Analysis I

Wintersemester 2015/2016

Ernst Kuwert

Mathematisches Institut
Universität Freiburg

Inhaltsverzeichnis

1	Körperaxiome und Anordnungsaxiome	1
2	Vollständige Induktion	5
3	Grenzwerte von Folgen	13
4	Vollständigkeit der reellen Zahlen	21
5	Teilmengen von $\mathbb{R}$ und von $\mathbb{R}^n$	31
6	Reihen	45
7	Stetigkeit	53
8	Zwischenwertsatz und monotone Funktionen	61
9	Die Ableitung	63
10	Mittelwertsatz	69
11	Die reelle Exponentialfunktion	77
12	Die trigonometrischen Funktionen	83
13	Konvergenz von Funktionenfolgen	91
14	Das Riemannsche Integral	101
15	Ableitung und Integral	109

1 Körperaxiome und Anordnungsaxiome

Den Ausgangspunkt dieser Vorlesung bilden die Axiome der reellen Zahlen. Danach sind die reellen Zahlen eine Menge, auf der folgende Strukturen gegeben sind:

eine Verknüpfung $+$*, die je zwei* $a, b \in \mathbb{R}$ *ein* $a + b \in \mathbb{R}$ *zuordnet (*Addition*),*
eine Verknüpfung $\cdot$*, die je zwei* $a, b \in \mathbb{R}$ *ein* $ab \in \mathbb{R}$ *zuordnet (*Multiplikation*),*
eine Relation $a > b$*, die für* $a, b \in \mathbb{R}$ *zutrifft oder nicht (*Anordnung*).*

Es sollen dabei folgende drei Gruppen von Axiomen gelten:

Die Körperaxiome (K)
Die Anordnungsaxiome (A1, A2 und A3)
Das Vollständigkeitsaxiom (V).

Cantor und Dedekind haben Verfahren gefunden, wie die reellen Zahlen aus den rationalen Zahlen *konstruiert* werden können, das heißt es gibt tatsächlich eine Menge $\mathbb{R}$ mit den geforderten Eigenschaften, insbesondere sind die Axiome widerspruchsfrei. Zusammen mit den Strukturen ist eine solche Menge $\mathbb{R}$ im wesentlichen auch eindeutig bestimmt, wobei wir auf die Präzisierung dieser Aussage hier verzichten. Beim Umgang mit den reellen Zahlen werden niemals die speziellen Konstruktionen von Cantor und Dedekind benutzt, sondern die gesamte Analysis wird allein auf den Axiomen aufgebaut, und so wollen auch wir vorgehen. Die neugierigen Hörer verweisen wir aber auf Kapitel 2 im Buch *Zahlen.*[1].

Wir wollen in der Folge die einzelnen Axiome vorstellen und beginnen mit den Körperaxiomen, die die arithmetischen Eigenschaften von $\mathbb{R}$ regeln, das heißt die Addition und die Multiplikation. Sie lauten wie folgt:

	$+$	$\bullet$
Assoziativgesetz:	$(a + b) + c = a + (b + c)$	$(a \cdot b) \cdot c = a \cdot (b \cdot c)$
Kommutativgesetz:	$a + b = b + a$	$a \cdot b = b \cdot a$
Neutrales Element:	*Es gibt Zahlen* $0 \in \mathbb{R}$ *und* $1 \in \mathbb{R}$ *mit* $1 \neq 0$*, so dass für alle* $a \in \mathbb{R}$ *gilt:*	
	$a + 0 = a$	$a \cdot 1 = a$
Inverses Element:	*Zu jedem* $a \in \mathbb{R}$ *gibt es Lösungen* $x, y \in \mathbb{R}$ *der Gleichungen*	
	$a + x = 0$	$a \cdot y = 1$ *falls* $a \neq 0$
Distributivgesetz:	$a \cdot (b + c) = a \cdot b + a \cdot c.$	

Eine Menge $\mathbb{K}$ mit Verknüpfungen $+$ und $\cdot$, so dass die obigen Axiome erfüllt sind, heißt Körper (englisch: *field*). In der Algebra treten viele verschiedene Körper auf, ein extremes Beispiel ist $\mathbb{K} = \{0, 1\}$ mit $1 + 1 = 0$ und den sonst üblichen Rechenregeln. Die neutralen Elemente sind durch die Axiome eindeutig bestimmt, denn wären zum Beispiel $0_1 \in \mathbb{R}$ und $0_2 \in \mathbb{R}$ neutrale Elemente für die Addition, so folgt mit dem Kommutativgesetz

$$0_1 = 0_1 + 0_2 = 0_2 + 0_1 = 0_2.$$

[1] Hrsgb. H.-D. Ebbinghaus, 3. Auflage, Springer 1992

Das Argument für die Eindeutigkeit von $1 \in \mathbb{R}$ ist natürlich ganz analog. Weiter sind die inversen Elemente ebenfalls eindeutig bestimmt, denn für zwei Lösungen $x_{1,2}$ der Gleichung $a + x = 0$ folgt mit dem Assoziativgesetz und dem Kommutativgesetz

$$x_1 = x_1 + 0 = x_1 + (a + x_2) = (x_1 + a) + x_2 = (a + x_1) + x_2 = 0 + x_2 = x_2.$$

Wieder ist das Argument für die Multiplikation analog. Wir bezeichnen die Lösung x der Gleichung $a + x = 0$ mit $-a$ sowie die Lösung y der Gleichung $ay = 1$ mit $1/a$ oder a^{-1}, und vereinbaren die Notation

$$a - b = a + (-b) \qquad \frac{a}{b} = a \cdot \frac{1}{b}.$$

Aus den Körperaxiomen leiten sich unter anderem folgende Rechengesetze ab.

Satz 1.1 (Rechnen in $\mathbb{R}$) *Für reelle Zahlen a, b gelten folgende Aussagen:*

$$\begin{cases} -(-a) = a & -(a+b) = (-a) + (-b), \\ (a^{-1})^{-1} = a & (ab)^{-1} = a^{-1}b^{-1} \text{ für } b \neq 0, \\ a \cdot 0 = 0 & a(-b) = -(ab), \\ (-a)(-b) = ab & a(b-c) = ab - ac. \end{cases} \tag{1.1}$$

$$ab = 0 \quad \Rightarrow \quad a = 0 \text{ oder } b = 0 \qquad \text{(Nullteilerfreiheit)}. \tag{1.2}$$

Beweis: Die erste Zeile von (1.1) folgt mit der Eindeutigkeit der Inversen aus

$$(-a) + a = a + (-a) = 0 \qquad (a + b) + ((-a) + (-b)) = (a + (-a)) + (b + (-b)) = 0 + 0 = 0.$$

Der Beweis der zweiten Zeile in (1.1) ist analog. Weiter rechnen wir

$$a \cdot 0 + a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 \quad \text{und} \quad a \cdot 0 + 0 = a \cdot 0.$$

Die Eindeutigkeit der Inversen liefert also $a \cdot 0 = 0$. Daraus ergibt sich

$$ab + a(-b) = a(b + (-b)) = a \cdot 0 = 0,$$

also $a(-b) = -ab$, und dann

$$(-a)(-b) = -((-a)b) = -(b(-a)) = -(-ba) = ba = ab.$$

Die letzte Aussage von (1.1) folgt mit

$$a(b - c) = a(b + (-c)) = ab + a(-c) = ab + (-ac) = ab - ac.$$

Ist in (1.2) $a \neq 0$, so folgt schließlich

$$0 = ab \cdot \frac{1}{a} = (a \cdot \frac{1}{a}) \cdot b = 1 \cdot b = b.$$

Damit sind alle Aussagen des Satzes gezeigt. □

Auch die folgenden Regeln der Bruchrechnung lassen sich aus den Körperaxiomen herleiten, dies sei jedoch den Lesern überlassen.

Satz 1.2 (Bruchrechnung) *Für $a, b, c, d \in \mathbb{R}$ mit $c, d \neq 0$ gilt:*

(1) $\dfrac{a}{c} + \dfrac{b}{d} = \dfrac{ad+bc}{cd}$,

(2) $\dfrac{a}{c} \cdot \dfrac{b}{d} = \dfrac{ab}{cd}$,

(3) $\dfrac{a/c}{b/d} = \dfrac{ad}{bc}$, *falls zusätzlich* $b \neq 0$.

Als nächstes formulieren wir zwei Anordnungsaxiome. Ein drittes, das Archimedische Axiom, werden wir erst später einführen, wenn es für den Grenzwertbegriff gebraucht wird.

(**A1**) Für jedes $a \in \mathbb{R}$ gilt genau eine der drei Aussagen $a > 0$, $a = 0$ oder $-a > 0$.

(**A2**) Aus $a, b > 0$ folgt $a + b > 0$ und $ab > 0$.

Statt $-a > 0$ schreiben wir auch $a < 0$, und statt $a - b > 0$ auch $a > b$. Hier einige Folgerungen aus den Anordnungsaxiomen (A1) und (A2).

Satz 1.3 (Rechnen mit Ungleichungen)

(1) *Für $a, b \in \mathbb{R}$ gilt genau eine der Relationen $a > b$, $a = b$ oder $a < b$.*

(2) *Aus $a > b$, $b > c$ folgt $a > c$ (*Transitivität*).*

(3) *Aus $a > b$ folgt*

$$\begin{cases} a + c > b + c \\ ac > bc, \ \text{wenn } c > 0 \\ ac < bc, \ \text{wenn } c < 0 \end{cases}$$

(4) *Aus $a > b$ und $c > d$ folgt*

$$\begin{cases} a + c > b + d, \\ ac > bd, \ \text{falls } b, d > 0 \end{cases}$$

(5) *Für $a \neq 0$ ist $a^2 > 0$.*

(6) *Aus $a > 0$ folgt $1/a > 0$.*

(7) *Aus $a > b$, $b > 0$ folgt $1/a < 1/b$.*

Beweis: (1) folgt direkt aus (A1) und der Definition von $a > b$. Für (2) rechnen wir

$$a - c = (a - b) + (b - c) > 0 \quad \text{nach (A2)}.$$

Ähnlich ergeben sich (3) und (4):

$$\begin{aligned} (a + c) - (b + c) &= a - b > 0, \\ ac - bc &= (a - b)c > 0 \quad \text{im Fall } c > 0 \text{ nach (A2)}, \\ bc - ac &= (a - b)(-c) > 0 \quad \text{im Fall } c < 0 \text{ nach (A2)}, \\ (a + c) - (b + d) &= (a - b) + (c - d) > 0 \quad \text{nach (A2)}, \\ ac - bd &= ac - bc + bc - bd \\ &= (a - b)c + b(c - d) > 0 \quad \text{nach (2) und (A2)}. \end{aligned}$$

Die Positivität von Quadraten folgt aus der Fallunterscheidung

$$a^2 = \begin{cases} a \cdot a > 0 & \text{im Fall } a > 0, \\ (-a) \cdot (-a) > 0 & \text{im Fall } -a > 0. \end{cases}$$

Dabei haben wir die Regel $(-a)(-b) = ab$ aus (1.1) benutzt. Nach (A1) ist $a^2 > 0$ für $a \neq 0$ bewiesen. Aussage (6) ergibt sich nun mit

$$\frac{1}{a} = \left(\frac{1}{a}\right)^2 \cdot a > 0 \text{ nach (5) und (A2).}$$

Zu guter Letzt haben wir für (7)

$$\frac{1}{b} - \frac{1}{a} = \frac{1}{ab}(a-b) > 0 \text{ mit (6) und (A2).}$$

□

Definition 1.1 (Betrag einer reellen Zahl) *Der Betrag von $a \in \mathbb{R}$ ist*

$$|a| = \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0. \end{cases} \tag{1.3}$$

Ein anschauliches Modell der reellen Zahlen ist die Zahlengerade. Ist $a < b$, so liegt b rechts von a im Abstand $|a-b|$. Insbesondere ist $|a|$ der Abstand zum Nullpunkt.

Satz 1.4 (Rechnen mit Beträgen) *Für $a, b \in \mathbb{R}$ gelten folgende Aussagen:*

(1) $|-a| = |a|$ *und* $a \leq |a|$.

(2) $|a| \geq 0$; *aus Gleichheit folgt* $a = 0$.

(3) $|ab| = |a| \cdot |b|$.

(4) $|a+b| \leq |a| + |b|$.

(5) $|a-b| \geq ||a| - |b||$.

Beweis: Aus Definition 1.1 folgt

$$|-a| = \begin{cases} -a & \text{falls } -a \geq 0 \\ -(-a) & \text{falls } -a \leq 0 \end{cases} = \begin{cases} -a & \text{falls } a \leq 0 \\ a & \text{falls } a \geq 0 \end{cases} = |a|.$$

Weiter folgt (2) aus

$$|a| - a = \begin{cases} 0 & \text{falls } a \geq 0, \\ -a-a \geq 0 & \text{falls } a \leq 0. \end{cases}$$

In (3) bleiben die linke und rechte Seite gleich, wenn wir a durch $-a$ ersetzen, dasselbe gilt bezüglich b. Also können wir $a, b \geq 0$ annehmen, und erhalten $|ab| = ab = |a| \cdot |b|$ wie verlangt. Für (4) schätzen wir mit (1) wie folgt ab:

$$|a+b| = \pm(a+b) = \pm a + (\pm b) \leq |a| + |b|.$$

Schließlich gilt $|a| = |a-b+b| \leq |a-b| + |b|$ nach (4), also $|a-b| \geq |a| - |b|$. Durch Vertauschen von a und b folgt (5). □

Wie wir später sehen werden, spielen Ungleichungen in der Analysis eine große Rolle.

2 Vollständige Induktion

Wir unterbrechen jetzt die Diskussion der Axiome der reellen Zahlen, um das Beweisverfahren der vollständigen Induktion kennenzulernen. Wir setzen voraus, dass die natürlichen Zahlen $\mathbb{N}$ als Folge 1, $1+1=2$, $2+1=3$, $\ldots$ gegeben sind. Ausgehend von $1 \in \mathbb{N}$ wird also jede natürliche Zahl erreicht, indem die 1 endlich oft addiert wird. Darauf beruht das

Induktionsprinzip: Sei $M \subset \mathbb{N}$ eine Menge mit den beiden Eigenschaften

(1) $1 \in M$,

(2) $n \in M \quad \Rightarrow \quad n+1 \in M$.

Dann gilt schon $M = \mathbb{N}$.

Die Beschreibung der natürlichen Zahlen als Folge $1, 2, 3, \ldots$ ist keine wirkliche Definition, da die Pünktchen nicht präzisiert werden. Demzufolge können wir auch das Induktionsprinzip nicht rigoros begründen, sondern nehmen es schlicht als gegeben hin. Es ergibt sich als Konsequenz:

Satz 2.1 (Beweisverfahren der vollständigen Induktion) *Gegeben sei eine Folge von Aussagen $A(n)$ für $n \in \mathbb{N}$. Es möge gelten:*

(1) *$A(1)$ ist wahr.*

(2) *$A(n)$ ist wahr $\quad \Rightarrow \quad$ $A(n+1)$ ist wahr.*

Dann sind alle Aussagen $A(n)$ wahr.

Beweis: Wir betrachten die Menge $M = \{n \in \mathbb{N} : A(n) \text{ ist wahr}\}$. Nach Voraussetzung gilt $1 \in M$, und mit $n \in M$ ist auch $n+1 \in M$. Das Induktionsprinzip ergibt $M = \mathbb{N}$, das heißt alle Aussagen $A(n)$ sind wahr. $\square$

Bevor wir das Beweisverfahren an Beispielen üben, brauchen wir noch das Summen- und Produktzeichen. Diese definieren wir induktiv durch

$$\sum_{k=1}^{n} a_k = \Big(\sum_{k=1}^{n-1} a_k\Big) + a_n \text{ für } n \geq 2, \quad \sum_{k=1}^{1} a_k = a_1.$$

Allgemeiner durchläuft k die ganzen Zahlen von einer unteren Grenze $k = m$, hier $m = 1$, bis zu einer oberen Grenze $k = n$. Der Laufindex kann substituiert werden, wobei die Grenzen anzupassen sind; zum Beispiel liefert $k = j+1$

$$\sum_{k=m}^{n} a_k = \sum_{j=m-1}^{n-1} a_{j+1} = a_m + \ldots + a_n.$$

Es ist praktisch den Fall zuzulassen, wenn die untere Grenze größer als die obere Grenze ist, in diesem Fall setzen wir die Summe gleich Null. Das Produktzeichen ist ganz analog erklärt:

$$\prod_{k=1}^{n} a_k = \Big(\prod_{k=1}^{n-1} a_k\Big) \cdot a_n \text{ für } n \geq 2, \quad \prod_{k=1}^{1} a_k = a_1.$$

Das leere Produkt wird gleich Eins gesetzt.

Ein Induktionsbeweis funktioniert immer in zwei Schritten: erst wird die Aussage $A(n)$ für den Fall $n = 1$ verifiziert (Induktionsanfang). Statt bei $n = 1$ kann die Induktion auch bei einer anderen Zahl starten. Im zweiten Schritt wird *vorausgesetzt*, dass $A(n)$ für ein $n \in \mathbb{N}$ richtig ist (Induktionsannahme), und daraus $A(n+1)$ gefolgert (Induktionsschluss).

Beispiel 2.1 (arithmetische Summe) Wir zeigen die Summenformel

$$A(n): \quad 1 + 2 + \ldots + n = \sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$

Für $n = 1$ ist sowohl die linke als auch die rechte Seite gleich Eins, also gilt der Induktionsanfang. Jetzt berechnen wir unter Verwendung von $A(n)$

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

Damit ist $A(n)$ für alle $n \in \mathbb{N}$ bewiesen. Nach dem neunjährigen Gauß kommen wir natürlich auch ohne Induktion zum Ziel:

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

Der Vorteil dieses Arguments ist, dass wir die Formel nicht vorher raten müssen. Das Argument kann auch veranschaulicht werden, anhand einer Treppe mit Stufen $1, \ldots, n$ und der umgedrehten Treppe mit Stufen $n, \ldots, 1$. $\square$

Beispiel 2.2 (geometrische Summe) Sei $x \in \mathbb{R}$, $x \neq 1$. Dann gilt für alle $n \in \mathbb{N}_0$

$$1 + x + \ldots + x^n = \sum_{k=0}^{n} x^k = \frac{1 - x^{k+1}}{1 - x}.$$

Wir zeigen das wieder durch vollständige Induktion, wobei wir bei $n = 0$ beginnen:

$$\sum_{k=0}^{0} x^k = x^0 = 1 = \frac{1 - x^{0+1}}{1 - x}.$$

Jetzt gelte die Formel für ein $n \in \mathbb{N}$. Dann folgt

$$\sum_{k=0}^{n+1} x^k = \left(\sum_{k=0}^{n} x^k\right) + x^{n+1} \overset{A(n)}{=} \frac{1 - x^{n+1}}{1 - x} + x^{n+1} = \frac{1 - x^{(n+1)+1}}{1 - x},$$

womit die Behauptung gezeigt ist. Auch hier haben wir ein alternatives Argument, nämlich den sogenannten Teleskopsummentrick:

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

$\square$

Ebenfalls mit dem Beweisverfahren der vollständigen Induktion zeigen wir folgende nützliche Ungleichung.

Satz 2.2 (Bernoullische Ungleichung) *Für $x \in \mathbb{R}$, $x \geq -1$, und $n \in \mathbb{N}_0$ gilt*

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

Beweis: Wir führen Induktion über $n \in \mathbb{N}_0$. Für $n = 0$ gilt nach Definition $(1+x)^0 = 1 = 1 + 0 \cdot x$. Wegen $1 + x \geq 0$ folgt weiter

$$\begin{aligned} (1+x)^{n+1} &= (1+x) \cdot (1+x)^n \\ &\geq (1+x) \cdot (1+nx) \quad \text{(nach Induktionsannahme)} \\ &= 1 + (n+1)x + nx^2 \\ &\geq 1 + (n+1)x. \end{aligned}$$

□

Als nächstes führen wir den Begriff der Abbildung ein (gleichbedeutend: Funktion). Wir schreiben eine Abbildung von einer Menge A in eine Menge B in der Form $f : A \to B$, $a \mapsto f(a)$; dabei heißt A Definitionsbereich und B Zielbereich von f. Die Abbildung f ist

$$\begin{aligned} \text{injektiv} &\Leftrightarrow \text{aus } f(a) = f(a') \text{ folgt } a = a', \\ \text{surjektiv} &\Leftrightarrow \text{zu jedem } b \in B \text{ gibt es ein } a \in A \text{ mit } f(a) = b, \\ \text{bijektiv} &\Leftrightarrow f \text{ ist injektiv und surjektiv.} \end{aligned}$$

Ist f bijektiv, so ist die Umkehrabbildung f^{-1} definiert, die jedem $b \in B$ sein eindeutig bestimmtes Urbild zuordnet:

$$f^{-1} : B \to A,\ f^{-1}(b) = a \quad \text{wobei } f(a) = b.$$

Die Verkettung von zwei Abbildungen $f : A \to B$ und $g : B \to C$ ist gegeben durch

$$g \circ f : A \to C,\ (g \circ f)(a) = g(f(a)).$$

Ist zum Beispiel f bijektiv, so gilt $f^{-1} \circ f = \mathrm{id}_A$ und $f \circ f^{-1} = \mathrm{id}_B$.

Manchmal ist es nützlich, sich den Graph von Abbildungen anzuschauen[2]. Das kartesische Produkt von zwei Mengen A und B ist $A \times B = \{(a,b) : a \in A,\ b \in B\}$. Wir haben dann die Projektion auf die Komponenten

$$\pi_A : A \times B \to A,\ \pi_A((a,b)) = a \quad \text{bzw.} \quad \pi_B : A \times B \to B,\ \pi_B((a,b)) = b.$$

Der Graph von $f : A \to B$ ist $\Gamma_f = \{(a, f(a)) : a \in A\} \subset A \times B$.

Satz 2.3 (Der Graph) *Eine Teilmenge $\Gamma \subset A \times B$ ist genau dann Graph einer Abbildung $f : A \to B$, wenn die Projektion $\pi_A|_\Gamma : \Gamma \to A$ bijektiv ist. Die Abbildung f lautet dann*

$$f = \pi_B \circ \pi_A|_\Gamma^{-1} : A \to B.$$

[2] das kann beim zweiten Lesen erfolgen

Beweis: Sei $\Gamma = \Gamma_f$ mit $f : A \to B$. Zu jedem $a \in A$ gibt es dann genau ein $b \in B$ mit $(a, b) \in \Gamma$, nämlich $b = f(a)$. Somit ist $\pi_A|_\Gamma$ bijektiv. Da $(a, f(a)) \in \Gamma$, folgt

$$f(a) = \pi_B\big((a, f(a)\big) = \pi_B\big(\pi_A|_\Gamma^{-1}(a)\big).$$

Sei umgekehrt $\pi_A|_\Gamma : \Gamma \to A$ bijektiv. Zu jedem $a \in A$ gibt es dann genau ein $b \in B$ mit $(a, b) \in \Gamma$, dieses hat die Darstellung

$$b = \pi_B\big((a, b)\big) = \pi_B\big(\pi_A|_\Gamma^{-1}(a)\big).$$

Somit ist $\Gamma = \Gamma_f$ mit $f = \pi_B \circ \pi_A|_\Gamma^{-1}$. □

Zum Beispiel folgt, indem wir $b = f(a)$ substituieren:

$$\Gamma_{f^{-1}} = \{(b, f^{-1}(b)) : b \in B\} = \{(f(a), a) : a \in A\} = \mathrm{S}(\Gamma_f),$$

mit $\mathrm{S} : A \times B \to B \times A$, $\mathrm{S}((a, b)) = (b, a)$.

Unser nächstes Ziel ist das Zählen von Elementen von gewissen Mengen, wozu wir den Abbildungsbegriff benutzen wollen.

Satz 2.4 (Schubfachprinzip) *Ist $f : \{1, \ldots, m\} \to \{1, \ldots, n\}$ injektiv, so folgt $m \leq n$.*

Beweis: Wir fassen die Behauptung als Aussage $A(n)$ auf, die jeweils für alle $m \in \mathbb{N}$ zu zeigen ist, und führen einen Induktionsbeweis. Für $n = 1$ haben wir eine injektive Abbildung $f : \{1, \ldots, m\} \to \{1\}$ und es folgt sofort $m = 1$, also der Induktionsanfang. Sei nun eine injektive Abbildung $f : \{1, \ldots, m\} \to \{1, \ldots, n + 1\}$ gegeben. Zu zeigen ist $m \leq n + 1$, was für $m = 1$ offensichtlich ist. Für $m \geq 2$ konstruieren wir eine injektive Abbildung

$$\tilde{f} : \{1, \ldots, m - 1\} \to \{1, \ldots, n\},\ k \mapsto \tilde{f}(k).$$

Mit der Induktionsannahme folgt dann $m - 1 \leq n$ beziehungsweise $m \leq n + 1$ wie gewünscht. Im Fall $f(k) \in \{1, \ldots, n\}$ für alle $k = 1, \ldots, m - 1$ können wir einfach $\tilde{f}(k) = f(k)$ setzen. Andernfalls gibt es genau ein $i \in \{1, \ldots, m - 1\}$ mit $f(i) = n + 1$. Da f nach Voraussetzung injektiv ist, folgt $f(m) \neq n + 1$, das heißt $f(m) \in \{1, \ldots, n\}$, und wir können setzen

$$\tilde{f}(k) = \begin{cases} f(k) & \text{für } k = 1, \ldots, m - 1,\ k \neq i, \\ f(m) & \text{für } k = i. \end{cases}$$

Es ist leicht zu sehen (*nachprüfen!*), dass in jedem der Fälle $\tilde{f}$ injektiv ist. □

Definition 2.1 (Zahl der Elemente) *Eine Menge M heißt endlich, wenn für ein $n \in \mathbb{N}$ eine Bijektion $f : \{1, \ldots, n\} \to M$ existiert, oder wenn M die leere Menge ist. Die Anzahl der Elemente ist dann n (Symbol: $\#M = n$), bzw. per Definition $\#\emptyset = 0$. Eine Menge heißt unendlich, wenn sie nicht endlich ist.*

Wesentlicher Punkt bei der Definition der Anzahl ist, dass die Zahl n mit einer Bijektion $f : \{1, \ldots, n\} \to M$ eindeutig bestimmt ist. Denn ist $\tilde{f} : \{1, \ldots, m\} \to M$ ebenfalls bijektiv, so haben wir die bijektiven, insbesondere injektiven, Abbildungen

$$\tilde{f}^{-1} \circ f : \{1, \ldots, n\} \to \{1, \ldots, m\} \quad \text{sowie} \quad f^{-1} \circ \tilde{f} : \{1, \ldots, m\} \to \{1, \ldots, n\}.$$

Aus dem Schubfachprinzip, Satz 2.4, folgt dann $n \leq m$ und $m \leq n$, also $m = n$. Der Satz garantiert also, dass die scheinbare Uneindeutigkeit in der Definition der Anzahl nicht vorhanden ist. Die Mathematiker haben dafür die schöne Formulierung, die Anzahl sei *wohldefiniert*.

Das Produkt $n! = \prod_{k=1}^{n} k = 1 \cdot 2 \cdot \ldots \cdot n$ wird als *n-Fakultät* bezeichnet; dabei ist per Definition $0! = 1$ (in Konsistenz mit unserer Vereinbarung zum leeren Produkt). Der folgende Satz beantwortet die Frage nach der Anzahl der möglichen Anordnungen (oder Umordnungen oder Permutationen) von n Dingen.

Satz 2.5 (Zahl der Permutationen) *Für $n \in \mathbb{N}$ sei S_n die Menge der bijektiven Abbildungen $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. Dann gilt $\#S_n = n!$.*

Beweis: Wir zeigen die Behauptung durch Induktion, wobei der Induktionsanfang $n = 1$ offensichtlich ist. Es ist praktisch, jedes $\sigma \in S_n$ mit dem n-Tupel $(\sigma_1, \ldots, \sigma_n)$ zu identifizieren, wobei $\sigma_i = \sigma(i)$. Die Menge S_{n+1} ist disjunkte Vereinigung der Teilmengen

$$S_{n+1,k} = \{\tau \in S_{n+1} : \tau_k = n+1\} \quad \text{für } k = 1, \ldots, n+1.$$

Beispielsweise ist in aufzählender Form

$$S_{4,2} = \{(1,4,2,3), (2,4,3,1), (3,4,1,2), (1,4,2,3), (2,4,1,3), (3,4,1,2)\}.$$

Jedem $\sigma = (\sigma_1, \ldots, \sigma_n) \in S_n$ können wir die Permutation $(\sigma_1, \ldots, \sigma_{k-1}, n+1, \sigma_k, \ldots, \sigma_n)$ in $S_{n+1,k}$ zuordnen, und diese Abbildung ist bijektiv (*nachprüfen!*). Also folgt aus der Induktionsannahme $\#S_{n+1,k} = \#S_n = n!$ und

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

□

Definition 2.2 (Binomialkoeffizienten) *Für $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}$ setzen wir*

$$\binom{\alpha}{k} = \frac{\alpha \cdot (\alpha - 1) \cdot \ldots \cdot (\alpha - k + 1)}{1 \cdot 2 \cdot \ldots \cdot k}, \qquad \textit{sowie} \qquad \binom{\alpha}{0} = 1.$$

Lemma 2.1 (Rekursionsformel für Binomialkoeffizienten) *Für $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}$ erfüllen die Binomialkoeffizienten die Formel*

$$\binom{\alpha+1}{k} = \binom{\alpha}{k} + \binom{\alpha}{k-1}.$$

Beweis: Für $k = 1$ ist leicht zu sehen, dass die Formel richtig ist. Für $k \geq 2$ berechnen wir

$$\begin{aligned}
\binom{\alpha}{k} + \binom{\alpha}{k-1} &= \frac{\alpha \cdot (\alpha-1) \cdot \ldots \cdot (\alpha-k+1)}{1 \cdot 2 \cdot \ldots \cdot k} + \frac{\alpha \cdot (\alpha-1) \cdot \ldots \cdot (\alpha-k+2)}{1 \cdot 2 \cdot \ldots \cdot (k-1)} \\
&= \frac{\alpha \cdot (\alpha-1) \cdot \ldots \cdot (\alpha-k+2) \cdot (\alpha-k+1+k)}{1 \cdot 2 \cdot \ldots \cdot k} \\
&= \frac{(\alpha+1) \cdot \alpha \cdot \ldots \cdot ((\alpha+1)-k+1)}{1 \cdot 2 \cdot \ldots \cdot k} \\
&= \binom{\alpha+1}{k}.
\end{aligned}$$

$\square$

Im Fall $\alpha = n \in \mathbb{N}_0$ erlaubt Lemma 2.1 die rekursive Berechnung der Binomialkoeffizienten $\binom{n}{k}$ nach dem Dreiecksschema von Blaise Pascal (1623-1662).

n=0							1						
n=1						1		1					
n=2					1		2		1				
n=3				1		3		3		1			
n=4			1		4		6		4		1		
n=5		1		5		10		10		5		1	
n=6	1		6		15		20		15		6		1

Ebenfalls für $\alpha = n \in \mathbb{N}_0$ folgt durch Erweitern der Binomialkoeffizienten mit $(n-k)!$ die alternative Darstellung

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} \text{ für } n \in \mathbb{N}_0,\ k \in \{0, 1, \ldots, n\}, \tag{2.1}$$

und daraus weiter die am Diagramm ersichtliche Symmetrieeigenschaft

$$\binom{n}{k} = \binom{n}{n-k} \text{ für } n \in \mathbb{N}_0,\ k \in \{0, 1, \ldots, n\}. \tag{2.2}$$

Satz 2.6 (Zahl der Kombinationen) *Sei $n \in \mathbb{N}_0$ und $k \in \{0, 1, \ldots, n\}$. Dann ist die Anzahl der k-elementigen Teilmengen von $\{1, \ldots, n\}$ gleich $\binom{n}{k}$.*

Beweis: Die Behauptung gilt für $k = 0$ und beliebiges n, denn die leere Menge ist die einzige null-elementige Teilmenge von $\{1, \ldots, n\}$, und nach Definition ist $\binom{n}{0} = 1$. Insbesondere gilt die Behauptung für $n = 0$. Wir führen nun Induktion über n, wobei wir die Behauptung jeweils für alle $k \in \{0, 1, \ldots, n\}$ zeigen. Im Induktionsschluss müssen wir die Anzahl der k-elementigen Teilmengen von $\{1, \ldots, n+1\}$ bestimmen, wobei wir $k \geq 1$ annehmen können. Diese Teilmengen zerfallen in zwei disjunkte Klassen:

Klasse 1: Die Menge enthält die Nummer $n + 1$ nicht.

Klasse 2: Die Menge enthält die Nummer $n + 1$.

Klasse 1 besteht genau aus den k-elementigen Teilmengen von $\{1, \ldots, n\}$, Klasse 2 ergibt sich durch Hinzufügen des Elements $n + 1$ zu jeder der $(k-1)$-elementigen Teilmengen von $\{1, \ldots, n\}$. Insgesamt ist die Zahl der k-elementigen Teilmengen von $\{1, \ldots, n+1\}$ nach Induktionsannahme also gleich

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

nach Lemma 2.1, womit der Satz bewiesen ist. $\square$

Satz 2.7 (Binomische Formel) *Für $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$ gilt*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}. \tag{2.3}$$

Das Produkt $(a+b)^n$ besteht aus n Faktoren. Beim Ausmultiplizieren ergeben sich Terme der Form $a^k b^{n-k}$ mit $k = 0, 1, \ldots, n$. Um die Potenz a^k zu erhalten, müssen wir in k Klammern den Faktor a wählen, in den übrigen Klammern b. Die Zahl dieser Terme ist also gleich der Anzahl der Möglichkeiten, aus den insgesamt n Klammern k Stück für a zu reservieren, also gleich dem Binomialkoeffizienten $\binom{n}{k}$. Diese Begründung ist sehr schlüssig, allerdings haben wir das Ausmultiplizieren nicht wirklich durchgeführt. Wir geben zur Sicherheit noch einen Induktionsbeweis.

Beweis: Wir zeigen eine Darstellung $(a+b)^n = \sum_{k=0}^{n} c(n,k)\, a^k b^{n-k}$, wobei die Koeffizienten $c(n,k)$ im Anschluss bestimmt werden. Für $n = 1$ gilt die Darstellung jedenfalls mit $c(1,0) = c(1,1) = 1$. Per Induktion folgt

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&= \sum_{k=0}^{n} c(n,k)\, a^{k+1} b^{n-k} + \sum_{k=0}^{n} c(n,k)\, a^k b^{n-k+1} \\
&= c(n,n)\, a^{n+1} + \sum_{k=1}^{n} c(n,k-1)\, a^k b^{n-(k-1)} + \sum_{k=1}^{n} c(n,k)\, a^k b^{n+1-k} + c(n,0)\, b^{n+1} \\
&= c(n,n)\, a^{n+1} + \sum_{k=1}^{n} \big(c(n,k) + c(n,k-1)\big)\, a^k b^{n+1-k} + c(n,0)\, b^{n+1} \\
&= \sum_{k=0}^{n+1} c(n+1,k)\, a^k b^{n+1-k},
\end{aligned}$$

wobei wir die Koeffizienten $c(n+1,k)$ wie folgt wählen:

$$c(n+1,k) = \begin{cases} c(n,k) + c(n,k-1) & \text{für } k = 1, \ldots, n \\ c(n,0) & \text{für } k = 0, \\ c(n,n) & \text{für } k = n+1. \end{cases}$$

Aber es gilt auch $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$ laut Rekursionsformel, Lemma 2.1, sowie $\binom{n}{0} = \binom{n}{n} = 1$ für alle $n \in \mathbb{N}$. Durch Induktion folgt also $c(n,k) = \binom{n}{k}$, und der Satz ist bewiesen. $\square$

3 Grenzwerte von Folgen

Die folgende Umformulierung des Induktionsprinzips erscheint absolut selbstverständlich; wir erinnern aber daran, dass wir keine strenge Definition der natürlichen Zahlen gegeben haben.

Satz 3.1 (Prinzip der kleinsten natürlichen Zahl) *Jede nichtleere Menge $M \subset \mathbb{N}$ besitzt ein kleinstes Element.*

Beweis: Wähle ein $N \in M$ und zerlege $M = M_0 \cup M_1$ mit

$$M_0 = \{n \in M : n \leq N\} \quad \text{und} \quad M_1 = \{n \in M : n > N\}.$$

M_0 ist Teilmenge von $\{1, \ldots, N\}$; sei $n_0 \in M_0$ die kleinste der endlich vielen Zahlen in M_0. Dann ist n_0 auch kleinstes Element in M, denn für $n \in M$ gilt

$$\begin{aligned} n \in M_0 \quad &\Rightarrow \quad n \geq n_0 \quad \text{nach Wahl von } n_0, \\ n \in M_1 \quad &\Rightarrow \quad n > N \geq n_0. \end{aligned}$$

□

Als Anwendung zeigen wir nun, dass die rationalen Zahlen auf der Zahlengeraden dicht verteilt sind. Dafür brauchen wir nun das noch fehlende, dritte Anordnungsaxiom.

A3 Zu jedem $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$, existiert ein $n \in \mathbb{N}$ mit $1/n < \varepsilon$ (*Archimedisches Axiom*).

Es gibt dann auch zu jedem $K \in \mathbb{R}$, $K > 0$, ein $n \in \mathbb{N}$ mit $n > K$, wie sich mit der Wahl $\varepsilon = 1/K > 0$ in A3 sofort ergibt.

Satz 3.2 ($\mathbb{Q}$ ist dicht in $\mathbb{R}$) *Zu je zwei reellen Zahlen $a, b \in \mathbb{R}$ mit $a < b$ gibt es eine rationale Zahl $q \in \mathbb{Q}$ mit $a < q < b$.*

Beweis: Wir können $b > 0$ annehmen, denn sonst gehen wir zu $a' = -b$, $b' = -a$ über. Ausserdem ist o.B.d.A. $a \geq 0$, da wir andernfalls einfach $q = 0$ setzen. Nach A3 existiert ein $n \in \mathbb{N}$ mit $1/n < b - a$. Die Zahlen k/n mit $k \in \mathbb{Z}$ sind rational und bilden ein Gitter, dessen Gitterlänge kleiner als der Abstand zwischen a und b ist. Daher sollte zwischen a und b ein Gitterpunkt liegen. Betrachte dazu die Menge

$$M = \{k \in \mathbb{N} : k/n > a\} \subset \mathbb{N}.$$

M ist nichtleer: falls $a = 0$ so ist $1 \in M$, andernfalls gibt es nach A3 ein $k \in \mathbb{N}$ mit $1/k < 1/na$ bzw. $k/n > a$. Sei $m \in M$ das kleinste Element nach Satz 3.1. $m \in M$ bedeutet $m/n > a$, und $m - 1 \notin M$ liefert $(m-1)/n \leq a$ (das gilt auch wenn $m = 1$). Also

$$\frac{m}{n} = \frac{m-1}{n} + \frac{1}{n} < a + (b - a) = b.$$

Die Zahl $q = m/n$ leistet somit das Verlangte. □

Definition 3.1 *Eine Folge reeller Zahlen ist eine Abbildung*

$$\mathbb{N} \longrightarrow \mathbb{R}, \quad n \longmapsto a_n.$$

Die Zahl a_n heißt das n-te Glied der Folge, die Folge insgesamt wird mit $(a_n)_{n\in\mathbb{N}}$ bzw. kurz mit (a_n) bezeichnet. Oft wird die Folge durch das Bildungsgesetz angegeben oder durch Aufzählen der ersten Folgenglieder definiert. Zum Beispiel ist die Folge der Quadratzahlen gegeben durch $a_n = n^2$, bzw. alternativ aufzählend $a_n = 1, 4, 9, 16, \ldots$.

Definition 3.2 (Konvergenz von Folgen) *Die Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert mit $n \to \infty$ gegen $a \in \mathbb{R}$, falls gilt:*

$$\text{Zu jedem } \varepsilon > 0 \text{ gibt es ein } K \in \mathbb{R}, \text{ so dass für alle } n > K \text{ gilt: } |a_n - a| < \varepsilon.$$

Die Zahl a heißt dann Grenzwert der Folge und wir schreiben kurz $\lim_{n\to\infty} a_n = a$ oder $a_n \to a$ für $n \to \infty$. Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt konvergent, wenn es ein $a \in \mathbb{R}$ gibt, das Grenzwert der Folge ist; andernfalls heißt die Folge divergent.

Mit den Quantoren $\forall$ (für alle), $\exists$ (existiert) und $\Rightarrow$ (daraus folgt) lässt sich die Definition der Konvergenz auch wie folgt fassen:

$$\forall \varepsilon > 0 \; \exists K \in \mathbb{R} : \quad \Big(n > K \quad \Rightarrow \quad |a_n - a| < \varepsilon \Big).$$

Beispiel 3.1 (Harmonische Folge) Die Folge $a_n = 1/n$ konvergiert gegen $a = 0$. Denn zu gegebenem $\varepsilon > 0$ wählen wir $K = 1/\varepsilon$, und es folgt für alle $n > K$

$$|a_n - a| = |1/n - 0| = 1/n < 1/K = \varepsilon.$$

□

In der Definition der Konvergenz gibt die Zahl $\varepsilon > 0$ gibt eine Genauigkeitsschranke vor, die die a_n erreichen sollen. Sei zum Beispiel $\varepsilon = 1$, es ist dann ein K_1 gesucht mit $|a_n - a| < 1$ für alle $n > K_1$. Entsprechend ist für $\varepsilon = 10^{-2}$ ein $K_{10^{-2}}$ gesucht mit $|a_n - a| < 10^{-2}$ für alle $n > K_{10^{-2}}$. Im Beispiel $a_n = \frac{1}{n}$ kann man $K_1 = 1$ wählen, während mindestens $K_{10^{-2}} = 100$ sein muss. Im allgemeinen gilt: je kleiner die geforderte Abweichung, umso später wird sie erreicht, umso größer muss demnach K gewählt werden. K hängt von $\varepsilon > 0$ ab, wir schreiben dafür manchmal $K = K_\varepsilon$. Eine Ausnahme bildet hier nur die konstante Folge.

Beispiel 3.2 (Konstante Folge) Ist $a_n = a$ für alle $n \in \mathbb{N}$, so folgt $\lim_{n\to\infty} a_n = a$. Denn für $\varepsilon > 0$ gilt $|a_n - a| = 0 < \varepsilon$ für alle $n > 0$, also können wir immer $K = 0$ wählen. □

Übrigens ist es egal, ob in der Definition des Grenzwerts statt $K \in \mathbb{R}$ die Bedingung $K \in \mathbb{N}$ verlangt wird, denn wir können statt $K \in \mathbb{R}$ ja immer die nächstgrößere natürliche Zahl nehmen. Überhaupt kann K immer vergrößert werden, und in der Regel besteht kein Interesse daran, dass kleinstmögliche $K_\varepsilon \in \mathbb{R}$ mit $|a_n - a| < \varepsilon$ für $n \geq K_\varepsilon$ zu finden. Dies ist natürlich anders, wenn ein Grenzwert numerisch berechnet werden soll, aber für den Nachweis der Konvergenz reicht es völlig, irgendeine Schranke zu finden, von der ab die Ungleichung gilt. Das vereinfacht oft die Bestimmung von K, weil wir die Differenz $|a_n - a|$ nur abschätzen müssen. Das sieht man deutlich an folgendem Beispiel.

Beispiel 3.3 (Geometrische Folge) Sei $q \in \mathbb{R}$ mit $|q| < 1$. Dann gilt $\lim_{n\to\infty} q^n = 0$. Um das zu zeigen, können wir $q \neq 0$ voraussetzen und haben dann $1/|q| > 1$, also gilt $1/|q| = 1 + x$ für ein $x > 0$. Es folgt mit der Bernoulli-Ungleichung, Satz 2.2,

$$|q^n - 0| = |q|^n = \frac{1}{(1+x)^n} \leq \frac{1}{1+nx} \leq \frac{1}{nx} < \varepsilon$$

für alle $n > 1/(\varepsilon x)$. Wir können also $K = 1/(\varepsilon x)$ wählen. $\square$

Beispiel 3.4 Die Folge $a_n = (-1)^n$, also $a_n = -1, 1, -1, \ldots$ ist nicht konvergent. Denn angenommen es wäre $\lim_{n\to\infty} a_n = a$ für ein $a \in \mathbb{R}$. Zu $\varepsilon = 1$ gibt es dann ein $K \in \mathbb{R}$ mit $|a_n - a| < 1$ für $n > K$, also gilt für $n > K$

$$2 = |a_n - a_{n+1}| = |a_n - a + a - a_{n+1}| \le |a_n - a| + |a - a_{n+1}| < 1 + 1 = 2,$$

ein Widerspruch. $\square$

Der Begriff des Grenzwerts wird etwas anschaulicher, indem wir folgende Teilmengen von $\mathbb{R}$ einführen.

Definition 3.3 (ε-Umgebung) *Die ε-Umgebung von $a \in \mathbb{R}$ ist die Menge*

$$U_\varepsilon(a) = \{x \in \mathbb{R} : |x - a| < \varepsilon\} = \{x \in \mathbb{R} : a - \varepsilon < x < a + \varepsilon\}.$$

Eine Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert genau dann gegen $a \in \mathbb{R}$, wenn die Folgenglieder ab einer gewissen Nummer in der ε-Umgebung von a liegen, egal wie klein $\varepsilon > 0$ gewählt ist.

Satz 3.3 (Eindeutigkeit des Grenzwerts) *Falls die Folge $(a_n)_{n\in\mathbb{N}}$ konvergent ist, so ist ihr Grenzwert eindeutig bestimmt.*

Beweis: Wir zeigen als erstes die Implikation

$$U_\varepsilon(a) \cap U_\varepsilon(a') \neq \emptyset \quad \Rightarrow \quad |a - a'| < 2\varepsilon. \tag{3.1}$$

Denn ist $x \in U_\varepsilon(a) \cap U_\varepsilon(a')$, so folgt $|a - a'| = |a - x + x - a'| \le |x - a| + |x - a'| < 2\varepsilon$.

Seien nun $a, a' \in \mathbb{R}$ Grenzwerte der Folge $(a_n)_{n\in\mathbb{N}}$. Zu jedem $\varepsilon > 0$ gibt es dann $K, K' \in \mathbb{R}$ mit $a_n \in U_\varepsilon(a)$ für $n > K$, sowie $a_n \in U_\varepsilon(a')$ für $n > K'$. Für $n > \max(K, K')$ folgt $a_n \in U_\varepsilon(a) \cap U_\varepsilon(a')$, also $|a - a'| < 2\varepsilon$ nach (3.1). Da $\varepsilon > 0$ beliebig war, folgt $a = a'$. $\square$

Wir haben hier benutzt: ist $a \ge 0$ und $a < \varepsilon$ für jedes $\varepsilon > 0$, so ist $a = 0$. Denn wäre $a > 0$, so könnten wir $\varepsilon = a$ wählen und hätten $a < a$, Widerspruch. Unser nächstes Ziel ist es, einige Rechenregeln für Grenzwerte zu erarbeiten. Wir beginnen mit der

Definition 3.4 (Beschränktheit von Folgen) *Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt*

beschränkt	*wenn es ein $C \ge 0$ gibt mit $|a_n| \le C$ für alle $n \in \mathbb{N}$,*
nach oben beschränkt	*wenn es ein $C \in \mathbb{R}$ gibt mit $a_n \le C$ für alle $n \in \mathbb{N}$,*
nach unten beschränkt	*wenn es ein $C \in \mathbb{R}$ gibt mit $a_n \ge C$ für alle $n \in \mathbb{N}$.*

Überlegen Sie: *beschränkt* ist äquivalent zu *nach oben und unten beschränkt.*

Beispiel 3.5 Die Folge $a_n = n$ ist nach unten beschränkt, denn es ist zum Beispiel $a_n \ge 0$ für alle n. Sie ist aber nicht nach oben beschränkt: angenommen, es gibt ein $C \in \mathbb{R}$ mit $a_n \le C$ für alle $n \in \mathbb{N}$. Dann ist $C \ge a_1 = 1 > 0$, also auch $1/C > 0$, und nach Archimedes gibt es ein $n \in \mathbb{N}$ mit $1/n < 1/C$ bzw. $a_n = n > C$, ein Widerspruch.

Satz 3.4 (konvergent $\Rightarrow$ beschränkt) *Jede konvergente Folge ist beschänkt.*

Beweis: Sei $\lim_{n\to\infty} a_n = a$. Wähle zu $\varepsilon = 1$ ein $N \in \mathbb{R}$ mit $|a_n - a| < 1$ für $n > N$. Wir können $N \in \mathbb{N}$ annehmen, andernfalls ersetzen wir N durch die nächstgrößere natürliche Zahl. Es gilt dann

$$\begin{aligned} n > N \quad &\Rightarrow \quad |a_n| = |a_n - a + a| \leq |a_n - a| + |a| < 1 + |a| \\ n \leq N \quad &\Rightarrow \quad |a_n| \leq \max(|a_1|, \ldots, |a_N|). \end{aligned}$$

Wir haben also $|a_n| \leq K$ für alle n, wobei $K = \max(|a_1|, \ldots, |a_N|, 1 + |a|)$. □

Satz 3.5 (Rechenregeln für Grenzwerte) *Es gelte $a_n \to a, b_n \to b$ mit $n \to \infty$.*

a) *Für alle $\lambda, \mu \in \mathbb{R}$ ist $(\lambda a_n + \mu b_n)_{n\in\mathbb{N}}$ konvergent mit Grenzwert $\lim_{n\to\infty}(\lambda a_n + \mu b_n) = \lambda a + \mu b$.*

b) *Die Folge $(a_n \cdot b_n)_{n\in\mathbb{N}}$ ist konvergent mit Grenzwert $\lim_{n\to\infty}(a_n \cdot b_n) = a \cdot b$.*

c) *Falls $b \neq 0$, so gibt es ein $K_0 \in \mathbb{R}$ mit $b_n \neq 0$ für $n > K_0$ und die Folge $(a_n/b_n)_{n>N_0}$ ist konvergent mit Grenzwert $\lim_{n\to\infty} a_n/b_n = a/b$.*

Beweis: Wir zeigen erst a) im Fall $\lambda = \mu = 1$. Zu $\varepsilon > 0$ gibt es ein $K \in \mathbb{R}$ mit $|a_n - a| < \varepsilon/2$ und $|b_n - b| < \varepsilon/2$ für $n > K$. Es folgt für $n > K$

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

Um b) zu zeigen, schätzen wir erst für $n \in \mathbb{N}$ ab:

$$|a_n b_n - ab| = |a_n b_n - a_n b + a_n b - ab| \leq |a_n| \cdot |b_n - b| + |a_n - a| \cdot |b|.$$

Nach Satz 3.4 gibt es ein $C > 0$ mit $|a_n| \leq C$ für alle $n \in \mathbb{N}$, und außerdem mit $|b| \leq C$. Also

$$|a_n b_n - ab| \leq C(|a_n - a| + |b_n - b|) \qquad \text{für alle } n \in \mathbb{N}.$$

Zu $\varepsilon > 0$ wähle $K \in \mathbb{R}$ mit $|a_n - a|$, $|b_n - b| < \varepsilon/(2C)$ für $n > K$. Es folgt

$$|a_n b_n - ab| < C\left(\frac{\varepsilon}{2C} + \frac{\varepsilon}{2C}\right) = \varepsilon \qquad \text{für } n > K.$$

In c) können wir $a_n = 1$ annehmen, der allgemeine Fall folgt dann aus b) mit

$$\frac{a_n}{b_n} = a_n \cdot \frac{1}{b_n}.$$

Als erstes müssen wir das Problem behandeln, dass der Nenner Null wird. Es gilt

$$|b_n| = |b - (b - b_n)| \geq |b| - |b_n - b|.$$

Wähle zu $|b|/2 > 0$ ein $K_0 \in \mathbb{R}$ mit $|b_n - b| < |b|/2$, also

$$|b_n| \geq |b| - \frac{|b|}{2} = \frac{|b|}{2} > 0 \qquad \text{für } n > K_0.$$

Jetzt schätze ab, für alle $n > K_0$,

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b_n - b|}{|b|\,|b_n|} \leq C|b_n - b| \qquad \text{mit } C = \frac{2}{|b|^2}.$$

Zu $\varepsilon > 0$ wähle $K \in \mathbb{R}$ mit $|b_n - b| < \varepsilon/C$, es folgt

$$|\frac{1}{b_n} - \frac{1}{b}| < \varepsilon \quad \text{für alle } n > \max(K_0, K).$$

$\square$

Beachten Sie, dass in c) die Information $b_n \neq 0$ nicht gereicht hätte. In der Abschätzung steht $b_n - b$ im Zähler, aber b_n im Nenner. Beide gehen gegen Null, es wäre nicht klar wer das Rennen gewinnt. Wir brauchen die untere Schranke $|b_n| \geq |b|/2 > 0$.

Die folgenden Bemerkungen können Sie zurückstellen, bis in der Linearen Algebra der Begriff des Vektorraums behandelt wurde. Allerdings ist sicher die Vektoraddition und Skalarmultiplikation im $\mathbb{R}^3$ bekannt, also

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \oplus \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix}, \quad \lambda \odot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \lambda a_3 \end{pmatrix}.$$

Für reelle Folgen hat man analog eine Addition und eine Skalarmultiplikation, nämlich

$$(a_n)_{n\in\mathbb{N}} \oplus (b_n)_{n\in\mathbb{N}} = (a_n + b_n)_{n\in\mathbb{N}}, \quad \lambda \odot (a_n)_{n\in\mathbb{N}} = (\lambda a_n)_{n\in\mathbb{N}}.$$

Salopp gesagt könnn wir die Folgen als Vektoren mit unendlich vielen Komponenten $a_1, a_2, \ldots$ ansehen; der Raum der Folgen ist nicht endlichdimensional. Die Unterräume des $\mathbb{R}^3$ sind Geraden und Ebenen durch den Nullpunkt, sowie der $\mathbb{R}^3$ selbst und der Nullvektor. Allgemein sind Unterräume gekennzeichnet durch die Eigenschaften

$$a, b \in U \quad \Rightarrow \quad a \oplus b,\, \lambda \odot a \in U.$$

Im Raum der reellen Folgen haben wir u.a. folgende Unterräume:

$$\{\text{Reelle Folgen}\} \supset \{\text{Beschränkte Folgen}\} \supset \{\text{Konvergente Folgen}\} \supset \{\text{Nullfolgen}\}.$$

Dabei sind Nullfolgen definiert durch $\lim_{n\to\infty} a_n = 0$. Man kann sich überlegen, dass alle Inklusionen echt sind und keiner der genannten Räume endlichdimensional ist. Damit beenden wir den Einschub.

Hier zwei Anwendungen der Rechenregeln für Grenzwerte.

Beispiel 3.6 (Grenzwerte rationaler Funktionen) Seien $p, q : \mathbb{R} \to \mathbb{R}$ Polynome vom Grad $m, n \in \mathbb{N}_0$, das heißt für alle $x \in \mathbb{R}$ gilt

$$p(x) = a_m x^m + a_{m-1} x^{m-1} + \ldots + a_0 \quad \text{und} \quad q(x) = b_n x^n + b_{n-1} x^{n-1} + \ldots + b_0,$$

wobei $a_i, b_j \in \mathbb{R}$ mit $a_m, b_n \neq 0$. Wir bestimmen im Fall $m \leq n$ das Verhalten von $p(k)/q(k)$ für $k \to \infty$, und zwar liefert Anwendung der Konvergenzregeln in Satz 3.5

$$\frac{p(k)}{q(k)} = k^{m-n} \frac{a_m + a_{m-1}k^{-1} + \ldots + a_0 k^{-m}}{b_n + b_{n-1}k^{-1} + \ldots + b_0 k^{-n}} \to \begin{cases} a_m/b_n & \text{falls } m = n, \\ 0 & \text{falls } m < n. \end{cases}$$

Beispiel 3.7 (geometrische Reihe) Für $-1 < q < 1$ betrachten wir die Folge

$$a_n = 1 + q + \ldots + q^n = \sum_{k=0}^{n} q^k.$$

Dann ergibt sich aus Beispiel 2.2, Beispiel 3.3 und Satz 3.5

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

Wir schreiben hierfür auch $\sum_{k=0}^{\infty} q^k = 1/(1-q)$. Folgen, deren Folgenglieder Summen sind, heißen Reihen. Sie spielen eine große Rolle in der Analysis und werden noch ausführlicher untersucht.

Satz 3.6 (Grenzwerte und Ungleichungen) *Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergent, mit Grenzwerten $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$. Dann gelten folgende Aussagen:*

a) *Ist $a_n \leq b_n$ für alle n, so folgt $a \leq b$.*

b) *Gilt $c \leq a_n \leq d$ für alle n mit $c, d \in \mathbb{R}$, so folgt $c \leq a \leq d$.*

c) *Ist $a_n \leq c_n \leq b_n$ und gilt $a = b$, so konvergiert auch die Folge $(c_n)_{n\in\mathbb{N}}$ gegen $a = b$.*

BEWEIS: Nach Voraussetzung gibt es zu jedem $\varepsilon > 0$ ein $K \in \mathbb{R}$ mit $a_n > a - \varepsilon$ und $b_n < b + \varepsilon$ für alle $n > K$. Die Voraussetzung in a) liefert dann $a - \varepsilon < b + \varepsilon$ beziehungsweis $(a-b)/2 < \varepsilon$ für jedes $\varepsilon > 0$, also $a \leq b$. Aussage b) folgt unmittelbar aus a), indem wir c, d als konstante Folgen auffassen. Unter den Voraussetzungen in c) folgt für $n > K$ die Ungleichungskette

$$a - \varepsilon < a_n \leq c_n \leq b_n < b + \varepsilon = a + \varepsilon,$$

also $\lim_{n\to\infty} c_n = a$ nach Definition des Grenzwerts. $\square$

Achtung: aus $a_n < b_n$ folgt *nicht* $a < b$, sondern nur $a \leq b$. Die Striktheit von Ungleichungen geht beim Übergang zu Grenzwerten im allgemeinen verloren. Zum Beispiel gilt $1/n > 0$ für alle $n \in \mathbb{N}$, aber $\lim_{n\to\infty} 1/n = 0$.

Definition 3.5 (Uneigentliche Konvergenz) *Die Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert uneigentlich (oder divergiert bestimmt) gegen $+\infty$, falls gilt:*

Zu jedem $K > 0$ gibt es ein $N \in \mathbb{R}$, so dass $a_n > K$ für alle $n > N$.

Wir schreiben $\lim_{n\to\infty} a_n = +\infty$ oder $a_n \to +\infty$ mit $n \to \infty$. Uneigentliche Konvergenz gegen $-\infty$ ist analog definiert.

Beispiel 3.8 Für $q > 1$ gilt $\lim_{n\to\infty} q^n = +\infty$. Denn zu gegebenem $C > 0$ gibt es nach Beispiel 3.3 ein $K \in \mathbb{R}$ mit $(1/q)^n < 1/C$ für $n > K$, also $q^n > C$ für $n > K$. Insgesamt haben wir für das Verhalten der Folge q^n mit $n \to \infty$ folgende Tabelle:

$$\begin{array}{rcl}
q > 1 & \Rightarrow & \lim_{n\to\infty} q^n = +\infty, \\
q = 1 & \Rightarrow & \lim_{n\to\infty} q^n = 1, \\
-1 < q < 1 & \Rightarrow & \lim_{n\to\infty} q^n = 0, \\
q \leq -1 & \Rightarrow & (q^n) \text{ nicht konvergent.}
\end{array}$$

Der Fall $-1 < q < 1$ wurde in Beispiel 3.3 behandelt, und der Fall $q \leq -1$ folgt mit etwas Überlegung aus Beispiel 3.4.

Beispiel 3.9 Betrachte die rekursiv definierte Folge

$$a_0 = 0, \quad a_{n+1} = a_n^2 + c \quad (c \in \mathbb{R} \text{ Parameter}).$$

Wir behaupten, dass die Folge für $c > \frac{1}{4}$ gegen $+\infty$ konvergiert. Und zwar gilt

$$a_{n+1} - a_n = a_n^2 - a_n + c = (a_n - \frac{1}{2})^2 + c - \frac{1}{4}.$$

Also folgt induktiv $a_n \geq n(c - \frac{1}{4}) \to \infty$. Die Frage, für welche $c \in \mathbb{R}$ die Folge divergiert bzw. konvergiert, ist nicht einfach zu beantworten (Stichwort Mandelbrotmenge).

Zum Schluss dieses Kapitels führen wir noch folgende Bezeichnungen ein:

$(a, b) = \{x \in \mathbb{R} : a < x < b\}$	offenes Intervall
$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$	abgeschlossenes Intervall
$[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$	rechtsseitig offen, linksseitig abgeschlossen
$(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$	linksseitig offen, rechtsseitig abgeschlossen
$\lvert I\rvert = b - a$ für ein Intervall I	Intervalllänge

Hierbei sind $+\infty$ und $-\infty$ als offene Intervallgrenzen zugelassen, zum Beispiel ist $(-\infty, 1] = \{x \in \mathbb{R} : -\infty < x \leq 1\}$. Den ε-Umgebungen bei der Definition der Konvergenz entsprechen bei uneigentlicher Konvergenz gegen $+\infty$ die Intervalle $(C, +\infty)$: für gegebenes $C > 0$ müssen ab einem gewissen Index alle Folgenglieder in $(C, +\infty)$ liegen.

Satz 3.7 (Konvergenz von Kehrwerten) *Für eine Folge $(a_n)_{n\in\mathbb{N}}$ gilt:*

(1) *Aus $a_n \to +\infty$ (bzw. $a_n \to -\infty$) folgt $1/a_n \to 0$.*

(2) *Aus $a_n \to 0$ und $a_n > 0$ (bzw. $a_n < 0$) folgt $1/a_n \to +\infty$ (bzw. $1/a_n \to -\infty$).*

BEWEIS: Übungsaufgabe $\square$

4 Vollständigkeit der reellen Zahlen

Die bisher eingeführten Axiome (K) sowie (A1) bis (A3) gelten selbstverständlich auch für die rationalen Zahlen. Dennoch sind die rationalen Zahlen für die Analysis ungeeignet. Wir beginnen mit folgender Beobachtung der Pythagoräer.

Satz 4.1 (Irrationalität von $\sqrt{2}$) *Die Gleichung $x^2 = 2$ ist in $\mathbb{Q}$ nicht lösbar.*

Beweis: (durch Widerspruch) Angenommen, die Gleichung $x^2 = 2$ hat eine rationale Lösung, also $x = p/q$ mit $p, q \in \mathbb{N}$. Durch fortgesetztes Kürzen können wir annehmen, dass höchstens eine der Zahlen p und q gerade ist. Nun gilt

$$p^2 = 2q^2 \quad \Rightarrow \quad p^2 \text{ gerade} \quad \Rightarrow \quad p \text{ gerade.}$$

Setze also $p = 2p_1$ mit $p_1 \in \mathbb{N}$, dann folgt weiter

$$2q^2 = 4p_1^2 \quad \Rightarrow \quad q^2 \text{ gerade} \quad \Rightarrow \quad q \text{ gerade .}$$

Also sind doch p, q beide gerade, ein Widerspuch. $\square$

In $\mathbb{R}$ ist die Gleichung $x^2 = 2$ und allgemeiner die Gleichung $x^n = a$ für beliebige $n \in \mathbb{N}$, $a > 0$ lösbar, wie wir in Kürze sehen werden. Das könnte man als Grund für die Erweiterung $\mathbb{Q} \subset \mathbb{R}$ anführen, ähnlich wie die Erweiterungen $\mathbb{N} \subset \mathbb{Z}$ beziehungsweise $\mathbb{Z} \subset \mathbb{Q}$ durch die Lösbarkeit von Gleichungen motiviert waren. Dies geht aber am Kern der Sache vorbei: einerseits bleibt die Gleichung $x^2 = -1$ auch in $\mathbb{R}$ unlösbar, andererseits bilden die reellen Nullstellen von beliebigen Polynomen mit rationalen Koeffizienten, die algebraischen Zahlen, nur eine relativ kleine Teilmenge von $\mathbb{R}$, wie wir auch noch zeigen werden.

Allgemein ist es das Ziel der Analysis, neue Objekte – Zahlen, Funktionen, Operationen – durch Grenzprozesse zu konstruieren. Unsere Definition des Grenzwerts setzt voraus, dass wir den Grenzwert der Folge bereits kennen. Das Vollständigkeitsaxiom muss die Existenz von Grenzwerten in Situationen garantieren, in denen der Grenzwert a priori nicht bekannt ist. Wir betrachten hierzu drei Beispiele.

Beispiel 4.1 (Verfahren von Heron) Für $a > 0$ wollen wir eine Lösung der Gleichung $x^2 = a$ als Grenzwert einer Folge konstruieren, und zwar setzen wir rekursiv

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right), \qquad \text{Startwert } x_0 > 0. \tag{4.1}$$

Zur Motivation der Formel: wir suchen die positive Nullstelle der Funktion $f(x) = x^2 - a$. Sei $x_n > 0$ schon berechnet, dann hat der Graph von f im Punkt $(x_n, f(x_n))$ die Tangente

$$y = f'(x_n)(x - x_n) + f(x_n).$$

Wir wählen für x_{n+1} den Schnittpunkt der Tangente mit der x-Achse, also

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - a}{2x_n} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right).$$

Was können wir über die Iteration aussagen? Zunächst sehen wir induktiv $x_n > 0$ für alle n, die Iteration bricht nicht ab. Weiter gilt

$$x_{n+1}^2 = \frac{1}{4}\left(x_n + \frac{a}{x_n}\right)^2 = \frac{1}{4}\left(x_n - \frac{a}{x_n}\right)^2 + a.$$

Also folgt $x_n^2 \geq a$ für $n \geq 1$. Weiter folgt

$$x_{n+1} = x_n - \frac{x_n^2 - a}{2x_n} \leq x_n \qquad \text{für } n \geq 1,$$

also $x_1 \geq x_2 \geq \ldots$. Hätten wir $x_n \to x$ mit $n \to \infty$, so folgt $x^2 \geq a > 0$ und

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) = \frac{1}{2}\left(x + \frac{a}{x}\right).$$

Die Lösung von $x^2 = a$ wäre damit gefunden. Die Konvergenz der Folge x_n können wir aber nicht mit Definition 3.2 zeigen, denn dafür müssten wir den Grenzwert $\sqrt{a}$ schon kennen, und den wollen wir ja gerade konstruieren.

Beispiel 4.2 (Zinseszinsrechnung) Wird ein Euro für x Jahre mit dem Zinsfaktor 1 angelegt, so ist die Auszahlung $f_1(x) = 1 + x$. Die Idee des Zinseszinses ist es, den Zeitraum in kürzere Intervalle zu unterteilen und den Zins anteilig anzurechnen mit dem Effekt, dass der schon gesparte Zinsanteil des Guthabens seinerseits Zinsen liefert. Wird der Zeitraum x in n Intervalle unterteilt, so ergibt das nach dem ersten Intervall $1 + \frac{x}{n}$, nach dem zweiten $(1 + \frac{x}{n})^2$, und schließlich nach dem n-ten Intervall, also insgesamt im Zeitraum x,

$$f_n(x) = \left(1 + \frac{x}{n}\right)^n \qquad \text{für } x \in \mathbb{R},\, n \in \mathbb{N}. \tag{4.2}$$

Es stellt sich die Frage nach kontinuierlicher Verzinsung, also dem Grenzwert $\lim_{n\to\infty} f_n(x)$.

Beispiel 4.3 (Dezimalbruchdarstellung) Für $a \in \mathbb{R}$ gibt es $k_0 \in \mathbb{Z}$ und $k_j \in \{0, 1, \ldots, 9\}$ für $j \in \mathbb{N}$, so dass folgende Darstellung als unendlicher Dezimalbruch gilt:

$$a = \lim_{n\to\infty} a_n \qquad \text{mit} \qquad a_n = \sum_{j=0}^{n} k_j \cdot 10^{-j} = k_0, k_1 k_2 \ldots k_n.$$

Um das zu zeigen, definieren wir induktiv $a_n = a_{n-1} + k_n \cdot 10^{-n}$ mit $a_n \leq a < a_n + 10^{-n}$. Für $n = 0$ setzen wir $k_0 = \max\{k \in \mathbb{Z} : k \leq a\}$ und haben wie gewünscht

$$a_0 = k_0 \leq a < k_0 + 1 = a_0 + 10^{-0}.$$

Sind $k_0, k_1, \ldots, k_{n-1}$ bereits gefunden für ein $n \in \mathbb{N}$, so definieren wir $k_n = \max\{k \in \mathbb{Z} : a_{n-1} + k \cdot 10^{-n} \leq a\}$. Nach Induktionsannahme gilt $a_{n-1} \leq a < a_{n-1} + 10 \cdot 10^{-n}$ und folglich $k_n \in \{0, 1, \ldots, 9\}$. Weiter liefert die Wahl von k_n

$$a_n = a_{n-1} + k_n \cdot 10^{-n} \leq a < a_{n-1} + (k_n + 1) \cdot 10^{-n} = a_n + 10^{-n}.$$

Die Darstellung von a als unendlicher Dezimalbruch folgt nun wegen

$$|a - a_n| \leq 10^{-n} \to 0 \text{ mit } n \to \infty.$$

Offen bleibt aber die umgekehrte Frage: seien $k_0, k_1, k_2, \ldots$ beliebig gegeben mit $k_0 \in \mathbb{Z}$ und $k_j \in \{0, 1, \ldots, 9\}$. Konvergiert dann die Dezimalbruchfolge $a_n = k_0, k_1 k_2 \ldots k_n$ gegen eine gewisse, reelle Zahl?

In diesen Beispielen brauchen wir wie gesagt eine Charakterisierung konvergenter Folgen, die ohne die Kenntnis des Grenzwerts auskommt. Die Idee von Augustin Louis Cauchy (1789–1857) besteht darin, die Glieder der Folge nicht mit dem unbekannten Grenzwert, sondern *untereinander* zu vergleichen.

Definition 4.1 (Cauchyfolge) *Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt Cauchyfolge, wenn gilt:*

$$\textit{Zu jedem } \varepsilon > 0 \textit{ gibt es ein } K \in \mathbb{R}, \textit{ so dass } |a_n - a_m| < \varepsilon \textit{ für alle } n, m > K.$$

Beim Nachweis dieser Eigenschaft reicht es aus, die Zahlen n, m mit $n < m$ zu betrachten, denn die Definition ist symmetrisch in n und m und für $n = m$ ist nichts zu tun.

(V) Vollständigkeitsaxiom: Jede Cauchyfolge ist konvergent.

Damit sind die Axiome (KAV) der reellen Zahlen komplett. Je nach Autor werden auch andere Aussagen als Vollständigkeitsaxiom zugrunde gelegt, die aber natürlich äquivalent sind und sich bei uns als Folgerungen ergeben werden.

Satz 4.2 *Eine Folge ist genau dann konvergent, wenn sie eine Cauchyfolge ist.*

Beweis: Eine Cauchyfolge ist konvergent nach dem Vollständigkeitsaxiom. Sei umgekehrt $\lim_{n\to\infty} a_n = a$. Zu $\varepsilon > 0$ gibt es dann ein $K \in \mathbb{R}$ mit $|a_n - a| < \varepsilon/2$ für $n > K$, und für $n, m > K$ folgt

$$|a_n - a_m| = |a_n - a + a - a_m| \leq |a_n - a| + |a_m - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Beispiel 4.4 (Konvergenz von Dezimalbrüchen) Seien $k_0 \in \mathbb{Z}$, $k_j \in \{0, 1, \ldots, 9\}$ gegeben für $j \in \mathbb{N}$. Dann konvergiert die Dezimalbruchfolge

$$a_n = \sum_{j=0}^{n} k_j \cdot 10^{-j}$$

gegen eine reelle Zahl. Für $n < m$ schätzen wir wie folgt ab, wobei wir die Formel für die geometrische Reihe, Beispiel 3.7, und Beispiel 3.3 verwenden:

$$|a_m - a_n| = \sum_{j=n+1}^{m} k_j \cdot 10^{-j} \leq 10^{-(n+1)} \sum_{j=0}^{\infty} 9 \cdot 10^{-j} \leq 10^{-n} < \varepsilon \qquad \text{für } n > K(\varepsilon).$$

Die Behauptung folgt also aus dem Vollständigkeitsaxiom.

Definition 4.2 (Monotone Folge) *Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt monoton wachsend, wenn*

$$a_{n+1} \geq a_n \qquad \textit{für alle } n \in \mathbb{N}.$$

Manche Autoren bezeichnen dies als nichtfallend, und reservieren den Begriff wachsend für eine Folge mit $a_{n+1} > a_n$; bei uns heißt das *streng monoton wachsend*.

Satz 4.3 (Konvergenzkriterium der Monotonie und Beschränktheit) *Jede monoton wachsende, nach oben beschränkte Folge ist eine Cauchyfolge und damit konvergent.*

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ monoton wachsend und $a_n \leq C < \infty$ für alle $n \in \mathbb{N}$. Für $\varepsilon > 0$ betrachten wir das Gitter $a_1 + \mathbb{Z}\cdot\varepsilon$, vgl. auch Satz 3.2, und setzen

$$M = \{j \in \mathbb{Z} : \text{es gibt ein } n \in \mathbb{N} \text{ mit } a_n \geq a_1 + j\varepsilon\}.$$

Offenbar ist $0 \in M$, und für $j \in M$ gilt $j \leq (C - a_1)/\varepsilon$. Sei $k \in M$ maximal. Dann gibt es ein $N \in \mathbb{N}$ mit $a_N \geq a_1 + k\varepsilon$, und für alle $n \geq N$ folgt

$$a_1 + k\varepsilon \leq a_N \leq a_n < a_1 + (k+1)\varepsilon.$$

Damit folgt $|a_n - a_m| \leq \varepsilon$ für $n, m \geq N$, das heißt (a_n) ist eine Cauchyfolge. □

Das Konvergenzkriterium der Monotonie und Beschränktheit ist nur hinreichend, denn es gibt auch konvergente Folgen die nicht monoton sind, etwa $a_n = (-1)^n/n$. Aber es ist ein enorm nützliches Kriterium, um Konvergenz zu zeigen.

Beispiel 4.5 (Konvergenz des Heronverfahrens) Die in Beispiel 4.1 konstruierte Folge x_n, $n \in \mathbb{N}_0$, erfüllte $x_n^2 \geq a$ und $x_1 \geq x_2 \geq \ldots$. Nach Satz 4.3 konvergiert die Folge gegen ein $x > 0$. Wir hatten bereits gesehen, dass dann die Gleichung $x^2 = a$ gelöst wird, das heißt wir haben eine Quadratwurzel von a gefunden.

Beispiel 4.6 (Die Zahl e und Zinsrechnung) Nach Beispiel 4.2 ergibt ein Euro mit Zinssatz Eins nach einem Jahr, wenn der Zins in n Intervallen berechnet wird,

$$f_n = \Big(1 + \frac{1}{n}\Big)^n.$$

Wir behaupten, dass die Folge f_n nach oben beschränkt und monoton wachsend ist, und damit konvergiert. Berechne dazu mit der binomischen Formel, siehe Satz 2.7,

$$f_n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} = \sum_{k=0}^{n} \frac{1}{k!}\,\frac{n}{n}\,\frac{n-1}{n}\cdot\ldots\cdot\frac{n-k+1}{n} \leq \sum_{k=0}^{n} \frac{1}{k!} =: e_n.$$

Nun ist $k! \geq 2^{k-1}$ für $k \geq 1$, also gilt

$$e_n \leq 1 + \sum_{k=1}^{n} 2^{1-k} \leq 1 + \sum_{k=1}^{\infty} 2^{1-k} = 3.$$

Also sind e_n und f_n nach oben beschränkt. Für die Monotonie verwende Bernoulli:

$$\begin{aligned}
\frac{f_n}{f_{n-1}} &= \frac{(n+1)^n}{n^n}\cdot\frac{(n-1)^{n-1}}{n^{n-1}} \\
&= \frac{n}{n-1}\left(\frac{(n+1)(n-1)}{n^2}\right)^n \\
&= \frac{n}{n-1}\big(1 - \frac{1}{n^2}\big)^n \\
&\geq \frac{n}{n-1}\big(1 - n\cdot\frac{1}{n^2}\big) = 1.
\end{aligned}$$

Also ist die Folge f_n konvergent.

Definition 4.3 (Eulersche Zahl) *Wir setzen*

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

Die Folge e_n ist offensichtlich monoton wachsend, auch sie hat damit einen Grenzwert. Wegen $e_n \geq f_n$ ist $\lim_{n\to\infty} e_n \geq e$. Andererseits haben wir für $m \leq n$, weil die Summanden nichtnegativ sind,

$$\sum_{k=0}^{m} \frac{1}{k!} \frac{n}{n} \frac{n-1}{n} \cdot \ldots \cdot \frac{n-k+1}{n} \leq f_n.$$

Die Produkte haben höchstens m Faktoren, die jeweils gegen Eins gehen. Mit $n \to \infty$ folgt

$$e_m = \sum_{k=0}^{m} \frac{1}{k!} \leq e, \quad \text{also} \lim_{m\to\infty} e_m \leq e.$$

Wir haben damit die alternative Formel

$$e = \sum_{k=0}^{\infty} \frac{1}{k!}. \tag{4.3}$$

Der Taschenrechner liefert zum Beispiel $e_3 = 2,666\ldots$ und $e_5 = 2,716\ldots$.

Satz 4.4 (Irrationalität der Eulerschen Zahl) $e = \sum_{k=0}^{\infty} 1/k!$ *ist nicht rational.*

BEWEIS: Wir zeigen, dass die e_n die Zahl e gut approximieren. Es gilt

$$\begin{aligned} e - e_n &= \sum_{k=n+1}^{\infty} \frac{1}{k!} \\ &= \frac{1}{(n+1)!} \left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \ldots\right) \\ &\leq \frac{1}{(n+1)!} \left(1 + 2^{-1} + 2^{-2} + \ldots\right) \\ &= \frac{2}{(n+1)!}. \end{aligned}$$

Nach Multiplikation mit $n!$ ergibt sich hieraus

$$0 < n!(e - e_n) \leq \frac{2}{n+1} < 1 \quad \text{für } n \geq 2.$$

Nun ist $n!e_n = \sum_{k=0}^{n} n!/k! \in \mathbb{N}$. Wäre e rational, so wäre der mittlere Term eine ganze Zahl für n hinreichend groß, ein Widerspruch. □

Als nächste Anwendung des Vollständigkeitsaxioms diskutieren wir nun das Intervallschachtelungsprinzip.

Definition 4.4 (Intervallschachtelung) *Eine Intervallschachtelung ist eine Folge von abgeschlossenen Intervallen* $I_n = [a_n, b_n] \subset \mathbb{R}$ *mit* $I_{n+1} \subset I_n$ *für alle* n *und* $|I_n| = b_n - a_n \to 0$ *mit* $n \to \infty$.

Satz 4.5 (Intervallschachtelungsprinzip) *Zu jeder Intervallschachtelung $(I_n)_{n\in\mathbb{N}}$ gibt es genau ein $x \in \mathbb{R}$ mit $x \in \bigcap_{n\in\mathbb{N}} I_n$, also $x \in I_n$ für alle $n \in \mathbb{N}$. Es gilt $x = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.*

Beweis: Nach Voraussetzung ist $a_1 \leq a_2 \leq \ldots$ und $b_1 \geq b_2 \geq \ldots$. Ferner gilt

$$a_n \leq b_n \leq b_1 \quad \text{und} \quad b_n \geq a_n \geq a_1 \quad \text{für alle } n \in \mathbb{N}.$$

Aus Satz 4.3 folgt die Existenz der Grenzwerte $a = \lim_{n\to\infty} a_n$ bzw. $b = \lim_{n\to\infty} b_n$. Dann folgt nach Satz 3.5 und Satz 3.6

$$0 \leq b - a = \lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = \lim_{n\to\infty} (b_n - a_n) = 0.$$

Setze $x := a = b$. Dann ist $a_n \leq a = x = b \leq b_n$, also $x \in I_n$ für alle n. Sei $y \in \mathbb{R}$ mit $y \in I_n$ für alle n, das heißt $a_n \leq y \leq b_n$. Durch Grenzübergang ergibt sich nach Satz 3.6 $a \leq y \leq b$, also $y = x$. $\square$

Wir wollen eine Intervallschachtelung verwenden, um für jedes $a \geq 0$ und $n \in \mathbb{N}$ die Zahl $a^{1/n}$ zu konstruieren, also die Lösung $x \geq 0$ der Gleichung $x^n = a$. Dazu eine Vorüberlegung: eine Funktion $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$, heißt *streng monoton wachsend*, wenn folgende Implikation gilt:

$$x_1 < x_2 \quad \Rightarrow \quad f(x_1) < f(x_2). \tag{4.4}$$

f ist dann injektiv, denn ist $f(x_1) = f(x_2)$, so kann weder $x_1 < x_2$ noch $x_1 > x_2$ sein, also $x_1 = x_2$. Die Funktion $f : I \to f(I)$ ist somit injektiv und surjektiv, das heißt es gibt die Umkehrfunktion $g : f(I) \to I$. Diese Umkehrfunktion ist dann ebenfalls streng monoton wachsend: wäre $a_1 < a_2$ mit $g(a_1) \geq g(a_2)$, so folgt aus der Monotonie von f

$$a_1 = f(g(a_1)) \geq f(g(a_2)) = a_2, \quad \text{Widerspruch.} \tag{4.5}$$

In unserem Fall ist $f : I = [0,\infty) \to \mathbb{R}$, $f(x) = x^n$, streng monoton wachsend nach Anordnungsaxiom (A2). Da f injektiv ist, gibt es also höchstens eine Lösung der Gleichung $f(x) = a$. Die Umkehrfunktion $g : f(I) \to I$ erfüllt $a = f(g(a)) = g(a)^n$, also ist $g(a) = a^{1/n}$. Allerdings ist g nur auf $f(I)$ definiert, wir müssen noch zeigen dass $f(I) = [0,\infty)$. Das ist genau das Problem der Existenz der Wurzel, was wir nun angehen.

Satz 4.6 (Existenz der n-ten Wurzel) *Zu jedem $a \geq 0$ und $n \in \mathbb{N}$ gibt es genau ein $x \geq 0$ mit $x^n = a$, Bezeichnung $\sqrt[n]{a}$ oder $a^{1/n}$. Die Funktion $a \mapsto a^{1/n}$ ist streng monoton wachsend auf $[0,\infty)$.*

Beweis: Wir konstruieren die Lösung mit dem Verfahren der fortgesetzten Intervallhalbierung: bestimme $I_k = [a_k, b_k]$ für $k = 1, 2, \ldots$, so dass mit $m_k = \frac{a_k + b_k}{2}$ gilt:

$$I_1 = [a_1, b_1] \text{ mit } a_1^n \leq a \leq b_1^n,$$

$$I_{k+1} = \begin{cases} [a_k, m_k] & \text{falls } m_k^n \geq a \\ [m_k, b_k] & \text{falls } m_k^n < a. \end{cases}$$

Es folgt $I_{k+1} \subset I_k$ für alle k und $|I_k| = 2^{1-k}|I_1| \to 0$ mit $k \to \infty$. Sei $x \in \mathbb{R}$ der durch die Intervallschachtelung definierte Punkt, siehe Satz 4.5. Aus der Definition von I_{k+1} ergibt sich per Induktion $a_k^n \leq a \leq b_k^n$ für alle $k \in \mathbb{N}$, und hieraus mit den Sätzen 3.5 und 3.6

$$x^n = \lim_{n\to\infty} a_k^n \leq a \leq \lim_{n\to\infty} b_k^n = x^n.$$

$\square$

Für rationale Exponenten $r = p/q$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ wird die Potenz erklärt durch $a^r = (a^p)^{1/q}$. Dies ist wohldefiniert, denn aus $p_1/q_1 = p_2/q_2$ folgt

$$\Big((a^{p_1})^{1/q_1}\Big)^{q_1 q_2} = \Big(((a^{p_1})^{1/q_1})^{q_1}\Big)^{q_2} = (a^{p_1})^{q_2} = a^{p_1 q_2} = a^{p_2 q_1} = \Big((a^{p_2})^{1/q_2}\Big)^{q_1 q_2},$$

also $(a^{p_1})^{1/q_1} = (a^{p_2})^{1/q_2}$. Weiter zeigt man leicht die Potenzgesetze (für ganzzahlige Exponenten sind diese Regeln klar und wurden schon benutzt)

$$(i)\ a^r a^s = a^{r+s} \quad (ii)\ (a^r)^s = a^{rs} \quad (iii)\ a^r b^r = (ab)^r.$$

Zum Beispiel gilt mit $r = k/m$ und $s = p/q$

$$(a^r a^s)^{mq} = (a^{k/m})^{mq}\,(a^{p/q})^{mq} = \Big(((a^k)^{1/m})^m\Big)^q \cdot \Big(((a^p)^{1/q})^q\Big)^m = (a^k)^q \cdot (a^p)^m = a^{kq+pm}.$$

Dies bedeutet

$$a^r a^s = \big(a^{kq+pm}\big)^{1/mq} = a^{(kq+pm)/mq} = a^{k/m+p/q} = a^{r+s}.$$

Die anderen beiden Potenzgesetze werden ähnlich verifiziert.

Beispiel 4.7 Ein Grenzwert, der häufiger auftritt, ist

$$\lim_{n\to\infty} a^{1/n} = 1 \quad \text{für } a > 0. \tag{4.6}$$

Betrachte erst den Fall $a \geq 1$, also auch $a^{1/n} \geq 1$ nach Monotonie der Wurzel. Setze $a^{1/n} = 1 + \xi_n$ mit $\xi_n \geq 0$. Dann folgt aus der Bernoulli-Ungleichung, Satz 2.2,

$$a = (1+\xi_n)^n \geq 1 + n\xi_n \quad \Rightarrow \quad 0 \leq \xi_n \leq (a-1)/n.$$

Also folgt $\lim_{n\to\infty} \xi_n = 0$ bzw. $\lim_{n\to\infty} a^{1/n} = 1$ mit Satz 3.6 c). Für $a < 1$ gilt

$$\Big(a^{1/n} \cdot (a^{-1})^{1/n}\Big)^n = a \cdot a^{-1} = 1 \quad \Rightarrow \quad a^{1/n} = \frac{1}{(a^{-1})^{1/n}},$$

und die Behauptung folgt aus dem vorigen Fall.

Definition 4.5 (Teilfolge) *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge und $(n_k)_{k\in\mathbb{N}}$ eine Folge natürlicher Zahlen mit $n_1 < n_2 < n_3 \ldots$ Dann heißt die Folge $(a_{n_k})_{k\in\mathbb{N}}$ Teilfolge von $(a_n)_{n\in\mathbb{N}}$.*

Durch Induktion ergibt sich sofort $n_k \geq k$: es ist $n_1 \geq 1$ und $n_{k+1} \geq n_k + 1 \geq k + 1$. Die Teilfolge entsteht aus der ursprünglichen Folge durch *Auswahl der Nummern n_k*. Da Folgen Abbildungen von $\mathbb{N}$ nach $\mathbb{R}$ sind, ist eine Teilfolge formal als Verkettung von zwei Abbildungen definiert: der Ausgangsfolge $a : \mathbb{N} \to \mathbb{R}$, $n \longmapsto a_n$ und der Folge $\mathbb{N} \to \mathbb{N}$, $k \longmapsto n_k$, die die Indizes auswählt. Am Beispiel $a_n = (-1)^n/n^3$ und $n_k = 2k-1$ sieht das wie folgt aus:

$$\begin{array}{ccccc}
\mathbb{N} & \stackrel{n_k=2k-1}{\longrightarrow} & \mathbb{N} & \stackrel{a_n=(-1)^n/n^3}{\longrightarrow} & \mathbb{R} \\
 & & & & \\
k & \longmapsto & n_k = 2k-1 & \longmapsto & a_{n_k} = -1/(2k-1)^3
\end{array}$$

Definition 4.6 (Häufungspunkt von Folgen) *$a \in \mathbb{R}$ heißt Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$, wenn eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ gibt, die mit $k \to \infty$ gegen a konvergiert.*

Beispielsweise hat die Folge $a_n = (-1)^n + 1/n^2$ den Häufungspunkt $+1$, denn mit $n_k = 2k$ gilt $a_{n_k} = a_{2k} = 1 + 1/(2k)^2 \to 1$ mit $k \to \infty$. Auch -1 ist ein Häufungspunkt der Folge, denn für $n_k = 2k - 1$ ist $a_{n_k} = a_{2k-1} = -1 + 1/(2k-1)^2 \to -1$ mit $k \to \infty$.

Lemma 4.1 *Die Zahl $a \in \mathbb{R}$ ist genau dann Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$, wenn für jedes $\varepsilon > 0$ die Menge $\{n \in \mathbb{N} : a_n \in U_\varepsilon(a)\}$ unendlich viele Elemente hat.*

Beweis: Wenn $a \in \mathbb{R}$ Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$ ist, so gilt nach Definition $a_{n_k} \to a$ mit $k \to \infty$ für eine Teilfolge n_k. Zu jedem $\varepsilon > 0$ gibt es dann ein $K \in \mathbb{R}$ mit $a_{n_k} \in U_\varepsilon(a)$ für alle $k > K$. Die Abbildung $k \mapsto n_k$ ist injektiv wegen $n_1 < n_2 < \ldots$, also ist die Menge $\{n_k : k > K\}$ nicht endlich nach dem Schubfachprinzip. Dies beweist die eine Richtung der Äquivalenz. Umgekehrt wählen wir induktiv n_k mit $n_1 < n_2 < \ldots$, so dass $a_{n_k} \in U_{1/k}(a)$. Die Induktion bricht nicht ab, da $a_n \in U_{1/k}(a)$ für unendlich viele n gilt. Es folgt dann $|a_{n_k} - a| < 1/k \to 0$ mit $k \to \infty$. □

Satz 4.7 (Bolzano-Weierstraß) *Jede beschränkte Folge reeller Zahlen hat eine konvergente Teilfolge, also mindestens einen Häufungspunkt.*

Beweis: Konstruktion durch fortgesetzte Intervallhalbierung: wähle eine obere Schranke b_1 und eine untere Schranke a_1 für $(x_n)_{n\in\mathbb{N}}$, also $x_n \in [a_1, b_1]$ für alle $n \in \mathbb{N}$. Wir nehmen nun induktiv an, dass $I_k = [a_k, b_k]$ schon gefunden ist mit der Eigenschaft

$$(*) \quad x_n \in I_k \quad \text{für unendlich viele } n.$$

Setze $m_k = \frac{1}{2}(a_k + b_k)$ und definiere

$$I_{k+1} = [a_{k+1}, b_{k+1}] = \begin{cases} [m_k, b_k], & \text{falls } x_n \in [m_k, b_k] \text{ für unendlich viele } n, \\ [a_k, m_k] & \text{sonst.} \end{cases}$$

Es ist offensichtlich, dass $(*)$ auch für das Intervall I_{k+1} gilt. Nach dem Intervallschachtelungsprinzip aus Satz 4.5 gibt es ein $x \in \mathbb{R}$ mit $x \in I_k$ für alle $k \in \mathbb{N}$. Zu $\varepsilon > 0$ gibt es nun ein $K \in \mathbb{R}$ mit $|I_k| < \varepsilon$ für $k > K$, also $I_k \subset U_\varepsilon(x)$ für $k > K$. Damit gilt auch $x_n \in U_\varepsilon(x)$ für unendlich viele n. Aus Lemma 4.1 schließen wir, dass x ein Häufungspunkt der Folge $(x_n)_{n\in\mathbb{N}}$ ist. □

Definition 4.7 (Limes superior/inferior) *Für eine reelle Folge $(x_n)_{n\in\mathbb{N}}$ und $x^* \in \mathbb{R} \cup \{\pm\infty\}$ gilt $\limsup_{n\to\infty} x_n = x^*$, falls folgende zwei Bedingungen erfüllt sind:*

(i) *es gibt eine Teilfolge x_{n_k} mit $x_{n_k} \to x^*$ für $k \to \infty$,*

(ii) *für alle $x > x^*$ ist die Menge $\{n \in \mathbb{N} : x_n > x\}$ endlich.*

Entsprechend bedeutet $\liminf_{n\to\infty} x_n = x_*$ mit $x_* \in \mathbb{R} \cup \{\pm\infty\}$:

(i) *es gibt eine Teilfolge x_{n_k} mit $x_{n_k} \to x_*$ für $k \to \infty$,*

(ii) *für alle $x < x_*$ ist die Menge $\{n \in \mathbb{N} : x_n < x\}$ endlich.*

Um zu sehen, wie die Definition funktioniert, betrachten wir ein einfaches

Beispiel 4.8 Sei $x_n = (-1)^n + \frac{1}{n^2}$. Dann gilt $x_{2k} = 1 + \frac{1}{4k^2} \to 1$, also Bedingung (i). Für $x > 1$ folgern wir

$$x_n > x \quad \Rightarrow \quad \frac{1}{n^2} > x - (-1)^n \geq x - 1 \quad \Rightarrow \quad n \leq n^2 < \frac{1}{x-1}.$$

Also ist Bedingung (ii) ebenfalls erfüllt, und $\limsup_{n\to\infty} x_n = 1$.

Ist $x^* = \limsup_{n\to\infty} x_n \in \mathbb{R}$, so ist x^* der größte Häufungspunkt. Denn zu $x > x^*$ gibt es ein $\varepsilon > 0$ mit $x - \varepsilon > x^*$, also nach (ii)

$$\{n \in \mathbb{N} : x_n \in U_\varepsilon(x)\} \subset \{n \in \mathbb{N} : x_n > x - \varepsilon\} = \text{ endlich.}$$

Im Fall $\limsup_{n\to\infty} x_n = -\infty$ gilt schon $\lim_{n\to\infty} x_n = -\infty$, wieder nach (ii). Dagegen muss im Fall $\limsup_{n\to\infty} x_n = +\infty$ nicht die ganze Folge konvergieren, etwa $x_n = (-1)^n n$. Während wir den Grenzwert nur für konvergente Folgen haben, sind lim sup bzw. lim inf für jede beliebige Folge x_n definiert. Das wollen wir nun zeigen, wobei wir uns auf den Limes superior beschränken.

Satz 4.8 (Existenz des Limes superior) *Für jede Folge $(x_n)_{n\in\mathbb{R}}$ gibt es genau ein $x^* \in \mathbb{R} \cup \{\pm\infty\}$ mit $\limsup_{n\to\infty} x_n = x^*$.*

Beweis:

Fall 1: (x_n) ist nicht nach oben beschränkt.
Dann ist $\{n : x_n \geq b\}$ unendlich für alle $b \in \mathbb{R}$. Bestimme induktiv $n_1 < n_2 < \ldots$ mit $x_{n_k} \geq k$. Es folgt $\limsup_{n\to\infty} x_n = +\infty$.

Fall 2: Es gibt ein $b_1 \in \mathbb{R}$ mit $x_n \leq b_1$ für alle $n \in \mathbb{N}$.

Fall 2.1: $\{n : x_n \geq a\}$ ist endlich für alle $a \in \mathbb{R}$.
Dann gilt $x_n \to -\infty$ und es folgt $\limsup_{n\to\infty} x_n = -\infty$.

Fall 2.2: Es gibt ein $a_1 \in \mathbb{R}$, so dass $\{n : x_n \in [a_1, b_1]\}$ unendlich ist.
Das Intervallhalbierungsverfahren aus Satz 4.7 liefert $x^* = \lim_{k\to\infty} b_k = \lim_{k\to\infty} a_k$. Wie dort bewiesen ist x^* ein Häufungspunkt, also gilt Bedingung (i). Wir behaupten

$$\{n \in \mathbb{N} : x_n > b_k\} \text{ ist endlich für jedes } k = 1, 2, \ldots.$$

Für $k = 1$ ist das richtig, da b_1 obere Schranke (Fall 2). Induktiv ergibt sich, je nachdem ob bei der Halbierung das rechte oder das linke Teilintervall gewählt wird:

$$\begin{aligned} b_{k+1} = b_k \quad &\Rightarrow \quad \{n : x_n > b_{k+1}\} = \{n : x_n = b_k\} = \text{ endlich (Induktion)}, \\ b_{k+1} = m_k \quad &\Rightarrow \quad \{n : x_n > b_{k+1}\} = \{n : x_n \in (m_k, b_k]\} \cup \{n : x_n > b_k\} = \text{ endlich}, \\ &\quad\quad \text{(Fallunterscheidung und Induktion)}. \end{aligned}$$

Für $x > x^*$ gilt $(x, \infty) \subset (b_k, \infty)$ für k hinreichend groß, also ist $\{n : x_n > x\}$ endlich. Dies zeigt $\limsup_{n\to\infty} x_n = x^*$, die Existenz.

Angenommen es gibt $x_1^* < x_2^*$ mit den Eigenschaften (i) und (ii). Wähle $x \in (x_1^*, x_2^*)$.

Wegen (ii) für x_1^* ist dann $\{n : x_n > x\}$ endlich. Dann kann aber (i) für x_2^* nicht gelten, ein Widerspruch. $\square$

Die Begriffe Häufungspunkt, Limes superior und Limes inferior sind gewöhnungsbedürftig, und wir werden bei Gelegenheit mehr Beispiele betrachten. Die logische Abfolge der *zentralen theoretischen Aussagen* in diesem Abschnitt war folgende:

Vollständigkeitsaxiom:
Cauchyfolgen sind konvergent
$\Downarrow$
Konvergenzkriterium der
Monotonie und Beschränktheit
$\Downarrow$
Intervallschachtelungsprinzip
$\Downarrow$
Bolzano-Weierstraß: Auswahlsatz
$\Downarrow$
Cauchyfolgen sind konvergent

Die Implikationen sind so zu verstehen, dass jeweils nur die jeweils vorangehende Eigenschaft von $\mathbb{R}$ im Beweis des darauf folgenden Resultats benutzt wurde. Die letzte Implikation werden wir dabei gleich noch zeigen. Es folgt, dass jede der vier Eigenschaften als Axiom für $\mathbb{R}$ benutzt werden könnte - die anderen Eigenschaften würden als Sätze folgen. Im nächsten Abschnitt werden wir eine weitere, äquivalente Eigenschaft kennenlernen, nämlich den Satz vom Supremum (Satz 5.1). In dieser Vorlesung wird die Konvergenz der Cauchyfolgen als grundlegendes Axiom gewählt.

BEWEIS: *Auswahlsatz von Bolzano-Weierstraß $\Rightarrow$ Cauchyfolgen sind konvergent*
Wir zeigen zunächst, dass jede Cauchyfolge $(a_n)_{n\in\mathbb{N}}$ beschränkt ist. Es gibt ein $n_0 \in \mathbb{N}$ mit $|a_n - a_m| \leq 1$ für $n, m \geq n_0$, also folgt

$$|a_n| \leq |a_n - a_{n_0}| + |a_{n_0}| \leq 1 + |a_{n_0}| \quad \text{für alle } n \geq n_0.$$

Nach Bolzano-Weierstraß gibt es eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ mit $a_{n_k} \to a$ für $k \to \infty$. Es gilt:

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| \quad \text{für alle } n, k \in \mathbb{N}.$$

Da a_n Cauchyfolge, ist $|a_n - a_m| < \varepsilon$ für $n, m > N$. Also folgt für $n > N$

$$\lim_{k\to\infty} |a_n - a_{n_k}| \leq \varepsilon \quad \text{und} \quad \lim_{k\to\infty} |a_{n_k} - a| = 0.$$

Dies zeigt $|a_n - a| \leq \varepsilon$ für $n > N$, das heißt die ganze Folge konvergiert gegen a. $\square$

5 Teilmengen von $\mathbb{R}$ und von $\mathbb{R}^n$

Für Mengen im $\mathbb{R}^n$ haben wir analoge Begriffe wie bei Folgen, mit kleinen Abwandlungen.

Definition 5.1 *Die Menge $M \subset \mathbb{R}$ heißt*

$$\begin{aligned} \textit{nach oben beschränkt} \quad &\Leftrightarrow \quad \exists\, K \in \mathbb{R} \textit{ mit } x \leq K \textit{ für alle } x \in M, \\ \textit{nach unten beschränkt} \quad &\Leftrightarrow \quad \exists\, K \in \mathbb{R} \textit{ mit } x \geq K \textit{ für alle } x \in M. \end{aligned}$$

Die Zahl K heißt dann obere bzw. untere Schranke. Weiter heißt

$$M \textit{ beschränkt} \quad \Leftrightarrow \quad \exists\, K \geq 0 \textit{ mit } |x| \leq K \textit{ für alle } x \in M.$$

Eine Menge ist genau dann beschränkt, wenn sie nach oben und unten beschränkt ist. Denn aus $|x| \leq K$ folgt $-K \leq x \leq K$, und aus $K_1 \leq x \leq K_2$ folgt umgekehrt $|x| \leq \max(|K_1|, |K_2|)$.

Beispiel 5.1 Die Menge $[0,1)$ ist nach oben beschränkt, eine obere Schranke ist zum Beispiel $K = 2015$. Es gibt in $[0,1)$ kein größtes Element, denn es gilt der Schluss

$$x \in [0,1) \quad \Rightarrow \quad x < \frac{x+1}{2} \in [0,1).$$

Unter den oberen Schranken von $[0,1)$ gibt es aber eine kleinste, nämlich die Zahl 1.

Definition 5.2 (Supremum/Infimum) *Die Zahl $a \in \mathbb{R} \cup \{\pm\infty\}$ heißt Supremum (bzw. Infimum) der Menge $M \subset \mathbb{R}$, wenn folgendes gilt:*

(1) *$x \leq a$ für alle $x \in M$ (bzw. $x \geq a$ für alle $x \in M$),*

(2) *Für alle $a' < a$ (bzw. $a' > a$) gibt es ein $x \in M$ mit $x > a'$ (bzw. $x < a'$).*

Notation: $a = \sup M$ (bzw. $a = \inf M$). Ist M nach oben (bzw. unten) beschränkt, so bezeichnen wir $\sup M$ auch als kleinste obere Schranke (bzw. $\inf M$ als größte untere Schranke).

Satz 5.1 *Jede Menge $M \subset \mathbb{R}$ hat genau ein Supremum $S \in \mathbb{R} \cup \{\pm\infty\}$.*

Beweis: Wir zeigen als erstes die Eindeutigkeit. Angenommen es gibt $S_{1,2} \in \mathbb{R} \cup \{\pm\infty\}$, die beide die Definition 5.2 erfüllen; wir können $S_1 < S_2$ annehmen. Nach Eigenschaft (2) bzgl. S_2 gibt es ein $x \in M$ mit $x > S_1$, im Widerspruch zur Eigenschaft (1) bezüglich S_1.

Man sieht leicht $\sup \emptyset = -\infty$, und $\sup M = +\infty$ falls M nicht nach oben beschränkt ist. Im verbleibenden Fall wählen wir ein Element a_1 von M und eine obere Schranke b_1 von M, und konstruieren wie folgt eine Intervallschachtelung $I_n = [a_n, b_n]$ für $n \geq 1$, wobei $m_n = (a_n + b_n)/2$:

$$I_{n+1} = [a_{n+1}, b_{n+1}] = \begin{cases} [m_n, b_n], & \text{falls } [m_n, b_n] \cap M \neq \emptyset \\ [a_n, m_n], & \text{sonst.} \end{cases}$$

Nach dem Intervallschachtelungsprinzip, Satz 4.5, gibt es genau ein $S \in \mathbb{R}$ mit $S \in I_n$ für alle n, und genauer gilt $a_n \to S$ sowie $b_n \to S$. Für $x \in M$ sieht man durch Induktion $x \leq b_n$ für alle n, also $x \leq \lim_{n \to \infty} b_n = S$. Andererseits gilt, ebenfalls induktiv, $M \cap I_n \neq \emptyset$ für alle n, das heißt es gibt $x_n \in M$ mit $a_n \leq x_n \leq b_n$. Ist $S' < S$, so gilt also $x_n > S'$ für hinreichend große n. Damit ist $\sup M = S$ gezeigt. $\square$

Folgerung 5.1 *Sei $M \subset \mathbb{R}$ nichtleer. Dann gibt es eine Folge $x_n \in M$ (bzw. $x'_n \in M$) mit $x_n \to \sup M$ (bzw. $x'_n \to \inf M$).*

Beweis: Wir zeigen die Aussage für das Supremum. Ist M nach oben beschränkt, so wurde eine solche Folge im Beweis des vorangehenden Satzes konstruiert. Ist M nicht nach oben beschränkt, so gibt es zu $n \in \mathbb{N}$ ein $x_n \in M$ mit $x_n \geq n$, also $x_n \to +\infty = \sup M$. □

Jetzt wollen wir uns einer neuen Frage zuwenden: wie kann man die unendlichen Mengen $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ der Größe nach vergleichen? Gibt es wirklich mehr rationale Zahlen als natürliche? Mehr reelle als rationale? Die Antwort lautet witzigerweise, dass es gleich viele natürliche, ganze und rationale Zahlen gibt, aber mehr reelle Zahlen. Alle genannten Mengen enthalten unendlich viele Elemente. Die Präzisierung der Begriffe *gleichviele* und *mehr* geht auf Georg Cantor (1872) zurück.

Definition 5.3 *Eine Menge A heißt gleichmächtig zur Menge B (Notation: $A \sim B$), wenn es eine bijektive Abbildung (Bijektion) $\varphi : A \to B$ gibt.*

Lemma 5.1 *Die Relation $A \sim B$ (A ist gleichmächtig zu B) ist eine Äquivalenzrelation, das heißt für alle Mengen A, B, C gilt:*

(i) $A \sim A$

(ii) $A \sim B \quad \Rightarrow \quad B \sim A$

(iii) $A \sim B$ und $B \sim C \quad \Rightarrow \quad A \sim C$.

Beweis:

(i) Wähle als Bijektion die Identität $id_A : A \to A$, $a \longmapsto a$.

(ii) Sei $\varphi : A \to B$ Bijektion. Wähle dann $\varphi^{-1} : B \to A$.

(iii) Seien $\varphi : A \to B$ und $\psi : B \to C$ bijektiv. Wähle dann $\psi \circ \varphi : A \to C$.

□

Durch die Relation *gleichmächtig* werden die Mengen in Äquivalenzklassen eingeteilt. Für jedes $n \in \mathbb{N}_0$ hat man die Klasse der Mengen mit n Elementen, ein Repräsentant ist die Menge $\{1, \ldots, n\}$. Eine weitere Äquivalenzklasse liefert die Menge $\mathbb{N}$, denn nach dem Schubfachprinzip, Satz 2.4, gibt es für $n \in \mathbb{N}$ keine Bijektion $\mathbb{N} \leftrightarrow \{1, \ldots, n\}$.

Definition 5.4 *Eine Menge A heißt*

- *abzählbar unendlich, wenn sie gleichmächtig zu $\mathbb{N}$ ist;*
- *abzählbar, wenn sie endlich oder abzählbar unendlich ist;*
- *überabzählbar, wenn sie nicht abzählbar ist.*

Lemma 5.2 *Falls eine surjektive Abbildung $\varphi : \mathbb{N} \to A$ existiert, so ist A abzählbar.*

Beweis: Sei $a_n = \varphi(n)$. Die naheliegende Idee ist, bei der Abzählung induktiv diejenigen a_n auszulassen, die bereits vorher auftraten, und die restlichen entsprechend neu zu nummerieren. Dazu setzen wir $n_1 = 1$ und konstruieren induktiv eine Teilfolge a_{n_k} durch die Vorschrift

$$n_{k+1} = \min\{n > n_k : a_n \neq a_{n_j} \text{ für } j = 1, \ldots, k\}.$$

Falls die Rekursion nach einem n_k abbricht, ist die Abbildung $f : \{1, \ldots, k\} \to A$, $k \mapsto a_{n_k}$ bijektiv und damit ist A endlich. Andernfalls ist die Abbildung $f : \mathbb{N} \to A$, $k \mapsto a_{n_k}$ bijektiv und A ist abzählbar unendlich. □

Satz 5.2 *Die Mengen $\mathbb{Z}$ und $\mathbb{Q}$ sind abzählbar.*

Beweis: Für $\mathbb{Z}$ wähle die surjektive (sogar bijektive) Abbildung

$$\varphi : \mathbb{N} \to \mathbb{Z},\ \varphi(n) = \begin{cases} (n-1)/2 & n \text{ ungerade} \\ -n/2 & n \text{ gerade.} \end{cases}$$

Das Argument zeigt, dass wir uns beim Beweis der Abzählbarkeit von $\mathbb{Q}$ auf die Menge $\mathbb{Q}^+ = \{p/q : p, q \in \mathbb{N}\}$ beschränken können. Die Idee ist dann, diagonal nach folgendem Schema abzuzählen:

$$\begin{array}{llcccccc}
 & p = & 1 & 2 & 3 & 4 & 5 & 6 \\
q = & & & & & & & \\
1 & & 1/1 & 2/1 & 3/1 & 4/1 & 5/1 & 6/1 \\
 & & \swarrow & \swarrow & \swarrow & \swarrow & \swarrow & \\
2 & & 1/2 & 2/2 & 3/2 & 4/2 & \cdot & \\
 & & \swarrow & \swarrow & \swarrow & \cdot & & \\
3 & & 1/3 & 2/3 & 3/3 & \cdot & & \\
 & & \swarrow & \swarrow & \cdot & & & \\
4 & & 1/4 & 2/4 & \cdot & & & \\
 & & \swarrow & \cdot & & & & \\
5 & & 1/5 & & & & &
\end{array}$$

Die k-te Diagonale enthält k Elemente, also enthalten die Diagonalen mit kleinerer Nummer insgesamt $1 + \ldots + (k-1) = (k-1)k/2$ Einträge. Jedes $n \in \mathbb{N}$ liegt in genau einem k-Abschnitt $\{(k-1)k/2 + i : i = 1, \ldots, k\}$, denn es gilt $(1-1)1/2 + 1 = 1$ und

$$(k-1)k/2 + k = k(k+1)/2.$$

Somit ist folgende Abbildung wohldefiniert:

$$\varphi : \mathbb{N} \to \mathbb{Q},\ \varphi(n) = \frac{k-i+1}{i} \qquad \text{für } n = \frac{(k-1)k}{2} + i \quad (k \in \mathbb{N},\ 1 \leq i \leq k).$$

Die Abbildung ist surjektiv, und zwar gilt $\varphi(n) = p/q$ für $i = q$ und $k = p + q - 1$. Die Abzählbarkeit von $\mathbb{Q}$ folgt nun mit Lemma 5.2. $\square$

Durch Anordnung in einem quadratischen Schema und analoge Abzählung läßt sich ganz analog zeigen, daß eine abzählbare Vereinigung von jeweils abzählbaren Mengen abermals abzählbar ist.

Satz 5.3 ($\mathbb{R}$ ist nicht abzählbar) *Es gibt keine surjektive Abbildung $\varphi : \mathbb{N} \to \mathbb{R}$.*

Beweis: Sei $\varphi : \mathbb{N} \to \mathbb{R}$ eine beliebige Abbildung und $\varphi(n) = x_n$. Wir konstruieren eine Intervallschachtelung $(I_n)_{n \in \mathbb{N}}$ mit $x_n \notin I_n$ für jedes n. Ist I_n schon bestimmt, so zerlege I_n in drei abgeschlossene Teilintervalle gleicher Länge und wähle für I_{n+1} ein Teilintervall, das x_{n+1} nicht enthält (im Zweifelsfall das rechte). Um I_1 zu definieren, wenden wir dieses Argument an auf $I_0 = [0, 1]$. Nach Satz 4.5 gibt es ein $x \in \bigcap_{n \in \mathbb{N}} I_n \in \mathbb{R}$, also $x \neq x_n$ für alle $n \in \mathbb{N}$. $\square$

Die Menge $\mathbb{R}$ repräsentiert also eine weitere Äquivalenzklasse der Relation gleichmächtig, die als Mächtigkeit des Kontinuums bezeichet wird. Cantor hat 1878 vermutet, dass jede Teilmenge von $\mathbb{R}$ entweder abzählbar oder gleichmächtig zu $\mathbb{R}$ ist. Es hat sich aber mit Arbeiten von Gödel (1938) und Cohen (1963) ergeben, dass diese sogenannte Kontinuumshypothese im Rahmen unserer Mengenaxiomatik (die wir nicht behandelt haben) nicht entschieden werden kann, ein merkwürdiges Ergebnis.

In Zukunft brauchen wir den Begriff der Konvergenz auch für Punkte im $\mathbb{R}^n$. Für $n = 1$ haben wir als Modell die Zahlengerade benutzt, entsprechend betrachten wir für $n = 2$ die Ebene mit kartesischen Koordinaten $z = (x, y)$ und für $n = 3$ den dreidimensionalen Raum mit Koordinaten $p = (x, y, z)$; dabei finde ich den zweidimensionalen Fall besonders anschaulich. Der $\mathbb{R}^n$ ist eine mathematische Verallgemeinerung:

$$\mathbb{R}^n = \{x = (x_1, \ldots, x_n) : x_i \in \mathbb{R}\} = \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n \text{ mal}}.$$

Die Vektoraddition und Skalarmultiplikation sind komponentenweise definiert:

$$\begin{aligned} x + y &= (x_1 + y_1, \ldots, x_n + y_n) \quad (x, y \in \mathbb{R}^n), \\ \lambda x &= (\lambda x_1, \ldots, \lambda x_n) \quad (x \in \mathbb{R}^n\ \lambda \in \mathbb{R}). \end{aligned}$$

Mit anderen Worten ist $(\lambda x + \mu y)_i = \lambda x_i + \mu y_i$ für $x, y \in \mathbb{R}^n$, $\lambda, \mu \in \mathbb{R}$. Die Verallgemeinerung des Betrags $|x|$ einer Zahl $x \in \mathbb{R}$ ist die Euklidische Länge oder Norm.

Definition 5.5 *Die Euklidische Norm eines Vektors $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ ist*

$$\|x\| = \sqrt{x_1^2 + \ldots + x_n^2}.$$

Der Euklidische Abstand von zwei Punkten $x, y \in \mathbb{R}^n$ ist $\|x - y\|$.

Das Argument der Wurzel ist als Summe von Quadraten nichtnegativ, also ist $\|x\|$ definiert. Der folgende Satz fasst wesentliche Eigenschaften der Euklidischen Norm zusammen.

Satz 5.4 *Für die Euklidische Norm $\|x\|$ gilt:*

(1) Positivität: $\|x\| \geq 0$ *mit Gleichheit genau wenn* $x = 0$.

(2) Halblinearität: $\|\lambda x\| = |\lambda| \, \|x\|$ *für alle* $\lambda \in \mathbb{R}$, $x \in \mathbb{R}^n$.

(3) Dreiecksungleichung: $\|x + y\| \leq \|x\| + \|y\|$ *für alle* $x, y \in \mathbb{R}^n$.

Gleichheit in (3) gilt genau wenn x und y gleichsinnig parallel sind.

Beweis: Nach Definition der Wurzel ist $\|x\| \geq 0$. Im Gleichheitsfall ist $x_1^2 + \ldots + x_n^2 = 0$, also $x_i = 0$ für alle $i = 1, \ldots, n$ und damit $x = 0$. Die Skalarmultiplikation im $\mathbb{R}^n$ ist komponentenweise definiert durch $\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)$, also folgt

$$\|\lambda x\| = \left(\sum_{i=1}^n (\lambda x_i)^2 \right)^{1/2} = \left(\lambda^2 \sum_{i=1}^n x_i^2 \right)^{1/2} = |\lambda| \left(\sum_{i=1}^n x_i^2 \right)^{1/2} = |\lambda| \, \|x\|.$$

Um (3) zu beweisen, zeigen wir erst die Ungleichung von Cauchy-Schwarz, und dazu brauchen wir das Euklidische Skalarprodukt.

Definition 5.6 *Das Standardskalarprodukt von $x, y \in \mathbb{R}^n$ ist*

$$\langle x, y\rangle = x_1 y_1 + \ldots + x_n y_n.$$

Offenbar gilt $\langle x, x\rangle = \|x\|^2$ für $x \in \mathbb{R}^n$.

Lemma 5.3 *Für das Standardskalarprodukt $\langle x, y\rangle$ gilt:*

(1) Positivität: $\langle x, x\rangle \geq 0$ *mit Gleichheit genau wenn $x = 0$.*

(2) Symmetrie: $\langle x, y\rangle = \langle y, x\rangle$ *für alle $x, y \in \mathbb{R}^n$.*

(3) Bilinearität: *Für alle $x, y, z \in \mathbb{R}^n$ und $\lambda, \mu \in \mathbb{R}$ gilt*

$$\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle \quad \textit{und} \quad \langle x, \lambda y + \mu z\rangle = \lambda\langle x, y\rangle + \mu\langle x, z\rangle.$$

Beweis: Die Positivität folgt aus der Positivität der Norm, und die Symmetrie ist offensichtlich. Auch die Bilinearität ergibt sich leicht aus der Definition:

$$\langle \lambda x + \mu y, z\rangle = \sum_{i=1}^n (\lambda x_i + \mu y_i) z_i = \lambda \sum_{i=1}^n x_i z_i + \mu \sum_{i=1}^n y_i z_i = \lambda\langle x, z\rangle + \mu\langle y, z\rangle.$$

□

Nach diesen Vorbereitungen kommt jetzt die

Satz 5.5 (Ungleichung von Cauchy-Schwarz) *Für $x, y \in \mathbb{R}^n$ gilt*

$$|\langle x, y\rangle| \leq \|x\|\,\|y\|. \tag{5.1}$$

Gleichheit gilt genau dann, wenn x und y linear abhängig sind.

Beweis: Für $a, b \in \mathbb{R}^n$ mit $\|a\| = \|b\| = 1$ gilt $|\langle a, b\rangle| \leq 1$, denn die Bilinearität ergibt

$$1 \pm \langle a, b\rangle = \frac{1}{2}\left(\|a\|^2 + \|b\|^2 \pm 2\langle a, b\rangle\right) = \frac{1}{2}\|a \pm b\|^2 \geq 0.$$

Ist $x = 0$ oder $y = 0$, so gilt $\langle x, y\rangle = \|x\|\,\|y\| = 0$. Sonst betrachten wir die Vektoren $a = x/\|x\|$ sowie $b = y/\|y\|$. Diese haben Länge Eins, denn nach Halblinearität der Norm gilt

$$\left\|\frac{x}{\|x\|}\right\| = \left\|\frac{1}{\|x\|}x\right\| = \frac{1}{\|x\|}\,\|x\| = 1.$$

Anwendung der gezeigten Ungleichung für a, b liefert

$$\frac{1}{\|x\|}\frac{1}{\|y\|}|\langle x, y\rangle| = \left|\frac{1}{\|x\|}\frac{1}{\|y\|}\langle x, y\rangle\right| = \left|\left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|}\right\rangle\right| = |\langle a, b\rangle| \leq 1,$$

was zu zeigen war. Für $x, y \neq 0$ kann Gleichheit kann nur dann eintreten, wenn $a \pm b = 0$, also wenn x und y parallel sind. □

Ersetzen wir in der Ungleichung x durch λx, so multiplizieren sich beide Seiten mit $|\lambda|$. Diese Skalierungseigenschaft legt nahe, im Beweis erst Einheitsvektoren zu betrachten. Wir können nun leicht den Beweis von Satz 5.4 abschließen:

Beweis der Dreiecksungleichung: Mit Satz 5.5 gilt

$$\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

Mit der Monotonie der Wurzel folgt $\|x+y\| \le \|x\| + \|y\|$ wie behauptet. Im Gleichheitsfall müssen zunächst x und y parallel sein, außerdem muss aber $\langle x, y\rangle \ge 0$ sein, also sind x, y sogar gleichsinnig parallel. $\square$

Zur Bezeichnung *Dreiecksungleichung*: für drei Punkte x, y, z folgt aus (3)

$$\|x-z\| = \|x-y+y-z\| \le \|x-y\| + \|y-z\|,$$

das heißt im Dreeick mit den Eckpunkten x, y, z ist jede Seite kürzer als die Summe der beiden anderen Seiten. Im Gleichheitsfall sind $x-y$ und $y-z$ parallel, das heißt die Punkte x, y, z liegen auf einer Geraden, genauer liegt y zwischen x und z.

Im Gegensatz zu $\mathbb{R}$ haben wir im $\mathbb{R}^n$ keine Anordnung zur Verfügung. Deshalb verwenden wir die Ungleichungszeichen $<, >, \le, \ge$, Formulierungen wie "nach oben bzw. unten beschränkt" und den Begriff der Monotonie *ausschließlich* für reelle Zahlen und *niemals* für Punkte im $\mathbb{R}^n$. Eine Ungleichung $a < b$ ist nur für reelle Zahlen sinnvoll (später auch mal für gewisse Matrizen, aber das tut jetzt nichts zur Sache). Ein Check zeigt jedoch, dass die Begriffe Konvergenz, ε-Umgebung, Beschränktheit, Cauchyfolge und Vollständigkeit in $\mathbb{R}$ alle nur die Betragsfunktion verwenden. Sie lassen sich dann auf den $\mathbb{R}^n$ verallgemeinern, indem wir statt der Betragsfunktion die Euklidische Norm einsetzen.

Definition 5.7 (Konvergenz im $\mathbb{R}^n$) *Die Folge $x_k \in \mathbb{R}^n$ konvergiert gegen $a \in \mathbb{R}^n$, falls gilt:*

Für alle $\varepsilon > 0$ gibt es ein $K \in \mathbb{R}$, so dass für alle $k > K$ gilt: $\|x_k - a\| < \varepsilon$.

Definition 5.8 (ε-Umgebung in $\mathbb{R}^n$) *Die ε-Umgebung von $a \in \mathbb{R}^n$ ist die Menge*

$$B_\varepsilon(a) = \{x \in \mathbb{R}^n : \|x-a\| < \varepsilon\}.$$

Sie wird auch als offene Kugel (für $n \ge 3$) bzw. offene Kreisscheibe (für $n = 2$) um a mit Radius $\varepsilon > 0$ bezeichnet. Die abgeschlossene Kugel ist

$$K_\varepsilon(a) = \{x \in \mathbb{R}^n : \|x-a\| \le \varepsilon\}.$$

Definition 5.9 (Beschränkte Teilmengen von $\mathbb{R}^n$) *Eine Menge $M \subset \mathbb{R}^n$ heißt beschränkt, wenn gilt:*

Es gibt ein $K \ge 0$ mit $\|x\| \le K$ für alle $x \in M$.

Zum Glück müssen wir nun nicht die ganze Theorie von vorne beginnen, sondern wir können alles zurückspielen auf die Resultate für $\mathbb{R}$, indem wir die einzelnen Koordinaten der Punkte $x \in \mathbb{R}^n$ betrachten.

Satz 5.6 (Norm versus Koordinaten) *Eine Folge $x_k \in \mathbb{R}^n$ ist genau dann (norm-) konvergent (beschränkt, Cauchyfolge), wenn für alle $i = 1, \ldots, n$ die Koordinatenfolgen $\big((x_k)_i\big)_{k\in\mathbb{N}}$ konvergent (beschränkt, Cauchyfolgen) sind.*

Beweis: Die Implikationen $\Rightarrow$ folgen alle direkt aus der Ungleichung $|x_i| \le \|x\|$. Seien nun die Koordinatenfolgen einzeln beschränkt, also $|(x_k)_i| \le K_i$. Dann gilt

$$\|x_k\| = \left(\sum_{i=1}^{n} (x_k)_i^2\right)^{1/2} \le \left(\sum_{i=1}^{n} K_i^2\right)^{1/2} =: K < \infty.$$

Das Argument für die Konvergenz geht wie folgt: sei $(x_k)_i \to a_i$ für $i = 1, \ldots, n$, das heißt zu $\varepsilon > 0$ gibt es $K_i \in \mathbb{R}$ mit $|(x_k)_i - a_i| < \varepsilon/\sqrt{n}$ für $k > K_i$. Es folgt mit $a := (a_1, \ldots, a_n)$

$$\|x_k - a\| = \left(\sum_{i=1}^{n} ((x_k)_i - a_i)^2\right)^{1/2} < \left(n\,\frac{\varepsilon^2}{n}\right)^{1/2} = \varepsilon \quad \text{für } k > K := \max_{i=1,\ldots,n} K_i.$$

Das Argument für die Cauchyfolgen ist analog: seien die $(x_k)_i$ Cauchyfolgen in $\mathbb{R}$ für jedes $i = 1, \ldots, n$. Zu $\varepsilon > 0$ gibt es dann $K_i \in \mathbb{R}$ mit $|(x_k)_i - (x_l)_i| < \varepsilon/\sqrt{n}$ für alle $k, l > K_i$. Es folgt für $k, l > \max_{i=1,\ldots,n} K_i$

$$\|x_k - x_l\| = \left(\sum_{i=1}^{n} ((x_k)_i - (x_l)_i)^2\right)^{1/2} \le \left(n\,\frac{\varepsilon^2}{n}\right)^{1/2} = \varepsilon.$$

□

Folgerung 5.2 *Mit der Euklidischen Norm ist $\mathbb{R}^n$ vollständig: zu jeder Cauchyfolge $(x_k)_{k\in\mathbb{N}}$ gibt es ein $a \in \mathbb{R}^n$, so dass $x_k \to a$ mit $k \to \infty$.*

Beweis: Nach Satz 5.6 sind die $\big((x_k)_i\big)_{k\in\mathbb{N}}$ Cauchyfolgen in $\mathbb{R}$. Laut Vollständigkeitsaxiom gibt es $a_i \in \mathbb{R}$ mit $\lim_{k\to\infty} (x_k)_i = a_i$ für alle $i \in \{1, \ldots, n\}$. Setze $a = (a_1, \ldots, a_n)$. Nach Satz 5.6 folgt $\lim_{k\to\infty} x_k = a$. □

Folgerung 5.3 (Bolzano-Weierstraß im $\mathbb{R}^n$) *Jede beschränkte Folge $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n$ besitzt eine konvergente Teilfolge.*

Beweis: Die Koordinatenfolgen $((x_k)_i)_{k\in\mathbb{N}}$ sind beschränkt, siehe Satz 5.6, haben also jeweils eine konvergente Teilfolge nach Satz 4.7. Das Problem ist, dass diese Teilfolgen a priori verschieden sind. Wir konstruieren sukzessive Teilfolgen, damit schließlich alle n Koordinaten konvergieren. Angenommen, für ein $j \in \{1, \ldots, n\}$ sei schon eine Teilfolge $k_1 < k_2 < \ldots$ gefunden, so dass

$$\lim_{l\to\infty} (x_{k_l})_i = a_i \quad \text{für alle } i = 1, \ldots, j-1.$$

Der Fall $j = 1$ ist dabei der Induktionsanfang. Es gibt nun eine weitere Teilfolge $l_1 < l_2 < \ldots$ und ein $a_j \in \mathbb{R}$, so dass zusätzlich

$$\lim_{m\to\infty} (x_{k_{l_m}})_j = a_j.$$

Nach n Schritten ist eine Teilfolge bestimmt, für die alle Koordinatenfolgen konvergieren. □

Definition 5.10 (Häufungspunkt von Mengen) *Ein Punkt $a \in \mathbb{R}^n$ heißt Häufungspunkt der Menge $M \subset \mathbb{R}^n$, falls für jedes $\varepsilon > 0$ die Menge $B_\varepsilon(a) \cap M$ unendlich viele Elemente hat.*

Ein Häufungspunkt von M ist nicht notwendig Element von M, zum Beispiel ist $0 \in \mathbb{R}$ Häufungspunkt von $M = \{1/n : n \in \mathbb{N}\}$.

Lemma 5.4 *Ein Punkt $a \in \mathbb{R}^n$ ist genau dann Häufungspunkt der Menge M, wenn es eine Folge von Punkten $x_k \in M\backslash\{a\}$ gibt mit $\lim_{k\to\infty} x_k = a$.*

Beweis: Das ist ähnlich zu Lemma 4.1. Sei a ein Häufungspunkt gemäß Definition 5.10. Zu $\varepsilon_k = 1/k$ gibt es dann ein $x_k \in B_{1/k}(a) \cap M\backslash\{a\}$. Die Folge $(x_k)_{k\in\mathbb{N}}$ ist wie verlangt. Sei umgekehrt eine Folge $x_k \in M\backslash\{a\}$ gegeben mit $\lim_{k\to\infty} x_k = a$. Angenommen, für ein $\varepsilon > 0$ ist die Menge $B_\varepsilon(a) \cap M$ endlich, also $B_\varepsilon(a) \cap M\backslash\{a\} = \{a_1, \ldots, a_m\}$ für ein $m \in \mathbb{N}_0$. Dann ist $\varrho := \min_{1\leq i\leq m} |a_i - a| > 0$. Wegen $\lim_{k\to\infty} \|x_k - a\| = 0$ folgt für k hinreichend groß

$$x_k \in B_\varrho(a) \cap M\backslash\{a\} = \emptyset,$$

ein Widerspruch. □

Beispiel 5.2 Die Menge $M = \{1/n : n \in \mathbb{N}\}$ hat $0 \in \mathbb{R}$ als einzigen Häufungspunkt. Dagegen ist die Menge der Häufungspunkte von $\mathbb{Q}$ gleich $\mathbb{R}$, denn in jeder Umgebung einer reellen Zahl gibt es eine rationale Zahl, siehe Satz 3.2. Es gibt eine kleine Differenz zwischen dem Begriff des Häufungspunkts für Folgen und für Mengen: die konstante Folge $x_k = a$, $k \in \mathbb{N}$, hat den Häufungspunkt a, dagegen hat die einelementige Menge $\{a\} \subset \mathbb{R}^n$ keinen Häufungspunkt. Überlegen Sie, dass für eine beliebige Folge $(x_k)_{k\in\mathbb{N}}$ im $\mathbb{R}^n$ die Häufungspunkte des Wertebereichs $M = \{x_k : k \in \mathbb{N}\}$ eine Teilmenge der Häufungspunkte der Folge bilden, die aber eine echte Teilmenge sein kann, wie das Beispiel der konstanten Folge zeigt.

Definition 5.11 *Eine Menge*

(i) *$U \subset \mathbb{R}^n$ heißt offen, falls es zu jedem $a \in U$ ein $\varepsilon > 0$ gibt mit $B_\varepsilon(a) \subset U$,*

(ii) *$A \subset \mathbb{R}^n$ heißt abgeschlossen, wenn folgende Implikation stets gilt:*

$$a = \lim_{n\to\infty} x_n \text{ mit } x_n \in A \quad \Rightarrow \quad a \in A.$$

Nach Lemma 5.4 ist eine Menge A genau dann abgeschlossen, wenn jeder Häufungspunkt von A schon Element von A ist.

Beispiel 5.3 Die Kugel $B_\varrho(a) = \{x \in \mathbb{R}^n : \|x - a\| < \varrho\}$ ist offen. Denn sei $x \in B_\varrho(a)$. Für $\varepsilon := \varrho - \|x - a\| > 0$ und $y \in B_\varepsilon(x)$ folgt aus der Dreiecksungleichung

$$\|y - a\| \leq \|y - x\| + \|x - a\| < \varepsilon + \|x - a\| = \varrho.$$

Also gilt $B_\varepsilon(x) \subset B_\varrho(a)$, das heißt $B_\varrho(a)$ ist offen. Die Kugel $K_\varrho(a) = \{x \in \mathbb{R}^n : \|x - a\| \leq \varrho\}$ ist dagegen abgeschlossen: sind $x_k \in K_\varrho(a)$ und $x_k \to x \in \mathbb{R}^n$, so folgt

$$\|x - a\| \leq \|x - x_k\| + \|x_k - a\| \leq \|x - x_k\| + \varrho \to \varrho,$$

das heißt $x \in K_\varrho(a)$.

Satz 5.7 *Eine Menge $M \subset \mathbb{R}^n$ ist genau dann offen, wenn $\mathbb{R}^n\backslash M$ abgeschlossen ist.*

BEWEIS: Sei M offen und $x_k \in \mathbb{R}^n \backslash M$ mit $\lim_{k\to\infty} x_k = a \in \mathbb{R}^n$. Wäre $a \in M$, so folgt $B_\varepsilon(a) \subset M$ für ein $\varepsilon > 0$, und somit $x_k \in M$ für hinreichend große k, ein Widerspruch. Also ist $a \in \mathbb{R}^n \backslash M$, das heißt $\mathbb{R}^n \backslash M$ ist abgeschlossen.

Sei nun $\mathbb{R}^n \backslash M$ abgeschlossen. Wäre M nicht offen, so gibt es ein $a \in M$ mit $B_\varepsilon(a) \backslash M \neq \emptyset$ für alle $\varepsilon > 0$. Finde induktiv $x_k \in B_{1/k}(a)$ mit $x_k \in \mathbb{R}^n \backslash M$. Es folgt $a = \lim_{k\to\infty} x_k \in \mathbb{R}^n \backslash M$, ein Widerspruch. $\square$

Satz 5.8 *Die offenen Teilmengen des $\mathbb{R}^n$ bilden eine Topologie, das heißt es gilt:*

(1) *$\emptyset$ und $\mathbb{R}^n$ sind offen.*

(2) *Jede Vereinigung $\bigcup_{\lambda\in\Lambda} U_\lambda$ von offenen Mengen $U_\lambda \subset \mathbb{R}^n$ ist offen.*

(3) *Der Durchschnitt $\bigcap_{i=1}^k U_i$ von endlich vielen offenen Mengen $U_i \subset \mathbb{R}^n$ ist offen.*

BEWEIS: Aussage (1) ist evident. Ist $x \in \bigcup_{\lambda\in\Lambda} U_\lambda$, so gilt $x \in U_\mu$ für ein $\mu \in \Lambda$. Da U_μ offen, gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset U_\mu \subset \bigcup_{\lambda\in\Lambda} U_\lambda$, was zu zeigen war.

Sei nun $x \in \bigcap_{i=1}^k U_i$ wie in (3). Da U_i offen, gibt es ein $\varepsilon_i > 0$ mit $B_{\varepsilon_i}(x) \subset U_i$ für $i = 1, \ldots, k$. Es folgt $\varepsilon := \min_{i=1,\ldots,k} \varepsilon_i > 0$ und $B_\varepsilon(x) \subset B_{\varepsilon_i}(x) \subset U_i$ für alle $i = 1, \ldots, k$, also $B_\varepsilon(x) \subset \bigcap_{i=1}^k U_i$. $\square$

Wir betonen, dass im Gegensatz zur Vereinigung der Durchschnitt von unendlich vielen offenen Mengen in der Regel nicht offen ist, zum Beispiel gilt

$$\bigcap_{n=1}^{\infty} B_{1/n}(a) = \{a\}.$$

Eine Menge $M \subset \mathbb{R}^n$ muss weder offen noch abgeschlossen sein; dies trifft zum Beispiel auf ein halboffenes Intervall $[a, b)$ zu. Die Mengen $\mathbb{R}$ und $\emptyset$ sind die einzigen Teilmengen von $\mathbb{R}$, die sowohl offen als auch abgeschlossen sind. Dies soll in den Übungen gezeigt werden. Durch Übergang zu den Komplementen sieht man, dass beliebige Durchschnitte und endliche Vereinigungen von abgeschlossenen Mengen abermals abgeschlossen sind.

Für uns hat der Fall $n = 2$ der Euklidischen Ebene eine besondere Bedeutung, denn aus dem Vektorraum $(\mathbb{R}^2, +)$ wird mit einer geeigneten Multiplikation der Körper $\mathbb{C} = (\mathbb{R}^2, +, \cdot)$ der komplexen Zahlen. Die Verknüpfungen sind dabei wie folgt definiert:

Die Addition ist die Vektorraumaddition von $\mathbb{R}^2$. Die Standardbasis wird mit $(1, 0) = 1$ und $(0, 1) = i$ bezeichnet, das heißt jedes $z = (x, y) \in \mathbb{R}^2$ besitzt die Basisdarstellung $z = x + iy$. Damit lautet die Addition von $x_k + iy_k$, $k = 1, 2$,

$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) = (x_1 + x_2) + i(y_1 + y_2).$$

Mit diesen Vereinbarungen gilt $x = x + i0 = (x, 0)$ für $x \in \mathbb{R}$, das heißt $\mathbb{R}$ wird mit der x-Achse in $\mathbb{R}^2$ identifiziert. Für $z = x + iy$ heißt $x =: \operatorname{Re} z \in \mathbb{R}$ der Realteil und $y = \operatorname{Im} z \in \mathbb{R}$ der Imaginärteil von z.

Die Multiplikation ergibt sich durch die Forderung $i^2 = -1$ und Ausmultiplizieren nach den Körpergesetzen:

$$(x_1 + iy_1)(x_2 + iy_2) = x_1x_2 + i^2y_1y_2 + i(x_1y_2 + x_2y_1) = x_1x_2 - y_1y_2 + i(x_1y_2 + x_2y_1).$$

Für $\lambda \in \mathbb{R}$ ist $\lambda(x+iy) = (\lambda + i0)(x+iy) = \lambda x + i\lambda y = (\lambda x, \lambda y) = \lambda(x,y)$, das heißt Multiplikation mit $\lambda \in \mathbb{R} \subset \mathbb{C}$ ist einfach Skalarmultiplikation im $\mathbb{R}$-Vektorraum $\mathbb{R}^2$. Dagegen liefert die Multiplikation mit i für $(x,y) = x+iy \in \mathbb{C}$

$$i(x,y) = i(x+iy) = -y + ix = (-y, x).$$

Der Vektor $(-y,x)$ entsteht anschaulich aus (x,y) durch Drehung um $90°$ im mathematisch positiven Sinn, das heißt gegen den Uhrzeigersinn.

Die Körpergesetze in $\mathbb{C}$ folgen leicht aus den Definitionen, nur die Bestätigung des Assoziativgesetzes der Multiplikation erfordert etwas Rechenarbeit. Um das inverse Element der Multiplikation anzugeben, ist ein weiterer Begriff nützlich: für $z = x+iy \in \mathbb{C}$ heißt $\overline{z} = x - iy \in \mathbb{C}$ die zu z konjugiert komplexe Zahl. Anschaulich ergibt sich $\overline{z}$ aus z durch Spiegelung an der x-Achse, insbesondere gilt $\overline{\overline{z}} = \overline{x-iy} = x+iy = z$.

Lemma 5.5 *Für die komplexe Konjugation gelten folgende Regeln:*

(1) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$, $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$,

(2) $\overline{z} = z \quad \Leftrightarrow \quad z \in \mathbb{R}$,

(3 *Für* $z = x+iy$ *ist* $\operatorname{Re} z = \frac{1}{2}(z+\overline{z})$ *und* $\operatorname{Im} z = \frac{1}{2i}(z - \overline{z})$.

Beweis: Die Beweise erfolgen alle durch Nachrechnen, zum Beispiel gilt

$$\overline{z_1}\,\overline{z_2} = (x_1 - iy_1)(x_2 - iy_2) = x_1x_2 - y_1y_2 - i(x_1y_2 + y_1x_2) = \overline{z_1 z_2}.$$

□

Man nennt die Euklidische Norm

$$|z| = \sqrt{x^2+y^2} \quad \text{für } z = x+iy,$$

auch den Betrag der komplexen Zahl z (Notation ohne Doppelstriche ist üblich).

Lemma 5.6 *Für den Betrag einer komplexen Zahl gelten folgende Regeln:*

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

(2) $|z_1 z_2| = |z_1||z_2|$.

(3) $|\operatorname{Re} z| \leq |z|$, $|\operatorname{Im}(z)| \leq |z|$.

Beweis: Für Aussage (1) berechnen wir mit $z = x+iy$

$$z\overline{z} = (x+iy)(x-iy) = x^2+y^2 = |z|^2.$$

Die Formel (2) folgt aus (1), Lemma 5.5(1) und den Körpergesetzen:

$$|z_1z_2|^2 = (z_1z_2)(\overline{z_1z_2}) = z_1\,\overline{z_1}\,z_2\overline{z_2} = |z_1|^2\,|z_2|^2.$$

Die Ungleichungen (3) sind offensichtlich. □

Damit können wir nun das inverse Element der Multiplikation leicht hinschreiben, und zwar folgt aus Lemma 5.6(1) für $z = x + iy \neq 0$

$$\frac{1}{z} = \frac{\overline{z}}{z\,\overline{z}} = \frac{x - iy}{x^2 + y^2}. \tag{5.2}$$

Wir wollen jetzt einige Grundtatsachen über Nullstellen von Polynomen erarbeiten. Im folgenden sei stets $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Definition 5.12 *Eine Funktion $p : \mathbb{K} \to \mathbb{K}$ heißt (reelles oder komplexes) Polynom vom Grad $n \in \mathbb{N}_0$, wenn es $a_0, a_1, \ldots, a_n \in \mathbb{K}$ mit $a_n \neq 0$ gibt, so dass gilt:*

$$p(z) = \sum_{k=0}^{n} a_k z^k \quad \text{für alle } z \in \mathbb{K}. \tag{5.3}$$

Aus der Definition ist nicht unmittelbar ersichtlich, daß der Grad n und die Koeffizienten a_i eindeutig durch die Funktion $p : \mathbb{K} \to \mathbb{K}$ bestimmt sind. Dies wird in Satz 5.10 gezeigt.

Satz 5.9 (Abspaltung von Linearfaktoren) *Sei $p : \mathbb{K} \to \mathbb{K}$ Polynom vom Grad $n \in \mathbb{N}$ mit Koeffizienten $a_0, \ldots, a_n$. Ist $p(\lambda) = 0$ für ein $\lambda \in \mathbb{K}$, so gibt es ein Polynom $q : \mathbb{K} \to \mathbb{K}$ vom Grad $n-1$ mit Koeffizienten $b_0, \ldots, b_{n-1}$, wobei $b_{n-1} = a_n$, so dass gilt:*

$$p(z) = (z - \lambda)\, q(z) \quad \text{für alle } z \in \mathbb{K}. \tag{5.4}$$

Beweis: Für $z, \lambda \in \mathbb{K}$ und $k \in \mathbb{N}$ gilt, vgl. Beispiel 2.2 zur geometrischen Summe,

$$z^k - \lambda^k = \sum_{j=1}^{k} z^j \lambda^{k-j} - \sum_{j=0}^{k-1} z^j \lambda^{k-j} = \sum_{j=0}^{k-1} (z^{j+1}\lambda^{k-j-1} - z^j \lambda^{k-j}) = (z - \lambda) \sum_{j=0}^{k-1} z^j \lambda^{k-j-1}.$$

Ist $p(\lambda) = 0$ so folgt

$$p(z) = p(z) - p(\lambda) = \sum_{k=1}^{n} a_k (z^k - \lambda^k) = (z - \lambda) \sum_{k=1}^{n} \sum_{j=0}^{k-1} a_k z^j \lambda^{k-j-1} = (z - \lambda) \sum_{j=0}^{n-1} b_j z^j,$$

wobei $b_j = \sum_{k=j+1}^{n} a_k \lambda^{k-j-1}$. Insbesondere $b_{n-1} = a_n$. □

Folgerung 5.4 *Sei $p : \mathbb{K} \to \mathbb{K}$ ein Polynom vom Grad $n \in \mathbb{N}_0$. Dann hat p höchstens n verschiedene Nullstellen.*

Beweis: Wir führen Induktion über $n \in \mathbb{N}_0$. Für $n = 0$ gilt $p(z) = a_0$ für alle $z \in \mathbb{K}$, wobei $a_0 \neq 0$, also hat p keine Nullstelle. Ist p Polynom vom Grad $n \in \mathbb{N}$, so hat entweder p keine Nullstelle oder es gilt nach Satz 5.9 $p(z) = (z - \lambda) q(z)$ für alle $z \in \mathbb{K}$, mit einem Polynom q vom Grad $n - 1$. Nach Induktion hat q höchstens $n - 1$ Nullstellen, also p höchstens n Nullstellen. □

Satz 5.10 (Koeffizientenvergleich) *Seien $p, q : \mathbb{K} \to \mathbb{K}$ Polynome vom Grad m bzw. n, das heißt es gilt mit $a_m, b_n \neq 0$*

$$p(z) = \sum_{i=0}^{m} a_i z^i \quad \textit{und} \quad q(z) = \sum_{i=0}^{n} b_i z^i \quad \textit{für alle } z \in \mathbb{K}.$$

Ist $p(\lambda_j) = q(\lambda_j)$ für paarweise verschiedene $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ mit $k > \max(n, m)$, so folgt $m = n$ und $a_i = b_i$ für $i = 0, \ldots, n$.

Beweis: Wäre $m \neq n$ oder $a_i \neq b_i$ für ein i, so wäre $p - q$ Polynom vom Grad höchstens $\max(n, m)$ mit $k > \max(n, m)$ Nullstellen, im Widerspruch zu Folgerung 5.4. □

Lemma 5.7 *Sei $p : \mathbb{K} \to \mathbb{K}$ Polynom vom Grad n. Dann besitzt p eine eindeutig bestimmte Zerlegung*

$$p(z) = (z - \lambda_1)^{\nu_1} \cdot \ldots \cdot (z - \lambda_r)^{\nu_r} q(z) \quad \textit{für alle } z \in \mathbb{K}. \tag{5.5}$$

Dabei sind $\{\lambda_1, \ldots, \lambda_r\} \subset \mathbb{K}$ die Nullstellen von p in $\mathbb{K}$ (eventuell $r = 0$), und $q : \mathbb{K} \to \mathbb{K}$ ist ein Polynom vom Grad $n - (\nu_1 + \ldots + \nu_r) \in \{0, \ldots, n\}$ mit $q(z) \neq 0$ für alle $z \in \mathbb{K}$. Die Zahlen $\nu_i \geq 1$ heißen Vielfachheiten der Nullstellen.

Beweis: Die Existenz der Zerlegung folgt induktiv aus Satz 5.9. Für die Eindeutigkeit betrachte eine zweite solche Zerlegung $p(z) = (z - \lambda_1)^{\mu_1} \cdot \ldots \cdot (x - \lambda_r)^{\mu_r} h(z)$, ohne Einschränkung mit $\nu_1 \geq \mu_1$. Indem wir durch $(z - \lambda_1)^{\mu_1}$ teilen, folgt für alle $z \neq \lambda_1$

$$(z - \lambda_1)^{\nu_1 - \mu_1} \cdot (z - \lambda_2)^{\nu_2} \cdot \ldots \cdot (z - \lambda_r)^{\nu_r} \cdot q(z) = (z - \lambda_2)^{\mu_2} \cdot \ldots \cdot (z - \lambda_r)^{\mu_r} \cdot h(z).$$

Beide Seiten sind Polynome, haben nach Satz 5.10 also dieselben Koeffizienten. Deshalb stimmen sie auch in $z = \lambda_1$ überein; es folgt $\nu_1 = \mu_1$ und analog $\nu_i = \mu_i$ für alle $i \in \{1, \ldots, r\}$. Nach Division haben wir schließlich $q(z) = h(z)$ für alle $z \in \mathbb{K} \backslash \{\lambda_1, \ldots, \lambda_r\}$, und Satz 5.10 liefert $q = h$. □

Das vorangegangene Lemma sagt natürlich gar nichts aus, wenn das gegebene Polynom keine Nullstelle hat, wie zum Beispiel $p : \mathbb{R} \to \mathbb{R}$, $p(x) = 1 + x^2$. Im Gegensatz zur Situation für $\mathbb{K} = \mathbb{R}$ hat im Fall $\mathbb{K} = \mathbb{C}$ jedes Polynom eine Nullstelle. Für Polynome vom Grad $n = 2$ kann dies leicht mit quadratischer Ergänzung gezeigt werden, für Polynome höheren Grades besagt dies der

Satz 5.11 (Fundamentalsatz der Algebra) *Jedes komplexe Polynom vom Grad $n \geq 1$ hat mindestens eine Nullstelle $\lambda \in \mathbb{C}$.*

Diesen Satz werden wir in Analysis II beweisen. Als Konsequenz ergibt sich jedenfalls

Folgerung 5.5 *Jedes Polynom $p : \mathbb{C} \to \mathbb{C}$ vom Grad n mit Koeffizienten $a_0, a_1, \ldots, a_n \in \mathbb{C}$ zerfällt über $\mathbb{C}$ in Linearfaktoren, das heißt es gilt*

$$p(z) = a_n (z - \lambda_1)^{\nu_1} \cdot \ldots \cdot (z - \lambda_k)^{\nu_k} \quad \textit{für alle } z \in \mathbb{C}. \tag{5.6}$$

Dabei sind $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ die Nullstellen von p, und die $\nu_i \geq 1$ sind die Vielfachheiten dieser Nullstellen mit $\nu_1 + \ldots + \nu_k = n$.

Beweis: Das Restpolynom q in (5.5) hat keine Nullstelle in $\mathbb{C}$, also hat q nach dem Fundamentalsatz der Algebra den Grad $n = 0$, das heißt $q(z) = a_n$ für alle $z \in \mathbb{C}$. $\square$

Mithilfe des Fundamentalsatzes der Algebra können wir auch die Situation in $\mathbb{R}$ analysieren. Jedes reelle Polynom $p(x) = a_0 + a_1 x + \ldots + a_n x^n$, das heißt $a_i \in \mathbb{R}$, kann nämlich als komplexes Polynom aufgefasst werden, indem wir für x auch komplexe Zahlen z einsetzen. Damit wird die reelle Funktion $p : \mathbb{R} \to \mathbb{R}$ zu einer komplexen Funktion $p : \mathbb{C} \to \mathbb{C}$ fortgesetzt.

Lemma 5.8 *Sei $p(x) = a_0 + a_1 x + \ldots + a_n x^n$ mit $a_0, a_1, \ldots, a_n \in \mathbb{R}$ ein reelles Polynom vom Grad n. Dann besitzt p, aufgefasst als komplexes Polynom, eine Faktorisierung der Form*

$$p(z) = a_n \prod_{i=1}^{r} (z - \alpha_i)^{\mu_i} \prod_{j=1}^{k} (z - \lambda_j)^{\nu_j} (z - \overline{\lambda_j})^{\nu_j} \quad \text{für alle } z \in \mathbb{C}.$$

Dabei sind $\{\alpha_1, \ldots, \alpha_r\} \subset \mathbb{R}$ die reellen Nullstellen von p mit Vielfachheiten $\mu_i \geq 1$, und $\{\lambda_1, \overline{\lambda_1}, \ldots, \lambda_k, \overline{\lambda_k}\} \subset \mathbb{C} \backslash \mathbb{R}$ die nichtreellen Nullstellen. Diese treten also in Paaren $\lambda_j, \overline{\lambda_j}$ mit gleicher Vielfachheit ν_j auf. Es gilt $\mu_1 + \ldots + \mu_r + 2(\nu_1 + \ldots + \nu_k) = n$.

Beweis: Es gilt für $z \in \mathbb{C}$, da $a_i \in \mathbb{R}$, wegen Lemma 5.5(i)

$$\overline{p(z)} = \overline{a_0 + a_1 z + \ldots + a_n z^n} = a_0 + a_1 \overline{z} + \ldots + a_n \overline{z}^n = p(\overline{z}).$$

Insbesondere folgt aus $p(z) = 0$, dass auch $p(\overline{z}) = 0$. Ist daher $\lambda \in \mathbb{C} \backslash \mathbb{R}$ eine nichtreelle Nullstelle, so erhalten wir nach Satz 5.9 die Zerlegung

$$p(z) = (z - \lambda)(z - \overline{\lambda})\, q(z) = (z^2 - 2(\operatorname{Re}\lambda) z + |\lambda|^2)\, q(z), \tag{5.7}$$

mit $q(z) = b_0 + b_1 z + \ldots + b_{n-2} z^{n-2}$. Beachten Sie, dass wir Satz 5.9 hier über $\mathbb{C}$ anwenden, das heißt die b_i sind zunächst komplexe Zahlen. Um zu sehen, dass die b_i doch reell sind, setzen wir $z := x \in \mathbb{R}$ in (5.7) ein und bilden die Imaginärteile beider Seiten:

$$\begin{aligned}
0 &= \operatorname{Im} p(x) \quad \text{(da } p \text{ reelle Koeffizienten hat)} \\
&= \operatorname{Im} \Big((x^2 - 2(\operatorname{Re}\lambda) x + |\lambda|^2)\, q(x) \Big) \\
&= (x^2 - 2(\operatorname{Re}\lambda) x + |\lambda|^2) \operatorname{Im} q(x) \\
&= (x^2 - 2(\operatorname{Re}\lambda) x + |\lambda|^2) \sum_{i=0}^{n-2} (\operatorname{Im} b_i) x^i.
\end{aligned}$$

Da der linke Faktor in der letzten Zeile für $x \in \mathbb{R}$ niemals Null ist, folgt

$$\sum_{i=0}^{n-2} (\operatorname{Im} b_i) x^i = 0 \quad \text{für alle } x \in \mathbb{R}.$$

Nach Satz 5.10 muss $\operatorname{Im} b_i = 0$ für alle i gelten, das heißt $b_i \in \mathbb{R}$. Das Polynom q ist also wieder ein reelles Polynom, und durch Induktion erhalten wir nun die gewünschte Zerlegung. $\square$

Folgerung 5.6 *Jedes Polynom $p : \mathbb{R} \to \mathbb{R}$ vom Grad n mit Koeffizienten $a_0, a_1, \ldots, a_n \in \mathbb{R}$ zerfällt über $\mathbb{R}$ in lineare und quadratische Faktoren. Genauer gilt*

$$p(x) = a_n \prod_{i=1}^{r} (x - \alpha_i)^{\mu_i} \prod_{j=1}^{k} (x^2 - 2(\operatorname{Re}\lambda_j)x + |\lambda_j|^2)^{\nu_j} \quad \text{für alle } x \in \mathbb{R}, \tag{5.8}$$

wobei die $\alpha_i \in \mathbb{R}$ die reellen und die $\lambda_j, \overline{\lambda_j} \subset \mathbb{C}\backslash\mathbb{R}$ die nichtreellen Nullstellen von p sind.

Komplexe Zahlen traten zuerst in der italienischen Renaissance in der ersten Hälfte des 16. Jahrhunderts auf, bei der Lösung von quadratischen und kubischen Gleichungen [3]. Besonders interessant war der Fall kubischer Polynome mit reellen Nullstellen, für die der Lösungsweg aber ins Komplexe führt (*casus irreducibilis*). Die Bezeichnung $i = \sqrt{-1}$ stammt von Euler (1777), die Bezeichnung *komplexe Zahl* von Gauß (1831). Der Fundamentalsatz der Algebra ist zuerst von Gauß in seiner Dissertation (Helmstedt 1799) bewiesen worden; wir werden den Satz im zweiten Semester zeigen. Trotz des Namens erfordert der Beweis Methoden der zweidimensionalen Analysis oder Topologie. Für Gleichungen fünften oder höheren Grades führen algebraische Auflösungsverfahren nämlich aus systematischen Gründen nicht zum Ziel, wie der norwegische Mathematiker Niels Henrik Abel 1825 erkannte.

[3] Hellmuth Gericke: Mathematik im Abendland, pp. 236-241, 3. Auflage, Fourier Verlag Wiesbaden 1994

6 Reihen

Viele Funktionen in der Analysis können als unendliche Reihen definiert beziehungsweise dargestellt werden. Zum Beispiel leiten wir in Kapitel 13 solche Darstellungen für die Exponentialfunktion und die trigonometrischen Funktionen her. Darum ist es wichtig, die Konvergenzfrage für Reihen zu untersuchen.

Definition 6.1 *Eine Folge $(S_n)_{n\in\mathbb{N}_0}$ heißt Reihe mit Gliedern $a_n \in \mathbb{C}$, falls gilt:*

$$S_n = \sum_{k=0}^{n} a_k \quad \text{für alle } n \in \mathbb{N}_0.$$

Die Reihe heißt konvergent mit Wert $S \in \mathbb{C}$, wenn die Folge $(S_n)_{n\in\mathbb{N}_0}$ gegen S konvergiert:

$$\sum_{k=0}^{\infty} a_k := \lim_{n\to\infty} \sum_{k=0}^{n} a_n = \lim_{n\to\infty} S_n = S.$$

Die Zahl S_n wird auch als n-te Partialsumme der Reihe bezeichnet. Oft wird die Reihe in der Form $a_0 + a_1 + a_2 + \ldots$ durch ihre ersten Glieder angegeben. Leider ist es auch üblich, die Reihe selbst - unabhängig von der Frage der Konvergenz - ebenfalls mit dem Symbol $\sum_{k=0}^{\infty} a_k$ zu bezeichnen.

Beispiel 6.1 Die Reihe $\dfrac{1}{1\cdot 2} + \dfrac{1}{2\cdot 3} + \dfrac{1}{3\cdot 4} + \ldots$ wird auch mit $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$ bezeichnet. Gemeint ist jeweils die Folge $(S_n)_{n\in\mathbb{N}}$ mit den Gliedern

$$S_1 = \frac{1}{1\cdot 2} = \frac{1}{2}, \quad S_2 = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} = \frac{2}{3}, \quad S_3 = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} = \frac{3}{4}, \ldots$$

Für die betrachtete Reihe gilt, für alle $n \in \mathbb{N}$,

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

Also ist die Reihe konvergent mit Wert

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

Die Regeln für die Addition von konvergenten Reihen sowie die Multiplikation von konvergenten Reihen mit reellen oder komplexen Zahlen ergeben sich direkt aus den entsprechenden Regeln für konvergente Folgen, siehe Satz 3.5 a) in Kapitel 1. So ist für konvergente Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ und $\lambda, \mu \in \mathbb{R}$ auch die Reihe $\sum_{k=0}^{\infty} (\lambda a_k + \mu b_k)$ konvergent und es gilt

$$\sum_{k=0}^{\infty} (\lambda a_k + \mu b_k) = \lambda \sum_{k=0}^{\infty} a_k + \mu \sum_{k=0}^{\infty} b_k.$$

Durch Abänderung, Hinzufügen oder Weglassen von endlich vielen Gliedern in einer Reihe $\sum_{k=0}^{\infty} a_k$ wird die Konvergenz oder Divergenz nicht beeinflusst, sondern nur der Wert der Reihe. Zum Beispiel haben wir für $n \in \mathbb{N}_0$ und $m \geq n$

$$\sum_{k=n}^{m} a_k = \sum_{k=0}^{m} a_k - \sum_{k=0}^{n-1} a_k,$$

und mit $m \to \infty$ folgt, falls die Reihe $\sum_{k=0}^{\infty} a_k$ konvergiert,

$$\sum_{k=n}^{\infty} a_k = \sum_{k=0}^{\infty} a_k - \sum_{k=0}^{n-1} a_k. \tag{6.1}$$

Wie bei Folgen kann der Anfangsindex einer Reihe statt $k = 0$ auch eine andere Zahl sein, zum Beispiel $k = 1$ wie in Beispiel 6.1 Allgemein ist eine Reihe nichts anderes als eine Folge $(S_n)_{n \in \mathbb{N}_0}$, die gegeben ist durch die Differenzen $a_n = S_{n+1} - S_n$ und das erste Folgenglied $S_0 = a_0$. Es stellen sich zwei Fragen:

- Wie kann ich den Gliedern a_n ansehen, ob die Reihe konvergiert bzw. divergiert?
- Im Fall der Konvergenz: welchen Wert hat die Reihe?

Bei der zweiten Frage ist zum Beispiel gemeint, ob eine bereits definierte Zahl wie $\sqrt{2}$, e, π, ... als Grenzwert einer Reihe dargestellt werden kann. Im Folgenden steht aber die erste Frage im Zentrum des Interesses. Dabei ist das nächste Beispiel fundamental.

Beispiel 6.2 (Geometrische Reihe) Die geometrische Reihe $\sum_{k=0}^{\infty} z^k$ mit $z \in \mathbb{C}$ konvergiert genau für $|z| < 1$. Der Konvergenzbeweis ist wie in Beispiel 3.7, und zwar gilt

$$\sum_{k=0}^{n} z^k = \frac{1 - z^{n+1}}{1 - z} \to \frac{1}{1 - z} \qquad \text{mit } n \to \infty.$$

Dagegen gilt für $|z| \geq 1$ mit $S_n = \sum_{k=0}^{n} z^k$

$$|S_{n+1} - S_n| = |z^{n+1}| = |z|^{n+1} \geq 1,$$

so dass $(S_n)_{n \in \mathbb{N}}$ keine Cauchyfolge sein kann.

Beispiel 6.3 (Unendliche Dezimalbrüche) Ist $(k_j)_{j \in \mathbb{N}_0}$ eine Folge von Ziffern $k_j \in \{0, 1, \ldots, 9\}$, so ist die Reihe $\sum_{j=0}^{\infty} k_j \, 10^{-j}$ konvergent, vgl. Beispiel 4.4.

Beispiel 6.4 (Harmonische Reihe) Die harmonische Reihe $\sum_{k=1}^{\infty} 1/k$ ist bestimmt divergent gegen $+\infty$. Dies zeigen wir, indem wir in den Partialsummen wie folgt Klammern setzen:

$$\underbrace{\left(\frac{1}{1}\right)}_{\geq 1/2} + \underbrace{\left(\frac{1}{2} + \frac{1}{3}\right)}_{\geq 1/2} + \underbrace{\left(\frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}\right)}_{\geq 1/2} + \underbrace{\left(\frac{1}{8} + \frac{1}{9} + \ldots + \frac{1}{15}\right)}_{\geq 1/2} + \ldots$$

Die Summe der $1/k$ mit $2^m \leq k < 2^{m+1}$ ist nach unten abgeschätzt durch $2^m \cdot 2^{-(m+1)} = 1/2$.

Nach diesen ersten Beispielen kommen nun zur allgemeinen Konvergenzfrage für Reihen, und beginnen mit einem notwendigen Kriterium.

Satz 6.1 (Nullfolgentest) *Ist die Reihe $\sum_{k=0}^{\infty} a_k$ konvergent, so folgt $\lim_{k\to\infty} a_k = 0$.*

Beweis: Nach Voraussetzung ist die Folge der Partialsummen $S_n = \sum_{k=0}^{n} a_k$ konvergent mit Grenzwert $S \in \mathbb{C}$, also folgt $a_n = S_n - S_{n-1} \to S - S = 0$. □

Als nächstes formulieren wir unsere Konvergenzkriterien für Folgen neu in der Situation von Reihen. Die Cauchyfolgeneigenschaft sieht wie folgt aus.

Satz 6.2 (Konvergenzkriterium von Cauchy) *Eine Reihe $\sum_{k=0}^{\infty} a_k$ konvergiert dann und nur dann, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{R}$ gibt, so dass gilt:*

$$\left|\sum_{k=n}^{m} a_k\right| < \varepsilon \qquad \text{für alle } n, m \in \mathbb{N}_0 \text{ mit } m \geq n > N.$$

Beweis: Satz 4.2. □

Satz 6.3 (Reihen mit Gliedern $a_k \geq 0$) *Eine reelle Reihe $\sum_{k=0}^{\infty} a_k$ mit $a_k \geq 0$ für alle k konvergiert genau dann, wenn die Folge der Partialsummen $S_n = \sum_{k=0}^{n} a_k$ nach oben beschränkt ist.*

Beweis: Da $a_k \geq 0$, ist die Folge (S_n) der Partialsummen monoton wachsend. Die Behauptung folgt aus Satz 4.3 und Satz 3.4. □

Beispiel 6.5 Für $s > 1$ ist die Reihe $\sum_{k=1}^{\infty} k^{-s}$ konvergent[4]. Denn Summation über den Abschnitt $2^m \leq k < 2^{m+1}$ ergibt, da $a \mapsto a^{-s}$ monoton fallend ist nach Satz 4.6,

$$\sum_{2^m \leq k < 2^{m+1}} k^{-s} \leq 2^m \cdot (2^m)^{-s} = (2^{1-s})^m.$$

Da $2^{1-s} < 1$, folgt für $n < 2^{M+1}$ mit der geometrischen Reihe, Beispiel 3.7,

$$\sum_{k=1}^{n} k^{-s} \leq \sum_{m=0}^{M} \left(\sum_{2^m \leq k < 2^{m+1}} k^{-s} \right) \leq \sum_{m=0}^{M} (2^{1-s})^m \leq \frac{1}{1 - 2^{1-s}}.$$

Die Behauptung folgt nun aus Satz 6.3. Wir haben damit eine wohldefinierte Funktion

$$\zeta : (1, \infty) \to \mathbb{R},\ \zeta(s) = \sum_{k=1}^{\infty} k^{-s},$$

die sogenannte Riemannsche Zetafunktion. Sie spielt bei der Untersuchung der Verteilung der Primzahlen eine fundamentale Rolle.

Die Reihe $\sum_{k=1}^{\infty} (-1)^{k-1}/k = 1 - 1/2 + 1/3 - + \ldots$ ist ein Beispiel für eine sogenannte alternierende Reihe. Während die entsprechende Reihe mit nur positiven Vorzeichen nicht konvergiert – es ist die harmonische Reihe aus Beispiel 6.4 – ist die Reihe mit dem Vorzeichenwechsel konvergent. Dies ergibt sich aus dem nächsten Satz, bei dessen Beweis wieder Monotonieargumente eine wesentliche Rolle spielen.

[4] k^{-s} ist bisher nur definiert für $s \in \mathbb{Q}$, vgl. aber Definition 11.1

Satz 6.4 (Leibniz-Kriterium für alternierende Reihen) *Sei a_k, $k \in \mathbb{N}_0$, eine reelle, monoton fallende Nullfolge (also insbesondere $a_k \geq 0$). Dann ist die Reihe $\sum_{k=0}^{\infty}(-1)^k a_k$ konvergent. Außerdem gilt für alle $n \in \mathbb{N}_0$ die Abschätzung*

$$0 \leq (-1)^n \sum_{k=n}^{\infty}(-1)^k a_k \leq a_n.$$

Beweis: Wir betrachten die Partialsummen $S_n = \sum_{k=0}^{n}(-1)^k a_k$. Für alle $n \in \mathbb{N}_0$ gilt:

$$\begin{aligned}
S_{2n+2} - S_{2n} &= (-1)^{2n+2}a_{2n+2} + (-1)^{2n+1}a_{2n+1} = a_{2n+2} - a_{2n+1} \leq 0,\\
S_{2n+3} - S_{2n+1} &= (-1)^{2n+3}a_{2n+3} + (-1)^{2n+2}a_{2n+2} = a_{2n+2} - a_{2n+3} \geq 0,\\
S_{2n+1} - S_{2n} &= (-1)^{2n+1}a_{2n+1} = -a_{2n+1} \leq 0.
\end{aligned}$$

Also ist $(S_{2n})_{n\in\mathbb{N}_0}$ monoton fallend, und $(S_{2n+1})_{n\in\mathbb{N}_0}$ monoton wachsend. Weiter ist

$$S_{2n} \geq S_{2n+1} \geq S_1 \quad \text{ und } S_{2n+1} \leq S_{2n} \leq S_0.$$

Nach Satz 4.3 sind die Folgen S_{2n} und S_{2n+1} konvergent. Aber $S_{2n+1} - S_{2n} \to 0$, das heißt beide Folgen, und damit die gesamte Folge, konvergieren gegen denselben Grenzwert $S \in \mathbb{R}$. Nun ist $S \in [S_1, S_0]$ mit $S_0 = a_0$ und $S_1 = a_0 - a_1 \geq 0$, also gilt die Abschätzung im Fall $n = 0$. Anwendung auf die Folge $b_k = a_{n+k}$ liefert die Abschätzung für beliebige $n \in \mathbb{N}_0$. □

Bei der alternierenden Reihe $1 - 1/2 + 1/3 - + \ldots$ stößt man auf Merkwürdigkeiten, wenn man die Summationsreihenfolge ändert. Während die Ausgangsreihe konvergent ist, ist die durch Umordnung entstehende Reihe

$$\left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{5} - \frac{1}{4}\right) + \left(\frac{1}{7} + \frac{1}{9} + \frac{1}{11} + \frac{1}{13} - \frac{1}{6}\right) + \ldots.$$

bestimmt divergent gegen $+\infty$, denn die Summe der positiven Zahlen in der m-ten Klammer ist mindestens $2^{m-1} \cdot 2^{-(m+1)} = 1/4$. Dies ist ein interessantes Phänomen, jedoch sind wir in erster Linie an Reihen interessiert, deren Konvergenz stabiler ist. Dabei ist der folgende Begriff zentral.

Definition 6.2 *Die Reihe $\sum_{k=0}^{\infty} a_k$ mit $a_k \in \mathbb{C}$ heißt absolut konvergent, wenn die Reihe $\sum_{k=0}^{\infty} |a_k|$ konvergiert, das heißt $\sum_{k=0}^{\infty} |a_k| < \infty$.*

Satz 6.5 (absolut konvergent ⇒ konvergent) *Wenn die Reihe $\sum_{k=0}^{\infty} a_k$ absolut konvergiert, so ist sie konvergent und es gilt*

$$\left|\sum_{k=0}^{\infty} a_k\right| \leq \sum_{k=0}^{\infty} |a_k|.$$

Beweis: Die Dreiecksungleichung besagt

$$\left|\sum_{k=n}^{m} a_k\right| \leq \sum_{k=n}^{m} |a_k| \quad \text{ für } m \geq n \geq 0.$$

Aus dem Cauchykriterium für $\sum_{k=0}^{\infty} |a_k|$ folgt deshalb das Cauchykriterium für $\sum_{k=0}^{\infty} a_k$, das heißt die Reihe konvergiert nach Satz 6.2. Die Abschätzung folgt, indem wir in der Dreiecksungleichung $n = 0$ setzen und $m \to \infty$ gehen lassen. □

Der folgende Satz fasst drei wichtige Kriterien für die absolute Konvergenz zusammen.

Satz 6.6 (Tests für absolute Konvergenz) *Sei* $\sum_{k=0}^{\infty} a_k$ *eine Reihe mit* $a_k \in \mathbb{C}$. *Ist eine der drei folgenden Bedingungen erfüllt, so ist* $\sum_{k=0}^{\infty} a_k$ *absolut konvergent:*

(a) Majorantenkriterium (M-Test): *Es gilt* $|a_k| \leq c_k \in [0, \infty)$ *mit* $\sum_{k=0}^{\infty} c_k < \infty$.

(b) Quotientenkriterium: *Es gibt ein* $\theta \in [0,1)$ *und ein* $n \in \mathbb{N}$ *mit*

$$\frac{|a_{k+1}|}{|a_k|} \leq \theta \textit{ für alle } k \geq n \quad (\textit{wobei } a_k \neq 0 \textit{ für } k \geq n).$$

(c) Wurzelkriterium: *Es gibt ein* $\theta \in [0,1)$ *und ein* $n \in \mathbb{N}$ *mit* $\sqrt[k]{|a_k|} \leq \theta$ *für alle* $k \geq n$.

Umgekehrt ist die Reihe divergent, wenn für ein $n \in \mathbb{N}$ *gilt:*

$$\frac{|a_{k+1}|}{|a_k|} \geq 1 \quad \textit{oder} \quad \sqrt[k]{|a_k|} \geq 1 \textit{ für alle } k \geq n.$$

Für die harmonische Reihe $\sum_{k=1}^{\infty} 1/k$ haben wir

$$\frac{|a_{k+1}|}{|a_k|} = \frac{k}{k+1} < 1 \text{ für alle } k \quad \text{und} \quad \lim_{k\to\infty} \frac{|a_{k+1}|}{|a_k|} = 1.$$

Also reicht es *nicht*, in (b) nur die Ungleichung $|a_{k+1}|/|a_k| < 1$ vorauszusetzen. Andererseits folgt aus $\lim_{k\to\infty} |a_{k+1}|/|a_k| = 1$ auch nicht notwendig Divergenz: die Reihe $\sum_{k=1}^{\infty} 1/k^2$ hat $|a_{k+1}|/|a_k| = (k/k{+}1)^2 \to 1$, aber sie konvergiert nach Beispiel 6.5. Eine ganz analoge Diskussion gilt für das Wurzelkriterium. Die Bedingungen (b) beziehungsweise (c) sind äquivalent zu den folgenden:

$$\limsup_{k\to\infty} \frac{|a_{k+1}|}{|a_k|} < 1 \quad \text{bzw.} \quad \limsup_{k\to\infty} \sqrt[k]{|a_k|} < 1.$$

Beweis des Satzes: (a) folgt aus Satz 6.3, denn für alle $n \in \mathbb{N}$ hat man die obere Schranke

$$\sum_{k=0}^{n} |a_k| \leq \sum_{k=0}^{n} c_k \leq \sum_{k=0}^{\infty} c_k < \infty.$$

Voraussetzung (b) liefert per Induktion

$$|a_k| \leq \theta^{k-n} |a_n| \quad \text{für } k \geq n.$$

Da die geometrische Reihe wegen $0 \leq \theta < 1$ nach Beispiel 3.7 konvergiert, folgt die Behauptung aus (a) und wir erhalten außerdem die Abschätzung

$$\sum_{k=n}^{\infty} |a_k| \leq \frac{|a_n|}{1-\theta}. \tag{6.2}$$

Unter der Voraussetzung (c) gilt

$$|a_k| \leq \theta^k \quad \text{für } k \geq n.$$

Wieder folgt die Behauptung durch durch M-Test mit der geometrischen Reihe. Die Abschätzung lautet hier

$$\sum_{k=n}^{\infty} |a_k| \leq \frac{\theta^n}{1-\theta}. \tag{6.3}$$

Die Divergenzaussagen folgen unmittelbar aus dem Nullfolgentest, Satz 6.1. □

Die folgenden zwei Sätze werden in der Vorlesung nicht gebraucht, auch wenn sie zum Standardrepertoire gerechnet werden.

Erstens geht es um die Multiplikation von zwei Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$. Wir schreiben die Produkte $a_k b_l$ mit $k, l \in \mathbb{N}_0$ in einem Schema auf, so dass $a_k b_l$ in der k-ten Zeile und l-ten Spalte steht. Es ist dann naheliegend, erst die $n+1$ Produkte $a_k b_l$ in jeder Diagonale $k + l = n$ zu addieren, und dann die Konvergenz der resultierenden Reihe zu studieren. Im folgenden schreiben wir kurz $\sum_{k+l=n}$ für die Summe über alle Paare $k, l \in \mathbb{N}_0$ mit $k + l = n$.

Satz 6.7 (Cauchyprodukt) *Die Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ seien absolut konvergent. Dann ist auch die Reihe*

$$\sum_{k=0}^{\infty} c_n \quad \textit{mit} \quad c_n = \sum_{k+l=n} a_k b_l = \sum_{k=0}^{n} a_k b_{n-k}$$

absolut konvergent, und es gilt

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} \left(\sum_{k+l=n} a_k\, b_l \right) = \left(\sum_{k=0}^{\infty} a_k \right) \cdot \left(\sum_{k=0}^{\infty} b_k \right).$$

Beweis: Wir setzen für $N \in \mathbb{N}_0$

$$A_N := \sum_{k=0}^{N} |a_k| \to \sum_{k=0}^{\infty} |a_k| =: A \quad \text{und} \quad B_N := \sum_{k=0}^{N} |b_k| \to \sum_{k=0}^{\infty} |b_k| =: B,$$

wobei nach Voraussetzung $A, B < \infty$. Es folgt

$$\sum_{n=0}^{N} |c_n| \leq \sum_{k+l \leq N} |a_k|\, |b_l| \leq \sum_{k,l \leq N} |a_k|\, |b_l| = A_N B_N \leq AB < \infty.$$

Nach Satz 6.3 ist $\sum_{n=0}^{\infty} c_n$ absolut konvergent und insbesondere konvergent (Satz 6.5). Um den Grenzwert zu identifizieren, reicht es also die geraden Partialsummen $\sum_{n=0}^{2N} c_n$ zu betrachten. Wir berechnen

$$\left| \left(\sum_{k=0}^{2N} a_k \right) \cdot \left(\sum_{l=0}^{2N} b_l \right) - \sum_{n=0}^{2N} c_n \right| = \left| \sum_{k,l \leq 2N} a_k b_l - \sum_{k+l \leq 2N} a_k b_l \right| \leq \sum_{k,l \leq 2N, \max(k,l) > N} |a_k|\, |b_l|,$$

Die rechte Seite ist gleich $A_{2N}\, B_{2N} - A_N\, B_N$, konvergiert also gegen Null für $N \to \infty$. Für die letzte Abschätzung ist es hilfreich, die Indexbereiche zu skizzieren. Jedenfalls ist damit die gewünschte Formel für das Cauchyprodukt bewiesen. □

Zweitens kommen wir zu der Frage der Umordnung zurück und zeigen, dass absolut konvergente Reihen beliebig umgeordnet werden können.

Satz 6.8 (Umordnungssatz) *Sei $\tau : \mathbb{N} \to \mathbb{N}$ bijektiv. Falls $\sum_{k=1}^{\infty} a_k$ absolut konvergent ist, so konvergiert auch die Reihe $\sum_{i=1}^{\infty} a_{\tau(i)}$ absolut und hat denselben Grenzwert.*

Beweis: Für $K \geq \max\{\tau(i) : 1 \leq i \leq m\}$ ist

$$\sum_{i=1}^{m} |a_{\tau(i)}| \leq \sum_{k=1}^{K} |a_k| \leq \sum_{k=1}^{\infty} |a_k| < \infty.$$

Nach dem Majorantenkriterium konvergiert die umgeordnete Reihe absolut. Sei nun $S = \sum_{k=1}^{\infty} a_k$. Für gegebenes $K \in \mathbb{N}$ und $m \geq \max\{\tau^{-1}(k) : 1 \leq k \leq K\}$, bzw. äquivalent dazu $\{\tau(1), \ldots, \tau(m)\} \supset \{1, \ldots, K\}$, gilt

$$\Big|S - \sum_{i=1}^{m} a_{\tau(i)}\Big| \leq \Big|S - \sum_{k=1}^{K} a_k\Big| + \Big|\sum_{k=1}^{K} a_k - \sum_{i=1}^{m} a_{\tau(i)}\Big| \leq \Big|S - \sum_{k=1}^{K} a_k\Big| + \sum_{k=K+1}^{\infty} |a_k|.$$

Die Behauptung folgt, indem wir erst $m \to \infty$ und dann $K \to \infty$ gehen lassen. $\square$

7 Stetigkeit

In diesem Kapitel beginnt das Studium von Funktionen mit den Methoden der Analysis. Eine Funktion auf einer Menge $D \subset \mathbb{R}^n$ mit Werten in $\mathbb{R}^m$ ist bekanntlich eine Abbildung $f : D \to \mathbb{R}^m$, $x \mapsto f(x)$. Im Fall $m = 1$ heißt die Funktion reellwertig.

Beispiel 7.1 Ist $I \subset \mathbb{R}$ ein Intervall, so bezeichnet man Funktionen $c : I \to \mathbb{R}^m$ auch als Wege oder Kurven ($c =$ *curva*). Oft wird die unabhängige Variable mit $t \in I$ (statt $x \in I$) bezeichnet. Diese Notation folgt Newton, der Bahnkurven eines Massenpunkts als Funktion der Zeit ($t =$ *tempus*) betrachtet und $c(t) = (x(t), y(t), z(t))$ schreibt. Ein explizites Beispiel ist der horizontale Wurf aus der Höhe h mit Geschwindigkeit v:

$$c : [0, \sqrt{2h/g}] \to \mathbb{R}^3,\ c(t) = \big(vt, 0, h - \frac{1}{2}gt^2\big) \quad (g = \text{Erdbeschleunigung}).$$

Beispiel 7.2 Die Abstandsfunktion einer Menge $E \subset \mathbb{R}^n$ ist

$$\operatorname{dist}_E : \mathbb{R}^n \to \mathbb{R},\ \operatorname{dist}_E(x) = \inf\{\|x - y\| : y \in E\}.$$

Das Konzept der stetigen Abbildung hat in der Mathematik fundamentale Bedeutung. In der Schule wird oft die anschauliche Charakterisierung gegeben, dass sich der Graph der Funktion zeichnen lässt, ohne mit dem Stift abzusetzen. Das ist gar nicht so schlecht, aber in mathematischen Argumenten können wir es nicht verwenden.

Definition 7.1 (Stetigkeit) *Die Funktion $f : D \to \mathbb{R}^m$ heißt stetig im Punkt $x_0 \in D$, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass gilt:*

$$\|f(x) - f(x_0)\| < \varepsilon \quad \textit{für alle } x \in D \textit{ mit } \|x - x_0\| < \delta. \tag{7.1}$$

f heißt stetig, falls f in allen $x_0 \in D$ stetig ist.

Mithilfe von Umgebungen lässt sich die Stetigkeit in x_0 auch so fassen:

> *Für alle $\varepsilon > 0$ gibt es ein $\delta > 0$ mit $f(D \cap B_\delta(x_0)) \subset B_\varepsilon(f(x_0))$.*

Wir wollen einige Beispiele betrachten.

Beispiel 7.3 Eine konstante Funktion $f : D \to \mathbb{R}^m$ ist stetig, denn es gilt

$$\|f(x) - f(x_0)\| = 0 \quad \text{für alle } x \in D.$$

Zu gegebenem $\varepsilon > 0$ können wir jedes $\delta > 0$ wählen.

Beispiel 7.4 Betrachte $f : \mathbb{R} \to \mathbb{R}$, $f(x) = ax$, mit $a \in \mathbb{R}\backslash\{0\}$. Es gilt

$$|f(x) - f(x_0)| = |ax - ax_0| = |a|\,|x - x_0|.$$

f ist stetig, wir können $\delta = \varepsilon/|a|$ wählen.

Beispiel 7.5 Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, ist stetig. Schätze ab

$$|f(x) - f(x_0)| \le (|x| + |x_0|)\,|x - x_0| \le (1 + 2|x_0|)\,|x - x_0| \quad \text{für } |x - x_0| \le 1.$$

Wir können also $\delta = \min(1, \varepsilon/(1 + 2|x_0|))$ wählen. In diesem Beispiel muss $\delta > 0$ von x_0 abhängen. Anschaulich wird der Graph von f zunehmend steiler für $|x_0|$ groß.

Beispiel 7.6 Die charakteristische Funktion (Indikatorfunktion) einer Menge $E \subset \mathbb{R}$ ist

$$\chi_E : \mathbb{R} \to \mathbb{R},\ \chi_E(x) = \begin{cases} 1 & \text{für } x \in E, \\ 0 & \text{sonst.} \end{cases}$$

Sei χ_E stetig im Punkt $x_0 \in \mathbb{R}$. Dann gibt es zu $\varepsilon = 1$ ein $\delta > 0$ mit

$$|\chi_E(x) - \chi_E(x_0)| < 1 \quad \text{für alle } x \in (x_0 - \delta, x_0 + \delta).$$

Ist $\chi_E(x_0) = 1$, so folgt $\chi_E(x) = 1$ und somit $(x_0 - \delta, x_0 + \delta) \subset E$. Ist $\chi_E(x_0) = 0$, so folgt analog $(x_0 - \delta, x_0 + \delta) \subset \mathbb{R}^n \backslash E$. Ein interessanter Spezialfall ist $E = \mathbb{Q}$. Die Funktion $\chi_{\mathbb{Q}}$ nennt man auch Dirichletfunktion, weil Dirichlet sie in seinen Vorlesungen als Beispiel eingeführt hat. $\mathbb{Q}$ ist dicht in $\mathbb{R}$ nach Satz 3.2. Aber $\mathbb{R}\backslash\mathbb{Q}$ ist ebenfalls dicht, zum Beispiel ist $\sqrt{2} + \mathbb{Q}$ dicht in $\mathbb{R}$ und Teilmenge von $\mathbb{R}\backslash\mathbb{Q}$ nach Satz 4.1. Also gibt es in jedem Intervall $(x_0 - \delta, x_0 + \delta)$ Punkte aus $\mathbb{Q}$ und Punkte aus $\mathbb{R}\backslash\mathbb{Q}$. Die Funktion $\chi_{\mathbb{Q}}$ ist daher in keinem Punkt $x_0 \in \mathbb{R}$ stetig.

Beispiel 7.7 Die Euklidische Norm auf $\mathbb{R}^n$ (Betragsfunktion im Fall $\mathbb{R}$ bzw. $\mathbb{C}$) ist

$$f : \mathbb{R}^n \to \mathbb{R},\ f(x) = \|x\| = (x_1^2 + \ldots + x_n^2)^{1/2}.$$

Diese Funktion ist stetig auf ganz $\mathbb{R}^n$, denn aus der Dreiecksungleichung folgt

$$|f(x) - f(x_0)| = |\|x\| - \|x_0\|| \leq \|x - x_0\|.$$

Wir können hier $\delta = \varepsilon$ nehmen.

Ein sehr nützliches hinreichendes Kriterium für die Stetigkeit ist das folgende.

Beispiel 7.8 $f : D \to \mathbb{R}^m$ heißt Lipschitzstetig mit Konstante $L \in [0, \infty)$, falls gilt:

$$\|f(x) - f(y)\| \leq L\|x - y\| \quad \text{für alle } x, y \in D.$$

Zum Beispiel ist die Euklidische Norm Lipschitzstetig mit Konstante $L = 1$, vgl. Beispiel 7.7. Es gilt allgemein: *jede Lipschitzstetige Funktion ist stetig*. Sei dazu $L > 0$ die Lipschitzkonstante. Zu gegebenem $\varepsilon > 0$ wählen wir $\delta = \varepsilon/L > 0$, und erhalten

$$\|x - x_0\| < \delta \quad \Rightarrow \quad \|f(x) - f(x_0)\| \leq L\|x - x_0\| < L\delta = \varepsilon.$$

Beispiel 7.9 Die Abstandsfunktion $\mathrm{dist}_E : \mathbb{R}^n \to \mathbb{R}$ einer Menge $E \subset \mathbb{R}^n$ ist Lipschitzstetig mit Konstante Eins. Denn es gilt nach der Dreiecksungleichung

$$\|y - x_2\| \leq \|y - x_1\| + \|x_1 - x_2\| \quad \text{für alle } x_{1,2} \in \mathbb{R}^n,\ y \in E.$$

Bilden wir auf beiden Seiten das Infimum über alle $y \in E$, so folgt

$$\mathrm{dist}_E(x_2) \leq \mathrm{dist}_E(x_1) + \|x_1 - x_2\|.$$

Durch Vertauschen von x_1 mit x_2 ergibt sich dist_E Lipschitzstetig, genauer

$$|\mathrm{dist}_E(x_1) - \mathrm{dist}_E(x_2)| \leq \|x_1 - x_2\|.$$

Beispiel 7.10 Lineare Abbildungen $A : \mathbb{R}^n \to \mathbb{R}^m$ sind Lipschitzstetig. Dazu verwenden wir für $x \in \mathbb{R}^n$ die Darstellung bezüglich der Standardbasis, und schätzen mit der Dreiecksungleichung und Cauchy-Schwarz ab:

$$\|Ax\|^2 = \|A\Big(\sum_{j=1}^n x_j e_j\Big)\|^2 \le \Big(\sum_{j=1}^n \|Ae_j\|\,|x_j|\Big)^2 \le \Big(\sum_{j=1}^n \|Ae_j\|^2\Big)\Big(\sum_{j=1}^n |x_j|^2\Big).$$

Also gilt die Abschätzung

$$\|Ax\| \le \|A\|\,\|x\| \quad \text{für alle } x \in \mathbb{R}^n,$$

wobei die Euklidische Norm der Matrix A definiert ist durch

$$\|A\| = \Big(\sum_{j=1}^n \|Ae_j\|^2\Big)^{1/2} = \Big(\sum_{j=1}^n \sum_{i=1}^m a_{ij}^2\Big)^{1/2}.$$

Es folgt weiter

$$\|Ax - Ay\| = \|A(x-y)\| \le \|A\|\,\|x-y\| \quad \text{für alle } x, y \in \mathbb{R}^n,$$

das heißt A ist Lipschitzstetig mit Konstante $\|A\|$. Die Verwendung der endlichen Basis e_j ist hier unverzichtbar, denn lineare Abbildungen $A : V \to W$ sind nicht notwendig stetig, wenn V unendlichdimensional ist.

Bisher haben wir direkt mit der Definition argumentiert. Jetzt wollen zeigen, dass die Stetigkeit unter gewissen Operationen erhalten bleibt. Dazu wollen wir die Konvergenzregeln für Folgen anwenden, siehe Satz 3.5; der Zusammenhang wird durch folgende Aussage geleistet.

Satz 7.1 (Folgenkriterium der Stetigkeit) *Für $f : D \to \mathbb{R}^m$ und $x_0 \in D$ sind folgende Aussagen äquivalent:*

(1) *f ist stetig in x_0.*

(2) *Für jede Folge $(x_k)_{k\in\mathbb{N}}$ mit $x_k \in D$ und $\lim_{k\to\infty} x_k = x_0$ gilt: $\lim_{k\to\infty} f(x_k) = f(x_0)$.*

Beweis: Für die Implikation (1) $\Rightarrow$ (2) sei $x_k \in D$ mit $x_k \to x_0$ für $k \to \infty$. Wähle zu $\varepsilon > 0$ ein $\delta > 0$ mit $\|f(x) - f(x_0)\| < \varepsilon$ für alle $x \in D$ mit $\|x - x_0\| < \delta$. Es gibt ein $K \in \mathbb{R}$ mit $\|x_k - x_0\| < \delta$ für $k > K$, also folgt $\|f(x_k) - f(x_0)\| < \varepsilon$ für $k > K$.

Jetzt gelte (2). Wir nehmen indirekt an, f sei nicht stetig in x_0. Es gibt dann ein $\varepsilon > 0$, so dass (7.1) für kein $\delta > 0$ erfüllt ist. Wählen wir $\delta_k = 1/k$ mit $k = 1, 2, \ldots$, so gibt es jeweils ein $x_k \in D$ mit $\|x_k - x_0\| < 1/k$, aber $\|f(x_k) - f(x_0)\| \ge \varepsilon$. Die Folge x_k konvergiert gegen x_0, aber $f(x_k)$ konvergiert nicht gegen $f(x_0)$, Widerspruch zu (2). Damit ist der Satz bewiesen. $\square$

Satz 7.2 (Verkettung stetiger Funktionen) *Seien $f : D \to \mathbb{R}^m$, $f(D) \subset E \subset \mathbb{R}^m$, und $g : E \to \mathbb{R}^k$. Ist f stetig in x_0 und g stetig in $y_0 = f(x_0)$, so ist $g \circ f : D \to \mathbb{R}^k$ stetig in x_0.*

Beweis: Wir verwenden Satz 7.1. Ist $x_n \in D$ eine Folge mit $\lim_{n\to\infty} x_n = x_0$, so folgt $f(x_n) \to f(x_0)$ aus der Stetigkeit von f in x_0, und weiter $g(f(x_n)) \to g(f(x_0))$ wegen der Stetigkeit von g in $y_0 = f(x_0)$. Nach Satz 7.1 ist die Stetigkeit von $g \circ f$ in x_0 gezeigt. $\square$

Beispiel 7.11 Seien $f, g : D \to \mathbb{R}$ stetig. Dann ist auch die Funktion $|f| : D \to \mathbb{R}$ stetig, und ebenso die Funktionen $\max(f, g) = (f + g + |f - g|)/2$ und $\min(f, g) = (f + g - |f - g|)/2$.

Lemma 7.1 *Eine Funktion $f : D \to \mathbb{R}^m$ ist genau dann stetig in $x_0 \in D$, wenn jede Koordinatenfunktion $f_i : D \to \mathbb{R}$ in x_0 stetig ist.*

Beweis: Nach Satz 5.6 ist eine Folge von Punkten im $\mathbb{R}^m$ genau dann konvergent, wenn die einzelnen Koordinatenfolgen konvergieren. Mit Satz 7.1 ergibt sich die Behauptung. □

Der Einfachheit halber beschränken wir uns im folgenden Satz auf reellwertige Funktionen.

Satz 7.3 (Stetigkeitsregeln) *Seien $f, g : D \to \mathbb{R}$ stetig in $x_0 \in D$. Dann gilt:*

(1) *Für beliebige $\lambda, \mu \in \mathbb{R}$ ist die Funktion $\lambda f + \mu g$ stetig in x_0.*

(2) *Die Funktion fg ist stetig in x_0.*

(3) *Ist $g(x_0) \neq 0$, so ist die Funktion $f/g : D \cap B_\delta(x_0) \to \mathbb{R}$ für $\delta > 0$ hinreichend klein definiert und stetig in x_0.*

Beweis: Die Funktionen sind grundsätzlich punktweise erklärt, das heißt für alle $x \in D$ gilt

$$(\lambda f + \mu g)(x) = \lambda f(x) + \mu g(x), \; (fg)(x) = f(x)g(x) \text{ und } (f/g)(x) = f(x)/g(x).$$

Wir führen die Aussagen auf die entsprechenden Regeln für Folgen zurück: sei $(x_n)_{n \in \mathbb{N}}$ eine beliebige Folge mit $x_n \in D$ und $\lim_{n \to \infty} x_n = x_0$. Nach Satz 7.1 gilt $f(x_n) \to f(x_0)$ und $g(x_n) \to g(x_0)$ mit $n \to \infty$. Aus Satz 3.5 folgt

$$\lambda f(x_n) + \mu g(x_n) \to \lambda f(x_0) + \mu g(x_0) \quad \text{sowie} \quad f(x_n)g(x_n) \to f(x_0)g(x_0).$$

Mit Satz 7.1 ergeben sich (1) und (2). Aussage (3) folgt ebenso, falls $g \neq 0$ auf D: dann ist $f/g : D \to \mathbb{R}$ definiert, und für jede Folge $x_n \to x_0$ gilt $f(x_n)/g(x_n) \to f(x_0)/g(x_0)$. Somit ist f/g stetig in x_0 nach Satz 7.1. Die Annahme $g \neq 0$ auf D ist aber zu stark, es wird ja nur eine lokale Aussage bei x_0 behauptet. Offenbar reicht $g \neq 0$ auf $D \cap B_\delta(x_0)$, wenn wir f/g nur auf $D \cap B_\delta(x_0)$ betrachten. Die Existenz dieser Umgebung wird im Anschluss bewiesen. □

Lemma 7.2 *Sei $g : D \to \mathbb{R}$ stetig in x_0 mit $g(x_0) \neq 0$. Dann gibt es ein $\delta > 0$ mit*

$$|g(x)| \geq \frac{1}{2}|g(x_0)| > 0 \quad \textit{für alle } x \in D \cap B_\delta(x_0).$$

Beweis: Da g stetig in x_0, gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit $|g(x) - g(x_0)| < \varepsilon$ für alle $x \in D \cap B_\delta(x_0)$, also folgt mit der Dreiecksungleichung für diese x

$$|g(x)| = |g(x_0) - (g(x_0) - g(x))| \geq |g(x_0)| - |g(x) - g(x_0)| > |g(x_0)| - \varepsilon.$$

Mit der Wahl $\varepsilon = \frac{1}{2}|g(x_0)| > 0$ folgt die Behauptung. □

Eine offensichtliche Konsequenz von Satz 7.3(1), sowie von Lemma 7.1 im Fall $m \geq 2$, ist

Folgerung 7.1 (C^0-Räume) *Die stetigen Abbildungen $C^0(D, \mathbb{R}^m)$ bilden einen Untervektorraum des Raums aller Abbildungen $f : D \to \mathbb{R}^m$, mit der punktweisen Addition und Skalarmultiplikation.*

Wir behandeln jetzt noch den Begriff des Grenzwerts einer Funktion. Die Unterschiede zu Folgen sind gering, deshalb fassen wir uns kurz.

Definition 7.2 (Grenzwert für Funktionen) *Sei $D \subset \mathbb{R}^n$ und x_0 ein Häufungspunkt von D. Die Funktion $f : D \to \mathbb{R}^m$ konvergiert für $x \to x_0$ gegen $a \in \mathbb{R}^m$, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass gilt:*

$$\|f(x) - a\| < \varepsilon \quad \textit{für alle } x \in D \textit{ mit } 0 < \|x - x_0\| < \delta.$$

Notation: $\lim_{x \to x_0} f(x) = a$ *oder* $f(x) \to a$ *für* $x \to x_0$.

Für die Existenz und den Wert von $\lim_{x \to x_0} f(x)$ ist es egal, ob die Funktion f in x_0 definiert ist bzw. welchen Funktionswert sie dort hat. Die Beziehung zwischen Grenzwert und Stetigkeit ist wie folgt:

Lemma 7.3 (Stetigkeit und Grenzwert) *Sei $x_0 \in D$ ein Häufungspunkt von D. Für die Funktion $f : D \to \mathbb{R}^m$ sind folgende Aussagen äquivalent:*

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

(2) *f ist stetig in x_0.*

Beweis: Die beiden Aussagen lauten:

(1) $\forall \varepsilon > 0 \, \exists \, \delta > 0 : \|f(x) - f(x_0)\| < \varepsilon \quad \forall x \in D \cap B_\delta(x_0) \backslash \{x_0\}$,

(2) $\forall \varepsilon > 0 \, \exists \, \delta > 0 : \|f(x) - f(x_0)\| < \varepsilon \quad \forall x \in D \cap B_\delta(x_0)$.

Da sowieso $\|f(x) - f(x_0)\| = 0$ für $x = x_0$, sind die Aussagen identisch. □

Für reellwertige Funktionen können wir den Konvergenzbegriff ausdehnen, indem wir (uneigentliche) Konvergenz gegen $\pm\infty$ zulassen.

Definition 7.3 (Uneigentlicher Grenzwert) *Sei $x_0 \in \mathbb{R}^n$ ein Häufungspunkt von D, und $f : D \to \mathbb{R}$. Dann gilt $\lim_{x \to x_0} f(x) = +\infty$, wenn es zu jedem $K > 0$ ein $\delta > 0$ gibt mit*

$$f(x) > K \textit{ für alle } x \in D \textit{ mit } 0 < \|x - x_0\| < \delta.$$

Entsprechend wird $\lim_{x \to x_0} f(x) = -\infty$ *erklärt.*

Im Fall $D \subset \mathbb{R}$ besteht weiter die Möglichkeit, Grenzwerte für $x \to \pm\infty$ zu betrachten.

Definition 7.4 (Grenzwert bei $\pm\infty$) *Sei $f : D \to \mathbb{R}^m$ mit $D \subset \mathbb{R}$ nicht nach oben beschränkt. Dann gilt $\lim_{x \to +\infty} f(x) = a \in \mathbb{R}^m$, wenn es zu $\varepsilon > 0$ ein $K \in \mathbb{R}$ gibt mit*

$$\|f(x) - a\| < \varepsilon \quad \textit{für alle } x \in D \textit{ mit } x > K.$$

Analog wird der Grenzwert für $x \to -\infty$ erklärt.

Sowohl bei der Stetigkeit als auch beim Grenzwert spielt der zugrundeliegende Definitionsbereich eine Rolle. Wir schreiben $\lim_{x \to x_0, x \in D} f(x)$, wenn wir den gewählten Definitionsbereich hervorheben möchten. In $\mathbb{R}$ werden oft einseitige Grenzwerte gebraucht:

$$\lim_{x \searrow x_0} f(x) = \lim_{x \to x_0, x > x_0} f(x) \quad \text{und} \quad \lim_{x \nearrow x_0} f(x) = \lim_{x \to x_0, x < x_0} f(x).$$

Zum Beispiel ist $\lim_{x \searrow 0} \operatorname{sign}(x) = +1$ und $\lim_{x \nearrow 0} \operatorname{sign}(x) = -1$, während der Grenzwert $\lim_{x \to 0} \operatorname{sign}(x)$ nicht existiert. Sind aber der links- und rechtsseitige Grenzwert einer Funktion $f(x)$ im Punkt x_0 gleich, so ist dies der Grenzwert von $f(x)$ für $x \to x_0$.

Der Grenzwertbegriff für Funktionen lässt sich wieder auf Folgen zurückführen.

Satz 7.4 (Definition von $\lim_{x \to x_0} f(x)$ mit Folgen) *Sei x_0 Häufungspunkt von $D \subset \mathbb{R}^n$, und $f : D \to \mathbb{R}^m$. Für $a \in \mathbb{R}^m$ sind äquivalent:*

(1) *$f(x) \to a$ für $x \to x_0$.*

(2) *$\lim_{k \to \infty} f(x_k) = a$ für jede Folge $x_k \in D \backslash \{x_0\}$ mit $x_k \to x_0$.*

Beweis: Der Beweis ist analog zu Satz 7.1. Einziger Unterschied: in Definition 7.2 kommen nur $x \in D$ vor mit $\|x - x_0\| > 0$. Entsprechend werden in (2) hier nur Folgen mit $x_k \neq x_0$ betrachtet. $\square$

Für Grenzwerte von Funktionen gelten Rechenregeln analog zu Satz 7.3. Der Beweis wird den Lesern/innen überlassen.

Satz 7.5 (Rechenregeln für Grenzwerte) *Sei x_0 Häufungspunkt von $D \subset \mathbb{R}^n$. Es gilt:*

(1) *Seien $f, g : D \to \mathbb{R}$ mit $f(x) \to a$, $g(x) \to b$ für $x \to x_0$. Dann folgt*

$$\begin{aligned} \lambda f(x) + \mu g(x) &\to \lambda a + \mu b \quad (\lambda, \mu \in \mathbb{R}), \\ f(x) g(x) &\to ab, \\ f(x)/g(x) &\to a/b, \quad \text{falls } b \neq 0. \end{aligned}$$

Im Fall $D \subset \mathbb{R}$ und $x_0 = \pm\infty$ gelten analoge Aussagen.

(2) *Seien $f : D \to \mathbb{R}^m$ mit $f(D) \subset E \subset \mathbb{R}^m$, und $g : E \to \mathbb{R}$. Gilt $f(x) \to y_0$ mit $x \to x_0$ und ist g stetig in y_0, so folgt $(g \circ f)(x) \to g(y_0)$ mit $x \to x_0$.*

(3) *Sei $f : D \to \mathbb{R}$ mit $f(x) \to a$ für $x \to x_0$. Ist $f \geq 0$ auf $B_\delta(x_0) \backslash \{x_0\}$, so folgt $a \geq 0$.*

(4) *Ist $f : D \to (0, \infty)$, so ist $\lim_{x \to x_0} f(x) = 0$ äquivalent zu $\displaystyle\lim_{x \to x_0} \frac{1}{f(x)} = +\infty$.*

Beispiel 7.12 (Rationale Funktionen) Seien p, q reelle Polynome vom Grad m bzw. n, also $p(x) = a_0 + a_1 x + \ldots + a_m x^m$ bzw. $q(x) = b_0 + b_1 x + \ldots + b_n x^n$ mit $a_i, b_j \in \mathbb{R}$ und $a_m, b_n \neq 0$. Setze $N = \{x \in \mathbb{R} : q(x) = 0\}$ und definiere

$$f : \mathbb{R} \backslash N \to \mathbb{R},\ f(x) = \frac{p(x)}{q(x)}.$$

Wann hat f eine stetige Fortsetzung bei $x_0 \in N$? Sei dazu x_0 eine ν-fache Nullstelle von q mit $\nu \geq 1$, und eine μ-fache Nullstelle von p, evtl. $\mu = 0$ falls $p(x_0) \neq 0$. Also gilt

$$p(x) = (x - x_0)^\mu \, \tilde{p}(x) \quad \text{bzw.} \quad q(x) = (x - x_0)^\nu \tilde{q}(x)$$

für Polynome $\tilde{p}, \tilde{q}$ mit $\tilde{p}(x_0), \tilde{q}(x_0) \neq 0$. Es folgt

$$f(x) = (x - x_0)^{\mu - \nu} \frac{\tilde{p}(x)}{\tilde{q}(x)} \qquad \text{für alle } x \in \mathbb{R} \backslash N,$$

und wir erhalten aus den Rechenregeln für Grenzwerte

$$\lim_{x \to x_0} f(x) = \begin{cases} 0 & \text{falls } \mu > \nu, \\ \dfrac{\tilde{p}(x_0)}{\tilde{q}(x_0)} \neq 0 & \text{falls } \mu = \nu. \end{cases}$$

In diesen beiden Fällen ist die Funktion stetig fortsetzbar nach Lemma 7.3, dagegen hat f im Fall $\mu < \nu$ in x_0 eine Polstelle. Sei zum Beispiel $\tilde{p}(x_0)/\tilde{q}(x_0) > 0$, dann ergibt sich für die einseitigen Grenzwerte bei x_0

	$x \searrow x_0$	$x \nearrow x_0$
$\nu - \mu$ gerade	$+\infty$	$+\infty$
$\nu - \mu$ ungerade	$+\infty$	$-\infty$

Das Verhalten von f für $x \to \pm\infty$ wurde bereits in Beispiel 3.6 untersucht.

8 Zwischenwertsatz und monotone Funktionen

In diesem Abschnitt haben wir es mit reellen Funktionen zu tun, die auf einem Intervall definiert sind. Eine Menge $I \subset \mathbb{R}$ ist genau dann ein Intervall, wenn $(a,b) \subset I$ mit $a = \inf I$ und $b = \sup I$. Hier sind also unendliche Intervallgrenzen zugelassen, aber das Intervall soll stets Teilmenge von $\mathbb{R}$ sein, das heißt unendliche Intervallgrenzen sind offen. Der Durchschnitt von zwei Intervallen $I_{1,2}$ mit den Grenzen $a_{1,2}$ und $b_{1,2}$ ist wieder ein Intervall mit Grenzen $a = \max(a_1, a_2)$ und $b = \min(b_1, b_2)$. Nach unserer Definition ist die leere Menge auch ein Intervall.

Satz 8.1 (Zwischenwertsatz) *Sei $f : [a,b] \to \mathbb{R}$ stetig. Dann gibt es zu jedem y_0 zwischen $f(a)$ und $f(b)$ ein $x_0 \in [a,b]$ mit $f(x_0) = y_0$.*

Bemerkung. Die Gleichung $f(x) = y_0$ kann mehrere Lösungen in $[a,b]$ besitzen, das heißt x_0 ist im allgemeinen nicht eindeutig bestimmt.

Beweis: Sei oBdA $f(a) \leq y_0 \leq f(b)$. Dann ist die Menge $M = \{x \in [a,b] : f(x) \leq y_0\}$ nichtleer, da $a \in M$. Wir behaupten $f(x_0) = y_0$ für $x_0 = \sup M$. Für $n \in \mathbb{N}$ ist $x_0 - 1/n$ keine obere Schranke von M, also gibt es $x_n \in M$ mit $x_0 - 1/n < x_n \leq x_0$. Da f stetig, folgt $f(x_0) = \lim_{n\to\infty} f(x_n) \leq y_0$. Andererseits ist x_0 obere Schranke von M, also gilt $f(x) > y_0$ für $x_0 < x \leq b$. Ist $x_0 < b$, so folgt $f(x_0) = \lim_{x \searrow x_0} f(x) \geq y_0$. Im Fall $x_0 = b$ gilt $f(x_0) \geq y_0$ sowieso nach Voraussetzung. □

Folgerung 8.1 *Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ stetig. Dann ist $f(I)$ ein Intervall mit Endpunkten $\alpha = \inf_{x\in I} f(x)$ und $\beta = \sup_{x \in I} f(x)$.*

Beweis: Zu $y \in (\alpha, \beta)$ gibt es $x_1, x_2 \in I$ mit $f(x_1) < y < f(x_2)$. Dann gibt es nach Satz 8.1 ein $x \in [x_1, x_2]$ (bzw. $x \in [x_2, x_1]$) mit $f(x) = y$. Es folgt $(\alpha, \beta) \subset f(I)$. □

Jede streng monotone Funktion $f : I \to \mathbb{R}$ ist injektiv und hat damit eine Umkehrfunktion $g : f(I) \to \mathbb{R}$. Es stellt sich die Frage nach der Stetigkeit von g. Zum nächsten Lemma vgl. das entsprechende Resultat für Folgen, Satz 4.3 in Kapitel 1.

Lemma 8.1 (Grenzwerte monotoner Funktionen) *Sei $f : I \to \mathbb{R}$ monoton wachsend, wobei I Intervall mit Endpunkten $a < b$. Dann hat f einseitige Grenzwerte, genauer gilt*

$$\begin{aligned} \lim_{x \nearrow x_0} f(x) &= \sup\{f(x) : x \in I,\, x < x_0\} \quad \textit{für } x_0 > a, \\ \lim_{x \searrow x_0} f(x) &= \inf\{f(x) : x \in I,\, x > x_0\} \quad \textit{für } x_0 < b. \end{aligned}$$

Die entsprechende Aussage gilt für f monoton fallend.

Beweis: Zu jedem $\alpha < \sup\{f(x) : x \in I,\, x < x_0\}$ gibt es ein $x' \in I$, $x' < x_0$, so dass $f(x') > \alpha$. Da f monoton wachsend, folgt für alle $x \in [x', x_0)$

$$\alpha < f(x') \leq f(x) \leq \sup\{f(x) : x \in I,\, x < x_0\}.$$

Dies zeigt die erste Behauptung, die zweite folgt analog. □

Satz 8.2 (Monotonie und Umkehrfunktion) *Sei I ein Intervall mit Endpunkten $a < b$, und $f : I \to \mathbb{R}$ sei streng monoton wachsend und stetig. Dann gilt:*

(1) *$f(I)$ ist ein Intervall mit Endpunkten $\alpha = \lim_{x \searrow a} f(x) < \lim_{x \nearrow b} f(x) = \beta$.*

(2) *Die Umkehrfunktion $g : f(I) \to \mathbb{R}$ ist ebenfalls streng monoton wachsend und stetig.*

(3) *$\lim_{y \searrow \alpha} g(y) = a$ und $\lim_{y \nearrow \beta} g(y) = b$.*

Beweis: (1) gilt nach Folgerung 8.1 und Lemma 8.1. Für g streng monoton wachsend wiederholen wir das Argument von Gleichung (4.5): wäre $y_1 < y_2$ aber $g(y_1) \geq g(y_2)$, so folgt

$$y_1 = f\big(g(y_1)\big) \geq f\big(g(y_2)\big) = y_2 \qquad \text{wegen } f \text{ monoton wachsend,}$$

Widerspruch. Damit hat g einseitige Grenzwerte nach Lemma 8.1; für die Stetigkeit bleibt zu zeigen, dass diese gleich dem jeweiligen Funktionswert sind. Wir betrachten oBdA die linksseitigen Grenzwerte; sei $y_0 = f(x_0) \in (\alpha, \beta]$ gegeben. Angenommen es ist

$$x_- := \lim_{y \nearrow y_0} g(y) < g(y_0) = x_0.$$

Da g streng monoton wachsend, ist $a < x_- < x_0$. Mit f stetig in x_- folgt

$$f(x_-) = \lim_{y \nearrow y_0} f(g(y)) = y_0 = f(x_0),$$

im Widerspruch zu f streng monoton wachsend. Damit ist die Stetigkeit von g bewiesen. Aussage (3) folgt aus (1) und (2), angewandt auf g statt f. $\square$

Der Beweis präzisiert die folgende Vorstellung: der Graph von f ergibt sich aus dem Graph von g durch Spiegelung an der Winkelhalbierenden. Ein Sprung von g würde dabei einem Intervall entsprechen, auf dem f konstant ist, im Widerspruch zur strengen Monotonie. Als Ergänzung zu Satz 8.2 bemerken wir noch

$$a \in I \Leftrightarrow \alpha \in f(I) \quad \text{und} \quad b \in I \Leftrightarrow \beta \in f(I). \tag{8.1}$$

Für $a \in I$ ist $\alpha = \inf\{f(x) : x \in I\} = f(a)$. Sei umgekehrt $\alpha = f(x)$ für ein $x \in I$. Wäre $x > a$, so ist $f(x') < f(x) = \alpha$ für $x' \in (a, x)$, Widerspruch. Also folgt $a = x \in I$.

Beispiel 8.1 $f : [0, \infty) \to \mathbb{R}$, $f(x) = x^n$, ist stetig und streng monoton wachsend mit

$$\lim_{x \searrow 0} f(x) = 0 \quad \text{und} \quad \lim_{x \nearrow \infty} f(x) = \infty.$$

Der Satz liefert $f([0, \infty)) = [0, \infty)$, also (erneut) die Existenz der n-ten Wurzel (vgl. Satz 4.6). Weiter: die Umkehrfunktion $g : [0, \infty) \to \mathbb{R}$, $g(y) = y^{1/n}$, ist stetig und es gilt

$$\lim_{y \searrow 0} y^{1/n} = 0 \quad \text{und} \quad \lim_{y \nearrow \infty} y^{1/n} = \infty.$$

Wir werden in Kürze weitere Anwendungen des Satzes sehen. Insbesondere werden wir die Logarithmusfunktion als Umkehrfunktion der Exponentialfunktion definieren.

9 Die Ableitung

Im diesem Abschnitt betrachten wir stets reellwertige oder vektorwertige Funktionen einer Variablen, die auf einem offenen Intervall $I \subset \mathbb{R}$ definiert sind.

Definition 9.1 (Ableitung) *Die Funktion $f : I \to \mathbb{R}^n$ hat in $x_0 \in I$ die Ableitung $a \in \mathbb{R}^n$ (Notation: $f'(x_0) = a$), falls gilt:*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = a. \tag{9.1}$$

Wir nennen f differenzierbar in x_0, falls es ein $a \in \mathbb{R}^n$ mit (9.1) gibt, falls also der in (9.1) betrachtete Grenzwert existiert.

Eine alternative Formulierung ergibt sich durch die Substitution $x = x_0 + h$:

$$f'(x_0) = a \quad \Leftrightarrow \quad \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = a.$$

Leibniz interessierte sich für die Definition der Ableitung im Zusammenhang mit dem Problem, die Tangente an eine ebene Kurve in einem gegebenen Punkt zu definieren. Nehmen wir dazu an, dass die Kurve als Graph einer Funktion $f : I \to \mathbb{R}$ gegeben ist, und dass die Tangente im Punkt $(x_0, f(x_0))$ gesucht ist. Der Differenzenquotient

$$\frac{f(x) - f(x_0)}{x - x_0} \quad (x_0, x \in I,\ x \neq x_0)$$

ist geometrisch die Steigung der Sekante durch die Punkte $(x_0, f(x_0))$ und $(x, f(x))$. Die Existenz der Ableitung bedeutet, dass die Sekantensteigungen für $x \to x_0$ gegen den Wert $f'(x_0)$ konvergieren. Die Tangente wird nun definiert als die Gerade, die durch den Punkt $(x_0, f(x_0))$ geht und die Steigung $f'(x_0)$ hat. Daraus ergibt sich ihre Gleichung

$$y = f(x_0) + f'(x_0)(x - x_0) \quad \text{für alle } x \in \mathbb{R}.$$

Im Fall von vektorwertigen Funktionen, also $n \geq 2$, ist der Grenzwert in (9.1) wie üblich bezüglich der Euklidischen Norm aufzufassen. Der in Kapitel 2 bewiesene Satz 5.6 über die Äquivalenz von Normkonvergenz und Konvergenz der Koordinaten besagt hier Folgendes:

Lemma 9.1 *Die Funktion $f : I \to \mathbb{R}^n$ ist genau dann in $x_0 \in I$ differenzierbar, wenn alle Koordinatenfunktionen $f_i : I \to \mathbb{R}$, $i = 1, \ldots, n$, in x_0 differenzierbar sind. Die Ableitung kann dann koordinatenweise berechnet werden, das heißt es gilt:*

$$f'(x_0) = ((f_1)'(x_0), \ldots, (f_n)'(x_0)) \in \mathbb{R}^n.$$

Newton entwickelte den Differentialkalkül (Englisch: Calculus) unter anderem um die Keplerschen Gesetze für die Planetenbewegung zu begründen, genauer konnte er diese Gesetze alle aus dem Gravitationsgesetz ableiten. Dazu wird die Bewegung eines Planeten durch eine Abbildung

$$f : I \to \mathbb{R}^3, \quad f(t) = \big(x(t), y(t), z(t)\big),$$

beschrieben, also durch dessen Koordinaten zur Zeit $t \in I$ bezüglich eines Euklidischen Koordinatensystems. Erstes Ziel ist dann die Definition der Momentangeschwindigkeit als Vektor in $\mathbb{R}^3$. Die vektorielle Durchschnittsgeschwindigkeit auf dem Zeitintervall $[t_0, t]$ ist der Quotient von Weg und Zeit, also gleich

$$\frac{f(t) - f(t_0)}{t - t_0} \in \mathbb{R}^3.$$

Die Momentangeschwindigkeit $v(t_0)$ zum Zeitpunkt $t = t_0$ ist deshalb als vektorielle Ableitung zu definieren, wobei Newton einen Punkt statt eines Strichs benutzt hat:

$$v(t_0) = f'(t_0) = (x'(t_0), y'(t_0), z'(t_0)) \in \mathbb{R}^3.$$

Definition 9.2 (Ableitungsfunktion) *Die Funktion $f : I \to \mathbb{R}^n$ heißt differenzierbar auf I (oder einfach differenzierbar), falls f in jedem Punkt $x_0 \in I$ differenzierbar ist. Die hierdurch gegebene Funktion*

$$f' : I \to \mathbb{R}^n, \quad x_0 \mapsto f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \in \mathbb{R}^n$$

heißt Ableitungsfunktion oder schlicht Ableitung von f.

Beispiel 9.1 Für eine konstante Funktion, also $f(x) = c \in \mathbb{R}^n$ für alle $x \in I$, gilt für $x \neq x_0$:

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{c - c}{x - x_0} = 0 \quad \Rightarrow \quad f'(x_0) = 0 \text{ bzw. } f' = 0.$$

Beispiel 9.2 Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x$, gilt

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x - x_0}{x - x_0} = 1 \quad \text{für alle } x \neq x_0,$$

also folgt $f'(x_0) = 1$.

Beispiel 9.3 Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$, ist nicht differenzierbar in $x_0 = 0$:

$$\lim_{x \searrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \searrow 0} \frac{x}{x} = 1 \quad \text{und} \quad \lim_{x \nearrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \nearrow 0} \frac{-x}{x} = -1.$$

Die rechts- und linksseitige Ableitung existieren in $x_0 = 0$, sie sind aber verschieden.

Satz 9.1 (differenzierbar $\Rightarrow$ stetig) *Sei I ein offenes Intervall. Ist $f : I \to \mathbb{R}^n$ differenzierbar in x_0, so ist f auch stetig in x_0.*

Beweis: Mit Satz 7.5(1) folgt für $x \to x_0$

$$f(x) = f(x_0) + \underbrace{\frac{f(x) - f(x_0)}{x - x_0}}_{\to f'(x_0)} \underbrace{(x - x_0)}_{\to 0} \longrightarrow f(x_0).$$

Dies zeigt $\lim_{x \to x_0} f(x) = f(x_0)$, also f stetig in x_0 nach Lemma 7.3. □

Satz 9.2 (Differentiationsregeln) *Seien $f, g : I \to \mathbb{R}$ differenzierbar in $x_0 \in I$. Dann sind auch die Funktionen $\alpha f + \beta g$ ($\alpha, \beta \in \mathbb{R}$), fg und f/g (im Fall $g(x_0) \neq 0$) in x_0 differenzierbar mit folgenden Ableitungen:*

(1) *Linearität:*

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0)$$

(2) *Produktregel:*

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$$

(3) *Quotientenregel:*

$$\Big(\frac{f}{g}\Big)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$$

Beweis: Wir müssen jeweils für $x \neq x_0$ die Differenzenquotienten bilden und zeigen, dass diese mit $x \to x_0$ gegen das gewünschte konvergieren. Für (1) haben wir

$$\begin{aligned}\frac{(\alpha f(x) + \beta g(x)) - (\alpha f(x_0) + \beta g(x_0))}{x - x_0} &= \alpha \frac{f(x) - f(x_0)}{x - x_0} + \beta \frac{g(x) - g(x_0)}{x - x_0}\\ &\to \alpha f'(x_0) + \beta g'(x_0).\end{aligned}$$

Natürlich gilt die Aussage mit demselben Argument auch für vektorwertige Funktionen. Die Produktregel folgt durch „Mischen der Terme“:

$$\begin{aligned}\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} &= \frac{f(x) - f(x_0)}{x - x_0}\, g(x) + f(x_0)\, \frac{g(x) - g(x_0)}{x - x_0}\\ &\to f'(x_0)g(x_0) + f(x_0)g'(x_0),\end{aligned}$$

wobei die Stetigkeit von g in x_0 benutzt wurde (Satz 9.1). Wir zeigen die Quotientenregel zunächst für die Funktion $1/g$, also $f \equiv 1$. Es gibt ein $\delta > 0$ mit $g(x) \neq 0$ für $|x - x_0| < \delta$ nach Lemma 7.2. Für diese $x \in I$ gilt

$$\begin{aligned}\frac{1}{x - x_0}\left(\frac{1}{g(x)} - \frac{1}{g(x_0)}\right) &= -\frac{1}{g(x)g(x_0)}\frac{g(x) - g(x_0)}{x - x_0}\\ &\to -\frac{1}{g(x_0)^2}\, g'(x_0) \quad \text{mit } x \to x_0.\end{aligned}$$

Für beliebiges f schreiben wir $f/g = f \cdot \dfrac{1}{g}$ und verwenden die Produktregel. □

Beispiel 9.4 Für $f_n(x) = x^n$ folgt aus Beispiel 9.2, also $f_1' = 1$, und der Produktregel

$$f_n'(x) = (f_1 f_{n-1})'(x) = f_1'(x) f_{n-1}(x) + f_1(x) f_{n-1}'(x) = x^{n-1} + x\, f_{n-1}'(x),$$

und damit per Induktion $f_n'(x) = nx^{n-1}$. Allgemeiner ergibt sich mit Satz 9.2(1) für Polynome $p(x) = \sum_{k=0}^n a_k x^k$ die Formel

$$p'(x) = \sum_{k=1}^n k a_k x^{k-1} = \sum_{j=0}^{n-1} (j+1) a_{j+1} x^j.$$

Beispiel 9.5 Für $f(x) = x^{-n}$, $n \in \mathbb{N}$, gilt $f'(x) = -nx^{-n-1}$ nach der Quotientenregel:

$$f'(x) = -\frac{nx^{n-1}}{(x^n)^2} = -nx^{-n-1}.$$

Satz 9.3 (Kettenregel) *Seien $f : I \to \mathbb{R}$, $g : J \to \mathbb{R}$ mit $f(I) \subset J$. Ist f in $x_0 \in I$ differenzierbar und g in $y_0 = f(x_0) \in J$ differenzierbar, so ist auch $g \circ f : I \to \mathbb{R}$ in x_0 differenzierbar und hat die Ableitung*

$$(g \circ f)'(x_0) = g'\big(f(x_0)\big)\, f'(x_0).$$

Beweis: Wir betrachten wieder für $x \neq x_0$ den Differenzenquotienten:

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \begin{cases} \dfrac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)}\, \dfrac{f(x) - f(x_0)}{x - x_0} & \text{falls } f(x) \neq f(x_0) \\ 0 & \text{falls } f(x) = f(x_0). \end{cases}$$

Wir definieren die Funktion

$$a : J \to \mathbb{R},\ a(y) = \begin{cases} \dfrac{g(y) - g(f(x_0))}{y - f(x_0)} & \text{für } y \neq f(x_0) \\ g'(f(x_0)) & \text{für } y = f(x_0). \end{cases}$$

Es ist a stetig in $f(x_0)$ nach Lemma 7.3. Mit $x \to x_0$ folgt, da $f(x) \to f(x_0)$,

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = a(f(x))\frac{f(x) - f(x_0)}{x - x_0} \to a(f(x_0))\, f'(x_0) = g'(f(x_0))f'(x_0).$$

□

Satz 9.4 (Differenzierbarkeit der Umkehrfunktion) *Sei $f : I \to \mathbb{R}$ streng monoton und stetig auf dem offenen Intervall I, und differenzierbar in x_0 mit $f'(x_0) \neq 0$. Dann ist $I^* = f(I)$ ein offenes Intervall, und die Umkehrfunktion $g : I^* \to I$ ist differenzierbar in $y_0 = f(x_0)$ mit Ableitung*

$$g'(y_0) = \frac{1}{f'\big(g(y_0)\big)}.$$

Beweis: Nach Satz 8.2 in Kapitel 3 mit Zusatz (8.1) ist I^* ein offenes Intervall und $g : I^* \to \mathbb{R}$ streng monoton und stetig, insbesondere $g(y) \to g(y_0) = x_0$ mit $y \to y_0$. Für $y \neq y_0$ folgt

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{g(y) - g(y_0)}{f(g(y)) - f(g(y_0))} = \frac{1}{\dfrac{f(g(y)) - f(g(y_0))}{g(y) - g(y_0)}} \to \frac{1}{f'(x_0)}.$$

□

Seien I, I^* offene Intervalle, $f : I \to I^*$ bijektiv und in x_0 differenzierbar. Ist die Umkehrfunktion $g : I^* \to I$ in $y_0 = f(x_0)$ auch differenzierbar, so folgt schon aus der Kettenregel

$$g(f(x)) = x \quad \Rightarrow \quad g'(f(x_0))f'(x_0) = 1.$$

Insbesondere muss $f'(x_0) \neq 0$ sein, und es gilt die Formel für die Ableitung der Umkehrfunktion. Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$, ist streng monoton wachsend und stetig; ihre Umkehrfunktion lautet

$$g : \mathbb{R} \to \mathbb{R},\ g(y) = \begin{cases} y^{1/3} & \text{für } y \geq 0 \\ -|y|^{1/3} & \text{für } y \leq 0. \end{cases}$$

Die Umkehrfunktion ist nicht differenzierbar in $y = 0$, denn es gilt

$$\frac{g(y) - g(0)}{y - 0} = |y|^{-2/3} \to +\infty \text{ mit } y \to 0.$$

Die Voraussetzung des Satzes ist hier verletzt, es ist $f'(0) = 0$.

Die beiden vorangegangenen Regeln sind in der von Leibniz eingeführten Notation besonders suggestiv. Er schreibt Funktionen in der Form $y = y(x)$ und bezeichnet die Ableitung mit dem Symbol $\frac{dy}{dx}$, das auch als *Differentialquotient* bezeichnet wird. Formal ergeben sich Kettenregel und Ableitung der Umkehrfunktion aus den Regeln der Bruchrechnung:

$$\begin{aligned} y = y(x),\ z = z(y) \quad &\Rightarrow \quad \frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}, \\ y = y(x),\ x = x(y) \quad &\Rightarrow \quad \frac{dx}{dy} = \left(\frac{dy}{dx}\right)^{-1}. \end{aligned}$$

Bei dieser saloppen Notation ist jedoch darauf zu achten, wo die jeweiligen Funktionen definiert sind. In jedem Fall ist die Bezeichnung $\frac{d}{dx}$ für den Ableitungsoperator üblich und praktisch.

Differenzierbarkeit kann als Approximierbarkeit durch eine affinlineare Funktion interpretiert werden, wobei der Fehler schneller als linear verschwindet. Diese Deutung wird uns bei Funktionen mehrerer Variabler erneut begegnen. Genauer ist Folgendes gemeint.

Lemma 9.2 *Genau dann hat $f : I \to \mathbb{R}^n$ in $x_0 \in I$ die Ableitung $a \in \mathbb{R}^n$, wenn gilt:*

$$\lim_{x \to x_0} \frac{f(x) - \big(f(x_0) + a(x - x_0)\big)}{x - x_0} = 0. \tag{9.2}$$

Beweis: Folgt sofort aus der Umformung

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

□

Um das asymptotische Verhalten von zwei Funktionen f und g für $x \to x_0$ zu vergleichen, werden oft die Landauschen Symbole benutzt:

$$\begin{aligned} f = o(g) \text{ für } x \to x_0 \quad &\Leftrightarrow \quad \lim_{x \to x_0} \frac{|f(x)|}{|g(x)|} = 0, \\ f = \mathcal{O}(g) \text{ für } x \to x_0 \quad &\Leftrightarrow \quad \limsup_{x \to x_0} \frac{|f(x)|}{|g(x)|} < \infty. \end{aligned}$$

Es handelt sich um eine symbolische Notation, die Terme $o(g)$ bzw. $\mathcal{O}(g)$ sind nicht wirklich Funktionen. Die Gleichung $f'(x_0) = a$ kann äquivalent wie folgt geschrieben werden:

$$f(x) = f(x_0) + a(x - x_0) + o(x - x_0) \qquad \text{für } x \to x_0.$$

Diese Notation ist auch in der Physik beliebt. Allerdings ist Vorsicht angesagt, wenn die Funktion f noch von anderen Parametern abhängt.

10 Mittelwertsatz

Es ist eine zentrales Problem der Analysis, aus Eigenschaften der Ableitung auf Eigenschaften der Funktion selbst zu schließen. Der Mittelwertsatz ist dafür ein einfaches und effektives Hilfsmittel. Für seinen Beweis müssen wir allerdings erst über Extremwerte – Maxima und Minima – von stetigen Funktionen sprechen. Da sich keine Unterschiede ergeben, betrachten wir dabei gleich reellwertige Funktionen auf einer Teilmenge des $\mathbb{R}^n$.

Wir benötigen einen Satz, der die Existenz von Extremalstellen für eine stetige Funktion $f : D \to \mathbb{R}$ allgemein garantiert. Oft wird das so formuliert, dass die Funktion ihr Minimum bzw. Maximum annimmt. Das ist allerdings misverständlich, denn diese Bezeichnungen implizieren schon die Existenz der Extremalstellen; die richtige Bezeichnung ist Infimum bzw. Supremum. Jedenfalls brauchen wir für die Existenz der Extremalstellen auch Voraussetzungen an den Definitionsbereich D, wie das folgende Beispiel zeigt:

$$f : D = (0, \infty) \to \mathbb{R},\ f(x) = \frac{1}{1+x}.$$

Hier wird weder das Supremum $\sup_{x \in D} f(x) = 1$ noch das Infimum $\inf_{x \in D} f(x) = 0$ durch die Funktion angenommen.

Satz 10.1 (Existenz von Extremalstellen) *Sei $D \subset \mathbb{R}^n$ nichtleer, abgeschlossen und beschränkt, und sei $f : D \to \mathbb{R}$ stetig. Dann ist f beschränkt und nimmt ihr Infimum und Supremum an, das heißt es gibt $x_0, x_1 \in D$ mit*

$$f(x_0) = \inf_{x \in D} f(x) \quad \text{und} \quad f(x_1) = \sup_{x \in D} f(x).$$

Beweis: Setze $\alpha = \inf_{x \in D} f(x) \in [-\infty, \infty)$. Wir zeigen die Existenz eines $x_0 \in D$ mit $f(x_0) = \alpha$, daraus folgt insbesondere $\alpha > -\infty$, das heißt f ist nach unten beschränkt. Für das Supremum kann analog argumentiert werden.

Nach Definition des Infimums gibt es eine Folge $x_k \in D$ mit $f(x_k) \to \alpha$. Da D beschränkt, ist die Folge x_k beschränkt. Nach Bolzano-Weierstraß, siehe Satz 4.7 bzw. Folgerung 5.3 in Kapitel 2, gibt es eine Teilfolge x_{k_j} und ein $x_0 \in \mathbb{R}^n$ mit $x_{k_j} \to x_0$ für $j \to \infty$. Aus D abgeschlossen, siehe Definition 5.11 in Kapitel 2, folgt $x_0 \in D$. Aber dann gilt wegen der Stetigkeit von f

$$-\infty < f(x_0) = \lim_{j \to \infty} f(x_{k_j}) = \alpha.$$

□

Im Beweis spielte die folgende Eigenschaft von D eine Rolle:

Definition 10.1 (Folgenkompaktheit) *Eine Menge $D \subset \mathbb{R}^n$ heißt folgenkompakt, wenn es zu jeder Folge $x_k \in D$ eine Teilfolge x_{k_j} gibt, die konvergiert mit Grenzwert in D:*

$$\lim_{j \to \infty} x_{k_j} = x_0 \in D.$$

Wir halten aus dem Beweis von Satz 10.1 folgende Beobachtung fest:

Satz 10.2 *Für $D \subset \mathbb{R}^n$ sind äquivalent:*

(1) *D ist folgenkompakt.*

(2) *D ist abgeschlossen und beschränkt.*

Beweis: Die Implikation (2) $\Rightarrow$ (1) wurde im Beweis von Satz 10.1 gezeigt. Jetzt setzen wir (1) voraus. Um die Abgeschlossenheit von D zu zeigen, sei $x_k \in D$ eine beliebige Folge mit $x_k \to x_0 \in \mathbb{R}^n$. Da D folgenkompakt, gibt es eine Teilfolge x_{k_j}, die gegen ein $x_0' \in D$ konvergiert. Aus der Eindeutigkeit des Grenzwerts folgt $x_0 = x_0' \in D$, also ist D abgeschlossen. Wäre D nicht beschränkt, so gibt es zu $k \in \mathbb{N}$ ein $x_k \in D$ mit $|x_k| \geq k$. Da D kompakt, existiert eine konvergente Teilfolge x_{k_j}, und diese ist nach Satz 3.4 in Kapitel 2 beschränkt, ein Widerspruch. $\square$

Wir kommen jetzt wieder zurück zur eindimensionalen Situation.

Definition 10.2 (Lokale Extrema) *Die Funktion $f : (a,b) \to \mathbb{R}$ hat in $x_0 \in (a,b)$ ein lokales Minimum, falls es ein $\delta > 0$ gibt, so dass gilt:*

$$f(x_0) \leq f(x) \quad \text{für alle } x \in U_\delta(x_0) = (x_0 - \delta, x_0 + \delta).$$

Ist sogar $f(x_0) < f(x)$ für $x \in U_\delta(x_0)$, $x \neq x_0$, so heißt das lokale Minimum isoliert. Ein (isoliertes) lokales Maximum ist entsprechend definiert.

Satz 10.3 (notwendige Bedingung für Extrema) *Die Funktion $f : (a,b) \to \mathbb{R}$ habe in $x_0 \in (a,b)$ ein lokales Extremum. Ist f in x_0 differenzierbar, so gilt $f'(x_0) = 0$.*

Beweis: Sei x_0 lokales Minimum von f. Dann gibt es ein $\delta > 0$ mit $f(x) \geq f(x_0)$ für $x \in U_\delta(x_0)$, also gilt

$$\frac{f(x) - f(x_0)}{x - x_0} \quad \begin{cases} \geq 0 & \text{für } x \in (x_0, x_0 + \delta), \\ \leq 0 & \text{für } x \in (x_0 - \delta, x_0). \end{cases}$$

Mit $x \searrow x_0$ folgt $f'(x_0) \geq 0$, mit $x \nearrow x_0$ folgt $f'(x_0) \leq 0$. $\square$

Die Funktion $f(x) = x^3$ erfüllt $f'(0) = 0$, aber in $x = 0$ liegt kein lokales Extremum vor. Die Bedingung $f'(x_0) = 0$ ist notwendig für eine lokale Extremalstelle einer differenzierbaren Funktion, aber sie ist nicht hinreichend. Im Beweis von Satz 10.3 wird tatsächlich eine Aussage für Extrema am Rand gezeigt. Hat zum Beispiel $f : [x_0, x_1] \to \mathbb{R}$ ein lokales Extremum im Punkt x_0, so folgt

$$f_+'(x_0) = \lim_{x \searrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \begin{cases} \geq 0 & \text{(Minimum)} \\ \leq 0 & \text{(Maximum).} \end{cases}$$

Satz 10.4 (Mittelwertsatz der Differentialrechnung) *Sei $f : [a,b] \to \mathbb{R}$ stetig, differenzierbar auf $(a,b) \neq \emptyset$. Dann gibt es ein $\xi \in (a,b)$ mit*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Beweis: Wir zeigen die Behauptung zuerst im Fall $f(a) = f(b) = 0$ (Satz von Rolle). Wir brauchen dann ein $\xi \in (a,b)$ mit $f'(\xi) = 0$. Nach Satz 10.1 gibt es $\xi_1, \xi_2 \in [a,b]$ mit mit

$$f(\xi_1) = \inf_{x \in [a,b]} f(x) \quad \text{und} \quad f(\xi_2) = \sup_{x \in [a,b]} f(x).$$

Ist $\xi_1 \in (a,b)$, so folgt $f'(\xi_1) = 0$ nach Satz 10.3 und wir können $\xi = \xi_1$ wählen. Analog, wenn $\xi_2 \in (a,b)$. Im verbleibenden Fall $\xi_1, \xi_2 \in \{a,b\}$ folgt $\inf f = \sup f = 0$ bzw. $f(x) = 0$ für alle $x \in [a,b]$, und damit auch $f'(x) = 0$ für alle x. Seien nun $f(a), f(b)$ beliebig. Definiere $h : [a,b] \to \mathbb{R}$ durch Abziehen der Sekante:

$$h(x) = f(x) - \left(f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \right).$$

Es gilt $h(a) = h(b) = 0$. Also existiert ein $\xi \in (a,b)$ mit

$$0 = h'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}.$$

□

Folgerung 10.1 (Monotoniekriterium) *Sei f differenzierbar auf (a,b), stetig auf $[a,b]$. Dann gelten folgende Aussagen:*

$$\begin{aligned}
f'(x) = 0 \text{ für alle } x \in (a,b) &\quad\Rightarrow\quad f \text{ ist konstant auf } [a,b] \\
f'(x) \geq 0 \text{ für alle } x \in (a,b) &\quad\Rightarrow\quad f \text{ ist wachsend auf } [a,b] \\
f'(x) \leq 0 \text{ für alle } x \in (a,b) &\quad\Rightarrow\quad f \text{ ist fallend auf } [a,b].
\end{aligned}$$

Bei strikter Ungleichung folgt strenge Monotonie auf $[a,b]$.

Beweis: Sei $a \leq x_1 < x_2 \leq b$. Nach dem Mittelwertsatz gibt es ein $\xi \in (x_1, x_2)$, so dass gilt:

$$f(x_2) - f(x_1) = f'(\xi) \underbrace{(x_2 - x_1)}_{>0} \begin{cases} = 0 & \text{wenn } f'(\xi) = 0 \\ \geq 0 & \text{wenn } f'(\xi) \geq 0 \\ > 0 & \text{wenn } f'(\xi) > 0 \\ \leq 0 & \text{wenn } f'(\xi) \leq 0 \\ < 0 & \text{wenn } f'(\xi) < 0 \end{cases}.$$

□

Folgerung 10.2 (Schrankensatz) *Ist $f : [a,b] \to \mathbb{R}$ stetig und differenzierbar auf (a,b), so gilt für $a \leq x_1 < x_2 \leq b$:*

$$\begin{aligned}
f'(x) \geq m \text{ für alle } x \in (a,b) &\quad\Rightarrow\quad \frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq m \\
f'(x) \leq M \text{ für alle } x \in (a,b) &\quad\Rightarrow\quad \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq M.
\end{aligned}$$

Beweis: Wir zeigen die erste Aussage. Mit $g(x) = mx$ gilt $(f-g)' = f' - g' \geq m - m = 0$. Nach Folgerung 10.1 ist $f-g$ wachsend, ds heißt $f(x_2) - m\,x_2 \geq f(x_1) - m\,x_1$. Die Behauptung folgt. $\square$

Der Mittelwertsatz gilt nicht für vektorwertige Funktionen $f : [a,b] \to \mathbb{R}^n$, $n \geq 2$. Ein ad hoc Beispiel ist $f(x) = (x^2, x^3)$ auf $[a,b] = [0,1]$. Angenommen es gibt ein $\xi \in (0,1)$ mit

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Dann folgt $2\xi = 1$ und $3\xi^2 = 1$, ein Widerspruch. Das Problem ist, dass für die einzelnen Koordinaten im allgemeinen nicht dieselbe Stelle ξ gewählt werden kann. Dennoch können wir auch im vektorwertigen Fall eine Version des Schrankensatzes beweisen.

Folgerung 10.3 (f' beschränkt $\Rightarrow$ f Lipschitz) *Sei $f : [a,b] \to \mathbb{R}^n$ stetig und auf (a,b) differenzierbar. Falls $\|f'(x)\| \leq L$ für alle $x \in (a,b)$, so folgt*

$$\|f(x_2) - f(x_1)\| \leq L|x_2 - x_1| \quad \text{für alle } x_1, x_2 \in [a,b].$$

Beweis: Um die Aussage auf den reellwertigen Fall zu reduzieren, betrachten wir für einen beliebigen Vektor $v \in \mathbb{R}^n$ mit $\|v\| = 1$ die Funktion

$$\varphi : [a,b] \to \mathbb{R},\ \varphi(x) = \langle v, f(x) \rangle = \sum_{i=1}^{n} v_i f_i(x).$$

Aus der Ungleichung von Cauchy-Schwarz, siehe Satz 5.5 in Kapitel 2, folgt

$$\varphi'(x) = \langle v, f'(x) \rangle \leq \|v\|\,\|f'(x)\| \leq L.$$

Für $x_1, x_2 \in [a,b]$ mit $x_1 < x_2$ erhalten wir aus Folgerung 10.2

$$\langle v, f(x_2) - f(x_1) \rangle = \varphi(x_2) - \varphi(x_1) \leq L(x_2 - x_1) = L|x_2 - x_1|.$$

Wir können $f(x_2) \neq f(x_1)$ annehmen, und wählen $v = (f(x_2) - f(x_1))/\|f(x_2) - f(x_1)\|$. Das ist ein Einheitsvektor, also folgt

$$\|f(x_2) - f(x_1)\| = \Big\langle \frac{f(x_2) - f(x_1)}{\|f(x_2) - f(x_1)\|}, f(x_2) - f(x_1) \Big\rangle = \langle v, f(x_2) - f(x_1) \rangle \leq L|x_1 - x_2|.$$

$\square$

Wir kommen jetzt zu höheren Ableitungen.

Definition 10.3 *Die k-te Ableitung von $f : (a,b) \to \mathbb{R}$ in x_0 ist induktiv definiert durch*

$$f^{(k)}(x_0) = \big(f^{(k-1)}\big)'(x_0).$$

Damit $f^{(k)}(x_0)$ definiert ist, müssen also die Ableitungen bis Ordnung $k-1$ in einer Umgebung von x_0 definiert sein, und $f^{(k-1)}$ muss in x_0 differenzierbar sein.

Natürlich schreiben wir f' und f'' statt $f^{(1)}$ bzw. $f^{(2)}$. Mit der zweiten Ableitung gewinnen wir genauere Informationen über lokale Extrema.

Satz 10.5 *Sei $f:(a,b)\to\mathbb{R}$ differenzierbar. In $x_0\in(a,b)$ gelte $f'(x_0)=0$, und $f''(x_0)$ sei definiert. Dann gilt:*

(1) *Ist $f''(x_0)>0$, so hat f in x_0 ein isoliertes, lokales Minimum.*

(2) *Hat f in x_0 ein lokales Minimum, so folgt $f''(x_0)\geq 0$.*

Analoge Aussagen gelten mit umgekehrten Ungleichungen für Maxima.

Beweis: Da $f'(x_0)=0$ nach Voraussetzung, gilt

$$f''(x_0)=\lim_{x\to x_0}\frac{f'(x)-f'(x_0)}{x-x_0}=\lim_{x\to x_0}\frac{f'(x)}{x-x_0}.$$

Unter der Voraussetzung (1) gibt es ein $\delta>0$ mit $f'(x)<0$ auf $(x_0-\delta,x_0)$ und $f'(x)>0$ auf $(x_0,x_0+\delta)$. Nach Folgerung 10.1 ist f dann streng monoton fallend auf $(x_0-\delta,x_0)$ und und streng monoton wachsend auf $(x_0,x_0+\delta)$, hat also in x_0 ein isoliertes, lokales Minimum. Wäre $f''(x_0)<0$ in (2), so hätte f in x_0 ein isoliertes lokales Maximum im Widerspruch zur Voraussetzung. □

Die Funktion $f(x)=x^4$ zeigt, dass in einem isolierten Minimum $f''(x_0)=0$ gelten kann. Häufig ist man nicht wirklich an den lokalen Minima, sondern am globalen Minimum interessiert. Dafür ist der Begriff der Konvexität relevant.

Definition 10.4 *Sei $I\subset\mathbb{R}$ ein Intervall. $f:I\to\mathbb{R}$ heißt konvex, falls gilt:*

$$f\big((1-t)x_0+tx_1\big)\leq(1-t)f(x_0)+tf(x_1)\quad\forall x_0,x_1\in I,\,t\in[0,1]. \tag{10.1}$$

Gilt dies mit $\geq$ statt mit $\leq$, so heißt f konkav.

Die Sekante durch $\big(x_0,f(x_0)\big)$ und $\big(x_1,f(x_1)\big)$ hat die Geradengleichung

$$y=f(x_0)+\frac{f(x_1)-f(x_0)}{x_1-x_0}(x-x_0)=:g(x).$$

An der Stelle $x(t)=(1-t)x_0+tx_1$ gilt

$$g\big(x(t)\big)=f(x_0)+\frac{f(x_1)-f(x_0)}{x_1-x_0}\,t(x_1-x_0)=(1-t)f(x_0)+tf(x_1).$$

Die Konvexität bedeutet also, dass der Graph von f stets unterhalb der Sekanten liegt,

Satz 10.6 (Konvexitätskriterien) *Sei $f:(a,b)\to\mathbb{R}$ differenzierbar. Dann sind folgende Aussagen äquivalent:*

(1) *f' ist monoton wachsend auf (a,b).*

(2) *f ist konvex.*

(3) *$f(x_1)\geq f(x_0)+f'(x_0)(x_1-x_0)$ für alle $x_0,x_1\in(a,b)$.*

Ist f zweimal differenzierbar auf (a,b), so ist außerdem äquivalent:

(4) *$f''\geq 0$.*

Beweis: Wir zeigen (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1). Sei (1) erfüllt. Angenommen f ist nicht konvex, das heißt es gibt $x_0, x_1 \in I$ und ein $t \in [0,1]$ mit

$$g(t) = f\big((1-t)x_0 + tx_1\big) - \big((1-t)f(x_0) + tf(x_1)\big) > 0.$$

Wir können $x_1 > x_0$ annehmen, sonst vertausche x_0, x_1 und ersetze t durch $1-t$. Die Funktion $g : [0,1] \to \mathbb{R}$ ist stetig mit $g(0) = g(1) = 0$. Nach Satz 10.1 nimmt g ihr strikt positives Maximum in einem $t_0 \in (a,b)$ an, und es gilt $g'(t_0) = 0$ nach Satz 10.3. Aber wegen f' wachsend ist auch g' wachsend:

$$g'(t) = f'\big(\underbrace{x_0 + t(x_1 - x_0)}_{\text{wachsend}}\big)\underbrace{(x_1 - x_0)}_{>0} - \big(\underbrace{f(x_1) - f(x_0)}_{\text{konstant}}\big).$$

Insbesondere folgt $g'(t) \geq 0$ für alle $t \geq t_0$, und daraus $g(1) \geq g(t_0) > 0$, Widerspruch.

Es gelte jetzt (2), also gilt für alle $t \in (0,1)$ und $x_0, x_1 \in (a,b)$ mit $x_0 \neq x_1$

$$(x_1 - x_0)\,\frac{f\big(x_0 + t(x_1 - x_0)\big) - f(x_0)}{t(x_1 - x_0)} \leq f(x_1) - f(x_0).$$

Mit $t \searrow 0$ folgt $(x_1 - x_0)f'(x_0) \leq f(x_1) - f(x_0)$, das ist Ungleichung (3). Schließlich folgt aus (3), indem wir dort x_0 und x_1 vertauschen und addieren,

$$\big(f'(x_1) - f'(x_0)\big)(x_1 - x_0) \geq 0,$$

womit wiederum (1) gezeigt ist. Schließlich: ist f zweimal differenzierbar, so impliziert $f'' \geq 0$ Bedingung (1) nach Folgerung 10.1, und umgekehrt folgt $f'' \geq 0$ aus (1) einfach durch Betrachtung des Differenzenquotienten. $\square$

Beispiel 10.1 (Youngsche Ungleichung) Für $x, y \geq 0$ gilt die Ungleichung

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q} \quad \text{falls } p,q \in (1,\infty) \text{ mit } \frac{1}{p} + \frac{1}{q} = 1.$$

Dazu betrachten wir für festes $y > 0$ die Funktion

$$f : (0,\infty) \to \mathbb{R},\ f(x) = xy - \frac{x^p}{p}.$$

Es gilt $f'(x) = y - x^{p-1}$ und $f''(x) = -(p-1)x^{p-2} \leq 0$, das heißt f ist konkav nach Satz 10.6. Aber $f'(x_0) = 0$ für $x_0 = y^{1/(p-1)}$, also folgt mit Satz 10.6(3) wegen $p/(p-1) = q$

$$xy - \frac{x^p}{p} = f(x) \leq f(x_0) + f'(x_0)(x - x_0) = y^{1/(p-1)}y - \frac{y^{p/(p-1)}}{p} = \frac{y^q}{q}.$$

Definition 10.5 *Sei $I \subset \mathbb{R}$ ein offenes Intervall und $k \in \mathbb{N}_0$. Wir bezeichnen mit $C^k(I)$ den $\mathbb{R}$-Vektorraum der k mal stetig differenzierbaren Funktionen $f : I \to \mathbb{R}$, das heißt*

$$C^k(I) = \{f : I \to \mathbb{R} : f^{(i)} : I \to \mathbb{R} \text{ sind definiert und stetig für } i = 0,1,\ldots,k\}.$$

Weiter definieren wir $C^\infty(I)$ als den $\mathbb{R}$-Vektorraum der unendlich oft differenzierbaren Funktionen, also ist

$$C^\infty(I) = \bigcap_{k \geq 0} C^k(I).$$

Der Umgang mit C^∞-Funktionen ist besonders angenehm, weil die Klasse im Gegensatz zu den Räumen $C^k(I)$ unter der Bildung von Ableitungen abgeschlossen ist. Es ist klar, dass Polynome unendlich oft differenzierbar sind; es gibt aber sehr viel mehr Funktionen in $C^\infty(I)$. Wir werden weitere im nächsten Abschnitt mithilfe der Exponentialfunktion konstruieren.

11 Die reelle Exponentialfunktion

In Beispiel 4.2 wurde das diskrete Modell des Zinseszins betrachtet. Das zugrundeliegende Prinzip war: der Zugewinn im Zeitintervall $[t, t+\Delta t]$ ist proportional zum Kontostand $f(t)$ und zum Zeitschritt Δt mit einer Rate $\lambda \in \mathbb{R}$, das heißt

$$f(t+\Delta t) = f(t)(1+\lambda \Delta t) \quad \text{bzw.} \quad \frac{f(t+\Delta t) - f(t)}{\Delta t} = \lambda f(t).$$

Mit $\Delta t \to 0$ ergibt sich formal das kontinuierliche Wachstumsgesetz

$$f'(t) = \lambda f(t).$$

Wachstums- oder Zerfallsprozesse von diesem Typ spielen in vielen Bereichen der Natur und der Anwendungen eine Rolle. Dementsprechend ist die zugrundeliegende Funktion von fundamentaler Bedeutung: es ist die Exponentialfunktion.

Satz 11.1 (reelle Exponentialfunktion) *Es gibt genau eine Funktion $f \in C^1(\mathbb{R})$ mit*

$$f' = f \quad \textit{auf } \mathbb{R}, \quad f(0) = 1. \tag{11.1}$$

Diese heißt Exponentialfunktion exp, *und sie hat folgende Eigenschaften:*

(a) exp *ist in $C^\infty(\mathbb{R})$, strikt positiv und streng monoton wachsend mit Grenzwerten*

$$\lim_{x\to\infty} \exp(x) = +\infty \quad \textit{und} \quad \lim_{x\to-\infty} \exp(x) = 0.$$

(b) *Es gilt die Funktionalgleichung*

$$\exp(x+y) = \exp(x)\exp(y) \quad \textit{für alle } x, y \in \mathbb{R},$$

insbesondere $\exp(x)\exp(-x) = 1$.

Hinweis. *Für die Existenz der Lösung von (11.1) siehe Satz 13.6.*

Beweis: Sei $\exp \in C^1(\mathbb{R})$ Lösung von (11.1). Ist $\exp \in C^k(\mathbb{R})$ für $k \geq 1$ schon gezeigt, so folgt aus der Differentialgleichung $\exp' = \exp \in C^k(\mathbb{R})$, also $\exp \in C^{k+1}(\mathbb{R})$. Somit ist exp beliebig oft stetig differenzierbar. Weiter gilt

$$\frac{d}{dx}\big(\exp(x)\exp(-x)\big) = \exp'(x)\exp(-x) - \exp(x)\exp'(-x) = 0.$$

Es folgt $\exp(x)\exp(-x) = \exp(0)^2 = 1$ für alle $x \in \mathbb{R}$. Nun ist $\exp(0) = 1$. Wäre $\exp(x) < 0$ für ein $x \in \mathbb{R}$, so hätte exp eine Nullstelle nach dem Zwischenwertsatz, Widerspruch. Somit gilt $\exp : \mathbb{R} \to (0, \infty)$. Aber dann ist $\exp' = \exp > 0$, also exp streng monoton wachsend. Für $x \geq 0$ folgt $\exp'(x) = \exp(x) \geq \exp(0) = 1$, also mit dem Schrankensatz

$$\exp(x) \geq 1 + x \to \infty \quad \text{mit } x \to \infty. \tag{11.2}$$

Für $x < 0$ folgt $\exp(x) = 1/\exp(-x) \to 0$ mit $x \to -\infty$. Damit sind alle Aussagen von Behauptung (a) bewiesen.

Wir zeigen nun, dass das Anfangswertproblem (11.1) höchstens eine Lösung hat. Für $g \in C^1(\mathbb{R})$ liefert die Quotientenregel

$$\left(\frac{g}{\exp}\right)' = \frac{g' \exp - g \exp'}{\exp^2}.$$

Somit gilt der Schluss

$$g \in C^1(\mathbb{R}) \text{ mit } g' = g \text{ auf } \mathbb{R} \quad \Rightarrow \quad g(x) = g(0)\exp(x) \text{ für alle } x \in \mathbb{R}. \tag{11.3}$$

Ist zusätzlich $g(0) = 1$, so folgt $g(x) = \exp(x)$, die Eindeutigkeit. Für (b) betrachte die Funktion $u(x) := \exp(x+y)$. Es gilt $u' = u$ mit $' = \frac{d}{dx}$, also folgt aus (11.3)

$$\exp(x+y) = u(x) = u(0)\exp(x) = \exp(x)\exp(y).$$

□

Es ist naheligend, das diskrete und das kontinuierliche Wachstumsmodell zu vergleichen. Unterteilen wir den Zeitraum von x Jahren in n Abschnitte der Dauer $\Delta t = x/n$, so liefert das diskrete Wachstum mit Rate Eins und Anfangswert $f_n(0) = 1$ per Induktion

$$f_n(x) = (1 + \Delta t)^n = \left(1 + \frac{x}{n}\right)^n.$$

Satz 11.2 (Eulerapproximation) *Es gilt:*

(a) $\exp(x) = \lim_{n\to\infty}(1 + \frac{x}{n})^n$ *für alle* $x \in \mathbb{R}$.

(b) $\exp(x) = e^x$ *für alle* $x \in \mathbb{Q}$, *insbesondere* $\exp(1) = e = 2,71828\ldots$.

Beweis: Wir bemerken vorab $\exp(nx) = \exp(x)^n$ für $n \in \mathbb{N}$, dies folgt aus der Funktionalgleichung durch Induktion. Sei nun zunächst $x \geq 0$. Dann gilt mit (11.2)

$$\exp(x) = \exp\left(\frac{x}{n}\right)^n \geq \left(1 + \frac{x}{n}\right)^n =: f_n(x).$$

Die f_n erfüllen andererseits für $x \in [0,k]$ die Abschätzung

$$f_n'(x) = \left(1 + \frac{x}{n}\right)^{-1} f_n(x) \geq \left(1 + \frac{k}{n}\right)^{-1} f_n(x) = \frac{n}{n+k} f_n(x).$$

Vergleich mit $g_n(x) = \exp(\frac{n}{n+k}\,x)$ liefert

$$\left(\frac{f_n}{g_n}\right)' = \frac{f_n' g_n - f_n g_n'}{g_n^2} \geq \frac{n}{n+k}\,\frac{f_n g_n - f_n g_n}{g_n^2} = 0.$$

Es folgt, da $f_n(0) = g_n(0) = 1$ und exp stetig in x,

$$f_n(x) \geq g_n(x) = \exp\left(\frac{nx}{n+k}\right) \to \exp(x) \quad \text{mit } n \to \infty.$$

Dies zeigt $\lim_{n\to\infty} f_n(x) = \exp(x)$ für $x \geq 0$. Für $n \geq |x|$ gilt nach Bernoulli, Satz 2.2,

$$1 \geq \left(1 + \frac{x}{n}\right)^n \left(1 - \frac{x}{n}\right)^n = \left(1 - \frac{x^2}{n^2}\right)^n \geq 1 - n \cdot \frac{x^2}{n^2} \to 1.$$

Also folgt mit Satz 11.1(b) für $x < 0$

$$f_n(x) = \frac{1}{f_n(-x)} \left(1 - \frac{x^2}{n^2}\right)^n \to \frac{1}{\exp(-x)} = \exp(x).$$

Damit ist Behauptung (a) gezeigt. Für $q \in \mathbb{N}$, $p \in \mathbb{Z}$ gilt nun mit der Funktionalgleichung

$$\big(\exp \frac{1}{q}\big)^q = \exp(1) \text{ bzw. } \exp \frac{1}{q} = \exp(1)^{\frac{1}{q}}, \quad \text{und} \quad \exp \frac{p}{q} = \big(\exp \frac{1}{q}\big)^p = \exp(1)^{\frac{p}{q}}.$$

Aber nach (a) ist $\exp(1) = \lim_{n\to\infty}(1 + \frac{1}{n})^n = e = 2,71828\ldots$, siehe Definition 4.3. □

Wir können die Funktionalgleichung auch anschaulich aus dem Zins-Modell verstehen: Anlegen des Startkapitals $f(0)$ mit jährlichem Zinssatz $x = 1$ ergibt nach der Zeit t

$$f(t) = f(0) \lim_{n\to\infty} (1 + \frac{t}{n})^n = f(0)\exp(t).$$

Bei konstantem Zinssatz muss es egal sein, ob das Geld erst für einen Zeitraum t_1 angelegt wird und dann die ersparte Summe wieder für einen Zeitraum t_2 , oder eben gleich für die Gesamtdauer $t_1 + t_2$. Genau das besagt die Funktionalgleichung:

$$f(0)\ \exp(t_1 + t_2) = f(t_1 + t_2) = f(t_1)\ \exp(t_2) = f(0)\exp(t_1)\exp(t_2).$$

In der Sprache der Algebra besagt die Funktionalgleichung, dass $\exp : (\mathbb{R}, +) \to (\mathbb{R}^+, \cdot)$ ein Gruppenhomomorphismus ist, wobei $\mathbb{R}^+ = \{r \in \mathbb{R} : r > 0\}$. Genauer ist exp ein Isomorphismus, wie wir gleich feststellen, der inverse Gruppenisomorphismus ist der Logarithmus.

Satz 11.3 (Definition des Logarithmus) $\exp : (-\infty, \infty) \to (0, \infty)$ *ist streng monoton wachsend und bijektiv. Die Umkehrfunktion* $\log : (0, \infty) \to (-\infty, \infty)$ *heißt (natürlicher) Logarithmus.* log *ist streng monoton wachsend und bijektiv, und hat die Grenzwerte*

$$\lim_{r\searrow 0} \log(r) = -\infty \quad \textit{und} \quad \lim_{r\to\infty} \log(r) = \infty. \tag{11.4}$$

Es gilt $\log(1) = 0$, $\log(e) = 1$, *und die Ableitung des Logarithmus ist*

$$\log'(r) = \frac{1}{r} \quad \textit{für alle } r \in (0, \infty). \tag{11.5}$$

Schließlich gilt die Funktionalgleichung

$$\log(rs) = \log(r) + \log(s) \quad \textit{für alle } r, s > 0. \tag{11.6}$$

Beweis: Nach Satz 8.2 ist exp bijektiv, die Umkehrfunktion log definiert und ebenfalls streng monoton wachsend. Die Grenzwerte (11.4) folgen aus Gleichung (8.1), die Funktionswerte aus $\exp(0) = 1$ bzw. $\exp(1) = e$. Wegen $\exp'(x) = \exp(x) > 0$ gilt nach Satz 9.4

$$\log'(r) = \frac{1}{\exp'(\log(r))} = \frac{1}{\exp(\log(r))} = \frac{1}{r}.$$

Die Funktionalgleichung (11.6) ergibt sich schließlich aus $\exp(x)\exp(y) = \exp(x + y)$, indem wir $x = \log r$, $y = \log(s)$ einsetzen und den Logarithmus nehmen. □

Bisher war die Potenz a^x nur für $x \in \mathbb{Q}$ definiert; mithilfe des Logarithmus können wir diese Einschänkung nun leicht abschaffen.

Definition 11.1 (Potenz mit reellen Exponenten) *Für $a > 0$, $x \in \mathbb{R}$ definieren wir*

$$a^x = \exp\left(x \log(a)\right).$$

Die Regeln der Potenzrechnung für $a, b > 0$

$$a^x\, a^y = a^{x+y} \qquad (a^x)^y = a^{xy}$$
$$\left(\frac{1}{a}\right)^x = a^{-x} \qquad a^x\, b^x = (ab)^x$$

ergeben sich leicht aus den Definitionen und den Funktionalgleichungen von exp bzw. log. Für $a > 0$, $x \in \mathbb{R}$ gelten die Ableitungsregeln

$$\frac{d}{dx} a^x = \log(a) a^x \quad \text{und} \quad \frac{d}{da} a^x = x a^{x-1}.$$

In Folgerung 10.2 haben wir gezeigt, dass Schranken an die Ableitung einer Funktion *integriert* werden können, sprich zu Abschätzungen der Funktion führen. Wir formulieren jetzt eine exponentielle Variante dieses Prinzips.

Folgerung 11.1 (exponentieller Schrankensatz) *Sei $f : [a, b] \to \mathbb{R}$ stetig, auf (a, b) differenzierbar, und sei $a \leq x_1 < x_2 \leq b$ sowie $\lambda \in \mathbb{R}$. Dann gilt:*

$$f' \leq \lambda f \ \textit{auf}\ (a,b) \quad \Rightarrow \quad f(x_2) \leq e^{\Lambda(x_2 - x_1)} f(x_1),$$
$$f' \geq \lambda f \ \textit{auf}\ (a,b) \quad \Rightarrow \quad f(x_2) \geq e^{\lambda(x_2 - x_1)} f(x_1).$$

Beweis: Wir berechnen mit der Produktregel, wobei $\frac{d}{dx}$ die Ableitung bezeichnet,

$$\frac{d}{dx}(e^{-\lambda x} f(x)) = -\lambda e^{-\lambda x} f(x) + e^{-\lambda x} f'(x) = e^{-\lambda x}(f'(x) - \lambda f(x)).$$

Verwende die Voraussetzung und Folgerung 10.1. □

Satz 11.4 (Wachstum von exp und log) *Es gelten folgende Aussagen:*

$$\lim_{x \to \infty} x^{-s} e^x = \infty \quad \textit{und} \quad \lim_{x \to \infty} x^s e^{-x} = 0 \quad \textit{für jedes } s > 0, \tag{11.7}$$
$$\lim_{y \to \infty} y^{-s} \log y = 0 \quad \textit{und} \quad \lim_{y \searrow 0} y^s \log y = 0 \quad \textit{für jedes } s > 0. \tag{11.8}$$

Beweis: Wir berechnen mit $f(x) = x^{-s} e^x$ für $x > 0$

$$f'(x) = -s x^{-s-1} e^x + x^{-s} e^x = x^{-s} e^x \big(1 - \frac{s}{x}\big) \geq \frac{1}{2} f(x) \quad \text{für } x \geq 2s =: x_0.$$

Mit Folgerung 11.1 ergibt sich

$$f(x) \geq f(x_0)\, e^{(x - x_0)/2} \to \infty \quad \text{für } x \to \infty.$$

Die zweite Aussage in (11.7) ergibt sich durch Übergang zum Kehrwert. Mit der Substitution $x = -s \log y$ folgt weiter für $y \searrow 0$

$$0 = \lim_{x \to \infty} e^{-x} x = \lim_{y \searrow 0} y^s (-s \log y).$$

Setzen wir $y = 1/z$ mit $z \to \infty$, so folgt auch der linke Grenzwert in (11.8). □

Folgerung 11.2 *Die Funktion*

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

ist unendlich oft differenzierbar.

Beweis: Wir zeigen durch Induktion, dass es Polynome p_n gibt mit

$$f^{(n)}(x) = \begin{cases} p_n(\frac{1}{x})\, e^{-1/x} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0. \end{cases}$$

Für $n = 0$ ist das richtig mit $p_0(s) \equiv 1$. Ist die Aussage für ein $n \in \mathbb{N}_0$ gezeigt, so folgt:

$$f^{(n+1)}(x) = \begin{cases} \big(\frac{1}{x^2}\, p_n(\frac{1}{x}) - \frac{1}{x^2}\, p_n'(\frac{1}{x})\big)\; e^{-1/x} & \text{für } x > 0, \\ 0 & \text{für } x < 0. \end{cases}$$

Für $x \neq 0$ gilt der Induktionsschluss also mit $p_{n+1}(s) = s^2\,(p_n(s) - p_n'(s))$. Zu zeigen bleibt $f^{(n+1)}(0) = 0$. Für den linksseitigen Differenzenquotienten ist das klar, für $x > 0$ berechnen wir mit (11.7)

$$\frac{f^{(n)}(x) - f^{(n)}(0)}{x} = \frac{1}{x}\, p_n\Big(\frac{1}{x}\Big)\, e^{-1/x} \to 0 \quad \text{mit } x \searrow 0.$$

□

12 Die trigonometrischen Funktionen

Die Differentialgleichung der harmonischen Schwingung lautet bei geeigneter Normierung $u'' + u = 0$. Physikalisch beschreibt $u : I = (t_1, t_2) \to \mathbb{R}$, $u = u(t)$, die Auslenkung eines schwingungsfähigen Systems (Oszillators) aus der Ruhelage $u = 0$, ein Beispiel ist die Schwingung einer elastischen Feder. Nach dem 2. Newtonschen Gesetz (1687) ist die Beschleunigung $u''(t)$ proportional zur Rückstellkraft des Oszillators, und diese ist im Fall der Feder proportional zur Auslenkung $u(t)$ nach dem Hookeschen Gesetz (1676). Mit geeigneter Wahl der Einheiten ergibt sich die Differentialgleichung. Bei einer Schwingung findet ein ständiger Austausch zwischen kinetischer und potentieller Energie statt. Bis auf Konstanten ist die kinetische Energie das Quadrat der Geschwindigkeit, die potentielle Energie das Quadrat der Auslenkung, dies wieder nach dem Hookeschen Gesetz. Die Gesamtenergie bleibt konstant. Das wollen wir jetzt zeigen, und auf die Frage der Eindeutigkeit anwenden.

Lemma 12.1 *Für Funktionen $u, v \in C^1(I)$ sind folgende Aussagen* (1) *und* (2) *äquivalent:*

(1) *$u \in C^2(I)$ ist Lösung von $u'' + u = 0$, und $v = -u'$.*

(2) *$c = u + iv \in C^1(I, \mathbb{C})$ ist Lösung von $c' = ic$, wobei $i = \sqrt{-1}$.*

Weiter: Für jede Lösung von (2) *ist $|c(t)|^2 = u(t)^2 + v(t)^2$ konstant.*

Beweis: Die Gleichung (2), also $c' = ic$, lautet in Koordinaten $c = u + iv$

$$u' + iv' = i(u + iv) \quad \Leftrightarrow \quad (u, v)' = (-v, u). \tag{12.1}$$

Aus (1) folgt $u' = -v$ und $v' = -u'' = u$, also (2). Aus (2) folgt wiederum $u' = -v$, insbesondere $u \in C^2(I)$, und $u'' = -v' = -u$. Schließlich berechnen wir mit (2)

$$\frac{d}{dt}|c|^2 = \frac{d}{dt}\big(u^2 + v^2\big) = 2uu' + 2vv' = 2u(-v) + 2vu = 0.$$

$\square$

Wir bezeichnen $c' = ic$ als Differentialgleichung der Kreisbewegung, denn nach Lemma 12.1 verlaufen die Lösungen auf einem Kreis mit einem gewissen Radius $R \geq 0$ um den Nullpunkt. Der Geschwindigkeitsvektor $c' = (-v, u)$ steht senkrecht auf $c = (u, v)$ und hat konstante Länge $|c'| = R$. Dadurch ist c' bis auf die Richtung eindeutig festgelegt, die ergibt sich aus

$$\det(c, c') = \det \begin{pmatrix} u & u' \\ v & v' \end{pmatrix} = uv' - u'v = u^2 + v^2 > 0 \quad \text{(außer wenn } R = 0\text{)}.$$

Anschaulich liegt die Kreisscheibe bei dieser Fahrtrichtung auf der linken Seite. Man sagt: der Kreis wird im mathematisch positiven Sinn durchlaufen, also gegen den Uhrzeigersinn. Wir zeigen jetzt, dass die Kreisbewegung durch die Anfangsposition eindeutig festgelegt ist, und folgern die Eindeutigkeit für das Anfagswertproblem der harmonischen Schwingung.

Satz 12.1 (Eindeutigkeit für Kreisbewegung/Schwingung) *Sei $t_0 \in I = (t_1, t_2)$.*

(1) *Zu $z_0 \in \mathbb{R}^2$ gibt es höchstens ein $c \in C^1(I, \mathbb{C})$ mit*

$$c' = ic \quad \text{auf } I, \quad c(t_0) = z_0. \tag{12.2}$$

(2) *Zu $x_0, y_0 \in \mathbb{R}$ gibt es höchstens ein $u \in C^2(I)$ mit*

$$u'' + u = 0 \quad \text{auf } I, \quad u(t_0) = x_0,\, u'(t_0) = y_0, \tag{12.3}$$

Beweis: Seien $c_{1,2} \in C^1(I,\mathbb{C})$ zwei Lösungen von (12.2). Dann gilt für $c = c_1 - c_2$

$$c' = (c_1 - c_2)' = i(c_1 - c_2) = ic, \quad c(t_0) = 0.$$

Aus Lemma 12.1 folgt $|c(t)|^2 = |c(t_0)|^2 = 0$ für alle t, also $c_1 = c_2$. Für (2) seien $u_{1,2} \in C^2(I)$ Lösungen von (12.3). Nach Lemma 12.1 sind dann $c_{1,2} = u_{1,2} - iu'_{1,2}$ Lösungen von $c'_{1,2} = ic_{1,2}$. Da $c_{1,2}(t_0) = u_{1,2}(t_0) - iu'_{1,2}(t_0) = x_0 - iy_0$, folgt $c_1 = c_2$ aus (1), also $u_1 = u_2$. □

Die reelle Differentialgleichung $f' = \lambda f$ mit $\lambda \in \mathbb{R}$ hat zum Anfangswert $f(0) = 1$ die Lösung $f(t) = e^{\lambda t}$. Die Gleichung $c' = ic$ mit Anfangswert $c(0) = 1$ ist formal identisch, allerdings ist der Wachstumskoeffizient imaginär und $c(t)$ komplexwertig. Jedenfalls ist die Bezeichnung $c(t) = e^{it}$ für die (gesuchte) Lösung sehr naheliegend. Wie im reellen Fall stellen wir das Problem der Existenz momentan zurück, und untersuchen erst die Eigenschaften der Funktion.

Satz 12.2 (Kreisbewegung) *Es gibt genau eine Funktion $c \in C^1(\mathbb{R},\mathbb{C})$ mit*

$$c' = ic \quad \text{auf } \mathbb{R}, \quad c(0) = 1. \tag{12.4}$$

Es ist $c \in C^\infty(\mathbb{R},\mathbb{C})$. Wir haben folgende Eigenschaften, wobei wir $c(t) =: e^{it}$ schreiben:

(a) $|e^{it}| = 1$ *für alle $t \in \mathbb{R}$.*

(b) $e^{i(s+t)} = e^{is}e^{it}$ *für alle $s, t \in \mathbb{R}$ (Funktionalgleichung).*

(c) $\overline{e^{it}} = e^{-it}$.

(d) $t \mapsto e^{it}$ *ist periodisch.*

Hinweis. *Für die Existenz der Lösung von (12.4) siehe Satz 13.6.*

Beweis: Die Eindeutigkeit gilt nach Satz 12.1. Ist schon $c \in C^k(\mathbb{R},\mathbb{C})$, so folgt $c' = ic \in C^k(\mathbb{R},\mathbb{C})$ und damit $c \in C^{k+1}(\mathbb{R},\mathbb{C})$, also $c \in C^\infty(\mathbb{R},\mathbb{C})$. Nach Lemma 12.1 gilt $|c(t)| = |c(0)| = 1$, das ist Behauptung (a). Für die Funktionalgleichung berechne

$$\begin{aligned}
\frac{d}{ds}c(s+t) &= c'(s+t) = ic(s+t), \quad c(s+t)|_{s=0} = c(t),\\
\frac{d}{ds}c(s)c(t) &= c'(s)c(t) = ic(s)c(t), \quad c(s)c(t)|_{s=0} = c(t).
\end{aligned}$$

Aus Satz 12.1 folgt $c(s+t) = c(s)c(t)$, insbesondere $c(t)c(-t) = c(0) = 1$. Andererseits gilt $c(t)\overline{c(t)} = |c(t)|^2 = 1$ nach (a), also $\overline{c(t)} = c(-t)$.

Es bleibt zu zeigen, dass $c(t)$ periodisch ist. Wir behaupten, dass $c(t) = u(t) + iv(t)$ den Punkt $i = (0,1)$ in endlicher Zeit erreicht. Betrachte dazu die Menge

$$M = \{t_1 > 0 : u(t) > 0 \text{ für } t \in [0,t_1]\}.$$

Da $u(0) = 1$ sowie $v(0) = 0$, $v'(0) = 1$, gibt es ein $t_0 \in M$ mit $v(t_0) > 0$. Sei $t_1 \in M$ mit $t_1 > t_0$. Dann ist $u(t) > 0$ auf $[0,t_1]$, und v ist wachsend auf $[0,t_1]$ wegen $v' = u$.

Insbesondere ist $v(t) \geq v(t_0) > 0$ für $t \geq t_0$, und wegen $u' = -v$ folgt mit Folgerung 10.2, dem Schrankensatz,

$$0 < u(t_1) \leq u(t_0) - v(t_0)(t_1 - t_0), \qquad \text{also} \qquad t_1 \leq t_0 + \frac{u(t_0)}{v(t_0)} < \infty.$$

Dies beweist $\tau = \sup M < \infty$. Aber wegen Stetigkeit ist dann $u(\tau) = 0$, und weiter $v(\tau) = \pm 1$ wegen $u(\tau)^2 + v(\tau)^2 = 1$. Da $v(t)$ wachsend, ist $v(\tau) = 1$ und damit $c(\tau) = i$ wie behauptet. Mit dem Eindeutigkeitssatz, Satz 12.1, folgt nun für alle $t \in \mathbb{R}$

$$c(t+\tau) = ic(t) \quad \text{folglich} \quad c(t+4\tau) = i^4 c(t) = c(t).$$

Also ist 4τ eine Periode von $c(t)$. $\square$

Wir zeigen jetzt, dass die Kreisbewegung längentreu ist. Zur Definition der Länge wählen wir die äquidistante Unterteilung $t_k = a + k\tau_n$, $k = 0, \ldots, n$, mit $\tau_n = (b-a)/n$, und setzen

$$L(c|_{[a,b]}) = \lim_{n\to\infty} L_n \quad \text{mit } L_n = \sum_{k=1}^{n} \big|c(t_k) - c(t_{k-1})\big| = \sum_{k=1}^{n} \big|e^{it_k} - e^{it_{k-1}}\big|.$$

Die Summe L_n ist die Länge des Polygonzugs, der durch die Punkte $c(t_k) = e^{it_k}$ gegeben ist. Der Grenzwert existiert, und zwar berechnen wir

$$L_n = \sum_{k=1}^{n} \big|e^{i(a+k\tau_n)} - e^{i(a+(k-1)\tau_n)}\big| = \sum_{k=1}^{n} \big|e^{i(a+(k-1)\tau_n)}(e^{i\tau_n} - 1)\big| = n\,|e^{i\tau_n} - 1| = \Big|\frac{e^{i\tau_n} - 1}{\tau_n}\Big|\,|b-a|.$$

Nun gilt $(e^{it} - 1)/t = (c(t) - c(0))/t \to c'(0) = i$ mit $t \to 0$, also folgt

$$L(c|_{[a,b]}) = \lim_{n\to\infty} L_n = |b-a|. \tag{12.5}$$

Das ist die Längentreue der Kreisbewegung. Wir können damit die Periode der Kreisbewegung angeben. Allgemein heißt $p \in \mathbb{R}\backslash\{0\}$ Periode einer Funktion $f : \mathbb{R} \to \mathbb{C}$, wenn gilt:

$$f(t+p) = f(t) \quad \text{für alle } t \in \mathbb{R}.$$

Mit p ist auch $-p$ Periode von f, und allgemeiner alle Zahlen $\mathbb{Z}^*p$, wobei $\mathbb{Z}^* = \mathbb{Z}\backslash\{0\}$. Eine Periode $p > 0$ heißt kleinste Periode, wenn es keine Periode $p' \in (0,p)$ gibt. In diesem Fall ist $\mathbb{Z}^*p$ die Menge aller Perioden: andernfalls gibt es eine Periode $p' \in (k \cdot p, (k+1) \cdot p)$ mit $k \in \mathbb{Z}$, und dann ist $p' - k \cdot p \in (0,p)$ eine kleinere Periode, Widerspruch.

Folgerung 12.1 (die Zahl π) *Die Kreisbewegung $c(t) = e^{it}$ hat die kleinste Periode 2π, die Länge des Einheitskreises.*

Beweis: Sei $\tau > 0$ wie im Beweis von Satz 12.2. Wegen $c(t+\tau) = ic(t)$ durchläuft $c(t)$ auf dem Intervall $[0, 4\tau]$ die Viertelkreise in den vier Quadranten der Reihe nach gegen den Uhrzeigersinn. Damit ist $c(4\tau) = c(0) = 1$, und c bildet das Intervall $[0, 4\tau)$ bijektiv auf den Kreis ab. Somit ist 4τ die kleinste Periode. Wir *definieren* π durch die Gleichung

$$2\pi = L(c|_{[0,4\tau]}). \tag{12.6}$$

Aus (12.5) folgt dann $4\tau = 2\pi$. $\square$

Satz 12.3 (Cosinus und Sinus) *Die Funktionen* cos, sin $\in C^\infty(\mathbb{R})$ *sind definiert durch*

$$e^{it} = \cos t + i \sin t \quad \textit{(Eulersche Formel)}, \qquad \textit{bzw. äquivalent} \tag{12.7}$$

$$\cos t = \frac{e^{it} + e^{-it}}{2}, \quad \sin t = \frac{e^{it} - e^{-it}}{2i}. \tag{12.8}$$

Die Funktionen haben folgende Eigenschaften:

(1) *Es gilt* $\cos' = -\sin$ *und* $\sin' = \cos$.

(2) $u(t) = \cos t$ *und* $v(t) = \sin t$ *sind die eindeutigen Lösungen der Anfangswertprobleme*

$$\begin{aligned} u'' + u = 0, \quad u(0) = 1, \quad u'(0) = 0,\\ v'' + v = 0, \quad v(0) = 0, \quad v'(0) = 1. \end{aligned}$$

(3) $\cos^2 t + \sin^2 t = 1$ *für alle* $t \in \mathbb{R}$.

(4) $\cos(-t) = \cos t$ *und* $\sin(-t) = -\sin(t)$ *für alle* $t \in \mathbb{R}$.

(5) *Für alle* $s, t \in \mathbb{R}$ *gelten die Additionstheoreme*

$$\begin{aligned} \cos(s+t) &= \cos s\, \cos t - \sin s\, \sin t,\\ \sin(s+t) &= \sin s\, \cos t + \cos s\, \sin t. \end{aligned}$$

(6) cos *und* sin *sind periodisch, jeweils mit kleinster Periode* 2π.

Beweis: Nach Satz 12.2 ist $c(t) = e^{it}$ in $C^\infty(\mathbb{R}, \mathbb{C})$, also sind cos und sin in $C^\infty(\mathbb{R})$. Behauptung (1) folgt mit (12.7) direkt aus (12.4). Behauptung (2) gilt dann nach Lemma 12.1. Die Gleichung (3) gilt nach Satz 12.2(a). Für (4) verwende Satz 12.2(c). Berechne für (5) mit der Funktionalgleichung, Satz 12.2(b),

$$\cos(s+t) + i \sin(s+t) = e^{i(s+t)} = e^{is} e^{it} = (\cos s + i \sin s)(\cos t + i \sin t).$$

Behauptung (5) folgt nun durch Vergleich von Real- und Imaginärteil. Nach Definition ist klar, dass jede Periode von e^{it} auch Periode von cos und sin ist. Es gilt aber auch die Umkehrung: sei zum Beispiel $\cos(t+p) = \cos t$ für alle $t \in \mathbb{R}$. Durch Ableiten folgt $\sin(t+p) = \sin t$, und damit $\exp i(t+p) = \exp it$. Da nach Definition 2π die kleinste Periode von $c(t) = e^{it}$ ist, gilt dasselbe auch für $\cos t$ und $\sin t$. $\square$

Mit der Formel $e^{i(t+\pi/2)} = ie^{it}$, die schon oben benutzt wurde, können wir folgende Wertetabelle der Funktionen Cosinus und Sinus aufstellen. Dabei sind die vier Quadranten entgegen dem Uhrzeigersinn mit I bis IV nummeriert.

t	0		$\pi/2$		π		$3\pi/2$		2π
e^{it}	1	I	i	II	-1	III	$-i$	IV	1
cos	1	$\searrow$	0	$\searrow$	-1	$\nearrow$	0	$\nearrow$	1
sin	0	$\nearrow$	1	$\searrow$	0	$\searrow$	-1	$\nearrow$	0

Die Nullstellen der trigonometrischen Funktionen sind wie folgt gegeben.

$$\begin{aligned} e^{it} = 1 \quad &\Leftrightarrow \quad t = 2k\pi \quad \text{mit } k \in \mathbb{Z},\\ \cos t = 0 \quad &\Leftrightarrow \quad t = \pi/2 + k\pi \quad \text{mit } k \in \mathbb{Z},\\ \sin t = 0 \quad &\Leftrightarrow \quad t = k\pi \quad \text{mit } k \in \mathbb{Z}. \end{aligned} \tag{12.9}$$

Die Funktionen cos und sin sind periodisch und damit nicht injektiv. Wir können sie aber auf ein Intervall einschränken, auf dem sie streng monoton sind und damit eine Umkehrfunktion haben. Man spricht genauer von einem Zweig des *arcus cosinus* bzw. *arcus sinus*,[5] weil Einschränkung auf andere Intervalle zu anderen Umkehrfunktionen führt.

Satz 12.4 (Arcusfunktionen) $\cos : [0,\pi] \to [-1,1]$ *und* $\sin : [-\frac{\pi}{2},\frac{\pi}{2}] \to [-1,1]$ *haben stetige Umkehrfunktionen, die wie folgt bezeichnet werden:*

(a) $\arccos : [-1,1] \to [0,\pi]$ *(arcus cosinus), mit Ableitung*

$$\arccos'(x) = -\frac{1}{\sqrt{1-x^2}} \qquad \textit{für } x \in (-1,1). \tag{12.10}$$

(b) $\arcsin : [-1,1] \to [-\frac{\pi}{2},\frac{\pi}{2}]$, *(arcus sinus), mit Ableitung*

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}} \qquad \textit{für } x \in (-1,1). \tag{12.11}$$

Beweis: Es gilt $\cos' = -\sin < 0$ auf $(0,\pi)$, und $\sin' = \cos > 0$ auf $[-\frac{\pi}{2},\frac{\pi}{2}]$. Also ist cos auf $[0,\pi]$ streng monoton fallend, und sin auf $[-\frac{\pi}{2},\frac{\pi}{2}]$ streng monoton wachsend, siehe Folgerung 10.1. Die Existenz und Stetigkeit der Umkehrfunktionen gilt nach Satz 8.2. Die Arcusfunktionen sind differenzierbar auf $(-1,1)$ nach Satz 9.4. Wir berechnen, wobei wir $\sin(\arccos x) > 0$ haben wegen $\arccos x \in (0,\pi)$:

$$\begin{aligned} 1 &= \frac{d}{dx}\cos(\arccos x) \\ &= -\sin(\arccos x)\arccos'(x) \\ &= -\sqrt{1-\cos^2(\arccos x)}\,\arccos'(x) \\ &= -\sqrt{1-x^2}\,\arccos'(x). \end{aligned}$$

Die Formel für $\arcsin'(x)$ folgt analog. □

Für $x + iy = e^{it}$ mit $0 \le t \le 2\pi$ sehen wir

$$\arccos x = \arccos(\cos t) = \begin{cases} t & \text{falls } t \in [0,\pi], \text{ also } y \ge 0, \\ 2\pi - t & \text{falls } t \in [\pi, 2\pi], \text{ also } y \le 0. \end{cases}$$

Somit ist $\arccos x$ die Länge des kürzeren Einheitskreisbogens zwischen den Punkten $(1,0)$ und (x,y) auf dem Einheitskreis. Daraus leitet sich folgende Definition ab.

Definition 12.1 (Winkel) *Der Winkel zwischen* $v, w \in \mathbb{R}^n$, $v, w \neq 0$, *ist*

$$\angle(v,w) = \arccos\left\langle \frac{v}{\|v\|}, \frac{w}{\|w\|} \right\rangle \in [0,\pi].$$

Das Skalarprodukt liegt in $[-1,1]$ wegen der Ungleichung von Cauchy-Schwarz, Satz 5.5. In den Grenzfällen sind v, w parallel bzw. antiparallel, der Winkel ist dann $\arccos 1 = 0$ bzw. $\arccos -1 = \pi$.

[5] Arcus = Bogen (Latein)

In der Schule werden die Funktionen Cosinus und Sinus am rechtwinkligen Dreieck betrachtet. Ist α ein Winkel des Dreiecks, gemessen in Bogenmaß, so setzt man

$$\cos\alpha = \frac{\text{Ankathete}}{\text{Hypothenuse}} \quad \text{und} \quad \sin\alpha = \frac{\text{Gegenkathete}}{\text{Hypothenuse}}.$$

Betrachten wir zuerst die Situation, dass das Dreeick durch seine Ecken bzw. Seitenlängen gegeben ist. Dann kann die rechte Seite ohne weiteres ausgerechnet werden. Um aber den Winkel α zu bestimmen, muss die Länge des zugehörigen Einheitskreisbogens ermittelt werden. Das ist ein Grenzwert, der durch die Funktion Arcus Cosinus berechnet wird:

$$\alpha = \arccos\left\langle \frac{b}{\|b\|}, \frac{c}{\|c\|}\right\rangle \qquad (b, c \text{ Seitenvektoren, die } \alpha \text{ einschließen}).$$

Für gewisse symmetrische Dreiecke ist die Bestimmung des Winkels elementar möglich, so im gleichschenkligen Fall ($\alpha = \frac{\pi}{4}$) oder beim halben gleichseitigen Dreieck ($\alpha = \frac{\pi}{3}$ bzw. $\alpha = \frac{\pi}{6}$). Zusammen mit $\alpha = 0$ und $\alpha = \frac{\pi}{2}$ ergibt sich so eine gewisse Wertetabelle. Allgemein kann der Winkel mit dem Geodreieck näherungsweise gemessen werden.

Um $\cos\alpha$ und $\sin\alpha$ in diesem Kontext zu *definieren*, müsste zu gegebenem $\alpha \in \mathbb{R}$ ein rechtwinkliges Dreieck mit Winkel α gefunden werden. Dazu wäre am Einheitskreis ein Bogen der Länge α abzutragen. Der gesuchte Punkt $e^{i\alpha}$ müsste ohne die Verwendung von Cosinus und Sinus berechnet werden, denn die sollen ja definiert werden. Das ist in der Schule nicht machbar, der Winkel kann aber mit dem Geodreieck näherungsweise konstruiert werden. In beiden Situationen kann man auch zum Taschenrechner greifen, jedenfalls wenn er die Tasten `cos` und `inv` hat. Das rechtwinklige Dreieck ist sicher die beste Methode, die trigonometrischen Funktionen einzuführen. Dabei kann es nicht schaden, die diskutierte Problematik im Hinterkopf zu haben.

Folgerung 12.2 (Polarkoordinaten) *Zu jedem $z \in \mathbb{C}\backslash\{0\}$ gibt es eindeutig bestimmte $r > 0$, $\vartheta \in [0, 2\pi)$ mit $z = re^{i\vartheta}$.*

Beweis: Sei zunächst $|z| = 1$. Wir definieren

$$\vartheta = \begin{cases} \arccos x \in [0, \pi] & \text{für } y \geq 0, \\ 2\pi - \arccos x \in (\pi, 2\pi) & \text{für } y < 0. \end{cases}$$

Verwende $\sin\vartheta = \sqrt{1 - \cos^2\vartheta}$ für $\vartheta \in [0, \pi]$, sowie $\sin(2\pi - \vartheta) = -\sin\vartheta$. Damit folgt

$$e^{i\vartheta} = \begin{cases} \cos(\arccos x) + i\sin(\arccos x) = x + iy & \text{falls } y \geq 0, \\ \cos(\arccos x) - i\sin(\arccos x) = x + iy & \text{falls } y < 0. \end{cases}$$

Für $z \in \mathbb{C}\backslash\{0\}$ beliebig setzen wir $r = |z| > 0$ und definieren $\vartheta \in [0, 2\pi]$ mit $z/r = e^{i\vartheta}$. Die Eindeutigkeit von $r = |z|$ ist klar, für ϑ folgt sie aus der Periodizität der Kreisbewegung, siehe Folgerung 12.1. □

Wir haben die Exponentialfunktion einerseits für reelle Argumente $\exp(x) = e^x$ und andererseits für rein imaginäre Argumente $\exp(it) = e^{it}$. Es ist nun fast zwingend, die Funktion auf ganz $\mathbb{C}$ zu erklären.

Satz 12.5 (komplexe Exponentialfunktion) *Wir definieren* $\exp : \mathbb{C} \to \mathbb{C}$ *durch*

$$\exp : \mathbb{C} \to \mathbb{C},\, e^z := e^x\, e^{iy} \qquad \textit{für } z = x + iy.$$

Die Funktion hat folgende Eigenschaften:

(1) $e^{z_1+z_2} = e^{z_1} e^{z_2}$ *für alle* $z_1, z_2 \in \mathbb{C}$ *(Funktionalgleichung).*

(2) *Es gilt* $\exp(\mathbb{C}) = \mathbb{C}\backslash\{0\}$, *und* $e^{z_1} = e^{z_2}$ *genau wenn* $z_1 - z_2 \in 2\pi i\mathbb{Z}$.

(3) *Zu gegebenen* $\lambda, z_0 \in \mathbb{C}$ *hat das Anfangswertproblem*

$$\gamma'(t) = \lambda\gamma(t) \textit{ für } t \in \mathbb{R}, \quad \gamma(0) = z_0,$$

genau eine Lösung $\gamma \in C^1(\mathbb{R}, \mathbb{C})$, *und zwar* $\gamma(t) = z_0 e^{\lambda t}$.

Beweis: Sei $z_{1,2} = x_{1,2} + iy_{1,2}$. Mit Satz 11.1(b) und Satz 12.2(b) folgt

$$e^{z_1+z_2} = e^{x_1+x_2+i(y_1+y_2)} = e^{x_1+x_2}\, e^{i(y_1+y_2)} = e^{x_1} e^{iy_1} e^{x_2} e^{iy_2} = e^{z_1} e^{z_2}.$$

Für $z = x + iy$ gilt $e^x > 0$ und $|e^{iy}| = 1$, siehe Satz 11.1(a) und Satz 12.2(a), also ist stets $e^z \neq 0$. Sei $w \in \mathbb{C}\backslash\{0\}$ gegeben. Nach Folgerung 12.2 gibt es ein $y \in [0, 2\pi)$ mit $e^{iy} = w/|w|$. Setzen wir $x = \log|w|$, so folgt

$$w = |w|\,\frac{w}{|w|} = e^x e^{iy} = e^z \qquad \text{mit } z = x + iy.$$

Weiter gilt, da 2π kleinste Periode von e^{iy} nach Folgerung 12.1 ist,

$$e^z = e^x e^{iy} = 1 \quad \Leftrightarrow \quad e^x = e^{iy} = 1 \quad \Leftrightarrow \quad x = 0,\, y \in 2\pi\mathbb{Z}.$$

Daraus folgt mit der Funktionalgleichung

$$e^{z_1} = e^{z_2} \quad \Leftrightarrow \quad e^{z_1 - z_2} = 1 \quad \Leftrightarrow \quad z_1 - z_2 \in 2\pi i\mathbb{Z}.$$

Die Exponentialabbildung hat also die Periode $2\pi i$, der Streifen $0 \leq \operatorname{Im} z < 2\pi$ wird bijektiv auf $\mathbb{C}\backslash\{0\}$ abgebildet. Damit ist (2) gezeigt.

Für (3) zeigen wir erst, dass $z_0 e^{\lambda t}$ eine Lösung ist. Es gilt $z_0 e^{\lambda t}|_{t=0} = z_0$, und mit $\lambda = \alpha + i\beta$ berechnen wir (verwende $e^{i\beta t} = c(\beta t)$ mit $c(t) = e^{it}$)

$$\frac{d}{dt} z_0 e^{\lambda t} = \frac{d}{dt}\big(z_0 e^{\alpha t} e^{i\beta t}\big) = z_0\,(\alpha + i\beta) e^{\alpha t} e^{i\beta t} = \lambda z_0 e^{\lambda t}.$$

Nun zur Eindeutigkeit: ist $\gamma(t)$ irgendeine Lösung, so gilt

$$\frac{d}{dt}\big(e^{-\lambda t}\gamma(t)\big) = -\lambda e^{-\lambda t}\gamma(t) + e^{-\lambda t}\lambda\gamma(t) = 0.$$

Es folgt $e^{-\lambda t}\gamma(t) = e^{-\lambda t}\gamma(t)|_{t=0} = z_0$, somit laut Funktionalgleichung $\gamma(t) = z_0 e^{\lambda t}$. □

In den letzten beiden Kapiteln wurden die Eigenschaften der Exponentialfunktion und der trigonometrischen Funktionen hergeleitet. Eine Formel für die Funktionen stand nicht zur Verfügung, sondern wir haben direkt die Differentialgleichungen benutzt. Für die meisten

Differentialgleichungen gibt es keine „explizite“ Lösungsformel – selbst wenn sind die Eigenschaften der Funktion oft schwierig daraus abzulesen. Darüber hinaus hat unser Weg den Vorteil, dass wir Informationen über alle möglichen Lösungen der Gleichung gewinnen, etwa Lösungen mit anderen Anfangswerten. Es bleibt aber das zentrale Problem, die Existenz der Lösungen von (11.1) und (12.4) zu zeigen. Wie wir das machen ist egal, und es gibt mehrere Zugänge. Im nächsten Kapitel verwenden wir einen Ansatz als Potenzreihe, der auf Cauchy zurückgeht. Von den meisten Autoren wird diese Reihe zur Definition der Exponentialfunktion benutzt. Auf der Basis der Vorlesung gibt es folgende Alternativen:

- Der Logarithmus kann als *Stammfunktion von* $1/r$ definiert werden, die Exponentialfunktion ergibt sich dann als Umkehrfunktion. Analog: der Arcus Cosinus ist Stammfunktion von $-(1-x^2)^{-1/2}$, der Kosinus ist dann die Umkehrfunktion.

- In Satz 11.2 wurde das *Eulerverfahren* betrachtet, eine Diskretisierung von $f' = f$. Die approximativen Lösungen $f_n(x)$ konvergieren gegen die Lösung des kontinuierlichen Anfangswertproblems. Solche Diskretisierungen werden auch zur numerischen Lösung benutzt.

- Für Differentialgleichungen der allgemeinen Form $f'(t) = \phi(f(t), t)$ haben *Picard* und *Lindelöf* um 1890 die Existenz von Lösungen des Anfangswertproblems gezeigt. Wir werden ihren Beweis später behandeln, er beruht auf einem Iterationsverfahren.

13 Konvergenz von Funktionenfolgen

In (11.1) bzw. (12.4) haben wir die Exponentialfunktion und die Kreisbewegung als Lösungen von Anfangswertproblemen definiert. In diesem Abschnitt zeigen wir die Existenz dieser Lösungen. Die Grundidee ist, die Funktionen als Grenzwerte von Reihen von Polynomen zu gewinnen. Dies führt auf den Begriff der Potenzreihe.

Definition 13.1 (Potenzreihen) *Eine komplexe Potenzreihe ist eine Reihe der Form*

$$P(z) = \sum_{k=0}^{\infty} a_k z^k \qquad \textit{mit } a_k \in \mathbb{C}.$$

Allgemeiner hat man Potenzreihen mit Entwicklungspunkt $z_0 \in \mathbb{C}$, das heißt Reihen der Form $P(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k$. Der Einfachheit halber beschränken wir uns auf $z_0 = 0$. Definition 13.1 lässt offen, ob beziehungsweise für welche $z \in \mathbb{C}$ die Reihe tatsächlich konvergiert. Klar ist $P(0) = a_0$, es kann aber sein, dass die Reihe für alle $z \neq 0$ divergiert, dann ist sie natürlich nicht relevant. Ein Beispiel ist $\sum_{k=0}^{\infty} k^k z^k$, wie der Nullfolgentest zeigt. Bei interessanten Reihen (was immer das ist) erwarten wir aber Konvergenz für Punkte z in einem gewissen Gebiet. Auf diesem Gebiet definiert die Reihe dann eine Funktion.

Konkreter betrachte $f' = f$ auf $\mathbb{R}$, $f(0) = 1$, das Anfangswertproblem für die reelle Exponentialfunktion. Mit Cauchy machen wir den Reihenansatz $f(x) = \sum_{k=0}^{\infty} a_k x^k$, und berechnen formal

$$f'(x) = \frac{d}{dx} \sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} \frac{d}{dx} a_k x^k = \sum_{k=1}^{\infty} k a_k x^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} x^k.$$

Koeffizientenvergleich in der Gleichung $f' = f$ ergibt $a_{k+1} = a_k/(k+1)$ für $k \geq 0$. Die Anfangsbedingung liefert $a_0 = f(0) = 1$, also folgt durch Induktion $a_k = 1/k!$. Wir kommen so zum *educated guess*, vgl. auch (4.3),

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \tag{13.1}$$

Das Problem ist, diese Rechnung rigoros zu machen. Wir beginnen dazu mit der allgemeinen Frage, für welche $z \in \mathbb{C}$ eine Potenzreihe konvergiert. Zuerst ein technisches Lemma.

Lemma 13.1 (von Abel) *Sei $P(z) = \sum_{k=0}^{\infty} a_k z^k$, $a_k \in \mathbb{C}$, eine Potenzreihe. Es gebe $z_0 \neq 0$ und $M \in [0, \infty)$ mit $|a_k|\,|z_0|^k \leq M$ für alle $k \in \mathbb{N}_0$. Dann ist $P(z)$ für $|z| < |z_0|$ absolut konvergent und für alle $k \in \mathbb{N}_0$ gilt*

$$|P(z) - P_k(z)| \leq \sum_{\ell=k+1}^{\infty} |a_\ell|\,|z|^\ell \leq \frac{M}{1 - \frac{|z|}{|z_0|}} \left(\frac{|z|}{|z_0|}\right)^{k+1}. \tag{13.2}$$

Hier ist $P_k(z) = \sum_{\ell=0}^{k} a_\ell z^\ell$ die k-te Partialsumme.

BEWEIS: Für $|z| < |z_0|$ gilt $|a_k z^k| = |a_k|\,|z_0|^k(|z|/|z_0|)^k \leq M(|z|/|z_0|)^k$. Wegen $|z|/|z_0| < 1$ folgt die absolute Konvergenz aus dem Majorantenkriterium durch Vergleich mit der geometrischen Reihe. Genauer gilt

$$|P(z) - P_k(z)| \leq \sum_{\ell=k+1}^{\infty} |a_\ell||z|^\ell \leq M \sum_{\ell=k+1}^{\infty} \left(\frac{|z|}{|z_0|}\right)^\ell = \frac{M}{1 - \frac{|z|}{|z_0|}} \left(\frac{|z|}{|z_0|}\right)^{k+1}.$$

□

Das Konvergenzgebiet einer Potenzreihe sieht allgemein wie folgt aus.

Satz 13.1 (vom Konvergenzradius) *Zu jeder Potenzreihe $P(z) = \sum_{k=0}^{\infty} a_k z^k$ gibt es genau ein $R \in [0, \infty]$, den Konvergenzradius, mit folgender Eigenschaft:*

$$P(z) \text{ ist } \begin{cases} \text{absolut konvergent} & \text{für } |z| < R, \\ \text{divergent} & \text{für } |z| > R. \end{cases}$$

BEWEIS: Die Eindeutigkeit von R ist klar. Zur Existenz definieren wir

$$R = \sup\{|z| : P(z) \text{ konvergiert}\} \in [0, \infty].$$

Ist $|z| < R$, so gibt es nach Definition ein $z_0 \in \mathbb{C}$ mit $|z| < |z_0| \leq R$, so dass $P(z_0)$ konvergiert. Es gilt dann $a_k z_0^k \to 0$ nach dem Nullfolgenkriterium, Satz 6.1. Damit gibt es ein $M \in [0, \infty)$ mit $|a_k||z_0|^k \leq M$, und $P(z)$ konvergiert absolut wegen Lemma 13.1. Andererseits ist die Reihe divergent für $|z| > R$ nach Definition von R. □

Beispiel 13.1 Für die Exponentialreihe $\sum_{k=0}^{\infty} z^k/k!$ gilt, außer im trivialen Fall $z = 0$,

$$\frac{|z^{k+1}/(k+1)!|}{|z^k/k!|} = \frac{|z|}{k+1} \to 0 \quad \text{mit } k \to \infty.$$

Nach dem Quotientenkriterium, Satz 6.6, konvergiert die Reihe absolut für alle $z \in \mathbb{C}$. Der Konvergenzradius der Reihe ist $R = \infty$.

Beispiel 13.2 Die Binomialreihe zum Parameter $\alpha \in \mathbb{C}$ lautet

$$B_\alpha(z) = \sum_{k=0}^{\infty} \binom{\alpha}{k} z^k = 1 + \alpha z + \frac{\alpha(\alpha - 1)}{2} z^2 + \ldots$$

Für $\alpha \in \mathbb{N}_0$ bricht die Reihe nach $k = \alpha$ ab, und die Binomische Formel aus Satz 2.7 liefert $B_\alpha(z) = (1 + z)^\alpha$. Im folgenden sei nun $\alpha \notin \mathbb{N}_0$. Für $z \neq 0$ ist dann $a_k = \binom{\alpha}{k} z^k \neq 0$ für alle $k \in \mathbb{N}_0$, und es gilt

$$\left|\frac{a_{k+1}}{a_k}\right| = \frac{|\alpha - k|}{k+1}|z| \to |z| \quad \text{mit } k \to \infty.$$

Nach dem Quotientenkriterium konvergiert die Reihe für $|z| < 1$ und divergiert für $|z| > 1$, das heißt der Konvergenzradius ist $R = 1$.

Setzen wir in eine Potenzreihe $P(z)$ ein $z \in \mathbb{C}$ ein, so haben wir eine Reihe von komplexen Zahlen. Aber eine Potenzreihe ist eine Folge von Funktionen, den Polynomen $P_k(z)$. Lassen sich Eigenschaften der $P_k(z)$, etwa die Stetigkeit, auf $P(z)$ übertragen? Das ist ein generelles und grundlegendes Problem: sei $f_k : D \to \mathbb{R}^m$ eine Folge von Funktionen auf $D \subset \mathbb{R}^n$, die punktweise gegen eine Grenzfunktion $f : D \to \mathbb{R}^m$ konvergiert, das heißt

$$f(x) = \lim_{k\to\infty} f_k(x) \quad \text{für alle } x \in D.$$

Folgt die Stetigkeit von f aus der Stetigkeit der f_k, unter geeigneten Voraussetzungen? Das folgende Beispiel zeigt, dass allein die punktweise Konvergenz im allgemeinen nicht die Stetigkeit der Grenzfunktion garantiert.

Beispiel 13.3 Betrachte die stetigen Funktionen $f_k : [0,1] \to \mathbb{R}$,

$$f_k(x) = \begin{cases} 1 - kx & \text{für } 0 \leq x < \frac{1}{k}, \\ 0 & \text{sonst.} \end{cases}$$

Die Folge konvergiert punktweise gegen $f : [0,1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 1 & \text{für } x = 0, \\ 0 & \text{sonst.} \end{cases}$$

Die Grenzfunktion ist im Punkt $x = 0$ nicht stetig.

Um die Stetigkeit der Grenzfunktion zu garantieren, brauchen wir also einen stärkeren Begriff von Konvergenz. Die Supremumsnorm ist naheliegend, weil sie alle Werte $f(x)$ gleichzeitig kontrolliert.

Definition 13.2 (Supremumsnorm) *Für $f : D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$, setzen wir*

$$\|f\|_D = \sup\{\|f(x)\| : x \in D\}.$$

Wir bezeichnen mit $\mathcal{B}(D, \mathbb{R}^m)$ den Vektorraum der beschränkten Funktionen $f : D \to \mathbb{R}^m$, also mit $\|f\|_D < \infty$. Die Supremumsnorm $\|\cdot\|_D : \mathcal{B}(D, \mathbb{R}^m) \to \mathbb{R}$ hat folgende Eigenschaften, analog zur Euklidischen Norm.

Positivität: $\|f\|_D \geq 0$ mit Gleichheit genau wenn $f = 0$,

Halblinearität: $\|\lambda f\|_D = |\lambda|\,\|f\|_D$ für $\lambda \in \mathbb{R}$.

Dreiecksungleichung: $\|f + g\|_D \leq \|f\|_D + \|g\|_D$.

Definition 13.3 (Gleichmäßige Konvergenz) *Eine Folge von Funktionen $f_k : D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$, konvergiert gleichmäßig gegen $f : D \to \mathbb{R}^m$, falls gilt:*

$$\|f - f_k\|_D \to 0 \quad \textit{mit } k \to \infty.$$

In Quantorensprache sieht punktweise bzw. gleichmäßige Konvergenz wie folgt aus:

$$\begin{array}{lllll} \forall \varepsilon > 0 & \forall x \in D & \exists K & \forall k > K: & \|f_k(x) - f(x)\| < \varepsilon \quad \text{(punktweise)}, \\ \forall \varepsilon > 0 & \exists K & \forall x \in D & \forall k > K: & \|f_k(x) - f(x)\| < \varepsilon \quad \text{(gleichmäßig)}. \end{array}$$

Der Unterschied: punktweise Konvergenz erlaubt, dass die Schranke K von $x \in D$ abhängt, also $K = K(x, \varepsilon)$, während gleichmäßige Konvergenz eine gemeinsame Schranke K für alle x verlangt. Im Beispiel 13.3 gilt etwa, falls $x > 0$ und damit $f(x) = 0$,

$$k < \frac{1}{2x} \quad \Rightarrow \quad x < \frac{1}{2k} \quad \Rightarrow \quad |f_k(x) - f(x)| = 1 - kx > \frac{1}{2}.$$

Also muss $K(x, \frac{1}{2}) \geq \frac{1}{2x}$ sein, eine gemeinsame Schranke für alle $x \in [0,1]$ gibt es nicht.

Klar ist, dass aus gleichmäßig punktweise folgt. Deshalb kann man in zwei Schritten vorgehen, um eine Funktionenfolge f_k auf gleichmäßige Konvergenz zu prüfen:

(1) Konvergiert die Folge punktweise? Wenn nicht, so erst recht nicht gleichmäßig. Wenn ja, so ist die punktweise Grenzfunktion die einzig mögliche Kandidatin für den gleichmäßigen Grenzwert.

(2) Nun bestimme $\|f_k - f\|_D$ bzw. schätze diese Norm ab. Gilt $\|f_k - f\|_D \to 0$ mit $k \to \infty$, so ist f_k gleichmäßig konvergent gegen f. Wenn nicht, so ist f_k zwar punktweise, aber nicht gleichmäßig konvergent.

Satz 13.2 (Gleichmäßige Konvergenz und Stetigkeit) *Seien $f_k : D \to \mathbb{R}^m$, $k \in \mathbb{N}$, stetige Funktionen auf $D \subset \mathbb{R}^n$, die gleichmäßig gegen $f : D \to \mathbb{R}^m$ konvergieren, also*

$$\|f_k - f\|_D \to 0 \quad \text{mit } k \to \infty.$$

Dann ist f ebenfalls stetig auf D.

Beweis: Sei $x_0 \in D$ gegeben. Für $x \in D$ beliebig und $k \in \mathbb{N}$ gilt

$$\begin{aligned} \|f(x) - f(x_0)\| &\leq \|f(x) - f_k(x)\| + \|f_k(x) - f_k(x_0)\| + \|f_k(x_0) - f(x_0)\| \\ &\leq \|f_k(x) - f_k(x_0)\| + 2\,\|f - f_k\|_D. \end{aligned}$$

Mit $x \to x_0$ ergibt sich, da f_k stetig ist,

$$\limsup_{x \to x_0} \|f(x) - f(x_0)\| \leq 2\,\|f - f_k\|_D \to 0 \quad \text{mit } k \to \infty.$$

Also folgt $\lim_{x \to x_0} \|f(x) - f(x_0)\| = 0$, die Stetigkeit in x_0. □

Satz 13.3 (Stetigkeit von Potenzreihen) *Sei $P(z) = \sum_{k=}^{\infty} a_k z^k$ eine komplexe Potenzreihe mit mit Konvergenzradius $R > 0$. Dann ist $P(z)$ stetig auf der Kreisscheibe $B_R(0) = \{z \in \mathbb{C} : |z| < R\}$.*

Beweis: Wir zeigen, dass die $P_k(z)$ auf jeder Kreisscheibe $B_\varrho(0)$ mit $\varrho < R$ gleichmäßig konvergieren. Die Behauptung folgt dann aus Satz 13.2. Sei $r \in (\varrho, R)$ beliebig gewählt. Da $P(r)$ konvergiert, gibt es ein $M \in [0, \infty)$ mit $|a_k| r^k \leq M$ für alle k. Nach Lemma 13.1 folgt

$$\|P - P_k\|_{B_\varrho(0)} \leq \frac{M}{1 - \frac{\varrho}{r}} \left(\frac{\varrho}{r}\right)^{k+1} \to 0 \quad \text{mit } k \to \infty.$$

Damit ist der Satz bewiesen. □

Nun zur Frage der Differenzierbarkeit der Grenzfunktion. Wir beschränken uns im folgenden auf Funktionen $f : I \to \mathbb{R}^m$ auf einem reellen Intervall, da wir nur für diese die Ableitung definiert haben.

Satz 13.4 (Vertauschung von Konvergenz und Ableitung) *Seien $f_k \in C^1(I, \mathbb{R}^m)$, $I = (a, b)$, punktweise konvergent gegen $f : I \to \mathbb{R}^m$. Konvergiert die Folge f_k' gleichmäßig gegen ein $g : I \to \mathbb{R}^m$, also*

$$\|f_k' - g\|_I \to 0 \quad \text{mit } k \to \infty,$$

so ist $f \in C^1(I, \mathbb{R}^m)$ und $f' = g$.

Beweis: Wir können $m = 1$ annehmen, andernfalls betrachte die Koordinatenfunktionen. Nach Satz 13.2 ist $g \in C^0(I)$; zu zeigen ist nur noch $f' = g$. Sei $x_0 \in I$ gegeben, mit der Dreiecksungleichung ist

$$\begin{aligned}
\Big|\frac{f(x) - f(x_0)}{x - x_0} - g(x_0)\Big| \quad &\leq \quad \Big|\frac{f(x) - f(x_0)}{x - x_0} - \frac{f_k(x) - f_k(x_0)}{x - x_0}\Big| \\
&+ \quad \Big|\frac{f_k(x) - f_k(x_0)}{x - x_0} - f_k'(x_0)\Big| + |f_k'(x_0) - g(x_0)|.
\end{aligned}$$

Der Schrankensatz, Folgerung 10.2 bzw. Folgerung 10.3, liefert

$$\Big|\frac{f_\ell(x) - f_\ell(x_0)}{x - x_0} - \frac{f_k(x) - f_k(x_0)}{x - x_0}\Big| = \Big|\frac{(f_\ell - f_k)(x) - (f_\ell - f_k)(x_0)}{x - x_0}\Big| \leq \|(f_\ell - f_k)'\|_I.$$

Mit $\ell \to \infty$ folgt wegen $\|(f_\ell - f_k)'\|_I \leq \|f_\ell' - g\|_I + \|f_k' - g\|_I$

$$\Big|\frac{f(x) - f(x_0)}{x - x_0} - \frac{f_k(x) - f_k(x_0)}{x - x_0}\Big| \leq \|f_k' - g\|_I.$$

Insgesamt ergibt sich die Abschätzung

$$\Big|\frac{f(x) - f(x_0)}{x - x_0} - g(x)\Big| \leq \Big|\frac{f_k(x) - f_k(x_0)}{x - x_0} - f_k'(x_0)\Big| + 2\|f_k' - g\|_I.$$

Mit $x \to x_0$ folgt daraus

$$\limsup_{x \to x_0} \Big|\frac{f(x) - f(x_0)}{x - x_0} - g(x)\Big| \leq 2\|f_k' - g\|_I \to 0 \quad \text{mit } k \to \infty.$$

Dies zeigt $f'(x_0) = g(x_0)$, die Behauptung. □

Um den Satz auf Potenzreihen anzuwenden, brauchen wir folgende Hilfsaussage.

Lemma 13.2 *Sei $P(z) = \sum_{k=0}^\infty a_k z^k$ eine Potenzreihe mit Konvergenzradius $R \in [0, \infty]$. Dann hat die formal differenzierte Reihe*

$$Q(z) = \sum_{k=1}^\infty k a_k z^{k-1} = \sum_{k=0}^\infty (k+1) a_{k+1} z^k$$

denselben Konvergenzradius R.

Beweis: $Q(z)$ hat höchstens den Konvergenzradius von $P(z)$, denn $|a_k z^k| \leq |k a_k z^{k-1}| \cdot |z|$ für $k \geq 1$. Zu $r \in (0,R)$ gibt es nun ein $M \in [0,\infty)$ mit $|a_k| r^k \leq M$, da $P(r)$ konvergiert. Es folgt für alle $z \in B_r(0)$

$$|k a_k z^{k-1}| = k|a_k| r^k \, \frac{|z|^{k-1}}{r^k} \leq \frac{kM|z|^{k-1}}{r^k}.$$

Die rechte Reihe konvergiert aber nach dem Quotientenkriterium, denn

$$\frac{(k+1)|z|^k}{r^{k+1}} \left(\frac{k|z|^{k-1}}{r^k} \right)^{-1} = \frac{(k+1)|z|}{kr} \to \frac{|z|}{r} < 1 \quad \text{mit } k \to \infty.$$

Also ist $Q(z)$ für $z \in B_R(0)$ absolut konvergent. □

Satz 13.5 (Differenzierbarkeit von Potenzreihen) *Die Potenzreihe $P(z) = \sum_{k=0}^\infty a_k z^k$ mit $a_k \in \mathbb{C}$ habe den Konvergenzradius $R > 0$. Dann ist die Funktion*

$$P : (-R,R) \to \mathbb{C},\ P(x) = \sum_{k=0}^\infty a_k x^k$$

stetig differenzierbar, und ihre Ableitung ergibt sich durch gliedweise Differentiation:

$$P'(x) = \sum_{k=1}^\infty k a_k x^{k-1} = \sum_{k=0}^\infty (k+1) a_{k+1} x^k \qquad \text{für alle } x \in (-R,R). \tag{13.3}$$

Hinweis. Durch Induktion folgt $P \in C^\infty((-R,R),\mathbb{C})$.

Beweis: Sei $\varrho \in (0,R)$. Nach Beweis von Satz 13.3 konvergieren die $P_k(x)$ auf $(-\varrho,\varrho)$ gleichmäßig gegen $P(x)$. Nach Lemma 13.2 konvergieren die $P_k'(x)$ auf $(-\varrho,\varrho)$ ebenfalls gleichmäßig gegen $Q(x)$, die gliedweise differenzierte Reihe. Satz 13.4 impliziert $P' = Q$ auf $(-\varrho,\varrho)$, also auf ganz $(-R,R)$. □

Satz 13.6 (Reihendarstellung von exp) *Es existieren Lösungen $\exp \in C^\infty(\mathbb{R})$ und $c \in C^\infty(\mathbb{R},\mathbb{C})$, $c(t) = e^{it}$, der Anfangswertprobleme (11.1) und (12.4), und zwar sind diese durch folgende Reihendarstellungen gegeben:*

$$\exp(x) = \sum_{k=0}^\infty \frac{x^k}{k!} \qquad \textit{und} \qquad c(t) = \sum_{k=0}^\infty \frac{(it)^k}{k!}. \tag{13.4}$$

Beweis: Nach Beispiel 13.1 hat die Exponentialreihe $P(x) = \sum_{k=0}^\infty x^k/k!$ den Konvergenzradius $R = \infty$. Aus Satz 13.5 folgt $P \in C^\infty(\mathbb{R})$, sowie nun rigoros

$$P'(x) = \sum_{k=1}^\infty \frac{x^{k-1}}{(k-1)!} = \sum_{k=0}^\infty \frac{x^k}{k!} = P(x), \quad P(0) = 1.$$

Das Argument für die Kreisbewegung ist analog: die Potenzreihe $P(t) = \sum_{k=0}^\infty i^k t^k/k!$ hat Konvergenzradius $R = \infty$, also ist $P \in C^\infty(\mathbb{R})$ und

$$P'(t) = \sum_{k=1}^\infty \frac{i^k t^{k-1}}{(k-1)!} = i \sum_{k=0}^\infty \frac{i^k t^k}{k!} = iP(t), \qquad \text{und } P(0) = 1.$$

□

Folgerung 13.1 (Reihendarstellung von $\cos$ **und** $\sin$**)** *Kosinus und Sinus haben für alle* $t \in \mathbb{R}$ *die Darstellungen*

$$\cos t = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!} \quad \textit{und} \quad \sin t = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!}. \tag{13.5}$$

BEWEIS: Aus Satz 13.6 und der Eulerschen Formel folgt mit $\ell = 2k$ bzw. $\ell = 2k+1$

$$\cos t + i \sin t = e^{it} = \sum_{\ell=0}^{\infty} \frac{i^\ell t^\ell}{\ell!} = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!}.$$

□

Beispiel 13.4 In Satz 12.5 hatten wir die komplexe Exponentialfunktion definiert durch $\exp(x+iy) = \exp(x)\exp(iy)$. Sie wird auch durch die Exponentialreihe $P(z)$ dargestellt, also

$$\exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!} \quad \text{für alle } z \in \mathbb{C}. \tag{13.6}$$

Es folgt nämlich durch gliedweise Differentiation wie oben, dass $P(tz)$ das Anfangswertproblem $\gamma'(t) = z\gamma(t)$, $\gamma(0) = 1$, löst. Nach Satz 12.5(c) folgt $P(tz) = \exp(tz)$, mit $t = 1$ also $P(z) = \exp(z)$.

Beispiel 13.5 Für die Potenzreihendarstellung des Logarithmus berechnen wir mit der geometrischen Reihe, für $x \in (-1,1)$,

$$\frac{d}{dx} \log(1+x) = \frac{1}{1+x} = \sum_{k=0}^{\infty} (-1)^k x^k =: Q(x).$$

$Q(x)$ ergibt sich durch gliedweise Differentiation der Reihe $P(x) = \sum_{k=1}^{\infty} (-1)^{k-1} x^k/k$. Nach Lemma 13.2 haben $P(x)$ und $Q(x)$ denselben Konvergenzradius, also $R = 1$, und Satz 13.5 ergibt $P'(x) = Q(x)$ für $x \in (-1,1)$. Nun ist $\log(1+x) = P(x) = 0$ für $x = 0$, also erhalten wir

$$\log(1+x) = P(x) = x - \frac{x^2}{2} + \frac{x^3}{3} - + \ldots \quad \text{für } x \in (-1,1).$$

Beispiel 13.6 Die Ableitung der Funktion $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ ist nach Quotientenregel

$$\tan' = \Big(\frac{\sin}{\cos}\Big)' = \frac{\cos^2 + \sin^2}{\cos^2} = 1 + \tan^2.$$

Für die Umkehrfunktion arctan folgt, für $x \in (-1,1)$,

$$\frac{d}{dx} \arctan x = \frac{1}{\tan'(\arctan x)} = \frac{1}{1+x^2} = \sum_{k=0}^{\infty} (-1)^k x^{2k} =: Q(x).$$

$Q(x)$ ist die gliedweise Ableitung von $P(x) = \sum_{k=0}^{\infty} (-1)^k x^{2k+1}/(2k+1)$. Der gemeinsame Konvergenzradius ist $R = 1$, und aus Satz 13.5 folgt $P'(x) = Q(x)$ für alle $x \in (-1,1)$. Da $P(0) = \arctan 0 = 0$, sehen wir

$$\arctan x = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - + \ldots \quad \text{für } x \in (-1,1).$$

Funktionen, die lokal durch eine Potenzreihe darstellbar sind, heißen analytisch. Es sind (bei weitem) nicht alle Funktionen analytisch, zum Beispiel hat die in Folgerung 11.2 konstruierte Funktion keine Darstellung als Potenzreihe in einer Umgebung des Nullpunkts. Dies zeigt der

Satz 13.7 (Identitätssatz für Potenzreihen) *Sei $P(z) = \sum_{k=0}^{\infty} a_k z^k$ eine Potenzreihe, die für ein $z_0 \neq 0$ konvergiert. Ist $0 \in \mathbb{C}$ Häufungspunkt der Menge $\{z \in \mathbb{C} : P(z) = 0\}$, so folgt $a_k = 0$ für alle $k \in \mathbb{N}_0$.*

Beweis: Wir nehmen induktiv an, dass schon $a_0 = a_1 = \ldots = a_{k-1} = 0$ gezeigt ist, wobei der Fall $k = 0$ den Induktionsanfang liefert. Für $|z| \leq |z_0|/2$ folgt aus (13.2) die Abschätzung

$$|P(z) - a_k z^k| = |P(z) - P_k(z)| \leq C|z|^{k+1} \qquad \text{wobei } C = 2M|z_0|^{-(k+1)}.$$

Nach Voraussetzung gilt $P(z_i) = 0$ für eine Folge $z_i \neq 0$ mit $z_i \to 0$, und Einsetzen von $z = z_i$ ergibt $|a_k|\,|z_i|^k \leq C|z_i|^{k+1}$, also $|a_k| \leq C|z_i| \to 0$ mit $i \to \infty$, das heißt $a_k = 0$. $\square$

Folgerung 13.2 (Koeffizientenvergleich) *Seien $P(z) = \sum_{k=0}^{\infty} a_k z^k$ und $Q(z) = \sum_{k=0}^{\infty} b_k z^k$ Potenzreihen mit positivem Konvergenzradius. Ist der Nullpunkt Häufungspunkt der Menge $\{z \in \mathbb{C} : P(z) = Q(z)\}$, so folgt $a_k = b_k$ für alle $k \in \mathbb{N}_0$.*

Beweis: Die Potenzreihe $F(z) = \sum_{k=0}^{\infty} c_k z^k$ mit $c_k = a_k - b_k$ hat positiven Konvergenzradius, und der Nullpunkt ist Häufungspunkt der Menge $\{z \in \mathbb{C} : F(z) = 0\}$. Die Behauptung folgt damit aus Satz 13.7. $\square$

Bisher haben wir uns immer in der offenen Kreissscheibe bewegt, auf der die Potenzreihe lokal gleichmäßig konvergiert. Zum Schluss des Kapitels betrachten wir die Situation am Rand.

Satz 13.8 (Abelscher Grenzwertsatz) *Ist die Potenzreihe $P(x) = \sum_{k=0}^{\infty} a_k x^k$ für $x = 1$ konvergent, so gilt*

$$\lim_{x \nearrow 1} P(x) = P(1).$$

Beweis: Nach Voraussetzung hat P Konvergenzradius $R \geq 1$. Berechne für $0 \leq x < 1$

$$\begin{aligned}
P_k(1) - P_k(x) &= \sum_{\ell=0}^{k} a_\ell (1 - x^\ell) \\
&= (1-x) \sum_{\ell=0}^{k} a_\ell \sum_{j=0}^{\ell-1} x^j \\
&= (1-x) \sum_{j=0}^{k-1} x^j \sum_{\ell=j+1}^{k} a_\ell \\
&= (1-x) \sum_{j=0}^{k-1} x^j \big(P_k(1) - P_j(1)\big).
\end{aligned}$$

Verwende $|P_k(1) - P_j(1)| < \varepsilon$ für $j, k \geq n$, und $|P_k(1)| \leq C$ für alle k. Es folgt

$$|P_k(1) - P_k(x)| \leq C(1-x) \sum_{j=0}^{n-1} x^j + \varepsilon(1-x) \sum_{j=n}^{k-1} x^j \leq C(1-x)n + \varepsilon,$$

wobei zuletzt die geometrische Reihe benutzt wurde. Mit $k \to \infty$ und $x \nearrow 1$ folgt

$$\limsup_{x \nearrow 1} |P(1) - P(x)| \leq \varepsilon.$$

$\square$

Beispiel 13.7 Nach dem Leibnizkriterium konvergiert die Reihe $x - \frac{1}{2}x^2 + \frac{1}{3}x^3 \pm \ldots$ auch für $x = 1$, also folgt aus Satz 13.8

$$\log 2 = \lim_{x \nearrow 1} \log(1+x) = 1 - \frac{1}{2} + \frac{1}{3} - + \ldots \quad \text{(Mercator 1668)}.$$

Ebenso konvergiert die Reihe des arctan auch für $x = 1$, und es ergibt sich die Darstellung

$$\frac{\pi}{4} = \arctan 1 = \lim_{x \nearrow 1} \arctan x = 1 - \frac{1}{3} + \frac{1}{5} - + \ldots \quad \text{(Gregory 1671, Leibniz 1674)}.$$

14 Das Riemannsche Integral

Das Integral einer nichtnegativen Funktion $f : I = [a,b] \to \mathbb{R}$ ist anschaulich der Flächeninhalt des Gebiets $\{(x,y) : x \in I, 0 < y < f(x)\}$. Allerdings haben wir den Flächeninhalt von Teilmengen des $\mathbb{R}^2$ noch gar nicht definiert, außer vielleicht von einfachen Gebieten wie Rechtecken, so dass die gegebene Beschreibung nicht zur Definition taugen kann. Dennoch lassen wir uns im Folgenden von dieser geometrischen Vorstellung leiten.

Definition 14.1 (Zerlegung) *Eine Zerlegung Z des Intervalls $I = [a,b]$ ist eine geordnete Menge von Punkten $a = x_0 \leq x_1 \leq \ldots \leq x_N = b$. Wir setzen $I_k = [x_{k-1}, x_k]$ sowie $\Delta x_k = x_k - x_{k-1}$ für $k = 1, \ldots, N$, und definieren die Feinheit von Z durch*

$$\Delta(Z) = \max_{1 \leq k \leq N} \Delta x_k. \tag{14.1}$$

Definition 14.2 (Riemannsche Summe) *Sei $f : I = [a,b] \to \mathbb{R}$. Die Riemannsche Summe von f zur Zerlegung Z und den Stützstellen $\xi_k \in I_k$ ist*

$$S_{Z,\xi}(f) = \sum_{k=1}^{N} f(\xi_k)\, \Delta x_k \in \mathbb{R}.$$

Die Riemannsche Summe ist ein Näherungswert für das noch zu definierende Integral. Eine konkrete Wahl der Zerlegung und der Stützstellen, zum Beispiel äquidistante Zerlegung und Intervallmittelpunkte, würde ein numerisches Verfahren zur Approximation des Integrals ergeben. Aber um das Integral stabil zu definieren, brauchen wir beliebige Approximationen, siehe Beispiel 14.2. Jedenfalls sollte bei Verfeinerung einer Zerlegung der Approximationsfehler kleiner werden. Dies führt auf B. Riemanns Begriff der Integrierbarkeit. [6]

Definition 14.3 (Riemann-Integral) *Eine beschränkte Funktion $f : I = [a,b] \to \mathbb{R}$ heißt (Riemann-)integrierbar mit Integral $S \in \mathbb{R}$, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für jede Zerlegung Z und jede Wahl ξ der Stützstellen gilt:*

$$\Delta(Z) < \delta \quad \Rightarrow \quad |S_{Z,\xi}(f) - S| < \varepsilon.$$

Wir nennen dann S das (bestimmte) Integral von f auf $[a,b]$ und schreiben

$$S = \int_I f \in \mathbb{R}.$$

Beispiel 14.1 Die konstante Funktion $f : I = [a,b] \to \mathbb{R}$, $f(x) = c$, ist integrierbar mit

$$\int_I f = c\,(b-a).$$

Denn für jede Zerlegung Z und jede Wahl der $\xi_k \in I_k$ gilt

$$S_{Z,\xi}(f) = \sum_{k=1}^{N} c\,\Delta x_k = c\,\sum_{k=1}^{N} (x_k - x_{k-1}) = c\,(b-a).$$

[6] ein anderer Zugang zum Integral stammt von H. Lebesgue, dieser wird in Analysis 3 behandelt.

Beispiel 14.2 Die Dirichletfunktion

$$\chi_{\mathbb{Q}} : [0,1] \to \mathbb{R},\ \chi_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q}, \\ 0 & \text{für } x \notin \mathbb{Q} \end{cases}$$

ist nicht (Riemann-) integrierbar. Sei zum Beispiel Z die Unterteilung mit $I_k = [\frac{k-1}{N}, \frac{k}{N}]$, also $\Delta(Z) = \frac{1}{N} \to 0$. Wählt man rationale Stützstellen ξ_k so ist $S_{Z,\xi}(\chi_{\mathbb{Q}}) = 1$, für ξ_k irrational ist dagegen $S_{Z,\xi}(\chi_{\mathbb{Q}}) = 0$.

Satz 14.1 (Linearität des Integrals) *Die Menge $\mathcal{R}(I)$ der Riemann-integrierbaren Funktionen $f : I \to \mathbb{R}$ ist ein Untervektorraum von $\mathcal{B}(I)$ und das Integral $\mathcal{R}(I) \to \mathbb{R}$, $f \mapsto \int_I f$, ist ein lineares Funktional. Es gilt also für $\lambda, \mu \in \mathbb{R}$*

$$\int_I (\lambda\, f + \mu\, g) = \lambda \int_I f + \mu \int_I g.$$

Beweis: Es gilt $S_{Z,\xi}(\lambda f + \mu g) = \lambda S_{Z,\xi}(f) + \mu S_{Z,\xi}(g)$, also folgt für $\Delta(Z)$ hinreichend klein

$$\left| S_{Z,\xi}(\lambda f + \mu g) - \left(\lambda \int_I f + \mu \int_I g\right)\right| \leq |\lambda| \left| S_{Z,\xi}(f) - \int_I f \right| + |\mu| \left| S_{Z,\xi}(g) - \int_I g \right| < \varepsilon.$$

□

Um zu zeigen, dass stetige Funktionen integrierbar sind, gehen wir in drei Schritten vor:

a) Stückweise konstante Funktionen (Treppenfunktionen) sind integrierbar.

b) Die Klasse der integrierbaren Funktionen ist abgeschlossen unter gleichmäßiger Konvergenz.

c) Stetige Funktionen lassen sich gleichmäßig durch Treppenfunktionen approximieren, und sind damit integrierbar.

Wir beginnen unser Programm, indem wir zunächst zwei Eigenschaften des Integrals zeigen.

Lemma 14.1 *Seien $f, \tilde{f} : I \to \mathbb{R}$ mit $\tilde{f}(x) = f(x)$ für alle $x \in I \backslash \{p_1, \ldots, p_r\}$. Mit $f \in \mathcal{R}(I)$ ist dann auch $\tilde{f} \in \mathcal{R}(I)$ und es gilt $\int_I \tilde{f} = \int_I f$.*

Beweis: Wir zeigen die Aussage für einen Ausnahmepunkt $p \in I$; der allgemeine Fall folgt daraus per Induktion. Es gilt für jede Zerlegung Z mit Stützstellen ξ_k

$$S_{Z,\xi}(\tilde{f}) - S_{Z,\xi}(f) = \sum_{k=1}^{N} (\tilde{f}(\xi_k) - f(\xi_k))\, \Delta x_k = (\tilde{f}(p) - f(p)) \sum_{\{k:\xi_k = p\}} \Delta x_k.$$

Es gibt höchstens zwei k mit $p \in I_k$ und $\Delta x_k > 0$. Damit schätzen wir ab

$$\begin{aligned} \left| S_{Z,\xi}(\tilde{f}) - \int_I f \right| &\leq \left| S_{Z,\xi}(\tilde{f}) - S_{Z,\xi}(f) \right| + \left| S_{Z,\xi}(f) - \int_I f \right| \\ &\leq 2 \left| \tilde{f}(p) - f(p) \right| \Delta(Z) + \left| S_{Z,\xi}(f) - \int_I f \right|, \end{aligned}$$

und die rechte Seite geht gegen Null mit $\Delta(Z) \to 0$. □

Wie Beispiel 14.2 zeigt, ist Lemma 14.1 nicht richtig für eine abzählbare Ausnahmemenge.

Lemma 14.2 *Sei $I = I_1 \cup \ldots \cup I_n$ eine Zerlegung von $I = [a,b]$ in Intervalle $I_k = [a_{k-1}, a_k]$. Ist $f : I \to \mathbb{R}$ auf jedem der Teilintervalle I_k integrierbar, so folgt $f \in \mathcal{R}(I)$ und*

$$\int_I f = \sum_{k=1}^n \int_{I_k} f.$$

Beweis: Es reicht den Fall $n = 2$ zu betrachten. Sei $I = I' \cup I''$ mit $I' = [a,p]$ und $I'' = [p,b]$, und sei $f : I \to \mathbb{R}$ integrierbar auf I' und I'', insbesondere $\|f\|_I = \max(\|f\|_{I'}, \|f\|_{I''}) < \infty$. Sei Z eine Zerlegung von I mit Punkten $a = x_0 \leq \ldots \leq x_N = b$ sowie Stützstellen $\xi_1, \ldots, \xi_N$. Wähle $r \in \{1, \ldots, N\}$ mit $p \in [x_{r-1}, x_r]$, und definiere Zerlegungen Z',Z'' von I', I'' mit Stützstellen ξ', ξ'' wie folgt:

$$\begin{aligned} Z' &= \{a = x_0 \leq \ldots \leq x_{r-1} \leq p\} \quad & \xi' &= \{\xi_1, \ldots, \xi_{r-1}, p\}, \\ Z'' &= \{p \leq x_r \leq \ldots \leq x_N = b\} \quad & \xi'' &= \{p, \xi_{r+1}, \ldots, \xi_N\}. \end{aligned}$$

Offenbar gilt $\Delta(Z'), \Delta(Z'') \leq \Delta(Z)$. Die Bilanz der Riemannschen Summen lautet

$$|S_{Z,\xi}(f) - \big(S_{Z',\xi'}(f) + S_{Z'',\xi''}(f)\big)| = |(f(\xi_r) - f(p))\,\Delta x_r| \leq 2\,\|f\|_I\,\Delta(Z).$$

Es folgt mit der Dreiecksungleichung

$$\Big|S_{Z,\xi}(f) - \Big(\int_{I'} f + \int_{I''} f\Big)\Big| \leq 2\,\|f\|_I\,\Delta(Z) + \Big|S_{Z',\xi'}(f) - \int_{I'} f\Big| + \Big|S_{Z'',\xi''}(f) - \int_{I''} f\Big|.$$

Die rechte Seite geht mit $\Delta(Z) \to 0$ gegen Null. $\square$

Folgerung 14.1 *Sei $f : I \to \mathbb{R}$ eine (Riemannsche) Treppenfunktion, d. h. es gibt eine Unterteilung $a = a_0 < a_1 < \ldots < a_n = b$ und $c_i \in \mathbb{R}$ $(i = 1, \ldots, n)$ mit $f(x) = c_i$ für alle $x \in (a_{i-1}, a_i)$. Dann ist f integrierbar und es gilt*

$$\int_I f = \sum_{i=1}^n (a_i - a_{i-1}) c_i.$$

Beweis: Nach Lemma 14.1 ist $f : [a_{i-1}, a_i] \to \mathbb{R}$ integrierbar für alle $i = 1, \ldots, n$. Aus Lemma 14.2 folgt die Behauptung. $\square$

Damit ist der erste Schritt unseres Programms erledigt. Im zweiten Schritt wollen wir für einen geeigneten Begriff von Konvergenz $f_k \to f$ folgende Aussage zeigen:

$$f_k \to f \text{ mit } f_k \text{ integrierbar} \quad \Rightarrow \quad f \text{ integrierbar und } \int_I f = \lim_{k\to\infty} \int_I f_k.$$

In Beispiel 13.3 hatten wir gesehen, dass die Stetigkeit unter punktweise Konvergenz nicht notwendig erhalten bleibt. Das gilt auch für die Integrierbarkeit.

Beispiel 14.3 Sei $q_1, q_2, \ldots$ eine Abzählung der rationalen Zahlen in $[0,1]$. Definiere

$$\chi_n : [0,1] \to \mathbb{R},\ \chi_n(x) = \begin{cases} 1 & \text{für } x \in \{q_1, \ldots, q_n\} \\ 0 & \text{sonst }. \end{cases}$$

Die Folge χ_n konvergiert punktweise gegen die Funktion $\chi_{\mathbb{Q}}$, die nach Beispiel 14.2 nicht integrierbar ist.

Beispiel 14.4 Auch wenn die punktweise Grenzfunktion integrierbar ist, lassen sich der Grenzwert und das Integral nicht notwendig vertauschen. Betrachte die Treppenfunktionen

$$f_k : [0,1] \to \mathbb{R},\ f_k(x) = \begin{cases} k & \text{für } 0 < x < \frac{1}{k} \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt $\lim_{k\to\infty} f_k(x) = 0$ für alle $x \in [0,1]$, denn

$$f_k(x) = 0 \quad \begin{cases} \text{für alle } k, & \text{falls } x = 0 \\ \text{für } k \geq \frac{1}{x}, & \text{falls } x > 0. \end{cases}$$

Also konvergiert f_k punktweise gegen $f \equiv 0$. Aber es ist

$$\lim_{k\to\infty} \int_{[0,1]} f_k = \lim_{k\to\infty} \frac{1}{k} \cdot k = 1 \neq 0 = \int_{[0,1]} f.$$

Nach den positiven Ergebnissen in Kapitel 13 ist zu hoffen, dass solche Probleme bei gleichmäßiger Konvergenz nicht auftreten. Für $I = [a,b]$ schreiben wir $|I| = b - a$.

Satz 14.2 (Standardabschätzung des Integrals) *Für $f \in \mathcal{R}(I)$ gilt*

$$\Big| \int_I f \Big| \leq |I|\, \|f\|_I.$$

Beweis: Für jede Zerlegung Z mit Stützstellen ξ_k gilt mit der Dreiecksungleichung

$$|S_{Z,\xi}(f)| \leq \sum_{k=1}^{N} |f(\xi_k)|\, \Delta x_k \leq \|f\|_I \sum_{k=1}^{N} \Delta x_k = |I|\, \|f\|_I. \tag{14.2}$$

Die Abschätzung für das Integral folgt. □

Für $f, f_k \in \mathcal{R}(I)$ haben wir nun

$$\Big| \int_I f_k - \int_I f \Big| = \Big| \int_I (f_k - f) \Big| \leq |I|\, \|f_k - f\|_I,$$

so dass aus $\|f_k - f\|_I \to 0$ auch die Konvergenz der Integrale folgt.

Satz 14.3 (Integral und gleichmäßige Konvergenz) *Konvergiert die Folge $f_k \in \mathcal{R}(I)$ gleichmäßig gegen $f : I \to \mathbb{R}$, also $\|f_k - f\|_I \to 0$, so ist $f \in \mathcal{R}(I)$ und es gilt*

$$\int_I f = \lim_{k\to\infty} \int_I f_k.$$

Beweis: Die Funktion f ist beschränkt wegen $\|f\|_I \leq \|f - f_k\|_I + \|f_k\|_I < \infty$. Mit $S_k = \int_I f_k$ gilt für k, l hinreichend groß nach Satz 14.2

$$|S_k - S_l| \leq |I|\, \|f_k - f_l\|_I \leq |I| \left(\|f_k - f\|_I + \|f - f_l\|_I \right) < \varepsilon.$$

Wir setzen $S = \lim_{k\to\infty} S_k$ und zeigen, dass f integrierbar ist mit $\int_I f = S$. Für jede Zerlegung Z mit Stützstellen ξ_j und jedes $k \in \mathbb{N}$ haben wir mit (14.2)

$$\begin{aligned} |S_{Z,\xi}(f) - S| &\leq |S_{Z,\xi}(f) - S_{Z,\xi}(f_k)| + |S_{Z,\xi}(f_k) - S_k| + |S_k - S| \\ &\leq |I|\, \|f - f_k\|_I + |S_{Z,\xi}(f_k) - S_k| + |S_k - S|. \end{aligned}$$

Zu $\varepsilon > 0$ wählen wir erst $k \in \mathbb{N}$ mit $|I|\,\|f - f_k\|_I < \varepsilon/3$ und $|S_k - S| < \varepsilon/3$. Für $\Delta(Z) < \delta$ ist dann auch $|S_{Z,\xi}(f_k) - S_k| < \varepsilon/3$, da f_k integrierbar ist. Damit ist der Satz bewiesen. □

Damit ist auch der zweite Schritt unseres Programms abgeschlossen. Es bleibt nachzuweisen, dass stetige Funktionen $f : I = [a,b] \to \mathbb{R}$ gleichmäßig durch Treppenfunktionen approximiert werden können.

Satz 14.4 *Sei $D \subset \mathbb{R}^n$ folgenkompakt und $f : D \to \mathbb{R}$ stetig. Dann ist f gleichmäßig stetig, das heißt zu $\varepsilon > 0$ gibt es ein $\delta > 0$ mit*

$$x, x' \in D, \quad |x - x'| < \delta \quad \Rightarrow \quad |f(x) - f(x')| < \varepsilon. \tag{14.3}$$

Beweis. Andernfalls gibt es für ein $\varepsilon > 0$ Punkte $x_n, x_n' \in D$, so dass gilt:

$$|x_n - x_n'| \to 0, \text{ aber } |f(x_n) - f(x_n')| \geq \varepsilon.$$

Da D kompakt, konvergiert die Folge x_n nach Übergang zu einer Teilfolge gegen ein $x_0 \in D$. Offenbar gilt dann auch $\lim_{n\to\infty} x_n' = x_0$. Da f stetig, ergibt sich der Widerspruch

$$\varepsilon \leq \lim_{n\to\infty} |f(x_n) - f(x_n')| = |f(x_0) - f(x_0)| = 0,$$

□

Beispiel 14.5 Ist $f : D \to \mathbb{R}$ stetig, aber D nicht kompakt, so muss f nicht gleichmäßig stetig sein. Betrachte zum Beispiel $f : (0, \frac{1}{\pi}] \to \mathbb{R}$, $f(x) = \sin\frac{1}{x}$. Mit $x_n = (n\pi)^{-1}$, $x_n' = (n\pi + \pi/2)^{-1}$ gilt $x_n, x_n' \to 0$, aber $|f(x_n) - f(x_n')| = 1$ für alle n.

Satz 14.5 (stetig ⇒ integrierbar) *Sei $I = [a,b]$ ein kompaktes Intervall. Dann ist jede stetige Funktion $f : I \to \mathbb{R}$ Riemann-integrierbar.*

Beweis: Wir konstruieren zu $\varepsilon > 0$ eine Treppenfunktion $\varphi : I \to \mathbb{R}$ mit $\|\varphi - f\|_I \leq \varepsilon$. Die Behauptung ergibt sich dann aus Folgerung 14.1 und Satz 14.3. Nach Satz 14.4 gibt es ein $\delta > 0$, so dass gilt:

$$x, x' \in I, \quad |x - x'| < \delta \quad \Rightarrow \quad |f(x) - f(x')| < \varepsilon.$$

Wähle eine beliebige Zerlegung Z mit Feinheit $\Delta(Z) < \delta$ und den Unterteilungspunkten $a = x_0 \leq \ldots \leq x_N = b$, und setze

$$\varphi(x) = \begin{cases} f(x_k) & \text{für } x \in (x_{k-1}, x_k] \text{ mit } k \in \{1, \ldots, N\}, \\ f(x_0) & \text{für } x = x_0. \end{cases}$$

Ist $x \in (x_{k-1}, x_k]$, so gilt $|x - x_k| < \delta$ und somit $|\varphi(x) - f(x)| = |f(x_k) - f(x)| < \varepsilon$ nach Wahl von δ. Da $\varphi(x_0) = f(x_0)$, folgt insgesamt $\|\varphi - f\|_I \leq \varepsilon$. □

Es ist nützlich, die Integrierbarkeit auch für stückweise stetige Funktionen $f : [a,b] \to \mathbb{R}$ zu haben, das heißt es gibt eine Zerlegung $a = a_0 < \ldots < a_N = b$, so dass f auf jedem Teilintervall $[a_{k-1}, a_k]$ nach eventueller Abänderung in den Endpunkten a_{k-1} und a_k stetig ist. Diese Verallgemeinerung folgt natürlich sofort aus Satz 14.5 und Lemma 14.2.

Während die bisherige Darstellung des Riemannintegrals genauso im vektorwertigen Fall zutrifft, spielt bei folgenden Aussagen die Anordnung von $\mathbb{R}$ eine Rolle.

Satz 14.6 (Monotonie des Integrals) *Sind $f, g \in \mathcal{R}(I)$, so gilt:*

$$f \le g \quad \Rightarrow \quad \int_I f \le \int_I g.$$

Insbesondere gilt $\int_I f \le \int_I |f|$, falls f, $|f| \in \mathcal{R}(I)$.

Beweis: Für jede Zerlegung Z mit Stützstellen ξ_k gilt

$$S_{Z,\xi}(f) = \sum_{k=1}^{N} f(\xi_k)\Delta x_k \le \sum_{k=1}^{N} g(\xi_k)\Delta x_k = S_{Z,\xi}(g).$$

□

Folgerung 14.2 (Mittelwertsatz der Integralrechnung) *Seien $f, \varphi : I = [a,b] \to \mathbb{R}$ stetig und $\varphi \ge 0$. Dann gibt es ein $\xi \in [a,b]$ mit*

$$\int_I f\varphi = f(\xi) \int_I \varphi.$$

Im Spezialfall $\varphi = 1$ gilt also $\int_I f = f(\xi)\,|I|$.

Beweis: Ist $\int_I \varphi = 0$, so ist $\varphi = 0$ (vergleiche Lemma 7.2), und es ist nichts zu zeigen. Nach Skalierung können wir daher annehmen, dass $\int_I \varphi = 1$. Setze $m = \min_{x\in I} f(x)$ und $M = \max_{x \in I} f(x)$. Dann gilt $m\varphi \le f\varphi \le M\varphi$, also

$$m = \int_I m\varphi \le \int_I f\varphi \le \int_I M\varphi = M.$$

Nach dem Zwischenwertsatz gibt es ein $\xi \in [a,b]$ mit $f(\xi) = \int_I f\varphi$. □

Der Vollständigkeit halber wollen wir noch eine alternative Definition des Riemann-Integrals erklären. Sei $f : I = [a,b] \to \mathbb{R}$ beschränkt und Z eine Zerlegung von I in Teilintervalle I_k der Länge Δx_k. Dann sind Ober- und Untersumme von f bzgl. Z definiert durch

$$\overline{S}_Z(f) = \sum_{k=1}^{N} (\sup_{I_k} f)\,\Delta x_k \quad \text{und} \quad \underline{S}_Z(f) = \sum_{k=1}^{N} (\inf_{I_k} f)\,\Delta x_k.$$

Nach Definition gilt $\underline{S}_Z(f) \le \overline{S}_Z(f)$. Nehmen wir zu Z den Unterteilungspunkt $\xi \in [x_{k-1}, x_k]$ hinzu, so folgt mit $I_k' = [x_{k-1}, \xi]$ und $I_k'' = [\xi, x_k]$

$$\begin{aligned} \underline{S}_{Z\cup\{\xi\}}(f) - \underline{S}_Z(f) &= (\inf_{I_k'} f)\,(\xi - x_{k-1}) + (\inf_{I_k''} f)\,(x_k - \xi) - (\inf_{I_k} f)\,(x_k - x_{k-1}) \\ &= (\inf_{I_k'} f - \inf_{I_k} f)\,(\xi - x_{k-1}) + (\inf_{I_k''} f - \inf_{I_k} f)\,(x_k - \xi). \end{aligned}$$

Mit $\inf_{I_k'} f$, $\inf_{I_k''} f$, $-\inf_{I_k} f \le \|f\|_I$ sowie $0 \le x_k - x_{k-1} \le \Delta(Z)$ erhalten wir

$$0 \le \underline{S}_{Z\cup\{\xi\}}(f) - \underline{S}_Z(f) \le 2\|f\|_I\,\Delta(Z),$$

und analog für die Obersummen

$$0 \ge \overline{S}_{Z\cup\{\xi\}}(f) - \overline{S}_Z(f) \ge -2\|f\|_I\,\Delta(Z).$$

Für beliebige Zerlegungen Z, Z' ergibt sich per Induktion

$$\underline{S}_Z(f) \leq \underline{S}_{Z\cup Z'}(f) \leq \overline{S}_{Z\cup Z'}(f) \leq \overline{S}_{Z'}(f). \tag{14.4}$$

Bezeichnet N die Zahl der Teilintervalle von Z', so folgt ebenfalls induktiv

$$\underline{S}_{Z\cup Z'}(f) - 2N\,\|f\|_I\,\Delta(Z) \leq \underline{S}_Z(f) \leq \overline{S}_Z(f) \leq \overline{S}_{Z\cup Z'}(f) + 2N\|f\|_I\,\Delta(Z). \tag{14.5}$$

Wir definieren nun das Ober- bzw. Unterintegral von f durch

$$\begin{aligned}\overline{S}(f) &= \inf\{\overline{S}_Z(f) : Z \text{ ist Zerlegung von } I\},\\ \underline{S}(f) &= \sup\{\underline{S}_Z(f) : Z \text{ ist Zerlegung von } I\}.\end{aligned}$$

Aus (14.4) folgt $\underline{S}(f) \leq \overline{S}(f)$.

Satz 14.7 *Eine beschränkte Funktion $f : I \to \mathbb{R}$ ist genau dann (Riemann-) integrierbar, wenn ihr Ober- und Unterintegral übereinstimmen, und es gilt dann*

$$\int_I f = \underline{S}(f) = \overline{S}(f).$$

Beweis: Ist f Riemannintegrierbar, so folgt für $\Delta(Z) < \delta$

$$\int_I f - \varepsilon \leq \inf_\xi S_{Z,\xi}(f) = \underline{S}_Z(f) \leq \underline{S}(f) \leq \overline{S}(f) \leq \overline{S}_Z(f) = \sup_\xi S_{Z,\xi}(f) \leq \int_I f + \varepsilon.$$

Umgekehrt sei Z' eine Zerlegung mit $\underline{S}_{Z'}(f) > \underline{S}(f) - \varepsilon/2$, und N sei die Zahl der Teilintervalle von Z'. Für jede Zerlegung Z gilt $\underline{S}_{Z\cup Z'}(f) \geq \underline{S}_{Z'}(f)$, also folgt mit (14.5)

$$S_{Z,\xi}(f) \geq \underline{S}_Z(f) \geq \underline{S}_{Z\cup Z'}(f) - 2N\|f\|_I\,\Delta(Z) > \underline{S}(f) - \varepsilon/2 - 2N\|f\|_I\,\Delta(Z).$$

Für $\Delta(Z) < \delta$ folgt $S_{Z,\xi}(f) > \underline{S}(f) - \varepsilon$. Die Abschätzung nach oben ist analog. $\square$

15 Ableitung und Integral

Wir kommen nun zu dem zentralen, von Newton und Leibniz studierten Zusammenhang zwischen Differentiation und Integration.

Definition 15.1 *Sei $f : (a,b) \to \mathbb{R}$. Eine differenzierbare Funktion $F : (a,b) \to \mathbb{R}$ heißt Stammfunktion von f, wenn gilt:*

$$F' = f \quad \Leftrightarrow \quad F'(x) = f(x) \quad \textit{für alle } x \in (a,b).$$

Satz 15.1 *Ist F eine Stammfunktion von f auf (a,b), so ist jede Stammfunktion von f auf (a,b) von der Form $F + c$, für eine Konstante $c \in \mathbb{R}$.*

Beweis: Sei G auch Stammfunktion von f auf (a,b). Es folgt

$$(G-F)' = G' - F' = f - f = 0.$$

Nach Folgerung 10.1 gibt es eine Konstante $c \in \mathbb{R}$ mit $G - F = c$, also $G = F + c$. $\square$

Die Gleichung $F' = f$ ist ein elementares Beispiel für eine Differentialgleichung. Folgerung 15.1 sagt aus, dass eine Lösung der Gleichung bis auf eine additive Konstante $c \in \mathbb{R}$ eindeutig bestimmt ist. Auch hier schließt sich die Frage nach der Existenz an:

Für welche f ist die Differentialgleichung $F' = f$ auf dem Intervall (a,b) lösbar?

Um eine Antwort zu geben, muss die Definition des Integrals noch etwas erweitert werden: sei I kompaktes Intervall mit Randpunktmenge $\{a,b\}$. Dann setzen wir

$$\int_a^b f(x)\,dx = \begin{cases} \int_I f & \text{falls } a \le b, \\ -\int_I f & \text{falls } a > b. \end{cases}$$

Es gilt dann für beliebige $a,b,c \in \mathbb{R}$, sofern f auf allen Intervallen Riemann-integrierbar ist,

$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx. \tag{15.1}$$

Für $a \le b \le c$ ist das Lemma 14.2, und allgemein folgt (15.1) dann durch Vertauschung von a, b und c. Die Notation $f(x)\,dx$ ist rein formal, ein dx wurde in der Vorlesung nicht definiert. Aber sie ist nützlich, zum Beispiel wenn f noch von weiteren Variablen $y, z, a, \ldots$ abhängt: es wird spezifiziert, bezüglich welcher Variablen integriert werden soll. Außerdem erinnert die Notation an die Riemannschen Summen, mit denen das Integral definiert wurde.

Satz 15.2 (Hauptsatz der Differential- und Integralrechnung) *Sei $f : I = [a,b] \to \mathbb{R}$ stetig. Dann ist für jedes $x_0 \in I$ die Funktion*

$$F : I \to \mathbb{R}, \quad F(x) = \int_{x_0}^x f(\xi)\,d\xi$$

eine Stammfunktion von f, das heißt es gilt $F'(x) = f(x)$ für alle $x \in I$.

Bemerkung. In den Endpunkten des Intervalls ist dies im Sinne der einseitigen Ableitungen $F'_+(a) = f(a)$ bzw. $F'_-(b) = f(b)$ zu verstehen.

Beweis: Die Funktion F ist definiert nach Satz 14.5. Wir berechnen für $x, x+h \in I$, $h \neq 0$, mit der Standardabschätzung, Satz 14.2,[7]

$$\begin{aligned}
\left|\frac{F(x+h)-F(x)}{h} - f(x)\right| &= \frac{1}{|h|}\left|\int_{x_0}^{x+h} f(\xi)\,d\xi - \int_{x_0}^{x} f(\xi)\,d\xi - hf(x)\right| \\
&= \frac{1}{|h|}\left|\int_{x}^{x+h} \big(f(\xi)-f(x)\big)\,d\xi\right| \\
&\leq \sup_{\xi\in[x,x+h]} |f(\xi)-f(x)|.
\end{aligned}$$

Da f stetig in x ist, geht die rechte Seite mit $h \to 0$ gegen Null. □

Folgerung 15.1 *Die Funktion $F \in C^0(I)$, $I = [a,b]$, sei Stammfunktion von $f \in C^0(I)$ auf (a,b). Dann gilt für beliebiges $x_0 \in I$*

$$F(x) = F(x_0) + \int_{x_0}^{x} f(\xi)\,d\xi \quad \textit{für alle } x \in I. \tag{15.2}$$

Beweis: Nach Satz 15.1 und Satz 15.2 gibt es ein $c \in \mathbb{R}$ mit

$$F(x) = c + \int_{x_0}^{x} f(\xi)\,d\xi \quad \text{für alle } x \in (a,b).$$

Durch Grenzübergang folgt das auch für $x = a$ bzw. $x = b$, zum Beispiel gilt

$$F(b) = \lim_{x\nearrow b} F(x) \quad \text{und} \quad \int_{x_0}^{b} f(\xi)\,d\xi = \lim_{x\nearrow b}\int_{x_0}^{x} f(\xi)\,d\xi.$$

Das erste gilt wegen $F \in C^0(I)$, das zweite folgt mit

$$\left|\int_{x_0}^{b} f(\xi)\,d\xi - \int_{x_0}^{x} f(\xi)\,d\xi\right| = \left|\int_{x}^{b} f(\xi)\,d\xi\right| \leq |b-x|\,\|f\|_I \to 0.$$

Setze nun in die Gleichheit $x = x_0$ ein, es folgt $c = F(x_0)$. □

Folgerung 15.2 (Berechnung von Integralen mit Stammfunktionen) *Die Funktion $F \in C^0(I)$, $I = [a,b]$, sei Stammfunktion von $f \in C^0(I)$ auf (a,b). Dann gilt*

$$\int_a^b f(x)\,dx = F(b) - F(a) =: \big[F(x)\big]_{x=a}^{x=b}$$

Beweis: Folgt aus (15.2) mit $x_0 = a$, $x = b$. □

Mit dem Hauptsatz bzw. mit Folgerung 15.2 lassen sich die Differentiationsregeln aus Kapitel 9 in Integrationsregeln übersetzen. Die resultierenden Integrationsregeln sind sowohl theoretisch wichtig als auch zur Berechnung konkreter Integrale.

[7]für $h < 0$ setze $[x, x+h] := [x+h, x]$.

Satz 15.3 (Partielle Integration) *Seien $f, g \in C^1(I)$ mit $I = [a, b]$. Dann gilt*

$$\int_a^b fg' = \big[f(x)\,g(x)\big]_{x=a}^{x=b} - \int_a^b f'g.$$

Beweis: Es gilt nach der Produktregel

$$(fg)' = f'g + fg' \quad \text{auf } I = [a, b].$$

Folgerung 15.2 liefert die Behauptung. □

Satz 15.4 (Substitutions- oder Transformationsregel) *Sei $I = [a, b]$, $I^* = [\alpha, \beta]$ und $\varphi \in C^1(I)$ mit $\varphi(I) \subset I^*$. Dann gilt für $f \in C^0(I^*)$*

$$\int_{\varphi(a)}^{\varphi(b)} f(y)\,dy = \int_a^b f\big(\varphi(x)\big)\varphi'(x)\,dx.$$

Beweis: Wähle nach Satz 15.2 eine Stammfunktion $F \in C^1(I^*)$ von f. Dann gilt

$$\begin{aligned}
\int_{\varphi(a)}^{\varphi(b)} f(y)\,dy &= \big[F(y)\big]_{y=\varphi(a)}^{y=\varphi(b)} \quad \text{(Folgerung 15.2)}\\
&= \big[F\big(\varphi(x)\big)\big]_{x=a}^{x=b}\\
&= \int_a^b (F \circ \varphi)'(x)\,dx \quad \text{(Folgerung 15.2)}\\
&= \int_a^b F'\big(\varphi(x)\big)\varphi'(x)\,dx \quad \text{(Kettenregel)}\\
&= \int_a^b f\big(\varphi(x)\big)\,\varphi'(x)\,dx.
\end{aligned}$$

□

Für die Anwendung der Substitutionsregel ist folgendes *Kochrezept* nützlich: bei einem gegebenen Integral $\int_\alpha^\beta f(y)\,dy$ möchten wir $y = y(x)$ substituieren. Berechne dazu

$$y = y(x) \quad \Rightarrow \quad dy = y'(x)\,dx.$$

Bestimme die neuen Intervallgrenzen a, b durch Auflösen der Gleichungen $\alpha = y(a)$, $\beta = y(b)$. Es ergibt sich die Substitutionsformel

$$\int_\alpha^\beta f(y)\,dy = \int_a^b f\big(y(x)\big)\,y'(x)\,dx.$$

Wir zeigen nun exemplarisch, wie die Integrationsregeln angewandt werden. Zunächst ergeben sich direkt Integrationsformeln, wenn die Stammfunktion bekannt ist.

Beispiel 15.1

$$\begin{aligned}
\int_a^b x^\alpha\,dx &= \left[\frac{x^{\alpha+1}}{\alpha+1}\right]_{x=a}^{x=b} \quad (\alpha \in \mathbb{R}\backslash\{-1\},\ a, b > 0),\\
\int_a^b \frac{dx}{x} &= \big[\log x\big]_{x=a}^{x=b} \quad (a, b > 0),\\
\int_a^b \frac{dx}{\sqrt{1-x^2}} &= \big[\arcsin x\big]_{x=a}^{x=b}, \quad (-1 < a, b < 1).
\end{aligned}$$

Ein Beispiel für die Anwendung der partiellen Integration ist

Beispiel 15.2

$$\int_1^x \log u\,du = \int_1^x 1\cdot\log u\,du = \big[u\log u\big]_{u=1}^{u=x} - \int_1^x u\frac{1}{u}\,du = x\log x-(x-1).$$

Eine schöne Anwendung von Satz 15.3 ist das

Beispiel 15.3 (Wallis-Produkt) Wir berechnen hier $A_n=\int_0^{\pi/2}\sin^n x\,dx$. Offenbar gilt

$$A_0=\pi/2 \quad \text{und } A_1=\big[-\cos x\big]_{x=0}^{x=\pi/2}=1.$$

Für $n\geq 1$ leiten wir durch partielle Integration eine Rekursionsformel her:

$$\begin{aligned} A_{n+1} &= \int_0^{\pi/2}\sin x\,\sin^n x\,dx \\ &= \big[\underbrace{-\cos x\,\sin^n x}_{=0}\big]_{x=0}^{x=\pi/2} + n\int_0^{\pi/2}\cos^2 x\,\sin^{n-1}x\,dx \\ &= n\int_0^{\pi/2}\sin^{n-1}x\,dx - n\int_0^{\pi/2}\sin^{n+1}x\,dx, \end{aligned}$$

wobei wir $\cos^2 x = 1-\sin^2 x$ benutzt haben. Es folgt

$$A_{n+1}=\frac{n}{n+1}A_{n-1}\quad (n\geq 1).$$

Durch Induktion erhalten wir

$$\begin{aligned} A_{2n} &= \frac{2n-1}{2n}\cdot\frac{2n-3}{2n-2}\cdot\ldots\cdot\frac{1}{2}\cdot A_0=\left(\prod_{k=1}^n\frac{2k-1}{2k}\right)\frac{\pi}{2}, \\ A_{2n+1} &= \frac{2n}{2n+1}\cdot\frac{2n-2}{2n-1}\cdot\ldots\cdot\frac{2}{3}\cdot A_1=\left(\prod_{k=1}^n\frac{2k}{2k+1}\right)1. \end{aligned}$$

Es folgt weiter

$$\frac{A_{2n}}{A_{2n+1}}=\frac{\pi}{2}\cdot\prod_{k=1}^n\frac{4k^2-1}{4k^2}.$$

Nun gilt $A_{n+1}=\int_0^{\pi/2}\sin^{n+1}x\,dx\leq\int_0^{\pi/2}\sin^n x\,dx=A_n$, und es folgt

$$1\leq\frac{A_n}{A_{n+1}}\leq\frac{A_{n-1}}{A_{n+1}}=\frac{n+1}{n}\longrightarrow 1.$$

Daraus ergibt sich die Produktdarstellung von Wallis

$$\frac{\pi}{2}=\prod_{k=1}^\infty\frac{4k^2}{4k^2-1}=\frac{2\cdot 2}{1\cdot 3}\cdot\frac{4\cdot 4}{3\cdot 5}\cdot\ldots$$

Als nächstes behandeln wir Beispiele zur Substitutionsregel.

Beispiel 15.4 (Lineare Parameterwechsel) Mit der Substitution $y = (x - x_0)/m$ haben wir $dy = 1/m\, dx$ und $x = x_0 + my$, also neue Grenzen $a = x_0 + m\alpha$ und $b = x_0 + m\beta$. Es folgt

$$\int_\alpha^\beta f(y)\, dy = \frac{1}{m} \int_{x_0+m\alpha}^{x_0+m\beta} f\Big(\frac{x - x_0}{m}\Big)\, dx.$$

Beispiel 15.5 (Integration von Ableitungen)

$$\begin{aligned}
\int_a^b \frac{f'(x)}{f(x)}\, dx &= \int_a^b (\log f)'(x)\, dx = \big[\log f(x)\big]_{x=a}^{x=b} \quad (f > 0),\\
\int_a^b u\sqrt{1+u^2}\, du &= \frac{1}{3} \int_a^b \big((1+u^2)^{\frac{3}{2}}\big)'\, du = \Big[\frac{1}{3}(1+u^2)^{\frac{3}{2}}\Big]_{u=a}^{u=b}.\\
\int_a^b F'\big(f(x)\big) f'(x)\, dx &= \big[F\big(f(x)\big)\big]_{x=a}^{x=b}.
\end{aligned}$$

Beispiel 15.6 (Flächeninhalt unter Hyperbel) Zu berechnen ist das Integral

$$A(x) = \int_1^x \sqrt{u^2 - 1}\, du \qquad \text{für } x \geq 1.$$

Wir substituieren $u = \cosh t$ und erhalten $du = \sinh t\, dt$, $t = \operatorname{Arcosh} u$, also

$$\begin{aligned}
A(x) &= \int_0^{\operatorname{Arcosh} x} \sinh^2 t\, dt\\
&= \frac{1}{4} \int_0^{\operatorname{Arcosh} x} (e^{2t} + e^{-2t} - 2)\, dt\\
&= \Big[\frac{1}{8}(e^{2t} - e^{-2t})\Big]_{t=0}^{t = \operatorname{Arcosh} x} - \frac{1}{2} \operatorname{Arcosh} x\\
&= \frac{1}{2}\left(\Big[\frac{e^t + e^{-t}}{2}\, \frac{e^t - e^{-t}}{2}\Big]_{t=0}^{t = \operatorname{Arcosh} x} - \operatorname{Arcosh} x\right)\\
&= \frac{1}{2}\big(x\sqrt{x^2 - 1} - \operatorname{Arcosh} x\big).
\end{aligned}$$

Für rationale Funktionen, also Quotienten von Polynomen, hat man ein spezielles Integrationsverfahren, die Partialbruchzerlegung, die wir hier nur an einem Beispiel vorführen:

Beispiel 15.7 (Partialbruchzerlegung) Um das Integral $\int_{-1/2}^{1/2} \frac{dx}{1-x^2}$ zu berechnen, machen wir den Ansatz

$$\frac{1}{1 - x^2} = \frac{1}{(1+x)(1-x)} \stackrel{!}{=} \frac{A}{1-x} + \frac{B}{1+x} = \frac{(A-B)x + (A+B)}{1 - x^2}.$$

Der Koeffizientenvergleich ergibt $A = B = 1/2$, also folgt

$$\int_{-1/2}^{1/2} \frac{dx}{1 - x^2} = \int_{-1/2}^{1/2} \left(\frac{1/2}{1-x} + \frac{1/2}{1+x}\right)\, dx = \frac{1}{2}\Big[\log \frac{1+x}{1-x}\Big]_{x=-\frac{1}{2}}^{x=\frac{1}{2}} = \frac{1}{2}(\log 3 - \log 1/3) = \log 3.$$

Bei der Definition des Riemannschen Integrals ist das Definitionsintervall $I = [a, b]$ nach Voraussetzung kompakt. Wir wollen ganz kurz erläutern, wie auch unendliche Integrationsintervalle $-\infty \le a < b \le \infty$ und in den Intervallgrenzen unbeschränkte Funktionen im Rahmen des Riemann-Integrals behandelt werden können.

Definition 15.2 (uneigentliches Riemann-Integral) *Sei $I = [a, b)$ mit $a < b \le \infty$. Die Funktion $f : I \to \mathbb{R}$ sei Riemann-integrierbar auf $[a, b']$ für alle $b' < b$. Falls $\lim_{x \nearrow b} \int_a^x f(\xi)\, d\xi$ existiert, so heißt das Integral $\int_a^b f(\xi)\, d\xi$ konvergent (oder existent) und wir setzen*

$$\int_a^b f(\xi)\, d\xi = \lim_{x \nearrow b} \int_a^x f(\xi)\, d\xi.$$

Die Konvergenzaussagen für das uneigentliche Integral sind analog zu den Konvergenzkriterien für Reihen. Das folgende Lemma entspricht dabei dem Cauchykriterium.

Lemma 15.1 *In der Situation von Definition 15.2 ist $\int_a^b f(x)\, dx$ genau dann konvergent, wenn es zu jedem $\varepsilon > 0$ ein $b' < b$ gibt mit*

$$\Big| \int_{x_1}^{x_2} f(x)\, dx \Big| < \varepsilon \quad \text{für alle } x_1, x_2 > b'.$$

Beweis: Setze $F(x) = \int_a^x f(\xi)\, d\xi$ für $x \in [a, b)$. Existiert das uneigentliche Integral, das heißt $F(x) \to S$ mit $x \nearrow b$, so gibt es ein $b' < b$ mit $|F(x) - S| < \varepsilon/2$ für $x > b'$ und es folgt

$$|F(x_1) - F(x_2)| \le |F(x_1) - S| + |F(x_2) - S| < \varepsilon \quad \text{für } x_{1,2} > b'.$$

Umgekehrt wählen wir eine Folge $x_k \nearrow b$. Nach Voraussetzung ist dann $F(x_k)$ Cauchyfolge, das heißt $S = \lim_{k \to \infty} F(x_k)$ existiert. Aber dann folgt sogar $\lim_{x \nearrow b} F(x) = S$. $\Box$

Hier sind einige Beispiele von uneigentlichen Riemann-Integralen.

Beispiel 15.8

$$\begin{aligned}
\int_1^\infty x^\alpha\, dx &= \begin{cases} \lim_{R \nearrow \infty} \left[\dfrac{x^{\alpha+1}}{\alpha+1}\right]_{x=1}^{x=R} = -\dfrac{1}{\alpha+1} & \text{für } \alpha < -1, \\ \text{divergent} & \text{für } \alpha \ge -1. \end{cases} \\
\int_0^1 x^\alpha\, dx &= \begin{cases} \text{divergent} & \text{für } \alpha \le -1, \\ \lim_{\varepsilon \searrow 0} \left[\dfrac{x^{\alpha+1}}{\alpha+1}\right]_{x=\varepsilon}^{x=1} = \dfrac{1}{\alpha+1} & \text{für } \alpha > -1. \end{cases} \\
\int_0^1 \frac{dx}{\sqrt{1-x^2}} &= \lim_{x \nearrow 1} \int_0^x \frac{d\xi}{\sqrt{1-\xi^2}} = \lim_{x \nearrow 1} \arcsin x = \pi/2.
\end{aligned}$$

In den vorangegangenen Beispielen sind die Integranden positiv. Das uneigentliche Integral $\int_a^b f(x)\, dx$ heißt absolut konvergent, wenn $\int_a^b |f(x)|\, dx$ konvergiert. Aus Lemma 15.1 folgt sofort, dass ein absolut konvergentes Integral konvergiert. Wie bei Reihen impliziert die Konvergenz aber nicht umgekehrt die absolute Konvergenz. Ein simples Beispiel ist das uneigentliche Integral $\int_1^\infty f(x)\, dx$, wobei

$$f(x) = \frac{(-1)^{k-1}}{k} \text{ für } k = [x].$$

Für die Existenz des uneigentlichen Integrals auf einem beidseitig offenen Intervall (a, b) wählt man einen Zwischenpunkt $c \in (a, b)$ und verlangt die Existenz der Integrale auf $(a, c]$ und $[c, b)$. Das Integral über (a, b) ergibt sich dann als Summe. Es ist leicht zu sehen, dass diese Definition nicht von der Wahl des Zwischenpunkts c abhängt. Ein Beispiel ist

$$\int_{-\infty}^{\infty} \frac{dx}{1 + x^2} = \pi.$$