagen wir f sei **abklingend** an a, falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit $|f(x)| < \varepsilon$ für alle $x \in X$ mit $0 < |x - a| < \delta$. Ist $c \in E$ so, dass $f - c$ abklingend an a ist, so ist wieder $c =: \lim_{x \to a} f(x)$ eindeutig bestimmt und heißt **Grenzwert** von f an a. Außerdem nennen wir f **lokal beschränkt** an a, falls $M, \delta > 0$ existieren mit $|f(x)| \leq M$ für alle $x \in X$ mit $0 < |x - a| < \delta$.

Bemerkung und Definition 14.2 Es seien $X \subset V$ und $f : X \to E$.
1. (vgl. Zerlegungsformel, Satz 10.5) Ist $a \in X^\circ$, also a innerer Punkt von X, so heißt f **(Fréchet-)differenzierbar** an der Stelle a, falls ein $A = A_{f,a} \in L(V, E)$ und eine an 0 abklingende Funktion $\varepsilon = \varepsilon_{f,a} : X_a := (X - a) \setminus \{0\} \to E$ existieren mit

$$(\tau_a f)(h) := f(a + h) - f(a) = A(h) + |h| \cdot \varepsilon(h) \quad (h \in X_a),$$

mit anderen Worten, falls $|\cdot|^{-1}(\tau_a f - A)$ abklingend an 0 ist. Man sieht leicht, dass die lineare Abbildung A im Falle der Differenzierbarkeit von f an a eindeutig bestimmt ist. Man schreibt $f'(a) := A$ und nennt $f'(a)$ die **(Fréchet-)Ableitung** von f an a. Dabei ist zu beachten, dass man im schon in Satz 10.5 betrachteten skalaren Fall $V = \mathbb{K}$ und $E = \mathbb{C}$ (als Vektorraum über $\mathbb{K}$) die Konstante $c \in \mathbb{C}$ mit der linearen Abbildung $A : \mathbb{K} \to \mathbb{C}$, definiert durch $A(h) := h \cdot c$, identifiziert.
2. Ist X offen und f differenzierbar an allen Stellen $a \in X$, so heißt f kurz **differenzierbar** (auf X). Die dann definierte Abbildung $f' : X \to L(V, E)$ mit heißt **Ableitung** von f (auf X). Weitere Schreibweisen sind wieder Df oder df oder auch df/dx. Ist $f' : X \to L(V, E)$ (wobei $L(V, E)$ mit der Operatornorm versehen ist) stetig, so sagen wir, f sei **stetig differenzierbar**. Im Falle $\mathbb{K} = \mathbb{R}$ schreiben wir $C^1(X, E)$ für die Menge aller stetig differenzierbaren $f : X \to E$.

Bemerkung 14.3 Es seien $X \subset V$, $a \in X^\circ$ und $f : X \to E$ differenzierbar an a. Dann gilt wie im skalaren Fall $X \subset \mathbb{K}$:
1. Ist f differenzierbar an a, so ist $|\cdot|^{-1}\tau_a f$ lokal beschränkt an 0 (da $|\cdot|^{-1}|A| \leq ||A||$) und damit $\tau_a f$ insbesondere abklingend an 0, also f stetig an a.
2. (Linearität der Ableitung): Ist $g : X \to E$ differenzierbar an a und ist $\lambda \in \mathbb{K}$, so ist auch $\lambda f + g$ differenzierbar an a mit $(\lambda f + g)'(a) = \lambda f'(a) + g'(a)$.
(Denn:

$$|\cdot|^{-1}(\tau_a(\lambda f + g) - \lambda f'(a) - g'(a)) = \lambda|\cdot|^{-1}(\tau_a f - f'(a)) + |\cdot|^{-1}(\tau_a g - g'(a))$$

ist abklingend an 0).
3. (Kettenregel) Ist $f(X) \subset Y \subset E$ und $g : Y \to W$ differenzierbar an $f(a)$, so ist $g \circ f$ differenzierbar an a mit

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

(Denn: Es sei zuächst speziell $a = 0$, $f(0) = 0$ und $= g(0) = 0$. Dann ist mit $\varepsilon := \varepsilon_{g,0}$ und $\varepsilon(0) := 0$

$$|f|(\varepsilon \circ f) + g'(0)(f - f'(0)) = g \circ f - g'(0)f'(0).$$

Da $\varepsilon \circ f$ sowie $|\cdot|^{-1}(f - f'(0))$ abklingend an 0 sind und $|\cdot|^{-1}f$ lokal beschränkt an 0 ist, ist $|\cdot|^{-1}(g \circ f - g'(0)f'(0))$ abklingend an 0. Also ist $(g \circ f)'(0) = g'(0)f'(0)$. Der allgemeine Fall ergibt sich daraus mit $\tau_{f(a)}g \circ \tau_a f = \tau_a(g \circ f)$.)

Ist V ein Banachraum, dessen Norm $|\cdot| = |\cdot|_V$ von einem Skalarprodukt induziert ist (gilt also $|\cdot|^2 = \langle\cdot,\cdot\rangle$), so nennt man $V = (V, \langle\cdot,\cdot\cdot\rangle)$ einen **Hilbertraum**. Wir verwenden im Weiteren die Cauchy-Schwarzsche Ungleichung (siehe Lineare Algebra), also

$$|\langle u, v\rangle| \leq |u| \cdot |v| \qquad (u, v \in V).$$

Beispiel 14.4 1. (affin-lineare Abbildungen) Es seien $A \in L(V, E)$ und $c \in E$. Ist $f : V \to E$ definiert durch

$$f(x) := A(x) + c \qquad (x \in V),$$

so ist $f'(x)(h) = A(h)$ für alle $x, h \in V$, also kurz $f'(x) = A$ für alle $x \in V$.
2. (quadratische Formen) Es seien V ein reeller Hilbertraum und $A \in L(V) := L(V, V)$. Ist $f : V \to \mathbb{R}$ definiert durch

$$f(x) := \langle x, Ax\rangle \qquad (x \in V),$$

so ist $f'(x)(h) = \langle h, A(x)\rangle + \langle x, A(h)\rangle$ für $x, h \in V$ ([Ü]), also kurz

$$f'(x) = \langle\cdot, A(x)\rangle + \langle x, A\rangle \qquad (x \in V).$$

Satz 14.5 *Es seien $X \subset V$ offen und $a, h \in V$ mit $[a, a + h] \subset X$.*

1. *(Mittelwertsatz) Ist $\mathbb{K} = \mathbb{R}$ und $f : X \to \mathbb{R}$ differenzierbar, so existiert ein $\xi \in (a, a + h)$ mit*
$$f(a + h) - f(a) = f'(\xi)(h).$$

2. *(Schrankensatz) Ist E ein Hilbertraum und ist $f : X \to E$ differenzierbar, so existiert ein $\xi \in (a, a + h)$ mit*
$$|f(a + h) - f(a)| \leq ||f'(\xi)|| \cdot |h|\ .$$

Beweis. Es sei $s := s_a^{a+h}$, also $s(t) := a + th$ für $t \in [0,1]$. Nach Beispiel 14.4.1 gilt $s'(t)(u) = uh$ für $t \in (0,1), u \in \mathbb{R}$ und damit nach der Kettenregel

$$(f \circ s)'(t)(u) = (f'(s(t))(uh) \qquad (t \in (0,1),\, u \in \mathbb{R}).$$

1. Ist $\mathbb{K} = \mathbb{R}$ und $f : V \to \mathbb{R}$, so existiert nach dem skalaren Mittelwertsatz (Satz 10.20.1) ein $\tau \in (0,1)$ mit

$$f(a+h) - f(a) = (f \circ s)(1) - (f \circ s)(0) = (f \circ s)'(\tau)(1) = (f'(s(\tau))(h).$$

Also folgt 1. mit $\xi := s(\tau)$.
2. Ist $\varphi \in L(E, \mathbb{K})$ und $g := \varphi \circ f \circ s : [0,1] \to \mathbb{K}$, so ist g stetig. Wieder nach der Kettenregel und Beispiel 14.4.1 gilt

$$g'(t)(u) = \varphi(f'(s(t)))(uh) \quad (t \in (0,1),\, u \in \mathbb{R}).$$

Nach dem skalaren Schrankensatz (Satz 10.20.2) existiert ein $\tau \in (0,1)$ mit

$$|\varphi(f(a+h) - f(a))| = |g(1) - g(0)| \leq |g'(\tau)(1)| \leq ||\varphi|| \cdot ||f'(s(\tau))|| \cdot |h|.$$

Ist $c := f(a+h) - f(a)$ und $\varphi := \langle \cdot, c\rangle$, so ist $||\varphi|| \leq |c|$ (nach der Cauchy-Schwarzschen Ungleichung) und $\varphi(c) = |c|^2$ (und damit auch $||\varphi|| = |c|$). Also folgt 2. wieder mit $\xi := s(\tau)$. $\square$

Bemerkung und Definition 14.6 Im Weiteren betrachten wir lineare Räume über $\mathbb{R}$, also $\mathbb{K} = \mathbb{R}$. Ein Vektor $\mathbf{v} \in V^*$ nennen wir eine **Richtung** in V. Für $a \in V$ ist dann $a + \mathbb{R}\mathbf{v}$ die Gerade durch den Punkt a in Richtung des Vektors $\mathbf{v}$.
Ist $X \subset V$ und $f : X \to E$ differenzierbar an a, so gilt für $t \in \mathbb{R}^*$ genügend klein

$$\Big|\frac{\tau_a f(t\mathbf{v})}{t} - f'(a)(\mathbf{v})\Big| = \frac{|\mathbf{v}|}{|t\mathbf{v}|}|(\tau_a f - f'(a))(t\mathbf{v})| \to 0 \quad (t \to 0)$$

und damit

$$\frac{f(a+t\mathbf{v}) - f(a)}{t} = \frac{\tau_a f(t\mathbf{v})}{t} \to f'(a)(\mathbf{v}) \quad (t \to 0).$$

Allgemein heißt f **richtungsdifferenzierbar** an der Stelle $a \in X$ in Richtung $\mathbf{v}$, falls

$$\partial_{\mathbf{v}} f(a) := \frac{\partial f}{\partial \mathbf{v}}(a) := D_{\mathbf{v}} f(a) := \lim_{t \to 0} \frac{f(a+t\mathbf{v}) - f(a)}{t}$$

existiert. In diesem Fall heißt $\partial_{\mathbf{v}} f(a)$ die **Richtungsableitung** von f an der Stelle a in Richtung $\mathbf{v}$. Die obige Überlegung zeigt, dass aus der Differenzierbarkeit von f an a insbesondere die Differenzierbarkeit in alle Richtungen $\mathbf{v}$ folgt, und dass dann

$$\partial_{\mathbf{v}} f(a) = f'(a)(\mathbf{v}) \tag{14.1}$$

gilt.
Ist speziell $V = \mathbb{R}^d$ und $\mathbf{v} = \mathbf{e}^{(k)}$ der k-te Einheitsvektor, so sagt man auch, f sei **partiell differenzierbar** an a nach der k-ten Variablen. Dann schreiben wir auch $\partial_k f(a)$ statt $\partial_{\mathbf{e}^{(k)}} f(a)$ und sprechen von der **partiellen Ableitung** von f an a nach der k-ten Variablen. Im Falle $d = 2$ schreibt man für die Variablen traditionell oft (x, y) statt (x_1, x_2). In diesem Falle spricht man von den partiellen Ableitungen nach x bzw. y und schreibt auch $\partial f/\partial x$ oder f_x sowie $\partial f/\partial y$ oder f_y. Entsprechend schreibt man im Falle $d = 3$ oft (x, y, z) statt (x_1, x_2, x_3) und $\partial f/\partial x$, $\partial f/\partial y$ sowie $\partial f/\partial z$ beziehungsweise f_x, f_y, f_z.
Die Definition zeigt, dass im Falle und $E = \mathbb{K}$ (als $\mathbb{R}$-Vektorraum) Richtungs- und partielle Ableitungen nichts anderes als Ableitungen von (reell oder komplexwertigen) Funktionen einer reellen Variable sind. Folglich stehen damit die Rechenregeln und Ergebnisse der eindimensionalen Differenzialrechnung zur Verfügung. Besonders einfach ist die Situation für partielle Ableitungen: $\partial_k f$ ist 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 als „Parameter“ aufgefasst).

Beispiel 14.7 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 allgemeines $(x, y) \in \mathbb{R}^2$

$$\partial_1 f(x, y) \left(= \frac{\partial f}{\partial x}(x, y) = f_x(x, y) \right) = 2x$$

und

$$\partial_2 f(x, y) \left(= \frac{\partial f}{\partial y}(x, y) = f_y(x, y) \right) = 1\,.$$

Weiter erhalten wir für $\mathbf{v} = (v_1, v_2)$ und $a = (0, 0)$

$$f\big(t(v_1, v_2)\big) = t^2 v_1^2 + t v_2 \qquad (t \in \mathbb{R})\,.$$

und damit

$$\partial_{\mathbf{v}} f(0, 0) = \left(\frac{\partial f}{\partial \mathbf{v}}(0, 0) = D_{\mathbf{v}} f(0, 0) \right) = \lim_{t \to 0} \frac{t^2 v_1 + t v_2}{t} = v_2\,.$$

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 $(x,y,z) \in \mathbb{R}^3$

$$\begin{aligned}
\partial_1 f(x,y,z) \quad &\Big(= \tfrac{\partial f}{\partial x}(x,y,z) \;=\; f_x(x,y,z)\Big) \;=\; y^2z^3 \\
\partial_2 f(x,y,z) \quad &\Big(= \tfrac{\partial f}{\partial y}(x,y,z) \;=\; f_y(x,y,z)\Big) \;=\; 2xyz^3 \\
\partial_3 f(x,y,z) \quad &\Big(= \tfrac{\partial f}{\partial z}(x,y,z) \;=\; f_2(x,y,z)\Big) \;=\; 3xy^2z^2 \, .
\end{aligned}$$

3. Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x,y) := \begin{cases} \dfrac{xy}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases} .$$

Dann gilt für $(x,y) \neq (0,0)$

$$\partial_1 f(x,y) = \frac{y}{x^2+y^2} - \frac{2x^2y}{(x^2+y^2)^2}$$

und

$$\partial_2 f(x,y) = \frac{x}{x^2+y^2} - \frac{2xy^2}{(x^2+y_0^2)^2} \, .$$

Aus $f(t,0) = f(0,t) = 0$ für alle $t \in \mathbb{R}$ folgt

$$\partial_1 f(0,0) = \partial_2 f(0,0) = 0 \, .$$

Also existieren die partiellen Ableitungen in allen Punkten (x,y). Die Funktion f ist allerdings nicht stetig an der Stelle $(0,0)$, denn für $(x_n,y_n) = (1/n, 1/n)$ gilt

$$f(x_n,y_n) = \frac{1/n^2}{1/n^2 + 1/n^2} = \frac{1}{2} \to \frac{1}{2} \neq 0 = f(0,0) \qquad (n \to \infty) \, .$$

Das letzte Beispiel zeigt, dass die Existenz der partiellen Ableitungen auf einer Umgebung von a im Allgemeinen noch nicht die Stetigkeit an a impliziert. Man beachte allerdings, dass dort die partiellen Ableitungen in der Nähe von $a = (0,0)$ unbeschränkt und damit insbesondere unstetig sind, denn für $x \neq 0$, $y \neq 0$ ist

$$\partial_1 f(0,y) = 1/y \, , \qquad \partial_2 f(x,0) = 1/x \, .$$

Sind die partiellen Ableitungen stetig, so verbessert sich die Situation deutlich:

Satz 14.8 *Es seien $X \subset \mathbb{R}^d$, $a \in X^\circ$ und $f : X \to \mathbb{R}$ so, dass die partiellen Ableitungen $\partial_1 f, \ldots, \partial_d f$ auf einer Umgebung von a existieren. Sind $\partial_1 f, \ldots, \partial_d f$ stetig an a, so ist f differenzierbar an a.*

Beweis. Ohne Einschränkung seien $a = 0$ und $f(0) = 0$ (der allgemeine Fall ergibt sich wieder mit $\tau_a f$ statt f). Wir setzen $A(h) := \sum_{k=1}^{d} \partial_k f(0) \cdot h_k$ für $h \in \mathbb{R}^d$.
Ist $\varepsilon > 0$ gegeben, so existiert ein $\delta > 0$ mit $\bigl|\partial_k f(x) - \partial_k f(0)\bigr| < \varepsilon/d$ für alle $x \in U_\delta(0)$ und $k = 1, \ldots, d$. Ist $h \in U_\delta(0), h \neq 0$, so setzen wir

$$v^{(k)} := \sum_{\ell=1}^{k} h_\ell \mathbf{e}^{(\ell)} \qquad (k = 0, \ldots, d).$$

Dann ist $v^{(0)} = 0$ und $v^{(d)} = \sum_{k=1}^{d} h_k \mathbf{e}^{(k)} = (h_1, \ldots, h_d) = h$, also (Teleskopsumme)

$$f(h) = \sum_{k=1}^{d} \bigl(f(v^{(k)}) - f(v^{(k-1)})\bigr).$$

Mit $|v^{(k)}| \le |h| < \delta$ und $v^{(k)} = v^{(k-1)} + h_k \mathbf{e}^{(k)}$ für $k = 0, \ldots, d$ folgt aus dem skalaren Mittelwertsatz (angewandt auf $g_k(t) := f(v^{(k-1)} + t\mathbf{e}^{(k)})$ für $t \in [0, h_k]$) die Existenz eines $\xi^{(k)} \in U_\delta(0)$ mit

$$f(v^{(k)}) - f(v^{(k-1)}) = \partial_k f(\xi^{(k)}) \cdot h_k\,.$$

Folglich erhalten wir

$$\begin{aligned}
\frac{1}{|h|}\bigl|f(h) - A(h)\bigr| &\le \frac{1}{|h|} \sum_{k=1}^{d} \bigl|f(v^{(k)}) - f(v^{(k-1)}) - \partial_k f(0) \cdot h_k\bigr| \\
&= \frac{1}{|h|} \sum_{k=1}^{d} \bigl|\bigl(\partial_k f(\xi^{(k)}) - \partial_k f(0)\bigr) h_k\bigr| \le \frac{1}{|h|} \frac{\varepsilon}{d} \sum_{k=1}^{d} |h_k| \le \varepsilon\,.
\end{aligned}$$

□

Beispiel 14.9 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 $\mathbb{R}^2$ (warum?). Also ist f nach Satz 14.8 differenzierbar auf $\mathbb{R}^2$.

Wir wollen zum Schluss des Abschnitts auf die geometrische Bedeutung der Ableitung im Mehrdimensionalen eingehen.

Bemerkung 14.10 1. Es seien $X \subset \mathbb{R}^d$, $a \in X$ und $f : X \to \mathbb{K}^m$. Ist f an a partiell differenzierbar nach allen Variablen, so heißt die Matrix $Jf(a) \in \mathbb{K}^{m\times d}$, deren k-te Spalte aus dem Vektor $\partial_k f(a)$ besteht, die **Jacobi-Matrix** von f an a. Ist f sogar differenzierbar an a, so gilt

$$\partial_k f(a) = f'(a)(\mathbf{e}^{(k)}) \qquad (k = 1, \ldots, d)$$

nach (14.1) und damit ist $Jf(a)$ die Matrix, die die lineare Abbildung $f'(a)$ in der Standardbasis darstellt, also

$$f'(a)(\mathbf{v}) = Jf(a)\mathbf{v}$$

für alle $\mathbf{v} = (v_1, \ldots, v_d)$ erfüllt. Ausserdem gilt dann nach (14.1) für alle Richtungen $\mathbf{v}$

$$\partial_{\mathbf{v}} f(a) = Jf(a)\mathbf{v}.$$

2. Ist speziell $m = 1$, so ist die Jacobi-Matrix der (Zeilen-)vektor $(\partial_1 f(a), \ldots, \partial_d f(a))$. Dann nennt man

$$\nabla f(a) := \mathrm{grad} f(a) := (\partial_1 f(a), \ldots, \partial_d f(a))^T$$

den **Gradient** von f an a. (Wir werden im Weiteren Spaltenvektoren meistens wieder als Zeilenvektoren schreiben, was immer dann egal ist, wenn wir keine Matrizenarithmetik betrieben.) Ist f differenzierbar an a, so gilt für jede Richtung $\mathbf{v}$

$$\partial_{\mathbf{v}} f(a) = (\nabla f)^T(a)\mathbf{v}$$

und damit für $\mathbb{K} = \mathbb{R}$ nach der Cauchy-Schwarzschen Ungleichung mit $\mathbf{v}^* := \nabla f(a)$

$$-|\mathbf{v}^*| \cdot |\mathbf{v}| \leq \partial_{\mathbf{v}} f(x) = (\mathbf{v}^*)^T \cdot \mathbf{v} \leq |\mathbf{v}^*| \cdot |\mathbf{v}| \ .$$

sowie im Falle $\mathbf{v}^* \neq 0$ für $\mathbf{v} = \mathbf{v}^*$ (die sogenannte Gradientenrichtung)

$$\partial_{\mathbf{v}^*} f(x) = (\mathbf{v}^*)^T \mathbf{v}^* = |\mathbf{v}^*|^2$$

und entsprechend

$$\partial_{-\mathbf{v}^*} f(x) = -|\mathbf{v}^*|^2 \ .$$

Angesichts des Mittelwertsatzes (also $f(x+\mathbf{v}) - f(x) = f'(\xi)(\mathbf{v}) = \partial_{\mathbf{v}} f(\xi)$ für ein $\xi \in (x, x+\mathbf{v})$) kann man daher – jedenfalls im Falle stetiger Richtungsableitungen – die Gradientenrichtung als „Richtung des steilsten Anstiegs“ von f und die negative Gradientenrichtung als „Richtung des steilsten Abstiegs“ von f ansehen. Außerdem gilt für $\mathbf{v} \perp \mathbf{v}^*$

$$\partial_{\mathbf{v}} f(x) = (\mathbf{v}^*)^T \cdot \mathbf{v} = 0 \, ,$$

d. h. die Richtungsableitungen der zur Gradientenrichtung senkrechten Richtungen verschwinden.

Beispiel 14.11 Es sei $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = x^2 + y^2 \qquad (x, y \in \mathbb{R}) \ .$$

Dann gilt

$$\nabla f(x,y) = (2x, 2y) \ .$$

Betrachtet man etwa $(x,y) = (1,1)$, so ist $\mathbf{v}^* = (2,2)$ Richtung des steilsten Anstiegs. Für die zu $\mathbf{v}^*$ senkrechte Richtung $\mathbf{v} = (1,-1)$ gilt $\partial_{\mathbf{v}} f(1,1) = 0$.

15 Extremstellen von Funktionen mehrerer Variablen

Wir beschäftigen uns zunächst mit Ableitungen höherer Ordnung.

Bemerkung und Definition 15.1 Es seien V, E reelle Banachräume, $X \subset V$ offen und $f : X \to E$. Sind $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \ldots$ Richtungen in V, so definiert man induktiv für $n \geq 2$ (soweit existent!)

$$\partial_{\mathbf{v}^{(n)}} \ldots \partial_{\mathbf{v}^{(1)}} f := \partial_{\mathbf{v}^{(n)}} \big(\partial_{\mathbf{v}^{(n-1)}} \ldots \partial_{\mathbf{v}^{(1)}} f \big).$$

Für $\mathbf{v}^{(1)} = \ldots = \mathbf{v}^{(n)} =: \mathbf{v}$ schreibt man kurz $\partial_{\mathbf{v}}^n f$ und spricht dann von der **Richtungsableitung der Ordnung** n von f in Richtung $\mathbf{v}$. Sind speziell $V = \mathbb{R}^d$ und $\mathbf{v}^{(1)} = \mathbf{e}^{(k_1)}, \ldots, \mathbf{v}^{(n)} = \mathbf{e}^{(k_n)}$, so schreibt man

$$\partial_{k_n} \ldots \partial_{k_1} f \quad \text{oder} \quad \frac{\partial^n f}{\partial x_{k_n} \ldots \partial x_{k_1}} \quad \text{oder} \quad f_{x_{k_1} \ldots x_{k_n}}$$

an Stelle von $\partial_{\mathbf{v}^{(n)}} \ldots \partial_{\mathbf{v}^{(1)}} f$ und spricht von den **partiellen Ableitungen der Ordnung** n. Für $k_1 = \ldots = k_n =: k$ schreibt man kurz $\partial_k^n f$ bzw. $\partial^n f / \partial x_k^n$. Schließlich setzt man 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 15.2 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

$$\partial_1 \partial_2 \partial_1 f(x,y) = 2 = \partial_2 \partial_1^2 f(x,y).$$

In Beispiel 15.2 ist es so, dass die Reihenfolge, in der die partiellen Ableitungen gebildet werden, vertauscht werden kann. Allgemein gilt

Satz 15.3 *(Schwarz)*
Es seien $X \subset \mathbb{R}^d$ offen und $p, q \in \{1, \ldots, d\}$. Ferner sei $f : X \to \mathbb{R}$ so, dass $\partial_p f$, $\partial_q f$ und $\partial_q \partial_p f$ auf X existieren. Ist $\partial_q \partial_p f$ stetig an der Stelle $a \in X$, so existiert auch $\partial_p \partial_q f(a)$ und es gilt

$$\partial_p \partial_q f(a) = \partial_q \partial_p f(a) .$$

Beweis. Ohne Einschränkung können wir $d = 2$ sowie $p = 1, q = 2$ annehmen. Außerdem sei dann ohne Einschränkung $a = (0, 0)$.
Da X offen ist, existiert ein $R > 0$ so, dass $(x, y) \in X$ 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 skalaren Mittelwertsatzes zeigen, dass ein $\xi = \xi(x, y) \in [0, x]$ und ein $\eta = \eta(x, y) \in [0, y]$ 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 sei $\varepsilon > 0$ gegeben. 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 $|v| < \delta$

$$|\partial_2 \partial_1 f(u, v) - \partial_2 \partial_1 f(0, 0)| < \varepsilon$$

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

$$\lim_{y \to 0} \frac{\Delta(x, y)}{xy} = \frac{\partial_2 f(x, 0) - \partial_2 f(0, 0)}{x}$$

gilt, 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 15.4 Die Aussage von Satz 15.3 wird im Allgemeinen falsch, wenn man auf die Voraussetzung der Stetigkeit der partiellen Ableitung $\partial_q \partial_p f$ an a verzichtet. So kann man etwa für die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x,y) := \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2} & , \quad (x,y) \neq (0,0) \\ 0 & , \quad (x,y) = (0,0) \end{cases}$$

zeigen ([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 $\partial_2\partial_1 f(0,0) = -1$ und $\partial_1\partial_2 f(0,0) = 1$.

Bemerkung und Definition 15.5 1. In den Beweisen zu Satz 14.8 und Satz 15.3 haben wir den Mittelwertsatz verwandt und deshalb reellwertige Funktionen betrachtet. Mithilfe von Bemerkung 8.9 kann man sich überlegen, dass die Aussagen der beiden Sätze allgemeiner für $\mathbb{K}^m$-wertige Funktionen gelten. Wir werden die entsprechenden Aussagen im Weiteren auch so verwenden.
2. Es sei $X \subset \mathbb{R}^d$ offen. Durch Anwendung der Ungleichungen

$$\max_{j,k} |a_{jk}| \leq \|A\| \leq d\sqrt{m} \max_{j,k} |a_{jk}|$$

für Matrizen $A = (a_{jk}) \in \mathbb{K}^{m\times d}$ und $x \mapsto Ax$ (siehe [Ü]) auf die Jacobi-Matrix sieht man, dass $f : X \to \mathbb{K}^m$ genau dann stetig differenzierbar auf X ist, wenn alle partiellen Ableitungen $\partial_1 f, \ldots, \partial_d f$ stetig auf X sind. Ist $n \in \mathbb{N}_0$, so bezeichnen wir mit $C^n(X, \mathbb{K}^m)$ die Menge aller Funktionen $f : X \to \mathbb{K}^m$ mit der Eigenschaft, dass alle partiellen Ableitungen der Ordnung $\leq n$ auf X existieren und dort stetig sind. Außerdem setzen wir $C^n(X) := C^n(X, \mathbb{C})$. Durch mehrfache Anwendung von Satz 15.3 sieht man: Ist $f \in C^n(X, \mathbb{K}^m)$ und sind $k_1, \ldots, k_n \in \{1, \ldots, d\}$, so gilt für jede Permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$:

$$\partial_{k_n} \ldots \partial_{k_1} f = \partial_{k_{\sigma(n)}} \ldots \partial_{k_{\sigma(1)}} f \ ,$$

d. h. die Reihenfolge, in der die partiellen Ableitungen $\partial_{k_1}, \ldots, \partial_{k_n}$ gebildet werden, spielt keine Rolle.

Bemerkung 15.6 Es seien $X \subset \mathbb{R}^d$ offen, $n \in \mathbb{N}_0$ und $\mathbf{v}$ eine Richtung. Weiter sei $f : X \to \mathbb{C}$ so, dass $\partial_{\mathbf{v}}^{n+1} f$ existiert und stetig ist. Ist $a \in X$ und $I \supset [0,1]$ ein offenes Intervall mit $a + I\mathbf{v} \subset X$, so betrachten wir die Funktion $g : I \to \mathbb{C}$, definiert durch

$$g(t) := f(a + t\mathbf{v}) \qquad (t \in I).$$

Aus der Definition der Richtungsableitungen ergibt sich $g^{(k)}(t) = \partial_{\mathbf{v}}^k f(a+t\mathbf{v})$ für $t \in I$ und $k \leq n+1$. Durch Anwendung des Satzes von Taylor (Satz 12.19) auf die Funktion g (mit $a = 0$ und $h = 1$) erhält man

$$f(a+\mathbf{v}) = \sum_{\nu=0}^{n} \frac{\partial_{\mathbf{v}}^{\nu} f(a)}{\nu!} + \frac{1}{n!} \int_0^1 (1-t)^n \partial_{\mathbf{v}}^{n+1} f(a+t\mathbf{v})\, dt$$

(Taylorformel für Richtugen). Ist f reellwertig, so existiert zudem nach Bemerkung 12.21 ein $\xi \in [a, a+\mathbf{v}]$ so, dass

$$\frac{1}{n!} \int_0^1 (1-t)^n \partial_{\mathbf{v}}^{n+1} f(a+t\mathbf{v})\, dt = \frac{1}{(n+1)!} \partial_{\mathbf{v}}^{n+1} f(\xi)$$

(Lagrangeform des Restgliedes).

Wir wollen nun zeigen, dass man die Richtungsableitungen $\partial_{\mathbf{v}}^n f$ durch die partiellen Ableitungen der Ordnung n von f ausdrücken kann. Dazu setzen wir für $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$

$$\alpha! = \prod_{k=1}^{d} \alpha_k!, \qquad |\alpha|_1 := \sum_{k=1}^{d} \alpha_k, \qquad \partial^{\alpha} := \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d}$$

und für $h = (h_1, \ldots, h_d) \in \mathbb{R}^d$

$$h^{\alpha} := \prod_{k=1}^{d} h_k^{\alpha_k}.$$

Satz 15.7 *Es seien $X \subset \mathbb{R}^d$ offen und $f \in C^n(X, \mathbb{K}^m)$. Dann ist $\partial_{\mathbf{v}}^n f$ stetig für alle Richtungen $\mathbf{v} = (v_1, \ldots, v_d)$ mit*

$$\partial_{\mathbf{v}}^n f = \sum_{(k_1,\ldots,k_n)\in\{1,\ldots,d\}^n} v_{k_1} \cdots v_{k_n} (\partial_{k_n} \ldots \partial_{k_1} f) = n! \sum_{|\alpha|_1 = n} \frac{1}{\alpha!} \mathbf{v}^{\alpha} \partial^{\alpha} f.$$

Beweis. 1. Wir beweisen den ersten Teil per Induktion nach n.
$n = 1$: Ist $f \in C^1(X, \mathbb{K}^m)$, so folgt aus Bemerkung 14.10

$$\partial_{\mathbf{v}}^1 f(x) = \partial_{\mathbf{v}} f(x) = Jf(x) \cdot \mathbf{v} = \sum_{k_1=1}^{d} v_{k_1} \partial_{k_1} f(x) \qquad (x \in X)\,.$$

$n \to n+1$: Ist $f \in C^{n+1}(X, \mathbb{K}^m)$, so folgt $\partial_{\mathbf{v}}^n f \in C^1(X, \mathbb{K}^m)$ 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

$$\begin{aligned} \partial_{\mathbf{v}}^{n+1} f(x) &= \partial_{\mathbf{v}}(\partial_{\mathbf{v}}^n f)(x) = (J\partial_{\mathbf{v}}^n f)(x) \cdot \mathbf{v} \\ &= \sum_{k_{n+1}=1}^{d} \sum_{k_1,\ldots,k_n=1}^{d} v_{k_1} \cdots v_{k_n} \cdot v_{k_{n+1}} \partial_{k_{n+1}}(\partial_{k_n} \ldots \partial_{k_1} f)(x) \quad (x \in X). \end{aligned}$$

Damit ist $\partial_{\mathbf{v}}^{n+1} f$ stetig und die erste Gleichung gilt für $n+1$.
2. Nach Bemerkung 15.5 kann man die Reihenfolgen der partiellen Ableitungen in der ersten Summe beliebig permutieren. Da $\frac{n!}{\alpha_1! \ldots \alpha_d!}$ Tupel $(k_1, \ldots, k_n) \in \{1, \ldots, d\}^n$ existieren, bei denen die Zahl $j \in \{1, \ldots, d\}$ genau α_j-mal vorkommt (siehe Bemerkung 3.10 für den Fall $d = 2$), ergibt sich auch die zweite Gleichung, also

$$\partial_{\mathbf{v}}^n f(x) = \sum_{|\alpha|_1 = n} \frac{n!}{\alpha_1! \cdots \alpha_d!} v_1^{\alpha_1} \cdots v_d^{\alpha_d} (\partial_1^{\alpha_1} \ldots \partial_d^{\alpha_d} f)(x) \qquad (x \in X)\ .$$

□

Bemerkung und Definition 15.8 Wir werden nun sehr bescheiden und betrachten speziell die Fälle $n = 0$ und $n = 1$ der Taylorformel:
Ist $f \in C^1(X, \mathbb{R})$, so ergibt sich für $a, h \in \mathbb{R}^d$ mit $[a, a+h] \subset X$ die Existenz eines $\xi \in [a, a+h]$ so, dass

$$f(a+h) = f(a) + \partial_h f(\xi) = f(a) + (\nabla f)^T(\xi) h,$$

also wieder die Aussage des Mittelwertsatzes. Für $f \in C^2(X, \mathbb{R})$ heißt die Matrix

$$(Hf)(x) := \big(\partial_k \partial_j f(x)\big)_{j,k=1,\ldots,d} \in \mathbb{R}^{d \times d}$$

die **Hesse-Matrix** von f an der Stelle x. Nach Satz 15.7 gilt

$$\partial_h^2 f(x) = \sum_{j,k=1}^{d} h_j h_k (\partial_k \partial_j f)(x) = h^T (Hf)(x) h$$

und nach der Taylorformel existiert ein $\xi \in [a, a+h]$ mit

$$f(a+h) = f(a) + \partial_h f(a) + \frac{1}{2} \partial_h^2 f(\xi) = f(a) + (\nabla f)^T(a) \cdot h + \frac{1}{2} h^T Hf(\xi) h\ .$$

Satz 15.9 *(Taylor)*
Es seien $X \subset \mathbb{R}^d$ offen und $f \in C^{n+1}(X, \mathbb{R})$ für ein $n \in \mathbb{N}_0$. Sind $a, h \in \mathbb{R}^d$ mit $[a, a+h] \subset X$, so existiert ein $\xi \in [a, a+h]$ so, dass

$$f(a+h) = \sum_{|\alpha|_1 \le n} \frac{\partial^\alpha f(a)}{\alpha!} h^\alpha + \sum_{|\alpha|_1 = n+1} \frac{\partial^\alpha f(\xi)}{\alpha!} h^\alpha.$$

Beweis. Ohne Einschränkung sei $h \neq 0$. Dann ergibt sich aus der Taylorformel für Richtungen und Satz 15.7 die Existenz eines $\xi \in [a, a+h]$ mit

$$f(a+h) - \sum_{|\alpha|_1 = n+1} \frac{\partial^\alpha f(\xi)}{\alpha!} h^\alpha = \sum_{\nu=0}^{n} \sum_{|\alpha|_1 = \nu} \frac{1}{\alpha!} \partial^\alpha f(a)\, h^\alpha = \sum_{|\alpha|_1 \le n} \frac{\partial^\alpha f(a)}{\alpha!} h^\alpha\ .$$

□

Bemerkung und Definition 15.10 Der erste Summand

$$T_{n,a}(h) := (T_{n,a} f)(h) := \sum_{|\alpha|_1 \le n} \frac{\partial^\alpha f(a)}{\alpha!} h^\alpha$$

ist bei festem a ein Polynom in den d Variablen $h_1, \ldots, h_d$ (vom Grad $\le n$). Dieses Polynom heißt wieder n-tes **Taylor-Polynom** von f bezüglich a.

Definition 15.11 Sind $X \subset V$ und $f : X \to E$ differenzierbar an $a \in X$, so heißt a **reguläre Stelle**, falls $f'(a)$ surjektiv ist. Ist $f'(a)$ nicht surjekiv, so heißt a **kritisch** (oder **singulär**). Im Falle $V = \mathbb{R}^d$ und $E = \mathbb{R}$ ist a genau dann kritisch, wenn $\nabla f(a) = 0$ gilt.

Als wesentliche Anwendung des Taylor-Satzes werden wir nun ein hinreichendes Kriterium für lokale Extrema herleiten. Zunächst gilt folgendes einfache *notwendige* Kriterium.

Satz 15.12 *Es seien $X \subset \mathbb{R}^d$ und a ein innerer Punkt von X. Ist $f : X \to \mathbb{R}$ differenzierbar an a und ist a eine Extremstelle von f, so ist a ein kritische Stelle.*

Beweis. Es sei $k \in \{1, \ldots, d\}$. Ist I ein offenes Intervall mit $0 \in I$ und $a + I\mathbf{e}^{(k)} \subset X$, und ist $g_k : I \to \mathbb{R}$ definiert durch

$$g_k(t) := f(a + t\mathbf{e}^{(k)}) \qquad (t \in I),$$

so ist 0 Extremstelle von g_k. Nach Satz 10.15 gilt $0 = g_k'(0) = \partial_k f(a)$. □

Beispiel 15.13 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 $\nabla f(x, y) = (2x, 2y)$, also tatsächlich $\nabla f(0, 0) = 0$.
2. Es sei $f : \mathbb{R} \to \mathbb{R}$ 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 $\nabla f(x, y) = (2x, 2y)$, also tatsächlich $\nabla f(0, 0) = 0$.
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 $\nabla f(x, y) = (2x, -2y)$, also $\nabla f(0, 0) = (0, 0)$. Allerdings ist $(0, 0)$ keine Extremstelle von f.

Das zweite Beispiel zeigt, dass auch im Höherdimensionalen an kritischen Stellen im Allgemeinen keine Extremstellen vorliegen. Um auf Extremstellen schließen zu können, bedarf es sogenannter Kriterien zweiter Ordnung, also Kriterien, die die Hesse-Matrix einbeziehen. Ist $f \in C^2(X, \mathbb{R})$ für eine offene Menge $X \subset \mathbb{R}^d$, so ist die Hesse-Matrix $Hf(x)$ nach dem Satz von Schwarz (Satz 15.3) für alle $x \in X$ symmetrisch.

Definition 15.14 Ist $A \in \mathbb{R}^{d \times d}$ symmetrisch, so heißt A

1. **positiv (semi-)definit**, falls $\mathbf{v}^T A\mathbf{v} > 0$ (≥ 0) für alle Richtungen $\mathbf{v}$ gilt,
2. **negativ (semi-)definit**, falls $-A$ positiv (semi-) definit ist,
3. **indefinit**, falls A weder positiv noch negativ semidefinit ist.

Damit gilt folgendes, für die mehrdimensionale Optimierung zentrale Resultat:

Satz 15.15 *Es seien $X \subset \mathbb{R}^d$ offen und $f \in C^2(X, \mathbb{R})$. Ist $a \in X$ ein kritischer Punkt von f, so gilt:*

1. *Ist $Hf(a)$ positiv (bzw. negativ) definit, so hat f an a ein striktes lokales Minimum (bzw. Maximum).*

2. *Hat f an a ein lokales Minimum (bzw. Maximum), so ist $Hf(a)$ positiv (bzw. negativ) semidefinit.*

Beweis. Wir setzen

$$S^{d-1} := \{\mathbf{v} \in \mathbb{R}^d : |\mathbf{v}| = 1\}$$

(die $(d-1)$-dimensionale **Einheitssphäre**). Dann ist S^{d-1} beschränkt und abgeschlossen, also kompakt nach dem Satz von Heine-Borel. Ist $B \in \mathbb{R}^{d\times d}$, so die Funktion $h \mapsto h^T Bh$ stetig auf $\mathbb{R}^d$ (da sogar differenzierbar nach Beispiel 14.4). Damit existieren $\min\limits_{\mathbf{v}\in S^{d-1}} \mathbf{v}^T B\mathbf{v}$ und $\max\limits_{\mathbf{v}\in S^{d-1}} \mathbf{v}^T B\mathbf{v}$. Außerdem folgt aus der Cauchy-Schwarzschen Ungleichung

$$|\mathbf{v}^T B\mathbf{v}| \leq ||B|| \qquad (\mathbf{v} \in S^{d-1})\,.$$

Es genügt jeweils, die ersten Behauptungen zu beweisen. Die zweiten ergeben sich dann durch Betrachtung von $-f$.

1. Es sei $A := Hf(a)$ positiv definit. Wir setzen

$$c := \min_{\mathbf{v}\in S^{d-1}} \mathbf{v}^T A\mathbf{v} (> 0)\,.$$

Aus der Stetigkeit von $\partial_k\partial_j f$ für alle j, k folgt die Stetigleit von $Hf : X \to \mathbb{R}^{d\times d}$ (versehen mit der Operatornorm) (vgl. Bemerkung 15.5). Daher existiert ein $\delta > 0$ mit $||Hf(x) - A|| < c$ für alle $x \in U_\delta(a)$ und damit auch

$$|\mathbf{v}^T (Hf(\xi) - A)\mathbf{v}| < c \qquad (\mathbf{v} \in S^{d-1}).$$

Es sei $h \in U_\delta(0)$, $h \neq 0$ und $\mathbf{v} := h/|h|$. Da a ein kritischer Punkt ist, existiert nach Bemerkung 15.8 existiert ein $\xi \in [a, a+h] \subset U_\delta(a)$ mit

$$\begin{aligned} \frac{2}{|h|^2}(f(a+h) - f(a)) &= \mathbf{v}^T Hf(\xi)\mathbf{v} = \mathbf{v}^T (Hf(\xi) - A)\mathbf{v} + \mathbf{v}^T A\mathbf{v} \\ &\geq \mathbf{v}^T A\mathbf{v} - \left|\mathbf{v}^T (Hf(\xi) - A)\mathbf{v}\right| > c - c = 0\,. \end{aligned}$$

Damit ist auch $f(a+h) > f(a)$.

2. Es sei $\mathbf{v}$ eine Richtung. Ist I ein offenes Intervall mit $0 \in I$ und $a + I\mathbf{v} \subset X$, und ist $g : I \to \mathbb{R}$ definiert durch $g(t) := f(a + t\mathbf{v})$ $(t \in I)$, so ist $g \in C^2(I, \mathbb{R})$ und hat an 0 ein lokales Minimum und damit kein striktes lokales Maximum. Aus Satz 11.2 und Bemerkung 15.8 folgt

$$0 \leq g''(0) = \partial_{\mathbf{v}}^2 f(a) = \mathbf{v}^T A \mathbf{v}.$$

□

Um Satz 15.15 anwenden zu können, ist es wichtig, Kriterien für die Definitheit symmetrischer Matrizen zur Verfügung zu haben.

Bemerkung 15.16 Es sei $A \in \mathbb{R}^{d\times d}$ symmetrisch.
1. Im Falle $d = 2$ ergibt sich ([Ü]): $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ ist genau dann positiv (bzw. negativ) definit, wenn $\det(A) > 0$ und $a > 0$ (bzw. $a < 0$) gilt. Entsprechend ist A genau dann positiv (bzw. negativ) semidefinit, wenn $\det(A) \geq 0$ und $a, c \geq 0$ (bzw. $a, c \leq 0$) gilt.
2. Allgemein gilt (siehe Lineare Algebra): Ist $\sigma(A) = \{\lambda_1, \ldots, \lambda_d\}$ die Menge der Eigenwerte von A (das **Spektrum** von A), so ist

$$\min_{\mathbf{v}\in S^{d-1}} \mathbf{v}^T A \mathbf{v} = \min \sigma(A) =: \lambda_{\min} \quad \text{und} \max_{\mathbf{v}\in S^{d-1}} \mathbf{v}^T A \mathbf{v} = \max \sigma(A) =: \lambda_{\max}\,.$$

Insbesondere folgt daraus: A ist genau dann positiv (semi-)definit, wenn $\lambda_{\min} > 0$ (≥ 0) gilt und genau dann negativ (semi-)definit wenn $\lambda_{\max} < 0$ (≤ 0) gilt.

Beispiel 15.17 1. Ist $f(x, y) = x^2 + y^2$ für $x, y \in \mathbb{R}$ (vgl. Beispiel 15.13.1), 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. Beispiel 15.13.2) so gilt

$$Hf(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Hier ist Hf stets indefinit. Also hat f nach Satz 15.15 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

$$\nabla f(x, y) = (6(x^2 - x)\ ,\ 6(y^2 + y)) = (0, 0)$$

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

$$\begin{aligned}
Hf(0,0) &= \begin{pmatrix} -6 & 0 \\ 0 & 6 \end{pmatrix} && \text{indefinit} \\
Hf(0,-1) &= \begin{pmatrix} -6 & 0 \\ 0 & -6 \end{pmatrix} && \text{negativ definit} \\
Hf(1,0) &= \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix} && \text{positiv definit} \\
Hf(1,-1) &= \begin{pmatrix} 6 & 0 \\ 0 & -6 \end{pmatrix} && \text{indefinit}
\end{aligned}$$

Damit ist f an $(0,-1)$ ein lokales Maximum, an $(1,0)$ ein lokales Minimum und ansonsten keine Extremstellen.

16 Hauptsätze der mehrdimensionalen Analysis

Wir starten wir mit einem allgemeinen Ergebnis über die Existenz von Fixpunkten

Bemerkung und Definition 16.1 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

$$d(\varphi(x), \varphi(x')) \leq \alpha d(x, x') \qquad (x, x' \in M) .$$

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 x_* ein weiterer Fixpunkt von φ, so ist $d(x_*, x^*) = d(\varphi(x_*), \varphi(x^*)) \leq \alpha d(x_*, x^*)$, also $d(x_*, x^*) = 0$.)

Satz 16.2 *(Banachscher Fixpunktsatz)*
Es seien (X, d) ein vollständiger metrischer Raum, $M \subset X$ abgeschlossen und $\varphi : M \to M$. 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 $n \in \mathbb{N}_0$

$$d(x_n, x^*) \leq \alpha^n d(x_0, x^*) \leq \frac{\alpha^n}{1-\alpha} d(x_1, x_0) .$$

Beweis. Für alle $k \in \mathbb{N}$ gilt

$$d(x_{k+1}, x_k) = d(\varphi(x_k), \varphi(x_{k-1})) \leq \alpha \cdot d(x_k, x_{k-1}) \leq \ldots \leq \alpha^k \cdot d(x_1, x_0) .$$

Also ist für $n, n' \in \mathbb{N}$ mit $n' > n$

$$d(x_n, x_{n'}) \leq \sum_{k=n}^{n'-1} d(x_{k+1}, x_k) \leq d(x_1, x_0) \sum_{k=n}^{\infty} \alpha^k = \frac{\alpha^n}{1-\alpha} d(x_1, x_0) .$$

Ist $\varepsilon > 0$, so existiert ein $R > 0$ mit

$$\frac{\alpha^n}{1-\alpha} d(x_1, x_0) < \varepsilon \qquad (n > R) ,$$

also auch $d(x_n, x_{n'}) < \varepsilon$ für $n' > n > R$. 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

$$(1 - \alpha) d(x_0, x^*) \leq d(x_0, x_1) \ .$$

□

Wir beschäftigen uns nun mit der „lokalen Umkehrbarkeit“ stetig differenzierbarer Funktionen f. Im skalaren Fall, also $X \subset \mathbb{R}$ offen und $f \in C^1(X, \mathbb{R})$ gilt:
Ist $a \in X$ regulär, also $f'(a) \neq 0$, so existiert ein $\delta > 0$ so, dass f' entweder > 0 oder < 0 auf $U := (a - \delta, a + \delta)$ ist. Dann ist f streng monoton auf U und damit $f_U : U \to f(U)$ mit $f_U(x) := f(x)$ bijektiv. Nach der Umkehrregel ist zudem $g := f_U^{-1}$ stetig differenzierbar mit $g'(y) = 1/f'(g(y))$ für $y \in f(U)$, also kurz

$$g' = 1/(f' \circ g).$$

Im Weiteren sei stets V *endlich-dimensional*. In diesem Fall ist jede lineare Abbildung A beschränkt und $A \in L(V)$ schon bijektiv mit $A^{-1} \in L(V)$, wenn A surjektiv oder injektiv ist (siehe Lineare Algebra; Dimensionssatz). Ist $V = \mathbb{R}^d$, so ist dies genau dann der Fall, wenn die Determinante $\det A$ nicht verschwindet. Außerdem ist

$$\mathrm{Aut}(V) := \{A \in L(V) : A \text{ bijektiv}\}$$

mit der Komposition $\circ$ eine Gruppe (die **Automorphismengruppe** von V).

Satz 16.3 *Ist* $A \in \mathrm{Aut}(V)$ *und ist* $B \in L(V)$ *mit* $||B - A|| < 1/||A^{-1}||$, *so ist auch* $B \in \mathrm{Aut}(V)$. *Außerdem ist die Abbildung* $\mathrm{Aut}(V) \ni A \mapsto A^{-1} \in \mathrm{Aut}(V)$ *stetig.*

Beweis. Es seien $I := \mathrm{id}_V$ und $0 < r < 1$. Ist $C \in L(V)$ mit $||C|| \leq r/||A^{-1}||$, so gilt $||CA^{-1}|| \leq ||C|| \cdot ||A^{-1}|| \leq r$. Hieraus ergibt sich $I - CA^{-1} \in \mathrm{Aut}(V)$ und

$$(I - CA^{-1})^{-1} = \sum_{\nu=0}^{\infty} (CA^{-1})^{\nu},$$

mit gleichmäßiger Konvergenz der Funktionenreihe $C \mapsto \sum_{\nu=0}^{\infty}(CA^{-1})^\nu$ (Neumannsche Reihe; [Ü]). Nach Satz 9.6 ist $C \mapsto \sum_{\nu=0}^{\infty}(CA^{-1})^\nu$ stetig an 0 mit Funktionswert I. Ist nun $||B-A|| \le 1/||A^{-1}||$ und $C := A-B$, so folgt

$$B = A - C = (I - CA^{-1})A \in \mathrm{Aut}(V)$$

und

$$B^{-1} = A^{-1}(I - CA^{-1})^{-1} = A^{-1}\sum_{\nu=0}^{\infty}(CA^{-1})^\nu.$$

Damit ergibt sich $B^{-1} \to A^{-1}$ für $B \to A$. □

Satz 16.4 *(Hauptsatz über lokale Umkehrbarkeit)*
Es seien $X \subset V$ offen, $f \in C^1(X,V)$ und $a \in X$ eine reguläre Stelle von f. Dann existiert eine offene Umgebung U von a mit folgenden Eigenschaften: Das Bild $f(U)$ ist offen, $f_U : U \to f(U)$ mit $f_U(x) := f(x)$ für $x \in U$ ist bijektiv und $g := f_U^{-1}$ ist stetig differenzierbar mit

$$g' = (f' \circ g)^{-1}.$$

Beweis. Wir setzen $A := f'(a)$. Nach Voraussetzung ist A surjektiv, also schon $A \in \mathrm{Aut}(V)$. Da f' stetig ist, existiert ein $\delta > 0$ so, dass mit $U := U_\delta(a)$

$$||f'(x) - A|| < c := (2||A^{-1}||)^{-1} \qquad (x \in U)\ .$$

Insbesondere ist damit $f'(x) \in \mathrm{Aut}(V)$ für $x \in U$ nach Satz 16.3.
Für $y \in V$ sei $\varphi = \varphi_y : U \to V$ 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. Weiter ergibt sich für $x \in U$

$$\varphi'(x) = I - A^{-1}f'(x) = A^{-1}(A - f'(x))$$

und damit $||\varphi'(x)|| \le ||A^{-1}|| \cdot ||A - f'(x)|| < 1/2$. Nach dem Schrankensatz gilt

$$|\varphi(x) - \varphi(x')| \le |x - x'|/2 \qquad (x, x' \in U).$$

Also ist φ eine 1/2-Kontraktion und damit hat φ höchstens einen Fixpunkt nach Bemerkung 16.1. Folglich ist $f|_U$ injektiv, d. h. die Umkehrfunktion $g : f(U) \to U$ von f_U existiert. Wir zeigen:

1. $f(U)$ ist offen.
2. $g \in C^1(f(U), V)$ mit $g' = (f' \circ g)^{-1}$.

Dazu seien $x \in U$ und $y := f(x)$.
1. Da U offen ist, existiert ein $\rho > 0$ mit $M := B_\rho(x) \subset U$. Ist $z \in U_{c\rho}(y)$ und $\varphi = \varphi_z$, so gilt $\varphi(M) \subset M$, denn für $u \in M$ ist

$$\begin{aligned} |\varphi(u) - x| &\leq |\varphi(u) - \varphi(x)| + |\varphi(x) - x| \\ &\leq |u - x|/2 + |A^{-1}(z - f(x))| \\ &\leq |u - x|/2 + ||A^{-1}|| \cdot |z - y| \leq \rho/2 + ||A||^{-1} c\rho = \rho \ , \end{aligned}$$

also $\varphi(u) \in M$. Da $M \subset V$ abgeschlossen ist, folgt aus dem Banachschen Fixpunktsatz, dass ein $u \in M$ existiert mit $\varphi(u) = u$, also $f(u) = z$. Damit ist $U_{c\rho}(y) \subset f(U)$.
2. Wir zeigen:

$$g'(y) = (f'(x))^{-1} = (f'(g(y))^{-1}.$$

Damit ist g' als Komposition der stetigen Abbildungen g, f' und $A \mapsto A^{-1}$ stetig auf $f(U)$, also g stetig differenzierbar.
Ohne Einschränkung können wir $x = y = 0$ annehmen. Der allgemeine Fall ergibt sich dann wieder durch Anwendung des Spezialfalles auf $\tau_x f$.
Da $\varphi = \varphi_0$ eine 1/2-Kontraktion mit $\varphi(0) = 0$ ist, gilt für alle $u \in U$

$$|u|/2 \geq |\varphi(u)| = |u - A^{-1} f(u)| \geq |u| - \|A^{-1}\| \cdot |f(u)|,$$

und damit $c|u| \leq |f(u)|$. Also gilt $|g(h)| \leq |h|/c$ für $h \in f(U)$ und insbesondere $g(h) \to 0$ für $h \to 0$. Aus der Differenzierbarkeit von f an 0 folgt mit $B := f'(0)$

$$\frac{c}{|h|} |g(h) - B^{-1}(h)| \leq \left\|B^{-1}\right\| \cdot \frac{|B(g(h)) - h|}{|g(h)|} = \|B^{-1}\| \cdot \frac{\left|f(g(h)) - Bg(h)\right|}{|g(h)|} \to 0$$

für $h \to 0$. Also ist auch g differenzierbar an 0 mit $g'(0) = B^{-1}$. □

Ist die Aussage von Satz 16.4 erfüllt, so sagt man, f sei **lokal C^1-umkehrbar** an der Stelle a.

Beispiel 16.5 Ist $V = \mathbb{R}^d$, so ist a eine reguläre Stelle genau dann, wenn $\det Jf(a) \neq 0$ gilt. Wir betrachten $f : (0, \infty) \times \mathbb{R} \to \mathbb{R}^2$ mit

$$f(r, \theta) := \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} \qquad (r > 0,\ \theta \in \mathbb{R}) \ .$$

Dann ist f stetig differenzierbar mit $\det Jf(r, \theta) = r > 0$. Also ist f nach Satz 16.4 lokal C^1-umkehrbar an allen Stellen (r, θ). Da $f(1, 2k\pi) = (1, 0)$ für alle $k \in \mathbb{Z}$ gilt, ist f jedoch nicht injektiv.

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

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

also mit $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ und $y = (y_1, \dots, y_m) \in \mathbb{R}^m$

$$\begin{aligned} F_1(x_1, \dots, x_d,\, y_1, \dots, y_m) &= 0 \\ &\vdots \\ F_m(x_1, \dots, x_d,\, y_1, \dots, y_m) &= 0 \end{aligned}$$

($d + m$ Unbekannte und m Gleichungen). Weiter sei $(\bar{x}, \bar{y}) \in M$ eine Lösung, also $F(\bar{x}, \bar{y}) = 0$. Unser Ziel ist, das Gleichungssystem „lokal nach y aufzulösen“, d. h. wir suchen Umgebungen U von $\bar{x}$ und W von $(\bar{x}, \bar{y})$ sowie eine Funktion $f : U \to \mathbb{R}^m$ so, dass für alle $(x, y) \in W$ gilt: Es ist $F(x, y) = 0$ genau dann, wenn $y = f(x)$, also

$$\{(x, y) \in W : F(x, y) = 0\} = \{(x, f(x)) : x \in U\}.$$

Die Lösungsmenge ist lokal der Graph der Funktion f. Wir orientieren uns an zwei einfachen Beispielen

Beispiel 16.6 1. Wir betrachten die Gleichung

$$F(x, y) := x^2 + y^2 - 1 = 0 \qquad (x, y \in \mathbb{R})\ .$$

Dann ist offenbar $(\bar{x}, \bar{y}) = (1/\sqrt{2}, 1/\sqrt{2})$ eine Lösung der Gleichung. Hier gilt etwa für $U := (-1, 1)$ und $W := (-1, 1) \times (0, \infty)$

$$F(x, y) = 0 \Leftrightarrow y = \sqrt{1 - x^2} =: f(x)\ .$$

Dies ist 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 das lineare Gleichungssystem $Ax + By = 0$). Dann gilt im Falle $\det B \neq 0$

$$F(x, y) = 0 \Leftrightarrow By = -Ax \Leftrightarrow y = -B^{-1}Ax =: f(x)$$

also $\{(x, y) : F(x, y) = 0\} = \{(x, f(x)) : x \in \mathbb{R}^d\}$.

Satz 16.7 *(Hauptsatz über implizite Funktionen)*
Es seien $\Omega \subset \mathbb{R}^d \times \mathbb{R}^m$ offen und $F \in C^1(\Omega, \mathbb{R}^m)$. Weiter sei $(\bar{x}, \bar{y}) \in \Omega$ mit $F(\bar{x}, \bar{y}) = 0$ und so, dass $\bar{y}$ regulär für $y \mapsto F(\bar{x}, y)$ ist. Dann existieren offene Umgebungen U von $\bar{x}$ und W von $(\bar{x}, \bar{y})$ sowie eine Funktion $f \in C^1(U, \mathbb{R}^m)$ so, dass

$$\{(x, y) \in W : F(x, y) = 0\} = \{(x, f(x)) : x \in U\}.$$

Mit $\partial F/\partial y := (\partial F_j/\partial y_k)_{j,k=1,\dots,m}$ und $\partial F/\partial x := (\partial F_j/\partial x_k)_{\substack{j=1,\dots,m \\ k=1,\dots,d}}$ ist zudem

$$Jf(x) = -\left(\frac{\partial F}{\partial y}(x, f(x))\right)^{-1} \frac{\partial F}{\partial x}(x, f(x)) \qquad (x \in U) .$$

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 $G(\bar{x}, \bar{y}) = (\bar{x}, 0)$ und $G \in C^1(\Omega, \mathbb{R}^d \times \mathbb{R}^m)$. Bezeichnet E_d die d-dimensionale Einheitsmatrix, so ist

$$JG = \begin{pmatrix} E_d & 0 \\ \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \end{pmatrix}$$

und damit $\det JG(\bar{x}, \bar{y}) = \det \frac{\partial F}{\partial y}(\bar{x}, \bar{y}) \neq 0$. Nach Satz 16.4 ist G lokal C^1-umkehrbar an der Stelle $(\bar{x}, \bar{y})$. Damit existieren ein $\delta > 0$ und eine offene Umgebung W von $(\bar{x}, \bar{y})$ so, dass $G_W : W \to U_\delta(\bar{x}) \times U_\delta(0)$ bijektiv ist mit stetig differenzierbarer Umkehrfunktion und $\det JG(x, y) \neq 0$ für $(x, y) \in W$.
Wir setzen $U := U_\delta(\bar{x})$ und schreiben im Weiteren kurz G statt G_W. Aus der Definition von G ergibt sich, dass G^{-1} mit einer geeigneten Funktion $H \in C^1(U \times U_\delta(0), \mathbb{R}^m)$ von der Form

$$G^{-1}(x, z) = (x, H(x, z)) \qquad ((x, z) \in U \times U_\delta(0))$$

ist. Definiert man $\pi : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^m$ durch $\pi(x, y) := y$, so gilt

$$\pi(G(x, y)) = \pi(x, F(x, y)) = F(x, y) \qquad ((x, y) \in W)$$

und folglich

$$F(x, H(x, z)) = \pi(G(x, H(x, z))) = \pi(x, z) = z \quad ((x, z) \in U \times U_\delta(0)) .$$

Nun definieren wir $f : U \to \mathbb{R}^m$ durch

$$f(x) := H(x, 0) \qquad (x \in U).$$

Aus $H \in C^1(U \times U_\delta(0), \mathbb{R}^m)$ ergibt sich $f \in C^1(U, \mathbb{R}^m)$. Außerdem gilt

$$F(x, f(x)) = 0 \qquad (x \in U)$$

und für $(x, y) \in W$ mit $F(x, y) = 0$ folgt

$$G(x, y) = (x, 0) = G(G^{-1}(x, 0)) = G(x, f(x))$$

und damit auch $y = f(x)$.
2. Aus $h(x) := F(x, f(x)) = 0$ für $x \in U$ ergibt sich mit der Kettenregel

$$0 = Jh(x) = JF(x, f(x)) \cdot \begin{pmatrix} E_d \\ Jf(x) \end{pmatrix} = \frac{\partial F}{\partial x}(x, f(x)) + \frac{\partial F}{\partial y}(x, f(x)) Jf(x)$$

(„implizites Differenzieren").
Aus $\det \frac{\partial F}{\partial y}(x, y) = \det JG(x, y) \neq 0$ für alle $(x, y) \in W$ ergibt sich die Zusatzbehauptung durch Auflösen der letzten Gleichung nach $Jf(x)$. □

Beispiel 16.8 (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 $0 = \frac{\partial F}{\partial y}(x, y)$ und $0 = F(x, y)$ genau dann, wenn $(x, y) = (0, 0)$ oder $(x, y) = (\pm\sqrt{2}, 0)$ ([Ü]). Nach Satz 16.7 ist für alle $(\bar{x}, \bar{y})$ mit $F(\bar{x}, \bar{y}) = 0$ und $(\bar{x}, \bar{y}) \notin \{(0, 0), (\pm\sqrt{2}, 0)\}$ die Gleichung $F(x, y) = 0$ auf einer Umgebung W von $(\bar{x}, \bar{y})$ auflösbar nach y. Außerdem ergibt sich für die Funktion $f = f_{(\bar{x}, \bar{y})}$

$$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))} = -\frac{x(x^2 + f^2(x) - 1)}{f(x)(x^2 + f^2(x) + 1)}$$

auf einer Umgebung U von $\bar{x}$. Also hat f Extremstellen höchstens in Punkten x mit

$$x^2 + f^2(x) = 1 \ .$$

Um eine Vorstellung von der Lösungsmenge zu bekommen, betrachten wir Polarkoordinaten: Es gilt für $(x, y) \neq (0, 0)$

$$0 = F(x, y) = F(r \cos\theta, r \sin\theta) = r^4 - 2r^2(\cos^2\theta - \sin^2\theta) = r^2(r^2 - 2\cos 2\theta)$$

genau dann, wenn $r = \sqrt{2\cos 2\theta}$, wobei θ so, dass $\cos(2\theta) > 0$ ist.

Eine wichtige Anwendung des Hauptsatzes über implizite Funktionen ergibt sich im Bereich der Optimierung unter Nebenbedingungen.

Definition 16.9 Es seien (X, d) ein metrischer Raum, E Vektorraum und $f : X \to \mathbb{R}$ sowie $g : X \to E$. Ist

$$L := \{x \in X : g(x) = 0\}$$

und ist $a \in L$, so heißt a eine **Extremstelle von f unter der Nebenbedingung** $g(x) = 0$, falls a Extremstelle von $f|_L$ ist. Man spricht dann natürlich auch wieder von Maximal- bzw. Minimalstelle, je nach dem ob a maxial oder minimal für $f|_L$ ist.

Satz 16.10 *(Lagrange)*
Es seien $X \subset \mathbb{R}^d$ offen und $f \in C^1(X, \mathbb{R})$. Weiter seien $g \in C^1(X, \mathbb{R}^m)$ mit $m < d$ und $a \in X$ eine reguläre Stelle von g. Ist a eine Extremstelle von f unter der Nebenbedingung $g(x) = 0$, so existiert ein $\lambda \in \mathbb{R}^m$ mit

$$\nabla f(a) = \sum_{j=1}^{m} \lambda_j \, \nabla g_j(a) = (Jg)^T(a)\lambda \, .$$

Beweis. Wir schreiben $x = (u, y)$ mit

$$u = (x_1, \ldots, x_{d-m}) \, , \; y = (x_{d-m+1}, \ldots, x_d)$$

für $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ und $(\bar{u}, \bar{y}) := a$. Nach geeigneter Permutation der Komponenten können wir annehmen, dass $\bar{y}$ regulär ist für die Abbildung $y \mapsto g(\bar{u}, y)$, also $\det \frac{\partial g}{\partial y}(a) \neq 0$ gilt (Lineare Algebra). Nach dem Hauptsatz über implizite Funktionen ist dann die Gleichung

$$g(x) = g(u, y) = 0$$

an a lokal nach y auflösbar. Insbesondere existieren eine offene Umgebung U von $\bar{u}$ sowie eine Funktion $\varphi \in C^1(U, \mathbb{R}^m)$ so, dass $\varphi(\bar{u}) = \bar{y}$ und

$$g(u, \varphi(u)) = 0 \qquad (u \in U)$$

gilt. Hieraus folgt durch implizites Differenzieren

$$0 = \frac{\partial g}{\partial u}(a) + \frac{\partial g}{\partial y}(a) J\varphi(\bar{u}).$$

Ist $h \in C^1(U, \mathbb{R})$ definiert durch

$$h(u) := f(u, \varphi(u)) \qquad (u \in U) \; ,$$

so ist $\bar{u}$ eine Extremstelle von h (ohne Nebenbedingung). Also gilt nach Satz 15.12

$$Jh(\bar{u}) = (\nabla h)^T(\bar{u}) = 0$$

und mit der Kettenregel folglich

$$0 = \frac{\partial f}{\partial u}(a) + \frac{\partial f}{\partial y}(a) J\varphi(\bar{u}) \, .$$

Wegen $\det \frac{\partial g}{\partial y}(a) \neq 0$ hat das lineare Gleichungssystem

$$\lambda^T \cdot \frac{\partial g}{\partial y}(a) = \frac{\partial f}{\partial y}(a)$$

genau eine Lösung $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$. Hieraus ergibt sich wiederum

$$\frac{\partial f}{\partial u}(a) = -\frac{\partial f}{\partial y}(a) J\varphi(\bar{u}) = -\lambda^T \frac{\partial g}{\partial y}(a) J\varphi(\bar{u}) = \lambda^T \frac{\partial g}{\partial u}(a)$$

und damit insgesamt $(\nabla f)^T(a) = \lambda^T \cdot Jg(a)$. □

Bemerkung 16.11 Ähnlich wie Satz 15.12 liefert Satz 16.10 lediglich eine notwendige Bedingung für das Vorliegen einer Extremstelle unter der Nebenbedingung $g(x) = 0$. Um die entsprechenden Punkte a zu bestimmen, hat man die Gleichungen

$$g_j(x_1, \ldots, x_d) = 0 \qquad (j = 1, \ldots, m)$$

und

$$\partial_k f(x_1, \ldots, x_d) = \sum_{j=1}^{m} \lambda_j \partial_k g_j(x_1, \ldots, x_d) \qquad (k = 1, \ldots, d)$$

zu lösen (also $(d+m)$ Gleichungen für die $(d+m)$ Unbekannten $x_1, \ldots, x_d$ und $\lambda_1, \ldots, \lambda_m$). Die Zahlen $\lambda_1, \ldots, \lambda_m$ heißen **Lagrange-Multiplikatoren**.

Beispiel 16.12 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 Satz 16.10 ist eine notwendige Bedingung gegeben durch

$$\nabla K(x, y) = \lambda \, \nabla P(x, y)$$

für ein $\lambda \in \mathbb{R}$, d. h. wir haben die drei Gleichungen

$$\begin{aligned} x^\alpha y^{1-\alpha} &= c \\ p &= \lambda \alpha x^{\alpha-1} y^{1-\alpha} \\ q &= \lambda (1-\alpha) x^\alpha y^{-\alpha} \end{aligned}$$

für die Unbekannten x, y, λ. Division der zweiten und dritten Gleichung ergibt

$$x = \frac{\alpha}{1-\alpha} \cdot \frac{q}{p} \cdot y$$

und mit der ersten 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

Die **natürlichen Zahlen** können axiomatisch beschrieben werden als ein Tripel $(\mathbb{N}, 1, \nu)$ mit den drei Eigenschaften (**Peano-Axiome**):

(N1) $\mathbb{N}$ ist eine Menge mit $1 \in \mathbb{N}$.

(N2) $\nu : \mathbb{N} \to \mathbb{N}$ ist eine injektive Funktion mit $1 \notin \nu(\mathbb{N})$.

(N3) (Prinzip der vollständigen Induktion) Ist $A \subset \mathbb{N}$ mit $1 \in A$ und $\nu(A) \subset A$, so ist $A = \mathbb{N}$.

Damit kann man zeigen: Auf $\mathbb{N}$ existiert genau eine assoziative und kommutative Verknüpfung $+$ mit

$$n + 1 = \nu(n) \quad \text{und} \quad n + \nu(m) = \nu(n + m)$$

für $n, m \in \mathbb{N}$. Unter Verwendung der Addition ist eine Ordnungrelation $<$ auf $\mathbb{N}$ definiert durch $n < m$ genau dann, wenn $m = n + k$ für ein $k \in \mathbb{N}$. Weiter kann man zeigen: Auf $\mathbb{N}$ existiert genau eine assoziative und kommutative Verknüpfung $\cdot$ so, dass 1 neutral bezüglich $\cdot$ ist und dass

$$m(n + 1) = mn + m$$

für $n, m \in \mathbb{N}$ gilt. Erweitert man $\mathbb{N}$ um ein Element 0 zu $\mathbb{N}_0$ mit $0 < n$ für alle $n \in \mathbb{N}$ und so, dass 0 neutral ist bezüglich $+$ und absorbierend bezüglich $\cdot$, so sind $(\mathbb{N}_0, +, 0)$ und $(\mathbb{N}_0, \cdot, 1)$ Monoide mit den **Kürzungsregeln**

$$n + m = n + k \;\Rightarrow\; m = k \quad \text{und} \quad n \cdot m = n \cdot k,\, n \neq 0 \;\Rightarrow\; m = k$$

Schließlich folgt aus dem Prinzip der vollständigen Induktion die wichtige Wohlordnungseigenschaft von $\mathbb{N}$: *Jede nichtleere Menge $M \subset \mathbb{N}$ hat ein Minimum.*
Hiermit kann man leicht zeigen : Zu jedem Paar $(n, p) \in \mathbb{N} \times \mathbb{N}$ existiert genau ein Paar $(a, r) \in \mathbb{N}_0 \times \mathbb{N}_0$ mit $r < p$ und $n = ap + r$ (Division mit Rest).
Ist $(M, +, 0)$ ein abelsches Monoid, so definieren wir für nicht notwendig endliche Indexmengen I und nichtleere $A \subset M$

$$A^{(I)} := \{x = (x_\alpha)_{\alpha \in I} \in A^I : I_x := \{\alpha \in I : x_\alpha \neq 0\} \text{ endlich}\}$$

sowie

$$\sum_{\alpha \in I} x_\alpha := \sum_{\alpha \in I_x} x_\alpha \quad (x = (x_\alpha)_{\alpha \in I} \in M^{(I)}).$$

Damit gilt folgende wichtige Aussage über die *Darstellung* natürlicher Zahlen:

Satz A.1 *Es seien $q \in \mathbb{N}$ mit $q \geq 2 := 1 + 1$ und $A := \{a \in \mathbb{N}_0 : a < q\}$. Dann existiert für jedes $n \in \mathbb{N}_0$ genau ein Tupel $a(n) = (a_j(n))_{j\in\mathbb{N}_0} \in A^{(\mathbb{N}_0)}$ mit*

$$n = \sum_{j\in\mathbb{N}_0} a_j(n)q^j.$$

Beweis. Die Eindeutigkeit folgt aus B. 2.7. Wir zeigen die Existenz per Induktion nach n.
1. Induktionsanfang $n = 0$: Man setze $a_j(0) := 0$ für $j \in \mathbb{N}_0$.
2. Induktionssschritt: Es sei $k \in \mathbb{N}_0$ mit $q^k \leq n < q^{k+1}$. Division mit Rest ergibt

$$n = aq^k + n'$$

mit $0 < a < q$ und $0 \leq n' < q^k$, also insbesondere $n' < n$.
Nach Induktionsvoraussetzung (Behauptung gilt für jedes $n' < n$) existiert eine Folge $\big(a_j(n')\big)$ mit

$$n' = \sum_{j\in\mathbb{N}_0} a_j(n')q^j.$$

Dabei ist $a_j(n') = 0$ für $j \geq k$, da $n' < q^k$. Setzt man

$$a_j(n) := \begin{cases} a_j(n') & \text{für } j \neq k \\ a & \text{für } j = k \end{cases},$$

so ist

$$n = aq^k + n' = \sum_{j\in\mathbb{N}_0} a_j(n)q^j.$$

□

Mit $d := d(n) := \max\{j \in \mathbb{N}_0 : a_j(n) \neq 0\}$ für $n \neq 0$ heißt $(a_d(n)a_{d-1}(n)\ldots a_0(n))_q$ die **q-adische Darstellung** von n. Die Zahlen $0, \ldots, q-1$ sind die Ziffern bezüglich q. Man setzt $3 := 2 + 1$, $4 := 3 + 1$, $5 := 4 + 1$, $6 := 5 + 1$, $7 := 6 + 1$, $8 := 7 + 1$ und $9 := 8 + 1$. Im Falle $q = 2 \cdot 5$ spricht man dann von der **Dezimal-**, im Falle $q = 2$ von der **Binär-** und im Falle $q = 2^4$ von der **Hexadezimaldarstellung**. Schließlich schreibt man im Dezimalfall auch kurz $a_d(n)\ldots a_0(n)$ statt $(a_d(n)\ldots a_0(n))_{2\cdot 5}$. So ist etwa für $n = 8 + 8 + 7$

$$n = 1 \cdot 2^4 + 0 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = (10111)_2$$

und

$$n = 2 \cdot (2 \cdot 5)^1 + 3 \cdot (2 \cdot 5)^0 = (23)_{2\cdot 5} = 23\,.$$

Definition A.2 Eine Relation $\sim$ in X heißt **Äquivalenzrelation** (auf X), falls für alle $x, y, z \in X$ gilt

(A1) $x \sim x$ (Reflexivität),

(A2) aus $x \sim y$ folgt $y \sim x$ (Symmetrie),

(A3) aus $x \sim y$ und $y \sim z$ folgt $x \sim z$ (Transitivität).

Ist $\sim$ eine Äquivalenzrelation, so heißt $[x] := [x]_\sim := \{x' \in X : x \sim x'\}$ die von x erzeugte **Äquivalenzklasse** und jedes $x \in [x]$ ein **Repräsentant** der Äquivalenzklasse $[x]$. Außerdem heißt $X/_\sim := \big\{[x] : x \in X\big\}$ **Quotientenmege** von X (**modulo** $\sim$).

Bemerkung und Definition A.3 1. Es sei $X = \mathbb{N}_0 \times \mathbb{N}_0$. Definiert man

$$(a, b) \sim (c, d) \quad :\Leftrightarrow \quad a + d = b + c,$$

so ist $\sim$ eine Äquivalenzrelation auf X (ergibt sich aus Rechenregeln für die Addition in $\mathbb{N}_0$).

$$
\begin{array}{c|ccccccccc}
 & & \vdots & & \vdots & & \vdots & & \vdots & \\
 & & & \swarrow & & \swarrow & & \swarrow & & \swarrow \\
-4 & & (0,3) & & (1,3) & & (2,3) & & (3,3) & \cdots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow \\
-3 & & (0,2) & & (1,2) & & (2,2) & & (3,2) & \cdots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow \\
-2 & & (0,1) & & (1,1) & & (2,1) & & (3,1) & \cdots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow \\
-1 & & (0,0) & & (1,0) & & (2,0) & & (3,0) & \cdots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow \\
\hline
0 & & 1 & & 2 & & 3 & & 4 &
\end{array}
$$

Sind $a, b \in \mathbb{N}_0$ mit $a \geq b$, so existiert (genau) ein $n \in \mathbb{N}_0$ mit $a = b + n$ und es gilt damit

$$[(a, b)] = \{(n + b, b) : b \in \mathbb{N}_0\} = [(n, 0)]$$

Ist $a < b$, so ist $b = a + m$ für ein $m \in \mathbb{N}$ und damit

$$[(a, b)] = \{(a, a + m) : a \in \mathbb{N}_0\} = [(0, m)].$$

Definiert man die Menge $\mathbb{Z}$ der ganzen Zahlen als

$$\mathbb{Z} := X/_{\sim} = (\mathbb{N}_0 \times \mathbb{N}_0)/_{\sim},$$

so sind durch

$$[(a, b)] + [(c, d)] := [(a + c, b + d)] \quad \text{und} \quad [(a, b)] \cdot [(c, d)] := [(ac + bd, ad + bc)]$$

Verknüpfungen $+$ und $\cdot$ auf $\mathbb{Z}$ (unabhängig von der Wahl der jeweiligen Repräsentanten) definiert, mit denen $(\mathbb{Z}, +, \cdot)$ zu einem kommutativen Ring wird. Dabei gilt $[(0, n)] = -[(n, 0)]$ für $n \in \mathbb{N}_0$. Indem man n mit $[(n, 0)]$ identifizert, ist $\mathbb{N}_0$ in $\mathbb{Z}$ eingebettet, und es ergibt sich

$$[(a, b)] = a - b \quad (a, b \in \mathbb{N}_0).$$

Außerdem ist damit durch

$$a - b < c - d :\Leftrightarrow a + d < b + c$$

für $a, b, c, d \in \mathbb{N}_0$ eine Erweiterung der Ordnung $<$ von $\mathbb{N}_0$ auf $\mathbb{Z}$ definiert.

Bemerkung und Definition A.4 Mit $\mathbb{Z}^* := Z \setminus \{0\}$ sei $X = \mathbb{Z} \times \mathbb{Z}^*$. Definiert man

$$(a, b) \sim (c, d) \quad :\Leftrightarrow \quad ad = bc,$$

so ist $\sim$ eine Äquivalenzrelation auf X (ergibt sich aus Rechenregeln für die Multiplikation in $\mathbb{Z}$). Mit etwas Zahlentheorie kann man zeigen: Für $(a, b) \in X$ existieren teilerfremde $p \in \mathbb{Z}$, $q \in \mathbb{N}$ mit

$$[(a, b)] = \{(mp, mq) : m \in \mathbb{Z}\} = [(p, q)].$$

Definiert man die Menge $\mathbb{Q}$ der rationalen Zahlen als

$$\mathbb{Q} := X/_{\sim} = \big(\mathbb{Z} \times \mathbb{Z}^*\big)/_{\sim}$$

und Verknüpfungen $+$ und $\cdot$ durch

$$[(a, b)] + [(c, d)] := [(ad + cb, bd)] \quad \text{und} \quad [(a, b)] \cdot [(c, d)] := [(ac, bd)],$$

so wird $(\mathbb{Q}, +, \cdot)$ zu einem Körper. Dabei gilt $1/[(a, b)] = [(b, a)]$ für $a \neq 0$. Durch Identifikation von a und $[(a, 1)]$ ist wieder $\mathbb{Z}$ in $\mathbb{Q}$ eingebettet und es gilt

$$[(a, b)] = a/b \quad ((a, b) \in X).$$

Schließlich erweitert die Setzung

$$\frac{a}{b} < \frac{c}{d} :\Leftrightarrow ad < bc$$

für $a, b \in \mathbb{Z}$ und $b, d \in \mathbb{N}$ (also $b, d > 0$) die Ordnung $<$ von $\mathbb{Z}$ auf $\mathbb{Q}$.

Bemerkung und Definition A.5 Es seien $q \in \mathbb{N}$, $q \geq 2$ und $A := \{0, 1, \ldots, q-1\}$. Wir betrachten die Menge

$$A^{\mathbb{Z})} := \{(a_j) = (a_j)_{j \in \mathbb{Z}} \in A^{\mathbb{Z}} : \text{es existiert ein } d \in \mathbb{Z} \text{ mit } a_j = 0 \text{ für } j > d\}.$$

Ein Tupel $(a_j) \in A^{\mathbb{Z})}$ nennen wir eine **q-adische Entwicklung**. Im Falle $q = 2$ sprechen wir von **Binärentwicklung** und im Falle $q = 10$ von **Dezimalentwicklung**. Die q-adische Entwicklung (a_j) heißt **abbrechend**, falls $(a_j) \in A^{(\mathbb{Z})}$ gilt, also falls ein $m \in \mathbb{Z}$ existiert mit $a_j = 0$ für $j < m$. Aus S. A.1 folgt, dass die Abbildung

$$A^{(\mathbb{Z})} \ni (a_j) \mapsto \sum_{j \in \mathbb{Z}} a_j q^j \in \bigcup_{m \in \mathbb{Z}} q^m \mathbb{N}_0 \subset \mathbb{Q}_+ \cup \{0\}$$

bijektiv ist. Wir identifizieren im Weiteren die abbechende Entwicklung (a_j) mit der entsprechenden rationalen Zahl. Da sowohl die Addition als auch die Multiplikation Verknüpfungen auf $\bigcup_{m \in \mathbb{Z}} q^m \mathbb{N}_0$ sind, können wir damit abbrechende Entwicklungen addieren und multiplizieren.
Weiter definieren wir auf $A^{\mathbb{Z})}$ eine Äquivalenzrelation durch $(a_j) \sim (b_j)$ genau dann, wenn $(a_j) = (b_j)$ oder wenn ein $m \in \mathbb{Z}$ so existiert, dass $a_j = b_j$ für $j > m$, $a_m = b_m + 1$ sowie $a_j = 0$ und $b_j = q - 1$ für $j < m$ (oder entsprechend mit vertauschten Rollen von a_j und b_j). Damit sind alle Äquivalenzklassen entweder ein- oder zweielementig, wobei im zweielementigen Fall eine der beiden Entwicklungen abbrechend ist.

Bemerkung A.6 Wir betrachten nun den Fall $q = 2$ und definieren

$$X := \{x = [(a_j)] : (a_j) \in \{0, 1\}^{\mathbb{Z})}\}.$$

Wählt man im Fall zweielementiger $[(a_j)]$ die abbrechende Entwicklung als Repräsentant, so entspricht jedem $x \in X$ genau eine Binärentwicklung (a_j). Wenn nichts anders gesagt ist, legen wir uns auf diese Darstellung fest und schreiben dann auch $x = (a_j)$. Ist $x = (a_j)$ (in diesem Sinne), so nennen wir für $m \in \mathbb{Z}$

$$\lfloor x \rfloor_m := \sum_{j \geq m} a_j 2^j \in 2^m \mathbb{N}_0$$

die m-te Abschneidung von x. Unter Verwendung der Relation $<$ auf $\mathbb{Q}_+ \cup \{0\}$ definieren wir für $x, y \in X$ mit $x \neq y$

$$x < y :\Leftrightarrow \lfloor x \rfloor_m < \lfloor y \rfloor_m \text{ für ein } m \in \mathbb{Z}.$$

Dann ist auch $\lfloor x \rfloor_n < \lfloor y \rfloor_n$ für $n \leq m$ (ergibt sich aus B. 2.7) und $\lfloor x \rfloor_k \leq \lfloor y \rfloor_k$ für alle $k \in \mathbb{Z}$. Man kann zeigen, dass damit $(X, <)$ geordnet und *vollständig* ist.
Wir wollen den Beweis der Vollständigkeit andeuten, indem wir eine Konstruktionsvorschrift für $\sup M$ angeben. Es sei also $M \subset X$ nichtleer und nach oben beschränkt. Wir definieren ein $s \in X$ rekursiv: Ohne Einschränkung können wir annehmen, dass ein $M \neq \{0\}$ ist. Zunächst folgt aus der Beschränktheit nach oben von M die Beschränktheit nach oben von

$$\{m \in \mathbb{Z} : \lfloor x \rfloor_m \neq 0 \text{ für ein } x \in M\}$$

in $\mathbb{Z}$. Ist d das Maximum dieser Menge, so setzen wir $\xi_d := 1$ und $\xi_j := 0$ $(j > d)$. Weiter definieren wir

$$\xi_{d-1} := \begin{cases} 1, & \text{falls } \lfloor x \rfloor_{d-1} > 2^d \text{ für ein } x \in M \\ 0, & \text{sonst} \end{cases}$$

und entsprechend für $n \in \mathbb{N}$

$$\xi_{d-n-1} := \begin{cases} 1, & \text{falls } \lfloor x \rfloor_{d-n-1} > 2^d + \xi_{d-1} 2^{d-1} + \ldots + \xi_{d-n} 2^{d-n} \text{ für ein } x \in M \\ 0, & \text{sonst} \end{cases}.$$

Hierdurch ist induktiv eine Binärentwicklung $(\xi_j)_{j \in \mathbb{Z}}$ definiert. Ist $s := [(\xi_j)]$, so ergibt sich $s = \sup M$ aus der Konstruktion von s.
Damit wird es ermöglicht, die Verknüpfungen $+$ und $\cdot$ von $A^{(\mathbb{Z})}$ auf X zu erweitern: Da $(X, <)$ vollständig ist, existieren nämlich für $x, y \in X$

$$x + y := \sup \left\{ \lfloor x \rfloor_m + \lfloor y \rfloor_m : m \in \mathbb{Z} \right\} \in X$$

und

$$x \cdot y := \sup \left\{ \lfloor x \rfloor_m \cdot \lfloor y \rfloor_m : m \in \mathbb{Z} \right\} \in X.$$

Das Hauptproblem besteht nun darin, zu zeigen, dass $(X, +, 0)$ ein Monoid mit der Kürzungseigenschaft und $(X \setminus \{0\}, \cdot, 1)$ eine Gruppe ist. Ist dies getan, so sieht man wie bei der Erweiterung von $\mathbb{N}_0$ zu $\mathbb{Z}$, dass durch

$$(a, b) \sim (c, d) \quad :\Leftrightarrow \quad a + d = b + c$$

für $(a, b), (c, d) \in X \times X$ eine Äquivalenzrelation auf $X \times X$ definiert ist. Damit setzt man

$$\mathbb{R} := (X \times X)/_{\sim}$$

und 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 $(\mathbb{R}, +, \cdot, <)$ ein vollständiger geordneter Körper wird. Dabei entspricht $\mathbb{Q}$ der Menge der periodischen Binärentwicklungen und ist in diesem Sinne eingebettet in $\mathbb{R}$

B Mächtigkeit von Mengen

Wir werden im Folgenden zeigen, dass in gewissem Sinne sehr viele reelle Zahlen irrational sind. Dazu beschäftigen wir uns zunächst mit dem Begriff der Mächtigkeit von Mengen.

Definition B.1 Es seien A, B beliebige Mengen.

1. A und B heißen **von gleicher Mächtigkeit** (oder kurz **gleichmächtig**), falls eine bijektive Abbildung $\varphi : A \to B$ existiert.

2. A hat **Mächtigkeit** $n \in \mathbb{N}$, falls A gleichmächtig zu $\{1, \ldots, n\}$ ist, d. h. falls ein Tupel $(a_1, \ldots, a_n)$ existiert mit $a_j \neq a_k$ für $j \neq k$ und $A = \{a_1, \ldots, a_n\}$ (man kann zeigen, dass n eindeutig ist). Die leere Menge hat die Mächtigkeit 0. Damit heißt A **endlich**, falls A eine Mächtigkeit $n \in \mathbb{N}_0$ hat. Wir schreiben dann $\#A := n$.

3. A heißt **abzählbar unendlich**, falls A gleichmächtig zu $\mathbb{N}$ ist, d. h. falls eine Folge $(a_n)_{n\in\mathbb{N}}$ existiert mit $a_j \neq a_k$ für $j \neq k$ und $A = \{a_n : n \in \mathbb{N}\}$.

4. A heißt **abzählbar**, falls A endlich oder abzählbar unendlich ist. Anderenfalls heißt A **überabzählbar**.

Satz B.2 *Eine nichtleere Menge A ist genau dann abzählbar, wenn eine Folge $(a_n)_{n\in\mathbb{N}}$ existiert mit $A = \{a_n : n \in \mathbb{N}\}$.*

Beweis. $\Rightarrow$: Ist A abzählbar unendlich, so existiert eine solche Folge nach Definition. Ist A endlich, etwa $A = \{a_1, \ldots, a_n\}$, so setze man $a_j := a_n$ für $j > n$.
$\Leftarrow$: Ist A endlich, so ist A abzählbar. Es sei also $A = \{a_n : n \in \mathbb{N}\}$ unendlich. Wir setzen $n_1 := 1$. Sind $n_1, \ldots, n_k$ bereits definiert, so existiert nach dem Wohlordnungsprinzip

$$n_{k+1} := \min\{n \in \mathbb{N} : a_n \notin \{a_{n_1}, \ldots, a_{n_k}\}\}.$$

Damit ist eine streng wachsend Folge (n_k) in $\mathbb{N}$ definiert. Setzt man $b_k = a_{n_k}$, so ist nach Konstruktion $b_j \neq b_k$ für $j \neq k$ und $\{b_k : k \in \mathbb{N}\} = A$. Also ist A abzählbar unendlich. $\square$

Bemerkung B.3 Aus S. B.2 folgt sofort, dass Bilder abzählbarer Mengen unter beliebigen Abbildungen wieder abzählbar sind und damit auch, dass Teilmengen abzählbarer Mengen abzählbar sind.

Satz B.4 *Es sei $J \neq \emptyset$ abzählbar. Ist $(A_j)_{j\in J}$ eine Familie abzählbarer Mengen, so ist auch $\bigcup\limits_{j\in J} A_j$ abzählbar.*

Beweis. Ohne Einschränkung können wir $A_j \neq \emptyset$ für alle $j \in J$ annehmen. Nach S. B.2 existieren Folgen $(a_n^{(j)})_{n\in\mathbb{N}}$ mit

$$A_j = \{a_n^{(j)} : n \in \mathbb{N}\} \qquad (j \in \mathbb{N}).$$

Außerdem existiert eine Folge (j_k) mit $J = \{j_k : k \in \mathbb{N}\}$. Wir setzen

$$B_m := \{a_n^{(j_k)} : k, n = 1, \dots, m\} \qquad (m \in \mathbb{N}).$$

Da $\#(B_m) \leq m^2$ und $B_m \subset B_{m+1}$ für alle $m \in \mathbb{N}$ gilt, ergibt sich induktiv die Existenz einer Folge $(b_\ell)_{\ell\in\mathbb{N}}$ mit $B_m = \{b_1, \dots, b_{m^2}\}$ für alle $m \in \mathbb{N}$. Nach Konstruktion ist

$$\bigcup_{j\in J} A_j = \bigcup_{m\in\mathbb{N}} B_m = \{b_\ell : \ell \in \mathbb{N}\}.$$

Mit S. B.2 folgt die Behauptung. □

Bemerkung B.5 Da $\mathbb{Z}$ abzählbar ist ([Ü]), ist $\mathbb{Z}\times\{n\}$ abzählbar für beliebiges $n \in \mathbb{Z}$ ($\mathbb{Z} \ni j \mapsto (j,n) \in \mathbb{Z}\times\{n\}$ ist bijektiv). Damit ist nach S. B.4 auch

$$\mathbb{Z}\times\mathbb{Z}^* = \bigcup_{n\in\mathbb{Z}^*} \big(\mathbb{Z}\times\{n\}\big)$$

abzählbar. Nach B. B.3 ist schließlich auch $\mathbb{Q}$ abzählbar (beachte: $\mathbb{Q} = \varphi(Z\times\mathbb{Z}^*)$ für $\varphi(a,b) := [(a,b)]$).

Wir wollen nun zeigen, dass $\mathbb{R}$ überabzählbar ist. Dazu beweisen wir zunächst:

Satz B.6 *(Intervallschachtelungsprinzip)*
Es sei (I_n) eine Folge von Intervallen der Form $I_n = [a_n, b_n]$ mit $I_{n+1} \subset I_n$ $(n \in \mathbb{N})$. Dann ist $c := \sup\{a_n : n \in \mathbb{N}\} \leq \inf\{b_n : n \in \mathbb{N}\} =: d$ und

$$\bigcap_{n\in\mathbb{N}} I_n = [c, d].$$

Gilt zusätzlich $b_n - a_n \to 0$ $(n \to \infty)$, so ist $c = d$.

Beweis. $\subset$: Ist $x \in \bigcap\limits_{n\in\mathbb{N}} I_n$, so folgt aus $a_n \le x \le b_n$ für alle n auch $c \le x \le d$ nach Definition des Supremums und des Infimums.
$\supset$: Ist $m \in \mathbb{N}$, so gilt nach Voraussetzung $a_n \le b_n \le b_m$ für $n \ge m$. Da (a_n) wachsend ist, folgt $c \le b_m$ und damit auch $c \le d$, also insgesamt

$$a_n \le c \le d \le b_n \qquad (n \in \mathbb{N}).$$

Folglich ist

$$[c, d] \subset \bigcap_{n\in\mathbb{N}} I_n \ .$$

Aus $0 \le d - c \le b_n - a_n$ ergibt sich $d - c = 0$ im Falle $b_n - a_n \to 0$. □

Satz B.7 *Jedes (nicht einpunktige) Intervall $I \subset \mathbb{R}$ 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 $I_0 := [0, 1]$ in die drei Intervalle $[0, 1/3], [1/3, 2/3]$ und $[2/3, 1]$ auf. Dann ist x_1 in einem oder zwei dieser Intervalle nicht enthalten. Wir wählen ein solches Intervall und nennen den Anfangspukt a_1 und den Endpunkt b_1. Nun teilen wir entsprechend $I_1 := [a_1, b_1]$ 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. Induktiv erhalten wir 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 $b_n - a_n = 1/3^n \to 0$ $(n \to \infty)$. Nach dem Intervallschachtelungsprinzip 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 $x_k \notin \bigcap\limits_{n\in\mathbb{N}} I_n$. Also ist $x_k \ne x$ für alle $k \in \mathbb{N}$. Widerspruch! □

Bemerkung B.8 Ist X überabzählbar und ist A abzählbar, so ist auch $X \setminus A$ überabzählbar (denn sonst wäre nach B. B.3 und S. B.4 auch $X = (X \cap A) \cup (X \setminus A)$ abzählbar). Insbesondere ist also nach S. B.5 und S. B.7 für jedes nicht einpunktige Intervall I die Menge $I \setminus \mathbb{Q}$ überabzählbar.

C Normierte Räume

Wir betrachten eine Klasse von Räumen, die eine lineare Struktur und eine metrische Struktur besitzen.

Bemerkung und Definition C.1 Es seien K ein Körper und $V \neq \emptyset$ eine Menge. Weiter seien $+$ eine Verknüpfung auf V und $s : K \times V \to V$. Dann heißt $V = (V, +, s)$ ein **K-Vektorraum** (oder **K-linearer Raum**), falls gilt

(V1) $(V, +)$ ist eine abelsche Gruppe.

(V2) Für alle $\lambda, \mu \in K$, $v \in V$ ist $s(\lambda, s(\mu, v)) = s(\lambda\mu, v)$.

(V3) Für alle $v \in V$ ist $s(1, v) = v$.

(V4) (Distributivgesetze) Für alle $\lambda, \mu \in K$, $u, v \in V$ ist

$$s(\lambda, u + v) = s(\lambda, u) + s(\lambda, v) \quad \text{und} \quad s(\lambda + \mu, v) = s(\lambda, v) + s(\mu, v).$$

Die Elemente von V heißen dabei **Vektoren**, die Elemente aus K **Skalare** und s **Sklarmultiplikation**. Man schreibt auch wieder kurz λv statt $s(\lambda, v)$.
Ist V ein K-Vektorraum und ist $M \subset V$, so heißt

$$\operatorname{span}(M) := \Big\{ \sum_{v \in E} \lambda_v v : E \subset M \text{ endlich}, (\lambda_v)_{v \in E} \in \mathbb{K}^E \Big\}$$

(linearer) Spann von M. Außerdem heißt M ein **Untervektorraum** oder **(linearer) Teilraum**, falls $\operatorname{span}(M) = M$ gilt. Man kann zeigen ([Ü]): $\emptyset \neq U \subset V$ ist genau dann Teilraum, wenn für alle $u, v \in U$, $\lambda \in K$ auch $u + v \in U$ und $\lambda u \in U$ gilt. Außerdem ist dann $(U, +, s|_{K \times U})$ ein Vektorraum.

Definition C.2 Ist $(V, +, s)$ ein K-Vektorraum und ist $(K, +, \cdot)$ ein Ring, so heißt $(V, +, s, \cdot)$ eine **(unitäre) K-Algebra**, falls für $\lambda \in K$ und $u, v \in V$

$$\lambda(u \cdot v) = (\lambda u) \cdot v = u \cdot (\lambda v)$$

gilt.

Beispiel C.3 Es seien K ein Körper, $X \neq \emptyset$ und $(K^X, +, \cdot)$ der kommutative Ring aus B. 2.12. Indem man λ mit $\lambda 1_{K^X}$ identifiziert, wird $(K^X, +, s, \cdot)$ (mit $s(\lambda, f) := \lambda \cdot f$) zu einer K-Algebra.

Definition C.4 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

(N1) (Definitheit) Es gilt $||0|| = 0$ und $||v|| > 0$ für alle $v \neq 0$.

(N2) (Homogenität) Für alle $v \in V$ und alle $\lambda \in \mathbb{K}$ ist $||\lambda v|| = |\lambda| \cdot ||v||$.

(N3) (Dreiecksungleichung) Für alle $u, v \in V$ gilt $||u + v|| \leq ||u|| + ||v||$.

Man nennt dann $(V, ||\cdot||)$ auch einen **normierten** Raum.

Beispiel C.5 Wir setzen

$$||x||_2 := \sqrt{\sum_{j=1}^{m} |x_j|^2} \qquad (x = (x_1, \ldots, x_m) \in \mathbb{K}^m).$$

Dann ist für $x, y \in \mathbb{K}^m \setminus \{0\}$ ist nach der Ungleichung zwischen geometrischem und arithmetischem Mittel

$$\sum_{j=1}^{m} \frac{|x_j|}{||x||_2} \cdot \frac{|y_j|}{||y||_2} \leq \frac{1}{2} \sum_{j=1}^{m} \Big(\frac{|x_j|^2}{||x||_2^2} + \frac{|y_j|^2}{||y||_2^2} \Big) = 1.$$

Also folgt (auch für $x = 0$ oder $y = 0$) die **Cauchy-Schwarzsche Ungleichung**

$$\sum_{j=1}^{m} |x_j y_j| \leq ||x||_2 ||y||_2.$$

Hieraus ergibt sich wiederum

$$||x + y||_2^2 \leq \sum_{j=1}^{m} |x_j| \cdot |x_j + y_j| + \sum_{j=1}^{m} |y_j| \cdot |x_j + y_j| \leq (||x||_2 + ||y||_2) \cdot ||x + y||_2,$$

also auch $||x + y||_2 \leq ||x||_2 + ||y||_2$. Damit ist (N3) für $||\cdot||_2$ erfüllt. Da – wie man sofort sieht – auch (N1) und (N2) erfüllt sind, ist $||\cdot||_2$ eine Norm auf $\mathbb{K}^m$. Im Fall $\mathbb{K} = \mathbb{R}$ nennt man $||x||_2$ die **euklidsche Länge** von x.
Wie man leicht sieht, sind durch

$$||x||_1 := \sum_{j=1}^{m} |x_j| \qquad (x = (x_1, \ldots, x_m) \in \mathbb{K}^m)$$

und

$$||x||_\infty := \max\{|x_j| : j = 1, \ldots, m\} \qquad (x = (x_1, \ldots, x_m) \in \mathbb{K}^m)$$

weitere Normen auf $\mathbb{K}^m$ gegeben. Aus

$$|x_k|^2 \leq \sum_{j=1}^{m} |x_j|^2 \leq \sum_{j,\ell=1}^{m} |x_j x_\ell| = \Big(\sum_{j=1}^{m} |x_j|\Big)^2$$

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

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

Falls nicht anders angegeben, soll $\mathbb{K}^m$ stets mit der Norm $|| \cdot ||_2$ versehen sein. Wir schreiben auch kurz $| \cdot | := || \cdot ||_2$.

D Fundamentalsatz der Algebra

In B./D. 7.25 haben 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 D.1 Sind $X \subset \mathbb{C}$ unbeschränkt und $f : X \to \mathbb{K}$, so schreiben wir

$$f(x) \to c \qquad (|x| \to \infty)$$

falls zu jedem $\varepsilon > 0$ ein $R = R_\varepsilon > 0$ so existiert, dass $|f(x) - c| < \varepsilon$ für alle $x \in X$ mit $|x| > R$. Dies ist genau dann der Fall, wenn $f(1/u) \to c$ $(u \to 0)$ gilt.
Ist $P : \mathbb{C} \to \mathbb{C}$ ein nichtkonstantes Polynom, so folgt aus B. 5.21

$$1/P(z) \to 0 \quad (|z| \to \infty).$$

Satz D.2 *(Fundamentalsatz der Algebra)*
Jedes nichtkonstante Polynom $P : \mathbb{C} \to \mathbb{C}$ hat P eine Nullstelle.

Beweis. Nach B./D. D.1 existiert ein $R > 0$ mit $|P(0)| < |P(z)|$ für alle $z \in \mathbb{C}$ mit $|z| > R$. Außerdem hat $|P|$ nach S. 8.33 ein Minimum auf der kompakten Menge $B_R(0)$, d.h. es existiert ein $z_0 \in B_R(0)$ mit

$$|P(z_0)| \leq |P(z)| \qquad (|z| \leq R)\,.$$

Dann ist auch $|P(z_0)| = \min\limits_{\mathbb{C}} |P|$. Angenommen, P hat keine Nullstelle, also $P(z_0) \neq 0$. Wir können ohne Einschränkung $z_0 = 0$ und $P(0) = 1$ annehmen (sonst betrachte man $P(z + z_0)/P(z_0)$). Dann existieren ein $m \in \mathbb{N}$, ein $c_m \neq 0$ und ein Polynom Q mit $Q(0) = 0$ und

$$P(z) = 1 + c_m z^m + z^m Q(z)\,.$$

Nach B./D. 7.25 existiert eine m-te Wurzel $\zeta \in \mathbb{S}$ aus $-|c_m|/c_m$. Wählt man $r > 0$ so klein, dass $|Q(r\zeta)| < |c_m|$ und $r^m|c_m| < 1$, so folgt

$$|P(r\zeta)| \leq \big|1 - r^m|c_m|\big| + r^m|Q(r\zeta))| = 1 - r^m|c_m| + r^m|Q(r\zeta)| < 1\,,$$

im Widerspruch zu $|P| \geq 1$. $\square$

Bemerkung D.3 Ist $d \in \mathbb{N}$ und ist P ein Polynom von Grad d mit Nullstelle $z_1 \in \mathbb{C}$, so existiert ein Polynom Q_1 vom Grad $d-1$ so, dass

$$P(z) = (z - z_1)Q_1(z) \quad (z \in \mathbb{C}).$$

Unter Verwendung von S. D.2 ergibt sich induktiv: Jedes Polynom $P : \mathbb{C} \to \mathbb{C}$ vom Grad d zerfällt in Linearfaktoren, d. h. es existieren $z_1, \dots, z_d \in \mathbb{C}$ und $c \in \mathbb{C}^*$ mit

$$P(z) = c \prod_{j=1}^{d} (z - z_j) \quad (z \in \mathbb{C}).$$

Index