

**Analysis**  
**Skriptum zur Vorlesung im Wintersemester 2015 &**  
**Sommersemester 2016**  
**Version vom 30. August 2016**

Bernhard Lamel



# Der Körper der reellen Zahlen

## 1. Das Axiomensystem

Wir führen den Körper der reellen Zahlen *axiomatisch* ein, das heisst, wir postulieren, dass es eine Menge  $\mathbb{R}$  (die Menge der *reellen Zahlen*) gibt, deren Elemente wir *reelle Zahlen* nennen, welche mit zwei binären Operationen ( $+$  und  $\cdot$ ) und einer Totalordnung  $\leq$  versehen ist, sodass die folgenden Eigenschaften erfüllt sind:

Axiom 1: (Das *Körperaxiom*)  $(\mathbb{R}, +, \cdot)$  ist ein Körper;

Axiom 2: (Das *Ordnungsaxiom*)  $(\mathbb{R}, +, \cdot, \leq)$  ist ein angeordneter Körper, d.h.

$$a + c \leq b + c, \text{ falls } a \leq b, \quad c \in \mathbb{R}$$

und

$$ac \leq bc, \text{ falls } a \leq b, \quad c \geq 0;$$

Axiom 3: (Das *Archimedische Axiom*) Für jedes  $M \in \mathbb{R}$  existiert ein  $N \in \mathbb{N}$  sodass  $M \leq N \cdot 1$ ;

Axiom 4: (Das *Vollständigkeitsaxiom*) Wenn  $S \subset \mathbb{R}$ ,  $S \neq \emptyset$ , nach oben beschränkt ist, so gibt es eine kleinste obere Schranke von  $S$  in  $\mathbb{R}$ .

Von diesen Axiomen ausgehend werden wir die Analysis mit reellen Zahlen entwickeln. Sie geben uns die Regeln, nach denen wir mit diesen Objekten hantieren werden, unabhängig von der Vorstellung, die wir von den reellen Zahlen haben mögen (zum Beispiel als Punkte des "Zahlenstrahls"). Wir werden nun damit beginnen, einige einfache Folgerungen zu ziehen, und die Axiome inhaltlich zu erläutern.

**1.1. Die Körperaxiome.** Ganz allgemein ist ein Körper ein Tripel  $(k, +, \cdot)$  gebildet aus einer Menge  $k$  und zwei binären Operationen  $+$  und  $\cdot$ , welche die folgenden Eigenschaften hat:

Körperaxiom 1:  $(k, +)$  ist eine abelsche Gruppe (mit einem neutralen Element, welches wir im folgenden mit  $0$  bezeichnen);

Körperaxiom 2:  $(k \setminus \{0\}, \cdot)$  ist eine abelsche Gruppe;

Körperaxiom 3:  $a(b + c) = ab + ac$  für alle  $a, b, c \in k$ .

Wir benutzen die üblichen abkürzenden Schreibweisen wie  $a + (-b) = a - b$  und  $ab^{-1} = \frac{a}{b}$ ; das neutrale Element bezüglich der Addition wird mit  $0$ , jenes bezüglich der Multiplikation mit  $1$  bezeichnet.

In einem Körper darf man rechnen, wie "man es gewohnt ist", und Faktoren, welche nicht  $0$  sind, kürzen, d.h. aus  $ab = ac$  mit  $a \neq 0$  folgt  $b = c$ .

**1.2. Die Ordnungsaxiome.** Wir haben gefordert, dass  $\leq$  eine Totalordnung auf  $\mathbb{R}$  ist, also:

i)  $\leq$  ist reflexiv, d.h.  $a \leq a$  für alle  $a$ ;

ii)  $\leq$  ist antisymmetrisch, d.h. wenn  $a \leq b$  und  $b \leq a$ , so ist  $a = b$ ;

iii)  $\leq$  ist transitiv, d.h. wenn  $a \leq b$  und  $b \leq c$ , so ist auch  $a \leq c$ ;

iv) Für zwei gegebene  $a, b$  gilt entweder  $a \leq b$  oder  $b \leq a$ .

Allgemein heissen Relationen, welche i)-iii) erfüllen, *Ordnungsrelationen*, und eine Totalordnung ist eine Ordnungsrelation, welche zusätzlich iv) erfüllt. Wie üblich schreiben wir  $a < b$ , wenn  $a \leq b$  aber  $a \neq b$  ist.

$M \in \mathbb{R}$  heisst *obere Schranke* von  $S \subset \mathbb{R}$ , wenn  $s \leq M$  für alle  $s \in S$  gilt;  $M$  heisst *untere Schranke* von  $S \subset \mathbb{R}$ , wenn  $s \geq M$  für alle  $s \in S$  gilt. Wir sagen,  $S \subset \mathbb{R}$  sei nach oben (bzw. nach unten beschränkt), wenn  $S$  eine obere (bzw. untere) Schranke besitzt.  $S \subset \mathbb{R}$  heisst beschränkt, wenn  $S$  sowohl nach oben als auch nach unten beschränkt ist.

Punkt 2 verbindet die Ordnungseigenschaften mit den Körperoperationen. Wir wollen einige einfache Folgerungen aus diesem Axiom ziehen, welche wir im folgenden ohne weiteren Kommentar anwenden werden.

LEMMA 1. Für jedes  $x \in \mathbb{R}$  ist  $x^2 \geq 0$ , insbesondere gilt  $0 < 1$ . Sind  $a, b$  in  $\mathbb{R}$  und  $a \leq b$ , so ist  $-b \leq -a$ , und allgemeiner gilt für  $c \leq 0$ , dass  $bc \leq ac$ . Sind für  $a, b, c, d \in \mathbb{R}$  sowohl  $a \leq b$  als auch  $c \leq d$ , so ist  $a + c \leq b + d$ ; falls sogar  $a < b$  ist, so ist  $a + c < b + d$ . Ist  $0 < a < b$  und  $0 < c < d$ , so ist  $ac < bd$ .

BEWEIS. Wir verwenden das Ordnungsaxiom für die Multiplikation mit  $a = 0$ ,  $b = x$ , und  $c = x$ , um zu sehen dass  $0 \leq x^2$  gilt. Da  $1 = 1^2$  und  $0 \neq 1$  gilt, folgt auch  $0 < 1$ .

Gilt  $a \leq b$ , so folgt aus dem Ordnungsaxiom für die Addition (mit  $c = -a - b$ ), dass  $-b \leq -a$  gilt. Insbesondere ist für  $c \leq 0$  die Ungleichung  $-c \geq 0$  erfüllt, womit  $-ac = a(-c) \leq b(-c) = -bc$  und damit aus dem soeben Bewiesenen  $bc \leq ac$  folgt.

Die letzten Behauptungen folgen aus dem Ordnungsaxiom für die Addition bzw. der Multiplikation sowie der Transitivität von  $\leq$ , da

$$a + c \leq b + c \leq b + d$$

gilt. Ist nun zusätzlich  $a < b$ , so ist  $a + c < b + c$  und somit die erste Ungleichung in der Kette strikt, also  $a + c < b + d$ . Ist  $0 < a < b$  und  $0 < c < d$ , so ist  $ac < bc$  und  $cb < db$ , also auch  $ac < bd$ .  $\square$

Für die Eigenschaften von Grössen von Produkten ist demnach vor allem das Verhalten auf *positiven* Zahlen ausschlaggebend. Wir schreiben

$$\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}.$$

LEMMA 2. Die natürlichen Zahlen  $\mathbb{N} = \{0, 1, 2, \dots\}$  sind in  $\mathbb{R}$  eingebettet, wenn man  $n \in \mathbb{N}$  mit  $n \cdot 1$  identifiziert.

BEWEIS. Wir schreiben  $\iota(n) = n \cdot 1$ . Aus den Körperaxiomen folgt, dass  $\iota(m + n) = \iota(m) + \iota(n)$  und  $\iota(mn) = \iota(m)\iota(n)$  für  $m, n \in \mathbb{N}$  gilt.

Um den Beweis abzuschliessen, müssen wir zeigen, dass  $\iota$  injektiv ist. Dazu beachten wir, dass  $\iota$  strikt monoton ist: Sind  $m, n \in \mathbb{N}$  mit  $m < n$ , so gilt  $\iota(m) < \iota(n)$ . Wir beweisen dies mit Hilfe von Induktion nach  $p = m + n$ : Für  $p = 1$  gibt es nur das Paar  $m = 0$ ,  $n = 1$ , und nach Lemma 1 ist  $0 < 1$ . Wir nehmen an, dass die Behauptung für alle  $m < n$  mit  $m + n \leq p$  erfüllt ist. Sei nun  $m_0 < n_0$  und  $m_0 + n_0 = p + 1$ . Dann gilt entweder  $m_0 = n_0 - 1$ , womit  $m_0 < n_0$  wegen  $0 < 1$  folgt, oder es ist  $m_0 < n_0 - 1$ , und wir können die Induktionsvoraussetzung auf  $m = m_0$ ,  $n = n_0 - 1$  anwenden, was  $\iota(m_0) < \iota(n_0 - 1)$  liefert. Da aber wiederum  $0 < 1$  gilt, ist  $\iota(m_0) < \iota(n_0)$ .  $\square$

Der Beweis scheint vielleicht unnötig kompliziert, doch die offensichtliche Aussage, dass in einem geordneten Körper  $m \cdot 1 < n \cdot 1$  für  $m < n$  ist, benötigt einen Induktionsbeweis!

Analog zum vorangehenden Lemma 2 lässt sich zeigen, dass die Abbildungen

$$\mathbb{Z} \rightarrow \mathbb{R}, \quad m \mapsto m \cdot 1, \quad \text{sowie} \quad \mathbb{Q} \rightarrow \mathbb{R}, \quad \frac{p}{q} \mapsto \frac{p \cdot 1}{q \cdot 1},$$

zu Identifikationen von  $\mathbb{Z}$  mit einem Teilring von  $\mathbb{R}$  bzw.  $\mathbb{Q}$  mit einem Teilkörper von  $\mathbb{R}$  führen. Wir werden deswegen im folgenden  $\mathbb{Z} \subset \mathbb{R}$  und  $\mathbb{Q} \subset \mathbb{R}$  schreiben.

**Übungsaufgabe 1.** Zeige die vorangehende Aussage für  $\mathbb{Q}$  und erläutere, wie die Aussage  $\mathbb{Q} \subset \mathbb{R}$  zu verstehen ist. Ist diese Identifikation von  $\mathbb{Q}$  mit einem Teilkörper von  $\mathbb{R}$  eindeutig?

Die Ordnung erlaubt es uns, *Intervalle* zu definieren: Wir setzen

$$\begin{aligned} (a, b) &:= \{x \in \mathbb{R} : a < x < b\}, \\ (a, b] &:= \{x \in \mathbb{R} : a < x \leq b\}, \\ [a, b) &:= \{x \in \mathbb{R} : a \leq x < b\}, \\ [a, b] &:= \{x \in \mathbb{R} : a \leq x \leq b\}. \end{aligned}$$

Es ist bequem, auch Intervalle von der Form

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

zu definieren.

Eine wichtige Methode, auf die Gültigkeit einer Ungleichung zu schliessen, ist, eine abgeschwächte Form dieser Ungleichung zu beweisen: Ist  $x \leq y$ , so ist für beliebiges  $\varepsilon > 0$  auch die Ungleichung  $x \leq y + \varepsilon$  erfüllt. Wenn diese Ungleichung nun für alle  $\varepsilon > 0$  erfüllt ist, können wir auch auf die Gültigkeit der ursprünglichen Ungleichung schliessen: So ist andererseits, wenn  $y < x$ , für  $\varepsilon_0 = \frac{x-y}{2}$  die Ungleichung  $x - y > \varepsilon_0$ , also  $x > y + \varepsilon_0$  erfüllt.

**1.3. Das Archimedische Axiom.** Das Archimedische Axiom kann auch etwas anders formuliert werden. Wenn  $k$  ein angeordneter Körper ist (d.h. Axiome 1 und 2 erfüllt), was wir für den Rest von Unterabschnitt 1.3 annehmen wollen, so zeigt der Beweis von Lemma 2, dass  $\mathbb{N} \subset k$  aufgefasst werden kann.

LEMMA 3.  $k$  erfüllt Axiom 3 genau dann, wenn  $\mathbb{N} \subset k$  unbeschränkt ist.

BEWEIS. Wenn Punkt 3 erfüllt ist, kann kein  $M \in k$  eine obere Schranke von  $\mathbb{N}$  sein. Wenn andererseits Axiom 3 nicht erfüllt ist, so gibt es ein  $M \in k$  sodass  $N \leq M$  für alle  $N \in \mathbb{N}$  gilt, also ist  $\mathbb{N}$  beschränkt.  $\square$

Die folgende Eigenschaft ist scheinbar stärker, aber äquivalent zum Archimedischen Axiom.

LEMMA 4.  $k$  erfüllt Axiom 3 genau dann, wenn für jedes  $q \in \mathbb{N}$ ,  $q \neq 0$  und jedes  $M \in k$  ein  $p \in \mathbb{N}$  existiert, sodass  $M \leq \frac{p}{q}$  gilt.

BEWEIS. Wir müssen nur zeigen, dass das Archimedische Axiom die stärkere Eigenschaft impliziert. Dazu wenden wir es auf die Zahl  $qM \in k$  an und erhalten ein  $p$  mit  $qM \leq p$ , also  $M \leq \frac{p}{q}$ .  $\square$

Das Archimedische Axiom ist so formuliert, dass man über natürliche Zahlen als beliebig groß denken kann. Das bedeutet, dass man über die Kehrwerte natürlicher Zahlen als beliebig klein denken kann.

LEMMA 5.  $k$  erfüllt Axiom 3 genau dann, wenn zu jedem  $\varepsilon > 0$ ,  $\varepsilon \in k$  ein  $N \in \mathbb{N}$  existiert welches  $\frac{1}{N} < \varepsilon$  erfüllt.

BEWEIS. Sei zunächst Axiom 3 erfüllt. Dann können wir für  $M = \varepsilon^{-1}$  ein  $N \in \mathbb{N}$  finden, welches  $\varepsilon^{-1} = M < N$  erfüllt. Durch Multiplikation mit  $\frac{\varepsilon}{N}$  erhalten wir also  $\frac{1}{N} < \varepsilon$ .

Sei nun die Bedingung des Lemmas erfüllt. Wenn  $M \in k$  mit  $M \leq 0$  gegeben ist, so ist die Folgerung des Archimedischen Axioms mit  $N = 0$  erfüllt. Ist andererseits  $M > 0$ , so erhalten wir für  $\varepsilon = M^{-1}$  ein  $N \in \mathbb{N}$  mit  $\frac{1}{N} < \varepsilon = M^{-1}$ , und es folgt  $M < N$ .  $\square$

LEMMA 6. Seien  $A, B \in \mathbb{R}$  mit  $A \leq B$  gegeben. Wenn  $B - A \leq \frac{1}{N}$  für alle  $N \in \mathbb{N}$  (mit  $N \neq 0$ ) gilt, so ist  $A = B$ .

**Übungsaufgabe 2.** Beweise Lemma 6.

LEMMA 7. Seien  $a, b \in \mathbb{R}$  mit  $a < b$  gegeben. Dann gibt es ein  $r \in \mathbb{Q}$  mit  $r \in (a, b)$ .

BEWEIS. Wir nehmen ohne Beschränkung der Allgemeinheit an, dass  $0 < a < b$  ist; andernfalls ersetzen wir  $a$  und  $b$  durch  $a + N$  bzw.  $b + N$  für ein genügend grosses  $N \in \mathbb{N}$ . Sei  $q \in \mathbb{N}$  so gross, dass  $\frac{1}{q} < b - a$  (ein solches  $q$  existiert nach Lemma 5). Wiederum nach Lemma 5 ist die Menge  $K = \{n \in \mathbb{N} : \frac{n}{q} < b\}$  nach oben beschränkt; sei  $p = \max K$ . Wir behaupten, dass  $\frac{p}{q} \in (a, b)$ : Da  $\frac{p}{q} < b$ , bleibt nur zu zeigen, dass  $a < \frac{p}{q}$ . Wenn  $\frac{p}{q} \leq a$  wäre, so wäre  $\frac{p+1}{q} \leq a + \frac{1}{q} < b$ , also  $p+1 \in K$ , im Widerspruch zu  $p = \max K$ .  $\square$

LEMMA 8. Sei  $k \in \mathbb{N}$ ,  $k \geq 2$ . Für jedes  $\varepsilon > 0$  gibt es ein  $n \in \mathbb{N}$ , sodass  $k^{-n} < \varepsilon$ .

BEWEIS. Wir zeigen zunächst mittels Induktion, dass  $k^n \geq nk$  für beliebiges  $n \in \mathbb{N}$  gilt: Für  $n = 1$  ist dies klar; wenn die Ungleichung für  $n = m - 1$  gilt, so folgt

$$k^m = k^{m-1}k \geq (m-1)k^2 \geq (m-1) \left(1 + \frac{1}{m-1}\right)k = (m-1) \frac{m}{m-1}k = mk.$$

Nach Lemma 5 gibt es ein  $n$ , sodass  $\frac{1}{n} \leq k\varepsilon$ . Es folgt

$$\frac{1}{k^n} \leq \frac{1}{nk} \leq \frac{1}{k}k\varepsilon = \varepsilon$$

wie behauptet.  $\square$

**1.4. Das Vollständigkeitsaxiom.** Axiom 4 ist, was die reellen Zahlen tatsächlich auszeichnet. So ist zum Beispiel  $\mathbb{Q}$  ein Körper, welcher auch Axiom 3 erfüllt. Wir werden bald sehen, dass  $\mathbb{Q}$  Axiom 4 *nicht* erfüllt, wollen aber zunächst einige hilfreiche Umformulierungen von Axiom 4 kennenlernen. Wir nehmen im folgenden an, dass  $k$  ein archimedischer Körper ist, also Axiome 1, 2, und 3 erfüllt.

Wir rufen in Erinnerung, dass eine kleinste obere Schranke einer Menge  $S$  eine obere Schranke  $A$  mit der Eigenschaft dass für jede obere Schranke  $M$  die Ungleichung  $A \leq M$  gilt. Eine kleinste obere Schranke ist notwendigerweise eindeutig bestimmt (wenn sie existiert) und wir schreiben  $A = \sup S$  und sagen,  $A$  sei das *Supremum* von  $S$ . Eine grösste untere Schranke einer Menge  $S$  ist eine untere Schranke  $B$  von  $S$  mit der Eigenschaft, dass für jede untere Schranke  $M$  von  $S$  die Ungleichung  $M \leq B$  erfüllt ist; auch die grösste untere Schranke (so sie existiert) ist eindeutig bestimmt, und wir schreiben  $B = \inf S$  und sagen,  $B$  sei das *Infimum* von  $S$ . Ob man die Existenz kleinster oberer oder grösster unterer Schranken fordert, ist irrelevant:

LEMMA 9.  $k$  erfüllt Axiom 4 genau dann, wenn jede nach unten beschränkte nichtleere Menge eine grösste untere Schranke hat.

BEWEIS. Sei zunächst Axiom 4 erfüllt, und  $S \subset k$  nach unten beschränkt, durch  $-M \leq s$  für alle  $s \in S$ . Dann ist  $M$  eine obere Schranke der Menge  $-S = \{-s : s \in S\}$ , und es gibt eine kleinste obere Schranke  $A$  von  $-S$ . Wir behaupten, dass  $-A$  die grösste untere Schranke von  $S$  ist:  $-A$  ist eine untere Schranke von  $S$ , da  $-s \leq A$  für alle  $s \in S$ , also  $s \geq -A$  für alle  $s \in S$  gilt. Ist nun  $-M$  eine beliebige untere Schranke von  $S$ , so ist nach dem obigen  $M$  eine obere Schranke von  $-S$ , also  $A \leq M$ . Es folgt  $-M \leq -A$ , also ist  $-A$  die grösste untere Schranke von  $S$ .

Die Gegenrichtung wird ganz entsprechend bewiesen. □

**Übungsaufgabe 3.** Beweise die Gegenrichtung in Lemma 9.

LEMMA 10. Sei  $S \subset \mathbb{R}$ . Dann ist  $A = \sup S$  genau dann, wenn  $A$  eine obere Schranke von  $S$  mit der Eigenschaft ist, dass für jedes  $\varepsilon > 0$  ein  $s = s(\varepsilon) \in S$  existiert sodass  $A - \varepsilon < s \leq A$  gilt. Weiters ist  $B = \inf S$  genau dann, wenn  $B$  eine untere Schranke von  $S$  mit der Eigenschaft ist, dass für jedes  $\varepsilon > 0$  ein  $s = s(\varepsilon) \in S$  existiert sodass  $B \leq s < B + \varepsilon$  gilt.

BEWEIS. Wenn  $A$  das Supremum von  $S$  ist, so ist für gegebenes  $\varepsilon > 0$  die Zahl  $A - \varepsilon < A$  keine obere Schranke von  $S$ . Also muss es ein  $s \in S$  geben, für welches  $A - \varepsilon < s$  gilt. □

**Übungsaufgabe 4.** Beweise die Aussage über das Infimum in Lemma 10.

In der Formulierung des folgenden Lemmas verwenden wir die Notation, welche im Beweis von Lemma 9 eingeführt wurde:  $-S = \{-s : s \in S\}$ . Allgemein schreiben wir für  $S, T \subset \mathbb{R}$  und  $c \in \mathbb{R}$

$$S + T = \{s + t : s \in S, t \in T\}, \quad cS = \{cs : s \in S\}.$$

LEMMA 11. Sei  $S \subset \mathbb{R}$ . Dann ist  $S$  nach oben beschränkt genau dann, wenn  $-S$  nach unten beschränkt ist, und es gilt  $\sup S = -\inf(-S)$ . Weiters ist  $S$  nach unten beschränkt genau dann, wenn  $-S$  nach oben beschränkt ist, und es gilt  $\inf S = -\sup(-S)$ .

BEWEIS. Es genügt, die erste Aussage zu beweisen, da die zweite aus der ersten angewendet auf die Menge  $-S$  folgt. Die erste Aussage wurde schon im Beweis von Lemma 9 bewiesen. □

LEMMA 12. Wenn  $S$  und  $T$  nach oben beschränkt sind, und  $c > 0$ , so sind auch  $S + T$  und  $cS$  nach oben beschränkt, und es gilt

$$\sup(S + T) = \sup(S) + \sup(T), \quad \sup(cS) = c \sup S.$$

BEWEIS. Die Summe oberer Schranken von  $S$  und von  $T$  liefert eine obere Schranke von  $S + T$ ; die Beschränktheit von  $cS$  wird entsprechend nachgewiesen.

Wir setzen nun  $A = \sup S + \sup T$ , was eine obere Schranke von  $S + T$  ist, und müssen zeigen, dass  $A = \sup(S + T)$  ist. Sei dazu  $\varepsilon > 0$  gegeben. Dann gibt es nach Lemma 10 ein  $s \in S$  und ein  $t \in T$  mit

$$\begin{aligned} \sup S - \frac{\varepsilon}{2} < s &\leq \sup S \\ \sup T - \frac{\varepsilon}{2} < t &\leq \sup T, \end{aligned}$$

nach Addition der Ungleichungen also  $A - \varepsilon < s + t \leq A$ . Da  $\varepsilon > 0$  beliebig war, ist die Bedingung in Lemma 10 erfüllt, und es folgt  $A = \sup(S + T)$ .

Ähnlich setzen wir zum Beweis der zweiten Gleichung  $A = c \sup S$  und wenden wieder Lemma 10 an, überlassen dies dem Leser aber als Übungsaufgabe.  $\square$

**Übungsaufgabe 5.** Beweise die zweite Aussage in Lemma 10, formuliere entsprechende Aussagen für die Infima, und skizziere, wie man diese beweisen kann.

Das Supremum hat auch ein Monotonieverhalten in Bezug auf die Ordnungsrelation der mengentheoretischen Inklusion.

LEMMA 13. Sind  $S \subset \mathbb{R}$  und  $T \subset \mathbb{R}$ ,  $S \neq \emptyset$  nach oben beschränkt und gilt  $S \subset T$ , so ist  $\sup S \leq \sup T$ .

BEWEIS. Die Ungleichung folgt, da jede obere Schranke von  $T$  auch eine obere Schranke von  $S$  ist.  $\sup T$  ist eine solche obere Schranke, also gilt, da  $\sup S$  die kleinste obere Schranke von  $S$  ist, auch  $\sup S \leq \sup T$ .  $\square$

**Übungsaufgabe 6.** Gilt in Lemma 13 auch die strikte Monotonie, i.e. impliziert  $S \subsetneq T$  die strikte Ungleichung  $\sup S < \sup T$ ?

Wir werden nun zeigen, dass in  $\mathbb{R}$  Wurzeln positiver Zahlen existieren. Wir erinnern daran, dass

$$\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}.$$

LEMMA 14. Sei  $y \in \mathbb{R}_+$ . Dann gibt es ein  $x \in \mathbb{R}_+$  mit  $x^2 = y$ , und wir können durch  $f(y) = \sqrt{y} = x$  eine Funktion  $\mathbb{R}_+ \rightarrow \mathbb{R}_+$  definieren, welche strikt monoton wachsend ist.

BEWEIS. Wir betrachten die Menge  $S_y = \{t \in \mathbb{R}_+ \cup \{0\} : 0 \leq t^2 < y\}$ . Diese Menge ist nach oben beschränkt, da für  $t > 1$  die Ungleichung  $t < t^2$  gilt, also auf jeden Fall  $t \leq \max(1, y)$  für jedes  $t \in S$  gilt. Daher gibt es  $x := \sup S$ , und wir behaupten, dass  $x^2 = y$  gilt. Zunächst bemerken wir, dass  $S_y = [0, x)$  gilt; wäre  $x \in S_y$ , also  $x^2 < y$ , so ist für  $\delta < \min(2x, \frac{y-x^2}{4x})$  die Ungleichung

$$(x + \delta)^2 = x^2 + 2x\delta + \delta^2 < x^2 + 4x\delta < x^2 + 4x \frac{y-x^2}{4x} = y$$

erfüllt, also  $x$  keine obere Schranke von  $S_y$ .

Somit gilt  $x^2 \geq y$ . Wäre nun  $x^2 > y$ , so wäre

$$a^2 = \left(x - \frac{x^2 - y}{2x}\right)^2 = x^2 - 2x \frac{x^2 - y}{2x} + \underbrace{\left(\frac{x^2 - y}{2x}\right)^2}_{>0} > y,$$

also  $a < x$ , aber  $a^2 > y$ , was einen Widerspruch zu  $a \in S_y$  darstellt.

Die Funktion  $f(y)$  ist damit durch

$$f(y) = \sup S_y, \quad y > 0$$

definiert, was wegen Lemma 13 die Monotonie liefert; die strikte Monotonie folgt dann wegen  $x^2 < x'^2$ , wenn  $0 < x < x'$ .  $\square$

Es ist sinnvoll,  $\sqrt{0} = 0$  zu definieren. Wir bemerken, dass die Folgerung über die strikte Monotonie der Wurzelfunktion nun auf  $\mathbb{R}_+ \cup \{0\}$  gilt. Wegen der strikten Monotonie ist insbesondere die Lösung  $x$  der Gleichung  $x^2 = y$ , welche  $x \geq 0$  erfüllt, eindeutig bestimmt. Eine weitere übliche Schreibweise ist  $\sqrt{x} = x^{\frac{1}{2}}$ .

LEMMA 15. Für nichtnegative reelle Zahlen  $y_1, y_2 \in \mathbb{R}_+ \cup \{0\}$  gilt  $\sqrt{y_1 y_2} = \sqrt{y_1} \sqrt{y_2}$ .

BEWEIS. Wir haben nur zu überprüfen, dass die rechte Seite der Gleichung—also die Zahl  $\sqrt{y_1} \sqrt{y_2}$ —die definierende Eigenschaft der linken Seite erfüllt. Zunächst ist  $\sqrt{y_1} \sqrt{y_2} \geq 0$ , und es gilt

$$(\sqrt{y_1} \sqrt{y_2})^2 = (\sqrt{y_1})^2 (\sqrt{y_2})^2 = y_1 y_2$$

wie gewünscht.  $\square$

**1.5. Potenzen mit rationalen Exponenten.** Wir haben in Unterabschnitt 1.4 die Existenz von Quadratwurzeln gezeigt. Wir können die dort vorgestellte Methode verallgemeinern, um Potenzen mit rationalen Exponenten zu definieren. Die entscheidende Ungleichung wollen wir zunächst extra formulieren.

LEMMA 16. *Seien  $x, \delta \in \mathbb{R}$  positive Zahlen. Dann gilt für  $q \in \mathbb{Z}$  mit  $q \geq 2$  die Ungleichung*

$$(x - \delta)^q > x^q - q\delta x^{q-1}.$$

BEWEIS. Die Aussage für  $q = 2$  wurde bereits im Beweis von Lemma 14 besprochen. Wir nehmen an, das Lemma sei für ein beliebiges, aber fixes  $q \in \mathbb{N}$  bereits bewiesen. Dann gilt

$$\begin{aligned} (x - \delta)^{q+1} &= (x - \delta)^q(x - \delta) \\ &\geq (x^q - q\delta x^{q-1})(x - \delta) \\ &= x^{q+1} - (q + 1)\delta x^q + \underbrace{q\delta^2 x^{q-1}}_{>0} \\ &> x^{q+1} - (q + 1)\delta x^q, \end{aligned}$$

womit das Lemma für alle  $q \in \mathbb{N}$  mit  $q \geq 1$  gilt. □

LEMMA 17. *Sei  $y \in \mathbb{R}_+$ ,  $p, q \in \mathbb{Z}$ ,  $q > 0$ . Dann gibt es ein  $x \in \mathbb{R}_+$ , welches die Gleichung  $x^q = y^p$  löst; dieses  $x$  wird mit  $y^{\frac{p}{q}}$  bezeichnet. Die Funktion  $f(y) = y^{\frac{p}{q}}$  ist strikt monoton fallend, wenn  $p < 0$  ist, und strikt monoton steigend, wenn  $p > 0$  ist.*

BEWEIS. Es ist genug, die Aussage des Lemmas für  $p = 1$  zu zeigen; die entsprechende Aussage für allgemeine  $p$  folgt aus der Anwendung des Lemmas auf  $y \mapsto (y^p)^{\frac{1}{q}}$ .

Wir betrachten die Menge  $S_y = \{t \in \mathbb{R}_+ \cup \{0\} : 0 \leq t^q < y\}$ . Diese Menge ist nach oben beschränkt, da für  $t > 1$  die Ungleichung  $t < t^q$  gilt, also auf jeden Fall  $t \leq \max(1, y)$  für jedes  $t \in S_y$  gilt. Daher gibt es  $x := \sup S$ , und wir behaupten, dass  $x^q = y$  gilt. Zunächst bemerken wir, dass  $S_y = [0, x)$  gilt; wäre  $x \in S_y$ , also  $x^q < y$ , so ist für  $0 < \delta < \min(x, \frac{y-x^q}{(2^q-1)x^{q-1}})$  die Ungleichung

$$\begin{aligned} (x + \delta)^q &= x^q + \sum_{j=1}^q \binom{q}{j} x^{q-j} \delta^j \\ &< x^q + \sum_{j=1}^q \binom{q}{j} x^{q-1} \delta \\ &< x^q + \sum_{j=1}^q \binom{q}{j} x^{q-1} \frac{y - x^q}{(2^q - 1)x^{q-1}} \\ &= x^q + \frac{y - x^q}{2^q - 1} \sum_{j=1}^q \binom{q}{j} \\ &= y \end{aligned}$$

erfüllt, also  $x$  keine obere Schranke von  $S_y$ ; im letzten Schritt haben wir die bekannte Folgerung

$$\sum_{j=0}^q \binom{q}{j} = 2^q$$

aus dem binomischen Lehrsatz verwendet.

Somit gilt  $x^q \geq y$ . Wäre nun  $x^q > y$ , so wäre wegen Lemma 16

$$a^2 = \left(x - \frac{x^q - y}{qx^{q-1}}\right)^q > x^q - q \frac{x^q - y}{qx^{q-1}} x^{q-1} = y,$$

also  $a < x$ , aber  $a^2 > y$ , was einen Widerspruch zu  $a \in S_y$  darstellt.

Die Funktion  $y^{\frac{1}{q}}$  ist damit durch

$$y^{\frac{1}{q}} = \sup S_y, \quad y > 0$$



definiert, was wegen Lemma 13 die Monotonie liefert; die strikte Monotonie folgt dann wegen  $x^q < x'^q$ , wenn  $0 < x < x'$ .  $\square$

**Übungsaufgabe 7.** Damit für  $r = \frac{p}{q} \in \mathbb{Q}$  durch  $y \mapsto y^r$  eine Funktion definiert wird, müssen wir nur noch zeigen, dass  $y^{\frac{p}{q}} = y^{\frac{p'}{q'}}$  gilt, wenn  $\frac{p}{q} = \frac{p'}{q'}$ . Führe dies durch und zeige, dass für beliebige  $r, r' \in \mathbb{Q}$  und  $y_1, y_2, y \in \mathbb{R}_+ \cup \{0\}$  die Gleichungen

$$(y_1 y_2)^r = y_1^r y_2^r, \quad y^r y^{r'} = y^{r+r'}, \quad (y^r)^{r'} = y^{rr'}$$

erfüllt sind.

**Übungsaufgabe 8.** Untersuche, für fixes  $y \in \mathbb{R}_+$ , die Monotonieeigenschaften von  $\mathbb{Q} \ni r \mapsto y^r$  (diese hängen von  $y$  ab).

## 2. Der Abstandsbegriff und topologische Grundbegriffe

**2.1. Der Betrag und der Abstand in  $\mathbb{R}$ .** Für eine reelle Zahl  $r$  definieren wir  $|r| = \sqrt{r^2}$ , oder (etwas konkreter)

$$|r| = \max(r, -r) = \begin{cases} r & r \geq 0 \\ -r & r \leq 0. \end{cases}$$

$|r|$  wird als *Betrag* oder *Absolutbetrag* von  $r$  bezeichnet. Mit Hilfe des Vorzeichens von  $r$ , welches durch

$$(1) \quad \operatorname{sgn} r = \begin{cases} +1 & r > 0 \\ 0 & r = 0 \\ -1 & r < 0 \end{cases}$$

definiert wird, können wir auch

$$|x| = (\operatorname{sgn} x)x$$

schreiben. Wir beachten gleich an dieser Stelle, dass  $|r| = 0$  genau für  $r = 0$  gilt, was oft in der Form

$$|x - y| = 0 \text{ genau dann, wenn } x = y, \quad x, y \in \mathbb{R},$$

verwendet wird.

Wir bezeichnen mit

$$d(x, y) = |x - y|, \quad x, y \in \mathbb{R}$$

den *Abstand* von  $x$  und  $y$ ;  $d$  ist dadurch als Funktion  $d: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  definiert. Aus Lemma 15 folgt sofort, dass  $|rs| = |r||s|$  gilt, was mit  $r = -1$ ,  $s = x - y$  die Symmetrie von  $d$  liefert:

$$d(x, y) = |x - y| = |-1||x - y| = |(-1)(x - y)| = |y - x| = d(y, x).$$

Eine weitere Eigenschaft des Betrags ist die *Dreiecksungleichung*:

LEMMA 18. Für  $x, y \in \mathbb{R}$  gilt

$$|x + y| \leq |x| + |y|.$$

Gleichheit gilt genau dann, wenn  $x = 0$  oder  $y = 0$ , oder, falls  $x \neq 0$ ,  $y \neq 0$ , wenn  $x$  und  $y$  dasselbe Vorzeichen haben. Für  $r, s, t \in \mathbb{R}$  gilt

$$d(r, t) \leq d(r, s) + d(s, t),$$

mit Gleichheit genau dann, wenn  $r \leq s \leq t$  oder  $t \leq s \leq r$  gilt.

Die zweite Formulierung erläutert besser den Namen Dreiecksungleichung: Der direkte Weg zwischen zwei Punkten ist nie länger als jener über einen Umweg.

BEWEIS. Es ist für beliebige  $x, y \in \mathbb{R}$  die Ungleichung  $xy \leq |x||y|$  erfüllt, wobei Gleichheit in genau zwei Fällen gilt: zunächst trivialerweise, wenn  $x = 0$  oder  $y = 0$ , oder wenn  $\operatorname{sgn} x = \operatorname{sgn} y$ . Damit ist

$$|x + y|^2 = (x + y)^2 = x^2 + 2xy + y^2 \leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2,$$

was wegen der Monotonie der Wurzelfunktion (Lemma 14)  $|x + y| \leq |x| + |y|$  impliziert. Die Behauptung über die Gleichheit folgt aus der schon beobachteten Gleichheit in der Ungleichung  $xy \leq |x||y|$ .

Wir können nun  $d(r, t)$  durch die Verwendung einer “klugen Null” mit Hilfe des soeben Bewiesenen abschätzen:

$$\begin{aligned} d(r, t) &= |r - t| \\ &= |r - s + s - t| \\ &= |(r - s) + (s - t)| \\ &\leq |r - s| + |s - t| \\ &= d(r, s) + d(s, t). \end{aligned}$$

Gleichheit gilt trivialerweise, wenn  $r = s$  oder  $t = s$  gilt, oder, wenn  $r - s$  und  $s - t$  dasselbe Vorzeichen haben: ist  $r - s > 0$  und  $s - t > 0$ , so ist  $r > s > t$ ; ist auf der anderen Seite  $r - s < 0$  und  $s - t < 0$ , so ist  $r < s < t$ .  $\square$

Oft benötigt man auch eine untere Schranke für  $|x - y|$ . Die Frage nach einer solchen ist oft delikater Natur; für die größeren Fälle genügt aber manchmal eine einfache Folgerung aus Lemma 18, welche gerne als *umgekehrte Dreiecksungleichung* bezeichnet wird.

LEMMA 19. Für  $x, y \in \mathbb{R}$  gilt

$$||x| - |y|| \leq |x - y|.$$

Für  $r, s, t \in \mathbb{R}$  gilt

$$|d(r, s) - d(s, t)| \leq d(r, t).$$

BEWEIS. Wir verwenden wieder die kluge Null: Es ist

$$\begin{aligned} |x| &= |x - y + y| = |(x - y) + y| \leq |x - y| + |y| \\ |y| &= |y - x + x| = |(y - x) + x| \leq |y - x| + |x|, \end{aligned}$$

also sowohl  $|x| - |y| \leq |x - y|$  als auch  $|y| - |x| \leq |x - y|$ . Die Behauptung für den Abstand folgt nun, indem man  $x = r - s$  und  $y = t - s$  setzt.  $\square$

**Übungsaufgabe 9.** Untersuche die Bedingungen, unter denen Gleichheit in Lemma 19 herrscht.

Um zu zeigen, dass zwei Punkte  $x, y$  gleich sind, ist es oft hilfreich zu zeigen dass sie beliebig nahe zueinander sind.

LEMMA 20. Es ist  $x = y$  genau dann, wenn für alle  $\varepsilon > 0$  die Ungleichung  $|x - y| < \varepsilon$  erfüllt ist.

**Übungsaufgabe 10.** Beweise Lemma 20.

Mit Hilfe des Betrags werden für  $x \in \mathbb{R}$  spezielle Mengen definiert, welche alle Punkte die nahe zu  $x$  sind, umfassen: Der  $\varepsilon$ -Ball  $B_\varepsilon(x)$  um  $x$  ist durch

$$B_\varepsilon(x) := \{y \in \mathbb{R} : d(x, y) < \varepsilon\}$$

definiert. Wir können natürlich auch

$$B_\varepsilon(x) = (x - \varepsilon, x + \varepsilon)$$

schreiben, wollen aber über  $B_\varepsilon(x)$  als die Menge aller Punkte denken, welche von  $x$  weniger als  $\varepsilon$  entfernt sind.

LEMMA 21. Sei  $y \in B_\varepsilon(x)$ . Dann gibt es ein  $\delta > 0$ , sodass  $B_\delta(y) \subset B_\varepsilon(x)$ .

BEWEIS. Wir setzen  $\delta = \varepsilon - d(x, y) > 0$ . Für beliebiges  $z \in B_\delta(y)$  ist dann nach Lemma 18

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

also  $z \in B_\varepsilon(x)$ . Es folgt die gewünschte Inklusion  $B_\delta(y) \subset B_\varepsilon(x)$ .  $\square$

LEMMA 22. Seien  $x_1, x_2 \in \mathbb{R}$ , und  $\varepsilon_1 > 0$ ,  $\varepsilon_2 > 0$ . Dann gibt es für jedes  $y \in B_{\varepsilon_1}(x_1) \cap B_{\varepsilon_2}(x_2)$  ein  $\delta > 0$  sodass

$$B_\delta(y) \subset B_{\varepsilon_1}(x_1) \cap B_{\varepsilon_2}(x_2)$$

gilt.

BEWEIS. Wir setzen  $\delta = \min(\varepsilon_1 - d(x_1, y), \varepsilon_2 - d(x_2, y))$ . Dann ist für beliebiges  $z \in B_\delta(y)$  sowohl

$$d(x_1, z) \leq d(x_1, y) + d(y, z) < d(x_1, y) + (\varepsilon_1 - d(x_1, y)) = \varepsilon_1$$

als auch

$$d(x_2, z) \leq d(x_2, y) + d(y, z) < d(x_2, y) + (\varepsilon_2 - d(x_2, y)) = \varepsilon_2,$$

also  $z \in B_{\varepsilon_1}(x_1) \cap B_{\varepsilon_2}(x_2)$ . Da  $z \in B_\delta(y)$  beliebig war, folgt  $B_\delta(y) \subset B_{\varepsilon_1}(x_1) \cap B_{\varepsilon_2}(x_2)$ .  $\square$

## 2.2. Der Umgebungsbegriff und topologische Grundbegriffe.

**Definition 1.** Eine Menge  $U$  ist eine *Umgebung* von  $x$ , wenn es ein  $\varepsilon > 0$  gibt, sodass  $B_\varepsilon(x) \subset U$ . Die Menge der Umgebungen  $\mathcal{N}(x)$  von  $x$  wird als das *Umgebungssystem* von  $x$  bezeichnet.

**Beispiel 1.** Die folgenden Mengen sind alle Beispiele von Umgebungen von  $0 \in \mathbb{R}$ :

- $(-1, 1)$ , da  $(-1, 1) = B_1(0)$ ;
- $(-1, \frac{1}{2})$ , da  $B_{\frac{1}{2}}(0) \subset (-1, \frac{1}{2})$ ;
- $[-1, 1]$ , wiederum da  $(-1, 1) = B_1(0)$ ;
- $[-1, 1] \cup \mathbb{Z}$ , wiederum da  $(-1, 1) = B_1(0)$ .

Im folgenden Lemma wollen wir einige einfache Eigenschaften von Umgebungen festhalten.

**LEMMA 23.** Seien  $U, V \in \mathcal{N}(x)$ , dann ist auch  $U \cap V \in \mathcal{N}(x)$ ; für beliebiges  $W$  mit  $U \subset W$  ist  $W \in \mathcal{N}(x)$ . Ist  $U_j \in \mathcal{N}(x)$ ,  $j \in J$ , eine beliebige Familie von Umgebungen von  $x$ , so ist auch  $\cup_j U_j \in \mathcal{N}(x)$ . Wenn  $V \in \mathcal{N}(x)$ , so gibt es ein  $U \in \mathcal{N}(x)$  mit  $U \subset V$  für welches  $U \in \mathcal{N}(y)$  für alle  $y \in U$  gilt.

**Übungsaufgabe 11.** Beweise Lemma 23.

**Definition 2.** Eine Menge  $U \subset \mathbb{R}$  heisst *offen*, wenn sie eine Umgebung jedes ihrer Punkte ist, d.h. wenn  $U \in \mathcal{N}(x)$  für alle  $x \in U$  ist. Eine Menge  $F \subset \mathbb{R}$  heisst *abgeschlossen*, wenn  $F^c$  offen ist.

**Beispiel 2.**

- $\emptyset$  und  $\mathbb{R}$  sind trivialerweise sowohl offen als auch abgeschlossen.
- Ein Intervall von der Form  $(a, b)$  mit  $a, b \in \mathbb{R}$  ist offen: Für  $x \in (a, b)$  ist  $\varepsilon = \min(x - a, b - x) > 0$  und  $B_\varepsilon(x) \subset (a, b)$ . Wir sagen deswegen,  $(a, b)$  sei ein *offenes Intervall*. Ähnliche Überlegungen zeigen, dass  $(-\infty, a)$  und  $(a, \infty)$  offen sind.

**LEMMA 24.** Für  $U, V \subset \mathbb{R}$  offen ist  $U \cap V$  offen. Ist  $U_j \subset \mathbb{R}$ ,  $j \in J$  eine beliebige Familie offener Mengen, so ist auch  $\cup_j U_j$  offen. Für  $F, G \subset \mathbb{R}$  abgeschlossen ist  $F \cup G$  abgeschlossen. Ist  $F_j \subset \mathbb{R}$ ,  $j \in J$  eine beliebige Familie abgeschlossener Mengen, so ist auch  $\cap_j F_j$  abgeschlossen.

BEWEIS. Sei  $x \in U \cap V$ . Da  $U \in \mathcal{N}(x)$  und  $V \in \mathcal{N}(x)$  ist nach Lemma 23 auch  $U \cap V \in \mathcal{N}(x)$ ; da  $x \in U \cap V$  beliebig war, ist  $U \cap V$  offen.

Sei  $x \in \cup_j U_j$ . Dann gibt es ein  $j_0 \in J$ , sodass  $x \in U_{j_0}$  ist. Da  $U_{j_0}$  offen ist, ist  $U_{j_0} \in \mathcal{N}(x)$ . Damit ist auch  $\cup_j U_j \supset U_{j_0}$  eine Umgebung von  $x$ ; also ist  $\cup_j U_j$  offen.  $\square$

**Beispiel 3.** Seien  $a, b \in \mathbb{R}$ , dann ist  $[a, b]$  abgeschlossen, da  $[a, b]^c = (-\infty, a) \cup (b, \infty)$  eine Vereinigung zweier nach Beispiel 2 offener Mengen, also nach Lemma 24 wieder offen. Wir sagen deswegen, das Intervall  $[a, b]$  sei ein *abgeschlossenes Intervall*.

Das Intervall  $(a, b]$  ist weder offen noch abgeschlossen, da es keine Umgebung von  $b \in (a, b]$  ist, und das Komplement  $(-\infty, a] \cup (b, \infty)$  keine Umgebung von  $a$  ist.

Beliebige Durchschnitte offener Mengen müssen nicht wieder offen sein, so ist zum Beispiel  $(a, b] = \cap_{j \in \mathbb{N}} (a, b + \frac{1}{j})$ .

**Übungsaufgabe 12.** Gib ein Beispiel einer abzählbaren Vereinigung abgeschlossener Mengen, die nicht abgeschlossen ist.

**Definition 3.**  $p \in \mathbb{R}$  ist ein *Häufungspunkt* von  $X \subset \mathbb{R}$  wenn für jedes  $U \in \mathcal{N}(p)$  die Menge  $(U \setminus \{p\}) \cap X \neq \emptyset$  ist.

**Beispiel 4.** Der Punkt  $a$  ist ein Häufungspunkt der Intervalle  $(a, b)$  und  $[a, b)$ . Nach Lemma 7 ist jede reelle Zahl Häufungspunkt von  $\mathbb{Q} \subset \mathbb{R}$ . Die Menge  $\mathbb{Z} \subset \mathbb{R}$  besitzt keine Häufungspunkte.

LEMMA 25. Sei  $X \subset \mathbb{R}$ . Dann gibt es eine kleinste abgeschlossene Menge, welche  $X$  enthält; wir bezeichnen sie mit  $\bar{X}$  und nennen sie den Abschluss von  $X$ .  $\bar{X}$  besteht aus  $X$  und allen Häufungspunkten von  $X$ .

BEWEIS. Da  $\mathbb{R}$  abgeschlossen ist, ist die Menge  $\mathcal{F} = \{F \subset \mathbb{R} : F \text{ abgeschlossen, } X \subset F\}$  nicht leer. Damit ist die Menge  $\bar{X} = \bigcap_{F \in \mathcal{F}} F$  die kleinste abgeschlossene Menge welche  $X$  enthält.

Wir behaupten nun, dass eine abgeschlossene Menge  $F$  jeden ihrer Häufungspunkte enthält: tatsächlich gibt es für  $p \in F^c$  (da  $F^c$  definitionsgemäß offen ist) eine Umgebung  $U \in \mathcal{N}(p)$  mit  $U \subset F^c$ , also ist  $U \cap F = \emptyset$ . Damit enthält jedes  $F \in \mathcal{F}$  alle Häufungspunkte von  $X$ .

Ist  $p \in X^c$  kein Häufungspunkt von  $X$ , so gibt es eine offene Umgebung  $U$  von  $p$  mit  $U \cap X = \emptyset$ . Damit ist  $U^c \in \mathcal{F}$ , also ist  $p \notin \bar{X}$ .  $\square$

PROPOSITION 1. Sei  $X \subset \mathbb{R}$  eine beschränkte, unendliche Menge (d.h., die Kardinalität von  $X$  ist zumindest abzählbar unendlich). Dann hat  $X$  zumindest einen Häufungspunkt.

BEWEIS. Da  $X$  beschränkt ist, gibt es ein  $M > 0$  mit  $X \subset [-M, M]$ . Wir definieren  $a_0 = -M$ ,  $b_0 = M$  und zerlegen das Intervall  $I_0 = [a_0, b_0]$  in zwei Teile:

$$[a_0, b_0] = \left[ a_0, \frac{a_0 + b_0}{2} \right] \cup \left[ \frac{a_0 + b_0}{2}, b_0 \right].$$

Mindestens einer dieser beiden Teile enthält unendlich viele der Elemente von  $X$ ; wir wählen diesen Teil und bezeichnen ihn mit  $[a_1, b_1]$ . Wir bemerken auch gleich, dass  $b_1 - a_1 = M$  ist.

Angenommen, wir haben ein Intervall  $I_j = [a_j, b_j]$  konstruiert, für das  $X \cap I_j$  eine unendliche Menge ist, und  $b_j - a_j = M2^{-j+1}$ . Wir zerlegen dieses Intervall  $I_j$  wiederum in zwei Hälften,

$$[a_j, b_j] = \left[ a_j, \frac{a_j + b_j}{2} \right] \cup \left[ \frac{a_j + b_j}{2}, b_j \right].$$

Mindestens einer der Teile hat die Eigenschaft, dass er unendlich viele Elemente von  $X$  enthält. Wir wählen diesen Teil und bezeichnen ihn mit  $I_{j+1} = [a_{j+1}, b_{j+1}]$ ; wir bemerken, dass  $b_{j+1} - a_{j+1} = \frac{1}{2}(b_j - a_j) = M2^{-j}$ .

Wir erhalten auf diese Weise eine absteigende Folge von Intervallen

$$I_0 \supset I_1 \supset I_2 \supset \dots$$

und behaupten zunächst, dass jeder Punkt  $p \in \bigcap_j I_j$  ein Häufungspunkt von  $X$  ist. Sei also  $U \subset \mathcal{N}(p)$  beliebig, und  $\varepsilon > 0$  so gewählt, dass  $B_\varepsilon(p) \subset U$ . Nach Lemma 8 können wir ein  $j > 0$  finden, sodass  $M2^{-j+2} < \varepsilon$  ist. Da  $p \in [a_j, b_j]$  ist, und  $b_j - a_j = M2^{-j+1}$  gilt, haben wir

$$|x - p| \leq |x - a_j| + |p - a_j| \leq M2^{-j+1} + M2^{-j+1} = M2^{-j+2}, \quad x \in [a_j, b_j],$$

also  $[a_j, b_j] \subset B_\varepsilon(p) \subset U$ ; und nachdem die Menge  $X \cap [a_j, b_j]$  unendlich ist, enthält sie zumindest einen Punkt, der von  $p$  verschieden ist.

Es stellt sich heraus, dass der Durchschnitt der Intervalle tatsächlich nur aus einem Punkt besteht (für unsere Proposition ist es nur wichtig, dass er nicht leer ist!). Zunächst sind die Zahlen  $A = \sup\{a_j : j \in \mathbb{N}\}$  und  $B = \inf\{b_j : j \in \mathbb{N}\}$  wohldefiniert, da beide der betrachteten Mengen beschränkt sind. Es gilt auch  $A \leq B$  (da  $A \leq b_j$  für alle  $j$  gilt); damit ist  $[A, B] \subset \bigcap_j [a_j, b_j]$ . Da nun aber  $B - A \leq b_j - a_j = M2^{-j+1}$  ist, muss nach Lemma 8 notwendigerweise  $A = B$  sein.  $\square$

**Definition 4.** Wir sagen,  $p \in \mathbb{R}$  ist ein *innerer Punkt* von  $X \subset \mathbb{R}$ , wenn es ein  $U \in \mathcal{N}(p)$  mit  $U \subset X$  gibt.

**Beispiel 5.** Jeder innere Punkt von  $X$  ist ein Häufungspunkt von  $X$ . Die inneren Punkte von  $[a, b]$  sind genau die Punkte in  $(a, b)$ .

LEMMA 26. Sei  $X \subset \mathbb{R}$ . Dann gibt es eine grösste offene Menge, welche in  $X$  enthalten ist; wir bezeichnen sie mit  $X^\circ$  und nennen sie das Innere von  $X$ .  $X^\circ$  besteht aus allen inneren Punkten von  $X$ .

BEWEIS. Da  $\emptyset$  offen ist, ist die Menge  $\mathcal{O} = \{U \subset \mathbb{R} : U \text{ offen, } U \subset X\}$  nicht leer. Damit ist  $X^\circ = \bigcup_{U \in \mathcal{O}} U$  die grösste offene Menge, welche in  $X$  enthalten ist.

Definitionsgemäß enthält  $X^\circ$  nur innere Punkte von  $X$ . Ist nun andererseits  $p$  kein innerer Punkt von  $X$ , so gibt es keine offene Menge  $U$ , welche  $p$  enthält und in  $X$  enthalten ist (wir verwenden hier die letzte Eigenschaft welche in Lemma 23 bewiesen wurde). Also ist  $p \notin X^\circ$ .  $\square$

### 3. Konvergenz von Folgen und Reihen

#### 3.1. Folgen.

**Definition 5.** Eine (reelle) *Folge* ist eine Abbildung  $a: \mathbb{N} \rightarrow \mathbb{R}$ . Die Bilder  $a(n) = a_n$  heissen die *Glieder* der Folge. Wir schreiben für die Folge  $a$  oft einfach  $(a_n)_{n=1}^\infty$ . Wir sagen, die Folge  $(a_n)_{n=1}^\infty$  *konvergiert* für  $n \rightarrow \infty$  gegen  $x \in \mathbb{R}$ , wenn für jede Umgebung  $U \in \mathcal{N}(x_0)$  von  $x_0$  ein  $N \in \mathbb{N}$  existiert sodass  $a_n \in U$  für alle  $n \geq N$ . Wenn  $(a_n)_{n=1}^\infty$  gegen  $x_0$  konvergiert, so schreiben wir

$$x = \lim_{n \rightarrow \infty} a_n, \quad \text{oder auch } a_n \rightarrow x \ (n \rightarrow \infty).$$

Die Menge der reellen Folgen  $\mathbb{R}^{\mathbb{N}}$  wird zu einem reellen Vektorraum, wenn man die Addition und Skalarmultiplikation der Folgen  $a$  und  $b$  gliedweise definiert, also

$$(a + \lambda b)(n) = a(n) + \lambda b(n) \in \mathbb{R}, \quad n \in \mathbb{N}$$

für  $a, b \in \mathbb{R}^{\mathbb{N}}$  und  $\lambda \in \mathbb{R}$  setzt.

**Beispiel 6.** Die Folge  $f$ , welche durch

$$a_n = \frac{1}{n}$$

definiert ist, konvergiert gegen 0: Wenn  $U$  eine beliebige Umgebung von 0 ist, so gibt es ein  $\varepsilon > 0$  sodass  $B_\varepsilon(0) \subset U$ . Nach Lemma 5 gibt es ein  $N \in \mathbb{N}$  sodass  $\frac{1}{N} < \varepsilon$  gilt. Für  $n \geq N$  ist aber  $0 < a_n \leq a_N < \varepsilon$ , also  $a_n \in B_\varepsilon(0)$ ; somit konvergiert  $a_n$  gegen 0.

**LEMMA 27.** Die Folge  $(a_n)_{n=1}^\infty$  konvergiert gegen  $x \in \mathbb{R}$  genau dann, wenn für jedes  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  existiert sodass  $a_n \in B_\varepsilon(x)$ , oder mit anderen Worten  $|a_n - x| < \varepsilon$ , für  $n \geq N$  gilt.

**BEWEIS.** Da  $B_\varepsilon(x) \in \mathcal{N}(x)$  gilt, ist die Bedingung des Lemmas im Falle der Konvergenz auf jeden Fall erfüllt. Ist auf der anderen Seite die Bedingung des Lemmas erfüllt, und  $U \in \mathcal{N}(x)$ , so gibt es nach der Definition einer Umgebung ein  $\varepsilon > 0$  für welches  $B_\varepsilon(x) \subset U$  ist. Nach der gegebenen Bedingung gibt es ein  $N \in \mathbb{N}$  gibt, sodass für  $n \geq N$  beliebig  $a_n \in B_\varepsilon(x) \subset U$  gilt; also ist  $\lim_{n \rightarrow \infty} a_n = x$ .  $\square$

Um Aussagen über Folgen wie

$$a_n = \frac{n^2 + 2n + 1}{3n^2 - 1}$$

(diese Folge konvergiert gegen  $1/3$ ) zu machen, ist es von Vorteil, von unserem Wissen über das Konvergenzverhalten von  $1/n$  auszugehen und eine gewisse Verträglichkeit der Konvergenz mit den Operationen der Addition und Multiplikation zu benutzen. Wir beginnen unsere Betrachtungen mit der Addition und Skalarmultiplikation.

**LEMMA 28.** Seien  $x, y, \lambda \in \mathbb{R}$  und  $U \in \mathcal{N}(x + \lambda y)$ . Dann gibt es ein  $V \in \mathcal{N}(x)$  und ein  $W \in \mathcal{N}(y)$  sodass  $V + \lambda W \subset U$  ist.

**BEWEIS.** Wir nehmen ohne Beschränkung der Allgemeinheit an, dass  $\lambda \neq 0$  ist (die Behauptung ist sonst trivial). Es gibt ein  $\varepsilon > 0$  mit  $B_\varepsilon(x + \lambda y) \subset U$ . Wir behaupten, dass mit

$$\delta_x = \frac{\varepsilon}{2}, \quad \delta_y = \frac{\varepsilon}{2|\lambda|},$$

$B_{\delta_x}(x) + \lambda B_{\delta_y}(y) \subset B_\varepsilon(x + \lambda y)$  gilt, und damit die Behauptung des Lemmas für  $V := B_{\delta_x}(x)$  und  $W := B_{\delta_y}(y)$  gilt.

Sei also  $p \in B_{\delta_x}(x)$  und  $q \in B_{\delta_y}(y)$ , dann ist

$$\begin{aligned} |(p + \lambda q) - (x + \lambda y)| &= |(p - x) + \lambda(q - y)| \\ &\leq |p - x| + |\lambda||q - y| \\ &\leq \delta_x + |\lambda|\delta_y \\ &\leq \frac{\varepsilon}{2} + |\lambda|\frac{\varepsilon}{2|\lambda|} \\ &= \varepsilon, \end{aligned}$$

also  $p + \lambda q \in B_\varepsilon(x + \lambda y) \subset U$ .  $\square$

SATZ 1. Seien  $(a_n)_{n=1}^\infty$  und  $(b_n)_{n=1}^\infty$  konvergente Folgen,  $\lambda \in \mathbb{R}$ . Dann ist auch  $(a_n + \lambda b_n)_{n=1}^\infty$  konvergent, und es gilt

$$\lim_{n \rightarrow \infty} (a_n + \lambda b_n) = \lim_{n \rightarrow \infty} a_n + \lambda \lim_{n \rightarrow \infty} b_n.$$

BEWEIS. Wir schreiben  $x = \lim_{n \rightarrow \infty} a_n$  und  $y = \lim_{n \rightarrow \infty} b_n$ . Sei  $U$  eine Umgebung von  $x + \lambda y$ . Nach Lemma 28 gibt es eine Umgebung  $V$  von  $x$  und eine Umgebung  $W$  von  $y$  sodass  $V + \lambda W \subset U$  gilt. Da  $a_n$  für  $n \rightarrow \infty$  gegen  $x$  konvergiert, gibt es ein  $N_x$  sodass  $a_n \in V$  für  $n \geq N_x$  gilt; da  $b_n$  für  $n \rightarrow \infty$  gegen  $y$  konvergiert, gibt es ein  $N_y$  sodass  $b_n \in W$  für  $n \geq N_y$  gilt. Ist also  $n \geq N := \max(N_x, N_y)$ , so ist  $a_n + \lambda b_n \in V + \lambda W \subset U$ . Damit ist gezeigt, dass  $a_n + \lambda b_n$  für  $n \rightarrow \infty$  gegen  $x + \lambda y$  konvergiert, wie behauptet.  $\square$

Um die Multiplikation zu behandeln, führen wir in Anlehnung an unsere Schreibweise für die Summe von Umgebungen für beliebige Mengen  $S, T \subset \mathbb{R}$  die Menge

$$ST = \{st : s \in S, t \in T\}$$

ein. Ist  $S \subset \mathbb{R} \setminus \{0\}$ , so definieren wir

$$S^{-1} = \{s^{-1} : s \in S\}.$$

LEMMA 29. Seien  $x, y \in \mathbb{R}$ , und  $U \in \mathcal{N}(xy)$ . Dann gibt es Umgebungen  $V \in \mathcal{N}(x)$  und  $W \in \mathcal{N}(y)$  sodass  $VW \subset U$  gilt.

BEWEIS. Wir nehmen ohne Beschränkung der Allgemeinheit an, Es gibt ein  $\varepsilon > 0$ , sodass  $B_\varepsilon(xy) \subset U$  gilt. Da

$$\begin{aligned} |pq - xy| &= |pq - py + py - xy| \\ &\leq |p||q - y| + |p - x||y| \end{aligned}$$

können wir

$$\delta_x = \min\left(\frac{\varepsilon}{2 \max(|y|, 1)}, 1\right), \quad \delta_y = \frac{\varepsilon}{2(|x| + 1)}$$

wählen um für  $p \in B_{\delta_x}(x)$  und  $q \in B_{\delta_y}(y)$  wie folgt abzuschätzen:

$$\begin{aligned} |pq - xy| &= |pq - py + py - xy| \\ &\leq |p||q - y| + |p - x||y| \\ &\leq |p|\delta_y + \delta_x|y| \\ &\leq (|x| + 1)\frac{\varepsilon}{2(|x| + 1)} + \frac{\varepsilon}{2 \max(|y|, 1)}|y| \leq \varepsilon; \end{aligned}$$

also ist  $B_{\delta_x}(x)B_{\delta_y}(y) \subset B_\varepsilon(xy)$  und die Behauptung des Lemmas mit  $V = B_{\delta_x}(x)$  und  $W = B_{\delta_y}(y)$  erfüllt.  $\square$

SATZ 2. Seien  $(a_n)_{n=1}^\infty$  und  $(b_n)_{n=1}^\infty$  konvergente Folgen. Dann ist auch  $(a_n b_n)_{n=1}^\infty$  konvergent, und es gilt

$$\lim_{n \rightarrow \infty} (a_n b_n) = \lim_{n \rightarrow \infty} a_n \lim_{n \rightarrow \infty} b_n.$$

BEWEIS. Wir schreiben  $x = \lim_{n \rightarrow \infty} a_n$  und  $y = \lim_{n \rightarrow \infty} b_n$ . Sei  $U$  eine Umgebung von  $xy$ . Nach Lemma 29 gibt es eine Umgebung  $V$  von  $x$  und eine Umgebung  $W$  von  $y$  sodass  $VW \subset U$  gilt. Da  $a_n$  für  $n \rightarrow \infty$  gegen  $x$  konvergiert, gibt es ein  $N_x$  sodass  $a_n \in V$  für  $n \geq N_x$  gilt; da  $b_n$  für  $n \rightarrow \infty$  gegen  $y$  konvergiert, gibt es ein  $N_y$  sodass  $b_n \in W$  für  $n \geq N_y$  gilt. Ist also  $n \geq N := \max(N_x, N_y)$ , so ist  $a_n b_n \in VW \subset U$ . Damit ist gezeigt, dass  $a_n b_n$  für  $n \rightarrow \infty$  gegen  $xy$  konvergiert, wie behauptet.  $\square$

Dem Leser wird aufgefallen sein, dass der Beweis von Theorem 2 aus Lemma 29 wortgleich wie der Beweis von Theorem 1 aus Lemma 28 geführt wurde. Die Bedingung, die dahinter steckt, verlangt deswegen danach, ihre allgemeine Formulierung in der Form der Stetigkeit zu finden. Bevor wir dies tun, wollen wir noch die Konvergenz von  $a_n^{-1}$  untersuchen.

LEMMA 30. Sei  $x \in \mathbb{R}$  mit  $x \neq 0$ , und  $U \in \mathcal{N}(x^{-1})$ . Dann gibt es eine Umgebung  $V \in \mathcal{N}(x)$  mit  $V^{-1} \subset U$ .

BEWEIS. Es gibt ein  $\varepsilon > 0$ , sodass  $B_\varepsilon(x^{-1}) \subset U$ . Wir setzen unsere übliche Abschätzung mit

$$\begin{aligned} \left| \frac{1}{p} - \frac{1}{x} \right| &= \left| \frac{x-p}{xp} \right| \\ &= \frac{|x-p|}{|xp|} \end{aligned}$$

an und wählen  $\delta_x = \min\left(\frac{|x|}{2}, \varepsilon \frac{|x|^2}{2}\right)$  um so für  $p \in B_{\delta_x}(x)$  zunächst nach Lemma 19

$$\frac{|x|}{2} \geq |x-p| \geq |x| - |p|,$$

also  $|p| \geq \frac{|x|}{2}$  abzuschätzen. Damit ist nun für  $p \in B_{\delta_x}(x)$

$$\begin{aligned} \left| \frac{1}{p} - \frac{1}{x} \right| &= \left| \frac{x-p}{xp} \right| \\ &= \frac{|x-p|}{|xp|} \\ &\leq \frac{2\delta_x}{|x|^2} \\ &= \varepsilon, \end{aligned}$$

also  $p^{-1} \in B_\varepsilon(x^{-1})$  und die Folgerung des Lemmas mit  $V = B_{\delta_x}(x)$  erfüllt.  $\square$

**SATZ 3.** Sei  $(a_n)_{n=1}^\infty$  eine konvergente Folge mit  $\lim_{n \rightarrow \infty} a_n^{-1} = x \neq 0$ . Dann gibt es ein  $k \in \mathbb{N}$  sodass  $a_n \neq 0$  für  $n \geq k$  ist, und es gilt

$$\lim_{n \rightarrow \infty} a_{n+k}^{-1} = \frac{1}{\lim_{n \rightarrow \infty} a_n}.$$

BEWEIS. Da  $x \neq 0$  ist, können wir durch  $B_{\frac{|x|}{2}}(x)$  eine Umgebung von  $x$  definieren. Da  $a_n$  für  $n \rightarrow \infty$  gegen  $x$  konvergiert, gibt es ein  $k$  sodass  $a_n \in B_{\frac{|x|}{2}}(x)$  für  $k \geq n$  ist. Damit ist  $\frac{|x|}{2} \geq |a_n - x| \geq |x| - |a_n|$ , also  $|a_n| \geq \frac{|x|}{2} > 0$  für  $n \geq k$ .

Sei nun  $U$  eine beliebige Umgebung von  $x^{-1}$ . Nach Lemma 30 gibt es eine Umgebung  $V$  von  $x$  mit  $V^{-1} \subset U$ , und da  $a_n$  für  $n \rightarrow \infty$  gegen  $x$  konvergiert, gibt es ein  $N \in \mathbb{N}$  sodass  $a_n \in V$  für alle  $n \geq N$  ist. Es folgt wieder, dass  $a_n^{-1} \in U$  für  $n \geq N$  ist, also  $\lim_{n \rightarrow \infty} a_{n+k}^{-1} = x^{-1}$ .  $\square$

Wir können nun zu unserem Beispiel von vorher zurückkehren und wie folgt rechnen:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{n^2 + 2n + 1}{3n^2 - 1} &= \lim_{n \rightarrow \infty} \frac{1 + \frac{2}{n} + \frac{1}{n^2}}{3 - \frac{1}{n^2}} \\ &= \frac{1 + \lim_{n \rightarrow \infty} \frac{2}{n} + \lim_{n \rightarrow \infty} \frac{1}{n^2}}{3 - \frac{1}{n^2}} \\ &= \frac{1 + 2 \lim_{n \rightarrow \infty} \frac{1}{n} + \lim_{n \rightarrow \infty} \frac{1}{n} \lim_{n \rightarrow \infty} \frac{1}{n}}{3 - \lim_{n \rightarrow \infty} \frac{1}{n} \lim_{n \rightarrow \infty} \frac{1}{n}} \\ &= \frac{1}{3} \end{aligned}$$

Zu beachten ist bei Rechnungen wie dieser immer, dass die Rechtfertigung der Gleichheitszeichen durch Anwendungen von Theorem 1, Theorem 2, und Theorem 3 erfolgt.

**Definition 6.** Ein Punkt  $x \in \mathbb{R}$  ist ein *Häufungswert* der Folge  $(a_n)_{n \in \mathbb{N}}$ , wenn für jede Umgebung  $U \in \mathcal{N}(x)$  die Menge  $\{n \in \mathbb{N} : a_n \in U\}$  unendlich ist.

**Beispiel 7.** 1 und  $-1$  sind Häufungswerte der Folge

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

Wir zeigen die Behauptung für 1; ähnliche Überlegungen zeigen, dass auch  $-1$  ein Häufungswert von  $(a_n)_{n \in \mathbb{N}}$  ist. Wenn  $U$  eine Umgebung von 1 ist, so gibt es ein  $\varepsilon > 0$  sodass  $B_\varepsilon(1) \subset U$  gilt. Nach Lemma 5 gibt es ein  $N \in \mathbb{N}$  mit  $1/N < \varepsilon$ . Damit ist für alle  $k \in \mathbb{N}$  mit  $2k \geq N$

$$|a_{2k} - 1| = \left| (-1)^{2k} + \frac{1}{2k} - 1 \right| = \frac{1}{2k} \leq \frac{1}{N} < \varepsilon,$$

also  $a_{2k} \in B_\varepsilon(1)$ .

**LEMMA 31.**  $x \in \mathbb{R}$  ist ein Häufungswert von  $(a_n)_{n \in \mathbb{N}}$  genau dann, wenn für jedes  $\varepsilon > 0$  und jedes  $N \in \mathbb{N}$  ein  $n \in \mathbb{N}$  existiert sodass  $n \geq N$  und  $|a_n - x| < \varepsilon$  gilt.

**BEWEIS.** Sei zunächst  $x$  ein Häufungswert von  $(a_n)_{n \in \mathbb{N}}$ , und  $\varepsilon > 0$  sowie  $N \in \mathbb{N}$  gegeben. Da  $B_\varepsilon(x)$  eine Umgebung von  $x$  ist, ist  $X = \{k \in \mathbb{N} : a_k \in B_\varepsilon(x)\}$  unendlich. Damit gibt es ein  $n \in X$  mit  $n \geq N$ , und für dieses  $n$  ist  $a_n \in B_\varepsilon(x)$ , also ist  $|a_n - x| < \varepsilon$ .

Wenn andererseits die Bedingung des Lemmas erfüllt ist, und  $U$  eine beliebige Umgebung von  $x$  ist, so gibt es ein  $\varepsilon > 0$  sodass  $B_\varepsilon(x) \subset U$  ist. Wir zeigen, dass  $\{k \in \mathbb{N} : a_k \in U\}$  unendlich ist, indem wir zeigen, dass  $X = \{k \in \mathbb{N} : a_k \in B_\varepsilon(x)\}$  eine strikt monotone Folge  $(k_\ell)_{\ell \in \mathbb{N}}$  enthält, also unendlich ist. Dazu wenden wir die Bedingung des Lemmas für  $N = 1$  an und erhalten ein  $k_1$  mit  $|a_{k_1} - x| < \varepsilon$ , also  $k_1 \in X$ . Angenommen, wir haben bereits  $k_\ell \in X$  gefunden, so gibt uns die Bedingung des Lemmas ein  $k_{\ell+1} \geq k_\ell + 1$  mit  $|a_{k_{\ell+1}} - x| < \varepsilon$ , also  $k_{\ell+1} \in X$ . Offensichtlich ist die so konstruierte Folge  $(k_\ell)_{\ell \in \mathbb{N}}$  strikt monoton wie behauptet.  $\square$

Im Beweis des obigen Lemmas wurde *fast* eine Folge konstruiert, die ein ‘‘Teil’’ der gegebenen Folge  $(a_n)_{n \in \mathbb{N}}$  ist und gegen den Häufungswert  $x$  *konvergiert*—wenn man das  $\varepsilon$  im zweiten Teil des Beweises nicht fix hält, sondern immer kleiner (z.B. als  $1/\ell$ ) wählt, sollten die so konstruierten  $a_{k_\ell}$  für  $\ell \rightarrow \infty$  gegen  $x$  konvergieren. Um diese Beobachtung zu präzisieren, benötigen wir den Begriff der *Teilfolge*.

**Definition 7.** Eine Folge  $(b_\ell)_{\ell \in \mathbb{N}}$  ist eine *Teilfolge* der Folge  $(a_n)_{n \in \mathbb{N}}$ , wenn es eine strikt monoton wachsende Folge  $(n_\ell)_{\ell \in \mathbb{N}}$  natürlicher Zahlen (d.h.  $n_\ell \in \mathbb{N}$  und  $n_\ell < n_{\ell+1}$  für alle  $\ell \in \mathbb{N}$ ), sodass  $b_\ell = a_{n_\ell}$  ist. Die Teilfolge  $(b_\ell)_{\ell \in \mathbb{N}}$  lässt sich also als  $(a_{n_\ell})_{\ell \in \mathbb{N}}$  schreiben.

Bevor wir uns der vorher besprochenen Charakterisierung von Häufungswerten durch Grenzwerte konvergenter Teilfolgen zuwenden, wollen wir ein kleines Lemma über Teilfolgen *konvergenter* Folgen formulieren.

**LEMMA 32.** Sei  $(a_n)_{n \in \mathbb{N}}$  eine konvergente Folge. Dann gilt für jede Teilfolge  $(a_{n_\ell})_{\ell \in \mathbb{N}}$  von  $(a_n)_{n \in \mathbb{N}}$ , dass  $(a_{n_\ell})_{\ell \in \mathbb{N}}$  auch konvergent ist, und es gilt  $\lim_{\ell \rightarrow \infty} a_{n_\ell} = \lim_{n \rightarrow \infty} a_n$ .

**BEWEIS.** Sei  $x = \lim_{n \rightarrow \infty} a_n$  und  $\varepsilon > 0$ . Dann gibt es ein  $N \in \mathbb{N}$  sodass  $|a_n - x| < \varepsilon$  für alle  $n \geq N$ ; da  $n_\ell$  strikt monoton steigend ist, also insbesondere  $n_\ell \geq \ell$  für alle  $\ell$  gilt, ist für  $\ell \geq N$   $|a_{n_\ell} - x| < \varepsilon$ . Also konvergiert, wie behauptet, auch  $a_{n_\ell} \rightarrow a(n \rightarrow \infty)$ .  $\square$

**Beispiel 8.** Sei  $k \in \mathbb{N}$ , dann ist durch  $n_\ell = k^\ell$  eine strikt monotone Folge natürlicher Zahlen definiert. Da wir schon wissen, dass  $1/n \rightarrow 0(n \rightarrow \infty)$ , folgt aus Lemma 32, dass

$$\frac{1}{k^\ell} \rightarrow 0, \quad (\ell \rightarrow \infty).$$

**SATZ 4.** Sei  $(a_n)_{n \in \mathbb{N}}$  eine Folge. Dann ist ein Punkt  $x \in \mathbb{R}$  genau dann ein Häufungswert von  $(a_n)_{n \in \mathbb{N}}$ , wenn es eine Teilfolge  $(a_{n_\ell})_{\ell \in \mathbb{N}}$  von  $(a_n)_{n \in \mathbb{N}}$  gibt, welche gegen  $x$  konvergiert.

**BEWEIS.** Wir wenden die Charakterisierung des Häufungswerts  $x$  in Lemma 31 an, um induktiv eine Teilfolge  $(a_{n_\ell})_{\ell \in \mathbb{N}}$  zu konstruieren. Dazu wählen wir  $n_1 \geq 1$ , sodass  $|a_{n_1} - x| < 1$  ist. Angenommen, wir haben bereits  $n_{\ell-1}$  gewählt, so wenden wir nun Lemma 31 mit  $N = n_{\ell-1} + 1$  und  $\varepsilon = 1/\ell$  an und erhalten ein  $n_\ell$  für welches

$$|a_{n_\ell} - x| < \frac{1}{\ell}$$

ist.



Wir behaupten nun, dass die Teilfolge  $a_{n_\ell}$  gegen  $x$  konvergiert. Sei dazu  $\varepsilon > 0$  beliebig, und  $N \in \mathbb{N}$  mit  $1/N < \varepsilon$  gewählt. Ist nun  $\ell \geq N$ , so folgt

$$|a_{n_\ell} - x| < \frac{1}{\ell} \leq \frac{1}{N} < \varepsilon,$$

also gilt  $a_{n_\ell} \rightarrow x$  für  $\ell \rightarrow \infty$ . □

**3.2. Monotone Folgen.** Eine Folge  $(a_n)_{n \in \mathbb{N}}$  ist *monoton steigend*, wenn  $a_n \leq a_{n+1}$  für alle  $n \in \mathbb{N}$  ist; sie ist *strikt monoton steigend*, wenn  $a_n < a_{n+1}$  für alle  $n \in \mathbb{N}$  ist. Eine Folge  $(a_n)_{n \in \mathbb{N}}$  ist *monoton fallend*, wenn  $a_n \geq a_{n+1}$  für alle  $n \in \mathbb{N}$  ist; sie ist *strikt monoton fallend*, wenn  $a_n > a_{n+1}$  für alle  $n \in \mathbb{N}$  ist.

Wir sagen, die Folge  $(a_n)_{n \in \mathbb{N}}$  ist *beschränkt*, wenn die Menge  $\{a_n : n \in \mathbb{N}\}$  (d.h. die Menge der Folgenglieder) beschränkt ist.

**SATZ 5.** Sei  $(a_n)_{n \in \mathbb{N}}$  eine beschränkte Folge. Wenn  $(a_n)_{n \in \mathbb{N}}$  monoton steigend ist, dann ist  $a_n$  konvergent, und es gilt

$$\lim_{n \rightarrow \infty} a_n = \sup\{a_n : n \in \mathbb{N}\}.$$

Wenn  $(a_n)_{n \in \mathbb{N}}$  monoton fallend ist, dann ist  $a_n$  konvergent, und es gilt

$$\lim_{n \rightarrow \infty} a_n = \inf\{a_n : n \in \mathbb{N}\}.$$

Es ist im folgenden einfacher,

$$\sup_{n \in \mathbb{N}} a_n := \sup\{a_n : n \in \mathbb{N}\}$$

zu schreiben.

**BEWEIS.** Wir behandeln zunächst den Fall, dass  $(a_n)_{n \in \mathbb{N}}$  monoton steigend ist. Sei  $\varepsilon > 0$ , und  $A := \sup_{n \in \mathbb{N}} a_n$ . Nach der Definition des Supremums gibt es ein  $N \in \mathbb{N}$  sodass  $a_N > A - \varepsilon$ . Ist nun  $n \geq N$ , so ist

$$|a_n - A| = A - a_n \leq A - a_N < \varepsilon.$$

Es folgt, dass  $\lim_{n \rightarrow \infty} a_n = A$  ist.

Ist  $(a_n)_{n \in \mathbb{N}}$  monoton fallend, so ist  $(-a_n)_{n \in \mathbb{N}}$  monoton steigend (und ebenfalls beschränkt). Nach dem ersten Teil ist  $(-a_n)_{n \in \mathbb{N}}$  also konvergent, mit

$$\lim_{n \rightarrow \infty} (-a_n) = \sup_{n \in \mathbb{N}} (-a_n) = - \inf_{n \in \mathbb{N}} a_n.$$

Nach Theorem 1 ist also auch  $a_n$  konvergent, und  $\lim_{n \rightarrow \infty} a_n = \inf_{n \in \mathbb{N}} a_n$ . □

Ist  $a_n$  eine beschränkte Folge, so ist die Folge

$$b_n = \sup_{\ell \geq n} a_\ell$$

monoton fallend, und nach Theorem 5 konvergent. Der Grenzwert

$$\lim_{\ell \rightarrow \infty} b_\ell = \lim_{\ell \rightarrow \infty} \sup_{\ell \leq n} a_n =: \limsup_{n \rightarrow \infty} a_n$$

wird als der *limes superior* von  $a_n$  bezeichnet. Der Grenzwert der monoton steigenden Folge

$$c_n = \inf_{\ell \geq n} a_\ell$$

wird dementsprechend als *limes inferior* von  $(a_n)_{n \in \mathbb{N}}$ , in Zeichen

$$\liminf_{n \rightarrow \infty} a_n := \lim_{\ell \rightarrow \infty} \inf_{\ell \leq n} a_n.$$

**LEMMA 33.** Sei  $(a_n)_{n \in \mathbb{N}}$  eine beschränkte Folge. Dann ist  $A = \limsup_{n \rightarrow \infty} a_n$  genau dann, wenn für jedes  $\varepsilon > 0$  die Menge  $\{n : a_n \geq A + \varepsilon\}$  endlich und die Menge  $\{n : a_n > A - \varepsilon\}$  unendlich ist. Weiters gilt  $B = \liminf_{n \rightarrow \infty} a_n$  genau dann, wenn für jedes  $\varepsilon > 0$  die Menge  $\{n : a_n \leq B - \varepsilon\}$  endlich und die Menge  $\{n : a_n < B + \varepsilon\}$  unendlich ist.

BEWEIS. Wir zeigen zunächst, dass die Bedingung des Lemmas für  $A = \limsup_{n \rightarrow \infty} a_n$  erfüllt ist. Sei also  $\varepsilon > 0$  beliebig. Dann gibt es ein  $N \in \mathbb{N}$ , sodass  $\sup_{\ell \geq k} a_\ell < A + \varepsilon$  wenn  $k \geq N$  ist. Damit ist für alle  $n$  mit  $a_n > A + \varepsilon$  die Ungleichung  $n \leq N$  erfüllt. Weiters ist  $A \leq \sup_{\ell \geq k} a_\ell$  für jedes  $k$ . Wir können deswegen eine monotone Folge  $n_k$  mit  $a_{n_k} > A - \varepsilon$  wie folgt konstruieren: Da  $A - \varepsilon < \sup_{\ell \geq 1} a_\ell$  ist, gibt es ein  $n_1$  mit  $A - \varepsilon < a_{n_1}$ . Angenommen, wir haben  $n_1, \dots, n_{k-1}$  mit  $a_{n_j} > A - \varepsilon$  für  $j = 1, \dots, k-1$  bereits konstruiert, so gibt es wegen

$$A - \varepsilon < \sup_{\ell \geq n_{k-1} + 1} a_\ell$$

ein  $n_k > n_{k-1}$  mit  $a_{n_k} > A - \varepsilon$ . Damit enthält die Menge  $\{n: a_n > A - \varepsilon\}$  die monotone Folge  $n_k$  und ist damit unendlich.

Die Behauptung über den limes inferior folgt durch Anwendung des ersten Teil des Beweises auf die Folge  $(-a_n)_{n \in \mathbb{N}}$ .  $\square$

SATZ 6. Sei  $(a_n)_{n \in \mathbb{N}}$  eine beschränkte Folge. Dann sind  $\limsup_{n \rightarrow \infty} a_n$  und  $\liminf_{n \rightarrow \infty} a_n$  Häufungswerte von  $(a_n)_{n \in \mathbb{N}}$ , und für jeden Häufungswert  $x$  von  $(a_n)_{n \in \mathbb{N}}$  gilt

$$\liminf_{n \rightarrow \infty} a_n \leq x \leq \limsup_{n \rightarrow \infty} a_n.$$

$a_n$  ist konvergent genau dann, wenn  $\liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n$ .

BEWEIS. Sei  $A := \limsup_{n \rightarrow \infty} a_n$ , und  $U \in \mathcal{N}(A)$  eine Umgebung von  $A$ . Wir wählen ein  $\varepsilon > 0$  mit  $B_\varepsilon(A) = (A - \varepsilon, A + \varepsilon) \subset U$ . Nach Lemma 33 ist die Menge  $\{n: a_n > A - \varepsilon\}$  unendlich, die Menge  $\{n: a_n \geq A + \varepsilon\}$  aber endlich; damit ist die Menge  $\{n: a_n \in B_\varepsilon\}$  unendlich, und damit  $A$  ein Häufungspunkt von  $(a_n)_{n \in \mathbb{N}}$ . Ganz analog zeigt man, dass  $B := \liminf_{n \rightarrow \infty} a_n$  ein Häufungswert von  $(a_n)_{n \in \mathbb{N}}$  ist.

Sei nun  $x \in \mathbb{R}$ . Wenn  $x > A$ , so ist nach Lemma 33, angewendet auf  $\varepsilon = \frac{x-A}{2}$  die Menge  $\{n: a_n > A + \varepsilon\}$  endlich. Da  $B_\varepsilon(x) \subset (A + \varepsilon, \infty)$ , ist also auch die Menge  $\{n: a_n \in B_\varepsilon(x)\}$  endlich, also  $x$  kein Häufungswert von  $(a_n)_{n \in \mathbb{N}}$ . Ganz analog zeigt man, dass ein  $x < B$  kein Häufungswert von  $(a_n)_{n \in \mathbb{N}}$  sein kann. Es folgt, dass jeder Häufungswert  $x$  die Ungleichung  $B \leq x \leq A$  erfüllt.  $\square$

**Beispiel 9.** Für ein beliebiges  $r \in \mathbb{Q}$  haben wir, für nichtnegatives  $y$ , bereits die Potenz  $y^r$  definiert. In den Übungen haben wir das Wachstumsverhalten der Funktion  $r \mapsto y^r$  untersucht—diese Funktion ist monoton wachsend, wenn  $y > 1$  ist, konstant, wenn  $y = 1$  oder  $y = 0$ , und monoton fallend für  $0 < y < 1$ . Wir können nun für ein beliebiges  $\alpha \in \mathbb{R}$  eine monotone Folge  $(r_n)_{n \in \mathbb{N}}$  wählen, welche  $\lim_{n \rightarrow \infty} r_n = \alpha$  erfüllt. Die Folge  $(y^{r_n})_{n \in \mathbb{N}}$  ist nun monoton und beschränkt, und wir definieren

$$y^\alpha := \lim_{n \rightarrow \infty} y^{r_n}.$$

Tatsächlich ist dieser Grenzwert von der Wahl der Folge  $r_n$  unabhängig, wir werden dieses Faktum aber jetzt noch nicht beweisen.

**3.3. Der Satz von Bolzano-Weierstrass.** Wir werden nun die Beziehung zwischen Folgen und den topologischen Begriffen aus Unterabschnitt 2.2 herstellen. Unsere ersten zwei Resultate stellen die Beziehungen zwischen Grenzwerten und Häufungspunkten klar. Wenn  $X \subset \mathbb{R}$  eine beliebige Teilmenge von  $\mathbb{R}$  ist, so sagen wir,  $(x_j)_{j \in \mathbb{N}}$  ist eine Folge in  $X$ , wenn  $x_j \in X$  für alle  $j \in \mathbb{N}$  gilt.

LEMMA 34. Sei  $X \subset \mathbb{R}$ , und  $p$  ein Häufungspunkt von  $X$ . Dann gibt es eine konvergente Folge  $(x_n)_{n \in \mathbb{N}}$  in  $X$  mit  $\lim_{n \rightarrow \infty} x_n = p$ .

BEWEIS. Da  $p$  ein Häufungspunkt von  $X$  ist, ist für jedes  $n \in \mathbb{N}$  die Menge  $X \cap (B_{\frac{1}{n}}(p) \setminus \{p\})$  nicht leer. Wir konstruieren uns eine Folge  $(x_j)_{j \in \mathbb{N}}$ , indem wir für jedes  $j \in \mathbb{N}$  ein  $x_j \in X \cap (B_{\frac{1}{j}}(p) \setminus \{p\})$  wählen. Diese Folge konvergiert gegen  $p$ : Wenn  $U \in \mathcal{N}(p)$  eine beliebige Umgebung von  $p$  ist, so gibt es ein  $N \in \mathbb{N}$  mit  $B_{\frac{1}{N}}(p) \subset U$ . Für  $j \geq N$  ist dann  $x_j \in B_{\frac{1}{j}}(p) \subset B_{\frac{1}{N}}(p) \subset U$ .  $\square$

Eine Folge  $(x_n)_{n \in \mathbb{N}}$  heisst *schliesslich konstant*, wenn es ein  $N \in \mathbb{N}$  gibt sodass  $x_j = x_k$  für alle  $j, k \geq N$  gilt.

LEMMA 35. Sei  $(x_n)_{n \in \mathbb{N}}$  eine konvergente Folge, welche nicht schliesslich konstant ist. Dann ist  $p = \lim_{n \rightarrow \infty} x_n$  ein Häufungspunkt von  $X$ .

BEWEIS. Sei  $U \in \mathcal{N}(p)$  eine beliebige Umgebung. Da  $p = \lim_{n \rightarrow \infty} x_n$ , gibt es ein  $N \in \mathbb{N}$ , sodass  $x_n \in U$  für alle  $n \geq N$  gilt. Da  $(x_n)_{n \in \mathbb{N}}$  nicht schliesslich konstant ist, gibt es ein  $n \geq N$  mit  $x_n \neq p$ . Also ist  $x_n \in X \cap (U \setminus \{p\})$ , also  $X \cap (U \setminus \{p\}) \neq \emptyset$ . Da  $U$  beliebig war, folgt, dass  $p$  ein Häufungspunkt von  $X$  ist.  $\square$

Wir können damit den Abschluss  $\bar{X}$  von  $X \subset \mathbb{R}$  (die kleinste abgeschlossene Teilmenge von  $\mathbb{R}$ , welche  $X$  enthält) auch mit Hilfe von Grenzwerten von Folgen in  $X$  beschreiben.  $\bar{X}$  besteht aus allen Grenzwerten von konvergenten Folgen in  $X$  beziehungsweise aus allen Häufungspunkten von Folgen in  $X$ .

PROPOSITION 2. Sei  $X \subset \mathbb{R}$ , dann gilt

$$\bar{X} = \left\{ \lim_{n \rightarrow \infty} x_n : (x_n)_{n \in \mathbb{N}} \text{ ist eine konvergente Folge in } X \right\}$$

und

$$\bar{X} = \{x : x \text{ ist eine Häufungswert einer Folge } (x_n)_{n \in \mathbb{N}} \text{ in } X\}$$

BEWEIS.  $\bar{X}$  besteht nach Lemma 25 aus  $X$  und allen Häufungspunkten von  $X$ . Ist nun  $p \in X$ , so konvergiert die konstante Folge  $(p)_{n \in \mathbb{N}}$  (welche nun eine Folge in  $X$  ist) gegen  $p$ , also tritt jedes  $p \in X$  als Grenzwert einer Folge in  $X$  auf. Ist  $p$  ein Häufungspunkt von  $X$ , so ist  $p$  ein Grenzwert einer konvergenten Folge in  $X$  nach Lemma 34.

Sei nun  $(x_n)_{n \in \mathbb{N}}$  eine konvergente Folge in  $X$ . Wenn  $(x_n)_{n \in \mathbb{N}}$  schliesslich konstant ist, so ist  $\lim_{n \rightarrow \infty} x_n \in X$ ; ist  $(x_n)_{n \in \mathbb{N}}$  nicht schliesslich konstant, so ist  $\lim_{n \rightarrow \infty} x_n$  ein Häufungspunkt von  $X$  nach Lemma 35.

Um die zweite Gleichheit zu zeigen, wenden wir Theorem 4 an, und die soeben gezeigte Gleichheit.  $\square$

Wir werden nun eine wichtige Eigenschaft von Teilmengen von  $\mathbb{R}$  kennenlernen, die *Kompaktheit*. Wir beginnen mit dem folgenden Satz:

SATZ 7 (Satz von Bolzano-Weierstrass). Sei  $X \subset \mathbb{R}$ , dann sind folgende Eigenschaften äquivalent:

(K1)  $X$  ist beschränkt und abgeschlossen.

(K2) Jede Folge  $(x_n)_{n \in \mathbb{N}}$  in  $X$  besitzt eine konvergente Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$  mit  $\lim_{k \rightarrow \infty} x_{n_k} \in X$ .

**Bemerkung 1.** Die Bedingung (K2) nennen wir *Folgenkompaktheit*:  $X$  ist folgenkompakt, wenn es die Bedingung (K2) erfüllt.

BEWEIS. ((K1) $\Rightarrow$ (K2)): Sei  $(x_n)_{n \in \mathbb{N}}$  eine Folge in  $X$ . Da  $X$  beschränkt ist, besitzt  $(x_n)_{n \in \mathbb{N}}$  einen Häufungswert  $x$  (Theorem 6); damit besitzt  $(x_n)_{n \in \mathbb{N}}$  eine Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$  mit  $\lim_{k \rightarrow \infty} x_{n_k} = x$  (Theorem 4), und nach Proposition 2 ist  $x \in X$ . ((K2) $\Rightarrow$ (K1)): Wenn  $X$  nicht beschränkt ist, so gibt es für jedes  $n \in \mathbb{N}$  ein  $x_n \in X$  mit  $|x_n| > n$ . Die Folge  $(x_n)_{n \in \mathbb{N}}$  ist eine unbeschränkte Folge in  $X$ , jede ihrer Teilfolgen ist auch unbeschränkt, und damit nicht konvergent. Wenn  $X$  nicht abgeschlossen ist, so gibt es einen Häufungspunkt  $p$  von  $X$  mit  $p \notin X$ . Nach Proposition 2 gibt es eine konvergente Folge  $(x_n)_{n \in \mathbb{N}}$  in  $X$  mit  $\lim_{n \rightarrow \infty} x_n = p$ . Dann ist aber jede Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$  konvergent (Lemma 32), mit Grenzwert  $\lim_{k \rightarrow \infty} x_{n_k} = p \notin X$ .  $\square$

Wir wollen noch eine weitere Eigenschaft einführen, die zu (K1) und (K2) äquivalent ist. Wir nennen eine Familie  $(U_j)_{j \in J}$ ,  $U_j \subset \mathbb{R}$ , (wo  $J$  eine beliebige Indexmenge ist) eine *offene Überdeckung* von  $X$ , wenn für jedes  $j \in J$  die Menge  $U_j \subset \mathbb{R}$  offen ist und

$$X \subset \bigcup_{j \in J} U_j$$

gilt.  $(U_j)_{j \in J}$  besitzt eine *endliche Teilüberdeckung*, wenn es ein  $N \in \mathbb{N}$  gibt und Indizes  $j_1, \dots, j_N \in J$ , sodass schon

$$X \subset U_{j_1} \cup \dots \cup U_{j_N}$$

gilt.

LEMMA 36. Sei  $(U_j)_{j \in J}$ ,  $U_j \subset \mathbb{R}$  eine unendliche Überdeckung von  $X \subset \mathbb{R}$ . Dann gibt es eine abzählbare Teilüberdeckung von  $(U_j)_{j \in J}$  beinhaltet, d.h. es gibt eine Abbildung  $\varphi: \mathbb{N} \rightarrow J$  gibt mit

$$X \subset \bigcup_{n \in \mathbb{N}} U_{\varphi(n)}.$$

BEWEIS. Zum Beweis des Lemmas betrachten wir die abzählbare Menge

$$\mathcal{B} = \{B_r(q) : q, r \in \mathbb{Q}, r > 0\}.$$

Wir behaupten, dass jede offene Menge  $U$  sich als

$$U = \bigcup_{\substack{B \in \mathcal{B} \\ B \subset U}} B \quad (*)$$

schreiben lässt: Ist  $x \in U$ , so gibt es ein  $\varepsilon > 0$  mit  $B_\varepsilon(x) \subset U$ . Wir wählen ein  $N \in \mathbb{N}$  mit  $1/N < \varepsilon/2$ ; nach Lemma 7 gibt es ein  $q \in \mathbb{Q}$  mit  $q \in B_{\frac{1}{N}}(x)$ . Es folgt, dass  $x \in B_{\frac{1}{N}}(q)$ ; da  $B_{\frac{1}{N}}(q) \in \mathcal{B}$  und  $B_{\frac{1}{N}}(q) \subset B_\varepsilon(x) \subset U$ , ist  $x$  also auch in der Vereinigung auf der rechten Seite von (\*).

Sei  $(U_j)_{j \in J}$  also eine offene Überdeckung von  $X$ . Wir definieren

$$\tilde{\mathcal{B}} = \{B \in \mathcal{B} : B \subset U_j \text{ für ein } j \in J\};$$

$\tilde{\mathcal{B}}$  ist (als Teilmenge einer abzählbaren Menge) wieder abzählbar. Da  $U = \bigcup_{j \in J} U_j$ , haben wir  $U = \bigcup_{B \in \tilde{\mathcal{B}}} B$ . Wir wählen für jedes  $B \in \tilde{\mathcal{B}}$  ein  $j(B) \in J$  mit  $B \subset U_{j(B)}$ , und erhalten

$$X \subset \bigcup_{B \in \tilde{\mathcal{B}}} U_{j(B)}.$$

Es folgt, dass  $(U_{j(B)})_{B \in \tilde{\mathcal{B}}}$  eine abzählbare Teilüberdeckung ist. □

Wir können nun eine zweite Charakterisierung geben.

SATZ 8. Sei  $X \subset \mathbb{R}$ . Dann erfüllt  $X$  die äquivalenten Bedingung (K1) beziehungsweise (K2) genau dann, wenn es die folgende Bedingung erfüllt:

(K3) Jede offene Überdeckung von  $X$  enthält eine endliche Teilüberdeckung.

BEWEIS. ((K2) $\Rightarrow$ (K3)): Angenommen, es gibt eine offene Überdeckung  $(U_j)_{j \in J}$  von  $X$  ohne endliche Teilüberdeckung. Nach Lemma 36 können wir ohne Beschränkung der Allgemeinheit annehmen, dass  $J = \mathbb{N}$  ist. Da es keine endliche Teilüberdeckung von  $(U_j)_{j \in \mathbb{N}}$  gibt, ist  $X \setminus (\bigcup_{1 \leq j \leq n} U_j) \neq \emptyset$  für jedes  $n \in \mathbb{N}$ . Wir wählen für jedes  $n \in \mathbb{N}$  ein  $x_n \in X \setminus (\bigcup_{1 \leq j \leq n} U_j)$ . Die Folge  $(x_n)_{n \in \mathbb{N}}$  besitzt dann keine Häufungswerte in  $X$ : Sei  $p \in X$ , dann gibt es ein  $N$  mit  $p \in U_N$ ; die Menge  $\{n \in \mathbb{N} : x_n \in U_N\}$  ist endlich, da  $x_n \notin U_N$  für  $n \geq N$  ist. Also ist  $p$  kein Häufungswert von  $(x_n)_{n \in \mathbb{N}}$ .

((K3) $\Rightarrow$ (K2)): Sei  $(x_n)_{n \in \mathbb{N}}$  eine Folge in  $X$  ohne Häufungswerte in  $X$ . Dann gibt es für jedes  $y \in X$  ein  $\varepsilon(y)$ , sodass die Menge  $\{x_j : j \in \mathbb{N}\} \cap B_{\varepsilon(y)}(y)$  endlich ist. Die Überdeckung  $(B_{\varepsilon(y)}(y))_{y \in X}$  von  $X$  enthält keine endliche Teilüberdeckung (wenn Sie eine endliche Teilüberdeckung enthalten würde, wäre der Wertebereich von  $(x_n)_{n \in \mathbb{N}}$  endlich; damit hätte  $(x_n)_{n \in \mathbb{N}}$  einen Häufungspunkt). □

Eine nette kleine Eigenschaft von kompakten Mengen sei noch erwähnt:

LEMMA 37. Sei  $K \subset \mathbb{R}$  kompakt. Dann gibt es  $a, b \in K$  mit  $K \subset [a, b]$ ; mit anderen Worten,  $K$  besitzt ein Minimum und ein Maximum.

**3.4. Das Cauchy-Kriterium.** Wir besitzen nach wie vor kein Mittel, die Konvergenz einer Folge festzustellen, ohne ihren Grenzwert zu kennen.

Sei  $(x_n)_{n \in \mathbb{N}}$  eine konvergente Folge,  $x = \lim_{n \rightarrow \infty} x_n$ . Sei  $U \in \mathcal{N}(0)$ ; dann gibt es nach Lemma 28 eine Umgebung  $V \in \mathcal{N}(x)$  mit der Eigenschaft, dass  $V - V \subset U$  ist. Sei  $N$  so gross gewählt, dass  $x_n \in V$  für alle  $n \geq N$  gilt. Sind dann  $k, \ell \geq N$ , so ist

$$x_k - x_\ell \in V - V \subset U.$$

Wir nennen Folgen mit dieser Eigenschaft Cauchyfolgen:

**Definition 8.** Sei  $(x_n)_{n \in \mathbb{N}}$  eine Folge. Wir sagen,  $(x_n)_{n \in \mathbb{N}}$  ist eine Cauchyfolge (oder auch einfach,  $(x_n)_{n \in \mathbb{N}}$  ist Cauchy), wenn für jede Umgebung  $U \in \mathcal{N}(0)$  ein  $N \in \mathbb{N}$  existiert, sodass  $x_k - x_\ell \in U$  für alle  $k, \ell \geq N$  ist.

Wir haben soeben gezeigt, dass jede konvergente Folge eine Cauchyfolge ist. Der grosse Vorteil von Cauchyfolgen ist es, dass man nur die Folge kennen muss, um festzustellen, ob sie eine Cauchyfolge ist. Die Cauchy-Eigenschaft liefert aber Konvergenz der Folge:

**SATZ 9.** Sei  $(x_n)_{n \in \mathbb{N}}$  eine Cauchyfolge. Dann ist  $(x_n)_{n \in \mathbb{N}}$  konvergent.

**BEWEIS.**  $(x_n)_{n \in \mathbb{N}}$  ist beschränkt: Es gibt ein  $N \in \mathbb{N}$  sodass  $x_k - x_\ell \in B_1(0)$  für alle  $k, \ell \geq N$  ist. Insbesondere ist  $|x_k - x_N| \leq 1$  für alle  $k \geq N$ , also nach Lemma 19  $|x_k| \leq 1 + |x_N|$  für alle  $k \geq N$ . Das heisst,  $M = \max\{|x_1|, \dots, |x_{N-1}|, 1 + |x_N|\}$  ist eine Schranke für  $\{x_j : j \in \mathbb{N}\}$ .

Da  $(x_n)_{n \in \mathbb{N}}$  beschränkt ist, gibt es einen Häufungswert  $p$  von  $(x_n)_{n \in \mathbb{N}}$  (Theorem 6). Wir behaupten, dass  $\lim_{n \rightarrow \infty} x_n = p$  gilt. Da  $p$  ein Häufungswert von  $(x_n)_{n \in \mathbb{N}}$  ist, gibt es eine konvergente Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$  von  $(x_n)_{n \in \mathbb{N}}$  mit  $\lim_{k \rightarrow \infty} x_{n_k} = p$ .

Sei nun  $V$  eine beliebige Umgebung von  $p$ . Wir wählen eine Umgebung  $U \in \mathcal{N}(0)$  und eine Umgebung  $W \in \mathcal{N}(p)$  mit  $U + W \subset V$ . Es gibt ein  $N \in \mathbb{N}$  mit  $x_k - x_\ell \in U$  für  $k, \ell \geq N$ , und ein  $K \in \mathbb{N}$  mit  $x_{n_j} \in W$  für alle  $j \geq K$ . Ist nun  $k \geq N$ , und  $n_j \geq N$ , so ist  $x_k = x_k - x_{n_j} + x_{n_j} \in V + W \subset U$ ; also konvergiert  $(x_n)_{n \in \mathbb{N}}$  gegen  $p$ .  $\square$

Eine weitere Umformulierung ist manchmal hilfreich. Unter einer *Nullfolge* verstehen wir eine konvergente Folge, die gegen 0 konvergiert.

**KOROLLAR 1.** Sei  $(x_n)_{n \in \mathbb{N}}$  eine Folge. Dann konvergiert  $(x_n)_{n \in \mathbb{N}}$  genau dann, wenn es eine Nullfolge  $(\varepsilon_n)_{n \in \mathbb{N}}$  gibt, sodass

$$|x_k - x_\ell| \leq \varepsilon_n$$

für alle  $k, \ell \geq n$ , für alle  $n \in \mathbb{N}$  gilt.

**BEWEIS.**  $(\Leftarrow)$  Sei  $U \in \mathcal{N}(0)$ . Dann gibt es ein  $\varepsilon > 0$  sodass  $B_\varepsilon(0) \subset U$  ist; nachdem  $\lim_{n \rightarrow \infty} \varepsilon_n = 0$ , gibt es ein  $N \in \mathbb{N}$  mit  $\varepsilon_n \in B_\varepsilon(0)$  für alle  $n \geq N$ , also gilt  $0 \leq \varepsilon_n \leq \varepsilon$  für alle  $n \geq N$ . Nach Voraussetzung ist nun  $|x_k - x_\ell| \leq \varepsilon_n$  für alle  $k, \ell \geq n$ ; es folgt  $x_k - x_\ell \in B_\varepsilon(0) \subset U$  für alle  $k, \ell \geq N$ . Es folgt, dass  $(x_n)_{n \in \mathbb{N}}$  eine Cauchyfolge ist, also konvergent nach Theorem 9.

$(\Rightarrow)$  Wenn  $(x_n)_{n \in \mathbb{N}}$  konvergiert, so ist  $(x_n)_{n \in \mathbb{N}}$  eine Cauchyfolge nach Theorem 9. Wir betrachten die Folge

$$\varepsilon_n = \sup\{|x_k - x_\ell| : k, \ell \geq n\},$$

und behaupten, dass  $\lim_{n \rightarrow \infty} \varepsilon_n = 0$ . Sei dazu  $U \in \mathcal{N}(0)$ . Es gibt ein  $\varepsilon > 0$  mit  $B_{2\varepsilon}(0) \subset U$ . Da  $(x_n)_{n \in \mathbb{N}}$  eine Cauchyfolge ist, gibt es ein  $N \in \mathbb{N}$  mit  $x_k - x_\ell \in B_\varepsilon(0)$  für alle  $k, \ell \geq N$ , oder  $|x_k - x_\ell| < \varepsilon$  für alle  $k, \ell \geq N$ . Es folgt, dass  $\varepsilon_n = \sup\{|x_k - x_\ell| : k, \ell \geq n\} \leq \varepsilon < 2\varepsilon$ . Es folgt  $\varepsilon_n \in B_{2\varepsilon}(0) \subset U$  für alle  $n \geq N$ . Wir schliessen, dass  $\lim_{n \rightarrow \infty} \varepsilon_n = 0$ .  $\square$

**3.5. Reihen.** Sei  $(a_k)_{k \in \mathbb{N}}$  eine Folge. Wir betrachten die Folge der *Partialsummen*

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

also die Folge  $(S_n)_{n \in \mathbb{N}}$ . Ist die Folge der Partialsummen gegeben, so erhält man die Folge  $a_n$  durch  $a_n = S_n - S_{n-1}$ , für  $n \in \mathbb{N}$ , wo wir  $S_0 = 0$  setzen.

**Definition 9.** Wir nennen die Folge der Partialsummen die *Reihe*, die zu  $(a_k)_{k \in \mathbb{N}}$  gehört, oder auch die aus  $(a_k)_{k \in \mathbb{N}}$  gebildete Reihe. Wir sagen, die Reihe  $\sum_{k=1}^{\infty} a_k$  konvergiert, wenn die Folge  $(S_n)_{n \in \mathbb{N}}$  konvergiert. In diesem Fall schreiben wir

$$\sum_{k=1}^{\infty} a_k := \lim_{n \rightarrow \infty} S_n.$$

Wir nennen die  $a_n$  die Glieder der Reihe.

Oft ist es so, dass die Indexmenge einer Folge (bzw. einer Reihe) sich nicht über  $\mathbb{N} = \{1, 2, \dots\}$  erstreckt, sondern über  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ , oder über  $\{n \in \mathbb{Z} : n \geq k\}$  für ein  $k \in \mathbb{Z}$ . Wir schreiben für solche Folgen/Reihen entsprechend  $(a_n)_{n \geq k}$  und verstehen darunter die Abbildung  $\varphi: \mathbb{N} \rightarrow \mathbb{R}$  welche durch  $\varphi(n) = a_{n+k-1}$  gegeben ist.

**Beispiel 10.** Die aus einer geometrischen Folge  $(q^n)_{n \geq 0}$  gebildete Reihe nennt man geometrische Reihe. Für ihre Partialsummen gilt

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

Ist nun  $|q| < 1$ , so gilt  $\lim_{n \rightarrow \infty} q^{n+1} = 0$ ; also konvergiert in diesem Fall die geometrische Reihe, und es gilt

$$\sum_{j=0}^{\infty} q^j = \lim_{n \rightarrow \infty} \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q}.$$

Die Glieder einer konvergenten Reihe werden klein:

**LEMMA 38.** Sei  $(a_n)_{n \in \mathbb{N}}$  eine Folge. Wenn die Reihe  $\sum_{n=1}^{\infty} a_n$  konvergiert, so ist die Folge  $(a_n)_{n \in \mathbb{N}}$  eine Nullfolge.

**BEWEIS.** Sei  $U$  eine beliebige Umgebung von 0. Nach Definition konvergiert die Folge der Partialsummen

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

ist also insbesondere nach Theorem 9 eine Cauchyfolge; es gibt also ein  $N \in \mathbb{N}$  mit  $S_k - S_\ell \in U$  für alle  $k, \ell \geq N$ . Setzen wir nun  $k = n$  und  $\ell = n - 1$  mit einem  $n \geq N + 1$  so folgt  $a_n = S_n - S_{n-1} \in U$  für alle  $n \geq N + 1$ ; also ist  $(a_n)_{n \in \mathbb{N}}$  eine Nullfolge.  $\square$

Eine Nullfolge zu sein ist nicht ausreichend für die Konvergenz der zugehörigen Reihe.

**Beispiel 11.** Wir betrachten nun die harmonische Reihe, deren Glieder durch die Folge  $1/n$  gegeben sind. Die harmonische Reihe konvergiert *nicht*: Wenn wir eine Partialsumme

$$H_n = \sum_{j=1}^n \frac{1}{j} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \dots + \frac{1}{n},$$

wo  $n = 2^{k+1}$  für ein  $k \in \mathbb{N}$  ist betrachten, so sehen wir, dass

$$\begin{aligned} \frac{1}{2} &\geq \frac{1}{2} \\ \frac{1}{3} + \frac{1}{4} &\geq 2 \cdot \frac{1}{4} = \frac{1}{2} \\ \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} &\geq 4 \cdot \frac{1}{8} = \frac{1}{2} \\ &\dots \\ \sum_{j=2^{2^k}+1}^{2^{2^{k+1}}} \frac{1}{j} &\geq \sum_{j=2^{2^k}+1}^{2^{2^{k+1}}} \frac{1}{2^{2^k+1}} = 2^{2^k} \frac{1}{2^{2^k+1}} = \frac{1}{2} \end{aligned}$$

Es folgt, dass  $H_{2^{k+1}} \geq 1 + 2^{k-1}$  ist.

Einen Spezialfall bilden *alternierende* Reihen, deren Glieder eine monotone Nullfolge bilden:

**SATZ 10.** Sei  $(a_j)_{j \in \mathbb{N}}$  eine monotone Nullfolge. Dann konvergiert die Reihe

$$\sum_{j=1}^{\infty} (-1)^{j-1} a_j.$$

BEWEIS. Wir betrachten die Folge der Partialsummen mit geraden Indizes und die Folge der Partialsummen mit ungeraden Indizes:

$$G_k = \sum_{j=1}^{2k} (-1)^{j-1} a_j$$

$$U_k = \sum_{j=1}^{2k-1} (-1)^{j-1} a_j.$$

Die Folge  $(G_k)_{k \in \mathbb{N}}$  ist monoton wachsend, da

$$G_{k+1} = G_k + a_{2k+1} - a_{2k+2} \geq G_k,$$

und die Folge  $(a_j)_{j \in \mathbb{N}}$  als monoton fallend angenommen war. Die Folge  $(U_k)_{k \in \mathbb{N}}$  ist aus demselben Grund monoton fallend, da

$$U_{k+1} = U_k - a_{2k+2} + a_{2k+1} \leq U_k.$$

Weiters ist  $U_k = G_{k-1} + a_{2k-1} \geq G_{k-1}$ , und  $G_k = U_k - a_{2k} \leq U_k$ , also auch  $U_k \geq G_1$  und  $G_k \leq U_1$ . Die Folgen  $(G_k)_{k \in \mathbb{N}}$  und  $(U_k)_{k \in \mathbb{N}}$  stellen sich damit als beschränkt heraus, und nach Theorem 5 konvergent. Da  $(G_k - U_k) = a_{2k}$ , ist  $\lim_{k \rightarrow \infty} (G_k - U_k) = \lim_{k \rightarrow \infty} a_{2k} = 0$ , also  $\lim_{k \rightarrow \infty} G_k = \lim_{k \rightarrow \infty} U_k =: L$ .

Die Folge der Partialsummen  $(S_n)_{n \in \mathbb{N}}$ , wo

$$S_n = \sum_{j=1}^n (-1)^{j-1} a_j$$

ist damit auch konvergent, mit Grenzwert  $L$ , da sie durch Mischung der Folgen  $U_k$  und  $G_k$  entsteht.  $\square$

**3.6. Reihen mit positiven Gliedern und absolute Konvergenz.** Ist  $(a_j)_{j \in \mathbb{N}}$  eine Folge mit positiven Gliedern, so konvergiert die Reihe  $\sum_{j=1}^{\infty} a_j$  nach dem Monotoniekriterium Theorem 5 genau dann, wenn es eine Konstante  $M > 0$  gibt, sodass

$$\sum_{j=1}^n a_j \leq M,$$

für alle  $n \in \mathbb{N}$ , gilt. Wenn eine Reihe mit positiven Gliedern  $(a_j)_{j \in \mathbb{N}}$  konvergiert, so schreiben wir

$$\sum_{j=1}^{\infty} a_j < \infty;$$

falls die Reihe nicht konvergiert, schreiben wir dementsprechend

$$\sum_{j=1}^{\infty} a_j = \infty.$$

**Definition 10.** Sei  $(a_n)_{n \in \mathbb{N}}$  eine Folge. Wir sagen, die Reihe  $\sum_{j=1}^{\infty} a_n$  konvergiert absolut, wenn

$$\sum_{n=1}^{\infty} |a_n| < \infty.$$

**Satz 11.** Sei  $(a_n)_{n \in \mathbb{N}}$  eine Folge. Wenn  $\sum_{j=1}^{\infty} a_j$  absolut konvergiert, so konvergiert  $\sum_{j=1}^{\infty} a_j$ .

BEWEIS. Wir zeigen, dass die Folge der Partialsummen

$$S_n = \sum_{j=1}^n a_j$$

eine Cauchyfolge ist; eine Anwendung von Theorem 9 liefert dann die Behauptung. Die Folge der Partialsummen der Reihe der Absolutbeträge bezeichnen wir mit

$$A_n = \sum_{j=1}^n |a_j|.$$

Sei  $U \in \mathcal{N}(0)$ , und  $\varepsilon > 0$  mit  $B_\varepsilon(0) \subset U$  gewählt. Da  $\sum_{j=1}^{\infty} |a_j| < \infty$  ist, gibt es ein  $N \in \mathbb{N}$  sodass  $A_k - A_\ell \in B_\varepsilon(0)$  für alle  $\ell, k \geq N$  ist, also für  $\ell \geq k \geq N$  die Ungleichung

$$\sum_{j=k+1}^{\ell} |a_j| = \sum_{j=1}^{\ell} |a_j| - \sum_{j=1}^k |a_j| = \left| \sum_{j=1}^{\ell} |a_j| - \sum_{j=1}^k |a_j| \right| = |A_\ell - A_k| < \varepsilon$$

erfüllt ist.

Es folgt

$$|S_k - S_\ell| = \left| \sum_{j=k+1}^{\ell} a_j \right| \leq \sum_{j=k+1}^{\ell} |a_j| < \varepsilon,$$

für alle  $\ell \geq k \geq N$ , also  $S_k - S_\ell \in B_\varepsilon(0) \subset U$  für alle  $k, \ell \geq N$ .  $\square$

**Beispiel 12.** Die Reihe  $\sum_{j=1}^{\infty} (-1)^{j-1} \frac{1}{j}$  konvergiert nach Theorem 10, konvergiert aber nicht absolut nach Beispiel 11.

Sei nun  $(a_{j,k})_{(j,k) \in \mathbb{N}^2}$  eine *Doppelfolge*. Wir nehmen an, dass einerseits für jedes  $j \in \mathbb{N}$  die Reihe  $A_j = \sum_{k=1}^{\infty} a_{j,k}$  konvergiert, und andererseits, dass für jedes  $k \in \mathbb{N}$  der Grenzwert  $\lim_{j \rightarrow \infty} a_{j,k} = a_k$  existiert. Wir fragen nun, unter welchen Bedingungen die Gleichung

$$\lim_{j \rightarrow \infty} \sum_{k=1}^{\infty} a_{j,k} = \sum_{k=1}^{\infty} \lim_{j \rightarrow \infty} a_{j,k} = \sum_{k=1}^{\infty} a_k$$

erfüllt ist.

**Satz 12.** Sei  $(a_{j,k})_{(j,k) \in \mathbb{N}^2}$  eine positive Doppelfolge, und für jedes  $j$  sei die Folge  $(a_{j,k})_{k \in \mathbb{N}}$  monoton wachsend und konvergent, mit  $\lim_{j \rightarrow \infty} a_{j,k} =: a_k$ . Dann gilt, dass die Reihe  $\sum_{k=1}^{\infty} a_k$  genau dann konvergiert, wenn es ein  $M > 0$  gibt, sodass

$$\sum_{k=1}^{\infty} a_{j,k} < M$$

für alle  $j \in \mathbb{N}$  gilt. In diesem Fall ist

$$\lim_{j \rightarrow \infty} \sum_{k=1}^{\infty} a_{j,k} = \sum_{k=1}^{\infty} \lim_{j \rightarrow \infty} a_{j,k}.$$

**BEWEIS.** Wenn  $\sum_{k=1}^{\infty} a_k < \infty$  gilt, so gibt es ein  $M > 0$  mit  $\sum_{k=1}^n a_k < M$  für alle  $n \in \mathbb{N}$ . Dann ist auch

$$\sum_{k=1}^n a_{j,k} \leq \sum_{k=1}^n a_k \leq M,$$

für alle  $j, n \in \mathbb{N}$ , und damit auch

$$\text{sum}_{k=1}^{\infty} a_{j,k} \leq M$$

für alle  $j \in \mathbb{N}$ .

Sei nun  $\sum_{k=1}^{\infty} a_{j,k} < M$  für alle  $j \in \mathbb{N}$ . Dann ist auch

$$\sum_{k=1}^n a_{j,k} \leq M$$

für alle  $j, n \in \mathbb{N}$ ; nach Grenzübergang  $j \rightarrow \infty$  also

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

und es folgt

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



Die Folge  $A_j = \sum_{k=1}^{\infty} a_{j,k}$  ist monoton steigend, und beschränkt, also existiert  $\lim_j A_j =: A$ . Es bleibt noch zu zeigen, dass  $A = \sum_{k=1}^{\infty} a_k$ . Die Ungleichung  $A \leq \sum_{k=1}^{\infty} a_k$  folgt, da  $A_j \leq \sum_{k=1}^{\infty} a_k$  für alle  $k \in \mathbb{N}$  gilt. Sei nun  $\varepsilon > 0$  beliebig. Es gibt ein  $N \in \mathbb{N}$ , sodass  $\sum_{k=1}^N a_k > \sum_{k=1}^{\infty} a_k - \frac{\varepsilon}{2}$  ist. Es gibt ein  $J \in \mathbb{N}$ , sodass

$$a_{j,k} > a_k - \frac{\varepsilon}{2N}, \quad j \geq J, \quad k \leq N,$$

ist.

Es folgt, dass

$$\begin{aligned} \sum_{k=1}^{\infty} a_{j,k} &\geq \sum_{k=1}^N a_{j,k} \\ &> \sum_{k=1}^N \left( a_k - \frac{\varepsilon}{2N} \right) \\ &= \sum_{k=1}^N a_k - \frac{\varepsilon}{2} \\ &> \sum_{k=1}^{\infty} a_k - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} = \sum_{k=1}^{\infty} a_k - \varepsilon, \end{aligned}$$

für alle  $j \in \mathbb{N}$  mit  $j \geq J$  gilt; nach Grenzübergang  $j \rightarrow \infty$  also  $A > \sum_{k=1}^{\infty} a_k - \varepsilon$ . Nachdem  $\varepsilon > 0$  beliebig war, folgt  $A \geq \sum_{k=1}^{\infty} a_k$ , und damit die gewünschte Gleichheit  $A = \sum_{k=1}^{\infty} a_k$ .  $\square$

**3.7. Rechnen mit Reihen.** Wenn  $(a_j)_{j \in \mathbb{N}}$  und  $(b_j)_{j \in \mathbb{N}}$  Folgen sind, und die daraus gebildeten Reihen konvergieren, so konvergiert wegen

$$\sum_{j=1}^n (a_j + \lambda b_j) = \left( \sum_{j=1}^n a_j \right) + \lambda \left( \sum_{j=1}^n b_j \right)$$

nach einer Anwendung von Theorem 1 auch die zu  $(a_j + \lambda b_j)_{j \in \mathbb{N}}$  gehörende Reihe, und es gilt

$$\sum_{j=1}^{\infty} (a_j + \lambda b_j) = \left( \sum_{j=1}^{\infty} a_j \right) + \lambda \left( \sum_{j=1}^{\infty} b_j \right).$$

Um auf diese Regel verweisen zu können, formulieren wir das noch einmal als

**SATZ 13.** *Seien  $(a_j)_{j \in \mathbb{N}}$  und  $(b_j)_{j \in \mathbb{N}}$  Folgen, und  $\lambda \in \mathbb{R}$ . Wenn die Reihen*

$$\sum_{j=1}^{\infty} a_j \quad \text{und} \quad \sum_{j=1}^{\infty} b_j$$

*konvergieren, so konvergiert auch*

$$\sum_{j=1}^{\infty} (a_j + \lambda b_j),$$

*und es gilt*

$$\sum_{j=1}^{\infty} (a_j + \lambda b_j) = \left( \sum_{j=1}^{\infty} a_j \right) + \lambda \left( \sum_{j=1}^{\infty} b_j \right).$$

Die Multiplikation von Reihen ist leider nicht ganz so unkompliziert: Wenn man

$$\left( \sum_{j=1}^{\infty} a_j \right) \left( \sum_{k=1}^{\infty} b_k \right)$$

zu bilden versucht, erhält man als Objekt

$$\sum_{j,k=1}^{\infty} a_j b_k;$$

wie man diesem einen Wert zuordnen kann, wissen wir (noch) nicht. Eine sinnvolle Definition ist zum Beispiel, dass

$$\sum_{j,k=1}^{\infty} a_j b_k = L$$

ist, wenn für jede Umgebung  $U \in \mathcal{N}(L)$  eine endliche Menge  $F \subset \mathbb{N}^2$  existiert, sodass

$$\sum_{(j,k) \in \tilde{F}} a_j b_k \in U$$

für alle endlichen Mengen  $\tilde{F} \supseteq F$  gilt. Man kann allerdings auch andere gerichtete Konvergenzbegriffe definieren. Wir umgehen dieses Problem, indem wir zunächst spezielle Summierungen betrachten, d.h. je endlich viele der Produkte  $a_j b_k$  zu gruppieren.

Ein erstes Beispiel ist die direkte Anwendung von Theorem 2: Wenn die Folgen  $(A_n)_{n \in \mathbb{N}}$  und  $(B_n)_{n \in \mathbb{N}}$  Partialsummen, d.h.

$$A_n = \sum_{j=1}^n a_j, \quad B_n = \sum_{k=1}^n b_k,$$

konvergieren, so konvergiert auch die Folge  $(A_n B_n)_{n \in \mathbb{N}}$ ; diese Folge ist die Folge der Partialsummen der Reihe

$$\sum_{\ell=1}^{\infty} \left( \sum_{j=1}^{\ell} a_j b_{\ell-j} + \sum_{k=1}^{\ell} a_{\ell-k} b_k - a_{\ell} b_{\ell} \right).$$

Für viele Zwecke ist eine andere Art der Summierung zu bevorzugen.

**Definition 11.** Seien  $(a_j)_{j \geq 0}$  und  $(b_k)_{k \geq 0}$  Folgen. Das Cauchyprodukt der Reihen  $\sum_{j=0}^{\infty} a_j$  und  $\sum_{k=0}^{\infty} b_k$  ist die Reihe

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

**Satz 14.** Seien  $(a_j)_{j \in \mathbb{N}}$  und  $(b_j)_{j \in \mathbb{N}}$  Folgen. Wenn  $\sum_{j=1}^{\infty} a_j$  konvergiert und  $\sum_{k=1}^{\infty} b_k$  absolut konvergiert, so konvergiert auch das Cauchyprodukt von  $\sum_{j=0}^{\infty} a_j$  und  $\sum_{k=0}^{\infty} b_k$ , und es gilt

$$\sum_{\ell=0}^{\infty} \left( \sum_{j=0}^{\ell} a_j b_{\ell-j} \right) = \left( \sum_{j=1}^{\infty} a_j \right) \left( \sum_{k=0}^{\infty} b_k \right).$$

**BEWEIS.** Wie oben schreiben wir

$$A_n = \sum_{j=0}^n a_j, \quad B_n = \sum_{k=0}^n b_k,$$

für die Folgen der Partialsummen, und setzen  $A = \lim_{n \rightarrow \infty} A_n$ ,  $B = \lim_{n \rightarrow \infty} B_n$ . Da  $\sum_{k=1}^{\infty} b_k$  absolut konvergiert und  $\sum_{j=1}^{\infty} a_j$  konvergiert, können wir ein  $M > 0$  mit  $\sum_{k=0}^n |b_k| < M$  und  $|\sum_{j=k}^{\ell} a_j| \leq M$  für alle  $k, \ell \in \mathbb{N}$  wählen.

Sei  $U \in \mathcal{N}(AB)$ , und  $\varepsilon > 0$  mit  $B_{\varepsilon}(AB) \subset U$  gewählt. Es gibt ein  $N \in \mathbb{N}$  sodass für alle  $n \geq N$  die Glieder der Produktfolge  $A_n B_n \in B_{\varepsilon/2}(AB)$  erfüllen, also  $|A_n B_n - AB| < \varepsilon/2$  für alle  $n \geq N$  gilt. Wenn wir  $N$  gross genug wählen, können wir auch für alle  $k, \ell \geq N$  die Abschätzung  $|A_k - A_{\ell}| \leq \frac{\varepsilon}{4M}$  fordern (da

$(A_n)_{n \in \mathbb{N}}$  konvergiert, also insbesondere eine Cauchyfolge ist). Da  $\lim_{k \rightarrow \infty} b_k = 0$  (Lemma 38), können wir ein  $\tilde{N} > 0$  wählen, sodass  $|b_k| \leq \frac{\varepsilon}{4(N-1)M}$  gilt. Wir schätzen ab:

$$\left| \sum_{\ell=0}^n \left( \sum_{j=0}^{\ell} a_j b_{\ell-j} \right) - AB \right| \leq \left| \sum_{\ell=0}^n \left( \sum_{j=0}^{\ell} a_j b_{\ell-j} \right) - A_n B_n \right| + |A_n B_n - AB|$$

und betrachten den ersten Summanden auf der rechten Seite der Ungleichung.

$$\begin{aligned} \left| \sum_{\ell=0}^n \left( \sum_{j=0}^{\ell} a_j b_{\ell-j} \right) - A_n B_n \right| &= \left| \sum_{\ell=n+1}^{2n} \left( \sum_{j=\ell-n}^n a_j b_{\ell-j} \right) \right| \\ &= \left| \sum_{k=0}^n \sum_{j=n-k+1}^n a_j b_k \right| \\ &\leq \left| \sum_{k=0}^n \sum_{j=\max(n-k+1, N+1)}^n a_j b_k \right| + \left| \sum_{k=n-N+1}^n \sum_{j=0}^N a_j b_k \right| \\ &\leq \sum_{k=0}^N \left( |b_k| \left| \sum_{j=\max(n-k+1, N+1)}^n a_j \right| \right) + \sum_{k=n-N+1}^n \left( |b_k| \left| \sum_{j=n-k+1}^n a_j \right| \right) \\ &\leq \sum_{k=0}^N \left( |b_k| \frac{\varepsilon}{4M} \right) + \sum_{k=n-N+1}^n |b_k| M \\ &\leq M \frac{\varepsilon}{4M} + (N-1)M \frac{\varepsilon}{4(N-1)M} = \frac{\varepsilon}{2}, \end{aligned}$$

wenn  $n \geq N + \tilde{N}$ . □

Ein bisschen einfacher gestaltet sich die Produktbildung von Reihen, wenn tatsächlich beide Reihen absolut konvergieren. Um die Situation zu durchleuchten, betrachten wir zunächst das Thema der Umordnungen.

### 3.8. Reihenummordnungen.

**Definition 12.** Sei  $(a_j)_{j \in \mathbb{N}}$  eine Folge, und  $\varphi: \mathbb{N} \rightarrow \mathbb{N}$  eine bijektive Funktion. Wir nennen die Reihe  $\sum_{k=1}^{\infty} a_{\varphi(k)}$  eine Umordnung der Reihe  $\sum_{j=1}^{\infty} a_j$ .

**PROPOSITION 3.** Sei  $\sum_{k=1}^{\infty} a_{\varphi(k)}$  eine Umordnung von  $\sum_{j=1}^{\infty} a_j$ . Wenn  $\sum_{j=1}^{\infty} a_j$  absolut konvergiert, so konvergiert  $\sum_{k=1}^{\infty} a_{\varphi(k)}$  absolut, und es gilt

$$\sum_{k=1}^{\infty} a_{\varphi(k)} = \sum_{j=1}^{\infty} a_j.$$

**BEWEIS.** Wenn  $\sum_{j=1}^{\infty} a_j$  absolut konvergiert, so gibt es eine Konstante  $M > 0$ , sodass

$$\sum_{j=1}^N |a_j| \leq M$$

für alle  $N \in \mathbb{N}$  gilt. Damit ist auch

$$\sum_{k=1}^K |a_{\varphi(k)}| \leq \sum_{j=1}^N |a_j| \leq M,$$

wo wir  $N = \max\{\varphi(k) : 1 \leq k \leq K\}$  wählen. Damit ist gezeigt, dass auch  $\sum_{k=1}^{\infty} a_{\varphi(k)}$  absolut konvergiert.

Wir bezeichnen mit

$$S_n = \sum_{j=1}^n a_j, \quad T_n = \sum_{k=1}^n a_{\varphi(k)}$$

die Partialsummen der ursprünglichen und der umgeordneten Reihe. Sei  $U \in \mathcal{N}(0)$  beliebig, und  $\varepsilon > 0$  mit  $B_\varepsilon(0) \subset U$  gewählt. Da  $\sum_{j=1}^{\infty} a_j$  absolut konvergiert, gibt es ein  $\mathbb{N}$  sodass für  $\ell \geq m \geq N$  die Ungleichung

$$\sum_{j=m}^{\ell} |a_j| < \varepsilon$$

gilt.

Wir betrachten die Differenz

$$S_n - T_n = \sum_{j=1}^n a_j - \sum_{k=1}^n a_{\varphi(k)}.$$

Wenn wir  $n \geq K = \max\{k: \varphi(k) \leq N\}$  wählen, so heben sich sämtliche Terme mit Indizes  $j \leq N$  in dieser Summe auf. Die Terme mit höheren Indizes können wir wie folgt abschätzen:

$$|S_n - T_n| \leq \sum_{j=N}^m |a_j| < \varepsilon$$

wo wir  $m = \max\{\varphi(k): k \leq N\}$  (oder einfach genügend gross) wählen. Es folgt, dass  $S_n - T_n \in U$  für  $n \geq K$  ist, und wir schliessen, dass  $\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} T_n$  ist, also

$$\sum_{k=1}^{\infty} a_{\varphi(k)} = \sum_{j=1}^{\infty} a_j.$$

□

Eine kleine Modifikation des obigen Beweises erlaubt es uns, auch “Klammersetzungen” zu betrachten. Dazu betrachten wir eine Funktion  $\varphi: \mathbb{N} \rightarrow \mathcal{P}(\mathbb{N})$ , wo  $\varphi(k) \subset \mathbb{N}$  eine endliche Teilmenge ist, und die  $(\varphi(k))_{k \in \mathbb{N}}$  eine Partition von  $\mathbb{N}$  bilden, also

$$\bigcup_{k \in \mathbb{N}} \varphi(k) = \mathbb{N}, \quad \text{und } \varphi(k) \cap \varphi(\ell) = \emptyset \text{ falls } \ell \neq k$$

gilt. Wir nennen die Reihe

$$\sum_{k=1}^{\infty} \left( \sum_{j \in \varphi(k)} a_j \right)$$

eine geklammerte Umordnung der Reihe

$$\sum_{j=1}^{\infty} a_j.$$

**PROPOSITION 4.** Sei  $\sum_{k=1}^{\infty} \left( \sum_{j \in \varphi(k)} a_j \right)$  eine geklammerte Umordnung von  $\sum_{j=1}^{\infty} a_j$ . Wenn  $\sum_{j=1}^{\infty} a_j$  absolut konvergiert, so konvergiert auch

$$\sum_{k=1}^{\infty} \left( \sum_{j \in \varphi(k)} a_j \right)$$

absolut, und es gilt

$$\sum_{k=1}^{\infty} \left( \sum_{j \in \varphi(k)} a_j \right) = \sum_{j=1}^{\infty} a_j.$$

**BEWEIS.** Wenn  $\sum_{j=1}^{\infty} a_j$  absolut konvergiert, so gibt es eine Konstante  $M > 0$ , sodass

$$\sum_{j=1}^N |a_j| \leq M$$

für alle  $N \in \mathbb{N}$  gilt. Damit ist auch

$$\sum_{k=1}^K \left| \sum_{j \in \varphi(k)} a_j \right| \leq \sum_{k=1}^K \sum_{j \in \varphi(k)} |a_j| \leq \sum_{j=1}^N |a_j| \leq M,$$

wo wir

$$N = \max_{k \leq K} \bigcup \varphi(k)$$

wählen. Damit ist gezeigt, dass auch  $\sum_{k=1}^{\infty} a_{\varphi(k)}$  absolut konvergiert.

Wir bezeichnen mit

$$S_n = \sum_{j=1}^n a_j, \quad T_n = \sum_{k=1}^n \left( \sum_{j \in \varphi(k)} a_j \right)$$

die Partialsummen der ursprünglichen und der geklammert ungeordneten Reihe. Sei  $U \in \mathcal{N}(0)$  beliebig, und  $\varepsilon > 0$  mit  $B_\varepsilon(0) \subset U$  gewählt. Da  $\sum_{j=1}^{\infty} a_j$  absolut konvergiert, gibt es ein  $\mathbb{N}$  sodass für  $\ell \geq m \geq N$  die Ungleichung

$$\sum_{j=m}^{\ell} |a_j| < \varepsilon$$

gilt.

Wir betrachten die Differenz

$$S_n - T_n = \sum_{j=1}^n a_j - \sum_{k=1}^n a_{\varphi(k)}.$$

Für jedes  $j \in \mathbb{N}$  gibt es ein (und nur ein)  $k(j)$  mit  $j \in \varphi(k(j))$ . Wenn wir

$$n \geq K = \max_{j=1}^N \varphi(k(j))$$

wählen, so heben sich sämtliche Terme mit Indizes  $j \leq N$  in der Differenz  $S_n - T_n$  auf. Die Terme mit höheren Indizes können wir wie folgt abschätzen:

$$|S_n - T_n| \leq \sum_{j=N}^m |a_j| < \varepsilon$$

wo wir

$$m = \max_{k=1}^N \varphi(k)$$

(oder einfach genügend gross) wählen. Es folgt, dass  $S_n - T_n \in U$  für  $n \geq K$  ist, und wir schliessen, dass  $\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} T_n$  ist, also

$$\sum_{k=1}^{\infty} a_{\varphi(k)} = \sum_{j=1}^{\infty} a_j.$$

□

Wir erinnern daran, dass wir eine Menge  $\Lambda$  als *abzählbar* bezeichnen, wenn es eine bijektive Abbildung  $\varphi: \mathbb{N} \rightarrow \Lambda$  gibt.

**Definition 13.** Sei  $\Lambda$  eine abzählbare Menge, und  $\lambda \mapsto a_\lambda$  eine Abbildung  $\Lambda \rightarrow \mathbb{R}$ . Wir sagen, die Reihe

$$\sum_{\lambda \in \Lambda} a_\lambda$$

konvergiert absolut, wenn es eine Konstante  $M$  gibt, sodass

$$\sum_{\lambda \in F} |a_\lambda| < M$$

für alle endlichen Teilmengen  $F \subset \Lambda$  gilt.

SATZ 15. Sei  $\Lambda$  eine abzählbare Menge, und

$$\sum_{\lambda \in \Lambda} a_\lambda$$

absolut konvergent. Dann gibt es eine Zahl  $L \in \mathbb{R}$  mit der Eigenschaft, dass für jede Abzählung (d.h. bijektive Abbildung)  $\varphi: \mathbb{N} \rightarrow \Lambda$

$$L = \sum_{j=1}^{\infty} a_{\varphi(j)}$$

gilt.

Wir bezeichnen die von der Abzählung unabhängige Zahl  $L$  in Theorem 15 als

$$L =: \sum_{\lambda \in \Lambda} a_\lambda.$$

BEWEIS. Sei  $\psi: \mathbb{N} \rightarrow \Lambda$  eine beliebige Abzählung von  $\Lambda$ . Die Reihe

$$\sum_{j=1}^{\infty} a_{\psi(j)}$$

konvergiert absolut; wir bezeichnen ihren Wert mit  $L$ . Ist  $\tilde{\psi}$  eine andere Abzählung von  $\Lambda$ , so gibt es eine bijektive Abbildung  $\varphi: \mathbb{N} \rightarrow \mathbb{N}$  mit  $\psi \circ \varphi = \tilde{\psi}$ ; also geht die Reihe

$$\sum_{k=1}^{\infty} a_{\tilde{\psi}(k)}$$

aus der ursprünglichen Reihe

$$\sum_{j=1}^{\infty} a_{\psi(j)}$$

durch Umordnung hervor. Nach Proposition 3 ist also

$$\sum_{k=1}^{\infty} a_{\tilde{\psi}(k)} = L.$$

□

Sei nun  $(\Lambda_\mu)_{\mu \in \Sigma}$  eine Partition von  $\Lambda$  in endliche Teilmengen  $\Lambda_\mu$ , indiziert durch die (notwendigerweise wieder abzählbare) Menge  $\Sigma$ . In diesem Fall liefert derselbe Beweis wie oben, unter Anwendung von Proposition 4, folgende Aussage:

SATZ 16. Sei

$$\sum_{\lambda \in \Lambda} a_\lambda,$$

und  $(\Lambda_\mu)_{\mu \in \Sigma}$  eine Partition von  $\Lambda$  in endliche Teilmengen  $\Lambda_\mu$ . Wenn

$$\sum_{\lambda \in \Lambda} a_\lambda$$

absolut konvergiert, so konvergiert auch

$$\sum_{\mu \in \Sigma} \left( \sum_{\lambda \in \Lambda_\mu} a_\lambda \right)$$

absolut, und es gilt

$$\sum_{\mu \in \Sigma} \left( \sum_{\lambda \in \Lambda_\mu} a_\lambda \right) = \sum_{\lambda \in \Lambda} a_\lambda.$$

Der vorangehende Satz gilt sogar für beliebige Partitionen  $(\Lambda_\mu)_{\mu \in \Sigma}$  von  $\Lambda$ , d.h. man kann die Beschränkung auf endliche Teilmengen wegfällen lassen.

Zunächst allerdings bemerken wir, dass wir die vorangehenden Betrachtungen auch auf die Produktreihe zweier absolut konvergenter Reihen anwenden können; die Reihe

$$\sum_{(j,k) \in \mathbb{N}^2} a_j b_k$$

konvergiert absolut, wenn die Reihen  $\sum_{j=1}^{\infty} a_j$  und  $\sum_{k=1}^{\infty} b_k$  absolut konvergieren, und es gilt

$$\sum_{(j,k) \in \mathbb{N}^2} a_j b_k = \sum_{\ell=1}^{\infty} \left( \sum_{k=1}^{\ell-1} a_k b_{k-\ell} \right) = \left( \sum_{j=1}^{\infty} a_j \right) \left( \sum_{k=1}^{\infty} b_k \right).$$

Die letzte Gleichheit folgt dabei aus Theorem 14, kann aber auch aus der bereits versprochenen Konvergenzaussage für beliebige Partitionen des Indexbereichs hergeleitet werden.

PROPOSITION 5. *Sei*

$$\sum_{(j,k) \in \mathbb{N}^2} a_{j,k}$$

eine absolut konvergente Reihe. Dann konvergiert für jedes  $k \in \mathbb{N}$  die Reihe

$$\sum_{j=1}^{\infty} a_{j,k},$$

für jedes  $j \in \mathbb{N}$  die Reihe

$$\sum_{k=1}^{\infty} a_{j,k}$$

absolut, und die Reihen

$$\sum_{k=1}^{\infty} \left( \sum_{j=1}^{\infty} a_{j,k} \right), \quad \sum_{j=1}^{\infty} \left( \sum_{k=1}^{\infty} a_{j,k} \right)$$

konvergieren absolut. Für die Summen gilt

$$\sum_{k=1}^{\infty} \left( \sum_{j=1}^{\infty} a_{j,k} \right) = \sum_{j=1}^{\infty} \left( \sum_{k=1}^{\infty} a_{j,k} \right) = \sum_{(j,k) \in \mathbb{N}^2} a_{j,k}.$$

BEWEIS. Da  $\sum_{(j,k) \in \mathbb{N}^2} a_{j,k}$  absolut konvergiert, gibt es ein  $M > 0$  mit

$$\sum_{(j,k) \in F} |a_{j,k}| < M$$

für beliebige endliche Teilmengen  $F \subset \mathbb{N}^2$ . Dann gilt für jedes  $k \in \mathbb{N}$  und für beliebiges  $N \in \mathbb{N}$ , dass

$$\sum_{j=1}^N |a_{j,k}| < M.$$

Wir schliessen, dass die Reihe

$$\sum_{j=1}^{\infty} a_{j,k}$$

absolut konvergiert.

Weiters gilt für beliebiges  $N \in \mathbb{N}$ , dass

$$\left| \sum_{j=1}^N a_{j,k} \right| \leq \sum_{j=1}^N |a_{j,k}|,$$

also für beliebiges  $K \in \mathbb{N}$

$$\sum_{k=1}^K \left| \sum_{j=1}^N a_{j,k} \right| \leq \sum_{k=1}^K \sum_{j=1}^N |a_{j,k}| < M,$$

und nach Grenzübergang  $N \rightarrow \infty$

$$\sum_{k=1}^K \left| \sum_{j=1}^{\infty} a_{j,k} \right| \leq \sum_{k=1}^K \sum_{j=1}^{\infty} |a_{j,k}| \leq M.$$

Es folgt, dass

$$\sum_{k=1}^{\infty} \left( \sum_{j=1}^{\infty} a_{j,k} \right)$$

absolut konvergiert.

Wir zeigen nun noch, dass

$$\sum_{k=1}^{\infty} \left( \sum_{j=1}^{\infty} a_{j,k} \right) = \sum_{\lambda \in \mathbb{N}^2} a_{\lambda}$$

gilt. Sei dafür eine beliebige Abzählung  $\ell \mapsto \lambda_{\ell}$  von  $\mathbb{N}^2$  gegeben. Wir schreiben  $S = \sum_{k=1}^{\infty} \left( \sum_{j=1}^{\infty} a_{j,k} \right)$ , und müssen zeigen, dass

$$S = \lim_{n \rightarrow \infty} \sum_{\ell=1}^n a_{\lambda_{\ell}}$$

gilt. Sei also  $U \in \mathcal{N}(S)$  eine beliebige Umgebung von  $S$ . Wir wählen ein  $\varepsilon > 0$  mit  $B_{3\varepsilon}(S) \subset U$ .

Nachdem die Reihe  $\sum_{\ell=1}^n a_{\lambda_{\ell}}$  absolut konvergiert, gibt es ein  $N \in \mathbb{N}$ , sodass  $\sum_{\ell=N+1}^m |a_{\lambda_{\ell}}| < \varepsilon$  für jedes  $m \geq N$ . Wir betrachten die endliche Teilmenge  $\Lambda_0 = \{\lambda_1, \dots, \lambda_N\} \subset \mathbb{N}^2$  und wählen  $K_0 \in \mathbb{N}$  und  $J_0 \in \mathbb{N}$  so gross, dass  $\Lambda_0 \subset \{1, \dots, J_0\} \times \{1, \dots, K_0\}$  ist.

Da  $\sum_{k=1}^{\infty} \left( \sum_{j=1}^{\infty} a_{j,k} \right) = S$  konvergiert, gibt es ein  $K_1 \geq K_0$  sodass die Ungleichung

$$\left| S - \sum_{k=1}^{K_1} \left( \sum_{j=1}^{\infty} a_{j,k} \right) \right| < \varepsilon$$

erfüllt ist. Weiters können wir ein  $L_1 \geq L_0$  finden, sodass

$$\left| \sum_{j=1}^{\infty} a_{j,k} - \sum_{j=1}^{L_1} a_{j,k} \right| \leq \frac{\varepsilon}{K_1}, \quad k = 1, \dots, K_1$$

gilt.

Somit gilt für  $n \geq N$ , dass

$$\begin{aligned} \left| S - \sum_{\ell=1}^n a_{\lambda_{\ell}} \right| &\leq \left| S - \sum_{k=1}^{K_1} \left( \sum_{j=1}^{\infty} a_{j,k} \right) \right| + \left| \sum_{k=1}^{K_1} \left( \sum_{j=1}^{\infty} a_{j,k} - \sum_{j=1}^{L_1} a_{j,k} \right) \right| + \left| \sum_{k=1}^{K_1} \sum_{j=1}^{L_1} a_{j,k} - \sum_{\ell=1}^n a_{\lambda_{\ell}} \right| \\ &< \varepsilon + \sum_{k=1}^{K_1} \left| \sum_{j=1}^{\infty} a_{j,k} - \sum_{j=1}^{L_1} a_{j,k} \right| + \varepsilon \\ &< \varepsilon + K_1 \frac{\varepsilon}{K_1} + \varepsilon = 3\varepsilon, \end{aligned}$$

also  $\sum_{\ell=1}^n a_{\lambda_{\ell}} \in U$  für alle  $n \geq N$ . □



### 3.9. Das Minoranten- und das Majorantenkriterium.

SATZ 17. Sei  $\sum_{k=1}^{\infty} \varepsilon_k$  eine absolut konvergente Reihe, mit nichtnegativen Gliedern  $\varepsilon_k \geq 0$ , und  $(a_n)_{n \in \mathbb{N}}$  eine Folge. Mit  $(S_n)_{n \in \mathbb{N}}$  bezeichnen wir die Folge der Partialsummen der Reihe, die aus der Folge  $(|a_n|)_{n \in \mathbb{N}}$  gebildet wird. Wenn es eine strikt monotone Folge  $(n_k)_{k \in \mathbb{N}}$  natürlicher Zahlen und ein  $K \in \mathbb{N}$  gibt, sodass für  $k \geq K$  die Abschätzung

$$S_{n_{k+1}} - S_{n_k} = \sum_{n=n_k+1}^{n_{k+1}} |a_n| \leq \varepsilon_k$$

erfüllt ist, so konvergiert die Reihe  $\sum_{n=1}^{\infty} a_n$  absolut.

Gilt insbesondere  $|a_n| \leq \varepsilon_n$ , so ist  $\sum_{n=1}^{\infty} a_n$  absolut konvergent.

**Beispiel 13.**  $\sum_{n=1}^{\infty} \frac{1}{n^{1+\varepsilon}} < \infty$  für  $\varepsilon > 0$ .

SATZ 18. Sei  $\sum_{k=1}^{\infty} \varepsilon_k$  eine Reihe mit nichtnegativen Gliedern  $\varepsilon_k \geq 0$ , welche

$$\sum_{n=1}^{\infty} \varepsilon_n = \infty$$

und  $(a_n)_{n \in \mathbb{N}}$  eine Folge. Wenn es ein  $N \in \mathbb{N}$  gibt, sodass für alle  $n \geq N$

$$|a_n| \geq \varepsilon_n$$

erfüllt ist, so konvergiert die Reihe  $\sum_{n=1}^{\infty} a_n$  nicht absolut.

### 3.10. Das Wurzel- und das Quotientenkriterium.

SATZ 19. Sei  $(a_n)_{n \in \mathbb{N}}$  eine Folge. Wenn

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

ist, so konvergiert

$$\sum_{n=1}^{\infty} a_n$$

absolut.

BEWEIS. Wir wählen ein  $q$  mit  $\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} < q < 1$ . Nach Definition des  $\limsup$  gibt es ein  $N \in \mathbb{N}$  sodass für alle  $n \geq N$  die Ungleichung  $\sqrt[n]{|a_n|} < q$ , also  $|a_n| < q^n$  erfüllt ist. Wir können also das Majorantenkriterium für die Reihe  $\sum_{n \geq N} a_n$  anwenden, da die geometrische Reihe  $\sum_{n=N}^{\infty} q^n$  absolut konvergiert, und folgern, dass die Reihe  $\sum_{n \geq N} a_n$  absolut konvergiert; damit ist aber auch die Reihe  $\sum_{n=1}^{\infty} a_n$  absolut konvergent.  $\square$

**Bemerkung 2.** Ist auf der anderen Seite  $\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1$ , so gibt es unendlich viele  $n \in \mathbb{N}$  welche  $|a_n| > 1$  erfüllen; also ist die Folge der Reihenglieder keine Nullfolge, und die Reihe damit nicht konvergent. Im Falle von  $\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = 1$  ist keine Aussage möglich, wie die Beispiele  $\sum_{n=1}^{\infty} 1$  und  $\sum_{n=1}^{\infty} \frac{1}{n^2}$  zeigen.

SATZ 20. Sei  $(a_n)_{n \in \mathbb{N}}$  eine Folge, mit  $a_n \neq 0$  für alle  $n \in \mathbb{N}$ . Wenn

$$\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1$$

ist, so konvergiert

$$\sum_{n=1}^{\infty} a_n$$

absolut.

## 4. Stetigkeit und Grenzwerte von Funktionen

### 4.1. Definitionen und einfache Eigenschaften.

**Definition 14.** Sei  $X \subset \mathbb{R}$ ,  $f: X \rightarrow \mathbb{R}$  eine Funktion,  $x_0 \in X$ . Wir sagen,  $f$  ist stetig im Punkt  $x_0$ , wenn für jede Umgebung  $V \in \mathcal{N}(f(x_0))$  eine Umgebung  $U \in \mathcal{N}(x_0)$  existiert sodass  $f(U \cap X) \subset V$  gilt. Wir sagen,  $f$  ist stetig auf  $X$ , wenn  $f$  in jedem Punkt  $x \in X$  stetig ist.

**Definition 15.** Sei  $x_0 \in \mathbb{R}$  ein Häufungspunkt von  $X \subset \mathbb{R}$ ,  $f: X \rightarrow \mathbb{R}$  eine Funktion, und  $L \in \mathbb{R}$ . Wir sagen,  $f(x)$  strebt gegen  $L$  für  $x \rightarrow x_0$ , und schreiben

$$\lim_{\substack{x \rightarrow x_0 \\ x \in X}} f(x) = L,$$

wenn für jede Umgebung  $V \in \mathcal{N}(L)$  eine Umgebung  $U \in \mathcal{N}(x_0)$  existiert sodass  $f(U \cap X \setminus \{x_0\}) \subset V$  gilt.

**Bemerkung 3.** Zwei typische Fälle für Grenzwerte von Funktionen treten auf, wenn man  $X = [a, b]$  betrachtet und den Grenzwert von  $f$  für  $x \rightarrow a$  bzw.  $x \rightarrow b$  betrachtet; in diesem Fall schreibt man

$$\lim_{\substack{x \rightarrow a \\ x \in [a, b]}} f(x) = \lim_{x \rightarrow a^+} f(x)$$

und

$$\lim_{\substack{x \rightarrow b \\ x \in [a, b]}} f(x) = \lim_{x \rightarrow b^-} f(x).$$

Eine ähnliche Notation wird auch oft verwendet, wenn  $f$  in einer Menge  $U$  definiert ist und  $a, b \in U$  ist; in diesem Fall schränkt man  $f$  einfach auf  $X = [a, \infty) \cap U$  bzw. auf  $X = (-\infty, b] \cap U$  ein und schreibt dann

$$\lim_{\substack{x \rightarrow a \\ x \in [a, \infty) \cap U}} f(x) = \lim_{x \rightarrow a^+} f(x)$$

und

$$\lim_{\substack{x \rightarrow b \\ x \in (-\infty, b] \cap U}} f(x) = \lim_{x \rightarrow b^-} f(x).$$

**Bemerkung 4.** Der Zusatz “ $x \in X$ ” erscheint zunächst nicht notwendig, da wir ja nur von Funktionen sprechen, welche auf  $X \subset \mathbb{R}$  definiert sind. Oft ist es aber so, dass wir—wie in der obigen Bemerkung—eine auf einer unter Umständen grösseren Menge  $U$  definierte Funktion entlang von verschiedenen “Richtungen” testen wollen, welche durch entsprechende Wahlen von  $X$  gegeben werden. In diesem Fall erspart die obige Notationskonvention das Einschränken der Funktion  $f$  auf die Menge  $X$ .

**Bemerkung 5.** Wenn es  $\lim_{\substack{x \rightarrow x_0 \\ x \in X}} f(x)$  gibt, so ist der Wert eindeutig bestimmt: Erfüllen sowohl  $L$  als auch  $M$  die Definition, und ist  $\varepsilon > 0$  beliebig, so gibt es Umgebungen  $U \in \mathcal{N}(x_0)$  und  $V \in \mathcal{N}(x_0)$  sodass für jedes  $x \in U \cap X$  die Ungleichung  $|L - f(x)| < \varepsilon/2$  und für jedes  $x \in V \cap X$  die Ungleichung  $|M - f(x)| < \varepsilon/2$  erfüllt ist. Da  $x_0$  ein Häufungspunkt von  $X$  ist, ist die Menge  $X \cap U \cap V \setminus \{x_0\} \neq \emptyset$ ; wir können daher ein  $x \in X \cap U \cap V \setminus \{x_0\}$  wählen und wie folgt abschätzen:

$$|L - M| \leq |L - f(x)| + |f(x) - M| < \varepsilon.$$

Nachdem  $\varepsilon > 0$  beliebig war, folgt  $L = M$ .

**Satz 21.** Sei  $f: X \rightarrow \mathbb{R}$  eine Funktion,  $x_0 \in X$  ein Häufungspunkt von  $X$ . Dann sind die folgenden Aussagen äquivalent:

- i)  $f$  ist stetig im Punkt  $x_0$ .
- ii)  $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ .
- iii) Für jede Folge  $(x_n)_{n=1}^{\infty}$  mit  $x_j \in (a, b)$  für alle  $j \in \mathbb{N}$  und  $\lim_{n \rightarrow \infty} x_n = x_0$  gilt

$$\lim_{n \rightarrow \infty} f(x_n) = f(x_0).$$

- iv) Für jedes  $\varepsilon > 0$  gibt es ein  $\delta > 0$  sodass  $|f(x) - f(x_0)| < \varepsilon$  für alle  $x \in X$  mit  $|x - x_0| < \delta$  ist.

BEWEIS. i)⇒ii): Wenn  $f$  im Punkt  $x_0$  stetig ist, so erfüllt  $f(x_0)$  genau die in Definition 15 geforderten Eigenschaften.

ii)⇒iii): Sei  $(x_n)_{n=1}^\infty$  eine solche Folge, und  $V \in \mathcal{N}(f(x_0))$  beliebig. Nachdem  $\lim_{x \rightarrow x_0, x \in X} f(x) = f(x_0)$  ist, gibt es eine Umgebung  $U \in \mathcal{N}(x_0)$  mit  $f(x) \in V$  für alle  $x \in U \cap X$ . Da  $\lim_{n \rightarrow \infty} x_n = x_0$ , gibt es ein  $N \in \mathbb{N}$  sodass  $x_n \in U$  für alle  $n \geq N$  gilt. Also ist  $f(x_n) \in V$  für alle  $n \geq N$ . Nachdem  $V \in \mathcal{N}(f(x_0))$  beliebig war, folgt  $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$ .

iii)⇒iv): Wenn iv) nicht erfüllt ist, so gibt es ein  $\varepsilon_0 > 0$  sodass für jedes  $\delta > 0$  ein  $y_\delta \in X$  existiert, welches  $|y_\delta - x_0| < \delta$  erfüllt, und sodass  $|f(x) - f(y_\delta)| \geq \varepsilon_0$  gilt. Wir erhalten somit eine Folge  $x_n = y_{1/n} \in X$ , welche  $|x_n - x_0| < 1/n$  erfüllt, also  $\lim_{n \rightarrow \infty} x_n = x_0$ , aber da  $|f(x_n) - f(x_0)| \geq \varepsilon_0$ , konvergiert  $(f(x_n))_{n=1}^\infty$  nicht gegen  $f(x_0)$ .

iv)⇒i): Sei  $V$  eine beliebige Umgebung von  $f(x_0)$ , so gibt es ein  $\varepsilon > 0$  mit  $B_\varepsilon(f(x_0)) \subset V$ . Wir wählen  $\delta > 0$  wie in iv) und erhalten mit  $U = B_\delta(x_0)$ , dass  $f(x) \in B_\varepsilon(f(x_0))$  für alle  $x \in U \cap X$  ist.  $\square$

**Beispiel 14.** Sei  $f(x) = a + bx$ , wo  $a, b \in \mathbb{R}$  Konstanten sind, eine affine Funktion,  $\neq 0$ . Dann ist  $f$  stetig auf ganz  $\mathbb{R}$ : Sei dazu  $x_0 \in \mathbb{R}$  beliebig, und  $\varepsilon > 0$  gegeben. Ist nun  $|x - x_0| < \varepsilon/|b|$ , so ist

$$|f(x) - f(x_0)| = |bx - bx_0| = |b||x - x_0| < |b|\varepsilon/|b| = \varepsilon,$$

und eine Anwendung von iv) liefert die Stetigkeit von  $f$ .

SATZ 22. Sei  $X \subset \mathbb{R}$ , und  $f: X \rightarrow \mathbb{R}$  eine Funktion,  $x_0 \in \mathbb{R} \setminus X$  ein Häufungspunkt von  $X$ . Wenn  $\lim_{x \rightarrow x_0, x \in X} f(x)$  existiert, so ist die Funktion  $\tilde{f}: X \cup \{x_0\} \rightarrow \mathbb{R}$ , welche durch

$$\tilde{f}(x) = \begin{cases} f(x) & x \in X \\ \lim_{x \rightarrow x_0, x \in X} f(x) & x = x_0 \end{cases}$$

definiert ist, stetig in  $x_0$ . Wir nennen dieses  $\tilde{f}$  die stetige Fortsetzung von  $f$  in  $x_0$  (oder auf  $X \cup \{x_0\}$ ).

BEWEIS. Bleibt als Übungsaufgabe.  $\square$

**Beispiel 15.** Die Funktion  $f: \mathbb{R} \setminus \{-1\} \rightarrow \mathbb{R}$ ,  $f(x) = \frac{x^2-1}{x+1}$ , besitzt eine stetige Fortsetzung auf  $\mathbb{R}$  bzw. in den Punkt  $-1$ , da

$$\lim_{\substack{x \rightarrow -1 \\ x \neq -1}} \frac{x^2-1}{x+1} = \lim_{\substack{x \rightarrow -1 \\ x \neq -1}} (x-1) = -2.$$

Wir fügen noch eine kleine Bemerkung zur Einschränkung von Funktionen und der Fortsetzung mit Hilfe von Grenzwerten an.

LEMMA 39. Wenn  $f$  stetig auf  $X$  ist, so ist  $f|_Y$  stetig auf  $Y$  für jede Teilmenge  $Y \subset X$ . Ist  $x_0$  ein Häufungspunkt von  $Y$ ,  $Y \subset X$ , und existiert

$$\lim_{\substack{x \rightarrow x_0 \\ x \in X}} f(x),$$

so existiert auch

$$\lim_{\substack{x \rightarrow x_0 \\ x \in Y}} f(x),$$

und es gilt

$$\lim_{\substack{x \rightarrow x_0 \\ x \in Y}} f(x) = \lim_{\substack{x \rightarrow x_0 \\ x \in X}} f(x).$$

Ist  $X = Y_1 \cup Y_2$  und  $x_0$  ein Häufungspunkt von sowohl  $Y_1$  als auch von  $Y_2$ , so existiert

$$\lim_{\substack{x \rightarrow x_0 \\ x \in X}} f(x)$$

genau dann, wenn die Grenzwerte

$$\lim_{\substack{x \rightarrow x_0 \\ x \in Y_1}} f(x) \text{ und } \lim_{\substack{x \rightarrow x_0 \\ x \in Y_2}} f(x)$$

existieren und gleich sind.

**Definition 16.** Wir sagen, eine Funktion  $f: X \rightarrow \mathbb{R}$  ist gleichmässig stetig auf  $X$ , wenn für jede Umgebung  $V \subset \mathcal{N}(0)$  eine Umgebung  $U \subset \mathcal{N}(0)$  existiert, sodass  $f(x) - f(y) \in V$  für alle  $x, y \in X$  mit  $x - y \in U$ .

SATZ 23. Sei  $X \subset \mathbb{R}$  kompakt, und  $f: X \rightarrow \mathbb{R}$  stetig auf  $X$ . Dann ist  $f$  gleichmässig stetig auf  $X$ .

BEWEIS. Sei  $\varepsilon > 0$ . Da  $f$  stetig in jedem Punkt  $x \in X$  ist, gibt es zu jedem  $z \in X$  ein  $\delta_z$  mit der Eigenschaft, dass  $|f(z) - f(x)| < \frac{\varepsilon}{2}$  wenn  $|z - x| < 2\delta_z$  und  $y \in X$  ist. Die Mengen  $B_{\delta_z}(z)$  überdecken  $X$ ; da  $X$  kompakt ist, gibt es endlich viele Punkte  $z_1, \dots, z_N$  mit

$$X \subset \bigcup_{j=1}^N B_{\delta_{z_j}}(z_j).$$

Wir wählen nun  $\delta = \min\{\delta_{z_1}, \dots, \delta_{z_N}\}$ . Für jeden Punkt  $x \in X$  gibt es also ein  $j \in \{1, \dots, N\}$  mit  $x \in B_{\delta_{z_j}}(z_j)$ . Ist  $y \in X$  mit  $|x - y| < \delta$  gegeben, so ist  $|z_j - y| \leq |z_j - x| + |x - y| < \delta_{z_j} + \delta \leq 2\delta_{z_j}$ , und wir können damit wie folgt abschätzen:

$$|f(x) - f(y)| \leq |f(x) - f(z_j)| + |f(z_j) - f(y)| \leq \varepsilon,$$

wenn  $|x - y| < \delta$ . Es folgt, dass  $f$  gleichmässig stetig auf  $X$  ist.  $\square$

## 4.2. Rechnen mit stetigen Funktionen.

SATZ 24. Seien  $f: X \rightarrow \mathbb{R}$  und  $g: X \rightarrow \mathbb{R}$  Funktionen,  $x_0 \in X$ , und  $\lambda \in \mathbb{R}$ . Wenn  $f$  und  $g$  stetig im Punkt  $x_0$  sind, so sind die auf  $X$  durch

$$(f + \lambda g)(x) = f(x) + \lambda g(x), \quad (fg)(x) = f(x)g(x)$$

definierten Funktionen  $f + \lambda g$  und  $fg$  stetig im Punkt  $x_0$ .

Ist  $f(x) \neq 0$  auf  $X$  und  $f$  stetig im Punkt  $x_0$ , so ist die durch  $(1/f)(x) = \frac{1}{f(x)}$  auf  $X$  definierte Funktion  $1/f$  stetig im Punkt  $x_0$ .

BEWEIS. Sei  $U \in \mathcal{N}((f + \lambda g)(x_0))$ . Dann gibt es nach Lemma 28 eine Umgebung  $V \in \mathcal{N}(f(x_0))$  und eine Umgebung  $W \in \mathcal{N}(g(x_0))$  mit  $V + \lambda W \subset U$ . Da  $f$  und  $g$  stetig im Punkt  $x_0$  sind, gibt es Umgebungen  $V_1 \in \mathcal{N}(x_0)$  und  $W_1 \in \mathcal{N}(x_0)$  mit  $f(x) \in V$  für alle  $x \in V_1 \cap X$  und  $g(x) \in W$  für alle  $x \in W_1 \cap X$ . Es folgt, dass  $(f + \lambda g)(x) = f(x) + \lambda g(x) \in V + \lambda W \subset U$  für alle  $x \in U_1 = V_1 \cap W_1$ . Also ist  $f + \lambda g$  stetig in  $x_0$ .

Die Aussagen über das Produkt bzw. die Inverse folgt mittels des analogen Beweises unter Verwendung von Lemma 29 bzw. Lemma 30.  $\square$

**Beispiel 16.** Jede Polynomfunktion ist stetig auf ganz  $\mathbb{R}$ . Eine rationale Funktion  $f(x) = \frac{p(x)}{q(x)}$ , wo  $p, q \in \mathbb{R}[x]$ , definiert auf  $\mathbb{R} \setminus \{q = 0\}$ , ist stetig in ihrem Definitionsbereich.

Die Menge der stetigen Funktionen auf der Menge  $X \subset \mathbb{R}$  wird mit  $C(X)$  bezeichnet. Der obige Satz besagt, dass  $C(X)$  abgeschlossen unter der Addition  $(f, g) \mapsto f + g$ , der Skalarmultiplikation  $(\lambda, f) \mapsto \lambda f$ , der Multiplikation  $(f, g) \mapsto fg$ , sowie der Bildung von multiplikativ Inversen (so diese gestattet ist)  $f \mapsto 1/f$ . Es lässt sich leicht überprüfen, dass  $C(X)$  mit der Addition und der Skalarmultiplikation versehen ein Vektorraum ist, und mit der Multiplikation eine Algebra.  $C(X)$  ist zumeist nicht endlichdimensional. Wir geben hier nur ein Beispiel:

**Beispiel 17.** Sei  $X = [0, 1]$ . Wir definieren die Funktionen  $f_n \in C(X)$ , wo  $n \in \mathbb{N}$  ist, durch

$$f_n(x) = \begin{cases} 0 & x \leq \frac{1}{n+1} \\ x - \frac{1}{n} & x \geq \frac{1}{n+1}. \end{cases}$$

Dann ist jede endliche Teilmenge von  $\{f_n : n \in \mathbb{N}\}$  linear unabhängig: Angenommen, dass  $\sum_{n=1}^N \lambda_n f_n = 0$  in  $C([0, 1])$  ist, also  $\sum_{n=1}^N \lambda_n f_n(x) = 0$  für alle  $x \in [0, 1]$  gilt. Dann ist für  $x = \frac{1}{N}$

$$0 = \sum_{n=1}^N \lambda_n f_n\left(\frac{1}{N}\right) = \lambda_N \left(\frac{1}{N} - \frac{1}{N+1}\right),$$

also  $\lambda_N = 0$ . Damit ist auch

$$0 = \sum_{n=1}^N \lambda_n f_n\left(\frac{1}{N-2}\right) = \lambda_{N-1} \left(\frac{1}{N-1} - \frac{1}{N}\right),$$

also  $\lambda_{N-1} = 0$ . Induktiv voranschreitend sehen wir, dass  $\lambda_N = \dots = \lambda_1 = 0$  ist.

**4.3. Konvergenz von Funktionenfolgen- und reihen.** Unter einer Folge von Funktionen  $(f_n)_{n=1}^\infty$  auf  $X$  verstehen wir eine Abbildung, die jedem  $n \in \mathbb{N}$  eine Funktion  $f_n: X \rightarrow \mathbb{R}$  zuordnet. Ist  $(g_j)_{j=1}^\infty$  eine Folge von Funktionen auf  $X$ , so verstehen wir unter der (aus den Gliedern  $g_j$  gebildeten) Funktionenreihe verstehen wir die Folge der Partialsummen  $S_n: X \rightarrow \mathbb{R}$ , welche durch

$$S_n(x) = \sum_{j=1}^n g_j(x)$$

definiert ist. Für jedes  $x \in X$  wissen wir, was unter der Konvergenz von

$$(f_n(x))_{n=1}^\infty$$

bzw. der (absoluten) Konvergenz von

$$\sum_{j=1}^\infty g_j(x)$$

zu verstehen ist.

Konvergiert  $(f_n(x))_{n=1}^\infty$  bzw.  $\sum_{j=1}^\infty g_j(x)$  für jedes  $x \in X$ , so existieren die *punktweisen Grenzfunktionen*  $f$  bzw.  $g$ , welche für  $x \in X$  durch  $f(x) = \lim_{n \rightarrow \infty} f_n(x)$  bzw.  $g(x) = \sum_{j=1}^\infty g_j(x)$  definiert sind.

Diese Art der punktweisen Konvergenz ist nicht genug, um Stetigkeit der Grenzfunktion einer Folge stetiger Funktionen zu garantieren.

**Beispiel 18.** Wir betrachten die Folge  $(f_n)_{n=1}^\infty$  von Funktionen auf  $X = [0, 1]$ , welche durch  $f_n(x) = x^n$  für  $x \in [0, 1]$  definiert ist. Jede der Funktionen  $f_n$  ist stetig. Die punktweise Grenzfunktion  $f: [0, 1] \rightarrow \mathbb{R}$  existiert, und ist durch

$$f(x) = \begin{cases} 0 & x < 1 \\ 1 & x = 1 \end{cases}$$

gegeben, also nicht stetig.

Das Problem ist, dass wir bei der Annäherung durch stetige Funktionen auch eine gute Annäherung von Werten "nahe" eines fixen Punktes  $x_0$  benötigen. Um dies zu erreichen, benötigt man eine strengere Konvergenz als die punktweise Konvergenz.

**Definition 17.** Wir sagen, die Folge  $(f_n)_{n=1}^\infty$  von Funktionen auf  $X$  konvergiert gleichmässig auf  $X$  gegen  $f: X \rightarrow \mathbb{R}$ , wenn für jede Umgebung  $U \in \mathcal{N}(0)$  ein  $N \in \mathbb{N}$  existiert, sodass  $f_n(x) - f(x) \in U$  für alle  $n \geq N$  und alle  $x \in X$  gilt. Wir sagen, die Folge  $(f_n)_{n=1}^\infty$  von Funktionen auf  $X$  ist eine gleichmässige Cauchyfolge auf  $X$ , wenn für jede Umgebung  $U \in \mathcal{N}(0)$  ein  $N \in \mathbb{N}$  existiert, sodass  $f_n(x) - f_m(x) \in U$  für alle  $n, m \geq N$  und alle  $x \in X$  gilt.

**Satz 25.** Seien  $f_n \in C(X)$  für alle  $n \in \mathbb{N}$ , und  $(f_n)_{n=1}^\infty$  eine gleichmässige Cauchyfolge auf  $X$ . Dann konvergiert  $(f_n)_{n=1}^\infty$  gleichmässig gegen eine Grenzfunktion  $f \in C(X)$ .

**BEWEIS.** Da  $(f_n)_{n=1}^\infty$  eine Cauchyfolge ist, ist für jedes  $x \in X$  die Folge  $(f_n(x))_{n=1}^\infty$  reeller Zahlen eine Cauchyfolge, die nach Theorem 9 einen Grenzwert besitzt, den wir mit  $f(x)$  bezeichnen; diese Festlegung definiert (wegen der Eindeutigkeit des Grenzwerts) eine Funktion  $f: X \rightarrow \mathbb{R}$ . Sei nun  $U \in \mathcal{N}(0)$  beliebig. Wir wählen eine Umgebung  $V \in \mathcal{N}(0)$  mit  $\bar{V} \subset U$ .

Es gibt ein  $N \in \mathbb{N}$  mit  $f_n(x) - f_m(x) \in V$  für alle  $m, n \geq N$ . Mittels Grenzübergang  $m \rightarrow \infty$  sehen wir, dass  $f_n(x) - f(x) \in \bar{V} \subset U$  für alle  $n \geq N$  gilt. Also konvergiert  $(f_n)_{n=1}^\infty$  gleichmässig auf  $X$  gegen  $f$ .

Sei  $x_0 \in X$  beliebig, aber fest gewählt, und  $V \subset \mathcal{N}(x_0)$ ; wir wählen ein  $W \in \mathcal{N}(0)$  mit  $x_0 + 3W \subset V$ . Da  $f$  der gleichmässige Grenzwert von  $(f_n)_{n=1}^\infty$  ist, gibt es ein  $N \in \mathbb{N}$ , sodass  $f_N(x) - f(x) \in W \cap (-W)$

für alle  $x \in X$  ist. Da  $f_N$  stetig im Punkt  $x_0$  ist, gibt es ein  $U \in \mathcal{N}(x_0)$  mit  $f_N(x) \in f_N(x_0) + W$  für alle  $x \in U \cap X$ . Dann ist

$$f(x) = f(x_0) + (f_N(x_0) - f(x_0)) + (f_N(x) - f_N(x_0)) + (f(x) - f_N(x)) \in f(x_0) + 3W \subset U$$

für alle  $x \in U \cap X$ , also  $f$  stetig im Punkt  $x_0$ . Nachdem  $x_0 \in X$  beliebig war, folgt  $f \in C(X)$ .  $\square$

**KOROLLAR 2.** Sei  $(e_n)_{n \in \mathbb{N}}$  eine Nullfolge. Wenn  $f_n \in C(X)$  ist, und  $|f_n(x) - f(x)| < e_n$  für alle  $n \in \mathbb{N}$  ist, so ist  $f \in C(X)$  und  $(f_n)_{n \in \mathbb{N}}$  konvergiert gleichmässig (auf  $X$ ) gegen  $f$ .

Sei  $g_n \in C(X)$  für alle  $n \in \mathbb{N}$ . Wenn  $|g_n(x)| \leq \varepsilon_n$  gilt und die Reihe  $\sum_{j \in \mathbb{N}} \varepsilon_j$  konvergiert, so konvergiert  $\sum g_n$  gleichmässig und absolut gegen eine Grenzfunktion  $g \in C(X)$ .

**BEWEIS.** Sei  $\varepsilon > 0$  beliebig, dann gibt es ein  $N \in \mathbb{N}$  sodass  $e_n < \frac{\varepsilon}{2}$  für alle  $n \geq N$  ist. Sind nun  $n, m \geq N$ , so ist

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \varepsilon,$$

für alle  $x \in X$ .

Es folgt, dass  $(f_n)_{n \in \mathbb{N}}$  eine gleichmässige Cauchyfolge auf  $X$  ist, die eine stetige Grenzfunktion  $g \in C(X)$  besitzt. Da der gleichmässige Grenzwert eindeutig bestimmt ist, ist  $f = g$  also auch stetig auf  $X$ .

Für die Aussage über die dominierende Reihe sei wieder  $\varepsilon > 0$  beliebig. Es gibt dann ein  $N \in \mathbb{N}$  sodass

$$\sum_{j=n+1}^m \varepsilon_j < \varepsilon$$

für alle  $N \leq n \leq m$  ist. Es folgt, dass für diese  $m, n$  auch

$$\begin{aligned} \sum_{j=1}^m |g_j(x)| - \sum_{j=1}^n |g_j(x)| &= \sum_{j=n+1}^m |g_j(x)| \\ &\leq \sum_{j=n+1}^m \varepsilon_j \\ &< \varepsilon, \end{aligned}$$

für alle  $x \in X$  gilt. Wie schon bei Reihen folgt auch bei Funktionenreihen die gleichmässige Konvergenz aus der gleichmässigen absoluten Konvergenz: Wenn wir mit  $S_n : X \rightarrow \mathbb{R}$ , wo  $S_n(x) = \sum_{j=1}^n g_j(x)$ , die  $n$ -te Partialsumme bezeichnen, so ist

$$\begin{aligned} |S_n(x) - S_m(x)| &\leq \left| \sum_{j=1}^n g_j(x) - \sum_{j=1}^m g_j(x) \right| \\ &= \left| \sum_{j=n+1}^m g_j(x) \right| \\ &\leq \sum_{j=n+1}^m |g_j(x)|; \end{aligned}$$

mit den Wahlen von  $n, m$  wie oben, und wir sehen, dass  $(S_n)_{n \in \mathbb{N}}$  eine gleichmässige Cauchyfolge bildet.  $\square$

**KOROLLAR 3.** Sei  $(g_n(x))_{n \in \mathbb{N}}$  eine Folge von stetigen Funktionen auf  $X$ . Wenn

$$\limsup_{n \rightarrow \infty} \sup_{x \in X} \sqrt[n]{|g_n(x)|} < 1$$

oder

$$\limsup_{n \rightarrow \infty} \sup_{x \in X} \frac{|g_{n+1}(x)|}{|g_n(x)|} < 1$$

gilt, so konvergiert  $\sum_j g_j$  gleichmässig auf  $X$ , und  $\sum_j g_j$  ist stetig auf  $X$ .

BEWEIS. Offensichtlich ist  $|g_j(x)| \leq \varepsilon_j = \sup_{y \in X} |g_j(y)|$  für alle  $x \in X$ . Die Voraussetzungen des Korollars implizieren nach Theorem 19 bzw. Theorem 20, dass die Reihe  $\sum_j \varepsilon_j$  absolut konvergiert. Korollar 2 impliziert, dass  $\sum_j g_j$  gleichmässig absolut konvergiert.  $\square$

**Beispiel 19.** Sei  $(a_n)_{n \geq 0}$  eine Folge reeller Zahlen. Wir definieren

$$R = \begin{cases} (\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|})^{-1} & 0 < \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} < \infty \\ \infty & \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = 0 \\ 0 & (\sqrt[n]{|a_n|})_{n \geq 0} \text{ nicht beschränkt.} \end{cases}.$$

Sei  $0 < r < R$ . Dann erfüllt die Folge  $g_j: [-r, r] \subset (-R, R)$ , welche durch  $g_j(x) = a_j x^j$  definiert ist, die Voraussetzungen von Korollar 3, da

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n x^n|} = |x| \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n x^n|} \begin{cases} \leq r R^{-1} < 1 & R \neq \infty \\ = 0 & R = \infty. \end{cases}$$

Es folgt, dass die Potenzreihe  $\sum_{j=0}^{\infty} a_j x^j$  für jedes  $x \in (-R, R)$  konvergiert, und die Konvergenz gleichmässig auf jedem Intervall  $[-r, r] \subset (-R, R)$  ist; die Grenzfunktion ist also stetig auf ganz  $(-R, R)$ . Wir nennen  $R$  den *Konvergenzradius* der Potenzreihe  $\sum_{j=0}^{\infty} a_j x^j$ .

**4.4. Vollständigkeit und Werte stetiger Funktionen.** Eine weitere wichtige Folgerung aus der Vollständigkeit ist der Zwischenwertsatz:

**SATZ 26** (Zwischenwertsatz). Sei  $f: \mathbb{R} \supset [a, b] \rightarrow \mathbb{R}$  stetig, und  $y \in \mathbb{R}$  mit  $f(a) \leq y \leq f(b)$  gegeben. Dann gibt es ein  $c \in [a, b]$  mit  $f(c) = y$ .

BEWEIS. Wir setzen  $S = \{x \in [a, b]: f(x) \leq y\} \neq \emptyset$  und somit gibt es  $c := \sup S \in [a, b]$ . Wir behaupten, dass  $f(c) = y$  ist. Dann betrachten wir zunächst die Folge  $c_j = \min\{c + 1/j, b\}$ . Da  $c_j > c$ , ist  $f(c_j) > y$ , und damit

$$f(c) = \lim_{j \rightarrow \infty} f(c_j) \geq y.$$

Auf der anderen Seite können wir eine Folge  $d_j$  mit  $c - 1/j \leq d_j \leq c$  betrachten, welche  $f(d_j) \leq y$  erfüllt. Damit sehen wir, dass

$$f(c) = \lim_{j \rightarrow \infty} f(d_j) \leq y$$

ist. Also folgt  $y = f(c)$  wie gewünscht.  $\square$

**Übungsaufgabe 13.** Zeige, dass ein Polynom ungeraden Grades  $p(x) = ax^{2k+1} + \dots, a \neq 0$ , als Funktion  $p: \mathbb{R} \rightarrow \mathbb{R}$  aufgefasst, surjektiv ist.

**Übungsaufgabe 14.** Sei  $f: \mathbb{R} \supset [a, b] \rightarrow [a, b]$  stetig. Dann hat  $f$  einen Fixpunkt in  $[a, b]$ , das heisst, es gibt ein  $x \in [a, b]$  mit  $f(x) = x$ .

Auf einer kompakten Menge werden das Maximum und das Minimum der Werte angenommen:

**SATZ 27.** Sei  $f: \mathbb{R} \supset X \rightarrow \mathbb{R}$  stetig auf der kompakten Menge  $X$ . Dann existieren  $\min_{x \in X} f(x)$  und  $\max_{x \in X} f(x)$ .

BEWEIS. Wir zeigen zunächst, dass  $f(X)$  beschränkt ist. Angenommen, das ist nicht der Fall: Dann gibt es eine Folge  $y_j = f(x_j)$  von Punkten in  $f(X)$  mit  $|y_j| > j$ . Da  $X$  kompakt ist, können wir nun eine Teilfolge  $(x_{j_k})_{k=1}^{\infty}$  der  $x_j$  wählen, welche gegen  $x_0 \in X$  konvergiert. Da  $f$  stetig auf  $X$  ist, erhalten wir einen Widerspruch, da die Folge  $(f(x_{j_k}))_{k=1}^{\infty}$  nicht beschränkt ist.

Nun existiert also  $\sup_{x \in X} f(x)$ , und für jedes  $n \in \mathbb{N}$  können wir ein  $x_n \in X$  wählen sodass  $f(x_n) > \sup_{x \in X} f(x) - 1/n$ . Wir wählen nun eine Teilfolge  $x_{n_k}$  welche gegen  $x_0 \in X$  konvergiert. Dann ist

$$\sup_{x \in X} f(x) \geq f(x_0) = \lim_{k \rightarrow \infty} f(x_{n_k}) \geq \lim_{k \rightarrow \infty} \left( \sup_{x \in X} f(x) - \frac{1}{n} \right) = \sup_{x \in X} f(x).$$

Es folgt, dass  $f(x_0) = \sup_{x \in X} f(x)$  ist, also wegen  $x_0 \in X$ , dass  $f(x_0) = \max_{x \in X} f(x)$ .

Der Beweis für das Minimum erfolgt entsprechend.  $\square$

SATZ 28. Sei  $f \in C([a, b])$  strikt monoton steigend (bzw. fallend). Dann ist  $f: [a, b] \rightarrow [f(a), f(b)]$  (bzw.  $f: [a, b] \rightarrow [f(b), f(a)]$ ) bijektiv, und die Umkehrfunktion  $f^{-1}$  wiederum stetig.

BEWEIS. Die Injektivität von  $f$  folgt direkt aus der strikten Monotonie, sowie, dass  $f([a, b]) \subset [f(a), f(b)]$ , wenn  $f$  monoton steigend ist (bzw.  $\subset [f(b), f(a)]$ , falls  $f$  monoton fallend ist). Die Surjektivität von  $f$  ist der Inhalt von Theorem 26.

Wir müssen nur noch zeigen, dass  $f^{-1}$  stetig ist; wir nehmen dazu an, dass  $f$  monoton steigend ist. Sei dazu  $y_0 \in [f(a), f(b)]$  gegeben, und  $y_j$  eine Folge in  $[f(a), f(b)]$  mit  $\lim_{j \rightarrow \infty} y_j = y_0$ . Die Folge  $x_j = f^{-1}(y_j)$  ist dann eine Folge in  $[a, b]$ , also beschränkt. Ist  $S$  ein beliebiger Häufungswert von  $(x_j)_{j=1}^{\infty}$ , so gibt es eine Teilfolge  $(x_{j_k})_{k \in \mathbb{N}}$  mit  $\lim_{k \rightarrow \infty} x_{j_k} = S$ . Wegen der Stetigkeit von  $f$  folgt nun, dass  $f(S) = \lim_{k \rightarrow \infty} f(x_{j_k}) = y_0$  ist, also  $S = f^{-1}(y_0)$ . Also ist  $f^{-1}$  stetig im Punkt  $y_0$ .  $\square$

**Beispiel 20.** Die stetigen Funktion  $f(x) = x^q$ , wo  $q \in \mathbb{Z} \setminus \{0\}$  ist, ist strikt monoton steigend (wenn  $q > 0$ ) bzw. strikt monoton fallend (wenn  $q < 0$ ) auf  $\mathbb{R}_+$ . Ihre inverse Funktion  $f^{-1}(y) =: y^{\frac{1}{q}}$  ist also auch stetig auf  $\mathbb{R}_+$ . Damit sind Potenzfunktionen mit rationalen Exponenten stetig auf  $\mathbb{R}_+$ .



## Stetigkeit und Vollständigkeit

### 1. Stetigkeit im Allgemeinen: Topologische Räume

Stetigkeit von Funktionen wird in einer “natürlichen” Art und Weise mit Hilfe einer *Topologie* auf einem Raum behandelt. Topologien sind durch Systeme *offener Mengen* gegeben, welche uns Umgebungen von Punkten liefern. Man definiert:

**Definition 18.** Eine Topologie  $\mathfrak{T}$  auf einer Menge  $X$  ist eine Familie  $\mathfrak{T} \subset 2^X$  von Teilmengen von  $X$ , welche abgeschlossen unter beliebigen Vereinigungen und endlichen Durchschnitten ist, d.h.

$$U_i \in \mathfrak{T}, i \in I \Rightarrow \bigcup_{i \in I} U_i \in \mathfrak{T}, \quad U, V \in \mathfrak{T} \Rightarrow U \cap V \in \mathfrak{T},$$

und welche  $\emptyset, X \in \mathfrak{T}$  erfüllt; die Elemente von  $\mathfrak{T}$  werden als *offene Mengen* bezeichnet. Ein topologischer Raum  $(X, \mathfrak{T})$  ist eine Menge  $X$  gemeinsam mit einer Topologie  $\mathfrak{T}$  auf  $X$ . Ist klar, welche Topologie auf  $X$  gemeint ist, sprechen wir auch oft von einem “topologischen Raum  $X$ ”.

Topologien sind sehr allgemeine Objekte (sie besitzen nur drei definierende Eigenschaften). Mit ihrer Hilfe kann man den Begriff der Stetigkeit und grundlegende Eigenschaften stetiger Funktionen für sehr grosse Klassen von Objekten beschreiben kann. Diese Macht kommt mit einem Preis: Man muss in Kauf nehmen, dass wichtige Eigenschaften (die für gewisse Objekte klarerweise erfüllt sind) hier erst definiert werden müssen. Auch ist es oft umständlich, spezifische Eigenschaften in diesem allgemeinen Kontext zu definieren und zu untersuchen. Wir werden deswegen relativ rasch zu weniger allgemeinen Räumen übergehen.

**Beispiel 21.** i)  $\mathfrak{T} = \{X, \emptyset\}$  ist eine Topologie auf jedem Raum  $X$ , die *triviale* Topologie.

ii)  $\mathfrak{T} = 2^X$ , die Potenzmenge von  $X$ , ist auch eine Topologie auf  $X$  für jede Menge  $X$ , die *diskrete* Topologie.

iii) Ist  $Y \subset X$ , so erhält  $Y$  eine Topologie (die *Teilraumtopologie*) indem man  $\mathfrak{T}_Y = \{V \cap Y : V \in \mathfrak{T}\}$  setzt.

iv) Die “gewöhnliche” Topologie auf  $\mathbb{R}$  ist jene, die wir schon vorher erklärt haben: Eine Menge  $U \subset \mathbb{R}$  ist offen, wenn für jedes  $x \in U$  ein  $\varepsilon = \varepsilon(x)$  existiert, sodass  $(x - \varepsilon, x + \varepsilon) \subset U$ .

Die Elemente von  $\mathfrak{T}$  werden die offenen Mengen in  $X$  genannt. Eine Menge  $C \subset X$  ist *abgeschlossen*, wenn  $C^c$  offen ist. Die abgeschlossenen Mengen erfüllen die nach den Gesetzen von de Morgan “dualen” Eigenschaften zu offenen Mengen, d.h. wenn  $F_\alpha, \alpha \in A$ , eine Familie abgeschlossener Mengen ist, so ist auch  $\bigcap_\alpha F_\alpha$  abgeschlossen, und sind  $F_1, F_2$  abgeschlossen, so ist auch  $F_1 \cup F_2$  abgeschlossen.

Für eine beliebige Teilmenge  $Y \subset X$  gibt es eine grösste offene Menge  $Y^\circ$ , das *Innere von Y*, welche in  $Y$  enthalten ist, und eine kleinste abgeschlossene Menge  $\bar{Y}$ , den *Abschluss von Y*, welche  $Y$  enthält; diese sind als die Vereinigung aller in  $Y$  enthaltenen offenen Mengen bzw. den Durchschnitt aller  $Y$  enthaltenden Mengen definiert. Die entsprechende Konstruktion findet sich in Unterabschnitt 2.2, siehe insbesondere Lemma 25 und Lemma 26.

**Übungsaufgabe 15.** Eine Menge  $D \subset X$  ist *dicht* in  $X$ , wenn für jede offene Menge  $U$  gilt, dass  $D \cap U \neq \emptyset$ . Zeige, dass  $D$  dicht in  $X$  ist genau dann, wenn  $\bar{D} = X$ .

**Definition 19.** Seien  $X, Y$  topologische Räume. Eine Abbildung  $f: X \rightarrow Y$  ist *stetig* wenn für jede offene Menge  $V \subset Y$  das Urbild  $f^{-1}(V) \subset X$  offen ist. Eine bijektive stetige Abbildung, deren Umkehrabbildung auch stetig ist, nennt man einen *Homöomorphismus*.

**Übungsaufgabe 16.** Seien  $X, Y, Z$  topologische Räume,  $f: X \rightarrow Y$  und  $g: Y \rightarrow Z$  stetig. Dann ist  $g \circ f: X \rightarrow Z$  stetig.

**Übungsaufgabe 17.** Die Abbildungen  $\mathbb{R}^2 \rightarrow \mathbb{R}$ , welche durch  $(x, y) \mapsto x + y$  und  $(x, y) \mapsto xy$  gegeben sind, sind stetig. Die Abbildung  $\mathbb{R}_+ \rightarrow \mathbb{R}_+$  welche durch  $x \mapsto x^{-1}$  gegeben ist, ist stetig.

Oft ist es bequemer, mit Umgebungen von Punkten zu hantieren. Eine Umgebung von  $x \in X$  ist eine Menge  $U \subset X$  für welche eine offene Menge  $V$  mit  $x \in V \subset U$  existiert. Wenn wir die Menge der Umgebungen von  $x \in X$  mit  $\mathcal{N}(x)$  bezeichnen, so ist  $\mathcal{N}(x)$  abgeschlossen unter beliebigen Vereinigungen und endlichen Durchschnitten, und  $x \in \mathcal{N}(x)$  für jedes  $x \in X$ . Eine *Umgebungsbasis*  $\mathcal{B}(x)$  ist eine Teilmenge von  $\mathcal{N}(x)$  mit der Eigenschaft, dass für jedes  $V \in \mathcal{N}(x)$  ein  $W \in \mathcal{B}(x)$  existiert sodass  $W \subset V$ . Mit anderen Worten, wenn eine Umgebungsbasis  $\mathcal{B}(x)$  gegeben ist, so ist  $\mathcal{N}(x) = \{V: V \supset W \in \mathcal{B}(x)\}$ . Solche Systeme von Umgebungen kann man auch verwenden, um eine Topologie zu *definieren*:

**SATZ 29.** Für jedes  $x \in X$  sei eine Menge  $\mathcal{N}(x)$  von Teilmengen von  $X$ , welche  $x$  enthalten, gegeben welche folgende Eigenschaften erfüllen:

(N1) Wenn  $U \in \mathcal{N}(x)$  und  $W \supset U$ , so gilt  $W \in \mathcal{N}(x)$ ;

(N2) Wenn  $U_i \in \mathcal{N}(x)$  für alle  $i \in I$ , so gilt  $\bigcup_{i \in I} U_i \in \mathcal{N}(x)$ .,

(N3) Wenn  $U, V \in \mathcal{N}(x)$ , so gilt  $U \cap V \in \mathcal{N}(x)$ ;

(N4) Für jede Menge  $W \in \mathcal{N}(x)$  gibt es eine Menge  $U \in \mathcal{N}(x)$ , welche  $U \in \mathcal{N}(y)$  für alle  $y \in U$  erfüllt. Dann gibt es eine (und nur eine) Topologie auf  $X$  für die  $\mathcal{N}(x)$  die Menge der Umgebungen des Punktes  $x$  für alle  $x \in X$  ist.

Wir wollen von einem Beweis dieses Satzes (der nicht allzu schwierig ist) absehen, da er für uns zwar vom Inhalt her interessant, aber vom Beweis her nicht interessant ist.

**Beispiel 22.** Seien  $X, Y$  topologische Räume. Die Produkttopologie auf  $X \times Y$  definiert man über Umgebungen, indem man die Umgebungen von  $(x, y)$  als Übermengen von Mengen der Form  $U_x \times V_y$  definiert, wo  $U_x$  eine Umgebung von  $x$  in  $X$  und  $V_y$  eine Umgebung von  $y$  in  $Y$  ist.

Mit Hilfe von Umgebungen können wir das Innere und den Abschluss einer Menge charakterisieren: Zunächst sagen wir,  $x \in X$  ist ein *Häufungspunkt* von  $Y \subset X$ , wenn für jede Umgebung  $U \in \mathcal{N}(x)$  der Schnitt  $Y \cap U$  Punkte, die von  $x$  verschieden sind, enthält. Der Abschluss  $\bar{Y}$  einer Menge  $Y \subset X$  besteht nun aus den Punkten der Menge  $Y$  und den Häufungspunkten von  $Y$ .

Für eine beliebige Teilmenge  $Y \subset X$  ist  $x \in Y^\circ$  genau dann, wenn es eine Umgebung  $U \subset \mathcal{N}(x)$  gibt, welche noch ganz in  $Y$  enthalten ist; das Komplement des Abschlusses  $\bar{Y}^c$  ist gegeben durch Punkte, welche eine Umgebung besitzen, die  $Y$  nicht trifft. Dazwischen liegen jene Punkte, für die jede Umgebung sowohl  $Y$  als auch  $Y^c$  trifft, also Punkte  $x \in X$  mit  $U \cap Y \neq \emptyset$  und  $U \cap Y^c \neq \emptyset$  für alle  $U \in \mathcal{N}(x)$ . Solche Punkte sind der *Rand* von  $Y$ , welcher gerne mit  $\partial Y$  bezeichnet wird.

**Übungsaufgabe 18.** Zeige, dass  $\partial Y = \bar{Y} \setminus Y^\circ$ .

Der Begriff der Umgebung erlaubt es uns, Stetigkeit an einer Stelle zu definieren:

**Definition 20.** Wir sagen, die Abbildung  $f: X \rightarrow Y$  ist stetig am Punkt  $x \in X$  wenn für jede Umgebung  $V$  von  $f(x)$  eine Umgebung  $U$  von  $x$  existiert sodass  $f(U) \subset V$ , oder äquivalent, wenn  $f^{-1}(V) \in \mathcal{N}(x)$  für jedes  $V \in \mathcal{N}(f(x))$  gilt.

**Beispiel 23.** Die Abbildungen  $\pi_X: X \times Y \rightarrow X$  und  $\pi_Y: X \times Y \rightarrow Y$ , welche durch

$$\pi_X(x, y) = x, \quad \pi_Y(x, y) = y$$

definiert sind, stetig (da  $\pi_X^{-1}(U_x) = U_x \times Y \in \mathcal{N}^{X \times Y}((x, y))$  für jedes  $U_x \in \mathcal{N}^X(x)$ ). Es gilt weiters, dass eine Abbildung  $f: Z \rightarrow X \times Y$  stetig ist genau dann, wenn die "Komponenten"  $\pi_X \circ f: Z \rightarrow X$  und  $\pi_Y \circ f: Z \rightarrow Y$  stetig sind. Zu zeigen ist dabei nur noch, dass wenn die Komponentenabbildungen stetig sind, auch  $f$  stetig ist.  $z \in Z$ ,  $(x, y) = f(z)$ , und sei  $U \in \mathcal{N}^{X \times Y}((x, y))$ , dann gibt es  $U_x \in \mathcal{N}^X(x)$  und  $V_y \in \mathcal{N}^Y(y)$  mit  $U_x \times V_y \subset U$ . Es folgt  $f^{-1}(U_x \times V_y) = (\pi_X \circ f)^{-1}(U_x) \cap (\pi_Y \circ f)^{-1}(V_y) \in \mathcal{N}(z)$ , da die Komponenten ja als stetig vorausgesetzt waren, also ist auch  $f^{-1}(U) \supset f^{-1}(U_x \times V_y)$  eine Umgebung von  $z$ .

**Übungsaufgabe 19.** Zeige, dass  $f: X \rightarrow Y$  stetig ist genau dann, wenn  $f$  stetig am Punkt  $x \in X$  für alle  $x \in X$  ist.

Wir hatten eine weitere Variante von Stetigkeit in Abschnitt 4 kennengelernt, die *Folgenstetigkeit*. Konvergenz von Folgen ist in topologischen Räumen ganz analog zur Konvergenz reeller Folgen definiert: Zunächst ist eine Folge  $(x_j)_{j \in \mathbb{N}}$  eine Abbildung  $\mathbb{N} \rightarrow X$ , die der natürlichen Zahl  $j$  den Punkt  $x_j \in X$  zuordnet. Wir sagen, die Folge  $(x_j)_{j \in \mathbb{N}}$  in  $X$  konvergiert gegen  $x_0$  wenn für jede Umgebung  $U$  von  $x_0$  ein  $N \in \mathbb{N}$  existiert,

sodass  $x_j \in U$  für alle  $j \geq N$  gilt. Eine Abbildung  $f: X \rightarrow Y$  ist folgenstetig am Punkt  $x$ , wenn für jede Folge  $(x_j)_{j \in \mathbb{N}}$  in  $X$ , welche gegen  $x \in X$  konvergiert, die Folge  $(f(x_j))_{j \in \mathbb{N}}$  gegen  $f(x)$  gilt.

**Übungsaufgabe 20.** Zeige, dass jede Abbildung, welche im Punkt  $x$  stetig ist, auch folgenstetig im Punkt  $x$  ist.

**Übungsaufgabe 21.** Wir versehen  $\mathbb{N}$  mit der diskreten Topologie und erklären eine Topologie auf dem Raum  $\mathbb{N} \cup \{\infty\}$  indem wir die Umgebungen von  $\infty$  als Komplemente endlicher Mengen in  $\mathbb{N}$  festlegen. Dies ergibt eine Topologie auf  $\mathbb{N}$ . Dann konvergiert die Folge  $x_j$  von Punkten im topologischen Raum  $X$  genau dann gegen  $x \in X$ , wenn die Funktion  $f: \mathbb{N} \cup \{\infty\} \rightarrow X$ ,  $f(j) = x_j$ ,  $f(\infty) = x$  stetig ist.

Grenzwerte von Folgen sind in allgemeinen topologischen Räumen nicht eindeutig. Eine Bedingung, welche die Eindeutigkeit garantiert, ist die folgende:

**Definition 21.** Ein topologischer Raum  $X$  ist ein *Hausdorffraum*, wenn für je zwei verschiedene Punkte  $x, y \in X$  Umgebungen von  $x$  und  $y$  existieren, welche sich nicht schneiden.

LEMMA 40. Sei  $X$  ein Hausdorffraum. Dann ist der Grenzwert einer konvergenten Folge  $(x_j)_{j \in \mathbb{N}}$  in  $X$  eindeutig bestimmt.

BEWEIS. Seien  $x$  ein Grenzwert der Folge  $(x_j)_{j \in \mathbb{N}}$ , und  $y \neq x$ ,  $y \in X$ . Dann gibt es Umgebungen  $U \in \mathcal{N}(x)$  und  $V \in \mathcal{N}(y)$  mit  $U \cap V = \emptyset$ . Da  $(x_j)_{j \in \mathbb{N}}$  gegen  $x$  konvergiert, gibt es ein  $N \in \mathbb{N}$  sodass  $x_j \in U$  für alle  $j \geq N$  gilt. Also ist  $y$  kein Grenzwert von  $(x_j)_{j \in \mathbb{N}}$ .  $\square$

**Übungsaufgabe 22.** Eine Menge  $X$  mit der trivialen Topologie ist Hausdorff genau dann, wenn  $X$  nur einen Punkt besitzt. Jede Menge  $X$  ist ein Hausdorffraum in der diskreten Topologie.  $\mathbb{R}$  mit der üblichen Topologie ist ein Hausdorffraum.

Die Grösse von Topologien ist sehr wichtig. Eine der wichtigsten Eigenschaften, die wir für Umgebungen reeller Zahlen verwendet haben, ist die Tatsache, dass es abzählbare Umgebungsbasen gibt, wie zum Beispiel die Basis

$$\mathfrak{B}(x) = \{B_{1/n}(x) : n \in \mathbb{N}\}.$$

Wenn eine Topologie “zu gross” wird, so kann man nicht hoffen, dass man Eigenschaften von Funktionen bzw. Mengen mit Folgen austesten kann. Wir definieren deswegen:

**Definition 22.** Ein topologischer Raum  $X$  erfüllt das *erste Abzählbarkeitsaxiom*, wenn jeder Punkt eine abzählbare Umgebungsbasis besitzt.

LEMMA 41. Sei  $X$  ein topologischer Raum,  $Y \subset X$ . Wenn es eine Folge  $(y_j)_{j \in \mathbb{N}}$  in  $Y$  gibt, die gegen  $x$  konvergiert, so ist  $x$  im Abschluss von  $Y$ . Wenn  $X$  das erste Abzählbarkeitsaxiom erfüllt, so ist ein Punkt  $x \in X$  im Abschluss  $\bar{Y}$  von  $Y$  genau dann, wenn es eine Folge  $(y_j)_{j \in \mathbb{N}}$  in  $Y$  gibt, die gegen  $x$  konvergiert.

BEWEIS. Sei  $x$  ein Grenzwert einer Folge  $(y_j)_{j \in \mathbb{N}}$  in  $Y$ . Wenn  $x = y_j$  für ein  $j \in \mathbb{N}$  gilt, so ist die Folgerung trivial. Wenn  $x \neq y_j$  für alle  $j \in \mathbb{N}$  ist und  $U \in \mathcal{N}(x)$  eine beliebige Umgebung von  $x$  ist, so gibt es ein  $N \in \mathbb{N}$ , sodass sogar  $y_j \in U \cap Y$  für  $j \geq N$  gilt. Es folgt, dass  $x$  ein Häufungspunkt von  $Y$  ist, sich also im Abschluss von  $Y$  befindet.

Angenommen,  $x$  besitzt eine abzählbare Umgebungsbasis  $\mathfrak{B}(x) = \{U_j : j \in \mathbb{N}\}$ . Wir können zu einer abzählbaren absteigenden Umgebungsbasis  $\tilde{\mathfrak{B}}(x) = \{V_k : k \in \mathbb{N}\}$  übergehen (d.h.  $V_{k+1} \subset V_k$  für alle  $k \in \mathbb{N}$ ), indem wir

$$V_k = \bigcap_{j=1}^k U_j \in \mathcal{N}(x)$$

definieren.

Sei  $x$  im Abschluss von  $Y$ . Für jedes  $k \in \mathbb{N}$  wählen wir ein  $y_k \in Y \cap V_k$  (diese Menge ist nicht leer). Die so konstruierte Folge  $(y_k)_{k \in \mathbb{N}}$  konvergiert gegen  $x$ : Sei nämlich  $U \in \mathcal{N}(x)$  eine beliebige Umgebung von  $x$ . Dann gibt es ein  $N \in \mathbb{N}$  mit  $V_N \subset U$ . Es folgt  $y_k \in V_k \subset V_N \subset U$  für alle  $k \geq N$ .  $\square$

LEMMA 42. Wenn  $X$  das erste Abzählbarkeitsaxiom erfüllt, so ist jede folgenstetige Abbildung  $f: X \rightarrow Y$  stetig.

BEWEIS. Angenommen,  $f$  ist nicht stetig; dann gibt es einen Punkt  $y = f(x) \in Y$  und eine Umgebung  $U$  von  $f(x)$  sodass  $f^{-1}(U)$  keine Umgebung von  $x$  ist. Sei  $\mathcal{B}(x) = \{V_1, V_2, \dots\}$  eine abzählbare Umgebungsbasis von  $x$  mit  $V_1 \supset V_2 \supset \dots$  (wir haben im Beweis von Lemma 41 gezeigt, dass wir das annehmen können). Da  $f^{-1}(U)$  keine Umgebung von  $x$  ist, gibt es für jedes  $j$  einen Punkt  $x_j \in V_j \setminus f^{-1}(U)$ . Dann konvergiert  $(x_j)_{j \in \mathbb{N}}$  gegen  $x$ , aber  $f(x_j) \notin U$ , somit kann  $f(x_j)$  nicht gegen  $f(x)$  konvergieren.  $\square$

Eine weitere wichtige topologische Eigenschaft ist die Kompaktheit. Für allgemeine topologische Räume wird die Kompaktheit durch die Überdeckungskompaktheit definiert:

**Definition 23.** Ein topologischer Raum  $X$  ist kompakt, wenn jede Überdeckung von  $X$  durch offene Mengen eine endliche Teilüberdeckung enthält. Eine Teilmenge  $Y \subset X$  heisst kompakt, wenn sie kompakt in der Teilraumtopologie ist.

**Übungsaufgabe 23.** Zeige, dass  $Y \subset X$  genau dann kompakt ist, wenn jede Überdeckung  $U_\alpha$ ,  $\alpha \in A$  von  $Y$  durch offene Mengen in  $X$  eine endliche Teilüberdeckung enthält.

Um kompakte Mengen durch Folgen zu charakterisieren, benötigen wir eine stärkere Bedingung als das erste Abzählbarkeitsaxiom. Eine Menge  $\mathcal{B} = \{U_\alpha : \alpha \in A\}$  offener Mengen in  $X$  heisst eine *Basis* der Topologie von  $X$ , wenn jede offene Menge in  $X$  als Vereinigung von Mengen in  $\mathcal{B}$  geschrieben werden kann.

**Definition 24.**  $X$  erfüllt das *zweite Abzählbarkeitsaxiom*, wenn  $X$  eine abzählbare Basis seiner Topologie besitzt.

**Übungsaufgabe 24.** Zeige, dass ein Raum, welcher das zweite Abzählbarkeitsaxiom erfüllt, auch das erste Abzählbarkeitsaxiom erfüllt.

**Übungsaufgabe 25.**  $X$  heisst separabel, wenn  $X$  eine abzählbare dichte Teilmenge besitzt. Zeige, dass ein Raum welcher das zweite Abzählbarkeitsaxiom erfüllt, separabel ist. Gilt die Umkehrung?

**Übungsaufgabe 26.**  $X$  ist ein *Lindelöfraum*, wenn jede offene Überdeckung eine *abzählbare* Teilüberdeckung enthält. Zeige, dass jeder Raum, welcher das zweite Abzählbarkeitsaxiom erfüllt, ein Lindelöfraum ist.

LEMMA 43. Sei  $X$  ein topologischer Raum, welcher das zweite Abzählbarkeitsaxiom erfüllt. Dann ist  $Y \subset X$  genau dann kompakt, wenn für jede Folge  $(x_j)_{j \in \mathbb{N}}$  in  $Y$  eine Teilfolge  $(x_{j_k})_{k \in \mathbb{N}}$  existiert, welche gegen ein  $y_0 \in Y$  konvergiert.

BEWEIS. Angenommen, es gibt eine Folge  $(x_j)_{j \in \mathbb{N}}$  in  $Y$ , welche keine konvergente Teilfolge enthält. Dann kann sich diese Folge an keinem Punkt von  $Y$  häufen, das heisst, es gibt für jedes  $y \in Y$  eine offene Umgebung  $U_y \in \mathcal{N}(y)$ , welche nur endlich viele Punkte aus  $\{x_j\}$  enthält (hier verwenden wir eigentlich nur das erste Abzählbarkeitsaxiom). Da  $(U_y)_{y \in Y}$  eine endliche Teilüberdeckung hat, ist die Wertemenge von  $x_j$  endlich, was ein Widerspruch ist.

Erfülle nun  $Y$  die Bedingung im Lemma, und sei  $U_\alpha$ ,  $\alpha \in A$  eine offene Überdeckung von  $Y$ . Da  $X$  das zweite Abzählbarkeitsaxiom erfüllt, können wir annehmen, dass  $A = \mathbb{N}$  abzählbar ist. Angenommen, die Überdeckung  $U_j$  hat keine endliche Teilüberdeckung. Dann wählen wir eine Folge  $x_j$  mit der Eigenschaft, dass  $x_j \in Y \setminus \bigcup_{i=1}^j U_i$ . Sei  $y_0$  ein Grenzwert einer Teilfolge dieser Folge  $x_j$ . Dann gibt es eine offene Menge  $U_N$  mit  $y_0 \in U_N$ ; nach der Konstruktion von  $x_j$  ist aber  $x_j \notin U_N$  für  $j \geq N$ , ein Widerspruch.  $\square$

LEMMA 44. Sei  $X$  ein kompakter topologischer Raum,  $f: X \rightarrow Y$  stetig. Dann ist  $f(X)$  kompakt. Insbesondere nimmt jede stetige reelle Funktion auf einem kompakten topologischen Raum ihr Minimum und ihr Maximum an.

BEWEIS. Seien  $(V_\alpha)_{\alpha \in A}$  eine offene Überdeckung von  $f(X)$ . Dann ist  $(f^{-1}(V_\alpha))_{\alpha \in A}$  eine offene Überdeckung von  $X$ . Also gibt es endlich viele Indizes  $\alpha_1, \dots, \alpha_N$  mit  $X = \bigcup_{j=1}^N f^{-1}(V_{\alpha_j})$ , und  $(V_{\alpha_j})_{j=1}^N$  ist eine endliche Teilüberdeckung von  $(V_\alpha)_{\alpha \in A}$ .  $\square$

## 2. Vollständigkeit im Allgemeinen: Metrische Räume

Um den Begriff der Vollständigkeit allgemeiner zu fassen, ist es von Vorteil nicht von allgemeinen topologischen Räumen auszugehen. Der Grund liegt darin, dass der Begriff der Cauchyfolge voraussetzt, dass die "Nähe" von Punkten an verschiedenen Stellen vergleichbar ist. Dies verlangt eine zusätzliche Struktur, und metrische Räume eignen sich hervorragend dafür, den Vollständigkeitsbegriff zu erläutern.

**Definition 25.** Ein metrischer Raum ist eine Menge  $X$  zusammen mit einer Abbildung (der Metrik)  $d: X \times X \rightarrow [0, \infty)$ , welche die folgenden Eigenschaften erfüllt:

- (M1)  $d(x, y) = d(y, x)$ ;
- (M2)  $d(x, y) = 0$  genau dann, wenn  $x = y$ ;
- (M3)  $d(x, z) \leq d(x, y) + d(y, z)$  (die *Dreiecksungleichung*).

**Bemerkung 6.** Wie schon zuvor beim topologischen Raum, werden wir nicht  $(X, d)$  für einen metrischen Raum schreiben, sondern einfach  $X$ , wenn die Metrik klar ist.

Eine Metrik induziert eine Topologie auf  $X$ , indem wir als Basis dieser Topologie die *offenen (metrischen) Bälle*  $B_\varepsilon(x) = \{y \in X : d(x, y) < \varepsilon\}$  festsetzen. Eine Umgebungsbasis ist dann durch  $\mathcal{B}(x) = \{B_\varepsilon(x) : \varepsilon > 0\}$  gegeben.

**Beispiel 24.** Wir können auf jeder Menge  $X$  die diskrete Metrik definieren:  $d(x, x) = 0$ ,  $d(x, y) = 1$ , wenn  $y \neq x$ . Zeige, dass die diskrete Metrik die diskrete Topologie induziert.

**Beispiel 25.** Sei  $M$  eine Menge. Wir bezeichnen mit  $B(M)$  die Menge der beschränkten reellwertigen Funktionen auf  $M$ . Die Festsetzung

$$d(f, g) = \sup_{m \in M} |f(m) - g(m)|$$

definiert eine Metrik auf  $M$ . (Wir werden diese Metrik für den (Teil)raum der stetigen Funktionen auf einer kompakten später noch genauer beleuchten).

**Übungsaufgabe 27.** Sei  $X$  ein metrischer Raum mit Metrik  $d$ . Zeige:

- (1) Eine Folge  $x_j$  konvergiert gegen  $x \in X$  genau dann, wenn  $d(x_j, x) \rightarrow 0$  ( $j \rightarrow \infty$ ).
- (2) Die Topologie auf  $X$  erfüllt das erste Abzählbarkeitsaxiom.
- (3)  $X$  ist separabel genau dann, wenn  $X$  das zweite Abzählbarkeitsaxiom erfüllt.
- (4)  $d: X^2 \rightarrow [0, \infty)$  ist eine stetige Abbildung ( $X^2$  wird hier mit der Produkttopologie versehen).
- (5) Sei  $f: [0, \infty) \rightarrow [0, \infty)$  eine strikt monoton wachsende, stetige, konkave Funktion (d.h.  $f((1-t)x + ty) \geq (1-t)f(x) + tf(y)$ ) mit  $f(0) = 0$ . Dann ist  $d_f(x, y) = f(d(x, y))$  auch eine Metrik auf  $X$ , und die Topologien, welche durch  $d$  und  $d_f$  induziert werden, sind gleich.

Ein topologischer Raum heisst *metrisierbar*, wenn eine Metrik existiert, welche die gegebene Topologie induziert. Wegen des vorangehenden Beispiels ist so eine Metrik alles andere als eindeutig, im allgemeinen definieren viele verschiedene Metriken dieselbe Topologie. Eine hinreichende Bedingung dafür, dass zwei Metriken dieselbe Topologie induzieren, ist ihre *Vergleichbarkeit*.

**Definition 26.** Seien  $d$  und  $\tilde{d}$  Metriken auf der Menge  $X$ . Wir sagen, dass  $d$  mit  $\tilde{d}$  *vergleichbar* ist, wenn es Konstanten  $c, C > 0$  gibt, sodass

$$cd(x, y) \leq \tilde{d}(x, y) \leq Cd(x, y)$$

für alle  $x, y \in X$  ist.

**Bemerkung 7.** Wenn  $d$  mit  $\tilde{d}$  vergleichbar ist, so ist auch  $\tilde{d}$  mit  $d$  vergleichbar, und wir sagen manchmal auch, dass  $d$  und  $\tilde{d}$  vergleichbar sind. Vergleichbarkeit

**Übungsaufgabe 28.** Zeige, dass Vergleichbarkeit eine Äquivalenzrelation auf der Menge aller Metriken auf  $X$  ist.

LEMMA 45. Wenn  $d$  und  $\tilde{d}$  vergleichbar sind, so sind die durch  $d$  und  $\tilde{d}$  induzierten Topologien auf  $X$  gleich.

BEWEIS. Wenn wir mit

$$B_\varepsilon(x) = \{y \in X : d(x, y) < \varepsilon\}, \quad \tilde{B}_\varepsilon(x) = \{y \in X : \tilde{d}(x, y) < \varepsilon\}$$

die metrischen Bälle bezeichnen, so sind die Umgebungssysteme für die durch die Metriken definierten Topologien durch

$$\begin{aligned} \mathcal{N}(x) &= \{U \subset X : \exists \varepsilon > 0, B_\varepsilon(x) \subset U\} \\ \tilde{\mathcal{N}}(x) &= \{U \subset X : \exists \varepsilon > 0, \tilde{B}_\varepsilon(x) \subset U\} \end{aligned}$$

definiert. Wir müssen zeigen, dass  $\mathcal{N}(x) = \tilde{\mathcal{N}}(x)$  ist; wegen Symmetriegründen genügt es zu zeigen, dass  $\mathcal{N}(x) \subset \tilde{\mathcal{N}}(x)$  gilt.

Sei also  $U \in \mathcal{N}(x)$ . Dann gibt es ein  $\varepsilon > 0$  mit  $B_\varepsilon(x) \subset U$ . Da  $d$  mit  $\tilde{d}$  vergleichbar ist, gibt es  $C > 0$  mit  $d(x, y) < C\tilde{d}(x, y)$ . Also gilt  $\tilde{B}_{\frac{\varepsilon}{C}} \subset B_\varepsilon(x) \subset U$ , und damit gilt  $\tilde{U} \in \mathcal{N}(x)$ .  $\square$

Stetige Abbildungen zwischen metrischen Räumen können durch die bekannte  $\varepsilon - \delta$ -Definition charakterisiert werden:

**LEMMA 46.** *Seien  $(X, d), (Y, e)$  metrische Räume. Eine Abbildung  $f: X \rightarrow Y$  ist stetig am Punkt  $x_0 \in X$  genau dann, wenn für jedes  $\varepsilon > 0$  ein  $\delta > 0$  existiert sodass  $e(f(x), f(x_0)) < \varepsilon$  für alle  $x \in X$  mit  $d(x, x_0) < \delta$  ist.*

**BEWEIS.** Sei zunächst die  $\varepsilon - \delta$ -Bedingung erfüllt. Wir müssen zeigen, dass jede offene Umgebung von  $f(x_0)$  ein Bild einer offenen Umgebung von  $x_0$  enthält. Wenn  $V \subset Y$  eine offene Umgebung von  $f(x_0)$  ist, so gibt es einen Ball  $B_\varepsilon(f(x_0)) = \{y \in Y: e(f(x_0), y) < \varepsilon\} \subset V$ . Nach Voraussetzung gibt es ein  $\delta > 0$  sodass  $e(f(x_0), f(x)) < \varepsilon$ , wenn  $d(x_0, x) < \delta$ ; das heisst aber, dass  $f(B_\delta(x_0)) \subset V$ , und  $B_\delta(x_0) = \{x \in X: d(x_0, x) < \delta\}$  ist definitionsgemäss eine offene Umgebung von  $x_0$ .

Für die Umkehrung sei  $\varepsilon > 0$ . Da  $B_\varepsilon(f(x_0))$  offen ist, gibt es eine offene Umgebung  $U$  von  $x_0$  mit  $f(U) \subset B_\varepsilon(f(x_0))$ , und da  $U$  eine offene Umgebung von  $x_0$  ist, ein  $\delta > 0$  mit  $B_\delta(x_0) \subset U$ . Also ist  $f(B_\delta(x_0)) \subset B_\varepsilon(f(x_0))$ .  $\square$

Zusammen mit der Tatsache, dass metrische Räume das erste Abzählbarkeitsaxiom erfüllen, und mit der Hilfe von Lemma 42 ergibt sich damit für Funktionen zwischen metrischen Räumen eine vollkommen analoge Charakterisierung der Stetigkeit wie in Theorem 21. Um Kompaktheit zu beschreiben, ist eine Version der *Beschränktheit* erforderlich.

**Definition 27.** Ein metrischer Raum  $X$  ist *beschränkt*, wenn das Bild von  $d$  beschränkt ist, also  $d(x, y) < R$  für ein  $R > 0$  und alle  $x, y \in X$  ist.  $X$  ist *total beschränkt* wenn für jedes  $r > 0$  eine endliche Menge  $\{x_1, \dots, x_n\} \subset X$  existiert sodass

$$X \subset \bigcup_{j=1}^n B_r(x_j).$$

**LEMMA 47.** *Ein total beschränkter Raum erfüllt das zweite Abzählbarkeitsaxiom.*

**BEWEIS.** Wir müssen zeigen, dass die Topologie, welche durch die Metrik induziert wird, eine abzählbare Basis besitzt. Für jedes  $N \in \mathbb{N}$  können wir eine endliche Menge  $E_n \subset X$  wählen, sodass

$$X \subset \bigcup_{x \in E_n} B_{1/n}(x)$$

ist. Wir behaupten, dass die Menge

$$\mathcal{B} = \{B_{1/n}(x): x \in E_n, n \in \mathbb{N}\}$$

eine Basis der Topologie von  $X$  ist. Sei  $x \in X$  beliebig, und  $U$  eine offene Umgebung von  $x$ . Es genügt zu zeigen, dass es ein  $n \in \mathbb{N}$  gibt und ein  $y \in E_n$  gibt, welches  $x \in B_{1/n}(y) \subset U$  erfüllt. Sei  $\varepsilon > 0$  mit  $U \supset B_\varepsilon(x)$  gewählt. Wir wählen nun  $n$  gross genug, dass  $2/n < \varepsilon$ , und  $y \in E_n$  sodass  $x \in B_{1/n}(y)$ . Dann können wir für alle  $z \in B_{1/n}(y)$  wie folgt abschätzen:

$$d(x, z) \leq d(x, y) + d(y, z) \leq \frac{1}{n} + \frac{1}{n} < \varepsilon,$$

und damit  $B_{1/n}(y) \subset U$ .  $\square$

Insbesondere ist nach Lemma 43 für einen total beschränkten metrischen Raum die Folgenkompaktheit mit der Kompaktheit im Sinne von Definition 23 äquivalent. Wir werden nun Kompaktheit für metrische Räume im Sinn des Satzes von Heine-Borel charakterisieren. Dazu benötigen wir eine Bedingung, die uns die Existenz von Grenzwerten garantiert (wir erinnern an den Beweis von Theorem 33, wo wir die Existenz des  $\limsup$  verwendeten). Hier tritt die Vollständigkeit wieder auf.

**Definition 28.** Eine Folge  $(x_n)_{n \in \mathbb{N}}$  in  $X$  ist eine *Cauchyfolge*, wenn für jedes  $\varepsilon > 0$  ein  $N = N(\varepsilon) \in \mathbb{N}$  existiert, sodass  $d(x_n, x_m) < \varepsilon$  für alle  $n, m \geq N$  ist.

**Übungsaufgabe 29.** Zeige, dass eine konvergente Folge eine Cauchyfolge ist.

**Definition 29.** Ein metrischer Raum  $X$  ist *vollständig*, wenn jede Cauchyfolge in  $X$  konvergiert.

LEMMA 48. *Ein metrischer Raum  $X$  ist kompakt genau dann, wenn er total beschränkt und vollständig ist.*

BEWEIS. Sei zunächst  $X$  kompakt. Dann ist  $X$  auch total beschränkt, und damit nach Lemma 47 folgenkompakt. Sei  $x_j$  eine Cauchyfolge in  $X$ . Dann gibt es eine Teilfolge  $x_{j_k}$ , welche gegen  $x \in X$  konvergiert. Wir behaupten, dass dann auch  $\lim_{j \rightarrow \infty} x_j = x$ . Sei also  $\varepsilon > 0$  beliebig. Dann gibt es ein  $N \in \mathbb{N}$  sodass  $d(x_m, x_n) < \varepsilon$  für  $m, n \geq N$  ist, und ein  $K \in \mathbb{N}$  sodass  $d(x_{j_k}, x) < \varepsilon$  für  $k \geq K$ . Ist nun  $n \geq \max(N, j_K)$ , so ist

$$d(x_n, x) \leq d(x_n, x_{j_K}) + d(x_{j_K}, x) \leq 2\varepsilon;$$

also ist  $(x_n)$  konvergent und damit  $X$  vollständig.

Sei nun  $X$  total beschränkt und vollständig. Damit ist wiederum nach Lemma 47 Folgenkompaktheit äquivalent zur Kompaktheit, und es genügt zu zeigen, dass jede Folge in  $X$  eine konvergente Teilfolge besitzt. Sei also  $x_n$  eine Folge in  $X$ . Wir behaupten, dass es eine Teilfolge  $x_{n_k}$  gibt, welche Cauchy ist.

Für jedes  $1/k$  gibt es endlich viele Bälle  $B_{1/k}(y)$ ,  $y \in E_k$  welche  $X$  überdecken. Wir extrahieren nun induktiv eine Folge aus  $x_j$  wie folgt. In einem der Bälle  $B_1(y)$ ,  $y \in E_1$ , gibt liegen  $x_j$  für unendlich viele  $j$ . Wir definieren nun  $j_1$  durch alle diese  $j$ . Dann extrahieren wir aus der Teilfolge  $x_{j_1}$  eine Teilfolge  $x_{j_2}$ , welche alle in einem Ball vom Radius  $1/2$  liegt, und so weiter. Wir behaupten, dass die Diagonalfolge  $x_{j_j}$  eine Cauchyfolge ist. Sei  $\varepsilon > 0$  beliebig, und  $j \in \mathbb{N}$  large enough so that  $j \leq 2\varepsilon$ . Wenn  $k, \ell \geq j$  ist, so ist  $x_{k_k}, x_{\ell_\ell} \in B_{1/j}(x_{j_j})$ . Damit ist  $d(x_{k_k}, x_{\ell_\ell}) \leq \varepsilon$ , also ist  $x_{m_m}$  eine Cauchyfolge.  $\square$

Wir können in einem metrischen Raum natürlich auch den Begriff der gleichmässigen Stetigkeit definieren:

**Definition 30.** Seien  $(X, d)$  und  $(Y, e)$  metrische Räume. Eine Abbildung  $f: X \rightarrow Y$  ist *gleichmässig stetig*, wenn für jedes  $\varepsilon > 0$  ein  $\delta > 0$  existiert, sodass  $e(f(x), f(y)) < \varepsilon$  wenn  $d(x, y) < \delta$ .

SATZ 30. *Sei  $f: X \rightarrow Y$  eine stetige Abbildung. Wenn  $X$  kompakt ist, so ist  $f$  gleichmässig stetig.*

BEWEIS. Sei  $\varepsilon > 0$  beliebig. Da  $f$  stetig ist, gibt es für jedes  $x \in X$  ein  $\delta(x)$  sodass  $e(f(x), f(y)) < \varepsilon$  wenn  $d(x, y) < 2\delta(x)$ . Die Bälle  $B_{\delta(x)}(x)$  überdecken  $X$ , also existieren endlich viele  $x_j$  sodass schon  $B_{\delta(x_j)}(x_j)$  ganz  $X$  überdecken. Wenn wir nun  $\delta = \min_j \delta(x_j) > 0$  wählen, so gilt für beliebige  $x, y \in X$ , dass aus  $x \in B_{\delta(x_j)}(x_j)$  auch  $y \in B_{2\delta(x_j)}(x_j)$ , da  $d(y, x_j) \leq d(y, x) + d(x, x_j) \leq \delta + \delta(x_j) \leq 2\delta(x_j)$ . Damit ist

$$e(f(x), f(y)) \leq e(f(x), f(x_j)) + e(f(x_j), f(y)) \leq 2\varepsilon,$$

und damit  $f$  gleichmässig stetig.  $\square$

In einem vollständigen metrischen Raum gilt der wichtige *Banach'sche Fixpunktsatz*.

**Definition 31.** Sei  $X$  ein metrischer Raum. Eine Abbildung  $f: X \rightarrow X$  ist eine *Kontraktion*, wenn es ein  $q < 1$  gibt, sodass  $d(f(x), f(y)) < qd(x, y)$  für alle  $x, y \in X$  ist.

**Bemerkung 8.** Insbesondere ist jede Kontraktion (sogar gleichmässig) stetig.

SATZ 31. *Sei  $X$  ein vollständiger metrischer Raum, und  $f: X \rightarrow X$  eine Kontraktion. Dann besitzt  $f$  einen Fixpunkt: Es gibt ein  $x \in X$  mit  $f(x) = x$ .*

BEWEIS. Sei  $x_0 \in X$  beliebig. Wir definieren eine Folge  $x_j$  durch  $x_{j+1} = f(x_j)$ . Wenn  $x_j$  konvergiert,  $\lim_j x_j = x$ , so folgt aus der Stetigkeit von  $f$  dass  $x = f(x)$ . Wir zeigen nun, dass  $x_j$  eine Cauchyfolge ist:

$$\begin{aligned} d(x_m, x_n) &\leq \sum_{j=1}^{n-m+1} d(x_{m+j-1}, x_{m+j}) \\ &= \sum_{j=1}^{n-m+1} d(f(x_{m+j-2}), f(x_{m+j-1})) \\ &\leq \sum_{j=1}^{n-m+1} qd(x_{m+j-2}, x_{m+j-1}) \\ &\dots \\ &\leq (q^m + q^{m+1} + \dots + q^n) d(x_1, x_0) \\ &\leq \frac{q^m}{1-q} d(x_1, x_0). \end{aligned}$$

Da  $q < 1$  ist, strebt der letzte Teil dieser Ungleichungskette gegen 0, wenn  $m \rightarrow \infty$ ; also ist  $x_j$  eine Cauchyfolge.  $\square$

### 3. Abstände in Vektorräumen

Bevor wir uns mit dem allgemeinen Abstandsbegriff in Vektorräumen auseinandersetzen, wollen wir den zunächst für uns wichtigsten Fall, den Abstand im  $\mathbb{R}^n$ , detailliert betrachten.

**3.1. Die Cauchy–Schwarz Ungleichung und der euklidische Abstand.** Der gängigste Abstandsbegriff im cartesischen Produkt  $\mathbb{R}^n$  beruht auf der *euklidischen Norm* eines Punktes  $x = (x_1, \dots, x_n)$ , mit  $x_j \in \mathbb{R}$  für  $j = 1, \dots, n$ , welche durch

$$\|x\|_2 = \sqrt{\sum_{j=1}^n x_j^2}$$

gegeben ist. Wir definieren das *innere Produkt* oder *Skalarprodukt* von zwei Punkten  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  und  $y = (y_1, \dots, y_n) \in \mathbb{R}^n$  durch

$$\langle x, y \rangle = \sum_{j=1}^n x_j y_j;$$

damit ist  $\|x\|_2^2 = \langle x, x \rangle$ . Das innere Produkt hat Symmetrie- und Linearitätseigenschaften, das heisst, es gilt

$$\langle x, y \rangle = \langle y, x \rangle, \quad \langle x, y + \lambda y' \rangle = \langle x, y \rangle + \lambda \langle x, y' \rangle, \quad \langle x + \lambda x', y \rangle = \langle x, y \rangle + \lambda \langle x', y \rangle, \quad x, x', y, y' \in \mathbb{R}^n, \lambda \in \mathbb{R}.$$

Um von der Norm–ähnlich wie vom Absolutbetrag–einen Abstandsbegriff abzuleiten, benötigen wir die *Dreiecksungleichung*, das heisst, wir müssen

$$\|x + y\|_2 \leq \|x\|_2 + \|y\|_2, \quad x, y \in \mathbb{R}^n,$$

zeigen. Folgen wir der Argumentation von Lemma 18, so berechnen wir

$$\begin{aligned} \|x + y\|_2^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\ &= \|x\|_2^2 + 2\langle x, y \rangle + \|y\|_2^2, \end{aligned}$$

und benötigen als entscheidende Ungleichung somit  $\langle x, y \rangle \leq \|x\|_2 \|y\|_2$ .



LEMMA 49 (Die Cauchy-Schwarz Ungleichung). Seien  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  und  $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ . Dann gilt

$$\left| \sum_{j=1}^n x_j y_j \right| \leq \sqrt{\sum_{j=1}^n x_j^2} \sqrt{\sum_{j=1}^n y_j^2}.$$

Falls  $x \neq 0$  ist, so gilt Gleichheit genau dann, wenn  $y = \lambda x$  für ein  $\lambda \in \mathbb{R}$  ist.

BEWEIS. Für jedes  $j = 1, \dots, n$  gilt  $(x_j - y_j)^2 \geq 0$ , also  $x_j y_j \leq \frac{x_j^2 + y_j^2}{2}$ . Wir addieren diese Ungleichungen und erhalten

$$\sum_j x_j y_j \leq \frac{1}{2} \left( \sum_j x_j^2 + \sum_j y_j^2 \right).$$

Ist nun  $\sum_j x_j^2 = \sum_j y_j^2 = 1$ , so folgt  $\sum_j x_j y_j \leq 1$ .

Wir verwenden diese Beobachtung nun, um unsere ursprüngliche Ungleichung zu zeigen. Dazu nehmen wir ohne Beschränkung der Allgemeinheit an, dass  $x \neq 0$  und  $y \neq 0$  gilt (die Ungleichung ist andernfalls trivialerweise erfüllt), und definieren  $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$  und  $\eta = (\eta_1, \dots, \eta_n) \in \mathbb{R}^n$  durch

$$\xi_k = \frac{x_k}{\sqrt{\sum_{j=1}^n x_j^2}}, \quad \eta_k = \frac{y_k}{\sqrt{\sum_{j=1}^n y_j^2}}, \quad k = 1, \dots, n.$$

Damit gilt

$$\sum_{k=1}^n \xi_k^2 = \sum_{k=1}^n \frac{x_k^2}{\sum_{j=1}^n x_j^2} = \frac{1}{\sum_{j=1}^n x_j^2} \sum_{k=1}^n x_k^2 = 1,$$

und

$$\sum_{k=1}^n \eta_k^2 = \sum_{k=1}^n \frac{y_k^2}{\sum_{j=1}^n y_j^2} = \frac{1}{\sum_{j=1}^n y_j^2} \sum_{k=1}^n y_k^2 = 1.$$

Wir können damit unsere Beobachtung von vorher auf  $\xi$  und  $\eta$  anwenden, und erhalten  $\sum_k \xi_k \eta_k \leq 1$ , also also

$$\sum_{j=1}^n x_j y_j \leq \sqrt{\sum_{j=1}^n x_j^2} \sqrt{\sum_{j=1}^n y_j^2}.$$

Falls die linke Seite dieser Ungleichung nicht positiv ist, so können wir die vorangegangene Argumentation auf die Vektoren  $-x$  und  $y$  anwenden und erhalten so

$$\left| \sum_{j=1}^n x_j y_j \right| = - \sum_{j=1}^n x_j y_j = \sum_{j=1}^n (-x_j) y_j \leq \sqrt{\sum_{j=1}^n x_j^2} \sqrt{\sum_{j=1}^n y_j^2}.$$

Gleichheit gilt, wenn in den zu Grunde gelegten Ungleichungen  $2\xi_k \eta_k \leq \xi_k^2 + \eta_k^2$  Gleichheit herrscht, was genau dann gilt, wenn  $\xi_k = \eta_k$  für alle  $k = 1, \dots, n$  ist. Wenn man nun die Schritte des vorigen Beweises weiter verfolgt, so bedeutet das

$$y = \pm \frac{\sqrt{\sum_{j=1}^n y_j^2}}{\sqrt{\sum_{j=1}^n x_j^2}} x = \lambda x.$$

□

LEMMA 50. Für  $x, y \in \mathbb{R}^n$  gilt  $\|x + y\|_2 \leq \|x\|_2 + \|y\|_2$ . Wenn  $x \neq 0$ , so gilt Gleichheit genau dann, wenn  $y = \lambda x$  mit  $\lambda \geq 0$ .

BEWEIS. Wir berechnen wie oben

$$\begin{aligned}\|x + y\|_2^2 &= \|x\|_2^2 + 2\langle x, y \rangle + \|y\|_2^2 \\ &\leq \|x\|_2^2 + 2|\langle x, y \rangle| + \|y\|_2^2 \\ &\leq \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2 \\ &= (\|x\|_2 + \|y\|_2)^2.\end{aligned}$$

Nach der Bedingung für Gleichheit in Lemma 49 gilt zunächst Gleichheit in der dritten Zeile der oberen Ungleichungskette, wenn  $y = \lambda x$  für ein  $\lambda \in \mathbb{R}$  gilt. In der zweiten Zeile herrscht dann Gleichheit, wenn  $\lambda \geq 0$  ist.  $\square$

Wir können nun durch

$$d_2(x, y) = \|x - y\|_2, \quad x, y \in \mathbb{R}^n,$$

eine Metrik auf  $\mathbb{R}^n$  definieren, d.h.  $d_2$  erfüllt die Bedingungen (M1)–(M3), welche wir zur Erinnerung noch einmal aufschreiben:

LEMMA 51.  $d_2: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  erfüllt folgende Eigenschaften:

- i)  $d_2(x, y) \geq 0$ , und  $d_2(x, y) = 0$  genau dann, wenn  $x = y$ ;
- ii)  $d_2(x, y) = d_2(y, x)$ ;
- iii)  $d_2(x, z) \leq d_2(x, y) + d_2(y, z)$ .

BEWEIS. Die Nichtnegativität von  $d_2$  ist offensichtlich. Wenn  $d_2(x, y) = 0$ , so ist

$$\sum_j (x_j - y_j)^2 = 0,$$

also  $x_j = y_j$  für  $j = 1, \dots, n$ .

Wir beachten, dass für  $\lambda \in \mathbb{R}$  und  $\xi \in \mathbb{R}^n$  die Gleichheit  $|\lambda| \|\xi\|_2 = \|\lambda\xi\|$  gilt, also ist

$$d_2(x, y) = \|x - y\|_2 = |-1| \|x - y\|_2 = \|y - x\|_2 = d_2(y, x).$$

Für die Dreiecksungleichung berechnen wir

$$\begin{aligned}d_2(x, z) &= \|x - z\|_2 \\ &= \|x - y + y - z\|_2 \\ &\leq \|x - y\|_2 + \|y - z\|_2 \\ &= d_2(x, y) + d_2(y, z).\end{aligned}$$

$\square$

**Übungsaufgabe 30.** Stelle fest, für welche  $y$  die Gleichheit  $d_2(x, z) = d_2(x, y) + d_2(y, z)$  gilt.

**3.2. Weitere Abstände und der allgemeine Abstands begriff.** Wir wollen nun zwei weitere Abstände auf  $\mathbb{R}^n$  betrachten, die wiederum auf Normen (wie der euklidische Abstand) beruhen. Diese sind definiert durch

$$(2) \quad \|x\|_1 = \sum_{j=1}^n |x_j|, \quad \|x\|_\infty = \max_j |x_j|.$$

Die notwendigen Dreiecksungleichungen sind für diese Normen einfache Folgerungen aus der Dreiecksungleichung Lemma 18: Zunächst ist

$$\begin{aligned}\|x + y\|_1 &= \sum_{j=1}^n |x_j + y_j| \\ &\leq \sum_{j=1}^n (|x_j| + |y_j|) \\ &= \sum_{j=1}^n |x_j| + \sum_{j=1}^n |y_j| \\ &= \|x\|_1 + \|y\|_1.\end{aligned}$$

Für die Maximumsnorm gilt nun für jedes  $j = 1, \dots, n$

$$|x_j + y_j| \leq |x_j| + |y_j| \leq \|x\|_\infty + \|y\|_\infty,$$

also auch

$$\|x + y\|_\infty = \max_j |x_j + y_j| \leq \|x\|_\infty + \|y\|_\infty.$$

Der Übergang von diesen Normen zu einem Abstand benötigt nur einige wenige ihrer Eigenschaften, welche wir in der folgenden Definition einer allgemeinen Norm zusammenfassen.

**Definition 32.** Wir sagen, eine Abbildung  $\mathbb{R}^n \ni x \mapsto \|x\| \in \mathbb{R}$  ist eine *Norm* auf  $\mathbb{R}^n$ , wenn sie folgende Eigenschaften erfüllt:

(N1)  $\|x\| \geq 0$  für alle  $x \in \mathbb{R}^n$ ,  $\|x\| = 0$  genau dann, wenn  $x = 0$ .

(N2)  $\|\lambda x\| = |\lambda| \|x\|$  für alle  $\lambda \in \mathbb{R}$  und  $x \in \mathbb{R}^n$ .

(N3)  $\|x + y\| \leq \|x\| + \|y\|$  für alle  $x, y \in \mathbb{R}^n$ .

LEMMA 52. Sei  $\|\cdot\|$  eine Norm auf  $\mathbb{R}^n$ . Dann wird durch  $d(x, y) = \|x - y\|$  eine Metrik auf  $\mathbb{R}^n$  definiert.

BEWEIS. Dass die durch  $d(x, y) = \|x - y\|$  definierte Abbildung (M1) erfüllt, folgt direkt aus (N1). (M2) ist erfüllt, da

$$d(x, y) = \|x - y\| = |-1| \|x - y\| = \|y - x\| = d(y, x).$$

Die Dreiecksungleichung der Metrik (M3) folgt aus der Dreiecksungleichung der Norm (N3):

$$\begin{aligned}d(x, z) &= \|x - z\| \\ &= \|x - y + y - z\| \\ &\leq \|x - y\| + \|y - z\| \\ &= d(x, y) + d(y, z).\end{aligned}$$

□

Bevor wir weitergehen, erinnern wir noch an die *umgekehrte Dreiecksungleichung*:

LEMMA 53. Sei  $\|\cdot\|$  eine Norm auf  $\mathbb{R}^n$ . Dann gilt

$$|\|v\| - \|w\|| \leq \|v - w\|.$$

Insbesondere ist  $\|\cdot\|$  stetig in Bezug auf die von  $\|\cdot\|$  induzierte Metrik.

BEWEIS. Es gilt

$$\begin{aligned}\|v\| &= \|v - w + w\| \leq \|v - w\| + \|w\| \\ \|w\| &= \|w - v + v\| \leq \|w - v\| + \|v\|,\end{aligned}$$

und damit sowohl  $\|v\| - \|w\| \leq \|v - w\|$  als auch  $\|w\| - \|v\| \leq \|v - w\|$ . Wir folgern, dass  $|\|v\| - \|w\|| \leq \|v - w\|$ .

Um zu sehen, dass  $\|\cdot\|$  stetig ist, wenden wir Lemma 46 an; die umgekehrte Dreiecksungleichung impliziert, dass die Wahl  $\delta = \varepsilon$  zulässig ist. □

Wir haben damit neben der euklidischen Metrik noch die folgenden Metriken definiert:

$$d_1(x, y) = \|x - y\|_1, \quad d_\infty(x, y) = \|x - y\|_\infty.$$

Wir können die definierten Normen miteinander *vergleichen*:

LEMMA 54. *Es gilt für  $x \in \mathbb{R}^n$ :*

$$\begin{aligned} \frac{1}{\sqrt{n}} \|x\|_1 &\leq \|x\|_2 \leq \|x\|_1 \\ \|x\|_\infty &\leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty \\ \|x\|_\infty &\leq \|x\|_1 \leq n \|x\|_\infty. \end{aligned}$$

BEWEIS. Wir beginnen mit dem Vergleich der 1- und der 2-Norm. Nach der Cauchy-Schwarz Ungleichung (Lemma 49) gilt

$$\begin{aligned} \|x\|_1 &= \sum_{j=1}^n |x_j| \\ &= \sum_{j=1}^n 1 \cdot |x_j| \\ &\leq \sqrt{\sum_{j=1}^n 1^2} \sqrt{\sum_{j=1}^n |x_j|^2} = \sqrt{n} \|x\|_2 \end{aligned},$$

was die erste Ungleichung impliziert. Auf der anderen Seite ist

$$\begin{aligned} \|x\|_1^2 &= \left( \sum_{j=1}^n |x_j| \right)^2 \\ &= \sum_{j,k=1}^n |x_j| |x_k| \\ &\geq \sum_{j=1}^n |x_j|^2 \\ &= \|x\|_2^2, \end{aligned}$$

was nach der Monotonie der Wurzelfunktion die zweite Ungleichung impliziert.

Um die  $\infty$ - und die 2-Norm zu vergleichen, bemerken wir, dass

$$|x_k| \leq \sqrt{\sum_{j=1}^n |x_j|^2}, \quad k = 1, \dots, n$$

gilt, also auch

$$\|x\|_\infty = \max_{k=1}^n |x_k| \leq \sqrt{\sum_{j=1}^n |x_j|^2} = \|x\|_2.$$

Die zweite Ungleichung folgt wegen

$$\|x\|_2 = \sqrt{\sum_{j=1}^n |x_j|^2} \leq \sqrt{\sum_{j=1}^n \|x\|_\infty^2} = \sqrt{n} \|x\|_\infty.$$

Der Vergleich der 1- und der  $\infty$ -Norm verläuft in ähnlichen Bahnen. Es ist zunächst

$$|x_k| \leq \sum_{j=1}^n |x_j|, \quad k = 1, \dots, n,$$

also auch

$$\|x\|_\infty = \max_{k=1}^n |x_k| \leq \sum_{j=1}^n |x_j| = \|x\|_1.$$

Auf der anderen Seite ist

$$\|x\|_1 = \sum_{j=1}^n |x_j| \leq \sum_{j=1}^n \|x\|_\infty = n \|x\|_\infty.$$

□

**SATZ 32.** Die von den Metriken  $d_1$ ,  $d_2$ , und  $d_\infty$  induzierten Topologien auf  $\mathbb{R}^n$  sind gleich, und stimmen mit der Produkttopologie überein. Insbesondere gilt: Eine Folge  $(x^j)_{j \in \mathbb{N}}$  von Punkten  $x^j = (x_1^j, \dots, x_n^j) \in \mathbb{R}^n$  konvergiert gegen  $x^0 \in \mathbb{R}^n$  genau dann, wenn für jedes  $k = 1, \dots, n$  die Folge  $(x_k^j)_{j \in \mathbb{N}}$  reeller Zahlen konvergiert.

**BEWEIS.** Wir haben in Lemma 54 gezeigt, dass die Normen  $\|\cdot\|_1$ ,  $\|\cdot\|_2$ , und  $\|\cdot\|_\infty$  miteinander vergleichbar sind. Also sind auch die Metriken  $d_1$ ,  $d_2$ , und  $d_\infty$  miteinander vergleichbar, und sie induzieren nach Lemma 45 dieselbe Topologie

Um zu sehen, dass die so beschriebene Topologie mit der Produkttopologie übereinstimmt, sei  $U \in \mathcal{N}^P(x)$  eine Umgebung von  $x = (x_1, \dots, x_n)$  in der Produkttopologie. Dann gibt es für jedes  $j = 1, \dots, n$  eine Umgebung  $U_j$  von  $x_j$  in  $\mathbb{R}$  sodass  $U \subset U_1 \times \dots \times U_n$ . Für jedes solche  $j$  gibt es damit ein  $\varepsilon_j > 0$  mit  $(x_j - \varepsilon_j, x_j + \varepsilon_j) \subset U_j$ . Also ist für  $\varepsilon = \min_{j=1}^n \varepsilon_j$  der metrische Ball  $B_\varepsilon^\infty(x) \subset U$ , und damit  $U \in \mathcal{N}^\infty(x)$ .

Ein ganz ähnliches Argument (jeder metrische Ball im Sinne von  $d_\infty$  ist eine Produktumgebung) liefert, dass  $\mathcal{N}^\infty(x) \subset \mathcal{N}^P(x)$ . □

Der obige Beweis regt an, über Metriken, welche von einer *beliebigen* Norm induziert werden, nachzudenken; stimmen die induzierten Topologien auch für solche mit der Produkttopologie überein? Bevor wir dies tun, benötigen (und beweisen) wir noch einige grundlegende Tatsachen über den  $\mathbb{R}^n$ .

**LEMMA 55.** Der metrische Raum  $(\mathbb{R}^n, d_2)$  ist vollständig.

**BEWEIS.** Sei  $(x^j)_{j \in \mathbb{N}}$  eine Folge in  $\mathbb{R}^n$ , und  $x^j = (x_1^j, \dots, x_n^j)$ . Für jedes  $k = 1, \dots, n$  ist  $(x_k^j)_{j \in \mathbb{N}}$  eine Folge reeller Zahlen; und wegen Lemma 54 ist

$$|x_k^j - x_k^\ell| \leq \|x^j - x^\ell\|_\infty \leq \|x^j - x^\ell\|_2,$$

also  $(x_k^j)_{j \in \mathbb{N}}$  eine Cauchyfolge in  $\mathbb{R}$  für jedes  $k = 1, \dots, n$ . Nach Theorem 9 gibt es also  $x_k^0 = \lim_{j \rightarrow \infty} x_k^j$ . An dieser Stelle können wir auf Theorem 32 verweisen, um zu sehen, dass  $(x^j)_{j \in \mathbb{N}}$  konvergiert, oder "zu Fuss" zeigen, dass  $\lim_{j \rightarrow \infty} x^j = x^0$ . Sei dazu  $\varepsilon > 0$  gegeben. Für jedes  $k = 1, \dots, n$  gibt es ein  $N_k$  sodass

$$|x_k^j - x_k^0| < \varepsilon, \quad j \geq N_k,$$

gilt. Für  $j \geq N_0 = \max\{N_1, \dots, N_n\}$  gilt damit

$$d_2(x^j, x^0) = \|x^j - x^0\|_2 \leq \|x^j - x^0\|_1 = \sum_{k=1}^n |x_k^j - x_k^0| < n\varepsilon.$$

Also konvergiert  $x^j$  gegen  $x^0$  im Sinne der Metrik  $d_2$ . □

**LEMMA 56.** Der metrische Raum  $(\mathbb{R}^n, d_2)$  erfüllt das zweite Abzählbarkeitsaxiom.

**BEWEIS.** Nachdem die Topologie von  $(\mathbb{R}^n, d_2)$  das erste Abzählbarkeitsaxiom erfüllt, genügt es zu zeigen, dass es eine abzählbare dichte Teilmenge  $D \subset \mathbb{R}^n$  gibt; wenn wir für jedes Element  $d$  dieser Teilmenge eine abzählbare Umgebungsbasis  $\mathcal{B}(d)$  offener Umgebungen wählen, so stellt sich

$$\mathcal{B} = \bigcup_{d \in D} \mathcal{B}(d)$$

als abzählbare Basis der Topologie heraus.

Wir behaupten, dass die abzählbare Menge  $\mathbb{Q}^n \subset \mathbb{R}^n$  dicht ist. Sei also  $U \subset \mathbb{R}^n$  offen (nicht leer); wir müssen zeigen, dass  $U \cap \mathbb{Q}^n \neq \emptyset$ . Sei  $x = (x_1, \dots, x_n) \in U$  beliebig. Dann gibt es ein  $\varepsilon > 0$  mit  $B_\varepsilon^\infty(x) \subset U$ .

Nachdem  $B_\varepsilon^\infty(x) = \prod_{j=1}^n (x_j - \varepsilon, x_j + \varepsilon)$  ist, gibt es nach Lemma 7 für jedes  $j = 1, \dots, n$  ein  $r_j \in \mathbb{Q}$  mit  $r_j \in (x_j - \varepsilon, x_j + \varepsilon)$ . Damit ist  $(r_1, \dots, r_n) \in \mathbb{Q}^n \cap U$ .  $\square$

**SATZ 33.**  $K \subset \mathbb{R}^n$  ist kompakt (in der von der Metrik  $d_2$  induzierten Topologie) genau dann, wenn  $K$  beschränkt und abgeschlossen ist.

**BEWEIS.** Nachdem  $\mathbb{R}^n$  das zweite Abzählbarkeitsaxiom erfüllt, genügt es zu zeigen, dass  $K$  genau dann folgenkompakt ist, wenn  $K$  beschränkt und abgeschlossen ist.

Wenn  $K$  nicht beschränkt ist, so gibt es für jedes  $n \in \mathbb{N}$  ein  $x_n \in K$  mit  $d(x_n, 0) \geq n$ , und die so konstruierte Folge  $(x_n)_{n \in \mathbb{N}}$  kann keine konvergente Teilfolge beinhalten: Wenn  $(x_{n_k})_{k \in \mathbb{N}}$  eine beliebige Teilfolge ist, so gibt es für jedes  $K \in \mathbb{N}$  ein  $L \in \mathbb{N}$  mit  $|x_\ell| \geq |x_{n_K}| + 1$ , für alle  $\ell \geq L$ , insbesondere ist  $|x_{n_\ell} - x_{n_K}| \geq 1$  für  $n_\ell \geq L$ . Die Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$  ist damit keine Cauchyfolge, also auch nicht konvergent.

Ist  $K$  nicht abgeschlossen, so gibt es einen Punkt  $p \in \overline{K} \setminus K$ . Damit gibt es nach Lemma 41 eine Folge  $(k_j)_{j \in \mathbb{N}}$  in  $K$ , welche gegen  $p$  konvergiert. Damit konvergiert aber auch jede Teilfolge von  $(k_j)_{j \in \mathbb{N}}$  gegen  $p \notin K$ .

Sei nun  $K$  beschränkt und abgeschlossen, und  $(x^j)_{j \in \mathbb{N}}$  eine Folge in  $K$ , wobei  $x^j = (x_1^j, \dots, x_n^j)$ . Dann sind die Folgen reeller Zahlen  $(x_k^j)_{j \in \mathbb{N}}$  beschränkt; nach Theorem 6 gibt es eine konvergente Teilfolge  $(x_1^{j_\ell})_{\ell \in \mathbb{N}}$  von  $(x_1^j)_{j \in \mathbb{N}}$ . Aus dem gleichen Grund finden wir eine weitere Teilfolge  $(x_2^{j_\ell})_{\ell \in \mathbb{N}}$  von  $(x_1^{j_\ell})_{\ell \in \mathbb{N}}$  mit der Eigenschaft, dass sowohl  $(x_1^{j_\ell})_{\ell \in \mathbb{N}}$  als auch  $(x_2^{j_\ell})_{\ell \in \mathbb{N}}$  konvergieren; wir fahren entsprechend fort und erhalten schlussendlich eine Teilfolge  $(x^{j_\ell})_{\ell \in \mathbb{N}}$  von  $(x^j)_{j \in \mathbb{N}}$  mit der Eigenschaft, dass es für jedes  $k = 1, \dots, n$  den Grenzwert  $\lim_{\ell \rightarrow \infty} x_k^{j_\ell} = x_k^0$  gibt. Nach Theorem 32 ist also  $\lim_{\ell \rightarrow \infty} x^{j_\ell} = x^0$ , und nachdem  $K$  abgeschlossen ist, ist  $x^0 \in K$ . Also enthält  $(x^j)_{j \in \mathbb{N}}$  eine Teilfolge, welche gegen einen Punkt in  $K$  konvergiert.  $\square$

**LEMMA 57.** Sei  $\|\cdot\|$  eine Norm auf  $\mathbb{R}^n$ . Dann ist  $\|\cdot\|$  mit  $\|\cdot\|_2$  vergleichbar, d.h. es gibt Konstanten  $c, C > 0$  sodass

$$c \|v\|_2 \leq \|v\| \leq C \|v\|_2.$$

**BEWEIS.** Wir behaupten zunächst, dass die norm  $\|\cdot\|$  stetig (im Sinne der von  $\|\cdot\|_2$  definierten Topologie) ist. Mit Hilfe der *Einheitsvektoren*

$$e^j = (0, \dots, 0, \underbrace{1}_{j\text{ter Eintrag}}, 0, \dots, 0), \quad j = 1, \dots, n,$$

können wir  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  als

$$x = \sum_{j=1}^n x_j e^j$$

schreiben. Seien nun  $x$  und  $x^0$  Punkte in  $\mathbb{R}^n$ . Dann ist nach Lemma 53

$$\begin{aligned}
 d^{\mathbb{R}}(\|x\|, \|x^0\|) &= \left| \|x\| - \|x^0\| \right| \\
 &\leq \|x - x^0\| \\
 &= \left\| \sum_{k=1}^n (x_k - x_k^0) e^k \right\| \\
 &\leq \sum_{k=1}^n |x_k - x_k^0| \|e^k\| \\
 &\leq \left( \sum_{k=1}^n |x_k - x_k^0|^2 \right)^{\frac{1}{2}} \left( \sum_{k=1}^n \|e^k\|^2 \right)^{\frac{1}{2}} \\
 &= \left( \sum_{k=1}^n \|e^k\|^2 \right)^{\frac{1}{2}} \|x - x^0\|_2 \\
 &= \left( \sum_{k=1}^n \|e^k\|^2 \right)^{\frac{1}{2}} d_2(x, x^0).
 \end{aligned}$$

Sei nun  $\varepsilon > 0$  gegeben, dann wählen wir

$$\delta = \frac{\varepsilon}{\left( \sum_{k=1}^n \|e^k\|^2 \right)^{\frac{1}{2}}},$$

und nach der Ungleichung oben ist, wenn  $d_2(x, x^0) < \delta$ ,

$$d^{\mathbb{R}}(\|x\|, \|x^0\|) \leq \left( \sum_{k=1}^n \|e^k\|^2 \right)^{\frac{1}{2}} d_2(x, x^0) < \left( \sum_{k=1}^n \|e^k\|^2 \right)^{\frac{1}{2}} \delta = \varepsilon.$$

Die Einheitskugel  $S^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$  ist beschränkt und abgeschlossen (als Urbild der abgeschlossenen Menge  $\{1\} \subset \mathbb{R}$  unter der stetigen Funktion  $\|\cdot\|_2$ ); also kompakt. Die stetige Funktion  $\|\cdot\|$  ist positiv auf  $S^{n-1}$  und nimmt ihr (positives) Minimum und Maximum an, es gibt also  $c, C > 0$  mit  $c \leq \|x\| \leq C$  für alle  $x \in S^{n-1}$ . Ist nun  $x \in \mathbb{R}^n$  beliebig, so gilt

$$\frac{1}{\|x\|_2} x \in S^{n-1},$$

also ist

$$c \leq \left\| \frac{1}{\|x\|_2} x \right\| \leq C,$$

und damit

$$c \|x\|_2 \leq \|x\| \leq C \|x\|_2.$$

□

**SATZ 34.** Sei  $\|\cdot\|$  eine Norm auf  $\mathbb{R}^n$ . Dann ist die von der Metrik  $d(x, y) = \|x - y\|$  induzierte Topologie auf  $\mathbb{R}^n$  die Produkttopologie. Wenn  $\tilde{d}$  eine andere durch eine Norm induzierte Metrik ist, so ist  $d$  mit  $\tilde{d}$  vergleichbar, d.h. es gibt Konstanten  $c, C > 0$  sodass

$$cd(x, y) \leq \tilde{d}(x, y) \leq Cd(x, y).$$

**BEWEIS.** Nach Lemma 57 sind die Normen  $\|\cdot\|$  und  $\|\cdot\|_2$  miteinander vergleichbar und induzieren damit dieselbe Topologie, welche nach Theorem 32 mit der Produkttopologie übereinstimmt.

Nachdem jede beliebige Norm mit der 2-Norm vergleichbar ist, sind je zwei beliebige Normen auch miteinander vergleichbar, damit sind die dadurch gegebenen Metriken vergleichbar, und induzieren nach Lemma 45 dieselbe Topologie. □

**3.3. Allgemeine Vektorräume.** Sei  $V$  ein Vektorraum über  $\mathbb{R}$ . Eine Abbildung  $\|\cdot\| : V \rightarrow \mathbb{R}$  heisst eine *Norm* auf  $V$ , wenn sie folgende Eigenschaften erfüllt: *Hier fehlt noch was!*

(N1)  $\|v\| \geq 0$  für alle  $v \in V$ ,  $\|v\| = 0$  genau dann, wenn  $v = 0$ .

(N2)  $\|\lambda v\| = |\lambda| \|v\|$  für alle  $\lambda \in \mathbb{R}$  und  $v \in V$ .

(N3)  $\|v + w\| \leq \|v\| + \|w\|$  für alle  $v, w \in V$ .

Wir sagen dann,  $(V, \|\cdot\|)$  ist ein *normierter Raum*. Eine Norm induziert durch  $d(u, v) = \|v - u\|$  eine Metrik, und damit eine Topologie auf  $V$ . Wenn wir von metrischen Eigenschaften eines normierten Raums sprechen, bzw. von topologischen Eigenschaften, so beziehen wir uns immer auf die von der Norm induzierte Metrik (bzw. Topologie). Ein normierter Raum  $(V, \|\cdot\|)$  heisst *Banachraum*, wenn er als metrischer Raum vollständig ist.

LEMMA 58. Die Abbildung  $\|\cdot\| : V \rightarrow \mathbb{R}$  ist stetig.

BEWEIS. Es gilt die umgekehrte Dreiecksungleichung:

$$d^{\mathbb{R}}(\|v\|, \|w\|) = \left| \|v\| - \|w\| \right| \leq \|v - w\| = d(v, w).$$

Ist also  $\varepsilon > 0$ , so gilt für  $d(v, w) < \delta := \varepsilon$  die Ungleichung  $d^{\mathbb{R}}(\|v\|, \|w\|) < \varepsilon$ . □

SATZ 35. Seien  $(V_1, \|\cdot\|_1)$  und  $(V_2, \|\cdot\|_2)$  normierte Räume. Eine lineare Abbildung  $A: V_1 \rightarrow V_2$  ist stetig genau dann, wenn es eine Konstante  $C > 0$  gibt, sodass  $\|Au\|_2 \leq C \|u\|_1$  für alle  $u \in V_1$  gilt. In diesem Fall wird die Zahl

$$\|A\| = \sup \left\{ \frac{\|Au\|_2}{\|u\|_1} : u \neq 0 \right\}$$

als die Operatornorm von  $A$  bezeichnet.

BEWEIS. Sei zunächst  $A$  stetig. Dann gibt es einen Ball  $B_\delta(0)$  sodass  $B_\delta^1(0) \subset A^{-1}(B_1^2(0))$  ist, d.h.

$$\|Au\|_2 \leq 1, \quad \text{wenn } \|u\|_1 \leq \delta.$$

Ist  $u \in V_1$  beliebig, so ist

$$\left\| \frac{\delta u}{2 \|u\|_1} \right\|_1 = \frac{\delta}{2} < \delta,$$

und damit

$$\left\| A \left( \frac{\delta u}{2 \|u\|_1} \right) \right\|_2 \leq 1,$$

also

$$\|Au\|_2 \leq \frac{2}{\delta} \|u\|_1.$$

Ist umgekehrt  $\|Au\|_2 \leq C \|u\|_1$ , und  $u_0 \in V$  sowie  $\varepsilon > 0$  gegeben, so ist für  $u \in B_\delta^1(u_0)$ , wo  $\delta = \varepsilon/C$  die Ungleichung

$$d^2(Au_0, Au) = \|Au_0 - Au\|_2 = \|A(u - u_0)\|_2 \leq C \|u - u_0\|_1 = Cd^1(u_0, u) < \varepsilon$$

erfüllt, also  $A$  stetig im Punkt  $u_0$ . Nachdem  $u_0$  beliebig war, ist  $A$  stetig. □

Sei nun  $(V, \|\cdot\|)$  ein normierter Raum. Dann können wir durch

$$(v, w) \mapsto \max\{\|v\|, \|w\|\}$$

eine Norm auf  $V^2$  definieren, welche die Produkttopologie auf  $V^2$  induziert (die Rechnungen sind ganz analog zum Fall des  $\mathbb{R}^n$ ); aus Lemma 54 folgt, dass diese Norm auf  $V^2$  vergleichbar mit der Norm, welche durch

$$(v, w) \mapsto \|v\| + \|w\|$$

definiert wird, ist. Es gilt, dass die (lineare) Abbildung

$$(v, w) \mapsto v + w$$

stetig (auf  $V^2$  versehen mit der Produkttopologie) ist, da ja  $\|v + w\| \leq \|v\| + \|w\| \leq 2 \max\{\|v\|, \|w\|\}$  gilt. Auch die skalare Multiplikation  $(\lambda, v) \mapsto \lambda v$  ist eine stetige Abbildung auf  $\mathbb{R} \times V$  (versehen mit der Produkttopologie) mit Werten in  $V$  (der Beweis kann entweder wie oben oder durch eine Adaption des Beweises von



Lemma 29 erfolgen). Mit anderen Worten: *Die Operationen der Addition und der Skalarmultiplikation sind stetig.*

SATZ 36. *Sei  $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$  eine lineare Abbildung. Dann ist  $A$  stetig.*

BEWEIS. Wir können den  $\mathbb{R}^n$  und den  $\mathbb{R}^m$  nach Lemma 57 mit einer beliebigen Norm versehen. Wir wählen die  $\infty$ -Norm, und versuchen, ein  $C > 0$  mit  $\|Au\|_\infty \leq \|u\|_\infty$  für alle  $u \in \mathbb{R}^n$  zu finden. Die lineare Abbildung  $A$  ist durch eine  $m \times n$  Matrix gegeben, d.h.

$$Au = \begin{pmatrix} (Au)_1 \\ \dots \\ (Au)_m \end{pmatrix} = \begin{pmatrix} A_1^1 & \dots & A_1^n \\ \vdots & & \vdots \\ A_m^1 & \dots & A_m^n \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix},$$

d.h.  $(Au)_j = \sum_{k=1}^n A_j^k u_k$ . Also ist

$$|(Au)_j| \leq \sum_{k=1}^n |A_j^k| |u_k| \leq \left( \sum_{k=1}^n |A_j^k| \right) \|u\|_\infty,$$

und damit

$$\|Au\|_\infty \leq \underbrace{\left( \max_{j=1}^m \sum_{k=1}^n |A_j^k| \right)}_{=:C} \|u\|_\infty.$$

□

Wir können nun ein weiteres wichtiges Beispiel eines Banachraums geben: für einen topologischen Raum  $X$  ist  $C(X)$  der Raum der stetigen reellwertigen Funktionen auf  $X$ . Wenn  $X$  kompakt ist, so ist  $C(X)$ , versehen mit der Supremumsnorm

$$\|\varphi\|_\infty = \sup_{x \in X} |\varphi(x)|,$$

welche nach Lemma 37 jene Norm ist, die von  $B(X)$  kommt, ein Banachraum. Bevor wir zum Beweis schreiten, schicken wir ein – zunächst unabhängiges – Lemma voran.

LEMMA 59. *Sei  $F \subset C(X)$  eine Menge von Funktionen mit folgender Eigenschaft: Für jedes  $x \in X$  und jedes  $\varepsilon > 0$  gibt es eine Umgebung  $U \in \mathcal{N}(x)$  sodass  $f(U) \subset B_\varepsilon(f(x))$  für jedes  $f \in F$  gilt.*

*Dann gibt es für jedes  $\varepsilon > 0$  eine endliche Überdeckung*

LEMMA 60. *Sei  $X$  ein kompakter topologischer Raum, und  $(\varphi_j)_{j \in \mathbb{N}} \subset C(X)$  eine Cauchyfolge im Sinne Supremumsnorm auf  $C(X)$ . Dann gibt es eine stetige Funktion  $\varphi \in C(X)$  sodass  $\varphi = \lim_{j \rightarrow \infty} \varphi_j$  im Sinne der Supremumsnorm ist.*

BEWEIS. Da  $\varphi_j$  eine Cauchyfolge bezüglich der Supremumsnorm ist, ist für beliebiges  $x \in X$  die Folge  $(\varphi_j(x))_{j \in \mathbb{N}} \subset \mathbb{R}$  konvergent, und wir definieren  $\varphi: X \rightarrow \mathbb{R}$  durch  $\varphi(x) = \lim_{j \rightarrow \infty} \varphi_j(x)$ . Wir behaupten, dass  $\varphi = \lim_j \varphi_j(x) \in C(X)$ . Sei dazu  $x_0 \in X$  und  $\varepsilon > 0$  beliebig. Wir wählen ein  $N \in \mathbb{N}$  mit  $|\varphi_j(x) - \varphi_k(x)| < \varepsilon$  für alle  $j, k \geq N$ . Da  $\varphi_N \in C(X)$ , gibt es eine Umgebung □

**Definition 33.** Eine Abbildung  $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$  heisst ein *inneres Produkt* auf  $V$ , wenn die folgenden Bedingungen erfüllt sind:

(P1)  $\langle v, v \rangle \geq 0$ , und  $\langle v, v \rangle = 0$  genau dann, wenn  $v = 0$ ;

(P2)  $\langle v, w \rangle = \langle w, v \rangle$  für alle  $v, w \in V$ ;

(P3)  $\langle u, v + \lambda w \rangle = \langle u, v \rangle + \lambda \langle u, w \rangle$  für alle  $u, v, w \in V$ ,  $\lambda \in \mathbb{R}$ .

Wir nennen das Paar  $(V, \langle \cdot, \cdot \rangle)$  einen Innenproduktraum, und schreiben  $\|v\| = \sqrt{\langle v, v \rangle}$ .

Es ist leicht zu zeigen, dass die so definierte Abbildung  $\|\cdot\|$  die Eigenschaften (N1) und (N2) einer Norm erfüllt. Für die Dreiecksungleichung benötigen wir die Cauchy-Schwarz-Ungleichung.

LEMMA 61. *Sei  $(V, \langle \cdot, \cdot \rangle)$  ein Innenproduktraum. Dann gilt  $\langle v, w \rangle \leq \|v\| \|w\|$  für alle  $v, w \in V$ .*

BEWEIS. Für jedes  $\lambda \in \mathbb{R}$  gilt

$$0 \leq \langle v - \lambda w, v - \lambda w \rangle = \|v\|^2 - 2\lambda \langle v, w \rangle + \lambda^2 \langle w, w \rangle,$$

also insbesondere, falls  $w \neq 0$ , für  $\lambda = \frac{\langle v, w \rangle}{\|w\|^2}$ :

$$0 \leq \|v\|^2 - 2 \frac{\langle v, w \rangle}{\|w\|^2} \langle v, w \rangle + \left( \frac{\langle v, w \rangle}{\|w\|^2} \right)^2 \langle w, w \rangle = \|v\|^2 - \frac{\langle v, w \rangle^2}{\|w\|^2},$$

oder gleichbedeutend

$$\langle v, w \rangle \leq |\langle v, w \rangle| \leq \|v\| \|w\|;$$

nachdem die Ungleichung trivialerweise auch für  $w = 0$  gilt, für alle  $v, w \in V$ . □

LEMMA 62. Die oben definierte Abbildung  $\|\cdot\|$  ist eine Norm auf  $V$ .

BEWEIS. Wir müssen nur noch die Dreiecksungleichung zeigen. Seien  $v, w \in V$  beliebig, dann ist

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle \\ &= \|v\|^2 + 2\langle v, w \rangle + \langle w, w \rangle \\ &\leq \|v\|^2 + 2\|v\| \|w\| + \langle w, w \rangle \\ &= (\|v\| + \|w\|)^2. \end{aligned}$$

□

# Differentialrechnung

## 1. Der Begriff der Differenzierbarkeit

**Definition 34.** Sei  $f: (a, b) \rightarrow \mathbb{R}$  eine reelle Funktion,  $x \in (a, b)$ . Wir sagen,  $f$  ist *differenzierbar* im Punkt  $x$ , wenn der Grenzwert der Differenzenquotienten

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} =: f'(x)$$

existiert. Der Grenzwert  $f'(x)$  wird als die *Ableitung* von  $f$  im Punkt  $x$  bezeichnet.

Ist  $f$  in jedem Punkt von  $(a, b)$  differenzierbar, so sagen wir,  $f$  ist auf  $(a, b)$  differenzierbar. Den Raum der differenzierbaren Funktionen auf  $(a, b)$  bezeichnen wir mit  $D((a, b))$ ; für  $f \in D((a, b))$  ist die Ableitung  $f': (a, b) \rightarrow \mathbb{R}$  eine auf  $(a, b)$  definierte Funktion.

**Beispiel 26.** Die Potenzfunktionen  $x \mapsto x^n$ ,  $n \in \mathbb{N}$  sind in jedem Punkt  $x \in \mathbb{R}$  differenzierbar, d.h.  $f(x) = x^n \in D(\mathbb{R})$ :

$$\lim_{h \rightarrow 0} \frac{(x+h)^n - x^n}{h} = \lim_{h \rightarrow 0} \frac{1}{h} \sum_{j=1}^n \binom{n}{j} x^{n-j} h^j = nx^{n-1},$$

also  $f'(x) = nx^{n-1}$ .

**Übungsaufgabe 31.** Zeige, dass die Funktionen  $x \mapsto x^{-n}$  für  $n \in \mathbb{N}$  differenzierbar auf  $\mathbb{R} \setminus \{0\}$  sind, und berechne ihre Ableitungen.

Eine einfache Umformulierung von Definition 34 ist die folgende:  $f: (a, b) \rightarrow \mathbb{R}$  ist differenzierbar in  $x \in (a, b)$  genau dann, wenn es eine Zahl  $\lambda \in \mathbb{R}$  sodass die durch

$$(3) \quad R(h) = f(x+h) - f(x) - \lambda h,$$

für  $h \in (-\delta, \delta)$  für ein  $\delta > 0$  definierte Funktion  $R$  die Eigenschaft hat, dass

$$\lim_{h \rightarrow 0} \frac{R(h)}{h} = 0 \text{ bzw. } R(h) = h\tilde{R}(h) \text{ mit } \lim_{h \rightarrow 0} \tilde{R}(h) = 0.$$

Die Zahl  $\lambda$  ist eindeutig bestimmt, und  $\lambda = f'(x)$ . Insbesondere ist  $\lim_{h \rightarrow 0} f(x+h) - f(x) = 0$ , also  $f$  stetig im Punkt  $x$ . Diese wichtige Folgerung aus der Differenzierbarkeit formulieren wir auch:

**LEMMA 63.** Sei  $f: (a, b) \rightarrow \mathbb{R}$  differenzierbar im Punkt  $x \in (a, b)$ . Dann ist  $f$  stetig im Punkt  $x$ . Insbesondere ist  $D((a, b)) \subset C((a, b))$ .

**Bemerkung 9.** Wir bemerken noch einmal explizit, dass Differenzierbarkeit wesentlich stärker als Stetigkeit ist. Eine Funktion  $f$  ist stetig im Punkt  $x \in (a, b)$ , wenn  $\lim_{h \rightarrow 0} f(x+h) - f(x) = 0$  ist. Wenn  $f$  differenzierbar im Punkt  $x$  ist, so gilt sogar dass es ein  $C > 0$  gibt sodass  $|f(x+h) - f(x)| < C|h|$  für kleine  $h$  ist: Da  $\lim_{h \rightarrow 0} \frac{R(h)}{h} = 0$  in (3) ist, sehen wir, dass es ein  $\delta > 0$  gibt mit

$$\left| \frac{R(h)}{h} \right| < 1.$$

Dann ist  $|f(x+h) - f(x)| \leq (|\lambda| + 1)|h|$  für  $|h| < \delta$ .

Ein explizites Beispiel einer Funktionen, welche stetig im Punkt 0 aber nicht differenzierbar im Punkt 0 ist, ist  $x \mapsto \sqrt{x}$ .

**Übungsaufgabe 32.** Zeige die letzte Behauptung:  $x \mapsto \sqrt{x}$  ist nicht differenzierbar im Punkt 0.

Eine graphische Interpretation der Ableitung kann man wie folgt erhalten: Der Differenzenquotient

$$\frac{f(x+h) - f(x)}{h}$$

ist der Anstieg der Geraden, welche die Punkte  $(x, f(x))$  und  $(x+h, f(x+h))$  im Graphen von  $f$  verbindet. Die Gerade durch  $(x, f(x))$  mit Anstieg  $f'(x)$  ist also die *Tangente* an den Graphen von  $f$  im Punkt  $(x, f(x))$ .

**Übungsaufgabe 33.** Zeige, dass die Betragsfunktion  $x \mapsto |x|$  auf  $\mathbb{R} \setminus \{0\}$  differenzierbar ist und berechne ihre Ableitung. Ist sie im Punkt 0 differenzierbar?

Oft ist es interessant, Ableitungen auch in *Randpunkten* von Intervallen zu betrachten; die dort gebildeten Grenzwerte sind natürlich einseitig zu bilden. Wir können auch ein wenig allgemeiner definieren:

**Definition 35.** Sei  $f: X \rightarrow \mathbb{R}$  eine reelle Funktion,  $x \in X$  ein Häufungspunkt von  $X$ . Wir sagen,  $f$  ist *differenzierbar* im Punkt  $x$ , wenn der Grenzwert der Differenzenquotienten

$$\lim_{\substack{h \rightarrow 0 \\ x+h \in X}} \frac{f(x+h) - f(x)}{h} =: f'(x)$$

existiert. Der Grenzwert  $f'(x)$  wird als die *Ableitung* von  $f$  im Punkt  $x$  bezeichnet.

Ist  $f$  in jedem Punkt von  $X$  differenzierbar, so sagen wir,  $f$  ist auf  $X$  differenzierbar. Den Raum der differenzierbaren Funktionen auf  $X$  bezeichnen wir mit  $D(X)$ ; für  $f \in D(X)$  ist die Ableitung  $f': X \rightarrow \mathbb{R}$  eine auf  $X$  definierte Funktion.

Diese Definition ist *sehr* allgemein; wir werden sie später nur in Spezialfällen (für kompakte Intervalle  $X$ ) genauer ausleuchten. Zuvor wollen wir einige Rechenregeln erkunden.

LEMMA 64. Seien  $f, g$  im Punkt  $x \in X$  differenzierbar, und  $\lambda \in \mathbb{R}$ . Dann sind auch die Funktionen  $\lambda f + g$ ,  $fg$ , sowie  $f/g$  (letztere allerdings nur, falls  $g(x) \neq 0$ ) im Punkt  $x$  differenzierbar, und für die Ableitungen gilt:

$$\begin{aligned} (\lambda f + g)'(x) &= \lambda f'(x) + g'(x) && \text{Summenregel} \\ (fg)'(x) &= f'(x)g(x) + f(x)g'(x) && \text{Produktregel} \\ \left(\frac{f}{g}\right)'(x) &= \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} && \text{Quotientenregel} \end{aligned}$$

BEWEIS. Die erste Behauptung des Lemmas wird als Übungsaufgabe dem Leser überlassen. Für die Produktregel rechnen wir wie folgt:

$$\begin{aligned} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} &= \frac{f(x+h)g(x+h) - f(x+h)g(x) + f(x+h)g(x) - f(x)g(x)}{h} \\ &= f(x+h) \frac{g(x+h) - g(x)}{h} + g(x) \frac{f(x+h) - f(x)}{h} \end{aligned}$$

und lassen nun  $h \rightarrow 0$ ; nach Lemma 63 ist  $f$  stetig im Punkt  $x$ , und damit ist, nachdem die Limiten der einzelnen Terme auf der rechten Seite der Gleichung alle existieren, auch  $fg$  im Punkt  $x$  differenzierbar, und  $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$  wie behauptet.

Für die Quotientenregel ist es nun genug, zu zeigen dass für  $g(x) \neq 0$  der Bruch  $1/g$  im Punkt  $x$  differenzierbar ist, und  $(1/g)'(x) = -g'(x)/g(x)^2$  gilt. Wir berechnen nun

$$\frac{1}{g(x+h)} - \frac{1}{g(x)} = \frac{1}{g(x+h)g(x)} \frac{g(x) - g(x+h)}{h}$$

und sehen, dass nach Voraussetzung der Grenzwert (wir verwenden wieder Lemma 63) für  $h \rightarrow 0$  der Terme auf der rechten Seite der Gleichung existieren, und somit  $(1/g)$  differenzierbar im Punkt  $x$  mit der behaupteten Ableitung ist.  $\square$

**Bemerkung 10.** Die Menge  $D(X)$  ist also ein *Ring*, genauer gesagt, eine *Algebra* über  $\mathbb{R}$ . Insbesondere folgt aus der Differenzierbarkeit der Funktion  $x$  auf  $\mathbb{R}$ , dass  $D(\mathbb{R})$  alle Polynomfunktionen enthält.

**Übungsaufgabe 34.** Die Ableitungen der Polynomfunktionen haben wir im Prinzip schon in Beispiel 26 berechnet. Zeige nochmals, diesmal unter Verwendung der Produktregel und mit Hilfe von Induktion, dass die Ableitung von  $f(x) = x^n$  durch  $f'(x) = nx^{n-1}$  für  $n \in \mathbb{N}$  gegeben ist.

Die letzte wichtige Rechenregel ist die Kettenregel.

LEMMA 65. Seien  $f, g$  gegeben, sodass die Zusammensetzung  $g \circ f$  in einer Umgebung von  $x$  Sinn macht. Ist nun  $f$  an der Stelle  $x$  differenzierbar, und  $g$  an der Stelle  $f(x)$  differenzierbar, so ist  $g \circ f$  an der Stelle  $x$  differenzierbar und es gilt

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

BEWEIS. Die Kettenregel folgt, indem man

$$\frac{g(f(x+h)) - g(f(x))}{h} = \frac{g(f(x+h)) - g(f(x))}{f(x+h) - f(x)} \frac{f(x+h) - f(x)}{h}$$

schreibt und den Grenzwert  $h \rightarrow 0$  bildet. Dabei muss natürlich das Problem, dass der Nenner des ersten Bruchs durchaus 0 sein kann, berücksichtigt werden. Es ist deswegen von Vorteil, entweder die Umformulierung in (3) zu verwenden (um die notwendig Fallunterscheidung zu vermeiden).

Wir schreiben deswegen

$$g(y+k) - g(y) - g'(y)k = k\tilde{S}(k), \quad f(x+h) - f(x) - f'(x)h = h\tilde{R}(h),$$

und rechnen:

$$\begin{aligned} g(f(x+h)) - g(f(x)) &= g'(f(x))(f(x+h) - f(x)) + (f(x+h) - f(x))\tilde{S}(f(x+h) - f(x)) \\ &= \left(g'(f(x)) + \tilde{S}(f(x+h) - f(x))\right) (f(x+h) - f(x)) \\ &= \left(g'(f(x)) + \tilde{S}(f(x+h) - f(x))\right) (f'(x)h + h\tilde{R}(h)) \\ &= \left(g'(f(x)) + \tilde{S}(f(x+h) - f(x))\right) (f'(x) + \tilde{R}(h))h. \end{aligned}$$

Damit ist

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{g(f(x+h)) - g(f(x))}{h} &= \lim_{h \rightarrow 0} \left(g'(f(x)) + \tilde{S}(f(x+h) - f(x))\right) (f'(x) + \tilde{R}(h)) \\ &= g'(f(x))f'(x). \end{aligned}$$

□

## 2. Der Mittelwertsatz und Folgerungen

Seien  $x, y \in [a, b]$ , und  $f$  eine Funktion auf  $[a, b]$ . Der Wert

$$\frac{f(x) - f(y)}{x - y}$$

ist kann als *mittlere Änderungsrate* von  $f$  zwischen  $x$  und  $y$  aufgefasst werden. Wenn nun  $y = x + h$  ist und wir den Grenzwert für  $y \rightarrow x$  (oder  $h \rightarrow 0$ ) dieses Differenzenquotienten betrachten, so macht es Sinn, die Ableitung  $f'(x)$  als *momentane Änderung* von  $f$  im Punkt  $x$  aufzufassen. Der Mittelwertsatz erlaubt es uns, diesen Zusammenhang zu präzisieren.

SATZ 37 (Mittelwertsatz). Sei  $f \in C([a, b]) \cap D((a, b))$ . Dann gibt es ein  $\xi \in (a, b)$  mit

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

BEWEIS. Sei

$$g(x) = (b-a)f(x) + (a-x)f(b) + (x-b)f(a).$$

Dann ist  $g(a) = g(b) = 0$ , und nach Lemma 64, wieder stetig auf  $[a, b]$  und differenzierbar auf  $(a, b)$ . Wir behaupten, dass es einen Punkt  $\xi \in (a, b)$  gibt, für welchen  $g'(\xi) = 0$  ist; für diesen Punkt  $\xi$  ist also

$$0 = (b-a)f'(\xi) + f(a) - f(b)$$

wie verlangt.

Wenn  $g(x)$  konstant auf  $[a, b]$  ist, so gibt es nichts zu zeigen. Ist  $g(x)$  nicht konstant auf  $[a, b]$ , so nehmen wir ohne Beschränkung der Allgemeinheit an, dass  $\max_{x \in [a, b]} g(x) > 0$ . Sei  $\xi$  ein Punkt mit  $g(\xi) = \max_{x \in [a, b]} g(x)$ . Dann ist auf der einen Seite, da  $g$  im Punkt  $\xi$  differenzierbar ist,

$$g'(\xi) = \lim_{h \searrow 0} \frac{g(\xi + h) - g(\xi)}{h} \leq 0,$$

und andererseits

$$g'(\xi) = \lim_{h \nearrow 0} \frac{g(\xi + h) - g(\xi)}{h} \geq 0,$$

also  $g'(\xi) = 0$  wie behauptet.  $\square$

Die im Beweis dieses Satzes gemachte Beobachtung (wenn  $g$  ein Maximum an der Stelle  $\xi$  hat, so ist  $g'(\xi) = 0$ ) wollen wir festhalten:

LEMMA 66. *Sei  $g$  eine stetige Funktion auf  $X$ . Wenn  $g$  ein (lokales) Maximum an der Stelle  $\xi$  hat, und an der Stelle  $\xi$  differenzierbar ist, so ist  $g'(\xi) = 0$ .*

Die Umkehrung dieses Satzes stimmt *nicht*, wie das Beispiel der Funktion  $x \mapsto x^3$  zeigt.

KOROLLAR 4. *Sei  $f \in C([a, b]) \cap D((a, b))$ , und  $f'(x) \geq 0$  (bzw.  $f'(x) > 0$ ) für  $x \in (a, b)$ . Dann ist  $f$  (strikt) monoton wachsend auf  $[a, b]$ .*

BEWEIS. Für beliebige  $x, y \in [a, b]$ ,  $x < y$ , gilt nach Theorem 37, dass

$$f(y) - f(x) = f'(\xi)(y - x) \geq 0 \text{ bzw. } > 0$$

nach Voraussetzung, also ist  $f$  (strikt) monoton.  $\square$

Nicht immer führt die Verwendung des Mittelwertsatzes zu optimalen Aussagen:

KOROLLAR 5. *Sei  $f \in C([a, b]) \cap D((a, b))$ ,  $f' \in C((a, b))$ , und  $f'(x) > 0$  für  $x \in (a, b)$ . Wir schreiben  $c = f(a)$ ,  $d = f(b)$ . Dann gibt es eine Funktion  $g \in C([a, b]) \cap D((c, d))$  welche  $f(g(y)) = y$  für  $y \in [c, d]$  und  $g(f(x)) = x$  für  $x \in [a, b]$  erfüllt. Die Ableitung von  $g$  ist gegeben durch*

$$g'(y) = \frac{1}{f'(g(y))}.$$

BEWEIS. Da  $f$  strikt monoton wachsend und stetig ist, gilt dass die inverse Funktion  $g$  definiert, stetig, und wiederum strikt monoton wachsend ist. Sei  $y \in (c, d)$ , dann ist  $h = f(g(y+h)) - f(g(y)) = f'(\xi_h)(g(y+h) - g(y))$  und damit

$$\frac{g(y+h) - g(y)}{h} = \frac{1}{f'(\xi_h)};$$

wenn wir  $h \rightarrow 0$  gehen lassen, folgt  $\xi_h \rightarrow g(y)$ , und damit (da  $f'$  stetig ist), dass  $g$  differenzierbar an der Stelle  $y$  ist, mit der behaupteten Formel für die Ableitung.  $\square$

**Übungsaufgabe 35.** Tatsächlich genügt es, anzunehmen, dass  $f$  an der Stelle  $x$  differenzierbar ist, dann gilt, dass  $f^{-1}$  an der Stelle  $f(x)$  differenzierbar ist, und  $(f^{-1})'(f(x)) = 1/(f'(x))$ . Zeige diese Behauptung direkt aus der Definition der Differenzierbarkeit und interpretiere die Aussage graphisch; genauer: Sei  $f: [a, b] \rightarrow [c, d]$  strikt monoton und stetig. Wenn  $f$  an der Stelle  $x$  differenzierbar ist, dann ist  $f^{-1}$  an der Stelle  $f(x)$  differenzierbar.

LEMMA 67. *Sei  $f \in C([a, b]) \cap D((a, b))$ . Angenommen  $f'(x)$  lässt sich stetig nach  $[a, b]$  (bzw.  $(a, b)$ ) fortsetzen. Dann ist  $f$  auch in  $a$  (bzw.  $b$ ) differenzierbar, und  $f'$  ist stetig auf  $[a, b]$  (bzw.  $(a, b)$ ).*

BEWEIS. Wir betrachten den Fall des linken Endpunkts; analoge Argumente geben das entsprechende Resultat im rechten Endpunkt des Intervalls. Mit Hilfe des Mittelwertsatzes erhalten wir für jedes  $h$  ein  $\xi_h \in (a, a+h)$  mit  $f(a+h) - f(a) = f'(\xi_h)h$ . Also ist

$$\lim_{h \searrow 0} \frac{f(a+h) - f(a)}{h} = \lim_{h \searrow 0} f'(\xi_h);$$

der Grenzwert auf der rechten Seite existiert aber nach Voraussetzung, da  $\xi_h \rightarrow a$  für  $h \searrow 0$ .  $\square$

Ableitungen von differenzierbaren Funktionen sind speziell (d.h. nicht jede Funktion tritt als Ableitung auf); eine interessante Eigenschaft ist die Zwischenwerteigenschaft, welche Ableitungen erfüllen.

**KOROLLAR 6.** Sei  $f \in D((a, b))$ ,  $c, d \in (a, b)$ ,  $f'(c) < y < f'(d)$ . Dann gibt es ein  $\xi \in (c, d)$  mit  $f'(\xi) = y$ .

**BEWEIS.** Sei  $g(x) = f(x) - yx$ . Da  $g$  stetig ist, nimmt es zwischen  $c$  und  $d$  sein Minimum auf  $[c, d]$  an, sagen wir, an der Stelle  $\xi$ . Nachdem  $g'(c) < 0$  und  $g'(d) > 0$  ist, muss dieser Punkt  $\xi \in (c, d)$  erfüllen. Nach Lemma 66 folgt  $g'(\xi) = 0$ , also  $f'(\xi) = y$ .  $\square$

Um Nullstellen von Funktionen zu finden, verwendet man oft Newtons Methode:

**KOROLLAR 7.** Sei  $f \in C([-a, a]) \cap D((-a, a))$  mit stetiger Ableitung  $f' \in C((-a, a))$ ,  $f'(0) \neq 0$ , und  $f(0) = 0$ . Dann gibt es ein  $\varepsilon > 0$  sodass für jedes  $x_0 \in [-\varepsilon, \varepsilon]$  die durch

$$x_{j+1} = x_j - \frac{f(x_j)}{f'(x_j)}$$

definierte Folge  $x_j$  gegen die Nullstelle 0 von  $f$  konvergiert.

**BEWEIS.** Wir wählen  $\varepsilon$  so klein, dass für beliebige  $\xi, \eta \in [-\varepsilon, \varepsilon]$

$$\left| 1 - \frac{f'(\xi)}{f'(\eta)} \right| < 1/2.$$

Wir definieren  $g(x) = x - \frac{f(x)}{f'(x)}$  und behaupten, dass  $|g(x)| \leq \frac{1}{2}|x|$  ist. Tatsächlich ist (da  $f(0) = 0$ ) nach dem Mittelwertsatz für ein  $\xi$  zwischen 0 und  $x$  die Gleichung  $f(x) = xf'(\xi)$  erfüllt und damit

$$|g(x)| = \left| x - \frac{f(x)}{f'(x)} \right| = \left| x - \frac{xf'(\xi)}{f'(x)} \right| = |x| \left| 1 - \frac{f'(\xi)}{f'(x)} \right| < \frac{1}{2}|x|.$$

Damit ist  $|x_n| \leq (1/2)^n|x_0|$ , also konvergiert  $x_n \rightarrow 0 (n \rightarrow \infty)$ .  $\square$

**Übungsaufgabe 36.** Zeige, dass man die Nullstelle 0 durch eine beliebige Zahl  $c \in [a, b]$  ersetzen kann; formuliere das entsprechende Ergebnis selber.

**Übungsaufgabe 37.** Analysiere den Beweis—kann man die Methode auch durch  $g(x) = x - Cf(x)$  für ein geeignetes  $C \in \mathbb{R}$  ersetzen? (Newton's Methode hat Vorteile, wenn  $f$  sogar 2-mal differenzierbar ist. Dazu aber später mehr.)

**Übungsaufgabe 38.** Seien  $f, g \in D((a, b)) \cap C([a, b])$ . Dann gibt es ein  $\xi \in (a, b)$  mit  $f'(\xi)(g(b) - g(a)) - g'(\xi)(f(b) - f(a)) = 0$ .

### 3. Grenzwertsätze

Wir werden nun vor allem auf kompakten Intervallen  $[a, b]$  arbeiten.

**Definition 36.** Für jedes  $k \in \mathbb{N}$  sagen wir, die Funktion  $f$  ist  $k$ -mal differenzierbar, wenn sie  $(k - 1)$ -mal differenzierbar und die  $(k - 1)$ -te Ableitung  $f^{(k-1)}$  erneut differenzierbar ist; die Ableitung von  $f^{(k-1)}$  ist die  $k$ -te Ableitung von  $f$  und wird mit  $f^{(k)}$  bezeichnet. Der Raum der  $k$ -mal differenzierbaren Funktionen wird mit  $D^k(X)$  bezeichnet. Ist die  $k$ -te Ableitung zusätzlich stetig auf  $X$ , so sagen wir,  $f$  ist  $k$ -mal stetig differenzierbar und schreiben  $f \in C^k(X)$ . Der Raum der stetigen Funktionen auf  $X$  wird auch mit  $C^0(X) = C(X)$  bezeichnet.

Eine Funktion  $f$ , die  $k$ -mal differenzierbar für jedes  $k \in \mathbb{N}$  ist, wird als *glatt* bezeichnet, und man schreibt  $f \in C^\infty(X)$ .

**Übungsaufgabe 39.** Zeige, dass  $D^k(X)$  und  $C^k(X)$  Algebren über  $\mathbb{R}$  sind, sowie die verallgemeinerte Produktregel: Wenn  $f, g \in D^k(X)$ , so ist die  $k$ -te Ableitung von  $fg$  gegeben durch

$$(fg)^{(k)} = \sum_{j=0}^k \binom{k}{j} f^{(j)} g^{(k-j)}.$$

Für endliches  $k$  wird der Raum  $C^k([a, b])$  mit Hilfe der Festsetzung

$$\|f\|_k = \sum_{j \leq k} \max_{x \in [a, b]} |f^{(j)}(x)|$$

zu einem normierten Raum. Die durch diese Norm induzierte Metrik  $d(f, g) = \|f - g\|_k$  erzeugt die Topologie der *gleichmässigen Konvergenz in den ersten  $k$  Ableitungen* (bzw. für  $k = 0$  die Topologie der *gleichmässigen Konvergenz*). Wir zeigen, dass  $C^k([a, b])$  *vollständig* ist, d.h.  $C^k([a, b])$  ist ein Banachraum:

LEMMA 68. *Der Raum  $C^k([a, b])$  ist für jedes  $k \geq 0$ ,  $k \in \mathbb{N}$  ein Banachraum. Die Abbildung  $D: C^k([a, b]) \rightarrow C^{k-1}([a, b])$  welche durch  $(Df)(x) = f'(x)$  definiert ist, ist stetig.*

BEWEIS. Sei zunächst  $k = 0$ . Eine Folge stetiger Funktionen  $f_n \in C([a, b])$  ist eine Cauchyfolge, wenn für jedes  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  existiert, sodass für alle  $m, n \geq N$

$$|f_m(x) - f_n(x)| < \varepsilon, \quad x \in [a, b]$$

gilt. Damit ist für jedes fixe  $x \in [a, b]$  die Folge  $f_n(x)$  eine Cauchyfolge, und besitzt damit einen Grenzwert, den wir mit  $f(x)$  bezeichnen. Wir behaupten, dass  $f$  eine stetige Funktion ist. Sei also  $\xi \in [a, b]$  fix, und  $\varepsilon > 0$ . Wir wählen  $N$  sodass

$$|f_m(x) - f_n(x)| \leq \varepsilon, \quad m, n \geq N, \quad x \in [a, b]$$

gilt.

Lassen wir nun  $n \rightarrow \infty$ , sehen wir, dass  $|f_m(x) - f(x)|$  für jedes  $m \geq N$  und  $x \in [a, b]$  gilt; insbesondere gilt, dass  $f_n \rightarrow f$  für  $n \rightarrow \infty$ , wenn wir nur zeigen können, dass  $f \in C(X)$ . Wir wählen  $\delta > 0$  sodass  $|f_N(\xi) - f_N(x)| < \varepsilon$  wenn  $|\xi - x| < \delta$ . Es folgt, dass

$$|f(\xi) - f(x)| \leq |f(\xi) - f_N(\xi)| + |f_N(\xi) - f_N(x)| + |f_N(x) - f(x)| < 3\varepsilon$$

wenn  $|\xi - x| < \delta$ . Also ist  $f$  stetig im Punkt  $\xi$ , und nachdem  $\xi$  beliebig war, auf ganz  $[a, b]$ .

Es genügt nun zu zeigen, dass wenn  $f_n$  eine Folge differenzierbarer Funktionen ist, welche gleichmässig auf  $[a, b]$  gegen eine stetige Funktion  $f$  konvergieren, und die Folge der Ableitungen  $f'_n$  gleichmässig gegen eine stetige Funktion  $g$  konvergieren, auch  $f$  differenzierbar ist und  $f' = g$  gilt. Sei also  $x \in [a, b]$  beliebig. Wir müssen den Grenzwert

$$\frac{f(x+h) - f(x)}{h}$$

für  $h \rightarrow 0$  berechnen und zeigen, dass dieser  $g(x)$  ist; mit anderen Worten, wir wollen zeigen, dass für ein geeignetes  $c > 0$  die Funktion

$$A(h) = \begin{cases} \frac{f(x+h) - f(x)}{h} & h \in [-c, c] \setminus \{0\} \\ g(x) & h = 0 \end{cases}$$

stetig ist. Nach dem ersten Teil des Beweises ist es genug, zu zeigen, dass  $A(h)$  der gleichmässige Grenzwert der stetigen Abbildungen

$$A_n(h) = \begin{cases} \frac{f_n(x+h) - f_n(x)}{h} & h \in [-c, c] \setminus \{0\} \\ f'_n(x) & h = 0 \end{cases}$$

ist. Falls  $A_n$  gleichmässig konvergiert,  $\lim_{n \rightarrow \infty} A_n = A$ ; es reicht also, wiederum weil wir schon wissen, dass  $C([-c, c])$  vollständig ist, zu zeigen, dass  $A_n$  eine Cauchyfolge ist.

Sei also  $\varepsilon > 0$  beliebig. Für jedes  $h$  und für jedes  $n$  existiert ein  $\xi_{n,h}$  mit  $|x - \xi_{n,h}| < |h|$  sodass  $A_n(h) = f'_n(\xi_{n,h})$ . Da die  $f'_n$  gleichmässig gegen  $g$  konvergieren, gibt es ein  $N \in \mathbb{N}$  mit  $|f'_m(t) - f'_n(t)| < \varepsilon$  für alle  $t \in [a, b]$  und alle  $n, m \geq N$ . Wir behaupten nun zunächst, dass es ein  $\delta > 0$  gibt, sodass für alle  $m \geq N$  und alle  $s, t \in [a, b]$  mit  $|s - t| < \delta$  die Ungleichung  $|f'_m(s) - f'_m(t)| < 3\varepsilon$  gilt. Da  $g$  gleichmässig stetig auf dem kompakten Intervall  $[a, b]$  ist, können wir ein  $\delta$  wählen, sodass  $|g(s) - g(t)| < \varepsilon$  für  $|s - t| < \delta$ ,  $s, t \in [a, b]$ , und wie folgt abschätzen:

$$|f'_m(s) - f'_m(t)| \leq |f'_m(s) - g(s)| + |g(s) - g(t)| + |g(t) - f'_m(t)| < 3\varepsilon.$$



Ist also  $|h| \leq \frac{\delta}{2}$ , so gilt

$$|A_n(h) - A_m(h)| = |f'_m(\xi_{m,h}) - f'_n(\xi_{n,h})| \leq |f'_m(\xi_{m,h}) - f'_m(\xi_{n,h})| + |f'_m(\xi_{n,h}) - f'_n(\xi_{n,h})| < 4\varepsilon,$$

da  $|\xi_{m,h} - \xi_{n,h}| < 2|h| \leq \delta$  ist. Auf der anderen Seite ist die Folge von Funktionen  $A_n$ , betrachtet auf dem Intervall  $[-c, -\delta]$  bzw.  $[c, \delta]$ , klarerweise gleichmässig konvergent (mit Grenzwert  $A$ ); wir können also ein  $\tilde{N}$  wählen, sodass auch  $|A_n(h) - A_m(h)| < 4\varepsilon$ , wenn  $h \in [-c, -\delta] \cup [c, \delta]$  und  $n, m \geq \tilde{N}$ . Zusammengenommen ist dann für  $n, m \geq \max(N, \tilde{N})$  und für jedes  $h \in [-c, c]$

$$|A_n(h) - A_m(h)| < 4\varepsilon,$$

also  $A_n$  eine Cauchyfolge wie behauptet.  $\square$

**Übungsaufgabe 40.** Ist die Konvergenz der abgeleiteten Funktionen eine notwendige Voraussetzung für die Differenzierbarkeit der Grenzfunktion? D.h. gibt es ein Beispiel einer Funktionenfolge  $f_n$ , für welche  $f'_n$  nicht gleichmässig konvergiert, und  $f = \lim_n f_n$  tatsächlich nicht differenzierbar ist?

#### 4. Implizite Funktionen

Oft haben wir Funktionen als die Lösung von Gleichungen gegeben; die Wurzelfunktion  $\sqrt{x}$  ist zum Beispiel als die Lösung  $y(x)$  der Gleichung  $y^2 = x$  gegeben. Der Satz über implizite Funktionen erlaubt es uns, Gleichungen von der Form  $F(x, y) = 0$  "nach  $y$ " aufzulösen, d.h. eine Funktion  $y(x)$  zu finden, für die  $F(x, y(x)) = 0$  gilt. Geometrisch interpretiert man das, indem man in der Nullstellenmenge von  $F$  den Graphen einer Funktion von  $x$  sucht. Nicht immer erwartet man, dass Lösungen impliziter Gleichungen wieder differenzierbar sind.

**Übungsaufgabe 41.** Zeige, dass die Gleichung  $y^2 = x^2$  eine Lösung besitzt, welche an der Stelle 0 nicht differenzierbar ist.

Um allgemeine Aussagen über die Lösungen impliziter Gleichungen machen zu können, müssen wir demnach Anforderungen an  $F$  stellen. Die Voraussetzungen, welche wir hier verwenden wollen, um  $y(x)$  zu finden, sind an die "partiellen" Ableitungen der Funktion  $F$  gestellt. Sei  $F(x, y)$  also eine Funktion von 2 reellen Variablen  $x, y$ , definiert auf dem Rechteck  $R = [a, b] \times [c, d]$ . Wir sagen,  $F$  ist auf  $R$  partiell nach  $x$  bzw.  $y$  differenzierbar, wenn der Grenzwert

$$\frac{\partial F}{\partial x}(x, y) = F_x(x, y) = \lim_{h \rightarrow 0} \frac{F(x+h, y) - F(x, y)}{h} \quad \text{bzw.} \quad \frac{\partial F}{\partial y}(x, y) = F_y(x, y) = \lim_{k \rightarrow 0} \frac{F(x, y+k) - F(x, y)}{k}$$

für jedes  $(x, y) \in R$  existiert. Die Funktion  $F_x$  (bzw.  $F_y$ ) ist die *partiellen Ableitung* von  $F$  nach  $x$  (bzw.  $y$ ).

**SATZ 38.** Sei  $F(x, y)$  auf einem Rechteck  $R = [-a, a] \times [-b, b]$  definiert und stetig **sowie**  $F(0, 0) = 0$ . Weiters existiere die partielle Ableitung  $F_y(x, y)$  auf  $R$  und es sei  $F_y$  stetig auf  $R$  sowie  $F_y(0, 0) \neq 0$ . Dann gibt es ein  $\varepsilon > 0$  und eine stetige Funktion  $y(x)$  auf  $(-\varepsilon, \varepsilon)$  sodass  $F(x, y(x)) = 0$  ist.

Ist darüberhinaus auch  $F_x$  auf  $R$  stetig, so ist  $y$  differenzierbar, und es gilt

$$y'(x) = -\frac{F_x(x, y(x))}{F_y(x, y(x))}.$$

**BEWEIS.** Wir können oBdA annehmen, dass  $F_y(0, 0) > 0$  ist. Da  $F_y$  stetig ist, ist auf einem möglicherweise kleineren Rechteck  $F_y(x, y) > 0$ ; wir ersetzen unser gegebenes Rechteck gegebenenfalls durch ein solches und nehmen an, dass  $F_y > 0$  auf  $R$  gilt. Für jedes  $x \in [-a, a]$  ist die Funktion  $y \mapsto F(x, y)$  dann strikt monoton steigend auf  $[-b, b]$ . Insbesondere gilt  $F(0, -b) < 0$  und  $F(0, b) > 0$ ; indem wir die Stetigkeit von  $F$  verwenden und uns falls notwendig wiederum auf ein kleineres Rechteck einschränken, können wir annehmen, dass  $F(x, -b) < 0$  und  $F(x, b) > 0$  für  $x \in [-a, a]$  gilt.

Wir definieren nun für fixes  $x \in [-a, a]$  den Wert  $y(x)$  als die eindeutige Lösung der Gleichung  $F(x, y) = 0$  welche nach Korollar 5 existiert, und zeigen, dass  $y$  stetig ist. Dazu genügt es zu zeigen, dass  $y(x)$  stetig am Punkt 0 ist; die anderen Punkte können durch die Anwendung desselben Arguments erhalten werden. Angenommen also,  $y(x)$  ist nicht stetig im Punkt 0. Da nach Voraussetzung  $y(0) = 0$  ist, gibt es deswegen ein  $\delta > 0$  sodass  $|y(x_j)| > \delta$  für eine Folge  $x_j$  mit  $x_j \rightarrow 0$  für  $j \rightarrow \infty$ . Die Folge  $(x_j, y(x_j))$  ist dann eine Folge im kompakten Rechteck  $R$ ; durch Übergang zu einer konvergenten Teilfolge können wir also annehmen, dass

$(0, y_0) = \lim_{j \rightarrow \infty} (x_j, y(x_j))$  ist, und  $|y_0| > \delta > 0$ . Da  $F$  stetig ist, ist  $F(0, y_0) = \lim_{j \rightarrow \infty} F(x_j, y_j) = 0$ ; nach Voraussetzung ist aber  $y = 0$  die eindeutige Lösung von  $F(0, y) = 0$ , was einen Widerspruch darstellt.

Um die Differenzierbarkeit von  $y$  zu zeigen, bemerken wir zunächst, dass

$$\lim_{h \rightarrow 0} \frac{F(x, y(x)) - F(x+h, y(x+h))}{h} = 0,$$

da  $f(x, y(x)) = f(x+h, y(x+h)) = 0$ , gilt. Die Differenz können wir mit Hilfe des Mittelwertsatzes als

$$\begin{aligned} F(x, y(x)) - F(x+h, y(x+h)) &= (F(x, y(x)) - F(x+h, y(x))) + (F(x+h, y(x)) - F(x+h, y(x+h))) \\ &= -F_x(\xi_h, y(x))h + F_y(x+h, \eta_h)(y(x) - y(x+h)), \end{aligned}$$

wo  $|x - \xi_h| < |h|$  und  $|\eta_h - y(x)| < |y(x) - y(x+h)|$  ist. Damit gilt, dass

$$\lim_{h \rightarrow 0} \frac{y(x) - y(x+h)}{h} = \lim_{h \rightarrow 0} \frac{F_x(\xi_h, y(x))}{F_y(x+h, \eta_h)} = \frac{F_x(x, y(x))}{F_y(x, y(x))},$$

da der Grenzwert auf der rechten Seite existiert; also ist  $y$  differenzierbar an der Stelle  $x$  mit der behaupteten Formel für die Ableitung.  $\square$

**Übungsaufgabe 42.** Im Beweis des Satzes haben wir Korollar 5 angewendet. Zeige, dass man eine lokale Form dieses Korollars auch aus dem Satz über implizite Funktionen folgern kann: Wenn  $f(x)$  eine Funktion ist, welche differenzierbar in einer Umgebung von 0 ist, und  $f'(0) > 0$ , dann gibt es ein  $\varepsilon > 0$  und ein  $g$ , welches auf  $[f(-\varepsilon), f(\varepsilon)]$  definiert ist und  $f \circ g = g \circ f = \text{id}$  erfüllt.

**Übungsaufgabe 43.** Die Funktionen  $x^r$ , wo  $r$  eine rationale Zahl ist, sind an allen Stellen  $x \neq 0$  differenzierbar. Finde auch eine Formel für die Ableitung.

REMINDER: vielleicht auch den Fixpunktsatz- Beweis??

## 5. Taylor's Satz und Folgerungen

Differenzierbarkeit bedeutet, dass  $f(x+h)$  gut durch die affine Funktion  $f(x) + hf'(x)$  approximiert wird, in dem Sinn, dass

$$R(x, h) = f(x+h) - f(x) - f'(x)h$$

die Eigenschaft hat, dass

$$\lim_{h \rightarrow 0} \frac{R(x, h)}{h} = 0.$$

Die genaue Geschwindigkeit, mit der  $R(x, h)$  gegen 0 geht, ist damit nicht bestimmt. Oft wünscht man sich, dass der Restterm  $R(x, h)$  die Eigenschaft hat, dass sogar  $h^{-2}R(x, h)$  für  $h \rightarrow 0$  beschränkt bleibt. Dies ist nicht immer der Fall, wie das Beispiel von  $x^\alpha$  für  $\alpha < 2$  (um  $x = 0$ ) zeigt. Diese Funktion ist allerdings auch nur 1-mal stetig differenzierbar; die zweite Ableitung ist in keiner Umgebung von 0 beschränkt und bei 0 nicht definiert. Höhere Differenzierbarkeit erlaubt es uns, die lineare Approximation zu verbessern.

Nach dem Mittelwertsatz ist (da  $R(x, 0) = 0$ )

$$R(x, h) = hR_h(x, k)$$

für ein  $k$  mit  $|k| < |h|$ . Nun ist  $R_h(x, h) = f'(x+h) - f'(x)$ , also  $R_h(x, 0) = 0$  und  $R_{h^2}(x, h) = f''(x+h)$ . Eine weitere Anwendung des Mittelwertsatzes zeigt, dass  $R(x, h) = hkf''(x+\ell)$  für ein  $\ell$  mit  $|\ell| < |k|$ , was zeigt, dass  $h^{-2}R(x, h)$  beschränkt ist (für  $h \rightarrow 0$ ). Wir formulieren eine Variante dieser letzten Beobachtung als

**LEMMA 69.** Sei  $R(h) \in C^k([a, b]) \cap D^{k+1}((a, b))$ ,  $a < 0 < b$ , und  $R(0) = R'(0) = \dots = R^{(k)}(0) = 0$ . Dann gibt es für jedes  $h \in [a, b]$  ein  $\xi_h$  mit  $|\xi_h| < |h|$  sodass

$$R(h) = \frac{h^{k+1}}{(k+1)!} R^{(k+1)}(\xi).$$

BEWEIS. Wir definieren

$$f(t) = R(t) + R'(t)(h-t) + R''(t)\frac{(h-t)^2}{2!} + \dots + R^{(k)}(t)\frac{(h-t)^k}{k!}, \quad g(t) = (t-h)^{k+1}.$$

Nach Übungsaufgabe 38 gibt es ein  $\xi$  mit  $|\xi| < |h|$  sodass

$$\begin{aligned} 0 &= f'(\xi)(g(h) - g(0)) - g'(\xi)(f(h) - f(0)) \\ &= R^{(k+1)}(\xi)\frac{(h-\xi)^k}{k!}(-(-h)^{k+1}) - (k+1)(\xi-h)^k R(h), \end{aligned}$$

was nach einer einfachen Umformung die Behauptung ergibt.  $\square$

Wir können nun unsere erste Form der Taylorentwicklung formulieren:

SATZ 39. Sei  $f \in C^k([a, b]) \cap D^{k+1}((a, b))$ ,  $a < x < b$ . Dann gibt es für jedes  $h$  mit  $x+h \in (a, b)$  ein  $\xi \in (x, x+h)$  sodass

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \dots + \frac{f^{(k)}(x)}{k!}h^k + \frac{f^{(k+1)}(\xi)}{(k+1)!}h^{k+1}.$$

BEWEIS. Wir betrachten die Differenz

$$R(h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \dots + \frac{f^{(k)}(x)}{k!}h^k - f(x+h).$$

Diese erfüllt die Voraussetzungen von Lemma 69, und  $R^{(k+1)}(h) = f^{(k+1)}(x+h)$ ; das ergibt die angegebene Form des Rests.  $\square$

Für die erste Anwendung wiederholen wir die Definition eines lokalen Extremums:

**Definition 37.** Sei  $f$  eine Funktion, welche in einer Umgebung von  $x$  definiert ist. Dann hat  $f$  ein *lokales Maximum (Minimum)* an der Stelle  $x$  wenn für genügend kleine  $h$   $f(x+h) \leq f(x)$  (bzw.  $f(x+h) \geq f(x)$ ) ist. Das Maximum ist *strikt*, wenn diese Ungleichungen für  $h \neq 0$  strikte Ungleichungen sind.

KOROLLAR 8. Sei  $f$  in einer Umgebung von  $x$  zweimal stetig differenzierbar,  $f'(x) = 0$ . Dann hat  $f$  ein *lokales Minimum* an der Stelle  $x$ , wenn  $f''(x) > 0$ , und ein *lokales Maximum*, wenn  $f''(x) < 0$ .

**Bemerkung 11.** Wie das Beispiel von  $x^3$  bzw.  $x^4$  zeigt, kann falls  $f''(0) = 0$  ist keine weitere Aussage ohne weitere Voraussetzungen möglich.

BEWEIS. Für kleine  $h > 0$  ist

$$f(x+h) = f(x) + \frac{f''(\xi)}{2}h^2 \begin{cases} > f(x), & f''(\xi) > 0 \\ < f(x), & f''(\xi) < 0, \end{cases}$$

wobei  $\xi \in (x, x+h)$ . Wenn  $h$  genügend klein gewählt ist, hat  $f''(\xi)$  dasselbe Vorzeichen wie  $f''(x)$ .  $\square$

**Übungsaufgabe 44.** Zeige die folgende Verallgemeinerung von Korollar 8: Sei  $f$  in einer Umgebung von  $x$   $k$ -mal stetig differenzierbar,  $f'(x) = \dots = f^{(k-1)}(x) = 0$ , und  $f^{(k)}(x) \neq 0$ . Dann ist  $x$  kein lokales Extremum, wenn  $k$  ungerade ist; wenn  $k$  gerade ist, so hat  $f$  ein lokales Minimum an der Stelle  $x$ , wenn  $f^{(k)}(x) > 0$ , und ein lokales Maximum, wenn  $f^{(k)}(x) < 0$ .

Unser nächstes Korollar zeigt, dass die Konvergenz in der Newton-Methode besser als in Korollar 7 angegeben ist.

KOROLLAR 9. Sei  $f$  in einer Umgebung von 0 zweimal stetig differenzierbar,  $f(0) = 0$ , und  $f'(0) \neq 0$ . Dann gibt es eine Umgebung von 0 und eine Konstante  $M > 0$  sodass die durch

$$g(x) = x - \frac{f(x)}{f'(x)}$$

definierte Funktion  $|g(x)| < M|x|^2$  erfüllt.

BEWEIS. Wir wenden Theorem 39 am Punkt  $x$  an und sehen, dass

$$0 = f(0) = f(x) + f'(x)(-x) + \frac{f''(\xi)}{2}x^2.$$

Also ist

$$|g(x)| = \left| x - \frac{f(x)}{f'(x)} \right| = \left| \frac{f''(\xi)}{2f'(x)} \right| |x|^2 \leq M|x|^2$$

wenn  $x$  (und damit auch  $\xi$ ) nahe genügend an 0 sind. □

**Übungsaufgabe 45.** Im Beweis des vorangehenden Satzes ist nicht besonders klar geworden, wie klein nun die Umgebung zu wählen ist, in der eine quadratische Ungleichung gilt. Versuche eine genauere Beschreibung einer solchen Umgebung zu geben.

**Definition 38.** Sei  $I$  ein (geschlossenes oder offenes) Intervall. Eine Funktion  $\varphi: I \rightarrow \mathbb{R}$  ist *konvex* (bzw. *konkav*) auf  $I$ , wenn für alle  $a, b \in I$  mit  $a < b$  und alle  $\lambda, \mu \in [0, 1]$  mit  $\lambda + \mu = 1$

$$f(\lambda a + \mu b) \leq \lambda f(a) + \mu f(b) \text{ bzw. } \geq \lambda f(a) + \mu f(b)$$

ist.  $\varphi$  ist *strikt konvex* bzw. *strikt konkav*, wenn die Ungleichung für  $\lambda\mu \neq 0$  strikt ist.

**KOROLLAR 10.** Sei  $\varphi: I \rightarrow \mathbb{R}$  zweimal stetig differenzierbar. Dann ist  $\varphi$  strikt konvex auf  $I$ , wenn  $\varphi''$  auf  $I$  positiv ist, und strikt konkav, wenn  $\varphi''$  auf  $I$  negativ ist.

BEWEIS. Seien  $a, b \in I$  fix,  $a < b$ . Wir betrachten die Funktion  $f(t): [0, 1] \rightarrow \mathbb{R}$ , welche durch

$$f(t) = \varphi((1-t)a + tb) - (1-t)\varphi(a) - t\varphi(b)$$

definiert ist. Dann ist  $f(0) = f(1) = 0$ . Wenn  $f$  nun an einer Stelle von  $(0, 1)$  nichtnegativ ist, so muss es in  $(a, b)$  eine Stelle eines lokalen Maximums geben. Es ist  $f'(t) = \varphi'((1-t)a + tb)(b-a) - \varphi(a) - \varphi(b)$ , und  $f''(t) = \varphi''((1-t)a + tb)(b-a)^2 > 0$ . Also kann  $f$  nach Korollar 8 nur lokale Minima haben, und  $f$  ist auf  $(0, 1)$  positiv.

Der Fall einer konkaven Funktion wird ganz analog behandelt. □

Wenn eine Funktion  $\varphi: [a, b] \rightarrow \mathbb{R}$  strikt konvex (konkav) ist, so gibt es genau einen Punkt  $x \in [a, b]$ , wo  $\varphi$  ein lokales (und in diesem Fall auch globales) Minimum (Maximum) hat; das hat zwar mit Taylor's Formel nicht viel zu tun, ist aber ein wichtiges Kriterium, weswegen wir es an dieser Stelle notieren:

**LEMMA 70.** Sei die Funktion  $\varphi: [a, b] \rightarrow \mathbb{R}$  strikt konvex und stetig. Dann gibt es genau einen Punkt  $x \in [a, b]$  wo  $\varphi$  sein Minimum auf  $[a, b]$  annimmt.

BEWEIS. Die Existenz des Minimums wird wegen der Stetigkeit von  $\varphi$  garantiert. Angenommen,  $x_0 \neq x_1$  sind zwei Punkte wo  $\varphi(x_0) = \varphi(x_1) = \min_{x \in [a, b]} \varphi(x)$ . Dann ist für  $\lambda, \mu \in [0, 1]$  mit  $\lambda + \mu = 1$  wegen der strikten Konvexität von  $\varphi$

$$\varphi(\lambda x_0 + \mu x_1) < \lambda \varphi(x_0) + \mu \varphi(x_1) = \min_{x \in [a, b]} \varphi(x),$$

was einen Widerspruch darstellt. □

**Bemerkung 12.** Wir bemerken, dass die Eindeutigkeitsaussage auch erhalten bleibt, wenn man die Existenz eines Minimums voraussetzt (ohne dabei Stetigkeit zu verlangen).

# Das Riemannintegral

## 1. Die Definition des Riemannintegrals

Sei  $[a, b]$  ein kompaktes Intervall. Eine Unterteilung  $U$  von  $[a, b]$  ist ein geordnetes  $n+1$ -tupel von Punkten

$$U = (x_0, x_1, \dots, x_n), \quad \text{wo } a = x_0 < x_1 < \dots < x_{n-1} < x_n = a,$$

ein Vektor von Zwischenpunkten  $\xi$  für  $U$  ist ein geordnetes  $n$ -tupel

$$\xi = (\xi_1, \dots, \xi_n), \quad x_j \leq \xi_{j+1} \leq x_{j+1}.$$

Wir definieren  $\delta(U) = \max_j(x_j - x_{j-1})$  als die *Feinheit* von  $U$ . Die *Riemannsumme* einer Funktion  $f: [a, b] \rightarrow \mathbb{R}$  (bezüglich  $U$  und  $\xi$ ) ist definiert durch

$$R(f, U, \xi) = \sum_{j=1}^n f(\xi_j)(x_j - x_{j-1}).$$

**Definition 39.** Die beschränkte Funktion  $f: [a, b] \rightarrow \mathbb{R}$  ist *integrierbar* (im Sinn von Riemann) wenn für jede Folge von Unterteilungen  $U^j$  und jede Folge von Zwischenpunkten  $\xi^j$  für  $U^j$  mit  $\delta(U^j) \rightarrow 0$  ( $j \rightarrow \infty$ ) die Folge der Riemannsummen  $R(f, U^j, \xi^j)$  konvergiert. Der (gemeinsame) Grenzwert dieser Folgen wird dann als das Riemannintegral von  $f$  über  $[a, b]$  bezeichnet und als

$$\lim_{j \rightarrow \infty} R(f, U^j, \xi^j) = \int_a^b f(x) dx$$

geschrieben; die Menge der im Sinne von Riemann integrierbaren Funktionen wird mit  $R([a, b])$  bezeichnet.

Die Definition des Riemannintegrals mit Hilfe von Folgenlimiten hat den Vorteil, dass wir einfache Eigenschaften desselben mit Hilfe der Sätze über das Verhalten von Folgenlimiten nachweisen können; sie hat den Nachteil, dass wir noch nicht viele Beispiele von Riemann-integrierbaren Funktionen angeben können. Zunächst wollen wir die Abgeschlossenheit unter einfachen algebraischen Operationen sowie das Monotonieverhalten des Integrals besprechen.

LEMMA 71. Seien  $f, g \in R([a, b])$ . Dann ist für jedes  $\lambda \in \mathbb{R}$  auch  $f + \lambda g \in R([a, b])$ , und es gilt

$$\int_a^b (f(x) + \lambda g(x)) dx = \int_a^b f(x) dx + \lambda \int_a^b g(x) dx.$$

Ist für jedes  $x \in [a, b]$  die Ungleichung  $f(x) \leq g(x)$  erfüllt, so gilt

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

Ist  $a \leq c \leq b$  und  $f \in R([a, b])$ , so sind die Einschränkungen von  $f$  auf  $[a, c]$  und auf  $[c, b]$  über die jeweiligen Intervalle Riemann-integrierbar, und es gilt

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

**Übungsaufgabe 46.** Beweise Lemma 71.

**Übungsaufgabe 47.** Zeige, dass jede Riemann-integrierbare Funktion  $f$  beschränkt ist.

Es ist bequem, für  $a < b$

$$\int_b^a f(x) dx = - \int_a^b f(x) dx$$

zu definieren.

In Wirklichkeit kommt es beim Riemannintegral nur auf die Feinheit der betreffenden Unterteilungen an, wie die nächsten Lemmata zeigen. Zunächst zeigen wir eine Art Cauchy Kriterium für die Riemann-Integrierbarkeit.

LEMMA 72. *Sei  $f$  beschränkt auf  $[a, b]$ . Dann ist  $f$  Riemann-integrierbar auf  $[a, b]$  genau dann, wenn für jedes  $\varepsilon > 0$  ein  $\delta > 0$  existiert, sodass für beliebige Unterteilungen  $U^1, U^2$  mit  $\delta(U^1) < \delta, \delta(U^2) < \delta$  und beliebige Vektoren von Zwischenpunkten  $\xi$  für  $U^1$  und  $\eta$  für  $U^2$  die Ungleichung*

$$|R(f, U^1, \xi) - R(f, U^2, \eta)| < \varepsilon$$

erfüllt ist.

BEWEIS. Ist die Bedingung des Lemmas erfüllt, so ist für jede Folge von Zerlegungen  $U^k$  und Vektoren von Zwischenpunkten  $\xi^k$  für  $U^k$  die Folge  $R(f, U^k, \xi^k)$  eine Cauchyfolge, also konvergent, und damit  $f$  Riemann integrierbar.

Wenn die Bedingung des Lemmas nicht erfüllt ist, so gibt es ein  $\varepsilon_0 > 0$  sodass für jedes  $k$  in  $\mathbb{N}$  eine Unterteilung  $U^k$  und ein Vektor von Zwischenpunkten  $\xi^k$  sowie eine Unterteilung  $V^k$  und ein Vektor von Zwischenpunkten  $\eta^k$  existiert, welche

$$\delta(U^k) \leq \frac{1}{k}, \quad \delta(V^k) \leq \frac{1}{k}, \quad |R(f, U^k, \xi^k) - R(f, V^k, \eta^k)| \geq \varepsilon_0$$

erfüllen. Die Mischung dieser Zerlegungsfolgen und Zwischenpunkten, d.h. die Folge  $W_k$  von Zerlegungen welche durch  $W_{2k} = U_k$  und  $W_{2k+1} = V_k$  und die Folge von Zwischenpunkten  $\zeta_k$  welche durch  $\zeta_{2k} = \xi_k$  und  $\zeta_{2k+1} = \eta_k$  gegeben ist erfüllt dann  $\delta(W_k) \rightarrow 0$ , ist aber offensichtlich keine Cauchyfolge, konvergiert also nicht; damit ist  $f$  also nicht Riemann-integrierbar.  $\square$

LEMMA 73. *Sei  $f$  beschränkt auf  $[a, b]$ . Dann ist  $f$  Riemann-integrierbar auf  $[a, b]$  genau dann, wenn es ein  $I \in \mathbb{R}$  gibt, sodass für jedes  $\varepsilon > 0$  ein  $\delta > 0$  existiert, welches die Eigenschaft hat, dass für jede Unterteilung  $U$  mit  $\delta(U) < \delta$  und jeden Vektor von Zwischenpunkten  $\xi$  für  $U$  die Ungleichung*

$$|R(f, U, \xi) - I| < \varepsilon$$

erfüllt ist. In diesem Fall ist

$$I = \int_a^b f(x) dx.$$

BEWEIS. Wenn die Bedingung des Lemmas erfüllt ist, ist es einfach zu zeigen, dass  $f$  Riemann-integrierbar ist. Ist auf der anderen Seite  $f$  Riemann integrierbar, so sei  $I = \int_a^b f(x) dx$ , und  $\varepsilon > 0$  beliebig. Wir wählen  $\delta$  mit Hilfe von Lemma 72. Sei nun  $U^k$  eine Folge von Unterteilungen und  $\xi^k$  eine Folge von Zwischenvektoren mit  $\delta(U_k) \rightarrow 0$  für  $k \rightarrow \infty$ , dann konvergiert nach Voraussetzung  $R(f, U^k, \xi^k) \rightarrow I$  für  $k \rightarrow \infty$ . Wir wählen ein  $N$  sodass  $|R(f, U^N, \xi^N) - I| < \varepsilon$  und  $\delta(U^N) < \delta$ . Dann ist für jede andere Zerlegung  $U$  und jeden Vektor von Zwischenpunkten  $\xi$  mit  $\delta(U) < \delta$

$$|R(f, U, \xi) - I| \leq |R(f, U, \xi) - R(f, U^N, \xi^N)| + |R(f, U^N, \xi^N) - I| < 2\varepsilon.$$

$\square$

Wir haben noch immer keine Beispiele Riemann-integrierbarer Funktionen; dies wollen wir nun beheben. Wir erinnern zunächst daran, dass die charakteristische Funktion einer Menge  $E \subset \mathbb{R}$  durch

$$\chi_E(x) = \begin{cases} 0 & x \notin E \\ 1 & x \in E \end{cases}$$

definiert ist.

**Definition 40.** Wir sagen, eine Funktion  $t$  ist eine Treppenfunktion, wenn  $t$  eine Linearkombination endlich vieler charakteristischer Funktionen von nach rechts halboffenen Intervallen ist; d.h. wenn es endlich viele  $\lambda_j \in \mathbb{R}$  und  $a_j < b_j \in \mathbb{R}$ ,  $j = 1, \dots, N$ , gibt sodass

$$t(x) = \sum_{j=1}^N \lambda_j \chi_{[a_j, b_j)}$$

ist. Der Raum aller Treppenfunktionen bildet einen Vektorraum über  $\mathbb{R}$ , den wir mit  $T(\mathbb{R})$  bezeichnen. Die Einschränkungen von Funktionen in  $T(\mathbb{R})$  auf ein Intervall  $[a, b]$  (oder eine beliebige Teilmenge  $E$  von  $\mathbb{R}$ ) wird mit  $T([a, b])$  (bzw.  $T(E)$ ) bezeichnet.

LEMMA 74. Jedes  $t \in T([a, b])$  ist Riemann-integrierbar. Wenn

$$t(x) = \sum_{j=1}^N \lambda_j \chi_{[a_j, b_j)}$$

mit  $a \leq a_j < b_j \leq b$  für alle  $j = 1, \dots, N$  ist, so ist

$$\int_a^b t(x) dx = \sum_{j=1}^N \lambda_j (b_j - a_j).$$

BEWEIS. Es genügt nach den Linearitätsaussagen von Lemma 71, den Fall  $N = 1$  zu betrachten, also  $t(x) = \lambda \chi_{[a_1, b_1)}$ . Sei  $U$  eine Unterteilung von  $[a, b]$  und  $\xi$  ein Zwischenvektor für  $U$ , dann ist

$$\begin{aligned} |R(t, U, \xi) - \lambda(a_1 - b_1)| &= \left| \sum_{j=1}^{n-1} \chi_{[a_1, b_1)}(\xi_j)(x_{j+1} - x_j) - \lambda(a_1 - b_1) \right| \\ &= \left| \sum_{a \leq \xi_j < b} \lambda(x_j - x_{j-1}) - \lambda(a_1 - b_1) \right| \\ &\leq 2\delta(U). \end{aligned}$$

Nach Lemma 73 ist also  $t(x)$  Riemann-integrierbar, mit dem gewünschten Integral.  $\square$

**1.1. Das Darboux-Integral.** Ein anderer Zugang zum Integralbegriff ist mit Hilfe von Ober- und Untersummen; hier wird nicht notwendigerweise mit Funktionswerten gewichtet:

**Definition 41.** Sei  $U$  eine Zerlegung von  $[a, b]$ , und  $f$  eine beschränkte Funktion auf  $[a, b]$ . Dann sind

$$R_+(f, U) = \sum_{j=1}^n \left( \sup_{\xi \in [x_{j-1}, x_j]} f(\xi) \right) (x_j - x_{j-1}), \quad R_-(f, U) = \sum_{j=1}^n \left( \inf_{\xi \in [x_{j-1}, x_j]} f(\xi) \right) (x_j - x_{j-1})$$

die *Riemannsche Ober- und Untersumme* von  $f$  auf  $[a, b]$ . Wir bezeichnen mit

$$D_+(f) = \inf_U R_+(f, U), \quad D_-(f) = \inf_U R_-(f, U)$$

das *obere und untere Darbouxintegral* von  $f$ .

LEMMA 75. Sei  $f$  beschränkt auf  $[a, b]$ . Dann gilt  $D_-(f) \leq D_+(f)$ , und  $D_-(f) = D_+(f)$  genau dann, wenn für jedes  $\varepsilon > 0$  eine Unterteilung  $U$  existiert mit

$$R_+(f, U) - R_-(f, U) < \varepsilon.$$

BEWEIS. Für jede Unterteilung  $U$  gilt offensichtlich  $R_-(f, U) \leq R_+(f, U)$ . Seien nun  $U^1$  und  $U^2$  beliebige Unterteilungen. Wir sagen, eine Unterteilung  $\tilde{U}$  ist eine *Verfeinerung* von  $U$ , wenn sie alle Unterteilungspunkte von  $U$  enthält. Es ist

$$R_-(f, U) \leq R_-(f, \tilde{U}), \quad R_+(f, \tilde{U}) \leq R_+(f, U)$$

für jede solche Verfeinerung  $\tilde{U}$ . Wir wählen  $\tilde{U}$  als gemeinsame Verfeinerung von  $U^1$  und  $U^2$  und erhalten so

$$R_-(f, U^1) \leq R_-(f, \tilde{U}) \leq R_+(f, \tilde{U}) \leq R_+(f, U^2).$$

Nachdem  $U^1$  und  $U^2$  beliebig waren, können wir links zum Supremum und rechts zum Infimum übergehen und erhalten  $D_-(f) \leq D_+(f)$ .

Ist  $D_-(f) = D_+(f) = I$ , so gibt es für jedes  $\varepsilon > 0$  eine Unterteilung  $U^1$  mit  $I - R_-(f, U^1) < \varepsilon$  und eine Unterteilung  $U^2$  welche  $R_+(f, U) - I < \varepsilon$  erfüllt. Damit ist für jede gemeinsame Verfeinerung  $\tilde{U}$  von  $U^1$  und  $U^2$

$$R_+(f, \tilde{U}) - R_-(f, \tilde{U}) \leq R_+(f, U^1) - R_-(f, U^2) < 2\varepsilon.$$

Die andere Richtung ist trivial.  $\square$

Unser nächstes Lemma erlaubt es uns, eine Verbindung zwischen Darboux- und Riemannintegralen herzustellen.

LEMMA 76. *Sei  $f$  beschränkt auf  $[a, b]$  und  $U$  eine Unterteilung von  $[a, b]$  mit  $R_+(f, U) - R_-(f, U) < \varepsilon$ . Dann gibt es ein  $\delta > 0$  sodass für jede Unterteilung  $V$  von  $[a, b]$  mit  $\delta(V) < \delta$ , dann gilt  $|R(f, V, \xi) - R_-(f, U)| < 2\varepsilon$  und  $|R(f, V, \xi) - R_+(f, U)| < 2\varepsilon$ .*

BEWEIS. Sei  $U = (x_0, x_1, \dots, x_n)$  und  $V = (y_0, y_1, \dots, y_m)$  Unterteilungen wie im Lemma. Wir zerlegen

$$\begin{aligned} R(f, V, \xi) - R_-(f, U) &= \sum_j f(\xi_j)(y_{j+1} - y_j) - \sum_k \inf_{x_k \leq \eta \leq x_{k+1}} f(\eta)(x_{k+1} - x_k) \\ &= \sum_{k=0}^{n-1} \left( \sum_{\xi_j \in [x_k, x_{k+1}]} f(\xi_j)(y_{j+1} - y_j) - \inf_{x_k \leq \eta \leq x_{k+1}} f(\eta)(x_{k+1} - x_k) \right), \end{aligned}$$

wobei wir im Summanden für  $k = n - 1$  statt  $[x_{n-1}, x_n)$  das abgeschlossene Intervall  $[x_{n-1}, x_n]$  verwenden wollen.

Nach Anwendung der Dreiecksungleichung sehen wir also, dass wir

$$\left| \sum_{\xi_j \in [x_k, x_{k+1}]} f(\xi_j)(y_{j+1} - y_j) - \inf_{x_k \leq \eta \leq x_{k+1}} f(\eta)(x_{k+1} - x_k) \right|$$

abschätzen müssen. Wir bezeichnen mit  $y_J$  das kleinste  $y_j \neq x_k$ , welches in der ersten Summe vorkommt, und mit  $y_K \neq x_{k+1}$  das grösste; es ist dann notwendigerweise  $|y_J - x_k| \leq \delta(V)$  und  $|y_K - x_{k+1}| \leq \delta(V)$ . Wir können also mit  $M > \sup_{x \in [a, b]} |f(x)|$  wie folgt abschätzen:

$$\begin{aligned} &\left| \sum_{j=J}^K f(\xi_j)(y_{j+1} - y_j) - \inf f(\eta)(x_{k+1} - x_k) \right| \\ &\leq 2M\delta(V) + \left| f(\xi_J)(y_{J+1} - x_k) + f(\xi_K)(x_{k+1} - y_K) + \sum_{j=J+1}^{K-1} f(\xi_j)(y_{j+1} - y_j) - \inf f(\eta)(x_{k+1} - x_k) \right| \\ &\leq 2M\delta(V) + (\sup f(\eta) - \inf f(\eta))(x_{k+1} - x_k). \end{aligned}$$

Summieren wir diese Abschätzung nun über  $k$ , so erhalten wir

$$|R(f, V, \xi) - R_-(f, U)| \leq 2nM\delta(V) + R_+(f, U) - R_-(f, U).$$

Damit gilt die Aussage des Lemmas für die Untersumme, wenn wir  $\delta > 0$  so klein wählen, dass  $\delta 2nM < \varepsilon$ . Die Aussage des Lemmas für die Obersumme wird ganz analog bewiesen.  $\square$

Der Zusammenhang zwischen den Darbouxintegralen und der Riemann-Integrierbarkeit wird im folgenden Satz zusammengefasst.

SATZ 40. *Sei  $f$  beschränkt auf  $[a, b]$ . Dann ist  $f$  integrierbar im Sinne von Riemann genau dann, wenn  $D_+(f) = D_-(f)$ , und in diesem Fall ist*

$$D_+(f) = D_-(f) = \int_a^b f(x) dx.$$



BEWEIS. Sei zunächst  $D_+(f) = D_-(f)$ , und  $\varepsilon > 0$  beliebig. Nach Lemma 75 gibt es eine Unterteilung  $U$  mit  $R_+(f, U) - R_-(f, U) < \varepsilon$ . Wir wählen nun  $\delta > 0$  wie in Lemma 76 und erhalten für beliebige Unterteilungen  $V, W$  mit  $\delta(V) < \delta$  und  $\delta(W) < \delta$  und Zwischenvektoren  $\xi$  zu  $V$  sowie  $\eta$  zu  $W$ , dass

$$|R(f, V, \xi) - R(f, W, \eta)| \leq |R(f, V, \xi) - R_+(f, U)| + R_+(f, U) - R_-(f, U) + |R_-(f, U) - R(f, W, \eta)| < 5\varepsilon.$$

Nach Lemma 72 ist  $f$  also Riemann-integrierbar.

Sei nun  $\varepsilon > 0$  wiederum beliebig, und  $\delta > 0$  so wie oben gewählt, aber auch mit der Eigenschaft, dass für jede Unterteilung mit  $\delta(V) < \delta$

$$\left| R(f, V, \xi) - \int_a^b f(x) dx \right| < \varepsilon$$

gilt. Dann ist

$$\left| D_+(f) - \int_a^b f(x) dx \right| \leq |D_+(f) - R_-(f, U)| + |R_-(f, U) - R(f, V, \xi)| + \left| R(f, V, \xi) - \int_a^b f(x) dx \right| < 4\varepsilon.$$

Nachdem  $\varepsilon > 0$  beliebig war, gilt also Gleichheit.

Die Gegenrichtung ist eine einfache Folgerung aus der Charakterisierung in Lemma 75 und der Cauchy-folgencharakterisierung der Riemann-Integrierbarkeit in Lemma 72.  $\square$

Zusammensetzungen von Funktionen werden im folgenden Lemma behandelt.

LEMMA 77. Sei  $f \in R([a, b])$ , und  $g$  eine Funktion, für die  $g \circ f$  definiert ist. Wenn es eine Konstante  $C > 0$  gibt, sodass  $|g(f(x)) - g(f(y))| \leq C|f(x) - f(y)|$  ist, so ist auch  $g \circ f \in R([a, b])$ .

BEWEIS. Wenn  $|g(f(x)) - g(f(y))| \leq C|f(x) - f(y)|$  ist, so gilt für beliebige Mengen  $E$ , dass

$$\sup_{x \in E} g \circ f(x) - \inf_{y \in E} g \circ f(y) \leq C \left( \sup_{x \in E} f(x) - \inf_{y \in E} f(y) \right).$$

Damit ist also für jede Unterteilung  $U$  auch

$$R_+(g \circ f, U) - R_-(g \circ f, U) \leq C(R_+(f, U) - R_-(f, U)),$$

und eine Anwendung der Charakterisierung in Theorem 40 gemeinsam mit Lemma 75 ergibt die Behauptung.  $\square$

## 2. Konvergenzsatz und Folgerungen

Sei  $f_n$  eine Folge von Funktionen auf  $[a, b]$ . Wir erinnern daran, dass  $f_n$  gleichmässig auf  $[a, b]$  gegen  $f$  konvergiert, wenn es für jedes  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  gibt, sodass

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

für alle  $n \geq N$  ist. Das Riemann-Integral ist "stabil" unter gleichmässiger Konvergenz:

SATZ 41. Sei  $f_n$  eine Folge Riemann-integrierbarer Funktionen auf  $[a, b]$ , welche gleichmässig gegen  $f$  konvergiert. Dann ist auch  $f$  Riemann-integrierbar, und es gilt

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

Wenn man die Aussage von Theorem 41 als

$$\int_a^b \lim_{n \rightarrow \infty} f_n(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx$$

schreibt, und sich in Erinnerung ruft, dass das Riemann-Integral als ein Grenzwert definiert ist, so wird klar, dass Theorem 41 eine Aussage über das Vertauschen von Grenzwertprozessen ist. Diese Vertauschbarkeit ist niemals selbstverständlich, und bedarf in jedem Fall eines Beweises.

BEWEIS. Um zu zeigen, dass  $f$  Riemann-integrierbar ist, bedienen wir uns zunächst Lemma 72. Sei also  $\varepsilon > 0$ . Wir wählen ein  $N \in \mathbb{N}$  mit  $|f_n(x) - f(x)| < \varepsilon$  für  $n \geq N$ . Da  $f_N$  Riemann-integrierbar ist, gibt es ein  $\delta > 0$  sodass

$$|R(f_N, U^1, \xi^1) - R(f_N, U^2, \xi^2)| < \varepsilon,$$

wenn  $U^1$  und  $U^2$  Unterteilungen von  $[a, b]$  mit  $\delta(U^j) < \delta$  sind. Dann ist

$$\begin{aligned} |R(f, U^1, \xi^1) - R(f, U^2, \xi^2)| &\leq |R(f, U^1, \xi^1) - R(f_N, U^1, \xi^1)| + |R(f_N, U^1, \xi^1) - R(f_N, U^2, \xi^2)| \\ &\quad + |R(f_N, U^2, \xi^2) - R(f, U^2, \xi^2)| \\ &\leq (b-a)\varepsilon + \varepsilon + (b-a)\varepsilon, \end{aligned}$$

und damit also  $f$  Riemann-integrierbar.

Um die Aussage über den Wert des Integrals zu beweisen, sei wieder  $\varepsilon > 0$ , und  $N \in \mathbb{N}$  so gewählt, dass  $|f_n(x) - f(x)| < \varepsilon$  für  $n \geq N$ . Wir wählen für jedes  $n$  eine Unterteilung  $U^n$  mit Zwischenvektor  $\xi^n$  für welche

$$\left| R(f_n, U^n, \xi^n) - \int_a^b f_n dx \right| < \varepsilon \quad \text{und} \quad \left| R(f, U^n, \xi^n) - \int_a^b f dx \right| < \varepsilon$$

ist. Dann gilt

$$\begin{aligned} \left| \int_a^b f_n(x) dx - \int_a^b f(x) dx \right| &\leq \left| \int_a^b f_n dx - R(f_n, U^n, \xi^n) \right| + |R(f_n, U^n, \xi^n) - R(f, U^n, \xi^n)| \\ &\quad + \left| R(f, U^n, \xi^n) - \int_a^b f dx \right| \\ &\leq \varepsilon + (b-a)\varepsilon + \varepsilon \\ &= (2 + (b-a))\varepsilon \end{aligned}$$

also gilt auch die Aussage über die Vertauschung der Grenzwerte. □

KOROLLAR 11. Sei  $f$  eine stetige Funktion auf  $[a, b]$ . Dann ist  $f$  über  $[a, b]$  Riemann-integrierbar.

BEWEIS. Da  $f$  stetig auf  $[a, b]$  ist, ist  $f$  nach Theorem 30 auch gleichmässig stetig. Für jedes  $n \in \mathbb{N}$  gibt es also ein  $k = k(n) \in \mathbb{N}$  sodass

$$|f(x) - f(y)| < \frac{1}{n}, \quad \text{wenn } |x - y| < \frac{1}{k}.$$

Wir definieren die Treppenfunktion

$$t_n(x) = f \left( a + \frac{j}{n}(b-a) \right), \quad j \frac{(b-a)}{n} \leq x - a < (j+1) \frac{(b-a)}{n}.$$

Dann ist  $|t_n(x) - f(x)| < 1/n$ , also konvergiert  $t_n$  gleichmässig auf  $[a, b]$  gegen  $f$ . Da  $t_n$  nach Lemma 74 Riemann-integrierbar ist, folgt aus Theorem 41, dass auch  $f$  Riemann-integrierbar ist. □

KOROLLAR 12. Sei  $f: [a, b] \times [c, d] \rightarrow \mathbb{R}$  stetig. Dann ist die Funktion

$$g: [c, d] \rightarrow \mathbb{R}, \quad y \mapsto \int_a^b f(x, y) dx$$

stetig auf  $[c, d]$ .

BEWEIS. Sei  $y_0 \in [c, d]$  beliebig. Dann konvergiert für jede Folge  $y_n$  mit  $y_n \rightarrow y_0$  für  $n \rightarrow \infty$  die Folge von Funktionen  $f_n(x) = f(x, y_n)$  gleichmässig gegen  $f_0(x) = f(x, y_0)$ . Also ist nach Theorem 41

$$\lim_{n \rightarrow \infty} g(y_n) = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b f_0(x) dx = g(y_0)$$

und damit  $g$  im Punkt  $y_0$  stetig. □

KOROLLAR 13. Sei  $f: [a, b] \times [c, d] \rightarrow \mathbb{R}$ , und für jedes fixe  $x \in [a, b]$  sei die Funktion  $y \mapsto f(x, y)$  differenzierbar auf  $[c, d]$ . Die Ableitung  $f_y(x, y)$ , die dann wiederum auf  $[a, b] \times [c, d]$  definiert ist, sei auch stetig. Dann ist die Funktion

$$g: [c, d] \rightarrow \mathbb{R}, \quad y \mapsto \int_a^b f(x, y) dx$$

differenzierbar auf  $[c, d]$ , und es gilt

$$g'(y) = \int_a^b f_y(x, y) dx.$$

BEWEIS. Sei  $y \in [c, d]$  fix, und  $h_j$  eine beliebige Folge mit  $h_j \rightarrow 0$  für  $j \rightarrow \infty$  und  $y + h_j \in [c, d]$  für alle  $j$ . Dann ist die Folge der Differenzenquotienten für  $g$ ,

$$h_n(y) = \frac{g(y + h_n) - g(y)}{h_n} = \int_a^b \frac{f(x, y + h_n) - f(x, y)}{h_n} dx = \int_a^b f_n(x, y) dx$$

durch Funktionen  $f_n$  gegeben, die auf  $[a, b]$  gleichmässig gegen  $f_y(x, y)$  konvergieren; dies wollen wir nun nachweisen.

Nach dem Mittelwertsatz der Differentialrechnung können wir  $f_n(x, y) = f_y(x, \xi_n)$  mit einem  $\xi_n$  welches  $|y - \xi_n| \leq |h_n|$  erfüllt. Nun ist  $f_y$  nach Voraussetzung stetig auf dem kompakten Intervall, also gleichmässig stetig; zu gegebenem  $\varepsilon > 0$  können wir demnach ein  $\delta > 0$  wählen sodass  $|f_y(x, y) - f_y(x', y')| < \varepsilon$  wenn  $|x - x'| < \delta$  und  $|y - y'| < \delta$ . Ist nun  $N \in \mathbb{N}$  so gross gewählt, dass  $|h_n| < \delta$  für  $n \geq N$ , so ist also  $|f_y(x, y) - f_n(x, y)| = |f_y(x, y) - f_y(x, \xi_n)| < \varepsilon$ .  $\square$

SATZ 42. Sei  $f$  stetig auf  $[a, b]$ . Dann gibt es ein  $\xi \in [a, b]$  mit

$$f(\xi) = \frac{1}{b-a} \int_a^b f(x) dx.$$

BEWEIS. Da  $[a, b]$  kompakt ist, gibt es  $x_0, x_1 \in [a, b]$  mit  $f(x_0) \leq f(x) \leq f(x_1)$  für  $x \in [a, b]$ . Dann ist

$$(b-a)f(x_0) \leq \int_a^b f(x) dx \leq (b-a)f(x_1),$$

und da  $f$  stetig ist, gibt es ein  $\xi$  mit

$$(b-a)f(\xi) = \int_a^b f(x) dx.$$

$\square$

### 3. Das Lebesgue-Kriterium

Die Charakterisierung der Riemann-Integrierbarkeit in Theorem 40 sagt aus, dass die Punkte an denen  $f$  nicht stetig ist, d.h. jene Punkte  $x_0$  für welche ein  $\varepsilon_0 > 0$  existiert sodass für jede Umgebung  $U$  die Ungleichung  $\sup_{x \in U} f(x) - \inf_{x \in U} f(x) > \varepsilon_0$  erfüllt ist, nicht "allzu viele" sein können. Diese Beobachtung wird im Kriterium von Lebesgue explizit gemacht. Zunächst benötigen wir den Begriff der *Oszillation* einer Funktion.

**Definition 42.** Sei  $f: \mathbb{R} \supset X \rightarrow \mathbb{R}$  eine Funktion, und  $E \subset X$  eine Menge. Dann bezeichnen wir mit  $\omega(f, E) = \sup_{x \in E} f(x) - \inf_{y \in E} f(y) = \sup_{x, y \in E} |f(x) - f(y)|$  die Oszillation von  $f$  auf  $E$ .

Für einen Punkt  $x \in X$  ist

$$\omega(f, x) = \inf_{x \in U} \omega(f, X \cap U),$$

wobei das Infimum über Umgebungen  $U$  von  $x$  geht, die Oszillation von  $f$  im Punkt  $x$ .

LEMMA 78.  $f: \mathbb{R} \supset X \rightarrow \mathbb{R}$  ist stetig im Punkt  $x \in X$  genau dann, wenn  $\omega(f, x) = 0$ .

**Übungsaufgabe 48.** Beweise Lemma 78.

Die Definition von "nicht allzu viele" ist etwas umständlicher.

**Definition 43.** Eine Teilmenge  $N \subset \mathbb{R}$  heisst *Nullmenge*, wenn es für jedes  $\varepsilon > 0$  eine abzählbare Familie  $I_k$ ,  $k \in \mathbb{N}$ , von Intervallen mit linkem Endpunkt  $a_k$  und rechten Endpunkt  $b_k$  (egal welchen Typs) mit

$$\sum_{k \in \mathbb{N}} (b_k - a_k) < \varepsilon \text{ und } N \subset \bigcup_{k \in \mathbb{N}} I_k$$

gibt. Wir nennen  $b_k - a_k =: |I_k|$  die Länge (oder den Inhalt) des Intervalls  $I_k$ .

**Beispiel 27.** Wenn  $N = \{q_j : j \in \mathbb{N}\} \subset \mathbb{R}$  abzählbar ist, so ist  $N$  eine Nullmenge. Sei  $\varepsilon > 0$  gegeben. Wir definieren

$$I_k = \left( q_k - \frac{\varepsilon}{2^{k+1}}, q_k + \frac{\varepsilon}{2^{k+1}} \right).$$

Dann ist  $|I_k| = \frac{\varepsilon}{2^k}$ , also

$$\sum_{k \geq 1} |I_k| = \sum_{k \geq 1} \frac{\varepsilon}{2^k} = \varepsilon.$$

**Beispiel 28.** Es gibt auch überabzählbare Nullmengen, wie zum Beispiel die Cantormenge  $C \subset [0, 1]$ . Wir definieren dafür für ein Intervall  $I = [a, b]$  das "innere Drittel"  $D(I) = (\frac{b+2a}{3}, \frac{2b-a}{3})$ . Nun konstruieren wir induktiv eine absteigende Folge von Teilmengen  $C_j = \bigcup_{k=1}^{2^j} I_j^k$  von  $[0, 1]$ , wobei jedes  $C_j$  eine Vereinigung von abgeschlossenen Intervallen ist, durch

$$C_1 = [0, 1] \setminus D([0, 1]) = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right] =: I_1^1 \cup I_1^2,$$

und

$$C_{j+1} = \bigcup_{k=1}^{2^{j-1}} (I_j^k \setminus D(I_j^k)) = \bigcup_{k=1}^{2^j} I_{j+1}^k,$$

wo  $I_j^k \setminus D(I_j^k) = I_{j+1}^{2k-1} \cup I_{j+1}^{2k}$ . Die Menge  $C$  ist definiert durch  $C = \bigcap_j C_j$ .

Die Menge  $C$  ist überabzählbar, da sie alle Zahlen  $r \in [0, 1]$  enthält, die eine unendliche triadische Entwicklung

$$r = \sum_{j=1}^{\infty} \frac{r_j}{3^j}, \quad r_j \in \{0, 2\}$$

besitzen; diese Menge ist überabzählbar nach dem Cantor'schen Diagonalprinzip (sie ist gleichmächtig mit  $\mathbb{N}^{\mathbb{N}}$ ). Genauer gesagt besteht  $C$  aus allen reellen Zahlen

$$r = \sum_{j=1}^{\infty} \frac{r_j}{3^j}, \quad r_j \in \{0, 1, 2\},$$

wo falls  $r_j = 1$  für ein  $j$  gilt, dann  $r_k = 0$  für  $k > j$  gilt.

Die Länge jedes Teilintervalls von  $I_j^k$  ist  $(\frac{1}{3})^j$ , und es gibt davon  $2^j$  viele. D.h. die Gesamtlänge der in  $C_j$  vertretenen Intervalle ist  $(\frac{2}{3})^j$ , und  $C$  erweist sich damit als Nullmenge (da es durch jedes  $C_j$  überdeckt wird).

**Beispiel 29.** Wir erinnern an die Definition der Cantormenge  $C \subset [0, 1]$  in Beispiel 28. Diese ist als

$$C = \bigcap_{j \in \mathbb{N}} C_j$$

definiert, und jedes  $C_j$  ist durch

$$C_j = \bigcup_{k=1}^{2^j} I_j^k$$

gegeben, wo die Intervalle  $I_j^k$  insbesondere die folgende Bedingung erfüllt:

$$I_{j+1}^k \subset I_j^{\lfloor \frac{k+1}{2} \rfloor}.$$

Wir definieren nun induktiv eine Folge von stückweise affinen Funktionen  $f_n$  wie folgt:

$$f_1(x) = \begin{cases} \frac{3}{2}x & x \in I_1^1 \\ \frac{1}{2} & x \in [\frac{1}{3}, \frac{2}{3}] \\ \frac{1}{2} + \frac{3}{2}(x - \frac{2}{3}) & x \in I_1^2. \end{cases}$$

Für  $n \geq 1$  legen wir fest, dass  $f_{n+1} = f_n$  auf  $(C_n)^c$  ist. Auf jedem der Intervalle  $I_n^j = [\frac{a(j)}{3^n}, \frac{a(j)+1}{3^n}]$ , wo  $1 \leq j \leq 2^n$ , legen wir  $f_{n+1}$  wie folgt fest:

$$f_{n+1}|_{I_n^j}(x) = \begin{cases} \left(\frac{3}{2}\right)^{n+1} \left(x - \frac{3a(j)+1}{3^{n+1}}\right) + \frac{2a(j)+1}{2 \cdot 3^n} & x \in I_{n+1}^{2j-1} \\ \frac{2a(j)+1}{2 \cdot 3^n} & x \in D(I_n^j) \\ \left(\frac{3}{2}\right)^{n+1} \left(x - \frac{3a(j)+2}{3^{n+1}}\right) + \frac{2a(j)+1}{2 \cdot 3^n} & x \in I_{n+1}^{2j} \end{cases}$$

Es folgt, dass  $f_n(x) - f_{n+1}(x)$  sich gleichmäßig nach oben abschätzen lässt:

$$\|f_n - f_{n+1}\|_\infty = \sup_{x \in [0,1]} |f_n(x) - f_{n+1}(x)| \leq \frac{1}{2^n};$$

Damit erweist sich  $(f_n)_{n \in \mathbb{N}}$  als Cauchyfolge in  $C([0,1])$ , da für  $m \geq n$ :

$$\|f_n - f_m\|_\infty \leq \sum_{k=0}^{m-n-1} \|f_{n+k+1} - f_{n+k}\|_\infty \leq \sum_{k=0}^{m-n-1} \left(\frac{1}{2}\right)^{n+k} \leq \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^{n+k} = \left(\frac{1}{2}\right)^{n-1}.$$

Es gibt daher  $f = \lim_n f_n$ , und wir bezeichnen  $f \in C([0,1])$  als die Cantorfunktion oder die *Teufelstreppe*. Die Funktion  $f: [0,1] \rightarrow [0,1]$  ist stetig, monoton, und  $f$  ist konstant auf  $C^c$ , also auf dem Komplement einer Nullmenge.

**Definition 44.** Wir sagen, eine Eigenschaft  $P(x)$  ist *fast überall* erfüllt, wenn die Menge der Punkte, wo sie nicht erfüllt ist, eine Nullmenge ist.

SATZ 43. Sei  $f: [a,b] \rightarrow \mathbb{R}$ . Dann ist  $f$  integrierbar im Sinn von Riemann genau dann, wenn  $f$  beschränkt auf  $[a,b]$  und fast überall stetig ist.

Bevor wir uns dem Beweis von Theorem 43 widmen können, benötigen wir einige Vorbereitungen.

LEMMA 79. Sei  $N_k, k \in \mathbb{N}$ , eine abzählbare Familie von Nullmengen. Dann ist  $N = \cup_{k \in \mathbb{N}} N_k$  eine Nullmenge.

BEWEIS. Sei  $\varepsilon > 0$  beliebig. Für jedes  $k$  gibt es dann eine abzählbare Familie von Intervallen  $I_j^k = (a_j^k, b_j^k)$  mit

$$\bigcup_{j \in \mathbb{N}} I_j^k \supset N_k, \quad \sum_{j \in \mathbb{N}} (b_j^k - a_j^k) \leq \frac{\varepsilon}{2^k}.$$

Dann ist  $I_j^k, j, k \in \mathbb{N}$ , wieder eine abzählbare Familie von Intervallen, welche

$$\bigcup_{j,k \in \mathbb{N}} I_j^k \supset \bigcup_{k \in \mathbb{N}} N_k = N, \quad \sum_{j,k \in \mathbb{N}} (b_j^k - a_j^k) \leq \sum_{k \in \mathbb{N}} \frac{\varepsilon}{2^k} = \varepsilon$$

erfüllt. □

LEMMA 80. Sei  $f: \mathbb{R} \supset X \rightarrow \mathbb{R}$  eine Funktion, und  $\varepsilon > 0$ . Dann ist die Menge  $\{x \in X: \omega(f, x) \geq \varepsilon\}$  abgeschlossen.

BEWEIS. Wenn  $y$  im Abschluss von  $E = \{x \in X: \omega(f, x) \geq \varepsilon\}$  liegt, so gibt es für jede Umgebung  $U$  von  $y$  einen Punkt  $x \in U \cap E$ . Damit ist

$$\omega(f, y) = \inf_{y \in V} \omega(f, V \cap X) \geq \omega(f, U \cap X) \geq \inf_{x \in W} \omega(f, W \cap X) = \omega(f, x) \geq \varepsilon,$$

also  $y \in E$ , und  $E$  erweist sich als abgeschlossen. □

BEWEIS VON THEOREM 43. Sei zunächst  $f$  Riemann-integrierbar. Dann ist  $f$  beschränkt nach Übungsaufgabe 47. Wenn wir mit  $S$  die Menge der Unstetigkeitsstellen von  $f$  bezeichnen, so ist nach Lemma 78

$$S = \bigcup_{k \in \mathbb{N}} \left\{ x \in [a, b] : \omega(f, x) \geq \frac{1}{k} \right\},$$

und nach Lemma 79 genügt es zu zeigen, dass jede der Mengen

$$S_\alpha = \{x \in [a, b] : \omega(f, x) \geq \alpha\}$$

eine Nullmenge ist.

Sei also  $\varepsilon > 0$  beliebig. Dann existiert eine Unterteilung  $U = (x_0, \dots, x_n)$  von  $[a, b]$  sodass  $R_+(f, U) - R_-(f, U) < \varepsilon$ . Seien  $J_k$  jene Teilintervalle von  $U$ ,  $J_k = I_{j_k} = [x_{j_k}, x_{j_k+1}]$  für welche  $I_{j_k} \cap S_\alpha \neq \emptyset$ . Dann ist

$$\alpha \sum_k (x_{j_k+1} - x_{j_k}) \leq \sum_k \omega(f, [x_{j_k}, x_{j_k+1}]) (x_{j_k+1} - x_{j_k}) \leq R_+(f, U) - R_-(f, U) < \varepsilon.$$

Also gilt

$$S_\alpha \subset \bigcup_k J_k, \text{ und } \sum_k (x_{j_k+1} - x_{j_k}) \leq \frac{\varepsilon}{\alpha},$$

und nachdem  $\varepsilon > 0$  beliebig war, erweist sich  $S_\alpha$  als Nullmenge.

Sei nun andererseits  $S$  (wie oben definiert) eine Nullmenge, und  $\varepsilon > 0$  beliebig. Dann ist  $S_\varepsilon$  eine Nullmenge, welche überdies eine abgeschlossene Teilmenge von  $[a, b]$  und damit kompakt ist. Überdecken wir also  $S_\varepsilon$  mit Intervallen  $I_k = (a_k, b_k)$ , so genügen schon endlich viele dieser Intervalle; nachdem  $S_\varepsilon$  eine Nullmenge ist, können wir dementsprechend eine solche endliche Überdeckung mit

$$\sum_k (b_k - a_k) < \varepsilon$$

wählen. Das Komplement  $K$  von  $\cup I_k$  ist eine abgeschlossene Teilmenge von  $[a, b]$ , also kompakt. Nachdem jeder Punkt  $x \in K$  eine Umgebung besitzt, für welche die Oszillation von  $f$  auf dieser Umgebung kleiner als  $\varepsilon$  ist, können wir eine Unterteilung  $U = (x_0, \dots, x_n)$  von  $[a, b]$  finden, welche die  $a_k$  und  $b_k$  an aneinanderfolgenden Stellen enthält und wo für jedes  $x_j$  welches kein  $a_k$  ist  $\omega(f, [x_j, x_{j+1}]) < \varepsilon$  gilt.

Damit ist

$$R_+(f, U) - R_-(f, U) = \sum_{j=0}^{n-1} \omega(f, [x_j, x_{j+1}]) (x_{j+1} - x_j) \leq 2\varepsilon \sup_{x \in [a, b]} |f(x)| + (b - a)\varepsilon.$$

Nachdem  $\varepsilon > 0$  beliebig war, folgt die Riemann-Integrierbarkeit von  $f$  aus Theorem 40. □

#### 4. Hauptsätze und Folgerungen

Der erste Hauptsatz sagt uns, dass Integration gefolgt von Differentiation die Identität darstellt.

SATZ 44. Sei  $f$  stetig auf  $[a, b]$ . Dann ist die Funktion

$$F: [a, b] \rightarrow \mathbb{R}, \quad F(x) = \int_a^x f(t) dt$$

auf  $[a, b]$  differenzierbar, mit  $F' = f$ .

BEWEIS. Sei  $x \in [a, b]$  fix. Für jedes  $h$  gibt es nach Theorem 42 ein  $\xi_h$  mit  $|x - \xi_h| < h$  sodass

$$f(\xi_h) = \frac{1}{h} \int_x^{x+h} f(t) dt$$

gilt. Somit ist

$$\lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} f(\xi_h) = f(x).$$

□

**Bemerkung 13.** Die lineare Abbildung  $D: C^1([a, b]) \rightarrow C([a, b])$  ist stetig, wie wir schon gesehen haben (dies war der Inhalt von Lemma 68). Der Kern von  $D$  besteht aus all jenen Funktionen, die konstant auf  $[a, b]$  sind; der Faktorraum

$$C^1([0, 1]) /_{\ker D}$$

kann mit jenen einmal stetig differenzierbaren Funktionen auf  $[a, b]$  identifiziert werden, welche in einem Punkt, zum Beispiel  $a$ , des Intervalls verschwinden:

$$C^1([0, 1]) /_{\ker D} = \{\varphi \in C^1([a, b]): \varphi(a) = 0\} =: C^1([a, b])_a.$$

Die Abbildung  $I: C([a, b]) \rightarrow C^1([a, b])_a$ , welche durch

$$I(f)(x) = \int_a^x f(t) dt$$

definiert ist, ist nach Theorem 44 eine linksinverse Abbildung zu  $D$ , d.h.  $D \circ I = \text{id}_{C([a, b])}$ .

Um die Funktion

$$F(x) = \int_a^x f(t) dt$$

zu definieren, genügt es vorauszusetzen, dass  $f$  Riemann-integrierbar ist. Auch für diesen Fall gilt eine Version des ersten Hauptsatzes:

SATZ 45. Sei  $f \in R([a, b])$ . Dann ist die Funktion

$$F: [a, b] \rightarrow \mathbb{R}, \quad F(x) = \int_a^x f(t) dt$$

auf  $[a, b]$  definiert. Wenn  $f$  stetig an einem Punkt  $x_0 \in [a, b]$  ist, so gilt  $F'(x_0) = f(x_0)$ .

BEWEIS. Für den Beweis rechnen wir:

$$\begin{aligned} |F(x_0 + h) - F(x_0) - hf(x_0)| &= \left| \int_a^{x_0+h} f(t) dt - \int_a^{x_0} f(t) dt - hf(x_0) \right| \\ &= \left| \int_{x_0}^{x_0+h} f(t) dt - \int_{x_0}^{x_0+h} f(x_0) dt \right| \\ &= \left| \int_{x_0}^{x_0+h} f(t) - f(x_0) dt \right| \\ &\leq \int_{x_0}^{x_0+h} |f(t) - f(x_0)| dt \\ &\leq |h| \sup_{|t-x_0| \leq |h|} |f(t) - f(x_0)|. \end{aligned}$$

Also ist

$$\lim_{h \rightarrow 0} \frac{|F(x_0 + h) - F(x_0) - hf(x_0)|}{|h|} = \lim_{h \rightarrow 0} \sup_{|t-x_0| \leq |h|} |f(t) - f(x_0)| = 0,$$

da  $f$  stetig im Punkt  $x_0$  ist. □

Wir wollen uns nun der Frage widmen, inwieweit Integration eine echte inverse Operation zur Differentiation darstellt. Diese wird durch den *zweiten Hauptsatz* beantwortet.

SATZ 46. Sei  $F: [a, b] \rightarrow \mathbb{R}$  stetig, und differenzierbar auf  $(a, b)$ , mit  $F' = f$ . Wenn  $f$  Riemann-integrierbar ist, so gilt

$$\int_a^b f(x) dx = F(b) - F(a).$$

BEWEIS. Wir wählen die Unterteilungen  $U^n = (x_0, x_1, \dots, x_n)$  von  $[a, b]$  welche durch  $x_j = a + \frac{j(b-a)}{n}$  gegeben ist. Nach dem Mittelwertsatz der Differentialrechnung gibt es  $\xi_j^n$  mit  $x_j \leq \xi_j^n \leq x_{j+1}$  welche  $F(x_{j+1}) - F(x_j) = f(\xi_j^n)(x_{j+1} - x_j)$  erfüllen. Es folgt

$$\begin{aligned} \int_a^b f(x) dx &= \lim_{n \rightarrow \infty} R(f, U^n, \xi^n) \\ &= \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} f(\xi_j^n)(x_{j+1} - x_j) \\ &= \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} (F(x_{j+1}) - F(x_j)) \\ &= F(b) - F(a). \end{aligned}$$

□

**Bemerkung 14.** Der zweite Hauptsatz sagt uns insbesondere, dass für  $F \in C^1([a, b])$  und für jedes  $x \in [a, b]$  die Gleichung

$$F(x) - F(a) = \int_a^x F'(t) dt$$

gilt, d.h. mit der Notation von Bemerkung 13, dass  $I \circ D = \text{id}_{C^1([a, b])_a}$  gilt.

Die Hauptsätze zeigen uns, in welchem Sinn die Integration und die Differentiation inverse Operationen sind: Will man eine stetige Funktion  $f$  integrieren, so verwendet man eine *Stammfunktion*  $F$ , d.h. eine Funktion  $F$  mit der Eigenschaft, dass ihre Ableitung  $F'$  mit der gegebenen Funktion  $f$  übereinstimmt. Stammfunktionen sind immer nur bis auf Konstanten bestimmt; man bezeichnet eine Stammfunktion  $F$  auch gerne als *unbestimmtes Integral* von  $f$  und schreibt

$$F(x) = \int f(x) dx.$$

wie schon gesagt, ist diese Schreibweise nicht exakt, da mit  $F(x)$  auch  $F(x) + C$  für jedes  $C \in \mathbb{R}$  auf der linken Seite dieser Gleichung stehen kann. Weiters gibt uns der erste Hauptsatz Theorem 44 eine “grosse” Klasse von Funktionen (d.h. stetige), welche Stammfunktion besitzen.

Die Rechenregeln der Differentialrechnung haben ihre Entsprechung in Regeln für die unbestimmte Integration:

LEMMA 81. *Wenn  $f, g$  Stammfunktionen besitzen, und  $\lambda \in \mathbb{R}$  beliebig ist, dann besitzt auch die Funktion  $\lambda f + g$  eine Stammfunktion, welche durch*

$$\int \lambda f(x) + g(x) dx = \lambda \int f(x) dx + \int g(x) dx$$

gegeben ist.

*Wenn  $f$  eine Stammfunktion  $F$  besitzt, und  $Fg'$  eine Stammfunktion besitzt, so besitzt auch  $fg$  eine Stammfunktion, welche durch*

$$\int f(x)g(x) dx = F(x)g(x) - \int F(x)g'(x) dx$$

gegeben ist (partielle Integration).

*Besitzt  $f(y)$  eine Stammfunktion  $F(y)$ , und ist  $\varphi(x)$  differenzierbar, so besitzt auch  $f'(\varphi(x))\varphi'(x)$  eine Stammfunktion, welche durch*

$$\int f'(\varphi(x))\varphi'(x) dx = F(\varphi(x))$$

gegeben ist (Substitutionsregel).

Mit Hilfe der Regeln in Lemma 81 kann man viele “elementare” Funktionen integrieren. Da wir an dieser Stelle noch nicht auf die noch einzuführenden elementaren transzendenten Funktionen vorgreifen wollen, wird eine genauere Behandlung von Techniken der Integration erst später erfolgen. Wir bemerken hier, dass



Lemma 81 mit Absicht etwas “unscharf” formuliert ist; sein Inhalt ist auch keine *analytische* Aussage über das Integral, sondern formale Rechenregeln, welche aus den algebraischen Regeln für die Ableitung folgen. Es macht deswegen keinen Sinn, sich im Rahmen dieses Lemmas die Fragen wo und wie sehr die entsprechenden Funktionen differenzierbar sind zu stellen.

**Übungsaufgabe 49.** Beweise die Aussagen von Lemma 81.

Anders ist dies, wenn wir die entsprechenden Aussagen für *bestimmte* (d.h. Riemann) Integrale beweisen wollen. Die Summenformel haben wir schon bewiesen; wie sieht es mit den entsprechenden Aussagen für die partielle Integration und die Substitutionsregel aus?

LEMMA 82. *Seien  $F$  und  $g$  auf  $[a, b]$  differenzierbar, und mit  $F' = f$  sowohl  $fg$  als auch  $Fg'$  Riemann-integrierbar auf  $[a, b]$ . Dann gilt*

$$\int_a^b f(x)g(x) dx = F(b)g(b) - F(a)g(a) - \int_a^b F(x)g'(x) dx.$$

BEWEIS. Da  $F(x)g(x)$  differenzierbar ist, gibt es nach dem Mittelwertsatz für beliebige  $x, y \in [a, b]$  ein  $\xi(x, y)$  mit  $x < \xi < y$  sodass  $F(y)g(y) - F(x)g(x) = (f(\xi)g(\xi) + F(\xi)g'(\xi))(y - x)$ . Für eine beliebige Unterteilung  $U = (x_0, \dots, x_n)$  können wir demnach mit  $\xi_j = \xi(x_{j-1}, x_j)$

$$\begin{aligned} F(b)g(b) - F(a)g(a) &= \sum_{j=1}^n F(x_j)g(x_j) - F(x_{j-1})g(x_{j-1}) \\ &= \sum_{j=1}^n (f(\xi_j)g(\xi_j) + F(\xi_j)g'(\xi_j))(x_j - x_{j-1}) \\ &= R(fg + Fg', U, \xi) \end{aligned}$$

schreiben. Da die beteiligten Funktionen alle nach Voraussetzung Riemann-integrierbar sind, folgt die Behauptung, wenn wir  $\delta(U) \rightarrow 0$  streben lassen.  $\square$

LEMMA 83. *Sei  $\varphi: [a, b] \rightarrow [\varphi(a), \varphi(b)] = [c, d]$  eine strikt monoton wachsende (bzw. monoton fallende) stetig differenzierbare Funktion. Dann gilt*

$$\int_c^d f(y) dy = \int_a^b f(\varphi(x))\varphi'(x) dx$$

für jede Riemann-integrierbare Funktion  $f$  auf  $[c, d]$ .

BEWEIS. Nach Theorem 43 ist  $f(\varphi(x))\varphi'(x)$  Riemann-integrierbar. Sei nun  $V = (y_0, \dots, y_n)$  eine Unterteilung von  $[c, d]$ , und  $y_j = \varphi(x_j)$ . Dann ist  $\varphi^{-1}(V) = (x_0, \dots, x_n)$  eine Unterteilung von  $[a, b]$ ; wir wählen für jedes  $j$  ein  $\xi_j$  mit  $x_{j-1} \leq \xi_j \leq x_j$  mit  $\varphi(x_{j+1}) - \varphi(x_j) = \varphi'(\xi_j)$ , und schreiben  $\varphi(\xi) = (\varphi(\xi_1), \dots, \varphi(\xi_n))$  für diesen Zwischenvektor gilt

$$R(f, U, \varphi(\xi)) = R((f \circ \varphi)\varphi', \varphi^{-1}(U), \xi);$$

wegen der gleichmässigen Stetigkeit von  $\varphi$  auf  $[a, b]$  gilt  $\delta(\varphi(U)) \rightarrow 0$  wenn  $\delta(U) \rightarrow 0$ , und der Übergang zum Grenzwert liefert die Behauptung.  $\square$

## 5. Der Logarithmus und die Exponentialfunktion

Wir betrachten die folgende auf  $(0, \infty)$  definierte Stammfunktion von  $x \mapsto x^{-1}$ ,

$$L(x) = \int_1^x \frac{dt}{t}.$$

Die so definierte Funktion  $L$  hat die folgenden Eigenschaften:

- (L1)  $L$  ist glatt auf  $(0, \infty)$
- (L2)  $L$  ist strikt monoton steigend und strikt konkav
- (L3)  $\lim_{x \rightarrow 0} \log(x) = -\infty$ , und  $\lim_{x \rightarrow +\infty} \log(x) = +\infty$
- (L4)  $L(xy) = L(x) + L(y)$  für  $x, y > 0$ .

(L1) und (L2) folgen direkt aus  $L'(x) = \frac{1}{x} \in C^\infty(0, \infty)$ , sowie aus  $L'(x) > 0$  und  $L''(x) < 0$  für  $x > 0$ . (L3) kann wie folgt gezeigt werden:

$$\begin{aligned} \int_1^x \frac{dt}{t} &\geq \int_1^{\lfloor x \rfloor} \frac{dt}{t} \\ &= \sum_{j=1}^{\lfloor x \rfloor - 1} \int_j^{j+1} \frac{dt}{t} \\ &\geq \sum_{j=1}^{\lfloor x \rfloor} \frac{1}{j+1} \rightarrow +\infty \quad (x \rightarrow +\infty). \end{aligned}$$

Die Abschätzung für  $x \rightarrow 0$  kann ganz analog erfolgen, oder man schätzt wie folgt ab:

$$\begin{aligned} \int_1^x \frac{dt}{t} &\leq \int_{\sqrt{x}}^x \frac{dt}{t} \\ &\leq \frac{x - \sqrt{x}}{x} \\ &= 1 - \frac{1}{\sqrt{x}} \rightarrow -\infty \quad (x \rightarrow 0). \end{aligned}$$

Zum Beweis von (L4) verwenden wir Lemma 83:

$$\begin{aligned} L(xy) &= \int_1^{xy} \frac{dt}{t} && t = xs, \quad dt = xds \\ &= \int_{x^{-1}}^y \frac{xds}{xs} \\ &= \int_{x^{-1}}^1 \frac{du}{u} + \int_1^y \frac{ds}{s} && u = x^{-1}t, \quad du = x^{-1}dt \\ &= \int_1^x \frac{dt}{t} + \int_1^y \frac{ds}{s} \\ &= L(x) + L(y). \end{aligned}$$

Wir bezeichnen  $L(x)$  als den *natürlichen Logarithmus* der Zahl  $x > 0$ , und schreiben  $L(x) =: \log(x) =: \ln x$ .

Wenn  $\tilde{L}: \mathbb{R}_+ \rightarrow \mathbb{R}$  eine Funktion ist, welche die *Funktionalgleichung*

$$\tilde{L}(xy) = \tilde{L}(x) + \tilde{L}(y)$$

für  $x, y > 0$  erfüllt, so ist

$$\tilde{L}(x+h) - \tilde{L}(x) = \tilde{L}\left(\frac{x+h}{x}\right).$$

Wenn  $\tilde{L}$  differenzierbar an der Stelle  $x$  ist, so ist  $\tilde{L}$  also differenzierbar an der Stelle 1; d.h.  $\tilde{L}$  ist entweder nirgends oder überall differenzierbar (der zweite Fall tritt z.Bsp. ein, wenn wir wissen, dass  $\tilde{L}$  strikt konkav ist). Im zweiten Fall können wir die obige Gleichung nach  $h$  ableiten, und an der Stelle  $h = 0$  auswerten, und erhalten:

$$\tilde{L}'(x) = \frac{\tilde{L}'(1)}{x}.$$

Wegen  $\tilde{L}(1) = \tilde{L}(1 \cdot 1) = 2\tilde{L}(1)$  ist  $\tilde{L}(1) = 0$ , und insgesamt folgt, dass

$$\tilde{L}(x) = \tilde{L}(0) + \tilde{L}'(1) \log(x) = \tilde{L}'(1) \log(x).$$

Wir fassen kurz zusammen:

**PROPOSITION 6.** *Sei  $\tilde{L}: \mathbb{R}_+ \rightarrow \mathbb{R}$  eine Funktion, welche (L1) und (L4) erfüllt. Dann gibt es ein  $c \in \mathbb{R}$  sodass  $\tilde{L}(x) = c \log(x)$ .*

Wegen (L1), (L2), und (L3) besitzt  $L$  nach Korollar 5 eine Umkehrfunktion  $E: \mathbb{R} \rightarrow \mathbb{R}_+$ , welche die folgenden Eigenschaften besitzt:

(E1)  $E \in C^\infty(\mathbb{R})$ ,

(E2)  $E$  ist strikt monoton steigend und strikt konkav

(E3)  $\lim_{x \rightarrow -\infty} E(x) = 0$ , und  $\lim_{x \rightarrow \infty} E(x) = \infty$ .

(E4)  $E(x + y) = E(x) + E(y)$  für  $x, y \in \mathbb{R}$ .

Wir müssen davon (die anderen Eigenschaften folgen direkt aus jenen des Logarithmus und da  $E$  seine Umkehrfunktion ist) nur noch (E4) nachweisen. Seien also  $x = L(s)$  und  $y = L(t)$  (oder gleichwertig:  $s = E(x)$  und  $t = E(y)$ ). Dann ist

$$E(x + y) = E(L(s) + L(t)) = E(L(st)) = st = E(x)E(y).$$

Wir bezeichnen die Funktion  $E$  als (reelle) Exponentialfunktion und schreiben  $E(x) =: \exp(x) =: e^x$ . Die Funktionalgleichung  $\exp(x + y) = \exp(x)\exp(y)$  hat auch eine infinitesimale Form:

$$\lim_{h \rightarrow 0} \frac{\exp(x + h) - \exp(x)}{h} = \lim_{h \rightarrow 0} \exp(x) \frac{\exp(h) - \exp(0)}{h} = \exp(x) \exp'(0) = \exp(x),$$

das heisst, die Exponentialfunktion erfüllt die Differentialgleichung

$$\exp' = \exp.$$

In einem gewissen Sinn ist die Exponentialfunktion die eindeutige Lösung dieser Differentialgleichung, und damit nach unserer Rechnung oben auch die eindeutige Lösung der Funktionalgleichung (E4):

PROPOSITION 7. Sei  $\lambda \in \mathbb{R}$ , und  $y: \mathbb{R} \rightarrow \mathbb{R}$  eine Funktion, welche  $y'(x) = \lambda y(x)$  für  $x \in \mathbb{R}$  erfüllt. Dann gilt

$$y(x) = y(0) \exp(\lambda x).$$

BEWEIS. Wir berechnen

$$\frac{d}{dx} \frac{y(x)}{\exp(\lambda x)} = \frac{\exp(\lambda x)y'(x) - \lambda \exp(\lambda x)y(x)}{\exp(2\lambda x)} = \frac{\exp(\lambda x)\lambda y(x) - \lambda \exp(\lambda x)y(x)}{\exp(2\lambda x)} = 0.$$

Also ist  $y(x) = C \exp(\lambda x)$  für ein  $C \in \mathbb{R}$ . Auswertung für  $x = 0$  liefert  $C = y(0)$  wie behauptet.  $\square$

Wir können nun auch für  $x > 0$  und  $\alpha \in \mathbb{R}$  die allgemeine Potenz

$$x^\alpha := \exp(\alpha \ln x)$$

definieren. Die Funktion  $P_\alpha: \mathbb{R}_+ \rightarrow \mathbb{R}$ ,  $P_\alpha(x) = x^\alpha$ , bzw.  $P: \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $P(x, \alpha) = P_\alpha(x)$ , heisst (allgemeine) Potenzfunktion. Offensichtlich ist  $P_\alpha$  auch wieder glatt, und  $P$  stetig. Für  $\alpha = \frac{p}{q} \in \mathbb{Q}$ , wo  $q \in \mathbb{N}$ , stimmt dies mit der früher in Lemma 17 gegebenen Definition überein, da die Funktionalgleichung der Exponentialfunktion in diesem Fall

$$\exp\left(\frac{p}{q}x\right)^q = \exp(px) = \exp(x)^p$$

ergibt, also  $\exp(\frac{p}{q}x) = y$  die (eindeutige) Lösung der Gleichung  $y^q = x^p$  ist. Die Rechenregeln für Potenzen, i.e.

$$x^\alpha x^\beta = x^{\alpha+\beta}, \quad (x^\alpha)^\beta = x^{\alpha\beta}, \quad x^\alpha y^\alpha = (xy)^\alpha,$$

gelten unverändert für allgemeine Potenzen. Zum Beweis zieht man entweder die Funktionalgleichung der Exponentialfunktion, so ergibt sich etwa

$$x^\alpha x^\beta = \exp(\alpha \ln x) \exp(\beta \ln x) = \exp((\alpha + \beta) \ln x) = x^{\alpha+\beta},$$

oder aber die bekannten Rechenregeln für rationales  $\alpha$  und die schon beobachtete Stetigkeit von  $P$  heran. Auch Potenzfunktionen kann man mit Hilfe von einer Differentialgleichung charakterisieren:

PROPOSITION 8. Sei  $\alpha \in \mathbb{R}$ , und  $y: \mathbb{R}_+ \rightarrow \mathbb{R}$  eine Funktion, welche  $xy'(x) = \alpha y(x)$  erfüllt. Dann gibt es ein  $C \in \mathbb{R}$  sodass  $y(x) = Cx^\alpha$  gilt.

BEWEIS. Ähnlich wie oben in Beweis von Proposition 7 berechnen wir

$$\frac{d}{dx} \frac{y(x)}{x^\alpha} = \frac{x^\alpha y'(x) - \alpha x^{\alpha-1} y(x)}{x^{2\alpha}} = \frac{x^\alpha \alpha \frac{y(x)}{x} - \alpha x^{\alpha-1} y(x)}{x^{2\alpha}} = \frac{x^{\alpha-1} y(x) - \alpha x^{\alpha-1} y(x)}{x^{2\alpha}} = 0,$$

woraus sich wie oben die Behauptung ergibt.  $\square$

## 6. Das Integralrestglied

Wir beginnen zunächst mit einer wichtigen Eigenschaft von stetig differenzierbaren Funktionen an Nullstellen.

**KOROLLAR 14.** *Sei  $f: [x_0 - \varepsilon, x_0 + \varepsilon] \rightarrow \mathbb{R}$  stetig differenzierbar. Dann gibt es eine stetige Funktion  $g: [-\varepsilon, \varepsilon] \rightarrow \mathbb{R}$  sodass  $f(x) - f(x_0) = (x - x_0)g(x)$ .*

**BEWEIS.** Wir betrachten die Funktion  $\varphi(t) = f(x_0 + t(x - x_0))$ . Dann ist  $\varphi(0) = f(x_0)$ ,  $\varphi(1) = f(x)$ , und damit

$$f(x) - f(x_0) = \varphi(1) - \varphi(0) = \int_0^1 \varphi'(t) dt = (x - x_0) \int_0^1 f'(x_0 + t(x - x_0)) dt.$$

Die Funktion

$$g(x) = \int_0^1 f'(x_0 + t(x - x_0)) dt$$

ist stetig nach Korollar 12. □

Die Idee des Beweis des vorangehenden Korollars lässt sich mit Hilfe von Induktion und partieller Integration dazu verwenden, um das Restglied im Taylor'schen Lehrsatz in einer geschlossenen, berechenbaren Form anzugeben:

$$\begin{aligned} f(x) &= f(x_0) + (f(x) - f(x_0)) \\ &= f(x_0) + \int_{x_0}^x f'(t) dt \\ &= f(x_0) + \left( -(x-t)f'(t) \Big|_{t=x_0}^{t=x} + \int_{x_0}^x (x-t)f''(t) dt \right) \\ &= f(x_0) + (x-x_0)f'(x_0) + \int_{x_0}^x (x-t)f''(t) dt \\ &= f(x_0) + (x-x_0)f'(x_0) + \left( -\frac{(t-x)^2}{2} f''(t) \Big|_{t=x_0}^{t=x} + \int_{x_0}^x \frac{(x-t)^2}{2} f'''(t) dt \right) \\ &= f(x_0) + (x-x_0)f'(x_0) + \frac{(x-x_0)^2}{2} f''(x_0) + \int_{x_0}^x \frac{(x-t)^2}{2} f'''(t) dt \\ &= \dots \\ &= (T_{x_0}^k f)(x) + \int_{x_0}^x \frac{(x-t)^k}{k!} f^{(k+1)}(t) dt \\ &= (T_{x_0}^k f)(x) + \int_{x_0}^x \frac{(x-t)^k}{k!} f^{(k+1)}(t) dt \\ &= (T_{x_0}^k f)(x) + \left( -\frac{(x-t)^{k+1}}{(k+1)!} f^{(k+1)}(t) \Big|_{t=x_0}^{t=x} + \int_{x_0}^x \frac{(x-t)^{k+1}}{(k+1)!} f^{(k+2)}(t) dt \right) \\ &= (T_{x_0}^{k+1} f)(x) + \int_{x_0}^x \frac{(x-t)^{k+1}}{(k+1)!} f^{(k+2)}(t) dt \end{aligned}$$

Wir fassen zusammen:

**SATZ 47.** *Sei  $f \in C^k([x_0, x]) \cap D^{k+1}((x_0, x))$ , mit  $f^{(k+1)} \in R([x_0, x])$ . Dann gilt*

$$f(x) = (T_{x_0}^k f)(x) + \int_{x_0}^x \frac{(x-t)^k}{k!} f^{(k+1)}(t) dt.$$

**Beispiel 30.** Für die Exponentialfunktion  $\exp$  gilt nach Theorem 47, dass für beliebiges  $k \in \mathbb{N}$  die Gleichung

$$\exp(x) = \sum_{j=0}^k \frac{x^j}{j!} + \int_0^x \frac{(x-t)^k}{k!} \exp(t) dt$$

gilt. Sei  $r > 0$ , und  $|x| \leq r$ , dann gilt

$$\begin{aligned} \left| \int_0^x \frac{(x-t)^k}{k!} \exp(t) dt \right| &\leq \int_0^{|x|} \frac{r^k}{k!} \exp r dt \\ &\leq \frac{r^{k+1}}{k!} \exp(r) \rightarrow 0 \quad (k \rightarrow \infty). \end{aligned}$$

Wir folgern also, dass

$$\exp(x) = \sum_{j=0}^{\infty} \frac{x^j}{j!}$$

gleichmässig auf kompakten Intervallen der Form  $[-r, r]$  - und damit auch gleichmässig auf beliebigen kompakten Teilmengen von  $\mathbb{R}$  - gilt.

## 7. Einige einfache Integralungleichungen

Um Integrale abzuschätzen, gibt es viele verschiedene Methoden. Wir geben hier nur kurz einige wichtige Ungleichungen an. Der Anfang ist die "Dreiecksungleichung":

LEMMA 84. Wenn  $f \in R([a, b])$  ist, so ist auch  $|f| \in R([a, b])$ , und es gilt

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

Insbesondere ist

$$\left| \int_a^b f(x) dx \right| \leq (b-a) \sup_{x \in [a, b]} |f(x)|.$$

Weitere Ungleichungen werden wir unter Ausnutzung von Konvexitätseigenschaften bekannter Funktionen erhalten. Wir erinnern zunächst an die Definition einer *konvexen Funktion* auf einem Intervall  $I \subset \mathbb{R}$ :

Die Funktion  $\varphi: I \rightarrow \mathbb{R}$  ist (strikt) konvex, wenn für beliebige  $x_1 \neq x_2 \in I$  und  $\lambda_1, \lambda_2 \in (0, 1)$  mit  $\lambda_1 + \lambda_2 = 1$  die Ungleichung

$$\varphi(\lambda_1 x_1 + \lambda_2 x_2) < \lambda_1 \varphi(x_1) + \lambda_2 \varphi(x_2)$$

erfüllt ist.

**Beispiel 31.** Seien  $a, b \in \mathbb{R}_+$ . Dann gilt für beliebige  $p, q \in \mathbb{R}_+$  mit  $1/p + 1/q = 1$ , dass

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Da  $a = e^{\frac{x}{p}}$  und  $b = e^{\frac{y}{q}}$  für gewisse  $x, y \in \mathbb{R}$ , und die Exponentialfunktion konvex ist, haben wir

$$ab = \exp\left(\frac{1}{p}x + \frac{1}{q}y\right) \leq \frac{1}{p}e^x + \frac{1}{q}e^y = \frac{a^p}{p} + \frac{b^q}{q}.$$

Unsere erste Beobachtung ist die folgende Verallgemeinerung auf beliebige Konvexkombinationen:

LEMMA 85 (Jensen's Ungleichung). Sei  $\varphi: I \rightarrow \mathbb{R}$  (strikt) konvex (bzw. konkav),  $x_1, \dots, x_n \in I$  paarweise verschieden, und  $\lambda_1, \dots, \lambda_n \in [0, 1]$  mit  $\lambda_1 + \dots + \lambda_n = 1$  gegeben. Dann gilt

$$\varphi\left(\sum_{j=1}^n \lambda_j x_j\right) \leq \sum_{j=1}^n \lambda_j \varphi(x_j),$$

bzw. im Falle einer konkaven Funktion  $\varphi$

$$\varphi\left(\sum_{j=1}^n \lambda_j x_j\right) \geq \sum_{j=1}^n \lambda_j \varphi(x_j),$$

und im Falle einer strikt konvexen (konkaven) Funktion  $\varphi$  gilt Gleichheit genau dann, wenn es ein  $j$ ,  $1 \leq j \leq n$  mit  $\lambda_j = 1$  gibt.

BEWEIS. Wir beweisen die Behauptung mit Hilfe von Induktion nach  $n$ , und beschränken uns im Beweis auf eine (strikt) konvexe Funktion  $\varphi$ . Für  $n = 2$  ist die Behauptung gerade die Definition einer (strikt) konvexen Funktion. Gelte also die Behauptung des Lemmas für ein beliebiges (aber fixes)  $n \in \mathbb{N}$ . Dann ist für beliebige  $x_1, \dots, x_{n+1}$  und  $\lambda_1, \dots, \lambda_{n+1}$  die Behauptung klar, wenn  $\lambda_{n+1} = 1$  ist. Ist  $\lambda_{n+1} < 1$ , so rechnen wir wie folgt:

$$\begin{aligned} \varphi\left(\sum_{j=1}^n \lambda_j x_j + \lambda_{n+1} x_{n+1}\right) &= \varphi\left((1 - \lambda_{n+1}) \sum_{j=1}^n \frac{\lambda_j}{1 - \lambda_{n+1}} x_j + \lambda_{n+1} x_{n+1}\right) \\ &\leq (1 - \lambda_{n+1}) \varphi\left(\sum_{j=1}^n \frac{\lambda_j}{1 - \lambda_{n+1}} x_j\right) + \lambda_{n+1} \varphi(x_{n+1}) \\ &\leq (1 - \lambda_{n+1}) \sum_{j=1}^n \frac{\lambda_j}{1 - \lambda_{n+1}} \varphi(x_j) + \lambda_{n+1} \varphi(x_{n+1}) \\ &= \sum_{j=1}^{n+1} \lambda_j \varphi(x_j); \end{aligned}$$

die Behauptung gilt also auch für  $n + 1$  paarweise verschiedene Punkte  $x_1, \dots, x_{n+1} \in I$ . □

**Beispiel 32.** Für beliebige  $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$  mit  $\sum_j \lambda_j = 1$  und  $a_1, \dots, a_n \in \mathbb{R}_+$  gilt

$$\prod_{j=1}^n a_j^{\lambda_j} \leq \sum_{j=1}^n \lambda_j a_j.$$

SATZ 48 (Jensen's Ungleichung). Sei  $f \in R([0, 1])$ ,  $I \supset f([0, 1])$  ein Intervall, und  $\varphi: I \rightarrow \mathbb{R}$  konvex (bzw. konkav). Dann gilt

$$\varphi\left(\int_0^1 f(x) dx\right) \leq \int_0^1 (\varphi \circ f)(x) dx,$$

bzw.

$$\varphi\left(\int_0^1 f(x) dx\right) \geq \int_0^1 (\varphi \circ f)(x) dx$$

wenn  $\varphi$  konkav ist.

BEWEIS. Da  $\varphi$  als konvexe Funktion stetig ist, ist  $\varphi \circ f$  auch Riemann-integrierbar. Für jede Riemannsumme gilt nach Lemma 85

$$\varphi(R(f, U, \xi)) = \varphi\left(\sum_{j=1}^n f(\xi_j)(x_j - x_{j-1})\right) \leq \sum_{j=1}^n \varphi(f(\xi_j))(x_j - x_{j-1}) = R(\varphi \circ f, U, \xi).$$

Indem wir  $\delta(U) \rightarrow 0$  gehen lassen, folgt die Behauptung. □

LEMMA 86 (Hölderungleichung). Seien  $a_j, b_j \in \mathbb{R}_+$  für  $j = 1, \dots, n$ , und  $p, q \in (1, \infty)$  mit  $\frac{1}{p} + \frac{1}{q} = 1$  gegeben. Dann ist

$$\sum_{j=1}^n a_j b_j \leq \left(\sum_{j=1}^n a_j^p\right)^{\frac{1}{p}} \left(\sum_{j=1}^n b_j^q\right)^{\frac{1}{q}}.$$

Gleichheit gilt genau dann, wenn es ein  $c > 0$  gibt, sodass  $a_j^p = c b_j^q$  für  $j = 1, \dots, n$ .

BEWEIS. Nach Jensen's Ungleichung, angewendet auf die konvexe Funktion  $x \mapsto x^p$ , ist

$$\begin{aligned}
 \left( \frac{1}{\sum_{k=1}^n b_k^q} \sum_{j=1}^n a_j b_j \right)^p &= \left( \sum_{j=1}^n \frac{b_j^q}{\sum_{k=1}^n b_k^q} \frac{a_j b_j}{b_j^q} \right)^p \\
 &\leq \sum_{j=1}^n \frac{b_j^q}{\sum_{k=1}^n b_k^q} \left( \frac{a_j b_j}{b_j^q} \right)^p \\
 &= \frac{1}{\sum_{k=1}^n b_k^q} \sum_{j=1}^n b_j^q \left( \frac{a_j b_j}{b_j^q} \right)^p \\
 &= \frac{1}{\sum_{k=1}^n b_k^q} \sum_{j=1}^n b_j^{q+p-pq} a_j^p \\
 &= \frac{1}{\sum_{k=1}^n b_k^q} \sum_{j=1}^n a_j^p.
 \end{aligned}$$

Nach Ziehen der  $p$ -ten Wurzel ergibt sich die gewünschte Ungleichung. Da alle Gewichte positiv sind, ist Gleichheit nur möglich, wenn für jedes  $j = 1, \dots, n$  die Gleichheit

$$a_1 b_1^{-\frac{q}{p}} = \frac{a_1 b_1}{b_1^q} = \frac{a_j b_j}{b_j^q} = a_j b_j^{-\frac{q}{p}}$$

gilt, also  $\frac{a_j^p}{b_j^q} = c > 0$  konstant ist. □

Ähnlich wie die Jensen'sche Ungleichung folgt die Integralform der Hölderungleichung durch Vergleich der Riemannsummen.

SATZ 49 (Hölderungleichung). Seien  $f, g \in R([a, b])$ , und  $p, q \in (1, \infty)$  mit  $\frac{1}{p} + \frac{1}{q} = 1$ . Dann gilt

$$\left| \int_a^b f(x)g(x) dx \right| \leq \left( \int_a^b |f(x)|^p dx \right)^{\frac{1}{p}} \left( \int_a^b |g(x)|^q dx \right)^{\frac{1}{q}}$$

BEWEIS. Wir wenden Lemma 86 an:

$$\begin{aligned}
 |R(fg, U, \xi)| &= \left| \sum_{j=1}^n f(\xi_j)g(\xi_j)(x_j - x_{j-1}) \right| \\
 &\leq \sum_{j=1}^n \left( |f(\xi_j)|(x_j - x_{j-1})^{\frac{1}{p}} \right) \left( |g(\xi_j)|(x_j - x_{j-1})^{\frac{1}{q}} \right) \\
 &\leq \left( \sum_{j=1}^n |f(\xi_j)|^p (x_j - x_{j-1}) \right)^{\frac{1}{p}} \left( \sum_{j=1}^n |g(\xi_j)|^q (x_j - x_{j-1}) \right)^{\frac{1}{q}} \\
 &= R(|f|^p, U, \xi)^{\frac{1}{p}} R(|g|^q, U, \xi)^{\frac{1}{q}}.
 \end{aligned}$$

Ein Grenzübergang mit  $\delta(U) \rightarrow 0$  liefert die Behauptung. □

Unsere letzte Ungleichung betrifft das Verhalten der sogenannten  $p$ -Normen (wo  $p \in \mathbb{R}_+$ ), die für Vektoren  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  durch

$$\|x\|_p = \left( \sum_{j=1}^n |x_j|^p \right)^{\frac{1}{p}}$$

und für  $f \in R([a, b])$  durch

$$\|f\|_p = \left( \int_a^b |f(x)|^p \right)^{\frac{1}{p}}$$

definiert sind. Um den Namen “Norm” zu rechtfertigen, sollten diese zumindest die Dreiecksungleichung erfüllen:

LEMMA 87 (Minkowski’s Ungleichung). *Seien  $a_j, b_j \in \mathbb{R}_+$  für  $j = 1, \dots, n$ , und  $p \in (1, \infty)$ . Dann gilt*

$$\left( \sum_{j=1}^n (a_j + b_j)^p \right)^{\frac{1}{p}} \leq \left( \sum_{j=1}^n a_j^p \right)^{\frac{1}{p}} + \left( \sum_{j=1}^n b_j^p \right)^{\frac{1}{p}}.$$

*Gleichheit gilt genau dann, wenn es ein  $c > 0$  gibt sodass  $a_j = cb_j$  für  $j = 1, \dots, n$ .*

BEWEIS. Wir wenden Hölder’s Ungleichung an; der “konjugierte Exponent” zu  $p > 1$  ist  $q = \frac{p}{p-1}$ :

$$\begin{aligned} \sum_{j=1}^n (a_j + b_j)^p &= \sum_{j=1}^n (a_j + b_j)(a_j + b_j)^{p-1} \\ &= \sum_{j=1}^n a_j (a_j + b_j)^{p-1} + \sum_{j=1}^n b_j (a_j + b_j)^{p-1} \\ &\leq \left( \sum_{j=1}^n a_j^p \right)^{\frac{1}{p}} \left( \sum_{j=1}^n (a_j + b_j)^p \right)^{\frac{p-1}{p}} + \left( \sum_{j=1}^n b_j^p \right)^{\frac{1}{p}} \left( \sum_{j=1}^n (a_j + b_j)^p \right)^{\frac{p-1}{p}} \\ &= \left( \left( \sum_{j=1}^n a_j^p \right)^{\frac{1}{p}} + \left( \sum_{j=1}^n b_j^p \right)^{\frac{1}{p}} \right) \left( \sum_{j=1}^n (a_j + b_j)^p \right)^{1 - \frac{1}{p}}, \end{aligned}$$

also nach Multiplikation mit

$$\left( \sum_{j=1}^n (a_j + b_j)^p \right)^{-1 + \frac{1}{p}}$$

die behauptete Ungleichung. Gleichheit gilt, wenn für beide der verwendeten Hölderungleichungen Gleichheit gibt, es also ein  $c_1 > 0$  und ein  $c_2 > 0$  mit

$$a_j^p = c_1 (a_j + b_j)^{\frac{p}{p-1}}, \quad b_j^p = c_2 (a_j + b_j)^{\frac{p}{p-1}};$$

also dann, wenn

$$a_j = \left( \frac{c_1}{c_2} \right)^{\frac{1}{p}} b_j$$

gilt. □

Die Integralversion folgt entweder mit derselben Rechnung wie im Beweis oben, oder mit Hilfe der diskreten Minkowskiungleichung Lemma 87 und Grenzübergang, wie schon oben zweimal durchgeführt:

SATZ 50 (Minkowski’s Ungleichung). *Seien  $f, g \in R([a, b])$ . Dann gilt*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$



BEWEIS. Für eine beliebige Riemannsumme gilt

$$\begin{aligned}
 R(|f + g|^p, U, \xi)^{\frac{1}{p}} &\leq R(|f| + |g|^p, U, \xi)^{\frac{1}{p}} \\
 &= \left( \sum_{j=1}^n (|f(\xi_j)| + |g(\xi_j)|)^p (x_j - x_{j-1}) \right)^{\frac{1}{p}} \\
 &= \left( \sum_{j=1}^n \left( |f(\xi_j)|(x_j - x_{j-1})^{\frac{1}{p}} + |g(\xi_j)|(x_j - x_{j-1})^{\frac{1}{p}} \right)^p \right)^{\frac{1}{p}} \\
 &\leq \left( \sum_{j=1}^n |f(\xi_j)|^p (x_j - x_{j-1}) \right)^{\frac{1}{p}} + \left( \sum_{j=1}^n |g(\xi_j)|^p (x_j - x_{j-1}) \right)^{\frac{1}{p}} \\
 &= R(|f|^p, U, \xi)^{\frac{1}{p}} + R(|g|^p, U, \xi)^{\frac{1}{p}},
 \end{aligned}$$

und ein Grenzübergang  $\delta(U) \rightarrow 0$  liefert die Behauptung.  $\square$

Die Grenzfälle sind  $p = 1$ ,  $q = \infty$ , und die obigen Ungleichungen gelten auch in diesen Fällen, wenn man

$$\|x\|_{\infty} = \max_{j=1, \dots, n} |x_j|, \quad \|f\|_{\infty} = \sup_{x \in [a, b]} |f(x)|$$

setzt.

Bei den Integralungleichungen haben wir noch nichts über Gleichheit ausgesagt; der Grund dafür ist zunächst der, dass die Gleichheit  $\int_a^b |f(x)| dx = 0$  "nur" bedeutet, dass  $f$  fast überall (also ausserhalb einer Nullmenge) mit der konstanten Funktion 0 übereinstimmt.



## Reell-analytische Funktionen

### 1. Formale Potenzreihen

Eine *formale Potenzreihe*  $f(x)$  im Entwicklungspunkt  $x_0 \in \mathbb{R}$  ist ein Ausdruck der Form

$$f(x) = \sum_{j=0}^{\infty} f_j(x - x_0)^j,$$

wo die  $f_j$  komplexe Zahlen (die Koeffizienten der Reihe) sind. Obwohl wir uns dieser bequemen Schreibweise bedienen wollen, betonen wir an dieser Stelle gleich, dass die Bezeichnung  $f(x)$  in keiner Weise bedeutet, dass wir die Reihe auf der rechten Seite der Gleichung oben an irgendeiner Stelle  $x \in \mathbb{R}$  auswerten können, und wir keinerlei Recht haben, den Ausdruck als Funktion zu deuten; in diesem Sinn ist  $f(x)$  einfach eine bequeme Schreibweise für die Koeffizientenfolge  $(f_j)_{j \in \mathbb{N}}$ . Eine einzige Ausnahme zu dieser Regel ist die Auswertung an der Entwicklungsstelle  $x_0$ , welche einfach den konstanten Term (d.h. den Term der Ordnung 0) der Potenzreihe reproduziert,  $f(x_0) = f_0$ .

Formale Potenzreihen  $f(x)$  und  $g(x)$ , gegeben durch

$$f(x) = \sum_{j=0}^{\infty} f_j(x - x_0)^j, \quad g(x) = \sum_{j=0}^{\infty} g_j(x - x_0)^j,$$

werden addiert und mit Skalaren  $\lambda \in \mathbb{R}$  multipliziert, indem man diese Operationen auf die Koeffizienten anwendet, d.h. wir definieren

$$(f(x) + \lambda g(x)) = \sum_{j=0}^{\infty} (f_j + \lambda g_j)(x - x_0)^j.$$

Die Multiplikation wird definiert, indem man Terme der richtigen Ordnung zusammenfasst, d.h.

$$f(x)g(x) = \sum_{k \in \mathbb{N}} \left( \sum_{j+l=k} f_j g_l \right) (x - x_0)^k.$$

Mit diesen Operationen wird der Raum der formalen Potenzreihen, welcher mit  $\mathbb{C}[[x - x_0]]$  bezeichnet wird, zu einer Algebra mit Einselement  $f(x) = 1$  (die formale Potenzreihe mit konstantem Glied 1) über  $\mathbb{R}$ .

**Übungsaufgabe 50.** Zeige die letzte Behauptung!

Unsere formalen Potenzreihen haben zwar komplexe Koeffizienten, aber die Variable wird als "reelle Veränderliche" aufgefasst werden (im Moment ist sie nur ein Buchstabe!). In diesem Sinn können und werden wir eine formale Potenzreihe  $f \in \mathbb{C}[[x]]$  in ihren Real- und Imaginärteil durch

$$f(x) = \operatorname{Re} f(x) + i \operatorname{Im} f(x) = \sum_j (\operatorname{Re} f_j) x^j + i \sum_j (\operatorname{Im} f_j) x^j,$$

wo nun  $\operatorname{Re} f, \operatorname{Im} f \in \mathbb{R}[[x]]$ , zerlegen.

Will man zwei formale Potenzreihen  $g(y) \in \mathbb{C}[[y - y_0]]$  und  $f(x) \in \mathbb{C}[[x - x_0]]$  zu einer formalen Potenzreihe  $g(f(x)) \in \mathbb{C}[[x - x_0]]$  zusammensetzen, so muss man im Allgemeinen verlangen, dass der konstante Term  $f(x_0) = y_0$  ist, und setzt in diesem Fall

$$g(f(x)) = \sum_k g_k (f(x) - y_0)^k.$$

**Übungsaufgabe 51.** Warum ist diese Definition erlaubt? Gibt es andere Fälle, in denen die Zusammensetzung erlaubt ist?

Wir werden im folgenden oft mit dem Entwicklungspunkt  $x_0 = 0$  arbeiten; dies ist keine besondere Einschränkung, da ja für  $f(x) \in \mathbb{C}[[x]]$  die Zusammensetzung  $f(x - x_0) \in \mathbb{C}[[x - x_0]]$  definiert ist, und die so definierte Abbildung  $f(x) \mapsto f(x - x_0)$  ein Algebraisomorphismus ist. Die Einheiten der Algebra  $\mathbb{C}[[x]]$ , also die bezüglich der Multiplikation invertierbaren Elemente, sind leicht zu bestimmen:

LEMMA 88. Die formale Potenzreihe  $f(x) \in \mathbb{C}[[x]]$  ist eine Einheit genau dann, wenn  $f(0) \neq 0$ . Damit gibt es ein eindeutig bestimmtes maximales Ideal  $\mathfrak{m} \subset \mathbb{C}[[x]]$ , welches durch

$$\mathfrak{m} = \{f(x) : f(0) = 0\}$$

gegeben ist.

BEWEIS. Die Notwendigkeit der Bedingung ist offensichtlich: Wenn es eine Potenzreihe  $g$  mit  $f(x)g(x) = 1$  gibt, so ist  $f(0)g(0) = 1$ , also  $f(0) \neq 0$ . Sei nun  $f(0) \neq 0$ ; ohne Beschränkung der Allgemeinheit können wir annehmen, dass  $f(0) = 1$  ist. Wir können also  $f(x) = 1 - h(x)$  mit  $h(0) = 0$  schreiben.

Es ist uns schon bekannt (oder leicht zu sehen), dass die Potenzreihe

$$\sum_{j \in \mathbb{N}} y^j = G(y) \in \mathbb{C}[[y]]$$

die Eigenschaft hat, dass  $(1 - y)G(y) = 1$  ist. Wir ersetzen nun  $y$  durch  $h(x)$  und sehen, dass  $f(x)G(h(x)) = (1 - h(x))G(h(x)) = 1$  ist. Die letzte Aussage des Lemmas folgt, da ein echtes Ideal keine Einheiten enthalten kann, und die Menge der Elemente, welche nicht Einheiten sind, durch die Bedingung  $f(0) = 0$  charakterisiert wird, welche diese Menge als Ideal erkennbar macht.  $\square$

Formale Ableitungen und Stammfunktionen sind einfach zu definieren; für eine formale Potenzreihe  $f(x)$  ist die Ableitung definiert durch

$$f'(x) = \sum_{j=0}^{\infty} j f_j x^{j-1}.$$

Offensichtlich ist  $f'(x) = 0$  genau dann, wenn  $f(x) = r \in \mathbb{R}$  eine konstante Potenzreihe ist. Eine formale Stammfunktion  $F(x)$  zu  $f(x)$  ist durch

$$F(x) = \int f(x) dx = \sum_{j=0}^{\infty} \frac{f_j}{j+1} x^{j+1}$$

gegeben; diese spezielle Wahl zeichnet sich durch  $F(0) = 0$  aus. Wir führen nun einige wichtige formale Potenzreihen ein; diese kennen wir zwar auch als Funktionen, hier gehen wir aber zunächst rein formal vor.

**Übungsaufgabe 52.** Rechenregeln für Ableitungen (Produktregel, Quotientenregel, Kettenregel) und unbestimmte Integration gelten auch im formalen Sinn.

LEMMA 89. Sei  $\lambda \in \mathbb{C}$ . Dann gibt es für jedes  $y_0 \in \mathbb{R}$  genau eine formale Potenzreihe  $y(x) \in \mathbb{C}[[x]]$ , welche die Differentialgleichung  $y'(x) = \lambda y(x)$  mit der Anfangsbedingung  $y(0) = y_0$  erfüllt. Diese Potenzreihe wird als Exponentialreihe bezeichnet und ist durch

$$y(x) = y_0 e^{\lambda x} = y_0 \sum_{j=0}^{\infty} \frac{(\lambda x)^j}{j!}$$

gegeben. Sie erfüllt die Funktionalgleichung

$$e^{(\lambda+\mu)x} = e^{\lambda x} e^{\mu x}, \quad \lambda, \mu \in \mathbb{C}.$$

BEWEIS. Der unbestimmte Ansatz  $y(x) = y_0 + \sum_{j>0} y_j x^j$  führt durch Koeffizientenvergleich zu der Familie von Gleichungen

$$j y_j = y_{j-1}, \quad j \geq 1,$$

welche eindeutig durch  $y_j = \frac{y_0}{j!}$  gelöst wird.

Die Funktionalgleichung überprüfen wir, indem wir beachten, dass sich die Ableitung der rechten Seite  $f(x) = e^{\lambda x} e^{\mu x}$  als  $f'(x) = \lambda e^{\lambda x} e^{\mu x} + \mu e^{\lambda x} e^{\mu x} = (\lambda + \mu)f(x)$  berechnet. Damit erfüllt  $f(x)$  die Differentialgleichung, welche nach dem ersten Teil durch  $e^{(\lambda + \mu)x}$  eindeutig gelöst wird.  $\square$

**Übungsaufgabe 53.** Löse die Differentialgleichungen  $y''(x) = y(x)$  und  $y''(x) = -y(x)$  in  $\mathbb{C}[[x]]$ .

**Beispiel 33.** Der Real- und Imaginärteil der Reihe  $e^{ix}$  werden als  $\cos(x)$  und  $\sin(x)$  bezeichnet:

$$\operatorname{Re} e^{ix} = \cos(x) + i \sin(x).$$

Diese Reihen sind also durch

$$\cos(x) = \sum_k \frac{(-1)^k}{(2k)!} x^{2k}, \quad \sin(x) = \sum_k \frac{(-1)^k}{(2k+1)!} x^{2k+1}$$

gegeben.

LEMMA 90. Die formale Stammfunktion zu  $f(x) = (1+x)^{-1} \in \mathbb{C}[[x]]$ , welche durch

$$\ln(1+x) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} x^j$$

gegeben ist, wird als die Logarithmusreihe bezeichnet, und erfüllt  $e^{\ln(1+x)} = 1+x$  und  $\ln(e^x) = x$ .

BEWEIS. Da

$$\frac{1}{1+x} = \sum_{j \geq 0} (-1)^j x^j,$$

erhalten wir mit formaler Integration

$$\int \frac{1}{1+x} dx = \sum_{j \geq 0} \frac{(-1)^{j+1}}{j} x^j.$$

Die Zusammensetzung  $f(x) = e^{\ln(1+x)}$  ist wegen  $\ln(1+0) = 0$  wohldefiniert und erfüllt  $f(0) = 1$ , und wir erhalten unter Verwendung der Differentialgleichung der Exponentialreihe und  $f'(x) = \frac{e^{\ln(1+x)}}{1+x} = \frac{f(x)}{1+x}$ . Diese Differentialgleichung besitzt für die Anfangsbedingung  $f(0) = 1$  die eindeutige Lösung  $f(x) = 1+x$ .

Auf der anderen Seite ist die Zusammensetzung  $g(x) = \ln(e^x)$  wegen  $e^0 = 1$  definiert (eigentlich steht hier  $\ln(1 + (e^x - 1))$ ), erfüllt  $g(0) = 0$  und wir können wiederum  $g'(x) = \frac{1}{e^x} e^x = 1$  berechnen, also  $g(x) = x$  wie behauptet.  $\square$

## 2. Konvergente Potenzreihen

Eine formale Potenzreihe  $f(x) \in \mathbb{C}[[x - x_0]]$  ist *konvergent*, wenn es ein  $x_1 \neq x_0$  gibt sodass die Reihe

$$f(x_1) = \sum_j f_j (x_1 - x_0)^j$$

konvergiert. In diesem Fall definieren wir den *Konvergenzradius* von  $f$  als

$$R(f) = \sup \left\{ |x_1 - x_0| : \sum_j f_j (x_1 - x_0)^j \text{ konvergiert} \right\},$$

und  $R(f) = 0$  wenn es kein  $x_1 \neq x_0$  gibt sodass  $f(x_1)$  konvergiert. Die Menge der konvergenten Potenzreihen wird mit  $\mathbb{C}\{x - x_0\}$  bezeichnet.

LEMMA 91. Sei  $f \in \mathbb{C}[[x - x_0]]$ , und  $K \subset (x_0 - R(f), x_0 + R(f)) =: D(f)$  kompakt. Dann konvergieren die Partialsummen

$$T_{x_0}^k f = \sum_{j=0}^k f_j (x - x_0)^j$$

absolut und gleichmäßig auf  $K$ . Die auf  $D(f)$  als Grenzwert dieser Partialsummen definierte Funktion, welche wir wiederum mit  $f$  bezeichnen, ist stetig.

BEWEIS. Wir können  $R(f) > 0$  annehmen. Da  $K$  kompakt ist, gibt es ein  $x_1 \in D(f)$  und ein  $q < 1$  mit  $q|x_1 - x_0| > \max_{x \in K} |x - x_0|$ . Weiters ist, da ja  $f(x_1)$  konvergiert, die Folge  $f_j(x_1 - x_0)^j$  eine Nullfolge, also insbesondere beschränkt; wir wählen ein  $M > 0$  mit  $|f_j(x_1 - x_0)^j| \leq M$  für alle  $j \in \mathbb{N}$ . Dann ist für beliebige  $x \in K$

$$\sum_{j=\ell+1}^k |f_j(x - x_0)^j| \leq \sum_{j=\ell+1}^k Mq^j \leq \frac{q^{\ell+1}}{1-q}.$$

Also konvergieren die Partialsummen gleichmässig und absolut auf  $K$ . Die Stetigkeit der Grenzfunktion folgt aus Lemma 68.  $\square$

Wir können nun sehen, dass  $\mathbb{C}\{x\} \subset \mathbb{C}[[x]]$  eine Teilalgebra ist:

LEMMA 92. Seien  $f, g \in \mathbb{C}[[x]]$ . Dann gilt

$$R(f + \lambda g) \geq \min(R(f), R(g)), \quad R(fg) \geq \min(R(f), R(g)).$$

Ist  $f(0) = 0$ , und  $R(f), R(g) > 0$ , so ist auch  $R(g \circ f) > 0$ .

**Übungsaufgabe 54.** Beweise Lemma 92.

Die Berechnung des Konvergenzradius kann zum Beispiel mit Hilfe des Wurzel- oder des Quotientenkriteriums erfolgen:

LEMMA 93 (Hadamardkriterium). Sei  $f(x) \in \mathbb{C}[[x]]$ . Dann ist

$$R(f) = \frac{1}{\limsup_{j \rightarrow \infty} |f_j|^{1/j}},$$

wo wir  $1/\infty = 0$  und  $1/0 = \infty$  vereinbaren.

BEWEIS. Wir schreiben zunächst  $S = \frac{1}{\limsup_{j \rightarrow \infty} |f_j|^{1/j}} \neq 0, \infty$ . Ist nun  $|x| < S$ , so schreiben wir  $|x| = qS$  mit einem  $q < 1$ . Da  $S \limsup_{j \rightarrow \infty} |f_j|^{1/j} = 1$ , ist für beliebiges  $\varepsilon > 0$  die Ungleichung  $S^j |f_j| \leq (1 + \varepsilon)^j$  für fast alle  $j$  erfüllt. Wir wählen  $\varepsilon$  so klein, dass  $q' = q(1 + \varepsilon) < 1$  ist. Damit ist für fast alle  $j$  die Ungleichung  $|f_j||x|^j = |f_j|q^j S^j \leq (q')^j$  erfüllt; damit ist die Reihe  $f(x)$  für diese  $x$  konvergent. Ist auf der anderen Seite  $|x| > S$ , also  $|x| = Sq$  mit einem  $q > 1$ , so ist für unendlich viele  $j$  die Ungleichung  $|f_j|^{1/j} > \frac{1}{Sq}$  erfüllt, also  $|f_j||x|^j > 1$  für diese  $j$ ; damit kann die Reihe  $f(x)$  für diese  $x$  nicht konvergieren, da ihre Glieder nicht gegen 0 gehen. Es folgt  $R(f) = S$ .

Die Fälle  $S = 0$  und  $S = \infty$  werden ganz ähnlich behandelt.  $\square$

LEMMA 94. Sei  $f(x) \in \mathbb{C}[[x]]$  mit  $|f_j| > 0$  für fast alle  $j$ . Dann ist

$$R(f) = \lim_{j \rightarrow \infty} \frac{|f_j|}{|f_{j+1}|}.$$

BEWEIS. Sei  $S = \lim_{j \rightarrow \infty} \frac{|f_j|}{|f_{j+1}|}$ ; wir nehmen ohne Beschränkung der Allgemeinheit an, dass  $f_j \neq 0$  für alle  $j \in \mathbb{N}$ . Nun ist

$$\lim_{j \rightarrow \infty} \frac{|f_{j+1}||x|^{j+1}}{|f_j||x|^j} = \frac{|x|}{S};$$

also können wir das übliche Quotientenkriterium für  $|x| < S$  anwenden, um auf die Konvergenz zu schliessen, und erhalten für  $|x| > S$  Divergenz.  $\square$

**Übungsaufgabe 55.** Berechne die Konvergenzradien von  $e^x$  und  $\ln(1+x)$ .

Manchmal ist es bequemer, mit Wachstumsbedingungen an die Koeffizienten einer konvergenten Reihe zu arbeiten.

LEMMA 95. Sei  $f \in \mathbb{C}[[x]]$ ,  $f = \sum_j f_j x^j$ , und für  $s \geq 0$  definiere

$$M_s(f) = \sup_j |f_j| s^j.$$

Dann ist  $R(f) = \sup\{s \geq 0: M_s(f) < \infty\}$ .

Aus dem Hadamardkriterium und der Tatsache, dass  $\lim_{n \rightarrow \infty} n^{1/n} = 1$  ist, können wir folgende Beobachtung herleiten:

LEMMA 96. *Ist  $f \in \mathbb{C}[[x]]$  und  $F$  eine formale Stammfunktion von  $f$ , so ist  $R(f') = R(F) = R(f)$ . Insbesondere ist  $f \in C^\infty(D(f))$ .*

Die letzte Behauptung folgt dabei direkt aus Lemma 91 und Lemma 68, da die Partialsummen (als Polynome) glatt sind.

Das Verhalten einer Potenzreihe in den Randpunkten des Konvergenzintervalls  $D(f)$  kann ganz unterschiedlich ausfallen. So ist die Reihe

$$\frac{1}{1-x} = \sum_j x^j$$

weder im linken noch im rechten Randpunkt konvergent; die Logarithmusreihe

$$\ln(1+x) = \sum_j \frac{(-1)^{j+1} x^j}{j}$$

konvergiert für  $x = 1$  nach dem Leibnizkriterium (aber nicht absolut) und divergiert für  $x = -1$  als harmonische Reihe; die Reihe

$$\sum_j \frac{x^j}{j^2}$$

konvergiert in beiden Randpunkten absolut. Eine befriedigende Antwort auf die sich hier aufdrängenden Fragen kann erst im Rahmen der komplexen Analysis gegeben werden.

Wir wollen nun versuchen, (konvergente) Potenzreihen bezüglich der Zusammensetzung zu invertieren; wir arbeiten der Einfachheit halber wieder in  $\mathbb{C}[[x]]$ . Seien  $f, g \in \mathbb{C}[[x]]$  mit  $f(0) = g(0) = 0$  gegeben, welche  $f(g(x)) = x$  erfüllen. Wir sehen direkt, dass  $f'(0)g'(0) = 1$  ist. Mit Hilfe einer nichtverschwindenden Ableitung können wir die Existenz einer inversen Potenzreihe nachweisen. Wir beweisen zunächst ein einfaches Lemma, welches uns später zum Nachweis der Konvergenz der inversen Potenzreihe dienen wird.

LEMMA 97. *Sei  $N > 1$ . Dann gibt es eine Konstante  $E$  mit*

$$\sum_{j+k=\ell} \frac{(1+\ell)^N}{(1+j)^N(1+k)^N} < E.$$

BEWEIS. Es ist

$$\begin{aligned} \sum_{j+k=\ell} \frac{(1+\ell)^N}{(1+j)^N(1+k)^N} &\leq 2 \sum_{j \leq \ell/2} \frac{(1+\ell)^N}{(1+j)^N(1+(\ell-j))^N} \\ &\leq 2 \sum_{j \leq \ell/2} \frac{(1+\ell)^N}{(1+\ell/2)^N(1+j)^N} \\ &\leq 2^{N+1} \sum_{j=0}^{\infty} \frac{1}{(1+j)^N} =: E. \end{aligned}$$

□

LEMMA 98. *Sei  $N > 1$ . Dann gibt es eine Konstante  $E$ , sodass die sodass für jedes  $m \in \mathbb{N}$  der Koeffizienten  $d_m$  von  $t^m$  in*

$$\left( \sum_k c_k t^k \right)^\ell = \sum_j d_j t^j$$

die Ungleichung

$$d_m \leq \frac{(ME)^\ell s^m}{(1+m)^N}$$

erfüllt, wenn für  $k \leq m$  die Ungleichung  $c_k < M \frac{s^k}{(1+k)^N}$  erfüllt ist.

BEWEIS. Wir beweisen die Gültigkeit der Ungleichung mit Induktion über  $\ell$ ; für  $\ell = 1$  ist sie trivialerweise erfüllt. Wissen wir nun, dass für

$$\sum_j e_j t^j = \left( \sum_k c_k t^k \right)^{\ell-1}$$

die Abschätzung

$$e_j \leq \frac{(ME)^{\ell-1} s^j}{(1+j)^N}, \quad j \leq m,$$

erfüllt ist, so folgt für den Koeffizienten  $d_m$ , dass

$$\begin{aligned} d_m &= \sum_{p+q=m} e_p c_q \\ &\leq \sum_{p+q=m} \frac{(ME)^{\ell-1} s^p}{(1+p)^N} \frac{M s^q}{(1+q)^N} \\ &\leq \frac{M(ME)^{\ell-1} s^m}{(1+m)^N} \sum_{p+q=m} \frac{(1+m)^N}{(1+p)^N (1+q)^N} \\ &\leq \frac{(ME)^{\ell} s^m}{(1+m)^N}. \end{aligned}$$

□

SATZ 51. Sei  $f(x) \in \mathbb{C}[[x - x_0]]$  mit  $f(x_0) = y_0$  und  $f'(x_0) \neq 0$  gegeben. Dann gibt es ein eindeutig bestimmtes  $g(y) \in \mathbb{C}[[y - y_0]]$ , welches  $f(g(y)) = y$  erfüllt; es gilt dann auch  $g(f(x)) = x$ . Ist  $f \in \mathbb{C}\{x - x_0\}$ , so ist auch  $g \in \mathbb{C}\{y - y_0\}$ .

BEWEIS. Zunächst bemerken wir, dass wir ohne Beschränkung der Allgemeinheit annehmen können, dass  $x_0 = y_0 = 0$ : Ist  $f$  wie in den Voraussetzungen, so betrachten wir  $\varphi(x) = f(x + x_0) - y_0 \in \mathbb{C}[[x]]$ ;  $\varphi(x)$  erfüllt  $\varphi(0) = 0$  und  $\varphi'(0) = f'(x_0) \neq 0$ . Kennen wir den Satz für  $x_0 = 0$ , so wissen wir, dass es ein  $\psi \in \mathbb{C}[[x]]$  mit  $\varphi(\psi(x)) = x$  gibt. Wir setzen  $g(y) = \psi(y - y_0) + x_0$ , dann ist  $f(g(y)) = f(\psi(y - y_0) + x_0) = \varphi(\psi(y - y_0)) + y_0 = y - y_0 + y_0 = y$ .

Wir setzen  $g = \sum_j g_j x^j$  an. Die Gleichung  $f(g(x)) = x$  übersetzt sich in  $g_0 = 0$ ,  $f_1 g_1 = 1$ , und für  $j > 1$

$$f_1 g_j + P_j(f_k, g_k : k < j) = 0,$$

wo  $P_j$  ein Polynom ist, genauer gesagt, der Koeffizient von  $t^j$  in dem Ausdruck

$$\sum_{k=2}^j f_k \left( \sum_{\ell=1}^{j-k} g_\ell t^\ell \right)^k.$$

Wir sehen also, dass wir die  $g_j$  induktiv eindeutig bestimmen können, was die Existenz einer formalen Potenzreihe mit der gewünschten Eigenschaft zeigt.

Weiters erfüllt das konstruierte  $g$  die Voraussetzungen des Satzes, es gibt also ein  $h(x)$  mit  $g(h(x)) = x$ ; dieses erweist sich aber wegen  $f(x) = f(g(h(x))) = h(x)$  als das gegebene  $f$ .

Für den Konvergenzbeweis wollen wir ohne Beschränkung der Allgemeinheit annehmen, dass  $f'(0) = 1$  ist: falls  $f'(0) = \lambda$ , so erfüllt die Funktion  $\varphi = \lambda^{-1} f$ , dass  $\varphi'(0) = 1$ . Wenn wir das Ergebnis für solche Funktionen bewiesen haben, erhalten wir ein  $\psi$  mit  $\varphi(\psi(x)) = x$ . Setzen wir nun  $g(x) = \psi(\lambda^{-1} x)$ , so ist  $f(g(x)) = \lambda \varphi(\psi(\lambda^{-1} x)) = x$ .

Sei im folgenden  $N > 1$  fix. Da  $f$  konvergent ist, können wir  $K, C > 0$  mit  $|f_j| \leq CK^j$  bestimmen.

Die Koeffizienten  $g_j$  sind nun induktiv durch

$$g_j = -P_j(f_k, g_k)$$

gegeben. Wir behaupten, dass es  $L, M > 0$  gibt, sodass

$$|g_j| \leq \frac{ML^j}{(1+j)^N}.$$



Indem wir  $L$  genügend gross wählen, können wir  $M$  beliebig klein machen, sodass die Abschätzung für  $j = 1$  erfüllt ist. Angenommen, dies ist für  $j < m$  der Fall. Dann ist

$$|g_m| \leq \sum_{k=2}^m CK^k \frac{(ME)^k L^m}{(1+m)^N} \leq \frac{L^m}{(1+m)^N} \sum_{k=2}^m C(KME)^k \leq \frac{L^m}{(1+m)^N},$$

wenn wir  $L$  genügend gross wählen, sodass

$$\sum_{k=2}^{\infty} (KME)^k = \frac{(KME)^2}{1 - KME} < \frac{1}{C}$$

ist. □

### 3. Reell-analytische Funktionen

Sei  $U \subset \mathbb{R}$  offen. Eine Funktion  $\varphi: R \supset U \rightarrow \mathbb{C}$  heisst *reell-analytisch* auf  $U$ , wenn es für jedes  $x_0 \in U$  eine Potenzreihe  $f(x) \in \mathbb{C}\{x-x_0\}$  gibt, welche  $R(f) > 0$  erfüllt und in einer Umgebung von  $x_0$  die Funktion  $f$  darstellt, also  $f(x) = \varphi(x)$  für alle  $x$  in einer Umgebung von  $x_0$ . Die Menge der reell-analytischen Funktionen auf  $U$  wird mit  $C^\omega(U)$  bezeichnet.

**Übungsaufgabe 56.** Welche algebraische Struktur besitzt  $C^\omega(U)$ ?

Nach Lemma 91 sind reell-analytische Funktionen insbesondere glatt. Eine glatte Funktion, welche reell-analytisch ist, muss an jeder Stelle ihres Funktionsbereichs durch ihre Taylorreihe an dieser Stelle gegeben sein: Wenn  $\varphi(x) = \sum_j f_j(x-x_0)^j$  ist, so erhalten wir durch Differenzieren, dass  $j!f_j = \varphi^{(j)}(x_0)$  für alle  $j$  ist. Allgemein schreiben wir

$$T_{x_0}\varphi = \sum_j \frac{\varphi^{(j)}}{j!} (x-x_0)^j \in \mathbb{C}[[x-x_0]]$$

für die (formale) Taylorreihe von einer glatten Funktion  $\varphi$ , welche in einer Umgebung von  $x_0$  definiert ist. Nicht alle glatten Funktionen sind reell-analytisch, und auch die Konvergenz der Taylorreihe einer glatten Funktion ist keine Garantie dafür, dass diese Funktion reell-analytisch ist (also die Funktionswerte in einer Umgebung eines Punktes  $x_0$  auch tatsächlich dem Grenzwert der Taylorreihe in diesem Punkt entsprechen). Während wir also reell-analytische Funktionen mit ihren Taylorreihen identifizieren dürfen, ist dies für glatte Funktionen im Allgemeinen falsch:

**Übungsaufgabe 57.** Zeige, dass die Funktion  $\varphi(x) = e^{-\frac{1}{x^2}}$  glatt auf  $\mathbb{R}$  ist, und  $\varphi^{(j)}(0) = 0$  für alle  $j$  gilt.

Eine grosse Klasse von Funktionen, welche reell-analytisch sind, ist durch konvergente Potenzreihen gegeben.

**LEMMA 99.** Sei  $f(x) \in \mathbb{C}\{x-x_0\}$  eine konvergente Potenzreihe. Dann ist  $f(x)$  reell-analytisch auf  $D(f)$ . Für die Ableitungen gelten die Cauchy-Abschätzungen: für jedes  $S < R(f)$  gibt es ein  $M > 0$  sodass für jedes  $x$  mit  $|x-x_0| < S$

$$|f^{(k)}(x)| \leq \frac{Mk!}{(S-|x-x_0|)^{k+1}}$$

gilt.

BEWEIS. Wir nehmen ohne Beschränkung der Allgemeinheit an, dass  $f \in \mathbb{C}\{x\}$ , und schreiben  $R = R(f)$ . Für ein beliebiges  $x_0$  mit  $|x_0| < R$  versuchen wir  $f(x)$  in eine Potenzreihe um  $x_0$  zu entwickeln:

$$\begin{aligned} f(x) &= \sum_j f_j x^j \\ &= \sum_j f_j (x - x_0 + x_0)^j \\ &= \sum_j f_j \left( \sum_{k=0}^j \binom{j}{k} x_0^{j-k} (x - x_0)^k \right) \\ &= \sum_k \left( \sum_{j=k}^{\infty} k! \binom{j}{k} f_j x_0^{j-k} \right) \frac{(x - x_0)^k}{k!} \\ &= \sum_k f^{(k)}(x_0) \frac{(x - x_0)^k}{k!}. \end{aligned}$$

Nun müssen wir offensichtlich in irgend einer Form die Vertauschung der Summation rechtfertigen, was dann auch zeigt, dass die Potenzreihe um  $x_0$  am Ende der Rechnung wiederum konvergiert.

Dazu beachten wir, dass nach Lemma 93 für ein beliebiges  $S < R$  die Folge  $|f_j|S^j$  beschränkt ist, sagen wir durch  $M$ . Damit ist mit  $|x_0| = qS$

$$\begin{aligned} |f^{(k)}(x_0)| &= \left| \sum_j k! \binom{j}{k} f_j x_0^{j-k} \right| \\ &\leq \sum_j k! \binom{j}{k} |f_j| (qS)^{j-k} \\ &\leq \frac{M}{S^k} \sum_j (j(j-1)\dots(j-k+1)) q^{j-k} \\ &= \frac{Mk!}{S^k(1-q)^{k+1}} \\ &= \frac{MSk!}{(S - |x_0|)^{k+1}}. \end{aligned}$$

Die obige Doppelsomme

$$\sum_k \left( \sum_{j=k}^{\infty} k! \binom{j}{k} f_j x_0^{j-k} \right) \frac{(x - x_0)^k}{k!}$$

erweist sich mit Hilfe der Abschätzung ab der zweiten Zeile dieser Ungleichung als absolut konvergent; man kann nämlich die Summation von beliebig vielen Absolutbeträgen ihrer Glieder für  $|x - x_0| < S - |x_0|$  durch

$$\sum_k |f^{(k)}(x_0)| |x - x_0|^k \leq \sum_k \frac{MS|x - x_0|^k}{(S - |x_0|)^{k+1}} \leq \frac{C}{1 - \left| \frac{x - x_0}{S - x_0} \right|}$$

mit einer von  $x$  unabhängigen Konstanten  $C$  abschätzen. Der Rest der Abschätzung liefert die behauptete Cauchy-Ungleichung.  $\square$

**Übungsaufgabe 58.** Gib ein Beispiel einer reell-analytischen Funktion auf  $\mathbb{R}$ , welche nicht durch eine Potenzreihe mit Konvergenzradius  $R = \infty$  gegeben ist (hier ist es gut, sich Gedanken darüber zu machen, welche reell-analytischen Funktionen Einheiten sind)!

SATZ 52. Eine Funktion  $f \in C^\infty(U)$  ist reell-analytisch auf  $U$  genau dann, wenn es für jede kompakte Teilmenge  $K \subset U$  Konstanten  $C, r > 0$  gibt, für die

$$\sup_{x \in K} |f^{(k)}(x)| \leq k! C r^k, \quad k \in \mathbb{N}$$

gilt.

BEWEIS. Um zu sehen, dass die Bedingung hinreichend ist, zeigen wir, dass jede glatte Funktion auf  $U$  welche sie erfüllt in einer Umgebung jedes Punktes  $x_0 \in U$  Grenzwert ihrer Taylorreihe ist. Nach Theorem 39 können wir mit

$$T_{x_0}^k f(x) = \sum_{j=1}^k \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j$$

die Abweichung von  $f$  von seiner Taylorreihe als

$$f(x) - T_{x_0}^k f(x) = \frac{f^{(k+1)}(\xi)}{(k+1)!} (x - x_0)^{(k+1)}$$

mit einem  $\xi$  mit  $|\xi - x_0| < |x - x_0|$  schreiben. Wir wählen ein  $c > 0$  mit  $[x_0 - c, x_0 + c] \subset U$  und erhalten  $C, r > 0$ , für die dann

$$|f(x) - T_{x_0}^k f(x)| = \frac{|f^{(k+1)}(\xi)|}{(k+1)!} |x - x_0|^{k+1} \leq Cr^{k+1} |x - x_0|^{k+1} \rightarrow 0$$

für  $k \rightarrow \infty$  ist, solange  $|x - x_0| \leq \min\{c, 1/2r\}$ .

Die Notwendigkeit folgt aus einem Kompaktheitsargument und der Abschätzung aus Lemma 99.  $\square$

Gleichungen zwischen reell-analytischen Funktionen auf offenen Intervallen kann man überprüfen, indem man die entsprechenden Gleichungen für die Potenzreihendarstellungen an einem beliebigen Punkt überprüft.

LEMMA 100. Wenn  $f, g \in C^\omega((a, b))$  die Eigenschaft haben, dass  $f(x) = g(x)$  auf einer nichtleeren offenen Teilmenge  $U \subset (a, b)$  ist, dann ist  $f(x) = g(x)$  für alle  $x \in (a, b)$ . Äquivalent dazu ist, dass für ein  $x_0 \in (a, b)$  die Ableitungen von  $f$  und  $g$  an der Stelle  $x_0$  übereinstimmen, also

$$f^{(k)}(x_0) = g^{(k)}(x_0), \quad k \in \mathbb{N}$$

gilt.

BEWEIS. Wir wählen ein beliebiges  $x_0 \in U$ ; für dieses gilt, nachdem wir beide Funktionen in ihre Taylorreihe entwickeln können und diese gegen die jeweilige Funktion konvergieren, dass die Gleichheit  $f(x) = g(x)$  in einer Umgebung von  $x_0$  genau dann gilt, wenn  $f^{(k)}(x_0) = g^{(k)}(x_0)$  ist. Wir betrachten nun

$$c = \inf\{y \in (a, b) : f(x) = g(x) \text{ für alle } x \in [y, x_0]\}$$

und müssen zeigen, dass  $c = a$  ist. Dazu zeigen wir, dass für jedes  $y_0$  mit  $x_0 > y_0 > a$  und  $f(x) = g(x)$  für  $x \in [y_0, x_0]$  ein  $\varepsilon > 0$  existiert, sodass  $f(x) = g(x)$  auch für  $x \in (y_0 - \varepsilon, y_0]$  gilt. Daraus folgt  $c = a$ .

An der Stelle  $y_0$  können wir  $f$  und  $g$  in ihre Taylorreihen entwickeln. Nachdem  $f$  und  $g$  rechts von  $y_0$  übereinstimmen, ist  $f^{(k)}(y_0) = g^{(k)}(y_0)$  für alle  $k \in \mathbb{N}$ . Damit stimmen die durch ihre Taylorreihen dargestellten Funktion  $f$  und  $g$  auf einer  $\varepsilon$ -Umgebung von  $y_0$  überein.

Dasselbe Argument zeigt, dass

$$d = \sup\{y \in (a, b) : f(x) = g(x) \text{ für alle } x \in [x_0, y]\}$$

notwendigerweise  $d = b$  erfüllt, und damit ist  $f(x) = g(x)$  für alle  $x \in (a, b)$ .  $\square$

Reell-analytische Funktionen sind unter der Bildung von inversen Funktionen abgeschlossen:

SATZ 53. Sei  $f \in C^\omega(a, b)$ ,  $f: (a, b) \rightarrow (c, d)$ , mit  $f'(x) \neq 0$  für  $x \in (a, b)$ . Dann gibt es eine reell-analytische Funktion  $g: (c, d) \rightarrow (a, b)$  mit  $f(g(x)) = x$  für  $x \in (c, d)$  und  $g(f(x)) = x$  für  $x \in (a, b)$ .

BEWEIS. Da  $f'(x)$  auf  $(a, b)$  nicht verschwindet, ist  $f$  strikt monoton steigend oder fallend; wir können ohne Beschränkung der Allgemeinheit das erstere annehmen. Damit gibt es eine (stetig differenzierbare) inverse Funktion  $g(x)$  auf  $(c, d)$ , welche sich aber nach Theorem 51 in jedem Punkt  $x_0 \in (c, d)$  in eine konvergente Potenzreihe entwickeln lässt, also reell-analytisch ist.  $\square$

**Beispiel 34.** Die Exponentialfunktion  $\exp: \mathbb{R} \rightarrow \mathbb{R}_+$  ist die Lösung der Differentialgleichung  $y' = y$ ,  $y(0) = 1$ , und wir haben gesehen, dass sie durch die auf ganz  $\mathbb{R}$  konvergente Potenzreihe  $e^x$  gegeben ist; sie ist also nach Lemma 99 eine reell-analytische Funktion auf  $\mathbb{R}$ , welche darüberhinaus die Funktionalgleichung  $e^{a+b} = e^a e^b$  erfüllt. Ihre inverse Funktion wird als der Logarithmus  $\ln x$  bezeichnet, und ist eine reell-analytische Funktion auf  $\mathbb{R}_+$ .

Um  $\ln x$  in eine Potenzreihe zu entwickeln, ist es allerdings einfacher, die Ableitung von  $\ln x$  implizit zu berechnen, also die Gleichung

$$e^{\ln x} = x$$

zu differenzieren. Dies liefert  $e^{\ln x} \frac{d}{dx} \ln x = 1$ , also

$$\frac{d}{dx} \ln x = \frac{1}{x}.$$

Die rechte Seite dieser Gleichung lässt sich mit Hilfe der geometrischen Reihe sehr leicht in eine Potenzreihe entwickeln, und Integration liefert dann die gewünschte Reihe für  $\ln x$ .

Die Funktionalgleichung des Logarithmus,  $\ln ab = \ln a + \ln b$ , folgt aus der Funktionalgleichung der Exponentialfunktion wegen  $ab = e^{\ln a} e^{\ln b} = e^{\ln a + \ln b}$ .

**Beispiel 35.** Nachdem wir nun die Logarithmusfunktion eingeführt haben, definieren wir für  $\alpha \in \mathbb{R}$  die Potenzfunktionen

$$x^\alpha: \mathbb{R}_+ \rightarrow \mathbb{R}_+, : x^\alpha := e^{\alpha \ln x}.$$

Die Funktionalgleichung der Exponentialfunktion zieht die wichtige Gleichung

$$x^{\alpha+\beta} = x^\alpha x^\beta, \quad x \in \mathbb{R}_+, \quad \alpha, \beta \in \mathbb{R}$$

nach sich.

**Übungsaufgabe 59.** Berechne die Ableitung(en) der Potenzfunktion  $x^\alpha$ , und zeige (Entwicklung um den Punkt 1), dass

$$(1+x)^\alpha = \sum_{j=0}^{\infty} \binom{\alpha}{j} x^j,$$

wobei

$$\binom{\alpha}{j} = \frac{\alpha(\alpha-1)\dots(\alpha-j+1)}{j!}$$

die verallgemeinerten Binomialkoeffizienten sind.

#### 4. Trigonometrische Funktionen

Wir haben bereits die Reihen  $\cos(x)$  und  $\sin(x)$  eingeführt,

$$\cos(x) = \sum_k \frac{(-1)^k}{(2k)!} x^{2k}, \quad \sin(x) = \sum_k \frac{(-1)^k}{(2k+1)!} x^{2k+1}.$$

Diese Potenzreihen haben unendlichen Konvergenzradius, stellen also reell-analytische Funktionen auf  $\mathbb{R}$  dar. Weiters gilt

$$\cos^2(x) + \sin^2(x) = (\cos(x) + i \sin(x))(\cos(x) - i \sin(x)) = e^{ix} e^{-ix} = e^0 = 1,$$

und wir wissen damit dass  $\cos, \sin: \mathbb{R} \rightarrow [-1, 1]$ , weiters ist  $\cos(0) = 1$ ,  $\sin(0) = 0$ . Es ist auch einfach zu sehen, dass  $\sin'(x) = \cos(x)$  und  $\cos'(x) = -\sin(x)$  gilt. Wir bemerken auch, dass  $\cos$  gerade ist (d.h.  $\cos(-x) = \cos(x)$ ) und  $\sin$  ungerade (d.h.  $\sin(-x) = -\sin(x)$ ).

LEMMA 101. *Es gibt eine kleinste reelle positive Nullstelle  $\xi$  von  $\cos$ ; wir definieren  $\pi = 2\xi$ .*

BEWEIS. Angenommen,  $\cos(x) > 0$  für alle positiven  $x$ ; dann ist wegen  $\cos''(x) = -\cos(x)$  die Funktion  $\cos$  strikt konkav auf  $\mathbb{R}_+$ . Das ist ein Widerspruch zur Beschränktheit von  $\cos(x)$ .  $\square$

Um die weiteren Nullstellen der trigonometrischen Funktionen zu finden, benötigen wir die Additionstheoreme.

LEMMA 102. Für  $x, y \in \mathbb{R}$  gilt

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y), \quad \sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y);$$

insbesondere gelten die Halbwinkelformeln

$$\cos(x)^2 = \frac{1 + \cos(2x)}{2}, \quad \sin(x)^2 = \frac{1 - \cos(2x)}{2}.$$

BEWEIS. Es ist

$$\begin{aligned} \cos(x+y) + i\sin(x+y) &= e^i(x+y) \\ &= e^{ix}e^{iy} \\ &= (\cos(x) + i\sin(x))(\cos(y) + i\sin(y)) \\ &= \cos(x)\cos(y) - \sin(x)\sin(y) + i(\sin(x)\cos(y) + \cos(x)\sin(y)); \end{aligned}$$

die Formel folgt also aus Vergleich von Real- und Imaginärteilen.  $\square$

Damit ist  $\sin \pi = \sin(\pi/2 + \pi/2) = 2\cos(\pi/2)\sin(\pi/2) = 0$  und auch  $\sin(2\pi) = 0$ ; an diesen Stellen ist notwendigerweise  $\cos(\pi) = -1$  und  $\cos(2\pi) = 1$  (wir verwenden hier, dass  $\sin^2(x) + \cos^2(x) = 1$  und die Monotonieeigenschaften, die wegen  $\cos'(x) = -\sin(x)$  gelten). Wir können nun auch die Periodizität der trigonometrischen Funktionen nachweisen:

$$\begin{aligned} \cos(x+2\pi) &= \cos(x)\cos(2\pi) + \sin(x)\sin(2\pi) = \cos(x), \\ \sin(x+2\pi) &= \sin(x)\cos(2\pi) + \cos(x)\sin(2\pi) = \sin(x), \\ \sin(x+\pi/2) &= \sin(x)\cos(\pi/2) + \cos(x)\sin(\pi/2) = \cos(x), \\ \cos(x+\pi/2) &= \cos(x)\cos(\pi/2) - \sin(x)\sin(\pi/2) = -\sin(x). \end{aligned}$$

Insgesamt ergibt sich damit folgendes Bild für die Nullstellen:

LEMMA 103. Die Nullstellen von  $\cos(x)$  sind durch

$$x = \frac{\pi}{2} + n\pi, \quad n \in \mathbb{Z}$$

gegeben, und jene von  $\sin(x)$  durch

$$x = n\pi, \quad n \in \mathbb{Z}.$$

An diesen Punkten gilt

$$\cos(n\pi) = (-1)^n, \quad \sin\left(\frac{\pi}{2} + n\pi\right) = (-1)^n.$$

**Übungsaufgabe 60.** Zeige, dass die in Lemma 103 angegebenen Nullstellen tatsächlich alle sind.

Der Tangens und der Cotangens,

$$\tan(x) = \frac{\sin(x)}{\cos(x)}, \quad \cot(x) = \frac{\cos(x)}{\sin(x)},$$

sind dementsprechend reell-analytische Funktionen auf  $\mathbb{R} \setminus \{\pi/2 + n\pi : n \in \mathbb{Z}\}$  beziehungsweise  $\mathbb{R} \setminus \{n\pi : n \in \mathbb{Z}\}$ .

Die inversen Funktionen von  $\sin$  und  $\cos$  können jeweils für die Einschränkung auf Intervalle zwischen zwei Nullstellen dieser Funktionen definiert werden; dies gibt die verschiedenen "Zweige" dieser Funktionen. Standardmässig setzt man

$$\arcsin: [-1, 1] \rightarrow [-\pi/2, \pi/2], \quad \arccos: [-1, 1] \rightarrow [0, \pi],$$

welche nach Theorem 53 reell-analytisch auf  $(-1, 1)$  sind. Andere Wahlen des  $\arcsin$  bzw. des  $\arccos$  unterscheiden sich von den Hauptzweigen um ein Vielfaches von  $2\pi$ .

Ähnlich verfährt man mit der Umkehrfunktion des Tangens; wegen  $\lim_{x \rightarrow -\pi/2} \tan(x) = -\infty$ ,  $\lim_{x \rightarrow \pi/2} \tan(x) = +\infty$  ist  $\tan: (-\pi/2, \pi/2) \rightarrow \mathbb{R}$ , weiters ist  $\tan'(x) = (\cos(x))^{-2} \neq 0$  für  $x$ ; nach Theorem 53 gibt es also die Umkehrfunktion

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

welche reell-analytisch auf  $\mathbb{R}$  ist.

Die Ableitungen der Umkehrfunktionen berechnet man mit Hilfe von impliziter Differentiation und unter Verwendung der trigonometrischen Formeln. So ist

$$\sin(\arcsin(x)) = x,$$

also nach Ableitung

$$\cos(\arcsin(x)) \arcsin'(x) = 1.$$

Nun ist aber  $\cos^2(\arcsin(x)) + x^2 = 1$ , und damit (da  $\cos$  auf  $(-\pi/2, \pi/2)$  positiv ist)  $\cos(\arcsin(x)) = \sqrt{1-x^2}$ , also

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}.$$

**Übungsaufgabe 61.** Zeige, dass  $\arccos'(x) = \frac{-1}{\sqrt{1-x^2}}$ .

Die Reihenentwicklung von  $\arcsin$  erhält man nun aus der Entwicklung von

$$(1-x^2)^{-\frac{1}{2}} = \sum_{j=0}^{\infty} \binom{-1/2}{j} (-1)^j x^{2j}$$

durch Integration.

Eine ähnliche Strategie kann man für die Reihe des Arcustangens verwenden: Aus  $\tan(\arctan(x)) = x$  folgt  $\arctan'(x) = \cos^2(\arctan(x))$ , also  $\arctan'(x) = (1+x^2)^{-1}$ . Die Reihenentwicklung ergibt sich damit als

$$\arctan(x) = \sum_{j \geq 0} \frac{(-1)^j}{2j+1} x^{2j+1}.$$

Wenn wir nun in diese Entwicklung  $x = 1$  einsetzen, erhalten wir

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \dots$$

Natürlich dürfen wir in die rechte Seite der Gleichung nicht einfach  $x = 1$  einsetzen, da wir uns an keinem inneren Punkt des Konvergenzintervalls der Reihe befinden. Wir können aber wie folgt argumentieren, um zu sehen dass

$$\lim_{x \rightarrow 1+} \sum_{j \geq 0} \frac{(-1)^j}{2j+1} x^{2j+1} = \sum_{j \geq 0} \frac{(-1)^j}{2j+1} :$$

Sei  $\varepsilon > 0$ , dann existiert ein gerades  $N$  mit

$$\left| \sum_{j \geq N} \frac{(-1)^j}{2j+1} x^{2j+1} \right| \leq \frac{1}{4N+1} < \varepsilon$$

für alle  $x \in [1/2, 1]$ . Damit ergibt sich

$$\left| \sum_j \frac{(-1)^j}{2j+1} x^{2j+1} - \sum_j \frac{(-1)^j}{2j+1} \right| \leq 2\varepsilon + \left| \sum_{j \leq N} \frac{(-1)^j}{2j+1} x^{2j+1} - \sum_{j \leq N} \frac{(-1)^j}{2j+1} \right|$$

Der letzte Summand kann nun für  $x \in [r_0, 1]$  auch kleiner als  $\varepsilon$  gemacht werden, und wir erhalten die behauptete Aussage.

Um die trigonometrischen Funktionen ihrer üblichen geometrischen Interpretation zuzuführen, ist es von Vorteil, das Konzept eines Wegs einzuführen. Ein Weg in  $\mathbb{R}^2 = \mathbb{C}$  ist eine stetige Abbildung  $\gamma: [a, b] \rightarrow \mathbb{R}^2$ , wobei wir sagen, dass  $\gamma(t) = (x(t), y(t))$  differenzierbar ist, wenn  $x$  und  $y$  differenzierbar sind, und wir schreiben  $\gamma'(t) = (x'(t), y'(t))$ . Wir nennen  $\gamma'(t)$  den *Tangentenvektor* an die Kurve  $\gamma$  im Punkt  $\gamma(t)$  und sagen, dass  $\gamma$  regulär ist, wenn  $\gamma'(t) \neq 0$  für alle  $t \in [a, b]$  gilt.

Die Funktion  $\gamma(t) = e^{it}$  ist ein Weg in  $\mathbb{R}$ , wobei wir  $t$  zunächst auf ein Intervall der Länge  $2\pi$  einschränken wollen. Wegen  $\gamma'(t) = ie^{it}$  ist  $|\gamma(t)| = |\gamma'(t)| = 1$ ; mit anderen Worten,  $\gamma(t)$  durchläuft den Einheitskreis mit gleichförmiger Geschwindigkeit 1. Der Punkt  $\gamma(t) = (\cos(t), \sin(t))$  ist damit der Punkt auf dem Einheitskreis, der Bogenlänge  $t$  von  $(1, 0)$  auf der Kreislinie entfernt liegt.

Damit erkennt man  $\cos(t)$  als die Länge der Ankathete und  $\sin(t)$  als die Länge der Gegenkathete eines Winkels  $t$  in einem rechtwinkligen Dreieck, dessen Hypotenuse Länge 1 besitzt. Der  $\tan(t)$  wird erhalten,

indem man das Dreieck so streckt, dass die Länge der Ankathete 1 wird, und bezeichnet dann die Länge der Gegenkathete. Die Länge der Hypothenuse im so entstandenen Dreieck ist der Sekans von  $t$ ,  $\sec(t) = \cos(t)^{-1}$ . Eine andere Interpretation ist, dass  $\tan(t)$  die Länge der Strecke von  $(1, 0)$  zum Schnittpunkt der Geraden, welche mit der  $x$ -Achse den Winkel  $t$  einschliesst, mit der Geraden  $x = 1$  darstellt (für  $t \in (-\pi/2, \pi/2)$ ).

Ähnliche Interpretationen besitzen der Cotangens

$$\cot(t) = \frac{\cos(t)}{\sin(t)},$$

wobei hier die Gegenkathete auf Länge 1 gebracht wird; der Cosecans  $\csc(t) = \sin(t)^{-1}$  entspricht nun der Geraden  $y = 1$ .

Der Tangens und der Cotangens können mit Hilfe der Exponentialfunktion in folgender Form geschrieben werden:

$$\tan(t) = \frac{1}{i} \frac{e^{it} - e^{-it}}{e^{it} + e^{-it}} = \frac{1}{i} \frac{e^{2it} - 1}{e^{2it} + 1}, \quad \cot(t) = i \frac{e^{it} + e^{-it}}{e^{it} - e^{-it}} = i \frac{e^{2it} + 1}{e^{2it} - 1}.$$

Auch Tangens und Cotangens erfüllen Additionstheoreme. Diese kann man aus der elementaren Identität

$$\frac{ab - 1}{ab + 1} = \frac{(a - 1)(b + 1) + (b - 1)(a + 1)}{(a + 1)(b + 1) + (a - 1)(b - 1)} = \frac{\frac{a-1}{a+1} + \frac{b-1}{b+1}}{1 + \frac{a-1}{a+1} \frac{b-1}{b+1}}$$

ableiten, indem man  $a = e^{2it}$  und  $b = e^{2is}$  setzt, und so

$$\tan(s + t) = \frac{\tan(s) + \tan(t)}{1 - \tan(s) \tan(t)}$$

erhält.

**Beispiel 36.** Wir berechnen eine Stammfunktion von  $\sqrt{1 - x^2}$  auf dem Intervall  $(-1, 1)$ . Mit Hilfe der Substitution  $x = \sin(t)$  sehen wir, dass

$$\begin{aligned} \int \sqrt{1 - x^2} dx &= \int \cos(t)^2 dt \\ &= \int \frac{1 + \cos(2t)}{2} dt \\ &= \frac{t}{2} + \frac{\sin(2t)}{4} \\ &= \frac{1}{2} \left( \arcsin(x) + x\sqrt{1 - x^2} \right) \end{aligned}$$

Die Wahl eines Zweiges des arcsin (also die Addition einer Vielfachen von  $\pi$ ) ändert nur den Wert der Stammfunktion an einer Stelle; wir haben hier die Stammfunktion, welche an der Stelle 0 verschwindet, angegeben.





## Elementare Fourierreihen

### 1. Definition, Motivation und Beispiele

Wir haben schon eine Methode kennengelernt, eine Funktion durch “elementare” Funktionen zu approximieren. Dabei haben wir, um eine Funktion  $\varphi$  durch *Polynome* zu approximieren, ein *formales Objekt*, die Taylorreihe von  $\varphi$  betrachtet. Für reell-analytische Funktionen ist die Übereinstimmung von dieser Potenzreihe und der durch sie dargestellten Funktion gegeben, und wir können uns in Rechnungen auf die Manipulation eines algebraischen Objekts beschränken.

Leider sind Ableitungen einerseits schwer zu berechnen, und andererseits ist es notwendig, dass Funktionen glatt sind um ihre Taylorreihe zu berechnen; damit die Taylorreihe dann auch tatsächlich die Funktion darstellt, ist auch noch ein kontrolliertes Wachstum der Ableitungen dieser Funktion notwendig. Dies macht es oft erstrebenswert, Funktionen auch auf andere Weise darzustellen.

Eine *formale Fourierreihe* ist ein Ausdruck der Form

$$\frac{a_0}{2} + \sum_{j=1}^{\infty} a_j \cos(jx) + \sum_{j=1}^{\infty} b_j \sin(jx),$$

wo  $a_j, b_j \in \mathbb{C}$ . Eine Fourierreihe stellt, so sie konvergiert, eine Funktion dar, welche  $\varphi(x + 2\pi) = \varphi(x)$  erfüllt—wir sagen,  $\varphi$  ist periodisch mit Periode  $2\pi$  oder einfach,  $\varphi$  ist  $2\pi$ -periodisch.

Für eine  $2\pi$ -periodische Funktion genügt es, die Werte  $f(x)$  für  $a \leq x < a + 2\pi$  zu kennen; genauer gesagt, kommt  $f$  von einer Funktion auf dem Quotientenraum  $\mathbb{R}/\sim$ , wo  $x \sim y$  durch  $y - x \in 2\pi\mathbb{Z}$  definiert wird. Dieser Raum ist der eindimensionale Torus—im folgenden wollen wir uns einfach auf Funktionen  $f$  beschränken, die auf  $[-\pi, \pi]$  definiert sind, und Werte  $f(x)$  für  $x \notin [-\pi, \pi]$  dadurch definiert werden, dass  $f(x) = f(x + 2k\pi)$  wo  $x + 2k\pi \in [-\pi, \pi]$ .

Oft ist es vorteilhaft, unter Verwendung von  $e^{ix} = \cos(x) + i\sin(x)$  die *komplexe Darstellung* der Fourierreihe

$$\sum_{j \in \mathbb{Z}} c_j e^{ijx}$$

zu verwenden. Dabei ist  $2c_0 = a_0$ , und

$$c_j = \frac{1}{2}(a_j - ib_j), \quad j > 0, \quad c_j = \frac{1}{2}(a_{-j} + ib_{-j}), \quad j < 0.$$

Eine formale Fourierreihe ist *reell*, wenn  $a_j, b_j \in \mathbb{R}$  oder äquivalent  $c_j = \overline{c_{-j}}$ .

Wir erinnern an einen Satz über die Konvergenz von Funktionenreihen, den wir schon oft verwendet haben, und nun auch allgemein formulieren wollen:

**SATZ 54.** *Sei  $X$  ein topologischer Raum, und  $f_j$  stetige Funktionen auf  $X$ , welche*

$$\|f_j\|_{\infty} = \sup_{x \in X} |f_j(x)| \leq \varepsilon_j$$

*erfüllen. Wenn  $\sum_j \varepsilon_j < \infty$ , dann ist*

$$f(x) = \sum_j f_j(x)$$

*eine stetige Funktion auf  $X$ .*

BEWEIS. Da die Folge der Partialsummen

$$g_n(x) = \sum_{j=1}^n f_j(x)$$

für fixes  $x \in X$  eine Cauchyfolge ist, gibt es die Grenzfunktion  $f(x) = \lim_{n \rightarrow \infty} g_n(x)$ —mehr noch: für jedes fixe  $\varepsilon > 0$  können wir ein  $N$  wählen, sodass für  $n \geq N$  sogar für jedes  $x \in X$  die Ungleichung  $|f(x) - g_n(x)| \leq \|f - g_n\|_\infty < \varepsilon$  erfüllt ist.

Sei nun  $x_0 \in X$  und  $\varepsilon > 0$  beliebig. Wir müssen zeigen, dass es eine Umgebung  $U$  von  $X$  gibt, sodass  $|f(y) - f(x_0)| < \varepsilon$  für alle  $y \in U$  gilt. Sei  $N$  so gewählt, dass  $\|f - g_N\| < \varepsilon/3$  ist. Da  $g_N$  stetig ist (als Summe endlich vieler stetiger Funktionen), gibt es eine Umgebung  $U$  von  $x_0$  mit  $|g_N(y) - g_N(x_0)| < \varepsilon/3$ . Insgesamt können wir also

$$\begin{aligned} |f(y) - f(x_0)| &\leq |f(y) - g_N(y)| + |g_N(y) - g_N(x_0)| + |g_N(x_0) - f(x_0)| \\ &\leq \|f - g_N\|_\infty + \varepsilon/3 + \|f - g_N\|_\infty \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon \end{aligned}$$

gilt. □

Wir erinnern auch an die entsprechende Aussage für differenzierbare Funktionen, welche aus Lemma 68 folgt:

SATZ 55. Seien  $f_j \in C^k([a, b])$  gegeben, und sei

$$\left\| f_j^{(\ell)} \right\|_\infty = \max_{x \in [a, b]} |f_j^{(\ell)}(x)| \leq \varepsilon_j, \quad \ell \leq k.$$

Wenn  $\sum_j \varepsilon_j$  konvergiert, so ist  $\sum_j f_j \in C^k([a, b])$ .

Wir können nun ein Kriterium angeben, um die gleichmässige Konvergenz von einer formalen Fourierreihe gegen eine stetige Grenzfunktion zu garantieren.

SATZ 56. Sei  $\sum_j \varepsilon_j$  konvergent. Wenn die Koeffizienten der Fourierreihe

$$\frac{a_0}{2} + \sum_j a_j \cos(jx) + \sum_j b_j \sin(jx)$$

die Ungleichungen

$$|a_j| \leq \varepsilon_j, \quad |b_j| \leq \varepsilon_j$$

erfüllt, so konvergieren die Partialsummen der Fourierreihe

$$S_N(x) = \frac{a_0}{2} + \sum_{j \leq N} a_j \cos(jx) + \sum_{j \leq N} b_j \sin(jx)$$

gleichmässig gegen eine  $2\pi$ -periodische stetige Funktion  $f(x)$ .

Allgemeiner ist, falls sogar  $\sum_j j^k \varepsilon_j$  konvergiert,  $f \in C^k(\mathbb{R})$ .

Der Beweis erfolgt durch Überprüfen der Voraussetzungen von Theorem 55 unter Beachtung der Tatsache, dass man sich auf das kompakte Intervall  $[-\pi, \pi]$  einschränken kann.

**Bemerkung 15.** Äquivalent zur Voraussetzung in Theorem 56 ist, dass die Koeffizienten der komplexen Darstellung der Fourierreihe die Abschätzung  $|c_j| \leq \varepsilon_j$  erfüllen. Die Partialsummen sind in dieser Form durch

$$S_N(x) = \sum_{j=-N}^N c_j e^{ijx}$$

gegeben.

Wir fragen uns nun, wie man die Koeffizienten der stetigen Grenzfunktion einer (gleichmässig) konvergenten Fourierreihe aus der Grenzfunktion berechnen kann. Der Schlüssel dazu liegt in den folgenden Identitäten:

LEMMA 104. *Es gelten die Orthogonalitätsrelationen*

$$\int_{-\pi}^{\pi} e^{imx} e^{-inx} dx = \begin{cases} 0 & m \neq n \\ 2\pi & m = n. \end{cases},$$

beziehungsweise

$$\int_{-\pi}^{\pi} \cos(mx) \cos(nx) dx = \int_{-\pi}^{\pi} \sin(mx) \sin(nx) dx = \begin{cases} 0 & m \neq n \\ \pi & m = n \neq 0, \end{cases}$$

und

$$\int_{-\pi}^{\pi} \cos(mx) \sin(nx) dx = 0.$$

BEWEIS. Wir beginnen mit der Aussage für die komplexe Exponentialfunktion. Es ist  $e^{imx} e^{-inx} = e^{ijx}$ , wo  $j = m - n$  ist. Falls  $j \neq 0$ , besitzt diese Funktion die Stammfunktion  $\frac{e^{ijx}}{j}$ , und wir erhalten

$$\int_{-\pi}^{\pi} e^{imx} e^{-inx} dx = \frac{e^{ijx}}{j} \Big|_{-\pi}^{\pi} = \frac{1}{j} (e^{ij\pi} - e^{-ij\pi}) = \frac{1}{j} ((-1)^j - (-1)^j) = 0.$$

Wir verwenden nun diese Aussage, um die reellen Orthogonalitätsrelationen herzuleiten. Wir haben

$$\begin{aligned} \int_{-\pi}^{\pi} \cos(mx) \cos(nx) dx &= \frac{1}{4} \int_{-\pi}^{\pi} (e^{imx} + e^{-imx})(e^{inx} + e^{-inx}) dx \\ &= \frac{1}{4} \int_{-\pi}^{\pi} (e^{imx} + e^{-imx})(e^{inx} + e^{-inx}) dx \\ &= \frac{1}{4} \int_{-\pi}^{\pi} (e^{i(m+n)x} + e^{i(n-m)x} + e^{i(m-n)x} + e^{-i(m+n)x}) dx \\ &= \begin{cases} 0 & m \neq n, \\ \pi & m = n \neq 0, \\ 2\pi & m = n = 0. \end{cases} \end{aligned}$$

Für die Aussage für die Sinusfunktion rechnet man ganz analog, unter Verwendung von  $2i \sin(x) = e^{ix} - e^{-ix}$ .  $\square$

Wir können nun für eine gleichmässig konvergente Fourierreihe

$$\sum_j c_j e^{ijx} = f(x)$$

wie folgt rechnen:

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) e^{-inx} dx &= \int_{-\pi}^{\pi} \left( \sum_j c_j e^{ijx} \right) e^{-inx} dx \\ &= \sum_j \int_{-\pi}^{\pi} c_j e^{ijx} e^{-inx} dx \\ &= c_j 2\pi. \end{aligned}$$

Ähnlich sehen wir, dass wir  $a_j$  und  $b_j$  wie folgt berechnen können:

$$a_j = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(jx) dx, \quad b_j = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(jx) dx.$$

**Definition 45.** Sei  $f \in R([-\pi, \pi])$ , dann definieren wir die Fourierkoeffizienten

$$c_n(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx,$$

beziehungsweise

$$a_n(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx, \quad b_n(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx.$$

Bevor wir einige Beispiele berechnen, bemerken wir folgende Vereinfachung:

LEMMA 105. Wenn  $f \in R([-\pi, \pi])$  gerade ist (d.h.  $f(x) = f(-x)$ ), so ist  $b_j(f) = 0$  und

$$a_j(f) = \frac{2}{\pi} \int_0^{\pi} f(x) \cos(jx) dx$$

für alle  $j$ ; wenn  $f$  ungerade ist (d.h.  $f(-x) = -f(x)$ ), so ist  $a_j(f) = 0$  für alle  $j$  und

$$b_j(f) = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(jx) dx.$$

**Beispiel 37.** Wir berechnen die Fourierreihe der Dreiecksfunktion  $|x|$ . Da  $x \mapsto |x|$  gerade ist, müssen wir nur  $a_n$  berechnen;  $a_0 = \pi$  folgt einfach, und für  $j > 0$  haben wir

$$\begin{aligned} \pi a_j(|x|) &= \int_{-\pi}^{\pi} |x| \cos(jx) dx \\ &= 2 \int_0^{\pi} x \cos(jx) dx \\ &= 2 \left( \frac{x \sin(jx)}{j} - \int_0^{\pi} \frac{\sin(jx)}{j} dx \right) \\ &= 2 \left( \frac{\cos(j\pi)}{j^2} + \frac{1}{j^2} \right) \\ &= 2 \frac{(-1)^j + 1}{j^2}. \end{aligned}$$

Also ergibt sich die Fourierreihe als

$$\frac{\pi}{2} - \frac{4}{\pi} \left( \cos(x) + \frac{\cos(3x)}{9} + \frac{\cos(5x)}{25} + \dots \right),$$

und diese konvergiert gleichmäßig gegen  $|x|$  auf  $[-\pi, \pi]$ . Eine Auswertung für  $x = 0$  liefert

$$\sum_j \frac{1}{(2j+1)^2} = \frac{\pi^2}{8}.$$

**Beispiel 38.** Nun wollen wir die Fourierreihe der Rechteckfunktion  $\operatorname{sgn} x$  berechnen. Diese ist ungerade, und wir erhalten

$$b_j(\operatorname{sgn} x) = \frac{2}{\pi} \int_0^{\pi} \sin(jx) dx = \frac{2}{\pi} \left( -\frac{\cos(j\pi)}{j} + \frac{1}{j} \right) = \frac{2(1 + (-1)^{j+1})}{\pi j}.$$

Die Fourierreihe ist damit

$$\frac{4}{\pi} \left( \sin(x) + \frac{\sin(3x)}{3} + \frac{\sin(5x)}{5} + \dots \right)$$

Hier können wir die Konvergenz noch nicht garantieren. Falls die Summe konvergiert, so erhalten wir für  $x = \pi/2$

$$1 - \frac{1}{3} + \frac{1}{5} - \dots = \frac{\pi}{4}.$$

**Beispiel 39.** Für die Sägezahnfunktion  $x$  (welche ungerade ist) erhalten wir

$$\begin{aligned} b_j(x) &= \frac{2}{\pi} \int_0^{\pi} x \sin(jx) dx \\ &= \frac{2}{\pi} \left( -\frac{x \cos(jx)}{j} \Big|_0^{\pi} + \int_0^{\pi} \frac{\cos(jx)}{j} dx \right) \\ &= \frac{2(-1)^{j+1}}{j}. \end{aligned}$$

Damit ergibt sich die Fourierreihe als

$$2 \left( \sin(x) - \frac{\sin(2x)}{2} + \frac{\sin(3x)}{3} - \dots \right).$$

## 2. Das Riemann-Lebesgue Lemma

Das nächste Lemma, das sogenannte Riemann-Lebesgue Lemma, zeigt uns, dass die Fourierkoeffizienten einer integrierbaren Funktion eine Nullfolge bilden.

LEMMA 106. Sei  $f \in R([-\pi, \pi])$ . Dann gilt

$$\lim_{n \rightarrow \infty} a_n(f) = \lim_{n \rightarrow \infty} b_n = 0.$$

BEWEIS. Wir zeigen die Behauptung für  $a_n$ . Zunächst zeigen wir, dass die Behauptung für beliebige Treppenfunktionen  $t(x)$  gilt; dazu genügt es zu zeigen, dass die Behauptung für eine charakteristische Funktion eines Intervalls  $[a, b] \subset [-\pi, \pi]$  gilt:

$$\pi |a_n(\chi_{[a,b]})| = \left| \int_a^b \cos(nx) dx \right| = \left| \frac{\sin(nb)}{n} - \frac{\sin(na)}{n} \right| \leq \frac{2}{n},$$

also ist  $a_n$  eine Nullfolge.

Sei nun  $\varepsilon > 0$  beliebig. Dann gibt es eine Treppenfunktion  $t$  mit

$$\frac{1}{\pi} \int_{-\pi}^{\pi} |t(x) - f(x)| dx \leq \varepsilon,$$

und wir können wie folgt abschätzen:

$$\begin{aligned} |a_n(f)| &= \left| \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx \right| \\ &\leq \frac{1}{\pi} \left( \int_{-\pi}^{\pi} |f(x) - t(x)| |\cos(nx)| dx \right) + |a_n(t)| \\ &\leq \varepsilon + |a_n(t)|. \end{aligned}$$

Da  $a_n(t) \rightarrow 0$  für  $n \rightarrow \infty$ , gibt es ein  $N \in \mathbb{N}$  sodass  $|a_n(t)| < \varepsilon$  für  $n \geq N$ . Für diese  $n$  ist also  $|a_n(f)| < 2\varepsilon$ ; und damit ist  $a_n(f)$  eine Nullfolge.  $\square$

Für stetige Funktionen wollen wir einen etwas anderen Beweis geben, der eine Abschätzung liefert, wie schnell die Fourierkoeffizienten gegen 0 gehen. Wenn  $f$  stetig auf  $[-\pi, \pi]$  ist, so ist mit Hilfe einer Koordinatenwechsels

$$\begin{aligned} 2\pi c_n(f) &= \int_{-\pi}^{\pi} f(x) e^{-inx} dx \\ &= \int_{-\pi(1+1/n)}^{\pi(1+1/n)} f(x - \pi/n) e^{-in(x-\pi/n)} dx \\ &= - \int_{-\pi}^{\pi} f(x - \pi/n) e^{-inx} dx, \end{aligned}$$

wobei wir unsere Verabredung über die Periodizität verwendet haben, sowie die Formel  $e^{i\pi} = -1$ . Damit ergibt sich

$$\begin{aligned} |c_n(f)| &= \left| \frac{1}{4\pi} \left( \int_{-\pi}^{\pi} f(x) e^{-inx} dx - \int_{-\pi}^{\pi} f(x - \pi/n) e^{-inx} dx \right) \right| \\ &\leq \frac{1}{2} \sup_{x \in [-\pi, \pi]} |f(x) - f(x - \pi/n)|. \end{aligned}$$

Der letzte Ausdruck geht für  $n \rightarrow \infty$  gegen 0, da  $f$  stetig auf dem kompakten Intervall  $[-\pi, \pi]$ , also gleichmäßig stetig ist.

Der letzte Ausdruck in der Ungleichungskette gibt uns auch einen Anhaltspunkt, um das Riemann-Lebesgue Lemma zu quantifizieren, also die Geschwindigkeit der Konvergenz der Koeffizienten zu bestimmen. Dazu definieren wir, dass eine stetige  $2\pi$ -periodische Funktion  $f$  Hölder-stetig von der Ordnung  $\alpha$  ist, wenn

$$|f(x) - f(y)| \leq C|x - y|^\alpha$$

für alle  $x, y \in [-\pi, \pi]$  ist. Die Menge der Hölder-stetigen Funktionen wird mit  $C^\alpha([-\pi, \pi])$  bezeichnet; sie wird mit der Festsetzung

$$\|f\|_\alpha = \|f\|_\infty + \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\alpha}$$

zu einem normierten Vektorraum (welcher sogar ein Banachraum ist). Allgemeiner definieren wir den Raum  $C^{k,\alpha}([-\pi, \pi])$  als den Raum der Funktionen  $f \in C^k([-\pi, \pi])$  mit  $f^{(k)} \in C^\alpha([-\pi, \pi])$ , und wird mit

$$\|f\|_{k,\alpha} = \sum_{j=1}^k \|f^{(j)}\|_\infty + \|f^{(k)}\|_\alpha$$

zu einem Banachraum.

Wir weisen die Vollständigkeit von  $C^\alpha([-\pi, \pi])$  nach. Sei also  $f_n$  eine Cauchyfolge in  $C^\alpha([-\pi, \pi])$ . Dann ist  $f_n$  eine Cauchyfolge in  $C([-\pi, \pi])$ , und damit gibt es eine stetige Funktion  $f$ , die der gleichmässige Grenzwert der  $f_n$  ist. Für diese erhalten wir für  $x \neq y$

$$\frac{|f(x) - f(y)|}{|x - y|^\alpha} = \lim_{n \rightarrow \infty} \frac{|f_n(x) - f_n(y)|}{|x - y|^\alpha} \leq \limsup_{n \rightarrow \infty} \|f_n\|_\alpha,$$

also  $f \in C^\alpha([-\pi, \pi])$ . Nun müssen wir noch nachweisen, dass  $f$  auch der Grenzwert der  $f_n$  in der  $C^\alpha([-\pi, \pi])$ -Norm ist:

$$\frac{|f(x) - f_n(x) - f(y) + f_n(y)|}{|x - y|^\alpha} = \lim_{m \rightarrow \infty} \frac{|f_m(x) - f_n(x) - f_m(y) + f_n(y)|}{|x - y|^\alpha} \leq \limsup_{m \rightarrow \infty} \|f_m - f_n\|_\alpha.$$

Also ist  $\|f - f_n\|_\alpha \leq \|f - f_n\|_\infty + \limsup_{m \rightarrow \infty} \|f_m - f_n\|_\alpha$ , und der letzte Ausdruck geht gegen 0 für  $n$  gegen  $\infty$ , was zeigt, dass  $\lim_{n \rightarrow \infty} f_n = f$  in  $C^\alpha([-\pi, \pi])$ .

Dieselbe Rechnung wie oben zeigt für  $f \in C^\alpha([-\pi, \pi])$ , dass

$$\begin{aligned} |c_n(f)| &= \left| \frac{1}{4\pi} \left( \int_{-\pi}^{\pi} f(x) e^{-inx} dx - \int_{-\pi}^{\pi} f(x - \pi/n) e^{-inx} dx \right) \right| \\ &\leq \frac{1}{2} \sup_{x \in [-\pi, \pi]} |f(x) - f(x - \pi/n)| \\ &\leq C|n|^{-\alpha}. \end{aligned}$$

Die Konstante  $C$  hängt hier nur von  $\|f\|_\alpha$  ab. Eine Ungleichung für die  $a_n(f)$  bzw.  $b_n(f)$  folgt nun, da  $2c_n = a_n - ib_n$ ,  $2c_{-n} = a_n + ib_n$ , also

$$|a_n(f)| \leq |c_n(f)| + |c_{-n}(f)| \leq 2C|n|^{-\alpha}.$$

Im Falle einer Funktion  $f \in C^{k,\alpha}$  können wir zunächst  $k$ -mal partiell integrieren und erhalten

$$c_n(f) = \frac{c_n(f^{(k)})}{n^k};$$

insgesamt haben wir damit folgendes gezeigt:

LEMMA 107. *Es gibt Konstanten  $C, K > 0$ , sodass für  $f \in C^{k,\alpha}([-\pi, \pi])$  die Ungleichungen*

$$\begin{aligned} |c_n(f)| &\leq \frac{C \|f\|_{k,\alpha}}{|n|^{k+\alpha}}, \quad n \in \mathbb{Z}, \\ |a_n(f)| &\leq \frac{K \|f\|_{k,\alpha}}{n^{k+\alpha}}, \quad |b_n(f)| \leq \frac{K \|f\|_{k,\alpha}}{n^{k+\alpha}}, \quad n \in \mathbb{N}, \end{aligned}$$

erfüllt sind.

### 3. Die Bessel-Ungleichung

Eine andere Methode, das Grössenwachstum von Fourierkoeffizienten zu kontrollieren, ist mit Hilfe von Integralen; der Ausgangspunkt ist die Ungleichung von Bessel. Wir erinnern an unsere Notation für die  $N$ -te Partialsumme einer Fourierreihe:

$$S_N f(x) = \sum_{j=-N}^N c_j(f) e^{ijx}.$$

Wir berechnen

$$\begin{aligned} |f(x) - S_N f(x)|^2 &= (f(x) - S_N f(x)) \overline{(f(x) - S_N f(x))} \\ &= |f(x)|^2 - f(x) \overline{S_N f(x)} - \overline{f(x)} S_N f(x) + |S_N f(x)|^2. \end{aligned}$$

Wenn wir diese Gleichung integrieren, erhalten wir

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |S_N f(x)|^2 dx = \sum_{j=-N}^N |c_j(f)|^2$$

und

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \sum_{j=-N}^N \overline{c_j(f)} e^{-ijx} dx = \sum_{j=-N}^N |c_j(f)|^2.$$

Zusammengenommen erhalten wir

$$0 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - S_N f(x)|^2 dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx - \sum_{j=-N}^N |c_j(f)|^2$$

Auch daraus folgt die Konvergenz der  $c_j$  gegen 0!

Wir fassen die Bessel-Ungleichung zusammen:

SATZ 57. Sei  $f \in R([a, b])$ . Dann gilt die folgende Gleichung:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - S_N f(x)|^2 dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx - \sum_{j=-N}^N |c_j(f)|^2;$$

insbesondere ist

$$\sum_{j=-\infty}^{\infty} |c_j(f)|^2 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx.$$

Die Gleichheit

$$\sum_{j=-\infty}^{\infty} |c_j(f)|^2 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx$$

gilt genau dann, wenn die Partialsummen  $S_N f$  im quadratischen Mittel gegen  $f$  konvergieren, also

$$\lim_{N \rightarrow \infty} \int_{-\pi}^{\pi} |f(x) - S_N f(x)|^2 dx = 0$$

ist. Insbesondere gilt also Gleichheit, wenn  $f$  der gleichmässige Grenzwert der  $S_N f$  ist, z.Bsp. wenn  $f$  selber der gleichmässige Grenzwert einer (seiner) Fourierreihe ist, also unter den Voraussetzungen von Theorem 56. Noch können wir diese Konvergenz einer Fourierreihe gegen die Funktion, welche sie erzeugt, nicht beweisen – sie gilt auch im allgemeinen nicht. Wir wollen uns nun auf einen Spezialfall konzentrieren.

#### 4. Konvergenz von Fourierreihen

Wir betrachten zunächst die Partialsummen einer Fourierreihe etwas genauer. Es ist

$$\begin{aligned}
 S_N f(x) &= \sum_{j=-N}^N c_j(f) e^{ijx} \\
 &= \sum_{j=-N}^N \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ij(t-x)} dt \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left( \sum_{j=-N}^N e^{ij(x-t)} \right) dt \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x+t) \left( \sum_{j=-N}^N e^{ijt} \right) dt \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x+t) D_N(t) dt
 \end{aligned}$$

Die Funktion  $D_N(t)$  heisst *Dirichletkern* und hat eine einfachere Darstellung wie folgt:

$$\begin{aligned}
 D_N(t) &= \sum_{j=-N}^N e^{ijt} \\
 &= e^{-iNt} \sum_{j=0}^{2N} e^{ijt} \\
 &= e^{-iNt} \frac{1 - e^{i(2N+1)t}}{1 - e^{it}} \\
 &= \frac{e^{-iNt} - e^{i(N+1)t}}{1 - e^{it}} \\
 &= \frac{e^{-\frac{it}{2}} e^{-iNt} - e^{i(N+1)t}}{e^{-\frac{it}{2}} (1 - e^{it})} \\
 &= \frac{e^{-i(N+\frac{1}{2})t} - e^{i(N+\frac{1}{2})t}}{e^{-i\frac{1}{2}t} - e^{i\frac{1}{2}t}} \\
 &= \frac{\sin\left(\left(N + \frac{1}{2}\right)t\right)}{\sin\left(\frac{1}{2}t\right)}.
 \end{aligned}$$

Es gilt

$$\int_{-\pi}^{\pi} D_N(t) dt = 2\pi$$

, und damit erhalten wir folgende Darstellung der Abweichung einer Funktion von einer Partialsumme ihrer Fourierreihe:

$$f(x) - S_N f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x) - f(x+t)) D_N(t) dt.$$

Wir sehen also, dass die Konvergenz einer Fourierreihe einer Funktion gegen diese Funktion an einem gegebenen Punkt  $x$  damit in Verbindung gesetzt werden kann, wie die Funktion  $f(x) - f(x+t)$  sich am Punkt  $t = 0$  verhält. Genauer wollen wir das für Funktionen formulieren, welche eine Hölder-Bedingung erfüllen:

**SATZ 58.** *Sei  $f \in C^\alpha([-\pi, \pi])$ . Dann konvergieren die Partialsummen der Fourierreihe von  $f$  gleichmässig gegen  $f$  auf  $[-\pi, \pi]$ .*



## Mehrdimensionale Differentialrechnung

### 1. Lineare Algebra

Wir arbeiten in endlich Vektorräumen über  $\mathbb{R}$ , das heisst, in Vektorräumen welche isomorph zu

$$\mathbb{R}^m = \{(x_1, \dots, x_m) : x_j \in \mathbb{R}\}$$

mit der üblichen komponentenweisen Addition und skalaren Multiplikation

$$(x_1, \dots, x_m) + (y_1, \dots, y_m) = (x_1 + y_1, \dots, x_m + y_m), \quad \lambda(x_1, \dots, x_m) = (\lambda x_1, \dots, \lambda x_m)$$

sind. Wir werden im folgenden diese Zeilenvektoren mit Spaltenvektoren nach Belieben identifizieren (eine Begründung dafür kommt noch). Die Vektoren

$$\mathbf{e}_i = (0, \dots, \underbrace{1}_{j\text{th spot}}, \dots, 0), \quad i = 1, \dots, m,$$

bilden eine Basis des  $\mathbb{R}^m$ , die Standardbasis. Wenn wir verschiedene Vektorräume  $\mathbb{R}^m$  und  $\mathbb{R}^n$  betrachten, werden wir dieselbe Notation für die Standardbasen in beiden Räumen verwenden (die Verwirrung, die dadurch entsteht, ist der Überladung der Notation vorzuziehen).

Der  $\mathbb{R}^m$  besitzt ein inneres Produkt, das Standardprodukt, definiert durch

$$\langle x, y \rangle = \sum_j x_j y_j,$$

und die damit verbundene euklidische Norm

$$\|x\|^2 = \langle x, x \rangle.$$

Eine Bemerkung ist, dass jede andere Norm mit der euklidischen Norm vergleichbar ist:

LEMMA 108. Sei  $\|\cdot\|_1$  eine Norm auf  $\mathbb{R}^m$ . Dann gibt es eine Konstante  $C > 0$  mit

$$\frac{1}{C} \|x\|_1 \leq \|x\| \leq C \|x\|_1.$$

BEWEIS. Sei  $\|\mathbf{e}_i\|_1 = \lambda_i > 0$ . Dann ist nach der Cauchy-Schwarzschen Ungleichung

$$\|x\|_1 = \left\| \sum_{j=1}^m x_j \mathbf{e}_j \right\|_1 \leq \sum_j |x_j| \lambda_j \leq \sqrt{\sum_j x_j^2} \sqrt{\sum_j \lambda_j^2} \leq L \|x\|.$$

Insbesondere ist die Identität als Abbildung von  $\mathbb{R}^m$  versehen mit der Topologie von  $\|\cdot\|_1$  nach  $\mathbb{R}^m$  versehen mit der Topologie von  $\|\cdot\|$  stetig. Damit ist die Einheitssphäre

$$S_1 = \{x : \|x\|_1 = 1\}$$

kompakt in  $\mathbb{R}^m$  mit der Topologie von  $\|\cdot\|$ . Damit ist

$$K = \min_{x \in S_1} \|x\| > 0$$

und

$$\left\| \frac{x}{\|x\|_1} \right\| \geq K, \quad x \in \mathbb{R}^m$$

also  $K \|x\|_1 \leq \|x\|$ . Wenn wir nun  $C = \max L, K^{-1}$  wählen, ist die Behauptung des Lemmas erfüllt.  $\square$

Insbesondere ist die durch eine beliebige Norm definierte Topologie auf  $\mathbb{R}^m$  eindeutig bestimmt und stimmt mit der Produkttopologie überein.

Eine Abbildung  $A: \mathbb{R}^m \rightarrow \mathbb{R}^n$  heisst linear, wenn sie mit den linearen Operationen vertauscht, also

$$A(x + \lambda y) = A(x) + \lambda A(y)$$

für alle  $x, y \in \mathbb{R}^m$  und  $\lambda \in \mathbb{R}$  gilt. Der Raum der linearen Abbildungen von  $\mathbb{R}^m$  nach  $\mathbb{R}^n$  wird mit der Festsetzung  $(A + \lambda B)(x) = A(x) + \lambda B(x)$  selber zu einem endlichdimensionalen  $\mathbb{R}$ -Vektorraum  $L(\mathbb{R}^m, \mathbb{R}^n)$ . Er wird mit dem Raum der  $n \times m$ -Matrizen ( $n$  Zeilen,  $m$  Spalten) mit reellen Einträgen identifiziert, also mit  $\mathbb{R}^{mn}$ . Dabei entspricht die lineare Abbildung  $A$  der Matrix mit Einträgen

$$A_{i,j} = \langle \mathbf{e}_i, A\mathbf{e}_j \rangle,$$

oder in Worten: der Eintrag in der  $i$ -ten Zeile und  $j$ -ten Spalte der Matrix zu  $A$  ist die  $i$ -te Koordinate des Bilds des  $j$ -ten Basisvektors unter  $A$ . Jede lineare Abbildung ist offensichtlich stetig.

## 2. Differenzierbarkeit

Sei  $f: \mathbb{R}^m \supset \Omega \rightarrow \mathbb{R}^n$  eine auf der offenen Menge  $\Omega$  definiert. Wir sagen,  $f$  ist differenzierbar an der Stelle  $x \in \Omega$ , wenn es eine lineare Abbildung  $A \in L(\mathbb{R}^m, \mathbb{R}^n)$  gibt, sodass für

$$R(h) = f(x + h) - f(x) - Ah$$

$$\lim_{h \rightarrow 0} \frac{R(h)}{\|h\|} = 0$$

gilt. Die lineare Abbildung  $A$  ist, falls sie existiert, durch diese Bedingung eindeutig bestimmt und wird die Ableitung von  $f$  an der Stelle  $x$  genannt; wir schreiben dafür  $f'(x) \in L(\mathbb{R}^m, \mathbb{R}^n)$ . Wenn  $f$  an jeder Stelle von  $\Omega$  differenzierbar ist, so sagen wir,  $f$  ist auf  $\Omega$  differenzierbar; die Menge der auf  $\Omega$  differenzierbaren Funktionen wird mit  $D(\Omega)$  bezeichnet; es ist einfach zu sehen, dass  $D(\Omega) \subset C(\Omega)$ . Die Ableitung ist dann selber eine Funktion  $f': \mathbb{R}^m \rightarrow L(\mathbb{R}^m, \mathbb{R}^n) = \mathbb{R}^{mn}$ ; ist sie wiederum auf  $\Omega$  differenzierbar, so sagen wir,  $f$  ist zweimal differenzierbar auf  $\Omega$ , und bezeichnen die Ableitung mit  $f''$ ; induktiv definieren wir, dass  $f$   $k$ -mal differenzierbar ist, wenn  $f$   $(k-1)$ -mal differenzierbar ist und  $f^{(k-1)}: \Omega \rightarrow \mathbb{R}^{m^{k-1}n}$  wiederum auf  $\Omega$  differenzierbar ist. Für jedes  $x \in \Omega$  ist also  $f^{(k)}(x) \in L(\mathbb{R}^m, L(\mathbb{R}^m, \dots, L(\mathbb{R}^m, \mathbb{R}^n) \dots)) = L^k(\mathbb{R}^m, \mathbb{R}^n)$ , wo  $L^k(\mathbb{R}^m, \mathbb{R}^n)$  der Raum der  $k$ -multilinearen Abbildungen  $A$  auf  $\mathbb{R}^m$  ist, also Abbildungen  $A: (\mathbb{R}^m)^k \rightarrow \mathbb{R}^n$ , welche in jede ihrer Variablen linear sind, beziehungsweise explizit

$$A(h^1, \dots, h^j + \lambda \tilde{h}^j, \dots, h^k) = A(h^1, \dots, h^j, \dots, h^k) + \lambda A(h^1, \dots, \tilde{h}^j, \dots, h^k)$$

erfüllen. Die Auswertung  $f^{(k)}(x)(h^1, \dots, h^k)$ , wo  $h^j \in \mathbb{R}^m$  für  $j = 1, \dots, k$  ist durch

$$f^{(k)}(x)(h^1, \dots, h^k) = \left( (f^{(k-1)})'(x)(h^k) \right) (h^1, \dots, h^{k-1})$$

definiert.

Ist die  $k$ -te Ableitung von  $f$  stetig auf  $\Omega$ , so sagen wir,  $f$  ist  $k$ -mal stetig differenzierbar auf  $\Omega$ ; die Menge der  $k$ -mal stetig differenzierbaren Funktionen auf  $\Omega$  wird mit  $C^k(\Omega)$  bezeichnet.

Ist  $K \subset \mathbb{R}^m$  kompakt, so definieren wir  $C^k(K)$  als die Menge aller Funktionen  $f$  auf  $K$  für welche es eine offene Menge  $U \supset K$  und eine Funktion  $\tilde{f} \in C^k(U)$  gibt, welche  $\tilde{f}|_K = f$  erfüllt. Die Mengen  $C^k(\Omega)$  und  $C^k(K)$  sind Vektorräume; dies folgt aus folgendem Lemma, welches man wie im eindimensionalen beweist.

LEMMA 109. Sind  $f, g$  im Punkt  $x$  differenzierbar, und  $\lambda \in \mathbb{R}$  so ist auch  $f + \lambda g$  im Punkt  $x$  differenzierbar, und  $(f + \lambda g)'(x) = f'(x) + \lambda g'(x)$ .

Weiters haben wir die Kettenregel:

LEMMA 110. Seien  $f: \mathbb{R}^m \supset U \rightarrow V \subset \mathbb{R}^n$ , und  $g: V \rightarrow \mathbb{R}^\ell$ . Wenn  $f$  im Punkt  $x \in U$  differenzierbar ist, und  $g$  im Punkt  $f(x)$ , so ist  $g \circ f$  im Punkt  $x$  differenzierbar, und  $(g \circ f)'(x) = g'(f(x)) \circ f'(x)$ .

BEWEIS. Es gilt

$$\begin{aligned} g(f(x+h)) - g(f(x)) &= g'(f(x))(f(x+h) - f(x)) + R_1(f(x+h) - f(x)) \\ &= g'(f(x))(f'(x)h) + g'(f(x))(R_2(h)) + R_1(f'(x)h + R_2(h)) \\ &= g'(f(x)) \circ f'(x)(h) + R_3(h). \end{aligned}$$

Das Restglied  $R_3$  erfüllt  $\|h\|^{-1} R_3(h) \rightarrow 0$  für  $h \rightarrow 0$ . □

Die Einträge der Matrix  $(A_{i,j})$ , welche zu  $f'(x)$  gehört, können wie folgt bestimmt werden:

$$A_{i,j} = \langle \mathbf{e}_i, \mathbf{A}\mathbf{e}_j \rangle = \lim_{h \rightarrow 0} \frac{f_i(x + h\mathbf{e}_j) - f_i(x)}{h} =: \frac{\partial f_i}{\partial x_j}(x) = f_{i,x_j}(x).$$

Wenn die partiellen Ableitungen einer Funktion  $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$  auf  $U$  existieren, so stellen sie wiederum Funktionen auf  $U$  dar, und so können iterierte partielle Ableitungen definiert werden. Diese sind, unter sinnvollen Voraussetzungen, von der Reihenfolge der Ableitung unabhängig.

LEMMA 111. Sei  $U: \mathbb{R}^2$  offen, und für  $f: U \rightarrow \mathbb{R}$  seien

$$\frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y \partial x}$$

definiert und stetig im Punkt  $(x^0, y^0) \in U$ . Dann gilt

$$\frac{\partial^2 f}{\partial x \partial y}(x^0, y^0) = \frac{\partial^2 f}{\partial y \partial x}(x^0, y^0).$$

BEWEIS. Sei  $D(h, k) = f(x^0 + h, y^0 + k) - f(x^0 + h, y^0) - f(x^0, y^0 + k) + f(x^0, y^0)$ . Dann ist

$$\begin{aligned} D(h, k) &= f(x^0 + h, y^0 + k) - f(x^0 + h, y^0) - f(x^0, y^0 + k) + f(x^0, y^0) \\ &= (f(x^0 + h, y^0 + k) - f(x^0 + h, y^0)) - (f(x^0, y^0 + k) - f(x^0, y^0)) \\ &= (f_x(\xi^1, y^0 + k) - f_x(\xi^1, y^0))h \\ &= f_{xy}(\xi^1, \eta^1)hk \end{aligned}$$

sowie

$$\begin{aligned} D(h, k) &= f(x^0 + h, y^0 + k) - f(x^0 + h, y^0) - f(x^0, y^0 + k) + f(x^0, y^0) \\ &= f(x^0 + h, y^0 + k) - f(x^0, y^0 + k) - (f(x^0 + h, y^0) - f(x^0, y^0)) \\ &= (f_y(x^0 + h, \eta^2) - f_y(x^0, \eta^2))k \\ &= f_{yx}(\xi^2, \eta^2)hk, \end{aligned}$$

also  $f_{xy}(\xi^1, \eta^1) = f_{yx}(\xi^2, \eta^2)$ , und durch Grenzübergang  $h, k \rightarrow 0$  erhalten wir die Behauptung wegen der Stetigkeit von  $f_{xy}$  und  $f_{yx}$  im Punkt  $(x^0, y^0)$ . □

Wir sehen mit Hilfe des vorangehenden Lemmas, dass die  $k$ -te Ableitung einer  $k$ -mal stetig differenzierbaren Funktion  $f \in C^{(k)}(\Omega)$ , ausgewertet an einer beliebigen Stelle  $x \in \Omega$  sogar eine *symmetrische*  $k$ -lineare Abbildung ist, also

$$f^{(k)}(x)(h^1, \dots, h^k) = f^{(k)}(x)(h^{\sigma(1)}, \dots, h^{\sigma(k)})$$

für jede Permutation  $\sigma: \{1, \dots, k\} \rightarrow \{1, \dots, k\}$  erfüllt. Es gilt nämlich mit Hilfe eines einfachen Induktionsbeweises nach  $k$  für  $h^j = (h_1^j, \dots, h_m^j)$ , dass

$$f^{(k)}(x)(h^1, \dots, h^k) = \sum_{1 \leq \ell_1, \dots, \ell_k \leq n} f_{x_{\ell_1} \dots x_{\ell_k}}(x) h_{\ell_1}^1 \dots h_{\ell_k}^k,$$

wobei die Summe über alle *geordneten*  $k$ -tupel in den "Buchstaben"  $x_1, \dots, x_n$  geht. Lemma 111 zeigt nun, dass die Koeffizienten in diesem Ausdruck symmetrisch unter Vertauschung der Reihenfolge ist, also ist es

auch der Ausdruck selber:

$$\begin{aligned}
 f^{(k)}(x)(h^{\sigma(1)}, \dots, h^{\sigma(k)}) &= \sum_{1 \leq \ell_1, \dots, \ell_k \leq n} f_{x_{\ell_1} \dots x_{\ell_k}}(x) h_{\ell_1}^{\sigma(1)} \dots h_{\ell_k}^{\sigma(k)} \\
 &= \sum_{1 \leq \ell_1, \dots, \ell_k \leq n} f_{x_{\ell_1} \dots x_{\ell_k}}(x) h_{\ell_{\sigma^{-1}(1)}}^1 \dots h_{\ell_{\sigma^{-1}(k)}}^k \\
 &= \sum_{1 \leq \ell_1, \dots, \ell_k \leq n} f_{x_{\ell_{\sigma^{-1}(1)}} \dots x_{\ell_{\sigma^{-1}(k)}}}(x) h_{\ell_{\sigma^{-1}(1)}}^1 \dots h_{\ell_{\sigma^{-1}(k)}}^k \\
 &= f^{(k)}(x)(h^1, \dots, h^k).
 \end{aligned}$$

Die Existenz partieller Ableitungen genügt nicht, um auf Differenzierbarkeit zu schliessen:

**Beispiel 40.** Die Funktion

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & (x, y) \neq 0 \\ 0 & (x, y) = (0, 0) \end{cases}$$

besitzt partielle Ableitungen  $f_x(0, 0) = 0$ ,  $f_y(0, 0) = 0$ , ist aber nicht stetig im Punkt  $(0, 0)$ .

SATZ 59. Sei  $f: \mathbb{R}^m \supset \Omega \rightarrow \mathbb{R}$  so, dass die partiellen Ableitungen

$$f_{x_j}: \Omega \rightarrow \mathbb{R}$$

definiert und stetig auf  $\Omega$  sind. Dann ist  $f \in C^1(\Omega)$ , und es gilt

$$f'(x)(h) = \sum_{k=1}^m f_{x_k}(x) h_k.$$

BEWEIS. Sei  $x \in \Omega$  beliebig. Dann ist

$$\begin{aligned}
 f(x+h) - f(x) &= f(x_1+h_1, \dots, x_m+h_m) - f(x_1, \dots, x_m) \\
 &= \sum_{k=1}^m f(x_1+h_1, \dots, x_k+h_k, x_{k+1}, \dots, x_m) - f(x_1+h_1, \dots, x_{k-1}+h_{k-1}, x_k, \dots, x_m) \\
 &= \sum_{k=1}^m f_{x_k}(x_1+h_1, \dots, \xi_k, x_{k+1}, \dots, x_m) h_k \\
 &= \sum_{k=1}^m f_{x_k}(\zeta_k) h_k
 \end{aligned}$$

mit aus Theorem 37 bestimmten Zwischenpunkten  $\xi_k$ , welche  $|\xi_k - x_k| < h_k$ ,  $k = 1, \dots, m$  erfüllen; wir schreiben in der letzten Zeile

$$\zeta_k = (x_1+h_1, \dots, x_{k-1}+h_{k-1}, \xi_k, x_{k+1}, \dots, x_m).$$

Damit erhalten wir

$$\left| f(x+h) - f(x) - \sum_{k=1}^m f_{x_k}(x) h_k \right| \leq \sum_{k=1}^m |f_{x_k}(x) - f_{x_k}(\zeta_k)| |h_k|.$$

Nachdem die  $f_{x_k}$  stetig sind, gilt damit für das Restglied

$$\begin{aligned}
 \frac{|R(h)|}{\|h\|} &\leq \sum_{k=1}^m |f_{x_k}(x) - f_{x_k}(\zeta_k)| \frac{|h_k|}{\|h\|} \\
 &\leq \max_{k=1, \dots, m} |f_{x_k}(x) - f_{x_k}(\zeta_k)| \sum_k \frac{|h_k|}{\|h\|} \\
 &\leq \max_{k=1, \dots, m} |f_{x_k}(x) - f_{x_k}(\zeta_k)| \rightarrow 0 \quad (h \rightarrow 0).
 \end{aligned}$$

Also ist  $f$  differenzierbar an der Stelle  $x$ , die Ableitung  $f'(x)$  ist durch die behauptete Formel gegeben, und damit nach Voraussetzung stetig.  $\square$

Eine Version des Mittelwertsatzes gilt auch in mehreren Dimensionen:

SATZ 60. Sei  $f: \Omega \rightarrow \mathbb{R}$  eine differenzierbare Funktion,  $x, y \in \Omega$  mit der Eigenschaft, dass die Verbindungsstrecke  $[x, y] = \{ty + (1-t)x: t \in [0, 1]\}$  noch ganz in  $\Omega$  liegt. Dann gibt es ein  $\xi \in (x, y) = \{ty + (1-t)x: t \in (0, 1)\}$  mit der Eigenschaft, dass

$$f(y) - f(x) = f'(\xi)(y - x)$$

gilt.

BEWEIS. Wir wenden Theorem 37 auf die Funktion  $\varphi(t) = f(ty + (1-t)x)$  an. Es ist

$$\varphi'(t) = f'(ty + (1-t)x)(y - x).$$

Damit gibt es ein  $s \in (0, 1)$  sodass  $\xi = sy + (1-s)x$

$$f(y) - f(x) = \varphi(1) - \varphi(0) = \varphi'(s) = f'(\xi)(y - x)$$

erfüllt. □

### 3. Taylorentwicklung

Um die Taylor'sche Formel im mehrdimensionalen anzuwenden, ist es notwendig, Multiindexnotation einzuführen. Für  $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m$  definieren wir

$$|\alpha| = \alpha_1 + \dots + \alpha_m, \quad \alpha! = \alpha_1! \dots \alpha_m!$$

$$\frac{\partial^{|\alpha|} f}{\partial x^\alpha} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \dots \frac{\partial^{\alpha_m}}{\partial x_m^{\alpha_m}} f, \quad x^\alpha = x_1^{\alpha_1} \dots x_m^{\alpha_m}.$$

Das Taylorpolynom  $k$ -ter Ordnung von  $f$  an der Stelle  $x_0$  ist durch

$$T_{x_0}^k f(x) = \sum_{|\alpha| \leq k} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x_0)(x - x_0)^\alpha$$

gegeben; in Bezug auf die Ableitungen von  $f$  kann man auch

$$T_{x_0}^k f(x) = \sum_{j=1}^k \frac{1}{j!} f^{(j)}(x_0) \underbrace{(h, \dots, h)}_{k\text{-mal}}.$$

Nachdem

$$\left. \frac{\partial^{|\alpha|} x^\beta}{\partial x^\alpha} \right|_{x=0} = \begin{cases} \alpha! & \alpha = \beta, \\ 0 & \alpha \neq \beta, \end{cases}$$

gilt, ist für  $f \in C^k(\Omega)$  und  $x_0 \in \Omega$  die Funktion

$$R(x) = f(x) - T_{x_0}^k f(x)$$

selber wieder in  $C^k(\Omega)$  und erfüllt  $R(x_0) = R'(x_0) = \dots = R^{(k)}(x_0) = 0$ .

Aus Theorem 39 folgt nun folgender Satz:

SATZ 61. Sei  $f \in C^k(\Omega)$ , sowie  $f \in D(k+1)\Omega$ . Wenn  $x, y \in \Omega$  mit  $[x, y] \subset \Omega$  gegeben sind, so gibt es ein  $\xi \in (x, y)$  mit

$$f(y) - T_x^k f(y) = \frac{f^{(k+1)}(\xi)}{(k+1)!} \underbrace{(h, \dots, h)}_{k+1\text{-mal}}.$$

Die erste Ableitung  $f'(x)$  ist eine lineare Funktion, welche durch die Matrix

$$(f_{x_1}(x), \dots, f_{x_m}(x)) \in L(\mathbb{R}^m, \mathbb{R})$$

dargestellt wird. Schreiben wir  $f'(x)(h) = \langle v(x), h \rangle$ , wo  $v(x)$  der entsprechende Spaltenvektor ist, so wird

$$v(x) = \begin{pmatrix} f_{x_1}(x) \\ \vdots \\ f_{x_m}(x) \end{pmatrix}$$

der *Gradient* von  $f$  an der Stelle  $x$  genannt und mit  $v(x) = \nabla f(x)$  bezeichnet.

Die zweite Ableitung ist eine quadratische Form, gegeben durch

$$f^{(2)}(x)(h, k) = (h_1, \dots, h_m) \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_m \partial x_m}(x) \end{pmatrix} \begin{pmatrix} k_1 \\ \vdots \\ k_m \end{pmatrix};$$

die Matrix

$$Hf(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_m \partial x_m}(x) \end{pmatrix},$$

welche nach Lemma 111 symmetrisch ist, wird als *Hesse'sche Matrix* (von  $f$  im Punkt  $x$ ) bezeichnet. Wir können also

$$f^{(2)}(x)(h, k) = h^t Hf(x) k = \sum_{p, q=1}^m \frac{\partial^2 f}{\partial x_p \partial x_q}(x) h_p k_q$$

schreiben, und haben folgende kompakte Formel für die Taylorentwicklung bis zum Grad 2:

$$f(x+h) - f(x) = \langle \nabla f(x), h \rangle + \frac{1}{2} h^t (Hf(x)) h + R(h).$$

wo  $R$  die Ungleichung  $|R(h)| \leq \|h\|^3$  erfüllt (wenn  $f$  zumindest  $C^3$  ist).

Eine  $m \times m$ -Matrix  $A$  ist positiv semidefinit, wenn  $h^t A h \geq 0$  für alle  $h \in \mathbb{R}^m$  ist, und positiv definit, wenn  $h^t A h > 0$  für alle  $h \in \mathbb{R}^m$  ist; ähnlich definiert man negativ definit bzw. semidefinit. Wir erhalten damit folgende hinreichende Bedingung zur Existenz eines lokalen Maximums/Minimums.

**SATZ 62.** Sei  $f \in C^2(\Omega)$ , und  $x_0 \in \Omega$  erfülle  $f'(x_0) = 0$ . Dann besitzt  $f$  an der Stelle  $x_0$  ein lokales Maximum, wenn  $Hf(x_0)$  negativ semidefinit ist, und ein lokales Minimum, wenn  $Hf(x_0)$  positiv semidefinit ist. Das Extremum ist strikt, wenn  $Hf(x_0)$  sogar definit ist.

Ähnlich wie im eindimensionalen muss man beachten, dass die Bedingung tatsächlich nur *hinreichend* ist.

**Beispiel 41.** Die Funktion  $f(x, y, z) = x^2 + y^2 + z^2$  erfüllt  $Hf(x, y, z) = 2I$  für alle  $x, y, z$ . Da  $f'(x, y, z)(h_1, h_2, h_3) = 2xh_1 + 2yh_2 + 2zh_3$  ist  $f'(0) = 0$ ;  $f$  hat ein striktes lokales Minimum an der Stelle 0. Für  $g(x, y, z) = x^2 + y^2 + z^4$  ist

$$Hg(x, y, z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2z \end{pmatrix},$$

also  $Hg(0)$  nur positiv semidefinit;  $g$  besitzt aber trotzdem ein striktes Minimum an der Stelle 0.

Allgemeiner gilt, dass die  $k$ -te Ableitung  $f^{(k)}$  einer  $k$ -mal stetig differenzierbaren Funktion durch eine  $k$ -lineare symmetrische Form gegeben ist, welche mit Hilfe der partiellen Ableitungen als

$$f^{(k)}(x)(h^1, \dots, h^k) = \sum_{\substack{1 \leq \ell_j \leq m \\ 1 \leq j \leq k}} f_{x_{\ell_1} \dots x_{\ell_m}} h_{\ell_1}^1 \dots h_{\ell_k}^k, \quad h^j \in \mathbb{R}^m, \quad j = 1, \dots, k$$

gegeben ist. Damit ist

$$f^{(k)}(x)(h, \dots, h) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} f_{x^\alpha}(x) h^\alpha.$$

#### 4. Richtungsableitungen

Die Funktion  $f: \mathbb{R}^m \rightarrow \mathbb{R}$  besitzt eine Richtungsableitung am Punkt  $x$  in Richtung  $v \in \mathbb{R}^m$ , wo  $\|v\| = 1$  ist, wenn die Funktion

$$t \mapsto f(x + tv)$$

am Punkt 0 differenzierbar ist. Der Wert

$$\lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t} = \frac{\partial f}{\partial v}(x)$$

wird die Richtungsableitung von  $f$  in Richtung  $v$  am Punkt  $x$  genannt.

Wenn  $f$  differenzierbar an der Stelle  $x$  ist, so besitzt  $f$  offensichtlich alle Richtungsableitungen, und diese lassen sich mit Hilfe von

$$\frac{\partial f}{\partial v}(x) = f'(x)(v) = \langle \nabla f(x), v \rangle$$

berechnen.

Die Existenz von Richtungsableitungen lässt allerdings nicht auf Differenzierbarkeit schliessen; so hat die Funktion

$$f(x, y) = \sqrt{x^2 + y^2} \psi \left( \frac{(x, y)}{\sqrt{x^2 + y^2}} \right)$$

sämtliche Richtungsableitungen am Punkt 0, da

$$\lim_{t \rightarrow 0} \frac{f(tv)}{t} = \psi(v);$$

aber eine entsprechend unstetige Wahl von  $\psi$  gibt ein nicht differenzierbares  $f$ .

Die Cauchy-Schwarz Ungleichung besagt, dass

$$\left| \frac{\partial f}{\partial v}(x) \right| \leq \|\nabla f(x)\| \|v\| = \|\nabla f(x)\|.$$

Weiters ist

$$\left| \frac{\partial f}{\partial v}(x) \right| < \|\nabla f(x)\|$$

wenn  $v$  nicht in Richtung von  $\nabla f(x)$  ist. In diesem Sinn bezeichnet man den Gradienten von  $f$  als die "Richtung des stärksten Anstiegs" von  $f$ .

#### 5. Implizite Funktionen

Wir schreiben  $(x, y) \in \mathbb{R}^p \times \mathbb{R}^q = \mathbb{R}^m$  für Koordinaten in  $\mathbb{R}^m$ , i.e.

$$x = (x_1, \dots, x_p), \quad y = (y_1, \dots, y_q).$$

Für eine Funktion  $f: \Omega \rightarrow \mathbb{R}^q$ , welche  $f(x_0, y_0) = 0$  erfüllt, wollen wir die Gleichung  $f(x, y) = 0$  nach  $y$  auflösen, also eine Umgebung  $U$  von  $x_0$  bestimmen und eine Funktion  $y: U \rightarrow \mathbb{R}^q$  sodass  $f(x, y(x)) = 0$  auf  $U$  ist.

Die hinreichende Bedingung, welche wir angeben wollen, ist eine an die Ableitung  $f'(x_0, y_0)$ , welche wir wie folgt zerlegen:

$$f'(x_0, y_0)(h, k) = f_x(x_0, y_0)h + f_y(x_0, y_0)k.$$

Wir können nun den Satz über implizite Funktionen formulieren:

**SATZ 63.** Sei  $f \in C^1(\Omega)$ , und an der Stelle  $(x_0, y_0)$  seien die Bedingungen

$$f(x_0, y_0) = 0, \quad \det f_y(x_0, y_0) \neq 0$$

erfüllt. Dann gibt es eine Umgebung  $U$  von  $x_0$  und eine Funktion  $y \in C^1(U)$  mit  $f(x, y(x)) = 0$ , und es gilt

$$y'(x) = -f_y(x, y)^{-1} f_x(x, y).$$

BEWEIS. Wir zeigen zunächst die Existenz einer stetigen Lösungsfunktion. Dazu wollen wir ohne Beschränkung der Allgemeinheit annehmen, dass

$$x_0 = y_0 = 0, \quad f_y(0, 0) = I$$

ist. Wir betrachten die Abbildung

$$A: \varphi \mapsto (x \mapsto \varphi(x) - f(x, \varphi(x)))$$

auf einem geeigneten Raum von Funktionen  $\varphi$ . Wir wählen dazu zunächst Umgebungen  $U = B_{\delta_1}(0)$  von 0 in  $\mathbb{R}^p$  und  $V = B_{\delta_2}(0)$  von 0 in  $\mathbb{R}^q$  sodass

$$\|(I - f_y(x, y))k\| \leq \frac{1}{4} \|k\|$$

für  $(x, y) \in \overline{U} \times \overline{V}$  und  $k \in \mathbb{R}^q$  ist. Dann ist für  $y \in \overline{V}$

$$y - f(x, y) = \int_0^1 (I - f_y(x, ty))y \, dt$$

und nach Anwendung der Dreiecksungleichung

$$\|y - f(x, y)\| \leq \int_0^1 \frac{1}{4} \|y\| \, dt = \frac{\|y\|}{4}.$$

Weiters gilt für beliebige  $y_1, y_2 \in \overline{V}$  dass

$$\begin{aligned} \|y_2 - f(x, y_2) - (y_1 - f(x, y_1))\| &= \left\| ty_2 + (1-t)y_1 - f(x, ty_2 + (1-t)y_1) \Big|_{t=0}^{t=1} \right\| \\ &= \left\| \int_0^1 (I - f_y(x, ty_2 + (1-t)y_1))(y_2 - y_1) \, dt \right\| \\ &\leq \int_0^1 \|(I - f_y(x, ty_2 + (1-t)y_1))(y_2 - y_1)\| \, dt \\ &\leq \frac{\|y_2 - y_1\|}{4}. \end{aligned}$$

Damit ist die Abbildung  $A$  auf dem Raum stetiger Funktionen auf der kompakten Menge  $\overline{U}$  mit Werten in  $\overline{V}$  definiert, also auf dem Raum

$$X = \{\varphi \in C(\overline{U})^q: \|\varphi\|_\infty \leq \delta_2\},$$

wobei

$$\|\varphi\|_\infty = \max_{x \in \overline{U}} \|\varphi(x)\|,$$

da für beliebiges  $x \in U$

$$\|A(\varphi)(x)\| = \|\varphi(x) - f(x, \varphi(x))\| \leq \frac{1}{4} \|\varphi(x)\| \leq \frac{1}{4} \|\varphi\|_\infty$$

ist. Wir definieren eine Metrik  $d$  auf  $X$  durch

$$d(\varphi_1, \varphi_2) = \|\varphi_2(x) - \varphi_1(x)\|_\infty,$$

diese macht  $X$  zu einem vollständigen metrischen Raum.

Weiters gilt für beliebige  $\varphi_1, \varphi_2 \in B_\varepsilon(0)$  die Ungleichung

$$\begin{aligned} \|A(\varphi_1)(x) - A(\varphi_2)(x)\| &= \|\varphi_1(x) - f(x, \varphi_1(x)) - (\varphi_2(x) - f(x, \varphi_2(x)))\| \\ &= \left\| t\varphi_1(x) + (1-t)\varphi_2(x) - f(x, t\varphi_1(x) + (1-t)\varphi_2(x)) \Big|_{t=0}^{t=1} \right\| \\ &= \left\| \int_0^1 (I - f_y(x, t\varphi_1(x) + (1-t)\varphi_2(x))) (\varphi_2(x) - \varphi_1(x)) \, dt \right\| \\ &\leq \int_0^1 \|(I - f_y(x, t\varphi_1(x) + (1-t)\varphi_2(x))) (\varphi_2(x) - \varphi_1(x))\| \, dt \\ &\leq \frac{\|\varphi_2(x) - \varphi_1(x)\|}{4}, \end{aligned}$$



solange Also ist die Abbildung  $A: B_\varepsilon(0) \rightarrow B_\varepsilon(0)$  eine Kontraktion und besitzt nach Theorem 31 einen Fixpunkt  $y(x)$  ( $B_\varepsilon(0)$  ist als abgeschlossene Teilmenge eines vollständigen metrischen Raums wieder ein vollständiger metrischer Raum). Diese Funktion  $y(x)$  erfüllt  $A(y)(x) = y(x) - f(x, y(x)) = y(x)$  für alle  $x \in \bar{U}$ , also  $f(x, y(x)) = 0$ .

Wir zeigen nun unter der Annahme, dass  $f \in C^1(U \times V)$ , dass  $y$  an einer beliebigen Stelle  $x_0$  (stetig) differenzierbar ist, mit Ableitung

$$y'(x_0) = -f_y(x_0, y(x_0))^{-1} f_x(x_0, y(x_0)).$$

(die Stetigkeit folgt aus der Formel für die Ableitung, die wir angegeben haben).

Wir können wieder annehmen, dass  $x_0 = 0$ , und  $y(0) = 0$ . Dann liegt für kleines  $h \in \mathbb{R}^p$  die Verbindungslinie von  $(0, 0)$  nach  $(h, y(h))$  ganz in  $U \times V$ . Dann ist

$$0 = f_j(h, y(h)) - f_j(0, 0) = f_{j,x}(\xi^j)h + f_{j,y}(\xi^j)y(h)$$

mit einem  $\xi^j \in ((0, 0), (h, y(h)))$ , für  $1 \leq j \leq q$ .

Mit den Matrizen

$$f_x(\xi) = \begin{pmatrix} f_{1,x}(\xi^1) \\ \vdots \\ f_{q,x}(\xi^1) \end{pmatrix}, \quad f_y(\xi) = \begin{pmatrix} f_{1,y}(\xi^1) \\ \vdots \\ f_{q,y}(\xi^1) \end{pmatrix}$$

gilt also  $y(h) = -f_y(\xi)^{-1} f_x(\xi)h$ . Wir können also das Restglied  $R(h)$  in

$$y(h) - y(0) - y'(0)h + R(h)$$

wie folgt anschreiben:

$$R(h) = y(h) + f_y(0, 0)^{-1} f_x(0, 0)h = (f_y(0, 0)^{-1} f_x(0, 0) - f_y(\xi)^{-1} f_x(\xi))h = B(\xi)h.$$

Da  $B(0) = 0$ , und  $B$  stetig auf  $\mathbb{R}^{mq}$  ist, gibt es für jedes  $\varepsilon > 0$  ein  $\delta > 0$  sodass für jedes  $h$  mit  $\|h\| < \delta$  die Ungleichung  $\|B(\xi)h\| \leq \varepsilon \|h\|$  ist. Für diese  $h$  ist also

$$\frac{\|R(h)\|}{\|h\|} < \varepsilon,$$

also ist  $y$  differenzierbar und die Ableitung durch die Formel

$$y'(x) = -f_y(x, y(x))^{-1} f_x(x, y(x))$$

gegeben, was (wie schon oben erwähnt) auch zeigt, dass  $y'$  stetig ist. □

## 6. Lagrange Multiplikatoren

Sei  $\varphi: W \rightarrow \mathbb{R}^d$  eine differenzierbare Funktion. Wir sind daran interessiert, folgendes Extremwertproblem zu lösen: Finde  $z \in W$ , für welche  $\varphi(z)$  unter der Nebenbedingung  $f(z) = 0$  ein Extremum hat.

Angenommen, die Matrix  $f_z$  hat konstanten Rang  $d$  auf  $W$ , dann können wir nach Ummumerierung der Koordinaten  $W = U \times V$  mit  $U \subset \mathbb{R}^{m-d}$ ,  $V \subset \mathbb{R}^d$  schreiben, und mit entsprechenden Koordinaten  $(x, y) = z$  annehmen, dass  $f_y$  invertierbar auf  $W$  ist. Wir haben dann folgendes Problem: Sei  $\varphi: U \times V \rightarrow \mathbb{R}^d$  eine differenzierbare Funktion. Finde  $(x, y) \in U \times V$ , für welche  $\varphi(x, y)$  unter der Nebenbedingung  $f(x, y) = 0$  ein Extremum hat.

Wenn, wie angenommen,  $f_y(x, y)$  invertierbar ist, so können wir in einer Umgebung von  $(x_0, y_0) \in U \times V$  eine Umgebung  $W$  von  $x_0$  und eine differenzierbare Funktion  $y(x)$  finden, sodass die Lösungen  $f(x, y) = 0$  mit  $x \in W$  und  $y$  nahe bei  $y_0$  von der Form  $y = y(x)$  sind.

Damit übersetzt sich die Lösung des Extremwertproblems in das Finden der Extremwerte von

$$\varphi(x, y(x))$$

übersetzen. Damit muss an der Stelle eines Extremums

$$\varphi_x(x, y(x)) + \varphi_y(x, y(x))y'(x) = \varphi_x(x, y(x)) + \varphi_y(x, y(x))(-f_y(x, y(x)))^{-1} f_x(x, y(x)) = 0$$

sein. Mit anderen Worten: An der Stelle eines Extremums erfüllt  $(\varphi_x, \varphi_y)$  die Gleichung

$$(\varphi_x, \varphi_y) \begin{pmatrix} I \\ -f_y^{-1} f_x \end{pmatrix} = 0;$$

die Matrix ist von Rang  $d$ , und  $(\varphi_x, \varphi_y)$  ist in ihrem (Links-)Nullraum. Man überprüft auch, dass die Vektoren  $(f_{j,x}, f_{j,y})$  für  $j = 1, \dots, d$  in ihrem Nullraum sind, und sie sind nach Voraussetzung linear unabhängig. Damit muss es also Konstanten  $\lambda_1, \dots, \lambda_d$  geben, welche  $\varphi' = \sum_j \lambda_j f'_j$  erfüllen.

Wir betrachten an Stelle von  $\varphi$  nun die Funktion

$$\Phi(z, \Lambda) = \varphi(z) - \sum_{k=1}^d \lambda_k f_k(z) = \varphi(z) - f(z) \cdot \Lambda,$$

welche auf  $W \times \mathbb{R}^d$  definiert ist, und wir  $\Lambda = (\lambda_1, \dots, \lambda_d)$  geschrieben haben. Dann ist  $\Phi'(z, \Lambda)$  durch folgende Matrix gegeben:

$$\Phi' = (\varphi_z - f_z \Lambda, f).$$

Also ist  $\Phi'(z, \Lambda) = 0$ , wenn  $f(z) = 0$  ist und es Konstanten  $\Lambda = (\lambda_1, \dots, \lambda_d)$  gibt, welche

$$\varphi' = \sum_j \lambda_j f'_j$$

erfüllen. An einer Stelle, wo  $\varphi$  ein Extremum mit Nebenbedingung  $f = 0$  hat, gilt also, dass die zugeordnete Funktion  $\Phi$  eine Nullstelle ihrer Ableitung hat, also  $\Phi' = 0$  ist.

**Beispiel 42.** Wir finden das Extremum von  $\varphi(x, y) = xy$  unter der Nebenbedingung  $x + y = 1$ . Elementar berechnen wir das, indem wir die Funktion  $\psi(x) = x(1 - x)$  betrachten und sehen, dass  $\psi$  ein Maximum an der Stelle  $x = 1/2$  annimmt.

Ohne diese Parametrisierung zu verwenden, können wir die Funktion  $\Phi(x, y, \lambda) = xy + \lambda(x + y - 1)$  betrachten. Dann müssen wir die Nullstellen von  $\Phi'$  bestimmen, also die Gleichungen

$$y + \lambda = 0, \quad x + \lambda = 0, \quad x + y = 1$$

lösen. Dieses lineare Gleichungssystem hat die Lösungen  $x = y = 1/2$ ,  $\lambda = -1/2$ .

Wir formulieren nochmals:

**SATZ 64.** Wenn die Funktion  $\varphi: U \rightarrow \mathbb{R}$  ein Extremum an der Stelle  $z_0$  unter der Nebenbedingung  $f(z) = 0$  hat, wo  $f: U \rightarrow \mathbb{R}^d$  von vollen Rang auf  $U$  ist (d.h. die Matrix  $f'(z)$  hat Rang  $d$  für alle  $z \in U$ ), dann gibt es ein  $\Lambda_0 \in \mathbb{R}^d$ , sodass die Funktion

$$\Phi(z, \Lambda) = \varphi(z) + \Lambda \cdot f(z)$$

$\Phi'(z_0, \Lambda_0) = 0$  erfüllt.