

Michael Růžička

Analysis

aufbauend auf die Skripte von
Prof. em. Dr. Dr. h.c. Rolf Schneider

Inhaltsverzeichnis

Analysis I

Sprech- und Schreibweisen	3
1 Die reellen Zahlen	5
1.1 Die Körperaxiome	5
1.2 Die Anordnungsaxiome	8
1.3 Das Vollständigkeitsaxiom	12
1.4 Natürliche Zahlen und vollständige Induktion	13
2 Abbildungen	23
2.1 Der Funktionsbegriff	23
2.2 Abzählbarkeit	27
3 Konvergenz	31
3.1 Konvergente Folgen	31
3.2 Reihen	42
3.3 Die Exponentialreihe	52
4 Topologie in \mathbb{R} und Stetigkeit	57
4.1 Topologische Eigenschaften	57
4.2 Grenzwerte von Funktionen und Stetigkeit	61
4.3 Eigenschaften stetiger Funktionen	69
5 Spezielle Funktionen	75
5.1 Logarithmus und allgemeine Potenz	75
5.2 Die Exponentialfunktion im Komplexen	78
5.3 Die trigonometrischen Funktionen	84
6 Differenzierbare Funktionen	91
6.1 Die Ableitung	91
6.2 Eigenschaften differenzierbarer Funktionen	98
6.3 Höhere Ableitungen und Taylorformel	104

7	Integration	115
7.1	Regelfunktionen	115
7.2	Das Integral einer Regelfunktion	119
7.3	Integration und Differentiation	125
7.4	Berechnung von Integralen	128
7.5	Parameterabhängige Integrale	140
7.6	Uneigentliche Integrale	145
8	Funktionenreihen	151
8.1	Konvergenz von Funktionenfolgen	151
8.2	Potenzreihen	157
8.3	Fourierreihen	166

Analysis I

Sprech- und Schreibweisen

Mathematische Sachverhalte werden mithilfe von Aussagesätzen (Aussagen) formuliert und benutzen oft die Mengenschreibweise. Dies gilt analog für das mathematische Argumentieren und Beweisen. Einige der dabei üblichen Ausdrucksweisen wollen wir jetzt zusammenstellen, weitere folgen bei Bedarf.

Unter einer Menge versteht man eine Zusammenfassung von unterscheidbaren mathematischen Objekten; diese Gesamtheit wird dann als neues Objekt angesehen. Diese plausibel klingende Erklärung pflegt man in der Mathematik als „naiv“ zu bezeichnen. Sie bedarf nämlich der Präzisierung. Eine strenge Begründung der Mengenlehre ist aber zum jetzigen Zeitpunkt weder notwendig noch angebracht. Erfahrungsgemäß kann man in der elementaren Analysis mit diesem naiven Gebrauch von Mengen sehr gut arbeiten. Wir werden also nicht genau definieren, was eine Menge sein soll, sondern lediglich gewisse Mengen als gegeben ansehen und gewisse Vorschriften, neue Mengen zu bilden, als zulässig betrachten. Mit anderen Worten, wir verwenden die Mengenlehre lediglich als eine bequeme und klare Sprechweise und nicht als das axiomatisch begründbare Fundament der Mathematik, als das sie sich in einem späteren Stadium darstellt.

Folgende Abkürzungen werden verwendet:

$x \in M$ bedeutet: x ist Element von M

$x \notin M$ bedeutet: x ist nicht Element von M

Nach Definition werden zwei Mengen M, N genau dann (dies heißt: dann und nur dann) als gleich angesehen, geschrieben $M = N$, wenn sie dieselben Elemente enthalten. Die Menge, die kein Element enthält, heißt *leere Menge*, sie wird mit \emptyset bezeichnet. Für konkrete Mengen, die im Verlauf einer mathematischen Darstellung auftreten, sind verschiedenartige Bezeichnungen üblich, zum Beispiel:

$\{a_1, \dots, a_n\}$ ist die Menge, deren Elemente genau die Objekte a_1, \dots, a_n sind (sie dürfen in beliebiger Reihenfolge hingeschrieben werden).

$\{x \in M \mid A(x)\}$ ist die Menge aller $x \in M$ (M eine gegebene Menge), die die Eigenschaft A besitzen (manchmal wird die Grundmenge M weggelassen).

Aus gegebenen Mengen kann man neue bilden nach den folgenden Vorschriften:

$$M_1 \cap M_2 := \{x \mid x \in M_1 \text{ und } x \in M_2\}$$

ist der *Durchschnitt* von M_1 und M_2 ,

$$M_1 \cup M_2 := \{x \mid x \in M_1 \text{ oder } x \in M_2\}$$

ist die *Vereinigung* von M_1 und M_2 ,

$$M_1 \setminus M_2 := \{x \mid x \in M_1 \text{ und } x \notin M_2\}$$

ist die (*mengentheoretische*) *Differenz* von M_1 und M_2 .

M_1 heißt *Teilmenge* von M_2 , geschrieben $M_1 \subset M_2$, wenn jedes Element von M_1 auch Element von M_2 ist. Offenbar gilt $M_1 = M_2$ genau dann, wenn $M_1 \subset M_2$ und $M_2 \subset M_1$ ist.

Mathematische Aussagen sind sprachliche Gebilde, von denen es sinnvoll ist zu fragen, ob sie wahr oder falsch sind. Aussagen sind entweder wahr (W) oder falsch (F). Aus gegebenen Aussagen A, B kann man mit den folgenden Verknüpfungen neue bilden:

$\neg A$	(nicht A)	Negation
$A \wedge B$	(A und B)	Konjunktion
$A \vee B$	(A oder B)	Adjunktion
$A \Rightarrow B$	(A impliziert B)	Implikation
$A \Leftrightarrow B$	(A und B sind äquivalent)	Äquivalenz

Die folgende Tafel legt fest, wie diese Verknüpfungen gebraucht werden.

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
W	W	F	W	W	W	W
W	F	F	F	W	F	F
F	W	W	F	W	W	F
F	F	W	F	F	W	W

Folgenden Bestandteile von Aussagen werden als Quantoren bezeichnet:

$\forall x$ bedeutet: für alle x
 $\exists x$ bedeutet: es gibt ein x

Beide Quantoren werden immer nur in Verbindung mit Aussagen verwendet, die von einer Variablen abhängen, z.B. $\exists x \in M : A(x)$ bedeutet: Es gibt ein Element x aus der Menge M , so dass die Aussage $A(x)$ wahr ist. Aussagen mit Quantoren können auch negiert werden. Diese Negationen sind wie folgt definiert:

$\neg(\forall x \in M : A(x))$ wird zu $\exists x \in M : \neg A(x)$
 $\neg(\exists x \in M : A(x))$ wird zu $\forall x \in M : \neg A(x)$

1 Die reellen Zahlen

Grundlage der Analysis sind die reellen Zahlen. Was sind Zahlen? Natürlich kann jeder mit Zahlen umgehen und hat genügend Erfahrung, um von ihrer Nützlichkeit überzeugt zu sein. Die Frage, was eine Zahl „wirklich“ ist, wird man schwerlich vernünftig beantworten können; das ist aber auch nicht notwendig, sondern es genügt zu wissen, wie man mit Zahlen umgeht. Wir wollen hier nicht den Versuch machen, explizit zu sagen, was reelle Zahlen sind, sondern lediglich gewisse Regeln festlegen, nach denen man mit reellen Zahlen rechnen und anderweitig umgehen kann. Wir werden dabei bemüht sein, so wenige Regeln wie möglich an den Anfang zu stellen und aus diesen alle anderen Regeln als Folgerungen herzuleiten. Dieses Vorgehen ist heute typisch für den Aufbau einer mathematischen Theorie: Gewisse Grundbegriffe werden undefiniert an den Anfang gestellt und gewisse Aussagen über diese Grundbegriffe als wahr angenommen. Weitere Begriffe dürfen dann nur durch Rückgriff auf bereits gegebene definiert werden, und weitere Aussagen müssen aus bereits bekannten hergeleitet werden. Die an den Anfang gestellten Aussagen bezeichnet man als *Axiome*. Sie haben in gewissem Sinn den Charakter von „Spielregeln“. In diesem Sinn wollen wir also jetzt ein Axiomensystem für reelle Zahlen angeben und erste Folgerungen daraus herleiten. Wir teilen das Axiomensystem in mehrere Axiomgruppen auf, die wir einzeln betrachten.

1.1 Die Körperaxiome

Gegeben sei eine nichtleere Menge \mathbb{R} . Ihre Elemente werden wir später *reelle Zahlen* nennen. Je zwei Elementen $a, b \in \mathbb{R}$ sei ein Element $a + b$, genannt ihre *Summe*, und ein Element $a \cdot b$ (meist ab geschrieben), ihr *Produkt*, eindeutig zugeordnet. Dabei sollen folgende Aussagen gelten:

Die Körperaxiome

$$(1.1) \quad \forall a, b, c \in \mathbb{R} : a + (b + c) = (a + b) + c$$

„Assoziativgesetz der Addition“

(Bedeutung der Beklammerung: $a + (b + c) := a + d$ mit $d := b + c$, und analog in anderen Fällen)

$$(1.2) \quad \forall a, b \in \mathbb{R} : a + b = b + a$$

„Kommutativgesetz der Addition“

$$(1.3) \quad \exists 0 \in \mathbb{R} \quad \forall a \in \mathbb{R} : a + 0 = a$$

„Existenz eines neutralen Elements der Addition“

$$(1.4) \quad \forall a \in \mathbb{R} \quad \exists -a \in \mathbb{R} : a + (-a) = 0$$

„Existenz eines inversen Elements der Addition“

.....

$$(1.5) \quad \forall a, b, c \in \mathbb{R} : a(bc) = (ab)c$$

$$(1.6) \quad \forall a, b \in \mathbb{R} : ab = ba$$

$$(1.7) \quad \exists 1 \in \mathbb{R} \setminus \{0\} \quad \forall a \in \mathbb{R} : a \cdot 1 = a$$

$$(1.8) \quad \forall a \in \mathbb{R} \setminus \{0\} \quad \exists a^{-1} \in \mathbb{R} : aa^{-1} = 1$$

.....

$$(1.9) \quad \forall a, b, c \in \mathbb{R} : a(b + c) = ab + ac$$

„Distributivgesetz“

Bemerkung. Wegen der Assoziativgesetze dürfen wir schreiben:

$$a + (b + c) = (a + b) + c =: a + b + c,$$

$$a(bc) = (ab)c =: abc.$$

Folgerungen aus den Axiomen

1.10 Behauptung. Die Zahl 0 ist eindeutig bestimmt.

Beweis. Sei $0' \in \mathbb{R}$ ein weiteres Element mit $a + 0' = a$ für alle $a \in \mathbb{R}$. Speziell gilt dann $0 + 0' = 0$. Nach (1.3) gilt $0' + 0 = 0'$. Nach (1.2) ist $0 + 0' = 0' + 0$. Somit folgt aus den letzten drei Gleichungen $0 = 0'$. ■

1.11 Behauptung. Zu $a \in \mathbb{R}$ ist $-a$ eindeutig bestimmt.

Beweis. Sei $a' \in \mathbb{R}$ ein Element mit $a + a' = 0$. Addition von $-a$ ergibt $-a + (a + a') = -a$. Aus (1.1) folgt $(-a + a) + a' = -a$. Wegen $-a + a \stackrel{(1.2)}{=} a + (-a) \stackrel{(1.4)}{=} 0$ und $0 + a' \stackrel{(1.2)}{=} a' + 0 \stackrel{(1.3)}{=} a'$ folgt $a' = -a$. ■

1.12 Behauptung. $-0 = 0$.

Beweis. Es gilt $0 \stackrel{(1.4)}{=} 0 + (-0) \stackrel{(1.2)}{=} -0 + 0 \stackrel{(1.3)}{=} -0$. ■

Bezeichnungen. Statt $a + (-b)$ schreibt man $a - b$, genannt *Differenz* von a und b .

1.13 Behauptung. Sei $a, b \in \mathbb{R}$. Die Gleichung $a + x = b$ hat eine eindeutig bestimmte Lösung $x \in \mathbb{R}$.

Beweis. Wegen $a + (b - a) \stackrel{(1.2)}{=} a + (-a + b) \stackrel{(1.1)}{=} (a + (-a)) + b = 0 + b = b + 0 = b$ ist $x = b - a$ eine Lösung. Sei $y \in \mathbb{R}$ ein Element mit $a + y = b$. Dann ist $b - a = -a + b = -a + (a + y) = (-a + a) + y = 0 + y = y$. ■

1.14 Behauptung. Für alle $a \in \mathbb{R}$ gilt $-(-a) = a$.

Beweis. Nach Definition ist $-(-a)$ das Inverse von $-a$, d.h. es gilt $(-a) + (-(-a)) = 0$. Andererseits ist $(-a) + a = a + (-a) = 0$. Aus Behauptung 1.11 folgt $-(-a) = a$. ■

1.15 Behauptung. Für alle $a, b \in \mathbb{R}$ gilt $-(a + b) = -a - b$.

Beweis. Es gilt $(a + b) - (a + b) = 0$ und $(a + b) + (-a - b) = a + (b - a - b) = a + (-a + b - b) = a + (-a) + 0 = 0$. Aus Behauptung 1.11 folgt die Behauptung. ■

Völlig analog wie die vorstehenden Behauptungen beweist man die folgenden Aussagen über die Multiplikation anstelle der Addition:

1.16 Behauptung. Die Zahl 1 ist eindeutig bestimmt.

1.17 Behauptung. Zu $a \in \mathbb{R} \setminus \{0\}$ ist a^{-1} eindeutig bestimmt.

1.18 Behauptung. Sei $a \in \mathbb{R} \setminus \{0\}, b \in \mathbb{R}$. Die Gleichung $ax = b$ hat eine eindeutig bestimmte Lösung x . Die Lösung ist $a^{-1}b =: \frac{b}{a} =: b/a$.

1.19 Behauptung. Für alle $a \in \mathbb{R} \setminus \{0\}$ gilt $(a^{-1})^{-1} = a$.

1.20 Behauptung. Für alle $a, b \in \mathbb{R} \setminus \{0\}$ gilt $(ab)^{-1} = b^{-1}a^{-1} = a^{-1}b^{-1}$.

Erst die nachstehenden Folgerungen benutzen auch das Distributivgesetz.

1.21 Behauptung. Für alle $a, b, c \in \mathbb{R}$ gilt $(a + b)c = ac + bc$.

Beweis. Es ist $(a + b)c \stackrel{(1.6)}{=} c(a + b) \stackrel{(1.9)}{=} ca + cb \stackrel{(1.6)}{=} ac + bc$. ■

1.22 Behauptung. Für alle $a \in \mathbb{R}$ gilt $a \cdot 0 = 0$.

Beweis. Wir haben $a \cdot 0 + a \cdot 0 \stackrel{(1.9)}{=} a(0+0) = a \cdot 0$. Da auch $a \cdot 0 + 0 = a \cdot 0$ ist, folgt $a \cdot 0 = 0$ aus Behauptung 1.13. ■

1.23 Behauptung. Für alle $a, b \in \mathbb{R}$ gilt $ab = 0$ genau dann, wenn $a = 0$ oder $b = 0$ ist.

Beweis. (\Rightarrow) Sei $ab = 0$ und etwa $a \neq 0$. Dann folgt $0 = a^{-1}(ab) \stackrel{(1.5)}{=} (a^{-1}a)b = 1 \cdot b = b$, also $b = 0$.

(\Leftarrow) Sei $a = 0$ oder $b = 0$, etwa $b = 0$. Dann ist $ab = a \cdot 0 = 0$ nach Behauptung 1.22. ■

1.24 Behauptung. Für alle $a, b \in \mathbb{R}$ gilt $(-a)b = -(ab)$.

Beweis. Es ist $ab + (-a)b \stackrel{1.21}{=} (a + (-a))b = 0 \cdot b \stackrel{1.22}{=} 0$ und $ab + (-ab) = 0$. Aus der Eindeutigkeit des Inversen folgt die Behauptung. ■

Speziell ist also $(-1)b = -b$ für alle $b \in \mathbb{R}$.

1.25 Behauptung. Für alle $a, b \in \mathbb{R}$ gilt $(-a)(-b) = ab$.

Beweis. Es ist $(-a)(-b) \stackrel{1.24}{=} -(a(-b)) \stackrel{1.24}{=} -(-ab) \stackrel{1.14}{=} ab$. ■

Bemerkung. Ein Tripel $(M, +, \cdot)$, bestehend aus einer nichtleeren Menge M und zwei Verknüpfungen $+$ und \cdot auf M , für die die Axiome (1.1)–(1.9) (mit M anstelle von \mathbb{R}) gelten, wird als (*kommutativer*) *Körper* bezeichnet.

Wie man leicht nachweist, gibt es einen Körper $(M, +, \cdot)$ mit $M = \{0, 1\}$ und der Eigenschaft $1 + 1 = 0$. Aus den Körperaxiomen läßt sich also noch nicht herleiten, dass $1 + 1 \neq 0$ ist.

1.2 Die Anordnungsaxiome

Zusätzlich nehmen wir nun an, dass auf der Menge \mathbb{R} eine Beziehung $<$ (gelesen „kleiner als“) gegeben ist, die zwischen zwei Zahlen bestehen kann. Für je zwei Zahlen $a, b \in \mathbb{R}$ ist also $a < b$ eine Aussage, die wahr oder falsch sein kann. Wir fordern die folgenden Regeln, die sogenannten „Anordnungsaxiome“.

(2.1) $\forall a, b \in \mathbb{R} : a \neq b \Leftrightarrow (a < b \vee b < a)$
„Vergleichbarkeit“

(2.2) $\forall a, b, c \in \mathbb{R} : (a < b \wedge b < c) \Rightarrow a < c$
„Transitivität“

(2.3) $\forall a, b, c \in \mathbb{R} : a < b \Rightarrow a + c < b + c$
 „Monotoniegesetz der Addition“

(2.4) $\forall a, b, c \in \mathbb{R} : (a < b \wedge 0 < c) \Rightarrow ac < bc$
 „Monotoniegesetz der Multiplikation“

Bezeichnungen.

$a > b \Leftrightarrow b < a$ ($>$ wird gelesen: „größer als“)

$a \leq b \Leftrightarrow a < b$ oder $a = b$

$a \geq b \Leftrightarrow a > b$ oder $a = b$

Statt $a < b$ und $b < c$ schreibt man kurz $a < b < c$. Hierfür sagt man auch, „ b liegt zwischen a und c “.

Definition. $a \in \mathbb{R}$ heißt *positiv* (*negativ*), wenn $a > 0$ (bzw. $a < 0$) ist.

Wir leiten nun aus den Anordnungsaxiomen (und den Körperaxiomen) einige Folgerungen her.

2.5 Behauptung. $\forall a, b \in \mathbb{R} : a < b \Rightarrow -b < -a$.

Beweis. Gelte $a < b$. Aus (2.3) (mit $c = -a - b$) folgt $a - a - b < b - a - b$, also $-b < -a$. ■

2.6 Behauptung. $\forall a, b, a', b' \in \mathbb{R} : (a < b \wedge a' < b') \Rightarrow a + a' < b + b'$.

Beweis. Aus $a < b$ und $a' < b'$ folgt nach (2.3) $a + a' < b + a' = a' + b < b' + b = b + b'$, nach (2.2) ist also $a + a' < b + b'$. ■

2.7 Behauptung. $\forall a, b, a', b' \in \mathbb{R} : (0 \leq a < b \wedge 0 \leq a' < b') \Rightarrow aa' < bb'$.

Beweis. Ist $a' = 0$, so ist $aa' = 0$ und $0 < bb'$ nach (2.4), also $aa' < bb'$. Sei jetzt $0 < a'$. Nach (2.4) ist $aa' < ba' < bb'$, nach (2.2) also $aa' < bb'$. ■

2.8 Behauptung. $\forall a, b, c \in \mathbb{R} : (a < b \wedge c < 0) \Rightarrow ac > bc$.

Beweis. Aus $c < 0$ folgt nach Behauptung 2.5 $0 < -c$ und damit nach (2.4) $a(-c) < b(-c)$, also $-(ac) < -(bc)$ und somit nach Behauptung 2.5 $ac > bc$. ■

2.9 Behauptung. $\forall a \in \mathbb{R} \setminus \{0\} : aa > 0$.

Beweis. Ist $a > 0$, so folgt die Behauptung aus (2.4). Andernfalls ist nach (2.1) $a < 0$, also $-a > 0$ und daher nach Behauptung 1.25 und (2.4) $aa = (-a)(-a) > 0$. ■

Insbesondere gilt $1 > 0$, da wegen (1.7) gilt $1 \neq 0$ und wegen 2.9 gilt $1 = 1 \cdot 1 > 0$. Aus $1 > 0$ und Behauptung 2.6 folgt $1 + 1 > 0 + 0 = 0$; wegen $1 + 1 > 0$ ist $1 + 1 \neq 0$, was allein aus den Körperaxiomen nicht gefolgert werden konnte.

2.10 Behauptung. $\forall a \in \mathbb{R} : a > 0 \Rightarrow a^{-1} > 0$.

Beweis. Sei $a > 0$. Es gilt $a^{-1} = (aa^{-1})a^{-1} = a(a^{-1}a^{-1}) > 0$ nach Behauptung 2.9 und (2.4) (denn wegen $aa^{-1} = 1$ und Behauptung 1.23 ist $a^{-1} \neq 0$). ■

Analog ergibt sich: $\forall a \in \mathbb{R} : a < 0 \Rightarrow a^{-1} < 0$.

2.11 Behauptung. $\forall a, b \in \mathbb{R} : 0 < a < b \Rightarrow a^{-1} > b^{-1}$.

Beweis. Aus $a > 0$ und $b > 0$ folgt $ab > 0$ und daraus $a^{-1}b^{-1} = (ab)^{-1} > 0$ nach Behauptung 2.10. Hieraus und aus $a < b$ folgt nach (2.4) $aa^{-1}b^{-1} < ba^{-1}b^{-1}$, also $b^{-1} < a^{-1}$. ■

Bezeichnungen. $\mathbb{R}^+ := \{x \in \mathbb{R} \mid x > 0\}$

Für $a, b \in \mathbb{R}$ mit $a \leq b$ wird definiert:

$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$	abgeschlossenes Intervall
$(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$	offenes Intervall
$[a, b) := \{x \in \mathbb{R} \mid a \leq x < b\}$	halboffenes Intervall
$(a, b] := \{x \in \mathbb{R} \mid a < x \leq b\}$	halboffenes Intervall

In jedem Fall heißt $b - a$ auch die *Länge* des Intervalls. Ferner wird definiert:

$$\begin{aligned} [a, \infty) &:= \{x \in \mathbb{R} \mid x \geq a\} \\ (a, \infty) &:= \{x \in \mathbb{R} \mid x > a\} \\ (-\infty, a] &:= \{x \in \mathbb{R} \mid x \leq a\} \\ (-\infty, a) &:= \{x \in \mathbb{R} \mid x < a\}. \end{aligned}$$

Das Symbol ∞ ist dabei nur zur Vereinfachung der Schreibweise gewählt; es ist *kein* Element von \mathbb{R} . Alle vorstehend definierten Mengen, dazu \emptyset und \mathbb{R} , heißen *Intervalle*.

Definition. Für $a \in \mathbb{R}$ sei

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

Die Zahl $|a|$ heißt der *Absolutbetrag* oder kurz der *Betrag* der Zahl a .

Folgerung (Eigenschaften des Absolutbetrages). Unmittelbar aus der Definition folgt für alle $a \in \mathbb{R}$:

$$\begin{aligned} |a| &\geq 0, \\ |a| &= 0 \Leftrightarrow a = 0, \\ -|a| &\leq a \leq |a|. \end{aligned}$$

2.12 Behauptung. $\forall a \in \mathbb{R} : |-a| = |a|$

Beweis. Fallunterscheidung:

1. Fall: $a \geq 0$. Dann ist $|a| = a$ und $|-a| = -(-a) = a$.
2. Fall: $a < 0$. Dann ist $|a| = -a$ und $|-a| = -a$. ■

2.13 Behauptung. Für alle $a, b \in \mathbb{R}$ gilt $|ab| = |a||b|$.

Beweis. Fallunterscheidung:

1. Fall: $a \geq 0$ und $b \geq 0$. Dann ist $ab \geq 0$, also $|ab| = ab$, und es ist $|a| = a$, $|b| = b$, also $ab = |a||b|$.
2. Fall: $a \geq 0$ und $b < 0$. Dann ist $ab \leq 0$, also $|ab| = -ab$, und es ist $|a| = a$, $|b| = -b$, also $-ab = |a||b|$.
3. Fall: $a < 0$ und $b \geq 0$. Analog.
4. Fall: $a < 0$ und $b < 0$. Analog. ■

2.14 Behauptung. $\forall a \in \mathbb{R} \setminus \{0\} : \left| \frac{1}{a} \right| = \frac{1}{|a|}$.

Beweis. Aus $1 = |1| = \left| a \cdot \frac{1}{a} \right| = |a| \cdot \left| \frac{1}{a} \right|$ folgt die Behauptung. ■

2.15 Behauptung (Dreiecksungleichung). $\forall a, b \in \mathbb{R} : |a + b| \leq |a| + |b|$.

Beweis. Wegen $a \leq |a|$ und $b \leq |b|$ ist nach Behauptung 2.6 $a + b \leq |a| + |b|$. Wegen $-a \leq |a|$ und $-b \leq |b|$ ist ebenso $-(a + b) \leq |a| + |b|$. Nach Definition des Absolutbetrages folgt $|a + b| \leq |a| + |b|$. ■

2.16 Behauptung. $\forall a, b \in \mathbb{R} : \left| |a| - |b| \right| \leq |a + b|$.

Beweis. Es ist $|a| = |(a + b) - b| \leq |a + b| + |b|$, also $|a| - |b| \leq |a + b|$. Vertauschung von a und b ergibt $-(|a| - |b|) = |b| - |a| \leq |a + b|$, insgesamt folgt die Behauptung. ■

Bemerkung. Ein Quadrupel $(M, +, \cdot, <)$, wobei M eine nichtleere Menge ist, $+$ und \cdot zwei Verknüpfungen und $<$ eine Beziehung auf M sind derart, dass die Axiome (1.1)–(1.9) und (2.1)–(2.4) (mit M statt \mathbb{R}) erfüllt sind, heißt *angeordneter Körper*.

Wir bemerken ohne Beweis, dass bis jetzt zum Beispiel die folgenden Aussagen noch nicht beweisbar sind:

- (a) Es gibt ein $x \in \mathbb{R}$ mit $x^2 = 2$ (nach Definition ist $x^2 := x \cdot x$ und $2 := 1+1$).
- (b) Für gegebenes $a \in \mathbb{R}$ ist schließlich $1 + 1 + \cdots + 1 > a$, wenn genügend oft 1 addiert wird.

1.3 Das Vollständigkeitsaxiom

Wie die Anmerkung am Ende des vorigen Paragraphen zeigt, reichen die Körper- und Anordnungsaxiome noch nicht aus, um alle Eigenschaften zu erfassen, die wir von den reellen Zahlen erwarten. Zur Formulierung des letzten Axioms dient die folgende Definition:

Definition. Sei $M \subseteq \mathbb{R}$. Eine Zahl $s \in \mathbb{R}$ heißt *obere Schranke* für M , wenn

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

gilt. M heißt *nach oben beschränkt*, wenn es in \mathbb{R} eine obere Schranke für M gibt.

Eine Zahl $s_0 \in \mathbb{R}$ heißt *kleinste obere Schranke* oder *Supremum* der Menge M , wenn s_0 obere Schranke für M ist und jede Zahl $s < s_0$ nicht obere Schranke für M ist.

Enthält die Menge M ein größtes Element m_0 (also ein $m_0 \in M$ mit $x \leq m_0$ für alle $x \in M$), so heißt m_0 *Maximum* der Menge M .

Ein Maximum einer Menge $M \subseteq \mathbb{R}$ ist also stets auch Supremum von M , aber i.a. nicht umgekehrt, denn ein Supremum von M braucht nicht Element von M zu sein.

Völlig analog definiert man die Begriffe untere Schranke, nach unten beschränkt, größte untere Schranke (Infimum), Minimum.

Die Menge $M \subseteq \mathbb{R}$ heißt *beschränkt*, wenn sie nach oben und nach unten beschränkt ist.

Beispiele. Für die Intervalle $[a, b]$, (a, b) ist jede Zahl s mit $s \geq b$ eine obere Schranke, und b ist kleinste obere Schranke. Für $[a, b]$ ist b das Maximum und a das Minimum, aber (a, b) besitzt kein Maximum.

Bezeichnungen.

$\sup M :=$ Supremum von M

$\inf M :=$ Infimum von M

$\max M :=$ Maximum von M

$\min M :=$ Minimum von M

Nun formulieren wir das Vollständigkeitsaxiom:

- (3.1) Zu jeder nichtleeren, nach oben beschränkten Teilmenge von \mathbb{R} gibt es in \mathbb{R} eine kleinste obere Schranke.

Als unmittelbare Folgerung haben wir:

3.2 Behauptung. *Zu jeder nichtleeren, nach unten beschränkten Teilmenge von \mathbb{R} gibt es in \mathbb{R} eine größte untere Schranke.*

Beweis. Sei $\emptyset \neq M \subseteq \mathbb{R}$. Setze $-M := \{x \in \mathbb{R} \mid -x \in M\}$. Für $s \in \mathbb{R}$ zeigt man leicht:

s untere Schranke für $M \Leftrightarrow -s$ obere Schranke für $-M$

s größte untere Schranke für $M \Leftrightarrow -s$ kleinste obere Schranke für $-M$.

Die Behauptung ergibt sich jetzt durch Anwendung des Vollständigkeitsaxioms auf $-M$. ■

Die Bedeutung des Axioms liegt darin, dass es die Existenz von Zahlen mit bestimmten Eigenschaften sichert. Zum Beispiel folgt die Existenz der Zahl $\sqrt{2}$ aus dem Vollständigkeitsaxiom.

Das Quadrupel $(\mathbb{R}, +, \cdot, <)$ mit (1.1)–(1.9), (2.1)–(2.4), (3.1) heißt *Körper der reellen Zahlen*. Eine solche Struktur ist „im wesentlichen“ eindeutig bestimmt, was wir hier aber nicht näher erläutern wollen. Die Elemente einer solchen Struktur nennen wir *reelle Zahlen*. Wenn wir im folgenden nur von „Zahlen“ sprechen, sind stets reelle Zahlen gemeint.

1.4 Natürliche Zahlen und vollständige Induktion

Im Körper \mathbb{R} der reellen Zahlen sind insbesondere die Zahlen

$$\begin{aligned}
&1 \\
&2 := 1 + 1 \\
&3 := 2 + 1 \\
&4 := 3 + 1 \\
&\text{u.s.w.}
\end{aligned}$$

enthalten. Sie heißen *natürliche Zahlen*. Das obenstehende „u.s.w.“ kann aber kaum als Definition dienen. Wir müssen also eine Definition der Menge der natürlichen Zahlen geben, und wir können das jetzt auch tun. Zur Menge der natürlichen Zahlen sollen selbstverständlich genau die Zahlen gehören, die man in der obigen Weise erhalten kann, daher die folgende

Definition. \mathbb{N} sei die kleinste Teilmenge von \mathbb{R} , die die Zahl 1 und mit x auch $x + 1$ enthält. Die Elemente von \mathbb{N} heißen *natürliche Zahlen*.

Die „kleinste“ Teilmenge von \mathbb{R} mit einer Eigenschaft E ist dabei zu verstehen als Durchschnitt aller Teilmengen von \mathbb{R} mit der Eigenschaft E . Der Durchschnitt einer Familie von Mengen ist völlig analog erklärt wie früher der Durchschnitt von zwei Mengen. Der Deutlichkeit halber wollen wir vorübergehend (d.h. nur in diesem Paragraphen) eine Teilmenge $M \subseteq \mathbb{R}$ *induktiv* nennen, wenn $1 \in M$ ist und mit $x \in M$ auch $x + 1 \in M$ gilt. Zum Beispiel sind \mathbb{R} und \mathbb{R}^+ induktive Teilmengen. Das System (synonym für „die Menge“) aller induktiven Teilmengen von \mathbb{R} sei mit I bezeichnet. Dann haben wir also nach Definition

$$\begin{aligned}
\mathbb{N} &:= \bigcap_{M \in I} M := \{a \in \mathbb{R} \mid a \in M \text{ für alle } M \in I\} \\
&= \{a \in \mathbb{R} \mid a \text{ ist Element jeder induktiven Teilmenge von } \mathbb{R}\}.
\end{aligned}$$

Definition.

$$\begin{aligned}
\mathbb{N}_0 &:= \mathbb{N} \cup \{0\} := \{0, 1, 2, 3, \dots\}, \\
\mathbb{Z} &:= \{a \in \mathbb{R} \mid a \in \mathbb{N}_0 \text{ oder } -a \in \mathbb{N}_0\}.
\end{aligned}$$

Die Elemente von \mathbb{Z} heißen *ganze Zahlen*.

$$\mathbb{Q} := \left\{ a \in \mathbb{R} \mid \exists m \in \mathbb{Z} \exists n \in \mathbb{N} : a = \frac{m}{n} \right\}.$$

Die Elemente von \mathbb{Q} heißen *rationale Zahlen* (oder *Brüche*). Die Elemente von $\mathbb{R} \setminus \mathbb{Q}$ heißen *irrationale Zahlen*.

Bemerkung. \mathbb{Q} mit den Verknüpfungen $+$ und \cdot und der Beziehung $<$ ist ebenfalls ein angeordneter Körper. In diesem Körper ist die Gleichung $x^2 = 2$ unlösbar.

Bemerkung. dass es überhaupt irrationale Zahlen gibt, wird erst später bewiesen.

Wir zeigen nun, wie sich einige an sich wohlvertraute Eigenschaften natürlicher Zahlen aus dieser Definition streng herleiten lassen.

4.1 Behauptung. *Die Menge \mathbb{N} der natürlichen Zahlen ist nicht nach oben beschränkt.*

Beweis. [Beweis durch Widerspruch: Wir zeigen, dass aus der Negation (dem logischen Gegenteil) der Behauptung ein Widerspruch folgt.] Angenommen, die Behauptung wäre falsch. Nach dem Vollständigkeitsaxiom gibt es dann für \mathbb{N} eine kleinste obere Schranke s . Da $s - 1$ keine obere Schranke für \mathbb{N} ist, existiert ein $n \in \mathbb{N}$ mit $s - 1 < n$, woraus $s < n + 1$ folgt. Da \mathbb{N} induktiv ist, ist mit $n \in \mathbb{N}$ auch $n + 1 \in \mathbb{N}$, also ist s nicht obere Schranke für \mathbb{N} , ein Widerspruch. ■

4.2 Behauptung. *Zu jedem $a \in \mathbb{R}^+$ und jedem $b \in \mathbb{R}$ existiert ein $n \in \mathbb{N}$ mit $na > b$.*

Beweis. Angenommen, das wäre falsch. Dann gibt es ein $a \in \mathbb{R}^+$ und ein $b \in \mathbb{R}$ mit $na \leq b$ für alle $n \in \mathbb{N}$. Es ist also $n \leq \frac{b}{a}$ für alle $n \in \mathbb{N}$, im Widerspruch zu Behauptung 4.1. ■

Behauptung 4.2 drückt eine besondere Eigenschaft der durch $<$ gegebenen Anordnung aus, die man auch die *archimedische* Eigenschaft nennt.

Die folgenden Konsequenzen werden später bei der Behandlung des Grenzwertbegriffes ständig benutzt.

4.3 Behauptung. $\forall a \in \mathbb{R}^+ \exists n \in \mathbb{N} : \frac{1}{n} < a$.

Beweis. Nach Behauptung 4.2 existiert ein $n \in \mathbb{N}$ mit $na > 1$, also mit $\frac{1}{n} < a$. ■

4.4 Behauptung. *Sei $a \in \mathbb{R}$, $a \geq 0$. Ist $a \leq \frac{1}{n}$ für alle $n \in \mathbb{N}$, so ist $a = 0$.*

Beweis. Andernfalls wäre $\frac{1}{a}$ eine obere Schranke für \mathbb{N} . ■

4.5 Behauptung. $\forall n \in \mathbb{N} : (n > 1 \Rightarrow n - 1 \in \mathbb{N})$.

Beweis. Die Menge $\{n \in \mathbb{N} \mid n = 1 \vee n - 1 \in \mathbb{N}\}$ ist induktiv, also gleich \mathbb{N} . ■

Hier wurde benutzt: Ist $M \subseteq \mathbb{N}$ eine induktive Teilmenge, so muß nach Definition von \mathbb{N} auch $\mathbb{N} \subseteq M$ gelten, also ist $M = \mathbb{N}$. Hierauf stützt sich auch das häufig verwendete

Beweisprinzip der vollständigen Induktion

Dieses läßt sich folgendermaßen beschreiben. Es bezeichne A eine Aussage über natürliche Zahlen mit folgenden Eigenschaften:

- (a) A gilt für die Zahl 1,
- (b) Wenn A für die natürliche Zahl n gilt, dann auch für $n + 1$.

Dann gilt A für alle $n \in \mathbb{N}$.

Um dies einzusehen, betrachtet man die Menge

$$M := \{n \in \mathbb{N} \mid A \text{ gilt für } n\}.$$

Sie ist induktiv, muß also nach der obigen Bemerkung gleich \mathbb{N} sein.

Um also zu beweisen, dass eine Behauptung für alle natürlichen Zahlen zutrifft, verfährt man folgendermaßen:

- (a) „Induktionsanfang“: Man beweist, dass die Behauptung für die Zahl 1 zutrifft.
- (b) „Induktionsschluß“: Unter der „Induktionsannahme“, die Behauptung treffe für ein $n \in \mathbb{N}$ zu, zeigt man, dass die Behauptung auch für $n + 1$ zutrifft.

Hierfür nun einige Beispiele.

4.6 Behauptung. $\forall n \in \mathbb{N} \forall x \in \mathbb{R} : (n < x < n + 1 \Rightarrow x \notin \mathbb{N})$.

Beweis. durch vollständige Induktion:

Induktionsanfang: Für $n = 1$ lautet die Behauptung:

$$\forall x \in \mathbb{R} : (1 < x < 2 \Rightarrow x \notin \mathbb{N})$$

Sie ist richtig, weil die Menge $M := \{x \in \mathbb{R} \mid x = 1 \vee x \geq 2\}$ induktiv ist und daher $\mathbb{N} \subseteq M$ gilt. Für alle $n \in \mathbb{N}$ ist also $n = 1$ oder $n \geq 2$.

Induktionsschluß: Wir machen die *Induktionsannahme*, die Behauptung gelte für die natürliche Zahl n , also

$$\forall x \in \mathbb{R} : (n < x < n + 1 \Rightarrow x \notin \mathbb{N}).$$

Sei jetzt $x \in \mathbb{R}$ eine Zahl mit $n + 1 < x < n + 2$. Dann folgt $n < x - 1 < n + 1$, nach Induktionsannahme also $x - 1 \notin \mathbb{N}$. Wegen $x > 1$ folgt aus Behauptung 4.5 jetzt $x \notin \mathbb{N}$. Die Behauptung ist also richtig für $n + 1$. Damit ist gezeigt, dass sie für alle $n \in \mathbb{N}$ gilt. ■

Hieraus können wir den folgenden, häufig nützlichen Satz herleiten:

4.7 Satz (Wohlordnung der natürlichen Zahlen). *In jeder nichtleeren Menge von natürlichen Zahlen gibt es ein kleinstes Element.*

Beweis. Sei $\emptyset \neq A \subseteq \mathbb{N}$. Setze $M := \{n \in \mathbb{N} \mid \forall a \in A : n \leq a\}$.

Behauptung. $\exists k \in M : k + 1 \notin M$.

Beweis. Wäre das falsch, so wäre M induktiv, also $M = \mathbb{N}$. Wegen $A \neq \emptyset$ existiert ein $a \in A$, wegen $a + 1 \in \mathbb{N}$ folgt $a + 1 \leq a$, ein Widerspruch. ■

Behauptung. k ist kleinstes Element von A .

Beweis. Wegen $k \in M$ ist $k \leq a$ für alle $a \in A$. Es bleibt $k \in A$ zu zeigen. Wegen $k + 1 \notin M$ existiert ein $a_0 \in A$ mit $k + 1 > a_0$. Aus $k \leq a_0 < k + 1$ folgt nach Behauptung 4.6 aber $k = a_0$. ■

Damit ist Satz 4.7 bewiesen. ■

Als wichtiges Anwendungsbeispiel haben wir den folgenden Satz:

4.8 Satz. *Zu beliebigen reellen Zahlen $a, b \in \mathbb{R}$ mit $a < b$ existiert eine rationale Zahl $r \in \mathbb{Q}$ mit $a < r < b$.*

Beweis. O.B.d.A. sei $a \geq 0$ (andernfalls ersetze man a, b durch $a + m, b + m$ mit genügend großem $m \in \mathbb{N}$). Nach Behauptung 4.3 gibt es eine Zahl $n \in \mathbb{N}$ mit $\frac{1}{n} < b - a$. Wähle ein solches n und setze $M := \{m \in \mathbb{N} \mid \frac{m}{n} > a\}$. Wegen Behauptung 4.2 ist $M \neq \emptyset$. Nach Satz 4.7 enthält M ein kleinstes Element k . Es ist also $a < \frac{k}{n}$. Wäre $\frac{k}{n} \geq b$, so wäre

$$\frac{k-1}{n} \geq b - \frac{1}{n} > b + a - b = a,$$

also $k - 1 \in M$, entgegen der Definition von k . Somit ist $a < \frac{k}{n} < b$. ■

Für die Aussage von Satz 4.8 sagt man auch: „ \mathbb{Q} liegt dicht in \mathbb{R} “.

Neben Beweisen durch Induktion sind auch Definitionen durch Induktion möglich. Wir erläutern dies an einem Beispiel:

Definition. Sei $a \in \mathbb{R}$. Setze $a^0 := 1$ und $a^1 := a$. Ist a^n für ein $n \in \mathbb{N}$ schon definiert, so setze $a^{n+1} := a \cdot a^n$. Damit ist a^n für alle $n \in \mathbb{N}$ definiert.

Man nennt ein solches Vorgehen „Definition durch vollständige Induktion“ oder „rekursive Definition“. Damit ist also folgendes gemeint. Sei $k \in \mathbb{N}$. Um für jedes $n \in \mathbb{N}$ mit $n \geq k$ eine Zahl a_n zu definieren, definiert man zuerst a_k . Dann gibt man an, wie a_n zu definieren ist, wenn $n > k$ ist und die Zahlen a_k, \dots, a_{n-1} schon definiert sind. dass damit a_n in der Tat für alle $n \in \mathbb{N}$ mit $n \geq k$ definiert ist, ist plausibel. Die Möglichkeit einer rekursiven Definition läßt sich durch Ausnutzung der Eigenschaften von \mathbb{N} streng begründen, was hier aber nicht geschehen soll. (Statt Zahlen können die a_n auch Elemente irgendwelcher Mengen sein.)

Über die so definierten *Potenzen* beweisen wir nun einige einfache Aussagen.

4.9 Behauptung. $\forall a \in \mathbb{R} \quad \forall m, n \in \mathbb{N} : a^m a^n = a^{m+n}$.

Beweis. Die Behauptung ist richtig für $n = 1$ (nach Definition). Sei sie bewiesen für ein $n \geq 1$. Für $m \in \mathbb{N}$ gilt dann

$$a^m a^{n+1} \stackrel{\text{Def.}}{=} a^m a^n a \stackrel{\text{Ind. - Ann.}}{=} a^{m+n} a \stackrel{\text{Def.}}{=} a^{m+n+1}. \quad \blacksquare$$

4.10 Behauptung (Bernoullische Ungleichung).

$$\forall a \in \mathbb{R} \quad \forall n \in \mathbb{N}_0 : (a > -1 \Rightarrow (1+a)^n \geq 1+na).$$

Beweis. Für $n = 0$ und $n = 1$ ist die Behauptung richtig. Sei sie bewiesen für ein $n \geq 1$. Dann folgt $(1+a)^{n+1} = (1+a)(1+a)^n \geq (1+a)(1+na) = 1 + (n+1)a + na^2 \geq 1 + (n+1)a$. \blacksquare

4.11 Behauptung. Sei $b \in \mathbb{R}$, $a \in \mathbb{R}^+$.

$$(1) \quad b > 1 \Rightarrow \exists n \in \mathbb{N} : b^n > a,$$

$$(2) \quad 0 < b < 1 \Rightarrow \exists n \in \mathbb{N} : b^n < a.$$

Beweis. (1) Nach Behauptung 4.2 existiert ein $n \in \mathbb{N}$ mit $n(b-1) > a-1$. Nach Behauptung 4.10 folgt

$$b^n = (1 + (b - 1))^n \geq 1 + n(b - 1) > 1 + a - 1 = a.$$

(2) Aus $0 < b < 1$ folgt $\frac{1}{b} > 1$, nach (1) existiert also ein $n \in \mathbb{N}$ mit $\left(\frac{1}{b}\right)^n > \frac{1}{a}$. Also gilt $\frac{1}{b^n} > \frac{1}{a}$ (denn $\left(\frac{1}{b}\right)^n = \frac{1}{b^n}$, wie sofort durch Induktion folgt). Behauptung 2.11 liefert also $b^n < a$. ■

Es folgen nun noch einige Aussagen über

Summen und Produkte

4.12 Behauptung. $\forall n \in \mathbb{N} \forall m \in \mathbb{N} : n + m \in \mathbb{N}$.

Beweis. Wir beweisen dies durch Induktion nach n :

Induktionsanfang: Zu zeigen ist die Behauptung

$$\forall m \in \mathbb{N} : 1 + m \in \mathbb{N}.$$

Sie ist richtig nach Definition von \mathbb{N} .

Induktionsschluß: Wir machen die Annahme

$$\forall m \in \mathbb{N} : n + m \in \mathbb{N}.$$

Für beliebiges $m \in \mathbb{N}$ gilt dann $(n + 1) + m = n + (m + 1) \in \mathbb{N}$ nach Induktionsannahme. Die Behauptung gilt also auch für $n + 1$. ■

4.13 Behauptung. $\forall n \in \mathbb{N} \forall m \in \mathbb{N} : nm \in \mathbb{N}$

Beweis. Analog, unter Verwendung von Behauptung 4.12. ■

4.14 Behauptung. $\forall n \in \mathbb{N} \forall m \in \mathbb{N} : (m > n \Rightarrow m - n \in \mathbb{N})$.

Beweis. Ähnlich; zum Induktionsanfang siehe Behauptung 4.5. ■

Durch rekursive Definition können wir auch Summen und Produkte von n reellen Zahlen erklären. Für $a_1, a_2 \in \mathbb{R}$ sind $a_1 + a_2$ und $a_1 \cdot a_2$ bereits erklärt. Nehmen wir an, für gegebenes $n \in \mathbb{N}$ und beliebige Zahlen $a_1, \dots, a_n \in \mathbb{R}$ seien die Summe $a_1 + \dots + a_n$ und das Produkt $a_1 \cdot \dots \cdot a_n$ bereits definiert. Dann definieren wir für $a_1, \dots, a_{n+1} \in \mathbb{R}$

$$\begin{aligned} a_1 + \cdots + a_{n+1} &:= (a_1 + \cdots + a_n) + a_{n+1}, \\ a_1 \cdots a_{n+1} &:= (a_1 \cdots a_n) \cdot a_{n+1}. \end{aligned}$$

Man kann nun leicht aus dem Assoziativgesetz durch vollständige Induktion herleiten, dass

$$(a_1 + \cdots + a_p) + (a_{p+1} + \cdots + a_n) = a_1 + \cdots + a_n$$

ist. Daraus kann man dann, ebenfalls durch Induktion, herleiten, dass man bei einem Ausdruck aus Summenzeichen und Klammern überhaupt die Klammern weglassen kann (allgemeines Assoziativgesetz). Aus dem Kommutativgesetz leitet man durch Induktion leicht her, dass es bei einer Summe $a_1 + \cdots + a_n$ nicht auf die Reihenfolge der Summanden ankommt (allgemeines Kommutativgesetz). Analoges gilt für die Multiplikation. Aus dem Distributivgesetz leitet man durch Induktion das allgemeine Distributivgesetz her:

$$a(a_1 + \cdots + a_n) = aa_1 + \cdots + aa_n.$$

Beim Arbeiten mit derartigen Ausdrücken ist die Verwendung von Summen- und Produktzeichen praktisch:

Bezeichnungen.

$$\begin{aligned} \sum_{i=k}^n a_i &:= a_k + a_{k+1} + \cdots + a_n, & \text{falls } n \geq k \\ \prod_{i=k}^n a_i &:= a_k a_{k+1} \cdots a_n, & \text{falls } n \geq k. \end{aligned}$$

Aus Assoziativ-, Kommutativ- und Distributivgesetz folgen einige Rechenregeln für Summen, z.B.

$$\begin{aligned} \sum_{i=k}^l \sum_{j=m}^n a_{ij} &= \sum_{j=m}^n \sum_{i=k}^l a_{ij}, \\ \left(\sum_{i=k}^l a_i \right) \left(\sum_{j=m}^n b_j \right) &= \sum_{i=k}^l \sum_{j=m}^n a_i b_j. \end{aligned}$$

Zum Abschluß dieses Paragraphen wollen wir noch zwei häufig auftretende Summen bzw. Produkte berechnen:

4.15 Behauptung (Summe einer endlichen geometrischen Reihe). *Für $q \in \mathbb{R} \setminus \{0, 1\}$ und $n \in \mathbb{N}$ gilt*

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

Beweis. Setze $\sum_{k=0}^n q^k =: x$. Dann ist

$$qx = q + q^2 + \cdots + q^{n+1} = x - 1 + q^{n+1},$$

also

$$x = \frac{1 - q^{n+1}}{1 - q}. \quad \blacksquare$$

Zur nun folgenden Berechnung der Potenz $(a + b)^n$ sind zunächst einige Definitionen erforderlich.

Definition.

$$\begin{aligned} n! &:= 1 \cdot 2 \cdot \cdots \cdot n = \prod_{k=1}^n k && \text{für } n \in \mathbb{N} \\ 0! &:= 1 \\ \binom{n}{k} &:= \frac{n(n-1) \cdots (n-k+1)}{1 \cdot 2 \cdots k} && \text{für } n, k \in \mathbb{N} \\ \binom{n}{0} &:= 1 && \text{für } n \in \mathbb{N}_0 \end{aligned}$$

Man liest $n!$ als „ n Fakultät“ und $\binom{n}{k}$ als „ n über k “. Die Größen $\binom{n}{k}$ heißen *Binomialkoeffizienten*. Offenbar gilt

$$\begin{aligned} \binom{n}{k} &= \frac{n!}{k!(n-k)!} && \text{für } k, n \in \mathbb{N}_0, n \geq k \\ \binom{n}{k} &= 0 && \text{für } k > n \\ \binom{n}{k} &= \binom{n}{n-k} \end{aligned}$$

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

Beweis der letzten Gleichung:

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

Jetzt können wir zeigen:

4.16 Behauptung (Binomische Formel). Für $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$ gilt

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

Beweis. (Induktion): Für $n = 1$ gilt die Formel. Sei sie bewiesen für ein $n \geq 1$. Dann folgt

$$\begin{aligned} (a + b)^{n+1} &= (a + b)(a + b)^n \stackrel{\text{Ind. - Ann.}}{=} (a + b) \sum_{k=0}^n \binom{n}{k} a^k b^{n-k} \\ &= \sum_{k=0}^n \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^n \binom{n}{k} a^k b^{n-k+1} \\ &= a^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=1}^n \binom{n}{k} a^k b^{n-k+1} + b^{n+1} \\ &= a^{n+1} + \sum_{j=1}^n \left[\binom{n}{j-1} + \binom{n}{j} \right] a^j b^{n+1-j} + b^{n+1} \\ &= \sum_{j=0}^{n+1} \binom{n+1}{j} a^j b^{n+1-j}. \quad \blacksquare \end{aligned}$$

2 Abbildungen

2.1 Der Funktionsbegriff

In diesem Paragraphen geht es hauptsächlich um die Vereinbarung einiger Sprechweisen und die Festlegung von Bezeichnungen. Es kommen also vorwiegend Definitionen vor und fast keine Sätze.

Der Begriff der Funktion, der im folgenden präzisiert werden soll, wird auch im täglichen Leben verwendet. Was meinen wir etwa, wenn wir sagen, beim Autofahren sei der Luftwiderstand eine Funktion der Geschwindigkeit? Gemeint ist doch, dass zu jedem Geschwindigkeitswert ein wohlbestimmter Wert des Luftwiderstandes gehört. Es ist aber nicht unbedingt gemeint, dass wir diesen durch eine explizite Formel oder Berechnungsvorschrift angeben können. Historisch mag es wohl so gewesen sein, dass man unter einer Funktion zunächst eine konkrete Berechnungsvorschrift verstanden hat, aber dies hat sich als zu eng erwiesen. Wesentlich an einer Funktion, wie man sie heute versteht, ist lediglich, dass durch jeden „Argumentwert“ der zugehörige „Funktionswert“ eindeutig bestimmt ist. Unter einer Funktion wird man also eine eindeutige Zuordnung verstehen. Zum Beispiel beschreibt auch jedes Telefonbuch eine Funktion: Jedem Inhaber eines Telefonanschlusses innerhalb eines gewissen Gebietes wird darin seine Telefonnummer zugeordnet. Wenn man dieses Beispiel im Sinn behält, wird man nicht in die Gefahr geraten, eine Funktion für eine explizite Berechnungsvorschrift zu halten.

Wir wollen eine formale Definition einer Funktion geben. Dazu müssen also eine Menge M_1 , die „Definitionsmenge“, und eine Menge M_2 , die „Zielmenge“ oder „Wertemenge“ der zu definierenden Funktion gegeben sein. Jedem $x \in M_1$ soll ein eindeutig bestimmter Funktionswert $y \in M_2$ zugeordnet sein. Die Zuordnung ist vollständig festgelegt, wenn wir zu jedem $x \in M_1$ und jedem $y \in M_2$ wissen, ob y dem Element x zugeordnet ist oder nicht. Wir können daher die Zuordnungsvorschrift geradezu identifizieren mit der Angabe aller geordneten Paare (x, y) , für die y dem Element x zugeordnet ist.

Hierzu zunächst eine Vorbemerkung. Sei $n \geq 2$ eine natürliche Zahl. Zu n Objekten x_1, \dots, x_n (die durch die Numerierung von 1 bis n mit einer Reihenfolge versehen sind) bilden wir ein neues Objekt (x_1, \dots, x_n) , das wir das (geordnete) n -Tupel von x_1, \dots, x_n nennen (2-Tupel heißen auch *Paare*, 3-Tupel *Tripel*). Wir wollen hierfür keine formale Definition geben (obwohl dies unter alleiniger Verwendung des Mengenbegriffs möglich wäre), sondern lediglich eine Gleichheitsdefinition festlegen:

$$\begin{aligned} (x_1, \dots, x_n) &= (y_1, \dots, y_n) \\ &:\Leftrightarrow x_1 = y_1 \quad \text{und} \quad x_2 = y_2 \quad \text{und} \quad \dots \quad \text{und} \quad x_n = y_n. \end{aligned}$$

Bemerkung. Mit (a, b) haben wir sowohl das offene Intervall von a bis b als auch das geordnete Paar mit Einträgen a und b bezeichnet. Diese (an sich unkorrekte) Mehrdeutigkeit gibt aber nicht Anlaß zu Verwechslungen, da aus dem Zusammenhang stets klar sein wird, was gemeint ist. In Zweifelsfällen wird angegeben, ob das geordnete Paar oder das Intervall gemeint ist.

Sind nun M_1, \dots, M_n Mengen, so nennt man die Menge

$$M_1 \times M_2 \times \dots \times M_n := \{(x_1, \dots, x_n) \mid x_1 \in M_1 \wedge \dots \wedge x_n \in M_n\}$$

das (*kartesische*) *Produkt* der Mengen M_1, \dots, M_n (in dieser Reihenfolge). Man schreibt auch

$$\underbrace{M \times \dots \times M}_{n \text{ - mal}} =: M^n.$$

Jetzt können wir definieren (M_1, M_2 seien Mengen):

Definition. Eine *Funktion* oder *Abbildung* von M_1 in M_2 ist eine Teilmenge f der Produktmenge $M_1 \times M_2$ derart, dass zu jedem $x \in M_1$ genau ein $y \in M_2$ existiert mit $(x, y) \in f$. M_1 heißt *Definitionsbereich* und M_2 *Wertebereich* von f .

Bezeichnungen. Statt $(x, y) \in f$ schreibt man $y = f(x)$ und nennt $f(x)$ das *Bild von x unter f* , auch den *Wert von f an der Stelle x* .

Die Abkürzungen

$$f : M_1 \rightarrow M_2 \quad \text{oder} \quad M_1 \xrightarrow{f} M_2$$

bedeuten, dass f eine Funktion von M_1 in M_2 ist. In konkreten Fällen schreibt man auch

$$f : M_1 \rightarrow M_2 : x \mapsto f(x) \quad \text{oder} \quad f : x \mapsto f(x) \quad \text{für } x \in M_1.$$

Bemerkung. Am Anfang von Abschnitt 1.1 hatten wir gesagt: „Je zwei Elementen $a, b \in \mathbb{R}$ sei ein Element $a+b$, ihre Summe, eindeutig zugeordnet.“ Wir können das jetzt so präzisieren, dass eine Abbildung $+$: $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ gegeben sein soll. Ebenso ist die Multiplikation \cdot eine Abbildung von $\mathbb{R} \times \mathbb{R}$ in \mathbb{R} .

Bezeichnungen. Sei $f : M_1 \rightarrow M_2$ eine Abbildung. Für $A \subseteq M_1$ heißt

$$\begin{aligned} f(A) &:= \{f(x) \mid x \in A\} \\ &= \{y \in M_2 \mid \exists x \in A : f(x) = y\} \end{aligned}$$

das *Bild* von A unter f . Die Menge $f(M_1)$ heißt *Bild* oder *Wertemenge* von f und wird auch mit Bild f bezeichnet. Für $B \subseteq M_2$ heißt

$$f^{-1}(B) := \{x \in M_1 \mid f(x) \in B\}$$

das *Urbild* von B bezüglich f .

f heißt *injektiv* (Injektion) $\Leftrightarrow \forall x, y \in M_1 : (x \neq y \Rightarrow f(x) \neq f(y))$

f heißt *surjektiv* (Surjektion) auf $M_2 \Leftrightarrow f(M_1) = M_2$

f heißt *bijektiv* (Bijektion) auf $M_2 \Leftrightarrow f$ injektiv und surjektiv auf M_2 .

Ist $f : M_1 \rightarrow M_2$ injektiv, so ist

$$f^{-1} := \{(f(x), x) \mid x \in M_1\}$$

eine bijektive Abbildung von Bild $f = f(M_1)$ auf M_1 ; f^{-1} heißt die *Umkehrabbildung* von f . Es gilt

$$\begin{aligned} f^{-1}(f(x)) &= x && \text{für alle } x \in M_1, \\ f(f^{-1}(y)) &= y && \text{für alle } y \in \text{Bild } f. \end{aligned}$$

Sind $f : M_1 \rightarrow M_2$ und $g : M_3 \rightarrow M_4$ Funktionen mit Bild $f \subseteq M_3$, so definiert man

$$g \circ f := \{(x, g(f(x))) \mid x \in M_1\}.$$

$g \circ f$ ist also eine Funktion von M_1 in M_4 . Es gilt $(g \circ f)(x) = g(f(x))$ für alle $x \in M_1$. $g \circ f$ heißt *Verkettung* oder *Komposition* von g und f .

Für eine beliebige Menge M ist die *Identität* auf M erklärt als die Abbildung $\text{id}_M : M \rightarrow M$ mit $\text{id}_M(x) := x$ für alle $x \in M$. Ist $f : M_1 \rightarrow M_2$ eine Bijektion auf M_2 und f^{-1} die Umkehrabbildung, so hat man offenbar

$$f^{-1} \circ f = \text{id}_{M_1}, \quad f \circ f^{-1} = \text{id}_{M_2}.$$

Ist $f : M_1 \rightarrow M_2$ eine Funktion und $M \subseteq M_1$, so heißt

$$f|_M := \{(x, f(x)) \mid x \in M\}$$

die *Einschränkung* oder *Restriktion* von f auf M .

Im weiteren Verlauf der Analysis werden wir es zunächst mit speziellen Funktionen, den Folgen, zu tun haben.

Definition. Sei M eine Menge. Eine *Folge in M* ist eine Abbildung $f: \mathbb{N} \rightarrow M$ (an die Stelle von \mathbb{N} kann auch $\{m \in \mathbb{Z} \mid m \geq m_0\}$ treten).

Bezeichnungen. Ist $f : \mathbb{N} \rightarrow M$ eine Folge, so schreibt man statt $f(n)$ meist f_n und statt f oft

$$(f_n)_{n \in \mathbb{N}} \quad \text{oder} \quad (f_1, f_2, f_3, \dots).$$

Man nennt f_n auch das *n -te Folgenglied*.

Bemerkung. Achtung! Man verwechsle nicht die Folge (f_1, f_2, f_3, \dots) mit der Menge $\{f_1, f_2, f_3, \dots\}$. Zum Beispiel ist $(-1, 1, -1, 1, -1, \dots)$ die Abbildung, die der Zahl $n \in \mathbb{N}$ die Zahl $(-1)^n$ zuordnet, und $\{-1, 1, -1, 1, \dots\}$ ist die Menge $\{1, -1\}$.

Später werden wir es vorwiegend mit reellen Funktionen zu tun haben.

Definition. Eine *reelle Funktion* ist eine Abbildung $f : D \rightarrow \mathbb{R}$ mit $D \subseteq \mathbb{R}$.

Einige mögliche Eigenschaften reeller Funktionen:

Definition. Sei $f : D \rightarrow \mathbb{R}$ eine reelle Funktion. Sie heißt *beschränkt*, wenn ihr Wertebereich beschränkt (d.h. nach oben und nach unten beschränkt) ist, also:

$$f \text{ heißt beschränkt} \Leftrightarrow \exists c \in \mathbb{R}^+ \quad \forall x \in D : |f(x)| \leq c.$$

Ferner:

$$f \text{ heißt monoton wachsend (streng monoton wachsend)} \Leftrightarrow \forall x, y \in D : (x < y \Rightarrow f(x) \leq f(y)) \text{ (bzw. } < \text{)}.$$

Analog werden definiert: monoton fallend, streng monoton fallend, monoton (= wachsend oder fallend), streng monoton.

Um in diesem Abschnitt wenigstens etwas zu beweisen, zeigen wir:

1.1 Satz. *Sei $f : D \rightarrow \mathbb{R}$ eine streng monotone reelle Funktion. Dann ist f injektiv und die Umkehrfunktion f^{-1} streng monoton.*

Beweis. Sei f etwa streng monoton wachsend (im anderen Fall schließt man analog). Sei $x, y \in D, x \neq y$. Aus $x < y$ folgt $f(x) < f(y)$, also $f(x) \neq f(y)$, analog für $x > y$. Also ist f injektiv. Somit existiert die Umkehrfunktion $f^{-1} : W \rightarrow D$ mit $W := \text{Bild } f$. Sei $u, v \in W, u < v$. Dann ist $u = f(x), v = f(y)$ mit geeigneten $x, y \in D$. Wäre $y \leq x$, so folgte $f(y) \leq f(x)$, also $v \leq u$, was ein Widerspruch ist. Somit ist $f^{-1}(u) = x < y = f^{-1}(v)$; also ist auch f^{-1} streng monoton wachsend. ■

2.2 Abzählbarkeit

Die nun zur Verfügung stehenden Begriffe der bijektiven und surjektiven Abbildung erlauben eine Unterscheidung der „Größe“ von Mengen durch Vergleich mit \mathbb{N} oder Teilmengen davon.

Bezeichnungen. Für $n \in \mathbb{N}_0$ sei $A_n := \{k \in \mathbb{N} \mid k \leq n\}$, also $A_0 = \emptyset, A_1 = \{1\}, A_2 = \{1, 2\}$, u.s.w. A_n heißt der n -te Abschnitt von \mathbb{N} .

2.1 Satz (Schubfachprinzip). *Für $n, k \in \mathbb{N}$ mit $n > k$ gibt es keine injektive Abbildung $f : A_n \rightarrow A_k$.*

Beweis. Angenommen, das wäre falsch. Dann gibt es eine Zahl $n \in \mathbb{N}$ und eine injektive Abbildung $f : A_{n+1} \rightarrow A_n$ (denn jede Abbildung $f : A_{n+1} \rightarrow A_k$ mit $k < n$ ist auch Abbildung von A_{n+1} in A_n). Sei n die (nach dem Wohlordnungssatz 4.7 aus Kapitel 1 existierende) kleinste derartige Zahl.

1. Fall: $n \notin \text{Bild } f$ oder $n = f(n+1)$.

Setze $g := f|_{A_n}$ (Einschränkung von f auf A_n .)

2. Fall: $n = f(k)$ für ein $k < n+1$.

Setze

$$g(m) := \begin{cases} f(m) & \text{für } m \in A_n \setminus \{k\} \\ f(n+1) & \text{für } m = k. \end{cases}$$

In jedem Fall ist g eine injektive Abbildung von A_n in A_{n-1} . Das widerspricht der Wahl von n als kleinster Zahl mit einer derartigen Abbildung. ■

Definition. Eine Menge M heißt *n-elementig*, wenn es eine Bijektion von A_n auf M gibt. M heißt *endlich*, wenn M n -elementig für ein $n \in \mathbb{N}$ oder $M = \emptyset$ ist. Andernfalls heißt M *unendlich*.

Aus Satz 2.1 folgt insbesondere, dass M n -elementig nur für ein n sein kann; ferner folgt, dass es für eine endliche Menge keine Bijektion auf eine echte Teilmenge geben kann. Bei unendlichen Mengen ist das anders, zum Beispiel ist

$$f : \mathbb{N} \rightarrow \mathbb{N} \\ n \mapsto n + 1$$

eine Bijektion von \mathbb{N} auf die echte Teilmenge $\{m \in \mathbb{N} \mid m \geq 2\}$.

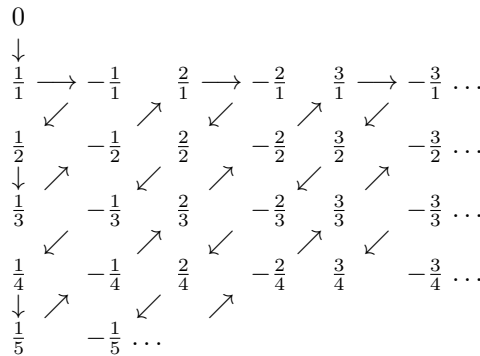
Definition. Eine Menge M heißt *abzählbar*, wenn es eine Surjektion von \mathbb{N} auf M gibt, und *überabzählbar*, wenn sie nicht abzählbar ist.

Man beachte, dass eine abzählbare Menge endlich oder unendlich sein kann. Im letzteren Fall heißt sie *abzählbar unendlich*.

Der folgende Satz ist in Anbetracht von Satz 4.8 aus Kapitel 1 zunächst vielleicht etwas überraschend:

2.2 Satz. Die Menge \mathbb{Q} der rationalen Zahlen ist abzählbar.

Beweis. Betrachte das untenstehende Schema. Jede rationale Zahl $\neq 0$ ist von der Form $\frac{m}{n}$ oder $-\frac{m}{n}$ mit $m, n \in \mathbb{N}$. Daher kommt im Schema jede rationale Zahl vor (in der Tat öfter, was aber nicht stört). Durchläuft man das Schema, indem man den Pfeilen folgt, so wird dadurch offenbar eine fortlaufende Numerierung erklärt, also eine Surjektion von \mathbb{N} auf \mathbb{Q} gegeben. (Man könnte die Abbildung explizit hinschreiben.)



■

Bemerkung. Mit derselben Beweisidee läßt sich offenbar zeigen: Die Vereinigung einer abzählbaren Familie (= Menge) von abzählbaren Mengen ist abzählbar.

Im Gegensatz hierzu ist die Menge \mathbb{R} der reellen Zahlen überabzählbar. Um das zu beweisen, zeigen wir zunächst das auch später sehr wichtige „Prinzip der Intervallschachtelung“:

2.3 Satz (Intervallschachtelungsprinzip). *Sei $(J_n)_{n \in \mathbb{N}}$ eine Folge abgeschlossener, beschränkter Intervalle (also $J_n = [a_n, b_n]$ mit $a_n, b_n \in \mathbb{R}$, $a_n \leq b_n$ für $n \in \mathbb{N}$) mit der Eigenschaft*

$$J_0 \supseteq J_1 \supseteq J_2 \supseteq \dots$$

Dann gibt es eine reelle Zahl $s \in \mathbb{R}$ mit $s \in \bigcap_{n \in \mathbb{N}} J_n$.

Beweis. Nach Voraussetzung gilt $a_k \leq a_n \leq b_n \leq b_k$ für $k < n$, also ist die Menge $A := \{a_n \mid n \in \mathbb{N}\}$ nach oben beschränkt. Nach dem Vollständigkeitsaxiom besitzt sie in \mathbb{R} eine kleinste obere Schranke s . Gäbe es ein $k \in \mathbb{N}$ mit $s \notin J_k$, so wäre $s < a_k$ oder $s > b_k$. Im ersten Fall ist s nicht obere Schranke für A , im zweiten Fall ist s nicht kleinste obere Schranke, da b_k obere Schranke für A ist, was ein Widerspruch ist. ■

Bemerkung. In Satz 2.3 ist die Voraussetzung, dass die Intervalle J_n abgeschlossen und beschränkt sind, nicht entbehrlich, z.B. gilt

$$\bigcap_{n \in \mathbb{N}} [n, \infty) = \emptyset, \quad \bigcap_{n \in \mathbb{N}} (0, \frac{1}{n}) = \emptyset.$$

2.4 Satz. *Die Menge \mathbb{R} der reellen Zahlen ist überabzählbar.*

Beweis. Angenommen, das wäre falsch. Dann gibt es eine Surjektion f von \mathbb{N} auf \mathbb{R} . Setze $f(n) = x_n$, also $\mathbb{R} = \{x_n \mid n \in \mathbb{N}\}$. Wähle ein Intervall $J_1 = [a_1, b_1]$ mit $a_1 < b_1$ und $x_1 \notin J_1$. Ist J_n schon definiert, so wähle ein Intervall $J_{n+1} = [a_{n+1}, b_{n+1}]$ mit $a_{n+1} < b_{n+1}$ und $J_{n+1} \subseteq J_n$, so dass $x_{n+1} \notin J_{n+1}$ ist. Damit ist rekursiv eine Folge $(J_n)_{n \in \mathbb{N}}$ definiert, die die Voraussetzungen von Satz 2.3 erfüllt. Es gibt also eine Zahl $s \in \mathbb{R}$ mit $s \in \bigcap_{n \in \mathbb{N}} J_n$. Wegen $\mathbb{R} = \{x_n \mid n \in \mathbb{N}\}$ gibt es eine Zahl $k \in \mathbb{N}$ mit $s = x_k$. Nach Konstruktion ist aber $x_k \notin J_k$, q.e.a. ■

Aus den Sätzen 2.2 und 2.4 folgt, dass die Menge der irrationalen Zahlen überabzählbar ist (wäre $\mathbb{R} \setminus \mathbb{Q}$ abzählbar, so wäre $\mathbb{R} = \mathbb{Q} \cup (\mathbb{R} \setminus \mathbb{Q})$ abzählbar). Insbesondere ist damit die Existenz von irrationalen Zahlen gezeigt.

3 Konvergenz

3.1 Konvergente Folgen

Unter Folgen sollen, solange nichts anderes gesagt ist, Folgen in \mathbb{R} verstanden werden.

Definition. Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} und sei $a \in \mathbb{R}$. Die Folge $(a_n)_{n \in \mathbb{N}}$ heißt *konvergent gegen a* , und a heißt *Grenzwert* oder *Limes* dieser Folge, wenn zu jedem $\varepsilon \in \mathbb{R}^+$ ein $n_0 \in \mathbb{N}$ existiert mit

$$|a_n - a| < \varepsilon \quad \text{für alle } n \geq n_0.$$

In Kurzschreibweise lautet die Definition:

$$(a_n)_{n \in \mathbb{N}} \text{ konvergiert gegen } a \\ \Leftrightarrow \forall \varepsilon \in \mathbb{R}^+ \exists n_0 \in \mathbb{N} \forall n \in \mathbb{N} : (n \geq n_0 \Rightarrow |a_n - a| < \varepsilon).$$

Eine Folge heißt *konvergent*, oder sie *konvergiert*, wenn sie konvergent gegen ein $a \in \mathbb{R}$ ist. Eine gegen 0 konvergente Folge heißt *Nullfolge*. Eine Folge heißt *divergent*, oder sie *divergiert*, wenn sie nicht konvergiert.

1.1 Satz. *Der Grenzwert einer konvergenten Folge ist eindeutig bestimmt.*

Beweis. Die Folge $(a_n)_{n \in \mathbb{N}}$ konvergiere gegen a und gegen b . Angenommen, es wäre $a \neq b$. Zu $\varepsilon := \frac{1}{2}|a - b|$ existieren ein $n_0 \in \mathbb{N}$ mit $|a_n - a| < \varepsilon$ für $n \geq n_0$ und ein $n_1 \in \mathbb{N}$ mit $|a_n - b| < \varepsilon$ für $n \geq n_1$. Für beliebiges $n \geq \max\{n_0, n_1\}$ gilt dann nach Behauptung 2.15 aus Kapitel 1 $|a - b| = |a - a_n - (b - a_n)| \leq |a - a_n| + |b - a_n| < \varepsilon + \varepsilon = |a - b|$, was ein Widerspruch ist. ■

Bezeichnungen. Für die Aussage „ $(a_n)_{n \in \mathbb{N}}$ konvergiert gegen a “ schreibt man auch

$$\lim_{n \rightarrow \infty} a_n = a \quad \text{oder} \quad a_n \rightarrow a \quad \text{für } n \rightarrow \infty.$$

Zur Vereinfachung der Sprechweise sind ferner die folgenden Bezeichnungen zweckmäßig:

Definition. Sei $a \in \mathbb{R}$, $\varepsilon \in \mathbb{R}^+$. Das Intervall $(a - \varepsilon, a + \varepsilon)$ heißt ε -Umgebung von a . Eine Teilmenge $U \subset \mathbb{R}$ heißt *Umgebung von a* , wenn sie eine ε -Umgebung von a für geeignetes $\varepsilon \in \mathbb{R}^+$ enthält.

„Fast alle“ heißt: alle, bis auf endlich viele.

Damit können wir auch formulieren:

$$\begin{aligned} \lim_{n \rightarrow \infty} a_n = a &\Leftrightarrow \text{Zu jeder Umgebung } U \text{ von } a \text{ existiert } n_0 \in \mathbb{N} \\ &\text{mit } a_n \in U \text{ für alle } n \geq n_0 \\ &\Leftrightarrow \text{Jede Umgebung von } a \text{ enthält } a_n \text{ für fast alle } n. \end{aligned}$$

Bemerkungen. (a) Aus $\lim_{n \rightarrow \infty} a_n = a$ und $b_n = a_n$ für fast alle n folgt $\lim_{n \rightarrow \infty} b_n = a$.
 (b) Aus $\lim_{n \rightarrow \infty} a_n = a$ folgt $\lim_{n \rightarrow \infty} a_{n+k} = a$ für $k \in \mathbb{N}$, und umgekehrt.

Manchen Folgen kann man schon aufgrund des folgenden Satzes ansehen, dass sie nicht konvergieren:

1.2 Satz. *Jede konvergente Folge ist beschränkt.*

Beweis. Sei $(a_n)_{n \in \mathbb{N}}$ eine konvergente Folge mit Grenzwert a . Es gibt ein $n_0 \in \mathbb{N}$ mit $|a_n - a| < 1$ für $n \geq n_0$. Für $n \geq n_0$ gilt also

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

Mit $c := \max\{|a_1|, \dots, |a_{n_0-1}|, 1 + |a|\}$ gilt $|a_n| \leq c$ für alle $n \in \mathbb{N}$. ■

Beispiele. (1) Sei $a_n = 1/n$ für $n \in \mathbb{N}$.

Behauptung. $\lim_{n \rightarrow \infty} a_n = 0$.

Beweis. Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Nach Behauptung 4.3 in Kapitel 1 gibt es ein $n_0 \in \mathbb{N}$ mit $1/n_0 < \varepsilon$. Für $n \geq n_0$ gilt dann

$$\left| \frac{1}{n} - 0 \right| = \frac{1}{n} \leq \frac{1}{n_0} < \varepsilon. \quad \blacksquare$$

(2) Sei $a_n = (-1)^n$ für $n \in \mathbb{N}$.

Behauptung. $(a_n)_{n \in \mathbb{N}}$ divergiert.

Beweis. Angenommen, $\lim_{n \rightarrow \infty} a_n = a$. Dann existiert ein $n_0 \in \mathbb{N}$ mit $|a_n - a| < 1$ für $n \geq n_0$. Für beliebiges $n \geq n_0$ gilt dann

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

was ein Widerspruch ist. ■

(3) Sei $a_n = \sqrt{1 + \frac{2}{n}}$ für $n \in \mathbb{N}$.

Behauptung. $\lim_{n \rightarrow \infty} a_n = 1$.

Beweis. Zu gegebenem $\varepsilon \in \mathbb{R}^+$ wähle $n_0 > \frac{2}{\varepsilon}$. Mithilfe der Binomischen Formel $(a+b)(a-b) = a^2 - b^2$ gilt dann für $n \geq n_0$

$$\left| \sqrt{1 + \frac{2}{n}} - 1 \right| = \frac{\frac{2}{n}}{\sqrt{1 + \frac{2}{n}} + 1} < \frac{2}{n} < \varepsilon. \quad \blacksquare$$

(4) Sei $a_n = b^n$ für $n \in \mathbb{N}$ ($b \in \mathbb{R}$ gegeben).

1. Fall: $|b| < 1$. Aus Behauptung 4.11 aus Kapitel 1 folgt $\lim_{n \rightarrow \infty} b^n = 0$.
2. Fall: $|b| > 1$. Die Folge $(b^n)_{n \in \mathbb{N}}$ ist wegen Behauptung 4.11 aus Kapitel 1 nicht beschränkt, also nach Satz 1.2 divergent.
3. Fall: $b = 1$. trivial.
4. Fall: $b = -1$. divergent (s.o.).

Der folgende Satz faßt einige wichtige Rechenregeln für Grenzwerte zusammen.

1.3 Satz. Seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ konvergente Folgen mit $\lim_{n \rightarrow \infty} a_n = a$ und $\lim_{n \rightarrow \infty} b_n = b$. Dann sind die Folgen $(a_n + b_n)_{n \in \mathbb{N}}$, $(a_n b_n)_{n \in \mathbb{N}}$, $(\lambda a_n)_{n \in \mathbb{N}}$ mit $\lambda \in \mathbb{R}$ konvergent, und es gilt

$$\lim_{n \rightarrow \infty} (a_n + b_n) = a + b, \quad \lim_{n \rightarrow \infty} (a_n b_n) = ab, \quad \lim_{n \rightarrow \infty} (\lambda a_n) = \lambda a.$$

Ist $a \neq 0$, so gibt es ein $m \in \mathbb{N}$ mit $a_n \neq 0$ für $n \geq m$, und die Folge $(1/a_n)_{n \in \mathbb{N}, n \geq m}$ ist konvergent gegen $1/a$.

Beweis. (1) $a_n + b_n$

Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Es gibt ein $n_0 \in \mathbb{N}$ mit $|a_n - a| < \varepsilon/2$ für $n \geq n_0$ und ein $n_1 \in \mathbb{N}$ mit $|b_n - b| < \varepsilon/2$ für $n \geq n_1$. Für $n \geq \max\{n_0, n_1\}$ gilt also

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

(2) $a_n b_n$

Sei $\varepsilon \in \mathbb{R}^+$. Nach Satz 1.2 existiert ein $c \in \mathbb{R}^+$ mit $|b_n| \leq c$ für $n \in \mathbb{N}$. Ferner existiert ein $n_0 \in \mathbb{N}$ mit

$$|a_n - a| < \frac{\varepsilon}{c + |a|}, \quad |b_n - b| < \frac{\varepsilon}{c + |a|} \quad \text{für } n \geq n_0.$$

Für alle $n \geq n_0$ gilt also

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

(3) λa_n . Folgt aus (2), wenn $b_n = \lambda$ gesetzt wird.

(4) $1/a_n$

Sei also $a \neq 0$. Sei $\varepsilon \in \mathbb{R}^+$. Es gibt ein $n_0 \in \mathbb{N}$ mit $|a_n - a| < \min\{\frac{1}{2}|a|, \frac{1}{2}|a|^2\varepsilon\}$ für $n \geq n_0$. Für $n \geq n_0$ gilt also

$$|a| = |a - a_n + a_n| \leq |a - a_n| + |a_n| < \frac{1}{2}|a| + |a_n|,$$

folglich $|a_n| > \frac{1}{2}|a|$, insbesondere $a_n \neq 0$, und daher

$$\left| \frac{1}{a_n} - \frac{1}{a} \right| = \frac{|a - a_n|}{|a_n||a|} < \frac{\frac{1}{2}|a|^2\varepsilon}{\frac{1}{2}|a|^2} = \varepsilon. \quad \blacksquare$$

Beispiel. $\lim_{n \rightarrow \infty} \frac{5n^3 + 2n^2 + 1}{2n^3 - 15n} = ?$ Es gilt

$$\frac{5n^3 + 2n^2 + 1}{2n^3 - 15n} = \frac{5 + \frac{2}{n} + \frac{1}{n^3}}{2 - \frac{15}{n^2}} \rightarrow \frac{5 + 0 + 0}{2 - 0} = \frac{5}{2}$$

für $n \rightarrow \infty$.

1.4 Satz. Sind $(a_n)_{n \in \mathbb{N}}$, $(b_n)_{n \in \mathbb{N}}$ konvergente Folgen mit $a_n \leq b_n$ für $n \in \mathbb{N}$, so ist $\lim a_n \leq \lim b_n$.

Beweis. Angenommen, es wäre $a := \lim a_n > \lim b_n =: b$. Es gibt ein $n_0 \in \mathbb{N}$ mit

$$|a_n - a| < \frac{1}{2}(a - b) \quad \text{und} \quad |b_n - b| < \frac{1}{2}(a - b) \quad \text{für } n \geq n_0.$$

Für solches n folgt

$$\left. \begin{array}{l} a - a_n < \frac{1}{2}(a - b) \Rightarrow a_n > \frac{1}{2}(a + b) \\ b_n - b < \frac{1}{2}(a - b) \Rightarrow b_n < \frac{1}{2}(a + b) \end{array} \right\} \Rightarrow a_n > b_n, \text{ q.e.a.}$$

■

Bemerkung. Achtung! Auch aus der schärferen Voraussetzung $a_n < b_n$ für $n \in \mathbb{N}$ folgt nur $\lim a_n \leq \lim b_n$.

Wir ergänzen die Definition der Konvergenz noch durch die folgende, manchmal bequeme Verabredung.

Definition. Die Folge $(a_n)_{n \in \mathbb{N}}$ heißt *bestimmt divergent gegen* ∞ , geschrieben

$$\lim_{n \rightarrow \infty} a_n = \infty,$$

wenn zu jedem $c \in \mathbb{R}$ ein $n_0 \in \mathbb{N}$ existiert mit

$$a_n > c \quad \text{für alle } n \geq n_0.$$

Beispiele. (1) Aus $\lim a_n = 0$ und $a_n > 0$ für $n \in \mathbb{N}$ folgt $\lim 1/a_n = \infty$.

(2) Sei $a_n = \sqrt{1 + \frac{2}{n}} - 1$ und $b_n = n$. Dann wissen wir, dass b_n bestimmt divergent gegen ∞ ist, d.h. $\lim_{n \rightarrow \infty} b_n = \infty$, und aus dem vorletzten Beispiel, dass $\lim_{n \rightarrow \infty} a_n = 0$. Also können wir Satz 1.3 auf die Folge $(a_n b_n)_{n \in \mathbb{N}}$ nicht anwenden, insbesondere wissen wir auch nicht was $0 \cdot \infty$ ist. Allerdings gilt, mithilfe der Rechnung aus obigen Beispiel,

$$a_n b_n = \frac{\frac{2}{n}}{\sqrt{1 + \frac{2}{n}} + 1} n = \frac{2}{\left(\sqrt{1 + \frac{2}{n}} - 1\right) + 2} \rightarrow \frac{2}{0 + 2} = 1$$

für $n \rightarrow \infty$.

Kriterien für Konvergenz

Häufig wird eine Folge vorliegen, von der man Konvergenz vermutet, ohne den eventuell existierenden Grenzwert zu kennen. Man benötigt daher Kriterien für Konvergenz, in denen nicht explizit der Grenzwert vorkommt.

1.5 Satz. *Jede monotone, beschränkte Folge in \mathbb{R} konvergiert.*

Beweis. Sei $(a_n)_{n \in \mathbb{N}}$ eine beschränkte Folge, etwa monoton wachsend. Nach dem Vollständigkeitsaxiom hat die Menge $\{a_n | n \in \mathbb{N}\}$ eine kleinste obere Schranke s . Sei $\varepsilon \in \mathbb{R}^+$ gegeben. Es gibt ein $n_0 \in \mathbb{N}$ mit $a_{n_0} > s - \varepsilon$, denn andernfalls wäre $s - \varepsilon$ obere Schranke. Da $(a_n)_{n \in \mathbb{N}}$ monoton wachsend ist, gilt $a_n \geq a_{n_0} > s - \varepsilon$ für $n \geq n_0$. Da s obere Schranke ist, gilt außerdem $a_n \leq s$, also $|a_n - s| = s - a_n < \varepsilon$ für $n \geq n_0$. Somit ist $\lim_{n \rightarrow \infty} a_n = s$. ■

Diese hinreichende Konvergenzbedingung ist sehr nützlich, aber natürlich nur begrenzt anwendbar. Für feinere Konvergenzbetrachtungen sind die folgenden Begriffsbildungen zweckmäßig.

Definition. Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge und $(n_k)_{k \in \mathbb{N}}$ eine streng monoton wachsende Folge natürlicher Zahlen. Dann heißt die Folge

$$(a_{n_k})_{k \in \mathbb{N}} = (a_{n_1}, a_{n_2}, a_{n_3}, \dots)$$

Teilfolge der Folge $(a_n)_{n \in \mathbb{N}}$.

Aus der Definition folgt unmittelbar: Konvergiert eine Folge gegen a , so konvergiert auch jede ihrer Teilfolgen gegen a .

Definition. Eine Zahl heißt *Häufungswert* (oder *Häufungspunkt*) einer Folge, wenn sie Grenzwert einer Teilfolge ist.

Beispiel. $a_n := (-1)^n + 1/n$. Die Teilfolge $(a_{2k})_{k \in \mathbb{N}}$ konvergiert gegen 1, die Teilfolge $(a_{2k+1})_{k \in \mathbb{N}}$ gegen -1 . Diese beiden Zahlen sind also Häufungswerte der Folge; andere Häufungswerte gibt es offenbar nicht.

Der folgende Satz (das Beste, was man als „Umkehrung“ von Satz 1.2 aussagen kann), ist sehr wichtig und wird an vielen Stellen benutzt.

1.6 Satz (von Bolzano-Weierstraß). *Jede beschränkte Folge in \mathbb{R} besitzt eine konvergente Teilfolge.*

Beweis. Sei $(x_n)_{n \in \mathbb{N}}$ eine beschränkte Folge. Wir definieren zunächst rekursiv eine Folge $(J_n)_{n \in \mathbb{N}_0}$ abgeschlossener Intervalle $J_n = [a_n, b_n]$ derart, dass für $n \in \mathbb{N}$ gilt:

- (a) $J_n \subseteq J_{n-1}$,
- (b) für unendlich viele $k \in \mathbb{N}$ gilt $x_k \in [a_n, b_n]$,
- (c) $b_n - a_n = 2^{-n}(b_0 - a_0)$.

Hierzu werde $[a_0, b_0]$ so gewählt, dass $x_k \in [a_0, b_0]$ für alle $k \in \mathbb{N}$ gilt (das ist möglich, da die Folge beschränkt ist). Sei $[a_n, b_n]$ schon definiert, so dass (a), (b), (c) erfüllt sind. Wir halbieren das Intervall $[a_n, b_n]$ durch seinen Mittelpunkt $z = (a_n + b_n)/2$. Wenigstens eines der Intervalle $[a_n, z]$, $[z, b_n]$ enthält x_k für unendlich viele k ; ein solches Teilintervall nehmen wir als $[a_{n+1}, b_{n+1}] = J_{n+1}$. Dann sind (a), (b), (c) auch für J_{n+1} erfüllt. Damit ist die Folge $(J_n)_{n \in \mathbb{N}}$ rekursiv definiert.

Nun definieren wir rekursiv eine Teilfolge $(x_{k_n})_{n \in \mathbb{N}}$ mit $x_{k_n} \in [a_n, b_n]$ für $n \in \mathbb{N}$. Wähle k_1 so, dass $x_{k_1} \in [a_1, b_1]$ gilt. Sei k_n mit $x_{k_n} \in [a_n, b_n]$ schon definiert. Das Intervall $[a_{n+1}, b_{n+1}]$ enthält x_k für unendlich viele k , also ist die Menge $\{k \in \mathbb{N} \mid k > k_n \text{ und } x_k \in [a_{n+1}, b_{n+1}]\}$ nicht leer. Ihr kleinstes Element nehmen wir als k_{n+1} . Damit ist die Folge rekursiv definiert.

Nach dem Intervallschachtelungsprinzip 2.3 aus Kapitel 2 gibt es eine reelle Zahl $x \in \bigcap_{n \in \mathbb{N}} J_n$. Für alle $n \in \mathbb{N}$ haben wir $x_{k_n} \in [a_n, b_n]$ und $x \in [a_n, b_n]$, woraus folgt

$$|x_{k_n} - x| \leq b_n - a_n = \frac{b_0 - a_0}{2^n}.$$

Daraus folgt $\lim_{n \rightarrow \infty} x_{k_n} = x$. ■

Konvergiert die Folge $(a_n)_{n \in \mathbb{N}}$ gegen a , so auch (wie früher bemerkt) jede Teilfolge, also ist a der einzige Häufungswert. Für beschränkte Folgen läßt sich dies umkehren:

1.7 Satz. *Hat eine beschränkte Folge genau einen Häufungswert, so konvergiert sie gegen diesen Häufungswert.*

Beweis. Sei $(a_n)_{n \in \mathbb{N}}$ eine beschränkte Folge und a ihr einziger Häufungswert. Angenommen, die Folge konvergiere nicht gegen a . Das bedeutet:

$$\exists \varepsilon \in \mathbb{R}^+ \quad \forall n_0 \in \mathbb{N} \quad \exists n \geq n_0 : |a_n - a| \geq \varepsilon.$$

Wir nehmen ein solches $\varepsilon > 0$. Es gibt dann ein $n_1 \in \mathbb{N}$ mit $|a_{n_1} - a| \geq \varepsilon$. Ist n_k schon definiert, so ist die Menge $\{n \in \mathbb{N} \mid n \geq n_k + 1 \text{ und } |a_n - a| \geq \varepsilon\}$

nicht leer; sei n_{k+1} ihr kleinstes Element. Damit ist rekursiv eine Teilfolge $(a_{n_k})_{k \in \mathbb{N}}$ definiert mit $|a_{n_k} - a| \geq \varepsilon$ für alle $k \in \mathbb{N}$. Da sie beschränkt ist, besitzt sie nach Satz 1.6 eine konvergente Teilfolge. Für deren Grenzwert b gilt $|b - a| \geq \varepsilon$, also hat die ursprüngliche Folge zwei verschiedene Häufungswerte, was ein Widerspruch ist. ■

Nun sind wir in der Lage, ein notwendiges und hinreichendes Kriterium für die Konvergenz einer Folge aufzustellen, das nicht explizit auf den Grenzwert Bezug nimmt. Eine notwendige Bedingung für Konvergenz liegt nahe, indem man den Grenzwert eliminiert: Die Folge $(a_n)_{n \in \mathbb{N}}$ konvergiere gegen a . Sei $\varepsilon \in \mathbb{R}^+$ gegeben. Es gibt ein $n_0 \in \mathbb{N}$ mit $|a_n - a| < \varepsilon/2$ für $n \geq n_0$. Für alle $m, n \geq n_0$ gilt also

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

Diese für die Konvergenz notwendige Eigenschaft erweist sich auch als hinreichend. Zwecks größerer Klarheit und vor allem im Hinblick auf spätere Verallgemeinerungen ist eine Definition angebracht:

Definition. Die Folge $(a_n)_{n \in \mathbb{N}}$ heißt *Cauchy-Folge*, wenn zu jedem $\varepsilon \in \mathbb{R}^+$ ein $n_0 \in \mathbb{N}$ existiert mit $|a_m - a_n| < \varepsilon$ für alle $m, n \geq n_0$.

1.8 Satz (Konvergenzkriterium von Cauchy). *Eine Folge in \mathbb{R} ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.*

Beweis. dass jede konvergente Folge eine Cauchy-Folge ist, wurde bereits gezeigt. Sei nun $(a_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge.

Behauptung. 1. $(a_n)_{n \in \mathbb{N}}$ ist beschränkt.

Beweis. Da $(a_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge ist, existiert ein $n_0 \in \mathbb{N}$ so, dass $|a_m - a_n| < 1$ für $m, n \geq n_0$. Für alle $n \geq n_0$ gilt also $|a_n| = |a_n - a_{n_0} + a_{n_0}| \leq |a_n - a_{n_0}| + |a_{n_0}| < 1 + |a_{n_0}|$. Mit $c := \max\{|a_1|, \dots, |a_{n_0-1}|, 1 + |a_{n_0}|\}$ gilt also $|a_n| \leq c$ für alle $n \in \mathbb{N}$. Somit ist $(a_n)_{n \in \mathbb{N}}$ beschränkt. ■

Nun folgt aus dem Satz von Bolzano-Weierstraß 1.6, dass eine konvergente Teilfolge $(a_{n_k})_{k \in \mathbb{N}}$ existiert. Sei a ihr Grenzwert.

Behauptung. 2. $\lim_{n \rightarrow \infty} a_n = a$.

Beweis. Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Da $(a_n)_{n \in \mathbb{N}}$ Cauchy-Folge ist, existiert ein $n_0 \in \mathbb{N}$ mit

$$|a_n - a_m| < \frac{\varepsilon}{2} \quad \text{für } m, n \geq n_0.$$

Wegen $\lim_{k \rightarrow \infty} a_{n_k} = a$ existiert ein $j \in \mathbb{N}$ mit $n_j \geq n_0$ und

$$|a_{n_j} - a| < \frac{\varepsilon}{2}.$$

Für alle $n \geq n_0$ gilt also

$$\begin{aligned} |a_n - a| &= |a_n - a_{n_j} + a_{n_j} - a| \leq |a_n - a_{n_j}| + |a_{n_j} - a| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Damit ist $a_n \rightarrow a$ für $n \rightarrow \infty$ gezeigt. ■

Damit ist Satz 1.8 gezeigt. ■

Wir gehen noch auf eine Begriffsbildung ein, die bei feineren Konvergenzuntersuchungen bzw. bei nicht konvergenten Folgen von Nutzen ist. Zunächst sei daran erinnert, dass wir für eine Teilmenge $M \subset \mathbb{R}$ die kleinste obere Schranke, falls sie existiert, auch mit

$$\sup M = \textit{Supremum} \text{ von } M$$

und die größte untere Schranke mit

$$\inf M = \textit{Infimum} \text{ von } M$$

bezeichnet haben.

Definition. Sei $(a_n)_{n \in \mathbb{N}}$ eine beschränkte Folge in \mathbb{R} . Dann heißt

$$\limsup_{n \rightarrow \infty} a_n := \lim_{n \rightarrow \infty} (\sup\{a_k \mid k \geq n\})$$

der *Limes superior* der Folge, und

$$\liminf_{n \rightarrow \infty} a_n := \lim_{n \rightarrow \infty} (\inf\{a_k \mid k \geq n\})$$

heißt *Limes inferior* der Folge.

Zu dieser Definition ist eine Erläuterung erforderlich. Setzt man

$$b_n := \sup\{a_k \mid k \geq n\} \quad \text{für } n \in \mathbb{N},$$

so ist die Folge $(b_n)_{n \in \mathbb{N}}$ monoton fallend; da sie beschränkt ist, existiert nach Satz 1.5 der Grenzwert $\lim_{n \rightarrow \infty} b_n$. Man kann also in der Tat wie oben definieren. Entsprechendes gilt für \liminf .

Beispiel. $a_n = (-1)^n(1 + \frac{1}{n})$. Es ist

$$\sup\{a_k | k \geq n\} = \begin{cases} 1 + \frac{1}{n}, & \text{falls } n \text{ gerade,} \\ 1 + \frac{1}{n+1}, & \text{falls } n \text{ ungerade.} \end{cases}$$

Also ist $\limsup_{n \rightarrow \infty} a_n = 1$. Ferner ist

$$\inf\{a_k | k \geq n\} = \begin{cases} -(1 + \frac{1}{n}), & \text{falls } n \text{ ungerade,} \\ -(1 + \frac{1}{n+1}), & \text{falls } n \text{ gerade,} \end{cases}$$

also $\liminf_{n \rightarrow \infty} a_n = -1$.

Der folgende Satz enthält eine äquivalente Definition, die oft besser zu handhaben ist als die ursprüngliche.

1.9 Satz. Sei $(a_n)_{n \in \mathbb{N}}$ eine beschränkte Folge und $a \in \mathbb{R}$. Dann gilt:

(1) $\limsup a_n \leq a \Leftrightarrow \forall \varepsilon \in \mathbb{R}^+$ ist die Menge $\{k \in \mathbb{N} | a_k \geq a + \varepsilon\}$ endlich.

(2) $\limsup a_n \geq a \Leftrightarrow \forall \varepsilon \in \mathbb{R}^+$ ist die Menge $\{k \in \mathbb{N} | a_k \geq a - \varepsilon\}$ unendlich.

Entsprechendes gilt für \liminf .

Beweis. (1) „ \Rightarrow “: Sei $\limsup a_n \leq a$. Sei $\varepsilon \in \mathbb{R}^+$. Wäre die Menge $\{k \in \mathbb{N} | a_k \geq a + \varepsilon\}$ unendlich, so wäre

$$\sup\{a_k | k \geq n\} \geq a + \varepsilon \quad \text{für alle } n \in \mathbb{N},$$

also $\limsup a_n \geq a + \varepsilon > a$, was ein Widerspruch ist.

„ \Leftarrow “: Sei $\{k \in \mathbb{N} | a_k \geq a + \varepsilon\}$ endlich für jedes $\varepsilon \in \mathbb{R}^+$. Dann gibt es zu gegebenem $\varepsilon \in \mathbb{R}^+$ ein $n_0 \in \mathbb{N}$ so, dass $a_k < a + \varepsilon$ für $n \geq n_0$, also

$$\sup\{a_k | k \geq n\} \leq a + \varepsilon \quad \text{für } n \geq n_0.$$

Es folgt $\limsup a_n \leq a + \varepsilon$; da $\varepsilon \in \mathbb{R}^+$ beliebig war, folgt $\limsup a_n \leq a$.

(2) „ \Rightarrow “: Sei $\limsup a_n \geq a$. Sei $\varepsilon \in \mathbb{R}^+$. Wäre die Menge $\{k \in \mathbb{N} | a_k \geq a - \varepsilon\}$ endlich, so existierte ein $n_0 \in \mathbb{N}$ so, dass $a_k < a - \varepsilon$ für $n \geq n_0$, also

$$\sup\{a_k | k \geq n\} \leq a - \varepsilon \quad \text{für } n \geq n_0.$$

Daraus folgt $\limsup a_n \leq a - \varepsilon < a$, was ein Widerspruch ist.

„ \Leftarrow “: Sei $\{k \in \mathbb{N} \mid a_k \geq a - \varepsilon\}$ unendlich für jedes $\varepsilon \in \mathbb{R}^+$. Dann gilt für jedes $\varepsilon \in \mathbb{R}^+$

$$\sup\{a_k \mid k \geq n\} \geq a - \varepsilon \quad \text{für alle } n \in \mathbb{N},$$

also ist $\limsup a_n \geq a - \varepsilon$. Da $\varepsilon \in \mathbb{R}^+$ beliebig war, folgt $\limsup a_n \geq a$. ■

Der folgende Satz macht die Bedeutung von Limes superior und Limes inferior deutlicher und erklärt zugleich die Wahl dieser Bezeichnungen.

1.10 Satz. *Sei $(a_n)_{n \in \mathbb{N}}$ eine beschränkte Folge. Dann ist $\limsup a_n$ der größte und $\liminf a_n$ der kleinste Häufungswert dieser Folge.*

Beweis. Es genügt, die Behauptung für \limsup zu beweisen. Sei H die Menge der Häufungswerte der Folge $(a_n)_{n \in \mathbb{N}}$. Nach Satz 1.6 ist $H \neq \emptyset$, und H ist beschränkt. Setze $s := \sup H$ und $\limsup a_n =: a$.

Nach Satz 1.9 (1) ist für jedes $\varepsilon \in \mathbb{R}^+$ die Menge $\{k \in \mathbb{N} \mid a_k \geq a + \varepsilon\}$ endlich. Also kann keine Zahl $\geq a + \varepsilon$ Grenzwert einer Teilfolge (d.h. Häufungswert) sein. Für alle $h \in H$ gilt also $h \leq a$, folglich ist $s \leq a$.

Behauptung. $a \in H$ [Daraus folgt dann $a \leq s$, also $a = s$.]

Beweis. Wir konstruieren rekursiv eine Teilfolge, die gegen a konvergiert. Setze $k_1 = 1$. Sei k_{n-1} schon definiert. Nach Satz 1.9 (2) ist die Menge

$$\left\{ k \in \mathbb{N} \mid a_k \geq a - \frac{1}{n} \right\}$$

unendlich, also existiert ein $k_n \in \mathbb{N}$ mit $k_n > k_{n-1}$ und $a_{k_n} \geq a - 1/n$. Damit ist die Teilfolge $(a_{k_n})_{n \in \mathbb{N}}$ definiert. Für alle $n \in \mathbb{N}$ gilt (wegen $k_n \geq n$)

$$\sup\{a_k \mid k \geq n\} \geq a_{k_n} \geq a - \frac{1}{n}.$$

Für $n \rightarrow \infty$ konvergieren beide Seiten gegen a , woraus $\lim_{n \rightarrow \infty} a_{k_n} = a$ folgt. ■

Damit ist Satz 1.10 gezeigt. ■

Wir können nun ein weiteres notwendiges und hinreichendes Kriterium für die Konvergenz einer Folge aufstellen.

1.11 Satz. *Die Folge $(a_n)_{n \in \mathbb{N}}$ konvergiert genau dann, wenn sie beschränkt ist und wenn $\limsup a_n = \liminf a_n$ ist.*

Das folgt sofort aus Satz 1.2, 1.7, 1.10.

3.2 Reihen

Unter Verwendung des Grenzwertbegriffes kann man auch für gewisse Folgen $(a_n)_{n \in \mathbb{N}}$ die „unendliche Summe“ $a_1 + a_2 + \dots$ erklären. Solche unendlichen Reihen sind wichtig, da man spezielle reelle Zahlen oder Funktionen häufig durch Reihen ausdrücken kann.

Definition. Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} . Man nennt die Zahl

$$s_n := \sum_{k=1}^n a_k \quad (n \in \mathbb{N})$$

die n -te *Partialsomme* dieser Folge. Die Folge $(s_n)_{n \in \mathbb{N}}$ heißt *unendliche Reihe* und wird mit

$$\sum_{k=1}^{\infty} a_k$$

bezeichnet. Ist die Folge $(s_n)_{n \in \mathbb{N}}$ konvergent, so wird ihr Grenzwert ebenfalls mit $\sum_{k=1}^{\infty} a_k$ bezeichnet.

Hiernach bedeutet $\sum_{k=1}^{\infty} a_k$ also zweierlei; das führt erfahrungsgemäß aber nicht zu Mißverständnissen. Nach Definition ist

$$\sum_{k=1}^{\infty} a_k = s \Leftrightarrow \lim_{n \rightarrow \infty} \sum_{k=1}^n a_k = s.$$

Bemerkung. Es ist klar, was unter $\sum_{k=m}^{\infty} a_k$ (mit $m \in \mathbb{Z}$) zu verstehen ist.

Beispiel (unendliche geometrische Reihe). Sei $q \in \mathbb{R}$, $|q| < 1$. Nach Behauptung 4.15 aus Kapitel 1 gilt

$$\sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q},$$

also

$$\left| \sum_{k=0}^n q^k - \frac{1}{1 - q} \right| = \frac{|q|^{n+1}}{1 - q} \rightarrow 0 \quad \text{für } n \rightarrow \infty,$$

wobei wir Behauptung 4.11 aus Kapitel 1 benutzt haben. Folglich ist

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

Konvergenz einer Reihe bedeutet nach Definition nichts anderes als die Konvergenz der zugehörigen Partialsummen-Folge. Die bisher für Folgenkonvergenz gewonnenen Ergebnisse übertragen sich also unmittelbar auf Ergebnisse über Konvergenz von Reihen. Wir wollen diese Resultate im folgenden zusammenstellen. Da es sich nur um Umformulierungen handelt, erübrigt es sich, auf die Beweise einzugehen.

Aus Satz 1.3 und Satz 1.4 folgt:

2.1 Satz. Seien $\sum_{k=1}^{\infty} a_k$, $\sum_{k=1}^{\infty} b_k$ konvergente Reihen. Dann ist die Reihe $\sum_{k=1}^{\infty} (a_k + b_k)$ konvergent, und es gilt

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

Gilt $a_k \leq b_k$ für $k \in \mathbb{N}$, so ist $\sum_{k=1}^{\infty} a_k \leq \sum_{k=1}^{\infty} b_k$.

Aus Satz 1.5 folgt:

2.2 Satz. Jede Reihe $\sum_{k=1}^{\infty} a_k$ mit $a_k \geq 0$ für alle $k \in \mathbb{N}$ ist konvergent oder bestimmt divergent.

Beispiel. An dieser Stelle können wir die gewohnte Praxis rechtfertigen, reelle Zahlen durch (im Prinzip unendliche) Dezimalbrüche darzustellen. Zum Beispiel bedeutet die Schreibweise

$$\pi = 3,141592653\dots$$

ja verabredungsgemäß nichts anderes, als dass

$$\pi = 3 + \frac{1}{10} + \frac{4}{100} + \frac{1}{1000} + \frac{5}{10000} + \dots$$

ist, also dass π durch eine gewisse Reihe mit nichtnegativen Gliedern dargestellt werden kann.

Insbesondere stellt also jeder „unendliche Dezimalbruch“ eine reelle Zahl dar. Es ist nicht schwer zu zeigen, dass umgekehrt jede reelle Zahl als Dezimalbruch dargestellt werden kann (man bestimme die a_k rekursiv).

Wir fahren fort, Ergebnisse für Folgen auf Reihen umzuformulieren. Ist

$$s_n = \sum_{k=1}^n a_k,$$

so ist für $m < n$

$$|s_n - s_m| = \left| \sum_{k=m+1}^n a_k \right|.$$

Aus Satz 1.8 erhalten wir also sofort die folgende wichtige Aussage.

2.3 Satz (Konvergenzkriterium von Cauchy). *Die Reihe $\sum_{k=1}^{\infty} a_k$ ist genau dann konvergent, wenn zu jedem $\varepsilon \in \mathbb{R}^+$ ein $n_0 \in \mathbb{N}$ existiert mit*

$$\left| \sum_{k=m}^{m+p} a_k \right| < \varepsilon \quad \text{für } m \geq n_0 \text{ und } p \in \mathbb{N}_0.$$

Da die Bedingung insbesondere für $p = 0$ erfüllt sein muß, haben wir:

2.4 Satz. *Notwendig für die Konvergenz der Reihe $\sum_{k=1}^{\infty} a_k$ ist die Bedingung $\lim_{k \rightarrow \infty} a_k = 0$.*

Diese Bedingung ist aber nicht hinreichend.

Beispiel. Die „harmonische Reihe“ $\sum_{k=1}^{\infty} \frac{1}{k}$ ist bestimmt divergent. Es ist nämlich

$$1 + \left(\frac{1}{2}\right) + \underbrace{\left(\frac{1}{3} + \frac{1}{4}\right)}_{> 2 \cdot \frac{1}{4} = \frac{1}{2}} + \underbrace{\left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right)}_{> 4 \cdot \frac{1}{8} = \frac{1}{2}} + \underbrace{\left(\frac{1}{9} + \dots + \frac{1}{16}\right)}_{> 8 \cdot \frac{1}{16} = \frac{1}{2}} + \dots$$

u.s.w., also $s_{2^n} \geq 1 + n \cdot \frac{1}{2}$, woraus die Behauptung folgt.

Beispiel. Für $\alpha > 1$ ist $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$ konvergent. [Dabei sei $\alpha \in \mathbb{N}$, denn sonst ist bis jetzt k^α noch nicht erklärt. Wenn aber später k^α für $\alpha \in \mathbb{R}$ erklärt sein wird, bleiben Behauptung und Beweis für $\alpha > 1$ richtig.] Sei $n \in \mathbb{N}$. Wähle $m \in \mathbb{N}$ mit $n \leq 2^{m+1} - 1$. Dann gilt

$$\begin{aligned} s_n &= \sum_{k=1}^n \frac{1}{k^\alpha} \leq \sum_{k=1}^{2^{m+1}-1} \frac{1}{k^\alpha} = 1 + \left(\frac{1}{2^\alpha} + \frac{1}{3^\alpha}\right) + \dots + \left(\sum_{k=2^m}^{2^{m+1}-1} \frac{1}{k^\alpha}\right) \\ &\leq \sum_{\ell=0}^m 2^\ell \frac{1}{(2^\ell)^\alpha} = \sum_{\ell=0}^m (2^{1-\alpha})^\ell < \sum_{\ell=0}^{\infty} (2^{1-\alpha})^\ell = \frac{1}{1-2^{1-\alpha}}. \end{aligned}$$

Dabei wurde $2^{1-\alpha} < 1$ (wegen $\alpha > 1$) benutzt. Die Konvergenz folgt jetzt aus Satz 2.2.

In Fällen, wo die Voraussetzung von Satz 2.2 nicht erfüllt ist, ist häufig das folgende Kriterium von Nutzen.

2.5 Satz (Leibniz-Kriterium für alternierende Reihen). *Ist $(a_n)_{n \in \mathbb{N}_0}$ eine monoton fallende Nullfolge, so ist die Reihe*

$$\sum_{k=0}^{\infty} (-1)^k a_k$$

konvergent.

Beweis. Setze $s_n := \sum_{k=0}^n (-1)^k a_k$. Es gilt $s_{2n+2} - s_{2n} = -a_{2n+1} + a_{2n+2} \leq 0$, also ist die Folge $(s_{2n})_{n \in \mathbb{N}}$ monoton fallend. Analog ist die Folge $(s_{2n+1})_{n \in \mathbb{N}}$ monoton wachsend. Wegen $s_{2n+1} - s_{2n} = -a_{2n+1} \leq 0$ gilt

$$s_{2n} \geq s_{2n+1} \geq s_1 \quad \text{und} \quad s_{2n+1} \leq s_{2n} \leq s_0.$$

Nach Satz 1.5 existieren also

$$a := \lim_{n \rightarrow \infty} s_{2n} \quad \text{und} \quad b := \lim_{n \rightarrow \infty} s_{2n+1}.$$

Es ist $a - b = \lim_{n \rightarrow \infty} (s_{2n} - s_{2n+1}) = \lim_{n \rightarrow \infty} a_{2n+1} = 0$, also $a = b$.

Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Es gibt ein $n_0 \in \mathbb{N}$ mit

$$|s_{2n} - a| < \varepsilon \quad \text{für } 2n \geq n_0,$$

und es gibt ein $n_1 \in \mathbb{N}$ mit

$$|s_{2n+1} - a| < \varepsilon \quad \text{für } 2n + 1 \geq n_1.$$

Für $m \geq \max\{n_0, n_1\}$ gilt also

$$|s_m - a| < \varepsilon.$$

Damit ist $\sum_{k=0}^{\infty} (-1)^k a_k = a$ gezeigt. ■

Absolute Konvergenz

Feinere Konvergenzbetrachtungen für Reihen erfordern eine neue Begriffsbildung. Hierauf werden wir beispielshalber bei der folgenden Fragestellung geführt. Gilt für unendliche Reihen auch ein „Kommutativgesetz“? Mit anderen Worten: Hängen Konvergenz und Grenzwert einer Reihe von der Reihenfolge der Summanden ab, oder darf man die Summanden beliebig umordnen? Mit einer „Umordnung“ ist folgendes gemeint:

Definition. Sei $\sum_{k=1}^{\infty} a_k$ eine Reihe und $\tau : \mathbb{N} \rightarrow \mathbb{N}$ eine Permutation (bijektive Abbildung auf sich) von \mathbb{N} . Dann heißt die Reihe $\sum_{k=1}^{\infty} a_{\tau(k)}$ eine *Umordnung* der ursprünglichen.

Beispiel. Die Reihe $\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n}$ ist nach dem Leibniz-Kriterium konvergent. Da $\sum_{n=1}^{\infty} \frac{1}{n}$ divergiert, läßt sich leicht eine Umordnung angeben, die divergiert: Für $n \in \mathbb{N}$ gilt

$$\frac{1}{2^n + 1} + \frac{1}{2^n + 3} + \cdots + \frac{1}{2^{n+1} - 1} > 2^{n-1} \cdot \frac{1}{2^{n+1}} = \frac{1}{4},$$

also

$$\begin{aligned} & \frac{1}{1} - \frac{1}{2} \\ & + \frac{1}{3} - \frac{1}{4} \\ & + \left(\frac{1}{5} + \frac{1}{7} \right) - \frac{1}{6} \\ & + \left(\frac{1}{9} + \frac{1}{11} + \frac{1}{13} + \frac{1}{15} \right) - \frac{1}{8} \\ & + \dots \\ & + \left(\frac{1}{2^n + 1} + \cdots + \frac{1}{2^{n+1} - 1} \right) - \frac{1}{2n + 2} \\ & \rightarrow \infty \text{ für } n \rightarrow \infty. \end{aligned}$$

Wie sich zeigt, kann etwas derartiges nicht passieren, wenn neben $\sum a_k$ auch die Reihe $\sum |a_k|$ konvergiert.

Definition. Die Reihe $\sum_{k=1}^{\infty} a_k$ heißt *absolut konvergent*, wenn die Reihe $\sum_{k=1}^{\infty} |a_k|$ konvergiert.

2.6 Satz. *Jede absolut konvergente Reihe ist konvergent.*

Beweis. Sei $\sum |a_k|$ konvergent. Wir zeigen, dass $\sum a_k$ die Voraussetzung des Konvergenzkriteriums von Cauchy erfüllt. Sei $\varepsilon \in \mathbb{R}^+$. Da $\sum |a_k|$ konvergiert, existiert ein $n_0 \in \mathbb{N}$ mit

$$\sum_{k=m}^{m+p} |a_k| < \varepsilon \quad \text{für } m \geq n_0, p \in \mathbb{N}_0.$$

Für diese m, p gilt also erst recht

$$\left| \sum_{k=m}^{m+p} a_k \right| \leq \sum_{k=m}^{m+p} |a_k| < \varepsilon.$$

Nach Satz 2.3 ist die Reihe $\sum_{k=1}^{\infty} a_k$ konvergent. ■

Hier wurde die Ungleichung $|b_1 + \dots + b_n| \leq |b_1| + \dots + |b_n|$ benutzt, die natürlich aus der Dreiecksungleichung durch Induktion herzuleiten ist.

2.7 Satz (Umordnungssatz). *Sei $\sum_{k=1}^{\infty} a_k$ eine absolut konvergente Reihe mit Grenzwert a . Dann konvergiert auch jede Umordnung dieser Reihe gegen a .*

Beweis. Sei $\tau : \mathbb{N} \rightarrow \mathbb{N}$ eine Permutation von \mathbb{N} . Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Da $\sum |a_k|$ konvergiert, gibt es ein $m \in \mathbb{N}$ mit

$$\sum_{k=m+1}^{\infty} |a_k| < \frac{\varepsilon}{2}.$$

Wählen wir ein $n_0 \in \mathbb{N}$ mit $\{1, \dots, m\} \subset \{\tau(1), \dots, \tau(n_0)\}$, so gilt für alle $n \geq n_0$

$$\begin{aligned} \left| \sum_{k=1}^n a_{\tau(k)} - a \right| &\leq \left| \sum_{k=1}^n a_{\tau(k)} - \sum_{k=1}^m a_k \right| + \left| \sum_{k=1}^m a_k - a \right| \\ &\leq \sum_{k=m+1}^{\infty} |a_k| + \sum_{k=m+1}^{\infty} |a_k| \\ &< \varepsilon. \end{aligned}$$

Hierbei haben wir benutzt, dass für $\ell > m$ gilt

$$\left| \sum_{k=1}^m a_k - a \right| = \left| \sum_{k=1}^{\ell} a_k - a - \sum_{k=m+1}^{\ell} a_k \right| \leq \left| \sum_{k=1}^{\ell} a_k - a \right| + \sum_{k=m+1}^{\infty} |a_k|$$

Der Grenzwert $\ell \rightarrow \infty$ liefert die obige Abschätzung des zweiten Terms. ■

Nachdem schon hierdurch (und erst recht im weiteren Verlauf) die Bedeutung von absoluter Konvergenz unterstrichen wird, wollen wir Kriterien für absolute Konvergenz zusammenstellen. Am häufigsten anwendbar ist wohl das folgende:

2.8 Satz (Majorantenkriterium). *Seien $\sum a_k, \sum c_k$ Reihen. Gilt*

$$|a_k| \leq c_k \quad \text{für alle } k \in \mathbb{N}$$

und ist die Reihe $\sum c_k$ konvergent, dann ist die Reihe $\sum a_k$ absolut konvergent.

Beweis. Sei $\varepsilon \in \mathbb{R}^+$. Da $\sum c_k$ konvergiert, existiert ein $n_0 \in \mathbb{N}$ mit

$$\sum_{k=m}^{m+p} c_k < \varepsilon \quad \text{für } m \geq n_0, p \in \mathbb{N}_0.$$

Für diese m, p gilt also

$$\sum_{k=m}^{m+p} |a_k| \leq \sum_{k=m}^{m+p} c_k < \varepsilon.$$

Aus Satz 2.3 folgt die Behauptung. ■

Gilt $|a_k| \leq c_k$ für $k \in \mathbb{N}$, so nennt man die Reihe $\sum c_k$ eine *Majorante* der Reihe $\sum a_k$. Eine Reihe ist also absolut konvergent, wenn sie eine konvergente Majorante besitzt. Häufig brauchbare Vergleichsreihen sind die geometrische Reihe ($c_k = q^k, q < 1$) und die ζ -Reihen ($c_k = \frac{1}{k^\alpha}$ mit einem $\alpha > 1$).

Gilt dagegen $a_k \geq c_k \geq 0$ für $k \in \mathbb{N}_0$, so nennt man die Reihe $\sum c_k$ eine *Minorante* der Reihe $\sum a_k$. Besitzt die Reihe $\sum a_k$ eine bestimmt divergente Minorante, so ist sie selbst bestimmt divergent. Eine manchmal brauchbare Vergleichsreihe ist die harmonische Reihe.

Wenn sich keines der obigen Kriterien unmittelbar anwenden läßt, versuche man es mit einem der folgenden.

2.9 Satz (Quotientenkriterium). Sei $\sum a_k$ eine Reihe mit $a_k \neq 0$ für $k \in \mathbb{N}$. Es gebe eine Zahl $q \in \mathbb{R}$ mit $0 < q < 1$ und

$$\left| \frac{a_{n+1}}{a_n} \right| \leq q \quad \text{für } n \in \mathbb{N}.$$

Dann ist die Reihe $\sum a_n$ absolut konvergent.

Beweis. Aus der Voraussetzung folgert man durch Induktion sofort

$$|a_{n+1}| \leq q^n |a_1|.$$

Da die Reihe $\sum q^n |a_1| = |a_1| \sum q^n$ wegen $|q| < 1$ konvergiert, folgt die Behauptung aus dem Majorantenkriterium. ■

Natürlich genügt es, wenn beim Quotientenkriterium die Voraussetzungen $a_n \neq 0$ und $\left| \frac{a_{n+1}}{a_n} \right| \leq q$ für fast alle n erfüllt sind.

Besonders bequem ist das Quotientenkriterium anzuwenden, wenn es gelingt, $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| =: a$ zu bestimmen, falls dieser Grenzwert existiert. Ist

$a < 1$, so gilt für beliebiges q mit $a < q < 1$ jedenfalls $\left| \frac{a_{n+1}}{a_n} \right| < q$ für fast alle n ; die Reihe ist also absolut konvergent. Gilt $a > 1$, so gilt

$$\left| \frac{a_{n+1}}{a_n} \right| \geq 1 \quad \text{für fast alle } n.$$

In diesem Fall ist die Reihe divergent, da $(a_n)_{n \in \mathbb{N}}$ keine Nullfolge ist. Im Fall $a = 1$ kann sowohl absolute Konvergenz als auch Divergenz vorliegen (z.B. $\sum \frac{1}{n^2}$ und $\sum \frac{1}{n}$).

Die Formulierung eines weiteren nützlichen Kriteriums erfordert eine Vorbemerkung. Sind Zahlen $n \in \mathbb{N}$ und $a \in \mathbb{R}^+$ gegeben, so gibt es eine Zahl $b \in \mathbb{R}^+$ mit $b^n = a$. Man schreibt dann $b = \sqrt[n]{a}$ und nennt b die n -te Wurzel aus a . Sie ist eindeutig bestimmt, denn aus $0 < b < c$ folgt $b^n < c^n$. Die Existenz der n -ten Wurzel könnten wir bereits an dieser Stelle leicht zeigen; wir verzichten aber darauf, da sie sich später in allgemeinerem Rahmen ergibt.

Ferner sei daran erinnert, dass wir in Abschnitt 3.1 für eine beschränkte Folge $(a_n)_{n \in \mathbb{N}}$ den Limes superior $\limsup_{n \rightarrow \infty} a_n$ definiert hatten. Wir ergänzen diese Definition noch:

Definition. Ist $(a_n)_{n \in \mathbb{N}}$ nicht nach oben beschränkt, so schreibt man

$$\limsup_{n \rightarrow \infty} a_n = \infty.$$

Man vereinbart noch $\infty > a$ für $a \in \mathbb{R}$.

2.10 Satz (Wurzelkriterium). Die Reihe $\sum_{n=1}^{\infty} a_n$ ist absolut konvergent, wenn

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} < 1$$

ist, und divergent, wenn

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1$$

ist.

Beweis. Sei $\limsup \sqrt[n]{|a_n|} =: a < 1$. Dann gibt es ein $q < 1$ mit $\sqrt[n]{|a_n|} \leq q$ und daher $|a_n| \leq q^n$ für fast alle n . Aus dem Majorantenkriterium folgt die absolute Konvergenz der Reihe $\sum a_n$.

Ist $\limsup \sqrt[n]{|a_n|} > 1$, so gilt $|a_n| \geq 1$ für unendlich viele n ; also ist $(a_n)_{n \in \mathbb{N}}$ keine Nullfolge und daher $\sum a_n$ divergent. ■

Bemerkung. Besonders bequem ist die Anwendung des Wurzelkriteriums, wenn $\lim \sqrt[n]{|a_n|}$ existiert und bestimmt werden kann.

Man beachte, dass das Wurzelkriterium im Fall $\limsup \sqrt[n]{|a_n|} = 1$ keine Aussage über Konvergenz oder Divergenz der Reihe $\sum a_n$ ermöglicht.

Beispiele.

$$\sum \frac{n^2}{2^n} \quad (\text{Quot.-Krit.}), \quad \sum \frac{n!}{3 \cdot 5 \cdot 7 \cdots (2n+1)} \quad (\text{Quot.-Krit.}),$$

$$\sum (\sqrt[n]{n} - 1)^n \quad (\text{Wurzel-Krit. und}$$

$$b_n := (\sqrt[n]{n} - 1) \Rightarrow n = (b_n + 1)^n \geq 1 + \binom{n}{2} b_n^2$$

$$\sum \left(\frac{n}{n+1} \right)^{n^2} \quad (\text{Wurzel-Krit. und Bernoullische Ungl.})$$

Absolute Konvergenz spielt auch eine Rolle, wenn wir versuchen, zwei konvergente Reihen gliedweise zu multiplizieren. Betrachten wir zunächst endliche Summen. Es ist

$$(a_1 + \cdots + a_n)(b_1 + \cdots + b_m) = \sum a_i b_j,$$

wo über alle Paare $(i, j) \in \{1, \dots, n\} \times \{1, \dots, m\}$ (in beliebiger Reihenfolge) zu summieren ist. Wenn wir dies auf unendliche Reihen übertragen wollen, müssen wir für die Paare $(i, j) \in \mathbb{N} \times \mathbb{N}$ eine Reihenfolge festlegen. Es ist übersichtlich, die Paare (i, j) nach wachsendem $i + j$ anzuordnen (und die mit festem $i + j$ beliebig). Man definiert daher folgendermaßen:

Definition. Seien $\sum_{k=0}^{\infty} a_k$, $\sum_{k=0}^{\infty} b_k$ Reihen. Das *Cauchy-Produkt* dieser Reihen ist die Reihe

$$\sum_{k=0}^{\infty} c_k \quad \text{mit} \quad c_k := \sum_{j=0}^k a_j b_{k-j} = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0.$$

2.11 Satz. Sind die Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ absolut konvergent, so ist auch ihr Cauchy-Produkt absolut konvergent, und es gilt

$$\left(\sum_{k=0}^{\infty} a_k \right) \left(\sum_{k=0}^{\infty} b_k \right) = \sum_{k=0}^{\infty} \left(\sum_{j=0}^k a_j b_{k-j} \right).$$

Beweis. Setze $\sum_{j=0}^k a_j b_{k-j} =: c_k$. Partialsummen bezeichnen wir mit den entsprechenden großen Buchstaben, also

$$A_n := \sum_{k=0}^n a_k, \quad B_n := \sum_{k=0}^n b_k, \quad C_n := \sum_{k=0}^n c_k.$$

Behauptung. 1. $\lim_{n \rightarrow \infty} (A_n B_n - C_n) = 0$.

Beweis. Sei $\varepsilon \in \mathbb{R}^+$ gegeben. Es ist

$$A_n B_n - C_n = \sum_{\substack{i \leq n \\ j \leq n}} a_i b_j - \sum_{i+j \leq n} a_i b_j = \sum_{(i,j) \in \Delta_n} a_i b_j$$

mit

$$\Delta_n := \{(i, j) \in \mathbb{N}_0 \times \mathbb{N}_0 \mid i \leq n, j \leq n, i + j > n\}.$$

Setze

$$D_n := \left(\sum_{k=0}^n |a_k| \right) \left(\sum_{k=0}^n |b_k| \right) = \sum_{\substack{i \leq n \\ j \leq n}} |a_i b_j|.$$

Nach Voraussetzung und Satz 1.3 ist die Folge $(D_n)_{n \in \mathbb{N}}$ konvergent, also existiert nach Satz 1.8 ein $n_0 \in \mathbb{N}$ mit

$$|D_n - D_{n_0}| < \varepsilon \quad \text{für } n \geq n_0.$$

Es ist

$$D_n - D_{n_0} = \sum_{(i,j) \in \Gamma_n} |a_i b_j|$$

mit

$$\Gamma_n := \{(i, j) \in \mathbb{N}_0 \times \mathbb{N}_0 \mid i \leq n, j \leq n\} \setminus \{(i, j) \in \mathbb{N}_0 \times \mathbb{N}_0 \mid i \leq n_0, j \leq n_0\}.$$

Für alle $n \geq 2n_0$ gilt $\Delta_n \subset \Gamma_n$ und daher

$$|A_n B_n - C_n| \leq \sum_{(i,j) \in \Delta_n} |a_i b_j| \leq \sum_{(i,j) \in \Gamma_n} |a_i b_j| = |D_n - D_{n_0}| < \varepsilon.$$

Damit ist die 1. Behauptung bewiesen. ■

Aus der 1. Behauptung folgt nach Satz 1.3

$$\lim_{n \rightarrow \infty} C_n = \left(\lim_{n \rightarrow \infty} A_n \right) \left(\lim_{n \rightarrow \infty} B_n \right).$$

Behauptung. 2. $\sum_{k=0}^{\infty} c_k$ konvergiert absolut.

Beweis. Setze $c_k^* := \sum_{j=0}^k |a_j| |b_{k-j}|$. Da die Reihen $\sum |a_k|$, $\sum |b_k|$ absolut konvergieren, konvergiert nach dem Vorstehenden die Reihe $\sum_{k=0}^{\infty} c_k^*$. Es gilt

$$|c_k| \leq \sum_{j=0}^k |a_j b_{k-j}| = c_k^*.$$

Nach dem Majorantenkriterium ist $\sum c_k$ absolut konvergent. ■

Damit ist Satz 2.11 gezeigt. ■

3.3 Die Exponentialreihe

Als erste Anwendung des Vorstehenden behandeln wir nun die Exponentialreihe. Viele für die Analysis und ihre Anwendungen wichtige reelle Funktionen werden durch unendliche Reihen definiert. Als erstes und besonders wichtiges Beispiel hierfür wollen wir jetzt die Exponentialfunktion einführen. Warum sie zum Beispiel bei der Beschreibung physikalischer Vorgänge, aber auch in anderen Anwendungsgebieten, oft auftritt, läßt sich allerdings erst gut erklären, wenn wir die Differentiation von Funktionen eingeführt haben. Einen Hinweis gibt aber schon Satz 3.4.

3.1 Satz. Für jedes $x \in \mathbb{R}$ ist die Reihe

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}$$

absolut konvergent.

Beweis. Wegen

$$\left| \frac{x^{n+1} n!}{(n+1)! x^n} \right| = \frac{|x|}{n+1} < \frac{1}{2} \quad \text{für } n \geq 2|x|$$

folgt dies aus dem Quotientenkriterium 2.9. ■

Definition. Die Reihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ heißt *Exponentialreihe*. Die durch

$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \text{für } x \in \mathbb{R}$$

definierte Funktion $\exp : \mathbb{R} \rightarrow \mathbb{R}$ heißt *Exponentialfunktion*. Man setzt

$$\sum_{n=0}^{\infty} \frac{1}{n!} = \exp(1) = e \quad (\text{Eulersche Zahl}).$$

Später werden wir $\exp(x) = e^x$ schreiben, aber zuvor muß gezeigt werden, dass in der Tat $\exp(n) = e^n$ für $n \in \mathbb{Z}$ gilt.

Will man $\exp(x)$ näherungsweise aus der Exponentialreihe berechnen, so muß man eine „Fehlerabschätzung“ haben.

3.2 Satz. Wird das „Restglied“ $R_{m+1}(x)$ durch

$$\exp(x) = \sum_{n=0}^m \frac{x^n}{n!} + R_{m+1}(x)$$

definiert, so gilt

$$|R_{m+1}(x)| \leq 2 \frac{|x|^{m+1}}{(m+1)!} \quad \text{für } m \geq 2|x| - 2.$$

Beweis. Es ist

$$\begin{aligned} |R_{m+1}(x)| &= \left| \sum_{n=m+1}^{\infty} \frac{x^n}{n!} \right| \leq \sum_{n=m+1}^{\infty} \frac{|x|^n}{n!} \\ &= \frac{|x|^{m+1}}{(m+1)!} \left(1 + \frac{|x|}{m+2} + \frac{|x|^2}{(m+2)(m+3)} + \dots \right) \\ &\leq \frac{|x|^{m+1}}{(m+1)!} \left(1 + \frac{|x|}{m+2} + \left(\frac{|x|}{m+2} \right)^2 + \dots \right). \end{aligned}$$

Wird also $\frac{|x|}{m+2} \leq \frac{1}{2}$ vorausgesetzt, so folgt

$$|R_{m+1}(x)| \leq \frac{|x|^{m+1}}{(m+1)!} \left(1 + \frac{1}{2} + \frac{1}{4} + \dots \right) = 2 \frac{|x|^{m+1}}{(m+1)!}. \quad \blacksquare$$

Wenn also eine Genauigkeitsschranke $\alpha > 0$ vorgeschrieben ist, so kann man aus der Abschätzung in Satz 3.2 zu gegebenem x eine Zahl m errechnen, so dass

$$\left| \exp(x) - \sum_{n=0}^m \frac{x^n}{n!} \right| < \alpha$$

ist. Zum Beispiel berechnet man so (man nehme $m = 14$):

$$e = 2,718281828459 \pm 2 \cdot 10^{-12}.$$

Bemerkung. Für $x = 1$ können wir auch wie folgt abschätzen:

$$0 < R_{m+1}(1) < \frac{1}{(m+1)!} \left(1 + \frac{1}{m+1} + \left(\frac{1}{m+1} \right)^2 + \dots \right) \\ = \frac{1}{(m+1)!} \frac{1}{1 - \frac{1}{m+1}} = \frac{1}{m \cdot m!}.$$

Hieraus können wir folgern, dass die Zahl e irrational ist: Andernfalls wäre nämlich $e = \frac{k}{m}$ mit passenden $k, m \in \mathbb{N}$, also

$$0 < \frac{k}{m} - \sum_{n=0}^m \frac{1}{n!} < \frac{1}{m \cdot m!},$$

folglich

$$0 < km! - m \sum_{n=0}^m \frac{m!}{n!} < 1,$$

was nicht sein kann, da es sich um eine ganze Zahl handelt.

Wir benutzen jetzt die im vorigen Abschnitt behandelte Multiplikation von Reihen, um die wichtige Funktionalgleichung der Exponentialfunktion herzuleiten.

3.3 Satz. Für alle $x, y \in \mathbb{R}$ gilt $\exp(x + y) = \exp(x) \exp(y)$.

Beweis. Da die Reihen $\sum \frac{x^n}{n!}$ und $\sum \frac{y^n}{n!}$ absolut konvergieren, ist das Produkt ihrer Grenzwerte nach Satz 2.11 gleich dem Grenzwert ihres Cauchy-Produktes, also ist

$$\exp(x) \exp(y) = \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right) \left(\sum_{n=0}^{\infty} \frac{y^n}{n!} \right) = \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!} \\ = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} = \sum_{n=0}^{\infty} \frac{1}{n!} (x+y)^n = \exp(x+y),$$

wobei die binomische Formel benutzt wurde. ■

Folgerungen

Für $n \in \mathbb{N}$ ist

$$\exp(n) = \exp(1 + \dots + 1) = \exp(1) \cdots \exp(1) = \exp(1)^n = e^n$$

nach Definition der Zahl e . Aus der Reihendarstellung ergibt sich $\exp(0) = 1 = e^0$.

Für $x \in \mathbb{R}$ ist

$$1 = \exp(0) = \exp(-x + x) = \exp(-x) \exp(x),$$

also

$$\exp(-x) = \frac{1}{\exp x}.$$

Von nun an schreiben wir

$$\exp(x) = e^x \quad \text{für } x \in \mathbb{R};$$

da wir $\exp(k) = e^k$ für $k \in \mathbb{Z}$ gezeigt haben, ist das in Einklang mit der bereits festgelegten Bedeutung von e^x für $x \in \mathbb{Z}$. Mit dieser Schreibweise ist also

$$e^{x+y} = e^x e^y \quad \text{für } x, y \in \mathbb{R}, \quad e^{-x} = \frac{1}{e^x}.$$

Für $x > 0$ ist

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots > 1.$$

Für $x < 0$ ist $-x > 0$, also $e^x = \frac{1}{e^{-x}} > 0$. Also ist $e^x > 0$ für alle $x \in \mathbb{R}$. Für $x > y$ gilt ferner

$$\frac{e^x}{e^y} = e^{x-y} > 1, \quad \text{also } e^x > e^y.$$

Die e -Funktion ist also streng monoton wachsend.

Zum Schluß geben wir noch eine Darstellung der Exponentialfunktion durch einen Grenzwert. Sie gibt einen ersten Hinweis auf das Auftreten der Exponentialfunktion in verschiedenen Anwendungen (Interpretation: Kontinuierliche Verzinsung, radioaktiver Zerfall).

3.4 Satz. Für alle $x \in \mathbb{R}$ gilt

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

Beweis. Sei $x \in \mathbb{R}$ gegeben. Wir setzen $a_n := \left(1 + \frac{x}{n}\right)^n$ (für gegebenes $x \in \mathbb{R}$). Nach der binomischen Formel ist

$$a_n = \sum_{k=0}^n \binom{n}{k} \frac{x^k}{n^k}.$$

Zunächst sei $x \geq 0$. Für $k \geq 1$ gilt

$$\binom{n}{k} \frac{1}{n^k} = \frac{1}{k!} \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n} \leq \frac{1}{k!},$$

also

$$a_n \leq \sum_{k=0}^n \frac{x^k}{k!} \leq \sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x.$$

Sei $\varepsilon \in \mathbb{R}^+$. Es gibt ein $m \in \mathbb{N}$ mit

$$\sum_{k=0}^m \frac{x^k}{k!} > e^x - \frac{\varepsilon}{2}.$$

Für $n \geq m$ gilt

$$a_n \geq \sum_{k=0}^m \binom{n}{k} \frac{x^k}{n^k} = \sum_{k=0}^m \frac{x^k}{k!} \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n}.$$

Für $n \rightarrow \infty$ konvergiert der k -te Summand der rechten Seite gegen $\frac{x^k}{k!}$, also die rechte Seite gegen $\sum_{k=0}^m \frac{x^k}{k!}$. Daher gibt es ein $n_0 \geq m$ mit

$$a_n > \sum_{k=0}^m \frac{x^k}{k!} - \frac{\varepsilon}{2} \quad \text{für } n \geq n_0.$$

Für alle $n \geq n_0$ gilt also

$$e^x \geq a_n > e^x - \varepsilon.$$

Somit gilt $\lim_{n \rightarrow \infty} a_n = e^x$.

Sei nun $x < 0$. Es gilt

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

Nach der Bernoullischen Ungleichung ist für $n > |x|$

$$\left(1 - \left(\frac{x}{n}\right)^2\right)^n \geq 1 - \frac{x^2}{n},$$

daraus folgt

$$\lim_{n \rightarrow \infty} \left(1 - \frac{x}{n}\right)^n \left(1 + \frac{x}{n}\right)^n = 1.$$

Wegen $-x > 0$ ist

$$\lim_{n \rightarrow \infty} \left(1 - \frac{x}{n}\right)^n = e^{-x},$$

also nach Satz 1.3

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = \frac{1}{e^{-x}} = e^x. \quad \blacksquare$$

4 Topologie in \mathbb{R} und Stetigkeit

4.1 Topologische Eigenschaften von Mengen reeller Zahlen

Die genauere Untersuchung von reellen Funktionen, wie sie in den folgenden Abschnitten beabsichtigt ist, erfordert zunächst einige Kenntnisse über mögliche Eigenschaften ihrer Definitionsbereiche. Eigenschaften der Art, wie sie hier betrachtet werden sollen, bezeichnet man als „topologische“ Eigenschaften (von griechisch *topos* = Ort), da es bei ihnen weniger um Größeneigenschaften als um Lage- und gestaltliche Eigenschaften geht. Im folgenden sprechen wir (zur Unterstützung der Anschauung) gerne von „Punktmengen“ anstatt von Zahlenmengen. Gemeint sind aber immer Teilmengen von \mathbb{R} .

Zunächst sei an den Umgebungsbegriff erinnert. Für $x \in \mathbb{R}$ und $\varepsilon \in \mathbb{R}^+$ hatten wir die Menge

$$U_\varepsilon(x) = (x - \varepsilon, x + \varepsilon) = \{y \in \mathbb{R} \mid |x - y| < \varepsilon\}$$

als die ε -Umgebung von x bezeichnet. Eine Menge $U \subseteq \mathbb{R}$ heißt *Umgebung von x* , wenn U eine passende ε -Umgebung von x enthält. Wir bemerken:

1.1 Behauptung. *Der Durchschnitt von endlich vielen Umgebungen von x ist Umgebung von x .*

Beweis. Seien U_1, \dots, U_n Umgebungen von x . Es gibt also Zahlen $\varepsilon_1, \dots, \varepsilon_n \in \mathbb{R}^+$ mit $U_{\varepsilon_i}(x) \subseteq U_i$ ($i = 1, \dots, n$). Mit $\varepsilon := \min\{\varepsilon_1, \dots, \varepsilon_n\}$ gilt $U_\varepsilon(x) \subseteq U_{\varepsilon_i}(x) \subseteq U_i$ für $i = 1, \dots, n$, also $U_\varepsilon(x) \subseteq U_1 \cap \dots \cap U_n$, folglich ist $U_1 \cap \dots \cap U_n$ Umgebung von x . ■

Nun kommen wir zu zwei grundlegenden Eigenschaften, die Mengen reeller Zahlen haben können.

Definition. Die Menge $M \subseteq \mathbb{R}$ heißt *offen*, wenn sie Umgebung jedes ihrer Punkte ist.

Mit anderen Worten:

$$M \text{ ist offen} \Leftrightarrow \forall x \in M \exists \varepsilon \in \mathbb{R}^+ : U_\varepsilon(x) \subseteq M.$$

Beispiele. offene Intervalle (Beweis trivial), \emptyset, \mathbb{R} . Weitere Beispiele ergeben sich aus dem folgenden Satz:

1.2 Satz. *Die Vereinigung von (beliebig vielen) offenen Mengen ist offen. Der Durchschnitt von endlich vielen offenen Mengen ist offen.*

Beweis. trivial bzw. Behauptung 1.1. ■

Definition. Sei $M \subseteq \mathbb{R}$ eine Teilmenge und $a \in \mathbb{R}$. Dann heißt a *Häufungspunkt* von M , wenn jede Umgebung von a einen von a verschiedenen Punkt von M enthält.

Mit anderen Worten:

$$\begin{aligned} a \text{ ist Häufungspunkt von } M \\ \Leftrightarrow \text{Für jede Umgebung } U \text{ von } a \text{ gilt } (U \cap M) \setminus \{a\} \neq \emptyset \\ \Leftrightarrow \forall \varepsilon \in \mathbb{R}^+ \exists x \in M : 0 < |x - a| < \varepsilon. \end{aligned}$$

Man muß also unterscheiden zwischen einem Häufungspunkt einer Folge und einem Häufungspunkt einer Menge. Nicht jeder Häufungspunkt einer Folge $(a_n)_{n \in \mathbb{N}}$ ist notwendig auch Häufungspunkt der Menge $\{a_n | n \in \mathbb{N}\}$. (Beispiel: $a_n = (-1)^n$).

Bemerkung. Ist a Häufungspunkt von M , so liegen in jeder Umgebung von a unendlich viele Elemente von M .

Analog zum Satz von Bolzano-Weierstraß können wir folgende Variante herleiten, die ebenfalls als Satz von Bolzano-Weierstraß bezeichnet wird:

1.3 Satz (Bolzano-Weierstraß). *Zu jeder beschränkten unendlichen Menge $M \subseteq \mathbb{R}$ gibt es in \mathbb{R} einen Häufungspunkt.*

Beweis. Man zeigt dies in offensichtlicher Weise völlig analog zum Beweis von Satz 1.6 aus Kapitel 3 durch Intervallhalbierungsverfahren und Intervallschachtelung. ■

Definition. Die Menge $M \subseteq \mathbb{R}$ heißt *abgeschlossen*, wenn sie alle ihre Häufungspunkte enthält.

Beispiele. abgeschlossene Intervalle, \emptyset, \mathbb{R} .

Bemerkung. Achtung! Eine Menge braucht weder offen noch abgeschlossen zu sein; eine Menge kann sowohl offen als auch abgeschlossen sein.

Es besteht jedoch ein einfacher Zusammenhang zwischen offenen und abgeschlossenen Mengen, siehe Satz 1.4.

Bezeichnungen. Ist $M \subseteq \mathbb{R}$, so nennt man die Menge

$$M^c := \mathbb{R} \setminus M$$

das *Komplement* von M . Für $x \in \mathbb{R}$ gilt also: $x \in M^c \Leftrightarrow x \notin M$. Es ist $(M^c)^c = M$.

1.4 Satz. Für $M \subseteq \mathbb{R}$ gilt:

$$M \text{ offen} \Leftrightarrow M^c \text{ abgeschlossen.}$$

Beweis. „ \Rightarrow “: Sei M offen. Sei x Häufungspunkt von M^c . Angenommen, es wäre $x \notin M^c$, also $x \in M$. Dann ist M Umgebung von x mit $M \cap M^c = \emptyset$, was ein Widerspruch ist. Also ist $x \in M^c$.

„ \Leftarrow “: Sei M^c abgeschlossen. Sei $x \in M$. Dann ist $x \notin M^c$, also x nicht Häufungspunkt von M^c . Folglich existiert eine Umgebung U von x mit $U \cap M^c = \emptyset$, also mit $U \subseteq M$. Da $x \in M$ beliebig war, ist M offen. ■

Für die Komplementbildung gelten die folgenden sogenannten *de Morgan'schen Regeln*, die leicht zu beweisen sind (hier ist \mathcal{M} ein beliebiges System von Mengen):

$$\begin{aligned} \left(\bigcup_{M \in \mathcal{M}} M \right)^c &= \bigcap_{M \in \mathcal{M}} M^c \\ \left(\bigcap_{M \in \mathcal{M}} M \right)^c &= \bigcup_{M \in \mathcal{M}} M^c. \end{aligned}$$

Berücksichtigt man sie, so folgert man aus Satz 1.4 und Satz 1.2:

1.5 Satz. Der Durchschnitt von (beliebig vielen) abgeschlossenen Mengen ist abgeschlossen. Die Vereinigung von endlich vielen abgeschlossenen Mengen ist abgeschlossen.

Unter den abgeschlossenen Mengen sind die beschränkten besonders wichtig, daher haben sie einen besonderen Namen:

Definition. Eine Menge $M \subseteq \mathbb{R}$ heißt *kompakt*, wenn sie beschränkt und abgeschlossen ist.

Aus Satz 1.3 sowie Satz 1.6 aus Kapitel 3 ergibt sich unmittelbar:

1.6 Satz. Sei $M \subseteq \mathbb{R}$ kompakt. Dann gilt:

- (a) Jede unendliche Teilmenge von M besitzt einen Häufungspunkt in M .
- (b) Jede Folge in M besitzt eine Teilfolge, die gegen ein Element von M konvergiert.

Eine weitere wichtige Eigenschaft kompakter Mengen ist die folgende:

1.7 Satz. Jede nichtleere kompakte Menge $M \subseteq \mathbb{R}$ enthält ein Maximum (größtes Element) und ein Minimum (kleinstes Element).

Beweis. Sei $M \subseteq \mathbb{R}$ kompakt, $M \neq \emptyset$. Dann existiert $s := \sup M$. Zu jedem $\varepsilon \in \mathbb{R}^+$ existiert ein $x \in M$ mit $x > s - \varepsilon$. Folglich gilt entweder $s \in M$, oder s ist Häufungspunkt von M , also gilt ebenfalls $s \in M$. Somit ist $s = \max M$. Analog schließt man für das Minimum. ■

Eine etwas weniger einfach zu formulierende, aber bei vielen feineren Untersuchungen besonders wichtige Eigenschaft kompakter Mengen bringt der „Überdeckungssatz von Heine-Borel“ zum Ausdruck.

Definition. Ein System (= Menge) \mathcal{M} von Teilmengen von \mathbb{R} heißt *Überdeckung* der Menge $A \subseteq \mathbb{R}$, wenn $A \subseteq \bigcup_{M \in \mathcal{M}} M$ gilt. \mathcal{M} heißt *offene Überdeckung*, wenn alle Elemente von \mathcal{M} offene Mengen sind.

1.8 Satz (Überdeckungssatz von Heine-Borel). Sei $A \subseteq \mathbb{R}$ kompakt und \mathcal{M} eine offene Überdeckung von A . Dann enthält \mathcal{M} eine endliche Überdeckung von A .

Beweis. Angenommen, die Behauptung wäre falsch. Dann gibt es eine offene Überdeckung \mathcal{M} von A ohne endliche Teilüberdeckung, d.h. je endlich viele Mengen von \mathcal{M} haben die Eigenschaft, dass ihre Vereinigung nicht ganz A enthält. Wir definieren rekursiv eine Intervallschachtelung $(J_n)_{n \in \mathbb{N}}$ mit folgender Eigenschaft: $J_n \cap A$ kann nicht durch endlich viele Mengen von \mathcal{M} überdeckt werden. Da A beschränkt ist, existiert ein Intervall $J_1 = [a_1, b_1]$ mit $A \subseteq J_1$. Sei $J_n = [a_n, b_n]$ schon definiert, so dass obige Eigenschaft besteht. Mindestens eine der Mengen $[a_n, \frac{1}{2}(a_n + b_n)] \cap A$, $[\frac{1}{2}(a_n + b_n), b_n] \cap A$ benötigt zu ihrer Überdeckung unendlich viele Mengen aus \mathcal{M} , sei J_{n+1} das

erste Intervall mit dieser Eigenschaft. Also müssen alle J_n unendlich viele Elemente von A enthalten. Damit ist die Folge $(J_n)_{n \in \mathbb{N}}$ definiert. Nach Satz 2.3 aus Kapitel 2 existiert ein $x \in \bigcap_{n \in \mathbb{N}} J_n$. Jede Umgebung von x enthält fast alle J_n , insbesondere unendlich viele Punkte aus A . Also ist x Häufungspunkt von A und daher $x \in A$. Da \mathcal{M} Überdeckung von A ist, existiert ein $M \in \mathcal{M}$ mit $x \in M$. Da M offen ist, existiert ein $\varepsilon \in \mathbb{R}^+$ mit $U_\varepsilon(x) \subseteq M$. Es gibt ein $n_0 \in \mathbb{N}$ mit $b_{n_0} - a_{n_0} < \varepsilon$. Für alle $y \in J_{n_0}$ gilt dann (wegen $x \in J_{n_0}$) $|x - y| < \varepsilon$, also $y \in U_\varepsilon(x)$. Somit ist $J_{n_0} \subseteq U_\varepsilon(x) \subseteq M$. J_{n_0} wird also sogar durch die eine Menge $M \in \mathcal{M}$ überdeckt, ein Widerspruch. ■

Satz 1.8 gestattet auch eine Umkehrung:

1.9 Satz. *Sei $A \subseteq \mathbb{R}$ eine Menge mit der Eigenschaft, dass jede offene Überdeckung von A eine endliche Überdeckung von A enthält. Dann ist A kompakt.*

Beweis. Da die Überdeckung $\{U_1(a) \mid a \in A\}$ eine endliche Überdeckung von A enthält, ist A beschränkt. Sei $x \in A^c$. Dann ist das System $\{M_n \mid n \in \mathbb{N}\}$ mit $M_n := \{y \in \mathbb{R} \mid |x - y| > \frac{1}{n}\}$ eine offene Überdeckung von A . Es gibt also ein $k \in \mathbb{N}$ mit

$$A \subseteq \bigcup_{n=1}^k M_n = M_k,$$

folglich mit $U_{\frac{1}{k}}(x) \subseteq A^c$. Also ist A^c offen und damit A abgeschlossen. ■

4.2 Grenzwerte von Funktionen und Stetigkeit

Wir betrachten im folgenden reelle Funktionen, worunter wir Funktionen $f : D \rightarrow \mathbb{R}$ mit einem nichtleeren Definitionsbereich $D \subseteq \mathbb{R}$ verstehen wollen. Von den Funktionen, die einfache physikalische Vorgänge beschreiben (z.B. „Weg als Funktion der Zeit“ bei der Bewegung eines Massenpunktes) ist man es gewöhnt, dass „kleinen“ Änderungen des Argumentes nur „kleine“ Änderungen der Funktionswerte entsprechen. Der Versuch, diese Vorstellung mathematisch zu präzisieren, führt zur Begriffsbildung der Stetigkeit, die für den weiteren Aufbau der Analysis von grundlegender Bedeutung ist. Zur bequemerem Behandlung dieser Eigenschaft von Funktionen führen wir zunächst einen Grenzwertbegriff für Funktionen ein, der in Analogie zum Grenzwert von Folgen gebildet ist.

Definition. Sei $f : D \rightarrow \mathbb{R}$ eine reelle Funktion ($D \subseteq \mathbb{R}$), a ein Häufungspunkt von D und $b \in \mathbb{R}$. Man schreibt

$$\lim_{x \rightarrow a} f(x) = b, \quad (2.1)$$

wenn zu jedem $\varepsilon \in \mathbb{R}^+$ ein $\delta \in \mathbb{R}^+$ existiert mit

$$|f(x) - b| < \varepsilon \quad \text{für alle } x \in D \setminus \{a\} \text{ mit } |x - a| < \delta.$$

Gilt (2.1), so sagt man, dass $\lim_{x \rightarrow a} f(x)$ (der Grenzwert von f bei Annäherung an a) existiert.

Kurz gefaßt lautet die Definition also:

$$\begin{aligned} \lim_{x \rightarrow a} f(x) = b \\ \Leftrightarrow \forall \varepsilon \in \mathbb{R}^+ \exists \delta \in \mathbb{R}^+ \forall x \in D : (0 < |x - a| < \delta \Rightarrow |f(x) - b| < \varepsilon). \end{aligned}$$

Bei der Anwendung des Grenzwertbegriffes für Funktionen ist es oft nützlich, dass man die Untersuchung von $\lim_{x \rightarrow a} f(x)$ zurückführen kann auf die Untersuchung von Folgen-Grenzwerten:

2.2 Satz. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion und a ein Häufungspunkt von D . Dann gilt:

$$\lim_{x \rightarrow a} f(x) = b \Leftrightarrow (*) \text{ Für jede Folge } (x_n)_{n \in \mathbb{N}} \text{ in } D \setminus \{a\} \text{ mit } \lim_{n \rightarrow \infty} x_n = a \text{ gilt } \lim_{n \rightarrow \infty} f(x_n) = b.$$

Beweis. „ \Rightarrow “: Gelte $\lim_{x \rightarrow a} f(x) = b$. Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in $D \setminus \{a\}$ mit $\lim_{n \rightarrow \infty} x_n = a$. Sei $\varepsilon \in \mathbb{R}^+$. Wegen $\lim_{x \rightarrow a} f(x) = b$ gibt es ein $\delta \in \mathbb{R}^+$ mit

$$|f(x) - b| < \varepsilon \quad \text{für alle } x \in D \setminus \{a\} \text{ mit } |x - a| < \delta.$$

Wegen $\lim_{n \rightarrow \infty} x_n = a$ gibt es ein $n_0 \in \mathbb{N}$ mit

$$|x_n - a| < \delta \quad \text{für alle } n \geq n_0.$$

Für $n \geq n_0$ gilt also $|f(x_n) - b| < \varepsilon$. Damit ist $\lim_{n \rightarrow \infty} f(x_n) = b$ gezeigt.

„ \Leftarrow “: Gelte (*). Angenommen, es wäre *nicht* $\lim_{x \rightarrow a} f(x) = b$. Das bedeutet:

$$\exists \varepsilon \in \mathbb{R}^+ \forall \delta \in \mathbb{R}^+ \exists x \in D \setminus \{a\} : (|x - a| < \delta \wedge |f(x) - b| \geq \varepsilon).$$

Insbesondere können wir zu jedem $n \in \mathbb{N}$ ein $x_n \in D \setminus \{a\}$ auswählen mit $|x_n - a| < \frac{1}{n}$ und $|f(x_n) - b| \geq \varepsilon$. Es gibt also eine Folge $(x_n)_{n \in \mathbb{N}}$ in $D \setminus \{a\}$ mit $\lim_{n \rightarrow \infty} x_n = a$, für die nicht $\lim_{n \rightarrow \infty} f(x_n) = b$ gilt, ein Widerspruch. ■

Bemerkung. Im letzten Teil des vorstehenden Beweises haben wir Gebrauch gemacht von der folgenden Tatsache:

(A) Ist $(M_n)_{n \in \mathbb{N}}$ eine Folge nichtleerer Mengen, so gibt es eine Abbildung $\alpha : \mathbb{N} \rightarrow \bigcup_{n \in \mathbb{N}} M_n$ mit $\alpha(n) \in M_n$ für alle $n \in \mathbb{N}$.

Obwohl wir in dieser Vorlesung die Mengenlehre nur in der sogenannten „naiven“ Weise benutzen, also ohne axiomatische Begründung, sei doch darauf hingewiesen, dass es sich hier um ein Axiom der Mengenlehre handelt (das abzählbare Auswahlaxiom). Es ist in der elementaren Analysis üblich, dieses Axiom stillschweigend zu benutzen.

Für die Existenz des Grenzwertes einer Funktion haben wir auch ein Cauchy-Kriterium zur Verfügung:

2.3 Satz. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion und a ein Häufungspunkt von D . Der Grenzwert $\lim_{x \rightarrow a} f(x)$ existiert genau dann, wenn gilt:

$$\begin{aligned} \forall \varepsilon \in \mathbb{R}^+ \quad \exists \delta \in \mathbb{R}^+ \quad \forall x, y \in D \setminus \{a\} : \\ (|x - a| < \delta \text{ und } |y - a| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon). \end{aligned} \quad (2.4)$$

Beweis. „ \Rightarrow “: Sei $\lim_{x \rightarrow a} f(x) = b$. Sei $\varepsilon \in \mathbb{R}^+$. Es gibt ein $\delta \in \mathbb{R}^+$ mit

$$|f(x) - b| < \frac{\varepsilon}{2} \quad \text{für alle } x \in D \setminus \{a\} \text{ mit } |x - a| < \delta.$$

Für alle $x, y \in D \setminus \{a\}$ mit $|x - a| < \delta$, $|y - a| < \delta$ gilt also

$$|f(x) - f(y)| \leq |f(x) - b| + |f(y) - b| < \varepsilon.$$

(2.4) ist also erfüllt.

„ \Leftarrow “: Gelte (2.4). Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in $D \setminus \{a\}$ mit $\lim_{n \rightarrow \infty} x_n = a$. Sei $\varepsilon \in \mathbb{R}^+$ gegeben; sei δ gemäß (2.4) gewählt. Es gibt ein $n_0 \in \mathbb{N}$ mit $|x_n - a| < \delta$ für $n \geq n_0$. Für $m, n \geq n_0$ gilt also $|f(x_m) - f(x_n)| < \varepsilon$. Somit ist die Folge $(f(x_n))_{n \in \mathbb{N}}$ eine Cauchyfolge und daher nach Satz 1.8 aus Kapitel 3 konvergent gegen eine Zahl b . Diese Zahl b hängt (wegen (2.4)) offenbar nicht von der speziellen Folge ab. Aus Satz 2.2 folgt also $\lim_{x \rightarrow a} f(x) = b$. ■

Zu dem eingeführten Grenzwertbegriff für Funktionen kann man einige naheliegende Varianten erklären, die manchmal nützlich sind.

Definition. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion und $a \in \mathbb{R}$. Ist a Häufungspunkt von $D \cap (-\infty, a)$, so schreibt man

$$\begin{aligned} \lim_{x \nearrow a} f(x) = b \\ \Leftrightarrow \forall \varepsilon \in \mathbb{R}^+ \quad \exists \delta \in \mathbb{R}^+ \quad \forall x \in D : (a - \delta < x < a \Rightarrow |f(x) - b| < \varepsilon). \end{aligned}$$

Ist a Häufungspunkt von $D \cap (a, \infty)$, so schreibt man

$$\lim_{x \searrow a} f(x) = b \\ \Leftrightarrow \forall \varepsilon \in \mathbb{R}^+ \exists \delta \in \mathbb{R}^+ \forall x \in D : (a < x < a + \delta \Rightarrow |f(x) - b| < \varepsilon).$$

Mit anderen Worten, es ist

$$\lim_{x \nearrow a} f(x) = b \Leftrightarrow \lim_{x \rightarrow a} f|_{(-\infty, a)}(x) = b,$$

wobei $f|_{(-\infty, a)}$ die Einschränkung von f auf $(-\infty, a)$ ist. Entsprechendes gilt für $\lim_{x \searrow a}$. Jetzt ist auch klar, dass sich die Sätze 2.2 und 2.3 sinngemäß auf diese „linksseitigen“ und „rechtsseitigen“ Grenzwerte übertragen lassen.

Definition. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion und a ein Häufungspunkt von D . Man schreibt

$$\lim_{x \rightarrow a} f(x) = \infty \\ \Leftrightarrow \forall c \in \mathbb{R}^+ \exists \delta \in \mathbb{R}^+ \forall x \in D \setminus \{a\} : (|x - a| < \delta \Rightarrow f(x) > c).$$

Analog definiert man $\lim_{x \rightarrow a} f(x) = -\infty$, $\lim_{x \nearrow a} f(x) = \infty$, u.s.w.

Definition. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion und D nicht nach oben beschränkt. Man schreibt

$$\lim_{x \rightarrow \infty} f(x) = b \\ \Leftrightarrow \forall \varepsilon \in \mathbb{R}^+ \exists c \in \mathbb{R} \forall x \in D : (x > c \Rightarrow |f(x) - b| < \varepsilon).$$

Analog definiert man $\lim_{x \rightarrow -\infty} f(x) = b$, $\lim_{x \rightarrow \infty} f(x) = \infty$, u.s.w.

Beispiele. Sei $D := \mathbb{R} \setminus \{0\}$, $f : D \rightarrow \mathbb{R}$ definiert durch $x \mapsto \frac{1}{x}$. Dann ist

$$\lim_{x \nearrow 0} f(x) = -\infty, \quad \lim_{x \searrow 0} f(x) = \infty.$$

Mit demselben D sei $f : D \rightarrow \mathbb{R}$ definiert durch $x \mapsto \frac{1}{x^2}$. Dann ist

$$\lim_{x \rightarrow 0} f(x) = \infty.$$

Sei $p : \mathbb{R} \rightarrow \mathbb{R}$ eine *Polynomfunktion*, das heißt

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \quad \text{für } x \in \mathbb{R}$$

mit gegebenen Zahlen $a_n, \dots, a_0 \in \mathbb{R}$. Sei etwa $a_n > 0$. Dann gilt (für $n \geq 1$)

$$\lim_{x \rightarrow \infty} p(x) = \infty,$$

$$\lim_{x \rightarrow -\infty} p(x) = \begin{cases} \infty, & \text{wenn } n \text{ gerade ist,} \\ -\infty, & \text{wenn } n \text{ ungerade ist.} \end{cases}$$

Dies liest man sofort an der für $x \neq 0$ gültigen Darstellung

$$p(x) = a_n x^n \left(1 + \frac{a_{n-1}}{a_n x} + \frac{a_{n-2}}{a_n x^2} + \dots + \frac{a_0}{a_n x^n} \right)$$

ab.

Stetigkeit

Die am Anfang dieses Paragraphen intuitiv erläuterte Stetigkeit einer Funktion läßt sich nun bequem mit Hilfe des Grenzwertbegriffes formulieren. Wir ziehen aber zunächst eine äquivalente Definition vor.

Definition. Sei $f : D \rightarrow \mathbb{R}$ eine reelle Funktion und $a \in D$. Die Funktion f heißt *stetig in a* , wenn zu jedem $\varepsilon \in \mathbb{R}^+$ ein $\delta \in \mathbb{R}^+$ existiert mit

$$|f(x) - f(a)| < \varepsilon \quad \text{für alle } x \in D \quad \text{mit } |x - a| < \delta.$$

Also kurz: f ist stetig in $a \Leftrightarrow$

$$\forall \varepsilon \in \mathbb{R}^+ \exists \delta \in \mathbb{R}^+ \forall x \in D : (|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon).$$

In gleichwertiger Weise können wir die Stetigkeit nun mit Verwendung des Grenzwertbegriffes ausdrücken, wie sich unmittelbar aus den Definitionen ergibt.

2.5 Satz. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion und $a \in D$ ein Häufungspunkt von D . Dann gilt

$$f \text{ stetig in } a \Leftrightarrow \lim_{x \rightarrow a} f(x) = f(a).$$

Bemerkung. Der Fall, dass $a \in D$, aber a nicht Häufungspunkt von D ist (a heißt dann *isolierter Punkt* von D), kann außer Betracht bleiben, da offenbar jede reelle Funktion in jedem isolierten Punkt ihres Definitionsbereiches stetig ist.

Unmittelbar aus Satz 2.5 und 2.2 erhalten wir:

2.6 Satz. Die Funktion $f : D \rightarrow \mathbb{R}$ ist in $a \in D$ genau dann stetig, wenn für jede Folge $(x_n)_{n \in \mathbb{N}}$ in D mit $\lim_{n \rightarrow \infty} x_n = a$ gilt: $\lim_{n \rightarrow \infty} f(x_n) = f(a)$.

Hieraus ergibt sich unmittelbar, dass Summen und Produkte von stetigen Funktionen stetig sind. Für Funktionen $f, g : D \rightarrow \mathbb{R}$ sind $f+g$ und fg erklärt durch $(f+g)(x) := f(x) + g(x)$ und $(fg)(x) := f(x)g(x)$ für alle $x \in D$.

2.7 Satz. Seien $f, g : D \rightarrow \mathbb{R}$ in $a \in D$ stetig. Dann sind die Funktionen $f+g$ und fg in a stetig. Ist $f(a) \neq 0$, so existiert eine Umgebung U von a mit $f(x) \neq 0$ für $x \in U \cap D$, und $1/f|_{U \cap D}$ ist stetig in a .

Beweis. Dies ergibt sich sofort aus Satz 2.6 und den Rechenregeln für Grenzwerte in Satz 1.3 aus Kapitel 3. Die Existenz von U folgt unmittelbar aus der Definition der Stetigkeit. ■

Bei Stetigkeitsbeweisen kann man wahlweise diejenige der äquivalenten Formulierungen benutzen, die gerade am bequemsten ist.

Beispiele. Zum Nachweis der Stetigkeit in a einer Funktion $f : D \rightarrow \mathbb{R}$ ist, wenn wir unmittelbar die gewählte Definition verwenden wollen, zu zeigen, dass zu gegebenem $\varepsilon \in \mathbb{R}^+$ ein $\delta \in \mathbb{R}^+$ existiert mit

$$x \in D \wedge |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

Wir geben im folgenden, soweit möglich, zu vorgegebenem ε ein solches δ explizit an.

- 1) Konstante Funktionen: δ beliebig.
- 2) $\text{id}_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}$ mit $x \mapsto x$, z.B. $\delta = \varepsilon$.
- 3) Hieraus und aus Satz 2.7 folgt, dass jede *rationale Funktion* in ihrem natürlichen Definitionsbereich stetig ist. Hierunter versteht man eine Funktion $f : D \rightarrow \mathbb{R}$ der Form $f(x) = \frac{p(x)}{q(x)}$ für $x \in D$, wobei p, q Polynomfunktionen sind und $D := \{x \in \mathbb{R} \mid q(x) \neq 0\}$ ist.
- 4) $\sqrt{\cdot} : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$, $x \mapsto \sqrt{x}$,
 1. Fall: $a = 0$. Wähle z.B. $\delta = \varepsilon^2$.
 2. Fall: $a > 0$. Wähle z.B. $\delta = \varepsilon\sqrt{a}$, denn es ist

$$|\sqrt{x} - \sqrt{a}| = \frac{|x - a|}{\sqrt{x} + \sqrt{a}} \leq \frac{|x - a|}{\sqrt{a}}.$$

5) $\text{abs} : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto |x|$

$$\delta = \varepsilon \text{ leistet das Gewünschte wegen } \left| |x| - |a| \right| \leq |x - a|$$

6) Die Exponentialfunktion: Nach Satz 3.2 aus Kapitel 3 gilt

$$|e^x - 1| \leq 2|x| \quad \text{für } |x| \leq 1, \text{ also}$$

$$|e^x - e^a| = e^a |e^{x-a} - 1| \leq e^a \cdot 2|x - a|.$$

Daher leistet $\delta := \min\{1, \frac{\varepsilon}{2e^a}\}$ das Gewünschte.

Nun zwei unstetige (d.h. in mindestens einem Punkt nicht stetige) Funktionen:

7) Definiere

$$f(x) := \begin{cases} 0, & \text{wenn } x < 0 \\ 1, & \text{wenn } x \geq 0 \end{cases} \quad \text{für } x \in \mathbb{R}.$$

Diese Funktion ist offenbar genau in 0 unstetig.

8) Die durch

$$f(x) := \begin{cases} 1, & \text{wenn } x \in \mathbb{Q} \\ 0, & \text{wenn } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

definierte sog. *Dirichlet-Funktion* ist offenbar in jedem Punkt $a \in \mathbb{R}$ unstetig. Denn in jeder Umgebung einer reellen Zahl liegen rationale und irrationale Zahlen.

Verschiedene äquivalente Formulierungen der Stetigkeit sind für die Anwendungen zweckmäßig. Zunächst eine besonders einprägsame unter Verwendung des Umgebungsbegriffes:

2.8 Satz. Die Funktion $f : D \rightarrow \mathbb{R}$ ist in $a \in D$ genau dann stetig, wenn zu jeder Umgebung V von $f(a)$ eine Umgebung U von a existiert mit $f(U \cap D) \subseteq V$.

Beweis. „ \Rightarrow “: Sei f stetig in a . Sei V eine Umgebung von $f(a)$. Dann existiert ein $\varepsilon \in \mathbb{R}^+$ mit $U_\varepsilon(f(a)) \subseteq V$. Zu diesem ε gibt es nach Voraussetzung ein $\delta \in \mathbb{R}^+$ mit

$$|f(x) - f(a)| < \varepsilon \quad \text{für alle } x \in D \text{ mit } |x - a| < \delta.$$

Setze $U := U_\delta(a)$. Dann gilt $f(U \cap D) \subseteq V$. In der Tat, sei $y \in f(U \cap D)$, also $y = f(x)$ mit $x \in U \cap D$. Wegen $x \in U_\delta(a)$ gilt $|x - a| < \delta$, also $|f(x) - f(a)| < \varepsilon$, d.h. $y = f(x) \in U_\varepsilon(f(a)) \subseteq V$.

„ \Leftarrow “: Zu jeder Umgebung V von $f(a)$ gebe es eine Umgebung U von a mit $f(U \cap D) \subseteq V$. Sei $\varepsilon \in \mathbb{R}^+$ gegeben. Zu $V := U_\varepsilon(f(a))$ existiert nach Voraussetzung eine Umgebung U von a mit $f(U \cap D) \subseteq V$. Es gibt ein $\delta \in \mathbb{R}^+$ mit $U_\delta(a) \subseteq U$. Sei jetzt $x \in D$ und $|x - a| < \delta$. Dann ist $x \in U_\delta(a) \subseteq U$, also $f(x) \in f(U \cap D) \subseteq V = U_\varepsilon(f(a))$, d.h. $|f(x) - f(a)| < \varepsilon$. ■

Mit dieser Formulierung der Stetigkeit ist bequem zu arbeiten. Hierfür ein Beispiel.

2.9 Satz. Sei $f : D \rightarrow \mathbb{R}$ stetig in einem $a \in D$, sei $g : D' \rightarrow \mathbb{R}$ mit $f(D) \subseteq D'$ stetig in $f(a)$. Dann ist $g \circ f$ stetig in a .

Beweis. Wir benutzen Satz 2.8. Sei W eine Umgebung von $(g \circ f)(a) = g(f(a))$. Da g stetig in $f(a)$ ist, existiert eine Umgebung V von $f(a)$ mit $g(V \cap D') \subseteq W$. Da f stetig in a ist, existiert eine Umgebung U von a mit $f(U \cap D) \subseteq V$. Also gilt $(g \circ f)(U \cap D) = g(f(U \cap D)) \subseteq g(V \cap D') \subseteq W$. ■

Schließlich wird noch naheliegenderweise definiert:

Definition. Die Funktion $f : D \rightarrow \mathbb{R}$ heißt *stetig*, wenn f stetig ist in a für alle $a \in D$.

Die Funktion im siebten der obigen Beispiele hat eine Unstetigkeit von besonders einfacher Art:

Definition. Die Funktion $f : D \rightarrow \mathbb{R}$ hat in $a \in D$ eine *Sprungstelle*, wenn die einseitigen Grenzwerte

$$\lim_{x \nearrow a} f(x) \quad \text{und} \quad \lim_{x \searrow a} f(x)$$

existieren und verschieden sind.

Ist f eine monotone Funktion, so ist leicht zu sehen, dass in jedem Punkt a des Definitionsbereiches die einseitigen Grenzwerte von $f(x)$ für $x \nearrow a$ und $x \searrow a$ existieren. Bei einer monotonen Funktion sind die Unstetigkeitsstellen also sämtlich Sprungstellen. Hieraus kann man folgern, dass bei einer monotonen Funktion höchstens abzählbar viele Unstetigkeitsstellen auftreten können.

Im übernächsten Paragraphen werden wir einige wichtige spezielle Funktionen als Umkehrfunktionen von gewissen streng monotonen Funktionen einführen. Nach Satz 1.1 aus Kapitel 2 ist eine streng monotone Funktion injektiv, und die daher existierende Umkehrfunktion f^{-1} ist ebenfalls streng monoton.

2.10 Satz. *Sei D ein Intervall, $f : D \rightarrow \mathbb{R}$ streng monoton. Dann ist die Umkehrfunktion $f^{-1} : f(D) \rightarrow \mathbb{R}$ stetig.*

Beweis. O.B.d.A. sei f streng monoton wachsend, und D enthalte mehr als einen Punkt. Seien $y \in f(D)$ und $\varepsilon \in \mathbb{R}^+$ vorgegeben. Es gilt $y = f(x)$ für ein (eindeutig bestimmtes) $x \in D$.

1. Fall: x ist nicht Endpunkt des Intervalls D . Dann existiert ein $\varepsilon' \in \mathbb{R}^+$ mit $\varepsilon' \leq \varepsilon$ und $[x - \varepsilon', x + \varepsilon'] \subseteq D$. Setze

$$\delta := \min\{y - f(x - \varepsilon'), f(x + \varepsilon') - y\}.$$

Wegen der strengen Monotonie von f ist $\delta > 0$. Sei jetzt $y_1 \in f(D)$, $|y_1 - y| < \delta$. Dann ist

$$f(x - \varepsilon') < y_1 < f(x + \varepsilon')$$

und daher nach Satz 1.1 aus Kapitel 2

$$x - \varepsilon' = f^{-1}(f(x - \varepsilon')) < f^{-1}(y_1) < f^{-1}(f(x + \varepsilon')) = x + \varepsilon',$$

wegen $x = f^{-1}(y)$ also $|f^{-1}(y_1) - f^{-1}(y)| < \varepsilon' \leq \varepsilon$.

2. Fall: x ist Endpunkt, etwa Maximum, des Intervalls D . Dann existiert ein $\varepsilon' \in \mathbb{R}^+$ mit $\varepsilon' \leq \varepsilon$ und $[x - \varepsilon', x] \subseteq D$. Setze $\delta := y - f(x - \varepsilon')$. Für $y_1 \in f(D)$ mit $|y_1 - y| < \delta$ gilt $f(x - \varepsilon') < y_1 \leq y$, woraus wie oben $|f^{-1}(y_1) - f^{-1}(y)| < \varepsilon$ folgt. ■

Man beachte, dass in Satz 2.10 nicht die Stetigkeit von f vorausgesetzt zu werden braucht. Andererseits läßt sich selbst bei stetigem f nicht auf die Stetigkeit von f^{-1} schließen, falls der Definitionsbereich von f kein Intervall ist. Dies läßt sich leicht durch Beispiele belegen.

4.3 Eigenschaften stetiger Funktionen

Wir stellen in diesem Abschnitt einige wichtige Sätze zusammen, die sich auf reelle Funktionen beziehen, die in ihrem ganzen Definitionsbereich stetig sind. Diese Sätze haben mannigfache Anwendungen. Aus der Stetigkeit

einer Funktion lassen sich besonders dann interessante Folgerungen ziehen, wenn der Definitionsbereich von spezieller Art ist. Zunächst betrachten wir kompakte Definitionsbereiche.

3.1 Satz. *Sei $D \subseteq \mathbb{R}$ kompakt und $f : D \rightarrow \mathbb{R}$ stetig. Dann ist $f(D)$ kompakt.*

Beweis. Wir benutzen die durch den Überdeckungssatz von Heine-Borel 1.8 und seine Umkehrung 1.9 gegebene Charakterisierung kompakter Mengen.

Sei \mathcal{M} eine offene Überdeckung von $f(D)$. Sei $x \in D$. Es gibt eine Menge $M \in \mathcal{M}$ mit $f(x) \in M$. Da M Umgebung von $f(x)$ und f stetig ist, können wir nach Satz 2.8 eine offene Umgebung U von x wählen mit $f(U \cap D) \subseteq M$. Das System

$$\{U \mid U \subseteq \mathbb{R} \text{ offen, } \exists M \in \mathcal{M} : f(U \cap D) \subseteq M\}$$

ist also eine offene Überdeckung von D und enthält daher nach Satz 1.8 eine endliche Teilüberdeckung $\{U_1, \dots, U_n\}$. Zu $i \in \{1, \dots, n\}$ gibt es ein $M_i \in \mathcal{M}$ mit $f(U_i \cap D) \subseteq M_i$. Es gilt $f(D) \subseteq \bigcup_{i=1}^n M_i$, also ist $\{M_1, \dots, M_n\}$ eine endliche Überdeckung von $f(D)$. Aus Satz 1.9 folgt die Kompaktheit von $f(D)$. ■

3.2 Satz. *Jede auf einer (nichtleeren) kompakten Menge stetige reelle Funktion nimmt dort ein Maximum (und analog ein Minimum) an.*

Genauer: Ist $D \subseteq \mathbb{R}$ kompakt und $f : D \rightarrow \mathbb{R}$ stetig, so existieren $a, b \in D$ mit $f(a) \leq f(x) \leq f(b)$ für alle $x \in D$. Der Beweis ergibt sich sofort aus Satz 3.1 und Satz 1.7, angewandt auf $f(D)$.

Nun betrachten wir stetige Funktionen, die auf Intervallen definiert sind. Der folgende Satz enthält eine der wichtigsten Eigenschaften stetiger Funktionen.

3.3 Satz. *Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, $f(a) < 0$ und $f(b) > 0$. Dann existiert ein $x_0 \in (a, b)$ mit $f(x_0) = 0$.*

Beweis. Wir definieren rekursiv eine Intervallschachtelung $(J_n)_{n \in \mathbb{N}}$, $J_n = [a_n, b_n]$, mit $f(a_n) < 0$ und $f(b_n) \geq 0$: Setze $J_1 := [a, b]$. Ist J_n schon definiert, so setze

$$J_{n+1} := \begin{cases} [a_n, \frac{1}{2}(a_n + b_n)], & \text{falls } f(\frac{1}{2}(a_n + b_n)) \geq 0, \\ [\frac{1}{2}(a_n + b_n), b_n], & \text{falls } f(\frac{1}{2}(a_n + b_n)) < 0. \end{cases}$$

Die Folge $(J_n)_{n \in \mathbb{N}}$ hat die angegebenen Eigenschaften. Nach dem Intervallschachtelungsprinzip existiert ein $x_0 \in \bigcap_{n \in \mathbb{N}} J_n$. Offenbar gilt $\lim_{n \rightarrow \infty} a_n =$

$x_0 = \lim_{n \rightarrow \infty} b_n$. Aus der Stetigkeit von f und Satz 2.7 folgt wegen $f(a_n) < 0$ und $f(b_n) \geq 0$

$$0 \geq \lim f(a_n) = f(x_0) = \lim f(b_n) \geq 0,$$

also $f(x_0) = 0$. ■

Bemerkung. Das Beweisverfahren ist konstruktiv, d.h. es kann zur näherungsweisen Berechnung einer Nullstelle verwendet werden.

3.4 Satz (Zwischenwertsatz). *Eine auf einem Intervall definierte stetige reelle Funktion nimmt jeden Wert zwischen zwei Funktionswerten an.*

Beweis. Sei D ein Intervall, $f : D \rightarrow \mathbb{R}$ stetig. Seien u, v zwei Werte von f , o.B.d.A. $u < v$. Es gibt also Elemente $a, b \in D$ mit $f(a) = u$, $f(b) = v$. Sei nun $u < z < v$. Im Fall $a < b$ wende man Satz 3.3 an auf die Funktion

$$\begin{aligned} g : [a, b] &\rightarrow \mathbb{R}, \\ x &\mapsto f(x) - z \end{aligned}$$

im Fall $a > b$ auf die Funktion

$$\begin{aligned} g : [a, b] &\rightarrow \mathbb{R}, \\ x &\mapsto -f(x) + z \end{aligned} \quad \blacksquare$$

Anwendungsbeispiel: Jede Polynomfunktion ungeraden Grades hat eine Nullstelle.

Gleichmäßige Stetigkeit

Werfen wir noch einmal einen Blick auf die Definition der (globalen) Stetigkeit. Dass $f : D \rightarrow \mathbb{R}$ stetig ist, bedeutet definitionsgemäß:

$$\forall x \in D \quad \forall \varepsilon \in \mathbb{R}^+ \quad \exists \delta \in \mathbb{R}^+ \quad \forall y \in D : (|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon).$$

Hier ist also zu gegebenem $x \in D$ und gegebenem $\varepsilon \in \mathbb{R}^+$ ein δ zu finden, so dass die rechts stehende Aussage gilt. Im allgemeinen wird (bei festem ε) dieses δ von dem vorgegebenem x abhängen. Beispiel:

$$\begin{aligned} f : (0, 1] &\rightarrow \mathbb{R} \\ x &\mapsto \frac{1}{x} \end{aligned}$$

Je näher x bei 0 liegt, umso kleiner muß (bei festem ε) das δ gewählt werden; es gibt kein $\delta \in \mathbb{R}^+$, mit dem man für alle $x \in D$ auskommt. Manche Konstruktionen und Beweise der Analysis sind aber nur möglich, wenn sich ein solches δ einheitlich für den ganzen Definitionsbereich wählen läßt. Funktionen, für die dies der Fall ist, heißen gleichmäßig stetig.

Definition. Die Funktion $f : D \rightarrow \mathbb{R}$ heißt *gleichmäßig stetig*, wenn zu jedem $\varepsilon \in \mathbb{R}^+$ ein $\delta \in \mathbb{R}^+$ existiert mit

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

für alle $x, y \in D$ mit $|x - y| < \delta$. Kurz: f ist gleichmäßig stetig \Leftrightarrow

$$\forall \varepsilon \in \mathbb{R}^+ \exists \delta \in \mathbb{R}^+ \forall x \in D \forall y \in D : (|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon).$$

Wie das obige Beispiel zeigt, ist nicht jede stetige Funktion gleichmäßig stetig. Dass diese Beispielfunktion unbeschränkt ist, ist nicht der entscheidende Punkt. Zum Beispiel wird eine beschränkte, stetige, aber nicht gleichmäßig stetige Funktion $f : (0, 1] \rightarrow \mathbb{R}$ erklärt durch

$$f(x) := \begin{cases} n(n+1)x - n & \text{für } x \in \left[\frac{1}{n+1}, \frac{1}{n}\right], n \in \mathbb{N} \text{ ungerade,} \\ -n(n+1)x + n + 1 & \text{für } x \in \left[\frac{1}{n+1}, \frac{1}{n}\right], n \in \mathbb{N} \text{ gerade.} \end{cases}$$

Wir wollen jetzt zeigen, dass so etwas nicht passieren kann, wenn der Definitionsbereich kompakt ist.

3.5 Satz. *Jede stetige Funktion mit kompaktem Definitionsbereich ist gleichmäßig stetig.*

Beweis. Sei D kompakt, $f : D \rightarrow \mathbb{R}$ stetig. Sei $\varepsilon \in \mathbb{R}^+$ gegeben. Zu jedem $a \in D$ existiert, da f in a stetig ist, ein $\delta(a) \in \mathbb{R}^+$ mit

$$|f(x) - f(a)| < \frac{\varepsilon}{2} \quad \text{für alle } x \in D \quad \text{mit } |x - a| < \delta(a).$$

Das System $\{U_{\frac{1}{2}\delta(a)}(a) \mid a \in D\}$ ist eine offene Überdeckung von D , enthält also nach Satz 1.8 eine endliche Teilüberdeckung $\{U_{\frac{1}{2}\delta(a_i)}(a_i) \mid i = 1, \dots, n\}$. Setze $\delta := \min\{\frac{1}{2}\delta(a_1), \dots, \frac{1}{2}\delta(a_n)\}$. Seien dann $x, y \in D$ Punkte mit $|x - y| < \delta$. Es gibt ein $i \in \{1, \dots, n\}$ mit $x \in U_{\frac{1}{2}\delta(a_i)}(a_i)$. Wegen $|a_i - x| < \frac{1}{2}\delta(a_i) < \delta(a_i)$ ist

$$|f(a_i) - f(x)| < \frac{\varepsilon}{2}.$$

Wegen $|a_i - y| \leq |a_i - x| + |x - y| \leq \frac{1}{2}\delta(a_i) + \delta \leq \delta(a_i)$ ist

$$|f(a_i) - f(y)| < \frac{\varepsilon}{2}.$$

Es folgt

$$|f(x) - f(y)| \leq |f(x) - f(a_i)| + |f(a_i) - f(y)| < \varepsilon.$$

Das angegebene δ leistet also das Gewünschte. ■

5 Spezielle Funktionen

In diesem Kapitel definieren und untersuchen wir einige spezielle Funktionen, die in der Analysis und ihren Anwendungen häufig verwendet werden.

5.1 Logarithmus und allgemeine Potenz

Die Exponentialfunktion $\exp : \mathbb{R} \rightarrow \mathbb{R}$ ist, wie in Abschnitt 3.3 gezeigt wurde, streng monoton wachsend. Ihre Umkehrfunktion existiert also, sie ist nach Satz 1.1 aus Kapitel 2 ebenfalls streng monoton wachsend und nach Satz 2.10 aus Kapitel 4 stetig. In Abschnitt 4.3 wurde gezeigt, dass \exp stetig ist. Aus der Reihendarstellung folgt

$$e^x \geq 1 + x \quad \text{für } x \geq 0$$

und daraus, ebenfalls für $x \geq 0$,

$$e^{-x} = \frac{1}{e^x} \leq \frac{1}{1+x},$$

also gilt $\lim_{x \rightarrow \infty} e^x = \infty$ und $\lim_{x \rightarrow -\infty} e^x = 0$. Aus dem Zwischenwertsatz 3.4 aus Kapitel 4 folgt, dass das Bild der Exponentialfunktion gleich \mathbb{R}^+ ist (beachte $e^x > 0$ für alle x). Wir können daher formulieren:

1.1 Definition. Die Umkehrfunktion der Exponentialfunktion ist auf \mathbb{R}^+ definiert, stetig und streng monoton wachsend; sie heißt (natürliche) Logarithmusfunktion und wird mit $\exp^{-1} = \ln$ bezeichnet.

1.2 Satz (Funktionalgleichung der \ln -Funktion). Für $x, y \in \mathbb{R}^+$ gilt

$$\ln xy = \ln x + \ln y.$$

Beweis. Setze $\ln x =: a$, $\ln y =: b$. Dann ist $x = e^a$, $y = e^b$, also nach Satz 3.3 aus Kapitel 3 $xy = e^a e^b = e^{a+b}$ und daher $\ln xy = a + b = \ln x + \ln y$. ■

Bemerkung. Für $x \in \mathbb{R}^+$ folgt

$$\ln \frac{1}{x} + \ln x = \ln \left(\frac{1}{x} \cdot x \right) = \ln 1 = 0$$

(wegen $e^0 = 1$), also

$$\ln \frac{1}{x} = -\ln x.$$

Definition. Für $a \in \mathbb{R}^+$ schreibt man

$$a^x := \exp_a x := e^{x \ln a} \quad \text{für } x \in \mathbb{R}.$$

Die Funktion $\exp_a : \mathbb{R} \rightarrow \mathbb{R}^+$ heißt *Exponentialfunktion* oder *Potenz zur Basis a* .

Für die allgemeine Potenz a^x stellen wir einige Rechenregeln zusammen:

1.3 Satz. Für $a, b \in \mathbb{R}^+$ und $x, y \in \mathbb{R}$ gilt

- (a) $a^x a^y = a^{x+y}$,
- (b) $(a^x)^y = a^{xy}$,
- (c) $a^x b^x = (ab)^x$,
- (d) $\left(\frac{1}{a}\right)^x = a^{-x} = \frac{1}{a^x}$.

Beweis. (a) $a^x a^y = e^{x \ln a} e^{y \ln a} = e^{(x+y) \ln a} = a^{x+y}$.

(b) Wegen $a^x = e^{x \ln a}$ ist $\ln a^x = x \ln a$, also $(a^x)^y = e^{y \ln a^x} = e^{yx \ln a} = a^{xy}$.

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

(d) $\left(\frac{1}{a}\right)^x = e^{x \ln \frac{1}{a}} = e^{-x \ln a} = a^{-x} = \frac{1}{e^{x \ln a}} = \frac{1}{a^x}$, wobei wir $\ln \frac{1}{a} = -\ln a$ und $e^{-y} = \frac{1}{e^y}$ benutzt haben. ■

Bemerkung. Für $a \in \mathbb{R}^+$ und $n \in \mathbb{N}$ schreibt man auch

$$a^{\frac{1}{n}} =: \sqrt[n]{a}$$

(n -te Wurzel aus a). Nach Satz 1.3 (b) ist

$$\left(\sqrt[n]{a}\right)^n = \left(a^{\frac{1}{n}}\right)^n = a^{\frac{1}{n} \cdot n} = a.$$

Die n -te Wurzel $\sqrt[n]{}$ ist also die Umkehrfunktion der n -ten Potenz $a \mapsto a^n$. (Wir haben die n -te Wurzel schon mehrfach benutzt, aber erst jetzt definiert.)

Offenbar ist auch die Exponentialfunktion zur Basis a stetig. Im Fall $a > 1$ ist sie (wegen $\ln a > 0$) streng monoton wachsend, im Fall $0 < a < 1$ ist

sie (wegen $\ln a < 0$) streng monoton fallend. Für $a \neq 1$ existiert also ihre Umkehrfunktion; sie heißt *Logarithmusfunktion zur Basis a* und wird mit \log_a bezeichnet. Sie unterscheidet sich nur um einen festen Faktor von der natürlichen Logarithmusfunktion: Setzen wir $\log_a x =: y$, so ist

$$x = a^y = e^{y \ln a},$$

also $\ln x = y \ln a$ und daher

$$\log_a x = \frac{\ln x}{\ln a}.$$

Es genügt daher, die natürliche Logarithmusfunktion (also die zur Basis e) zu kennen.

Bemerkung. Die Schreibweisen für die Logarithmusfunktionen sind in der Literatur nicht einheitlich. So wird für \ln gelegentlich auch \log geschrieben; oft ist mit \log auch \log_{10} gemeint.

Einige Grenzwerte

Wir wollen einige mit der Exponential- und Logarithmusfunktion gebildete Grenzwerte bestimmen, die häufig auftreten.

1.4 Behauptung. $\forall a \in \mathbb{R}^+ : \lim_{n \rightarrow \infty} \sqrt[n]{a} = 1.$

Beweis. $\sqrt[n]{a} = e^{\frac{1}{n} \ln a} \rightarrow e^0 = 1$ für $n \rightarrow \infty$ wegen der Stetigkeit der e -Funktion und Satz 2.6 aus Kapitel 4. ■

1.5 Behauptung. $\forall k \in \mathbb{N} : \lim_{x \rightarrow \infty} \frac{e^x}{x^k} = \infty.$

Beweis. Für alle $x > 0$ gilt

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} > \frac{x^{k+1}}{(k+1)!},$$

woraus die Behauptung folgt. ■

1.6 Behauptung. $\forall k \in \mathbb{N} : \lim_{x \rightarrow \infty} x^k e^{-x} = 0.$

Beweis. $x^k e^{-x} = \left(\frac{e^x}{x^k}\right)^{-1}$ und Behauptung 1.5. ■

1.7 Behauptung. $\forall k \in \mathbb{N} : \lim_{x \searrow 0} x^k e^{\frac{1}{x}} = \infty.$

Beweis.

$$\lim_{x \searrow 0} x^k e^{\frac{1}{x}} = \lim_{y \rightarrow \infty} \left(\frac{1}{y}\right)^k e^y = \lim_{y \rightarrow \infty} \frac{e^y}{y^k} = \infty. \quad \blacksquare$$

1.8 Behauptung. $\lim_{x \rightarrow \infty} \ln x = \infty$, $\lim_{x \rightarrow 0} \ln x = -\infty$.

Beweis. Zu gegebenem $c \in \mathbb{R}$ setze $c_0 := e^c$. Für alle $x \geq c_0$ gilt dann $\ln x \geq c$. Also ist $\lim_{x \rightarrow \infty} \ln x = \infty$. Ferner gilt

$$\lim_{x \rightarrow 0} \ln x = \lim_{y \rightarrow \infty} \ln \frac{1}{y} = - \lim_{y \rightarrow \infty} \ln y = -\infty. \quad \blacksquare$$

1.9 Behauptung. $\forall \alpha \in \mathbb{R}^+ : \lim_{x \searrow 0} x^\alpha = 0$.

Beweis. Zu gegebenem $\varepsilon \in \mathbb{R}^+$ existiert wegen Behauptung 1.8 ein $\delta \in \mathbb{R}^+$ mit $\alpha \ln x < \ln \varepsilon$ für $0 < x < \delta$. Aus $0 < x < \delta$ folgt dann

$$0 < x^\alpha = e^{\alpha \ln x} < e^{\ln \varepsilon} = \varepsilon. \quad \blacksquare$$

1.10 Behauptung. $\forall \alpha \in \mathbb{R}^+ : \lim_{x \rightarrow \infty} \frac{\ln x}{x^\alpha} = 0$.

Beweis. Sei $\varepsilon \in \mathbb{R}^+$ gegeben. Wegen Behauptung 1.5 existiert ein $c_0 \in \mathbb{R}$ mit

$$0 < \frac{1}{\alpha} \frac{y}{e^y} < \varepsilon \quad \text{für } y > c_0.$$

Wegen Behauptung 1.8 existiert ein $c \in \mathbb{R}$ mit $\alpha \ln x > c_0$ für $x > c$. Für alle $x > c$ gilt also

$$0 < \frac{\ln x}{x^\alpha} = \frac{1}{\alpha} \frac{\alpha \ln x}{e^{\alpha \ln x}} < \varepsilon. \quad \blacksquare$$

1.11 Behauptung. $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$.

Beweis. $\sqrt[n]{n} = e^{\frac{1}{n} \ln n} \rightarrow e^0 = 1$ für $n \rightarrow \infty$ wegen Behauptung 1.10. \blacksquare

5.2 Die Exponentialfunktion im Komplexen

Um bequem die trigonometrischen Funktionen einführen und untersuchen zu können, ist es zweckmäßig, die Exponentialfunktion auch für komplexe Argumente zur Verfügung zu haben. Wir werden daher jetzt die hierzu erforderlichen Grundbegriffe und Sätze der Analysis unter Zugrundelegung des Körpers der komplexen Zahlen entwickeln.

Ein Anlaß zur Einführung der komplexen Zahlen ist zum Beispiel die Tatsache, dass im Bereich der reellen Zahlen nicht jede „quadratische Gleichung“ lösbar ist. So gibt es etwa keine reelle Zahl x mit $x^2 = -1$ (denn es ist $-1 < 0$, und nach Behauptung 2.9 aus Kapitel 1 gilt $x^2 \geq 0$ für alle $x \in \mathbb{R}$). Man erweitert daher den Bereich der reellen Zahlen in geeigneter Weise, was man formal folgendermaßen tun kann.

\mathbb{C} sei die Menge $\mathbb{R} \times \mathbb{R}$ der geordneten Paare reeller Zahlen. Addition und Multiplikation der Elemente von \mathbb{C} werden erklärt durch

$$\begin{aligned}(a, b) + (c, d) &:= (a + c, b + d) \\ (a, b) \cdot (c, d) &:= (ac - bd, ad + bc)\end{aligned}$$

für $(a, b), (c, d) \in \mathbb{C}$. Das Tripel $(\mathbb{C}, +, \cdot)$ erfüllt die in Abschnitt 1.1 zusammengestellten Körperaxiome, wie man durch elementare Rechnungen leicht nachweist. Man bezeichnet $(\mathbb{C}, +, \cdot)$ oder kurz \mathbb{C} als den *Körper der komplexen Zahlen* und die Elemente als komplexe Zahlen.

Wegen

$$\begin{aligned}(a, 0) + (c, 0) &= (a + c, 0) \\ (a, 0) \cdot (c, 0) &= (ac, 0)\end{aligned}$$

kann man die Teilmenge $\{(a, 0) \mid a \in \mathbb{R}\}$ mit den darauf eingeschränkten Verknüpfungen identifizieren mit dem Körper der reellen Zahlen. Statt $(a, 0)$ schreibt man daher a . Ferner setzt man zur Abkürzung

$$(0, 1) =: i$$

und nennt diese komplexe Zahl die *imaginäre Einheit*. Dann ist $i^2 = -1$.

Mit diesen Vereinbarungen ist jetzt für $(a, b) \in \mathbb{C}$

$$(a, b) = (a, 0)(1, 0) + (b, 0)(0, 1) = a + bi.$$

Jede komplexe Zahl $z \in \mathbb{C}$ kann also in der Form $z = a + ib$ mit $a, b \in \mathbb{R}$ dargestellt werden, und zwar eindeutig, denn aus $a + ib = a' + ib'$ mit $a, b, a', b' \in \mathbb{R}$, also $(a, b) = (a', b')$, folgt $a = a'$ und $b = b'$ nach der Gleichheitsdefinition für geordnete Paare.

Definition. Sei $z \in \mathbb{C}$, $z = a + ib$ mit $a, b \in \mathbb{R}$. Dann heißt

$$\begin{aligned}\operatorname{Re}(z) &:= a && \text{der Realteil und} \\ \operatorname{Im}(z) &:= b && \text{der Imaginärteil}\end{aligned}$$

von z . Die komplexe Zahl

$$\bar{z} := a - ib$$

heißt die zu z *konjugiert-komplexe* Zahl.

2.1 Behauptung. Für $z, z_1, z_2 \in \mathbb{C}$ gilt

- (a) $\operatorname{Re}(z) = \frac{1}{2}(z + \bar{z})$, $\operatorname{Im}(z) = \frac{1}{2i}(z - \bar{z})$,
- (b) $\overline{\bar{z}} = z$
- (c) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$
- (d) $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$

Beweis. trivial (nachrechnen) ■

Nachdem wir den Körper der reellen Zahlen ausgedehnt haben zum Körper der komplexen Zahlen, stellt sich die Frage, ob auch die Beziehung $<$ von den reellen Zahlen auf die komplexen Zahlen fortgesetzt werden kann. Dabei würde man natürlich fordern, dass die Anordnungsaxiome aus Abschnitt 1.2, die die Verträglichkeit dieser Kleinerbeziehung mit den Körperverknüpfungen regeln, erfüllt sind. Dies ist aber nicht möglich, da in einem angeordneten Körper nach Behauptung 2.9 aus Kapitel 1 für jedes Körperelement a die Beziehung $a^2 \geq 0$ gilt, während doch $i^2 = -1 < 0$ ist.

Eine eingehendere Durchsicht der Konvergenzbetrachtungen in den vorhergehenden Abschnitten zeigt nun aber, dass dabei gar nicht so sehr von der Größerbeziehung Gebrauch gemacht wurde, sondern vorwiegend von den daraus abgeleiteten Eigenschaften des Absolutbetrages. Ein Absolutbetrag mit ähnlichen Eigenschaften läßt sich nun auch für komplexe Zahlen erklären, und dies hat zur Folge, dass man auch im Bereich der komplexen Zahlen eine weitgehend analoge Konvergenztheorie aufbauen kann.

Definition. Für $z = a + ib$ ($a, b \in \mathbb{R}$) heißt

$$|z| := \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$$

der Betrag von z .

Der Betrag einer komplexen Zahl ist also eine nichtnegative reelle Zahl. Für $z \in \mathbb{R}$ stimmt diese Definition mit der früher für den Absolutbetrag gegebenen überein. Für $z \in \mathbb{C}$ ist $|\bar{z}| = |z|$, ferner $|\operatorname{Re}(z)| \leq |z|$, $|\operatorname{Im}(z)| \leq |z|$.

2.2 Satz. Für $z, z_1, z_2 \in \mathbb{C}$ gilt

- (a) $|z| \geq 0$; $|z| = 0 \Leftrightarrow z = 0$,
- (b) $|z_1 z_2| = |z_1| |z_2|$,
- (c) $|z_1 + z_2| \leq |z_1| + |z_2|$ (Dreiecksungleichung)

Beweis. (a) ist trivial.

(b) $|z_1 z_2|^2 = (z_1 z_2) \overline{(z_1 z_2)} = z_1 z_2 \bar{z}_1 \bar{z}_2 = (z_1 \bar{z}_1)(z_2 \bar{z}_2) = |z_1|^2 |z_2|^2$.

(c) Unter Verwendung von

$$\operatorname{Re}(z_1 \bar{z}_2) \leq |z_1 \bar{z}_2| = |z_1| |\bar{z}_2| = |z_1| |z_2|$$

folgt

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)(\bar{z}_1 + \bar{z}_2) = z_1 \bar{z}_1 + z_1 \bar{z}_2 + z_2 \bar{z}_1 + z_2 \bar{z}_2 \\ &= |z_1|^2 + 2\operatorname{Re}(z_1 \bar{z}_2) + |z_2|^2 \\ &\leq |z_1|^2 + 2|z_1| |z_2| + |z_2|^2 = (|z_1| + |z_2|)^2. \quad \blacksquare \end{aligned}$$

Wir sind nun in der Lage, die Konvergenzbetrachtungen der vorhergehenden Paragraphen weitgehend vom Reellen ins Komplexe zu übertragen. Alle Behauptungen und Beweise, die nicht von der Anordnung der reellen Zahlen, sondern (außer von den Körperaxiomen) nur vom Absolutbetrag und seinen Eigenschaften Gebrauch machten, bleiben ohne Änderungen gültig. Die Resultate, deren Beweise sich wörtlich übertragen lassen, werden wir nur auflisten, ohne die Begründungen zu wiederholen.

Definition. Sei $(z_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{C} und $z \in \mathbb{C}$. Die Folge (z_n) heißt *konvergent* gegen z , und z heißt *Grenzwert* dieser Folge, geschrieben $\lim_{n \rightarrow \infty} z_n = z$, wenn gilt

$$\forall \varepsilon \in \mathbb{R}^+ \exists n_0 \in \mathbb{N} \forall n \in \mathbb{N} : (n \geq n_0 \Rightarrow |z_n - z| < \varepsilon).$$

Die Sätze 1.1 (Eindeutigkeit des Grenzwertes), 1.2 (Beschränktheit konvergenter Folgen; $(z_n)_{n \in \mathbb{N}}$ heißt *beschränkt*, wenn $c \in \mathbb{R}$ existiert mit $|z_n| \leq c \forall n \in \mathbb{N}$), 1.3 (Rechenregeln für Grenzwerte) und ihre Beweise aus Kapitel 3 lassen sich wörtlich übertragen.

Andere Aussagen von früher lassen sich nicht verallgemeinern, weil sie in \mathbb{C} sinnlos sind. Zum Beispiel kann man nicht von einer monotonen Folge komplexer Zahlen sprechen.

Der folgende Satz gestattet es, Konvergenz im Komplexen zurückzuführen auf Konvergenz im Reellen:

2.3 Satz. Für jede Folge $(z_n)_{n \in \mathbb{N}}$ in \mathbb{C} gilt:

$$(z_n)_{n \in \mathbb{N}} \text{ konvergent} \Leftrightarrow (\operatorname{Re}(z_n))_{n \in \mathbb{N}} \text{ und } (\operatorname{Im}(z_n))_{n \in \mathbb{N}} \text{ konvergent.}$$

Ist $(z_n)_{n \in \mathbb{N}}$ konvergent, so gilt

$$\lim_{n \rightarrow \infty} z_n = \lim_{n \rightarrow \infty} \operatorname{Re}(z_n) + i \lim_{n \rightarrow \infty} \operatorname{Im}(z_n).$$

Beweis. Unter Verwendung der Abschätzungen

$$|a|, |b| \leq |a + ib| \leq |a| + |b|$$

ergibt sich dies folgendermaßen. Setze $z_n = a_n + ib_n, z = a + ib$.

„ \Rightarrow “: Sei $\lim_{n \rightarrow \infty} z_n = z$. Sei $\varepsilon \in \mathbb{R}^+$. Es gibt ein $n_0 \in \mathbb{N}$ mit $|z_n - z| < \varepsilon$ für $n \geq n_0$. Für $n \geq n_0$ gilt also $|a_n - a| \leq |z_n - z| < \varepsilon$. Somit gilt $\lim_{n \rightarrow \infty} a_n = a$. Analog folgt $\lim_{n \rightarrow \infty} b_n = b$.

„ \Leftarrow “: Gelte $\lim a_n = a$ und $\lim b_n = b$. Sei $\varepsilon \in \mathbb{R}^+$. Es gibt ein $n_1 \in \mathbb{N}$ mit $|a_n - a| < \frac{\varepsilon}{2}$ für $n \geq n_1$ und ein $n_2 \in \mathbb{N}$ mit $|b_n - b| < \frac{\varepsilon}{2}$ für $n \geq n_2$. Für $n \geq n_0 := \max\{n_1, n_2\}$ gilt also $|z_n - z| \leq |a_n - a| + |b_n - b| < \varepsilon$ mit $z = a + ib$. Daraus folgt die Behauptung. ■

Definition. $(z_n)_{n \in \mathbb{N}}$ heißt *Cauchy-Folge*

$$\Leftrightarrow \forall \varepsilon \in \mathbb{R}^+ \exists n_0 \in \mathbb{N} \forall m, n \geq n_0 : |z_m - z_n| < \varepsilon.$$

2.4 Satz. Die Folge $(z_n)_{n \in \mathbb{N}}$ in \mathbb{C} ist genau dann eine Cauchy-Folge, wenn $(\operatorname{Re}(z_n))_{n \in \mathbb{N}}$ und $(\operatorname{Im}(z_n))_{n \in \mathbb{N}}$ Cauchy-Folgen sind.

Beweis. Analog zum Beweis von Satz 2.3. ■

Aus den beiden vorhergehenden Sätzen und dem Cauchy-Kriterium 1.8 aus Kapitel 3 im Reellen folgt jetzt sofort:

2.5 Satz. Eine Folge in \mathbb{C} ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.

Reihen komplexer Zahlen und ihre Konvergenz werden genau wie im Reellen erklärt. So bezeichnet $\sum_{k=1}^{\infty} z_k$ sowohl die Folge $(s_n)_{n \in \mathbb{N}}$ der Partialsummen $s_n = \sum_{k=1}^n z_k$ als auch, wenn die Reihe konvergiert, ihren Grenzwert, und $\sum_{k=1}^{\infty} z_k = z$ bedeutet $\lim_{n \rightarrow \infty} \sum_{k=1}^n z_k = z$.

Definition. Die Reihe $\sum_{k=1}^{\infty} z_k$ heißt *absolut konvergent*, wenn die Reihe

$$\sum_{k=1}^{\infty} |z_k|$$

konvergiert.

Die folgenden Sätze über Konvergenz und ihre Beweise aus Kapitel 3 lassen sich nun wörtlich übertragen.

Satz 2.3 Konvergenzkriterium von Cauchy für Reihen

Satz 2.6 Absolute Konvergenz \Rightarrow Konvergenz

Satz 2.7 Umordnungssatz

Satz 2.8 Majorantenkriterium

Satz 2.9 Quotientenkriterium

Satz 2.10 Wurzelkriterium

Satz 2.11 Cauchy-Produkt

Nun können wir durch wörtliche Übertragung der Beweise die Betrachtungen aus Abschnitt 3.3 über die Exponentialreihe ins Komplexe ausdehnen.

2.6 Satz. Für jedes $z \in \mathbb{C}$ ist die Reihe

$$\sum_{n=0}^{\infty} \frac{z^n}{n!}$$

absolut konvergent.

Beweis wie für Satz 3.1 aus Kapitel 3.

Definition. $\exp(z) := e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!}$ für $z \in \mathbb{C}$.

2.7 Satz. Für alle $z, w \in \mathbb{C}$ gilt $e^{z+w} = e^z e^w$.

Beweis wie für Satz 3.3 aus Kapitel 3.

Folgerung. $e^z \neq 0$ für alle $z \in \mathbb{C}$, denn $e^z e^{-z} = 1$.

2.8 Behauptung. $e^{\bar{z}} = \overline{e^z}$ für $z \in \mathbb{C}$.

Beweis. Mit $s_n(z) := \sum_{k=0}^n \frac{z^k}{k!}$ gilt nach Behauptung 2.1

$$s_n(\bar{z}) = \sum_{k=0}^n \frac{\bar{z}^k}{k!} = \overline{\sum_{k=0}^n \frac{z^k}{k!}} = \overline{s_n(z)},$$

also wegen Satz 2.3

$$\begin{aligned} \overline{e^z} &= \overline{\lim \operatorname{Re}(s_n(z)) + i \lim \operatorname{Im}(s_n(z))} \\ &= \lim \operatorname{Re}(s_n(z)) - i \lim \operatorname{Im}(s_n(z)) \\ &= \lim(\overline{s_n(z)}) = \lim s_n(\bar{z}) = e^{\bar{z}}. \end{aligned}$$

■

Folgerung. Ist $z \in \mathbb{C}$ rein imaginär, d.h. $z = ix$ mit $x \in \mathbb{R}$, so ist $|e^z| = 1$.
Denn für $x \in \mathbb{R}$ gilt $|e^{ix}|^2 = e^{ix} \overline{e^{ix}} = e^{ix} e^{-ix} = e^{ix} e^{-ix} = e^0 = 1$.

5.3 Die trigonometrischen Funktionen

Das Ziel unseres Ausfluges ins Komplexe war es, jetzt in bequemer Weise die Funktionen Sinus und Cosinus einführen zu können.

Definition. Für $x \in \mathbb{R}$ setzt man

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

Es gilt also

$$\begin{aligned} e^{ix} &= \cos x + i \sin x, \\ \cos x &= \frac{1}{2}(e^{ix} + e^{-ix}), \\ \sin x &= \frac{1}{2i}(e^{ix} - e^{-ix}). \end{aligned}$$

Daraus folgt sofort für $x \in \mathbb{R}$

$$\begin{aligned} \cos(-x) &= \cos x, \\ \sin(-x) &= -\sin x, \\ \sin^2 x + \cos^2 x &= |e^{ix}|^2 = 1. \end{aligned}$$

Aus der Funktionalgleichung der Exponentialfunktion ergeben sich die Additionstheoreme der „trigonometrischen Funktionen“ sin und cos:

3.1 Satz (Additionstheoreme). *Für $x, y \in \mathbb{R}$ gilt*

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

Beweis. Es ist

$$\begin{aligned} \cos(x + y) + i \sin(x + y) &= e^{i(x+y)} = e^{ix} e^{iy} \\ &= (\cos x + i \sin x)(\cos y + i \sin y) \\ &= (\cos x \cos y - \sin x \sin y) + i(\sin x \cos y + \cos x \sin y). \end{aligned}$$

Vergleich von Real- und Imaginärteil ergibt die Behauptung. ■

3.2 Satz. Für $x \in \mathbb{R}$ gilt

$$\begin{aligned}\cos x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} \dots, \\ \sin x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} \dots\end{aligned}$$

Beide Reihen sind absolut konvergent.

Beweis. Die absolute Konvergenz folgt nach dem Majorantenkriterium aus der absoluten Konvergenz der Exponentialreihe. Es ist

$$\begin{aligned}\cos x + i \sin x &= e^{ix} = \sum_{n=0}^{\infty} \frac{i^n x^n}{n!} \\ &= \frac{1}{0!} + i \frac{x}{1!} - \frac{x^2}{2!} - i \frac{x^3}{3!} + \frac{x^4}{4!} + i \frac{x^5}{5!} - \frac{x^6}{6!} - i \frac{x^7}{7!} + \dots \\ &= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \right) + i \left(\frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \right),\end{aligned}$$

wobei Satz 2.3 benutzt wurde. Vergleich von Real- und Imaginärteil ergibt die Behauptung. ■

3.3 Satz. Für die durch

$$\begin{aligned}\cos x &= \sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!} + R_{2n+2}(x), \\ \sin x &= \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!} + R_{2n+3}(x)\end{aligned}$$

definierten Restglieder gilt die Abschätzung

$$|R_m(x)| \leq \frac{|x|^m}{m!} \quad \text{für } |x| \leq m + 1.$$

Beweis.

$$\begin{aligned}|R_{2n+2}(x)| &= \frac{|x|^{2n+2}}{(2n+2)!} \left| 1 - \frac{x^2}{(2n+3)(2n+4)} + \frac{x^4}{(2n+3)\dots(2n+6)} - \dots \right| \\ &= \frac{|x|^{2n+2}}{(2n+2)!} |1 - a_1 + a_2 - \dots|\end{aligned}$$

mit

$$a_k := \frac{x^{2k}}{(2n+3)\cdots(2n+2k+2)}.$$

Wegen

$$a_k = a_{k-1} \frac{x^2}{(2n+2k+1)(2n+2k+2)}$$

gilt für $|x| \leq 2n+3$ die Ungleichungskette $1 \geq a_1 \geq a_2 \geq \dots$, also $0 \leq 1 - a_1 + a_2 - a_3 + \dots \leq 1$. Analog schließt man für $R_{2n+3}(x)$. ■

Als Folgerung notieren wir:

3.4 Behauptung.

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1.$$

Beweis. Nach Satz 3.3 ist $|\sin x - x| \leq \frac{|x|^3}{6}$ für $|x| \leq 4$, woraus nach Division durch x die Behauptung folgt. ■

3.5 Satz. \cos und \sin sind stetig.

Beweis. Satz 3.2 aus Kapitel 3 gilt mit demselben Beweis auch für $x \in \mathbb{C}$. Für $z, w \in \mathbb{C}$ ist also $|e^z - e^w| \leq 2|e^w||z - w|$ für $|z| \leq 1$. Für $x, a \in \mathbb{R}$ mit $|x - a| \leq 1$ folgt also

$$|\cos x - \cos a| = |\operatorname{Re}(e^{ix} - e^{ia})| \leq |e^{ix} - e^{ia}| \leq 2|x - a|,$$

woraus die Stetigkeit von \cos folgt. Analog für \sin . ■

Die Zahl π

Die Zahl π , klassisch definiert als das Verhältnis des Kreisumfangs zum Kreisdurchmesser, werden wir hier auf ganz andere Weise einführen. Erst wesentlich später ergibt sich der Zusammenhang mit dem Kreisumfang. Die nachfolgenden Behauptungen dienen zur Vorbereitung.

3.6 Behauptung. $\sin x > 0$ für $x \in (0, 2)$.

Beweis. Nach Satz 3.3 ist

$$\sin x = x \left(1 + \frac{R_3(x)}{x} \right)$$

mit

$$\left| \frac{R_3(x)}{x} \right| \leq \frac{|x|^2}{6} \quad \text{für } |x| \leq 4.$$

Für $x \in (0, 2]$ gilt also $\left| \frac{R_3(x)}{x} \right| \leq \frac{2}{3}$ und daher

$$\sin x = x \left(1 + \frac{R_3(x)}{x} \right) \geq x \cdot \frac{1}{3} > 0. \quad \blacksquare$$

3.7 Behauptung. *In $[0, 2]$ ist \cos streng monoton fallend.*

Beweis. Mit $\frac{x+y}{2} =: u$, $\frac{x-y}{2} =: v$ folgt aus Satz 3.1

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

also

$$-\cos x + \cos y = 2 \sin \frac{x+y}{2} \sin \frac{x-y}{2}.$$

Für $0 \leq y < x \leq 2$ ist nach Behauptung 3.6 die rechte Seite positiv. \blacksquare

3.8 Behauptung. $\cos 2 < 0$.

Beweis. Nach Satz 3.3 ist

$$\cos x = 1 - \frac{x^2}{2} + R_4(x) \quad \text{mit } |R_4(x)| \leq \frac{|x|^4}{4!} \quad \text{für } |x| \leq 5,$$

also insbesondere

$$\cos 2 = 1 - 2 + R_4(2) \quad \text{mit } |R_4(2)| \leq \frac{16}{24} = \frac{2}{3},$$

woraus die Behauptung folgt. \blacksquare

Über die Funktion $f := \cos|_{[0, 2]}$ wissen wir nun: f ist stetig, $f(0) = 1$, $f(2) < 0$. Nach dem Zwischenwertsatz 3.4 aus Kapitel 4 existiert ein $x_0 \in (0, 2)$ mit $f(x_0) = 0$. Wegen 3.7 gibt es nur ein solches x . Wir können also definieren:

Definition. Die eindeutig bestimmte Nullstelle der Funktion \cos im Intervall $(0, 2)$ wird mit $\frac{\pi}{2}$ bezeichnet.

Näherungsweise Berechnung ergibt

$$\pi = 3,141592653 \pm 10^{-9}.$$

Aus $\cos \frac{\pi}{2} = 0$ ergeben sich einige spezielle Werte für die Funktionen \sin , \cos , \exp : Es ist

$$\sin^2 \frac{\pi}{2} = 1 - \cos^2 \frac{\pi}{2} = 1,$$

wegen Behauptung 3.6 also

$$\sin \frac{\pi}{2} = 1,$$

folglich

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

Daraus folgt $e^{\pi i} = -1$, $e^{2\pi i} = 1$, $e^{\frac{3}{2}\pi i} = -i$ und hieraus

x	0	$\pi/2$	π	$3\pi/2$	2π
$\sin x$	0	1	0	-1	0
$\cos x$	1	0	-1	0	1

Zusammen mit den Additionstheoremen 3.1 liefert das:

3.9 Satz. Für alle $x \in \mathbb{R}$ gilt

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

Insbesondere ist

$$\cos\left(\frac{\pi}{2} - x\right) = \sin x, \quad \sin\left(\frac{\pi}{2} - x\right) = \cos x.$$

3.10 Satz.

$$\begin{aligned} \{x \in \mathbb{R} \mid \sin x = 0\} &= \{k\pi \mid k \in \mathbb{Z}\}, \\ \{x \in \mathbb{R} \mid \cos x = 0\} &= \left\{ \left(k + \frac{1}{2}\right)\pi \mid k \in \mathbb{Z} \right\}. \end{aligned}$$

Beweis. $\sin k\pi = 0$ für $k \in \mathbb{Z}$ ist klar. Sei jetzt $x \in \mathbb{R}$ eine Zahl mit $\sin x = 0$. Wir können $x = k\pi + x_1$ mit $k \in \mathbb{Z}$ und $x_1 \in [0, \pi)$ schreiben. Dann ist

$$\sin x_1 = \sin(x - k\pi) = \sin x \cos k\pi - \cos x \sin k\pi = 0.$$

Nun gilt nach Behauptung 3.6: $\sin x_1 > 0$ für $x_1 \in (0, \frac{\pi}{2})$.

Für $x_1 \in [\frac{\pi}{2}, \pi)$ ist $\sin x_1 = \cos(\frac{\pi}{2} - x_1) = \cos(x_1 - \frac{\pi}{2}) > 0$ wegen $x_1 - \frac{\pi}{2} \in [0, \frac{\pi}{2})$ und der Definition von $\pi/2$. Also muß $x_1 = 0$ sein, d.h. $x = k\pi$.

Wegen $\cos x = \sin(x + \frac{\pi}{2})$ folgt die zweite Gleichheit aus der ersten. ■

Nachdem wir jetzt die Nullstellen von \sin und \cos kennen, können wir definieren:

Definition.

$$\begin{aligned} \text{Tangens :} \quad \tan x &:= \frac{\sin x}{\cos x} && \text{für } x \in \mathbb{R} \setminus \left\{ \left(k + \frac{1}{2}\right)\pi \mid k \in \mathbb{Z} \right\} \\ \text{Cotangens :} \quad \cot x &:= \frac{\cos x}{\sin x} && \text{für } x \in \mathbb{R} \setminus \{k\pi \mid k \in \mathbb{Z}\}. \end{aligned}$$

Weitere wichtige spezielle Funktionen ergeben sich als Umkehrfunktionen der bisher eingeführten Funktionen. Da die trigonometrischen Funktionen aber nicht injektiv sind, existieren Umkehrfunktionen nur für geeignete Einschränkungen.

3.11 Satz. (a) Die Einschränkung $\cos|_{[0, \pi]}$ ist streng monoton fallend und eine bijektive Abbildung auf $[-1, 1]$. Die Umkehrfunktion wird mit

$$\arccos : [-1, 1] \rightarrow \mathbb{R} \quad (\text{Arcus Cosinus})$$

bezeichnet.

(b) $\sin|_{[-\frac{\pi}{2}, \frac{\pi}{2}]}$ ist streng monoton wachsend und bijektiv auf $[-1, 1]$. Die Umkehrfunktion wird mit

$$\arcsin : [-1, 1] \rightarrow \mathbb{R} \quad (\text{Arcus Sinus})$$

bezeichnet.

(c) $\tan|_{(-\frac{\pi}{2}, \frac{\pi}{2})}$ ist streng monoton wachsend und bijektiv auf \mathbb{R} . Die Umkehrfunktion wird mit

$$\arctan : \mathbb{R} \rightarrow \mathbb{R} \quad (\text{Arcus Tangens})$$

bezeichnet.

(d) $\cot|_{(0, \pi)}$ ist streng monoton fallend und bijektiv auf \mathbb{R} . Die Umkehrfunktion wird mit

$$\operatorname{arccot} : \mathbb{R} \rightarrow \mathbb{R} \quad (\text{Arcus Cotangens})$$

bezeichnet.

Wir verzichten auf die Beweise, da sie mit Hilfe der bisher bekannten Ergebnisse sehr leicht zu erbringen sind.

Als Anwendung leiten wir eine besondere Darstellung der komplexen Zahlen her:

3.12 Satz. *Jede komplexe Zahl $z \in \mathbb{C}$ ist darstellbar in der Form*

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

mit $r = |z|$ und einer reellen Zahl $\varphi \in [0, 2\pi)$ (genannt das Argument von z). Für $z \neq 0$ ist φ eindeutig bestimmt.

Beweis. Für $z = 0$ ist $z = 0 \cdot e^{i\varphi}$ mit beliebigem φ . Sei $z \neq 0$, $r := |z|$ und $\frac{z}{r} = x + iy$ mit $x, y \in \mathbb{R}$. Wegen $|\frac{z}{r}| = 1$ ist $x^2 + y^2 = 1$, also $x \in [-1, 1]$. Daher ist $\alpha := \arccos x$ definiert. Wegen $\cos \alpha = x$ ist $\sin^2 \alpha = 1 - \cos^2 \alpha = 1 - x^2 = y^2$, also $\sin \alpha = \pm y$. Setze

$$\varphi := \begin{cases} \alpha, & \text{falls } \sin \alpha = y, \\ 2\pi - \alpha, & \text{falls } \sin \alpha = -y \text{ und } \alpha \neq 0. \end{cases}$$

Dann gilt $\sin \varphi = y$ und $\cos \varphi = \cos \alpha = x$, also

$$e^{i\varphi} = \cos \varphi + i \sin \varphi = x + iy = \frac{z}{|z|}.$$

Aus $e^{i\varphi} = e^{i\psi}$ folgt $e^{i(\varphi-\psi)} = 1$, also $\cos(\varphi - \psi) = 1$ und $\sin(\varphi - \psi) = 0$. Ist o.B.d.A. $\varphi \geq \psi$, so folgt aus Satz 3.10 $\varphi - \psi = 0$ oder π , wegen $\cos \pi = -1$ also $\varphi - \psi = 0$. ■

6 Differenzierbare Funktionen

6.1 Die Ableitung

Ein nützliches allgemeines Prinzip bei der Untersuchung komplizierter Funktionen besteht darin, diese durch einfache, besser bekannte Funktionen zu approximieren (d.h. anzunähern). Der Begriff der Approximation kann auf verschiedene Weisen verstanden werden. Im folgenden interessieren wir uns für Approximation durch affine Funktionen in der Nähe eines festen Punktes. Dieses Verfahren ist heutzutage schon aus dem täglichen Leben geläufig, zum Beispiel wenn man von „Geschwindigkeit“ spricht. Bei einer gleichförmigen Bewegung, das heißt einer solchen, bei der der zurückgelegte Weg proportional zur Zeit ist, versteht man unter der Geschwindigkeit das Verhältnis von Weg zu Zeit. Bei einer nicht gleichförmigen Bewegung kann man nur von einer momentanen Geschwindigkeit reden, und man versteht hierunter die Geschwindigkeit der gleichförmigen Bewegung, die der betrachteten im gegebenen Zeitpunkt besonders nahe kommt. dass es eine solche gibt, läuft mathematisch auf die Forderung der Differenzierbarkeit hinaus. Eine Funktion soll im Punkt x differenzierbar heißen, wenn sie in x von erster Ordnung durch eine affine Funktion approximiert werden kann, d.h. genauer:

Definition. Die Funktion $f : D \rightarrow \mathbb{R}$ heißt *differenzierbar in $x \in D$* , wenn x Häufungspunkt von D ist und es eine Zahl $c \in \mathbb{R}$ gibt mit

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x) - ch}{h} = 0. \quad (1.1)$$

Gibt es eine solche Zahl c , so ist sie eindeutig bestimmt; sie wird mit $f'(x)$ bezeichnet und heißt *Ableitung von f an der Stelle x* .

Die Gleichung (1.1) ist äquivalent mit

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = c.$$

Damit ist auch die Eindeutigkeit von c klar. Die Ableitung von f an der Stelle x ist also definiert durch

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{y \rightarrow x} \frac{f(y) - f(x)}{y - x},$$

falls dieser Grenzwert existiert.

Bemerkung. Der letzte Limes ergibt die bekannte geometrische Interpretation der Ableitung als „Steigung“ der „Tangente“ an den Graphen („Limes“ von Sekanten) von f .

Bezeichnungen. Üblich (aber anfechtbar) ist auch die Schreibweise

$$f'(x) = \frac{df}{dx}.$$

Besser wäre dann schon die Schreibweise

$$f'(x) = \left. \frac{df(y)}{dy} \right|_{y=x}.$$

1.2 Satz. Ist $f : D \rightarrow \mathbb{R}$ in x differenzierbar, so ist f in x stetig.

Beweis. Aus

$$\lim_{y \rightarrow x} \frac{f(y) - f(x)}{y - x} = f'(x)$$

und $\lim_{y \rightarrow x} (y - x) = 0$ folgt $\lim_{y \rightarrow x} [f(y) - f(x)] = f'(x) \cdot 0$. Also ist $\lim_{y \rightarrow x} f(y) = f(x)$. Nach Satz 2.6 aus Kapitel 4 ist f stetig in x . ■

Bemerkung. Es gibt stetige Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$, die in keinem Punkt differenzierbar sind.

Definition. $f : D \rightarrow \mathbb{R}$ heißt *differenzierbar*, wenn f in jedem $x \in D$ differenzierbar ist. In diesem Fall heißt die Funktion $f' : D \rightarrow \mathbb{R}$, $x \mapsto f'(x)$, die *Ableitung* von f .

Beispiele. (1) Sei $c \in \mathbb{R}$ und $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto c$. Wegen

$$\frac{f(x+h) - f(x)}{h} = 0$$

ist f differenzierbar und $f'(x) = 0$ für alle $x \in \mathbb{R}$.

(2) Sei $c \in \mathbb{R}$ und $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto cx$. Wegen

$$\frac{f(x+h) - f(x)}{h} = c$$

ist f differenzierbar und $f'(x) = c$ für alle $x \in \mathbb{R}$.

(3) exp. Es ist

$$\frac{e^{x+h} - e^x}{h} = e^x \frac{e^h - 1}{h}.$$

Nach Satz 3.2 aus Kapitel 3 ist

$$|e^h - 1 - h| = |R_2(h)| \leq 2 \frac{|h|^2}{2!} \quad \text{für } |h| \leq \frac{3}{2},$$

also

$$\lim_{h \rightarrow 0} \frac{e^h - 1}{h} = 1$$

und folglich

$$\exp'(x) = \exp(x) \quad \text{für alle } x \in \mathbb{R}.$$

Die Exponentialfunktion reproduziert sich also bei Differentiation. Dies ist eine Erklärung (von mehreren) für ihr häufiges Auftreten in Anwendungen.

(4) cos. Im Beweis von Behauptung 3.7 aus Kapitel 5 wurde berechnet

$$\cos y - \cos x = 2 \sin \frac{x+y}{2} \sin \frac{x-y}{2},$$

also ist

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

Nun ist $\lim_{h \rightarrow 0} \sin \left(x + \frac{h}{2}\right) = \sin x$ wegen der Stetigkeit von sin und

$$\lim_{h \rightarrow 0} \frac{\sin \frac{h}{2}}{\frac{h}{2}} = 1 \quad \text{nach Beh. 3.4 aus Kapitel 5,}$$

also

$$\lim_{h \rightarrow 0} \frac{\cos(x+h) - \cos x}{h} = -\sin x.$$

Somit gilt $\cos' x = -\sin x$ für $x \in \mathbb{R}$.

(5) sin. Analog wie in (4) findet man $\sin' x = \cos x$.

(6) Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto |x|$ ist in 0 nicht differenzierbar, denn es ist

$$\lim_{h \searrow 0} \frac{|0+h| - |0|}{h} = 1 \neq -1 = \lim_{h \nearrow 0} \frac{|0+h| - |0|}{h}.$$

Hier existieren aber einseitige Ableitungen; sie sind naheliegenderweise wie folgt definiert.

Definition. Die Funktion $f : D \rightarrow \mathbb{R}$ heißt *in x von rechts differenzierbar*, wenn

$$f'_r(x) := \lim_{h \searrow 0} \frac{f(x+h) - f(x)}{h}$$

existiert, und *in x von links differenzierbar*, wenn

$$f'_l(x) := \lim_{h \nearrow 0} \frac{f(x+h) - f(x)}{h}$$

existiert.

Die Berechnung der Ableitung von komplizierteren, „zusammengesetzten“ Funktionen wird ermöglicht durch einige allgemeine Sätze, die sich auf die Differentiation von Summen, Produkten, Quotienten, Umkehrfunktionen und Kompositionen differenzierbarer Funktionen beziehen.

1.3 Satz. Die Funktionen $f, g : D \rightarrow \mathbb{R}$ seien differenzierbar in einem Punkt $x \in D$. Dann sind die Funktionen $f + g$, fg , λf ($\lambda \in \mathbb{R}$) in x differenzierbar, und es gilt

$$\begin{aligned} (f + g)'(x) &= f'(x) + g'(x), \\ (fg)'(x) &= f'(x)g(x) + f(x)g'(x) \quad (\text{Produktregel}), \\ (\lambda f)'(x) &= \lambda f'(x). \end{aligned}$$

Ist $g(x) \neq 0$, so ist die Funktion $\frac{f}{g}$ differenzierbar in x , und es gilt

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} \quad (\text{Quotientenregel}).$$

Beweis. Die Behauptung für $f + g$ folgt unmittelbar aus den Definitionen. Es ist

$$\begin{aligned} \frac{(fg)(x+h) - (fg)(x)}{h} &= \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\ &= \frac{f(x+h) - f(x)}{h}g(x) + f(x+h)\frac{g(x+h) - g(x)}{h}. \end{aligned}$$

Daraus und aus der Stetigkeit von f folgt

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x).$$

Als Spezialfall ergibt sich $(\lambda f)'(x) = \lambda f'(x)$.

Sei $g(x) \neq 0$. Dann ist $g(y) \neq 0$ für alle y aus einer Umgebung von x . Es ist

$$\begin{aligned} \frac{\left(\frac{1}{g}\right)(x+h) - \left(\frac{1}{g}\right)(x)}{h} &= \frac{1}{h} \left(\frac{1}{g(x+h)} - \frac{1}{g(x)} \right) \\ &= -\frac{1}{g(x+h)g(x)} \frac{g(x+h) - g(x)}{h}, \end{aligned}$$

woraus

$$\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g(x)^2}$$

folgt. Sodann ergibt sich

$$\begin{aligned} \left(\frac{f}{g}\right)'(x) &= \left(f \cdot \frac{1}{g}\right)'(x) = f'(x)\frac{1}{g}(x) + f(x) \left(-\frac{g'(x)}{g(x)^2}\right) \\ &= \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}. \quad \blacksquare \end{aligned}$$

Beispiele. (1) Sei $f(x) = x^n$ für $x \in \mathbb{R}$ ($n \in \mathbb{N}$).

Behauptung. $f'(x) = nx^{n-1}$.

Beweis. durch Induktion. Für $n = 1$ ist die Behauptung richtig. Sei sie bewiesen für n . Sei $f(x) = x^{n+1}$, also $f(x) = x^n \cdot x$ und daher

$$f'(x) = nx^{n-1} \cdot x + x^n \cdot 1 = (n+1)x^n. \quad \blacksquare$$

(2) Sei $f(x) = x^{-n}$ für $x \in \mathbb{R}$, ($n \in \mathbb{N}$). Nach der Quotientenregel und Beispiel (1) folgt

$$f'(x) = \frac{-nx^{n-1}}{(x^n)^2} = -nx^{-n-1}.$$

(3) Aus $\tan x = \frac{\sin x}{\cos x}$ folgt

$$\begin{aligned} \tan'(x) &= \frac{\sin' x \cos x - \sin x \cos' x}{\cos^2 x} \\ &= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}. \end{aligned}$$

Der folgende Satz zeigt, wie man Umkehrfunktionen differenzierbarer Funktionen differenziert.

1.4 Satz. Sei $f : D \rightarrow \mathbb{R}$ injektiv, in $a \in D$ differenzierbar, und $f'(a) \neq 0$. Ist die Umkehrfunktion $\varphi : f(D) \rightarrow D$ von f in $b = f(a)$ stetig, so ist φ in b differenzierbar, und es gilt

$$\varphi'(b) = \frac{1}{f'(a)} = \frac{1}{f'(\varphi(b))}.$$

Beweis. Setze

$$g(x) := \begin{cases} \frac{f(x) - f(a)}{x - a} & \text{für } x \in D \setminus \{a\}, \\ f'(a) & \text{für } x = a. \end{cases}$$

Dann ist g an der Stelle a stetig, also ist nach Satz 2.8 aus Kapitel 4 und Voraussetzung die Funktion $g \circ \varphi$ an der Stelle b stetig. Somit gilt

$$\lim_{y \rightarrow b} (g \circ \varphi)(y) = (g \circ \varphi)(b) = g(\varphi(b)) = g(a) = f'(a) \neq 0$$

und daher

$$\lim_{y \rightarrow b} \frac{1}{(g \circ \varphi)(y)} = \frac{1}{f'(a)}.$$

Für $y \in f(D)$ mit $y \neq b$ ist $\varphi(y) \neq a$, also

$$\frac{1}{(g \circ \varphi)(y)} = \frac{1}{\frac{f(\varphi(y)) - f(a)}{\varphi(y) - a}} = \frac{\varphi(y) - a}{f(\varphi(y)) - f(a)},$$

woraus die Behauptung folgt. ■

Beispiele. In den folgenden Beispielen sind Differenzierbarkeit der Funktion f und Stetigkeit der Umkehrfunktion φ jeweils bereits bekannt. Wir benutzen

$$\varphi'(x) = \frac{1}{f'(\varphi(x))}.$$

(1) \ln ist die Umkehrfunktion von \exp , und es ist $\exp' x = \exp x \neq 0$, also

$$\ln' x = \frac{1}{\exp(\ln x)} = \frac{1}{x}.$$

- (2) \arccos ist die Umkehrfunktion von \cos $| [0, \pi]$. Für $y \in (0, \pi)$ gilt $\cos' y = -\sin y \neq 0$, also für $x \in (-1, 1)$

$$\arccos' x = \frac{1}{\cos'(\arccos x)} = \frac{-1}{\sin(\arccos x)}.$$

Setzen wir $\arccos x = y$, so ist $\cos y = x$, also $\sin^2 y = 1 - \cos^2 y = 1 - x^2$, folglich $\sin y = \sqrt{1 - x^2}$ ($-\sqrt{1 - x^2}$ kommt nicht in Frage wegen $\sin y > 0$). Es folgt

$$\arccos' x = \frac{-1}{\sqrt{1 - x^2}} \quad \text{für } -1 < x < 1.$$

- (3) Analog findet man

$$\arcsin' x = \frac{1}{\sqrt{1 - x^2}} \quad \text{für } -1 < x < 1.$$

- (4) \arctan ist die Umkehrfunktion von \tan $| (-\frac{\pi}{2}, \frac{\pi}{2})$. Für $y \in (-\frac{\pi}{2}, \frac{\pi}{2})$ ist $\tan' y = \frac{1}{\cos^2 y} \neq 0$, also für $x \in \mathbb{R}$

$$\arctan' x = \cos^2(\arctan x).$$

Mit $y := \arctan x$ ist $x = \tan y = \frac{\sin y}{\cos y}$. Daraus folgt $\cos^2 y = 1 - \sin^2 y$, was man auch als $\cos^2 y = \frac{1}{1+x^2}$ schreiben kann. Somit haben wir

$$\arctan' x = \frac{1}{1+x^2}.$$

1.5 Satz (Kettenregel). Sei $f : D \rightarrow \tilde{D}$ differenzierbar in $a \in D$ und sei $g : \tilde{D} \rightarrow \mathbb{R}$ differenzierbar in $f(a)$. Dann ist $g \circ f$ differenzierbar in a , und es gilt

$$(g \circ f)'(a) = g'(f(a))f'(a).$$

Beweis. Setze

$$h(y) := \begin{cases} \frac{g(y) - g(f(a))}{y - f(a)} & \text{für } y \in \tilde{D} \setminus \{f(a)\}, \\ g'(f(a)) & \text{für } y = f(a). \end{cases}$$

Dann ist h an der Stelle $f(a)$ stetig, also ist die Funktion $h \circ f$ an der Stelle a stetig. Für $x \in D \setminus \{a\}$ gilt

$$\begin{aligned}
& \frac{(g \circ f)(x) - (g \circ f)(a)}{x - a} \\
&= \begin{cases} \frac{g(f(x)) - g(f(a))}{f(x) - f(a)} \frac{f(x) - f(a)}{x - a}, & \text{falls } f(x) \neq f(a) \\ 0, & \text{falls } f(x) = f(a) \end{cases} \\
&= h(f(x)) \frac{f(x) - f(a)}{x - a}.
\end{aligned}$$

Aus $\lim_{x \rightarrow a} h(f(x)) = h(f(a)) = g'(f(a))$ und

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = f'(a)$$

folgt also die Behauptung. ■

Beispiele. (1) Sei $\alpha \in \mathbb{R}$, $f(x) = x^\alpha$ für $x > 0$. Es ist $f(x) = e^{\alpha \ln x} = \exp(g(x))$ mit $g(x) = \alpha \ln x$. Es folgt

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

(2) $f(x) = \ln \sin x$ für $0 < x < \pi$. Es ist $f'(x) = \ln'(\sin x) \cdot \sin' x = \frac{1}{\sin x} \cdot \cos x = \cotan x$.

(3) $f(x) = \cos e^{(x^2)}$. Es ist

$$f'(x) = -\left(\sin e^{(x^2)}\right) \cdot e^{(x^2)} \cdot 2x.$$

(4) $f(x) = x^{\frac{1}{x}}$ für $x > 0$. Es ist $f(x) = e^{\frac{1}{x} \ln x}$, also

$$f'(x) = e^{\frac{1}{x} \ln x} \left(-\frac{1}{x^2} \ln x + \frac{1}{x} \cdot \frac{1}{x}\right) = x^{\frac{1}{x}-2}(1 - \ln x).$$

6.2 Eigenschaften differenzierbarer Funktionen

Bei differenzierbaren Funktionen steht man oft vor dem Problem, von Eigenschaften der Ableitung auf Eigenschaften der Funktion selbst schließen zu müssen. Zentrales Hilfsmittel hierbei ist der Mittelwertsatz der Differentialrechnung mit seinen Folgerungen. Der wesentliche Kern dieses Satzes steckt bereits in der folgenden Behauptung.

2.1 Satz (von Rolle). Sei $a < b$, $f : [a, b] \rightarrow \mathbb{R}$ stetig, $f|_{(a, b)}$ differenzierbar und $f(a) = f(b)$. Dann existiert ein $c \in (a, b)$ mit $f'(c) = 0$.

Beweis. Nach Satz 3.2 aus Kapitel 4 nimmt f ein Maximum und ein Minimum an. Mindestens eines von beiden, o.B.d.A. ein Maximum, wird an einer Stelle $c \in (a, b)$ angenommen (wegen $f(a) = f(b)$). Es ist

$$f'(c) = \lim_{h \searrow 0} \frac{f(c+h) - f(c)}{h} \leq 0$$

und

$$f'(c) = \lim_{h \nearrow 0} \frac{f(c+h) - f(c)}{h} \geq 0,$$

also $f'(c) = 0$. ■

2.2 Satz (Mittelwertsatz der Differentialrechnung). Sei $a < b$, $f : [a, b] \rightarrow \mathbb{R}$ stetig und $f|_{(a, b)}$ differenzierbar. Dann existiert ein $c \in (a, b)$ mit

$$\frac{f(b) - f(a)}{b - a} = f'(c).$$

Beweis. Setze

$$g(x) := f(x) - \frac{f(b) - f(a)}{b - a}(x - a) \quad \text{für } x \in [a, b]$$

und wende Satz 2.1 an. ■

Bemerkung. Wir können das Ergebnis auch etwas anders schreiben: Ist f in $[x - |h|, x + |h|]$ differenzierbar, so gilt

$$f(x+h) = f(x) + f'(x + \vartheta h)h \quad \text{mit einem } \vartheta \in (0, 1).$$

Wir können also den „Zuwachs“ $f(x+h) - f(x)$ nicht nur näherungsweise durch $f'(x)h$, sondern exakt durch $f'(x + \vartheta h)h$ ausdrücken. Allerdings wissen wir von ϑ nur, dass es in $(0, 1)$ liegt.

Wir kommen zu einigen Folgerungen aus dem Mittelwertsatz. Von Interesse in Anwendungen ist statt der schärferen Aussage des Mittelwertsatzes oft nur die folgende unmittelbare Konsequenz.

2.3 Folgerung. Sei D ein Intervall und $f : D \rightarrow \mathbb{R}$ differenzierbar mit $|f'(x)| \leq c$ für $x \in D$ ($c \in \mathbb{R}$ fest). Dann gilt

$$|f(x) - f(y)| \leq c|x - y| \quad \text{für } x, y \in D. \quad (2.4)$$

Bemerkung. Eine Funktion f mit der Eigenschaft (2.4) nennt man auch “Lipschitz stetig“.

Anwendung mit $c = 0$ ergibt:

2.5 Satz. Sei D ein Intervall, $f : D \rightarrow \mathbb{R}$ differenzierbar und $f'(x) = 0$ für alle $x \in D$. Dann ist f konstant.

Bemerkung. Die Voraussetzung, dass der Definitionsbereich von f ein Intervall ist, ist offenbar wesentlich.

Die folgende Beobachtung ist in den Anwendungen wichtig:

Behauptung. Ist $f : [0, b] \rightarrow \mathbb{R}$ eine differenzierbare Funktion mit

$$f'(x) = cf(x) \quad \text{für } x \in [0, b] \quad \text{und } f(0) = a$$

($a, b, c \in \mathbb{R}$ gegeben), so ist

$$f(x) = ae^{cx} \quad \text{für } x \in [0, b].$$

Beweis. Setze $g(x) := f(x)e^{-cx}$ für $x \in [0, b]$. Dann ist

$$g'(x) = f'(x)e^{-cx} - f(x)ce^{-cx} = (f'(x) - cf(x))e^{-cx} = 0$$

für alle $x \in [0, b]$, also ist g konstant, und zwar gleich $g(0) = a$. ■

2.6 Satz. Sei D ein Intervall und $f : D \rightarrow \mathbb{R}$ differenzierbar. Dann ist f genau dann monoton wachsend, wenn $f'(x) \geq 0$ für $x \in D$ gilt.

Beweis. „ \Rightarrow “: Sei f monoton wachsend. Dann gilt für $x \in D$, $h \neq 0$ und $x + h \in D \setminus \{x\}$:

$$\frac{f(x+h) - f(x)}{h} \geq 0,$$

also $f'(x) \geq 0$.

„ \Leftarrow “: Nach Satz 2.2 gilt für $x, y \in D$ mit $y > x$

$$f(y) - f(x) = f'(c)(y - x) \geq 0$$

mit einem geeigneten $c \in (x, y)$. ■

Bemerkung. Aus $f'(x) > 0$ für alle $x \in D$ folgt offenbar, dass f streng monoton wachsend ist (aber nicht umgekehrt).

Wir beweisen noch eine Verallgemeinerung des Mittelwertsatzes, die vor allem bei der Ermittlung von Grenzwerten gute Dienste leistet:

2.7 Satz (Zweiter Mittelwertsatz der Differentialrechnung). *Sei $a < b$, seien $f, g : [a, b] \rightarrow \mathbb{R}$ stetig, $f|_{(a,b)}$ und $g|_{(a,b)}$ differenzierbar. Dann existiert ein $c \in (a, b)$ mit*

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

Beweis. Setze

$$h(x) := (g(x) - g(a))(f(b) - f(a)) - (f(x) - f(a))(g(b) - g(a))$$

für $x \in [a, b]$ und wende den Satz von Rolle 2.1 an. ■

Als Anwendung haben wir die folgende, oft nützliche Regel.

2.8 Satz (Regel von de l'Hospital). *Seien $f, g : (a, b] \rightarrow \mathbb{R}$ differenzierbar. Es gelte*

$$\begin{aligned} \lim_{x \searrow a} f(x) = \lim_{x \searrow a} g(x) = 0, \\ g(x) \neq 0, \quad g'(x) \neq 0 \quad \text{für } x \in (a, b], \\ \lim_{x \searrow a} \frac{f'(x)}{g'(x)} = c. \end{aligned}$$

Dann ist

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = c.$$

Beweis. Wir setzen noch $f(a) = g(a) = 0$; dann sind f und g stetige Funktionen auf $[a, b]$. Sei $\varepsilon \in \mathbb{R}^+$. Nach Voraussetzung existiert ein $\delta \in \mathbb{R}^+$ mit

$$\left| \frac{f'(y)}{g'(y)} - c \right| < \varepsilon \quad \text{für } a < y < a + \delta.$$

Sei jetzt $a < x < a + \delta$. Nach Satz 2.7 existiert ein $y \in (a, x)$ mit

$$\frac{f(x)}{g(x)} = \frac{f'(y)}{g'(y)}.$$

Es folgt

$$\left| \frac{f(x)}{g(x)} - c \right| = \left| \frac{f'(y)}{g'(y)} - c \right| < \varepsilon.$$

Damit ist

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = c$$

gezeigt. ■

Bemerkung. Die analoge Aussage für $\lim_{x \nearrow a}$ liegt auf der Hand. Aus beiden folgt eine entsprechende Aussage über $\lim_{x \rightarrow a}$. Durch eine naheliegende Abwandlung des Beweises zeigt man auch eine entsprechende Aussage für $\lim_{x \rightarrow \infty}$ bzw. $\lim_{x \rightarrow -\infty}$.

Nützlich ist oft auch die folgende Version der Regel von de l'Hospital, bei der Zähler und Nenner des fraglichen Quotienten nicht gegen 0 konvergieren, sondern bestimmt divergieren.

2.9 Satz. Seien $f, g : (a, b] \rightarrow \mathbb{R}$ differenzierbar. Es gelte

$$\begin{aligned} \lim_{x \searrow a} f(x) = \infty, \quad \lim_{x \searrow a} g(x) = \infty, \\ g(x) \neq 0, \quad g'(x) \neq 0 \quad \text{für } x \in (a, b], \\ \lim_{x \searrow a} \frac{f'(x)}{g'(x)} = c. \end{aligned}$$

Dann ist

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = c.$$

Beweis. Sei $\varepsilon \in \mathbb{R}^+$. Nach Voraussetzung existiert ein $d > a$ mit

$$\left| \frac{f'(y)}{g'(y)} - c \right| < \frac{\varepsilon}{4} \quad \text{für } a < y < d. \quad (\star)$$

Es gibt ein $\delta \in \mathbb{R}^+$ mit $a + \delta \leq d$, so dass für alle x mit $a < x < a + \delta$ gilt:

$$\begin{aligned} 0 \neq g(x) \neq g(d), \quad 0 \neq f(x) \neq f(d), \\ \left| \frac{1 - \frac{g(d)}{g(x)}}{1 - \frac{f(d)}{f(x)}} \right| < 2, \quad |c| \left| \frac{1 - \frac{g(d)}{g(x)}}{1 - \frac{f(d)}{f(x)}} - 1 \right| < \frac{\varepsilon}{2}. \end{aligned}$$

Für alle x mit $a < x < a + \delta$ gilt dann wegen

$$\begin{aligned} \frac{f(x)}{g(x)} - c &= \left(\frac{f(x) - f(d)}{g(x) - g(d)} - c \right) \frac{f(x)}{f(x) - f(d)} \frac{g(x) - g(d)}{g(x)} \\ &\quad + c \left(\frac{f(x)}{f(x) - f(d)} \frac{g(x) - g(d)}{g(x)} - 1 \right) \end{aligned}$$

die Abschätzung

$$\left| \frac{f(x)}{g(x)} - c \right| \leq \left| \frac{f(x) - f(d)}{g(x) - g(d)} - c \right| \underbrace{\left| \frac{1 - \frac{g(d)}{g(x)}}{1 - \frac{f(d)}{f(x)}} \right|}_{< 2} + |c| \underbrace{\left| \frac{1 - \frac{g(d)}{g(x)}}{1 - \frac{f(d)}{f(x)}} - 1 \right|}_{< \frac{\varepsilon}{2}}$$

Nach Satz 2.7 existiert ein $y \in (x, d)$ mit

$$\frac{f(x) - f(d)}{g(x) - g(d)} = \frac{f'(y)}{g'(y)}.$$

Dies und die Abschätzung (*) ergeben also

$$\left| \frac{f(x) - f(d)}{g(x) - g(d)} - c \right| < \frac{\varepsilon}{4}.$$

Damit ist

$$\left| \frac{f(x)}{g(x)} - c \right| < \varepsilon$$

gezeigt, woraus die Behauptung folgt. ■

Bemerkung. Analoge Aussagen gelten wieder für $\lim_{x \nearrow a}$ etc.

Beispiele. (1) $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$, denn $\lim_{x \rightarrow 0} \sin x = \lim_{x \rightarrow 0} x = 0$ und $\lim_{x \rightarrow 0} \frac{\cos x}{1} = 1$.

(2) $\lim_{x \rightarrow 0} \left(\frac{1}{\sin x} - \frac{1}{x} \right) = ?$ Es ist

$$\frac{x - \sin x}{x \sin x} =: \frac{f(x)}{g(x)} \quad \text{mit} \quad \lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 0} g(x) = 0,$$

$$f'(x) = 1 - \cos x, \quad g'(x) = \sin x + x \cos x,$$

$$\lim_{x \rightarrow 0} f'(x) = 0, \quad \lim_{x \rightarrow 0} g'(x) = 0,$$

$$(f')'(x) = \sin x, \quad (g')'(x) = 2 \cos x - x \sin x,$$

$$\lim_{x \rightarrow 0} \frac{(f')'(x)}{(g')'(x)} = 0 \Rightarrow \lim_{x \rightarrow 0} \frac{f'(x)}{g'(x)} = 0 \Rightarrow \lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = 0.$$

(3) $\lim_{x \searrow 0} x^\alpha \ln x = ?$ ($\alpha > 0$) Es ist

$$-x^\alpha \ln x = \frac{\ln \frac{1}{x}}{x^{-\alpha}} =: \frac{f(x)}{g(x)},$$

$$\lim_{x \searrow 0} f(x) = \infty, \quad \lim_{x \searrow 0} g(x) = \infty,$$

$$f'(x) = -\frac{1}{x}, \quad g'(x) = -\alpha x^{-\alpha-1},$$

$$\frac{f'(x)}{g'(x)} = \frac{1}{\alpha} x^\alpha, \quad \lim_{x \searrow 0} \frac{f'(x)}{g'(x)} = 0 \Rightarrow \lim_{x \searrow 0} x^\alpha \ln x = 0.$$

6.3 Höhere Ableitungen und Taylorformel

Definition. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion, $a \in D$. Falls f in einer Umgebung von a (geschnitten mit D) differenzierbar und f' in a differenzierbar ist, heißt f *zweimal differenzierbar in a* , und

$$(f')'(a) =: f''(a) =: f^{(2)}(a) =: \left. \frac{d^2 f(x)}{dx^2} \right|_{x=a}$$

heißt *zweite Ableitung von f in a* .

Rekursive Definition: Falls f in einer Umgebung von a (geschnitten mit D) k -mal differenzierbar und $f^{(k)}$ in a differenzierbar ist, heißt f *$(k+1)$ -mal differenzierbar in a* , und

$$f^{(k)'}(a) =: f^{(k+1)}(a) =: \left. \frac{d^{k+1} f(x)}{dx^{k+1}} \right|_{x=a}$$

heißt *$(k+1)$ -te Ableitung von f in a* . Setze noch $f^{(0)} = f$, $f^{(1)} = f'$.

f heißt *k -mal differenzierbar*, wenn f k -mal differenzierbar in a für alle $a \in D$ ist.

f heißt *k -mal stetig differenzierbar*, wenn f k -mal differenzierbar und $f^{(k)}$ stetig ist.

Beispiele. Die bisher betrachteten speziellen Funktionen, also rationale Funktionen, exp, ln, sin, cos, tan, arcsin etc. und alle Funktionen, die aus

solchen durch rationale Rechenoperationen und Komposition zusammengesetzt sind, führen bei Differentiation auf Funktionen vom selben Typ; sie sind daher beliebig oft differenzierbar (im jeweiligen Definitionsbereich).

Wir bringen noch ein weniger triviales Beispiel, das verschiedentlich von Nutzen ist. Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$f(x) := \begin{cases} e^{-\frac{1}{x}} & \text{für } x > 0, \\ 0 & \text{für } x \leq 0. \end{cases}$$

Behauptung. f ist beliebig oft differenzierbar.

Beweis. An jeder Stelle $x \neq 0$ ist f beliebig oft differenzierbar (trivial für $x < 0$, klar nach vorstehender Bemerkung für $x > 0$). Für $x > 0$ gilt

$$f^{(k)}(x) = e^{-\frac{1}{x}} \frac{p_k(x)}{x^{2k}} \quad (3.1)$$

mit einer gewissen Polynomfunktion p_k . Dies beweisen wir durch Induktion. Zunächst ist

$$f^{(1)}(x) = f'(x) = e^{-\frac{1}{x}} \cdot \frac{1}{x^2}.$$

Für eine Zahl $k \geq 1$ sei die Darstellung (3.1) bereits gezeigt. Dann ist

$$\begin{aligned} f^{(k+1)}(x) &= e^{-\frac{1}{x}} \left(\frac{1}{x^2} \frac{p_k(x)}{x^{2k}} + \frac{p'_k(x)x^{2k} - p_k(x)2kx^{2k-1}}{x^{4k}} \right) \\ &= e^{-\frac{1}{x}} \frac{p_k(x) + p'_k(x)x^2 - 2kp_k(x)x}{x^{2(k+1)}}, \end{aligned}$$

woraus die Behauptung folgt.

Nun zeigen wir durch Induktion, dass f in 0 k -mal differenzierbar ist. Für $x > 0$ gilt

$$\frac{f(x) - f(0)}{x} = \frac{1}{x} e^{-\frac{1}{x}} \rightarrow 0$$

für $x \searrow 0$ nach Behauptung 1.7 aus Kapitel 5, also $f'_r(0) = 0$. Da auch $f'_\ell(0) = 0$ ist, ist $f'(0) = 0$.

Sei bereits bewiesen, dass f in 0 k -mal differenzierbar ist. Dann existiert also die k -te Ableitung $f^{(k)} : \mathbb{R} \rightarrow \mathbb{R}$. Für $x > 0$ gilt

$$\frac{f^{(k)}(x) - f^{(k)}(0)}{x} = \frac{p_k(x)}{x^{2k+1}} e^{-\frac{1}{x}} \rightarrow 0$$

für $x \searrow 0$ nach Behauptung 1.7 aus Kapitel 5, also $f^{(k)'}_r(0) = 0$. Da auch $f^{(k)'}_\ell(0) = 0$ ist, ist $f^{(k)'}(0) = 0$. ■

Differenzierbarkeit einer Funktion bedeutet, grob gesagt, dass sich die Funktion durch eine affine Funktion, also eine Polynomfunktion vom Grad 1, gut approximieren lässt. Analog kann man k -malige Differenzierbarkeit so interpretieren, dass die Funktion durch eine Polynomfunktion vom Grad k gut approximiert werden kann, und zwar „von k -ter Ordnung“. Zunächst zeigen wir unter einer schärferen Voraussetzung eine genauere Aussage.

3.2 Satz (Taylorscher Satz). *Sei $a < b$, $n \in \mathbb{N}_0$, $f : [a, b] \rightarrow \mathbb{R}$ n -mal stetig differenzierbar und f auf (a, b) $(n+1)$ -mal differenzierbar. Dann existiert ein $c \in (a, b)$ mit*

$$f(b) = f(a) + \frac{1}{1!} f'(a)(b-a) + \frac{1}{2!} f''(a)(b-a)^2 + \dots + \frac{1}{n!} f^{(n)}(a)(b-a)^n + \frac{1}{(n+1)!} f^{(n+1)}(c)(b-a)^{n+1}.$$

Beweis. Setze

$$g(x) := f(b) - f(x) - f'(x)(b-x) - \dots - \frac{f^{(n)}(x)}{n!} (b-x)^n - \alpha \frac{(b-x)^{n+1}}{(n+1)!}$$

für $a \leq x \leq b$. Dann ist $g(b) = 0$, und α sei so gewählt, dass auch $g(a) = 0$ ist. Dann kann man auf g den Satz von Rolle 2.1 anwenden. Es folgt die Existenz eines $c \in (a, b)$ mit $g'(c) = 0$. Nun ist

$$\begin{aligned} g'(x) &= 0 - f'(x) - [f''(x)(b-x) - f'(x)] - \dots \\ &\quad - \left[\frac{f^{(n+1)}(x)}{n!} (b-x)^n - \frac{f^{(n)}(x)}{(n-1)!} (b-x)^{n-1} \right] + \alpha \frac{(b-x)^n}{n!} \\ &= - \frac{f^{(n+1)}(x)}{n!} (b-x)^n + \alpha \frac{(b-x)^n}{n!}. \end{aligned}$$

Einsetzen von $x = c$ ergibt $f^{(n+1)}(c) = \alpha$. Setzt man jetzt $x = a$ in der Definitionsgleichung für g , so folgt die Behauptung. ■

Bemerkung. Für $n = 0$ reduziert sich der Taylorsche Satz 3.2 gerade auf den Mittelwertsatz 2.2.

Analog wie Satz 3.2 kann man den Fall $b < a$ behandeln; aus beiden Fällen ergibt sich Satz 3.3.

3.3 Satz. *Sei D ein Intervall und $f : D \rightarrow \mathbb{R}$ $(n+1)$ -mal differenzierbar, sei $a \in D$. Dann gilt für beliebige $x \in D$*

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + \frac{f^{(n+1)}(a + \vartheta(x-a))}{(n+1)!} (x-a)^{n+1}$$

(„Taylorformel“) mit einem (von a und x abhängenden) $\vartheta \in (0, 1)$.

Die Funktion

$$x \mapsto \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k$$

nennt man auch das *Taylorpolynom n -ter Ordnung von f zur Stelle a* .

Ist $f^{(n+1)}$ in D beschränkt, so folgt aus Satz 3.3 unter den dortigen Voraussetzungen insbesondere

$$\lim_{x \rightarrow a} \frac{1}{(x-a)^n} \left| f(x) - \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k \right| = 0.$$

Diese Aussage läßt sich auch schon unter schwächeren Voraussetzungen beweisen:

3.4 Satz. *Sei $a \in \mathbb{R}$, $n \in \mathbb{N}$. Die Funktion f sei in einer Umgebung von a $(n-1)$ -mal differenzierbar und in a n -mal differenzierbar. Dann gilt*

$$\lim_{x \rightarrow a} \frac{1}{(x-a)^n} \left| f(x) - \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k \right| = 0.$$

Beweis. Im Fall $n = 1$ lautet die Behauptung

$$\lim_{x \rightarrow a} \left| \frac{f(x) - f(a) - f'(a)(x-a)}{x-a} \right| = 0;$$

dies ist wegen der Differenzierbarkeit von f in a erfüllt.

Sei jetzt $n \geq 2$. Es gibt nach Voraussetzung ein $\delta \in \mathbb{R}^+$, so dass f in $[a-\delta, a+\delta]$ $(n-1)$ -mal differenzierbar ist. Setze

$$g(x) := f(x) - \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k \quad \text{für } x \in [a-\delta, a+\delta].$$

Sei $x \in [a-\delta, a+\delta]$. Nach dem Taylorschen Satz 3.3 (für $n-2$ statt n), angewendet auf die Funktion g , existiert ein (i.a. von x abhängendes) c mit $|c-a| \leq |x-a|$, so dass

$$g(x) = \sum_{k=0}^{n-2} \frac{g^{(k)}(a)}{k!} (x-a)^k + \frac{g^{(n-1)}(c)}{(n-1)!} (x-a)^{n-1}$$

ist. Nun ist $g(a) = g'(a) = \dots = g^{(n-2)}(a) = 0$ und

$$g^{(n-1)}(c) = f^{(n-1)}(c) - f^{(n-1)}(a) - f^{(n)}(a)(c-a),$$

also

$$\begin{aligned} |g(x)| &= \frac{1}{(n-1)!} \left| f^{(n-1)}(c) - f^{(n-1)}(a) - f^{(n)}(a)(c-a) \right| |x-a|^{n-1} \\ &\leq \frac{1}{(n-1)!} \left| \frac{f^{(n-1)}(c) - f^{(n-1)}(a)}{c-a} - f^{(n)}(a) \right| |x-a|^n. \end{aligned}$$

Da $f^{(n-1)}$ nach Voraussetzung in a differenzierbar ist, gilt

$$\lim_{c \rightarrow a} \left| \frac{f^{(n-1)}(c) - f^{(n-1)}(a)}{c-a} - f^{(n)}(a) \right| = 0.$$

Wegen $|c-a| \leq |x-a|$ folgt

$$\lim_{x \rightarrow a} \frac{|g(x)|}{|x-a|^n} = 0,$$

wie behauptet. ■

Taylorreihen

Ist die Funktion f in einer Umgebung von a beliebig oft differenzierbar, so gilt für sie die Taylorformel für jede natürliche Zahl n . Dies legt es nahe, neben den Taylorpolynomen

$$\sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k$$

auch die unendliche Reihe

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k$$

zu betrachten.

Definition. Sei $D \subset \mathbb{R}$ ein Intervall und $f : D \rightarrow \mathbb{R}$ beliebig oft differenzierbar, sei $a \in D$. Dann heißt die Reihe

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k$$

die *Taylorreihe* von f an der Stelle x zur Entwicklungsstelle a .

Bemerkung. Achtung! Die Taylorreihe einer Funktion f an der Stelle x braucht nicht zu konvergieren, und wenn sie konvergiert, braucht sie nicht gegen $f(x)$ zu konvergieren. Zum Beispiel sei

$$f(x) := \frac{1}{1-x} \quad \text{für } x \neq 1.$$

Man berechnet

$$f^{(k)}(x) = \frac{k!}{(1-x)^{k+1}}, \quad \text{also } \frac{f^{(k)}(0)}{k!} = 1.$$

Die Taylorreihe von f zur Entwicklungsstelle $a = 0$ ist also gegeben durch

$$\sum_{k=0}^{\infty} x^k.$$

Diese Reihe konvergiert, wie wir wissen, für $|x| < 1$, und zwar gegen $f(x)$. Aber für $|x| \geq 1$ ist die Reihe divergent. – Es gibt sogar beliebig oft differenzierbare Funktionen, deren Taylorreihe zur Entwicklungsstelle a für jedes $x \neq a$ divergiert.

Als zweites Beispiel betrachten wir

$$f(x) := \begin{cases} e^{-\frac{1}{x}} & \text{für } x > 0, \\ 0 & \text{für } x \leq 0. \end{cases}$$

Wie früher gezeigt, ist f beliebig oft differenzierbar und $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}_0$. Die Taylorreihe dieser Funktion zur Entwicklungsstelle $a = 0$ konvergiert also trivialerweise, aber für $x > 0$ nicht gegen $f(x)$.

Man muß also in jedem Einzelfall untersuchen, für welche x die Taylorreihe einer gegebenen Funktion f an der Stelle x wirklich gegen $f(x)$ konvergiert. Definieren wir das $(n+1)$ -te Restglied R_{n+1} durch

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + R_{n+1}(x),$$

so ist die Konvergenz der Taylorreihe gegen $f(x)$ also gleichbedeutend mit

$$\lim_{n \rightarrow \infty} R_{n+1}(x) = 0.$$

Hier ist nun die Taylorformel von Nutzen, die uns eine Darstellung dieses Restgliedes angibt, nämlich

$$R_{n+1}(x) = \frac{f^{(n+1)}(c)}{(n+1)!} (x-a)^{n+1}$$

mit einer Zahl c zwischen a und x . In manchen, aber nicht in allen Fällen kann man hiermit die gewünschte Konvergenz der Taylorreihe gegen f zeigen.

Als Beispiel betrachten wir die Exponentialfunktion und eine beliebige Entwicklungsstelle a . Das Restglied lautet

$$R_{n+1}(x) = \frac{f^{(n+1)}(c_n)}{(n+1)!} (x-a)^{n+1} = \frac{e^{c_n}}{(n+1)!} (x-a)^{n+1}.$$

Die Zahl c_n hängt zwar von n ab, aber es ist

$$e^{c_n} \leq \max\{e^a, e^x\},$$

also gilt $\lim_{n \rightarrow \infty} R_{n+1}(x) = 0$. Daher ist

$$e^x = \sum_{k=0}^{\infty} \frac{e^a}{k!} (x-a)^k = e^a e^{x-a},$$

womit sich wieder die Funktionalgleichung der Exponentialfunktion ergeben hat.

Die genauere Behandlung von Taylorreihen erfolgt erst in Abschnitt 8.2, da es nützlich ist, hierzu die Integralrechnung zur Verfügung zu haben.

Wir beschließen dieses Kapitel mit zwei Anwendungen von Ableitungen zweiter und höherer Ordnung.

Extremwerte differenzierbarer Funktionen

Die Taylorformel liefert Informationen über das Verhalten hinreichend oft differenzierbarer Funktionen in der Umgebung eines Punktes. Als Anwendung ergeben sich zum Beispiel Aussagen über Extremwerte.

Definition. Die Funktion $f : D \rightarrow \mathbb{R}$ hat im Punkt $a \in D$ ein *lokales Maximum (starkes lokales Maximum)*, wenn es eine Umgebung U von a gibt mit $f(a) \geq f(x)$ (bzw. $f(a) > f(x)$) für alle $x \in (U \cap D) \setminus \{a\}$.

Analog: Minimum, Oberbegriff: Extremum

Der Punkt $a \in D$ heißt *innerer Punkt* von D , wenn D Umgebung von a ist, also wenn ein $\varepsilon > 0$ existiert mit $(a - \varepsilon, a + \varepsilon) \subset D$.

3.5 Satz. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion und a innerer Punkt von D .

(a) Ist f in a differenzierbar und hat f in a ein lokales Extremum, so ist $f'(a) = 0$.

(b) Sei f n -mal stetig differenzierbar und

$$f'(a) = \dots = f^{(n-1)}(a) = 0, \quad f^{(n)}(a) \neq 0.$$

Ist n ungerade, so hat f in a kein lokales Extremum. Sei n gerade. Im Fall $f^{(n)}(a) > 0$ hat f in a ein starkes lokales Minimum, im Fall $f^{(n)}(a) < 0$ ein starkes lokales Maximum.

Beweis. Die Behauptung (a) wurde bereits im Beweis von Satz 2.1 gezeigt.

(b) Da a innerer Punkt von D ist, folgt aus der Taylorformel und den Voraussetzungen

$$f(x) = f(a) + \frac{f^{(n)}(c)}{n!}(x-a)^n$$

mit $|c-a| \leq |x-a|$ für alle x aus einer Umgebung von a . Da $f^{(n)}$ stetig und $f^{(n)}(a) \neq 0$ ist, existiert eine Umgebung U von a mit $f^{(n)}(c)f^{(n)}(a) > 0$ für alle $c \in U$. Jetzt liest man die Behauptung ab. ■

Konvexe Funktionen

Wechselt die erste Ableitung einer Funktion auf einem Intervall nicht das Vorzeichen, so ist die Funktion dort monoton. Die Bedingung, dass die zweite Ableitung nicht das Vorzeichen ändern soll, führt auf die wichtige Klasse der konvexen bzw. konkaven Funktionen.

Definition. Sei D ein Intervall. Die Funktion $f : D \rightarrow \mathbb{R}$ heißt *konvex*, wenn

$$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2)$$

für alle $x_1, x_2 \in D$ und alle $\lambda \in [0, 1]$ gilt. f heißt *konkav*, wenn $-f$ konvex ist.

3.6 Satz. Sei D ein Intervall und $f : D \rightarrow \mathbb{R}$ differenzierbar. Dann ist f konvex genau dann, wenn f' monoton wachsend ist.

Beweis. Die Konvexitätsbedingung ist äquivalent mit: Für alle $x_1, x_2 \in D$ mit $x_1 < x_2$ und alle $x \in (x_1, x_2)$ gilt

$$f(x) \leq \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2). \quad (3.7)$$

Seien x_1, x_2 wie oben. Ist f konvex, so folgt aus (3.7) durch Umrechnung

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x}$$

und daraus durch Grenzübergang $x \searrow x_1$ bzw. $x \nearrow x_2$ die Ungleichung $f'(x_1) \leq f'(x_2)$.

Umgekehrt folgt aus dem Mittelwertsatz 2.2 die Existenz von Zahlen $c_1 \in (x_1, x)$ und $c_2 \in (x, x_2)$ mit

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(c_1), \quad \frac{f(x_2) - f(x)}{x_2 - x} = f'(c_2).$$

Ist nun f' monoton wachsend, so ist $f'(c_1) \leq f'(c_2)$, woraus die Ungleichung (3.7) sich durch Umrechnung ergibt. ■

Als Folgerung ergibt sich wegen Satz 2.6:

3.8 Satz. Sei D ein Intervall und $f : D \rightarrow \mathbb{R}$ zweimal differenzierbar. Dann ist f genau dann konvex, wenn $f''(x) \geq 0$ ist für alle $x \in D$.

Als Anwendung beweisen wir eine nützliche Ungleichung. Dazu betrachten wir zunächst die Funktion $f(x) = \ln x$ ($x > 0$) und berechnen $f''(x) = -\frac{1}{x^2} < 0$, also ist \ln konkav. Für beliebige $x, y > 0$ und $p, q \in (1, \infty)$ mit $\frac{1}{p} + \frac{1}{q} = 1$ gilt daher

$$\ln \left(\frac{1}{p}x + \frac{1}{q}y \right) \geq \frac{1}{p} \ln x + \frac{1}{q} \ln y,$$

folglich

$$\frac{1}{p}x + \frac{1}{q}y \geq e^{\frac{1}{p} \ln x} e^{\frac{1}{q} \ln y} = x^{\frac{1}{p}} y^{\frac{1}{q}},$$

also

$$x^{\frac{1}{p}} y^{\frac{1}{q}} \leq \frac{x}{p} + \frac{y}{q}.$$

Dies gilt trivialerweise auch für $x = 0$ oder $y = 0$.

Hieraus können wir herleiten:

3.9 Satz. Für $p, q \in (1, \infty)$ mit $\frac{1}{p} + \frac{1}{q} = 1$ und beliebige $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{R}$ gilt

$$\sum_{i=1}^n |x_i y_i| \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}}$$

(Höldersche Ungleichung).

Beweis. O.B.d.A. sei die rechte Seite $\neq 0$. Setze

$$a_i := \frac{|x_i|^p}{\sum_{j=1}^n |x_j|^p}, \quad b_i = \frac{|y_i|^q}{\sum_{j=1}^n |y_j|^q}.$$

Aus der obigen Ungleichung folgt

$$\frac{|x_i y_i|}{(\sum |x_j|^p)^{\frac{1}{p}} (\sum |y_j|^q)^{\frac{1}{q}}} = a_i^{\frac{1}{p}} b_i^{\frac{1}{q}} \leq \frac{a_i}{p} + \frac{b_i}{q},$$

also

$$\frac{\sum_i |x_i y_i|}{(\sum |x_j|^p)^{\frac{1}{p}} (\sum |y_i|^q)^{\frac{1}{q}}} \leq \frac{\sum a_i}{p} + \frac{\sum b_i}{q} = \frac{1}{p} + \frac{1}{q} = 1. \quad \blacksquare$$

7 Integration

Notwendigkeit des Integralbegriffes und Hinweise zu seiner Präzisierung liegen auf der Hand. Betrachten wir etwa den physikalischen Begriff der Arbeit, die im einfachsten Fall, nämlich bei konstanter Kraft, als das Produkt „Kraft mal Weg“ definiert ist. Bei stückweise konstanter Kraft wird man sie entsprechend als Summe definieren. Stellen wir die Kraft als Funktion des Weges graphisch dar, so veranschaulicht sich die Arbeit als der „Flächeninhalt unter dem Graphen“. Ist nun die Kraft nicht mehr stückweise konstant, so wird man intuitiv immer noch den „Flächeninhalt unter dem Graphen“ als wohldefiniert ansehen und als Maß für die entsprechende Arbeit nehmen. Um aus dieser anschaulichen Vorstellung wirklich eine Definition zu machen, werden wir versuchen, eine gegebene Funktion durch stückweise konstante Funktionen zu approximieren. Für jede der approximierenden Funktionen ist dann die Fläche unter dem Graphen bzw. die Arbeit erklärt, und wir werden hoffen, dass diese Werte einem Grenzwert zustreben. Damit dies wirklich der Fall ist und der Grenzwert nicht von der Auswahl der approximierenden Folge abhängt, müssen wir aber eine besonders gute Art der Approximation wählen. Unser Integral wird daher nur für solche Funktionen erklärt sein, die sich in dieser Weise durch stückweise konstante Funktionen approximieren lassen. Das ist aber zunächst ausreichend.

7.1 Regelfunktionen

Im folgenden liegt stets ein festes kompaktes Intervall $[a, b]$ zugrunde.

$B([a, b])$ sei die Menge aller beschränkten Funktionen $f : [a, b] \rightarrow \mathbb{R}$. Für reelle Funktionen f, g mit demselben Definitionsbereich D erklärt man, wie schon erwähnt, $f + g$ und λf ($\lambda \in \mathbb{R}$) durch $(f + g)(x) := f(x) + g(x)$ und $(\lambda f)(x) := \lambda f(x)$ für $x \in D$. Mit $f, g \in B([a, b])$ ist dann auch $f + g \in B([a, b])$ und $\lambda f \in B([a, b])$ für $\lambda \in \mathbb{R}$. Da die Vektorraumaxiome trivialerweise erfüllt sind, ist $B([a, b])$ also ein reeller Vektorraum. Der Nullvektor ist die Funktion, die identisch gleich 0 ist. Wir wollen sie ebenfalls mit 0 bezeichnen.

Definition. Für $f \in B([a, b])$ sei

$$\|f\| := \sup_{x \in [a, b]} |f(x)| := \sup\{|f(x)| \mid x \in [a, b]\}.$$

1.1 Satz. Für $f, g \in B([a, b])$ gilt

- (a) $\|f\| \geq 0$, und $\|f\| = 0$ nur für $f = 0$ (Nullfunktion),
- (b) $\|\lambda f\| = |\lambda| \|f\|$ für $\lambda \in \mathbb{R}$,
- (c) $\|f + g\| \leq \|f\| + \|g\|$.

Beweis. (a) und (b) sind trivial, und (c) folgt aus

$$|(f + g)(x)| = |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\| + \|g\| \blacksquare$$

Eine auf einem reellen Vektorraum V erklärte Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$ mit den Eigenschaften (a), (b), (c) aus Satz 1.1 nennt man eine *Norm* auf V . Die hier auf $B([a, b])$ erklärte Norm heißt *Supremumsnorm*.

Diese Norm verwenden wir nun zur Erklärung eines für unsere Zwecke geeigneten Approximationsbegriffes.

Definition. Seien $f_n, f \in B([a, b])$ ($n \in \mathbb{N}$). Die Folge $(f_n)_{n \in \mathbb{N}}$ konvergiert *gleichmäßig gegen f* , wenn

$$\lim_{n \rightarrow \infty} \|f_n - f\| = 0$$

gilt.

Es gilt also:

$$(f_n)_{n \in \mathbb{N}} \text{ konvergiert gleichmäßig gegen } f \Leftrightarrow$$

$$\forall \varepsilon \in \mathbb{R}^+ \exists n_0 \in \mathbb{N} \forall n \geq n_0 \forall x \in [a, b] : |f_n(x) - f(x)| \leq \varepsilon.$$

Dies kann man sich leicht veranschaulichen: Zu jedem $\varepsilon > 0$ müssen schließlich alle Funktionen der Folge im „ ε -Streifen um f “ liegen.

Die hier erklärte gleichmäßige Konvergenz einer Funktionenfolge ist zu unterscheiden von der Konvergenz schlechthin, unter der man punktweise Konvergenz versteht. Man sagt, dass die Folge $(f_n)_{n \in \mathbb{N}}$ (punktweise) gegen f konvergiert, wenn für jedes $x \in [a, b]$ die Folge $(f_n(x))_{n \in \mathbb{N}}$ gegen $f(x)$ konvergiert. Es gilt also:

$$(f_n)_{n \in \mathbb{N}} \text{ konvergiert gegen } f \Leftrightarrow$$

$$\forall x \in [a, b] \quad \forall \varepsilon \in \mathbb{R}^+ \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : |f_n(x) - f(x)| \leq \varepsilon.$$

Der wesentliche Unterschied liegt darin, dass das n_0 (bei gegebenem ε) hier von x abhängen kann, bei gleichmäßiger Konvergenz (daher der Name) dagegen einheitlich für alle $x \in [a, b]$ wählbar ist.

Definition. $f \in B([a, b])$ heißt *Treppenfunktion*, wenn es Punkte

$$a = x_0 < x_1 < \dots < x_n = b$$

gibt, so dass f auf jedem offenen Teilintervall (x_{k-1}, x_k) konstant ist ($k = 1, \dots, n$), und (x_0, \dots, x_n) heißt dann eine *zu f gehörige Unterteilung* von $[a, b]$. Sei $T([a, b])$ die Menge der Treppenfunktionen auf $[a, b]$.

$T([a, b])$ ist ein Untervektorraum von $B([a, b])$. Zum Beweis ist nur zu zeigen, dass mit $f, g \in T([a, b])$, $\lambda \in \mathbb{R}$, auch $f + g \in T([a, b])$ und $\lambda f \in T([a, b])$ ist. Letzteres ist trivial. Zum Beweis des ersteren sei (x_0, \dots, x_n) eine zu f und (y_0, \dots, y_m) eine zu g gehörige Unterteilung von $[a, b]$. Sei (z_0, \dots, z_k) die Unterteilung, die durch Vereinigung beider Teilpunktfolgen und Ordnen nach der Größe entsteht. Auf jedem Intervall (z_{j-1}, z_j) ist $f + g$ konstant, also ist $f + g \in T([a, b])$.

Definition. Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt *Regelfunktion*, wenn es eine gleichmäßig gegen f konvergierende Folge von Treppenfunktionen auf $[a, b]$ gibt. $R([a, b])$ sei die Menge der Regelfunktionen auf $[a, b]$.

$R([a, b])$ ist ein Untervektorraum von $B([a, b])$. Zunächst ist klar, dass jede Regelfunktion beschränkt ist und dass mit $f \in R([a, b])$ auch $\lambda f \in R([a, b])$ für $\lambda \in \mathbb{R}$ gilt. Seien jetzt $f, g \in R([a, b])$. Es gibt also zwei Folgen $(f_n), (g_n)$ in $T([a, b])$, die gleichmäßig gegen f bzw. g konvergieren. Aus

$$\|(f + g) - (f_n + g_n)\| \leq \|f - f_n\| + \|g - g_n\|$$

folgt dann, dass die Folge $(f_n + g_n)$ von Treppenfunktionen gleichmäßig gegen $f + g$ konvergiert, also ist $f + g \in R([a, b])$.

Es fragt sich, wie man einer gegebenen Funktion „ansehen“ kann, ob sie Regelfunktion ist. Dies wird durch folgenden Satz beantwortet:

1.2 Satz. *Die Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist genau dann Regelfunktion, wenn für alle $c \in [a, b]$ die einseitigen Grenzwerte*

$$\lim_{x \nearrow c} f(x) \quad \text{und} \quad \lim_{x \searrow c} f(x)$$

existieren.

Beweis. „ \Rightarrow “: Sei f Regelfunktion. Es gibt also eine Folge $(t_n)_{n \in \mathbb{N}}$ in $T([a, b])$ mit $\lim_{n \rightarrow \infty} \|t_n - f\| = 0$. Eine Treppenfunktion hat offenbar überall rechtsseitige und linksseitige Grenzwerte. Die Behauptung ergibt sich daher aus der folgenden allgemeineren:

Behauptung. Konvergiert die Folge $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f und hat jedes f_n überall einseitige Grenzwerte, so gilt dies auch für f .

Beweis. (etwa für rechtsseitige Grenzwerte). Sei $c \in [a, b)$. Zu $\varepsilon \in \mathbb{R}^+$ existiert ein $n \in \mathbb{N}$ mit $\|f - f_n\| < \varepsilon/3$. Weiter existiert nach dem Cauchy-Kriterium 2.3 aus Kapitel 4 (das völlig analog für einseitige Grenzwerte gilt) ein $\delta \in \mathbb{R}^+$ mit

$$|f_n(x) - f_n(y)| < \frac{\varepsilon}{3} \quad \text{für alle } x, y \in [a, b] \text{ mit } c < x, y < c + \delta.$$

Für diese x, y gilt also

$$|f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| < \varepsilon.$$

Aus Satz 2.3 aus Kapitel 4 (mit $D := (c, b)$) folgt die Existenz von $\lim_{x \searrow c} f(x)$. ■

„ \Leftarrow “: Es existiere $\lim_{x \searrow c} f(x)$ für alle $c \in [a, b)$ und $\lim_{x \nearrow c} f(x)$ für alle $c \in (a, b]$.

Sei $n \in \mathbb{N}$. Zu jedem $c \in [a, b]$ existiert nach Voraussetzung und nach Satz 2.3 aus Kapitel 4 ein $\delta_c \in \mathbb{R}^+$ mit

$$|f(x) - f(y)| < \frac{1}{n}$$

für alle $x, y \in [a, b]$ mit $c < x, y < c + \delta_c$ oder $c - \delta_c < x, y < c$. Das System $\{U_{\delta_c}(c) \mid c \in [a, b]\}$ ist eine offene Überdeckung von $[a, b]$, enthält also nach dem Überdeckungssatz von Heine-Borel 1.8 aus Kapitel 4 eine endliche Teilüberdeckung. Es gibt also endlich viele Punkte $c_1, \dots, c_k \in [a, b]$ mit zugehörigen Zahlen $\delta_j := \delta_{c_j}$, so dass $[a, b] \subset \bigcup_{j=1}^k U_{\delta_j}(c_j)$ ist. Wir ordnen die Zahlen $c_1, \dots, c_k, c_1 \pm \delta_1, \dots, c_k \pm \delta_k$ (soweit sie in $[a, b]$ liegen) nach der Größe und erhalten eine Unterteilung

$$a = x_0 < x_1 < \dots < x_m = b$$

von $[a, b]$. Wähle $z_j \in (x_{j-1}, x_j)$ ($j = 1, \dots, m$) und setze

$$t_n(x) := \begin{cases} f(z_j) & \text{falls } x \in (x_{j-1}, x_j) \text{ für ein } j \in \{1, \dots, m\}, \\ f(x_j) & \text{falls } x = x_j \text{ für ein } j \in \{0, \dots, m\}. \end{cases}$$

Dann ist $t_n \in T([a, b])$. Sei $x \in [a, b]$. Ist $x = x_j$ für ein j , so gilt $|f(x) - t_n(x)| = 0$. Andernfalls gibt es genau ein j mit $x \in (x_{j-1}, x_j)$. Da (x_{j-1}, x_j) Teilmenge eines offenen Intervalls $(c_i - \delta_i, c_i + \delta_i)$ ist und entweder links oder rechts von c_i liegt, folgt

$$|f(x) - t_n(x)| = |f(x) - f(z_j)| < \frac{1}{n}.$$

Da $x \in [a, b]$ beliebig war, ist $\|f - t_n\| \leq \frac{1}{n}$. Da $n \in \mathbb{N}$ beliebig war, ist damit eine Folge $(t_n)_{n \in \mathbb{N}}$ in $T([a, b])$ gefunden, die gleichmäßig gegen f konvergiert. ■

7.2 Das Integral einer Regelfunktion

Zuerst erklären wir das Integral einer Treppenfunktion:

Definition. Sei $f \in T([a, b])$. Sei durch

$$a = x_0 < x_1 < \dots < x_n = b$$

eine Unterteilung von $[a, b]$ gegeben, so dass f auf (x_{j-1}, x_j) den Wert c_j annimmt. Dann ist die Zahl

$$I(f) := \sum_{j=1}^n c_j (x_j - x_{j-1})$$

unabhängig von der Wahl der Unterteilung; sie heißt *Integral* von f über $[a, b]$ und wird mit

$$\int_a^b f \quad \text{oder} \quad \int_a^b f(x) dx$$

bezeichnet.

Die Behauptung, dass $I(f)$ nur von f und nicht von der speziellen Unterteilung abhängt, ist leicht zu sehen: Ist noch eine andere Unterteilung zu f gegeben, so bilde man die gemeinsame Verfeinerung. Sie entsteht aus der ursprünglichen Unterteilung durch Einführung weiterer Teilpunkte. Es genügt zu zeigen, dass die Summe sich nicht ändert, wenn man einen weiteren Teilpunkt einfügt. Da dieser ein Intervall zerlegt, auf dem f konstant ist, ist das aber klar: ein Summand $c_j(x_j - x_{j-1})$ wird ersetzt durch $c_j(x - x_{j-1}) + c_j(x_j - x)$, was dasselbe ist.

Der folgende Satz bringt die wichtigsten Eigenschaften des Integrals (d.h. der Abbildung $I : T([a, b]) \rightarrow \mathbb{R}$) zum Ausdruck.

2.1 Satz. Die Abbildung $I : T([a, b]) \rightarrow \mathbb{R} : f \mapsto \int_a^b f$, ist ein lineares Funktional, d.h. es gilt

$$(a) \int_a^b (f + g) = \int_a^b f + \int_a^b g \text{ für alle } f, g \in T([a, b]),$$

$$(b) \int_a^b (\lambda f) = \lambda \int_a^b f \text{ für alle } f \in T([a, b]) \text{ und alle } \lambda \in \mathbb{R}.$$

Dieses Funktional ist monoton, d.h. es gilt

$$(c) \text{ Aus } f \leq g \text{ (d.h. } f(x) \leq g(x) \forall x \in [a, b]) \text{ folgt } \int_a^b f \leq \int_a^b g.$$

Ferner gilt

$$(d) \left| \int_a^b f \right| \leq (b - a) \|f\| \text{ für } f \in T([a, b]),$$

das Funktional ist also beschränkt.

Ein lineares Funktional φ auf einem normierten Vektorraum $(V, \|\cdot\|)$ heißt *beschränkt*, wenn es eine Konstante $K \in \mathbb{R}$ gibt mit $|\varphi(x)| \leq K\|x\|$ für alle $x \in V$.

Beweis. Sind $f, g \in T([a, b])$ gegeben, so wähle man zu jeder dieser Funktionen eine zugehörige Unterteilung von $[a, b]$ und bilde dann die gemeinsame Verfeinerung. Sie ist eine zu f, g und $f + g$ gehörige Unterteilung. Die Behauptungen (a) und (c) folgen jetzt sofort aus der Definition, (b) ist trivial. Schließlich gilt

$$\begin{aligned} \left| \int_a^b f \right| &= \left| \sum_{j=1}^n c_j (x_j - x_{j-1}) \right| \leq \sum_{j=1}^n |c_j| (x_j - x_{j-1}) \\ &\leq \max\{|c_1|, \dots, |c_n|\} \sum_{j=1}^n (x_j - x_{j-1}) \leq \|f\| (b - a). \quad \blacksquare \end{aligned}$$

Nun setzen wir die Abbildung I unter Beibehaltung ihrer Eigenschaften fort auf den Raum $R([a, b])$ der Regelfunktionen.

Definition. Sei $f \in R([a, b])$. Sei $(t_n)_{n \in \mathbb{N}}$ eine Folge in $T([a, b])$, die gleichmäßig gegen f konvergiert. Dann existiert

$$\int_a^b f := \lim_{n \rightarrow \infty} \int_a^b t_n$$

und ist unabhängig von der Wahl der Folge (t_n) . $\int_a^b f$ heißt das *Integral* von f über $[a, b]$ und wird auch mit

$$\int_a^b f(x) dx$$

bezeichnet.

Beweis der Wohldefiniertheit. Sei $\varepsilon \in \mathbb{R}^+$. Es gibt ein $n_0 \in \mathbb{N}$ mit $\|t_n - f\| < \varepsilon/2(b-a)$ für $n \geq n_0$. Für $m, n \geq n_0$ gilt also

$$\|t_m - t_n\| \leq \|t_m - f\| + \|t_n - f\| < \frac{\varepsilon}{b-a}$$

und daher nach Satz 2.1

$$\left| \int_a^b t_m - \int_a^b t_n \right| = \left| \int_a^b (t_m - t_n) \right| \leq \|t_m - t_n\|(b-a) < \varepsilon.$$

Also ist $(\int_a^b t_n)_{n \in \mathbb{N}}$ eine Cauchyfolge reeller Zahlen und daher konvergent.

Sei $(u_n)_{n \in \mathbb{N}}$ eine weitere Folge in $T([a, b])$, die gleichmäßig gegen f konvergiert. Sei $\varepsilon \in \mathbb{R}^+$. Es gibt ein $n_0 \in \mathbb{N}$ mit

$$\|f - t_n\| < \frac{\varepsilon}{2(b-a)} \quad \text{und} \quad \|f - u_n\| < \frac{\varepsilon}{2(b-a)} \quad \text{für } n \geq n_0.$$

Für $n \geq n_0$ gilt also

$$\|t_n - u_n\| \leq \|t_n - f\| + \|f - u_n\| < \frac{\varepsilon}{b-a}$$

und daher nach Satz 2.1

$$\left| \int_a^b t_n - \int_a^b u_n \right| \leq \|t_n - u_n\|(b-a) < \varepsilon.$$

Also ist $\left(\int_a^b t_n - \int_a^b u_n \right)_{n \in \mathbb{N}}$ eine Nullfolge und somit

$$\lim_{n \rightarrow \infty} \int_a^b t_n = \lim_{n \rightarrow \infty} \int_a^b u_n. \quad \blacksquare$$

Ist speziell f eine Treppenfunktion, so liefert die neue Definition denselben Integralwert wie die alte, wie sich aus der Unabhängigkeit des Grenzwertes von der approximierenden Folge ergibt.

2.2 Satz. Für alle $f, g \in R([a, b])$ und $\lambda \in \mathbb{R}$ gilt

$$\begin{aligned} (a) \quad & \int_a^b (f + g) = \int_a^b f + \int_a^b g, \\ (b) \quad & \int_a^b (\lambda f) = \lambda \int_a^b f, \\ (c) \quad & f \leq g \Rightarrow \int_a^b f \leq \int_a^b g, \\ (d) \quad & \left| \int_a^b f \right| \leq \int_a^b |f| \leq (b - a) \|f\|. \end{aligned}$$

Beweis. Seien $(t_n)_{n \in \mathbb{N}}, (u_n)_{n \in \mathbb{N}}$ Folgen in $T([a, b])$, die gleichmäßig gegen f bzw. g konvergieren. Dann konvergiert $(t_n + u_n)_{n \in \mathbb{N}}$ gleichmäßig gegen $f + g$, also gilt

$$\begin{aligned} \int_a^b (f + g) &= \lim_{n \rightarrow \infty} \int_a^b (t_n + u_n) = \lim_{n \rightarrow \infty} \left[\int_a^b t_n + \int_a^b u_n \right] \\ &= \lim_{n \rightarrow \infty} \int_a^b t_n + \lim_{n \rightarrow \infty} \int_a^b u_n = \int_a^b f + \int_a^b g. \end{aligned}$$

Damit ist (a) bewiesen. (b) folgt analog.

Ist $f \leq g$, so gelten für jedes $\varepsilon \in \mathbb{R}^+$ für fast alle n die Ungleichungen $t_n \leq f + \varepsilon$ und $u_n \geq g - \varepsilon$, also $t_n \leq u_n + 2\varepsilon$, woraus

$$\begin{aligned} \int_a^b f &= \lim_{n \rightarrow \infty} \int_a^b t_n \leq \lim_{n \rightarrow \infty} \int_a^b (u_n + 2\varepsilon) \\ &= \lim_{n \rightarrow \infty} \left(\int_a^b u_n + 2\varepsilon(b - a) \right) = \int_a^b g + 2\varepsilon(b - a) \end{aligned}$$

folgt. Da $\varepsilon > 0$ beliebig war, folgt (c).

Zu (d) ist zunächst zu bemerken, dass mit f auch $|f|$ eine Regelfunktion ist. Für Treppenfunktionen ist das klar, und daraus folgt es allgemein: Wenn (t_n) gleichmäßig gegen f konvergiert, konvergiert $(|t_n|)$ wegen

$$\left| |t_n| - |f| \right| \leq \|t_n - f\|$$

(klar wegen $||x| - |y|| \leq |x - y|$) gleichmäßig gegen $|f|$.

Ferner folgt aus

$$||t_n| - |f|| \leq \|t_n - f\|$$

(wegen $\|t_n\| \leq \|t_n - f\| + \|f\|$ etc.) dass $(\|t_n\|)$ gegen $\|f\|$ konvergiert. Damit ergibt sich

$$\left| \int_a^b f \right| = \left| \lim \int_a^b t_n \right| = \lim \left| \int_a^b t_n \right| \leq \lim \int_a^b |t_n| = \int_a^b |f|$$

und

$$\lim \int_a^b |t_n| \leq \lim (b-a) \|t_n\| = (b-a) \|f\|. \quad \blacksquare$$

Ist $f \in R([a, b])$ und $[c, d] \subset [a, b]$, so ist die Einschränkung $f|_{[c, d]}$ offenbar eine Regelfunktion. Ihr Integral ist also definiert; wir bezeichnen es mit

$$\int_c^d f.$$

Zur Ergänzung definiert man für $c \leq d$

$$\int_d^c f := - \int_c^d f, \quad \int_c^c f := 0.$$

Mit diesen Festsetzungen gilt dann allgemein die im folgenden Satz ausgedrückte Additivitätseigenschaft des Integrals hinsichtlich der Integrationsbereiche.

2.3 Satz. Für $f \in R([a, b])$ und $u, v, w \in [a, b]$ gilt

$$\int_u^v f + \int_v^w f = \int_u^w f.$$

Beweis. Zunächst sei $u < v < w$. Die Behauptung folgt für Treppenfunktionen aus der Definition des Integrals, allgemein dann durch Approximation. Die übrigen Fälle ergeben sich hieraus und mit den obigen Festsetzungen. Ist z.B. $u < w < v$, so ist

$$\int_u^v f + \int_v^w f = \int_u^v f - \int_w^v f = \int_u^w f.$$

Analog schließt man in den anderen Fällen. ■

Wir schließen diesen Abschnitt mit zwei manchmal nützlichen weiteren Eigenschaften des Integrals.

2.4 Satz. Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge von Regelfunktionen, die gleichmäßig gegen eine Funktion f konvergiert. Dann ist f Regelfunktion, und es gilt

$$\int_a^b f = \lim_{n \rightarrow \infty} \int_a^b f_n.$$

Beweis. Es ist leicht zu sehen, dass f Regelfunktion ist. Aus Satz 2.2 folgt

$$\left| \int_a^b f - \int_a^b f_n \right| \leq (b-a) \|f - f_n\|$$

und daraus die Behauptung. ■

Beispiel. Die Aussage von Satz 2.4 wird i.a. falsch, wenn „gleichmäßig“ ersetzt wird durch „punktweise“: Setze

$$f_n(x) := \begin{cases} n & \text{für } 0 < x < \frac{1}{n}, \\ 0 & \text{für } x = 0 \text{ und für } \frac{1}{n} \leq x \leq 1. \end{cases}$$

Dann ist f_n eine Treppenfunktion auf $[0, 1]$, und die Folge $(f_n)_{n \in \mathbb{N}}$ konvergiert punktweise gegen 0. Es gilt

$$\lim_{n \rightarrow \infty} \int_0^1 f_n = 1 \neq 0 = \int_0^1 \lim_{n \rightarrow \infty} f_n.$$

2.5 Satz (Mittelwertsatz der Integralrechnung). Seien $f, g \in R([a, b])$, sei f stetig und $g \geq 0$. Dann existiert ein $c \in [a, b]$ mit

$$\int_a^b fg = f(c) \int_a^b g.$$

(Speziell für $g \equiv 1$ also $\int_a^b f = f(c)(b-a)$.)

Beweis. Die stetige Funktion f nimmt nach Satz 3.2 aus Kapitel 4 auf dem kompakten Intervall $[a, b]$ ein Minimum m und ein Maximum M an. Aus $m \leq f \leq M$ und $g \geq 0$ folgt $mg \leq fg \leq Mg$ und daraus nach Satz 2.2

$$m \int_a^b g \leq \int_a^b fg \leq M \int_a^b g.$$

Es ist also $\int_a^b fg = \alpha \int_a^b g$ mit einer Zahl $\alpha \in [m, M]$. Nach dem Zwischenwertsatz 3.4 aus Kapitel 4 gibt es eine Zahl $c \in [a, b]$ mit $f(c) = \alpha$. ■

7.3 Integration und Differentiation

Die zu Anfang dieses Kapitels erläuterte Interpretation des Integrals als verallgemeinerte Summe ist nicht die einzige Motivation für die Einführung des Integralbegriffs. Eine andere liegt darin, dass man Integration als Umkehrung von Differentiation ansehen kann. dass ein solcher Zusammenhang bestehen muß, macht man sich leicht schon anschaulich klar, wenn man das Integral einer etwa stetigen positiven Funktion f als „Fläche unter dem Graphen“ interpretiert und dann für kleine positive h die Größen

$$\int_a^{x+h} f(t) dt - \int_a^x f(t) dt \quad \text{und} \quad f(x)h$$

miteinander vergleicht. Der Zusammenhang zwischen Integration und Differentiation ist Gegenstand des Hauptsatzes der Differential- und Integralrechnung.

Definition. Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Eine Funktion $F : [a, b] \rightarrow \mathbb{R}$ heißt *Stammfunktion* von f , wenn F differenzierbar und $F' = f$ ist.

Nicht jede Funktion besitzt eine Stammfunktion, wohl aber jede stetige Funktion.

3.1 Satz (Hauptsatz der Differential- und Integralrechnung). *Die Funktion $f : [a, b] \rightarrow \mathbb{R}$ sei stetig. Die durch*

$$F(x) := \int_a^x f = \int_a^x f(t) dt, \quad x \in [a, b],$$

erklärte Funktion F ist Stammfunktion von f .

Jede Stammfunktion G von f ist von der Form $G = F + c$ mit einer Konstanten c .

Beweis. Sei $x \in [a, b]$. Für $h \in \mathbb{R}$ mit $x + h \in [a, b]$ gilt

$$\begin{aligned} |F(x+h) - F(x) - f(x)h| &= \left| \int_a^{x+h} f(t) dt - \int_a^x f(t) dt - f(x)h \right| \\ &= \left| \int_x^{x+h} f(t) dt - \int_x^{x+h} f(x) dt \right| = \left| \int_x^{x+h} (f(t) - f(x)) dt \right| \\ &\leq \int_x^{x+h} |f(t) - f(x)| dt, \end{aligned}$$

wobei Satz 2.3 benutzt worden ist. Sei nun $\varepsilon > 0$ gegeben. Da f auf $[a, b]$ nach Satz 3.5 Kapitel 4 gleichmäßig stetig ist, existiert ein $\delta > 0$, so dass $|f(t) - f(x)| < \varepsilon$ für alle $x, t \in [a, b]$ mit $|t - x| < \delta$. Für $|h| < \delta$ folgt also

$$|F(x+h) - F(x) - f(x)h| \leq |h|\varepsilon.$$

Damit ist nach Definition der Ableitung

$$F'(x) = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = f(x)$$

gezeigt, also ist F Stammfunktion von f .

Ist auch G Stammfunktion von f , so ist $F' = f = G'$, also $(F - G)' = 0$ auf $[a, b]$; nach Satz 2.5 aus Kapitel 6 ist also $F - G$ konstant. ■

Der Hauptsatz der Differential- und Integralrechnung enthält zwei wichtige Aussagen: Die erste Aussage ist, dass das Integral einer stetigen Funktion nach der oberen Grenze differenzierbar ist und dass die Ableitung den Integranden ergibt. Die zweite Aussage ist, dass die Stammfunktion eindeutig ist. Beides zusammen hat die folgende wichtige Konsequenz. Ist F eine Stammfunktion der stetigen Funktion f , so ist mit einer Konstanten c

$$F(b) - F(a) = \left(c + \int_a^b f \right) - \left(c + \int_a^a f \right) = \int_a^b f.$$

Man schreibt auch $F(b) - F(a) = [F(x)]_{x=a}^{x=b} = [F(x)]_a^b$. Wir können für stetiges f also das Integral $\int_a^b f$ berechnen, wenn wir eine Stammfunktion

von f kennen. Wir kommen hierauf im nächsten Abschnitt zurück. Zunächst, ebenfalls im Hinblick auf die Berechnung von Integralen, noch eine wichtige Folgerung.

Bemerkung. Die Funktion $h : [a, b] \rightarrow \mathbb{R}$ sei stetig differenzierbar und es sei $H(x) := \int_a^x h'(t) dt$. Dann folgt aus dem Hauptsatz, dass $H' = h'$ und somit gibt es eine Konstante c mit $H = h + c$. Dies und die Folgerung des Hauptsatzes implizieren

$$\int_a^b h'(t) dt = H(b) - H(a) = h(b) - h(a) = [h(x)]_a^b.$$

3.2 Satz (Partielle Integration oder Produktintegration). *Für stetig differenzierbare Funktionen $f, g : [a, b] \rightarrow \mathbb{R}$ gilt*

$$\int_a^b f(x)g'(x) dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x) dx.$$

Beweis. Nach der Produktregel gilt

$$(fg)' = f'g + fg',$$

also gilt nach obiger Bemerkung

$$\int_a^b fg' + \int_a^b f'g = \int_a^b (fg)' = [f(x)g(x)]_a^b. \quad \blacksquare$$

Hiermit sind wir insbesondere in der Lage, einen neuen Beweis der Taylorformel mit einer anderen nützlichen Darstellung des Restgliedes anzugeben:

3.3 Satz (Taylorsche Formel mit Restglied in Integralform). *Sei $D \subset \mathbb{R}$ ein Intervall und $f : D \rightarrow \mathbb{R}$ $(n + 1)$ -mal stetig differenzierbar. Dann gilt für $a, x \in D$*

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x - a)^k + R_{n+1}(x)$$

mit

$$R_{n+1}(x) = \frac{1}{n!} \int_a^x f^{(n+1)}(t)(x - t)^n dt.$$

Beweis. Wir beweisen die Behauptung durch Induktion nach n . Für $n = 0$ lautet die Behauptung

$$f(x) = f(a) + \int_a^x f'(t) dt,$$

sie ist richtig nach Satz 3.1. Sei die Behauptung bewiesen für $n - 1$, also

$$R_n(x) = \frac{1}{(n-1)!} \int_a^x f^{(n)}(t)(x-t)^{n-1} dt = \int_a^x f^{(n)}(t)g'(t) dt$$

mit

$$g(t) := -\frac{(x-t)^n}{n!}.$$

Partielle Integration ergibt

$$\begin{aligned} R_n(x) &= \left[-f^{(n)}(t) \frac{(x-t)^n}{n!} \right]_{t=a}^{t=x} + \frac{1}{n!} \int_a^x f^{(n+1)}(t)(x-t)^n dt \\ &= \frac{f^{(n)}(a)}{n!} (x-a)^n + \frac{1}{n!} \int_a^x f^{(n+1)}(t)(x-t)^n dt, \end{aligned}$$

woraus die Behauptung folgt. ■

7.4 Berechnung von Integralen

Der Hauptsatz der Differential- und Integralrechnung bietet die Möglichkeit, die Integrale vieler spezieller Funktionen explizit anzugeben. In anderen Fällen wird man sich zur numerischen Berechnung von Integralen verschiedener Näherungsmethoden bedienen müssen. Wir wollen im folgenden einige Verfahren zur Integralberechnung und Rechenregeln für Integrale zusammenstellen.

Prinzipiell kann man natürlich zur näherungsweise Integralberechnung unmittelbar die Integraldefinition heranziehen. In manchen einfachen Fällen erhält man so sogar den genauen Wert. Eine ungleich bequemere Methode der Integralberechnung liefert aber der Hauptsatz der Differential- und Integralrechnung. Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig und kennen wir eine Stammfunktion F von f , so folgt aus der Folgerung von Satz 3.1

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

Um zum Beispiel $\int_0^a \cos x \, dx$ zu berechnen, erinnern wir uns, dass $\sin' = \cos$ ist und erhalten

$$\int_0^a \cos x \, dx = \sin a - \sin 0 = \sin a.$$

Allgemein können wir uns eine Liste machen von allen stetigen Funktionen, die uns als Ableitungen anderer Funktionen begegnet sind. Von diesen Funktionen kennen wir also Stammfunktionen und können daher die Integrale explizit angeben.

Beispiele (F Stammfunktion von f).

$f(x)$	$F(x)$
$e^{ax} \ (a \neq 0)$	$\frac{1}{a}e^{ax}$
$\frac{1}{x}$	$\ln x$
$x^\alpha \ (\alpha \neq -1)$	$\frac{1}{\alpha+1}x^{\alpha+1}$
$a^x \ (a \neq 1)$	$\frac{1}{\ln a}a^x$
$\cos x$	$\sin x$
$\sin x$	$-\cos x$
$\frac{1}{\cos^2 x}$	$\tan x$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$
$\frac{1}{1+x^2}$	$\arctan x$

Viele weitere Stammfunktionen findet man (unter der Überschrift „Unbestimmte Integrale“) in den Formelsammlungen und Integraltafeln. Hat man durch eigene Berechnungen eine Stammfunktion explizit ermittelt, so sollte man stets die Probe durch Differenzieren machen. Es sei darauf hingewiesen, dass häufig für Stammfunktionen eine bequeme, aber nicht ganz korrekte Schreibweise üblich ist, wie zum Beispiel

$$\int \frac{1}{1+x^2} dx = \arctan x.$$

Korrekt wäre

$$\int_a^x \frac{1}{1+t^2} dt = \arctan x + c,$$

aber die obige Schreibweise ist so verbreitet, dass wir sie auch benutzen werden. Man bezeichnet Ausdrücke wie

$$\int \frac{1}{1+x^2} dx$$

auch als „unbestimmte Integrale“; das sind also nichts anderes als Stammfunktionen.

Häufig kann man unbekannte Integrale durch Umformungen auf bekannte zurückführen. Diese Umformungsregeln erhält man, wenn man Differentiationsregeln mit Hilfe des Hauptsatzes der Differential- und Integralrechnung in Integrationsregeln übersetzt. Die erste solche Regel ist in Satz 3.2 enthalten: Für stetig differenzierbare Funktionen f, g gilt

$$\int_a^b f g' = [f g]_a^b - \int_a^b f' g.$$

Vor der Behandlung von Beispielen eine Bemerkung: Wenn wir bei gegebenem f das Integral

$$\int_a^b f$$

für alle b (des Definitionsbereiches von f) bestimmen können, haben wir damit natürlich eine Stammfunktion von f gefunden. Dies wird bei den Beispielen stets der Fall sein.

Beispiele. (1)

$$\begin{aligned} \int_a^b x e^x dx &= \int_a^b f(x) g'(x) dx && \begin{cases} f(x) = x, \\ g(x) = e^x \end{cases} \\ &= [x e^x]_a^b - \int_a^b 1 \cdot e^x dx = [x e^x]_a^b - [e^x]_a^b. \end{aligned}$$

Mit der üblichen Schreibweise für unbestimmte Integrale, die wir von nun an verwenden wollen, lautet das Ergebnis also

$$\int x e^x dx = x e^x - e^x.$$

Das bedeutet, wie gesagt, dass $F(x) := x e^x - e^x$ eine Stammfunktion von $f(x) := x e^x$ ist. Die Probe durch Differenzieren (die man immer machen sollte) bestätigt das.

Ganz ähnlich:

$$\begin{aligned} \int x \sin x \, dx & \quad \left[\begin{array}{l} f(x) = x \\ g(x) = -\cos x \end{array} \right] \\ &= -x \cos x + \int \cos x \, dx \\ &= -x \cos x + \sin x. \end{aligned}$$

Zwei Kunstgriffe sind im Zusammenhang mit partieller Integration oft nützlich: Man kann immer den Faktor 1 einfügen und als g' ansehen. Zweitens kann man eventuell durch partielle Integration eine Gleichung für das fragliche Integral erhalten und hieraus das Integral berechnen.

$$\begin{aligned} (2) \quad \int \ln x \, dx &= \int f(x)g'(x) \, dx \quad \left[\begin{array}{l} f(x) = \ln x \\ g(x) = x \end{array} \right] \\ &= x \ln x - \int \frac{1}{x} x \, dx = x \ln x - x. \\ (3) \quad \int \frac{1}{x} \ln x \, dx &= \int f(x)g'(x) \, dx \quad \left[\begin{array}{l} f(x) = \ln x \\ g(x) = \ln x \end{array} \right] \\ &= (\ln x)^2 - \int \frac{1}{x} \ln x \, dx. \end{aligned}$$

Also ist

$$\int \frac{1}{x} \ln x \, dx = \frac{1}{2}(\ln x)^2.$$

Manchmal muß man partielle Integration mehrfach anwenden, um zu einer Gleichung zu kommen, aus der das Integral bestimmbar ist:

$$\begin{aligned} (4) \quad \int e^x \sin x \, dx &= -e^x \cos x + \int e^x \cos x \, dx \\ &= -e^x \cos x + e^x \sin x - \int e^x \sin x \, dx. \end{aligned}$$

Also ist

$$\int e^x \sin x \, dx = \frac{1}{2}e^x(\sin x - \cos x).$$

Manchmal erfordert die „partielle Integration“, also die Bestimmung von g zu g' in $\int fg'$, ihrerseits partielle Integration:

$$(5) \quad \int (\ln x)^2 \, dx = \int \ln x \ln x \, dx.$$

Wir wissen nach Beispiel (2), dass

$$\frac{d}{dx}(x \ln x - x) = \ln x$$

ist, also ergibt sich

$$\begin{aligned} \int \ln x \ln x \, dx &= \ln x(x \ln x - x) - \int \frac{1}{x}(x \ln x - x) \, dx \\ &= \ln x(x \ln x - x) - \int \ln x \, dx + \int 1 \, dx \\ &= \ln x(x \ln x - x) - (x \ln x - x) + x \\ &= x(\ln x)^2 - 2x(\ln x) + 2x. \end{aligned}$$

Schließlich ermöglicht es partielle Integration bisweilen, zu einer Rekursionsformel zu kommen. Als Beispiel betrachten wir

$$(6) \quad I_m := \int_a^b \sin^m x \, dx \quad \text{für } m \in \mathbb{N}.$$

Für $m \geq 2$ ist

$$\begin{aligned} I_m &= - \int_a^b \sin^{m-1} x \cos' x \, dx \\ &= [-\cos x \sin^{m-1} x]_a^b + (m-1) \int_a^b \sin^{m-2} x \cos^2 x \, dx \\ &= [-\cos x \sin^{m-1} x]_a^b + (m-1) \int_a^b \sin^{m-2} x (1 - \sin^2 x) \, dx \\ &= [-\cos x \sin^{m-1} x]_a^b + (m-1)I_{m-2} - (m-1)I_m, \end{aligned}$$

also

$$I_m = -\frac{1}{m} [\cos x \sin^{m-1} x]_a^b + \frac{m-1}{m} I_{m-2}.$$

Dies ist eine Rekursionsformel, mit deren Hilfe sich I_m berechnen läßt, denn I_0 und I_1 sind wohlbekannt.

Die zweite wichtige (und vielseitiger anwendbare) Integralumformung ist die Substitutionsregel. Man erhält sie, kurz gesagt, durch Integration der Kettenregel.

4.1 Satz (Substitutionsregel). *Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und $g : [c, d] \rightarrow [a, b]$ stetig differenzierbar. Dann gilt*

$$\int_c^d f(g(t))g'(t) dt = \int_{g(c)}^{g(d)} f(x) dx.$$

Ist $g : [c, d] \rightarrow [a, b]$ außerdem bijektiv, so gilt

$$\int_a^b f(x) dx = \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(t))g'(t) dt.$$

Beweis. Zu f gibt es nach Satz 3.1 eine Stammfunktion F auf $[a, b]$. Nach der Kettenregel 1.5 aus Kapitel 6 gilt

$$(F \circ g)'(t) = F'(g(t))g'(t) = f(g(t))g'(t)$$

für $t \in [c, d]$, also

$$\begin{aligned} \int_c^d f(g(t))g'(t) dt &= \int_c^d (F \circ g)'(t) dt = (F \circ g)(d) - (F \circ g)(c) \\ &= F(g(d)) - F(g(c)) = \int_{g(c)}^{g(d)} F'(x) dx = \int_{g(c)}^{g(d)} f(x) dx. \end{aligned}$$

Sei jetzt $g : [c, d] \rightarrow [a, b]$ bijektiv. Aus den Eigenschaften von g folgert man leicht, dass entweder $g(c) = a$, $g(d) = b$ oder $g(c) = b$, $g(d) = a$ gelten muß. Im ersten Fall ist dann die Behauptung klar; im zweiten Fall ergibt sie sich aus

$$\int_b^a f(x) dx = \int_{g(c)}^{g(d)} f(x) dx = \int_{g^{-1}(b)}^{g^{-1}(a)} f(g(t))g'(t) dt$$

und Satz 2.3. ■

Die Anwendung dieser Regel wollen wir zunächst an einem sehr einfachen Beispiel studieren. Es sei das Integral

$$\int_c^d \sin^3 t \cos t dt$$

zu berechnen. Es fällt sofort ins Auge, dass \cos die Ableitung von \sin ist, also wird man versuchen, den Integranden in der Form $f(g(t))g'(t)$ mit $g(t) = \sin t$

zu schreiben. Man hat also $f(x) = x^3$ zu wählen. Dann ergibt die Substitutionsregel

$$\begin{aligned} \int_c^d \sin^3 t \cos t \, dt &= \int_c^d f(g(t))g'(t) \, dt \\ &= \int_{\sin c}^{\sin d} f(x) \, dx = \int_{\sin c}^{\sin d} x^3 \, dx = \left[\frac{1}{4}x^4 \right]_{\sin c}^{\sin d} \\ &= \frac{1}{4} \sin^4 d - \frac{1}{4} \sin^4 c. \end{aligned}$$

Dies war die korrekte und einwandfrei geschriebene Anwendung der Substitutionsregel. In der Praxis verwendet man eine nicht ganz einwandfreie, aber leicht zu merkende Schreibweise, die wir jetzt an demselben Beispiel erläutern wollen. Man würde x statt g schreiben und sagen, dass man „die Substitution

$$\sin t = x$$

macht“. Hierbei muß man im Kopf haben, dass x als Funktion von t aufzufassen ist. Man kann sie nach t differenzieren:

$$\frac{dx}{dt} = \cos t.$$

Schreibt man dies formal in der Weise

$$dx = \cos t \, dt$$

und setzt dies in das Integral ein, so erhält man

$$\int_c^d \sin^3 t \cos t \, dt = \int x^3 \, dx,$$

wobei noch die Grenzen einzusetzen sind nach der Regel: Ist $t = c$, so ist $x = \sin c$ usw., also

$$= \int_{\sin c}^{\sin d} x^3 \, dx.$$

Man kommt bei diesen formalen Umformungen also zum richtigen Ergebnis. Bezeichnet nun F eine Stammfunktion des Integranden (nämlich $F(x) = \frac{1}{4}x^4$), so ist dieses Integral

$$= [F(x)]_{x=\sin c}^{x=\sin d} = [F(\sin t)]_{t=c}^{t=d}$$

Man erhält also das richtige Ergebnis, wenn man am Ende wieder $x = \sin t$ einsetzt. Da hier d beliebig ist, haben wir in Wirklichkeit nicht nur ein bestimmtes Integral gefunden, sondern eine Stammfunktion.

Die folgenden Beispiele zur Anwendung der Substitutionsregel behandeln wir immer in dieser sehr einprägsamen Form. Es sollte keine Mühe bereiten, die Umformungen auch in korrekter Weise darzustellen. Die Merkregel ist aber zur Auffindung passender Substitutionen (für die es kein Patentrezept gibt) sehr hilfreich.

Es versteht sich in den folgenden Beispielen von selbst, dass die Integranden i.a. nur auf passenden Intervallen erklärt sind; dies wird nicht jeweils im Einzelnen angegeben. Da wir in Wahrheit immer Stammfunktionen bestimmen, geben wir, der Konvention entsprechend, keine Integrationsgrenzen an.

$$(1) \int \tan t \, dt = \int \frac{\sin t}{\cos t} \, dt \quad \left[\text{Substitution: } \begin{array}{l} x = \cos t \\ dx = -\sin t \, dt \end{array} \right]$$

$$\int \tan t \, dt = - \int \frac{1}{x} \, dx = -\ln |x| = -\ln |\cos t|.$$

Das läßt sich verallgemeinern. Zu bestimmen sei etwa

$$\int \frac{g'(t)}{g(t)} \, dt \quad \text{mit } g > 0. \quad \left[\text{Substitution: } \begin{array}{l} x = g(t) \\ dx = g'(t) \, dt \end{array} \right]$$

$$\int \frac{g'(t)}{g(t)} \, dt = \int \frac{1}{x} \, dx = \ln x = \ln g(t).$$

$$(2) \int \frac{x}{(1+x^2)^2} \, dx \quad \left[\text{Substitution } \begin{array}{l} 1+x^2 = t \\ 2x \, dx = dt \end{array} \right]$$

$$= \frac{1}{2} \int \frac{1}{t^2} \, dt = -\frac{1}{2} \frac{1}{t} = -\frac{1}{2} \frac{1}{1+x^2}.$$

(3) Ein etwas schwierigeres Beispiel, bei dem das vorstehende nach algebraischer Umformung und partieller Integration benutzt wird:

$$\int \frac{1}{(1+x^2)^2} \, dx = \int \left(\frac{1}{1+x^2} - \frac{x^2}{(1+x^2)^2} \right) dx$$

$$= \arctan x - \int \frac{x^2}{(1+x^2)^2} \, dx,$$

$$\begin{aligned}
\int \frac{x^2}{(1+x^2)^2} dx &= \int x \cdot \frac{x}{(1+x^2)^2} dx && \text{(vgl. (2))} \\
&= -\frac{1}{2} \frac{x}{1+x^2} + \frac{1}{2} \int \frac{1}{1+x^2} dx \\
&= -\frac{1}{2} \frac{x}{1+x^2} + \frac{1}{2} \arctan x \\
\Rightarrow \int \frac{1}{(1+x^2)^2} dx &= \frac{1}{2} \frac{x}{1+x^2} + \frac{1}{2} \arctan x.
\end{aligned}$$

$$(4) \int_a^b \sqrt{1-x^2} dx \quad \text{mit } [a, b] \subset [-1, 1]$$

Hier substituieren wir $x = \sin t$. Im Integranden fassen wir also x auf als $x = \sin(\arcsin x)$ und substituieren dann $\arcsin x = t$. Hierbei ist wesentlich, dass die Abbildung

$$\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow [-1, 1]$$

bijektiv ist (und somit eine Umkehrabbildung besitzt). Im Grunde benutzen wir also die zweite, speziellere Form der Substitutionsregel 4.1. Wir erhalten

$$\begin{aligned}
\int \sqrt{1-x^2} dx &= \int \sqrt{1-\sin^2 t} \cos t dt \\
&= \int \cos^2 t dt && \text{[partielle Integration]} \\
&= \frac{1}{2} \sin t \cos t + \frac{1}{2} t \\
&= \frac{1}{2} \sin t \sqrt{1-\sin^2 t} + \frac{1}{2} t \\
&= \frac{1}{2} x \sqrt{1-x^2} + \frac{1}{2} \arcsin x.
\end{aligned}$$

Speziell ergibt sich (mit einem Grenzübergang)

$$\int_{-1}^1 \sqrt{1-x^2} dx = \left[\frac{1}{2} x \sqrt{1-x^2} + \frac{1}{2} \arcsin x \right]_{-1}^1 = \frac{\pi}{2}.$$

Damit ist die Deutung der Zahl π als Flächeninhalt des Kreises vom Radius 1 gewonnen.

- (5) Häufig muß man Substitution und partielle Integration kombinieren, wie im folgenden Beispiel.

$$\begin{aligned} \int \arctan x \, dx &= x \arctan x - \int \frac{x}{1+x^2} \, dx \\ \int \frac{x}{1+x^2} \, dx &\quad \left[\text{Substitution } \begin{array}{l} x^2 = t \\ 2x \, dx = dt \end{array} \right] \\ &= \frac{1}{2} \int \frac{1}{1+t} \, dt = \frac{1}{2} \ln(1+t) = \frac{1}{2} \ln(1+x^2), \end{aligned}$$

also

$$\int \arctan x \, dx = x \arctan x - \frac{1}{2} \ln(1+x^2).$$

Numerische Integration

Schon für relativ einfache Funktionen, die aus den speziellen, elementaren Funktionen zusammengesetzt sind, ist es oft prinzipiell unmöglich, explizit die Stammfunktion durch elementare Funktionen auszudrücken. Integrale über solche Funktionen können nur näherungsweise (mit beliebiger Genauigkeit) berechnet werden. Hierzu gibt es verschiedene Verfahren. Natürlich kann man grundsätzlich auf die Integraldefinition zurückgehen, also die gegebene Funktion durch Treppenfunktionen approximieren. Das wäre jedoch meist mit unnötig viel Rechenaufwand verbunden. Es gibt bessere Verfahren, die sich vor allem dadurch auszeichnen, dass bei Kenntnis von Ableitungen der Funktion günstige Fehlerabschätzungen möglich sind. Wir wollen ein solches Verfahren betrachten.

Zur Berechnung von $\int_a^b f(x) \, dx$ wird man im allgemeinen so vorgehen, dass man zuerst das Intervall $[a, b]$ in genügend kleine Teilintervalle zerlegt und dann in jedem Teilintervall die Funktion f ersetzt durch eine Funktion, deren Integral man berechnen kann. Die naheliegendste Ersatzfunktion ist eine affine. Wir wollen zunächst nur ein Teilintervall betrachten, etwa das Intervall $[-h, h]$. Wir ersetzen f in $[-h, h]$ durch die affine Funktion, die in den Endpunkten des Intervalls dieselben Werte annimmt wie f . Deren Integral ist offenbar

$$h(f(-h) + f(h)) =: A.$$

Das Problem ist jetzt, eine Fehlerabschätzung, d.h. eine obere Schranke für

$$\left| \int_{-h}^h f(x) \, dx - A \right|$$

zu finden. Ohne weitere Information über f können wir natürlich keine solche Schranke aufstellen. Nun ist die obige Differenz gleich Null, wenn f eine affine

Funktion ist. Wir hoffen daher, dass die Differenz klein ist, wenn f nur wenig von einer affinen Funktion abweicht. Da affine Funktionen durch $f'' = 0$ gekennzeichnet sind, können wir das Abweichen von einer affinen Funktion durch die Größe $\|f''\|$ messen, vorausgesetzt, dass f'' existiert und beschränkt ist.

4.2 Satz (Sehnen- oder Trapezregel). *Sei $h > 0$ und $f : [-h, h] \rightarrow \mathbb{R}$ zweimal stetig differenzierbar, sei*

$$A := h(f(-h) + f(h)).$$

Dann gilt

$$\int_{-h}^h f(x) dx = A - \frac{2}{3}h^3 f''(c)$$

mit passendem $c \in [-h, h]$, also

$$\left| \int_{-h}^h f(x) dx - A \right| \leq \frac{2}{3}h^3 \|f''\|.$$

Beweis. Durch zweimalige partielle Integration erhalten wir

$$\begin{aligned} \int_{-h}^h f(x) dx &= [xf(x)]_{-h}^h - \int_{-h}^h xf'(x) dx \\ &= A - \left[\frac{1}{2}(x^2 - h^2)f'(x) \right]_{-h}^h + \int_{-h}^h \frac{1}{2}(x^2 - h^2)f''(x) dx. \end{aligned}$$

Hier ist

$$\left[\frac{1}{2}(x^2 - h^2)f'(x) \right]_{-h}^h = 0$$

(deshalb wurde bei der zweiten partiellen Integration zu x die Stammfunktion $\frac{1}{2}(x^2 - h^2)$ gewählt). Wegen $h^2 - x^2 \geq 0$ für $x \in [-h, h]$ folgt aus dem Mittelwertsatz der Integralrechnung 2.5 mit passendem $c \in [-h, h]$

$$\begin{aligned} \int_{-h}^h (h^2 - x^2)f''(x) dx &= f''(c) \int_{-h}^h (h^2 - x^2) dx \\ &= f''(c) \left[h^2x - \frac{x^3}{3} \right]_{-h}^h = f''(c) \frac{4}{3}h^3, \end{aligned}$$

also

$$\int_{-h}^h f(x) dx = A - \frac{2}{3}h^3 f''(c)$$

und daher die Behauptung. \blacksquare

Jetzt wollen wir, wie oben angedeutet, das Intervall $[a, b]$ in Teilintervalle zerlegen und die Sehnenregel auf die Teilintervalle anwenden.

Sei $n \in \mathbb{N}$. Wir zerlegen $[a, b]$ in n Intervalle gleicher Länge und setzen dazu

$$h := \frac{b-a}{n}, \quad x_j := a + jh \quad \text{für } j = 0, \dots, n.$$

Anwendung der Sehnenregel auf das Intervall $[x_{j-1}, x_j]$ ($\frac{h}{2}$ spielt also jetzt die Rolle des früheren h) ergibt

$$\int_{x_{j-1}}^{x_j} f(x) dx = \frac{h}{2} [f(x_{j-1}) + f(x_j)] - \frac{h^3}{12} f''(c_j)$$

mit $c_j \in [x_{j-1}, x_j]$ ($j = 1, \dots, n$). Summieren wir über j und setzen

$$A_n := h \left[\frac{1}{2} f(a) + f(a+h) + \dots + f(a+(n-1)h) + \frac{1}{2} f(b) \right],$$

so erhalten wir

$$\int_a^b f(x) dx = A_n - \frac{h^3}{12} \sum_{j=1}^n f''(c_j),$$

also

$$\left| \int_a^b f(x) dx - A_n \right| \leq \frac{h^3}{12} n \|f''\| = \frac{h^2}{12} (b-a) \|f''\|.$$

Wir sehen also, dass der Fehler, den wir bei der Approximation des Integrals $\int_a^b f(x) dx$ durch die endliche Näherungssumme A_n machen, wie das Quadrat der Schrittweite h gegen Null geht.

7.5 Parameterabhängige Integrale

In der Taylorschen Formel 3.3 hatten wir das Restglied in der Form

$$R_{n+1}(x) = \frac{1}{n!} \int_a^x f^{(n+1)}(t)(x-t)^n dt$$

angegeben. Dies ist ein Beispiel für ein „parameterabhängiges Integral“. Im Integranden und in der oberen Grenze kommt ein „Parameter“ vor, d.h. eine Zahl, die aus einer gegebenen Teilmenge von \mathbb{R} gewählt werden kann, so dass also durch das Integral eine auf dieser Teilmenge definierte Funktion erklärt wird. Wir wollen Eigenschaften einer so erklärten Funktion betrachten; dabei beschränken wir uns auf den Fall, dass der Parameter im Integranden vorkommt. Zur exakten Formulierung benötigen wir den Begriff einer „Funktion von zwei Veränderlichen“.

Wir bezeichnen das kartesische Produkt $\mathbb{R} \times \mathbb{R}$, also die Menge aller geordneten Paare reeller Zahlen, mit \mathbb{R}^2 .

Definition. Sei $A \subset \mathbb{R}^2$. Eine Funktion $f : A \rightarrow \mathbb{R}$ heißt „Funktion von zwei (reellen) Veränderlichen“. Statt $f((x, y))$ schreibt man $f(x, y)$.

Die Funktion $f : A \rightarrow \mathbb{R}$ heißt *gleichmäßig stetig* \Leftrightarrow

$$\begin{aligned} \forall \varepsilon \in \mathbb{R}^+ \quad \exists \delta \in \mathbb{R}^+ \quad \forall (x, t), (x', t') \in A : \\ |x - x'| < \delta \wedge |t - t'| < \delta \Rightarrow |f(x, t) - f(x', t')| < \varepsilon. \end{aligned}$$

Ist $D, [a, b] \subset \mathbb{R}$ und $f : D \times [a, b] \rightarrow \mathbb{R}$ gleichmäßig stetig, so ist insbesondere für jedes $x \in D$ die Funktion

$$f(x, \cdot) : [a, b] \rightarrow \mathbb{R} : t \mapsto f(x, t)$$

stetig und daher eine Regelfunktion. Wir können also

$$F(x) := \int_a^b f(x, t) dt \quad \text{für } x \in D$$

definieren. Wir fragen jetzt nach Eigenschaften dieser Funktion F wie Stetigkeit und Differenzierbarkeit. Es sei darauf hingewiesen, dass die im folgenden getroffenen Voraussetzungen wesentlich stärker sind als notwendig wäre. (Für allgemeinere Aussagen siehe z.B. Barner-Flohr.) Wir begnügen uns jedoch mit diesen einfach zu beweisenden Resultaten, die für viele Anwendungen ausreichen.

5.1 Satz. Sei $D, [a, b] \subset \mathbb{R}$, sei $f : D \times [a, b] \rightarrow \mathbb{R}$ eine gleichmäßig stetige Funktion. Dann ist die durch

$$F(x) := \int_a^b f(x, t) dt \quad \text{für } x \in D$$

definierte Funktion F stetig.

Beweis. Sei $x \in D$ und $\varepsilon \in \mathbb{R}^+$ gegeben. Wegen der gleichmäßigen Stetigkeit von f gibt es ein $\delta \in \mathbb{R}^+$ mit

$$|f(x, t) - f(x', t)| < \frac{\varepsilon}{b - a}$$

für alle $t \in [a, b]$ und alle $x, x' \in D$ mit $|x - x'| < \delta$. Für $x, x' \in D$ mit $|x - x'| < \delta$ folgt also

$$|F(x) - F(x')| \leq \int_a^b |f(x, t) - f(x', t)| dt \leq \varepsilon. \quad \blacksquare$$

Jetzt fragen wir nach der Differenzierbarkeit von F . Sie erfordert natürlich eine geeignete Differenzierbarkeitsvoraussetzung an f .

Definition. Sei $f : D \times [a, b] \rightarrow \mathbb{R}$ eine Funktion. Sie heißt *partiell differenzierbar nach der ersten Veränderlichen*, wenn für alle $t \in [a, b]$ die Funktion

$$f(\cdot, t) : D \rightarrow \mathbb{R} : x \mapsto f(x, t)$$

differenzierbar ist. Ihre Ableitung an der Stelle x wird dann mit $\partial_1 f(x, t)$ bezeichnet. Die Funktion

$$\partial_1 f : D \times [a, b] \rightarrow \mathbb{R} : (x, t) \mapsto \partial_1 f(x, t)$$

heißt *partielle Ableitung von f nach der ersten Veränderlichen*.

5.2 Satz. Sei $[c, d], [a, b] \subset \mathbb{R}$, sei $f : [c, d] \times [a, b] \rightarrow \mathbb{R}$ eine gleichmäßig stetige Funktion, deren partielle Ableitung $\partial_1 f$ existiert und gleichmäßig stetig ist. Dann ist die durch

$$F(x) := \int_a^b f(x, t) dt \quad \text{für } x \in [c, d]$$

definierte Funktion F differenzierbar, und es gilt

$$F'(x) = \int_a^b \partial_1 f(x, t) dt$$

(„Vertauschbarkeit von Differentiation und Integration“).

Beweis. Sei $x \in [c, d]$ und $\varepsilon \in \mathbb{R}^+$ gegeben. Für $h \neq 0$ mit $x + h \in [c, d]$ gilt

$$\frac{F(x+h) - F(x)}{h} - \int_a^b \partial_1 f(x, t) dt = \int_a^b \left(\frac{f(x+h, t) - f(x, t)}{h} - \partial_1 f(x, t) \right) dt.$$

Für jedes $t \in [a, b]$ wenden wir auf die Funktion $f(\cdot, t)$ den Mittelwertsatz der Differentialrechnung 2.2 aus Kapitel 6 an. Danach existiert ein $\vartheta_t \in (0, 1)$ mit

$$\frac{f(x+h, t) - f(x, t)}{h} = \partial_1 f(x + \vartheta_t h, t).$$

Nach Voraussetzung ist $\partial_1 f$ gleichmäßig stetig, also existiert ein $\delta \in \mathbb{R}^+$ mit

$$|\partial_1 f(x, t) - \partial_1 f(x', t)| < \frac{\varepsilon}{b-a}$$

für alle $t \in [a, b]$ und alle $x, x' \in [c, d]$ mit $|x - x'| < \delta$. Sei $|h| < \delta$. Dann gilt $|x - (x + \vartheta_t h)| = \vartheta_t |h| < \delta$, also

$$\left| \frac{f(x+h, t) - f(x, t)}{h} - \partial_1 f(x, t) \right| = |\partial_1 f(x + \vartheta_t h, t) - \partial_1 f(x, t)| < \frac{\varepsilon}{b-a}$$

und daher

$$\left| \frac{F(x+h) - F(x)}{h} - \int_a^b \partial_1 f(x, t) dt \right| < \varepsilon. \quad \blacksquare$$

Nun fragen wir analog, ob man „unter dem Integralzeichen integrieren“ darf.

5.3 Satz. Sei $f : [c, d] \times [a, b] \rightarrow \mathbb{R}$ gleichmäßig stetig. Dann gilt

$$\int_c^d \left(\int_a^b f(x, t) dt \right) dx = \int_a^b \left(\int_c^d f(x, t) dx \right) dt$$

(„Vertauschbarkeit der Integrationsreihenfolge“).

Beweis. Setze

$$G(y) := \int_c^y \left(\int_a^b f(x, t) dt \right) dx \quad \text{für } y \in [c, d].$$

Nach Satz 5.1 ist der Integrand des äußeren Integrals stetig; nach dem Hauptsatz der Differential- und Integralrechnung 3.1 ist daher G differenzierbar und

$$G'(y) = \int_a^b f(y, t) dt.$$

Setze

$$H(y) := \int_a^b \left(\int_c^y f(x, t) dx \right) dt \quad \text{für } y \in [c, d].$$

Die durch

$$h(y, t) := \int_c^y f(x, t) dx \quad \text{für } (y, t) \in [c, d] \times [a, b]$$

definierte Funktion h ist (wieder nach Satz 3.1) partiell differenzierbar nach der ersten Veränderlichen, und es gilt

$$\partial_1 h = f.$$

Also ist $\partial_1 h$ gleichmäßig stetig. Nach Satz 5.2 folgt die Differenzierbarkeit der Funktion H und die Gleichung

$$H'(y) = \int_a^b \partial_1 h(y, t) dt = \int_a^b f(y, t) dt.$$

Es ist also $(G - H)' = 0$ und daher $G - H = \text{const.}$ Wegen $G(c) = 0 = H(c)$ ist $G = H$. Daraus folgt die Behauptung. ■

Durch ein Beispiel wollen wir noch zeigen, dass man die Integrationsreihenfolge keineswegs immer vertauschen darf.

Beispiel. Die Funktion $f : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ sei definiert durch

$$f(x, y) := \begin{cases} \frac{1}{y^2} & \text{für } 0 < x < y < 1, \\ -\frac{1}{x^2} & \text{für } 0 < y < x < 1, \\ 0 & \text{sonst.} \end{cases}$$

Ist $y = 0$ oder $y = 1$, so ist $f(x, y) = 0$ für alle $x \in [0, 1]$. Sei $y \in (0, 1)$. Dann ist

$$f(x, y) = \begin{cases} 0 & \text{für } x = 0, \\ \frac{1}{y^2} & \text{für } 0 < x < y, \\ 0 & \text{für } x = y, \\ -\frac{1}{x^2} & \text{für } y < x < 1, \\ 0 & \text{für } x = 1. \end{cases}$$

Also ist $f(\cdot, y)$ auf dem Intervall $(0, y)$ konstant und auf dem Intervall $(y, 1)$ Einschränkung einer auf $[y, 1]$ stetigen Funktion. Es folgt, dass $f(\cdot, y)$ Regelfunktion ist. Daher ist das Integral

$$F(y) := \int_0^1 f(x, y) dx$$

definiert. Es ist $F(0) = F(1) = 0$. Sei $y \in (0, 1)$. Dann ist

$$F(y) = \int_0^y \frac{1}{y^2} dx + \int_y^1 \left(-\frac{1}{x^2}\right) dx = \left[\frac{1}{y^2}x\right]_{x=0}^{x=y} + \left[\frac{1}{x}\right]_{x=y}^{x=1} = 1.$$

Also ist F Regelfunktion und

$$\int_0^1 F(y) dy = 1.$$

Analog findet man, dass für jedes $x \in [0, 1]$ die Funktion $f(x, \cdot)$ Regelfunktion ist und dass

$$\begin{aligned} G(x) &:= \int_0^1 f(x, y) dy \\ &= \int_0^x \left(-\frac{1}{x^2}\right) dy + \int_x^1 \frac{1}{y^2} dy = -1 \end{aligned}$$

für $x \in (0, 1)$ sowie $G(0) = G(1) = 0$ gilt. Also ist G Regelfunktion und $\int_0^1 G(x) dx = -1$. Damit ist

$$\int_0^1 \int_0^1 f(x, y) dx dy = 1 \neq -1 = \int_0^1 \int_0^1 f(x, y) dy dx$$

gezeigt.

7.6 Uneigentliche Integrale

Bisher können wir nur Regelfunktionen integrieren. Eine solche Funktion ist stets auf einem kompakten Intervall definiert und beschränkt. Es kommt jedoch häufig vor, dass Funktionen über unendliche Intervalle oder unbeschränkte Funktionen zu integrieren sind. In naheliegender Weise können solche Integrale durch Grenzübergänge definiert werden. Wir betrachten zunächst den Fall eines unbeschränkten Integrationsintervalls.

Definition. Sei $a \in \mathbb{R}$, sei $f : [a, \infty) \rightarrow \mathbb{R}$ eine Funktion derart, dass für jedes $b \in [a, \infty)$ die Einschränkung $f|_{[a,b]}$ eine Regelfunktion ist. Falls der Grenzwert

$$\lim_{b \rightarrow \infty} \int_a^b f(x) dx$$

existiert, bezeichnet man ihn mit $\int_a^\infty f(x) dx$ und sagt, dass das (uneigentliche) Integral $\int_a^\infty f(x) dx$ *konvergiert*.

Ganz analog definiert man $\int_{-\infty}^a f(x) dx$ und

$$\int_{-\infty}^{\infty} f(x) dx := \int_{-\infty}^a f(x) dx + \int_a^{\infty} f(x) dx,$$

wobei $a \in \mathbb{R}$ beliebig und die rechte Seite offenbar unabhängig von a ist.

Im folgenden wird von allen auftretenden Funktionen stillschweigend vorausgesetzt, dass ihre Einschränkungen auf kompakte Intervalle Regelfunktionen sind.

Für die Konvergenz uneigentlicher Integrale gilt das folgende Cauchy-Kriterium.

6.1 Satz (Cauchy-Kriterium). *Das Integral $\int_a^\infty f(x) dx$ konvergiert*

$$\Leftrightarrow \forall \varepsilon \in \mathbb{R}^+ \quad \exists u \in \mathbb{R} \quad \forall x_1, x_2 \geq u : \left| \int_{x_1}^{x_2} f(x) dx \right| < \varepsilon.$$

Der Satz ist analog zum Cauchy-Kriterium 2.3 aus Kapitel 3 für Reihen und zum Cauchy-Kriterium 2.3 aus Kapitel 4 für Grenzwerte von Funktionen; man beweist ihn analog wie letzteres.

Beispiele. (1) $\int_1^\infty \frac{1}{x^s} dx$. Für $b \geq 1$ und $s \neq 1$ ist

$$\int_1^b \frac{1}{x^s} dx = \left[\frac{1}{1-s} \frac{1}{x^{s-1}} \right]_1^b = \frac{1}{s-1} \left(1 - \frac{1}{b^{s-1}} \right).$$

Also ist das Integral $\int_1^\infty \frac{1}{x^s} dx$ konvergent für $s > 1$, nicht konvergent für $s < 1$. Für $s = 1$ ist es wegen $\int_1^b \frac{1}{x} dx = \ln b$ ebenfalls nicht konvergent.

(2) $\int_{-\infty}^\infty \frac{1}{1+x^2} dx$. Für $b \geq 0$ gilt

$$\int_{-b}^0 \frac{1}{1+x^2} dx = [\arctan x]_{-b}^0 = -\arctan(-b) \rightarrow \frac{\pi}{2} \quad \text{für } b \rightarrow \infty,$$

analog

$$\int_0^b \frac{1}{1+x^2} dx \rightarrow \frac{\pi}{2}, \quad \text{also } \int_{-\infty}^\infty \frac{1}{1+x^2} dx = \pi.$$

(3) $\int_0^\infty \frac{\sin x}{x} dx$ ist konvergent (an der Stelle 0 ist der Integrand gleich 1 zu setzen; er ist dann stetig). In der Tat, für $b > a > 0$ ergibt partielle Integration

$$\int_a^b \frac{\sin x}{x} dx = \left[-\frac{1}{x} \cos x \right]_a^b - \int_a^b \frac{\cos x}{x^2} dx.$$

Sei $\varepsilon \in \mathbb{R}^+$. Für $x_2 > x_1 > \frac{1}{\varepsilon}$ gilt

$$\left| \int_{x_1}^{x_2} \frac{\cos x}{x^2} dx \right| \leq \int_{x_1}^{x_2} \frac{1}{x^2} dx = \left[-\frac{1}{x} \right]_{x_1}^{x_2} = \frac{1}{x_1} - \frac{1}{x_2} < \frac{1}{x_1} < \varepsilon.$$

Nach dem Cauchy-Kriterium 6.1 ist also $\int_a^\infty \frac{\cos x}{x^2} dx$ konvergent und daher auch $\int_a^\infty \frac{\sin x}{x} dx$ konvergent.

Das Integralkriterium für Reihen

An dieser Stelle können wir ein Konvergenzkriterium für Reihen nachtragen, dessen Formulierung den Begriff des uneigentlichen Integrals erfordert.

6.2 Satz (Integralkriterium für Reihen). Sei $f : [1, \infty) \rightarrow \mathbb{R}^+$ eine monoton fallende Funktion. Dann ist die Reihe

$$\sum_{n=1}^{\infty} f(n)$$

genau dann konvergent, wenn das uneigentliche Integral

$$\int_1^{\infty} f(x) dx$$

konvergent ist.

Beweis. Für $k \in \mathbb{N}$ und $k-1 \leq x \leq k$ gilt $f(k) \leq f(x) \leq f(k-1)$ und daher

$$f(k) \leq \int_{k-1}^k f(x) dx \leq f(k-1).$$

Summation ergibt

$$\sum_{k=2}^n f(k) \leq \int_1^n f(x) dx \leq \sum_{k=1}^{n-1} f(k).$$

Wegen $f \geq 0$ liest man hieran die Behauptung ab. \blacksquare

Beispiele. (1) $\sum_{n=1}^{\infty} \frac{1}{n^s}$ konvergiert für $s > 1$ und divergiert für $s \leq 1$ (vgl. Beispiel (1) nach Satz 6.1)

(2) $\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^s}$ ist für $s > 1$ konvergent, für $s \leq 1$ divergent. In der Tat, erhalten wir für $b \geq 2$ und $s \neq 1$ [Substitution $x = e^t$]

$$\int_2^b \frac{1}{x(\ln x)^s} dx = \int_{\ln 2}^{\ln b} \frac{1}{t^s} dt = \left[\frac{1}{1-s} \frac{1}{t^{s-1}} \right]_{\ln 2}^{\ln b} = \frac{1}{s-1} \left[\frac{1}{(\ln 2)^{s-1}} - \frac{1}{(\ln b)^{s-1}} \right].$$

Für $s > 1$ ist also $\int_2^{\infty} \frac{1}{x(\ln x)^s} dx$ konvergent, für $s < 1$ divergent. Für $s = 1$ gilt

$$\int_2^b \frac{1}{x \ln x} dx = \int_{\ln 2}^{\ln b} \frac{1}{t} dt = [\ln t]_{\ln 2}^{\ln b} = \ln \ln b - \ln \ln 2,$$

also ist $\int_2^{\infty} \frac{1}{x \ln x} dx$ ebenfalls divergent.

Wir kommen zum zweiten Typ uneigentlicher Integrale, bei denen der Integrand an einer Integrationsgrenze nicht erklärt ist und auch nicht so erklärt werden kann, dass eine Regelfunktion entsteht.

Definition. Sei $f : (a, b] \rightarrow \mathbb{R}$ eine Funktion derart, dass für jedes $\varepsilon \in \mathbb{R}^+$ (mit $\varepsilon < b - a$) die Einschränkung $f|_{[a+\varepsilon, b]}$ eine Regelfunktion ist. Man sagt, „das Integral $\int_a^b f(x) dx$ konvergiert“, wenn

$$\lim_{\varepsilon \searrow 0} \int_{a+\varepsilon}^b f(x) dx$$

existiert und bezeichnet dann diesen Grenzwert mit $\int_a^b f(x) dx$.

Analog wird $\int_a^b f(x) dx$ behandelt, wenn f nur auf $[a, b)$ erklärt ist. Ist f nur auf (a, b) erklärt, so bezeichnet man $\int_a^b f(x) dx$ als konvergent, wenn für ein (und dann für jedes a) $c \in (a, b)$ die Integrale

$$\int_a^c f(x) dx \quad \text{und} \quad \int_c^b f(x) dx$$

konvergieren; in diesem Fall setzt man

$$\int_a^b f(x) dx := \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Beispiele. (1) Wegen

$$\int_{\varepsilon}^1 \frac{1}{x^s} dx = \begin{cases} \frac{1}{1-s}(1 - \varepsilon^{1-s}) & \text{für } s \neq 1, \\ -\ln \varepsilon & \text{für } s = 1 \end{cases}$$

ist $\int_0^1 \frac{1}{x^s} dx$ für $s < 1$ konvergent, für $s \geq 1$ jedoch nicht.

(2) Wegen $\int_{\varepsilon}^1 \ln x dx = [x \ln x - x]_{\varepsilon}^1 = -1 - \varepsilon \ln \varepsilon + \varepsilon$ und $\lim_{\varepsilon \searrow 0} \varepsilon \ln \varepsilon = 0$ (wie aus Behauptung 1.10 aus Kapitel 5 folgt) ist $\int_0^1 \ln x dx = -1$.

(3) $\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} dx$ konvergiert: Es ist

$$\int_0^{1-\varepsilon} \frac{1}{\sqrt{1-x^2}} dx = [\arcsin x]_0^{1-\varepsilon} = \arcsin(1-\varepsilon)$$

$$\Rightarrow \lim_{\varepsilon \rightarrow 0} \int_0^{1-\varepsilon} \frac{1}{\sqrt{1-x^2}} dx = \frac{\pi}{2}.$$

Analog $\int_{-1}^0 \frac{1}{\sqrt{1-x^2}} dx = \frac{\pi}{2}$.

- (4) Im Beispiel $\int_0^1 \sin \frac{1}{x} dx$ ist der Integrand beschränkt, aber $\lim_{x \searrow 0} \sin \frac{1}{x}$ existiert nicht; daher ist der Integrand nicht Einschränkung einer auf $[0, 1]$ erklärten Regelfunktion.

Die Substitution $x = \frac{1}{t}$ ergibt

$$\int_{\varepsilon}^1 \sin \frac{1}{x} dx = \int_1^{\frac{1}{\varepsilon}} \frac{\sin t}{t^2} dt.$$

Da das Integral $\int_1^{\infty} \frac{\sin t}{t^2} dt$ konvergiert (z.B. nach dem Cauchy-Kriterium), ist $\int_0^1 \sin \frac{1}{x} dx$ konvergent.

8 Funktionenreihen

In diesem Kapitel wollen wir etwas ausführlicher die Möglichkeit studieren, Funktionen durch unendliche Reihen darzustellen. Wir hatten schon im Anschluß an die Taylorformel kurz die Möglichkeit erörtert, bei beliebig oft differenzierbaren Funktionen von den Taylorpolynomen zur Taylorreihe überzugehen. Schon wesentlich früher hatten wir unendliche Reihen benutzt, um gewisse spezielle Funktionen einzuführen, zum Beispiel die Exponentialfunktion durch

$$\exp x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Hierbei ergibt sich unter anderem die Frage, wie man von Eigenschaften der Reihenglieder auf Eigenschaften der Funktion schließen kann. Wir können zum Beispiel zur Berechnung der Ableitung versuchen, „gliedweise“ zu differenzieren. Formal vorgehend, erhält man

$$\exp' x = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{dx^k}{dx} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \exp x.$$

Das Ergebnis ist richtig; aber führt ein solches Vorgehen in jedem Fall zum richtigen Ergebnis? Diese und ähnliche Fragen werden im folgenden beantwortet. Wir betrachten zunächst allgemein konvergente Folgen von Funktionen und erst dann spezielle Reihen.

8.1 Konvergenz von Funktionenfolgen

Wir kennen bereits zwei verschiedene Konvergenzbegriffe für Funktionenfolgen, und an diese sei zunächst erinnert.

Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge von reellen Funktionen, die sämtlich denselben Definitionsbereich D haben. Für jedes $x \in D$ ist dann $(f_n(x))_{n \in \mathbb{N}}$ eine Folge

reeller Zahlen, und für diese ist ein Konvergenzbegriff wohldefiniert. Wenn für jedes $x \in D$ die Folge $(f_n(x))_{n \in \mathbb{N}}$ konvergiert, ist durch

$$f(x) := \lim_{n \rightarrow \infty} f_n(x) \quad (x \in D)$$

eine neue Funktion f auf D erklärt, die wir als *Grenzfunktion* der Folge bezeichnen können.

Definition. Seien f, f_n ($n \in \mathbb{N}$) reelle Funktionen auf D . Die Folge $(f_n)_{n \in \mathbb{N}}$ *konvergiert (punktweise)* gegen f , geschrieben

$$\lim_{n \rightarrow \infty} f_n = f \quad \text{oder} \quad f_n \rightarrow f \quad (n \rightarrow \infty),$$

wenn $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ für alle $x \in D$ gilt.

Beispiele. (1) Sei $D = \mathbb{R}$ und

$$f_n(x) := \sum_{k=0}^n \frac{x^k}{k!}.$$

Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ gegen \exp .

(2) Sei $D = [0, 1]$ und

$$f_n(x) := x^n.$$

Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ gegen die durch

$$f(x) := \begin{cases} 1 & \text{für } x = 1, \\ 0 & \text{für } 0 \leq x < 1. \end{cases}$$

erklärte Funktion f .

Bei diesem Beispiel fällt auf, dass zwar jede Funktion f_n der Folge stetig ist, dass aber die Grenzfunktion unstetig ist. Stetigkeit überträgt sich also bei punktweiser Konvergenz i.a. nicht auf die Grenzfunktion. Um den oft wünschenswerten Schluß von der Stetigkeit der Folgenglieder auf die Stetigkeit der Grenzfunktion zu ermöglichen, braucht man einen Konvergenzbegriff, der schärfer ist als punktweise Konvergenz. Dies ist die bereits in Abschnitt 7.2 benutzte gleichmäßige Konvergenz, die wir jetzt etwas allgemeiner definieren wollen.

Definition. Seien f, f_n ($n \in \mathbb{N}$) reelle Funktionen auf D , sei $D' \subset D$. Die Folge $(f_n)_{n \in \mathbb{N}}$ *konvergiert gleichmäßig in D' gegen f* , wenn gilt

$$\forall \varepsilon \in \mathbb{R}^+ \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 \quad \forall x \in D' : |f_n(x) - f(x)| \leq \varepsilon.$$

Statt „gleichmäßig in D' “ sagt man kurz „gleichmäßig“.

Ist ein fester Definitionsbereich D gegeben, so können wir wie früher für Intervalle die *Supremumsnorm* einer beschränkten Funktion $f : D \rightarrow \mathbb{R}$ erklären durch

$$\|f\| := \sup_{x \in D} |f(x)|.$$

Dann gilt also (für Funktionen f, f_n auf D):

$$\begin{aligned} (f_n)_{n \in \mathbb{N}} &\text{ konvergiert gleichmäßig gegen } f \\ \Leftrightarrow \forall \varepsilon \in \mathbb{R}^+ \exists n_0 \in \mathbb{N} \forall n \geq n_0 : \|f_n - f\| &\leq \varepsilon \\ \Leftrightarrow \lim_{n \rightarrow \infty} \|f_n - f\| &= 0. \end{aligned}$$

Man beachte, dass $\|f_n - f\| < \varepsilon$ impliziert, dass $\|f_n - f\|$ definiert, also $f_n - f_m$ beschränkt ist. Es wird aber nicht vorausgesetzt, dass f, f_n beschränkt sind.

Für die Konvergenz von Folgen reeller Zahlen kennen wir das Kriterium von Cauchy. Ein ganz analoges Kriterium gilt auch für die gleichmäßige Konvergenz von Funktionenfolgen. Im folgenden liege stets ein fester Definitionsbereich D zugrunde.

Definition. Die Folge $(f_n)_{n \in \mathbb{N}}$ von Funktionen auf D heißt *Cauchy-Folge* genau dann, wenn

$$\forall \varepsilon \in \mathbb{R}^+ \exists n_0 \in \mathbb{N} \forall n, m \geq n_0 : \|f_n - f_m\| < \varepsilon.$$

1.1 Satz. Die Folge $(f_n)_{n \in \mathbb{N}}$ konvergiert genau dann gleichmäßig, wenn sie eine Cauchy-Folge ist.

Beweis. „ \Rightarrow “: Sei $(f_n)_{n \in \mathbb{N}}$ gleichmäßig konvergent gegen f . Sei $\varepsilon \in \mathbb{R}^+$. Es gibt ein $n_0 \in \mathbb{N}$ mit $\|f_n - f\| < \varepsilon/2$ für $n \geq n_0$. Für alle $n, m \geq n_0$ gilt also

$$\|f_n - f_m\| \leq \|f_n - f\| + \|f - f_m\| < \varepsilon,$$

wobei die Dreiecksungleichung für die Supremumsnorm benutzt wurde.

„ \Leftarrow “: Sei $(f_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge. Sei $\varepsilon \in \mathbb{R}^+$. Nach Voraussetzung existiert ein $n_0 \in \mathbb{N}$ mit $\|f_n - f_m\| < \varepsilon$ für alle $n, m \geq n_0$. Insbesondere gilt also für jedes $x \in D$

$$|f_n(x) - f_m(x)| < \varepsilon \quad \text{für } n, m \geq n_0.$$

Die Folge $(f_n(x))_{n \in \mathbb{N}}$ ist also eine Cauchy-Folge reeller Zahlen und daher nach dem gewöhnlichen Cauchy-Kriterium konvergent gegen eine Zahl, die wir $f(x)$ nennen. Da $x \in D$ beliebig war, ist damit eine Funktion $f : D \rightarrow \mathbb{R}$ erklärt. Wir behaupten, dass $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f konvergiert.

Sei $\varepsilon \in \mathbb{R}^+$ und dazu n_0 wie oben. Für beliebiges $x \in D$ gilt dann

$$|f_n(x) - f_m(x)| < \varepsilon \quad \text{für } n, m \geq n_0.$$

Der Grenzübergang $m \rightarrow \infty$ liefert

$$|f_n(x) - f(x)| \leq \varepsilon \quad \text{für } n \geq n_0.$$

Da dies für alle $x \in D$ gilt, folgt $\|f_n - f\| \leq \varepsilon$ für alle $n \geq n_0$. ■

Der folgende Satz rückt die Bedeutung der gleichmäßigen Konvergenz ins rechte Licht.

1.2 Satz. *Seien f_n, f Funktionen auf D ($n \in \mathbb{N}$), sei $a \in D$. Sind alle f_n stetig in a und konvergiert $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f , so ist f stetig in a .*

Beweis. Sei $\varepsilon \in \mathbb{R}^+$. Es gibt ein $m \in \mathbb{N}$ mit $\|f_m - f\| < \varepsilon/3$, also

$$|f_m(x) - f(x)| < \frac{\varepsilon}{3} \quad \text{für alle } x \in D.$$

Da f_m in a stetig ist, existiert ein $\delta \in \mathbb{R}^+$ mit

$$|f_m(x) - f_m(a)| < \frac{\varepsilon}{3} \quad \text{für alle } x \in D \text{ mit } |x - a| < \delta.$$

Sei jetzt $x \in D$ und $|x - a| < \delta$. Dann gilt

$$\begin{aligned} |f(x) - f(a)| &\leq |f(x) - f_m(x)| + |f_m(x) - f_m(a)| + |f_m(a) - f(a)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned} \quad \blacksquare$$

Bemerkung. Natürlich kann man nicht umgekehrt schließen, d.h. wenn $(f_n)_{n \in \mathbb{N}}$ punktweise gegen f konvergiert und alle f_n sowie f stetig sind, braucht keineswegs die Konvergenz gleichmäßig zu sein.

Kann man in Satz 1.2 „stetig“ durch „differenzierbar“ ersetzen? Das ist nicht der Fall.

Beispiel. Sei $D = \mathbb{R}$, $f_n(x) := \sqrt{x^2 + \frac{1}{n}}$ ($n \in \mathbb{N}$) und $f(x) := |x|$ für $x \in \mathbb{R}$. Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f , wie aus

$$|f_n(x) - f(x)| = \left| \sqrt{x^2 + \frac{1}{n}} - \sqrt{x^2} \right| = \frac{x^2 + \frac{1}{n} - x^2}{\sqrt{x^2 + \frac{1}{n}} + \sqrt{x^2}} \leq \frac{1}{\sqrt{n}}$$

folgt. Jede Funktion f_n ist differenzierbar, aber f ist in 0 nicht differenzierbar.

Es kann auch sein, dass zwar die Grenzfunktion f der Folge $(f_n)_{n \in \mathbb{N}}$ differenzierbar ist, aber die Folge $(f'_n)_{n \in \mathbb{N}}$ nicht gegen f' konvergiert.

Beispiel. Sei $f_n(x) = \frac{1}{n} \sin nx$ und $f(x) = 0$ für $x \in \mathbb{R}$. Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f , alle f_n sowie f sind differenzierbar, aber $(f'_n)_{n \in \mathbb{N}}$ konvergiert nicht punktweise gegen f' . Es ist nämlich $f'_n(x) = \cos nx$ und z.B. $\lim_{n \rightarrow \infty} f'_n(0) = 1$, $f'(0) = 0$.

Um wirklich Grenzübergang und Differentiation vertauschen zu können, braucht man stärkere Voraussetzungen, z.B. die folgenden:

1.3 Satz. Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge differenzierbarer Funktionen auf $[a, b]$. Die Folge $(f'_n)_{n \in \mathbb{N}}$ konvergiere gleichmäßig, und $(f_n)_{n \in \mathbb{N}}$ konvergiere an wenigstens einer Stelle $x_0 \in [a, b]$. Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen eine differenzierbare Funktion f und $(f'_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f' .

Beweis. Wir zeigen zuerst, dass $(f_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge ist. Sei $\varepsilon \in \mathbb{R}^+$ gegeben. Nach Voraussetzung und Satz 1.1 existiert ein $n_1 \in \mathbb{N}$ mit

$$\|f'_m - f'_n\| < \frac{\varepsilon}{2(b-a)} \quad \text{für alle } m, n \geq n_1.$$

Da $(f_n(x_0))_{n \in \mathbb{N}}$ eine Cauchy-Folge ist, existiert ein $n_2 \in \mathbb{N}$ mit

$$|f_m(x_0) - f_n(x_0)| < \frac{\varepsilon}{2} \quad \text{für } m, n \geq n_2.$$

Sei $x \in [a, b]$. Es ist

$$|f_m(x) - f_n(x)| \leq |(f_m - f_n)(x) - (f_m - f_n)(x_0)| + |f_m(x_0) - f_n(x_0)|.$$

Nach dem Mittelwertsatz 2.2 aus Kapitel 6 existiert ein $z \in [a, b]$ mit

$$(f_m - f_n)(x) - (f_m - f_n)(x_0) = (f'_m(z) - f'_n(z))(x - x_0).$$

Für alle $m, n \geq n_0 := \max\{n_1, n_2\}$ folgt

$$\begin{aligned} |f_m(x) - f_n(x)| &\leq |f'_m(z) - f'_n(z)| |x - x_0| + \frac{\varepsilon}{2} \\ &\leq \|f'_m - f'_n\| (b-a) + \frac{\varepsilon}{2} < \varepsilon. \end{aligned}$$

Da $x \in [a, b]$ beliebig war, folgt $\|f_m - f_n\| \leq \varepsilon$ für $m, n \geq n_0$. Also ist $(f_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge und daher nach Satz 1.1 gleichmäßig konvergent gegen eine Funktion f .

Wir zeigen jetzt die Differenzierbarkeit von f . Sei $c \in [a, b]$. Setze

$$g_n(x) := \begin{cases} \frac{f_n(x) - f_n(c)}{x - c} - f'_n(c) & \text{für } x \in [a, b] \setminus \{c\}, \\ 0 & \text{für } x = c. \end{cases}$$

Wegen der Differenzierbarkeit von f_n in c ist g_n in c stetig. Wir zeigen zunächst, dass $(g_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge ist. Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Nach Voraussetzung und Satz 1.1 existiert ein $n_0 \in \mathbb{N}$ mit

$$\|f'_m - f'_n\| < \frac{\varepsilon}{2} \quad \text{für alle } m, n \geq n_0.$$

Sei $x \in [a, b]$. Im Fall $x \neq c$ ist

$$g_m(x) - g_n(x) = \frac{(f_m - f_n)(x) - (f_m - f_n)(c)}{x - c} - (f_m - f_n)'(c).$$

Nach dem Mittelwertsatz existiert ein $z \in [a, b]$ mit

$$\frac{(f_m - f_n)(x) - (f_m - f_n)(c)}{x - c} = (f_m - f_n)'(z).$$

Für $m, n \geq n_0$ folgt

$$\begin{aligned} |g_m(x) - g_n(x)| &\leq |f'_m(z) - f'_n(z)| + |f'_m(c) - f'_n(c)| \\ &\leq 2\|f'_m - f'_n\| \leq \varepsilon. \end{aligned}$$

Dies gilt trivialerweise auch für $x = c$. Da $x \in [a, b]$ beliebig war, ist also $\|g_m - g_n\| \leq \varepsilon$ für $m, n \geq n_0$. Also ist $(g_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge.

Nach Satz 1.1 ist $(g_n)_{n \in \mathbb{N}}$ gleichmäßig konvergent gegen eine Funktion g , die $g(c) = 0$ erfüllt und nach Satz 1.2 in c stetig ist. Nun gilt für alle $x \in [a, b] \setminus \{c\}$

$$\frac{f_n(x) - f_n(c)}{x - c} - f'_n(c) = g_n(x),$$

woraus durch Grenzübergang

$$\frac{f(x) - f(c)}{x - c} - \lim_{n \rightarrow \infty} f'_n(c) = g(x)$$

folgt. Wegen $\lim_{x \rightarrow c} g(x) = g(c) = 0$ folgt

$$\lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c} = \lim_{n \rightarrow \infty} f'_n(c).$$

Also ist f in c differenzierbar und $f'(c) = \lim_{n \rightarrow \infty} f'_n(c)$. Damit ist die Differenzierbarkeit von f gezeigt, ferner die punktweise Konvergenz von $(f'_n)_{n \in \mathbb{N}}$ gegen f' . Diese Konvergenz ist nach Voraussetzung gleichmäßig. ■

8.2 Potenzreihen

Nachdem wir früher Reihen reeller Zahlen und jetzt Funktionenfolgen betrachtet haben, ist klar, was allgemein unter einer Funktionenreihe zu verstehen ist. Sei $(f_k)_{k \in \mathbb{N}_0}$ eine Funktionenfolge auf D . Dann verstehen wir unter

$$\sum_{k=0}^{\infty} f_k$$

die Funktionenfolge der Partialsummen $(\sum_{k=0}^n f_k)_{n \in \mathbb{N}}$, und wir bezeichnen $\sum_{k=0}^{\infty} f_k$ als Funktionenreihe. Ist die Folge punktweise konvergent, so bezeichnen wir mit $\sum_{k=0}^{\infty} f_k$ auch die Grenzfunktion. Die Schreibweise

$$\sum_{k=0}^{\infty} f_k = f \quad \text{in } D$$

bedeutet also definitionsgemäß:

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n f_k(x) = f(x) \quad \text{für alle } x \in D.$$

Konvergiert die Folge $(\sum_{k=0}^n f_k)_{n \in \mathbb{N}}$ in $D' \subset D$ gleichmäßig (gegen f), so sagen wir, dass die Reihe $\sum_{k=0}^{\infty} f_k$ in D' gleichmäßig (gegen f) konvergiert.

Die Begriffe der punktweisen oder gleichmäßigen Konvergenz von Reihen sind also nichts Neues gegenüber den Funktionenfolgen. Auch die Sätze aus Abschnitt 8.1 gelten natürlich sinngemäß für Funktionenreihen. Neu gegenüber der Folgenkonvergenz ist lediglich - wie schon bei Zahlenreihen - der Begriff der absoluten Konvergenz. Definitionsgemäß konvergiert die Reihe $\sum f_k$ absolut, wenn die Reihe $\sum |f_k|$ konvergiert.

In Verallgemeinerung von Satz 2.8 aus Kapitel 3 haben wir das folgende wichtige Kriterium für gleichmäßige und absolute Konvergenz.

2.1 Satz (Majorantenkriterium). *Sei $f_k : D \rightarrow \mathbb{R}$, $c_k \in \mathbb{R}^+$ ($k \in \mathbb{N}_0$) gegeben. Gilt*

$$\|f_k\| \leq c_k \quad \text{für alle } k \in \mathbb{N}_0$$

und ist die Reihe $\sum_{k=0}^{\infty} c_k$ konvergent, so konvergiert die Reihe $\sum_{k=0}^{\infty} f_k$ absolut und gleichmäßig.

Beweis. Sei $\varepsilon \in \mathbb{R}^+$. Da $\sum c_k$ konvergiert, existiert ein $n_0 \in \mathbb{N}$ mit

$$\sum_{k=m}^{m+p} c_k < \varepsilon$$

für alle $m \geq n_0, p \in \mathbb{N}_0$. Für diese m, p gilt also

$$\left\| \sum_{k=m}^{m+p} f_k \right\| \leq \left\| \sum_{k=m}^{m+p} |f_k| \right\| \leq \sum_{k=m}^{m+p} \|f_k\| \leq \sum_{k=m}^{m+p} c_k < \varepsilon.$$

Die Folge der Partialsummen von $\sum f_k$ und die Folge der Partialsummen von $\sum |f_k|$ sind also Cauchyfolgen und daher nach Satz 1.1 gleichmäßig konvergent. ■

Das Vorstehende und die Ergebnisse aus Abschnitt 8.1 wollen wir jetzt anwenden auf den besonders wichtigen Spezialfall der Potenzreihen.

Unter einer *Potenzreihe zur Stelle a* versteht man eine Funktionenreihe

$$\sum_{k=0}^{\infty} f_k \quad \text{mit } f_k(x) = a_k(x-a)^k \quad \text{für } x \in \mathbb{R},$$

wo $a \in \mathbb{R}$ und $(a_k)_{k \in \mathbb{N}_0}$ eine Folge reeller Zahlen ist. Hier ist die folgende ungenaue, aber bequeme Sprechweise üblich. Man sagt

$$\text{„die Potenzreihe } \sum_{k=0}^{\infty} a_k(x-a)^k \text{“}$$

statt

$$\text{„die Potenzreihe } \sum_{k=0}^{\infty} f_k \text{ mit } f_k(x) = a_k(x-a)^k \text{ für } x \in \mathbb{R} \text{“}.$$

Wir fragen jetzt nach der Menge aller $x \in \mathbb{R}$, für die eine gegebene Potenzreihe

$$\sum_{k=0}^{\infty} a_k(x-a)^k \tag{2.2}$$

konvergiert. Auf jeden Fall konvergiert sie trivialerweise für $x = a$ (es gibt Potenzreihen, die für kein anderes x konvergieren). Allgemein gibt das Wurzelkriterium 2.10 aus Kapitel 3 Auskunft. Nach ihm ist (für gegebenes $x \in \mathbb{R}$) die Reihe (2.2) absolut konvergent, wenn

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n(x-a)^n|} = |x-a| \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} < 1$$

ist, und sie ist divergent, wenn $|x-a| \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1$ ist.

Zur Vermeidung lästiger Fallunterscheidungen schreiben wir $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ und definieren

$$-\infty < y < \infty \quad \text{für } y \in \mathbb{R}.$$

Außerdem definieren wir $\frac{1}{\infty} = 0$ und $\frac{1}{0} = \infty$, ferner $y + \infty = \infty$, $y - \infty = -\infty$ für $y \in \mathbb{R}$. Es sei daran erinnert, dass wir in Abschnitt 3.2 $\limsup \sqrt[n]{|a_n|} = \infty$ definiert hatten im Fall, dass die Folge $(\sqrt[n]{|a_n|})_{n \in \mathbb{N}}$ nicht beschränkt ist.

Für unsere gegebene Potenzreihe setzen wir jetzt

$$r := \frac{1}{\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}}.$$

Dann gilt also: Für alle $x \in \mathbb{R}$ mit $|x - a| < r$ ist (2.2) absolut konvergent, für alle $x \in \mathbb{R}$ mit $|x - a| > r$ ist (2.2) divergent. Das Intervall $(a - r, a + r)$ heißt daher *Konvergenzintervall* der Potenzreihe.

Wie steht es mit gleichmäßiger Konvergenz? Sei (2.2) absolut konvergent für ein x_0 . Für alle $x \in \mathbb{R}$ mit $|x - a| \leq |x_0 - a|$ gilt dann

$$|a_k(x - a)^k| \leq |a_k||x_0 - a|^k \quad \text{für } k \in \mathbb{N}_0,$$

und die Reihe $\sum_{k=0}^{\infty} |a_k||x_0 - a|^k$ ist konvergent. Nach dem Majorantenkriterium 2.1 folgt, dass die Reihe (2.2) in $[a - |x_0 - a|, a + |x_0 - a|]$ absolut und gleichmäßig konvergiert. Wir fassen zusammen:

2.3 Satz. *Zu der Potenzreihe*

$$\sum_{k=0}^{\infty} a_k(x - a)^k \tag{2.4}$$

gibt es ein $r \in \overline{\mathbb{R}}$, $r \geq 0$, so dass die Reihe (2.4) für $|x - a| < r$ absolut konvergiert und für $|x - a| > r$ divergiert. Die Zahl r heißt Konvergenzradius der Reihe (2.4) und ist gegeben durch

$$r = \frac{1}{\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}}.$$

Das offene Intervall $(a - r, a + r)$ heißt Konvergenzintervall der Reihe (2.4). In jedem kompakten Teilintervall des Konvergenzintervalls konvergiert die Reihe (2.4) gleichmäßig.

Bemerkung. Achtung! Über die Konvergenz in den Endpunkten des Konvergenzintervalls wird hier nichts ausgesagt (und läßt sich auch allgemein nichts sagen). Es gibt Potenzreihen, die in keinem, einem oder beiden Endpunkten des Konvergenzintervalls konvergieren.

Ferner beachte man, dass i.a. nicht im ganzen Konvergenzintervall gleichmäßige Konvergenz vorliegt, sondern nur in kompakten Teilintervallen.

Beispiel.

$$\sum_{k=1}^{\infty} \frac{1}{k^m} x^k \quad \text{mit einem } m \in \mathbb{N}_0. \quad (2.5)$$

Wegen $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$ (vgl. Behauptung 1.11 aus Kapitel 5) ist der Konvergenzradius = 1. Ferner gilt:

- für $m = 0$ ist (2.5) divergent in 1 und -1
- für $m = 1$ ist (2.5) konvergent in -1 , divergent in 1
- für $m = 2$ ist (2.5) konvergent in -1 und 1.

Nehmen wir an, die Potenzreihe

$$\sum_{k=0}^{\infty} a_k (x - a)^k \quad (2.6)$$

habe den Konvergenzradius $0 < r < \infty$ und konvergiere etwa auch noch im Endpunkt $a + r$ des Konvergenzintervalls. Wir wollen zeigen, dass sie dann im Intervall $[a, a + r]$ gleichmäßig konvergiert. Der Unterschied zur früheren Argumentation ist, dass jetzt nicht notwendig absolute Konvergenz vorliegt, daher ist das Majorantenkriterium nicht anwendbar. O.B.d.A. können wir uns auf den Fall $a = 0, r = 1$ beschränken, der durch eine einfache Transformation erreichbar ist.

2.7 Satz. *Ist $\sum_{k=0}^{\infty} a_k$ konvergent, so ist $\sum_{k=0}^{\infty} a_k x^k$ gleichmäßig konvergent in $[0, 1]$.*

Beweis. Betrachte die Restglieder $b_k := \sum_{j=k+1}^{\infty} a_j$ für $k \in \mathbb{N}_0$. Dann ist $(b_k)_{k \in \mathbb{N}}$ eine Nullfolge, und es ist $b_{k-1} - b_k = a_k$. Also ergibt sich für $n > m$

$$\sum_{k=m+1}^n a_k x^k = b_m x^{m+1} - b_n x^n + \sum_{k=m+1}^{n-1} b_k (x^{k+1} - x^k).$$

Sei $\varepsilon \in \mathbb{R}^+$. Es gibt ein $n_0 \in \mathbb{N}$ mit $|b_k| < \varepsilon/3$ für $k \geq n_0$. Sei jetzt $m \geq n_0, p \in \mathbb{N}$. Dann gilt für beliebiges $x \in [0, 1]$

$$\begin{aligned} \left| \sum_{k=m+1}^{m+p} a_k x^k \right| &\leq |b_m| + |b_{m+p}| + \sum_{k=m+1}^{m+p-1} |b_k| |x^{k+1} - x^k| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \underbrace{\sum_{k=m+1}^{m+p-1} (x^k - x^{k+1})}_{= x^{m+1} - x^{m+p}} \leq \varepsilon. \end{aligned}$$

Aus dem Cauchy-Kriterium 1.1 folgt jetzt die behauptete gleichmäßige Konvergenz. ■

Aus Satz 2.7 und Satz 1.2 folgt insbesondere:

2.8 Folgerung. *Ist*

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \quad \text{für } x \in [0, 1],$$

so ist f in $[0, 1]$ stetig.

Eine durch eine konvergente Potenzreihe dargestellte Funktion ist also stetig. Allgemein sagen wir, falls

$$f(x) = \sum_{k=0}^{\infty} a_k (x-a)^k \quad \text{für } x \in (a-r, a+r) \quad (2.9)$$

mit $r > 0$ gilt, die Funktion f sei durch die obige Potenzreihe *dargestellt*, oder sie sei in eine Potenzreihe um a *entwickelt*.

Wir untersuchen jetzt die Differenzierbarkeit einer derart dargestellten Funktion. Dazu betrachten wir die durch gliedweise Differentiation von (2.9) entstehende Potenzreihe

$$\sum_{k=1}^{\infty} k a_k (x-a)^{k-1} \quad (2.10)$$

Wegen $\limsup_{n \rightarrow \infty} \sqrt[n]{(n+1)|a_{n+1}|} = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}$ hat sie denselben Konvergenzradius wie (2.9). In jedem kompakten Teilintervall des Konvergenzintervalls $(a-r, a+r)$ ist also (2.10) nach Satz 2.3 gleichmäßig konvergent; aus Satz 1.3 folgt daher die Differenzierbarkeit von f und die Konvergenz von (2.10) gegen f' . Auf f' kann man natürlich wieder denselben Schluß anwenden, usw. Auf diese Weise erhalten wir den folgenden Satz:

2.11 Satz. *Sei*

$$f(x) = \sum_{k=0}^{\infty} a_k (x-a)^k \quad \text{für } x \in (a-r, a+r)$$

mit positivem Konvergenzradius r . Dann ist f beliebig oft differenzierbar, und für $n \in \mathbb{N}$ gilt

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1) \cdots (k-n+1) a_k (x-a)^{k-n},$$

wobei die rechts stehende Reihe denselben Konvergenzradius r hat. Speziell ist

$$a_n = \frac{f^{(n)}(a)}{n!}.$$

Die letzte Aussage zeigt insbesondere, dass zwei verschiedene Potenzreihen (zur selben Stelle a) nicht dieselbe Funktion darstellen können. Mit anderen Worten: Aus

$$\sum_{k=0}^{\infty} a_k(x-a)^k = \sum_{k=0}^{\infty} b_k(x-a)^k$$

mit Konvergenz in einem Intervall um a folgt $a_k = b_k$ für $k \in \mathbb{N}_0$ (Eindeutigkeitssatz für Potenzreihen).

Taylorreihen

Wir haben eben gezeigt: Wenn die Funktion f durch eine (in einer Umgebung von a konvergente) Potenzreihe um a dargestellt wird, so ist diese gegeben durch

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k.$$

Hier sei kurz an Abschnitt 6.3 erinnert: Ist f in einer Umgebung von a beliebig oft differenzierbar, so hatten wir die Reihe

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k$$

als die *Taylorreihe* von f zur Stelle a bezeichnet. Wenn also f überhaupt durch eine Potenzreihe um a dargestellt werden kann, dann nur durch die Taylorreihe. Im konkreten Fall kommt es also darauf an, den Konvergenzradius der Taylorreihe zu ermitteln. Ist er positiv, so folgt aber allein daraus noch nicht, dass die Taylorreihe gegen die Funktion konvergiert, wie ein Beispiel in Abschnitt 6.3 zeigte. Um zu zeigen, dass eine gegebene Funktion im Konvergenzintervall wirklich durch ihre Taylorreihe dargestellt wird, muß man also entweder zeigen, dass das Restglied gegen Null konvergiert, oder, falls dies nicht gelingt, auf andere Weise schließen. Hierfür im folgenden einige Beispiele. Für die Abschätzung des Restgliedes haben wir die durch die Taylorformel gegebenen Darstellungen zur Verfügung: Wird

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + R_{n+1}(x)$$

gesetzt (und ist f auf einem Intervall definiert), so gilt nach Satz 3.2 aus Kapitel 6

$$R_{n+1}(x) = \frac{f^{(n+1)}(c)}{(n+1)!} (x-a)^{n+1}$$

mit einem (von n und x abhängenden) c zwischen a und x , und nach Satz 3.3 aus Kapitel 7 gilt

$$R_{n+1}(x) = \frac{1}{n!} \int_a^x f^{(n+1)}(t)(x-t)^n dt.$$

In manchen, aber nicht in allen Fällen kann man hiermit die gewünschte Konvergenz der Taylorreihe gegen f zeigen.

Wir wollen als Beispiele die Taylorreihen für einige der im Kapitel über spezielle Funktionen betrachteten Funktionen untersuchen. Für die Funktionen \exp , \sin , \cos wurde (für $a = 0$) bereits früher gezeigt, dass das Restglied gegen Null geht.

Wir betrachten die Logarithmus-Funktion. Als Entwicklungsstelle kommt nur ein Punkt des Definitionsbereiches \mathbb{R}^+ in Frage. Wir wählen $a = 1$. Für $f = \ln$ beweist man leicht durch vollständige Induktion

$$f^{(k)}(x) = (-1)^{k-1}(k-1)!x^{-k}, \quad (k \in \mathbb{N})$$

speziell $f^{(k)}(1) = (-1)^{k-1}(k-1)!$. Die Taylorreihe zur Stelle 1 lautet also

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

Der Konvergenzradius dieser Reihe ist offenbar 1; sie stellt also in (0,2) eine Funktion g dar. Eine Abschätzung des Restgliedes stößt auf Schwierigkeiten: Da man nichts über c_n weiß, kann man nicht ausschließen, dass $c_n = 1/2$ für alle n ist. Wegen

$$R_{n+1}(x) = \frac{(-1)^n (x-1)^{n+1}}{n+1 c_n^{n+1}}$$

würde dann aber für $0 < x < 1/2$ gelten:

$$\left| \frac{x-1}{c_n} \right| > 1,$$

also $R_{n+1}(x) \not\rightarrow 0$. Man kann aber auf andere Weise leicht zeigen, dass g in (0,2) mit der Funktion \ln übereinstimmt. Dazu differenzieren wir die Funktion $\ln - g$: Es ist

$$\ln' x - g'(x) = \frac{1}{x} - \sum_{k=1}^{\infty} (-1)^{k-1} (x-1)^{k-1} = \frac{1}{x} - \frac{1}{1+(x-1)} = 0.$$

Die Funktion $\ln -g$ ist also in $(0,2)$ konstant; da sie an der Stelle $x = 1$ gleich Null ist, ist sie überall Null. Damit ist

$$\ln x = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (x-1)^k \quad \text{für } x \in (0, 2)$$

gezeigt. Die rechts stehende Reihe konvergiert nach dem Leibnizkriterium auch für $x = 2$. Da die dargestellte Funktion nach Folgerung 2.8 (passend transformiert) in 2 noch stetig ist, stimmt sie dort mit $\ln 2$ überein. Wir können das Ergebnis auch in der Form

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k \quad \text{für } -1 < x \leq 1$$

schreiben.

Als nächstes Beispiel betrachten wir die Funktion $f = \arctan$, die wir in eine Potenzreihe um 0 entwickeln wollen. Zur Berechnung der n -ten Ableitung an der Stelle 0 kann man den Ansatz

$$f^{(n)}(x) = \frac{P_n(x)}{(1+x^2)^n}$$

machen und findet für P_n eine Rekursionsformel, aus der sich herleiten läßt, dass

$$f^{(2n)}(0) = 0, \quad f^{(2n+1)}(0) = (-1)^n (2n)!$$

ist. Die Taylorreihe der Funktion \arctan lautet also

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

Der Konvergenzradius ist offenbar 1. Sei f die in $(-1, 1)$ dargestellte Funktion. Es gilt

$$f'(x) = \sum_{k=0}^{\infty} (-1)^k x^{2k} = \frac{1}{1+x^2} = \arctan' x,$$

ferner $f(0) = 0 = \arctan 0$. Also ist $f(x) = \arctan x$. Wir haben also

$$\begin{aligned} \arctan x &= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} \quad \text{in } (-1, 1). \end{aligned}$$

Nach dem Leibniz-Kriterium konvergiert die Reihe auch für $x = 1$ und $x = -1$. Nach Folgerung 2.8 ist die dargestellte Funktion dort stetig, stimmt also mit der Funktion \arctan überein.

Als letztes Beispiel wollen wir die allgemeine Potenz $x \mapsto x^\alpha$ betrachten. Als Entwicklungsstelle wählen wir ebenfalls 1, aber es schreibt sich bequemer, die Funktion $x \mapsto (1+x)^\alpha$ zu nehmen und um 0 zu entwickeln. Für $f(x) = (1+x)^\alpha$ ($x > -1$) berechnet man

$$f^{(k)}(x) = \alpha(\alpha-1) \cdots (\alpha-k+1)(1+x)^{\alpha-k}.$$

Zur übersichtlicheren Schreibweise wollen wir wie früher den Binomialkoeffizienten $\binom{\alpha}{k}$ erklären durch

$$\binom{\alpha}{k} := \frac{\alpha(\alpha-1) \cdots (\alpha-k+1)}{k!}.$$

Dann ist die Taylorreihe der Funktion f zur Stelle 0 gegeben durch

$$\sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.$$

Im Fall $\alpha \in \mathbb{N}_0$ sind fast alle Koeffizienten 0, es liegt also ein Polynom vor. Wir setzen jetzt $\alpha \notin \mathbb{N}_0$ voraus.

Aus dem Quotientenkriterium 2.9 aus Kapitel 3 folgt sofort, dass der Konvergenzradius gleich 1 ist. Die Reihe stellt also in $(-1, 1)$ eine Funktion g dar. Setzen wir

$$h(x) := \frac{g(x)}{(1+x)^\alpha} \quad \text{für } -1 < x < 1,$$

so gilt

$$h'(x) = \frac{g'(x)(1+x)^\alpha - g(x)\alpha(1+x)^{\alpha-1}}{(1+x)^{2\alpha}} = \frac{g'(x)(1+x) - \alpha g(x)}{(1+x)^{\alpha+1}}.$$

Nun ist

$$\begin{aligned} (1+x)g'(x) &= (1+x) \sum_{k=1}^{\infty} \binom{\alpha}{k} k x^{k-1} \\ &= \sum_{k=0}^{\infty} \binom{\alpha}{k+1} (k+1) x^k + \sum_{k=1}^{\infty} \binom{\alpha}{k} k x^k \\ &= \sum_{k=0}^{\infty} \left\{ \binom{\alpha}{k+1} (k+1) + \binom{\alpha}{k} k \right\} x^k \\ &= \alpha \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k = \alpha g(x), \end{aligned}$$

also $h' = 0$ und daher $h = \text{const.}$ Wegen $h(0) = 1$ folgt

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \quad \text{für } -1 < x < 1.$$

Man nennt diese Reihe auch die „binomische Reihe“. Das Konvergenzverhalten an den Stellen ± 1 ist etwas schwieriger zu beurteilen:

Bemerkung. Sei $\alpha \in \mathbb{R} \setminus \mathbb{N}_0$. Die binomische Reihe $\sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$ ist für

	in $x = 1$	in $x = -1$
$\alpha > 0$	konvergent	konvergent
$-1 < \alpha < 0$	konvergent	divergent
$\alpha \leq -1$	divergent	divergent

8.3 Fourierreihen

Ein wichtiger Typ von Reihenentwicklungen wird gegeben durch die Fourierreihen. Fourier war Physiker, und wir wollen zunächst kurz die physikalische Fragestellung, bei deren Behandlung Fourier (1807) erstmals die heute nach ihm benannten Reihen verwendete, heuristisch erläutern.

Fourier befaßte sich mit Problemen der Wärmeleitung in homogenen Medien. Als besonders einfaches (idealisiertes) Wärmeleitungsproblem betrachten wir einen geraden Stab der Länge l , den wir als „unendlich dünn“ annehmen. Seine Punkte können also durch eine Koordinate x beschrieben werden, die das Intervall $[0, l]$ durchläuft. In diesem Stab herrsche an der Stelle x zur Zeit $t \geq 0$ die Temperatur $T(x, t)$. Wir nehmen an, die beiden Enden des Stabes würden von außen ständig auf der Temperatur Null gehalten, und im übrigen sei der Stab vollkommen isoliert. Zur Zeit $t = 0$ herrsche eine bekannte Temperaturverteilung, gegeben durch eine Funktion $f : [0, l] \rightarrow \mathbb{R}$. Von der Temperaturverteilung $T(x, t)$ ist also bekannt, dass

$$T(x, 0) = f(x) \quad \text{für } x \in [0, l]$$

und

$$T(0, t) = T(l, t) = 0 \quad \text{für } t \geq 0$$

ist. Im Lauf der Zeit werden sich nun auf Grund der Wärmeleitung die Temperaturunterschiede ausgleichen. Die Frage ist, welche Temperaturverteilung zu einer Zeit $t > 0$ herrscht, also wie man die Funktion $T(x, t)$ aus der Anfangsverteilung $T(x, 0)$ berechnen kann.

Physikalische Grundannahmen und Überlegungen führen zu der Folgerung, dass die Funktion T (unter, wie üblich, geeigneten Differenzierbarkeitsannahmen) der partiellen Differentialgleichung

$$\frac{\partial T}{\partial t} = \mu \frac{\partial^2 T}{\partial x^2}$$

genügen muß, wobei $\mu > 0$ eine Materialkonstante ist. Wir stellen nun zunächst die Frage der Anfangsverteilung zurück und suchen allgemeiner eine Lösung $\varphi : [0, l] \times [0, \infty) \rightarrow \mathbb{R}$ des Problems

$$\frac{\partial \varphi}{\partial t} = \mu \frac{\partial^2 \varphi}{\partial x^2}, \quad \varphi(0, t) = \varphi(l, t) = 0. \quad (3.1)$$

Um spezielle Lösungen zu finden, kann man untersuchen, ob es Lösungen der Form

$$\varphi(x, t) = h(t)g(x)$$

gibt („Separationsansatz“, „Trennung der Veränderlichen“). Hierzu muß wegen

$$\frac{\partial \varphi}{\partial t} = h'(t)g(x), \quad \frac{\partial^2 \varphi}{\partial x^2} = h(t)g''(x)$$

also

$$\frac{h'(t)}{h(t)} = \mu \frac{g''(x)}{g(x)}$$

sein. Da die linke Seite nicht von x und die rechte nicht von t abhängt, handelt es sich um eine Konstante c . Somit werden wir auf die Gleichungen

$$\begin{aligned} h'(t) &= ch(t), \\ \mu g''(x) &= cg(x) \end{aligned}$$

geführt. Die erste Differentialgleichung läßt sich sofort lösen:

$$h(t) = be^{ct}$$

mit einer Konstanten b . Bei der zweiten ist die Randbedingung

$$g(0) = g(l) = 0$$

zu berücksichtigen. Es liegt daher der Ansatz $g(x) = \sin ax$ nahe. Die Differentialgleichung ist erfüllt, wenn $al = k\pi$ mit ganzzahligem k ist. Für die Konstanten ergibt sich also

$$a = \frac{k\pi}{l}, \quad c = -\frac{\mu k^2 \pi^2}{l^2}.$$

Insgesamt erhalten wir, dass bei beliebigem $k \in \mathbb{N}$ durch

$$\varphi_k(x, t) = b_k e^{-\frac{\mu k^2 \pi^2}{l^2} t} \sin \frac{k\pi x}{l}$$

mit $b_k = \text{const.}$ eine Lösung des Problems (3.1) gegeben ist.

Natürlich kann hierdurch i.a. noch nicht die gesuchte Temperaturverteilung $T(x, t)$ gegeben sein, denn es sollte ja $T(x, 0)$ eine vorgeschriebene Funktion $f(x)$ sein. Wir erhalten aber

$$\varphi_k(x, 0) = b_k \sin \frac{k\pi x}{l},$$

also für jedes k eine sehr spezielle Funktion.

Nun beachte man aber, dass unser Problem „linear“ ist. Dies hat zur Folge, dass die Summe von Lösungen auch eine Lösung ist, also ist durch

$$\varphi(x, t) := \sum_{k=1}^n \varphi_k(x, t) = \sum_{k=1}^n b_k e^{-\frac{\mu k^2 \pi^2}{l^2} t} \sin \frac{k\pi x}{l}$$

($n \in \mathbb{N}$) ebenfalls eine Lösung gegeben. Für diese Lösung ist

$$\varphi(x, 0) = \sum_{k=1}^n b_k \sin \frac{k\pi x}{l},$$

und im allgemeinen wird es immer noch nicht möglich sein, die Zahl n und die Koeffizienten b_k so zu wählen, dass $\varphi(x, 0) = f(x)$ ist. Hier setzt nun Fouriers entscheidende (und damals kühne) Überlegung ein. Er geht davon aus, dass man die weitgehend willkürliche Funktion f durch die *unendliche* Reihe

$$\sum_{k=1}^{\infty} b_k \sin \frac{k\pi x}{l}$$

darstellen kann, wenn man die Koeffizienten b_k geeignet bestimmt. Alsdann setze man

$$T(x, t) := \sum_{k=1}^{\infty} b_k e^{-\frac{\mu k^2 \pi^2}{l^2} t} \sin \frac{k\pi x}{l}.$$

Wenn die obere Reihe für jedes $x \in [0, l]$ absolut konvergiert, dann nach dem Majorantenkriterium auch die untere. Die „Anfangsbedingung“ $T(x, 0) = f(x)$ (für $x \in [0, l]$) und die „Randbedingung“ $T(0, t) = T(l, t) = 0$ (für $t \geq 0$) sind dann erfüllt. Falls es erlaubt ist, die Reihe gliedweise zu differenzieren (einmal nach t , zweimal nach x), erhält man, dass T auch der

vorgeschriebenen partiellen Differentialgleichung genügt. Das Problem der Temperaturverteilung ist damit als gelöst anzusehen.

Der springende Punkt bei dieser Argumentation ist natürlich die Frage, ob es möglich ist, eine gegebene Funktion $f : [0, l] \rightarrow \mathbb{R}$ mit $f(0) = f(l) = 0$ darzustellen als Reihe der Form

$$f(x) = \sum_{k=1}^{\infty} b_k \sin \frac{k\pi x}{l}.$$

Mit einer Variante dieser Fragestellung befassen wir uns im folgenden. Es bedeutet dabei keine Einschränkung der Allgemeinheit, $l = \pi$ anzunehmen. Andererseits soll die Voraussetzung $f(0) = f(l) = 0$ fallengelassen werden, und wir fragen allgemeiner nach der Entwicklungsmöglichkeit in Reihen der Form

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx).$$

Es ist bequem, Funktionen zu betrachten, die auf ganz \mathbb{R} definiert sind. Wenn eine solche Funktion durch die obige Reihe dargestellt wird, ist $f(x + 2\pi) = f(x)$ für $x \in \mathbb{R}$. Funktionen mit dieser Eigenschaft nennen wir 2π -periodisch.

Wir nehmen nun an, für eine gegebene Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ existiere in der Tat eine Darstellung der Form

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \quad x \in \mathbb{R},$$

wobei die Reihe sogar gleichmäßig konvergent sei. Dann lassen sich die Koeffizienten a_k, b_k folgendermaßen berechnen. Multiplikation mit $\cos mx$ und Integration über $[0, 2\pi]$ (beachte, dass f wegen der gleichmäßigen Konvergenz stetig ist) ergibt

$$\begin{aligned} & \int_0^{2\pi} f(x) \cos mx \, dx \\ &= \frac{a_0}{2} \int_0^{2\pi} \cos mx \, dx + \int_0^{2\pi} \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) \cos mx \, dx. \end{aligned}$$

Die rechts stehende Reihe ist ebenfalls gleichmäßig konvergent (nämlich gegen $f(x) \cos mx$; beachte $\|fg\| \leq \|f\| \|g\|$), darf also nach Satz 2.4 aus Kapitel 7 gliedweise integriert werden. Es ist also

$$\int_0^{2\pi} f(x) \cos mx \, dx$$

$$= \frac{a_0}{2} \int_0^{2\pi} \cos mx \, dx + \sum_{k=1}^{\infty} \left(a_k \int_0^{2\pi} \cos kx \cos mx \, dx + b_k \int_0^{2\pi} \sin kx \cos mx \, dx \right).$$

Nun findet man leicht durch partielle Integration:

$$\int_0^{2\pi} \cos kx \cos mx \, dx = \int_0^{2\pi} \sin kx \sin mx \, dx = 0 \quad \text{für } k \neq m$$

$$\int_0^{2\pi} \sin kx \cos mx \, dx = 0,$$

$$\int_0^{2\pi} \cos^2 kx \, dx = \int_0^{2\pi} \sin^2 kx \, dx = \pi \quad \text{für } k \geq 1.$$

Damit erhält man

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx \, dx \quad \text{für } k \in \mathbb{N}_0.$$

Ganz analog beweist man

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx \, dx \quad \text{für } k \in \mathbb{N}.$$

Man bezeichnet nun ganz allgemein für jede 2π -periodische Funktion f und für $k \in \mathbb{N}_0$ die Zahlen

$$a_k := \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx \, dx,$$

$$b_k := \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx \, dx,$$

($k \in \mathbb{N}_0$) falls diese Integrale existieren, als die *Fourierkoeffizienten* von f und nennt die Funktionenreihe

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$$

die *Fourierreihe* von f . Dabei ist (analog wie bei Taylorreihen) nichts darüber ausgesagt, ob diese Reihe überhaupt (und für welche x bzw. in welchem Sinne) konvergiert und ob sie im Konvergenzfall gegen $f(x)$ konvergiert. Die Beantwortung dieser Fragen erfordert eingehendere Untersuchungen.

Bevor wir hierauf eingehen, wollen wir aber die hier sehr zweckmäßige komplexe Schreibweise einführen. Es ist ja

$$e^{ix} = \cos x + i \sin x, \quad \cos x = \frac{1}{2}(e^{ix} + e^{-ix}), \quad \sin x = \frac{1}{2i}(e^{ix} - e^{-ix}).$$

Damit ergibt sich für $k \in \mathbb{N}_0$

$$\begin{aligned} a_k \cos kx + b_k \sin kx &= a_k \cdot \frac{1}{2}(e^{ikx} + e^{-ikx}) - b_k \cdot \frac{i}{2}(e^{ikx} - e^{-ikx}) \\ &= \frac{1}{2}(a_k - ib_k)e^{ikx} + \frac{1}{2}(a_k + ib_k)e^{-ikx} \\ &= c_k e^{ikx} + c_{-k} e^{-ikx} \end{aligned}$$

mit

$$c_k := \frac{1}{2}(a_k - ib_k), \quad c_{-k} := \frac{1}{2}(a_k + ib_k),$$

also für $n \in \mathbb{N}$

$$\begin{aligned} \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx) \\ &= c_0 + \sum_{k=1}^n (c_k e^{ikx} + c_{-k} e^{-ikx}) \\ &= \sum_{k=-n}^n c_k e^{ikx} \end{aligned}$$

Die Folge $(\sum_{k=-n}^n c_k e^{ikx})_{n \in \mathbb{N}}$ kürzt man ab mit

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}.$$

Dies ist also ebenfalls die Fourierreihe von f , nur in komplexer Schreibweise.

Es ist weiter zweckmäßig, auch komplexwertige Funktionen $f : \mathbb{R} \rightarrow \mathbb{C}$ zuzulassen. Jede solche Funktion läßt sich eindeutig darstellen in der Form $f = u + iv$ mit reellen Funktionen u, v . Man definiert dann (falls u und v , eingeschränkt auf $[a, b]$, Regelfunktionen sind)

$$\int_a^b f(x) dx := \int_a^b u(x) dx + i \int_a^b v(x) dx.$$

Mit dieser Konvention ist für $k \geq 0$

$$\begin{aligned} c_k &= \frac{1}{2}(a_k - ib_k) \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(x)(\cos kx - i \sin kx) dx \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(x)e^{-ikx} dx \end{aligned}$$

und

$$c_{-k} = \frac{1}{2}(a_k + ib_k) = \frac{1}{2\pi} \int_0^{2\pi} f(x)e^{ikx} dx,$$

somit

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x)e^{-ikx} dx \quad \text{für } k \in \mathbb{Z}.$$

Wir wollen hier auch komplexwertige Funktionen f zulassen, erklären dann die Fourierkoeffizienten c_k von f durch diese Gleichung (falls die Integrale existieren) und nennen die Reihe

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

die *Fourierreihe* von f .

Die Konvergenztheorie der Fourierreihen wird dann besonders einfach und übersichtlich, wenn man nicht nach der (i.a. selbst bei stetigen reellen Funktionen nicht vorliegenden) punktwweisen Konvergenz fragt, sondern nach Konvergenz in einem schwächeren Sinne. Mit dieser auch für Anwendungen wichtigen Konvergenz wollen wir uns hier befassen.

Konvergenz im quadratischen Mittel

Wir betrachten die Menge V der Funktionen $f : \mathbb{R} \rightarrow \mathbb{C}$ mit $f(x+2\pi) = f(x)$ für $x \in \mathbb{R}$ und der Eigenschaft, dass Real- und Imaginärteil von f , eingeschränkt auf $[0, 2\pi]$, Regelfunktionen sind. Offenbar ist V (mit den üblichen Verknüpfungen) ein Vektorraum über \mathbb{C} .

Definition. Für $f, g \in V$ sei

$$\langle f, g \rangle := \frac{1}{2\pi} \int_0^{2\pi} f(x) \overline{g(x)} dx.$$

Die komplexe Zahl $\langle f, g \rangle$ heißt das *Skalarprodukt* von f und g .

3.2 Satz. Die Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ hat die folgenden Eigenschaften (für $f, g, h \in V$ und $\lambda \in \mathbb{C}$):

- (1) $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$
- (2) $\langle f, g + h \rangle = \langle f, g \rangle + \langle f, h \rangle$
- (3) $\langle \lambda f, g \rangle = \lambda \langle f, g \rangle$
- (4) $\langle f, \lambda g \rangle = \overline{\lambda} \langle f, g \rangle$
- (5) $\langle g, f \rangle = \overline{\langle f, g \rangle}$
- (6) $\langle f, f \rangle \geq 0$ (d.h. $\langle f, f \rangle \in \mathbb{R}$ und ≥ 0)
- (7) $|\langle f, g \rangle|^2 \leq \langle f, f \rangle \langle g, g \rangle$
- (8) $\sqrt{\langle f + g, f + g \rangle} \leq \sqrt{\langle f, f \rangle} + \sqrt{\langle g, g \rangle}$.

Beweis. (1) – (6) folgen unmittelbar aus der Definition; bei (6) beachte man $f\bar{f} = |f|^2 \geq 0$. Wegen (6) sind die Wurzeln in (8) im Reellen erklärt.

(7) Nach (6) gilt für $f, g \in V$ und $\lambda \in \mathbb{C}$

$$\langle f + \lambda g, f + \lambda g \rangle \geq 0,$$

also unter Verwendung von (1) – (5)

$$\langle f, f \rangle + \overline{\lambda} \langle f, g \rangle + \lambda \overline{\langle f, g \rangle} + \lambda \overline{\lambda} \langle g, g \rangle \geq 0.$$

Ist $\langle g, g \rangle \neq 0$, so kann man

$$\lambda = -\frac{\langle f, g \rangle}{\langle g, g \rangle}$$

einsetzen und erhält die behauptete Ungleichung. Ist $\langle g, g \rangle = 0$, so setze man $\lambda = -n \langle f, g \rangle$; es folgt $|\langle f, g \rangle|^2 \leq 1/2n \langle f, f \rangle$, also (da $n \in \mathbb{N}$ beliebig ist) $|\langle f, g \rangle| \leq 0 = \langle f, f \rangle \langle g, g \rangle$.

(8) Aus (7) folgt

$$\operatorname{Re} \langle f, g \rangle \leq |\langle f, g \rangle| \leq \sqrt{\langle f, f \rangle \langle g, g \rangle},$$

also

$$\begin{aligned}
\langle f + g, f + g \rangle &= \langle f, f \rangle + \langle f, g \rangle + \overline{\langle f, g \rangle} + \langle g, g \rangle \\
&= \langle f, f \rangle + \langle g, g \rangle + 2\operatorname{Re}\langle f, g \rangle \\
&\leq \langle f, f \rangle + \langle g, g \rangle + 2\sqrt{\langle f, f \rangle \langle g, g \rangle} \\
&= \left(\sqrt{\langle f, f \rangle} + \sqrt{\langle g, g \rangle} \right)^2,
\end{aligned}$$

woraus (8) folgt. ■

Definition. Für $f \in V$ ist

$$\|f\|_2 := \sqrt{\langle f, f \rangle}$$

die L_2 -Norm oder *Hilbert-Norm* von f .

Die Ungleichungen (7) und (8) aus Satz 3.2 schreiben sich damit in der Form

$$|\langle f, g \rangle| \leq \|f\|_2 \|g\|_2$$

(*Cauchy-Schwarzsche Ungleichung*) und

$$\|f + g\|_2 \leq \|f\|_2 + \|g\|_2;$$

letzteres ist die Dreiecksungleichung für die L_2 -Norm.

Wir erinnern uns daran, dass wir schon früher Konvergenz im Sinne einer Norm erklärt hatten, und definieren:

Definition. Seien $f, f_n \in V$ ($n \in \mathbb{N}$). Die Folge $(f_n)_{n \in \mathbb{N}}$ heißt *konvergent im quadratischen Mittel* (oder konvergent in der L_2 -Norm) gegen f , wenn

$$\lim_{n \rightarrow \infty} \|f - f_n\|_2 = 0$$

ist.

Die Bezeichnung „im quadratischen Mittel“ ist naheliegend wegen

$$\|f - f_n\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x) - f_n(x)|^2 dx;$$

der rechts stehende Ausdruck läßt sich als Mittelwert der quadratischen Abweichung von f und f_n ansehen. Man beachte, dass eine im quadratischen Mittel konvergierende Funktionenfolge nicht punktweise konvergent zu sein braucht.

Wir betrachten jetzt speziell die durch

$$e_k(x) := e^{ikx}$$

definierten Elemente $e_k \in V$ ($k \in \mathbb{Z}$) und berechnen $\langle e_k, e_m \rangle$. Nach Definition ist

$$\begin{aligned} \langle e_k, e_m \rangle &= \frac{1}{2\pi} \int_0^{2\pi} e^{ikx} \overline{e^{imx}} dx = \frac{1}{2\pi} \int_0^{2\pi} e^{i(k-m)x} dx \\ &= \frac{1}{2\pi} \left(\int_0^{2\pi} \cos(k-m)x dx + i \int_0^{2\pi} \sin(k-m)x dx \right) \\ &= \begin{cases} 1 & \text{für } k = m, \\ 0 & \text{für } k \neq m. \end{cases} \end{aligned}$$

Allgemein nennt man zwei Elemente $f, g \in V$ mit $\langle f, g \rangle = 0$ *orthogonal*, und eine Folge $(b_k)_{k \in \mathbb{Z}}$ in V heißt *Orthonormalsystem*, wenn

$$\langle b_k, b_m \rangle = \begin{cases} 1 & \text{für } k = m \\ 0 & \text{für } k \neq m \end{cases}$$

ist.

Im folgenden sei jetzt zunächst ein beliebiges Orthonormalsystem $(b_k)_{k \in \mathbb{Z}}$ gegeben. Ist $f \in V$, so nennt man die durch

$$c_k := \langle f, b_k \rangle$$

definierten komplexen Zahlen die *Fourierkoeffizienten* von f bezüglich des gegebenen Orthonormalsystems. Die früher definierten speziellen Fourierkoeffizienten sind also genau diejenigen bezüglich des speziellen Orthonormalsystems $(e_k)_{k \in \mathbb{Z}}$. Die Reihe

$$\sum_{k \in \mathbb{Z}} c_k b_k$$

nennt man die *Fourierreihe* von f bezüglich des Orthonormalsystems $(b_k)_{k \in \mathbb{Z}}$.

Sei jetzt $f \in V$ gegeben, und seien $\alpha_{-n}, \dots, \alpha_n$ beliebige komplexe Zahlen. Wir wollen

$$\left\| f - \sum_{k=-n}^n \alpha_k b_k \right\|_2^2$$

berechnen und aus dem Ergebnis wichtige Folgerungen ziehen. Es ist (Summation jeweils von $-n$ bis n)

$$\begin{aligned} \left\| f - \sum \alpha_k b_k \right\|_2^2 &= \left\langle f - \sum \alpha_k b_k, f - \sum \alpha_j b_j \right\rangle \\ &= \langle f, f \rangle - \sum \alpha_k \underbrace{\langle b_k, f \rangle}_{\bar{c}_k} - \sum \bar{\alpha}_j \underbrace{\langle f, b_j \rangle}_{c_j} + \left\langle \sum \alpha_k b_k, \sum \alpha_j b_j \right\rangle \end{aligned}$$

und

$$\left\langle \sum \alpha_k b_k, \sum \alpha_j b_j \right\rangle = \sum_{k,j} \alpha_k \bar{\alpha}_j \langle b_k, b_j \rangle = \sum_k \alpha_k \bar{\alpha}_k.$$

Ferner ist

$$\begin{aligned} \sum |c_k - \alpha_k|^2 &= \sum (c_k - \alpha_k)(\bar{c}_k - \bar{\alpha}_k) \\ &= \sum c_k \bar{c}_k - \sum \alpha_k \bar{c}_k - \sum \bar{\alpha}_k c_k + \sum \alpha_k \bar{\alpha}_k. \end{aligned}$$

Damit ergibt sich

$$\left\| f - \sum \alpha_k b_k \right\|_2^2 = \|f\|_2^2 - \sum |c_k|^2 + \sum |c_k - \alpha_k|^2.$$

Hieran lesen wir folgendes ab:

- (1) $\|f - \sum_{k=-n}^n \alpha_k b_k\|_2^2$ wird genau dann minimal, wenn $\alpha_k = c_k$ für $k = -n, \dots, n$ ist. Dies ist also eine Kennzeichnung der Fourierkoeffizienten durch eine Extremaleigenschaft: f wird durch eine Linearkombination der Vektoren b_{-n}, \dots, b_n im Sinne der L_2 -Norm genau dann am besten approximiert, wenn man als Koeffizienten die Fourierkoeffizienten wählt.
- (2) Wählen wir jetzt $\alpha_k = c_k$, so lautet die obige Gleichung

$$\|f\|_2^2 - \sum_{k=-n}^n |c_k|^2 = \left\| f - \sum_{k=-n}^n c_k b_k \right\|_2^2,$$

und dies ist ≥ 0 . Es folgt die Konvergenz der Reihe $\sum_{k=-\infty}^{\infty} |c_k|^2$ und die Ungleichung

$$\sum_{k=-\infty}^{\infty} |c_k|^2 \leq \|f\|_2^2.$$

Genau dann gilt hier das Gleichheitszeichen, wenn

$$\lim_{n \rightarrow \infty} \left\| f - \sum_{k=-n}^n c_k b_k \right\|_2 = 0$$

ist.

Wir fassen die erhaltenen Ergebnisse zusammen und ergänzen sie durch einige neue Bezeichnungen.

3.3 Satz. Sei $(b_k)_{k \in \mathbb{Z}}$ ein Orthonormalsystem in V , sei $f \in V$, und seien $c_k = \langle f, b_k \rangle$, $k \in \mathbb{Z}$, die Fourierkoeffizienten von f bezüglich des Orthonormalsystems $(b_k)_{k \in \mathbb{Z}}$. Dann gilt

$$\sum_{k=-\infty}^{\infty} |c_k|^2 \leq \|f\|_2^2 \quad (\text{Besselsche Ungleichung}).$$

Genau dann konvergiert die Fourierreihe von f im quadratischen Mittel gegen f , d.h. es gilt

$$\lim_{n \rightarrow \infty} \left\| f - \sum_{k=-n}^n c_k b_k \right\|_2 = 0,$$

wenn

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \|f\|_2^2 \quad (\text{Parsevalsche Vollständigkeitsrelation}) \quad (3.4)$$

gilt. Das Orthonormalsystem $(b_k)_{k \in \mathbb{Z}}$ heißt vollständig, wenn (3.4) für jedes $f \in V$ gilt.

Nach diesen allgemeinen Betrachtungen kehren wir zum speziellen Orthonormalsystem $(e_k)_{k \in \mathbb{Z}}$ zurück und zeigen:

3.5 Satz. Das Orthonormalsystem $(e_k)_{k \in \mathbb{Z}}$ ist vollständig.

Für jede Funktion $f \in V$ konvergiert also die Fourierreihe

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

mit

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx$$

im quadratischen Mittel gegen $f(x)$, d.h. es ist

$$\lim_{n \rightarrow \infty} \int_0^{2\pi} \left| f(x) - \sum_{k=-n}^n c_k e^{ikx} \right|^2 dx = 0$$

oder, anders geschrieben,

$$\lim_{n \rightarrow \infty} \|f - S_n[f]\|_2 = 0,$$

wenn wir (wie auch im folgenden) mit

$$S_n[f] := \sum_{k=-n}^n c_k e^{ikx}$$

die n -te Partialsumme der Fourierreihe von f bezeichnen. Nach Satz 3.3 ist dies gleichwertig mit

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx. \quad (3.6)$$

von Satz 3.5 bzw. (3.6). Zum Beweis von (3.6) nehmen wir zunächst f von einer sehr speziellen Gestalt an. Sei $0 \leq a \leq b \leq 2\pi$ und

$$f(x) := \begin{cases} 1 & \text{für } a \leq x \leq b, \\ 0 & \text{für } 0 \leq x < a \text{ und } b < x \leq 2\pi; \end{cases}$$

im übrigen sei $f(x+2\pi) = f(x)$. Für die komplexen Fourierkoeffizienten von f erhalten wir

$$c_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx = \frac{b-a}{2\pi}$$

und für $k \neq 0$

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx = \frac{1}{2\pi} \left[\frac{-1}{ik} e^{-ikx} \right]_a^b = \frac{i}{2\pi k} (e^{-ikb} - e^{-ika}),$$

also

$$|c_k|^2 = \frac{1 - \cos k(b-a)}{2\pi^2 k^2}.$$

Somit ist

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \frac{(b-a)^2}{4\pi^2} + \frac{1}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} - \frac{1}{\pi^2} \sum_{k=1}^{\infty} \frac{\cos k(b-a)}{k^2}.$$

Zur Berechnung der letzten Summe folgt eine längere Zwischenrechnung.

Behauptung.

$$\sum_{k=1}^{\infty} \frac{\sin kx}{k} = \frac{\pi - x}{2} \quad \text{für } 0 < x < 2\pi.$$

Beweis.

$$\begin{aligned} 1 + 2 \sum_{k=1}^n \cos kt &= \sum_{k=-n}^n e^{ikt} = e^{-int} \sum_{k=0}^{2n} e^{ikt} \\ &= e^{-int} \frac{e^{i(2n+1)t} - 1}{e^{it} - 1} = \frac{e^{i(n+\frac{1}{2})t} - e^{-i(n+\frac{1}{2})t}}{e^{\frac{i}{2}t} - e^{-\frac{i}{2}t}} \\ &= \frac{\sin(n + \frac{1}{2})t}{\sin \frac{1}{2}t}, \end{aligned}$$

also

$$\sum_{k=1}^n \cos kt = \frac{\sin(n + \frac{1}{2})t}{2 \sin \frac{1}{2}t} - \frac{1}{2}.$$

Integration von π bis x ergibt

$$\sum_{k=1}^n \frac{\sin kx}{k} = \int_{\pi}^x \frac{\sin(n + \frac{1}{2})t}{2 \sin \frac{1}{2}t} dt - \frac{x - \pi}{2}.$$

Für $n \rightarrow \infty$ konvergiert das Integral gegen 0, da die Fourierkoeffizienten einer Regelfunktion wegen der Besselschen Ungleichung gegen 0 konvergieren. ■

Behauptung. Für $\delta \in (0, \pi)$ ist die Reihe $\sum_{k=1}^{\infty} \sin kx/k$ gleichmäßig konvergent in $[\delta, 2\pi - \delta]$.

Beweis. Setze

$$s_n(x) := \sum_{k=1}^n \sin kx = \operatorname{Im} \left(\sum_{k=1}^n e^{ikx} \right).$$

Für $\delta \leq x \leq 2\pi - \delta$ gilt

$$\begin{aligned} |s_n(x)| &\leq \left| \sum_{k=1}^n e^{ikx} \right| = |e^{ix}| \left| \frac{e^{inx} - 1}{e^{ix} - 1} \right| \\ &\leq \frac{2}{\left| e^{\frac{ix}{2}} - e^{-\frac{ix}{2}} \right|} = \frac{1}{\sin \frac{x}{2}} \leq \frac{1}{\sin \frac{\delta}{2}}. \end{aligned}$$

Für $m > n > 0$ folgt

$$\begin{aligned} \left| \sum_{k=n}^m \frac{\sin kx}{k} \right| &= \left| \sum_{k=n}^m \frac{s_k(x) - s_{k-1}(x)}{k} \right| \\ &= \left| \sum_{k=n}^m s_k(x) \left(\frac{1}{k} - \frac{1}{k+1} \right) + \frac{s_m(x)}{m+1} - \frac{s_{n-1}(x)}{n} \right| \\ &\leq \frac{1}{\sin \frac{\delta}{2}} \left(\frac{1}{n} - \frac{1}{m+1} + \frac{1}{m+1} + \frac{1}{n} \right) = \frac{2}{n \sin \frac{\delta}{2}}. \end{aligned}$$

Nach dem Cauchy-Kriterium 1.1 folgt die gleichmäßige Konvergenz. ■

Behauptung.

$$\sum_{k=1}^{\infty} \frac{\cos kx}{k^2} = \left(\frac{x - \pi}{2} \right)^2 - \frac{\pi^2}{12} \quad \text{für } x \in [0, 2\pi].$$

Beweis. Setze $F(x) := \sum_{k=1}^{\infty} \cos kx/k^2$. Diese Reihe ist nach dem Majorantenkriterium gleichmäßig konvergent. Für beliebiges $\delta > 0$ ist die Reihe $-\sum_{k=1}^{\infty} \sin kx/k$ auf $[\delta, 2\pi - \delta]$ nach der 2. Behauptung gleichmäßig konvergent. Nach Satz 1.3 und der 1. Behauptung folgt

$$F'(x) = \frac{x - \pi}{2}, \quad \text{also } F(x) = \left(\frac{x - \pi}{2} \right)^2 + c.$$

Nun ist

$$\int_0^{2\pi} F(x) dx = \frac{\pi^3}{6} + 2\pi c,$$

und dies ist andererseits nach Satz 2.4 aus Kapitel 7

$$= \int_0^{2\pi} \sum_{k=1}^{\infty} \frac{\cos kx}{k^2} dx = \sum_{k=1}^{\infty} \frac{1}{k^2} \int_0^{2\pi} \cos kx dx = 0.$$

Also ist $c = -\frac{\pi^2}{12}$, womit die 3. Behauptung bewiesen ist. ■

Jetzt können wir unsere oben begonnene Rechnung fortsetzen und erhalten

$$\begin{aligned} \sum_{k=-\infty}^{\infty} |c_k|^2 &= \frac{(b-a)^2}{4\pi^2} + \frac{1}{\pi^2} \cdot \frac{\pi^2}{6} - \frac{1}{\pi^2} \left(\frac{b-a-\pi}{2} \right)^2 + \frac{1}{\pi^2} \cdot \frac{\pi^2}{12} \\ &= \frac{b-a}{2\pi} = \|f\|_2^2. \end{aligned}$$

Die Gleichung (3.6) ist für diese spezielle Funktion f also bewiesen.

Nun sei $f : [0, 2\pi] \rightarrow \mathbb{R}$ eine Treppenfunktion. Dann gibt es endlich viele Funktionen f_1, \dots, f_r der oben betrachteten Art und reelle Zahlen $\alpha_1, \dots, \alpha_r$ mit $f = \sum_{j=1}^r \alpha_j f_j$. Allgemein sei mit $S_n[g]$ die n -te Partialsumme der Fourierreihe von g bezeichnet, also

$$S_n[g](x) := \sum_{k=-n}^n c_k[g] e^{ikx}, \quad c_k[g] := \frac{1}{2\pi} \int_0^{2\pi} g(x) e^{-ikx} dx.$$

Dann gilt

$$\begin{aligned} \|f - S_n[f]\|_2 &= \left\| \sum \alpha_j f_j - \sum \alpha_j S_n[f_j] \right\|_2 \\ &\leq \sum |\alpha_j| \|f_j - S_n[f_j]\|_2 \rightarrow 0 \quad \text{für } n \rightarrow \infty \end{aligned}$$

nach dem bereits Bewiesenen und Satz 3.3.

Schließlich sei $f : [0, 2\pi] \rightarrow \mathbb{R}$ eine Regelfunktion. Zu $\varepsilon \in \mathbb{R}^+$ existiert eine Treppenfunktion t auf $[0, 2\pi]$ mit $|f(x) - t(x)| \leq \varepsilon/2$ für $x \in [0, 2\pi]$. Für $g := f - t$ folgt

$$\|g - S_n[g]\|_2^2 = \|g\|_2^2 - \sum_{k=-n}^n |c_k[g]|^2 \leq \|g\|_2^2 \leq \left(\frac{\varepsilon}{2}\right)^2.$$

Nach dem bereits Bewiesenen existiert ein $n_0 \in \mathbb{N}$ mit $\|t - S_n[t]\|_2 < \varepsilon/2$ für $n \geq n_0$. Für $n \geq n_0$ gilt also

$$\|f - S_n[f]\|_2 \leq \|g - S_n[g]\|_2 + \|t - S_n[t]\|_2 < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Somit ist

$$\lim_{n \rightarrow \infty} \|f - S_n[f]\|_2 = 0.$$

Die Ausdehnung auf $f \in V$ ist jetzt trivial. ■

Gegenüber der Konvergenz im quadratischen Mittel ist die Beurteilung der punktweisen Konvergenz einer Fourierreihe wesentlich schwieriger. Selbst für stetiges f braucht die Folge $(S_n[f])_{n \in \mathbb{N}}$ der Partialsummen der Fourierreihe von f nicht punktweise gegen f zu konvergieren. Erstaunlicherweise gilt aber, dass die arithmetischen Mittel dieser Partialsummen sogar gleichmäßig gegen f konvergieren, wie der folgende Satz zeigt.

3.7 Satz (Fejér). *Ist $f : \mathbb{R} \rightarrow \mathbb{R}$ 2π -periodisch und stetig und wird*

$$\sigma_n[f] := \frac{1}{n} (S_0[f] + \dots + S_{n-1}[f])$$

gesetzt, so konvergiert die Folge $(\sigma_n[f])_{n \in \mathbb{N}}$ gleichmäßig gegen f .

Beweis. Es ist

$$\begin{aligned}
 S_n[f](x) &= \sum_{k=-n}^n \left(\frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt \right) e^{ikx} \\
 &= \frac{1}{2\pi} \int_0^{2\pi} f(t) \sum_{k=-n}^n e^{ik(x-t)} dt \\
 &= \frac{1}{2\pi} \int_0^{2\pi} f(t) D_n(x-t) dt \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_n(t) dt \quad [\text{Substitution und Periodizität}]
 \end{aligned}$$

mit

$$D_n(x) := \sum_{k=-n}^n e^{ikx}.$$

Wir bemerken

$$\frac{1}{2\pi} \int_0^{2\pi} D_n(x) dx = 1$$

und

$$(e^{ix} - 1)D_n(x) = e^{i(n+1)x} - e^{-inx}.$$

Multiplikation mit $e^{-\frac{ix}{2}}$ ergibt

$$D_n(x) = \frac{\sin(n + \frac{1}{2})x}{\sin \frac{x}{2}}.$$

Setze

$$F_n := \frac{1}{n}(D_0 + \cdots + D_{n-1}),$$

dann ist

$$\frac{1}{2\pi} \int_0^{2\pi} F_n(x) dx = 1$$

und

$$F_n(x) = \frac{1}{n} \sum_{k=0}^{n-1} \frac{\sin(k + \frac{1}{2})x}{\sin \frac{x}{2}}.$$

Aus

$$\begin{aligned} \sum_{k=0}^{n-1} e^{i(k+\frac{1}{2})x} &= e^{\frac{ix}{2}} \sum_{k=0}^{n-1} e^{ikx} = e^{\frac{ix}{2}} \frac{e^{inx} - 1}{e^{ix} - 1} \\ &= \frac{e^{inx} - 1}{e^{\frac{ix}{2}} - e^{-\frac{ix}{2}}} = \frac{\cos nx - 1 + i \sin nx}{2i \sin \frac{x}{2}} \end{aligned}$$

folgt

$$\sum_{k=0}^{n-1} \sin \left(k + \frac{1}{2} \right) x = \frac{1 - \cos nx}{2 \sin \frac{x}{2}} = \frac{\sin^2 \frac{nx}{2}}{\sin \frac{x}{2}},$$

also

$$F_n(x) = \frac{1}{n} \left(\frac{\sin \frac{nx}{2}}{\sin \frac{x}{2}} \right)^2.$$

Nun ist

$$\begin{aligned} \sigma_n[f](x) &= \sum_{k=0}^{n-1} \frac{1}{n} S_k[f](x) = \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_k(t) dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) F_n(t) dt, \end{aligned}$$

also

$$\begin{aligned} &|f(x) - \sigma_n[f](x)| \\ &= \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} F_n(t) dt \cdot f(x) - \frac{1}{2\pi} \int_0^{2\pi} f(x-t) F_n(t) dt \right| \\ &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x-t) - f(x)| F_n(t) dt. \end{aligned}$$

Zu gegebenem $\varepsilon > 0$ existiert ein $\delta > 0$ mit $\delta < \pi$ und

$$|f(x-t) - f(x)| \leq \varepsilon \quad \text{für } |t| < \delta,$$

da f nach Satz 3.5 aus Kapitel 4 gleichmäßig stetig ist; ferner ist

$$|f(x-t) - f(x)| \leq M \quad \text{für alle } x \text{ und } t$$

mit passendem M . Es folgt

$$\begin{aligned} & \int_{-\pi}^{\pi} |f(x-t) - f(x)| F_n(t) dt \\ &= \int_{|t| \leq \delta} |f(x-t) - f(x)| F_n(t) dt + \int_{|t| \geq \delta} |f(x-t) - f(x)| F_n(t) dt \\ &\leq \varepsilon + M \int_{|t| \geq \delta} \frac{1}{n} \left[\frac{\sin \frac{nt}{2}}{\sin \frac{t}{2}} \right]^2 dt \leq \varepsilon + \frac{M}{n} \frac{2\pi}{(\sin \frac{\delta}{2})^2}. \end{aligned}$$

Für $n \geq n_0$ mit passendem n_0 ist der letzte Summand kleiner als ε , also ist $|f(x) - \sigma_n[f](x)| < \varepsilon$ für $n \geq n_0$. ■

Wir wollen an den Satz von Fejér noch zwei Bemerkungen anfügen.

Zunächst soll noch einmal kurz das Wesen der dort verwendeten Summierung herausgestellt werden. Für eine Reihe

$$\sum_{k=0}^{\infty} a_k$$

bildet man die Mittel der Partialsummen, also

$$\begin{aligned} \sigma_n &= \frac{s_0 + s_1 + \cdots + s_n}{n+1} \\ &= \frac{a_0 + (a_0 + a_1) + (a_0 + a_1 + a_2) + \cdots + (a_0 + a_1 + \cdots + a_n)}{n+1} \\ &= a_0 + \frac{n}{n+1} a_1 + \frac{n-1}{n+1} a_2 + \cdots + \frac{1}{n+1} a_n \\ &= \sum_{k=0}^n \left(1 - \frac{k}{n+1} \right) a_k. \end{aligned}$$

Gegenüber der Bildung der Partialsumme $\sum_{k=0}^n a_k$ werden hierbei also die Summanden a_k mit Gewichten versehen, die von n abhängen und mit k abnehmen.

Sodann wollen wir darauf hinweisen, dass man aus dem Satz von Fejér leicht einen wichtigen Satz über die gleichmäßige Approximierbarkeit stetiger reeller Funktionen durch Polynome herleiten kann.

3.8 Satz (Approximationssatz von Weierstraß). *Sei $g : [a, b] \rightarrow \mathbb{R}$ stetig und $\varepsilon > 0$. Dann existiert ein Polynom P mit*

$$|g(x) - P(x)| \leq \varepsilon \quad \text{für alle } x \in [a, b].$$

Beweis. O.B.d.A. sei $a = -1$ und $b = 1$. Setze

$$f(\varphi) := g(\cos \varphi) \quad \text{für } -\pi \leq \varphi \leq \pi$$

und ergänze f zu einer 2π -periodischen Funktion. Wegen $f(-\pi) = f(\pi)$ ist f stetig. Nach dem Satz von Fejér existiert ein $n \in \mathbb{N}$ mit

$$|f(\varphi) - \sigma_n[f](\varphi)| \leq \varepsilon \quad \text{für alle } \varphi \in \mathbb{R}.$$

Die reelle Fourierreihe von f ist von der Form

$$f(\varphi) \sim \frac{a_0}{2} + \sum_{k=0}^{\infty} a_k \cos k\varphi,$$

denn wegen $f(\varphi) = f(-\varphi)$ verschwinden alle b_k . Da $\cos k\varphi$ sich als Polynom in $\cos \varphi$ ausdrücken läßt, ist

$$\sigma_n[f](\varphi) = P(\cos \varphi)$$

mit einem Polynom P . Es ist also

$$|g(\cos \varphi) - P(\cos \varphi)| \leq \varepsilon \quad \text{für alle } \varphi$$

und daher

$$|g(x) - P(x)| \leq \varepsilon \quad \text{für alle } x \in [-1, 1]. \quad \blacksquare$$