Analysis 1 und 2

Dr. Maximilian Hanusch*

Institut für Mathematik
Universität Paderborn
Warburger Straße 100
33098 Paderborn
Deutschland

3. September 2023

Inhaltsverzeichnis

*`mhanusch@math.upb.de`

Vorwort

Das vorliegende Skript richtet sich strukturell nach einer Skriptvorlage von Herrn Prof. Dr. Karl Hermann Neeb (aktuell nicht öffentlich verfügbar), welche dem Autor dankenswerterweise im Rohformat (.tex file) zur Verfügung gestellt wurde. Einige Abschnitte sind bis auf leichte Anpassungen und Umformulierungen deckungsgleich, andere selbst konzipiert, und wiederum andere Teile sind Erweiterungen entsprechender Abschnitte in besagtem Skript. In indirekter Form sind ebenfalls Inhalte aus dem Analysis 2 Skript von Herrn Prof. Dr. Helge Glöckner eingeflossen (dies betrifft die Teile ab Kapitel 10). Die Abschnitte 10.3.1 und 10.3.3 sind (inhaltlich) einem Zusatzmaterial aus der Analysis 1 Vorlesung im WS 18/19 von Herrn Glöckner entnommen. Abschnitt 4.4.2 ist dem Analysis 1 Skript von Herrn Prof. Dr. Wolfgang Soergel entlehnt.

- Die mit „$*$“ markierten Teile (Abschnitte/Bemerkungen/Beweise etc.) sind als ergänzendes Material gedacht, und sind daher nicht Teil des Stoffes der Vorlesung.
- Die mit „$(*)$“ markierten Abschnitte wurden in der Vorlesung nur in verkürzter Form behandelt.

1 Einführung

Das Kernobjekt der Analysis sind die reellen Zahlen $\mathbb{R}$ und die hierauf definierten Funktionen (wie z.B. der Sinus, der Kosinus, und die Exponentialfunktion). Bevor wir uns diesen Begrifflichkeiten zuwenden, wollen wir zunächst die Logik und die Mengenlehre in den Grundzügen diskutieren. Die logischen Grundlagen werden in Kapitel 2 behandelt. Mengen und Abbildung werden in Kapitel 3 genauer diskutiert.

Dem Begriff der Menge stellen wir uns naiv gegenüber, d.h., wir stellen uns auf den Standpunkt, dass wir eine Menge kennen, wenn uns gesagt wird, welche Elemente sie enthält.

Definition 1 (Georg Cantor (1845-1918))**.** *Unter einer Menge verstehen wir eine Zusammenfassung gewisser wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens zu einem Ganzen. Die in einer Menge zusammengefassten Objekte werden als die Elemente dieser Menge bezeichnet.*

Bemerkung 1. *Obige Definition ist eigentlich keine Definition im engeren mathematischen Sinne, da der zu definierende Begriff Menge nicht auf andere (bereits definierte Begriffe) zurückgeführt wird. Es handelt sich um eine Beschreibung, welcher Sachverhalt vorliegen muss, damit von einer Menge gesprochen werden kann. Geht man mit obigem Mengenbegriff nicht sorgsam genug um, dann kommt es zu Widersprüchen – z.B. wenn man die Menge aller Mengen betrachtet (welche nicht existieren kann, siehe Beispiel 1 und Übung 5). Das Auftreten derartiger Paradoxien kann durch den Übergang zu einer axiomatisierten[1] Mengenlehre (der Zermelo-Fraenkel-Mengenlehre) vermieden werden. Diese im Detail abzuhandeln ist aufwendig, und normalerweise Gegenstand einer eigenständigen Logikvorlesung. Daher bleiben wir im Wesentlichen bei der (mehr als ausreichenden) Definition Cantors, wobei die Objekte der Anschauung oder des Denkens meist Zahlenmengen, Mengen von Symbolfolgen, oder Mengen von Funktionen sind.*

Sei X eine Menge und x ein Objekt. Man schreibt

- $x \in X$, wenn x ein Element der Menge X ist (wenn x in X enthalten ist).
- $x \notin X$, wenn x kein Element der Menge X ist (wenn x nicht in X enthalten ist).

[1] Im allgemeinen versteht man unter einem Axiom einen als wahr angenommenen Grundsatz einer Theorie, der innerhalb dieser Theorie weder begründet noch abgeleitet werden kann. Beispiele aus der Physik (in welcher Axiome oft auch als Postulate bezeichnet werden) sind die Newtonschen Axiome der Mechanik, die Maxwell-Gleichungen der Elektrodynamik, die Einsteinschen Feldgleichungen der allgemeinen Relativitätstheorie, oder die Schrödingergleichung der Quantenmechanik (wobei jede dieser Theorien in einen entsprechenden mathematischen Rahmen eingebettet ist).

Gegeben Mengen X, Y, so schreibt man

- $X \subseteq Y$, wenn jedes Element von X auch ein Element von Y ist. X wird in diesem Fall als Teilmenge von Y bezeichnet.
- $X \subset Y$, wenn jedes Element von X auch ein Element von Y ist, aber Y mindestens ein zusätzliches Element enthält, welches nicht in X enthalten ist. X wird in diesem Fall als echte Teilmenge von Y bezeichnet.
- $X = Y$, wenn X und Y die selben Elemente haben. In diesem Fall sagt man, dass die Mengen X und Y gleich sind.

Eine Menge heißt endlich, wenn sie nur endlich viele Elemente hat. Die Anzahl der Elemente einer endlichen Menge X nennen wir ihre *Kardinalität* oder *Mächtigkeit*, und notieren diese mit $|X|$. Endliche Mengen kann man durch eine vollständige Liste ihrer Elemente in geschweiften Klammern (Mengenklammern) angeben – zum Beispiel in der Form

$$X = \{x_1, x_2, \dots, x_n\} \qquad \text{mit} \qquad n \geq 1.$$

Die Elemente dürfen mehrfach genannt werden, und es kommt nicht auf die Reihenfolge an, in der diese genannt werden. Wir haben also $\{1,1,2\} = \{2,1\}$, und beispielsweise $1 \in \{2,1\}$ und $3 \notin \{2,1\}$. Die leere Menge $\emptyset = \{\}$ ist die Menge, die gar kein Element enthält (ihre Existenz wird in der axiomatischen Mengenlehre als Axiom vorausgesetzt). Sie ist Teilmenge jeder Menge (es gilt also $\emptyset \subseteq X$ für jede Menge X). Wir vereinbaren, dass die leere Menge endlich ist, mit $|\emptyset| = 0$. Mit dieser Konvention ist jede Teilmenge einer endlichen Menge wieder endlich. Man kann auch unendliche Mengen mit Hilfe von Mengenklammern angeben, wie z.B.

$$\begin{aligned}
\mathbb{N} &= \{0, 1, 2, 3, \dots\} && \text{die Menge der natürlichen Zahlen} \\
\mathbb{N}_{>0} &= \{1, 2, 3, \dots\} && \text{die Menge der natürlichen Zahlen ohne die 0} \\
\mathbb{Z} &= \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\} && \text{die Menge der ganzen Zahlen,}
\end{aligned}$$

welche Sie ja bereits aus der Schule kennen. Die natürlichen Zahlen $\mathbb{N}$ können durch die sogenannten Peano-Axiome charakterisiert werden (inklusive der Rechenoperationen). Teil dieser Charakterisierung (bzw. unmittelbare Folge dieser) ist das *Wohlordnungsprinzip*, welches den (offensichtlichen) Umstand zum Ausdruck bringt, dass jede nichtleere Teilmenge $M \subseteq \mathbb{N}$ ein kleinstes Element besitzt (z.B. besitzt $M = \{4, 8, 9, 10\}$ das kleinste Element 4). Im Rahmen der Zermelo-Fraenkel-Mengenlehre (ZF) erhält man die Menge der natürlichen Zahlen (und dann im Prinzip auch $\mathbb{Z}$) vermöge der folgenden iterativen Konstruktion:

$$\begin{aligned}
0 &:= \{\} = \emptyset \\
1 &:= \{0\} = \{\emptyset\} \\
2 &:= \{0, 1\} = \{\emptyset, \{\emptyset\}\} \\
3 &:= \{0, 1, 2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \\
&\vdots \\
n + 1 &:= \{0, 1, 2, \dots, n\}.
\end{aligned}$$

Dass dieses induktive Vorgehen eine (unendliche) Menge definiert, sichert das sogenannte Unendlichkeitsaxiom von ZF.

Das Symbol „:=“ bedeutet hier (und sinngemäß auch im Folgenden), dass die linke Seite jeweils durch die rechte Seite definiert (festgelegt) wird.

Mengen können auch durch Ausdrücke der Form

$$\{x \mid E(x)\} \tag{1}$$

definiert werden, wobei E eine Eigenschaft bezeichnet, die ein Objekt (eines gewissen Grundbereiches) entweder besitzt oder nicht besitzt. Beispielsweise sind die rationalen Zahlen gegeben durch

$$\mathbb{Q} = \left\{ \frac{p}{q} \,\middle|\, p \in \mathbb{Z} \ \text{ und } \ q \in \mathbb{N}_{>0} \right\}, \tag{2}$$

wobei hier allerdings rationale Zahlen mehrfach aufgeführt werden: Es beschreiben p/q und p'/q' die selbe rationale Zahl (den selben Anteil eines Ganzen) genau dann, wenn diese Brüche durch kürzen ineinander übergehen, also $pq' = qp'$ gilt.[2] Wir werden die rationalen Zahlen in Abschnitt 3.2.1 (Beispiel 9) noch formeller definieren, sobald wir die nötigen mengentheoretischen Operationen zur Verfügung haben. Die reellen Zahlen $\mathbb{R}$ werden in Abschnitt 4.3 konstruiert; und die komplexen Zahlen $\mathbb{C}$ werden in Kapitel 4.5 definiert. Wir haben die Inklusionen

$$\mathbb{N}_{>0} \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Es sei an dieser Stelle darauf hingewiesen, dass der allzu sorglose Umgang mit Mengendefinitionen der Form (1) zu Widersprüchen führen kann.

Beispiel 1. *Wir betrachten die Menge $R := \{X \mid X \notin X\}$ (die Russelsche Unmenge) – also die Menge aller Mengen, die sich nicht selbst als Element enthalten. Es gibt die beiden Möglichkeiten $R \in R$ und $R \notin R$, welche aber einander selbst widersprechen:*

- *Ist $R \in R$, so gilt $R \notin R$ nach der Definition von R.*
- *Ist $R \notin R$, so gilt $R \in R$ nach der Definition von R.*

Um das nächste Kapitel mit Beispielen füllen zu können, benötigen wir noch die folgende Begrifflichkeit:

Definition 2. *Eine ganze Zahl $n \in \mathbb{Z}$ heißt gerade genau dann, wenn sie durch 2 teilbar ist; wenn also ein $m \in \mathbb{N}$ existiert, sodass $n = 2m$ gilt. Eine ganze Zahl $n \in \mathbb{Z}$ heißt ungerade genau dann, wenn sie nicht gerade ist.*

Die Menge der ganzen Zahlen zerfällt daher in die zwei Teilmengen

$$\begin{aligned} \{\dots, -6, -4, -2, 0, 2, 4, 6, \dots\} &= \{n \in \mathbb{Z} \mid n = 2m \text{ für ein } m \in \mathbb{Z}\} && \text{der geraden Zahlen,} \\ \{\dots, -5, -3, -1, 0, 1, 3, 5, \dots\} &= \{n \in \mathbb{Z} \mid n = 2m+1 \text{ für ein } m \in \mathbb{Z}\} && \text{der ungeraden Zahlen.} \end{aligned}$$

2 Logik und Beweistechniken

2.1 Aussagenlogik

Eine Aussage ist eine sprachliches bzw. in mathematischen Symbolen formuliertes Satzgebilde, dem in eindeutiger Weise einer der Wahrheitswerte „wahr“ oder „falsch“ zugeordnet werden kann.

Bemerkung 2. *Eine Aussage ist also entweder wahr oder falsch. Eine dritte Möglichkeit gibt es nicht – Dies wird als das Prinzip des ausgeschlossenen Dritten bezeichnet.*

Man sagt, dass eine Aussage gilt/nicht gilt, wenn sie wahr/falsch ist. Beispiele für Aussagen sind:

- $2 + 2 = 4$.
- Jetzt regnet es.
- $1 + 1$ ergibt 3.

[2] Es ist $pq' = qp'$ gleichbedeutend mit $\frac{p}{q} - \frac{p'}{q'} = \frac{pq' - qp'}{qq'} = 0$.

Bemerkung 3.

- *Der Wahrheitswert einer Aussage kann Situationsabhängig sein (z.B. von der Erfülltheit gewisser Voraussetzungen abhängen). So sind zum Beispiel Aussagen gewisser mathematischer Sätze nur unter Zusatzvoraussetzungen wahr (die im Vorfeld natürlich explizit angegeben werden müssen). Ein weiteres Beispiel ist die Aussage „Dieser Planet ist blau.“, deren Wahrheitsgehalt natürlich davon abhängt, auf welchem Planeten man sich gerade befindet.*
- *Essentiell bei einer Aussage ist natürlich deren Sinn. D.h., zwei Aussagen A und B sind „gleich“, wenn sie durch sinnerhaltende sprachliche (mathematisch symbolische) Umformungen ineinander überführt werden können. In diesem Falle schreiben wir $A \equiv B$.*[3]

Im Folgenden seien A und B Aussagen. Es können die folgenden neuen Aussagen gebildet werden:

a) Die Negation $\neg A$ (nicht A) der Aussage A ist die Aussage „Es gilt nicht A.“ Die Aussage $\neg A$ ist genau dann wahr/falsch, wenn die Aussage A falsch/wahr ist. Die zugehörige Wahrheitstafel ist:

A	$\neg A$
w	f
f	w

Man erhält durch sinnerhaltende Umformulierungen

$$\begin{aligned} \neg(\neg A) &\equiv \text{„Es gilt nicht, dass die Aussage A nicht gilt.“} \\ &\equiv \text{„Die Aussage A gilt.“} \\ &\equiv A, \end{aligned}$$

was sich in der zugehörigen Wahrheitstafel widerspiegelt:

A	$\neg A$	$\neg(\neg A)$
w	f	w
f	w	f

b) Die Konjunktion $A \wedge B$ (A und B) der Aussagen A und B ist die Aussage „A und B sind beide wahr.“ Daher ist die Aussage $A \wedge B$ genau dann wahr, wenn die Aussagen A und B beide wahr sind. Die zugehörige Wahrheitstafel ist:

A	B	$A \wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

c) Die Disjunktion $A \vee B$ (A oder B) der Aussagen A und B ist die Aussage „Mindestens eine der beiden Aussagen A oder B ist wahr.“ Die zugehörige Wahrheitstafel ist:

A	B	$A \vee B$
w	w	w
w	f	w
f	w	w
f	f	f

Terminologie 1.

- *Eine Aussage, die immer den Wert wahr annimmt, bezeichnet man als Tautologie. Z.B. ist $A \vee \neg A$ für jede Aussage A eine Tautologie („Es regnet oder es regnet nicht.“).*

[3]Das Symbol $\equiv$ wird im Rahmen von Aussagen nur in diesem Kapitel verwendet werden.

- *Eine Aussage, die immer den Wert falsch annimmt, bezeichnet man als Widerspruch. Z.B. ist $A \wedge \neg A$ für jede Aussage A ein Widerspruch („Es regnet und es regnet nicht.“).*

Anwendung sinnerhaltender sprachlicher Umformulierungen liefert die De Morganschen Regeln

$$\neg(A \wedge B) \equiv \neg A \vee \neg B \qquad \text{sowie} \qquad \neg(A \vee B) \equiv \neg A \wedge \neg B. \tag{3}$$

In der Tat erhalten wir

$$\begin{aligned} \neg(A \wedge B) &\equiv \text{„Es gilt nicht, dass die Aussagen } A \text{ und } B \text{ beide wahr sind.“} \\ &\equiv \text{„Mindestens eine der Aussagen } A \text{ oder } B \text{ ist falsch.“} \\ &\equiv \text{„Mindestens eine der Aussagen } \neg A \text{ oder } \neg B \text{ ist wahr.“} \\ &\equiv \neg A \vee \neg B \end{aligned}$$

sowie

$$\begin{aligned} \neg(A \vee B) &\equiv \text{„Es gilt nicht, dass mindestens eine der Aussagen } A \text{ oder } B \text{ wahr ist.“} \\ &\equiv \text{„Beide Aussagen } A \text{ und } B \text{ sind falsch.“} \\ &\equiv \text{„Beide Aussagen } \neg A \text{ und } \neg B \text{ sind wahr.“} \\ &\equiv \neg A \wedge \neg B. \end{aligned}$$

Dieser Sachverhalt spiegelt sich auch in folgender Wahrheitstafel wieder, die man leicht nachprüft:

A	B	$A \wedge B$	$A \vee B$	$\neg A$	$\neg B$	$\neg(A \wedge B)$	$\neg A \vee \neg B$	$\neg(A \vee B)$	$\neg A \wedge \neg B$
w	w	w	w	f	f	f	f	f	f
w	f	f	w	f	w	w	w	f	f
f	w	f	w	w	f	w	w	f	f
f	f	f	f	w	w	w	w	w	w

(4)

Die Implikation $A \Rightarrow B$ ist die Aussage, die definiert ist durch

$$A \Rightarrow B := \neg A \vee B \overset{(3)}{\equiv} \neg(A \wedge \neg B) \tag{5}$$

mit

$$A \wedge \neg B \equiv \text{„A ist wahr und B ist falsch.“} \tag{6}$$

$$\text{also} \qquad A \Rightarrow B \equiv \text{„Es gilt nicht: A ist wahr und B ist falsch.“} \tag{7}$$

Die zugehörige Wahrheitstafel ist

A	B	$A \Rightarrow B$	
w	w	w	**(1)**
w	f	f	**(2)**
f	w	w	**(3)**
f	f	w	**(4)**

(8)

Terminologie 2.

- *Im mathematischen Sprachgebrauch sagt man, dass eine Implikation $A \Rightarrow B$ gilt, wenn aus der Wahrheit der Aussage A die Wahrheit der Aussage B folgt.*

- *Angenommen, die Implikation $A \Rightarrow B$ gilt. Dann ist die Aussage (Implikation) $A \Rightarrow B$ immer wahr – nämlich wegen den Zeilen* **(3)** *und* **(4) in** (8) *auch dann, wenn die Aussage A falsch ist.*
- *Angenommen, sowohl A als auch die Aussage (Implikation) $A \Rightarrow B$ ist wahr. Laut Wahrheitstafel muss dann auch die Aussage B wahr sein (Zeile* **1***).*

Es sei explizit darauf hingewiesen, dass die Aussage (Implikation) $A \Rightarrow B$ natürlich auch wahr sein kann, wenn die Wahrheit der Aussage A nicht die Wahrheit der Aussage B erzwingt. A und B könnten zufällig immer gleichzeitig wahr sein, z.B. wenn irgendeine Aussage gilt, die sowohl A als auch B impliziert. Weiterhin könnten A und B beides Tautologien sein, oder A ein Widerspruch. Es ist daher wichtig, zwischen dem Terminus „Eine Implikation gilt." und dem Terminus „Eine Implikation ist wahr." strikt zu unterscheiden.

- *Eine Implikation $A \Rightarrow B$ zu zeigen, bedeutet (im mathematischen Sprachgebrauch) nachzuweisen, dass aus der Wahrheit der Aussage A die Wahrheit der Aussage B folgt.*

  In der mathematischen Praxis wird A oft auch als Voraussetzung formuliert sein, die dann einfach als wahre Aussage interpretiert wird.

- *Abweichend von der zweiten Zeile in* (6)*, spricht man „$A \Rightarrow B$" in der mathematischen Umgangssprache oft auch wie folgt aus:*

  „A impliziert B.",
  „Aus A folgt B."
  oder „Wenn A gilt, dann gilt auch B."

  Zeile **(2)** *wird auch intepretiert als „Aus Wahrem folgt nichts Falsches."; und die Zeilen* **(3)** *und* **(4)** *als „Aus Falschem folgt alles Mögliche.".*

Übung 1. *Seien X, Y, Z Mengen. Zeigen Sie die folgenden Implikationen:*

$$\begin{aligned} X \subseteq Y \ \wedge \ Y \subseteq Z \quad &\Rightarrow \quad X \subseteq Z \\ X \subset Y \ \wedge \ Y \subseteq Z \quad &\Rightarrow \quad X \subset Z \\ X \subseteq Y \ \wedge \ Y \subset Z \quad &\Rightarrow \quad X \subset Z. \end{aligned}$$

Die Äquivalenz $A \Leftrightarrow B$ ist die Aussage, die definiert ist durch

$$A \Leftrightarrow B := (A \Rightarrow B) \wedge (B \Rightarrow A).$$

Die zugehörige Wahrheitstafel ist

A	B	$A \Leftrightarrow B$
w	w	w
w	f	f
f	w	f
f	f	w

(9)

Die Aussage $A \Leftrightarrow B$ ist also genau dann wahr, wenn A und B beide wahr oder beide falsch sind.

Terminologie 3.

- *Im mathematischen Sprachgebrauch sagt man, dass eine Äquivalenz $A \Leftrightarrow B$ gilt, wenn die Implikationen $A \Rightarrow B$ und $B \Rightarrow A$ gelten.*

- *Angenommen, die Äquivalenz $A \Leftrightarrow B$ gilt. Dann gilt auch, dass A genau dann falsch ist, wenn B falsch ist. Daher ist in diesem Falle die Aussage (Äquivalenz) $A \Leftrightarrow B$ wegen obiger Wahrheitstafel immer wahr.*
- *Angenommen, die Aussage (Äquivalenz) $A \Leftrightarrow B$ ist wahr. Laut Wahrheitstafel ist dann A/B genau dann wahr, wenn B/A wahr ist.*

- *Eine Äquivalenz $A \Leftrightarrow B$ zu zeigen, bedeutet (im mathematischen Sprachgebrauch) die Implikationen $A \Rightarrow B$ und $B \Rightarrow A$ zu zeigen.*

Natürlich gilt $\tilde{A} \Leftrightarrow \tilde{B}$, wenn $\tilde{A}$ und $\tilde{B}$ die gleichen Aussagen sind. In Formeln gilt:[4]

$$\tilde{A} \equiv \tilde{B} \quad \Rightarrow \quad \text{(es gilt) } \tilde{A} \Leftrightarrow \tilde{B}. \tag{10}$$

Aus (3) und dem allgemeinen Umstand (10) folgt beispielsweise, dass gilt:

$$\neg(A \wedge B) \quad \Leftrightarrow \quad \neg A \vee \neg B \qquad \text{sowie} \qquad \neg(A \vee B) \quad \Leftrightarrow \quad \neg A \wedge \neg B,$$

was allerdings auch unmittelbar aus der Wahrheitstafel (4) ersichtlich ist.

Ein weiteres wichtiges Beispiel für Äquivalenzen von Aussagen ist die Kontraposition $\neg B \Rightarrow \neg A$ einer vorgegebenen Implikation $A \Rightarrow B$. Wir rechnen[5]

$$A \Rightarrow B \overset{(5)}{=} \neg A \vee B \equiv B \vee \neg A \equiv \neg(\neg B) \vee \neg A \overset{(5)}{=} \neg B \Rightarrow \neg A; \tag{11}$$

und erhalten wieder wegen (10)

$$A \Rightarrow B \quad \Leftrightarrow \quad \neg B \Rightarrow \neg A, \tag{12}$$

was man anhand von Wahrheitstafeln ebenfalls leicht nachrechnet:

A	B	$\neg A$	$\neg B$	$A \Rightarrow B$	$\neg B \Rightarrow \neg A$	
w	w	f	f	w	w	**(1)**
w	f	f	w	f	f	**(2)**
f	w	w	f	w	w	**(3)**
f	f	w	w	w	w	**(4)**

(13)

Notation 1.

- *Benutzten wir im Folgenden anstatt des kurzen Implikationspfeiles „$\Rightarrow$“ den langen Implikationspfeil „$\Longrightarrow$“, so geschieht dies nur aus ästhetischen Gründen, und macht keine bedeutungstechnischen Unterschied. Genauso verhält es sich mit den den Äquivalenzpfeilen „$\Leftrightarrow$“ und „$\Longleftrightarrow$“.*

  Man schreibt gegebenenfalls auch „$B \Leftarrow A$“ anstatt „$A \Rightarrow B$“, wenn dies die Lesbarkeit verbessert.
- *Der Implikationspfeil „$\Rightarrow$“ wird oft in mathematischen Beweisen (meisten in abgesetzten Formeln) in der Form „hieraus folgt“ benutzt. Schreibt man z.B. $A \Rightarrow B$, wobei aus dem Kontext heraus klar ist, dass die Aussage A wahr ist, so meint man einfach, dass die Wahrheit der Aussage B in offensichtlicher Weise aus bereits Bewiesenem (oder im Vorfeld Angenommenem) sowie dem Gelten der Aussage A folgt.*
- *Der Äquivalenzpfeil „$\Leftrightarrow$“ wird oft auch als definierendes Symbol benutzt. Ist zum Beispiel X eine Menge, so kann die Aussage $x \notin X$ (also die Bedeutung des Symbols $\notin$), auch in der Form*

  $$x \notin X \quad \Longleftrightarrow \quad \neg(x \in X) \qquad \textit{oder unmissverständlicher} \qquad x \notin X \quad \overset{\text{def.}}{\Longleftrightarrow} \quad \neg(x \in X)$$

  durch die Aussage $\neg(x \in X)$ (deren Bedeutung man bereits kennt) definiert werden.

[4]Man mache sich an einem Beispiel klar, das die Umkehrung $(\tilde{A} \Leftrightarrow \tilde{B}) \Rightarrow (\tilde{A} \equiv \tilde{B})$ üblicherweise nicht gilt.
[5]Man macht sich diese Rechnung auch leicht anhand sinnerhaltender sprachlicher Umformulierung klar.

Übung 2. *Die Kontravalenz $A \dot\vee B$ der Aussagen A und B, ist die Aussage „Genau eine der beiden Aussagen A und B ist wahr.“ Weisen Sie die (das Gelten der) Äquivalenz*

$$A \dot\vee B \qquad \Longleftrightarrow \qquad (A \wedge \neg B) \vee (\neg A \wedge B)$$

anhand einer Wahrheitstafel nach.

Übung 3. *Seien A, B, C Aussagen. Zeigen Sie anhand einer Wahrheitstafel:*

$$\begin{aligned} A \vee (B \vee C) \qquad &\Longleftrightarrow \qquad (A \vee B) \vee C \\ A \wedge (B \wedge C) \qquad &\Longleftrightarrow \qquad (A \wedge B) \wedge C \\ A \vee (B \wedge C) \qquad &\Longleftrightarrow \qquad (A \vee B) \wedge (A \vee C) \\ A \wedge (B \vee C) \qquad &\Longleftrightarrow \qquad (A \wedge B) \vee (A \wedge C). \end{aligned} \tag{14}$$

2.2 Beweistechniken für Implikationen

Wir besprechen nun einige Beweistechniken für Implikationen, die in der Praxis natürlich auch kombiniert auftreten. Seien hierfür A und B Aussagen.

2.2.1 Beweis durch Kontraposition

Um die Implikation $A \Rightarrow B$ zu zeigen reicht es aus, die Kontraposition (Implikation) $\neg B \Rightarrow \neg A$ zu zeigen: Gilt nämlich die Implikation $\neg B \Rightarrow \neg A$, so ist die Aussage $\neg B \Rightarrow \neg A$ immer wahr. Wegen (11) (oder (12)) ist daher auch die Aussage $A \Rightarrow B$ immer wahr, womit die Implikation $A \Rightarrow B$ gilt.

Bemerkung 4. *Das Prinzip des Beweises durch Kontraposition lässt sich auch anders begründen:*

- *Angenommen A ist wahr und es gilt $\neg B \Rightarrow \neg A$. Dann sind die Aussagen A und $\neg B \Rightarrow \neg A$ beide wahr – also muss laut Wahrheitstafel* (13) *auch B wahr sein (man kann sich nur noch in Zeile* **(1)** *befinden). Daher gilt $A \Rightarrow B$.*
- *Angenommen es wurde gezeigt, dass aus der Falschheit von B ($\neg B$ ist wahr) die Falschheit von A ($\neg A$ ist wahr) folgt. Weiß man dann, dass A wahr ist, so kann B nicht falsch sein (da ja die Falschheit von B die Falschheit von A erzwingen würde).*

Manche mathematischen Sachverhalte lassen sich „anscheinend nur“ durch (bzw. nur durch erheblichen Mehraufwand ohne) Kontraposition beweisen.

Beispiel 2. *Sei $n \in \mathbb{Z}$ fixiert. Wir betrachten die Aussagen $A =$ „n^2 ist ungerade.“ und $B =$ „n ist ungerade.“. Wir beweisen durch Kontraposition, dass die Implikation $A \Rightarrow B$ gilt:*

Beweis. Sei $\neg B$ wahr, also n gerade. Dann ist $n = 2m$ für ein $m \in \mathbb{Z}$, und wir erhalten

$$n^2 = (2m)^2 = 4m^2 = 2(2m^2).$$

Dies zeigt, dass n^2 gerade ist, dass also $\neg A$ gilt. □

2.2.2 Beweis durch Widerspruch

Um eine Implikation $A \Rightarrow B$ durch Widerspruch zu beweisen, nimmt man an, dass diese nicht gilt (dass also A wahr und B falsch ist), und leitet dann einen Widerspruch zu bereits bekannten Tatsachen her. In Formeln zeigt man also, dass

$$A \wedge \neg B \quad \Rightarrow \quad C \tag{15}$$

gilt, für eine falsche Aussage C. Dies zeigt dann die Implikation $A \Rightarrow B$.

Beweis der Behauptung: Da C falsch ist, kann $A \wedge \neg B$ wegen (15) sowie der Wahrheitstabelle (8) (für die Implikation) nicht wahr sein, muss also falsch sein. Da A per Voraussetzung wahr ist, muss dann $\neg B$ falsch sein, also B wahr sein. Daher gilt $A \Rightarrow B$. □

Beispiel 3. *Sei $n \in \mathbb{Z}$ gegeben. Wir betrachten die Aussage $A =$ „n ist gerade." sowie die Aussage $B =$ „$n+1$ ist ungerade.". Wir beweisen durch Widerspruch, dass die Implikation $A \Rightarrow B$ gilt:*

Beweis. Seien A und $\neg B$ wahr; also $n = 2m$ für ein $m \in \mathbb{Z}$, sowie $n+1 = 2k$ für ein $k \in \mathbb{Z}$. Wir erhalten

$$2k = n + 1 = 2m + 1 \qquad \Longrightarrow \qquad 1 = 2(k - m),$$

was wegen $m, k \in \mathbb{Z}$ unmöglich ist. □

Bemerkung 5. *Sowohl das Prinzip des Beweises durch Kontraposition, als auch das Prinzip des Beweises durch Widerspruch, wird als indirekte Beweistechnik bezeichnet.*

Übung 4. *Seien X, Y Mengen. Zeigen Sie die Äquivalenz*

$$X = Y \qquad \Longleftrightarrow \qquad X \subseteq Y \ \wedge \ Y \subseteq X.$$

Hinweis: Die eine Implikation ist trivial. Beweisen Sie die andere Implikation durch Widerspruch.

Übung 5. *Zeigen Sie, dass die Menge aller Mengen*

$$M = \{X \mid X \text{ ist Menge}\}$$

nicht existieren kann. Nehmen Sie hierfür an, dass M existiert, und leiten Sie hieraus einen Widerspruch her, indem Sie die Teilmenge $N := \{x \in M \mid x \notin x\}$ betrachten.

2.2.3 Direkter Beweis

a) Ein direkter Beweis einer Implikation $A \Rightarrow B$ erfolgt im Allgemeinen durch mehrere Beweisschritte. Konkreter zeigt man alle Implikationen in einer Implikationskette

$$A \Rightarrow C_1 \Rightarrow C_2 \Rightarrow \ldots \Rightarrow C_{n-1} \Rightarrow C_n \Rightarrow B$$

mit Zwischenaussagen $C_1, \ldots, C_n$ $(n \geq 1)$. Dies zeigt dann die Implikation $A \Rightarrow B$.

b) Ebenso kann man eine Äquivalenz $A \Leftrightarrow B$ zeigen, indem man alle Äquivalenzen in einer Äquivalenzkette

$$A \Leftrightarrow C_1 \Leftrightarrow C_2 \Leftrightarrow \ldots \Leftrightarrow C_{n-1} \Leftrightarrow C_n \Leftrightarrow B$$

mit Zwischenaussagen $C_1, \ldots, C_n$ $(1 \geq 2)$ zeigt. Dies zeigt dann die Äquivalenz $A \Leftrightarrow B$.

c) Will man die paarweise Äquivalenz gewisser Aussagen $A_1, \ldots, A_n$ $(n \geq 3)$ zeigen, so kann man dies in Form eines Ringschlusses tun. Dies bedeutet, das man alle Implikationen in der Implikationskette

$$A_1 \Rightarrow A_2 \Rightarrow \ldots \Rightarrow A_{n-1} \Rightarrow A_n \Rightarrow A_1$$

zeigt. Wegen a) folgt dann $A_k \Leftrightarrow A_\ell$ für alle $1 \leq k, \ell \leq n$. Natürlich kann man in obiger Kette auch wieder Implikationen bezüglich gewisser Zwischenaussagen einfügen.

Beispiel 4 (Zu Punkt a)). *Sei $n \in \mathbb{Z}$ gegeben, sowie*

$$\begin{aligned} A &= \text{„}n \text{ ist ungerade.“} \\ C_1 &= \text{„}n+1 \text{ ist gerade.“} \\ C_2 &= \text{„Es gilt } n+1 = 2m \text{ für ein } m \in \mathbb{Z}.\text{“} \\ B &= \text{„Es gilt } n = 2m+1 \text{ für ein } m \in \mathbb{Z}.\text{“} \end{aligned}$$

Dann gilt $A \Rightarrow C_1$ wegen Beispiel 3, $C_1 \Rightarrow C_2$ per Definition, und $C_2 \Rightarrow B$ vermöge Umformung.

Übung 6. *Seien X und Y Mengen. Zeigen Sie die Äquivalenz der folgenden Aussagen:*

i) $X \cap Y = X$,

ii) $X \subseteq Y$,

iii) $X \cup Y = Y$.

Hinweis: Weisen Sie die Äquivalenz durch einen Ringschluss nach, beweisen Sie also die Implikationen i) $\Rightarrow$ ii), ii) $\Rightarrow$ iii), und iii) $\Rightarrow$ i). Gegebenenfalls ist Übung 4 nutzbringend einsetzbar.

2.2.4 Beweis durch Fallunterscheidung

Um eine Implikation $A \Rightarrow B$ durch Fallunterscheidung zu zeigen, wählt man Aussagen C_1, C_2, sodass die folgenden Implikationen gelten:

$$A \;\Rightarrow\; C_1 \vee C_2, \qquad C_1 \Rightarrow B, \qquad C_2 \Rightarrow B. \tag{16}$$

Nun gilt aber auch die folgende Implikation (siehe Übung 7):

$$((C_1 \vee C_2) \wedge ((C_1 \Rightarrow B) \wedge (C_2 \Rightarrow B))) \;\Rightarrow\; B. \tag{17}$$

Wegen (16) ist die linke Seite von (17) wahr, sodass B auch wahr sein muss.

Übung 7. *Zeigen Sie die Implikation* (17) *anhand einer Wahrheitstafel.*

Übung 8. *Seien $n, k \in \mathbb{N}$ gegeben mit $n+3 = k^2$. Zeigen Sie durch Fallunterscheidung, dass n nicht durch* 4 *Teilbar ist (dass also $n \neq 4m$ für alle $m \in \mathbb{N}$ gilt). Unterscheiden Sie hierfür die beiden Fälle, dass k gerade bzw. ungerade ist.*

2.3 Quantoren

Eine Aussageform A über einer Menge X ist eine Menge $A = \{A(x) \,|\, x \in X\}$ von Aussagen, d.h., für jedes $x \in X$ ist $A(x)$ eine Aussage. Mit Hilfe der folgenden Quantoren können hieraus neue Aussagen gebildet werden:

a) Das Symbol „$\forall$“ („für alle“) wird als Allquantor bezeichnet. Dann ist

$$\forall\, x \in X \colon A(x) \tag{18}$$

die symbolische Darstellung der Aussage

„Für alle $x \in X$ gilt $A(x)$.“ $\equiv$ „Die Aussage $A(x)$ ist für alle $x \in X$ wahr.“

Die Aussageform A heißt allgemeingültig, falls (18) eine wahre Aussage ist (falls also $A(x)$ für jedes $x \in X$ wahr ist). Um eine Allaussage (18) zu widerlegen, reicht es aus, ein Gegenbeispiel anzugeben – also ein $x \in X$ zu finden, sodass $A(x)$ falsch ist.

b) Das Symbol „∃“ (ës existiert”) wird als Existenzquantor bezeichnet. Dann ist

$$\exists\, x \in X\colon A(x) \tag{19}$$

die symbolische Darstellung der Aussage

„Es existiert ein $x \in X$, sodass die Aussage $A(x)$ wahr ist.“

Die Aussageform A heißt erfüllbar, falls (19) eine wahre Aussage ist (falls also $A(x)$ für ein $x \in X$ wahr ist).

c) Das Symbol „∄“ („es existiert nicht“) ist die Negation des Existenzquantors. Dann ist

$$\nexists\, x \in X\colon A(x) \tag{20}$$

die symbolische Darstellung der Aussage

„Es existiert kein $x \in X$, sodass die Aussage $A(x)$ wahr ist.“

Die Aussageform A heißt unerfüllbar, falls (20) eine wahre Aussage ist (falls also $A(x)$ für kein $x \in X$ wahr ist).

d) Das Symbol „∃!“ („es existiert genau ein“) wird als Eindeutigkeitsquantor bezeichnet. Dann ist

$$\exists!\, x \in X\colon A(x)$$

die symbolische Darstellung der Aussage

„Es existiert genau ein $x \in X$, sodass die Aussage $A(x)$ wahr ist.“

Es gelten die folgenden Rechenregeln:[6]

$$\begin{aligned} \neg(\forall\, x \in X\colon A(x)) &\equiv \exists\, x \in X\colon \neg A(x) \\ \neg(\exists\, x \in X\colon A(x)) &\equiv \forall\, x \in X\colon \neg A(x) \\ &\equiv \nexists\, x \in X\colon A(x). \end{aligned}$$

2.4 Das Induktionsprinzip

Die letzte Beweistechnik, die wir besprechen wollen, ist der Beweis per Induktion. Wir setzen $\mathbb{N}_\infty := \mathbb{N} \cup \{\infty\}$ für das Unendlichkeitssymbol ∞, und vereinbaren $n < \infty$ und $n \leq \infty$ für alle $n \in \mathbb{N}$.

Satz 1. *Seien $\mathbb{N} \ni k < p \in \mathbb{N}_\infty$ vorgegeben. Eine Aussageform A über $N := \{n \in \mathbb{N} \mid k \leq n < p\}$ ist allgemeingültig (es ist also $A(n)$ wahr für jedes $n \in N$), falls die folgenden zwei Bedingungen erfüllt sind:*

a) $A(k)$ ist wahr. *(Induktionsanfang)*

b) Für alle $n \in N$ gilt *(Induktionsschritt)*

$$\underbrace{A(n)}_{\text{Induktionsvoraussetzung}} \quad \Longrightarrow \quad \underbrace{A(n+1).}_{\text{Induktionsbehauptung}}$$

Oder alternativ:

a') $A(k)$ ist wahr. *(Induktionsanfang)*

[6] Das Symbol „≡“ kann hier natürlich auch wieder durch das Symbol „⇔“ ersetzt werden.

b') Für alle $n \in N$ gilt *(Induktionsschritt)*

$$\underbrace{\forall\, k \le q \le n\colon\ A(q)}_{\text{Induktionsvoraussetzung}} \qquad \Longrightarrow \qquad \underbrace{A(n+1)}_{\text{Induktionsbehauptung}}.$$

Beweis. Wir beweisen den Satz durch Widerspruch (indirekter Beweis).

Version 1: Angenommen a) und b) sind erfüllt, aber $A(q)$ ist falsch für ein $q \in N$. Dann gilt

$$\emptyset \ne M := \{\ell \in N \mid A(\ell) \text{ ist falsch}\} \subseteq \mathbb{N}.$$

Nach dem Wohlordnungsprinzip besitzt M ein kleinstes Element m. Wegen a) gilt dann $m \ge k+1$, also $m-1 \ge k$. Wegen b) ist dann auch $A(m-1)$ falsch, was der Minimalität (der Wahl) von m widerspricht.

Version 2: Angenommen a') und b') sind erfüllt, aber $A(q)$ ist falsch für ein $q \in N$. Dann gilt

$$\emptyset \ne M := \{\ell \in N \mid A(\ell) \text{ ist falsch}\} \subseteq \mathbb{N}.$$

Nach dem Wohlordnungsprinzip besitzt M ein kleinstes Element m. Wegen a') gilt dann $m \ge k+1$, also $m-1 \ge k$. Wegen b') ist dann auch $A(p)$ falsch für ein $k \le p \le m-1$, was der Minimalität (der Wahl) von m widerspricht. □

Bemerkung 6. *Es ist unerlässlich, sowohl den Induktionsanfang, als auch den Induktionsschritt zu zeigen. Für sich genommen sagt keine dieser Aussagen etwas über die Allgemeingültigkeit von A aus. Das Induktionsprinzip wird im Folgenden zumeist für denn Fall $k = 0, 1$ und $p = \infty$ angewandt werden.*

Beispiel 5. *Wir zeigen per Induktion, dass $n^2 + n$ für jedes $n \in \mathbb{N}_{>0}$ gerade (durch 2 teilbar) ist.*

(IA) Induktionsanfang: $1^2 + 1 = 2$ ist gerade.

(IV) Induktionvoraussetzung: Es sei $n^2 + n$ gerade für ein $n \in \mathbb{N}_{>0}$.

(IS) Induktionsschritt: Es gilt

$$(n+1)^2 + (n+1) = n^2 + 2n + 1 + n + 1 = (n^2 + n) + 2(n+1),$$

was offensichtlich durch 2 teilbar ist, da $n^2 + n$ per Induktionsvoraussetzung durch 2 teilbar ist.

Bemerkung 7. *In einem Induktionsbeweis ist stets darauf zu achten, den Induktionsanfang (abgekürzt IA), die Induktionsvoraussetzung (abgekürzt IV), sowie den Induktionsschritt (abgekürzt IS) klar kenntlich zu machen. Die formell sicherste Variante wurde in Beispiel 5 dargestellt. Besondere Vorsicht ist auch bei der Induktionsvoraussetzung geboten. Zum Beispiel wäre es grob fahrlässig bzw. falsch, unter der Induktionsvoraussetzung in Beispiel 5 zu schreiben: „Es sei $n^2 + n$ für $n \in \mathbb{N}_{>0}$ gerade.“. In der Tat lässt sich dieser Satz sehr leicht als „Es sei $n^2 + n$ gerade für alle $n \in \mathbb{N}_{>0}$.“ interpretieren, was ja aber gerade das ist, was erst noch gezeigt werden muss.*

Übung 9. *Zeigen Sie per Induktion, dass $n^3 + 2n$ für alle $n \in \mathbb{N}_{>0}$ durch 3 teilbar ist.*

Wir definieren die Fakultät einer Natürlichen Zahl durch $0! := 1$, sowie

$$n\text{-Fakultät:} \qquad n! := 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n \qquad \forall\, n \in \mathbb{N}_{>0}.$$

Formeller lässt die Fakultät auch induktiv definieren durch

$$0! := 1 \qquad \text{und} \qquad (n+1)! := n! \cdot (n+1) \qquad \forall\, n \in \mathbb{N}_{>0}.$$

In der Tat impliziert dann Satz 1, dass der Ausdruck „$n!$“ für alle $n \in \mathbb{N}$ definiert ist:

(IA): $0! = 1$ ist definiert.

(IV): Es sei $n!$ definiert für ein $n \in \mathbb{N}_{>0}$.

(IS): Es ist $(n+1)! = n! \cdot (n+1)$ definiert wegen (IV).

Wir werden die Methode der induktiven Definition von Ausdrücken in dieser Vorlesung noch öfter anwenden. Allgemein lässt diese sich wie folgt beschreiben:

Prinzip 1. *Sei $k \in \mathbb{N}$, sowie T_n für jedes $n \geq k$ eine Menge von mathematischen Termen (Ausdrücken). Es seien weiterhin alle Terme $\tau \in T_k$ definiert (IA), und eine der beiden Implikationen*

$$\underbrace{\text{Alle Terme in } T_n \text{ sind definiert.}}_{\text{(IV)}} \quad \overset{\text{(IS)}}{\Longrightarrow} \quad \text{Alle Terme in } T_{n+1} \text{ sind definiert.}$$

$$\underbrace{\text{Alle Terme in } \textstyle\bigcup_{k\leq \ell \leq n} T_\ell \text{ sind definiert.}}_{\text{(IV)}} \quad \overset{\text{(IS)}}{\Longrightarrow} \quad \text{Alle Terme in } T_{n+1} \text{ sind definiert.}$$

gelte für alle $n \geq k$. Dann sind alle Terme in $\bigcup_{n\geq k} T_n$ definiert.

Notation 2. *Gegeben $n \in \mathbb{N}_{>0}$ und $x_1, \ldots, x_n \in \mathbb{Q}$, so definieren wir induktiv (Prinzip 1)*

$$x_1 + \ldots + x_n := (x_1 + \ldots + x_{n-1}) + x_n \qquad \text{und} \qquad x_1 \cdot \ldots \cdot x_n := (x_1 \cdot \ldots \cdot x_{n-1}) \cdot x_n.$$

Sei nun $n \in \mathbb{N}$, und $J = \{j_1, \ldots, j_n\}$ eine n-elementige Menge, also $J = \emptyset$ für $n = 0$. Es sei für jedes $j \in J$ ein $x_j \in \mathbb{Q}$ gegeben. Wir definieren

$$\begin{array}{llll} \sum_{j\in J} x_j := 0 & \text{sowie} & \prod_{j\in J} x_j := 1 & \text{für} \quad J = \emptyset, \\ \sum_{j\in J} x_j := x_{j_1} + \ldots + x_{j_n} & \text{sowie} & \prod_{j\in J} x_j := x_{j_1} \cdot \ldots \cdot x_{j_n} & \text{für} \quad J \neq \emptyset. \end{array}$$

Sind $m, n \in \mathbb{Z}$ mit $m \leq n$, sowie $x_m, \ldots, x_n \in \mathbb{Q}$, so setzt man

$$\sum_{k=m}^{n} x_k := \sum_{j\in\{m,\ldots,n\}} x_j \qquad \text{und} \qquad \prod_{k=m}^{n} x_k := \prod_{j\in\{m,\ldots,n\}} x_j.$$

Vermöge Induktionsbeweis kann man dann (offensichtliche) Identitäten, wie z.B.

$$\sum_{k=m}^{n} x_k = \sum_{k=m}^{p} x_k + \sum_{k=p}^{n} x_k \qquad \text{für} \qquad m \leq p \leq n$$

formell strikt zeigen. Wir werden in Abschnitt 4.1 (siehe Notation 12) obige Notationen später noch allgemeiner fassen, und entsprechende Identitäten nachweisen.

Beispiel 6 (Kleiner Gauß)**.** *Wir zeigen per Induktion, dass*

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2} \qquad \forall\, n \in \mathbb{N}. \tag{21}$$

(IA): Für $n = 0$ ist (21) offensichtlich wahr, da dann auf beiden Seiten 0 steht.

(IV): Sei (21) erfüllt für ein $n \in \mathbb{N}$.

(IS): Es gilt

$$\sum_{k=0}^{n+1} k = (n+1) + \sum_{k=0}^{n} k \overset{\text{(IV)}}{=} \frac{2(n+1)}{2} + \frac{n(n+1)}{2} = \frac{(n+1)\cdot(n+2)}{2}.$$

3 Mengen, Relationen, und Abbildungen

Gegenstand dieses Kapitels sind Operationen mit Mengen, Relationen, und Abbildungen zwischen Mengen.

Terminologie 4. *Sei J eine Menge. Für jedes $j \in J$, sei O_j ein mathematisches Objekt (z.B. eine Zahl oder eine Menge etc.). Wir sprechen dann von der Familie von Objekten $(O_j)_{j\in J}$.*

3.1 Operationen mit Mengen

Im Folgenden seien $J \neq \emptyset$ sowie X, Y, Z Mengen; und es sei $(Z_j)_{j\in J}$ eine Familie von Mengen.

Definition 3. *Die Vereinigung der Mengen $(Z_j)_{j\in J}$ ist die Menge*

$$\bigcup_{j\in J} Z_j := \{z \mid \exists\, j \in J \colon z \in Z_j\}.$$

Sie besteht aus genau den Elementen, die in mindestens einer der Mengen Z_j $(j \in J)$ enthalten ist.

Notation 3. *Sind die Z_j paarweise disjunkt – gilt also $Z_i \cap Z_j = \emptyset$ für $J \ni i \neq j \in J$ – so schreibt man auch $\dot{\bigcup}_{j\in J} Z_j$ anstelle $\bigcup_{j\in J} Z_j$.*

Es handelt sich bei Definition 3 um eine Definition des Typs (1). Wir hatten bereits in Kapitel 1 darauf hingewiesen, dass derartige Definition zu Widersprüchen führen können – obige Konstruktion wird aber durch das sogenannte Vereinigungsaxiom der Zermelo-Fraenkel-Mengenlehre (ZF) legitimiert. Eine Konstruktion von Mengen, die immer funktioniert, ist der Übergang zu Teilmengen einer bereits vorhandenen Menge X vermöge Aussonderung durch eine Aussageformen A, also

$$\{x \in X \mid A(x)\} \qquad - \qquad \text{Die Menge aller } x \in X \text{, für die die Aussage } A(x) \text{ gilt.}$$

Beispiel hierfür sind die folgenden Konstruktionen:

Definition 4.

1) *Der Schnitt der Mengen $(Z_j)_{j\in J}$ ist definiert durch*

$$\bigcap_{j\in J} Z_j := \big\{z \in \bigcup_{j\in J} Z_j \mid \forall\, j \in J \colon z \in Z_j\big\},$$

besteht also aus allen Elementen, die in jeder der Mengen Z_j mit $j \in J$ enthalten ist.

2) *Die Differenz der Mengen X und Y ist definiert durch (X ohne Y)*

$$X \setminus Y := \{x \in X \mid x \notin Y\},$$

besteht also aus genau den Elementen von X, die nicht in Y enthalten sind. Zum Beispiel ist $\{1, 2\} \setminus \{0, 1, 3\} = \{2\}$. Ist $Y \subseteq X$ eine Teilmenge, so wird $X \setminus Y$ auch als das Komplement von Y in X bezeichnet.

Ist $J = \{j_1, \ldots, j_n\}$ $(n \geq 2)$ endlich, so schreibt man auch

$$\begin{array}{lcl} Z_{j_1} \cup \ldots \cup Z_{j_n} & \text{anstelle} & \bigcup_{j\in J} Z_j \\ Z_{j_1} \,\dot\cup\, \ldots \,\dot\cup\, Z_{j_n} & \text{anstelle} & \dot{\bigcup}_{j\in J} Z_j \\ Z_{j_1} \cap \ldots \cap Z_{j_n} & \text{anstelle} & \bigcap_{j\in J} Z_j. \end{array}$$

Beispielsweise ist $\{1, 2\} \cup \{0, 1, 3\} = \{0, 1, 2, 3\}$, sowie $\{1, 2\} \cap \{0, 1, 3\} = \{1\}$.

Bemerkung 8. *Seien X, Y, Z Mengen, $(Z_j)_{j\in J}$ eine Familie von Mengen, und $M \subseteq X$ eine Teilmenge von X. Dann gelten die folgenden Identitäten (Übung 10):*

- $X \cap (Y \cap Z) = (X \cap Y) \cap Z$
- $X \cup (Y \cup Z) = (X \cup Y) \cup Z$
- $M \setminus (X \setminus Y) = M \cap Y$
- $(X \setminus Y) \setminus Z = X \setminus (Y \cup Z)$
- $X \cap \bigcup_{j \in J} Z_j = \bigcup_{j \in J} X \cap Z_j$
- $X \cup \bigcap_{j \in J} Z_j = \bigcap_{j \in J} X \cup Z_j$
- $X \setminus \bigcup_{j \in J} Z_j = \bigcap_{j \in J} X \setminus Z_j$
- $X \setminus \bigcap_{j \in J} Z_j = \bigcup_{j \in J} X \setminus Z_j$

Man erhält diese Aussagen zum Beispiel unter Zuhilfenahme der logischen Operationen $\neg$, $\wedge$, $\vee$ durch sukzessive Äquivalenzumformungen aus den Regeln (3) *und* (14)*:*

$$\begin{aligned}
z \in M \setminus (X \setminus Y) \quad &\Longleftrightarrow \quad z \in M \;\wedge\; z \notin (X \setminus Y) \\
&\Longleftrightarrow \quad z \in M \;\wedge\; \neg(z \in X \setminus Y) \\
&\Longleftrightarrow \quad z \in M \;\wedge\; \neg(z \in X \;\wedge\; z \notin Y) \\
&\Longleftrightarrow \quad z \in M \;\wedge\; (z \notin X \;\vee\; z \in Y) \\
&\Longleftrightarrow \quad (z \in M \;\wedge\; z \notin X) \;\vee\; (z \in M \;\vee\; z \in Y) \\
&\Longleftrightarrow \quad z \in M \;\wedge\; z \in Y \\
&\Longleftrightarrow \quad z \in M \cap Y.
\end{aligned}$$

Übung 10. *Zeigen Sie die Identitäten in Bemerkung 8.*

Definition 5. *Die Potenzmenge von X ist die Menge aller Teilmengen von X, also die Menge*

$$\mathcal{P}(X) := \{Y \mid Y \subseteq X\}$$

deren Elemente genau die Teilmengen Y von X sind. Die Existenz der Potenzmenge $\mathcal{P}(X)$ zu einer gegebenen Menge X ist Teil der Axiomatik von ZF.

Zum Beispiel ist (Erinnerung: $\emptyset$ ist Teilmenge jeder Menge)

$$\mathcal{P}(\emptyset) = \{\emptyset\} \qquad \text{und} \qquad \mathcal{P}(\{0,2\}) = \{\emptyset, \{0\}, \{2\}, \{0,2\}\}.$$

Übung 11. *Bestimmen Sie die Mengen $\mathcal{P}(X)$ und $\mathcal{P}(\mathcal{P}(X))$ für die Menge $X = \{a\}$. Bestimmen Sie weiterhin $|\mathcal{P}(X)|$ und $|\mathcal{P}(\mathcal{P}(X))|$; und prüfen Sie hieran die Formel $|\mathcal{P}(Y)| = 2^{|Y|}$ für eine endliche Menge Y nach, die wir später noch behandeln werden.*

Definition 6. *Das kartesische Produkt der Mengen X und Y ist definiert durch*

$$X \times Y := \{(x,y) \mid x \in X \;\wedge\; y \in Y\}.$$

Ein Ausdruck der Form (x,y), mit $x \in X$ und $y \in Y$, wird als geordnetes Paar bezeichnet. Man definiert die Gleichheit zweier geordneter Paare durch

$$(x,y) = (x',y') \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad x = x' \;\wedge\; y = y' \tag{22}$$

für $x, x' \in X$ und $y, y' \in Y$.

Bemerkung 9.

1) *Für $X \neq Y$ ist $X \times Y \neq Y \times X$. Bei der Angabe von geordneten Paaren (x,y), ist also stets auf die Reihenfolge zu achten ($(x,y) = (y,x) \;\Rightarrow\; x = y \;\Rightarrow\; X \cap Y \neq \emptyset$).*

2) Sei $X = \emptyset$, und Y eine Menge. Dann gilt $X \times Y = \emptyset$; denn da $X = \emptyset$ kein Element enthält, existiert auch kein geordnetes Paar (x, y) mit $x \in X$ (und $y \in Y$).

3) Seien $X = \{x_1, \ldots, x_m\}$ mit $m \geq 1$, und $Y = \{y_1, \ldots, y_n\}$ mit $n \geq 1$ nichtleere endliche Mengen. Dann sind (x_i, y_j) für $i = 1, \ldots, m$ und $j = 1, \ldots, n$ wegen (22) paarweise unterschiedlich. Also hat $X \times Y$ genau $m \cdot n$ Elemente; sodass also $|X \times Y| = |X| \cdot |Y|$ gilt.

Beispiel 7.

- $\{0,1\} \times \{a,b\} = \{(0,a),(0,b),(1,a),(1,b)\}$ *hat* $4 = 2 \cdot 2 = |\{0,1\}| \cdot |\{a,b\}|$ *Elemente.*
- $\{a,b\} \times \{0,1\} = \{(a,0),(a,1),(b,0),(b,1)\}$ *hat* $4 = 2 \cdot 2 = |\{a,b\}| \cdot |\{0,1\}|$ *Elemente.*
- $\mathbb{Z}^2 = \mathbb{Z} \times \mathbb{Z} = \{(m,n) \mid m,n \in \mathbb{Z}\}$ *(oder* $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x,y) \mid x,y \in \mathbb{R}\}$ *die Tafelebene).*
- $*$ *Die disjunkte Vereinigung einer Familie von nichtleeren Mengen* $(Z_j)_{j \in J}$ *ist definiert durch*

$$\bigsqcup_{j \in J} Z_j := \{(j,z) \in J \times \bigcup_{j \in J} Z_j \mid z \in Z_j\} \qquad \Big(= \dot{\bigcup}_{j \in J} \{j\} \times Z_j \Big).$$

Der Sinn hinter dieser Definition ist natürlich der, dass die Mengen Z_j nicht paarweise disjunkt sein müssen – dass also durchaus $Z_i \cap Z_j \neq \emptyset$ gelten kann. Manchmal will man aber eine Menge haben, in welcher die Mengen Z_j sozusagen separiert voneinander als Teilmengen enthalten sind. Dies wird in obiger Konstruktion erreicht, indem man jedes Element einer der Mengen Z_j mit dem Index „j“ versieht: In der Tat sind ja die Mengen $\{j\} \times Z_j$ ($j \in J$) wegen (22) paarweise disjunkt.

Zum Beispiel: Eine Bank wird neu eröffnet, und n Personen ($n \geq 2$) zahlen einen Betrag ein – Person j zahlt den Betrag $b_j \in \mathbb{Q}$ ein. Die Bank ist unerfahren, und erstellt die Liste B aller eingezahlten Beträge in der Form $B := \{b_1\} \cup \cdots \cup \{b_n\}$. Dies führt natürlich zu informationstechnischen Problemen falls mehrere Personen den gleichen Betrag einzahlen (diese Beträge werden nämlich in der Liste B zu einem Element zusammengefasst). Die richtige Art und Weise die Liste B zu bilden, wäre natürlich in der Form $B := \{b_1\} \sqcup \cdots \sqcup \{b_n\}$ – denn dann wird zusätzlich zum eingezahlten Betrag auch der Name des Einzahlenden (in Form des Index j) festgehalten, sodass zwei eingezahlte Beträge immer unterschiedliche Elemente von B sind.

Bemerkung 10. *Die Konstruktion in Definition 6 ist erlaubt ($X \times Y$ ist tatsächlich eine Menge), denn $X \times Y$ lässt sich auch als Teilmenge von $\mathcal{P}(\mathcal{P}(X \cup Y))$ auffassen; und zwar*

$$X \times Y := \big\{z \in \mathcal{P}(\mathcal{P}(X \cup Y)) \mid \exists\, x \in X,\ y \in Y \colon z = \{\{x\},\{x,y\}\}\big\}.$$

Für $x \in X$ und $y \in Y$ kann man nun das Symbol „(x,y)“ einfach durch $(x,y) := \{\{x\},\{x,y\}\}$ definieren. Per Konstruktion gilt dann (vgl. (22))

$$(x,y) = (x',y') \qquad \Longleftrightarrow \qquad x = x' \quad \wedge \quad y = y',$$

denn für $x, x' \in X$ und $y, y' \in Y$ weist man die folgende Äquivalenz leicht nach (siehe Übung 12):

$$\{\{x\},\{x,y\}\} = \{\{x'\},\{x',y'\}\} \qquad \Longleftrightarrow \qquad x = x' \ \wedge \ y = y'. \tag{23}$$

Übung 12. *Zeigen Sie die Äquivalenz (23).*

Hinweis: Benutzen Sie, dass zwei Mengen gleich sind, wenn diese aus den selben Elementen bestehen. Die eine Richtung ist dann klar. Für die andere Richtung sind verschiedene Fälle zu beachten. Es könnte z.B. $\{\{x\},\{x,y\}\}$ durchaus einelementig sein, nämlich wenn $x = y$ gilt (falls $X \cap Y \neq \emptyset$).

Lösung: Gilt die rechte Seite von (23), so bestehen die Mengen $\{\{x\},\{x,y\}\}$ und $\{\{x'\},\{x',y'\}\}$ aus den selben Elementen, sind also gleich.

Es gelte nun die Gleichheit auf der linken Seite von (23). Wir machen eine Fallunterscheidung:

Fall 1: Es gelte $|\{\{x\},\{x,y\}\}| = 1$, also zwangsläufig auch $|\{\{x'\},\{x',y'\}\}| = 1$.

- Dann ist $\{x\} = \{x,y\}$, also $\{x,y\}$ einelementig, also $x = y$, also $\{\{x\},\{x,y\}\} = \{\{x\}\}$.
- Dann ist $\{x'\} = \{x',y'\}$, also $\{x',y'\}$ einelementig, also $x' = y'$, also $\{\{x'\},\{x',y'\}\} = \{\{x'\}\}$.

Es gilt dann die Implikation (vgl. linke Seite von (23))

$$\{\{x\}\} = \{\{x'\}\} \qquad \Longrightarrow \qquad \{x\} = \{x'\} \qquad \Longrightarrow \qquad x = x'$$

also auch $y = x = x' = y'$.

Fall 2: Es gelte $|\{\{x\},\{x,y\}\}| = 2$, also zwangsläufig auch $|\{\{x'\},\{x',y'\}\}| = 2$. Dann gilt $x \neq y$ und $x' \neq y'$. Dies erzwingt $\{x\} = \{x'\}$ (der andere Fall $\{x\} = \{x',y'\}$ kann wegen $x' \neq y'$ nicht gelten). Es folgt $x = x'$, also $\{x,y\} = \{x,y'\}$, also $y = y'$, was die rechte Seiten von (23) zeigt. □

Wir benötigen Definition 6 in noch allgemeinerer Form:

Definition 7. *Das n-fache Produkt von Mengen $X_1,\ldots,X_n$ ($n \geq 2$), ist definiert durch*

$$X_1 \times \cdots \times X_n := \{(x_1,\ldots,x_n) \mid x_i \in X_i \text{ für } i = 1,\ldots,n\}. \tag{24}$$

Ein Ausdruck der Form $(x_1,\ldots,x_n)$ mit $x_i \in X_i$ für $i = 1,\ldots,n$ wird als n-Tupel (geordnete Liste) bezeichnet. Man definiert die Gleichheit zweier n-Tupel durch

$$(x_1,\ldots,x_n) = (x_1',\ldots,x_n') \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad x_j = x_j' \quad \forall\, j = 1,\ldots,n.$$

- *Für X eine Menge setzt man $X^0 := \emptyset$, sowie*
$$X^n := \underbrace{X \times \ldots \times X}_{\text{n−mal}} \qquad \forall\, n \in \mathbb{N}_{>0}. \tag{25}$$
- *Auch bei einem n-Tupel kommt es wieder auf die Reihenfolge an, in der die Elemente angegeben werden.*
- *Sind $X_1,\ldots,X_n \neq \emptyset$ endliche Mengen, so gilt offensichtlich (vgl. Bemerkung 9)*
$$|X_1 \times \ldots \times X_n| = |X_1| \cdot \ldots \cdot |X_n|. \tag{26}$$
- *Wir bemerken nur am Rande, dass sich das n-fache Produkt auch durch wiederholtes bilden des kartesischen Produktes (Definition 6) erhalten lässt; nämlich in der Form*
$$X_1 \times (X_2 \times (\cdots \times (X_{n-2} \times (X_{n-1} \times X_n))\ldots)).$$
(Man bildet also das Produkt $X_{n-1} \times X_n$, dann das Produkt $X_{n-2} \times (X_{n-1} \times X_n)$, ... und so weiter.)

Definition 7 lässt sich nochmals verallgemeinern:

Definition 8. *Sei $J \neq \emptyset$. Das Produkt einer Familie von Mengen $(X_j)_{j\in J}$, ist definiert durch*

$$\textstyle\prod_{j\in J} X_j := \{(x_j)_{j\in J} \mid x_j \in X_j \text{ für alle } j \in J\}, \tag{27}$$

besteht also aus allen Familien $(x_j)_{j\in J}$ mit $x_j \in X_j$ für alle $j \in J$, d.h.,

$$(x_j)_{j\in J} \in \textstyle\prod_{j\in J} X_j \qquad \Longleftrightarrow \qquad x_j \in X_j \quad \forall\, j \in J.$$

Für $(x_j)_{j\in J} \in \prod_{j\in J} X_j$ und $(x_j')_{j\in J} \in \prod_{j\in J} X_j$, setzt man

$$(x_j)_{j\in J} = (x_j')_{j\in J} \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad x_j = x_j' \quad \forall\, j \in J.$$

Bemerkung 11.

- *Wir erhalten die Konstruktion aus Definition 7, indem wir in Definition 8 die Indexmenge* $J = \{1, \ldots, n\}$ *betrachten, und dann eine Familie* $(x_j)_{j\in J} \in \prod_{j\in J} X_j$ *als* n*-Tupel* $(x_1, \ldots, x_n)$ *auffassen.*
- *Ähnlich wie in Definition 6 ist eigentlich wieder unklar, ob es sich bei (27) um eine richtige Mengendefinition handelt (es handelt sich hier um eine Abart der Definition vom Typus (1)). Wir wollen an dieser Stelle lediglich anmerken, dass sich sich* $\prod_{j\in J} X_j$ *auch als die Teilmenge*

  $$\big\{z \in \mathcal{P}(\textstyle\bigsqcup_{j\in J} X_j) \mid z = \bigcup_{j\in J}\{(j, x_j)\} \text{ mit } x_j \in X_j \text{ für alle } j \in J\big\}$$

  von $\mathcal{P}(\bigsqcup_{j\in J} X_j)$ *auffassen lässt.*

Notation 4. *Sei* X *eine Menge. In der Situation von Definition 8, betrachten wir den Spezialfall, dass* $X_j = X$ *für alle* $j \in J$ *gilt. Wir definieren*

$$X^J := \prod_{j\in J} X_j = \prod_{j\in J} X. \tag{28}$$

Beispiel: *Ist* $J = \mathsf{N}$ *von der Form* $\mathsf{N} = \{n \in \mathbb{Z} \mid n \geq \mathsf{p}\}$ *für ein* $\mathsf{p} \in \mathbb{Z}$ *(meistens mit* $\mathsf{p} = 0$ *oder* $\mathsf{p} = 1$*), so wird* X^{N} *auch als Folgenraum bezeichnet. Ein Elemente* $(x_n)_{n\in\mathsf{N}} \in X^{\mathsf{N}}$ *wird dann auch als Folge in* X *(bzw.* X*-Folge) bezeichnet. Beispiele sind:*

- *Die* $\mathbb{Q}$*-Folge* $(\frac{1}{n})_{n\in\mathbb{N}_{>0}} \in \mathbb{Q}^{\mathbb{N}_{>0}}$. *(*$X = \mathbb{Q} \;\wedge\; \mathsf{N} = \mathbb{N}_{>0}$*)*
- *Die* $\mathbb{N}$*-Folge* $(n^2)_{n\in\mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$. *(*$X = \mathbb{N} \;\wedge\; \mathsf{N} = \mathbb{N}$*)*
- *Die Fibonacci-Folge* $(f_n)_{n\in\mathbb{N}_{>0}} \in (\mathbb{N}_{>0})^{\mathbb{N}_{>0}}$ *(*$\mathbb{N}_{>0}$*-Folge). Diese ist rekursiv definiert durch* $f_1 := 1$, $f_2 := 1$, *sowie* $f_n := f_{n-1} + f_{n-2}$ *für* $n \geq 3$.

3.2 Ordnungen und Relationen

Terminologie 5. *Eine* Relation *zwischen Mengen* X *und* Y *ist eine Teilmenge* $R \subseteq X \times Y$. *Gegeben* $x \in X$ *und* $y \in Y$, *so schreibt man*

$$xRy \quad \overset{\text{def.}}{\iff} \quad (x, y) \in R.$$

Ist $X = Y$, *so heißt* R *eine Relation auf* X.

3.2.1 Äquivalenzklassen

Definition 9. *Eine Äquivalenzrelation auf einer Menge* X *ist eine Relation* R *auf* X, *sodass für alle* $x, y, z \in X$ *die folgenden Bedingungen gelten:*

1) Transitivität: $xRy \wedge yRz \;\Rightarrow\; xRz$,
2) Symmetrie: $xRy \;\Rightarrow\; yRx$,
3) Reflexivität: xRx.

Äquivalenzrelationen notiert man üblicherweise mit $\sim$ anstatt mit R. Man schreibt dann $x \sim y$ anstelle xRy für $x, y \in X$ und sagt, dass x und y äquivalent zueinander sind (wegen 2) ist dieser Wortgebrauch nicht missverständlich).

Mit Hilfe von Äquivalenzrelationen ist es möglich, Mengen in disjunkte Teilmengen zu zerlegen: Sei X eine Menge, und $\sim$ eine Äquivalenzrelation auf X.

- Gegeben $x \in X$, so wird die Teilmenge

$$[x]_\sim := \{y \in X \mid x \sim y\} \subseteq X$$

als *Äquivalenzklasse von* x bezeichnet (sie besteht also aus allen Elementen von X, die äquivalent zu x sind). Wegen 3) gilt

$$x \in [x]_\sim \neq \emptyset \quad \forall\, x \in X \qquad \Longrightarrow \qquad X = \bigcup_{x \in X} [x]_\sim. \tag{29}$$

- Die Menge aller Äquivalenzklassen wird notiert mit

$$\begin{aligned} X/\!\sim\, &= \{[x]_\sim \mid x \in X\} \subseteq \mathcal{P}(X) \\ &\quad \big(= \{\alpha \in \mathcal{P}(X) \mid \exists\, x \in X \colon \alpha = [x]_\sim\}\big). \end{aligned}$$

  - Es wird $X/\!\sim$ wird oft auch als *Quotientenraum* (Quotient von X bezüglich $\sim$) bezeichnet.
  - Die Elemente $x \in \alpha$ einer Äquivalenzklasse $\alpha \in X/\!\sim$, werden auch als die *Repräsentanten* dieser Äquivalenzklasse bezeichnet.

Mit (29) erhalten wir

$$\alpha \neq \emptyset \quad \text{für jedes} \quad \alpha \in X/\!\sim \qquad \text{sowie} \qquad X = \bigcup_{\alpha \in X/\sim} \alpha. \tag{30}$$

Lemma 1. *Gegeben eine Äquivalenzrelation $\sim$ auf einer Menge X, so ist X die disjunkte Vereinigung der nichtleeren Mengen $\alpha \in X/\!\sim$. In Formeln*

$$X = \dot{\bigcup}_{\alpha \in X/\sim} \alpha \quad \text{mit} \quad \alpha \neq \emptyset \quad \text{für alle} \quad \alpha \in X/\sim .$$

Beweis. Wegen 30 bleibt nur noch nachzuweisen, dass gilt:

$$\alpha \cap \beta \neq \emptyset \quad \text{für} \quad \alpha, \beta \in X/\!\sim \qquad \Longrightarrow \qquad \alpha = \beta.$$

Seien daher $\alpha, \beta \in X/\!\sim$ mit $z \in \alpha \cap \beta \neq \emptyset$. Wir wählen $x, y \in X$ mit $\alpha = [x]_\sim$ und $\beta = [y]_\sim$.

- Per definitionem gilt $x \sim z$ sowie $y \sim z$, also wegen 2) auch $z \sim y$.
- Mit 1) folgt $x \sim y$, also gilt wegen 2) auch $y \sim x$.

Mit 1) folgt (zweite Implikation):

$$\begin{array}{ccccccccccc} x' \in \alpha & \overset{2)}{\Longrightarrow} & x' \sim x & \overset{x \sim y}{\Longrightarrow} & x' \sim y & \overset{2)}{\Longrightarrow} & y \sim x' & \Longrightarrow & x' \in \beta = [y]_\sim \\ y' \in \beta & \overset{2)}{\Longrightarrow} & y' \sim y & \overset{y \sim x}{\Longrightarrow} & y' \sim x & \overset{2)}{\Longrightarrow} & x \sim y' & \Longrightarrow & y' \in \alpha = [x]_\sim. \end{array}$$

Dies zeigt $\alpha \subseteq \beta$ sowie $\beta \subseteq \alpha$, also $\alpha = \beta$ (wegen Übung 4). □

Beispiel 8. *Sei $S = \{s_1, \ldots, s_n\}$ ($n \geq 2$) eine Packung Smarties. Wir definieren*

$$s_i \sim s_j \quad \text{für} \quad 1 \leq i, j \leq n \qquad \Longleftrightarrow \qquad s_i \text{ und } s_j \text{ haben die gleiche Farbe.}$$

Die Äquivalenzklasse $[s_i]_\sim \in S/\!\sim$ für $1 \leq i \leq n$ besteht dann aus allen Smarties (in S), die die gleiche Farbe wie s_i haben. Insbesondere zeigt dieses Beispiel, dass Äquivalenzklassen im Allgemeinen nicht die gleiche Anzahl von Elemente enthalten müssen.

***Beispiel 1.** *Für $k \in \mathbb{N}$ fixiert, erhalten wir eine Äquivalenzrelation $\sim_k$ auf $\mathbb{Z}$ (Übung 13), vermöge*

$$m \sim_k n \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad m - n \in k\mathbb{Z} := \{n \cdot k \mid n \in \mathbb{Z}\}. \tag{31}$$

(Bemerkung: Gilt $m \sim_k n$ für $m, n \in \mathbb{Z}$, so sagt man auch, dass m und n kongurent modulo k sind, und schreibt $m \equiv n \mod k$.)

Man prüft leicht nach (Übung 13), dass gilt:

$$[m]_{\sim_k} = m + k\mathbb{Z} := \{m + n \cdot k \mid n \in \mathbb{Z}\} \qquad \forall\, m \in \mathbb{Z}. \tag{32}$$

- *Für $k = 0$, ist dann $[m]_{\sim_k} = \{m\}$ einelementig für alle $m \in \mathbb{Z}$; also ist $\mathbb{Z}/{\sim_0}$ (bis auf die Identifikation $m \equiv \{m\}$ für alle $m \in \mathbb{Z}$) identisch mit $\mathbb{Z}$.*
- *Für $k \geq 1$, gilt dann*

  $$\mathbb{Z}/{\sim_k} = \{[0]_{\sim_k}, \dots, [k-1]_{\sim_k}\} \qquad \textit{wegen} \qquad \mathbb{Z} = \dot{\bigcup}_{p=0}^{k-1} (p + k\mathbb{Z}),$$

  und daher $\mathbb{Z}/{\sim_k}| = k$ k.

Für $\alpha, \beta \in \mathbb{Z}/{\sim_k}$, sind dann

$$\begin{aligned} \alpha + \beta &:= [m+n]_{\sim_k} && \textit{für} && m \in \alpha,\ n \in \beta \\ -\,\alpha &:= [-m]_{\sim_k} && \textit{für} && m \in \alpha \\ \alpha \cdot \beta &:= [m \cdot n]_{\sim_k} && \textit{für} && m \in \alpha,\ n \in \beta \end{aligned}$$

wohldefinierte Operationen auf $\mathbb{Z}/{\sim_k}$; denn es gilt

$$\begin{aligned} [m + p \cdot k]_{\sim_k} + [n + q \cdot k]_{\sim_k} &= [m + p \cdot k + n + q \cdot k]_{\sim_k} \\ &= [(m+n) + (p+q) \cdot k]_{\sim_k} \\ &= [m+n]_{\sim_k} \\ &= [m]_{\sim_k} + [n]_{\sim_k} \\ -[n + p \cdot k]_{\sim_k} &= -[n]_{\sim_k} && (\textit{Übung 13}) \\ [m + p \cdot k]_{\sim_k} \cdot [n + q \cdot k]_{\sim_k} &= [m]_{\sim_k} \cdot [n]_{\sim_k} && (\textit{Übung 13}) \end{aligned} \tag{33}$$

für alle $p, q \in \mathbb{Z}$.

Übung 13. *Zeigen Sie, dass* (31) *in Beispiel 1 eine Äquivalenzrelation definiert. Machen Sie sich weiterhin* (32) *klar, und zeigen Sie die letzten beiden Identitäten in* (33).

In Kapitel 1 wurden die rationalen in der Form (2) angegeben. Wir hatten uns dort auf den Standpunkt gestellt, dass wir die rationalen Zahlen bereits kennen – insbesondere also, dass zwei Brüche p/q und p'/q' die selbe rationale Zahl darstellen (also in der Mengendefinition (2) zum selben Element zusammengefasst werden), wenn diese sich durch kürzen ineinander überführen lassen. Mit Hilfe von Äquivalenzrelationen lassen sich die rationalen Zahlen nun präziser wie folgt definieren:

Beispiel 9. *Sei $\mathbb{Z}_{\neq 0} := \mathbb{Z} \setminus \{0\}$, und $Q := \mathbb{Z} \times \mathbb{Z}_{\neq 0}$, und bezeichne $\sim$ die Äquivalenzrelation*

$$(p, q) \sim (p', q') \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad pq' = qp' \tag{34}$$

für $p, p' \in \mathbb{Z}$ und $q, q' \in \mathbb{Z}_{\neq 0}$. Wir setzen

$$\frac{p}{q} := [(p,q)]_{\sim} \qquad \forall\, p \in \mathbb{Z},\ q \in \mathbb{Z}_{\neq 0},$$

sodass also $\frac{p}{q} = \frac{p'}{q'}$ *per definitionem genau dann gilt, wenn* $pq' = qp'$ *erfüllt ist. Die Menge der rationalen Zahlen ist definiert als der Quotient (vgl.* (2)*)*

$$\mathbb{Q} := Q/\sim = \left\{ \frac{p}{q} \,\middle|\, p \in \mathbb{Z},\ q \in \mathbb{Z}_{\neq 0} \right\}.$$

Man identifiziert $n := \frac{n}{1}$ *für alle* $n \in \mathbb{Z}$ *(insbesondere ist dann* $1 = \frac{1}{1}$ *sowie* $0 = \frac{0}{1}$*), und definiert die Rechenoperationen auf* $\mathbb{Q}$ *durch*

$$\frac{p}{q} + \frac{p'}{q'} := \frac{pq' + qp'}{qq'}, \qquad -\frac{p}{q} := \frac{-p}{q}, \qquad \frac{p}{q} \cdot \frac{p'}{q'} := \frac{pp'}{qq'}, \qquad \left(\frac{p''}{q}\right)^{-1} := \frac{q}{p''} \tag{35}$$

für $p, p' \in \mathbb{Z}$, $p'' \in \mathbb{Z}_{\neq 0}$, *und* $q, q' \in \mathbb{Z}_{\neq 0}$*. Unsere Notation täuscht darüber hinweg, dass hier noch etwas nachzuweisen ist, und zwar die Wohldefiniertheit dieser Operationen. Präziser ließt sich z.B. die Definition der Addition nämlich*

$$\alpha + \beta := [(pq' + qp', qq')]_\sim \quad \text{für} \quad \alpha, \beta \in Q/\sim \quad \text{mit} \quad (p, q) \in \alpha \quad \text{und} \quad (p', q') \in \beta.$$

Die rechte Seite obiger Definition (also die Äquivalenzklasse $[(pq' + qp', qq')]_\sim$*) hängt nun aber möglicherweise von der expliziten Wahl der Repräsentanten* $(p, q) \in \alpha$ *und* $(p', q') \in \beta$ *ab. Das dies jedoch nicht der Fall ist, sieht man wie folgt:*

Seien $\tilde{p}, \tilde{p}' \in \mathbb{Z}$ *und* $\tilde{q}, \tilde{q}' \in \mathbb{N}_{>0}$ *gegeben, mit*

$$(\tilde{p}, \tilde{q}) \in \alpha = [(p, q)]_\sim \qquad \text{und} \qquad (\tilde{p}', \tilde{q}') \in \beta = [(p', q')]_\sim.$$

Wegen (34) *gilt dann* $p\tilde{q} = q\tilde{p}$ *sowie* $p'\tilde{q}' = q'\tilde{p}'$*; und wir müssen hieraus Schlussfolgern, dass*

$$(\tilde{p}\tilde{q}' + \tilde{q}'\tilde{p}, \tilde{q}\tilde{q}') \sim (pq' + q'p, qq')$$

gilt, bzw. umgeschrieben

$$\frac{\tilde{p}\tilde{q}' + \tilde{q}'\tilde{p}}{\tilde{q}\tilde{q}'} = \frac{pq' + q'p}{qq'}.$$

Hierfür rechnen wir

$$\begin{aligned}
(\tilde{p}\tilde{q}' + \tilde{q}'\tilde{p}) \cdot (qq') &= \tilde{p}\tilde{q}'qq' + \tilde{q}'\tilde{p}qq' \\
&= 2(q\tilde{p})q'\tilde{q}' \\
&= 2(p\tilde{q})q'\tilde{q}' \\
&= 2pq'\tilde{q}\tilde{q}' \\
&= pq'\tilde{q}\tilde{q}' + q'p\tilde{q}\tilde{q}' \\
&= (pq' + q'p) \cdot (\tilde{q}\tilde{q}'),
\end{aligned}$$

was die Behauptung zeigt.

Übung 14. *Zeigen Sie, dass* (34) *eine Äquivalenzrelation definiert, und weisen Sie die Wohldefiniertheit der übrigen Rechenoperationen in* (35) *nach.*

3.2.2 Ordnungen

Definition 10. *Eine Ordnungsrelation auf einer Menge* X *ist eine Relation* R *auf* X*, sodass für alle* $x, y, z \in X$ *die folgenden Bedingungen gelten:*

1) Transitivität: $\quad xRy \wedge yRz \;\Rightarrow\; xRz,$

2) Antisymmetrie: $xRy \wedge yRx \;\Rightarrow\; x = y,$

3) Reflexivität: $xRx.$

Die Ordnungsrelation R heißt Anordnung, falls zusätzlich folgende Bedingung gilt:

4) Totalität: Für alle $x, y \in X$ gilt xRy *oder* yRx.

Ordnungsrelationen notiert man üblicherweise mit $\leq$ anstatt mit R. Man schreibt $x \leq y$ oder $y \geq x$ anstatt xRy. Weiter kürzt man

$$\begin{array}{llll} (x \leq y \;\wedge\; x \neq y) & \text{mit} & x < y & \text{ab} \\ (x \geq y \;\wedge\; x \neq y) & \text{mit} & x > y & \text{ab.} \end{array}$$

Terminologie 6.

- *Ist $\leq$ eine Ordnungsrelation auf X, so spricht man von der teilgeordneten Menge $(X, \leq)$. Zwei Elemente $x, y \in X$ heißen miteinander vergleichbar, wenn $x \leq y$ oder $y \leq x$ gilt.*
- *Ist $\leq$ eine Anordnung, so spricht man von der angeordneten Menge $(X, \leq)$.*

Beispiel 10. *Ist Z eine Menge, so ist $(\mathcal{P}(Z), \subseteq)$ eine teilgeordnete Menge.*

(In obiger Notation ist also $X = \mathcal{P}(Z)$ und $\leq = \subseteq$, sodass $x \leq y$ für $x, y \in X$ (also $x, y \subseteq Z$) genau dann gilt, wenn $x \subseteq y$ erfüllt ist.)

Beispiel 11. *Sei $X = \mathbb{Q}$ die Menge der rationalen Zahlen. Wir setzen*

$$\frac{p}{q} \geq 0 \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad pq \geq 0 \qquad\qquad \textit{für alle } \frac{p}{q} \in \mathbb{Q}.$$

Dies ist wohldefiniert (vgl. Beispiel 9), denn für $p, p' \in \mathbb{Z}$ und $q, q' \in \mathbb{Z}_{\neq 0}$ mit $pq' = qp'$ gilt

$$pq \geq 0 \qquad \Longleftrightarrow \qquad pq(q')^2 \geq 0 \qquad \Longleftrightarrow \qquad p'q'q^2 \geq 0 \qquad \Longleftrightarrow \qquad p'q' \geq 0.$$

Für $x, y \in \mathbb{Q}$, definiert man

$$x \leq y \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad y - x \geq 0.$$

Dann ist $(\mathbb{Q}, \leq)$ eine angeordnete Menge, und für alle $x, y, z \in \mathbb{Q}$ gilt

$$x \leq y \quad \Longrightarrow \quad x + z \leq y + z \qquad \textit{sowie} \qquad x, y \geq 0 \quad \Longrightarrow \quad xy \geq 0. \tag{36}$$

Der Betrag einer rationalen Zahl $x \in \mathbb{Q}$ wird definiert durch *($|x| = \max(x, -x)$)*

$$|x| := \begin{cases} x & \textit{für } x \geq 0 \\ -x & \textit{für } x < 0. \end{cases} \tag{37}$$

Übung 15. *Zeigen Sie, dass $(\mathbb{Q}, \leq)$ eine angeordnete Menge ist, und weisen Sie die Implikationen in (36) nach. Zeigen Sie weiterhin, dass $x^2 = 0$ für $x \in \mathbb{Q}$ genau dann gilt, wenn $x = 0$ ist.*

Übung 16. *Benutzen Sie Übung 15, um die folgenden Aussagen für $x, y, a, b \in \mathbb{Q}$ nachzuweisen:*

1) $x \leq y \;\wedge\; a \leq b \quad \Longrightarrow \quad x + a \leq y + b$
2) $x < y \;\wedge\; a \leq b \quad \Longrightarrow \quad x + a < y + b$
3) $x \leq y \;\wedge\; a \geq 0 \quad \Longrightarrow \quad ax \leq ay$
4) $x < y \;\wedge\; a > 0 \quad \Longrightarrow \quad ax < ay$

$$
\begin{array}{lrcl}
5) & 0 \le x < y \ \wedge\ 0 \le a < b & \Longrightarrow & 0 \le ax < by \\
6) & x < y & \Longrightarrow & -y < -x \\
7) & y \le x \ \wedge\ a \le 0 & \Longrightarrow & ax \le ay \\
8) & x \ne 0 & \Longrightarrow & x^2 > 0 \\
9) & 1 > 0 & & \\
10) & x > 0 & \Longrightarrow & x^{-1} > 0 \\
11) & 0 < x < y & \Longrightarrow & 0 < y^{-1} < x^{-1}
\end{array}
$$

Hinweis: Gehen Sie in der gegebenen Reihenfolge vor.

Notation 5. *Sei $(X, \le)$ eine teilgeordnete Menge, $x \in X$ eine Element, und $Y \subseteq X$ eine Teilmenge.*

- *Wir schreiben $Y \le x$ (oder $x \ge Y$) genau dann, wenn $y \le x$ für alle $y \in Y$ gilt.*
- *Wir schreiben $x \le Y$ (oder $Y \ge x$) genau dann, wenn $x \le y$ für alle $y \in Y$ gilt.*

Bemerkung 12. *Sei $(X, \le)$ eine teilgeordnete Menge, und $Z, \tilde{Z} \subseteq X$ Teilmengen mit $Z \subseteq \tilde{Z}$. Offensichtlich gilt dann für jedes $x \in X$:*

$$
\begin{array}{ccc}
\tilde{Z} \le x & \Longrightarrow & Z \le x \\
x \le \tilde{Z} & \Longrightarrow & x \le Z.
\end{array}
\tag{38}
$$

Definition 11. *Sei $(X, \le)$ eine teilgeordnete Menge, und $x \in X$ ein Element.*

1) *Gilt $X \le x$, so heißt x größtes Element von X.*
2) *Existiert kein $y \in X$ mit $y > x$, so heißt x maximal (bzw. maximales Element von X).*
3) *Gilt $x \le X$, so heißt x kleinstes Element von X.*
4) *Existiert kein $y \in X$ mit $x > y$, so heißt x minimal (bzw. minimales Element von X).*

Bemerkung 13. *Sei $(X \le)$ eine teilgeordnete Menge.*

a) *Ein größtes/kleinstes Element von $(X, \le)$ ist notwendigerweise eindeutig. Es ist daher legitim von dem größten/kleinsten Element von X zu sprechen (sofern es existiert).*

   Beweis der Eindeutigkeit: Seien x, y beides größte/kleinste Elemente von X. Dann gilt (in beiden Fällen) sowohl $x \le y$ als auch $y \le x$, also $x = y$ nach 2) in Definition 10. ☐

b) *Ein größtes/kleinstes Element ist automatisch maximal/minimal.*[7] *Existiert ein größtes/kleinstes Element, so kann es keine weiteren maximalen/minimalen Elemente geben.*

   Beweis der Behauptungen:

   - Sei $x \in X$ größtes Element von X. Gilt $x < y$ für ein $y \in X$, so auch $x \le y$; und per Annahme aber auch $y \le x$. Wegen 2) folgt $x = y$.
   - Sei $x \in X$ kleinstes Element von X. Gilt $y < x$ für ein $y \in X$, so auch $y \le x$; und per Annahme aber auch $x \le y$. Wegen 2) folgt $x = y$. ☐

c) *Es sei an dieser Stelle explizit darauf hingewiesen, dass in teilgeordneten Mengen nicht alle Elemente miteinander vergleichbar sein müssen. Existieren z.B. mehrere maximale (oder mehrere minimale) Elemente, so können diese auch nicht miteinander vergleichbar sein:*

   Beweis der Behauptungen:

[7] Erinnerung: Nach unseren Konventionen impliziert $y > x$ bzw. $y < x$ für $x, y \in X$ insbesondere, dass $y \ge x$ bzw. $y \le x$ gilt.

- Seien $x, y \in X$ maximal. Aus $x \leq y$ (bzw. $y \leq x$) folgt dann bereits $x = y$, weil $x < y$ (bzw. $y < x$) nicht gelten kann.
- Seien $x, y \in X$ minimal. Aus $x \leq y$ (bzw. $y \leq x$) folgt dann bereits $x = y$, weil $x < y$ (bzw. $y < x$) nicht gelten kann. □

Existiert allerdings ein größtes/kleinstes Element, so ist dieses per Definition mit allen anderen Elementen von X vergleichbar.

d) *Ein Element ist automatisch sowohl maximal als auch minimal, wenn es mit keinem anderen Element vergleichbar ist.*

Beispiel 12.

- *Sei Z eine Menge, und $\mathcal{P}(Z)$ teilgeordnet durch die Inklusionsrelation – wir haben also $X = \mathcal{P}(Z)$ sowie $\leq = \subseteq$. Dann ist Z das größte Element von X, und $\emptyset$ das kleinste Element von X.*
- *Sei $Z = \{1, 2, 3, 4, 5, 6\}$, sowie*

$$X = \{\underbrace{\{1,3,4\}}_{x_1}, \underbrace{\{1,2,3\}}_{x_2}, \underbrace{\{1,4\}}_{x_3}, \underbrace{\{1,3\}}_{x_4}, \underbrace{\{5\}}_{x_5}\} \subset \mathcal{P}(Z) \qquad \textit{und} \qquad \leq = \subseteq .$$

Die Menge X hat kein größtes und kein kleinstes Element. Die Elemente x_1, x_2, x_5 sind maximal, und die Elemente x_3, x_4, x_5 sind minimal.

Lemma 2. *Sei $(X, \leq)$ eine angeordnete Menge. Dann ist $x \in X$ größtes/kleinstes Element von X genau dann, wenn x maximal/minimal ist.*

Beweis. Ist x größtes/kleinstes Element von X, so ist x maximal/minimal wegen Bemerkung 13.b).

- Sei x maximal. Für $y \in X$ gilt dann $x < y$ oder $y \leq x$, und zwar wegen 4). Da $x < y$ nicht gelten kann, folgt $y \leq x$. Dies zeigt $X \leq x$.
- Sei x minimal. Für $y \in X$ gilt dann $x \leq y$ oder $y < x$, und zwar wegen 4). Da $y < x$ nicht gelten kann, folgt $x \leq y$. Dies zeigt $x \leq X$. □

Definition 12 (Induzierte Ordnung)**.** *Sei $(X, \leq)$ eine teilgeordnete/angeordnete Menge, und $Y \subseteq X$ eine Teilmenge. Dann ist Y in natürlicher Weise teilgeordnet/angeordnet (Übung) durch*

$$y \leq_Y \tilde{y} \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad y \leq \tilde{y} \qquad\qquad \forall\, y, \tilde{y} \in Y,$$

wobei wir auf der rechten Seite $y, \tilde{y}$ als Elemente in X aufgefasst haben. Ist aus dem Kontext heraus klar, als Teilmenge welcher teilgeordneten Menge $(X, \leq)$ die Menge Y zu verstehen ist, so schreibt man auch einfach $\leq$ anstatt $\leq_Y$. Unter diesem Gesichtspunkt ist dann auch der folgende Sprachgebrauch zu verstehen:

- *Ein kleinstes/größtes Element von Y, ist ein kleinstes/größtes Element von $(Y, \leq_Y)$.*
- *Ein minimales/maximales Element von Y, ist ein minimales/maximales Element von $(Y, \leq_Y)$.*

(Bemerkung: Formeller ausgedrückt, ist die Relation (Teilordnung/Anordnung) $\leq_Y \subseteq Y \times Y$ gegeben als der Schnitt $\leq \cap (Y \times Y)$ der Relation (Teilordnung/Anordnung) $\leq \subseteq X \times X$ mit der Teilmenge $Y \times Y \subseteq X \times X$.)

Bemerkung 14. *Sei $(X, \leq)$ eine teilgeordnete Menge, sowie $A, B \subseteq X$ Teilmengen mit $A \subseteq B$.*

a) *Sei $b \in B$ größtes Element von $(B, \leq_B)$; und sei $a \in A$ größtes Element von $(A, \leq_A)$. Wegen $a \in A \subseteq B \leq b$, gilt dann $a \leq b$.*

b) Sei $b \in B$ kleinstes Element von $(B, \leq_B)$; und sei $a \in A$ kleinstes Element von $(A, \leq_A)$. Wegen $b \leq B \supseteq A \ni a$, gilt dann $b \leq a$.

Terminologie 7 (Maximum und Minimum)**.** *Sei $(X, \leq)$ eine angeordnete Menge, und $Y \subseteq X$ eine Teilmenge. Dann ist $(Y, \leq_Y)$ angeordnet; also fallen gemäß Lemma 2 die Begriffe maximales/minimales und größtes/kleinstes Element von Y zusammen.*

- *Existiert ein maximales (größtes) Element von $(Y, \leq_Y)$, so wird dieses mit* $\max(Y)$ *notiert, und als das Maximum von Y bezeichnet.*
- *Existiert ein minimales (kleinstes) Element von $(Y, \leq_Y)$, so wird dieses mit* $\min(Y)$ *notiert, und als das Minimum von Y bezeichnet.*

Ist $Y = \{y_1, \dots, y_n\}$ mit $n \geq 1$ eine nichtleere endliche Teilmenge, so schreibt man auch

$$\begin{aligned} \max(y_1, \dots, y_n) &:= \max(\{y_1, \dots, y_n\}) = \max(Y) \\ \min(y_1, \dots, y_n) &:= \min(\{y_1, \dots, y_n\}) = \min(Y). \end{aligned} \tag{39}$$

Lemma 3.1) stellt sicher, das die rechte Seite von (39) *(im angeordneten Fall) immer existiert.*

Lemma 3. *Sei $(X, \leq)$ eine angeordnete Menge.*

1) Ist $Y \subseteq X$ nichtleer und endlich, so existieren $\min(Y)$ *und* $\max(Y)$.

2) Seien $A, B \subseteq X$ Teilmengen.

- *Existieren* $\max(A)$ *und* $\max(B)$*, so gilt die Implikation*

$$A \subseteq B \qquad \Longrightarrow \qquad \max(A) \leq \max(B).$$

- *Existieren* $\min(A)$ *und* $\min(B)$*, so gilt die Implikation*

$$A \subseteq B \qquad \Longrightarrow \qquad \min(B) \leq \min(A).$$

3) Seien $A, B \subseteq X$ Teilmengen.

- *Existieren* $\max(A)$ *und* $\max(B)$*, dann existiert* $\max(A \cup B)$*, und es gilt*

$$\max(A \cup B) = \max(\max(A), \max(B)).$$

- *Existieren* $\min(A)$ *und* $\min(B)$*, dann existiert* $\min(A \cup B)$*, und es gilt*

$$\min(A \cup B) = \min(\min(A), \min(B)).$$

Beweis. 1) (IA): Ist $Y = \{y\}$ einelementig, so gilt $\min(Y) = y = \max(Y)$. (IV): Sei $n \geq 1$ gegeben, sodass die Behauptung für alle n-elementigen Teilmengen von X gilt. (IS): Sei $Y = \{y_0, \dots, y_n\}$ eine $(n+1)$-elementige Teilmengen von X.

- Da $\leq$ eine Anordnung ist, gilt $y_0 \geq y_i$ oder $y_0 \leq y_i$ für ein gegebenes $1 \leq i \leq n$.
  - Ist $y_0 \geq y_1, \dots, y_n$, so gilt $\max(Y) = y_0$.
  - Existiert ein $1 \leq i \leq n$ mit $y_0 \leq y_i$, so gilt $y_0 \leq y_i \leq \max(y_1, \dots, y_n)$ (existiert nach (IV)). Es folgt $y_0, \dots, y_n \leq \max(y_1, \dots, y_n)$, also $\max(Y) = \max(y_1, \dots, y_n)$.
- Da $\leq$ eine Anordnung ist, gilt $y_0 \leq y_i$ oder $y_0 \geq y_i$ für ein gegebenes $1 \leq i \leq n$.
  - Ist $y_0 \leq y_1, \dots, y_n$, so gilt $\min(Y) = y_0$.
  - Existiert ein $1 \leq i \leq n$ mit $y_0 \geq y_i$, so gilt $y_0 \geq y_i \geq \min(y_1, \dots, y_n)$ (existiert nach (IV)). Es folgt $y_0, \dots, y_n \geq \min(y_1, \dots, y_n)$, also $\min(Y) = \min(y_1, \dots, y_n)$.

2) Klar wegen Bemerkung 14 und Lemma 2.

3) • Wegen Teil 1) existiert

$$\alpha := \max(\max(A), \max(B)) \in \{\max(A), \max(B)\} \subseteq A \cup B.$$

Dann gilt $A \leq \max(A) \leq \alpha$ sowie $B \leq \max(B) \leq \alpha$, also $A \cup B \leq \alpha \in A \cup B$.

• Wegen Teil 1) existiert

$$\alpha := \min(\min(A), \min(B)) \in \{\min(A), \min(B)\} \subseteq A \cup B.$$

Dann gilt $A \geq \min(A) \geq \alpha$ sowie $B \geq \min(B) \geq \alpha$, also $A \cup B \geq \alpha \in A \cup B$. □

Übung 17. *Entscheiden Sie jeweils, ob Maximum bzw. Minimum der folgenden Teilmengen der angeordneten Menge $(\mathbb{Q}, \leq)$ existieren, und bestimmen Sie diese im Falle der Existenz:*

$$\begin{gathered} M_0 = \left\{\tfrac{1}{n} \mid n \in \mathbb{N}_{>0}\right\}, \quad M_1 = \left\{n - \tfrac{1}{n} \mid n \in \mathbb{N}_{>0}\right\}, \quad M_2 = \left\{\tfrac{(-1)^n}{n} \mid n \in \mathbb{N}_{>0}\right\}, \\ M_3 = \{x \in \mathbb{Q} \mid 0 < x \leq 1\}, \quad M_4 = \emptyset. \end{gathered} \tag{40}$$

Hinweis: Sie können die Aussagen aus Übung 16 frei verwenden.

Definition 13. *Sei $(X, \leq)$ eine teilgeordnete Menge, und $Y \subseteq X$ eine Teilmenge.*

1) Ein Element $o \in X$ heißt obere Schranke der Menge Y, wenn $Y \leq o$ gilt. Wir notieren die Menge aller oberen Schranken von Y in X mit

$$O_X(Y) := \{o \in X \mid Y \leq o\}.$$

- *Y heißt nach oben beschränkt (in X), wenn $O_X(Y) \neq \emptyset$ gilt.*
- *Besitzt $O_X(Y)$ ein kleinstes Element (bezüglich der induzierten Ordnung auf $O_X(Y)$), so wird dieses als das Supremum (oder die kleinste obere Schranke) von Y bezeichnet, und mit $\sup_X(Y)$ notiert (dieses ist eindeutig wegen Bemerkung 13.a)).*

  (Ist aus dem Kontext heraus klar, bezüglich welcher Menge X das Supremum von Y gebildet wird, so schreibt man auch einfach $\sup(Y)$ anstelle $\sup_X(Y)$.)

2) Ein Element $u \in X$ heißt untere Schranke der Menge Y, wenn $u \leq Y$ gilt. Wir notieren die Menge aller unteren Schranken von Y in X mit

$$U_X(Y) := \{u \in X \mid u \leq Y\}.$$

- *Y heißt nach unten beschränkt (in X), wenn $U_X(Y) \neq \emptyset$ gilt.*
- *Besitzt $U_X(Y)$ ein größtes Element (bezüglich der induzierten Ordnung auf $U_X(Y)$), so wird dieses als das Infimum (oder die größte untere Schranke) von Y bezeichnet, und mit $\inf_X(Y)$ notiert (dieses ist eindeutig wegen Bemerkung 13.a)).*

  (Ist aus dem Kontext heraus klar, bezüglich welcher Menge das Infimum von Y gebildet wird, so schreibt man auch einfach $\inf(Y)$ anstelle $\inf_X(Y)$.)

3) Die Menge Y heißt beschränkt (in X), wenn $U_X(Y) \neq \emptyset \neq O_X(Y)$ gilt.

Beispiel 13. *Wie in Beispiel 12 (zweiter Teil) sei $Z = \{1, 2, 3, 4, 5, 6\}$, sowie*

$$Y = \{\{1,3,4\}, \{1,2,3\}, \{1,4\}, \{1,3\}, \{5\}\} \subset X = \mathcal{P}(Z) \qquad \textit{und} \qquad \leq = \subseteq.$$

Dann ist (vgl. Übung 20 unten)

$$\sup_X(Y) = \{1, 2, 3, 4, 5\} \qquad \textit{sowie} \qquad \inf_X(Y) = \{\emptyset\}.$$

Lemma 4. *Sei $(X, \leq)$ eine teilgeordnete Menge, und $Y, Z \subseteq X$ Teilmengen.*

1) Existieren $\sup_X(Y)$ und $\sup_X(Z)$, so gilt

$$Y \subseteq Z \qquad \Longrightarrow \qquad \sup_X(Y) \leq \sup_X(Z).$$

2) Existieren $\inf_X(Y)$ und $\inf_X(Z)$, so gilt

$$Y \subseteq Z \qquad \Longrightarrow \qquad \inf_X(Z) \leq \inf_X(Y).$$

Beweis. 1) Es gilt $O_X(Z) \subseteq O_X(Y)$ wegen (38). Die Behauptung folgt nun unmittelbar aus Bemerkung 14.b).

2) Es gilt $U_X(Z) \subseteq U_X(Y)$ wegen (38). Die Behauptung folgt nun unmittelbar aus Bemerkung 14.a). $\square$

Lemma 5. *Sei $(X, \leq)$ eine teilgeordnete Menge, und $Y \subseteq X$ eine Teilmenge. Ist y größtes (kleinstes) Element von $(Y, \leq_Y)$, so ist y Supremum (Infimum) von Y in X, also $\sup_X(Y) = y$ ($\inf_Y(x) = y$).*

Beweis. Sei y größtes Element von $(Y, \leq_Y)$. Dann ist y eine obere Schranke von Y in X. Sei nun $o \in O_X(Y)$ eine beliebige obere Schranke von Y in X. Wegen $y \in Y$ gilt dann $y \leq o$; also ist y tatsächlich das kleinste Element von $O_X(Y)$. Analog zeigt man, dass das kleinste Element von Y das Infimum von Y in X ist. $\square$

Übung 18. *Sei $(X, \leq)$ eine teilgeordnete Menge, und $Y \subseteq X$ eine Teilmenge. Zeigen Sie:*

- *Gilt $z := \sup_X(Y) \in Y$, so ist z größtes Element von Y.*
- *Gilt $z := \inf_X(Y) \in Y$, so ist z kleinstes Element von Y.*

Beispiel 14. *Sei $X = \mathbb{Q}$, sowie*

$$Y = \{q \in \mathbb{Q} \mid q < 1\} \qquad \text{und} \qquad \tilde{Y} = \{q \in \mathbb{Q} \mid q \leq 1\}.$$

Die Teilmenge $Y \subseteq \mathbb{Q}$ hat kein größtes Element, aber es gilt $\sup_{\mathbb{Q}}(Y) = 1$. Die Teilmenge $\tilde{Y} \subseteq \mathbb{Q}$ besitzt das größtes Element 1, und somit gilt auch $\sup_{\mathbb{Q}}(\tilde{Y}) = 1$. Beide Teilmengen Y und $\tilde{Y}$ haben keine untere Schranke in $\mathbb{Q}$, und daher auch keine größten untere Schranke.

Übung 19. *Entscheiden Sie, welche der Mengen in (40) ein Supremum bzw. ein Infimum in $\mathbb{Q}$ haben, und bestimmen Sie diese im Falle der Existenz.*

Übung 20. *Sei Z eine Menge, und $\mathcal{P}(Z)$ teilgeordnet durch die Inklusionsrelation (vgl. Beispiel 12). Gegeben eine Teilmenge $Y \subseteq \mathcal{P}(Z)$. Zeigen Sie*

$$\sup_{\mathcal{P}(Z)}(Y) = \bigcup_{A \in Y} A \qquad \text{sowie} \qquad \inf_{\mathcal{P}(Z)}(Y) = \bigcap_{A \in Y} A.$$

Wir wollen abschließend das Konzept des Intervalls behandeln.

Definition 14. *Sei $(X, \leq)$ eine teilgeordnete Menge. Eine Teilmenge $I \subseteq X$ heißt Intervall (in X), falls die folgende Implikation gilt:*

$$x \leq y \leq z \quad \text{für} \quad x, z \in I,\ y \in X \qquad \Longrightarrow \qquad y \in I.$$

(Mit je zwei Punkten, enthält I auch jeden dazwischenliegenden Punkt.)

Sei $(X, \leq)$ eine teilgeordnete Menge. Die leere Menge $\emptyset$ ist ein Intervall in X; und genauso jede einelementige Teilmenge $\{x\} \subseteq X$ (sofern $X \neq \emptyset$ gilt). Allgemeiner sind für $x, z \in X$ vorgegeben, die folgenden Teilmengen Intervalle in X:

$$\begin{aligned}
[x,z] &:= \{y \in X \mid x \leq y \leq z\} && \text{(abgeschlossenes Intervall)}\\
(x,z) &:= \{y \in X \mid x < y < z\} && \text{(offenes Intervall)}\\
(x,z] &:= \{y \in X \mid x < y \leq z\} && \text{(linksoffenes Intervall)}\\
[x,z) &:= \{y \in X \mid x \leq y < z\} && \text{(rechtsoffenes Intervall)}\\
(x,\infty) &:= \{y \in X \mid x < y\}\\
[x,\infty) &:= \{y \in X \mid x \leq y\}\\
(-\infty,z) &:= \{y \in X \mid y < z\}\\
(-\infty,z] &:= \{y \in X \mid y \leq z\}\\
(-\infty,\infty) &:= X.
\end{aligned} \tag{41}$$

Die Intervalleigenschaft folgt hier unmittelbar aus der Transitivitätseigenschaft 1).[8] Das Unendlichkeitssymbol ∞ ist hier lediglich als Notationshilfe zu interpretieren.

Übung 21. *Sei $(X, \leq)$ eine teilgeordnete Menge, und seien $I, J \subseteq X$ Intervalle in X. Zeigen Sie, dass $I \cap J \subseteq X$ ein Intervall in X ist.*

Hinweis: Unterscheiden Sie die Fälle $I \cap J = \emptyset$ und $I \cap J \neq \emptyset$.

Übung 22. *Sei $(X, \leq)$ eine angeordnete Menge, $J \neq \emptyset$ eine nichtleere Indexmenge, und $(I_j)_{j \in J}$ eine Familien von Intervallen in X. Zeigen Sie, dass $\tilde{I} := \bigcup_{j \in J} I_j \subseteq X$ ein Intervall in X ist, sofern $\bigcap_{j \in J} I_j \neq \emptyset$ gilt.*

Hinweis: Betrachten Sie $x, z \in \tilde{I}$ und $y \in X$ mit $x \leq y \leq z$. Wählen Sie $j, j' \in J$ mit $x \in I_j$ und $y \in I_{j'}$. Benutzen Sie nun die Existenz eines Elements in $I_j \cap I_{j'}$ (Warum ist diese Menge nicht leer?) sowie die Anordnungseigenschaft von $(X, \leq)$, um $y \in I_j \ \vee \ y \in I_{j'}$ zu zeigen (hierbei ist es sinnvoll, zwischen den beiden Fällen $j \neq j'$ und $j = j'$ zu unterscheiden).

Übung 23. *Sei $(X, \leq)$ eine nichtleere angeordnete Menge. Ein Intervall $I \subseteq X$ heißt maximal, wenn für jedes Intervall $J \subseteq X$ mit $I \cap J \neq \emptyset$, bereits $J \subseteq I$ gilt. Es bezeichne $\mathcal{I} \subseteq \mathcal{P}(x)$ die Menge aller maximalen Intervalle in X. Für $x \in X$, bezeichne $\mathcal{I}_x \subseteq \mathcal{P}(X)$ die Menge aller Intervalle in X, die x als Element enthalten.*

- *Zeigen Sie, dass zwei maximale Intervalle entweder gleich oder disjunkt sind.*
- *Sei $x \in X$. Folgern Sie aus Übung 22, dass $I_x := \bigcup_{I \in \mathcal{I}_x} I \subseteq X$ ein Intervall ist, das x enthält (Warum gilt $\mathcal{I}_x \neq \emptyset$?). Zeigen Sie nun, dass I_x maximal ist – z.B. indem Sie geschickt durch Widerspruch argumentieren und Übung 22 anwenden.*
- *Folgern Sie aus dem bereits Gezeigten, dass $X = \dot{\bigcup}_{I \in \mathcal{I}} I$ gilt – dass also X die disjunkte Vereinigung aller in X enthaltenen maximalen Intervalle ist.*

3.3 Abbildungen zwischen Mengen

In diesem Abschnitt behandeln wir Abbildung zwischen Mengen, sowie das Konzept der Mächtigkeit von Mengen.

[8] Gilt beispielsweise $a, b \in [x, z]$ sowie $a \leq c \leq b$ für $c \in X$, so folgt bereits $x \leq a \leq c \leq b \leq z$, also $c \in [x, z]$.

3.3.1 Abbildungen

Seien X und Y Mengen. Eine Abbildung $f\colon X \to Y$ von X nach Y ist eine Vorschrift, die jedem Element $x \in X$ genau ein Element $f(x) \in Y$ zuordnet. Abbildungen werden gleichwertig oft auch als Funktionen bezeichnet. Es scheint allerdings auch eine weit verbreitete Konvention zu sein, den Begriff Funktion hauptsächlich für Abbildungen zwischen Zahlenmengen zu verwenden.

Formeller definiert man das Konzept der Abbildungen wie folgt:

Definition 15. *Eine Abbildung f ist ein Tripel (X, Y, Γ_f) bestehend aus Mengen X und Y sowie einer Teilmenge $\Gamma_f \subseteq X \times Y$ (Relation zwischen X und Y), sodass gilt:*

$$\forall\, x \in X\colon \exists!\, y \in Y\colon (x, y) \in \Gamma_f. \tag{42}$$

Wir haben die folgenden Notationen und Terminologien:

1) Die Menge X wird als Definitionsbereich *der Abbildung f bezeichnet. Man notiert sie auch mit* $\mathrm{dom}(f)$ *(domain).*

Die Menge Y wird als Werte- oder Bildbereich *von f bezeichnet.*

Die Relation Γ_f heißt der Graph *von f.*

Zwei Abbildungen sind daher genau dann gleich, wenn sowohl deren Definitionsbereiche, Wertebereiche, als auch deren Graphen miteinander übereinstimmen.

2) Das dem Element $x \in X$ durch die Bedingung (42) *eindeutig zugeordnete Element $y \in Y$ wird mit $f(x)$ notiert, und als das Bild von x unter f bezeichnet, bzw. als der Wert von f an der Stelle x. Man spricht dann auch vom Auswerten der Funktion f an der Stelle x, oder vom Einsetzen von x in f. Mit besagter Notation gilt dann*

$$\Gamma_f = \{(x, y) \in X \times Y \mid f(x) = y\}.$$

3) Das Bild $f(Z)$ einer Teilmenge $Z \subseteq X$ unter f, ist definiert durch

$$\begin{aligned} f(Z) &:= \{f(z) \mid z \in Z\} \\ &= \{y \in Y \mid \exists\, z \in Z\colon (z, y) \in \Gamma_f\} \subseteq Y. \end{aligned}$$

Das Bild von f ist definiert durch $\mathrm{im}(f) := f(X) \subseteq Y$ *(image).*

4) Das Urbild $f^{-1}(W)$ einer Teilmenge $W \subseteq Y$ unter f, ist definiert durch

$$\begin{aligned} f^{-1}(W) &:= \{x \in X \mid f(x) \in W\} \\ &= \{x \in X \mid \exists\, w \in W\colon (x, w) \in \Gamma_f\} \subseteq X. \end{aligned}$$

Für ein Element $y \in Y$, setzt man $f^{-1}(y) := f^{-1}(\{y\})$. Ist $f^{-1}(\{y\}) = \{x\}$ einelementig, so identifiziert man $f^{-1}(\{y\}) = \{x\}$ auch gerne mit dem Element x. (Dies geschieht aus notationellen Konsistenzgründen im Hinblick auf die Definition der Umkehrabbildung – siehe Definition 20.3) und Notation 9.)

Bemerkung 15. *Sei $f = (X, Y, \Gamma_f)$ eine Abbildung.*

- *Ist $Y = \emptyset$, so kann* (42) *nur für $X = \emptyset$ erfüllt sein, sodass dann also $f = (\emptyset, \emptyset, \emptyset)$ gilt.*
- *Ist $X = \emptyset$, so ist* (42) *automatisch erfüllt. Also ist $f = (\emptyset, Y, \emptyset)$ eine Abbildung (man beachte $\emptyset \times Y = \emptyset$ wegen Bemerkung 9.1)).*

Notation 6. *Wir notieren Abbildungen oft in der Form*

$$\begin{aligned} f\colon X &\to Y \\ x &\mapsto f(x), \end{aligned}$$

sowie in verschiedenen Variationen und Verkürzungen dieser Notation. Beispielsweise schreiben wir

$$f\colon X \to Y, \qquad x \mapsto f(x),$$

oder noch prägnanter: $f\colon X \ni x \mapsto f(x) \in Y$*. Manchmal sprechen wir auch einfach von einer Abbildung* $f\colon X \to Y$*, oder sogar nur von einer Abbildung* $X \to Y$*. In allen diesen Fällen sind* X *und* Y *automatisch als Mengen zu verstehen, auch wenn dies im Vorfeld nicht explizit erwähnt wurde.*

Notation 7. *Gegeben Mengen* X *und* Y*, so wird der Raum aller Abbildungen* $X \to Y$ *mit* $\mathsf{Abb}(X,Y)$ *notiert. Vermöge der Bedingung (42) lässt sich* $\mathsf{Abb}(X,Y)$ *auch als Teilmenge von* $\mathcal{P}(X \times Y)$ *auffassen, ist daher also selbst eine Menge.*

Beispiel 15.

- *Die Funktion* $f\colon \mathbb{Z} \ni x \mapsto x^2 \in \mathbb{Z}$ *hat den Graphen* $\Gamma_f = \{(x,x^2) \in \mathbb{Z}^2 \mid x \in \mathbb{Z}\}$*, und es gilt* $f(\mathbb{Z}) = \operatorname{im}(f) \subseteq \mathbb{N} = f(\mathbb{N})$.
- *Es gibt keine Funktion* $g\colon \mathbb{Z} \to \mathbb{Z}$ *mit* $\Gamma_g = \{(x^2,x) \in \mathbb{Z}^2 \mid x \in \mathbb{Z}\}$*. In der Tat gilt ja für* $x \neq 0$*, sowohl* $(x^2,x) \in \Gamma_g$ *als auch* $(x^2,-x) = ((-x)^2,-x) \in \Gamma_g$*, sodass (42) nicht erfüllt sein kann.*

  (Man mache sich anhand einer Skizze klar, dass Γ_g *durch Rotation von* Γ_f *um* $90°$ *im Uhrzeigersinn hervorgeht.)*

Allgemeinere Beispiele sind:

Beispiel 16. *Seien* X *und* Y *Mengen.*

- *Die Identitätsabbildung auf* X *(die Identität auf* X*), ist definiert durch*

  $$\mathrm{id}_X\colon X \to X, \qquad x \mapsto x.$$

  Wir haben also $\Gamma_{\mathrm{id}_X} = \{(x,x) \mid x \in X\}$*, sowie* $\operatorname{dom}(\mathrm{id}_X) = X = \operatorname{im}(\mathrm{id}_X)$.
- *Gegeben ein Element* $y \in Y$*, so ist die konstante Abbildung von* X *auf* y*, definiert durch*

  $$f[y]\colon X \to Y, \qquad x \mapsto y$$

  Wir haben also $\Gamma_{f[y]} = \{(x,y) \mid x \in X\}$*, sowie* $\operatorname{dom}(f[y]) = X$ *und* $\operatorname{im}(f[y]) = \{y\}$.
- *Gegeben eine Abbildung* $f\colon X \to Y$*, so existieren auch die folgenden beiden Abbildungen auf den dazugehörigen Potenzmengen:*

  $$\begin{aligned} f_*\colon \mathcal{P}(X) &\to \mathcal{P}(Y), & A &\mapsto f(A) \\ f^*\colon \mathcal{P}(Y) &\to \mathcal{P}(X), & B &\mapsto f^{-1}(B). \end{aligned}$$

  Wir haben dann $\Gamma_{f_*} = \{(A, f(A)) \mid A \subseteq X\}$ *sowie* $\Gamma_{f^*} = \{(B, f^{-1}(B)) \mid B \subseteq Y\}$.

Übung 24. *Sei* $f\colon X \to Y$ *eine Abbildung; sowie* $A, \tilde{A} \subseteq X$ *und* $B, \tilde{B} \subseteq Y$ *Teilmengen. Zeigen Sie die folgenden Implikationen:*

$$\begin{aligned} A \subseteq \tilde{A} \qquad &\Longrightarrow \qquad f(A) \subseteq f(\tilde{A}) \\ B \subseteq \tilde{B} \qquad &\Longrightarrow \qquad f^{-1}(B) \subseteq f^{-1}(\tilde{B}). \end{aligned} \tag{43}$$

Übung 25. *Seien X, Y, J Mengen, sowie $f\colon X \to Y$ eine Abbildung. Seien weiterhin $A \subseteq X$ und $B, C \subseteq Y$ Teilmengen, sowie $Y_j \subseteq Y$ für jedes $j \in J$ eine Teilmenge.*

a) Zeigen Sie, dass $f(f^{-1}(B)) \subseteq B$ sowie $A \subseteq f^{-1}(f(A))$ gilt.

b) Zeigen Sie die folgenden Identitäten:

- $f^{-1}\big(\bigcap_{j\in J} Y_j\big) = \bigcap_{j\in J} f^{-1}(Y_j)$,
- $f^{-1}\big(\bigcup_{j\in J} Y_j\big) = \bigcup_{j\in J} f^{-1}(Y_j)$,
- $f^{-1}\big(B \setminus C\big) = f^{-1}(B) \setminus f^{-1}(C)$.

Hinweis: Die Behauptungen folgen zwanglos, wenn Sie jeweils mit der linken Seite starten, und dann sukzessive Äquivalenzumformungen (die Definitionen) anwenden.

Definition 16. *Seien X, Y Mengen, und $f\colon X \to Y$ eine Abbildung.*

1) Die Einschränkung von f auf eine Teilmenge $A \subseteq X$, ist die Abbildung

$$f|_A\colon A \to Y, \qquad x \mapsto f(x).$$

Wir haben also $\Gamma_{f|_A} = \{(x, f(x)) \mid x \in A\}$, sowie $\mathrm{dom}(f|_A) = A$.

2) Die Koeinschränkung von f auf eine Teilmenge $B \subseteq Y$ mit $f(X) \subseteq B$, ist die Abbildung

$$f|^B\colon X \to B, \qquad x \mapsto f(x).$$

Wir haben also $\Gamma_{f|^B} = \{(x, f(x)) \mid x \in X\} \subseteq X \times B$, sowie $\mathrm{im}(f|^B) = \mathrm{im}(f)$.

Übung 26. *Seien X, Y Mengen, $f\colon X \to Y$ eine Abbildung, sowie $A \subseteq X$ und $B \subseteq Y$ Teilmengen mit $f(X) \subseteq B$. Zeigen Sie die folgenden Aussagen:*

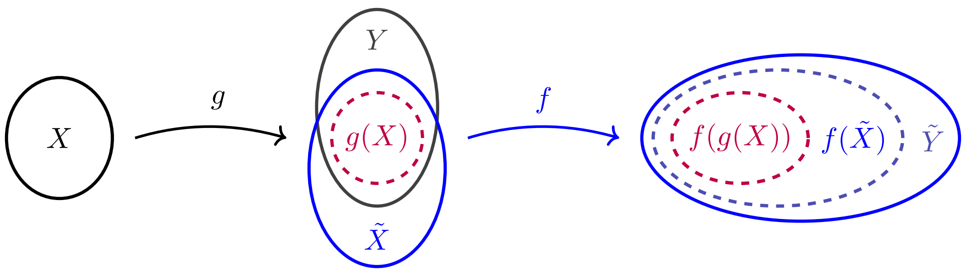
a) Es gilt $(f|_A)^{-1}(Z) = f^{-1}(Z) \cap A$ für alle $Z \subseteq Y$.

b) Es gilt $(f|^B)^{-1}(Z \cap B) = f^{-1}(Z \cap B) = f^{-1}(Z)$ für alle $Z \subseteq Y$.

Definition 17. *Seien $X, \tilde{X}, Y, \tilde{Y}$ Mengen, sowie $f\colon \tilde{X} \to \tilde{Y}$ und $g\colon X \to Y$ Abbildungen. Gilt $\mathrm{im}(g) = g(X) \subseteq \tilde{X}$, so ist die Verkettung (Komposition) von f mit g definiert durch*

$$f \circ g\colon X \to \tilde{Y}, \qquad x \mapsto f(g(x)).$$

Wir haben also $\Gamma_{f\circ g} = \{(x, f(g(x))) \mid x \in X\}$.



Bemerkung 16. *Ist $f\colon X \to Y$ eine Abbildung, so gilt natürlich*

$$\mathrm{id}_Y \circ f = f = f \circ \mathrm{id}_X.$$

Beispiel 17. *Sei* $f\colon \mathbb{N} \ni y \mapsto \frac{1}{1+y} \in \mathbb{Q}$ *und* $g\colon \mathbb{Z} \ni x \mapsto 1 + x^2 \in \mathbb{Z}$. *Dann gilt*

$$\operatorname{im}(g) \subset \mathbb{N} = \operatorname{dom}(f),$$

sodass die folgende Verkettung gebildet werden kann:

$$f \circ g\colon \mathbb{Z} \to \mathbb{Q}, \qquad x \mapsto \frac{1}{2 + x^2}.$$

Übung 27. *Seien* $f\colon Y \to Z$ *und* $g\colon X \to Y$ *Abbildungen. Zeigen Sie*

$$(f \circ g)^{-1}(A) = g^{-1}(f^{-1}(A)) \qquad \forall\, A \in \mathcal{P}(Z).$$

Definition 18. *Seien* X *und* Y *Mengen, und* $f\colon X \to Y$ *eine Abbildung.*

1) *Die Abbildung* f *heißt injektiv, wenn* $f^{-1}(y)$ *für jedes* $y \in Y$ *leer oder einelementig ist. Dies ist gleichbedeutend damit, dass die folgende Implikation gilt:*

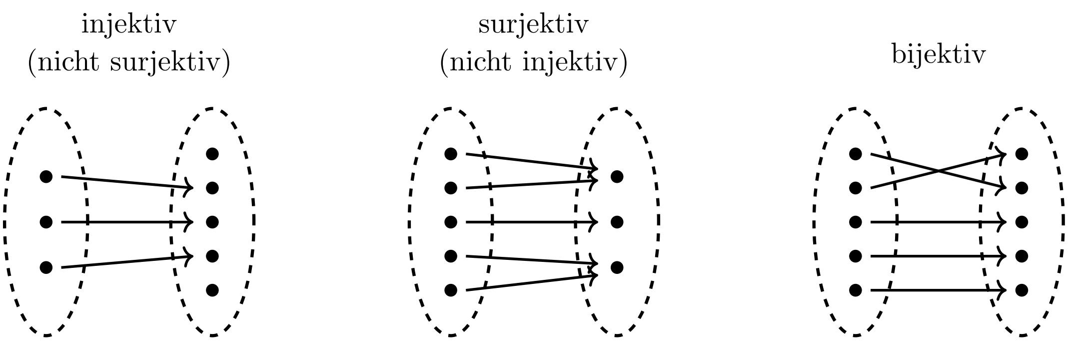
$$f(x) = f(\tilde{x}) \quad \text{für} \quad x, \tilde{x} \in X \qquad \Longrightarrow \qquad x = \tilde{x}.$$

2) *Die Abbildung* f *heißt surjektiv, wenn* $f^{-1}(y) \neq \emptyset$ *für jedes* $y \in Y$ *gilt. Dies ist gleichbedeutend damit, dass gilt:*

$$\operatorname{im}(f) = Y \qquad \textit{also} \qquad \forall\, y \in Y\colon \exists\, x \in X\colon f(x) = y.$$

3) *Die Abbildung* f *heißt bijektiv, wenn sie sowohl injektiv als auch surjektiv ist; wenn also die folgende Bedingung gilt:*

$$\forall\, y \in Y\colon \exists!\, x \in X\colon f(x) = y.$$



Übung 28. *Sei* $f\colon X \to Y$ *eine Abbildung. Zeigen Sie:*

- *Ist* f *surjektiv, so gilt* $f(f^{-1}(B)) = B$ *für jede Teilmenge* $B \subseteq Y$.
- *Ist* f *injektiv, so gilt* $f^{-1}(f(A)) = A$ *für jede Teilmenge* $A \subseteq X$.

Machen Sie sich Anhand der obigen Grafiken klar, dass diese Aussagen ohne die gegeben Zusatzvoraussetzungen an f *im Allgemeinen nicht gelten.*

Übung 29. *Seien* X, Y, J *Mengen, sowie* $f\colon X \to Y$ *eine Abbildung. Sei weiterhin* $X_j \subseteq X$ *für jedes* $j \in J$ *eine Teilmenge. Zeigen Sie, dass gilt:*

$$f(\textstyle\bigcup_{j \in J} X_j) = \bigcup_{j \in J} f(X_j)$$

Machen Sie sich zudem an einem Beispiel klar, dass die entsprechenden Identität für Schnitte (anstelle Vereinigungen) im Allgemeinen nicht erfüllt sein muss (Hinweis: Injektivität).

Notation 8. *Seien X und Y Mengen. Wir definieren die folgenden Teilräume (Teilmengen) von* $\mathsf{Abb}(X,Y)$ *(vgl. Notation 7):*

- $\mathsf{Inj}(X,Y)$ *– Die Menge aller Injektionen $X \to Y$.*
- $\mathsf{Surj}(X,Y)$ *– Die Menge aller Surjektionen $X \to Y$.*
- $\mathsf{Bij}(X,Y)$ *– Die Menge aller Bijektionen $X \to Y$.*

Es gilt $\mathsf{Bij}(X,Y) \subseteq \mathsf{Inj}(X,Y), \mathsf{Surj}(X,Y) \subseteq \mathsf{Abb}(X,Y)$.

Beispiel 18.

a) Sei $f = (\emptyset, Y, \emptyset)$ eine Abbildung (vgl. Bemerkung 15). Dann ist f injektiv. Zudem ist f surjektiv genau dann, wenn $Y = \emptyset$ gilt (also $f = (\emptyset, \emptyset, \emptyset)$).

b) Seien $X_1, \ldots, X_n$ mit $n \geq 1$ Mengen. Für jedes $1 \leq j \leq n$, ist dann die Projektion

$$\mathrm{pr}_j \colon X_1 \times \cdots \times X_n \to X_j, \qquad (x_1, \ldots, x_n) \mapsto x_j$$

auf den j-ten Faktor surjektiv.

c) Sei X eine Menge, und $A \subseteq X$ eine Teilmenge. Dann ist die Inklusion

$$\iota_A^X = \mathrm{id}|_A \colon A \to X, \qquad x \mapsto x$$

von A in X injektiv. (Wir bemerken, dass $f|_A = f \circ \iota_A^X$ für jede Abbildung $f \colon X \to Y$ gilt).

d) Sei $f \colon \mathbb{N} \to \mathbb{Z}$ gegeben durch

$$f(n) := \begin{cases} \frac{n}{2} & \text{falls } n \text{ gerade} \\ -\frac{n+1}{2} & \text{falls } n \text{ ungerade.} \end{cases}$$

Dann ist f eine Bijektion zwischen $\mathbb{N}$ und $\mathbb{Z}$ (Übung).

Lemma 6. *Seien $Z, \tilde{Z}$ Mengen, sowie $f \colon \tilde{Z} \to Z$ und $g \colon Z \to \tilde{Z}$ Abbildungen. Gilt $f \circ g = \mathrm{id}_Z$, dann ist f surjektiv und g ist injektiv.*

Beweis. Wir erhalten

$$Z = \mathrm{id}_Z(Z) = (f \circ g)(Z) = f(g(Z)) \overset{(43)}{\subseteq} f(\tilde{Z}) \qquad \overset{f(\tilde{Z}) \subseteq Z}{\Longrightarrow} \qquad f(\tilde{Z}) = Z,$$

also ist f surjektiv. Sind $z, z' \in Z$ vorgegeben mit $g(z) = g(z')$, so folgt

$$z = (f \circ g)(z) = f(g(z)) = f(g(z')) = (f \circ g)(z') = z',$$

also ist g injektiv. □

Definition 19. *Seien $(X_j)_{j \in J}$ und $(Y_j)_{j \in J}$ Familien von nichtleeren Mengen, sowie $(f_j)_{j \in J}$ eine Familie von Abbildungen $f_j \colon X_j \to Y_j$ für $j \in J$. Die zugehörige Produktabbildung ist gegeben durch*

$$\prod_{j \in J} f_j \colon \prod_{j \in J} X_j \to \prod_{j \in J} Y_j, \qquad (x_j)_{j \in J} \mapsto (f_j(x_j))_{j \in J}.$$

Ist $J = \{1, \ldots, n\}$ mit $n \geq 1$ endlich, so notiert man obige Abbildung auch mit $f_1 \times \ldots \times f_n$.

Lemma 7. *Seien $(X_j)_{j \in J}$ und $(Y_j)_{j \in J}$ Familien von nichtleeren Mengen, sowie $(f_j)_{j \in J}$ eine Familie von Abbildungen $f_j \colon X_j \to Y_j$ alle $j \in J$. Es gelten die folgen Aussagen für die Produktabbildung $f := \prod_{j \in J} f_j$:*

1) Ist f_j injektiv für alle $j \in J$, so ist auch f injektiv.

2) Ist f_j surjektiv für alle $j \in J$, so ist auch f surjektiv.

3) Ist f_j bijektiv für alle $j \in J$, so ist auch f bijektiv.

Beweis. 1) Seien $(x_j)_{j\in J}, (\tilde{x}_j)_{j\in J} \in \prod_{j\in J} X_j$ gegeben mit $f((x_j)_{j\in J}) = f((\tilde{x}_j)_{j\in J})$. Dann folgt aus der Injektivitätsvoraussetzung (erste Implikation)

$$f_j(x_j) = f_j(\tilde{x}_j) \quad \forall\, j \in J \qquad \Longrightarrow \qquad x_j = \tilde{x}_j \quad \forall\, j \in J \qquad \Longrightarrow \qquad (x_j)_{j\in J} = (\tilde{x}_j)_{j\in J},$$

was die Injektivität von f zeigt.

2) Gegeben $(y_j)_{j\in J} \in \prod_{j\in J} Y_j$, so existiert für jedes $j \in J$, ein $x_j \in X_j$ mit $f_j(x_j) = f_j(y_j)$. Dann gilt $f((x_j)_{j\in J}) = (y_j)_{j\in J}$, was die Surjektivität zeigt.

3) Die folgt sofort aus 1) und 2). $\square$

Konsequentes anwenden der Definitionen, liefert die Aussagen der folgenden Übungsaufgaben:

Übung 30. *Seien X und Y Mengen, und $f\colon X \to Y$ eine Abbildung. Zeigen Sie:*

- *Ist f injektiv, so ist für jede Teilmenge $A \subseteq X$ auch $f|_A$ injektiv.*
- *Ist f injektiv, so ist für jede Teilmenge $B \subseteq Y$ mit $\mathrm{im}(f) \subseteq B$ auch $f|^B$ injektiv.*
- *Die Koeinschränkung $f|^{\mathrm{im}(f)}$ ist surjektiv. Ist f injektiv, so ist $f|^{\mathrm{im}(f)}$ bijektiv.*

Übung 31. *Seien X, Y, Z Mengen, sowie $f\colon X \to Y$ und $g\colon Y \to Z$ Abbildungen. Zeigen Sie:*

- *Sind f und g injektiv, so ist auch $f \circ g$ injektiv.*
- *Sind f und g surjektiv, so ist auch $f \circ g$ surjektiv.*

Folgern Sie:

- *Sind f und g bijektiv, so ist auch $f \circ g$ bijektiv.*
- *Ist f bijektiv und g injektiv/surjektiv, so ist auch $f \circ g$ injektiv/surjektiv.*
- *Ist f injektiv/surjektiv und g bijektiv, so ist auch $f \circ g$ injektiv/surjektiv.*

Übung 32. *Seien W, X, Y, Z Mengen, sowie $f\colon Y \to Z$, $g\colon X \to Y$, $h\colon W \to X$ Abbildungen. Zeigen Sie die Assoziativität*

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

Definition 20. *Seien X und Y Mengen, und $f\colon X \to Y$ eine Abbildung.*

1) Eine Abbildung $f_L^{-1}\colon Y \to X$ heißt linksinvers zu f (oder Linksinverses von f), wenn gilt:

$$f_L^{-1} \circ f = \mathrm{id}_X. \tag{44}$$

2) Eine Abbildung $f_R^{-1}\colon Y \to X$ heißt rechtsinvers zu f (oder Rechtsinverses von f), wenn gilt:

$$f \circ f_R^{-1} = \mathrm{id}_Y. \tag{45}$$

3) Eine Abbildung $f^{-1}\colon Y \to X$ heißt invers zu f (oder Inverses bzw. Umkehrabbildung von f), wenn f^{-1} sowohl links- als auch rechtsinvers zu f ist – wenn also gilt:

$$f^{-1} \circ f = \mathrm{id}_X \qquad \textit{sowie} \qquad f \circ f^{-1} = \mathrm{id}_Y. \tag{46}$$

Lemma 8. *Sei $f\colon X \to Y$ eine Abbildung.*

1) Existiert ein Linksinverses f_L^{-1} von f, dann ist f injektiv und f_L^{-1} ist surjektiv.

2) Existiert ein Rechtsinverses f_R^{-1} von f, dann ist f surjektiv und f_R^{-1} ist injektiv.

3) Existiert ein Inverses f^{-1} von f, so sind f und f^{-1} beide bijektiv.

4) Existiert ein Inverses f^{-1} von f, so ist dieses eindeutig bestimmt.

Bemerkung: *Die Existenz vorausgesetzt, ist es also legitim, von der Umkehrabbildung von f zu sprechen.*

5) Existiert ein Linksinverses f_L^{-1} von f sowie ein Rechtsinverses f_R^{-1} von f, so ist $f_L^{-1} = f_R^{-1}$ die Umkehrabbildung von f.

Beweis. 1) Dies folgt wegen (44) sofort aus Lemma 6.

2) Dies folgt wegen (45) sofort aus Lemma 6.

3) Dies folgt wegen (46) sofort Lemma 6 (bzw. aus Teil 1) und Teil 2)).

4) Sei $g\colon Y \to X$ invers zu f. Aus Bemerkung 16 (erster und fünfter Schritt) sowie Übung 32 (Assoziativität der Verkettung) folgt

$$g = g \circ \mathrm{id}_Y = g \circ (f \circ f^{-1}) = (g \circ f) \circ f^{-1} = \mathrm{id}_X \circ f^{-1} = f^{-1}.$$

Dies zeigt die Eindeutigkeit des Inversen.

5) Aus Bemerkung 16 (und Übung 32) folgt

$$f_L^{-1} = f_L^{-1} \circ \mathrm{id}_Y = f_L^{-1} \circ (f \circ f_R^{-1}) = (f_L^{-1} \circ f) \circ f_R^{-1} = \mathrm{id}_X \circ f_R^{-1} = f_R^{-1}.$$

Per definitionem ist daher $f_L^{-1} = f_R^{-1}$ invers zu f, und wegen Teil 4) dann das Inverse von f. □

Notation 9. *Für jede Abbildung $f\colon X \to Y$ und jede Teilmenge $Z \subseteq Y$ ist die Urbildmenge $f^{-1}(Z)$ definiert. In diesem Sinne verwenden wir das Symbol f^{-1} also auch dann, wenn keine Umkehrabbildung zu f existiert. Existiert die Umkehrabbildung f^{-1} von f jedoch, so ist für jedes $Z \subseteq Y$ natürlich $f^{-1}(Z)$ identisch mit dem Urbild von Z unter f. Insbesondere ist dies dann auch mit unserer Konvention verträglich, ein einelementiges Urbilder einer einelementigen Menge mit dessen einzigem Element zu identifizieren (vgl. Definition 15.4)).*

Die ersten beiden Aussagen in Lemma 8, lassen sich wie folgt umkehren:

Lemma 9. *Sei $f\colon X \to Y$ eine Abbildung.*

1) Ist f injektiv mit $X \neq \emptyset$, so existiert ein (surjektives) Linksinverses von f.

2) Ist f surjektiv, so existiert ein (injektives) Rechtsinverses von f.

Beweis. 1) Sei $z \in X \neq \emptyset$ fixiert. Zu jedem $y \in f(X)$, existiert ein $h(y) \in X$ mit $y = f(h(y))$. Wir definieren die Abbildung $g\colon Y \to X$ durch

$$g(y) := \begin{cases} z & \text{für} \quad y \in Y \setminus f(X) \\ h(y) & \text{für} \quad y \in f(X). \end{cases}$$

Für jedes $x \in X$ gilt dann $(g \circ f)(x) = h(f(x))$, und daher:

$$f((g \circ f)(x)) = f(h(\underbrace{f(x)}_{y})) = \underbrace{f(x)}_{y} \quad \overset{f \text{ injektiv}}{\Longrightarrow} \quad (g \circ f)(x) = x.$$

Dies zeigt, dass g linksinvers zu f ist.

2) Da f surjektiv ist, existiert für jedes $y \in Y$ ein $g(y) \in X$ mit $y = f(g(y))$. Dies definiert eine Abbildung $g\colon Y \to X$ mit

$$(f \circ g)(y) = f(g(y)) = y \qquad \forall\, y \in Y.$$

Also ist g rechtsinvers zu f. □

Übung 33. *Seien X und Y Mengen, und $f\colon X \to Y$ eine Abbildung. Zeigen Sie (vgl. Beispiel 16):*

- *f_* ist injektiv/surjektiv genau dann, wenn f injektiv/surjektiv ist.*
- *f^* ist injektiv/surjektiv genau dann, wenn f surjektiv/injektiv ist.*

Hinweis: Für die eine Richtung ist es nützlich mit einelementigen Mengen zu argumentieren. Die andere Richtung folgt zwanglos aus Übung 28.

Bemerkung 17. *Wir erinnern an die Konvention* (28) *für Produkträume aus Notation 4.*

1) *Gegeben Mengen $X, Y \neq \emptyset$, so lässt sich $\mathsf{Abb}(X,Y)$ auch als Produktraum Y^X auffassen. Genauer hat man die natürlichen Bijektion (mit Umkehrabbildung)*

$$\Psi\colon \mathsf{Abb}(X,Y) \to Y^X, \qquad f \mapsto (f(x))_{x\in X}$$
$$\Psi^{-1}\colon Y^X \to \mathsf{Abb}(X,Y), \qquad (y_x)_{x\in X} \mapsto \big[f\colon X \ni x \mapsto y_x \in Y\big].$$

2) *Gegeben eine Menge $X \neq \emptyset$, so betrachten wir den Produktraum $\mathbf{2}^X := \{0,1\}^X$. Dann ist*

$$\Omega\colon \mathbf{2}^X \to \mathcal{P}(X), \qquad (o_x)_{x\in X} \mapsto \{x \in X \mid o_x = 1\}$$

eine Bijektion. Ihre Umkehrabbildung

$$\Omega^{-1}\colon \mathcal{P}(X) \to \mathbf{2}^X, \qquad Z \mapsto (\alpha(Z)_x)_{x\in X}$$

ist gegeben durch

$$\alpha(Z)_x := \begin{cases} 1 & \text{falls} \quad x \in Z \\ 0 & \text{falls} \quad x \in Z \setminus X \end{cases}$$

für alle $x \in X$. Die Potenzmenge einer Menge X wird daher oft auch mit $\mathbf{2}^X$ notiert.

3.3.2 Mächtigkeit von Mengen

Wir behandeln nun das Konzept der Mächtigkeit von Mengen.

Definition 21. *Seien X, Y Mengen.*

1) *X heißt endlich, wenn $X = \emptyset$ gilt oder wenn eine Surjektion $\{1, \ldots, n\} \to X$ mit $n \geq 1$ existiert.*
2) *X heißt abzählbar, wenn $X = \emptyset$ gilt oder wenn eine Surjektion $\mathbb{N} \to X$ existiert.*
3) *X heißt überabzählbar, wenn X nicht abzählbar ist.*
4) *X und Y heißen gleichmächtig, wenn eine Bijektion $X \to Y$ existiert.*

Bemerkung 18. *Seien X und Y Mengen.*

1) *Sei $X \neq \emptyset$.*
   - *X ist genau dann endlich, wenn eine Injektion $\iota\colon X \to \{1, \ldots, n\}$ mit $n \in \mathbb{N}_{>0}$ existiert; und zwar wegen Lemma 9*

- *X ist genau dann endlich, wenn eine Bijektion $\kappa\colon X \to \{1,\dots,m\}$ mit $m \in \mathbb{N}_{>0}$ ($m \le n$) existiert:*
  - *Ist $\kappa\colon X \to \{1,\dots,m\}$ eine Bijektion, so setze $\iota := \kappa$ sowie $n := m$.*
  - *Ist $\iota\colon X \to \{1,\dots,n\}$ (mit $n \in \mathbb{N}_{>0}$) injektiv, so ist die Koeinschränkung*

$$\iota|^{\mathrm{im}(\iota)}\colon X \to \mathrm{im}(\iota) \subseteq \{1,\dots,n\}$$

bijektiv (Übung 30). Es gilt dann $\mathrm{im}(\iota) = \{\ell_1,\dots,\ell_m\}$ für ein $1 \le m \le n$, und gewisse $1 \le \ell_1 < \ldots < \ell_m \le n$. Wir definieren $\alpha\colon \mathrm{im}(\iota) \to \{1,\dots,m\}$ durch $\alpha(\ell_j) := j$ für $j = 1,\dots,m$. Dann ist α bijektiv, also auch dei verkettung

$$\kappa := \alpha \circ \iota|^{\mathrm{im}(\iota)}\colon X \to \{1,\dots,m\}.$$

a) *Es ist nun unmittelbar einsichtig, dass X genau dann endlich ist, wenn X aus endlich vielen Elementen besteht (letzteres entspricht unserer ursprünglichen Definition einer nichtleeren endlichen Menge aus Kapitel 1):*
   - *Ist $X \neq \emptyset$ endlich sowie $\psi_X\colon X \to \{1,\dots,n\}$ ($n \in \mathbb{N}_{>0}$) eine Bijektion, so folgt unmittelbar $|X| = n$.*
   - *Ist $X \neq \emptyset$ gegeben mit $|X| = n \in \mathbb{N}_{>0}$, so gilt $X = \{x_1,\dots,x_n\}$ für paarweiese unetrschiedliche Elemente x_j mit $1 \le j \le n$. Dann ist $\kappa\colon X \ni x_j \mapsto j \in \{1,\dots,n\}$ offensichtlich bijektiv.*

   Unsere Definition der Kardinalität einer endlichen Menge besteht somit weiterhin.

b) *Sind $X, Y \neq \emptyset$ endlich, so gilt $|X| = |Y|$ genau dann, wenn X und Y gleichmächtig sind.*

   Beweis der Äquivalenz: Seien $\psi_X\colon X \to \{1,\dots,n\}$ mit $n \in \mathbb{N}_{>0}$, sowie $\psi_Y\colon Y \to \{1,\dots,m\}$ mit $m \in \mathbb{N}_{>0}$ Bijektionen.
   - Ist $\psi\colon X \to Y$ eine Bijektion, so ist

$$\psi_Y \circ \psi \circ \psi_X^{-1}\colon \{1,\dots,n\} \to \{1,\dots,m\}$$

   eine Bijektion (nach Übung 31); also gilt $|X| = n = m = |Y|$.
   - Gilt umgekehrt $|X| = |Y|$, so folgt $n = m$, also $\mathrm{im}(\psi_X) = \mathrm{im}(\psi_Y)$. Daher ist $\psi := \psi_Y^{-1} \circ \psi_X\colon X \to Y$ eine Bijektion. □

2) *Endliche Mengen sind abzählbar.*

Beweis der Behauptung: Für $X = \emptyset$ ist die Behauptung klar. Sei also $\chi\colon \{1,\dots,n\} \to X$ mit $n \in \mathbb{N}_{>0}$ eine Surjektion. Dann ist $\tilde{\chi}\colon \mathbb{N} \to X$ gegeben durch

$$\tilde{\chi}(p) := \begin{cases} \chi(p) & \text{für } 1 \le p \le n \\ \chi(1) & \text{für } p \in \mathbb{N} \setminus \{1,\dots,n\}, \end{cases}$$

ebenfalls eine Surjektion. □

3) *Ist X abzählbar und nicht endlich, so ist X gleichmächtig zu $\mathbb{N}$.*

Beweis der Implikation: Wegen Lemma 9, existiert eine Injektion $\iota\colon X \to \mathbb{N}$. Nach Übung 30 ist

$$\iota|^{\mathrm{im}(\iota)}\colon X \to \mathrm{im}(\iota) \subseteq \mathbb{N}$$

eine Bijektion (also $\mathrm{im}(\iota)$ nicht endlich). Dann ist $\alpha\colon \mathrm{im}(\iota) \to \mathbb{N}$, definiert durch[9]

$$\alpha(p) := |\{q \in \mathrm{im}(\iota) \mid q \le p\}| \qquad \forall\, p \in \mathrm{im}(\iota),$$

bijektiv. In der Tat, für $p, p' \in \mathrm{im}(\iota)$ mit $p < p'$, gilt automatisch $\alpha(p) < \alpha(p')$:

[9] Die rechte Seite bezeichnet die Anzahl der Elemente in $\mathrm{im}(\iota)$, die kleiner oder gleich p sind.

- Die Injektivität von α folgt nun unmittelbar.
- Die Surjektivität von α folgt durch Widerspruch ($\operatorname{im}(\iota)$ ist nicht endlich). $\square$

4) Jede Teilmenge $Y \subseteq X$ einer abzählbaren Menge X ist abzählbar.

Beweis der Behauptung: Die Aussage ist klar für $Y = \emptyset$. Ist nun $Y \neq \emptyset$, so gilt notwendigerweise $X \neq \emptyset$, und es existiert eine Surjektion $\chi\colon \mathbb{N} \to X$. Wir fixieren ein $p \in \mathbb{N}$ mit $\chi(p) \in Y$, und definieren $\tilde{\chi}\colon \mathbb{N} \to Y$ durch

$$\tilde{\chi}(n) := \begin{cases} \chi(n) & \text{für } \chi(n) \in Y, \\ \chi(p) & \text{für } \chi(n) \in X \setminus Y. \end{cases}$$

Offensichtlich ist $\tilde{\chi}$ surjektiv, also ist Y abzählbar. $\square$

5) Seien X und Y gleichmächtig, mit Bijektion $\kappa\colon X \to Y$.

- *Ist X endlich, so auch Y.*

  (Ist $\chi\colon \{1, \ldots, n\} \to X$ eine Surjektion, so ist auch $\kappa \circ \chi\colon \{1, \ldots, n\} \to Y$ eine Surjektion.)
- *Ist X abzählbar, so auch Y.*

  (Ist $\chi\colon \mathbb{N} \to X$ eine Surjektion, so ist auch $\kappa \circ \chi\colon \mathbb{N} \to Y$ eine Surjektion.)

Beispiel 19. *Seien X und Y Mengen.*

- *Gilt $X, Y \neq \emptyset$, so sind Y^X und $\mathsf{Abb}(X, Y)$ gleichmächtig nach Bemerkung 17.*1).
- *Sind X, Y beide endliche, so gilt $|\mathsf{Abb}(X, Y)| = |Y|^{|X|}$ (mit der Konvention $1^0 = 1 = 0^0$).*

  Beweis der Behauptung: Ist $X = \emptyset$, so gilt $|X| = 0$, und $(\emptyset, Y, \emptyset)$ ist die einzige Abbildung für eine gegebene endliche Menge Y. Daher gilt $|\mathsf{Abb}(X, Y)| = 1 = |Y|^0 = |Y|^{|X|}$. Diese Formel ist dann auch für $Y = \emptyset$ korrekt, denn nach Bemerkung 15 erzwingt $Y = \emptyset$ bereits $X = \emptyset$.

  Sei nun $X, Y \neq \emptyset$. Dann gilt $|\mathsf{Abb}(X, Y)| = |Y^X|$ wegen dem vorherigen Punkt, sowie Bemerkung 18.1)b). Weiterhin gilt $|Y^X| = |Y|^{|X|}$ wegen (26) in Definition 7. $\square$

Beispiel 20. *Sei $X \neq \emptyset$ eine Menge.*

- *Die Mengen $\mathcal{P}(X)$ und $\mathbf{2}^X = \{0, 1\}^X$ sind gleichmächtig nach Bemerkung 17.*2).
- *Ist X eine endliche Menge, so gilt $|\mathcal{P}(X)| = 2^{|X|}$.*

  Beweis. Es gilt $|\mathcal{P}(X)| = |\mathbf{2}^X|$ wegen dem vorherigen Punkt, sowie Bemerkung 18.1)b). Weiterhin gilt $|\mathbf{2}^X| = |\{0, 1\}|^{|X|} = \mathbf{2}^{|X|}$ wegen (26) in Definition 7. $\square$

Übung 34. *Sei $X \neq \emptyset$ eine endliche Menge, und $f \in \mathsf{Abb}(X, X)$ eine Selbstabbildung* ($\operatorname{dom}(f) = X = \operatorname{im}(f)$)*. Zeigen Sie*

$$f \in \mathsf{Bij}(X, X) \qquad \Longleftrightarrow \qquad f \in \mathsf{Inj}(X, X) \qquad \Longleftrightarrow \qquad f \in \mathsf{Surj}(X, X).$$

Gemäß Definition 21.4) sind zwei Mengen X und Y genau dann gleichmächtig, wenn eine Paarbildung $\Gamma \subseteq X \times Y$ mit

$$\forall\, x \in X : \exists!\, y \in Y : (x, y) \in \Gamma \qquad \wedge \qquad \forall\, y \in Y : \exists!\, x \in X : (x, y) \in \Gamma$$

existiert. Die Anschauung ist, dass die beiden Mengen X und Y „gleichviele“ Elemente enthalten, wieviele es auch sein mögen. Beispielsweise zeigt der letzte Punkt in Beispiel 18, dass die Mengen $\mathbb{N}$ und $\mathbb{Z}$ gleichmächtig sind. Ein weiteres Beispiel liefert der folgende Satz.

Satz 2. *Die Mengen* $\mathbb{N}$ *und* $\mathbb{N} \times \mathbb{N}$ *sind gleichmächtig. Insbesondere ist also* $\mathbb{N} \times \mathbb{N}$ *abzählbar.*

$*$*Beweis.* Wir definieren eine Bijektion $\alpha\colon \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ wie folgt: Sei

$$p_0 := 0 \qquad \text{sowie} \qquad p_n := 1 + \ldots + n \qquad \forall\, n \in \mathbb{N}_{>0}$$

(also $p_n = p_{n-1} + n$ für alle $n \geq 1$). Für jedes $m \in \mathbb{N}$, existiert dann ein eindeutiges $n(m) \in \mathbb{N}$ mit $p_{n(m)} \leq m < p_{n(m)+1}$, und wir setzen

$$\alpha(m) := (m - p_{n(m)}, n(m) - (m - p_{n(m)})).$$

- Die Abbildung α ist injektiv. Gilt nämlich $\alpha(m) = \alpha(m')$ für $m, m' \in \mathbb{N}$, so folgt

$$\begin{aligned} m - p_{n(m)} &= m' - p_{n(m')} \\ n(m) - (m - p_{n(m)}) &= n(m') - (m' - p_{n(m')}) \end{aligned}$$

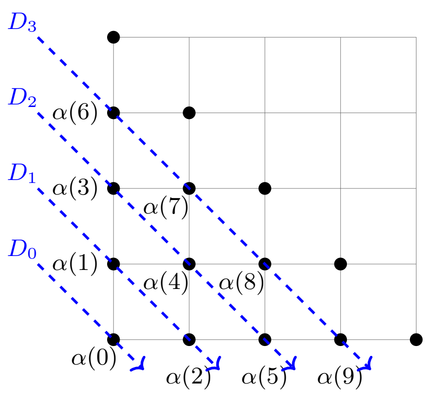
  Einsetzen der ersten in die zweite Zeile liefert $n(m) = n(m')$, was zusammen mit der ersten Zeile $m = m'$ impliziert.

- Die Abbildung α ist surjektiv. Für $(a, b) \in \mathbb{N} \times \mathbb{N}$ sei nämlich $n := a + b$ und $m := p_n + a$. Dann gilt $n(m) = n$, also

$$\alpha(m) = (m - p_n, n - (m - p_n)) = (a, n - a) = (a, b).$$

  Dies zeigt die Behauptung. $\square$

Bemerkung 19. *Die Bijektion* $\alpha\colon \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ *im Beweis von Satz 2 durchläuft nacheinander die blau gekennzeichneten Diagonalen* $(D_0, D_2, D_3, \ldots)$ *in der angegebenen Richtung. Auf der* n*-ten Diagonale* ($n \in \mathbb{N}$) *befinden sich* n*-Elemente, nämlich* $\alpha(p_n), \ldots, \alpha(p_n + n)$*. Für alle Elemente* (a, b) *auf der* n*-ten Diagonale gilt* $a + b = n$*.*



Korollar 1. *Jede abzählbare Vereinigung abzählbarer Mengen ist abzählbar, d.h. ist* $J \neq \emptyset$ *abzählbar und* $(X_j)_{j \in J}$ *eine Familie abzählbarer Mengen, so ist* $\bigcup_{j \in J} X_j$ *eine abzählbare Menge.*

Beweis. Gilt $X_j = \emptyset$ für alle $j \in J$, so ist die Behauptung klar. Andernfalls ist die Teilmenge

$$\emptyset \neq \tilde{J} := \{j \in J \mid X_j \neq \emptyset\} \subseteq J$$

nichtleer und abzählbar. Wir können daher ohne Beschränkung annehmen, dass $X_j \neq \emptyset$ für alle $j \in J$ gilt. Sei $\chi\colon \mathbb{N} \to J$ eine Surjektion. Per Voraussetzung, existiert für jedes $j \in J$ eine Surjektion $\beta_j\colon \mathbb{N} \to X_j$, und dann ist

$$\beta\colon \mathbb{N} \times \mathbb{N} \to \textstyle\bigcup_{j \in J} X_j, \qquad (n, m) \mapsto \beta_{\chi(n)}(m)$$

ebenfalls surjektiv. Wegen Satz 2 existiert eine Bijektion $\alpha\colon \mathbb{N} \to \mathbb{N} \times \mathbb{N}$, und dann ist $\beta \circ \alpha\colon \mathbb{N} \to \bigcup_{j \in J} X_j$ ebenfalls surjektiv (Übung 31). $\square$

Korollar 2. *Die Menge $\mathbb{Q}$ der rationalen Zahlen ist abzählbar.*

Beweis. Wegen Beispiel 18, ist $\mathbb{Z}$ abzählbar (also auch $\mathbb{Z}_{\neq 0}$ wegen Bemerkung 18.4)). Für jedes $n \in \mathbb{Z}_{\neq 0}$ ist also $X_n := \{\frac{m}{n} \mid m \in \mathbb{Z}\}$ abzählbar (mit Surjektion $\chi_j \colon \mathbb{Z} \ni m \mapsto \frac{m}{n} \in \mathbb{Q}$). Wegen Korollar 1 ist dann auch $\mathbb{Q} = \bigcup_{n \in \mathbb{Z}_{\neq 0}} X_n$ abzählbar. □

Übung 35. *Zeigen Sie per Induktion, dass $\mathbb{N}^n$ (vgl. (25)) für alle $n \in \mathbb{N}_{>0}$ abzählbar ist. Folgern Sie, dass X^n für jede abzählbare Menge X und jedes $n \in \mathbb{N}_{>0}$ abzählbar ist.*

Hinweis: Benutzen Sie Satz 2, Lemma 7, und Übung 31.

Übung 36. *Sei $X \neq \emptyset$ eine abzählbare Menge. Zeigen Sie, dass für jedes $n \in \mathbb{N}_{>0}$, die Menge*

$$\mathcal{E}_n(X) := \{Y \subseteq X : 1 \leq |Y| \leq n\}$$

der maximal n-elementigen nichtleeren Teilmengen von X abzählbar ist. Folgern Sie, dass die Menge $\mathcal{E}(X)$ aller endlichen Teilmengen von X abzählbar ist.

Hinweis: Sie dürfen Übung 35 benutzen: Gibt es eine natürliche Surjektion $X^n \to \mathcal{E}_n(X)$?

Satz 3 (Cantor-Russel)**.** *Sei X eine Menge. Dann existiert keine Surjektion $X \to \mathcal{P}(X)$, insbesondere also auch keine Bijektion zwischen X und $\mathcal{P}(X)$.*

Beweis durch Widerspruch: Sei $f \colon X \to \mathcal{P}(X)$ eine surjektive Abbildung. Wir betrachten

$$Z := \{x \in X \mid x \notin f(x)\} \in \mathcal{P}(X).$$

Wegen der Surjektivität von f existiert ein $x \in X$ mit $f(x) = Z$. Es gilt dann entweder $x \in f(x)$ oder $x \notin f(x)$, wobei sich beide Annahmen jeweils selbst widersprechen:

- Sei $x \in f(x)$. Dann folgt $x \in f(x) = Z$; also $x \notin f(x)$, was der Annahme widerspricht.
- Sei $x \notin f(x)$. Dann folgt $x \in Z$; also $x \in Z = f(x)$, was der Annahme widerspricht. □

Korollar 3. *Die Menge $\mathcal{P}(\mathbb{N})$ ist überabzählbar.*

Beweis. Klar wegen Satz 3. □

Bemerkung 20. *Seien X und Y Mengen. Man schreibt*

- *$|X| \leq |Y|$ (bzw. $|Y| \geq |X|$) genau dann, wenn eine Surjektion $Y \to X$ existiert.*
- *$|X| = |Y|$ genau dann, wenn $|X| \leq |Y|$ und $|Y| \leq |X|$ gilt.*
- *$|X| < |Y|$ genau dann, wenn $|X| \leq |Y|$ aber nicht $|Y| \leq |X|$ gilt.*

Mit dem bisher Gezeigtem macht man sich leicht klar, dass obige Definitionen im Falle endlicher Mengen X und Y mit unserer Definition der Kardinalität im Einklang steht. Selbiges gilt für die folgenden beiden Grundlegenden Sätze der Mengenlehre:

- **Vergleichbarkeitssatz:**

  Gegeben zwei Mengen X, Y, so gilt stets $|X| \leq |Y|$ oder $|Y| \leq |X|$.
- **Satz von Cantor-Bernstein-Schröder:**

  Gegeben Mengen X, Y mit $|X| = |Y|$, so sind X und Y gleichmächtig (sind also in Bijektion zueinander).

Satz 3 impliziert, dass keine „größte Menge“ existieren kann:

Korollar 4. *Es gilt $|X| < |\mathcal{P}(X)|$ für jede Menge $X \neq \emptyset$.*

Beweis. • Nach Satz 3 gilt **nicht** $|\mathcal{P}(X)| \leq |X|$.

• Es gilt $|X| \leq |\mathcal{P}(X)|$; denn fixiert man $x_0 \in X$ und definiert

$$\chi(Z) := \begin{cases} x & \text{für } Z = \{x\} \text{ mit } x \in X \\ x_0 & \text{sonst} \end{cases}$$

für alle $Z \in \mathcal{P}(X)$, so ist $\chi\colon \mathcal{P}(X) \to X$ offensichtlich surjektiv. □

Wir wollen schließlich noch einige Aussagen über endliche Mengen beweisen.

Satz 4. *Seien X und Y endliche Mengen, mit $|X|, |Y| = n \in \mathbb{N}$. Dann existieren $n!$ verschiedene bijektive Abbildung $X \to Y$ (also $|\mathsf{Bij}(X,Y)| = n!$).*

$*$*Beweis.* Die Aussage ist klar für $n \in \{0,1\}$ (IA). Es gelte nun die Aussage für ein $n \geq 1$ (IV). Sei

$$X = \{x_0, \ldots, x_n\} \qquad \text{und} \qquad Y = \{y_0, \ldots, y_n\},$$

mit $x_i \neq x_j$ und $y_i \neq y_j$ für alle $0 \leq i \neq j \leq n$. Ist $f\colon X \to Y$ eine Bijektion, so kann $f(x_0)$ genau $n+1$ mögliche Werte in Y annehmen. Ist nun $f(x_0) \in Y$ fixiert, so ist die eingeschränkte Abbildung $f|_{\{x_1,\ldots,x_n\}} \to Y \setminus \{f(x_0)\}$ ebenfalls eine Bijektion. Nach (IV) gibt es hierfür genau $n!$ Möglichkeiten, denn es gilt

$$|\{x_1, \ldots, x_n\}| = n = |N \setminus \{f(x_0)\}|.$$

Insgesamt ergeben sich daher $(n+1) \cdot n! = (n+1)!$ Möglichkeiten für eine Bijektion $X \to Y$. □

Terminologie 8. *Seien X und Y Mengen.*

1) *Eine Bijektion von einer Menge auf sich selbst wird als Permutation dieser Menge bezeichnet.*
2) *Sei $X = \{1, \ldots, n\}$ mit $n \geq 1$, so kann eine Bijektion $\{1, \ldots, n\} \to Y$ auch als Aufzählung der Menge Y in der Form $f(1), \ldots, f(n)$ aufgefasst werden. Satz 4 besagt dann, dass genau $n!$ Aufzählungen von Y existieren. Ist Beispielsweise Y eine Menge von $n \geq 1$ Büchern, so gibt es genau $n!$ Möglichkeiten, diese nebeneinander ins Regal zu stellen.*

Für $n, k \in \mathbb{N}$, definieren wir den zugehörigen Binomialkoeffizienten durch $(0! = 1)$

$$\binom{n}{k} := \begin{cases} \frac{n!}{k! \cdot (n-k)!} & \text{für} \quad 0 \leq k \leq n \\ 0 & \text{für} \quad k \geq n+1. \end{cases}$$

Nachrechnen liefert (Übung)

$$\begin{aligned} \binom{n}{k} &= \binom{n}{n-k} && \text{für} \quad 0 \leq k \leq n \\ \binom{n}{k} &= \binom{n-1}{k-1} + \binom{n-1}{k} && \text{für} \quad 1 \leq k \leq n. \end{aligned} \tag{47}$$

Mit Hilfe der Formel in der zweiten Zeile, lassen sich die Binomialkoeffizienten sukzessive berechnen. Es ergibt sich ein Wert in dem Pascalschen Dreieck (siehe unten) durch Addition der beiden links und rechts schräg darüber stehenden Werte:

$$\begin{array}{c}
\binom{0}{0}\\
\binom{1}{0}\binom{1}{1}\\
\binom{2}{0}\binom{2}{1}\binom{2}{2}\\
\binom{3}{0}\binom{3}{1}\binom{3}{2}\binom{3}{3}\\
\binom{4}{0}\binom{4}{1}\binom{4}{2}\binom{4}{3}\binom{4}{4}\\
\binom{5}{0}\binom{5}{1}\binom{5}{2}\binom{5}{3}\binom{5}{4}\binom{5}{5}\\
\binom{6}{0}\binom{6}{1}\binom{6}{2}\binom{6}{3}\binom{6}{4}\binom{6}{5}\binom{6}{6}\\
\binom{7}{0}\binom{7}{1}\binom{7}{2}\binom{7}{3}\binom{7}{4}\binom{7}{5}\ \dots\\
\vdots
\end{array}
\qquad
\begin{tikzcd}
\binom{n-1}{k-1} \arrow[dr, dashed] & + & \binom{n-1}{k} \arrow[dl, dashed] \\
 & \binom{n}{k} &
\end{tikzcd}
\qquad
\begin{array}{c}
1\\
1\ \ 1\\
1\ \ 2\ \ 1\\
1\ \ 3\ \ 3\ \ 1\\
1\ \ 4\ \ 6\ \ 4\ \ 1\\
1\ \ 5\ \ 10\ \ 10\ \ 5\ \ 1\\
1\ \ 6\ \ 15\ \ 20\ \ 15\ \ 6\ \ 1\\
1\ \ 7\ \ 21\ \ 35\ \ 35\ \ 21\ \dots\\
\vdots
\end{array}$$

Eine Interpretation der Binomialkoeffizienten liefert der folgende Satz.

Satz 5. *Eine Menge X mit $n \in \mathbb{N}$ Elementen hat $\binom{n}{k}$ Teilmengen mit $0 \leq k \leq n$ Elementen.*

$*$*Beweis.* (IA): Die Formel gilt für $n = 0$; denn für $n = 0$ ist $k = 0$, also $\binom{n}{k} = 1$. Weiterhin ist $\emptyset$ ist die einzige Teilmenge von X die 0 Elemente enthält.

(IV): Es gelte die Formel für alle Mengen mit n Elementen, für ein $n \in \mathbb{N}$.

(IS): Sei X eine Menge mit $n + 1$ Elementen.

- Die Formel gilt für $k = 0$: Es ist dann $\binom{n+1}{k} = 1$; und $\emptyset$ ist die einzige Teilmenge von X, die 0 Elemente enthält.
- Die Formel gilt für $k = n+1$: Es ist dann $\binom{n+1}{k} = 1$; und X ist die einzige Teilmenge von X, die $n + 1$ Elemente enthält.

Sei nun also $1 \leq k \leq n$. Wir fixieren ein $x_0 \in X$, und erhalten

$$X = \{x_0\} \,\dot{\cup}\, (X \setminus \{x_0\}) \qquad \text{mit} \qquad |(X \setminus \{x_0\})| = n.$$

Ist nun $Y \subseteq X$ eine k-elementige Teilmenge, so sind die folgenden beiden Fälle möglich:

- Es gilt $x_0 \in Y$: Dann ist $Y \cap (X \setminus \{x_0\})$ eine $k-1$-elementige Teilmenge von $X \setminus \{x_0\}$. Nach (IV) gibt es hierfür $\binom{n}{k-1}$ Möglichkeiten (beachte $1 \leq k \leq n$).
- Es gilt $x_0 \notin Y$: Dann ist $Y \cap (X \setminus \{x_0\})$ eine k-elementige Teilmenge von $(X \setminus \{x_0\})$. Nach (IV) gibt es hierfür $\binom{n}{k}$ Möglichkeiten (beachte $1 \leq k \leq n$).

Insgesamt existieren also (beachte $1 \leq k \leq n$)

$$\binom{n}{k-1} + \binom{n}{k} = \binom{(n+1)-1}{k-1} + \binom{(n+1)-1}{k} \overset{(47)}{=} \binom{n+1}{k}$$

unterschiedliche k-elementige Teilmengen von X. $\square$

Bemerkung* 1. *Alternativ lässt sich die Aussage in Satz 5 auch wie folgt herleiten:*

Sei $n \geq 2$, $1 \leq k \leq n-1$ (die anderen Fälle sind trivial), sowie X eine n-elementige Menge. Man erhält jede k-elementigen Teilmenge von X, indem man alle Mengen der Form

$$M_\alpha := \{\alpha(1), \ldots, \alpha(k)\} \qquad \textit{mit} \qquad \alpha \in \mathsf{Bij}(\{1, \ldots, n\}, X) \quad \textit{(eine Aufzählung)}$$

betrachtet. Nach Satz 4 existieren $n!$ derartige Aufzählungen – allerdings liefern zwei Aufzählungen $\alpha, \beta \in \mathsf{Bij}(\{1, \ldots, n\}, X)$ die gleiche k-elementige Menge ($M_\alpha = M_\beta$), wenn

$$\alpha(\{1, \ldots, k\}) = \beta(\{1, \ldots, k\}) \tag{48}$$

gilt. Aus einem fixierten α, erhält nun alle β mit (48), indem man α mit allen Bijektionen

$$\gamma\colon \{1,\ldots,k\} \to \{1,\ldots,k\} \qquad \text{und} \qquad \delta\colon \{k+1,\ldots,n\} \to \{k+1,\ldots,n\}$$

kombiniert; und zwar in der Form $\alpha \circ f[\gamma,\delta]\colon \{1,\ldots,n\} \to X$ mit

$$f[\gamma,\delta](i) := \begin{cases} \gamma(i) & \text{für} \quad 1 \le i \le k \\ \delta(i) & \text{für} \quad k+1 \le i \le n. \end{cases}$$

Satz 4 zeigt, dass es für derartige Verkettungen genau $k! \cdot (n-k)!$ Möglichkeiten gibt. Hieraus ergibt sich dann die gewünschte Anzahl der k-elementigen Teilmengen von X zu

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

Wir ziehen die folgenden Schlussfolgerungen.

Korollar 5. *Es gilt $\binom{n}{k} \in \mathbb{N}_{>0}$ für alle $n \in \mathbb{N}$ und $0 \le k \le n$.*

Beweis. Per definitionem gilt $0 \le \binom{n}{k} \in \mathbb{Q}$. Satz 5 zeigt nun weiterhin $\binom{n}{k} \in \mathbb{N}$. □

***Korollar 1.** *Seien X, Y endliche Mengen mit $1 \le |X| \le |Y|$. Dann gilt*

$$|\mathsf{Inj}(X,Y)| = \frac{|Y|!}{(|Y|-|X|)!}.$$

Beweis. Sei $k := |X|$ und $n := |Y|$. Nach Satz 5, existieren $p := \binom{n}{k}$ unterschiedliche k-elementige Teilmengen $Y_1,\ldots,Y_p$ von Y. Nach Satz 4 gilt $|\mathsf{Bij}(X,Y_\ell)| = k!$ für jedes $1 \le \ell \le p$. Weiterhin gilt $\mathsf{Bij}(X,Y_\ell) \cap \mathsf{Bij}(X,Y_{\ell'}) = \emptyset$ für $1 \le \ell \ne \ell' \le p$ (einfach wegen $Y_\ell \ne Y_{\ell'}$). Jedes Element von $\dot\bigcup_{1\le\ell\le p} \mathsf{Bij}(X,Y_\ell)$, definiert dann eine andere Injektion $X \to Y$; also gilt

$$|\mathsf{Inj}(X,Y)| \le \tbinom{n}{k} \cdot k! = \tfrac{|Y|!}{(|Y|-|X|)!}. \tag{49}$$

Schließlich gilt $f|^{\mathrm{im}(f)} \in \dot\bigcup_{1\le\ell\le p} \mathsf{Bij}(X,Y_\ell)$ für alle $f \in \mathsf{Inj}(X,Y)$; also gilt in (49) sogar Gleichheit. □

4 Körper

In diesem Kapitel behandeln wir das Konzept des Körpers, und führen die reellen Zahlen ein.

4.1 Mengen mit Verknüpfungen

Wir diskutieren nun zunächst einige grundlegende algebraische Strukturen, sowie deren wichtigsten Eigenschaften.

Definition 22. *Eine Verknüpfung $*$ auf einer Menge X ist eine Abbildung*

$$*\colon X \times X \to X, \qquad (x,y) \mapsto x * y := *(x,y).$$

Sind $Y, Z \subseteq X$ Teilmengen, so definieren wir

$$Y * Z := *(Y,Z) = \{y * z \mid y \in Y \ \wedge \ z \in Z\}. \tag{50}$$

Terminologie 9. *Sei $*$ eine Verknüpfung auf der Menge X.*

1) *Das Paar $(X, *)$ heißt Halbgruppe, wenn die Verknüpfung $*$ assoziativ ist:*

$$x * (y * z) = (x * y) * z \qquad \forall\, x, y, z \in X. \tag{51}$$

Die Halbgruppe $(X, *)$ heißt abelsch (kommutativ), wenn

$$x * y = y * x \qquad \forall\, x, y \in X. \tag{52}$$

Eine Teilmenge $Y \subseteq X$ heißt Unterhalbgruppe von X, wenn $Y * Y \subseteq Y$ gilt, und somit $(Y, *|_{Y \times Y})$ selbst eine Halbgruppe ist.

2) *Ein neutrales Element einer Halbgruppe $(X, *)$ ist ein Element $e \in X$, sodass*

$$e * x = x = x * e \qquad \forall\, x \in X. \tag{53}$$

Sind $f, e \in X$ beides neutrale Elemente von $(X, *)$, so folgt bereits $f = f * e = e$. Es ist daher legitim von dem neutralen Element einer Halbgruppe zu sprechen, sofern es existiert.

3) *Ein Monoid ist ein Tripel $(X, *, e)$, sodass $(X, *)$ eine Halbgruppe mit neutralem Element e ist. Sei $x \in X$ ein Element:*

   a) *Ein Element $x_L^{-1} \in X$ heißt Linksinverses von (linksinvers zu) x, wenn $x_L^{-1} * x = e$ gilt.*
      - *Sei $y_L^{-1} \in X$ linksinvers zu $y \in X$, und $z_L^{-1} \in X$ linksinvers zu $z \in X$. Dann ist $z_L^{-1} * y_L^{-1}$ linksinvers zu $y * z$, wegen*

$$(z_L^{-1} * y_L^{-1}) * (y * z) \stackrel{(51)}{=} ((z_L^{-1} * y_L^{-1}) * y) * z \stackrel{(51)}{=} (z_L^{-1} * (y_L^{-1} * y)) * z = (z_L^{-1} * e) * z = e.$$

   b) *Ein Element $x_R^{-1} \in X$ heißt Rechtsinverses von (rechtsinvers zu) x, wenn $x * x_R^{-1} = e$ gilt.*
      - *Sei $y_R^{-1} \in X$ rechtsinvers zu $y \in X$, und $z_R^{-1} \in X$ rechtsinvers zu $z \in X$. Dann ist $z_R^{-1} * y_R^{-1}$ rechtsinvers zu $y * z$, wegen*

$$(y * z) * (z_R^{-1} * y_R^{-1}) \stackrel{(51)}{=} ((y * z) * z_R^{-1}) * y_R^{-1} \stackrel{(51)}{=} (y * (z * z_R^{-1})) * y_R^{-1} = (y * e) * y_R^{-1} = e.$$

   c) *Ein Element $x^{-1} \in X$ heißt Inverses von (invers zu) x, wenn x^{-1} sowohl links- als auch rechtsinvers zu x ist; wenn also gilt:*

$$x^{-1} * x = e = x * x^{-1}. \tag{54}$$

   Wir bemerken folgendes:

   - *Existiert ein Inverses von x, so ist dieses eindeutig bestimmt, und wird mit x^{-1} notiert.*

     Beweis der Eindeutigkeit: Sind $z, \tilde{z} \in X$ beide invers zu x, so folgt

$$z = z * (x * \tilde{z}) \stackrel{(51)}{=} (z * x) * \tilde{z} \stackrel{(51)}{=} \tilde{z},$$

     was die Behauptung zeigt □

     Wegen $e * e = e$, gilt dann insbesondere $e^{-1} = e$.
   - *Sei $x_L^{-1} \in X$ linksinvers zu x, sowie $x_R^{-1} \in X$ rechtsinvers zu x. Dann ist $x_L^{-1} = x_R^{-1}$ das Inverse von x; und zwar wegen*

$$x_L^{-1} = x_L^{-1} * (x * x_R^{-1}) \stackrel{(51)}{=} (x_L^{-1} * x) * x_R^{-1} = x_R^{-1}.$$

   - *Ist x^{-1} invers zu x, so ist x invers zu x^{-1}; es gilt also $(x^{-1})^{-1} = x$.*

– *Seien $y, z \in X$ gegeben mit Inversen $y^{-1}, z^{-1} \in X$. Dann gilt*

$$(y * z)^{-1} = z^{-1} * y^{-1}. \tag{55}$$

Dies folgt unmittelbar aus den entsprechenden Aussagen in a) und b) (Übung).

Ein Monoid heißt abelsch (kommutativ) genau dann, wenn die zugehörige Halbgruppe abelsch (kommutativ) ist. Eine Teilmenge $Y \subseteq X$ heißt Untermonoid von X, wenn $e \in Y$ sowie $Y * Y \subseteq Y$ gilt, und somit $(Y, *|_{Y\times Y}, e)$ selbst ein Monoid ist.

4) *Eine Gruppe ist ein Monoid $(X, *, e)$, sodass jedes $x \in X$ ein (notwendigerweise eindeutiges) Inverses $x^{-1} \in X$ besitzt. Eine Gruppe heißt abelsch (kommutativ), wenn der zugehörige Monoid abelsch (kommutativ) ist. Eine Teilmenge $Y \subseteq X$ heißt Untergruppe von X, wenn $(Y, *|_{Y\times Y}, e)$ selbst ein Gruppe ist – wenn also gilt:*

$$e \in Y, \qquad Y * Y \subseteq Y, \qquad y \in Y \implies y^{-1} \in Y.$$

5) *Seien $(X, *, e)$ und $(\tilde{X}, \tilde{*}, \tilde{e})$ Gruppen. Eine Abbildung $\Psi\colon X \to \tilde{X}$ heißt Gruppenhomomorphismus, wenn*

$$\Psi(x * y) = \Psi(x) \mathbin{\tilde{*}} \Psi(y) \qquad \forall\, x, y \in X. \tag{56}$$

Der Kern von Ψ, ist die Teilmenge

$$\ker(\Psi) := \Psi^{-1}(\tilde{e}) = \{x \in G \mid \Psi(x) = e\} \subseteq G.$$

Ist Ψ ein Gruppenhomomorphismus, so gelten die folgenden Aussagen:

ã) $\Psi(e) = \tilde{e}$, wegen

$$\begin{aligned}
\tilde{e} &= \Psi(e) \mathbin{\tilde{*}} \Psi(e)^{-1} \\
&= \Psi(e * e) \mathbin{\tilde{*}} \Psi(e)^{-1} \\
&\overset{(56)}{=} (\Psi(e) \mathbin{\tilde{*}} \Psi(e)) \mathbin{\tilde{*}} \Psi(e)^{-1} \\
&= \Psi(e) \mathbin{\tilde{*}} (\Psi(e) \mathbin{\tilde{*}} \Psi(e)^{-1}) \\
&= \Psi(e) \mathbin{\tilde{*}} \tilde{e} \\
&= \Psi(e).
\end{aligned}$$

$\tilde{b}$) $\Psi(x^{-1}) = \Psi(x)^{-1}$ für alle $x \in X$, denn es gilt

$$\tilde{e} \overset{\tilde{a})}{=} \Psi(e) = \Psi(x \cdot x^{-1}) \overset{(56)}{=} \Psi(x) \mathbin{\tilde{*}} \Psi(x^{-1})$$
$$\tilde{e} \overset{\tilde{a})}{=} \Psi(e) = \Psi(x^{-1} \cdot x) \overset{(56)}{=} \Psi(x^{-1}) \mathbin{\tilde{*}} \Psi(x)$$

wegen der Eindeutigkeit des Inversen $\Psi(x)^{-1}$ von $\Psi(x)$.

$\tilde{c}$) Ψ ist injektiv genau dann, wenn $\ker(\Psi) = \Psi^{-1}(\tilde{e}) = \{e\}$.

Beweis der Äquivalenz: Die eine Richtung ist klar wegen ã). Es gelte nun also $\Psi^{-1}(\tilde{e}) = \{e\}$. Es folgt mit (56) und $\tilde{b}$) (erster Schritt)

$$\begin{aligned}
\Psi(x) = \Psi(y) \quad &\implies \quad \Psi(x \mathbin{\tilde{*}} y^{-1}) = \Psi(e) \overset{\tilde{a})}{=} \tilde{e} \quad \implies \quad x * y^{-1} \in \Psi^{-1}(\tilde{e}) = \{e\} \\
&\implies \quad x * y^{-1} = e \quad \implies \quad x = y
\end{aligned}$$

für $x, y \in X$.

Beispiel 21.

- *Wir haben die folgenden Verknüpfungen auf* $\mathbb{Q}$ *(ebenfalls auf* $\mathbb{Z}$, $\mathbb{N}$, $\mathbb{N}_{>0}$*):*

$$+ \colon \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}, \qquad (q,p) \mapsto q + p$$
$$\cdot \colon \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}, \qquad (q,p) \mapsto q \cdot p.$$

  - *Es sind* $(\mathbb{Z}, +, 0)$, $(\mathbb{Q}, +, 0)$ *abelsche Gruppen, wobei dann* $\mathbb{Z} \subseteq \mathbb{Q}$ *eine Untergruppe ist.*
  - *Es ist* $(\mathbb{Q} \setminus \{0\}, \cdot, 1)$ *eine abelsche Gruppe.*[10]
  - *Es ist* $(\mathbb{N}, +, 0)$ *eine abelscher Monoid, und* $(\mathbb{N}_{>0}, +)$ *eine abelsche Halbgruppe (kein Monoid).*
  - *Es sind* $(\mathbb{N}, \cdot, 1)$, $(\mathbb{Z}, \cdot, 1)$, $(\mathbb{Q}, \cdot, 1)$ *abelsche Monoide, wobei* $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$ *Untermonoide sind.*

- *Ist* X *eine Menge, so ist* $(\mathsf{Abb}(X,X), \circ, \mathrm{id}_X)$ *ein Monoid (vgl. Übung 32). Dann ist* $\mathsf{Bij}(X,X) \subseteq \mathsf{Abb}(X,X)$ *ein Untermonoid, und sogar eine Gruppe (mit Umkehrabbildungen als Inverse).*

  (Ist X *endlich mit* $|X| \leq 1$*, so ist* $\mathsf{Abb}(X,X) = \mathsf{Bij}(X,X)$ *abelsch (Übung). Die Gruppe* $\mathsf{Bij}(X,X)$ *ist auch noch abelsch für* $|X| = 2$*, aber nicht mehr für* $|X| \geq 3$ *(Übung).)*

***Beispiel 2.** *Wir betrachten die Verknüpfung*

$$* \colon \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}, \qquad (x,y) \mapsto x - y.$$

Dann ist $*$ *nicht assoziative, also* $(\mathbb{Q}, *)$ *keine Halbgruppe; denn für* $x, y, z \in \mathbb{Q}$ *mit* $z \neq 0$ *gilt*

$$x * (y * z) = x - (y * z) = x - (y - z) = (x - y) + z \neq (x - y) - z = (x - y) * z = (x * y) * z.$$

Notation 10. *Sei* $(X, *)$ *eine Halbgruppe. Für* $x_1, \ldots, x_n$ *mit* $n \geq 3$*, definieren wir*

$$x_1 * \ldots * x_n := x_1 * (x_2 * (\ldots (x_{n-1} * x_n) \ldots)).$$

Diese Ausdrücke lassen sich auch induktiv mit Hilfe von Prinzip 1 definieren, nämlich durch

$$\begin{aligned} x_1 * x_2 * x_3 &:= x_1 * (x_2 * x_3) && \forall\, x_1, x_2, x_3 \in X, \\ x_0 * \ldots * x_n &:= x_0 * (x_1 * \ldots * x_n) && \forall\, x_0, \ldots, x_n \in X,\ n \geq 3. \end{aligned}$$

Die nächsten beiden Lemmata stellen formell klar, dass wegen der Assoziativität von $*$ die explizite Form der Klammerung bei Mehrfachprodukten keine Rolle spielt.

Lemma 10. *Sei* $(X, *)$ *eine Halbgruppe, und* $m, n \geq 1$*. Dann gilt*

$$(x_1 * \ldots * x_m) * (y_1 * \ldots * y_n) = x_1 * \ldots * x_m * y_1 * \ldots * y_n \tag{57}$$

für alle $x_1, \ldots, x_m, y_1, \ldots, y_n \in X$.

$*$*Beweis.* Wir beweisen die Aussage per Induktion über $m \geq 1$:

(IA): (57) ist klar für $m = 1$ und $n \geq 1$.

(IV): Es gelte (57) für ein $m \geq 1$ und alle $n \geq 1$.

(IS): Sei $n \geq 1$, und $x_0, \ldots, x_m, y_1, \ldots, y_n \in X$. Dann folgt

$$\begin{aligned} (x_0 * \ldots * x_m) * (y_1 * \ldots * y_n) &\overset{\text{def.}}{=} (x_0 * (x_1 * \ldots * x_m)) * (y_1 * \ldots * y_n) \\ &\overset{(51)}{=} x_0 * ((x_1 * \ldots * x_m) * (y_1 * \ldots * y_n)) \\ &\overset{\text{(IV)}}{=} x_0 * (x_1 * \ldots * x_m * y_1 * \ldots * y_n) \\ &\overset{\text{def.}}{=} x_0 * \ldots * x_m * y_1 * \ldots * y_n. \end{aligned}$$

Die Behauptung folgt nun per Induktion. □

[10]Beachte: Das multiplikative Inverse $\frac{1}{0}$ der 0 existiert nicht (ist nicht definiert).

Lemma 11. *Sei $(X, *)$ eine Halbgruppe. Gegeben $p \geq 2$, $n_1, \ldots, n_p \geq 1$, sowie $x[\ell]_1, \ldots, x[\ell]_{n_\ell} \in X$ für $\ell = 1, \ldots, p$. Dann gilt*

$$(x[1]_1 * \ldots * x[1]_{n_1}) * \ldots * (x[p]_1 * \ldots * x[p]_{n_p}) = x[1]_1 * \ldots * x[1]_{n_1} * \ldots * x[p]_1 * \ldots * x[p]_{n_p}.$$

$*$*Beweis.* Wegen Lemma 10 gilt die Behauptung für $p = 2$ (IA). Gilt nun die Behauptung für ein $p \geq 2$ (IV), so folgt für $n_1, \ldots, n_{p+1} \geq 1$, sowie $x[\ell]_1, \ldots, x[\ell]_{n_\ell} \in X$ für $\ell = 1, \ldots, p+1$, dass

$$\begin{aligned}
&(x[1]_1 * \ldots * x[1]_{n_1}) * \ldots * (x[p+1]_1 * \ldots * x[p+1]_{n_{p+1}}) \\
&\quad \overset{\text{def.}}{=} ((x[1]_1 * \ldots * x[1]_{n_1}) * \ldots * (x[p]_1 * \ldots * x[p]_{n_p})) * (x[p+1]_1 * \ldots * x[p+1]_{n_{p+1}}) \\
&\quad \overset{\text{(IV)}}{=} (x[1]_1 * \ldots * x[1]_{n_1} * \ldots * x[p]_1 * \ldots * x[p]_{n_p}) * (x[p+1]_1 * \ldots * x[p+1]_{n_{p+1}}) \\
&\quad \overset{\text{(IA)}}{=} x[1]_1 * \ldots * x[1]_{n_1} * \ldots * x[p+1]_1 * \ldots * x[p+1]_{n_{p+1}}.
\end{aligned}$$

Die Behauptung folgt somit per Induktion. □

Übung 37. *Sei $(X, *)$ eine abelsche Halbgruppe, und $n \geq 2$. Machen Sie sich klar:*

- *Für alle $1 \leq i \leq j \leq n$ und $x_1, \ldots, x_n \in X$ gilt*

$$\begin{aligned}
&x_1 * \ldots * x_{i-1} * x_{\boldsymbol{i}} * x_{i+1} * \ldots * x_{j-1} * x_{\boldsymbol{j}} * x_{j+1} \ldots * x_n \\
&\qquad = x_1 * \ldots * x_{i-1} * x_{\boldsymbol{j}} * x_{i+1} * \ldots * x_{j-1} * x_{\boldsymbol{i}} * x_{j+1} \ldots * x_n.
\end{aligned} \tag{58}$$

  Hinweis: Lemma 11 erlaubt es, den Ausdruck auf der linke Seite von (58) durch Einfügen von Klammern in geeignete Häppchen derart aufzuteilen, dass (52) iterativ anwendbar ist.

- *Für jede Permutation (bijektive Abbildung) $\iota \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ gilt*

$$x_1 * \ldots * x_n = x_{\iota(1)} * \ldots * x_{\iota(n)} \qquad \forall\, x_1, \ldots, x_n \in X.$$

  Hinweis: Verwenden Sie den ersten Punkt, um $x_{\iota(1)}$ an die erste Stelle zu schieben, dann $x_{\iota(2)}$ an die zweite Stelle ... usw.

Korollar 6. *Sei $(X, *, e)$ eine Gruppe. Für alle $n \geq 2$ gilt*

$$(x_1 * \ldots * x_n)^{-1} = x_n^{-1} * \ldots * x_1^{-1} \qquad \forall\, x_1, \ldots, x_n \in X.$$

$*$*Beweis.* Wegen (55) gilt die Behauptung für $n = 2$ (IA). Gilt nun die Behauptung für ein $n \geq 2$ (IV), so folgt

$$\begin{aligned}
(x_0 * \ldots * x_n)^{-1} &\overset{(51)}{=} (x_0 * (x_1 * \ldots * x_n))^{-1} \\
&\overset{(55)}{=} (x_1 * \ldots * x_n)^{-1} * x_0^{-1} \\
&\overset{\text{(IV)}}{=} (x_n^{-1} * \ldots * x_1^{-1}) * x_0^{-1} \\
&= x_n^{-1} * \cdots * x_0^{-1},
\end{aligned}$$

wobei wir im letzten Schritt Lemma 10 benutzt haben. □

Notation 11. *Sei $(X, *)$ eine Halbgruppe. Für jedes $x \in X$ definieren wir*

$$x^1 := x \quad \textit{sowie} \quad x^n := \underbrace{x * \ldots * x}_{n\text{-mal}} \quad \textit{für alle} \quad n \geq 2. \tag{59}$$

Ist e ein neutrales Element von $(X,*)$, ist also $(X,*,e)$ ein Monoid, so definieren wir zusätzlich $x^0 := e$. Ist $(X,*,e)$ sogar eine Gruppe, so definieren wir zudem

$$x^{-n} := (x^{-1})^n = \underbrace{x^{-1} * \ldots * x^{-1}}_{n\text{-mal}} \quad \textit{für alle} \quad n \geq 2. \tag{60}$$

Induktiv erhält man $e^n = e$ für alle $n \in \mathbb{Z}$. Ist schließlich $Z \subseteq X$ eine Teilmenge, so setzen wir

$$Z^{-1} := \{z^{-1} \mid z \in Z\}. \tag{61}$$

Wir erhalten die folgenden Rechenregeln, die Sie für die rationalen Zahlen bereits aus der Schule kennen.

Lemma 12. *Sei $(X,*,e)$ eine Gruppe. Für alle $x \in X$, und $m,n \in \mathbb{Z}$ gilt*

$$x^{-n} = (x^n)^{-1}, \qquad x^m * x^n = x^{m+n}, \qquad (x^m)^n = x^{m \cdot n}. \tag{62}$$

$*$*Beweis.* Korollar 6 zeigt $x^{-n} = (x^n)^{-1}$ für alle $n \in \mathbb{N}$; Lemma 10 zeigt $x^m * x^n = x^{m+n}$ für alle $m,n \in \mathbb{N}$; und Lemma 11 zeigt $(x^m)^n = x^{m \cdot n}$ für alle $m,n \in \mathbb{N}$.

- Für $n \in \mathbb{Z}$ mit $n \leq 0$ ist $n = -|n|$, und wir erhalten

  $$x^{-n} = x^{|n|} = ((x^{|n|})^{-1})^{-1} \overset{(60)}{=} (x^{-|n|})^{-1} = (x^n)^{-1},$$

  was die erste Gleichheit in (62) für alle $n \in \mathbb{Z}$ zeigt.
- Für $m,n \in \mathbb{Z}$ mit $m,n \leq 0$ ist nach dem bereits Gezeigten

  $$x^m * x^n = x^{-|m|} * x^{-|n|} = (x^{-1})^{|m|} * (x^{-1})^{|n|} = (x^{-1})^{|m|+|n|} = x^{-(|m|+|n|)} = x^{m+n}.$$

  Weiterhin folgt für $p,q \geq 0$, dass

  $$\begin{aligned} x^{-p+(p+q)} &= x^q = x^{-p} * (x^p * x^q) = x^{-p} * x^{p+q} \\ x^{(-p-q)+p} &= x^{-q} = (x^q)^{-1} = (x^{-p} * (x^p * x^q))^{-1} = (x^{-p} * x^{p+q})^{-1} \overset{(55)}{=} x^{-p-q} * x^p \end{aligned}$$

  womit die zweite Gleichheit in (62) dann auch für die Fälle $m \leq 0$ und $n \geq 0$, sowie $m \geq 0$ und $n \leq 0$ folgt.
- Ist $m,n \leq 0$, so folgt aus dem bereits Gezeigten

  $$x^{m \cdot n} = x^{|m| \cdot |n|} = (x^{|m|})^{|n|} = ((x^m)^{-1})^{|n|} = (x^m)^{-|n|} = (x^m)^n.$$

  Ist $m \leq 0$ und $n \geq 0$, so folgt hieraus

  $$(x^m)^n = ((x^{-1})^{|m|})^n = (x^{-1})^{|m| \cdot n} = x^{-|m| \cdot n} = x^{m \cdot n}.$$

  Ist schließlich $m \geq 0$ und $n \leq 0$, so folgt weiterhin

  $$(x^m)^n = (x^m)^{-|n|} = ((x^m)^{|n|})^{-1} = (x^{m \cdot |n|})^{-1} = x^{m \cdot n},$$

  was die dritte Gleichheit in (62) zeigt. □

Notation 12. *Sei $(X,*,e)$ eine abelsche Gruppe. Sei $n \in \mathbb{N}$, und J eine n-elementige Menge, also $J = \emptyset$ für $n = 0$. Es sei für jedes $j \in J$ ein $x_j \in X$ vorgegeben. Wir definieren*

$$\begin{aligned} \mathop{\Large *}_{j \in J} x_j &:= e && \textit{für} \quad J = \emptyset \\ \mathop{\Large *}_{j \in J} x_j &:= x_{\alpha(1)} * \ldots * x_{\alpha(n)} && \textit{für} \quad J \neq \emptyset, \end{aligned} \tag{63}$$

für eine Aufzählung (Bijektion) $\alpha\colon \{1,\ldots,n\} \to J$ im Falle $n \geq 1$.

- *Die erste Zeile in* (63) *wird als leere Verknüpfung bezeichnet.*
- *Wegen Übung 37 ist die zweite Zeile in* (63) *wohldefiniert, also unabhängig von der konkreten Wahl der Aufzählung* α.

Für jede Teilemenge $I \subseteq J$, *folgt dann aus Lemma 10, dass*

$$\mathop{\Large *}_{j\in J} x_j = \Big(\mathop{\Large *}_{j\in I} x_j\Big) * \Big(\mathop{\Large *}_{j\in J\setminus I} x_j\Big).$$

Für $m, n \in \mathbb{Z}$ *mit* $m \leq n$, *sowie* $x_m, \ldots, x_n \in X$, *setzten wir*

$$\mathop{\Large *}_{k=p}^{q} x_k := e \quad \text{für} \quad q < p \qquad \text{sowie} \qquad \mathop{\Large *}_{k=p}^{q} x_k := \mathop{\Large *}_{j\in\{p,\ldots,q\}} x_j \quad \text{für} \quad m \leq p \leq q \leq n.$$

Notation 13. *In Anlehnung an die Gruppen* $(\mathbb{Q}, +, 0)$ *und* $(\mathbb{Q}, \cdot, 1)$, *notiert man Gruppen im abelschen Fall oft auch wie folgt:*

a) $(X, \cdot, 1)$ *(multiplikative Gruppe mit multiplikativem neutralen Element* 1*):*

In diesem Fall benutzt man anstatt dem Symbol „$\ast$*“ das Produktzeichen „*$\prod$*“ (vgl. Notation 2). Die leere Verknüpfung (erste Zeile in* (63)*) wird dann als leeres Produkt bezeichnet.*

b) $(X, +, 0)$ *(additive Gruppe mit additivem neutralen Element* 0*):*

In diesem Fall benutzt man anstatt dem Symbol „$\ast$*“ das Summenzeichen „*$\sum$*“ (vgl. Notation 2); und die leere Verknüpfung (erste Zeile in* (63)*) wird als leere Summe bezeichnet. Weiterhin schreibt man* $-x$ *anstatt* x^{-1} *für* $x \in X$, *und dann* $x - y$ *anstatt* $x + (-y)$ *für* $x, y \in X$. *Für eine Teilmenge* $Z \subseteq X$, *schreibt man* $-Z$ *anstelle* Z^{-1}. *Zudem schreibt man* nx *anstatt* x^n *für* $x \in X$ *und* $n \in \mathbb{Z}$*; sodass dann also nach Notation 11*

$$0x = 0, \qquad 1x = x \qquad (\pm n)x = n(\pm x)\ \Big(= \sum_{k=1}^{n} \pm x\Big) \tag{64}$$

für $x \in X$ *und* $n \geq 2$ *gilt, wobei* $(\pm n)0 = 0$. *Lemma 12 liest sich dann*

$$(-n)x = -(nx), \qquad mx + nx = (m+n)x, \qquad n(mx) = (nm)x \tag{65}$$

für alle $x \in X$ *und* $m, n \in \mathbb{Z}$.

4.2 Körper

Die nächsten beiden Definitionen formalisieren die Eigenschaften der rationalen Zahlen.

Definition 23. *Sei* $(\mathbb{K}, +, 0_{\mathbb{K}}, \cdot, 1_{\mathbb{K}})$ *ein Tupel, bestehend aus den Folgenden Objekten:*

- *Eine Menge* $\mathbb{K}$ *mit ausgezeichneten Elementen* $\mathbb{K} \ni 0_{\mathbb{K}} \neq 1_{\mathbb{K}} \in \mathbb{K}$ *(*$\mathbb{K}$ *hat also mindestens zwei unterschiedliche Elemente).*
- *Zwei assoziative Verknüpfungen*

$$\begin{aligned} +\colon \mathbb{K}\times\mathbb{K} &\to \mathbb{K}, \qquad (x,y) \mapsto x + y \\ \cdot\colon \mathbb{K}\times\mathbb{K} &\to \mathbb{K}, \qquad (x,y) \mapsto x \cdot y; \end{aligned}$$

wobei die zweite Verknüpfung auf $\mathbb{K}_\times := \mathbb{K} \setminus \{0_{\mathbb{K}}\}$ *einschränkt, also* $\mathbb{K}_\times \cdot \mathbb{K}_\times \subseteq \mathbb{K}_\times$ *gilt.*

Dann heißt $(\mathbb{K}, +, 0_{\mathbb{K}}, \cdot, 1_{\mathbb{K}})$ *Körper, wenn die folgenden drei Bedingungen erfüllt sind:*

a) $(\mathbb{K}, +, 0_{\mathbb{K}})$ *ist eine abelsche Gruppe.* *(additive Gruppe)*

b) $(\mathbb{K}_\times, \cdot, 1_{\mathbb{K}})$ *ist eine abelsche Gruppe.* *(multiplikative Gruppe).*

c) Es gilt das Distributivgesetz

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z) \qquad \forall\, x, y, z \in \mathbb{K}. \tag{66}$$

Lemma 13. *Ist $(\mathbb{K}, +, 0_\mathbb{K}, \cdot, 1_\mathbb{K})$ ein Körper, so ist $(\mathbb{K}, \cdot, 1_\mathbb{K})$ ist ein abelscher Monoid. Genauer gilt*

$$x \cdot 0_\mathbb{K} = 0_\mathbb{K} = 0_\mathbb{K} \cdot x \qquad \forall\, x \in \mathbb{K}, \tag{67}$$

also auch $1_\mathbb{K} \cdot 0_\mathbb{K} = 0_\mathbb{K} = 0_\mathbb{K} \cdot 1_\mathbb{K}$.

Beweis. Wir erhalten aus (66) (jeweils dritter Schritt)

$$\begin{aligned}
0_\mathbb{K} &= x \cdot 0_\mathbb{K} - x \cdot 0_\mathbb{K} = x \cdot (0_\mathbb{K} + 0_\mathbb{K}) - x \cdot 0_\mathbb{K} = (x \cdot 0_\mathbb{K} + x \cdot 0_\mathbb{K}) - x \cdot 0_\mathbb{K} = x \cdot 0_\mathbb{K} \\
0_\mathbb{K} &= 0_\mathbb{K} \cdot x - 0_\mathbb{K} \cdot x = (0_\mathbb{K} + 0_\mathbb{K}) \cdot x - 0_\mathbb{K} \cdot x = (0_\mathbb{K} \cdot x + 0_\mathbb{K} \cdot x) - 0_\mathbb{K} \cdot x = 0_\mathbb{K} \cdot x,
\end{aligned}$$

was die Behauptung zeigt. □

Bemerkung 21. *In der Definition eines Körpers (Definition 23), wurde wir explizit $0_\mathbb{K} \neq 1_\mathbb{K}$ gefordert. Der Grund hierfür ist, dass man den Trivialfall ausschließen möchte, in dem $\mathbb{K}$ nur aus einem Element besteht. In der Tat gilt ja wegen (67) die Implikation*[11]

$$1_\mathbb{K} = 0_\mathbb{K} \qquad \Longrightarrow \qquad x = 1_\mathbb{K} \cdot x = 0_\mathbb{K} \cdot x \overset{(67)}{=} 0_\mathbb{K} \quad \forall\, x \in \mathbb{K}.$$

Notation 14. *Sei $(\mathbb{K}, +, 0_\mathbb{K}, \cdot, 1_\mathbb{K})$ ein Körper.*

- *Für jedes $x \in \mathbb{K}$, wird das additive Inverse (also das Inverse bezüglich $(\mathbb{K}, +, 0_\mathbb{K})$) mit $-x$ notiert. Für Teilmengen $Y, Z \subseteq \mathbb{K}$ ist $-Z = \{-z \mid z \in Z\}$ und $Y + Z = \{y + z \mid y \in Y \;\wedge\; z \in Z\}$ (vgl. (61) und (50)). Für $x, y \in \mathbb{K}$ schreibt man üblicherweise $x - y$ anstatt $x + (-y)$. Schreibt man $+x$ ohne weiteren Summanden, so ist einfach x gemeint – z.B. bedeutet dann $\pm x$ (oder $\mp x$) wie in (64) einfach die entsprechende Fallunterscheidung zwischen x und $-x$. Ist ein Körper fixiert, und keine Verwechslung zu befürchten, so schreibt man auch 0 anstatt $0_\mathbb{K}$.*
- *Für jedes $x \in \mathbb{K}_\times$ wird das multiplikative Inverse (also das Inverse bezüglich $(\mathbb{K}, \cdot, 1_\mathbb{K})$) mit x^{-1} bzw. $1/x$ oder $\frac{1}{x}$ notiert. Für $y \in \mathbb{K}$ und $z \in \mathbb{K}_\times$, schreibt man dann auch y/z oder $\frac{y}{z}$ anstatt $y \cdot z^{-1}$ bzw. $z^{-1} \cdot y$. Für Teilmengen $X \subseteq \mathbb{K}_\times$ und $Y, Z \subseteq \mathbb{K}$, ist $X^{-1} = \{x^{-1} \mid x \in X\}$ sowie $Y \cdot Z = \{y \cdot z \mid y \in Y \;\wedge\; z \in Z\}$. Anstatt $x \cdot y$ für $x, y \in \mathbb{K}$, schreibt man oft auch einfach xy. Ist ein Körper fixiert, und keine Verwechslung zu befürchten, so schreibt man auch 1 anstatt $1_\mathbb{K}$.*

Definition 24. *Sei $(\mathbb{K}, +, 0_\mathbb{K}, \cdot, 1_\mathbb{K})$ ein Körper, und $\leq$ eine Anordnung auf $\mathbb{K}$. Man bezeichnet das Tupel $(\mathbb{K}, \leq, +, 0_\mathbb{K}, \cdot, 1_\mathbb{K})$ als angeordneten Körper, wenn die folgenden beiden Implikationen für alle $x, y, z \in \mathbb{K}$ gelten:*

$$x \leq y \qquad \Longrightarrow \qquad x + z \leq y + z \tag{68}$$

$$x, y \geq 0_\mathbb{K} \qquad \Longrightarrow \qquad x \cdot y \geq 0_\mathbb{K}. \tag{69}$$

Beispiel 22. *Die rationalen Zahlen $\mathbb{Q}$ bilden zusammen mit den in Beispiel 9 definierten Rechenoperationen, und der in Beispiel 11 definierten Anordnung einen angeordneten Körper (Übung).*

Bemerkung 22. *In der Situation von Definition 24, impliziert Bedingung (68) natürlich*

$$x < y \quad \text{für} \quad x, y \in \mathbb{K} \qquad \Longrightarrow \qquad x + z < y + z \quad \forall\, z \in \mathbb{K}, \tag{70}$$

denn $x + z = y + z$ impliziert ja $x = (x + z) - z = (y + z) - z = y$.

Notation 15. *Sind Uneindeutigkeiten ausgeschlossen, so spricht man vereinfachend von einem Körper $\mathbb{K}$, anstatt $(\mathbb{K}, +, 0_\mathbb{K}, \cdot, 1_\mathbb{K})$ zu notieren. Weiterhin spricht man vereinfachend von einem angeordneten Körper $(\mathbb{K}, \leq)$, anstatt $(\mathbb{K}, \leq, +, 0_\mathbb{K}, \cdot, 1_\mathbb{K})$ zu notieren.*

[11] Die Identität (67) folgt in der gleichen Weise auch dann, wenn man $0_\mathbb{K} \neq 1_\mathbb{K}$ nicht fordert.

4.2.1 Körperhomomorphismen

Terminologie 10. *Seien* $\mathbb{K}, \tilde{\mathbb{K}}$ *Körper, und* $\Psi\colon \mathbb{K} \to \tilde{\mathbb{K}}$ *eine Abbildung.*

- Ψ *heißt Körperhomomorphismus, wenn für alle* x, y *gilt*

  $$\Psi(x+y) = \Psi(x) + \Psi(y) \qquad \textit{sowie} \qquad \Psi(x \cdot y) = \Psi(x) \cdot \Psi(y). \tag{71}$$

  Dies ist gleichbedeutend damit, dass Ψ *ein Gruppenhomomorphismus, sowohl zwischen den zugehörigen additiven, als auch den zugehörigen Multiplikativen Gruppen von* $\mathbb{K}$ *und* $\tilde{\mathbb{K}}$ *induziert.*

  Beweis der Äquivalenz: Es ist klar, dass (71) die besagten Gruppenhomomorphismeneigenschaften impliziert. Gelten umgekehrt die Gruppenhomomorphismeneigenschaften, so folgt wegen Lemma 13 und $\Psi(0_{\mathbb{K}}) = 0_{\tilde{\mathbb{K}}}$ auch (71). □

- *Sind* $(\mathbb{K}, \leq)$ *und* $(\tilde{\mathbb{K}}, \tilde{\leq})$ *beides angeordnete Körper, so fordert man von einem Körperhomomorphismus zusätzlich zu* (71) *noch*

  $$x \leq y \qquad \Longrightarrow \qquad \Psi(x) \mathrel{\tilde{\leq}} \Psi(y) \qquad \forall\, x, y \in \mathbb{K}. \tag{72}$$

Ein Körperhomomorphismus ist automatisch injektiv, und wird deshalb auch als Körpereinbettung bezeichnet. Ist ein Körperhomomorphismus bijektiv, so heißt er Körperisomorphismus.

Beweis der Injektivität. Wir fassen Ψ als Gruppenhomomorphismus zwischen den additiven Gruppen von $\mathbb{K}$ und $\tilde{\mathbb{K}}$ auf. Gemäß Terminologie 9.5).$\tilde{c}$), müssen wir nur nachweisen, dass $\Psi(x) = 0_{\tilde{\mathbb{K}}}$ bereits $x = 0_{\mathbb{K}}$ impliziert. Angenommen, die Behauptung ist falsch. Dann existiert ein $x \in \mathbb{K}_\times$ mit $\Psi(x) = 0_{\tilde{\mathbb{K}}}$, und wir erhalten mit Terminologie 9.5).$\tilde{c}$) (erster Schritt)

$$1_{\tilde{\mathbb{K}}} = \Psi(1_{\mathbb{K}}) = \Psi(x \cdot x^{-1}) = \Psi(x) \cdot \Psi(x)^{-1} = 0_{\tilde{\mathbb{K}}} \cdot \Psi(x)^{-1} \overset{(67)}{=} 0_{\tilde{\mathbb{K}}},$$

was $1_{\tilde{\mathbb{K}}} \neq 0_{\tilde{\mathbb{K}}}$ widerspricht. □

Lemma 14. *Seien* $(\mathbb{K}, \leq)$ *und* $(\tilde{\mathbb{K}}, \tilde{\leq})$ *angeordnete Körper, und* $\Psi\colon \mathbb{K} \to \tilde{\mathbb{K}}$ *ein Körperhomomorphismus. Dann gilt die folgende Implikation:*

$$x < y \quad \textit{für} \quad x, y \in \mathbb{K} \qquad \Longrightarrow \qquad \Psi(x) \mathrel{\tilde{<}} \Psi(y). \tag{73}$$

Beweis. Für $\mathbb{K} \ni x < y \in \mathbb{K}$, gilt zunächst $\Psi(x) \mathrel{\tilde{\leq}} \Psi(y)$. Da Ψ injektiv ist (Terminologie 10), gilt $\Psi(x) \neq \Psi(y)$ wegen $x \neq y$, also $\Psi(x) \mathrel{\tilde{<}} \Psi(y)$. □

Lemma 15. *Seien* $\mathbb{K}, \tilde{\mathbb{K}}$ *Körper, und* $\Psi\colon \mathbb{K} \to \tilde{\mathbb{K}}$ *ein Körperisomorphismus. Dann ist* Ψ^{-1} *ein Körperisomorphismus.*

Beweis. Wegen Lemma 8.3) ist Ψ^{-1} bijektiv. Sind nun $\tilde{x}, \tilde{y} \in \tilde{\mathbb{K}}$ gegeben, so folgt

$$\Psi^{-1}(\tilde{x}) + \Psi^{-1}(\tilde{y}) = (\Psi^{-1} \circ \Psi)\big(\Psi^{-1}(\tilde{x}) + \Psi^{-1}(\tilde{y})\big) = \Psi^{-1}\big(\Psi\big(\Psi^{-1}(\tilde{x})\big) + \Psi\big(\Psi^{-1}(\tilde{y})\big)\big) = \Psi^{-1}(\tilde{x} + \tilde{y})$$

$$\Psi^{-1}(\tilde{x}) \,\cdot\, \Psi^{-1}(\tilde{y}) = (\Psi^{-1} \circ \Psi)\big(\Psi^{-1}(\tilde{x}) \,\cdot\, \Psi^{-1}(\tilde{y})\big) = \Psi^{-1}\big(\Psi\big(\Psi^{-1}(\tilde{x})\big) \,\cdot\, \Psi\big(\Psi^{-1}(\tilde{y})\big)\big) = \Psi^{-1}(\tilde{x} \,\cdot\, \tilde{y}),$$

was die Behauptung zeigt. □

Korollar 7. *Seien* $(\mathbb{K}, \leq)$ *und* $(\tilde{\mathbb{K}}, \tilde{\leq})$ *angeordnete Körper, und* $\Psi\colon \mathbb{K} \to \tilde{\mathbb{K}}$ *ein Körperisomorphismus. Dann ist* Ψ^{-1} *ein Körperisomorphismus.*

Beweis. Seien $\tilde{\mathbb{K}} \ni \tilde{x} \mathrel{\tilde{\leq}} \tilde{y} \in \tilde{\mathbb{K}}$ gegeben. Gilt $\Psi^{-1}(\tilde{x}) > \Psi^{-1}(\tilde{y})$, so zeigt Lemma 14

$$\tilde{x} = \Psi(\Psi^{-1}(\tilde{x})) \mathrel{\tilde{>}} \Psi(\Psi^{-1}(\tilde{y})) = \tilde{y},$$

was der Annahme widerspricht. Daher gilt $\Psi^{-1}(\tilde{x}) \leq \Psi^{-1}(\tilde{y})$. Der Rest folgt aus Lemma 15. □

4.2.2 Grundlegendes zu Körpern

Wir wollen nun zunächst einige Grundlegende Eigenschaften von Körpern zusammentragen. Angeordnete Körper werden in Abschnitt 4.2.3 weitergehend diskutiert.

Übung 38 (Erweitertes Distributivgesetz)**.** *Sei $\mathbb{K}$ ein Körper. Zeigen Sie:*

$$y \cdot \left(\sum_{k=1}^{n} x_k\right) = \sum_{k=1}^{n} y \cdot x_k \qquad \forall\, n \geq 2,\ x_1, \dots, x_n, y \in \mathbb{K}.$$

Hinweis: Induktion über n unter Verwendung von (66)*.*

Lemma 16. *Sei $\mathbb{K}$ ein Körper, und $x, y \in \mathbb{K}$ und $u, v \in \mathbb{K}_\times$. Es gelten die folgenden Aussagen:*

1) $x \cdot y = 0_{\mathbb{K}} \quad \Longrightarrow \quad x = 0_{\mathbb{K}} \ \vee \ y = 0_{\mathbb{K}}$,

2) $-1_{\mathbb{K}} \cdot x = -x$,

3) $-(x \cdot y) = (-x) \cdot y$,

4) $(-u)^{-1} = -u^{-1}$,

5) $(-x) \cdot (-y) = x \cdot y$,

6) $\frac{-x}{u} = -\frac{x}{u} = \frac{x}{-u}$,

7) $\frac{x}{u} \cdot \frac{y}{v} = \frac{x \cdot y}{u \cdot v}$, $\frac{x \cdot v}{u \cdot v} = \frac{x}{u}$, $(\frac{u}{v})^{-1} = \frac{v}{u}$,

8) $\frac{x}{u} + \frac{y}{v} = \frac{x \cdot v + u \cdot y}{u \cdot v}$,

9) Gemäß der Konventionen für additive abelsche Gruppen aus Notation 13.b), gilt

$$n(x \cdot y) = (nx) \cdot y \qquad \forall\, n \in \mathbb{Z}. \tag{74}$$

Beweis. 1) Klar wegen $\mathbb{K}_\times \cdot \mathbb{K}_\times \subseteq \mathbb{K}_\times$.

2) $(-1_{\mathbb{K}}) \cdot x$ ist das eindeutige additive Inverse $-x$ von x, wegen

$$x + (-1_{\mathbb{K}}) \cdot x = (1_{\mathbb{K}} \cdot x) + ((-1_{\mathbb{K}}) \cdot x) \overset{(66)}{=} (1_{\mathbb{K}} - 1_{\mathbb{K}}) \cdot x = 0_{\mathbb{K}} \cdot x \overset{(67)}{=} 0_{\mathbb{K}}.$$

3) Es gilt $-(x \cdot y) \overset{2)}{=} -1_{\mathbb{K}} \cdot (x \cdot y) = (-1_{\mathbb{K}} \cdot x) \cdot y = (-x) \cdot y$.

4) Wir erhalten $(-1_{\mathbb{K}}) \cdot (-1_{\mathbb{K}}) \overset{2)}{=} -(-1_{\mathbb{K}}) = 1_{\mathbb{K}}$. Hiermit folgt

$$(-x) \cdot (-y) = ((-1_{\mathbb{K}}) \cdot x) \cdot ((-1_{\mathbb{K}}) \cdot y) = x \cdot (-1_{\mathbb{K}}) \cdot (-1_{\mathbb{K}}) \cdot y = x \cdot y.$$

5) Es gilt $-u^{-1} \cdot (-u) \overset{5)}{=} u^{-1} \cdot u = e$.

6) Wir erhalten

$$\frac{-x}{u} = -x \cdot u^{-1} \overset{3)}{=} -(x \cdot u^{-1}) = -\frac{x}{u} = -(x \cdot u^{-1}) \overset{3)}{=} x \cdot (-u^{-1}) \overset{4)}{=} x \cdot (-u)^{-1} = \frac{x}{-u}.$$

7) Klar.

8) Der vorherige Punkt zusammen mit dem Distributivgesetz liefert

$$\frac{x}{u} + \frac{y}{v} \overset{7)}{=} \frac{x \cdot v}{u \cdot v} + \frac{u \cdot y}{u \cdot v} \overset{(66)}{=} (x \cdot v + u \cdot y) \cdot (v \cdot u)^{-1} = \frac{x \cdot v + u \cdot y}{u \cdot v}.$$

9) Übung 38 liefert (vorletzter Schritt) für $m \in \mathbb{N}$

$$\begin{aligned}(\pm m)(x \cdot y) &\overset{\text{def.}}{=} \textstyle\sum_{k=1}^{m} \pm(x \cdot y) \overset{2)}{=} \sum_{k=1}^{m} (\pm 1_{\mathbb{K}}) \cdot (x \cdot y) \\ &\overset{(51)}{=} \textstyle\sum_{k=1}^{m} (\pm 1_{\mathbb{K}} \cdot x) \cdot y \overset{2)}{=} \sum_{k=1}^{m} (\pm x) \cdot y = (\sum_{k=1}^{m} \pm x) \cdot y \overset{\text{def.}}{=} ((\pm m)x) \cdot y.\end{aligned}$$

Dies beweist das Lemma. □

Übung 39. *Sei $\mathbb{K}$ ein Körper. Zeigen Sie, dass die Gleichung $x^2 = x$ genau zwei Lösungen in $\mathbb{K}$ hat. Welche sind das?*

Satz 6 (Binomischer Lehrsatz)**.** *Sei $\mathbb{K}$ ein Körper. Dann gilt*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k \cdot y^{n-k} \qquad \forall\, x, y \in \mathbb{K},\ n \in \mathbb{N}. \tag{75}$$

Beweis. Zunächst sind die Summanden in (75) definiert, und zwar wegen Korollar 5 (sowie Lemma 16.9)). Wir beweisen die Behauptung nun per Induktion.

(IA): Für $n = 0$ gilt $(x+y)^0 = 1_{\mathbb{K}} = \binom{0}{0}(x^0 \cdot y^0) = \sum_{k=0}^{n} \binom{n}{k} x^k \cdot y^{n-k}$.

(IV): Es gelte (75) für ein $n \in \mathbb{N}$.

(IS): Wir erhalten

$$\begin{aligned}(x+y)^{n+1} &= (x+y) \cdot (x+y)^n \\ &\overset{(\text{IV})}{=} (x+y) \cdot \textstyle\sum_{k=0}^{n} \binom{n}{k} x^k \cdot y^{n-k} \\ &= \textstyle\sum_{k=0}^{n} \binom{n}{k} x^{k+1} \cdot y^{n-k} + \sum_{k=0}^{n} \binom{n}{k} x^k \cdot y^{n+1-k} \\ &= \textstyle\binom{n}{n} x^{n+1} \cdot y^0 + \sum_{k=0}^{n-1} \binom{n}{k} x^{k+1} \cdot y^{n-k} + \sum_{k=1}^{n} \binom{n}{k} x^k \cdot y^{n+1-k} + \binom{n}{0} x^0 \cdot y^{n+1} \\ &= \textstyle x^{n+1} + \sum_{k=1}^{n} \binom{n}{k-1} x^k \cdot y^{n+1-k} + \sum_{k=1}^{n} \binom{n}{k} x^k \cdot y^{n+1-k} + y^{n+1} \\ &= \textstyle x^{n+1} + \sum_{k=1}^{n} \left(\binom{n}{k-1} + \binom{n}{k}\right) x^k \cdot y^{n+1-k} + y^{n+1} \\ &\overset{(47)}{=} \textstyle\binom{n+1}{n+1} x^{n+1} \cdot y^0 + \sum_{k=1}^{n} \binom{n+1}{k} x^k \cdot y^{n+1-k} + \binom{n+1}{0} x^0 \cdot y^{n+1} \\ &= \textstyle\sum_{k=0}^{n+1} \binom{n+1}{k} x^k \cdot y^{n+1-k}.\end{aligned}$$

Die Behauptung folgt nun per Induktion. □

Bemerkung 23. *Man kann Satz 6 auch anschaulich begründen: Ausmultiplizieren von*

$$(x+y)^n = \underbrace{(x+y)}_{\mathrm{T}_1} \cdot \ldots \cdot \underbrace{(x+y)}_{\mathrm{T}_n}$$

ergibt eine Summe von 2^n Termen der Form $x^k \cdot y^{n-k}$, für gewisse $0 \leq k \leq n$. Ist $0 \leq k \leq n$ fixiert, so entspricht die Häufigkeit des Auftretens des Terms $x^k \cdot y^{n-k}$ genau der Anzahl der Möglichkeiten, aus k der n „Töpfe“ $\mathrm{T}_1, \ldots, \mathrm{T}_n$ das Element x „herauszunehmen“, was wiederum der Anzahl der k-elementigen Teilmengen der Menge $\{1, \ldots, n\}$ entspricht, also $\binom{n}{k}$ nach Satz 5.

Bemerkung 24 (Einbettung von $\mathbb{Q}$)**.** *Sei $\mathbb{K}$ ein Körper.*

1) Wir definieren $n_{\mathbb{K}} := n 1_{\mathbb{K}}$ für alle $n \in \mathbb{Z}$. Wir erhalten für $m, n \in \mathbb{Z}$ aus (65)*, dass*

$$(-n)_{\mathbb{K}} = -n_{\mathbb{K}}, \qquad m_{\mathbb{K}} + n_{\mathbb{K}} = (m+n)_{\mathbb{K}}, \qquad n(m_{\mathbb{K}}) = (nm)_{\mathbb{K}} \tag{76}$$

gilt. Wir erhalten aus (74) *(Lemma 16.9)), dass*

$$n_{\mathbb{K}} \cdot x = (n1_{\mathbb{K}}) \cdot x = n(1_{\mathbb{K}} \cdot x) = nx \qquad \forall\, n \in \mathbb{Z} \tag{77}$$

für jedes $x \in \mathbb{K}$ *gilt; also insbesondere*

$$n_{\mathbb{K}} \cdot m_{\mathbb{K}} \overset{(77)}{=} n(m_{\mathbb{K}}) \overset{(76)}{=} (nm)_{\mathbb{K}} \qquad \forall\, m, n \in \mathbb{Z}. \tag{78}$$

2) Wir nehmen nun an, dass die folgende Äquivalenz gilt

$$n_{\mathbb{K}} = 0_{\mathbb{K}} \quad \textit{für} \quad n \in \mathbb{Z} \qquad \Longleftrightarrow \qquad n = 0; \tag{79}$$

und bemerken folgendes:

- *Für alle* $p \in \mathbb{Z}$ *und* $q \in \mathbb{Z}_{\neq 0}$*, ist das Körperelement*

$$\frac{p}{q}_{\mathbb{K}} := \frac{p_{\mathbb{K}}}{q_{\mathbb{K}}} = \frac{p1_{\mathbb{K}}}{q1_{\mathbb{K}}} = p1_{\mathbb{K}} \cdot (q1_{\mathbb{K}})^{-1} \in \mathbb{K}$$

  definiert (es gilt $q1_{\mathbb{K}} \in \mathbb{K}_{\times}$*); und wegen Lemma 16.1) folgt*

$$0_{\mathbb{K}} = \frac{p}{q}_{\mathbb{K}} \in \mathbb{K} \quad \textit{für} \quad p \in \mathbb{Z},\ q \in \mathbb{Z}_{\neq 0} \qquad \Longleftrightarrow \qquad p = 0 \qquad \Longleftrightarrow \qquad 0 = \frac{p}{q} \in \mathbb{Q}. \tag{80}$$

- *Mit Lemma 16 erhalten wir für* $p, p' \in \mathbb{Z}$*,* $p'' \in \mathbb{Z}_{\neq 0}$ *und* $q \in \mathbb{Z}_{\neq 0}$*, die folgenden Identitäten:*

$$\begin{aligned}
\frac{p}{q}_{\mathbb{K}} \cdot \frac{p'}{q'}_{\mathbb{K}} &= \frac{p_{\mathbb{K}}}{q_{\mathbb{K}}} \cdot \frac{p'_{\mathbb{K}}}{q'_{\mathbb{K}}} = \frac{p_{\mathbb{K}} \cdot p'_{\mathbb{K}}}{q_{\mathbb{K}} \cdot q'_{\mathbb{K}}} \overset{(78)}{=} \frac{(pp')_{\mathbb{K}}}{(qq')_{\mathbb{K}}} = \frac{pp'}{qq'}_{\mathbb{K}} \\
\frac{p}{q}_{\mathbb{K}} + \frac{p'}{q'}_{\mathbb{K}} &= \frac{p_{\mathbb{K}}}{q_{\mathbb{K}}} + \frac{p'_{\mathbb{K}}}{q'_{\mathbb{K}}} = \frac{p_{\mathbb{K}} \cdot q'_{\mathbb{K}} + q_{\mathbb{K}} \cdot p'_{\mathbb{K}}}{q_{\mathbb{K}} \cdot q'_{\mathbb{K}}} \overset{(76),(78)}{=} \frac{(pq' + qp')_{\mathbb{K}}}{(qq')_{\mathbb{K}}} = \frac{(pq' + qp')}{qq'}_{\mathbb{K}} \\
-\frac{p}{q}_{\mathbb{K}} &= -\frac{p_{\mathbb{K}}}{q_{\mathbb{K}}} = \frac{-p_{\mathbb{K}}}{q_{\mathbb{K}}} \overset{(76)}{=} \frac{(-p)_{\mathbb{K}}}{q_{\mathbb{K}}} = \frac{-p}{q}_{\mathbb{K}} \\
\left(\frac{p''}{q}_{\mathbb{K}}\right)^{-1} &= \left(\frac{(p'')_{\mathbb{K}}}{q_{\mathbb{K}}}\right)^{-1} = \frac{q_{\mathbb{K}}}{(p'')_{\mathbb{K}}} = \frac{q}{p''}_{\mathbb{K}}.
\end{aligned} \tag{81}$$

- *Seien* $p, p' \in \mathbb{Z}$ *und* $q, q' \in \mathbb{N}_{>0}$ *gegeben. Dann gilt*

$$\frac{p}{q} = \frac{p'}{q'} \quad \Leftrightarrow \quad \frac{pq' - qp'}{qq'} = 0 \quad \overset{(80)}{\Leftrightarrow} \quad \frac{pq' + qp'}{qq'}_{\mathbb{K}} = 0_{\mathbb{K}} \quad \Leftrightarrow \quad \frac{p_{\mathbb{K}}}{q_{\mathbb{K}}} = \frac{p'_{\mathbb{K}}}{q'_{\mathbb{K}}} \quad \Leftrightarrow \quad \frac{p}{q}_{\mathbb{K}} = \frac{p'}{q'}_{\mathbb{K}}.$$

  Daher ist die Abbildung

$$\Psi_{\mathbb{K}} \colon \mathbb{Q} \to \mathbb{K}, \qquad \frac{p}{q} \mapsto \frac{p}{q}_{\mathbb{K}} \tag{82}$$

  wohldefiniert und injektiv. Die ersten beiden Zeilen in (81) *zeigen dann bereits, dass* $\Psi_{\mathbb{K}}$ *ein Körperhomomorphismus ist.*

Gilt also die Bedingung (79)*, so ist* $\mathbb{Q}$ *eingebettet in* $\mathbb{K}$*, und zwar vermöge der Körpereinbettung* (82)*. Insbesondere kann dann* $\mathbb{K}$ *nicht endlich sein; genauer gilt* $|\mathbb{K}| \geq |\mathbb{Q}|$ *(vgl. Bemerkung 20 und Lemma 9.1)).*

Übung 40 (Endliche Körper)**.** *Sei* $k \in \mathbb{N}$*, sowie* $\mathbb{Z}/\!\sim_k$ *und* $+, \cdot, -$ *wie in Beispiel 1 definiert. Zeigen Sie die folgenden Aussagen:*

a) Für alle $k \in \mathbb{N}$, ist $(\mathbb{Z}/{\sim_k}, +, [0]_{\sim_k})$, ist eine abelsche Gruppe.

b) Für $k \geq 2$, ist $(\mathbb{Z}/{\sim_k}, \cdot, [1]_{\sim_k})$ ein abelscher Monoid; und es gilt das Distributivgesetz

$$\alpha \cdot (\beta + \gamma) = (\alpha \cdot \beta) + (\alpha \cdot \gamma) \qquad \forall\, \alpha, \beta, \gamma \in \mathbb{Z}/{\sim_k}.$$

c) Für $k = 2, 3, 5$, ist

$$(\mathbb{Z}/{\sim_k} \setminus \{[0]_{\sim_k}\}, \cdot, [1]_{\sim_k})$$

ist eine abelsche Gruppe (multiplikative Inverse); also $\mathbb{Z}/{\sim_k}$ ein Körper.

d) Gilt $k = p \cdot q$ für gewisse $\mathbb{N} \ni p, q \geq 2$, so ist $\mathbb{Z}/{\sim_k}$ kein Körper.

Hinweis: Lemma 16.1).

Bemerkung: Wie Teil c) bereits vermuten lässt, kann man zeigen, dass $\mathbb{Z}/{\sim_k}$ für jede Primzahl $k \geq 2$ ein Körper ist – dieser wird gemeinhin mit $\mathbb{F}_k$ notiert. Aus dem Satz über die Primfaktorzerlegung folgt nun, dass jedes $\mathbb{N} \ni k \geq 2$ welches keine Primzahl ist, als $k = p \cdot q$ für gewisse $\mathbb{N} \ni p, q \geq 2$ geschrieben werden kann. Punkt d) zeigt daher, dass $\mathbb{Z}/{\sim_k}$ auch nur für Primzahlen ein Körper ist.

4.2.3 Angeordnete Körper

Das nächste Lemma formalisiert die Aussagen für die rationalen Zahlen aus Übung 16.

Lemma 17. *Sei $(\mathbb{K}, \leq)$ ein angeordneter Körper. Dann gelten die folgenden Aussagen für alle $x, y, a, b \in \mathbb{K}$:*

$$
\begin{array}{lrcl}
1) & x \leq y \ \wedge \ a \leq b & \Longrightarrow & x + a \leq y + b \\
2) & x < y \ \wedge \ a \leq b & \Longrightarrow & x + a < y + b \\
3) & x \leq y \ \wedge \ a \geq 0_{\mathbb{K}} & \Longrightarrow & a \cdot x \leq a \cdot y \\
4) & x < y \ \wedge \ a > 0_{\mathbb{K}} & \Longrightarrow & a \cdot x < a \cdot y \\
5) & 0_{\mathbb{K}} \leq x < y \ \wedge \ 0_{\mathbb{K}} \leq a < b & \Longrightarrow & 0_{\mathbb{K}} \leq a \cdot x < b \cdot y \\
6) & x < y & \Longrightarrow & -y < -x \\
7) & y \leq x \ \wedge \ a \leq 0_{\mathbb{K}} & \Longrightarrow & a \cdot x \leq a \cdot y \\
8) & x \neq 0_{\mathbb{K}} & \Longrightarrow & x^2 > 0_{\mathbb{K}} \\
9) & 1_{\mathbb{K}} > 0_{\mathbb{K}} & & \\
10) & x > 0_{\mathbb{K}} & \Longrightarrow & x^{-1} > 0_{\mathbb{K}} \\
11) & 0_{\mathbb{K}} < x < y & \Longrightarrow & 0_{\mathbb{K}} < y^{-1} < x^{-1} \\
12) & y < x < 0_{\mathbb{K}} & \Longrightarrow & x^{-1} < y^{-1} < 0_{\mathbb{K}}
\end{array}
$$

Erinnerung: Für $x, y, z \in \mathbb{K}$ gelten die Implikationen:

$$x \leq y \quad \Longrightarrow \quad x + z \leq y + z \tag{68}$$

$$x < y \quad \Longrightarrow \quad x + z < y + z \tag{70}$$

$$x, y \geq 0_{\mathbb{K}} \quad \Longrightarrow \quad x \cdot y \geq 0_{\mathbb{K}}. \tag{69}$$

Beweis. 1) Zweimaliges anwenden von (68), liefert $x + a \leq y + a \leq y + b$.

2) Anwenden von (70) und dann (68), liefert $x + a < y + a \leq y + b$.

3) Addition von $-x$ liefert $0_{\mathbb{K}} \leq y - x$ wegen (68). Anwendung von (69) liefert $0_{\mathbb{K}} \leq a \cdot (y - x) = a \cdot y - a \cdot x$. Addition von $a \cdot x$ liefert $a \cdot x \leq a \cdot y$ (wegen (68)).

4) Wegen Teil 3) gilt $a \cdot x \leq a \cdot y$. Nun kann $a \cdot x = a \cdot y$ wegen $a \in \mathbb{K}_\times$ nicht gelten, da ansonsten Multiplikation mit a^{-1}, die Gleichheit $x = y$ impliziert. Dies zeigt $a \cdot x \leq a \cdot y$.

5) Aus Teil 3) folgt $0_\mathbb{K} = a \cdot 0_\mathbb{K} \leq a \cdot x \leq a \cdot y < b \cdot y$. Für den letzten Schritt beachte man, dass $a \cdot y = b \cdot y$ wegen $y \in \mathbb{K}_\times$, bereits $a = b$ implizieren würde, was den Voraussetzungen widerspricht.

6) Dies folgt durch Addition von $-x - y$ auf beiden Seiten.

7) Wegen Teil 6) gilt $-a \geq 0_\mathbb{K}$; also $(-a) \cdot y \leq (-a) \cdot x$ wegen Teil 3); also $a \cdot x \leq a \cdot y$ wegen Teil 6).

8) Wegen Lemma 16.5) gilt $x^2 = (-x)^2$. Ist $x > 0_\mathbb{K}$, so folgt $x^2 > 0_\mathbb{K}$ aus Teil 5). Ist $x < 0_\mathbb{K}$, so ist $(-x) > 0_\mathbb{K}$ wegen Teil 6); und es folgt ebenfalls $x = (-x)^2 > 0_\mathbb{K}$.

9) Wegen 8) und $1_\mathbb{K} \neq 0_\mathbb{K}$ gilt $1_\mathbb{K} = (1_\mathbb{K})^2 > 0_\mathbb{K}$.

10) Es ist $x^{-2} = (x^{-1})^2 > 0_\mathbb{K}$ wegen Teil 8); also folgt die Behauptung aus Teil 3) durch Multiplikation von $x > 0_\mathbb{K}$ mit $(x^{-1})^2$.

11) Dies folgt durch Multiplikation mit $y^{-1} \cdot x^{-1}$. □

12) Multiplikation mit -1 liefert $0_\mathbb{K} < (-x) < (-y)$ wegen Teil 6). Teil 11) zeigt dann $0_\mathbb{K} < (-y)^{-1} < (-x)^{-1}$; und Multiplikation mit -1 liefert wegen Teil 6) und Lemma 16.4) die Behauptung.

Übung 41. *Sei $\mathbb{K}$ ein angeordneter Körper. Zeigen Sie, dass die Gleichung $x = -x$ genau eine Lösungen in $\mathbb{K}$ hat. Welche ist das?*

Lemma 18. *Ist $(\mathbb{K}, \leq)$ ein angeordneter Körper, so gilt $n_\mathbb{K} > 0_\mathbb{K}$ für alle $n \in \mathbb{N}_{>0}$.*

Beweis. Nach Lemma 17.9) gilt $1_\mathbb{K} > 0_\mathbb{K}$ (IA). Gilt nun $n_\mathbb{K} > 0$ für ein $n \in \mathbb{N}_{>0}$ (IV), so folgt $(n+1)_\mathbb{K} = n_\mathbb{K} + 1_\mathbb{K} > 0_\mathbb{K} + 0_\mathbb{K} = 0_\mathbb{K}$ aus Lemma 17.1) . Per Induktion folgt daher $n_\mathbb{K} > 0_\mathbb{K}$ für alle $n \in \mathbb{N}_{>0}$. □

Wir erinnern an die Diskussionen in Bemerkung 24.2). Der nächste Satz zeigt, dass der Körper der rationalen Zahlen vermöge der Abbildung (82) in jeden angeordneten Körper eingebettet ist.

Satz 7. *Sei $(\mathbb{K}, \leq)$ ein angeordneter Körper. Dann ist die Abbildung $\Psi_\mathbb{K} \colon \mathbb{Q} \ni \frac{p}{q} \mapsto \frac{p}{q}_\mathbb{K} \in \mathbb{K}$ definiert und eine Körpereinbettung.*

Beweis. • Wegen Lemma 18, gilt $n_\mathbb{K} > 0_\mathbb{K}$ für alle $n \in \mathbb{N}_{>0}$; und aus (76) sowie Lemma 17.6) (zweiter Schritt) folgt dann

$$(-n)_\mathbb{K} \overset{(76)}{=} -n_\mathbb{K} < 0_\mathbb{K} \qquad \forall\, n \in \mathbb{N}_{>0}.$$

Daher gilt (79) (also: $n_\mathbb{K} = 0_\mathbb{K}$ für $n \in \mathbb{Z} \quad \Leftrightarrow \quad n = 0$), sodass $\Psi_\mathbb{K}$ nach Bemerkung 24.2) ein wohldefinierter injektiver Körperhomomorphismus ist (wenn $\mathbb{Q}, \mathbb{K}$ als Körper ohne Anordnung aufgefasst).

• Es bleibt noch (72) nachzuweisen. (also: $x \leq y$ für $x, y \in \mathbb{Q} \quad \Rightarrow \quad \Psi(x) \leq \Psi(y)$)

- Hierfür genügt es zu zeigen, dass $\Psi_\mathbb{K}(z) \geq 0_\mathbb{K}$ für alle $\mathbb{Q} \ni z \geq 0$ gilt; denn dann folgt

$$x \leq y \quad \Rightarrow \quad y - x \geq 0 \quad \Rightarrow \quad \Psi_\mathbb{K}(y) - \Psi_\mathbb{K}(x) = \Psi_\mathbb{K}(y - x) \geq 0_\mathbb{K} \quad \Rightarrow \quad \Psi_\mathbb{K}(x) \leq \Psi_\mathbb{K}(y).$$

- Sei nun $\mathbb{Q} \ni \frac{p}{q} \geq 0$ vorgegeben. Wir können ohne Einschränkung $p \in \mathbb{N}$ und $q \in \mathbb{N}_{>0}$ annehmen. Dann gilt $p_\mathbb{K} \geq 0_\mathbb{K}$ und $q_\mathbb{K} > 0_\mathbb{K}$ wegen Lemma 18; also auch $(q_\mathbb{K})^{-1} > 0_\mathbb{K}$ nach Lemma 17.10).
- Mit (69) folgt ($x, y \geq 0_\mathbb{K}$ für $x, y \in \mathbb{K} \quad \Rightarrow \quad x \cdot y \geq 0_\mathbb{K}$)

$$\Psi_\mathbb{K}\Big(\frac{p}{q}\Big) = \frac{p}{q}_\mathbb{K} = p_\mathbb{K} \cdot (q_\mathbb{K})^{-1} \geq 0_\mathbb{K},$$

was die Behauptung zeigt. □

Korollar 8. *Sei* $(\mathbb{K}, \leq)$ *ein angeordneter Körper. Für* $x, y \in \mathbb{K}$ *mit* $x < y$*, gilt* $x < \frac{1}{2}_{\mathbb{K}} \cdot (x + y) < y$.

Beweis. Es gilt $0 < \frac{1}{2} < 1$ in $\mathbb{Q}$, also $0_{\mathbb{K}} < \frac{1}{2}_{\mathbb{K}} < 1_{\mathbb{K}}$ nach Satz 7. Weiterhin gilt $(y - x) > 0_{\mathbb{K}}$ (wegen (68)); also folgt aus Lemma 17.5), dass gilt:

$$0_{\mathbb{K}} < \frac{1}{2}_{\mathbb{K}} \cdot (y - x) < 1_{\mathbb{K}} \cdot (y - x).$$

Addition mit $x = 1_{\mathbb{K}} \cdot x$ liefert

$$x < \underbrace{1_{\mathbb{K}} \cdot x + \frac{1}{2}_{\mathbb{K}} \cdot (y - x)}_{\alpha} < y \qquad \text{mit} \qquad \alpha \overset{(81)}{=} \frac{1}{2}_{\mathbb{K}} \cdot (x + y),$$

was die Behauptung zeigt. □

Lemma 19. *Sei* $(\mathbb{K}, \leq)$ *ein angeordneter Körper, sowie* $x, y \in \mathbb{K}$.

1) Ist $y \leq x + \varepsilon$ *für alle* $\varepsilon > 0_{\mathbb{K}}$*, so gilt* $y \leq x$.

2) Ist $x - \varepsilon \leq y$ *für alle* $\varepsilon > 0_{\mathbb{K}}$*, so gilt* $x \leq y$.

Beweis. 1) Beweis durch Kontraposition: Sei $x < y$. Nach Korollar 8 ist dann

$$x < \frac{1}{2}_{\mathbb{K}} \cdot (x + y) < y \qquad \Longrightarrow \qquad y > x + \underbrace{\frac{1}{2}_{\mathbb{K}} \cdot (x + y) - x}_{=: \varepsilon > 0_{\mathbb{K}}}.$$

2) Dies folgt aus 1), wenn man dort die Rollen von x und y vertauscht. □

Lemma 19 wird oft auch in der folgenden Form angewandt:

Korollar 9. *Sei* $(\mathbb{K}, \leq)$ *ein angeordneter Körper, sowie* $x, y \in \mathbb{K}$ *und* $0_{\mathbb{K}} < \tau \in \mathbb{K}$.

1) Ist $y \leq x + \varepsilon$ *für alle* $0_{\mathbb{K}} < \varepsilon < \tau$*, so gilt* $y \leq x$.

2) Ist $x - \varepsilon \leq y$ *für alle* $0_{\mathbb{K}} < \varepsilon < \tau$*, so gilt* $x \leq y$.

Beweis. Die Behauptungen folgen umgehend aus den entsprechenden Behauptungen in Lemma 19; denn wir haben die folgen Implikationen:

- Gilt $y \leq x + \varepsilon$ für alle $0_{\mathbb{K}} < \varepsilon < \tau$,[12] so gilt auch $y \leq x + \varepsilon$ für alle $\varepsilon > 0_{\mathbb{K}}$:
  Ist nämlich $\tilde{\varepsilon} \geq \tau$ vorgegeben, so gilt $0_{\mathbb{K}} < \varepsilon := \frac{1}{2}_{\mathbb{K}} \cdot \tau < \tau$ nach Korollar 8; und wir erhalten
  $$y \leq x + \varepsilon \qquad \Longrightarrow \qquad y \leq x + \varepsilon < x + \tau \leq x + \tilde{\varepsilon}.$$
- Gilt $x - \varepsilon \leq y$ für alle $0_{\mathbb{K}} < \varepsilon < \tau$, so gilt auch $x - \varepsilon \leq y$ für alle $\varepsilon > 0_{\mathbb{K}}$:
  Ist nämlich $\tilde{\varepsilon} \geq \tau$ vorgegeben, so gilt $0_{\mathbb{K}} < \varepsilon := \frac{1}{2}_{\mathbb{K}} \cdot \tau < \tau$ nach Korollar 8; und wir erhalten
  $$x - \varepsilon \leq y \qquad \Longrightarrow \qquad y \geq x - \varepsilon > x - \tau \geq x - \tilde{\varepsilon}.$$

Die zeigt die Behauptung. □

Satz 8 (Die Bernoullische Ungleichung)**.** *Sei* $(\mathbb{K}, \leq)$ *ein angeordneter Körper. Dann gilt*

$$1_{\mathbb{K}} + nx \leq (1_{\mathbb{K}} + x)^n \qquad \forall\, n \in \mathbb{N},\ x \geq -1_{\mathbb{K}}.$$

[12]Prinzipiell könnte dies zunächst eine Nullbedingung sein, da ad hoc kein derartiges $0_{\mathbb{K}} < \varepsilon < \tau$ existieren muss.

Beweis. Die Aussage ist klar für $n = 0$ (IA). Gilt nun die Behauptung für ein $n \in \mathbb{N}$ (IV), so folgt

$$\begin{aligned}
(1_{\mathbb{K}} + x)^{n+1} &= (1_{\mathbb{K}} + x)^n \cdot (1_{\mathbb{K}} + x) \\
&\overset{\text{(IV)}}{\geq} (1_{\mathbb{K}} + nx) \cdot (1_{\mathbb{K}} + x) \\
&= 1_{\mathbb{K}} + nx + x + nx^2 \\
&\overset{(76),(77)}{=} 1_{\mathbb{K}} + (n+1)x + n_{\mathbb{K}} \cdot x^2 \\
&\geq 1_{\mathbb{K}} + (n+1)x.
\end{aligned}$$

Im zweiten Schritt haben wir zusätzlich (69) sowie $(1_{\mathbb{K}} + x) \geq 0_{\mathbb{K}}$ angewandt (beachte $(1_{\mathbb{K}} + x)^n - (1_{\mathbb{K}} + nx) \geq 0_{\mathbb{K}}$ nach (IV)). Im letzten Schritt haben wir $x^2 \geq 0_{\mathbb{K}}$ (Lemma 17.8)), $n_{\mathbb{K}} \geq 0_{\mathbb{K}}$ (Lemma 18), sowie (69) benutzt. □

Wir erinnern an die Begrifflichkeiten des Supremums und des Infimums aus Definition 13, sowie an die Äquivalenzen der Begriffe größtes/kleinstes bzw. maximales/minimales Element im angeordneten Fall (siehe Lemma 2).

Lemma 20. *Sei* $(\mathbb{K}, \leq)$ *ein angeordneter Körper, und* $\emptyset \neq X \subseteq \mathbb{K}$ *eine nichtleere Teilmenge. Dann gelten die folgenden Äquivalenzen:*

$$\mathbb{K} \ni o = \sup_{\mathbb{K}}(X) \iff X \leq o \quad \wedge \quad \forall\, \mathbb{K} \ni \varepsilon > 0_{\mathbb{K}} \colon \exists\, x \in X \colon x > o - \varepsilon \tag{83}$$

$$\mathbb{K} \ni u = \inf_{\mathbb{K}}(X) \iff u \leq X \quad \wedge \quad \forall\, \mathbb{K} \ni \varepsilon > 0_{\mathbb{K}} \colon \exists\, x \in X \colon x < u + \varepsilon. \tag{84}$$

Beweis. Wir zeigen nur (83). Die Äquivalenz (84) folgt auf die gleiche Weise.

- Sei $o = \sup_{\mathbb{K}}(X)$. Gemäß Definition 13, ist dann o das kleinste Element der Menge $O_{\mathbb{K}}(X) = \{o \in \mathbb{K} \mid X \leq o\}$ aller oberen Schranken von X. Daher gilt auch $X \leq o$. Ist nun weiterhin $\varepsilon > 0_{\mathbb{K}}$ vorgegeben, so gilt $o - \varepsilon < o$, also $(o - \varepsilon) \notin O_{\mathbb{K}}(X)$ (wegen $o \leq O_{\mathbb{K}}(X)$).
  - Dann gilt **nicht** $X \leq o - \varepsilon$; also existiert ein $x \in X$, sodass **nicht** $x \leq (o - \varepsilon)$ gilt.
  - Da $\mathbb{K}$ angeordnet ist, gilt somit $x > o - \varepsilon$.
- Es gelte die rechte Seite von (83).
  - Wegen $X \leq o$ gilt $o \in O_{\mathbb{K}}(X)$.
  - Angenommen, es existiert ein $s \in O_{\mathbb{K}}(X)$ mit $s < o$. Dann gilt $\varepsilon := o - s > 0_{\mathbb{K}}$. Wegen der rechten Seite von (83), existiert somit ein $x \in X$ mit $x > o - \varepsilon = s$. Dies widerspricht der Annahme $X \leq s$. Da $\mathbb{K}$ angeordnet ist, zeigt dies $o \leq O_{\mathbb{K}}(X)$.

  Folglich ist o das kleinste Element von $O_{\mathbb{K}}(X)$. □

Übung 42. *Zeigen Sie die Äquivalenz* (84).

Proposition 1. *Sei* $(\mathbb{K}, \leq)$ *ein angeordneter Körper, und* $\emptyset \neq X, Y \subseteq \mathbb{K}$ *nichtleere Teilmengen. Dann gelten die folgenden Aussagen:*

1) $\sup_{\mathbb{K}}(X + Y) = \sup_{\mathbb{K}}(X) + \sup_{\mathbb{K}}(Y)$, *falls* $\sup_{\mathbb{K}}(X)$ *und* $\sup_{\mathbb{K}}(Y)$ *existieren.*

2) $\inf_{\mathbb{K}}(X + Y) = \inf_{\mathbb{K}}(X) + \inf_{\mathbb{K}}(Y)$, *falls* $\inf_{\mathbb{K}}(X)$ *und* $\inf_{\mathbb{K}}(Y)$ *existieren.*

3) $\sup_{\mathbb{K}}(X \cdot Y) = \sup_{\mathbb{K}}(X) \cdot \sup_{\mathbb{K}}(Y)$ *für* $X, Y \geq 0_{\mathbb{K}}$, *falls* $\sup_{\mathbb{K}}(X)$ *und* $\sup_{\mathbb{K}}(Y)$ *existieren.*

4) $\inf_{\mathbb{K}}(X \cdot Y) = \inf_{\mathbb{K}}(X) \cdot \inf_{\mathbb{K}}(Y)$ *für* $X, Y \geq 0_{\mathbb{K}}$, *falls* $\inf_{\mathbb{K}}(X)$ *und* $\inf_{\mathbb{K}}(Y)$ *beide existieren.*

5) $\inf_{\mathbb{K}}(X) = -\sup_{\mathbb{K}}(-X)$, *falls* $\sup_{\mathbb{K}}(-X)$ *existiert.*

6) $\sup_{\mathbb{K}}(X) = -\inf_{\mathbb{K}}(-X)$*, falls* $\inf_{\mathbb{K}}(-X)$ *existiert.*

7) $\sup_{\mathbb{K}}(X \cup Y) = \max(\sup_{\mathbb{K}}(X), \sup_{\mathbb{K}}(Y))$*, falls* $\sup_{\mathbb{K}}(X)$ *und* $\sup_{\mathbb{K}}(Y)$ *existieren.*

8) $\inf_{\mathbb{K}}(X \cup Y) = \min(\inf_{\mathbb{K}}(X), \inf_{\mathbb{K}}(Y))$*, falls* $\inf_{\mathbb{K}}(X)$ *und* $\inf_{\mathbb{K}}(Y)$ *existieren.*

9) $\sup_{\mathbb{K}}(X) \leq \sup_{\mathbb{K}}(Y)$ *für* $X \subseteq Y$*, falls* $\sup_{\mathbb{K}}(X)$ *und* $\sup_{\mathbb{K}}(Y)$ *existieren.*

10) $\inf_{\mathbb{K}}(Y) \leq \inf_{\mathbb{K}}(X)$ *für* $X \subseteq Y$*, falls* $\inf_{\mathbb{K}}(X)$ *und* $\inf_{\mathbb{K}}(Y)$ *existieren.*

Beweis. Im Falle der jeweiligen Existenz sei

$$o_X := \sup_{\mathbb{K}}(X), \qquad o_Y := \sup_{\mathbb{K}}(Y), \qquad u_X := \inf_{\mathbb{K}}(X), \qquad u_Y := \inf_{\mathbb{K}}(Y)$$

Wir beweisen die ersten vier Aussagen mit Hilfe der Äquivalenz (83) in Lemma 20.

1) Sei $o := o_X + o_Y$.

- Es gilt $X + Y \leq o$, wegen Lemma 17.1). $\qquad (x + y \leq o_X + o_Y = o)$
- Sei nun $\varepsilon > 0_{\mathbb{K}}$ vorgegeben; sowie $\tilde{\varepsilon} := \frac{1}{2}_{\mathbb{K}} \cdot \varepsilon$.
  - Es gilt $\tilde{\varepsilon} + \tilde{\varepsilon} = \varepsilon$, wegen (81) und (66).
  - Es gilt $0_{\mathbb{K}} < \tilde{\varepsilon} < \varepsilon$, wegen Korollar 8.

  Gemäß (83), existieren $x \in X$ und $y \in Y$ mit $x > o_X - \tilde{\varepsilon}$ und $y > o_Y - \tilde{\varepsilon}$. Wir erhalten aus Lemma 17.2), dass gilt:

$$X + Y \ni x + y > (o_X - \tilde{\varepsilon}) + (o_Y - \tilde{\varepsilon}) = o_X + o_Y - (\tilde{\varepsilon} + \tilde{\varepsilon}) = o - \varepsilon.$$

Die Behauptung folgt nun aus (83).

2) Dies folgt analog zu Teil 1) (Übung 43).

3) Sei $o := o_X \cdot o_Y$; und beachte $o_X, o_Y \geq 0_{\mathbb{K}}$ (wegen $X, Y \geq 0_{\mathbb{K}}$). Die Behauptung ist klar, wenn $o_X = 0_{\mathbb{K}}$ oder $o_Y = 0_{\mathbb{K}}$ gilt; denn dann ist $X \cdot Y = \{0_{\mathbb{K}}\}$, also $\sup_{\mathbb{K}}(X \cdot Y) = 0_{\mathbb{K}} = 0_{\mathbb{K}} + 0_{\mathbb{K}} = o$. Wir können daher annehmen, das $\delta := \min(o_X, o_Y) > 0_{\mathbb{K}}$ gilt:

- Es gilt $X \cdot Y \leq o$ wegen Lemma 17.3). $\qquad (x \cdot y \leq x \cdot o_Y \leq o_X \cdot o_Y)$
- Sei nun $\varepsilon > 0_{\mathbb{K}}$ vorgegeben. Mit (69) und Lemma 17 folgt

$$\tilde{\varepsilon} := \min\left(\tfrac{\varepsilon}{o_X + o_Y}, \tfrac{1}{2}_{\mathbb{K}} \cdot \delta\right) > 0_{\mathbb{K}}.$$

  - Mit Lemma 17 und Korollar 8 (letzter Schritt) folgt

$$\begin{aligned} o_X - \tilde{\varepsilon} &\geq o_X - \tfrac{1}{2}_{\mathbb{K}} \cdot \delta \geq o_X - \tfrac{1}{2}_{\mathbb{K}} \cdot o_X = \tfrac{1}{2}_{\mathbb{K}} \cdot o_X \geq 0_{\mathbb{K}} \\ o_Y - \tilde{\varepsilon} &\geq o_Y - \tfrac{1}{2}_{\mathbb{K}} \cdot \delta \geq o_Y - \tfrac{1}{2}_{\mathbb{K}} \cdot o_Y = \tfrac{1}{2}_{\mathbb{K}} \cdot o_Y \geq 0_{\mathbb{K}}. \end{aligned}$$

  - Gemäß (83) existieren $x \in X$ und $y \in Y$, mit $x > o_X - \tilde{\varepsilon}$ und $y > o_Y - \tilde{\varepsilon}$.

  Wegen Lemma 17.5) (erster Schritt), $\tilde{\varepsilon}^2 > 0_{\mathbb{K}}$ (dritter Schritt), und $-(o_X + o_Y) \cdot \tilde{\varepsilon} \geq -\varepsilon$ (letzter Schritt), folgt

$$X \cdot Y \ni x \cdot y > (o_X - \tilde{\varepsilon}) \cdot (o_Y - \tilde{\varepsilon}) = o_X \cdot o_Y - (o_X + o_Y) \cdot \tilde{\varepsilon} + \tilde{\varepsilon}^2 > o - (o_X + o_Y) \cdot \tilde{\varepsilon} \geq o - \varepsilon.$$

Die Behauptung folgt nun aus (83).

4) Sei $u := u_X \cdot u_Y$; und beachte $u_X, u_Y \geq 0_{\mathbb{K}}$ (wegen $X, Y \geq 0_{\mathbb{K}}$). Die Behauptung ist klar, wenn $u_X = 0_{\mathbb{K}}$ oder $u_Y = 0_{\mathbb{K}}$ gilt; denn dann ist $X \cdot Y = \{0_{\mathbb{K}}\}$, also $\inf_{\mathbb{K}}(X \cdot Y) = 0_{\mathbb{K}} = u$. Wir können daher annehmen, das $u_X, u_Y > 0_{\mathbb{K}}$ gilt:

- Es gilt $u \leq X \cdot Y$, wegen Lemma 17.3).

- Sei nun $\varepsilon > 0_{\mathbb{K}}$ vorgegeben; und sei $\tilde{\varepsilon} := \min\left(\frac{\varepsilon}{u_X+u_Y+1}, 1\right) > 0_{\mathbb{K}}$. Nach (83) existieren $x \in X$ und $y \in Y$, mit $0_{\mathbb{K}} \leq x < o_X + \tilde{\varepsilon}$ und $0_{\mathbb{K}} \leq y < o_Y + \tilde{\varepsilon}$. Wegen Lemma 17.5) (erster Schritt), $\tilde{\varepsilon} \leq 1$ (dritter Schritt), und $\tilde{\varepsilon} \leq u_X + u_Y + 1$ (letzter Schritt), folgt

$$\begin{aligned} X \cdot Y \ni x \cdot y &< (u_X + \tilde{\varepsilon}) \cdot (u_Y + \tilde{\varepsilon}) \\ &= u_X \cdot u_Y + (u_X + u_Y + \tilde{\varepsilon}) \cdot \tilde{\varepsilon} \\ &\leq u + (u_X + u_Y + 1) \cdot \tilde{\varepsilon} \\ &\leq u + \varepsilon. \end{aligned}$$

Die Behauptung folgt nun aus (83).

5) Wir zeigen die Behauptung unter Verwendung von Lemma 17.6):

- Es ist $-\sup_{\mathbb{K}}(-X) \in U_{\mathbb{K}}(X) = \{u \in \mathbb{K} \mid u \leq X\}$ eine untere Schranke von X, wegen

$$\begin{aligned} -x \leq \sup\nolimits_{\mathbb{K}}(-X) \quad \forall\, x \in X \qquad &\Longrightarrow \qquad -\sup\nolimits_{\mathbb{K}}(-X) \leq x \quad \forall\, x \in X \\ &\Longrightarrow \qquad -\sup\nolimits_{\mathbb{K}}(-X) \leq X. \end{aligned}$$

- Sei $U_{\mathbb{K}}(X) \ni u \leq X$ eine untere Schranke von X. Dann folgt

$$-X \leq -u \quad \overset{\text{def.}}{\Longleftrightarrow} \quad -u \in O_{\mathbb{K}}(-X) \quad \Longrightarrow \quad \sup\nolimits_{\mathbb{K}}(-X) \leq -u \quad \Longrightarrow \quad u \leq -\sup\nolimits_{\mathbb{K}}(-X).$$

Zusammen zeigt dies $\inf_{\mathbb{K}}(X) = -\sup_{\mathbb{K}}(-X)$.

6) Dies folgt analog zu Teil 5) (Übung 43).

7) Wir argumentieren wie folgt:

- Wegen $X \cup Y \leq \max(o_X, o_Y)$, ist $\max(o_X, o_Y) \in O_{\mathbb{K}}(X)$ eine obere Schranke von $X \cup Y$.
- Ist $X \cup Y \leq o \in O_{\mathbb{K}}(X \cup Y)$ eine obere Schranke von $X \cup Y$, so folgt

$$X \leq o \quad \Longrightarrow \quad o_X \leq o \qquad \text{sowie} \qquad Y \leq o \quad \Longrightarrow \quad o_Y \leq o,$$

also $\max(o_X, o_Y) \leq o$.

Dies zeigt $\sup(X \cup Y) = \max(o_X, o_Y) = \max(\sup_{\mathbb{K}}(X), \sup_{\mathbb{K}}(Y)$.

8) Dies folgt analog zu Teil 7) (Übung 43).

9) Gilt $X \subseteq Y$, so ist $\sup_{\mathbb{K}}(Y)$ eine obere Schranke von X; also gilt $\sup_{\mathbb{K}}(X) \leq \sup_{\mathbb{K}}(Y)$.

10) Gilt $X \subseteq Y$, so ist $\inf_{\mathbb{K}}(Y)$ eine untere Schranke von X; also gilt $\inf_{\mathbb{K}}(Y) \leq \inf_{\mathbb{K}}(X)$. □

Bemerkung* 2. *Proposition 1 gilt wegen Lemma 2, Lemma 5 und Übung 18 in der gleichen Weise, wenn man überall* $\sup_{\mathbb{K}}$ *durch* max*, sowie* $\inf_{\mathbb{K}}$ *durch* min *ersetzt (vgl. auch Lemma 3.3)).*

Exemplarisch für Proposition 1.1):

- *Es gelte* $\max(X), \max(Y) \in \mathbb{K}$. *Wegen Lemma 2 ist* $\max(X)$ *größtes Element von* X*, sowie* $\max(Y)$ *größtes Element von* Y. *Wegen Lemma 5 gilt dann* $\sup_{\mathbb{K}}(X) = \max(X)$ *und* $\sup_{\mathbb{K}}(Y) = \max(Y)$.
- *Mit Proposition 1.1) folgt*

$$\sup\nolimits_{\mathbb{K}}(X + Y) = \sup\nolimits_{\mathbb{K}}(X) + \sup\nolimits_{\mathbb{K}}(Y) = \max(X) + \max(Y) \in X + Y.$$

Übung 18 zeigt dann, dass $\sup_{\mathbb{K}}(X + Y)$ *größtes Element von* $X + Y$ *ist; sodass* $\max(X + Y) = \sup_{\mathbb{K}}(X + Y)$ *wegen Lemma 2 gilt.*

- *Zusammen folgt* $\max(X+Y) = \max(X) + \max(Y)$.

Übung 43. *Zeigen Sie die Aussagen 2), 6), 8) in Proposition 1.*

Bemerkung 25 (Die Betragsfunktion)**.**

1) *Gegeben ein angeordneter Körper* $(\mathbb{K}, \leq)$, *so definiert man die zugehörige Betragsfunktion durch* $|x| := \max(x, -x)$. *Wegen Lemma 17.6) gilt dann (vgl.* (37) *für* $\mathbb{Q}$*)*

$$|x| = \begin{cases} x & \text{für } x \geq 0_{\mathbb{K}}, \\ -x & \text{für } x < 0_{\mathbb{K}}. \end{cases} \tag{85}$$

Vermöge Fallunterscheidung, erhält man aus Lemma 16 und Lemma 17 weiterhin, dass

$$x \leq |x|, \qquad |x| = |-x|, \qquad |x \cdot y| = |x| \cdot |y|, \qquad |x+y| \leq |x| + |y| \tag{86}$$

für alle $x, y \in \mathbb{K}$ *gilt. (Übung 44)*

2) *Gegeben angeordnete Körper* $(\mathbb{K}, \leq)$ *und* $(\tilde{\mathbb{K}}, \tilde{\leq})$ *sowie ein Körperhomomorphismus* $\Psi\colon \mathbb{K} \to \tilde{\mathbb{K}}$, *gilt automatisch (beachte* $\Psi(0_{\mathbb{K}}) = 0_{\tilde{\mathbb{K}}}$*)*

$$|\Psi(x)| = \Psi(|x|) \qquad \forall\, x \in \mathbb{K}. \tag{87}$$

Übung 44. *Zeigen Sie die Abschätzungen und Identitäten in* (85), (86) *und* (87).

Übung 45. *Sei* $(\mathbb{K}, \leq)$ *ein angeordneter Körper. Zeigen Sie die folgenden Aussagen:*

a) Seien $x, y \in \mathbb{K}$ *gegeben mit* $0_{\mathbb{K}} < x < y$. *Dann gilt* $0_{\mathbb{K}} < x^n < y^n$ *für alle* $n \in \mathbb{N}_{>0}$.

Hinweis: Induktion.

b) Seien $0_{\mathbb{K}} \leq x, y \in \mathbb{K}$ *und* $k \in \mathbb{N}_{>0}$ *gegeben, mit* $x^k = y^k$. *Dann gilt* $x = y$.

Übung 46. *Sei* $(\mathbb{K}, \leq)$ *ein angeordneter Körper, und* $x \in \mathbb{K}$. *Zeigen Sie:*

a) Ist $0_{\mathbb{K}} < x < 1_{\mathbb{K}}$, *so gilt* $0_{\mathbb{K}} < x^n < 1_{\mathbb{K}}$ *für alle* $n \in \mathbb{N}_{>0}$.

Folgern Sie aus Lemma 17.4), dass $x^n > x^{n+\ell}$ *für alle* $n, \ell \in \mathbb{N}$ *gilt.*

b) Ist $1_{\mathbb{K}} < x$, *so gilt* $1_{\mathbb{K}} < x^n$ *für alle* $n \in \mathbb{N}_{>0}$.

Folgern Sie aus Lemma 17.4), dass $x^{n+\ell} > x^n$ *für alle* $n, \ell \in \mathbb{N}$ *gilt.*

Übung 47. *Sei* $(\mathbb{K}, \leq)$ *ein angeordneter Körper; sowie* $n \in \mathbb{N}_{>0}$ *und* $x, y \in \mathbb{K}$ *mit* $x, y \geq 0_{\mathbb{K}}$. *Zeigen Sie mit Hilfe von Satz 6, dass*

$$(x+y)^n \geq x^n + y^n \qquad \textit{sowie} \qquad (x+y)^n \geq x^n + n(x^{n-1} \cdot y) \qquad \textit{gilt.}$$

Übung 48. *Sei* $(\mathbb{K}, \leq)$ *ein angeordneter Körper. Zeigen Sie die folgenden Aussagen:*

a) Ist $I \subseteq \mathbb{K}$ *ein Intervall, so ist* $-I = \{-x \mid x \in I\} \subseteq \mathbb{K}$ *ein Intervall.*

b) Ist $I \subseteq \mathbb{K}_{\times}$ *ein Intervall, so ist* $I^{-1} = \{x^{-1} \mid x \in I\} \subseteq \mathbb{K}$ *ein Intervall.*

Hinweis: Zeigen Sie zunächst, dass $I \geq 0_{\mathbb{K}}$ *oder* $I \leq 0_{\mathbb{K}}$ *gilt (ist* $I = \emptyset$, *so gilt natürlich beides); und benutzen Sie dann Teil a), sowie die bisher hergeleiteten Rechenregeln.*

4.2.4 Archimedisch angeordnete Körper

Definition 25. *Ein angeordneter Körper* $(\mathbb{K}, \leq)$ *heißt archimedisch angeordnet, wenn zu jedem* $x \in \mathbb{K}$*, ein* $n \in \mathbb{Z}$ *mit* $x < n_{\mathbb{K}}$ *existiert.*

Beispiel 23. *Der Körper der rationalen Zahlen* $(\mathbb{Q}, \leq)$ *ist archimedisch angeordnet; denn für* $\frac{p}{q} \in \mathbb{Q}$ *mit* $p \in \mathbb{Z}$ *und* $q \in \mathbb{N}_{>0}$ *ist* $(|p|+1) \cdot q - p \geq (|p|+1) - p \geq 1 > 0$*, also*

$$(|p|+1) - \frac{p}{q} = \frac{(|p|+1) \cdot q - p}{q} > 0 \qquad \Longrightarrow \qquad \frac{p}{q} < |p|+1.$$

Proposition 2. *Ist* $(\mathbb{K}, \leq)$ *archimedisch angeordnet, so gelten die folgenden Aussagen:*

1) Zu jedem $x \in \mathbb{K}$ *existiert ein* $n \in \mathbb{Z}$ *mit* $n_{\mathbb{K}} < x$*.*

2) Zu jedem $\mathbb{K} \ni \varepsilon > 0_{\mathbb{K}}$ *existiert ein* $n \in \mathbb{N}_{>0}$ *mit* $0_{\mathbb{K}} < \frac{1}{n}_{\mathbb{K}} < \varepsilon$*.*

3) Für alle $x, y \in \mathbb{K}$ *mit* $y > 0_{\mathbb{K}}$ *existiert ein* $n \in \mathbb{N}_{>0}$ *mit* $x < ny$*.*

4) Für alle $x, y \in \mathbb{K}$ *mit* $x < y$ *existiert ein* $\frac{p}{q} \in \mathbb{Q}$ *mit* $x < \frac{p}{q}_{\mathbb{K}} < y$*.*

Beweis. Wir bemerken zunächst, dass wegen Satz 7 (Einbettung von $\mathbb{Q}$ in $\mathbb{K}$) gilt:

$$p_{\mathbb{K}} < q_{\mathbb{K}} \quad \text{für} \quad p, q \in \mathbb{Z} \qquad \Longrightarrow \qquad p < q. \tag{88}$$

1) Sei $n \in \mathbb{Z}$ mit $-x < n_{\mathbb{K}}$. Dann gilt $(-n)_{\mathbb{K}} = -n_{\mathbb{K}} < x$ mit $(-n) \in \mathbb{Z}$.

2) Sei $n \in \mathbb{Z}$ mit $0_{\mathbb{K}} < \varepsilon < n_{\mathbb{K}}$ (also $n \geq 1$ wegen (88)). Dann gilt $0_{\mathbb{K}} < \frac{1}{n}_{\mathbb{K}} < \varepsilon$ wegen Lemma 17.11).

3) Per Voraussetzung existiert ein $n \in \mathbb{Z}$ mit $\frac{x}{y} = x \cdot y^{-1} < n_{\mathbb{K}}$. Wegen $y > 0_{\mathbb{K}}$ und Lemma 17.4), folgt $x = y \cdot \frac{x}{y} < n_{\mathbb{K}} \cdot y = ny$.

4) Wegen Teil 2), existiert ein $n \in \mathbb{N}_{>0}$ mit $0_{\mathbb{K}} < \frac{1}{n}_{\mathbb{K}} < y - x$. Multiplikation mit $n_{\mathbb{K}}$, liefert (wegen Lemma 17):

$$n_{\mathbb{K}} \cdot x + 1_{\mathbb{K}} < n_{\mathbb{K}} \cdot y. \tag{89}$$

Wir argumentieren wie folgt:

- Wegen Teil 1) existieren $p, q \in \mathbb{Z}$ mit

$$p_{\mathbb{K}} < n_{\mathbb{K}} \cdot y < q_{\mathbb{K}} \qquad \stackrel{(88)}{\Longrightarrow} \qquad p < q. \tag{90}$$

  Daher ist die Menge

$$M := \{m \in \mathbb{Z} \mid p \leq m < q \ \wedge \ m_{\mathbb{K}} \leq n_{\mathbb{K}} \cdot y\}$$

  nichtleer und endlich. Wegen Lemma 3.1), existiert somit $m := \max(M)$.

- Offensichtlich gilt $m + 1 \leq q$. Wegen der Maximalität von m gilt weiterhin (erster Schritt)

$$n_{\mathbb{K}} \cdot y < (m+1)_{\mathbb{K}} = m_{\mathbb{K}} + 1_{\mathbb{K}}. \tag{91}$$

  (In der Tat, gilt nämlich $(m+1)_{\mathbb{K}} \leq n_{\mathbb{K}} \cdot y$, so folgt mit der linken Seite von (90)

$$(m+1)_{\mathbb{K}} \leq n_{\mathbb{K}} \cdot y \stackrel{(90)}{<} q_{\mathbb{K}} \quad \text{also} \quad m+1 < q \qquad \Longrightarrow \qquad (m+1) \in M.)$$

Wir erhalten mit $m_{\mathbb{K}} \leq n_{\mathbb{K}} \cdot y$ (dritte Abschätzung)

$$\begin{aligned} n_{\mathbb{K}} \cdot x + 1_{\mathbb{K}} \stackrel{(89)}{<} n_{\mathbb{K}} \cdot y \stackrel{(91)}{<} m_{\mathbb{K}} + 1_{\mathbb{K}} \leq n_{\mathbb{K}} \cdot y + 1_{\mathbb{K}} \qquad &\stackrel{-1_{\mathbb{K}}}{\Longrightarrow} \qquad n_{\mathbb{K}} \cdot x < m_{\mathbb{K}} < n_{\mathbb{K}} \cdot y \\ &\Longrightarrow \qquad x < \frac{m}{n}_{\mathbb{K}} < y \end{aligned}$$

gilt, was die Behauptung zeigt. □

***Korollar 2.** *Sei $0 < x < 1_{\mathbb{K}}$. Für jedes $\varepsilon > 0_{\mathbb{K}}$, existiert ein $N_\varepsilon \in \mathbb{N}_{>0}$ mit*

$$0_{\mathbb{K}} < x^n < \varepsilon \qquad \forall\, n \geq N_\varepsilon.$$

Beweis. Zunächst gilt $0_{\mathbb{K}} < x^n < 1_{\mathbb{K}}$ für alle $n \in \mathbb{N}_{>0}$, wegen Übung 46.a). Weiterhin gilt $0_{\mathbb{K}} < 1_{\mathbb{K}} < x^{-1}$, wegen Lemma 17.11). Wir setzen $h := x^{-1} - 1_{\mathbb{K}} > 0_{\mathbb{K}}$; sodass dann also $x = (1_{\mathbb{K}} + h)^{-1}$ gilt. Satz 8 (zweiter Schritt) und Lemma 17.2) (letzter Schritt) liefern

$$(x^n)^{-1} \overset{(62)}{=} (x^{-1})^n = (1_{\mathbb{K}} + h)^n \geq 1_{\mathbb{K}} + nh > nh > 0 \qquad \forall\, n \in \mathbb{N}.$$

Also gilt $x^n \leq (nh)^{-1}$ wegen Lemma 17.11). Für $\varepsilon > 0_{\mathbb{K}}$ vorgegeben, sei nun $N_\varepsilon \in \mathbb{N}_{>0}$ mit $\varepsilon^{-1} < N_\varepsilon h$ (Proposition 2.3)); also (Lemma 17.11))

$$\varepsilon^{-1} < nh \quad \forall\, n \geq N_\varepsilon \qquad \Longrightarrow \qquad x^n \leq (nh)^{-1} < \varepsilon \quad \forall\, n \geq N_\varepsilon,$$

was die Behauptung zeigt. □

Bemerkung* 3. *Proposition 2.4) verschärft die Aussage in Korollar 8, sofern man Letztere auf die Aussage „$\forall\, x, y \in \mathbb{K}\colon \exists\, z \in \mathbb{K}\colon x < z < y$“ reduziert. Wegen Satz 7 könnte man evtl. dennoch auf den Gedanken kommen, dass jeder angeordnete Körper automatisch archimedisch ist – da er ja in einer gewissen Form „beliebig große“ natürliche Zahlen enthält. Das dies nicht der Fall ist, wollen wir durch folgendes Gegenbeispiel grob plausibel machen:*

Wir betrachten den Folgenraum $\mathbb{Z}^{\mathbb{N}}$ (vgl. Notation 4), und darin die Teilmenge

$$\mathsf{Z} := \big\{(x_n)_{n\in\mathbb{N}} \in \mathbb{Z}^{\mathbb{N}} \,\big|\, \exists\, \ell \in \mathbb{N}\colon x_n = 0 \quad \forall\, n \geq \ell\big\}$$

(„aller $\mathbb{N}$-Tupel $(x_n)_{n\in\mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$ mit nur endlich vielen Einträgen x_n ungleich 0“). Wir setzen

$$\mathsf{n} := (\delta_\ell)_{\ell\in\mathbb{N}} \quad \textit{für} \quad (\delta_\ell)_{\ell\in\mathbb{N}} \in \mathsf{Z} \quad \textit{gegeben durch} \quad \delta_0 = n \quad \textit{und} \quad \delta_\ell = 0 \quad \forall\, \ell \geq 1$$

für alle $n \in \mathbb{Z}$; und definieren weiterhin

$$\begin{aligned} \mathsf{grad}((x_n)_{n\in\mathbb{N}}) &:= \max(\{n \in \mathbb{N} \mid x_n \neq 0\}) && \forall\, (x_n)_{n\in\mathbb{N}} \in \mathsf{Z} \\ \mathsf{abs}((x_n)_{n\in\mathbb{N}}) &:= x_{\mathsf{grad}((x_n)_{n\in\mathbb{N}})} && \forall\, (x_n)_{n\in\mathbb{N}} \in \mathsf{Z}. \end{aligned}$$

Wir definieren die Abbildungen

$$\begin{aligned} +\colon \mathsf{Z} \times \mathsf{Z} &\to \mathsf{Z}, && ((x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}) \mapsto (x_n + y_n)_{n\in\mathbb{N}} \\ -\colon \mathsf{Z} &\to \mathsf{Z}, && (x_n)_{n\in\mathbb{N}} \mapsto (-x_n)_{n\in\mathbb{N}}; \end{aligned}$$

sowie $\boldsymbol{\cdot}\colon \mathsf{Z} \times \mathsf{Z} \to \mathsf{Z}$ mit

$$\big((x_m)_{m\in\mathbb{N}} \boldsymbol{\cdot} (y_n)_{n\in\mathbb{N}}\big)_\ell := \textstyle\sum_{m+n=\ell} x_m \cdot y_n \qquad \forall\, \ell \in \mathbb{N}.$$

Beispielsweise ist also

$$\begin{aligned} \big((x_m)_{m\in\mathbb{N}} \boldsymbol{\cdot} (y_n)_{n\in\mathbb{N}}\big)_0 &= x_0 \cdot y_0, \\ \big((x_m)_{m\in\mathbb{N}} \boldsymbol{\cdot} (y_n)_{n\in\mathbb{N}}\big)_1 &= x_0 \cdot y_1 + x_1 \cdot y_0, \\ \big((x_m)_{m\in\mathbb{N}} \boldsymbol{\cdot} (y_n)_{n\in\mathbb{N}}\big)_2 &= x_0 \cdot y_2 + x_1 \cdot y_1 + x_2 \cdot y_0, \\ \big((x_m)_{m\in\mathbb{N}} \boldsymbol{\cdot} (y_n)_{n\in\mathbb{N}}\big)_3 &= x_0 \cdot y_3 + x_1 \cdot y_2 + x_2 \cdot y_1 + x_3 \cdot y_0, \\ &\ \ \vdots \end{aligned}$$

Man prüft nach, dass $(\mathsf{Z},+,0)$ *und* $(\mathsf{Z},\cdot,1)$ *abelsche Monoide sind, mit*

$$\alpha\cdot(\beta+\gamma)=(\alpha\cdot\beta)+(\alpha\cdot\gamma)\qquad\forall\,\alpha,\beta,\gamma\in\mathsf{Z}.$$

Anschaulich kann man Z *einfach als eine (gewisse) Vergrößerung von* $\mathbb{Z}$ *verstehen, mit Einbettung (injektive Abbildung, die mit den Rechenoperationenverträglich ist)*

$$\mathbb{Z}\to\mathsf{Z},\qquad n\mapsto\mathsf{n}.$$

Analog zur Konstruktion von $\mathbb{Q}$ *in Beispiel 9, setzen wir* $\mathsf{Q}:=\mathsf{Z}\times\mathsf{Z}_{\neq 0}$ *mit* $\mathsf{Z}_{\neq 0}:=\mathsf{Z}\setminus\{0\}$*; und definieren*

$$(\alpha,\beta)\sim(\alpha',\beta')\qquad\overset{\text{def.}}{\Longleftrightarrow}\qquad\alpha\cdot\beta'=\beta\cdot\alpha'$$

für $\alpha,\alpha'\in\mathsf{Z}$ *und* $\beta,\beta'\in\mathsf{Z}_{\neq 0}$*, sowie*

$$\frac{\alpha}{\beta}:=[(\alpha,\beta)]_\sim\qquad\forall\,\alpha\in\mathsf{Z},\ \beta\in\mathsf{Z}_{\neq 0}.$$

Wir betrachten den Quotienten

$$\mathbf{Q}:=\mathsf{Q}/\sim\,=\Big\{\frac{\alpha}{\beta}\,\Big|\,\alpha\in\mathsf{Z},\ \beta\in\mathsf{Z}_{\neq 0}\Big\},$$

und setzen $\boldsymbol{n}:=\frac{\mathsf{n}}{\mathsf{n}}$ *für alle* $n\in\mathbb{Z}$ *(also insbesondere* $\mathbf{0}=\frac{0}{1}$ *und* $\mathbf{1}=\frac{1}{1}$*). Die Rechenoperationen auf* $\mathbf{Q}$ *sind definiert durch*

$$\frac{\alpha}{\beta}+\frac{\alpha'}{\beta'}:=\frac{\alpha\beta'+\beta\alpha'}{\beta\beta'},\qquad-\frac{\alpha}{\beta}:=\frac{-\alpha}{\beta},\qquad\frac{\alpha}{\beta}\cdot\frac{\alpha'}{\beta'}:=\frac{\alpha\alpha'}{\beta\beta'},\qquad\Big(\frac{\alpha''}{\beta}\Big)^{-1}:=\frac{\beta}{\alpha''}$$

für $\alpha,\alpha'\in\mathsf{Z}$*,* $\alpha''\in\mathsf{Z}_{\neq 0}$*, und* $\beta,\beta'\in\mathsf{Z}_{\neq 0}$*. Schließlich definieren wir*

$$\frac{\alpha}{\beta}\geq\mathbf{0}\qquad\overset{\text{def.}}{\Longleftrightarrow}\qquad\mathsf{abs}(\alpha\cdot\beta)\geq 0$$

für $\alpha\in\mathsf{Z}$ *und* $\beta\in\mathsf{Z}\setminus\{0\}$*; und dann*

$$x\geq y\quad\text{für}\quad x,y\in\mathbf{Q}\qquad\Longleftrightarrow\qquad x-y\geq\mathbf{0}.$$

Man prüft nach, dass diese Operationen auf Äquivalenzklassen wohldefiniert sind; und weiterhin, dass $(\mathbf{Q},\leq,+,\mathbf{0},\cdot,\mathbf{1})$ *ein angeordneter Körper ist, mit Körpereinbettung (mit* $\Psi_{\mathbf{Q}}(n)=\boldsymbol{n}$ *für alle* $n\in\mathbb{Z}$*)*

$$\Psi_{\mathbf{Q}}\colon\mathbb{Q}\to\mathbf{Q},\qquad\frac{p}{q}\mapsto\frac{\mathsf{p}}{\mathsf{q}}.\tag{92}$$

Für $\alpha\in\mathsf{Z}$ *mit* $\mathsf{grad}(\alpha)\geq 1$*, gilt dann* $\mathsf{abs}\big(\frac{\alpha}{1}-\boldsymbol{n}\big)=\mathsf{abs}(\alpha)\geq 1$ *für alle* $n\in\mathbb{Z}$*, also* $\boldsymbol{n}\leq\frac{\alpha}{1}$ *für alle* $n\in\mathbb{Z}$*. Daher ist* $(\mathbf{Q},\leq)$ *nicht archimedisch.*

Beweis der Transitivität von $\leq$*.* Wir wollen schließlich die Transitivität von $\leq$ behandeln, da diese am schwierigsten zu verifizieren scheint: Es ist unmittelbar einsichtig, dass

$$\begin{aligned}\mathsf{grad}(\alpha\cdot\beta)&=\mathsf{grad}(\alpha)+\mathsf{grad}(\beta)\\\mathsf{abs}(\alpha\cdot\beta)&=\mathsf{abs}(\alpha)\cdot\mathsf{abs}(\beta)\end{aligned}\tag{93}$$

für alle $\alpha, \beta \in \mathsf{Z}$ gilt. Seien nun $\alpha, \alpha' \in \mathsf{Z}$ sowie $\beta, \beta' \in \mathsf{Z}_{\neq 0}$ gegeben mit $\frac{\alpha}{\beta}, \frac{\alpha'}{\beta'} \geq \mathbf{0}$; also $\mathsf{abs}(\alpha) \cdot \mathsf{abs}(\beta) \geq 0$ und $\mathsf{abs}(\alpha') \cdot \mathsf{abs}(\beta') \geq 0$. Dann gilt $\frac{\alpha}{\beta} + \frac{\alpha'}{\beta'} \geq \mathbf{0}$, wegen

$$\begin{aligned}(\mathsf{abs}(\alpha) \cdot \mathsf{abs}(\beta') \; + \; \mathsf{abs}(\beta) \cdot \mathsf{abs}(\alpha')) \cdot \mathsf{abs}(\beta) \cdot \mathsf{abs}(\beta') \\ = \mathsf{abs}(\alpha) \cdot \mathsf{abs}(\beta) \cdot \mathsf{abs}(\beta')^2 + \mathsf{abs}(\alpha') \cdot \mathsf{abs}(\beta') \cdot \mathsf{abs}(\beta)^2 \geq 0.\end{aligned}$$

Sind nun $\alpha, \alpha', \alpha'' \in \mathsf{Z}$ sowie $\beta, \beta', \beta'' \in \mathsf{Z}_{\neq 0}$ gegeben, mit $\frac{\alpha}{\beta} \leq \frac{\alpha'}{\beta'}$ und $\frac{\alpha'}{\beta'} \leq \frac{\alpha''}{\beta''}$, dann gilt $\frac{\alpha'}{\beta'} - \frac{\alpha}{\beta} \geq \mathbf{0}$ sowie $\frac{\alpha''}{\beta''} - \frac{\alpha'}{\beta'} \geq \mathbf{0}$; also

$$\frac{\alpha''}{\beta''} - \frac{\alpha}{\beta} = \Big(\frac{\alpha''}{\beta''} - \frac{\alpha'}{\beta'}\Big) + \Big(\frac{\alpha'}{\beta'} - \frac{\alpha}{\beta}\Big) \geq \mathbf{0}$$

nach dem soeben Gezeigten. □

4.2.5 Das Vollständigkeitsaxiom

Die letzte Körpereigenschaft, die wir besprechen wollen, ist die der Vollständigkeit.

Notation 16. *Gegeben ein angeordneter Körper* $(\mathbb{K}, \leq)$ *und eine Teilmenge* $X \subseteq \mathbb{K}$*, so schreiben wir im Folgenden vereinfachen* $\sup(X)$ *anstelle* $\sup_{\mathbb{K}}(X)$*, sowie* $\inf(X)$ *anstelle* $\inf_{\mathbb{K}}(X)$*.*

Definition 26. *Ein angeordneter Körper* $(\mathbb{K}, \leq)$ *heißt vollständig (oder vollständig angeordnet), wenn jede nichtleere nach oben beschränkte Teilmenge von* $\mathbb{K}$ *eine kleinste obere Schranke hat. Ist also* $\emptyset \neq X \subseteq \mathbb{K}$ *gegeben mit* $X \leq o$ *für ein* $o \in \mathbb{K}$*, so existiert* $\sup(X) \in \mathbb{K}$*.*

Korollar 10. *Ein angeordneter Körper* $(\mathbb{K}, \leq)$ *ist vollständig genau dann, wenn jede nichtleere nach unten beschränkte Teilmenge von* $\mathbb{K}$ *eine größte untere Schranke hat (für* $\emptyset \neq X \subseteq \mathbb{K}$ *mit* $u \leq X$ *für ein* $u \in \mathbb{K}$*, existiert* $\inf(X) \in \mathbb{K}$*).*

Beweis. Die Aussage folgt aus Proposition 1:

- Sei $(\mathbb{K}, \leq)$ vollständig angeordnet; sowie $\emptyset \neq X \subseteq \mathbb{K}$ und $u \in \mathbb{K}$ mit $u \leq X$. Dann gilt $-X \leq -u$. Daher existiert $\sup(-X)$ wegen der Vollständigkeitsannahme. Proposition 1.5) zeigt $\inf(X) = -\sup(-X)$, also insbesondere die Existenz von $\inf(X)$.
- Es gelte, dass jede nichtleere nach unten beschränkte Teilmenge von $\mathbb{K}$ eine größte untere Schranke hat. Seien nun $\emptyset \neq X \subseteq \mathbb{K}$ und $o \in \mathbb{K}$ mit $X \leq o$. Dann gilt $-o \leq -X$; also existiert $\inf(-X)$ per Annahme. Proposition 1.6) zeigt $\sup(X) = -\inf(-X)$, also insbesondere die Existenz von $\sup(X)$. □

Notation 17. *Sei* $(\mathbb{K}, \leq)$ *ein vollständig angeordneter Körper, und* $\emptyset \neq X \subseteq \mathbb{K}$ *eine nichtleere Teilmenge von* $\mathbb{K}$*.*

- *Wir schreiben* $\sup(X) < \infty$*, wenn* $\sup(X)$ *existiert (also* X *nach oben beschränkt ist); sowie* $\sup(X) = \infty$*, falls* $\sup(X)$ *nicht existiert (also* X *nach oben unbeschränkt ist).*
- *Wir schreiben* $-\infty < \inf(X)$*, wenn* $\inf(X)$ *existiert (also* X *nach unten beschränkt ist); sowie* $\inf(X) = -\infty$*, falls* $\inf(X)$ *nicht existiert (also* X *nach unten unbeschränkt ist).*

Bemerkung 26 (Intervalle in vollständig angeordneten Körpern)**.** *Ist* $(\mathbb{K}, \leq)$ *vollständig angeordnet, so sind die nichtleeren Intervalle in* $\mathbb{K}$*, genau die Teilmengen der Form* (41)*; also mit* $x, y \in \mathbb{K}$*:*

$$\begin{array}{lrll}
(x \leq z) & [x,z] & := \{y \in \mathbb{K} \mid x \leq y \leq z\} & \textit{(abgeschlossenes Intervall)} \\
(x < z) & (x,z) & := \{y \in \mathbb{K} \mid x < y < z\} & \textit{(offenes Intervall)} \\
(x < z) & (x,z] & := \{y \in \mathbb{K} \mid x < y \leq z\} & \textit{(linksoffenes Intervall)} \\
(x < z) & [x,z) & := \{y \in \mathbb{K} \mid x \leq y < z\} & \textit{(rechtsoffenes Intervall)} \\
 & (x,\infty) & := \{y \in \mathbb{K} \mid x < y\} & \\
 & [x,\infty) & := \{y \in \mathbb{K} \mid x \leq y\} & \\
 & (-\infty,z) & := \{y \in \mathbb{K} \mid y < z\} & \\
 & (-\infty,z] & := \{y \in \mathbb{K} \mid y \leq z\} & \\
 & (-\infty,\infty) & = \mathbb{K} &
\end{array} \tag{94}$$

Sei hierfür $\emptyset \neq D \subseteq \mathbb{K}$ *ein Intervall:*

- *Sei* D *weder nach oben noch nach unten beschränkt. Dann gilt* $D = (-\infty, \infty) = \mathbb{K}$*:*

  Beweis. Für $c \in \mathbb{K}$ vorgegeben, gilt weder $c \leq D$ noch $D \leq c$. Da $\mathbb{K}$ angeordnet ist, existieren daher gewisse $a, b \in D$ mit $a < c < b$. Aus der Intervalleigenschaft von D folgt $c \in D$. □

- *Sei* D *nach oben aber nicht nach unten beschränkt. Dann existiert* $z = \sup(D)$*; und es sind die folgenden beiden Fälle möglich:*

  - $z \in D \quad \Rightarrow \quad D = (-\infty, z]$
    (Für alle $c \in D$ *gilt* $c \leq z$*, und somit* $D \subseteq (-\infty, z]$*. Für* $c \in (-\infty, z]$ *vorgegeben, existiert (wegen* $c \not\leq D$*) ein* $a \in D$ *mit* $a < c$*. Es folgt* $a < c \leq z$ *mit* $a, z \in D$*, also* $c \in D$*.)*
  - $z \notin D \quad \Rightarrow \quad D = (-\infty, z)$.
    (Für alle $c \in D$ *gilt* $c < z$*, und somit* $D \subseteq (-\infty, z)$*. Für* $c \in (-\infty, z)$ *vorgegeben, existiert (wegen* $c \not\leq D$*) ein* $a \in D$ *mit* $a < c$*. Es folgt* $a < c < z$ *mit* $a, z \in D$*, also* $c \in D$*.)*

  Ebenso argumentiert man für den Fall, dass D *nach unten aber nicht nach oben beschränkt ist; und erhält die beiden Fälle* $D = [x, \infty)$ *sowie* $D = (x, \infty)$ *mit* $x := \inf(D)$*.*

- *Eine analoge Fallunterscheidung liefert die ersten vier Intervalltypen in* (94)*; und zwar für den Fall, dass* D *sowohl nach oben als auch nach unten beschränkt ist.*

Übung 49. *Seien* $(\mathbb{K}, \leq)$ *ein vollständig angeordneter Körper. Seien* $I, J \neq \emptyset$*, und* $(s_{ij})_{(i,j) \in I \times J}$ *eine Familie von Elementen aus* $\mathbb{K}$*, mit* $s := \sup(\{s_{ij} \mid (i,j) \in I \times J\}) < \infty$*. Zeigen Sie:*

a) Für alle $i \in I$*, gilt* $s_i := \sup(\{s_{ij} \mid j \in J\}) \leq s$*.*

b) Für alle $(i', j') \in I \times J$ *gilt* $s_{i'j'} \leq \sup(\{s_i \mid i \in I\})$*.*

c) Es gilt $s = \sup(\{s_i \mid i \in I\})$*.*

Lemma 21. *Ist* $(\mathbb{K}, \leq)$ *vollständig angeordnet, so ist* $(\mathbb{K}, \leq)$ *archimedisch angeordnet.*

Beweis. Sei $(\mathbb{K}, \leq)$ vollständig angeordnet. Ist $(\mathbb{K}, \leq)$ nicht archimedisch angeordnet, so existiert ein $o \in \mathbb{K}$ mit $Z := \{n_{\mathbb{K}} \mid n \in \mathbb{Z}\} \leq o$. Daher existiert $\sup(Z) \in \mathbb{K}$. Nun gilt

$$\begin{array}{llcrl}
n_{\mathbb{K}} \leq \sup(Z) & \forall\, n \in \mathbb{Z} & \Longleftrightarrow & n_{\mathbb{K}} + 1_{\mathbb{K}} = (n+1)_{\mathbb{K}} \leq \sup(Z) & \forall\, n \in \mathbb{Z} \\
 & & \Longleftrightarrow & n_{\mathbb{K}} \leq \sup(Z) - 1_{\mathbb{K}} & \forall\, n \in \mathbb{Z} \\
 & & \Longleftrightarrow & Z \leq \sup(Z) - 1_{\mathbb{K}}. &
\end{array}$$

Wir erhalten $Z \leq \sup(Z) - 1_{\mathbb{K}} < \sup(Z)$ (wegen $-1_{\mathbb{K}} < 0_{\mathbb{K}}$), also eine Widerspruch zur der Definition von $\sup(Z)$. □

Bemerkung* 4. *Insbesondere zeigt Lemma 21, dass der Körper* **Q** *aus Bemerkung 3 nicht vollständig angeordnet ist (sein kann).*

Wir kommen nun zum Wurzelziehen (siehe auch Beispiel 61 in Abschnitt 7.2.2).

Satz 9. *Sei* $(\mathbb{K}, \leq)$ *ein vollständig angeordneter Körper. Gegeben* $k \in \mathbb{N}_{>0}$ *und* $x \geq 0_{\mathbb{K}}$*, so existiert genau ein* $\mathbb{K} \ni z \geq 0_{\mathbb{K}}$ *mit* $z^k = x$*.*

$*$*Beweis.* Die Eindeutigkeitsaussage ist klar wegen Übung 45.a) und $(0_{\mathbb{K}})^k = 0_{\mathbb{K}}$. Wegen $(0_{\mathbb{K}})^k = 0_{\mathbb{K}}$, können wir weiterhin $x > 0_{\mathbb{K}}$ annehmen.

Für die Existenz, setzen wir

$$C[y] := k \cdot \max\left(\left\{\tbinom{k}{\ell}\, y^{\ell} \;\middle|\; 0 \leq \ell \leq k-1\right\}\right) \geq 1_{\mathbb{K}} \qquad \forall\, y \geq 0_{\mathbb{K}}. \tag{95}$$

Für $0_{\mathbb{K}} < \delta \leq 1_{\mathbb{K}}$ und $y \geq 0_{\mathbb{K}}$, liefert der Binomische Lehrsatz (Satz 6):

$$\begin{aligned}(y+\delta)^k &= y^k + \left(\textstyle\sum_{\ell=0}^{k-1} \binom{k}{\ell}\, y^{\ell} \cdot \delta^{k-\ell-1}\right) \cdot \delta \leq y^k + C[y] \cdot \delta \\ (y-\delta)^k &= y^k - \left(\textstyle\sum_{\ell=0}^{k-1} \binom{k}{\ell}\, (-1)^{k-\ell} \cdot y^{\ell} \cdot \delta^{k-\ell-1}\right) \cdot \delta \geq y^k - C[y] \cdot \delta.\end{aligned} \tag{96}$$

Sei nun $x \geq 0_{\mathbb{K}}$, und

$$X := \{0_{\mathbb{K}} \leq y \in \mathbb{K} \mid y^k < x\}.$$

Wegen $0_{\mathbb{K}} \in X$, gilt $X \neq \emptyset$. Weiterhin ist $X \leq \max(1_{\mathbb{K}}, x)$ nach oben beschränkt; denn für $y \in X$ erhalten wir:

- Ist $0_{\mathbb{K}} \leq y \leq 1_{\mathbb{K}}$, so gilt $y^k \leq 1_{\mathbb{K}}$ wegen Übung 46.
- Ist $1_{\mathbb{K}} < y$, so gilt $y < y^k < x$ wegen Übung 46.

Wegen der Vollständigkeit von $\mathbb{K}$, existiert dann $z := \sup(X) \geq 0_{\mathbb{K}}$. Wir argumentieren nun wie folgt:

- Es gelte $z^k < x$. Für $\varepsilon := \min\left(1_{\mathbb{K}}, x - z^k\right)$ und $\delta := \frac{1}{2}_{\mathbb{K}} \cdot \frac{\varepsilon}{C[z]} > 0_{\mathbb{K}}$, folgt dann wegen (96)

  $$(z+\delta)^k \leq z^k + \frac{1}{2}_{\mathbb{K}} \cdot \varepsilon < z^k + \varepsilon \leq x \qquad \Longrightarrow \qquad (z+\delta) \in X,$$

  was $X \leq z$ widerspricht. Daher gilt $x \leq z^k$.
- Für $0_{\mathbb{K}} < \varepsilon < 1_{\mathbb{K}}$ vorgegeben, sei $\delta := \frac{\varepsilon}{C[z]} > 0_{\mathbb{K}}$. Nach Lemma 20, existiert ein $y \in X$ mit $y > z - \delta$. Es gilt dann wegen (96) und Übung 45.a)

  $$z^k - \varepsilon = z^k - C[z] \cdot \delta \leq (z-\delta)^k < y^k < x \qquad \Longrightarrow \qquad z^k \leq x + \varepsilon.$$

  Da dies für alle $\varepsilon > 0_{\mathbb{K}}$ gilt, folgt $z^k \leq x$ aus Korollar 9.1).

Dies zeigt $z^k = x$, also die Behauptung. □

Bemerkung 27. *Für die Eindeutigkeitsaussage in Satz 9, ist die Bedingung* $z \geq 0_{\mathbb{K}}$ *unabdingbar wenn* $x > 0_{\mathbb{K}}$ *gilt und* k *gerade ist. Gilt nämlich* $z^{2n} = x$ *für* $x, z > 0_{\mathbb{K}}$ *und* $n \in \mathbb{N}_{>0}$*, so gilt ebenfalls*

$$(-z)^{2n} = ((-z)^2)^n = (z^2)^n = z^{2n} = x$$

mit $-z \neq z$ *wegen* $z > 0$ *(vgl. Übung 41).*

Definition 27 (k-te Wurzel)**.** *Sei* $(\mathbb{K}, \leq)$ *vollständig angeordnet,* $k \in \mathbb{N}_{>0}$ *und* $0_{\mathbb{K}} \leq x \in \mathbb{K}$*. Die* k*-te Wurzel* $\sqrt[k]{x}$ *von* x*, ist definiert als das (nach Satz 9) eindeutige Elemente* $0_{\mathbb{K}} \leq z \in \mathbb{K}$ *mit* $z^k = x$*.*

Übung 50. *Sei* $(\mathbb{K}, \leq)$ *ein vollständig angeordneter Körper, und* $x \in \mathbb{K}$*. Zeigen Sie:*

a) Ist $0_\mathbb{K} < x < 1_\mathbb{K}$*, so gilt* $\sqrt[n+\ell]{x} > \sqrt[n]{x}$ *für alle* $n, \ell \in \mathbb{N}_{>0}$*.*

b) Ist $1_\mathbb{K} < x$*, so gilt* $\sqrt[n]{x} > \sqrt[n+\ell]{x}$ *für alle* $n, \ell \in \mathbb{N}_{>0}$*.*

Hinweis: Beweis durch Kontraposition, mit Hilfe der dritten Identität in (62) *sowie Übung 46.a).*

Definition 28 (Rationale Potenzen)**.** *Sei* $(\mathbb{K}, \leq)$ *ein vollständig angeordneter Körper. Für* $\frac{p}{q} \in \mathbb{Q}$ *mit* $p \in \mathbb{Z}$ *und* $q \in \mathbb{N}_{>0}$*, definieren wir*

$$x^{\frac{p}{q}} := (\sqrt[q]{x})^p = \sqrt[q]{x^p} \qquad \forall\, \mathbb{K} \ni x \geq 0_\mathbb{K}. \tag{97}$$

Für $\alpha \in \mathbb{Q}$ *beliebig, wählen wir* $p \in \mathbb{Z}$ *und* $q \in \mathbb{N}_{>0}$ *mit* $\alpha = \frac{p}{q}$*, und setzen* $x^\alpha := x^{\frac{p}{q}}$*.*

Das folgende Lemma klärt die Wohldefiniertheit der rationalen Potenz.

Lemma 22. *Sei* $(\mathbb{K}, \leq)$ *ein vollständig angeordneter Körper, und* $0_\mathbb{K} \leq x \in \mathbb{K}$*.*

a) Es gilt $(\sqrt[q]{x})^p = \sqrt[q]{x^p}$ *für alle* $p \in \mathbb{Z}$ *und* $q \in \mathbb{N}_{>0}$*.*

b) Ist $\frac{p}{q} = \frac{p'}{q'}$ *für* $p, p' \in \mathbb{Z}$ *und* $q, q' \in \mathbb{N}_{>0}$*, so gilt* $\sqrt[q]{x^p} = \sqrt[q']{x^{p'}}$*.*

Beweis. Dies folgt aus die Eindeutigkeitsaussage in Satz 9, sowie der dritte Identität in (62). □

Übung 51. *Beweisen Sie Lemma 22.*

Übung 52. *Sei* $(\mathbb{K}, \leq)$ *ein vollständig angeordneter Körper, und seien* $0_\mathbb{K} < x, y \in \mathbb{K}$*. Zeigen Sie:*

a) Es gilt $(x \cdot y)^\alpha = x^\alpha \cdot y^\alpha$ *für alle* $\alpha \in \mathbb{Q}$*.*

   Hinweis: Zeigen Sie die Aussage zunächst per Induktion für alle $\alpha \in \mathbb{N}$*; und wenden Sie dann* (62) *in Lemma 12 an, um die Aussage für alle* $\alpha \in \mathbb{Z}$ *zu zeigen. Folgern Sie hieraus, dass die Aussage für alle* $\alpha \in \{\frac{1}{n} \mid n \in \mathbb{N}_{>0}\}$ *gilt (Eindeutigkeit in Satz 9). Folgern Sie schließlich die Behauptung.*

b) Ist $0_\mathbb{K} < x < y$*, so gilt* $\sqrt[n]{x} < \sqrt[n]{y}$ *für alle* $n \in \mathbb{N}_{>0}$*.*

c) Ist $0_\mathbb{K} < x < y$*, so gilt* $x^\alpha < y^\alpha$ *für alle* $0 < \alpha \in \mathbb{Q}$*.*

Sie dürfen die Aussagen in Übung 45 frei benutzen.

Wir schließen diesen Abschnitt mit dem folgenden wichtigen Korollar zu Satz 9.

Korollar 11. *Der archimedisch angeordnete Körper* $(\mathbb{Q}, \leq)$ *ist nicht vollständig.*

Beweis. Wir nehmen an, dass $\mathbb{Q}$ vollständig angeordnet ist. Nach Satz 9, existiert dann ein $0 < x \in \mathbb{Q}$ mit $x^2 = 2$. Wir schreiben $x = \frac{p}{q}$ mit $p, q \in \mathbb{N}_{>0}$ teilerfremd; also

$$\nexists\, (\mathbb{N} \ni n \geq 2 \;\wedge\; a, b \in \mathbb{N}_{>0}) \colon \quad (\, p = na \quad \wedge \quad q = nb\,). \tag{98}$$

Einschub: Um eine derartige Darstellung von x zu erhalten, fixiere man $\tilde{p}, \tilde{q} \in \mathbb{N}_{>0}$ mit $x = \frac{\tilde{p}}{\tilde{q}}$, und betrachte die endliche Menge

$$N := \big\{(\mathbb{N} \ni n \geq 2 \,\big|\, \exists\, p[n], q[n] \in \mathbb{N}_{>0} \colon (\, \tilde{p} = n \cdot p[n] \;\wedge\; \tilde{q} = n \cdot q[n]\,)\big\}.$$

- Gilt $N = \emptyset$, so ist $\frac{\tilde{p}}{\tilde{q}}$ die gesuchte Darstellung.

- Gilt $N \neq \emptyset$, so existiert $n := \max(N)$. Die gesuchte Darstellung ist dann gegeben durch $\frac{p[n]}{q[n]}$:

  Sind nämlich $p[n]$ und $q[n]$ nicht teilerfremd, so existieren $\mathbb{N} \ni m \geq 2$ sowie $a, b \in \mathbb{N}_{>0}$ mit $p[n] = m \cdot a$ und $q[n] = m \cdot b$; und wir erhalten $\tilde{p} = (nm) \cdot a$ und $\tilde{q} = (nm) \cdot b$ mit $nm > n$, was der Maximalität von n widerspricht.

Wir erhalten

$$2 = x^2 = \tfrac{p^2}{q^2} \qquad \Longrightarrow \qquad p^2 = 2q^2 \qquad \Longleftrightarrow \qquad q^2 = \tfrac{q^2}{2}.$$

Somit ist p^2 gerade; und daher ist auch p gerade (Beweis durch Kontraposition). Folglich gilt $p = 2a$ für ein $a \in \mathbb{N}_{>0}$, also $p^2 = 4a^2$. Es folgt $q^2 = 4a^2/2 = 2a^2$. Daher ist q^2 gerade; also ist $q = 2b$ (für ein $b \in \mathbb{N}_{>0}$) ebenfalls gerade. Daher gilt $p = 2a$ und $q = 2b$, was (98) widerspricht. □

Bemerkung* 5 (Approximation von Wurzeln in archimedisch angeordneten Körpern)**.** *Wir wollen an dieser Stelle noch zwei Eigenschaften diskutieren, die wir im nächsten Abschnitt zur Veranschaulichung der Konstruktion der reellen Zahlen (vermöge Dedekindscher Schnitte) benötigen werden. Sei hierfür* $(\mathbb{K}, \leq)$ *ein angeordneter Körper,* $0_{\mathbb{K}} \leq x \in \mathbb{K}$*, und*

$$X := \{0_{\mathbb{K}} \leq y \in \mathbb{K} \mid y^k < x\} \tag{99}$$

wie im Beweis von Satz 9. Es gilt dann Folgendes:

1) *Für jedes* $y \in X$*, existiert ein* $z \in X$ *mit* $y < z$*.*
2) *Ist* $(\mathbb{K}, \leq)$ *archimedisch angeordnet, so existiert zu jedem* $0 \leq \tau < x$ *ein* $y \in X$ *mit* $\tau < y^k < x$*.*
3) *Anschaulich bedeutet die Eigenschaft in* 2)*, dass – obwohl die* k*-te Wurzel von* x *in* $\mathbb{K}$ *möglicherweise nicht existiert – die Teilmenge* X *besagte Wurzel dennoch darstellt/beschreibt bzw. „von unten approximiert“.*

   Auf die Annahme, dass $(\mathbb{K}, \leq)$ *archimedisch angeordnet ist, kann hierbei ohne Weiteres nicht verzichtet werden.*

Beweis der Aussagen:

1) Sei $y \geq 0_{\mathbb{K}}$ mit $y^k < x$; und sei $\varepsilon := \min\big(1, x - y^k\big)$, sowie $\delta := \frac{1}{2}_{\mathbb{K}} \cdot \frac{\varepsilon}{C[y]} > 0$ mit $C[y]$ wie in (95). Dann gilt

$$(y + \delta)^k \overset{(96)}{\leq} y^k + \frac{1}{2}_{\mathbb{K}} \cdot \varepsilon < x \qquad \Longrightarrow \qquad (y + \delta) \in X,$$

   was die Behauptung zeigt.

2) Wie im Beweis von Satz 9 folgt $X \leq o$ für $o := \max(1_{\mathbb{K}}, x)$. Für $C[o]$ wie in (95), folgt aus dem Binomischen Lehrsatz (vgl. (96))

$$(y + \delta)^k - y^k = \big(\textstyle\sum_{\ell=0}^{k-1} \binom{k}{\ell}\, y^\ell \cdot \delta^{k-\ell}\big) \cdot \delta \leq C[o] \cdot \delta \qquad \forall\, 0_{\mathbb{K}} \leq y \leq o,\ \delta > 0_{\mathbb{K}}. \tag{100}$$

   Weiterhin gilt $x \leq o^k$ wegen Übung 46. Wir nehmen nun an, dass die Bahauptung falsch ist. Dann existiert ein $0_{\mathbb{K}} \leq \tau < x$ mit $0_{\mathbb{K}} \leq y^k \leq \tau$ für alle $y \in X$. Nun existiert wegen Teil 1) und $0_{\mathbb{K}} \in X$, ein $0_{\mathbb{K}} < u \in X$. Insbesondere gilt also $0_{\mathbb{K}} < u^k \leq \tau$. Wegen Proposition 2.2), existiert ein $n \in \mathbb{N}$ mit $0_{\mathbb{K}} < \frac{1}{n}_{\mathbb{K}} < \min\big(u, \frac{x-\tau}{C[o]}\big)$; also

$$\big(\tfrac{1}{n}_{\mathbb{K}}\big)^k < u^k \leq \tau \qquad \text{sowie} \qquad \big(y + \tfrac{1}{n}_{\mathbb{K}}\big)^k - y^k \overset{(100)}{<} x - \tau \qquad \forall\, y \in X. \tag{101}$$

Wegen Proposition 2.3), existiert ein $m \in \mathbb{N}_{>0}$ mit $\frac{m}{n}_{\mathbb{K}} \geq o$, also

$$(\tfrac{m}{n}_{\mathbb{K}})^k \geq o^k \geq x > \tau.$$

Wegen Satz 7 gilt $0_{\mathbb{K}} < \frac{1}{n}_{\mathbb{K}} < \frac{2}{n}_{\mathbb{K}} < \ldots < \frac{m}{n}_{\mathbb{K}}$. Somit ist dann

$$\ell := \max\left(\left\{1 \leq p \leq m-1 \,\middle|\, \left(\tfrac{p}{n}_{\mathbb{K}}\right)^k \leq \tau\right\}\right)$$

definiert, und es gilt $\left(\frac{\ell}{n}_{\mathbb{K}}\right)^k \leq \tau < \left(\frac{\ell+1}{n}_{\mathbb{K}}\right)^k$. Aus der Wahl von τ folgt dann

$$\left(\tfrac{\ell}{n}_{\mathbb{K}}\right)^k \leq \tau < x \leq \left(\tfrac{\ell+1}{n}_{\mathbb{K}}\right)^k \qquad \Longrightarrow \qquad \left(\tfrac{\ell+1}{n}_{\mathbb{K}}\right)^k - \left(\tfrac{\ell}{n}_{\mathbb{K}}\right)^k \geq x - \tau.$$

Aus der rechten Seite von (101) folgt nun aber

$$\left(\tfrac{\ell+1}{n}_{\mathbb{K}}\right)^k - \left(\tfrac{\ell}{n}_{\mathbb{K}}\right)^k = \left(\tfrac{\ell}{n}_{\mathbb{K}} + \tfrac{1}{n}_{\mathbb{K}}\right)^k - \left(\tfrac{\ell}{n}_{\mathbb{K}}\right)^k < x - \tau,$$

was der vorherigen Abschätzung widerspricht.

3) Um die letzte Aussage in 3) einzusehen, betrachten wir den nicht archimedisch angeordneten Körper $(\mathbf{Q}, \leq)$ aus Bemerkung 3; und darin die Grad 1-Elemente $\boldsymbol{z}[n] = (0, n, 0, 0, \ldots)$ für $n \in \mathbb{N}_{>0}$ – also

$$\boldsymbol{x}[n] = (\delta_\ell)_{\ell \in \mathbb{N}} \quad \text{für} \quad (\delta_\ell)_{\ell \in \mathbb{N}} \in \mathsf{Z} \quad \text{gegeben durch} \quad \delta_1 = n \quad \text{und} \quad \delta_\ell = 0 \quad \forall\, \ell \in \mathbb{N} \setminus \{1\}.$$

Für $n \in \mathbb{N}_{>0}$, sei $\boldsymbol{X}[n] := \left\{\mathbf{0} \leq \boldsymbol{y} \in \mathbf{Q} \,\middle|\, \boldsymbol{y}^2 \leq \boldsymbol{x}[n]\right\}$, wie in (99) mit $k = 2$ definiert. Es gilt dann

$$\boldsymbol{X}[n] = \boldsymbol{W}[n] := \left\{\tfrac{\mathsf{p}}{\mathsf{q}} = \left(\tfrac{p}{q}, 0, 0, \ldots\right) \,\middle|\, p \in \mathbb{N} \ \wedge \ q \in \mathbb{N}_{>0}\right\} \qquad \forall\, n \in \mathbb{N}_{>0},$$

was man wie folgt einsieht:

- Für jedes $\boldsymbol{y} \in \boldsymbol{W}[n]$ gilt

$$\begin{aligned} \mathsf{grad}(\boldsymbol{y}) = 0 = \mathsf{grad}(\boldsymbol{y}^2) \quad &\Longrightarrow \quad \mathsf{abs}(\boldsymbol{x}[n] - \boldsymbol{y}^2) = \mathsf{abs}(\boldsymbol{x}[n]) \geq 0 \\ &\Longrightarrow \quad \boldsymbol{y}^2 \leq \boldsymbol{x}[n] \\ &\Longrightarrow \quad \boldsymbol{y} \in \boldsymbol{X}[n]. \end{aligned}$$

- Gilt umgekehrt $\boldsymbol{y}^2 \leq \boldsymbol{x}[n]$ für ein $\mathbf{0} \leq \boldsymbol{y} \in \mathbf{Q}$, so folgt $\mathsf{grad}(\boldsymbol{y}^2) \leq \mathsf{grad}(\boldsymbol{x}[n]) = 1$; also $\mathsf{grad}(\boldsymbol{y}) = 0$, wegen (93). Daher gilt $\boldsymbol{y} \in \boldsymbol{W}[n]$.

Offensichtlich besitzt $\boldsymbol{X}[n]$ (für $n \in \mathbb{N}_{>0}$) die Eigenschaft in 1). Allerdings besitzt zum Beispiel $\boldsymbol{X} := \boldsymbol{X}[2]$ (es ist also $\boldsymbol{x} = \boldsymbol{x}[2]$) nicht die Eigenschaft in 2), denn für $\boldsymbol{\tau} := \boldsymbol{x}[1]$ gilt $\boldsymbol{X} \leq \boldsymbol{\tau} < \boldsymbol{x}$. (Selbiges gilt auch für $X = \boldsymbol{X}[n+1]$ mit $n \in \mathbb{N}_{>0}$). □

4.3 Die Reellen Zahlen – vgl. Kaptiel 1 in [11]

Korollar 11 zeigt, dass der Körper der rationalen Zahlen eine wichtige Eigenschaft nicht besitzt, nämlich die der Vollständigkeit. Beispielsweise garantiert ja diese Eigenschaft gemäß Satz 9, dass positive Wurzeln positiver Körperelemente existieren. Anschaulich kann man sagen, dass die rationalen Zahlen in einer gewissen Form Lücken aufweisen, die nun gerade durch den Übergang zu den reellen Zahlen geschlossen werden. Hauptgegenstand dieses Abschnittes ist der Beweis des folgenden Satzes:

Satz 10 (Charakterisierung der reellen Zahlen)**.**

1) Es existiert ein vollständig angeordneter Körper $(\mathbb{R}, \leq_{\mathbb{R}})$.

2) Ist $(\mathbb{K}, \leq)$ *ein vollständig angeordneter Körper, so existiert ein Körperisomorphismus* $\Phi\colon \mathbb{R} \to \mathbb{K}$.

Notation 18. *Obgleich wir um der Klarheit willen, in diesem Abschnitt den Index „$_\mathbb{R}$“ an dem Ordnungssymbol $\leq$, sowie an den noch zu definierenden Rechenoperationen und neutralen Elementen von $\mathbb{R}$, konsequent mitnotieren (also $\mathbb{R}, \leq_\mathbb{R}, +_\mathbb{R}, 0_\mathbb{R}, 1_\mathbb{R}, \cdot_\mathbb{R}$ schreiben), werden wir dies in den in den Folgeabschnitten und Kapiteln nicht mehr tun. Wir betrachten dann $\mathbb{Q}$ einfach als Teilmenge von $\mathbb{R}$, womit alle Konventionen und Notationen gelten, die Sie bereits aus der Schule kennen (also $\mathbb{R}, \leq, +, 0, 1, \cdot$). Die Elemente von $\mathbb{R} \setminus \mathbb{Q}$ werden dann als* **irrationale Zahlen** *bezeichnet.*

Definition 29. *Eine Teilmenge $\alpha \subseteq \mathbb{Q}$ heißt Dedekindscher Schnitt, wenn sie die folgenden Bedingungen erfüllt:*

a) $\alpha \neq \emptyset$ *(α ist nicht leer)*

b) $\alpha \leq o$ *für ein* $o \in \mathbb{Q}$ *(α ist nach oben beschränkt)*

c) $x < y \in \alpha \implies x \in \alpha$ *(ist y in α, so auch jedes kleinere Element)*

d) $x \in \alpha \implies \exists\, y \in \alpha \colon x < y$ *(α besitzt kein größtes Element)*

Die Menge der reellen Zahlen, ist definiert als die Menge aller Dedekindschen Schnitte; also

$$\mathbb{R} := \{\alpha \subseteq \mathbb{Q} \mid \alpha \text{ ist Dedekindscher Schnitt}\} \subset \mathcal{P}(\mathbb{Q}).$$

Bemerkung 28.

- *Jedes $\alpha \in \mathbb{R}$ ist ein Intervall in $\mathbb{Q}$:*

  In der Tat, für $x \in \mathbb{Q}$ mit $\alpha \ni z < x < y \ni \alpha$, folgt $x \in \alpha$ wegen c) *(da $x < y$).*

- *Jedes $\alpha \in \mathbb{R}$ ist „offen“ in dem Sinne, dass für jedes $x \in \alpha$ ein $0 < \varepsilon \in \mathbb{Q}$ existiert, sodass $\mathbb{Q} \supseteq (-\infty, x + \varepsilon) \subseteq \alpha$ gilt:*

  In der Tat, wegen d) *existiert ein $x < y \in \alpha$. Für $0 < \varepsilon := \frac{y-x}{2} \in \mathbb{Q}$ gilt dann $(-\infty, x+\varepsilon) < y \in \alpha$; also $(-\infty, x + \varepsilon) \subseteq \alpha$ wegen* c).

(Beachte: Nicht jedes Intervall in $\mathbb{Q}$ ist (notwendigerweise) von der Form (94); denn $\mathbb{Q}$ ist ja nicht vollständig angeordnet.)

Beispiel 24.

- *Für jedes $z \in \mathbb{Q}$, ist das Intervall*

$$\mathbb{R} \ni \alpha[z] := (-\infty, z) \tag{102}$$

  ein Dedekindscher Schnitt.

  Beweis der Behauptung: Die Eigenschaften a) und c) sind klar per Definition. Weiterhin gilt b), wegen $\alpha[z] \leq z$. Schließlich gilt d) wegen $x < x + \frac{z-x}{2} < z$, für alle $x \in \alpha[z]$. □

- *Gegeben $k \in \mathbb{N}_{>0}$ und $0 < x \in \mathbb{Q}$, so ist die Menge*

$$\mathbb{R} \ni \alpha = (-\infty, 0) \cup \{0 \leq y \in \mathbb{Q} \mid y^k < x\} \tag{103}$$

  ein Dedekindscher Schnitt.

  Beweis der Behauptung: Die Eigenschaften a) ist evident, und b) gilt wegen $\alpha \leq \max(1, x)$ (vgl. Beweis von Satz 9). Weiterhin gelten c) und d) wegen Übung 45.a) und Bemerkung 5.1). □

Die reellen Zahlen werden zu einer angeordneten Menge, vermöge der Relation

$$\alpha \leq_\mathbb{R} \beta \quad \text{für} \quad \alpha, \beta \in \mathbb{R} \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad \alpha \subseteq \beta.$$

Nachweis der Anordnungseigenschaft: Die Ordnungsaxiome 1)-3) in Definition 10 sind unmittelbar einsichtig. Für das Axiom 4) (also die Totalität), seien $\alpha, \beta \in \mathbb{R}$ vorgegeben.

Angenommen, es gilt $\beta \nsubseteq \alpha$.

- Dann existiert ein $x \in \beta$ mit $x \notin \alpha$.
- Es folgt $\alpha \leq x$; denn ansonsten existiert ein ja ein $y \in \alpha$ mit $x < y$, also $x \in \alpha$ nach c).
- Aus c) (für β) folgt $\alpha \subseteq \beta$, was die Behauptung zeigt. □

Lemma 23. *Ist $\emptyset \neq X \subseteq \mathbb{R}$ nichtleer und nach oben beschränkt, so existiert $\sup_{\mathbb{R}}(X) \in \mathbb{R}$; und zwar gilt $\sup_{\mathbb{R}}(X) = \alpha := \bigcup_{\beta \in X} \beta$.*

Beweis. Es ist α ein Dedekindscher Schnitt:

- Es gilt $\alpha \neq \emptyset$ (also Eigenschaften a)):

  Wegen $X \neq \emptyset$, existiert ein $\beta \in X \subseteq \mathbb{R}$. Es gilt $\beta \neq \emptyset$ wegen a) (für β); also $\emptyset \neq \beta \subseteq \alpha$.
- Es gilt die Eigenschaften b):

  Per Annahme ist X nach oben beschränkt. Es existiert also ein $\gamma \in \mathbb{R}$ mit $X \leq_{\mathbb{R}} \gamma$. Wegen Eigenschaft b) (für γ), existiert ein $o \in \mathbb{Q}$ mit $\gamma \leq o$. Es folgt

$$\beta \leq_{\mathbb{R}} \gamma \quad \forall \beta \in X \qquad \Longrightarrow \qquad \beta \subseteq \gamma \leq o \quad \forall \beta \in X \qquad \Longrightarrow \qquad \alpha = \bigcup_{\beta \in X} \beta \leq o.$$

- Es gilt die Eigenschaften c):

  Sei $x < y \in \alpha$. Per definitionem existiert ein $\beta \in X$ mit $y \in \beta$. Wegen c) für β, folgt $x \in \beta \subseteq \alpha$; was die Eigenschaft c) für α zeigt.
- Es gilt die Eigenschaften d):

  Gilt $x \in \alpha$, so gilt $x \in \beta$ für ein $\beta \in X$. Wegen d) für β, existiert ein $x < y \in \beta \subseteq \alpha$, was die Eigenschaft d) für α zeigt.

Es gilt offensichtlich

$$\beta \subseteq \alpha \quad \forall \beta \in X \qquad \Longleftrightarrow \qquad \beta \leq_{\mathbb{R}} \alpha \quad \forall \beta \in X \qquad \Longleftrightarrow \qquad X \leq_{\mathbb{R}} \alpha.$$

Sei nun $X \leq_{\mathbb{R}} \gamma \in \mathbb{R}$ eine obere Schranke von X. Dann gilt

$$\beta \subseteq \gamma \quad \forall \beta \in X \qquad \Longrightarrow \qquad \alpha = \bigcup_{\beta \in X} \beta \subseteq \gamma \quad \Leftrightarrow \quad \alpha \leq_{\mathbb{R}} \gamma.$$

Dies zeigt $\alpha = \sup_{\mathbb{R}}(X)$. □

Um $(\mathbb{R}, \leq_{\mathbb{R}})$ zu einem vollständig angeordneten Körper zu machen, müssen wir noch die entsprechend Rechenoperationen definieren. Dies geschieht im Abschnitt 4.3.1, und beweist den ersten Teil von Satz 10. Wir werden zudem die zugehörige Körpereinbettung aus Satz 7 explizit mituntersuchen (siehe (104)), um die Definitionen anhand von Beispielen plausibler zu machen.

4.3.1 Definition der Rechenoperationen auf den reellen Zahlen (*)

(In der Vorlesung Addition im Detail behandelt (also die erste Seite). Den Rest nur grob erklärt.)

Wir definieren die neutralen Elemente durch (vgl. (102))

$$0_{\mathbb{R}} := \alpha[0] = (-\infty, 0) \qquad \text{und} \qquad 1_{\mathbb{R}} := \alpha[1] = (-\infty, 1).$$

Wir betrachten zudem die Abbildung

$$\Psi_{\mathbb{R}} \colon \mathbb{Q} \to \mathbb{R}, \qquad x \mapsto \alpha[x] = (-\infty, x), \tag{104}$$

welche die folgenden Eigenschaften besitzt:

- $\Psi_\mathbb{R}$ ist injektiv, denn für $x, y \in \mathbb{Q}$ mit $x < y$ gilt $\Psi_\mathbb{R}(x) \notin x + \frac{y-x}{2} \in \Psi_\mathbb{R}(y)$.
- $\Psi_\mathbb{R}$ ist ordnungserhaltend, denn für $x, y \in \mathbb{Q}$ gilt

$$x \leq y \qquad \Longleftrightarrow \qquad \Psi_\mathbb{R}(x) = (-\infty, x) \subseteq (-\infty, y) = \Psi_\mathbb{R}(y) \qquad \Longleftrightarrow \qquad \Psi_\mathbb{R}(x) \leq_\mathbb{R} \Psi_\mathbb{R}(y).$$

- Es gilt $\Psi_\mathbb{R}(0) = 0_\mathbb{R}$, sowie $\Psi_\mathbb{R}(1) = 1_\mathbb{R}$.

Die Abbildung (104) entspricht der Körpereinbettung aus Satz 7. Wir definieren nun die Rechenoperationen auf $\mathbb{R}$ in der Form, dass (104) ein Körperhomomorphismus wird (also eine Körpereinbettung):

- Wir definieren die Addition, sowie die additive Inversenbildung durch

$$\begin{aligned} \alpha +_\mathbb{R} \beta &:= \{v + w \mid v \in \alpha \quad \wedge \quad w \in \beta\} \\ -_\mathbb{R}\, \alpha &:= \{v \in \mathbb{Q} \mid \exists\, 0 < \tau \in \mathbb{Q}\colon\ -v - \tau \notin \alpha\} \\ \alpha -_\mathbb{R} \beta &:= \alpha +_\mathbb{R} (-_\mathbb{R}\, \beta). \end{aligned}$$

Nachweis der Schnitteigenschaften von $\alpha +_\mathbb{R} \beta$ sowie $-_\mathbb{R}\, \alpha$:

- Sei $\gamma := \alpha +_\mathbb{R} \beta$, sowie $o, \tilde{o} \in \mathbb{Q}$ mit $\alpha \leq o$ und $\beta \leq \tilde{o}$. Offensichtlich gilt a) für γ; sowie auch b), wegen $\gamma \leq o + \tilde{o}$. Sei nun $x < y \in \gamma$ mit $y = v + w$ für $v \in \alpha$ und $w \in \beta$. Dann gilt

$$\begin{aligned} \tilde{v} := (v - \tfrac{y-x}{2}) < v \quad \wedge \quad \tilde{w} := (w - \tfrac{y-x}{2}) < w \quad &\overset{c)}{\Longrightarrow} \quad \tilde{v} \in \alpha \quad \wedge \quad \tilde{w} \in \beta \\ &\Longrightarrow \quad x = \big((v - \tfrac{y-x}{2}) + (w - \tfrac{y-x}{2})\big) \in \gamma, \end{aligned}$$

was c) zeigt. Sei schließlich $x = v + w \in \gamma$ mit $v \in \alpha$ und $w \in \beta$. Wegen d) existieren $v < \tilde{v} \in \alpha$ und $w < \tilde{w} \in \beta$, also $x < y := \tilde{v} + \tilde{w} \in \gamma$, was d) für γ zeigt.

- Sei $\gamma := -_\mathbb{R}\, \alpha$, sowie $o \in \mathbb{Q}$ mit $\alpha \leq o$. Dann gilt $-y \in \gamma$ für $y := o + 2$, wegen $y - 1 = o + 1 \notin \alpha$. Es folgt $\gamma \neq \emptyset$, also gilt a) für γ. Ist weiterhin $x \in \alpha$ ein fixiertes Element, so folgt (wegen c) für α) für alle $y \in \mathbb{Q}$ und $0 < \tau \in \mathbb{Q}$ mit $y - \tau \notin \alpha$, dass $x \in \alpha \leq y$ gilt; also $-y \leq -x$. Dies zeigt $\gamma \leq -x$, also b) für γ. Sei nun $x < y$ für $y \in \mathbb{Q}$, mit $-y - \tau \notin \alpha$ für ein $0 < \tau \in \mathbb{Q}$. Dann gilt $-y - \tau < -x - \tau$; also $-x - \tau \notin \alpha$ wegen c) für α. Dies zeigt $x \in \gamma$, also c) für γ. Sei schließlich $x \in \mathbb{Q}$ und $0 < \tau \in \mathbb{Q}$ mit $x - \tau \notin \alpha$. Dann gilt $y - \frac{\tau}{2} \notin \alpha$ für $y := x - \frac{\tau}{2} > x$, also $\gamma \ni -x < -y \in \gamma$, was d) für γ zeigt. □

Beispiel 25.

- *Für $x, y \in \mathbb{Q}$ gilt*

$$\alpha[x] +_\mathbb{R} \alpha[y] = \alpha[x + y] \qquad \textit{und} \qquad -_\mathbb{R}\, \alpha[x] = \alpha[-x]. \tag{105}$$

Beweis von (105)*:* Wir erhalten

$$z \in \alpha[-x] \quad \Leftrightarrow \quad z < -x \quad \Leftrightarrow \quad x < -z \quad \Leftrightarrow \quad \exists\, \tau > 0\colon x < -z - \tau \quad \Leftrightarrow \quad z \in -_\mathbb{R}\alpha[x],$$

was die rechte Seiten von (105) zeigt. Die linke Seite von (105) erhält man wie folgt:

* Wegen $\alpha[x] < x$ und $\alpha[y] < y$ gilt $\alpha[x] + \alpha[y] < x + y$, also $\alpha[x] +_\mathbb{R} \alpha[y] \subseteq \alpha[x + y]$.
* Ist $z \in \alpha[x + y]$, also $z < x + y$, so gilt[13]

$$z < \underbrace{x - \varepsilon}_{=:\, v} + \underbrace{y - \varepsilon}_{=:\, w} \qquad \text{für} \qquad \varepsilon := \frac{x + y - z}{4},$$

mit $v \in \alpha[x]$ wegen $v < x$, sowie $w \in \alpha[y]$ wegen $w < y$. Daher gilt $z < v + w \in \alpha[x] +_\mathbb{R} \alpha[y]$; also $z \in \alpha[x] +_\mathbb{R} \alpha[y]$, wegen c) für den Schnitt $\alpha[x] +_\mathbb{R} \alpha[y]$. □

[13] Es gilt $x - \varepsilon + y - \varepsilon = x + y - 2\varepsilon = z + (x + y - z) - 2\varepsilon = z + \frac{x+y-z}{2} > z$.

– *Ist* $\alpha = (-\infty, 0) \cup \{0 \leq y \in \mathbb{Q} \mid y^k < x\}$ *wie in* (103), *so gilt* $-_{\mathbb{R}}\, \alpha = \{-y \mid y^k > x\}$*:*

  Sei $y \in \mathbb{Q}$ *mit* $y^k > x$. *Ist* $0 < \tau \in \mathbb{Q}$ *genügend klein, so folgt aus der zweiten Zeile in* (96), *dass* $y - \tau \notin \alpha$ *gilt; also* $-y \in -_{\mathbb{R}}\alpha$. *Gilt umgekehrt* $y \in -_{\mathbb{R}}\alpha$, *also* $y - \tau \notin \alpha$ *für* $y \in \mathbb{Q}$ *und* $0 < \tau \in \mathbb{Q}$, *so folgt* $y^k > (y-\tau)^k \geq x$.

Es ist $(\mathbb{R}, +_{\mathbb{R}}, 0_{\mathbb{R}})$ eine abelsche Gruppe:

– Offensichtlich gilt $\alpha +_{\mathbb{R}} \beta = \beta +_{\mathbb{R}} \alpha$ für alle $\alpha, \beta \in \mathbb{R}$, also (52).
– Wir zeigen $\alpha +_{\mathbb{R}} 0_{\mathbb{R}} = \alpha$, also (53):
  Sei $\alpha \in \mathbb{R}$. Für $x \in \alpha$ und $y \in 0_{\mathbb{R}}$ $(y < 0)$ gilt $x + y < x$, also $x + y \in \alpha$ wegen c). Dies zeigt $\alpha +_{\mathbb{R}} 0_{\mathbb{R}} \subseteq \alpha$. Ist nun $x \in \alpha$ vorgegeben, so existiert nach d) ein $x < y \in \alpha$. Es folgt $x = x + (y - x)$ mit $y - x < 0$, also $y - x \in 0_{\mathbb{R}}$. Dies zeigt $\alpha \subseteq \alpha +_{\mathbb{R}} 0_{\mathbb{R}}$.
– Für $\alpha, \beta, \gamma \in \mathbb{R}$ gilt

$$\begin{aligned} \alpha +_{\mathbb{R}} (\beta +_{\mathbb{R}} \gamma) &= \{x + \tau \mid x \in \alpha \ \wedge\ \tau \in (\beta +_{\mathbb{R}} \gamma)\} \\ &= \{x + (y + z) \mid x \in \alpha \ \wedge\ y \in \beta \ \wedge\ z \in \gamma)\} \\ &= \{(x + y) + z) \mid x \in \alpha \ \wedge\ y \in \beta \ \wedge\ z \in \gamma)\} \\ &= \{\tau + z \mid \tau \in (\alpha +_{\mathbb{R}} \beta) \ \wedge\ z \in \gamma\} \\ &= (\alpha +_{\mathbb{R}} \beta) +_{\mathbb{R}} \gamma, \end{aligned}$$

  was (51) zeigt.
– Sei $\alpha \in \mathbb{R}$. Wir zeigen $\alpha -_{\mathbb{R}} \alpha = 0_{\mathbb{R}}$, also (54):
  * Sei zunächst $x \in \alpha$, sowie $y \in \mathbb{Q}$ und $0 < \tau \in \mathbb{Q}$ mit $y - \tau \notin \alpha$. Es gilt dann $x < y - \tau < y$ (wegen c)); also $x + (-y) < 0$, und daher $x + (-y) \in 0_{\mathbb{R}}$. Dies zeigt $\alpha -_{\mathbb{R}} \alpha \subseteq 0_{\mathbb{R}}$.
  * Sei nun $w \in 0_{\mathbb{R}}$, also $\mathbb{Q} \ni w < 0$ vorgegeben. Wir setzen $v := -\frac{w}{2} > 0$; und wählen $n \in \mathbb{Z}$ maximal mit $n \cdot v \in \alpha$, also $(n+1) \cdot v \notin \alpha$.

    Beweis der Existenz: Wegen b) gilt $\alpha \leq o$ für ein $o \in \mathbb{Q}$; und wegen Proposition 2.3) gilt dann $\alpha \leq o \leq q \cdot v$ (also $q \cdot v \notin \alpha$) für ein $q \in \mathbb{N} \subseteq \mathbb{Z}$. Ähnlich erhält man $p \cdot v \in \alpha$ für ein $p \in \mathbb{Z}$, indem man ein Element in α fixiert, und Proposition 2.3) in entsprechender Weise anwendet. Es gilt dann $n = \max(\{p \leq \ell \leq q \mid \ell \cdot v \in \alpha\})$. □

    Wir setzen $x := n \cdot v \in \alpha$, $y := (n+2) \cdot v \notin \alpha$ und $\tau := v > 0$. Dann gilt $y - \tau = (n+1) \cdot v \notin \alpha$, also $-y \in -_{\mathbb{R}}\, \alpha$. Per Konstruktion gilt dann $x + (-y) = w$, also $w \in \alpha -_{\mathbb{R}} \alpha$, was $0_{\mathbb{R}} \subseteq \alpha -_{\mathbb{R}} \alpha$ zeigt.

Wir erhalten für $\alpha, \beta, \gamma \in \mathbb{R}$:

$$\begin{aligned} \alpha \leq_{\mathbb{R}} \beta \qquad &\Longrightarrow \qquad \alpha \subseteq \beta \\ &\Longrightarrow \qquad v + w \in \beta +_{\mathbb{R}} \gamma \qquad \forall\, v \in \alpha,\ w \in \gamma \\ &\Longrightarrow \qquad \alpha +_{\mathbb{R}} \gamma \subseteq \beta +_{\mathbb{R}} \gamma \\ &\Longrightarrow \qquad \alpha +_{\mathbb{R}} \gamma \leq_{\mathbb{R}} \beta +_{\mathbb{R}} \gamma, \end{aligned}$$

was (68) zeigt. Weiterhin gilt die linke Seite von (71), wegen (der linken Seite von) (105).

• Wir definieren für $\mathbb{R} \ni \alpha, \beta >_{\mathbb{R}} 0_{\mathbb{R}}$:

$$\begin{aligned} \alpha \cdot_{\mathbb{R}} \beta &:= \{x \in \mathbb{Q} \mid \exists\, (0 < v \in \alpha \ \wedge\ 0 < w \in \beta) \colon x \leq v \cdot w\} \\ \alpha^{-1} &:= (-\infty, 0] \cup \{0 < x \in \mathbb{Q} \mid \exists\, 0 < \tau \in \mathbb{Q} \colon x^{-1} - \tau \notin \alpha\}; \end{aligned}$$

und dann $\alpha \cdot_{\mathbb{R}} 0_{\mathbb{R}} := 0_{\mathbb{R}} =: 0_{\mathbb{R}} \cdot_{\mathbb{R}} \alpha$, $\alpha^{-1} := -_{\mathbb{R}}(-_{\mathbb{R}}\alpha)^{-1}$ für $\alpha <_{\mathbb{R}} 0$, sowie

$$\alpha \cdot_{\mathbb{R}} \beta = \begin{cases} (-_{\mathbb{R}}\alpha) \cdot_{\mathbb{R}} (-_{\mathbb{R}}\beta) & \text{für } \alpha, \beta <_{\mathbb{R}} 0 \\ -_{\mathbb{R}}((-_{\mathbb{R}}\alpha) \cdot_{\mathbb{R}} \beta) & \text{für } \alpha <_{\mathbb{R}} 0,\ \beta >_{\mathbb{R}} 0 \\ -_{\mathbb{R}}(\alpha \cdot_{\mathbb{R}} (-_{\mathbb{R}}\beta)) & \text{für } \alpha >_{\mathbb{R}} 0,\ \beta <_{\mathbb{R}} 0. \end{cases}$$

Nachweis der Schnitteigenschaft von $\alpha \cdot_{\mathbb{R}} \beta$ und α^{-1}:

- Sei $\gamma := \alpha \cdot_{\mathbb{R}} \beta$. Es genügt den Fall $\alpha, \beta >_{\mathbb{R}} 0$ abzuhandeln. Es gilt dann $\alpha \setminus 0_{\mathbb{R}} \neq \emptyset$, $\beta \setminus 0_{\mathbb{R}} \neq \emptyset$; also existieren $0 < x \in \alpha$ und $0 < y \in \beta$, womit $x \cdot y \in \gamma \neq \emptyset$ gilt, also a). Sind weiterhin $o, \tilde{o} > 0$ mit $\alpha \leq o$ und $\beta \leq \tilde{o}$, so folgt unmittelbar $\gamma \leq o \cdot \tilde{o}$, also b). Sei nun $x < y \in \gamma$, also $y \leq v \cdot w$ für gewisse $v \in \alpha$ und $w \in \beta$. Dann gilt ebenfalls $y < v \cdot w$, also $y \in \gamma$, was c) zeigt. Sei schließlich $x \in \gamma$ gegeben, mit $x \leq v \cdot w$ für $v \in \alpha$ und $w \in \beta$. Wegen d) (für α und β), existieren $v < \tilde{v} \in \alpha$ und $w < \tilde{w} \in \beta$; und wir erhalten $y \leq \tilde{v} \cdot \tilde{w}$ für $y := \tilde{v} \cdot \tilde{w}$, was d) zeigt.
- Sei $\gamma := \alpha^{-1}$. Es genügt ebenfalls den Fall $\alpha >_{\mathbb{R}} 0$ abzuhandeln. Es gilt $\alpha \leq o$ für ein $o \in \mathbb{Q}$, also $x - \tau \notin \alpha$ für $x := o + 2$ und $\tau := 1$; und daher $x^{-1} \in \gamma \neq \emptyset$, also a). Ist weiterhin $0 < x \in \alpha$ ein fixiertes Element, so folgt (wegen c) für α) für alle $y \in \mathbb{Q}$ und $0 < \tau \in \mathbb{Q}$ mit $y - \tau \notin \alpha$, dass $x \in \alpha \leq y$ gilt; also $y^{-1} \leq x^{-1}$. Dies zeigt $\gamma \leq x^{-1}$, also b) für γ. Sei nun $x < y$ für $y \in \mathbb{Q}$, mit $y^{-1} - \tau \notin \alpha$ für ein $0 < \tau \in \mathbb{Q}$. Ist $x \leq 0$, so folgt $x \in \gamma$ sofort aus der Definition von γ. Ist $x > 0$, so folgt $y^{-1} - \tau < x^{-1} - \tau$; also $x^{-1} - \tau \notin \alpha$ wegen c) für α. Dies zeigt $x \in \gamma$, also c) für γ. Sei schließlich $0 < x \in \mathbb{Q}$ und $0 < \tau \in \mathbb{Q}$ mit $x^{-1} - \tau \notin \alpha$ (also $0 < x \in \alpha$). Dann gilt $y - \frac{\tau}{2} \notin \alpha$ für $y := x^{-1} - \frac{\tau}{2} > x^{-1}$, also $\gamma \ni x < y^{-1} \in \gamma$, was d) für γ zeigt. □

Beispiel 26. *Für $x, y \in \mathbb{Q} \setminus \{0\}$ gilt*

$$\alpha[x] \cdot_{\mathbb{R}} \alpha[y] = \alpha[x \cdot y] \qquad \textit{und} \qquad \alpha[x]^{-1} = \alpha[x^{-1}]. \tag{106}$$

Beweis von (106)*:* Wegen der rechten Seite von (105) reicht es aus, die Aussagen für $x, y > 0$ zu zeigen. Es gilt $\alpha[x]^{-1} \supseteq (-\infty, 0] \subseteq \alpha[x^{-1}]$; und für $z > 0$ erhalten wir

$$z \in \alpha[x^{-1}] \quad \Leftrightarrow \quad z < x^{-1} \quad \Leftrightarrow \quad x < z^{-1} \quad \Leftrightarrow \quad \exists\, \tau > 0\colon x < y^{-1} - \tau \quad \Leftrightarrow \quad y \in \alpha[x]^{-1},$$

was die rechte Seiten von (106) zeigt. Die linke Seite von (105) erhält man wie folgt:

- ∗ Wegen $\alpha[x] < x$ und $\alpha[y] < y$ gilt $\alpha[x] \cdot_{\mathbb{R}} \alpha[y] < x \cdot y$, also $\alpha[x] \cdot_{\mathbb{R}} \alpha[y] \subseteq \alpha[x \cdot y]$.
- ∗ Ist $z \in \alpha[x \cdot y]$, also $z < x \cdot y$, so gilt[14]

$$z \leq \underbrace{(x - \varepsilon)}_{=:\, v} \cdot \underbrace{(y - \varepsilon)}_{=:\, w} = x \cdot y - \varepsilon(x + y) + \varepsilon^2 \qquad \text{für} \qquad \varepsilon := \min\left(\tfrac{x}{2}, \tfrac{y}{2}, \tfrac{x \cdot y - z}{2 \cdot (x+y)}\right),$$

  mit $0 < v \in \alpha[x]$ wegen $v < x$, sowie $0 < w \in \alpha[y]$ wegen $w < y$; also $z \in \alpha[x] \cdot_{\mathbb{R}} \alpha[y]$.[15] □

Es ist $(\mathbb{R}_\times, \cdot_{\mathbb{R}}, 1_{\mathbb{R}})$ eine abelsche Gruppe: $(\mathbb{R}_\times = \mathbb{R} \setminus \{0_{\mathbb{R}}\})$

- Offensichtlich gilt $\alpha \cdot_{\mathbb{R}} \beta = \beta \cdot_{\mathbb{R}} \alpha$ für alle $0_{\mathbb{R}} <_{\mathbb{R}} \alpha, \beta \in \mathbb{R}$. Eine einfache Fallunterscheidung, zeigt dann selbige Aussage für alle $\alpha, \beta \in \mathbb{R}_\times$, also (52).
- Wir zeigen $\alpha \cdot_{\mathbb{R}} 1_{\mathbb{R}} = \alpha$ für $0_{\mathbb{R}} <_{\mathbb{R}} \alpha \in \mathbb{R}$. Eine Fallunterscheidung zeigt dann $\alpha \cdot_{\mathbb{R}} 1_{\mathbb{R}} = \alpha$ für alle $\alpha \in \mathbb{R}_\times$, also (53).

  Sei also $0 <_{\mathbb{R}} \alpha \in \mathbb{R}$. Für $x \in \alpha$ und $0 < y \in 1_{\mathbb{R}}$ gilt $x \cdot y < x$, also $x \cdot y \in \alpha$ wegen c) (für α). Dies zeigt $\alpha \cdot_{\mathbb{R}} 1_{\mathbb{R}} \subseteq \alpha$. Ist nun $0 < x \in \alpha$ vorgegeben, so existiert nach d) ein $x < v \in \alpha$. Es folgt $x = v \cdot w$ mit $0 < w := \frac{x}{y} < 1$, also $w \in 1_{\mathbb{R}}$. Dies zeigt $x \in \alpha \cdot_{\mathbb{R}} 1_{\mathbb{R}}$; und es folgt $\alpha \subseteq \alpha \cdot_{\mathbb{R}} 1_{\mathbb{R}}$.

[14] Es gilt $\varepsilon(x+y) \leq \frac{x \cdot y - z}{2}$, also $x \cdot y - \varepsilon(x+y) + \varepsilon^2 > x \cdot y - \varepsilon(x+y) \geq x \cdot y - \frac{x \cdot y - z}{2} = z + (x \cdot y - z) - \frac{x \cdot y - z}{2} = z + \frac{x \cdot y - z}{2} > z$.
[15] Es gilt $x > v = x - \varepsilon \geq x - \frac{x}{2} = \frac{x}{2} > 0$; sowie $y > w = y - \varepsilon \geq y - \frac{y}{2} = \frac{y}{2} > 0$.

– Für $0_{\mathbb{R}} <_{\mathbb{R}} \alpha, \beta, \gamma \in \mathbb{R}$ gilt

$$\begin{aligned}
\alpha \cdot_{\mathbb{R}} (\beta \cdot_{\mathbb{R}} \gamma) &= \{x \mid \exists\, (0 < u \in \alpha \quad \wedge \quad 0 < \tilde{u} \in \beta \cdot_{\mathbb{R}} \gamma) \colon x \leq u \cdot \tilde{u}\} \\
&= \{x \mid \exists\, (0 < u \in \alpha \quad \wedge \quad 0 < v \in \beta \quad \wedge \quad 0 < w \in \gamma) \colon x \leq u \cdot v \cdot w\} \\
&= \{x \mid \exists\, (0 < u \in \alpha \cdot_{\mathbb{R}} \beta \quad \wedge \quad 0 < \tilde{u} \in \gamma) \colon x \leq u \cdot \tilde{u}\} \\
&= (\alpha \cdot_{\mathbb{R}} \beta) \cdot_{\mathbb{R}} \gamma.
\end{aligned}$$

Eine Fallunterscheidung[16] zeigt dann $\alpha \cdot_{\mathbb{R}} (\beta \cdot_{\mathbb{R}} \gamma) = (\alpha \cdot_{\mathbb{R}} \beta) \cdot_{\mathbb{R}} \gamma$ für alle $\alpha, \beta, \gamma \in \mathbb{R}$, also (51).

– Sei $\alpha \in \mathbb{R}$. Wir zeigen $\alpha \cdot_{\mathbb{R}} \alpha^{-1} = 1_{\mathbb{R}}$ für $0_{\mathbb{R}} <_{\mathbb{R}} \alpha \in \mathbb{R}$. Es folgt dann unmittelbar auch $\alpha \cdot_{\mathbb{R}} \alpha^{-1} = 1_{\mathbb{R}}$ für $\mathbb{R} \ni \alpha <_{\mathbb{R}} 0_{\mathbb{R}}$; also (54):

  * Sei zunächst $x \in \alpha$, sowie $0 < y \in \mathbb{Q}$ und $0 < \tau \in \mathbb{Q}$ mit $y - \tau \notin \alpha$ (also $y^{-1} \in \alpha$). Es gilt dann $x < y - \tau < y$ (wegen c)); also $x \cdot y^{-1} < 1$, und daher $x \cdot y^{-1} \in 1_{\mathbb{R}}$. Dies zeigt $\alpha \cdot_{\mathbb{R}} \alpha^{-1} \subseteq 1_{\mathbb{R}}$.
  * Sei $0 < z < 1$ vorgegeben, $\varepsilon := 1 - z$, sowie $n[\varepsilon] \in \mathbb{N}_{>0}$ mit $\frac{2}{n[\varepsilon]+2} < \varepsilon$ (Proposition 2). Wir fixieren $\delta > 0$ mit $n[\varepsilon] \cdot \delta \in \alpha$ (Proposition 2 und $\alpha >_{\mathbb{R}} 0_{\mathbb{R}}$); und wählen $m \geq n[\varepsilon]$ maximal mit $m \cdot \delta \in \alpha$, also $(m+1) \cdot \delta \geq \alpha$. Dann gilt $y := (m+2) \cdot \delta > \alpha$ mit $y - 1 \notin \alpha$, also $y^{-1} \in \alpha^{-1}$. Wir erhalten

$$\alpha \cdot_{\mathbb{R}} \alpha^{-1} \ni x \cdot y^{-1} = \frac{m}{m+2} = 1 - \frac{2}{m+2} \geq 1 - \frac{2}{n[\varepsilon]+2} > 1 - \varepsilon = z,$$

  also $z \in \alpha \cdot_{\mathbb{R}} \alpha^{-1}$ wegen c). Dies zeigt $1_{\mathbb{R}} \subseteq \alpha \cdot_{\mathbb{R}} \alpha^{-1}$.

Per definitionem gilt $\alpha \cdot_{\mathbb{R}} \beta = 0_{\mathbb{R}}$, falls $\alpha = 0_{\mathbb{R}}$ oder $\beta = 0_{\mathbb{R}}$. Gilt weiterhin $\mathbb{R} \ni \alpha, \beta >_{\mathbb{R}} 0_{\mathbb{R}}$, so existieren $0 < x \in \alpha$ und $0 < y \in \beta$, also $0 < x \cdot y \in \alpha \cdot_{\mathbb{R}} \beta$. Daher gilt $0_{\mathbb{R}} \subseteq \alpha \cdot_{\mathbb{R}} \beta$; also $\alpha \cdot_{\mathbb{R}} \beta >_{\mathbb{R}} 0_{\mathbb{R}}$, was (69) zeigt. Zudem gilt die rechte Seite von (71), wegen (der linken Seite von) (106).

• Es gilt das Distributivgesetz (66); also

$$\alpha \cdot_{\mathbb{R}} (\beta +_{\mathbb{R}} \gamma) = \alpha \cdot_{\mathbb{R}} \beta +_{\mathbb{R}} \alpha \cdot_{\mathbb{R}} \gamma \qquad \forall\, \alpha, \beta, \gamma \in \mathbb{R}. \tag{107}$$

Beweis von (107)*:* Wir beweisen die Aussage in mehreren Schritten:

– Es gilt (107) für $\alpha, \beta, \gamma >_{\mathbb{R}} 0_{\mathbb{R}}$:

  Seien hierfür $\alpha, \beta, \gamma >_{\mathbb{R}} 0_{\mathbb{R}}$ fixiert, und setze $\delta_1 := \alpha \cdot_{\mathbb{R}} (\beta +_{\mathbb{R}} \gamma)$ und $\delta_2 := \alpha \cdot_{\mathbb{R}} \beta +_{\mathbb{R}} \alpha \cdot_{\mathbb{R}} \gamma$:

  * Für $x \in \delta_1$, existieren $0 < u \in \alpha$ sowie $v \in \beta$ und $w \in \gamma$ mit $0 < v+w$, sodass $x \leq u \cdot (v+w) = u \cdot v + u \cdot w$ gilt. Wegen $\beta, \gamma >_{\mathbb{R}} 0_{\mathbb{R}}$, existieren $0 < \tilde{v} \in \beta$ und $0 < \tilde{w} \in \gamma$, mit $v \leq \tilde{v}$ und $w \leq \tilde{w}$. Daher gilt $x \leq z := u \cdot \tilde{v} + u \cdot \tilde{w}$ mit $u \cdot \tilde{v} \in \alpha \cdot_{\mathbb{R}} \beta$ und $u \cdot \tilde{w} \in \alpha \cdot_{\mathbb{R}} \gamma$; also $x \leq z$ mit $z \in \delta_2$; und daher $x \in \delta_2$ wegen c) für δ_2. Dies zeigt $\delta_1 \subseteq \delta_2$
  * Für $x \in \delta_2$, existieren $0 < u, \tilde{u} \in \alpha$, $0 < v \in \beta$ und $0 < w \in \gamma$ mit $y = u \cdot v + \tilde{u} \cdot w$; also $y \leq o \cdot (v + w)$ mit $0 < o := \max(u, \tilde{u}) \in \alpha$, und $0 < v + w \in \beta +_{\mathbb{R}} \gamma$. Daher gilt $x \in \delta_1$; also folgt $\delta_2 \subseteq \delta_1$.

– Es gilt (107) für $\alpha, \beta >_{\mathbb{R}} 0_{\mathbb{R}}$, $\gamma <_{\mathbb{R}} 0_{\mathbb{R}}$, sowie $\beta +_{\mathbb{R}} \gamma >_{\mathbb{R}} 0_{\mathbb{R}}$:

  In der Tat erhalten wir aus dem ersten Punkt (zweiter Schritt), sowie der Definition der Multiplikation und $-_{\mathbb{R}} (-_{\mathbb{R}} \delta) = \delta$ für alle $\delta \in \mathbb{R}$ (letzter Schritt), dass

$$\alpha \cdot_{\mathbb{R}} \beta = \alpha \cdot_{\mathbb{R}} (\beta +_{\mathbb{R}} \gamma +_{\mathbb{R}} (-_{\mathbb{R}} \gamma)) = \alpha \cdot_{\mathbb{R}} (\beta +_{\mathbb{R}} \gamma) +_{\mathbb{R}} \alpha \cdot_{\mathbb{R}} (-_{\mathbb{R}} \gamma) = \alpha \cdot_{\mathbb{R}} (\beta +_{\mathbb{R}} \gamma) -_{\mathbb{R}} (\alpha \cdot_{\mathbb{R}} \gamma)$$

  gilt. Addition von $\alpha \cdot_{\mathbb{R}} \gamma$ liefert die Behauptung.

– Es gilt (107) für $\alpha, \gamma >_{\mathbb{R}} 0_{\mathbb{R}}$, $\beta <_{\mathbb{R}} 0_{\mathbb{R}}$, sowie $\beta +_{\mathbb{R}} \gamma >_{\mathbb{R}} 0_{\mathbb{R}}$.

  (Gleicher Beweis wie im vorangegangenen Punkt.)

[16] Man benutze $-_{\mathbb{R}} (-_{\mathbb{R}} \alpha) = \alpha$ für $\alpha \in \mathbb{R}$; sowie $-_{\mathbb{R}} \alpha <_{\mathbb{R}} 0_{\mathbb{R}}$ für $0_{\mathbb{R}} <_{\mathbb{R}} \alpha \in \mathbb{R}$.

Bisher wurde gezeigt, dass (107) für $\alpha >_{\mathbb{R}} 0_{\mathbb{R}}$, sowie $\mathbb{R} \ni \beta, \gamma \neq 0_{\mathbb{R}}$ mit $\beta +_{\mathbb{R}} \gamma >_{\mathbb{R}} 0_{\mathbb{R}}$ gilt. Es folgt:

– Es gilt (107) für $\alpha >_{\mathbb{R}} 0_{\mathbb{R}}$, sowie $\mathbb{R} \ni \beta, \gamma \neq 0_{\mathbb{R}}$ mit $\beta +_{\mathbb{R}} \gamma \neq 0_{\mathbb{R}}$:

  Es bleibt der Fall $\beta +_{\mathbb{R}} \gamma <_{\mathbb{R}} 0_{\mathbb{R}}$ abzuklären. Es gilt $(-_{\mathbb{R}}\beta) +_{\mathbb{R}} (-_{\mathbb{R}}\gamma) >_{\mathbb{R}} 0_{\mathbb{R}}$; also folgt

$$\begin{aligned}\alpha \cdot_{\mathbb{R}} (\beta +_{\mathbb{R}} \gamma) &= -_{\mathbb{R}}(\alpha \cdot_{\mathbb{R}} (-_{\mathbb{R}}(\beta +_{\mathbb{R}} \gamma))) = -_{\mathbb{R}}(\alpha \cdot_{\mathbb{R}} ((-_{\mathbb{R}}\beta) +_{\mathbb{R}} (-_{\mathbb{R}}\gamma))) \\ &= -_{\mathbb{R}}(\alpha \cdot_{\mathbb{R}} (-_{\mathbb{R}}\beta) +_{\mathbb{R}} \alpha \cdot_{\mathbb{R}} (-_{\mathbb{R}}\gamma)) = \alpha \cdot_{\mathbb{R}} \beta +_{\mathbb{R}} \alpha \cdot_{\mathbb{R}} \gamma.\end{aligned}$$

– Es gilt (107) für $\alpha \neq 0_{\mathbb{R}}$, sowie $\mathbb{R} \ni \beta, \gamma \neq 0_{\mathbb{R}}$ mit $\beta +_{\mathbb{R}} \gamma \neq 0_{\mathbb{R}}$:

  Es bleibt der Fall $\alpha < 0_{\mathbb{R}}$ abzuklären. Dieser folgt aber problemlos aus der dem vorherigen Punkt (wieder unter Verwendung der Regel $-_{\mathbb{R}}(-_{\mathbb{R}}\delta) = \delta$ für $\delta \in \mathbb{R}$).

Weiterhin ist klar, dass (107) immer für den Fall $\alpha = 0_{\mathbb{R}}$ bzw. immer für den Fall $\beta, \gamma = 0_{\mathbb{R}}$ gilt. Schließlich sieht man dann auch leicht ein, dass (107) für $\alpha \in \mathbb{R}$ und $0_{\mathbb{R}} \neq \beta, \gamma \in \mathbb{R}$ mit $\beta +_{\mathbb{R}} \gamma = 0_{\mathbb{R}}$ gilt:

Es folgt $\beta = -_{\mathbb{R}}\gamma$ bzw. $\gamma = -_{\mathbb{R}}\beta$. Sei ohne Beschränkung $\beta >_{\mathbb{R}} 0_{\mathbb{R}}$. Dann gilt

$$\begin{aligned}\alpha \cdot_{\mathbb{R}} (\beta +_{\mathbb{R}} \gamma) = \alpha \cdot_{\mathbb{R}} 0_{\mathbb{R}} = 0_{\mathbb{R}} &= \alpha \cdot_{\mathbb{R}} \beta -_{\mathbb{R}} (\alpha \cdot_{\mathbb{R}} \beta) = \alpha \cdot_{\mathbb{R}} \beta -_{\mathbb{R}} (-_{\mathbb{R}}(\alpha \cdot_{\mathbb{R}} (-_{\mathbb{R}}\beta))) \\ &= \alpha \cdot_{\mathbb{R}} \beta -_{\mathbb{R}} (-_{\mathbb{R}}(\alpha \cdot_{\mathbb{R}} \gamma)) = \alpha \cdot_{\mathbb{R}} \beta +_{\mathbb{R}} \alpha \cdot_{\mathbb{R}} \gamma,\end{aligned}$$

was die Behauptung zeigt. □

Insgesamt zeigt dies Satz 10.1); und wir haben zudem nachgewiesen, dass die Abbildung (104) eine Körpereinbettung ist.

Bemerkung* 6. *Bevor wir nun mit dem Beweis von Satz 10 fortfahren, wollen wir noch plausibel machen, dass die reelle Zahl (der Dedekindsche Schnitt) (vgl.* (103)*)*

$$\mathbb{R} \ni \alpha = (-\infty, 0) \cup \{0 \leq y \in \mathbb{Q} \mid y^k < x\} \qquad \textit{für} \qquad 0 < x \in \mathbb{Q},\ k \in \mathbb{N}_{>0}$$

tatsächlich die k-te Wurzel des Dedekindschen Schnittes $\alpha[x] = (-\infty, x)$ ist. Es gilt

$$\alpha^k = \{z \in \mathbb{Q} \mid \exists\, (0 < v_i \in \alpha \quad \textit{für} \quad i = 1, \dots, k) \colon z \leq v_1 \cdot \ldots \cdot v_k\} = \{z \in \mathbb{Q} \mid \exists\, 0 < v \in \alpha \colon z \leq v^k\}.$$

*Wegen $v^k < x$ für $0 < v \in \alpha$, gilt dann $\alpha^k \subseteq (-\infty, x)$. Ist nun $0 < z < x$ gegeben, so existiert nach Bemerkung 5.*2) *ein $v \in \alpha$ mit $z < v^k$; also gilt $z \in \alpha^k$. Dies zeigt $(-\infty, x) \subseteq \alpha^k$; also gilt $\alpha^k = (-\infty, x)$.*

4.3.2 Eindeutigkeit des vollständig angeordneten Körpers (*)

(In der Vorlesung nur kurze Erklärung der Beweisidee.)

Wir haben nun alles was wir benötigen, um auch den zweiten Teil von Satz 10 zu beweisen:

Beweis von Satz 10.2). Sei $(\mathbb{K}, \leq)$ vollständig angeordnet, sowie $\Psi_{\mathbb{K}} \colon \mathbb{Q} \to \mathbb{K}$ die zugehörige Körpereinbettung aus Satz 7. Wir definieren

$$\Phi \colon \mathbb{R} \to \mathbb{K}, \qquad \alpha \mapsto \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\alpha)).$$

Das Supremum auf der rechten Seite existiert; denn für $\alpha \in \mathbb{R}$ mit $\alpha \leq o \in \mathbb{Q}$ (vgl. Eigenschaft b)), ist $\emptyset \neq \Psi_{\mathbb{K}}(\alpha) \leq \Psi_{\mathbb{K}}(o)$ nicht leer, und nach oben beschränkt.

- Φ ist ordnungserhaltend und injektiv; denn für $\alpha, \beta \in \mathbb{R}$ gilt

  $$\alpha <_{\mathbb{R}} \beta \quad\Leftrightarrow\quad \alpha \subset \beta \quad\Rightarrow\quad \Psi_{\mathbb{K}}(\alpha) \subset \Psi_{\mathbb{K}}(\beta) \quad\Rightarrow\quad \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\alpha)) < \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\beta)) \quad\Rightarrow\quad \Phi(\alpha) < \Phi(\beta).$$

  Die zweite Implikation ist klar wegen der Injektivität von $\Psi_{\mathbb{K}}$. Für die dritte Implikation, wählen wir[17] $x, y \in \mathbb{Q}$ mit $\alpha < x < y \in \beta$, und erhalten

  $$\Psi_{\mathbb{K}}(\alpha) < \Psi_{\mathbb{K}}(x) < \Psi_{\mathbb{K}}(y) \leq \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\beta)) \quad\Longrightarrow\quad \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\alpha)) \leq \Psi_{\mathbb{K}}(x) < \Psi_{\mathbb{K}}(y) \leq \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\beta)).$$

- Φ ist surjektiv:

  Für $z \in \mathbb{K}$ fixiert, sei $\alpha := \{x \in \mathbb{Q} \mid \Psi_{\mathbb{K}}(x) < z\}$:

  - Es gilt $\alpha \in \mathbb{R}$. In der Tat, die Eigenschaften a) und b) folgen aus Lemma 21 und Proposition 2. Weiterhin ist c) klar, weil $\Psi_{\mathbb{K}}$ ein Körperhomomorphismus ist. Schließlich folgt d) aus der Körpereinbettungseigenschaft von $\Psi_{\mathbb{K}}$, sowie Proposition 2.4) (und Lemma 21).
  - Es gilt $\Phi(\alpha) = z$. In der Tat, sei $0 < \varepsilon \in \mathbb{K}$ vorgegeben. Wegen Proposition 2.4), existiert ein $x \in \mathbb{Q}$ mit $z - \varepsilon < \Psi_{\mathbb{K}}(x) < z$ (also $x \in \alpha$). Wegen (83) in Lemma 20 folgt $\Phi(\alpha) = \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\alpha)) = z$.

- Φ ist ein Körperhomomorphismus:

  - Für $\alpha, \beta \in \mathbb{R}$ erhalten wir

    $$\begin{aligned} \Psi_{\mathbb{K}}(\alpha +_{\mathbb{R}} \beta) &= \{\Psi_{\mathbb{K}}(v + w) \mid v \in \alpha \ \wedge\ w \in \beta\} \\ &= \{\Psi_{\mathbb{K}}(v) + \Psi_{\mathbb{K}}(w) \mid v \in \alpha \ \wedge\ w \in \beta\} \\ &= \{x + y \mid x \in \Psi_{\mathbb{K}}(\alpha) \ \wedge\ y \in \Psi_{\mathbb{K}}(\beta)\} \\ &= \Psi_{\mathbb{K}}(\alpha) + \Psi_{\mathbb{K}}(\beta). \end{aligned} \tag{108}$$

    Proposition 1.1) (dritter Schritt) zeigt dann

    $$\begin{aligned} \Phi(\alpha +_{\mathbb{R}} \beta) = \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\alpha +_{\mathbb{R}} \beta)) &\overset{(108)}{=} \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\alpha) + \Psi_{\mathbb{K}}(\beta)) \\ &= \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\alpha)) + \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\beta)) = \Phi(\alpha) + \Phi(\beta). \end{aligned}$$

    Wegen Terminologie 9.5), gilt dann ebenfalls

    $$\Phi(0_{\mathbb{R}}) = 0_{\mathbb{K}} \qquad \text{sowie} \qquad \Phi(-_{\mathbb{R}}\alpha) = -\Phi(\alpha) \quad \forall\, \alpha \in \mathbb{R}. \tag{109}$$

  - Für $0_{\mathbb{R}} <_{\mathbb{R}} \alpha, \beta \in \mathbb{R}$, erhalten wir

    $$\Psi_{\mathbb{K}}(\alpha \cdot_{\mathbb{R}} \beta) = \{\Psi_{\mathbb{K}}(x) \mid x \in \mathbb{Q}\colon \exists\, (0 < v \in \alpha \ \wedge\ 0 < w \in \beta)\colon x \leq v \cdot w\}.$$

    Offensichtlich gilt dann wegen der Homomorphismeneigenschaft von $\Psi_{\mathbb{K}}$, dass

    $$\Phi(\alpha \cdot_{\mathbb{R}} \beta) = \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\alpha \cdot_{\mathbb{R}} \beta)) = \sup_{\mathbb{K}}(Z) \quad \text{für} \quad Z := \{\Psi_{\mathbb{K}}(v) \cdot \Psi_{\mathbb{K}}(w) \mid 0 < v \in \alpha \ \wedge\ 0 < w \in \beta\}.$$

    Nun gilt weiterhin

    $$Z = A \cdot B \qquad \text{für} \qquad A := \{\Psi_{\mathbb{K}}(v) \mid 0 < v \in \alpha\} \qquad \text{und} \qquad B := \{\Psi_{\mathbb{K}}(w) \mid 0 < w \in \beta\},$$

    mit $\sup_{\mathbb{K}}(A) = \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\alpha))$ und $\sup_{\mathbb{K}}(B) = \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\beta))$ (da $\Psi_{\mathbb{K}}$ ordnungserhaltend ist). Proposition 1.3) (zweiter Schritt) liefert schließlich

    $$\Phi(\alpha \cdot_{\mathbb{R}} \beta) = \sup_{\mathbb{K}}(Z) = \sup_{\mathbb{K}}(A) \cdot \sup_{\mathbb{K}}(B) = \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\alpha)) \cdot \sup_{\mathbb{K}}(\Psi_{\mathbb{K}}(\beta)) = \Phi(\alpha) \cdot \Phi(\beta).$$

    Es folgt nun $\Phi(\alpha \cdot_{\mathbb{R}} \beta) = \Phi(\alpha) \cdot \Phi(\beta)$ für alle $\alpha, \beta \in \mathbb{R}$ mit Hilfe einer einfachen Fallunterscheidung, und zwar unter Verwendung der Rechenregeln in (109). $\square$

[17] Man benutze $\beta \setminus \alpha \neq \emptyset$, sowie Eigenschaft d) für β.

4.3.3 Die Dezimalbruchdarstellung reeller Zahlen – Kurzfassung

Von nun an gelten die Konventionen aus Notation 18. Wir betrachten also $\mathbb{Q}$ als Teilmenge von $\mathbb{R}$, verwenden die unindizierten Rechensymbole $\leq, \geq, +, -, \cdot$, und notieren das additive neutrale Element mit 0, sowie das multiplikative neutrale Element mit 1. Weiterhin notiert $|x|$ den Betrag von $x \in \mathbb{R}$; also (vgl. Bemerkung 25)

$$|x| = \begin{cases} x & \text{für } x \geq 0, \\ -x & \text{für } x < 0. \end{cases}$$

Reelle Zahlen lassen sich als unendliche Gleitkommazahlen (unendliche Dezimalbrüche) darstellen. Diese Darstellung ist (unter gewissen Umständen) uneindeutig – ähnlich wie bei den rationalen Zahlen, wo zwei Brüche die gleiche rationale Zahl darstellen, wenn diese durch kürzen ineinander übergehen. Um die Notation zu entlasten, diskutieren wir diesen Sachverhalt im Folgenden nur für den Fall nichtnegativer reeller Zahlen $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \mid x \geq 0\}$. Für negative reelle Zahlen folgen die entsprechenden Aussagen ganz analog (bzw. durch anwenden der additiven Inversion, aus den hergeleiteten Aussagen für den positiven Fall).

Ein unendlicher Dezimalbruch ist ein Ausdruck der Form

$$z = z_{\mathsf{ord}(z)} \dots z_0, z_1 z_2 z_3 \dots$$

mit $\mathbb{Z} \ni \mathsf{ord}(z) \leq 0$, $z_{\mathsf{ord}(z)} \neq 0$ und $z_j \in \{0, \dots, 9\}$ für alle $\mathsf{ord}(z) \leq j \in \mathbb{Z}$; also beispielsweise

$$\begin{aligned} 1234,85847\dots \qquad & (\mathsf{ord}(z) = -3) \\ 0,23968\dots \qquad & (\mathsf{ord}(z) = 0) \\ 0,00168\dots \qquad & (\mathsf{ord}(z) = 0). \end{aligned}$$

Wir bezeichnen die Menge dieser Ausdrücke mit $\mathbf{R}_+$.

Jedem Element in $\mathbf{R}_+$, kann nun eine reelle Zahl wie folgt zugeordnet werden:

$$\Lambda \colon \mathbf{R}_+ \to \mathbb{R}_{\geq 0}, \qquad z \mapsto \sup_{\mathbb{R}} \Big(\underbrace{\Big\{ \textstyle\sum_{j=\mathsf{ord}(z)}^{k} z_j \cdot 10^{-j} \;\Big|\; z_{\mathsf{ord}(z)} \leq k \in \mathbb{Z} \Big\}}_{=:\, M[z]} \Big).$$

Dies ist wegen der Vollständigkeit von $\mathbb{R}$ wohldefiniert, denn es gilt $M[z] \leq 10^{-\mathsf{ord}(z)+1}$.

Beispiel 27. *Sei $z = 4,8847\dots$. Dann gilt*

$$\Lambda(z) = \sup_{\mathbb{R}} \left(\left\{4, 4 + \tfrac{8}{10}, 4 + \tfrac{8}{10} + \tfrac{8}{100}, \dots\right\}\right) = \sup_{\mathbb{R}} \left(\left\{4, \tfrac{48}{10}, \tfrac{488}{100}, \tfrac{4887}{1000}, \dots\right\}\right).$$

Es lässt sich Folgendes zeigen:

- Λ ist surjektiv. (Jedes $x \in \mathbb{R}_{\geq 0}$ hat eine unendliche Dezimalbruchdarstellung.)

  Beweisidee: Sukzessive Teilung mit Rest durch Potenzen von 10.

  (Funktioniert auch für alle anderen $\mathbb{N} \ni b \geq 2$; nicht nur für $b = 10$. Für $b = 2$, erhält man beispielsweise einen unendlichen Dualbruch in der Form $10101,011110101\dots$.)
- Für jedes $x \in \mathbb{R}_{\geq 0} \setminus \mathbb{Q}$ (irrational), existiert **genau ein** $z \in \mathbf{R}_+$ mit $\Lambda(z) = x$.
- Es ist $z \in \mathbf{R}_+$ *periodisch* genau dann, wenn $\Lambda(z) \in \mathbb{Q}$ gilt.

  Beispiele für *periodische* Dezimalbrüche:

$$\begin{aligned} 1623,85385385\dots \\ 123,23232323\dots \\ 1,29999999\dots \\ 2,10000000\dots . \end{aligned}$$

- Ist $\Lambda^{-1}(x)$ mehrelementig ($x \in \mathbb{R}_{\geq 0}$), so gilt Folgendes: (Uneindeutigkeit)
  - $\Lambda^{-1}(x) = \{z, \tilde{z}\}$ ist zweielementig $\quad \wedge \quad x \in \mathbb{Q}$.
  - Es gilt $\mathsf{ord}(z) = m = \mathsf{ord}(\tilde{z})$; und es existiert $\ell \geq m$ mit:

$$\begin{aligned} &z_j = \tilde{z}_j && \forall\, m \leq j \leq \ell - 1 \\ &z_\ell = \tilde{z}_\ell + 1 && \text{und} \qquad z_p = 0 \ \wedge \ \tilde{z}_p = 9 \quad \text{für alle} \quad p \geq \ell + 1. \end{aligned}$$

  Beispielsweise gilt $\Lambda(z) = \Lambda(\tilde{z})$ für:

$$\begin{array}{lll} z = \underline{1}, 000000\ldots & z = 1234, 5\underline{8}00000\ldots & z = 1\underline{2}00, 0000000\ldots \\ \tilde{z} = \underline{0}, 999999\ldots & \tilde{z} = 1234, 5\underline{7}99999\ldots & z = 1\underline{1}99, 9999999\ldots \end{array}$$

  Man macht sich z.B. leicht klar, dass

$$\Lambda(z) = \sup_{\mathbb{R}}(\{1\}) = 1 = \sup_{\mathbb{R}}\left(\{0, \tfrac{9}{10}, \tfrac{99}{100}, \tfrac{999}{1000}, \tfrac{9999}{10000}, \ldots\}\right) = \Lambda(\tilde{z})$$

  gilt; denn 1 ist eine obere Schranke von $M[z]$; und für $1 - \varepsilon$ mit $\varepsilon > 0$ gilt dies nicht (Übung).

Mit Hilfe der unendlichen Dezimalbruchdarstellung, lässt sich nun folgender Satz beweisen:

Satz 11 (Cantor)**.** $\mathbb{R}$ *ist überabzählbar.*

Sei $\mathcal{I} \subseteq \mathbf{R}_+$ die Teilmenge aller $z \in \mathbf{R}_+$, sodass die folgenden Bedingungen gelten:

$$\mathsf{ord}(z) = 0, \qquad \exists\, j \geq 1\colon\ z_j \neq 0, \qquad \forall\, \ell \in \mathbb{N}\colon \exists\, j \geq \ell\colon\ z_j \neq 9. \tag{110}$$

Es folgt aus den oben diskutierten Eigenschaften, dass $\Lambda|_{\mathcal{I}}\colon \mathcal{I} \to (0, 1)$ bijektiv ist.

Beweisskizze zu Satz 11. Angenommen $\mathbb{R}$ ist abzählbar.

- Wegen Bemerkung 18.4), ist das Intervall $(0, 1) \subseteq \mathbb{R}$ abzählbar.
- Verkettung einer Surjektion $\mathbb{N} \to (0, 1)$ mit dem Inversen von $\Lambda|_{\mathcal{I}}$, liefert eine Aufzählung $(\delta_\ell)_{\ell\in\mathbb{N}}$ von $\mathcal{I}$ (also $\mathcal{I} = \{\delta_\ell \mid \ell \in \mathbb{N}\}$).

Wir definieren $z \in \mathbf{R}_+$ durch $\mathsf{ord}(z) := 0$, sowie

$$z_j := \begin{cases} 5 & \text{falls} \ \ (\delta_j)_j \neq 5 \\ 4 & \text{falls} \ \ (\delta_j)_j = 5 \end{cases}$$

für alle $j \in \mathbb{N}_{>0}$. Wir erhalten einen Widerspruch, denn:

- Per Konstruktion gilt $z \neq \delta_\ell$ für alle $\ell \in \mathbb{N}$; also gilt $z \notin \mathcal{I}$.
- Per Konstruktion erfüllt z die Bedingungen in (110); also gilt $z \in \mathcal{I}$. □

4.4 Der Euklidische Abstand und die Kreiszahl π

In diese Abschnitt behandeln wir den euklidischen Abstand und das euklidische Skalarprodukt, welche Sie bereits aus der Schule kennen. Wir geben weiterhin eine erste Anschauung für die Kreiszahl π.

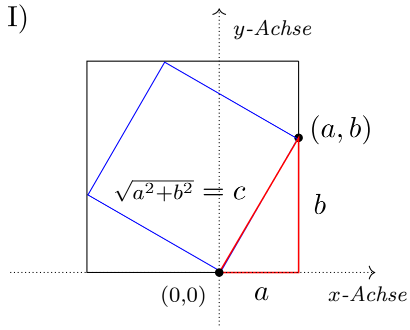
4.4.1 Der Euklidische Abstand

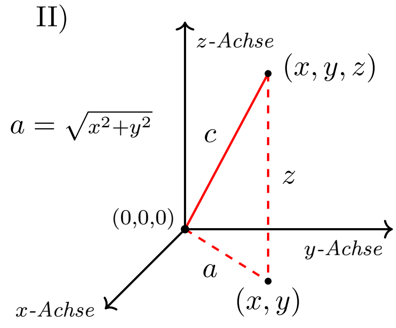
Gegeben $n \in \mathbb{N}_{>0}$, so definieren wir die euklidische Norm (die euklidischen Länge) $\|\cdot\|_{\text{euk}} \colon \mathbb{R}^n \to [0, \infty)$ auf dem n-dimensionalen euklidische Raum ($\mathbb{R}$-Vektorraum[18]) $\mathbb{R}^n$ durch

$$\|x\|_{\text{euk}} := \sqrt{(x_1)^2 + \ldots + (x_n)^2} \geq 0 \qquad \forall\, x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Dies ist wohldefiniert, da $x_i^2 \geq 0$ für alle $1 \leq i \leq n$; also auch $(x_1)^2 + \cdots + (x_n)^2 \geq 0$ gilt. Wir notieren den Nullvektor in $\mathbb{R}^n$ oft auch einfach durch 0. Offensichtlich gilt: $\|x\|_{\text{euk}} = 0 \;\Leftrightarrow\; x = 0$.

Bemerkung* 7 (Satz des Pythagoras)**.** *Wir wollen nun anschaulich machen, dass $\|x\|_{\text{euk}}$ für $x \in \mathbb{R}^n$ (mit $n = 1, 2, 3$), den Abstand des Punktes x zum Koordinatenursprung (die Länge der Verbindungsgeraden zwischen 0 und x) misst:*





- *Für $n = 1$, gilt $\|x\|_{\text{euk}} = \sqrt{x^2} = |x|$ für alle $x \in \mathbb{R}$.*
- *Sei $n = 2$, und $(a, b) \in \mathbb{R}^2$ (ein Punkt in der Tafelebene/Anschauungsebene). Bild I) skizziert den Fall $a, b \geq 0$ (die anderen Fälle folgen ganz analog). Der Flächeninhalt des äußeren Quadrates ist $F_1 = (a+b)^2$, die des inneren Quadrates $F_2 = c^2$. Der Flächeninhalt jedes der vier rechtwinkligen Dreiecke ist $F_3 = \frac{a \cdot b}{2}$. Daher gilt*

$$(a+b)^2 = F_1 = F_2 + 4 \cdot F_3 = c^2 + 2ab \quad \Longrightarrow \quad c^2 = a^2 + b^2 \quad \Longrightarrow \quad \|(a, b)\|_{\text{euk}} = c.$$

- *Sei $n = 3$, und $(x, y, z) \in \mathbb{R}^3$ ein Element (also ein Punkt im Anschauungsraum). Bild II) skizziert den Fall $x, y, z \geq 0$ (die anderen Fälle folgen ganz analog). Zweimaliges anwenden des zweidimensionalen Pythagoras, liefert*

$$c = \sqrt{a^2 + z^2} = \sqrt{\left(\sqrt{x^2 + y^2}\right)^2 + z^2} = \|(x, y, z)\|_{\text{euk}}.$$

Das Standardskalarprodukt auf $\mathbb{R}^n$ für $n \in \mathbb{N}_{>0}$, ist definiert durch

$$\langle x, y\rangle_{\text{euk}} := \textstyle\sum_{i=1}^{n} x_i \cdot y_i \qquad \forall\, x = (x_1, \ldots, x_n) \in \mathbb{R}^n,\; x = (y_1, \ldots, y_n) \in \mathbb{R}^n.$$

Man prüft leicht nach, dass für alle x, y, z und $\lambda \in \mathbb{R}$ gilt:

a) $\langle x, x\rangle_{\text{euk}} = \|x\|_{\text{euk}}^2 \geq 0$,

b) $\langle x, y\rangle_{\text{euk}} = \langle y, x\rangle_{\text{euk}}$,

c) $\langle x + y, z\rangle_{\text{euk}} = \langle x, z\rangle_{\text{euk}} + \langle y, z\rangle_{\text{euk}}$,

d) $\langle x, y + z\rangle_{\text{euk}} = \langle x, y\rangle_{\text{euk}} + \langle x, z\rangle_{\text{euk}}$,

e) $\langle \lambda \cdot x, y\rangle_{\text{euk}} = \lambda \cdot \langle x, y\rangle_{\text{euk}} = \langle x, \lambda \cdot y\rangle_{\text{euk}}$.

[18] Eine kurze Zusammenfassung über Vektorräume finden Sie im Zusatzabschnitt 4.6. Die dort dargelegten Sachverhalte sollten allerdings bereits aus der linearen Algebra bekannt sein.

Lemma 24 (Cauchy-Schwarzsche Ungleichung)**.** *Für $n \geq 1$ gilt*

$$|\langle x, y\rangle_{\text{euk}}| \leq \|x\|_{\text{euk}} \cdot \|y\|_{\text{euk}} \qquad \forall\, x, y \in \mathbb{R}^n.$$

$*$*Beweis.* Für $x, y \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$, erhalten wir aus den obigen Identitäten a)-e), dass gilt:

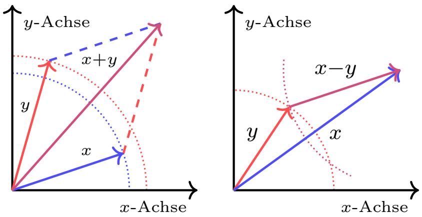
$$\begin{aligned} 0 &\leq \|x - \lambda \cdot y\|_{\text{euk}}^2 \\ &= \langle x - \lambda \cdot y, x - \lambda \cdot y\rangle_{\text{euk}} \\ &= \langle x, x\rangle_{\text{euk}} - 2\lambda \cdot \langle x, y\rangle_{\text{euk}} + \lambda^2 \cdot \langle y, y\rangle_{\text{euk}} \\ &= \|x\|_{\text{euk}}^2 - 2\lambda \cdot \langle x, y\rangle_{\text{euk}} + \lambda^2 \cdot \|y\|_{\text{euk}}^2. \end{aligned} \tag{111}$$

- Ist $y = 0$, so gilt die Aussage wegen $\langle x, y\rangle_{\text{euk}} = \langle x, 0 \cdot y\rangle_{\text{euk}} = 0 \cdot \langle x, y\rangle_{\text{euk}} = 0$ (vgl. e)).
- Ist $y \neq 0$, so gilt $\langle y, y\rangle_{\text{euk}} = \|y\|_{\text{euk}}^2 \neq 0$. Somit ist $\lambda := \langle x, y\rangle_{\text{euk}} \cdot \|y\|_{\text{euk}}^{-2}$ definiert, und wir erhalten

$$0 \overset{(111)}{\leq} \|x\|_{\text{euk}}^2 - 2 \cdot \langle x, y\rangle_{\text{euk}}^2 \cdot \|y\|_{\text{euk}}^{-2} + \langle x, y\rangle_{\text{euk}}^2 \cdot \|y\|_{\text{euk}}^{-2} = \|x\|_{\text{euk}}^2 - |\langle x, y\rangle_{\text{euk}}|^2 \cdot \|y\|_{\text{euk}}^{-2}.$$

Umstellen und Wurzelziehen liefert die Behauptung. □

Lemma 25. *Sei $n \in \mathbb{N}_{>0}$. Für $\lambda \in \mathbb{R}$ und $x, y \in \mathbb{R}^n$, gilt:*



1) $\|x\|_{\text{euk}} = 0 \quad \Leftrightarrow \quad x = 0$,

2) $\|\lambda \cdot x\|_{\text{euk}} = |\lambda| \cdot \|x\|_{\text{euk}}$,

3) $\|x + y\|_{\text{euk}} \leq \|x\|_{\text{euk}} + \|y\|_{\text{euk}}$, *(Dreiecksungleichung)*

4) $|\|x\|_{\text{euk}} - \|y\|_{\text{euk}}| \leq \|x - y\|_{\text{euk}}$. *(inverse Dreiecksungl.)*

Beweis. Wir zeigen nur die letzten beiden Punkte. Die ersten beiden Punkte sind Übung 53.

3) Es folgt mit Lemma 24 (vierter Schritt)

$$\begin{aligned} \|x + y\|_{\text{euk}}^2 &= \langle x + y, x + y\rangle_{\text{euk}} \\ &= \|x\|_{\text{euk}}^2 + 2 \cdot \langle x, y\rangle_{\text{euk}} + \|y\|_{\text{euk}}^2 \\ &\leq \|x\|_{\text{euk}}^2 + 2 \cdot |\langle x, y\rangle_{\text{euk}}| + \|y\|_{\text{euk}}^2 \\ &\leq \|x\|_{\text{euk}}^2 + 2 \cdot \|x\|_{\text{euk}} \cdot \|y\|_{\text{euk}} + \|y\|_{\text{euk}}^2 \\ &= (\|x\|_{\text{euk}} + \|y\|_{\text{euk}})^2. \end{aligned}$$

Die Behauptung folgt nun durch Wurzelziehen auf beiden Seiten (vgl. Übung 52.b)).

4) Wegen der Dreiecksungleichung 3), sowie 2) in der Form $\|-z\|_{\text{euk}} = \|z\|_{\text{euk}}$ für $z \in \mathbb{R}^n$, gilt

$$\begin{aligned} \|x\|_{\text{euk}} &= \|x - y + y\|_{\text{euk}} \leq \|x - y\|_{\text{euk}} + \|y\|_{\text{euk}} &&\Longrightarrow &&\|x\|_{\text{euk}} - \|y\|_{\text{euk}} \leq \|x - y\|_{\text{euk}} \\ \|y\|_{\text{euk}} &= \|y - x + x\|_{\text{euk}} \leq \|y - x\|_{\text{euk}} + \|x\|_{\text{euk}} &&\Longrightarrow &&\|y\|_{\text{euk}} - \|x\|_{\text{euk}} \leq \|x - y\|_{\text{euk}} \end{aligned}$$

für alle $x, y \in \mathbb{R}^n$. Ist $\|x\|_{\text{euk}} - \|y\|_{\text{euk}} \geq 0$, so folgt die Behauptung aus der ersten Zeile; und ist $\|x\|_{\text{euk}} - \|y\|_{\text{euk}} \leq 0$, so folgt die Behauptung aus der zweiten Zeile. □

Übung 53. *Zeigen Sie die ersten beiden Aussagen in Lemma 25.*

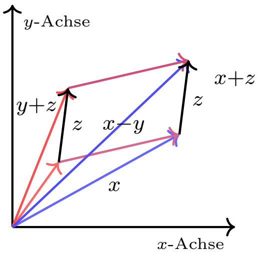
Hinweis: Zeigen und benutzen Sie $\sqrt{\mu^2} = |\mu|$ für alle $\mu \in \mathbb{R}$.

Der euklidische Abstand (die euklidische Metrik) auf $\mathbb{R}^n$ ($n \geq 1$) ist definiert durch

$$\mathrm{d}_{\text{euk}} \colon \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty), \qquad (x, y) \mapsto \|x - y\|_{\text{euk}}. \tag{112}$$

Aus Lemma 25 (bzw. direkt aus der Definition), erhält man unmittelbar:

Korollar 12. *Für $n \geq 1$, und $\lambda \in \mathbb{R}$ und $x, y, z \in \mathbb{R}^n$, gilt:*



1) $\mathrm{d}_{\mathrm{euk}}(x, y) = 0 \quad \Leftrightarrow \quad x = y$,

2) $\mathrm{d}_{\mathrm{euk}}(x, y) = \mathrm{d}_{\mathrm{euk}}(y, x)$, *(Symmetrie)*

3) $\mathrm{d}_{\mathrm{euk}}(x, z) \leq \mathrm{d}_{\mathrm{euk}}(x, y) + \mathrm{d}_{\mathrm{euk}}(y, z)$, *(Dreiecksungleichung)*

4) $|\mathrm{d}_{\mathrm{euk}}(x, y) - \mathrm{d}_{\mathrm{euk}}(y, z)| \leq \mathrm{d}_{\mathrm{euk}}(x, z)$, *(inverse Dreiecksungl.)*

5) $\mathrm{d}_{\mathrm{euk}}(z + x, z + y) = \mathrm{d}_{\mathrm{euk}}(x, y)$. *(Translationsinvarianz)*

$*$*Beweis.* 1) Wir erhalten für $x, y \in \mathbb{R}^n$ mit Lemma 25.1):

$$\mathrm{d}_{\mathrm{euk}}(x, y) = 0 \quad \Longleftrightarrow \quad \|x - y\|_{\mathrm{euk}} = 0 \quad \Longleftrightarrow \quad x = y.$$

2) Wir erhalten für $x, y \in \mathbb{R}^n$ mit Lemma 25.1):

$$\mathrm{d}_{\mathrm{euk}}(x, y) = \|x - y\|_{\mathrm{euk}} = \| - 1 \cdot (y - x)\|_{\mathrm{euk}} = | - 1| \cdot \|y - x\|_{\mathrm{euk}} = \|y - x\|_{\mathrm{euk}} = \mathrm{d}_{\mathrm{euk}}(y, x).$$

3) Wir erhalten für $x, y, z \in \mathbb{R}^n$ mit Lemma 25.3):

$$\mathrm{d}_{\mathrm{euk}}(x, z) = \|x - z\|_{\mathrm{euk}} = \|(x - y) + (y - z)\|_{\mathrm{euk}} \leq \|x - y\|_{\mathrm{euk}} + \|y - z\|_{\mathrm{euk}} = \mathrm{d}_{\mathrm{euk}}(x, y) + \mathrm{d}_{\mathrm{euk}}(y, z).$$

4) Für $x, y, z \in \mathbb{R}^n$, sei $v := x - y$ sowie $w := z - y$. Wir erhalten mit Lemma 25.4):

$$|\mathrm{d}_{\mathrm{euk}}(x, y) - \mathrm{d}_{\mathrm{euk}}(y, z)| = |\|v\|_{\mathrm{euk}} - \|w\|_{\mathrm{euk}}| \leq \|v - w\|_{\mathrm{euk}} = \|x - z\|_{\mathrm{euk}} = \mathrm{d}_{\mathrm{euk}}(x, z). \qquad \square$$

5) Wir erhalten für $x, y, z \in \mathbb{R}^n$:

$$\mathrm{d}_{\mathrm{euk}}(z + x, z + y) = \|(z + x) - (z + y)\|_{\mathrm{euk}} = \|(x - y)\|_{\mathrm{euk}} = \mathrm{d}_{\mathrm{euk}}(x, y).$$

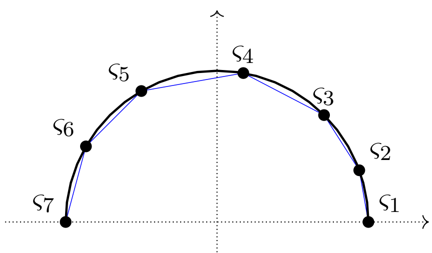
4.4.2 Die Kreiszahl π

Der Einheitskreis $\mathrm{K} \subseteq \mathbb{R}^2$ ist definiert als die Menge aller Punkte $\varsigma \in \mathbb{R}^2$ mit euklidischem Abstand 1 zum Ursprung $0 \in \mathbb{R}^2$, d.h.,

$$\mathrm{K} := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.$$

Es bezeichne $\mathcal{P}$ die Menge aller Längen von Polygonzügen von $(1, 0)$ nach $(-1, 0)$, entlang des oberen Einheitshalbkreisbogens $\mathrm{H} := \{(x, y) \in \mathrm{K} \mid y \geq 0\}$; also

$$\mathcal{P} := \left\{ \textstyle\sum_{j=1}^{n-1} \mathrm{d}_{\mathrm{euk}}(\varsigma_{j+1}, \varsigma_j) \;\middle|\; \varsigma_j = (x_j, y_j) \in \mathrm{H} \text{ für } j = 1, \ldots, n \text{ mit } -1 = x_n < \ldots < x_1 = 1 \right\}.$$



$(1, 0) = \varsigma_1 < \ldots < \varsigma_7 = (-1, 0)$

Man erhält $\mathcal{P} \leq 4$; und zwar mit Hilfe der Abschätzung

$$|a| + |b| = \sqrt{(|a| + |b|)^2} = \sqrt{|a|^2 + 2 \cdot |a| \cdot |b| + |b|^2} \geq \sqrt{a^2 + b^2} = \|(a, b)\|_{\text{euk}} \qquad \forall\, a, b \in \mathbb{R}. \tag{113}$$

In der Tat, folgt zunächst für $\varsigma_j = (x_j, y_j) \in \mathrm{H}$ für $j = 1, \ldots, n$ mit $-1 = x_n < \ldots < x_1 = 1$, dass

$$\textstyle\sum_{j=1}^{n-1} \mathrm{d}_{\text{euk}}(\varsigma_{j+1}, \varsigma_j) = \sum_{j=1}^{n-1} \|\varsigma_{j+1} - \varsigma_j\|_{\text{euk}} \leq \sum_{j=1}^{n-1} |x_{j+1} - x_j| + \sum_{j=1}^{n-1} |y_{j+1} - y_j|$$

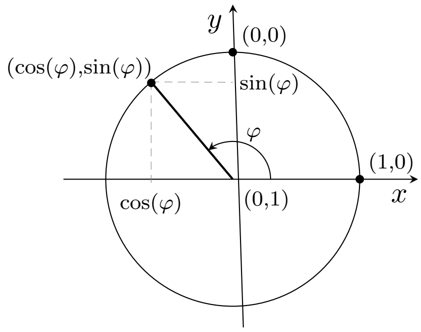
gilt. Hierbei ist $\sum_{j=1}^{n-1} |x_{j+1} - x_j| = \sum_{j=1}^{n-1} x_j - x_{j+1} = 2$; und ähnlich folgt $\sum_{j=1}^{n-1} |y_{j+1} - y_j| \leq 2$ durch geschicktes Aufteilen letzterer Summe.

Definition 30. *Die Zahl Kreiszahl π ist definiert als das Supremum $\pi := \sup_{\mathbb{R}}(\mathcal{P})$; anschaulich also als die Länge des oberen Halbkreisbogens. Die unendliche Dezimalbruchdarstellung von π lässt sich auf beliebig viele Nachkommastellen genau berechnen. Bis zur siebenten Nachkommastelle gilt $\pi = 3,1415926\ldots$. Es lässt sich zeigen, dass $\pi \notin \mathbb{Q}$ nicht rational (also irrational ist). Es lässt sich weiterhin zeigen, dass π auch nicht* algebraisch *ist; also keine Gleichung der Gestalt*

$$\pi^n \cdot q_n + \pi^{n-1} \cdot q_{n-1} + \ldots + q_1 \cdot \pi + q_0 = 0 \quad \textit{mit} \quad q_0, \ldots, q_n \in \mathbb{Q} \quad \textit{und} \quad n \geq 1$$

gilt. Reelle zahlen, die nicht algebraisch sind, heißen transzendent.

Bemerkung 29. *Wir werden die Kreiszahl π im Späteren mit Hilfe der Kosinusfunktion, noch auf eine andere Art definieren. Erst in der Analysis 2 Vorlesung werden wir dann im Kontext raumwertiger Kurven strikt einsehen können, dass diese beiden Definitionen wirklich übereinstimmen; bzw., dass die Definitionen von Sinus und Kosinus, die wir im Rahmen dieser Analysis 1 Vorlesung in Termen von Potenzreihen geben werden, mit den Definitionen übereinstimmen, die Sie aus der Schule kennen:*



4.5 Die Komplexen Zahlen

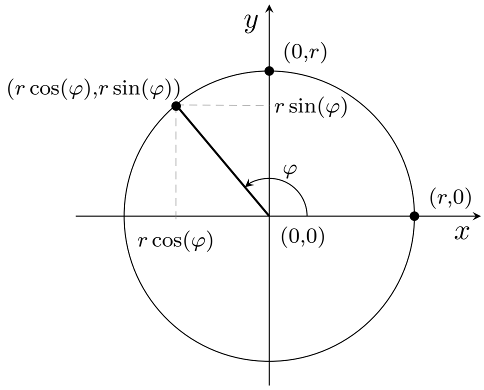
In diesem Abschnitt behandeln wir die komplexen Zahlen $\mathbb{C}$. Als Menge sind diese einfach gegeben durch den $\mathbb{R}^2$; also $\mathbb{C} = \mathbb{R}^2$. Die Addition und die additive Inversenbildung auf $\mathbb{C}$ ist komponentenweise definiert durch

$$\begin{aligned} z + \tilde{z} &:= (x + \tilde{x}, y + \tilde{y}) \\ -z &:= (-x, -y) \end{aligned} \tag{114}$$

für alle $z = (x, y) \in \mathbb{C}$ und $\tilde{z} = (\tilde{x}, \tilde{y}) \in \mathbb{C}$. Man sieht problemlos, dass $(\mathbb{C}, +, (0, 0))$ eine abelsche Gruppe ist. Die Multiplikation ist definiert durch (man schreibt auch $z\tilde{z}$ anstelle $z \cdot \tilde{z}$)

$$z \cdot \tilde{z} := (x \cdot \tilde{x} - y \cdot \tilde{y}, x \cdot \tilde{y} + \tilde{x} \cdot y) \qquad \forall\, z = (x, y) \in \mathbb{C},\ \tilde{z} = (\tilde{x}, \tilde{y}) \in \mathbb{C}. \tag{115}$$

Bemerkung 30. *Die Definition der Multiplikation wirkt möglicherweise zunächst fremdartig; hat aber einen geometrischen Hintergrund:*



Schreibt man nämlich mit Hilfe der trigonometrischen Funktionen Sinus und Kosinus[19]

$$z = (r \cdot \cos(\varphi), r \cdot \sin(\varphi)) \qquad \text{und} \qquad \tilde{z} = (\tilde{r} \cdot \cos(\tilde{\varphi}), \tilde{r} \cdot \sin(\tilde{\varphi}))$$

für $r, \tilde{r} > 0$ *und* $\phi, \tilde{\phi} \in \mathbb{R}$*, so folgt aus den Additionstheoremen (Schule):*

$$z \cdot \tilde{z} = (r\tilde{r} \cdot \cos(\varphi + \tilde{\varphi}), r\tilde{r} \cdot \sin(\varphi + \tilde{\varphi})).$$

Anschaulich bewirkt also die Multiplikation von $\tilde{z}$ *mit* z*, eine Rotation von* $\tilde{z}$ *um den Winkel* φ *entgegen dem Uhrzeigersinn, bei gleichzeitiger Streckung (Skalierung der Länge) um den Faktor* r*.*

Man rechnet leicht nach, dass $(\mathbb{C}, \cdot, (1, 0))$ ein abelscher Monoid ist (Assoziativität), und dass das Distributivgesetz (66) gilt. Um die multiplikative Inversion definieren zu können – um also $(\mathbb{C}_\times, \cdot, (1, 0))$ zu einer abelschen Gruppe, und somit $(\mathbb{C}, +, (0, 0), \cdot, (1, 0))$ zu einem Körper (kein angeordneter Körper) zu machen – ist noch etwas Vorarbeit von Nöten. Zunächst definieren wir den Realteil sowie den Imaginärteil einer komplexen Zahl durch

$$\mathrm{Re}(z) := x \qquad \text{sowie} \qquad \mathrm{Im}(z) := y \qquad \text{für} \qquad z = (x, y) \in \mathbb{C}.$$

Weiterhin definieren wir die komplexe Konjugation auf $\mathbb{C}$ durch

$$\overline{z} := (x, -y) \qquad \forall\, z = (x, y) \in \mathbb{C}. \tag{116}$$

Wir definieren den Betrag eine komplexen Zahl durch

$$|z| := \|z\|_{\text{euk}} = \sqrt{x^2 + y^2} \qquad \forall\, z = (x, y) \in \mathbb{C}. \tag{117}$$

Wir erhalten mit (115), dass $\qquad (\sqrt{\mathrm{Re}(z \cdot \overline{z})} = |z|)$

$$z \cdot \overline{z} = (x^2 + y^2, 0) = (|z|^2, 0) \qquad \forall\, z = (x, y) \in \mathbb{C}. \tag{118}$$

Das multiplikative Inverse eines Elementes $z = (x, y) \in \mathbb{C}_\times$ ist nun gegeben durch

$$z^{-1} = \overline{z} \cdot \big(\tfrac{1}{|z|^2}, 0\big) = \big(\tfrac{x}{|z|^2}, -\tfrac{y}{|z|^2}\big). \tag{119}$$

In der Tat folgt ja mit den bisher erhaltenen Rechenregeln

$$z \cdot z^{-1} = (z \cdot \overline{z}) \cdot \big(\tfrac{1}{|z|^2}, 0\big) = (|z|^2, 0) \cdot \big(\tfrac{1}{|z|^2}, 0\big) = (1, 0).$$

Wir notieren wieder $\frac{\tilde{z}}{z} := \tilde{z} \cdot z^{-1}$ für alle $\tilde{z} \in \mathbb{C}$ und $z \in \mathbb{C}_\times$. Wir haben nun nachgewiesen, dass $(\mathbb{C}, +, (0, 0), \cdot, (1, 0))$ ein Körper ist.

[19] Wir werden die Funktionen Sinus und Kosinus später noch im Rahmen der komplexen Exponentialreihe exakt definieren und untersuchen. An dieser Stelle gehen wir davon aus, dass Sie mit den grundlegenden Eigenschaften dieser Funktionen bereits aus der Schule vertraut sind.

Lemma 26. *Für alle $z, w \in \mathbb{C}$ gilt*

$$\overline{z} \cdot \overline{w} = \overline{zw} \qquad \text{und} \qquad |zw| = |z| \cdot |w|. \tag{120}$$

Beweis. Die linke Seite von (120) rechnet man problemlos nach (Übung). Für die rechte Seite von (120) bemerken wir zunächst, dass für alle $a, b \in \mathbb{R}$ gilt:

$$(a,0) \cdot (b,0) = (ab,0) \qquad \Longrightarrow \qquad \mathrm{Re}((a,0) \cdot (b,0)) = \mathrm{Re}((a,0)) \cdot \mathrm{Re}((b,0)). \tag{121}$$

Wir erhalten mit Hilfe der linken Seite von (120) (zweiter Schritt), unmittelbar aus Übung 52.a) (vierter Schritt):

$$\begin{aligned} |zw| &\overset{(118)}{=} \sqrt{\mathrm{Re}(zw \cdot \overline{zw})} \\ &= \sqrt{\mathrm{Re}(z\overline{z} \cdot w\overline{w})} \\ &\overset{(121)}{=} \sqrt{\mathrm{Re}(z\overline{z}) \cdot \mathrm{Re}(w\overline{w})} \\ &= \sqrt{\mathrm{Re}(z\overline{z})} \cdot \sqrt{\mathrm{Re}(w\overline{w})} \\ &\overset{(118)}{=} |z| \cdot |w|. \end{aligned}$$

Dies zeigt die Behauptung. □

Wir wollen nun die Notation entlasten, indem wir komplexe Zahlen mit Hilfe der imaginären Einheit darstellen. Hierfür bemerken wir, dass $\mathbb{R}$ in natürliche Weise in $\mathbb{C}$ eingebettet ist, nämlich vermöge der Körpereinbettung

$$\Phi\colon \mathbb{R} \to \mathbb{C}, \qquad x \mapsto (x,0). \tag{122}$$

Übung 54. *Machen Sie sich klar, dass* (122) *eine Körpereinbettung ist; also injektiv mit*

$$\Phi(x+y) = \Phi(x) + \Phi(y) \qquad \text{sowie} \qquad \Phi(x \cdot y) = \Phi(x) \cdot \Phi(y) \qquad \forall\, x, y \in \mathbb{R}.$$

Bemerkung 31. Φ *ist keine ordnungserhaltende Körpereinbettung, da* $\mathbb{C}$ *nicht zu einem angeordneten Körper gemacht werden kann:*

Beweis. Angenommen, es existiert eine Ordnung $\leq$ auf $\mathbb{C}$, sodass $(\mathbb{C}, \leq)$ ein angeordneter Körper ist. Wir erhalten aus Lemma 17.8) (erste Abschätzung), sowie Lemma 17.9) und Lemma 17.6) (letzte Abschätzung)

$$(0,0) < (0,1)^2 \overset{(115)}{=} (-1,0) \overset{(114)}{=} -(1,0) < (0,0), \tag{123}$$

also einen Widerspruch. □

Wir betrachten nun im Folgenden $\mathbb{R}$ als Teilmenge von $\mathbb{C}$, vermöge der Identifikation von $\mathbb{R}$ mit der Teilmenge $\mathbb{R} \times \{0\} = \Phi(\mathbb{R}) \subseteq \mathbb{C}$. Konkret identifizieren wir dann also das Element (Symbol) $x \in \mathbb{R}$ mit dem Element (Symbol) $(x,0)$ in $\mathbb{C}$, sodass dann insbesondere $1 = (1,0)$ und $0 = (0,0)$ gilt. Weiterhin definieren wir $\mathrm{i} := (0,1)$, womit nun (vgl. (123) in Bemerkung 31)

$$\mathrm{i}^2 = (0,1)^2 = (-1,0) = -(1,0) = -1$$

gilt (genau so, wie Sie das bereits aus der Schule kennen). Eine komplexe Zahl $z = (x,y) \in \mathbb{C}$ wird dann üblicherweise in der Form $z = x + \mathrm{i}y$ dargestellt[20] – nämlich als Vereinfachung der formell korrekteren Schreibweise

$$(x,0) \cdot 1 + (y,0) \cdot \mathrm{i} = (x,0) \cdot (1,0) + (y,0) \cdot (0,1) = (x,0) + (0,y) = (x,y).$$

[20]Oder $x + \mathrm{i} \cdot y$ oder $x + y\mathrm{i}$ oder $x + y \cdot \mathrm{i}$.

Offensichtlich gilt $\overline{x + \mathrm{i}y} = x - \mathrm{i}y$; und daher

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

für alle $z \in \mathbb{C}$. Wir bemerken weiterhin:

- Es liest sich (123) nun einfach in der Form

$$0 < \mathrm{i}^2 = -1 < 0.$$

- Naives anwenden der üblichen Rechenregeln (Distributivgesetz), liefert den korrekten Ausdruck für das Produkt zweier komplexer Zahlen (vgl. (115)):

$$(x + \mathrm{i}y) \cdot (\tilde{x} + \mathrm{i}\tilde{y}) = x\tilde{x} + x\tilde{y} \cdot \mathrm{i} + y\tilde{x} \cdot \mathrm{i} + y\tilde{y} \cdot \mathrm{i}^2 = (x\tilde{x} - y\tilde{y}) + (x\tilde{y} + \tilde{x}y)\mathrm{i}.$$

4.6 Vektorräume*

Vektorräume sind Kerngegenstand der linearen Algebra, und werden dort auch im Detail behandelt. Wir werden daher nur die für uns unmittelbar relevanten Sachverhalte darlegen.

Definition 31. *Sei $(\mathbb{K}, +, 0_{\mathbb{K}}, \cdot, 1_{\mathbb{K}})$ ein Körper. Ein $\mathbb{K}$-Vektorraum ist ein Tupel $(V, \oplus, 0_V, \odot)$, bestehend aus einer (nichtleeren) Menge V, einem ausgezeichneten Element $0_V \in V$ (Nullvektor), einer Verknüpfung $\oplus \colon V \times V \to V$ (Vektoraddition), und einer Abbildungen (Skalarmultiplikation)*

$$\odot \colon \mathbb{K} \times V \to V, \qquad (\lambda, v) \mapsto \lambda \odot v,$$

sodass $(V, \oplus, 0_V)$ eine abelsche Gruppe ist, und zusätzlich gilt:

1) $\lambda \odot (v \oplus w) = (\lambda \odot v) \oplus (\lambda \odot w)$ *für alle* $\lambda \in \mathbb{K}$ *und* $v, w \in V$, *(Distributivgesetz)*
2) $(\lambda + \mu) \odot v = (\lambda \odot v) \oplus (\mu \odot v)$ *für alle* $\lambda, \mu \in \mathbb{K}$ *und* $v \in V$, *(Distributivgesetz)*
3) $(\lambda \cdot \mu) \odot v = \lambda \odot (\mu \odot v)$ *für alle* $\lambda, \mu \in \mathbb{K}$ *und* $v \in V$, *(Assoziativgesetz)*
4) $1_{\mathbb{K}} \cdot v = v$ *für alle* $v \in V$. *(Neutralität des Einselements)*

Die Elemente von V werden Vektoren genannt.

Bemerkung 32. *In der Situation von Definition 31 gelten automatisch die folgenden Rechenregeln für alle $\lambda \in \mathbb{K}$ und $v \in V$:*

i) $0_{\mathbb{K}} \odot v = 0_V$, *($0_{\mathbb{K}} = 0_{\mathbb{K}} + 0_{\mathbb{K}}$)*
ii) $\lambda \odot 0_V = 0_V$, *($0_V = 0_V + 0_V$)*
iii) $\lambda \odot (-v) = (-\lambda) \odot v = -(\lambda \odot v)$, *(wegen $0_V = 0_V + 0_V$)*
iv) $\lambda \odot v = 0_V \quad \Longrightarrow \quad \lambda = 0_{\mathbb{K}} \quad \vee \quad v = 0_V$.

Insbesondere folgt hiermit, dass das additive Inverse $-v \in V$ von $v \in V$ automatisch mit dem Element $(-1_{\mathbb{K}}) \odot v \in V$ übereinstimmt.

Notation 19. *Sind Uneindeutigkeiten ausgeschlossen, so spricht man vereinfachend von einem $\mathbb{K}$-Vektorraum V anstatt $(V, \oplus, 0_V, \odot)$ zu notieren. Man notiert dann auch $\odot$ mit $\cdot$, und $\oplus$ mit $+$, benutzt also die selben Symbole wie für die Rechenoperationen auf $\mathbb{K}$. Damit es hierbei nicht zu Uneindeutigkeiten kommt ist stets darauf zu achten, aus dem Kontext heraus klarzustellen, welcher Menge die zu verknüpfenden Elemente angehören.*

Sei $\mathbb{K}$ ein Körper, und V ein $\mathbb{K}$-Vektorraum.

- Eine Basis von V ist eine Teilmenge $\mathfrak{B} \subseteq V$, sodass jedes Element von V eindeutig als endliche Linearkombination von Elementen in $\mathfrak{B}$ geschrieben werden kann. In Formeln bedeutet dies:

  Ist $v \in V$ gegeben, so existiert eine eindeutige nichtleere endliche Teilmenge

  $$\{(\lambda_1, b_1), \ldots, (\lambda_n, b_n)\} \subseteq \mathbb{K} \times \mathfrak{B} \quad \text{mit} \quad n \in \mathbb{N}, \quad \text{sodass} \quad v = \sum_{i=1}^{n} \lambda_i \cdot b_i \quad \text{gilt.} \tag{124}$$

  - Gilt also $v = \sum_{i=1}^{m} \mu_i \cdot b_i$ für $\{(\mu_1, b_1), \ldots, (\mu_m, b_m)\} \subseteq \mathbb{K}$ mit $m \geq 1$, so folgt $m = n$, und es existiert eine Permutation τ auf $\{1, \ldots, n\}$ mit $(\mu_i, a_i) = (\lambda_{\tau(i)}, b_{\tau(i)})$ für $i = 1, \ldots, n$.
  - Es kann in (124) auch der Fall $n = 0$ auftreten, nämlich wenn $v = 0_V$ gilt. Nach unserer Konvention für die leere Summe (in abelschen Gruppen) ist dann in der Tat $\sum_{i=1}^{n} \lambda_i \cdot b_i = 0_V$.

- V heißt endlichdimensional, wenn V eine endliche Basis $\mathfrak{B}$ besitzt ($\mathfrak{B}$ besteht aus endlich vielen Elementen). Die Anzahl der Elemente von $\mathfrak{B}$ wird als die Dimension von V bezeichnet, und mit $\dim(V)$ notiert. Die Dimension von V ist in dieser Form wohldefiniert, denn ist $\tilde{\mathfrak{B}}$ eine weitere Basis von V, so folgt bereits (lineare Algebra) $|\tilde{\mathfrak{B}}| = |\mathfrak{B}|$. Ein Vektorraum heißt unendlichdimensional genau dann, wenn er nicht endlichdimensional ist. Das sogenannte Lemma von Zorn impliziert, dass jeder Vektorraum eine Basis besitzt.

Beispiel 28. *Zu einem Körper $(\mathbb{K}, +, 0_{\mathbb{K}}, \cdot, 1_{\mathbb{K}})$ existieren automatisch auch immer die folgenden $\mathbb{K}$-Vektorräume:*

a) Der Nullvektorraum $\{0_{\mathbb{K}}\}$ ist ein $\mathbb{K}$-Vektorraum mit Basis $\emptyset$; ist also Nulldimensional.

b) Für $n \geq 1$ ist das n-fache Produkt $\mathbb{K}^n$ ein $\mathbb{K}$-Vektorraum, vermöge den Operationen

$$\begin{aligned} +\colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}^n, &\qquad ((x_1, \ldots, x_n), (y_1, \ldots, y_n)) \mapsto (x_1 + y_1, \ldots, x_n + y_n) \\ \cdot : \mathbb{K} \;\; \times \mathbb{K}^n \to \mathbb{K}^n, &\qquad (\lambda, (x_1, \ldots, x_n)) \mapsto (\lambda \cdot x_1, \ldots, \lambda \cdot x_n). \end{aligned}$$

Der Nullvektor ist gegeben durch $(0_{\mathbb{K}}, \ldots, 0_{\mathbb{K}}) \in \mathbb{K}^n$. Die kanonische (natürliche) Basis von V ist $\{e_1, \ldots, e_n\}$ mit $e_i = ((e_i)_1, \ldots, (e_i)_n) \in \mathbb{K}^n$, für $1 \leq i \leq n$, das Element

$$(e_i)_j := \begin{cases} 0_{\mathbb{K}} & \text{für } i \neq j \\ 1_{\mathbb{K}} & \text{für } i = j. \end{cases}$$

Die Dimension von $\mathbb{K}^n$ ist also n. Beispielsweise ist $\mathbb{K}$ ein eindimensionaler $\mathbb{K}$-Vektorraum.

c) Sind $V_1, \ldots, V_n$ mit $n \geq 1$ $\mathbb{K}$-Vektorräume, so ist das n-fache Produkt $V := V_1 \times \cdots \times V_n$ ein $\mathbb{K}$-Vektorraum, vermöge den Operationen

$$\begin{aligned} +\colon V \times V \to V, &\qquad ((v_1, \ldots, v_n), (w_1, \ldots, w_n)) \mapsto (v_1 + w_1, \ldots, v_n + w_n) \\ \cdot : \mathbb{K} \times V \to V, &\qquad (\lambda, (v_1, \ldots, v_n)) \mapsto (\lambda \cdot v_1, \ldots, \lambda \cdot v_n). \end{aligned}$$

Der Nullvektor ist gegeben durch $(0_{V_1}, \ldots, 0_{V_n}) \in V$. Ist nun $\mathfrak{B}_j$ eine Basis von V_j für $j = 1, \ldots, n$, dann ist $\bigcup_{1 \leq j \leq n} \tilde{\mathfrak{B}}_j$ eine Basis von V mit

$$\tilde{\mathfrak{B}}_j := \{0_{V_1}\} \times \cdots \times \{0_{V_{j-1}}\} \times \mathfrak{B}_j \times \{0_{V_{j+1}}\} \times \cdots \times \{0_{V_n}\} \qquad \forall\, 1 \leq j \leq n.$$

Sind $V_1, \ldots, V_n$ endlichdimensional, so ist V endlichdimensional, mit $\dim(V) = \sum_{j=1}^{n} \dim(V_j)$. Für $V_1, \ldots, V_n = \mathbb{K}$ erhalten wir die Konstruktion aus Punkt b) zurück.

d) Für jede Menge $X \neq \emptyset$ und jeden $\mathbb{K}$-Vektorraum V ist $\mathsf{Abb}(X, V)$ ein $\mathbb{K}$-Vektorraum, vermöge den (punktweisen) Operationen

$$\begin{aligned} +\colon \mathsf{Abb}(X, V) \times \mathsf{Abb}(X, V) &\to \mathsf{Abb}(X, V) \\ (f, g) &\mapsto \big[f + g\colon X \to V, \quad x \mapsto f(x) + g(x)\big], \\ \cdot : \mathbb{K} \;\; \times \mathsf{Abb}(X, V) &\to \mathsf{Abb}(X, V) \\ (\lambda, f) &\mapsto \big[\lambda \cdot f\colon X \to V, \quad x \mapsto \lambda \cdot f(x)\big]. \end{aligned}$$

Der Nullvektor ist gegeben durch die konstante Funktion 0_V auf X, also $f[0_V]\colon X \ni x \mapsto 0_V \in \mathbb{K}$. Ist X nicht endlich und $\dim(V) \geq 1$, so ist $\mathsf{Abb}(X,V)$ unendlichdimensional. Ist X endlich mit $n := |X| \geq 1$, so erhalten wir die Konstruktion aus Punkt c) für den Spezialfall $V_1,\dots,V_n = V$ zurück – also $\mathsf{Abb}(X,V) = V^n$ (vgl. Bemerkung 17.1)).

5 Metrische und Normierte Räume

Im Folgenden sei immer $\mathbb{K} \in \{\mathbb{Q},\mathbb{R},\mathbb{C}\}$.

5.1 Normierte Vektorräume

Definition 32. *Ein normierter Raum (Vektorraum) ist ein Tupel $(V, \|\cdot\|)$, bestehend aus einem $\mathbb{K}$-Vektorraum V und einer Abbildung $\|\cdot\|\colon V \to [0,\infty)$, sodass für alle $x,y \in V$ und $\lambda \in \mathbb{K}$ gilt:*

V1) $\|x\| = 0 \quad \Leftrightarrow \quad x = 0$,

V2) $\|\lambda \cdot x\| = |\lambda| \cdot \|x\|$,

V3) $\|x+y\| \leq \|x\| + \|y\|$. *(Dreiecksungleichung)*

Es wird dann $\|\cdot\|$ als Norm auf V bezeichnet; und es gilt automatisch für alle $x,y \in V$:

V4) $|\|x\| - \|y\|| \leq \|x - y\|$. *(inverse Dreiecksungleichung)*

Beweis von V4)*:* Der Beweis ist analog zu dem Beweis von Lemma 25.4): Wegen V3), sowie V2) in der Form $\|-z\| = \|z\|$ für $z \in V$ (vgl. Bemerkung 32), gilt

$$\begin{aligned} \|x\| = \|x-y+y\| \leq \|x-y\| + \|y\| \qquad &\Longrightarrow \qquad \|x\| - \|y\| \leq \|x-y\| \\ \|y\| = \|y-x+x\| \leq \|y-x\| + \|x\| \qquad &\Longrightarrow \qquad \|y\| - \|x\| \leq \|x-y\| \end{aligned}$$

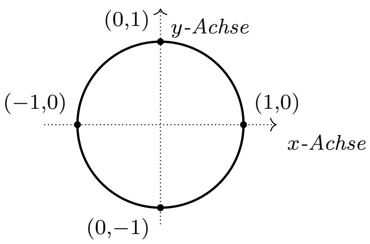
für alle $x,y \in V$. Ist $\|x\|-\|y\| \geq 0$, so folgt die Abschätzung aus der ersten Zeile; und ist $\|x\|-\|y\| \leq 0$, so folgt die Abschätzung aus der zweiten Zeile. □

Beispiel 29. *Wir betrachten den $\mathbb{R}$-Vektorraum $V = \mathbb{R}^n$, mit $n \in \mathbb{N}_{>0}$.*

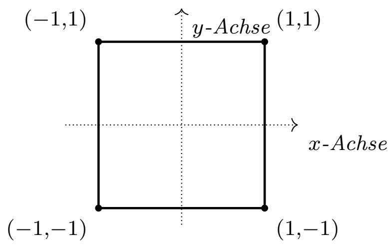
- *Es ist $(\mathbb{R}^n, \|\cdot\|_{\mathrm{euk}})$ ein normierter Raum, nach den ersten drei Punkten in Lemma 25.*
- *Es ist $(\mathbb{R}^n, \|\cdot\|_{\max})$ ein normierter Raum, mit der Maximumsnorm (Übung 55)*

$$\|x\|_{\max} := \max(|x_1|,\dots,|x_n|) \qquad \forall\, x = (x_1,\dots,x_n) \in \mathbb{R}^n. \tag{125}$$

Die beiden Grafiken, zeigen die beiden Niveaumengen $(\|\cdot\|_{\mathrm{euk}})^{-1}(1)$ (links) und $(\|\cdot\|_{\max})^{-1}(1)$ (rechts), im Falle $n = 2$:



$\{(x,y)\in\mathbb{R}^2 \mid \|(x,y)\|_{\mathrm{euk}}=1\}$



$\{(x,y)\in\mathbb{R}^2 \mid \|(x,y)\|_{\max}=1\}$

Generell erhält man die folgenden Abschätzungen für alle $n \in \mathbb{N}_{>0}$ und $x \in \mathbb{R}^n$:

$$\|x\|_{\max} \leq \|x\|_{\mathrm{euk}} \qquad \textit{sowie} \qquad \|x\|_{\mathrm{euk}} \leq n \cdot \|x\|_{\max}. \tag{126}$$

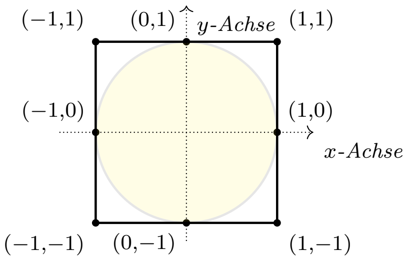
Beweis der Äquivalenz: Sei $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ mit $n \in \mathbb{N}_{>0}$ vorgegeben.

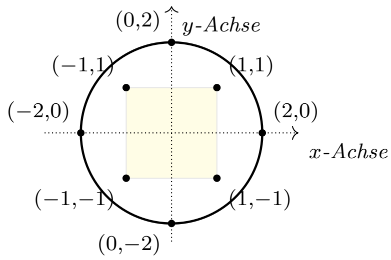
- Offensichtlich gilt $x_1^2 + \ldots + x_n^2 \leq (|x_1| + \ldots + |x_n|)^2$; also
$$\|x\|_{\text{euk}}^2 \leq (|x_1| + \ldots + |x_n|)^2 \leq (n \cdot \|x\|_{\max})^2 \qquad \Longrightarrow \qquad \|x\|_{\text{euk}} \leq n \cdot \|x\|_{\max}.$$
- Offensichtlich gilt $|x_i| = \sqrt{x_i^2} \leq \|x\|_{\text{euk}}$ für $i = 1, \ldots, n$; also auch $\|x\|_{\max} \leq \|x\|_{\text{euk}}$. $\square$

Anschaulich bedeutet für $n = 2$:

– *die linke Seite von* (126)*, das Enthaltensein der Einheitskreisscheibe, in dem bei* 0 *zentrierten Quadrat mit Seitenlänge* 2 *– siehe linkes Bild.*

– *die rechte Seite von* (126)*, das Enthaltensein des bei* 0 *zentrierten Quadrates mit Seitenlänge* 2*, in der Kreisscheibe mit dem Radius* 2 *– siehe rechtes Bild.*





Übung 55. *Zeigen Sie, dass* (125) *eine Norm auf* $\mathbb{R}^n$ *definiert.*

Hinweis: Gegebenenfalls sind die Eigenschaften (86) *der Betragsfunktion gewinnbringend einsetzbar.*

5.2 Metrische Räume

In Verallgemeinerung von normierten Vektorräumen, haben wir das folgende Konzept:

Definition 33. *Ein metrischer Raum ist ein Tupel (X, d), bestehend aus einer nichtleeren Menge X, und einer Abbildung $\mathrm{d}\colon X \times X \to [0, \infty)$, sodass für alle $x, y \in X$ gilt:*

M1) $\mathrm{d}(x, y) = 0 \quad \Leftrightarrow \quad x = y$,

M2) $\mathrm{d}(x, y) = \mathrm{d}(y, x)$, *(Symmetrie)*

M3) $\mathrm{d}(x, z) \leq \mathrm{d}(x, y) + \mathrm{d}(y, z)$, *(Dreiecksungleichung)*

Es wird dann d *als Metrik auf X bezeichnet; und es gilt automatisch für alle $x, y, z \in X$:*

M4) $|\mathrm{d}(x, y) - \mathrm{d}(y, z)| \leq \mathrm{d}(x, z)$, *(inverse Dreiecksungleichung)*

Beweis der Eigenschaft M4)*:* Für alle $x, y, z \in X$, gilt
$$\begin{aligned} \mathrm{d}(x, y) &\overset{\text{M3)}}{\leq} \mathrm{d}(x, z) + \mathrm{d}(z, y) &\overset{\text{M2)}}{\Longrightarrow}& \qquad \mathrm{d}(x, y) - \mathrm{d}(y, z) \leq \mathrm{d}(x, z) \\ \mathrm{d}(y, z) &\overset{\text{M3)}}{\leq} \mathrm{d}(y, x) + \mathrm{d}(x, z) &\overset{\text{M2)}}{\Longrightarrow}& \qquad \mathrm{d}(y, z) - \mathrm{d}(x, y) \leq \mathrm{d}(x, z); \end{aligned}$$
also $|\mathrm{d}(x, y) - \mathrm{d}(y, z)| \leq \mathrm{d}(x, z)$. $\square$

Bemerkung 33. *Ist $(V, \|\cdot\|)$ ein normierter Raum, so ist*
$$\mathrm{d}_{\|\cdot\|}\colon V \times V \to [0, \infty), \qquad (x, y) \mapsto \|x - y\| \tag{127}$$

eine translationsinvariante Metrik auf $X = V$ (Übung).

Ein Beispiel hierfür ist die euklidische Metrik (12) *aus Abschnitt 4.4.1 – siehe Korollar 12. Insbesondere sind also auch $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ metrische Räume, vermöge der zugehörigen Betragsmetriken (vgl.* (117) *für $\mathbb{C}$)*

$$\mathrm{d}_{|\cdot|} \colon \mathbb{K} \times \mathbb{K} \to [0, \infty), \qquad (x, y) \mapsto |x - y|. \tag{128}$$

Terminologie 11 (Metrischer Unterraum)**.** *Sei (X, d) ein metrischer Raum; sowie $Y \subseteq X$ eine Teilmenge. Dann wird Y zu einem metrischen Raum, vermöge der eingeschränkten Metrik*

$$\mathrm{d}_Y := \mathrm{d}|_{Y \times Y} \colon Y \times Y \to [0, \infty), \qquad (x, y) \mapsto \mathrm{d}(x, y). \tag{129}$$

Im Folgenden werden wir uns Teilmengen von metrischen Räumen stillschweigend immer mit der eingeschränkten Metrik ausgestattet denken. Insbesondere betrifft dies Teilmengen von $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, spezieller beispielsweise die Teilmenge $\mathbb{K}_\times$ aller invertierbaren Elemente von $\mathbb{K}$. Derartige Teilmengen sind dann also mit der entsprechend eingeschränkten Betragsmetrik (128) *ausgestattet.*

Definition 34. *Sei (X, d) ein metrischer Raum.*

1) *Für $x \in X$ und $\varepsilon > 0$, definieren wir den offenen (metrische) ε-Ball um x durch*

$$\mathsf{B}_\varepsilon(x) := \{y \in X \mid \mathrm{d}(x, y) < \varepsilon\}. \tag{130}$$

(Wegen M1) *gilt $x \in \mathsf{B}_\varepsilon(x)$.)*

2) *Für $x \in X$ und $\varepsilon \geq 0$, definieren wir den abgeschlossenen (metrische) ε-Ball um x durch*

$$\overline{\mathsf{B}}_\varepsilon(x) := \{y \in X \mid \mathrm{d}(x, y) \leq \varepsilon\}.$$

(Wegen M1) *gilt $x \in \overline{\mathsf{B}}_\varepsilon(x)$, sowie $\overline{\mathsf{B}}_0(x) = \{x\}$.)*

3) *Eine Teilmenge $U \subseteq X$ heißt Umgebung von $x \in X$, wenn ein $\varepsilon > 0$ existiert mit $\mathsf{B}_\varepsilon(x) \subseteq U$.*

4) *Eine Teilmenge $O \subseteq X$ heißt offen, wenn sie eine Umgebung aller ihrer Punkte ist, d.h.,*

$$\forall\, x \in O \colon \exists\, \varepsilon[x] > 0 \colon \mathsf{B}_{\varepsilon[x]}(x) \subseteq O. \tag{131}$$

Offensichtlich gilt dann

$$O = \textstyle\bigcup_{x \in O} \mathsf{B}_{\varepsilon[x]}(x). \tag{132}$$

5) *Eine Teilmenge $A \subseteq X$ heißt abgeschlossen, wenn $X \setminus A$ offen ist.*

Notation 20. *Sei (X, d) ein metrischer Raum.*

- *Wir notieren wir die Menge aller offenen Teilmengen von X mit $\mathcal{O}(X)$.*
- *Wir notieren wir die Menge aller abgeschlossenen Teilmengen von X mit $\mathcal{A}(X)$.*

Notation 21 (Erweiterte Notation)**.** *Haben wir es im Späteren mit mehreren metrischen Räumen gleichzeitig zu arbeiten, so schreiben wir – für einen metrischen Raum (X, d) – die offenen bzw. abgeschlossenen metrischen Bälle auch in der Form*

$$\begin{aligned} \mathsf{B}[\mathrm{d}]_\varepsilon(x) &= \{y \in X \mid \mathrm{d}(x, y) < \varepsilon\} \qquad && \forall\, \varepsilon > 0,\ x \in X \\ \overline{\mathsf{B}}[\mathrm{d}]_\varepsilon(x) &= \{y \in X \mid \mathrm{d}(x, y) < \varepsilon\} \qquad && \forall\, \varepsilon \geq 0,\ x \in X. \end{aligned} \tag{133}$$

Weiterhin schreiben wir $\mathcal{O}[\mathrm{d}](X)$ anstelle $\mathcal{O}(X)$, sowie $\mathcal{A}[\mathrm{d}](X)$ anstelle $\mathcal{A}(X)$ (vgl. Notation 20).

Notation 22. *Ist $(V, \|\cdot\|)$ ein normierter Raum, so verstehen wir Definition 34 im Sinne des metrischen Raumes $(V, \mathrm{d}_{\|\cdot\|})$, wobei $\mathrm{d}_{\|\cdot\|}$ wie in (127) in Bemerkung 33 definiert ist. Es ist dann also*

$$\begin{aligned} \mathsf{B}_\varepsilon(x) &= \{y \in V \mid \|x-y\| < \varepsilon\} && \forall\, \varepsilon > 0,\ x \in X \\ \overline{\mathsf{B}}_\varepsilon(x) &= \{y \in V \mid \|x-y\| \leq \varepsilon\} && \forall\, \varepsilon \geq 0,\ x \in X. \end{aligned}$$

Beispiel 30 (zu Notation 22)**.**

- *In dem normierten Raum $(\mathbb{R}, |\cdot|)$ gilt*

$$\begin{aligned} \mathsf{B}_\varepsilon(x) &= (x-\varepsilon, x+\varepsilon) && \forall\, \varepsilon > 0,\ x \in X \\ \overline{\mathsf{B}}_\varepsilon(x) &= [x-\varepsilon, x+\varepsilon] && \forall\, \varepsilon \geq 0,\ x \in X. \end{aligned} \tag{134}$$

- *Wir betrachten den normierten Raum $(\mathbb{R}^2, \|\cdot\|_{\mathrm{euk}})$ (also $(\mathbb{C}, |\cdot|)$); und fixieren $\varepsilon > 0$ sowie $x \in \mathbb{R}^2$:*
  - *$\mathsf{B}_\varepsilon(x)$ ist die, bei x zentrierte Einheitskreisscheibe mit Radius ε, und zwar ohne den Rand $R := \{x \in \mathbb{R}^2 \mid \|x\|_{\mathrm{euk}} = \varepsilon\}$.*
  - *$\overline{\mathsf{B}}_\varepsilon(x)$ ist die, bei x zentrierte Einheitskreisscheibe (mit Rand), mit Radius ε.*

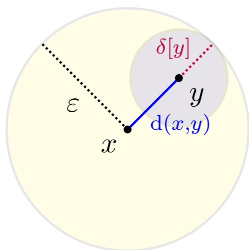
Wie die Bezeichnung „offener metrischer Ball“ in Definition 34.2) bereits erahnen lässt, gilt der folgenden Sachverhalt:

Lemma 27. *Sei (X, d) ein metrischer Raum. Es gilt $\mathsf{B}_\varepsilon(x) \in \mathcal{O}(X)$ für alle $x \in X$ und $\varepsilon > 0$.*

Beweis. Seien $x \in X$ und $\varepsilon > 0$ fixiert. Für $y \in \mathsf{B}_\varepsilon(x)$ setzen wir $\delta[y] := \varepsilon - \mathrm{d}(x,y)$, und erhalten

$$\begin{aligned} z \in \mathsf{B}_{\delta[y]}(y) \qquad &\Longrightarrow \qquad \mathrm{d}(y,z) < \delta[y] \\ &\Longrightarrow \qquad \mathrm{d}(x,z) \leq \mathrm{d}(x,y) + \mathrm{d}(y,z) < \mathrm{d}(x,y) + \varepsilon - \mathrm{d}(x,y) = \varepsilon \\ &\Longrightarrow \qquad z \in \mathsf{B}_\varepsilon(x), \end{aligned}$$

also $\mathsf{B}_{\delta[y]}(y) \subseteq \mathsf{B}_\varepsilon(x)$.



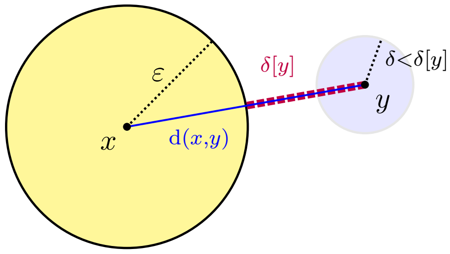
$(X, \mathrm{d}) = (\mathbb{R}^2, \mathrm{d}_{\mathrm{euk}})$

Dies zeigt, dass der metrische Ball $\mathsf{B}_\varepsilon(x)$ eine Umgebung aller seiner Punkte ist, also offen nach Definition 34.4). $\square$

Ebenfalls sind die abgeschlossenen metrischen Bälle abgeschlossen:

Lemma 28. *Seien (X, d) ein metrischer Raum. Es gilt $\overline{\mathsf{B}}_\varepsilon(x) \in \mathcal{A}(X)$ für alle $x \in X$ und $\varepsilon \geq 0$.*

Beweis. Seien $x \in X$ und $\varepsilon \geq 0$ vorgegeben. Wir müssen zeigen, dass $O := X \setminus \overline{\mathsf{B}}_\varepsilon(x)$ offen ist; dass also für jedes $y \in O$ ein $\delta[y] > 0$ existiert, mit $\mathsf{B}_{\delta[y]}(y) \subseteq O$:



$$(X, \mathrm{d}) = (\mathbb{R}^2, \mathrm{d}_{\mathrm{euk}})$$

Sei hierfür $y \in O = X \setminus \overline{\mathsf{B}}_\varepsilon(x)$ vorgegeben, gilt $\mathrm{d}(x, y) > \varepsilon$; und wir setzen $\delta[y] := \mathrm{d}(x, y) - \varepsilon > 0$. Wir erhalten mit der inversen Dreiecksungleichung:

$$\begin{aligned} z \in \mathsf{B}_{\delta[y]}(y) \qquad &\Rightarrow \qquad \mathrm{d}(y, z) < \delta[y] < \mathrm{d}(x, y) \\ &\Rightarrow \qquad \mathrm{d}(x, z) \overset{\mathsf{M4)}}{\geq} |\mathrm{d}(x, y) - \mathrm{d}(y, z)| = \mathrm{d}(x, y) - \mathrm{d}(y, z) > \mathrm{d}(x, y) - \delta[y] = \varepsilon \\ &\Rightarrow \qquad z \notin \overline{\mathsf{B}}_\varepsilon(x); \end{aligned}$$

also $\mathsf{B}_{\delta[y]}(y) \subseteq X \setminus \overline{\mathsf{B}}_\varepsilon(x) = O$. □

Bemerkung 34. *Wegen $\overline{\mathsf{B}}_0(x) = \{x\}$ für alle $x \in X$, sind gemäß Lemma 28 insbesondere also auch alle einelementigen Teilmengen von X abgeschlossen.*

Proposition 3. *Sei (X, d) ein metrischer Raum. Es gelten die folgenden Aussagen:*

OM1) $\emptyset, X \in \mathcal{O}(X)$,

OM2) $O_1, \ldots, O_n \in \mathcal{O}(X) \quad \Longrightarrow \quad O_1 \cap \ldots \cap O_n \in \mathcal{O}(X)$,

OM3) $\bigcup_{j \in J} O_j \in \mathcal{O}(X)$, *für jede Familie $(O_j)_{j \in J}$ von Elementen in $\mathcal{O}(X)$.*

Beweis. OM1) Klar.

OM2) Sei $x \in O_1 \cap \ldots \cap O_n$ vorgegeben. Wegen $O_\ell \in \mathcal{O}(X)$ für $\ell = 1, \ldots, n$, existiert für jedes $1 \leq \ell \leq n$ ein $\varepsilon_\ell > 0$ mit $\mathsf{B}_{\varepsilon_\ell}(x) \subseteq O_\ell$. Wir setzen $\varepsilon := \min(\varepsilon_1, \ldots, \varepsilon_n) > 0$, und erhalten

$$\mathsf{B}_\varepsilon(x) \subseteq \mathsf{B}_{\varepsilon_\ell}(x) \subseteq O_\ell \quad \text{für} \quad \ell = 1, \ldots, n \qquad \Longrightarrow \qquad \mathsf{B}_\varepsilon(x) \subseteq O_1 \cap \ldots \cap O_n.$$

Dies zeigt die Behauptung.

OM3) Sei $x \in \bigcup_{j \in J} O_j$ vorgegeben. Dann existiert ein $j \in J$ mit $x \in O_j$. Wegen $O_j \in \mathcal{O}(X)$, existiert ein $\varepsilon > 0$ mit $\mathsf{B}_\varepsilon(x) \subseteq O_j \subseteq \bigcup_{j \in J} O_j$; was die Behauptung zeigt. □

Wegen dem dritten Punkt in Bemerkung 8 gilt

$$M \cap Y = M \setminus (X \setminus Y) \qquad \text{für} \qquad M, Y \subseteq X. \tag{135}$$

Beispiel 31. *Sei (X, d) ein metrischer Raum. Dann gilt*

$$O \setminus A \in \mathcal{O}(X) \qquad \forall\, O \in \mathcal{O}(X),\ A \in \mathcal{A}(X).$$

Beweis der Behauptung: Wegen (135) gilt $A = X \setminus (X \setminus A)$. Nochmaliges anwenden von (135) (mit $M = O$ und $Y = X \setminus A$) liefert

$$O \cap (X \setminus A) = O \setminus (X \setminus (X \setminus A)) = O \setminus A.$$

Es gilt $X \setminus A \in \mathcal{O}(X)$ per Definition, sowie $O \in \mathcal{O}(X)$ per Annahme. Daher folgt die Behauptung aus OM2) in Proposition 3. □

Lemma 29. *Sei* (X, d) *ein metrischer Raum. Für* $Y \subseteq X$ *gilt*

$$Y \in \mathcal{O}(X) \qquad \Longleftrightarrow \qquad X \setminus Y \in \mathcal{A}(X).$$

Beweis. Es gilt $Y = X \setminus (X \setminus Y)$ wegen (135) ($M = X$), also

$$Y \in \mathcal{O}(X) \quad \overset{(135)}{\Longleftrightarrow} \quad X \setminus (X \setminus Y) \in \mathcal{O}(X) \quad \overset{\text{def.}}{\Longleftrightarrow} \quad X \setminus Y \in \mathcal{A}(X). \qquad \square$$

Korollar 13. *Sei* (X, d) *ein metrischer Raum. Es gelten die folgenden Aussagen:*

AM1) $\emptyset, X \in \mathcal{A}(X)$,

AM2) $A_1, \ldots, A_n \in \mathcal{A}(X) \quad \Longrightarrow \quad A_1 \cup \ldots \cup A_n \in \mathcal{A}(X)$,

AM3) $\bigcap_{j \in J} A_j \in \mathcal{A}(X)$, *für jede Familie* $(A_j)_{j \in J}$ *von Elementen in* $\mathcal{A}(X)$.

$*$*Beweis.* AM1) Wegen Proposition 3.OM1) gilt $X \setminus X = \emptyset \in \mathcal{O}(X)$ sowie $X \setminus \emptyset = X \in \mathcal{O}(X)$; also $\emptyset, X \in \mathcal{A}(X)$.

AM2) Es gilt $O_\ell := X \setminus A_\ell \in \mathcal{O}(X)$ für $\ell = 1, \ldots, n$, sowie $O_1 \cap \ldots \cap O_n \in \mathcal{O}(X)$ wegen Proposition 3.OM2). Dann gilt $X \setminus (O_1 \cap \ldots \cap O_n) \in \mathcal{A}(X)$ wegen Lemma 29; und Bemerkung 8 liefert

$$A_1 \cup \ldots \cup A_n = (X \setminus O_1) \cup \ldots \cup (X \setminus O_n) = X \setminus (O_1 \cap \ldots \cap O_n) \in \mathcal{A}(X).$$

AM3) Es gilt $O_j := X \setminus A_j \in \mathcal{O}(X)$ für alle $j \in J$, sowie $\bigcup_{j \in J} O_j \in \mathcal{O}(X)$ wegen Proposition 3.OM3). Dann gilt $X \setminus \bigcup_{j \in J} O_j \in \mathcal{A}(X)$ wegen Lemma 29; und Bemerkung 8 liefert

$$\bigcap_{j \in J} A_j = \bigcap_{j \in J} X \setminus O_j = X \setminus \bigcup_{j \in J} O_j \in \mathcal{A}(X). \qquad \square$$

***Beispiel 3.** *Sei* (X, d) *ein metrischer Raum. Dann gilt*

$$A \setminus O \in \mathcal{A}(X) \qquad \forall\, A \in \mathcal{A}(X),\ O \in \mathcal{O}(X).$$

Beweis der Behauptung: Wegen (135) gilt $O = X \setminus (X \setminus O)$. Nochmaliges anwenden von (135) (mit $M = A$ und $Y = X \setminus O$), liefert

$$A \cap (X \setminus O) = A \setminus (X \setminus (X \setminus O)) = A \setminus O.$$

Es gilt $X \setminus O \in \mathcal{A}(X)$ nach Lemma 29, sowie $A \in \mathcal{A}(X)$ per Annahme. Daher folgt die Behauptung aus AM3) in Korollar 13. $\square$

Bemerkung 35. *Wegen Proposition 3.*OM1) *und Korollar 13.*AM1)*, sind die beiden Mengen* $\emptyset$ *und* X *sowohl offen als auch abgeschlossen: Dies ist keine Widersprüchlichkeit in den Definitionen – siehe auch Beispiel 32 und Beispiel 33.*

Übung 56. *Wir betrachten den metrischen Raum* $(\mathbb{R}, \mathrm{d}_{|\cdot|})$*, also (vgl.* (128)*)*

$$\mathrm{d}_{|\cdot|} \colon \mathbb{R} \times \mathbb{R} \to [0, \infty), \qquad (x, y) \mapsto |x - y|.$$

Machen Sie sich zunächst klar, dass

$$(x, y) = \mathsf{B}_{\frac{y-x}{2}}\left(x + \tfrac{y-x}{2}\right) \qquad \text{für alle} \qquad \mathbb{R} \ni x < y \in \mathbb{R} \tag{136}$$

gilt; dass also wegen Lemma 27 das Intervall $(x, y) \subseteq \mathbb{R}$ *offen bezüglich der Betragsmetrik* $\mathrm{d}_{|\cdot|}$ *ist.*

a) *Zeigen Sie, dass für $x \in \mathbb{R}$ auch die Intervalle $(-\infty, x)$ und (x, ∞) offen sind.*

*Hinweis: Proposition 3.*OM3) *bzw.* (131).

b) *Folgern Sie aus a) durch Bildung geeigneter Komplemente (von geeigneten Vereinigungen), dass die folgenden Mengen abgeschlossen sind:*

$$\begin{aligned} [x,y] \quad &\textit{für} \quad \mathbb{R} \ni x \leq y \in \mathbb{R}, \\ (-\infty, x] \quad &\textit{für} \quad x \in \mathbb{R}, \\ [y, \infty) \quad &\textit{für} \quad y \in \mathbb{R}, \\ (-\infty, x] \cup [y, \infty) \quad &\textit{für} \quad \mathbb{R} \ni x < y \in \mathbb{R}. \end{aligned} \tag{137}$$

c) *Zeigen Sie, dass die Mengen in* (137) *nicht offen sind.*

Hinweis: Betrachten Sie z.B. für die Menge $(-\infty, x]$, die offenen Kugeln $\mathsf{B}_\varepsilon(x)$ mit $\varepsilon > 0$ – Was gilt für diese Kugeln nicht?

d) *Folgern Sie aus c), dass die folgenden Mengen nicht abgeschlossen sind:*

$$\begin{aligned} (-\infty, x) \cup (y, \infty) \quad &\textit{für} \quad \mathbb{R} \ni x \leq y \in \mathbb{R}, \\ (x, \infty) \quad &\textit{für} \quad x \in \mathbb{R}, \\ (-\infty, y) \quad &\textit{für} \quad y \in \mathbb{R}, \\ (x, y) \quad &\textit{für} \quad \mathbb{R} \ni x < y \in \mathbb{R}. \end{aligned} \tag{138}$$

Übung 57. *Sei (X, d) ein metrischer Raum. Zeigen Sie die folgenden Aussagen:*

a) *Für jedes $x \in X$, gilt*

$$\textstyle\bigcap_{n \in \mathbb{N}_{>0}} \mathsf{B}_{\frac{1}{n}}(x) = \{x\}.$$

Der Schnitt von abzählbar vielen offenen Mengen ist also nicht notwendigerweise ebenfalls offen. (Teil b) von Übung 56 zeigt ja, dass $\{x\} = [x, x]$ im Falle $(\mathbb{R}, \mathrm{d}_{|\cdot|})$ nicht offen ist.)

*Dies ist kein Widerspruch zu Proposition 3.*OM2)*, da dort nur endlich viele Schnitt auftauchen.*

b) *Für jedes $x \in X$, gilt*

$$\textstyle\bigcup_{n \in \mathbb{N}_{>0}} \overline{\mathsf{B}}_{1-\frac{1}{n}}(x) = \mathsf{B}_1(x).$$

Die Vereinigung von abzählbar vielen abgeschlossenen Mengen ist also nicht notwendigerweise ebenfalls abgeschlossen. (Teil c) von Übung 56 zeigt ja, dass das Intervall $(x - \varepsilon, x + \varepsilon)$ im Falle $(\mathbb{R}, \mathrm{d}_{|\cdot|})$ nicht abgeschlossen ist.)

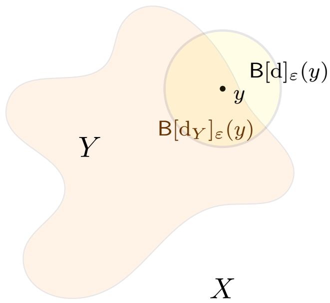
Dies ist kein Widerspruch zu AM2)*, da dort nur von endlichen Vereinigungen die Rede ist.*

Bemerkung 36 (Metrischer Unterraum)**.** *Sei (X, d) ein metrischer Raum, und $Y \subseteq X$ eine Teilmenge. Wir betrachten den metrischen Unterraum (Y, d_Y) mit der eingeschränkten Metrik (vgl. Terminologie 11)*

$$\mathrm{d}_Y := \mathrm{d}|_{Y \times Y} \colon Y \times Y \to [0, \infty), \qquad (x, y) \mapsto \mathrm{d}(x, y).$$

Es folgt zwanglos, dass gilt:

$$\mathsf{B}[\mathrm{d}_Y]_\varepsilon(y) = \mathsf{B}[\mathrm{d}]_\varepsilon(y) \cap Y \qquad \forall\, \varepsilon > 0,\ y \in Y. \tag{139}$$



Beweis von (139)*:* Wir erhalten

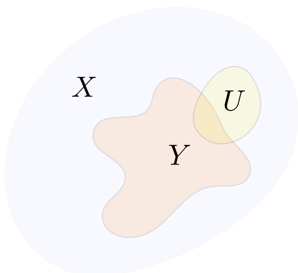
$$\begin{aligned}\mathsf{B}[\mathrm{d}_Y]_\varepsilon(y) &= \{z \in Y \mid \mathrm{d}_Y(y,z) < \varepsilon\}\\ &= \{z \in Y \mid \mathrm{d}(y,z) < \varepsilon\}\\ &= \{z \in X \mid z \in Y \ \wedge\ \mathrm{d}(y,z) < \varepsilon\}\\ &= \{z \in X \mid \mathrm{d}(y,z) < \varepsilon\} \cap Y\\ &= \mathsf{B}[\mathrm{d}]_\varepsilon(y) \cap Y,\end{aligned}$$

was die Behauptung zeigt. □

Man erhält weiterhin:

a) Eine Teilmenge $O \subseteq Y$ ist offen im metrischen (Unter)Raum (Y, d_Y) genau dann, wenn eine im metrischen Raum (X, d) offene Teilmenge $U \subseteq X$ existiert, sodass $O = U \cap Y$ gilt – In Formeln:

$$O \in \mathcal{O}[\mathrm{d}_Y](Y) \qquad \Longleftrightarrow \qquad \exists\, U \in \mathcal{O}[\mathrm{d}](X)\colon O = U \cap Y. \tag{140}$$



Beweis der Behauptung:

⇒) Ist $O \in \mathcal{O}[\mathrm{d}_Y](Y)$, so existieren (vgl. (132)) $\varepsilon[y] > 0$ für $y \in O$, mit $O = \bigcup_{y \in O} \mathsf{B}[\mathrm{d}_Y]_{\varepsilon[y]}(y)$. Wegen Lemma 27 und Proposition 3.OM3) gilt dann

$$U := \bigcup_{y \in O} \mathsf{B}[\mathrm{d}]_{\varepsilon[y]}(y) \in \mathcal{O}[\mathrm{d}](X).$$

Wir erhalten

$$U \cap Y = Y \cap \bigcup_{y \in O} \mathsf{B}[\mathrm{d}]_{\varepsilon[y]}(y) = \bigcup_{y \in O} \big(\mathsf{B}[\mathrm{d}]_{\varepsilon[y]}(y) \cap Y\big) \overset{(139)}{=} \bigcup_{y \in O} \mathsf{B}[\mathrm{d}_Y]_{\varepsilon[y]}(y) = O.$$

⇐) Sei $U \in \mathcal{O}[\mathrm{d}](X)$. Wir wollen zeigen, dass $O := U \cap Y \in \mathcal{O}[\mathrm{d}_Y](Y)$ gilt. Sei hierfür $y \in O$ vorgegeben. Wegen $y \in U$, existiert ein $\varepsilon > 0$ mit $\mathsf{B}[\mathrm{d}]_\varepsilon(y) \subseteq U$. Wir erhalten

$$\mathsf{B}[\mathrm{d}_Y]_\varepsilon(y) \overset{(139)}{=} \mathsf{B}[\mathrm{d}]_\varepsilon(y) \cap Y \subseteq U \cap Y = O.$$

Da $y \in O$ beliebig war, folgt hiermit $O \in \mathcal{O}[\mathrm{d}_Y](Y)$. □

b) **(∗)** *Eine Teilmenge $A \subseteq Y$ ist abgeschlossen im metrischen (Unter)Raum (Y, d_Y) genau dann, wenn eine im metrischen Raum (X, d) abgeschlossene Teilmenge $B \subseteq X$ existiert, sodass $A = B \cap Y$ gilt – In Formeln:*

$$A \in \mathcal{A}[\mathrm{d}_Y](Y) \qquad \Longleftrightarrow \qquad \exists\, B \in \mathcal{A}[\mathrm{d}](X) \colon A = B \cap Y. \tag{141}$$

Beweis der Behauptung:

$\Rightarrow$) Sei $A \in \mathcal{A}[\mathrm{d}_Y](Y)$ vorgegeben. Per Definition gilt $O := Y \setminus A \in \mathcal{O}[\mathrm{d}_Y](Y)$. Wegen (140), existiert ein $U \in \mathcal{O}[\mathrm{d}](X)$ mit $O = U \cap Y$. Wegen Lemma 29 gilt $B := X \setminus U \in \mathcal{A}[\mathrm{d}](X)$; und wir erhalten

$$B \cap Y = (X \setminus U) \cap Y = Y \setminus U = Y \setminus (U \cap Y) = Y \setminus O = Y \setminus (Y \setminus A) = A.$$

$\Leftarrow$) Sei $B \in \mathcal{A}[\mathrm{d}](X)$, und $A := B \cap Y$. Per Definition gilt $U := X \setminus B \in \mathcal{O}[\mathrm{d}](X)$; und wir erhalten

$$Y \setminus A = Y \setminus (B \cap Y) = Y \setminus B = (X \setminus B) \cap Y = U \cap Y \overset{(140)}{\in} \mathcal{O}[\mathrm{d}_Y](Y).$$

Dies zeigt $A \in \mathcal{A}[\mathrm{d}_Y](Y)$. □

Die Punkte b) und c) in Übung 56 könnten den Trugschluss verursachen, dass in einem metrischen Raum (X, d) Mengen generell immer entweder offen oder abgeschlossen sind – Gemäß Bemerkung 35 ist dies z.B. für die Teilmengen $\emptyset, X \subseteq X$ nie der Fall, da diese immer sowohl offen als auch abgeschlossen sind. Nichttrivialere Gegenbeispiele erhält man z.B., wenn man zu „gewissen unzusammenhängenden“ Teilmengen („Teilmengen mit Lücken“) von metrischen Räumen übergeht; und diese dann mit der eingeschränkten Metrik ausstattet:

Beispiel 32. *Sei $(X, \mathrm{d}) = (\mathbb{R}, \mathrm{d}_{|\cdot|})$, sowie $Y := \{-3\} \cup [-1, 1) \subseteq X$. Wir betrachten den metrischen Unterraum (Y, d_Y) – d.h.*

$$\mathrm{d}_Y \colon Y \times Y \to [0, \infty), \qquad (x, y) \mapsto |x - y|.$$

Dann gelten die folgenden Aussagen (Übung 58):

a) Die Mengen $\{-3\}, [-1, 1)$ sind offen im metrischen Raum (Y, d_Y).

Hinweis: Betrachten Sie offene Bälle um die Punkte -3 und 0.

b) Die Mengen $\{-3\}, [-1, 1)$ sind abgeschlossen im metrischen Raum (Y, d_Y).

Hinweis: Komplemente.

Übung 58. *Zeigen Sie die Aussagen a) und b) in Beispiel 32.*

Beispiel 33. *Sei $(X, \mathrm{d}) = (\mathbb{R}, \mathrm{d}_{|\cdot|})$. Wir betrachten den metrischen Unterraum $(Y, \mathrm{d}_Y) = (\mathbb{Q}, \mathrm{d}_{|\cdot|})$; also $\mathbb{Q}$ aufgefasst als Teilmenge von $\mathbb{R}$, mit der Betragsmetrik.*

- *Wie in Übung 56 sieht man:*
  - *Die folgenden Intervalle sind offen aber nicht abgeschlossen in $(\mathbb{Q}, \mathrm{d}_{|\cdot|})$:*

$$\begin{aligned} (x, y) \quad &\text{für} \quad \mathbb{Q} \ni x < y \in \mathbb{Q}, \\ (x, \infty) \quad &\text{für} \quad x \in \mathbb{Q}, \\ (-\infty, y) \quad &\text{für} \quad y \in \mathbb{Q}. \end{aligned}$$

– *Die folgenden Intervalle sind abgeschlossen aber nicht offen in* $(\mathbb{Q}, \mathrm{d}_{|\cdot|})$*:*

$$\begin{aligned} [x,y] \quad &\textit{für} \quad \mathbb{Q} \ni x \le y \in \mathbb{Q}, \\ (-\infty, x] \quad &\textit{für} \quad x \in \mathbb{Q}, \\ [y, \infty) \quad &\textit{für} \quad y \in \mathbb{Q}. \end{aligned}$$

- *Wegen der „Lückenhaftigkeit“ von* $\mathbb{Q}$*, existieren nun aber auch noch andere Intervalle in* $\mathbb{Q}$*, die nicht in obiger Form geschrieben werden können:*

  Sei $\tau \in \mathbb{R} \setminus \mathbb{Q}$ *irrational, sowie* $U_- := (-\infty, \tau) \in \mathcal{O}(\mathbb{R})$ *und* $U_+ := (\tau, \infty) \in \mathcal{O}(\mathbb{R})$*. Wegen Bemerkung 36.a) gilt*

$$\begin{aligned} \{x \in \mathbb{Q} \mid x < \tau\} &:= O_- := U_- \cap \mathbb{Q} \in \mathcal{O}(\mathbb{Q}) \\ \{x \in \mathbb{Q} \mid \tau < x\} &:= O_+ := U_+ \cap \mathbb{Q} \in \mathcal{O}(\mathbb{Q}). \end{aligned}$$

  Wegen $\tau \in \mathbb{R} \setminus \mathbb{Q}$ *gilt nun*

$$\mathbb{Q} = \mathbb{Q} \cap \mathbb{R} = \mathbb{Q} \cap (\mathbb{R} \setminus \{\tau\}) = \mathbb{Q} \cap \big((-\infty, \tau) \cup (\tau, \infty)\big),$$

  und daher

$$\begin{aligned} \mathbb{Q} = O_- \cup O_+ \quad \textit{mit} \quad O_- \cap O_+ = \emptyset \qquad &\Longrightarrow \qquad O_\pm = \mathbb{Q} \setminus O_\mp \\ &\Longrightarrow \qquad \mathcal{O}(\mathbb{Q}) \ni O_\pm \in \mathcal{A}(\mathbb{Q}). \end{aligned}$$

  Die Intervalle $O_\pm$ *sind also beide sowohl offen als auch abgeschlossen im metrischen (Unter)Raum* $(\mathbb{Q}, \mathrm{d}_{|\cdot|})$.

Übung 59. *Ein metrischer Raum* (X, d) *heißt diskret, wenn jeder Punkt in* X *isoliert liegt – wenn also für jedes* $x \in X$ *ein* $\varepsilon > 0$ *existiert, sodass* $\mathsf{B}_\varepsilon(x) \cap X \setminus \{x\} = \emptyset$ *gilt. Zeigen Sie die Äquivalenz der folgenden Aussagen:*

a) (X, d) *ist diskret,*

b) $\{x\} \in \mathcal{O}(X)$ *für alle* $x \in X$,

c) $\mathcal{O}(X) = \mathcal{P}(X)$,

d) $\mathcal{A}(X) = \mathcal{P}(X)$.

Beispiele für solche diskreten metrischen Räume sind:

– $(\mathbb{Z}, \mathrm{d})$ *mit*

$$\mathrm{d} \colon \mathbb{Z} \times \mathbb{Z} \to [0, \infty), \qquad (m,n) \mapsto |m - n|.$$

– (X, d) *mit einer beliebigen Menge* X*, sowie*

$$\mathrm{d}(x,y) = \begin{cases} 1 & \textit{für} \quad x \neq y \\ 0 & \textit{für} \quad x = y. \end{cases}$$

6 Konvergenz von Folgen und Reihen

In diesem Kapitel behandeln wir den Begriff der Konvergenz von Folgen und Reihen.

6.1 Folgen

Sei (X, d) ein metrischer Raum. Unter einer Folge in X, verstehen wir eine Familie $(x_n)_{n\in\mathsf{N}}$ von Elementen in X, wobei $\mathsf{N} = \{n \in \mathbb{Z} \mid n \geq \mathsf{p}\} \subseteq \mathbb{Z}$ für ein $\mathsf{p} \in \mathbb{Z}$ gilt (vgl. Notation 4). Alternativ notieren wir auch $(x_n)_{n\geq\mathsf{p}}$ anstelle $(x_n)_{n\in\mathsf{N}}$; sowie $(x_n)_{n\in\mathbb{N}}$ für $\mathsf{p} = 0$ und $(x_n)_{n\in\mathbb{N}_{>0}}$ für $\mathsf{p} = 1$.

Gilt $(X, \mathrm{d}) = (\mathbb{K}, \mathrm{d}_{|\cdot|})$ mit $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, so wird $(x_n)_{n\in\mathsf{N}}$ auch als *Zahlenfolge* bezeichnet – Beispielsweise (vgl. Beispiel 4):

- Die $\mathbb{Q}$-Folge $(\frac{1}{n})_{n\in\mathbb{N}_{>0}}$. $\qquad (x_n = \frac{1}{n}$ für alle $n \in \mathsf{N} = \mathbb{N}_{>0})$
- Die $\mathbb{N}$-Folge $(n^2)_{n\in\mathbb{N}}$. $\qquad (x_n = n^2$ für alle $n \in \mathsf{N} = \mathbb{N})$
- Die $\mathbb{C}$-Folge $(z^n)_{n\in\mathbb{N}}$ mit $z \in \mathbb{C}$. $\qquad (x_n = z^n$ für alle $n \in \mathsf{N} = \mathbb{N})$

Definition 35 (Teilfolgen)**.** *Sei $(x_n)_{n\in\mathsf{N}}$ eine Folge in X. Gegeben eine Abbildung $\iota\colon \mathbb{N} \to \mathsf{N}$ mit*

$$\mathsf{N} \ni m > n \in \mathsf{N} \qquad \Longrightarrow \qquad \iota(m) > \iota(n).$$

(ι ist dann notwendig injektiv), so wird die Folge $(x_{\iota(n)})_{n\in\mathbb{N}}$ als Teilfolge von $(x_n)_{n\in\mathsf{N}}$ bezeichnet.

Beispielsweise:

- *Sei $\mathsf{N} = \{n \in \mathbb{Z} \mid n \geq \mathsf{p}\}$ und $\iota\colon \mathbb{N} \ni n \mapsto n + \mathsf{p} \in \mathsf{N}$; also $(x_{\iota(n)})_{n\in\mathbb{N}} = (x_{n+\mathsf{p}})_{n\in\mathbb{N}}$.*

  Die Abbildung ι verschiebt daher nur den Anfangsindex p auf 0.
- *Sei $\mathsf{N} = \mathbb{N}$ und $\iota\colon \mathbb{N} \ni n \mapsto 2n \in \mathbb{N}$; also $(x_{\iota(n)})_{n\in\mathbb{N}} = (x_{2n})_{n\in\mathbb{N}}$.*

  Die Abbildung ι terminiert daher alle ungeraden Indizes (bzw. die dazugehörigen Folgenglieder).

Der erste Punkt legt nahe, dass man sich eigentlich immer auf die Indexmenge $\mathsf{N} = \mathbb{N}$ festlegen könnte. Dies bringt aber lästige technische Schwierigkeiten mit sich; sowohl in Beweisen, als auch bereits in der Angabe von einfachen Folgen wie $(\frac{1}{n})_{n\in\mathbb{N}_{>0}}$.

Bemerkung 37 (Teilfolgen von Teilfolgen sind Teilfolgen)**.** *Sei (X, d) ein metrischer Raum, sowie $(x_n)_{n\in\mathsf{N}}$, $(y_n)_{n\in\mathbb{N}}$, $(z_n)_{n\in\mathbb{N}}$ Folgen in X.*

- *Sei $(y_n)_{n\in\mathbb{N}}$ eine Teilfolge von $(x_n)_{n\in\mathsf{N}}$; also $(y_n)_{n\in\mathbb{N}} = (x_{\iota(n)})_{n\in\mathbb{N}}$ für $\iota\colon \mathbb{N} \to \mathsf{N}$ mit $\iota(m) > \iota(n)$ für alle $\mathsf{N} \ni m > n \in \mathsf{N}$.*
- *Sei $(z_n)_{n\in\mathbb{N}}$ eine Teilfolge von $(y_n)_{n\in\mathbb{N}}$; also $(z_n)_{n\in\mathbb{N}} = (y_{\kappa(n)})_{n\in\mathbb{N}}$ für ein $\kappa\colon \mathbb{N} \to \mathbb{N}$ mit $\kappa(m) > \kappa(n)$ für alle $\mathbb{N} \ni m > n \in \mathbb{N}$.*

Dann ist auch $(z_n)_{n\in\mathbb{N}} = (y_{\kappa(n)})_{n\in\mathbb{N}} = (x_{(\kappa\circ\iota)(n))})_{n\in\mathbb{N}}$ eine Teilfolge von $(x_n)_{n\in\mathsf{N}}$, wegen

$$\mathsf{N} \ni m > n \in \mathsf{N} \quad \Longrightarrow \quad \iota(m) > \iota(n) \quad \Longrightarrow \quad (\kappa \circ \iota)(m) = \kappa(\iota(m)) > \kappa(\iota(n)) = (\kappa \circ \iota)(n).$$

6.1.1 Grenzwerte von Folgen

Wir kommen nun zu einem zentralen Begriff der Analysis.

Definition 36 (Konvergenz von Folgen)**.** *Sei (X, d) ein metrischer Raum.*

a) Eine Folge $(x_n)_{n\in\mathsf{N}}$ in X heißt konvergent gegen $x \in X$ genau dann, wenn gilt:

$$\forall\, \varepsilon > 0\colon\ \exists\, N_\varepsilon \in \mathsf{N}\colon\ \forall\, n \geq N_\varepsilon\colon\ \mathrm{d}(x, x_n) < \varepsilon \tag{142}$$

(wenn also $x_n \in \mathsf{B}_\varepsilon(x)$ für alle $n \geq N_\varepsilon$, bzw. $\{x_n \mid \mathbb{N} \ni n \geq N_\varepsilon\} \subseteq \mathsf{B}_\varepsilon(x)$ gilt). Man sagt dann auch, dass fast alle (alle bis auf endlich viele) Folgenglieder von $(x_n)_{n\in\mathsf{N}}$ in $\mathsf{B}_\varepsilon(x)$ liegen.

Man verwendet eine der folgenden Schreibweisen um anzugeben, dass $(x_n)_{n\in\mathsf{N}}$ *gegen* x *konvergiert*

$$(x_n)_{n\in\mathsf{N}} \to x, \qquad x_n \to x, \qquad \lim_{n\to\infty} x_n = x, \qquad \lim_{n\to\infty} x_n = x, \qquad \lim_n x_n = x.$$

Es wird dann x *als* Grenzwert *(oder* Limes*) der Folge* $(x_n)_{n\in\mathsf{N}}$ *bezeichnet.*

Beachte: *Gilt* $\mathrm{d}(x,x_n) < \varepsilon$ *für ein* $\varepsilon > 0$*, so gilt auch* $\mathrm{d}(x,x_n) < \varepsilon'$ *für alle* $\varepsilon < \varepsilon'$*. Um* $\lim_n x_n = x$ *nachzuweisen, reicht es daher Folgendes zu zeigen:*

$$\forall\, m \in \mathbb{N}\colon\ \exists N_m \in \mathsf{N}\colon\ \forall\, n \geq N_m\colon\ \mathrm{d}(x,x_n) < \tfrac{1}{m}. \tag{143}$$

In der Tat existiert gemäß Proposition 2.2) zu jedem $\varepsilon > 0$ *ja ein* $m \in \mathbb{N}_{>0}$ *mit* $0 < \frac{1}{m} < \varepsilon$.

b) Eine Folge $(x_n)_{n\in\mathsf{N}}$ *in* X *heißt* divergent*, wenn sie keinen Grenzwert (in* X*) hat – also gegen kein* $x \in X$ *konvergiert – wenn also gilt:*

$$\forall\, x \in X\colon\ \exists \varepsilon > 0\colon\ \forall\, N \in \mathsf{N}\colon \exists\, n \geq N\colon\ \mathrm{d}(x,x_n) \geq \varepsilon. \tag{144}$$

(Für jedes $N \in \mathsf{N}$ *existiert also ein* $n \geq N$ *mit* $x_n \notin \mathsf{B}_\varepsilon(x)$*; und zwar für alle fixierten* $\varepsilon > 0$ *und alle* $x \in X$*.)*

Bemerkung 38. *Sei* $(V, \|\cdot\|)$ *ein normierter Raum. Wir verstehen dann den Konvergenzbegriff im Sinne des metrischen Raumes* $(V, \mathrm{d}_{\|\cdot\|})$ *(vgl. Bemerkung 33). Es liest sich dann also* (142) *als*

$$\forall\, \varepsilon > 0\colon\ \exists N_\varepsilon \in \mathsf{N}\colon\ \forall\, n \geq N_\varepsilon\colon\ \|x - x_n\| < \varepsilon \tag{145}$$

sowie (144) *in der Form*

$$\forall\, x \in X\colon\ \exists \varepsilon > 0\colon\ \forall\, N \in \mathsf{N}\colon \exists\, n \geq N\colon\ \|x - x_n\| \geq \varepsilon \tag{146}$$

Im Falle $V = \mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$*, ist dann wiederum einfach* $\|x - x_n\|$ *durch* $|x_n - x|$ *zu ersetzen.*

Beispiel 34.

a) Es gilt $(\frac{1}{n})_{n\in\mathbb{N}_{>0}} \to 0$. *(*$x = 0$ *sowie* $x_n = \frac{1}{n}$ *für alle* $n \in \mathsf{N} = \mathbb{N}_{>0}$*)*

Sei hierfür $\varepsilon > 0$ *vorgegeben. Wegen Proposition 2.2), existiert ein* $N_\varepsilon \in \mathbb{N}_{>0}$ *mit* $0 < \frac{1}{N_\varepsilon} < \varepsilon$*. Daher gilt*

$$|x - x_n| = |0 - \tfrac{1}{n}| = \tfrac{1}{n} \leq \tfrac{1}{N_\varepsilon} < \varepsilon \qquad \forall\, n \geq N_\varepsilon.$$

b) Es gilt $(\frac{n}{n+1})_{n\in\mathbb{N}} \to 1$. *(*$x = 1$ *sowie* $x_n = \frac{n}{n+1}$ *für alle* $n \in \mathsf{N} = \mathbb{N}$*)*

Sei $\varepsilon > 0$ *vorgegeben. Wie in Teil a), existiert dann ein* $N_\varepsilon \in \mathbb{N}_{>0}$ *mit* $\frac{1}{n+1} < \varepsilon$ *für alle* $n \geq N_\varepsilon$*. Wir erhalten*

$$|x - x_n| = |1 - \tfrac{n}{n+1}| = |\tfrac{n+1-n}{n+1}| = \tfrac{1}{n+1} < \varepsilon \qquad \forall\, n \geq N_\varepsilon.$$

c) Es gilt $(\frac{n}{2^n})_{n\in\mathbb{N}} \to 0$. *(*$x = 0$ *sowie* $x_n = \frac{n}{2^n}$ *für alle* $n \in \mathsf{N} = \mathbb{N}$*)*

Der Binomische Lehrsatz (Satz 6) liefert für $n \in \mathbb{N}$

$$2^n = (1+1)^n = \textstyle\sum_{k=0}^n \binom{n}{k} = 1 + n + \binom{n}{2} + \ldots \geq n + \frac{n\cdot(n-1)}{2} = \frac{n\cdot(n+1)}{2} \geq \frac{n^2}{2}$$

und daher $\frac{n}{2^n} \leq \frac{2}{n}$*. Für* $\varepsilon > 0$ *vorgegeben, sei* $N_\varepsilon \in \mathbb{N}$ *mit* $\frac{2}{n} < \varepsilon$ *für alle* $n \geq N_\varepsilon$*. Dann gilt*

$$|x - x_n| = |0 - \tfrac{n}{2^n}| = \tfrac{n}{2^n} \leq \tfrac{2}{n} < \varepsilon \qquad \forall\, n \geq N_\varepsilon.$$

Beispiel 35 (Konstante Folge)**.** *Sei (X, d) ein metrischer Raum, $(x_n)_{n\in\mathbb{N}}$ eine Folge in X, sowie $x \in X$. Existiert ein $N \in \mathsf{N}$ mit $x_n = x$ für alle $n \geq N$, so gilt*

$$\mathrm{d}(x, x_n) = 0 \qquad \forall\, n \geq N \qquad\qquad \Longrightarrow \qquad\qquad \lim_n x_n = x.$$

Übung 60. *Zeigen Sie $(\frac{1}{m^n})_{n\in\mathbb{N}_{>0}} \to 0$ für alle $m \geq 2$.*
Hinweis: Zeigen Sie beispielsweise zunächst per Induktion $m^n > n$ für alle $n \in \mathbb{N}_{>0}$, falls $m \geq 2$.

Satz 12. *Sei (X, d) ein metrischer Raum. Existiert der Grenzwert einer Folge in X, so ist dieser eindeutig bestimmt (jede Folge hat höchstens einen Grenzwert).*

Beweis. Wir nehmen an, die Behauptung wäre falsch. Dann existiert eine Folge $(x_n)_{n\in\mathsf{N}}$ in X, sowie $X \ni x \neq \tilde{x} \in X$ derart, dass sowohl $\lim_n x_n = x$ als auch $\lim_n x_n = \tilde{x}$ gilt. Wegen $x \neq \tilde{x}$ ist $\varepsilon := \mathrm{d}(x, \tilde{x}) > 0$. Per definitionem (vgl. (142)) existieren $N_1, N_2 \in \mathsf{N}$ mit

$$\mathrm{d}(x, x_n) < \frac{\varepsilon}{2} \quad \forall\, n \geq N_1 \qquad\qquad \text{sowie} \qquad\qquad \mathrm{d}(\tilde{x}, x_n) < \frac{\varepsilon}{2} \quad \forall\, n \geq N_2.$$

Für $N := \max(N_1, N_2)$ und alle $n \geq N$ gilt daher $\mathrm{d}(x, x_n) < \frac{\varepsilon}{2}$ sowie $\mathrm{d}(\tilde{x}, x_n) < \frac{\varepsilon}{2}$. Wir erhalten

$$\mathrm{d}(x, \tilde{x}) \overset{\mathsf{M3)}}{\leq} \mathrm{d}(x, x_N) + \mathrm{d}(x_N, \tilde{x}) \overset{\mathsf{M2)}}{=} \mathrm{d}(x, x_N) + \mathrm{d}(\tilde{x}, x_N) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon = \mathrm{d}(x, \tilde{x}),$$

also einen Widerspruch. □

Bemerkung* 8. *Der geometrische Hintergrund zu Satz 12 ist der, dass zu $X \ni x \neq y \in X$ vorgegeben, ein $\varepsilon > 0$ mit $\mathsf{B}_\varepsilon(x) \cap \mathsf{B}_\varepsilon(y) = \emptyset$ existiert (es können dann nicht gleichzeitig fast alle Folgenglieder sowohl in $\mathsf{B}_\varepsilon(x)$ als auch in $\mathsf{B}_\varepsilon(y)$ liegen).*

Beweis der Behauptung: Angenommen, die Behauptung ist falsch. Für jedes $\varepsilon > 0$, existiert dann ein $z \in \mathsf{B}_{\frac{\varepsilon}{2}}(x) \cap \mathsf{B}_{\frac{\varepsilon}{2}}(y)$; womit gilt:

$$\mathrm{d}(x, y) \leq \mathrm{d}(x, z) + \mathrm{d}(z, y) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, gilt $\mathrm{d}(x, y) = 0$ wegen Lemma 19. Wegen $\mathsf{M1)}$ gilt also $x = y$, was im Widerspruch zur Annahme $x \neq y$ steht. □

Lemma 30 (Erhalt von Ungleichungen unter Limiten)**.** *Seien $(a_n)_{n\in\mathsf{N}} \to a$ und $(b_n)_{n\in\mathsf{N}} \to b$ konvergente Folgen in $\mathbb{R}$. Existiert ein $N \in \mathsf{N}$ mit $a_n \leq b_n$ für alle $n \geq N$, so gilt $a \leq b$.*

Beweis. Angenommen es gilt $b < a$. Wir setzen $\varepsilon := \frac{a-b}{2} > 0$, und finden (wegen $\lim_n a_n = a$ und $\lim_n b_n = b$) ein $m \geq N$, mit $|a - a_m| < \varepsilon$ und $|b - b_m| < \varepsilon$.

- Es gilt $a_m > a - \varepsilon$ wegen:

$$\begin{aligned} a \geq a_m \qquad &\Longrightarrow \qquad a - a_m = |a - a_m| < \varepsilon \qquad \Longrightarrow \qquad a_m > a - \varepsilon \\ a_m > a \qquad &\Longrightarrow \qquad a_m > a > a - \varepsilon. \end{aligned}$$

- Es gilt $b_m < b + \varepsilon$ wegen:

$$\begin{aligned} b > b_m \qquad &\Longrightarrow \qquad b_m < b \leq b + \varepsilon \\ b_m > b \qquad &\Longrightarrow \qquad b_m - b = |b - b_m| < \varepsilon \qquad \Longrightarrow \qquad b_m < b + \varepsilon. \end{aligned}$$

Wir erhalten mit $a - b = 2\varepsilon$ (also $b + \varepsilon = a - \varepsilon$), dass

$$b_m < b + \varepsilon = a - \varepsilon < a_m$$

gilt, was wegen $m \geq N$ der Voraussetzung $a_m \leq b_m$ widerspricht. □

Korollar 14. *Sei* $(a_n)_{n\in\mathsf{N}} \to a$ *eine konvergente Folgen in* $\mathbb{R}$*, sowie* $c \in \mathbb{R}$*.*

- *Existiert ein* $N \in \mathsf{N}$ *mit* $a_n \leq c$ *für alle* $n \geq N$*, so gilt* $a \leq c$*.*
- *Existiert ein* $N \in \mathsf{N}$ *mit* $c \leq a_n$ *für alle* $n \geq N$*, so gilt* $c \leq a$*.*

Beweis. Dies folgt sofort aus Lemma 30, wenn man c als konstante Folge auffasst. $\square$

Bemerkung 39. *Für konvergente* $\mathbb{R}$*-Folgen* $(a_n)_{n\in\mathsf{N}} \to a$ *und* $(b_n)_{n\in\mathsf{N}} \to b$*, folgt aus* $a_n < b_n$ *für alle* $n \in \mathsf{N}$ *im Allgemeinen nicht* $a < b$*, sondern auch nur* $a \leq b$*. Ist beispielsweise* $\mathsf{N} = \mathbb{N}_{>0}$*, sowie* $a_n = 0$ *und* $b_n = \frac{1}{n}$ *für alle* $n \in \mathsf{N}$*, so gilt zwar* $a_n < b_n$ *für alle* $n \in \mathsf{N}$*, aber es ist* $\lim_n a_n = 0 = \lim_n b_n$*.*

Korollar 15. *Seien* $(a_n)_{n\in\mathsf{N}} \to a$*,* $(b_n)_{n\in\mathsf{N}} \to b$*,* $(c_n)_{n\in\mathsf{N}} \to c$ *konvergente Folgen in* $\mathbb{R}$*. Existiert ein* $N \in \mathsf{N}$ *mit* $a_n \leq b_n \leq c_n$ *für alle* $n \geq N$*, so gilt* $a \leq b \leq c$*. Insbesondere impliziert* $a = c$*, bereits* $a = b = c$*.*

Beweis. Dies folgt sofort aus Lemma 30. $\square$

Übung 61. *Seien* $\mathbb{R} \ni a \leq b \in \mathbb{R}$*,* $(c_n)_{n\in\mathsf{N}} \to c \in \mathbb{R}$ *eine konvergente Folge in* $\mathbb{R}$*, sowie* $N \in \mathsf{N}$*. Zeigen Sie die Implikation*

$$c_n \in [a,b] \qquad \forall\, n \geq N \qquad\Longrightarrow\qquad c \in [a,b].$$

Lemma 31 (Quetschlemma)**.** *Seien* $(a_n)_{n\in\mathsf{N}}$*,* $(b_n)_{n\in\mathsf{N}}$*,* $(c_n)_{n\in\mathsf{N}}$ *Folgen in* $\mathbb{R}$*. Es gelte* $(a_n)_{n\in\mathsf{N}} \to \tau$ *sowie* $(c_n)_{n\in\mathsf{N}} \to \tau$*, für ein* $\tau \in \mathbb{R}$*; und weiterhin existiere ein* $N \in \mathsf{N}$ *mit* $a_n \leq b_n \leq c_n$ *für alle* $n \geq N$*. Dann gilt* $\lim_n b_n = \tau$*.*

Beweis. Für $\varepsilon > 0$ vorgegeben, existiert ein $N_1 \in \mathsf{N}$ mit $a_n \in (\tau - \varepsilon, \tau + \varepsilon)$ für alle $n \geq N_1$, sowie ein $N_2 \in \mathsf{N}$ mit $c_n \in (\tau - \varepsilon, \tau + \varepsilon)$ für alle $n \geq N_2$. Für alle $n \geq N_\varepsilon := \max(N, N_1, N_2)$, gilt dann

$$\tau - \varepsilon < a_n \leq b_n \leq c_n < \tau + \varepsilon \qquad\Longrightarrow\qquad b_n \in (\tau - \varepsilon, \tau + \varepsilon).$$

Dies zeigt $\lim_n b_n = \tau$. $\square$

Proposition 4. *Sei* (X, d) *ein metrischer Raum, und* $(x_n)_{n\in\mathsf{N}} \to x \in X$ *eine konvergente Folge in* X*. Sei zudem* $\iota\colon \mathbb{N} \to \mathsf{N}$ *injektiv. Dann gilt* $\lim_n x_{\iota(n)} = x$*.*

Beweis. Sei $\varepsilon > 0$ vorgegeben, und $\mathsf{N} = \{n \in \mathbb{Z} \,|\, n \geq \mathsf{p}\}$.

- Wegen $(x_n)_{n\in\mathsf{N}} \to x$, existiert ein $N_\varepsilon \in \mathsf{N}$ mit $x_n \in \mathsf{B}_\varepsilon(x)$ für alle $n \geq N_\varepsilon$.
- Da ι injektiv ist, ist die Menge $M := \iota^{-1}(\{\mathsf{p}, \dots, N_\varepsilon\}) \subseteq \mathbb{N}$ endlich (es gilt $|M| \leq N_\varepsilon + 1 - \mathsf{p}$).
- Sei $\tilde{N}_\varepsilon := 1 + \max(M)$. Für $n \geq \tilde{N}_\varepsilon$, gilt dann $n \notin M$; also $\iota(n) \geq N_\varepsilon + 1$, und somit $x_{\iota(n)} \in \mathsf{B}_\varepsilon(x)$ wegen dem ersten Punkt.[21]

Dies zeigt $(x_{\iota(n)})_{n\in\mathbb{N}} \to x$. $\square$

Korollar 16. *Sei* (X, d) *ein metrischer Raum und* $(x_n)_{n\in\mathsf{N}} \to x \in X$ *eine konvergente Folge in* X*. Dann konvergiert auch jede Teilfolge* $(x_{\iota(n)})_{n\in\mathbb{N}}$ *von* $(x_n)_{n\in\mathsf{N}}$ *gegen* x*.*

Beweis. Klar wegen Proposition 4; denn für eine Teilfolge $(x_{\iota(n)})_{n\in\mathsf{N}}$ von $(x_n)_{n\in\mathsf{N}}$, ist $\iota\colon \mathsf{N} \to \mathbb{N}$ notwendig injektiv ($\mathsf{N} \in m > n \in \mathsf{N} \quad\Rightarrow\quad \iota(m) > \iota(n)$ – siehe Definition 35). $\square$

Übung 62. *Sei* (X, d) *ein metrischer Raum,* $x \in X$*, und* $(x_n)_{n\in\mathsf{N}}$ *eine Folge in* X*. Zeigen Sie die Äquivalenz der folgenden Aussagen:*

a) Es gilt $(x_n)_{n\in\mathsf{N}} \to x$*.*

[21] Beachte, dass nicht notwendigerweise $\iota(\max(M)) = N_\varepsilon$ gelten muss, da ι nur als injektiv vorausgesetzt wurde.

b) Es gilt $(\mathrm{d}(x, x_n))_{n \in \mathsf{N}} \to 0$.

c) Es existiert eine Folge $(a_n)_{n \in \tilde{\mathsf{N}}}$ *in* $[0, \infty)$ *mit* $\lim_n a_n = 0$*, sodass ein* $N \in \mathsf{N} \cap \tilde{\mathsf{N}}$ *existiert mit*

$$\mathrm{d}(x, x_n) \le a_n \qquad \forall\, n \ge N.$$

Übung 63. *Zeigen Sie die folgenden Aussagen:*

a) Für eine komplexe Zahl $z = x + \mathrm{i}y$ *gilt*

$$\max(|x|, |y|) \le |z| \le |x| + |y|.$$

Hinweis: Übung 52 bzw. Beweis von (126).

b) Sei $(z_n)_{n \in \mathsf{N}} = (x_n + \mathrm{i}y_n)_{n \in \mathsf{N}}$ *eine komplexe Zahlenfolge. Dann gilt*

$$(z_n)_{n \in \mathsf{N}} \to z = (x, y) \qquad \Longleftrightarrow \qquad ((x_n)_{n \in \mathsf{N}} \to x) \quad \wedge \quad ((y_n)_{n \in \mathsf{N}} \to y).$$

c) Zeigen Sie: $\lim_n \left(\frac{n}{n+1} + \mathrm{i}\frac{n}{2^n+1}\right) = 1$.

Übung 64.

a) Sei $M \supseteq \mathbb{R} \le x \in \mathbb{R}$ *mit* $M \neq \emptyset$*. Zeigen Sie, dass* $\sup_{\mathbb{R}}(M) = x$ *genau dann gilt, wenn eine Folge* $(x_n)_{n \in \mathbb{N}}$ *in* M *existiert, sodass* $\lim_n x_n = x$ *gilt.*

b) Sei $\mathbb{R} \ni x \le M \subseteq \mathbb{R}$ *mit* $M \neq \emptyset$*. Zeigen Sie, dass* $\inf_{\mathbb{R}}(M) = x$ *genau dann gilt, wenn eine Folge* $(x_n)_{n \in \mathbb{N}}$ *in* M *existiert, sodass* $\lim_n x_n = x$ *gilt.*

6.1.2 Beschränktheit

Terminologie 12. *Sei* (X, d) *ein metrischer Raum. Eine Teilmenge* $Y \subseteq X$ *heißt beschränkt, wenn ein* $C \ge 0$ *existiert, mit*

$$\mathrm{d}(z, y) \le C \qquad \forall\, z, y \in Y. \tag{147}$$

(Wegen der Dreiecksungleichung, ist es äquivalent zu fordern, dass ein $\tau \in X$ *und ein* $C \ge 0$ *existiert, sodass* $\mathrm{d}(\tau, y) \le C$ *für alle* $y \in Y$ *gilt – also* $Y \subseteq \overline{\mathsf{B}}_C(\tau)$*.)*

- *Ist* $(V, \|\cdot\|)$ *ein normierter Raum, so verstehen wir obige Definition wieder im Sinne des metrischen Raumes* $(V, \mathrm{d}_{\|\cdot\|})$*. In diesem Fall ist* $Y \subseteq V$ *genau dann beschränkt, wenn ein* $C \ge 0$ *existiert, mit*

$$\|y\| \le C \qquad \forall\, y \in Y. \tag{148}$$

  Beweis der Äquivalenz:

  – Es gelte (147) für $\mathrm{d} = \mathrm{d}_{\|\cdot\|}$. Wir fixieren $z \in Y$, und erhalten

$$\|y\| = \|y - z + z\| \le \|y - z\| + \|z\| = \mathrm{d}_{\|\cdot\|}(z, y) + \|z\| \le C + \|z\| =: \tilde{C} \qquad \forall\, y \in V.$$

  – Gilt (148), so folgt

$$\mathrm{d}_{\|\cdot\|}(z, y) = \|y - z\| = \|y + (-z)\| \le \|y\| + \|z\| \le 2 \cdot C =: \tilde{C} \qquad \forall\, z, y \in Y. \qquad \square$$

- *Ist* $V = \mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ *und* $\|\cdot\| = |\cdot|$*, so ließt sich* (148) *in der Form*

$$|y| \le C \qquad \forall\, y \in Y. \tag{149}$$

  (Beachte: Für $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}\}$ *ist* (149) *gleichbedeutend mit der Beschränktheit von* Y *in* $\mathbb{K}$ *im Sinne von Definition 13 (Übung).)*

Terminologie 13. *Sei* (X, d) *ein metrischer Raum. Eine Folge* $(x_n)_{n\in\mathsf{N}}$ *in* X *heißt beschränkt, wenn die Teilmenge* $\{x_n \mid n \in \mathsf{N}\} \subseteq X$ *beschränkt ist, d.h.,*

- *Allgemeiner Fall: Es existiert ein* $C \geq 0$ *mit*

$$\mathrm{d}(x_m, x_n) \leq C \qquad \forall\, m, n \in \mathsf{N}. \tag{150}$$

- $(X, \mathrm{d}) = (V, \mathrm{d}_{\|\cdot\|})$*: Es existiert ein* $C \geq 0$ *mit*

$$\|x_n\| \leq C \qquad \forall\, n \in \mathsf{N}. \tag{151}$$

- $(X, \mathrm{d}) = (\mathbb{K}, \mathrm{d}_{|\cdot|})$*: Es existiert ein* $C \geq 0$ *mit*

$$|x_n| \leq C \qquad \forall\, n \in \mathsf{N}. \tag{152}$$

Satz 13. *Sei* (X, d) *ein metrischer Raum, und* $(x_n)_{n\in\mathsf{N}} \to x \in X$ *eine konvergente Folge in* X*. Dann ist* $(x_n)_{n\in\mathsf{N}}$ *beschränkt.*

Beweis allgemeiner Fall: Wegen $(x_n)_{n\in\mathsf{N}} \to x$, existiert ein $N \in \mathsf{N}$ mit

$$\mathrm{d}(x, x_n) \leq 1 \qquad \forall\, n \geq N.$$

Sei $C := 1 + \max(\{\mathrm{d}(x, x_m) \mid \mathsf{N} \ni m \leq N\})$.

- Für alle $\mathsf{N} \ni m \leq N$, gilt $\mathrm{d}(x, x_n) \leq \max(\{\mathrm{d}(x, x_m) \mid \mathsf{N} \ni m \leq N\}) \leq C$.
- Für alle $n \geq N$, gilt $\mathrm{d}(x, x_n) \leq 1 \leq C$.

Für alle $m, n \in \mathsf{N}$ folgt dann mit M2) und M3) (Dreiecksungleichung)

$$\mathrm{d}(x_m, x_n) \overset{\text{M3)}}{\leq} \mathrm{d}(x_m, x) + \mathrm{d}(x, x_n) \overset{\text{M2)}}{=} \mathrm{d}(x, x_m) + \mathrm{d}(x, x_n) \leq C + C = 2C.$$

Dies zeigt die Beschränktheit von $\{x_n \mid n \in \mathsf{N}\}$. □

∗*Beweis für* $(\mathbb{K}, \mathrm{d}_{|\cdot|})$*:* Wegen $(x_n)_{n\in\mathsf{N}} \to x$, existiert ein $N \in \mathsf{N}$ mit

$$|x - x_n| \leq 1 \qquad \forall\, n \geq N.$$

Sei $C := 1 + |x| + \max(\{|x_m| \mid \mathsf{N} \ni m \leq N\})$.

- Für alle $\mathsf{N} \ni m \leq N$, gilt $|x_m| \leq \max(\{|x_m| \mid \mathsf{N} \ni m \leq N\}) \leq C$.
- Für alle $n \geq N$, folgt mit der Dreiecksungleichung

$$|x_n| = |(x - x_n) + x| \leq |x - x_n| + |x| \leq 1 + |x| \leq C.$$

Dies zeigt die Beschränktheit von $\{x_n \mid n \in \mathsf{N}\}$. □

Die nächste Proposition zeigt insbesondere, dass die beschränkte Folge $((-1)^n)_{n\in\mathbb{N}}$ divergent ist; dass also die Umkehrung von Satz 13 im Allgemeinen nicht gilt.

Proposition 5. *(Konvergenzverhalten der geometrische Folge) Sei* $z \in \mathbb{C}$*, und* $(z^n)_{n\in\mathbb{N}}$ *die zugehörige geometrische Folge. Dann gilt*

$$(z^n)_{n\in\mathbb{N}} \to \begin{cases} 1 & \textit{für } z = 1 \\ 0 & \textit{für } |z| < 1, \end{cases}$$

und $(z^n)_{n\in\mathbb{N}}$ *ist divergent falls* $z \neq 1$ *mit* $|z| \geq 1$ *gilt.*

Beweis. • Sei $z = 1$. Dann gilt $z^n = 1$ für alle $n \in \mathbb{N}$; also $(z^n)_{n\in\mathbb{N}} \to 1$.

- Sei $|z| > 1$. Die Bernoullische Ungleichung (Satz 8) liefert dann

$$|z^n| = |z|^n = (1 + (|z| - 1))^n \geq 1 + n \cdot (|z| - 1).$$

  Daher ist $(z^n)_{n\in\mathbb{N}}$ nicht beschränkt wegen Proposition 2.3) (und $|z| - 1 > 0$); kann also nach Satz 13 nicht konvergent sein.

- Sei $|z| < 1$. Ist $z = 0$, so gilt $z^n = 0$ für alle $n \in \mathbb{N}$; also $(z^n)_{n\in\mathbb{N}} \to 0$. Sei also $z \neq 0$ und $\varepsilon > 0$ vorgegeben. Dann gilt $|z|^{-1} > 1$. Wir erhalten wie im vorherigen Fall

$$|z^n|^{-1} = (|z|^n)^{-1} = (|z|^{-1})^n = (1 + (|z|^{-1} - 1))^n \geq 1 + n \cdot (|z|^{-1} - 1).$$

  Wegen Proposition 2.3) existiert ein $N_\varepsilon \in \mathbb{N}$ mit $1 + n \cdot (|z|^{-1} - 1) > \frac{1}{\varepsilon}$ für alle $n \geq N_\varepsilon$; also

$$|0 - z^n| = |z^n| \leq \frac{1}{1 + n \cdot (|z|^{-1} - 1)} < \varepsilon \qquad \forall\, n \geq N_\varepsilon.$$

  Dies zeigt $(z^n)_{n\in\mathbb{N}} \to 0$.

- Sei $|z| = 1$ sowie $z \neq 1$. Wir argumentieren durch Widerspruch; nehmen also an, dass $(z^n)_{n\in\mathbb{N}} \to w$ für ein $w \in \mathbb{C}$ gilt. Dann existiert ein $N \in \mathbb{N}$ mit $|z^n - w| < \frac{1}{2} \cdot |z - 1|$ für alle $n \geq N$ (beachte $z - 1 \neq 0$). Wir erhalten (wegen $|z| = 1$)

$$\begin{aligned}|z - 1| = |z|^N \cdot |z - 1| &= |z^{N+1} - z^N| = |z^{N+1} - w + w - z^N| \\ &\leq |z^{N+1} - w| + |w - z^N| < \tfrac{1}{2} \cdot |z - 1| + \tfrac{1}{2} \cdot |z - 1| = |z - 1|,\end{aligned}$$

  also einen Widerspruch. □

Terminologie 14 (Bestimmte Divergenz)**.** *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$.*

- *Man schreibt $(a_n)_{n\in\mathbb{N}} \to \infty$, wenn für jedes $C \geq 0$ ein $N \in \mathbb{N}$ existiert, sodass $a_n \geq C$ für alle $n \geq N$ gilt. In diesem Falle sagt man, dass $(a_n)_{n\in\mathbb{N}}$ bestimmt gegen ∞ divergiert.*
- *Man schreibt $(a_n)_{n\in\mathbb{N}} \to -\infty$, wenn für jedes $C \leq 0$ ein $N \in \mathbb{N}$ existiert, sodass $a_n \leq C$ für alle $n \geq N$ gilt. In diesem Falle sagt man, dass $(a_n)_{n\in\mathbb{N}}$ bestimmt gegen $-\infty$ divergiert.*

Definition 37 (Erweiterte reelle Zahlen)**.** *Die* erweiterten reellen Zahlen *sind definiert durch* $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$.

- *Divergiert eine reelle Folge bestimmt gegen $\pm\infty$, so sagen wir auch, dass sie gegen $\pm\infty$ konvergiert.*
- *Wir sagen, dass eine reelle Folge konvergent in $\overline{\mathbb{R}}$ ist, wenn sie gegen ein $\tau \in \overline{\mathbb{R}}$ konvergiert.*

Beispiel 36.

a) *Es gilt $(\pm n)_{n\in\mathbb{N}} \to \pm\infty$ da $\mathbb{R}$ archimedisch angeordnet ist.*

b) *Es gilt $(\frac{2^n}{n})_{n\in\mathbb{N}_{>0}} \to \infty$.*

   Beweis. Wegen Beispiel 34.c), gilt $(\frac{n}{2^n})_{n\in\mathbb{N}} \to 0$. Für $C > 0$ vorgegeben, setzte $\varepsilon := \frac{1}{C}$. Dann existiert ein $N \in \mathbb{N}$ mit

$$\tfrac{n}{2^n} = |0 - \tfrac{n}{2^n}| < \varepsilon \quad \forall\, n \geq N \qquad \Longrightarrow \qquad \tfrac{2^n}{n} > \tfrac{1}{\varepsilon} = C \quad \forall\, n \geq N.$$

   Dies zeigt die Behauptung. □

Übung 65. *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $(0, \infty)$. Zeigen Sie die folgenden Implikationen:*

- $(a_n)_{n\in\mathsf{N}} \to 0 \quad \Longrightarrow \quad \left(\frac{1}{a_n}\right)_{n\in\mathsf{N}} \to \infty.$
- $(a_n)_{n\in\mathsf{N}} \to \infty \quad \Longrightarrow \quad \left(\frac{1}{a_n}\right)_{n\in\mathsf{N}} \to 0.$

Hinweis: Verallgemeinern Sie die Argumentation in Beispiel 36.b).

Übung 66. *Zeigen Sie* $\lim_n \sqrt[k]{a_n} = \infty$*, für jedes* $k \in \mathbb{N}_{>0}$ *und jede Folge* $(a_n)_{n\in\mathbb{N}}$ *in* $[0,\infty)$ *mit* $\lim_n a_n = \infty$.

Hinweis: Übung 52.

6.1.3 Grenzwertsätze für Folgen

In diesem Abschnitt betrachten wir den Spezialfall normierter Räume; sowie der Körper $\mathbb{Q}, \mathbb{R}, \mathbb{C}$. Im Folgenden sei also immer $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$.

Proposition 6. *Sei* $(V, \|\cdot\|)$ *ein normierter* $\mathbb{K}$*-Vektorraum, sowie* $\lambda, \mu \in \mathbb{K}$*. Seien weiterhin* $(v_n)_{n\in\mathsf{N}} \to v \in V$ *und* $(w_n)_{n\in\mathsf{N}} \to w \in V$ *konvergente Folgen in* V*, sowie* $(\lambda_n)_{n\in\mathsf{N}} \to \lambda$ *eine konvergente Folge in* $\mathbb{K}$*. Dann gelten die folgenden Aussagen:*

1) $(v_n + w_n)_{n\in\mathsf{N}} \to v + w$.

2) $(-v_n)_{n\in\mathsf{N}} \to -v$.

3) $(\lambda_n \cdot v_n)_{n\in\mathsf{N}} \to \lambda \cdot v$.

4) $(\lambda \cdot v_n + \mu \cdot w_n)_{n\in\mathsf{N}} \to \lambda \cdot v + \mu \cdot w$.

Beweis. 1) Sei $\varepsilon > 0$ vorgegeben, sowie $N_1, N_2 \in \mathsf{N}$ mit

$$\|v - v_n\| < \tfrac{\varepsilon}{2} \quad \forall\, n \geq N_1 \qquad \text{und} \qquad \|w - w_n\| < \tfrac{\varepsilon}{2} \quad \forall\, n \geq N_2.$$

Dann folgt mit der Dreiecksungleichung

$$\|(v + w) - (v_n + w_n)\| \leq \|v - v_n\| + \|w - w_n\| < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon \qquad \forall\, n \geq N_\varepsilon = \max(N_1, N_2),$$

was die Behauptung zeigt.

2) Dies ist unmittelbar klar wegen $\|-v - (-v_n)\| = \|v - v_n\|$ (vgl. V2)).

3) Zunächst erhalten wir aus V2) und V3), dass

$$\begin{aligned} \|\lambda \cdot v - \lambda_n \cdot v_n\| &= \|(\lambda \cdot v - \lambda_n \cdot v) + (\lambda_n \cdot v - \lambda_n \cdot v_n)\| \\ &\leq \|(\lambda - \lambda_n) \cdot v\| + \|\lambda_n \cdot (v - v_n)\| \\ &= |\lambda - \lambda_n| \cdot \|v\| + |\lambda_n| \cdot \|v - v_n\| \end{aligned} \tag{153}$$

für alle $n \in \mathsf{N}$ gilt. Wegen Satz 13, existiert ein $C > 0$ mit $|\lambda_n| < C$ für alle $n \in \mathsf{N}$. Für $\varepsilon > 0$ vorgegeben, wählen wir nun $N_1, N_2 \in \mathsf{N}$ mit (beachte, dass $v = 0$ gelten kann)

$$\|\lambda - \lambda_n\| < \frac{\varepsilon}{2 \cdot \max(1, \|v\|)} \quad \forall\, n \geq N_1 \qquad \text{sowie} \qquad \|v - v_n\| < \frac{\varepsilon}{2C} \quad \forall\, n \geq N_2.$$

Für $N_\varepsilon := \max(N_1, N_2)$ und $n \geq N_\varepsilon$, erhalten wir

$$\begin{aligned} \|\lambda \cdot v - \lambda_n \cdot v_n\| &\overset{(153)}{\leq} |\lambda - \lambda_n| \cdot \|v\| + |\lambda_n| \cdot \|(v - v_n)\| \\ &< \frac{\varepsilon}{2 \cdot \max(1, \|v\|)} \cdot \|v\| + C \cdot \frac{\varepsilon}{2C} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

4) Dies sofort folgt aus Teil 1) und Teil 3), wenn man λ und μ als konstante Folgen auffasst. $\square$

Fasst man $\mathbb{K}$ als eindimensionalen $\mathbb{K}$-Vektorraum auf, so ist dieser normiert vermöge $|\cdot|$. Daher erhält man aus Proposition 6 unmittelbar:[22]

Korollar 17. *Seien $\lambda, \mu \in \mathbb{K}$, sowie $(a_n)_{n\in\mathsf{N}} \to a \in \mathbb{K}$, $(b_n)_{n\in\mathsf{N}} \to b \in \mathbb{K}$ konvergente Folgen in $\mathbb{K}$. Dann gelten die folgenden Aussagen:*

1) $(a_n + b_n)_{n\in\mathsf{N}} \to a + b$.

2) $(-a_n)_{n\in\mathsf{N}} \to -a$.

3) $(a_n \cdot b_n)_{n\in\mathsf{N}} \to a \cdot b$.

4) $(\lambda \cdot a_n + \mu \cdot b_n)_{n\in\mathsf{N}} \to \lambda \cdot a + \mu \cdot b$.

Beweis. Folgt unmittelbar aus Proposition 6. □

In $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, gilt nun weiterhin:[23]

Lemma 32. *Sei $(a_n)_{n\in\mathsf{N}} \to a \in \mathbb{K}_\times$ eine konvergente Folge in $\mathbb{K}_\times$.*

a) Dann gilt $\left(\frac{1}{a_n}\right)_{n\in\mathsf{N}} \to \frac{1}{a}$.

b) Ist $(b_n)_{n\in\mathsf{N}} \to b \in \mathbb{K}$ eine konvergent Folge in $\mathbb{K}$, so gilt $\left(\frac{b_n}{a_n}\right)_{n\in\mathsf{N}} \to \frac{b}{a}$.

Beweis. Der Teil b) folgt unmittelbar aus Teil a), sowie Korollar 17.3). Wir erhalten die Aussage in Teil a) wie folgt:

- Sei zunächst $a = 1$. Wir erhalten (beachte: $1 = |x \cdot x^{-1}| = |x| \cdot |x^{-1}|$ für $x \in \mathbb{K}_\times$)

$$|1 - a_n^{-1}| = |(a_n - 1) \cdot a_n^{-1}| = |a_n - 1| \cdot |a_n^{-1}| = |a_n - 1| \cdot |a_n|^{-1}. \tag{154}$$

  Sei nun $\varepsilon > 0$ vorgegeben, und $N_\varepsilon \in \mathsf{N}$ mit $|1 - a_n| < \min(\frac{\varepsilon}{2}, \frac{1}{2})$ für alle $n \geq N_\varepsilon$. Dann folgt für alle $n \geq N_\varepsilon$ (wegen $1 = |1 - a_n + a_n| \leq |1 - a_n| + |a_n|$)

$$1 \leq |1 - a_n| + |a_n| \quad \Rightarrow \quad |a_n| \geq 1 - |1 - a_n| > \frac{1}{2} \quad \Rightarrow \quad |a_n|^{-1} < 2 \quad \overset{(154)}{\Rightarrow} \quad |1 - a_n^{-1}| < \frac{\varepsilon}{2} \cdot 2 = \varepsilon.$$

- Gilt $a \neq 1$, so setzten wir $b_n := a^{-1} \cdot a_n \in \mathbb{K}_\times$ für alle $n \in \mathsf{N}$. Dann gilt $(b_n)_{n\in\mathsf{N}} \to a^{-1} \cdot a = 1$, wegen Korollar 17.4). Aus dem vorhergehenden Punkt folgt somit $\lim_n b_n^{-1} = 1$; und nochmaliges anwenden von Korollar 17.4) liefert

$$a^{-1} = \lim_n (a^{-1} \cdot b_n^{-1}) = \lim_n (a^{-1} \cdot a \cdot a_n^{-1}) = \lim_n a_n^{-1},$$

  was die Behauptung zeigt. □

Beispiel 37. *Sei $(a_n)_{n\in\mathsf{N}}$ eine $\mathbb{K}$-Folge der Gestalt*[24]

$$a_n = \frac{x_k \cdot n^k + x_{k-1} \cdot n^{k-1} + \ldots + x_0}{y_m \cdot n^m + x_{m-1} \cdot n^{m-1} + \ldots + y_0} \qquad \forall\, n \in \mathsf{N}$$

für $k, m \in \mathbb{N}$, sowie $x_0, \ldots, x_k, y_0, \ldots, y_m \in \mathbb{K}$ mit $x_k, y_m \neq 0$. Dann gilt

$$a_n = n^{k-m} \cdot \frac{\overbrace{x_k + \frac{x_{k-1}}{n} + \ldots + \frac{x_0}{n^k}}^{=:\, b_n}}{\underbrace{y_m + \frac{y_{m-1}}{n} + \ldots + \frac{y_0}{n^m}}_{=:\, c_n}} \qquad \forall\, n \in \mathsf{N}.$$

[22] Um Korollar 17 zu erhalten, kann man im Beweis von Proposition 6 alternativ auch einfach $\|\cdot\|$ durch $|\cdot|$ ersetzen.
[23] Gemäß Terminologie 11, ist hier $\mathbb{K}_\times$ als metrischer Raum bezüglich der eingeschränkten Betragsmetrik aufzufassen.
[24] Gemäß Satz 33, besitzt ein Polynom n-ten Grades höchstens n Nullstellen. Wählt man daher $\mathsf{p} \geq 1$ groß genug, so kann man gewährleisten, dass der Nenner ungleich 0 ist.

– *Wegen* $\lim_n \frac{1}{n} = 0$ *und Korollar 17.3), gilt* $\lim_n \frac{1}{n^p} = 0$ *für alle* $p \in \mathbb{N}_{>0}$.

– *Mit Korollar 17 folgt somit* $\lim_n b_n = x_k$ *und* $\lim_n c_n = y_m$.

– *Wegen Lemma 32.b) gilt dann* $\lim_n \frac{b_n}{c_n} = \frac{x_k}{y_m}$.

Es sind die folgenden 3 Fälle zu unterscheiden:

- *Ist* $k = m$, *so gilt* $a_n \to \frac{x_k}{y_m}$.
- *Ist* $k < m$, *so gilt* $n^{k-m} \to 0$ – *also* $a_n \to 0$ *(Korollar 17.3)).*
- *Ist* $k > m$, *so divergiert* $(a_n)_{n\in\mathbb{N}}$.

  In der Tat, andernfalls gilt $a_n \to a$ *für ein* $a \in \mathbb{K}$*; und somit*

  $$\frac{x_k}{y_m} = \lim_n \frac{b_n}{c_n} = \lim_n (n^{m-k} \cdot a_n) = 0 \cdot a = 0$$

  (wegen $\lim_n n^{m-k} = 0$ *und wegen Korollar 17.3)), was* $\frac{x_k}{y_m} \neq 0$ *widerspricht.*

6.1.4 Monotone Folgen

Wir betrachten nun Folgen in $\mathbb{R}$.

Terminologie 15. *Eine Folge* $(a_n)_{n\in\mathsf{N}}$ *in* $\mathbb{R}$ *heißt*

- monoton wachsend $\iff$ $a_n \le a_{n+1}$ $\forall\, n \in \mathsf{N}$,
- streng monoton wachsend $\iff$ $a_n < a_{n+1}$ $\forall\, n \in \mathsf{N}$,
- monoton fallend $\iff$ $a_n \ge a_{n+1}$ $\forall\, n \in \mathsf{N}$,
- streng monoton fallend $\iff$ $a_n > a_{n+1}$ $\forall\, n \in \mathsf{N}$.

Die Folge $(a_n)_{n\in\mathsf{N}}$ *heißt* monoton, *wenn sie monoton wachsend oder monoton fallend ist; sowie* streng monoton, *wenn sie streng monoton wachsend oder streng monoton fallend ist.*

Weiterhin heißt die Folge $(a_n)_{n\in\mathsf{N}}$

- nach unten beschränkt, *wenn ein* $C \le 0$ *mit* $C \le \{a_n \mid n \in \mathbb{N}\}$ *existiert.*
- nach oben beschränkt, *wenn ein* $C \ge 0$ *mit* $\{a_n \mid n \in \mathbb{N}\} \le C$ *existiert.*
- beschränkt, *wenn sie sowohl nach unten, als auch nach oben beschränkt ist.*

  (Dies ist gleichbedeutend zu (152) *in Terminologie 13.)*

Der folgende Satz ist grundlegend in der Analysis.

Satz 14 (Monotone Konvergenz)**.** *Sei* $(a_n)_{n\in\mathsf{N}}$ *eine Folge in* $\mathbb{R}$*. Es gelten die folgenden Aussagen:*

1) *Sei* $(a_n)_{n\in\mathsf{N}}$ nach oben beschränkt *und* monoton wachsend*. Dann ist* $(a_n)_{n\in\mathsf{N}}$ *konvergent in* $\mathbb{R}$*, mit* $\lim_n a_n = \sup_{\mathbb{R}}(\{a_n \mid n \in \mathsf{N}\})$.
2) *Sei* $(a_n)_{n\in\mathsf{N}}$ nach unten beschränkt *und* monoton fallen*. Dann ist* $(a_n)_{n\in\mathsf{N}}$ *konvergent in* $\mathbb{R}$*, mit* $\lim_n a_n = \inf_{\mathbb{R}}(\{a_n \mid n \in \mathsf{N}\})$.
3) *Sei* $(a_n)_{n\in\mathsf{N}}$ beschränkt *und* monoton*. Dann ist* $(a_n)_{n\in\mathsf{N}}$ *konvergent in* $\mathbb{R}$.

Beweis. Der dritte Teil folgt dann unmittelbar aus den ersten beiden Teilen. Wir zeigen hier nur den ersten Teil – Der zweite Teil folgt analog; bzw. aus 1), wenn man $(a_n)_{n\in\mathsf{N}}$ durch $(-a_n)_{n\in\mathsf{N}}$ ersetzt und Proposition 1 anwendet.

Wir argumentieren nun wie folgt:

- Da $(a_n)_{n\in\mathsf{N}}$ nach oben beschränkt ist, existiert $a := \sup_{\mathbb{R}}(\{a_n \mid n \in \mathsf{N}\}) \in \mathbb{R}$. ($\mathbb{R}$ ist vollständig)
- Für $\varepsilon > 0$ vorgegeben, existiert nach Lemma 20 ein $N \in \mathsf{N}$ mit $a_N > a - \varepsilon$. Da $(a_n)_{n\in\mathsf{N}}$ monoton wachsen ist, mit $a_n \le a$ für alle $n \in \mathsf{N}$ (per Definition des Supremums), erhalten wir für alle $n \ge N$, dass $a - \varepsilon < a_N \le a_n \le a$ gilt (also $|a - a_n| < \varepsilon$ für alle $n \ge N$). □

Beispiel 38 (Babylonisches Wurzelziehen). *Für $a > 0$ und $k \in \mathbb{N}_{>0}$, betrachten wir die wie folgt induktiv (rekursiv) definierte $\mathbb{R}$-Folge $(a_n)_{n\in\mathbb{N}}$:*

$$a_0 := a + 1 \qquad \text{und} \qquad a_{n+1} := a_n \cdot \left(1 + \frac{a - (a_n)^k}{k \cdot (a_n)^k}\right) \qquad \forall\, n \in \mathbb{N}.$$

Dann gilt für alle $n \in \mathbb{N}$ (siehe Beweis weiter unten):

$$(i)\ a_n > 0, \qquad (ii)\ (a_n)^k \ge a, \qquad (iii)\ a_{n+1} \le a_n. \tag{155}$$

Daher ist $(a_n)_{n\in\mathbb{N}}$ monoton fallend, sowie nach unten beschränkt durch 0. Satz 14 zeigt $\lim_n a_n \to b$ für ein $b \in \mathbb{R}$; wobei $b \ge 0$ wegen Korollar 14 und (i) gilt.

- *Aus Korollar 17, Korollar 14, und (ii) folgt, dass $b^k = \lim_n (a_n)^k \ge a > 0$ gilt (also auch $b > 0$).*
- *Durch Limesbildung (auf beiden Seiten) erhalten wir aus den Grenzwertsätzen*

$$a_{n+1} = a_n \cdot \left(1 + \frac{a - (a_n)^k}{k \cdot (a_n)^k}\right) \qquad \Longrightarrow \qquad b = b \cdot \left(1 + \frac{a - b^k}{k \cdot b^k}\right).$$

Kürzen und umstellen liefert $1 = 1 + \frac{a-b^k}{k \cdot b^k}$; also $b^k = a$. Dies zeigt $b = \sqrt[k]{a}$.

Beweis von (155) *per Induktion:*

(IA) Es gilt $a_0 = a + 1 > 0$ (wegen $a > 0$), also *(i).*
Zudem gilt $(a_0)^k = (a+1)^k > a + 1 > a$ (wegen Übung 46.b) und $a + 1 > 1$), *(ii).*
Schließlich gilt *(iii)*, wegen ($a_0 \ge 0$ und $a/(a_0)^k \le 1$)

$$a_1 = a_0 \cdot \left(1 + \frac{a - (a_0)^k}{k \cdot (a_0)^k}\right) = a_0 \cdot \left(1 - \frac{1}{k} + \frac{1}{k} \cdot \frac{a}{(a_0)^k}\right) \le a_0 \cdot \left(1 - \frac{1}{k} + \frac{1}{k}\right) = a_0.$$

(IS) Es gelte nun (155) für ein $n \in \mathbb{N}$. Dann folgt:

(i) Wir haben $a_{n+1} > 0$, wegen $a_n > 0$ sowie (beachte $k \ge 1$ und $a/(a_n)^k \ge 1 \ge 0$)

$$1 + \frac{a - (a_n)^k}{k \cdot (a_n)^k} \ge \frac{(k-1) \cdot (a_n)^k + a}{k \cdot (a_n)^k} = \frac{k-1}{k} > 0.$$

(ii) Wir erhalten aus Satz 8 (Bernoulli-Ungleichung)

$$\begin{aligned}(a_{n+1})^k &= (a_n)^k \cdot \left(1 + \frac{a - (a_n)^k}{k \cdot (a_n)^k}\right)^k \\ &\ge (a_n)^k \cdot \left(1 + k \cdot \frac{a - (a_n)^k}{k \cdot (a_n)^k}\right) \\ &= (a_n)^k + a - (a_n)^k = a;\end{aligned}$$

denn nach (IV) gilt $\frac{a-(a_n)^k}{k \cdot (a_n)^k} = \frac{a}{k \cdot (a_n)^k} - \frac{(a_n)^k}{k \cdot (a_n)^k} \ge -\frac{(a_n)^k}{k \cdot (a_n)^k} \ge -1$ (vgl. Voraussetzungen von Satz 8).

(iii) Wir erhalten aus (i) (also $a_{n+1} > 0$) und (ii) (also $(a_{n+1})^k \geq a$), dass

$$a_{n+2} = a_{n+1} \cdot \left(1 + \frac{\overbrace{a - (a_{n+1})^k}^{\leq 0}}{k \cdot (a_{n+1})^k}\right) \leq a_{n+1}$$

gilt, was die Behauptung zeigt. □

Übung 67. *Die $\mathbb{R}$-Folge $(a_n)_{n\in\mathbb{N}}$ sei rekursiv definiert durch*

$$a_0 := 1 \qquad \textit{und} \qquad a_{n+1} = 1 + \frac{a_n}{2} \quad \forall\, n \in \mathbb{N}. \tag{156}$$

Zeigen Sie:

a) Es gilt $1 \leq a_n \leq 2$ sowie $a_n \leq a_{n+1}$ für alle $n \in \mathbb{N}$.

b) Folgern Sie, dass $(a_n)_{n\in\mathbb{N}}$ konvergiert.

c) Bestimmen Sie den Grenzwert von $(a_n)_{n\in\mathbb{N}}$.

Lemma 33. *Jede Folge in $\mathbb{R}$ besitzt eine monotone Teilfolge.*

Beweis. Sei $(a_n)_{n\in\mathsf{N}}$ eine Folgen in $\mathbb{R}$ mit $\mathsf{N} = \{n \in \mathbb{Z} \,|\, n \geq \mathsf{p}\}$. Wir setzen

$$\mathcal{N} := \{n \in \mathsf{N} \,|\; a_n \geq a_{n+m} \;\text{ für alle }\; m \in \mathbb{N}\,\} \subseteq \mathsf{N}.$$

- Sei $\mathcal{N}$ nicht endlich:

  Wir erhalten eine streng monotone[25] Injektion $\iota\colon \mathbb{N} \to \mathcal{N}$, indem wir induktiv definieren:

  $$\iota(0) := \min(\mathcal{N}) \qquad \text{sowie} \qquad \iota(n+1) := \min(\{p \in \mathcal{N} \,|\, p > \iota(n)\}) \quad \text{für alle} \quad n \in \mathbb{N}.$$

  Die Teilfolge $(a_{\iota(n)})_{n\in\mathsf{N}}$ ist dann per Konstruktion monoton fallend.

- Sei $\mathcal{N}$ endlich: Wir setzen

  $$\iota_0 := \begin{cases} \mathsf{p} & \text{für } \mathcal{N} = \emptyset \\ \max(\mathcal{N}) + 1 & \text{für } \mathcal{N} \neq \emptyset. \end{cases}$$

  - Wegen $\iota_0 \notin \mathcal{N}$, existiert ein $\mathsf{N} \ni \iota_1 > \iota_0$ mit $a_{\iota_1} > a_{\iota_0}$.
  - Wir wenden nun das gleiche Argument auf ι_1 an (beachte $\iota_1 \notin \mathcal{N}$ wegen $\mathcal{N} < \iota_0 < \iota_1$), und erhalten $\mathcal{N} < \iota_0 < \iota_1 < \iota_2$ mit $a_{\iota_0} < a_{\iota_1} < a_{\iota_2}$.
  - Induktive Anwendung des gleichen Argumentes, liefert $\mathcal{N} < \iota_0 < \iota_1 < \iota_2 < \dots$ mit $a_{\iota_0} < a_{\iota_1} < a_{\iota_2} < \dots$.

  Sei $\iota\colon \mathbb{N} \ni n \mapsto \iota_n \in \mathsf{N}$. Dann ist $(a_{\iota(n)})_{n\in\mathsf{N}}$ eine monoton wachsende Teilfolge von $(a_n)_{n\in\mathsf{N}}$. □

Satz 15 (Bolzano-Weierstrass)**.** *Jede beschränkte Folge in $\mathbb{R}$, besitzt eine konvergente Teilfolge.*

Beweis. Sei $(a_n)_{n\in\mathsf{N}}$ eine beschränkte Folge in $\mathbb{R}$. Mit Lemma 33 existiert eine monotone Teilfolge $(a_{\iota(n)})_{n\in\mathbb{N}}$ von $(a_n)_{n\in\mathsf{N}}$. Offensichtlich ist dann $(a_{\iota(n)})_{n\in\mathbb{N}}$ ebenfalls beschränkt; konvergiert also gegen ein Element in $\mathbb{R}$ wegen Satz 14.3) (Monotone Konvergenz). □

Korollar 18. *Sei $a, b \in \mathbb{R}$ mit $a \leq b$, sowie $(x_n)_{n\in\mathsf{N}}$ eine Folge in $[a, b]$. Dann besitzt $(x_n)_{n\in\mathsf{N}}$ eine in $[a, b]$ konvergente Teilfolge.*

[25]Für $\mathbb{N} \ni m > n \in \mathsf{N}$ gilt $\iota(m) > \iota(n)$.

Beweis. Wegen Satz 15 besitzt $(x_n)_{n\in\mathsf{N}}$ eine in $\mathbb{R}$ konvergente Teilfolge $(x_{\iota(n)})_{n\in\mathbb{N}} \to x \in \mathbb{R}$. Wegen Übung 61 (siehe auch Korollar 14), gilt dann bereits $x \in [a, b]$. $\square$

Bemerkung 40. *Sei $(x_n)_{n\in\mathsf{N}}$ eine Folge in $\mathbb{R}$, mit $\mathsf{N} = \{n \in \mathbb{Z} \mid n \geq \mathsf{p}\}$.*

- *Ist $(x_n)_{n\in\mathsf{N}}$* nicht *nach oben beschränkt, so existiert zu $C \geq 0$ und $N \in \mathsf{N}$, ein $\ell > N$ mit $x_\ell > C$.*

  (Andernfalls existieren $C \geq 0$ und $N \in \mathsf{N}$, mit $x_\ell \leq C$ für alle $\ell > N$; also $x_n \leq \max(x_{\mathsf{p}}, \ldots, x_N, C)$ für alle $n \in \mathsf{N}$, was der Beschränktheit widerspricht.)

- *Ist $(x_n)_{n\in\mathsf{N}}$* nicht *nach unten beschränkt, so existiert zu $C \leq 0$ und $N \in \mathsf{N}$, ein $\ell > N$ mit $x_\ell < C$.*

  (Andernfalls existieren $C \leq 0$ und $N \in \mathsf{N}$, mit $x_\ell \geq C$ für alle $\ell > N$; also $x_n \geq \min(x_{\mathsf{p}}, \ldots, x_N, C)$ für alle $n \in \mathsf{N}$, was der Beschränktheit widerspricht..)

Unter Zuhilfenahme von Bemerkung 40, erhalten wir aus Satz 15:

Korollar 19. *Jede reelle Folge besitzt eine in $\overline{\mathbb{R}}$ konvergente Teilfolge.*

Beweis. Sei $(a_n)_{n\in\mathsf{N}}$ eine Folge in $\mathbb{R}$ mit $\mathsf{N} = \{n \in \mathbb{Z} \mid n \geq \mathsf{p}\}$.

- Ist $(a_n)_{n\in\mathsf{N}}$ beschränkt, so folgt die Behauptung aus Satz 15.
- Ist $(a_n)_{n\in\mathsf{N}}$ nicht nach oben beschränkt, so definieren wir induktiv $\iota(0) := \mathsf{p}$ und

$$\iota(n+1) := \min(\{\ell > \iota(n) \mid a_\ell \geq n\}) \qquad \forall\, n \in \mathbb{N}.$$

  Dann ist $(a_{\iota(n)})_{n\in\mathbb{N}}$ eine Teilfolge von $(a_n)_{n\in\mathsf{N}}$ mit $(a_{\iota(n)})_{n\in\mathbb{N}} \to \infty$.

- Ist $(a_n)_{n\in\mathsf{N}}$ nicht nach unten beschränkt, so definieren wir induktiv $\iota(0) := \mathsf{p}$ und

$$\iota(n) := \min(\{\ell > \iota(n) \mid a_\ell \leq -n\}) \qquad \forall\, n \in \mathbb{N}.$$

  Dann ist $(a_{\iota(n)})_{n\in\mathbb{N}}$ eine Teilfolge von $(a_n)_{n\in\mathsf{N}}$ mit $(a_{\iota(n)})_{n\in\mathbb{N}} \to -\infty$. $\square$

6.1.5 Häufungspunkte von Folgen

Definition 38 (Häufungspunkt einer Folge)**.** *Sei (X, d) ein metrischer Raum, und $(x_n)_{n\in\mathsf{N}}$ eine Folge in X. Ein Punkt $x \in X$ heißt* Häufungspunkt *(Limespunkt) von $(x_n)_{n\in\mathsf{N}}$, wenn eine Teilfolge von $(x_n)_{n\in\mathsf{N}}$ existiert, die gegen x konvergiert.*

(Die Menge der Häufungspunkte zu einer gegebenen Folge, ist also die Menge aller Grenzwerte von Teilfolgen dieser Folge.)

Beispiel 39. *Sei (X, d) ein metrischer Raum, und $(x_n)_{n\in\mathsf{N}} \to x \in X$ eine konvergente Folge in X. Dann hat $(x_n)_{n\in\mathsf{N}}$ genau einen Häufungspunkt, und zwar x (siehe Korollar 16: Jede Teilfolge einer konvergenten Folge konvergiert gegen den gleichen Grenzwert).*

Wir betrachten nun reelle Zahlenfolgen, und beachten, dass dann Häufungspunkte per definitionem Elemente von $\mathbb{R}$ sind (nicht von $\overline{\mathbb{R}}$):

Beispiel 40. *Die Folge $(a_n)_{n\in\mathbb{N}} = ((-1)^n)_{n\in\mathbb{N}}$ besitzt die Häufungspunkte 1 und -1; denn es gilt $(a_{2n})_{n\in\mathbb{N}} = (1)_{n\in\mathbb{N}} \to 1$ und $(a_{2n+1})_{n\in\mathbb{N}} = (-1)_{n\in\mathbb{N}} \to -1$. Zudem existiert kein weiterer Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$; denn für $a \notin \{-1, 1\}$ gilt*

$$\varepsilon := \min(|1-a|, |-1-a|) > 0 \qquad \textit{mit} \qquad \pm\, 1 \notin (a-\varepsilon, a+\varepsilon),$$

also $a_n \notin (a-\varepsilon, a+\varepsilon)$ für alle $n \in \mathbb{N}$.

Der folgende Satz klärt einige grundlegende Eigenschaften von Häufungspunkten von reellen Zahlenfolgen:

Satz 16. *Für reelle Zahlenfolgen gelten die folgenden Aussagen:*

1) Jede beschränkte reelle Folge hat mindestens einen Häufungspunkt.

2) Eine konvergente Folge hat genau einen Häufungspunkt.

3) Eine Folge ist genau dann konvergent, wenn sie beschränkte ist und genau einen Häufungspunkt besitzt.

4) Ein Punkt $a \in \mathbb{R}$ ist genau dann ein Häufungspunkt der Folge $(a_n)_{n\in\mathsf{N}}$, wenn jede Umgebung von a unendlich viele Folgenglieder enthält.

Erinnerung: Eine Umgebung $U \subseteq \mathbb{R}$ von a, enthält per definitionem einen metrischen Ball um a, d.h., es gilt $\mathsf{B}_\varepsilon(x) = (a-\varepsilon, a+\varepsilon) \subseteq U$ für ein $\varepsilon > 0$. Insbesondere ist jeder metrische Ball $\mathsf{B}_\varepsilon(x)$ um x $(\varepsilon > 0)$ eine Umgebung von x, da er sich ja selbst enthält.

Beweis. 1) Das ist die Aussagen von Satz 15 (Bolzano-Weierstrass). (Jede beschränkte reelle Folge hat eine reell konvergente Teilfolge.)

2) Das ist die Aussage von Beispiel 39.

3) Gemäß Satz 13, ist jede konvergente Folge beschränkt; und wegen 2) hat diese genau einen Häufungspunkt. Es bleibt daher nur noch nachzuweisen, dass eine beschränkte Folge $(a_n)_{n\in\mathsf{N}}$ $(\mathsf{N} = \{n \in \mathbb{Z} \mid n \geq \mathsf{p}\})$ konvergiert, wenn sie nur einen Häufungspunkt a hat:

Angenommen, es gilt nicht $(a_n)_{n\in\mathsf{N}} \to a$. Dann existiert ein $\varepsilon > 0$ mit

$$\forall\, m \in \mathsf{N}\colon\ \exists\, N[m] > m\colon\ a_{N[m]} \notin (a-\varepsilon, a+\varepsilon). \tag{157}$$

Dann tritt (mindestens) einer der folgenden beiden Fälle auf:[26]

$$\begin{aligned} &\forall\, m \in \mathsf{N}\colon\ \exists\, N[m] > m\colon\ a_{N[m]} \geq a+\varepsilon \\ &\forall\, m \in \mathsf{N}\colon\ \exists\, N[m] > m\colon\ a_{N[m]} \leq a-\varepsilon. \end{aligned}$$

Im ersten Fall (der zweite Fall folgt analog), erhalten wir eine Teilfolge $(a_{\iota(n)})_{n\in\mathbb{N}}$ in $[a+\varepsilon, \infty)$, indem wir induktiv definieren:

$$\iota(0) := N[\mathsf{p}] \qquad \text{sowie} \qquad \iota(n+1) := N[\iota(n)] \qquad \forall\, n \in \mathbb{N}.$$

Dann ist $(a_{\iota(n)})_{n\in\mathbb{N}}$ beschränkt, weil $(a_n)_{n\in\mathsf{N}}$ beschränkt ist; hat also nach Satz 15 eine konvergente Teilfolge $(a_{\kappa(\iota(n))})_{n\in\mathbb{N}} \to b \in \mathbb{R}$:

- Wegen Korollar 14 gilt $b \geq a+\varepsilon \neq a$.
- Wegen Bemerkung 37 ist $(a_{(\kappa\circ\iota)(n)})_{n\in\mathbb{N}}$ auch eine Teilfolge von $(a_n)_{n\in\mathsf{N}}$. Daher ist b ein Häufungspunkt von $(a_n)_{n\in\mathsf{N}}$. Wegen $b \neq a$, widerspricht dies nun der Voraussetzung, dass $(a_n)_{n\in\mathsf{N}}$ nur einen Häufungspunkt besitzt.

4) • Sei a ein Häufungspunkt von $(a_n)_{n\in\mathsf{N}}$, sowie $U \subseteq \mathbb{R}$ eine Umgebung von a, und $\varepsilon > 0$ mit $(a-\varepsilon, a+\varepsilon) = \mathsf{B}_\varepsilon(a) \subseteq U$. Per definitionem, existiert eine Teilfolge von $(a_n)_{n\in\mathsf{N}}$ mit $(a_{\iota(n)})_{n\in\mathsf{N}} \to a$. Daher existiert ein $N_\varepsilon \in \mathbb{N}$ mit $a_{\iota(n)} \in \mathsf{B}_\varepsilon(a) = (a-\varepsilon, a+\varepsilon)$ für alle $n \geq N_\varepsilon$.

[26] Gelten beide Zeilen nicht, so existieren $m_1, m_2 \in \mathsf{N}$ mit $a_p \geq a+\varepsilon$ für alle $p > m_1$, sowie $a_q \leq a-\varepsilon$ für alle $q > m_2$; also $a_\ell \in (a-\varepsilon, a+\varepsilon)$ für alle $\ell > N := \max(m_1, m_2)$, was (157) widerspricht.

- Es enthalte jede Umgebung U von a unendlich viele Folgenglieder von $(a_n)_{n\in\mathsf{N}}$. Dann gilt

$$\forall\,(m \in \mathsf{N} \ \wedge\ n \in \mathbb{N}_{>0})\colon\ \exists\, N[m,n] > m\colon\ a_{N[m,n]} \in \mathsf{B}_{\frac{1}{n}}(a) = \big(a - \tfrac{1}{n}, a + \tfrac{1}{n}\big).$$

(*Beachte: Die Negation obiger Aussage ist*

$$\exists\,(m \in \mathsf{N} \ \wedge\ n \in \mathbb{N}_{>0})\colon\ \forall\, N > m\colon\ a_N \notin \mathsf{B}_{\frac{1}{n}}(a) = \big(a - \tfrac{1}{n}, a + \tfrac{1}{n}\big). \quad)$$

Wir erhalten eine Teilfolge $(a_{\iota(n)})_{n\in\mathbb{N}}$, indem wir induktiv definieren: $\big(a_{\iota(n+1)} \in \mathsf{B}_{\frac{1}{n}}(a)\big)$

$$\iota(0) := N[\mathsf{p}, 1] \qquad \text{sowie} \qquad \iota(n+1) := N[\iota(n), n] \qquad \forall\, n \in \mathbb{N}.$$

Dann gilt $(a_{\iota(n)})_{n\in\mathbb{N}} \to a$; denn ist $\varepsilon > 0$ vorgegeben, so gilt für $N_\varepsilon \in \mathbb{N}_{>0}$ mit $\frac{1}{N_\varepsilon} < \varepsilon$:

$$a_{\iota(n+1)} = a_{N[\iota(n),n]} \in \mathsf{B}_{\frac{1}{n}}(a) \subseteq \mathsf{B}_{\frac{1}{N_\varepsilon}}(a) \subseteq \mathsf{B}_\varepsilon(a) \qquad \forall\, n \geq N_\varepsilon.$$

Dies zeigt, dass a ein Häufungspunkt von $(a_n)_{n\in\mathsf{N}}$ ist. □

Bemerkung* 9. *Satz 16 gilt in der gleichen Form auch für Folgen in $\mathbb{C}$. Die Beweise von Teil 2) und Teil 4) sind komplett analog; und die anderen beiden Teile erhält man wie folgt:*

Beweis von Satz 16.1) für $\mathbb{C}$: Sei $(z_n)_{n\in\mathsf{N}} = (x_n + \mathrm{i}y_n)_{n\in\mathsf{N}}$ eine beschränkte Folge in $\mathbb{C}$. Dann sind $(x_n)_{n\in\mathsf{N}}$ und $(y_n)_{n\in\mathsf{N}}$ beschränkte Folgen in $\mathbb{R}$ (vgl. Übung 63.a)). Wegen Satz 15, existiert eine konvergente Teilfolge $(x_{\iota(n)})_{n\in\mathbb{N}} \to x \in \mathbb{R}$ (von $(x_n)_{n\in\mathsf{N}}$). Nach selbigem Satz, existiert eine konvergente Teilfolge $(y_{\kappa(\iota(n))})_{n\in\mathbb{N}} \to y \in \mathbb{R}$ (von $(y_{\iota(n)})_{n\in\mathbb{N}}$). Wegen Korollar 16, gilt dann ebenfalls $(x_{\kappa(\iota(n))})_{n\in\mathbb{N}} \to x$; und daher $(z_{\kappa(\iota(n))})_{n\in\mathbb{N}} \to x + \mathrm{i}y \in \mathbb{C}$ wegen Übung 63.c). □

Beweis von Satz 16.3) für $\mathbb{C}$: Sei $(z_n)_{n\in\mathsf{N}} = (x_n + \mathrm{i}y_n)_{n\in\mathsf{N}}$ eine Folge in $\mathbb{C}$.

- Konvergiert $(z_n)_{n\in\mathsf{N}}$ in $\mathbb{C}$, so ist sie beschränkt wegen Satz 13; und hat genau einen Häufungspunkt wegen Beispiel 39.
- Es sei $(z_n)_{n\in\mathsf{N}}$ beschränkt, mit genau einen Häufungspunkt. Dann sind $(x_n)_{n\in\mathsf{N}}$ und $(y_n)_{n\in\mathsf{N}}$ beschränkte Folgen in $\mathbb{R}$ (vgl. Übung 63). Wegen Satz 16.3) (reeller Fall) und Übung 63.c), reicht es nun zu zeigen, dass $(x_n)_{n\in\mathsf{N}}$ und $(y_n)_{n\in\mathsf{N}}$ ebenfalls genau einen Häufungspunkt haben:

  Angenommen, $(x_n)_{n\in\mathsf{N}}$ besitzt die Häufungspunkte $\mathbb{R} \ni x_1 \neq x_2 \in \mathbb{R}$ (für $(y_n)_{n\in\mathsf{N}}$ argumentiert man analog). Dann existieren Teilfolgen $(x_{\iota_p(n)})_{n\in\mathbb{N}} \to x_p$ mit $p = 1, 2$. Die Teilfolgen $(y_{\iota_p(n)})_{n\in\mathbb{N}}$ (von $(y_n)_{n\in\mathsf{N}}$) sind beschränkt, und haben daher konvergente Teilfolgen $(y_{\kappa_p(\iota_p(n))})_{n\in\mathbb{N}} \to y_p \in \mathbb{R}$ $(p = 1, 2)$. Wegen Korollar 16, gilt nun ebenfalls $(x_{\kappa_p(\iota_p(n))})_{n\in\mathbb{N}} \to x_p \in \mathbb{R}$ für $p = 1, 2$; und mit Übung 63.c) folgt

$$(z_{\kappa_p(\iota_p(n))})_{n\in\mathbb{N}} = (x_{\kappa_p(\iota_p(n))} + \mathrm{i}y_{\kappa_p(\iota_p(n))})_{n\in\mathbb{N}} \to x_p + \mathrm{i}y_p \qquad \text{für} \qquad p = 1, 2.$$

  Die widerspricht der Voraussetzung , dass $(z_n)_{n\in\mathsf{N}}$ genau einen Häufungspunkt hat. □

Beispiel 41.

a) Die Folge $(a_n)_{n\in\mathbb{N}_{>0}} = ((-1)^n + \frac{1}{n})_{n\in\mathbb{N}_{>0}}$ besitzt die Häufungspunkte 1 und -1; denn es gilt

$$(a_{2n})_{n\in\mathbb{N}} = (1 + \tfrac{1}{2n})_{n\in\mathbb{N}} \to 1 \qquad \textit{und} \qquad (a_{2n+1})_{n\in\mathbb{N}} = (-1 + \tfrac{1}{2n+1})_{n\in\mathbb{N}} \to -1.$$

Eine leichte Variation des Arguments in Beispiel 40 zeigt, dass auch $(a_n)_{n\in\mathbb{N}}$ keine weiteren Häufungspunkte hat (Übung 68).

b) Die Folge $(a_n)_{n\in\mathbb{N}} = (n)_{n\in\mathbb{N}}$ besitzt keinen Häufungspunkt, da jede Teilfolge unbeschränkt ist; und daher wegen Satz 13 nicht konvergieren kann.

c) Die Folge

$$a_n = \begin{cases} n & \textit{falls } n \textit{ gerade} \\ \frac{1}{n} & \textit{falls } n \textit{ ungerade} \end{cases}$$

ist unbeschränkt, hat aber wegen $(a_{2n+1})_{n\in\mathbb{N}} = (\frac{1}{2n+1})_{n\in\mathbb{N}} \to 0$ *den Häufungspunkt* 0.

Übung 68. *Machen Sie sich klar, dass die Folge aus Beispiel 41.a) nur die Häufungspunkte* $-1, 1$ *hat. Machen Sie sich weiterhin klar, dass die Folge aus Beispiel 41.c) nur den Häufungspunkt* 0 *hat.*

Terminologie 16 (Erweiterterte reelle Zahlen)**.** *Wir erinnern an die erweiterten reellen Zahlen* $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ *aus Definition 37.*

- *Wir definieren eine Anordnung auf* $\overline{\mathbb{R}}$*, indem wir setzen:*

  $$-\infty < x < \infty \qquad \forall\, x \in \mathbb{R} \qquad \textit{sowie} \qquad -\infty < \infty.$$

  Weiterhin setzen wir $[-\infty, \infty] := \overline{\mathbb{R}}$*, sowie für* $z \in \mathbb{R}$*:*

  $$\begin{aligned} [-\infty, z) &:= \{x \in \overline{\mathbb{R}} \mid -\infty \le x < z\} = \{-\infty\} \cup (-\infty, z) \\ (z, \infty] &:= \{x \in \overline{\mathbb{R}} \mid z < x \le \infty\} \quad = (z, \infty) \cup \{\infty\}. \end{aligned}$$

- *Hiermit gelten die folgenden Regeln für Suprema und Infima:*
  Für eine Teilmenge $\emptyset \neq M \subseteq \mathbb{R}$*, gilt (vgl. Notation 17)*

  $$\sup_{\overline{\mathbb{R}}}(M) = \sup_{\mathbb{R}}(M) = \begin{cases} \sup_{\mathbb{R}}(M) & \textit{falls } M \textit{ nach oben beschränkt ist,} \\ \infty & \textit{falls } M \textit{nicht nach oben beschränkt ist.} \end{cases}$$

  $$\inf_{\overline{\mathbb{R}}}(M) = \inf_{\mathbb{R}}(M) = \begin{cases} \inf_{\mathbb{R}}(M) & \textit{falls } M \textit{ nach unten beschränkt ist,} \\ -\infty & \textit{falls } M \textit{nicht nach unten beschränkt ist.} \end{cases}$$

  Weiterhin gilt

  $$\begin{aligned} \sup_{\overline{\mathbb{R}}}(\{-\infty\}) &= -\infty & \textit{sowie} \qquad & \inf_{\overline{\mathbb{R}}}(\{\infty\}) = \infty \\ \sup_{\overline{\mathbb{R}}}(\{-\infty\} \cup M) &= \sup_{\overline{\mathbb{R}}}(M) & \textit{sowie} \qquad & \inf_{\overline{\mathbb{R}}}(M \cup \{\infty\}) = \inf_{\overline{\mathbb{R}}}(M) \end{aligned}$$

  für $\emptyset \neq M \subseteq \mathbb{R}$*; sowie*

  $$\begin{aligned} \sup_{\overline{\mathbb{R}}}(M) &= \infty & \textit{falls} \qquad & \infty \in M \subseteq \overline{\mathbb{R}} \\ \inf_{\overline{\mathbb{R}}}(M) &= -\infty & \textit{falls} \qquad & -\infty \in M \subseteq \overline{\mathbb{R}}. \end{aligned}$$

Definition 39 (Limes superior und Limes inferior)**.** *Für eine reelle Folge* $(a_n)_{n\in\mathsf{N}}$*, definieren wir*

$$\begin{aligned} &\text{Limes superior:} \qquad \overline{\lim}_n\, a_n := \inf_{\overline{\mathbb{R}}}\big(\{x \in \overline{\mathbb{R}} \mid \exists\, N \in \mathsf{N} \mid a_n \le x \quad \forall\, n \ge N\}\big) \in \overline{\mathbb{R}} \\ &\text{Limes inferior:} \qquad \underline{\lim}_n\, a_n := \sup_{\overline{\mathbb{R}}}\big(\{x \in \overline{\mathbb{R}} \mid \exists\, N \in \mathsf{N} \mid x \le a_n \quad \forall\, n \ge N\}\big) \in \overline{\mathbb{R}}. \end{aligned} \tag{158}$$

In anderen Worten:

- *Es ist* $\overline{\lim}_n\, a_n$ *das Infimum (größte untere Schranke) in* $\overline{\mathbb{R}}$*, der Menge*

  $$O := \{x \in \overline{\mathbb{R}} \mid a_n \le x \ \textit{ für fast alle } n \in \mathsf{N}\} \tag{159}$$

  der quasi-oberen Schranken von $(a_n)_{n\in\mathsf{N}}$.
  Wegen $\infty \in O \neq \emptyset$ *(*$a_n < \infty$ *für alle* $n \in \mathsf{N}$*), ist dann* $\overline{\lim}_n\, a_n = \inf_{\overline{\mathbb{R}}}(O) \le \infty$ *definiert.*

  Beispielsweise:

– $(a_n)_{n\in\mathsf{N}} = (\frac{1}{n})_{n\in\mathbb{N}_{>0}}$: *Es gilt* $O = (0,\infty]$, *also* $\overline{\lim}_n a_n = \inf_{\overline{\mathbb{R}}}(O) = 0$.
– $(a_n)_{n\in\mathsf{N}} = (n)_{n\in\mathbb{N}}$: *Es gilt* $O = \{\infty\}$, *also* $\overline{\lim}_n a_n = \inf_{\overline{\mathbb{R}}}(O) = \infty$.
– $(a_n)_{n\in\mathsf{N}} = (-n)_{n\in\mathbb{N}}$: *Es gilt* $O = (-\infty,\infty]$, *also* $\overline{\lim}_n a_n = \inf_{\overline{\mathbb{R}}}(O) = -\infty$.

- *Es ist* $\underline{\lim}_n a_n$ *das Supremum (die kleinste obere Schranke) in* $\overline{\mathbb{R}}$, *der Menge*

$$U := \{x \in \overline{\mathbb{R}} \mid x \leq a_n \ \text{für fast alle } n \in \mathsf{N}\}. \tag{160}$$

der quasi-unteren Schranken von $(a_n)_{n\in\mathsf{N}}$.
Wegen $-\infty \in U \neq \emptyset$ *(*$-\infty < a_n$ *für alle* $n \in \mathsf{N}$*), ist dann* $\underline{\lim}_n a_n = \sup_{\overline{\mathbb{R}}}(U) \geq -\infty$ *definiert.*

Beispielsweise:

– $(a_n)_{n\in\mathsf{N}} = (-\frac{1}{n})_{n\in\mathbb{N}_{>0}}$: *Es gilt* $U = [-\infty,0)$, *also* $\underline{\lim}_n a_n = \sup_{\overline{\mathbb{R}}}(U) = 0$.
– $(a_n)_{n\in\mathsf{N}} = (-n)_{n\in\mathbb{N}}$: *Es gilt* $U = \{-\infty\}$, *also* $\underline{\lim}_n a_n = \sup_{\overline{\mathbb{R}}}(U) = -\infty$.
– $(a_n)_{n\in\mathsf{N}} = (n)_{n\in\mathbb{N}}$: *Es gilt* $U = [-\infty,\infty)$, *also* $\underline{\lim}_n a_n = \sup_{\overline{\mathbb{R}}}(O) = \infty$.

Beispiel 42.

a) *Für* $(a_n)_{n\in\mathbb{N}} = ((-1)^n)_{n\in\mathbb{N}}$ *gilt* $\overline{\lim}_n a_n = 1$ *und* $\underline{\lim}_n a_n = -1$.
b) *Für* $(a_n)_{n\in\mathbb{N}} = ((-1)^n + \frac{1}{n})_{n\in\mathbb{N}}$ *gilt* $\overline{\lim}_n a_n = 1$ *und* $\underline{\lim}_n a_n = -1$.

Bemerkung* 10. *Eine alternative Definition des Limes superior und des Limes inferior ist*

$$\overline{\lim}_n a_n := \lim_n \sup_{\overline{\mathbb{R}}}\big(\{m \geq n \mid a_m\}\big) \in \overline{\mathbb{R}}$$
$$\underline{\lim}_n a_n := \lim_n \ \inf_{\overline{\mathbb{R}}}\big(\{m \geq n \mid a_m\}\big) \in \overline{\mathbb{R}}.$$

Hierfür muss man dann noch zusätzliche Konvention bezüglich der Konvergenz von Folgen in $\overline{\mathbb{R}}$ *einführen; denn beispielsweise gilt für die Folge* $(a_n)_{n\in\mathbb{N}} = (n)_{n\in\mathbb{N}}$ *ja* $\sup_{\mathbb{R}}(\{m \geq n \mid a_m\}) = \infty$ *für alle* $n \geq \mathbb{N}$.

Übung 69. *Sei* $(a_n)_{n\in\mathsf{N}}$ *eine reelle Folge, sowie* $c \geq 0$. *Zeigen Sie, dass*

$$\overline{\lim}_n(c\cdot a_n) = c\cdot\overline{\lim}_n a_n \qquad \text{sowie} \qquad \underline{\lim}_n(c\cdot a_n) = c\cdot\underline{\lim}_n a_n$$

gilt; sofern man $0\cdot\pm\infty := 0$, *sowie* $c\cdot\pm\infty := \pm\infty$ *für* $c > 0$ *setzt.*

Lemma 34. *Für eine reelle Folge* $(a_n)_{n\in\mathsf{N}}$ *gelten die folgenden Implikationen:*

a) $\overline{\lim}_n a_n \in \mathbb{R} \implies \overline{\lim}_n a_n$ *ist größter Häufungspunkt von* $(a_n)_{n\in\mathsf{N}}$.
b) $\underline{\lim}_n a_n \in \mathbb{R} \implies \underline{\lim}_n a_n$ *ist kleinster Häufungspunkt von* $(a_n)_{n\in\mathsf{N}}$.

Beweis. a) Sei $o := \overline{\lim}_n a_n$, und O wie in (159). Wir bemerken zunächst, dass für $\tau \in \mathbb{R}$ gilt:

$$\forall\, N \in \mathsf{N}\colon \exists\, n \geq N\colon \tau \leq a_n \qquad\implies\qquad \tau \leq O. \tag{161}$$

(Denn für jedes $x \in O$, existiert ja gemäß (158) ein $N \in \mathsf{N}$ mit $a_n \leq x$ für alle $n \geq N$.)

- Angenommen, o ist kein Häufungspunkt von $(a_n)_{n\in\mathsf{N}}$. Wegen Satz 16.4), existiert dann eine Umgebung von o, die nur endlich viele Folgenglieder a_n enthält. Somit existiert ein $\varepsilon > 0$, sodass $(o-\varepsilon, o+\varepsilon)$ nur endlich viele Folgenglieder a_n enthält. Daher existiert ein $\tilde{N} \in \mathsf{N}$ mit

$$(\, a_n \leq o-\varepsilon \quad\vee\quad a_n \geq o+\varepsilon\,) \qquad \forall\, n \geq \tilde{N}.$$

Es gibt nun zwei Möglichkeiten:

– Es existiert ein $N \geq \tilde{N}$ mit $a_n \leq o - \varepsilon$ für alle $n \geq N$. Wir erhalten

$$a_n \leq o - \varepsilon \quad \forall\, n \geq N \qquad \Longrightarrow \qquad o - \varepsilon \in O \qquad \Longrightarrow \qquad o = \inf_{\overline{\mathbb{R}}}(O) \leq o - \varepsilon < o,$$

also den Widerspruch $o < o$.

– Für jedes $N \geq \tilde{N}$, existiert ein $n \geq N$ mit

$$a_n \not\leq o - \varepsilon \qquad \overset{n \geq N \geq \tilde{N}}{\Longleftrightarrow} \qquad a_n \geq o + \varepsilon \qquad \Longleftrightarrow \qquad \tau := o + \varepsilon \leq a_n.$$

Dann gilt die linke Seiten von (161), also $o + \varepsilon = \tau \leq O$. Wir erhalten den Widerspruch $o < o + \varepsilon \leq \inf_{\overline{\mathbb{R}}}(O) = o$.

Dies zeigt, dass o ein Häufungspunkt von $(a_n)_{n\in\mathsf{N}}$ ist.

- Sei $(a_{\iota(n)})_{n\in\mathsf{N}} \to a \in \mathbb{R}$ eine konvergente Teilfolge. Angenommen, es gilt $o < a$. Dann gilt $\varepsilon := \frac{a-o}{2} > 0$; und wir finden ein $N_\varepsilon \in \mathbb{N}$ mit $a_{\iota(n)} \in (a-\varepsilon, a+\varepsilon)$ für alle $n \geq N_\varepsilon$. Wir erhalten

$$\begin{aligned} \tau := o + \varepsilon = a - \varepsilon < a_{\iota(n)} \quad \forall\, n \geq N_\varepsilon \qquad &\overset{(161)}{\Longrightarrow} \qquad o + \varepsilon \leq O \\ &\Longrightarrow \qquad o < o + \varepsilon \leq \inf_{\overline{\mathbb{R}}}(O) = o, \end{aligned}$$

also den Widerspruch $o < o$.

Dies zeigt, dass o der größte Häufungspunkt von $(a_n)_{n\in\mathsf{N}}$ ist.

b) Folgt analog zu Teil a). □

Korollar 20. *Sei $(a_n)_{n\in\mathsf{N}}$ eine reelle Folge. Es gelten die folgenden Implikationen:*

a) $\lim_n a_n \in \overline{\mathbb{R}} \qquad \Longrightarrow \qquad \overline{\lim}_n a_n = \lim_n a_n = \underline{\lim}_n a_n$

b) $\overline{\lim}_n a_n = \underline{\lim}_n a_n \qquad \Longrightarrow \qquad \overline{\lim}_n a_n = \lim_n a_n = \underline{\lim}_n a_n$

Beweis. Seien U, O wie in (159) und (160).

a) - Sei $\lim_n a_n = a \in \mathbb{R}$. Wegen Satz 16.2), ist $a \in \mathbb{R}$ der einzige Häufungspunkt von $(a_n)_{n\in\mathsf{N}}$. Weiterhin existiert ein $N \in \mathsf{N}$ mit

$$a - 1 < a_n < a + 1 \quad \forall\, n \geq N_\varepsilon \qquad \Longrightarrow \qquad O \geq a - 1 \in U \quad \wedge \quad U \leq a + 1 \in O.$$

Hiermit folgt $a-1 \leq \inf_{\overline{\mathbb{R}}}(O) \leq a+1$ sowie $a-1 \leq \sup_{\overline{\mathbb{R}}}(U) \leq a+1$, also $\overline{\lim}_n a_n, \underline{\lim}_n a_n \in \mathbb{R}$. Daher folgt die Behauptung sofort aus Lemma 34.

- Sei $\lim_n a_n = \infty$. Für $C \geq 0$, existiert dann ein $N \in \mathsf{N}$ mit

$$\begin{aligned} C \leq a_n \quad \forall\, n \geq N \qquad &\overset{(161)}{\Longrightarrow} \qquad C \leq O \quad \wedge \quad C \in U \\ &\Longrightarrow \qquad C \leq \overline{\lim}_n a_n \quad \wedge \quad C \leq \underline{\lim}_n a_n. \end{aligned}$$

Da $C \geq 0$ beliebig war, folgt $\overline{\lim}_n a_n = \infty = \underline{\lim}_n a_n$.

Analog sieht man $\overline{\lim}_n a_n = -\infty = \underline{\lim}_n a_n$ im Falle $\lim_n a_n = -\infty$.

b) Sei $a := \overline{\lim}_n a_n = \underline{\lim}_n a_n \in \overline{\mathbb{R}}$:

- Sei $a \in \mathbb{R}$. Für $\varepsilon > 0$ vorgegeben, existieren gemäß Lemma 20 gewisse $o \in O$ und $u \in U$ mit $o < a + \varepsilon$ und $u > a - \varepsilon$. Weiterhin existieren $N_1, N_2 \in \mathsf{N}$ mit

$$a_n \leq o < a + \varepsilon \quad \forall\, n \geq N_1 \qquad \text{und} \qquad a - \varepsilon < u \leq a_n \quad \forall\, n \geq N_2.$$

Für $n \geq N_\varepsilon := \max(N_1, N_2)$, gilt dann $a_n \in (a - \varepsilon, a + \varepsilon)$, was $\lim_n a_n = a$ zeigt.

- Sei $a = \infty$, also $\sup_{\overline{\mathbb{R}}}(U) = \underline{\lim}_n a_n = \infty$. Dann ist ∞ die einzige obere Schranke von U. Für $\mathbb{R} \ni C \geq 0$ vorgegeben, existiert daher ein $u \in U$ mit $C < u$. Per definitionem, existiert nun ein $N \in \mathsf{N}$ mit $C < u \leq a_n$ für alle $n \geq N$. Dies zeigt $\lim_n a_n = \infty$.
- Sei $a = -\infty$, also $\inf_{\overline{\mathbb{R}}}(O) = \overline{\lim}_n a_n = -\infty$. Dann ist $-\infty$ die einzige untere Schranke von O. Für $\mathbb{R} \ni C \leq 0$ vorgegeben, existiert daher ein $o \in O$ mit $o < C$. Per definitionem, existiert nun ein $N \in \mathsf{N}$ mit $a_n \leq o < C$ für alle $n \geq N$. Dies zeigt $\lim_n a_n = -\infty$. □

Übung 70. *Sei (X, d) ein metrischer Raum, $z \in X$, sowie $(x_n)_{n\in\mathsf{N}}$ und $(y_n)_{n\in\tilde{\mathsf{N}}}$ Folgen in X. Es existiere ein $N \in \mathsf{N} \cap \tilde{\mathsf{N}}$ mit $x_n = y_n$ für alle $n \geq N$. Zeigen Sie:*

$$(x_n)_{n\in\mathsf{N}} \to z \qquad \Longleftrightarrow \qquad (y_n)_{n\in\mathsf{N}} \to z. \tag{162}$$

Lösung: Es gelte $(x_n)_{n\in\mathsf{N}} \to z$. Für $\varepsilon > 0$ vorgegeben, existiert dann ein $N_\varepsilon \in \mathsf{N}$ mit $\mathrm{d}(z, x_n) < \varepsilon$ für alle $n \geq N_\varepsilon$. Für $\tilde{N}_\varepsilon := \max(N_\varepsilon, N)$ und $n \geq \tilde{N}_\varepsilon$, gilt dann

$$\mathrm{d}(z, y_n) \overset{n\geq N}{=} \mathrm{d}(z, x_n) \overset{n\geq N_\varepsilon}{<} \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt hiermit $(y_n)_{n\in\mathsf{N}} \to z$. Die umgekehrte Implikation in (162), zeigt man analog. □

Übung 71. *Seien $(a_n)_{n\in\mathsf{N}}$ und $(b_n)_{n\in\tilde{\mathsf{N}}}$ Folgen in $\mathbb{R}$. Es existierte ein $N \in \mathsf{N} \cap \tilde{\mathsf{N}}$ mit $x_n = y_n$ für alle $n \geq N$. Zeigen Sie:*

$$\overline{\lim}_n a_n = \overline{\lim}_n b_n \qquad \textit{sowie} \qquad \underline{\lim}_n a_n = \underline{\lim}_n b_n$$

Hinweis: Was gilt für die entsprechenden Mengen in (159) *sowie* (160)*.*

Übung 72. *Sei $(a_n)_{n\in\mathsf{N}}$ eine Folge in $\mathbb{R}$. Zeigen Sie $\underline{\lim}_n a_n \leq \overline{\lim}_n a_n$.*

Hinweis: Seien U, O wie in (159) *und* (160)*. Zeigen Sie, dass $x \leq y$ für alle $x \in U$ und $y \in O$ gilt. Folgern Sie hieraus die Behauptung.*

Übung 73. *Sei $(a_n)_{n\in\mathsf{N}}$ eine Folge in $\mathbb{R}$. Zeigen Sie:*

a) Ist $(a_n)_{n\in\mathsf{N}}$ nach oben beschränkt, so gilt $\underline{\lim}_n a_n, \overline{\lim}_n a_n < \infty$.

b) Ist $(a_n)_{n\in\mathsf{N}}$ nach unten beschränkt, so gilt $\underline{\lim}_n a_n, \overline{\lim}_n a_n > -\infty$.

c) Ist $(a_n)_{n\in\mathsf{N}}$ beschränkt, so gilt $\underline{\lim}_n a_n, \overline{\lim}_n a_n \in \mathbb{R}$.

Übung 74. *Seien $(a_n)_{n\in\mathsf{N}}$ und $(b_n)_{n\in\tilde{\mathsf{N}}}$ Folgen in $\mathbb{R}$, sodass ein $N \in \mathsf{N} \cap \tilde{\mathsf{N}}$ existiert, mit $a_n \leq b_n$ für alle $n \geq N$. Zeigen Sie*

$$\overline{\lim}_n a_n \leq \overline{\lim}_n b_n \qquad \textit{und} \qquad \underline{\lim}_n a_n \leq \underline{\lim}\, b_n.$$

6.1.6 Cauchy-Folgen

Definition 40. *Sei (X, d) ein metrischer Raum. Eine Folge $(x_n)_{n\in\mathsf{N}}$ in X heißt Cauchy-Folge, wenn*

$$\forall\, \varepsilon\colon\ \exists\, N_\varepsilon \in \mathsf{N}\colon\ \forall\, m, n \geq N_\varepsilon\colon \quad \mathrm{d}(x_m, x_n) < \varepsilon \tag{163}$$

gilt (anschaulich also, „wenn die Folgenglieder für große Indizes immer näher zusammenrücken“).

Bemerkung 41. *Ist $(V, \|\cdot\|)$ ein normierter Raum, so ließt sich* (163) *in der Form*

$$\forall\, \varepsilon\colon\ \exists\, N_\varepsilon \in \mathsf{N}\colon\ \forall\, m, n \geq N_\varepsilon\colon \quad \|x_m - x_n\| < \varepsilon. \tag{164}$$

Im Falle $V = \mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, ist dann wieder $\|x - x_n\|$ durch $|x_n - x|$ zu ersetzen.

Lemma 35. *Sei (X, d) ein metrischer Raum.*

1) Jede konvergente Folge in X ist eine Cauchy-Folge.

2) Jede Cauchy-Folge ist beschränkt.

3) Konvergiert eine Teilfolge einer Cauchy-Folge, so konvergiert auch die Cauchy-Folge; und zwar gegen den gleichen Grenzwert.

(Eine Cauchy-Folge hat höchstens einen Häufungspunkt.)

Beweis. Sei $(x_n)_{n\in\mathsf{N}}$ eine Folge in X.

1) Es gelte $(x_n)_{n\in\mathsf{N}} \to x \in X$. Für $\varepsilon > 0$ vorgegeben, sei $N_\varepsilon \in \mathsf{N}$ mit $\mathrm{d}(x, x_n) < \frac{\varepsilon}{2}$ für alle $n \geq N_\varepsilon$. Mit M2) und M3) (Dreiecksungleichung) folgt

$$\mathrm{d}(x_m, x_n) \leq \mathrm{d}(x, x_m) + \mathrm{d}(x, x_n) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon \qquad \forall\, m, n \geq N_\varepsilon,$$

was die Behauptung zeigt.

2) Sei $(x_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge; sowie $N \in \mathsf{N}$ mit $\mathrm{d}(x_m, x_n) < 1$ für alle $m, n \geq N$. Wir setzen

$$C := 1 + \max(\{\mathrm{d}(x_n, x_N) \mid \mathsf{N} \ni n \leq N\}).$$

Dann gilt $\mathrm{d}(x_\ell, x_N) \leq C$ für alle $\ell \in \mathsf{N}$; und wir erhalten

$$\mathrm{d}(x_m, x_n) \leq \mathrm{d}(x_m, x_N) + \mathrm{d}(x_n, x_N) \leq C + C = 2 \cdot C \qquad \forall\, m, n \in \mathsf{N},$$

was die Beschränktheit von $(x_n)_{n\in\mathsf{N}}$ zeigt.

3) Sei $(x_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge, und $(x_{\iota(n)})_{n\in\mathbb{N}} \to x \in X$ eine konvergente Teilfolge. Sei $\varepsilon > 0$ vorgegeben:

- Wir wählen $N_\varepsilon \in \mathsf{N}$ mit $\mathrm{d}(x_m, x_n) < \frac{\varepsilon}{2}$ für alle $m, n \geq N_\varepsilon$. (Cauchy-Bedingung (163))
- Wir wählen $\ell \in \mathbb{N}$ mit $\iota(\ell) \geq N_\varepsilon$ (Injektivität von ι); sowie $L \geq \ell$ mit $\mathrm{d}(x, x_{\iota(L)}) < \frac{\varepsilon}{2}$.

Für alle $n \geq N_\varepsilon$, gilt dann wegen $\iota(L) \geq \iota(\ell) \geq N_\varepsilon$:

$$\mathrm{d}(x, x_n) \leq \mathrm{d}(x, x_{\iota(L)}) + \mathrm{d}(x_{\iota(L)}, x_n) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon,$$

was die Behauptung zeigt. □

Definition 41.

- *Ein metrischer Raum (X, d) heißt* vollständig, *wenn jede Cauchy-Folge in X (in X) konvergiert.*
- *Ein normierter Raum $(V, \|\cdot\|)$ heißt* Banachraum, *wenn $(V, \mathrm{d}_{\|\cdot\|})$ vollständig ist.*

Satz 17. *Der metrische Raum $(\mathbb{R}, \mathrm{d}_{|\cdot|})$ ist vollständig (im Sinne von Definition 41), d.h., eine Folge reeller Zahlen konvergiert genau dann in $\mathbb{R}$, wenn sie eine Cauchy-Folge ist.*

Beweis. Wegen Lemma 35.1) bleibt nur noch nachzuweisen, dass jede Cauchy-Folge in $\mathbb{R}$ auch in $\mathbb{R}$ konvergiert.

Sei hierfür $(a_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge in $\mathbb{R}$. Dann ist $(a_n)_{n\in\mathsf{N}}$ beschränkt wegen Lemma 35.2); hat also eine in $\mathbb{R}$ konvergente Teilfolge wegen Satz 15 (Bolzano-Weierstrass). Wegen Lemma 35.3) konvergiert somit auch die Folge $(a_n)_{n\in\mathsf{N}}$, und zwar gegen den gleichen Grenzwert. □

Korollar 21 (Cauchy-Kriterium für komplexe Folgen)**.** *Der metrische Raum $(\mathbb{C}, \mathrm{d}_{|\cdot|})$ist vollständig, d.h., eine Folge komplexer Zahlen konvergiert genau dann, wenn sie eine Cauchy-Folge ist.*

Beweis. Wegen Lemma 35.1) bleibt nur noch nachzuweisen, dass jede Cauchy-Folge in $\mathbb{C}$ auch in $\mathbb{C}$ konvergiert.

Sei $(z_n)_{n\in\mathsf{N}} = (x_n + \mathrm{i}y_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge komplexer Zahlen. Wegen (vgl. (126))

$$|x_m - x_n|,\ |y_m - y_n| \leq \max(|x_m - x_n|, |y_m - y_n|) \leq |z_m - z_n| \qquad \forall\, m, n \in \mathsf{N},$$

sind dann $(x_n)_{n\in\mathsf{N}}$ und $(y_n)_{n\in\mathsf{N}}$ Cauchy-Folgen in $\mathbb{R}$, also nach Satz 17 konvergent in $\mathbb{R}$, d.h., $(x_n)_{n\in\mathsf{N}} \to x \in \mathbb{R}$ und $(y_n)_{n\in\mathsf{N}} \to y \in \mathbb{R}$. Nach Übung 63.b) gilt dann $(z_n)_{n\in\mathsf{N}} \to x + \mathrm{i}y \in \mathbb{C}$, was die Behauptung zeigt. □

Bemerkung 42. *Der metrische Raum* $(\mathbb{Q}, \mathrm{d}_{|\cdot|})$ *ist nicht vollständig; denn für* $a = 2$ *und* $k = 2$, *konvergiert die rekursiv definierte (in diesem Fall)* $\mathbb{Q}$*-Folge* $(a_n)_{n\in\mathbb{N}}$ *aus Beispiel 38 (Babylonisches Wurzelziehen) gegen* $\sqrt{2} \in \mathbb{R}$ *(und sonst gegen keinen anderen Grenzwert wegen Satz 12). Gemäß Lemma 35.1) ist* $(a_n)_{n\in\mathbb{N}}$ *eine Cauchy-Folge im metrischen Raum* $(\mathbb{R}, \mathrm{d}_{|\cdot|})$*; und dann klarerweise auch in* $\mathbb{Q}$. *Wegen* $\sqrt{2} \notin \mathbb{Q}$ *(vgl. Korollar 11), ist somit* $(\mathbb{Q}, \mathrm{d})$ *nicht vollständig.*

Bemerkung* 11. *Satz 17 und Bemerkung 42 lassen bereits erahnen, dass der Begriff des vollständigen metrischen Raumes im Kontext angeordneter Körper mit der Vollständigkeit im Sinne von Definition 26 übereinstimmen. Weiß man nämlich, dass* $(\mathbb{R}, \mathrm{d}_{|\cdot|})$ *ein vollständiger metrischer Raum ist, so kann man hieraus Schlussfolgern, dass eine nichtleere nach oben beschränkte Teilmenge* $\emptyset \neq M \leq C \in \mathbb{R}$ *von* $\mathbb{R}$, *ein Supremum in* $\mathbb{R}$ *hat:*

Für jedes $n \in \mathbb{N}$, *wählt man* $\alpha[n] \in \mathbb{N}$ *maximal mit* $M \leq a_n := C - \alpha[n] \cdot 10^{-n}$. *Man überlegt sich dann, dass* a_n *monoton fallend und (wegen* $M \neq \emptyset$*) nach unten beschränkt ist – genauer folgt per Induktion, dass ein* $L \geq 0$ *existiert, sowie „Ziffern"* $o_j \in \{0, \ldots, 9\}$ *für* $j \in \mathbb{N}$, *sodass gilt:*

$$\alpha[n] = L + \textstyle\sum_{\ell=0}^{n} o_n \cdot 10^{-\ell} \qquad \forall\, \ell \in \mathbb{N}$$

Hiermit folgt, dass $(a_n)_{n\in\mathbb{N}}$ *eine Cauchy-Folge ist, also wegen der Vollständigkeit gegen ein* $a \in \mathbb{R}$ *konvergiert. Wegen Korollar 14 gilt dann* $M \leq a$*; und aus der Definition von* $(a_n)_{n\in\mathbb{N}}$ *schlussfolgert man weiterhin, dass für jedes* $\tilde{C} \in \mathbb{R}$ *mit* $M \leq \tilde{C}$, *bereits* $a \leq \tilde{C}$ *gilt.*

Bemerkung* 12 (alternative Konstruktion von $\mathbb{R}$)**.** *Es bezeichne* $\boldsymbol{R}$ *die Menge aller Cauchy-Folgen in* $\mathbb{Q}$ *mit Indexmenge* $\mathbb{N}$. *Man erhält eine alternative Konstruktion von* $\mathbb{R}$, *indem man setzt* $\mathbb{R} := \boldsymbol{R}/\sim$ *mit*

$$(q_n)_{n\in\mathbb{N}} \sim (\tilde{q}_n)_{n\in\mathbb{N}} \quad \textit{für} \quad (q_n)_{n\in\mathsf{N}}, (\tilde{q}_n)_{n\in\mathbb{N}} \in \boldsymbol{R} \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad \lim_n (q_n - \tilde{q}_n) = 0.$$

Die Operationen sind dann punktweise auf den Repräsentanten definiert – Beispielsweise:

$$\begin{aligned}
[(q_n)_{n\in\mathbb{N}}]_\sim + [(\tilde{q}_n)_{n\in\mathbb{N}}]_\sim &:= [(q_n + \tilde{q}_n)_{n\in\mathbb{N}}]_\sim \\
[(q_n)_{n\in\mathbb{N}}]_\sim \cdot [(\tilde{q}_n)_{n\in\mathbb{N}}]_\sim &:= [(q_n \cdot \tilde{q}_n)_{n\in\mathbb{N}}]_\sim.
\end{aligned}$$

Die Anordnung erhält man, indem man $[(q_n)_{n\in\mathbb{N}}]_\sim > 0$ *definiert, falls* $0 < q \in \mathbb{Q}$ *und* $N \in \mathbb{N}$ *mit* $q_n \geq q$ *für alle* $n \geq N$ *gilt. Es lässt sich dann nachweisen, dass man hiermit einen vollständig angeordneten Körper erhält, bzw., vermöge der zugehörigen Betragsmetrik, einen vollständigen metrischen Raum. Im letzteren Falle muss man dann beispielsweise zeigen, dass dass jede Cauchy-Folge* $(\gamma_n)_{n\in\mathbb{N}}$ *von Äquivalenzklassen* $\gamma_n \in \boldsymbol{R}/\sim$ *von Cauchy-Folgen in* $\mathbb{Q}$, *gegen eine Äquivalenzklassen* $\gamma \in \boldsymbol{R}/\sim$ *von Cauchy-Folgen in* $\mathbb{Q}$ *konvergiert (dies ist womöglich nicht weniger technisch und abstrakt, als unsere Konstruktion von* $\mathbb{R}$ *in Form von Dedekindscher Schnitten).*

6.1.7 Folgenabgeschlossenheit

Definition 42. *Sei (X, d) ein metrischer Raum. Eine Teilmenge $Y \subseteq X$ heißt* folgenabgeschlossen*, wenn die folgende Implikation gilt:*

$$(y_n)_{n\in\mathsf{N}} \ \textit{Folge in } Y \ \textit{mit} \ (y_n)_{n\in\mathsf{N}} \to x \in X \qquad \Longrightarrow \qquad x \in Y. \tag{165}$$

Lemma 36. *Sei (X, d) ein metrischer Raum, und $Y \subseteq X$ eine Teilmenge. Dann gilt*

$$Y \ \textit{folgenabgeschlossen} \qquad \Longleftrightarrow \qquad Y \in \mathcal{A}(X).$$

Beweis. • Sei Y abgeschlossen, aber nicht folgenabgeschlossen:

- Dann gilt **nicht** die Implikation (165); d.h., es existiert eine Folge $(y_n)_{n\in\mathsf{N}}$ in Y, mit $(y_n)_{n\in\mathsf{N}} \to x \in X \setminus Y$.
- Da $X \setminus Y$ per Definition offen ist, existiert ein $\varepsilon > 0$ mit $\mathsf{B}_\varepsilon(x) \subseteq X \setminus Y$.
- Wegen $(y_n)_{n\in\mathsf{N}} \to x$, enthält $\mathsf{B}_\varepsilon(x)$ mindestens ein Folgenglied y_m mit $m \in \mathsf{N}$.
- Wir erhalten den Widerspruch $Y \ni y_m \in \mathsf{B}_\varepsilon(x) \subseteq X \setminus Y$.

• Sei Y folgenabgeschlossen, aber nicht abgeschlossen:

- Dann ist $X \setminus Y$ nicht offen; d.h., es existiert ein $x \in X \setminus Y$ mit $\mathsf{B}_\varepsilon(x) \cap Y \neq \emptyset$ ($\mathsf{B}_\varepsilon(x) \not\subseteq X \setminus Y$) für alle $\varepsilon > 0$.
- Zu jedem $n \in \mathbb{N}_{>0}$ existiert daher ein $y_n \in \mathsf{B}_{\frac{1}{n}}(x) \cap Y \subseteq Y$; und es gilt dann $(y_n)_{n\in\mathbb{N}_{>0}} \to x$: Für $\varepsilon > 0$ vorgegeben, sowie $N_\varepsilon \in \mathbb{N}_{>0}$ mit $0 < \frac{1}{N_\varepsilon} < \varepsilon$, folgt $\qquad (\frac{1}{n} < \frac{1}{N_\varepsilon} < \varepsilon)$

$$y_n \in \mathsf{B}_{\frac{1}{n}}(x) \subseteq \mathsf{B}_{\frac{1}{N_\varepsilon}}(x) \subseteq \mathsf{B}_\varepsilon(x) \qquad \forall\, n \geq N_\varepsilon.$$

- Da Y folgenabgeschlossen ist, gilt $x \in Y$; was der Wahl $x \in X \setminus Y$ widerspricht. □

6.2 Reihen

In diesem Abschnitt behandeln wir Reihen in normierten Vektorräumen; hauptsächlich in den Körpern $\mathbb{R}$ und $\mathbb{C}$.

6.2.1 Grundlegendes zu Reihen

Sei $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, sowie $(V, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum.

Reihen sind besondere Folgen:

Definition 43. *Sei $(v_k)_{k\in\mathsf{N}}$ eine Folge in V:*

- *Die n-te Partialsumme der Folge $(v_k)_{k\in\mathsf{N}}$, ist definiert durch $S_n := \sum_{k=\mathsf{p}}^n v_k$ für alle $n \in \mathsf{N}$.*
- *Die Folge $(S_n)_{n\in\mathsf{N}}$ der (n-ten) Partialsummen heißt unendliche Reihe (in V) mit den Gliedern v_k. Man notiert sie auch mit*

$$\textstyle\sum_{k=\mathsf{p}}^\infty v_k := (S_n)_{n\in\mathsf{N}} = (\sum_{k=\mathsf{p}}^n v_k)_{n\in\mathsf{N}}.$$

- *Man sagt, dass die Reihe $\sum_{k=\mathsf{p}}^\infty v_k$ konvergiert, wenn die Folge der Partialsummen konvergiert; wenn also $(S_n)_{n\in\mathsf{N}} \to S \in V$ gilt. In diesem Falle schreibt man trotz der Doppeldeutigkeit, für den Grenzwert auch $S = \sum_{k=\mathsf{p}}^\infty v_k$.*
- *Man sagt, dass die Reihe $\sum_{k=\mathsf{p}}^\infty v_k$ divergiert, wenn sie nicht konvergiert, d,h, wenn die Folge der Partialsummen $(S_n)_{n\in\mathsf{N}}$ divergiert.*

Beispiel 43. *Es gilt* $\sum_{k=1}^{\infty} \frac{1}{k\cdot(k+1)} = 1$, *denn für* $n \geq 1$ *erhalten wir*

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

also $\sum_{k=1}^{\infty} \frac{1}{k\cdot(k+1)} = \lim_n S_n = 1$.

Übung 75. *Sei* $\sum_{k=\mathsf{p}}^{\infty} v_k$ *konvergent mit Grenzwert* S. *Sei* $\sum_{k=\mathsf{q}}^{\infty} w_k$ *eine Reihe, sowie* $N \geq \mathsf{p}, \mathsf{q}$ *mit* $v_k = w_k$ *für alle* $k \geq N$. *Dann konvergiert* $\sum_{k=\mathsf{q}}^{\infty} w_k$, *und hat den Grenzwert*

$$R := S - \sum_{k=\mathsf{p}}^{N} v_k + \sum_{k=\mathsf{q}}^{N} w_k.$$

Wir haben die folgende notwendige (aber nicht hinreichende) Bedingung für die Konvergenz von Reihen:

Lemma 37. *Konvergiert die Reihe* $\sum_{k=\mathsf{p}}^{\infty} v_k$, *so gilt* $(v_k)_{k\in\mathsf{N}} \to 0$.

Beweis. Sei $S_n := \sum_{k=\mathsf{p}}^{n} v_k$ für alle $n \in \mathsf{N}$. Nach Lemma 35.1) ist $(S_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge. Für jedes $\varepsilon > 0$ existiert daher ein $N_\varepsilon \in \mathsf{N}$ mit

$$\|S_m - S_n\| < \varepsilon \quad \forall\, m, n \geq N_\varepsilon \qquad \Longrightarrow \qquad \|v_n - 0\| = \|v_n\| = \|S_n - S_{n-1}\| < \varepsilon \quad \forall\, n \geq N_\varepsilon + 1,$$

was die Behauptung zeigt. ☐

Proposition 5 liefert das folgende Beispiel:

Korollar 22 (Die geometrische Reihe $\sum_{n=0}^{\infty} z^n$)**.** *Sei* $z \in \mathbb{C}$ *vorgegeben.*

1) *Für* $z \neq 1$, *gilt die geometrische Summenformel*

$$\sum_{k=0}^{n} z^k = \frac{1 - z^{n+1}}{1 - z} \qquad \forall\, n \in \mathbb{N}.$$

2) *Für* $|z| < 1$, *konvergiert die geometrische Reihe mit*

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

3) *Für* $|z| \geq 1$, *divergiert die geometrische Reihe.*

Beweis. 1) Für alle $n \in \mathbb{N}$, gilt

$$(1 - z) \cdot \sum_{k=0}^{n} z^k = \sum_{k=0}^{n} z^k - \overbrace{\sum_{k=0}^{n} z^{k+1}}^{\sum_{k=1}^{n+1} z^k} = 1 - z^{n+1} \qquad \Longrightarrow \qquad \sum_{k=0}^{n} z^k = \frac{1 - z^{n+1}}{1 - z}.$$

2) Wegen Proposition 5 gilt $(z^n)_{n\in\mathbb{N}} \to 0$ für $|z| < 1$; also wegen Teil 1)

$$\sum_{k=0}^{\infty} z^k = \lim_n \sum_{k=0}^{n} z^k = \lim_n \frac{1-z^{n+1}}{1-z} = \frac{1}{1-z}.$$

3) Es ist $(z^n)_{n\in\mathbb{N}}$ konstant 1 für $z = 1$; sowie gemäß Proposition 5 divergent für $z \neq 1$ mit $|z| \geq 1$. Daher folgt die Behauptung aus Lemma 37. ☐

Proposition 6 liefert die folgende Rechenregel:

Korollar 23. *Seien* $\sum_{n=\mathsf{p}}^{\infty} v_n$ *und* $\sum_{n=\mathsf{p}}^{\infty} w_n$ *konvergente Reihen in* V, *sowie* $\lambda, \mu \in \mathbb{K}$. *Dann gilt*

$$\sum_{n=\mathsf{p}}^{\infty} (\lambda \cdot v_n + \mu \cdot w_n) = \lambda \cdot \sum_{n=\mathsf{p}}^{\infty} v_n + \mu \cdot \sum_{n=\mathsf{p}}^{\infty} w_n.$$

(Selbiges gilt natürlich insbesondere im Falle $V = \mathbb{K}$.)

Beweis. Für $n \in \mathsf{N}$, sei $S_n := \sum_{k=\mathsf{p}}^n v_k$ sowie $R_n := \sum_{k=\mathsf{p}}^n w_k$. Es gilt

$$\textstyle\sum_{k=\mathsf{p}}^n (\lambda \cdot v_k + \mu \cdot w_k) = \lambda \cdot \sum_{k=\mathsf{p}}^n v_k + \mu \cdot \sum_{k=\mathsf{p}}^n w_k = \lambda \cdot S_n + \mu \cdot R_n \qquad \forall\, n \in \mathsf{N}. \tag{166}$$

Proposition 6.4) liefert daher (dritter Schritt)

$$\begin{aligned} \textstyle\sum_{n=\mathsf{p}}^\infty (\lambda \cdot v_n + \mu \cdot w_n) &= \textstyle\lim_n \sum_{k=\mathsf{p}}^n (\lambda \cdot v_k + \mu \cdot w_k) \\ &\overset{(166)}{=} \lim_n \big(\lambda \cdot S_n + \mu \cdot R_n\big) \\ &= \lambda \cdot \lim_n S_n + \mu \cdot \lim_n R_n \\ &= \textstyle\lambda \cdot \sum_{k=\mathsf{p}}^\infty v_k + \mu \cdot \sum_{k=\mathsf{p}}^\infty w_k, \end{aligned}$$

was die Behauptung zeigt. □

6.2.2 Monotone Konvergenz und das Leibniz-Kriterium

Wir betrachten nun den Fall $V = \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Satz 14 (monotone Konvergenz) motiviert die folgenden Definitionen:

Terminologie 17.

1) *Eine $\mathbb{K}$-Reihe (Reihe in $\mathbb{K}$) $\sum_{n=\mathsf{p}}^\infty c_n$ heißt* beschränkt*, wenn die Folge der Partialsummen $(S_n)_{n\in\mathsf{N}} = (\sum_{k=\mathsf{p}}^n c_k)_{n\in\mathsf{N}}$ beschränkt ist; wenn also ein $C \geq 0$ existiert, mit*

$$|S_n| = |\textstyle\sum_{k=\mathsf{p}}^n c_k| \leq C \qquad \forall\, n \in \mathsf{N}.$$

Beachte: *Ist eine $\mathbb{K}$-Reihe $\sum_{n=\mathsf{p}}^\infty c_n$ konvergent, so auch beschränkt; und zwar wegen Satz 13, angewandt auf die Folge der Partialsummen $(S_n)_{n\in\mathsf{N}}$.*

Die $\mathbb{K}$-Reihe $\sum_{n=\mathsf{p}}^\infty c_n$ heißt unbeschränkt*, wenn sie nicht beschränkt ist; wenn also die Folge der Partialsummen $(S_n)_{n\in\mathsf{N}}$ unbeschränkt (nicht beschränkt) ist.*

2) *Eine $\mathbb{R}$-Reihe $\sum_{n=\mathsf{p}}^\infty a_n$ heißt* positiv*, wenn $a_n \geq 0$ für alle $n \in \mathsf{N}$ gilt.*

Beachte: *Eine $\mathbb{R}$-Reihe $\sum_{n=\mathsf{p}}^\infty a_n$ ist genau dann positiv, wenn die Folge $S_n = \sum_{k=\mathsf{p}}^n a_k$ (für $n \in \mathsf{N}$) monoton wachsend ist mit $a_\mathsf{p} \geq 0$:*

- *Ist $\sum_{n=\mathsf{p}}^\infty a_n$ positiv, dann gilt $a_\mathsf{p} \geq 0$, sowie*

$$S_{n+1} = \textstyle\sum_{k=\mathsf{p}}^{n+1} a_k = S_n + a_{n+1} \geq S_n \qquad \forall\, n \in \mathsf{N}.$$

- *Ist $(S_n)_{n\in\mathsf{N}}$ monoton wachsend mit $a_\mathsf{p} \geq 0$, so gilt $a_{n+1} = S_{n+1} - S_n \geq 0$ für alle $n \in \mathsf{N}$.*

3) *Eine $\mathbb{R}$-Reihe $\sum_{n=\mathsf{p}}^\infty c_n$ in $\mathbb{R}$ heißt* alternierend*, wenn eine Folge $(a_n)_{n\in\mathsf{N}}$ in $[0, \infty)$ existiert mit*

$$\big(\, c_n = (-1)^n \cdot a_n \quad \forall\, n \in \mathsf{N} \,\big) \qquad \vee \qquad \big(\, c_n = (-1)^{n+1} \cdot a_n \quad \forall\, n \in \mathsf{N} \,\big).$$

Satz 18 (Monotone Konvergenz für Reihen)**.** *Eine positive Reihe ($\mathbb{R}$-Reihe) $\sum_{k=\mathsf{p}}^\infty a_k$ konvergiert genau dann (in $\mathbb{R}$), wenn sie beschränkt ist; und dann gilt $\sum_{k=\mathsf{p}}^\infty a_k = \sup_\mathbb{R}(\{\sum_{k=\mathsf{p}}^n a_k \mid n \in \mathsf{N}\})$.*

Beweis. • Konvergiert die Reihe, so ist sie beschränkt wegen Terminologie 17.1).

- Ist die Reihe $\sum_{k=\mathsf{p}}^{\infty} a_k$ beschränkt, so ist per definitionem die Folge der Partialsummen $(S_n)_{n\in\mathsf{N}} = (\sum_{k=\mathsf{p}}^{n} a_k)_{n\in\mathsf{N}}$ beschränkt.

  Gemäß Terminologie 17.2) ist nun $(S_n)_{n\in\mathsf{N}}$ monoton wachsend. Satz 14.1) zeigt daher

  $$\textstyle\sum_{k=\mathsf{p}}^{\infty} a_k = \lim_n S_n = \sup_{\mathbb{R}}(\{S_n \mid n \in \mathsf{N}\}) = \sup_{\mathbb{R}}(\{\sum_{k=\mathsf{p}}^{n} a_k \mid n \in \mathsf{N}\}). \qquad \square$$

Notation 23. *Sei* $\sum_{k=\mathsf{p}}^{\infty} a_k$ *eine unbeschränkte positive Reihe; sowie* $S_n := \sum_{k=\mathsf{p}}^{n} a_k$ *für alle* $n \in \mathsf{N}$. *Dann gilt* $\lim_n S_n = \infty$*; und wir setzen* $\sum_{k=\mathsf{p}}^{\infty} a_k := \infty$.

Gemäß Satz 18, gilt für eine positive Reihe $\sum_{k=\mathsf{p}}^{\infty} a_k$ *also immer* $\sum_{k=\mathsf{p}}^{\infty} a_k \in [0, \infty]$.

Beweis von $\lim_n S_n = \infty$*:* $(S_n)_{n\in\mathsf{N}}$ ist nicht nach oben beschränkt, da $(S_n)_{n\in\mathsf{N}}$ durch 0 nach unten beschränkt ist. Zu jedem $C \geq 0$ existiert daher ein $N \in \mathsf{N}$ mit $C \leq S_N$. Da $(S_n)_{n\in\mathsf{N}}$ gemäß Terminologie 17.2) monoton wachsend ist, erhalten wir $C \leq S_N \leq S_n$ für alle $n \geq N$. Dies zeigt $\lim_n S_n = \infty$. $\square$

Beispiel 44. *Es gilt* $\sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$. *In der Tat, wegen Beispiel 43 und Satz 18 gilt*

$$\textstyle\sum_{k=1}^{n} \frac{1}{k\cdot(k+1)} \leq \sup_{\mathbb{R}}(\{\sum_{k=1}^{n} \frac{1}{k\cdot(k+1)} \mid n \in \mathsf{N}\}) = \sum_{k=1}^{\infty} \frac{1}{k\cdot(k+1)} = 1.$$

Für $n \in \mathbb{N}$*, erhalten wir*

$$\textstyle S_n := \sum_{k=1}^{n} \frac{1}{k^2} = 1 + \sum_{k=2}^{n} \frac{1}{k^2} \leq 1 + \sum_{k=2}^{n} \frac{1}{(k-1)\cdot k} = 1 + \sum_{k=1}^{n-1} \frac{1}{k\cdot(k+1)} \leq 1 + 1 = 2.$$

Daher ist die Reihe $\sum_{k=1}^{\infty} \frac{1}{k^2}$ *beschränkt, und konvergiert somit nach Satz 18.*

Satz 19 (Leibniz-Kriterium). *Sei* $(a_n)_{n\in\mathbb{N}}$ *eine monoton fallende Folge in* $[0, \infty)$. *Es konvergiert die (alternierende) Reihe* $\sum_{n=0}^{\infty}(-1)^n \cdot a_n$ *genau dann, wenn* $(a_n)_{n\in\mathbb{N}} \to 0$ *gilt.*

Beweis. • Konvergiert die Reihe $\sum_{n=0}^{\infty}(-1)^n \cdot a_n$, so gilt $((-1)^n \cdot a_n)_{n\in\mathbb{N}} \to 0$ wegen Lemma 37. Offensichtlich impliziert dies $(a_n)_{n\in\mathbb{N}} \to 0$. $(|a_n - 0| = |(-1)^n \cdot a_n - 0|)$

- Es gelte $(a_n)_{n\in\mathbb{N}} \to 0$. Wir setzen $S_n = \sum_{k=0}^{n}(-1)^n \cdot a_n$ für alle $n \in \mathbb{N}$. Da $(a_n)_{n\in\mathbb{N}}$ monoton fallend ist, gilt für alle $n \in \mathbb{N}$:

  (1) $S_{2n+2} = S_{2n} - a_{2n+1} + a_{2n+2} \leq S_{2n}$,

  (2) $S_{2n+3} = S_{2n+1} + a_{2n+2} - a_{2n+3} \geq S_{2n+1}$,

  (3) $S_{2n} \geq S_{2n} - a_{2n+1} = S_{2n+1} \geq S_1$, (letzter Schritt folgt induktiv aus (2))

  (4) $S_{2n} - S_{2n+1} = a_{2n+1}$.

  Wegen (1) ist $(S_{2n})_{n\in\mathbb{N}}$ ist monoton fallend; und wegen (3) ist $(S_{2n})_{n\in\mathbb{N}}$ durch S_1 nach unten beschränkt. Wegen Satz 14.2) (monotone Konvergenz) gilt daher

  $$S := \lim_n S_{2n} = \inf_{\mathbb{R}}(\{S_{2n} \mid n \in \mathbb{N}\}).$$

  Für $\varepsilon > 0$, existiert somit ein $N_1 \in \mathbb{N}$ mit $|S - S_{2n}| < \frac{\varepsilon}{2}$ für alle $n \geq N_1$. Wegen $(a_n)_{n\in\mathbb{N}} \to 0$, existiert weiterhin ein $N_2 \in \mathbb{N}$ mit $|a_n| = |a_n - 0| < \frac{\varepsilon}{2}$ für alle $n \geq N_2$. Für $N := \max(N_1, N_2)$ und $n \geq N$, gilt somit

  $$|S - S_{2n}| < \frac{\varepsilon}{2} < \varepsilon$$
  $$|S - S_{2n+1}| \leq |S - S_{2n}| + |S_{2n} - S_{2n+1}| \overset{(4)}{=} |S - S_{2n}| + |a_{2n+1}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

  Wir erhalten $|S - S_\ell| < \varepsilon$ für alle $\ell \geq 2N + 1 =: N_\varepsilon$.[27] $\square$

Beispiel 45. *Die folgenden beiden Reihen sind konvergent:*

- $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} + \ldots (= \frac{\pi}{4})$
- $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \ldots (= \log(2))$

[27] Für $N_\varepsilon \leq \ell = 2n$ $(n \in \mathbb{N})$ gilt $n \geq (N_\varepsilon - 1)/2 = N$; und auch für $N_\varepsilon \leq \ell = 2n + 1$ $(n \in \mathbb{N})$ gilt $n \geq (N_\varepsilon - 1)/2 = N$.

6.2.3 Cauchy-Kriterium und absolute Konvergenz von Reihen

Im Folgenden sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Satz 20 (Cauchy-Kriterium)**.** *Eine $\mathbb{K}$-Reihe $\sum_{n=\mathsf{p}}^{\infty} c_n$ konvergiert genau dann (in $\mathbb{K}$), wenn gilt:*

$$\forall\, \varepsilon > 0\colon\ \exists\, N_\varepsilon \in \mathsf{N}\colon\ \forall\, m, n \geq N_\varepsilon\colon\ |\textstyle\sum_{k=m}^{n} c_k| < \varepsilon. \tag{167}$$

Beweis. Sei $S_n = \sum_{k=\mathsf{p}}^{n} c_k$ für alle $n \in \mathsf{N}$. Dann gilt

$$|S_n - S_{m-1}| = |\textstyle\sum_{k=m}^{n} c_k| \qquad \forall\, \mathsf{p} + 1 \leq m, n \in \mathsf{N}.$$

Die Behauptung folgt nun aus dem entsprechenden Cauchy-Kriterium für Konvergenz von Folgen in $\mathbb{R}$ bzw. $\mathbb{C}$; also aus Satz 17 bzw. Korollar 21. □

Beispiel 46 (Divergenz der harmonische Reihe)**.** *Die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ divergiert. Für die zugehörige Folge der Partialsummen $(S_n)_{n \in \mathbb{N}_{>0}}$, gilt nämlich*

$$|S_{2\ell} - S_\ell| = \textstyle\sum_{k=\ell+1}^{2\ell} \frac{1}{k} \geq \sum_{k=\ell+1}^{2\ell} \frac{1}{2\ell} = \frac{\ell}{2\ell} = \frac{1}{2} \qquad \forall\, \ell \in \mathbb{N}_{>0},$$

womit (167) *nicht erfüllt sein kann.*

Definition 44. *Eine $\mathbb{K}$-Reihe $\sum_{n=\mathsf{p}}^{\infty} c_n$ heißt* absolut konvergent*, wenn die positive Reihe $\sum_{n=\mathsf{p}}^{\infty} |c_n|$ der Beträge (in $\mathbb{R}$) konvergiert (d.h., $\sum_{n=\mathsf{p}}^{\infty} |c_n| < \infty$ gemäß Notation 23).*

Satz 21. *Jede absolut konvergente Reihe konvergiert.*

Beweis. Sei $\sum_{n=\mathsf{p}}^{\infty} c_n$ absolut konvergent; also $\sum_{n=\mathsf{p}}^{\infty} |c_n| < \infty$. Satz 20, angewandt auf die Reihe $\sum_{n=\mathsf{p}}^{\infty} |c_n|$ liefert

$$\forall\, \varepsilon > 0\colon\ \exists\, N_\varepsilon \in \mathsf{N}\colon\ \forall\, m, n \geq N_\varepsilon\colon\ \textstyle\sum_{k=m}^{n} |c_k| = |\sum_{k=m}^{n} |c_k|| < \varepsilon. \tag{168}$$

Wegen der Dreiecksungleichung gilt nun

$$|\textstyle\sum_{k=m}^{n} c_k| \leq \sum_{k=m}^{n} |c_k| \qquad \forall\, m, n \in \mathsf{N};$$

sodass (168), die Bedingung (167) impliziert. Satz (20) zeigt nun die Behauptung. □

Reihen, die nicht absolut konvergieren, zeigen ein etwas merkwürdiges Konvergenzverhalten. Sei beispielsweise

$$S := \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} \pm \dots$$

die Summe der alternierenden harmonischen Reihe, die nach dem Leibniz-Kriterium konvergiert, aber gemäß Beispiel 46 nicht absolut konvergiert. Wir erhalten dann (Konvergenzsätze)

$$\begin{aligned} S &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} \pm \dots \\ \frac{1}{2} \cdot S &= 0 + \frac{1}{2} + 0 - \frac{1}{4} + 0 + \frac{1}{6} + 0 - \frac{1}{8} \pm \dots \\ \frac{3}{2} \cdot S &= 1 + 0 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + 0 + \frac{1}{7} - \frac{1}{4} \pm \dots, \end{aligned}$$

wobei die zweite Zeile durch Multiplikation der ersten Zeile mit dem Faktor $\frac{1}{2}$ und dem Einfügen von Nullen entsteht; sowie die letzte Zeile durch aufaddieren der ersten beiden Zeilen. Lässt man die Nullen in der dritten Zeile weg, so entsteht nach geeigneter Umordnung aus der Reihe in der letzten Zeile wieder die alternierende harmonische Reihe. Die führt zu dem scheinbaren Widerspruch $S = \frac{3}{2} \cdot S$, welcher sich durch die Feststellung auflösen lässt, dass sich im Allgemeinen der Grenzwert einer Reihe bei einer Umordnung ändert (vgl. Satz 23).

Wie der folgende Satz zeigt, treten derartige Effekte bei absolut konvergenten Reihen nicht auf:

Satz 22. *Die $\mathbb{K}$-Reihe $\sum_{n=\mathsf{p}}^{\infty} c_n$ sei absolute konvergent, und es sei $\alpha\colon \mathsf{N} \to \mathsf{N}$ eine Bijektion. Dann konvergiert auch die (umgeordnete) Reihe $\sum_{n=\mathsf{p}}^{\infty} c_{\alpha(n)}$ absolut, und hat den selben Grenzwert.*

(Es gilt sowohl $\sum_{n=\mathsf{p}}^{\infty} c_n = \sum_{n=\mathsf{p}}^{\infty} c_{\alpha(n)}$, als auch $\sum_{n=\mathsf{p}}^{\infty} |c_n| = \sum_{n=\mathsf{p}}^{\infty} |c_{\alpha(n)}|$.)

$*$*Beweis.* Für $n \in \mathsf{N}$, sei $S_n := \sum_{k=\mathsf{p}}^{n} c_k$ sowie $R_n := \sum_{k=\mathsf{p}}^{n} c_{\alpha(k)}$. Sei zudem $\varepsilon > 0$ vorgegeben.

- Wegen Satz 20, angewandt auf $\sum_{n=\mathsf{p}}^{\infty} |c_n|$, existiert ein $M \in \mathsf{N}$ mit (vgl.(168))

$$\textstyle\sum_{k=M+1}^{n} |c_k| = \left| \sum_{k=M+1}^{n} |c_k| \right| < \frac{\varepsilon}{2} \qquad \forall\, n \geq M+1. \tag{169}$$

- Wegen der Surjektivität (Bijektivität) von α, existiert ein $N \geq M+1$ mit

$$\{\mathsf{p}, \dots, M\} \subseteq \alpha(\{\mathsf{p}, \dots, n\}) \quad \text{also} \quad \max(\alpha(\{\mathsf{p}, \dots, n\})) \geq M+1 \qquad \forall\, n \geq N$$

  Für jedes $n \geq N$, gilt dann

$$\textstyle\sum_{k=\mathsf{p}}^{n} c_{\alpha(k)} = \sum_{k \in \alpha(\{\mathsf{p},\dots,n\})} c_k = \overbrace{\sum_{k \in \{\mathsf{p},\dots,M\}} c_k}^{\sum_{k=\mathsf{p}}^{M} c_k} + \sum_{k \in \alpha(\{\mathsf{p},\dots,n\}) \setminus \{\mathsf{p},\dots,M\}} c_k \tag{170}$$

$$\text{mit} \qquad \textstyle\sum_{k \in \alpha(\{\mathsf{p},\dots,n\}) \setminus \{\mathsf{p},\dots,M\}} |c_k| \overset{(169)}{\leq} \sum_{k=M+1}^{\max(\alpha(\{\mathsf{p},\dots,n\}))} |c_k| < \frac{\varepsilon}{2}. \tag{171}$$

Für $n \geq N$, erhalten wir

$$\begin{aligned}
|S_n - R_n| &= \textstyle\left| \sum_{k=\mathsf{p}}^{n} c_k - \sum_{k=\mathsf{p}}^{n} c_{\alpha(k)} \right| \\
&\overset{(170)}{=} \textstyle\left| \sum_{k=M+1}^{n} c_k - \sum_{k \in \alpha(\{\mathsf{p},\dots,n\}) \setminus \{\mathsf{p},\dots,M\}} c_k \right| \\
&\leq \textstyle\left| \sum_{k=M+1}^{n} c_k \right| + \left| \sum_{k \in \alpha(\{\mathsf{p},\dots,n\}) \setminus \{\mathsf{p},\dots,M\}} c_k \right| \\
&\leq \textstyle\sum_{k=M+1}^{n} |c_k| + \sum_{k \in \alpha(\{\mathsf{p},\dots,n\}) \setminus \{\mathsf{p},\dots,M\}} |c_k| \\
&\overset{(169),\,(171)}{<} \textstyle\frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Dies zeigt $\lim_n (R_n - S_n) = 0$; und wir erhalten mit Korollar 23 (zweiter Schritt)

$$\textstyle\sum_{n=\mathsf{p}}^{\infty} c_n = \lim_n S_n = \lim_n (S_n + (R_n - S_n)) = \lim_n R_n = \sum_{n=\mathsf{p}}^{\infty} c_{\alpha(n)}.$$

Wendet man nun die gleiche Argumentation auf die Reihe $\sum_{n=\mathsf{p}}^{\infty} |c_n|$ an, so folgt gleichsam $\sum_{n=\mathsf{p}}^{\infty} |c_n| = \sum_{n=\mathsf{p}}^{\infty} |c_{\alpha(n)}|$; also die absolute Konvergenz von $\sum_{n=\mathsf{p}}^{\infty} c_{\alpha(n)}$. □

Der folgende Satz zeigt auf, wie instabil konvergente, aber nicht absolut konvergente Reihen gegenüber Umordnungen sind:

Satz 23 (Riemannscher Umordnungssatz)**.** *Ist die $\mathbb{R}$-Reihe $\sum_{k=0}^{\infty} c_k$ konvergent aber nicht absolut konvergent, so existiert zu jedem beliebigen $x \in \mathbb{R}$, eine Bijektion $\alpha : \mathbb{N} \to \mathbb{N}$ mit $\sum_{k=\mathsf{p}}^{\infty} c_{\alpha(k)} = x$.*

$*$*Beweisidee:* • Da $\sum_{k=0}^{\infty} c_k$ nicht absolut konvergiert, existieren unendlich viele $k \in \mathbb{N}$ mit $c_k < 0$, sowie unendlich viele $k \in \mathbb{N}$ mit $c_k > 0$.

  - Existieren nur endlich viele $k \in \mathbb{N}$ mit $c_k < 0$, so existiert ein $N \in \mathbb{N}$ mit ($c_k = |c_k|$) $c_k \geq 0$ für alle $k \geq N$. Dann konvergiert $\sum_{k=0}^{\infty} |c_k|$ wegen Übung 75, im Widerspruch zur Voraussetzung.
  - Existieren nur endlich viele $k \in \mathbb{N}$ mit $c_k > 0$, so existiert ein $N \in \mathbb{N}$ mit ($c_k = -|c_k|$) $c_k \leq 0$ für alle $k \geq N$. Dann konvergiert $\sum_{k=0}^{\infty} -|c_k|$ wegen Übung 75; also auch $-\sum_{k=0}^{\infty} -|c_k| = \sum_{k=0}^{\infty} |c_k|$ wegen Korollar 23, im Widerspruch zur Voraussetzung.

- Wir setzen $c[+]_n := \max(c_n, 0)$ und $c[-]_n := -\min(c_n, 0)$ für alle $n \in \mathbb{N}$. Dann sind $(c[\pm]_n)_{n\in\mathbb{N}}$ (positive) Folgen in $[0, \infty)$, mit $c_n = c[+]_n - c[-]_n$ und $c[+]_n = 0 \;\vee\; c[-]_n = 0$, für alle $n \in \mathbb{N}$. Wegen Lemma 37, gilt $\lim_n c_n = 0$; und somit auch $\lim_n c[\pm]_n = 0$.
- Aus der Annahme, dass $\sum_{k=0}^{\infty} c_k$ nicht absolut konvergiert, folgt

$$\textstyle\sum_{n=0}^{\infty} c[-]_n = \infty \qquad \text{also auch} \qquad \sum_{n=N}^{\infty} c[+]_n = \infty \quad \forall\, N \in \mathbb{N}. \tag{172}$$

  (Es gelte beispielsweise $\sum_{n=0}^{\infty} c[+]_n \neq \infty$, also $\sum_{n=0}^{\infty} c[+]_n < \infty$ (vgl. Notation 23). Da $\sum_{k=0}^{\infty} c_k$ per Annahme konvergiert, folgt aus Korollar 23, dass auch

$$\textstyle\sum_{k=0}^{\infty} c[+]_n - \sum_{k=0}^{\infty} c_n = \sum_{k=0}^{\infty} \underbrace{(c[+]_n - c_n)}_{= c[-]_n} = \sum_{k=0}^{\infty} c[-]_n$$

  konvergiert; und somit $\sum_{k=0}^{\infty} c[+]_k + \sum_{k=0}^{\infty} c[-]_k = \sum_{k=0}^{\infty} c[+]_k + c[-]_k = \sum_{k=0}^{\infty} |c_k|$.

Man summiert nun abwechselnd (in steigender Reihenfolge) immer Folgenglieder von $(c_n)_{n\in\mathbb{N}}$ mit $c_n = c[+]_n$ (positive Folgenglieder) und $c_n = -c[-]_n$ (negative Folgenglieder) auf; und zwar derart, dass die Summe im positiven/negativen Fall gerade wieder $>$ / $<$ x wird. Dies ist wegen (172) möglich; und definiert eine Bijektion $\alpha\colon \mathbb{N} \to \mathbb{N}$. Aus der Eigenschaft $\lim_n c[\pm]_n = 0$, folgt dann $\sum_{k=\mathsf{p}}^{\infty} c_{\alpha(k)} = x$. $\square$

Satz 24 (Die Cauchy-Produktformel)**.** *Es seien* $\sum_{n=0}^{\infty} a_n$ *und* $\sum_{n=0}^{\infty} b_n$ *zwei absolut konvergente* $\mathbb{K}$*-Reihen. Wir setzen*

$$c_n := \textstyle\sum_{k=0}^{n} a_{n-k} \cdot b_k \qquad \forall\, n \in \mathbb{N}. \tag{173}$$

Dann ist die Reihe $\sum_{n=0}^{\infty} c_n$ *absolut konvergent, und es gilt*

$$\textstyle\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n\right) \cdot \left(\sum_{n=0}^{\infty} b_n\right). \tag{174}$$

Beweis. Wir setzen

$$A_n := \textstyle\sum_{k=0}^{n} a_k, \qquad B_n := \sum_{k=0}^{n} b_k, \qquad C_n := \sum_{k=0}^{n} c_k, \qquad D_n := A_n \cdot B_n$$
$$A := \lim_n A_n, \qquad B := \lim_n B_n;$$

d.h., $\lim_n D_n = A \cdot B$ gemäß Korollar 17.3), sowie

$$C_n = \textstyle\sum_{k=0}^{n} \sum_{\ell=0}^{k} a_{k-\ell} \cdot b_\ell = \sum_{\{0\leq i,j\leq n \,|\, i+j\leq n\}} a_i \cdot b_j \tag{175}$$

- Wir zeigen zunächst (174), also $\sum_{n=0}^{\infty} c_n = \lim_n C_n = A \cdot B$:

  Hierfür genügt es, $\lim_n(C_n - D_n) = 0$ nachzuweisen; denn dann zeigt Korollar 23 (dritter Schritt):

$$A \cdot B = \lim_n D_n = \lim_n(C_n - D_n) + \lim_n D_n = \lim_n((C_n - D_n) + D_n) = \lim_n C_n.$$

  Wir argumentieren wie folgt:

  - Für jedes $n \in \mathbb{N}$, sei

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

    Wegen der absoluten Konvergenz der Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$, existiert nun gemäß Korollar 17.3) der Grenzwert

$$P := \lim_n P_n = \left(\textstyle\sum_{k=0}^{\infty} |a_k|\right) \cdot \left(\sum_{k=0}^{\infty} |b_k|\right) \in \mathbb{K}.$$

    Für $\varepsilon > 0$ vorgegeben, existiert daher wegen dem Cauchy-Kriterium Satz 20 (oder auch Lemma 35.1)), ein $N_\varepsilon \in \mathbb{N}$ mit

$$|P_n - P_m| < \varepsilon \qquad \forall\, n, m \geq N_\varepsilon. \tag{176}$$

– Für $n \in \mathbb{N}$, setzen wir $J[n] := \{(i,j) \in \{0,\dots,n\}^2 \mid i+j \geq n+1\}$.
Für $n \geq 2N_\varepsilon$, gilt dann

$$0 \leq i,j \leq N_\varepsilon \quad \Longrightarrow \quad i+j \leq 2N_\varepsilon \leq n \quad \Longrightarrow \quad (i,j) \notin J[n],$$

und daher

$$J[n] \cap \{0,\dots,N_\varepsilon\}^2 = \emptyset \qquad \text{also} \qquad J[n] \subseteq \{0,\dots,n\}^2 \setminus \{0,\dots,N_\varepsilon\}^2 \tag{177}$$

$$\{0,\dots,N_\varepsilon\}^2 \subseteq \{0,\dots,n\}^2. \tag{178}$$

Hiermit erhalten wir für $n \geq 2N_\varepsilon$

$$\begin{aligned} \varepsilon \overset{(176)}{>} \left|P_n - P_{N_\varepsilon}\right| &= \left|\textstyle\sum_{0\leq i,j\leq n} |a_i|\cdot|b_j| - \sum_{0\leq i,j\leq N_\varepsilon} |a_i|\cdot|b_j|\right| \\ &= \left|\textstyle\sum_{(i,j)\in\{0,\dots,n\}^2} |a_i|\cdot|b_j| - \sum_{(i,j)\in\{0,\dots,N_\varepsilon\}^2} |a_i|\cdot|b_j|\right| \\ &\overset{(178)}{=} \textstyle\sum_{(i,j)\in\{0,\dots,n\}^2\setminus\{0,\dots,N_\varepsilon\}^2} |a_i|\cdot|b_j| \\ &\overset{(177)}{\geq} \textstyle\sum_{(i,j)\in J[n]} |a_i|\cdot|b_j|. \end{aligned} \tag{179}$$

– Wir erhalten für $n \geq 2N_\varepsilon$, dass (vgl. (175))

$$\begin{aligned} |D_n - C_n| = |A_n \cdot B_n - C_n| &= \left| \underbrace{\left(\sum_{k=0}^{n} a_k\right)\cdot\left(\sum_{k=0}^{n} b_k\right)}_{\sum_{(i,j)\in\{0,\dots,n\}^2} a_i\cdot b_j} - \sum_{\{0\leq i,j\leq n \mid i+j\leq n\}} a_i \cdot b_j \right| \\ &= \left| \sum_{(i,j)\in\{0,\dots,n\}^2} a_i\cdot b_j - \sum_{(i,j)\in\{0,\dots,n\}^2\setminus J[n]} a_i\cdot b_j \right| \\ &= \left| \sum_{(i,j)\in J[n]} a_i\cdot b_j \right| \leq \sum_{(i,j)\in J[n]} |a_i|\cdot|b_j| < \varepsilon \end{aligned}$$

gilt; und somit folgt $\lim_n(C_n - D_n) = 0$.

- Es bleibt die absolute Konvergenz von $\sum_{n=0}^{\infty} c_n$ nachzuweisen. Hierfür betrachten wir die Absolutreihen $\sum_{n=0}^{\infty} |a_n|$ und $\sum_{n=0}^{\infty} |b_n|$, die per Annahme ebenfalls konvergieren (und da positiv, ebenfalls absolut konvergieren). Für alle $n \in \mathbb{N}$, setzen wir

$$\tilde{a}_n := |a_n|, \qquad \tilde{b}_n := |b_n|, \qquad \tilde{c}_n := \textstyle\sum_{k=0}^{n} \tilde{a}_{n-k}\cdot\tilde{b}_k = \sum_{k=0}^{n} |a_{n-k}|\cdot|b_k| \geq 0;$$

sodass also $\tilde{c}_n$ wie in (173) definiert ist, sofern man dort die Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ durch die Reihen $\sum_{n=0}^{\infty} \tilde{a}_n = \sum_{n=0}^{\infty} |a_n|$ und $\sum_{n=0}^{\infty} \tilde{b}_n = \sum_{n=0}^{\infty} |b_n|$ ersetzt.

– Da die Reihen $\sum_{n=0}^{\infty} \tilde{a}_n$, $\sum_{n=0}^{\infty} \tilde{b}_n$ absolut konvergieren, konvergiert nach dem bereits Gezeigten auch $\sum_{n=0}^{\infty} \tilde{c}_n$. Da die Reihe $\sum_{n=0}^{\infty} \tilde{c}_n$ positiv ist, zeigt Satz 18:

$$\sup_{\mathbb{R}}(\{\textstyle\sum_{k=0}^{n} \tilde{c}_k \mid n \in \mathbb{N}\}) = \sum_{n=0}^{\infty} \tilde{c}_n < \infty.$$

– Für alle $n \in \mathbb{N}$, gilt nun

$$|c_n| = |\textstyle\sum_{k=0}^{n} a_{n-k}\cdot b_k| \leq \sum_{k=0}^{n} |a_{n-k}|\cdot|b_k| = \tilde{c}_n; \tag{180}$$

also erhalten wir für die n-te Partialsumme von $\sum_{n=0}^{\infty} |c_n|$:

$$|\textstyle\sum_{k=0}^{n} |c_n|| = \sum_{k=0}^{n} |c_n| \leq \sum_{k=0}^{n} \tilde{c}_n \leq \sup_{\mathbb{R}}(\{\sum_{k=0}^{n} \tilde{c}_k \mid n \in \mathbb{N}\}) < \infty.$$

Daher ist die positive Reihe $\sum_{k=0}^{\infty} |c_n|$ beschränkt, konvergiert also gemäß Satz 18. $\square$

Bemerkung 43. *Die Voraussetzung der absoluten Konvergenz der Reihen ist essentiell für die Gültigkeit der Cauchyschen Produktformel. Als Beispiel betrachten wir die durch*

$$a_n = b_n = \tfrac{(-1)^n}{\sqrt{n+1}} \qquad \forall\, n \in \mathbb{N}$$

gegebenen Reihen $\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} b_n$*, die nach Satz 19 (Leibnizkriterium) und Übung 66 konvergieren.*[28] *Andererseits ergibt sich*

$$c_n := \sum_{k=0}^{n} a_{n-k} \cdot b_k = (-1)^n \cdot \underbrace{\sum_{k=0}^{n} \tfrac{1}{\sqrt{n-k+1}} \cdot \tfrac{1}{\sqrt{k+1}}}_{=: \tilde{c}_n}.$$

Um zu zeigen, dass die Reihe $\sum_{n=0}^{\infty} c_n$ *divergiert, genügt nach Satz 19 (Leibnizkriterium) nachzuweisen, dass* $(c_n)_{n\in\mathbb{N}}$ *keine Nullfolge ist: Für alle* $n \in \mathbb{N}$ *gilt*

$$\tfrac{1}{\sqrt{n-k+1}} \cdot \tfrac{1}{\sqrt{k+1}} = \tfrac{1}{\sqrt{(n-k+1)\cdot(k+1)}} \geq \tfrac{1}{\sqrt{(n+1)^2}} = \tfrac{1}{n+1} \qquad \forall\, 0 \leq k \leq n$$

$$\Longrightarrow \qquad c_n = \sum_{k=0}^{n} \tfrac{1}{\sqrt{n-k+1}} \cdot \tfrac{1}{\sqrt{k+1}} \geq \sum_{k=0}^{n} \tfrac{1}{n+1} = \tfrac{n+1}{n+1} = 1,$$

und daher ist $(c_n)_{n\in\mathbb{N}}$ *keine Nullfolge.*

6.3 Konvergenzkriterien für Reihen

Im Folgenden sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Das Majorantenkriterium

Definition 45. *Eine positive Reihe* $\sum_{n=\mathsf{q}}^{\infty} a_n$ *heißt* Majorante *der* $\mathbb{K}$*-Reihe* $\sum_{n=\mathsf{p}}^{\infty} c_n$*, wenn ein* $N \geq \max(\mathsf{p}, \mathsf{q})$ *mit* $|c_n| \leq a_n$ *für alle* $n \geq N$ *existiert.*

Satz 25 (Majorantenkriterium)**.** *Eine* $\mathbb{K}$*-Reihe mit konvergenter Majorante konvergiert absolut.*

Beweis. Sei $\sum_{k=\mathsf{q}}^{\infty} a_k$ eine konvergente Majorante der $\mathbb{K}$-Reihe $\sum_{n=\mathsf{p}}^{\infty} c_n$; und sei $C := \sum_{k=\mathsf{p}}^{N} |c_k|$ für $N \geq \max(\mathsf{p}, \mathsf{q})$ wie in Definition 45. Wir erhalten mit Satz 18 (letzter Schritt):

$$\sum_{k=\mathsf{p}}^{n} |c_k| \leq C + \sum_{k=N+1}^{n} a_k \leq C + \sum_{k=\mathsf{q}}^{n} a_k \leq C + \underbrace{\sup_{\mathbb{R}}(\{\sum_{k=\mathsf{q}}^{n} a_k \mid n \in \mathsf{N}\})}_{\sum_{k=\mathsf{q}}^{\infty} a_k} < \infty \qquad \forall\, n \in \mathsf{N}.$$

Daher ist die positive Reihe $\sum_{n=\mathsf{p}}^{\infty} |c_n|$ beschränkt, also konvergent nach Satz 18. ☐

Beispiel 47 (Minorantenkriterium)**.** *Sei* $\sum_{n=\mathsf{q}}^{\infty} a_n$ *eine* Majorante *der* $\mathbb{K}$*-Reihe* $\sum_{n=\mathsf{p}}^{\infty} c_n$*. Konvergiert* $\sum_{n=\mathsf{p}}^{\infty} c_n$ *nicht absolut, so divergiert* $\sum_{n=\mathsf{q}}^{\infty} a_n$ *wegen Satz 25.*

(Beispielsweise könnte $\sum_{n=\mathsf{p}}^{\infty} c_n$ *eine divergente positive Reihe sein, d.h.,* $\sum_{n=\mathsf{p}}^{\infty} c_n = \infty$*.)*

Satz 26 (Verdichtungskriterium von Cauchy)**.** *Sei* $(c_n)_{n\in\mathbb{N}}$ *eine monoton fallende Folge in* $[0, \infty)$*. Dann gilt die Äquivalenz:*

$$\sum_{n=1}^{\infty} c_n \ \textit{ konvergiert} \qquad \Longleftrightarrow \qquad \sum_{n=1}^{\infty} 2^n \cdot c_{2^n} \ \textit{ konvergiert.}$$

Anschaulich folgt diese Aussage, wenn man die Glieder der Reihe wie angedeutet gruppiert:

$$c_1 + \overbrace{c_2 + c_3}^{\leq 2\cdot c_2} + \overbrace{c_4 + \ldots + c_7}^{\leq 4\cdot c_4} + \overbrace{c_8 + \ldots + c_{15}}^{\leq 8\cdot c_8} + \ldots$$

$$c_1 + c_2 + \underbrace{c_3 + c_4}_{\geq 2\cdot c_4} + \underbrace{c_5 + \ldots + c_8}_{\geq 4\cdot c_8} + \underbrace{c_9 + \ldots + c_{16}}_{\geq 8\cdot c_{16}} + \ldots.$$

[28] Wie wir in Proposition 7 weiter unten sehen werden, konvergieren diese Reihen nicht absolut.

$*$Beweis. Für $k \geq 1$ gilt offensichtlich

$$\sum_{n=2}^{2^{k+1}-1} c_n = \sum_{\ell=1}^{k} \sum_{n=2^\ell}^{2^{\ell+1}-1} c_n \qquad \text{sowie} \qquad \sum_{n=3}^{2^k} c_n = \sum_{\ell=1}^{k} \sum_{n=2^{\ell-1}+1}^{2^\ell} c_n.$$

- Da $(c_n)_{n\in\mathbb{N}}$ monoton fallend ist, gilt

$$\sum_{n=2^\ell}^{2^{\ell+1}-1} c_n \leq 2^\ell \cdot c_{2^\ell} \quad \forall\, 1 \leq \ell \leq k \qquad \Longrightarrow \qquad \sum_{n=2}^{2^{k+1}-1} c_n \leq \sum_{\ell=1}^{k} 2^\ell \cdot c_{2^\ell}. \tag{181}$$

  Ist nun $\sum_{n=1}^{\infty} 2^n \cdot c_{2^n}$ konvergent, so wegen Satz 18 auch beschränkt. Dann ist $\sum_{n=0}^{\infty} c_n$ wegen (181) ebenfalls beschränkt, sodass Satz 18 die Konvergenz von $\sum_{n=0}^{\infty} c_n$ zeigt.

- Umgekehrt gilt

$$2 \cdot \sum_{n=2^{\ell-1}+1}^{2^\ell} c_n \geq 2 \cdot 2^{\ell-1} \cdot c_{2^\ell} = 2^\ell \cdot c_{2^\ell} \quad \forall\, 1 \leq \ell \leq k \qquad \Longrightarrow \qquad 2 \cdot \sum_{n=3}^{2^k} c_n \geq \sum_{\ell=1}^{k} 2^\ell \cdot c_{2^\ell}.$$

  Konvergiert $\sum_{n=0}^{\infty} c_n$, so auch $\sum_{n=1}^{\infty} 2^n \cdot c_{2^n}$; und zwar nach der gleichen Argumentation wie im vorherigen Punkt. □

Proposition 7. *Sei $\alpha \in \mathbb{Q}$. Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ konvergiert genau dann, wenn $\alpha > 1$ gilt.*

$*$Beweis.
- Für $\alpha < 0$ und $n \in \mathbb{N}_{>0}$ gilt $\frac{1}{n^\alpha} = n^{-\alpha} \geq 1^{-\alpha} = 1$ wegen Übung 52.c), also konvergiert $\left(\frac{1}{n^\alpha}\right)_{n\in\mathbb{N}_{>0}}$ nicht gegen Null. Wegen Lemma 37 ist die Reihe $\sum_{n=0}^{\infty} \frac{1}{n^\alpha}$ daher divergent.

- Sei $\alpha \geq 0$. Für alle $n \in \mathbb{N}_{>0}$, gilt dann $n^\alpha \leq (n+1)^\alpha$ gemäß Übung 52.c), also $\frac{1}{(n+1)^\alpha} \leq \frac{1}{n^\alpha}$. Satz 26 zeigt daher

$$\sum_{n=0}^{\infty} \frac{1}{n^\alpha} < \infty \qquad \Longleftrightarrow \qquad \sum_{n=0}^{\infty} \frac{2^n}{(2^n)^\alpha} < \infty.$$

  Für $\mathbb{R} \ni x > 0$ und $\alpha, \beta \in \mathbb{Q}$, gilt nun (siehe Übung 76 oder Bemerkung 56.c) in Abschnitt 7.3.2)

$$x^\alpha \cdot x^\beta = x^{\alpha+\beta} \qquad \text{sowie} \qquad (x^\alpha)^\beta = x^{\alpha\cdot\beta}.$$

  Hiermit erhalten wir

$$\sum_{n=1}^{\infty} \frac{2^n}{(2^n)^\alpha} = \sum_{n=0}^{\infty} 2^{n\cdot(1-\alpha)} = \sum_{n=1}^{\infty} (2^{1-\alpha})^n.$$

  Dies ist eine geometrische Reihe, die gemäß Korollar 22 genau dann konvergiert, wenn $2^{1-\alpha} < 1$ gilt; also $1 - \alpha < 0$ wegen Übung 52.c)[29], und daher $\alpha > 1$. □

Mit Hilfe der Theorie der Fourierreihen, kann man zeigen:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}, \qquad \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}, \qquad \sum_{n=1}^{\infty} \frac{1}{n^6} = \frac{\pi^6}{945}.$$

Übung 76. *Sei $\mathbb{R} \ni x > 0$, sowie $\alpha, \beta \in \mathbb{Q}$. Zeigen Sie $x^\alpha \cdot x^\beta = x^{\alpha+\beta}$ sowie $(x^\alpha)^\beta = x^{\alpha\cdot\beta}$.*

Das Quotienten- und das Wurzelkriterium

Im Beweis von Proposition 7, hatte wir bereits die Konvergenzeigenschaften der geometrische Reihe verwendet. Ein weiteres Konvergenzkriterium, welches aus den Eigenschaften der geometrischen Reihe folgt, ist das sogenannte Quotientenkriterium:

Satz 27 (Quotientenkriterium)**.** *Sei $\sum_{n=\mathsf{p}}^{\infty} c_n$ eine $\mathbb{K}$-Reihe, sodass $N \in \mathsf{N}$ und $0 \leq \alpha < 1$ existieren, mit $|c_{n+1}| \leq \alpha \cdot |c_n|$ für alle $n \geq N$. Dann konvergiert die Reihe $\sum_{n=\mathsf{p}}^{\infty} c_n$ absolut.*

[29] Ist $1 - \alpha > 0$, so folgt $2^{1-\alpha} > 1^{1-\alpha} = 1$ wegen Übung 52.c). Ist $1 - \alpha = 0$, so folgt $2^{1-\alpha} = 1$.

Beweis. Für $m \in \mathbb{N}$, folgt induktiv $|c_{N+m}| \leq \alpha^m \cdot |c_N|$. Wegen Korollar 22 (geometrische Reihe) sowie Korollar 23 (Rechenregel für Limiten), gilt

$$\sum_{k=N}^{\infty} |c_N| \cdot \alpha^{k-N} = |c_N| \cdot \sum_{m=0}^{\infty} \alpha^m = |c_N| \cdot \tfrac{1}{1-\alpha} \qquad \text{mit} \qquad c_k \leq |c_N| \cdot \alpha^{k-N} \quad \forall\, k \geq N.$$

Somit ist $\sum_{k=N}^{\infty} |c_N| \cdot \alpha^{k-N}$ eine konvergente Majorante von $\sum_{n=\mathsf{p}}^{\infty} c_n$; und die Behauptung folgt aus Satz 25. □

Korollar 24. *Sei* $\sum_{n=\mathsf{p}}^{\infty} c_n$ *eine* $\mathbb{K}$*-Reihe, sodass ein* $N \in \mathsf{N}$ *existiert mit* $c_n \neq 0$ *für alle* $n \geq N$*.*

a) Gilt $\overline{\lim}_n \frac{|c_{n+1}|}{|c_n|} < 1$*, so konvergiert die Reihe* $\sum_{n=\mathsf{p}}^{\infty} c_n$ *absolut.*

b) Gilt $\underline{\lim}_n \frac{|c_{n+1}|}{|c_n|} > 1$*, so divergiert die Reihe* $\sum_{n=\mathsf{p}}^{\infty} c_n$*.*

(Wegen Übung 71, hängen die entsprechenden Bedingungen nicht von der expliziten Wahl des Index $N \in \mathsf{N}$ ab.)

Beweis. Sei $\tilde{\mathsf{N}} := \{n \in \mathbb{Z} \mid n \geq N\}$, und $a_n := \frac{|c_{n+1}|}{|c_n|}$ für alle $n \in \tilde{\mathsf{N}}$. Seien weiterhin U, O wie in (159) und (160) (für die Folge $(a_n)_{n \in \tilde{\mathsf{N}}}$), d.h., also

$$\begin{aligned} \mu &:= \inf\nolimits_{\overline{\mathbb{R}}}(O) = \overline{\lim}_n\, a_n = \overline{\lim}_n \tfrac{|c_{n+1}|}{|c_n|} \\ \nu &:= \sup\nolimits_{\overline{\mathbb{R}}}(U) = \underline{\lim}_n\, a_n = \underline{\lim}_n \tfrac{|c_{n+1}|}{|c_n|}. \end{aligned}$$

a) Sei $\overline{\lim}_n \frac{|c_{n+1}|}{|c_n|} = \mu < \alpha < 1$ fixiert, und setzte $\varepsilon := \alpha - \mu > 0$.

Wegen Lemma 20, existiert ein $x \in O$ mit $x < \mu + \varepsilon = \alpha$. Wegen $x \in O$, existiert weiterhin ein $\tilde{N} \geq N$ mit

$$\frac{|c_{n+1}|}{|c_n|} = a_n \leq x < \alpha \quad \forall\, n \geq \tilde{N} \qquad \Longrightarrow \qquad |c_{n+1}| \leq \alpha \cdot |c_n| \quad \forall\, n \geq \tilde{N}.$$

Daher folgt die Behauptung aus Satz 27.

b) Sei $\underline{\lim}_n \frac{|c_{n+1}|}{|c_n|} = \nu > 1$, und setzte $\varepsilon := \nu - 1 > 0$.

Wegen Lemma 20, existiert $x \in U$ mit $x > \nu - \varepsilon = 1$. Wegen $x \in U$, existiert ein $\tilde{N} \geq N$, mit

$$\frac{|c_{n+1}|}{|c_n|} = a_n \geq x > 1 \quad \forall\, n \geq \tilde{N} \qquad \Longrightarrow \qquad |c_{n+1}| \geq |c_n| \quad \forall\, n \geq \tilde{N}.$$

Induktiv erhalten wir $|c_{\tilde{N}+m}| \geq |c_{\tilde{N}}| > 0$ für alle $m \in \mathbb{N}$; sodass die Folge $(c_n)_{n \in \mathsf{N}}$ nicht gegen 0 konvergieren kann. Die Behauptung folgt nun aus Lemma 37. □

Beispiel 48. *Die Reihe* $\sum_{n=0}^{\infty} \frac{n^2}{2^n}$ *konvergiert. Sei nämlich* $c_n := \frac{n^2}{2^n} \neq 0$ *für alle* $n \in \mathbb{N}_{>0}$*. Dann gilt wegen Beispiel 34.b), Lemma 32 und Korollar 17.4) (dritter Schritt):*

$$\lim_n \frac{|c_{n+1}|}{|c_n|} = \lim_n \frac{\frac{(n+1)^2}{2^{n+1}}}{\frac{n^2}{2^n}} = \lim_n \frac{(n+1)^2}{n^2} \cdot \frac{1}{2} = \frac{1}{2} < 1.$$

Korollar 20.a) zeigt $\overline{\lim}_n \frac{|c_{n+1}|}{|c_n|} = \lim_n \frac{|c_{n+1}|}{|c_n|} < 1$*. Die Behauptung folgt daher aus Korollar 24.a).*

Bemerkung 44.

a) Für die (divergente) harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$*, gilt* $\frac{|c_{n+1}|}{|c_n|} = \frac{n}{n+1} < 1$ *für alle* $n \in \mathbb{N}_{>0}$*:*

- *Satz 27, liefert keine Konvergenzaussage; denn wegen* $\lim_n \frac{n}{n+1} = 1$ *(Beispiel 34.b)), existiert kein* $0 \leq \alpha < 1$ *mit* $\frac{1}{n+1} \leq \alpha \cdot \frac{1}{n}$ *(bzw.* $\frac{n}{n+1} \leq \alpha$*) für fast alle* $n \in \mathbb{N}_{>0}$.
- *Das Quotientenkriterium in Form von Korollar 24 ist ebenfalls nicht anwendbar; bzw., liefert keine Information über Konvergenz oder Divergenz; denn es gilt (Korollar 20.a)):*

$$1 = \lim_n \tfrac{n}{n+1} = \underline{\lim}_n \tfrac{n}{n+1} = \overline{\lim}_n \tfrac{n}{n+1}.$$

Die Bedingung $\underline{\lim}_n \frac{|c_{n+1}|}{|c_n|} > 1$ *ist daher* nur hinreichend *für die Divergenz der Reihe, aber* nicht notwendig.

b) *Trotzdem die Reihe* $\sum_{n=1}^{\infty} \frac{1}{n^2}$ *konvergiert (vgl. Beispiel 44 bzw. Proposition 7), gilt*

$$\overline{\lim}_n \tfrac{|c_{n+1}|}{|c_n|} = \overline{\lim}_n \left(\tfrac{n}{n+1}\right)^2 = \lim_n \left(\tfrac{n}{n+1}\right)^2 = 1.$$

Die Bedingung $\overline{\lim}_n \frac{|c_{n+1}|}{|c_n|} < 1$ *ist daher* nur hinreichend *für die Konvergenz der Reihe, aber* nicht notwendig.

Das nächste Kriterium ist häufiger anwendbar (vgl. Bemerkung 46.b)), allerdings oft schwieriger auszuwerten.

Satz 28 (Wurzelkriterium)**.** *Sei* $\sum_{n=\mathsf{p}}^{\infty} c_n$ *eine* $\mathbb{K}$*-Reihe.*

a) *Die Reihe* $\sum_{n=\mathsf{p}}^{\infty} c_n$ *konvergiert absolut, falls gilt*

$$\overline{\lim}_n \sqrt[n]{|c_n|} < 1. \tag{182}$$

b) *Gilt* $\sqrt[n]{|c_n|} \geq 1$ *für unendlich viele* $n \in \mathsf{N}$*, so divergiert die Reihe* $\sum_{n=\mathsf{p}}^{\infty} c_n$.

Insbesondere ist dies der Fall für $\overline{\lim}_n \sqrt[n]{|c_n|} > 1$.

Beweis. Sei $a_n := \sqrt[n]{|c_n|}$ für alle $n \in \mathsf{N}$, sowie O wie in (159); also

$$\mu := \inf_{\overline{\mathbb{R}}}(O) = \overline{\lim}_n a_n = \overline{\lim}_n \sqrt[n]{|c_n|}.$$

a) Sei $\overline{\lim}_n \sqrt[n]{|c_n|} = \mu < \alpha < 1$ fixiert, und setzte $\varepsilon := \alpha - \mu > 0$.

Wegen Lemma 20, existiert ein $x \in O$ mit $x < \mu + \varepsilon = \alpha$. Wegen $x \in O$, existiert ein $N \in \mathsf{N}$ mit

$$\sqrt[n]{|c_n|} = a_n \leq x < \alpha \quad \forall\, n \geq N \qquad \Longrightarrow \qquad |c_n| \leq \alpha^n \quad \forall\, n \geq N.$$

Daher ist die geometrische Reihe $\sum_{n=0}^{\infty} \alpha^n$ eine konvergente Majorante von $\sum_{n=\mathsf{p}}^{\infty} c_n$; also folgt die Behauptung aus Satz 25 (Majorantenkriterium).

b)
- Gilt $\sqrt[n]{|c_n|} \geq 1$ für unendlich viele $n \in \mathsf{N}$, so gilt (wegen Übung 45) auch $|c_n| \geq 1$ für unendlich viele $n \in \mathsf{N}$. Daher kann die Folge $(c_n)_{n\in\mathsf{N}}$ nicht gegen 0 konvergieren; also divergiert die Reihe $\sum_{n=1}^{\infty} c_n$ gemäß Lemma 37.
- Es gelte $\overline{\lim}_n \sqrt[n]{|c_n|} > 1$. Dann gilt $1 < \inf_{\overline{\mathbb{R}}}(O) \leq O$, also $1 \notin O$. Daher existiert für jedes $m \in \mathsf{N}$, ein $\ell \geq m$ mit $\sqrt[n]{|c_\ell|} > 1$. Also gibt es unendlich viele $n \in \mathsf{N}$ mit $\sqrt[n]{|c_n|} \geq 1$. □

Bemerkung 45. *Man beachte den Unterschied zwischen den Divergenzkriterien in Korollar 27.b) und Satz 28.b) (*$\underline{\lim}$ *bzw.* $\overline{\lim}$*).*

Um das Wurzelkriteriun anwenden zu können, ist es wichtig Grenzwerte von Folgen der Art $\sqrt[n]{|c_n|}$ zu kennen. Das nächste Lemma ist hierfür sehr hilfreich:

Lemma 38. *Es gilt*

$$\lim_n \sqrt[n]{n} = 1 \qquad \textit{sowie} \qquad \lim_n \sqrt[n]{c} = 1 \quad \forall\, c > 0,$$

und daher wegen Korollar 17

$$\lim_n \sqrt[n]{n^k} = \lim_n(\sqrt[n]{n})^k = (\lim_n \sqrt[n]{n})^k = 1 \qquad \forall\, k \in \mathbb{Z}.$$

$*$*Beweis:* • Wir zeigen $\lim_n \sqrt[n]{n} = 1$:

Sei $\varepsilon > 0$ vorgegeben. Wir müssen zeigen, dass ein $N_\varepsilon \in \mathbb{N}_{>0}$ existiert, sodass $|1 - \sqrt[n]{n}| < \varepsilon$ für alle $n \geq N_\varepsilon$ gilt.

- Wegen $1 \leq \sqrt[n]{n}$ für alle $n \in \mathbb{N}_{>0}$ (Übung 52) genügt es nachzuweisen, dass ein $N_\varepsilon \in \mathbb{N}_{>0}$ mit $\sqrt[n]{n} < 1 + \varepsilon$ für alle $n \geq N_\varepsilon$ existiert.
- Es genügt dann $\lim_n \frac{n}{(1+\varepsilon)^n} = 0$ zu zeigen; denn dann existiert ein $N_\varepsilon \in \mathbb{N}$ mit

$$\begin{aligned} \tfrac{n}{(1+\varepsilon)^n} < 1 \quad \forall\, n \geq N_\varepsilon \quad &\iff \quad n < (1+\varepsilon)^n \quad \forall\, n \geq N_\varepsilon \\ &\iff \quad \sqrt[n]{n} < 1 + \varepsilon \quad \forall\, n \geq N_\varepsilon. \end{aligned}$$

- Satz 6 (Binomischer Lehrsatz) liefert

$$(1+\varepsilon)^n = \sum_{k=0}^n \binom{n}{k} \cdot \varepsilon^k \geq \binom{n}{2} \cdot \varepsilon^2 = \frac{n\cdot(n-1)}{2} \cdot \varepsilon^2$$

(da alle Summanden nichtnegativ sind). Damit erhalten wir

$$0 \leq \frac{n}{(1+\varepsilon)^n} \leq \frac{2n}{n\cdot(n-1)\cdot\varepsilon^2} = \frac{2}{(n-1)\cdot\varepsilon^2} \qquad \text{mit} \qquad \lim_n \frac{2}{(n-1)\cdot\varepsilon^2} = 0.$$

Das Quetschlemma (Lemma 31) zeigt nun $\lim_n \frac{n}{(1+\varepsilon)^n} = 0$.

• Wir zeigen $\lim_n \sqrt[n]{c} = 1$:

- Sei zunächst $c \geq 1$; sowie $N \in \mathbb{N}$ mit $N \geq c$. Für alle $n \geq N$ gilt dann $1 \leq c \leq n$, d.h., $1 \leq \sqrt[n]{c} \leq \sqrt[n]{n}$ gemäß Übung 52. Wegen $\lim_n \sqrt[n]{n} = 1$ (vorheriger Punkt), zeigt Lemma 31 (Quetschlemma) $\lim_n \sqrt[n]{c} = 1$.
- Ist $0 < c < 1$, so gilt $c^{-1} > 1$ mit $\sqrt[n]{c} = (\sqrt[n]{c^{-1}})^{-1}$ wegen Lemma 22.a). Die Behauptung folgt nun aus dem bereits Bewiesenen, sowie Lemma 32.a). □

Bemerkung 46.

a) *Das Wurzelkriterium ist nur* hinreichend *für Konvergenz bzw. Divergenz, aber* nicht notwendig*:*

- *Wir betrachten die (divergente) harmonische Reihe $\sum_{n=1}^\infty \frac{1}{n}$, d.h., $c_n = \frac{1}{n}$ für alle $n \in \mathbb{N}_{>0}$. Satz 28 ist in diesem Fall nicht anwendbar:* *(Lemma 38)*

$$\begin{aligned} \overline{\lim}_n \sqrt[n]{|c_n|} = \lim_n \sqrt[n]{|c_n|} &= \lim_n(\sqrt[n]{n})^{-1} = 1 \\ \sqrt[n]{|c_n|} &= (\sqrt[n]{n})^{-1} < 1 \qquad \forall\, n \in \mathbb{N}_{>0}. \end{aligned}$$

- *Wir betrachte die (konvergente) Reihe $\sum_{n=1}^\infty \frac{1}{n^2}$, d.h., $c_n = \frac{1}{n^2}$ für alle $n \in \mathbb{N}_{>0}$. Satz 28 ist in diesem Fall nicht anwendbar:* *(Lemma 38)*

$$\begin{aligned} \overline{\lim}_n \sqrt[n]{|c_n|} = \lim_n \sqrt[n]{|c_n|} &= \lim_n(\sqrt[n]{n})^{-2} = 1 \\ \sqrt[n]{|c_n|} &= (\sqrt[n]{n})^{-2} < 1 \qquad \forall\, n \in \mathbb{N}_{>0}. \end{aligned}$$

b) *Ist das Quotientenkriterium (Satz 27) anwendbar, so auch das Wurzelkriterium (Satz 28.a)):*

Sei $\sum_{n=\mathsf{p}}^{\infty} c_n$ *eine* $\mathbb{K}$*-Reihe; sowie* $0 \leq \alpha < 1$ *und* $N \in \mathsf{N}$ *mit* $|c_{n+1}| \leq \alpha \cdot |c_n|$ *für alle* $n \geq N$ *(Voraussetzungen von Satz 27). Dann gilt* $\overline{\lim}_n \sqrt[n]{|c_n|} \leq \alpha < 1$ *(Voraussetzung Satz 28.a)).*

Beweis der Behauptung: Wir erhalten induktiv

$$|c_{N+m}| \leq \alpha^m \cdot |c_N| = \alpha^{N+m} \cdot \frac{|c_N|}{\alpha^N} \qquad \forall\, m \in \mathbb{N}.$$

Übung 74 (zweiter Schritt), Übung 69 (dritter Schritt), und Lemma 38 zeigt (vierter Schritt)

$$\begin{aligned}\overline{\lim}_n \sqrt[n]{|c_n|} &= \overline{\lim}_m \sqrt[N+m]{|c_{N+m}|} \leq \overline{\lim}_m \left(\alpha \cdot \sqrt[N+m]{\tfrac{|c_N|}{\alpha^N}}\right) \\ &= \alpha \cdot \overline{\lim}_m \sqrt[N+m]{\tfrac{|c_N|}{\alpha^N}} = \alpha \cdot \lim_m \sqrt[N+m]{\tfrac{|c_N|}{\alpha^N}} = \alpha < 1.\end{aligned}$$

□

c) *Für* $x \in \mathbb{R}$*, betrachten wir die Reihe* $\sum_{n=1}^{\infty} c_n$*, mit*

$$c_n := \begin{cases} x^n & \text{für } n \text{ gerade} \\ 2x^n & \text{für } n \text{ ungerade.} \end{cases}$$

- *Es gilt* $\overline{\lim}_n \sqrt[n]{|c_n|} = \lim_n \sqrt[n]{|c_n|} = |x|$ *wegen* $\lim \sqrt[n]{2} = 1$ *(Lemma 38). Das Wurzelkriterium liefert daher Konvergenz für* $|x| < 1$ *sowie Divergenz für* $|x| > 1$*.*
- *Es gilt*

$$\frac{|c_{n+1}|}{|c_n|} = \begin{cases} 2 \cdot |x| & \text{falls } n \text{ gerade} \\ \dfrac{|x|}{2} & \text{falls } n \text{ ungerade,} \end{cases}$$

  also wegen Lemma 34 und Übung 73

$$\overline{\lim}_n \frac{|c_{n+1}|}{|c_n|} = 2 \cdot |x| \qquad \text{sowie} \qquad \underline{\lim}_n \frac{|c_{n+1}|}{|c_n|} = \frac{|x|}{2}.$$

  Das Quotientenkriterium liefert daher Konvergenz für $|x| < \frac{1}{2}$ *und Divergenz für* $|x| > 2$*.*

Wie wir in Punkt b) bereits gesehen hatten, lässt das Quotientenkriterium im Allgemeinen also weniger scharfe Schlüsse zu als das Wurzelkriterium.

6.4 Potenzreihen und die Exponentialfunktion

Im Folgenden sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

6.4.1 Potenzreihen

Definition 46. *Eine Potenzreihe* P *in* $\mathbb{K}$*, ist eine Reihe der Gestalt* $\mathsf{P}(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$ *mit einer fixierte* $\mathbb{K}$*-Folge* $(c_n)_{n\in\mathbb{N}}$*, sowie* $z \in \mathbb{K}$*. Konvergiert für ein* $z \in \mathbb{K}$*, die Reihe* $\sum_{n=0}^{\infty} c_n \cdot z^n$*, so wird mit auch der entsprechenden Grenzwert mit* $\mathsf{P}(z)$ *bezeichnet.*

(Formell präziser ist P *eine Abbildung von* $\mathbb{K}$ *in den Raum der* $\mathbb{K}$*-Reihen, die jedem* $z \in \mathbb{K}$ *die* $\mathbb{K}$*-Reihe* $\mathsf{P}(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$ *zuordnet.)*

Beispiel 49. *Sei* $(c_n)_{n\in\mathbb{N}}$ *eine beschränkte* $\mathbb{K}$*-Folge. Dann konvergiert* $\mathsf{P}(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$ *für alle* $z \in \mathbb{K}$ *mit* $|z| < 1$ *absolut.*

Beweis. Sei $C \geq 0$ mit $|c_n| \leq C$ für alle $n \in \mathbb{N}$. Dann gilt

$$|c_n \cdot z^n| \leq C \cdot |z|^n \qquad \forall\, n \in \mathbb{N}.$$

Da die geometrische Reihe $\sum_{n=0}^{\infty} |z|^n$ gemäß Korollar 22 konvergiert, zeigt Korollar 23, dass gilt:

$$\textstyle\sum_{n=0}^{\infty} C \cdot |z|^n = C \cdot \sum_{n=0}^{\infty} |z|^n < \infty.$$

Daher ist $\sum_{n=0}^{\infty} C \cdot |z|^n$ eine konvergente Majorante von $\sum_{n=0}^{\infty} c_n \cdot z^n$, womit die Behauptung aus Satz 25 folgt. $\square$

Definition 47. *Sei* $\mathsf{P}(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$ *eine Potenzreihe in* $\mathbb{K}$.

- *Der* Kovergenzbereich *von* P*, ist definiert durch* *(beachte* $0 \in \Theta[\mathsf{P}] \neq \emptyset$*)*

$$\Theta[\mathsf{P}] := \left\{ w \in \mathbb{K} \,\middle|\, \textstyle\sum_{n=0}^{\infty} c_n \cdot w^n \ \textit{ konvergiert in } \ \mathbb{K} \right\}.$$

- *Der* Kovergenzradius *von* P*, ist definiert durch*

$$\mathsf{R}[\mathsf{P}] := \sup_{\overline{\mathbb{R}}}(\{\, |w| \,\big|\, w \in \Theta[\mathsf{P}]\}) \in [0, \infty].$$

- *Der* Konvergenzkreis *von* P*, ist definiert durch* $\mathrm{K}[\mathsf{P}] := \{z \in \mathbb{K} \mid |z| < \mathsf{R}[\mathsf{P}]\}$*, d.h.,*

$$\mathrm{K}[\mathsf{P}] = \begin{cases} \mathsf{B}_{\mathsf{R}[\mathsf{P}]}(0) & \textit{für} \quad \mathsf{R}[\mathsf{P}] \in (0,\infty) \\ \mathbb{K} & \textit{für} \quad \mathsf{R}[\mathsf{P}] = \infty \\ \emptyset & \textit{für} \quad \mathsf{R}[\mathsf{P}] = 0. \end{cases}$$

Wir setzen $\overline{\mathrm{K}}[\mathsf{P}] := \{z \in \mathbb{K} \mid |z| \leq \mathsf{R}[\mathsf{P}]\}$*, d.h.,*

$$\overline{\mathrm{K}}[\mathsf{P}] = \begin{cases} \overline{\mathsf{B}}_{\mathsf{R}[\mathsf{P}]}(0) & \textit{für} \quad \mathsf{R}[\mathsf{P}] \in (0,\infty) \\ \mathbb{K} & \textit{für} \quad \mathsf{R}[\mathsf{P}] = \infty \\ \{0\} & \textit{für} \quad \mathsf{R}[\mathsf{P}] = 0. \end{cases}$$

Satz 29. *Sei* $\mathsf{P}(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$ *eine Potenzreihe in* $\mathbb{K}$. *Es gelten die folgenden Aussagen:*

a) $\mathsf{P}(z)$ divergiert *für alle* $z \in \mathbb{K}$ *mit* $|z| > \mathsf{R}[\mathsf{P}]$. *(*$\Theta[\mathsf{P}] \subseteq \overline{\mathrm{K}}[\mathsf{P}]$*)*

b) $\mathsf{P}(z)$ konvergiert absolut *für alle* $z \in \mathrm{K}[\mathsf{P}]$. *(*$\mathrm{K}[\mathsf{P}] \subseteq \Theta[\mathsf{P}]$ *– Satz 21)*

c) Es gilt die Formel von Hadamard*:*

$$\mathsf{R}[\mathsf{P}] = \frac{1}{\overline{\lim}_n \sqrt[n]{|c_n|}} \in [0, \infty] \qquad \textit{mit} \qquad \frac{1}{0} := \infty \qquad \textit{und} \qquad \frac{1}{\infty} := 0. \tag{183}$$

Beweis. a) Sei $z \in \mathbb{K}$ mit $|z| > \mathsf{R}[\mathsf{P}]$. Dann gilt

$$|z| > \{|w| \mid w \in \Theta[\mathsf{P}]\} \qquad \Longrightarrow \qquad z \notin \Theta[\mathsf{P}].$$

Daher konvergiert die Reihe $\sum_{n=0}^{\infty} c_n \cdot z^n$ nicht – divergiert also.

b) Sei $z \in \mathrm{K}[\mathsf{P}]$ ($0 \leq |z| < \mathsf{R}[\mathsf{P}]$). Dann existiert ein $w \in \Theta[\mathsf{P}]$ mit $|z| < |w|$, und wir setzen

$$a_n := c_n \cdot z^n \qquad \text{sowie} \qquad b_n := c_n \cdot w^n \qquad \text{für alle} \qquad n \in \mathbb{N}.$$

Beweis der Existenz: Sei $\varepsilon := \mathsf{R}[\mathsf{P}] - |z| > 0$. Gemäß Lemma 20, existiert ein $\lambda \in \{\, |w| \,\big|\, w \in \Theta[\mathsf{P}]\}$ mit $\lambda > \mathsf{R}[\mathsf{P}] - \varepsilon = |z|$. Per definitionem gilt dann $\lambda = |w|$ für ein $w \in \Theta[\mathsf{P}]$, also $|z| < |w|$. $\square$

Wegen $w \in \Theta[\mathsf{P}]$, konvergiert $\mathsf{P}(w) = \sum_{n=0}^{\infty} b_n$ (divergiert also nicht). Wegen Satz 28.b) (Kontraposition), gilt daher $\overline{\lim}_n \sqrt[n]{|b_n|} \le 1$. Wir erhalten mit Übung 52.a) (linke Seite), sowie Übung 69 und $\frac{|z|}{|w|} < 1$ (rechte Seite):

$$\sqrt[n]{|a_n|} = \frac{|z|}{|w|} \cdot \sqrt[n]{|b_n|} \quad \forall\, n \in \mathbb{N} \qquad \Longrightarrow \qquad \overline{\lim}_n \sqrt[n]{|a_n|} = \frac{|z|}{|w|} \cdot \overline{\lim}_n \sqrt[n]{|b_n|} < 1.$$

Somit zeigt Satz 28.a), dass $\mathsf{P}(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$ absolut konvergiert.

c) Für $z \in \mathbb{K}$, gelten wegen Übung 52.a), Lemma 26, und Übung 69 (linke Seite), die Implikationen:

$$|z| \cdot \overline{\lim}_n \sqrt[n]{|c_n|} = \overline{\lim}_n \sqrt[n]{|c_n \cdot z^n|} \le 1 \qquad \Longrightarrow \qquad |z| \le \frac{1}{\overline{\lim}_n \sqrt[n]{|c_n|}} \tag{184}$$

$$|z| \cdot \overline{\lim}_n \sqrt[n]{|c_n|} = \overline{\lim}_n \sqrt[n]{|c_n \cdot z^n|} < 1 \qquad \Longleftarrow \qquad |z| < \frac{1}{\overline{\lim}_n \sqrt[n]{|c_n|}} \tag{185}$$

(unter Berücksichtigung der Rechenregeln für 0 und ∞ in (183)). Wir zeigen nun $\mathsf{R}[\mathsf{P}] \le (\overline{\lim}_n \sqrt[n]{|c_n|})^{-1}$ (erster Punkt), sowie $(\overline{\lim}_n \sqrt[n]{|c_n|})^{-1} \le \mathsf{R}[\mathsf{P}]$ (zweiter Punkt):

- Für $z \in \Theta[\mathsf{P}]$, konvergiert $\mathsf{P}(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$ per Definition. Wegen Satz 28.b) (Kontraposition), gilt daher die linke Seite von (184) – also auch die rechte. Wir erhalten:

$$|w| \le \tfrac{1}{\overline{\lim}_n \sqrt[n]{|c_n|}} \quad \forall\, w \in \Theta[\mathsf{P}] \qquad \Longrightarrow \qquad \mathsf{R}[\mathsf{P}] = \sup_{\mathbb{R}}(\{\, |w| \;\big|\; w \in \Theta[\mathsf{P}]\}) \le \tfrac{1}{\overline{\lim}_n \sqrt[n]{|c_n|}}.$$

- Für $z \in \mathbb{K}$ gelte die rechte Seite von (185) – also auch die linke. Dann zeigt Satz 28.a), dass die Reihe $\mathsf{P}(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$ absolut konvergiert; also konvergent ist nach Satz 21. Wegen Teil a) (Kontraposition), muss dann $|z| \le \mathsf{R}[\mathsf{P}]$ gelten. Wir erhalten daher:

$$\begin{aligned} &\Big(z \in \mathbb{K} \quad \wedge \quad |z| < \tfrac{1}{\overline{\lim}_n \sqrt[n]{|c_n|}} \qquad \Longrightarrow \qquad |z| \le \mathsf{R}[\mathsf{P}]\Big) \\ \Longrightarrow \qquad & \tfrac{1}{\overline{\lim}_n \sqrt[n]{|c_n|}} - \varepsilon \le \mathsf{R}[\mathsf{P}] \qquad \forall\ 0 < \varepsilon < \tfrac{1}{\overline{\lim}_n \sqrt[n]{|c_n|}} \\ \Longrightarrow \qquad & \tfrac{1}{\overline{\lim}_n \sqrt[n]{|c_n|}} \le \mathsf{R}[\mathsf{P}], \end{aligned}$$

wobei wir in der letzten Implikation Korollar 9 (für den Fall $\mathsf{R}[\mathsf{P}] < \infty$) benutzt haben. □

Bemerkung 47.

a) *Jede Potenzreihe $\mathsf{P}(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$ in $\mathbb{K}$, definiert eine Abbildung*

$$\mathsf{P}\colon \Theta[\mathsf{P}] \to \mathbb{K}, \qquad z \mapsto \mathsf{P}(z),$$

indem man $\mathsf{P}(z)$ (für $z \in \Theta[\mathsf{P}]$) als den Grenzwert $\sum_{n=0}^{\infty} c_n \cdot z^n \in \mathbb{K}$ auffasst.

(In der Praxis ist eher die Einschränkung $\mathsf{P}|_{\mathrm{K}[\mathsf{P}]}\colon \mathrm{K}[\mathsf{P}] \to \mathbb{K}$ von Belang, da hier sogar absolute Konvergenz der entsprechenden Reihe vorliegt.)

b) *Jede Potenzreihe $\mathsf{P}(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$ in $\mathbb{R}$ (d.h., $c_n \in \mathbb{R}$ für alle $n \in \mathsf{N}$), lässt sich auch als Potenzreihe $\tilde{\mathsf{P}}$ in $\mathbb{C}$ auffassen; und zwar indem man $(c_n)_{n\in\mathbb{N}}$ als komplexe Zahlenfolge interpretiert. Dann gilt (Übung 77)*

$$\mathsf{R}[\mathsf{P}] = \mathsf{R}[\tilde{\mathsf{P}}] \quad \textit{also} \quad \mathrm{K}[\mathsf{P}] = \mathbb{R} \cap \mathrm{K}[\tilde{\mathsf{P}}] \qquad \textit{sowie} \qquad \Theta[\mathsf{P}] = \mathbb{R} \cap \Theta[\tilde{\mathsf{P}}]. \tag{186}$$

Übung 77. *Zeigen Sie die Identitäten in* (186).

Hinweis: Für die linke Seite ist Satz 29.c) nützlich. Für die rechte Seite ist Übung 63 hilfreich. (Es folgt zwanglos, dass $\Theta[\mathsf{P}] \subseteq \mathbb{R} \cap \Theta[\tilde{\mathsf{P}}]$ *gilt. Für die umgekehrte Inklusion, müssen Sie* $\tilde{\mathsf{P}}(z) \in \mathbb{R}$ *für alle* $z \in \mathbb{R} \cap \Theta[\tilde{\mathsf{P}}]$ *nachweisen.)*

Bemerkung 48. *Für* $z \in \mathbb{K}$ *mit* $|z| = \mathsf{R}[\mathsf{P}] > 0$*, liefert Satz 29 (bzw. das Wurzelkriterium) keine Konvergenzaussage. Daher müssen diese Fälle jeweils gesondert betrachtet werden:*[30]

a) Sei $\mathsf{P}(z) = \sum_{k=0}^{\infty} z^k$ *die geometrische Reihe (also* $c_n = 1 \in \mathbb{K}$ *für alle* $n \in \mathbb{N}$*). Die Formel von Hadamard liefert* $\mathsf{R}[\mathsf{P}] = 1$*, und Korollar 22 zeigt dann:* $\Theta[\mathsf{P}] = \{z \in \mathbb{K} \mid |z| < 1\} = \mathrm{K}[\mathsf{P}]$.

(Es sei angemerkt, dass die Abbildung $z \mapsto \frac{1}{1-z}$ *natürlich sogar auf* $\mathbb{C} \setminus \{1\}$ *definiert ist, obwohl die geometrische Reihe für alle* $|z| > 1$ *nicht konvergiert.)*

b) Sei $\mathsf{P}(z) = \sum_{n=0}^{\infty} \frac{z^n}{n+1}$*, also* $c_n = \frac{1}{n+1}$ *für alle* $n \in \mathbb{N}$*. Dann gilt*

$$\overline{\lim}_n \sqrt[n]{|c_n|} = \lim_n \sqrt[n]{|c_n|} = \lim_n \frac{1}{\sqrt[n]{n+1}} = 1 \quad \overset{\text{Satz 29.c)}}{\Longrightarrow} \quad \mathsf{R}[\mathsf{P}] = \frac{1}{\overline{\lim}_n \sqrt[n]{|c_n|}} = 1.$$

(Die linke Seite gilt wegen Lemma 31 (Quetschlemma); denn Lemma 38 zeigt $\lim_n \sqrt[n]{n} = 1 = \lim_n \sqrt[n]{2}$*, und es gilt*

$$1 \le \sqrt[n]{n} \le \sqrt[n]{n+1} \le \sqrt[n]{2n} = \sqrt[n]{2} \cdot \sqrt[n]{n}. \;)$$

In diesem Fall gilt $\mathrm{K}[\mathsf{P}] \subset \Theta[\mathsf{P}] \subset \overline{\mathrm{K}}[\mathsf{P}]$*, wegen* $\mathrm{K}[\mathsf{P}] \not\ni -1 \in \Theta[\mathsf{P}]$ *und* $\Theta[\mathsf{P}] \not\ni 1 \in \overline{\mathrm{K}}[\mathsf{P}]$*:*

- *Für* $z = -1$*, konvergiert* $\mathsf{P}(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}$ *nach dem Leibnizkriterium (Satz 19).*
- *Für* $z = 1$*, ist* $\mathsf{P}(z) = \sum_{n=1}^{\infty} \frac{1}{n}$ *die divergente harmonische Reihe (Beispiel 46).*

c) Sei $\mathsf{P}(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^2}$*, also* $c_n = \frac{1}{n^2}$ *für alle* $n \in \mathbb{N}$*. Dann gilt*

$$\mathsf{R}[\mathsf{P}] = 1 \qquad \textit{sowie} \qquad \{z \in \mathbb{K} \mid |z| < 1\} = \mathrm{K}[\mathsf{P}] \subset \Theta[\mathsf{P}] = \overline{\mathrm{K}}[\mathsf{P}] = \{z \in \mathbb{K} \mid |z| \le 1\}.$$

In der Tat, gilt

$$\overline{\lim}_n \sqrt[n]{|c_n|} = \overline{\lim}_n \sqrt[n]{n^{-2}} = \lim_n \big(\sqrt[n]{n}\big)^{-2} = 1 \quad \overset{\text{Satz 29.c)}}{\Longrightarrow} \quad \mathsf{R}[\mathsf{P}] = \frac{1}{\overline{\lim}_n \sqrt[n]{|c_n|}} = 1.$$

Für alle $z \in \mathbb{K}$ *mit* $|z| \le 1$*, ist zudem* $\sum_{n=1}^{\infty} \frac{1}{n^2}$ *eine konvergente Majorante von* $\mathsf{P}(z)$ *(Proposition 7 bzw. Beispiel 44); also gilt* $\overline{\mathrm{K}}[\mathsf{P}] = \{z \in \mathbb{K} \mid |z| \le 1\} \subseteq \Theta[\mathsf{P}]$ *(wegen Satz 25).*

Die bisher besprochenen Sachverhalte, lassen sich wie folgt verallgemeinern:

Terminologie 18 (Potenzreihe in einem Entwicklungspunkt)**.** *Eine Potenzreihe* Q *in* $\mathbb{K}$ *im Entwicklungspunkt* $p \in \mathbb{K}$*, ist eine Reihe der Gestalt* $\mathsf{Q}(z) = \sum_{n=0}^{\infty} c_n \cdot (z-p)^n$*, mit einer* $\mathbb{K}$*-Folge* $(c_n)_{n \in \mathbb{N}}$ *sowie* $z \in \mathbb{K}$*. Wir definieren den Konvergenzbereich von* Q *durch*

$$\Theta[\mathsf{Q}] := \big\{ w \in \mathbb{K} \;\big|\; \textstyle\sum_{n=0}^{\infty} c_n \cdot (w-p)^n \;\; \textit{konvergiert in } \mathbb{K} \big\}.$$

Sei $\mathsf{P}(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$ *für alle* $z \in \mathbb{K}$*. Offensichtlich gilt dann* $\Theta[\mathsf{Q}] = p + \Theta[\mathsf{P}]$*; und wir setzen:*

- $\mathsf{R}[\mathsf{Q}] := \mathsf{R}[\mathsf{P}]$ *(Konvergenzradius von* Q*)*
- $\mathrm{K}[\mathsf{Q}] := p + \mathrm{K}[\mathsf{P}]$ *(Konvergenzkreis von* Q*)*
- $\overline{\mathrm{K}}[\mathsf{Q}] := p + \overline{\mathrm{K}}[\mathsf{P}]$

Satz 29 zeigt:

a) $\mathsf{Q}(z)$ *divergiert für alle* $z \in \mathbb{K}$ *mit* $|z-p| > \mathsf{R}[\mathsf{Q}]$. *(*$\Theta[\mathsf{Q}] \subseteq \overline{\mathrm{K}}[\mathsf{Q}]$*)*

b) $\mathsf{Q}(z)$ *konvergiert absolut für alle* $z \in \mathrm{K}[\mathsf{Q}]$. *(*$\mathrm{K}[\mathsf{Q}] \subseteq \Theta[\mathsf{Q}]$*)*

c) $\mathsf{R}[\mathsf{Q}] = \big(\overline{\lim}_n \sqrt[n]{|c_n|}\big)^{-1} \in [0, \infty]$.

[30] Die folgenden Beispiele kann man sowohl als Beispiele in $\mathbb{K} = \mathbb{R}$ als auch als Beispiele in $\mathbb{K} = \mathbb{C}$ auffassen.

6.4.2 Die Exponentialreihe

Definition 48. *Die Exponentialreihe ist die komplexe Potenzreihe*

$$\mathrm{e}^z := \exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!} \qquad \forall\, z \in \mathbb{C}. \tag{187}$$

Satz 30. *Die Exponentialreihe konvergiert absolut für alle $z \in \mathbb{C}$ (hat also notwendig Konvergenzradius ∞). Die Exponentialfunktion*

$$\exp\ \mathbb{C} \to \mathbb{C}, \quad z \mapsto \mathrm{e}^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

hat die folgende Eigenschaften:

a) $\mathrm{e}^{z+w} = \mathrm{e}^z \cdot \mathrm{e}^w$ *für alle* $z, w \in \mathbb{C}$, *(Funktionalgleichung)*

b) $\mathrm{e}^{-z} = (\mathrm{e}^z)^{-1}$ *für alle* $z \in \mathbb{C}$, *(*$\exp\colon \mathbb{C} \to \mathbb{C}_\times$*)*

c) $\overline{\mathrm{e}^z} = \mathrm{e}^{\overline{z}}$ *für alle* $z \in \mathbb{C}$,

d) $\mathrm{e}^x > 0$ *für alle* $x \in \mathbb{R}$.

Beweis. Sei $z \in \mathbb{C}$, sowie $c_n := \frac{z^n}{n!}$ für alle $n \in \mathbb{N}$. Dann gilt

$$\overline{\lim}_n \tfrac{|c_{n+1}|}{|c_n|} = \overline{\lim}_n \tfrac{|z|}{n+1} = \lim_n \frac{|z|}{n+1} = 0.$$

Daher konvergiert $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ gemäß dem Quotientenkriterium (Satz 28) absolut.

a) Wegen der absoluten Konvergenz der Reihen $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ und $\sum_{n=0}^{\infty} \frac{w^n}{n!}$, gilt nach Satz 24

$$\mathrm{e}^z \cdot \mathrm{e}^w = \big(\textstyle\sum_{n=0}^{\infty} \frac{z^n}{n!}\big) \cdot \big(\sum_{n=0}^{\infty} \frac{w^n}{n!}\big) = \sum_{n=0}^{\infty} \big(\sum_{k=0}^{n} \frac{1}{(n-k)!\cdot k!} \cdot z^{n-k} \cdot w^k\big)$$

(Cauchy-Produktformel). Nun folgt aus dem binomischen Lehrsatz (zweiter Schritt)

$$\textstyle\sum_{k=0}^{n} \frac{1}{(n-k)!\cdot k!} \cdot z^{n-k} \cdot w^k = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \cdot z^{k-n} \cdot w^k = \frac{1}{n!} \cdot (z+w)^n \qquad \forall\, n \in \mathbb{N}.$$

Daher gilt $\mathrm{e}^z \cdot \mathrm{e}^w = \sum_{n=0}^{\infty} \frac{1}{n!} \cdot (z+w)^n = \mathrm{e}^{z+w}$.

b) Dies folgt wegen a), einfach aus $\mathrm{e}^{-z} \cdot \mathrm{e}^z = \mathrm{e}^{-z+z} = \mathrm{e}^0 = 1$.

c) Für $u \in \mathbb{C}$ und $n \in \mathbb{N}$, sei $S_n[u] := \sum_{k=0}^{n} \frac{u^k}{k!}$, d.h.,

$$\mathrm{e}^u = \lim_n S_n[u] \qquad \forall\, u \in \mathbb{C}.$$

Per Induktion folgt (mit Lemma 26, sowie $\overline{v+w} = \overline{v} + \overline{w}$ für $v, w \in \mathbb{C}$):

$$\overline{S_n[z]} = \textstyle\sum_{k=0}^{n} \overline{\frac{z^k}{k!}} = \sum_{k=0}^{n} \frac{\overline{z}^k}{k!} = S_n[\overline{z}] \qquad \forall\, n \in \mathbb{N}.$$

Wir erhalten (mit $|\overline{u}| = |u|$ für $u \in \mathbb{C}$ im dritten Schritt):

$$\begin{aligned} \big|\overline{\mathrm{e}^z} - \mathrm{e}^{\overline{z}}\big| &= \big|\overline{\mathrm{e}^z} - \overline{S_n[z]} + \overline{S_n[z]} - \mathrm{e}^{\overline{z}}\big| \\ &= \big|\overline{\mathrm{e}^z - S_n[z]} + \overline{S_n[z]} - \mathrm{e}^{\overline{z}}\big| \\ &\le \big|\overline{\mathrm{e}^z - S_n[z]}\big| + \big|\overline{S_n[z]} - \mathrm{e}^{\overline{z}}\big| \\ &= \big|\mathrm{e}^z - S_n[z]\big| + \big|S_n[\overline{z}] - \mathrm{e}^{\overline{z}}\big|. \end{aligned}$$

Wegen $\lim_n S_n[z] = \mathrm{e}^z$ und $\lim_n S_n[\overline{z}] = \mathrm{e}^{\overline{z}}$, folgt $\big|\overline{\mathrm{e}^z} - \mathrm{e}^{\overline{z}}\big| < \varepsilon$ für alle $\varepsilon > 0$, also $\big|\overline{\mathrm{e}^z} - \mathrm{e}^{\overline{z}}\big| = 0$ wegen Lemma 19. Dies zeigt die Behauptung.

d) Wegen b) gilt $\operatorname{im}(\exp) \subseteq \mathbb{C}_\times$; also $e^x \neq 0$. Offensichtlich gilt weiterhin $\exp(\mathbb{R}) \subseteq \mathbb{R}$; also zeigt a), dass $\mathrm{e}^x = \mathrm{e}^{\frac{x}{2}} \cdot \mathrm{e}^{\frac{x}{2}} = (\mathrm{e}^{\frac{x}{2}})^2 > 0$ gilt (vgl. Lemma 17.8)). □

Bemerkung 49 (Eulersche Zahl)**.** *Die Zahl* $\mathrm{e} := \mathrm{e}^1 = \sum_{n=0}^{\infty} \frac{1}{n!}$ *heißt* Eulersche Zahl*. Wegen*

$$\mathrm{e} = 1 + 1 + \tfrac{1}{2} + \tfrac{1}{6} + \dots \qquad \textit{gilt} \qquad \mathrm{e} > 2{,}6.$$

Etwas genauer ist $\mathrm{e} = 2{,}7182871826198\dots$.

Eine andere Darstellung der Zahl e, liefert der folgenden Satz:

Satz 31. *Für alle* $z \in \mathbb{C}$ *gilt* $\mathrm{e}^z = \lim_n (1 + \frac{z}{n})^n$*, also insbesondere* $\mathrm{e} = \lim_n (1 + \frac{1}{n})^n$.

∗*Beweis.* Sie $z \in \mathbb{C}$, sowie $x_n := (1 + \frac{z}{n})^n$ und $y_n := \sum_{k=0}^{n} \frac{z^k}{k!}$ für alle $n \in \mathbb{N}$. Sei weiterhin $\varepsilon > 0$ vorgegeben. Wir zeigen, dass ein $N_\varepsilon \in \mathbb{N}$ existiert, sodass $|x_n - y_n| < \varepsilon$ für alle $n \geq N_\varepsilon$ gilt. Hiermit folgt $\lim_n (x_n - y_n) = 0$; und dann $\lim_n x_n = \lim_n y_n = \mathrm{e}^z$.

- Wegen Satz 30 (absolute Konvergenz) und Satz 20 (Cauchy-Kriterium), existiert ein $N \in \mathbb{N}$ mit

$$\sum_{k=N+1}^{n} \frac{|z|^k}{k!} = \Big| \sum_{k=0}^{n} \Big|\frac{z^k}{k!}\Big| - \sum_{k=0}^{N} \Big|\frac{z^k}{k!}\Big| \Big| < \frac{\varepsilon}{2} \qquad \forall\, n \geq N.$$

- Aus Satz 6 (Binomischer Lehrsatz) folgt

$$\begin{aligned} x_n = \left(1 + \tfrac{z}{n}\right)^n &= \sum_{k=0}^{n} \binom{n}{k} \cdot \frac{z^k}{n^k} \\ &= \sum_{k=0}^{n} \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{k!} \cdot \frac{z^k}{n^k} \\ &= \sum_{k=0}^{n} \frac{z^k}{k!} \cdot \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{n^k} \\ &= \sum_{k=0}^{n} \frac{z^k}{k!} \underbrace{\prod_{\ell=0}^{k-1} \left(1 - \tfrac{\ell}{n}\right)}_{\in [0,1]}, \end{aligned}$$

  und daher gilt

$$y_n - x_n = \sum_{k=0}^{n} \frac{z^k}{k!} \overbrace{\Big[1 - \prod_{\ell=0}^{k-1} \left(1 - \tfrac{\ell}{n}\right)\Big]}^{=:\, d_{k,n} \in [0,1]}.$$

  Für jedes $0 \leq k \leq N$, gilt dann $\lim_n d_{k,n} = 0$. Daher existiert ein $N_\varepsilon \geq N$, sodass $d_{k,n} < \frac{\varepsilon}{2(N+1)}$ für alle $0 \leq k \leq N$ gilt, sofern $n \geq N_\varepsilon$.

Für $n \geq N_\varepsilon$, erhalten wir (mit $d_{k,n} \in [0,1]$ im zweiten Schritt)

$$\begin{aligned} |y_n - x_n| &= \sum_{k=N+1}^{n} \frac{|z|^k}{k!} \cdot d_{k,n} + \sum_{k=0}^{N} \frac{|z|^k}{k!} \cdot d_{k,n} \\ &\leq \sum_{k=N+1}^{n} \frac{|z|^k}{k!} \qquad + \sum_{k=0}^{N} \frac{|z|^k}{k!} \cdot d_{k,n} \\ &< \frac{\varepsilon}{2} + (N+1) \cdot \frac{\varepsilon}{2(N+1)} \\ &= \varepsilon, \end{aligned}$$

was die Behauptung zeigt. □

Bemerkung 50. *Eine Möglichkeit, den Grenzwert aus Satz 31 zu interpretieren, stellt die kontinuierliche Verzinsung dar. Wir stellen uns z als jährlichen Zinssatz vor, und bezeichnen mit k_a das Anfangskapital:*

- *Bei jährlicher Verzinsung ergibt sich nach einem Jahr das neue Kaptial zu* $k_1 = k_a \cdot (1 + z)$.
- *Bei halbjährlicher Verzinsung ergibt sich nach einem Jahr das neue Kaptial zu* $k_2 = k_a \cdot (1 + \frac{z}{2})^2$.
- *Bei monatliche Verzinsung ergibt sich nach einem Jahr das neue Kaptial zu* $k_{12} = k_a \cdot (1 + \frac{z}{12})^{12}$.
- *Bei* $\frac{1}{n}$*-jähriger Verzinsung ergibt sich nach einem Jahr das neue Kaptial zu* $k_n = k_a \cdot (1 + \frac{z}{n})^n$.
- *Im Grenzwert* $n \to \infty$ *(kontinuierliche Verzinsung), ergibt sich nach einem Jahr das neue Kaptial zu* $k_\infty = k_a \cdot \mathrm{e}^z$.

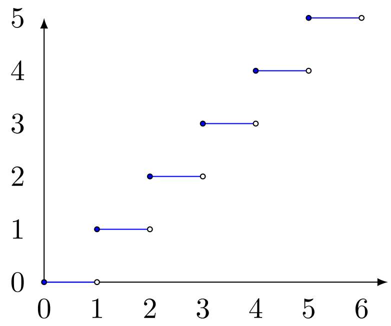
7 Stetigkeit

In diesem Kapitel behandeln den Begriff der stetigen Abbildungen zwischen metrischen Räumen, im spezielleren wieder $\mathbb{R}$ und $\mathbb{C}$. Im Folgenden sei zunächst $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$.

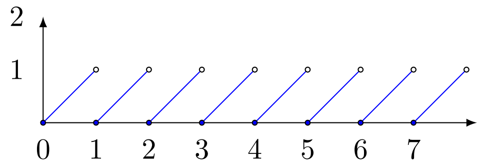
7.1 Stetige Abbildungen

Stetige Funktionen $\mathbb{R} \to \mathbb{R}$ haben Sie bereits in der Schule kennengelernt. Anschaulich gesprochen, ist eine Funktion stetig, wenn ihr Graph in einem Zuge zeichenbar ist (also ohne absetzen des Stiftes). Vorab einige Beispiele:

- Für $n \in \mathbb{N}$, ist $\mathbb{R} \ni x \mapsto x^n \in \mathbb{R}$ stetig.
- Die stückweise konstante Funktion $\mathbb{R} \ni x \mapsto [x] := \max(\{k \in \mathbb{Z} \mid k \leq x\}) \in \mathbb{Z} \subseteq \mathbb{R}$, wird als *Gaußklammer* bezeichnet. Sie ist in allen Punkten $x \in \mathbb{Z}$ unstetig.



- Die *Sägezahnfunktion* $\mathbb{R} \ni x \mapsto x - [x] \in \mathbb{R}$, ist ebenfalls in allen Punkten $x \in \mathbb{Z}$ unstetig.



- Die Funktion

$$\mathbb{R} \to \mathbb{R}, \qquad x \mapsto \begin{cases} 1 & \text{falls } |x| \geq 1, \\ \frac{1}{n} & \text{falls } \frac{1}{n} \leq |x| < \frac{1}{n-1} \text{ für } \mathbb{N} \ni n \neq 2, \\ 0 & \text{falls } x \leq 0, \end{cases}$$

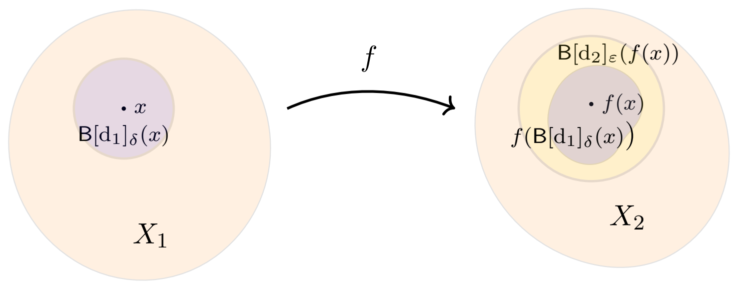
  ist in allen Zahlen der Form $\frac{1}{n}$ mit $\mathbb{Z} \ni n \neq 0$ unstetig, aber stetig in 0.

Die Idee hinter der Stetigkeit ist der, dass sich bei genügend kleiner Variation des Argumentes um einen vorgegebenen Punkt, der Funktionswert auch nur wenig ändert. Die exakte Definition lautet wie folgt:

Definition 49. *Seien (X_1, d_1), (X_2, d_2) metrische Räume, sowie $f\colon X_1 \to X_2$ eine Abbildung.*

1) f heißt stetig in dem Punkt $x \in X_1$ genau dann, wenn gilt:

$$\forall\, \varepsilon > 0\colon\ \exists\, \delta > 0\colon\ f(\mathsf{B}[\mathrm{d}_1]_\delta(x)) \subseteq \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x)). \tag{188}$$



2) f heißt stetig genau dann, wenn f stetig in jedem Punkt ist; wenn also gilt:

$$\forall\,(\,x \in X_1 \;\wedge\; \varepsilon > 0\,)\colon\; \exists\,\delta > 0\colon\; f(\mathsf{B}[\mathrm{d}_1]_\delta(x)) \subseteq \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x)). \tag{189}$$

Notation 24 (Raum der Stetigen Abbildungen)**.** *Gegeben metrische Räume (X_1, d_1), (X_2, d_2), so wird der Raum aller stetigen Abbildungen $X_1 \to X_2$ mit $C(X_1, X_2)$ notiert.*

Terminologie 19. *Die Formulierung des Stetigkeitsbegriffes in Definition 49, wird auch als das ε-δ-Kriterium für Stetigkeit bezeichnet. Der Hintergrund ist, dass auch alternative bzw. gleichwertige Formulierungen des Stetigkeitsbegriffes existieren, welche wir in Abschnitt 7.1.1 noch behandeln werden.*

Terminologie 20 (Unstetigkeit)**.** *Seien (X_1, d_1), (X_2, d_2) metrische Räume, sowie $f\colon X_1 \to X_2$ eine Abbildung.*

- *f heißt unstetig in $x \in X_1$, wenn f nicht stetig in x ist; wenn also ein $\varepsilon > 0$ existiert mit*

$$\forall\,\delta > 0\colon\; f(\mathsf{B}[\mathrm{d}_1]_\delta(x)) \not\subseteq \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x)).$$
$$\big(\;\forall\,\delta > 0\colon\; \exists\, y \in \mathsf{B}[\mathrm{d}_1]_\delta(x)\colon\; f(y) \notin \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x))\;\big)$$

  Sei beispielsweise $f\colon \mathbb{R} \to \mathbb{R}$ gegeben durch

$$f(x) = \begin{cases} 1 & \textit{für } x = 0 \\ 0 & \textit{für } x \neq 0 \end{cases}$$

  für alle $x \in \mathbb{R}$. Dann ist f unstetig in 0; denn ist $0 < \varepsilon < 1$ vorgegeben, so gilt für alle $\delta > 0$:

$$f(\mathsf{B}_\delta(0)) = f((-\delta, \delta)) \ni f(\delta/2) = 0 \notin (1 - \varepsilon, 1 + \varepsilon) = \mathsf{B}_\varepsilon(f(0)).$$

- *f heißt unstetig, wenn f nicht stetig ist; wenn also ein $x \in X_1$ existiert, sodass f nicht stetig in x ist.*

Bemerkung 51 (ε-δ-Kriterium)**.** *Die Bedingung* (188) *in Definition 49.1) (Stetigkeit in $x \in X_1$), lässt sich offensichtlich auch wie folgt formulieren (analog für Bedingung* (189) *in Definition 49.2)):*

Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$, sodass gilt:

$$\mathrm{d}_1(x, y) < \delta \quad \textit{für} \quad y \in X_1 \qquad \Longrightarrow \qquad \mathrm{d}_2(f(x), f(y)) < \varepsilon. \tag{190}$$

- *Sei $(V_1, \|\cdot\|_1)$ ein normierter Raum; sowie $X_1 \subseteq V_1$, und d_1 die auf X_1 eingeschränkte Metrik $\mathrm{d}_{\|\cdot\|_1}$). Dann liest sich die linke Seite von* (190) *auch in der Form $\|x - y\|_1 < \delta$.*
- *Sei $(V_2, \|\cdot\|_2)$ ein normierter Raum; sowie $X_2 \subseteq V_2$, und d_2 die auf X_2 eingeschränkte Metrik $\mathrm{d}_{\|\cdot\|_2}$. Dann liest sich die rechte Seite von* (190) *auch in der Form $\|f(x) - f(y)\|_2 < \varepsilon$.*

Ist beispielsweise $V_1 = \mathbb{K} = V_2$, *so liest sich* (190) *in der Form*

$$|x-y| < \delta \quad \textit{für} \quad y \in X_1 \qquad \Longrightarrow \qquad |f(x) - f(y)| < \varepsilon.$$

Terminologie 21 (Lipschitz-Stetigkeit). *Seien* (X_1, d_1) *und* (X_2, d_2) *metrische Räume. Eine Abbildung* $f\colon X_1 \to X_2$ *heißt Lipschitz-stetig, wenn ein* $L > 0$ *(Lipschitz-Konstante) existiert mit*

$$\mathrm{d}_2(f(x), f(y)) \leq L \cdot \mathrm{d}_1(x,y) \qquad \forall\, x, y \in X_1. \tag{191}$$

In diesem Fall ist f *stetig; denn für* $x \in X_1$ *und* $\varepsilon > 0$ *vorgegeben, erhalten wir für* $\delta := \frac{\varepsilon}{L}$*:*

$$\mathrm{d}_1(x,y) < \delta \quad \textit{für} \quad y \in X_1 \qquad \Longrightarrow \qquad \mathrm{d}_2(f(x), f(y)) \leq L \cdot \mathrm{d}_1(x,y) < L \cdot \delta = \varepsilon.$$

Beispiel 50.

a) Seien (X_1, d_1) *und* (X_2, d_2) *metrische Räume, und* $z \in X_2$. *Dann ist die konstante Abbildung*

$$f[z]\colon X_1 \ni x \mapsto z \in X_2$$

Lipschitz-stetig (also stetig). In der Tat ist ja (191) *sogar für alle* $L > 0$ *erfüllt, da die linke Seite immer gleich* 0 *ist.*

b) Sei (X, d) *ein metrischer Raum. Die Identität* $\mathrm{id}_X\colon X \ni x \mapsto x \in X$ *ist Lipschitz-stetig (also stetig), da* (191) *für* $L = 1$ *gilt (auch für alle* $L \geq 1$*).*

c) Sei (X, d) *ein metrischer Raum, und* $z \in X$ *fixiert. Dann ist die Abstandsfunktion*

$$\mathrm{d}_z\colon X \to [0, \infty), \qquad x \mapsto \mathrm{d}(z,x)$$

Lipschitz-stetig (also stetig); denn aus der inversen Dreiecksungleichung M4) *folgt*

$$\mathrm{d}_{|\cdot|}(\mathrm{d}_z(x), \mathrm{d}_z(y)) = |\mathrm{d}_z(x) - \mathrm{d}_z(y)| = |\mathrm{d}(z,x) - \mathrm{d}(z,y)| \leq \mathrm{d}(x,y),$$

womit (191) *für* $L = 1$ *gilt.*

d) Sei $(V, \|\cdot\|)$ *ein normierter Raum. Mit der Notation aus Teil c), gilt dann für* $\mathrm{d}_{\|\cdot\|}$ *und* $z = 0$*:*

$$\mathrm{d}_z(x) = \mathrm{d}_0(x) = \mathrm{d}_{\|\cdot\|}(x, 0) = \|x - 0\| = \|x\| \qquad \forall\, x \in X = V.$$

Gemäß Teil c), ist daher

$$\|\cdot\|\colon V \to [0, \infty), \qquad x \mapsto \|x\|$$

Lipschitz-stetig. Wählt man nun $V = \mathbb{K}$ *mit* $\|\cdot\| = |\cdot|$, *so folgt hiermit auch die Stetigkeit der Betragsfunktion.*

Beispiel 51. *Sei* $X_1 = \mathbb{K} = X_2$ *(mit der Betragsmetrik). Für jedes* $w \in \mathbb{K}$ *ist*

$$f_w\colon \mathbb{K} \to \mathbb{K}, \qquad z \mapsto w \cdot z$$

Lipschitz-stetig (also stetig), wegen $(L = |w|)$

$$\mathrm{d}_{|\cdot|}(f_w(x), f_w(y)) = |f_w(x) - f_w(y)| = |w \cdot (x-y)| = |w| \cdot |x-y| = |w| \cdot \mathrm{d}_{|\cdot|}(x,y) \qquad \forall\, x, y \in \mathbb{K}.$$

Übung 78. *Zeigen Sie, dass die komplexe Konjugationsabbildung* $\mathbb{C} \ni z \mapsto \overline{z} \in \mathbb{C}$ *eine Lipschitz-stetige Abbildung ist.*

7.1.1 Alternative Formulierungen von Stetigkeit

In diesem Abschnitt behandeln wir zwei alternative Formulierungen von Stetigkeit; welche in der ein oder anderen Situation sehr nützlich sind, um Stetigkeiten von gegebenen Abbildungen nachzuweisen.

Definition 50 (Folgenstetigkeit)**.** *Seien (X_1, d_1), (X_2, d_2) metrische Räume, und $f\colon X_1 \to X_2$ eine Abbildung.*

1) f heißt folgenstetig in dem Punkt $x \in X_1$, wenn für jede Folge $(x_n)_{n\in\mathsf{N}}$ in X_1 gilt:

$$(x_n)_{n\in\mathsf{N}} \to x \qquad \Longrightarrow \qquad (f(x_n))_{n\in\mathsf{N}} \to f(x).$$

2) f heißt folgenstetig genau dann, wenn f folgenstetig in jedem Punkt ist.

Proposition 8. *Seien (X_1, d_1), (X_2, d_2) metrische Räume, sowie $f\colon X_1 \to X_2$ eine Abbildung. Dann gelten die folgenden Äquivalenzen:*

$$\begin{aligned} &f \ \text{ stetig in } x \in X_1 && \Longleftrightarrow && f \ \text{ folgenstetig in } x \in X_1 \\ &f \ \text{ stetig} && \Longleftrightarrow && f \ \text{ folgenstetig.} \end{aligned}$$

Beweis. Die zweite Äquivalenz folgt sofort aus der ersten Äquivalenz. Wir zeigen nun die erste Äquivalenz:

- Sei f stetig in $x \in X_1$; sowie $(x_n)_{n\in\mathsf{N}}$ eine Folge in X_1, mit $(x_n)_{n\in\mathsf{N}} \to x$. Für $\varepsilon > 0$ vorgegeben, sei $\delta > 0$ mit $f(\mathsf{B}[\mathrm{d}_1]_\delta(x)) \subseteq \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x))$. Wir wählen $N_\delta \in \mathsf{N}$ mit $x_n \in \mathsf{B}[\mathrm{d}_1]_\delta(x)$ für alle $n \geq N_\delta$, und erhalten

  $$f(x_n) \in f(\mathsf{B}[\mathrm{d}_1]_\delta(x)) \subseteq \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x)) \qquad \forall\, n \geq N_\delta.$$

  Dies zeigt $(f(x_n))_{n\in\mathsf{N}} \to f(x)$.
- Wir beweisen die Implikation „$\Leftarrow$“ durch Kontraposition:

  Ist f nicht stetig in x, so existiert ein $\varepsilon > 0$ mit

  $$f(\mathsf{B}[\mathrm{d}_1]_\delta(x)) \not\subseteq \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x)) \qquad \forall\, \delta > 0.$$

  Für jedes $n \in \mathbb{N}_{>0}$, existiert daher ein $x_n \in \mathsf{B}[\mathrm{d}_1]_{\frac{1}{n}}(x)$ mit $f(x_n) \notin \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x))$.

  - Es gilt $(x_n)_{n\in\mathbb{N}_{>0}} \to x$; denn für $\tilde{\varepsilon} > 0$ vorgegeben, existiert ein $N \in \mathbb{N}_{>0}$ mit $0 < \frac{1}{N} < \tilde{\varepsilon}$, d.h.,

    $$x_n \in \mathsf{B}[\mathrm{d}_1]_{\frac{1}{n}}(x) \subseteq \mathsf{B}[\mathrm{d}_1]_{\frac{1}{N}}(x) \subseteq \mathsf{B}[\mathrm{d}_1]_{\tilde{\varepsilon}}(x) \qquad \forall\, n \geq N.$$

  - Es konvergiert $(f(x_n))_{n\in\mathbb{N}_{>0}}$ nicht gegen $f(x)$, wegen $f(x_n) \notin \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x))$ für alle $n \in \mathbb{N}_{>0}$.

  Dies zeigt, dass f nicht folgenstetig in x ist. □

Wir erinnern an die Begrifflichkeiten aus Definition 34.

Proposition 9. *Seien (X_1, d_1), (X_2, d_2) metrische Räume, und $f\colon X_1 \to X_2$ eine Abbildung.*

1) f ist genau dann stetig in $x \in X_1$, wenn gilt:

$$U_2 \subseteq X_2 \ \text{ ist Umgebung von } f(x) \qquad \Longrightarrow \qquad f^{-1}(U_2) \subseteq X_1 \ \text{ ist Umgebung von } x. \tag{192}$$

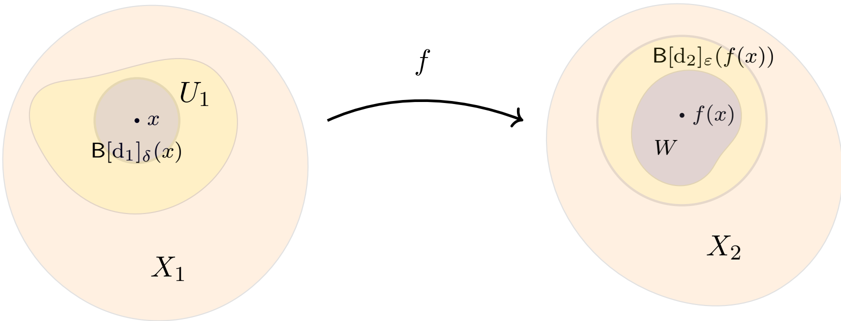
2) f ist genau dann stetig, wenn gilt: *(Urbilder offener Mengen sind offen)*

$$f^{-1}(O) \in \mathcal{O}(X_1) \qquad \forall\, O \in \mathcal{O}(X_2). \tag{193}$$

Beweis. 1) • Es gelte (192). Für $\varepsilon > 0$, ist dann $\mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x))$ eine Umgebung von $f(x)$. Per Voraussetzung, ist dann $U_1 := f^{-1}(\mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x)))$ eine Umgebung von x. Per definitionem, existiert daher ein $\delta > 0$ mit $\mathsf{B}[\mathrm{d}_1]_\delta(x) \subseteq U_1$; und es folgt

$$\underbrace{f(\mathsf{B}[\mathrm{d}_1]_\delta(x))}_{=:W} \subseteq f(U_1) = \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x)).$$

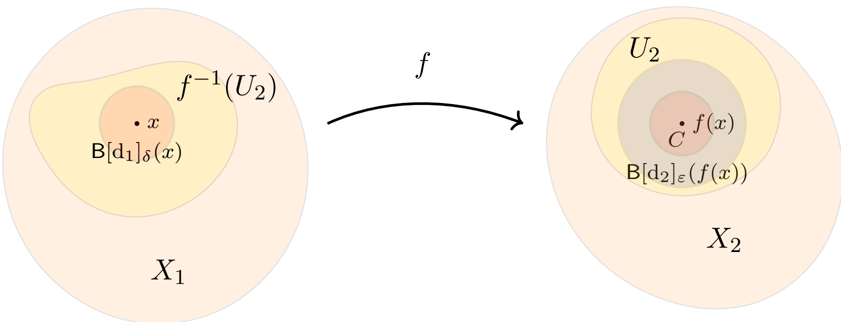
Dies zeigt, dass f stetig in x ist.



• Es sei f stetig in $x \in X_1$, und es sei $U_2 \subseteq X_2$ eine Umgebung von $f(x)$. Per definitionem, existiert ein $\varepsilon > 0$ mit $\mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x)) \subseteq U_2$. Per Annahme, existiert ein $\delta > 0$ mit

$$\underbrace{f(\mathsf{B}[\mathrm{d}_1]_\delta(x))}_{C} \subseteq \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x)) \quad \overset{f^{-1}(\cdot)}{\Longrightarrow} \quad \mathsf{B}[\mathrm{d}_1]_\delta(x) \subseteq f^{-1}(C) \subseteq f^{-1}(\mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x))) \subseteq f^{-1}(U_2),$$

womit $f^{-1}(U_2)$ eine Umgebung von x ist.



2) • Es sei f stetig, sowie $O \in \mathcal{O}(X_2)$ vorgegeben. Per definitionem

– ist f stetig in jedem $x \in X_1$.
– ist O eine Umgebung aller ihrer Punkte.

Gegeben $x \in f^{-1}(O) \subseteq X_1$, so ist daher O eine Umgebung von $f(x) \in O$. Wegen Teil 1) ist somit $f^{-1}(O)$ eine Umgebung von x. Da $x \in f^{-1}(O)$ beliebig war, ist die Teilmenge $f^{-1}(O)$ eine Umgebung aller ihrer Punkt, also per Definition offen.

• Es gelte (193); und es sei $x \in X_1$, sowie $\varepsilon > 0$ vorgegeben.

– Wegen Lemma 27, ist $\mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x)) \in \mathcal{O}(X_2)$ offen.
– Per Annahme, ist dann $\tilde{O} := f^{-1}(\mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x))) \in \mathcal{O}(X_1)$ offen mit $x \in \tilde{O}$.
– Per definitionem, existiert somit ein $\delta > 0$ mit $\mathsf{B}[\mathrm{d}_1]_\delta(x) \subseteq \tilde{O}$.

– Wir erhalten $f(\mathsf{B}[\mathrm{d}_1]_\delta(x)) \subseteq f(\tilde{O}) = \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x))$.

Da $x \in X_1$ und $\varepsilon > 0$ beliebig waren, folgt hiermit die Stetigkeit von f in jedem $x \in X_1$. $\square$

Übung 79. *Seien (X, d) und (X', d') metrische Räume, sowie $f\colon X \to X'$ eine Abbildung. Zeigen Sie die folgenden Aussagen:*

a) Sei U eine Umgebung von $x \in X$, sodass $f|_U$ bezüglich der eingeschränkten Metrik $\mathrm{d}|_U$ stetig in x ist. Dann ist auch f stetig in x.

b) Sei $O \in \mathcal{O}(X)$ offen, und $f|_O$ stetig bezüglich der eingeschränkten Metrik d_O. Dann ist f stetig in jedem $x \in O$.

c) Sei $X = \bigcup_{j\in J} O_j$ mit $O_j \in \mathcal{O}(X)$ offen für alle $j \in J$. Ist $f|_{O_j}$ für jedes $j \in J$ stetig bezüglich der eingeschränkten Metrik $\mathrm{d}|_{O_j}$, so ist f stetig.

Übung 80. *Seien (X, d) und (X', d') metrische Räume, sowie $f\colon X \to X'$ eine Abbildung. Zeigen Sie die folgenden Aussagen:*

a) f ist genau dann stetig, wenn $f^{-1}(A') \in \mathcal{A}(X)$ für alle $A' \in \mathcal{A}(X')$ gilt.

b) **(∗)** *Sei $X = \bigcup_{j=0}^{n} A_j$ mit $A_j \in \mathcal{A}(X)$ für alle $j = 0, \ldots, n \in \mathbb{N}$. Ist $f|_{A_j}$ für jedes $j \in J$ stetig bezüglich der eingeschränkten Metrik $\mathrm{d}|_{A_j}$, so ist f stetig.*

*Hinweis: Bemerkung 36.b) und Korollar 13.*AM3).

Übung 81. *Seien (X, d) und (X', d') metrische Räume, $x \in X$, und $f\colon X \to X'$ eine Abbildung. Seien $Y_1, \ldots, Y_n \subseteq X$ Teilmengen, mit $X = \bigcup_{i=1}^{n} Y_i$ und $x \in Y_1 \cap \ldots \cap Y_n$, sodass $f|_{Y_i}$ für jedes $1 \leq i \leq n$ stetig (bezüglich d_{Y_i}) ist. Zeigen Sie, dass f stetig in x ist.*

Hinweis: Gegeben $\varepsilon > 0$, so existiert zu jedem $1 \leq i \leq n$ ein $\delta_i > 0$ mit $f(\mathsf{B}[\mathrm{d}_{Y_i}]_{\delta_i}(x)) \subseteq \mathsf{B}[\mathrm{d}']_\varepsilon(f(x))$. Betrachten Sie den Ball $\mathsf{B}[\mathrm{d}]_\delta(x)$ für $\delta := \min(\delta_1, \ldots, \delta_n)$, und benutzen Sie (139) *sowie Übung 29.*

7.1.2 Stetigkeitskriterien

In diesem Abschnitt behandeln wir einige Stetigkeitskriterien.

Lemma 39 (Verkettung stetiger Abbildungen)**.** *Seien (X, d), (X', d'), (X'', d'') metrische Räume, sowie $f\colon X' \to X''$, $g\colon X \to X'$ Abbildungen.*

1) Sei g stetig in $x \in X$, sowie f stetig in $g(x) \in X'$. Dann ist $f \circ g$ stetig in x.

2) Sind f und g stetig, so ist auch $f \circ g$ stetig.

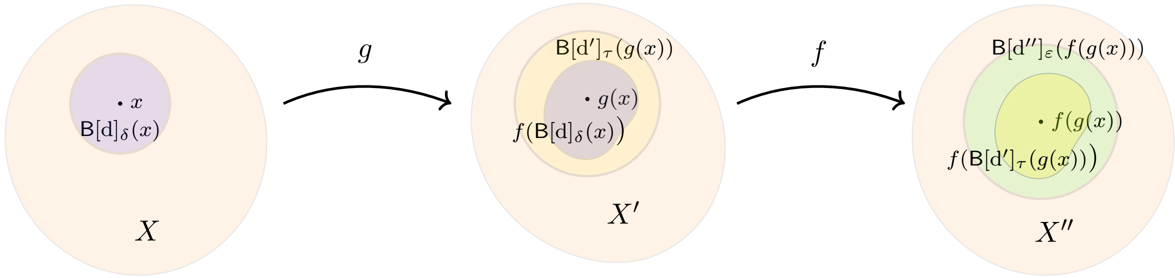
Beweis. 1) Sei $\varepsilon > 0$ vorgegeben.

- Da f stetig in $g(x)$ ist, existiert ein $\tau > 0$ mit $f(\mathsf{B}[\mathrm{d}']_\tau(g(x))) \subseteq \mathsf{B}[\mathrm{d}'']_\varepsilon(f(g(x)))$.
- Da g stetig in x ist, existiert ein $\delta > 0$ mit $g(\mathsf{B}[\mathrm{d}]_\delta(x)) \subseteq \mathsf{B}[\mathrm{d}']_\tau(g(x))$.

Wir erhalten

$$(f \circ g)(\mathsf{B}[\mathrm{d}]_\delta(x)) = f(g(\mathsf{B}[\mathrm{d}]_\delta(x))) \subseteq f(\mathsf{B}[\mathrm{d}']_\tau(g(x))) \subseteq \mathsf{B}[\mathrm{d}'']_\varepsilon(f(g(x))).$$

Da $\varepsilon > 0$ beliebig war, folgt hiermit die Stetigkeit von $f \circ g$ in $x \in X$.



2) Klar wegen Teil 1). $\square$

Übung 82. *Beweisen Sie Lemma 39 nochmals, und zwar jeweils mit Hilfe der Stetigkeitskriterien aus Proposition 8 sowie Proposition 9.*

Lemma 40 (Einschränkung und Koeinschränkung)**.** *Seien (X, d), (X', d') metrische Räume, sowie $f\colon X \to X'$ eine Abbildung. Seien weiterhin $Y \subseteq X$ und $Y' \subseteq X'$ Teilmengen, sowie (Y, d_Y) und $(Y', \mathrm{d}'_{Y'})$ die zugehörigen metrischen Unterräume.*

1) Ist f stetig in $y \in Y$, so ist $f|_Y\colon Y \to X'$ stetig in y (bezüglich d_Y).

2) Ist f stetig, so ist $f|_Y\colon Y \to X'$ stetig (bezüglich d_Y).

Es sei nun $\mathrm{im}(f) = f(X) \subseteq Y'$. Dann gilt weiterhin:

3) Ist f stetig in $x \in X$, so ist $f|^{Y'}\colon X \to Y'$ stetig in x (bezüglich $\mathrm{d}'_{Y'}$).

4) Ist f stetig, so ist $f|^{Y'}\colon X \to Y'$ stetig (bezüglich $\mathrm{d}'_{Y'}$).

Zudem gilt für die Koeinschränkung auch die Umkehrung der beiden vorhergehenden Aussagen:

5) Ist $f|^{Y'}\colon X \to Y'$ stetig in x (bezüglich $\mathrm{d}'_{Y'}$), so ist f stetig in $x \in X$ (bezüglich d').

6) Ist $f|^{Y'}\colon X \to Y'$ stetig (bezüglich $\mathrm{d}'_{Y'}$), so ist f stetig (bezüglich d').

Beweis. Es folgt jeweils 2) aus 1), 4) aus 3), sowie 6) aus 5). Weiterhin folgen 1), 3), 5) unmittelbar aus der Charakterisierung (190), sowie der Definition der eingeschränkten Metrik (129):

1) Sei $\varepsilon > 0$ vorgegeben. Wegen (190), existiert ein $\delta > 0$ mit

$$\mathrm{d}(y, z) < \delta \quad \text{für} \quad z \in X \qquad \Longrightarrow \qquad \mathrm{d}'(f(y), f(z)) < \varepsilon.$$

Wir erhalten

$$\mathrm{d}_Y(y, z) < \delta \quad \text{für} \quad z \in Y \qquad \Longrightarrow \qquad \mathrm{d}(y, z) < \delta \qquad \Longrightarrow \qquad \mathrm{d}'(f(y), f(z)) < \varepsilon,$$

also (190) für die metrischen Räume (Y, d_Y) und (X', d').

3) Sei $\varepsilon > 0$ vorgegeben. Wegen (190), existiert ein $\delta > 0$ mit

$$\mathrm{d}(x, y) < \delta \quad \text{für} \quad y \in X \qquad \Longrightarrow \qquad \mathrm{d}'(f(x), f(y)) < \varepsilon.$$

Wegen $f(X) \subseteq Y'$, gilt nun

$$\mathrm{d}'_{Y'}\big(f|^{Y'}(x), f|^{Y'}(y)\big) = \mathrm{d}'_{Y'}(f(x), f(y)) = \mathrm{d}'(f(x), f(y)) \qquad \forall\, x, y \in X; \tag{194}$$

Dies zeigt die Implikation

$$\mathrm{d}(x, y) < \delta \quad \text{für} \quad y \in Y \qquad \Longrightarrow \qquad \mathrm{d}'_{Y'}\big(f|^{Y'}(x), f|^{Y'}(y)\big) < \varepsilon,$$

also (190) für die metrischen Räume (X, d) und $(Y', \mathrm{d}'_{Y'})$.

5) Sei $\varepsilon > 0$ vorgegeben. Wegen (190) (für die metrischen Räume (X, d) und $(Y', \mathrm{d}'_{Y'})$), existiert ein $\delta > 0$ mit

$$\mathrm{d}(x,y) < \delta \quad \text{für} \quad y \in X \qquad \Longrightarrow \qquad \mathrm{d}'_{Y'}\big(f|^{Y'}(x), f|^{Y'}(y)\big) < \varepsilon.$$

Wegen (194) gilt dann ebenfalls

$$\mathrm{d}(x,y) < \delta \quad \text{für} \quad y \in Y \qquad \Longrightarrow \qquad \mathrm{d}'(f(x), f(y)) < \varepsilon,$$

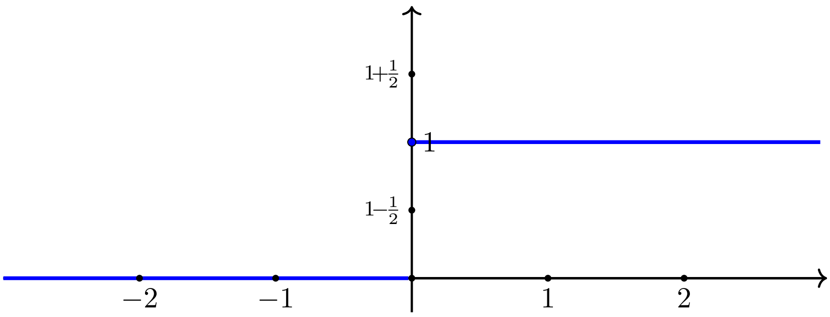
also (190) für die metrischen Räume (X, d) und (X', d').

□

Übung 83. *Machen Sie sich klar, dass im Kontext der ersten beiden Teile von Lemma 40 im Allgemeinen nicht die Umkehrungen gelten – dass also beispielsweise die Stetigkeit von $f|_Y$ in $y \in Y$, nicht notwendigerweise die Stetigkeit von f in y impliziert. Betrachten Sie hierfür zum Beispiel die Abbildung*

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

sowie die Menge $Y = [0, \infty)$ und den Punkt $y = 0$.



7.1.3 Operationen mit stetigen Abbildungen

Notation 25. *Sei X eine Menge, und V ein $\mathbb{K}$-Vektorraum.*

1) *Gegeben $f, g \in \mathsf{Abb}(X, V)$, $h \in \mathsf{Abb}(X, \mathbb{K})$, $\lambda \in \mathbb{K}$, so definieren wir*

$$\begin{aligned} \mathsf{Abb}(X,V) \ni (f+g)&\colon X \to V, \qquad x \mapsto f(x) + g(x), \\ \mathsf{Abb}(X,V) \ni (h \cdot f)&\colon X \to V, \qquad x \mapsto h(x) \cdot f(x), \\ \mathsf{Abb}(X,V) \ni (-f)&\colon X \to V, \qquad x \mapsto -f(x), \\ \mathsf{Abb}(X,V) \ni (\lambda \cdot f)&\colon X \to V, \qquad x \mapsto \lambda \cdot f(x). \end{aligned} \tag{195}$$

Die letzte Operation ist ein Spezialfall der zweiten Operation (konstante Abbildung).

2) *Für $f \in \mathsf{Abb}(X, \mathbb{K}_\times)$, $g \in \mathsf{Abb}(X, \mathbb{K})$ und $h \in \mathsf{Abb}(X, \mathbb{C})$, definieren wir (siehe (116) für die dritte Zeile)*

$$\begin{aligned} \mathsf{Abb}(X,\mathbb{K}_\times) \ni \frac{1}{f}&\colon X \to \mathbb{K}_\times, \qquad x \mapsto \frac{1}{f(x)} \\ \mathsf{Abb}(X,\mathbb{K}) \ni \frac{g}{f}&\colon X \to \mathbb{K}, \qquad x \mapsto \left(g \cdot \frac{1}{f}\right)(x) = \frac{g(x)}{f(x)} \\ \mathsf{Abb}(X,\mathbb{C}) \ni \overline{h}&\colon X \to \mathbb{C}, \qquad x \mapsto \overline{h(x)}. \end{aligned} \tag{196}$$

Es gilt der folgende Satz:

Satz 32. *Sei* (X, d) *ein metrischer Raum, und* $(V, \|\cdot\|)$ *ein normierter* $\mathbb{K}$*-Vektorraum.*

1) Sind $f, g \in \mathsf{Abb}(X, V)$ *sowie* $h \in \mathsf{Abb}(X, \mathbb{K})$ *stetig in* $x \in X$*, so sind die Abbildungen in* (195) *ebenfalls stetig in* x*. Insbesondere ist* $(\lambda \cdot f \pm \mu \cdot g)$ *stetig in* x *für alle* $\lambda, \mu \in \mathbb{K}$*.*

2) Sind $f \in \mathsf{Abb}(X, \mathbb{K}_\times)$*,* $g \in \mathsf{Abb}(X, \mathbb{K})$ *und* $h \in \mathsf{Abb}(X, \mathbb{C})$ *stetig in* $x \in X$*, so sind die Abbildungen in* (196) *ebenfalls stetig in* x*.*

Selbiges gilt natürlich insbesondere für den Spezialfall $V = \mathbb{K}$.

Beweis. 1) Sei $(x_n)_{n \in \mathsf{N}}$ eine Folge in X, mit $\lim_n x_n = x$. Wegen Proposition 8 gilt

$$\lim_n f(x_n) = f(x), \qquad \lim_n g(x_n) = g(x), \qquad \lim_n h(x_n) = h(x).$$

Wegen Proposition 6 gilt dann

$$\begin{aligned}
\lim_n (f+g)(x_n) &= \lim_n (f(x_n) + g(x_n)) = \lim_n f(x_n) + \lim_n g(x_n) = f(x) + g(x) = (f+g)(x) \\
\lim_n (h \cdot g)(x_n) &= \lim_n (h(x_n) \cdot g(x_n)) = \lim_n h(x_n) \cdot \lim_n g(x_n) = h(x) \cdot g(x) = (h \cdot g)(x) \\
\lim_n (-f)(x_n) &= \lim_n -f(x_n) = -f(x) = (-f)(x).
\end{aligned}$$

Daher zeigt Proposition 8 die Stetigkeit besagter Abbildungen im Punkte x. Wählen wir $h = f[\lambda] \colon \mathbb{K} \ni x \mapsto \lambda \in \mathbb{K}$ konstant, so folgt nun auch die Stetigkeit von $\lambda \cdot f$.

2)
- $\overline{h}$ stetig in x, wegen Lemma 39 (Verkettungen) und Übung 78 (Stetigkeit der komplexen Konjugation).
- Sei $(x_n)_{n \in \mathsf{N}}$ eine Folge in X mit $\lim_n x_n = x$. Wegen Proposition 8 gilt $\lim_n f(x_n) = x$; und Lemma 32, zeigt dann

$$\lim_n \left(\frac{1}{f}\right)(x_n) = \lim_n \frac{1}{f(x_n)} = \frac{1}{f(x)} = \left(\frac{1}{f}\right)(x).$$

  Proposition 8 zeigt daher die Stetigkeit von $\frac{1}{f}$ in x.
- Wegen dem vorherigen Punkt ist $1/f$ stetig in x. Wegen Lemma 40.5) ist dies immer noch der Fall, wenn wir $1/f$ als Abbildung $X \to \mathbb{K}$ auffassen. Die Stetigkeit von g/f in x, folgt daher aus Teil 1) mit $V = \mathbb{K}$. □

Korollar 25. *Sei* (X, d) *ein metrischer Raum, und* $(V, \|\cdot\|)$ *ein normierter* $\mathbb{K}$*-Vektorraum.*

1) Gegeben $f, g \in C(X, V)$*,* $h \in C(X, \mathbb{K})$ *und* $\lambda \in \mathbb{K}$*, so gilt*

$$(f+g) \in C(X, V), \qquad (h \cdot f) \in C(X, V), \qquad -f \in C(X, V), \qquad \lambda \cdot f \in C(X, V)$$

Insbesondere gilt $(\lambda \cdot f \pm \mu \cdot g) \in C(X, V)$ *für alle* $\lambda, \mu \in \mathbb{K}$*.*

2) Gegeben $f \in C(X, \mathbb{K}_\times)$*, so gilt*

$$\frac{1}{f} \in C(X, \mathbb{K}_\times) \qquad \textit{und somit} \qquad \frac{g}{f} \in C(X, \mathbb{K}) \qquad \forall\, g \in C(X, \mathbb{K}).$$

Selbiges gilt natürlich insbesondere für den Spezialfall $V = \mathbb{K}$.

Beweis. Klar wegen Satz 32. □

Beispiel 52.

a) *Eine Funktion $P\colon \mathbb{K} \to \mathbb{K}$ heißt Polynomfunktion (Polynom) vom Grad $n \in \mathbb{N}$, wenn $a_0, \ldots, a_n \in \mathbb{K}$ mit $a_n \neq 0$ existieren, sodass gilt:* $\big(P = \sum_{k=0}^{n} a_k \cdot (\mathrm{id}_{\mathbb{K}})^k\big)$

$$P(z) = \textstyle\sum_{k=0}^{n} a_k \cdot z^k = a_0 + a_1 \cdot z + \ldots + a_{n-1} \cdot z^{n-1} + a_n \cdot z^n \qquad \forall\, z \in \mathbb{K}.$$

Wir bezeichnen die Menge aller Polynomfunktionen auf $\mathbb{K}$ mit $\mathsf{Pol}(\mathbb{K})$. Per Induktion folgt aus Korollar 25.1) sowie den ersten beiden Teilen von Beispiel 50, dass jedes $P \in \mathsf{Pol}(\mathbb{K})$ stetig ist.

b) *Gegeben zwei Polynomfunktionen ($z \in \mathbb{K}$)*

$$P(z) = a_0 + a_1 \cdot z + \cdots + a_n \cdot z^n \qquad \textit{und} \qquad Q(z) = b_0 + b_1 \cdot z + \cdots + b_m \cdot z^m \qquad \textit{mit} \qquad b_m \neq 0$$

auf $\mathbb{K}$, so ist die Menge $\mathcal{Z} := Q^{-1}(0)$ endlich (Satz 33), und die Funktion

$$\frac{P}{Q}\colon \mathbb{K} \setminus \mathcal{Z} \to \mathbb{K}, \qquad z \mapsto \frac{P(z)}{Q(z)} = \frac{a_0 + a_1 \cdot z + \cdots + a_n \cdot z^n}{b_0 + b_1 \cdot z + \cdots + b_m \cdot z^m}$$

ist stetig (Teil a), Lemma 40.2), und Korollar 25.2)). Beispielsweise ist die Funktion $f : \mathbb{C}\setminus\{1\} \ni z \mapsto \frac{z+1}{z-1} \in \mathbb{C}$ stetig.

c) *Unter Zuhilfenahme von Korollar 25.2) erhält man, dass jede Funktion der Form*

$$\mathbb{K}_\times \ni z \mapsto \textstyle\sum_{k=0}^{n} a_k \cdot z^{-k} \in \mathbb{K},$$

mit $a_0, \ldots, a_n \in \mathbb{K}$ für $n \in \mathbb{N}$, stetig ist.[31]

Übung 84. *Machen Sie sich die Stetigkeitsaussage in Beispiel 52 klar.*

Satz 33. *Eine Polynomfunktion vom Grad $n \in \mathbb{N}_{>0}$ besitzt höchstens n Nullstellen.*[32]

Beweis. Wir zeigen die Behauptung per Induktion nach dem Grad des Polynoms:

- Ist $n = 1$, so ist die Behauptung klar: $a_0 + a_1 \cdot z = 0 \quad \Leftrightarrow \quad z = -a_0/a_1$.
- Es gelte nun die Behauptung für ein $n \in \mathbb{N}_{>0}$; und es sei $P(z) = \sum_{k=0}^{n+1} a_k \cdot z^k$ eine Polynomfunktion vom Grad $n+1$. Hat P keine Nullstelle, so ist die Behauptung klar. Sei andernfalls z_0 eine Nullstelle von P. Der Binomische Lehrsatz (zweiter Schritt) liefert

$$\begin{aligned} P(z) &= \textstyle\sum_{k=0}^{n+1} a_k \cdot (z_0 + (z - z_0))^k \\ &= \textstyle\sum_{k=0}^{n+1} a_k \cdot \sum_{j=0}^{k} \binom{k}{j} \cdot (z_0)^{k-j} \cdot (z - z_0)^j \\ &= \textstyle\sum_{0 \leq j \leq k \leq n+1} a_k \cdot \binom{k}{j} \cdot (z_0)^{k-j} \cdot (z - z_0)^j \\ &= \textstyle\sum_{j=0}^{n+1} \underbrace{\big(\sum_{k=j}^{n+1} a_k \cdot \binom{k}{j} \cdot (z_0)^{k-j}\big)}_{b_j :=} \cdot (z - z_0)^j \end{aligned}$$

für alle $z \in \mathbb{K}$. Wir erhalten

$$0 = P(z_0) = \textstyle\sum_{j=0}^{n+1} b_j \cdot (z_0 - z_0)^j = b_0 \qquad \Longrightarrow \qquad \begin{aligned} P(z) &= (z - z_0) \cdot \textstyle\sum_{j=1}^{n+1} b_j \cdot (z - z_0)^{j-1} \\ &= (z - z_0) \cdot \underbrace{\textstyle\sum_{j=0}^{n} b_{j+1} \cdot (z - z_0)^j}_{Q(z) :=} \end{aligned}$$

[31] Man kombiniere beispielweise Korollar 25.2) mit Beispiel 50.b) sowie Lemma 39.2) (Verkettungen) mit Teil a).

[32] Für $n = 0$ kann diese Aussage nicht gelten, da die konstante Funktion 0 unendlich viele Nullstellen hat.

für alle $z \in \mathbb{K}$. Die Funktion $Q\colon \mathbb{K} \ni z \to Q(z) \in \mathbb{K}$ ist eine Polynomfunktion vom Grad n; denn der Binomische Lehrsatz (Satz 6) liefert

$$\begin{aligned} Q(z) &= \textstyle\sum_{j=0}^{n} b_{j+1} \cdot \sum_{k=0}^{j} \binom{j}{k} \cdot (-z_0)^{j-k} \cdot z^k \\ &= \textstyle\sum_{0\le k\le j\le n} b_{j+1} \cdot \binom{j}{k} \cdot (-z_0)^{j-k} \cdot z^k \\ &= \textstyle\sum_{k=0}^{n} \underbrace{\left(\sum_{j=k}^{n} b_{j+1} \cdot \binom{j}{k} \cdot (-z_0)^{j-k} \right)}_{c_k} \cdot z^k \end{aligned}$$

für alle $z \in \mathbb{K}$, mit $c_n = b_{n+1} = a_{n+1} \neq 0$. Da Q nach Induktionsvoraussetzung höchstens n Nullstellen besitzt, kann P somit höchstens $n+1$ Nullstellen besitzen; denn aus $P(z) = 0$, folgt $z = z_0$ oder $Q(z) = 0$ wegen Lemma 16.1). $\square$

7.1.4 Gleichmäßige Stetigkeit

Definition 51 (Gleichmäßige Stetigkeit)**.** *Seien* (X_1, d_1) *und* (X_2, d_2) *metrische Räume. Eine Abbildung* $f\colon X_1 \to X_2$ *heißt gleichmäßig stetig, wenn gilt:*

$$\begin{gathered} \forall\, \varepsilon > 0\colon\ \exists\, \delta > 0\colon\ \forall\, x \in X_1\colon\ \ f(\mathsf{B}[\mathrm{d}_1]_\delta(x)) \subseteq \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x)) \\ \Big(\forall\, \varepsilon > 0\colon\ \exists\, \delta > 0\colon \quad \big(\mathrm{d}_1(x,y) < \delta \quad \textit{für} \quad x, y \in X_1 \quad \Longrightarrow \quad \mathrm{d}_2(f(x), f(y)) < \varepsilon \,\big) \Big) \end{gathered} \tag{197}$$

Bemerkung 52 (Gleichmäßige Stetigkeit)**.**

a) Ist $f\colon X_1 \to X_2$ *Lipschitz-stetig (mit Lipschitz-Konstante* $L > 0$*), so ist* f *auch gleichmäßig stetig. Ist nämlich* $\varepsilon > 0$ *vorgegeben, so gilt für* $\delta := \frac{\varepsilon}{L}$*:*

$$\mathrm{d}_1(x,y) < \delta \quad \textit{für} \quad x, y \in X_1 \qquad \Longrightarrow \qquad \mathrm{d}_2(f(x), f(y)) \le L \cdot \mathrm{d}_1(x,y) < L \cdot \delta = \varepsilon.$$

Insbesondere sind also die Abbildung aus Beispiel 50 und Beispiel 51 stetig.

b) Für $X_1, X_2 \subseteq \mathbb{K}$ *(mit der eingeschränkten Betragsmetrik), liest sich* (197) *auch in der Form:*

$$\forall\, \varepsilon > 0\colon\ \exists\, \delta > 0\colon \quad \big(|x-y| < \delta \quad \textit{für} \quad x, y \in X_1 \quad \Longrightarrow \quad |f(x) - f(y)| < \varepsilon \,\big)$$

c) Aus den Bedingungen (197) *und* (188)*, sieht man sofort, dass gleichmäßig stetige Abbildungen insbesondere auch stetig sind. In der Tat ist gleichmäßige Stetigkeit eine stärkere Bedingung als Stetigkeit, da in* (188)*, das zu einem vorgegebenen* $\varepsilon > 0$ *zugehörige* $\delta > 0$*, auch explizit vom Punkt* $x \in X_1$ *abhängen kann. In* (197) *ist es allerdings so, dass das zu einem vorgegebenen* $\varepsilon > 0$ *gehörige* $\delta > 0$*, für alle* $x \in X_1$ *gleichzeitig die Inklusion* $f(\mathsf{B}[\mathrm{d}_1]_\delta(x)) \subseteq \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x))$ *erzwingt.*

Sei Beispielsweise $X_1 = \mathbb{R} = X_2$ *(mit der Betragsmetrik):*

Die Abbildung $f\colon \mathbb{R} \ni x \mapsto x^2 \in \mathbb{R}$ *ist zwar stetig wegen Beispiel 52, aber nicht gleichmäßig stetig (Übung 85 und Teil d)). Anschaulich gesprochen liegt das daran, dass der Anstieg der Parabel* x^2 *(Schule), für große Werte von* $|x|$ *immer größer wird – je größer also* $|x|$ *ist, desto mehr weicht für festgehaltenes* $\delta > 0$ *der Funktionswert* $f(x+\delta)$ *vom Funktionswert* $f(x)$ *ab.*

d) Will man nachweisen, dass eine Abbildung $f\colon X_1 \to X_2$ *zwischen den metrischen Räumen* (X_1, d_1) *und* (X_2, d_2) *nicht gleichmäßig stetig ist, so reicht es per Definition aus ein* $\varepsilon > 0$ *zu finden, sodass gilt:*

$$\begin{gathered} \forall\, \delta > 0\colon\ \exists\, x \in X_1\colon\ f(\mathsf{B}[\mathrm{d}_1]_\delta(x)) \not\subseteq \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x)). \\ \big(\iff \qquad \forall\, \delta > 0\colon\ \exists\, x_\delta, y_\delta \in X_1\colon \quad \mathrm{d}_1(x_\delta, y_\delta) < \delta \quad \wedge \quad \mathrm{d}_2(f(x_\delta), f(y_\delta)) \ge \varepsilon \,\big) \end{gathered} \tag{198}$$

Übung 85. *Zeigen Sie, dass die Abbildung $\mathbb{R} \ni x \mapsto x^2 \in \mathbb{R}$ nicht gleichmäßig stetig ist. Gehen Sie beispielsweisewie in Bemerkung 52.d) beschrieben vor.*

Es gilt der folgenden Sachverhalt:

Satz 34. *Seien $a, b \in \mathbb{R}$ mit $a < b$, sowie $f\colon [a,b] \to \mathbb{R}$ stetig. Dann ist f gleichmäßig stetig.*

Beweis. Wir nehmen an, dass f nicht gleichmäßig stetig ist. Dann existiert ein $\varepsilon > 0$, sodass (197) nicht gilt, d.h. (vgl. (198))

$$\forall\, \delta > 0\colon\ \exists\, x_\delta, y_\delta \in [a,b]\colon \quad (\, |x_\delta - y_\delta| < \delta \quad \wedge \quad |f(x_\delta) - f(y_\delta)| \geq \varepsilon\,).$$

Für jedes $n \in \mathbb{N}_{>0}$, existieren somit $x_n, y_n \in [a,b]$ mit

$$|x_n - y_n| < \tfrac{1}{n} \qquad \text{sowie} \qquad |f(x_n) - f(y_n)| \geq \varepsilon.$$

Dies definiert uns Folgen $(x_n)_{n\in\mathbb{N}_{>0}}$ und $(y_n)_{n\in\mathbb{N}_{>0}}$ in $[a,b]$. Gemäß Korollar 18, existiert eine Teilfolge $(x_{\iota(n)})_{n\in\mathbb{N}}$, die gegen ein $z \in [a,b]$ konvergiert. Dann gilt ebenfalls $(y_{\iota(n)})_{n\in\mathbb{N}} \to z$; und zwar wegen

$$|z - y_{\iota(n)}| \leq |z - x_{\iota(n)}| + \overbrace{|x_{\iota(n)} - y_{\iota(n)}|}^{\leq \frac{1}{\iota(n)}} \qquad \forall\, n \in \mathbb{N}.$$

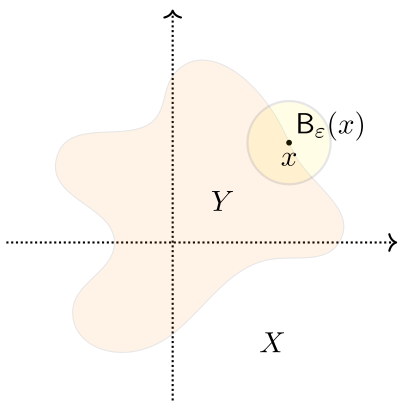
Aus der Stetigkeit der Betragsfunktion (siehe Beispiel 50.d)), erhalten wir mit Proposition 8 (Folgenstetigkeit), dass

$$\lim_n |f(x_{\iota(n)}) - f(y_{\iota(n)})| = |f(z) - f(z)| = 0$$

gilt, was der Annahme $|f(x_n) - f(y_n)| \geq \varepsilon$ für alle $n \in \mathbb{N}_{>0}$ widerspricht (beachte $\iota\colon \mathbb{N} \to \mathbb{N}_{>0}$). $\square$

7.1.5 Grenzwerte von Abbildungen

Definition 52 (Berührungspunkt)**.** *Sei (X, d) ein metrischer Raum, und $Y \subseteq X$ eine Teilmenge. Ein Element $x \in X$ heißt Berührungspunkt von Y, wenn $\mathsf{B}_\varepsilon(x) \cap Y \neq \emptyset$ für alle $\varepsilon > 0$ gilt. Wir notieren die Menge aller Berührungspunkte von Y mit $\mathcal{B}(Y)$.*



Bemerkung 53. *Sei (X, d) ein metrischer Raum, und $Y \subseteq X$ eine Teilmenge.*

a) Es gilt $Y \subseteq \mathcal{B}(Y)$, wegen $y \in \mathsf{B}_\varepsilon(y)$ für alle $y \in Y$ und $\varepsilon > 0$.

b) Es gilt $x \in \mathcal{B}(Y)$ genau dann, wenn eine Folge $(y_n)_{n\in\mathbb{N}}$ in $Y \subseteq X$ existiert mit $\lim_n y_n = x$.

Beweis der Äquivalenz:

- Sei $x \in \mathcal{B}(Y)$. Für jedes $n \in \mathbb{N}_{>0}$, existiert dann ein $y_n \in \mathsf{B}_{\frac{1}{n}}(x) \cap Y \neq \emptyset$. Per Konstruktion gilt $(y_n)_{n\in\mathbb{N}_{>0}} \to x$; denn für $\varepsilon > 0$ und $N \in \mathbb{N}_{>0}$ mit $0 < 1/N < \varepsilon$, folgt

$$y_n \in \mathsf{B}_{\frac{1}{n}}(x) \subseteq \mathsf{B}_{\frac{1}{N}}(x) \subseteq \mathsf{B}_\varepsilon(x) \qquad \forall\, n \geq N.$$

- Sei $x \in X$, sowie $(y_n)_{n\in\mathsf{N}}$ eine Folge in Y mit $\lim_n y_n = x$. Für $\varepsilon > 0$ vorgegeben, existiert dann ein $N \in \mathsf{N}$ mit $Y \ni y_N \in \mathsf{B}_\varepsilon(x)$; also $y_N \in Y \cap \mathsf{B}_\varepsilon(x) \neq \emptyset$. Folglich gilt $x \in \mathcal{B}(Y)$. □

Beispiel 53. *Wir betrachten die folgenden Spezialfälle:*

- *Für $Y = (-1, 1) \subseteq \mathbb{R} = X$, ist $\mathcal{B}(Y) = [-1, 1]$.*
- *Für $Y = \{z \in \mathbb{C} \mid |z| < 1\} \subseteq \mathbb{C} = X$ die offene Einheitskreisscheibe, ist $\mathcal{B}(Y) = \{z \in \mathbb{C} \mid |z| \leq 1\}$ die abgeschlossene Einheitskreisscheibe.*
- *In Verallgemeinerung der beiden vorherigen Beispiele kann man zeigen:*

  Ist $(V, \|\cdot\|)$ ein normierter Raum, so gilt $\mathcal{B}(\mathsf{B}_\varepsilon(x)) = \overline{\mathsf{B}}_\varepsilon(x)$ für alle $x \in X$ und $\varepsilon > 0$.

  (Siehe auch Beispiel 96 und Satz 85.(1).)

Man könnten nun denken, dass auch in metrischen Räumen immer $\mathcal{B}(\mathsf{B}_\varepsilon(x)) = \overline{\mathsf{B}}_\varepsilon(x)$ gilt. Dies ist im Allgemeinen allerdings nur für normierte Räume korrekt. Der Grund ist grob gesprochen der, dass ein normierter $\mathbb{K}$-Vektorraum mit je einem Element v, auch die ganze Gerade $\mathbb{K} \cdot v$ enthält (in metrischen Räumen sind ja derartige Vektorraumoperationen üblicherweise nicht verfügbar). Als Gegenbeispiel betrachten wir den (diskreten) metrischen Raum $(\mathbb{Z}, \mathrm{d}_\mathbb{Z})$ aus Übung 59 (also $\mathbb{Z}$ mit der eingeschränkte Betragsmetrik). Dann gilt

$$\mathcal{B}(\mathsf{B}_1(0)) = \mathsf{B}_1(0) = \{0\} \qquad \textit{sowie} \qquad \overline{\mathsf{B}}_1(0) = \{-1, 0, 1\}.$$

(Wegen $\mathsf{B}_1(x) = \{x\}$ für alle $x \in \mathbb{Z}$, gilt nämlich sogar $\mathsf{B}_1(z) \cap \mathsf{B}_1(0) = \emptyset$ für jedes $z \in \mathbb{Z} \setminus \{0\}$.)

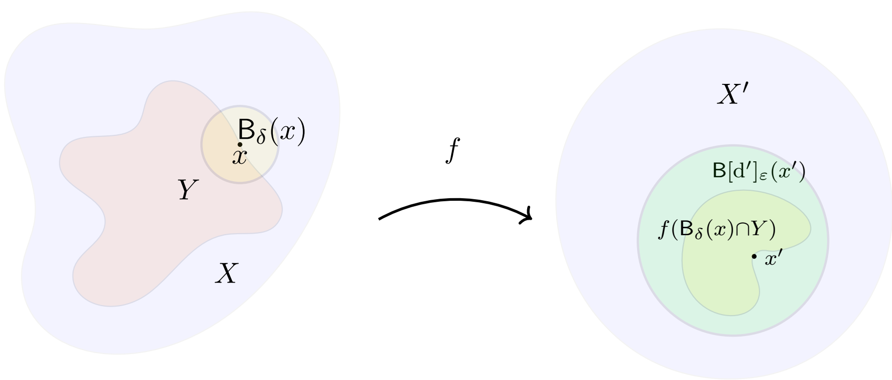
Definition 53 (Grenzwert einer Abbildung)**.** *Seien (X, d), (X', d') metrische Räume, $Y \subseteq X$ eine Teilmenge, und $f\colon Y \to X'$ eine Abbildung. Gegeben $x \in \mathcal{B}(Y)$ und $x' \in X'$, so schreiben wir*

$$\lim_{y\to x} f(y) = x' \tag{199}$$

(f konvergiert für $y \to x$ gegen x') genau dann, wenn gilt:

$$\begin{aligned} &\forall\, \varepsilon > 0\colon\ \exists\, \delta > 0\colon\ f(\mathsf{B}[\mathrm{d}]_\delta(x) \cap Y) \subseteq \mathsf{B}[\mathrm{d}']_\varepsilon(x') \\ \big(\iff \quad &\forall\, \varepsilon > 0\colon\ \exists\, \delta > 0\colon\ \mathsf{B}[\mathrm{d}]_\delta(x) \cap Y \subseteq f^{-1}(\mathsf{B}[\mathrm{d}']_\varepsilon(x'))\big) \end{aligned} \tag{200}$$

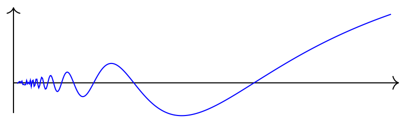
Es wird x' auch als Grenzwert von f in x *bezeichnet.*



Beispiel 54. *Sei $X = \mathbb{R} = X'$, sowie $Y = (0, \infty)$, $x = 0 = x'$, und (Schule)*

$$f\colon Y = (0, \infty) \to \mathbb{R} = X', \qquad y \mapsto y \cdot \sin\left(\tfrac{1}{y}\right).$$

Dann gilt $\lim_{y \to x} f(y) = x'$ *(also* $\lim_{y \to 0} f(y) = 0$*):*



In der Tat, für $\varepsilon > 0$ vorgegeben, sei $\delta := \varepsilon$. Dann gilt (beachte$|\sin| \le 1$):

$$0 < y < \delta \qquad \Longrightarrow \qquad |f(y)| = |y \cdot \sin(\tfrac{1}{y})| = |y| \cdot |\sin(\tfrac{1}{y})| \le y < \delta = \varepsilon$$

also

$$f(\mathsf{B}_\delta(x) \cap Y) = f((-\delta, \delta) \cap (0, \infty)) = f((0, \delta)) \subseteq (-\varepsilon, \varepsilon) = \mathsf{B}_\varepsilon(x').$$

Beispiel 55. *In der Situation von Definition 53, sei $x \in Y$. Dann sind die folgenden Aussagen Äquivalent zueinander:*

A) *Es gilt* $\lim_{y \to x} f(y) = x'$.

B) *$f\colon Y \to X'$ ist stetig in x (bezüglich d_Y) mit $f(x) = x'$.*

Beweis der Äquivalenz: Wegen (139) und $x \in Y$ gilt

$$\mathsf{B}[\mathrm{d}_Y]_\delta(x) = \mathsf{B}[\mathrm{d}]_\delta(x) \cap Y \qquad \forall\, \delta > 0.$$

Daher liest sich (200) (also A)) in diesem Fall auch in der Form:

$$\forall\, \varepsilon > 0\colon\ \exists\, \delta > 0\colon\ f(\mathsf{B}[\mathrm{d}_Y]_\delta(x)) \subseteq \mathsf{B}[\mathrm{d}']_\varepsilon(x'). \tag{201}$$

Zudem liest sich (188) (also B)) in der Form:

$$\forall\, \varepsilon > 0\colon\ \exists\, \delta > 0\colon\ f(\mathsf{B}[\mathrm{d}_Y]_\delta(x)) \subseteq \mathsf{B}[\mathrm{d}']_\varepsilon(f(x)) \tag{202}$$

(Stetigkeit von f in x bezüglich $(X_1, \mathrm{d}_1) = (Y, \mathrm{d}_Y)$ und $(X_2, \mathrm{d}_2) = (X', \mathrm{d}')$).

- Gilt nun B), so gilt (202) mit $f(x) = x'$. Daher gilt (201), also A).
- Gilt nun A), so gilt (201); und Übung 57.a) zeigt (zweite Implikation)

$$\begin{aligned} f(x) \in \mathsf{B}[\mathrm{d}']_\varepsilon(x') \qquad \varepsilon > 0 \qquad &\Longrightarrow \qquad f(x) \in \textstyle\bigcap_{n \in \mathbb{N}_{>0}} \mathsf{B}[\mathrm{d}']_{\frac{1}{n}}(x') = \{x'\} \\ &\Longrightarrow \qquad f(x) = x'. \end{aligned}$$

Somit gilt (202), also B). □

Proposition 10 (Stetige Fortsetzung in einem Punkt)**.** *Seien (X, d), (X', d') metrische Räume, $Y \subseteq X$ eine Teilmengen, und $f\colon Y \to X'$ eine Abbildung. Seien weiterhin $x \in \mathcal{B}(Y) \setminus Y$ und $x' \in X'$ vorgegeben. Wir betrachten den metrischen Raum $(\hat{Y}, \mathrm{d}_{\hat{Y}})$ mit $\hat{Y} := Y \cup \{x\}$; sowie die Abbildung $\hat{f}\colon \hat{Y} \to X'$, definiert durch*

$$\hat{f}(z) := \begin{cases} f(z) & \textit{für} \quad z \in Y \\ x' & \textit{für} \quad z = x \end{cases}$$

für alle $z \in \hat{Y}$. Es gelten die folgenden Aussagen:

1) Es ist $\lim_{y\to x} f(y) = x'$ *äquivalent dazu, dass* $\hat{f}$ *stetig in* x *ist.*

2) $*$ *Es gelte* $\lim_{y\to x} f(y) = x'$*. Dann ist* f *stetig (bzgl.* d_Y*) genau dann, wenn* $\hat{f}$ *stetig (bzgl.* $\mathrm{d}_{\hat{Y}}$*) ist.*

Beweis. 1) • Ist $\hat{f}\colon \hat{Y} \to X'$ stetig in x, so erhalten wir

$$\begin{aligned}
& \forall\,\varepsilon > 0\colon\ \exists\,\delta > 0\colon\ \hat{f}\big(\mathsf{B}[\mathrm{d}_{\hat{Y}}]_\delta(x)\big) \subseteq \mathsf{B}[\mathrm{d}']_\varepsilon(\hat{f}(x)) = \mathsf{B}[\mathrm{d}']_\varepsilon(x') \\
\overset{(139)}{\iff}\quad & \forall\,\varepsilon > 0\colon\ \exists\,\delta > 0\colon\ \hat{f}\big(\mathsf{B}[\mathrm{d}]_\delta(x) \cap \hat{Y}\big) \subseteq \mathsf{B}[\mathrm{d}']_\varepsilon(x') \\
\overset{Y\subseteq\hat{Y}}{\implies}\quad & \forall\,\varepsilon > 0\colon\ \exists\,\delta > 0\colon\ \hat{f}\big(\mathsf{B}[\mathrm{d}]_\delta(x) \cap Y\big) \subseteq \mathsf{B}[\mathrm{d}']_\varepsilon(x') \\
\overset{\hat{f}|_Y = f}{\iff}\quad & \forall\,\varepsilon > 0\colon\ \exists\,\delta > 0\colon\ f\big(\mathsf{B}[\mathrm{d}]_\delta(x) \cap Y\big) \subseteq \mathsf{B}[\mathrm{d}']_\varepsilon(x'),
\end{aligned}$$

also $\lim_{y\to x} f(y) = x'$.

• Gilt $\lim_{y\to x} f(y) = x'$, so erhalten wir $\big($beachte $\mathsf{B}[\mathrm{d}]_\delta(x) \cap \hat{Y} = \{x\} \cup (\mathsf{B}[\mathrm{d}]_\delta(x) \cap Y)\ \big)$

$$\begin{aligned}
& \forall\,\varepsilon > 0\colon\ \exists\,\delta > 0\colon\ f\big(\mathsf{B}[\mathrm{d}]_\delta(x) \cap Y\big) \subseteq \mathsf{B}[\mathrm{d}']_\varepsilon(x') = \mathsf{B}[\mathrm{d}']_\varepsilon(\hat{f}(x)) \\
\implies\quad & \forall\,\varepsilon > 0\colon\ \exists\,\delta > 0\colon\ \hat{f}\big(\mathsf{B}[\mathrm{d}]_\delta(x) \cap \hat{Y}\big) \subseteq \mathsf{B}[\mathrm{d}']_\varepsilon(\hat{f}(x)) \\
\overset{(139)}{\iff}\quad & \forall\,\varepsilon > 0\colon\ \exists\,\delta > 0\colon\ \hat{f}\big(\mathsf{B}[\mathrm{d}_{\hat{Y}}]_\delta(x)\big) \subseteq \mathsf{B}[\mathrm{d}']_\varepsilon(\hat{f}(x)).
\end{aligned}$$

Dies zeigt die Stetigkeit von $\hat{f}$ in x.

2) • Ist $\hat{f}$ stetig, so ist wegen Lemma 40.2) auch f stetig; denn es gilt $\hat{f}|_Y = f$, und die Einschränkung von $\mathrm{d}_{\hat{Y}}$ auf Y ist gegeben durch d_Y.

• Es sei f stetig (bezüglich d_Y).

– Wegen Teil 1) ist $\hat{f}$ stetig in x.

– Sei $y \in Y$, sowie $\varepsilon > 0$ vorgegeben. Da f stetig ist, existiert ein $\delta > 0$ mit

$$\mathrm{d}_Y(y,z) < \delta \quad \text{für} \quad z \in Y \qquad \implies \qquad \mathrm{d}'(f(y), f(z)) < \varepsilon.$$

Wegen $x \neq y$, gilt $\delta' := \mathrm{d}_{\hat{Y}}(x,y) > 0$. Für $\hat{\delta} := \min(\delta, \delta')$, gilt dann

$$\begin{aligned}
\mathrm{d}_{\hat{Y}}(y,z) < \hat{\delta} \quad \text{für} \quad z \in \hat{Y} \qquad & \implies \qquad x \neq z \in Y \ \wedge\ \mathrm{d}_Y(y,z) < \delta \\
& \implies \qquad \mathrm{d}'(f(y), f(z)) < \varepsilon,
\end{aligned}$$

was die Stetigkeit von $\hat{f}$ in x zeigt. □

Proposition 11. *Seien* (X, d), (X', d') *metrische Räume,* $Y \subseteq X$ *eine Teilmenge,* $x \in \mathcal{B}(Y)$, $x' \in X'$, *und* $f\colon Y \to X'$ *eine Abbildung. Die folgenden Aussagen sind äquivalent zueinander:*

a) Es gilt $\lim_{y\to x} f(y) = x'$.

b) Es gilt die Implikation:

$$(y_n)_{n\in\mathsf{N}} \ \textit{Folge in}\ \ Y \subseteq X \ \ \textit{mit} \ \ (y_n)_{n\in\mathsf{N}} \to x \qquad \implies \qquad (f(y_n))_{n\in\mathsf{N}} \to x'. \tag{203}$$

Beweis. Der Beweis verläuft analog zum Beweis der Äquivalenz von Stetigkeit in einem Punkt zur Folgenstetigkeit in einem Punkt – siehe Proposition 8:

- Es gelte $\lim_{y\to x} f(y) = x'$, also (200). Sei $(y_n)_{n\in\mathsf{N}}$ eine Folge in Y mit $(y_n)_{n\in\mathsf{N}} \to x$. Für $\varepsilon > 0$ vorgegeben, sei $\delta > 0$ wie in (200), d.h.,
$$\mathsf{B}[\mathrm{d}]_\delta(x) \cap Y \subseteq f^{-1}(\mathsf{B}[\mathrm{d}']_\varepsilon(x')). \tag{204}$$
Dann existiert ein $N \in \mathsf{N}$ mit
$$\begin{aligned} y_n \in \mathsf{B}[\mathrm{d}]_\delta(x) \qquad \forall\, n \geq N \qquad &\overset{(204)}{\Longrightarrow} \qquad y_n \in \mathsf{B}[\mathrm{d}]_\delta(x) \cap Y \subseteq f^{-1}(\mathsf{B}[\mathrm{d}']_\varepsilon(x')) \qquad \forall\, n \geq N \\ &\Longrightarrow \qquad f(y_n) \in \mathsf{B}[\mathrm{d}']_\varepsilon(x') \qquad \forall\, n \geq N. \end{aligned}$$
Dies zeigt $(f(y_n))_{n\in\mathsf{N}} \to x'$.
- Wir zeigen die Implikation „*b*) $\Rightarrow$ *a*)“ durch Kontraposition: Es gelte nicht $\lim_{y\to x} f(y) = x'$. Dann existiert ein $\varepsilon > 0$ mit
$$\mathsf{B}[\mathrm{d}]_\delta(x) \cap Y \not\subseteq f^{-1}(\mathsf{B}[\mathrm{d}']_\varepsilon(x')) \qquad \forall\, \delta > 0.$$
Für jedes $n \in \mathbb{N}_{>0}$, existiert daher ein
$$Y \ni y_n \in \big(\mathsf{B}[\mathrm{d}]_{\frac{1}{n}}(x) \cap Y\big) \setminus f^{-1}(\mathsf{B}[\mathrm{d}']_\varepsilon(x')).$$
Per Konstruktion gilt dann $(y_n)_{n\in\mathbb{N}_{>0}} \to x$, sowie
$$y_n \notin f^{-1}(\mathsf{B}[\mathrm{d}']_\varepsilon(x')) \qquad \forall\, n \in \mathbb{N}_{>0} \qquad \Longleftrightarrow \qquad f(y_n) \notin \mathsf{B}[\mathrm{d}']_\varepsilon(x') \qquad \forall\, n \in \mathbb{N}_{>0},$$
womit $(f(y_n))_{n\in\mathbb{N}_{>0}}$ nicht gegen x' konvergieren kann. Somit gilt (203) nicht. $\square$

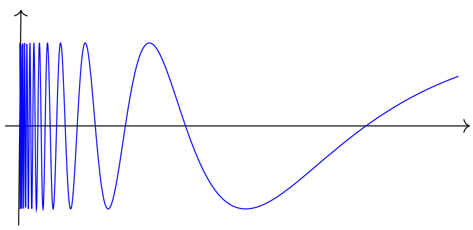
Zusammen mit Satz 12, folgt:

Korollar 26. *Seien (X, d), (X', d') metrische Räume, $Y \subseteq X$ eine Teilmenge, $x \in \mathcal{B}(Y)$, und $f\colon Y \to X'$ eine Abbildung. Der Grenzwert von f in x ist eindeutig, sofern er existiert.*

Beweis. Es gelte $\lim_{y\to x} f(y) = x'_1$ sowie $\lim_{y\to x} f(y) = x'_2$, für gewisse $x'_1, x'_2 \in X'$.

- Wegen Bemerkung 53.b), existiert eine Folge $(y_n)_{n\in\mathsf{N}}$ in $Y \subseteq X$ mit $\lim_n y_n = x$.
- Wegen Proposition 11 (also (203)), gilt $x'_1 = \lim_n f(y_n) = x'_2$.
- Satz 12 (Eindeutigkeit des Grenzwertes von Folgen) zeigt nun $x'_1 = x'_2$. $\square$

Beispiel 56. *Sei $X = \mathbb{R} = X'$, $Y = (0, \infty)$, $x = 0$, und $f\colon (0,\infty) \ni y \mapsto \sin(\frac{1}{y}) \in \mathbb{R}$ (Schule). Dann existiert $\lim_{y\to x} f(y)$ nicht:*



Für jedes $n \in \mathbb{N}_{>0}$, sei
$$y_n := 1/(2\pi n + \tfrac{\pi}{2}) \qquad \textit{und} \qquad y'_n := 1/(2\pi n + \tfrac{3\pi}{2}).$$
Dann gilt $\lim_n y_n = 0 = \lim_n y'_n$ sowie
$$\lim_n f(y_n) = \lim_n \sin(2\pi n + \tfrac{\pi}{2}) = \lim_n 1 = 1 \neq -1 = \lim_n -1 = \lim_n \sin(2\pi n + \tfrac{3\pi}{2}) = \lim_n f(y'_n),$$
womit (203) für kein $x' \in \mathbb{R}$ gelten kann.

Beispiel 57 (Erhalt von Ungleichungen unter Limiten)**.** *Sei* (X, d) *ein metrischer Raum,* $Y \subseteq X$ *eine Teilmenge,* $x \in \mathcal{B}(Y)$*, sowie* $f\colon Y \to \mathbb{R} = X'$ *eine Abbildung mit* $\lim_{y\to x} f(y) = x' \in \mathbb{R}$*. Es gelten die folgenden Implikationen:*

$$\exists\,(c \in \mathbb{R} \ \wedge\ \delta > 0)\colon \quad f|_{\mathsf{B}[\mathrm{d}]_\delta(x)\,\cap\, Y} \leq c \qquad\Longrightarrow\qquad x' = \lim\nolimits_{y\to x} f(y) \leq c \tag{205}$$

$$\exists\,(c \in \mathbb{R} \ \wedge\ \delta > 0)\colon \quad f|_{\mathsf{B}[\mathrm{d}]_\delta(x)\,\cap\, Y} \geq c \qquad\Longrightarrow\qquad x' = \lim\nolimits_{y\to x} f(y) \geq c. \tag{206}$$

Beweis. Wegen Bemerkung 53.b), existiert ein Folge $(y_n)_{n\in\mathsf{N}}$ in Y mit $\lim_n y_n = x$. Somit existiert ein $N \in \mathsf{N}$ mit $y_n \in \mathsf{B}[\mathrm{d}]_\delta(x)$ für alle $n \geq N$. Wegen der linken Seite von (205), gilt dann $f(y_n) \leq c$ für alle $n \geq N$. Mit Proposition 11 (zweiter Schritt) und Korollar 14 (dritter Schritt) folgt

$$x' = \lim\nolimits_{y\to x} f(y) = \lim\nolimits_n f(y_n) \leq c.$$

Die Implikation (206) zeigt man analog. □

Übung*. *Seien* (X, d), (X', d') *metrische Räume,* $Y \subseteq X$ *eine Teilmenge,* $x \in \mathcal{B}(Y)$*, sowie* $f\colon Y \to X'$ *eine Abbildung mit* $\lim_{y\to x} f(y) = x' \in X'$*. Zeigen Sie, dass für* $\varepsilon, \delta > 0$ *gilt:*

$$f(\mathsf{B}[\mathrm{d}]_\delta(x) \cap Y) \subseteq \overline{\mathsf{B}}[\mathrm{d}']_\varepsilon(x') \qquad\Longrightarrow\qquad \lim\nolimits_{y\to x} f(y) = x' \in \overline{\mathsf{B}}[\mathrm{d}']_\varepsilon(x').$$

Hinweis: Verallgemeinern Sie das Argument in Beispiel 57, indem Sie anstelle Korollar 14, Lemma 36 und Lemma 28 benutzen.

Aus Proposition 11, sowie Proposition 6, Lemma 32, und der Folgenstetigkeit der komplexen Konjugation (Übung und Proposition 8), erhalten wir umgehend die folgenden Rechenregeln:[33]

Korollar 27. *Sei* (X, d) *ein metrischer Raum,* $(V, \|\cdot\|)$ *ein normierter* $\mathbb{K}$*-Vektorraum,* $Y \subseteq X$ *eine Teilmengen, und* $x \in \mathcal{B}(Y)$ *ein Punkt.*

1) Seien $f, g \in \mathsf{Abb}(Y, V)$ *und* $h \in \mathsf{Abb}(Y, \mathbb{K})$*, sodass* $\lim_{y\to x} f(y) \in V$, $\lim_{y\to x} g(y) \in V$*, sowie* $\lim_{y\to x} h(y) \in \mathbb{K}$ *existiert. Dann gilt*

$$\lim_{y\to x}(\lambda\cdot f + \mu\cdot g)(y) = \lambda \cdot \lim_{y\to x} f(y) + \mu\cdot \lim_{y\to x} f(y) \qquad \forall\, \lambda, \mu \in \mathbb{K}$$

$$\lim_{y\to x}(h\cdot f)(y) = \lim_{y\to x} h(y)\ \cdot\ \lim_{y\to x} f(y).$$

2) Seien $f \in \mathsf{Abb}(X, \mathbb{K}_\times)$, $g \in \mathsf{Abb}(X, \mathbb{K})$ *und* $h \in \mathsf{Abb}(X, \mathbb{C})$*, sodass* $\lim_{y\to x} f(y) \in \mathbb{K}_\times$, $\lim_{y\to x} g(y) \in \mathbb{K}$*, sowie* $\lim_{y\to x} h(y) \in \mathbb{C}$ *existiert. Dann gilt*

$$\lim_{y\to x}\left(\frac{1}{f}\right)(y) = \frac{1}{\lim_{y\to x} f(y)}, \qquad \lim_{y\to x}\left(\frac{g}{f}\right)(y) = \frac{\lim_{y\to x} g(y)}{\lim_{y\to x} f(y)}, \qquad \lim_{y\to x} \overline{h}(y) = \overline{\lim_{y\to x} f(y)}.$$

Wir betrachten nun abschließend den Spezialfall reellwertiger Funktion:

Terminologie 22. *Sei* $D \subseteq \mathbb{R}$ *eine Teilmenge.*

1) *Für* $\varepsilon > 0$ *setzen wir*

$$\mathsf{B}_\varepsilon(-\infty) := \{x \in \mathbb{R} \mid x < -\varepsilon^{-1}\} \qquad \textit{sowie} \qquad \mathsf{B}_\varepsilon(\infty) := \{x \in \mathbb{R} \mid \varepsilon^{-1} < x\}.$$

Hiermit gilt für eine Folge $(x_n)_{n\in\mathsf{N}}$ *in* $\mathbb{R}$ *(vgl. Terminologie 14):*

$$\lim\nolimits_n x_n = \pm\infty \qquad\Longleftrightarrow\qquad \forall\,\varepsilon > 0\colon \exists\, N_\varepsilon \in \mathsf{N}\colon \quad x_n \in \mathsf{B}_\varepsilon(\pm\infty) \quad \forall\, n \geq N_\varepsilon. \tag{207}$$

[33] Beispielsweise gilt für jede Folge $(y_n)_{n\in\mathsf{N}}$ in Y mit $\lim_n y_n = x$, dass $\lim_n(\lambda\cdot f + \mu\cdot g)(y_n) = \lim_n(\lambda\cdot f(y_n) + \mu\cdot g(y_n)) = \lambda\cdot\lim_n f(y_n) + \mu\cdot\lim_n g(y_n) = \lambda\cdot\lim_{y\to x} f(y) + \mu\cdot\lim_{y\to x} f(y)$, wegen Proposition 11 im letzten Schritt und Proposition 6 im zweiten Schritt. Da die Folge $(y_n)_{n\in\mathsf{N}}$ beliebig war, folgt die erste Identität in Teil 1) aus Proposition 11.

2) *In Erweiterung von Definition 52 setzen wir*

$$\overline{\mathcal{B}}(D) := \left\{ x \in \overline{\mathbb{R}} \mid \mathsf{B}_\varepsilon(x) \cap D \neq \emptyset \ \ \textit{für alle} \ \ \varepsilon > 0 \right\} \subseteq \overline{\mathbb{R}}.$$

Hiermit gilt:

- $\overline{\mathcal{B}}(D) = \mathcal{B}(D) \subseteq \mathbb{R} \quad \Leftrightarrow \quad D$ *ist beschränkt*
- $\infty \in \overline{\mathcal{B}}(D) \quad \Leftrightarrow \quad D$ *ist nach oben unbeschränkt.*
- $-\infty \in \overline{\mathcal{B}}(D) \quad \Leftrightarrow \quad D$ *ist nach unten unbeschränkt.*

Insbesondere folgt mit Bemerkung 53.b), dass $x \in \overline{\mathcal{B}}(D)$ genau dann gilt, wenn eine Folge $(x_n)_{n\in\mathsf{N}}$ in D mit $\lim_n x_n = x$ existiert.

3) *Sei $f\colon \mathbb{R} \supseteq D \to \mathbb{R}$ eine Abbildung, sowie $p \in \overline{\mathcal{B}}(D)$ und $a \in \overline{\mathbb{R}}$. Wir vereinbaren die folgenden Erweiterung von* (199)*:*

- *Wir schreiben*

$$\lim_{x\to p} f(x) = a \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad \forall\, \varepsilon > 0\colon\ \exists\, \delta > 0\colon\ f(\mathsf{B}_\delta(p) \cap D) \subseteq \mathsf{B}_\varepsilon(a). \tag{208}$$

Der gleiche Beweis[34] *wie in Proposition 11 zeigt dann, dass* (208) *äquivalent ist zu:*

$$(x_n)_{n\in\mathsf{N}} \ \ \textit{Folge in} \ \ D \ \ \textit{mit} \ \ (x_n)_{n\in\mathsf{N}} \to p \qquad \Longrightarrow \qquad (f(x_n))_{n\in\mathsf{N}} \to a. \tag{209}$$

- *Wir setzen $(\infty,\infty) := \emptyset =: (-\infty,-\infty)$, und schreiben*

$$\lim_{x\uparrow p} f(x) = a \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad \forall\, \varepsilon > 0\colon\ \exists\, \delta > 0\colon\ f\big(\mathsf{B}_\delta(p) \cap D \cap (-\infty,p)\big) \subseteq \mathsf{B}_\varepsilon(a) \tag{210}$$

$$\lim_{x\downarrow p} f(x) = a \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad \forall\, \varepsilon > 0\colon\ \exists\, \delta > 0\colon\ f\big(\mathsf{B}_\delta(p) \cap D \cap (p,\infty)\big) \subseteq \mathsf{B}_\varepsilon(a). \tag{211}$$

Im ersten/zweiten Fall wird a als linksseitiger/rechtsseitige Grenzwert *von f in p bezeichnet.*

Beispiel: *Sei $f\colon D = \mathbb{R}_\times \to \mathbb{R}$ gegeben durch*

$$f(x) = \begin{cases} -1 & \textit{für} \ \ x < 0 \\ \ \ 1 & \textit{für} \ \ x > 0. \end{cases}$$

Dann gilt $\lim_{x\uparrow 0} f(x) = -1$ und $\lim_{x\downarrow 0} f(x) = 1$. □

Es ist (210) *äquivalent zu* (212)*, sowie* (211) *äquivalent zu* (213)*:*

$$(x_n)_{n\in\mathsf{N}} \ \ \textit{Folge in} \ \ D \cap (-\infty,p) \ \ \textit{mit} \ \ (x_n)_{n\in\mathsf{N}} \to p \qquad \Longrightarrow \qquad (f(x_n))_{n\in\mathsf{N}} \to a \tag{212}$$

$$(x_n)_{n\in\mathsf{N}} \ \ \textit{Folge in} \ \ D \cap \ (p,\infty) \ \ \textit{mit} \ \ (x_n)_{n\in\mathsf{N}} \to p \qquad \Longrightarrow \qquad (f(x_n))_{n\in\mathsf{N}} \to a \tag{213}$$

Beweis der Äquivalenz: Wir zeigen nur die Äquivalenz von (211) und (213) (den anderen Fall behandelt man analog). Sei hierfür $\tilde{D} = D \cap (p,\infty)$.

– Gilt $p \notin \overline{\mathcal{B}}(\tilde{D})$, so sind die Bedingungen (211) und (213) beide trivialerweise erfüllt.

  (Es existiert ein $\delta > 0$ mit $\mathsf{B}_\delta(p) \cap \tilde{D} = \emptyset$; und es existiert keine Folge $(x_n)_{n\in\mathsf{N}}$ in $\tilde{D}$ mit $\lim_n x_n = p$.)

– Gilt $p \in \overline{\mathcal{B}}(\tilde{D})$, so folgt die Äquivalenz von (211) und (213) sofort aus der Äquivalenz von (208) und (209), wenn man dort jeweils D durch $\tilde{D}$ ersetzt. □

[34] Man ersetze dort einfach Y durch D, x durch p, x' durch a, $(y_n)_{n\in\mathsf{N}}$ durch $(x_n)_{n\in\mathsf{N}}$, sowie $\mathsf{B}[\mathrm{d}], \mathsf{B}[\mathrm{d}']$ durch B. Man beachte weiterhin (207).

Übung 86. *Machen Sie sich die Aussagen in Terminologie 22.1) und 22.2) klar. Bestimmen Sie zudem $\overline{\mathcal{B}}(D)$ für die folgenden Wahlen von $D \subseteq \mathbb{R}$:*

$$(-1,1), \quad [-1,1], \quad (0,\infty), \quad (-\infty,0), \quad \mathbb{N}, \quad \mathbb{Z}.$$

Lemma 41. *Sei $f\colon \mathbb{R} \supseteq D \to \mathbb{R}$ eine Abbildung, $p \in \overline{\mathcal{B}}(D) \setminus D$, und $a \in \overline{\mathbb{R}}$. Dann gilt*

$$\lim_{x\to p} f(x) = a \qquad \Longleftrightarrow \qquad \lim_{x\uparrow p} f(x) = a \quad \wedge \quad \lim_{x\downarrow p} f(x) = a. \tag{214}$$

Beweis. • Gilt $\lim_{x\to p} f(x) = a$, so wegen

$$\mathsf{B}_\delta(p) \cap D \cap (-\infty, p),\ \mathsf{B}_\delta(p) \cap D \cap (p,\infty) \ \subseteq\ \mathsf{B}_\delta(p) \cap D \qquad \forall\, \delta > 0$$

auch $\lim_{x\uparrow p} f(x) = a = \lim_{x\downarrow p} f(x)$.

• Es gelte die rechte Seite von (214). Für $\varepsilon > 0$ vorgegeben, existieren dann $\delta_\pm > 0$ mit

$$f\big(\mathsf{B}_{\delta_-}(p) \cap D \cap (-\infty,p)\big) \subseteq \mathsf{B}_\varepsilon(a) \qquad \text{sowie} \qquad f\big(\mathsf{B}_{\delta_+}(p) \cap D \cap (p,\infty)\big) \subseteq \mathsf{B}_\varepsilon(a).$$

Für $\delta := \min(\delta_-, \delta_+)$, gilt nun

$$\begin{aligned} \mathsf{B}_\delta(p) \cap D &\overset{p\notin D}{=} \big(\ \mathsf{B}_\delta(p) \cap D \cap (-\infty,p)\big) \cup \big(\ \mathsf{B}_\delta(p) \cap D \cap (p,\infty)\big) \\ &\subseteq \big(\mathsf{B}_{\delta_-}(p) \cap D \cap (-\infty,p)\big) \cup \big(\mathsf{B}_{\delta_+}(p) \cap D \cap (p,\infty)\big), \end{aligned}$$

und wir erhalten somit (vgl. Übung 29)

$$f\big((\mathsf{B}_\delta(p) \cap D\big) \subseteq f\big(\mathsf{B}_{\delta_-}(p) \cap D \cap (-\infty,p)\big) \cup f\big(\mathsf{B}_{\delta_+}(p) \cap D \cap (p,\infty)\big) \ \subseteq\ \mathsf{B}_\varepsilon(a). \qquad \square$$

Lemma 42. *Sei $f\colon \mathbb{R} \supseteq D \to \mathbb{R}$ eine Abbildung, $p \in D \subseteq \mathcal{B}(D)$, und $a \in \mathbb{R}$. Es gilt*

$$\lim_{x\to p} f(x) = a \qquad \Longleftrightarrow \qquad f(p) = a \quad \wedge \quad \lim_{x\uparrow p} f(x) = a \quad \wedge \quad \lim_{x\downarrow p} f(x) = a. \tag{215}$$

Beweis. • Gilt $\lim_{x\to p} f(x) = a$, so auch $\lim_{x\uparrow p} f(x) = a = \lim_{x\downarrow p} f(x)$ (gleiches Argument wie in Lemma 41). Weiterhin gilt $a = \lim_n f(p) = f(p)$ (konstante Folge $(p)_{n\in\mathbb{N}}$) wegen (209).[35]

• Es gelte die rechte Seite von (215), und es sei $\varepsilon > 0$ vorgegeben. Mit der gleichen Wahl von $\delta_\pm$ sowie $\delta := \min(\delta_-, \delta_+)$ wie im Beweis von Lemma 41, gilt zunächst

$$f\big(\mathsf{B}_{\delta_-}(p) \cap D \cap (-\infty,p)\big) \subseteq \mathsf{B}_\varepsilon(a) \qquad \text{sowie} \qquad f\big(\mathsf{B}_{\delta_+}(p) \cap D \cap (p,\infty)\big) \subseteq \mathsf{B}_\varepsilon(a)$$

und

$$\begin{aligned} \mathsf{B}_\delta(p) \cap D &= \{p\} \cup \big(\ \mathsf{B}_\delta(p) \cap D \cap (-\infty,p)\big) \cup \big(\ \mathsf{B}_\delta(p) \cap D \cap (p,\infty)\big) \\ &\subseteq \{p\} \cup \big(\mathsf{B}_{\delta_-}(p) \cap D \cap (-\infty,p)\big) \cup \big(\mathsf{B}_{\delta_+}(p) \cap D \cap (p,\infty)\big). \end{aligned}$$

Wegen $f(p) = a$ erhalten wir

$$f\big((\mathsf{B}_\delta(p) \cap D\big) \subseteq \underbrace{\{f(p)\}}_{=\{a\}} \cup f\big(\mathsf{B}_{\delta_-}(p) \cap D \cap (-\infty,p)\big) \cup f\big(\mathsf{B}_{\delta_+}(p) \cap D \cap (p,\infty)\big) \ \subseteq\ \mathsf{B}_\varepsilon(a). \qquad \square$$

Beispiel 58.

a) Sei $m \in \mathbb{N}_{>0}$ und $f\colon \mathbb{R} \ni x \mapsto x^m \in \mathbb{R}$. Dann gilt $\lim_{x\to\infty} f(x) = \infty$; denn für jede Folge $(x_n)_{n\in\mathsf{N}}$ in $\mathbb{R}$ mit $\lim_n x_n = \infty$, gilt auch $\lim_n (x_n)^m = \infty$ (wegen $x^m \geq x$ für $x \geq 1$). Analog folgt

$$\lim_{x\to-\infty} f(x) = \begin{cases} \infty & \textit{für } m \textit{ gerade} \\ -\infty & \textit{für } m \textit{ ungerade.} \end{cases}$$

[35] Alternativ erhält man $f(p) = a$ aus Beispiel 55 und zwar wegen $p \in D \subseteq \mathcal{B}(D)$.

b) *Sei* $m \in \mathbb{N}_{>0}$ *und* $f\colon \mathbb{R}_\times \ni x \mapsto \frac{1}{x^m} \in \mathbb{R}$. *Dann gilt* $\lim_{x\to\pm\infty} f(x) = 0$. *In der Tat, für jede Folge* $(x_n)_{n\in\mathsf{N}}$ *in* $\mathbb{R}_\times$ *mit* $\lim_n x_n = \pm\infty$, *gilt* $\lim_n \frac{1}{x_n} = 0$. *Wegen Korollar 17.3) gilt dann auch* $\lim_n f(x_n) = \lim_n \left(\frac{1}{x_n}\right)^m = 0$.

c) $*$ *Für die Gaußklammer* $f\colon \mathbb{R} = D \ni x \mapsto [x] = \max(\{k \in \mathbb{Z} \mid k \leq x\}) \in \mathbb{Z} \subseteq \mathbb{R}$, *gilt*

$$\lim_{x\to p} f(x) = f(p) \quad \textit{für alle} \quad p \notin \mathbb{Z}.$$

Für $p \in \mathbb{Z}$, *gilt* $\lim_{x\uparrow p} f(x) = p - 1 \neq p = f(p) = p = \lim_{x\downarrow p} f(x)$. *Also existiert* $\lim_{x\to p} f(x)$ *gemäß Lemma 42 nicht. Gemäß Beispiel 55 ist weiterhin* f *genau dann stetig in* $p \in D = \mathbb{R}$, *wenn* $p \notin \mathbb{Z}$ *gilt.*

Übung 87. *Zeigen Sie* $\lim_{x\to\infty} \mathrm{e}^x = \infty$ *sowie* $\lim_{x\to-\infty} \mathrm{e}^x = 0$.

Übung 88. *Sei* $f\colon D \to (0,\infty)$ *eine Abbildung, und* $p \in \overline{\mathfrak{B}}(D)$. *Zeigen Sie:*

- $\lim_{x\to p} f(x) = 0 \quad \Longrightarrow \quad \lim_{x\to p} \left(\frac{1}{f}\right)(x) = \infty$
- $\lim_{x\to p} f(x) = \infty \quad \Longrightarrow \quad \lim_{x\to p} \left(\frac{1}{f}\right)(x) = 0$

Übung 89. *Sei* $f\colon D \to \mathbb{R}$ *eine Abbildung, und* $p \in \overline{\mathfrak{B}}(D)$. *Zeigen Sie für* $\lambda > 0$ *und* $\mu < 0$:

- $\lim_{x\to p} f(x) = \pm\infty \quad \Longrightarrow \quad \lim_{x\to p} (\lambda \cdot f)(x) = \pm\infty$
- $\lim_{x\to p} f(x) = \pm\infty \quad \Longrightarrow \quad \lim_{x\to p} (\mu \cdot f)(x) = \mp\infty$

7.2 Eigenschaften stetiger reeller Funktionen

In diesem Abschnitt diskutieren wir einige Eigenschaften stetiger Funktionen $f\colon D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}$ ein Intervall. Einige dieser Eigenschaften gelten in verallgemeinerter Form auch in höheren Dimensionen, und sind dann Gegenstand der Analysis 2 Vorlesung.

7.2.1 Der Satz vom Maximum

Satz 35 (Satz vom Maximum)**.** *Sei* $f\colon \mathbb{R} \supseteq [a,b] \to \mathbb{R}$ *mit* $a \leq b$ *eine stetige Funktion. Dann gilt*

$$\inf_{\mathbb{R}}(\mathrm{im}(f)) \in \mathrm{im}(f) \ni \sup_{\mathbb{R}}(\mathrm{im}(f)).$$

Insbesondere ist $\mathrm{im}(f)$ *beschränkt.*

(Jede stetige Funktion auf einem abgeschlossenen und beschränkten Intervall nimmt sowohl ihr Maximum als auch ihr Minimum an, vgl. Übung 18 und Lemma 2.)

Beweis. Wir zeigen nur $\sup_{\mathbb{R}}(\mathrm{im}(f)) \in \mathrm{im}(f)$; die Aussage $\inf_{\mathbb{R}}(\mathrm{im}(f)) \in \mathrm{im}(f)$ folgt analog. Sei hierfür $o := \sup_{\overline{\mathbb{R}}}(\mathrm{im}(f)) \in \mathbb{R} \cup \{\infty\}$:

- Aus der Definition des Supremums, erhalten wir die Existenz einer Folge $(z_n)_{n\in\mathbb{N}_{>0}}$ in $\mathrm{im}(f)$, mit $(z_n)_{n\in\mathbb{N}_{>0}} \to o$. In der Tat:
  - Gilt $o \in \mathbb{R}$, so existiert gemäß Lemma 20 zu jedem $n \in \mathbb{N}_{>0}$ ein $z_n \in \mathrm{im}(f)$ mit $z_n > o - \frac{1}{n}$, d.h., $|o - z_n| = o - z_n < \frac{1}{n}$.
  - Gilt $o = \infty$, so existiert zu jedem $n \in \mathbb{N}_{>0}$ ein $z_n \in \mathrm{im}(f)$ mit $z_n \geq n$ (ansonsten wäre $\mathrm{im}(f)$ nach oben beschränkt).
- Betrachten der Urbilder $f^{-1}(z_n) \subseteq [a,b]$, liefert eine Folge $(x_n)_{n\in\mathbb{N}_{>0}}$ in $[a,b]$, mit $f(x_n) = z_n$ für alle $n \in \mathbb{N}_{>0}$.
- Wegen Korollar 18, existiert eine in $[a,b]$ konvergente Teilfolge $(x_{\iota(n)})_{n\in\mathbb{N}} \to x \in [a,b]$. Wegen Korollar 16 (zweiter Schritt) und Proposition 8 (vierter Schritt), gilt dann

$$o = \lim_n z_n = \lim_n z_{\iota(n)} = \lim_n f(x_{\iota(n)}) = f(x) \in \mathrm{im}(f),$$

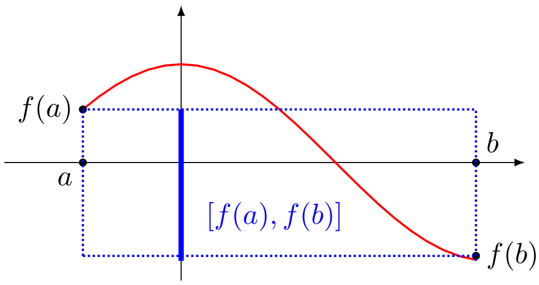
was die Behauptung zeigt. □

7.2.2 Der Zwischenwertsatz

In diesem Abschnitt beweisen wir den Zwischenwertsatz; und behandeln ein Stetigkeitsresultat für Umkehrabbildungen stetiger Funktionen auf $\mathbb{R}$.

Satz 36 (Zwischenwertsatz)**.** *Sei $f\colon \mathbb{R} \supseteq [a,b] \to \mathbb{R}$ mit $a \leq b$ eine stetige Funktion. Dann nimmt f jeden Wert zwischen $f(a)$ und $f(b)$ an – d.h., es gilt:*

- $f(a) \leq f(b) \quad \Rightarrow \quad [f(a), f(b)] \subseteq \operatorname{im}(f) = f([a,b])$,
- $f(b) \leq f(a) \quad \Rightarrow \quad [f(b), f(a)] \subseteq \operatorname{im}(f) = f([a,b])$.



Beweis. Wir zeigen nur den Fall $f(a) \leq f(b)$. (Der Fall $f(b) \leq f(a)$ folgt analog, bzw. durch anwenden der gezeigten Aussage auf die Funktion $-f$.).

Zunächst dürfen wir annehmen, dass $f(a) < f(b)$ gilt; denn die Aussage ist ja klar für $f(a) = f(b)$. Es reicht dann zu zeigen, dass für ein vorgegebenes $z \in (f(a), f(b))$ ein $x \in [a,b]$ mit $f(x) = z$ existiert (für $z \in \{f(a), f(b)\}$ ist die Aussage evident). Hierfür betrachten wir die (durch b nach oben beschränkte) Teilmenge

$$[a,b] \supseteq M := \{y \in [a,b] \mid f(y) \leq z\} \leq b.$$

Dann gilt $a \in M \neq \emptyset$ wegen $f(a) < z \in (f(a), f(b))$; also existiert $a \leq x := \sup_{\mathbb{R}}(M) \leq b$. Wir zeigen nun, dass $f(x) = z$ gilt:

Wegen der Stetigkeit von f in x, existiert für jedes $\varepsilon > 0$ ein $\delta > 0$ mit

$$|x - y| < \delta \quad \text{für} \quad y \in [a,b] \qquad \Longrightarrow \qquad |f(x) - f(y)| < \varepsilon. \tag{216}$$

- Es gilt $f(x) \leq z$: Angenommen, es gilt $z < f(x)$, d.h., $\varepsilon := \frac{f(x)-z}{2} > 0$.[36]
  - Wir wählen $\delta > 0$ wie in (216). Gemäß Lemma 20, existiert ein $y_\delta \in M$ mit $y_\delta > x - \delta$.
  - Wegen $y_\delta \in M \leq x = \sup_{\mathbb{R}}(M)$, folgt $0 \leq x - y_\delta < \delta$, d.h.

$$|x - y_\delta| < \delta \quad \text{mit} \quad y_\delta \in M \subseteq [a,b] \qquad \overset{(216)}{\Longrightarrow} \qquad |f(x) - f(y_\delta)| < \varepsilon.$$

  - Hiermit gilt $f(y_\delta) \in (f(x) - \varepsilon, f(x) + \varepsilon)$, und somit

$$f(y_\delta) > f(x) - \varepsilon = f(x) - \tfrac{f(x)-z}{2} = z + \tfrac{f(x)-z}{2} = z + \varepsilon > z.$$

  Wegen $y_\delta \in M$ (also $f(y_\delta) \leq z$), widerspricht dies der Definition von M.

- Es gilt $z \leq f(x)$: In der Tat, angenommen es gilt $f(x) < z$, also $\varepsilon := \frac{z-f(x)}{2} > 0$.[37]

[36] Wir konstruieren nun einen Widerspruch, indem wir die Existenz eines $\delta_y \in M$ mit $z < f(y_\delta)$ zeigen.
[37] Wir konstruieren nun einen Widerspruch, indem wir die Existenz eines $y \in M$ mit $y > x = \sup_{\mathbb{R}}(M)$ zeigen.

– Es gilt $x < b$, wegen $f(x) < z \in (f(a), f(b))$.
– Wir setzen $\delta' := b - x > 0$, und wählen $\delta > 0$ mit $\delta < \delta'$, sodass die Implikation (216) gilt.
– Wir setzen $y := x + \frac{\delta}{2}$, und erhalten

$$a \leq x < y = x + \frac{\delta}{2} \leq x + \frac{\delta'}{2} \leq b \qquad \text{also} \qquad x < y \in [a, b].$$

Wegen $|x - y| = \frac{\delta}{2} < \delta$ und $y \in [a, b]$, zeigt (216) $|f(x) - f(y)| < \varepsilon$. Hiermit folgt

$$f(y) < f(x) + \varepsilon = f(x) + \tfrac{z - f(x)}{2} = z - \tfrac{z - f(x)}{2} = z - \varepsilon < z.$$

Wir erhalten somit den Widerspruch $M \leq x < y \in M$. □

Wie die folgenden beiden Beispiele zeigen, kann der Zwischenwertsatz dazu verwendet werden, die Existenz von Nullstellen gewisser Abbildungen zu zeigen.

Beispiel 59. *Der Zwischenwertsatz zeigt, dass die stetige Funktion*

$$f \colon \mathbb{R} \supseteq [0, 2] \to \mathbb{R}, \qquad x \mapsto x^2 - 2$$

eine Nullstelle in $[0, 2]$ *hat. In der Tat gilt ja* $f(0) = -2$ *sowie* $f(2) = 2$*; und wegen* $0 \in [-2, 2] = [f(0), f(2)]$ *zeigt dann Satz 36 die Existenz eines* $x \in [0, 2]$ *mit* $f(x) = 0$*. Es gilt natürlich* $x = \sqrt{2}$*.*

Wegen $\sqrt{2} \notin \mathbb{Q}$ *zeigt dieses Beispiel insbesondere, dass der Zwischenwertsatz in obiger Form nicht gelten kann, wenn man* $\mathbb{R}$ *durch* $\mathbb{Q}$ *ersetzt. Dieser Umstand kann nicht verwundern, da wir ja im Beweis von Satz 36 expliziten Gebrauch der Existenz von Suprema von nach oben beschränkten (nicht leereren) Teilmengen gemacht haben.*

Beispiel 60 (Nullstellen ungerader Polynome)**.** *Jede Polynomfunktion* $\mathsf{Pol}(\mathbb{R}) \ni P \colon x \to \sum_{j=0}^{n} a_j \cdot x^j \in \mathbb{R}$*, ungeraden Grades* $n \in \mathbb{N}$ *(d.h.,* $n \geq 1$*) hat mindestens eine reelle Nullstelle.*

Beweis der Behauptung: Es hat P genau dann eine reelle Nullstelle, wenn die Polynomfunktion $a_n^{-1} \cdot P$ eine reelle Nullstelle hat. Wir dürfen daher annehmen, dass $a_n = 1$ gilt. Wir erhalten

$$P(x) = x^n \cdot \underbrace{\left(1 + \frac{a_{n-1}}{x} + \ldots + \frac{a_0}{x^n}\right)}_{=:\,\eta(x)} \qquad \forall\, x \neq 0.$$

Wir wählen $c > 0$ mit $\kappa := |\frac{a_{n-1}}{\pm c}| + \cdots + |\frac{a_0}{(\pm c)^n}| < 1$, also $\eta(\pm c) - 1 > -\kappa > -1 \;\Rightarrow\; \eta(\pm c) > 0$. Da n ungerade ist, gilt

$$P(-c) = -c^n \cdot \eta(-c) < 0 \qquad \text{sowie} \qquad P(c) = c^n \cdot \eta(c) > 0.$$

Die Einschränkung $f := P|_{[-c,c]}$ ist stetig (Beispiel 52.a) und Lemma 40.2)), mit $f(-c) < 0$ und $f(c) > 0$. Der Zwischenwertsatz (Satz 36) zeigt daher die Existenz eines $x \in [-c, c]$ mit $P(x) = f(x) = 0 \in [f(-c), f(c)]$. □

Satz 36 impliziert die folgende Aussage:

Korollar 28. *Sei* $D \subseteq \mathbb{R}$ *ein Intervall, und* $f \colon D \to \mathbb{R}$ *stetig. Dann ist* $f(D)$ *ebenfalls ein Intervall.*

Beweis. Ist $D = \emptyset$, so ist $f(D) = \emptyset$ ein Intervall. Sei also $D \neq \emptyset$, und setzte $\hat{D} := f(D)$. Um nachzuweisen, dass $\hat{D}$ ein Intervall ist, müssen wir für je zwei vorgegebene Element $\hat{D} \ni x_1 \leq x_2 \in \hat{D}$, die folgende Implikation nachweisen:

$$x_1 \leq y \leq x_2 \quad \text{für} \quad y \in \mathbb{R} \qquad \Longrightarrow \qquad y \in \hat{D}. \tag{217}$$

Wir wählen $c_1, c_2 \in D$ mit $f(c_1) = x_1$ und $f(c_2) = x_2$.

- Gilt $c_1 \leq c_2$, so setzen wir $a := c_1$ und $b := c_2$.
- Gilt $c_1 > c_2$, so setzen wir $a := c_2$ und $b := c_1$.

Dann gilt $a \leq b$, sowie $[a, b] \subseteq D$ wegen der Intervalleigenschaft von D. Weiterhin ist $\tilde{f} := f|_{[a,b]} : [a, b] \to \mathbb{R}$ stetig wegen Lemma 40.2). Gilt nun die linke Seite von (217), so zeigt Satz 36:

- Im Falle $c_1 \leq c_2$: $\quad y \in [x_1, x_2] = [\tilde{f}(a), \tilde{f}(b)] \subseteq \mathrm{im}(\tilde{f}) \subseteq \mathrm{im}(f) = \hat{D}$
- Im Falle $c_1 > c_2$: $\quad y \in [x_1, x_2] = [\tilde{f}(b), \tilde{f}(a)] \subseteq \mathrm{im}(\tilde{f}) \subseteq \mathrm{im}(f) = \hat{D}$ $\square$

Beispiel 61 (Existenz k-ter Wurzeln). *Sei $k \in \mathbb{N}_{>0}$ und $f : [0, \infty) \ni z \mapsto z^k \in [0, \infty)$.*

- *f ist stetig wegen Beispiel 52.a). Gemäß Korollar 28, ist somit $D := \mathrm{im}[f] \subseteq \mathbb{R}$ ein Intervall.*

  Nun gilt $0 = f(0) \in D$. Weiterhin ist D nach oben unbeschränkt; denn für $c \in (0, \infty)$ vorgegeben, gilt $c \leq f(\max(1, c)) \in D$. Da D ein Intervall ist, folgt $D = [0, \infty)$. Somit ist f surjektiv.
- *f ist injektiv wegen Übung 45.a). Daher gilt für die Umkehrabbildung*

$$f^{-1} : [0, \infty) \to [0, \infty), \qquad x \mapsto f^{-1}(x) = \sqrt[k]{x}.$$

Dies liefert einen alternativen Beweis für die Existenz k-ter Wurzeln in $[0, \infty) \subseteq \mathbb{R}$.

Definition 54. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{Q}\}$, sowie $M, N \subseteq \mathbb{K}$. Eine Funktion $f : M \to N$ heißt*

1) monoton wachsend*, wenn gilt* $\quad x \leq y \;\Rightarrow\; f(x) \leq f(y)$
2) streng monoton wachsend*, wenn gilt* $\quad x < y \;\Rightarrow\; f(x) < f(y)$
3) monoton fallend*, wenn gilt* $\quad x \leq y \;\Rightarrow\; f(x) \geq f(y)$
4) streng monoton fallend*, wenn gilt* $\quad x < y \;\Rightarrow\; f(x) > f(y)$

jeweils für alle $x, y \in M$.

Die Funktion f heißt monoton*, wenn sie monoton wachsend oder monoton fallend ist; sowie* streng monoton*, wenn sie streng monoton wachsend oder streng monoton fallend ist.*

Beispiel 62. *Die in 0 unstetige Funktion*

$$f : (-1, 1] \to \mathbb{R}, \qquad x \mapsto \begin{cases} x & \text{für} \quad 0 \leq x \leq 1 \\ -1 - x & \text{für} \quad -1 < x < 0 \end{cases}$$

ist injektiv aber nicht streng monoton. Das folgende Lemma zeigt, dass so etwas für stetige Funktionen auf Intervallen nicht vorkommt.

Lemma 43. *Sei $\emptyset \neq D \subseteq \mathbb{R}$ ein Intervall. Eine stetige Funktion $f : D \to \mathbb{R}$ ist genau dann injektiv, wenn sie streng monoton ist.*

Beweis. • Ist f streng monoton und $D \ni x \neq y \in D$, so gilt $f(x) < f(y)$ oder $f(x) > f(y)$. In jedem Fall gilt $f(x) \neq f(y)$; also ist f injektiv.

- Sei f injektiv. Ist f nicht streng monoton, so existieren $x_1, x_2, y_1, y_2 \in D$ mit

$$x_1 < y_1 \quad \wedge \quad f(x_1) < f(y_1) \qquad \text{sowie} \qquad x_2 < y_2 \quad \wedge \quad f(x_2) > f(y_2). \tag{218}$$

(Wegen der Injektivität von f, ist dies die Negation von:

$$\forall\, D \ni x_1 < y_1 \in D\colon\; f(x_1) > f(y_1) \qquad \vee \qquad \forall\, D \ni x_2 < y_2 \in D\colon\; f(x_2) < f(y_2) \quad)$$

Als Zusammensetzung stetiger Abbildungen, ist nun

$$g\colon [0,1] \to \mathbb{R}, \qquad t \mapsto f(t \cdot x_1 + (1-t) \cdot x_2) - f(t \cdot y_1 + (1-t) \cdot y_2)$$

stetig; und es gilt $g(0) = f(x_2) - f(y_2) > 0$ sowie $g(1) = f(x_1) - f(y_1) < 0$. Gemäß Satz 36, existiert somit ein $\tau \in (0,1)$ mit $g(\tau) = 0$, d.h.,

$$f(\underbrace{\tau \cdot x_1 + (1-\tau) \cdot x_2}_{=:\, x}) = f(\underbrace{\tau \cdot y_1 + (1-\tau) \cdot y_2}_{=:\, y}) \qquad \overset{f \text{ injektiv}}{\Longrightarrow} \qquad x = y.$$

Wegen $0 < \tau < 1$, gilt nun aber $\tau > 0$ sowie $(1-\tau) > 0$. Daher zeigt (218), dass

$$x = \tau \cdot x_1 + (1-\tau) \cdot x_2 < \tau \cdot y_1 + (1-\tau) \cdot y_2 = y$$

gilt, im Widerspruch zu $x = y$. Somit war die Annahme falsch, dass f nicht streng monoton ist. □

Satz 37 (Satz über die Umkehrfunktion)**.** *Sei $\emptyset \neq D \subseteq \mathbb{R}$ ein Intervall, sowie $f\colon D \to \mathbb{R}$ stetig und injektiv. Dann ist $\hat{D} := f(D) \subseteq \mathbb{R}$ ein Intervall, und f ist streng monoton.*

Weiterhin besitzt $f|^{\hat{D}}$ eine stetige und streng monotone Umkehrfunktion $(f|^{\hat{D}})^{-1}\colon \hat{D} \to D \subseteq \mathbb{R}$. (Ist f streng monoton wachsend/fallend, so ist $(f|^{\hat{D}})^{-1}$ streng monoton wachsend/fallend.)

Beweis. Wegen Korollar 28 ist $\hat{D}$ ein Intervall; und wegen Lemma 43 ist f streng monoton. Dies zeigt den ersten Teil. Wir zeigen den zweiten Teil nur den Fall, dass f streng monoton wachsend ist (der andere Fall folgt analog):

Wir betrachten nun f als Abbildung $D \to \hat{D}$ (wir ersetzen also f durch die Koeinschränkung $f|^{\hat{D}}$). Dann ist f surjektiv und injektiv (vgl. Übung 30), sodass $f^{-1}\colon \hat{D} \to D$ existiert).

- f^{-1} ist streng monoton wachsend:

  Gegeben $x, y \in \hat{D}$ mit $x < y$, so existieren $a, b \in D$ mit $f(a) = x < y = f(b)$. Da f streng monoton wachsend ist, gilt notwendig $a < b$, d.h., $f^{-1}(x) = a < b = f^{-1}(y)$.
- f^{-1} ist stetig:

  Seien $x \in \hat{D}$ und $\varepsilon > 0$ vorgegeben. Wir zeigen, dass ein $\delta_L > 0$ existiert mit

  $$x - y < \delta_L \quad \text{für} \quad \hat{D} \ni y < x \qquad \Longrightarrow \qquad |f^{-1}(x) - f^{-1}(y)| < \varepsilon. \tag{219}$$

  Analog folgt dann auch die Existenz eines $\delta_R > 0$ mit

  $$y - x < \delta_R \quad \text{für} \quad x < y \in \hat{D} \qquad \Longrightarrow \qquad |f^{-1}(x) - f^{-1}(y)| < \varepsilon.$$

  Setzt man $\delta := \min(\delta_L, \delta_R)$, so folgt mit diesen beiden Abschätzungen:

  $$|x - y| < \delta \quad \text{für} \quad y \in \hat{D} \qquad \Longrightarrow \qquad |f^{-1}(x) - f^{-1}(y)| < \varepsilon,$$

  also die Stetigkeit von f^{-1} in x. Wir zeigen nun (219). Hierfür unterscheiden wir die folgenden beiden Fälle:

  – Es gilt $x = \min(\hat{D})$ ($\hat{D}$ ist von der Form $[x, \infty)$ oder $[x, z)$ oder $[x, z]$ für ein $z \in \mathbb{R}$). Dann existiert kein $y \in \hat{D}$ mit $y < x$; sodass (219) für jedes $\delta_L > 0$ erfüllt ist.
  – Es existiert ein $z \in \hat{D}$ mit $z < x$. Dann gilt $f^{-1}(z) < f^{-1}(x)$, da f^{-1} streng monoton wachsend ist. Wir fixieren $\max(f^{-1}(z), f^{-1}(x) - \varepsilon) < \tau < f^{-1}(x)$. Dann gilt $\tau \in D$, da D ein Intervall ist.

Da f streng monoton wachsend ist, gilt somit $f(\tau) < x$; und da f^{-1} streng monoton wachsend ist, erhalten wir:

$$\begin{aligned} f(\tau) < y < x \quad \text{für} \quad y \in \hat{D} \quad &\overset{f^{-1}(\cdot)}{\Longrightarrow} & \tau < f^{-1}(y) < f^{-1}(x) \\ &\Longrightarrow & f^{-1}(x) - \varepsilon < \tau < f^{-1}(y) < f^{-1}(x) \\ &\Longrightarrow & |f^{-1}(y) - f^{-1}(x)| < \varepsilon. \end{aligned} \tag{220}$$

Für $\hat{D} \ni y < x$ mit $x - y < \delta_L := x - f(\tau) > 0$, gilt nun aber die linke Seite von (220); sodass also (219) unmittelbar aus (220) folgt. □

Beispiel 63 (Stetigkeit der Wurzelfunktion)**.**

a) *Für $\frac{p}{q} \in \mathbb{Q}$ mit $p \in \mathbb{N}$ und $q \in \mathbb{N}_{>0}$, ist die folgende Abbildung stetig (vgl.* (97) *in Definition 28):*

$$\alpha[p,q]\colon [0,\infty) \to [0,\infty), \qquad x \mapsto x^{\frac{p}{q}} = (\sqrt[q]{x})^p$$

Beweis der Behauptung: Zunächst ist $\alpha[n,1]\colon [0,\infty) \ni x \mapsto x^n \in [0,\infty)$ für alle $n \in \mathbb{N}$ stetig wegen Beispiel 52.a); sowie bijektiv für $n \in \mathbb{N}_{>0}$ nach Beispiel 61. Somit ist auch $\alpha[1,n] = \alpha[n,1]^{-1}$ für alle $n \in \mathbb{N}_{>0}$ stetig, nach Satz 37. Dann ist $\alpha[p,q] = \alpha[p,1] \circ \alpha[1,q]$ für $p \in \mathbb{N}$ und $q \in \mathbb{N}_{>0}$ stetig nach Lemma 39. □

b) *Für $\frac{p}{q} \in \mathbb{Q}$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}_{>0}$ ist die folgende Abbildung ebenfalls stetig:*

$$\beta[p,q]\colon (0,\infty) \to (0,\infty), \qquad x \mapsto x^{\frac{p}{q}} = (\sqrt[q]{x})^p.$$

Beweis der Behauptung: Für $p \geq 0$, ist $\beta[p,q] = \alpha[p,q]|_{(0,\infty)}$ als Einschränkung einer stetigen Funktion stetig. Für $p \leq -1$, ist dann auch $\beta[p,q] = 1/\beta[|p|,q]$ stetig gemäß Korollar 25.2). □

c) *Ist $n \in \mathbb{N}_{>0}$ ungerade, so ist $\gamma[n]\colon \mathbb{R} \ni x \mapsto x^n \in \mathbb{R}$ stetig und bijektiv. Daher ist $\gamma[n]^{-1}$ ebenfalls stetig; und wir setzen $\sqrt[n]{x} := \gamma[n]^{-1}(x)$ für alle $x \in \mathbb{R}$ ($n \in \mathbb{N}_{>0}$ ungerade).*

d) *Seien $\mathbb{R} \ni a < b \in \mathbb{R}$. Wir konstruieren eine stetige bijektive Funktion $f\colon (a,b) \to \mathbb{R}$, die streng monoton wachsend ist:*

   Wir betrachten die stetigen und streng monoton wachsenden Abbildungen

$$\begin{aligned} f_1\colon (a,b) &\to (\tfrac{1}{b-a},\infty), & x &\mapsto \tfrac{1}{b-x} \\ f_2\colon (a,b) &\to (-\infty,-\tfrac{1}{b-a}), & x &\mapsto \tfrac{1}{a-x}. \end{aligned}$$

   Dann ist $f := f_1 + f_2$ ebenfalls stetig und streng monoton wachsenden, also injektiv wegen Lemma 43. Weiterhin ist $\operatorname{im}(f)$ *ein Intervall wegen Korollar 28, und es gilt*

$$\lim_{x\downarrow a} f(x) = -\infty \qquad \textit{sowie} \qquad \lim_{x\uparrow b} f(x) = \infty.$$

   Hiermit folgt $\operatorname{im}(f) = \mathbb{R}$, *also die Surjektivität von f.*

7.3 Folgen und Reihen von Funktionen

In diesem Abschnitt behandeln wir Konvergenzeigenschaften von Folgen von Funktionen.

7.3.1 Allgemeine Konvergenzaussagen

Definition 55. *Sei $Z \neq \emptyset$ eine Menge, (X, d) ein metrischer Raum, $(f_n)_{n\in\mathsf{N}}$ eine Folgen in $\mathsf{Abb}(Z, X)$ (Funktionenfolge), und $f \in \mathsf{Abb}(Z, X)$.*

1) *Die Funktionenfolge $(f_n)_{n\in\mathsf{N}}$* konvergiert punktweise *gegen f, wenn gilt:*

$$\begin{gathered}(f_n(z))_{n\in\mathsf{N}} \to f(z) \qquad \forall\, z \in Z \\ \Big(\ \forall\,(z \in Z \ \wedge\ \varepsilon > 0)\colon \quad \exists N_{z,\varepsilon} \in \mathsf{N}\colon\ \forall\, n \geq N_{z,\varepsilon}\colon \quad \mathrm{d}(f(z), f_n(z)) < \varepsilon\ \Big)\end{gathered} \tag{221}$$

2) *Die Funktionenfolge $(f_n)_{n\in\mathsf{N}}$* konvergiert gleichmäßig *gegen f, wenn gilt:*

$$\forall\, \varepsilon > 0\colon\ \exists N_\varepsilon \in \mathsf{N}\colon\ \forall\, n \geq N_\varepsilon\colon \qquad \mathrm{d}(f(z), f_n(z)) < \varepsilon \quad \forall\, z \in Z. \tag{222}$$

Offensichtlich impliziert gleichmäßige Konvergenz auch punktweise Konvergenz. Warum man sich nicht mit dem Konzept der punktweisen Konvergenz zufrieden gibt, sondern gleichmäßige Konvergenz einführt, klären Bemerkung 54 und Proposition 12.

Bemerkung 54. *Der punktweise Grenzwert stetiger Funktionen ist nicht notwendigerweise ebenfalls stetig. Sei beispielsweise $C([0,1], \mathbb{R}) \ni f_n\colon [0,1] \ni x \mapsto x^n \in \mathbb{R}$ für alle $n \in \mathbb{N}$, und*

$$f\colon [0,1] \to \mathbb{R}, \qquad x \mapsto \begin{cases} 1 & \textit{für} \quad x = 1 \\ 0 & \textit{für} \quad x \in [0,1). \end{cases}$$

Offensichtlich ist f unstetig in 1, und wegen Proposition 5 (geometrische Folge) gilt

$$\lim_{n\to\infty} f_n(x) = \lim_n x^n = f(x) \qquad \forall\, x \in [0,1].$$

Die Folge stetiger Funktion $(f_n)_{n\in\mathbb{N}}$, konvergiert also punktweise gegen die unstetige Funktion f.

Die nächste Proposition zeigt, dass der in Bemerkung 54 angesprochene Sachverhalt bei gleichmäßiger Konvergenz nicht auftritt:

Proposition 12. *Seien (X_1, d_1), (X_2, d_2) metrische Räume, $(f_n)_{n\in\mathsf{N}}$ eine Folge in $C(X_1, X_2)$, und $f \in \mathsf{Abb}(X_1, X_2)$. Konvergiert $(f_n)_{n\in\mathsf{N}}$ gleichmäßig gegen f, so gilt $f \in C(X_1, X_2)$.*

(Gleichmäßige Grenzwerte stetiger Abbildungen sind stetig.)

Beweis. Seien $x \in X_1$ und $\varepsilon > 0$ vorgegeben.

- Wegen der gleichmäßigen Konvergenz, existiert ein $N \in \mathsf{N}$ mit $\mathrm{d}_2(f(z), f_N(z)) < \frac{\varepsilon}{3}$ für alle $z \in X_1$.
- Da f_N stetig ist, existiert ein $\delta > 0$ mit $\mathrm{d}_2(f_N(x), f_N(y)) < \frac{\varepsilon}{3}$ für alle $y \in \mathsf{B}[\mathrm{d}_1]_\delta(x)$.
- Für $y \in \mathsf{B}[\mathrm{d}_1]_\delta(x)$ folgt mit zweifacher Anwendung der Dreiecksungleichung M3) (sowie M2)), dass

$$\begin{aligned}\mathrm{d}_2(f(x), f(y)) &\leq \mathrm{d}_2(f(x), f_N(x)) + \mathrm{d}_2(f_N(x), f(y)) \\ &\leq \mathrm{d}_2(f(x), f_N(x)) + \mathrm{d}_2(f_N(x), f_N(y)) + \mathrm{d}_2(f_N(y), f(y)) < 3 \cdot \frac{\varepsilon}{3} = \varepsilon\end{aligned}$$

  für alle $y \in \mathsf{B}[\mathrm{d}_1]_\delta(x)$ gilt, d.h., $f(\mathsf{B}[\mathrm{d}_1]_\delta(x)) \subseteq \mathsf{B}[\mathrm{d}_2]_\varepsilon(f(x))$.

Dies zeigt die Stetigkeit von f in x, und somit die Behauptung. □

Definition 56. *Sei $Z \neq \emptyset$ eine Menge, und $(V, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum. Wir definieren*

$$\|f\|_\infty := \sup_{\overline{\mathbb{R}}}(\{\|f(z)\| \mid z \in Z\}) \in [0, \infty] \qquad \forall\, f \in \mathsf{Abb}(Z, V). \tag{223}$$

Für $f \in \mathsf{Abb}(Z, V)$, gilt dann die Äquivalenz:

$$\|f\|_\infty < \infty \qquad \Longleftrightarrow \qquad \exists\, C \in [0, \infty)\colon\ \|f(z)\| \leq C \quad \forall\, z \in Z. \tag{224}$$

(Gilt nämlich die linke Seite von (224), so gilt die rechte Seite für $C = \|f\|_\infty$. Gilt umgekehrt die rechte Seite von (224), so gilt $\|f\|_\infty \leq C < \infty$.)

Terminologie 23. *Sei $Z \neq \emptyset$ eine Menge, und $(V, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum.*

1) *Der Raum der* beschränkten Abbildungen *auf Z mit Werten in V, ist definiert durch*

$$\begin{aligned} \mathrm{B}(Z, V) \ &:= \ \{f \in \mathsf{Abb}(Z, V) \mid \exists\, C \geq 0\colon\ \|f(z)\| \leq C \text{ für alle } z \in Z\} \\ &\overset{(224)}{=} \{f \in \mathsf{Abb}(Z, V) \mid \|f\|_\infty < \infty\}. \end{aligned}$$

   Es ist $\mathrm{B}(Z, V)$ ein $\mathbb{K}$-Vektorraum, vermöge den Operationen in der ersten, dritten, und vierten Zeile in (195), mit der Nullabbildung $f[0]\colon Z \ni z \mapsto 0 \in V$ als additiven neutralem Element. Zudem ist $\|\cdot\|_\infty$ eine Norm auf $\mathrm{B}(Z, V)$, die sogenannte Supremumsnorm.

   Beweis der Behauptungen: Seien $f, g \in \mathrm{B}(Z, V)$ und $\lambda \in \mathbb{K}$. Dann gilt (beachte $-f = (-1) \cdot f$)

$$\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty, \qquad \|\lambda \cdot f\|_\infty = |\lambda| \cdot \|f\|_\infty, \qquad \|f\|_\infty = 0 \ \Leftrightarrow\ f = f[0], \tag{225}$$

   denn für alle $z \in Z$ erhalten wir

$$\begin{aligned} \|(f + g)(z)\| = \|f(z) + g(z)\| \leq \|f(z)\| + \|g(z)\| &\leq \|f\|_\infty + \|g\|_\infty \\ \|(\lambda \cdot f)(z)\| = \|\lambda \cdot f(z)\| &= |\lambda| \cdot \|f(z)\| \\ 0 \leq \|f(z)\| &\leq \|f\|_\infty, \end{aligned}$$

   womit (225) problemlos folgt (Proposition 1.3) für V2)). □

2) *Sei (Z, d) ein metrischer Raum. Der Raum aller* beschränkten stetigen Abbildungen von Z nach V*, ist definiert durch*

$$C_b(Z, V) := C(Z, V) \cap \mathrm{B}(Z, V).$$

   Wegen Korollar 25, ist $C_b(Z, V)$ ebenfalls ein $\mathbb{K}$-Vektorraum (Untervektorraum von $\mathrm{B}(Z, V)$), und $\|\cdot\|_\infty$ ist eine Norm auf $C_b(Z, V)$.

In beiden Fällen, liest sich (222) (gleichmäßige Konvergenz) in der Form:

$$\forall\, \varepsilon > 0\colon\ \exists N_\varepsilon \in \mathsf{N}\colon\ \forall\, n \geq N_\varepsilon\colon \qquad \|f(z) - f_n(z)\| < \varepsilon \quad \forall\, z \in Z. \tag{226}$$

Offensichtlich[38] *ist diese Bedingung äquivalent zu der folgenden Bedingung:*

$$\forall\, \varepsilon > 0\colon\ \exists N_\varepsilon \in \mathsf{N}\colon\ \forall\, n \geq N_\varepsilon\colon \qquad \|f - f_n\|_\infty < \varepsilon. \tag{227}$$

Somit gilt in den entsprechenden Fällen:

[38] Die Implikation (227) $\Rightarrow$ (226) ist unmittelbar klar. Für die umgekehrten Implikation, ersetze man auf der echten Seite von (226) „$< \varepsilon$“ durch „$< \frac{\varepsilon}{2}$“, und bilde dann das Supremum.

(A) *Eine Folge* $(f_n)_{n\in\mathsf{N}}$ *in* $\mathrm{B}(Z,V)$ *konvergiert gleichmäßig gegen ein* $f\in\mathrm{B}(Z,V)$ *genau dann, wenn* $(f_n)_{n\in\mathsf{N}}$ *im normierten Raum* $(\mathrm{B}(Z,V),\|\cdot\|_\infty)$ *gegen* f *konvergiert.*

(B) *Eine Folge* $(f_n)_{n\in\mathsf{N}}$ *in* $C_b(Z,V)$ *konvergiert gleichmäßig gegen ein* $f\in C_b(Z,V)$ *genau dann, wenn* $(f_n)_{n\in\mathsf{N}}$ *im normierten Raum* $(C_b(Z,V),\|\cdot\|_\infty)$ *gegen* f *konvergiert.*

Lemma 44. *Sei* $Z\neq\emptyset$ *eine Menge,* $(V,\|\cdot\|)$ *ein normierter Raum, und* $(f_n)_{n\in\mathsf{N}}$ *eine Folge in* $\mathrm{B}(Z,V)$*, die gleichmäßig gegen ein* $f\in\mathsf{Abb}(Z,V)$ *konvergiert.*

1) Es gilt $f\in\mathrm{B}(Z,V)$.

($(f_n)_{n\in\mathsf{N}}$ *konvergiert im normierten Raum* $(\mathrm{B}(Z,V),\|\cdot\|_\infty)$ *gegen* $f\in\mathrm{B}(Z,V)$*, vgl.* (A).*)*

2) Ist (Z,d) *ein metrischer Raum, und gilt* $f_n\in C_b(Z,V)$ *für alle* $n\in\mathsf{N}$*, so gilt auch* $f\in C_b(Z,V)$.

($(f_n)_{n\in\mathsf{N}}$ *konvergiert im normierten Raum* $(C_b(Z,V),\|\cdot\|_\infty)$ *gegen* $f\in C_b(Z,V)$*, vgl.* (B).*)*

Beweis. 1) Per Annahme existiert ein $N\in\mathsf{N}$ mit $\|f(z)-f_N(z)\|<1$ für alle $z\in Z$. Es folgt

$$\|f(z)\|\leq\|f(z)-f_N(z)\|+\|f_N(z)\|<1+\|f_N\|_\infty\qquad\forall\,z\in Z,$$

und daher $\|f\|_\infty\leq 1+\|f_N\|_\infty<\infty$.

2) Teil 1) zeigt $f\in\mathrm{B}(Z,V)$, und Proposition 12 zeigt $f\in C(Z,V)$. Daher gilt $f\in C(Z,V)\cap\mathrm{B}(Z,V)=C_b(Z,V)$. □

Bemerkung 55 (Cauchy-Folge)**.** *Sei* $Z\neq\emptyset$ *eine Menge, und* $(V,\|\cdot\|)$ *ein normierter Raum. Gemäß Definition 40 (vgl.* (164)*) gilt Folgendes:*

- *Eine Folge* $(f_n)_{n\in\mathsf{N}}$ *im normierten Raum* $(\mathrm{B}(Z,V),\|\cdot\|_\infty)$ *heißt Cauhy-Folge, wenn gilt:*

$$\forall\,\varepsilon>0\colon\ \exists N_\varepsilon\in\mathsf{N}\colon\ \forall\,m,n\geq N_\varepsilon\colon\qquad\|f_m-f_n\|_\infty<\varepsilon.\tag{228}$$

- *Sei* (Z,d) *ein metrischer Raum. Eine Folge* $(f_n)_{n\in\mathsf{N}}$ *im normierten Raum* $(C_b(Z,V),\|\cdot\|_\infty)$ *heißt Cauchy-Folge, wenn* (228) *gilt. Wegen* $C_b(Z,V)\subseteq\mathrm{B}(Z,V)$*, ist somit jede Cauchy-Folge in* $(C_b(Z,V),\|\cdot\|_\infty)$ *automatisch auch eine Cauchy-Folge in* $(\mathrm{B}(Z,V),\|\cdot\|_\infty)$.

Ist $(V,\|\cdot\|)$ ein Banachraum (jede Cauchy-Folge in $(V,\|\cdot\|)$ konvergiert in V), so erhalten wir das folgende Resultat:

Lemma 45. *Sei* $Z\neq\emptyset$ *eine Menge, und* $(V,\|\cdot\|)$ *ein Banachraum. Ist* $(f_n)_{n\in\mathsf{N}}$ *eine Cauchy-Folge im normierten Raum* $(\mathrm{B}(Z,V),\|\cdot\|_\infty)$*, so konvergiert* $(f_n)_{n\in\mathsf{N}}$ *gleichmäßig gegen*

$$f\colon Z\to V,\qquad z\mapsto\lim_n f_n(z).$$

Beweis. Es gilt (228). Für jedes $z\in Z$, ist dann $(f_n(z))_{n\in\mathsf{N}}$ eine Cauchy-Folge in V; und zwar wegen

$$\|f_n(z)-f_m(z)\|\leq\|f_n-f_m\|_\infty\qquad\forall\,z\in Z,\ m,n\in\mathsf{N}.$$

Da V vollständig ist, konvergiert somit $(f_n(z))_{n\in\mathsf{N}}$ für jedes $z\in Z$ gegen ein Element in V, welches wir mit $f(z)$ bezeichnen. Dies definiert die Abbildung $f\colon Z\to V$. Wir müssen nun noch nachweisen, dass $(f_n)_{n\in\mathsf{N}}$ gleichmäßig gegen f konvergiert. Sei hierfür $\varepsilon>0$ vorgegeben:

- Wegen (228), existiert ein $N\in\mathbb{N}$ mit

$$\|f_m-f_n\|_\infty<\tfrac{\varepsilon}{2}\qquad\forall\,m,n\geq N.$$

- Für jedes $z\in Z$, existiert wegen $\lim_n f_n(z)=f(z)$, ein $N_z\geq N$ mit $\|f(z)-f_{N_z}(z)\|<\frac{\varepsilon}{2}$.

Wir erhalten für $n \geq N$, dass

$$\|f(z) - f_n(z)\| \leq \|f(z) - f_{N_z}(z)\| + \overbrace{\|f_{N_z}(z) - f_n(z)\|}^{\leq \|f_{N_z} - f_n\|_\infty} < 2 \cdot \frac{\varepsilon}{2} = \varepsilon \qquad \forall\, z \in Z$$

gilt, was die Behauptung zeigt. □

Wir erhalten schließlich den folgenden Satz:

Satz 38. *Sei $(V, \|\cdot\|)$ ein Banachraum.*

1) Ist $Z \neq \emptyset$ eine Menge, so ist $(\mathrm{B}(Z,V), \|\cdot\|_\infty)$ ein Banachraum.

2) Ist (Z, d) ein metrischer Raum, so ist $(C_b(Z,V), \|\cdot\|_\infty)$ ein Banachraum.

In beiden Fällen ist die Grenzfunktion zu einer entsprechenden Cauchy-Folge, gegeben durch

$$f\colon Z \to V, \qquad z \mapsto \lim_n f_n(z).$$

Beweis. 1) Sei $(f_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge in $(\mathrm{B}(Z,V), \|\cdot\|_\infty)$. Wegen Lemma 45 konvergiert $(f_n)_{n\in\mathsf{N}}$ gleichmäßig gegen f. Somit zeigt Lemma 44.1) die Behauptung.

2) Ist $(f_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge in $(C_b(Z,V), \|\cdot\|_\infty)$, so auch in $(\mathrm{B}(Z,V), \|\cdot\|_\infty)$. Wegen Lemma 45 konvergiert $(f_n)_{n\in\mathsf{N}}$ gleichmäßig gegen f. Somit zeigt Lemma 44.2) die Behauptung. □

Korollar 29 (Konvergenzsatz von Weierstraß)**.** *Sei $Z \neq \emptyset$ eine Menge, $(V, \|\cdot\|)$ ein Banachraum, sowie $\sum_{n=\mathsf{p}}^\infty f_n$ eine Reihe in $\mathrm{B}(Z,V)$ mit $\sum_{k=\mathsf{p}}^\infty \|f_n\|_\infty < \infty$.*

1) $\sum_{n=\mathsf{p}}^\infty f_n$ konvergiert gleichmäßig (bzgl. $\|\cdot\|_\infty$) gegen

$$\mathrm{B}(Z,V) \ni f\colon Z \to V, \qquad z \mapsto \textstyle\sum_{k=\mathsf{p}}^\infty f_n(z).$$

2) Ist (Z, d) ein metrischer Raum mit $f_n \in C_b(Z,V)$ für alle $n \in \mathsf{N}$, so gilt $f \in C_b(Z,V)$.

Beweis. Setze $\tilde{f}_n := \sum_{k=\mathsf{p}}^n f_k$ für alle $n \in \mathsf{N}$. Wegen Satz 38, reicht zu zeigen, dass $(\tilde{f}_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge in $(\mathrm{B}(Z,V), \|\cdot\|_\infty)$ ist.[39] Sei hierfür $\varepsilon > 0$ vorgegeben:

- Da $\sum_{k=\mathsf{p}}^\infty \|f_n\| < \infty$ nach Annahme konvergiert, ist $(\sum_{k=\mathsf{p}}^n \|f_k\|_\infty)_{n\in\mathsf{N}}$ gemäß Lemma 35.1) eine Cauchy-Folge in $\mathbb{R}$, d.h., es existiert ein $N_\varepsilon \in \mathsf{N}$ mit (vgl. Bemerkung 41)

$$\textstyle\sum_{k=m+1}^n \|f_k\|_\infty = |\sum_{k=\mathsf{p}}^n \|f_k\|_\infty - \sum_{k=\mathsf{p}}^m \|f_k\|_\infty| < \varepsilon \qquad \forall\, N_\varepsilon \leq m \leq n \leq N_\varepsilon.$$

- Die Dreiecksungleichung liefert

$$\|\tilde{f}_n - \tilde{f}_m\|_\infty = \underbrace{\|\textstyle\sum_{k=\mathsf{p}}^n f_k - \sum_{k=\mathsf{p}}^m f_k\|_\infty}_{\|\sum_{k=m+1}^n f_k\|_\infty} \leq \textstyle\sum_{k=m+1}^n \|f_k\|_\infty < \varepsilon \qquad \forall\, N_\varepsilon \leq m \leq n \leq N_\varepsilon,$$

was die Cauchy-Eigenschaft zeigt. □

[39] In der Situation von Teil 2), ist dann $(\tilde{f}_n)_{n\in\mathsf{N}}$ automatisch eine Cauchy-Folge in $(C_b(Z,V), \|\cdot\|_\infty)$.

7.3.2 Potenzreihen und die Exponentialfunktion

Wir betrachten nun den Spezialfall $Z = V = \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.[40]

Satz 39. *Sei* $\mathsf{Q}(z) = \sum_{n=0}^{\infty} c_n \cdot (z-p)^n$ *eine Potenzreihe in* $\mathbb{K}$ *im Entwicklungspunkt* $p \in \mathbb{K}$*, mit Konvergenzradius* $\mathsf{R}[\mathsf{Q}] > 0$*. Für* $n \in \mathsf{N}$*, sei*

$$g_n \colon \mathbb{K} \ni z \mapsto c_n \cdot (z-p)^n \in \mathbb{K}.$$

Dann gelten die folgenden Aussagen:

a) Sei $0 < r < \mathsf{R}[\mathsf{Q}]$*, sowie* $f_n := g_n|_{\mathsf{B}_r(p)}$ *für alle* $n \in \mathsf{N}$*. Dann gilt* $f_n \in C_b(\mathsf{B}_r(p), \mathbb{K})$ *für alle* $n \in \mathsf{N}$*. Zudem konvergiert* $\sum_{n=0}^{\infty} f_n$ *gleichmäßig gegen die stetige und beschränkte Abbildung*

$$\mathsf{Q}|_{\mathsf{B}_r(p)} \colon \mathsf{B}_r(p) \to \mathbb{K}, \qquad z \mapsto \textstyle\sum_{n=0}^{\infty} c_n \cdot (z-p)^n,$$

d.h., $\mathsf{Q}|_{\mathsf{B}_r(p)} = \sum_{n=0}^{\infty} f_n \in C_b(\mathsf{B}_r(p), \mathbb{K})$*.*

b) Die Abbildung $\mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]} \colon \mathrm{K}[\mathsf{Q}] \ni z \mapsto \mathsf{Q}(z) \in \mathbb{K}$ *ist stetig.*

Beweis. Es genügt die Aussagen für $p = 0$ zu zeigen. Der allgemeine Fall folgt dann einfach durch Verkettung mit der stetigen Abbildung $\mathbb{K} \ni z \mapsto z - p \in \mathbb{K}$.

a) Für jedes $n \in \mathsf{N}$, ist g_n eine Polynomfunktion, also stetig wegen Beispiel 52.a). Hiermit folgt nun $f_n \in C_b(\mathsf{B}_r(0), \mathbb{K})$ für jedes $n \in \mathsf{N}$:

- Wegen Lemma 40.2) ist somit $f_n = g_n|_{\mathsf{B}_r(0)}$ stetig.
- Es gilt $|f_n(z)| \leq |c_n| \cdot r^n$ für alle $z \in \mathsf{B}_r(0)$.

Wegen $0 < r < \mathsf{R}[\mathsf{Q}]$ und Satz 29.b), gilt $C := \sum_{n=0}^{\infty} |c_n| \cdot r^n < \infty$ (absolute Konvergenz von $\mathsf{Q}(r)$) also wegen Satz 18 auch $\sum_{k=0}^{n} |c_k| \cdot r^k \leq C$ für alle $n \in \mathsf{N}$. Wir erhalten

$$\textstyle\sum_{k=0}^{n} |f_k|_\infty \leq \sum_{k=0}^{n} |c_k| \cdot r^k \leq C \quad \forall\, n \in \mathsf{N} \qquad \overset{\text{Satz 18}}{\Longrightarrow} \qquad \sum_{n=0}^{\infty} |f_n|_\infty \leq C < \infty.$$

Korollar 29 ($Z = \mathsf{B}_r(0)$ und $V = \mathbb{K}$) zeigt somit die Behauptung.

b)
- Wegen Teil a) ist $(\mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]})|_{\mathsf{B}_r(0)} = \mathsf{Q}|_{\mathsf{B}_r(0)}$ stetig für jedes $0 < r < \mathsf{R}[\mathsf{Q}]$.
- Es gilt $\mathrm{K}[\mathsf{Q}] = \mathsf{B}_{\mathsf{R}[\mathsf{Q}]}(0) = \bigcup_{0<r<\mathsf{R}[\mathsf{Q}]} \mathsf{B}_r(0)$.[41]

Gemäß Übung 79.c), ist somit $\mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]}$ ebenfalls stetig.[42] □

Korollar 30. *Sei* $\mathsf{Q}(x) = \sum_{n=0}^{\infty} c_n \cdot (x-p)^n$ *eine Potenzreihe in* $\mathbb{K}$ *im Entwicklungspunkt* $p \in \mathbb{K}$*, mit Konvergenzradius* $\mathsf{R}[\mathsf{Q}] > 0$*. Wir setzen* $f := \mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]}$*. Die folgenden Aussagen äquivalent zueinander:*

1) *Es gilt* $c_n = 0$ *für alle* $n \in \mathbb{N}$*.*
2) *Es gilt* $f = 0$*.*
3) *Es existiert eine Folge* $(x_n)_{n\in\mathsf{N}}$ *in* $\mathrm{K}[\mathsf{Q}] \setminus \{p\}$ *mit* $\lim_n x_n = p$*, sowie* $f(x_n) = 0$ *für alle* $n \in \mathsf{N}$*.*

Beweis. Die Implikation 1) $\Rightarrow$ 2) ist klar; und für die Implikation 2) $\Rightarrow$ 3) betrachte die Folge $(p + \frac{1}{n})_{n \geq k}$ für $k \in \mathbb{N}_{>0}$ mit $\frac{1}{k} < \mathsf{R}[\mathsf{Q}]$. Die Implikation 3) $\Rightarrow$ 1) folgt per Induktion:

- Wegen Satz 39.b) ist f stetig, also folgenstetig wegen Proposition 8. Wir erhalten $c_0 = f(p) = \lim_n f(x_n) = 0$.

[40] Erinnerung: Gemäß Satz 17 und Korollar 21 sind $\{\mathbb{R}, \mathbb{C}\}$ Banachräume (vollständige metrische Räume).
[41] Offensichtlich gilt $\bigcup_{0<r<\mathsf{R}[\mathsf{Q}]} \mathsf{B}_r(0) \subseteq \mathsf{B}_{\mathsf{R}[\mathsf{Q}]}(0)$; und für $x \in \mathsf{B}_{\mathsf{R}[\mathsf{Q}]}(0)$, gilt $x \in \mathsf{B}_r(0)$ mit $r := |x| + (\mathsf{R}[\mathsf{Q}] - |x|)/2 < \mathsf{R}[\mathsf{Q}]$.
[42] Für $0 < r < \mathsf{R}[\mathsf{Q}]$, ist $\mathsf{B}_r(0)$ offen im metrischen Raum $\mathrm{K}[\mathsf{Q}]$ (eingeschränkte Betragsmetrik).

- Es gelte $c_0, \dots, c_n = 0$ für ein $n \in \mathbb{N}$. Dann gilt für alle $p \neq x \in \Theta[\mathsf{Q}]$

$$\frac{f(x)}{(x-p)^{n+1}} = \frac{1}{(x-p)^{n+1}} \cdot \sum_{\ell=n+1}^{\infty} c_\ell \cdot (x-p)^\ell = \sum_{k=0}^{\infty} c_{k+n+1} \cdot (x-p)^k =: \mathsf{P}(x).$$

Somit gilt $\Theta[\mathsf{P}] \subseteq \Theta[\mathsf{Q}]$, also $\mathsf{R}[\mathsf{Q}] \leq \mathsf{R}[\mathsf{P}]$, und daher $\mathrm{K}[\mathsf{Q}] \subseteq \mathrm{K}[\mathsf{P}]$. Wegen Satz 39.b) ist $\tilde{f} := \mathsf{P}|_{\mathrm{K}[\mathsf{P}]}$ stetig (folgenstetig), mit

$$\tilde{f}(x_n) = \frac{f(x_n)}{(x_n-p)^{n+1}} = 0 \qquad \forall\, n \in \mathsf{N} \qquad \Longrightarrow \qquad c_{n+1} = \tilde{f}(p) = \lim_n \tilde{f}(x_n) = 0. \qquad \square$$

Proposition 13 (Eigenschaften der Exponentialfunktion)**.**

1) *Die Exponentialfunktion* $\exp\colon \mathbb{C} \ni z \mapsto \sum_{k=0}^{\infty} \frac{z^k}{k!} \in \mathbb{C}$ *ist stetig.*
2) $\exp_{\mathbb{R}} := \exp|_{\mathbb{R}}\colon \mathbb{R} \to (0,\infty)$ *ist stetig, bijektiv, und streng monoton wachsend.*
3) *Die Logarithmusfunktion*

$$\log := (\exp_{\mathbb{R}})^{-1}\colon (0,\infty) \to \mathbb{R}$$

ist stetig, bijektiv, und streng monoton wachsend.

Beweis. Wegen Satz 18 gilt

$$\mathrm{e}^x = \textstyle\sum_{k=0}^{\infty} \frac{x^k}{k!} \geq \sum_{k=0}^{n} \frac{x^k}{k!} \qquad \forall\, n \in \mathbb{N},\ x \geq 0. \tag{229}$$

1) Klar wegen Satz 39.b) (exp hat Konvergenzradius ∞).
2)
   - Wegen Lemma 40.2) und Teil 1) ist $\exp_{\mathbb{R}}$ stetig.
   - Wegen Satz 30.d) ($\mathrm{e}^x > 0$ für $x \in \mathbb{R}$) gilt $D := \exp_{\mathbb{R}}(\mathbb{R}) \subseteq (0,\infty)$; und wegen Korollar 28 ist D ein Intervall. Wegen (229) gilt $x := \mathrm{e}^1 - 1 > 0$; und dann zeigen Satz 30.a) (Funktionalgleichung) sowie Satz 8 (Bernoullische Ungleichung):

   $$\mathrm{e}^n = (\mathrm{e}^1)^n = (1+x)^n \geq 1 + n \cdot x \qquad \forall\, n \in \mathbb{N}_{>0}.$$

   Wegen Satz 30.b) gilt dann weiterhin $\mathrm{e}^{-n} \leq \frac{1}{1+n\cdot x}$ für alle $n \in \mathbb{N}_{>0}$. Aus der Intervalleigenschaft von D folgt nun $D = (0,\infty)$; also ist $\exp_{\mathbb{R}}\colon \mathbb{R} \to (0,\infty)$ surjektiv.
   - Für $\mathbb{R} \ni x < y \in \mathbb{R}$, gilt $\mathrm{e}^{y-x} \geq 1 + (y-x) > 1$ wegen (229). Die Funktionalgleichug liefert (beachte $\mathrm{e}^x > 0$)

   $$\mathrm{e}^y = \mathrm{e}^{x+(y-x)} = \mathrm{e}^x \cdot \mathrm{e}^{y-x} > \mathrm{e}^x.$$

   Somit ist die Exponentialfunktion streng monoton wachsend, und daher auch injektiv gemäß Lemma 43.
3) Klar wegen Teil 2) und Satz 37 (Satz über die Umkehrfunktion). $\square$

Bemerkung 56 (Die allgemeine Potenzfunktion)**.** *Wir setzen*

$$x^\alpha := \mathrm{e}^{\alpha \log(x)} \qquad \forall\, \alpha \in \mathbb{C},\ x > 0. \tag{230}$$

a) *Die neue Definition stimmt für $\alpha \in \mathbb{Q}$ mit der alten Definition 28 überein:*
   - *Im Falle $\alpha = \frac{1}{n}$ mit $n \in \mathbb{N}_{>0}$, erhalten wir aus der Eindeutigkeit der n-ten Wurzel (Implikation)*

   $$(\sqrt[n]{x})^n = x \overset{(*)}{=} \left(\mathrm{e}^{\frac{1}{n}\log(x)}\right)^n \qquad \Longrightarrow \qquad \sqrt[n]{x} = \mathrm{e}^{\frac{1}{n}\log(x)}. \tag{231}$$

   $(*)$ = (*Satz 30.a)* $\wedge$ $\mathrm{e}^{\log(x)} = x$ *für* $x > 0$)

- *Im Falle $\alpha = \frac{m}{n} \in \mathbb{Q}$ mit $m \in \mathbb{Z}$ und $n \in \mathbb{N}_{>0}$, erhalten wir aus Satz 30 (zweiter Schritt)*

$$(\sqrt[n]{x})^m \overset{(231)}{=} \left(\mathrm{e}^{\frac{1}{n}\log(x)}\right)^m = \mathrm{e}^{\frac{m}{n}\log(x)} = x^{\frac{m}{n}}. \tag{232}$$

b) Gegeben $\alpha \in \mathbb{C}$, so ist die Potenzfunktion $\eta_\alpha\colon (0,\infty) \ni x \mapsto x^\alpha \in \mathbb{C}$ stetig; denn sie ist eine Verkettung stetiger Funktionen: $\quad \eta_\alpha = \exp \circ \lambda_\alpha \circ \log$ *mit* $\lambda_\alpha\colon \mathbb{C} \ni y \mapsto \alpha \cdot y \in \mathbb{C}$.

Für $\alpha > 0$ ($\alpha < 0$), ist dann $\eta_\alpha\colon (0,\infty) \ni x \mapsto x^\alpha \in (0,\infty)$ streng monoton wachsend (fallend), weil $\lambda_\alpha|_{(0,\infty)}$ streng monoton wachsend (fallend) ist; und weil exp *sowie* log *beide streng monoton wachsend sind. Weiterhin ist η_α bijektiv weil $\lambda_\alpha|_{(0,\infty)}$ bijektiv ist.*

c) Für $x \in (0,\infty)$ gilt

$$x^\alpha \cdot x^\beta = x^{\alpha+\beta} \quad \forall\, \alpha,\beta \in \mathbb{C} \qquad \textit{sowie} \qquad (x^\alpha)^\beta = x^{\alpha\cdot\beta} \quad \forall\, \alpha,\beta \in \mathbb{R}. \tag{233}$$

In der Tat folgt mit Satz 30.a) (erste Zeile) sowie $x^\alpha \in (0,\infty)$ für $\alpha \in \mathbb{R}$ (zweite Zeile):

$$\begin{aligned} x^\alpha \cdot x^\beta &= \mathrm{e}^{\alpha\log(x)} \cdot \mathrm{e}^{\beta\log(x)} = \mathrm{e}^{(\alpha+\beta)\log(x)} = x^{\alpha+\beta} && \forall\, \alpha,\beta \in \mathbb{C} \\ (x^\alpha)^\beta &= \mathrm{e}^{\beta\log(x^\alpha)} = \mathrm{e}^{\beta(\alpha\log(x))} = \mathrm{e}^{(\alpha\cdot\beta)\log(x)} = x^{\alpha\cdot\beta} && \forall\, \alpha,\beta \in \mathbb{R}. \end{aligned}$$

Insbesondere gilt (zweiten Zeile) $\eta_\alpha^{-1} = \eta_{\alpha^{-1}}$ für alle $\alpha \in \mathbb{R}_\times$.

d) Für $\alpha \in [0,\infty)$, lässt sich η_α stetig auf $[0,\infty)$ fortsetzen. Wegen $\lim_{x\downarrow 0}\log(x) = -\infty$, gilt nämlich

$$\lim_{x\downarrow 0} x^\alpha = \lim_{x\downarrow 0} \mathrm{e}^{\alpha\log(x)} = \begin{cases} 0 & \textit{für } \alpha > 0 \\ 1 & \textit{für } \alpha = 0 \\ \infty & \textit{für } \alpha < 0. \end{cases}$$

8 Differentialrechnung

Ausgehend von der Frage nach der Approximierbarkeit von Funktionen durch affine Funktionen, d.h., Funktionen, deren Graph eine Gerade ist, werden wir in diesem Kapitel zu dem Begriff der Differenzierbarkeit einer Funktion $f\colon D \to \mathbb{R}$ auf einem reellen Intervall D geführt. Wir werden sehen, dass das lokale Verhalten einer differenzierbaren Funktion sehr eng mit dem Verhalten ihrer Ableitung korrelliert ist (Abschnitt 8.2). Hieraus ergeben sich insbesondere einfache Kriterien für lokale Extremwerte von Funktionen, die man zur Lösung vieler praktischer Probleme verwenden kann. Schließlich führen wir in Abschnitt 8.3 die wichtigsten trigonometrischen Funktionen, wie die Sinus- und die Cosinusfunktion ein, und definieren die Zahl π.

8.1 Die Ableitung einer Funktion

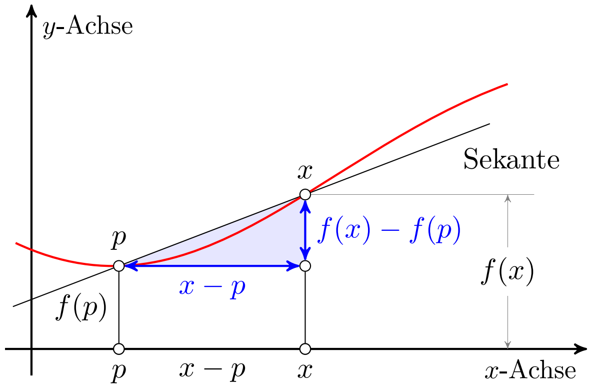
In diesem Abschnitt behandeln wir Differenzierbarkeit von Abbildungen $f\colon D \to \mathbb{R}$, wobei $D \subseteq \mathbb{R}$ im Folgenden immer als **nichtentartetes Intervall** (mindestens zweielementiges Intervall) angenommen wird: Es ist also D von der Form (26), mit $x < z$ im ersten Fall. Für $p \in D$ vorgegeben, setzen wir $D[p] := D \setminus \{p\}$, und definieren wir den zugehörigen **Differenzenquotienten** durch

$$\Delta[f,p]\colon D[p] \to \mathbb{R}, \qquad x \mapsto \frac{f(x) - f(p)}{x - p}.$$

Es beschreibt $\Delta[f,p](x)$ für $x \in D[p]$, die Steigung der *Sekanten* des Graphen

$$\Gamma_f = \{(x, f(x)) \mid x \in D\}$$

von f, die durch die beiden Punkte $(p, f(p))$ und $(x, f(x))$ verläuft.



(Quelle: https://tex.stackexchange.com/questions/168297/plot-to-illustrate-secant-lines)

Beispiel 64. *Sei $f\colon \mathbb{R} \to \mathbb{R}$ eine Abbildung.*

a) f heißt linear, wenn ein $a \in \mathbb{R}$ existiert, mit $f(x) = a \cdot x$ für alle $x \in \mathbb{R}$. Es gilt

$$\Delta[f,p](x) = \frac{a \cdot x - a \cdot p}{x - p} = a \qquad \forall\, p \in \mathbb{R},\ x \in \mathbb{R}.$$

b) f heißt affin, wenn $a, b \in \mathbb{R}$ existieren, mit $f(x) = a \cdot x + b$ für alle $x \in \mathbb{R}$. Es gilt

$$\Delta[f,p](x) = \frac{a \cdot x + b - (a \cdot p + b)}{x - p} = a \qquad \forall\, p \in \mathbb{R},\ x \in \mathbb{R}.$$

Übung 90. *Sei $D \subseteq \mathbb{R}$ ein nichterntartetes Intervall, und $p \in D$ ein Punkt. Machen Sie sich (durch Fallunterscheidungen) klar, dass p ein Berührungspunkt von $D[p]$ ist (also $p \in \mathcal{B}(D[p])$ gilt).*

Definition 57. *Sei $f\colon D \to \mathbb{R}$ eine Abbildung, und $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall.*

1) f heißt differenzierbar *bei (in) $p \in D$ mit Ableitung $f'(p) \in \mathbb{R}$, wenn gilt*

$$f'(p) = \lim_{x \to p} \Delta[f,p](x) = \lim_{x \to p} \frac{f(x) - f(p)}{x - p} = \lim_{h \to 0} \frac{f(p+h) - f(p)}{h}. \tag{234}$$

- *Gemäß Lemma 41 ist* (234) *gleichbedeutend zu* *(beachte* $\mathrm{dom}(\Delta[f,p]) = D[p]$*)*

$$\lim_{h \uparrow 0} \frac{f(p+h) - f(p)}{h} = \lim_{x \uparrow p} \Delta[f,p](x) = f'(p) = \lim_{x \downarrow p} \Delta[f,p](x) = \lim_{h \downarrow 0} \frac{f(p+h) - f(p)}{h}. \tag{235}$$

- *Gemäß Definition 53 (bzw.* (208)*) schreibt sich* (234) *auch wie folgt:*

$$\begin{gathered} \forall\, \varepsilon > 0\colon \exists\, \delta > 0\colon \quad (\, x \in D[p] \ \wedge\ |x - p| < \delta\,) \ \Rightarrow\ \left|\tfrac{f(x)-f(p)}{x-p} - f'(p)\right| < \varepsilon \\ (\ \forall\, \varepsilon > 0\colon \exists\, \delta > 0\colon \quad (\, p + h \in D[p] \ \wedge\ |h| < \delta\,) \ \Rightarrow\ \left|\tfrac{f(p+h)-f(p)}{h} - f'(p)\right| < \varepsilon\ ) \end{gathered} \tag{236}$$

- *Alternative Schreibweisen für die Ableitung $f'(p)$ sind:*

$$\tfrac{\mathrm{d}f}{\mathrm{d}x}(p), \qquad \tfrac{\mathrm{d}}{\mathrm{d}x} f(p), \qquad \tfrac{\mathrm{d}f(x)}{\mathrm{d}x}\Big|_{x=p}, \qquad \tfrac{\mathrm{d}}{\mathrm{d}x}\Big|_{x=p} f(x).$$

2) f heißt differenzierbar auf einer Teilmenge *$A \subseteq D$, wenn f in allen $p \in A$ differenzierbar ist.*

3) f heißt differenzierbar*, wenn f auf D differenzierbar ist (also differenzierbar in allen $p \in D$). In diesem Fall wird die Abbildung*

$$f'\colon D \to \mathbb{R}, \qquad p \mapsto f'(p)$$

als die Ableitung von f bezeichnet.

Beispiel 65. *Seien $a, b \in \mathbb{R}$, und $f\colon \mathbb{R} \ni x \mapsto a \cdot x + b \in \mathbb{R}$ eine affine Abbildung. Beispiel 64.b) zeigt*

$$\Delta[f,p](x) = a \qquad \forall\, p \in \mathbb{R} \;\wedge\; x \in \mathbb{R} \setminus \{p\} \qquad \Longrightarrow \qquad f'(p) = \lim_{x \to p} \Delta[f,p](x) = a \qquad \forall\, p \in \mathbb{R},$$

d.h., $f' = f[a]\colon \mathbb{R} \ni x \mapsto a \in \mathbb{R}$ ist konstant a.

Lemma 46. *Seien $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall, $p \in D$ ein Punkt, und $f\colon D \to \mathbb{R}$ differenzierbar in p. Ist $\tilde{D} \subseteq \mathbb{R}$ ein nichtentartetes Intervall mit $p \in \tilde{D} \subseteq D$, so ist auch $\tilde{f} := f|_{\tilde{D}}$ differenzierbar in p mit $\tilde{f}'(p) = f'(p)$.*

Beweis. Per Annahme gilt (236). Setzen wir $\tilde{f}'(p) := f'(p)$, so folgt wegen $\tilde{D}[p] \subseteq D[p]$

$$\forall\, \varepsilon > 0\colon \exists\, \delta > 0\colon \qquad (\, x \in \tilde{D}[p] \;\wedge\; |x - p| < \delta\,) \quad \Rightarrow \quad \left| \tfrac{\tilde{f}(p) - \tilde{f}(x)}{x - p} - \tilde{f}'(p) \right| < \varepsilon,$$

also gilt (236) für $\tilde{f}$, $\tilde{D}[p]$ mit $\tilde{f}'(p) = f'(p)$, d.h., $\lim_{x \to p} \tilde{f}(x) = \tilde{f}'(p) = f'(p)$. □

Wir erhalten unmittelbar:

Korollar 31. *Ist $f\colon D \to \mathbb{R}$ differenzierbar (D nichtentartet), so ist auch $f|_{\tilde{D}}$ für jedes nichtentartete Intervall $\tilde{D} \subseteq D$ differenzierbar mit $(f|_{\tilde{D}})' = f'|_{\tilde{D}}$.*

Beweis. Klar wegen Lemma 46. □

Übung 91. *Sei $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall, $p \in D$ ein Punkt, $I := (a, b)$ für $\overline{\mathbb{R}} \ni a < p < b \in \overline{\mathbb{R}}$, und $\tilde{D} := D \cap I$.*

a) Machen Sie sich klar, dass $\tilde{D}$ ebenfalls ein nichtentartetes Intervall ist (Fallunterscheidungen).

b) Zeigen Sie, dass ein $\tau > 0$ existiert, sodass für alle $x \in \mathbb{R}$ und $0 < \delta \leq \tau$ gilt:

$$x \in D[p] \;\wedge\; |x - p| < \delta \qquad \Longleftrightarrow \qquad x \in \tilde{D}[p] \;\wedge\; |x - p| < \delta. \tag{237}$$

Hinweis: Man wähle $\tau > 0$ mit $(p - \tau, p + \tau) \subseteq I$, und stelle beide Seiten von (237) als Schnitt von Mengen dar.

Übung 92. *Sei $f\colon D \to \mathbb{R}$ (D nicht entartet) eine Abbildung, und $p \in D$ ein Punkt. Sei $I = (a, b)$ für $\overline{\mathbb{R}} \ni a < p < b \in \overline{\mathbb{R}}$, sowie $\tilde{f} := f|_{I \cap D}$. Zeigen Sie:*[43]

a) Ist f differenzierbar in $p \in D$, so ist auch $\tilde{f}$ differenzierbar in p mit $\tilde{f}'(p) = f'(p)$.

Hinweis: Lemma 46 und Übung 91.a).

b) Ist $\tilde{f}$ differenzierbar in p, so ist auch f differenzierbar in p mit $f'(p) = \tilde{f}'(p)$.

Hinweis: Verwenden Sie das Kriterium (236) sowie Übung 91.b).

Lemma 47. *Ist $f\colon D \to \mathbb{R}$ differenzierbar in $p \in D$ (D nichtentartet), so ist f stetig in $p \in D$.*

Beweis. Wegen $\lim_{x \to p} \Delta[f,p](x) = f'(p)$ und Lemma 41 gilt (vgl. (235))

$$\lim_{h \uparrow 0} \tfrac{f(p+h) - f(p)}{h} = f'(p) \qquad \text{sowie} \qquad \lim_{h \downarrow 0} \tfrac{f(p+h) - f(p)}{h} = f'(p).$$

Mit Korollar 17.3), sowie der Äquivalenzen (210) ⇔ (212) und (211) ⇔ (213), folgt

$$\lim_{h \uparrow 0} f(p+h) - f(p) = \lim_{h \uparrow 0}(f(p+h) - f(p)) = \lim_{h \uparrow 0} \left(\tfrac{f(p+h) - f(p)}{h} \cdot h \right) = f'(p) \cdot \lim_{h \uparrow 0} h = 0$$
$$\lim_{h \downarrow 0} f(p+h) - f(p) = \lim_{h \downarrow 0}(f(p+h) - f(p)) = \lim_{h \downarrow 0} \left(\tfrac{f(p+h) - f(p)}{h} \cdot h \right) = f'(p) \cdot \lim_{h \downarrow 0} h = 0.$$

Lemma 42 zeigt somit $\lim_{h \to 0} f(p+h) = f(p)$, sodass f wegen Beispiel 55 stetig in p ist. □

[43]Teil b) ist artverwandt mit Übung 79.a).

$*$ *Alternativer Beweis:* Sei $\varepsilon > 0$ vorgegeben. Wegen (236) existiert ein $\delta > 0$ mit

$$|\Delta[f,p](x) - f'(p)| < \tfrac{\varepsilon}{2} \qquad \forall\, x \in (p-\delta, p+\delta) \cap D[p].$$

Indem wir δ durch $\min(\delta, \varepsilon/(2 \cdot |f'(p)|))$ ersetzen, können wir annehmen, dass zusätzlich gilt:

$$|x-p| \cdot |f'(p)| < \tfrac{\varepsilon}{2} \qquad \forall\, x \in (p-\delta, p+\delta) \cap D[p].$$

Für $x \in (p-\delta, p+\delta) \cap D[p]$, erhalten wir aus der Dreiecksungleichung:

$$\begin{aligned} |f(x) - f(p)| &= \big|\Delta[f,p](x) \cdot (x-p) - f'(p) \cdot (x-p) + f'(p) \cdot (x-p)\big| \\ &\le |x-p| \cdot |\Delta[f,p](x) - f'(p)| + |x-p| \cdot |f'(p)| \\ &\le \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Dies zeigt die Implikation (für $x = p$ ist die rechte Seite 0)

$$|x-p| < \delta \quad \text{für} \quad x \in D \qquad \Longrightarrow \qquad |f(x) - f(p)| < \varepsilon,$$

womit f stetig in p ist. $\square$

Beispiel 66.

a) *Gemäß Beispiel 50.d), ist die Betragsfunktion $f\colon \mathbb{R} \ni x \mapsto |x| \in [0,\infty)$ stetig. Allerdings ist f in 0 nicht differenzierbar; denn da die Grenzwerte*

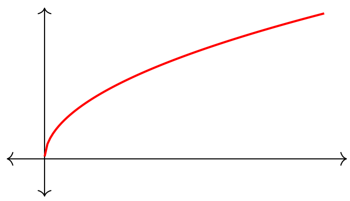
$$\lim_{h\uparrow 0} \tfrac{f(h)-f(0)}{h} = \lim_{h\uparrow 0} \tfrac{-h}{h} = \lim_{h\uparrow 0} -1 = -1 \quad \textit{und} \quad \lim_{h\downarrow 0} \tfrac{f(h)-f(0)}{h} = \lim_{h\downarrow 0} \tfrac{h}{h} = \lim_{h\uparrow 0} 1 = 1.$$

nicht übereinstimmen, existiert $f'(0) = \lim_{h\to 0} \frac{f(h)-f(0)}{h}$ gemäß (235) nicht.

b) *Gemäß Beispiel 63, ist die Wurzelfunktion $f\colon [0,\infty) \ni x \mapsto \sqrt{x} \in [0,\infty)$ stetig. Allerdings ist f in 0 nicht differenzierbar; denn es gilt*

$$\lim_{h\downarrow 0} \tfrac{f(h)-f(0)}{h} = \lim_{h\downarrow 0} \tfrac{\sqrt{h}}{h} = \lim_{h\downarrow 0} \tfrac{1}{\sqrt{h}} = \infty.$$

Geometrisch bedeutet dies, dass der Graph von f im Punkt $(p, f(p)) = (0,0)$ eine vertikale Tangente hat:



Wir kommen nun zu alternativen Charakterisierung von Differenzierbarkeit.

Lemma 48. *Sei $f\colon D \to \mathbb{R}$ eine Abbildung, $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall, $b \in \mathbb{R}$, $p \in D$, und $D - p = \{x - p \mid x \in D\}$. Die folgenden Aussagen sind äquivalent zueinander:*

1) *Es ist f differenzierbar in p mit Ableitung $f'(p) = b$.*
2) *Es existiert eine in 0 stetige Abbildung $\varepsilon\colon D - p \to \mathbb{R}$ mit $\varepsilon(0) = 0$, sodass gilt:*

$$f(p+h) = \underbrace{f(p) + h \cdot b}_{\textit{affine Funktion}} + \underbrace{h \cdot \varepsilon(h)}_{\textit{Restglied } R(h)} \qquad \forall\, h \in D - p. \tag{238}$$

(Beachte: Für das Restglied gilt $\lim_{h\to 0} \frac{R(h)}{h} = 0$.)

3) Es existiert eine in 0 *stetige Abbildung* $\Phi\colon D-p\to\mathbb{R}$ *mit* $\Phi(0)=b$*, sodass gilt:*

$$f(p+h)-f(p)=\Phi(h)\cdot h \qquad \forall\, h\in D-p. \tag{239}$$

Beweis. • 1) $\Rightarrow$ 2): Wir definieren

$$\varepsilon(h):=\begin{cases}\frac{f(p+h)-f(p)}{h}-b & \text{für } h\neq 0\\ 0 & \text{für } h=0\end{cases}$$

für alle $h\in D-p$. Wegen (236) (bzw. Proposition 10), ist ε stetig in 0 mit $\varepsilon(0)=0$, und es gilt

$$f(p+h)=f(p)+h\cdot b+h\cdot\varepsilon(h) \qquad \forall\, h\in D-p.$$

• 2) $\Rightarrow$ 3): Sei $\Phi(h):=b+\varepsilon(h)$ für alle $h\in D-p$. Dann ist Φ stetig in 0 mit $\Phi(0)=b$, und es gilt

$$f(p+h)-f(p)=\Phi(h)\cdot h \qquad \forall\, h\in D-p.$$

• 3) $\Rightarrow$ 1): Es gilt $\lim_{h\to 0}\frac{f(p+h)-f(p)}{h}=\lim_{h\to 0}\Phi(h)=\Phi(0)=b$. □

Satz 40. *Seien* $f,g\colon D\to\mathbb{R}$ *differenzierbar in* $p\in D$ *(*D *nichtentartet), sowie* $\lambda,\mu\in\mathbb{R}$*. Dann sind* $\lambda\cdot f+\mu\cdot g$ *und* $f\cdot g$ *differenzierbar in* p*. Gilt* $f(x)\neq 0$ *für alle* $x\in D$*, so sind auch auch* $\frac{1}{f}$ *und* $\frac{g}{f}$ *differenzierbar in* p*. In den entsprechenden Fällen gilt:*

1) $(\lambda\cdot f+\mu\cdot g)'(p)=\lambda\cdot f'(p)+\mu\cdot g'(p)$, *(Linearität)*

2) $(f\cdot g)'(p)=f'(p)\cdot g(p)+f(p)\cdot g'(p)$, *(Produktregel)*

3) $\left(\frac{1}{f}\right)'(p)=-\frac{f'(p)}{f(p)^2}$,

4) $\left(\frac{g}{f}\right)'(p)=\frac{g'(p)\cdot f(p)-g(p)\cdot f'(p)}{f(p)^2}$. *(Quotientenregel)*

Beweis. Wegen Lemma 48, existieren Funktionen $\Phi,\Psi\colon D-p\to\mathbb{R}$, mit

$$\begin{aligned} f(p+h)-f(p)&=\Phi(h)\cdot h, &\quad \Phi\ \text{ stetig in } 0, &\quad \Phi(0)=f'(p)\\ g(p+h)-g(p)&=\Psi(h)\cdot h, &\quad \Psi\ \text{ stetig in } 0, &\quad \Psi(0)=g'(p)\end{aligned}$$

für alle $h\in D-p$. Wir argumentieren mit Hilfe von Lemma 48, indem wir die Bedingung (239) für die entsprechenden Abbildungen nachweisen:

1) Für alle $h\in D-p$ gilt

$$\begin{aligned}(\lambda\cdot f+\mu\cdot g)&(p+h)-(\lambda\cdot f+\mu\cdot g)(p)\\ &=\lambda\cdot\big(f(p+h)-f(p)\big)+\mu\cdot\big(g(p+h)-g(p)\big)\\ &=\lambda\cdot\Phi(h)\cdot h+\mu\cdot\Psi(h)\cdot h=\underbrace{(\lambda\cdot\Phi+\mu\cdot\Psi)}_{=:\,\Theta}(h)\cdot h,\end{aligned}$$

wobei $\Theta\colon D-p\to\mathbb{R}$ stetig in 0 ist mit $\Theta(0)=\lambda\cdot f'(p)+\mu\cdot g'(p)$.

2) Für alle $h\in D-p$ gilt

$$\begin{aligned}(f\cdot g)(p+h)&=f(p+h)\cdot g(p+h)\\ &=\big(f(p)+\Phi(h)\cdot h\big)\cdot\big(g(p)+\Psi(h)\cdot h\big)\\ &=f(p)\cdot g(p)+\underbrace{\big(\Phi(h)\cdot g(p)+f(p)\cdot\Psi(h)+\Phi(h)\cdot\Psi(h)\cdot h\big)}_{=:\,\Theta(h)}\cdot h,\end{aligned}$$

wobei $\Theta\colon (D-p)\to\mathbb{R}$ stetig in 0 ist mit $\Theta(0)=f'(p)\cdot g(p)+f(p)\cdot g'(p)$.

3) Wegen Lemma 47 ist f stetig in p. Für alle $h \in D - p$ gilt nun

$$\left(\frac{1}{f}\right)(p+h) - \left(\frac{1}{f}\right)(p) = \frac{1}{f(p+h)} - \frac{1}{f(p)} = \frac{f(p) - f(p+h)}{f(p+h) \cdot f(p)} = \underbrace{\left(\frac{-\Phi(h)}{f(p+h) \cdot f(p)}\right)}_{=:\, \Theta(h)} \cdot h,$$

wobei $\Theta\colon D - p \to \mathbb{R}$ stetig in 0 ist mit $\Theta(0) = -\frac{f'(p)}{f(p)^2}$. □

4) Dies folgt sofort aus 2) und 3).

Übung 93. *Zeigen Sie Satz 40.4). Benutzen Sie zudem Übung 92 und Lemma 47 (f ist stetig in p)) um nachzuweisen, dass dass die Teile 3) und 4) von Satz 40 auch noch gelten, wenn man nur $f(p) \neq 0$ (anstelle $f(x) \neq 0$ für alle $x \in D$) fordert.*

Übung 94. *Sei $n \in \mathbb{N}_{>0}$ und $a, a_0, \ldots, a_n \in \mathbb{R}$. Zeigen Sie die folgenden Aussagen:*

a) Es ist $f\colon \mathbb{R} \ni x \mapsto a \in \mathbb{R}$ differenzierbar, mit $f' = 0$.

b) Es ist $f\colon \mathbb{R} \ni x \mapsto x^n \in \mathbb{R}$ differenzierbar, mit $f'\colon \mathbb{R} \ni x \mapsto n \cdot x^{n-1} \in \mathbb{R}$.

c) Es ist $f\colon \mathbb{R} \ni x \mapsto \sum_{k=0}^{n} a_k \cdot x^k \in \mathbb{R}$ differenzierbar, mit $f'\colon \mathbb{R} \ni x \mapsto \sum_{k=1}^{n} a_k \cdot k \cdot x^{k-1} \in \mathbb{R}$.

Hinweis: Induktion und Beispiel 66.

Übung 95. *Sei $n \in \mathbb{N}_{>0}$ und $c_0, \ldots, c_n \in \mathbb{R}$, sowie $f\colon \mathbb{R} \ni x \mapsto \sum_{k=0}^{n} c_k \cdot x^k \in \mathbb{R}$. Zeigen Sie, dass f genau dann konstant 0 ist, wenn $c_0, \ldots, c_n = 0$ gilt.*

Folgern Sie, dass die Darstellung einer Polynomfunktion (vgl. Beispiel 52) $P\colon \mathbb{R} \to \mathbb{R}$ in der Form $P = \sum_{k=0}^{n} a_k \cdot (\mathrm{id}_{\mathbb{R}})^k$ mit $n \in \mathbb{N}_{>0}$ und $a_0, \ldots, a_n \in \mathbb{R}$ eindeutig ist, d.h., gilt $P = \sum_{k=0}^{m} b_k \cdot (\mathrm{id}_{\mathbb{R}})^k$ für $m \in \mathbb{N}_{>0}$ und $b_0, \ldots, b_m \in \mathbb{R}$, so folgt bereits $m = n$ und $b_i = a_i$ für $i = 0, \ldots, n$.

Satz 41 (Kettenregel)**.** *Seien $f\colon \tilde{D} \to \mathbb{R}$ und $g\colon D \to \mathbb{R}$ Abbildungen ($D, \tilde{D} \subseteq \mathbb{R}$ nichtentartet), sodass $\mathrm{im}(g) \subseteq \tilde{D}$ gilt. Sei g differenzierbar in $p \in D$ und f differenzierbar in $g(p) \in \tilde{D}$. Dann ist $f \circ g$ differenzierbar in $p \in D$ mit*

$$(f \circ g)'(p) = f'(g(p)) \cdot g'(p).$$

Beweis. Wegen Lemma 48, existieren Funktionen $\Phi\colon D - p \to \mathbb{R}$ und $\Psi\colon \tilde{D} - g(p) \to \mathbb{R}$ mit

$$\begin{aligned} g(p+h) - g(p) &= \Phi(h) \cdot h, && \Phi \text{ stetig in } 0, && \Phi(0) = g'(p) \\ f(g(p) + \tilde{h}) - f(g(p)) &= \Psi(\tilde{h}) \cdot \tilde{h}, && \Psi \text{ stetig in } 0, && \Psi(0) = f'(g(p)) \end{aligned}$$

für alle $h \in D - p$ und $\tilde{h} \in \tilde{D} - g(p)$. Für $h \in D - p$, gilt dann $\Phi(h) \cdot h \in \tilde{D} - g(p)$ wegen $\mathrm{im}(g) \subseteq \tilde{D}$; und wir erhalten

$$(f \circ g)(p+h) = f(g(p+h)) = f\big(g(p) + \Phi(h) \cdot h\big) = \underbrace{f(g(p))}_{(f \circ g)(p)} + \underbrace{\Psi\big(\Phi(h) \cdot h\big) \cdot \Phi(h)}_{=:\, \Theta(h)} \cdot h,$$

wobei $\Theta\colon D - p \to \mathbb{R}$ stetig in 0 ist mit $\Theta(0) = \Psi(0) \cdot \Phi(0) = f'(g(p)) \cdot g'(p)$. Die Behauptung folgt daher aus Lemma 48. □

Übung 96. *Sei $n \in \mathbb{N}_{>0}$. Zeigen Sie die folgenden Aussagen:*

a) Es ist $f\colon \quad (0, \infty) \ni z \mapsto z^{-1} \in \mathbb{R}$ differenzierbar mit $f'(z) = -z^{-2}$ für alle $z \in (0, \infty)$.

b) Es ist $f\colon (-\infty, 0) \ni z \mapsto z^{-1} \in \mathbb{R}$ differenzierbar mit $f'(z) = -z^{-2}$ für alle $z \in (-\infty, 0)$.

c) Es ist $f\colon \quad (0, \infty) \ni z \mapsto z^{-n} \in \mathbb{R}$ differenzierbar mit $f'(z) = -n \cdot z^{-(n+1)}$ für alle $z \in (0, \infty)$.

d) Es ist $f\colon (-\infty,0) \ni z \mapsto z^{-n} \in \mathbb{R}$ differenzierbar mit $f'(z) = -n \cdot z^{-(n+1)}$ für alle $z \in (-\infty,0)$.

Beispiel 67. *Man kann mit Hilfe von Satz 41 die Ableitung der Funktion $h(x) = (x^3+1)^2$ ausrechnen, ohne auszumultiplizieren zu müssen:*

Es sind $f\colon \mathbb{R} \ni x \mapsto x^2 \in \mathbb{R}$ und $g\colon \mathbb{R} \ni x \mapsto x^3+1 \in \mathbb{R}$ beide differenzierbar gemäß Übung 94. Dann ist $h = f \circ g\colon \mathbb{R} \to \mathbb{R}$ differenzierbar nach Satz 41, mit

$$h'(x) = f'(g(x)) \cdot g'(x) = 2 \cdot g(x) \cdot 3x^2 = 6 \cdot (x^3+1) \cdot x^2$$

wegen $f'(x) = 2x$ und $g'(x) = 3x^2$ für alle $x \in \mathbb{R}$.

Der folgende Satz erweitert die Aussage von Satz 37 (Satz über die Umkehrfunktion) auf den differenzierbaren Fall.

Satz 42 (Differenzierbarkeit der Umkehrfunktionen)**.** *Sei $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall, sowie $f\colon D \to \mathbb{R}$ stetig und injektiv sowie differenzierbar in $p \in D$ mit $f'(p) \neq 0$. Dann ist $\hat{D} := f(D)$ ein nichtentartetes Intervall; und die Koeinschränkung $(f|^{\hat{D}})^{-1}\colon \hat{D} \to D$ ist stetig, sowie differenzierbar in $f(p)$ mit*

$$((f|^{\hat{D}})^{-1})'(f(p)) = \frac{1}{f'(p)}. \tag{240}$$

Notation 26. *Gegeben eine injektive Abbildung $f\colon \mathbb{R} \supseteq D \to \mathbb{R}$ mit $\hat{D} := f(D)$, so schreiben wir im Folgenden auch vereinfachen f^{-1} anstelle $(f|^{\hat{D}})^{-1}$.*

Beweis. Gemäß Satz 37 ist $\hat{D} = f(D)$ ein Intervall, und $(f|^{\hat{D}})^{-1}\colon \hat{D} \to D$ ist stetig. Zudem ist $\hat{D}$ nichtentartet; denn f ist injektiv, und D enthält mehr als einen Punkt (da nichtentartet). Für die Differenzierbarkeit in $f(p)$ argumentieren wir nun wie folgt:

- Gemäß Lemma 48, existiert eine Abbildung $\Phi\colon D - p \to \mathbb{R}$ mit

  $$f(p+h) - f(p) = \Phi(h) \cdot h, \qquad \Phi \text{ stetig in } 0, \qquad \Phi(0) = f'(p) \tag{241}$$

  für alle $h \in D - p$.
- Per Voraussetzung, gilt $\Phi(0) = f'(p) \neq 0$; und wegen der Injektivität von f gilt weiterhin

  $$f(p+h) \neq f(p) \quad \forall\, 0 \neq h \in D-p \qquad \Rightarrow \qquad \Phi(h) = \frac{f(p+h)-f(p)}{h} \neq 0 \quad \forall\, 0 \neq h \in D-p.$$

  Dies zeigt $\Phi(h) \neq 0$ für alle $h \in D - p$.
- Sei $q := f(p)$ und $\hat{h} \in \hat{D} - q$. Wir erhalten mit (241) (dritter Schritt)

  $$\begin{aligned} q + \hat{h} &= f\big(f^{-1}(\underbrace{q+\hat{h}}_{\in \hat{D}})\big) \\ &= f\big(p + \underbrace{(f^{-1}(q+\hat{h}) - p)}_{h \in D-p}\big) \\ &= f(p) + \Phi\big(f^{-1}(q+\hat{h}) - p\big) \cdot \big(f^{-1}(q+\hat{h}) - p\big) \\ &= q + \underbrace{\Phi\big(f^{-1}(q+\hat{h}) - p\big)}_{\neq 0} \cdot \big(f^{-1}(q+\hat{h}) - f^{-1}(q)\big). \end{aligned}$$

  Umstellen liefert

  $$(f|^{\hat{D}})^{-1}(q+\hat{h}) - (f|^{\hat{D}})^{-1}(q) = \overbrace{\frac{1}{\Phi\big(f^{-1}(q+\hat{h}) - p\big)}}^{=:\,\Theta(\hat{h})} \cdot \hat{h}.$$

- Nun ist $(f|^{\hat{D}})^{-1}$ stetig; und Φ ist stetig in 0. Daher zeigt Lemma 39 (Verkettungen), dass auch $\Theta\colon \hat{D} - f(p) \to \mathbb{R}_\times$ stetig in 0 ist (für $\hat{h} = 0$ ist $f^{-1}(q+\hat{h}) - p = f^{-1}(q) - p = 0$). Wegen Lemma 48, ist daher $(f|^{\hat{D}})^{-1}$ differenzierbar in $q = f(p)$, mit

$$((f|^{\hat{D}})^{-1})'(q) = \Theta(0) = \frac{1}{\Phi\big(f^{-1}(q) - p\big)} = \frac{1}{\Phi(0)} = \frac{1}{f'(p)}. \qquad \square$$

Beispiel 68. *Wir betrachten die affine Abbildung $f\colon \mathbb{R} \ni x \mapsto a \cdot x + b \in \mathbb{R}$ mit $a \neq 0$ (vgl. Beispiel 65). Dann ist $f' = a \neq 0$ konstant, und die Umkehrabbildung ist gegeben durch*

$$f^{-1}(y) = \frac{y-b}{a} \qquad \textit{mit} \qquad (f^{-1})'(f(x)) = \frac{1}{a} = \frac{1}{f'(x)}.$$

$*$Dieses Beispiel zeigt,[44] *dass man durch Spiegelung einer Geraden f mit der Steigung $a \neq 0$ (Graph von f) an der Diagonalen $\mathrm{D} := \{(x,x) \mid x \in \mathbb{R}\}$ in $\mathbb{R}^2$, eine Gerade mit der Steigung a^{-1} erhält (Graph von f^{-1}). Für allgemeine (nicht notwendigerweise affine) Funktionen besagt* (240) *daher, dass die Tangente an den Graphen von f^{-1} im Punkt $(f(p), p)$, durch Spiegelung an der Diagonalen* D *der Tangente an den Graph von f im Punkt $(p, f(p))$ entsteht.*

Bemerkung 57. *Die wesentliche Erkenntnis von Satz 42 ist, dass (unter den egegebenen Voraussetzungen) die Umkehrfunktion f^{-1} in $f(p)$ automatisch differenzierbar ist. Wüssten wir nämlich bereits, dass f^{-1} in $f(p)$ differenzierbar ist, so erhielten wir direkt aus der Kettenregel (Satz 41)*

$$(f^{-1})'(f(p)) \cdot f'(p) = (f^{-1} \circ f)'(p) = (\mathrm{id}_{\mathbb{R}})'(p) = 1$$

$$\Longrightarrow \qquad (f^{-1})'(f(p)) \neq 0 \qquad \wedge \qquad (f^{-1})'(f(p)) = \frac{1}{f'(p)}.$$

(Wir werden später sehen (Korollar 32.3)), dass die Voraussetzung der Injektivität (strenge Monotonie – siehe Lemma 43) insbesondere dann erfüllt ist, wenn $f'(x) > 0$ oder $f'(x) < 0$ für alle $x \in D$ gilt.)

Lemma 49. *Für die komplexe Exponentialfunktion* $\exp\colon \mathbb{C} \ni z \mapsto \sum_{n=0}^{\infty} \frac{z^n}{n!} \in \mathbb{C}_\times$ *gilt*

$$\lim_{h\to 0} \frac{\exp(h) - 1}{h} = 1, \tag{242}$$

d.h., $\lim_n \frac{\exp(h_n)-1}{h_n} = 1$ *für jede Folge* $(h_n)_{n\in\mathsf{N}}$ *in* $\mathbb{C}_\times$ *mit* $\lim_n h_n = 0$.

(Ausführlicher liest (242) sich in der Form

$$\lim_{h\to 0} \alpha(h) = 1 \qquad \text{mit} \qquad \alpha\colon \mathbb{C}_\times \ni h \mapsto (\exp(h) - 1) \cdot h^{-1} \in \mathbb{C}. \text{)}$$

Beweis. Für $h \in \mathbb{C}_\times$, gilt wegen Übung 75 (zweite Identität) und Korollar 17 (vierte Identität):

$$\alpha(h) = \frac{\mathrm{e}^h - 1}{h} = \frac{1}{h} \cdot \sum_{k=1}^{\infty} \frac{h^k}{k!} = \frac{1}{h} \cdot \lim_n \sum_{k=1}^{n} \frac{h^k}{k!} = \lim_n \sum_{k=1}^{n} \frac{h^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{h^{k-1}}{k!} = \overbrace{\sum_{k=0}^{\infty} \frac{h^k}{(k+1)!}}^{=:\,\mathsf{P}(h)}. \tag{243}$$

[44] Man überlegt sich allgemein, dass man den Graph $\Gamma_{f^{-1}}$ der Umkehrabbildung zu einer Bijektion $f\colon \mathbb{R} \to \mathbb{R}$, durch Spiegelung von Γ_f an der Diagonalen D erhält: Für $x \in \mathbb{R}$ steht die Verbindungsgerade zwischen $p_1 := (x, f(x))$ und $p_2 := (f(x), x) = (f(x), f^{-1}(f(x)))$ senkrecht auf der Diagonalen D; und schneidet diese in dem Punkt $p := (\frac{x+f(x)}{2}, \frac{x+f(x)}{2})$ (Übung), wobei p den selben euklidischen Abstand zu p_1 und p_2 hat.

- Die Potenzreihe P konvergiert wegen (243) auf ganz $\mathbb{C}$, d.h., $\mathbb{C} \subseteq \Theta[\mathsf{P}]$, also $\mathsf{R}[\mathsf{P}] = \infty$.

  (Alternativ erhält man die Konvergenz auch aus Satz 25, da $\mathrm{e}^{|h|} = \sum_{k=0}^{\infty} \frac{|h|^k}{k!}$, für jedes $h \in \mathbb{C}$, eine konvergente Majorante von $\mathsf{P}(h)$ ist.)
- Gemäß Satz 39 ist $\mathsf{P}\colon \mathbb{C} \ni z \mapsto \mathsf{P}(z) \in \mathbb{C}$ stetig; und wegen (243) gilt $\alpha = \mathsf{P}|_{\mathbb{C}_\times}\colon \mathbb{C}_\times \to \mathbb{C}$.
- Für jede Folge $(h_n)_{n\in\mathsf{N}}$ in $\mathbb{C}_\times$ mit $\lim_n h_n = 0$, gilt nun (P ist folgenstetig wegen Proposition 8)

$$\lim_n \alpha(h_n) = \lim_n \mathsf{P}(h_n) = \mathsf{P}(0) = \frac{1}{1!} = 1,$$

  sodass (242) aus Proposition 11 folgt. □

Beispiel 69. *Wir betrachten die reelle Exponentialfunktion* $\exp_{\mathbb{R}}\colon \mathbb{R} \to (0,\infty)$ *und ihre Umkehrabbildung* $\log\colon (0,\infty) \to \mathbb{R}$ *(vgl. Proposition 13).*

1) *Wegen Lemma 49, gilt*

$$\exp_{\mathbb{R}}'(0) = \lim_{h\to 0} \frac{\mathrm{e}^h - 1}{h} = \lim_{h\to 0} \frac{\mathrm{e}^h - \mathrm{e}^0}{h} = 1.$$

Dies folgert man entweder direkt aus dem Umstand, dass $\mathbb{R} \subseteq \mathbb{C}$ *ein metrischer Unterraum ist, oder man zunächst bemerkt, dass wegen Lemma 49 ebenfalls*

$$\lim_n \frac{\mathrm{e}^{h_n} - \mathrm{e}^0}{h_n} = \frac{\mathrm{e}^{h_n} - 1}{h_n} = 1$$

für jede Folge $(h_n)_{n\in\mathsf{N}}$ *in* $\mathbb{R}_\times$ *mit* $\lim_n h_n = 0$ *gilt, und dann Proposition 11 (bzw. die Äquivalenz* (208) $\Leftrightarrow$ (209) *aus Terminolgie 22.*3)*) anwendet.*

Aus der Funktionalgleichung $\mathrm{e}^{p+h} = \mathrm{e}^p \cdot \mathrm{e}^h$ *für* $p, h \in \mathbb{R}$ *(Satz 30.a)), erhalten folgt nun allgemein*

$$\exp_{\mathbb{R}}'(p) = \lim_{h\to 0} \frac{\mathrm{e}^{p+h} - \mathrm{e}^p}{h} = \lim_{h\to 0} \frac{\mathrm{e}^p \cdot \mathrm{e}^h - \mathrm{e}^p}{h} = \mathrm{e}^p \cdot \lim_{h\to 0} \frac{\mathrm{e}^h - \mathrm{e}^0}{h} = \mathrm{e}^p \cdot 1 = \exp_{\mathbb{R}}(p).$$

Dies zeigt, dass $\exp_{\mathbb{R}}$ *differenzierbar ist, und die folgende* Differentialgleichung *erfüllt*

$$\exp_{\mathbb{R}}' = \exp_{\mathbb{R}}. \tag{244}$$

2) *Mit Satz 42 und* $\mathrm{im}(\exp_{\mathbb{R}}') \overset{(244)}{=} \mathrm{im}(\exp_{\mathbb{R}}) = (0,\infty) \subseteq \mathbb{R}_\times$*, folgt nun auch die Differenzierbarkeit der Logarithmus* $\log\colon (0,\infty) \to \mathbb{R}$*, mit*

$$\log'(\exp_{\mathbb{R}}(p)) = \frac{1}{\exp_{\mathbb{R}}'(p)} \overset{(244)}{=} \frac{1}{\exp_{\mathbb{R}}(p)} \qquad \forall\, p \in (0,\infty).$$

Es gilt also die Beziehung (wähle $p = \log(x)$*)*

$$\log'(x) = \frac{1}{x} \qquad \forall\, x \in (0,\infty). \tag{245}$$

∗Eine Anwendung des Obigen ist ein neuer Beweis von Satz 31 im positiven reellen Falle; also der Formel $\mathrm{e}^x = \lim_n (1 + \frac{x}{n})^n$ *für alle* $x \in (0,\infty)$*:*

- *Wir erhalten mit Bemerkung 56*

$$\Big(1 + \frac{x}{n}\Big)^n \overset{(232)}{=} \exp\Big(n \cdot \log\Big(1 + \frac{x}{n}\Big)\Big) \qquad \forall\, n \in \mathbb{N},\ x \in \mathbb{R}. \tag{246}$$

- *Umformen des Exponenten liefert*

$$\lim_n n \cdot \log\left(1+\frac{x}{n}\right) = \lim_n \frac{\log(1+\frac{x}{n})}{\frac{1}{n}} = x \cdot \lim_n \frac{\log(1+\frac{x}{n}) - \log(1)}{\frac{x}{n}} = x \cdot \log'(1) \overset{(245)}{=} x.$$

- *Aus der Stetigkeit von* exp *und* (246)*, folgt daher* $\mathrm{e}^x = \lim_n (1+\frac{x}{n})^n$.

3) *Sei $\alpha \in \mathbb{R}$. Wir betrachten die Potenzfunktion (vgl. Bemerkung 56.b))*

$$\eta_\alpha \colon (0,\infty) \to (0,\infty), \qquad x \mapsto x^\alpha = \mathrm{e}^{\alpha \cdot \log(x)}.$$

Es gilt $\eta_\alpha = \exp \circ g$, für die differenzierbare Abbildung $g \colon (0,\infty) \ni x \mapsto \alpha \cdot \log(x) \in \mathbb{R}$. Wegen der Kettenregel ist daher auch η_α differenzierbar, mit

$$\eta_\alpha'(x) = \exp_{\mathbb{R}}'(g(x)) \cdot g'(x) = \exp_{\mathbb{R}}(g(x)) \cdot \alpha \cdot \frac{1}{x} = x^\alpha \cdot \frac{\alpha}{x} \overset{(233)}{=} \alpha \cdot x^{\alpha-1} \tag{247}$$

für alle $x \in (0,\infty)$.

4) *Seien $g \colon D \to (0,\infty)$ und $h \colon D \to \mathbb{R}$ differenzierbar (D nichtentartet). Dann ist*

$$f := g^h \colon D \to \mathbb{R}, \qquad x \mapsto g(x)^{h(x)} = \mathrm{e}^{h(x)\cdot \log(g(x))}$$

gemäß den bisher gezeigten Kriterien differenzierbar, und es folgt

$$f' = g^{h-1} \cdot (h' \cdot g \cdot (\log \circ g) + h \cdot g'). \tag{248}$$

- *Für $g \colon (0,\infty) \ni x \mapsto x \in (0,\infty)$ und $h \colon (0,\infty) \ni x \mapsto \alpha \in \mathbb{R}$, erhalten wir* (247)*.*
- *Für $g = \mathrm{id}_{(0,\infty)} = h$, also $f \colon (0,\infty) \ni x \mapsto x^x \in \mathbb{R}$, erhalten wir*

$$f'(x) = x^{x-1} \cdot (x \cdot \log(x) + x) = f(x) \cdot (\log(x) + 1) \qquad \forall\, x \in (0,\infty).$$

Übung 97. *Zeigen Sie die Identität in* (248)*.*

Terminologie 24. *Sei $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall, und $f \colon D \to \mathbb{R}$ eine Abbildung.*

- *f heißt n-*mal differenzierbar *für $n \in \mathbb{N}_{>0}$, wenn die Ableitungsoperation*

$$\partial \colon \{\, h \in C(D,\mathbb{R}) \mid h' \in \mathsf{Abb}(D,\mathbb{R}) \ \textit{existiert} \,\} \to \mathsf{Abb}(D,\mathbb{R}), \qquad h \mapsto h'$$

n-mal auf f anwendbar ist, d.h., es existiert

$$f^{(k)} := \underbrace{(\partial \circ \ldots \circ \partial)}_{k\text{-}mal}(f) = \partial(\partial(\ldots \partial(f)\ldots)) \qquad \textit{für alle} \qquad 1 \le k \le n.$$

Wir setzen $f^{(0)} := f$ (beachte $f^{(1)} = f'$). Man schreibt auch f'' anstelle $f^{(2)}$, f''' anstelle $f^{(3)}$ ... .

- *f heißt 0-mal stetig differenzierbar, wenn f stetig ist. Wir setzen $C^0(D,\mathbb{R}) := C(D,\mathbb{R})$.*
- *f heißt* stetig differenzierbar*, wenn f differenzierbar ist und f' zudem stetig ist. Die Menge der stetig differenzierbaren Funktionen $D \to \mathbb{R}$ wird mit $C^1(D,\mathbb{R})$ notiert.*
- *f heißt n-*mal stetig differenzierbar *für $n \in \mathbb{N}_{>0}$, wenn f eine n-mal differenzierbare Funktion ist, und $f^{(n)}$ zudem stetig ist. Die Menge der n-mal stetig differenzierbaren Funktionen $D \to \mathbb{R}$ wird mit $C^n(D,\mathbb{R})$ notiert.*
- *Es bezeichnet $C^\infty(D,\mathbb{R}) := \bigcap_{n \in \mathbb{N}_{>0}} C^n(D,\mathbb{R})$ die Menge der beliebig oft differenzierbaren (glatten) Abbildungen $D \to \mathbb{R}$. (Die Forderung nach der Stetigkeit der Ableitungen ist hier redundant; denn aus der Differenzierbarkeit von $f^{(n)}$ für $n \in \mathbb{N}$, folgt wegen Lemma 47 automatisch die Stetigkeit von $f^{(n)}$.)*

Bemerkung 58. *Die Ableitung einer differenzierbaren Funktion ist nicht immer differenzierbar. Beispielsweise ist die Funktion*

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad x \mapsto x \cdot |x| \tag{249}$$

differenzierbar mit $f'(x) = 2 \cdot |x|$ für alle $x \in \mathbb{R}$. Weiterhin ist f' im Nullpunkt nicht differenzierbar wegen Beispiel 66.b). Insbesondere gilt also $f \in C^1(\mathbb{R}, \mathbb{R}) \setminus C^2(\mathbb{R}, \mathbb{R})$.

Übung 98. *Zeigen Sie, dass f in (249) differenzierbar ist, mit $f'(x) = 2 \cdot |x|$ für alle $x \in \mathbb{R}$.*

Hinweis: Betrachten Sie die Fälle $x < 0$, $x = 0$, $x > 0$ jeweils separat.

Übung 99. *Zeigen Sie, dass $\exp_{\mathbb{R}}$ und $\log$ glatt sind, d.h., es gilt $\exp_{\mathbb{R}} \in C^\infty(\mathbb{R}, (0, \infty))$ sowie $\log \in C^\infty((0, \infty), \mathbb{R})$.*

Übung 100. *Seien $D, \hat{D} \subseteq \mathbb{R}$ nichtentartete Intervalle, sowie $k \in \mathbb{N}_{>0}$. Zeigen Sie (in der gegebenen Reihenfolge):*

a) Seien $f, g \in C^k(D, \mathbb{R})$ und $\lambda, \mu \in \mathbb{R}$. Dann gilt $(\lambda \cdot f + \mu \cdot g) \in C^k(D, \mathbb{R})$.

b) Gilt $f, g \in C^k(D, \mathbb{R})$, so gilt auch $(f \cdot g) \in C^k(D, \mathbb{R})$.

c) Gilt $g \in C^k(D, \mathbb{R})$ und $f \in C^k(\hat{D}, \mathbb{R})$ mit $\mathrm{im}(g) \subseteq \hat{D}$, so gilt $(f \circ g) \in C^k(D, \mathbb{R})$.

Analoge Aussagen gelten für k-mal differenzierbare Abbildungen, d.h., sind f und g beide k-mal differenzierbar, so sind in alle Fällen die entsprechenden Kombinationen ebenfalls k-mal differenzierbar.

8.2 Das lokale Verhalten von Funktionen

In diesem Abschnitt untersuchen wir, inwiefern die Eigenschaften einer differenzierbaren Funktion bereits durch die Eigenschaften ihrer Ableitung festgelegt sind. Beispielsweise ist eine Funktion, deren Ableitung 0 ist, bereits konstant. Ist die Ableitung stets nichtnegativ (≥ 0), so ist die Funktion monoton wachsend. Der Schlüssel hierzu ist der Mittelwertsatz der Differentialrechnung.

8.2.1 Der Mittelwertsatz

Wir beginnen mit dem folgenden Spezialfall des Mittelwertsatzes:

Satz 43 (Satz von Rolle)**.** *Sei $f\colon \mathbb{R} \supseteq [a, b] \to \mathbb{R}$ mit $a < b$ eine stetige Funktion, die auf (a, b) differenzierbar ist. Gilt $f(a) = 0 = f(b)$, so existiert ein $c \in (a, b)$ mit $f'(c) = 0$.*

Beweis. Ist f konstant, so gilt $f'(c) = 0$ für alle $c \in (a, b)$ (siehe Beispiel 65). Ist f nicht konstant, so existiert ein $x \in [a, b]$ mit $f(x) \neq 0$ (beachte $0 \in \mathrm{im}(f)$ wegen $f(a) = 0 = f(b)$). Wir nehmen an, dass $f(x) < 0$ gilt (den Fall $f(x) > 0$ behandelt man analog).

- Gemäß Satz 35, nimmt f ihr Minimum in einem Punkt $c \in [a, b]$ an.
- Es gilt $c \in (a, b)$; und zwar wegen $f(c) \leq f(x) < 0$ und $f(a) = 0 = f(b)$.
- Per Voraussetzung ist f in c differenzierbar. Wir erhalten mit $f(c) \leq \mathrm{im}(f)$, Lemma 41, sowie Korollar 14 (bzw. Beispiel 57 angewandt auf $\Delta[f, c]|_{[c,b)}$ und $\Delta[f, c]|_{[a,c)}$), dass

$$f'(c) = \lim_{h \downarrow 0} \underbrace{\frac{f(c + h) - f(c)}{h}}_{\geq 0} \geq 0 \qquad \text{und} \qquad f'(c) = \lim_{h \uparrow 0} \underbrace{\frac{f(c + h) - f(c)}{h}}_{\leq 0} \leq 0$$

gilt, also $f'(c) = 0$. □

Satz 44 (Mittelwertsatz)**.** *Sei $f\colon \mathbb{R} \supseteq [a,b] \to \mathbb{R}$ mit $a < b$ eine stetige Funktion, die auf (a,b) differenzierbar ist. Dann existiert ein $c \in (a,b)$ mit*

$$\frac{f(b)-f(a)}{b-a} = f'(c). \tag{250}$$

Beweis. Wir betrachten die stetige, und auf (a,b) differenzierbare Abbildung[45]

$$g\colon [a,b] \to \mathbb{R}, \qquad x \mapsto f(x) - \Big(f(a) + (x-a)\cdot \tfrac{f(b)-f(a)}{b-a}\Big),$$

für die $g(a) = 0 = g(b)$ gilt. Nach dem Satz von Rolle (Satz 43), existiert ein $c \in (a,b)$ mit

$$0 = g'(c) = f'(c) - \tfrac{f(b)-f(a)}{b-a} \qquad \Longrightarrow \qquad f'(c) = \tfrac{f(b)-f(a)}{b-a}.$$ □

Korollar 32 (Monotonie und Ableitung)**.** *Sei $f\colon D \to \mathbb{R}$ differenzierbar (D nichtentartet). Dann gelten die folgenden Aussagen:*

1) f ist monoton wachsend (fallend) genau dann, wenn $f' \geq 0$ ($f' \leq 0$) gilt.

2) f ist konstant genau dann, wenn $f' = 0$ gilt (f' ist konstant 0).

3) Ist $f' > 0$ (< 0), so ist f streng monoton wachsend (fallend).

Im Allgemeinen gilt die Umkehrung von Teil 3) **nicht**. Beispielsweise ist $f\colon \mathbb{R} \ni x \mapsto x^3 \in \mathbb{R}$ streng monoton wachsend, mit $f'\colon \mathbb{R} \to 3\cdot x^2 \in \mathbb{R}$, d.h., also $f'(0) = 0$.

Beweis. 1) Wir zeigen nur, dass f genau dann monoton wachsend ist, wenn $f' \geq 0$ gilt (der andere Fall folgt analog):

- Sei f monoton wachsend, sowie $p \in D$ vorgegeben. Dann gilt

$$\begin{aligned} f(p+h) - f(p) &\leq 0 \qquad && \forall\, (D[p]-p) \ni h < 0 \\ f(p+h) - f(p) &\geq 0 \qquad && \forall\, 0 < h \in (D[p]-p), \end{aligned}$$

  und somit $\frac{f(p+h)-f(p)}{h} \geq 0$ für alle $h \in D[p]$. Beispiel 57 zeigt

$$f'(p) = \lim_{h\to 0} \frac{f(p+h)-f(p)}{h} \geq 0.$$

- Sei $f' \geq 0$ auf D, und seien $D \ni x < y \in D$ vorgegeben. Nach Satz 44 (Mittelwertsatz), angewandt auf die stetige und differenzierbare Einschränkung $f|_{[x,y]}$ (Lemma 40.2) und Korollar 31), existiert ein $z \in (x,y)$ mit

$$f(y) - f(x) = \overbrace{f'(z)}^{\geq 0} \cdot \overbrace{(y-x)}^{>0} \geq 0. \tag{251}$$

  Somit ist f monoton wachsend.

2) Die Funktion f ist genau dann konstant, wenn f sowohl monoton wachsend als auch monoton fallend ist. Gemäß Teil 1) ist dies genau dann der Fall, wenn sowohl $f' \geq 0$ als auch $f' \leq 0$ gilt, also $f' = 0$.

3) Gilt $f' > 0$, so zeigt (251) im Beweis von Teil 1), dass f sogar streng monoton wachsend ist. Den Fall $f' < 0$ behandelt man analog. □

[45] Wir ziehen also von f die affine Funktion ab, deren Graph die beiden Endpunkte $(a, f(a))$ und $(b, f(b))$ des Graphen von f verbindet.

Korollar 33. *Sei D ein nichtentartetes Intervall, $z \in D$, und $\lambda, c \in \mathbb{R}$. Dann existiert genau eine differenzierbare Abbildung $f\colon D \to \mathbb{R}$, sodass $f' = \lambda \cdot f$ und $f(z) = c$ gilt. Diese Abbildung ist gegeben durch $f\colon D \ni x \mapsto c \cdot \mathrm{e}^{\lambda\cdot(x-z)} \in \mathbb{R}$.*

Beweis. Offensichtlich ist $f\colon D \ni x \mapsto c \cdot \mathrm{e}^{\lambda\cdot(x-z)} \in \mathbb{R}$ differenzierbar, mit $f' = \lambda \cdot f$ und $f(z) = c$.

Für die Eindeutigkeit, sei $g\colon D \to \mathbb{R}$ differenzierbar mit $g' = \lambda \cdot g$ und $g(z) = c$. Dann ist auch $h\colon D \ni x \mapsto g(x) \cdot \mathrm{e}^{-\lambda\cdot x} \in \mathbb{R}$ differenzierbar (Produkt und Kettenregel) mit

$$h'(x) = g'(x) \cdot \mathrm{e}^{-\lambda\cdot x} - g(x) \cdot \lambda \cdot \mathrm{e}^{-x} = 0 \qquad \forall\, x \in D.$$

Wegen Korollar 32.2), ist somit h konstant; und wir erhalten

$$h(x) = h(z) = c \cdot \mathrm{e}^{-\lambda\cdot z} \quad \forall\, x \in D \qquad \Longrightarrow \qquad g(x) = h(x) \cdot \mathrm{e}^{\lambda\cdot x} = c \cdot \mathrm{e}^{\lambda\cdot(x-z)} \quad \forall\, x \in D,$$

was die Eindeutigkeit zeigt. □

Übung 101. *Sei $D \subseteq \mathbb{R}$ nichtentartet, $f : \mathbb{R} \supseteq D \to \mathbb{R}$ stetig, und $p \in D$. Sei weiterhin f differenzierbar auf $D[p]$ mit $\lim_{x\to p} f'|_{D[p]} = a$. Zeigen Sie, dass f differenzierbar in p ist mit $f'(p) = a$.*

Hinweis: Machen Sie sich durch Fallunterscheidung klar, dass wegen Satz 44 zu jedem $x \in D[p]$ ein $\vartheta_x \in D[p] \cap (p - |x-p|, p + |x-p|)$ existiert mit

$$\frac{f(x) - f(p)}{x - p} = f'(\vartheta_x).$$

8.2.2 Extremwerte

Im Folgenden sei $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall.

Definition 58. *Sei $f\colon D \to \mathbb{R}$ eine Abbildung.*

a) *Ein Punkt $p \in D$ heißt* lokales Maximum von f, *wenn ein $\delta > 0$ existiert mit*

$$f(p) \geq f(x) \qquad \forall\, x \in (p-\delta, p+\delta) \cap D.$$

Es heißt $p \in D$ isoliertes lokales Maximum von f, *wenn ein $\delta > 0$ existiert mit*

$$f(p) > f(x) \qquad \forall\, x \in (p-\delta, p+\delta) \cap D[p].$$

b) *Ein Punkt $p \in D$ heißt* lokales Minimum von f, *wenn ein $\delta > 0$ existiert mit*

$$f(p) \leq f(x) \qquad \forall\, x \in (p-\delta, p+\delta) \cap D.$$

Es heißt $p \in D$ isoliertes lokales Minimum von f, *wenn ein $\delta > 0$ existiert mit*

$$f(p) < f(x) \qquad \forall\, x \in (p-\delta, p+\delta) \cap D[p].$$

c) *Ein Punkt $p \in D$ heißt* (isoliertes) lokales Extremum, *wenn p ein (isoliertes) lokales Maximum oder ein (isoliertes) lokales Minimum ist.*

d) *Ein Punkt $p \in D$ heißt*
   - globales Maximum, *wenn $f(p) = \max(f(D))$ gilt.*
   - globales Minimum, *wenn $f(p) = \min(f(D))$ gilt.*

Offensichtlich ist jedes globale Maximum/Minimum auch ein lokales Maximum/Minimum.

e) f heißt lokal streng monoton wachsend (fallend) um $p \in D$*, wenn ein* $\delta > 0$ *existiert, sodass* $f|_{(p-\delta,p+\delta)\,\cap\, D}$ *streng monoton wachsend (fallend) ist.*

Lemma 50 (Bedingungen für Extremwerte)**.** *Sei $f\colon \mathbb{R} \supseteq [a,b] \to \mathbb{R}$ mit $a < b$ eine differenzierbare Abbildung. Es gelten die folgenden Aussagen:*

1)
- *Ist a ein lokales Maximum von f, so gilt $f'(a) \leq 0$.*
- *Gilt $f'(a) < 0$, so ist a ein isoliertes lokales Maximum von f.*

2)
- *Ist b ein lokales Maximum von f, so gilt $f'(b) \geq 0$.*
- *Gilt $f'(b) > 0$, so ist b ein isoliertes lokales Maximum von f.*

3)
- *Ist a ein lokales Minimum von f, so gilt $f'(a) \geq 0$.*
- *Gilt $f'(a) > 0$, so ist a ein isoliertes lokales Minimum von f.*

4)
- *Ist b ein lokales Minimum von f, so gilt $f'(b) \leq 0$.*
- *Gilt $f'(b) < 0$, so ist b ein isoliertes lokales Minimum von f.*

5) Ist $p \in (a,b)$ ein lokales Extremum von f, so gilt $f'(p) = 0$.

Beweis. 1)
- Sei a ein lokales Maximum von f. Dann existiert ein $0 < \delta \leq b - a$ mit

$$\begin{aligned} & f(a+h) \leq f(a) && \forall\, (a+h) \in (a-\delta, a+\delta) \cap [a,b] \\ \Longrightarrow \quad & \frac{f(a+h)-f(a)}{h} \leq 0 && \forall\, (a+h) \in (a-\delta, a+\delta) \cap [a,b] = [a, a+\delta). \end{aligned}$$

  Beispiel 57 zeigt somit $f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \leq 0$.
- Gilt $0 > f'(a) = \lim_{h\downarrow 0} \frac{f(a+h)-f(a)}{h}$, so existiert ein $\delta > 0$ mit $\frac{f(a+h)-f(a)}{h} > 0$ für alle $0 < h < \delta$, d.h., $f(a+h) > f(a)$ für alle $0 < h < \delta$.

2) Analog zu Punkt 1).

3) Analog zu Punkt 1).

4) Analog zu Punkt 1).

5) Sei p ein lokales Maximum von f (der andere Fall folgt analog). Dann gilt

$$\begin{aligned} \frac{f(p+h)-f(p)}{h} &\geq 0 && \forall\, a-p \leq h < 0 \\ \frac{f(p+h)-f(p)}{h} &\leq 0 && \forall\, 0 > h \leq b-p. \end{aligned}$$

Hiermit folgt (wegen Beispiel 57 angewandt auf $\Delta[f,c]|_{[p,b]}$ und $\Delta[f,c]|_{[a,p]}$)

$$0 \geq \lim_{h\downarrow 0} \frac{f(p+h)-f(p)}{h} = f'(p) = \lim_{h\uparrow 0} \frac{f(p+h)-f(p)}{h} \geq 0,$$

also $f'(p) = 0$. □

Definition 59. *Sei $\emptyset \neq I \subseteq \mathbb{R}$ ein offenes Intervall, also von der Form (a,b) mit $-\infty \leq a < b \leq \infty$. Sei weiterhin $f\colon I \to \mathbb{R}$ eine differenzierbare Abbildung. Ein Punkt $p \in I$ heißt* kritischer Punkt von f*, wenn $f'(p) = 0$ gilt.*

Lemma 50.5) (angewandt auf $f|_{[x,y]}$ mit $a < x < p < y < b$) zeigt, dass jedes (lokale) Extremum notwendigerweise auch ein kritischer Punkt ist. Wir wollen nun noch untersuchen, inwiefern auch die Umkehrung gilt.

Lemma 51 (Kritische Punkte)**.** *Sei $\emptyset \neq I \subseteq \mathbb{R}$ ein offenes Intervall; und $f\colon I \to \mathbb{R}$ eine differenzierbare Abbildung mit kritischem Punkt $p \in I$ (also $f'(p) = 0$).*

1) *Existiert ein $\delta > 0$ mit $f'(p+h) \cdot h > 0$ für alle $h \neq 0$ mit $|h| < \delta$,[46] so ist p ein isoliertes lokales Minimum von f. Dies gilt insbesondere, wenn eine der folgenden Bedingungen erfüllt ist:*
   - *f' ist lokal streng monoton wachsend um p.*
   - *Es existiert $f''(p)$, und es gilt $f''(p) > 0$.*
2) *Existiert ein $\delta > 0$ mit $f'(p+h) \cdot h < 0$ für alle $h \neq 0$ mit $|h| < \delta$,[47] so ist p ein isoliertes lokales Maximum von f. Dies gilt insbesondere, wenn eine der folgenden Bedingungen erfüllt ist:*
   - *f' ist lokal streng monoton fallend um p.*
   - *Es existiert $f''(p)$, und es gilt $f''(p) < 0$.*

Beweis. 1) Es gelte $f'(p+h) \cdot h > 0$ für alle $0 < |h| < \delta$:

- Es gilt $f'(p+h) < 0$ für alle $-\delta < h < 0$. Wegen Korollar 32.3), also ist $f|_{(p-\delta,p)}$ streng monoton fallend.
- Es gilt $f'(p+h) > 0$ für alle $0 < h < \delta$. Wegen Korollar 32.3), ist daher $f|_{(p,p+\delta)}$ streng monoton wachsend.

Aus der Stetigkeit von f (Lemma 47) folgt dann[48]

$$f(x) > f(p) < f(y) \quad \text{für alle} \quad x, y \in (p-\delta, p+\delta) \quad \text{mit} \quad x < p < y,$$

also ist p ein isoliertes lokales Minimum von f.

- Sei nun f' lokal um p streng monoton wachsend. Da I offen ist, existiert ein $\delta > 0$ mit $(p-\delta, p+\delta) \subseteq I$, sodass $f'|_{(p-\delta,p+\delta)}$ streng monoton wachsend ist. Wegen $f'(p) = 0$ folgt

$$f'(p+h) > f'(p) = 0 \quad \text{für} \quad 0 < h < \delta \qquad \wedge \qquad f'(p+h) < f'(p) = 0 \quad \text{für} \quad -\delta < h < 0,$$

  d.h., $f'(p+h) \cdot h > 0$ für alle $0 < |h| < \delta$.
- Wegen $f'(p) = 0$ und $f''(p) > 0$, gilt

$$0 < f''(p) = \lim_{h\to 0} \frac{f'(p+h) - f'(p)}{h} = \lim_{h\to 0} \frac{f'(p+h)}{h}.$$

  Daher existiert ein $\delta > 0$, mit $\frac{f'(p+h)}{h} > 0$ für alle $0 < |h| < \delta$; und wir erhalten

$$f'(p+h) \cdot h = \frac{f'(p+h)}{h} \cdot h^2 > 0 \qquad \forall\, 0 < |h| < \delta.$$

2) Analog zu Teil 1). □

Übung 102. *Sei $\mathbb{R} \ni a < p < b \in \mathbb{R}$ und $f\colon (a,b) \to \mathbb{R}$ stetig. Zeigen Sie, dass f streng monoton wachsend ist, wenn die Einschränkungen $f|_{(a,p)}$ und $f|_{(p,b)}$ streng monoton wachsen sind.*

Lemma 52. *Sei $\emptyset \neq I \subseteq \mathbb{R}$ ein offenes Intervall, $p \in I$, und $f\colon I \to \mathbb{R}$ differenzierbar mit $f'(p) = 0$.*

[46] Dies ist gleichbedeutend zu: $f'(p+h) > 0$ für $0 < h < \delta$ $\wedge$ $f'(p+h) < 0$ für $-\delta < h < 0$

[47] Dies ist gleichbedeutend zu: $f'(p+h) < 0$ für $0 < h < \delta$ $\wedge$ $f'(p+h) > 0$ für $-\delta < h < 0$

[48] Sei $(z_n)_{n\in\mathbb{N}}$ eine streng monoton wachsende Folge in (x,p) mit $\lim_n z_n = p$. Da f stetig ist, gilt $\lim_n f(z_n) = f(p)$. Wegen der Monotonieeigenschaft gilt weiterhin $f(x) > f(z_0) \geq f(z_n)$ für alle $n \in \mathbb{N}$. Korollar 14 liefert $f(x) > f(z_0) \geq \lim_n f(z_n) = f(p)$. Die Ungleichung $f(p) < f(y)$ zeigt man analog.

1) Ist p ein isoliertes lokales Minimum von f', so ist f lokal streng monoton wachsend um p.

2) Ist p ein isoliertes lokales Maximum von f', so ist f lokal streng monoton fallend um p.

Beweis. 1) Sei p ein isoliertes lokales Minimum von f'.

- Wegen der Offenheit von I, existiert ein $\delta > 0$ mit $(p-\delta, p+\delta) \subseteq I$ und $f'(p+h) > f'(p) = 0$ für alle $0 < |h| < \delta$.
- Wegen Korollar 32.3), sind $f|_{(p-\delta,p)}$ und $f|_{(p,p+\delta)}$ streng monoton wachsend. Da $f|_{(p-\delta,p+\delta)}$ stetig ist, zeigt Übung 102, dass $f|_{(p-\delta,p+\delta)}$ streng monoton wachsend ist.

2) Analog zu Teil 1). $\square$

Satz 45. *Sei $\emptyset \neq I \subseteq \mathbb{R}$ offen, $n \geq 2$, und $f\colon I \to \mathbb{R}$ eine $(n-1)$-mal differenzierbare Funktion. Für einen Punkt $p \in I$ existiere auch $f^{(n)}(p)$, und es gelte*

$$f^{(n)}(p) \neq 0 \qquad \textit{sowie} \qquad f^{(k)}(p) = 0 \quad \textit{für alle} \quad 1 \leq k \leq n-1.$$

Dann tritt genau einer der folgenden vier Fälle auf:

- *n ist gerade mit $f^{(n)}(p) > 0$: Dann ist p ein isoliertes lokales Minimum von f.*
- *n ist gerade mit $f^{(n)}(p) < 0$: Dann ist p ein isoliertes lokales Maximum von f.*
- *n ist ungerade mit $f^{(n)}(p) > 0$: Dann wächst f lokal um p streng monoton.*
- *n ist ungerade mit $f^{(n)}(p) < 0$: Dann fällt f lokal um p streng monoton.*

Beweis. Die Fälle mit $f^{(n)}(p) < 0$ führt man durch Multiplikation mit -1 auf die jeweiligen anderen Fälle zurück. Wir nehmen daher $f^{(n)}(p) > 0$ an. Nach Voraussetzung ist

$$\left(f^{(n-2)}\right)''(p) = f^{(n)}(p) > 0.$$

Wegen Lemma 51.1) (zweiter Unterpunkt) hat $f^{(n-2)}$ in p ein isoliertes lokales Minimum:

- Ist $n = 2$, so gilt die Aussage daher wegen $f = f^{(n-2)}$ (d.h., f hat in p ein isoliertes lokales Minimum).
- Ist $n = 3$, so zeigt Lemma 52.1), dass $f = f^{(n-3)}$ lokal streng monoton wachsend um p ist. Also ist die Aussage auch in diesem Fall korrekt.

Es sei nun $n \geq 4$:

- $f^{(n-3)}$ ist lokal streng monoton wachsend um p (gleiches Argument wie im Fall $n = 3$).
- $f^{(n-4)}$ hat daher wegen Lemma 51.1) (erster Unterpunkt) sowie $f^{(n-3)}(p) = 0$, ein isoliertes lokales Minimum in p.

Die Behauptung folgt nun für beliebige $n \geq 2$ durch iteratives anwenden von Lemma 52.1) und Lemma 51.1) (erster Unterpunkt) in obiger Weise. $\square$

Bemerkung 59 (Bestimmung von Extremwerten – Zusammenfassung)**.** *Sei $f\colon [a,b] \to \mathbb{R}$ mit $a < b$ eine differenzierbare Abbildung.*

1) Nach Satz 35 existieren Maxima und Minima von f in $[a,b]$, da f nach Lemma 47 stetig ist. Insbesondere existiert ein globales Maximum und ein globales Minimum von f.[49]

Nach Lemma 50 liegen alle lokalen Extrema in der Menge

$$\{a,b\} \cup \{x \in (a,b) \mid f'(x) = 0\}.$$

[49] Beachte, dass weder globale Maxima noch globale Minima eindeutig sein sein müssen. Beispielsweise hat die Sinusfunktion die globalen Maxima $\{\frac{\pi}{2} + 2\pi n \mid n \in \mathbb{Z}\}$ und die globalen Minima $\{-\frac{\pi}{2} + 2\pi n \mid n \in \mathbb{Z}\}$.

2) Ist $p \in (a,b)$ ein lokales Extremum, so gilt $f'(p) = 0$. Dies ist notwendig, aber nicht hinreichend. Gilt $f''(p) \neq 0$ (Existenz vorausgesetzt), so liegt ein isoliertes lokales Extremum vor (Lemma 51):

- *Gilt $f''(p) > 0$, so ist p ein isoliertes lokales Minimum von f.*
- *Gilt $f''(p) < 0$, so ist p ein isoliertes lokales Maximum von f.*

3) Ist a ein lokales Maximum (Minimum), so ist $f'(a) \leq 0$ $(f'(a) \geq 0)$. Dies ist notwendig, aber nicht hinreichend. Ist $f'(a) < 0$ $(f'(a) > 0)$, so ist a ein isoliertes lokales Maximum (Minimum).

Ist b ein lokales Maximum (Minimum), so ist $f'(b) \leq 0$ $(f'(b) \geq 0)$. Dies ist notwendig, aber nicht hinreichend. Ist $f'(b) > 0$ $(f'(b) < 0)$, so ist b ein isoliertes lokales Maximum (Minimum).

Beispiel 70.

a) Sei $f\colon [1,2] \ni x \mapsto x^2 \in \mathbb{R}$. Dann gilt $f'(x) = 2x > 0$ für alle $x \in [1,2]$; also ist f streng monoton wachsend gemäß Korollar 32. Folglich liegt bei $p = 1$ ein globales Minimum vor (insbesondere ein isoliertes lokales Minimum), und bei $p = 2$ ein globales Maximum (insbesondere ein isoliertes lokales Maximum).

b) Sei $f\colon [0,2] \ni x \mapsto x^3 - x \in \mathbb{R}$. Die einzige Nullstelle von $f'\colon [0,2] \ni x \mapsto 3x^2 - 1 \in \mathbb{R}$ in $[0,2]$, ist gegeben durch $x_0 = \frac{1}{\sqrt{3}} \in (0,2)$. Es gilt $f''(x_0) = 6 \cdot x_0 = \frac{6}{\sqrt{3}} > 0$; also liegt gemäß Bemerkung 59.2) ein isoliertes lokales Minimum vor. An den Intervallrändern haben wir

$$f'(0) = -1 < 0 \qquad \text{und} \qquad f'(2) = 3 \cdot 4 - 1 > 0.$$

Gemäß Bemerkung 59.3) sind daher $x_1 = 0$ und $x_2 = 2$ beides isolierte lokale Maxima. Für die Funktionswerte erhalten wir

$$f(x_1) = f(0) = 0, \quad f(x_2) = f(2) = 8 - 2 = 6 \quad \text{und} \quad f(x_0) = \tfrac{1}{\sqrt{3}} \cdot \left(\tfrac{1}{3} - 1\right) = -\tfrac{2}{3 \cdot \sqrt{3}} < 0.$$

Somit ist x_0 das einzige lokale Minimum von f, mithin das einzige globale Minimum (vgl. Bemerkung 59.1)). Weiterhin ist x_2 das einzige globale Maximum von f, d.h., das lokale Maximum x_1 nicht global.

Übung 103. *Sei $f\colon \mathbb{R} \supseteq (a,b) \to \mathbb{R}$ mit $a < b$ differenzierbar. Dann besitzt f' die Zwischenwerteigenschaft, d.h.,*

$$f'(x) \leq c \leq f'(y) \quad \text{für} \quad a < x < y < b \quad \text{und} \quad c \in \mathbb{R} \qquad \Longrightarrow \qquad \exists\, z \in [x,y]\colon f'(z) = c.$$

Hinweis: Ersetzt man f durch $(a,b) \ni x \mapsto f(x) - c \cdot x$, so kann man ohne Einschränkung $c = 0$ annehmen (Nachweis!). Die Aussage ist klar für $f'(x) = f'(y)$; also kann man zudem $f'(x) < 0 < f'(y)$ annehmen, indem man notfalls f durch $-f$ ersetzt (Nachweis!). Ist nun z eine Minimalstelle von f auf dem Intervall $[x,y]$ (Nachweis der Existenz!), so zeige man $f'(z) = 0$ (warum liegt kein Minimum am Rand?).

8.3 Die Regeln von de l'Hospital

In diesem Abschnitt behandeln wir eine effiziente Methode zur Berechnung von Grenzwerte von Quotienten zweier differenzierbarer Funktionen, die man nicht direkt auswerten kann, da sie vom Typ $\frac{0}{0}$ oder $\frac{\infty}{\infty}$ sind.

Satz 46 (Allgemeiner Mittelwertsatz der Differentialrechnung)**.** *Seien $f, g\colon [a,b] \to \mathbb{R}$ mit $a < b$ stetige Abbildungen, die auf dem Intervall (a,b) differenzierbar sind. Dann existiert ein $\xi \in (a,b)$ mit*

$$f'(\xi) \cdot (g(b) - g(a)) = g'(\xi) \cdot (f(b) - f(a)). \tag{252}$$

Hat g' keine Nullstelle in (a,b), so gilt $g(b) \neq g(a)$ und daher

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}. \tag{253}$$

Beweis. Sei $F : [a,b] \to \mathbb{R}$, definiert durch

$$F(x) := (f(x) - f(a)) \cdot (g(b) - g(a)) - (g(x) - g(a)) \cdot (f(b) - f(a)) \qquad \forall\, x \in [a,b].$$

Dann ist F stetige, sowie differenzierbar auf (a,b). Weiterhin gilt $F(a) = 0 = F(b)$, also existiert nach dem Satz von Rolle ein $\xi \in (a,b)$ mit

$$0 = F'(\xi) = f'(\xi) \cdot (g(b) - g(a)) - g'(\xi) \cdot (f(b) - f(a)),$$

was (252) zeigt. Es geltet nun zusätzlich $g'(x) \neq 0$ für alle $x \in (a,b)$. Wegen Satz 44 (Mittelwertsatz), existiert ein $\tau \in (a,b)$ mit

$$\tfrac{g(b)-g(a)}{b-a} = g'(\tau) \neq 0 \qquad \Longrightarrow \qquad g(b) - g(a) = (b-a) \cdot g'(\tau) \neq 0,$$

womit (253) aus (252) folgt. □

Aus dem allgemeinen Mittelwertsatz erhalten wir nun die folgenden Rechenregeln für Grenzwerte:

Satz 47 (1. Regel von de l'Hospital)**.** *Seien $f, g\colon [a,b] \to \mathbb{R}$ mit $a < b$ stetige Abbildungen, die auf (a,b) differenzierbar sind, mit $g'(x) \neq 0$ für alle $x \in (a,b)$. Dann gelten die folgenden Aussagen:*[50]

1) Sei $f(a) = 0 = g(a)$, und es existiere $\lim_{x\downarrow a} \frac{f'(x)}{g'(x)} \in \overline{\mathbb{R}}$. Dann gilt

$$\lim_{x\downarrow a} \frac{f(x)}{g(x)} = \lim_{x\downarrow a} \frac{f'(x)}{g'(x)}. \tag{254}$$

(Beachte $g(x) \neq g(a) = 0$ für alle $x \in (a,b)$, wegen Satz 46) angewandt auf $g|_{[a,x]}$.)

2) Sei $f(b) = 0 = g(b)$, und es existiere $\lim_{x\uparrow b} \frac{f'(x)}{g'(x)} \in \overline{\mathbb{R}}$. Dann gilt

$$\lim_{x\uparrow b} \frac{f(x)}{g(x)} = \lim_{x\uparrow b} \frac{f'(x)}{g'(x)}. \tag{255}$$

(Beachte $g(x) \neq g(a) = 0$ für alle $x \in (a,b)$, wegen Satz 46) angewandt auf $g|_{[a,x]}$.)

Beweis. Wir zeigen nur den ersten Teil (der zweite Teil folgt analog). Für jedes $x \in [a,b]$, existiert gemäß Satz 46 (allgemeiner Mittelwertsatz für $f|_{[a,x]}$ und $g|_{[a,x]}$) ein $\xi[x] \in (a,x)$ mit

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi[x])}{g'(\xi[x])}.$$

Sei nun $(x_n)_{n\in\mathsf{N}}$ eine Folge in (a,b) mit $\lim_n x_n = a$. Dann gilt

$$a < \xi[x_n] < x_n \qquad \forall\, n \in \mathsf{N} \qquad \overset{\text{Lemma 31}}{\Longrightarrow} \qquad \lim\nolimits_n \xi[x_n] = a.$$

Da $\lim_{x\downarrow a} \frac{f'(x)}{g'(x)} \in \overline{\mathbb{R}}$ per Voraussetzung existiert, erhalten wir (mit der Äquivalenz (211) $\Leftrightarrow$ (213))

$$\lim_n \frac{f(x_n)}{g(x_n)} = \lim_n \frac{f'(\xi[x_n])}{g'(\xi[x_n])} = \lim_{x\downarrow a} \frac{f'(x)}{g'(x)}.$$

Da die Folge $(x_n)_{n\in\mathsf{N}}$ beliebig war, folgt die Behauptung aus der Äquivalenz (211) $\Leftrightarrow$ (213). □

Beispiel 71.

[50] Wir fassen $\frac{f}{g}$ und $\frac{g'}{f'}$ als Abbildungen $(a,b] \to \mathbb{R}$ auf.

a) Sei $a = 0$, $b = 1$. Es ist $\exp_{\mathbb{R}}$ differenzierbar, mit $\exp'_{\mathbb{R}} = \exp_{\mathbb{R}}$ und $\exp_{\mathbb{R}}(0) = 1$. Daher sind auch

$$f := \exp_{\mathbb{R}}|_{[0,1]} - 1 - \mathrm{id}_{[0,1]} \qquad \textit{und} \qquad g := \mathrm{id}_{[0,1]} \colon x \mapsto x$$

differenzierbar, mit $f(0) = 0 = g(0)$ und $g' = 1 \neq 0$. Da $\exp_{\mathbb{R}}$ stetig ist (vierter Schritt), zeigt Satz 47 (zweiter Schritt):

$$\lim_{x\downarrow 0} \frac{e^x - 1 - x}{x} = \lim_{x\downarrow 0} \frac{f(x)}{g(x)} = \lim_{x\downarrow 0} \frac{f'(x)}{g'(x)} = \lim_{x\downarrow 0} \frac{e^x - 1}{1} = 0.$$

b) $*$ Sei $t \in \mathbb{R}$, $a = 0$, und

$$0 < b < |t|^{-1} \quad \textit{für} \quad t \neq 0 \qquad \textit{sowie} \qquad b = 1 \quad \textit{für} \quad t = 0.$$

Dann erfüllen die Abbildungen

$$f \colon [0, b] \ni x \mapsto \log(1 + t \cdot x) \in \mathbb{R} \qquad \textit{und} \qquad g := \mathrm{id}_{[0,b]} \colon x \mapsto x$$

die Voraussetzungen in Satz 47.1):

- *Es gilt $f(0) = \log(1) = 0 = g(0)$.*
- *f, g sind differenzierbar mit (Beispiel 50, Beispiel 65, Proposition 13, Beispiel 69.2), Satz 41)*

$$f'(x) = \tfrac{t}{1+t\cdot x} \qquad \textit{sowie} \qquad g'(x) = 1 \qquad \textit{für alle} \qquad x \in [0, b].$$

Satz 47 liefert daher

$$\lim_{x\downarrow 0} \frac{\log(1 + t \cdot x)}{x} = \lim_{x\downarrow 0} \frac{f(x)}{g(x)} = \lim_{x\downarrow 0} \frac{f'(x)}{g'(x)} = \lim_{x\downarrow 0} \frac{\frac{t}{1+t\cdot x}}{1} = t.$$

Insbesondere erhalten wir hiermit einen neuen Beweis von

$$\lim_{n\to\infty} \left(1 + \frac{t}{n}\right)^n = \lim_{x\downarrow 0} (1 + x \cdot t)^{\frac{1}{x}} = \lim_{x\downarrow 0} \mathrm{e}^{\frac{\log(1+t\cdot x)}{x}} = \mathrm{e}^t.$$

für $t \in \mathbb{R}$ (vgl. Satz 31 und Beispiel 69.2)), wobei wir im letzten Schritt die Stetigkeit der Exponentialfunktion verwendet haben.

c) Satz 47 lässt sich auch induktiv anwenden: $\quad$ ($a = 0$ und $b = 1$)

$$\lim_{x\downarrow 0} \frac{\log(1 + x) - x + \frac{1}{2}x^2}{x^3} = \lim_{x\downarrow 0} \frac{\frac{1}{1+x} - 1 + x}{3x^2} = \lim_{x\downarrow 0} \frac{1 - \frac{1}{(1+x)^2}}{6x} = \lim_{x\downarrow 0} \frac{\frac{2}{(1+x)^3}}{6} = \frac{2}{6} = \frac{1}{3}.$$

Beachte: *Alle Voraussetzungen von Satz 47 müssen bei jeder Anwendung der de l'Hospitalschen Regel verifiziert werden. Die Existenz der Grenzwerte folgt dann sukzessive (und zwar in umgekehrter Reihenfolge) aus der Existenz des allerletzten Grenzwertes.*

Korollar 34. *Seien $f, g \colon [a, \infty) \to \mathbb{R}$ mit $a \in \mathbb{R}$ stetig Abbildung, die auf (a, ∞) differenzierbar sind. Es gelte weiterhin*

$$\lim\nolimits_{x\to\infty} f(x) = 0 = \lim\nolimits_{x\to\infty} g(x) \qquad \textit{sowie} \qquad g'(x) \neq 0 \qquad \textit{für alle} \qquad x \in (a, \infty).$$

Existiert der Grenzwert $\lim_{x\to\infty} \frac{f'(x)}{g'(x)} \in \overline{\mathbb{R}}$, so gilt

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)}.$$

$*$*Beweis.* Sei o.B.d.A. $a > 0$ (falls nicht, ersetze a durch 1). Wir setzen

$$\tilde{f} : [0, \tfrac{1}{a}] \to \mathbb{R}, \qquad \tilde{f}(x) = \begin{cases} f(\frac{1}{x}) & \text{falls } \; x \in (0, \frac{1}{a}] \\ 0 & \text{falls } \; x = 0 \end{cases}$$

$$\tilde{g} : [0, \tfrac{1}{a}] \to \mathbb{R}, \qquad \tilde{g}(x) = \begin{cases} g(\frac{1}{x}) & \text{falls } \; x \in (0, \frac{1}{a}] \\ 0 & \text{falls } \; x = 0 \end{cases}.$$

Dann sind $\tilde{f}$ und $\tilde{g}$ stetig mit $\tilde{f}(0) = 0 = \tilde{g}(0)$, sowie differenzierbar auf $(0, \frac{1}{a})$

$$\tilde{f}'(t) = f'(\tfrac{1}{t}) \cdot (-\tfrac{1}{t^2}) \qquad \text{sowie} \qquad g'(t) = g'(\tfrac{1}{t}) \cdot (-\tfrac{1}{t^2}) \qquad \text{für alle} \qquad t \in (0, \tfrac{1}{a}).$$

Existiert nun der Grenzwert $\lim_{x\to\infty} \frac{f'(x)}{g'(x)} \in \overline{\mathbb{R}}$, so folgt mit Satz 47 (dritter Schritt)

$$\lim_{x\to\infty} \frac{f'(x)}{g'(x)} = \lim_{t\downarrow 0} \frac{f'(\frac{1}{t}) \cdot (-\frac{1}{t^2})}{g'(\frac{1}{t}) \cdot (-\frac{1}{t^2})} = \lim_{t\downarrow 0} \frac{\tilde{f}'(t)}{\tilde{g}'(t)} = \lim_{t\downarrow 0} \frac{\tilde{f}(t)}{\tilde{g}(t)} = \lim_{x\to\infty} \frac{f(x)}{g(x)}.$$ □

Satz 48 (2. Regel von de l'Hospital)**.** *Seien $f, g \colon [a, b) \to \mathbb{R}$ mit $\mathbb{R} \ni a < b \in \overline{\mathbb{R}}$ stetige Abbildungen, die auf (a, b) differenzierbar sind, mit*

$$\lim_{x\uparrow b} g(x) = \infty \qquad \text{und} \qquad g'(x) \neq 0 \qquad \text{für alle} \qquad x \in (a, b).$$

Existiert $\lim_{x\uparrow b} \frac{f'(x)}{g'(x)} \in \overline{\mathbb{R}}$, so gilt[51]

$$\lim_{x\uparrow b} \frac{f(x)}{g(x)} = \lim_{x\uparrow b} \frac{f'(x)}{g'(x)}.$$

$*$*Beweis.* Wegen $\lim_{x\to b} g(x) = \infty$, existiert ein $C \in (a, b)$ mit $g(x) > 0$ für alle $x \in (C, b)$. Wegen $\lim_{x\to b} g(x) = \infty$, ist $g|_{(C,b)}$ zudem nicht monoton fallend. Gemäß Korollar 32, existiert somit ein $x \in (C, b)$ mit $g'(x) \not< 0$; also $g'(x) > 0$, da g' per Annahme keine Nullstellen besitzt. Mit Übung 103 folgt nun $g'(x) > 0$ für alle $x \in (a, b)$, da g' keine Nullstelle besitzt. Wir dürfen daher im Folgenden $g > 0$ und $g' > 0$ annehmen. Wir setzen $q := \lim_{x\uparrow b} \frac{f'(x)}{g'(x)} \in \overline{\mathbb{R}}$.

- Sei $q \in \mathbb{R}$. Für $\varepsilon > 0$ vorgegeben, existiert dann ein $C \in (a, b)$ mit

$$\frac{f'(y)}{g'(y)} \in (q - \varepsilon, q + \varepsilon) \qquad \forall\, y \in (C, b). \tag{256}$$

  Sei $p \in (C, p)$ fixiert. Für $C < p < x < b$, gilt dann $g(p) < g(x)$ (also $g(p) \neq g(x)$) wegen $g' > 0$ und Korollar 32; also

$$q - \varepsilon < \frac{f(x) - f(p)}{g(x) - g(p)} < q + \varepsilon$$

  wegen Satz 46 und (256). Wir erhalten

$$(q - \varepsilon) \cdot g(x) + c < f(x) < (q + \varepsilon) \cdot g(x) + d \qquad \forall\, p < x < b,$$

  mit Konstanten $c, d \in \mathbb{R}$. Wegen $g > 0$, gilt daher

$$q - \varepsilon + \frac{c}{g(x)} < \frac{f(x)}{g(x)} < q + \varepsilon + \frac{d}{g(x)}.$$

[51]Man beachte, dass keine Bedingung an den Grenzwert $\lim_{x\uparrow b} f(x)$ gestellt wird.

Wegen $\lim_{x\uparrow b} g(x) = \infty$, gilt $\lim_{x\uparrow b} \frac{c}{g(x)} = 0 = \lim_{x\uparrow b} \frac{d}{g(x)}$. Daher existiert ein $p < C' < b$, mit

$$q - 2\varepsilon < \frac{f(x)}{g(x)} < q + 2\varepsilon \qquad \forall\, x \in (C', b).$$

Da $\varepsilon > 0$ beliebig war, folgt $\lim_{x\to b} \frac{f(x)}{g(x)} = q$.

- Sei $q = \infty$. Analog zum vorherigen Punkt, erhalten wir für $R > 0$ vorgegeben, $a < p < b < x$ mit

$$R < \frac{f(x) - f(p)}{g(x) - g(p)} \qquad \forall\, p < x.$$

Hieraus erhalten wir $R \cdot g(x) + c < f(x)$ für alle $p < x < b$, für eine Konstante $c \in \mathbb{R}$; und somit

$$R + \frac{c}{g(x)} < \frac{f(x)}{g(x)} \qquad \forall\, p < x < b.$$

Dann existiert ein $p < C' < b$ mit Hiermit folgt $\frac{f(x)}{g(x)} > \frac{R}{2}$ für alle $c' < x < b$; also $\lim_{x\to\infty} \frac{f(x)}{g(x)} = q = \infty$, da $r \geq 0$ beliebig war.

Für $q = -\infty$ argumentiert man analog. □

Beispiel 72 (∗).

a) Seien $f(x) = \sum_{k=0}^{n} a_k \cdot x^k$ *und* $g(x) = \sum_{k=0}^{n} b_k \cdot x^k$ *Polynome vom Grad* n, *mit* $b_n > 0$. *Dann gilt* $\lim_{x\to\infty} g^{(k)}(x) = \infty$ *für alle* $0 \leq k \leq n-1$, *also*

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} = \ldots = \lim_{x\to\infty} \frac{f^{(n)}(x)}{g^{(n)}(x)} = \frac{a_n}{b_n}.$$

Siehe hierzu auch Beispiel 37.

b) $\lim_{x\to\infty} \frac{\log x}{x^2} = \lim_{x\to\infty} \frac{1/x}{2x} = \lim_{x\to\infty} \frac{1}{2x^2} = 0.$

c) $\lim_{x\to\infty} \frac{e^x}{x} = \lim_{x\to\infty} e^x = \infty.$

8.4 Trigonometrische Funktionen

In diesem Abschnitt führen wir die trigonometrischen Funktionen mittels der Eulerschen Formel

$$\mathrm{e}^{\mathrm{i}x} = \cos(x) + \mathrm{i}\sin(x) \quad \text{für} \quad x \in \mathbb{R} \tag{EF}$$

ein; wobei auf der linken Seite, die komplexe Exponentialfunktion[52]

$$\exp\ \mathbb{C} \to \mathbb{C}, \qquad z \mapsto \textstyle\sum_{n=0}^{\infty} \frac{z^n}{n!}$$

auf die komplexe Zahl $\mathrm{i}x$ ($x \in \mathbb{R}$) angewendet wird. Hieraus leiten wir ihre wichtigsten Eigenschaften, wie Periodizität, Reihenentwicklungen und die Additionstheoreme ab.

Definition 60. *Wir definieren den Sinus und den Kosinus durch*[53]

$$\sin\colon \mathbb{R} \to \mathbb{R}, \qquad x \mapsto \frac{\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{-\mathrm{i}x}}{2\mathrm{i}} = \mathrm{Im}(\mathrm{e}^{\mathrm{i}x}) \qquad \textit{(Sinusfunktion)}$$

$$\cos\colon \mathbb{R} \to \mathbb{R}, \qquad x \mapsto \frac{\mathrm{e}^{\mathrm{i}x} + \mathrm{e}^{-\mathrm{i}x}}{2} = \mathrm{Re}(\mathrm{e}^{\mathrm{i}x}) \qquad \textit{(Cosinusfunktion)},$$

womit (EF) *per konstruktionem gilt. Diese Funktionen sind stetig, da die Exponentialfunktion gemäß Proposition 13 stetig ist, und da die Abbildungen* $\mathbb{R} \ni x \mapsto \pm\mathrm{i}x \in \mathbb{C}$ *stetig sind.*

[52]Siehe Definition 48, Satz 30, und Proposition 13.

[53]Erinnerung: Gemäß Proposition 30.c), gilt $\overline{\mathrm{e}^{\mathrm{i}x}} = \mathrm{e}^{\overline{\mathrm{i}x}} = \mathrm{e}^{-\mathrm{i}x}$ für alle $x \in \mathbb{R}$.

Bemerkung 60.

a) *Wegen* $\mathrm{e}^0 = 1$, *gilt* $\cos(0) = 1$ *und* $\sin(0) = 0$.

b) *Wegen* $|\mathrm{e}^{\mathrm{i}x}| \overset{(118)}{=} \mathrm{e}^{\mathrm{i}x} \cdot \overline{\mathrm{e}^{\mathrm{i}x}} = \mathrm{e}^{\mathrm{i}x} \cdot \mathrm{e}^{-\mathrm{i}x} = \mathrm{e}^{\mathrm{i}(x-x)} = \mathrm{e}^0 = 1$, *ist*

$$\Psi\colon \mathbb{R} \ni x \mapsto \mathrm{e}^{\mathrm{i}x} \in \mathrm{K} = \{z \in \mathbb{C} \mid |z| = 1\} \tag{257}$$

eine Abbildung von $\mathbb{R}$ *auf den Einheitskreis. Insbesondere gilt (vgl.* (126)*)*

$$\operatorname{im}(\sin), \operatorname{im}(\cos) \subseteq [-1,1] \qquad \textit{also} \qquad \sin, \cos\colon \mathbb{R} \to [-1,1]. \tag{258}$$

Gemäß (EF), *geben dann* $\cos(x)$ *bzw.* $\sin(x)$ *die Projektion des Punktes* $\Psi(x)$ *auf die* x*-Achse bzw. die* y*-Achse an.*

$\big($ *Da* Ψ *stetig ist, ist gemäß Lemma 40.3) auch die Koeinschränkung* $\Psi|^{\mathrm{K}}\colon \mathbb{R} \to \mathrm{K}$ *stetig bezüglich der auf* K *eingeschränkten komplexen Betragsmetrik. Wegen der Funktionalgleichung ist dann* Ψ *ein stetiger Gruppenhomomorphismus von der additiven Gruppe* $(\mathbb{R}, +, 0)$ *der reellen Zahlen in die Gruppe* $(\mathrm{K}, \cdot, 1)$ *(Untergruppe der multiplikativen Gruppe* $(\mathbb{C}_\times, \cdot, 1)$*). Satz 52 und Lemma 54 (siehe Abschnitt 8.4.2) zeigen dann weiterhin, dass* Ψ *surjektiv ist mit*

$$\ker(\Psi) := \{x \in \mathbb{R} \mid \Psi(x) = 1\} = 2\pi \cdot \mathbb{Z}. \big)$$

Lemma 53. *Für alle* $x, y \in \mathbb{R}$ *gilt:*

1) $\sin(-x) = -\sin(x)$ *(ungerade Funktion)*
   $\cos(-x) = \cos(x)$ *(gerade Funktion)*
2) $\cos(x)^2 + \sin(x)^2 = 1$
3) $\cos(x+y) = \cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y)$ *(Additionstheorem des Cosinus)*
4) $\sin(x+y) = \cos(x) \cdot \sin(y) + \sin(x) \cdot \cos(y)$ *(Additionstheorem des Sinus)*

Beweis. 1) Dies folgt sofort aus der Definition.

2) Wir erhalten

$$\cos(x)^2 + \sin(x)^2 = \operatorname{Re}(\mathrm{e}^{\mathrm{i}x})^2 + \operatorname{Im}(\mathrm{e}^{\mathrm{i}x})^2 = |\mathrm{e}^{\mathrm{i}x}|^2 \overset{(118)}{=} \mathrm{e}^{\mathrm{i}x} \cdot \overline{\mathrm{e}^{\mathrm{i}x}} = \mathrm{e}^{\mathrm{i}x} \cdot \mathrm{e}^{-\mathrm{i}x} = \mathrm{e}^0 = 1.$$

3) und 4): Wir erhalten aus der Funktionalgleichung (Satz (30).(a)))

$$\begin{aligned}
&\cos(x+y) \ + \ \mathrm{i}\sin(x+y) \\
&\quad \overset{(\mathrm{EF})}{=} \mathrm{e}^{\mathrm{i}(x+y)} = \mathrm{e}^{\mathrm{i}x} \cdot \mathrm{e}^{\mathrm{i}y} \\
&\quad \overset{(\mathrm{EF})}{=} (\cos(x) + \mathrm{i}\sin(x)) \cdot (\cos(y) + \mathrm{i}\sin(y)) \\
&\quad = (\cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y)) + \mathrm{i}(\sin(x) \cdot \cos(y) + \sin(y) \cdot \cos(x)).
\end{aligned}$$

Die Behauptungen folgen nun durch Vergleich von Real- und Imaginärteil. □

Um die diffenrenzierbarkeit von Sinus Cosinus nachzuweisen, und um deren Ableitungen auszurechnen, benötigen wir das folgende Korollar zu Lemma 49.

Korollar 35. *Es gilt*

$$\lim_{x \to 0} \frac{\sin(x)}{x} = 1 \qquad \textit{sowie} \qquad \lim_{x \to 0} \frac{\cos(x) - 1}{x} = 0.$$

Beweis. Für jede Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}_\times$ mit $\lim_n x_n = 0$, gilt $\lim_n \pm\mathrm{i}x_n = 0$. Daher zeigt Lemma 49 (sowie Lemma 53.1) im letzten Schritt), dass

$$1 = \lim_n \frac{\mathrm{e}^{\pm\mathrm{i}x_n} - 1}{\pm\mathrm{i}x_n} \overset{\text{(EF)}}{=} \lim_n \frac{\cos(\pm x_n) - 1 + \mathrm{i}\sin(\pm x_n)}{\pm\mathrm{i}x_n} = \lim_n \left(\frac{\sin(x_n)}{x_n} \mp \mathrm{i}\,\frac{\cos(x_n) - 1}{x_n} \right)$$

gilt. Mit Korollar 17.1) folgt $\lim_n \frac{\sin(x_n)}{x_n} = 1$ sowie $\lim_n \frac{\cos(x_n)-1}{x_n} = 0$; und da die Folge $(x_n)_{n\in\mathbb{N}}$ beliebig war, zeigt Proposition 11 die Behauptung (alternativ wieder die Äquivalenz (208) $\Leftrightarrow$ (209) in Terminoligie 22.3)). $\square$

Satz 49. *Die Funktionen* $\sin\colon \mathbb{R} \to [-1,1]$ *und* $\cos\colon \mathbb{R} \to [-1,1]$ *sind glatt, und es gilt*

$$\sin' = \cos \qquad \textit{sowie} \qquad \cos' = -\sin\,. \tag{259}$$

Beweis. Für $x \in \mathbb{R}$ und $h \in \mathbb{R}_\times$, folgt mit Lemma 53 (Additionstheoreme), Korollar 27 (Grenzwertsätze für Abbildungen), sowie Korollar 35:

$$\begin{aligned}
\lim_{h\to 0} \frac{\cos(x+h) - \cos(x)}{h} &= \lim_{h\to 0} \frac{\cos(x)\cdot\cos(h) - \sin(x)\cdot\sin(h) - \cos(x)}{h} \\
&= \cos(x)\cdot \lim_{h\to 0}\frac{\cos(h)-1}{h} - \sin(x)\cdot\lim_{h\to 0}\frac{\sin(h)}{h} = -\sin(x), \\
\lim_{h\to 0} \frac{\sin(x+h) - \sin(x)}{h} &= \lim_{h\to 0} \frac{\cos(x)\cdot\sin(h) + \sin(x)\cdot\cos(h) - \sin(x)}{h} \\
&= \sin(x)\cdot \lim_{h\to 0}\frac{\cos(h)-1}{h} + \cos(x)\cdot\lim_{h\to 0}\frac{\sin(h)}{h} = \cos(x).
\end{aligned}$$

Daher sind Sinus und Kosinus differenzierbar, und es gilt (259). Es folgt nun per Induktion, dass Sinus und Kosinus unendlich oft differenzierbar sind; denn sind Sinus und Kosinus k-mal differenzierbar für ein $k \in \mathbb{N}_{>0}$, so wegen (259) und Übung 100 automatisch auch $(k+1)$-mal differenzierbar. $\square$

8.4.1 Schwingungsgleichungen

Proposition 14 (Schwingungsgleichung)**.** *Eine zweimal differenzierbare Funktion* $u : \mathbb{R} \to \mathbb{R}$ *löst (erfüllt) die Schwingungsgleichung*

$$u'' + u = 0$$

genau dann, wenn $u = \alpha\cdot\cos + \beta\cdot\sin$ *für gewisse* $\alpha, \beta \in \mathbb{R}$ *gilt. Es gilt dann notwendig:*

a) $\alpha = u(0)$ *und* $\beta = u'(0)$, *wegen* $\cos(0) = 1$ *und* $\sin(0) = 0$.

b) $u \in C^\infty(\mathbb{R},\mathbb{R})$, *wegen Satz 49 und Übung 100.*

Beweis. • Für $\alpha, \beta \in \mathbb{R}$ vorgegeben, sei $u := \alpha\cdot\cos + \beta\cdot\sin$. Wegen Satz 49 und Übung 100 gilt dann $u \in C^\infty(\mathbb{R},\mathbb{R})$ mit $u'' = -u$.

- Sei nun $u\colon \mathbb{R} \to \mathbb{R}$ zweimal differenzierbar, mit $u'' + u = 0$. Wir setzen $\tilde{u} := \alpha\cdot\cos + \beta\cdot\sin$, mit $\alpha := u(0)$ und $\beta := u'(0)$. Für $f := u - \tilde{u}$, gilt dann $f(0) = 0 = f'(0)$ sowie $f'' = -f$. Für $F := f^2 + (f')^2$ erhalten wir $F(0) = 0$, mit

$$F' = 2\cdot f\cdot f' + 2\cdot f'\cdot f'' = 2\cdot f\cdot f' - 2\cdot f'\cdot f = 0.$$

  Wegen Korollar 32.2) ist daher F konstant, d.h., $F(t) = F(0) = 0$ für alle $t \in \mathbb{R}$. Wegen $f^2, (f')^2 \geq 0$, gilt dann notwendig bereits $f = 0$. Dies zeigt $u = \tilde{u} = \alpha\cdot\cos + \beta\cdot\sin$. $\square$

Korollar 36 (Allgemeine Schwingungsgleichung). *Sei $\omega \in \mathbb{R}_\times$ vorgegeben. Eine zweimal differenzierbare Funktion $u : \mathbb{R} \to \mathbb{R}$ löst (erfüllt) die allgemeine Schwingungsgleichung*

$$u'' + \omega^2 \cdot u = 0$$

genau dann, wenn $u\colon \mathbb{R} \ni t \mapsto \alpha \cdot \cos(\omega \cdot t) + \beta \cdot \sin(\omega \cdot t) \in \mathbb{R}$ für gewisse $\alpha, \beta \in \mathbb{R}$ gilt. Es gilt dann notwendig:

a) $\alpha = u(0)$ und $\beta = \frac{u'(0)}{\omega}$, wegen $\cos(0) = 1$ und $\sin(0) = 0$.

b) $u \in C^\infty(\mathbb{R}, \mathbb{R})$, wegen Satz 49 und Übung 100.

Die Zahl ω heißt die *Winkelgeschwindigkeit* der Schwingung.

Beweis. • Für $\alpha, \beta \in \mathbb{R}$ vorgegeben, sei $u\colon \mathbb{R} \ni t \mapsto \alpha \cdot \cos(\omega \cdot t) + \beta \cdot \sin(\omega \cdot t) \in \mathbb{R}$. Wegen Satz 49 und Übung 100 gilt dann $u \in C^\infty(\mathbb{R}, \mathbb{R})$. mit $u'' + \omega^2 \cdot u = 0$.

• Sei nun $u\colon \mathbb{R} \to \mathbb{R}$ zweimal differenzierbar, mit $u'' + \omega^2 \cdot u = 0$. Dann ist $\tilde{u}\colon \mathbb{R} \ni t \mapsto u\big(\frac{t}{\omega}\big) \in \mathbb{R}$ ebenfalls zweimal differenzierbar (Kettenregel – Satz 41), mit $\tilde{u}'' + \tilde{u} = 0$ wegen

$$\tilde{u}'(t) = \tfrac{1}{\omega} \cdot u'\big(\tfrac{t}{\omega}\big) \qquad \forall\, t \in \mathbb{R} \qquad \Longrightarrow \qquad \tilde{u}''(t) = \tfrac{1}{\omega^2} \cdot u''\big(\tfrac{t}{\omega}\big) = -u\big(\tfrac{t}{\omega}\big) = -\tilde{u}(t) \qquad \forall\, t \in \mathbb{R}.$$

Proposition 14 zeigt, dass $\tilde{u} = \alpha \cdot \cos + \beta \cdot \sin$ für gewisse $\alpha, \beta \in \mathbb{R}$ gilt, d.h.,

$$u(t) = \tilde{u}(\omega \cdot t) = \alpha \cdot \cos(\omega \cdot t) + \beta \cdot \sin(\omega \cdot t) \qquad \forall\, t \in \mathbb{R},$$

was die Behauptung zeigt. □

8.4.2 Die Zahl π

Satz 50 (Reihendarstellung von Sinus und Kosinus). *Für alle $x \in \mathbb{R}$ gilt*

$$\cos(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \qquad \textit{sowie} \qquad \sin(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Beweis. Wegen Lemma 12 gilt $\mathrm{i}^{2k} = (\mathrm{i}^2)^k = (-1)^k$ für alle $k \in \mathbb{N}$. Weiterhin gilt

$$\sum_{n=0}^{2\ell} \mathrm{i}^n \frac{1 + (-1)^n}{2} \frac{x^n}{n!} = \sum_{k=0}^{\ell} \mathrm{i}^{2k} \frac{x^{2k}}{(2k)!} \qquad \text{sowie} \qquad \sum_{n=0}^{2\ell} \mathrm{i}^n \frac{1 - (-1)^n}{2\mathrm{i}} \frac{x^n}{n!} = \sum_{k=0}^{\ell-1} \mathrm{i}^{2k} \frac{x^{2k+1}}{(2k+1)!}$$

für alle $\ell \in \mathbb{N}$. Wir erhalten mit Korollar 23 (jeweils zweiter Schritt)[54]

$$\cos(x) = \frac{\mathrm{e}^{\mathrm{i}x} + \mathrm{e}^{-\mathrm{i}x}}{2} = \sum_{n=0}^{\infty} \frac{(\mathrm{i}x)^n + (-\mathrm{i}x)^n}{2 \cdot n!} = \sum_{n=0}^{\infty} \mathrm{i}^n \frac{1 + (-1)^n}{2} \frac{x^n}{n!} = \sum_{k=0}^{\infty} \mathrm{i}^{2k} \frac{x^{2k}}{(2k)!} = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$$

$$\sin(x) = \frac{\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{-\mathrm{i}x}}{2\mathrm{i}} = \sum_{n=0}^{\infty} \mathrm{i}^n \frac{1 - (-1)^n}{2\mathrm{i}} \frac{x^n}{n!} = \sum_{k=0}^{\infty} \mathrm{i}^{2k} \frac{x^{2k+1}}{(2k+1)!} = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!},$$

was die Behauptung zeigt. □

Korollar 37. *Es gilt $\cos(2) < 0$ sowie $\cos|_{[0,1]} \geq \frac{1}{2}$.*

[54] Für den vierter Schritt in der ersten Zeile, beachte beispielsweise, dass für $\alpha := \sum_{n=0}^{\infty} \mathrm{i}^n \frac{1+(-1)^n}{2} \frac{x^n}{n!}$ und $\varepsilon > 0$, ein $N_\varepsilon \in \mathbb{N}$ existiert, sodass $|\alpha - \sum_{k=0}^{\ell} \mathrm{i}^{2k} \frac{x^{2k}}{(2k)!}| = |\alpha - \sum_{n=0}^{2\ell} \mathrm{i}^n \frac{1+(-1)^n}{2} \frac{x^n}{n!}| < \varepsilon$ für alle $\ell \geq N_\varepsilon$ gilt.

Beweis. Sei $-3 < x < 3$. Dann ist $\left(\frac{x^{2n}}{(2n)!}\right)_{n\in\mathbb{N}_{>0}}$ monoton fallend, wegen

$$\frac{\frac{x^{2(k+1)}}{(2(k+1))!}}{\frac{x^{2k}}{(2k)!}} = \frac{x^2}{(2k+2)(2k+1)} \leq \frac{x^2}{4\cdot 3} < \frac{9}{12} < 1 \qquad \forall\, k \in \mathbb{N}_{>0}.$$

Wir erhalten mit Satz 50

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \sum_{m=1}^{\infty} \underbrace{\frac{x^{2(2m+1)}}{(2(2m+1))!} - \frac{x^{2(2m+2)}}{(2(2m+2))!}}_{\geq 0} \leq 1 - \frac{x^2}{2} + \frac{x^4}{4!}$$

$$\cos(x) = 1 - \frac{x^2}{2} + \sum_{m=1}^{\infty} \underbrace{\frac{x^{2(2m)}}{(2(2m))!} - \frac{x^{2(2m+1)}}{(2(2m+1))!}}_{\geq 0} \geq 1 - \frac{x^2}{2},$$

d.h., $\cos(2) \leq 1-2+\frac{16}{24} = -1+\frac{2}{3} = -\frac{1}{3} < 0$, sowie $\cos(x) \geq 1-\frac{x^2}{2} \geq 1-\frac{1}{2} = \frac{1}{2}$ für alle $x \in [0,1]$. □

Definition 61. *Wir definiern die Zahl π durch*

$$\pi := 2 \cdot \inf_{\mathbb{R}}(\underbrace{\{x \in [0,\infty) \mid \cos(x) = 0\}}_{=:\,\mathcal{N}}) \geq 0. \tag{260}$$

- *Dies ist wohldefiniert; denn wegen $\cos|_{[0,1]} \geq \frac{1}{2}$ und $\cos(2) < 0$ (Korollar 37), zeigt der Zwischenwertsatz (Satz 36), dass die Kosinusfunktion eine Nullstelle $\lambda \in (1,2)$ besitzt, aber nicht in $[0,1]$. Somit gilt $1 < \mathcal{N} \neq \emptyset$ mit $\lambda \in \mathcal{N} \cap (1,2)$. Es folgt $1 \leq \frac{\pi}{2} = \inf_{\mathbb{R}}(\mathcal{N}) \leq \lambda < 2$, also $2 \leq \pi = 2\cdot\inf_{\mathbb{R}}(\mathcal{N}) < 4$.*
- *Satz 51 zeigt nun weiterhin, dass $\cos(\frac{\pi}{2}) = 0$ gilt, also $\mathcal{N} \ni \frac{\pi}{2} = \inf_{\mathbb{R}}(\mathcal{N})$ und daher $\frac{\pi}{2} = \min(\mathcal{N})$. Wegen $1 \leq \frac{\pi}{2}$ und $\cos(1) \geq \frac{1}{2} > 0$, folgt hiermit nun auch $1 < \frac{\pi}{2} < 2$, also $2 < \pi < 4$.*

Satz 51. *Es gilt $e^{i\frac{\pi}{2}} = i$, d.h., $\cos(\frac{\pi}{2}) = 0$ und $\sin(\frac{\pi}{2}) = 1$ wegen* (EF).

Beweis. • Wegen Übung 64.b), existiert eine Folge $(x_n)_{n\in\mathsf{N}}$ in $\mathcal{N}$ (d.h. $\cos(x_n) = 0$ für alle $n \in \mathsf{N}$) mit $\lim_n x_n = \frac{\pi}{2}$. Wegen der Stetigkeit der Kosinusfunktion gilt daher $\cos(\frac{\pi}{2}) = \lim_n \cos(x_n) = 0$.

- Der Zwischenwertsatz (Satz 36) impliziert $\cos|_{[0,\frac{\pi}{2})} \geq 0$: Gilt nämlich $\cos(x) < 0$ für ein $x \in (0,\frac{\pi}{2})$, so existiert wegen $\cos(0) = 1$ und Satz 36 (angewandt auf $\cos|_{[0,x]}$), ein $0 < y < x < \frac{\pi}{2}$ mit $\cos(y) = 0$, was der Definition von π widerspricht.

  Wir erhalten mit Satz 49

$$\sin(\tfrac{\pi}{2}) = -\cos'(\tfrac{\pi}{2}) = -\lim_{h\uparrow 0} \frac{\cos(\frac{\pi}{2}+h)-\cos(\frac{\pi}{2})}{h} = \lim_{h\uparrow 0} \frac{\overbrace{\cos(\frac{\pi}{2}-|h|)}^{\geq 0}}{|h|} \geq 0.$$

  Wegen $\cos^2+\sin^2 = 1$ (Lemma 53.2)), gilt daher $\sin(\frac{\pi}{2})^2 = 1$ und somit $\sin(\frac{\pi}{2}) = 1$. □

Quadrieren der Identität aus Satz 51 liefert die bemerkenswerte Gleichung

$$e^{i\pi} + 1 = 0,$$

in der die 5 wichtigsten Zahlen der Analysis $0, 1, \pi, e$ und i vorkommen.

Korollar 38 (Periodizität)**.** *Für alle $x \in \mathbb{R}$, gilt:*

a) $\cos(x+2\pi) = \cos(x)$ *und* $\sin(x+2\pi) = \sin(x)$, *(2π-Periodizität)*

b) $\cos(x+\pi) = -\cos(x)$ *und* $\sin(x+\pi) = -\sin(x)$, *c)* $\cos(x \pm \frac{\pi}{2}) = \mp \sin(x)$ *und* $\sin(x \pm \frac{\pi}{2}) = \pm \cos(x)$.

Beweis. Wegen Lemma 53.3), Satz 51, und Lemma 53.1) gilt

$$\begin{aligned} \cos(x \pm \tfrac{\pi}{2}) &= \cos(x) \cdot \cos(\pm\tfrac{\pi}{2}) - \sin(x) \cdot \sin(\pm\tfrac{\pi}{2}) = \mp \sin(x) && \forall\, x \in \mathbb{R} \\ \sin(x \pm \tfrac{\pi}{2}) &= \cos(x) \cdot \sin(\pm\tfrac{\pi}{2}) + \sin(x) \cdot \cos(\pm\tfrac{\pi}{2}) = \pm \cos(x) && \forall\, x \in \mathbb{R}. \end{aligned}$$

Dies zeigt c); und weiterhin folgt

$$\begin{aligned} \cos(x+\pi) &= \cos((x+\tfrac{\pi}{2})+\tfrac{\pi}{2}) = -\sin(x+\tfrac{\pi}{2}) = -\cos(x) \\ \sin(x+\pi) &= \sin((x+\tfrac{\pi}{2})+\tfrac{\pi}{2}) = \quad \cos(x+\tfrac{\pi}{2}) = -\sin(x) \end{aligned}$$

für alle $x \in \mathbb{R}$, also b). Hieraus erhalten wir

$$\begin{aligned} \cos(x+2\pi) &= \cos((x+\pi)+\pi) = -\cos(x+\pi) = \cos(x) \\ \sin(x+2\pi) &= \sin((x+\pi)+\pi) = -\sin(x+\pi) = \sin(x) \end{aligned}$$

für alle $x \in \mathbb{R}$, also a). □

Korollar 39 (Nullstellen von Sinus und Kosinus)**.** *Es gilt*

$$\{x \in \mathbb{R} \mid \sin(x) = 0\} = \mathbb{Z} \cdot \pi \qquad \textit{und} \qquad \{x \in \mathbb{R} \mid \cos(x) = 0\} = \tfrac{\pi}{2} + \mathbb{Z} \cdot \pi.$$

Beweis. Sei $\mathcal{Z} := \{x \in \mathbb{R} \mid \cos(x) = 0\}$. Aus Korollar 38.b) folgt induktiv

$$\cos(x + k \cdot \pi) = (-1)^k \cdot \cos(x) \qquad \forall\, k \in \mathbb{Z},\ x \in \mathbb{R}. \tag{261}$$

- Aus Satz 51 und (261), folgt unmittelbar $(\frac{\pi}{2} + \mathbb{Z} \cdot \pi) \subseteq \mathcal{Z}$.
- Offensichtlich gilt $\mathbb{R} = \bigcup_{k \in \mathbb{Z}}(k \cdot \pi + [-\frac{\pi}{2}, \frac{\pi}{2}])$. Für $z \in \mathcal{Z}$ vorgegeben, existiert daher ein $k \in \mathbb{Z}$ mit $z \in \big(k \cdot \pi + [-\frac{\pi}{2}, \frac{\pi}{2}]\big)$. Wegen (261) gilt dann

  $$x := z - k \cdot \pi \in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}] \cap \mathcal{Z}.$$

  Aus der Definition von π (vgl. (260)) und der Symmetrie $\cos(x) = \cos(-x)$ für alle $x \in \mathbb{R}$ (vgl. Lemma 53.1)), folgt nun $x \in \{-\frac{\pi}{2}, \frac{\pi}{2}\}$; also $z = x + k \cdot \pi \in (\frac{\pi}{2} + \mathbb{Z} \cdot \pi)$.

Dies zeigt $\mathcal{Z} = \{x \in \mathbb{R} \mid \cos(x) = 0\} = \frac{\pi}{2} + \mathbb{Z} \cdot \pi$. Die Nullstellenmenge der Sinusfunktion erhalten wir nun mit Korollar 38.c). □

Gemäß der Funktionalgleichung (Satz 30.a)), ist die Exponentialfunktion $\exp\colon \mathbb{C} \to \mathbb{C}_\times$ ein Gruppenhomomorphismus der additiven Gruppe $(\mathbb{C}, +, 0)$ in die multiplikative Gruppe $(\mathbb{C}_\times, \cdot, 1)$. Wir bestimmen nun den Kern $\ker(\exp) = \exp^{-1}(1)$ dieses Gruppenhomomorphismus.

Lemma 54. *Es gilt* $\ker(\exp) = \{z \in \mathbb{C} \mid \exp(z) = 1\} = 2\pi\mathrm{i} \cdot \mathbb{Z}$.

Beweis. Für alle $n \in \mathbb{Z}$, folgt induktiv mit Korollar 38.a) (zweiter Schritt)

$$\exp(2\pi\mathrm{i} \cdot n) \overset{\text{(EF)}}{=} \cos(2\pi \cdot n) + \mathrm{i}\sin(2\pi \cdot n) = \cos(0) + \mathrm{i}\sin(0) \overset{\text{(EF)}}{=} \exp(0) = 1,$$

d.h., $2\pi\mathrm{i} \cdot \mathbb{Z} \subseteq \ker(\exp)$. Wir zeigen nun $\ker(\exp) \subseteq 2\pi\mathrm{i} \cdot \mathbb{Z}$. Sei hierfür $z = x + \mathrm{i}y \in \ker(\exp)$ mit $x, y \in \mathbb{R}$. Dann liefert die Funktionalgleichung

$$1 = \exp(z) = \exp(x + \mathrm{i}y) = \exp(x) \cdot \exp(\mathrm{i}y) \overset{\text{(EF)}}{=} \mathrm{e}^x \cdot \cos(y) + \mathrm{e}^x \cdot \mathrm{i}\sin(y).$$

Wegen $\mathrm{e}^x > 0$ (Satz 30.c)) folgt $\sin(y) = 0$ und $1 = \mathrm{e}^x \cdot \cos(y)$, d.h.,

$$\begin{aligned}
\sin(y) = 0 \quad &\overset{\text{Korollar 39}}{\Longrightarrow} \quad y = n\pi \text{ für ein } n \in \mathbb{Z} \\
&\overset{\text{Korollar 38.}b)}{\Longrightarrow} \quad 1 = \mathrm{e}^x \cdot \cos(y) = \mathrm{e}^x \cdot (-1)^n \\
&\overset{\text{Proposition 13.2)}}{\Longrightarrow} \quad x = 0 \quad \wedge \quad n = 2k \text{ für ein } k \in \mathbb{Z} \\
&\Longrightarrow \quad z = in\pi = 2\pi\mathrm{i} \cdot k \in 2\pi\mathrm{i} \cdot \mathbb{Z}.
\end{aligned}$$

Dies zeigt $\ker(\exp) \subseteq 2\pi\mathrm{i} \cdot \mathbb{Z}$. □

Satz 52 (Polardarstellung)**.** *Zu jedem $z \in \mathbb{C}_\times$ existiert genau ein $\phi \in [0, 2\pi)$ mit $z = |z| \cdot \mathrm{e}^{\mathrm{i}\phi}$.*

(Wegen Lemma 54 und Satz 30.a) (Funktionalgleichung), gilt dann ebenfalls $z = |z| \cdot \mathrm{e}^{\mathrm{i}(\phi+2\pi n)}$ für alle $n \in \mathbb{Z}$.)

Beweis. Die Eindeutigkeitsaussage erhalten wir mit

$$\begin{aligned}
|z| \cdot \mathrm{e}^{\mathrm{i}\phi} = |z| \cdot \mathrm{e}^{\mathrm{i}\psi} \quad \text{für} \quad \phi, \psi \in [0, 2\pi) \quad &\overset{|z|>0}{\Longrightarrow} \quad \mathrm{e}^{\mathrm{i}\phi} = \mathrm{e}^{\mathrm{i}\psi} \\
&\overset{\text{Satz 30}}{\Longrightarrow} \quad 1 = \mathrm{e}^{\mathrm{i}\phi} \cdot (\mathrm{e}^{\mathrm{i}\psi})^{-1} = \mathrm{e}^{\mathrm{i}\phi} \cdot \mathrm{e}^{-\mathrm{i}\psi} = \mathrm{e}^{\mathrm{i}(\phi-\psi)} \\
&\overset{\text{Lemma 54}}{\Longrightarrow} \quad (\phi - \psi) \in (-2\pi, 2\pi) \cap 2\pi \cdot \mathbb{Z} = \{0\},
\end{aligned}$$

wobei für die zweite Implikation Satz 30.b) und Satz 30.a) benutzt haben.

Um die Existenz zu zeigen, setzen wir $w := \frac{z}{|z|}$, d.h.,

$$z = |z| \cdot w \quad \text{und} \quad w = a + \mathrm{i}b \quad \text{für} \quad a, b \in \mathbb{R} \quad \text{mit} \quad 1 = |w| = a^2 + b^2.$$

Wir müssen nun zeigen, dass ein $\phi \in [0, 2\pi)$ existiert, sodass $w = a + \mathrm{i}b = \mathrm{e}^{\mathrm{i}\phi} \overset{\text{(EF)}}{=} \cos(\phi) + \mathrm{i}\sin(\phi)$ gilt. Mit Bemerkung 60.a) und Korollar 38 (Periodizitäten), erhalten wir

$$\cos(0) = 1 \qquad \wedge \qquad \cos(\pi) = -\cos(0) = -1.$$

Wegen $-1 \leq a \leq 1$ ($|a| = \sqrt{a^2} \leq \sqrt{a^2 + b^2} = 1$), existiert gemäß Satz 36 (Zwischenwertsatz) angewandt auf $\cos|_{[0,\pi]}$ ein $\phi \in [0, \pi]$ mit $a = \cos(\phi)$. Wir erhalten mit Lemma 53.2) (rechte Seite)

$$\begin{aligned}
b^2 = 1 - a^2 \quad &\Longrightarrow \quad |b| = \sqrt{1 - a^2} = \sqrt{1 - \cos(\phi)^2} = |\sin(\phi)| \\
&\Longrightarrow \quad b \in \{\sin(\phi), -\sin(\phi)\}.
\end{aligned}$$

- Sei $b = \sin(\phi)$. Dann gilt $a + \mathrm{i}b = \cos(\phi) + \mathrm{i}\sin(\phi)$ mit $\phi \in [0, \pi] \subseteq [0, 2\pi)$.
- Sei $b = -\sin(\phi)$. Dann zeigt Lemma 53.1), dass

$$a = \cos(\phi) = \cos(-\varphi) \qquad \text{und} \qquad b = -\sin(\phi) = \sin(-\phi)$$

also

$$a + \mathrm{i}b = \cos(-\phi) + \mathrm{i}\sin(-\phi) = \cos(2\pi - \phi) + \mathrm{i}\sin(2\pi - \phi)$$

wegen Korollar 38.a), mit $2\pi - \phi \in [\pi, 2\pi) \subseteq [0, 2\pi)$. □

Bemerkung 61 (Multiplikation in der Polardarstellung)**.** *Mit Hilfe der Polardarstellung lassen sich Produkte komplexer Zahlen sehr leicht berechnen. Sind nämlich $z = |z| \cdot \mathrm{e}^{\mathrm{i}\phi} \neq 0$ und $w = |w| \cdot \mathrm{e}^{\mathrm{i}\psi} \neq 0$ mit $\phi, \psi \in [0, 2\pi)$ gegeben, so gilt wegen der Funktionalgleichung:*

$$z \cdot w = |z| \cdot |w| \cdot \mathrm{e}^{\mathrm{i}\phi} \cdot \mathrm{e}^{\mathrm{i}\psi} = |zw| \cdot \mathrm{e}^{\mathrm{i}(\phi+\psi)}.$$

Es kann dann natürlich $2\pi \leq \phi + \psi < 4\pi$ gelten; aber dann erhalten wir

$$\mathrm{e}^{\mathrm{i}(\phi+\psi)} = \mathrm{e}^{\mathrm{i}(\phi+\psi)} \cdot \mathrm{e}^{-2\pi\mathrm{i}} = \mathrm{e}^{\mathrm{i}(\phi+\psi-2\pi)}$$

mit der Funktionalgleichung und Lemma 54), d.h., die Polardarstellung von $z \cdot w$ ist in diesem Fall gegeben durch $|zw| \cdot \mathrm{e}^{\mathrm{i}\chi}$ mit $\chi = \phi + \psi - 2\pi$.

Korollar 40 (n-te Einheitswurzeln)**.** *Für $n \in \mathbb{N}_{>0}$, hat die Gleichung $z^n = 1$ genau die n komplexen Lösungen*

$$\left\{ \mathrm{e}^{\frac{k}{n}\cdot 2\pi\mathrm{i}} \,\middle|\, k = 0, \ldots, n-1 \right\}. \tag{262}$$

Beweis. Offensichtlich gilt $\frac{k}{n} \cdot 2\pi \in [0, 2\pi)$ für $k = 0, \ldots, n-1$; und weiterhin gilt $\frac{p}{n} \cdot 2\pi \neq \frac{q}{n} \cdot 2\pi$ für $0 \leq p \neq q \leq n-1$. Wegen der Eindeutigkeitsaussage in Satz 52, ist daher die Menge (262) n-elementig. Zudem zeigen die Funktionalgleichung und Lemma 54, dass

$$\left(\mathrm{e}^{\frac{k}{n}\cdot 2\pi\mathrm{i}}\right)^n = \mathrm{e}^{2\pi\mathrm{i}\cdot k} = 1$$

gilt; also sind alle Elemente von (262), alle Lösungen der Gleichung $z^n = 1$.

Ist nun umgekehrt $z \in \mathbb{C}$ mit $z^n = 1$ vorgegeben, so gilt $|z|^n = |z^n| = 1$, d.h., $|z| = 1$. Gemäß Satz 52, existiert daher ein $\phi \in [0, 2\pi)$ mit $z = \mathrm{e}^{\mathrm{i}\phi}$. Die Funktionalgleichung liefert

$$1 = z^n = (\mathrm{e}^{\mathrm{i}\phi})^n = \mathrm{e}^{\mathrm{i}n\cdot\phi}.$$

Gemäß Korollar 54, gilt daher $n \cdot \phi = 2\pi \cdot k$ für ein $k \in \mathbb{Z}$. Wegen $n \cdot \phi \geq 0$, gilt dann $k \geq 0$; und wegen $n \cdot \phi \in [0, 2\pi \cdot n)$, gilt $k \leq n-1$. Folglich gilt $\phi = \frac{k}{n} \cdot 2\pi$ mit $0 \leq k \leq n-1$. □

Übung 104. *Sei $n \in \mathbb{N}_{>0}$ und $w \in \mathbb{C}_\times$ vorgegeben. Bestimmen Sie alle komplexen Lösungen der Gleichung $z^n = w$.*

8.4.3 Mehr über trigonometrische Funktionen

Übung 105. *Zeigen Sie, dass $\sin(x) > 0$ für alle $x \in (0, \pi)$ gilt.*

Hinweis: Wegen Korollar 39 hat die Sinusfunktion keine Nullstelle in $(0, \pi)$. Nehmen Sie nun an, dass ein $x \in (0, \pi)$ mit $\sin(x) < 0$ existiert; und verwenden Sie Satz 51 und Satz 36, um einen Widerspruch herzuleiten.

Terminologie 25 (Tangens und Arkustangens)**.** *Die Tangensfunktion ist definiert durch*

$$\tan\colon \mathcal{B} \to \mathbb{R}, \qquad x \mapsto \frac{\sin(x)}{\cos(x)}$$

mit $\mathcal{B} := \mathbb{R} \setminus (\frac{\pi}{2} + \pi \cdot \mathbb{Z})$ (vgl. Korollar 39). Es gelten die folgenden Sachverhalte:

(1) *Die Tangensfunktion ist stetig.* *(Satz 32.2))*

(2) $\tan(x + n \cdot \pi) = \tan(x)$ *für alle $n \in \mathbb{Z}$ und $x \in \mathcal{B}$.* *(wegen Korollar 38.b))*

(3) *Die Tangensfunktion ist differenzierbar, mit $\tan(x)' = \frac{1}{\cos(x)^2} > 0$ für alle $x \in \mathcal{B}$:*

Wegen Satz 49, Korollar 31 und Satz 40.4), ist $\tan$ *differenzierbar mit*

$$\tan' = \frac{\sin' \cdot \cos - \sin \cdot \cos'}{\cos^2} = \frac{\cos^2 + \sin^2}{\cos^2} = \frac{1}{\cos^2}.$$

Wegen $1 + \tan^2 = \frac{\cos^2 + \sin^2}{\cos^2} = \frac{1}{\cos^2}$ genügt die Tangensfunktion somit der Differentialgleichung

$$\tan'(x) = 1 + \tan(x)^2 \qquad \forall\, x \in \mathcal{B}.$$

(4) $\tan(x) \cdot \tan(x \pm \frac{\pi}{2}) = -1$ *für alle* $x \in \mathcal{B}$. *(wegen Korollar 38.c))*

(5) $\tan(-x) = -\tan(x)$ *für alle* $x \in \mathcal{B}$. *(wegen Lemma 53.*1*))*

(6) $\tan|_{(-\frac{\pi}{2},\frac{\pi}{2})} \colon (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ *ist bijektiv:*

- *Wegen Lemma 46 (Einschränkungen) ist* $\tan_0 := \tan|_{(-\frac{\pi}{2},\frac{\pi}{2})}$ *differenzierbar, mit* $\tan_0' > 0$ *wegen* (3). *Wegen Korollar 32.3) ist daher* $\tan_0$ *streng monoton wachsend*
- *Wegen Lemma 47 (Einschränkungen) ist* $\tan_0$ *stetig, also injektiv nach Lemma 43 (da* $\tan_0$ *monoton wachsend ist).*
- $I := \mathrm{im}(\tan_0)$ *ist ein Intervall wegen Korollar 28 (da* $\tan_0$ *stetig ist).*

Mit (4) *erhalten wir (Übung 105 im letzten Schritt)*

$$\lim_{x\downarrow -\frac{\pi}{2}} \tan_0(x) = \lim_{x\downarrow 0} \tan(x - \tfrac{\pi}{2}) = \lim_{x\downarrow 0} -\tfrac{1}{\tan(x)} = \lim_{x\downarrow 0} -\tfrac{\cos(x)}{\sin(x)} = -\infty,$$

da $\cos(0) = 1$ *und* $\sin(0) = 0$ *gilt. Wegen* $\tan_0(0) = 0$*, folgt nun* $(-\infty, 0] \subseteq I$*; und wegen* $\tan_0(-x) = -\tan_0(x)$ *für alle* $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$*, gilt dann ebenfalls* $[0, \infty) \subseteq I$*. Dies zeigt* $I = \mathbb{R}$*.*

(7) *Mit* (6)*,* (3)*, sowie Satz 42 (Differenzierbarkeit der Umkehrfunktion), erhalten wir die differenzierbare Umkehrabbildung*

$$\arctan := \tan_0^{-1} \colon \mathbb{R} \to \big(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\big),$$

mit $\arctan(0) = 0$ *und*

$$\arctan'(x) = \frac{1}{\tan_0'(\arctan(x))} \overset{(3)}{=} \frac{1}{1 + \tan_0(\arctan(x))^2} = \frac{1}{1 + x^2} \qquad \forall\, x \in \mathbb{R}.$$

Bemerkung 62 (Arkussinus und Arkuskosinus)**.**

a) Wir betrachten die stetige Abbildung $\cos_0 := \cos|_{[0,\pi]}$*:*

- *Es gilt* $\{-1, 1\} = \{\cos(\pi), \cos(0)\} \subseteq \mathrm{im}(\cos_0) \subseteq [-1, 1]$*. (Bemerkung 60 und Korollar 38.b))*
- *Wegen Korollar 28, ist* $\mathrm{im}(\cos_0)$ *ein Intervall, d.h., es gilt* $\mathrm{im}(\cos_0) = [-1, 1]$*.*
- *Wegen Satz 49 und Lemma 46 ist* $\cos_0$ *differenzierbar mit* $\cos_0' = -\sin|_{[0,\pi]}$*, also* $\cos_0'|_{(0,\pi)} < 0$ *wegen Übung 105. Wegen Korollar 32.3) ist daher* $\cos_0$ *streng monoton fallend auf* $(0, \pi)$*; und somit auch auf* $[0, \pi]$*.*

  Beweis der Behauptung: Es gilt $\cos(0) = 1 > -1 = \cos(\pi)$. Gilt nun beispielsweise $1 = \cos(0) \le \cos(x)$ für ein $x \in (0, \pi)$, so folgt $\cos(y) > \cos(x) \ge 1$ für alle $0 < y < x$, was $\mathrm{im}(\cos_0) = [-1, 1]$ widerspricht. Analog führt man die Annahme $-1 = \cos(\pi) \ge \cos(x)$ für ein $x \in (0, \pi)$ zum Widerspruch. □

  Wegen Lemma 43 ist $\cos_0$ *daher injektiv.*
- *Wegen Satz 37, ist die Arkuskosinusfunktion (Umkehrfunktion von* $\cos_0$*)*

$$\arccos := \cos_0^{-1}\, [-1, 1] \to [0, \pi]$$

streng monoton fallend; und wegen Satz 42 differenzierbar auf $(-1, 1)$ *(beachte* $\cos_0'|_{(0,\pi)} \neq 0$ *wegen* $\cos_0'|_{(0,\pi)} < 0$*), mit Ableitung (beachte* $\sin|_{(0,\pi)} > 0$ *wegen Übung 105)*

$$\arccos'(x) = \frac{1}{\cos'(\arccos(x))} = -\frac{1}{\underbrace{\sin(\arccos(x))}_{>0}} = -\frac{1}{\sqrt{1 - \cos(\arccos(x))^2}} = -\frac{1}{\sqrt{1 - x^2}}$$

für alle $x \in (-1, 1)$*.*

b) Wir betrachten die stetige Abbildung $\sin_0 := \sin|_{[-\frac{\pi}{2},\frac{\pi}{2}]}$*:*

- *Es gilt* $\{-1,1\} = \{\sin(-\frac{\pi}{2}), \sin(\frac{\pi}{2})\} \subseteq \operatorname{im}(\sin_0) \subseteq [-1,1]$*. (Bemerkung 60 und Korollar 38.c))*
- *Wegen Korollar 28 ist* $\operatorname{im}(\sin_0)$ *ein Intervall, d.h., es gilt* $\operatorname{im}(\sin_0) = [-1,1]$.
- *Wegen Satz 49 und Lemma 46 ist* $\sin_0$ *differenzierbar mit* $\sin_0' = \cos|_{[-\frac{\pi}{2},\frac{\pi}{2}]}$*. Wegen Korollar 38.c) und Übung 105 gilt*

$$\cos|_{(-\frac{\pi}{2},\frac{\pi}{2})}(x) = \sin(x+\tfrac{\pi}{2}) > 0 \qquad \forall\, x \in (-\tfrac{\pi}{2},\tfrac{\pi}{2}).$$

Gemäß Satz 32.3) ist somit $\sin_0$ *streng monoton wachsend auf* $(-\frac{\pi}{2},\frac{\pi}{2})$*, und somit auch auf* $[-\frac{\pi}{2},\frac{\pi}{2}]$ *(gleiches Argument wie in a)). Wegen Lemma 43 ist* $\sin_0$ *somit injektiv.*
- *Wegen Satz 37 ist die Arkussinusfunktion (Umkehrfunktion von* $\sin_0$*)*

$$\arcsin := \sin_0^{-1}\, [-1,1] \to [-\tfrac{\pi}{2},\tfrac{\pi}{2}]$$

streng monoton wachsend. Mit Korollar 38.c) erhalten wir

$$\sin_0(x) = \cos(x-\tfrac{\pi}{2}) = \cos(\tfrac{\pi}{2}-x) = \cos_0(\tfrac{\pi}{2}-x) \qquad \forall\, x \in [-\tfrac{\pi}{2},\tfrac{\pi}{2}].$$

Für $x = \arcsin(y)$ *mit* $y \in [-1,1]$*, folgt hiermit*

$$\begin{aligned} y = \cos(\tfrac{\pi}{2} - \arcsin(y)) \qquad &\Longrightarrow \qquad \arccos(y) = \tfrac{\pi}{2} - \arcsin(y) \\ &\Longrightarrow \qquad \arcsin(y) = \tfrac{\pi}{2} - \arccos(y). \end{aligned}$$

Wegen a) ist daher $\arcsin$ *differenzierbar auf* $(-1,1)$*, mit*

$$\arcsin'(x) = -\arccos'(x) = \frac{1}{\sqrt{1-x^2}} \qquad \forall\, x \in (-1,1).$$

Beispiel 73 (Trigonometrische Grenzwerte $*$)**.** *a) Es existiert* $\lim_{x\to\pm\infty}\cos(x)$ *nicht, da die Folge* $\cos(\pm n\pi) = (-1)^n$ *nicht konvergiert. Daher existiert der Grenzwert* $\lim_{x\to 0}\cos(\frac{1}{x})$ *ebenfalls nicht.*

b) Wir betrachten die Funktion

$$f : \mathbb{R} \to \mathbb{R}, \qquad x \mapsto \begin{cases} x \cdot \cos(\frac{1}{x}) & \text{für} \quad x \neq 0 \\ 0 & \text{für} \quad x = 0. \end{cases}$$

Es ist $f|_{(-\infty,0)}$ *sowie* $f|_{(0,\infty)}$ *differenzierbar als Produkt und Verkettung differenzierbarer Abbildungen. Gemäß Übung 92.b) ist daher* f *differenzierbar in jedem* $x \in \mathbb{R}_\times$*. Im Nullpunkt ergibt sich aus* $|f(x)| \leq |x|$ *für* $x \in \mathbb{R}$*, sofort*

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

Also ist f *auf ganz* $\mathbb{R}$ *stetig. Allerdings ist* f *in* 0 *nicht differenzierbar, denn*

$$\lim_{h\to 0}\frac{f(h)-f(0)}{h} = \lim_{h\to 0}\cos\left(\frac{1}{h}\right)$$

existiert nicht wegen Teil a).

c) Die Funktion

$$f : \mathbb{R} \to \mathbb{R}, \qquad x \mapsto \begin{cases} x^2 \cdot \sin\left(\frac{1}{x}\right) & \text{für} \quad x \neq 0 \\ 0 & \text{für} \quad x = 0 \end{cases}$$

ist differenzierbare, hat aber eine (in 0*) unstetige Ableitung:*

- *Für $x \neq 0$ ist*

$$f'(x) = 2x \cdot \sin\left(\frac{1}{x}\right) - \frac{x^2}{x^2} \cdot \cos\left(\frac{1}{x}\right) = 2x \cdot \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right).$$

- *Für $x = 0$ ist*

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{f(h)}{h} = \lim_{h\to 0} \frac{1}{h} \cdot h^2 \cdot \sin\left(\frac{1}{h}\right) = \lim_{h\to 0} h \cdot \sin\left(\frac{1}{h}\right) = 0,$$

wegen $|\sin\left(\frac{1}{h}\right)| \leq 1$.

Somit ist f auf ganz $\mathbb{R}$ differenzierbar mit

$$f'(x) = \begin{cases} 2x \cdot \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right) & \textit{für} \quad x \neq 0 \\ 0 & \textit{für} \quad x = 0. \end{cases}$$

Wir betrachten die Nullfolge $(x_n)_{n\in\mathbb{N}_{>0}} \to 0$, gegeben durch $x_n := \frac{1}{2\pi n}$ für alle $n \in \mathbb{N}_{>0}$; und erhalten (mit den bisher hergeleiteten Eigenschaften von Sinus und Kosinus)

$$\lim_n f'(x_n) = \lim_n f'\left(\tfrac{1}{2\pi n}\right) = \lim_n \left(\tfrac{2}{2\pi n} \cdot \sin(2\pi n) - \cos(2\pi n)\right) = \lim_n -1 = -1 \neq 0 = f'(0),$$

womit f' im Nullpunkt unstetig ist.

Bemerkung 63 (*). *Die Funktionenfolge $(f_n)_{n\in\mathbb{N}_{>0}}$ gegeben durch $f_n(x) = \frac{1}{n} \cdot \sin(n^2 \cdot x)$ für alle $n \in \mathbb{N}_{>0}$, konvergiert auf $\mathbb{R}$ gleichmäßig gegen die Nullfunktion; denn es gilt*

$$\lim_n |f_n|_\infty = \lim_n \sup\{|f_n(x)| \mid x \in \mathbb{R}\} = \lim_n \tfrac{1}{n} = 0.$$

Wegen $f_n'(0) = n \cdot \cos(n^2 \cdot 0) = n$ für alle $n \in \mathbb{N}_{>0}$, divergiert zudem die Folge $(f_n'(0))_{n\in\mathbb{N}_{>0}}$. Gleichmäßige Konvergenz erzwingt daher nicht einmal die punktweise Konvergenz der Folge der Ableitungen.

9 Integralrechnung

Im letzten Kapitel haben wir die Differentialrechnung kennengelernt, die es erlaubt die Änderungsrate einer Funktion durch deren Ableitung mathematisch präzise zu beschreiben. In diesem Kapitel wenden wir uns der Frage zu, wie man für eine Abbildung $f\colon [a,b] \to [0,\infty]$ $(a < b)$ und ein gegebenes $z \in [a,b]$ die Fläche

$$F[z] = \{(x,y) \mid x \in [a,z] \quad \wedge \quad 0 \leq y \leq f(x)\}$$

berechnen kann, die der Funktionsgraph von $f|_{[a,z]}$ mit der x-Achse einschließt. Der Hauptsatz der Differential- und Integralrechnung besagt $F'(z) = f(z)$, d.h. die Änderungsrate der Fläche $F[z]$ bei fortschreitender rechter Grenze des Definitionsbereichs ist einfach gegeben durch die Funktion selbst. Dies ist eines der zentralen Resultate der Analysis von Funktionen einer Veränderlichen. Wir werden sehen, wie sich dieser Satz dazu verwenden lässt, viele konkrete Integrale explizit zu berechnen. Über dies hinaus werden wir sehen, dass die Berechnung des Flächeninhaltes für stetige Funktionen f immer möglich ist. Als problematisch erweisen sich Funktionen, die „zu viele Unstetigkeitsstellen“ haben, wie zum Beispiel die *Dirichletfunktion*

$$f : [0,1] \to \mathbb{R}, \qquad x \mapsto \begin{cases} 1 & \text{für} \quad x \in \mathbb{Q} \cap [0,1] \\ 0 & \text{für} \quad x \in [0,1] \setminus \mathbb{Q}, \end{cases}$$

die in keinem Punkt stetig ist.

9.1 Das Integral von Treppenfunktionen

Wir beginnen unsere Untersuchungen mit den sogenannten Treppenfunktionen:

Definition 62. *Seien Seien $a, b \in \mathbb{R}$ mit $a < b$ vorgegeben.*

1) *Eine* Zerlegung des Intervalls $[a, b]$ *ist ein* $(n+1)$*-Tupel*

$$Z = (z_0, \ldots, z_n) \qquad \text{mit} \qquad a = z_0 < z_1 < z_2 < \ldots < z_n = b \qquad \text{für} \qquad n \in \mathbb{N}_{>0}.$$

Die Zerlegung $Z = (z_0, \ldots, z_n)$ heißt feiner *als eine Zerlegung $W = (w_0, \ldots, w_m)$ (von $[a, b]$), wenn $\{w_0, \ldots, w_m\} \subseteq \{z_0, \ldots, z_n\}$ gilt, d.h., wenn jeder Unterteilungspunkt von W auch ein Unterteilungspunkt von Z ist. Es wir dann Z auch als Verfeinerung von W bezeichnet.*

Sind Z und Z' zwei Zerlegungen von $[a, b]$, so notiert $Z \cup Z'$ diejenige Zerlegung, die durch Vereinigung der Mengen der Unterteilungspunkte von Z und Z' entsteht. Sie ist eine Verfeinerung sowohl von Z als auch von Z'.

2) *Eine Abbildung $\chi\colon [a, b] \to \mathbb{R}$ heißt* Treppenfunktion *auf $[a, b]$, wenn eine Zerlegung $Z = (z_0, \ldots, z_n)$ von $[a, b]$ existiert, sowie (notwendig eindeutige) $o_1, \ldots, o_n \in \mathbb{R}$ mit*[55]

$$\chi|_{(z_{k-1}, z_k)} = o_k \quad \text{für} \quad 1 \leq k \leq n. \tag{263}$$

Wir sprechen dann von einer Treppenfunktion zur Zerlegung Z (von $[a, b]$). Von den Funktionswerten an den Unterteilungspunkten $z_0, \ldots, z_n$ wird nichts verlangt. Wir notieren den Raum aller Treppenfunktionen auf $[a, b]$ mit

$$\mathcal{T}_a^b := \{\chi\colon [a, b] \to \mathbb{R} \mid \chi \text{ ist Treppenfunktion auf } [a, b]\}.$$

Offensichtlich gilt $\mathcal{T}_a^b \subseteq \mathrm{B}([a, b], \mathbb{R})$, d.h., jede Treppenfuktion ist beschränkt.

Beispiel 74. *Für $\mathbb{R} \ni a < b \in \mathbb{R}$ und $c \in \mathbb{R}$ ist die konstante Funktion $\chi_c\colon [a, b] \ni x \mapsto c \in \mathbb{R}$ eine Treppenfunktion, d.h., $\chi_c \in \mathcal{T}_a^b$.*

Bemerkung 64. *Sei $\mathbb{R} \ni a < b \in \mathbb{R}$, $Z = (z_0, \ldots, z_n)$ eine Zerlegungen von $[a, b]$ und $\chi \in \mathcal{T}_a^b$ eine Treppenfunktion bezüglich Z mit $o_1, \ldots, o_n \in \mathbb{R}$ wie in* (263)*.*

a) Sei $z_{\ell-1} < u < z_\ell$ für ein $1 \leq \ell \leq n$:

- *$U := (z_0, \ldots, z_{\ell-1}, u, z_\ell, \ldots, z_n)$ ist eine Verfeinerung von Z.*
- *χ ist eine Treppenfunktion zu U, mit Konstanten $o_1, \ldots, o_\ell, o_\ell, o_{\ell+1}, \ldots, o_n \in \mathbb{R}$.*

b) Seien $a \leq a' < b' \leq b$ vorgegeben. Dann existiert $0 \leq \ell_- \leq n$ maximal mit $z_{\ell_-} \leq a'$, sowie $0 \leq \ell_+ \leq n$ minimal mit $b' \leq z_{\ell_+}$. Wegen $a' < b'$, gilt dann $\ell_- < \ell_+$:

- *$W := (a', z_{\ell_-+1}, \ldots, z_{\ell_+-1}, b')$ ist eine Zerlegung von $[a', b']$.*
- *$\chi|_{[a',b']} \in \mathcal{T}_{a'}^{b'}$ ist eine Treppenfunktion zu W, mit Konstanten $o_{\ell_-+1}, \ldots, o_{\ell_+}$.*

Übung 106. *Sei $\mathbb{R} \ni a < b \in \mathbb{R}$, sowie Z, Z', U Zerlegungen von $[a, b]$. Sei weiterhin $\chi \in \mathcal{T}_a^b$ eine Treppenfunktion zu Z. Zeigen Sie (in der gegeben Reihenfolge):*

a) Ist U eine Verfeinerung von Z, so ist χ auch eine Treppenfunktion zu U.

Hinweis: Induktion und Bemerkung 64.a).

b) Ist χ eine Treppenfunktion zur Zerlegung Z', so auch zur Zerlegung $Z \cup Z'$.

[55] Ist $t \in (z_{k-1}, z_k)$ für $1 \leq k \leq n$, so gilt $\chi(t) = o_k$.

Folgern Sie, dass zu je zwei Treppenfunktionen $\phi, \psi \in \mathcal{T}_a^b$ eine Zerlegung von $[a,b]$ existiert, bezüglich derer sowohl ϕ als auch ψ eine Treppenfunktion ist

Lemma 55. *Sei $\mathbb{R} \ni a < c < b \in \mathbb{R}$.*

a) $\mathcal{T}_a^b$ ist ein $\mathbb{R}$-Vektorraum (Untervektorraum von $\mathrm{B}([a,b],\mathbb{R})$), d.h., für $\phi, \psi \in \mathcal{T}_a^b$ und $\lambda \in \mathbb{R}$, sind $\phi + \psi$ und $\lambda \cdot \phi$ wieder Elemente von $\mathcal{T}_a^b$.

b) Für $\chi \in \mathcal{T}_a^b$ gilt $\chi|_{[a,c]} \in \mathcal{T}_a^c$ sowie $\chi|_{[c,b]} \in \mathcal{T}_c^b$.

c) Seien $\chi_- \in \mathcal{T}_a^c$ und $\chi_+ \in \mathcal{T}_c^b$ sowie $\tau \in \mathbb{R}$. Dann gilt $\chi \in \mathcal{T}_a^b$, für $\chi\colon [a,b] \to \mathbb{R}$ gegeben durch

$$\chi(x) := \begin{cases} \chi_-(x) & \text{für} \quad x \in [a,c) \\ \tau & \text{für} \quad x \in c \\ \chi_+(x) & \text{für} \quad x \in (c,b]. \end{cases}$$

Beweis. a) Sei ϕ eine Treppenfunktionen zu $Z = (z_0, \ldots, z_n)$, mit $o_1, \ldots o_n \in \mathbb{R}$ wie in (263). Sei zudem ψ eine Treppenfunktionen zu Z'.

- Es gilt $(\lambda \cdot \phi)|_{(z_{k-1},z_k)} = \lambda \cdot o_k$ für $1 \leq k \leq n$; also ist $\lambda \cdot \phi$ eine Treppenfunktionen zu Z, mit Konstanten $\lambda \cdot o_1, \ldots, \lambda \cdot o_n \in \mathbb{R}$.
- Wegen Übung 106.b) sind ϕ, ψ Treppenfunktionen zu $Z \cup Z' = (u_0, \ldots, u_m)$ $(m \in \mathbb{N}_{>0})$, d.h., es existieren $o_1, \ldots, o_m, \tilde{o}_1, \ldots, \tilde{o}_m \in \mathbb{R}$ mit

$$\phi|_{(u_{k-1},u_k)} = o_k \qquad \forall\, 1 \leq k \leq m \qquad \wedge \qquad \psi|_{(u_{k-1},u_k)} = \tilde{o}_k \qquad \forall\, 1 \leq k \leq m$$

also

$$(\phi + \psi)|_{(u_{k-1},u_k)} = o_k + \tilde{o}_k \qquad \forall\, 1 \leq k \leq m.$$

Daher ist $\phi + \psi$ eine Treppenfunktion zu $Z \cup Z'$, mit Konstanten $o_1 + \tilde{o}_1, \ldots, o_m + \tilde{o}_m \in \mathbb{R}$.

b) Klar wegen Bemerkung 64.b).

c) Sei χ_- eine Treppenfunktion zur Zerlegung $Z_- = (z_0^-, \ldots, z_n^-)$ von $[a,c]$, mit den Konstanten $(o_1^-, \ldots, o_n^-)$. Sei zudem χ_+ eine Treppenfunktion zur Zerlegung $Z_+ = (z_0^+, \ldots, z_m^+)$ von $[c,b]$, mit Konstanten $(o_1^+, \ldots, o_m^+)$. Dann gilt $z_n^- = c = z_0^+$; und χ ist eine Treppenfunktion zur Zerlegung $U := (z_0^-, \ldots, z_{n-1}^-, c, z_1^+, \ldots, z_m^+)$ von $[a,b]$, mit Konstanten $o_1^-, \ldots, o_n^-, o_1^+, \ldots, o_m^+$. $\square$

Sei $\mathbb{R} \ni a < b \in \mathbb{R}$, sowie $\chi \in \mathcal{T}_a^b$ eine Treppenfunktion zur Zerlegung $Z = (z_0, \ldots, z_n)$ von $[a,b]$, mit $o_1, \ldots, o_n$ wie in (263). In diesem Fall definieren wir

$$\mathcal{S}_Z(\chi) := \textstyle\sum_{k=1}^n o_k \cdot (z_k - z_{k-1}).$$

Lemma 56. *Sei $\mathbb{R} \ni a < b \in \mathbb{R}$, und $\chi \in \mathcal{T}_a^b$ eine Treppenfunktion zur Zerlegungen Z von $[a,b]$. Ist U eine Verfeinerung von Z, so gilt $\mathcal{S}_U(\chi) = \mathcal{S}_Z(\chi)$.*

($\mathcal{S}_U(\chi)$ ist definiert; denn wegen Übung 106.a) ist χ eine Treppenfunktion zu U.)

Beweis. Sei nun $Z = (z_0, \ldots, z_n)$, sowie $o_1, \ldots, o_n$ wie in (263). Die allgemeine Aussage folgt induktiv aus dem Fall, dass U durch Hinzufügen eines Punktes zu Z entsteht, d.h.

$$U = (z_0, \ldots, z_{\ell-1}, u, z_\ell, \ldots, z_n) \qquad \text{mit} \qquad 1 \leq \ell \leq n \qquad \text{und} \qquad z_{\ell-1} < u < z_\ell.$$

In diesem Fall ist, gemäß Bemerkung 64.a), χ eine Treppenfunktion zu U mit den Konstanten $o_1, \ldots, o_\ell, o_\ell, o_{\ell+1}, \ldots, o_n \in \mathbb{R}$; und es folgt

$$\mathcal{S}_U(\chi) = \textstyle\sum_{k=1}^{\ell-1} o_k \cdot (z_k - z_{k-1}) + \underbrace{o_\ell \cdot (u - z_{\ell-1}) + o_\ell \cdot (z_\ell - u)}_{o_\ell \cdot (z_\ell - z_{\ell-1})} + \textstyle\sum_{k=\ell+1}^n o_k \cdot (z_k - z_{k-1}) = \mathcal{S}_Z(\chi),$$

was die Behauptung zeigt. $\square$

Lemma 57. *Sei* $\mathbb{R} \ni a < b \in \mathbb{R}$, *und* $\chi \in \mathcal{T}_a^b$ *eine Treppenfunktion zu den Zerlegungen* Z, Z' *von* $[a,b]$. *Dann gilt* $\mathcal{S}_Z(\chi) = \mathcal{S}_{Z'}(\chi)$.

Beweis. Gemäß Übung 106.b) ist χ eine Treppenfunktion zu $Z \cup Z'$. Es reicht daher

$$\mathcal{S}_{Z\cup Z'}(\chi) = \mathcal{S}_Z(\chi) \qquad \text{sowie} \qquad \mathcal{S}_{Z\cup Z'}(\chi) = \mathcal{S}_{Z'}(\chi).$$

nachzuweisen. Dies ist aber klar wegen Lemma 56; denn $Z \cup Z'$ ist sowohl eine Verfeinerung von Z als auch von Z'. □

Definition 63. $\mathbb{R} \ni a < b \in \mathbb{R}$, *und* $\chi \in \mathcal{T}_a^b$ *eine Treppenfunktion zur Zerlegungen* Z *von* $[a,b]$. *Wir definieren das* Riemann-Integral *von* χ *durch*

$$\int_a^b \chi := \mathcal{S}_Z(\chi). \tag{264}$$

Wegen Lemma 57 ist diese Definition unabhängig von der expliziten Wahl der Zerlegung Z. *Die Zahl* a *heißt die untere Grenze des Integrals (untere Integrationsgrenze), die Zahl* b *heißt die obere Grenze des Integrals (obere Integrationsgrenze). Eine weitere Schreibweise für das Integral von* χ *ist* $\int_a^b \chi(x)\,\mathrm{d}x$, *wobei hier die explizite Bezeichnung der Integrationsvariable* x *keine Rolle spielt.*[56] *Wir definieren*

$$\int_c^c \chi := 0 \qquad \textit{sowie} \qquad \int_b^a \chi := -\int_a^b \chi \qquad \textit{für alle} \qquad a \le c \le b. \tag{265}$$

Zudem definieren wir *(beachte* $\chi|_{[a',b']} \in \mathcal{T}_{a'}^{b'}$ *gemäß Bemerkung 64.b).)*

$$\int_{a'}^{b'} \chi := \int_{a'}^{b'} \chi|_{[a',b']} \qquad \forall\, a \le a' < b' \le b. \tag{266}$$

(Gegeben eine Menge Z und Abbildungen $\phi, \psi \colon Z \to \mathbb{R}$, so schreiben wir $\phi \le \psi$ genau dann, wenn $\phi(z) \le \psi(z)$ für alle $z \in Z$ gilt.)

Proposition 15. *Sei* $\mathbb{R} \ni a < b \in \mathbb{R}$. *Das Integral von Treppenfunktionen hat die folgende Eigenschaften:*

(I1) *Intervalladditivität:* $\int_a^b \chi = \int_a^c \chi + \int_c^b \chi$ *für alle* $a \le c \le b$ *und* $\chi \in \mathcal{T}_a^b$.

(I2) *Monotonie:* $\int_a^b \phi \le \int_a^b \psi$ *für* $\mathcal{T}_a^b \ni \phi \le \psi \in \mathcal{T}_a^b$.

(I3) *Linearität:* $\int_a^b (\lambda\cdot\phi + \mu\cdot\psi) = \lambda\cdot\int_a^b \phi + \mu\cdot\int_a^b \psi$ *für alle* $\lambda, \mu \in \mathbb{R}$ *und* $\phi, \psi \in \mathcal{T}_a^b$.

(I4) *Normierung:* $\int_a^b \chi_c = c\cdot(b-a)$ *für alle* $c \in \mathbb{R}$.

Beweis. I1) Sei χ eine Treppenfunktion zur Zerlegung $Z = (z_0, \dots, z_n)$ von $[a,b]$, mit $o_1, \dots, o_n \in \mathbb{R}$ wie in (263). Es können die folgenden beiden Fälle auftreten:

- Sei $c = z_\ell$ für ein $0 \le \ell \le n$. Die Behauptung ist klar, falls $\ell = 0$ oder $\ell = n$ gilt. Andernfalls gilt $n \ge 2$ mit $0 < \ell < n$; und es sind $Z_- := (z_0, \dots, z_\ell)$ bzw. $Z_+ := (z_\ell, \dots, z_n)$ Zerlegungen von $[a,c]$ bzw. $[c,b]$. Weiterhin sind $\chi|_{[a,c]}$ bzw. $\chi|_{[c,b]}$ Treppenfunktionen zu Z_- bzw. Z_+, mit Konstanten $o_1, \dots, o_\ell \in \mathbb{R}$ bzw. $o_{\ell+1}, \dots, o_n \in \mathbb{R}$, d.h.

$$\begin{aligned}\textstyle\int_a^b \chi &= \underbrace{\textstyle\sum_{k=1}^n o_k\cdot(z_k - z_{k-1})}_{\mathcal{S}_Z(\chi)}\\ &= \underbrace{\textstyle\sum_{k=1}^\ell o_k\cdot(z_k - z_{k-1})}_{\mathcal{S}_{Z_-}(\chi|_{[a,c]})} + \underbrace{\textstyle\sum_{k=\ell+1}^n o_k\cdot(z_k - z_{k-1})}_{\mathcal{S}_{Z_+}(\chi|_{[c,b]})} = \textstyle\int_a^c \chi + \int_c^b \chi.\end{aligned}$$

[56]Man kann beispielsweise also auch t oder s anstatt x benutzen.

- Es gilt $z_\ell < c < z_{\ell+1}$ für ein $0 \leq \ell \leq n$. Es sind $Z_- := (z_0, \ldots, z_\ell, c)$ bzw. $Z_+ := (c, z_{\ell+1}, \ldots, z_n)$ Zerlegungen von $[a,c]$ bzw. $[c,b]$. Weiterhin sind $\chi|_{[a,c]}$ bzw. $\chi|_{[c,b]}$ Treppenfunktionen zu Z_- bzw. Z_+, mit Konstanten $o_1, \ldots, o_\ell, o_{\ell+1} \in \mathbb{R}$ bzw. $o_{\ell+1}, \ldots, o_n \in \mathbb{R}$, d.h.,

$$\begin{aligned}\int_a^b \chi &= \underbrace{\textstyle\sum_{k=1}^n o_k \cdot (z_k - z_{k-1})}_{\mathcal{S}_Z(\chi)} \\ &= \textstyle\sum_{k=1}^\ell o_k \cdot (z_k - z_{k-1}) + \overbrace{o_{\ell+1} \cdot (c - z_\ell) \quad + \quad o_{\ell+1} \cdot (z_{\ell+1} - c)}^{o_\ell \cdot (z_{\ell+1} - z_\ell)} + \textstyle\sum_{k=\ell+2}^n o_k \cdot (z_k - z_{k-1}) \\ &= \qquad \mathcal{S}_{Z_-}(\chi|_{[a,c]}) \qquad + \qquad \mathcal{S}_{Z_+}(\chi|_{[c,b]}) \\ &= \textstyle\int_a^c \chi|_{[a,c]} + \int_c^b \chi|_{[c,b]} = \int_a^c \chi + \int_c^b \chi.\end{aligned}$$

I2) und (I3) Sei $Z = (z_0, \ldots, z_n)$ eine Zerlegung von $[a,b]$, bezüglich derer ϕ und ψ Treppenfunktionen sind (Übung 106), mit entsprechenden Konstanten $o_1, \ldots, o_n$ und $\tilde{o}_1, \ldots, \tilde{o}_n$, sodass gilt:

$$\phi|_{(z_{k-1}, z_k)} = o_k \qquad \forall\, 1 \leq k \leq m \qquad \wedge \qquad \psi|_{(z_{k-1}, z_k)} = \tilde{o}_k \qquad \forall\, 1 \leq k \leq m$$

also

$$(\lambda \cdot \phi + \mu \cdot \psi)|_{(z_{k-1}, z_k)} = \lambda \cdot o_k + \mu \cdot \tilde{o}_k \qquad \forall\, 1 \leq k \leq m.$$

Wir erhalten

$$\begin{aligned}\textstyle\int_a^b (\lambda \cdot \phi + \mu \cdot \psi) &= \overbrace{\textstyle\sum_{k=1}^n (\lambda \cdot o_k + \mu \cdot \tilde{o}_k) \cdot (z_k - z_{k-1})}^{\mathcal{S}_Z(\lambda \cdot \phi + \mu \cdot \psi)} \\ &= \lambda \cdot \underbrace{\textstyle\sum_{k=1}^n o_k \cdot (z_k - z_{k-1})}_{\mathcal{S}_Z(\phi)} + \mu \cdot \underbrace{\textstyle\sum_{k=1}^n \tilde{o}_k \cdot (z_k - z_{k-1})}_{\mathcal{S}_Z(\psi)} = \lambda \cdot \textstyle\int_a^b \phi + \mu \cdot \int_a^b \psi.\end{aligned}$$

Gilt nun $\phi \leq \psi$, so auch $o_k \leq \tilde{o}_k$ für $1 \leq k \leq n$; und somit

$$\textstyle\int_a^b \phi \;=\; \underbrace{\textstyle\sum_{k=1}^n o_k \cdot (z_k - z_{k-1})}_{\mathcal{S}_Z(\phi)} \;\leq\; \underbrace{\textstyle\sum_{k=1}^n \tilde{o}_k \cdot (z_k - z_{k-1})}_{\mathcal{S}_Z(\psi)} \;=\; \textstyle\int_a^b \psi.$$

(I4) Dies folgt sofort aus den Definitionen. □

9.2 Das Riemann-Integral

Das Riemann-Integral ist wie folgt definiert:

Definition 64 (Riemann-Integral)**.** *Seien $a, b \in \mathbb{R}$ mit $a < b$ vorgegeben.*

1) *Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Abbildung, d.h., $f \in \mathrm{B}([a,b], \mathbb{R})$. Wir definieren das* Oberintegral *bzw. das* Unterintegral *von f durch*

$$\overline{\int_a^b} f := \inf_{\mathbb{R}} \Big(\underbrace{\Big\{ \int_a^b \psi \;\Big|\; f \leq \psi \in \mathcal{T}_a^b \Big\}}_{=: O_f} \Big) \qquad \text{bzw.} \qquad \underline{\int_a^b} f := \sup_{\mathbb{R}} \Big(\underbrace{\Big\{ \int_a^b \phi \;\Big|\; \mathcal{T}_a^b \ni \phi \leq f \Big\}}_{=: U_f} \Big). \quad (267)$$

Die Beschränktheit von f zusammen mit Proposition 15 impliziert:

- *Es gilt $O_f \neq \emptyset \neq U_f$; und O_f/U_f ist nach unten/oben beschränkt, d.h. sowohl Ober- als auch Unterintegral existieren in $\mathbb{R}$.* $(\inf_{\mathbb{R}}(O_f) \in \mathbb{R} \ni \sup_{\mathbb{R}}(U_f))$
- *Es gilt*

$$\overline{\int_a^b} f \geq \underline{\int_a^b} f. \qquad (268)$$

Beweis der Behauptungen:

- Wegen $f \in \mathrm{B}([a,b],\mathbb{R})$, gilt $|f| \leq c$ für ein $0 \leq c \in \mathbb{R}$. Sei $\mathcal{T}_a^b \ni \chi_{\pm c}\colon [a,b] \ni x \mapsto \pm c \in \mathbb{R}$ die konstante Treppenfunktion $\pm c$. Dann gilt $\chi_{-c} \leq f \leq \chi_c$, also $O_f \neq \emptyset \neq U_f$.
  Weiterhin folgt mit Teil (I2) und Teil (I4) von Proposition 15 (jeweils letzte Implikation):
  $$\begin{array}{ccccc} f \leq \psi \in \mathcal{T}_a^b & \Longrightarrow & \chi_{-c} \leq f \leq \psi & \Longrightarrow & -c\cdot(b-a) = \int_a^b \chi_{-c} \leq \int_a^b \psi \\ \mathcal{T}_a^b \ni \phi \leq f & \Longrightarrow & \phi \leq f \leq \chi_c & \Longrightarrow & \int_a^b \phi \leq \int_a^b \chi_c = c\cdot(b-a) \end{array}$$
  Somit ist O_f ist nach unten beschränkt, sowie U_f nach oben beschränkt.
- Proposition 15.(I2) (zweiter Schritt) zeigt
  $$\begin{array}{rcl} \alpha \in U_f \quad\wedge\quad \beta \in O_f & \Longleftrightarrow & \exists\, \mathcal{T}_a^b \ni \phi \leq f \leq \psi \in \mathcal{T}_a^b\colon \quad \alpha = \int_a^b \phi \quad\wedge\quad \beta = \int_a^b \psi \\ & \Longrightarrow & \alpha = \int_a^b \phi \leq \int_a^b \psi = \beta. \end{array}$$
  Wir erhalten
  $$\begin{array}{rcl} U_f \leq \beta \quad \forall\, \beta \in O_f & \Longrightarrow & \sup_{\mathbb{R}}(U_f) \leq \beta \qquad \forall\, \beta \in O_f \\ & \Longrightarrow & \underline{\int_a^b} f = \sup_{\mathbb{R}}(U_f) \leq \inf_{\mathbb{R}}(O_f) = \overline{\int_a^b} f. \end{array}$$
  □

2) *Eine Abbildung $f\colon [a,b] \to \mathbb{R}$ heißt Riemann-integrierbar (Riemann-integrabel), wenn sie beschränkt ist und $\underline{\int_a^b} f = \overline{\int_a^b} f$ gilt.*

Bemerkung: *Wegen Lemma 20 und (268) ist f Riemann-integrierbar genau dann, wenn gilt:*[57]
$$\forall\, \varepsilon > 0\colon\ \exists\, \phi_\varepsilon, \psi_\varepsilon \in \mathcal{T}_a^b\colon \qquad \phi_\varepsilon \leq f \leq \psi_\varepsilon \quad\wedge\quad \int_a^b \psi_\varepsilon - \int_a^b \phi_\varepsilon < \varepsilon. \tag{269}$$

Beweis der Äquivalenz:

- Gilt (269), so ist f beschränkt (beachte $\phi_\varepsilon, \psi_\varepsilon \in \mathcal{T}_a^b \subseteq \mathrm{B}([a,b],\mathbb{R})$); und wir erhalten
  $$0 \overset{(268)}{\leq} \overline{\int_a^b} f - \underline{\int_a^b} f \leq \int_a^b \psi_\varepsilon - \int_a^b \phi_\varepsilon < \varepsilon \quad \forall\, \varepsilon > 0 \qquad \overset{\text{Lemma 19}}{\Longrightarrow} \qquad \overline{\int_a^b} f - \underline{\int_a^b} f = 0,$$
  wegen $U_f \ni \int_a^b \phi_\varepsilon \leq \sup_{\mathbb{R}}(U_f) = \underline{\int_a^b} f$ und $\overline{\int_a^b} f \leq \inf_{\mathbb{R}}(O_f) = \int_a^b \psi_\varepsilon \in O_f$.
- Sei $\varepsilon > 0$ vorgegeben. Wegen Lemma 20, existieren $\mathcal{T}_a^b \ni \phi_\varepsilon \leq f \leq \psi_\varepsilon \in \mathcal{T}_a^b$ mit
  $$U_f \ni \int_a^b \phi_\varepsilon > \underline{\int_a^b} f - \tfrac{\varepsilon}{2} \qquad \text{und} \qquad O_f \ni \int_a^b \psi_\varepsilon < \overline{\int_a^b} f + \tfrac{\varepsilon}{2}.$$
  Gilt nun $\underline{\int_a^b} f = \overline{\int_a^b} f$, so erhalten wir
  $$\int_a^b \psi_\varepsilon - \int_a^b \phi_\varepsilon = \Big(\int_a^b \psi_\varepsilon - \overline{\int_a^b} f\Big) + \Big(\underline{\int_a^b} f - \int_a^b \phi_\varepsilon\Big) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$
  □

Ist f Riemann-integrierbar, so definieren wir das Riemann-Integral von f durch
$$\int_a^b f := \overline{\int_a^b} f = \underline{\int_a^b} f. \tag{270}$$

Alternativ schreiben wir auch $\int_a^b f(x)\,\mathrm{d}x$ anstelle (270) (wobei die explizite Bezeichnung der Integrationsvariable x wieder keine Rolle spielt). Schließlich setzen wir
$$\int_c^c f := 0 \qquad \textit{sowie} \qquad \int_b^a f := -\int_a^b f \qquad \textit{für alle} \qquad a \leq c \leq b. \tag{271}$$

[57] Beachte $0 < \int_a^b \psi_\varepsilon - \int_a^b \phi_\varepsilon = \int_a^b (\psi_\varepsilon - \phi_\varepsilon) < \varepsilon$ wegen Proposition 15.(I3).

3) *Wir notieren die Menge aller Riemann-integrierbaren Funktionen* $[a,b] \to \mathbb{R}$ *mit* $\mathcal{R}_a^b \subseteq \mathrm{B}([a,b],\mathbb{R})$.

Es folgt zwanglos aus den Definitionen, dass $\mathcal{T}_a^b \subseteq \mathcal{R}_a^b$ *gilt; und dass das Riemann-integral* (270) *auf* $\mathcal{T}_a^b$ *mit dem Integral für Treppenfunktionen* (264) *übereinstimmt. Insbesondere sind somit die Konventionen* (265) *und* (271) *konsistent miteinander.*

Beweis der Behauptung: Sei $f \in \mathcal{T}_a^b$ vorgegeben; und es bezeichne α_f das Integral (264). Wegen $f \leq f$, gilt dann offensichtlich $O_f \ni \alpha_f \in U_f$; und Proposition 15.(I2) zeigt weiterhin

$$\big(\, f \leq \psi \in \mathcal{T}_a^b \quad \Rightarrow \quad \alpha_f \leq \textstyle\int_a^b \psi \,\big) \qquad \text{sowie} \qquad \big(\, \mathcal{T}_a^b \ni \phi \leq f \quad \Rightarrow \quad \textstyle\int_a^b \phi \leq \alpha_f \,\big).$$

Somit gilt $\alpha_f \leq O_f \ni \alpha_f \in U_f \leq \alpha_f$, also

$$\overline{\int_a^b} f = \inf_{\mathbb{R}}(O_f) = \alpha_f = \sup_{\mathbb{R}}(U_f) = \underline{\int_a^b} f \qquad \Longrightarrow \qquad \alpha_f = \int_a^b f,$$

wobei auf der rechten Seite, „ $\int_a^b f$ “ das Riemann-Integral (270) von f bezeichnet. □

Übung 107. *Sei* $\emptyset \neq I \subseteq \mathbb{R}$ *ein offenes Intervall in* $\mathbb{R}$. *Zeigen Sie* $I \cap \mathbb{Q} \neq \emptyset$ *sowie* $I \setminus \mathbb{Q} \neq \emptyset$.

Übung 108. *Sei* $\lambda \geq 0$ *und* $\emptyset \neq M \subseteq \mathbb{R}$. *zeigen Sie:*

- *Ist* M *nach oben beschränkt, so gilt* $\sup_{\mathbb{R}}(\lambda \cdot M) = \lambda \cdot \sup_{\mathbb{R}}(M)$.
- *Ist* M *nach unten beschränkt, so gilt* $\inf_{\mathbb{R}}(\lambda \cdot M) = \lambda \cdot \inf_{\mathbb{R}}(M)$.

Beispiel 75. *Ein wichtiges Beispiel, das die Subtilitäten des Integrierbarkeitsbegriff verdeutlicht, ist die* Dirichletfunktion

$$f : [0,1] \to \mathbb{R}, \qquad x \mapsto \begin{cases} 1 & \textit{falls} \quad x \in \mathbb{Q} \cap [0,1] \\ 0 & \textit{falls} \quad x \in [0,1] \setminus \mathbb{Q}. \end{cases}$$

Diese Funktion ist zwar beschränkt, aber in keinem Punkt stetig; und weiterhin nicht Riemann-integrierbar wegen

$$\underline{\int_0^1} f = 0 < 1 = \overline{\int_0^1} f.$$

Beweis der Behauptung: Sei $\mathcal{T}_a^b \ni \chi_{0/1} \colon [a,b] \ni x \mapsto 0/1 \in \mathbb{R}$ die konstante Treppenfunktion $0/1$.

- Es gilt $\chi_0 \leq f \leq \chi_1$, mit $U_f \ni \int_0^1 \chi_0 = 0$ und $O_f \ni \int_0^1 \chi_1 = 1$; also

$$\underline{\int_0^1} f = \sup_{\mathbb{R}}(U_f) \geq \int_0^1 \chi_0 = 0 \qquad \wedge \qquad 1 = \int_0^1 \chi_1 \geq \inf_{\mathbb{R}}(O_f) = \overline{\int_0^1} f.$$

- Sei $\mathcal{T}_0^1 \ni \phi \leq f$ eine Treppenfunktion zu der Zerlegung $Z = (z_0, \dots, z_n)$ von $[0,1]$, mit Konstante $o_1, \dots, o_n$. Übung 107 zeigt $o_1, \dots, o_n \leq 0$, also $U_f \ni \int_0^1 \phi \leq 0$. Es folgt $\underline{\int_0^1} f = \sup_{\mathbb{R}}(U_f) \leq 0$.
- Sei $f \leq \psi \in \mathcal{T}_0^1$ eine Treppenfunktion zu der Zerlegung $Z = (z_0, \dots, z_n)$ von $[0,1]$, mit Konstante $o_1, \dots, o_n$. Übung 107 zeigt $o_1, \dots, o_n \geq 1$, also $O_f \ni \int_0^1 \psi \geq 1$. Es folgt $\overline{\int_0^1} f = \inf_{\mathbb{R}}(O_f) \geq 1$. □

Satz 53. *Sei* $\mathbb{R} \ni a < b \in \mathbb{R}$. *Das Riemann-Integral hat die folgende Eigenschaften:*

(I1) *Intervalladditivität: Für* $a < c < b$ *gilt* $f \in \mathcal{R}_a^b$ *genau dann, wenn* $f|_{[a,c]} \in \mathcal{R}_a^c$ *und* $f|_{[c,b]} \in \mathcal{R}_c^b$ *gilt. In diesem Fall ist*

$$\int_a^b f = \int_a^c f + \int_c^b f. \tag{272}$$

(Wegen (271), *gilt* (272) *natürlich auch im Falle* $c \in \{a,b\}$.*)*

(I2) *Monotonie: Es gilt* $\int_a^b f \leq \int_a^b g$, *für* $\mathcal{R}_a^b \ni f \leq g \in \mathcal{R}_a^b$.

(I3) *Linearität:* *Für $\lambda, \mu \in \mathbb{R}$ und $f, g \in \mathcal{R}_a^b$, gilt $(\lambda \cdot f + \mu \cdot g) \in \mathcal{R}_a^b$ mit*

$$\int_a^b (\lambda \cdot f + \mu \cdot g) = \lambda \cdot \int_a^b f + \mu \cdot \int_a^b g.$$

(I4) *Normierung:* $\int_a^b \chi_c = c \cdot (b-a)$ *für alle* $c \in \mathbb{R}$.

Beweis. Für $h \in \mathcal{R}_a^b$, setzen wir (vgl. (267))

$$O_h := \{\int_a^b \psi \mid h \leq \psi \in \mathcal{T}_a^b\} \qquad \text{sowie} \qquad U_h := \{\int_a^b \phi \mid \mathcal{T}_a^b \ni \phi \leq h\}.$$

I1) Sei $f_- := f|_{[a,c]}$ und $f_+ := f|_{[c,b]}$. Wir zeigen zunächst $O_f = O_{f_-} + O_{f_+}$ sowie $U_f = U_{f_-} + U_{f_+}$:

- Es gilt $O_f \subseteq O_{f_-} + O_{f_+}$ sowie $U_f \subseteq U_{f_-} + U_{f_+}$:
  Sei $\alpha \in O_f$, also $\alpha = \int_a^b \psi$ mit $f \leq \psi \in \mathcal{T}_a^b$. Dann gilt $f_- \leq \psi|_{[a,c]} \in \mathcal{T}_a^c$ und $f_+ \leq \psi|_{[c,b]} \in \mathcal{T}_c^b$ (Bemerkung 64.b)). Wegen Proposition 15 gilt

  $$\alpha = \int_a^b \psi \overset{(I1)}{=} \underbrace{\int_a^c \psi|_{[a,c]}}_{\in O_{f_-}} + \underbrace{\int_c^b \psi|_{[c,b]}}_{\in O_{f_+}} \in O_{f_-} + O_{f_+}.$$

  In der gleichen Weise folgt $U_f \subseteq U_{f_-} + U_{f_+}$.
- Es gilt $O_{f_-} + O_{f_+} \subseteq O_f$ sowie $U_{f_-} + U_{f_+} \subseteq U_f$:
  Sei $\alpha = \alpha_- + \alpha_+$ mit $\alpha_\pm \in O_{f_\pm}$ vorgegeben. Dann existieren $f_- \leq \chi[\alpha]_- \in \mathcal{T}_a^c$ und $f_+ \leq \chi[\alpha]_+ \in \mathcal{T}_c^b$ mit $\alpha_- = \int_a^c \chi[\alpha]_-$ und $\alpha_+ = \int_c^b \chi[\alpha]_+$.
  - Wir können ohne Einschränkung $\chi[\alpha]_\pm(c) = f(c)$ annehmen, einfach indem wir notfalls den Funktionswert von $\chi[\alpha]_\pm$ bei c in $f(c)$ abändern.
    (Offensichtlich gilt dann weiterhin $f_\pm \leq \chi[\alpha]_\pm$ sowie $\alpha_- = \int_a^c \chi[\alpha]_-$ und $\alpha_+ = \int_c^b \chi[\alpha]_+$.)
  - Wir definieren $\chi[\alpha] \in \mathcal{T}_a^b$ wie in Lemma 55.c), d.h.

  $$\chi[\alpha](x) = \begin{cases} \chi[\alpha]_-(x) & \text{für} \quad x \in [a,c) \\ f(c) & \text{für} \quad x = c \\ \chi[\alpha]_+(x) & \text{für} \quad x \in (c,b]. \end{cases}$$

  Offensichtlich gilt dann $f \leq \chi[\alpha]$, und Proposition 15 liefert

  $$\int_a^b \chi[\alpha] \overset{(I1)}{=} \int_a^c \chi[\alpha]|_{[a,c]} + \int_c^b \chi[\alpha]|_{[c,b]} = \int_a^c \chi[\alpha]_- + \int_c^b \chi[\alpha]_+ = \overbrace{\alpha_- + \alpha_+}^{\alpha} \in O_f.$$

  Dies zeigt $O_{f_-} + O_{f_+} \subseteq O_f$; und ganz analog folgt auch $U_{f_-} + U_{f_+} \subseteq U_f$.

Die ersten beiden Teile von Proposition 1 (zweiter Schritt) liefern:

$$\begin{aligned} \underline{\int_a^b} f &= \sup\nolimits_{\mathbb{R}}(U_f) = \sup\nolimits_{\mathbb{R}}(U_{f_-} + U_{f_+}) = \sup\nolimits_{\mathbb{R}}(U_{f_-}) + \sup\nolimits_{\mathbb{R}}(U_{f_+}) = \underline{\int_a^c} f_- + \underline{\int_c^b} f_+ \\ \overline{\int_a^b} f &= \inf\nolimits_{\mathbb{R}}(O_f) = \inf\nolimits_{\mathbb{R}}(O_{f_-} + O_{f_+}) = \inf\nolimits_{\mathbb{R}}(O_{f_-}) + \inf\nolimits_{\mathbb{R}}(O_{f_+}) = \overline{\int_a^c} f_- + \overline{\int_c^b} f_+ \end{aligned} \tag{273}$$

- Es gelte $f_- \in \mathcal{R}_a^c \ \wedge \ f_+ \in \mathcal{R}_c^b$; also $\underline{\int_a^c} f_- = \int_a^b f_- = \overline{\int_a^c} f_-$ sowie $\underline{\int_c^b} f_+ = \int_a^b f_+ = \overline{\int_c^b} f_+$. Wir erhalten

  $$\underline{\int_a^b} f \overset{(273)}{=} \underline{\int_a^c} f_- + \underline{\int_c^b} f_+ = \int_a^c f_- + \int_c^b f_+ = \overline{\int_a^c} f_- + \overline{\int_c^b} f_+ \overset{(273)}{=} \overline{\int_a^b} f,$$

  also $f \in \mathcal{R}_a^b$ mit $\int_a^b f = \int_a^c f_- + \int_c^b f_+$.

- Es gelte $f \in \mathcal{R}_a^b$ also $\underline{\int}_a^b f = \int_a^b f = \overline{\int}_a^b f$. Es folgt

$$\underline{\int}_a^c f_- + \underline{\int}_c^b f_+ \overset{(273)}{=} \underline{\int}_a^b f = \int_a^b f = \overline{\int}_a^b f \overset{(273)}{=} \overline{\int}_a^c f_- + \overline{\int}_c^b f_+, \tag{274}$$

  und wir erhalten

$$0 \overset{(268)}{\leq} \overline{\int}_a^c f_- - \underline{\int}_a^c f_- \overset{(274)}{=} \underline{\int}_c^b f_+ - \overline{\int}_c^b f_+ \overset{(268)}{\leq} 0 \quad \text{ also } \quad \overline{\int}_a^c f_- = \underline{\int}_a^c f_- \ \wedge \ \overline{\int}_c^b f_+ = \underline{\int}_c^b f_+.$$

  Dies zeigt $f_- \in \mathcal{R}_a^c$ und $f_+ \in \mathcal{R}_c^b$ mit $\int_a^b f = \int_a^c f_- + \int_c^b f_+$.

I2) Sei $\mathcal{R}_a^b \ni f \leq g \in \mathcal{R}_a^b$. Dann gilt $O_g \subseteq O_f$ wegen

$$\alpha \in O_g \qquad \Longrightarrow \qquad \exists\, g \leq \psi \in \mathcal{T}_a^b\colon\ \alpha = \int_a^b \psi \qquad \overset{f \leq g \leq \psi}{\Longrightarrow} \qquad \alpha \in O_f.$$

Proposition 1.10) zeigt $\int_a^b f = \overline{\int}_a^b f = \inf_{\mathbb{R}}(O_f) \leq \inf_{\mathbb{R}}(O_g) = \overline{\int}_a^b g = \int_a^b g$.

I3)
- Wir zeigen $\int_a^b (f+g) = \int_a^b f + \int_a^b g$:
  Es gilt $O_f + O_g \subseteq O_{f+g}$ und $U_f + U_g \subseteq U_{f+g}$; denn Proposition 15.(I3) zeigt (jeweils zweite Implikation)

$$\begin{aligned}
\alpha \in O_f \quad \wedge \quad \beta \in O_g \qquad &\Longrightarrow \qquad \exists\, f \leq \phi \in \mathcal{T}_a^b\colon \alpha = \int_a^b \phi \quad \wedge \quad \exists\, g \leq \psi \in \mathcal{T}_a^b\colon \beta = \int_a^b \psi \\
&\Longrightarrow \qquad (f+g) \leq (\phi + \psi) \in \mathcal{T}_a^b \quad \wedge \quad \int_a^b (\phi+\psi) = \alpha + \beta \\
&\Longrightarrow \qquad \alpha + \beta \in O_{f+g}, \\
\alpha \in U_f \quad \wedge \quad \beta \in U_g \qquad &\Longrightarrow \qquad \exists\, \mathcal{T}_a^b \ni \phi \leq f\colon \alpha = \int_a^b \phi \quad \wedge \quad \exists\, \mathcal{T}_a^b \ni \psi \leq g\colon \beta = \int_a^b \psi \\
&\Longrightarrow \qquad (\phi+\psi) \leq (f+g) \in \mathcal{T}_a^b \quad \wedge \quad \int_a^b (\phi+\psi) = \alpha + \beta \\
&\Longrightarrow \qquad \alpha + \beta \in U_{f+g}.
\end{aligned}$$

  Sei $\gamma := \int_a^b f + \int_a^b g$ Wir erhalten mit den ersten- sowie den letzten beiden Teilen von Proposition 1, dass

$$\begin{aligned}
\overline{\int}_a^b (f+g) &= \inf_{\mathbb{R}}(O_{f+g}) \leq \inf_{\mathbb{R}}(O_f + O_g) = \inf_{\mathbb{R}}(O_f) + \inf_{\mathbb{R}}(O_g) = \overline{\int}_a^b f + \overline{\int}_a^b g = \gamma \\
\underline{\int}_a^b (f+g) &= \sup_{\mathbb{R}}(U_{f+g}) \geq \sup_{\mathbb{R}}(U_f + U_g) = \sup_{\mathbb{R}}(U_f) + \sup_{\mathbb{R}}(U_g) = \underline{\int}_a^b f + \underline{\int}_a^b g = \gamma
\end{aligned}$$

  gilt; also $\underline{\int}_a^b (f+g) \geq \gamma \geq \overline{\int}_a^b (f+g) \overset{(268)}{\geq} \underline{\int}_a^b (f+g)$, d.h. $\underline{\int}_a^b (f+g) = \gamma = \overline{\int}_a^b (f+g)$.
- Wir zeigen $\int_a^b (\lambda \cdot f) = \lambda \cdot \int_a^b f$:
  - Für $\lambda = 0$, gilt $\int_a^b (\lambda \cdot f) = \int_a^b \chi_0 = 0 = 0 \cdot \int_a^b f = \lambda \cdot \int_a^b f$ wegen Proposition 15.(I4).
  - Für $\lambda > 0$ und $\chi \in \mathcal{T}_a^b$ gilt

$$\chi \leq f \quad \Leftrightarrow \quad \lambda \cdot \chi \leq \lambda \cdot f \qquad \text{ sowie } \qquad f \leq \chi \quad \Leftrightarrow \quad \lambda \cdot f \leq \lambda \cdot \chi,$$

    also $U_{\lambda \cdot f} = \lambda \cdot U_f$ sowie $O_{\lambda \cdot f} = \lambda \cdot O_f$ wegen Proposition 15.(I3). Übung 108 zeigt nun

$$\begin{aligned}
\overline{\int}_a^b \lambda \cdot f &= \inf_{\mathbb{R}}(O_{\lambda \cdot f}) = \inf_{\mathbb{R}}(\lambda \cdot O_f) = \lambda \cdot \inf_{\mathbb{R}}(O_f) = \lambda \cdot \overline{\int}_a^b f = \lambda \cdot \int_a^b f \\
\underline{\int}_a^b \lambda \cdot f &= \sup_{\mathbb{R}}(U_{\lambda \cdot f}) = \sup_{\mathbb{R}}(\lambda \cdot U_f) = \lambda \cdot \sup_{\mathbb{R}}(U_f) = \lambda \cdot \underline{\int}_a^b f = \lambda \cdot \int_a^b f,
\end{aligned}$$

    also $\int_a^b (\lambda \cdot f) = \lambda \cdot \int_a^b f$.

– Für $\lambda < 0$, folgt mit dem bereits Gezeigten:

$$0 = \int_a^b \chi_0 = \int_a^b (\lambda \cdot f + (-\lambda) \cdot f) = \int_a^b (\lambda \cdot f) + \int_a^b (-\lambda \cdot f)$$
$$\Longrightarrow \qquad \int_a^b (\lambda \cdot f) = -\int_a^b (-\lambda) \cdot f = -(-\lambda) \cdot \int_a^b f = \lambda \cdot \int_a^b f.$$

I4) Wegen $f \in \mathcal{T}_a^b$, folgt die Behauptung sofort aus Proposition 15.(I4), sowie der Gleichheit der Integrale (264) und (270) auf Treppenfunktionen vgl. (Definition 64.3)). $\square$

Beispiel 76. *Für $\mathbb{R} \ni a < b \in \mathbb{R}$ und $\mathcal{R}_a^b \ni f \geq 0$, gilt*

$$\int_a^b f \overset{(\mathrm{I2})}{\geq} \int_a^b \chi_0 \overset{(\mathrm{I4})}{=} 0 \cdot (b-a) = 0.$$

9.3 Integrierbarkeit Stetiger und Monotoner Funktionen

Wir zeigen nun, dass stetige-, sowie auch monotone Funktionen $\mathbb{R} \supseteq [a,b] \to \mathbb{R}$ $(a < b)$ Riemann-integrierbar sind. Hierfür benutzen wir Satz 34, welcher uns die gleichmäßige Stetigkeit stetiger Funktionen $[a,b] \to \mathbb{R}$ liefert.

Lemma 58. *Sei $f\colon \mathbb{R} \supseteq [a,b] \to \mathbb{R}$ stetig, mit $a < b$. Zu jedem $\varepsilon > 0$, existiert ein $\chi \in \mathcal{T}_a^b$ mit*

$$|f - \chi|_\infty = \sup\{|f(x) - \chi(x)| \mid x \in [a,b]\} \leq \varepsilon.$$

Beweis. Wegen Satz 34 ist f gleichmäßig stetig. Für $\varepsilon > 0$ vorgegeben, existiert also ein $\delta > 0$ mit

$$|f(x) - f(y)| < \varepsilon \quad \text{für alle} \quad x, y \in [a,b] \quad \text{mit} \quad |x - y| < \delta.$$

Wir wählen nun eine Zerlegung $Z = (z_0, \dots, z_n)$ von $[a,b]$ mit $|z_{k+1} - z_k| < \delta$ für alle $k = 0, \dots, n-1$,[58] und definieren die Treppenfunktion χ durch

$$\chi(t) := \begin{cases} f(z_k) & \text{für} \quad t \in [z_k, z_{k+1}) \;\text{ mit }\; 1 \leq k \leq n-1 \\ f(b) & \text{für} \quad t = b. \end{cases}$$

Dann gilt $|f(b) - \chi(b)| = 0$, sowie

$$|f(t) - \chi(t)| = |f(t) - f(z_k)| < \varepsilon \quad \text{für} \quad t \in [z_k, z_{k+1}) \quad \text{mit} \quad 0 \leq k \leq n-1.$$

Dies zeigt $|f - \chi|_\infty \leq \varepsilon$. $\square$

Korollar 41. *Sei $f\colon \mathbb{R} \supseteq [a,b] \to \mathbb{R}$ stetig, mit $a < b$. Zu jedem $\varepsilon > 0$ existieren $\phi_\varepsilon, \psi_\varepsilon \in \mathcal{T}_a^b$ mit $\phi_\varepsilon \leq f \leq \psi_\varepsilon$ und $\psi_\varepsilon - \phi_\varepsilon = \varepsilon$. Insbesondere gilt*

$$f - \varepsilon \leq \psi_\varepsilon - \varepsilon = \phi_\varepsilon \leq f \leq \psi_\varepsilon = \phi_\varepsilon + \varepsilon \leq f + \varepsilon.$$

Beweis. Wegen Lemma 58, existiert ein $\chi \in \mathcal{T}_a^b$ mit $|f - \chi|_\infty \leq \frac{\varepsilon}{2}$. Dann gilt (Lemma 55.a))

$$\mathcal{T}_a^b \ni \phi_\varepsilon := \chi - \frac{\varepsilon}{2} \leq f \qquad \text{und} \qquad f \leq \chi + \frac{\varepsilon}{2} =: \psi_\varepsilon \in \mathcal{T}_a^b \qquad \text{sowie} \qquad \psi_\varepsilon - \phi_\varepsilon = \varepsilon.$$ $\square$

Satz 54. *Sei $\mathbb{R} \ni a < b \in \mathbb{R}$.*

a) Jede stetige Funktion $\mathbb{R} \supseteq [a,b] \to \mathbb{R}$ ist Riemann-integrierbar. $\quad (C([a,b],\mathbb{R}) \subseteq \mathcal{R}_a^b)$

b) Jede monotone Funktion $\mathbb{R} \supseteq [a,b] \to \mathbb{R}$ ist Riemann-integrierbar.

[58] Man wähle $n \in \mathbb{N}_{>0}$ mit $\tau := (b-a)/n < \delta$, und setze $z_k := a + k \cdot \tau$ für $k = 0, \dots, n$.

Beweis. a) Ist $f : [a,b] \to \mathbb{R}$ stetig, so auch beschränkt gemäß Satz 35 (Satz vom Maximum). Weiterhin existieren gemäß Korollar 41 zu jedem $\varepsilon > 0$ Treppenfunktionen $\phi_\varepsilon, \psi_\varepsilon \in \mathcal{T}_a^b$ mit $\phi_\varepsilon \leq f \leq \psi_\varepsilon$ und $\psi_\varepsilon - \phi_\varepsilon = \frac{\varepsilon}{2(b-a)}$, also

$$\int_a^b \psi_\varepsilon - \int_a^b \phi_\varepsilon = \int_a^b (\psi_\varepsilon - \phi_\varepsilon) = \int_a^b \frac{\varepsilon}{2(b-a)} \overset{(\mathsf{I4})}{=} \frac{\varepsilon}{2} \cdot \frac{b-a}{b-a} < \varepsilon.$$

Somit gilt (269), und daher $f \in \mathcal{R}_a^b$.

b) Es genügt die Aussage für monoton wachsendes f zu zeigen: Ist nämlich f monoton fallend, so ist $-f$ monoton wachsend, und dann Riemann-integrierbar. Dann ist aber auch $f = -(-f)$ Riemann-integrierbar, wegen Satz 53.(I3).

Sei also $f\colon [a,b] \to \mathbb{R}$ monoton wachsend. Dann gilt $f(a) \leq f(t) \leq f(b)$ für alle $t \in [a,b]$; also ist f beschränkt. Ist f konstant, so ist die Behauptung klar. Wir können daher $f(a) < f(b)$ annehmen. Sei nun $\varepsilon > 0$ vorgegeben. Wir wählen eine Zerlegung $Z = (z_0, \ldots, z_n)$ von $[a,b]$ mit

$$z_{k+1} - z_k < \frac{\varepsilon}{f(b) - f(a)} \qquad \text{für} \quad 0 \leq k \leq n-1.$$

Wir definieren $\phi_\varepsilon, \psi_\varepsilon \in \mathcal{T}_a^b$, durch $\phi_\varepsilon(b) := f(b) =: \psi_\varepsilon(b)$ sowie

$$\phi_\varepsilon(t) := f(z_k) \quad \text{sowie} \quad \psi_\varepsilon(t) := f(z_{k+1}) \quad \text{für} \quad t \in [z_k, z_{k+1}) \quad \text{mit} \quad 0 \leq k \leq n-1.$$

Offensichtlich gilt $\phi_\varepsilon \leq f \leq \psi_\varepsilon$, und wir erhalten

$$\begin{aligned} \int_a^b \psi_\varepsilon - \int_a^b \phi_\varepsilon &= \sum_{k=0}^{n-1} (f(z_{k+1}) - f(z_k)) \cdot (z_{k+1} - z_k) \\ &< \frac{\varepsilon}{f(b)-f(a)} \cdot \sum_{k=0}^{n-1} (f(z_{k+1}) - f(z_k)) \\ &= \frac{\varepsilon}{f(b)-f(a)} \cdot (f(b) - f(a)) \\ &= \varepsilon. \end{aligned}$$

Wegen (269) ist f somit Riemann-integrierbar. □

Übung 109. *Für $x \in \mathbb{R}$, sei $x_+ := \max(x,0) \geq 0$ und $x_- = -\min(x,0) \geq 0$. Zeigen Sie:*

a) $x = x_+ - x_-$, $|x| = x_+ + x_-$, $x_+ \cdot x_- = 0$ *für alle* $x \in \mathbb{R}$.

b) $x_+ \leq y_+$ *sowie* $y_- \leq x_-$ *für alle* $\mathbb{R} \ni x \leq y \in \mathbb{R}$.

c) $y_+ - x_+ \leq y - x$ *für alle* $\mathbb{R} \ni x \leq y \in \mathbb{R}$.

9.4 Der Mittelwertsatz der Integralrechnung

Notation 27. *Sei Z eine Menge. Für $f, g \in \mathsf{Abb}(Z, \mathbb{R})$, setzen wir*

$$\begin{aligned} \max(f,g)\colon Z \to \mathbb{R}, &\qquad z \mapsto \max(f(z), g(z)) \\ \min(f,g)\colon Z \to \mathbb{R}, &\qquad z \mapsto \min(f(z), g(z)). \end{aligned}$$

Für $M \in \mathbb{R}$, sind dann mit $\max(f,M)$ bzw. $\min(f,M)$ obige Ausdrücke für die konstante Abbildung $g = M$ gemeint. Eine einfache Fallunterscheidung zeigt

$$\max(f,g) = \tfrac{1}{2} \cdot (f+g) + \tfrac{1}{2} \cdot |f-g| \qquad \textit{sowie} \qquad \min(f,g) = \tfrac{1}{2} \cdot (f+g) - \tfrac{1}{2} \cdot |f-g|.$$

Übung 110. *Seien $\chi, \chi' \in \mathcal{T}_a^b$ mit $a < b$. Zeigen sie, dass gilt:*

$$|\chi| \in \mathcal{T}_a^b, \qquad \chi^2 \in \mathcal{T}_a^b, \qquad \max(\chi, \chi') \in \mathcal{T}_a^b, \qquad \min(\chi, \chi') \in \mathcal{T}_a^b.$$

Lemma 59. *Gegeben $f, g \in \mathcal{R}_a^b$ mit $a < b$, so sind auch die Abbildungen*

$$f_+ := \max(f, 0), \qquad f_- := -\min(f, 0), \qquad \max(f, g), \qquad \min(f, g), \qquad |f|$$

Riemann-integrierbar, also Elemente von $\mathcal{R}_a^b$.

Beweis. Für $\alpha, \beta \in \mathsf{Abb}([a, b], \mathbb{R})$, sei $\alpha_+ := \max(\alpha, 0)$ und $\alpha_- := -\min(\alpha, 0)$. Dann gilt

$$\alpha_- = \alpha_+ - \alpha \qquad \text{sowie} \qquad |\alpha| = \alpha_+ + \alpha_-, \tag{275}$$

wegen Übung 109.a); und wegen Notation 27 gilt

$$\max(\alpha, \beta) = \tfrac{1}{2} \cdot (\alpha + \beta) + \tfrac{1}{2} \cdot |\alpha - \beta| \qquad \text{sowie} \qquad \min(\alpha, \beta) = \tfrac{1}{2} \cdot (\alpha + \beta) - \tfrac{1}{2} \cdot |\alpha - \beta|. \tag{276}$$

Die Behauptung folgt somit aus Satz 53, sobald wir nachgewiesen haben, dass für ein gegebenes $f \in \mathcal{R}_a^b$, auch $f_+ \in \mathcal{R}_a^b$ gilt:

- Für $f \in \mathcal{R}_a^b$, gilt dann $f_\pm \in \mathcal{R}_a^b$ wegen der linken Seite von (275) und Satz 53; also auch $|f| \in \mathcal{R}_a^b$ wegen der rechten Seite von (275) und Satz 53.
- Für $f, g \in \mathcal{R}_a^b$, gilt nun $\gamma^\pm := f \pm g \in \mathcal{R}_a^b$ wegen Satz 53; und dann auch $|\gamma^-| \in \mathcal{R}_a^b$ wegen dem vorhergehenden Punkt. Daher gilt $\max(f, g) \in \mathcal{R}_a^b$ sowie $\min(f, g) \in \mathcal{R}_a^b$ wegen (276) und Satz 53.

Sei also $f \in \mathcal{R}_a^b$ vorgegeben. Für $\varepsilon > 0$ vorgegeben, existieren gemäß (269) $\phi, \psi \in \mathcal{T}_a^b$ mit

$$\phi \leq f \leq \psi \qquad \text{sowie} \qquad \textstyle\int_a^b (\psi - \phi) = \int_a^b \psi - \int_a^b \phi < \varepsilon.$$

Wegen Übung 109 (Teil b) und c)), gilt dann

$$\phi_+ \leq f_+ \leq \psi_+ \qquad \text{sowie} \qquad \psi_+ - \phi_+ \leq \psi - \phi,$$

mit $\phi_+, \psi_+ \in \mathcal{T}_a^b$ wegen Übung 110. Wir erhalten mit der Monotonie des Integrals

$$\textstyle\int_a^b \psi_+ - \int_a^b \phi_+ = \int_a^b (\psi_+ - \phi_+) \leq \int_a^b (\psi - \phi) < \varepsilon.$$

Also ist f_+ Riemann-integrierbar wegen (269). □

Satz 55 (Integralabschätzung)**.** *Sei $\mathbb{R} \ni a < b \in \mathbb{R}$. Dann gilt*

$$\left| \int_a^b f \right| \leq \int_a^b |f| \qquad \forall\, f \in \mathcal{R}_a^b.$$

Beweis. Lemma 59 zeigt $|f| \in \mathcal{R}_a^b$. Wir erhalten mit Satz 53

$$\pm f \leq |f| \qquad \Longrightarrow \qquad \pm \textstyle\int_a^b f \overset{(13)}{=} \int_a^b \pm f \overset{(12)}{\leq} \int_a^b |f| \qquad \Longrightarrow \qquad \left| \int_a^b f \right| \leq \int_a^b |f|.$$ □

Korollar 42. *Sei $\mathbb{R} \ni a \neq b \in \mathbb{R}$. Dann gilt*

$$\left| \int_a^b f \right| \leq \left| \int_a^b |f| \right| \qquad \forall\, f \in \mathcal{R}_{\min(a,b)}^{\max(a,b)}. \tag{277}$$

Beweis. Für $a < b$ ist die Behauptung klar wegen Satz 55 und Beispiel 76. Für $\mathbb{R} \ni b < a \in \mathbb{R}$, erhalten wir mit Satz 55 (dritter Schritt), sowie Beispiel 76 (vierter Schritt):

$$\left| \textstyle\int_a^b f \right| \overset{(271)}{=} \left| -\int_b^a f \right| = \left| \int_b^a f \right| \leq \int_b^a |f| = \left| \int_b^a |f| \right| \overset{(271)}{=} \left| -\int_a^b |f| \right| = \left| \int_a^b |f| \right|.$$ □

Lemma 60. *Für $f, g \in \mathcal{R}_a^b$, gilt auch auch $f \cdot g \in \mathcal{R}_a^b$.*

Beweis. Es genügt die Behauptung für den Fall $f = g$ zu zeigen, dass also mit $f \in \mathcal{R}_a^b$ auch $f^2 \in \mathcal{R}_a^b$ gilt. In der Tat folgt dann die allgemeine Aussage sofort mit Satz 53.(I3) aus

$$f \cdot g = \tfrac{1}{2} \cdot \big((f+g)^2 - f^2 - g^2\big).$$

Wegen $|f| \in \mathcal{R}_a^b$ (gemäß Lemma 59) und $f^2 = |f|^2$, dürfen wir dann weiterhin $f \geq 0$ annehmen. Sei nun $\varepsilon > 0$ vorgegeben; sowie $C > 0$ mit $f \leq C$, d.h., $0 \leq f \leq C$ (beachte $f \in \mathrm{B}([a,b],\mathbb{R})$):

- Wegen (269), existieren $\phi, \psi \in \mathcal{T}_a^b$ mit $\phi \leq f \leq \psi$ sowie $\int_a^b (\psi - \phi) = \int_a^b \psi - \int_a^b \phi < \frac{\varepsilon}{2C}$.
- Wegen Übung 110, gilt $\tilde{\phi} := \max(\phi, 0) \in \mathcal{T}_a^b$ sowie $\tilde{\psi} := \min(\psi, C) \in \mathcal{T}_a^b$ mit

$$0 \leq \tilde{\phi} \leq f \leq \tilde{\psi} \leq C \qquad \text{und} \qquad \int_a^b \big(\tilde{\psi} - \tilde{\phi}\big) \overset{\text{(I2)}}{\leq} \int_a^b (\psi - \phi) < \frac{\varepsilon}{2C}. \tag{278}$$

(Für die linke Seite beachte $0 \leq f \leq C$, und für die rechte Seite beachte $\tilde{\psi} \leq \psi$ sowie $\tilde{\phi} \geq \phi$.)

Somit gilt $\tilde{\phi}^2 \leq f^2 \leq \tilde{\psi}^2$ (linke Seite von (278)), mit $\tilde{\phi}^2, \tilde{\psi}^2 \in \mathcal{T}_a^b$ wegen Übung 110. Wir erhalten

$$\int_a^b \tilde{\psi}^2 - \int_a^b \tilde{\phi}^2 = \int_a^b \big(\tilde{\psi}^2 - \tilde{\phi}^2\big) = \int_a^b \underbrace{(\tilde{\psi} + \tilde{\phi})}_{\leq 2C} \cdot \underbrace{(\tilde{\psi} - \tilde{\phi})}_{\geq 0} \overset{\text{(I2)}}{\leq} \int_a^b 2C \cdot \big(\tilde{\psi} - \tilde{\phi}\big) = 2C \cdot \int_a^b \underbrace{(\tilde{\psi} - \tilde{\phi})}_{< \frac{\varepsilon}{2C}} < \varepsilon.$$

Somit folgt $f^2 \in \mathcal{R}_a^b$ aus (269). □

Satz 56 (Mittelwertsatz der Integralrechnung)**.** *Sei $f\colon \mathbb{R} \supseteq [a,b] \to \mathbb{R}$ mit $a < b$ stetig, und sei $0 \leq g \in \mathcal{R}_a^b$. Dann existiert ein $\xi \in [a,b]$ mit*

$$\int_a^b f \cdot g = f(\xi) \cdot \int_a^b g.$$

Insbesondere existiert ein $\xi \in [a,b]$ mit $\int_a^b f = f(\xi) \cdot (b-a)$. *(setze $g = 1$)*

Beweis. Wegen Satz 54.a) und Lemma 60, ist $f \cdot g \in \mathcal{R}_a^b$. Wegen Satz 35 (Satz vom Maximum), existieren $u, w \in [a,b]$ mit $m := \min(\mathrm{im}(f)) = f(u)$ und $M := \max(\mathrm{im}(f)) = f(w)$. Wir erhalten wegen $g \geq 0$

$$m \cdot g \leq f \cdot g \leq M \cdot g \qquad \overset{\text{Satz 53}}{\Longrightarrow} \qquad m \cdot \int_a^b g \leq \int_a^b f \cdot g \leq M \cdot \int_a^b g. \tag{279}$$

Nun ist die Funktion $F\colon [a,b] \ni x \mapsto f(x) \cdot \int_a^b g \in \mathbb{R}$ stetig, mit $F(u) = m \cdot \int_a^b g$ und $F(w) = M \cdot \int_a^b g$. Wegen (279) und Satz 36 (Zwischenwertsatz), existiert daher ein $\xi \in [a,b]$ mit[59]

$$\int_a^b f \cdot g = F(\xi) = f(\xi) \cdot \int_a^b g.$$

□

9.5 Der Hauptsatz der Differential- und Integralrechnung

Der Hauptsatz der Differential- und Integralrechnung zeigt, dass Ableiten und Integrieren zueinander inverse Operationen sind. Da Ableitungen von Funktionen in der Regel leicht zu berechnen sind, stellt er ein wichtiges Werkzeug zur Berechnung von Integralen dar.

Im Folgenden sei $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall.

Notation 28. *Sei $D \subseteq \mathbb{R}$ nichtentartet.*

[59] Genauer: Für $u < w$ zeigen Satz 36 und (279), die Existenz von $\xi \in [u,w] \subseteq [a,b]$ mit $F(\xi) = \int_a^b f \cdot g$; und für $u < w$, zeigen Satz 36 und (279), die Existenz von $\xi \in [w,u] \subseteq [a,b]$ mit $F(\xi) = \int_a^b f \cdot g$.

- *Sei $f\colon D \to \mathbb{R}$ stetig. Für $D \ni a < b \in D$ ist dann $f|_{[a,b]}$ ebenfalls stetig wegen Lemma 40.2), also Riemann-integrierbar wegen Satz 54.a). Wir definieren*

$$\int_a^b f := \int_a^b f|_{[a,b]}, \qquad \int_b^a f := -\int_a^b f \qquad \textit{sowie} \qquad \int_c^c f := 0 \quad \textit{für} \quad c \in [a,b]. \tag{280}$$

Die Formeln in Satz 53, gelten dann in entsprechend gleicher Form.

- *Für eine Abbildung $\alpha\colon D \to \mathbb{R}$ definieren wir* $\big($*alternativ* $\big[\alpha(t)\big]_{t=x}^{t=y}\big)$

$$[\alpha]_x^y := \alpha(y) - \alpha(x) \qquad \forall\, x, y \in D. \tag{281}$$

Übung 111. *Sei $D \subseteq \mathbb{R}$ nichtentartet, und $f\colon D \to \mathbb{R}$ stetig. Zeigen Sie*

$$\textstyle\int_a^b f = -\int_b^a f \qquad \textit{sowie} \qquad \int_a^b f = \int_a^c f + \int_c^b f \qquad \forall\, a,b,c \in D. \tag{282}$$

Hinweis: Fleiß und Fallunterscheidungen.

Terminologie 26 (Stammfunktion)**.** *Sei $f\colon D \to \mathbb{R}$ stetig (D nichtentartet).*

a) Eine Funktion $F\colon D \to \mathbb{R}$ heißt Stammfunktion *von f, wenn F differenzierbar ist mit $F' = f$.*

b) Es bezeichne $\int f \subseteq C^1(D,\mathbb{R})$ die Menge aller Stammfunktionen von f. Dann gilt

$$\textstyle\int f = \{F + c \mid c \in \mathbb{R}\} \qquad \textit{für jedes} \qquad F \in \int f, \tag{283}$$

*wobei Satz 57.*1) *(weiter unten) $\int f \neq \emptyset$ garantiert.* *($\int f \ni F_a\colon D \ni x \mapsto \int_a^x f \quad \forall\, a \in D$)*
Beweis von (283)*:*

- Ist F eine Stammfunktion von f, so für jedes $c \in \mathbb{R}$ auch $F + c$ wegen $(F + c)' = F' = f$.
- Sind $F, \tilde{F}$ Stammfunktionen von f, so gilt $(F - \tilde{F})' = f - f = 0$. Gemäß Korollar 32.2) ist somit $F - \tilde{F}$ konstant. □

Insbesondere gilt für je zwei Stammfunktionen $F, \tilde{F}$ von f

$$[F]_x^y = F(y) - F(x) = \tilde{F}(y) - \tilde{F}(x) = [\tilde{F}]_x^y. \tag{284}$$

Beispiel 77 (Stammfunktionen)**.**

a) Für eine Polynomfunktion $f : \mathbb{R} \ni x \mapsto \sum_{k=0}^n a_k \cdot x^k \in \mathbb{R}$, ist

$$F\colon \mathbb{R} \to \mathbb{R}, \qquad x \mapsto \textstyle\sum_{k=0}^n \frac{a_k}{k+1} \cdot x^{k+1}$$

eine Stammfunktion von f (es gilt $F' = f$ wegen Übung 94.c)).

b) Für $f = \exp_{\mathbb{R}}$ ist $F = \exp_{\mathbb{R}}$ wegen (244) *eine Stammfunktion.*

*c) Sei $\alpha \in \mathbb{R}$, und $f = \eta_\alpha\colon (0,\infty) \ni x \mapsto x^\alpha \in \mathbb{R}$ die zugehörige Potenzfunktion (Beispiel 69.*3)*):*

- *Für $\alpha \neq -1$ ist $F(x) = \frac{x^{\alpha+1}}{\alpha+1}$ eine Stammfunktion von f.* *(Beispiel 69.*3)*)*
- *Für $\alpha = -1$ ist $F(x) = \log(x)$ eine Stammfunktion von f.* *(Beispiel 69.*3)*)*

Speziell ist für alle $x > 0$: *(Satz 57.*2)*)*

$$\log(x) = \log(x) - \log(1) = [\log]_1^x = \textstyle\int_1^x \frac{1}{t}\,\mathrm{d}t.$$

Satz 57 (Hauptsatz der Differential- und Integralrechnung)**.** *Sei $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall, $a \in D$, und $f\colon D \to \mathbb{R}$ eine stetige Abbildung.*

1) *Die Abbildung*

$$F\colon D\to\mathbb{R},\qquad x\mapsto\int_a^x f$$

ist eine Stammfunktion von f.

2) *Ist $\tilde F\in C^1(D,\mathbb{R})$ eine Stammfunktion von f, so gilt*

$$\tilde F(x)-\tilde F(a)=[\tilde F]_a^x=\int_a^x f\qquad\forall\,x\in D.$$

Beweis. 1) Sei $x\in D$ und $h\neq 0$ mit $x+h\in D$. Wir erhalten mit Satz 53 und (282):

$$\left(\int_a^{x+h} f-\int_a^x f\right)\overset{(282)}{=}\int_x^{x+h} f\qquad\wedge\qquad f(x)\overset{(\text{I4})}{=}\tfrac{1}{h}\cdot\int_x^{x+h} f(x)$$

(auf der rechten Seite ist $f(x)$ als Konstante aufzufassen). Hiermit folgt

$$\frac{F(x+h)-F(x)}{h}-f(x)=\frac{1}{h}\cdot\left(\int_a^{x+h} f-\int_a^x f\right)-f(x)=\frac{1}{h}\cdot\int_x^{x+h}(f-f(x)).\tag{285}$$

Für $\varepsilon>0$ vorgegeben, sei nun $\delta>0$ mit $|f|_{(x-\delta,t+\delta)\cap D}-f(x)|<\varepsilon$ (Stetigkeit). Für $0<|h|\leq\delta$, erhalten wir mit Korollar 42 (zweiter Schritt)

$$\begin{aligned}\left|\frac{F(x+h)-F(x)}{h}-f(x)\right| &\overset{(285)}{=}\frac{1}{|h|}\cdot\left|\int_x^{x+h}(f-f(x))\right|\\ &\overset{(277)}{\leq}\frac{1}{|h|}\cdot\left|\int_x^{x+h}|f-f(x)|\right|\\ &\overset{(\text{I2})}{\leq}\frac{1}{|h|}\cdot\left|\int_x^{x+h}\varepsilon\right|\\ &\overset{(\text{I4})}{=}\varepsilon.\end{aligned}$$

Dies zeigt

$$\lim_{h\to 0}\frac{F(x+h)-F(x)}{h}=f(x)\qquad\text{also}\qquad F'(x)=f(x).$$

Da $x\in D$ beliebig war, ist F eine Stammfunktion von f.

2) Wir erhalten mit (284), Teil 1), und der dritten Konvention in (280)

$$[\tilde F]_a^x\overset{(284)}{=}[F]_a^x\overset{\text{Teil 1)}}{=}\int_a^x f-\int_a^a f\overset{(280)}{=}\int_a^x f.\qquad\square$$

Bemerkung 65 (Unbestimmtes Integral)**.** *Sei $f\colon D\to\mathbb{R}$ eine stetige Abbildung (D nichtentartet). Wegen Satz 57 haben wir für $x,y\in D$ vorgegeben, die wohldefinierte Abbildung*

$$[\,\cdot\,]_x^y\colon\int f\to\mathbb{R},\qquad\int f\ni F\mapsto[F]_x^y=\int_x^y f.\tag{286}$$

Die Menge $\int f$ *wird daher auch als das* unbestimmte Integral von f *bezeichnet (die Elemente von $\int f$ unterscheiden sich zwar durch additive Konstanten, liefern aber unter Anwendung von $[\,\cdot\,]_x^y$, alle den selben Integrationswert $\int_x^y f$).*

9.6 Integrationsregeln

Der wesentliche Vorteil des Hauptsatzes ist der, dass er es erlaubt, Integrale auszurechnen ohne explizit Ober- und Unterintegrale bestimmen zu müssen (was im Einzelnen sehr kompliziert sein kann). Man muss zu einer gegebenen stetigen Funktion $f\colon D \to \mathbb{R}$, nur irgendwie eine $F \in C^1(D, \mathbb{R})$ mit $F' = f$ finden (Stammfunktion), und diese dann entsprechend an den Integrationsgrenzen $x, y \in D$ in der Form $[F]_x^y$ auswerten. Ableitungen von konkreten Funktionen sind wegen der einfachen Rechenregeln aus Satz 53 und Satz 42 eben oft sehr viel leichter zu bestimmen als Integrale.

Jede der genannte Ableitungsregel zieht nun aber trotzdem eine entsprechende Integrationsregel nach sich, die in der ein oder anderen Situation nutzbringend angewendet werden kann. Das einfachste Beispiel liefert bereits Satz 53.(I3), der es erlaubt, Integrale von Linearkombinationen von Funktionen umgehen anzugeben, sofern man die entsprechenden Integrale der Summanden kennt. Diese Rechenregel entspricht natürlich der Ableitungsregel in Satz 40.1). Wir wollen nun noch die restlichen Ableitungsregeln auf das Riemann-Integral übertragen.

Im Folgenden sei wieder $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall.

9.6.1 Partielle Integration

Aus der Produktregel erhalten wir den Satz über die partielle Integration:

Satz 58 (Partielle Integration)**.** *Sind $f, g\colon D \to \mathbb{R}$ stetig differenzierbar ($f, g \in C^1(D, \mathbb{R})$), so gilt*

$$\int_a^b f \cdot g' = [f \cdot g]_a^b - \int_a^b f' \cdot g \qquad \forall\, a, b \in D.$$

Beweis. Gemäß Satz 40.1) ist $H := f \cdot g$ eine Stammfunktion von $h := f \cdot g' + f' \cdot g$. Satz 57.2) zeigt

$$[f \cdot g]_a^b = [H]_a^b = \textstyle\int_a^b h = \int_a^b (f \cdot g' + f' \cdot g) \overset{\text{(I3)}}{=} \int_a^b f \cdot g' + \int_a^b f' \cdot g. \qquad \square$$

Beispiel 78. *Seien $f, g \in C^1(D, \mathbb{R})$ und $a, b \in D$ wie in Satz 58. Für $g = \mathrm{id}_D$, erhalten wir (in vereinfachter Schreibweise)*

$$\int_a^b f(t)\,\mathrm{d}t = \big[f(t) \cdot t\big]_a^b - \int_a^b f'(t) \cdot t\,\mathrm{d}t. \tag{287}$$

a) Speziell ergibt sich aus (287) *für $f = \log$ auf $D = (0, \infty)$ mit $b = x$:*

$$\begin{aligned}\int_a^x \log(t)\;\mathrm{d}t &= \big[t \cdot \log(t)\big]_a^x - \int_a^x \log'(t) \cdot t\,\mathrm{d}t = \big[t \cdot \log(t)\big]_a^x - \int_a^x \frac{t}{t}\,\mathrm{d}t \\ &= \big[t \cdot \log(t)\big]_a^x - (x - a) = \big[t \cdot \log(t) - t\big]_a^x = \big[t \cdot (\log(t) - 1)\big]_a^x.\end{aligned} \tag{288}$$

Dies zeigt, dass $\tilde{F}\colon (0, \infty) \ni x \mapsto x \cdot (\log(x) - 1) \in \mathbb{R}$ eine Stammfunktion von $\log$ ist:

- *Gemäß Satz 57.1), ist $F\colon (0, \infty) \ni x \mapsto \int_a^x \log \in \mathbb{R}$ eine Stammfunktion von $\log$, d.h., $F' = \log$.*
- *Wegen* (288)*, gilt $F = \tilde{F} - \tilde{F}(a)$, also ebenfalls $\tilde{F}' = F' = \log$.*

b) Wir suchen eine Formel für die Integrale

$$A_m(x) := \int_0^x \frac{\mathrm{d}t}{(1 + t^2)^m} \qquad \forall\, x \in \mathbb{R},\ m \in \mathbb{N}_{>0}$$

bzw. eine Stammfunktion von $f_m\colon \mathbb{R} \ni x \mapsto (1+x^2)^{-m} \in (0,\infty)$ *für alle* $m \in \mathbb{N}_{>0}$. *In derartigen Fällen gelingt es manchmal eine rekursive Formel zu finden, indem man die Produktregel verwendet: Wir erhalten aus* (287)

$$\begin{aligned} A_m(x) &= \left[\frac{t}{(1+t^2)^m}\right]_0^x + \int_0^x \frac{t\cdot 2t\cdot m}{(1+t^2)^{m+1}}\,\mathrm{d}t \\ &= \frac{x}{(1+x^2)^m} + 2m\int_0^x \frac{1+t^2-1}{(1+t^2)^{m+1}}\,\mathrm{d}t \\ &= \frac{x}{(1+x^2)^m} + 2m\cdot A_m(x) - 2m\cdot A_{m+1}(x). \end{aligned}$$

Hieraus ergibt sich eine Rekursionsformel zur Berechnung dieser Stammfunktionen:

$$A_{m+1}(x) = \frac{2m-1}{2m}A_m(x) + \frac{x}{2m\cdot(1+x^2)^m}$$

(beachte $m > 0$*), wobei*

$$A_1(x) = \int_0^x \frac{\mathrm{d}t}{(1+t^2)} = [\arctan]_0^x = \arctan(x) - \arctan(0) = \arctan(x) \qquad \forall\, x\in\mathbb{R}$$

wegen Terminologie 25.(7) *und Satz 57.*2) *gilt.*

9.6.2 Das Integral der Umkehrfunktion

Aus dem Satz über die Umkehrabbildung (Satz 37), erhalten wir die folgende Integrationsregel:

Satz 59. *Sei* $f\colon \mathbb{R}\supseteq [a,b]\to\mathbb{R}$ *(*$a<b$*) differenzierbar und injektiv. Dann gilt* $\mathrm{im}(f) = [a',b']$ *mit*

- $a' = f(a) < f(b) = b'$, *wenn* f *streng monoton wachsend ist.*
- $a' = f(b) < f(a) = b'$, *wenn* f *streng monoton fallend ist.*

Es existiert die stetige Umkehrfunktion $f^{-1}\colon [a',b']\to[a,b]$ *(Satz 37), und es gilt*

$$\int_{f(a)}^{f(b)} f^{-1} = b\cdot f(b) - a\cdot f(a) - \int_a^b f. \tag{289}$$

(Beachte, dass f und f^{-1} wegen Lemma 47 und Satz 37 beide stetig (und monoton) sind (mit der gleichen Monotonieeigenschaft), also Riemann-interierbar gemäß Satz 54. Zudem ist $\mathrm{im}(f)$ ein Intervall, also zwangsweise von der angegebenen Gestalt.)

Beweis. Gemäß Satz 57.1) ist

- $F\colon [a\,,b\,] \ni x \mapsto \int_a^x f \in \mathbb{R}$ eine Stammfunktion von f, d.h., $F' = f$.
- $H\colon [a',b'] \ni y \mapsto \int_{f(a)}^y f^{-1} \in \mathbb{R}$ eine Stammfunktion von f^{-1}, d.h., $H' = f^{-1}$.

Wir betrachten die Abbildung

$$g\colon [a,b]\to\mathbb{R}, \qquad x\mapsto \underbrace{\int_a^x f}_{F(x)} + \underbrace{\int_{f(a)}^{f(x)} f^{-1}}_{(H\circ f)(x)} - x\cdot f(x) + a\cdot f(a).$$

Wegen Satz 40.2) (Produktregel), und Satz 41 (Kettenregel) ist g differenzierbar, mit

$$\begin{aligned} g'(x) &= f(x) + f^{-1}(f(x))\cdot f'(x) - 1\cdot f(x) - x\cdot f'(x) \\ &= f(x) + x\cdot f'(x) - f(x) - x\cdot f'(x) = 0. \end{aligned}$$

Gemäß Korollar 32.2) ist somit g konstant; also gilt $g(b) = g(a) = 0$, was (289) zeigt. $\square$

9.6.3 Die Substitutionsregel

Aus der Kettenregel erhalten wir die Substitutionsregel (auch Transformationssatz genannt):

Satz 60 (Transformationssatz)**.** *Sei $f\colon D \to \mathbb{R}$ stetig, sowie $\phi\colon \mathbb{R} \supseteq [a,b] \to D \subseteq \mathbb{R}$ stetig differenzierbar ($a < b$). Dann gilt $(f \circ \phi) \cdot \phi' \in C(D,\mathbb{R})$ mit*[60]

$$\int_a^b (f \circ \phi) \cdot \phi' = \int_{\phi(a)}^{\phi(b)} f.$$

Beweis. Es ist $(f \circ \phi) \cdot \phi'$ stetig als Verkettung und Produkt stetiger Funktionen ($\phi \in C^1([a,b], D)$), also Riemann-integrabel.

- Gemäß Satz 57.1) ist $F\colon D \ni x \mapsto \int_{\phi(a)}^x f \in \mathbb{R}$ eine Stammfunktion von f, d.h., $F' = f$.
- Gemäß Satz 41 (Kettenregel) ist $G := F \circ \phi\colon [a,b] \to \mathbb{R}$ differenzierbar mit

$$G' = (F' \circ \phi) \cdot \phi' = (f \circ \phi) \cdot \phi',$$

  d.h. G ist eine Stammfunktion von $(f \circ \phi) \cdot \phi'$.
- Wir erhalten mit Satz 57.2) (erster und letzter Schritt)

$$\int_a^b (f \circ \phi) \cdot \phi' = [G]_a^b = G(b) - G(a) = F(\phi(b)) - F(\phi(a)) = [F]_{\phi(a)}^{\phi(b)} = \int_{\phi(a)}^{\phi(b)} f. \qquad \square$$

Beispiel 79 ($*$)**.** *Wir wollen $\int_0^y x \cdot \sqrt{1+x}\,\mathrm{d}x$ für $y > 0$ berechnen, d.h. das Integral $\int_0^y \alpha$ für die stetige Abbildung $\alpha\colon [0,\infty) \ni x \mapsto x \cdot \sqrt{1+x} \in \mathbb{R}$. Seien hierfür*

$$\begin{aligned} \phi\colon [0,y] &\to [0,\infty), & x &\mapsto \sqrt{1+x} \\ f\colon \mathbb{R} &\to \mathbb{R}, & u &\mapsto 2 \cdot u^2 \cdot (u^2 - 1) = 2 \cdot (u^4 - u^2). \end{aligned}$$

- *$F\colon \mathbb{R} \ni u \mapsto 2 \cdot (\frac{u^5}{5} - \frac{u^3}{3}) \in \mathbb{R}$ ist eine Stammfunktion von f.*
- *Es gilt $\phi^2(x) - 1 = x$ sowie $\phi'(x) = \frac{1}{2\sqrt{1+x}} = \frac{1}{2 \cdot \phi(x)}$ (Kettenregel) für alle $x \in [0,y]$, also*

$$\alpha = (\phi^2 - 1) \cdot \phi = (\phi^2 - 1) \cdot \phi \cdot \underbrace{2 \cdot \phi(x) \cdot \phi'(x)}_{=1} = (f \circ \phi) \cdot \phi'.$$

Wir erhalten mit Satz 60 (dritte Identität)

$$\begin{aligned} \int_0^y x \cdot \sqrt{1+x}\,\mathrm{d}x &= \int_0^y \alpha = \int_0^y (f \circ \phi) \cdot \phi' = \int_{\phi(0)}^{\phi(y)} f = \big[F\big]_{\phi(0)}^{\phi(y)} = \big[F\big]_1^{\sqrt{1+y}} \\ &= \frac{2}{5} \cdot (1+y)^{\frac{5}{2}} - \frac{2}{3} \cdot (1+y)^{\frac{3}{2}} - \frac{2}{5} + \frac{2}{3}. \end{aligned}$$

Übung 112. *Sei $\mathbb{R} \ni a < b \in \mathbb{R}$, sowie $f\colon D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}$ nichtentartet. Zeigen Sie:*

a) $\int_a^b f(x+c)\,\mathrm{d}x = \int_{a+c}^{b+c} f(u)\,\mathrm{d}u$ *für* $c \in \mathbb{R}$ *mit* $[a+c, b+c] \subseteq D$.

b) $c \cdot \int_a^b f(c \cdot x)\,\mathrm{d}x = \int_{c \cdot a}^{c \cdot b} f(u)\,\mathrm{d}u$ *für* $c \in \mathbb{R}$ *mit* $[c \cdot a, c \cdot b] \subseteq D$.

c) $\int_a^b x^{n-1} \cdot f(x^n)\,\mathrm{d}x = \frac{1}{n} \cdot \int_{a^n}^{b^n} f(u)\,\mathrm{d}u$ *falls* $[a^n, b^n] \subseteq D$.

[60] In anderer Notation: $\int_a^b f(\phi(t)) \cdot \phi'(t)\,\mathrm{d}t = \int_{\phi(a)}^{\phi(b)} f(u)\,\mathrm{d}u$

9.7 Integrale und Funktionenfolgen

In der Integrationstheorie ist es von grundlegender Bedeutung, das Konvergenzverhalten von Folgen der Gestalt $(\int_a^b f_n)_{n\in\mathsf{N}}$ zu kennen, wobei $a < b$ und $f_n \in \mathcal{R}_a^b$ für alle $n \in \mathsf{N}$ gilt. Man ist also an möglichst einfachen Kriterien interessiert, die die Existenz des Grenzwertes $\lim_n \int_a^b f_n \in \overline{\mathbb{R}}$ sicherstellen, bzw. es sogar erlauben, diesen explizit anzugeben. Spezieller ist man hier zunächst an der Vertauschbarkeit von Grenzwertbildung und Integration interessiert, d.h. man will wissen, ob

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b \lim_{n\to\infty} f_n \tag{290}$$

gilt, wobei $(f_n)_{n\in\mathsf{N}}$ in einem „geeigneten Sinne“ gegen eine Grenzfunktion $f \in \mathcal{R}_a^b$ konvergiert. Das nächste Beispiel zeigt, dass hiermit nicht die punktweise Konvergenz von $(f_n)_{n\in\mathsf{N}}$ gegen f gemeint sein kann – zumindest dann nicht, wenn man keine weiteren Bedingungen an die Funktionenfolge $(f_n)_{n\in\mathsf{N}}$ stellen will:

Beispiel 80. *Für $n \in \mathbb{N}_{>0}$, definieren wir die stetigen Funktionen ($a = 0$ und $b = 1$)*

$$f_n : [0,1] \to \mathbb{R}, \qquad f_n(x) = \begin{cases} n^2 x & \textit{für} \quad 0 \le x \le \frac{1}{n} \\ 2n - n^2 x & \textit{für} \quad \frac{1}{n} < x \le \frac{2}{n} \\ 0 & \textit{für} \quad x > \frac{2}{n} \end{cases}$$

(beachte $f_n(\frac{1}{n}) = n = 2n - n^2 \cdot \frac{1}{n}$ sowie $f_n(\frac{2}{n}) = 2n - n^2 \cdot \frac{2}{n} = 0$).

- *$(f_n)_{n\in\mathsf{N}}$ konvergiert punktweise gegen $f = 0$.*

  Beweis der Behauptung: Es gilt $f_n(0) = 0$ für alle $n \in \mathbb{N}_{>0}$. Für $x \in (0,1]$ gilt zudem $f_n(x) = 0$ für alle $n \in \mathbb{N}_{>0}$ mit $n > \frac{2}{x}$. □

- *Für jedes $n \in \mathbb{N}_{>0}$, gilt $\int_0^1 f_n = 1$ wegen*

$$\begin{aligned} \textstyle\int_0^1 f_n &= \textstyle\int_0^{\frac{1}{n}} n^2 \cdot x \,\mathrm{d}x + \int_{\frac{1}{n}}^{\frac{2}{n}} n^2 \cdot (\frac{2}{n} - x) \,\mathrm{d}x + \int_{\frac{2}{n}}^1 0 \,\mathrm{d}x \\ &= n^2 \cdot \Big[\tfrac{x^2}{2}\Big]_0^{\frac{1}{n}} + n^2 \cdot \Big[-\tfrac{1}{2} \cdot (\tfrac{2}{n} - x)^2\Big]_{\frac{1}{n}}^{\frac{2}{n}} \\ &= n^2 \cdot \tfrac{1}{2n^2} + n^2 \cdot \tfrac{1}{2n^2} = 1. \end{aligned}$$

Somit gilt: $\lim_n \int_0^1 f_n = 1 \neq 0 = \int_0^1 f$

9.7.1 Vertauschbarkeit von Grenzwertbildung und Integration

Satz 61. *Sei $(f_n)_{n\in\mathsf{N}}$ eine Folge in $\mathcal{R}_a^b$ mit $a < b$, die gleichmäßig gegen ein $f \in \mathsf{Abb}([a,b],\mathbb{R})$ konvergiert. Dann gilt $f \in \mathcal{R}_a^b$ sowie*[61]

$$\int_a^b f = \lim_{n\to\infty} \int_a^b f_n.$$

Beweis. Wegen der gleichmäßigen Konvergenz gilt:

$$\begin{gathered} \forall\, \varepsilon > 0\colon\ \exists\, N_\varepsilon \in \mathsf{N}\colon\ \forall\ (\ x \in [a,b] \ \ \wedge\ \ n \ge N_\varepsilon\ )\colon \qquad |f(x) - f_n(x)| < \varepsilon \\ \text{also} \\ f_n - \varepsilon \le f \le f_n + \varepsilon \qquad \forall\, n \ge N_\varepsilon. \end{gathered} \tag{291}$$

[61]Natürlich gilt dann ebenfalls $\int_b^a f = -\int_a^b f = -\lim_{n\to\infty} \int_a^b f_n = \lim_{n\to\infty} \int_b^a f_n$.

Gilt $f \in \mathcal{R}_a^b$, so folgt umgehend aus Satz 53 und Satz 55

$$\left|\int_a^b f - \int_a^b f_n\right| \overset{(\text{I3})}{=} \left|\int_a^b (f - f_n)\right| \overset{\text{Satz 55}}{\leq} \int_a^b |f - f_n| \overset{(\text{I2})}{\leq} \int_a^b \chi_\varepsilon \overset{(\text{I4})}{=} \varepsilon \cdot (b-a) \qquad \forall\, n \geq N_\varepsilon,$$

was $\lim_n \int_a^b f_n = \int_a^b f$ zeigt. Es bleibt daher nur noch $f \in \mathcal{R}_a^b$ nachzuweisen, also das Kriterium (269).

Sei hierfür $\varepsilon > 0$ vorgegeben; und setze $\tilde{\varepsilon} := \frac{\varepsilon}{2(b-a)+1}$ sowie $N := N_{\tilde{\varepsilon}}$.

- Gemäß (269) (für $f_N \in \mathcal{R}_a^b$), existieren $\tilde{\phi}, \tilde{\psi} \in \mathcal{T}_a^b$ mit

$$\tilde{\phi} \leq f_N \leq \tilde{\psi} \qquad \wedge \qquad \int_a^b \tilde{\psi} - \int_a^b \tilde{\phi} < \tilde{\varepsilon}.$$

- Aus (291) erhalten wir

$$\mathcal{T}_a^b \ni \phi \;:=\; \tilde{\phi} - \tilde{\varepsilon} \;\leq\; f_N - \tilde{\varepsilon} \;\overset{(291)}{\leq}\; f \;\overset{(291)}{\leq}\; f_N + \tilde{\varepsilon} \;\leq\; \tilde{\psi} + \tilde{\varepsilon} \;=:\; \psi \in \mathcal{T}_a^b$$

mit

$$\int_a^b \psi - \int_a^b \phi = 2 \cdot \int_a^b \tilde{\varepsilon} + \int_a^b \tilde{\psi} - \int_a^b \tilde{\phi} < (2(b-a)+1) \cdot \tilde{\varepsilon} = \varepsilon,$$

was wegen (269) die Behauptung zeigt. □

9.7.2 Vertauschbarkeit von Grenzwertbildung und Ableitung

Satz 61 impliziert eine entsprechende Aussage für Ableitungen.

Satz 62. *Sei $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall, $p \in D$ ein Punkt; und $(f_n)_{n \in \mathsf{N}}$ eine Folge in $C^1(D, \mathbb{R})$ (d.h. $f_n\colon D \to \mathbb{R}$ ist stetig differenzierbar für alle $n \in \mathsf{N}$), sodass gilt:*

a) $f_p := \lim_n f_n(p) \in \mathbb{R}$,

b) $(f_n')_{n \in \mathsf{N}}$ *konvergiert gleichmäßig gegen ein* $\alpha \in \mathsf{Abb}(D, \mathbb{R})$.

1) *Es gilt $\alpha \in C(D, \mathbb{R})$ wegen Proposition 12; und $(f_n)_{n \in \mathsf{N}}$ konvergiert punktweise gegen die Stammfunktion (vgl. Satz 57.1))*

$$f\colon D \to \mathbb{R}, \qquad x \mapsto f_p + \int_p^x \alpha$$

von α, d.h. $f \in C^1(D, \mathbb{R})$ mit $f' = \alpha = \lim_n f_n' \in C(D, \mathbb{R})$.

(*Beachte, dass wegen Satz 57.2) gilt:*

$$f_n(x) = f_n(p) + \int_p^x f_n'(t)\,\mathrm{d}t \qquad \forall\, x \in D,\ n \in \mathsf{N}\ \Big) \tag{292}$$

2) *Ist $B \subseteq D$ ein beschränktes Teilintervall, so konvergiert $(f_n|_B)_{n \in \mathsf{N}} \to f|_B$ gleichmäßig.*

Beweis. 1) Satz 61, angewandt auf

- $(f_n'|_{[p,x]})_{n \in \mathsf{N}}$ und $\alpha|_{[p,x]}$ für $p < x \in D$,
- $(f_n'|_{[x,p]})_{n \in \mathsf{N}}$ und $\alpha|_{[x,p]}$ für $D \in x < p$,

liefert zusammen mit a):

$$f(x) = f_p + \int_p^x \alpha \overset{a),\,\text{Satz 61}}{=} \lim_n f_n(p) + \lim_n \int_p^x f_n' \overset{(292)}{=} \lim_n f_n(x).$$

2) Sei $d \geq 1$ mit $-d \leq B \leq d$, also $-c \leq B - p \leq c$ für $c := d + |p| \geq 1$.

Für $\varepsilon > 0$ vorgegeben, existiert nun wegen a) und Teil b) ein $N_\varepsilon \in \mathsf{N}$, sodass für $n \geq N_\varepsilon$ gilt:

$$|f_n(p) - f_p| < \tfrac{\varepsilon}{2} \qquad \wedge \qquad |f_n'(x) - \alpha(x)| < \tfrac{\varepsilon}{2c} \quad \forall\, x \in D$$

Für $n \geq N_\varepsilon$ und alle $x \in B$, folgt mit Teil 1) und (292) (erster Schritt), Korollar 42 (zweiter Schritt), sowie Satz 53 (dritter Schritt), dass

$$\begin{aligned} |f_n(x) - f(x)| &\leq |f_n(p) - f_p| \ + \left| \textstyle\int_p^x (f_n'(x) - \alpha(x)) \right| \\ &\leq |f_n(p) - f_p| + \left| \textstyle\int_p^x |f_n'(x) - \alpha(x)| \right| \\ &< \frac{\varepsilon}{2} + \underbrace{|x - p|}_{\leq c} \cdot \frac{\varepsilon}{2c} < \varepsilon \end{aligned}$$

gilt, was die Behauptung zeigt. □

9.7.3 Ableitung und Integration von Potenzreihen

Terminologie 27. *Sei* $\mathsf{Q}(x) = \sum_{n=0}^{\infty} c_n \cdot (x-p)^n$ *eine reelle Potenzreihe im Entwicklungspunkt* $p \in \mathbb{R}$. *Die* formale Ableitung *und das* formale Integral *von* Q *sind definiert als die Potenzreihen*

$$\underbrace{\mathsf{Q}_\downarrow(x) = \sum_{n=1}^{\infty} n \cdot c_n \cdot (x-p)^{n-1}}_{\text{formale Ableitung}} \qquad \textit{und} \qquad \underbrace{\mathsf{Q}^\uparrow(x) = \sum_{n=0}^{\infty} \frac{c_n}{n+1} \cdot (x-p)^{n+1}}_{\text{formales Integral}}.$$

(*Formell korrekter definiert man diese durch*

$$\mathsf{Q}_\downarrow(x) = \sum_{n=0}^{\infty} (n+1) \cdot c_{n+1} \cdot (x-p)^n \qquad \textit{und} \qquad \mathsf{Q}^\uparrow(x) = \sum_{n=1}^{\infty} \frac{c_{n-1}}{n} \cdot (x-p)^n.\,)$$

Bemerkung 66. *Seien* $(\lambda_n)_{n \in \mathsf{N}}$, $(\mu_n)_{n \in \mathsf{N}}$ *Folgen in* $[0, \infty)$ *mit* $\lim_n \lambda_n = \lambda \in [0, \infty)$. *Dann gilt*

$$\overline{\lim}_n (\lambda_n \cdot \mu_n) = \lambda \cdot \overline{\lim}_n \, \mu_n.$$

Beweis der Behauptung: Für $\varepsilon > 0$ vorgegeben, existiert ein $N_\varepsilon \in \mathsf{N}$ mit $(\lambda - \varepsilon) \cdot \mu_n \leq \lambda_n \cdot \mu_n \leq (\lambda + \varepsilon) \cdot \mu_n$ für alle $n \geq N_\varepsilon$. Wir erhalten mit Übung 69 und Übung 74, dass

$$(\lambda - \varepsilon) \cdot \overline{\lim}_n \, \mu_n = \overline{\lim}_n \big((\lambda - \varepsilon) \cdot \mu_n\big) \leq \overline{\lim}_n (\lambda_n \cdot \mu_n) \leq \overline{\lim}_n \big((\lambda + \varepsilon) \cdot \mu_n\big) = (\lambda + \varepsilon) \cdot \overline{\lim}_n \, \mu_n$$

gilt. Da $\varepsilon > 0$ beliebig war, folgt hiermit die Behauptung. □

Lemma 61. *Sei* $\mathsf{Q}(x) = \sum_{n=0}^{\infty} c_n \cdot (x-p)^n$ *eine reelle Potenzreihe im Entwicklungspunkt* $p \in \mathbb{R}$. *Dann haben* $\mathsf{Q}_\downarrow$ *und* $\mathsf{Q}^\uparrow$ *den gleichen Konvergenzradius wie* Q.

Beweis. Für $p \neq x \in \mathbb{R}$ gilt:

$$\mathsf{Q}_\downarrow(x) \ \text{konvergiert} \quad \Longleftrightarrow \quad (x-p) \cdot \mathsf{Q}_\downarrow(x) = \sum_{n=1}^{\infty} \underbrace{n \cdot c_n}_{=:\, a_n} \cdot (x-p)^n =: \tilde{\mathsf{Q}}_\downarrow(x) \ \text{konvergiert}$$

$$\mathsf{Q}^\uparrow(x) \ \text{konvergiert} \quad \Longleftrightarrow \quad (x-p)^{-1} \cdot \mathsf{Q}^\uparrow(x) = \sum_{n=0}^{\infty} \underbrace{\frac{c_n}{n+1}}_{=:\, b_n} \cdot (x-p)^n =: \tilde{\mathsf{Q}}^\uparrow(x) \ \text{konvergiert}$$

Somit haben $\mathsf{Q}_\downarrow$ und $\tilde{\mathsf{Q}}_\downarrow$ bzw. $\mathsf{Q}^\uparrow$ und $\tilde{\mathsf{Q}}^\uparrow$ jeweils den gleichen den gleichen Konvergenzbereich, also auch den gleichen Konvergenzradius. Es bezeichne $\mathsf{R}, \tilde{\mathsf{R}}_\downarrow, \tilde{\mathsf{R}}^\uparrow$ die Konvergenzradien von $\mathsf{Q}, \tilde{\mathsf{Q}}_\downarrow, \tilde{\mathsf{Q}}^\uparrow$ (in der gegeben Reihenfolge). Nach der Formel von Hadamard ((183) in Satz 29.c)), gilt

$$\mathsf{R} = \frac{1}{\overline{\lim}_n \sqrt[n]{|c_n|}}, \qquad \tilde{\mathsf{R}}_\downarrow = \frac{1}{\overline{\lim}_n \sqrt[n]{|a_n|}}, \qquad \tilde{\mathsf{R}}^\uparrow = \frac{1}{\overline{\lim}_n \sqrt[n]{|b_n|}}.$$

Wegen Lemma 38 gilt $\lim_n \sqrt[n]{n} = 1$; und in Bemerkung 48.b) hatten wir gesehen, dass ebenfalls $\lim_n \sqrt[n]{n+1} = 1$ gilt. Wir erhalten mit Bemerkung 66

$$\overline{\lim}_n \sqrt[n]{|a_n|} = \overline{\lim}_n \sqrt[n]{n} \cdot \sqrt[n]{|c_n|} = \overline{\lim}_n \sqrt[n]{|c_n|} = \overline{\lim}_n \tfrac{1}{\sqrt[n]{n+1}} \cdot \sqrt[n]{|c_n|} = \overline{\lim}_n \sqrt[n]{|b_n|},$$

also $\tilde{\mathsf{R}}_\downarrow = \mathsf{R} = \tilde{\mathsf{R}}^\uparrow$. □

Satz 63. *Sei* $\mathsf{Q}(x) = \sum_{n=0}^\infty c_n \cdot (x-p)^n$ *eine reelle Potenzreihe im Entwicklungspunkt* $p \in \mathbb{R}$, *mit Konvergenzradius* $\mathsf{R}[\mathsf{Q}] > 0$. *Dann ist* $f := \mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]}$ *differenzierbar mit Ableitung* $f' = \mathsf{Q}_\downarrow|_{\mathrm{K}[\mathsf{Q}]}$, *und* $F := \mathsf{Q}^\uparrow|_{\mathrm{K}[\mathsf{Q}]}$ *ist eine Stammfunktion von* f *(genauer* $F(x) = \int_p^x f$ *für alle* $x \in \mathrm{K}[\mathsf{Q}]$*).*

Beachte: Wegen Lemma 61 sind $\mathsf{Q}_\downarrow|_{\mathrm{K}[\mathsf{Q}]}$ und $\mathsf{Q}^\uparrow|_{\mathrm{K}[\mathsf{Q}]}$ als Abbildungen definiert.

Beweis. Für $n \in \mathbb{N}$ definieren wir

$$\begin{aligned} g_n &\colon \mathbb{R} \to \mathbb{R}, & x &\mapsto c_n \cdot (x-p)^n \\ G_n &\colon \mathbb{R} \to \mathbb{R}, & x &\mapsto \frac{c_n}{n+1} \cdot (x-p)^{n+1}. \end{aligned}$$

Dann gilt $G_n' = g_n$, d.h. G_n ist eine Stammfunktion von g_n mit (Satz 57.2))

$$\int_p^x g_n = [G_n]_p^x = \frac{c_n}{n+1} \cdot (x-p)^{n+1} \qquad \forall\, x \in \mathbb{R}. \tag{293}$$

Gemäß Satz 39, konvergiert nun für $0 < r < \mathsf{R}[\mathsf{Q}]$, die Folge $(\sum_{k=0}^n g_n|_{\mathsf{B}_r(p)})_{n \in \mathbb{N}}$ (stetiger beschränkter Abbildungen) gleichmäßig gegen die (stetige und beschränkte Abbildung) $\mathsf{Q}|_{\mathsf{B}_r(p)}$. Wir erhalten für $p - r < x < p + r$ mit Satz 61 (dritter Schritt)[62] sowie Satz 53 (vierter Schritt)

$$\begin{aligned} \int_p^x f &= \int_p^x f|_{\mathsf{B}_r(p)} = \int_p^x \mathsf{Q}|_{\mathsf{B}_r(p)} \\ &= \lim_n \int_p^x \sum_{k=0}^n g_n|_{\mathsf{B}_r(p)} \overset{(13)}{=} \lim_n \sum_{k=0}^n \int_p^x g_n|_{\mathsf{B}_r(p)} \overset{(293)}{=} \lim_n \sum_{k=0}^n \frac{c_n}{n+1} \cdot (x-p)^{n+1} \\ &= \mathsf{Q}^\uparrow(x) = F(x). \end{aligned}$$

Wegen Satz 57.1) gilt somit $F'(x) = f(x)$ für alle $x \in \mathsf{B}_r(p)$. Da $0 < r < \mathsf{R}[\mathsf{Q}]$ beliebig war, folgt hiermit $F' = f$.

Die gleiche Argumentation zeigt nun auch, dass f (genauer $\tilde{F} := f - c_0$) eine Stammfunktion von $\tilde{f} := \mathsf{Q}_\downarrow|_{\mathrm{K}[\mathsf{Q}]}$ ist, d.h., dass $f' = \tilde{F}' = \tilde{f} = \mathsf{Q}_\downarrow|_{\mathrm{K}[\mathsf{Q}]}$ gilt.[63] □

Bemerkung 67. *Sei* $\mathsf{Q}(x) = \sum_{n=0}^\infty c_n \cdot (x-p)^n$ *eine reelle Potenzreihe im Entwicklungspunkt* $p \in \mathbb{R}$, *mit Konvergenzradius* $\mathsf{R}[\mathsf{Q}] > 0$. *Per Induktion folgt aus Satz 63 und Lemma 61:*

[62] Formell korrekter, wendet man hier Satz 61 wieder auf die entsprechenden Einschränkungen auf $[p, x]$ für $p < x$ bzw. $[x, p]$ für $x < p$ an.

[63] Alternativ kann man auch Satz 62 bemühen.

a) *Die Abbildung* $f := \mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]}$ *ist glatt, mit*

$$f^{(k)} = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} \cdot c_n \cdot (x-p)^{(n-k)} \qquad \forall\, x \in \mathrm{K}[\mathsf{Q}],\ k \in \mathbb{N}_{>0}.$$

b) *Für jedes* $k \in \mathbb{N}$, *ist*

$$F[k]\colon \mathrm{K}[\mathsf{Q}] \to \mathbb{R}, \qquad x \mapsto \overbrace{\sum_{n=0}^{\infty} \frac{n!}{(n+k)!} \cdot c_n \cdot (x-p)^{(n+k)}}^{=\sum_{n=k}^{\infty} \frac{(n-k)!}{n!} \cdot c_{n-k} \cdot (x-p)^n} \qquad \forall\, x \in \mathrm{K}[\mathsf{Q}]$$

definiert und glatt, mit $F[k]^{(k)} = f$ *sowie* $F[k]^{(\ell)}(p) = 0$ *für* $0 \leq \ell \leq k-1$.

Korollar 43. *Sei* $\mathsf{Q}(x) = \sum_{n=0}^{\infty} c_n \cdot (x-p)^n$ *eine reelle Potenzreihe im Entwicklungspunkt* $p \in \mathbb{R}$ *mit Konvergenzradius* $\mathsf{R}[\mathsf{Q}] > 0$. *Dann ist* $f := \mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]}$ *glatt mit*

$$c_k = \frac{f^{(k)}(p)}{k!} \qquad \forall\, k \in \mathbb{N}. \tag{294}$$

Insbesondere gilt $f = 0$ *genau dann, wenn* $c_k = 0$ *für alle* $k \in \mathbb{N}$ *gilt.*

Beweis. Die Glattheit von f sowie die Identität (294) folgt sofort aus Bemerkung 67.a).

- Gilt $c_k = 0$ für alle $k \in \mathbb{N}$, so offensichtlich auch $f = 0$.
- Gilt $f = 0$, so gilt auch $f^{(k)} = 0$ für alle $k \in \mathbb{N}$, also $c_k = \frac{f^{(k)}(p)}{k!} = 0$ für alle $k \in \mathbb{N}$. $\square$

Bemerkung 68 (Abelscher Grenzwertsatz $*$). *Konvergiert eine reelle Potenzreihe* $\mathsf{Q}(x) = \sum_{n=0}^{\infty} c_n \cdot (x-p)^n$ *auf* $[p, p+R]$ *(analog* $[p-R, p]$*) für ein* $R > 0$, *so besagt der* Abelsche Grenzwertsatz, *dass sie dort sogar gleichmäßig konvergiert, gemäß Proposition 12 also auf* $[p, R]$ *stetig ist.*

Beispiel 81.

a) *Reihenentwicklung von* $\arctan|_{(-1,1)}$: *Für* $x \in (-1,1)$ *zeigt Terminologie 25.*(7) *(erster Schritt) und Korollar 22.2) (geometrische Reihe):*

$$\arctan'(x) = \frac{1}{1+x^2} = \frac{1}{1-(-x^2)} = \sum_{n=0}^{\infty} (-x^2)^n = \sum_{n=0}^{\infty} (-1)^n \cdot x^{2n} \qquad \forall\, x \in (-1,1),$$

wobei der Konvergenzradius der Reihe auf der rechten Seite 1 *ist (geometrische Reihe). Mit Satz 63 erhalten wir durch Integration die Reihenentwicklung (beachte* $\arctan(0) = 0$*)*

$$\arctan(x) = \int_0^x \frac{\mathrm{d}t}{1+t^2} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \cdot x^{2n+1} \qquad \forall\, x \in (-1,1) \tag{295}$$

der Arkustangensfunktion auf $(-1,1)$ *(gleicher Konvergenzradius).*

$*$ *Nach dem Leibnizkriterium ist nun die Reihe* $\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}$ *konvergent (also die rechte Seite von* (295) *für* $x = 1$*). Mit dem Abelschen Grenzwertsatz (Bemerkung 68) und der Stetigkeit von* arctan *erhält man (vgl. Beispiel 45)*

$$\begin{aligned} \frac{\pi}{4} &= \arctan(1) = \lim_{x \uparrow 1} \arctan(x) = \lim_{x \uparrow 1} \Big(\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \cdot x^{2n+1} \Big) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \\ &= 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} \pm \dots . \end{aligned}$$

b) Reihenentwicklung von $\log(1+\cdot)|_{(-1,1)}$*:* *Für* $x \in (-1,1)$ *zeigt Beispiel 69.*2) *und Korollar 22.2)*

$$\log'(1+x) = \frac{1}{1+x} = \frac{1}{1-(-x)} = \sum_{n=0}^{\infty} (-1)^n \cdot x^n,$$

wobei der Konvergenzradius der Reihe auf der rechten Seite 1 *ist. Wie in Teil a), erhalten aus Satz 63 (beachte* $\log(1) = 0$*)*

$$\log(1+x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} \cdot x^{n+1} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \cdot x^n. \tag{296}$$

$*$ *Mit dem Leibnizkriterium und dem Abelschen Grenzwertsatz folgt (vgl. Beispiel 45)*

$$\log(2) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} \pm \dots .$$

10 Taylorpolynome und Taylorreihen

In diesem Kapitel behandeln wir die Taylorentwicklung von Funktion. Diese erlaubt es, C^n-Funktionen ($n \geq 1$) lokal durch Polynome zu approximieren, bzw. im glatten Fall ($n = \infty$) manchmal sogar durch Reihen exakt darzustellen (vgl. Beispiel 81). Im gleichen Sinne, wie die Differenzierbarkeit einer Funktion es erlaubt, diese lokal durch eine affine Funktion anzunähern, liefert die n-malige Differenzierbarkeit die lokale Approximierbarkeit durch Polynome n-ten Grades. Die Taylorformel ist die Grundlage vieler Anwendungen in der Analysis im Bereich der Angewandten Mathematik; und insbesondere in der Physik, da sie es erlaubt mit Näherungen zu rechnen, wenn die exakten Formeln zu kompliziert werden.

Im Folgenden sei $D \subseteq \mathbb{R}$ immer ein nichtentartetes Intervall.

10.1 Die Taylorentwicklung

Um die Grundidee der Taylorentwicklung zu verstehen, betrachten wir zunächst Polynomfunktionen.

Bemerkung 69 (Darstellungen von Polynomfunktionen)**.** *Wir werden im Folgenden auch Funktionen der Gestalt* $f\colon \mathbb{R} \ni x \mapsto \sum_{k=0}^{n} a_k \cdot (x-p)^k \in \mathbb{R}$*, für* $p \in \mathbb{R}$ *sowie* $a_0, \dots, a_n \in \mathbb{R}$ *mit* $a_n \neq 0$*, als* Polynomfunktionen vom Grad n *bezeichnen:*

a) Mit Hilfe des Binomischen Lehrsatzes hatten wir im Beweis von Satz 33 bereits eingesehen:

- *Sei* $f\colon \mathbb{R} \ni x \mapsto \sum_{k=0}^{n} c_k \cdot x^k \in \mathbb{R}$ *eine Polynomfunktion vom Grad* $n \in \mathbb{N}$ *(d.h.* $c_0, \dots, c_n \in \mathbb{R}$ *mit* $c_n \neq 0$*). Gegeben ein* $p \in \mathbb{R}$*, so gilt:* Insbesondere gilt: $a_n = c_n \neq 0$

  $$f(x) = \textstyle\sum_{k=0}^{n} a_k \cdot (x-p)^k \quad \forall\, x \in \mathbb{R} \qquad \textit{mit} \qquad a_k = \sum_{\ell=k}^{n} c_\ell \cdot \binom{\ell}{k} \cdot p^{\ell-k} \quad \forall\, 0 \leq k \leq n.$$

- *Sei* $f\colon \mathbb{R} \ni x \mapsto \sum_{k=0}^{n} a_k \cdot (x-p)^k \in \mathbb{R}$ *eine Polynomfunktion vom Grad* $n \in \mathbb{N}$ *(d.h.,* $p \in \mathbb{R}$ *und* $a_0, \dots, a_n \in \mathbb{R}$ *mit* $a_n \neq 0$*). Dann gilt:*[64] Insbesondere gilt: $c_n = a_n \neq 0$

  $$f(x) = \textstyle\sum_{k=0}^{n} c_k \cdot x^k \quad \forall\, x \in \mathbb{R} \qquad \textit{mit} \qquad c_k = \sum_{\ell=k}^{n} a_\ell \cdot \binom{\ell}{k} \cdot (-p)^{\ell-k} \quad \forall\, 0 \leq k \leq n.$$

[64] D.h., f ist eine Polynomfunktion vom Grad n in unserem ursprünglichen Sinne.

b) Sei $f\colon \mathbb{R} \to \mathbb{R}$ eine Polynomfunktion vom Grad $n \in \mathbb{N}$, und sei $p \in \mathbb{R}$ fixiert. Gemäß Teil a) existieren $a_0, \dots, a_n \in \mathbb{R}$ mit $a_n \neq 0$, sodass:

$$f(x) = \sum_{k=0}^{n} a_k \cdot (x-p)^k \qquad \forall\, x \in \mathbb{R}. \tag{297}$$

Für $0 \le m \le n$ erhalten wir durch m-faches Ableiten:

$$\begin{aligned} f^{(m)}(x) &= \sum_{k=m}^{n} a_k \cdot k \cdot (k-1) \cdot \ldots \cdot (k-m+1) \cdot (x-p)^{k-m} \\ &= \sum_{k=m}^{n} a_k \cdot \frac{k!}{(k-m)!} \cdot (x-p)^{k-m}. \end{aligned} \tag{298}$$

- *Für $m = n$ erhalten wir $f^{(n)} = a_n \cdot n!$ – also gilt $f^{(n+\ell)} = 0$ für alle $\ell \in \mathbb{N}_{>0}$.*
- *Auswerten von (298) an der Stelle $x = p$ liefert:*

$$a_m = \frac{f^{(m)}(p)}{m!} \qquad \forall\, 0 \le m \le n. \tag{299}$$

Somit sind die Koeffizienten $a_0, \dots, a_n$ eindeutig bestimmt durch die ersten n Ableitungen von f in p, und somit

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(p)}{k!} \cdot (x-p)^k \qquad \forall\, x \in \mathbb{R} \tag{300}$$

die notwendig eindeutige *Darstellung von f in der Form (297) (also mit $a_n \neq 0$).*[65]

Zusammen genommen zeigen nun die Teile a) und b) sogar, dass eine Polynomfunktion vom Grad $n \in \mathbb{N}$ eindeutig durch ihre Ableitungen bis zur Ordnung n in jedem beliebig vorgegebenen (aber fixierten) Punkt $p \in \mathbb{R}$ bestimmt ist.

Definition 65. *Sei D nichtentartet, $f\colon D \to \mathbb{R}$ eine Funktion, $p \in D$ ein Punkt, und $n \in \mathbb{N}$.*

- *Ist f n-mal differenzierbar, so ist das* n-te Taylorpolynom von f in p *definiert durch*

$$\mathrm{T}_p^n(f)\colon \mathbb{R} \to \mathbb{R}, \qquad x \mapsto \sum_{k=0}^{n} \frac{f^{(k)}(p)}{k!} \cdot (x-p)^k. \tag{301}$$

  Beachte: *Der Grad von $\mathrm{T}_p^n(f)$ ist immer $\le n$. Ist beispielsweise $f = 0$ die Nullfunktion, so ist der Grad von $\mathrm{T}_p^n(f)$ für jedes $n \in \mathbb{N}$ sogar gleich 0.*

- *Ist f glatt, so wird die Potenzreihe*

$$\mathrm{T}_p^\infty(f)(x) := \sum_{k=0}^{\infty} \frac{f^{(k)}(p)}{k!} \cdot (x-p)^k \tag{302}$$

  als die Taylorreihe von f in p *bezeichnet.* *($\mathrm{T}_p^\infty(f)(x) = \lim_n \mathrm{T}_p^n(x)$ im Existenzfall)*

Bemerkung 70. *Sei $D \subseteq \mathbb{R}$ nichtentartet, und $p \in D$ ein Punkt.*

a) • *Sei $f\colon \mathbb{R} \to \mathbb{R}$ eine Polynomfunktion vom Grad $n \in \mathbb{N}$ in $p \in \mathbb{R}$. Wegen (300) gilt $f = \mathrm{T}_p^n(f)$.*

[65] Wegen $f^{(n+\ell)}(p) = 0$ für $\ell \ge 1$, gilt natürlich auch $f(x) = \sum_{k=0}^{n+d} \frac{f^{(k)}(p)}{k!} \cdot (x-p)^k$ für alle $x \in \mathbb{R}$, für jedes $d \in \mathbb{N}_{>0}$. Hierbei ist dann aber zu beachten, dass der Koeffizient mit dem größten Index nicht ungleich Null ist, d.h. $n + d$ entspricht nicht dem Grad der Polynomfunktion f.

- *Sei $f := \mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]}$, für eine Potenzreihe $\mathsf{Q}(x) = \sum_{k=0}^{\infty} c_k \cdot (x-p)^k$ (in $\mathbb{R}$) in p mit Konvergenzradius $\mathsf{R}[\mathsf{Q}] > 0$. Dann gilt $\mathrm{T}_p^{\infty}(f) = \mathsf{Q}$.*

  Beweis. Wegen Korollar 43 ist $f = \mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]}$ glatt mit

  $$c_k = \frac{f^{(k)}(p)}{k!} \qquad \forall\, k \in \mathbb{N} \qquad \overset{(302)}{\Longrightarrow} \qquad \mathrm{T}_p^{\infty}(f) = \mathsf{Q}. \qquad \square$$

b) Sei $f\colon D \to \mathbb{R}$ eine n-mal differenzierbare Funktion mit $n \in \mathbb{N}$.

- *Ableiten liefert*[66]

  $$\mathrm{T}_p^n(f)^{(k)}(p) = f^{(k)}(p) \quad \forall\, 0 \le k \le n \qquad \textit{sowie} \qquad \mathrm{T}_p^n(f)^{(n+\ell)} = 0 \quad \forall\, \ell \in \mathbb{N}_{>0}. \tag{303}$$

  Es ist $\mathrm{T}_p^n(f)$ eine Polynomfunktion vom Grad $m \le n$; also gemäß Bemerkung 69.b) eindeutig bestimmt durch die ersten n (genauer die ersten m) Ableitungen in p. Wegen (303) ist somit $\mathrm{T}_p^n(f)$ die eindeutig bestimmte Polynomfunktion $g\colon \mathbb{R} \to \mathbb{R}$ vom Grad $\le n$, mit $g^{(k)}(p) = f^{(k)}(p)$ für alle $0 \le k \le n$.

- *Insbesondere folgt mit* (303) *(Übung):*

  $$\mathrm{T}_p^n\big(\mathrm{T}_p^{\ell}(f)\big) = \mathrm{T}_p^{\ell}(f) \qquad \forall\, 0 \le \ell \le n. \tag{304}$$

- *Für $n = 1$ ist*

  $$\mathrm{T}_p^1(f)(x) = f(p) + (x-p) \cdot f'(p) \qquad \forall\, x \in D$$

  diejenige affine Funktion, die wir bereits in der Charakterisierung von Differenzierbarkeit in Lemma 48.2) kennengelernt hatten.[67]

Übung 113. *Seien $f, g\colon D \to \mathbb{R}$ n-mal differenzierbar mit $n \in \mathbb{N} \cup \{\infty\}$. Zeigen Sie:*

$$\mathrm{T}_p^n(\lambda \cdot f + \mu \cdot g) = \lambda \cdot \mathrm{T}_p^n(f) + \mu \cdot \mathrm{T}_p^n(g) \qquad \forall\, \lambda, \mu \in \mathbb{R}.$$

Übung 114. *Seien $f, g\colon D \to \mathbb{R}$ n-mal differenzierbar mit $n \in \mathbb{N}$. Zeigen Sie die Leibnizformel:*

$$(f \cdot g)^{(n)}(p) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(p) \cdot g^{(n-k)}(p) \qquad \forall\, p \in D. \tag{305}$$

Hinweis: Induktion analog zum Beweis von Satz 6 (Binomischer Lehrsatz).

10.1.1 Der Satz von Taylor

Terminologie 28. *Sei $f\colon D \to \mathbb{R}$ eine n-mal differenzierbar Abbildung für $n \in \mathbb{N}$, sowie $p \in D$ ein Punkt. Die Abbildung*

$$\mathrm{R}_p^n(f)\colon D \to \mathbb{R}, \qquad x \mapsto f(x) - \mathrm{T}_p^n(f)(x) \tag{306}$$

heißt n-tes Restglied von f in p.

- *Per definitionem gilt $f = \mathrm{T}_p^n(f)|_D + \mathrm{R}_p^n(f)$.*
- *Ist f k-mal (stetig) differenzierbar, so auch $\mathrm{R}_p^n(f)$ (Übung 100). Zudem zeigt* (303)*:*

  $$\mathrm{R}_p^n(f)^{(k)}(p) = 0 \qquad \textit{für} \qquad k = 0, \dots, n. \tag{307}$$

[66] Es ist $\mathrm{T}_p^n(f)$ eine Polynomfunktion in p. Wegen (299) ist $\mathrm{T}_p^n(f)^{(k)}(p)$ nun gerade das $k!$-fache des k-ten Polynomkoeffizienten $\frac{f^{(k)}(p)}{k!}$ von $\mathrm{T}_p^n(f)^{(k)}(p)$, also gegeben durch $f^{(k)}(p)$.

[67] Also die affine Funktion, die sich in p am besten an f in dem Sinne anschmiegt, dass sie in p den gleichen Wert und die gleiche Ableitung besitzt. Ihr Graph ist die Tangente an den Graphen $\Gamma(f)$ durch den Punkt $(p, f(p))$.

Satz 64 (Taylorformel)**.** *Sei $f \in C^{n+1}(D,\mathbb{R})$ mit $n \in \mathbb{N}$. Für alle $x, p \in D$ gilt*

$$f(x) = \mathrm{T}_p^n(f)(x) + \mathrm{R}_p^n(f)(x) \qquad \textit{mit} \qquad \mathrm{R}_p^n(f)(x) = \frac{1}{n!} \cdot \int_p^x (x-t)^n \cdot f^{(n+1)}(t)\,\mathrm{d}t. \tag{308}$$

$\big($ Beachte: Für $n = 0$, erhalten wir aus obiger Formel in Übereinstimmung mit Satz 57 (Hauptsatz):

$$f(x) = \mathrm{T}_p^0(f)(x) + \mathrm{R}_p^0(f)(x) = f(p) + \textstyle\int_p^x f' \qquad \forall\, x \in D. \big)$$

Beweis. Die linke Seite von (308) ist klar. Es bleibt somit noch die Integraldarstellung des Restglieds auf der rechten Seite von (308) zu zeigen. Zunächst folgt mit (306) und (303)

$$\mathrm{R}_p^n(f)^{(n+1)} \overset{(306)}{=} f^{(n+1)} - \underbrace{\mathrm{T}_p^n(f)^{(n+1)}}_{\overset{(303)}{=}\,0} = f^{(n+1)}. \tag{309}$$

- Für $n = 0$, folgt mit dem Hauptsatz (zweiter Schritt) wegen $\mathrm{R}_p^0(f)(p) = 0$ (dritter Schritt)

$$\mathrm{R}_p^n(f)(x) = \mathrm{R}_p^0(f)(x) = \mathrm{R}_p^0(f)(p) + \int_p^x \mathrm{R}_p^0(f)'(t)\,\mathrm{d}t \overset{(309)}{=} \int_p^x f'(t)\,\mathrm{d}t = \frac{1}{n!} \cdot \int_p^x (x-t)^n \cdot f^{(n+1)}(t)\,\mathrm{d}t.$$

- Für $n \geq 1$, erhalten wir mit partieller Integration $\big(\int_a^b u \cdot v' = [u \cdot v]_a^b - \int_a^b u' \cdot v\big)$

$$\begin{aligned}\int_p^x (x-t)^n \cdot f^{(n+1)}(t)\,\mathrm{d}t &\overset{(309)}{=} \int_p^x (x-t)^n \cdot \mathrm{R}_p^n(f)^{(n+1)}(t)\,\mathrm{d}t \\ &= \underbrace{\big[(x-t)^n \cdot \mathrm{R}_p^n(f)^{(n)}(t)\big]_p^x}_{\overset{(*)}{=}\,0} + n \cdot \int_p^x (x-t)^{n-1} \cdot \mathrm{R}_p^n(f)^{(n)}(t)\,\mathrm{d}t.\end{aligned}$$

Die Gleichheit $(*)$ gilt wegen $(x-x)^n = 0$, sowie $\mathrm{R}_p^n(f)^{(n)}(p) = 0$ gemäß (307). Für $0 < \ell \leq n-1$, zeigt eine analoge Rechnung:

$$\int_p^x (x-t)^\ell \cdot \mathrm{R}_p^n(f)^{(\ell+1)}(t)\,\mathrm{d}t = \overbrace{\big[(x-t)^\ell \cdot \mathrm{R}_p^n(f)^{(\ell)}(t)\big]_p^x}^{=\,0} + \ell \cdot \int_p^x (x-t)^{\ell-1} \cdot \mathrm{R}_p^n(f)^{(\ell)}(t)\,\mathrm{d}t,$$

wegen $(x-x)^\ell = 0$, sowie $\mathrm{R}_p^n(f)^{(\ell)}(p) = 0$ gemäß (307). Induktiv folgt (erster Schritt)

$$\int_p^x (x-t)^n \cdot f^{(n+1)}(t)\,\mathrm{d}t = n! \cdot \int_p^x \mathrm{R}_p^n(f)'(t)\,\mathrm{d}t = n! \cdot \big(\mathrm{R}_p^n(f)(x) - \mathrm{R}_p^n(f)(p)\big) = n! \cdot \mathrm{R}_p^n(f)(x),$$

mit dem Hauptsatz im zweiten Schritt, sowie $\mathrm{R}_p^n(f)(p) = 0$ wegen (307) im dritten Schritt. □

Wir wollen nun noch zwei weitere Darstellungen der Restgliedes besprechen. Die einfachste Darstellung ist die folgende, welche für viele Abschätzungen sehr wichtig ist.

Satz 65 (Restglieddarstellung nach Lagrange)**.** *Sei $f \in C^{n+1}(D,\mathbb{R})$ mit $n \in \mathbb{N}$. Zu jedem $x \in D$ existiert ein ξ_x zwischen*[68] *x und p mit*

$$\mathrm{R}_p^n(f)(x) = \frac{(x-p)^{n+1}}{(n+1)!} \cdot f^{(n+1)}(\xi_x).$$

[68] D.h. $x \leq \xi_x \leq p$ für $x < p$, $p \leq \xi_x \leq$ für $p < x$, und $\xi_x = p$ für $x = p$.

Beweis. Für $x = p$ gilt die Aussage offensichtlich für $\xi_x := p$, da dann beide Seiten Null sind (siehe (307) mit $\ell = 0$). Sei nun $p < x$. Nach Satz 56 (Mittelwertsatz der Integralrechnung), existiert ein $\xi_x \in [p, x]$ mit (zweiter Schritt)

$$\begin{aligned} \mathrm{R}_p^n(f)(x) &\overset{\text{Satz 64}}{=} \frac{1}{n!} \cdot \int_p^x \overbrace{(x-t)^n}^{\geq 0} \cdot f^{(n+1)}(t)\,\mathrm{d}t \quad = \quad \frac{1}{n!} \cdot f^{(n+1)}(\xi_x) \cdot \int_p^x (x-t)^n\,\mathrm{d}t \\ &= \frac{1}{n!} \cdot f^{(n+1)}(\xi_x) \cdot \left[-\frac{(x-\cdot)^{n+1}}{n+1} \right]_p^x = f^{(n+1)}(\xi_x) \cdot \frac{(x-p)^{n+1}}{(n+1)!}. \end{aligned}$$

Für $x < p$ ist $(t - x)^n \geq 0$ für $x \leq t \leq p$, und wir erhalten analog für ein $\xi_x \in [x, p]$ (Satz 56 im dritten Schritt)

$$\begin{aligned} \mathrm{R}_p^n(f)(x) &= \frac{1}{n!} \cdot \int_p^x (x-t)^n \cdot f^{(n+1)}(t)\,\mathrm{d}t \\ &= (-1)^{n+1} \cdot \frac{1}{n!} \cdot \int_x^p \overbrace{(t-x)^n}^{\geq 0} \cdot f^{(n+1)}(t)\,\mathrm{d}t = (-1)^{n+1} \cdot \frac{1}{n!} \cdot f^{(n+1)}(\xi_x) \cdot \overbrace{\int_x^p (t-x)^n\,\mathrm{d}t}^{\left[\frac{(\cdot - x)^{n+1}}{n+1}\right]_x^p} \\ &= (-1)^{n+1} \cdot f^{(n+1)}(\xi_x) \cdot \frac{(p-x)^{n+1}}{(n+1)!} = f^{(n+1)}(\xi_x) \cdot \frac{(x-p)^{n+1}}{(n+1)!}. \end{aligned}$$

□

Korollar 44. *Sei $f \in C^n(D, \mathbb{R})$ mit $n \in \mathbb{N}$. Für jedes $p \in D$ gilt*

$$\lim_{x \to p} \frac{\mathrm{R}_p^n(f)(x)}{(x-p)^n} = 0.$$

Beweis.
- Für $n = 0$ gilt $\mathrm{R}_p^n(f)(x) = f(x) - f(p)$ für $x \in D$, also wegen der Stetigkeit von f in p:

$$\lim_{x \to p} \frac{\mathrm{R}_p^n(f)(x)}{(x-p)^n} = \lim_{x \to p} f(x) - f(p) = 0.$$

- Sei $n \geq 1$. Per definitionem gilt (beachte $C^{n-1}(D, \mathbb{R}) \subseteq C^n(D, \mathbb{R})$)

$$\mathrm{T}_p^n(f)(x) + \mathrm{R}_p^n(f)(x) = f(x) = \mathrm{T}_p^{n-1}(f)(x) + \mathrm{R}_p^{n-1}(f)(x)$$

für $x \in D$, und umstellen führt auf

$$\mathrm{R}_p^n(f)(x) = \mathrm{R}_p^{n-1}(f)(x) + \big(\mathrm{T}_p^{n-1}(f)(x) - \mathrm{T}_p^n(f)(x)\big) = \mathrm{R}_p^{n-1}(f)(x) - \frac{f^{(n)}(p)}{n!} \cdot (x-p)^n.$$

Für $p \neq x \in D$, liefert nun Satz 65 ein $\xi_x \in D$ mit[69]

$$|p - \xi_x| \leq |p - x| \qquad \text{sowie} \qquad \mathrm{R}_p^{n-1}(f)(x) = \frac{(x-p)^n}{n!} \cdot f^{(n)}(\xi_x),$$

also

$$\frac{\mathrm{R}_p^n(f)(x)}{(x-p)^n} = \frac{1}{(x-p)^n} \cdot \left(\frac{(x-p)^n}{n!} \cdot f^{(n)}(\xi_x) - \frac{f^{(n)}(p)}{n!} \cdot (x-p)^n \right) = \frac{f^{(n)}(\xi_x) - f^{(n)}(p)}{n!}.$$

Für jede Folge $(x_\ell)_{\ell \in \mathsf{N}}$ in $D \setminus \{p\}$ mit $\lim_\ell x_\ell = p$ gilt nun ebenfalls $\lim_\ell \xi_{x_\ell} = p$, also

$$\lim_\ell \frac{\mathrm{R}_p^n(f)(x_\ell)}{(x_\ell - p)^n} = \lim_\ell \left(\frac{f^{(n)}(\xi_{x_\ell}) - f^{(n)}(p)}{n!} \right) = \left(\frac{f^{(n)}(p) - f^{(n)}(p)}{n!} \right) = 0$$

wegen der Stetigkeit (Folgenstetigkeit) von $f^{(n)}$. Da die Folge $(x_\ell)_{\ell \in \mathsf{N}}$ beliebig war, folgt hiermit die Behauptung (Proposition 11). □

[69] Anders formuliert gilt $\xi_x \in D \cap (p - |x-p|, p + |x-p|)$.

Bemerkung 71. *Sei $f\colon D \to \mathbb{R}$ eine Abbildung (D nichtentartet), und $p \in D$ ein Punkt.*

a) Sei $f \in C^n(D, \mathbb{R})$ mit $n \in \mathbb{N}$. Wegen Korollar 44 und Proposition 10, ist dann die Abbildung

$$\Theta_n\colon D \to \mathbb{R}, \qquad x \mapsto \begin{cases} \dfrac{\mathrm{R}_p^n(f)(x)}{(x-p)^n} & \textit{für} \quad x \neq p \\ 0 & \textit{für} \quad x = p. \end{cases}$$

stetig in p (sogar stetig wegen Proposition 10.2), bzw. Proposition 10.1) und Übung 79.a)). Wir erhalten hiermit (beachte $\mathrm{R}_p^n(p) = 0$)

$$f(x) = \mathrm{T}_p^n(f)(x) + \overbrace{(x-p)^n \cdot \Theta_n(x)}^{=\mathrm{R}_p^n(x)} \qquad \forall\, x \in D. \tag{310}$$

Für $n = 1$ und $h \in D - p$, erhalten wir $f(p+h) = f(p) + f'(p) \cdot h + h \cdot \Theta_1(p+h)$, also die Charakterisierung von Differenzierbarkeit von f in p, die wir bereits in Lemma 48.2) kennengelernt hatten (setze dort $\varepsilon\colon h \mapsto \Theta_1(p+h)$).

b) Sei $f \in C^{n+1}(D, \mathbb{R})$ mit $n \in \mathbb{N}$. Wegen Satz 65 und der Stetigkeit von $f^{(n+1)}$ in p, gilt

$$\lim_{x\to p} \frac{\mathrm{R}_p^n(f)(x)}{(x-p)^{n+1}} = \lim_{x\to p} \frac{f^{(n+1)}(\xi_x)}{(n+1)!} = \frac{f^{(n+1)}(p)}{(n+1)!}.$$

Wegen Proposition 10 ist dann die Abbildung

$$\Psi_n\colon D \to \mathbb{R}, \qquad x \mapsto \begin{cases} \dfrac{\mathrm{R}_p^n(f)(x)}{(x-p)^{n+1}} & \textit{für} \quad x \neq p \\ \dfrac{f^{(n+1)}(p)}{(n+1)!} & \textit{für} \quad x = p, \end{cases}$$

stetig in p (sogar stetig wegen Proposition 10.2), bzw. Proposition 10.1) und Übung 79.a)). Wir erhalten hiermit (beachte $\mathrm{R}_p^n(p) = 0$)

$$f(x) = \mathrm{T}_p^n(f)(x) + \overbrace{(x-p)^{n+1} \cdot \Psi_n(x)}^{=\mathrm{R}_p^n(x)} \qquad \forall\, x \in D. \tag{311}$$

Für $n = 0$ und $h \in D - p$, erhalten wir $f(p+h) = f(p) + h \cdot \Psi_0(p+h)$ mit $\Psi_0(p) = f'(p)$, also die Charakterisierung von Differenzierbarkeit von f in p, die wir bereits in Lemma 48.3) kennengelernt hatten (setze dort $\Phi\colon h \mapsto \Psi_0(p+h)$).

Beispiel 82. *Die Taylorentwicklung kann insbesondere dazu verwendet werden, Grenzwerte effizient auszurechnen. Gesucht sei der Grenzwert*

$$\lim_{x\to 0} \frac{1-\cos(x)}{x^2}.$$

Die Funktion $f := 1 - \cos\colon \mathbb{R} \to \mathbb{R}$ ist glatt. Dann gilt ($f(0) = 0 = \sin(0) = f'(0)$, $f''(0) = \cos(0) = 1$)

$$\mathrm{T}_0^1(f)(x) = 0 \qquad \textit{sowie} \qquad \mathrm{T}_0^2(f)(x) = \tfrac{x^2}{2} \qquad \textit{für alle} \qquad x \in \mathbb{R}.$$

Für $n = 2$ bzw. $n = 1$, liefern (310) bzw. (311), die Darstellungen

$$1 - \cos(x) = f(x) = \tfrac{1}{2} \cdot x^2 + x^2 \cdot \Theta_2(x) \quad \textit{bzw.} \quad 1 - \cos(x) = f(x) = x^2 \cdot \Psi_1(x) \quad \textit{für alle} \quad x \in \mathbb{R},$$

mit (in 0) stetigen Abbildungen $\Theta_2, \Psi_1\colon \mathbb{R} \to \mathbb{R}$ mit $\Theta_2(0) = 0$ bzw. $\Psi_1(0) = \frac{f''(0)}{2!} = \frac{1}{2}$. Es folgt

$$\lim_{x\to 0} \frac{1-\cos(x)}{x^2} = \frac{1}{2} + \lim_{x\to 0} \Theta_2(x) = \frac{1}{2} + \Theta_2(0) = \mathbf{\frac{1}{2}} = \Psi_1(0) = \lim_{x\to 0} \Psi_1(x) = \lim_{x\to 0} \frac{1-\cos(x)}{x^2}.$$

Übung 115. *Sei $f\colon D \to \mathbb{R}$ n-mal differenzierbar mit $n \in \mathbb{N}_{>0}$, sowie $\tilde{D} \subseteq \mathbb{R}$ nichtentartet mit $\tilde{D} \subseteq D$. Zeigen Sie, dass $\tilde{f} := f|_{\tilde{D}}$ ebenfalls n-mal differenzierbar ist, mit*

$$\tilde{f}^{(k)} = f^{(k)}|_{\tilde{D}} \qquad \forall\, 0 \leq k \leq n.$$

Beachte: *Ist $p \in D$ ein Punkt und $I \subseteq \mathbb{R}$ ein offenes Intervall mit $p \in I$, so gilt obige Aussage wegen Übung 91.a) insbesondere für $\tilde{D} = I \cap D$.*

Hinweis: Induktion mit Hilfe von Korollar 31.

Der nächste Satz ist eine Verschärfung der Restglieddarstellung von Lagrange; denn es wird $f^{(n+1)}$ nicht als stetig vorausgesetzt, und der ξ_x entsprechende Wert liegt im offenen Intervall $(p, p+h)$:

Satz 66 (Verschärfte Restglieddarstellung von Lagrange)**.** *Sei $p \in \mathbb{R}$ und $h > 0$, sowie $f\colon [p, p+h] \to \mathbb{R}$ eine $(n+1)$-mal differenzierbare Abbildung. Dann existiert ein $\theta_h \in (0,1)$, mit*

$$f(p+h) = \mathrm{T}_p^n(f)(h) + \frac{f^{(n+1)}(p + \theta_h \cdot h)}{(n+1)!} \cdot h^{n+1}.$$

Bemerkung:

- Satz 66 gilt in entsprechender Form (mit analogem Beweis) auch für eine $(n+1)$-mal differenzierbare Abbildung $f\colon [p-h, p] \to \mathbb{R}$ mit $p \in \mathbb{R}$ und $h > 0$.
- Ist $f\colon D \to \mathbb{R}$ eine $(n+1)$-mal differenzierbare Abbildung mit $n \in \mathbb{N}$, sowie $p \in D$ und $h > 0$ mit $[p, p+h] \subseteq D$, so lässt sich Satz 66 wegen Übung 115 auf die Einschränkung $f|_{[p,p+h]}$ anwenden.

Beweis. Es gilt $f \in C^n([p, p+h], \mathbb{R})$ wegen Lemma 47, denn f ist $(n+1)$-mal differenzierbar. Dann sind auch die folgenden beiden Abbildungen n-mal stetig differenzierbar:

$$\begin{aligned} \mathrm{R}\colon [0,h] \to \mathbb{R}, &\qquad y \mapsto \mathrm{R}_p^n(f)(p+y) = f(p+y) - \mathrm{T}_p^n(f)(p+y) \\ g\colon [0,h] \to \mathbb{R}, &\qquad y \mapsto y^{n+1}. \end{aligned}$$

Es gilt $\mathrm{R}^{(n+1)} = f^{(n+1)}(p + \cdot)$ wegen $\mathrm{T}_p^n(f)^{(n+1)} = 0$, sowie $g^{(n+1)} = (n+1)!$. Weiterhin gilt

$$\mathrm{R}^{(\ell)}(0) = 0 = g^{(\ell)}(0) \qquad \text{für} \qquad 0 \leq \ell \leq n \qquad \text{wobei} \qquad g^{(\ell)}|_{(0,h]} \qquad \text{keine Nullstelle hat.}$$

Iteratives Anwenden des allgemeinen Mittelwertsatzes (genauer (253) in Satz 46), auf Einschränkungen (beachte Übung 115) obiger Ableitungen auf geeignete Intervalle $[0, \mu]$ für gewisse $0 \leq \mu < h$, liefert Faktoren $\theta_1, \dots, \theta_{n+1} \in (0,1)$ mit

$$\begin{aligned} \frac{f(p+h) - \mathrm{T}_p^n(f)(p+h)}{h^{n+1}} = \frac{\mathrm{R}(h)}{g(h)} = \frac{\mathrm{R}(h) - \mathrm{R}(0)}{g(h) - g(0)} &= \frac{\mathrm{R}^{(1)}(\theta_1 \cdot h)}{g^{(1)}(\theta_1 \cdot h)} \qquad = \frac{\mathrm{R}^{(1)}(\theta_1 \cdot h) - \mathrm{R}^{(1)}(0)}{g^{(1)}(\theta_1 \cdot h) - g^{(1)}(0)} \\ &= \frac{\mathrm{R}^{(2)}(\theta_2 \cdot \theta_1 \cdot h)}{g^{(2)}(\theta_2 \cdot \theta_1 \cdot h)} = \frac{\mathrm{R}^{(2)}(\theta_2 \cdot \theta_1 \cdot h) - \mathrm{R}^{(2)}(0)}{g^{(2)}(\theta_2 \cdot \theta_1 \cdot h) - g^{(2)}(0)} \\ &= \dots \\ &= \frac{\mathrm{R}^{(n+1)}(\overbrace{\theta_{n+1} \cdot \ldots \cdot \theta_1}^{=:\,\theta_h} \cdot h)}{(n+1)!} = \frac{f^{(n+1)}(p + \theta_h \cdot h)}{(n+1)!}. \end{aligned}$$ □

Übung 116. *Sei $f\colon D \to \mathbb{R}$ n-mal differenzierbar, mit $n \in \mathbb{N}_{>0}$. Zeigen Sie:*

$$\mathrm{T}_p^{n-k}\big(f^{(k)}\big) = \mathrm{T}_p^n(f)^{(k)} \qquad \textit{sowie} \qquad \mathrm{R}_p^{n-k}\big(f^{(k)}\big) = \mathrm{R}_p^n(f)^{(k)} \qquad \textit{für alle} \qquad 0 \leq k \leq n.$$

10.1.2 Taylorreihen

Bemerkung 72. *Das Konvergenzverhalten von Taylorreihen ist in der Regel* **nicht** *besonders gut.*

a) *Jede Potenzreihe (ungeachtet des Konvergenzradius) tritt als Taylorreihe einer geeigneten glatten Funktion auf. Genauer gilt (siehe bspw. [1, Satz 4.5]):*

   Satz (Peano-Borel): *Für jede Folge reeller Zahlen* $(c_n)_{n\in\mathbb{N}}$ *existiert eine Funktion* $f \in C^\infty(\mathbb{R},\mathbb{R})$ *mit* $f^{(n)}(0) = n! \cdot c_n$ *für alle* $n \in \mathbb{N}$*, d.h.* $\mathrm{T}_0^\infty(f)(x) = \sum_{n=0}^\infty c_n \cdot x^n$.

b) *Auch wenn zu einem* $f \in C^\infty(D,\mathbb{R})$ *und* $p \in D$ *die Taylorreihe* $\mathrm{T}_p^\infty(f)$ *Konvergenzradius* $\mathsf{R} > 0$ *hat, so muss sie nicht notwendigerweise* f *darstellen. Konkreter kann sogar* $\mathrm{T}_p^\infty(f)(x) \neq f(x)$ *für alle* $p \neq x \in D \cap (p+\mathsf{R}, p-\mathsf{R})$ *gelten:*

   Man betrachte beispielsweise die Taylorreihe $\mathrm{T}_0^\infty(f)$ *der Funktion*

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad x \mapsto \begin{cases} \mathrm{e}^{-\frac{1}{x}} & \text{für } \ x > 0 \\ 0 & \text{für } \ x = 0 \\ \mathrm{e}^{\frac{1}{x}} & \text{für } \ x < 0, \end{cases} \tag{312}$$

   für die sich nachweisen lässt, dass $f \in C^\infty(\mathbb{R},\mathbb{R})$ *mit* $f^{(n)}(0) = 0$ *für alle* $n \in \mathbb{N}$ *gilt (Übung 117). In diesem Fall gilt* $\mathrm{T}_0^\infty(f) = 0$ *sowie* $f|_{\mathbb{R}\setminus\{0\}} > 0$*, d.h.* $\mathrm{T}_0^\infty(f)(x) = 0 \neq f(x)$ *für alle* $x \in \mathbb{R}\setminus\{0\}$.

Übung 117. *Zeigen Sie, dass die Funktion* (312) *glatt ist, mit* $f^{(n)}(0) = 0$ *für alle* $n \in \mathbb{N}$.

Lemma 62. *Sei* $f\colon D \to \mathbb{R}$ *eine* n*-mal differenzierbare Funktion, mit* $n \in \mathbb{N} \cup \{\infty\}$*. Sei zudem* $\tilde{D} \subseteq \mathbb{R}$ *nichtentartet mit* $\tilde{D} \subseteq D$*. Dann gilt*

$$\mathrm{T}_p^n(f|_{\tilde{D}}) = \mathrm{T}_p^n(f) \qquad \forall\, p \in \tilde{D}.$$

Ist insbesondere $I \subseteq \mathbb{R}$ *ein offenes Intervall mit* $I \cap D \neq \emptyset$*, so gilt*

$$\mathrm{T}_p^n(f|_{D\cap I}) = \mathrm{T}_p^n(f) \qquad \forall\, p \in I \cap D.$$

Beweis. Wegen Übung 115 gilt $(f|_{\tilde{D}})^{(k)} = f^{(k)}|_{\tilde{D}}$ für alle $k \in \mathbb{N}$; womit die erste Behauptung unmittelbar folgt. Die zweite Behauptung folgt nun sofort aus Übung 91.a). □

Notation 29. *Für* $D \subseteq \mathbb{R}$ *nichtentartet,* $p \in D$ *und* $\varepsilon > 0$*, notieren wir vereinfachend*

$$D[p,\varepsilon] := D \cap (p-\varepsilon, p+\varepsilon).$$

Gemäß Übung 91.a) ist $D[p,\varepsilon]$ *nichtentartet mit* $D[p,\varepsilon] \subseteq D$.

Trotz des schlechten Konvergenzverhaltens von Taylorreihen im Allgemeinen, gibt es dennoch Kriterien, die deren Konvergenz sicherstellen. Der nächste Satz zeigt beispielsweise, dass glatte Funktionen, die lokal durch konvergente Potenzreihen dargestellt werden können, lokal mit ihrer Taylorreihe übereinstimmen.

Satz 67. *Sei* $f \in C^\infty(D,\mathbb{R})$*, sowie* $\mathsf{Q}(x) = \sum_{n=0}^\infty c_n \cdot (x-p)^n$ *eine Potenzreihe in* $p \in D$ *mit Konvergenzradius* $\mathsf{R}[\mathsf{Q}] > 0$*. Es sei zudem* $0 < \varepsilon \leq \mathsf{R}[\mathsf{Q}]$ *mit*

$$f|_{D[p,\varepsilon]} = \mathsf{Q}|_{D[p,\varepsilon]}. \tag{313}$$

Dann gilt $\mathrm{T}_p^\infty(f) = \mathsf{Q}$*, also* $c_n = \frac{f^{(n)}(p)}{n!}$ *für alle* $n \in \mathbb{N}$.

Beweis. Gemäß Bemerkung 70.a), gilt $\mathrm{T}_p^\infty(\mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]}) = \mathsf{Q}$. Wir erhalten mit Lemma 62 im zweiten und im letzten Schritt (es ist $\mathrm{K}[\mathsf{Q}] \supseteq D[p,\varepsilon] \subseteq D$ nichtentartet):

$$\mathsf{Q} = \mathrm{T}_p^\infty\big(\mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]}\big) = \mathrm{T}_p^\infty\big((\mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]})|_{D[p,\varepsilon]}\big) = \mathrm{T}_p^\infty\big(\mathsf{Q}|_{D[p,\varepsilon]}\big) \overset{(313)}{=} \mathrm{T}_p^\infty\big(f|_{D[p,\varepsilon]}\big) = \mathrm{T}_p^\infty(f). \qquad \square$$

Beispiel 83. *Sei* $\mathsf{Q}(x) = \sum_{n=0}^\infty \frac{x^n}{n!}$, *d.h.* $p = 0$, $\mathsf{R}[\mathsf{Q}] = \infty$, $f := \exp_{\mathbb{R}} = \mathsf{Q}$. *Wir erhalten wegen* $\exp'_{\mathbb{R}} = \exp_{\mathbb{R}}$ *(vgl.* (244)*), in Übereinstimmung mit Satz 67:*

$$f^{(n)}(0) \overset{(244)}{=} \exp_{\mathbb{R}}(0) = 1 \qquad \forall\, n \in \mathbb{N} \qquad \Longrightarrow \qquad \mathrm{T}_0^\infty(f)(x) = \textstyle\sum_{n=0}^\infty \frac{x^n}{n!}.$$

Wir haben das folgende hinreichende (nicht notwendige) Konvergenzkriterium für Taylorreihen.

Satz 68 (Konvergenzkriterium)**.** *Sei* $f \in C^\infty(D,\mathbb{R})$, *sowie* $M \in [0,\infty)$ *und* $N \in \mathbb{N}$ *mit*

$$|f^{(n)}(x)| \leq M^n \qquad \forall\, x \in D,\ n \geq N. \tag{314}$$

Dann gilt $f = \mathrm{T}_p^\infty(f)|_D$ *für jedes* $p \in D$.

Beachte: Für $p \in D$ gilt dann insbesondere

$$D \subseteq \Theta[\mathrm{T}_p^\infty(f)] \qquad \Longrightarrow \qquad \mathsf{R}[\mathrm{T}_p^\infty(f)] \geq \sup_{\overline{\mathbb{R}}}(D - p).$$

Beweis. Seien $p, x \in D$ vorgegeben. Wir zeigen, dass dann $\lim_n \mathrm{T}_p^n(f)(x) = f(x)$ gilt. Dann gilt automatisch

$$\mathrm{T}_p^\infty(f)(x) = \lim_n \textstyle\sum_{k=0}^n \frac{f^{(k)}(p)}{k!} \cdot (x-p)^k = \lim_n \mathrm{T}_p^n(f)(x) = f(x).$$

Für $n \geq N$, folgt nun mit der Restglieddarstellung von Lagrange (Satz 65) im zweiten Schritt

$$0 \leq \lambda_n := \big|f(x) - \mathrm{T}_p^n(f)(x)\big| = |\mathrm{R}_p^n(f)(x)| = \frac{|x-p|^{n+1}}{(n+1)!} \cdot \big|f^{(n+1)}(\xi[n]_x)\big| \overset{(314)}{\leq} \frac{(M \cdot |x-p|)^{n+1}}{(n+1)!} =: \kappa_n$$

($\xi[n]_x \in D$ zwischen x und p). Da die Reihe $\exp(M \cdot |x-p|)$ konvergiert, gilt $\lim_n \kappa_n = 0$ (Lemma 37). Lemma 31 (Quetschlemma) zeigt somit $\lim_n \lambda_n = 0$, also $\lim_n \mathrm{T}_p^n(f)(x) = f(x)$. $\square$

Übung 118. *Sei* $f \in C^\infty(D,\mathbb{R})$ *sowie* $M > 0$ *und* $N \in \mathbb{N}$ *mit*

$$|f^{(n)}(x)| \leq M^n \cdot n! \qquad \forall\, x \in D,\ n \geq N.$$

Dann gilt $f|_{D[p,\varepsilon]} = \mathrm{T}_p^\infty(f)|_{D[p,\varepsilon]}$ *für* $\varepsilon := M^{-1}$.

Beispiel 84. *a) Sei* $f = \cos\colon \mathbb{R} \to [-1,1]$ *die Kosinusfunktion. Mit Satz 49 erhalten wir*

$$f^{[4n]}(x) = \cos(x), \quad f^{[4n+1]}(x) = -\sin(x), \quad f^{[4n+2]}(x) = -\cos(x), \quad f^{[4n+3]}(x) = \sin(x) \tag{315}$$

für alle $n \in \mathbb{N}$. *Für* $M = 1$ *und* $N = 0$ *gilt dann*

$$|f^{(n)}(x)| \leq M^n \qquad \forall\, x \in \mathbb{R},\ n \geq N,$$

also (314)*. Satz 68 liefert somit für* $p = 0$ *(erster Schritt):* $\qquad (\cos(0) = 1 \ \wedge\ \sin(0) = 0\,)$

$$\cos(x) = \mathrm{T}_0^\infty(\cos)(x) = \sum_{n=0}^\infty \frac{\cos^{(n)}(0)}{n!} \cdot x^n = \sum_{n=0}^\infty \frac{\cos^{(2n)}(0)}{(2n)!} \cdot x^{2n} = \sum_{n=0}^\infty \frac{(-1)^n}{(2n)!} \cdot x^{2n}.$$

Im Einklang mit Satz 67, reproduziert diese Diskussion unsere Potenzreihendarstellung aus Satz 50, die wir unmittelbar aus der Definition der Cosinusfunktion erhielten.

(Eine analoge Rechnung reproduziert natürlich auch die entsprechende Reihenentwicklung aus Satz 50 für die Sinusfunktion.)

b) $*$ *In Satz 68 ist es wichtig, dass die Konstante M nicht vom Index n abhängt. Sei beispielsweise $D = [-2, 2]$ und $f\colon D \ni x \mapsto \frac{1}{1+x^2} \in \mathbb{R}$. In Beispiel 81.a) hatten wir mit Hilfe der geometrischen Reihe eingesehen, dass*

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n \cdot x^{2n} =: \mathsf{P}(x) \qquad \forall\, x \in (-1, 1)$$

gilt mit $\mathsf{R}[\mathsf{P}] = 1$. Satz 67 zeigt nun, dass $\mathrm{T}_0^\infty(f) = \mathsf{P}$ gilt, womit $\mathrm{T}_0^\infty(f)$ nicht auf ganz $[-2, 2]$ konvergieren (also dort auch nicht f darstellen) kann. Wegen $f \in C^\infty(D, \mathbb{R})$ sind nun aber alle Ableitungen von f stetig, also gilt $M_n := |f^{(n)}|_\infty < \infty$ für alle $n \in \mathbb{N}$ (wegen Satz 35 – Satz vom Maximum). Satz 68 zeigt somit, dass für jedes $M \geq 1$ und jedes $N \in \mathbb{N}$, ein $n \geq N$ mit $M_n \geq M^n \geq M$ existiert, dass also die Folge $(M_n)_{n\in\mathbb{N}}$ insbesondere unbeschränkt ist.[70]

Satz 69 (Binomische Reihe). *Für $x \in (-1, 1)$ und $\alpha \in \mathbb{R}$ gilt* *(beachte $\binom{\alpha}{0} = 1$ auch für $\alpha = 0$)*

$$(1+x)^\alpha = \underbrace{\sum_{n=0}^{\infty} \binom{\alpha}{n} \cdot x^n}_{=:\ \mathsf{P}_\alpha(x)} \quad \text{mit} \quad \binom{\alpha}{n} := \frac{\alpha \cdot (\alpha-1) \cdot \ldots \cdot (\alpha - n + 1)}{n!} \quad \text{für} \quad n \geq 1, \quad \binom{\alpha}{0} := 1.$$

- P_α *ist die Taylorreihe von $(1 + \cdot\,)^\alpha\colon (-1, \infty) \to \mathbb{R}$ in 0, wegen Satz 67.*
- *Für $\alpha \in \mathbb{N}$ gilt die Behauptung wegen dem binomischen Lehrsatz für jedes $x \in \mathbb{R}$; denn dann gilt $\binom{\alpha}{n} = 0$ für alle $n > \alpha$ $\big(\binom{\alpha}{n}$ ist der übliche Binomialkoeffizient$\big)$.*
- *Für $\alpha \in \mathbb{R} \setminus \mathbb{N}$, gilt $\mathsf{R}[\mathsf{P}_\alpha] = 1$.*

Beweis. Sei $\alpha \in \mathbb{R} \setminus \mathbb{N}$; und für $x \in \mathbb{R}_\times$ sei $a[x]_n := \binom{\alpha}{n} \cdot x^n \neq 0$ für alle $n \in \mathbb{N}$. Dann gilt[71]

$$\lim_n \left| \frac{a[x]_{n+1}}{a[x]_n} \right| = \lim_n \left| \frac{\alpha - n}{n+1} \right| \cdot |x| = |x| \qquad \Longrightarrow \qquad \overline{\lim}_n \left| \frac{a[x]_{n+1}}{a[x]_n} \right| = |x| = \underline{\lim}_n \left| \frac{a[x]_{n+1}}{a[x]_n} \right|$$

Daher konvergiert $\mathsf{P}(x)$ nach dem Quotientenkriterium (Korollar 24) für alle $x \in (-1, 1)$, und divergiert für alle $x \in \mathbb{R} \setminus [-1, 1]$. Dies zeigt $(-1, 1) \subseteq \Theta[\mathsf{P}_\alpha] \subseteq [-1, 1]$, also $\mathsf{R}[\mathsf{P}_\alpha] = 1$. Wir definieren

$$\begin{aligned} &f\colon (-1,1) \to \mathbb{R}, && x \mapsto \textstyle\sum_{n=0}^\infty \binom{\alpha}{n}\, x^n \\ &g\colon (-1,1) \to (0, \infty), && x \mapsto (1+x)^\alpha = \mathrm{e}^{\alpha \cdot \log(1+x)}. \end{aligned}$$

- g ist differenzierbar (Beispiel 69.3)), mit $\mathrm{im}(g) \subseteq (0, \infty)$ und $g(0) = 1$. $(\mathrm{e}^{\alpha \cdot \log(1)} = \mathrm{e}^0 = 1)$
- f ist glatt (Bemerkung 67) mit $f(0) = \binom{\alpha}{0} = 1$, und es gilt (Nachweis siehe unten)

$$(1 + x) \cdot f'(x) = \alpha \cdot f(x) \qquad \forall\, x \in (-1, 1). \tag{316}$$

Dann ist $h := \frac{f}{g}\colon (-1, 1) \to \mathbb{R}$ definiert mit $h(0) = 1$; sowie differenzierbar nach Satz 40 (Quotientenregel), mit

$$h'(x) = \frac{f'(x) \cdot (1+x)^\alpha - f(x) \cdot \alpha \cdot (1+x)^{\alpha-1}}{((1+x)^\alpha)^2} \overset{(316)}{=} \frac{\alpha \cdot f(x) \cdot (1+x)^{\alpha-1} - \alpha \cdot f(x) \cdot (1+x)^{\alpha-1}}{((1+x)^\alpha)^2} = 0.$$

Gemäß Korollar 32.2) ist daher h konstant $h(0) = 1$; also gilt $f(x) = g(x)$ für alle $x \in (-1, 1)$.

[70] Selbiges gilt für jede Einschränkung Funktion (312), auf ein Intervall $[-R, R]$ für $R > 0$.

[71] Beachte $\lim_n \frac{\alpha}{n+1} = 0$, sowie $\lim_n \frac{n}{n+1} = 1$ wegen $\lim_n \frac{n+1}{n} = 1$.

Wir erhalten (316) durch gliedweises differenzieren (Bemerkung 67):

$$\begin{aligned}
(1+x)\cdot f'(x) &= (1+x)\cdot \sum_{n=1}^{\infty}\binom{\alpha}{n}\, n\cdot x^{n-1}\\
&= (1+x)\cdot \sum_{n=1}^{\infty}\frac{\alpha\cdot(\alpha-1)\cdot\ldots\cdot(\alpha-n+1)}{(n-1)!}\cdot x^{n-1}\\
&= \alpha\cdot(1+x)\cdot\sum_{n=1}^{\infty}\binom{\alpha-1}{n-1}\cdot x^{n-1}\\
&= \alpha\cdot\left(\sum_{n=0}^{\infty}\binom{\alpha-1}{n}\cdot x^n+\sum_{n=1}^{\infty}\binom{\alpha-1}{n-1}\cdot x^n\right)\\
&= \alpha\cdot\left(1+\sum_{n=1}^{\infty}\left(\binom{\alpha-1}{n}+\binom{\alpha-1}{n-1}\right)\cdot x^n\right)\\
&\overset{(*)}{=} \alpha\cdot\left(1+\sum_{n=1}^{\infty}\binom{\alpha}{n}\cdot x^n\right)=\alpha\cdot f(x).
\end{aligned}$$

Im letzten Schritt $(*)$ haben wir das Additionstheorem für Binomialkoeffizienten benutzt: $(n\geq 1)$

$$\begin{aligned}
\binom{\alpha-1}{n}+\binom{\alpha-1}{n-1} &= \frac{(\alpha-1)\cdot\ldots\cdot\overbrace{((\alpha-1)-n+1)}^{\alpha-n}}{n!}\cdot\left(1+\frac{n}{\alpha-n}\right)\\
&= \frac{(\alpha-1)\cdot\ldots\cdot(\alpha-n+1)\cdot(\alpha-n)}{n!}\cdot\frac{\alpha}{\alpha-n}=\binom{\alpha}{n}.
\end{aligned}$$

(Beachte $\binom{\alpha-1}{n-1}=\frac{(\alpha-1)\cdot\ldots\cdot(\alpha-n+1)}{(n-1)!}$ wegen $(\alpha-1)-(n-1)+1=\alpha-n+1$.) $\square$

Übung 119 (Binomische Reihe)**.** *Sei $x\in(-1,1)$ vorgegeben. Zeigen Sie (jeweils dritter Schritt):*

$$\frac{1}{1+x}=(1+x)^{-1}=\sum_{n=0}^{\infty}\binom{-1}{n}\cdot x^n=\sum_{n=0}^{\infty}(-1)^n\cdot x^n=1-x+x^2-x^3+\ldots \tag{a}$$

$$\frac{1}{(1+x)^2}=(1+x)^{-2}=\sum_{n=0}^{\infty}\binom{-2}{n}\cdot x^n=\sum_{n=0}^{\infty}(-1)^n\cdot(n+1)\cdot x^n=1-2x+3x^3-4x^4+\ldots \tag{b}$$

$$\sqrt{1+x}=(1+x)^{\frac{1}{2}}=\sum_{n=0}^{\infty}\binom{\frac{1}{2}}{n}\cdot x^n=1+\sum_{n=1}^{\infty}(-1)^{n-1}\cdot\frac{(2n-3)\cdot(2n-5)\cdot\ldots\cdot 3\cdot 1}{\underbrace{(2n)\cdot(2n-2)\cdot(2n-4)\cdot\ldots\cdot 2}_{2^n\cdot n!}}\cdot x^n \tag{c}$$

$$=1+\frac{1}{2}\cdot x-\frac{1}{2\cdot 4}\cdot x^2+\frac{1\cdot 3}{2\cdot 4\cdot 6}\cdot x^3-\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6\cdot 8}\cdot x^4\pm\ldots .$$

$$\frac{1}{\sqrt{1+x}}=(1+x)^{-\frac{1}{2}}=\sum_{n=0}^{\infty}\binom{-\frac{1}{2}}{n}\cdot x^n=1+\sum_{n=1}^{\infty}(-1)^n\cdot\frac{(2n-1)\cdot(2n-3)\cdot\ldots\cdot 3\cdot 1}{\underbrace{(2n)\cdot(2n-2)\cdot(2n-4)\cdot\ldots\cdot 2}_{2^n\cdot n!}}\cdot x^n \tag{d}$$

$$=1-\frac{1}{2}\cdot x+\frac{2}{2\cdot 4}\cdot x^2-\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot x^3+\frac{1\cdot 3\cdot 5\cdot 7}{2\cdot 4\cdot 6\cdot 8}\cdot x^4\mp\ldots .$$

Aus der Formel, (c) erhält man eine brauchbare Näherungsformel für die Wurzelfunktion:

$$\sqrt{1+x}\approx 1+\frac{x}{2}\qquad\text{für kleine}\qquad x.$$

In der Speziellen Relativitätstheorie, verwendet man oft Näherungen des Typs (Formel (d))

$$\left(1-\frac{v^2}{c^2}\right)^{-\frac{1}{2}} \approx 1-\frac{1}{2}\cdot\left(-\frac{v^2}{c^2}\right) = 1+\frac{1}{2}\cdot\frac{v^2}{c^2}.$$

(Die linke Seite heißt Lorentzfaktor, und wird mit γ abgekürzt.)

Beispiel 85 (Taylorreihe des Arcussinus)**.** *Wir betrachten die Arkussinusfunktion*

$$\arcsin = \sin^{-1} : [-1,1] \to \left[-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right]$$

Bemerkung 62.b) (erster Schritt) und Satz 69 (dritter Schritt)[72] *liefern für $x \in (-1,1)$:*

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}} = \left(1+(-x^2)\right)^{-\frac{1}{2}} = \sum_{n=0}^{\infty}\binom{-\frac{1}{2}}{n}\cdot(-1)^n\cdot x^{2n} = \sum_{n=0}^{\infty}\binom{n-\frac{1}{2}}{n}\cdot x^{2n}.$$

Wegen $\arcsin(0)=0$*, liefert Bemerkung 67 die Entwicklung*

$$\arcsin(x) = \sum_{n=0}^{\infty}\binom{n-\frac{1}{2}}{n}\cdot\frac{1}{2n+1}\cdot x^{2n+1} \qquad \forall\, x\in(-1,1),$$

wobei die Reihe auf der rechten Seite ebenfalls den Konvergenzradius 1*. hat. Konkreter gilt für $n \geq 1$:*

$$\binom{n-\frac{1}{2}}{n}\cdot\frac{1}{2n+1} = \frac{(n-\frac{1}{2})\cdot(n-\frac{3}{2})\cdot\ldots\cdot\frac{1}{2}}{n!}\cdot\frac{1}{2n+1} = \frac{(2n-1)\cdot(2n-3)\cdot\ldots\cdot 3\cdot 1}{\underbrace{(2n)\cdot(2n-2)\cdot\ldots\cdot 4\cdot 2}_{2^n\cdot n!}}\cdot\frac{1}{2n+1},$$

also

$$\arcsin(x) = x + \frac{1}{2\cdot 3}\cdot x^3 + \frac{1\cdot 3}{2\cdot 4\cdot 5}\cdot x^5 + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6\cdot 7}\cdot x^7 + \ldots\ .$$

10.2 Rechenregeln für Taylorpolynome und Taylorreihen

Wir wollen nun einige Rechenregeln für Taylorpolynome und Taylorreihen zusammentragen.

Satz 70. *Seien $f,g\colon D\to\mathbb{R}$ n-mal differenzierbar, mit $n\in\mathbb{N}\cup\{\infty\}$.*

1) Für $\lambda,\mu\in\mathbb{R}$ gilt: $\quad \mathrm{T}_p^n(\lambda\cdot f+\mu\cdot g) = \lambda\cdot\mathrm{T}_p^n(f)+\mu\cdot\mathrm{T}_p^n(g)$

2) Für $n\in\mathbb{N}$ gilt: $\quad \mathrm{T}_p^n(f\cdot g) = \mathrm{T}_p^n\big(\mathrm{T}_p^n(f)\cdot\mathrm{T}_p^n(g)\big)$

3) Es gilt

$$\mathrm{T}_p^n(f\cdot g) = \sum_{m=0}^{n}\left(\sum_{k=0}^{m}\frac{f^{(k)}(p)}{k!}\cdot\frac{g^{(m-k)}(p)}{(m-k)!}\right)\cdot\underbrace{(x-p)^m}_{(x-p)^k\cdot(x-p)^{m-k}}. \tag{317}$$

Bemerkung 73.

a) Anschaulich bedeutet Satz 70.2), dass man das Taylorpolynom von $f\cdot g$ erhält, indem man die Taylorpolynome $\mathrm{T}_p^n(f)$ und $\mathrm{T}_p^n(g)$ ausmultipliziert, und anschließend alle Terme der Ordnung $\geq n+1$ wegfallen lässt (die ersten n Ableitungen besagter Terme sind ja in p gleich 0*.*

b) In Satz 70.3), sei $n\equiv\infty$. Dann ist $\mathrm{T}_p^\infty(f\cdot g)$ wegen (317) *gegeben durch das formelle Cauchy-Produkt - siehe Cauchy-Produktformel in Satz 24 – der Taylorreihen*

$$\mathrm{T}_p^\infty(f)(x) = \sum_{\ell=0}^{\infty}\underbrace{\frac{f^{(\ell)}(p)}{\ell!}\cdot(x-p)^\ell}_{(=b_\ell)} \qquad \text{und} \qquad \mathrm{T}_p^\infty(g)(x) = \sum_{\ell=0}^{\infty}\underbrace{\frac{g^{(\ell)}(p)}{\ell!}\cdot(x-p)^\ell}_{(=a_\ell)}.$$

Es gelte $\mathsf{R} := \min(\mathsf{R}[\mathrm{T}_p^\infty(f)],\mathsf{R}[\mathrm{T}_p^\infty(g)]) > 0$*:* *(für $\varepsilon>0$ ist $D[p,\varepsilon] = D\cap(p-\varepsilon,p+\varepsilon)$)*

[72]Man könnte hier im dritten Schritt auch Formel (d) anwenden, sieht dann aber die elegante Umformung im letzten Schritt nicht mehr so gut.

- *Wegen Satz 29.b) konvergieren* $\mathrm{T}_p^\infty(f)(x)$ *und* $\mathrm{T}_p^\infty(g)(x)$ *für jedes* $x \in (p-\mathsf{R}, p+\mathsf{R}) = \mathsf{B}_\mathsf{R}(p)$ *absolut. Wegen Satz 24 und* (317)*, konvergiert dann auch* $\mathrm{T}_p^\infty(f\cdot g)$ *auf* $\mathsf{B}_\mathsf{R}(p)$ *absolut, mit*

$$\mathrm{T}_p^\infty(f\cdot g)|_{\mathsf{B}_\mathsf{R}(p)} = \mathrm{T}_p^\infty(f)|_{\mathsf{B}_\mathsf{R}(p)} \cdot \mathrm{T}_p^\infty(g)|_{\mathsf{B}_\mathsf{R}(p)}. \tag{318}$$

  Bilden der Taylorreihe in p beider Seiten, liefert Satz 70.2) nun sogar für $n = \infty$.[73]
- *Existiert ein* $0 < \varepsilon \le \mathsf{R}$ *mit* $f|_{D[p,\varepsilon]} = \mathrm{T}_p^\infty f|_{D[p,\varepsilon]}$ *und* $g|_{D[p,\varepsilon]} = \mathrm{T}_p^\infty g|_{D[p,\varepsilon]}$*, so folgt aus dem vorangegangenen Punkt auch automatisch:*

$$\mathrm{T}_p^\infty(f\cdot g)|_{D[p,\varepsilon]} \overset{(318)}{=} \mathrm{T}_p^\infty f|_{D[p,\varepsilon]} \cdot \mathrm{T}_p^\infty g|_{D[p,\varepsilon]} = f|_{D[p,\varepsilon]} \cdot g|_{D[p,\varepsilon]} = (f\cdot g)|_{D[p,\varepsilon]}. \tag{319}$$

Beweis. 1) Siehe Übung 113.

3) Die Leibnizformel (305) (zweiter Schritt) liefert

$$\begin{aligned}\mathrm{T}_p^n(f\cdot g)(x) &= \sum_{m=0}^n \frac{(f\cdot g)^{(m)}(p)}{m!}\cdot(x-p)^m = \sum_{m=0}^\infty \frac{1}{m!}\cdot\left(\sum_{k=0}^m \binom{m}{k}\cdot f^{(k)}(p)\cdot g^{(m-k)}(p)\right)\cdot(x-p)^m\\ &= \sum_{m=0}^n\left(\sum_{k=0}^m \frac{f^{(k)}(p)}{k!}\cdot\frac{g^{(m-k)}(p)}{(m-k)!}\right)\cdot(x-p)^m,\end{aligned}$$

was die Behauptung zeigt.

2) Für $x \in \mathbb{R}$ erhalten wir (Leibnizformel (305) bzw. Teil 3) in der letzten Identität)

$$\begin{aligned}&\mathrm{T}_p^n(f)(x)\cdot\mathrm{T}_p^n(g)(x)\\ &= \left(\sum_{k=0}^n \frac{f^{(k)}(p)}{k!}\cdot(x-p)^k\right)\cdot\left(\sum_{\ell=0}^n \frac{g^{(\ell)}(p)}{\ell!}\cdot(x-p)^\ell\right)\\ &= \sum_{m=0}^{2n}\sum_{k,\ell\in\mathbb{N}:\, k+\ell=m} \frac{f^{(k)}(p)}{k!}\cdot\frac{g^{(\ell)}(p)}{\ell!}\cdot(x-p)^{k+\ell}\\ &= \sum_{m=0}^{n}\sum_{k,\ell\in\mathbb{N}:\, k+\ell=m} \frac{f^{(k)}(p)}{k!}\cdot\frac{g^{(\ell)}(p)}{\ell!}\cdot(x-p)^{k+\ell} + \underbrace{\sum_{n<k+\ell\le 2n} \frac{f^{(k)}(p)}{k!}\cdot\frac{g^{(\ell)}(p)}{\ell!}\cdot(x-p)^{k+\ell}}_{=:\,\mathcal{O}[n+1](x)}\\ &= \underbrace{\sum_{m=0}^n\left(\sum_{k=0}^m \frac{f^{(k)}(p)}{k!}\cdot\frac{g^{(m-k)}(p)}{(m-k)!}\right)\cdot(x-p)^m}_{\mathrm{T}_p^n(f\cdot g)(x)} + \mathcal{O}[n+1](x).\end{aligned}$$

Aus der Definition von $\mathcal{O}[n+1]$ ist nun klar, dass gilt:

$$\mathcal{O}[n+1]^{(j)}(p) = 0 \qquad \forall\, 0 \le j \le n \qquad \text{also} \qquad \mathrm{T}_p^n\big(\mathcal{O}[n+1]\big) = 0. \tag{320}$$

Wir erhalten somit aus Teil 1) und (304), dass

$$\mathrm{T}_p^n\big(\mathrm{T}_p^n(f)\cdot\mathrm{T}_p^n(g)\big) \overset{1)}{=} \mathrm{T}_p^n\big(\mathrm{T}_p^n(f\cdot g)\big) + \mathrm{T}_p^n\big(\mathcal{O}[n+1]\big) \overset{(320)}{=} \mathrm{T}_p^n\big(\mathrm{T}_p^n(f\cdot g)\big) \overset{(304)}{=} \mathrm{T}_p^n(f\cdot g). \qquad \square$$

[73]Beide Seiten sind glatt, mit $\mathrm{T}_p^\infty\big(\mathrm{T}_p^\infty(f\cdot g)\big) = \mathrm{T}_p^\infty(f\cdot g)$ wegen Bemerkung 70.a) bzw. Satz 67.

Beispiel 86. *Wir suchen die Taylorreihe* $\mathrm{T}_0^\infty(h)$ *in* $p=0$*, der Funktion*

$$h\colon (-1,1)\to\mathbb{R},\qquad x\mapsto \frac{\log(1+x)}{1+x}.$$

Aus Beispiel 81.b) und Korollar 22 (geometrische Reihe mit $z=-x$*) wissen wir, dass*

$$f(x):=\log(1+x)=\underbrace{\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\cdot x^k}_{\overset{(*)}{=}\mathrm{T}_0^\infty(f)(x)}\qquad \text{und}\qquad g(x):=\frac{1}{1+x}=\underbrace{\sum_{k=0}^{\infty}(-1)^k\cdot x^k}_{\overset{(*)}{=}\mathrm{T}_0^\infty(g)(x)} \tag{321}$$

für alle $x\in(-1,1)$ *gilt. Die Gleichheiten* $(*)$ *werden durch Satz 67 gewährleistet, und liefern*

$$f(0)=0,\quad \tfrac{f^{(k)}(0)}{k!}=\tfrac{(-1)^{k-1}}{k}\quad \text{für}\quad k\in\mathbb{N}_{>0}\qquad \text{sowie}\qquad \tfrac{g^{(k)}(0)}{k!}=(-1)^k\quad \text{für}\quad k\in\mathbb{N}.$$

Mit (317) *für* $n\equiv\infty$ *und* $p\equiv 0$ *(erster Schritt), folgt hiermit:*

$$\mathrm{T}_0^\infty(h)(x)=\mathrm{T}_0^\infty(f\cdot g)(x)=\sum_{n=1}^{\infty}\left(\sum_{k=1}^{n}\frac{(-1)^{k-1}}{k}\cdot(-1)^{n-k}\right)\cdot x^n=\sum_{n=1}^{\infty}(-1)^{n-1}\left(\sum_{k=1}^{n}\frac{1}{k}\right)\cdot x^n.$$

Darüber hinaus zeigt Bemerkung 73.b), dass $\mathrm{T}_p^\infty(h)$ *auf* $(-1,1)$ *(absolut) konvergiert, mit*

$$\mathrm{T}_p^\infty(h)|_{(-1,1)}=\mathrm{T}_p^\infty(f)|_{(-1,1)}\cdot\mathrm{T}_p^\infty(g)|_{(-1,1)}=f\cdot g=h.$$

Wir wollen schließlich noch eine Kettenregel für Taylorpolynome nachweisen:

Satz 71 (Kettenregel für Taylorpolynome)**.** *Seien* $D,\hat{D}\subseteq\mathbb{R}$ *nichtentartete Intervalle, sowie* $g\colon D\to\mathbb{R}$ *und* $f\colon\hat{D}\to\mathbb{R}$ *zwei* n*-mal differenzierbare Abbildungen für* $n\in\mathbb{N}$*, mit* $\operatorname{im}(g)\subseteq\hat{D}$*. Dann gilt*

$$\mathrm{T}_p^n(f\circ g)=\mathrm{T}_p^n\big(\mathrm{T}_{g(p)}^n(f)\circ\mathrm{T}_p^n(g)\big)\qquad\forall\, p\in D. \tag{322}$$

Beispiel 87. *Für* $n=1$ *liefert Satz 71 die Kettenregel* $(f\circ g)'=f'(g(p))\cdot g'(p)$*. In der Tat gilt*

$$\begin{aligned}\mathrm{T}_p^1(g)(x)&=g(p)+g'(p)\cdot(x-p)\\ \mathrm{T}_{g(p)}^1(f)(x)&=f(g(p))+f'(g(p))\cdot(x-g(p))\\ \mathrm{T}_p^1(f\circ g)(x)&=f(g(p))+(f\circ g)'(p)\cdot(x-p)\end{aligned}$$

für alle $x\in\mathbb{R}$*. Einsetzen der ersten Identität in die zweite Identität liefert*

$$\underbrace{\big(\mathrm{T}_{g(p)}^1(f)\circ\mathrm{T}_p^1(g)\big)(y)}_{=:\,\alpha(y)}=f(g(p))+f'(g(p))\cdot\big(\underbrace{\mathrm{T}_p^1(g)(y)-g(p)}_{g'(p)\cdot(y-p)}\big)\qquad\forall\, y\in\mathbb{R},$$

also $\alpha(p)=f(g(p))$ *und* $\alpha'(p)=f'(g(p))\cdot g'(p)$*. Somit gilt für alle* $x\in\mathbb{R}$*:*

$$\begin{aligned}\mathrm{T}_p^1(\alpha)(x)&=\alpha(p)\qquad+\alpha'(p)\cdot(x-p)\\ &=f(g(p))+f'(g(p))\cdot g'(p)\cdot(x-p).\end{aligned}$$

Satz 71 besagt $\mathrm{T}_p^1(f\circ g)=\mathrm{T}_p^1(\alpha)$*, also für* $x\neq p$*:* $\qquad ((x-p)\neq 0)$

$$\cancel{f(g(p))}+(f\circ g)'(p)\cdot\cancel{(x-p)}=\cancel{f(g(p))}+f'(g(p))\cdot g'(p)\cdot\cancel{(x-p)}.$$

Für den Beweis von Satz 71 benötigen wir die folgenden Hilfsaussagen:

Übung 120. *Sei $\alpha\colon D \to \mathbb{R}$ n-mal differenzierbar mit $n \in \mathbb{N}$, sowie $p \in D$. Zeigen Sie:*

$$\mathrm{T}_p^n\big(\alpha^\ell\big) = \mathrm{T}_p^n\big(\mathrm{T}_p^n(\alpha)^\ell\big) \qquad \forall\, \ell \in \mathbb{N}. \tag{323}$$

Beweis. Für $\ell = 0$ gilt $\mathrm{T}_p^n\big(\alpha^\ell\big) = \mathrm{T}_p^n(1) = 1 = \mathrm{T}_p^n(1) = \mathrm{T}_p^n\big(\mathrm{T}_p^n(\alpha)^\ell\big)$. Gilt nun (323) für ein $\ell \in \mathbb{N}$, so folgt ($\mathrm{T}_p^n(\mathrm{T}_p^n(\alpha)) = \mathrm{T}_p^n(\alpha)$ im vierten Schritt)

$$\begin{aligned}\mathrm{T}_p^n\big(\alpha^{\ell+1}\big) &= \mathrm{T}_p^n\big(\alpha\cdot\alpha^\ell\big) \overset{\text{Satz 70.2)}}{=} \mathrm{T}_p^n\big(\mathrm{T}_p^n(\alpha)\cdot\mathrm{T}_p^n\big(\alpha^\ell\big)\big) \overset{(323)}{=} \mathrm{T}_p^n\big(\mathrm{T}_p^n(\alpha)\cdot\mathrm{T}_p^n\big(\mathrm{T}_p^n(\alpha)^\ell\big)\big)\\ &\overset{(304)}{=} \mathrm{T}_p^n\big(\mathrm{T}_p^n(\mathrm{T}_p^n(\alpha))\cdot\mathrm{T}_p^n\big(\mathrm{T}_p^n(\alpha)^\ell\big)\big) \overset{\text{Satz 70.2)}}{=} \mathrm{T}_p^n\big(\mathrm{T}_p^n(\alpha)\cdot\mathrm{T}_p^n(\alpha)^\ell\big) = \mathrm{T}_p^n\big(\mathrm{T}_p^n(\alpha)^{\ell+1}\big).\end{aligned}$$

□

Bemerkung 74. *Sei $g\colon D \to \mathbb{R}$ n-mal differenzierbar mit $n \in \mathbb{N}$, und $p \in D$ ein Punkt.*

a) *Seien $a_0, \ldots, a_n \in \mathbb{R}$, und somit*

$$\psi\colon \mathbb{R} \to \mathbb{R}, \qquad x \mapsto \textstyle\sum_{\ell=0}^n a_\ell \cdot (x - g(p))^\ell \tag{324}$$

eine Polynomfunktion vom Grad $\leq n$. Es folgt mit (323) sowie Satz 70.1)

$$\mathrm{T}_p^n(\psi\circ g) = \mathrm{T}_p^n\big(\psi\circ\mathrm{T}_p^n(g)\big) \quad \big(= \mathrm{T}_p^n\big(\mathrm{T}_p^n(\psi)\circ\mathrm{T}_p^n(g)\big)\big). \tag{325}$$

Beweis. Sei $\alpha := g - g(p)$. Wir erhalten mit (323) sowie Satz 70.1):

$$\begin{aligned}\mathrm{T}_p^n(\psi\circ g) &= \mathrm{T}_p^n\left(\sum_{\ell=0}^n a_\ell\cdot(g-g(p))^\ell\right) \overset{\text{Satz 70.1)}}{=} \sum_{\ell=0}^n a_\ell\cdot\mathrm{T}_p^n\big(\alpha^\ell\big) \overset{(323)}{=} \sum_{\ell=0}^n a_\ell\cdot\mathrm{T}_p^n\big(\mathrm{T}_p^n(\alpha)^\ell\big)\\ &\overset{\text{Satz 70.1)}}{=} \mathrm{T}_p^n\left(\sum_{\ell=0}^n a_\ell\cdot\mathrm{T}_p^n(\alpha)^\ell\right) = \mathrm{T}_p^n\left(\sum_{\ell=0}^n a_\ell\cdot\mathrm{T}_p^n(g-g(p))^\ell\right)\\ &\overset{\text{Satz 70.1)}}{=} \mathrm{T}_p^n\left(\sum_{\ell=0}^n a_\ell\cdot(\mathrm{T}_p^n(g)-g(p))^\ell\right) = \mathrm{T}_p^n\big(\psi\circ\mathrm{T}_p^n(g)\big),\end{aligned}$$

wobei im vorletzten Schritt Satz 70.1) und (304) in der Form $\mathrm{T}_p^n(g(p)) = g(p)$ benutzt wurde. □

b) *Sei $h\colon \hat{D} \to \mathbb{R}$ eine k-mal differenzierbare Abbildung, mit $0 \leq k \leq n$ und $\operatorname{im}(g) \subseteq \hat{D}$. Dann gilt:*

$$h^{(\ell)}(g(p)) = 0 \quad \forall\, 0 \leq \ell \leq k \qquad \Longrightarrow \qquad (h\circ g)^{(\ell)}(p) = 0 \quad \forall\, 0 \leq \ell \leq k. \tag{326}$$

Beweis. • (IA): Für $k = 0$ ist $(h\circ g)^{(0)}(p) = h(g(p)) = 0$.

- (IV): Sei $0 \leq k < n$, sodass für jede k-mal differenzierbare Abbildung $\tilde{h}\colon \hat{D} \to \mathbb{R}$ gilt:

$$\tilde{h}^{(\ell)}(g(p)) = 0 \quad \forall\, 0 \leq \ell \leq k \qquad \Longrightarrow \qquad (\tilde{h}\circ g)^{(\ell)}(p) = 0 \quad \forall\, 0 \leq \ell \leq k. \tag{327}$$

- (IS): Sei $h\colon \hat{D} \to \mathbb{R}$ $(k+1)$-mal differenzierbar, mit $h^{(\ell)}(g(p)) = 0$ für alle $0 \leq \ell \leq k+1$.
  - Es erfüllt $\tilde{h} \equiv h$ die linke Seite von (327), also auch die rechte (IV). Wir müssen daher nur noch $(h\circ g)^{(k+1)}(p) = 0$ nachweisen.
  - Es erfüllt $\tilde{h} := h'$ die linke Seite von (327), also auch die rechte (IV). Die Kettenregel (zweiter Schritt) und die Leibnizregel (305) (dritter Schritt) liefern

$$\begin{aligned}\big(h\circ g\big)^{(k+1)}(p) &= \big((h\circ g)'\big)^{(k)}(p) = \big((h'\circ g)\cdot g'\big)^{(k)}(p)\\ &= \sum_{\ell=0}^k \binom{k}{\ell}\cdot\underbrace{(\tilde{h}\circ g)^{(\ell)}(p)}_{\overset{(\mathrm{IV})}{=}0}\cdot g^{(k+1-\ell)}(p) = 0.\end{aligned}$$

□

Beweis von Satz 71. Wir schreiben $f = \psi + h$ mit $\psi := \mathrm{T}^n_{g(p)}(f)$ und $h := \mathrm{R}^n_{g(p)}(f)$. Dann gilt $f \circ g = \psi \circ g + h \circ g$, und wir erhalten hiermit (erster Schritt)

$$\mathrm{T}^n_p(f \circ g) \overset{\text{Satz 70.1)}}{=} \mathrm{T}^n_p\big(\psi \circ g\big) + \mathrm{T}^n_p(h \circ g) \overset{(325)}{=} \underbrace{\mathrm{T}^n_p\big(\psi \circ \mathrm{T}^n_p(g)\big)}_{\mathrm{T}^n_p\left(\mathrm{T}^n_{g(p)}(f)\ \circ\ \mathrm{T}^n_p(g)\right)} + \mathrm{T}^n_p(h \circ g),$$

mit $\mathrm{T}^n_p(h \circ g) = 0$ wegen

$$h^{(\ell)}(g(p)) = \mathrm{R}^n_{g(p)}(f)^{(\ell)}(g(p)) \overset{(307)}{=} 0 \qquad \forall\, 0 \leq \ell \leq n \qquad \overset{(326)}{\Longrightarrow} \qquad (h \circ g)^{(\ell)}(p) = 0 \qquad \forall\, 0 \leq \ell \leq n$$

$$\Longrightarrow \qquad \mathrm{T}^n_p(h \circ g) = 0. \qquad \square$$

10.3 Reell Analytische Funktionen

Definition 66 (Reell Analytische Funktionen)**.** *Eine glatte Funktion $f\colon D \to \mathbb{R}$ heißt reell analytisch, wenn sie lokal in jedem Punkt durch ihre Taylorreihe dargestellt wird, d.h., wenn zu jedem $p \in D$ ein $0 < \varepsilon \leq \mathsf{R}[\mathrm{T}^\infty_p(f)]$ existiert mit* *($D[p,\varepsilon] \subseteq \mathrm{K}[\mathrm{T}^n_p(f)]$)*

$$f|_{D[p,\varepsilon]} = \mathrm{T}^\infty_p(f)|_{D[p,\varepsilon]} \quad \textit{mit} \quad D[p,\varepsilon] = D \cap (p-\varepsilon, p+\varepsilon). \tag{328}$$

Bemerkung: *Gemäß Satz 67 ist dies gleichbedeutend damit, dass zu jedem fixierten $p \in D$ eine Potenzreihe Q in p mit Konvergenzradius $\mathsf{R}[\mathsf{Q}] > 0$ existiert, sodass $f|_{D[p,\varepsilon]} = \mathsf{Q}|_{D[p,\varepsilon]}$ für ein $0 < \varepsilon \leq \mathsf{R}[\mathsf{Q}]$ gilt.*

Übung 121. *Seien $f, g\colon D \to \mathbb{R}$ reell analytisch, sowie $\lambda, \mu \in \mathbb{R}$. Folgern Sie aus Bemerkung 73, sowie den entsprechenden Rechenregeln für Potenzreihen, dass auch die Funktionen $(f \cdot g)$ und $(\lambda \cdot f + \mu \cdot g)$ reell analytisch sind.*

Am Ende dieses Abschnittes werden die folgenden beiden allgemeinen Sätze bewiesen (Satz 72 und Satz 73). Hierfür benötigen wir zunächst den technisch aufwendigen Doppelreihensatz (Satz 75), weswegen wir an dieser Stelle erst einmal einige Konsequenzen dieser Sätze diskutieren.

Satz 72. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Gegeben seien Potenzreihen*

$$\mathsf{P}(x) = \sum_{n=0}^{\infty} c_n \cdot (x-q)^n \qquad \textit{und} \qquad \mathsf{Q}(x) = \sum_{n=0}^{\infty} b_n \cdot (x-p)^n$$

in den Entwicklungspunkten $q \in \mathbb{K}$ und $p \in \mathbb{K}$, mit $\mathsf{R}[\mathsf{P}] > 0 < \mathsf{R}[\mathsf{Q}]$ sowie $b_0 = \mathsf{Q}(p) = q$. Es existiert ein $0 < r \leq \mathsf{R}[\mathsf{Q}]$ mit $\mathsf{Q}(\mathsf{B}_r(p)) \subseteq \mathrm{K}[\mathsf{P}]$, sowie eine Potenzreihe H in $\mathbb{K}$ im Entwicklungspunkt p

$$\textit{mit} \qquad \mathsf{R}[\mathsf{H}] \geq r \qquad \textit{sowie} \qquad \mathsf{P} \circ \mathsf{Q}|_{\mathsf{B}_r(p)} = \mathsf{H}|_{\mathsf{B}_r(p)}.$$

Genauer ist H gegeben durch

$$\mathsf{H}(x) = c_0 + \sum_{m=1}^{\infty} \left(\sum_{n=1}^{m} c_n \cdot \sum_{p_1,\ldots,p_n \in \mathbb{N}_{>0}\colon\, \sum_{j=1}^n p_j = m} b_{p_1} \cdot \ldots \cdot b_{p_n} \right) \cdot (x-p)^m.$$

Korollar 45 (Verkettung Analytischer Funktionen)**.** *Seien $g\colon D \to \mathbb{R}$ und $f\colon \hat{D} \to \mathbb{R}$ reell analytisch mit $\mathrm{im}(g) \subseteq \hat{D}$. Dann ist auch $f \circ g$ reell analytisch.*

Beweis. Für $\varepsilon > 0$ klein genug, existiert wegen Satz 72 eine Potenzreihe H im Entwicklungspunkt p mit $\mathsf{R}[\mathsf{H}] \geq \varepsilon$, sodass: (beachte $D[p,\varepsilon] = D \cap \mathsf{B}_\varepsilon(p)$)

$$\mathrm{T}^\infty_{g(p)} f \circ \mathrm{T}^\infty_p g|_{\mathsf{B}_\varepsilon(p)} = \mathsf{H}|_{\mathsf{B}_\varepsilon(p)} \quad \overset{\varepsilon \text{ verkleinern}}{\Longrightarrow} \quad \overbrace{f \circ g|_{D[p,\varepsilon]}}^{(f\circ g)|_{D[p,\varepsilon]}} \overset{(*)}{=} \mathrm{T}^\infty_{g(p)} f \circ \mathrm{T}^\infty_p g|_{D[p,\varepsilon]} = \mathsf{H}|_{D[p,\varepsilon]}$$

$$\overset{\text{Satz 67}}{\Longrightarrow} \quad \mathrm{T}^\infty_p(f \circ g) = \mathsf{H}.$$

Erste Implikation (Gleichheit $(*)$):

- Durch Verkleinern von $\varepsilon > 0$, können wir $g|_{D[p,\varepsilon]} = \mathrm{T}_p^\infty g|_{D[p,\varepsilon]}$ erreichen. (Analytizität von g)
- Per Annahme, existiert ein $\hat{\varepsilon} > 0$ mit $f|_{\hat{D}[g(p),\hat{\varepsilon}]} = \mathrm{T}_{g(p)}^\infty f|_{\hat{D}[g(p),\hat{\varepsilon}]}$. (Analytizität von f)
- Durch Verkleinern von ε, können wir $g(D[p,\varepsilon]) \subseteq \hat{D}[g(p),\hat{\varepsilon}]$ erreichen (g ist stetig). □

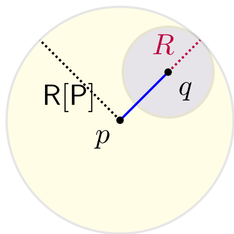
Satz 73. *Sei $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$, und $\mathsf{P}(x) = \sum_{n=0}^\infty c_k \cdot (x-p)^n$ eine Potenzreihe in $\mathbb{K}$ im Entwicklungspunkt $p \in \mathbb{K}$, mit $\mathsf{R}[\mathsf{P}] > 0$. Zu jedem $q \in \mathrm{K}[\mathsf{P}]$ existiert eine Potenzreihe Q in $\mathbb{K}$ im Entwicklungspunkt q*[74]

$$\textit{mit} \qquad \mathsf{R}[\mathsf{Q}] \geq R := \mathsf{R}[\mathsf{P}] - |q-p| > 0 \qquad \textit{sowie} \qquad \mathsf{P}|_{\mathsf{B}_R(q)} = \mathsf{Q}|_{\mathsf{B}_R(q)}. \tag{329}$$

Genauer ist Q gegeben durch

$$\mathsf{Q}(x) = \sum_{m=0}^\infty \underbrace{\left(\sum_{n=m}^\infty c_n \cdot \binom{n}{m} \cdot (q-p)^{n-m} \right)}_{=:\, B_m} \cdot (x-q)^m$$

wobei die Koeffizientenreihe B_m für jedes $m \in \mathbb{N}$ absolut konvergiert.



$\mathbb{K} = \mathbb{C}$

Wir wollen nun zunächst einigen Konsequenzen von Satz 73 diskutieren.

Korollar 46. *Jede reelle Potenzreihe mit positivem Konvergenzradius (> 0) stellt auf ihrem Konvergenzkreis eine reell analytische Funktion dar.*

Beweis. Dies folgt sofort aus Satz 73 für $\mathbb{K} = \mathbb{R}$, sowie Satz 67:

Für $f := \mathsf{P}|_{\mathrm{K}[\mathsf{P}]}$ und $D := \mathrm{K}[\mathsf{P}]$, liefert (329): $(D[q,R] = \mathsf{B}_R(q) \subseteq \mathsf{B}_{\mathsf{R}[\mathsf{P}]}(p) = \mathrm{K}[\mathsf{P}])$

$$f|_{D[q,R]} = (\mathsf{P}|_{\mathrm{K}[\mathsf{P}]})|_{D[q,R]} = \mathsf{P}|_{D[q,R]} = \mathsf{P}|_{\mathsf{B}_R(q)} \overset{(329)}{=} \mathsf{Q}|_{\mathsf{B}_R(q)} = \mathsf{Q}|_{D[q,R]}.$$

Wegen Satz 67 gilt $\mathrm{T}_q^\infty(f) = \mathsf{Q}$, womit die Behauptung folgt. □

Proposition 16 (Analytische Fortsetzung)**.** *Sei $f\colon D \to \mathbb{R}$ reell analytisch. Dann gilt $f = \tilde{f}|_D$ für eine reell analytische Funktion $\tilde{f}\colon I \to \mathbb{R}$, wobei $I \subseteq \mathbb{R}$ ein offenes Intervall mit $D \subseteq I$ ist.*

Beweisskizze. Ist beispielsweise $D = [a,b)$ für $\mathbb{R} \ni a < b \in \overline{\mathbb{R}}$, so definiert man

$$\tilde{f}\colon \underbrace{(a - \overbrace{\mathsf{R}[\mathrm{T}_a^\infty(f)]}^{=:\,\mathsf{R}}, b)}_{=:\,I}, \qquad x \mapsto \begin{cases} \mathrm{T}_a^\infty(f)(x) & \text{für} \quad x \in (a-\mathsf{R}, a) \\ f(x) & \text{für} \quad x \in [a,b). \end{cases}$$

Man sieht leicht ein:

- Lokal um a, wird $\tilde{f}$ durch $\mathrm{T}_a^\infty(f)$ dargestellt; und zwar weil f für ein $\varepsilon > 0$ auf $[a, a+\varepsilon)$ durch $\mathrm{T}_a^\infty(f)$ dargestellt wird, und weil $\tilde{f}$ definitionsgemäß auf $(a-\mathsf{R}, a)$ durch $\mathrm{T}_a^\infty(f)$ dargestellt wird.

[74] Es gilt $\mathsf{B}_R(q) \subseteq \mathrm{K}[\mathsf{P}] = \mathsf{B}_{\mathsf{R}[\mathsf{P}]}(p)$: $|x-q| < R \;\Rightarrow\; |x-p| \leq |x-q| + |q-p| < R + |q-p| = \mathsf{R}[\mathsf{P}]$.

- Lokal um $p \in (a,b)$ wird $\tilde f$ durch $\mathrm{T}_p^\infty(f)$ dargestellt, weil $\tilde f|_{(a,b)} = f$ gilt, und weil f lokal um p durch $\mathrm{T}_p^\infty(f)$ dargestellt wird.
- Lokal um $p \in (a-\mathsf{R}, a)$ wird $\tilde f$ durch $\mathrm{T}_p^\infty(\tilde f)$ dargestellt, weil $\tilde f|_{(a,b)} = \mathrm{T}_a^\infty(f)$ gilt, und weil $\mathrm{T}_a^\infty(f)$ wegen Korollar 46 lokal um p durch $\mathrm{T}_p^\infty\big(\mathrm{T}_a^\infty(f)\big)$ dargestellt wird.

Somit wird $\tilde f$ in jedem Punkt lokal durch eine Potenzreihe in diesem Punkt dargestellt, womit die Behauptung aus Satz 67 folgt. □

$*$*Beweis.* Ist D offen, so ist die Behauptung evident. Sei also D nicht offen. Dann ist D von der Form $[a,b]$ mit $\mathbb{R} \ni a < b \in \mathbb{R}$, $(a,b]$ mit $\overline{\mathbb{R}} \ni a < b \in \mathbb{R}$, oder $[a,b)$ mit $\mathbb{R} \ni a < b \in \overline{\mathbb{R}}$ (vgl. Bemerkung 26 und Übung 56). Wir behandeln nur den dritten Fall (die anderen Fälle folgen analog). Sei also $D = [a,b)$ mit $\mathbb{R} \ni a < b \in \overline{\mathbb{R}}$. Wir definieren

$$\tilde f\colon \underbrace{\big(a - \overbrace{\mathsf{R}[\mathrm{T}_a^\infty(f)]}^{=:\,\mathsf{R}}, b\big)}_{=:I}, \qquad x \mapsto \begin{cases} \mathrm{T}_a^\infty(f)(x) & \text{für} \quad x \in (a-\mathsf{R}, a) \\ f(x) & \text{für} \quad x \in [a,b). \end{cases}$$

Per Konstruktion gilt $f = \tilde f|_D$, und zudem ist $\tilde f$ analytisch:

- Sei $p \in (a,b)$. Es existiert $\varepsilon > 0$ mit $f|_{D[p,\varepsilon]} = \mathrm{T}_p^\infty(f)|_{D[p,\varepsilon]}$. Indem wir $\varepsilon > 0$ verkleinern, können wir $\mathsf{B}_\varepsilon(p) \subseteq (a,b) \subseteq I$ annehmen, also $I[p,\varepsilon] = \mathsf{B}_\varepsilon(p) = D[p,\varepsilon]$ $(\varepsilon < \mathsf{R}[\mathrm{T}_a^\infty(f)])$. Dann gilt

  $$\tilde f|_{I[p,\varepsilon]} = \tilde f|_{D[p,\varepsilon]} = \mathrm{T}_p^\infty(f)|_{D[p,\varepsilon]} = \mathrm{T}_p^\infty(f)|_{I[p,\varepsilon]}.$$

  (Wegen Satz 67 gilt dann auch $\mathrm{T}_p^\infty(\tilde f) = \mathrm{T}_p^\infty(f)$.)
- Sei $p = a$. Es existiert $\varepsilon > 0$ mit $f|_{D[a,\varepsilon]} = \mathrm{T}_a^\infty(f)|_{D[a,\varepsilon]}$. Indem wir $\varepsilon > 0$ verkleinern, können wir $\mathsf{B}_\varepsilon(a) \subseteq I$ annehmen, also $I[a,\varepsilon] = \mathsf{B}_\varepsilon(a)$ und $D[a,\varepsilon] = [a, a+\varepsilon)$. Dann gilt:

  - $\tilde f|_{[a,a+\varepsilon)} = f|_{[a,a+\varepsilon)} = f|_{D[a,\varepsilon]} = \mathrm{T}_a^\infty(f)|_{D[a,\varepsilon]} = \mathrm{T}_a^\infty(f)|_{[a,a+\varepsilon)}$.
  - $\tilde f|_{(a-\varepsilon,a)} = \mathrm{T}_a^\infty(f)|_{(a-\varepsilon,a)}$.

  Dies zeigt $\tilde f|_{I[p,\varepsilon]} = \tilde f|_{(a-\varepsilon,a+\varepsilon)} = \mathrm{T}_a^\infty(f)|_{(a-\varepsilon,a+\varepsilon)} = \mathrm{T}_a^\infty(f)|_{I[a,\varepsilon]}$.

  (Wegen Satz 67 gilt dann auch $\mathrm{T}_a^\infty(\tilde f) = \mathrm{T}_a^\infty(f)$.)
- Sei $p \in (a-\mathsf{R}, a)$. Gemäß Korollar 46, existiert ein $0 < \varepsilon \leq \min(|p-a|, \mathsf{R})$ mit

  $$\mathrm{T}_a^\infty(f)|_{\mathsf{B}_\varepsilon(p)} = \mathrm{T}_p^\infty(\mathrm{T}_a^\infty(f))|_{\mathsf{B}_\varepsilon(p)} \quad \overset{I[p,\varepsilon]=\mathsf{B}_\varepsilon(p)}{\Longrightarrow} \quad \tilde f|_{I[p,\varepsilon]} = \mathrm{T}_a^\infty(f)|_{\mathsf{B}_\varepsilon(p)} = \mathrm{T}_p^\infty(\mathrm{T}_a^\infty(f))|_{\mathsf{B}_\varepsilon(p)}.$$

  (Wegen Satz 67 gilt dann auch $\mathrm{T}_p^\infty(\tilde f) = \mathrm{T}_p^\infty(\mathrm{T}_a^\infty(f))|_{\mathsf{B}_\varepsilon(p)}$.) □

Aus Proposition 16 und Korollar 30, erhalten wir die folgende fundamentale Aussage über reell analytische Funktionen:

Satz 74 (Lokale Bestimmtheit)**.** *Sei $f\colon D \to \mathbb{R}$ reell analytisch. Dann sind äquivalent:*

i) *Es gilt $f = 0$.*

ii) *Es existiert eine Folge $(x_n)_{n\in\mathsf{N}}$ in $D \setminus \{p\}$ mit $\lim_n x_n = p$, sowie $f(x_n) = 0$ für alle $n \in \mathsf{N}$.*

Beweis. • Es gelte die Bedingung i), also $f = 0$. Da D nichtentartet ist, existiert ein $p \in D$ und eine Folge $(x_n)_{n\in\mathbb{N}}$ in $D \setminus \{p\}$ mit $\lim_n x_n = p$ (Übung 90 und Bemerkung 53.b)). Dann gilt natürlich $f(x_n) = 0$ für alle $n \in \mathbb{N}$.

- Es gelte die Bedingung ii). Wegen Proposition 16 dürfen wir (der Einfachheit halber) annehmen, dass $I \equiv D$ offen ist (einfach indem wir notfalls zu einer analytischen Fortsetzung von f übergehen, für welche die Voraussetzungen dann ja offensichtlich auch erfüllt sind):

  Wir wählen $\varepsilon > 0$, sodass $f|_{\mathsf{B}_\varepsilon(p)} = \mathrm{T}_p^\infty(f)|_{\mathsf{B}_\varepsilon(p)}$ gilt (vgl. (328)), insbesondere $\mathsf{B}_\varepsilon(p) \subseteq \mathrm{K}[\mathrm{T}_p^\infty(f)]$. Korollar 30 angewandt auf $\mathsf{Q} \equiv \mathrm{T}_p^\infty(f)$ zeigt $\mathrm{T}_p^\infty(f)|_{\mathrm{K}[\mathrm{T}_p^\infty(f)]} = 0$, also $f|_{\mathsf{B}_\varepsilon(p)} = \mathrm{T}_p^\infty|_{\mathsf{B}_\varepsilon(p)} = 0$.

  Es bezeichne nun M die Menge aller offenen Intervalle $J \subseteq \mathbb{R}$ mit $p \in J \subseteq I$, sodass $f|_J = 0$ gilt:

  - Es gilt $(p-\varepsilon, p+\varepsilon) = \mathsf{B}_\varepsilon(p) \in M \neq \emptyset$.
  - Es ist $O := \bigcup_{J \in M} J \subseteq I$ als Vereinigung offener Intervalle offen wegen Proposition 3.OM3), und zudem ein Intervall wegen Übung 22 (beachte $p \in J$ für alle $J \in M$).
  - Es gilt $f|_O = 0$, denn jedes $x \in O$ ist in einem $J \in M$ enthalten.

  Wir wollen nun zeigen, dass $O = I$ gilt; und schreiben hierfür:

  $$I = (\iota_-, \iota_+) \qquad \text{sowie} \qquad O = (o_-, o_+) \qquad \text{mit} \qquad \overline{\mathbb{R}} \ni \iota_- \leq o_- < o_+ \leq \iota_+ \in \overline{\mathbb{R}}.$$

  Angenommen es gilt $o_+ < \iota_+$ (den Fall $\iota_- < o_-$ behandelt man analog). Wir wählen eine Folge $(x_n)_{n\in\mathbb{N}}$ in O mit $\lim_n x_n = o_+ =: \tilde{p}$. Dann gilt $f(x_n) = 0$ für alle $n \in \mathbb{N}$.

  - Es existiert ein $0 < \tilde{\varepsilon} \leq \mathsf{R}[\mathrm{T}_{\tilde{p}}^\infty(f)]$ mit $f|_{\mathsf{B}_{\tilde{\varepsilon}}(\tilde{p})} = \mathrm{T}_{\tilde{p}}^\infty(f)|_{\mathsf{B}_{\tilde{\varepsilon}}(\tilde{p})}$. (Analytizität von f).
  - Es existiert ein $N \in \mathbb{N}$ mit $x_n \in \mathsf{B}_{\tilde{\varepsilon}}(\tilde{p}) \subseteq \mathrm{K}[\mathrm{T}_{\tilde{p}}^\infty(f)]$ für alle $n \in \mathbb{N}$, d.h.,

  $$\begin{aligned} \mathrm{T}_{\tilde{p}}^\infty(f)(x_n) = f(x_n) = 0 \qquad \forall\, n \geq N \qquad &\overset{\text{Korollar 30}}{\Longrightarrow} \qquad \mathrm{T}_{\tilde{p}}^\infty(f)|_{\mathrm{K}[\mathrm{T}_{\tilde{p}}^\infty(f)]} = 0 \\ &\Longrightarrow \qquad f|_{\mathsf{B}_{\tilde{\varepsilon}}(\tilde{p})} = \mathrm{T}_{\tilde{p}}^\infty(f)|_{\mathsf{B}_{\tilde{\varepsilon}}(\tilde{p})} = 0. \end{aligned}$$

  Offensichtlich $\mathsf{B}_{\tilde{\varepsilon}}(\tilde{p}) \cap O \neq \emptyset$, also ist $\tilde{J} := \mathsf{B}_{\tilde{\varepsilon}}(\tilde{p}) \cup O \subseteq I$ gemäß Übung 22 ein Intervall, sowie offen wegen Proposition 3.OM3). Dann gilt $f|_{\tilde{J}} = 0$, also $\tilde{J} \in M$ wegen $p \in \tilde{J}$. Wegen $o_+ = \tilde{p} \in \tilde{J} \backslash O \neq \emptyset$, widerspricht dies der Definition von O. $\square$

Übung 122. *Sei $\emptyset \neq I \subseteq \mathbb{R}$ ein offenes Intervall, und $f\colon I \to \mathbb{R}$ eine Abbildung. Zeigen Sie:*

a) Ist f reell analytisch und $J \subseteq \mathbb{R}$ ein offenes Intervall mit $J \subseteq I$, so ist $f|_J$ ebenfalls reell analytisch.

b) Es ist f reell analytisch, wenn zu jedem $p \in I$ ein offenes Intervall $J \subseteq \mathbb{R}$ mit $p \in J \subseteq I$ existiert, sodass $f|_J$ analytisch ist.

Hinweis: Die Behauptungen folgen zwanglos, wenn Sie die Bemerkung in Definition 66 berücksichtigen (Sie dürfen diese also frei verwenden).

Übung 123. *Seien $\emptyset \neq I, J \subseteq \mathbb{R}$ offene Intervalle mit $I \cap J \neq \emptyset$. Seien $f\colon I \to \mathbb{R}$ und $g\colon J \to \mathbb{R}$ analytische Funktionen. Zeigen Sie, dass $f|_{I\cap J} = g|_{I\cap J}$ gilt, sofern ein $p \in I \cap J$ und eine Folge $(x_n)_{n\in\mathsf{N}}$ in $I \cap J \setminus \{p\}$ existiert, mit $f(x_n) = g(x_n)$ für alle $n \in \mathsf{N}$.*

Hinweis: Zeigen Sie zunächst, dass die Einschränkungen $f|_{I\cap J}$ und $g|_{I\cap J}$ analytisch sind. Benutzen Sie dann Satz 74 und Übung 121.

Übung 124 (Maximale Analytische Fortsetzung)**.** *Sei $f\colon D \to \mathbb{R}$ analytisch, und es bezeichne M die Menge aller analytische Fortsetzungen von f, d.h., die Menge aller reell analytischen Funktionen $g\colon I \to \mathbb{R}$ mit $I \subseteq \mathbb{R}$ ein offenes Intervall, sodass $D \subseteq I$ und $g|_D = f$ gilt. Wir definieren $\tilde{f}\colon \tilde{I} := \bigcup_{g\in M} \mathrm{dom}(g) \to \mathbb{R}$ durch die Vorschrift*

$$\tilde{f}(x) := g(x) \quad \textit{für ein} \quad g \in M \quad \textit{mit} \quad x \in \mathrm{dom}(g)$$

für alle $x \in \tilde{I}$. Zeigen Sie in der gegeben Reihenfolge:

- $\tilde{I}$ *ist ein offenes Intervall.*
- $\tilde{f}$ *ist wohldefiniert mit* $g = \tilde{f}|_{\mathrm{dom}(g)}$ *für alle* $g \in M$.
- $\tilde{f}$ *ist analytisch.*

Wegen dem dritten Punkt wird $\tilde{f}$ *auch als die maximale analytische Fortsetzung von* f *bezeichnet.*

10.3.1 Der Doppelreihensatz (*)

(In der Vorlesung Satz 75 kurz erklärt. In der Zentralübung bewiesen.)

Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Wir erinnern an die folgenden Sachverhalte:

Bemerkung* 13.

a) Sei $\sum_{n=\mathsf{p}}^{\infty} c_n$ *eine absolut konvergente* $\mathbb{K}$*-Reihe, d.h.* $\sum_{n=\mathsf{p}}^{\infty} |c_n| < \infty$*. Dann konvergiert wegen Satz 21 auch* $\sum_{n=\mathsf{p}}^{\infty} c_n$*, und wegen der Dreiecksungleichung sowie Lemma 30 (Erhalt von Ungleichungen unter Limiten – angewandt auf die entsprechenden Folgen der Partialsummen) gilt*

$$\left|\sum_{n=\mathsf{p}}^{\infty} c_n\right| \leq \sum_{n=\mathsf{p}}^{\infty} |c_n|.$$

b) Seien $\sum_{n=0}^{\infty} a_n$, $\sum_{n=0}^{\infty} b_n$ *positive Reihen, mit* $\sum_{n=0}^{\infty} b_n < \infty$ *konvergent und* $a_n \leq b_n$ *für alle* $n \in \mathbb{N}$*. Dann konvergiert* $\sum_{n=0}^{\infty} a_n$ *nach dem Majorantenkriterium (Satz 25), und Lemma 30 (angewandt auf die entsprechenden Folgen der Partialsummen) zeigt*

$$\sum_{n=0}^{\infty} a_n \leq \sum_{n=0}^{\infty} b_n.$$

c) Sei $\sum_{n=0}^{\infty} c_n < \infty$ *eine positiv und konvergent* $\mathbb{K}$*-Reihe. Dann gilt wegen Satz 18 (monotone Konvergenz für Reihen)*

$$0 \leq \sum_{\ell=0}^{m} c_\ell \leq \sup_{\mathbb{R}}(\{\sum_{k=0}^{n} c_k \mid n \in \mathsf{N}\}) = \sum_{n=0}^{\infty} c_n \qquad \forall\, m \in \mathbb{N}.$$

Satz 75 (Doppelreihensatz)**.** *Sei* $(a_{n,m})_{(n,m)\in\mathbb{N}\times\mathbb{N}}$ *eine Familie von Elementen in* $\mathbb{K}$ *mit*

$$\sum_{m=0}^{\infty} |a_{n,m}| < \infty \quad \forall\, n \in \mathbb{N} \qquad \textit{sowie} \qquad \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} |a_{n,m}|\right) < \infty. \tag{330}$$

Es gelten die folgenden Aussagen:

1) Für jedes $n \in \mathbb{N}$*, konvergiert die Reihe* $\sum_{m=0}^{\infty} a_{n,m}$ *absolut; und die Reihe* $\sum_{n=0}^{\infty} (\sum_{m=0}^{\infty} a_{n,m})$ *ist ebenfalls absolut konvergent (also auch konvergent).*

2) Für jedes $m \in \mathbb{N}$*, konvergiert die Reihe* $\sum_{n=0}^{\infty} a_{n,m}$ *absolut; und die Reihe* $\sum_{m=0}^{\infty} (\sum_{n=0}^{\infty} a_{n,m})$ *ist ebenfalls absolut konvergent (also auch konvergent), mit*

$$\sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a_{n,m}\right) = \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a_{n,m}\right). \tag{331}$$

Beachte: *Die Teile 1) und 2) gelten dann automatisch auch für die Familie* $(\hat{a}_{n,m})_{(n,m)\in\mathbb{N}\times\mathbb{N}}$ *der Absolutwerte* $\hat{a}_{n,m} = |a_{n,m}|$ *für alle* $m, n \in \mathbb{N}$*, da für diese Familie die entsprechenden Voraussetzungen ebenfalls erfüllt sind.*

$*$*Beweis.* 1) Wegen der linken Seiten von (330) konvergiert $\sum_{m=0}^{\infty} a_{n,m}$ für jedes $n \in \mathbb{N}$ absolut. Wegen Bemerkung 13.a), existiert somit auch $\sum_{m=0}^{\infty} a_{n,m} \in \mathbb{K}$ für jedes $n \in \mathbb{N}$, und es gilt

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

Die rechte Seite von (330) zusammen mit Bemerkung 13.b) liefert

$$\sum_{n=0}^{\infty}\left|\sum_{m=0}^{\infty} a_{n,m}\right| \leq \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} |a_{n,m}|\right) < \infty.$$

2) Sei zunächst $a_{n,m} \geq 0$ für alle $n, m \in \mathbb{N}$, d.h.

$$\sum_{m=0}^{\infty} a_{n,m} < \infty \qquad \forall\, n \in \mathbb{N} \qquad \text{sowie} \qquad \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a_{n,m}\right) < \infty \tag{332}$$

wegen (330), da $|a_{n,m}| = a_{n,m}$ für alle $n, m \in \mathbb{N}$. Wegen der rechten Seite von (332) gilt $M := \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a_{n,m}\right) < \infty$. Bemerkung 13.c) liefert

$$\sum_{m=c}^{p}\sum_{n=0}^{q} a_{n,m} \leq \sum_{m=0}^{p}\sum_{n=0}^{q} a_{n,m} = \sum_{n=0}^{q}\sum_{m=0}^{p} a_{n,m} \leq \sum_{n=0}^{q}\left(\sum_{m=0}^{\infty} a_{n,m}\right) \leq M \tag{333}$$

für alle $c, p, q \in \mathbb{N}$ mit $c \leq p$.

- Wählen wir in (333) $m = c = p$, so erhalten wir mit Satz 18 (monotone Konvergenz für Reihen)

$$\sum_{n=0}^{q} a_{n,m} \leq M \qquad \forall\, m, q \in \mathbb{N} \qquad \overset{\text{Satz 18}}{\Longrightarrow} \qquad \sum_{n=0}^{\infty} a_{n,m} \leq M < \infty \qquad \forall\, m \in \mathbb{N}. \tag{334}$$

- Wählen wir in (333) $c = 0$, so erhalten wir mit Lemma 30 im letzten Schritt (die linke Seite existiert wegen der rechten Seite von (334))

$$\sum_{m=0}^{p}\left(\sum_{n=0}^{\infty} a_{n,m}\right) = \lim_{q} \sum_{m=0}^{p}\sum_{n=0}^{q} a_{n,m} \leq M \qquad \forall\, p \in \mathbb{N}.$$

  Satz 18 (monotone Konvergenz für Reihen) zeigt daher die Konvergenz

$$\sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a_{n,m}\right) \leq M = \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a_{n,m}\right) < \infty. \tag{335}$$

  Insbesondere zeigt dies die absolute Konvergenz von $\sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a_{n,m}\right)$ (beachte $a_{n,m} \geq 0$ für alle $n, m \in \mathbb{N}$, also $\left|\sum_{n=0}^{\infty} a_{n,m}\right| = \sum_{n=0}^{\infty} a_{n,m}$).
- Wegen (334) (rechte Seite) und (335) (linke Seite) erfüllt nun die Familie $(\tilde{a}_{n,m})_{(n,m)\in\mathbb{N}\times\mathbb{N}}$ mit $\tilde{a}_{n,m} := a_{m,n} \geq 0$ für alle $m, n \in \mathbb{N}$ die Voraussetzungen unseres Satzes. Die selbe Argumentation wie gerade eben liefert daher (im zweiten Schritt wird (335) für $\tilde{a}_{n,m}$ benutzt)

$$\sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a_{m,n}\right) = \sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} \tilde{a}_{n,m}\right) \leq \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} \tilde{a}_{n,m}\right) = \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a_{m,n}\right).$$

  Indem wir die Summationsindizes in der Form $m \leftrightarrow n$ umbenennen, erhalten wir

$$\sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a_{n,m}\right) \leq \sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a_{n,m}\right).$$

  Zusammen mit (335) zeigt dies (331).

Sei nun $a_{n,m} \in \mathbb{K}$ für alle $n, m \in \mathbb{N}$ beliebig. Dann gilt (332) für die Familie $(\hat{a}_{m,n})_{(n,m)\in\mathbb{N}\times\mathbb{N}}$ mit $\hat{a}_{m,n} = |a_{m,n}|$ für alle $m, n \in \mathbb{N}$. Wir erhalten aus dem bereits Bewiesenen (sowie den Voraussetzungen):

$$\begin{gathered} \sum_{m=0}^{\infty} |a_{n,m}| < \infty \quad \forall\, n \in \mathbb{N} \qquad \text{sowie} \qquad \sum_{n=0}^{\infty} |a_{n,m}| < \infty \quad \forall\, m \in \mathbb{N} \\ \mathbb{K} \ni \sum_{n=0}^{\infty} \left(\sum_{m=0}^{\infty} |a_{n,m}| \right) = \sum_{m=0}^{\infty} \left(\sum_{n=0}^{\infty} |a_{n,m}| \right) \in \mathbb{K}. \end{gathered} \tag{336}$$

Insbesondere wissen wir nun bereits (rechte Seite in der ersten Zeile), dass die Reihe $\sum_{n=0}^{\infty} a_{n,m}$ für alle $m \in \mathbb{N}$ absolut konvergiert. Weiterhin sind wegen (336) die Voraussetzungen unseres Satzes für die Familie $(\tilde{a}_{n,m})_{(n,m)\in\mathbb{N}\times\mathbb{N}}$ mit $\tilde{a}_{n,m} = a_{m,n}$ für alle $n, m \in \mathbb{N}$ erfüllt. Somit zeigt Teil 1) unseres Satzes, angewandt auf diese Familie, dass die Reihe $\sum_{m=0}^{\infty} \left(\sum_{n=0}^{\infty} a_{n,m}\right) = \sum_{m=0}^{\infty} \left(\sum_{n=0}^{\infty} \tilde{a}_{m,n}\right)$ absolut konvergiert, also insbesondere auch konvergiert. Es bleibt daher nur noch (331) nachzuweisen. Sei hierfür

$$\begin{aligned} \alpha(p,q) &:= \sum_{n=p+1}^{\infty} \left(\sum_{m=0}^{\infty} |a_{n,m}| \right) + \sum_{n=0}^{p} \left(\sum_{m=q+1}^{\infty} |a_{n,m}| \right) \\ \beta(p,q) &:= \sum_{m=p+1}^{\infty} \left(\sum_{n=0}^{\infty} |a_{n,m}| \right) + \sum_{m=0}^{p} \left(\sum_{n=q+1}^{\infty} |a_{n,m}| \right) \end{aligned}$$

für alle $p, q \in \mathbb{N}$. Es gilt

$$\alpha(p', q') \leq \alpha(p, q) \qquad \text{für} \qquad \beta(p', q') \leq \beta(p, q) \tag{337}$$

für $p, q, p', q' \in \mathbb{N}$ mit $p \leq p'$ und $q \leq q'$.[75] Weiterhin erhalten wir für $p, q \in \mathbb{N}$ mit der Dreiecksungleichung und Bemerkung 13.a):

$$\begin{aligned} \left| \sum_{n=0}^{\infty} \left(\sum_{m=0}^{\infty} a_{n,m} \right) - \sum_{n=0}^{p} \left(\sum_{m=0}^{q} a_{n,m} \right) \right| &= \left| \sum_{n=p+1}^{\infty} \left(\sum_{m=0}^{\infty} a_{n,m} \right) + \sum_{n=0}^{p} \left(\sum_{m=q+1}^{\infty} a_{n,m} \right) \right| \leq \alpha(p,q) \\ \left| \sum_{m=0}^{\infty} \left(\sum_{n=0}^{\infty} a_{n,m} \right) - \sum_{m=0}^{p} \left(\sum_{n=0}^{q} a_{n,m} \right) \right| &= \left| \sum_{m=p+1}^{\infty} \left(\sum_{n=0}^{\infty} a_{n,m} \right) + \sum_{m=0}^{p} \left(\sum_{n=q+1}^{\infty} a_{n,m} \right) \right| \leq \beta(p,q). \end{aligned}$$

Sei nun $\varepsilon > 0$ vorgegeben:

- Wegen der zweiten Zeile in (336), existiert ein $P \in \mathbb{N}$ mit

$$\sum_{n=P+1}^{\infty} \left(\sum_{m=0}^{\infty} |a_{n,m}| \right) < \frac{\varepsilon}{4} > \sum_{m=P+1}^{\infty} \left(\sum_{n=0}^{\infty} |a_{n,m}| \right).$$

- Wegen der ersten Zeile in (336), existiert ein $Q \in \mathbb{N}$ mit

$$\begin{aligned} \sum_{m=Q+1}^{\infty} |a_{n,m}| &< \frac{\varepsilon}{4 \cdot (P+1)} \qquad \forall\, 0 \leq n \leq P \\ \sum_{n=Q+1}^{\infty} |a_{n,m}| &< \frac{\varepsilon}{4 \cdot (P+1)} \qquad \forall\, 0 \leq m \leq P. \end{aligned}$$

[75] Vergrößert man q, so ändert sich am ersten Summanden nichts, und der zweite Summand verkleinert sich. Vergrößert man p, so verkleinert sich der Gesamtterm wegen $\sum_{m=0}^{\infty} |a_{n,m}| \geq \sum_{m=q+1}^{\infty} |a_{n,m}|$ für alle $n \in \mathbb{N}$.

Wir erhalten

$$\alpha(P,Q) < \frac{\varepsilon}{4} + (P+1)\cdot\frac{\varepsilon}{4\cdot(P+1)} = \frac{\varepsilon}{2} \qquad \text{sowie} \qquad \beta(P,Q) < \frac{\varepsilon}{4} + (P+1)\cdot\frac{\varepsilon}{4\cdot(P+1)} = \frac{\varepsilon}{2}.$$

Wegen (337) gilt dann $\alpha(N,N), \beta(N,N) < \frac{\varepsilon}{2}$ für $N := \max(P,Q)$, also erhalten wir

$$\begin{aligned}
&\left| \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a_{n,m}\right) - \sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a_{n,m}\right) \right| \\
&\qquad \le \left| \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a_{n,m}\right) - \sum_{n=0}^{N}\left(\sum_{m=0}^{N} a_{n,m}\right) \right| + \left| \sum_{m=0}^{N}\left(\sum_{n=0}^{N} a_{n,m}\right) - \sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a_{n,m}\right) \right| \\
&\qquad \le \alpha(N,N) + \beta(N,N) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt hiermit (331). □

Wir können nun die Sätze 72 und 73 beweisen:

10.3.2 Der Beweis von Satz 72 (*)

(In der Vorlesung nur Lemma 63 kurz erläutert.)

Übung 125. *Sei $\ell \in \mathbb{N}_{>0}$, und zudem $\sum_{n=0}^{\infty} c[j]_n$ eine absolut konvergente $\mathbb{K}$-Reihe für $1 \le j \le \ell$, mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Zeigen Sie per Induktion mit Hilfe der Cauchy-Produktformel (Satz 24), dass*

$$\left(\sum_{n=0}^{\infty} c[1]_n\right)\cdot\ldots\cdot\left(\sum_{n=0}^{\infty} c[\ell]_n\right) = \sum_{n=0}^{\infty}\underbrace{\left(\sum_{p_1,\ldots,p_\ell\in\mathbb{N}:\ \sum_{j=1}^{\ell} p_j = n} c[1]_{p_1}\cdot\ldots\cdot c[\ell]_{p_\ell}\right)}_{=:\,\Lambda_n} \tag{338}$$

gilt, wobei die Reihe $\sum_n \Lambda_n$ absolut konvergiert.

Lemma 63. *Sei $\ell \in \mathbb{N}_{>0}$, $q \in \mathbb{K}$, und zudem $\mathsf{P}_j(x) = \sum_{n=0}^{\infty} c[j]_n\cdot(x-p)^n$ für $1 \le j \le \ell$ eine Potenzreihe in $\mathbb{K}$ im Entwicklungspunkt p mit $\mathsf{R}[\mathsf{P}_j] > 0$. Dann gilt für $\mathsf{R} := \min(\mathsf{R}[\mathsf{P}_1],\ldots,\mathsf{R}[\mathsf{P}_\ell])$:*

$$\mathsf{P}_1(x)\cdot\ldots\cdot\mathsf{P}_\ell(x) = \sum_{n=0}^{\infty}\left(\sum_{p_1,\ldots,p_\ell\in\mathbb{N}:\ \sum_{j=1}^{\ell} p_j = n} c[1]_{p_1}\cdot\ldots\cdot c[\ell]_{p_\ell}\cdot(x-q)^n\right) \qquad \forall\, x \in \mathsf{B}_{\mathsf{R}}(p).$$

Ebenfalls gilt für die entsprechenden Absolutreihen $\overline{\mathsf{P}}_j(x) := \sum_{n=0}^{\infty} |c[j]_n\cdot(x-p)^n|$ für $1 \le j \le n$:

$$\overline{\mathsf{P}}_1(x)\cdot\ldots\cdot\overline{\mathsf{P}}_\ell(x) = \sum_{n=0}^{\infty}\left(\sum_{p_1,\ldots,p_\ell\in\mathbb{N}:\ \sum_{j=1}^{\ell} p_j = n} |c[1]_{p_1}\cdot\ldots\cdot c[\ell]_{p_\ell}\cdot(x-p)^n|\right) \qquad \forall\, x \in \mathsf{B}_{\mathsf{R}}(p).$$

Beweis. Dies folgt sofort aus Übung 125, denn für jedes $x \in \mathsf{B}_{\mathsf{R}}(q)$, sind wegen $\mathsf{B}_{\mathsf{R}}(q) \subseteq \mathrm{K}[\mathsf{P}_1] \cap \ldots \cap \mathrm{K}[\mathsf{P}_\ell]$, die Reihen $\mathsf{P}_j(x)$ und $\overline{\mathsf{P}}_j(x)$ für $1 \le j \le \ell$, wegen Satz 29.b) absolut konvergent. □

Beweis von Satz 72. Da $\mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]}$ stetig ist mit $\mathsf{Q}(p) = q$, existiert $0 < R' \le \mathsf{R}[\mathsf{Q}]$ mit $\mathsf{Q}(\mathsf{B}_{R'}(p)) \subseteq \mathrm{K}[\mathsf{P}]$.

- Wir betrachten die Potenzreihe $\tilde{\mathsf{Q}}(x) := \sum_{m=1}^{\infty} |b_m|\cdot x^m < \infty$. Wegen der absoluten Konvergenz von Q auf $\mathrm{K}[\mathsf{Q}]$, existiert ein $r > 0$ mit $\tilde{\mathsf{Q}}(r) = \sum_{m=1}^{\infty} |b_m|\cdot r^m < \infty$. Somit gilt $\mathsf{R}[\tilde{\mathsf{Q}}] \ge r > 0$. Dann ist $\tilde{\mathsf{Q}}|_{\mathsf{B}_r(0)}$ stetig, und wegen $\tilde{\mathsf{Q}}(0) = 0$ existiert somit ein $0 < R \le \min(r, R')$ mit

  $$\tilde{\mathsf{Q}}(|x-p|) = \sum_{m=1}^{\infty} |b_m\cdot(x-p)^n| < \mathsf{R}[\mathsf{P}] \qquad \forall\, x \in \mathsf{B}_R(p).$$

  Zudem gilt $\mathsf{Q}(\mathsf{B}_R(p)) \subseteq \mathrm{K}[\mathsf{P}]$ wegen $R \le R'$.

- Wegen der absoluten Konvergenz von P auf $\mathrm{K}[\mathsf{P}]$, gilt dann für $x \in \mathsf{B}_R(p)$:

$$
\begin{aligned}
\infty > \sum_{n=0}^{\infty} \left| c_n \cdot \left(\sum_{m=1}^{\infty} |b_m \cdot (x-p)^n| \right)^n \right| &\geq \sum_{n=1}^{\infty} \left| c_n \cdot \left(\sum_{m=1}^{\infty} |b_m \cdot (x-p)^n| \right)^n \right| \\
&= \sum_{n=1}^{\infty} |c_n| \cdot \left(\sum_{m=1}^{\infty} |b_m \cdot (x-p)^n| \right)^n \\
&= \sum_{n=1}^{\infty} |c_n| \cdot \left(\sum_{m=1}^{\infty} \sum_{p_1,\dots,p_n \in \mathbb{N}_{>0}\colon\ \sum_{j=1}^n p_j = m} |b_{p_1} \cdot \ldots \cdot b_{p_n} \cdot (x-p)^m| \right) \\
&= \sum_{n=1}^{\infty} \left(\sum_{m=1}^{\infty} \underbrace{|c_n| \cdot \sum_{p_1,\dots,p_n \in \mathbb{N}_{>0}\colon\ \sum_{j=1}^n p_j = m} |b_{p_1} \cdot \ldots \cdot b_{p_n} \cdot (x-p)^m|}_{=:\,\tau_{n,m}} \right),
\end{aligned}
$$

  wobei wir im vierten Schritt für jedes $n \in \mathbb{N}_{>0}$, separat Lemma 63 für den Fall angewandt haben: $\mathsf{P}_j(x) = \mathsf{P}(x) = \sum_{m=0}^{\infty} |\tilde{b}_m \cdot (x-p)^m|$ für alle $1 \leq j \leq \ell = n$, mit $\tilde{b}_0 = 0$ sowie $\tilde{b}_m = b_m$ für $m \in \mathbb{N}_{>0}$. Insbesondere gilt

$$
\sum_{m=1}^{\infty} \tau_{n,m} < \infty \quad \forall\, n \in \mathbb{N} \qquad \text{sowie} \qquad \sum_{n=1}^{\infty} \left(\sum_{m=1}^{\infty} \tau_{n,m} \right) < \infty. \tag{339}
$$

- Analog erhalten wir für $x \in \mathsf{B}_R(p)$:

$$
\begin{aligned}
\mathsf{P}(\mathsf{Q}(x)) = \sum_{n=0}^{\infty} c_n \cdot (\mathsf{Q}(x) - q)^n &= \sum_{n=0}^{\infty} c_n \cdot \left(\sum_{m=1}^{\infty} b_m \cdot (x-p)^m \right)^n \\
&= c_0 + \sum_{n=1}^{\infty} c_n \cdot \left(\sum_{m=1}^{\infty} \sum_{p_1,\dots,p_n \in \mathbb{N}_{>0}\colon\ \sum_{j=1}^n p_j = m} b_{p_1} \cdot \ldots \cdot b_{p_n} \cdot (x-p)^m \right) \\
&= c_0 + \sum_{n=1}^{\infty} \left(\sum_{m=1}^{\infty} \underbrace{c_n \cdot \sum_{p_1,\dots,p_n \in \mathbb{N}_{>0}\colon\ \sum_{j=1}^n p_j = m} b_{p_1} \cdot \ldots \cdot b_{p_n} \cdot (x-p)^m}_{=:\,a_{n,m}} \right),
\end{aligned}
$$

  wobei wir im dritten Schritt für jedes $n \in \mathbb{N}_{>0}$ separat Lemma 63 für den Fall angewandt haben: $\mathsf{P}_j(x) = \mathsf{P}(x) = \sum_{m=0}^{\infty} \tilde{b}_m \cdot (x-p)^m$ für alle $1 \leq j \leq \ell = n$, mit $\tilde{b}_0 = 0$ sowie $\tilde{b}_m = b_m$ für $m \in \mathbb{N}_{>0}$.

Wir setzen $a_{0,m} := 0$ für alle $m \in \mathbb{N}$, sowie $a_{n,0} := 0$ für alle $n \in \mathbb{N}$. Dann gilt $|a_{n,m}| \leq \tau_{n,m}$ für alle $n, m \in \mathbb{N}_{>0}$, sodass die Voraussetzungen (330) von Satz 75 wegen (339) erfüllt sind (Majorantenkriterium). Wir erhalten aus besagtem Satz (zweiter Schritt):

$$
\begin{aligned}
\mathsf{P}(\mathsf{Q}(x)) = c_0 + \sum_{n=0}^{\infty} \left(\sum_{m=0}^{\infty} a_{n,m} \right) &= c_0 + \sum_{m=0}^{\infty} \left(\sum_{n=0}^{\infty} a_{n,m} \right) \\
&= c_0 + \sum_{m=1}^{\infty} \left(\sum_{n=1}^{\infty} c_n \cdot \sum_{p_1,\dots,p_n \in \mathbb{N}_{>0}\colon\ \sum_{j=1}^n p_j = m} b_{p_1} \cdot \ldots \cdot b_{p_n} \right) \cdot (x-p)^m \\
&= c_0 + \sum_{m=1}^{\infty} \left(\sum_{n=1}^{m} c_n \cdot \sum_{p_1,\dots,p_n \in \mathbb{N}_{>0}\colon\ \sum_{j=1}^n p_j = m} b_{p_1} \cdot \ldots \cdot b_{p_n} \right) \cdot (x-p)^m
\end{aligned}
$$

für alle $x \in \mathsf{B}_R(p)$. $\square$

10.3.3 Der Beweis von Satz 73

Beweis von Satz 73. Für $n, m \in \mathbb{N}$ und $x \in \mathsf{B}_R(q)$ definieren wir

$$a_{n,m} := \begin{cases} c_n \cdot \binom{n}{m} \cdot (q-p)^{n-m} \cdot (x-q)^m & \text{für} \quad m \leq n \\ 0 & \text{für} \quad m > n. \end{cases}$$

Der Binomische Lehrsatz liefert:

$$\mathsf{P}(x) = \sum_{n=0}^{\infty} c_n \cdot (\overbrace{(q-p)+(x-q)}^{=(x-p)})^n = \sum_{n=0}^{\infty} c_n \cdot \sum_{m=0}^{n} \binom{n}{m} (q-p)^{n-m} \cdot (x-q)^m = \sum_{n=0}^{\infty} \left(\sum_{m=0}^{\infty} a_{n,m} \right).$$

Wir zeigen nun, dass die Voraussetzungen (330) von Satz 75 erfüllt sind:

- Offensichtlich gilt $\sum_{m=0}^{\infty} |a_{n,m}| = \sum_{m=0}^{n} |a_{n,m}| < \infty$ für jedes $n \in \mathbb{N}$ (die Summe ist endliche).
- $(|q-p| + |x-q|) < |q-p| + R = |q-p| + \mathsf{R}[\mathsf{P}] - |q-p| = \mathsf{R}[\mathsf{P}]$:
  Mit der absoluten Konvergenz von P auf K[P] (Satz 29.b)), folgt

$$\begin{aligned} \sum_{n=0}^{\infty} \left(\sum_{m=0}^{\infty} |a_{n,m}| \right) &= \sum_{n=0}^{\infty} \left(\sum_{m=0}^{n} \left| c_n \cdot \binom{n}{m} (q-p)^{n-m} \cdot (x-q)^m \right| \right) \\ &= \sum_{n=0}^{\infty} \left(|c_n| \cdot \sum_{m=0}^{n} \binom{n}{m} \cdot |(q-p)|^{n-m} \cdot |x-q|^m \right) \\ &= \sum_{n=0}^{\infty} |c_n| \cdot (|q-p| + |x-q|)^n < \infty. \end{aligned}$$

Wir erhalten mit der ersten Aussage in Satz 75.2) (Doppelreihensatz) für $m \in \mathbb{N}$, mit einer Wahl $q \neq x \in \mathsf{B}_R(q)$:

$$\infty > |x-q|^{-m} \cdot \sum_{n=0}^{\infty} |a_{n,m}| = \sum_{n=m}^{\infty} \left| c_n \cdot \binom{n}{m} \cdot (q-p)^{n-m} \right|,$$

was die absolute Konvergenz (also auch die Konvergenz) von $B_m = \sum_{n=m}^{\infty} c_n \cdot \binom{n}{m} \cdot (q-p)^{n-m}$ zeigt. Weiterhin zeigt nun Satz 75.2) (zweiter Schritt):

$$\begin{aligned} \mathsf{P}(x) = \sum_{n=0}^{\infty} \left(\sum_{m=0}^{\infty} a_{n,m} \right) &= \sum_{m=0}^{\infty} \left(\sum_{n=0}^{\infty} a_{n,m} \right) \\ &= \sum_{m=0}^{\infty} \left(\sum_{n=0}^{\infty} c_n \cdot \binom{n}{m} \cdot (q-p)^{n-m} \cdot (x-q)^m \right) \\ &= \sum_{m=0}^{\infty} \underbrace{\left(\sum_{n=m}^{\infty} c_n \cdot \binom{n}{m} \cdot (q-p)^{n-m} \right)}_{B_m} \cdot (x-q)^m = \mathsf{Q}(x). \end{aligned}$$

Dies zeigt $\mathsf{Q}(x) = \mathsf{P}(x)$ für alle $x \in \mathsf{B}_R(q)$, und hiermit folgt nun automatisch $\mathsf{R}[\mathsf{Q}] \geq R$. □

11 Mehr zu Normierten Räumen

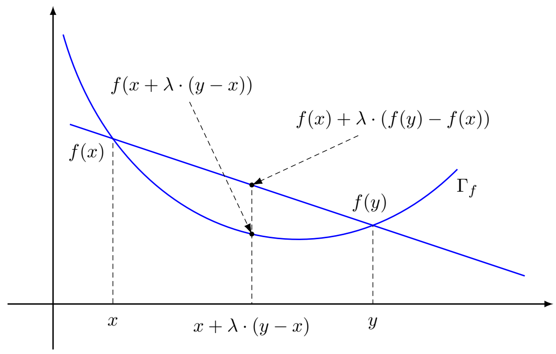
Im Folgenden sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $n \in \mathbb{N}_{>0}$. In diesem Kapitel lernen wir den Begriff der Äquivalenz von Normen kennen, den wir im Falle des $\mathbb{K}^n$ näher untersuchen werden. Zudem werden wir weitere interessante Normen auf dem $\mathbb{K}^n$ kennenlernen, deren unendlichdimensionalen Analoga eine wichtige Rolle in der Maßtheorie spielen. Hierfür benötigen wir zunächst den Begriff der konvexen Funktion.

11.1 Konvexe Funktionen

Definition 67. *Sei $D \subseteq \mathbb{R}$ ein Intervall. Eine Abbildung $f\colon D \to \mathbb{R}$ heißt*

- konvex, *wenn für alle $x, y \in D$ gilt:*

$$f(\underbrace{x + \lambda \cdot (y-x)}_{\lambda \cdot y + (1-\lambda)\cdot x}) \leq \underbrace{f(x) + \lambda \cdot (f(y) - f(x))}_{\lambda \cdot f(y) + (1-\lambda)\cdot f(x)} \qquad \forall\, 0 < \lambda < 1. \tag{340}$$



Beachte: *Es ist* $[0,1] \ni \lambda \mapsto f(x) + \lambda \cdot (f(y) - f(x)) \in \mathbb{R}$ *die Sekante des Graphen Γ_f von f, die durch die beiden Punkte $(x, f(x))$ und $(y, f(y))$ verläuft.*

- konkav, *wenn $-f$ konvex ist, d.h., wenn für alle $x, y \in D$ gilt:*

$$f(x + \lambda \cdot (y-x)) \geq f(x) + \lambda \cdot (f(y) - f(x)) \qquad \forall\, 0 < \lambda < 1. \tag{341}$$

Bemerkung 75. *Die Bedingung* (340) *bzw.* (341) *gilt für alle $x, y \in D$ genau dann, wenn sie für alle $D \ni x < y \in D$ gilt; und zwar wegen*[76]

$$u + \lambda \cdot (v - u) = v + (1-\lambda) \cdot (u - v) \qquad \forall\, u, v, \lambda \in \mathbb{R},$$

wobei $0 < \lambda < 1$ genau dann gilt, wenn $0 < 1 - \lambda < 1$ gilt.

Übung 126. *Zeigen Sie, dass jede affine Funktion sowohl konvex als auch konkav ist.*

Proposition 17. *Eine differenzierbare Abbildung $f\colon D \to \mathbb{R}$ ($D \subseteq \mathbb{R}$ nichtentartet) ist genau dann konvex/konkav, wenn f' monoton wachsend/fallend ist.*

Beachte: Ist f zweimal differenzierbar, so zeigt nun Korollar 32.1), dass f genau dann konvex (konkav) ist, wenn $f'' \geq 0$ ($f'' \leq 0$) gilt.

Beweis. Wir behandeln nur den konvexen Fall (der konkave Fall folgt analog, bzw. durch Multiplikation mit -1).

- Sei f konvex und $D \ni x < y \in D$:
  - Addition von (340) mit $-f(x)$ und anschließende Division durch $\lambda \cdot (y-x) > 0$ liefert

$$\frac{f(x + \lambda \cdot (y-x)) - f(x)}{\lambda \cdot (y-x)} \leq \frac{f(y) - f(x)}{y - x} \qquad \forall\, 0 < \lambda < 1.$$

Wir erhalten aus den Grenzwertsätzen: $\quad f'(x) = \lim_{\lambda \downarrow 0} \frac{f(x+\lambda\cdot(y-x)) - f(x)}{\lambda\cdot(y-x)} \leq \frac{f(y)-f(x)}{y-x}$

[76] Für $x = y$ sind (340) und (341) offensichtlich immer erfüllt.

– Addition von (340) mit $-f(y)$ und anschließende Division durch $(1-\lambda)\cdot(x-y)<0$ liefert:

$$\frac{\overbrace{f(x+\lambda\cdot(y-x))}^{f(y+(1-\lambda)\cdot(x-y))}-f(y)}{(1-\lambda)\cdot(x-y)} \geq \frac{f(x)-f(y)}{x-y} = \frac{f(y)-f(x)}{y-x} \qquad \forall\, 0<\lambda<1.$$

Wir erhalten aus den Grenzwertsätzen: $\frac{f(y)-f(x)}{y-x} \leq \lim_{\lambda\uparrow 1} \frac{f(y+(1-\lambda)\cdot(x-y))-f(y)}{(1-\lambda)\cdot(x-y)} = f'(y)$

Zusammen zeigt dies $f'(x) \leq \frac{f(y)-f(x)}{y-x} \leq f'(y)$.

- Sei f' monoton wachsend. Gemäß Bemerkung 75 genügt es, die Bedingung (340) für den Fall $D \ni x < y \in D$ nachzuweisen. Sei hierfür $0<\lambda<1$ und $c := x+\lambda\cdot(y-x)$. Der Mittelwertsatz (Satz 44) liefert reelle Zahlen $x<\xi_x<c<\xi_y<y$ mit (es gilt $f'(\xi_x)\leq f'(\xi_y)$ per Annahme)

$$\frac{f(c)-f(x)}{c-x} = f'(\xi_x) \leq f'(\xi_y) = \frac{f(y)-f(c)}{y-c}$$

also

$$(y-c)\cdot(f(c)-f(x)) \leq (c-x)\cdot(f(y)-f(c))$$

mit

$$(y-c) = (1-\lambda)\cdot(y-x) \qquad \text{und} \qquad (c-x) = \lambda\cdot(y-x).$$

Division der zweiten Ungleichung durch $(y-x)>0$, liefert nun

$$f(c)-f(x) \;-\; \cancel{\lambda\cdot f(c)} + \lambda\,\cdot f(x) \leq \lambda\cdot f(y) \;-\; \cancel{\lambda\cdot f(c)}$$

also

$$f(x+\lambda\cdot(y-x)) = f(c) \leq f(x)+\lambda\cdot(f(y)-f(x)). \qquad \square$$

Beispiel 88.

a) Die reelle Exponentialfunktion $\exp_{\mathbb{R}}\colon \mathbb{R}\to(0,\infty)$ *ist konvex wegen* $\exp''_{\mathbb{R}} = \exp_{\mathbb{R}} > 0$ *(bzw. da* $\exp'_{\mathbb{R}} = \exp_{\mathbb{R}}$ *wegen Proposition 13.2) (streng) monoton wachsend ist).*

b) Für $\alpha\in\mathbb{R}$*, sei* $\eta_\alpha\colon (0,\infty)\ni x\mapsto x^\alpha = \mathrm{e}^{\alpha\cdot\log(x)}\in(0,\infty)$ *die zugehörige Potenzfunktion (vgl. Bemerkung 56). Dann gilt wegen (247) in Beispiel 69.3)*

$$\eta''_\alpha\colon (0,\infty)\ni x\mapsto \alpha\cdot(\alpha-1)\cdot x^{\alpha-2}\in\mathbb{R}.$$

Gemäß Proposition 17 ist somit η_α *genau dann konvex, wenn* $\alpha\in\mathbb{R}\setminus(0,1)$ *gilt (*$\eta''_\alpha\geq 0$*); sowie genau dann konkav, wenn* $\alpha\in[0,1]$ *gilt (*$\eta''_\alpha\leq 0$*). Für* $\alpha\in\{0,1\}$ *ist* η_α *affin (*$\eta_0 = 1$ *und* $\eta_1 = \mathrm{id}_{\mathbb{R}}$*), also gemäß Übung 126 sowohl konvex als auch konkav.*

Notation 30. *Wir setzen* $0^\alpha := 0$ *für alle* $\alpha\in(0,\infty)$.

Beachte: *Dies ist konsistent mit der Definition der allgemeinen Potenzfunktion, denn gemäß Bemerkung 56.d) gilt* $\lim_{x\downarrow 0} x^\alpha = \lim_{x\downarrow 0}\mathrm{e}^{\alpha\cdot\log(x)} = 0$ *für* $\alpha>0$ *(beachte* $\lim_{x\downarrow 0}\log(x) = -\infty$*).*

Aus Proposition 17 erhalten wir das folgende Korollar.

Korollar 47. *Seien* $p,q\in(1,\infty)$ *mit* $\frac{1}{p}+\frac{1}{q}=1$*. Dann gilt*

$$x^{\frac{1}{p}}\cdot y^{\frac{1}{q}} \leq \frac{x}{p}+\frac{y}{q} \qquad \forall\, x,y\geq 0.$$

Beweis. Wegen Notation 30 ist die Behauptung klar, falls $x=0$ oder $y=0$ gilt. Seien daher $x,y>0$. Aus der Konvexität von $\exp_{\mathbb{R}}$ (vgl. Beispiel 88.b)) sowie der Funktionalgleichung, erhalten wir aus (340) mit $\lambda\equiv\frac{1}{p}$, $(1-\lambda)\equiv\frac{1}{q}$, $f=\exp_{\mathbb{R}}$, $x\equiv\log(x)$, $y\equiv\log(y)$:

$$x^{\frac{1}{p}}\cdot y^{\frac{1}{q}} = \mathrm{e}^{\frac{1}{p}\cdot\log(x)}\cdot\mathrm{e}^{\frac{1}{q}\cdot\log(y)} = \mathrm{e}^{\frac{1}{p}\cdot\log(x)+\frac{1}{q}\cdot\log(y)} \leq \frac{1}{p}\cdot\mathrm{e}^{\log(x)}+\frac{1}{q}\cdot\mathrm{e}^{\log(y)} = \frac{x}{p}+\frac{y}{q}. \qquad \square$$

11.2 Einige Normen auf dem $\mathbb{K}^n$

Sei wieder $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ sowie $n \in \mathbb{N}_{>0}$. In diesem Abschnitt wollen wir einige Abschätzungen und Normen behandeln, die Entsprechungen im unendlich-dimensionalen Rahmen haben, die in der Maßtheorie eine wichtige Rolle spielen.

Terminologie 29. *Für $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ definieren wir*[77]

$$\|x\|_p := \big(|x_1|^p + \ldots + |x_n|^p\big)^{\frac{1}{p}} \qquad \forall\, p \in [1, \infty). \tag{342}$$

Weiterhin definieren wir

$$\|x\|_\infty := \|x\|_{\max} = \max(|x_1|, \ldots, |x_n|) \qquad \forall\, x = (x_1, \ldots, x_n) \in \mathbb{K}^n, \tag{343}$$

mit der Maximumsnorm (125) *aus Beispiel 29.*

Bemerkung: *Eine Plausibilisierung der Notation in* (343) *liefert die Formel (Übung* (128)*)*

$$\lim_{p\uparrow\infty} \|x\|_p = \|x\|_\infty \qquad \forall\, x \in \mathbb{K}^n. \tag{344}$$

Man kann auch jedes $x \in \mathbb{K}^n$ als Abbildung $f_x \colon \{1, \ldots, n\} \ni \ell \mapsto x_\ell \in \mathbb{K}$ auffassen. Dann ist $\|\cdot\|_\infty$ gerade die Supremumsnorm (223)*.*

Wir wollen als nächstes zeigen, dass $\|\cdot\|_p$ für jedes $p \in [1, \infty]$ eine Norm ist. Für den Fall $p = \infty$ und $\mathbb{K} = \mathbb{R}$ hatten wir dies bereits in Übung 55 eingesehen (der Beweis im Falle $\mathbb{K} = \mathbb{C}$ ist genau der gleiche). Ebenfalls unmittelbar folgt die Normeigenschaft im Falle $p = 1$ sofort aus der Normeigenschaft der entsprechenden Betragsfunktion auf $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Übung 127. *Zeigen Sie, dass für $x, y \in (0, \infty)$ und $\alpha \in \mathbb{R}$ gilt:*

$$\log(x \cdot y) = \log(x) + \log(y) \qquad \textit{sowie} \qquad (x \cdot y)^\alpha = x^\alpha \cdot y^\alpha. \tag{345}$$

Zeigen Sie weiterhin, dass die rechte Seite von (345) *wegen Notation 30 ebenfalls für $x, y \in [0, \infty)$ und $\alpha \in (0, \infty)$ gilt.*

Bemerkung 76. *Gemäß Bemerkung 56.b) ist η_α für $\alpha > 0$ streng monoton wachsend, d.h., es gilt (beachte Notation 30 für den Fall $x = 0$):*

$$x^\alpha < y^\alpha \qquad \forall\, 0 \leq x < y < \infty,\ \alpha \in (0, \infty). \tag{346}$$

Hiermit erhalten wir die folgenden Aussagen für $p \in [1, \infty]$:

a) Für $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ und $p \in [1, \infty]$ gilt

$$|x_j| \leq \|x\|_p \qquad \forall\, 1 \leq j \leq n \qquad \textit{also} \qquad \|x\|_\infty \leq \|x\|_p. \tag{347}$$

Beweis. Offensichtlich folgt die rechte Seite aus der linken Seite. Zudem ist die linke Seite für $p = \infty$ klar. Für $p \in [1, \infty)$, liefert (346) mit $\alpha \equiv \frac{1}{p}$:

$$|x_j| \overset{(233)}{=} (|x_j|^p)^{\frac{1}{p}} \overset{(346)}{\leq} \big(|x_1|^p + \ldots + |x_n|^p\big)^{\frac{1}{p}} = \|x\|_p \qquad \forall\, 1 \leq j \leq n. \qquad \square$$

b) Für $p \in [1, \infty]$ gilt (mit $\frac{1}{\infty} := 0$)

$$\|x\|_p \leq n^{\frac{1}{p}} \cdot \|x\|_\infty \qquad \forall\, x \in \mathbb{K}^n. \tag{348}$$

[77] Beachte $\|\cdot\|_1 = |\cdot|$ für $n = 1$; sowie $\|\cdot\|_2 = \|\cdot\|_{\text{euk}}$ im Falle $\mathbb{K} = \mathbb{R}$.

Beweis. Für $p = \infty$ ist die Aussage klar. Für $p \in [1, \infty)$ erhalten wir aus (346) mit $\alpha \equiv p$:

$$(\|x\|_p)^p \overset{(233)}{=} |x_1|^p + \ldots + |x_n|^p \leq n \cdot (\|x\|_\infty)^p \qquad \forall\, x = (x_1, \ldots, x_n) \in \mathbb{K}^n.$$

Die Behauptung folgt somit aus (346) und (345) mit $\alpha \equiv \frac{1}{p}$. □

c) *Mit* (347)*, erhalten wir für* $p \in [1, \infty]$ *die Äquivalenz:*

$$\|x\|_p = 0 \quad \textit{für} \quad x = (x_1, \ldots, x_n) \qquad \Longleftrightarrow \qquad x = 0 \quad (\textit{d.h.}, x_j = 0 \quad \forall\, 1 \leq j \leq n), \tag{349}$$

Sei nun $\lambda \in \mathbb{K}$ *und* $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$*. Für* $p \in (1, \infty)$ *folgt*

$$\|\lambda \cdot x\|_p = \big(\textstyle\sum_{\ell=1}^n |\lambda \cdot x_\ell|^p\big)^{\frac{1}{p}} \overset{(345)}{=} \big(\textstyle\sum_{\ell=1}^n |\lambda|^p \cdot |x_\ell|^p\big)^{\frac{1}{p}} \overset{(345)}{=} |\lambda| \cdot \big(\textstyle\sum_{\ell=1}^n |x_\ell|^p\big)^{\frac{1}{p}} = |\lambda| \cdot \|x\|_p.$$

Im Falle $p \in \{1, \infty\}$ *ist* $\|\lambda \cdot x\|_p = |\lambda| \cdot \|x\|_p$ *ebenfalls klar.*

Übung 128. *Folgern Sie* (344) *aus* (347) *und* (348).

Bemerkung 76.c) zeigt die Normeigenschaften V1) und V2) im Falle $p \in [1, \infty]$. Um nun noch die Dreiecksungleichung nachweisen zu können, benötigen wir zunächst die Hölder-Ungleichung:

Satz 76 (Hölder-Ungleichung)**.** *Seien* $p, q \in (1, \infty)$ *mit* $\frac{1}{p} + \frac{1}{q} = 1$*. Dann gilt*

$$\sum_{j=1}^n |x_j \cdot y_j| \leq \|x\|_p \cdot \|y\|_q = \left(\sum_{\ell=1}^n |x_\ell|^p\right)^{\frac{1}{p}} \cdot \left(\sum_{\ell=1}^n |y_\ell|^q\right)^{\frac{1}{q}} \tag{350}$$

für alle $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ *und* $y = (y_1, \ldots, y_n) \in \mathbb{K}^n$.

Beachte: Die Hölder-Ungleichung (350) gilt auch für den Fall $p = 1$ und $q = \infty$, sofern man die Notation (343) mit der Konvention $\frac{1}{\infty} := 0$ benutzt:

$$\sum_{j=1}^n |x_j \cdot y_j| \leq \|x\|_\infty \cdot \sum_{j=1}^n |y_j| = \|x\|_\infty \cdot \|y\|_1.$$

Beweis. Wegen (349) ist die Behauptung klar, wenn $x = 0$ ($\|x\|_p = 0$) oder $y = 0$ ($\|y\|_p = 0$) gilt. Es seien daher $x \neq 0$ ($\|x\|_p \neq 0$) und $y \neq 0$ ($\|y\|_q \neq 0$). Wir setzen

$$\alpha_j := \frac{|x_j|^p}{(\|x\|_p)^p} > 0 \qquad \text{sowie} \qquad \beta_j := \frac{|y_j|^q}{(\|y\|_q)^q} > 0.$$

Dann gilt $\sum_{j=1}^n \alpha_j = 1 = \sum_{j=1}^n \beta_j$, und weiterhin liefert Korollar 47

$$\frac{|x_j|}{\|x\|_p} \cdot \frac{|y_j|}{\|y\|_q} = \alpha_j^{\frac{1}{p}} \cdot \beta_j^{\frac{1}{q}} \leq \frac{\alpha_j}{p} + \frac{\beta_j}{q} \qquad \forall\, 1 \leq j \leq n.$$

Summation über j liefert

$$\frac{1}{\|x\|_p \cdot \|y\|_q} \cdot \sum_{j=1}^n |x_j| \cdot |y_j| \leq \frac{1}{p} \cdot \sum_j \alpha_j + \frac{1}{q} \cdot \sum_j \beta_j = \frac{1}{p} + \frac{1}{q} = 1.$$

□

Beispiel 89. *Für* $p = q = 2$*, erhalten wir die* Cauchy-Schwarzsche Ungleichung*:*

$$\left|\sum_{j=1}^n x_j \cdot y_j\right| \leq \sum_{j=1}^n |x_j| \cdot |y_j| \leq \|x\|_2 \cdot \|y\|_2.$$

Im Falle $\mathbb{K} = \mathbb{R}$*, wird obige Gleichung zur Cauchy-Schwarzen Ungleichung* $|\langle x, y\rangle_{\text{euk}}| \leq \|x\|_{\text{euk}} \cdot \|y\|_{\text{euk}}$ *für* $x, y \in \mathbb{R}^n$*, welche wir bereits in Lemma 24 bewiesen hatten.*

Aus der Hölder-Ungleichung erhalten wir nun die gewünschte Dreiecksungleichung für die Normen $\|\cdot\|_p$, welche auch als Minkowski-Ungleichung bezeichnet wird.

Satz 77 (Minkowski-Ungleichung)**.** *Für alle* $p \in [1, \infty]$ *gilt*

$$\|x+y\|_p \le \|x\|_p + \|y\|_p \qquad \forall\, x, y \in \mathbb{K}^n. \tag{351}$$

Beweis. • Für $p = \infty$, ist $\|\cdot\|_p$ die Maximumsnorm; also gilt auch die Dreiecksungleichung (351).

- Für $p = 1$ und $z = (z_1, \ldots, z_n) \in \mathbb{K}^n$, ist $\|z\|_p = |z_1| + \ldots + |z_n|$, also folgt (351) unmittelbar aus der Dreiecksungleichung für die Betragsfunktion.
- Sei nun $p \in (1, \infty)$, und setze $q := \frac{p}{p-1} = \left(1 - \frac{1}{p}\right)^{-1} \in (1, \infty)$. Dann gilt

$$\frac{1}{p} + \frac{1}{q} = 1, \qquad p = q \cdot (p-1), \qquad p - \frac{p}{q} = 1. \tag{352}$$

Seien nun $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ und $y = (y_1, \ldots, y_n) \in \mathbb{K}^n$ vorgegeben. Die Behauptung ist klar, wenn $x + y = 0$ gilt. Wir können daher $x + y \neq 0$, also $\|x + y\|_p > 0$ annehmen. Sei $z = (z_1, \ldots, z_n) \in (0, \infty)^n \subseteq \mathbb{K}^n$ definiert durch $z_j := |x_j + y_j|^{p-1} \ge 0$ für $1 \le j \le n$. Dann gilt[78]

$$\begin{aligned} |z_j|^q \overset{(233)}{=} |x_j + y_j|^{q\cdot(p-1)} \overset{(352)}{=} |x_j + y_j|^p \quad \forall\, 1 \le j \le n \quad &\Rightarrow \quad (\|z\|_q)^q = (\|x+y\|_p)^p \\ &\Rightarrow \quad \|z\|_q = (\|x+y\|_p)^{\frac{p}{q}}. \end{aligned} \tag{353}$$

Die Dreiecksungleichung für die Betragsfunktion (dritter Schritt) sowie die Hölder-Ungleichung (vierter Schritt) liefern:

$$\begin{aligned} (\|x+y\|_p)^p &= \sum_{j=1}^n |x_j + y_j|^p \overset{(350)}{=} \sum_{j=1}^n |x_j + y_j| \cdot |z_j| \le \overbrace{\sum_{j=1}^n |x_j| \cdot |z_j|}^{\sum_{j=1}^n |x_j \cdot z_j|} + \overbrace{\sum_{j=1}^n |y_j| \cdot |z_j|}^{\sum_{j=1}^n |y_j \cdot z_j|} \\ &\overset{(350)}{\le} \|x\|_p \cdot \|z\|_q + \|y\|_p \cdot \|z\|_q = (\|x\|_p + \|y\|_p) \cdot \|z\|_q \\ &\overset{(353)}{=} (\|x\|_p + \|y\|_p) \cdot (\|x+y\|_p)^{\frac{p}{q}}. \end{aligned} \tag{354}$$

Wegen $\|x + y\|_p > 0$ erhalten wir mit (233) im zweiten und letzten Schritt, sowie durch Multiplikation beider Seiten von (354) mit $(\|x + y\|_p)^{-\frac{p}{q}} > 0$ im dritten Schritt:

$$\begin{aligned} \|x+y\|_p &\overset{(352)}{=} (\|x+y\|_p)^{p-\frac{p}{q}} \overset{(233)}{=} (\|x+y\|_p)^p \cdot (\|x+y\|_p)^{-\frac{p}{q}} \\ &\overset{(354)}{\le} (\|x\|_p + \|y\|_p) \cdot \underbrace{(\|x+y\|_p)^{\frac{p}{q}} \cdot (\|x+y\|_p)^{-\frac{p}{q}}}_{\overset{(233)}{=} 1} = \|x\|_p + \|y\|_p. \end{aligned}$$

□

Wir erhalten nun schließlich:

Korollar 48. *Für jedes* $p \in [1, \infty]$ *ist* $\|\cdot\|_p \colon \mathbb{K}^n \to \mathbb{R}$ *eine Norm auf* $\mathbb{K}^n$*.*

Beweis. Klar wegen Bemerkung 76.c) und Satz 77. □

[78] Die erste Gleichheit ist wegen unserer Notation 30 (beachte $q, p - 1 > 0$) ebenfalls im Falle $x_j + y_j = 0$ korrekt.

11.3 Äquivalenz von Normen

Definition 68. *Sei V ein $\mathbb{K}$-Vektorraum. Zwei Normen $\|\cdot\|_1$, $\|\cdot\|_2$ auf V heißen zueinander äquivalent, wenn Konstanten $C_1, C_2 > 0$ existieren, mit*

$$\|x\|_1 \leq C_2 \cdot \|x\|_2 \qquad \forall\, x \in V \qquad \textit{sowie} \qquad \|x\|_2 \leq C_1 \cdot \|x\|_1 \qquad \forall\, x \in V. \tag{355}$$

In diesem Fall schreiben wir $\|\cdot\|_1 \sim_{\mathrm{N}} \|\cdot\|_2$.

Übung 129. *Sei V ein $\mathbb{K}$-Vektorraum. Zeigen Sie, dass $\sim_{\mathrm{N}}$ eine Äquivalenzrelation auf der Menge aller Normen auf V definiert.*

Beispiel 90. *Die Normen $\|\cdot\|_p$ auf $\mathbb{K}^n$ mit $n \in \mathbb{N}_{>0}$ und $p \in [1, \infty]$ sind alle äquivalent zueinander, d.h., es gilt $\|\cdot\|_p \sim_{\mathrm{N}} \|\cdot\|_q$ für alle $p, q \in [1, \infty]$.*

Beweis. Es gilt $\|\cdot\|_p \sim_{\mathrm{N}} \|\cdot\|_\infty \sim_{\mathrm{N}} \|\cdot\|_q$ für alle $p, q \in [1, \infty]$ wegen (347) und (348). Also folgt die Behauptung sofort aus Übung 129. $\square$

Wir werden in Abschnitt (11.3.2) weiter unten zeigen, dass alle Normen auf dem $\mathbb{K}^n$ äquivalent zueinander sind (dies gilt in unendlich-dimensionale Vektorräumen im Allgemeinen nicht – siehe Bemerkung 80.b)), dass also der in Beispiel 90 geschilderte Sachverhalt kein Zufall ist. Die Beweisstrategie wird dabei genau die gleich sein, wie in Beispiel 90; nämlich werden wir zeigen, dass jede Norm auf $\mathbb{K}^n$ äquivalent zur Maximumsnorm $\|\cdot\|_\infty$ ist, und dann Übung 129 anwenden. Wir wollen nun aber zunächst einige Grundlegende Sachverhalte darlegen, die aus der Äquivalenz von Normen unmittelbar folgen.

11.3.1 Allgemeine Eigenschaften Äquivalenter Normen

Sei V ein $\mathbb{K}$-Vektorraum mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Wir beginnen unsere Untersuchungen mit dem folgenden offensichtlichen Sachverhalt:

Bemerkung 77 (Beschränktheit)**.** *Sei V ein $\mathbb{K}$-Vektorraum und $\|\cdot\|_1 \sim_{\mathrm{N}} \|\cdot\|_2$ zwei zueinander äquivalente Normen auf V. Eine Teilmenge $Y \subseteq V$ ist beschränkt bezüglich $\|\cdot\|_1$ genau dann, wenn sie beschränkt bezüglich $\|\cdot\|_2$ ist:*

Beweis der Behauptung. Sei Y beschränkt bezüglich $\|\cdot\|_1$ (die andere Richtung folgt analog), d.h., es existiert ein $D \geq 0$ mit $\|y\|_1 \leq D$ für alle $y \in Y$ (vgl. (148)). Wegen (355) existiert ein $C_1 > 0$ mit $\|\cdot\|_2 \leq C_1 \cdot \|\cdot\|_1$; also gilt $\|y\|_2 \leq C_1 \cdot \|y\|_1 \leq C_1 \cdot D \geq 0$ für alle $y \in Y$. $\square$

Proposition 18. *Sei V ein $\mathbb{K}$-Vektorraum und $\|\cdot\|_1 \sim_{\mathrm{N}} \|\cdot\|_2$ zwei zueinander äquivalente Normen auf V. Sei zudem $x \in V$, und $(x_n)_{n\in\mathsf{N}}$ eine Folge in V.*

1) Die Folge $(x_n)_{n\in\mathsf{N}}$ konvergiert bezüglich $\|\cdot\|_1$ gegen x genau dann, wenn sie bezüglich $\|\cdot\|_2$ gegen x konvergiert.

2) Die Folge $(x_n)_{n\in\mathsf{N}}$ ist eine Cauchy-Folge bezüglich $\|\cdot\|_1$ genau dann, wenn sie eine Cauchy-Folge bezüglich $\|\cdot\|_2$ ist.

3) Es ist $(V, \|\cdot\|_1)$ vollständig (ein Banachraum) genau dann, wenn $(V, \|\cdot\|_2)$ vollständig (ein Banachraum) ist.

Beweis. Seien $C_1, C_2 > 0$ wie in (355), d.h., es gilt $\|\cdot\|_1 \leq C_2 \cdot \|\cdot\|_2$ und $\|\cdot\|_2 \leq C_1 \cdot \|\cdot\|_1$.

1) Es gelte $\lim_n x_n = x$ bezüglich $\|\cdot\|_1$. Für $\varepsilon > 0$ vorgegeben, existiert dann ein $N_\varepsilon \in \mathsf{N}$ mit $\|x - x_n\|_1 < \frac{\varepsilon}{C_1}$ für alle $n \geq N_\varepsilon$, d.h.

$$\|x - x_n\|_2 \overset{(355)}{\leq} C_1 \cdot \|x - x_n\|_1 < C_1 \cdot \frac{\varepsilon}{C_1} = \varepsilon \qquad \forall\, n \geq N_\varepsilon.$$

Die zeigt $\lim_n x_n = x$ bezüglich $\|\cdot\|_2$. Die umgekehrte Implikation folgt nun analog.

2) Es sei $(x_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge bezüglich $\|\cdot\|_1$. Für $\varepsilon > 0$ vorgegeben, existiert dann ein $N_\varepsilon \in \mathsf{N}$ mit $\|x_m - x_n\|_1 < \frac{\varepsilon}{C_1}$ für alle $m, n \geq N_\varepsilon$, d.h.

$$\|x_m - x_n\|_2 \overset{(355)}{\leq} C_1 \cdot \|x_m - x_n\|_1 < C_1 \cdot \frac{\varepsilon}{C_1} = \varepsilon \qquad \forall\, m, n \geq N_\varepsilon.$$

Somit ist $(x_n)_{n\in\mathsf{N}}$ ist eine Cauchy-Folge bezüglich $\|\cdot\|_2$. Die umgekehrte Implikation folgt nun analog.

3) Dies folgt sofort aus Teil 1) und Teil 2):

Sei beispielsweise $(V, \|\cdot\|_2)$ vollständig, und $(x_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge in $(V, \|\cdot\|_1)$. Dann ist $(x_n)_{n\in\mathsf{N}}$ wegen Teil 2) auch eine Cauchy-Folge in $(V, \|\cdot\|_2)$; also konvergent gegen ein $x \in V$ bezüglich $\|\cdot\|_2$, wegen der Vollständigkeit von $(V, \|\cdot\|_2)$. Wegen Teil 1) konvergiert dann aber $(x_n)_{n\in\mathsf{N}}$ auch bezüglich $\|\cdot\|_1$ gegen x. Dies zeigt die Vollständigkeit von $(V, \|\cdot\|_1)$. □

Bemerkung 78. *Sei V ein $\mathbb{K}$-Vektorraum mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, und seien $\|\cdot\|_1 \sim_{\mathrm{N}} \|\cdot\|_2$ zueinander äquivalente Normen auf V. Sei weiterhin $Y \subseteq V$ eine Teilmenge, sowie*

$$\mathrm{d}_1 := (\mathrm{d}_{\|\cdot\|_1})_Y = \mathrm{d}_{\|\cdot\|_1}|_{Y\times Y} \qquad \textit{und} \qquad \mathrm{d}_2 := (\mathrm{d}_{\|\cdot\|_2})_Y = \mathrm{d}_{\|\cdot\|_2}|_{Y\times Y}$$

die, zu den Normmetriken $\mathrm{d}_{\|\cdot\|_1}$ und $\mathrm{d}_{\|\cdot\|_2}$ gehörigen Unterraummetriken. Dann gilt:

1) Eine Folge $(y_n)_{n\in\mathsf{N}}$ in Y konvergiert bezüglich d_1 gegen $y \in Y$ genau dann, wenn sie bezüglich d_2 gegen y konvergiert. *(Beweis analog zu Proposition 18.1))*

2) Eine Folge $(y_n)_{n\in\mathsf{N}}$ in Y ist eine Cauchy-Folge bezüglich d_1 genau dann, wenn sie eine Cauchy-Folge bezüglich d_2 ist. *(Beweis analog zu Proposition 18.2))*

3) (Y, d_1) vollständig $\quad\Leftrightarrow\quad$ (Y, d_2) vollständig *(Beweis analog zu Proposition 18.2))*

Wir erinnern an den Begriff der offenen Menge sowie der Umgebung aus Definition 34,[79] sowie unserer Konvention, dass wir einen normierten Raume $(V, \|\cdot\|)$ als metrischen Raum bezüglich der Normmetrik $\mathrm{d}_{\|\cdot\|}$ (siehe (127) in Bemerkung 33) auffassen. Notation 21 folgend, notieren wir mit $\mathsf{B}[\mathrm{d}_{\|\cdot\|}]_\varepsilon(x)$ ($\varepsilon > 0$ und $x \in V$) die entsprechenden offenen ε-Bälle um x, sowie mit $\mathcal{O}[\mathrm{d}_{\|\cdot\|}](V)$ die Menge aller offenen Teilmengen von V. In Beispiel 29 hatten wir anhand der euklidischen Norm und der Maximusnorm auf $\mathbb{R}^n$ bereits schemenhaft angedeutet, dass die Äquivalenz zweier Normen auf einem $\mathbb{K}$-Vektorraum, das gegenseitige Enthaltensein gewisser Normbälle um den Ursprung impliziert. Wir wollen dies nun weiter ausführen, indem wir zeigen, dass zwei zueinander äquivalente Normen auf einem $\mathbb{K}$-Vektorraum das gleiche System von offenen Mengen definieren:

Satz 78. *Sei V ein $\mathbb{K}$-Vektorraum mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, und seien $\|\cdot\|_1 \sim_{\mathrm{N}} \|\cdot\|_2$ zueinander äquivalente Normen auf V. Dann gilt $\mathcal{O}[\mathrm{d}_{\|\cdot\|_1}](V) = \mathcal{O}[\mathrm{d}_{\|\cdot\|_2}](V)$.*

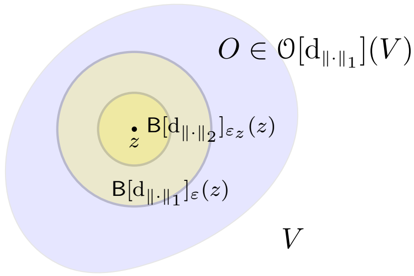
Beweis. Wir zeigen nur die Inklusion $\mathcal{O}[\mathrm{d}_{\|\cdot\|_1}](V) \subseteq \mathcal{O}[\mathrm{d}_{\|\cdot\|_2}](V)$, denn die umgekehrte Inklusion folgt analog. Sei hierfür $O \in \mathcal{O}[\mathrm{d}_{\|\cdot\|_1}](V)$ vorgegeben. Wir müssen nun zeigen:

$$O \in \mathcal{O}[\mathrm{d}_{\|\cdot\|_2}](V) \qquad \overset{\text{def.}}{\Longleftrightarrow} \qquad \forall\, z \in O\colon\ \exists\, \varepsilon_z > 0\colon\ \mathsf{B}[\mathrm{d}_{\|\cdot\|_2}]_{\varepsilon_z}(z) \subseteq O.$$

Sei daher $z \in O$ vorgegeben; sowie $C_2 > 0$ mit (vgl. Definition 68)

$$\|x\|_1 \leq C_2 \cdot \|x\|_2 \qquad \forall\, x \in V.$$

[79] Erinnerung: Eine Teilmenge eines metrischen Raumes heißt Umgebung eines Punktes, wenn sie einen ganzen offenen metrischen Ball um diesen Punkt enthält. Eine Teilmenge heißt offen, wenn sie eine Umgebung aller ihrer Punkte ist.



Per definitionem existiert ein $\varepsilon > 0$ mit $\mathsf{B}[\mathrm{d}_{\|\cdot\|_1}]_\varepsilon(z) \subseteq O$ $\big($es gilt ja $O \in \mathcal{O}[\mathrm{d}_{\|\cdot\|_1}](X)\big)$. Setzen wir nun $\varepsilon_z := \varepsilon \cdot C_2^{-1}$, so erhalten wir

$$\begin{aligned} x \in \mathsf{B}[\mathrm{d}_{\|\cdot\|_2}]_{\varepsilon_z}(z) \quad &\overset{\text{def.}}{\iff} \quad \|z-x\|_2 < \varepsilon_z = \varepsilon \cdot C_2^{-1} \\ &\implies \quad \|z-x\|_1 \leq C_2 \cdot \|z-x\|_2 < \varepsilon \\ &\implies \quad x \in \mathsf{B}[\mathrm{d}_{\|\cdot\|_1}]_\varepsilon(z) \subseteq O, \end{aligned}$$

und daher $\mathsf{B}[\mathrm{d}_{\|\cdot\|_2}]_{\varepsilon_z}(z) \subseteq \mathsf{B}[\mathrm{d}_{\|\cdot\|_1}]_\varepsilon(z) \subseteq O$. □

Wegen Bemerkung 36.a) gilt Satz 78 ebenfalls, wenn man V durch eine Teilmenge $Y \subseteq V$ ersetzt, sowie die auftauchenden Normmetriken durch ihre entsprechenden Einschränkungen auf Y:

Korollar 49. *Sei V ein $\mathbb{K}$-Vektorraum mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, und seien $\|\cdot\|_1 \sim_{\mathrm{N}} \|\cdot\|_2$ zueinander äquivalente Normen auf V. Sei weiterhin $Y \subseteq V$ eine Teilmenge, sowie*

$$\mathrm{d}_1 := (\mathrm{d}_{\|\cdot\|_1})_Y = \mathrm{d}_{\|\cdot\|_1}|_{Y\times Y} \qquad \textit{und} \qquad \mathrm{d}_2 := (\mathrm{d}_{\|\cdot\|_2})_Y = \mathrm{d}_{\|\cdot\|_2}|_{Y\times Y}$$

die, zu den Normmetriken $\mathrm{d}_{\|\cdot\|_1}$ und $\mathrm{d}_{\|\cdot\|_2}$ gehörigen Unterraummetriken. Dann gilt:

1) *Eine Teilmenge $O \subseteq Y$ ist offen bezüglich d_1 genau dann, wenn sie offen bezüglich d_2 ist, d.h.*

$$\mathcal{O}[\mathrm{d}_1](Y) = \mathcal{O}[\mathrm{d}_2](Y).$$

2) *Eine Teilmenge $U \subseteq Y$ ist genau dann eine Umgebung von $p \in Y$ bezüglich d_1, wenn sie eine Umgebung von $p \in Y$ bezüglich d_2 ist.*

Bemerkung 79 (Teilräume und Stetigkeit)**.** *In der Situation von Korollar 49, sei (X, d) ein weiterer metrischer Raum. Zusammen mit den Stetigkeitscharakterisierungen in Proposition 9 folgt:*

a) Es ist $f: Y \to X$ stetig in $p \in Y$ bezüglich d_1 genau dann, wenn f stetig in p bezüglich d_2 ist. Entsprechend ist f genau dann stetig bezüglich d_1, wenn f stetig bezüglich d_2 ist.

b) Es ist $f: X \to Y$ stetig in $p \in X$ bezüglich d_1 genau dann, wenn f stetig in p bezüglich d_2 ist. Entsprechend ist f genau dann stetig bezüglich d_1, wenn f stetig bezüglich d_2 ist.

Beweis von Korollar 49. 1) Wegen Bemerkung 36.a) gilt (erste und letzte Äquivalenz)

$$\begin{aligned} O \in \mathcal{O}[\mathrm{d}_1](Y) \quad &\iff \quad \exists\, U \in \underbrace{\mathcal{O}[\mathrm{d}_{\|\cdot\|_1}](V)}_{\|\ \text{Satz 78}} : O = U \cap Y \\ &\iff \quad \exists\, U \in \overbrace{\mathcal{O}[\mathrm{d}_{\|\cdot\|_2}](V)} : O = U \cap Y \\ &\iff \quad O \in \mathcal{O}[\mathrm{d}_2](Y). \end{aligned}$$

2) Sei U eine Umgebung von p bezüglich d_1 (die andere Implikation folgt analog). Per definitionem existiert ein $\varepsilon > 0$ mit $O := \mathsf{B}[\mathrm{d}_1]_\varepsilon(p) \subseteq U$. Wegen Lemma 27 ist O offen bezüglich d_1, also wegen Teil 1) ebenfalls bezüglich d_2. Somit ist O auch eine Umgebung von p bezüglich d_2, und wegen $O \subseteq U$ dann ebenfalls U. □

11.3.2 Äquivalenz aller Normen auf dem $\mathbb{K}^n$

Sei $n \in \mathbb{N}_{>0}$ und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Wir wollen nun zeigen, dass alle Normen auf dem $\mathbb{K}^n$ äquivalent zueinander sind, d.h

Satz 79. *Je zwei Normen auf $\mathbb{K}^n$ sind zueinander äquivalent.*

Wir beweisen Satz 79, indem wir schrittweise Folgendes zeigen:

(A) Ist $\|\cdot\|$ eine Norm auf $\mathbb{K}^n$ für $n \in \mathbb{N}_{>0}$, so existiert ein $C > 0$ mit $\|\cdot\| < C \cdot \|\cdot\|_\infty$.

(B) Ist $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$ für $n \in \mathbb{N}_{>0}$, so existiert ein $C' > 0$ mit $\|\cdot\|_\infty < C' \cdot \|\cdot\|$.[80]

Zusammen mit (A) und Übung 129, zeigt dies Satz 79 für $\mathbb{K} = \mathbb{R}$.

(C) Gilt (B), so gilt auch die analoge Aussage für $\mathbb{C}$, d.h., ist $\|\cdot\|$ eine Norm auf $\mathbb{C}^n$ für $n \in \mathbb{N}_{>0}$, so existiert ein $C'' > 0$ mit $\|\cdot\|_\infty < C'' \cdot \|\cdot\|$.

Zusammen mit (A) und Übung 129, zeigt dies Satz 79 für $\mathbb{K} = \mathbb{C}$.

Wir zeigen nun erst (A), dann (C), dann (B).

Der Beweis von (A):

Für $1 \leq j \leq n$, bezeichne $e_j \in \mathbb{K}^n$ den j-ten Standardeinheitvektor, d.h. $e_j = (z_1, \ldots, z_n)$ mit

$$z_\ell = \begin{cases} 0 & \text{für} \quad \ell \neq j \\ 1 & \text{für} \quad \ell = j \end{cases} \tag{356}$$

für alle $1 \leq \ell \leq n$. Offensichtlich gilt $x = \sum_{j=1}^n x_j \cdot e_j$ für jedes $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$. Ist nun $\|\cdot\|$ eine Norm auf $\mathbb{K}^n$, so erhalten wir

$$\|x\| = \left\| \sum_{j=1}^n x_j \cdot e_j \right\| \leq \sum_{j=1}^n \|x_j \cdot e_j\| = \sum_{j=1}^n |x_j| \cdot \|e_j\| \leq \max(|x_1|, \ldots, |x_n|) \cdot \overbrace{\sum_{j=1}^n \|e_j\|}^{=: C > 0} = C \cdot \|x\|_\infty. \quad \square$$

Der Beweis von (C) unter der Voraussetzung (B):

Gegeben eine Norm $\|\cdot\|_\mathbb{C}$ auf $\mathbb{C}^n$, so definieren wir eine Norm $\|\cdot\|_\mathbb{R}$ auf $\mathbb{R}^{2n}$ durch

$$\|\cdot\|_\mathbb{R} \colon \mathbb{R}^{2n} \to [0, \infty), \qquad (x_1, \ldots, x_n, y_1, \ldots, y_n) \mapsto \|((x_1 + \mathrm{i} y_1), \ldots, (x_n + \mathrm{i} y_n))\|_\mathbb{C}. \tag{357}$$

Übung 130. *Weisen Sie nach, dass* (357) *eine Norm auf $\mathbb{R}^{2n}$ ist.*

Mit (348) für $n, p \equiv 2$, erhalten wir

$$|x + \mathrm{i}y| = \|(x, y)\|_2 \leq \sqrt{2} \cdot \|(x, y)\|_\infty = \max(|x|, |y|) \qquad \forall\, x, y \in \mathbb{R}.$$

Gegeben $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$, so folgt hiermit

$$\begin{aligned} \|((x_1 + \mathrm{i}y_1), \ldots, (x_n + \mathrm{i}y_n))\|_\infty &= \max(|x_1 + \mathrm{i}y_1|, \ldots, |x_n + \mathrm{i}y_n|) \\ &= \max(\|(x_1, y_1)\|_2, \ldots, \|(x_n, y_n)\|_2) \\ &\leq \sqrt{2} \cdot \max\big(\max(|x_1|, |y_1|), \ldots \max(|x_n|, |y_n|)\big) \\ &= \sqrt{2} \cdot \max\big(|x_1|, \ldots, |x_n|, |y_1|, \ldots, |y_n|\big) \\ &= \sqrt{2} \cdot \|(x_1, \ldots, x_n, y_1, \ldots, y_n)\|_\infty. \end{aligned} \tag{358}$$

[80] Im Beweis wird an einer Stelle auch Aussage (A) benutzt.

Wegen (B), existiert ein $C' > 0$ mit $\|\cdot\|_\infty \leq C' \cdot \|\cdot\|_{\mathbb{R}}$ (für $\|\cdot\|_\infty$ die Maximumsnorm auf $\mathbb{R}^{2n}$). Wir setzen $C'' := \sqrt{2} \cdot C'$, und erhalten

$$\begin{aligned}\|z\|_\infty &= \max(|x_1 + \mathrm{i}y_1|, \ldots, |x_n + \mathrm{i}y_n|) \overset{(358)}{\leq} \sqrt{2} \cdot \|(x_1, \ldots, x_n, y_1, \ldots, y_n)\|_\infty \\ &\leq C'' \cdot \|(x_1, \ldots, x_n, y_1, \ldots, y_n)\|_{\mathbb{R}} = C'' \cdot \|((x_1 + \mathrm{i}y_1), \ldots, (x_n + \mathrm{i}y_n))\|_{\mathbb{C}} = \|z\|_{\mathbb{C}}\end{aligned}$$

für $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$ mit $z_j = x_j + \mathrm{i}y_j$ für $x_j, y_j \in \mathbb{R}$, für $1 \leq j \leq n$. Dies zeigt (C). $\square$

Der Beweis von (B):

Wir erinnern an die Projektionsabbildungen $\mathrm{pr}_j \colon \mathbb{K}^n \ni (y_1, \ldots, y_n) \mapsto y_j \in \mathbb{K}$ für $1 \leq j \leq n$ aus Beispiel 18.b).

Lemma 64. *Eine Folge in $(\mathbb{K}^n, \|\cdot\|_\infty)$ konvergiert genau dann, wenn sie komponentenweise konvergiert, d.h., ist $(x_\ell)_{\ell\in\mathsf{N}}$ eine Folge in $\mathbb{K}^n$ und $x \in \mathbb{K}^n$, dann gilt:*

$$\lim_\ell x_\ell = x \quad \textit{bezüglich} \quad \|\cdot\|_\infty \qquad \Longleftrightarrow \qquad \lim_\ell \mathrm{pr}_j(x_\ell) = \mathrm{pr}_j(x) \qquad \forall\, 1 \leq j \leq n.$$

Beweis. • Es gelte $\lim_\ell x_\ell = x$ bezüglich $\|\cdot\|_\infty$, d.h. $\lim_\ell \|x - x_\ell\|_\infty = 0$ (Übung 62). Wir erhalten

$$0 \leq |\mathrm{pr}_j(x) - \mathrm{pr}_j(x_\ell)| \leq \|x - x_\ell\|_\infty \qquad \forall\, \ell \in \mathsf{N},\ 1 \leq j \leq n.$$

Somit zeigt Lemma 31 (Quetschlemma), dass $\lim_\ell |\mathrm{pr}_j(x) - \mathrm{pr}_j(x_\ell)| = 0$ für alle $1 \leq j \leq n$ gilt.

- Es gelte $\lim_\ell \mathrm{pr}_j(x_\ell) = \mathrm{pr}_j(x)$ für $1 \leq j \leq n$. Zu $\varepsilon > 0$, existieren dann $N[1]_\varepsilon, \ldots, N[n]_\varepsilon \in \mathsf{N}$ mit $|\mathrm{pr}_j(x_\ell) - \mathrm{pr}_j(x)| < \varepsilon$ für alle $\ell \geq N[j]_\varepsilon$ und $1 \leq j \leq n$, d.h.

$$\|x - x_\ell\|_\infty = \max\big(1 \leq j \leq n \,\big|\, |\mathrm{pr}_j(x) - \mathrm{pr}_j(x_\ell)|\big) < \varepsilon \qquad \forall\, \ell \geq N_\varepsilon := \max(N[1]_\varepsilon, \ldots, N[n]_\varepsilon),$$

also $\lim_\ell x_\ell = x$ bezüglich $\|\cdot\|_\infty$. $\square$

Für $n \in \mathbb{N}_{>0}$ und $c \geq 0$, definieren wir den Quader

$$\mathcal{Q}_c := [-c, c]^n = \{x \in \mathbb{R}^n \mid \|x\|_\infty \leq c\} = \overline{\mathsf{B}}[\mathrm{d}_{\|\cdot\|_\infty}]_c(0) \subseteq \mathbb{R}^n.$$

Lemma 65. *Sei $n \in \mathbb{N}_{>0}$ und $c \geq 0$. Jede Folge in $\mathcal{Q}_c$ hat eine Teilfolge, die bezüglich $\|\cdot\|_\infty$ gegen ein Element in $\mathcal{Q}_c$ konvergiert.*

Beweis. Sei $(x_\ell)_{\ell\in\mathsf{N}}$ eine Folge in $\mathcal{Q}_c$. Wir schreiben $x_\ell = (x[1]_\ell, \ldots, x[n]_\ell)$ für alle $\ell \in \mathsf{N}$, d.h.

$$(x[j]_\ell)_{\ell\in\mathsf{N}} \quad \text{ist Folge in} \quad [-c, c] \quad \text{mit} \quad x[j]_\ell = \mathrm{pr}_j(x_\ell) \qquad \forall\, 1 \leq j \leq n.$$

Wir wenden nun Korollar 18 induktiv an:

- Wegen Korollar 18 existiert eine konvergente Teilfolge $(x[1]_{\iota_1(\ell)})_{\ell\in\mathbb{N}} \to x[1] \in [-c, c]$.
- Die Teilfolge $(x[2]_{\iota_1(\ell)})_{\ell\in\mathbb{N}}$ ist ebenfalls eine Folge in $[-c, c]$:
  - Korollar 18 liefert eine konvergente Teilfolge $(x[2]_{(\iota_2\circ\iota_1)(\ell)})_{\ell\in\mathbb{N}} \to x[2] \in [-c, c]$.
  - Wegen Korollar 16 gilt dann ebenfalls $(x[1]_{(\iota_2\circ\iota_1)(\ell)})_{\ell\in\mathbb{N}} \to x[1] \in [-c, c]$.
- Die Teilfolge $(x[3]_{(\iota_2\circ\iota_1)(\ell)})_{\ell\in\mathbb{N}}$ ist ebenfalls eine Folge in $[-c, c]$.
  - Korollar 18 liefert eine konvergente Teilfolge $(x[3]_{(\iota_3\circ\iota_2\circ\iota_1)(\ell)})_{\ell\in\mathbb{N}} \to x[3] \in [-c, c]$.
  - Wegen Korollar 16 gilt dann ebenfalls

$$(x[1]_{(\iota_3\circ\iota_2\circ\iota_1)(\ell)})_{\ell\in\mathbb{N}} \to x[1] \in [-c, c] \qquad \text{sowie} \qquad (x[2]_{(\iota_3\circ\iota_2\circ\iota_1)(\ell)})_{\ell\in\mathbb{N}} \to x[2] \in [-c, c].$$

Indem wir in dieser Weise induktiv fortfahren, erhalten wir eine Teilfolge $(x_{\iota(\ell)})_{\ell\in\mathbb{N}}$ von $(x_\ell)_{\ell\in\mathbb{N}}$ ($\iota = \iota_n \circ \ldots \circ \iota_1$), sowie $x[1], \ldots, x[n] \in [-c, c]$ mit

$$(x[j]_{\iota(\ell)})_{\ell\in\mathbb{N}} \to x[j] \qquad \forall\, 1 \leq j \leq n.$$

Gemäß Lemma 64 gilt dann $(x_{\iota(\ell)})_{\ell\in\mathbb{N}} \to x := (x[1], \ldots, x[n]) \in \mathfrak{Q}_c$ bezüglich $\|\cdot\|_\infty$. □

Beweis von Aussage (B). Sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$. Wir setzen

$$A := \underbrace{\{x \in \mathbb{R}^n \mid \|x\|_\infty = 1\}}_{\neq \emptyset} \subseteq \mathfrak{Q}_1, \qquad \Theta := \underbrace{\{\|x\| \mid x \in A\}}_{\subseteq (0,\infty)} \neq \emptyset \qquad \tilde{C} := \inf(\Theta) \in [0, \infty).$$

Für $0 \neq x \in \mathbb{R}^n$, gilt $\frac{x}{\|x\|_\infty} \in A$ mit $\left\|\frac{x}{\|x\|_\infty}\right\| > 0$:[81]

- Dies zeigt $A \neq \emptyset$, also auch $\Theta \neq \emptyset$. Zudem gilt $\Theta \subseteq (0, \infty)$ wegen: $x \in A \;\Rightarrow\; x \neq 0 \;\Rightarrow\; \|x\| > 0$. Somit ist $\tilde{C} \in [0, \infty)$ definiert.
- Es genügt $\tilde{C} > 0$ zu zeigen; denn dann folgt mit $C' := \tilde{C}^{-1}$ (dritte Implikation):

$$0 \neq x \in \mathbb{R}^n \quad\Rightarrow\quad \frac{x}{\|x\|_\infty} \in A \quad\Rightarrow\quad \|x\| = \|x\|_\infty \cdot \underbrace{\left\|\frac{x}{\|x\|_\infty}\right\|}_{\in\,\Theta} \geq \|x\|_\infty \cdot \tilde{C} \quad\Rightarrow\quad \|x\|_\infty \leq C' \cdot \|x\|.$$

Wir nehmen nun $\tilde{C} = 0$ an; und wählen eine Folge $(\vartheta_\ell)_{\ell\in\mathbb{N}}$ in Θ mit $\lim_\ell \vartheta_\ell = 0$ (Übung 64). Per definitionem, existiert für jedes $\ell \in \mathbb{N}$, ein $x_\ell = (x[1]_\ell, \ldots, x[n]_\ell) \in A$ mit $\vartheta_\ell = \|x_\ell\|$. Wegen $A \subseteq \mathfrak{Q}_1$, liefert Lemma 65 eine bezüglich $\|\cdot\|_\infty$ konvergente Teilfolge $(x_{\iota(\ell)})_{\ell\in\mathbb{N}} \to x \in \mathfrak{Q}_1$:

- Wegen $0 \leq \big|\|x\|_\infty - \|x_{\iota(\ell)}\|_\infty\big| \leq \|x - x_{\iota(\ell)}\|_\infty$ für alle $\ell \in \mathbb{N}$ (inverse Dreiecksungleichung), gilt:

$$\|x\|_\infty = \lim\nolimits_\ell \|x_{\iota(\ell)}\|_\infty = 1 \qquad\Longrightarrow\qquad A \ni x \neq 0 \qquad\Longrightarrow\qquad \|x\| \neq 0. \tag{359}$$

- Wegen der Aussage (A) und der inversen Dreiecksungleichung, gilt $\qquad (C > 0$ wie in (A))

$$0 \leq \big|\|x\| - \|x_{\iota(\ell)}\|\big| \leq \|x - x_{\iota(\ell)}\| \overset{(A)}{\leq} C \cdot \underbrace{\|x - x_{\iota(\ell)}\|_\infty}_{\overset{\ell\to\infty}{\longrightarrow} 0} \qquad \forall\, \ell \in \mathbb{N}, \tag{360}$$

  und daher $\lim_\ell \|x_{\iota(\ell)}\| = \|x\| \neq 0$ wegen (359).
- Jede Teilfolge von $(\vartheta_\ell)_{\ell\in\mathbb{N}}$ konvergiert ebenfalls gegen 0, und wir erhalten hiermit den Widerspruch

$$0 = \lim\nolimits_\ell \vartheta_{\iota(\ell)} = \lim\nolimits_\ell \|x_{\iota(\ell)}\| = \|x\| \neq 0.$$

□

Wir erhalten das folgende Korollar zu Satz 79, Proposition 18 und Lemma 64.

Korollar 50. *Sei $\|\cdot\|$ eine Norm auf $\mathbb{K}^n$.*

1) Eine Folge in $(\mathbb{K}^n, \|\cdot\|)$ konvergiert genau dann, wenn sie komponentenweise konvergiert, d.h., ist $(x_\ell)_{\ell\in\mathbb{N}}$ eine Folge in $\mathbb{K}^n$ und $x \in \mathbb{K}^n$, dann gilt:

$$\lim\nolimits_\ell x_\ell = x \quad \textit{bezüglich} \quad \|\cdot\| \quad \overset{\substack{\text{Satz 79}\\ \text{Proposition 18}}}{\Longleftrightarrow} \quad \lim\nolimits_\ell x_\ell = x \quad \textit{bezüglich} \quad \|\cdot\|_\infty$$
$$\overset{\text{Lemma 64}}{\Longleftrightarrow} \quad \lim\nolimits_\ell \operatorname{pr}_j(x_\ell) = \operatorname{pr}_j(x) \quad \forall\, 1 \leq j \leq n.$$

2) Es ist $(\mathbb{K}^n, \|\cdot\|)$ ein Banachraum, d.h. jede Cauchy-Folge konvergiert.

[81] Hier und im Folgenden setzen wir $\frac{w}{\lambda} := \frac{1}{\lambda} \cdot w$ für alle $\lambda \in \mathbb{R}$ und $w \in \mathbb{R}^n$.

Beweis. 1) Wegen Satz 79 gilt $\|\cdot\| \sim_{\mathsf{N}} \|\cdot\|_\infty$.

2) Sei $(x_\ell)_{\ell\in\mathsf{N}}$ eine Cauchy-Folge in $(\mathbb{K}^n, \|\cdot\|)$. Wegen Satz 79 und Proposition 18.2) ist $(x_\ell)_{\ell\in\mathsf{N}}$ eine Cauchy-Folge in $(\mathbb{K}^n, \|\cdot\|_\infty)$, d.h., zu jedem $\varepsilon > 0$ existiert ein $N_\varepsilon \in \mathbb{N}$ mit

$$\begin{aligned} & \|x_\ell - x_k\|_\infty < \varepsilon && \forall\, \ell, k \geq N_\varepsilon \\ \Longrightarrow \qquad & |\mathrm{pr}_j(x_\ell) - \mathrm{pr}_j(x_k)| \leq \|x_\ell - x_k\|_\infty < \varepsilon && \forall\, \ell, k \geq N_\varepsilon,\ 1 \leq j \leq n. \end{aligned}$$

Daher ist für jedes $1 \leq j \leq n$ die Komponentenfolge $(\mathrm{pr}_j(x_\ell))_{\ell\in\mathsf{N}}$ eine Cauchy-Folge in $\mathbb{K}$, konvergiert also gegen ein $x[j] \in \mathbb{K}$ wegen Satz 17 und Korollar 21. Wegen Teil 1) konvergiert dann also $(x_\ell)_{\ell\in\mathsf{N}}$ bezüglich $\|\cdot\|$ gegen $(x[1], \ldots, x[n])$). □

Übung 131. *Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum, mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $n \geq 1$. Zeigen Sie, dass alle Normen auf V äquivalent zueinander sind. Wählen Sie hierfür einen Isomorphismus (lineare Bijektion mit notwendigerweise linearer Umkehrabbildung) $\alpha\colon V \to \mathbb{K}^n$; und beachten Sie, dass dann beispielsweise $\|\cdot\| \circ \alpha$ eine Norm auf V ist wenn $\|\cdot\|$ eine Norm auf $\mathbb{K}^n$ ist (Nachweis!).*

Bemerkung 80. *Satz 79 ist in dem Sinne beachtlich und wichtig, als dass auf einem $\mathbb{K}$-Vektorraum V üblicherweise sehr viele verschiedene Normen existieren können (siehe* a)*). Er ist im unendlich-dimensionalen Falle im allgemeinen nicht mehr korrekt (siehe* b)*):*

a) *Im Falle $V = \mathbb{K}^n$, hatten wir in Abschnitt 11.2 bereits die Normen $\|\cdot\|_p$ mit $p \in [1, \infty]$ kennengelernt (Korollar 48). Ganz allgemein gilt nun auch Folgendes. Ist V ein $\mathbb{K}$-Vektorraum, und sind $\|\cdot\|_1, \ldots, \|\cdot\|_n$ Normen auf V, so auch (Übung)*

$$\textstyle\sum_{k=1}^n \lambda_k \cdot \|\cdot\|_k \qquad \textit{und} \qquad \max\big(\|\cdot\|_1, \ldots, \|\cdot\|_n\big).$$

b) *Wir betrachten den $\mathbb{R}$-Vektorraum $V := C^1([0,1], \mathbb{R})$, und hierauf die Normen (Übung)*

$$\|f\|_0 := |f|_\infty \qquad \textit{und} \qquad \|f\|_1 := \max\big(|f|_\infty, |f'|_\infty\big) \quad \textit{für} \quad f \in V.$$

Erinnerung: *Es ist*

$$|f|_\infty = \sup_{\mathbb{R}}\{|f(x)| \mid x \in [0,1]\} \in [0, \infty) \qquad \forall\, f \in \mathrm{B}([0,1], f)$$

die zum Betrag auf $\mathbb{R}$ gehörige Supremumsnorm aus Terminologie 23. Wegen der Stetigkeit von f und f', sind die Ausdrücke $|f|_\infty$ und $|f'|_\infty$ gemäß Satz 35 (Satz vom Maximum) in $\mathbb{R}$ definiert (anders ausgedrückt gilt $C([0,1], \mathbb{R}) \subseteq \mathrm{B}([0,1], \mathbb{R})$).

Dann existiert kein $C > 0$ mit $\|\cdot\|_1 \leq C \cdot \|\cdot\|_0$; denn für $C > 0$ vorgegeben, gilt ja beispielsweise $V \ni f\colon [0,1] \ni x \mapsto \sin(2C \cdot x) \in \mathbb{R}$, mit

$$\begin{aligned} & |f'|_\infty = \sup_{\mathbb{R}}\{|2C \cdot \cos(2C \cdot x)| \mid x \in [0,1]\} \overset{\cos(0)=1}{=} 2C \\ & |f|_\infty = \sup_{\mathbb{R}}\{|\sin(2C \cdot x)| \mid x \in [0,1]\} \leq 1 \\ \textit{also} \qquad & \|f\|_1 = \max(|f|_\infty, |f'|_\infty) \geq 2C > C \geq C \cdot |f|_\infty = C \cdot \|f\|_0. \end{aligned}$$

12 Kompaktheit in Metrischen Räumen

In diesem Abschnitt behandeln wir den Begriff der kompakten Teilmenge eines metrischen Raumes und das Verhalten von stetigen Funktionen auf kompakten Mengen (Satz vom Maximum (Satz 4.2.3), gleichmäßige Stetigkeit). Wir erhalten hierbei von einem abstrakten Standpunkt Sätze, die wir schon für den Spezialfall von Funktionen auf abgeschlossenen beschränkten Intervallen in der Analysis I kennengelernt haben.

12.1 Der Kompaktheitsbegriff

Definition 69. *Sei (X, d) ein metrischer Raum und $Y \subseteq X$ eine Teilmenge.*

a) Eine offene Überdeckung *von Y (in (X, d)) ist eine Familie $(U_j)_{j\in J}$ offener Teilmengen von X, sodass $Y \subseteq \bigcup_{j\in J} U_j$ gilt.*

b) Y heißt kompakt *(in (X, d)), wenn zu jeder offenen Überdeckung $(U_j)_{j\in J}$ von Y (in (X, d)) eine endliche Teilmenge $F \subseteq J$ mit $Y \subseteq \bigcup_{j\in F} U_j$ existiert. In diesem Fall wird $(U_j)_{j\in F}$ als* endliche Teilüberdeckung *von Y bezeichnet.*

(D.h., Y ist kompakt, wenn jede offene Überdeckung von Y eine endliche Teilüberdeckung hat.)

Der metrische Raum (X, d) heißt kompakt*, wenn X (aufgefasst als Teilmenge von X) kompakt ist, d.h., wenn jede offene Überdeckung von X (in (X, d)) eine endliche Teilüberdeckung besitzt.*

Bemerkung 81. *Die Eigenschaft, die unter b) von einer kompakten Menge gefordert wird, heißt die* Heine-Borel–Überdeckungseigenschaft*. Die hier gegebene Formulierung von Kompaktheit wird manchmal auch Überdeckungskompaktheit genannt.*

Man beachte, dass in b) nicht *gefordert wird, dass Y eine endliche offene Überdeckung besitzt, sondern dass jede offene Überdeckung eine solche enthält (jede Teilmenge von X besitzt ja die endliche offene Überdeckung, bestehend aus der offenen Menge X).*

Beispiel 91. *Sei (X, d) ein metrischer Raum und $(x_n)_{n\in\mathsf{N}}$ eine Folge in X, die gegen $x \in X$ konvergiert. Dann ist die folgende Teilmenge kompakt:*

$$Y := \{x_n \mid n \in \mathsf{N}\} \cup \{x\} \subseteq X.$$

Beweis. Sei $(U_j)_{j\in J}$ eine offene Überdeckung von Y. Dann existiert ein $j_x \in J$ mit $x \in U_{j_x}$. Da U_{j_x} eine Umgebung von x ist (Definition 34.4)), existiert ein $\varepsilon > 0$ mit $\mathsf{B}_\varepsilon(x) \subseteq U_{j_x}$, und wegen $\lim_n x_n = x$ ein $N_\varepsilon \in \mathsf{N}$ mit $x_n \in \mathsf{B}_\varepsilon(x) \subseteq U_{j_x}$ für alle $n \geq N_\varepsilon$. Wir wählen nun zu jedem $\mathsf{N} \ni n < N_\varepsilon$ ein $j_n \in J$ mit $x_n \in U_{j_n}$. Für $F := \{j_n \mid \mathsf{N} \ni n \leq N_\varepsilon\} \cup \{j_x\}$ ist dann $Y \subseteq \bigcup_{j\in F} U_j$. □

Beispielsweise ist $\{\frac{1}{n} | n \in \mathbb{N}_{>0}\} \cup \{0\}$ kompakt im metrischen Raum $(\mathbb{R}, \mathrm{d}_{|\cdot|})$. Man beachte hierbei, dass die Teilmenge $Z := \{\frac{1}{n} \mid n \in \mathbb{N}_{>0}\} \subseteq \mathbb{R}$ nicht *kompakt ist, da $((\varepsilon, 1))_{0<\varepsilon<1}$ eine offene Überdeckung von Z ohne endliche Teilüberdeckung ist (Übung).*

Bemerkung 82. *Sei V ein $\mathbb{K}$-Vektorraum mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, und seien $\|\cdot\|_1 \sim_{\mathrm{N}} \|\cdot\|_2$ zueinander äquivalente Normen auf V. Sei weiterhin $Y \subseteq V$ eine Teilmenge, sowie*

$$\mathrm{d}_1 := (\mathrm{d}_{\|\cdot\|_1})_Y = \mathrm{d}_{\|\cdot\|_1}|_{Y\times Y} \qquad \text{und} \qquad \mathrm{d}_2 := (\mathrm{d}_{\|\cdot\|_2})_Y = \mathrm{d}_{\|\cdot\|_2}|_{Y\times Y}$$

die, zu den Normmetriken $\mathrm{d}_{\|\cdot\|_1}$ und $\mathrm{d}_{\|\cdot\|_2}$ gehörigen Unterraummetriken.

Eine Teilmenge $K \subseteq Y$ ist kompakt in (Y, d_1) genau dann, wenn sie kompakt in (Y, d_2) ist.

Beweis der Behauptung. Wegen Korollar 49 definieren beide Metriken die gleichen offenen Teilmengen von Y, also auch die gleichen offenen Überdeckungen. □

12.1.1 Grundlegende Eigenschaften Kompakter Mengen

Sei (X, d) ein metrischer Raum. Wir erinnern an den in Bemerkung* 8 gezeigten Sachverhalt, dass zu $X \ni x \neq y \in X$ vorgegeben, ein $\varepsilon > 0$ existiert mit[82]

$$\mathsf{B}_\varepsilon(x) \cap \mathsf{B}_\varepsilon(y) = \emptyset. \tag{361}$$

[82]Die Aussage gilt z.B. für $\varepsilon := \mathrm{d}(x, y)/2$; denn ist $z \in \mathsf{B}_\varepsilon(x)$, so liefert die inverse Dreiecksungleichung M4): $\mathrm{d}(y, z) \geq |\mathrm{d}(y, x) - \mathrm{d}(x, z)| > 2\varepsilon - \varepsilon = \varepsilon$, also $z \notin \mathsf{B}_\varepsilon(y)$.

Lemma 66. *Sei* (X, d) *ein metrischer Raum. Ist* $Y \subseteq X$ *kompakt, so auch abgeschlossen.*

Beweis. Wir müssen nachweisen, dass $U := X \setminus Y$ offen ist. Sei hierfür $p \in U$ vorgegeben. Wir müssen nun zeigen, dass ein $\varepsilon > 0$ mit $\mathsf{B}_\varepsilon(p) \subseteq U$ existiert:

Wegen (361), existiert zu jedem $y \in Y$ ein $\varepsilon_y > 0$ mit $\mathsf{B}_{\varepsilon_y}(p) \cap \mathsf{B}_{\varepsilon_y}(y) = \emptyset$. Nun ist $(\mathsf{B}_{\varepsilon_y}(y))_{y \in Y}$ eine offene Überdeckung von Y, hat also per Annahme eine endliche Teilüberdeckung $(\mathsf{B}_{\varepsilon_y}(y))_{y \in F}$, d.h., $Y \subseteq \bigcup_{y \in F} \mathsf{B}_{\varepsilon_y}(y)$. Für $\varepsilon := \min(\{\varepsilon_y \mid y \in F\})$, gilt dann

$$\begin{aligned} & \mathsf{B}_\varepsilon(p) \cap \mathsf{B}_{\varepsilon_y}(y) \overset{\varepsilon \le \varepsilon_y}{\subseteq} \mathsf{B}_{\varepsilon_y}(p) \cap \mathsf{B}_{\varepsilon_y}(y) = \emptyset \qquad \forall\, y \in F \\ \Longrightarrow \quad & \mathsf{B}_\varepsilon(p) \cap Y \subseteq \mathsf{B}_\varepsilon(p) \cap \big(\textstyle\bigcup_{y \in F} \mathsf{B}_{\varepsilon_y}(y)\big) = \emptyset \\ \Longrightarrow \quad & \mathsf{B}_\varepsilon(p) \subseteq U. \end{aligned}$$

□

Lemma 67. *Sei* (X, d) *ein metrischer Raum,* $Y \subseteq X$ *kompakt, und* $A \subseteq X$ *abgeschlossen in* (X, d)*. Gilt* $A \subseteq Y$*, so ist auch* A *kompakt.*

Beweis. Sei $(U_j)_{j \in J}$ eine offene Überdeckung von A. Wegen der Abgeschlossenheit von A in (X, d), ist $U := X \setminus A$ offen. Wegen $A \subseteq \bigcup_{j \in J} U_j$, gilt dann $Y \subseteq X = U \cup \bigcup_{j \in J} U_j$. Wegen der Kompaktheit von Y, existiert nun eine endliche Teilmenge $F \subseteq J$ mit

$$A \subseteq Y \subseteq \big(U \cup \textstyle\bigcup_{j \in F} U_j\big) \qquad \overset{A \cap U = \emptyset}{\Longrightarrow} \qquad A \subseteq \bigcup_{j \in F} U_j,$$

womit $(U_j)_{j \in F}$ eine endliche Teilüberdeckung von A ist. □

Bemerkung* 14. *In der Situation von Lemma 67 ist es wegen der Kompaktheit von* Y *(sowie Lemma 66) und* $A \subseteq Y$ *prinzipiell egal, ob wir* A *als abgeschlossen in* X *oder als abgeschlossen im metrische Unterraum* (Y, d_Y) *fordern. Dies folgt aus Bemerkung 36.b), welche besagte:*

$$A \text{ abgeschlossen in } (Y, \mathrm{d}_Y) \quad \Leftrightarrow \quad A = B \cap Y \text{ für ein } B \subseteq X \text{ abgeschlossen in } (X, \mathrm{d}) \qquad (*)$$

Beweis der Behauptung. • Sei A abgeschlossen in (X, d). Wegen $A \subseteq Y$ gilt $A = B \cap Y$ für $B \equiv A$, also die rechte Seite von $(*)$, und somit auch die linke Seite.

- Sei A abgeschlossen in (Y, d_Y), d.h. es gilt die linke Seite von $(*)$. Wegen Lemma 66 ist Y abgeschlossen in (X, d) da kompakt. Somit ist A wegen der rechten Seite von $(*)$, als Schnitt der in (X, d) abgeschlossenen Mengen B und Y (gemäß Korollar 13.AM3) ebenfalls abgeschlossen in (X, d). □

Artverwandt mit dem in Bemerkung 14 dargelegten Sachverhalt ist das folgende Lemma, welches besagt, dass man sich bei der allgemeinen Untersuchung des Kompaktheitsbegriffes im Prinzip auf den Fall kompakter metrischer Räume beschränken kann.

Lemma 68. *Sei* (X, d) *ein metrischer Raum. Eine Teilmenge* $Y \subseteq X$ *ist genau dann kompakt, wenn der metrische (Unter)Raum* (Y, d_Y) *kompakt ist.*

Beweis. • Sei Y kompakt in (X, d). Sei $(\tilde{U}_j)_{j \in J}$ eine offene Überdeckung von Y in (Y, d_Y). Bemerkung 36.a) zeigt:

$$\forall\, j \in J\colon\ \exists\, U_j \in \mathcal{O}[\mathrm{d}](X)\colon\ \tilde{U}_j = \overbrace{U_j \cap Y}^{\subseteq U_j}.$$

Dann ist $(U_j)_{j \in J}$ eine offene Überdeckung von Y in (X, d), hat also per Annahme eine endliche Teilüberdeckung $(U_j)_{j \in F}$. Dann ist $(\tilde{U}_j)_{j \in F}$ eine endliche Teilüberdeckung von Y in (Y, d_Y), wegen

$$Y = Y \cap \bigcup_{j \in F} U_j = \bigcup_{j \in F} \tilde{U}_j.$$

- Sei (Y, d_Y) kompakt. Sei $(U_j)_{j\in J}$ eine offene Überdeckung von Y in (X, d). Wegen Bemerkung 36.a) ist dann

$$\tilde{U}_j := U_j \cap Y \in \mathcal{O}[\mathrm{d}_Y](Y) \ \text{ offen in } \ (Y, \mathrm{d}_Y) \ \text{ für alle } \ j \in J \ \text{ mit } \ Y = Y \cap \bigcup_{j\in J} U_j = \bigcup_{j\in J} \tilde{U}_j.$$

  Daher ist $(\tilde{U}_j)_{j\in J}$ eine offene Überdeckung von Y in (Y, d_Y), hat also per Annahme eine endliche Teilüberdeckung $(\tilde{U}_j)_{j\in F}$. Dann ist $(U_j)_{j\in F}$ eine endliche Teilüberdeckung von Y in (X, d). □

Übung 132. *Sei (X, d) ein metrischer Raum, und $K_1, \ldots, K_n \subseteq X$ kompakte Teilmengen. Zeigen Sie, dass die endliche Vereinigung $K := K_1 \cup \ldots \cup K_n$ ebenfalls kompakt ist.*

Übung 133. *Zeigen Sie, dass jede endliche Teilmenge eines metrischen Raumes kompakt ist.*

Übung 134. *Sei (X, d) ein metrischer Raum, und $(K_j)_{j\in J}$ eine Familie kompakter Teilmengen von X. Zeigen Sie, dass der Schnitt $K := \bigcap_{j\in J} K_j$ ebenfalls kompakt ist.*

Übung 135. *Sei (X, d) ein metrischer Raum, und $(K_n)_{n\in\mathbb{N}}$ eine abzählbare Familie kompakter Teilmengen von X, mit $K_{n+1} \subseteq K_n \neq \emptyset$ für alle $n \in \mathbb{N}$. Zeigen Sie, dass $K := \bigcap_{n\in\mathbb{N}} K_n \neq \emptyset$ gilt.*

Übung 136. *Sei (X, d) ein metrischer Raum, und $(K_n)_{n\in\mathbb{N}}$ eine abzählbare Familie kompakter Teilmengen von X mit $\bigcap_{n\in\mathbb{N}} K_n = \emptyset$. Zeigen Sie, dass ein $m \in \mathbb{N}$ existiert mit $K_0 \cap \ldots \cap K_m = \emptyset$.*

Übung 137. *Sei (X, d) ein metrischer Raum, sowie $Z \subseteq Y \subseteq X$ Teilmengen. Folgern Sie aus Lemma 68, dass Z genau dann kompakt in (X, d) ist, wenn Z kompakt in (Y, d_Y) ist.*

Bemerkung: *Die obige Aussage zeigt, dass Kompaktheit* keine *relative sondern eine absolute Eigenschaft ist. Genauer ist es ja so, dass die Offen- bzw. Abgeschlossenheit einer Teilmenge in einem metrischen Unterraum keinesfalls die Offen- bzw. Abgeschlossenheit im übergeordneten metrischen Raum impliziert, d.h., Offen- bzw. Abgeschlossenheit von Teilmengen sind relative Eigenschaften, die explizit davon abhängen, als Teilmengen welchen Unterraumes man sie auffasst.*

12.1.2 Kompaktheit und Stetige Abbildungen

Lemma 69. *Seien (X, d) und (X', d') metrische Räume, und $f\colon X \to X'$ eine stetige Abbildung. Ist $Y \subseteq X$ kompakt in (X, d), so ist $f(Y) \subseteq X'$ kompakt in (X', d').*

(Bilder von Kompakta unter stetigen Abbildungen sind kompakt.[83])

Beweis. Sei $(U'_j)_{j\in J}$ eine offene Überdeckung von $f(Y)$ in (X', d'):

- $(f^{-1}(U'_j))_{j\in J}$ ist eine offene Überdeckung von Y:
  - Wegen Proposition 9.2) ist $U_j := f^{-1}(U'_j)$ offen für alle $j \in J$.
  - Es gilt (Übung 25 im ersten und dritten Schritt)

$$Y \subseteq f^{-1}(f(Y)) \subseteq f^{-1}\big(\bigcup_{j\in J} U'_j\big) = \bigcup_{j\in J} f^{-1}(U'_j) = \bigcup_{j\in J} U_j.$$

- Da Y kompakt ist, existiert $F \subseteq J$ endlich mit $Y \subseteq \bigcup_{j\in F} U_j$, d.h. (Übung 29 im zweiten Schritt)

$$f(Y) \subseteq f\big(\bigcup_{j\in F} U_j\big) = \bigcup_{j\in F} f(U_j) = \bigcup_{j\in F} f(f^{-1}(U'_j)) \subseteq \bigcup_{j\in F} U'_j.$$

  Somit ist $\bigcup_{j\in F} U'_j$ eine endliche Teilüberdeckung von $f(Y)$. □

[83] Beachte: Urbilder offener/abgeschlossener Teilmengen unter stetigen Abbildungen sind offen/abgeschlossen. „Bei Kompakta ist die Richtung umgekehrt."

Satz 80. *Seien* (X, d) *und* (X', d') *metrische Räume, und* $f\colon X \to X'$ *eine stetige Abbildung. Ist* (X, d) *kompakt und* f *bijektiv, so ist* $f^{-1}\colon X' \to X$ *stetig.*

Terminologie: Eine bijektive stetige Abbildung zwischen metrischen Räumen, deren Umkehrabbildung ebenfalls stetig ist, wird als Homöomorphismus bezeichnet.

Beweis. Wegen Übung 80.a) müssen wir nur nachweisen, dass gilt:

$$A \subseteq X \text{ abgeschlossen in } (X, \mathrm{d}) \qquad \Longrightarrow \qquad \left(f^{-1}\right)^{-1}(A) = f(A) \subseteq X' \text{ abgeschlossen in } (X', \mathrm{d}').$$

(In der Gleichheit auf der echten Seite, haben wir die Bijektivität von f benutzt.) Sei nun $A \subseteq X$ abgeschlossen in (X, d):

- A ist kompakt wegen Lemma 67. Daher ist $f(A)$ gemäß Lemma 69 kompakt in (X', d').
- Wegen Lemma 66 ist $f(A)$ abgeschlossen in (X', d'). □

Übung 138. *Seien* (X, d) *und* (X', d') *metrische Räume, und* $f\colon X \to X'$ *eine stetige Abbildung. Sei* (X, d) *kompakt und* f *injektiv. Folgern Sie aus Satz 80, dass* f *eine* Einbettung *ist, d.h., dass die Koeinschränkung* $f|^{\mathrm{im}(f)}\colon X' \to \mathrm{im}(f)$ *– aufgefasst als Abbildung zwischen den metrischen Räumen* (X, d) *und* $\left(\mathrm{im}(f), \mathrm{d}'_{\mathrm{im}(f)}\right)$ *– ein Homöomorphismus ist.*

Im nächsten Abschnitt werden wir sehen, dass die abgeschlossenen Intervallen $[a, b] \subseteq \mathbb{R}$ mit $a \leq b$ kompakte Teilmengen von $\mathbb{R}$ sind. Der folgende Satz verallgemeinert daher unseren Satz 34:

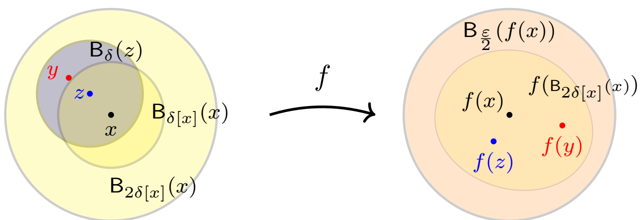
Satz 81. *Seien* (X, d) *und* (X', d') *metrische Räume, und* $f\colon X \to X'$ *eine Abbildung. Ist* f *stetig und* (X, d) *kompakt, so ist* f *gleichmäßig stetig.*

Beweis. Sei $\varepsilon > 0$ vorgegeben. Wir müssen zeigen, dass ein $\delta > 0$ existiert, sodass $f(\mathsf{B}_\delta(z)) \subseteq \mathsf{B}_\varepsilon(f(z))$ für alle $z \in X$ gilt (vgl. (197)):

- Da f stetig ist, existiert zu jedem $x \in X$ ein $\delta[x] > 0$ mit $f(\mathsf{B}_{2\delta[x]}(x)) \subseteq \mathsf{B}_{\frac{\varepsilon}{2}}(f(x))$.
- $(\mathsf{B}_{\delta[x]}(x))_{x \in X}$ ist eine offene Überdeckung von X, hat also eine endliche Teilüberdeckung $(\mathsf{B}_{\delta[x]}(x))_{x \in F}$. Wir setzen $\delta := \min(\delta[x] \mid x \in F)$.

Für $z \in X$ vorgegeben, gilt $z \in \mathsf{B}_{\delta[x]}(x)$ für ein $x \in F$ (Überdeckungseigenschaft). Wir erhalten $\mathsf{B}_\delta(z) \subseteq \mathsf{B}_{2\delta[x]}(x)$ wegen

$$\mathrm{d}(y, x) \leq \mathrm{d}(y, z) + \mathrm{d}(z, x) < \delta + \delta[x] \leq 2\delta[x] \qquad \forall\, y \in \mathsf{B}_\delta(z).$$



Hiermit folgt $f(\mathsf{B}_\delta(z)) \subseteq \mathsf{B}_\varepsilon(f(z))$, wegen $\qquad \left(z, y \in \mathsf{B}_{2\delta[x]}(x)\right)$

$$\mathrm{d}(f(z), f(y)) \leq \mathrm{d}(f(z), f(x)) + \mathrm{d}(f(x), f(y)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \qquad \forall\, y \in \mathsf{B}_\delta(z).$$ □

12.2 Folgenkompaktheit

Die Definition der Kompaktheit durch die Überdeckungseigenschaft ist nicht so leicht intuitiv zu erfassen. Wir diskutieren nun eine alternative Definition, die in Termen von Folgen gegeben ist.

Definition 70. *Ein metrischer Raum* (X, d) *heißt* folgenkompakt, *wenn jede Folge* $(x_n)_{n\in\mathsf{N}}$ *in* X *eine in* X *konvergente Teilfolge besitzt.*

Beispiel 92. *Sei* $\mathbb{R} \ni a < b \in \mathbb{R}$. *Wegen Korollar 18 ist* $[a, b] \subseteq \mathbb{R}$ *(ausgestattet mit der eingeschränkten Betragsmetrik) folgenkompakt.*

Bemerkung 83. *Sei* V *ein* $\mathbb{K}$*-Vektorraum mit* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, *und seien* $\|\cdot\|_1 \sim_{\mathrm{N}} \|\cdot\|_2$ *zueinander äquivalente Normen auf* V. *Sei weiterhin* $Y \subseteq V$ *eine Teilmenge, sowie*

$$\mathrm{d}_1 := (\mathrm{d}_{\|\cdot\|_1})_Y = \mathrm{d}_{\|\cdot\|_1}|_{Y\times Y} \qquad \textit{und} \qquad \mathrm{d}_2 := (\mathrm{d}_{\|\cdot\|_2})_Y = \mathrm{d}_{\|\cdot\|_2}|_{Y\times Y}$$

die, zu den Normmetriken $\mathrm{d}_{\|\cdot\|_1}$ *und* $\mathrm{d}_{\|\cdot\|_2}$ *gehörigen Unterraummetriken.*

Beachte: (Y, d_1) *ist genau dann folgenkompakt, wenn* (Y, d_2) *folgenkompakt ist.*

Beweis der Behauptung. Wegen Bemerkung 78.1) hat eine gegebene Folgen in Y eine bezüglich d_1 konvergente Teilfolge genau dann, wenn sie eine bezüglich d_2 konvergente Teilfolge hat. □

12.2.1 Eigenschaften Folgenkompakter Mengen

Lemma 70. *Ein folgenkompakter metrischer Raum ist vollständig.*

Beweis. Ist $(x_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge im folgenkompakten metrischen Raum (X, d), so existiert eine Teilfolge, die gegen ein $x \in X$ konvergiert. Wegen Lemma 35.3) gilt dann ebenfalls $(x_n)_{n\in\mathsf{N}} \to x$. □

Beispiel 93. *Das rationale Intervall* $[0, 2] \subseteq \mathbb{Q}$ *(ausgestattet mit der eingeschränkten Betragsmetrik) ist* nicht *folgenkompakt:*

- *Wegen Übung 107 existiert eine Folge* $(q_n)_{n\in\mathbb{N}}$ *in* $[0, 2] \subseteq \mathbb{Q}$, *die im übergeordneten metrischen Raum* $(\mathbb{R}, \mathrm{d}_{|\cdot|})$ *gegen* $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$ *konvergiert; was dann wegen Korallar 16 ebenfalls für jede Teilfolge von* $(q_n)_{n\in\mathbb{N}}$ *gilt.*
- *Hiermit kann auch keine Teilfolge von* $(q_n)_{n\in\mathbb{N}}$ *gegen ein Element aus* $\mathbb{Q}$ *konvergieren, also auch nicht in* $[0, 2] \subseteq \mathbb{Q}$.

Die Unvollständigkeit von $\mathbb{Q}$ *überträgt sich sozusagen auf den (zwar in* $\mathbb{Q}$ *abgeschlossenen, aber eben nicht kompakten) Teilraum* $[0, 2] \subseteq \mathbb{Q}$.

Lemma 71. *Sei* (X, d) *ein folgenkompakter metrischer Raum. Dann ist* X *beschränkt, d.h., es existiert ein* $C \geq 0$ *mit* $\mathrm{d}(x, y) \leq C$ *für alle* $x, y \in X$.

Beweis. Sei X unbeschränkt, und $z \in X$ fixiert (der Fall $X = \emptyset$ ist trivial):

- Für jedes $n \in \mathbb{N}$, existiert ein $z_n \in X$ mit $\mathrm{d}(z, z_n) \geq n$. In der Tat, andernfalls erhalten wir

  $$\mathrm{d}(x, y) \leq \mathrm{d}(x, z) + \mathrm{d}(z, y) \leq 2n \qquad \forall\, x, y \in X$$

  im Widerspruch zur Annahme, dass X unbeschränkt ist.
- Für $x \in X$ fixiert, liefert die inverse Dreiecksungleichung

  $$\mathrm{d}(x, z_n) \overset{\mathsf{M4)}}{\geq} |\mathrm{d}(x, z) - \mathrm{d}(z, z_n)| \qquad \forall\, n \in \mathbb{N}.$$

  Wir wählen $N \in \mathbb{N}$ mit $N \geq 1 + \mathrm{d}(x, z)$, und erhalten mit unserer Wahl von z_n:

  $$\mathrm{d}(x, z_n) \geq |\mathrm{d}(x, z) - n| \geq 1 \qquad \forall\, n \geq N.$$

  Dies zeigt, dass weder $(z_n)_{n\in\mathbb{N}}$ noch irgendeine Teilfolge von $(z_n)_{n\in\mathbb{N}}$ gegen x konvergieren kann. Da $x \in X$ beliebig war, widerspricht dies der Annahme, dass (X, d) folgenkompakt ist. □

12.2.2 Äquivalenz von Kompaktheit und Folgenkompaktheit

Der nächste Satz (Satz 82) zeigt, dass die Begriffe Kompakt und Folgenkompakt für metrische Räume äquivalent zueinander sind.[84] Dies liefert neue Möglichkeiten Kompaktheit zu verifizieren. Insbesondere erlaubt dies aber auch, Aussagen für folgenkompakte metrische Räume (wie Lemma 71) auf kompakte Teilmengen von metrischen Räumen zu übertragen (vgl. Korollar 51, Korollar 53 und Korollar 52).

Wir erinnern zunächst an den Begriff des Häufungspunktes:

Bemerkung 84. *Sei (X, d) ein metrischer Raum, $x \in X$, und $(x_n)_{n\in\mathsf{N}}$ eine Folge in X.*

- *Gemäß Definition 38 heißt x ein Häufungspunkt von $(x_n)_{n\in\mathsf{N}}$ genau dann, wenn eine Teilfolge von $(x_n)_{n\in\mathsf{N}}$ existiert, die gegen x konvergiert.*
- *Der gleiche Beweis wie in Satz 16.4) zeigt, dass x genau dann ein Häufungspunkt von $(x_n)_{n\in\mathsf{N}}$ ist, wenn jede Umgebung von x unendliche viele Folgenglieder von $(x_n)_{n\in\mathsf{N}}$ enthält.*

Satz 82. *Ein metrischer Raum (X, d) ist genau dann kompakt, wenn er folgenkompakt ist.*

Beweis.

- Sei (X, d) kompakt. Ist (X, d) nicht folgenkompakt, so existiert eine Folge $(x_n)_{n\in\mathsf{N}}$ in X, die keine konvergente Teilfolge besitzt, d.h., kein $x \in X$ ist ein Häufungspunkt von $(x_n)_{n\in\mathsf{N}}$:
  - Wegen Bemerkung 84 existiert zu jedem $x \in X$ eine offene[85] Umgebung U_x von x, die nur endlich viele Folgenglieder von $(x_n)_{n\in\mathsf{N}}$ enthält (es ist also $\{n \in \mathsf{N} \mid x_n \in U_x\} \subseteq \mathsf{N}$ endlich).
  - $(U_x)_{x\in X}$ ist eine offene Überdeckung von X, hat also eine endliche Teilüberdeckung $(U_x)_{x\in F}$.
  - $\bigcup_{x\in F} U_x$ enthält per konstruktionem nur endlich Folgenglieder von $(x_n)_{n\in\mathsf{N}}$ (es ist also $\{n \in \mathsf{N} \mid x_n \in \bigcup_{x\in F} U_x\} \subseteq \mathsf{N}$ endlich). Wegen $X \subseteq \bigcup_{x\in F} U_x$, widerspricht dies der Unendlichkeit der Indexmenge N.

- Sei (X, d) folgenkompakt. Ist (X, d) nicht kompakt, so existiert eine offene Überdeckung $(U_j)_{j\in J}$ von X, die keine endliche Teilüberdeckung hat. Für jedes $x \in X$, setzen wir

$$\varepsilon[x] := \sup_{\mathbb{R}}(\underbrace{\{0 < r \leq 1 \mid \exists\, j \in J \text{ mit } \mathsf{B}_{2r}(x) \subseteq U_j\}}_{=: B[x]}) \in (0, 1]. \tag{362}$$

  Beachte: Die Menge $B[x]$ ist offensichtlich nach oben beschränkt; und weiterhin nichtleer, denn wegen der Überdeckungseigenschaft von $(U_j)_{j\in J}$ gilt $x \in U_j$ für ein $j \in J$, also auch $\mathsf{B}_\varepsilon(x) \subseteq U_j$ für ein $\varepsilon > 0$ wegen der Offenheit von U_j.

  Gemäß Lemma 20 existiert zu jedem $x \in X$ ein $B[x] \ni r[x] > \varepsilon[x] - \frac{\varepsilon[x]}{2} = \frac{\varepsilon[x]}{2}$; und somit auch ein $j[x] \in J$ mit $\mathsf{B}_{\varepsilon[x]}(x) \subseteq \mathsf{B}_{2r[x]}(x) \subseteq U_{j[x]}$ (beachte $\varepsilon[x] < 2r[x]$).

  - Per Annahme hat $(U_j)_{j\in J}$ keine endliche Teilüberdeckung von X. Also gilt $X \setminus \bigcup_{\ell=0}^{n} U_{j[x_\ell]} \neq \emptyset$ für jedes $n \in \mathbb{N}$ und jede Wahl $x_0, \ldots, x_n \in X$.
  - Sei $x_0 \in X$ fixiert. Wir finden induktiv eine Folge $(x_n)_{n\in\mathbb{N}}$ in X mit $x_n \in X \setminus \bigcup_{\ell=0}^{n-1} U_{j[x_\ell]}$ für alle $n \in \mathbb{N}_{>0}$.
  - Für $n, m \in \mathbb{N}$ mit $n > m$, erhalten wir

$$x_n \in X \setminus \textstyle\bigcup_{\ell=0}^{n-1} U_{j[x_\ell]} \overset{m\leq n-1}{\subseteq} X \setminus U_{j[x_m]} \subseteq X \setminus \mathsf{B}_{2r[x_m]}(x_m) \subseteq X \setminus \mathsf{B}_{\varepsilon[x_m]}(x_m),$$

  also $x_n \notin \mathsf{B}_{\varepsilon[x_m]}(x_m)$, und daher $\varepsilon[x_m] \leq \mathrm{d}(x_n, x_m)$.

[84] In beliebigen topologischen Räumen (Folgevorlesungen) ist dies nicht der Fall. Hier sind Kompaktheit und Folgenkompaktheit üblicherweise zueinander disjunkte Begrifflichkeiten.

[85] Wir dürfen U_x als offen annehmen, da jede Umgebung von x eine offene Umgebung von x (sogar einen offenen ε-Ball um x) enthält.

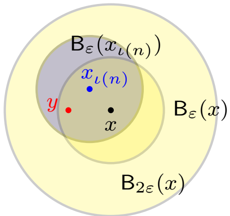
Da (X, d) folgenkompakt ist, besitzt $(x_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge $(x_{\iota(n)})_{n\in\mathbb{N}} \to x \in X$. Wegen Lemma 35.1) ist $(x_{\iota(n)})_{n\in\mathbb{N}}$ eine Cauchy-Folge, und wir erhalten mit obiger Abschätzung:

$$0 \le \varepsilon[x_{\iota(n)}] \le \mathrm{d}(x_{\iota(n+1)}, x_{\iota(n)}) \qquad \forall\, n \in \mathbb{N} \qquad \Longrightarrow \qquad \lim_n \varepsilon[x_{\iota(n)}] = 0. \tag{363}$$

- Es existiert ein $j \in J$ und ein $\varepsilon > 0$ mit $\mathsf{B}_{2\varepsilon}(x) \subseteq U_j$. $\quad ((U_j)_{j\in J}$ offene Überdeckung von $X)$
- Es existiert ein $N_\varepsilon \in \mathbb{N}$ mit $x_{\iota(n)} \in \mathsf{B}_\varepsilon(x)$ für alle $n \ge N_\varepsilon$. $\quad$ (beachte $(x_{\iota(n)})_{n\in\mathbb{N}} \to x$)
- Für $n \ge N_\varepsilon$ gilt:

$$y \in \mathsf{B}_\varepsilon(x_{\iota(n)}) \qquad \Longrightarrow \qquad \mathrm{d}(x, y) \le \mathrm{d}(x, x_{\iota(n)}) + \mathrm{d}(x_{\iota(n)}, y) < 2\varepsilon \qquad \Longrightarrow \qquad y \in \mathsf{B}_{2\varepsilon}(x),$$

also $\mathsf{B}_\varepsilon(x_{\iota(n)}) \subseteq \mathsf{B}_{2\varepsilon}(x) \subseteq U_j$. Gemäß (362) gilt somit $\frac{\varepsilon}{2} \in B[x_{\iota(n)}]$, also $\varepsilon[x_{\iota(n)}] \ge \frac{\varepsilon}{2}$.



Mit Korollar 14, erhalten wir den Widerspruch $0 \overset{(363)}{=} \lim_n \varepsilon[x_{\iota(n)}] \ge \frac{\varepsilon}{2} > 0$. $\square$

12.2.3 Zusammenfassung: Eigenschaften von Kompakta

Korollar 51. *Ein kompakter metrischer Raum ist vollständig und beschränkt.*

Beweis. Klar wegen Satz 82, sowie Lemma 70 und Lemma 71. $\square$

Korollar 52. *Sei (X, d) ein metrischer Raum, und $Y \subseteq X$ eine Teilmenge. Dann gilt:*

$$Y \text{ kompakt in } X \quad \overset{\text{Lemma 68}}{\Longleftrightarrow} \quad (Y, \mathrm{d}_Y) \text{ kompakt} \quad \overset{\text{Satz 82}}{\Longleftrightarrow} \quad (Y, \mathrm{d}_Y) \text{ folgenkompakt.}$$

Beachte: *Die letzte Bedingung ist gleichbedeutend damit, dass jede Folge in Y eine Teilfolge besitzt, die im metrischen Raum (X, d) gegen ein $y \in Y$ konvergiert (Definition von d_Y).*

Korollar 53. *Sei (X, d) ein metrischer Raum. Ist $Y \subseteq X$ kompakt, so auch beschränkt und abgeschlossen. Weiterhin ist (Y, d_Y) vollständig.*

Beweis. • Wegen Lemma 67 ist Y abgeschlossen in (X, d).

- Wegen Korollar 52 ist (Y, d_Y) folgenkompakt:
  - Wegen Lemma 70 ist (Y, d_Y) vollständig.
  - Lemma 71 zeigt, dass (Y, d_Y) beschränkt ist, dass also ein $C \ge 0$ existiert mit

$$\mathrm{d}(x, y) = \mathrm{d}_Y(x, y) \le C \qquad \forall\, x, y \in Y. \qquad \square$$

Beispiel 94. *Wir betrachten den metrischen Raum $(\mathbb{R}^n, \mathrm{d}_{\|\cdot\|})$, mit $n \ge 1$ und einer Norm $\|\cdot\|$ auf $\mathbb{R}^n$. Für $c \ge 0$, ist $\mathcal{Q}_c = [-c, c]^n$ eine kompakte Teilmenge von $(\mathbb{R}^n, \mathrm{d}_{\|\cdot\|})$, bzw. selbst kompakt als metrischer (Unter)Raum $\big(\mathcal{Q}_c, (\mathrm{d}_{\|\cdot\|})_{\mathcal{Q}_c}\big)$. Daher wird $\mathcal{Q}_c$ auch als* kompakter Quader *in $\mathbb{R}^n$ bezeichnet.*

Beweis. • Wegen Lemma 65 ist $(\mathcal{Q}_c, (\mathrm{d}_{\|\cdot\|_\infty})_{\mathcal{Q}_c})$ folgenkompakt, also kompakt wegen Korollar 52.

- Wegen Satz 79 $(\|\cdot\| \sim_{\mathrm{N}} \|\cdot\|_\infty)$ und Bemerkung 82, ist $\big(\mathcal{Q}_c, (\mathrm{d}_{\|\cdot\|})_{\mathcal{Q}_c}\big)$ ebenfalls kompakt.
- Wegen Korollar 52 ist $\mathcal{Q}_c$ eine kompakte Teilmenge von $(\mathbb{R}^n, \mathrm{d}_{\|\cdot\|})$. $\square$

Im Falle $n = 1$, zeigen natürlich analog Beispiel 92 und Korollar 52, dass jedes Intervall $[a, b] \subseteq \mathbb{R}$ mit $\mathbb{R} \ni a \leq b \in \mathbb{R}$ eine kompakte Teilmenge von $(\mathbb{R}, \mathrm{d}_{|\cdot|})$ ist (bzw. $[a, b]$ mit der eingeschränkten Betragsmetrik ein kompakter metrischer Raum ist). Diese Intervalle werden daher auch als kompakte Intervalle *bezeichnet.*

***Beispiel 4.** *Das rationale Intervall $[0, 2] \subseteq \mathbb{Q}$ ist nicht kompakt, da nicht folgenkompakt wegen Beispiel 93. Wir können hier auch anders argumentieren:*

Wäre $Y := [0, 2] \subseteq \mathbb{Q}$ (mit der eingeschränkten Betragsmetrik) ein kompakter metrischer Raum, so wäre Y eine kompakte Teilmenge von $\mathbb{R}$ (natürlich auch von $\mathbb{Q}$) wegen Korollar 52, mithin abgeschlossen in $\mathbb{R}$ wegen Korollar 53. Wegen $0 < \sqrt{2} < 2$ hat nun aber gemäß Übung 107 jede Umgebung von $\sqrt{2} \in \mathbb{R} \setminus Y$ nichtleeren Schnitt mit Y, womit $\mathbb{R} \setminus Y$ nicht offen, also Y nicht abgeschlossen sein kann.

Wegen Beispiel 94 (Kompaktheit der Intervalle $[a, b]$), stellt der folgende Satz eine direkte Verallgemeinerung von Satz 35 (Satz vom Maximum) dar (gleicher Beweis).

Satz 83 (Allgemeiner Satz vom Maximum)**.** *Sei (X, d) ein kompakter metrischer Raum, und $f : X \to \mathbb{R}$ eine stetige Abbildung. Dann gilt*

$$\inf_{\mathbb{R}}(\mathrm{im}(f)) \in \mathrm{im}(f) \ni \sup_{\mathbb{R}}(\mathrm{im}(f)).$$

(Jede reellwertige stetige Funktion auf einem kompakten metrischen Raum, nimmt sowohl ihr Maximum als auch ihr Minimum an.)

Beweis. Wir zeigen nur $\sup_{\mathbb{R}}(\mathrm{im}(f)) \in \mathrm{im}(f)$ (die Aussage $\inf_{\mathbb{R}}(\mathrm{im}(f)) \in \mathrm{im}(f)$ folgt analog). Sei hierfür $o := \sup_{\overline{\mathbb{R}}}(\mathrm{im}(f)) \in \mathbb{R} \cup \{\infty\}$:

- Wir wählen eine Folge $(z_n)_{n\in\mathbb{N}}$ in $\mathrm{im}(f)$ mit $(z_n)_{n\in\mathbb{N}} \to o$. (Übung 64)
- Wir wählen eine Folge $(x_n)_{n\in\mathbb{N}}$ in X mit $f(x_n) = z_n$ für alle $n \in \mathbb{N}$.
- Wegen Satz 82 existiert eine konvergente Teilfolge $(x_{\iota(n)})_{n\in\mathbb{N}} \to x \in X$; und wir erhalten mit der Folgenstetigkeit von f (letzter Schritt) sowie Korollar 16 (zweiter Schritt):

$$o = \lim_n z_n = \lim_n z_{\iota(n)} = \lim_n f(x_{\iota(n)}) = f(x) \in \mathrm{im}(f). \qquad \square$$

Übung 139. *Sei (X, d) ein metrischer Raum, $f : X \to \mathbb{R}$ stetig, und $Y \subseteq X$ kompakt. Zeigen Sie*

$$\inf_{\mathbb{R}}(f(Y)) \in f(Y) \ni \sup_{\mathbb{R}}(f(Y)).$$

12.3 Der Satz von Heine-Borel

Sei $n \in \mathbb{N}_{>0}$ und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. In diesem Abschnitt wollen wir schließlich den Satz von Heine-Borel beweisen, der die kompakten Teilmengen des normierten Raumes $\mathbb{K}^n$ charakterisiert:

Satz 84 (Heine-Borel)**.** *Eine Teilmenge von $\mathbb{K}^n$ ist kompakt genau dann, wenn sie abgeschlossen und beschränkt ist.*

Die Aussage in Satz 84 ist wohlformuliert, da es wegen Satz 79 bei den verwendeten Begrifflichkeiten nicht auf die explizite Wahl der Norm auf $\mathbb{K}^n$ ankommt:

- Wegen Satz 79 und Bemerkung 82 ist der Kompaktheitsbegriff unabhängig von der expliziten Wahl der Norm auf $\mathbb{K}^n$.
- Wegen Satz 79 und Bemerkung 77 ist der Beschränktheitsbegriff unabhängig von der expliziten Wahl der Norm auf $\mathbb{K}^n$.

- Wegen Satz 79 und Korollar 49 ist der Offenheits- und somit auch der Abgeschlossenheitsbegriff unabhängig von der expliziten Wahl der Norm auf $\mathbb{K}^n$.

Bemerkung 85. *Satz 84 gilt für beliebige metrische Räume im allgemeinen nicht; noch nicht einmal für metrische Unterräume von $\mathbb{R}$:*

*Sei $X = (0,1) \subseteq \mathbb{R}$, sowie d die auf $(0,1)$ eingeschränkte Betragsmetrik $\mathrm{d}_{|\cdot|}$. Dann ist (X, d) ein (durch $C = 1$) beschränkter metrischer Raum, und X ist abgeschlossen in (X, d) wegen Korollar 13.*AM1). *Es ist nun aber $((\frac{1}{n}, 1))_{n\in\mathbb{N}_{>0}}$ eine Überdeckung von X, die keine endliche Teilüberdeckung besitzt; also kann (X, d) nicht kompakt sein.*

Beweis von Satz 84. Wegen Korollar 53 ist eine kompakte Teilmenge von $\mathbb{K}^n$ ist abgeschlossen und beschränkt. Wir zeigen die umgekehrte Richtung in zwei Teilschritten:

a) Eine abgeschlossene und beschränkte Teilmenge $Y \subseteq \mathbb{R}^n$ mit $n \in \mathbb{N}_{>0}$ ist kompakt:

   Beweis. Da Y beschränkt ist, gilt $\|Y\|_\infty \leq c$ für ein $c \geq 0$, also $Y \subseteq [-c, c]^n = \mathcal{Q}_c$. Wegen Beispiel 94 ist $\mathcal{Q}_c \subseteq \mathbb{R}^n$ kompakt; also ist Y als abgeschlossene Teilmenge einer kompakten Menge, wegen Lemma 67 ebenfalls kompakt. □

b) Eine abgeschlossene und beschränkte Teilmenge $Y \subseteq \mathbb{C}^n$ mit $n \in \mathbb{N}_{>0}$ ist kompakt:

   Beweis. Wir betrachten die zueinander inversen $\mathbb{R}$-linearen Abbildungen:

$$\begin{aligned} \alpha\colon \mathbb{R}^{2n} \to \mathbb{C}^n, \qquad & (x_1, \dots, x_n, y_1, \dots, y_n) \mapsto (x_1 + \mathrm{i}y_1, \dots, x_n + \mathrm{i}y_n) \\ \beta\colon \mathbb{C}^n \to \mathbb{R}^{2n}, \qquad & (x_1 + \mathrm{i}y_1, \dots, x_n + \mathrm{i}y_n) \mapsto (x_1, \dots, x_n, y_1, \dots, y_n). \end{aligned}$$

   Wir fixieren eine Norm $\|\cdot\|_\mathbb{C}$ auf $\mathbb{C}^n$, und setzen $\|\cdot\|_\mathbb{R} := \|\cdot\|_\mathbb{C} \circ \alpha$ (dies liefert genau die Norm aus (357)). Per Konstruktion gilt dann:

$$\begin{aligned} \|\beta(x) - \beta(y)\|_\mathbb{R} &= \|\beta(x-y)\|_\mathbb{R} = \|(\alpha \circ \beta)(x-y)\|_\mathbb{C} = \|x-y\|_\mathbb{C} && \forall\, x, y \in \mathbb{C}^n \\ \|\alpha(x) - \alpha(y)\|_\mathbb{C} &= \|\alpha(x-y)\|_\mathbb{C} = \|x-y\|_\mathbb{R} && \forall\, x, y \in \mathbb{R}^{2n}. \end{aligned}$$

   - Somit bilden α, β beschränkte Teilmengen auf beschränkte Teilmengen ab.
   - Weiterhin sind α, β Lipschitzstetig, also stetig. Wegen $\alpha^{-1} = \beta$ (bzw. $\beta = \alpha^{-1}$), sowie Proposition 9, Übung 80.a), Lemma 69, bilden α, β somit offene/abgeschlossene/kompakte Teilmengen auf offene/abgeschlossene/kompakte Teilmengen ab.

   Folglich ist $\beta(Y) \subseteq \mathbb{R}^{2n}$ abgeschlossen und beschränkt, also kompakt wegen Teil a). Also ist $Y = \alpha(\beta(Y)) \subseteq \mathbb{C}^n$ ebenfalls kompakt. □

Gemäß Korollar 51 ist ein kompakter metrischer Raum vollständig. Man mag sich also durchaus wundern, warum diese Bedingung in Satz 84 nicht mehr vorkommt. Der Grund hierfür ist die Vollständigkeit des $\mathbb{K}^n$, die wir ja in Korollar 50.2) gezeigt hatten. Wie das nächste Lemma zeigt, impliziert nämlich die Abgeschlossenheit einer Teilmenge in einem vollständigen metrischen Raum, bereits die Vollständigkeit des zugehörigen metrischen Unterraumes:

Lemma 72 (Kriterium für Vollständigkeit)**.** *Sei (X, d) ein vollständiger metrischer Raum und $Y \subseteq X$ eine Teilmenge. Der metrische Unterraum (Y, d_Y) ist genau dann vollständig, wenn Y abgeschlossen in (X, d) ist, d.h., es gilt:*

$$(X, \mathrm{d}) \text{ vollständig und } Y \subseteq X\colon \quad (Y, \mathrm{d}_Y) \text{ vollständig} \quad \Leftrightarrow \quad Y \text{ abgeschlossen in } (X, \mathrm{d}).$$

Wir erinnern an Lemma 36, welches besagt, dass eine Teilmenge $Y \subseteq X$ genau dann abgeschlossen in (X, d) ist, wenn sie folgenabgeschlossen ist, d.h., für jede Folge $(y_n)_{n\in\mathbb{N}}$ in Y mit $\lim_n y_n = x \in X$ gilt bereits $x \in Y$.

Beweis. • Sei Y abgeschlossen. Ist $(y_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge in (Y, d_Y), so auch in (X, d). Daher existiert der Grenzwert $\lim_n y_n = x \in X$. Dann gilt $x \in Y$, da Y wegen Lemma 36 folgenabgeschlossen ist.

- Sei (Y, d_Y) vollständig. Sei $(y_n)_{n\in\mathsf{N}}$ eine Folge in Y mit $\lim_n y_n = x \in X$ bezüglich d.
  - Dann ist $(y_n)_{n\in\mathsf{N}}$ eine Cauchy-Folge in (X, d), also auch in (Y, d_Y) wegen $\{y_n \mid n \in \mathsf{N}\} \subseteq Y$.
  - Per Annahme konvergiert $(y_n)_{n\in\mathsf{N}}$ bezüglich d_Y gegen ein $y \in Y$, also auch bezüglich d.
  - Wegen der Eindeutigkeit des Grenzwertes, gilt dann bereits $x = y \in Y$.

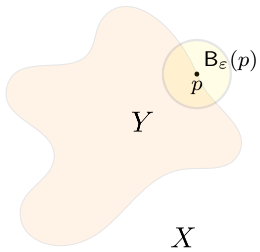
  Somit ist Y folgenabgeschlossen, also abgeschlossen wegen Lemma 36. □

12.4 Weitere Begrifflichkeiten: Abschluss, Inneres, und der Rand

In diesem Abschnitt wollen wir noch einige Grundbegriffe im Rahmen metrischer Räume diskutieren, die auch im allgemeinen Rahmen der sogenannten topologischen Räume (Folgevorlesungen) auftauchen, und eine zentrale Rolle spielen. Im Folgenden sei (X, d) ein metrischer Raum; und es bezeichne $\mathcal{O}(X)$ die Menge aller offenen Teilmengen von X, sowie $\mathcal{A}(X)$ die Menge aller abgeschlossenen Teilmengen von X (vgl. Notation 20).

Bemerkung 86 (Berührungspunkte)**.** *Sei (X, d) ein metrischer Raum, und $Y \subseteq X$ eine Teilmenge. Wir erinnern an die Definition der Menge aller Berührungspunkte von Y:*

$$\mathcal{B}(Y) = \big\{\, p \in X \;\big|\; Y \cap \mathsf{B}_\varepsilon(p) \neq \emptyset \;\; \textit{für alle} \;\; \varepsilon > 0 \,\big\}.$$



Bemerkung 53.b)) besagte:

$$Y \subseteq \mathcal{B}(Y) \qquad \wedge \qquad p \in \mathcal{B}(Y) \quad \Leftrightarrow \quad \exists \; \textit{Folge } (y_n)_{n\in\mathsf{N}} \textit{ in } Y \textit{ mit } \lim_n y_n = p. \tag{364}$$

Die folgende etwas abstraktere Charakterisierung wird auch im Rahmen topologischer Räume benutzt:

$$\begin{aligned} \mathcal{B}(Y) &= \overbrace{\big\{\, p \in X \,\big|\, Y \cap U \neq \emptyset \;\; \textit{für jede Umgebung } U \textit{ von } p \,\big\}}^{=:\, \mathcal{B}_{\mathsf{u}}(Y)} \\ &= \underbrace{\big\{\, p \in X \,\big|\, Y \cap O \neq \emptyset \;\; \textit{für jede offene Umgebung } O \textit{ von } p \,\big\}}_{=:\, \mathcal{B}_{\mathsf{o}}(Y)} \end{aligned} \tag{365}$$

Beweis von (365)*.* Wir zeigen $\mathcal{B}(Y) \subseteq \mathcal{B}_{\mathsf{u}}(Y) \subseteq \mathcal{B}_{\mathsf{o}}(Y) \subseteq \mathcal{B}(Y)$:

- $\mathcal{B}(Y) \subseteq \mathcal{B}_{\mathsf{u}}(Y)$: Sei $p \in \mathcal{B}(Y)$ und U eine Umgebung von p. Dann existiert ein $\varepsilon > 0$ mit $\mathsf{B}_\varepsilon(p) \subseteq U$. Per Annahme gilt $\emptyset \neq Y \cap \mathsf{B}_\varepsilon(p) \subseteq Y \cap U$. Dies zeigt $p \in \mathcal{B}_{\mathsf{u}}(Y)$.
- $\mathcal{B}_{\mathsf{u}}(Y) \subseteq \mathcal{B}_{\mathsf{o}}(Y)$: Eine offene Umgebung von $p \in X$ ist insbesondere eine Umgebung von p.
- $\mathcal{B}_{\mathsf{o}}(Y) \subseteq \mathcal{B}(Y)$: Für jedes $\varepsilon > 0$ und $p \in X$ ist $\mathsf{B}_\varepsilon(p)$ eine offene Umgebung von p. □

Terminologie 30. *Sei (X, d) ein metrischer Raum, und $Y \subseteq X$ eine Teilmenge.*

1) Der Abschluss von Y in (X, d), ist definiert durch

$$\overline{Y} := \bigcap\nolimits_{Y \subseteq A \in \mathcal{A}(X)} A.$$

$\overline{Y}$ ist die kleinste abgeschlossene Teilmenge von X, die Y enthält:

- *$Y \subseteq \overline{Y}$ (offensichtlich),*
- *$\overline{Y} \in \mathcal{A}(X)$ ist abgeschlossen (Korollar 13.AM3)), und es gilt:*

$$Y \subseteq A \in \mathcal{A}(X) \qquad \Longrightarrow \qquad \overline{Y} \subseteq A\,. \tag{366}$$

Insbesondere gilt $\overline{Y} = Y$ genau dann, wenn $Y \in \mathcal{A}(X)$ abgeschlossen ist (Übung).

2) Das Innere von Y in (X, d), ist definiert durch

$$Y^\circ := \bigcup\nolimits_{\mathcal{O}(X) \ni O \subseteq Y} O.$$

Y° ist die größte offene Menge, die in Y enthalten ist:

- *$Y^\circ \subseteq Y$ (offensichtlich),*
- *$Y^\circ \in \mathcal{O}(X)$ ist offen (Proposition 3.OM3)), und es gilt:*

$$\mathcal{O}(X) \ni O \subseteq Y \qquad \Longrightarrow \qquad O \subseteq Y^\circ. \tag{367}$$

Insbesondere gilt $Y^\circ = Y$ genau dann, wenn $Y \in \mathcal{O}(X)$ offen ist (Übung).

3) Der Rand von Y in (X, d) ist definiert durch

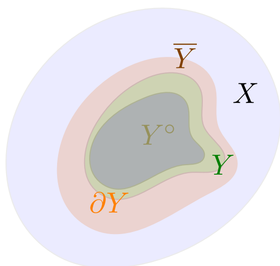
$$\partial Y := \overline{Y} \setminus Y^\circ \qquad \textit{also} \qquad \partial Y \subseteq \overline{Y} \quad \wedge \quad Y^\circ \cap \partial Y = \emptyset. \tag{368}$$

$\partial Y \in \mathcal{A}(X)$ ist abgeschlossen wegen Lemma 29 (als Differenz einer offenen mit einer abgeschlossenen Teilmenge).

(Beachte: $Y = \partial Y \,\dot\cup\, Y^\circ$, wenn Y abgeschlossen ist ($\overline{Y} = Y$).)

Zusammenfassend haben wir:

$$Y^\circ \subseteq Y \subseteq \overline{Y} = \partial Y \,\dot\cup\, Y^\circ \qquad \text{also} \qquad Y = (Y \cap \partial Y) \,\dot\cup\, Y^\circ.$$



D.h., Y ist die disjunkte Vereinigung von Y° mit der Teilmenge $Y \cap \partial Y \subseteq \partial Y$ – Letztere kann durchaus auch ganz ∂Y sein ($\mathcal{A}(X) \ni Y = \overline{Y}$), oder aber auch leer ($\mathcal{O}(X) \ni Y = Y^\circ$).[86]

Wir haben die folgenden Charakterisierungen der Mengen in Terminologie 30:

Satz 85. *Sei (X, d) ein metrischer Raum, und $Y \subseteq X$ eine Teilmenge. Es gilt:*

[86] Aus obiger Grafik heraus könnte man vermuten, dass ∂Y durchaus einen offenen metrischen Ball enthalten könnte, welcher leeren Schnitt mit Y (dieser wäre komplett im hellroten Bereich enthalten) oder mit $X \setminus Y$ (dieser wäre komplett im hellgrünen Bereich enthalten) hat. Satz 85.(3) zeigt allerdings, dass dies nicht der Fall ist – „Der Rand ist in obiger Grafik sozusagen viel zu dick gemalt“.

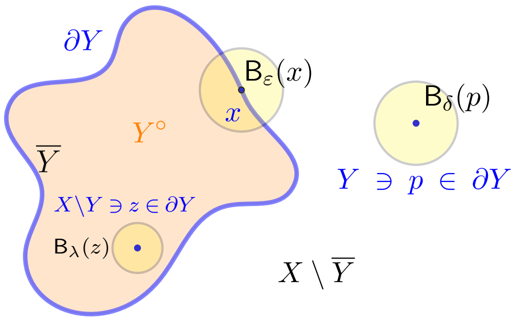
(1) $\overline{Y} \;=\; \mathcal{B}(Y) \qquad\qquad \big(\overset{(365)}{=} \{\, p \in X \,|\, Y \cap U \neq \emptyset \;\; \textit{für jede Umgebung } U \textit{ von } p \,\}\big)$

(2) $Y^\circ = \{\, p \in X \,|\, Y \textit{ ist Umgebung von } p \,\} \overset{\text{def.}}{=} \{\, p \in X \,|\, \mathsf{B}_\varepsilon(p) \subseteq Y \;\; \textit{für ein } \varepsilon > 0 \,\}$

(3)
$$\begin{aligned}
\partial Y &= \{\, p \in X \,|\; Y \cap U \neq \emptyset \neq (X \setminus Y) \cap U \quad \textit{für jede Umgebung } U \textit{ von } p \,\} &&=: \partial_{\mathsf{u}} Y \\
&= \{\, p \in X \,|\; Y \cap O \neq \emptyset \neq (X \setminus Y) \cap O \quad \textit{für jede offene Umgebung } O \textit{ von } p \,\} &&=: \partial_{\mathsf{o}} Y \\
&= \{\, p \in X \,|\, Y \cap \mathsf{B}_\varepsilon(p) \neq \emptyset \neq (X \setminus Y) \cap \mathsf{B}_\varepsilon(p) \quad \textit{für jedes } \varepsilon > 0 \,\} &&=: \partial_\epsilon Y
\end{aligned}$$

Es gilt also:

$$p \in \partial Y \quad \textit{für} \quad p \in X \qquad \Leftrightarrow \qquad Y \cap U \neq \emptyset \neq (X \setminus Y) \cap U \quad \textit{für jede Umgebung } U \textit{ von } p.$$



Die Blaue Menge ∂Y kann sowohl Elemente von Y als auch von $X \setminus Y$ enthalten;
und nicht jeder metrische Ball um einen Randpunkt schneidet notwendig Y°.

Beachte: Die obige Grafik deutet an, dass nicht jeder metrische Ball um ein $p \in \partial Y$ zwangsläufig nichtleeren Schnitt mit Y° hat. Sei beispielsweise $(X, \mathrm{d}) = (\mathbb{R}, \mathrm{d}_{|\cdot|})$:

- $Y = \mathbb{Q}$: Es gilt $\partial Y = \mathbb{R} = \overline{Y}$, also $Y^\circ = \emptyset$ (Übung).
- $Y = [0,1) \cup \{2\}$: Es gilt $Y^\circ = (0,1)$, $\overline{Y} = [0,1] \cup \{2\}$ und $\partial Y = \{0,1,2\}$ (Übung).

Beweis von Satz 85. (1) Wir verwenden $\mathcal{B}(Y) = \mathcal{B}_{\mathsf{o}}(Y)$:

- $\overline{Y} \subseteq \mathcal{B}(Y)$: Wegen (366) genügt es $Y \subseteq \mathcal{B}(Y) \in \mathcal{A}(X)$ nachzuweisen. Nun gilt $Y \subseteq \mathcal{B}(Y)$ wegen (364); und zudem ist $\mathcal{B}(Y)$ abgeschlossen, denn $X \setminus \mathcal{B}(Y)$ ist offen:

  Wir zeigen, dass $X \setminus \mathcal{B}(Y)$ mit jedem p auch eine offene Umgebung von p enthält, womit dann $X \setminus \mathcal{B}(Y)$ als Vereinigung offener Mengen offen ist. Hierfür beachten wir:

$$\begin{aligned}
X \setminus \mathcal{B}(Y) &\overset{(365)}{=} X \setminus \mathcal{B}_{\mathsf{o}}(Y) \\
&= X \setminus \{\, p \in X \,|\, Y \cap O \neq \emptyset \;\; \text{für jede offene Umgebung } O \text{ von } p \,\} \\
&= \{\, p \in X \,|\, Y \cap O = \emptyset \;\; \text{für eine offene Umgebung } O \text{ von } p \,\} \\
&= \{\, p \in X \,|\, \exists\; O \subseteq X \setminus Y \text{ offen mit } p \in O \}. \qquad (369)
\end{aligned}$$

  Sei $p \in X \setminus \mathcal{B}(Y)$ und $O \subseteq X \setminus Y$ offen mit $p \in O$. Dann gilt bereits $O \subseteq X \setminus \mathcal{B}(Y)$, wegen:

$$q \in O \quad \Longrightarrow \quad O \subseteq X \setminus Y \;\text{ offen mit }\; q \in O \quad \overset{(369)}{\Longrightarrow} \quad q \in X \setminus \mathcal{B}(Y).$$

- $\mathcal{B}(Y) \subseteq \overline{Y}$: Andernfalls existiert ein $p \in \mathcal{B}(Y) \setminus \overline{Y} \subseteq X \setminus \overline{Y} =: O$, wobei O wegen der Abgeschlossenheit von $\overline{Y}$ offen ist. Wir erhalten den Widerspruch $\emptyset = Y \cap O \neq \emptyset$:
  – $\emptyset = \overline{Y} \cap O \supseteq Y \cap O$: Es gilt $Y \subseteq \overline{Y}$.

– $\emptyset \neq Y \cap O$: Es gilt $p \in \mathcal{B}(Y) = \mathcal{B}_{\mathsf{o}}(Y)$, und O ist eine offene Umgebung von p.

(2) Die zweite Gleichheit gilt per Definition; denn Y ist eine Umgebung von $p \in X$ genau dann, wenn Y einen offenen metrischen Ball um p enthält. Wir bleibt daher $Y^\circ = O := \{\, p \in X \mid \mathsf{B}_\varepsilon(p) \subseteq Y \text{ für ein } \varepsilon > 0 \,\}$ nachweisen:

- $O \subseteq Y^\circ$: Wegen Lemma 27 sind die offenen metrischen Bälle offen. Also zeigt (367):

$$p \in O \quad \overset{\text{def.}}{\Longrightarrow} \quad \mathsf{B}_\varepsilon(p) \subseteq Y \text{ für ein } \varepsilon > 0 \quad \overset{(367)}{\Longrightarrow} \quad p \in \mathsf{B}_\varepsilon(p) \subseteq Y^\circ.$$

- $Y^\circ \subseteq O$: Y° ist offen mit $Y^\circ \subseteq Y$. Für $p \in Y^\circ$, existiert daher $\varepsilon > 0$ mit $\mathsf{B}_\varepsilon(p) \subseteq Y^\circ \subseteq Y$ (Y° ist Umgebung aller ihrer Punkte, da offen).

(3) Wie in Bemerkung 86 folgt $\partial_\epsilon Y \subseteq \partial_{\mathsf{u}} Y \subseteq \partial_{\mathsf{o}} Y \subseteq \partial_\epsilon Y$, also $\partial_{\mathsf{u}} Y = \partial_{\mathsf{o}} Y = \partial_\epsilon Y$.

- $\partial Y \subseteq \partial_\epsilon Y$: Sei $p \in \partial Y$ und $\varepsilon > 0$.
  - Wegen $p \in \partial Y \subseteq \overline{Y}$ zeigt Teil (1) ($\overline{Y} = \mathcal{B}(Y)$), dass $Y \cap \mathsf{B}_\varepsilon(p) \neq \emptyset$ gilt.
  - Wegen $p \notin Y^\circ$ zeigt Teil (2), dass $\mathsf{B}_\varepsilon(p) \not\subseteq Y$ gilt, d.h., $(X \setminus Y) \cap \mathsf{B}_\varepsilon(p) \neq \emptyset$.
- $\partial_\epsilon Y \subseteq \partial Y$: Sei $p \in \partial_\epsilon Y$, also $Y \cap \mathsf{B}_\varepsilon(p) \neq \emptyset \neq (X \setminus Y) \cap \mathsf{B}_\varepsilon(p)$ für alle $\varepsilon > 0$.
  - Es gilt $p \in \mathcal{B}(Y) = \overline{Y}$ (Teil (1)), wegen $Y \cap \mathsf{B}_\varepsilon(p) \neq \emptyset$ für alle $\varepsilon > 0$.
  - Es gilt $p \notin Y^\circ$; denn Y ist keine Umgebung von p (Teil (2)), wegen $\mathsf{B}_\varepsilon(p) \not\subseteq Y$ für $\varepsilon > 0$. ☐

Übung 140. *Sei (X, d) ein metrischer Raum, und $Y \subseteq X$ eine Teilmenge. Zeigen Sie:*

$$X \setminus \overline{Y} = (X \setminus Y)^\circ \qquad \textit{sowie} \qquad X \setminus Y^\circ = \overline{X \setminus Y}.$$

Beispiel 95. *Sei $(X, \mathrm{d}) = (\mathbb{R}, \mathrm{d}_{|\cdot|})$, sowie $\mathbb{R} \ni a < b \in \mathbb{R}$ und $p \in \mathbb{R}$:*

a) *Für $Y = [a, b)$ gilt:* $Y^\circ = (a, b)$, $\overline{Y} = [a, b]$, $\partial Y = \{a, b\}$.
b) *Für $Y = (a, b)$ gilt:* $Y^\circ = (a, b)$, $\overline{Y} = [a, b]$, $\partial Y = \{a, b\}$.
c) *Für $Y = [a, b]$ gilt:* $Y^\circ = (a, b)$, $\overline{Y} = [a, b]$, $\partial Y = \{a, b\}$.
d) *Für $Y = \mathbb{R}$ gilt:* $Y^\circ = \mathbb{R}$, $\overline{Y} = \mathbb{R}$, $\partial Y = \emptyset$.
e) *Für $Y = \{p\}$ gilt:* $Y^\circ = \emptyset$, $\overline{Y} = \{p\}$, $\partial Y = \{p\}$.
f) *Für $Y = \{\frac{1}{n} \mid n \in \mathbb{N}_{>0}\}$ gilt:* $Y^\circ = \emptyset$, $\overline{Y} = Y \cup \{0\} = \partial Y$.

Beachte: *Die Begriffe Abschluss, Inneres und Rand hängen empfindlich davon ab, als Teilmenge welchen metrischen Raumes man Y auffasst. Betrachten wir beispielsweise den metrische Unterraum (Z, d_Z) von (X, d) mit $Z := [a, b)$, so erhalten wir:*

a') *Für $Y = [a, b)$ gilt:* $Y^\circ = [a, b)$, $\overline{Y} = [a, b)$, $\partial Y = \emptyset$. *(Z ist offen/abgeschlossen in (Z, d))*
b') *Für $Y = (a, b)$ gilt:* $Y^\circ = (a, b)$, $\overline{Y} = [a, b)$, $\partial Y = \{a\}$.

Beispiel 96. *Sei $(V, \|\cdot\|)$ ein normierter Raum. Für $x \in X$ und $\varepsilon > 0$ gilt*

$$\overline{\mathsf{B}_\varepsilon(x)} = \overline{\mathsf{B}}_\varepsilon(x) \qquad \textit{sowie} \qquad \overline{\mathsf{B}}_\varepsilon(x)^\circ = \mathsf{B}_\varepsilon(x). \tag{370}$$

Beweis der Behauptungen. Es gilt $\mathsf{B}_\varepsilon(x) \subseteq \overline{\mathsf{B}}_\varepsilon(x)$. Somit zeigen Lemma 28 ($\overline{\mathsf{B}}_\varepsilon(x)$ ist abgeschlossen) und (366) (linke Seite) sowie Lemma 27 ($\mathsf{B}_\varepsilon(x)$ ist offen) und (367) (rechte Seite):[87]

$$\overline{\mathsf{B}_\varepsilon(x)} \subseteq \overline{\mathsf{B}}_\varepsilon(x) \qquad \text{sowie} \qquad \overline{\mathsf{B}}_\varepsilon(x)^\circ \supseteq \mathsf{B}_\varepsilon(x).$$

Wir benutzen nun die Vektorraumoperationen auf V, um die Gleichheiten nachzuweisen:

[87] Diese beiden Inklusionen gelten auch in allgemeinen metrischen Räumen immer.

- Angenommen, es existiert ein $z \in \overline{\mathsf{B}}_\varepsilon(x) \setminus \overline{\mathsf{B}_\varepsilon(x)}$.
  - Wegen $\mathsf{B}_\varepsilon(x) \subseteq \overline{\mathsf{B}_\varepsilon(x)}$ gilt dann notwendig $\|z - x\| = \varepsilon$.
  - Wegen $X \setminus \overline{\mathsf{B}_\varepsilon(x)} \in \mathcal{O}(V)$, existiert ein $0 < \delta \leq \varepsilon$ mit $\mathsf{B}_\delta(z) \subseteq X \setminus \overline{\mathsf{B}_\varepsilon(x)}$.

  Wir setzen $y := z - \frac{\delta}{2\|z-x\|} \cdot (z - x)$, und erhalten den Widerspruch: $\left(0 < \frac{\delta}{2\|z-x\|} = \frac{\delta}{2\varepsilon} \leq \frac{1}{2}\right)$

$$\begin{aligned} \|y - z\| &= \tfrac{\delta}{2} && \Longrightarrow && y \in \mathsf{B}_\delta(z) \subseteq X \setminus \overline{\mathsf{B}_\varepsilon(x)} \\ \|y - x\| &\overset{\text{V2)}}{=} \left|1 - \tfrac{\delta}{2\|z-x\|}\right| \cdot \|z - x\| < \varepsilon && \Longrightarrow && y \in \mathsf{B}_\varepsilon(x) \subseteq \overline{\mathsf{B}_\varepsilon(x)}. \end{aligned}$$

- Angenommen, es existiert ein $z \in \overline{\mathsf{B}}_\varepsilon(x)^\circ \setminus \mathsf{B}_\varepsilon(x)$.
  - Wegen $\overline{\mathsf{B}}_\varepsilon(x)^\circ \subseteq \overline{\mathsf{B}}_\varepsilon(x)$, gilt dann notwendig $\|z - x\| = \varepsilon$.
  - Wegen $\overline{\mathsf{B}}_\varepsilon(x)^\circ \in \mathcal{O}(V)$, existiert ein $\delta > 0$ mit $\mathsf{B}_\delta(z) \subseteq \overline{\mathsf{B}}_\varepsilon(x)^\circ$.

  Wir setzen $y := z + \frac{\delta}{2\|z-x\|} \cdot (z - x)$, und erhalten den Widerspruch:

$$\begin{aligned} \|y - z\| &= \tfrac{\delta}{2} && \Longrightarrow && y \in \mathsf{B}_\delta(z) \subseteq \overline{\mathsf{B}}_\varepsilon(x)^\circ \subseteq \overline{\mathsf{B}}_\varepsilon(x) \\ \|y - x\| &\overset{\text{V2)}}{=} \left|1 + \tfrac{\delta}{2\|z-x\|}\right| \cdot \|z - x\| > \varepsilon && \Longrightarrow && y \notin \overline{\mathsf{B}}_\varepsilon(x). \end{aligned}$$ □

Übung 141. *Machen Sie sich anhand eines Gegenbeispiels explizit klar, dass die Identitäten* (370) *in Beispiel 96 für metrische Räume im allgemeinen nicht erfüllt sein müssen. Betrachten Sie hierfür beispielsweise den diskreten metrischen Raum* $(\mathbb{Z}, \mathrm{d})$ *mit* $\mathrm{d}(m, n) = |m - n|$ *für* $m, n \in \mathbb{Z}$ *– vgl. Übung 59) – und hierin entsprechende Bälle mit Radius* 1. *Machen Sie sich ebenfalls klar, dass die Identitäten* (370) *tatsächlich darauf basieren, dass man in normierten Räumen, die (mit der Norm kompatiblen) Vektorraumoperationen zur Verfügung hat, mit denen man aus gegebenen Elementen (Vektoren), neue Elemente (Vektoren) mit gewissen Zusatzeigenschaften konstruieren kann.*

13 Produkträume und Multilineare Abbildungen

In diesem Abschnitt behandeln wir Produkte metrischer Räume, und diskutieren einige Grundlegende Eigenschaften von multilinearen Abbildungen. Sei hierfür im Folgenden immer $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

13.1 Produkte Metrischer Räume

Terminologie 31 (Produkträume)**.** *Seien* $(X_1, \mathrm{d}_1), \dots, (X_n, \mathrm{d}_n)$ *metrische Räume. Wir versehen den Produktraum* $X := X_1 \times \ldots \times X_n$ *immer mit der Maximumsmetrik (Übung 142)*

$$\mathrm{d}_\infty \colon X \times X \to [0, \infty), \qquad ((x_1, \dots, x_n), (y_1, \dots, y_n)) \mapsto \max(\mathrm{d}_1(x_1, y_1), \dots, \mathrm{d}_n(x_n, y_n)). \tag{371}$$

- *Die Projektion auf den j-ten Faktor für $1 \leq j \leq n$, ist gegeben durch* *(vgl. Beispiel 18.b))*

$$\mathrm{pr}_j \colon X_1 \times \cdots \times X_n \to X_j, \qquad (x_1, \dots, x_n) \mapsto x_j.$$

- *Ist $f \colon Y \to X$ eine Abbildung (Y Menge), so setzen wir:*

$$f_j := \mathrm{pr}_j \circ f \colon Y \to X_j \qquad \forall\, 1 \leq j \leq n \qquad \textit{d.h.} \qquad f \colon Y \ni y \mapsto (f_1(y), \dots, f_n(y)) \in X.$$

  Wir notieren im Folgenden auch $f = (f_1, \dots, f_n)$.

Der Umgebungsbegriff sowie Offenheit und Abgeschlossenheit von Teilmengen; und weiterhin Konvergenz von Folgen sowie Stetigkeit von Abbildungen sind dann immer bezüglich der Maximumsmetrik d_∞ *zu verstehen. Beispielsweise:*

- *Eine Folge* $(x_\ell)_{\ell\in\mathsf{N}}$ *in* X *konvergiert gegen ein* $x \in X$*, wenn sie bezüglich* d_∞ *konvergiert.*

  Beachte: *Gemäß Übung 143 (unten) ist dies gleichbedeutend mit der Konvergenz von* $(\mathrm{pr}_j(x_\ell))_{\ell\in\mathsf{N}}$ *bezüglich* d_j *gegen* $\mathrm{pr}_j(x)$ *für alle* $1 \le j \le n$.

- *Sei* (Y, d) *ein metrischer Raum:*
  - *Eine Abbildung* $f : X \to Y$ *heißt stetig (in einem Punkt), wenn sie stetig (in einem Punkt) bezüglich* d_∞ *ist.*
  - *Eine Abbildung* $f : Y \to X$ *heißt stetig (in einem Punkt), wenn sie stetig (in einem Punkt) bezüglich* d_∞ *ist.*

    Beachte: *Lemma 73 weiter unten zeigt, dass dies gleichbedeutend mit der Stetigkeit (in einem Punkt) aller Komponentenfunktionen* $f_1, \ldots, f_n$ *von* f *ist.*

Für (offene) Umgebungen erhalten wir die folgenden Aussagen:

1) $U \subseteq X$ *ist genau dann eine Umgebung von* $x = (x_1, \ldots, x_n) \in U$*, wenn gilt:*

$$\exists\, \varepsilon > 0 : \quad \mathsf{B}[\mathrm{d}_1]_\varepsilon(x_1) \times \ldots \times \mathsf{B}[\mathrm{d}_n]_\varepsilon(x_n) \overset{(*)}{=} \mathsf{B}[\mathrm{d}_\infty]_\varepsilon(x) \subseteq U.$$

Die Gleichheit $(*)$ *erhalten wir mit:*

$$\begin{aligned} y \in \mathsf{B}[\mathrm{d}_1]_\varepsilon(x_1) \times \ldots \times \mathsf{B}[\mathrm{d}_n]_\varepsilon(x_n) \quad &\Longleftrightarrow \quad \mathrm{d}_j(x_j, y_j) < \varepsilon \quad \text{für} \quad 1 \le j \le n \\ &\Longleftrightarrow \quad \mathrm{d}_\infty(x, y) = \max(\mathrm{d}_1(x_1, y_1), \ldots, \mathrm{d}_n(x_n, y_n)) < \varepsilon \\ &\Longleftrightarrow \quad y \in \mathsf{B}[\mathrm{d}_\infty]_\varepsilon(x). \end{aligned}$$

2) $O \subseteq X$ *ist genau dann offen, wenn zu jedem* $x \in O$ *ein* $\varepsilon > 0$ *mit* $\mathsf{B}[\mathrm{d}_\infty]_\varepsilon(x) \subseteq O$ *existiert.*

Insbesondere: *Gegeben nichtleere Teilmengen* $\emptyset \neq Y_j \subseteq X_j$ *für* $1 \le j \le n$*, so gilt:*[88]

$$\begin{aligned} & Y_j \subseteq X_j \quad \text{offen in} \quad (X_j, \mathrm{d}_j) \quad \text{für} \quad 1 \le j \le n \\ \Longleftrightarrow \qquad & Y := Y_1 \times \ldots \times Y_n \subseteq X \quad \text{offen in} \quad (X, \mathrm{d}_\infty). \end{aligned} \tag{372}$$

Beachte: *Gilt* $Y_j = \emptyset$ *für ein* $1 \le j \le n$*, so ist* $Y = \emptyset$ *automatisch offen. Allerdings lässt sich dann aus der Offenheit von* $Y = \emptyset$ *nicht die Offenheit aller* Y_j *für* $1 \le j \le n$ *Schlussfolgern.*

Beweis von (372). • Es gelte die erste Zeile in (372). Für $y = (y_1, \ldots, y_n) \in Y$ vorgegeben, existieren dann $\varepsilon_1, \ldots, \varepsilon_n > 0$ mit $\mathsf{B}[\mathrm{d}_j]_{\varepsilon_j}(y_j) \subseteq Y_j$ für $1 \le j \le n$. Für $\varepsilon := \min(\varepsilon_1, \ldots, \varepsilon_n)$ gilt dann

$$\mathsf{B}[\mathrm{d}_\infty]_\varepsilon(y) \overset{(*)}{=} \mathsf{B}[\mathrm{d}_1]_\varepsilon(y_1) \times \ldots \times \mathsf{B}[\mathrm{d}_n]_\varepsilon(y_n) \subseteq \mathsf{B}[\mathrm{d}_1]_{\varepsilon_1}(y_1) \times \ldots \times \mathsf{B}[\mathrm{d}_n]_{\varepsilon_n}(y_n) \subseteq Y_1 \times \ldots \times Y_n = Y.$$

- Es gelte die zweite Zeile in (372); und sei $1 \le j \le n$ sowie $y_j \in Y_j$ fixiert. Wir wählen $y_i \in Y_i$ für $1 \le i \neq j \le n$ (möglich wegen $Y_i \neq \emptyset$ für $1 \le i \le n$), und setzen $y := (y_1, \ldots, y_n) \in Y$. Per Annahme existiert ein $\varepsilon > 0$ mit

$$\mathsf{B}[\mathrm{d}_1]_\varepsilon(y_1) \times \ldots \times \mathsf{B}[\mathrm{d}_n]_\varepsilon(y_n) \overset{(*)}{=} \mathsf{B}[\mathrm{d}_\infty]_\varepsilon(y) \subseteq Y_1 \times \ldots \times Y_n = Y \quad \Rightarrow \quad \mathsf{B}[\mathrm{d}_j]_\varepsilon(y_j) \subseteq Y_j. \quad \square$$

[88]Beachte: Die Äquivalenz (372) bedeutet nicht, dass jede offene Teilmenge von X von der Form $O_1 \times \ldots \times O_n$ mit $O_j \subseteq X_j$ offen für $1 \le j \le n$ ist; denn beliebige Vereinigungen derartiger Produkte offener Mengen, sind ja gemäß Proposition 3.OM3) ebenfalls wieder offen.

Übung 142. *Zeigen Sie, dass* (371) *eine Metrik ist.*

Übung 143. *Seien* $(X_1, \mathrm{d}_1), \ldots, (X_n, \mathrm{d}_n)$ *metrische Räume, und* $X = X_1 \times \ldots \times X_n$*. Zeigen Sie, dass eine Folge* $(x_\ell)_{\ell \in \mathsf{N}}$ *in* (X, d_∞) *genau dann gegen ein* $x \in X$ *konvergiert, wenn* $(\mathrm{pr}_j(x_\ell))_{\ell \in \mathsf{N}} \to \mathrm{pr}_j(x)$ *bezüglich* d_j *für alle* $1 \leq j \leq n$ *gilt.*

Hinweis: Beweis von Lemma 64.

Bemerkung 87.

a) *Seien* $(V_1, \|\cdot\|_1), \ldots, (V_n, \|\cdot\|_n)$ *normierte* $\mathbb{K}$*-Vektorräume. Dann ist* $V := V_1 \times \ldots \times V_n$ *ein normierter* $\mathbb{K}$*-Vektorraum vermöge der Maximumsnorm*

$$\|\cdot\|_\infty \colon V \to [0, \infty), \qquad (v_1, \ldots, v_n) \mapsto \max(\|v_1\|_1, \ldots, \|v_n\|_n).$$

Setzen wir $(X_j, \mathrm{d}_j) = (V_j, \mathrm{d}_{\|\cdot\|_j})$ *für* $1 \leq j \leq n$*, so gilt* $\mathrm{d}_\infty = \mathrm{d}_{\|\cdot\|_\infty}$*.*

Beachte: *Sei* $\|\cdot\|'_j$ *eine Norm auf* V_j *für* $1 \leq j \leq n$*, sowie* $\|\cdot\|'_\infty$ *die zugehörige Maximumsnorm. Gilt* $\|\cdot\|_j \sim_{\mathrm{N}} \|\cdot\|'_j$ *für alle* $1 \leq j \leq n$*, so auch* $\|\cdot\|_\infty \sim_{\mathrm{N}} \|\cdot\|'_\infty$*.*[89]

b)
- *Sei* $(V_j, \|\cdot\|_j) = (\mathbb{K}, |\cdot|)$ *für alle* $1 \leq j \leq n$*. Dann ist* $\|\cdot\|_\infty$ *gerade die übliche Maximumsnorm auf* $V = \mathbb{K}^n$*, also* d_∞*, die zur Maximumsnorm gehörige Normmetrik.*

  Insbesondere zeigt (372)*, dass ein Produkt* $O_1 \times \ldots \times O_n$ *nichtleerer offener Teilmengen* $O_1, \ldots, O_n \subseteq \mathbb{K}$ *(bspw. offene Intervalle in* $\mathbb{R}$ *im Falle* $\mathbb{K} = \mathbb{R}$*) offen in* $\mathbb{K}^n$ *ist.*
- *Sei allgemeiner* $(V_j, \|\cdot\|_j) = (\mathbb{K}^{m_j}, \|\cdot\|^j_\infty)$ *mit* $m_j \in \mathbb{N}_{>0}$ *und* $\|\cdot\|^j_\infty$ *die Maximumsnorm auf* $\mathbb{K}^{m_j}$*, für* $1 \leq j \leq n$*. Dann ist* $\|\cdot\|_\infty$ *ebenfalls die übliche Maximumsnorm auf* $\mathbb{K}^{m_1 + \cdots + m_n}$*.*

  Ist also $O_j \subseteq \mathbb{K}^{m_j}$ *nichtleer und offen für* $1 \leq j \leq n$*, so ist* $O_1 \times \ldots \times O_n$ *offen in* $\mathbb{K}^{m_1 + \cdots + m_n}$*.*[90]

Übung 144. *Seien* $(X_1, \mathrm{d}_1), (X_2, \mathrm{d}_2), (Y, \mathrm{d})$ *metrische Räume, und sei* $f \colon X_1 \times X_2 \to Y$ *stetig (bezüglich* d_∞*). Zeigen Sie, dass für jedes fixierte* $x_1 \in X_1$ *und* $x_2 \in X_2$ *die Abbildungen*

$$f_{x_2} \colon X_1 \ni z \mapsto f(z, x_2) \in Y \qquad \textit{und} \qquad f_{x_1} \colon X_2 \ni z \mapsto f(x_1, z) \in Y$$

ebenfalls stetig sind.

Lemma 73. *Seien* $(Y, \mathrm{d}), (X_1, \mathrm{d}_1), \ldots, (X_n, \mathrm{d}_n)$ *metrische Räume, und* $f \colon Y \to X_1 \times \ldots \times X_n$ *eine Abbildung. Es ist* f *genau dann stetig in* $y \in Y$ *(bezüglich der Maximumsmetrik* d_∞*), wenn die Komponentenfunktionen* $f_1, \ldots, f_n$ *stetig in* y *sind.*

Beweis. Sei $\varepsilon > 0$ vorgegeben.

- Ist f stetig in y, so existiert ein $\delta > 0$ mit

$$\begin{aligned} \mathrm{d}(y, y') < \delta \quad \text{für} \quad y' \in Y \qquad &\Longrightarrow \qquad \mathrm{d}_\infty\big(f(y), f(y')\big) < \varepsilon \\ &\Longrightarrow \qquad \mathrm{d}_j\big(f_j(y), f_j(y')\big) \leq \mathrm{d}_\infty\big(f(y), f(y')\big) < \varepsilon \quad \text{für} \quad 1 \leq j \leq n. \end{aligned}$$

  Dies zeigt die Stetigkeit der Komponentenfunktionen in y.

[89] Sei $C_1, \ldots, C_n > 0$ mit $\|\cdot\|'_j \leq C_j \cdot \|\cdot\|_j$ für $1 \leq j \leq n$. Für $C := \max(C_1, \ldots, C_n)$, gilt dann offensichtlich $\|\cdot\|'_\infty \leq C \cdot \|\cdot\|_\infty$. Die umgekehrte Abschätzung folgt analog.

[90] Sei beispielsweise $n = 2$, $m_1 = 2$, $m_2 = 1$, sowie $O_1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + x^2 < 1\}$ und $O_2 = (0, 1)$. Dann ist $O_1 \times O_2$ ein offener Zylinder in $\mathbb{R}^3$ mit Radius 1 und Länge 1.

- Sind die Komponentenfunktionen stetig in y, so existieren $\delta_1, \ldots, \delta_n > 0$ mit

$$\mathrm{d}(y, y') < \delta_j \quad \text{für} \quad y' \in Y \qquad \Longrightarrow \qquad \mathrm{d}_j\big(f_j(y), f_j(y')\big) < \varepsilon$$

für $1 \leq j \leq n$. Für $\delta := \min(\delta_1, \ldots, \delta_n)$, gilt dann offensichtlich

$$\mathrm{d}(y, y') < \delta \quad \text{für} \quad y' \in Y \qquad \Longrightarrow \qquad \mathrm{d}_\infty\big(f(y), f(y')\big) < \varepsilon,$$

was die Stetigkeit von f in y zeigt. □

Sei (Y, d) ein metrischer Raum und $n \in \mathbb{N}_{>0}$. Wegen der Äquivalenz aller Normen auf $\mathbb{K}^n$ ist die Stetigkeit einer Abbildung $f\colon Y \to \mathbb{K}^n$ nicht von der expliziten Wahl der Norm $\|\cdot\|$ auf $\mathbb{K}^n$ abhängig (siehe Bemerkung 79). Konkreter gilt:

Lemma 74. *Sei (Y, d) ein metrischer Raum. Eine Abbildung $f\colon Y \to \mathbb{K}^n$ ist genau dann stetig in $y \in Y$, wenn die Komponentenfunktionen $f_1, \ldots, f_n$ stetig in y sind.*

Beweis. Wegen Bemerkung 79 können wir ohne Einschränkung die Maximumsnorm auf $\mathbb{K}^n$ betrachten; und dann folgt die Behauptung sofort aus Bemerkung 87.b) sowie Lemma 73. □

*Bemerkung: Alternativ kann man auch mit Korollar 50 sowie der Äquivalenz von Stetigkeit und Folgenstetigkeit (Proposition 8) argumentieren:

$$\begin{array}{lcll}
f \text{ stetig in } y & \overset{\text{Proposition 8}}{\Longleftrightarrow} & f \text{ folgenstetig in } y & \\
& \overset{\text{def.}}{\Longleftrightarrow} & (y_n)_{n\in\mathsf{N}} \to y \quad \Rightarrow & f(y_n) \to f(y) \\
& \overset{\text{Korollar50}}{\Longleftrightarrow} & (y_n)_{n\in\mathsf{N}} \to y \quad \Rightarrow & f_j(y_n) \to f(y) \text{ für } 1 \leq j \leq n \\
& \overset{\text{def.}}{\Longleftrightarrow} & f_j \text{ folgenstetig in } y \text{ für } 1 \leq j \leq n & \\
& \overset{\text{Proposition 8}}{\Longleftrightarrow} & f_j \text{ stetig in } y \text{ für } 1 \leq j \leq n. &
\end{array}$$

Bemerkung* 15. *Wir wollen das wichtige Lemma 74 noch einmal etwas konkreter begründen: Sind $\|\cdot\|_1$ und $\|\cdot\|_2$ Normen auf $\mathbb{K}^n$, so gilt $\|\cdot\|_1 \leq C \cdot \|\cdot\|_2$ für ein $C > 0$ (Satz 79), also*

$$\|f(x) - f(y)\|_1 \leq C \cdot \|f(x) - f(y)\|_2 \qquad \forall\, x, y \in X. \tag{373}$$

Sei $\|\cdot\|$ eine Norm auf $\mathbb{K}^n$:

- *Wählen wir in (373) $\|\cdot\|_1 \equiv \|\cdot\|$ sowie $\|\cdot\|_2 \equiv \|\cdot\|_\infty$, so erhalten wir*

$$\|f(x) - f(y)\| \leq C \cdot \max(\{|f_j(x) - f_j(y)| \mid 1 \leq j \leq n\}) \qquad \forall\, x, y \in Y.$$

Hiermit folgt, dass f bezüglich $\|\cdot\|$ stetig in $y \in Y$ ist, wenn f_j für alle $1 \leq j \leq n$ stetig in y ist.

- *Wählen wir in (373) $\|\cdot\|_1 \equiv \|\cdot\|_\infty$ sowie $\|\cdot\|_2 \equiv \|\cdot\|$, so erhalten wir*

$$|f_j(x) - f_j(y)| \leq \|f(x) - f(y)\|_\infty \leq C \cdot \|f(x) - f(y)\| \qquad \forall\, x, y \in Y,\ 1 \leq j \leq n.$$

Hiermit folgt, dass f_j für alle $1 \leq j \leq n$ stetig in $y \in Y$ ist, wenn f bezüglich $\|\cdot\|$ stetig in y ist.

Beispiel 97 (Allgemeine Polynomfunktionen)**.** *Sei $m, n \in \mathbb{N}$. Es gilt*

$$\mathrm{id}_{\mathbb{K}^m} = (\mathrm{pr}_1, \ldots, \mathrm{pr}_m)\colon \mathbb{K}^m \to \mathbb{K}^m$$

für die Identitätsabbildung auf $\mathbb{K}^m$.

- *Wegen der Stetigkeit von* $\mathrm{id}_{\mathbb{K}^m}$*, sind gemäß Lemma 74 auch die Komponentenfunktionen, also die Projektionen* pr_j *für* $1 \leq j \leq m$ *stetig.*
- *Wegen Korollar 25 sind dann alle endliche Produkte der Form* $\mathrm{pr}_{j_1} \cdot \ldots \cdot \mathrm{pr}_{j_\ell} \colon \mathbb{K}^m \to \mathbb{K}$ *mit* $\ell \in \mathbb{N}_{>0}$ *und* $1 \leq j_1, \ldots, j_\ell \leq n$ *stetig.*

Für $\alpha = (\alpha_1, \ldots, \alpha_m) \in \mathbb{N}^m$*, sei* $|\alpha| := \alpha_1 + \ldots + \alpha_m$*. Für* $\alpha \in \mathbb{N}^m$ *ist dann das Monom*

$$\mathrm{pr}_1^{\alpha_1} \cdot \ldots \cdot \mathrm{pr}_n^{\alpha_m} \colon \mathbb{K}^m \ni x \mapsto x^\alpha := (x_1)^{\alpha_1} \cdot \ldots \cdot (x_n)^{\alpha_m} \in \mathbb{K}$$

vom Exponenten α *stetig (Korollar 25). Eine Funktion der Gestalt*

$$P \colon \mathbb{K}^m \to \mathbb{K}^n, \qquad x \mapsto \sum_{\alpha \in \mathbb{N}^m \colon |\alpha| \leq k} x^\alpha \cdot u_\alpha$$

mit $k \in \mathbb{N}$ *und* $u_\alpha \in \mathbb{K}^n$ *für alle* $|\alpha| \leq k$*, heißt Polynomfunktion vom Grad* k*, wenn* $u_\alpha \neq 0$ *für ein* $\alpha \in \mathbb{N}^m$ *mit* $|\alpha| = k$ *gilt. Als Linearkombination von (stetigen) Monomen ist* P *gemäß Korollar 25 ebenfalls stetig.*

13.2 Multilineare Abbildungen

Im Folgenden sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $n \in \mathbb{N}_{>0}$.

Terminologie 32. *Gegeben* $\mathbb{K}$*-Vektorräume* $V_1, \ldots, V_n, W$*. Eine Abbildung* $\phi \colon V_1 \times \ldots \times V_n \to W$ *heißt* $\mathbb{K}$-multilinear *(oder einfach multilinear), wenn sie* $\mathbb{K}$*-linear in jedem Faktor ist, d.h., wenn für* $1 \leq j \leq n$ *sowie* $u, v \in V_j$ *und* $\lambda \in \mathbb{K}$ *gilt:*

$$\begin{aligned} &\phi(v_1, \ldots, v_{j-1}, \lambda \cdot u + v, v_{j+1}, \ldots, v_n) \\ &\qquad\qquad = \lambda \cdot \phi(v_1, \ldots, v_{j-1}, u, v_{j+1}, \ldots, v_n) + \phi(v_1, \ldots, v_{j-1}, v, v_{j+1}, \ldots, v_n). \end{aligned}$$

Beachte: *Ist* $v_j = 0$ *für ein* $1 \leq j \leq n$*, so folgt mit Bemerkung 32.i):*

$$\phi(v_1, \ldots, v_n) \overset{i)}{=} \phi(v_1, \ldots, v_{j-1}, 0 \cdot v_j, v_{j+1}, \ldots, v_n) = 0 \cdot \phi(v_1, \ldots, v_n) \overset{i)}{=} 0. \tag{374}$$

- *Ist* $n = 1$*, also* $V \equiv V_1$ *und*

  $$\phi(\lambda \cdot v + u) = \lambda \cdot \phi(v) + \phi(u) \qquad \forall\, \lambda \in \mathbb{K},\ u, v \in V,$$

  so heißt ϕ $\mathbb{K}$-linear *(oder einfach lineare Abbildung).*
  - *Die Menge aller lineare Abbildungen* $V \to W$ *wird mit* $\mathrm{Hom}(V, W)$ *notiert (Homomorphismen).*

    Beachte: *Ist* $\phi \in \mathrm{Hom}(V, W)$ *bijektiv, so folgt automatisch* $\phi^{-1} \in \mathrm{Hom}(W, V)$ *(vgl. Beweis von Lemma 15).*
  - *Ist* $V = W$*, so notiert man* $\mathrm{End}(V) := \mathrm{Hom}(V, V)$ *(Endomorphismen).*
  - *Die bijektiven (also invertierbaren) Elemente von* $\mathrm{End}(V)$ *werden Automorphismen genannt und mit* $\mathsf{Aut}(V)$ *notiert. Wegen dem ersten Punkt gilt dann:* $\phi \in \mathsf{Aut}(V) \Leftrightarrow \phi^{-1} \in \mathsf{Aut}(V)$
- *Ist* $n = 2$*, also*

  $$\begin{aligned} \phi(\lambda \cdot v + u, v_2) &= \lambda \cdot \phi(v, v_2) + \phi(u, v_2) \qquad && \forall\, \lambda \in \mathbb{K},\ u, v \in V_1,\ v_2 \in V_2 \\ \phi(v_1, \lambda \cdot v + u) &= \lambda \cdot \phi(v_1, v) + \phi(v_1, u) \qquad && \forall\, \lambda \in \mathbb{K},\ u, v \in V_2,\ v_1 \in V_1, \end{aligned}$$

  so heißt ϕ $\mathbb{K}$-bilinear *(oder einfach bilineare Abbildung).*

Beispiel: *Sei* $V_1, \dots, V_n = V := \mathbb{K}^n$ *für ein* $n \in \mathbb{N}_{>0}$, *sowie* $W := \mathbb{K}$. *Die Determinantenfunktion*

$$\det\colon V^n \to W, \qquad (v_1, \dots, v_n) \mapsto \sum_{\sigma \in \mathrm{S}_n} \big(\operatorname{sgn}(\sigma) \cdot \prod_{j=1}^n v_{\sigma(j),j} \big)$$

($v_j = (v_{j,1}, \dots, v_{j,n}) \in \mathbb{K}^n$ *für* $1 \leq j \leq n$*) ist eine* $\mathbb{K}$*-multilineare Abbildung. Hier bezeichnet* S_n *die Menge aller Bijektionen* $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$ *(Permutationen von* $\{1, \dots, n\}$*), wobei* $\operatorname{sgn}(\sigma)$ *das Signum (siehe lineare Algebra) der Permutation* σ *notiert.*

Lemma 75. *Gegeben normierte* $\mathbb{K}$*-Vektorräume* $(V_1, \|\cdot\|)_1, \dots, (V_n, \|\cdot\|_n), (W, \|\cdot\|)$ *und eine* $\mathbb{K}$*-multilineare Abbildung* $\phi\colon V_1 \times \ldots \times V_n \to W$, *so sind die folgenden Aussage äquivalent zueinander:*

a) ϕ *ist stetig.*

b) ϕ *ist stetig in* 0.

c) Es existiert ein $C \geq 0$ *mit*

$$\|\phi(v_1, \dots, v_n)\| \leq C \cdot \|v_1\|_1 \cdot \ldots \cdot \|v_n\|_n \qquad \forall\, v_j \in V_j \quad \textit{für} \quad j = 1, \dots, n. \tag{375}$$

Bemerkung.
- *Wir versehen natürlich* $V_1 \times \ldots \times V_n$ *mit der, zu den Normmetriken* $\mathrm{d}_{\|\cdot\|_1}, \dots, \mathrm{d}_{\|\cdot\|_n}$ *gehörigen Maximumsmetrik* d_∞, *d.h.* $\mathrm{d}_\infty = \mathrm{d}_{\|\cdot\|_\infty}$ *wegen Bemerkung 87.*a).
- *Offensichtlich ist die Charakterisierung in Teil c) invariant unter dem Austausch von zueinander äquivalenten Norm auf jedem der auftauchenden Vektorräumen. Selbiges gilt für die ersten beiden Bedingungen, wegen Bemerkung 87.*a).

Beweis. Wir zeigen *a)* $\Rightarrow$ *b)* $\Rightarrow$ *c)* $\Rightarrow$ *a)*, wobei die Implikation *a)* $\Rightarrow$ *b)* evident ist:

- Es gelte *b)*. Wegen (374) müssen wir nur die Existenz eines $C \geq 0$ nachweisen, sodass (375) mit $V_j \ni v_j \neq 0$ für alle $1 \leq j \leq n$ gilt. Per Annahme existiert ein $\delta > 0$ mit (vgl. (371))

  $$\phi(\mathsf{B}[\mathrm{d}_\infty]_\delta(0)) \subseteq \mathsf{B}[\mathrm{d}_{\|\cdot\|}]_1(\phi(0)) \overset{(374)}{=} \mathsf{B}[\mathrm{d}_{\|\cdot\|}]_1(0).$$

  Ausgeschrieben bedeutet dies: $\big($beachte $\mathrm{d}_\infty = \mathrm{d}_{\|\cdot\|_\infty}$ wegen Bemerkung 87.a)$\big)$

  $$\|u_1\|_1, \dots, \|u_n\|_n < \delta \quad \text{für} \quad u_1 \in V_1, \dots, u_n \in V_n \qquad \Longrightarrow \qquad \|\phi(u_1, \dots, u_n)\| < 1.$$

  Sei nun $V_j \ni v_j \neq 0$ für $1 \leq j \leq n$. Wir setzen $u_j := \frac{\delta}{2 \cdot \|v_j\|_j} \cdot v_j$ für $1 \leq j \leq n$. Dann gilt $\|u_1\|_1, \dots, \|u_n\|_n = \frac{\delta}{2} < \delta$, also wegen der Multilinearität von ϕ sowie V1) (erste Gleichheit)

  $$\begin{aligned} \tfrac{\delta}{2 \cdot \|v_1\|_1} \cdot \ldots \cdot \tfrac{\delta}{2 \cdot \|v_n\|_n} \cdot \|\phi(v_1, \dots, v_n)\| &= \|\phi(u_1, \dots, u_n)\| < 1 \\ \Longrightarrow \qquad \|\phi(v_1, \dots, v_n)\| &< \underbrace{\big(\tfrac{2}{\delta}\big)^n}_{=:\,C} \cdot \|v_1\|_1 \cdot \ldots \cdot \|v_n\|_n. \end{aligned}$$

- Es gelte c). Sei $v = (v_1, \dots, v_n) \in V := V_1 \times \ldots \times V_n$ fixiert; sowie $Q := \max(1, \|v_1\|_1, \dots, \|v_n\|_n)$. Für $\sigma = (\sigma_1, \dots, \sigma_n) \in V$, ist dann $\phi(v + \sigma) - \phi(v)$ wegen der Multilinearität von ϕ eine endliche Summe der Form ($q \in \mathbb{N}_{>0}$)[91]

  $$\phi(v + \sigma) - \phi(v) = \sum_{p=1}^q \phi(u[p]_1, \dots, u[p]_n) \quad \text{mit} \quad u[p]_j \in \{v_j, \sigma_j\} \quad \text{für} \quad 1 \leq j \leq n,$$

  wobei für jedes $1 \leq p \leq q$ mindestens ein $1 \leq j_p \leq n$ existiert, sodass $u[p]_j = \sigma_j$ gilt.

[91] Für $n = 1$, ist $\phi(v + \sigma) - \phi(v) = \phi(\sigma)$; und für $n = 2$, ist $\phi(v + \sigma) - \phi(v) = \phi(v_1, \sigma_2) + \phi(\sigma_1, v_2) + \phi(\sigma_1, \sigma_2)$.

Für $\sigma \in V$ mit $\|\sigma\|_\infty \leq 1$ (d.h., $\|\sigma_j\|_j \leq 1$ für $1 \leq j \leq n$), gilt dann $\|u[p]_1\|_1 \cdot \ldots \cdot \|u[p]_n\|_n \leq Q^{n-1} \cdot \|\sigma\|_\infty$ für $1 \leq p \leq q$, also

$$\begin{aligned}\|\phi(v+\sigma) - \phi(v)\| &\leq \sum_{p=1}^{q} \|\phi(u[p]_1, \ldots, u[p]_n)\| \leq C \cdot \sum_{p=1}^{q} \|u[p]_1\|_1 \cdot \ldots \cdot \|u[p]_n\|_n \\ &\leq C \cdot \sum_{p=1}^{q} Q^{n-1} \cdot \|\sigma\|_\infty = q \cdot C \cdot Q^{n-1} \cdot \|\sigma\|_\infty.\end{aligned}$$

Für $\varepsilon > 0$ vorgegeben und $0 < \delta < \min\big(1, \varepsilon/(q \cdot Q^{n-1})\big)$, gilt somit

$$\|\sigma\|_\infty < \delta \quad \text{für} \quad \sigma \in V \qquad \Longrightarrow \qquad \|\phi(v+\sigma) - \phi(v)\| < \varepsilon$$

also $\phi(\mathsf{B}[\mathrm{d}_\infty]_\delta(v)) \subseteq \mathsf{B}[\mathrm{d}_{\|\cdot\|}]_\varepsilon(\phi(v))$. $\square$

Der nächste Satz und das folgende Korollar zeigen, dass im endlichdimensionalen Fall jede multilineare Abbildung automatisch stetig ist. In beliebigen (unendlichdimensonalen) normierten Räumen muss dies allerdings nicht mehr gelten:

- Ein unendlichdimensionales stetige Beispiel liefert für fixierte $\mathbb{R} \ni a < b \in \mathbb{R}$ die lineare Integralabbildung

$$\mathfrak{I}(a,b) \colon C^0([a,b],\mathbb{R}) \to \mathbb{R}, \qquad f \mapsto \textstyle\int_a^b f,$$

  wobei wir $V = C^0([a,b],\mathbb{R})$ als normierten $\mathbb{R}$-Vektorraum mit der Supremumsnorm $|\cdot|_\infty$ auffassen (vgl. (223)). In der Tat gilt ja wegen Satz 55 (erster Schritt) und der Monotonie der Riemann-Integrals (zweiter Schritt):

$$|\mathfrak{I}(a,b)(f)| = \big|\textstyle\int_a^b f\big| \leq \int_a^b |f| \leq \int_a^b |f|_\infty = (b-a) \cdot |f|_\infty \qquad \forall\, f \in V = C^0([a,b],\mathbb{R}).$$

- Ein unendlichdimensionales unstetiges Beispiel (Übung) liefert für fixierte $\mathbb{R} \ni a < b \in \mathbb{R}$ die lineare Ableitungsabbildung

$$\mathfrak{D}(a,b) \colon C^1([a,b],\mathbb{R}) \to C^0([a,b],\mathbb{R}), \qquad f \mapsto f',$$

  wobei wir $V = C^1([a,b],\mathbb{R})$ sowie $W = C^0([a,b],\mathbb{R})$ beide als normierten $\mathbb{R}$-Vektorräume jeweils mit der Supremumsnorm $|\cdot|_\infty$ auffassen.

Satz 86. *Sei $(W, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum, sowie $n \geq 1$ und $m[1], \ldots, m[n] \in \mathbb{N}_{>0}$. Jede $\mathbb{K}$-multilineare Abbildung $\phi \colon \mathbb{K}^{m[1]} \times \ldots \times \mathbb{K}^{m[n]} \to W$ ist stetig.*

Beweis. Für $1 \leq j \leq n$, bezeichne $e[j]_1, \ldots, e[j]_{m[j]}$ die kanonischen Basisvektoren des $\mathbb{K}^{m[j]}$, d.h.

$$v[j] = \textstyle\sum_{\ell=1}^{m[j]} v[j]_\ell \cdot e[j]_\ell \qquad \text{für} \qquad v[j] = (v[j]_1, \ldots, v[j]_{m[j]}) \in \mathbb{K}^{m[j]}.$$

Wir erhalten aus der Multilinearität von ϕ sowie der Dreiecksungleichung:[92]

$$\begin{aligned}
\left\| \phi(v[1],\ldots,v[n]) \right\| &= \left\| \phi\big(\textstyle\sum_{\ell_1=1}^{m[1]} v[1]_{\ell_1} \cdot e[1]_{\ell_1}, v[2],\ldots,v[n]\big) \right\| \\
&= \left\| \textstyle\sum_{\ell_1=1}^{m[1]} v[1]_{\ell_1} \cdot \phi\big(e[1]_{\ell_1}, v[2],\ldots,v[n]\big) \right\| \\
&\vdots \\
&= \left\| \textstyle\sum_{\ell_1=1}^{m[1]} v[1]_{\ell_1} \cdot \ldots \cdot \sum_{\ell_n=1}^{m[n]} v[n]_{\ell_n} \cdot \phi\big(e[1]_{\ell_1},\ldots,e[n]_{\ell_n}\big) \right\| \\
&\leq \textstyle\sum_{\ell_1=1}^{m[1]} \underbrace{|v[1]_{\ell_1}|}_{\leq \|v[1]\|_\infty} \cdot \ldots \cdot \sum_{\ell_n=1}^{m[n]} \underbrace{|v[n]_{\ell_n}|}_{\leq \|v[n]\|_\infty} \cdot \left\|\phi\big(e[1]_{\ell_1},\ldots,e[n]_{\ell_n}\big)\right\| \\
&\leq C \cdot \|v[1]\|_\infty \cdot \ldots \cdot \|v[n]\|_\infty
\end{aligned}$$

für $C := (m[1] \cdot \ldots \cdot m[n]) \cdot \max\big(\|\phi(e[1]_{\ell_1},\ldots,e[n]_{\ell_n})\| \mid 1 \leq \ell_j \leq m[j] \text{ für } 1 \leq j \leq n\big)$. □

Korollar 54. *Sei $(W, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum, sowie $V_1,\ldots,V_n$ endlichdimensionale $\mathbb{K}$-Vektorräume mit Dimension ≥ 1. Jede $\mathbb{K}$-multilineare Abbildung $\phi\colon V_1 \times \ldots \times V_n \to W$ ist stetig.*

Beweis. Wegen Übung 131 sind auf den Vektorräumen $V_1,\ldots,V_n$ jeweils alle Normen äquivalent zueinander. Wie in Lemma 75 bereits angemerkt, müssen wir daher (375) nur für irgendeine Wahl von Normen auf den Vektorräumen $V_1,\ldots,V_n$ nachweisen. Wir wählen:

- Isomorphismen $\alpha_j\colon \mathbb{K}^{m[j]} \to V_j$ mit $m_j \in \mathbb{N}_{>0}$ für $1 \leq j \leq n$.
- Normen $\|\cdot\|_j$ auf $\mathbb{K}^{m[j]}$ für $1 \leq j \leq n$.

Wir definieren die Normen $\|\cdot\|'_j := \|\cdot\|_j \circ \alpha_j^{-1}$ auf V_j für $1 \leq j \leq n$. Die Verkettung

$$\tilde{\phi}\colon \mathbb{K}^{m[1]} \times \ldots \times \mathbb{K}^{m[n]} \to W, \qquad (w_1,\ldots,w_n) \mapsto \phi\big(\alpha_1(w_1),\ldots,\alpha_n(w_n)\big)$$

ist $\mathbb{K}$-multilinear, also stetig wegen Satz 86. Wegen Lemma 75 (zweiter Schritt), existiert $C \geq 0$ mit

$$\begin{aligned}
\|\phi(v_1,\ldots,v_n)\| = \big\|\tilde{\phi}(\alpha_1^{-1}(v_1),\ldots,\alpha_n^{-1}(v_n))\big\| &\leq C \cdot \|\alpha_1^{-1}(v_1)\|_1 \cdot \ldots \cdot \|\alpha_n^{-1}(v_n)\|_n \\
&= C \cdot \|v_1\|'_1 \cdot \ldots \cdot \|v_n\|'_n
\end{aligned}$$

für alle $(v_1,\ldots,v_n) \in V_1 \times \ldots \times V_n$. Die Behauptung folgt somit aus Lemma 75. □

Wir wollen nun noch den Fall linearer Abbildung etwas genauer untersuchen.

Terminologie 33 (Operatornorm)**.** *Seien $(V_1, \|\cdot\|_1)$ und $(V_2, \|\cdot\|_2)$ normierte $\mathbb{K}$-Vektorräume.*

1) *Wir bezeichnen die Menge aller* **stetigen** *linearen Abbildungen $\phi\colon V_1 \to V_2$, mit $\mathcal{L}(V_1, V_2)$.*

   Die Menge $\mathcal{L}(V_1, V_2)$ ist ein $\mathbb{K}$-Vektorraum bezüglich der punktweisen Addition, sowie der punktweisen Multiplikation mit Körperelementen (vgl. erste und letzte Zeile in (195)*).*

2) *Wir definieren die (zu den Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ gehörige)* Operatornorm[93] *durch*

$$\|\cdot\|_{\mathrm{op}}\colon \mathcal{L}(V_1,V_2) \to [0,\infty), \qquad \phi \mapsto \sup\nolimits_{\mathbb{R}}\Big(\underbrace{\big\{\|\phi(v)\|_2 \,\big|\, v \in V_1\colon \|v\|_1 \leq 1\big\}}_{=:\, B[\phi]}\Big). \tag{376}$$

[92] Das Kernargument des Beweises ist, dass multilineare Abbildung bereits komplett durch Ihre Werte auf allen Kombinationen von Basisvektoren auf den jeweiligen Faktoren bestimmt sind. Da jeder Faktor endlichdimensional ist, gibt es nur endlich viele solcher Kombinationen. Daher kann $\|\phi\|$ wie angegeben abgeschätzt werden.

[93] Eine Lineare Abbildung wird oft auch Operator genannt.

- $\|\cdot\|_{\mathrm{op}}$ *ist definiert:* $\qquad (\phi \in \mathcal{L}(V_1, V_2))$
  - $\emptyset \neq B[\phi] \ni \|\phi(0)\|_2 = 0$ *gilt wegen* $\|0\|_1 = 0 \leq 1$.
  - $B[\phi]$ *ist nach oben beschränkt wegen Lemma 75. Genauer gilt:*

$$\begin{aligned} \|\phi\|_2 \leq C \cdot \|\cdot\|_1 \ \text{ für } \ C \geq 0 \qquad &\Rightarrow \qquad \|\phi(v)\|_2 \leq C \cdot \|v\|_1 \leq C \ \text{ für } v \in V_1 \ \text{ mit } \|v\|_1 \leq 1 \\ &\Rightarrow \qquad B[\phi] \leq C \\ &\Rightarrow \qquad \|\phi\|_{\mathrm{op}} = \sup_{\mathbb{R}}(B[\phi]) \leq C, \end{aligned} \tag{377}$$

  wobei Lemma 75 die Existenz eines derartigen $C \geq 0$ *garantiert.*
- *Die Normeigenschaften* V1), V2), V3) *folgen problemlos aus den Rechenregeln für Suprema.*
- $\|\phi\|_{\mathrm{op}}$ *für* $\phi \in \mathcal{L}(V_1, V_2)$ *ist die kleinste Konstante* $C \geq 0$*, sodass die linke Seite von* (377) *gilt. Dies folgt unmittelbar aus* (377)*, wegen*

$$\|\phi(v)\|_2 \leq \|\phi\|_{\mathrm{op}} \cdot \|v\|_1 \qquad \forall\, v \in V_1. \tag{378}$$

  Beweis von (378). Sei $0 \neq v \in V_1$ (für $v = 0$ ist die Abschätzung klar). Aus der Linearität von ϕ (erster Schritt), sowie $\big\|\frac{v}{\|v\|_1}\big\|_1 = 1$ also $\big\|\phi\big(\frac{v}{\|v\|_1}\big)\big\|_2 \in B[\phi]$ (zweiter Schritt), erhalten wir:

$$\|\phi(v)\|_2 = \left\|\phi\left(\frac{v}{\|v\|_1}\right)\right\|_2 \cdot \|v\|_1 \leq \|\phi\|_{\mathrm{op}} \cdot \|v\|_1. \qquad \square$$

- *Ist* $V_1 \neq \{0\}$ *(d.h.* V_1 *ist nicht der Nullvektorraum), so gilt*

$$\|\phi\|_{\mathrm{op}} = \|\phi\|^{\sim}_{\mathrm{op}} := \sup_{\mathbb{R}}\big(\underbrace{\big\{\|\phi(v)\|_2 \,\big|\, v \in V_1 \colon \|v\|_1 = 1\big\}}_{=:\, \tilde{B}[\phi]}\big) \qquad \forall\, \phi \in \mathcal{L}(V_1, V_2). \tag{379}$$

  Beweis von (379). Es gilt $\tilde{B}[\phi] \neq \emptyset$; denn es existiert ein $0 \neq v \in V_1$, also $\phi\big(\frac{v}{\|v\|_1}\big) \in \tilde{B}[\phi]$ wegen $\big\|\frac{v}{\|v\|_1}\big\|_1 = 1$. Wir erhalten:
  - $\|\phi\|^{\sim}_{\mathrm{op}} = \sup_{\mathbb{R}}(\tilde{B}[\phi]) \overset{\tilde{B}[\phi] \subseteq B[\phi]}{\leq} \sup_{\mathbb{R}}(B[\phi]) = \|\phi\|_{\mathrm{op}}$,
  - $\|\phi\|_2 \leq \|\phi\|^{\sim}_{\mathrm{op}} \cdot \|\cdot\|_1$ folgt wie im Beweis von (378). Also gilt $\|\phi\|_{\mathrm{op}} \leq \|\phi\|^{\sim}_{\mathrm{op}}$ wegen (377). $\square$

3) *Im Folgenden werden wir* $\mathcal{L}(V_1, V_2)$ *immer mit der Operatornorm* (376) *ausstatten. D.h. ist* (X, d) *ein metrischer Raum, so gilt:*

   - *Eine Abbildung* $X \to \mathcal{L}(V_1, V_2)$ *heißt stetig (in einem Punkt), wenn sie stetig (in einem Punkt) bezüglich* $\|\cdot\|_{\mathrm{op}}$ *ist.*
   - *Eine Abbildung* $\mathcal{L}(V_1, V_2) \to X$ *heißt stetig (in einem Punkt), wenn sie stetig (in einem Punkt) bezüglich* $\|\cdot\|_{\mathrm{op}}$ *ist.*

   Dies führt zu keinen Mehrdeutigkeiten; denn sind $\|\cdot\|_1 \sim_{\mathrm{N}} \|\cdot\|'_1$ *bzw.* $\|\cdot\|_2 \sim_{\mathrm{N}} \|\cdot\|'_2$ *äquivalente Normen auf* V_1 *bzw.* V_2*, so gilt auch* $\|\cdot\|_{\mathrm{op}} \sim_{\mathrm{N}} \|\cdot\|'_{\mathrm{op}}$ *für die zugehörigen Operatornormen.*

Übung 145. *Machen Sie sich klar, dass* $\mathcal{L}(V_1, V_2)$ *ein* $\mathbb{K}$*-Vektorraum ist, und weisen Sie die Normeigenschaft der Operatornorm* (376) *nach. Zeigen Sie weiterhin die Äquivalenzaussage bezüglich der Operatornormen in Terminologie 33.*3).

Beispiel 98 (Die Kompositionsabbildung)**.** *Seien* $(V_1, \|\cdot\|_1)$, $(V_2, \|\cdot\|_2)$, $(V_3, \|\cdot\|_3)$ *normierte* $\mathbb{K}$*-Vektorräume. Wir betrachten die normierten* $\mathbb{K}$*-Vektorräume*

$$(\mathcal{L}(V_1, V_2), \|\cdot\|_{12,\mathrm{op}}), \qquad (\mathcal{L}(V_2, V_3), \|\cdot\|_{23,\mathrm{op}}), \qquad (\mathcal{L}(V_1, V_3), \|\cdot\|_{13,\mathrm{op}}),$$

wobei $\|\cdot\|_{12,\mathrm{op}}$ *bzw.* $\|\cdot\|_{23,\mathrm{op}}$ *bzw.* $\|\cdot\|_{13,\mathrm{op}}$*, die zu den Normen* $\|\cdot\|_1, \|\cdot\|_2$ *bzw.* $\|\cdot\|_2$*,* $\|\cdot\|_3$ *bzw.* $\|\cdot\|_1$*,* $\|\cdot\|_3$ *gehörige Operatornorm bezeichnet. Die Kompositionsabbildung*

$$\circ\colon \mathcal{L}(V_2, V_3) \times \mathcal{L}(V_1, V_2) \to \mathcal{L}(V_1, V_3), \qquad (\phi, \psi) \mapsto \phi \circ \psi \tag{380}$$

ist offensichtlich $\mathbb{K}$*-bilinear (Übung), und weiterhin stetig wegen* $(\phi \circ \psi = \circ(\phi, \psi))$

$$\|\phi \circ \psi\|_{13,\mathrm{op}} \leq \|\phi\|_{23,\mathrm{op}} \cdot \|\psi\|_{12,\mathrm{op}}. \tag{381}$$

Beweis. Wegen Lemma 75 genügt es (381) nachzuweisen. Nun erhalten wir durch zweimaliges Anwenden von (378) (zweiter und dritter Schritt), dass

$$\|(\phi \circ \psi)(v)\|_3 = \|\phi(\psi(v))\|_3 \leq \|\phi\|_{23,\mathrm{op}} \cdot \|\psi(v)\|_2 \leq \|\phi\|_{23,\mathrm{op}} \cdot \|\psi\|_{12,\mathrm{op}} \cdot \|v\|_1 = \|\phi\|_{23,\mathrm{op}} \cdot \|\psi\|_{12,\mathrm{op}}$$

für alle $v \in V_1$ mit $\|v\|_1 \leq 1$ gilt, womit (381) nun unmittelbar folgt. □

Bemerkung 88. *Seien* V_1, V_2 *endlichdimensionale* $\mathbb{K}$*-Vektorräume, d.h.,* $V_1 \cong \mathbb{K}^m$ *und* $V_2 \cong \mathbb{K}^n$ *für* $m, n \in \mathbb{N}_{>0}$*. Korollar 54 zeigt, dass* $\mathcal{L}(V_1, V_2) = \mathrm{Hom}(V_1, V_2)$ *gilt; und wegen Übung 131 (Äquivalenz aller Normen auf einem endlichdimensionalen Vektorraum) und Übung 145 sind alle* Operatornormen *auf* $\mathrm{Hom}(V_1, V_2)$ *äquivalent zueinander.*

Vielmehr gilt wegen Übung 131 sogar, dass alle Normen auf $\mathrm{Hom}(V_1, V_2)$ *äquivalent zueinander sind; und zwar weil* $\mathrm{Hom}(V_1, V_2)$ *ein* $(m \cdot n)$*-dimensionaler (endlichdimensionaler)* $\mathbb{K}$*-Vektorraum ist, da man diesen mit dem* $\mathbb{K}^{m \cdot n}$ *vermöge einer* $\mathbb{K}$*-lineare Bijektion (Isomorphismus) identifizieren kann. Beispielsweise stelle man ein Element in* $\mathrm{Hom}(V_1, V_2)$ *zunächst bezüglich einer Wahl von Basen durch eine* $n \times m$*-Matrix*

$$\begin{pmatrix} A_{11} & \cdots & A_{1m} \\ A_{21} & \cdots & A_{2m} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nm} \end{pmatrix}$$

dar, und ordne dann alle Zeilen nebeneinander zu einem Element $(A_{11}, \dots A_{1m}, \dots, A_{n1}, \dots, A_{nm}) \in \mathbb{K}^{m \cdot n}$ *an (Details weiter unten).*

In diesem Sinne folgt die Stetigkeit der $\mathbb{K}$*-bilinearen Kompositionsabbildung*

$$\circ\colon \mathrm{Hom}(U_2, U_3) \times \mathrm{Hom}(U_1, U_2) \to \mathrm{Hom}(U_1, U_3)$$

aus Beispiel 98, für den endlichdimensionalen Fall $U_1 \cong \mathbb{K}^{m_1}$*,* $U_2 \cong \mathbb{K}^{m_2}$*,* $U_3 \cong \mathbb{K}^{m_3}$ *mit* $m_1, m_2, m_3 \in \mathbb{N}_{>0}$ *nun alternativ auch sofort aus Korollar 54 (mit* $n \equiv 2$*,* $V_1 \equiv \mathrm{Hom}(U_2, U_3)$*,* $V_2 \equiv \mathrm{Hom}(U_1, U_2)$*,* $W = \mathrm{Hom}(U_1, U_3)$*).*

Notation 31. *Sei* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$*.*

- *Für* $m, n \geq 1$*, definieren wir die Indexmenge*

$$\mathcal{I}(n, m) := \{1, \dots, n\} \times \{1, \dots, m\}.$$

- *Für* $m, n \geq 1$*, definieren wir den Isomorphismus*

$$\Theta[m, n]\colon \mathrm{M}_{n,m}(\mathbb{K}) \to \mathbb{K}^{m \cdot n}, \qquad \begin{pmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nm} \end{pmatrix} \mapsto (A_{11}, \dots A_{1m}, \dots, A_{n1}, \dots, A_{nm}). \tag{382}$$

- *Für $n \in \mathbb{N}_{>0}$, notieren wir die $n \times n$-Einheitsmatrix mit*

$$\mathbf{1}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathrm{M}_{n,n}(\mathbb{K}).$$

- *Für eine $n \times m$-Matrix $A = (A_{ij})_{(i,j)\in\mathcal{I}(n,m)}$, notieren wir die zugehörige transponierte Matrix ($m \times n$-Matrix) mit $A^\top = (A_{ji})_{(i,j)\in\mathcal{I}(m,n)}$, d.h.*

$$A = \begin{pmatrix} A_{11} & \cdots & A_{m1} \\ A_{12} & \cdots & A_{m2} \\ \vdots & & \vdots \\ A_{1n} & \cdots & A_{mn} \end{pmatrix} \qquad \textit{und} \qquad A^\top = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ A_{21} & \cdots & A_{2n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix}.$$

Insbesondere gilt diese Konvention auch für den Fall $n = 1$ bzw. $m = 1$, d.h. Zeilenvektoren werden zu Spaltenvektoren und umgekehrt:

$$(A_1, \dots, A_m)^\top = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} \qquad \textit{und} \qquad \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix}^\top = (A_1, \dots, A_n).$$

Terminologie 34 (Matrixdarstellung bezüglich der kanonischen Basen)**.** *Seien $m, n \geq 1$ und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ vorgegeben, und es bezeichne $(e_1, \dots, e_m)$ die kanonische Basis von $\mathbb{K}^m$ sowie $(\tilde{e}_1, \dots, \tilde{e}_n)$ die kanonische Basis von $\mathbb{K}^n$. Für $1 \leq i \leq n$, bezeichne weiterhin $\tilde{\mathrm{pr}}_i \colon \mathbb{K}^n \ni (\tilde{w}_1, \dots, \tilde{w}_n) \mapsto \tilde{w}_i \in \mathbb{K}$ die Projektion auf die i-te Koordinate in $\mathbb{K}^n$. Die folgenden Abbildungen sind $\mathbb{K}$-linear und invers zueinander (Isomorphismen):*

$$\Omega[m,n]\colon \qquad \mathrm{Hom}(\mathbb{K}^m, \mathbb{K}^n) \ni \phi \mapsto \big(\tilde{\mathrm{pr}}_i\big(\phi(e_j)\big)\big)_{(i,j)\in\mathcal{I}(n,m)} \in \mathrm{M}_{n,m}(\mathbb{K}) \tag{383}$$

$$\mho[m,n]\colon \quad \mathrm{M}_{n,m}(\mathbb{K}) \ni (A_{ij})_{(i,j)\in\mathcal{I}(n,m)} \mapsto \Big[\phi\colon v \mapsto \underbrace{\textstyle\sum_{i=1}^n \big(\sum_{j=1}^m A_{ij} \cdot v_j\big) \cdot \tilde{e}_i}_{= (A \cdot v^\top)^\top}\Big] \in \mathrm{Hom}(\mathbb{K}^m, \mathbb{K}^n).$$

Der Isomorphismus $\Omega[m,n]$ stellt ein Element $\phi \in \mathrm{Hom}(\mathbb{K}^m, \mathbb{K}^n)$ als Matrix bezüglich der kanonischen Basen dar:

$$\phi(v) = \big(\Omega[m,n](\phi) \cdot v^\top\big)^\top \qquad \forall\, v \in \mathbb{K}^m. \tag{384}$$

Beweis der Behauptungen.

- Es gilt (384), denn für alle $v \in \mathbb{K}^m$ erhalten wir:

$$\begin{aligned} \phi(v) &= \phi\big(\textstyle\sum_{j=1}^m v_j \cdot e_j\big) \\ &= \textstyle\sum_{j=1}^m v_j \cdot \phi(e_j) \\ &= \textstyle\sum_{i=1}^n \tilde{\mathrm{pr}}_i\big(\sum_{j=1}^m v_j \cdot \phi(e_j)\big) \cdot \tilde{e}_i \\ &= \textstyle\sum_{i=1}^n \big(\sum_{j=1}^m v_j \cdot \tilde{\mathrm{pr}}_i(\phi(e_j))\big) \cdot \tilde{e}_i \\ &= \textstyle\sum_{i=1}^n \underbrace{\big(\sum_{j=1}^m \Omega[m,n](\phi)_{ij} \cdot v_j\big)}_{\tilde{\mathrm{pr}}_i\left(\left(\Omega[m,n](\phi) \cdot v^\top\right)^\top\right)} \cdot \tilde{e}_i \\ &= \big(\Omega[m,n](\phi) \cdot v^\top\big)^\top. \end{aligned} \tag{385}$$

- Aus (385) und der Definition (383) von $\mho[m,n]$ (zweiter Schritt), folgt nun unmittelbar:

$$\mho[m,n]\big(\Omega[m,n](\phi)\big)(v) \overset{(383)}{=} \big(\Omega[m,n](\phi) \cdot v^\top\big)^\top \overset{(384)}{=} \phi(v) \qquad \forall\, v \in \mathbb{K}^m.$$

- Offensichtlich gilt für $A = (A_{ij})_{(i,j)\in\mathfrak{I}(n,m)} \in \mathrm{M}_{n,m}(\mathbb{K})$

$$\Omega[m,n]\big(\mho[m,n](A)\big)_{ij} = \tilde{\mathrm{pr}}_i\big(\mho[m,n](A)(e_j)\big) = \tilde{\mathrm{pr}}_i\big(\textstyle\sum_{\ell=1}^n A_{\ell j}\cdot \tilde{e}_\ell\big) = A_{ij}$$

für alle $(i,j) \in \mathfrak{I}(n,m)$. □

Bemerkung 89 (Stetige Abbildungen und Homomorphismen). *Sei $m,n \in \mathbb{N}_{>0}$, $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$, und (X,d) ein metrischer Raum.*

– *In Bemerkung 88 hatten wir bereits klargestellt, dass alle Normen auf* $\mathrm{Hom}(\mathbb{K}^m,\mathbb{K}^n)$ *äquivalent zueinander sind. Der Isomorphismus (vgl.* (382) *und* (383)*)*

$$\Upsilon[m,n] := \Theta[m,n]\circ\Omega[m,n] \colon \mathrm{Hom}(\mathbb{K}^m,\mathbb{K}^n) \to \mathbb{K}^{m\cdot n}$$

identifiziert ja $\mathrm{Hom}(\mathbb{K}^m,\mathbb{K}^n)$ *direkt mit dem (endlichdimensionalen Vektorraum)* $\mathbb{K}^{m\cdot n}$.

– *Als lineare Abbildung ist nun $\Upsilon[m,n]$ gemäß Korollar 54 automatisch stetig; und dasselbe trifft auf die (notwendig lineare) Umkehrabbildung*

$$\Upsilon[m,n]^{-1} = \mho[m,n]\circ\Theta[m,n]^{-1} \colon \mathbb{K}^{m\cdot n} \to \mathrm{Hom}(\mathbb{K}^m,\mathbb{K}^n) \qquad \textit{von} \qquad \Upsilon[m,n] \qquad \textit{zu.}$$

Hiermit erhalten wir die folgenden Sachverhalte:

- *Eine Abbildung $f\colon \mathrm{Hom}(\mathbb{R}^m,\mathbb{R}^n) \to X$ ist genau dann stetig (in einem Punkt $\phi \in \mathrm{Hom}(\mathbb{R}^m,\mathbb{R}^n)$), wenn $f\circ\Upsilon[m,n]^{-1}\colon \mathbb{K}^{m\cdot n}\to X$ stetig (in $\Upsilon[m,n](\phi)$) ist.*
- *Eine Abbildung $f\colon X \to \mathrm{Hom}(\mathbb{R}^m,\mathbb{R}^n)$ ist genau dann stetig (in einem Punkt $x \in X$), wenn $\widetilde{f} := \Upsilon[m,n]\circ f\colon X \to \mathbb{K}^{m\cdot n}$ stetig (in x) ist.*

  Beachte: *Gemäß Lemma 74, ist die Stetigkeit von $\tilde{f}$ (in x) gleichbedeutend damit, dass die Komponentenfunktionen*

$$\widetilde{f}_\ell = \mathrm{pr}_\ell \circ \big(\overbrace{\Upsilon[m,n]\circ f}^{=\widetilde{f}}\big)\colon X \mapsto \mathbb{K} \qquad\qquad \forall\, 1 \le \ell \le m\cdot n \tag{386}$$

  stetig (in x) sind. Dies wiederum ist offensichtlich gleichbedeutend mit der Stetigkeit (in x) der Matrixeintragsfunktionen $\qquad$ *(*$\Upsilon[m,n] = \Theta[m,n]\circ\Omega[m,n]$*)*

$$\widehat{f}_{ij} = \big(\overbrace{\Omega[m,n]\circ f}^{=:\,\widehat{f}}\big)_{ij}\colon X \mapsto \mathbb{K} \qquad\qquad \forall\,(i,j)\in\mathfrak{I}(n,m), \tag{387}$$

  denn die Komponentenfunktionen von $\widetilde{f}$ sind ja die Matrixeintragsfunktionen von $\widehat{f}$.

Terminologie 35 (Affine Abbildung). *Seien V und W beides $\mathbb{K}$-Vektorräume. Eine* affine Abbildung *ist eine Abbildung der Gestalt (vgl. Beispiel 64.b))*

$$f\colon V \to W, \qquad v \mapsto w + \phi(v)$$

mit $w \in W$ und $\phi \in \mathrm{Hom}(V,W)$.

Übung 146 (Zeilensummennorm). *Seien $m,n \ge 1$, sowie $\|\cdot\|_{\infty,\mathrm{op}}$ die, zu den Maximumsnormen auf $\mathbb{K}^m$ und $\mathbb{K}^n$ gehörige Operatornorm auf $\mathrm{Hom}(\mathbb{K}^m,\mathbb{K}^n)$. Zeigen Sie:*

$$\|\phi\|_{\infty,\mathrm{op}} = \max\big(\textstyle\sum_{j=1}^n |\Omega[m,n](\phi)_{ij}| \,\big|\, 1 \le i \le n\big) \qquad \forall\,\phi\in\mathrm{Hom}(\mathbb{R}^m,\mathbb{R}^n).$$

Der Ausdruck auf der rechten Seite wird als Zeilensummennorm *bezeichnet.*

14 Differentialrechnung in mehreren Veränderlichen

In diesem Kapitel wenden wir uns der Differentialrechnung von vektorwertigen Funktionen zu, die zudem von mehreren Argumenten abhängen können, also Funktionen $f\colon \mathbb{R}^m \supseteq U \to \mathbb{R}^n$ für $m, n \in \mathbb{N}_{>0}$, mit einer offenen Teilmenge $U \subseteq \mathbb{R}^n$. In Abschnitt 14.1 diskutieren wir zunächst differenzierbare Kurven, also den Fall $m = 1$; und in Abschnitt 14.2 wenden wir uns dem allgemeinen Falle zu.

14.1 Kurven im $\mathbb{R}^n$

Nach den eher abstrakten Überlegungen der vorangegangenen Kapitel, betrachten wir nun konkrete geometrischen Objekten, nämlich Kurven im $\mathbb{R}^n$ mit $n \in \mathbb{N}_{>0}$. Wir behandeln die Ableitung, das Integral, und die Bogenlänge einer Kurve.

14.1.1 Grundlegende Definitionen

Definition 71.

- *Eine Kurve ist eine Abbildung $\gamma\colon D \to \mathbb{R}^n$, mit $n \in \mathbb{N}_{>0}$ sowie einem nichtentarteten Intervall $D \subseteq \mathbb{R}$. Wir schreiben $\gamma = (\gamma_1, \ldots, \gamma_n)$ mit $\gamma_j = \mathrm{pr}_j \circ \gamma$ für alle $1 \leq j \leq n$ (Terminologie 31).*
- *Eine stetige Kurve wird oft auch als Weg bezeichnet. Gemäß Lemma 74 ist γ genau dann stetig, wenn die Komponentenfunktionen $\gamma_1, \ldots, \gamma_n$ stetig sind.*

Beispiel 99.

a) Sei $p = (p_1, \ldots, p_n) \in \mathbb{R}^n$ und $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$. Dann ist die (affine) Gerade

$$\gamma_{p,v}\colon \mathbb{R} \to \mathbb{R}^n, \qquad t \mapsto p + t \cdot v = (p_1 + t \cdot v_1, \ldots, p_n + t \cdot v_n)$$

eine stetige Kurve.

b) Ist $f\colon D \to \mathbb{R}$ eine (stetige) Abbildung, so ist

$$\gamma_f\colon D \to \mathbb{R}^2, \qquad x \mapsto (x, f(x))$$

eine (stetige) Kurve im $\mathbb{R}^2$, die den Graphen Γ_f von f durchläuft.

c) Die stetige Kurve

$$\gamma_\Psi\colon \mathbb{R} \to \mathbb{R}^2, \qquad x \mapsto (\cos(x), \sin(x))$$

entspricht dem stetigen Gruppenhomomorphismus $\Psi\colon \mathbb{R} \ni x \mapsto \mathrm{e}^{\mathrm{i}x} \in \mathrm{K} = \{v \in \mathbb{R}^2 \mid \|v\|_{\mathrm{euk}} = 1\}$ aus Bemerkung 60.b) (vgl. (257) mit $\mathbb{C} = \mathbb{R}^2$), bildet also $\mathbb{R}$ auf den Einheitskreis K ab. Wegen Satz 52 (Polardarstellung) ist die Einschränkung

$$\gamma_\circ := \gamma_\psi|_{[0,2\pi)}\colon [0, 2\pi) \to \mathrm{K}$$

bijektiv (insbesondere ist γ_ψ auch surjektiv). Mit den in Abschnitt 8.4.2 hergeleiteten Eigenschaften folgt nun bereits, dass $\gamma_\circ$ den Einheitskreis entgegen dem Uhrzeigersinn (im mathematisch positiven Drehsinn) durchläuft, mit

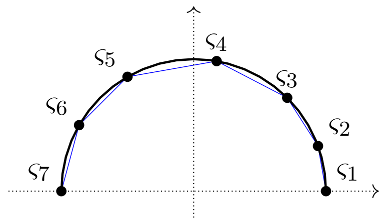
$$\gamma_\circ(0) = (1,0), \quad \gamma_\circ(\tfrac{\pi}{2}) = (0,1), \quad \gamma_\circ(\pi) = (-1,0), \quad \gamma_\circ(\tfrac{3\pi}{2}) = (0,-1), \quad \lim_{x \uparrow 2\pi} \gamma_\circ(x) = (1,0)$$

(wegen Bemerkung 60.a), Satz 51, Korollar 38 sowie der Stetigkeit von γ_ψ). In der Tat hatten wir ja bereits in Bemerkung 62.a) eingesehen, dass $\cos_0 = \cos|_{[0,\pi]}$ streng monoton fallend ist. Aus den Periodizitätseigenschaften von Sinus und Kosinus folgt hiermit:

- $(\mathrm{pr}_1 \circ \gamma_\psi)|_{[0,\pi]} \quad = \cos|_{[0,\pi]}$ *ist streng monoton fallend,*

- $(\mathrm{pr}_1 \circ \gamma_\psi)|_{[\pi,2\pi]} \;= \cos|_{[\pi,2\pi]}$ *ist streng monoton wachsend,*
- $(\mathrm{pr}_2 \circ \gamma_\psi)|_{[-\frac{\pi}{2},\frac{\pi}{2}]} = \sin|_{[-\frac{\pi}{2},\frac{\pi}{2}]}$ *ist streng monoton wachsend,*[94]
- $(\mathrm{pr}_2 \circ \gamma_\psi)|_{[\frac{\pi}{2},\frac{3\pi}{2}]} \;= \sin|_{[\frac{\pi}{2},\frac{3\pi}{2}]}$ *ist streng monoton fallend.*

Insbesondere ist hiermit nun klar, dass für jeden Polygonzug entlang des oberen Halbkreisbogens wie in Abschnitt 4.4.2, die zugehörigen Unterteilungspunkten $\varsigma_1, \ldots, \varsigma_n \in \mathrm{H} = \{(x,y) \in \mathrm{K} \,|\, y \geq 0\}$ *gegeben sind durch* $\varsigma_j = \gamma_\circ(\lambda_j)$ *für* $1 \leq j \leq n$, *für gewisse* $0 = \lambda_1 < \ldots < \lambda_n = \pi$.



$(1,0) = \varsigma_1 < \ldots < \varsigma_7 = (-1,0)$

In Beispiel 101 in Abschnitt 14.1.4, werden wir hiermit schließlich schlussfolgern, dass π *tatsächlich, wie wir in Definition 30 bereits vorgegriffen hatten, gegeben ist als das Supremum* $\pi := \sup_{\mathbb{R}}(\mathcal{P})$ *der Menge* $\mathcal{P}$ *alle Längen derartiger Polygonzüge. Insbesondere wird auch klar werden, dass* $\phi = \mu - \lambda$ *für* $0 < \lambda < \mu < 2\pi$, *tatsächlich die Bogenlänge des Kreissegments (welches den Punkt* $(1,0)$ *nicht enthält) zwischen* $\gamma_\circ(\lambda)$ *und* $\gamma_\circ(\mu)$ *ist.*

d) Die Ellipse um den Punkt $p = (p_1, p_2) \in \mathbb{R}^2$ *mit den Halbachsen* a *und* b, *wird durch die Kurve*

$$\gamma\colon [0,2\pi) \to \mathbb{R}^2, \qquad t \mapsto p + (a \cdot \cos(t), b \cdot \sin(t))$$

beschrieben. Das Bild dieser Kurve ist die Ellipse, die durch die Gleichung

$$\frac{(x_1 - p_1)^2}{a^2} + \frac{(x_2 - p_2)^2}{b^2} = 1$$

beschrieben wird.

e) Eine Schraubenlinie im $\mathbb{R}^3$ *lässt sich durch die folgende Kurve beschreiben:*

$$\gamma\colon \mathbb{R} \to \mathbb{R}^3, \qquad t \mapsto (\cos(t), \sin(t), t)$$

14.1.2 Differenzierbare Kurven

Terminologie 36 (Differenzierbare Kurven)**.** *Sei* $\gamma\colon D \to \mathbb{R}^n$ *eine Kurve* ($D \subseteq \mathbb{R}$ *nichtentartet).*

- γ *heißt differenzierbar in* $p \in D$, *wenn alle Komponentenfunktionen* $\gamma_1, \ldots, \gamma_n$ *differenzierbar in* p *sind. In diesem Fall wird* $\gamma'(p) := (\gamma_1'(p), \ldots, \gamma_n'(p))$ *als die Ableitung von* γ *in* p *bezeichnet, oder auch als der Geschwindigkeitsvektor von* γ *in* p.[95] *Die Zahl*

$$\|\gamma'(p)\|_{\text{euk}} = \sqrt{\sum_{j=1}^{n} |\gamma_j'(p)|^2}$$

heißt Geschwindigkeit γ *in* p *bezeichnet.*

[94] Somit ist sowohl $\gamma_\circ|_{[0,\frac{\pi}{2}]}$ als auch $\gamma_\circ|_{[\frac{3\pi}{2},2\pi)}$ (2π-Periodizität) streng monoton wachsend.
[95] Eine Anschauung liefert Lemma 76.

- *γ heißt k-mal (stetig) differenzierbar für $k \in \mathbb{N} \cup \{\infty\}$, wenn alle Komponentenfunktionen k-mal (stetig) differenzierbar sind. In diesem Fall setzen wir*

$$\gamma^{(\ell)} := \big(\gamma_1^{(\ell)}, \dots, \gamma_n^{(\ell)}\big) \qquad \forall\, \mathbb{N} \ni \ell \leq k.$$

Den Raum aller k-mal stetig differenzierbaren Kurven notieren wir mit $C^k(D, \mathbb{R}^n)$.

Eine Kurve heißt (stetig) differenzierbar, wenn sie 1-mal (stetig) differenzierbar ist. In diesem Fall wird $\gamma' := \gamma^{(1)} = (\gamma_1', \dots, \gamma_n')$ als die Ableitung von γ bezeichnet. Eine C^∞-Kurve wird auch als glatte Kurve bezeichnet.

Bemerkung:

a) *Ist γ k-mal differenzierbar mit $k \in \mathbb{N} \cup \{\infty\}$, so gilt für alle $\mathbb{N} \ni p, q \leq k$ mit $p + q \leq k$:*

$$\big(\gamma^{(p)}\big)^{(q)} = \big(\gamma_1^{(p)}, \dots, \gamma_n^{(p)}\big)^{(q)} = \big(\big(\gamma_1^{(p)}\big)^{(q)}, \dots, \big(\gamma_n^{(p)}\big)^{(q)}\big) = \big(\gamma_1^{(p+q)}, \dots, \gamma_n^{(p+q)}\big) = \gamma^{(p+q)}.$$

b) *Ist γ $(\ell + 1)$-mal differenzierbar mit $\ell \in \mathbb{N}$, so ist $\gamma^{(\ell)}$ stetig.*

Beweis. Die Komponentenfunktionen $\gamma_1^{(\ell)}, \dots, \gamma_n^{(\ell)}$ von $\gamma^{(\ell)}$ sind per Annahme differenzierbar, also stetig wegen Lemma 47. Lemma 74 zeigt somit, dass $\gamma^{(\ell)}$ stetig ist. □

c) *Gilt $\gamma \in C^k(D, \mathbb{R}^n)$ für $k \in \mathbb{N} \cup \{\infty\}$, so ist $\gamma^{(\ell)}$ stetig für alle $\mathbb{N} \ni \ell \leq k$.*

Beweis.
- Sei $k \equiv \infty$: Per Definitionem sind die Komponentenfunktionen von $\gamma^{(\ell)}$ für alle $\ell \in \mathbb{N}$ stetig, also auch $\gamma^{(\ell)}$ wegen Lemma 74.
- Sei $k \in \mathbb{N}$: $\gamma^{(k)}$ ist stetig per Annahme; und für $\mathbb{N} \ni \ell < k$ ist γ $(\ell + 1)$-mal differenzierbar, also $\gamma^{(\ell)}$ stetig wegen Teil b). □

Beispiel 100.

a) Die affine Kurve $\gamma_{p,v} \colon \mathbb{R} \ni t \mapsto p + t \cdot v \in \mathbb{R}^n$ aus Beispiel 99.a) ist glatt, mit

$$\gamma_{p,v}' = v \quad \textit{also} \quad \|\gamma_{p,v}'\|_{\text{euk}} = \|v\|_{\text{euk}} \quad \textit{sowie} \quad {\gamma_{p,v}}^{(\ell)} = 0 \quad \textit{für alle} \quad \ell \geq 2.$$

b) Ist $f \colon D \to \mathbb{R}$ eine k-mal (stetig) differenzierbare Abbildung mit $k \in \mathbb{N}$, so auch die zugehörige „Graphenkurve" $\gamma_f \colon D \ni x \mapsto (x, f(x)) \in \mathbb{R}^2$ aus Beispiel 99.b). Im Falle $k \geq 1$ ist

$$\gamma_f' = (1, f') \quad \textit{also} \quad \|\gamma_f'\|_{\text{euk}} = \sqrt{1 + |f'|^2} \quad \textit{sowie} \quad \gamma_f^{(\ell)} = \big(0, f^{(\ell)}\big) \quad \textit{für alle} \quad 2 \leq \ell \leq k.$$

c) Die Kreiskurve $\gamma_\Psi \colon \mathbb{R} \ni x \mapsto (\cos(x), \sin(x)) \in \mathbb{R}^2$ ist glatt. Unter der Identifikation $\mathbb{R}^2 = \mathbb{C}$, erhält man aus Satz 49 die einfache Formel (Übung)

$$\gamma_\Psi^{(n)} \colon \mathbb{R} \ni x \mapsto \mathrm{i}^n \cdot \mathrm{e}^{\mathrm{i}x} \in \mathbb{C} = \mathbb{R}^2 \qquad \forall\, n \in \mathbb{N},$$

also auch $\|\gamma_\Psi^{(n)}(x)\|_{\text{euk}} = |\mathrm{i}^n \cdot \mathrm{e}^{\mathrm{i}x}| = 1$ für alle $x \in \mathbb{R}$ und $n \in \mathbb{N}$.

Beispielsweise ist:

$$\begin{aligned} \gamma_\Psi' &= \mathrm{i} \cdot (\cos + \mathrm{i} \sin) = -\sin + \mathrm{i} \cos = (-\sin, \cos) \qquad \textit{mit} \qquad \langle \gamma_\psi, \gamma_\psi' \rangle_{\text{euk}} = 0, \\ \gamma_\psi'' &= \mathrm{i}^2 \cdot (\cos + \mathrm{i} \sin) = -\gamma_\psi. \end{aligned}$$

Der Geschwindigkeitsvektor γ_ψ' steht also senkrecht auf dem Ortsvektor γ_ψ, und der „Beschleunigungsvektor" γ_ψ'' ist das Negative des Ortsvektors. Interpretiert man γ_ψ als die Bahnkurve eines Teilchens, welches sich mit konstanter Winkelgeschwindigkeit 1 auf einer Kreisbahn bewegt, so besagt $\gamma_\psi'' = -\gamma_\psi$, dass es sich hierbei um eine beschleunigte Bewegung handelt; denn der Geschwindigkeitvektor γ_ψ' ist ja zeitlich nicht konstant (zwar ist die Geschwindigkeit $\|\gamma_\psi'\|_{\text{euk}} = 1$ konstant, aber γ_ψ' selbst nicht).

Bemerkung 90. *Aus den entsprechenden Resultaten für Funktionen in einer Variablen erhält man unmittelbar die folgenden Rechenregeln (Übung):*

a) Sind $\gamma, \delta \colon D \to \mathbb{R}^n$ *differenzierbar in* $p \in D$*, so auch* $\lambda \cdot \gamma + \mu \cdot \delta$ *für* $\lambda, \mu \in \mathbb{R}$ *mit*

$$(\lambda \cdot \gamma + \mu \cdot \delta)'(p) = \lambda \cdot \gamma'(p) + \mu \cdot \delta'(p).$$

b) Sind $\gamma \colon D \to \mathbb{R}^n$ *und* $f \colon D \to \mathbb{R}$ *differenzierbar in* $p \in D$*, so auch* $f \cdot \gamma$ *mit*

$$(f \cdot \gamma)'(p) = f'(p) \cdot \gamma(p) + f(p) \cdot \gamma'(p).$$

c) Seien $\gamma \colon \tilde{D} \to \mathbb{R}^n$ *und* $\rho \colon D \to \mathbb{R}$ *Abildungen mit* $\mathrm{im}(\rho) \subseteq \tilde{D}$*. Ist* ρ *differenzierbar in* $p \in D$ *und* γ *differenzierbar in* $\rho(p)$*, so ist* $\gamma \circ \rho$ *differenzierbar in* p *mit*

$$(\gamma \circ \rho)'(p) = \rho'(p) \cdot \gamma'(\rho(p)).$$

d) Sind $\gamma, \delta \colon D \to \mathbb{R}^n$ *differenzierbar in* $p \in D$*, so auch* $\langle \gamma, \delta \rangle_{\mathrm{euk}}$ *mit*

$$\langle \gamma, \delta \rangle'_{\mathrm{euk}}(p) = \langle \gamma', \delta \rangle_{\mathrm{euk}}(p) + \langle \gamma, \delta' \rangle_{\mathrm{euk}}(p).$$

Sind die jeweiligen Abbildungen n*-mal differenzierbar (*C^n*-Abbildungen) für ein* $n \in \mathbb{N}$*, so sind wegen Übung 100 auch die entsprechenden Kombinationen* n*-mal differenzierbar (*C^n*-Abbildungen). Speziell erhalten wir in der Situation von Teil b), aus Übung 114 die allgemeinere Leibniz-Regel*

$$(f \cdot \gamma)^{(n)}(p) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(p) \cdot \gamma^{(n-k)}(p) \qquad \forall\, p \in D. \tag{388}$$

Lemma 76 (Geometrische Anschauung)**.** *Sei* $\|\cdot\|$ *eine Norm auf* $\mathbb{R}^n$ *mit* $n \in \mathbb{N}_{>0}$*. Eine Kurve* $\gamma \colon D \to \mathbb{R}^n$ *ist genau dann differenzierbar in* $p \in D$ *mit Ableitung* $\gamma'(p) \in \mathbb{R}^n$*, wenn gilt:*

$$\lim_{h \to 0} \frac{\gamma(p+h) - \gamma(p)}{h} = \gamma'(p) \qquad \textit{also} \qquad \lim_{h \to 0} \left\| \frac{\gamma(p+h) - \gamma(p)}{h} - \gamma'(p) \right\| = 0. \tag{389}$$

Beweis. Sind $\|\cdot\|_1$ und $\|\cdot\|_2$ Normen auf $\mathbb{R}^n$, so gilt $\|\cdot\|_1 \leq C \cdot \|\cdot\|_2$ für ein $C > 0$ (Satz 79), also

$$\|\lambda \cdot (\gamma(x) - \gamma(y)) - w\|_1 \leq C \cdot \|\lambda \cdot (\gamma(x) - \gamma(y)) - w\|_2 \qquad \forall\, x, y \in D,\ \lambda \in \mathbb{R},\ w \in \mathbb{R}^n. \tag{390}$$

- Sei γ differenzierbar in p mit Ableitung $\gamma'(p) = (\gamma_1'(p), \dots, \gamma_n'(p))$ in p. Wir erhalten aus (390) mit $\lambda \equiv \frac{1}{h}$, $w \equiv \gamma'(p)$, $x \equiv p + h$, $y \equiv p$, sowie $\|\cdot\|_1 \equiv \|\cdot\|$ und $\|\cdot\|_2 \equiv \|\cdot\|_\infty$:

$$0 \leq \left\| \frac{\gamma(p+h) - \gamma(p)}{h} - \gamma'(p) \right\| \leq C \cdot \max_{1 \leq j \leq n} \left\{ \left| \frac{\gamma_j(p+h) - \gamma_j(p)}{h} - \gamma_j'(p) \right| \right\} \xrightarrow{h \to 0} 0.$$

  Hiermit folgt (389) unmittelbar.

- Es gelte (389) für ein Element $\gamma'(p) \in \mathbb{R}^n$. Wir erhalten für $1 \leq j \leq n$, aus (390) mit $\lambda \equiv \frac{1}{h}$, $w \equiv \gamma'(p)$, $x \equiv p + h$, $y \equiv p$, sowie $\|\cdot\|_1 \equiv \|\cdot\|_\infty$ und $\|\cdot\|_2 \equiv \|\cdot\|$:

$$\begin{aligned} 0 \leq \left| \frac{\gamma_j(p+h) - \gamma_j(p)}{h} - \mathrm{pr}_j(\gamma'(p)) \right| \leq &\left\| \frac{\gamma(p+h) - \gamma(p)}{h} - \gamma'(p) \right\|_\infty \\ \leq C \cdot &\left\| \frac{\gamma(p+h) - \gamma(p)}{h} - \gamma'(p) \right\| \xrightarrow{h \to 0} 0, \end{aligned}$$

  also $\gamma_j'(p) = \mathrm{pr}_j(\gamma'(p))$. Per definitionem ist somit $\gamma'(p) = (\mathrm{pr}_1(\gamma'(p)), \dots, \mathrm{pr}_n(\gamma'(p))) \in \mathbb{R}^n$ die Ableitung von γ in p. $\square$

Bemerkung* 16 (Differenzierbare Kurven in normierten Räumen). *Die Charakterisierung von Differenzierbarkeit in Lemma 76 verallgemeinert umgehend auf Kurven in normierte Räumen:*

Sei $(V, \|\cdot\|)$ *ein normierter Raum,* $D \subseteq \mathbb{R}$ *nichtentartet, und* $\gamma\colon D \to V$ *eine Abbildung (Kurve):*

- γ *heißt differenzierbar in* $p \in D$ *mit Ableitung* $\gamma'(p) \in V$, *wenn* $\lim_{h\to 0} \frac{\gamma(p+h)-\gamma(p)}{h} = \gamma'(p)$ *gilt.*
- γ *heißt* γ *differenzierbar, wenn* γ *differenzierbar in jedem* $p \in D$ *ist. Dann wird* $\gamma'\colon D \ni t \mapsto \gamma'(t) \in V$ *Ableitung von* γ *genannt.*
- γ *heißt stetig differenzierbar, wenn* γ *differenzierbar und* γ' *stetig ist.*
- *Höhere Ableitungen definiert man induktiv.*

Prinzipiell hätten wir im Falle $V = \mathbb{R}^n$ *Differenzierbarkeit von Kurven auch komplett in dieser Form definieren können, also ohne die Differenzierbarkeit der Komponentenfunktionen zu thematisieren. Dann hätten wir aber die Rechenregeln in Bemerkung 90 mühsam per Hand beweisen müssen.*

Übung*. *Eine Kurve* $\gamma\colon D \to \mathbb{R}^n$ *heißt analytisch, wenn die Komponentenfunktionen analytisch sind.*[96] *Zeigen Sie, dass* $\gamma = 0$ *genau dann gilt, wenn zu jedem* $1 \leq j \leq n$ *ein* $p[j] \in D$ *und eine Folge* $(x[j]_n)_{n\in\mathsf{N}[j]} \in D \setminus \{p[j]\}$ *mit* $(x[j]_n)_{n\in\mathsf{N}[j]} \to p[j]$ *existiert, sodass* $\gamma(x[j]_n) = 0$ *für alle* $n \in \mathsf{N}[j]$ *gilt.*

14.1.3 Riemann-integrierbare Kurven:

Notation 32. *Wir notieren die Menge aller beschränkten Kurven* $\gamma\colon D \to \mathbb{R}^n$ $(D \subseteq \mathbb{R}$ *nichtentartet) mit* $\mathrm{B}(D, \mathbb{R}^n)$.

Beachte: *Wegen Satz 79 (vgl. Bemerkung 77) ist eine Kurve* $\gamma\colon D \to \mathbb{R}^n$ *genau dann beschränkt, wenn ein* $C \geq 0$ *existiert mit* $\|\gamma(t)\|_\infty \leq C$ *für alle* $t \in D$. *Hiermit folgt unmittelbar, dass* γ *genau dann beschränkt ist, wenn alle Komponentenfunktionen* $\gamma_1, \dots, \gamma_n$ *beschränkt sind.*[97]

Definition 72 (Integrierbare Kurven). *Sei* $\mathbb{R} \ni a < b \in \mathbb{R}$.

- *Ein beschränkte Kurve* $\gamma\colon [a,b] \to \mathbb{R}^n$ *(d.h.* $\gamma \in \mathrm{B}([a,b], \mathbb{R}^n)$*) heißt* Riemann-integrierbar, *wenn* $\gamma_1, \dots, \gamma_n \in \mathcal{R}_a^b$ *gilt.*[98]
- *Es bezeichnet* $\mathcal{R}_a^b[n] \subseteq \mathrm{B}([a,b], \mathbb{R}^n)$ *den Raum aller Riemann-integrierbaren Kurven* $\gamma\colon [a,b] \to \mathbb{R}^n$. *Insbesondere also* $\mathcal{R}_a^b[1] = \mathcal{R}_a^b$.

Für $\gamma \in \mathcal{R}_a^b[n]$ *definieren wir das* Riemann-integral *komponentenweise durch:*

$$\int_a^b \gamma := \left(\int_a^b \gamma_1, \dots, \int_a^b \gamma_n\right) \in \mathbb{R}^n$$

(alternativ wieder $\int_a^b \gamma(s)\,\mathrm{d}s$*). Zudem setzen wir*

$$\int_c^c \gamma := 0 \qquad \textit{sowie} \qquad \int_b^a \gamma := -\int_a^b \gamma \qquad \textit{für alle} \qquad a \leq c \leq b. \tag{391}$$

Bemerkung 91. *Durch die komponentenweisen Definition des Riemann-Integrals für Kurven, übertragen sich die meisten Rechenregel aus Kapitel 9 umgehend auf die allgemeine Situation (Übung):*

[96] Alle Kurven in Beispiel 99 analytisch, sofern man in Teil b) die Funktion f als analytisch voraussetzt.

[97] Ist γ beschränkt, so wegen obiger Abschätzung offensichtlich auch alle Komponentenfunktionen von γ. Ist jede Komponentenfunktion von γ beschränkt, existiert also für jedes $1 \leq j \leq n$ ein $C_j \geq 0$ mit $|\gamma_j(t)| \leq C_j$ für alle $t \in D$, so gilt offensichtlich $\|\gamma(t)\|_\infty \leq \max(C_1, \dots, C_n)$ für alle $t \in D$.

[98] Beachte: Gemäß Definition 64 ist jedes Element von $\mathcal{R}_a^b$ beschränkt. Wegen der Anmerkung in Notation 32 ist daher die Forderung $\gamma \in \mathrm{B}([a,b], \mathbb{R}^n)$ eigentlich redundant.

(a) *Intervalladditivität: Für $a < c < b$ gilt $\gamma \in \mathcal{R}_a^b[n]$ genau dann, wenn $\gamma|_{[a,c]} \in \mathcal{R}_a^c[n]$ und $\gamma|_{[c,b]} \in \mathcal{R}_c^b[n]$ gilt. In diesem Fall ist* *(Satz 53.*(I1)*)*

$$\int_a^b \gamma = \int_a^c \gamma + \int_c^b \gamma. \tag{392}$$

(Wegen (391) gilt (392) natürlich auch im Falle $c \in \{a,b\}$.)

(b) *Linearität: Für $\lambda, \mu \in \mathbb{R}$ und $\gamma, \delta \in \mathcal{R}_a^b[n]$, gilt $(\lambda \cdot \gamma + \mu \cdot \delta) \in \mathcal{R}_a^b[n]$ mit* *(Satz 53.*(I3)*)*

$$\int_a^b (\lambda \cdot \gamma + \mu \cdot \delta) = \lambda \cdot \int_a^b \gamma + \mu \cdot \int_a^b \delta.$$

(c) *Normierung: $\int_a^b w = (b-a) \cdot w$ für alle $w \in \mathbb{R}^n$.* *(Satz 53)*

(d) *Sei $\gamma \in \mathcal{R}_c^d[n]$ mit $\mathbb{R} \ni c < d \in \mathbb{R}$, sowie $a, b \in \{c,d\}$ mit $a \neq b$. Lemma 59 zeigt $\|\gamma\|_\infty \in \mathcal{R}_c^d$. Mit Korollar 42 im ersten Schritt, sowie Satz 53.*(I2) *(Monotonie) im zweiten Schritt folgt:*

$$\Big\| \int_a^b \gamma \Big\|_\infty \leq \max\Big(\big| \int_a^b |\gamma_1| \big|, \dots, \big| \int_a^b |\gamma_n| \big| \Big) \leq \big| \int_a^b \|\gamma\|_\infty \big|. \tag{393}$$

(e) *Für $g \in \mathcal{R}_a^b$ und $\gamma \in \mathcal{R}_a^b[n]$ gilt $g \cdot \gamma \in \mathcal{R}_a^b[n]$.* *(Lemma 60)*

Sei nun $\gamma\colon [a,b] \to \mathbb{R}^n$ stetig mit $\mathbb{R} \in a < b \in \mathbb{R}$. Dann sind auch alle Komponentenfunktionen von γ stetig (Lemma 74), also Riemann-integrierbar wegen Satz 54. Somit gilt $C^0([a,b],\mathbb{R}^n) \subseteq \mathcal{R}_a^b[n]$ für alle $\mathbb{R} \in a < b \in \mathbb{R}$. Die Begrifflichkeiten und Integrationsresultate für stetige Funktionen aus Abschnitt 9.5 verallgemeinern sich nun umgehend:

Terminologie 37. *Sei $\gamma\colon D \to \mathbb{R}^n$ stetig ($D \subseteq \mathbb{R}$ nichtentartet).*

- *Für $D \ni a < b \in D$ ist $\gamma|_{[a,b]}$ ebenfalls stetig, also Riemann-integrierbar da hiermit die Komponenten stetig sind. Wir definieren*

$$\int_a^b \gamma := \int_a^b \gamma|_{[a,b]}, \qquad \int_b^a \gamma := -\int_a^b \gamma \qquad \textit{sowie} \qquad \int_c^c \gamma := 0 \quad \textit{für} \quad c \in [a,b]. \tag{394}$$

  Die Aussagen in Bemerkung 91 gelten dann in entsprechend gleicher Form.

- *$\Gamma \in C^1(D,\mathbb{R}^n)$ heißt Stammfunktion von γ, wenn $\Gamma' = \gamma$ gilt.*

  Beachte: *Γ ist genau dann eine Stammfunktion von γ, wenn Γ_j für jedes $1 \leq j \leq n$ eine Stammfunktion von γ_j ist.*

- *Für jede Kurve $\alpha\colon D \to \mathbb{R}^n$, setzen wir*

$$[\alpha]_x^y := \alpha(y) - \alpha(x) \qquad \forall\, x, y \in D.$$

- *Sind Γ_1, Γ_2 Stammfunktionen von γ, so gilt $\Gamma_1 - \Gamma_2 = w$ für ein $w \in \mathbb{R}^n$ (vgl. Terminologie 26). Ist weiterhin Γ eine Stammfunktion von γ, so auch $\Gamma + w$ für alle $w \in \mathbb{R}^n$. Für die Menge $\int \gamma$ aller Stammfunktionen von γ gilt somit*

$$\int \gamma = \{\Gamma + w \mid w \in \mathbb{R}^n\} \tag{395}$$

  für jede fixierte Stammfunktion Γ von γ. Insbesondere gilt $[\Gamma_1]_x^y = [\Gamma_2]_x^y$ für alle $\Gamma_1, \Gamma_2 \in \int \gamma$.

Der verallgemeinerte Hauptsatz garantiert, dass die Menge (395) nicht leer ist:

Korollar 55 (Verallgemeinerter Hauptsatz)**.** *Sei $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall, $a \in D$, und $\gamma\colon D \to \mathbb{R}$ eine stetige Kurve (ein Weg).*

1) *Die Abbildung*

$$\Gamma\colon D \to \mathbb{R}^n, \qquad x \mapsto \int_a^x \gamma$$

ist eine Stammfunktion von γ.

2) *Ist $\tilde{\Gamma} \in C^1(D,\mathbb{R})$ eine Stammfunktion von γ, so gilt*

$$\tilde{\Gamma}(x) - \tilde{\Gamma}(a) = [\tilde{\Gamma}]_a^x = [\Gamma]_a^x = \int_a^x \gamma \qquad \forall\, x \in D.$$

Beweis. Folgt sofort aus Satz 57 (Hauptsatz), angewandt auf die Komponentenfunktionen von γ. $\square$

Korollar 56 (Partielle Integration)**.** *Für $g \in C^1(D,\mathbb{R})$ und $\gamma \in C^1(D,\mathbb{R}^n)$, gilt*

$$\int_a^b g \cdot \gamma' = [g \cdot \gamma]_a^b - \int_a^b g' \cdot \gamma \qquad \forall\, a,b \in D.$$

Beweis. Folgt sofort durch komponentenweises Anwenden von Satz 58. $\square$

Korollar 57 (Substitutionsformel)**.** *Sei $\gamma \in C^0(D,\mathbb{R}^n)$ und $\phi \in C^1([a,b],\mathbb{R})$ mit $a < b$ und $\operatorname{im}(\phi) \subseteq D$. Dann gilt $(\gamma \circ \phi) \cdot \phi' \in C^0(D,\mathbb{R}^n)$ mit*

$$\int_a^b (\gamma \circ \phi) \cdot \phi' \;=\; \int_{\phi(a)}^{\phi(b)} \gamma.$$

Beweis. Folgt sofort durch komponentenweises Anwenden von Satz 60 (beachte $((\gamma \circ \phi) \cdot \phi')_j = (\gamma_j \circ \phi) \cdot \phi'$ für $1 \leq j \leq n$). $\square$

Aus der Intervalladditivität (392) erhalten wir eine veranschaulichende und nützliche Formel für das Riemann-integral einer stetigen Kurve:

Lemma 77. *Sei $\gamma \in C^0([a,b],\mathbb{R}^n)$ mit $\mathbb{R} \ni a < b \in \mathbb{R}$. Dann gilt*

$$\int_a^b \gamma = \lim_\ell \sum_{j=0}^{\ell-1} \frac{b-a}{\ell} \cdot \gamma\Big(\underbrace{a + j \cdot \frac{b-a}{\ell}}_{=:\,\tau[\ell]_j}\Big). \tag{396}$$

Beachte: *Für jedes $\ell \in \mathbb{N}_{>0}$ ist $Z_\ell = (\tau[\ell]_0, \ldots, \tau[\ell]_\ell)$ eine Unterteilung von $[a,b]$ mit*

$$\tau[\ell]_{j+1} - \tau[\ell]_j = \frac{b-a}{\ell} \qquad \textit{für alle} \qquad 0 \leq j \leq \ell - 1. \tag{397}$$

Beweis. Sei $\varepsilon > 0$ vorgegeben:

- Wegen Satz 81 (Satz 34) ist γ gleichmäßig stetig (bezüglich $\|\cdot\|_\infty$), d.h. es existiert $\delta > 0$ mit:

$$|s-t| < \delta \quad \text{für} \quad a \leq s,t \leq b \qquad \Longrightarrow \qquad \|\gamma(s) - \gamma(t)\|_\infty < \frac{\varepsilon}{b-a}.$$

- Sei $\mathbb{N} \ni N_\delta > \frac{b-a}{\delta}$. Dann gilt wegen (397):

$$\ell \geq N_\delta\colon\ 0 \leq j \leq \ell - 1\colon \quad |\tau[\ell]_j - \tau| \leq |\tau[\ell]_j - \tau[\ell]_{j+1}| \leq \frac{b-a}{N_\delta} < \delta \quad \text{für} \quad \tau \in [\tau[\ell]_j, \tau[\ell]_{j+1}]$$
$$\text{also} \tag{398}$$
$$\|\gamma(\tau[\ell]_j) - \gamma(\tau)\|_\infty < \frac{\varepsilon}{b-a}.$$

Mit der Intervalladditivität (a) und der Normierung (I4) (erster Schritt), der Linearität (b) und der Dreiecksungleichung (zweiter Schritt), der Abschätzung (393) (dritter Schritt), sowie der Monotonie (I2) und (398) (vierter Schritt), erhalten wir für alle $\ell \geq N_\delta$:

$$\begin{aligned}
\Big\| \int_a^b \gamma - \sum_{j=0}^{\ell-1} \frac{b-a}{\ell} \cdot \gamma(\tau[\ell]_j) \Big\|_\infty &= \Big\| \sum_{j=0}^{\ell-1} \int_{\tau[\ell]_j}^{\tau[\ell]_{j+1}} \gamma - \sum_{j=0}^{\ell-1} \int_{\tau[\ell]_j}^{\tau[\ell]_{j+1}} \gamma(\tau[\ell]_j) \Big\|_\infty \\
&\leq \sum_{j=0}^{\ell-1} \Big\| \int_{\tau[\ell]_j}^{\tau[\ell]_{j+1}} \big(\gamma - \gamma(\tau[\ell]_j)\big) \Big\|_\infty \\
&\overset{(393)}{\leq} \sum_{j=0}^{\ell-1} \Big| \int_{\tau[\ell]_j}^{\tau[\ell]_{j+1}} \|\gamma - \gamma(\tau[\ell]_j)\|_\infty \Big| \\
&\overset{(398)}{\leq} \sum_{j=0}^{\ell-1} \Big| \int_{\tau[\ell]_j}^{\tau[\ell]_{j+1}} \frac{\varepsilon}{b-a} \Big| = \sum_{j=0}^{\ell-1} \frac{b-a}{\ell} \cdot \frac{\varepsilon}{b-a} = \varepsilon. \qquad \square
\end{aligned}$$

Übung 147. *Sei $\mathbb{R} \ni c < d \in \mathbb{R}$, $a, b \in \{c, d\}$ mit $a \neq b$, sowie $\gamma \in C^0([a,b], \mathbb{R}^n)$. Zeigen Sie, dass für jede Norm $\|\cdot\|$ auf $\mathbb{R}^n$ gilt:*

$$\big\| \textstyle\int_a^b \gamma \big\| \leq \big| \textstyle\int_a^b \|\gamma\| \big|. \tag{399}$$

Hinweis: Sie müssen hier nur den Fall $a = c$ und $b = d$ abhandeln, da der andere Fall dann genau wie in Korollar 42 folgt. Benutzen Sie hierfür (396)*, die Dreicksungleichung, sowie die Stetigkeiten der involvierten Abbildungen.*

14.1.4 Die Bogenlänge von Kurven

Terminologie 38 (Bogenlänge)**.** *Sei $\gamma\colon D \to \mathbb{R}^n$ stetig differenzierbar (d.h. $\gamma \in C^1(D, \mathbb{R}^n)$). Für $D \ni a < b \in D$, definieren wir die* Bogenlänge zwischen γ(a) und γ(b) *durch*[99]

$$\mathsf{arc}[\gamma](a,b) := \int_a^b \|\gamma'\|_{\mathrm{euk}} \in [0, \infty)$$

Aus den üblichen Integrationsregeln erhalten wir die folgenden Eigenschaften:

1) *Wegen der Intervalladditivität des Riemann-integrals gilt:*

$$\mathsf{arc}[\gamma](a,b) = \mathsf{arc}[\gamma](a,c) + \mathsf{arc}[\gamma](c,b) \qquad \forall\, D \ni a \leq c \leq b \in D. \tag{400}$$

- *Für $D \ni a \leq a' < b' \leq b$ gilt somit*

$$\mathsf{arc}[\gamma](a,b) = \underbrace{\mathsf{arc}[\gamma](a,a')}_{\geq 0} + \mathsf{arc}[\gamma](a',b') + \underbrace{\mathsf{arc}[\gamma](b',b)}_{\geq 0} \geq \mathsf{arc}[\gamma](a',b'). \tag{401}$$

- *Ist $\gamma'|_{(a,b)} \neq 0$, so gilt $\mathsf{arc}[\gamma](a,b) > 0$.*

  Beweis. Per Annahme existiert ein $p \in (a,b)$ mit $\gamma'(p) \neq 0$, d.h., $C := \|\gamma'(p)\|_{\mathrm{euk}} > 0$. Wegen der Stetigkeit von $\|\gamma'\|_{\mathrm{euk}}$, existiert ein $\delta > 0$ mit $[p-\delta, p+\delta] \subseteq [a,b]$, sodass $\|\gamma'(\cdot)\|_{\mathrm{euk}}|_{[p-\delta,p+\delta]} \geq C/2$ gilt. Die Monotonie des Riemann-integrals (dritter Schritt) liefert

$$\mathsf{arc}[\gamma](a,b) \overset{(401)}{\geq} \mathsf{arc}[\gamma](p-\delta, p+\delta) = \textstyle\int_{p-\delta}^{p+\delta} \|\gamma'\|_{\mathrm{euk}} \geq \int_{p-\delta}^{p+\delta} C/2 = 2\delta \cdot C/2 = C \cdot \delta > 0. \qquad \square$$

[99]Beachte, dass $\|\gamma'\|_{\mathrm{euk}}$ stetig ist weil $\|\cdot\|_{\mathrm{euk}}$ stetig ist; und dass $\mathsf{arc}[\gamma](a,b) \in [0,\infty)$ wegen Beispiel 76 gilt.

Beispielsweise gilt $\gamma'|_{(a,b)} \neq 0$, *wenn* $[a,b]$ *nur abzählbar viele Nullstellen von* γ' *enthält (Übung).*

2) *Die Bogenlänge ist invariant unter* Umparametrisierungen*:*

Sei $\phi\colon [a',b'] \to [a,b]$ *eine stetig differenzierbare surjektive Abbildung mit* $\mathbb{R} \ni a' < b' \in \mathbb{R}$*:*

- *Sei* $\phi' \geq 0$ *mit* $\phi(a') = a$ *und* $\phi(b') = b$*. Wir erhalten mit Satz 60 (vorletzter Schritt):*

$$\begin{aligned}\mathsf{arc}[\gamma\circ\phi](a',b') &= \textstyle\int_{a'}^{b'} \|(\gamma\circ\phi)'(s)\|_{\text{euk}}\,\mathrm{d}s = \int_{a'}^{b'} \|\gamma'(\phi(s))\cdot\phi'(s)\|_{\text{euk}}\,\mathrm{d}s\\ &= \textstyle\int_{a'}^{b'} \|\gamma'(\phi(s))\|_{\text{euk}}\cdot|\phi'(s)|\,\mathrm{d}s \overset{\phi'\geq 0}{=} \int_{a'}^{b'} \underbrace{\|\gamma'(\phi(s))\|_{\text{euk}}}_{(\|\cdot\|_{\text{euk}}\circ\gamma')(\phi(s))}\cdot\phi'(s)\,\mathrm{d}s\\ &= \textstyle\int_a^b \|\gamma'(u)\|_{\text{euk}}\,\mathrm{d}u = \mathsf{arc}[\gamma](a,b).\end{aligned}$$

  Beachte: *Wegen* $\phi' \geq 0$ *und Korollar 32.1) ist* ϕ *monoton wachsend. Die Voraussetzung* $\phi(a') = a$ *und* $\phi(b') = b$ *folgt somit auch automatisch aus der Surjektivität von* ϕ *(Übung).*

- *Sei* $\phi' \leq 0$ *mit* $\phi(a') = b$ *und* $\phi(b') = a$*. Wir erhalten mit Satz 60 (fünfter Schritt):*

$$\begin{aligned}\mathsf{arc}[\gamma\circ\phi](a',b') &= \textstyle\int_{a'}^{b'} \|(\gamma\circ\phi)'(s)\|_{\text{euk}}\,\mathrm{d}s = \int_{a'}^{b'} \|\gamma'(\phi(s))\cdot\phi'(s)\|_{\text{euk}}\,\mathrm{d}s\\ &= \textstyle\int_{a'}^{b'} \|\gamma'(\phi(s))\|_{\text{euk}}\cdot|\phi'(s)|\,\mathrm{d}s \overset{\phi'\leq 0}{=} -\int_{a'}^{b'} \|\gamma'(\phi(s))\|_{\text{euk}}\cdot\phi'(s)\,\mathrm{d}s\\ &= \textstyle -\int_b^a \|\gamma'(u)\|_{\text{euk}}\,\mathrm{d}u = \int_a^b \|\gamma'(u)\|_{\text{euk}}\,\mathrm{d}u = \mathsf{arc}[\gamma](a,b).\end{aligned}$$

  Beachte: *Wegen* $\phi' \leq 0$ *und Korollar 32.1) ist* ϕ *monoton fallend. Die Voraussetzung* $\phi(a') = b$ *und* $\phi(b') = a$ *folgt somit auch automatisch aus der Surjektivität von* ϕ *(Übung).*

3) *Ist schließlich* D *von der Form* $[a,b]$ *mit* $\mathbb{R} \ni a < b \in \mathbb{R}$*, so definieren wir die* Gesamtbogenlänge von γ *durch* $\mathsf{Arc}[\gamma] := \mathsf{arc}[\gamma](a,b)$*. Zudem definieren wir die* Bogenlängenfunktion *durch*

$$\mathsf{arc}[\gamma]\colon [a,b] \to [0,\infty), \qquad t \mapsto \mathsf{arc}[\gamma](a,t) = \int_a^t \|\gamma'\|_{\text{euk}}.$$

   a) *Wegen dem Hauptsatz gilt dann* $\mathsf{arc}[\gamma] \in C^1([a,b],\mathbb{R})$ *mit* $\mathsf{arc}[\gamma]' = \|\gamma'\|_{\text{euk}}$.
   b) *Wegen* 1) *ist* $\mathsf{arc}[\gamma]$ *monoton wachsend, d.h.* $\mathrm{im}(\mathsf{arc}[\gamma]) = [0,\mathsf{Arc}[\gamma]]$.

      Hat γ' *nur abzählbar viele Nullstellen, so ist* $\mathsf{arc}[\gamma]$ *sogar streng monoton wachsend (Übung).*
   c) *Wegen* 2) *gilt für jede surjektive* C^1*-Abbildung* $\phi\colon [a',b'] \to [a,b]$ *mit* $\phi' \geq 0$*:*

$$\mathsf{Arc}[\gamma\circ\phi] = \mathsf{Arc}[\gamma] \qquad \textit{sowie} \qquad \mathsf{arc}[\gamma\circ\phi] = \mathsf{arc}[\gamma]\circ\phi.$$

Übung 148. *Machen Sie sich die Aussagen* a), b), c) *in Terminologie 38.*3) *klar. Zeigen Sie zudem, dass im Kontext von Terminologie 38.*2) *die Formel* $\mathsf{arc}[\gamma\circ\phi](a',b') = \mathsf{arc}[\gamma](a,b)$ *mit* $\phi\colon [a',b'] \to [a,b]$ *stetig differenzierbar und surjektiv, ohne eine der Voraussetzung* $\phi' \geq 0$ *bzw.* $\phi' \leq 0$ *im Allgemeinen nicht gilt.*

Die nächste Proposition (genauer Korollar 59) liefert eine Anschauung für den Begriff der Bogenlänge: Für $\ell \in \mathbb{N}_{>0}$, betrachten wir die Unterteilung $Z_\ell = (\tau[\ell]_0,\ldots,\tau[\ell]_\ell)$ von $[a,b]$ aus Lemma 77, d.h.

$$\tau[\ell]_j := a + j\cdot\frac{b-a}{\ell} \quad \text{für} \quad 0 \leq j \leq \ell \qquad \text{mit} \qquad \tau[\ell]_{j+1} - \tau[\ell]_j = \frac{b-a}{\ell} \qquad \text{für} \qquad 0 \leq j \leq \ell-1.$$

Proposition 19. *Sei* $\gamma\colon [a,b] \to \mathbb{R}^n$ *stetig differenzierbar. Dann gilt*

$$\mathsf{Arc}[\gamma] = \lim_\ell \underbrace{\sum_{j=0}^{\ell-1} \|\gamma(\tau[\ell]_{j+1}) - \gamma(\tau[\ell]_j)\|_{\text{euk}}}_{=:\,\Omega[\ell]}. \tag{402}$$

(Die Summe auf der rechten Seite von (402) ist die Länge des Polygonzuges (entlang γ) zwischen den Verbindungspunkten $\gamma(a) = \gamma(\tau[\ell]_0)$, $\gamma(\tau[\ell]_1)$, $\ldots$, $\gamma(\tau[\ell]_{\ell-1})$, $\gamma(\tau[\ell]_\ell) = \gamma(b)$.)

Beweis. Lemma 77, angewandt auf die stetige Abbildung (Kurve) $\|\cdot\|_{\mathrm{euk}} \circ \gamma' \colon [a,b] \to \mathbb{R}$, zeigt

$$\mathsf{Arc}[\gamma] = \lim_\ell \underbrace{\sum_{j=0}^{\ell-1} \frac{b-a}{\ell} \cdot \|\gamma'(\tau[\ell]_j)\|_{\mathrm{euk}}}_{=: \Lambda[\ell]} . \tag{403}$$

Die Behauptung folgt somit umgehend aus der Dreicksungleichung sobald wir nachgewiesen haben, dass zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}_{>0}$ existiert mit $\quad (|\mathsf{Arc}[\gamma] - \Omega[\ell]| \leq |\mathsf{Arc}[\gamma] - \Lambda[\ell]| + |\Lambda[\ell] - \Omega[\ell]|)$

$$|\Omega[\ell] - \Lambda[\ell]| < \varepsilon \qquad \forall\, \ell \geq N_\varepsilon.$$

Dies folgt nun aber aus dem Mittelwertsatz:

Wir wählen $C > 0$ mit $\|\cdot\|_{\mathrm{euk}} \leq C \cdot \|\cdot\|_\infty$, schreiben $\gamma = (\gamma_1, \ldots, \gamma_n)$, und fixieren $\varepsilon > 0$.

- Seien $\ell \in \mathbb{N}_{>0}$, $0 \leq j \leq \ell - 1$, sowie $1 \leq p \leq n$ vorgegeben. Der Mittelwertsatz (Satz 44) liefert ein

$$\tau[\ell]_j < \xi[\ell,j,p] < \tau[\ell]_{j+1} \qquad \text{mit} \qquad \gamma_p'(\xi[\ell,j,p]) \;=\; \underbrace{\frac{\gamma_p(\tau[\ell]_{j+1}) - \gamma_p(\tau[\ell]_j)}{\tau[\ell]_{j+1} - \tau[\ell]_j}}_{\frac{\ell}{b-a} \cdot (\gamma_p(\tau[\ell]_{j+1}) - \gamma_p(\tau[\ell]_j))} . \tag{404}$$

- Wegen Satz 34 ist γ_p' gleichmäßig stetig für $1 \leq p \leq n$. Somit existiert ein $\delta > 0$ mit

$$\forall\, 1 \leq p \leq n\colon \qquad |x - y| < \delta \quad \text{für} \quad x, y \in [a,b] \qquad \Longrightarrow \qquad |\gamma_p'(x) - \gamma_p'(y)| < \frac{\varepsilon}{C \cdot (b-a)}.$$

  Wir wählen $N_\varepsilon \in \mathbb{N}_{>0}$ mit $\frac{b-a}{N_\varepsilon} < \delta$, und erhalten

$$|\gamma_p'(\tau[\ell]_j) - \gamma_p'(\xi[\ell,j,p])| < \frac{\varepsilon}{C \cdot (b-a)} \qquad \forall\, 1 \leq p \leq n,\ \ell \geq N_\varepsilon,\ 0 \leq j \leq \ell - 1. \tag{405}$$

Für $\ell \geq N_\varepsilon$ folgt mit der Dreiecksungleichung für $|\cdot|$ (zweiter Schritt) sowie der inversen Dreiecksungleichung für $\|\cdot\|_{\mathrm{euk}}$ (dritter Schritt):

$$\begin{aligned}
|\Omega[\ell] - \Lambda[\ell]| \;&=\; \Big|\textstyle\sum_{j=0}^{\ell-1} \big\|\gamma(\tau[\ell]_{j+1}) - \gamma(\tau[\ell]_j)\big\|_{\mathrm{euk}} - \big\|\frac{b-a}{\ell} \cdot \gamma'(\tau[\ell]_j)\big\|_{\mathrm{euk}}\Big| \\
&\leq\; \textstyle\sum_{j=0}^{\ell-1} \Big|\big\|\gamma(\tau[\ell]_{j+1}) - \gamma(\tau[\ell]_j)\big\|_{\mathrm{euk}} - \big\|\frac{b-a}{\ell} \cdot \gamma'(\tau[\ell]_j)\big\|_{\mathrm{euk}}\Big| \\
&\leq\; \textstyle\sum_{j=0}^{\ell-1} \big\|\gamma(\tau[\ell]_{j+1}) - \gamma(\tau[\ell]_j) - \frac{b-a}{\ell} \cdot \gamma'(\tau[\ell]_j)\big\|_{\mathrm{euk}} \\
&\leq\; C \cdot \textstyle\sum_{j=0}^{\ell-1} \big\|\gamma(\tau[\ell]_{j+1}) - \gamma(\tau[\ell]_j) - \frac{b-a}{\ell} \cdot \gamma'(\tau[\ell]_j)\big\|_\infty \\
&\overset{(404)}{=}\; C \cdot \textstyle\sum_{j=0}^{\ell-1} \frac{b-a}{\ell} \cdot \underbrace{\max\Big(\big|\gamma_p'(\xi[\ell,j,p]) - \gamma_p'(\tau[\ell]_j)\big| \;\Big|\; 1 \leq p \leq n\Big)}_{\overset{(405)}{<} \frac{\varepsilon}{C \cdot (b-a)}} < \varepsilon.
\end{aligned}$$

□

Korollar 58. *Sei $\gamma\colon [a,b] \to \mathbb{R}^n$ stetig differenzierbar. Dann gilt*

$$\|\gamma(a) - \gamma(b)\|_{\mathrm{euk}} \leq \mathsf{Arc}[\gamma].$$

Beweis. Wegen Proposition 19 gilt

$$\mathsf{Arc}[\gamma] = \lim_{\ell} \overbrace{\textstyle\sum_{j=0}^{\ell-1} \|\gamma(\tau[\ell]_{j+1}) - \gamma(\tau[\ell]_j)\|_{\mathrm{euk}}}^{\Omega[\ell]}.$$

Für $\varepsilon > 0$ vorgegeben, existiert somit ein $\ell_\varepsilon \in \mathbb{N}_{>0}$ mit $\Omega[\ell_\varepsilon] \leq \mathsf{Arc}[\gamma] + \varepsilon$. Die Dreiecksungleichung (zweiter Schritt) liefert

$$\|\gamma(a) - \gamma(b)\|_{\mathrm{euk}} = \Big\| \big(\textstyle\sum_{j=0}^{\ell_\varepsilon-1} \gamma(\tau[\ell_\varepsilon]_{j+1}) - \gamma(\tau[\ell_\varepsilon]_j) \big) \Big\|_{\mathrm{euk}} \leq \Omega[\ell_\varepsilon] \leq \mathsf{Arc}[\gamma] + \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt die Behauptung aus Lemma 19. □

Korollar 59. *Sei $\gamma\colon [a,b] \to \mathbb{R}^n$ stetig differenzierbar. Dann gilt*

$$\mathsf{Arc}[\gamma] = \sup_{\mathbb{R}} \Big(\underbrace{\big\{ \textstyle\sum_{j=0}^{\ell-1} \|\gamma(\tau_{j+1}) - \gamma(\tau_j)\|_{\mathrm{euk}} \,\big|\, (\tau_0, \dots, \tau_\ell) \text{ mit } \ell \in \mathbb{N}_{>0} \text{ ist Zerlegung von } [a,b] \big\}}_{=:\, \mathcal{P} \;\equiv\; \text{Menge aller Längen von Polygonzügen entlang } \gamma.} \Big).$$

Beweis. Offensichtlich gilt $\mathcal{P} \neq \emptyset$. Wir erhalten mit der Additivität (400) sowie Korollar 58, dass

$$\mathsf{Arc}[\gamma] \overset{(400)}{=} \sum_{j=0}^{\ell-1} \underbrace{\mathsf{arc}[\gamma](\tau_j, \tau_{j+1})}_{\mathsf{Arc}[\gamma|_{[\tau_j,\tau_{j+1}]}]} \geq \sum_{j=0}^{\ell-1} \|\gamma(\tau_{j+1}) - \gamma(\tau_j)\|_{\mathrm{euk}}$$

für jede Zerlegung $(\tau_0, \dots, \tau_\ell)$ von $[a,b]$ gilt; also $\mathcal{P} \leq \mathsf{Arc}[\gamma]$, und somit $\sup_{\mathbb{R}}(\mathcal{P}) \leq \mathsf{Arc}[\gamma]$. Wegen Proposition 19 existiert nun die Folge $(\Omega[\ell])_{\ell \in \mathbb{N}_{>0}}$ in $\mathcal{P}$ mit $\lim_\ell \Omega[\ell] = \mathsf{Arc}[\gamma]$, womit $\sup_{\mathbb{R}}(\mathcal{P}) = \mathsf{Arc}[\gamma]$ wegen Übung 64 gilt. □

Beispiel 101. *Wie in Beispiel 99.c) bereits angekündigt, erhalten wir nun schließlich die Interpretation der Kreiszahl π als die Länge des oberen Halbkreisbogens. Wir betrachten wieder die Kreiskurve*

$$\gamma_\Psi\colon \mathbb{R} \to \mathbb{R}^2, \qquad x \mapsto (\cos(x), \sin(x))$$

aus Beispiel 99.c), sowie ihre Einschränkung $\gamma_\circ = \gamma_\Psi|_{[0,2\pi)}$. Dann gilt für alle $\mathbb{R} \ni a < b \in \mathbb{R}$

$$\mathsf{arc}[\gamma_\psi](a,b) = \textstyle\int_a^b \|\gamma_\psi'\|_{\mathrm{euk}} = \int_a^b \|(-\sin, \cos)\|_{\mathrm{euk}} = \int_a^b \sqrt{\sin^2 + \cos^2} = \int_a^b 1 = b - a.$$

Gemäß Korollar 59 ist nun aber

$$\mathsf{Arc}[\gamma_\circ|_{[0,\pi]}] = \mathsf{arc}[\gamma_\psi](0,\pi) = \pi - 0 = \pi$$

gerade das Supremum über alle Längen von Polygonzügen entlang der Parametrisierung $\gamma_\circ|_{[0,\pi]}$ des oberen Halbkreisbogens, welche diesen gemäß Beispiel 99.c) entgegen dem Uhrzeigersinn durchläuft.

Terminologie 39 (Parametrisierung nach Bogenlänge)**.**

1) *Eine differenzierbare Kurve $\gamma\colon D \to \mathbb{R}^n$ heißt* immersiv*, wenn $\gamma'(t) \neq 0$ für alle $t \in D$ gilt, d.h. wenn $\|\gamma'\|_{\mathrm{euk}}$ keine Nullstellen besitzt.*

   Beispiel: *Alle Kurven in Beispiel 99 sind immersiv, sofern man in Teil a) $v \neq 0$, sowie in Teil b) f differenzierbar wählt.*

2) *Eine immersive Kurve $\gamma\colon D \to \mathbb{R}^n$ heißt* nach Bogenlänge parametrisiert*, wenn $\|\gamma'\|_{\mathrm{euk}} = 1$ gilt.*

   Behauptung: *Für $\gamma \in C^1([a,b], \mathbb{R}^n)$ mit $a < b$, ist dies gleichbedeutend mit $\mathsf{arc}[\gamma] = \mathrm{id}_{[a,b]} - a$:*

   Beweis. • Gilt $\|\gamma'\|_{\mathrm{euk}} = 1$, so folgt

   $$\mathsf{arc}[\gamma](t) = \textstyle\int_a^t \|\gamma'\|_{\mathrm{euk}} = \int_a^t 1 = t - a = (\mathrm{id}_{[a,b]} - a)(t) \qquad \forall\, t \in [a,b].$$

- Gilt $\mathsf{arc}[\gamma] = \mathrm{id}_{[a,b]} - a$, so folgt $\|\gamma'\|_{\mathrm{euk}} = \mathsf{arc}[\gamma]' = 1$ aus Terminologie 38.3).a). □

Beispielsweise ist die Kreiskurve γ_ψ aus Beispiel 101 nach Bogenlänge parametrisiert.

Proposition 20 (Parametrisierung nach Bogenlänge)**.** *Sei $\gamma \in C^1([a,b],\mathbb{R}^n)$ immersiv.*

1) $\mathsf{arc}[\gamma]\colon [a,b] \to [0,\mathsf{Arc}[\gamma]]$ *ist streng monoton wachsend, surjektiv und stetig differenzierbar mit* $\mathsf{arc}[\gamma]' = \|\gamma'\|_{\mathrm{euk}} > 0$.
2) $\phi := \mathsf{arc}[\gamma]^{-1}\colon [0,\mathsf{Arc}[\gamma]] \to [a,b]$ *ist streng monoton wachsend, surjektiv und stetig differenzierbar; und $\gamma \circ \phi$ ist nach Bogenlänge parametrisiert.*
3) *Sei $\psi\colon [0,c] \to [a,b]$ $(c > 0)$ streng monoton wachsend, surjektiv und stetig differenzierbar, sodass $\gamma \circ \psi$ nach Bogenlänge parametrisiert ist. Dann gilt bereits $\psi = \mathsf{arc}[\gamma]^{-1}$ (also auch $c = \mathsf{Arc}[\gamma]$).*

Beweis. 1) Wegen Terminologie 38.3).a) ist $\mathsf{arc}[\gamma]$ stetig differenzierbar, mit $\mathsf{arc}[\gamma]' = \|\gamma'\|_{\mathrm{euk}} > 0$ (beachte $\gamma' \neq 0$). Wegen Korollar 32.3) ist daher $\mathsf{arc}[\gamma]$ streng monoton wachsend, und somit gilt zwangsweise $\mathsf{arc}[\gamma]([a,b]) = [\mathsf{arc}[\gamma](a),\mathsf{arc}[\gamma](b)] = [0,\mathsf{Arc}[\gamma]]$ (alternativ Terminologie 38.3).b)).

2) Wegen Teil 1) und Lemma 43 ist $\mathsf{arc}[\gamma]$ injektiv; und es gilt $\mathsf{arc}[\gamma]' > 0$. Wegen Satz 42 hat $\mathsf{arc}[\gamma]$ daher eine differenzierbare (also stetige) Umkehrabbildung $\phi := \mathsf{arc}[\gamma]^{-1}\colon [0,\mathsf{Arc}[\gamma]] \to [a,b]$, mit

$$\phi' = \frac{1}{\mathsf{arc}[\gamma]'(\phi)} = \frac{1}{\|\gamma' \circ \phi\|_{\mathrm{euk}}} > 0. \tag{406}$$

- Korollar 32.3) zusammen mit (406) zeigt, dass ϕ streng monoton wachsend ist.[100]
- ϕ' ist stetig, weil alle Abbildungen auf der rechten Seite von (406) stetig sind. Daher ist ϕ stetig differenzierbar mit

$$\|(\gamma \circ \phi)'\|_{\mathrm{euk}} = \|(\gamma' \circ \phi) \cdot \phi'\|_{\mathrm{euk}} = \|\gamma' \circ \phi\|_{\mathrm{euk}} \cdot |\phi'| = \|\gamma' \circ \phi\|_{\mathrm{euk}} \cdot \phi' = 1.$$

3) Terminologie 39.2) (erster Schritt) und Terminologie 38.3).c) (zweiter Schritt) liefern

$$\mathrm{id}_{[0,c]} = \mathsf{arc}[\gamma \circ \psi] = \mathsf{arc}[\gamma] \circ \psi.$$

Somit ist ψ injektiv; also bijektiv da per Voraussetzung surjektiv. Daher gilt $\psi = \mathsf{arc}[\gamma]^{-1}$. □

Übung 149. *Sei $\gamma\colon [a,b] \to \mathbb{R}^n$ eine nichtimmersive differenzierbare Kurve. Zeigen Sie, dass keine differenzierbare surjektive Abbildung $\phi\colon [a',b'] \to [a,b]$ mit $\mathbb{R} \ni a' < b' \in \mathbb{R}$ existieren kann, sodass $\|(\gamma \circ \phi)'\|_{\mathrm{euk}} = 1$ gilt. (Eine nichtimmersive Kurve kann also nicht nach Bogenlänge parametrisiert werden).*

14.2 Differenzierbare Abbildungen

In diesem Abschnitt erweitern wir das Konzept der differenzierbaren Abbildung in einer Variablen (differenzierbare Kurve) auf den Fall von Abbildungen in mehrerern Veränderlichen. Genauer betrachten wir Abbildungen $f\colon \mathbb{R}^m \supseteq U \to \mathbb{R}^n$ mit $m,n \in \mathbb{N}_{>0}$, wobei nun $\emptyset \neq U \subseteq \mathbb{R}^m$ immer als nichtleer und offen vorausgesetzt wird. Grob gesprochen garantiert die Offenheit von U, dass f in jedem fixierten $p \in U$ in jede Richtung $v \in \mathbb{R}^m$ abgeleitet werden kann, bzw. dass zunächst für jedes $v \in \mathbb{R}^m$ ein $\delta > 0$ existiert, sodass der Differenzenquotient

$$\frac{f(p + t \cdot v) - f(p)}{t} \qquad \forall\, 0 \neq t \in (-\delta,\delta) \tag{407}$$

definiert ist. Dies ist auch essentiell für die Eindeutigkeit der Ableitung (Existenz vorausgesetzt), die im mehrdimensionalen Fall nun keine Zahl mehr ist, sondern ein Element in $\mathrm{Hom}(\mathbb{R}^m,\mathbb{R}^n)$, also eine Lineare Abbildung, die die Funktion lokal um einen Punkt approximiert.

[100]Alternativ folgt das auch aus Satz 37, da $\mathsf{arc}[\gamma]$ streng monoton wachsend ist.

14.2.1 Richtungableitungen und das Differential einer Abbildung

Im Folgenden sei $m, n \in \mathbb{N}_{>0}$.

Bemerkung 92. *Sei $U \subseteq \mathbb{R}^m$ offen und $p \in U$ ein Punkt.*

a) *Es gilt:* $$\forall\, v \in \mathbb{R}^m \colon \exists\, \delta_v > 0 \colon \quad p + (-\delta_v, \delta_v) \cdot v \subseteq U. \tag{408}$$

Beweis. Sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^m$. Da U offen ist, existiert ein $\varepsilon > 0$ mit $\mathsf{B}_\varepsilon(p) \subseteq U$.

- Für $v = 0$ gilt die Behauptung für jedes $\delta_v > 0$.
- Für $v \neq 0$ gilt die Behauptung für $\delta_v := \frac{\varepsilon}{\|v\|}$. □

b) *Seien $I, \tilde{I} \subseteq \mathbb{R}$ offene Intervalle mit $t \in I \cap \tilde{I}$, sowie $\gamma \colon I \to \mathbb{R}^n$ und $\tilde{\gamma} \colon \tilde{I} \to \mathbb{R}^n$ Kurven mit $\gamma|_{I \cap \tilde{I}} = \tilde{\gamma}|_{I \cap \tilde{I}}$. Übung 92 angewandt auf die Komponentenfunktionen beider Kurven zeigt:*

$$\gamma \text{ differenzierbar in } t \qquad \Longleftrightarrow \qquad \tilde{\gamma} \text{ differenzierbar in } t$$

mit

$$\lim_{h \to 0} \frac{\gamma(t+h) - \gamma(t)}{h} \overset{\text{Lemma 76}}{=} \gamma'(t) \overset{\text{Übung 92}}{=} \tilde{\gamma}'(t) \overset{\text{Lemma 76}}{=} \lim_{h \to 0} \frac{\tilde{\gamma}(t+h) - \tilde{\gamma}(t)}{h}.$$

Übung 150. *Machen Sie sich die Aussage in Bemerkung 92.b) klar.*

Terminologie 40 (Richtungsableitung)**.** *Sei $U \subseteq \mathbb{R}^m$ offen, $f \colon U \to \mathbb{R}^n$ eine Abbildung, sowie $p \in U$ und $v \in \mathbb{R}^m$. Die Existenz vorausgesetzt, heißt*

$$D_v f(p) := \lim_{t \to 0} \frac{f(p + t \cdot v) - f(p)}{t} \in \mathbb{R}^n \tag{409}$$

die Ableitung von f in p in Richtung v*. Die rechte Seite ist formal exakt definiert durch:*

$$\lim_{t \to 0} \frac{f(p + t \cdot v) - f(p)}{t} := \lim_{t \to 0} \frac{\gamma[v](t) - \gamma[v](0)}{t} \overset{\text{Lemma 76}}{=} \gamma[v]'(0) \tag{410}$$

mit der Kurve

$$\gamma[v] \colon (-\delta_v, \delta_v) \to \mathbb{R}^n, \qquad t \mapsto f(p + t \cdot v) \tag{411}$$

wobei $\delta_v > 0$ derart gewählt ist, dass $p + (-\delta_v, \delta_v) \cdot v \subseteq U$ gilt.

Beachte: *Wegen Bemerkung 92.a) existiert ein derartiges $\delta_v > 0$; und wegen Bemerkung 92.b) ist die rechte Seite von (410) unabhängig von der Wahl von $\delta_v > 0$.*

Existiert $D_v f(p)$ für alle $p \in U$ für ein vorgegebenes $v \in \mathbb{R}^m$, so wird die Abbildung

$$D_v f \colon U \to \mathbb{R}^n, \qquad p \mapsto D_v f(p) \tag{412}$$

als die Ableitung von f in Richtung v bezeichnet.

(1) **Rechenregeln:** *Für $p \in U$ und $v \in \mathbb{R}^m$ sei $\gamma[v]$ wie in (411).*

a) *Es existiert $D_v f(p)$ genau dann, wenn $D_v f_j(p)$ für jedes $1 \leq j \leq n$ existiert; und dann gilt*

$$D_v f(p) = (D_v f_1(p), \ldots, D_v f_n(p)). \tag{413}$$

Beweis. Sei $\delta_v > 0$ wie in Bemerkung 92.a). Dann gilt für alle $t \in (-\delta_v, \delta_v)$:

$$(\gamma[v]_1(t), \ldots, \gamma[v]_n(t)) = \gamma[v](t) = f(p + t \cdot v) = (f_1(p + t \cdot v), \ldots, f_n(p + t \cdot v))$$

also

$$D_v f(p) \overset{(410)}{=} \gamma[v]'(0) = (\gamma[v]_1'(0), \ldots, \gamma[v]_n'(0)) \overset{(410)}{=} (D_v f_1(p), \ldots, D_v f_n(p))$$

im jeweiligen Falle der Existenz einer der beiden Seiten (in der letzten Zeile), wobei (410) auf der rechten Seite komponentenweise angewendet wird. □

b) *Existiert $D_v f(p)$, so liefert die Kettenregel (Bemerkung 90.c)):*

$$D_{\lambda \cdot v} f(p) = \lambda \cdot D_v f(p) \qquad \forall\, \lambda \in \mathbb{R}. \tag{414}$$

Die Richtungsableitung $D_v f(p)$ existiert also genau dann, wenn für alle $\lambda \in \mathbb{R}$ die Richtungsableitungen $D_{\lambda \cdot v} f(p)$ existieren.

Beweis von (414). Für $\lambda \in \mathbb{R}$, sei $\delta_{\lambda \cdot v} = \frac{\delta_v}{1+|\lambda|}$ sowie

$$\gamma[\lambda \cdot v] \colon (-\delta_{\lambda \cdot v}, \delta_{\lambda \cdot v}) \ni t \mapsto f(p + t \cdot \lambda \cdot v) \in \mathbb{R}^n.$$

Dann gilt $\gamma[\lambda \cdot v] := \gamma[v] \circ \phi_\lambda$ für $\phi_\lambda \colon (-\delta_{\lambda \cdot v}, \delta_{\lambda \cdot v}) \ni t \mapsto \lambda \cdot t \in (-\delta_v, \delta_v)$, und wir erhalten mit der Kettenregel (Bemerkung 90.c)) im dritten Schritt:

$$\lambda \cdot D_v f(p) = \lambda \cdot \gamma[v]'(0) = \phi_\lambda'(0) \cdot \gamma[v]'(\phi_\lambda(0)) = \gamma[\lambda \cdot v]'(p) \overset{(410)}{=} D_{\lambda \cdot v} f(p). \qquad \square$$

c) *Seien $h, g \colon U \to \mathbb{R}^n$ Abbildungen, sodass die Richtingsableitungen $D_v h(p), D_v g(p)$ existieren. Wegen Bemerkung 90.a) existiert dann auch die Richtungsableitung $D_v(\lambda \cdot h + \mu \cdot g)(p)$, und es gilt*

$$D_v(\lambda \cdot h + \mu \cdot g)(p) = \lambda \cdot D_v h(p) + \mu \cdot D_v g(p) \qquad \forall\, \lambda, \mu \in \mathbb{R}.$$

d) *Seien $h \colon U \to \mathbb{R}$ und $g \colon U \to \mathbb{R}^n$ Abbildungen, sodass die Richtungsableitungen $D_v h(p), D_v g(p)$ existieren. Wegen Bemerkung 90.c) existiert dann auch die Richtungsableitung $D_v(h \cdot g)(p)$, und es gilt*

$$D_v(h \cdot g)(p) = D_v h(p) \cdot g(p) + h(p) \cdot D_v g(p).$$

(2) **Partielle Ableitungen:** *Seien $e_1, \ldots, e_m$ die kanonischen Basisvektoren von $\mathbb{R}^m$. Die* j-te partielle Ableitung von f in p *ist definiert (Existenz vorausgesetzt) als die Richtungsableitung*

$$\partial_j f(p) := D_{e_j} f(p) \qquad \text{für} \qquad 1 \le j \le m. \tag{415}$$

Es existiert $\partial_j f(p)$ genau dann, wenn $\partial_j f_i(p)$ für $1 \le i \le n$ existiert, und es gilt

$$\partial_j f(p) \overset{(413)}{=} (\partial_j f_1(p), \ldots, \partial_j f_n(p)). \tag{416}$$

Existiert für ein $1 \le j \le m$ die partielle Ableitung $\partial_j f(p)$ für alle $p \in U$, so wird die Abbildung

$$\partial_j f \colon U \to \mathbb{R}^n, \qquad p \mapsto \partial_j f(p)$$

als die j-te partielle Ableitung von f *bezeichnet.*

Eine alternative Notation für die j-te partielle Ableitung ist:

$$\partial_j f(p) \equiv \frac{\partial f}{\partial x_j}(p) \qquad \text{bzw.} \qquad \partial_j f \equiv \frac{\partial f}{\partial x_j} \qquad \text{für} \qquad 1 \le j \le m.$$

Übung 151. *Zeigen Sie die Aussagen* c) *und* d) *in Terminologie 40.*(1).

Übung 152 (Einschränkungen)**.** *Seien* $V, U \subseteq \mathbb{R}^m$ *offen mit* $V \subseteq U$. *Sei* $f\colon U \to \mathbb{R}^n$ *eine Abbildung, sowie* $p \in V$ *und* $v \in \mathbb{R}^m$. *Zeigen Sie, dass* $D_v f(p) \in \mathbb{R}^n$ *genau dann existiert, wenn* $D_v(f|_V)(p) \in \mathbb{R}^n$ *existiert, mit* $D_v f(p) = D_v(f|_V)(p)$.

Terminologie 41 (Differential einer Abbildung)**.** *Sei* $U \subseteq \mathbb{R}^m$ *offen, und* $f\colon U \to \mathbb{R}^n$ *eine Abbildung. Sei weiterhin* $\|\cdot\|$ *eine Norm auf* $\mathbb{R}^m$.

(1) f *heißt* differenzierbar in p ∈ U, *wenn ein* $\phi \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ *existiert sowie eine in* 0 *stetige Abbildung* $\varepsilon\colon U - p \to \mathbb{R}^n$ *mit* $\varepsilon(0) = 0$, *sodass gilt:*

$$f(p+h) = \underbrace{f(p) + \phi(h)}_{\text{affine Abbildung}} + \underbrace{\|h\| \cdot \varepsilon(h)}_{\text{Restglied } R(h)} \qquad \forall\, h \in U - p. \tag{417}$$

(*Beachte: Für das Restglied gilt* $\lim_{h\to 0} \frac{R(h)}{\|h\|} = 0$.)

Es gelten die folgenden Notationen und Sachverhalte:

a) **Eindeutigkeit des Differentials:** *Die Abbildung* ϕ *ist eindeutig bestimmt; und zwar durch:*

$$\phi(v) = D_v f(p) \qquad \forall\, v \in \mathbb{R}^m. \tag{418}$$

Sie heißt die Ableitung- *bzw.* das Differential von f in p *und wird mit* $\mathrm{d}_p f$ *notiert, d.h.*

$$\mathrm{d}_p f(v) = D_v f(p) \qquad \forall\, v \in \mathbb{R}^m. \tag{419}$$

Beachte: *Wegen* $\varepsilon|_{U-p} = \left(\frac{f - f(p) - \phi}{\|\cdot\|}\right)\Big|_{U-p}$ *und* $\varepsilon(0) = 0$, *ist dann auch* ε *eindeutig bestimmt.*

Beweis der Eindeutigkeitsbedingung (418)*.* Für $v \in \mathbb{R}^m$, sei $\delta_v > 0$ wie (408). Dann gilt

$$\frac{f(p+t\cdot v) - f(p)}{t} = \underbrace{\frac{1}{t}\cdot\phi(t\cdot v)}_{\phi(v)} + \frac{|t|}{t}\cdot\|v\|\cdot\varepsilon(t\cdot v) \qquad \forall\, 0 \neq t \in (-\delta_v, \delta_v)$$

$$\Longrightarrow \qquad D_v f(p) = \lim_{t\to 0}\frac{f(p+t\cdot v) - f(p)}{t} = \lim_{t\to 0}\left(\phi(v) + \frac{|t|}{t}\cdot\|v\|\cdot\varepsilon(t\cdot v)\right) = \phi(v). \qquad \square$$

b) **Unabhängigkeit von der Wahl der Norm:** *Die Bedingung* (417) *ist unabhängig von der Wahl der Norm* $\|\cdot\|$ *auf* $\mathbb{R}^m$.[101]

Beweis. Es gelte (417), und es sei $\|\cdot\|'$ eine weitere Norm auf $\mathbb{R}^m$. Dann gilt $\|\cdot\| \leq C\cdot\|\cdot\|'$ für ein $C > 0$ wegen Satz 79. Wir setzen

$$\varepsilon'\colon U - p \to \mathbb{R}^n, \qquad h \mapsto \begin{cases} \frac{\|h\|}{\|h\|'}\cdot\varepsilon(h) & \text{für} \quad 0 \neq h \in (U-p) \\ 0 & \text{für} \quad h = 0. \end{cases}$$

Dann gilt $\|h\|'\cdot\varepsilon'(h) = \|h\|\cdot\varepsilon(h)$, sowie $\lim_{h\to 0}\varepsilon'(h) = 0$ wegen $\frac{\|h\|}{\|h\|'} \leq C$. $\square$

c) **Lokalität:** *Sei* $V \subseteq \mathbb{R}^m$ *offen mit* $p \in V \subseteq U$. *Es ist* f *genau dann differenzierbar in* p, *wenn die Einschränkung* $\tilde{f} := f|_V$ *differenzierbar in* p *ist.*

[101] Die Bedingung (417) ist natürlich auch unabhängig von der Wahl der Norm auf $\mathbb{R}^n$; denn die Stetigkeit von ε in 0 ist ja unabhängig von dieser Wahl.

Beweis. • Ist f differenzierbar in p, so gilt (417) auch für $\tilde{f} = f|_V$, sofern man ε durch $\tilde{\varepsilon} := \varepsilon|_{V-p}$ ersetzt (wegen Lemma 40.1) ist $\tilde{\varepsilon}$ ebenfalls stetig in p).

• Sei $\tilde{f} = f|_V$ differenzierbar in p, d.h. es gilt (417) für $\tilde{f}$ mit einem zugehörigen $\tilde{\varepsilon}$. Dann gilt (417) für f und

$$\varepsilon\colon U-p \to \mathbb{R}^n, \qquad h \mapsto \begin{cases} \dfrac{f(p+h)-f(p)-\phi(h)}{\|h\|} & \text{für} \quad h \in (U-p)\setminus(V-p) \\ \tilde{\varepsilon}(h) & \text{für} \quad h \in V-p, \end{cases}$$

wobei ε stetig in 0 ist wegen Übung 79.a) (wegen (140) in Bemerkung 36.a) ist $V-p$ offen im metrischen Raum $(U-p, \mathrm{d}_{U-p})$ mit einer Normmetrik d auf $\mathbb{R}^m$, denn $V-p$ ist offen in $\mathbb{R}^m$ mit $V-p = (V-p)\cap(U-p)$). □

d) **Jacobimatrix:** *Sei $\varsigma_1, \ldots, \varsigma_m$ eine Basis von $\mathbb{R}^m$, und sei $u = \sum_{j=1}^m u_j \cdot \varsigma_j \in \mathbb{R}^m$ mit $u_1, \ldots, u_m \in \mathbb{R}$ ein Vektor. Dann gilt wegen der Linearität von $\mathrm{d}_p f$ und (419):*

$$\mathrm{d}_p f(u) = \mathrm{d}_p f\left(\sum_{j=1}^m u_j \cdot \varsigma_j\right) = \sum_{j=1}^m u_j \cdot \mathrm{d}_p f(\varsigma_j) = \sum_{j=1}^m u_j \cdot D_{\varsigma_j} f(p). \tag{420}$$

• *Für die kanonische Basis $e_1, \ldots, e_m$ von $\mathbb{R}^m$ erhalten wir die Identität:*

$$\mathrm{d}_p f(v) = \sum_{j=1}^m v_j \cdot \partial_j f(p) \qquad \forall\, v = (v_1, \ldots, v_m) \in \mathbb{R}^m. \tag{421}$$

• *Schreiben wir $f = (f_1, \ldots, f_n)$ in Komponentenform, und bezeichnet $(\tilde{e}_1, \ldots, \tilde{e}_n)$ die kanonische Basis von $\mathbb{R}^n$, so erhalten wir:*

$$\mathrm{d}_p f(v) = \sum_{i=1}^n \left(\sum_{j=1}^m v_j \cdot \partial_j f_i(p)\right) \cdot \tilde{e}_i \tag{422}$$

$$\overset{\text{Terminologie 34}}{\Longrightarrow} \qquad \Omega[m,n](\mathrm{d}_p f)_{ij} = \tilde{\mathrm{pr}}_i\big(\mathrm{d}_p f(e_j)\big) = \partial_j f_i(p) \quad \textit{für} \quad (i,j) \in \mathcal{I}(n,m).$$

Beweis von (422). Wir erhalten mit (421), sowie (416) – also $\partial_j f(p) = (\partial_j f_1(p), \ldots, \partial_j f_n(p))$:

$$\begin{aligned} \mathrm{d}_p f(v) &\overset{(421)}{=} \textstyle\sum_{j=1}^m v_j \cdot \partial_j f(p) \overset{(416)}{=} \sum_{j=1}^m v_j \cdot \big(\sum_{i=1}^n \partial_j f_i(p) \cdot \tilde{e}_i\big) \\ &= \textstyle\sum_{i=1}^n \big(\sum_{j=1}^m v_j \cdot \partial_j f_i(p)\big) \cdot \tilde{e}_i. \end{aligned}$$ □

Somit ist die Darstellungsmatrix $\widehat{\mathrm{d}_p f} = \Omega[m,n](\mathrm{d}_p f)$ von $\mathrm{d}_p f$ (Terminologie 34) gegeben durch die sogenannte Jacobimatrix von f in p*:*

$$J_p(f) := \big(\partial_j f_i(p)\big)_{(i,j)\,\in\,\mathcal{I}(n,m)} = \begin{pmatrix} \partial_1 f_1(p) & \cdots & \partial_m f_1(p) \\ \partial_1 f_2(p) & \cdots & \partial_m f_2(p) \\ \vdots & & \vdots \\ \partial_1 f_n(p) & \cdots & \partial_m f_n(p). \end{pmatrix} \tag{423}$$

• *Gemäß Notation 31 gilt dann*

$$\Theta[m,n](J_p(f)) \overset{(382)}{=} (\partial_1 f_1(p), \ldots, \partial_m f_1(p), \ldots, \partial_1 f_m(p), \ldots, \partial_m f_n(p)) \in \mathbb{R}^{m\cdot n}. \tag{424}$$

- *Gemäß (416) ist die j-te Spalte von $J_p(f)$ für $1 \leq j \leq n$ gerade gegeben durch die j-t partielle Ableitung $\partial_j f(p)$ von f in p, sofern man diese als Spaltenvektor auffasst:*

$$\big(\underbrace{\partial_j f_1(p), \ldots, \partial_j f_n(p)}_{\partial_j f(p)}\big)^\top = \begin{pmatrix} \partial_j f_1(p) \\ \vdots \\ \partial_j f_n(p) \end{pmatrix}.$$

- *Im Falle $n = 1$ ($f\colon U \to \mathbb{R}$) besteht die Jacobimatrix aus einer Zeile und m Spalten, und wird als der* Gradient $\nabla f(p)$ von f in p *bezeichnet, d.h.*

$$\nabla f(p) := J_p(f) = (\partial_1 f(p), \ldots, \partial_m f(p)) \in \mathbb{R}^m. \tag{425}$$

In Termen des euklidischen Skalarprodunktes schreibt sich dann die Richtungsableitung:

$$D_v f(p) \overset{(419)}{=} \mathrm{d}_p f(v) \overset{(421)}{=} \sum_{j=1}^{m} v_j \cdot \partial_j f(p) = \langle \nabla f(p)\,, v\rangle_{\text{euk}} \qquad \forall\, v \in \mathbb{R}^m. \tag{426}$$

Existiert $\nabla f(p)$ für alle $p \in U$, so wird die Abbildung

$$\nabla f\colon U \to \mathbb{R}^m, \qquad p \mapsto \nabla f(p)$$

als der Gradienten von f *bezeichnet.*

Bemerkung: Man sagt, dass der Gradient in die Richtung der größten Änderung von f in p zeigt:

Wegen der Cauchy-Schwarzschen Ungleichung (Lemma 24) gilt:

$$|D_v f(p)| = |\langle \nabla f(p), v\rangle_{\text{euk}}| \leq \|\nabla f(p)\|_{\text{euk}} \quad \textit{für alle} \quad v \in \mathbb{R}^m \quad \textit{mit} \quad \|v\|_{\text{euk}} = 1.$$

(∗) *Gilt $\nabla f(p) = 0$, so gilt natürlich auch $D_v f(p) = \langle \nabla f(p)\,, v\rangle_{\text{euk}} = 0$ für alle $v \in \mathbb{R}^m$.*

In diesem Fall ist also obige Aussage zu salopp formuliert.

(∗) *Ist $\nabla f(p) \neq 0$, so gilt für die speziellen Einheitsvektoren $v_\pm := \pm\frac{\nabla f(p)}{\|\nabla f(p)\|_{\text{euk}}}$:*

$$D_{v_\pm} f(p) = \langle \nabla f(p), v_\pm\rangle_{\text{euk}} = \pm\tfrac{\|\nabla f(p)\|^2_{\text{euk}}}{\|\nabla f(p)\|_{\text{euk}}} = \pm\|\nabla f(p)\|_{\text{euk}}.$$

Weiterhin gilt für $v \in \mathbb{R}^m$ mit $\|v\|_{\text{euk}} = 1$:

$$\begin{aligned}
\underbrace{D_v f(p)}_{\langle \nabla f(p)\,,\, v\rangle_{\text{euk}}} = \pm\|\nabla f(p)\|_{\text{euk}} \quad &\Rightarrow \quad \langle \nabla f(p), v\rangle_{\text{euk}} = \langle \nabla f(p), v_\pm\rangle_{\text{euk}} \\
&\Rightarrow \quad \langle \nabla f(p), v_\pm - v\rangle_{\text{euk}} = 0 \\
&\Rightarrow \quad \langle v_\pm, v - v_\pm\rangle_{\text{euk}} = 0 \\
&\Rightarrow \quad 0 \leq \underbrace{\langle v - v_\pm, v - v_\pm\rangle_{\text{euk}}}_{\|v - v_\pm\|_{\text{euk}}} = \langle v, v - v_\pm\rangle_{\text{euk}} \\
&\qquad\qquad = \|v\|^2_{\text{euk}} - \langle v, v_\pm\rangle_{\text{euk}} \\
&\qquad\qquad = 1 \mp \|\nabla f(p)\|^{-1}_{\text{euk}} \cdot \langle \nabla f(p), v\rangle_{\text{euk}} \\
&\qquad\qquad = 1 \mp \|\nabla f(p)\|^{-1}_{\text{euk}} \cdot \pm\|\nabla f(p)\|_{\text{euk}} \\
&\qquad\qquad = 0 \\
&\Rightarrow \quad v = v_\pm.
\end{aligned}$$

e) *Wir definieren die Jacobimatrix $J_p(f)$ im Folgenden auch immer dann durch (423), wenn wir erst einmal nur wissen, dass die partiellen Ableitungen $\big(\partial_j f_i(p)\big)_{(i,j)\,\in\,\mathfrak{I}(n,m)}$ existieren.*

Beachte: *Wie wir in Bemerkung 94 und Übung 154 sehen werden, ist die Existenz besagter partieller Ableitungen zwar notwendig für die Existenz des Differentials $\mathrm{d}_p f$ von f in p, aber im Allgemeinen keinesfalls hinreichend; und zwar einfach deswegen, weil die zu $J_p(f)$ gehörige lineare Abbildung $\phi \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ nicht zwangsweise die Bedingung (417) erfüllen muss. Beispielsweise:*

- *Es kann vorkommen (Bemerkung 94), dass zwar alle partiellen Ableitungen von f in p existieren, aber $D_v f(p)$ für ein $0 \neq v \in \mathbb{R}^m$* **nicht**. *Würde nun (417) gelten, so erhielten wir $\phi(v) = D_v f(p)$ wegen (418), im Widerspruch zur Nichtexistenz von $D_v f(p)$.*
- *Es kann vorkommen (Übung 154), dass alle Richtungsableitungen von f in p existieren, trotzdem aber* **nicht** *die Linearitätsbedingung $D_{v+w} f(p) = D_v f(p) + D_w f(p)$ für alle $v, w \in \mathbb{R}^m$ erfüllt ist. Würde aber (417) gelten, so erhielten wir*

$$D_{v+w} f(p) = \phi(v + w) = \phi(v) + \phi(w) = D_v f(p) + D_w f(p)$$

für alle $v, w \in \mathbb{R}^m$ wegen (418).

f) *Sei $m = 1$ und $U \equiv I \subseteq \mathbb{R}$ ein offenes Intervall. Dann ist $\gamma \equiv f$ differenzierbar in $p \in I$ genau dann, wenn γ differenzierbar im Sinne von Definition 36 ist, mit ($1 = e_1$ in $\mathbb{R} = \mathbb{R}^1$)*

$$\mathrm{d}_p\gamma(1) = D_1\gamma(p) = \gamma'(p). \tag{427}$$

Beweis. Zunächst gilt im Falle der Existenz (einer der beiden Seiten):

$$D_1\gamma(p) = \lim_{h\to 0} \frac{\gamma(t + h \cdot 1) - \gamma(t)}{h} \overset{\text{Lemma 76}}{=} \gamma'(t) \qquad \forall\, t \in I.$$

- Existiert $\mathrm{d}_p\gamma$, so gilt $\mathrm{d}_p\gamma(1) = D_1\gamma(p)$ wegen (419), und somit auch (427).
- Existiert $D_1\gamma(p)$, so existiert auch $\mathrm{d}_p\gamma$; also gilt (427) wegen dem vorherigen Punkt.
  In der Tat ist ja $\phi\colon \mathbb{R} \ni h \mapsto h \cdot D_1 f(p) \in \mathbb{R}$ linear; und es gilt

$$\gamma(p + h) = \gamma(p) + \phi(h) + |h| \cdot \varepsilon(h) \quad \text{mit} \quad \varepsilon(0) = 0 \quad \text{und} \quad \lim\nolimits_{h\to 0} \varepsilon(h) = 0,$$

für

$$\varepsilon\colon I - p \to \mathbb{R}, \qquad h \mapsto \begin{cases} \frac{h}{|h|} \cdot \Big(\frac{\gamma(p+h) - \gamma(h)}{h} - \underbrace{\phi(1)}_{D_1 f(p)} \Big) & \text{für} \quad 0 \neq h \in I - p \\ 0 & \text{für} \quad h = 0. \end{cases} \qquad \square$$

(2) • *f heißt* differenzierbar*, wenn f in jedem Punkt $p \in U$ differenzierbar ist. In diesem Fall wird*

$$\mathrm{d}f\colon U \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n), \qquad p \mapsto \mathrm{d}_p f$$

als die Ableitung- *bzw.* das Differential von f *bezeichnet. Wir notieren im Folgenden meistens $\mathrm{d}_p f$ anstatt $\mathrm{d}f(p)$ bzw. $\mathrm{d}_p f(v)$ anstatt $\mathrm{d}f(p)(v)$ für $v \in \mathbb{R}^m$.*

• *f heißt* stetig differenzierbar*, wenn f differenzierbar sowie $\mathrm{d}f$ stetig ist. Die Menge aller stetig differenzierbaren Abbildungen $U \to \mathbb{R}^n$ notieren wir mit $C^1(U, \mathbb{R}^n)$.*

Stetigkeitskriterium: *Wegen Bemerkung 89 ist $\mathrm{d}f$ genau dann stetig, wenn alle Matrixeintragsfunktionen von*

$$\widehat{\mathrm{d}f} = \Omega[m, n] \circ \mathrm{d}f\colon U \to \mathrm{M}_{n,m}(\mathbb{R}), \qquad p \mapsto J_p(f)$$

stetig sind,[102] *bzw. alle Komponentenfunktionen von*

$$\widetilde{\mathrm{d}f} = \Upsilon[m,n] \circ \mathrm{d}f \colon U \to \mathbb{R}^{m\cdot n}$$

$$p \mapsto \Theta[m,n](J_p(f)) \overset{(424)}{=} \big(\partial_1 f_1(p), \dots, \partial_m f_1(p), \dots, \partial_1 f_n(p), \dots, \partial_m f_n(p)\big).$$

Wir können daher $\mathrm{d}f$ *auch als stetige Abbildung* $U \to \mathbb{R}^{m\cdot n}$ *auffassen.*

Bemerkung* 17. *Für* $m = 1 = n$*, wird* (417) *zu der Bedingung* (238) *in Lemma 48.2); und zwar wenn man in* (417) $\phi \colon \mathbb{R} \ni h \mapsto h \cdot b \in \mathbb{R}$ *wählt, sowie in* (238) $D \equiv I$ *als offen voraussetzt und dort* $h \cdot \varepsilon(h)$ *für* $0 \neq h \in I - p$ *in der Form* $|h| \cdot \tilde{\varepsilon}(h)$ *schreibt mit*

$$\tilde{\varepsilon} \colon I - p \to \mathbb{R}^n, \qquad h \mapsto \begin{cases} \frac{h}{|h|} \cdot \varepsilon(h) & \text{für} \quad 0 \neq h \in (I-p) \\ 0 & \text{für} \quad h = 0. \end{cases}$$

Beispiel 102 (Jacobimatrix)**.** *Wir betrachten die Funktion*

$$f \colon \mathbb{R}^3 \to \mathbb{R}^2, \qquad (x,y,z) \mapsto \big(f_1(x,y,z), f_2(x,y,z)\big) = \big(x \cdot y, \sin(x) + \cos(z)\big).$$

Die partiellen Ableitungen der Komponenten von f *im Punkt* $(x,y,z) \in \mathbb{R}^3$ *sind gegeben durch:*

$$\begin{aligned} &\partial_1 f_1(x,y,z) = y, && \partial_2 f_1(x,y,z) = x, && \partial_3 f_1(x,y,z) = 0, \\ &\partial_1 f_2(x,y,z) = \cos(x), && \partial_2 f_2(x,y,z) = 0, && \partial_3 f_2(x,y,z) = -\sin(z). \end{aligned}$$

Beispielsweise ist ja[103]

$$\partial_1 f_1(x,y,z) = \lim_{h\to 0} \frac{f_1(x+h,y,z) - f_1(x,y,z)}{h} = \lim_{h\to 0} \frac{(x+h)\cdot y - x \cdot y}{h} = \lim_{h \to 0} \frac{h \cdot y}{h} = y.$$

Die Jacobimatrix von f *in* (x,y,z) *ist daher gegeben durch*

$$J_{(x,y,z)}(f) = \begin{pmatrix} \partial_1 f_1(x,y,z) & \partial_2 f_1(x,y,z) & \partial_3 f_1(x,y,z) \\ \partial_1 f_2(x,y,z) & \partial_2 f_2(x,y,z) & \partial_3 f_2(x,y,z) \end{pmatrix} = \begin{pmatrix} y & x & 0 \\ \cos(x) & 0 & -\sin(z) \end{pmatrix}.$$

Übung 153. *Sei* $u \in \mathbb{R}^m$*,* $\phi \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$*, sowie*

$$f \colon \mathbb{R}^m \to \mathbb{R}^n, \qquad v \mapsto u + \phi(v)$$

die zugehörige affine Abbildung. Zeigen Sie, dass für jede nichtleere offene Teilmenge $U \subseteq \mathbb{R}^m$*, die Einschränkung* $g := f|_U \colon U \to \mathbb{R}^n$ *differenzierbar ist, mit* $\mathrm{d}_p g = \phi$ *für alle* $p \in U$*. Folgern Sie, dass die Identitätsabbildung* id_U *auf* U *differenzierbar ist, mit* $\mathrm{d}_p \mathrm{id}_U = \mathrm{id}_{\mathbb{R}^m}$ *also* $J_p(\mathrm{id}_U) = \mathbf{1}_m$ *für alle* $p \in U$*.*

Wir haben das folgende Analogon zu Lemma 47:

Lemma 78. *Sei* $U \subseteq \mathbb{R}^m$ *offen und* $f \colon U \to \mathbb{R}^n$ *differenzierbar in* $p \in U$*. Dann ist* f *stetig in* p*.*

Beweis. Wegen (417) gilt ($\|\cdot\|$ eine Norm auf $\mathbb{R}^m$)

$$f(p+h) = f(p) + \mathrm{d}_p f(h) + \|h\| \cdot \varepsilon(h) \qquad \forall\, h \in U - p$$

mit $\lim_{h\to 0} \varepsilon(h) = 0 = \varepsilon(0)$. Zudem gilt $\lim_{h \to 0} \mathrm{d}_p f(h) = \mathrm{d}_p f(0) = 0$, da gemäß Satz 86 jedes Element von $\mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ stetig ist. Somit gilt $\lim_{h\to 0} f(p+h) = f(p)$.

Alternativ: Sei $\|\cdot\|'$ eine Norm auf $\mathbb{R}^n$. Da $\mathrm{d}_p f \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ wegen Satz 86 stetig (in 0) ist mit $\mathrm{d}_p f(0) = 0$, und da ε stetig in 0 ist mit $\varepsilon(0) = 0$, existiert zu jedem $\varepsilon > 0$ ein $0 < \delta \leq 1$ mit $\|\mathrm{d}_p f(h)\|', \|\varepsilon(h)\|' < \frac{\varepsilon}{2}$ für alle $h \in U - p$ mit $\|h\| < \delta$. Die Dreiecksungleichung zusammen mit (417) (erster Schritt) liefert:

$$\|f(p+h) - f(p)\|' \leq \|\mathrm{d}_p f(h)\|' + \|h\|' \cdot \|\varepsilon(h)\|' \leq \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon \quad \text{für alle} \quad h \in U - p \quad \text{mit} \quad \|h\| < \delta. \quad \square$$

[102] Also die Funktionen $U \ni p \mapsto J_p(f)_{ij} = \partial_j f_i(p) \in \mathbb{R}$ für $(i,j) \in \mathcal{I}(n,m)$.

[103] Man erhält also die partielle Ableitung von f_1 an der Stelle (x,y,z), indem man die Funktion $g_{y,z} \colon \mathbb{R} \ni t \mapsto f_1(t,y,z) \in \mathbb{R}$ an der Stelle $t = x$ ableitet.

14.2.2 Charakterisierungen von Differenzierbarkeit

Im Folgenden bezeichne $\|\cdot\|$ immer eine Norm auf dem entsprechenden $\mathbb{R}^\ell$ für $\ell \in \mathbb{N}_{>0}$. Genauso wie bei der Stetigkeit ist auch Differenzierbarkeit eine Eigenschaft, die sich unmittelbar an den Komponentenfunktionen ablesen lässt. Dies hatten wir für Richtungsableitungen auch bereits in Terminologie 40.(1).a) gesehen.

Lemma 79. *Sei $U \subseteq \mathbb{R}^m$ offen und $p \in U$. Eine Abbildung $f\colon U \to \mathbb{R}^n$ ist genau dann differenzierbar in p, wenn alle Komponentenfunktionen $f_1, \ldots, f_n\colon U \to \mathbb{R}$ differenzierbar in p sind, und dann gilt:*

$$\mathrm{d}_p f(v) = (\mathrm{d}_p f_1(v), \ldots, \mathrm{d}_p f_n(v)) \qquad \forall\, v \in \mathbb{R}^m. \tag{428}$$

Beweis. Seien $\phi = (\phi_1, \ldots, \phi_n)\colon \mathbb{R}^m \to \mathbb{R}^n$ und $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)\colon U - p \to \mathbb{R}^n$ Abbildung. Dann gilt:

- $\phi \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \quad \Leftrightarrow \quad \phi_j \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R})$ für alle $1 \leq j \leq n$
- ε stetig in 0 mit $\varepsilon(0) = 0 \quad \Leftrightarrow \quad \varepsilon_j$ stetig in 0 mit $\varepsilon_j(0) = 0$ für alle $1 \leq j \leq n$ (Lemma 74)
- Für jedes $h \in U - p$ gilt:

$$f(p+h) = f(p) + \phi(h) + \|h\| \cdot \varepsilon(h) \quad \Leftrightarrow \quad f_j(p+h) = f_j(p) + \phi_j(h) + \|h\| \cdot \varepsilon_j(h) \quad \forall\, 1 \leq j \leq n.$$

Somit ist f genau dann differenzierbar in p mit $\mathrm{d}_p f = \phi$, wenn f_j für jedes $1 \leq j \leq n$ differenzierbar in p ist mit $\mathrm{d}_p f_j = \phi_j$; und dann gilt $\mathrm{d}_p f = \phi = (\phi_1, \ldots, \phi_n) = (\mathrm{d}_p f_1, \ldots, \mathrm{d}_p f_n)$, also (428). □

Lemma 79 verdeutlicht, dass die höhere Komplexität des Differenzierbarkeitsbegriffes für Funktionen in mehreren Veränderlichen weniger von der Dimension n des Bildbereiches herrührt, sondern vielmehr von Anzahl der Variablen m im Definitionsbereich. Im stetig differenzierbaren Fall gilt dennoch eine zu Lemma 79 komplementäre Aussage betreffend partieller Ableitungen, welche wir gleich noch in Satz 87 bzw. Proposition 21 formulieren und beweisen werden. Wir bemerken zunächst Folgendes:

Bemerkung 93. *Sei $U \subseteq \mathbb{R}^m$ offen, sowie $v \in \mathbb{R}^m$ und $f\colon U \to \mathbb{R}^n$ eine Abbildung, sodass die Richtungsableitung $D_v f\colon U \to \mathbb{R}^n$ existiert.*[104]

a) *Es gilt $D_{\lambda \cdot v} f = \lambda \cdot D_v f$ für alle $\lambda \in \mathbb{R}$ wegen Terminologie 40.(1).*
($D_{\lambda \cdot v} f(p)$ existiert für jedes $p \in U$ und $\lambda \in \mathbb{R}$, mit $D_{\lambda \cdot v} f(p) = \lambda \cdot D_v f(p)$.)

b) *Sei $p \in U$ und $I \subseteq \mathbb{R}$ ein offenes Intervall mit $p + I \cdot v \subseteq U$ (die Existenz garantiert (408) in Bemerkung 92.a)). Die Kurve*

$$\gamma\colon I \to \mathbb{R}^n, \qquad t \mapsto f(p + t \cdot v)$$

ist differenzierbar mit $\gamma'(t) = D_v f(p + t \cdot v)$ für alle $t \in I$.

Beweis der Differenzierbar. Für $t \in I$ sei $\tilde{p} := p + t \cdot v$. Da I offen ist, existiert ein $\delta_v > 0$ mit $t + (-\delta_v, \delta_v) \subseteq I$, also $\tilde{p} + (-\delta_v, \delta_v) \cdot v = p + (t + (-\delta_v, \delta_v)) \cdot v \subseteq p + I \cdot v \subseteq U$. Sei

$$\gamma[v]\colon (-\delta_v, \delta_v) \to \mathbb{R}^n, \qquad t \mapsto f(\tilde{p} + t \cdot v)$$

wie in (411) für $p \equiv \tilde{p}$. Dann gilt $\gamma[v] = \gamma(t + \cdot)$, also (Kettenregel im letzten Schritt)

$$D_v f(p + t \cdot v) = D_v f(\tilde{p}) = \gamma[v]'(0) = \gamma'(t).$$ □

[104] Wir werden die folgenden Aussagen weiter unten für den Spezialfall $n = 1$ anwenden.

Lemma 80. *Sei $U \subseteq \mathbb{R}^m$ offen, $p \in U$, und $f\colon U \to \mathbb{R}^n$ differenzierbar, sodass $\mathrm{d}f$ stetig in p ist. Für jedes $v \in \mathbb{R}^m$, ist dann die zugehörige Richtungsableitung ebenfalls stetig in p:*

$$D_v f\colon U \to \mathbb{R}^n, \qquad q \mapsto D_v f(q) \overset{(419)}{=} \mathrm{d}_q f(v). \tag{429}$$

Insbesondere sind alle Richtungsableitungen (existent und) stetig, wenn f stetig differenzierbar ist.

Beweis. Wir fixieren Normen auf $\mathbb{R}^m$ und $\mathbb{R}^n$, und notieren die zugehörige Operatornorm mit $\|\cdot\|_{\mathrm{op}}$. Für $\varepsilon > 0$, existiert $\delta > 0$ mit $\|\mathrm{d}_{p+h}f - \mathrm{d}_p f\|_{\mathrm{op}} < \frac{\varepsilon}{\max(1,\|v\|)}$ für $h \in \mathsf{B}_\delta(p)$, also

$$\begin{aligned}\|D_v f(p+h) - D_v f(p)\| &\overset{(429)}{=} \|\mathrm{d}_{p+h}f(v) - \mathrm{d}_p f(v)\| = \|(\mathrm{d}_{p+h}f - \mathrm{d}_p f)(v)\| \\ &\overset{(378)}{\leq} \|\mathrm{d}_{p+h}f - \mathrm{d}_p f\|_{\mathrm{op}} \cdot \|v\| < \tfrac{\varepsilon}{\max(1,\|v\|)} \cdot \|v\| \leq \varepsilon.\end{aligned}$$ □

Die Stetigkeitsaussage in Lemma 80 gilt nun natürlich auch für die partiellen Ableitungen $\partial_j f = D_{e_j} f$ mit $1 \leq j \leq n$. Hiermit erhalten wir bereits die eine Implikation in dem folgenden Satz, welcher Lemma 79 im stetig differenzierbaren Fall komplementiert.

Proposition 21. *Sei $U \subseteq \mathbb{R}^m$ offen, $p \in U$, und $f\colon U \to \mathbb{R}^n$ eine Abbildung, sodass die partiellen Ableitungen $\partial_1 f, \dots, \partial_m f$ existieren und stetig in p sind. Dann ist f differenzierbar in p.*

(Insbesondere wird also das Differential von f in p durch die Jacobimatrix dargestellt, die sich in der Form $J_p(f) = \big((\partial_1 f(p))^\top, \dots, (\partial_m f(p))^\top\big)$ schreiben lässt.)

Beweis. Wir definieren (vgl. (421))

$$\mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \ni \Lambda(p)\colon (v_1, \dots, v_m) \mapsto \textstyle\sum_{j=1}^m v_j \cdot \partial_j f(p).$$

Weiterhin definieren wir $\varepsilon\colon U - p \to \mathbb{R}^n$, durch $\varepsilon(0) := 0$ sowie

$$\varepsilon(h) := \frac{f(p+h) - f(p) - \Lambda(p)(h)}{\|h\|_\infty} \qquad \forall\, 0 \neq h \in U - p.$$

Per Konstruktionem gilt dann $f(p+h) = f(p) + \Lambda(p)(h) + \|h\|_\infty \cdot \varepsilon(h)$ für alle $h \in U - p$; und wir müssen nur noch $\lim_{h\to 0} \varepsilon(h) = 0$ nachweisen (ε ist stetig in 0). Wegen Lemma 74 genügt es hierfür, die Stetigkeit in 0 für jede Komponente von ε separate zu zeigen. Wir fixieren $1 \leq \ell \leq n$, und erhalten

$$\varepsilon_\ell(0) = 0 \quad \text{ mit } \quad \varepsilon_\ell(h) = \frac{f_\ell(p+h) - f_\ell(p) - \overbrace{\textstyle\sum_{j=1}^m v_j \cdot \partial_j f_\ell(p)}^{= \Lambda(p)_\ell(h)}}{\|h\|_\infty} \quad \text{ für } \quad 0 \neq h \in U - p.$$

Wir wählen $0 < \tilde\epsilon < 1$ mit $\mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_{2\tilde\epsilon}(p) \subseteq U \subseteq \mathbb{R}^m$, sowie

$$0 \neq h = (h_1, \dots, h_m) = \textstyle\sum_{i=1}^m h_i \cdot e_i \in \mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_{\tilde\epsilon}(0).$$

- Wir definieren $\qquad (h[0] = 0 \quad \text{und} \quad h[m] = h)$

$$h[j] := (h_1, \dots, h_j, 0 \dots, 0) = \textstyle\sum_{i=1}^j h_i \cdot e_i \in \mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_{\tilde\epsilon}(0) \subseteq U - p \qquad \forall\, 0 \leq j \leq m,$$

- Für $1 \leq j \leq m$ definieren wir[105] $\qquad (\alpha_j(0) = p + h[j-1] \quad \text{und} \quad \alpha_j(1) = p + h[j])$

$$\begin{aligned}\alpha_j &\colon (-\tilde\epsilon, 1+\tilde\epsilon) \to \mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_{2\tilde\epsilon}(p), && t \mapsto \quad p + h[j-1] + t \cdot h_j \cdot e_j \\ \gamma_j := f_\ell \circ \alpha_j &\colon (-\tilde\epsilon, 1+\tilde\epsilon) \to \mathbb{R}, && t \mapsto f_\ell\big(p + h[j-1] + t \cdot h_j \cdot e_j\big).\end{aligned}$$

[105] Beachte $|(1+\tilde\epsilon) \cdot h_j| \leq |h_j| + \tilde\epsilon \cdot |h_j| < \tilde\epsilon + \tilde\epsilon^2 \leq 2\tilde\epsilon$ wegen $\tilde\epsilon < 1$.

Wegen Bemerkung 93 ist γ_j differenzierbar mit

$$\begin{aligned}\gamma_j'(t) &\overset{\text{b)}}{=} D_{h_j\cdot e_j} f_\ell(p+h[j-1]+t\cdot h_j\cdot e_j)\\ &\overset{\text{a)}}{=} h_j\cdot D_{e_j} f_\ell(p+h[j-1]+t\cdot h_j\cdot e_j)\\ &= h_j\cdot \partial_j f_\ell(p+h[j-1]+t\cdot h_j\cdot e_j)\\ &= h_j\cdot \partial_j f_\ell(\alpha_j(t))\end{aligned}$$

für alle $t\in(-\tilde\epsilon,1+\tilde\epsilon)$. Der Mittelwertsatz (zweite Identität) liefert ein $\xi_{h,j}\in(0,1)$ mit

$$f_\ell(\alpha_j(1))-f_\ell(\alpha_j(0)) = \frac{\gamma_j(1)-\gamma_j(0)}{1-0} \overset{\text{Satz 44}}{=} \gamma_j'(\xi_{h,j}) = h_j\cdot\partial_j f_\ell(\alpha_j(\xi_{h,j})).$$

Durch Bildung einer Teleskopsumme erhalten wir $\qquad (\alpha_m(1)=p+h \quad \text{und} \quad \alpha_1(0)=p)$

$$f_\ell(p+h)-f_\ell(p)-\Lambda(p)_\ell(h) = \sum_{j=1}^m \underbrace{f_\ell(\overbrace{p+h[j]}^{\alpha_j(1)}) - f_\ell(\overbrace{p+h[j-1]}^{\alpha_j(0)})}_{h_j\ \cdot\ \partial_j f_\ell(\alpha_j(\xi_{h,j}))} - h_j\cdot\partial_j f_\ell(p).$$

Hiermit folgt aus der Dreiecksungleichung

$$|\varepsilon_\ell(h)| \le \sum_{j=1}^m \overbrace{\frac{|h_j|}{\|h\|_\infty}}^{\le 1}\cdot\big|\partial_j f_\ell(\alpha_j(\xi_{h,j}))-\partial_j f_\ell(p)\big|.$$

Wegen $\alpha_j(\xi_{h,j})\in \mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_{2\tilde\epsilon}(p)$ für $1\le j\le m$, folgt nun $\lim_{h\to 0}\varepsilon(h)=0$ aus der Stetigkeit der partiellen Ableitungen von f (bzw. deren Komponenten). ☐

Satz 87. *Sei $U\subseteq\mathbb{R}^m$ offen. Eine Abbildung $f\colon U\to\mathbb{R}^n$ ist genau dann stetig differenzierbar, wenn die partiellen Ableitungen $\partial_1 f,\dots,\partial_n f$ existieren und stetig sind.*

(Insbesondere wird also das Differential von f in p durch die Jacobimatrix dargestellt, die sich in der Form $J_\bullet(f)=\big((\partial_1 f)^\top,\dots,(\partial_m f)^\top\big)$ schreiben lässt.)

Beweis. • Ist f stetig differenzierbar, so existieren wegen Lemma 80 alle partiellen Ableitungen von f und sind stetig.

• Existieren alle partiellen Ableitungen von f und sind stetig, so ist f gemäß Proposition 21 differenzierbar mit

$$\mathrm{d}f\colon U\to \operatorname{Hom}(\mathbb{R}^m,\mathbb{R}^n),\qquad p\mapsto \big[\,\Lambda(p)\colon (v_1,\dots,v_m)\mapsto \textstyle\sum_{j=1}^m v_j\cdot\partial_j f(p)\,\big].$$

Es bleibt die Stetigkeit von $\mathrm{d}f$ nachzuweisen. Es bezeichne hierfür $\|\cdot\|_{\infty,\mathrm{op}}$ die, zu den Maximumsnormen auf $\mathbb{R}^m$ und $\mathbb{R}^n$ gehörige Operatornorm. Sei nun $p\in U$ und $\varepsilon>0$. Per Voraussetzung existiert ein $\delta>0$ mit

$$\|\partial_j f(p+h)-f(p)\|_\infty < \varepsilon/2m \qquad \forall\, h\in \mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_\delta(0),\ 1\le j\le n.$$

Für $h\in\mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_\delta(0)$ und $v=(v_1,\dots,v_m)\in\mathbb{R}^m$ mit $\|v\|_\infty=1$, erhalten wir

$$\begin{aligned}\|(\Lambda(p+h)-\Lambda(p))(v)\|_\infty &= \|\Lambda(p+h)(v)-\Lambda(p)(v)\|_\infty = \|\textstyle\sum_{j=1}^m v_j\cdot(\partial_j f(p+h)-\partial_j f(p))\|_\infty\\ &\le \textstyle\sum_{j=1}^m |v_j|\cdot\|\partial_j f(p+h)-\partial_j f(p)\|_\infty\\ &\le m\cdot\|v\|_\infty\cdot\frac{\varepsilon}{2m} \le \|v\|_\infty\cdot\varepsilon/2 = \varepsilon/2.\end{aligned}$$

Dies zeigt $\|\Lambda(p+h)-\Lambda(p)\|_{\infty,\mathrm{op}}\le\varepsilon/2<\varepsilon$ für alle $h\in\mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_\delta(0)$. ☐

Beispiel 103. *Wir betrachten die Abbildung*

$$f\colon \mathbb{R}^3 \to \mathbb{R}^2, \qquad (x,y,z) \mapsto \big(xy, \sin(x) + \cos(z)\big)$$

aus Beispiel 102. Die Richtungsableitungen in Spaltenform, sind dann gegeben durch die entsprechenden Spalten der Jacobimatrix

$$J_{(x,y,z)}(f) = \begin{pmatrix} y & x & 0 \\ \cos(x) & 0 & -\sin(z) \end{pmatrix} \qquad \forall\, (x,y,z) \in \mathbb{R}^3,$$

also offensichtlich stetig. Somit ist $J_p(f)$ auch die Matrixdarstellung des Differentials $\mathrm{d}_p f$ von f in p, welches ja wegen Satz 87 existiert.

Bemerkung 94. *In Satz 87 (Proposition 21) ist die Forderung nach der Stetigkeit (in p) der partiellen Ableitungen zwingend erforderlich, d.h., es ist im Allgemeinen nicht korrekt, dass f differenzierbar (in p) ist wenn alle partiellen Ableitungen existieren. Die Existenz der partiellen Ableitungen ist somit zwar notwendig für die Existenz des Differentials aber nicht hinreichend:*

Wir betrachten die Funktion

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \qquad (x,y) \mapsto \begin{cases} 0 & \textit{falls} \quad (x,y) = (0,0) \\ \dfrac{xy}{x^2+y^2} & \textit{sonst.} \end{cases}$$

- *Lemma 78 zeigt, dass f in $(0,0)$ nicht differenzierbar ist, denn f ist unstetig in $(0,0)$:*[106]

$$f(t,t) = \tfrac{t^2}{t^2+t^2} = \tfrac{1}{2} \qquad \forall\, t \in \mathbb{R}_\times \qquad \Longrightarrow \qquad \lim_{0 \neq t \to 0} f(t,t) = \tfrac{1}{2} \neq 0 = f(0,0).$$

- *Dennoch existieren alle partiellen Ableitungen von f:*
  - *Für $(0,0) \neq (x,y) \in \mathbb{R}^2$ erhalten wir aus den üblichen Ableitungsregeln:*

$$\partial_1 f(x,y) = \frac{y}{x^2+y^2} - \frac{2x^2 y}{(x^2+y^2)^2}, \qquad \partial_2 f(x,y) = \frac{x}{x^2+y^2} - \frac{2xy^2}{(x^2+y^2)^2}.$$

  - *Für $(x,y) = (0,0)$ erhalten wir:*

$$\partial_1 f(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} 0 = 0 = \lim_{h\to 0} 0 = \frac{f(0,h) - f(0,0)}{h} = \partial_2 f(0,0).$$

Im Einklang mit Proposition 21 sind die partiellen Ableitungen in $(0,0)$ unstetig;[107] *denn beispielsweise existieren die folgenden beiden Grenzwerte nicht:*

$$\lim_{0 \neq t \to 0} \partial_1 f(0,t) = \lim_{0 \neq t \to 0} \frac{t}{t^2} = \underbrace{\lim_{0 \neq t \to 0} \frac{1}{t}}_{\text{existiert nicht}} = \lim_{0 \neq t \to 0} \frac{t}{t^2} = \lim_{0 \neq t \to 0} \partial_2 f(t,0).$$

Zudem kann f auch deswegen nicht in $(0,0)$ differenzierbar sein, weil die Richtungsableitung

$$D_{e_1+e_2} f(0,0) = \lim_{h\to 0} \frac{f(h,h) - f(0,0)}{h} = \lim_{h\to 0} \frac{1}{2h}$$

schlichtweg nicht existiert.

[106] Die Schreibweise $\lim_{0 \neq t \to 0} f(t,t)$ steht natürlich abkürzend für $\lim_{t\to 0} g(t)$ für $g\colon \mathbb{R}_\times \ni t \mapsto f(t,t)$; analog auch weiter unten.

[107] Genauer erzwingt Proposition 21 die Unstetigkeit mindestens einer partiellen Ableitung.

Übung 154. *Wir betrachten die Funktion*

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \qquad (x,y) \mapsto \begin{cases} 0 & \textit{falls} \quad x = 0 \\ \dfrac{xy^2}{x^2+y^4} & \textit{falls} \quad x \neq 0. \end{cases}$$

Zeigen Sie:

- $D_v f(0,0)$ *existiert für alle* $v \in \mathbb{R}^2$,
- f *ist in* $(0,0)$ *nicht differenzierbar.*

14.2.3 Differentiationsregeln

In diesem Abschnitt werden wir einige grundlegende Differenzierbarkeitsregeln wie die Produkt- und die Kettenregel beweisen. Hierfür ist das folgende Analogon zu Lemma 48.3) für den eindimensioanlen Fall sehr hilfreich.

Proposition 22 (Differenzierbarkeitskriterium)**.** *Sei* $U \subseteq \mathbb{R}^m$ *offen,* $p \in U$*,* $\phi \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$*, und* $f\colon U \to \mathbb{R}^n$ *eine Abbildung. Die folgenden Aussagen sind zueinander äquivalent:*

1) f *ist differenzierbar in* p *mit* $\mathrm{d}_p f = \phi$.

2) Es existiert eine in 0 *stetige Abbildung* $\Phi\colon U - p \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ *mit* $\Phi(0) = \phi$*, sodass gilt:*

$$f(p+h) - f(p) = \Phi(h)(h) \qquad \forall\, h \in U - p. \tag{430}$$

Beweis. Wir betrachten auf $\mathbb{R}^m$ und $\mathbb{R}^n$ jeweils die euklidische Norm, und bezeichnen mit $\|\cdot\|_{\text{euk,op}}$ die zugehörige Operatornorm auf $\mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$.

- 1) $\Rightarrow$ 2): Sei ε wie in (417) für $\|\cdot\| = \|\cdot\|_{\text{euk}} = \langle \cdot, \cdot \rangle_{\text{euk}}$ die euklidischen Norm auf $\mathbb{R}^m$, d.h.,

  $$f(p+h) = f(p) + \phi(h) + \|h\|_{\text{euk}} \cdot \varepsilon(h) \qquad \forall\, h \in U - p. \tag{431}$$

  Wir definieren $\Phi\colon U - p \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ durch

  $$\Phi(h)\colon \mathbb{R}^m \to \mathbb{R}^n, \qquad v \mapsto \begin{cases} \phi(v) & \text{für} \quad h = 0 \\ \phi(v) + \dfrac{\langle h, v \rangle_{\text{euk}}}{\|h\|_{\text{euk}}} \cdot \varepsilon(h) & \text{für} \quad 0 \neq h \in U - p. \end{cases}$$

  Dann gilt $\Phi(0) = \phi$ also $\Phi(0)(0) = 0 = f(p+0) - f(p)$, und weiterhin

  $$\Phi(h)(h) = \phi(h) + \|h\|_{\text{euk}} \cdot \varepsilon(h) \overset{(431)}{=} f(p+h) - f(p) \qquad \forall\, 0 \neq h \in U - p.$$

  Zudem ist Φ stetig in 0; denn für $v \in \mathbb{R}^m$ mit $\|v\|_{\text{euk}} = 1$ und $0 \neq h \in U - p$, liefert die Cauchy-Schwarzsche Ungleichung (letzter Schritt)

  $$\begin{aligned} \|(\Phi(h) - \Phi(0))(v)\|_{\text{euk}} &= \|\Phi(h)(v) - \Phi(0)(v)\|_{\text{euk}} \\ &= \left\| \frac{\langle h, v \rangle_{\text{euk}}}{\|h\|_{\text{euk}}} \cdot \varepsilon(h) \right\|_{\text{euk}} = \frac{|\langle h, v \rangle_{\text{euk}}|}{\|h\|_{\text{euk}}} \cdot \|\varepsilon(h)\|_{\text{euk}} \\ &\overset{(111)}{\leq} \frac{\|h\|_{\text{euk}} \cdot \|v\|_{\text{euk}}}{\|h\|_{\text{euk}}} \cdot \|\varepsilon(h)\|_{\text{euk}} = \|\varepsilon(h)\|_{\text{euk}}. \end{aligned}$$

  Somit gilt für die zugehörige Operatornorm

  $$\|\Phi(h) - \Phi(0)\|_{\text{euk,op}} \leq \|\varepsilon(h)\|_{\text{euk}} \qquad \forall\, 0 \neq h \in U - p,$$

  womit die Stetigkeit von Φ in 0 nun aus der Stetigkeit von ε in 0 folgt.

- 2) $\Rightarrow$ 1): Wir definieren $\varepsilon\colon U-p \to \mathbb{R}^n$ durch

$$\varepsilon(h) := \begin{cases} 0 & \text{für} \quad h = 0 \\ \dfrac{\Phi(h)(h) - \phi(h)}{\|h\|_{\text{euk}}} & \text{für} \quad 0 \neq h \in U - p. \end{cases}$$

Dann gilt (417) (also (431)) per Konstruktion. Nun ist $\Phi\colon U-p \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ stetig in 0 mit $\Phi(0) = \phi$. Zu $\tilde{\epsilon} > 0$ vorgegeben, existiert somit ein $\tilde{\delta} > 0$ mit $\|\Phi(h) - \phi\|_{\text{euk,op}} < \tilde{\epsilon}$ für alle $h \in U-p$ mit $\|h\|_{\text{euk}} < \tilde{\delta}$. Mit (378) erhalten wir für $0 \neq h \in U-p$ mit $\|h\|_{\text{euk}} < \tilde{\delta}$:

$$\|\varepsilon(h)\|_{\text{euk}} = \frac{1}{\|h\|_{\text{euk}}} \cdot \|(\Phi(h) - \phi)(h)\|_{\text{euk}} \overset{(378)}{\leq} \|\Phi(h) - \phi\|_{\text{euk,op}} < \tilde{\epsilon},$$

was die Stetigkeit von ε in 0 zeigt (da $\varepsilon(0) = 0$ per definitionem gilt). □

Mit Hilfe von Proposition 22 erhalten wir nun die folgenden Rechenregeln für Ableitungen.

Satz 88 (Summenregel und Kettenregel)**.**

1) Sei $U \subseteq \mathbb{R}^m$ offen, $f, g\colon U \to \mathbb{R}^n$ differenzierbar in $p \in U$, und $\lambda, \mu \in \mathbb{R}$. Dann ist auch die Funktion $(\lambda \cdot f + \mu \cdot g)$ differenzierbar in p, mit

$$\mathrm{d}_p(\lambda \cdot f + \mu \cdot g) = \lambda \cdot \mathrm{d}_p f + \mu \cdot \mathrm{d}_p g.$$

2) Seien $U \subseteq \mathbb{R}^m$ und $V \subseteq \mathbb{R}^n$ offen, sowie $g\colon U \to \mathbb{R}^n$ und $f\colon V \to \mathbb{R}^k$ Abbildungen mit $\mathrm{im}(g) \subseteq V$, für $m, n, k \in \mathbb{N}_{>0}$. Ist g differenzierbar in $p \in U$ und f differenzierbar in $g(p) \in V$, so ist $(f \circ g)\colon U \to \mathbb{R}^k$ differenzierbar in p mit *(Kettenregel)*

$$\mathrm{d}_p(f \circ g) = \mathrm{d}_{g(p)} f \circ \mathrm{d}_p g.$$

(Die Beweise verlaufen komplett analog zu den Beweisen von Satz 40.1) und Satz 41.)

Beweis. 1) Wegen Proposition 22 existieren Abbildungen $\Phi, \Psi\colon U - p \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$, mit

$$\begin{aligned} f(p+h) - f(p) &= \Phi(h)(h), \qquad &\Phi \text{ stetig in } 0, \qquad &\Phi(0) = \mathrm{d}_p f \\ g(p+h) - g(p) &= \Psi(h)(h), \qquad &\Psi \text{ stetig in } 0, \qquad &\Psi(0) = \mathrm{d}_p g \end{aligned}$$

für alle $h \in U - p$. Für $h \in U - p$ erhalten wir

$$(\lambda \cdot f + \mu \cdot g)(p+h) - (\lambda \cdot f + \mu \cdot g)(p) = \overbrace{(\lambda \cdot \Phi + \mu \cdot \Psi)}^{=:\,\Theta}(h)(h),$$

wobei die Abbildung $\Theta\colon U - p \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ als Linearkombination von in 0 stetigen Abbildungen gemäß Satz 32 (mit $(V, \|\cdot\|) \equiv (\mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n), \|\cdot\|_{\text{op}})$) ebenfalls in 0 stetig ist. Proposition 22 zeigt somit die Differenzierbarkeit von $(\lambda \cdot f + \mu \cdot g)$ in p, mit

$$\mathrm{d}_p(\lambda \cdot f + \mu \cdot g) = \Theta(0) = \lambda \cdot \Phi(0) + \mu \cdot \Psi(0) = \lambda \cdot \mathrm{d}_p f + \mu \cdot \mathrm{d}_p g.$$

2) Wegen Proposition 22 existieren Abbildungen $\Phi\colon U - p \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ und $\Psi\colon V - g(p) \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^k)$ mit

$$\begin{aligned} g(p+h) - g(p) \quad &= \Phi(h)(h), \qquad &\Phi \text{ stetig in } 0, \qquad &\Phi(0) = \mathrm{d}_p g \\ f(g(p) + \tilde{h}) - f(g(p)) &= \Psi(\tilde{h})(\tilde{h}), \qquad &\Psi \text{ stetig in } 0, \qquad &\Psi(0) = \mathrm{d}_{g(p)} f \end{aligned}$$

für alle $h \in U - p$ und $\tilde{h} \in V - g(p)$. Dann ist

$$\alpha\colon U - p \to \mathbb{R}^n, \qquad h \mapsto \Phi(h)(h) = f(p+h) - f(p)$$

stetig in 0 wegen Lemma 78 (g ist stetig in p); und wir definieren

$$\Theta\colon U - p \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^k), \qquad h \mapsto \big[\, \Psi(\underbrace{\Phi(h)(h)}_{=\alpha(h)}) \circ \Phi(h)\colon v \mapsto \Psi\big(\Phi(h)(h)\big)\big(\Phi(h)(v)\big)\,\big].$$

Für $h \in U - p$ erhalten wir

$$(f\circ g)(p+h) = f(g(p+h)) = f\big(g(p) + \Phi(h)(h)\big) = \overbrace{f(g(p))}^{(f\circ g)(p)} + \overbrace{\Psi\big(\Phi(h)(h)\big)\big(\Phi(h)(h)\big)}^{\Theta(h)(h)}.$$

Zudem gilt $\Theta(0) = (\mathrm{d}_{g(p)}f \circ \mathrm{d}_p g)$, denn für $v \in \mathbb{R}^m$ ist

$$\Theta(0)(v) = \Psi(\overbrace{\Phi(0)(0)}^{0})(\overbrace{\Phi(0)}^{\mathrm{d}_p g}(v)) = \overbrace{\Psi(0)}^{\mathrm{d}_{g(p)}f}(\mathrm{d}_p g(v)) = \mathrm{d}_{g(p)}f(\mathrm{d}_p g(v)) = \big(\mathrm{d}_{g(p)}f \circ \mathrm{d}_p g\big)(v).$$

Wegen Proposition 22 folgt hiermit die Behauptung sobald wir die Stetigkeit von Θ in 0 nachgewiesen haben. Zunächst ist $\Psi \circ \alpha\colon U - p \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^k)$ stetig in 0 wegen Lemma 39.1) (α stetig in 0 mit $\alpha(0) = 0$); also ist[108]

$$\Lambda := (\Psi \circ \alpha, \Phi)\colon U - p \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^k) \times \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$$

stetig in 0 wegen Lemma 73 (Φ ist ja ebenfalls stetig in 0). Wegen der Stetigkeit der Kompositionsabbildung (siehe (380) in Beispiel 98)

$$\circ\colon \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^k) \times \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^k), \qquad (\phi, \psi) \mapsto \phi \circ \psi,$$

ist dann $\Theta = \circ \circ \Lambda\colon U - p \ni h \mapsto \circ(\Psi(\alpha(h)), \Phi(h)) = \Psi(\alpha(h)) \circ \Phi(h) \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^k)$ als Verkettung von $\circ$ mit Λ ebenfalls stetig in 0. $\square$

Bemerkung 95. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Aus der Linearen Algebra kennen Sie die folgenden Sachverhalte:*

a) *Seien $\phi\colon \mathbb{K}^n \to \mathbb{K}^k$ und $\psi\colon \mathbb{K}^m \to \mathbb{K}^n$ lineare Abbildungen. Dann gilt*

$$\underbrace{\Omega[m,k](\phi \circ \psi)}_{C} = \underbrace{\Omega[n,k](\phi)}_{A} \cdot \underbrace{\Omega[m,n](\psi)}_{B}.$$

D.h., die Darstellungmatrix (bzgl. der kanonischen Basen) der Verkettung $\phi \circ \psi$ ist gegeben durch das Matrixprodukt der Darstellungsmatrizen (bzgl. der entsprechenden Standardbasen) von ϕ und ψ. Ist nämlich $(e_1, \dots, e_m)$ die kanonische Basis von $\mathbb{K}^m$, $(\tilde{e}_1, \dots, \tilde{e}_n)$ die kanonische Basis von $\mathbb{K}^n$, und $\hat{\mathrm{pr}}_i\colon \mathbb{K}^k \to \mathbb{K}$ für $1 \leq i \leq k$ die Projektion auf die i-te Koordinate in $\mathbb{K}^k$, so gilt für $(i,j) \in \mathfrak{I}(k,m)$:

$$C_{ij} = \hat{\mathrm{pr}}_i\big((\phi\circ\psi)(e_j)\big) \overset{(385)}{=} \hat{\mathrm{pr}}_i\big(\phi(\overbrace{\textstyle\sum_{\ell=1}^n B_{\ell j} \cdot \tilde{e}_\ell}^{\psi(e_j)})\big) = \textstyle\sum_{\ell=1}^n \hat{\mathrm{pr}}_i\big(\phi(\tilde{e}_\ell)\big) \cdot B_{\ell j} = \sum_{\ell=1}^n A_{i\ell} \cdot B_{\ell j}.$$

b) *Ist $\phi\colon V \to W$ eine injektive lineare Abbildung zischen den endlichdimensionalen $\mathbb{K}$-Vektorräumen V und W, so gilt notwendig $\dim(V) \leq \dim(W)$; denn ϕ bildet jede Basis von V auf eine Familie linear unabhängiger Vektoren ab.*

Bemerkung 96 (Jacobimatrix und Verkettungen)**.** *Seien $U \subseteq \mathbb{R}^m$ und $V \subseteq \mathbb{R}^n$ offene Mengen, sowie $g\colon U \to \mathbb{R}^n$ und $f\colon V \to \mathbb{R}^k$ Abbildungen mit $\mathrm{im}(g) \subseteq V$, für $m,n,k \in \mathbb{N}_{>0}$. Sei g differenzierbar in $p \in V$ und f differenzierbar in $g(p) \in V$:*

[108] Unseren Konventionen aus Kapitel 13.1 folgend, versehen wir den Produktraum auf der rechten Seite mit der Maximumsmetrik welche zu den durch die entsprechenden Operatornormen gehörigen Normmetriken gehört.

1) *Wegen Terminologie 41.(1).d) ist die Darstelungsmatrix bezüglich der kanonischen Basen, des Differentials einer Abbildung in einem Punkt gegeben durch die Jacobimatrix in diesem Punkt.*

Wegen Satz 88.2) und Bemerkung 95.a) gilt somit der Zusammenhang:

$$J_p(f \circ g) = J_{g(p)}(f) \cdot J_p(g). \tag{432}$$

Für die partiellen Ableitungen (Einträge der Jacobimatrix) erhalten wir

$$\partial_j(f \circ g)_i(p) \overset{(423)}{=} J_p(f \circ g)_{ij} \overset{(432)}{=} \sum_{\ell=1}^{n} J_{g(p)}(f)_{i\ell} \cdot J_p(g)_{\ell j} \overset{(423)}{=} \sum_{\ell=1}^{n} \partial_\ell f_i(g(p)) \cdot \partial_j g_\ell(p) \tag{433}$$

für alle $(i,j) \in \mathfrak{I}(k,m)$, *also auch*

$$\partial_j(f \circ g)(p) = \sum_{\ell=1}^{n} \partial_\ell f(g(p)) \cdot \partial_j g_\ell(p) \qquad \forall\ 1 \leq j \leq m. \tag{434}$$

- *Sei* $m = 1$ *und* $U \equiv I$ *ein offenes Intervall. Dann können wir* g *als eine in* p *differenzierbare Kurve* $\gamma := g \colon I \to \mathbb{R}^n$ *auffassen (Terminologie 41.(1).f)); und selbiges trifft dann auch auf* $f \circ \gamma \colon I \to \mathbb{R}^k$ *zu. Wir erhalten den Zusammenhang*

$$\begin{aligned}(f \circ \gamma)'(p) \overset{(427)}{=} \partial_1(f \circ g)(p) &\overset{(434)}{=} \textstyle\sum_{\ell=1}^{n} \partial_\ell f(g(p)) \cdot \partial_1 g_\ell(p) \\ &\overset{(427)}{=} \textstyle\sum_{\ell=1}^{n} \partial_\ell f(\gamma(p)) \cdot \gamma_\ell'(p) \overset{(426)}{=} D_{\gamma'(p)} f(\gamma(p)),\end{aligned} \tag{435}$$

wobei auf der rechten Seite die Richtungsableitung von f *in* p *in Richtung* $\gamma'(p) \in \mathbb{R}^m$ *steht.*

- *In der Situation des vorherigen Punktes sei* $k = 1$. *Man bezeichnet die Teilmengen*

$$H_c := \{x \in V \mid f(x) = c\} \quad \textit{mit} \quad c \in \mathbb{R}$$

auch als „Höhenlinien der Funktion f*“;*[109] *und zwar in Anlehnung an den Fall* $n = 2$, *in welchem man sich* f *als Höhenfunktion über dem „Gelände“* V *vorstellen kann, d.h.* $f(x)$ *ist die Höhe des Punktes* $x \in V$ *bspw. über dem Meeresspiegel.*

Behauptung: *Sei* $I \subseteq \mathbb{R}$ *ein Intervall. Eine differenzierbare Kurve* $\gamma \colon I \to V$ *verläuft genau dann in einer „Höhenlinie“* H_c *mit* $c \in \mathbb{R}$ $(\gamma(I) \subseteq H_c)$, *wenn* $\langle \nabla f \circ \gamma, \gamma' \rangle_{\text{euk}} = 0$ *gilt.*

(Geometrisch interpretiert man dies so, dass der Geschwindigkeitsvektor $\gamma'(t)$ *der Kurve* γ *in* t *senkrecht zum Gradienten* $\nabla f(\gamma(t))$ *im Punkt* $\gamma(t)$ *steht. „Die Höhenlinien verlaufen also in jedem Punkt senkrecht zum Gradienten.“)*

Beweis der Behauptung. Wir erhalten mit Korollar 32.2) und (435):

$$\exists\, c \in \mathbb{R} \colon\ \gamma(I) \subseteq H_c \qquad \overset{\text{Korollar 32}}{\Longleftrightarrow} \qquad 0 = (f \circ \gamma)' \overset{(435)}{=} D_{\gamma'} f(\gamma) \overset{(426)}{=} \langle \nabla f \circ \gamma, \gamma' \rangle_{\text{euk}}. \quad \square$$

2) *Es gelte* $f \circ g = \mathrm{id}_U$, *d.h.* g *ist rechtsinvers zu* f *(bzw.* f *ist linksinvers zu* g*) mit* $k = m$. *Aus Satz 88.2) und Übung 153 erhalten wir*

$$\mathrm{id}_{\mathbb{R}^m} = \mathrm{d}_{g(p)} f \circ \mathrm{d}_p g \qquad \textit{und} \qquad \mathbf{1}_m = J_{g(p)}(f) \cdot J_p(g).$$

Daher ist $\mathrm{d}_p g \colon \mathbb{R}^m \to \mathbb{R}^n$ *rechtsinvers zu* $\mathrm{d}_{g(p)} f$, *also injektiv. Bemerkung 95.b) zeigt* $m \leq n$.[110]

[109] Die Menge H_c kann natürlich auch einelementig sein, bzw. ganz V sofern f konstant ist.

[110] Sind insbesondere $f \colon \mathbb{R}^m \supseteq U \to V \subseteq \mathbb{R}^n$ und $g \colon \mathbb{R}^n \supseteq V \to U \subseteq \mathbb{R}^m$ zueinander inverse differenzierbare Abbildung, so muss bereits $m = n$ gelten. Wir kommen hierauf in Abschnitt 16.3 (Korollar 64) zurück.

Proposition 23. *Sei $\Psi\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^k$ bilinear mit $m, n, k \in \mathbb{N}_{>0}$. Dann ist Ψ differenzierbar mit*

$$\mathrm{d}_{(x,y)}\Psi\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^k, \qquad (v,w) \mapsto \Psi(v,y) + \Psi(x,w) \tag{436}$$

für alle $(x,y) \in \mathbb{R}^m \times \mathbb{R}^n$.

Wir statten $\mathbb{R}^m$, $\mathbb{R}^n$ und $\mathbb{R}^m \times \mathbb{R}^n$ mit der jeweiligen Maximumsnormen aus.[111]

Beweis. Sei $p = (x,y) \in \mathbb{R}^m \times \mathbb{R}^n$ fixiert, und sei $\phi\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^k$ durch die rechte Seite von (436) definiert. Offensichtlich ist ϕ linear;[112] und es gilt für $(0,0) \neq h = (v,w) \in \mathbb{R}^m \times \mathbb{R}^n$:

$$\begin{aligned}
\Psi(p+h) = \Psi((x,y)+(v,w)) &= \Psi(x+v, y+w) \\
&= \underbrace{\Psi(x,y)}_{\Psi(p)} + \underbrace{\Psi(v,y) + \Psi(x,w)}_{\phi(v,w) = \phi(h)} + \underbrace{\Psi(v,w)}_{\Psi(h)} \\
&= \Psi(p) + \phi(h) + \|h\|_\infty \cdot \underbrace{\frac{\Psi(h)}{\|h\|_\infty}}_{=:\,\varepsilon(h)}.
\end{aligned}$$

Wir setzen $\varepsilon(0,0) := 0$ und müssen noch $\lim_{h\to 0} \varepsilon(h) = 0$ zeigen. Wegen Satz 86 ist Ψ stetig; also existiert gemäß Lemma 75 ein $C \geq 0$ mit $\|\Psi(v,w)\|_\infty \leq C \cdot \|v\|_\infty \cdot \|w\|_\infty$ für alle $(v,w) \in \mathbb{R}^m \times \mathbb{R}^n$. Somit gilt

$$\|\varepsilon(v,w)\|_\infty \leq C \cdot \frac{\|v\|_\infty \cdot \|w\|_\infty}{\|(v,w)\|_\infty} \leq C \cdot \frac{\|(v,w)\|_\infty^2}{\|(v,w)\|_\infty} = C \cdot \|(v,w)\|_\infty$$

für alle $(v,w) \in \mathbb{R}^m \times \mathbb{R}^n$. Also ist ε stetig in 0 mit $\varepsilon(0,0) = 0$. $\square$

Übung 155 (Allgemeine Produktregel)**.** *Sei $U \subseteq \mathbb{R}^\ell$ offen, sowie $f\colon U \to \mathbb{R}^m$ und $g\colon U \to \mathbb{R}^n$ differenzierbar in $p \in U$. Sei zudem $\Psi\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^k$ bilinear. Zeigen Sie, dass die Abbildung $\Lambda\colon \mathbb{R}^\ell \ni x \mapsto \Psi(f(x), g(x)) \in \mathbb{R}^k$ in p differenzierbar ist, und berechnen Sie $\mathrm{d}_p\Lambda$. Vergleichen Sie das Ergebnis mit der Produktregel für Funktionen in einer veränderlichen (Satz 40), bzw. b) und d) in Bemerkung 90.*

Übung 156 (Version der Kettenregel)**.** *Seien $U \subseteq \mathbb{R}^m$ und $V \subseteq \mathbb{R}^n$ offen, sowie $g\colon U \to \mathbb{R}^n$ und $f\colon V \to \mathbb{R}^k$ Abbildungen mit $\mathrm{im}(g) \subseteq V$, für $m, n, k \in \mathbb{N}_{>0}$. Sie weiterhin f differenzierbar in $g(p)$ für $p \in U$. Zeigen Sie, dass*

$$D_v(f \circ g)(p) = \mathrm{d}_{g(p)} f(D_v g(p)) = \textstyle\sum_{\ell=1}^n \partial_\ell f(g(p)) \cdot D_v g_\ell(p)$$

für $v \in \mathbb{R}^m$ gilt, sofern $D_v g(p)$ existiert.

Hinweis: Die erste Identität folgt problemlos aus (417) *sowie der Stetigkeit der linearen Abbildung $\mathrm{d}_{g(p)} f$. Für die zweite Identität benutzen Sie die Linearität von $\mathrm{d}_{g(p)} f$ sowie* (413) *für g.*

[111] Erinnerung an Bemerkung 87.b): Es macht keinen Unterschied, ob wir $\mathbb{R}^m \times \mathbb{R}^n$ mit $\mathbb{R}^{m+n}$ identifizieren und hierauf die Maximumsnorm wählen, oder ob wir auf $\mathbb{R}^m \times \mathbb{R}^n$ die zu den Maximumsnormen auf $\mathbb{R}^m$ und $\mathbb{R}^n$ zugehörige Maximumsnorm wählen.

[112] Wegen der Bilinearität von Ψ gilt: $\phi(\lambda \cdot (v,w)) = \phi(\lambda \cdot v, \lambda \cdot w) = \Psi(\lambda \cdot v, y) + \Psi(x, \lambda \cdot w) = \lambda \cdot \Psi(v,y) + \lambda \cdot \Psi(x,w) = \lambda \cdot \phi(v,w)$ für $\lambda \in \mathbb{R}$ und $(v,w) \in \mathbb{R}^m \times \mathbb{R}^n$; sowie $\phi((v,w)+(v',w')) = \phi(v+v', w+w') = \Psi(v+v', y) + \Psi(x, w+w') = \Psi(v,y) + \Psi(x,w) + \Psi(v',y) + \Psi(x,w') = \phi(v,w) + \phi(v',w')$ für $(v,w), (v',w') \in \mathbb{R}^m \times \mathbb{R}^n$.

14.3 Der Schrankensatz (Mittelwertsatz in mehreren Veränderlichen)

Im Folgenden seien $m, n \in \mathbb{N}_{>0}$.

Bemerkung 97. *Sei $U \subseteq \mathbb{R}^m$ offen, $f\colon U \to \mathbb{R}^n$ eine Abbildung, sowie $p \in U$, $v \in \mathbb{R}^m$ und $\emptyset \neq I \subseteq \mathbb{R}$ ein offenes Intervall mit $p + I \cdot v \subseteq U$, sodass $D_v f(p + t \cdot v)$ für alle $t \in I$ existiert. Dann ist*

$$\gamma\colon I \ni t \mapsto f(p+t\cdot v) \in \mathbb{R}^n \quad \textit{differenzierbar} \quad \textit{mit} \quad \gamma'\colon I \ni t \mapsto D_v f(p+t\cdot v) \in \mathbb{R}^n.$$

Beweis der Behauptung. Gleicher Beweis wie in Bemerkung 93.b). □

Der Mittelwertsatz (Satz 44) liefert die folgende Aussage:

Lemma 81. *Sei $U \subseteq \mathbb{R}^m$ offen, $f\colon U \to \mathbb{R}$ eine Abbildung, sowie $p \in U$, $v \in \mathbb{R}^m$ und $\emptyset \neq I \subseteq \mathbb{R}$ ein offenes Intervall mit $p + I \cdot v \subseteq U$, sodass $D_v f(p + t \cdot v)$ für alle $t \in I$ existiert. Für $I \ni a < b \in I$ vorgegeben, existiert ein $\varsigma \in (a, b)$ mit*

$$\frac{f(p+b\cdot v) - f(p+a\cdot v)}{b-a} = D_v f(p + \varsigma \cdot v). \tag{437}$$

Beweis. Wegen Bemerkung 97 (für $n \equiv 1$) ist $\gamma\colon I \ni t \mapsto f(p + t \cdot v) \in \mathbb{R}$ differenzierbar (also stetig) mit $\gamma'\colon I \ni t \mapsto D_v f(p+t\cdot v) \in \mathbb{R}$. Somit erfüllt $g := \gamma|_{[a,b]}$ die Voraussetzungen von Satz 44 (Mittelwertsatz); also existiert ein $\varsigma \in (a, b)$ mit

$$\frac{f(p+b\cdot v) - f(p+a\cdot v)}{b-a} = \frac{g(b)-g(a)}{b-a} = g'(\varsigma) = D_v f(p+\varsigma\cdot v).$$ □

Hiermit erhalten wir umgehend den folgenden Satz.

Satz 89 (Mittelwertsatz für reellwertige Funktionen). *Sei $U \subseteq \mathbb{R}^m$ offen und $f\colon U \to \mathbb{R}$ eine differenzierbare Abbildung. Seien $p \in U$, $\mathbb{R} \ni a < b \in \mathbb{R}$, sowie $v \in \mathbb{R}^m$ gegeben mit $p + [a, b] \cdot v \subseteq U$. Dann existiert ein $\varsigma \in (a, b)$ mit*

$$\frac{f(p+b\cdot v) - f(p+a\cdot v)}{b-a} = \mathrm{d}_{p+\varsigma\cdot v} f(v).$$

Beweis. Da U offen ist, existiert ein $\epsilon > 0$ mit $p+I\cdot v \subseteq U$ für $I := (a-\epsilon, b+\epsilon)$. Wegen (419) ($D_v f(\bullet) = \mathrm{d}_\bullet f(v)$) gelten die Voraussetzungen in Lemma 81, also (437) mit $D_v f(p+\varsigma\cdot v) = \mathrm{d}_{p+\varsigma\cdot v} f(v)$.

(Begründung der Existenz von $\epsilon > 0$ wie oben: Sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^m$. Da U offen ist, existiert ein $\delta > 0$ mit $\mathsf{B}_\delta(p + a \cdot v) \subseteq U$ sowie $\mathsf{B}_\delta(p + b \cdot v) \subseteq U$. Die Behauptung gilt dann für $\epsilon := \delta/\max(1, \|v\|)$.) □

Bemerkung 98. *Für differenzierbare Abbildungen $f\colon U \to \mathbb{R}^n$ mit $n \geq 2$, gelten die Aussagen in Lemma 81 bzw. Satz 89 im Allgemeinen nicht mehr. Als Beispiel betrachten wir die Spiralkurve*

$$\gamma \equiv f\colon \mathbb{R} \to \mathbb{R}^3, \qquad t \mapsto (\cos(t), \sin(t), t)$$

sowie das Intervall $[0, 2\pi]$, d.h. $m = 1$, $n = 3$, $U = \mathbb{R}$, $p = a = 0$, $b = 2\pi$ und $v = 1$.

- *Es gilt $\gamma(0) = (1, 0, 0)$ und $\gamma(2\pi) = (1, 0, 2\pi)$, also $\gamma(2\pi) - \gamma(0) = (0, 0, 2\pi)$.*
- *Weiterhin gilt nun ((427))*

$$\mathrm{d}_t\gamma(1) = D_1\gamma(t) = \gamma'(t) = (-\sin(t), \cos(t), 1) \qquad \forall\, t \in \mathbb{R},$$

womit die ersten beiden Komponenten von γ' niemals gleichzeitig 0 sein können, d.h.

$$(0,0,1) = \frac{\gamma(2\pi) - \gamma(0)}{2\pi - 0} \neq \mathrm{d}_t\gamma(1) \qquad \forall\, t \in \mathbb{R}.$$

Wegen Bemerkung 98 kann man nicht erwarten, dass Lemma 81 bzw. Satz 89 in allgemeinerer Form auch für $n \geq 2$ korrekt ist. Allerdings erhält man in diesem Fall aus dem Hauptsatz trotzdem eine gute und nützliche Verallgemeinerung, den sogenannten Schrankensatz (Satz 90). Es bezeichne $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$:

Lemma 82. *Sei $U \subseteq \mathbb{R}^m$ offen, $f\colon U \to \mathbb{R}^n$ eine Abbildung, sowie $p \in U$, $v \in \mathbb{R}^m$ und $\emptyset \neq I \subseteq \mathbb{R}$ ein offenes Intervall mit $p + I \cdot v \subseteq U$, sodass $\alpha\colon I \ni t \mapsto D_v f(p + t \cdot v) \in \mathbb{R}^n$ definiert und stetig ist. Dann gilt*

$$f(p + b \cdot v) - f(p + a \cdot v) = \int_a^b \underbrace{D_v f(p + t \cdot v)}_{\alpha(t)} \,\mathrm{d}t \qquad \forall\, I \ni a \neq b \in I. \tag{438}$$

Ist $L \geq 0$ mit $\|D_v f(p + t \cdot v)\| \leq L$ für alle $t \in [\min(a, b), \max(a, b)]$ für $I \ni a \neq b \in I$, so gilt

$$\|f(p + b \cdot v) - f(p + a \cdot v)\| \leq |b - a| \cdot L. \tag{439}$$

Beweis. Es genügt den Fall $a < b$ zu zeigen.[113]

- Wegen Bemerkung 97 ist $\gamma\colon I \ni t \mapsto f(p + t \cdot v) \in \mathbb{R}^n$ differenzierbar mit $\gamma' = \alpha$, also stetig differenzierbar. Korollar 55.2) (verallgemeinerter Hauptsatz) liefert im zweiten Schritt:

$$f(p + b \cdot v) - f(p + a \cdot v) = \gamma(b) - \gamma(a) = \textstyle\int_a^b \gamma' = \int_a^b \alpha = \int_a^b D_v f(p + t \cdot v)\,\mathrm{d}t.$$

- Wir erhalten aus Übung 147 (zweiter Schritt) und den elementaren Integrationsregeln:

$$\|f(p + b \cdot v) - f(p + a \cdot v)\| \overset{(438)}{=} \|\textstyle\int_a^b \alpha\| \leq \int_a^b \|\alpha\| \leq \int_a^b L = (b - a) \cdot L. \qquad \square$$

Satz 90 (Schrankensatz)**.** *Sei $U \subseteq \mathbb{R}^m$ offen und $f\colon U \to \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Seien $p \in U$, $\mathbb{R} \ni a < b \in \mathbb{R}$, sowie $v \in \mathbb{R}^m$ gegeben mit $p + [a, b] \cdot v \subseteq U$. Dann gilt*

$$f(p + b \cdot v) - f(p + a \cdot v) = \int_a^b \mathrm{d}_{p+t\cdot v} f(v)\,\mathrm{d}t.$$

Ist $L \geq 0$ mit $\|\mathrm{d}_{p+t\cdot v} f(v)\| \leq L$ für alle $t \in [a, b]$, so gilt

$$\|f(p + b \cdot v) - f(p + a \cdot v)\| \leq (b - a) \cdot L.$$

Beweis. Wegen Lemma 80 existiert $D_v f$ und ist stetig, mit $D_v f(\bullet) = \mathrm{d}_\bullet f(v)$.

- Sei $\epsilon > 0$ mit $p + (a - \epsilon, b + \epsilon) \cdot v \subseteq U$, sowie $I := (a - \epsilon, b + \epsilon)$. (vgl. Beweis von Satz 89)
- Als Verkettung stetiger Abbildungen ist $\alpha\colon I \ni t \mapsto D_v f(p + t \cdot v) = \mathrm{d}_{p+t\cdot v} f(v) \in \mathbb{R}^n$ (definiert und) stetig; also folgt die Behauptung unmittelbar aus Lemma 82 . $\square$

Übung 157. *Sei $m, n \in \mathbb{N}_{>0}$, sowie $\|\cdot\|_m$ eine Norm auf $\mathbb{R}^m$ und $\|\cdot\|_n$ eine Norm auf $\mathbb{R}^n$. Sei zudem $U \subseteq \mathbb{R}^m$ offen und $f \in C^1(U, \mathbb{R}^n)$. Zeigen Sie, dass f* lokal Lipschitzstetig *ist, d.h., für jedes $p \in U$ existiert ein $\mathsf{L} \geq 0$ und eine offene Umgebung V von p mit $V \subseteq U$, sodass gilt:*

$$\|f(q) - f(q')\|_m \leq \mathsf{L} \cdot \|q - q'\|_n \qquad \forall\, q, q' \in V. \tag{440}$$

Bemerkung* 18. *Die zweite Aussage in Satz 90 bzw. Lemma 82 gilt auch noch (sogar in allgemeinerer Form), wenn man f nur als differenzierbar voraussetzt:*

[113] Der Fall $b < a$ folgt dann sofort aus dem Bewiesenen. Für (439) ist dies offensichtlich, und für (438) folgt sie sofort aus unserer Konvention (394).

- *Eine Teilmenge $C \subseteq \mathbb{R}^n$ heißt konvex, wenn für jedes $\ell \in \mathbb{N}_{>0}$ gilt:*

$$\textstyle\sum_{j=0}^{\ell-1} \lambda_\ell \cdot w_\ell \in C \qquad \forall\, w_0, \dots, w_{\ell-1} \in C;\ \ 0 \leq \lambda_0, \dots, \lambda_{\ell-1} \leq 1\colon\ \sum_{j=0}^{\ell-1} \lambda_j = 1.$$

- *Sei beispielsweise $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$. Dann ist für $x \in \mathbb{R}^n$ und $\varepsilon > 0$, der offene Ball $\mathsf{B}_\varepsilon(x)$ konvex; und zwar wegen*

$$\big\| \textstyle\sum_{j=0}^{\ell-1} \lambda_\ell \cdot w_\ell - x\big\| = \big\| \sum_{j=0}^{\ell-1} \lambda_\ell \cdot w_\ell - \overbrace{\sum_{j=0}^{\ell-1} \lambda_\ell}^{=1} \cdot x\big\| \leq \sum_{j=0}^{\ell-1} \lambda_j \cdot \|w_j - x\| \leq \varepsilon \cdot \sum_{j=0}^{\ell-1} \lambda_j = \varepsilon$$

für $\ell \in \mathbb{N}_{>0}$ sowie $w_0, \dots, w_{\ell-1} \in \mathsf{B}_\varepsilon(x)$ und $0 \leq \lambda_1, \dots, \lambda_{\ell-1} \leq 1$ mit $\sum_{j=0}^{\ell-1} \lambda_\ell = 1$.

Lemma* 1. *Sei $C \subseteq \mathbb{R}^n$ eine offene konvexe Teilmenge, und $\gamma\colon [a,b] \to \mathbb{R}^n$ eine differenzierbare Kurve mit $\mathrm{im}[\gamma'] \subseteq C$. Dann gilt*

$$\frac{\gamma(b) - \gamma(a)}{b-a} \in C. \tag{441}$$

Beweis. Zu jedem $s \in [a,b]$ existiert ein $\delta_s > 0$ mit

$$\frac{\gamma(s+h)\ -\ \gamma(s)}{h} \in C \quad \text{für alle} \quad h \in (\delta_s, \delta_s) \quad \text{mit} \quad s+h \in [a,b]; \tag{442}$$

denn C ist offen, und es gilt $\lim_{h\to 0} \frac{\gamma(s+h)\ -\ \gamma(s)}{h} = \gamma'(s) \in C$.

- Insbesondere ist: $\quad M := \big\{a < \tau \leq b \mid \frac{\gamma(t)\ - \gamma(a)}{t-a} \in C \quad \text{für alle} \quad a \leq t \leq \tau\big\} \neq \emptyset$

  nichtleer ((442) für $s \equiv a$); also existiert $a < s := \sup_{\mathbb{R}}(M) \leq b$ (beachte $M \leq b$).
- Sei $\tau \in (s - \delta_s, s) \cap [a,b]$ fixiert. Für $t \in [s, s+\delta_s) \cap [a,b]$ folgt: $\qquad (\lambda_1 + \lambda_2 + \lambda_3 = 1)$

$$\frac{\gamma(t) - \gamma(a)}{t-a} = \underbrace{\frac{t-s}{t-a}}_{\lambda_1} \cdot \underbrace{\frac{\gamma(t) - \gamma(s)}{t-s}}_{\in C} + \underbrace{\frac{s-\tau}{t-a}}_{\lambda_2} \cdot \underbrace{\frac{\gamma(s) - \gamma(\tau)}{s-\tau}}_{\in C} + \underbrace{\frac{\tau-a}{t-a}}_{\lambda_3} \cdot \underbrace{\frac{\gamma(\tau) - \gamma(a)}{\tau-a}}_{\in C} \in C. \tag{443}$$

Wegen (443) mit $t \equiv s$, gilt $s \in M$. Gilt nun $s < b$, so existiert wegen (443) ein $s < t \leq b$ mit $t \in M$, was der Definition von s widerspricht. Somit gilt $b = s \in M$, also (441). □

14.4 Vertauschbarkeit von Integral und Ableitung

In diesem kurzen Abschnitt lernen wir eine wichtige Rechenmethoden kennen, die das Vertauschen von Ableiten und Integration betrifft, sofern verschiedene Variablen betroffen sind. Wir erinnern an die Maximumsmetrik d_∞ auf Produkten metrischer Räume aus Terminologie 31:

- Wir betrachten zunächst den Fall $[a,b] \times X$ mit einem metrischen Raum (X, d), wobei $[a,b] \subseteq \mathbb{R}$ mit $a < b$ natürlich mit der Betragsmetrik versehen ist (siehe Proposition 24).
- Danach betrachten wir den Fall $X \equiv U \subseteq \mathbb{R}^m$ für $m \in \mathbb{N}_{>0}$, wobei wir $\mathbb{R}^m$ und $\mathbb{R} \times \mathbb{R}^m \cong \mathbb{R}^{m+1}$ mit irgendeiner Normmetrik ausstatten (siehe Satz 91). Wegen Satz 79 (Äquivalenz von Normen) sind dann die Begriffe Stetigkeit, Differenzierbarkeit und Offenheit unabhängig von dieser Wahl.

  (Beispielweise können wir für d einfach die Einschränkung auf U der, durch die Maximumsnorm auf $\mathbb{R}^m$ induzierte Metrik wählen. In diesem Fall ist dann die entsprechende Maximumsmetrik d_∞ auf $Y := [a,b] \times U \subseteq \mathbb{R}^{m+1}$ einfach gegeben durch die Einschränkung der, durch die Maximumsnorm $\|\cdot\|_\infty$ auf $\mathbb{R}^{m+1}$ induzierten Metrik (vgl. Bemerkung 87.b)), d.h. $\mathrm{d}_\infty = \mathrm{d}_{\|\cdot\|_\infty}|_{Y\times Y}$.)

Proposition 24. *Seien (X, d) ein metrischer Raum, sowie $\mathbb{R} \ni a < b \in \mathbb{R}$, $k \in \mathbb{N}_{>0}$, und $f\colon [a,b] \times X \to \mathbb{R}^k$ eine stetige Abbildung. Dann ist die folgende Abbildung ebenfalls stetig:*

$$F\colon X \to \mathbb{R}^k, \qquad x \mapsto \int_a^b f(t,x)\,\mathrm{d}t.$$

Beachte: Wegen Übung 144 ist $\gamma_x\colon [a,b] \ni t \mapsto f(t,x) \in \mathbb{R}^k$ für jedes fixierte $x \in X$ eine stetige Kurve, also auch Riemann-integrierbar.

Beweis. Wir zeigen die Stetigkeit von F in einem fixierten Punkt $x \in X$. Sei hierfür $\varepsilon > 0$ vorgegeben. Da f stetig ist, existiert zu jedem $s \in [a,b]$ ein $\delta_s > 0$ mit ($\|\cdot\|$ eine Norm auf $\mathbb{R}^k$)

$$\mathrm{d}_\infty((t,y),(s,x)) < \delta_s \quad \text{für} \quad (t,y) \in [a,b]\times X \qquad \Longrightarrow \qquad \|f(t,y) - f(s,x)\| < \tfrac{\varepsilon}{2(b-a)}. \tag{444}$$

Nun ist $\big((s-\delta_s, s+\delta_s)\big)_{s\in[a,b]}$ eine offene Überdeckung von $[a,b]$, hat also wegen der Kompaktheit von $[a,b]$ eine endliche Teilüberdeckung, d.h.

$$\exists\, s_1,\dots,s_n \in [a,b]\colon \quad [a,b] \subseteq \bigcup_{\ell=1}^{n} (s_\ell - \delta_{s_\ell}, s_\ell + \delta_{s_\ell}).$$

Wir setzen $\delta := \min(\delta_{s_1},\dots,\delta_{s_n})$. Für $y \in \mathsf{B}_\delta(x)$ und $t \in [a,b] \cap (s_\ell - \delta_{s_\ell}, s_\ell + \delta_{s_\ell})$ mit $1 \le \ell \le n$, gilt

$$\mathrm{d}_\infty((t,x),(s_\ell,x)) = |t-s_\ell| < \delta_{s_\ell} \qquad \text{sowie} \qquad \mathrm{d}_\infty((t,y),(s_\ell,x)) = \max(|t-s_\ell|, \mathrm{d}(y,x)) < \delta_{s_\ell},$$

und wir erhalten mit der Dreiecksungleichung

$$\|f(t,x) - f(t,y)\| \le \|f(t,x) - f(s_\ell,x)\| + \|f(s_\ell,x) - f(t,y)\| \overset{(444)}{<} 2\cdot \tfrac{\varepsilon}{2(b-a)} = \tfrac{\varepsilon}{b-a}. \tag{445}$$

Da jedes $t \in [a,b]$ in einem $(s_\ell - \delta_{s_\ell}, s_\ell + \delta_{s_\ell})$ mit $1 \le \ell \le n$ enthalten ist, gilt (445) für alle $t \in [a,b]$ und alle $y \in \mathsf{B}_\delta(x)$. Wir erhalten aus der Linearität, Monotonie, und Normierung der Riemann-Integrals, sowie Übung 147 (Integralabschätzung)[114]

$$\|F(x) - F(y)\| = \left\| \textstyle\int_a^b (f(t,x) - f(t,y))\,\mathrm{d}t \right\| \overset{(399)}{\le} \textstyle\int_a^b \|f(t,x) - f(t,y)\|\,\mathrm{d}t \le \int_a^b \frac{\varepsilon}{b-a}\,\mathrm{d}t = \varepsilon$$

für alle $y \in \mathsf{B}_\delta(x)$, was die Behauptung zeigt. □

Satz 91. *Sei $\emptyset \neq I \subseteq \mathbb{R}$ ein offenes Intervall und $\emptyset \neq U \subseteq \mathbb{R}^m$ eine offene Teilmenge. Sei $v \in \mathbb{R}^m$ sowie $f\colon I \times U \to \mathbb{R}^k$ eine stetige Abbildung, sodass die Richtungsableitung*[115]

$$D_{(0,v)}f\colon I\times U \to \mathbb{R}^k, \qquad (t,x) \mapsto \lim_{h\to 0} \frac{f(t,x+h\cdot v) - f(t,x)}{h}$$

existiert und stetig ist. Für $I \ni a < b \in I$ ist dann die Abbildung

$$F\colon U \to \mathbb{R}^k, \qquad x \mapsto \int_a^b f(t,x)\,\mathrm{d}t$$

stetig, mit stetiger Richtungsableitung

$$D_vF\colon U \to \mathbb{R}^k, \qquad x \mapsto \int_a^b D_{(0,v)}f(t,x)\,\mathrm{d}t.$$

(Beachte, dass $I\times U \subseteq \mathbb{R}^{m+1}$ wegen Bemerkung 87 offen ist (wegen Satz 79 ungeachtet der expliziten Wahl der Norm auf $\mathbb{R}^{m+1}$), also unsere Definitionen aus Terminologie 40 anwendbar sind. Man beachte weiterhin, dass obige Aussage für $a = b$ trivialerweise ebenfalls gilt.)

[114] Alternativ kann man natürlich auch einfach $\|\cdot\| = \|\cdot\|_\infty$ wählen und Bemerkung 91.(d) verwenden.

[115] Wir schreiben im Folgenden auch vereinfachend $(s, x+\lambda\cdot v)$ anstelle $(s,x) + \lambda\cdot(0,v)$ für $s \in I$ und $(\lambda,x) \in \mathbb{R}\times U$.

Beweis. Wir fixieren Normen auf $\mathbb{R}^m, \mathbb{R}^{m+1}, \mathbb{R}^k$, wobei wir die Norm auf $\mathbb{R}^k$ mit $\|\cdot\|$ bezeichnen. Zunächst sind nun wegen Proposition 24 beide Integrale definiert, und stellen stetige Funktionen $U \to \mathbb{R}^k$ dar. Sei $p \in U$ vorgegeben, sowie $\epsilon > 0$ mit $p + (-\epsilon, \epsilon) \cdot v \subseteq U$. Wir setzen

$$\Lambda\colon [a,b] \times (-\epsilon, 0) \cup (0, \epsilon) \to \mathbb{R}^k, \qquad (t,h) \mapsto \underbrace{\frac{f(\overbrace{(t,p) + s \cdot (0,v)}^{(t,\, p + s\cdot v)}) - f(t,p)}{h}}_{=:\, \Omega(t,h)} - D_{(0,v)} f(t,p).$$

Wegen Lemma 82 gilt: $\quad (\alpha\colon (-\epsilon, \epsilon) \ni s \mapsto D_{(0,v)} f((t,p) + s \cdot (0,v)) \in \mathbb{R}^k$ ist definiert und stetig)

$$\Omega(t,h) = \frac{1}{h} \cdot \int_0^h D_{(0,v)} f(\overbrace{(t,p) + s \cdot (0,v)}^{(t,\, p + s\cdot v)}) \, \mathrm{d}s \qquad \forall\, t \in [a,b],\ 0 \neq h \in (-\epsilon, \epsilon).$$

- Sei $\varepsilon > 0$ vorgegeben. Es existiert $\delta > 0$ mit $\|\Lambda(t,h)\| < \frac{\varepsilon}{b-a}$ für alle $t \in [a,b]$ und $0 \neq h \in (-\delta, \delta)$:

  In der Tat, wegen der Stetigkeit von $D_{(0,v)} f$ und der Kompaktheit von $[a,b]$, folgt genau wie in Proposition 24 die Existenz eines $\delta > 0$ mit (vgl. (445))

  $$\|D_{(0,v)} f(t, p + s \cdot v) - D_{(0,v)} f(t,p)\| < \frac{\varepsilon}{b-a} \qquad \forall\, t \in [a,b],\ s \in (-\delta, \delta).$$

  Wir erhalten für $0 \neq h \in (-\delta, \delta)$: (Übung 147 im dritten Schritt)

  $$\begin{aligned} \|\Lambda(t,h)\| &= \left\| \Omega(t,h) - D_{(0,v)} f(t,p) \right\| = \left\| \frac{1}{h} \cdot \int_0^h D_{(0,v)} f(t, p + s \cdot v) \, \mathrm{d}s - \frac{1}{h} \cdot \int_0^h D_{(0,v)} f(t,p) \, \mathrm{d}s \right\| \\ &\overset{(399)}{\leq} \frac{1}{|h|} \cdot \left| \int_0^h \left\| D_{(0,v)} f(t, p + s \cdot v) - D_{(0,v)} f(t,p) \right\| \mathrm{d}s \right| \leq \frac{\varepsilon}{b-a}. \end{aligned}$$

- Wir erhalten für $0 \neq h \in (-\delta, \delta)$:

  $$\Bigg\| \overbrace{\underbrace{\frac{F(p + h \cdot v) - F(p)}{h}}_{\int_a^b \frac{f(t,p+h\cdot v) - f(t,p)}{h} \mathrm{d}t} - \int_a^b D_{(0,v)} f(t,p) \, \mathrm{d}t}^{\int_a^b \left(\frac{f(t,p+h\cdot v) - f(t,p)}{h} - D_{(0,v)} f(t,p) \right) \mathrm{d}t} \Bigg\| \overset{(399)}{\leq} \int_a^b \|\Lambda(t,h)\| \, \mathrm{d}t \leq (b-a) \cdot \frac{\varepsilon}{b-a} = \varepsilon. \qquad \square$$

15 Höhere Ableitungen, Taylorrentwicklung und Extremwerte

In diesem Abschnitt wenden wir uns partiellen Ableitungen höherer Ordnung zu. Wir klären zunächst deren allgemeine Eigenschaften, und diskutieren dann Taylorpolynome von C^k-Abbildungen $\mathbb{R}^n \supseteq U \to \mathbb{R}^m$. Schließlich leiten wir mit Hilfe des Taylorpolynomes zweiter Ordnung hinreichende Bedingungen für lokale Extrema für reellwertige Funktionen mehrerer Veränderlicher her.

15.1 Höhere Ableitungen und der Satz von Schwarz

Terminologie 42. *Sei $\emptyset \neq U \subseteq \mathbb{R}^m$ offen und $f\colon U \to \mathbb{R}^n$ eine Abbildung mit $m, n \in \mathbb{N}_{>0}$.*

1) *f heißt* k-mal partiell differenzierbar *für $k \in \mathbb{N}_{>0}$, wenn für alle $1 \leq p \leq k$ und alle $1 \leq j_1, \ldots, j_p \leq m$, die iterierte partielle Ableitung*

$$\partial_{j_1} \ldots \partial_{j_p} f\colon U \to \mathbb{R}^m, \qquad x \mapsto \partial_{j_1}(\partial_{j_2}(\ldots \partial_{j_{p-1}}(\partial_{j_p} f) \ldots))(x)$$

existiert. Alternativ notiert man auch

$$\partial_{j_1} \dots \partial_{j_p} f \equiv \frac{\partial f}{\partial_{x_{j_1}} \dots \partial_{x_{j_p}}} \qquad \forall\, 1 \leq p \leq k,\ 1 \leq j_1, \dots, j_p \leq m. \tag{446}$$

Ein Ausdruck der Form (446) *wird* partielle Ableitung der Ordnung p *genannt. (f selbst fassen wir als partielle Ableitung von f der Ordnung 0 auf.)*

2) *f heißt* k-mal stetig partiell differenzierbar *für $k \in \mathbb{N}_{>0}$, wenn f stetig sowie k-mal partiell differenzierbar ist, und alle partiellen Ableitungen $\partial_{j_1} \dots \partial_{j_p} f$ mit $1 \leq p \leq k$ und $1 \leq j_1, \dots, j_p \leq m$ stetig sind.*

   Der Raum aller k-mal stetig partiell differenzierbaren Abbildungen von U nach $\mathbb{R}^n$ wird mit $C^k(U, \mathbb{R}^n)$ notiert. (Für $k = 1$ ist diese Definition wegen Satz 87 konsistent mit der entsprechenden Notation aus Terminologie 41.(2)*.)*

   - *Ist f stetig, so sagen wir, dass f 0-mal stetig partiell differenzierbar ist.*
   - *f heißt glatt, wenn $f \in C^\infty(U, \mathbb{R}^n) := \bigcap_{k \in \mathbb{N}} C^k(U, \mathbb{R}^n)$ gilt, d.h. wenn f k-mal stetig partiell differenzierbar ist für jedes $k \in \mathbb{N}$.*

Beispiel 104. *Wir betrachten die Funktion $f\colon \mathbb{R}^2 \ni (x, y) \mapsto 3x^2 y + y^3 \in \mathbb{R}$. Dann ist f zweimal stetig partiell differenzierbar (sogar glatt), mit den partiellen Ableitungen:*

$$\partial_1 f(x, y) = 6xy, \qquad \partial_2 f(x, y) = 3x^2 + 3y^2,$$
$$\partial_1 \partial_1 f(x, y) = 6y, \qquad \partial_2 \partial_2 f(x, y) = 6y, \qquad \partial_1 \partial_2 f(x, y) = 6x = \partial_2 \partial_1 f(x, y).$$

Wir werden in Satz 92 (Satz von Schwarz) sehen, dass es an dieser Stelle kein Zufall ist, dass die gemischten partiellen Abbildungen übereinstimmen ($\partial_1 \partial_2 f = \partial_2 \partial_1 f$).

Wir werden das nächste Lemma im Folgenden frei verwenden.

Lemma 83. *Seien $m, n \in \mathbb{N}_{>0}$, $k \in \mathbb{N}_{>0} \cup \{\infty\}$, $\emptyset \neq U \subseteq \mathbb{R}^m$ offen, und $f\colon U \to \mathbb{R}^n$ eine Abbildung.*

1) *Ist f eine k-mal (stetig) partiell differenzierbare Abbildung, so auch $f|_V$ für jede offene Teilmenge $V \subseteq \mathbb{R}^m$ mit $V \subseteq U$, und es gilt:*

$$\partial_{j_1} \dots \partial_{j_p} (f|_V) = (\partial_{j_1} \dots \partial_{j_p} f)|_V \qquad \forall\, 1 \leq p \leq k,\ 1 \leq j_1, \dots, j_p \leq m. \tag{447}$$

2) *Sei $(O_\alpha)_{\alpha \in I}$ eine offene Überdeckung von U, sodass $f|_{O_\alpha}$ für jedes $\alpha \in I$ eine k-mal (stetig) partiell differenzierbare Abbildung ist. Dann ist auch f eine k-mal (stetig) partiell differenzierbare Abbildung mit*

$$(\partial_{j_1} \dots \partial_{j_p} f)|_{O_\alpha} = \partial_{j_1} \dots \partial_{j_p} (f|_{O_\alpha}) \qquad \forall\, 1 \leq p \leq k,\ 1 \leq j_1, \dots, j_p \leq m,\ \alpha \in I. \tag{448}$$

Beweis. 1) Wir müssen nur die Differenzierbarkeitsaussage zeigen, denn die Stetigkeitsaussage folgt dann umgehend aus Lemma 40.2) und (447). Für $p = 1$ ist (447) klar wegen Übung 152, was auch bereits den Fall $k = 1$ abdeckt. Sei also $k \geq 2$; und es gelte (447) für ein $1 \leq p \leq k - 1$. Für $1 \leq j_0, \dots, j_p \leq m$, erhalten wir aus Übung 152 (zweiter Schritt) sowie der Induktionsvoraussetzung (dritter Schritt):

$$\begin{aligned}(\partial_{j_0} \dots \partial_{j_p} f)|_V = \big(\partial_{j_0} (\partial_{j_1} \dots \partial_{j_p} f)\big)|_V &= \partial_{j_0} \big((\partial_{j_1} \dots \partial_{j_p} f)|_V\big) \\ = \partial_{j_0} \big(\partial_{j_1} \dots \partial_{j_p} (f|_V)\big) &= \partial_{j_0} \dots \partial_{j_p} (f|_V).\end{aligned}$$

2) Wir müssen nur die Differenzierbarkeitsaussage zeigen, denn die Stetigkeitsaussage folgt dann umgehend aus Übung 79.c). Für $p = 1$ ist (448) klar wegen Übung 152, was auch bereits den Fall $k = 1$ abdeckt. Sei also $k \geq 2$; und es gelte (448) für ein $1 \leq p \leq k-1$. Für $1 \leq j_0, \dots, j_p \leq m$, erhalten wir aus der Induktionsvoraussetzung (zweiter Schritt) sowie Übung 152 (letzter Schritt):

$$\partial_{j_0}\dots\partial_{j_p}(f|_{O_\alpha}) = \partial_{j_0}(\partial_{j_1}\dots\partial_{j_p}(f|_{O_\alpha})) = \partial_{j_0}((\partial_{j_1}\dots\partial_{j_p}f)|_{O_\alpha}) = \underbrace{\partial_{j_0}\dots\partial_{j_p}(f|_{O_\alpha})}_{(\partial_{j_0}(\partial_{j_1}\dots\partial_{j_p}f)|_{O_\alpha})} \quad \forall\, \alpha \in I. \quad \square$$

Lemma 84. *Sei $\emptyset \neq U \subseteq \mathbb{R}^m$ offen, und $f = (f_1, \dots, f_n)\colon U \to \mathbb{R}^n$ eine Abbildung mit $m, n \in \mathbb{N}_{>0}$.*

1) Sei $p \in \mathbb{N}_{>0}$. Gegeben $1 \leq j_1, \dots, j_p \leq m$, so existiert $\partial_{j_1} \dots \partial_{j_p} f$ (und ist stetig) genau dann, wenn $\partial_{j_1} \dots \partial_{j_p} f_\ell$ für alle $1 \leq \ell \leq n$ existiert (und stetig ist); und dann gilt

$$\partial_{j_1} \dots \partial_{j_p} f = ((\partial_{j_1} \dots \partial_{j_p} f_1), \dots, (\partial_{j_1} \dots \partial_{j_p} f_n)). \tag{449}$$

Beachte: *Wegen Lemma 74 gilt selbiges für die 0-ten partiellen Ableitungen von f, also die Komponentenfunktionen $f_1, \dots, f_n$.*

2) Sei f k-mal partiell differenzierbar für ein $k \in \mathbb{N}_{>0}$. Sind alle partiellen Ableitungen vom Grad k stetig, so ist f bereits k-mal stetig partiell differenzierbar, d.h., alle partiellen Ableitungen vom Grad $0 \leq p \leq k$ sind stetig (insbesondere ist f stetig).

Beweis. 1) Wir müssen nur die Differenzierbarkeitsaussage zeigen, denn die Stetigkeitsaussage folgt dann umgehend aus Lemma 74 und (449). Für $p = 1$ ist (449) klar wegen Terminologie 40.(1).a). Wir können daher annehmen, dass (449) für ein $p \in \mathbb{N}_{>0}$ gilt. Für $1 \leq j_0, \dots, j_p \leq m$ gilt dann im jeweiligen Falle der Existenz der linken bzw. der rechten Seite die Gleichungskette

$$\underbrace{\partial_{j_0} \dots \partial_{j_p} f}_{\partial_{j_0}\big(\underbrace{\partial_{j_1} \dots \partial_{j_p} f}_{g}\big)} \overset{(*)}{=} \partial_{j_0}\big(\underbrace{(\partial_{j_1} \dots \partial_{j_p} f_1)}_{g_1}, \dots, \underbrace{(\partial_{j_1} \dots \partial_{j_p} f_n)}_{g_n}\big) \overset{(**)}{=} \underbrace{((\partial_{j_0} \dots \partial_{j_p} f_1), \dots, (\partial_{j_0} \dots \partial_{j_p} f_n))}_{(\partial_{j_0}(\underbrace{\partial_{j_1} \dots \partial_{j_p} f_1}_{g_1}),\ \dots,\ \partial_{j_0}(\underbrace{\partial_{j_1} \dots \partial_{j_p} f_n}_{g_n}))}$$

wobei in $(*)$ die Induktionvoraussetzung (449) und in $(**)$ Terminologie 40.(1).a) verwendet wird. Hierbei ist zu beachten:

- Existiert die Linke Seite, so definieren wir $g := \partial_{j_1} \dots \partial_{j_p} f$, und es bezeichnen $g_1, \dots, g_n$ die zugehörigen Komponentenfunktionen.
- Existiert die rechte Seite, so definieren wir $g_\ell := \partial_{j_1} \dots \partial_{j_p} f_\ell$ für $1 \leq \ell \leq n$, und setzen $g := (g_1, \dots, g_n)$.

2) Per Voraussetzung existiert ein $1 \leq p \leq k$, sodass für jedes $p \leq \ell \leq k$ alle partiellen Ableitungen der Ordnung ℓ stetig sind.

- $p = 1$: f ist differenzierbar wegen Satz 87, und somit stetig wegen Lemma 78.
- $p \geq 2$: Für $1 \leq j_1, \dots, j_{p-1} \leq m$ vorgegeben, ist $\partial_{j_1} \dots \partial_{j_{p-1}} f$ einmal stetig partiell differenzierbar; mithin differenzierbar wegen Satz 87, also stetig wegen Lemma 78.

Hiermit folgt die Behauptung induktiv. $\square$

Bemerkung 99. *Im Kontext von Lemma 84.2) könnte man sich von Lemma 47 für den eindimensionalen Fall ($m = 1$) evtl. dazu verleiten lassen anzunehmen, dass $f \in C^k(U, \mathbb{R}^n)$ gilt falls f $(k+1)$-mal partiell differenzierbar ist. In Bemerkung 94 (das ist der Fall $k = 0$) hatte wir allerdings bereits Gegenteiliges festgestellt.*

Gemäß Terminologie 41.(2) können wir das Differential einer (stetig) differenzierbaren Abbildung $f\colon \mathbb{R}^m \supseteq U \to \mathbb{R}^n$ auch als (stetige) Abbildung

$$\widetilde{\mathrm{d}f}\colon U \to \mathbb{R}^{m\cdot n}, \qquad p \mapsto \big(\partial_1 f_1(p), \dots, \partial_m f_1(p), \dots, \partial_1 f_n(p), \dots, \partial_m f_n(p)\big)$$

auffassen, d.h., die Komponenten von $\widetilde{\mathrm{d}f}$ sind stetig. Mit dieser Identifikation gilt die folgende Aussage, welche wir in Abschnitt 16.3 zur Anwendung bringen werden:

Lemma 85. *Sei $m, n, k \in \mathbb{N}_{>0}$, $\emptyset \neq U \subseteq \mathbb{R}^m$ offen. Eine Abbildung $f\colon U \to \mathbb{R}^n$ ist genau dann C^k (k-mal stetig differenzierbar), wenn f differenzierbar und $\widetilde{\mathrm{d}f}$ eine C^{k-1}-Abbildung ist ($k-1$-mal stetig differenzierbar).*

Beweis. • Es existiere $\mathrm{d}f$, und es gelte $\widetilde{\mathrm{d}f} \in C^{k-1}(U, \mathbb{R}^{m\cdot n})$.

  – Wegen Lemma 84.1) angewandt auf $\widetilde{\mathrm{d}f}$, existiert $\partial_{j_1} \dots \partial_{j_k} f_\ell$ für alle $1 \leq \ell \leq n$, $1 \leq j_1, \dots, j_k \leq m$ und ist stetig.
  – Wegen Lemma 84.1) angewandt auf f, existiert $\partial_{j_1} \dots \partial_{j_k} f$ für alle $1 \leq j_1, \dots, j_k \leq m$ und ist stetig. Lemma 84.2) zeigt nun $f \in C^k(U, \mathbb{R}^n)$.

• Sei $f \in C^k(U, \mathbb{R}^m)$. Wegen Lemma 84.1) sind alle Komponenten (f_ℓ mit $1 \leq \ell \leq n$) von f k-mal stetig partiell differenzierbar; also sind alle Komponenten von $\widetilde{\mathrm{d}f}$ $(k-1)$-mal stetig partiell differenzierbar. Lemma 84.1) zeigt nun $\widetilde{\mathrm{d}f} \in C^{k-1}(U, \mathbb{R}^n)$. □

Übung 158. *Seien $k, m, n \in \mathbb{N}_{>0}$, $\emptyset \neq U \subseteq \mathbb{R}^m$ offen, $\lambda, \mu \in \mathbb{R}$, sowie $f, g\colon U \to \mathbb{R}^n$ beide k-mal (stetig) partiell differenzierbar. Zeigen Sie per Induktion, mit Hilfe der Summenregel in Terminologie 40.*(1).c)*, dass $(\lambda \cdot f + \mu \cdot g)$ ebenfalls k-mal partiell differenzierbar ist, mit*

$$\partial_{j_1} \dots \partial_{j_\ell}(\lambda \cdot f + \mu \cdot g) = \lambda \cdot \partial_{j_1} \dots \partial_{j_\ell} f + \mu \cdot \partial_{j_1} \dots \partial_{j_\ell} g$$

für alle $1 \leq j_1, \dots, j_\ell \leq m$ und $1 \leq \ell \leq k$. Folgern Sie, dass $(\lambda \cdot f + \mu \cdot g) \in C^k(U, \mathbb{R}^n)$ gilt, falls f und g beides C^k-Abbildungen sind.

Übung 159. *Seien $k, m, n \in \mathbb{N}_{>0}$, $\emptyset \neq U \subseteq \mathbb{R}^m$ offen, sowie $f\colon U \to \mathbb{R}$ und $g\colon U \to \mathbb{R}^n$ beide k-mal (stetig) partiell differenzierbar. Zeigen Sie, dass dann $(f \cdot g)$ ebenfalls k-mal (stetig) partiell differenzierbar ist.*

Hinweis: *Zeigen Sie per Induktion, mit Hilfe der Produktregel in Terminologie 40.*(1).d)*, dass für jedes $1 \leq p \leq k$ und alle $1 \leq j_1, \dots, j_p \leq m$, die partielle Ableitung $\partial_{j_1} \dots \partial_{j_p}(f \cdot g)$ als eine Linearkombination von Ausdrücken der Form*

$$\partial_{r_1} \dots \partial_{r_\ell} f \cdot \partial_{s_1} \dots \partial_{s_{p-\ell}} g \qquad \text{mit} \qquad 0 \leq \ell \leq p \qquad \text{und} \qquad r_1, \dots, r_\ell, s_1, \dots, s_{p-\ell} \in \{j_1, \dots, j_p\}$$

geschrieben werden kann; wobei für $\ell = 0$ der linke Faktor gleich f zu setzen ist, sowie der rechte Faktor gleich g falls $\ell = p$ gilt. Wegen Lemma 84 können Sie der Einfachheit halber $n = 1$ annehmen.

Übung 160. *Seien $m, n \in \mathbb{N}_{>0}$, $\emptyset \neq U \subseteq \mathbb{R}^m$ offen, $v \in \mathbb{R}^m$, $p \in U$, sowie $f\colon U \to \mathbb{R}_\times$ und $g\colon U \to \mathbb{R}^n$ Abbildungen, sodass die Richtungsableitungen $D_v f(p)$ und $D_v g(p)$ existieren. Zeigen Sie:*

$$D_v\left(\frac{g}{f}\right)(p) = \frac{D_v g(p) \cdot f(p) - g(p) \cdot D_v f(p)}{f(p)^2}.$$

Folgern Sie, dass $\frac{g}{f}$ (stetig) partiell differenzierbar ist, wenn f, g (stetig) partiell differenzierbar sind.

Übung 161. *Seien $k, m, n \in \mathbb{N}_{>0}$, $\emptyset \neq U \subseteq \mathbb{R}^m$ offen, sowie $f\colon U \to \mathbb{R}_\times$ und $g\colon U \to \mathbb{R}^n$ beide k-mal (stetig) partiell differenzierbar. Zeigen Sie, dass $\frac{g}{f}$ k-mal (stetig) partiell differenzierbare ist.*

Hinweis: Zeigen Sie per Induktion über $1 \leq p \leq k$: Für $1 \leq j_1, \ldots, j_p \leq m$ vorgegeben, existiert ein $z \in \mathbb{N}_{>0}$ und eine $(k-p)$-mal (stetig) partiell differenzierbare Abbildung $\alpha\colon U \to \mathbb{R}^n$ mit

$$\partial_{j_1} \ldots \partial_{j_p}\left(\frac{g}{f}\right) = \frac{\alpha}{f^z}. \tag{450}$$

Der folgende wichtige Satz zeigt, dass es kein Zufall ist, dass im obigem Beispiel 104 die „gemischten" partiellen Ableitungen $\partial_1\partial_2 f$ und $\partial_2\partial_1 f$ miteinander übereinstimmen.

Satz 92 (Satz von Schwarz)**.** *Sei $\emptyset \neq U \subseteq \mathbb{R}^m$ offen und $f \in C^2(U, \mathbb{R}^n)$ mit $m, n \in \mathbb{N}_{>0}$. Dann gilt:*

$$\partial_i\partial_j f = \partial_j\partial_i f \qquad \forall\, 1 \leq i, j \leq m. \tag{451}$$

Beweis. Seien $1 \leq i, j \leq m$ und $p \in U$ vorgegeben, sowie $\varepsilon > 0$ mit $\mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_{2\varepsilon}(p) \subseteq U$:

- Es gilt $p + (-\varepsilon, \varepsilon) \cdot e_j + (-\varepsilon, \varepsilon) \cdot e_i \subseteq \mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_\varepsilon(p) \subseteq U$. Dann ist

$$\phi\colon (-\varepsilon, \varepsilon) \times (-\varepsilon, \varepsilon) \to \mathbb{R}^n, \qquad (s, t) \mapsto \partial_j f(p + s \cdot e_j + t \cdot e_i)$$

definiert und stetig,[116] mit stetiger Richtungsableitung[117]

$$\begin{aligned}
D_{(0,1)}\phi\colon (-\varepsilon, \varepsilon) \times (-\varepsilon, \varepsilon) &\to \mathbb{R}^n \\
(s, t) &\mapsto \lim_{h\to 0} \frac{\phi(s, t + h \cdot 1) - \phi(s, t)}{h} \\
&= \lim_{h\to 0} \frac{\partial_j f(p + s \cdot e_j + (t+h) \cdot e_i) - \partial_j f(p + s \cdot e_j + t \cdot e_i)}{h} \\
&= \partial_i(\partial_j f)(p + s \cdot e_j + t \cdot e_i).
\end{aligned} \tag{452}$$

- Es gilt $p + (-\varepsilon, \varepsilon) \cdot e_j + \mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_\varepsilon(0) \subseteq \mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_{2\varepsilon}(p) \subseteq U$. Für jedes $\mu \in [0, \varepsilon)$ ist dann

$$\eta_\mu\colon \mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_\varepsilon(0) \to \mathbb{R}^n, \qquad y \mapsto \int_0^\mu \partial_j f(p + s \cdot e_j + y)\,\mathrm{d}s$$

definiert und stetig, denn wegen Lemma 82 gilt:

$$\eta_\mu(y) = f(p + \mu \cdot e_j + y) - f(p + y) \qquad \forall\, y \in \mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_\varepsilon(0). \tag{453}$$

Zudem erhalten wir für $h \neq 0$ mit $t + h \in (-\varepsilon, \varepsilon)$:

$$\begin{aligned}
&\frac{\int_0^\mu \phi(s, t+h)\,\mathrm{d}s - \int_0^\mu \phi(s, t)\,\mathrm{d}s}{h} \\
&\qquad = \frac{\int_0^\mu \partial_j f(p + s \cdot e_j + (t+h) \cdot e_i)\,\mathrm{d}s - \int_0^\mu \partial_j f(p + s \cdot e_j + t \cdot e_i)\,\mathrm{d}s}{h} \\
&\qquad = \frac{\eta_\mu(t \cdot e_i + h \cdot e_i) - \eta_\mu(t \cdot e_i)}{h}.
\end{aligned} \tag{454}$$

[116] Per Annahme ist $\partial_j f\colon U \to \mathbb{R}^n$ stetig. Weiterhin ist $\Theta\colon (-\varepsilon, \varepsilon)^2 \ni (s, t) \mapsto p + s \cdot e_j + t \cdot e_i \in \mathbb{R}^n$ stetig bezüglich der jeweiligen Maximumsmetriken (Maximumsnormen) auf $\mathbb{R} \cong \mathbb{R}^2$ und $\mathbb{R}^n$; denn es gilt ja $\|\Theta(s', t') - \Theta(s, t)\|_\infty = \max(|s - s'|, |t - t'|)$ für alle s, s', t, t'.

[117] Wir halten uns an dieser Stelle strikt an die Notation aus Satz 91 mit $m \equiv 1$ und $k \equiv n$; d.h., $I, U \equiv (-\varepsilon, \varepsilon)$ und $v = 1 \in \mathbb{R} = \mathbb{R}^m$.

Für $\mu \in [0,\varepsilon)$ und $t \in (-\varepsilon,\varepsilon)$ liefert Satz 91:

$$\int_0^\mu D_{(0,1)}\phi(s,t)\,\mathrm{d}s \overset{\text{Satz 91}}{=} \underbrace{\lim_{h\to 0} \frac{\int_0^\mu \phi(s,t+h)\,\mathrm{d}s - \int_0^\mu \phi(s,t)\,\mathrm{d}s}{h}}_{D_1\left(\int_0^\mu \phi(s,\bullet)\,\mathrm{d}s\right)(t)} \overset{(454)}{=} \partial_i \eta_\mu(t\cdot e_i). \tag{455}$$

Wir erhalten für $\mu \in [0,\varepsilon)$ mit der Lokalität der Richtungsableitungen (Übung 152) im ersten Schritt:[118]

$$\underbrace{\underbrace{\partial_i f(p+\mu\cdot e_j+\bullet)(0)}_{\partial_i f(p+\mu\cdot e_j)}}_{=:\,\alpha(\mu)} \overset{(453)}{=} \partial_i(\eta_\mu(\bullet)+f(p+\bullet))(0) = \partial_i\eta_\mu(0)+\partial_i f(p) \overset{(455)}{=} \underbrace{\int_0^\mu D_{(0,1)}\phi(s,0)\,\mathrm{d}s \,+\, \partial_i f(p)}_{=:\,\beta(\mu)}.$$

Ableiten nach μ an der Stelle $\mu = 0$, liefert gemäß dem Hauptsatz

$$\partial_j(\partial_i f)(p) = \lim_{\mu\downarrow 0}\frac{\alpha(\mu)-\alpha(0)}{\mu} = \lim_{\mu\downarrow 0}\frac{\beta(\mu)-\beta(0)}{\mu} = D_{(0,1)}\phi(0,0) \overset{(452)}{=} \partial_i(\partial_j f)(p). \qquad \square$$

Korollar 60. *Sei $\emptyset \neq U \subseteq \mathbb{R}^m$ offen sowie $f \in C^k(U,\mathbb{R}^n)$ mit $m,n \in \mathbb{N}_{>0}$ und $k \geq 2$. Seien $1 \leq j_1,\dots,j_p \leq m$ mit $2 \leq p \leq k$, und $\sigma\colon \{1,\dots,p\} \to \{1,\dots,p\}$ eine Bijektion. Dann gilt*

$$\partial_{j_1}\dots\partial_{j_p} f = \partial_{j_{\sigma(1)}}\dots\partial_{j_{\sigma(p)}} f.$$

Beweis. Dies folgt induktiv aus Satz 92 (Übung). $\square$

Beispiel 105 (Gradientenfeld)**.** *Sei $\mathcal{V} \in C^1(\mathbb{R}^2,\mathbb{R}^2)$ (ein stetig differenzierbares Vektorfeld). Gesucht ist ein $\phi\colon C^2(\mathbb{R}^2,\mathbb{R})$ mit $J_\bullet(\phi) = \nabla\phi = \mathcal{V}$ (d.h. $(\partial_1\phi,\partial_2\phi) = (\mathcal{V}_1,\mathcal{V}_2)$).*

- *Existiert eine derartige „Potentialfunktion“ ϕ, so erfüllt $\mathcal{V}$ nach dem Satz von Schwarz (Satz 92) notwendigerweise die Bedingung:*

$$\partial_2\mathcal{V}_1 = \partial_2\partial_1\phi = \partial_1\partial_2\phi = \partial_1\mathcal{V}_2. \tag{456}$$

  Beispiel: *Für $\mathcal{V}\colon \mathbb{R}^2 \ni (x,y) \mapsto (-y,x) \in \mathbb{R}^2$ ist diese Bedingung wegen $\partial_2\mathcal{V}_1 = -1 \neq 1 = \partial_1\mathcal{V}_2$ nicht erfüllt, sodass $\mathcal{V} = \nabla\phi$ für* **kein** *$\phi \in C^2(\mathbb{R}^2,\mathbb{R})$ gelten kann.*

- *Sei $\mathcal{V} \in C^1(\mathbb{R}^2,\mathbb{R}^2)$ vorgegeben mit $\partial_2\mathcal{V}_1 = \partial_1\mathcal{V}_2$, und setze*

$$\phi\colon \mathbb{R}^2 \to \mathbb{R}, \qquad (x,y) \mapsto \textstyle\int_0^y \mathcal{V}_2(0,t)\,\mathrm{d}t + \int_0^x \mathcal{V}_1(s,y)\,\mathrm{d}s.$$

Dann gilt $\phi \in C^1(\mathbb{R}^2,\mathbb{R})$ mit $\nabla\phi = \mathcal{V}$, also $\phi \in C^2(\mathbb{R}^2,\mathbb{R})$ (Übung 162).

Übung 162. *Zeigen Sie die Behauptung im zweiten Punkt von Beispiel 105.*

Übung 163. *Die Abbildung*

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \qquad (x,y) \mapsto \begin{cases} \dfrac{xy\cdot(x^2-y^2)}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

ist zweimal partiell differenzierbar, aber keine C^2-Abbildung. Berechnen Sie deren erste partielle Ableitungen $\partial_1 f$ und $\partial_2 f$, und zeigen Sie $\partial_1\partial_2 f(0,0) \neq \partial_2\partial_1 f(0,0)$.

[118] Die Ausdrücke auf der linken und der rechten Seite sind für sich genommen natürlich sogar für $\mu \in (-\varepsilon,\varepsilon)$ definiert.

15.2 Polynomfunktionen und Allgemeine Rechenregeln

Terminologie 43.

a) *Ein* Multiindex *(vom Grad $\ell \in \mathbb{N}_{>0}$) ist ein n-Tupel $\alpha = (\alpha_1, \ldots, \alpha_\ell) \in \mathbb{N}^\ell$.*

Für $\alpha = (\alpha_1, \ldots, \alpha_\ell) \in \mathbb{N}^\ell$ und $\beta = (\beta_1, \ldots, \beta_\ell) \in \mathbb{N}^\ell$, heißt

- $|\alpha| := \alpha_1 + \ldots + \alpha_\ell$ *die* Ordnung *von α.*
- $\alpha! := \alpha_1! \cdot \ldots \cdot \alpha_\ell!$ *die* Fakultät *von α (α-Fakultät)*
- $\binom{\beta}{\alpha} := \prod_{j=1}^{\ell} \binom{\beta_j}{\alpha_j}$ *der zugehörige* Binomialkoeffizient.

Wir setzen:

- $\beta \leq \alpha \quad \Leftrightarrow \quad \beta_j \leq \alpha_j$ *für alle* $1 \leq j \leq \ell$. *(Ordnung auf $\mathbb{N}^\ell$)*
- $\alpha \pm \beta := (\alpha_1 \pm \beta_1, \ldots, \alpha_\ell \pm \beta_\ell)$, *wobei* $\alpha - \beta$ *nur für* $\beta \leq \alpha$ *definiert wird.*

Für $q \in \mathbb{N}$ setzen wir $\mathcal{B}(q) := \{\alpha \in \mathbb{N}^\ell \mid |\alpha| = q\}$.

b) *Sei $m, n \in \mathbb{N}_{>0}$. Für $x \in \mathbb{R}^m$ und $\alpha \in \mathbb{N}^m$, setzen wir (vgl. Beispiel 97)*

$$x^\alpha := x_1^{\alpha_1} \cdot \ldots \cdot x_m^{\alpha_m} \in \mathbb{R}.$$

- *Für $\alpha \in \mathbb{N}^m$, $p \in \mathbb{R}^m$ und $u \in \mathbb{R}^n$, heißt die stetige Funktion (vgl. Beispiel 97)*[119]

$$M_{\alpha,p,u} := \mathrm{pr}_1^{\alpha_1} \cdot \ldots \cdot \mathrm{pr}_m^{\alpha_m} \colon \mathbb{R}^m \to \mathbb{R}^n, \qquad x \mapsto u \cdot (x-p)^\alpha$$

  Monom vom Exponenten α in p *bzw.* Monom vom Grad $|\alpha|$ in p.

- *Sei $p \in \mathbb{R}^m$ und $\{u_\alpha \mid \alpha \in \mathbb{N}^m\} \subseteq \mathbb{R}^n$, sodass $\{\alpha \in \mathbb{N}^m \mid u_\alpha \neq 0\}$ endlich ist. Dann heißt (beachte die Endlichkeit der Summe)*

$$P = \textstyle\sum_{\alpha \in \mathbb{N}^m} M_{\alpha,p,u_\alpha} \colon \mathbb{R}^m \to \mathbb{R}^n, \qquad x \mapsto \textstyle\sum_{\alpha \in \mathbb{N}^m} u_\alpha \cdot (x-p)^\alpha$$

  Polynomfunktion vom Grad $\mathrm{grad}(\mathsf{P}) := \max\big(\{0\} \cup \{|\alpha| \mid \alpha \in \mathbb{N}^m \colon u_\alpha \neq 0\}\big)$ in p.

  Beachte: *Die Komponentenfunktionen $P_1, \ldots, P_n$ von P sind ebenfalls Polynomfunktionen:*

$$P_j \colon \mathbb{R}^m \to \mathbb{R}, \qquad x \mapsto \sum_{\alpha \in \mathbb{N}^m} M_{\alpha,p,(u_\alpha)_j} = \sum_{\alpha \in \mathbb{N}^m} (u_\alpha)_j \cdot (x-p)^\alpha \qquad \forall\, 1 \leq j \leq n$$

  mit $u_\alpha = ((u_\alpha)_1, \ldots, (u_\alpha)_n)$ für alle $\alpha \in \mathbb{N}^m$, also $\mathrm{grad}(P_j) \leq \mathrm{grad}(P)$ für $1 \leq j \leq n$ (Übung).

c) *Sei $U \subseteq \mathbb{R}^m$ offen, $k \in \mathbb{N}_{>0}$, und $f \colon U \to \mathbb{R}^n$ eine k-mal partiell differenzierbare Abbildung.*

- *Die k-fache partielle Ableitung von f in Richtung e_j mit $1 \leq j \leq m$ bezeichnen wir mit*

$$\partial_j^k f \equiv D_{e_j}^k f \equiv \frac{\partial^k f}{\partial_{x_j}^k},$$

  also $\partial_j^k f := \partial_{j_1} \ldots \partial_{j_k} f$ mit $j_1, \ldots, j_k = j$. Wir setzen $\partial_j^0 f := f$.

- *Für $\alpha \in \mathbb{N}^m$ definieren wir*

$$\partial^\alpha f := \partial_1^{\alpha_1} \ldots \partial_m^{\alpha_m} f =: \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots \partial x_m^{\alpha_m}} f.$$

  Beachte:

[119] Die korrekte Notation für das Produkt von $\lambda \in \mathbb{R}$ und $u \in \mathbb{R}^n$ ist normalerweise $\lambda \cdot u$. Wir schreiben im Folgenden aber dennoch eher $u \cdot \lambda$, damit sich die Notation mit unserer Notation für reellwertige Polynome deckt.

- *Wegen Lemma 84.1) gilt*

$$\partial^\alpha f = (\partial^\alpha f_1, \ldots, \partial^\alpha f_n) \qquad \forall\, \alpha \in \mathbb{N}^n \quad \textit{mit} \quad |\alpha| \leq k. \tag{457}$$

- *Ist $f \in C^k(U, \mathbb{R}^n)$, so lässt sich wegen Korollar 60 jede partielle Ableitung der Ordnung $\leq k$ als $\partial^\alpha f$ für ein $\alpha \in \mathbb{N}^m$ mit $|\alpha| \leq k$ schreiben (Übung).*

15.2.1 Einige Rechenregeln

Lemma 86. *Seien $k, m, n \in \mathbb{N}_{>0}$, $U \subseteq \mathbb{R}^m$ offen, und $f\colon \mathbb{R}^m \supseteq U \to \mathbb{R}^n$ eine k-mal partiell differenzierbare Abbildung. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $p + I \cdot e_j \subseteq U$. Dann ist*

$$\gamma\colon I \to \mathbb{R}^n, \qquad t \mapsto f(p + t \cdot e_j)$$

k-mal differenzierbar, mit $\gamma^{(\ell)}\colon I \ni t \mapsto \partial_j^\ell f(p + t \cdot e_j) \in \mathbb{R}$ für alle $1 \leq \ell \leq k$.

- *Insbesondere gilt $\gamma \in C^k(I, \mathbb{R}^n)$, wenn $f \in C^k(U, \mathbb{R}^n)$ gilt.*
- *Wegen Lemma 84 und unserer Komponentenweisen Definition der Ableitung einer Kurve, gilt:*

$$\gamma_i^{(\ell)}(t) = \partial_j^\ell f_i(p + t \cdot e_j) \qquad \forall\, t \in I,\ 0 \leq \ell \leq k,\ 1 \leq i \leq n.$$

Beweis. (IA) Die Behauptung ist klar für $\ell = 0$. (IV) Es gelte also die Behauptung für ein $0 \leq \ell < k$. (IS) Per Voraussetzung ist $\tilde{f} := \partial_j^\ell f$ partiell differenzierbar mit $\partial_j \tilde{f}$, und wir setzen $\tilde{\gamma}\colon I \ni t \mapsto \tilde{f}(p + t \cdot e_j) \in \mathbb{R}$, d.h. $\tilde{\gamma} = \gamma^{(\ell)}$ per (IV). Bemerkung 97 angewandt auf $\tilde{f}$ und $\tilde{\gamma}$ zeigt (dritter Schritt)

$$\partial_j^{\ell+1} f(p + t \cdot e_j) = \partial_j\big(\partial_j^\ell f\big)(p + t \cdot e_j) = \partial_j \tilde{f}(p + t \cdot e_j) = \tilde{\gamma}'(t) = \big(\gamma^{(\ell)}\big)'(t) = \gamma^{(\ell+1)}(t)$$

für alle $t \in I$, also folgt die Behauptung per Induktion. $\square$

Lemma 87 (Allgemeine Leibniz-Regel)**.** *Sei $U \subseteq \mathbb{R}^m$ offen, sowie $f\colon U \to \mathbb{R}$ und $g\colon U \to \mathbb{R}^n$ zwei k-mal partiell differenzierbare Funktion für ein $k \in \mathbb{N}_{>0}$. Dann gilt die allgemeine Leibniz-Formel:*

$$\partial^\alpha (f \cdot g) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^\beta f \cdot \partial^{\alpha - \beta} g \qquad \forall\, \alpha \in \mathbb{N}^m \quad \textit{mit} \quad |\alpha| \leq k. \tag{458}$$

Beweis. • Sei zunächst $1 \leq j \leq m$ und $0 \leq \ell \leq k$. Für $p \in U$ vorgegeben, wählen wir ein offenes Intervall $I \subseteq \mathbb{R}$ mit $0 \in I$ und $p + I \cdot e_j \subseteq U$, und betrachten die Kurven:

$$\delta_1\colon I \ni t \mapsto f(p + t \cdot e_j) \in \mathbb{R}, \quad \delta_2\colon I \ni t \mapsto g(p + t \cdot e_j) \in \mathbb{R}^n, \quad \gamma\colon I \ni t \mapsto (f \cdot g)(p + t \cdot e_j) \in \mathbb{R}^n,$$

also $\gamma = \delta_1 \cdot \delta_2$. Wir erhalten mit Lemma 86 im ersten und im letzten Schritt, sowie der Leibniz-Formel (388) für Kurven im zweiten Schritt:

$$\partial_j^\ell (f \cdot g)(p) = \gamma^{(\ell)}(0) = (\delta_1 \cdot \delta_2)^{(\ell)}(0) = \sum_{q=0}^{\ell} \binom{\ell}{q} \delta_1^{(q)}(0) \cdot \delta_2^{(\ell - q)}(0) = \sum_{q=0}^{\ell} \binom{\ell}{q} \partial_j^q f(p) \cdot \partial_j^{\ell - q} g(p).$$

• Durch iteratives Anwenden des vorhergehenden Punktes sowie Übung 158 , erhalten wir

$$\begin{aligned}
\partial^\alpha (f \cdot g) &= \partial_1^{\alpha_1}(\partial_2^{\alpha_2}(\ldots \partial_m^{\alpha_m}(f \cdot g))) = \partial_1^{\alpha_1} \ldots \partial_{m-1}^{\alpha_{m-1}} \textstyle\sum_{\beta_m = 0}^{\alpha_m} \binom{\alpha_m}{\beta_m} \partial_m^{\beta_m} f \cdot \partial_m^{\alpha_m - \beta_m} g \\
&= \partial_1^{\alpha_1} \ldots \partial_{m-2}^{\alpha_{m-2}} \textstyle\sum_{\beta_{m-1} = 0}^{\alpha_{m-1}} \sum_{\beta_m = 0}^{\alpha_m} \binom{\alpha_{m-1}}{\beta_{m-1}} \cdot \binom{\alpha_m}{\beta_m} \partial_{m-1}^{\beta_{m-1}} \partial_m^{\beta_m} f \cdot \partial_{m-1}^{\alpha_{m-1} - \beta_{m-1}} \partial_m^{\alpha_m - \beta_m} g \\
&\vdots \\
&= \textstyle\sum_{\beta_1 = 0}^{\alpha_1} \cdots \sum_{\beta_m = 0}^{\alpha_m} \binom{\alpha_1}{\beta_1} \cdot \ldots \cdot \binom{\alpha_m}{\beta_m} \partial_1^{\beta_1} \ldots \partial_m^{\beta_m} f \cdot \partial_1^{\alpha_1 - \beta_1} \ldots \partial_m^{\alpha_m - \beta_m} g \\
&= \textstyle\sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^\beta f \cdot \partial^{\alpha - \beta} g.
\end{aligned}$$

$\square$

Lemma 88 (Verkettungen). *Seien $m,n,\ell \in \mathbb{N}_{>0}$, $k \in \mathbb{N}$, $U \subseteq \mathbb{R}^m$ und $V \subseteq \mathbb{R}^n$ offen, sowie $f \in C^k(V,\mathbb{R}^\ell)$ und $g \in C^k(U,\mathbb{R}^n)$ mit $g(U) \subseteq V$. Dann gilt $(f \circ g) \in C^k(U,\mathbb{R}^\ell)$.*

Beweis. Offensichtlich gilt die Behauptung für $k = 0$. Wir können daher annehmen, dass die Behauptung für ein $k \in \mathbb{N}$ gilt. Seien nun f, g beides C^{k+1}-Abbildungen. Für $1 \leq j \leq m$ erhalten wir aus der Kettenregel (434) (beachte, dass f und g wegen Satz 87 beide differenzierbar sind)

$$\partial_j(f \circ g) \overset{(434)}{=} \sum_{i=1}^n (\partial_i f \circ g) \cdot \partial_j g_i,$$

wobei $\partial_i f \circ g \in C^k(U,\mathbb{R}^\ell)$ per Induktionsvoraussetzung gilt, sowie $\partial_j g_i \in C^k(U,\mathbb{R})$ für $1 \leq i \leq n$ wegen Lemma 84.1). Wegen Übung 158 und Übung 159 gilt somit $\partial_j(f \circ g) \in C^k(U,\mathbb{R}^\ell)$ für alle $1 \leq j \leq m$, also $(f \circ g) \in C^{k+1}(U,\mathbb{R}^\ell)$. □

Im nächsten Abschnitt benötigen wir die folgende Rechenregel:

Lemma 89. *Sei $U \subseteq \mathbb{R}^m$ offen, $k \in \mathbb{N}_{>0}$ und $f \in C^k(U,\mathbb{R}^n)$. Seien $p \in U$, $h \in \mathbb{R}^m$ und $\epsilon > 0$ vorgegeben mit $p + (-\epsilon, 1+\epsilon) \cdot h \subseteq U$. Dann ist*

$$g\colon I := (-\epsilon, 1+\epsilon) \to \mathbb{R}^n, \qquad t \mapsto f(p + t \cdot h)$$

k-mal stetig partiell differenzierbar, mit

$$g^{(\nu)}(t) = \sum_{|\alpha|=\nu} \frac{\nu!}{\alpha!} \cdot \partial^\alpha f(p + t \cdot h) \cdot h^\alpha \qquad \forall\, t \in I,\ 0 \leq \nu \leq k. \tag{459}$$

Beweis. Es genügt die Differenzierbarkeitsaussage zu zeigen; denn die Stetigkeitsaussage folgt dann umgehend aus der C^k-Eigenschaft von f, sowie der expliziten Formel (459). Offensichtlich gilt die Behauptung für $\nu = 0$. Also können wir annehmen, dass (459) für ein $0 \leq \nu \leq k-1$ mit $k \geq 1$ gilt:

Für $j = 1, \dots, m$ und $\alpha \in \mathcal{B}(\nu)$, sei $\alpha[j] := \alpha + e_j$. Für jedes $\alpha \in \mathcal{B}(\nu)$ ist $\partial^\alpha f$ stetig partiell differenzierbar, also (stetig) differenzierbar wegen Satz 87. Wir erhalten somit aus der Induktionsvoraussetzung (459), der Linearitätsaussage in Übung 158 sowie der Kettenregel (434)[120] (zweiter Schritt):

$$\begin{aligned}
g^{(\nu+1)}(t) &= \big(g^{(\nu)}\big)'(t) \\
&= \sum_{|\alpha|=\nu} \frac{\nu!}{\alpha!} \cdot h^\alpha \cdot \sum_{j=1}^m \partial_j \partial^\alpha f(p + t \cdot h) \cdot h_j \\
&= \sum_{|\alpha|=\nu} \sum_{j=1}^m \frac{\nu!}{\alpha!} \cdot \partial^{\alpha[j]} f(p + t \cdot h) \cdot h^{\alpha[j]} \\
&= \sum_{|\alpha|=\nu} \sum_{j=1}^m \binom{\overbrace{\alpha_j + 1}^{\alpha[j]_j}}{\nu + 1} \cdot \frac{(\nu+1)!}{\alpha[j]!} \cdot \partial^{\alpha[j]} f(p + t \cdot h) \cdot h^{\alpha[j]} \\
&\overset{(*)}{=} \sum_{|\beta|=\nu+1} \frac{\nu!}{\beta!} \cdot \partial^\beta f(p + t \cdot h) \cdot h^\beta,
\end{aligned} \tag{460}$$

wobei $(*)$ durch Umformung folgt: Für $\beta \in \mathbb{N}^m$ mit $|\beta| = \nu + 1$, sei $(\beta \in \mathcal{B}(\nu+1))$

$$\begin{aligned}
\Lambda(\beta) &:= \{(\alpha, j) \in \mathcal{B}(\nu) \times \{1, \dots, m\} \mid \alpha[j] = \beta\} \\
&= \{(\beta - e_j, j) \mid j \in \{1, \dots, m\} : \beta_j \geq 1\}.
\end{aligned} \tag{461}$$

[120]Genauer (434) angewandt auf $\tilde{f} := \partial^\alpha f$ und $\tilde{g}\colon (-\epsilon, 1+\epsilon) \ni t \mapsto p + t \cdot h \in \mathbb{R}^m$.

Die vorletzten Zeile von (460) schreibt sich dann in der Form:

$$g^{(\nu+1)}(h) = \sum_{|\beta|=\nu+1} \underbrace{\sum_{(\alpha,j)\in\Lambda(\beta)} \left(\frac{\overbrace{\alpha[j]_j}^{\beta_j}}{\nu+1}\right)}_{=:\,\eta(\beta)} \cdot \underbrace{\frac{(\nu+1)!}{\alpha[j]} \cdot \partial^{\alpha[j]} f(p+t\cdot h)\cdot h^{\alpha[j]}}_{\frac{(\nu+1)!}{\beta!}\cdot \partial^{\beta} f(p+t\cdot h)\cdot h^{\beta}}.$$

Für $\beta \in \mathcal{B}(\nu+1)$ gilt nun aber $\eta(\beta)=1$ wegen:

$$(\nu+1)\cdot\eta(\beta) = \textstyle\sum_{(\alpha,j)\in\Lambda(\beta)} \beta_j \overset{(461)}{=} \sum_{j\in\{1,\dots,m\}\,:\,\beta_j\geq 1} \beta_j = |\beta| = (\nu+1). \qquad \square$$

15.2.2 Polynomfunktionen

Beispiel 106 (Partielle Ableitungen von Monomen)**.** *Sei $\beta\in\mathbb{N}^m$, $p\in\mathbb{R}^m$, $u\in\mathbb{R}^n$, sowie*

$$M_{\beta,p,u}\colon \mathbb{R}^m\to\mathbb{R}^n, \qquad x\mapsto u\cdot(x-p)^\beta = u\cdot(x_1-p_1)^{\beta_1}\cdot\ldots\cdot(x_m-p_m)^{\beta_m}.$$

Dann gilt für $\alpha\in\mathbb{N}^m$:

$$\partial^\alpha M_{\beta,p,u}(x) = \begin{cases} u\cdot\dfrac{\beta!}{(\beta-\alpha)!}\cdot(x-p)^{\beta-\alpha} & \textit{für}\quad \alpha\leq\beta \\ 0 & \textit{sonst.}\end{cases}$$

Um dies einzusehen, wendet man iterativ die Produktregel zusammen mit

$$\begin{aligned} &\partial_j \mathrm{pr}_i(\cdot-p)^q = 0 && \textit{für} \quad 1\leq i\neq j\leq m \\ &\partial_j \mathrm{pr}_j(\cdot-p)^q = q\cdot \mathrm{pr}_j(\cdot-p)^{q-1} && \textit{für} \quad 1\leq j\leq m \end{aligned}$$

für $q\in\mathbb{N}$ an (erster Schritt):

$$\begin{aligned}\partial^\alpha M_{\beta,p,u} &= u\cdot\partial_1^{\alpha_1}\ldots\partial_m^{\alpha_m}\big(\mathrm{pr}_1(\cdot-p)^{\beta_1}\cdot\ldots\cdot\mathrm{pr}_m(\cdot-p)^{\beta_m}\big)\\ &= u\cdot\big(\partial_1^{\alpha_1}\mathrm{pr}_1(\cdot-p)^{\beta_1}\big)\cdot\ldots\cdot\big(\partial_m^{\alpha_m}\mathrm{pr}_m(\cdot-p)^{\beta_m}\big)\\ &= \begin{cases} u\cdot\dfrac{\beta_1!}{(\beta_1-\alpha_1)!}\cdot\mathrm{pr}_1(\cdot-p)^{\beta_1-\alpha_1}\cdot\ldots\cdot\dfrac{\beta_m!}{(\beta_m-\alpha_m)!}\cdot\mathrm{pr}_m(\cdot-p)^{\beta_m-\alpha_m} & \textit{für}\quad\alpha\leq\beta\\ 0 & \textit{sonst.}\end{cases}\end{aligned}$$

An der Stelle $x=p$ gilt dann insbesondere

$$\big(\partial^\alpha M_{\beta,p,u}\big)(p) = \begin{cases} 0 & \textit{für}\quad \alpha\neq\beta\\ u\cdot\alpha! & \textit{für}\quad\alpha=\beta.\end{cases} \tag{462}$$

Bemerkung 100. *Gegeben seien zwei Polynomfunktionen*

$$\begin{aligned} &P\colon\mathbb{R}^m\to\mathbb{R}^n, && x\mapsto \textstyle\sum_{\alpha\in\mathbb{N}^m} u_\alpha\cdot(x-p)^\alpha\\ &Q\colon\mathbb{R}^m\to\mathbb{R}^{\tilde n}, && x\mapsto \textstyle\sum_{\alpha\in\mathbb{N}^m} v_\alpha\cdot(x-p)^\alpha\end{aligned}$$

im Punkt $p\in\mathbb{R}^m$ mit $n,\tilde n\in\mathbb{N}_{>0}$ (d.h., $u_\alpha, v_\alpha\in\mathbb{R}^n$ für alle $\alpha\in\mathbb{N}^m$, sodass $\{\alpha\in\mathbb{N}^m\mid u_\alpha\neq 0\}$ und $\{\alpha\in\mathbb{N}^m\mid v_\alpha\neq 0\}$ endlich sind). Wir bemerken Folgendes:

- *Beispiel 106 zeigt*

$$\partial^\alpha P(p) \overset{(462)}{=} u_\alpha \cdot \alpha! \qquad \forall\, \alpha \in \mathbb{N}^n.$$

Somit sind alle auftretenden Koeffizienten u_α für $\alpha \in \mathbb{N}^n$, also die explizite Darstellung der Poynomfunktion P (und somit auch P als Funktion selbst), bereits eindeutig bestimmt durch die partiellen Ableitungen $\{\partial^\alpha P(p) \mid |\alpha| \leq \mathrm{grad}(P)\}$.

- *P lässt sich auch als Polynomfunktion in jedem anderen Entwicklungspunkt $q \in \mathbb{R}^m$ auffassen, und zwar mit dem gleichen Grad $\mathrm{grad}(P)$. Hierfür schreibt man*

$$P(x) = \textstyle\sum_{\alpha\in\mathbb{N}^m} u_\alpha \cdot ((x-q)+(q-p))^\alpha \qquad \forall\, x \in \mathbb{R}^n,$$

und wendet dann, analog zum eindimensionalen Fall, den binomischen Lehrsatz an. Der vorherigen Punkt zeigt dann, dass P auch durch die partiellen Ableitungen $\{\partial^\alpha P(q) \mid |\alpha| \leq \mathrm{grad}(P)\}$ in q bereits eindeutig bestimmt sind.

- *Für $n = \tilde{n}$ und $\lambda, \mu \in \mathbb{R}$ ist*

$$\lambda \cdot P + \mu \cdot Q \colon x \mapsto \textstyle\sum_{\alpha\in\mathbb{N}^m} (\lambda \cdot u_\alpha + \mu \cdot v_\alpha) \cdot (x-p)^\alpha$$

ebenfalls eine Polynomfunktion in p, mit $\mathrm{grad}(\lambda \cdot P + \mu \cdot Q) \leq \max(\mathrm{grad}(P), \mathrm{grad}(Q))$.

- *Für $\tilde{n} = 1$ ist $Q \cdot P \colon \mathbb{R}^m \to \mathbb{R}^n$ eine Polynomfunktion in p, mit $\mathrm{grad}(Q \cdot P) \leq \mathrm{grad}(Q) \cdot \mathrm{grad}(P)$.*

15.3 Taylorentwicklung in Mehreren Veränderlichen

Terminologie 44. *Sei $m, n, k \in \mathbb{N}_{>0}$, $U \subseteq \mathbb{R}^m$ offen, und $f \colon U \to \mathbb{R}^n$ eine k-mal partiell differenzierbare Abbildung:*

- *Das* k-te Taylorpolynom *(Taylorpolynom der Ordnung k) von f in $p \in U$ ist die Polynomfunktion:*

$$\begin{aligned}
\mathrm{T}_p^k(f)(x) &:= \sum_{|\alpha|\leq k} \frac{\partial^\alpha f(p)}{\alpha!} \cdot (x-p)^\alpha \\
&= \sum_{\nu=0}^{k} \sum_{\alpha\in\mathcal{B}(\nu)} \frac{\partial^\alpha f(p)}{\alpha!} \cdot (x-p)^\alpha \\
&= \underbrace{f(p)}_{\nu=0} + \underbrace{\overbrace{\textstyle\sum_{j=1}^m \partial_j f(p) \cdot (x_j - p_j)}^{\mathrm{d}_p f(x-p) \text{ falls } f \text{ differenzierbar, bspw. } C^1}}_{\nu=1} \\
&\quad + \underbrace{\tfrac{1}{2} \cdot \textstyle\sum_{j=1}^m \partial_j^2 f(p) \cdot (x_j-p_j)^2 + \sum_{1\leq i<j\leq n} \partial_i\partial_j f(p) \cdot (x_i - p_i) \cdot (x_j - p_j)}_{\nu=2} \\
&\quad + \sum_{\nu=3}^{k} \sum_{\alpha\in\mathcal{B}(\nu)} \frac{\partial^\alpha f(p)}{\alpha!} \cdot (x-p)^\alpha.
\end{aligned} \tag{463}$$

Beachte: *Ein Multiindex der Ordnung 1 hat die Gestalt $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ (eine 1 an der j-ten Stelle) für ein $1 \leq j \leq n$; und jeder Multiindex der Ordnung 2 hat die Gestalt $2e_j = (0, \ldots, 0, 2, 0, \ldots, 0)$ für ein $1 \leq j \leq n$ oder $e_i + e_j = (0, \ldots, 0, 1, 0, \ldots, 0, 1, 0, \ldots, 0)$ für ein $1 \leq i \neq j \leq n$.*

- *Das* k-te Restglied *von f in p ist definiert durch:*

$$\mathrm{R}_p^k(f) \colon U \to \mathbb{R}, \qquad x \mapsto f(x) - \mathrm{T}_p^k(f)(x). \tag{464}$$

Beachte: *Für jedes auftretende α ist $\partial^\alpha f(p) \in \mathbb{R}^n$ ein Vektor und $(x-p)^\alpha \in \mathbb{R}$ eine Zahl. Wegen (457) kann man die obigen Definitionen dann auch wieder komponentenweise verstehen:*

Für $f = (f_1, \ldots, f_n)$ und $1 \le j \le n$, erhalten wir aus (457):

$$\begin{aligned} \overbrace{\mathrm{pr}_j\big(\mathrm{T}^k_p(f)\big)}^{\mathrm{T}^k_p(f)_j} &= \mathrm{pr}_j\left(\sum_{|\alpha|\le k} \frac{\partial^\alpha f(p)}{\alpha!}\cdot(\cdot-p)^\alpha\right) = \sum_{|\alpha|\le k}\frac{\partial^\alpha f_j(\cdot-p)}{\alpha!} = \mathrm{T}^k_p(f_j) \\ \Longrightarrow \qquad \overbrace{\mathrm{pr}_j\big(\mathrm{R}^k_p(f)\big)}^{\mathrm{R}^k_p(f)_j} &= \mathrm{pr}_j\big(f - \mathrm{T}^k_p(f)\big) = f_j - \mathrm{T}^k_p(f_j) = \mathrm{R}^k_p(f_j). \end{aligned} \tag{465}$$

Das k-te Taylorpolynom von f ist also komponentenweise gegeben durch die k-ten Taylorpolynome der Komponentenfunktionen von f. Weiterhin sei angemerkt (Übung):

- $\mathrm{T}^k_p(f)$ *ist das eindeutige Polynom P vom Grad $\le k$ in p, mit $\partial^\alpha P(p) = \partial^\alpha f(p)$ für alle $|\alpha| \le k$.*
- *Es gilt:*

$$\mathrm{T}^{k+\ell}_p\big(\mathrm{T}^k_p(f)\big) = \mathrm{T}^k_p(f) \quad \forall\, \ell \in \mathbb{N}, \qquad \partial^\alpha \mathrm{R}^k_p(p) = 0 \quad \forall\, |\alpha| \le k, \qquad \mathrm{T}^k_p\big(\mathrm{R}^k_p(f)\big) = 0. \tag{466}$$

Satz 93 (Taylorformel)**.** *Sei $m, n \in \mathbb{N}_{>0}$, $k \in \mathbb{N}$, $U \subseteq \mathbb{R}^m$ offen, und $f \in C^{k+1}(U, \mathbb{R}^n)$. Gegeben $p, x \in U$ mit $p + [0,1]\cdot(x-p) \subseteq U$, so gilt*

$$\begin{aligned} f(x) = {} & \underbrace{\sum_{\ell=0}^{k}\sum_{|\alpha|=\ell}\frac{\partial^\alpha f(p)}{\alpha!}\cdot(x-p)^\alpha}_{\mathrm{T}^k_p(f)(x)} \\ & + \underbrace{(k+1)\cdot\sum_{|\alpha|=k+1}(x-p)^\alpha\cdot\int_0^1(1-t)^k\cdot\frac{\partial^\alpha f(p+t\cdot(x-p))}{\alpha!}\,\mathrm{d}t}_{\mathrm{R}^k_p(f)(x)}, \end{aligned} \tag{467}$$

wobei $0 \le \xi_1, \ldots, \xi_n \le 1$ existieren mit[121]

$$\underbrace{\mathrm{R}^k_p(f)_j(x)}_{\overset{(465)}{=}\mathrm{R}^k_p(f_j)(x)} = \sum_{|\alpha|=k+1}\frac{\partial^\alpha f_j(p+\xi_j\cdot(x-p))}{\alpha!}\cdot(x-p)^\alpha \qquad \forall\, 1 \le j \le n. \tag{468}$$

Beweis. Wegen der Offenheit von U existiert ein $\epsilon > 0$ mit $p + (-\epsilon, 1+\epsilon)\cdot(x-p) \subseteq U$. Wir definieren

$$C^{k+1}(I, \mathbb{R}^n) \ni g\colon I := (-\epsilon, 1+\epsilon) \to \mathbb{R}^n, \qquad t \mapsto f(p + t\cdot(x-p))$$

wie in Lemma 89. Indem wir nun Satz 64 komponentenweise anwenden (zweiter Schritt), erhalten

[121] Für den Beweis dieser Aussage werden wir Satz 65 verwenden. Unter Verwendung des allgemeineren Satzes 66, lässt sich eigentlich sogar $0 < \xi_1, \ldots, \xi_n < 1$ zeigen.

wir mit Lemma 89 (vierter Schritt)

$$
\begin{aligned}
f_j(x) = g_j(1) &= \mathrm{T}_0^k(g_j)(1) + \mathrm{R}_0^k(g_j)(1) \\
&= \sum_{\ell=0}^{k} \frac{g_j^{(\ell)}(0)}{\ell!} \cdot (1-0)^\ell \;+\; \frac{1}{k!} \cdot \int_0^1 (1-t)^k \cdot g_j^{(k+1)}(t)\,\mathrm{d}t \\
&= \sum_{\ell=0}^{k} \frac{1}{\ell!} \cdot \sum_{|\alpha|=\ell} \frac{\ell!}{\alpha!} \cdot \partial^\alpha f_j(p) \cdot (x-p)^\alpha \\
&\quad + \frac{1}{k!} \cdot \int_0^1 (1-t)^k \cdot \sum_{|\alpha|=k+1} \frac{(k+1)!}{\alpha!} \cdot \partial^\alpha f_j(p + t\cdot(x-p)) \cdot (x-p)^\alpha \,\mathrm{d}t \\
&= \sum_{\ell=0}^{k} \sum_{|\alpha|=\ell} \frac{\partial^\alpha f_j(p)}{\alpha!} \cdot (x-p)^\alpha \\
&\quad + (k+1) \cdot \sum_{|\alpha|=k+1} (x-p)^\alpha \cdot \int_0^1 (1-t)^k \cdot \frac{\partial^\alpha f_j(p+t\cdot(x-p))}{\alpha!}\,\mathrm{d}t
\end{aligned}
$$

für $1 \leq j \leq m$; also folgt nun (467) aus (465). In selbiger Weise, liefert komponentenweise Anwendung von Satz 65 gewisse $0 \leq \xi_1, \dots, \xi_n \leq 1$ mit

$$
\begin{aligned}
f_j(x) = g_j(1) &= \mathrm{T}_0^k(g_j)(1) + \frac{(1-0)^{k+1}}{(k+1)!} \cdot g_j^{(k+1)}(\xi_j) \\
&= \underbrace{\sum_{\ell=0}^{k} \sum_{|\alpha|=\ell} \frac{\partial^\alpha f_j(p)}{\alpha!} \cdot (x-p)^\alpha}_{= \mathrm{T}_p^k(f_j)(x)} + \sum_{|\alpha|=k+1} \frac{\partial^\alpha f_j(p+\xi_j\cdot(x-p))}{\alpha!} \cdot (x-p)^\alpha
\end{aligned}
$$

für $1 \leq j \leq n$, was (468) zeigt. □

Korollar 61. *Sei $m, n \in \mathbb{N}_{>0}$, $\tilde{k} \in \mathbb{N}$, $U \subseteq \mathbb{R}^m$ offen, $p \in U$, und $f \in C^{\tilde{k}}(U, \mathbb{R}^n)$. Dann gilt*

$$
f = \mathrm{T}_p^{\tilde{k}}(f) + \mathrm{R}_p^{\tilde{k}}(f) \qquad \textit{mit} \qquad \lim_{x\to p} \frac{\|\mathrm{R}_p^{\tilde{k}}(f)(x)\|_n}{(\|x-p\|_m)^{\tilde{k}}} = 0. \tag{469}
$$

(Es notiert $\|\cdot\|_{m/n}$ fixierte Normen auf $\mathbb{R}^{m/n}$.)

$*$*Beweis analog zu Korollar 44.* Für $\tilde{k} = 0$ folgt die Aussage sofort aus der Stetigkeit von f. Die linke Seite von (469) gilt einfach per Definition des Restgliedes. Für die rechte Seite bemerken wir zunächst, dass wegen der Offenheit von U ein $\varepsilon > 0$ mit $\mathsf{B}_\varepsilon(p) \subseteq U$ existiert; womit dann $p + [0,1] \cdot (x-p) \subseteq \mathsf{B}_\varepsilon(p) \subseteq U$ für alle $x \in \mathsf{B}_\varepsilon(p)$ gilt. Somit gelten die Identitäten (467) und (468) in Satz 93 für $k \equiv \tilde{k} - 1$ und jedes $x \in \mathsf{B}_\varepsilon(p)$. Wir erhalten nun zunächst aus der Definition des $\tilde{k}$-ten und des $(\tilde{k}-1)$-ten Restgliedes:

$$
f(x) = \mathrm{T}_p^{\tilde{k}}(f)(x) + \mathrm{R}_p^{\tilde{k}}(f)(x) \;=\; \overbrace{\underbrace{\mathrm{T}_p^{\tilde{k}-1}(f)}_{f(x) - \mathrm{R}_p^{\tilde{k}-1}(f)(x)} \;+\; \sum_{|\alpha|=\tilde{k}} \frac{\partial^\alpha f(p)}{\alpha!} \cdot (x-p)^\alpha}^{\mathrm{T}_p^{\tilde{k}}(f)(x)} \;+\; \mathrm{R}_p^{\tilde{k}}(f)(x).
$$

Umstellen dieser Gleichung liefert mit (468) im zweiten Schritt:

$$\mathrm{R}_p^{\tilde{k}}(f)_j(x) = \mathrm{R}_p^{\tilde{k}-1}(f)_j(x) - \sum_{|\alpha|=\tilde{k}} \frac{\partial^\alpha f_j(p)}{\alpha!} \cdot (x-p)^\alpha = \sum_{|\alpha|=\tilde{k}} \frac{\partial^\alpha f_j(p+\xi_j \cdot (x-p)) - \partial^\alpha f_j(p)}{\alpha!} \cdot (x-p)^\alpha.$$

Sei nun $C, C' > 0$ mit $\|\cdot\|_\infty \leq C \cdot \|\cdot\|_m$ und $\|\cdot\|_n \leq C' \cdot \|\cdot\|_\infty$. Dann gilt

$$|(x-p)^\alpha| = |x_1 - p_1|^{\alpha_1} \cdot \ldots \cdot |x_m - p_m|^{\alpha_m} \leq \|x-p\|_\infty^{|\alpha|} \leq C^{|\alpha|} \cdot \|x-p\|_m^{|\alpha|} = C^{\tilde{k}} \cdot \|x-p\|_m^{\tilde{k}}$$

für $\alpha \in \mathbb{N}^m$ mit $|\alpha| = \tilde{k}$, also

$$\left|\big(\partial^\alpha f_j(p+\xi_j \cdot (x-p)) - \partial^\alpha f_j(p)\big) \cdot (x-p)^\alpha\right| \leq C^{\tilde{k}} \cdot |\partial^\alpha f_j(p+\xi_j \cdot (x-p)) - \partial^\alpha f_j(p)| \cdot \|x-p\|_m^{\tilde{k}}$$

für $1 \leq j \leq m$. Wir erhalten somit

$$\frac{\|\mathrm{R}_p^{\tilde{k}}(f)(x)\|_n}{(\|x-p\|_m)^{\tilde{k}}} \leq C' \cdot \frac{\|\mathrm{R}_p^{\tilde{k}}(f)(x)\|_\infty}{(\|x-p\|_m)^{\tilde{k}}} \leq C'C^{\tilde{k}} \cdot \sum_{|\alpha|=\tilde{k}} \max_{1\leq j\leq n} \left(\frac{|\partial^\alpha f_j(p+\xi_j \cdot (x-p)) - \partial^\alpha f_j(p)|}{\alpha!}\right) \cdot \overbrace{\frac{\|x-p\|_m^{\tilde{k}}}{\|x-p\|_m^{\tilde{k}}}}^{=1}.$$

Da die Summe endlich ist, und da alle auftretenden partiellen Ableitungen stetig sind, folgt hiermit die rechte Seite von (469). □

Bemerkung 101 (Rechenregeln für Taylorpolynome)**.** *Sei $U \subseteq \mathbb{R}^m$ offen, $p \in U$ und $k \in \mathbb{N}$. Analog zum eindimensionalen Fall gelten die folgenden Rechneregeln:*

a) *Sind $f, g\colon U \to \mathbb{R}^n$ k-mal partiell differenzierbar, $\lambda, \mu \in \mathbb{R}$ und $p \in U$, so gilt*

$$\mathrm{T}_p^k(\lambda \cdot f + \mu \cdot g) = \lambda \cdot \mathrm{T}_p^k(f) + \mu \cdot \mathrm{T}_p^k(g). \tag{470}$$

Dies folgt sofort aus Übung 158.

b) *Sind $f\colon U \to \mathbb{R}$ und $g\colon U \to \mathbb{R}^n$ k-mal partiell differenzierbar, so gilt die Produktregel:*

$$\mathrm{T}_p^k(f \cdot g) = \mathrm{T}_p^k\big(\mathrm{T}_p^k(f) \cdot \mathrm{T}_p^k(g)\big). \tag{471}$$

(∗)*Beweis.* Wir erhalten mit Teil a) (dritter Schritt)

$$\begin{aligned}
\mathrm{T}_p^k(f \cdot g) &= \mathrm{T}_p^k\Big(\big(\mathrm{T}_p^k(f) + \mathrm{R}_p^k(f)\big) \cdot \big(\mathrm{T}_p^k(g) + \mathrm{R}_p^k(g)\big)\Big) \\
&= \mathrm{T}_p^k\Big(\mathrm{T}_p^k(f) \cdot \mathrm{T}_p^k(g) + \underbrace{\mathrm{T}_p^k(f) \cdot \mathrm{R}_p^k(g) + \mathrm{R}_p^k(f) \cdot \mathrm{T}_p^k(g) + \mathrm{R}_p^k(f) \cdot \mathrm{R}_p^k(g)}_{=:\,\Lambda}\Big) \\
&= \mathrm{T}_p^k\big(\mathrm{T}_p^k(f) \cdot \mathrm{T}_p^k(g)\big) + \mathrm{T}_p^k(\Lambda).
\end{aligned}$$

Wegen $\partial^\alpha \mathrm{R}_p^k(f)(p) = 0 = \partial^\alpha \mathrm{R}_p^k(g)(p)$ für alle $|\alpha| \leq k$ (vgl. (466)), folgt $\mathrm{T}_p^k(\Lambda) = 0$ aus der allgemeinen Leibniz-Regel (458) in Lemma 87. □

c) *Sei $f \in C^k(U, \mathbb{R}^n)$ und $g \in C^k(V, \mathbb{R}^m)$ mit $V \subseteq \mathbb{R}^\ell$ und $U \subseteq \mathbb{R}^m$ offen für $m, n, \ell \in \mathbb{N}_{>0}$, $k \in \mathbb{N}$, sowie $g(V) \subseteq U$. Dann gilt für jedes $p \in U$:*

$$\mathrm{T}_p^k(f \circ g) = \mathrm{T}_p^k\big(\mathrm{T}_{g(p)}^k(f) \circ \mathrm{T}_p^k(g)\big). \tag{472}$$

(∗)*Beweis.* • Für $q \in \mathbb{N}$ sowie $\tilde{f} \in C^q(U, \mathbb{R}^n)$ und $\tilde{g} \in C^q(V, \mathbb{R}^m)$ mit $g(V) \subseteq U$ gilt:

$$\partial^\alpha \tilde{f}(\tilde{g}(p)) = 0 \quad \forall\, |\alpha| \leq q \qquad \Longrightarrow \qquad \partial^\alpha(\tilde{f} \circ \tilde{g})(p) = 0 \quad \forall\, |\alpha| \leq q. \tag{473}$$

Beweis. Für $q = 0$ gilt die Behauptung wegen $(\tilde{f} \circ \tilde{g})(p) = f(\tilde{g}(p)) = 0$. Wir können daher annehmen, dass ein $z \in \mathbb{N}$ existiert, sodass (473) für alle $0 \leq q \leq z$ gilt.

Seien nun f, g beides C^{z+1}-Abbildungen, sodass die linke Seite von (473) gilt. Offensichtlich gilt dann die rechte Seite für $|\alpha| = 0$. Sei also $1 \leq |\alpha| \leq z + 1$ und $1 \leq j \leq m$ mit $\alpha_j \geq 1$. Wir erhalten aus der Kettenregel (434) sowie der allgemeinen Leibniz-Regel (458)

$$\begin{aligned} \partial^\alpha\big(\tilde{f} \circ \tilde{g}\big)(p) &= \partial^{\alpha - e_j}\partial_j\big(\tilde{f} \circ \tilde{g}\big)(p) \overset{(434)}{=} \textstyle\sum_{i=1}^m \partial^{\alpha - e_j}\big((\partial_i \tilde{f} \circ \tilde{g}) \cdot \partial_j \tilde{g}_i\big)(p) \\ &\overset{(458)}{=} \textstyle\sum_{i=1}^m \sum_{\beta \leq \alpha - e_j} \binom{\alpha - e_j}{\beta} \underbrace{\partial^\beta(\overbrace{\partial_i \tilde{f}}^{=:\tilde{\tilde{f}}} \circ \tilde{g})(p)}_{\overset{(\mathrm{IV})\text{ für } \tilde{\tilde{f}}}{=} 0} \cdot \partial^{\alpha - \beta}\tilde{g}_i(0) = 0, \end{aligned}$$

was die Behauptung zeigt. □

- Wegen (466) gilt $\partial^\alpha \mathrm{R}_p^k(f)(p) = 0$ für alle $|\alpha| \leq k$, also $\mathrm{T}_p^k\big(\mathrm{R}_p^k(f) \circ g\big) = 0$ wegen dem vorhergehenden Punkt. Wir erhalten wegen $f = \mathrm{T}_p^k(f) + \mathrm{R}_p^k(f)$:

$$\mathrm{T}_p^k(f \circ g) = \mathrm{T}_p^k\big(\mathrm{T}_p^k(f) \circ g\big) + \underbrace{\mathrm{T}_p^k\big(\mathrm{R}_p^k(f) \circ g\big)}_{=0}.$$

Da $\mathrm{T}_p^k(f)$ ein Polynom ist, folgt durch iteratives Anwenden von a) und b) unter Berücksichtigung von (465), dass $\mathrm{T}_p^k\big(\mathrm{T}_p^k(f) \circ g\big) = \mathrm{T}_p^k\big(\mathrm{T}_{g(p)}^k(f) \circ \mathrm{T}_p^k(g)\big)$ gilt (vgl. Bemerkung 74.a)). □

15.4 Das lokale Verhalten von Funktionen

Im Folgenden sei $n \in \mathbb{N}_{>0}$ und $U \subseteq \mathbb{R}^n$ offen. In diesem Abschnitt studieren wir Extrema von Abbildungen $U \to \mathbb{R}$. Wir werden sehen, dass für $f \in C^2(U, \mathbb{R})$ das Extremalverhalten in der Regel durch das Taylorpolynom zweiter Ordnung bestimmt wird. Hierfür beobachten wir zunächst, dass für $f \in C^2(U, \mathbb{R})$, wegen Satz 92 (Satz von Schwarz) und (463), das Taylorpolynom zweiter Ordnung in $p \in U$ von f gegeben ist durch

$$\mathrm{T}_p^2(f)(x) = f(p) + \underbrace{\sum_{j=1}^n \partial_j f(p) \cdot (x_j - p_j)}_{=:\Lambda} + \frac{1}{2} \cdot \sum_{i,j=1}^n \partial_i\partial_j f(p) \cdot (x_i - p_i) \cdot (x_j - p_j) \tag{474}$$

mit $\Lambda = \mathrm{d}_p f(x - p) = \langle \nabla f(p), (x - p)\rangle_{\text{euk}} = J_p(f) \cdot (x - p)^\top$.

Terminologie 45 (Hessematrix und Hesseform)**.** *Sei $f \in C^2(U, \mathbb{R})$ und $p \in U$. Die* Hessematrix von f in p *ist definiert durch:*

$$H_p(f) := (\partial_i\partial_j f(p))_{1 \leq i,j \leq n} \in \mathrm{M}_{n,n}(\mathbb{R}).$$

Nach dem Satz von Schwarz (Satz 92) ist $H_p(f)$ eine symmetrische Matrix*, d.h. $H_p(f)_{ij} = H_p(f)_{ji}$ für alle $1 \leq i, j \leq n$ (bzw. $H_p(f)^\top = H_p(f)$).*

Notation: *Für eine $n \times n$-Matrix $A \in \mathrm{M}_{n,n}(\mathbb{R})$ und ein Element $v \in \mathbb{R}^n$ schreiben wir im Folgenden vereinfachend $A \cdot v \equiv (A \cdot v^\top)^\top$.*

Die Abbildung:

$$\mathcal{H}_p(f) \colon \mathbb{R}^n \to \mathbb{R}, \qquad v \mapsto \langle v, H_p(f) \cdot v\rangle_{\text{euk}} = \textstyle\sum_{i,j=1}^n \partial_i\partial_j f(p) \cdot v_i \cdot v_j,$$

heißt Hesseform von f in p*. Sie ist eine sogenannte quadratische Form auf $\mathbb{R}^n$, also eine Summe von Monomen vom Grad 2.*

- *Es ließt sich* (474) *in der Form*

$$\mathrm{T}_p^2(f)(x) = f(p) \;+\; \mathrm{d}_p f(x-p) \;+\; \frac{1}{2}\cdot \mathcal{H}_p(f)(x-p). \tag{475}$$

- *Gemäß Korollar 61 gilt* *($\|\cdot\|$ Norm auf $\mathbb{R}^n$)*

$$f = \mathrm{T}_p^2(f) + \mathrm{R}_p^2(f) \qquad \textit{mit} \qquad \lim_{x\to p} \frac{|\mathrm{R}_p^2(f)(x)|}{\|x-p\|^2} = 0. \tag{476}$$

Definition 73. *Sei $f\colon \mathbb{R}^n \supseteq U \to \mathbb{R}$ differenzierbar (U offen). Ein Punkt $p \in U$ heißt* kritischer Punkt von f, *wenn $\mathrm{d}_p f = 0$ gilt. In diesem Fall heißt $f(p)$* kritischer Wert von f.

Beachte: *Wegen* (421) *ist $\mathrm{d}_p f = 0$ gleichbedeutend mit $\partial_j f(p) = 0$ für alle $1 \le j \le n$, also $\nabla f(p) = (\partial_1 f(p), \dots, \partial_n f(p)) = 0$.*

Definition 74 (Maxima und Minima)**.** *Sei $f\colon \mathbb{R}^n \supseteq M \to \mathbb{R}$ eine Abbildung.*

a) Ein Punkt $p \in M$ heißt lokales Maximum *von f, wenn ein $\varepsilon > 0$ existiert mit*

$$f(x) \le f(p) \qquad \forall\, x \in \mathsf{B}_\varepsilon(p) \subseteq M.$$

Es heißt p ein isoliertes lokales Maximum *von f, wenn ein $\varepsilon > 0$ existiert mit*

$$f(x) < f(p) \qquad \forall\, x \in \mathsf{B}_\varepsilon(p) \subseteq M.$$

b) Ein Punkt $p \in M$ heißt lokales Minimum *von f, wenn ein $\varepsilon > 0$ existiert mit*

$$f(p) \le f(x) \qquad \forall\, x \in \mathsf{B}_\varepsilon(p) \subseteq M.$$

Es heißt p ein isoliertes lokales Minimum *von f, wenn ein $\varepsilon > 0$ existiert mit*

$$f(p) < f(x) \qquad \forall\, x \in \mathsf{B}_\varepsilon(p) \subseteq M.$$

c) Ein Punkt $p \in M$ heißt (isoliertes) lokales Extremum von f, *wenn p ein (isoliertes) lokales Maximum oder ein (isoliertes) lokales Minimum von f ist.*

d) Ein Punkt $p \in U$ heißt (isoliertes) globales Maximum *bzw.* (isoliertes) globales Minimum, *wenn ($f(x) < f(p)$) $f(x) \le f(p)$ bzw. ($f(x) > f(p)$) $f(x) \ge f(p)$ für alle $x \in M$ gilt. (Jedes globale Maximum/Minimum ist auch ein lokales Maximum bzw. Minimum.)*

Ein kritischer Punkt der kein lokales Extremum (lokales Maximum/Minimum) ist heißt Sattelpunkt, *d.h., p ist genau dann ein Sattelpunk von f, wenn gilt:*

$$\forall\, \varepsilon > 0\colon\;\; \exists\, x_\pm \in \mathsf{B}_\varepsilon(p)\colon\;\; f(x_-) < f(p) < f(x_+). \tag{477}$$

Lemma 90 (Notwendige Bedingung für Extrema)**.** *Sei $f\colon \mathbb{R}^n \supseteq U \to \mathbb{R}$ differenzierbar (U offen). Hat f in p ein lokales Extremum, so ist p ein kritischer Punkt.*

Beweis. Sei $\varepsilon > 0$ mit $\mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_\varepsilon(p) \subseteq U$.

- Für $1 \le j \le n$ ist gemäß Bemerkung 93.b) und (419)

  $$g[j]\colon (-\varepsilon,\varepsilon) \ni t \to f(p + t\cdot e_j) \in \mathbb{R} \quad \text{differenzierbar mit} \quad g[j]'\colon (-\varepsilon,\varepsilon) \ni t \mapsto \partial_j f(p + t\cdot e_j) \in \mathbb{R}.$$

- Für jedes $1 \le j \le n$ ist nun 0 eine Extremstelle von $g[j]$, also gilt $\partial_j f(p) = g[j]'(0) = 0$ wegen Lemma 50.5). Dies zeigt $\mathrm{d}_p f = 0$ (vgl. Definition 73). □

Beispiel 107. • *Sei* $f\colon \mathbb{R}^2 \ni (x,y) \mapsto x^2+y^2 \in \mathbb{R}$. *Dann gilt*

$$\nabla f(x,y) = (2x,2y) = 0 \qquad \Longleftrightarrow \qquad (x,y)=(0,0).$$

Daher ist der Nullpunkt der einzige kritische Punkt, und dort liegt ein isoliertes globales Minimum vor ($f(0,0) < f(x,y)$ *für alle* $(x,y) \neq (0,0)$).

Entsprechend hat $f\colon \mathbb{R}^2 \ni (x,y) \mapsto -(x^2+y^2) \in \mathbb{R}$ *im Nullpunkt ein isoliertes globales Maximum* ($f(0,0) > f(x,y)$ *für alle* $(x,y) \neq (0,0)$).

• *Die Abbildung* $f\colon \mathbb{R}^2 \ni (x,y) \mapsto x^2-y^2 \in \mathbb{R}$ *hat wegen*

$$\nabla f(x,y) = (2x,-2y) = 0 \qquad \Longleftrightarrow \qquad (x,y)=(0,0).$$

in $(0,0)$ *den einzigen kritischen Punkt. Dieser ist aber kein lokales Extremum wegen*

$$f(x,0) = x^2 > 0 = f(0,0) \quad \forall\, \mathbb{R} \ni x \neq 0 \qquad \textit{sowie} \qquad f(0,y) = -y^2 < 0 = f(0,0) \qquad \forall\, \mathbb{R} \ni y \neq 0,$$

also ein Sattelpunkt.

Definition 75 (Linearen Algebra). *Sei* $A \in \mathrm{M}_{n,n}(\mathbb{R})$ *eine symmetrische* $(n \times n)$*-Matrix.*

a) *A heißt* positiv definit, *wenn* $\langle x, A\cdot x\rangle_{\text{euk}} > 0$ *für alle* $0 \neq x \in \mathbb{R}^n$ *gilt.*

b) *A heißt* positiv semidefinit, *wenn* $\langle x, A\cdot x\rangle_{\text{euk}} \geq 0$ *für alle* $0 \neq x \in \mathbb{R}^n$ *gilt.*

c) *A heißt* negativ (semi-)definit, *wenn* $-A$ positiv (semi-)definit *ist.*[122]

d) *A heißt* indefinit, *wenn es* $x, y \in \mathbb{R}^n$ *gibt, sodass* $\langle x, A\cdot x\rangle_{\text{euk}} > 0$ *und* $\langle y, A\cdot y\rangle_{\text{euk}} < 0$ *gilt.*

Bemerkung 102. *Sei* $A \in \mathrm{M}_{n,n}(\mathbb{R})$ *eine symmetrischen* $n \times n$*-Matrix* ($A^\top = A$) *mit* $n \in \mathbb{N}_{>0}$:

a) *Aus der Linearen Algebra ist bekannt, dass eine Orthonormalbasis* $v_1, \ldots, v_n$ *von* $\mathbb{R}^n$ *aus Eigenvektoren von A existiert, d.h.* $v_1, \ldots, v_n$ *ist eine Basis von* $\mathbb{R}^n$ *mit* $\langle v_i, v_j\rangle_{\text{euk}} = \delta_{ij}$ *für* $1 \leq i, j \leq n$ *und es existieren* $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ *mit*

$$A \cdot v_j = \lambda_j \cdot v_j \qquad \forall\, 1 \leq j \leq n.$$

Für einen Vektor $x = \sum_{j=1}^n \nu_j \cdot v_j \in \mathbb{R}^n$ *gilt dann*

$$\langle x, A\cdot x\rangle_{\text{euk}} = \textstyle\sum_{j=1}^n \lambda_j \cdot \nu_j^2. \tag{478}$$

Hiermit folgt durch einsetzen der Basisvektoren $v_1, \ldots, v_n$ *(Übung):*

- *A ist genau dann positiv (semi)definit, wenn* ($\lambda_1, \ldots, \lambda_n \geq 0$) $\lambda_1, \ldots, \lambda_n > 0$ *gilt.*
- *A ist genau dann negativ (semi)definit, wenn* ($\lambda_1, \ldots, \lambda_n \leq 0$) $\lambda_1, \ldots, \lambda_n < 0$ *gilt.*
- *A ist genau dann indefinit, wenn sowohl positive* (> 0) *als auch negative* (< 0) *Eigenwerte existieren.*

Beachte: *Für die Matrix* $S := \left(v_1^\top \ldots v_n^\top\right) \in \mathrm{M}_{n,n}(\mathbb{R})$ *gilt dann*

$$\underbrace{S \cdot S^\top = \mathbf{1}_n = S^\top S}_{S \text{ ist orthogonal}} \qquad \textit{sowie} \qquad S^\top A S = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix},$$

und mit der Multiplikativität der Determinantenfunktion[123]

$$\det\colon \mathrm{M}_{n,n}(\mathbb{R}) \to \mathbb{R}, \qquad (A_{ij}) \mapsto \textstyle\sum_{\sigma \in \mathsf{S}_n} \operatorname{sgn}(\sigma) \cdot A_{1,\sigma(1)} \cdots A_{n,\sigma(n)}$$

erhalten wir

$$\det(A) = \det(S^\top S) \cdot \det(A) = \det(S^\top) \cdot \det(A) \cdot \det(S) = \det(S^\top A S) = \lambda_1 \cdot \ldots \cdot \lambda_n. \tag{479}$$

[122] Dies ist gleichbedeutend dazu, dass in den entprechenden Ungleichungen „$<$“ bzw. „$\leq$“ anstelle „$>$“ bzw. „$\geq$“ steht.

[123] D.h., $\det(B \cdot C) = \det(B) \cdot \det(C)$ für $B, C \in \mathrm{M}_{n,n}(\mathbb{R})$.

b) (∗) *Ist $B \in \mathrm{GL}_n(\mathbb{R})$ eine invertierbare Matrix, so ist A genau dann positiv definit, wenn $B^\top A \cdot B$ positiv definit ist.*

Beweis. Eine einfache Rechnung zeigt

$$\langle v, B^\top AB \cdot v\rangle_{\text{euk}} = \langle B \cdot v, AB \cdot v\rangle_{\text{euk}} \qquad \forall\, n \in \mathbb{R}^n.$$

Da die Multiplikation mit B (also $\mathbb{R}^n \ni v \mapsto B \cdot v \in \mathbb{R}^n$) bijektiv ist, ist die rechte Seite (also auch die linke Seite) genau dann für alle $0 \neq v \in \mathbb{R}^n$ positiv (> 0), wenn $\langle w, Aw\rangle > 0$ für alle $0 \neq w \in \mathbb{R}^n$ gilt. Also ist A genau dann positiv definit, wenn dies für $B^\top AB$ der Fall ist. □

Satz 94 (Hurwitzkriterium)**.** *Eine symmetrische $n \times n$-Matrix $A = (A_{ij})_{1\le i,j\le n}$ ist genau dann positiv definit, wenn alle Hauptminoren positiv sind, d.h. wenn gilt:*

$$\det(A_k) > 0 \qquad \textit{mit} \qquad A_k := \begin{pmatrix} A_{11} & \dots & A_{1k} \\ \vdots & & \vdots \\ A_{k1} & \dots & A_{kk} \end{pmatrix} \qquad \textit{für alle} \qquad 1 \le k \le n.$$

∗*Beweis.* • Ist A positiv definit, so sind auch alle Matrizen A_k positiv definit: Für $0 \neq x \in \mathbb{R}^k$ setzen wir $\tilde{x} = (x_1, \dots, x_k, 0, \dots, 0) \in \mathbb{R}^n$. Dann gilt[124]

$$\langle x, A_k \cdot x\rangle_{\text{euk}} = \langle \tilde{x}, A \cdot \tilde{x}\rangle_{\text{euk}} > 0.$$

Insbesondere sind alle Eigenwerte von A_k positiv (Bemerkung 102.a)); also gilt $\det(A_k) > 0$ wegen (479).

• Es gelte $\det(A_k) > 0$ für alle $1 \le k \le n$. Wir zeigen durch Induktion nach $n \in \mathbb{N}_{>0}$, dass dann A positiv definit ist. Für $n = 1$ ist die Behauptung klar. Es gelte nun die Behauptung für ein $n \geq 1$; und es sei A eine symmetrische $(n+1) \times (n+1)$-Matrix mit $\det(A_k) > 0$ für alle $1 \le k \le n+1$.

Für $\tilde{A} := A_n$, gilt dann ebenfalls $\det(\tilde{A}_k) > 0$ für alle $1 \le k \le n$. Per Induktionsvoraussetzung ist somit $\tilde{A} = A_n$ positiv definit, d.h, es existiert eine Orthonormalbasis von Eigenvektoren $v_1, \dots, v_n$ von $\tilde{A}$ mit entsprechenden positiven Eigenwerten $\lambda_1, \dots, \lambda_n > 0$. Für $\tilde{S} := (v_1^\top \dots v_n^\top)$ gilt dann

$$\underbrace{\tilde{S} \cdot \tilde{S}^\top = \mathbf{1}_n = \tilde{S}^\top \tilde{S}}_{\tilde{S}\text{ ist orthogonal}} \qquad \text{sowie} \qquad \tilde{S}^\top \tilde{A} \tilde{S} = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}.$$

Dann ist $S := \begin{pmatrix} \tilde{S} & 0 \\ 0 & 1 \end{pmatrix}$ ebenfalls orthogonal, und wir erhalten

$$B := S^\top AS = \begin{pmatrix} \lambda_1 & \dots & 0 & b_1 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & \lambda_n & b_n \\ b_1 & \dots & b_n & b_{n+1} \end{pmatrix}$$

mit gewissen $b_1, \dots, b_{n+1} \in \mathbb{R}$. Nach Voraussetzung ist (vgl. (479)) $\det(B) = \det(A_{n+1}) = \det(A) > 0$. Wir setzen

$$T := \begin{pmatrix} 1 & \dots & 0 & c_1 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 1 & c_n \\ 0 & \dots & 0 & 1 \end{pmatrix} \qquad \text{mit} \qquad c_j := -\frac{b_j}{\lambda_j} \quad \text{für} \quad j = 1, \dots, n$$

[124] Links steht das euklidische Skalarprodukt auf $\mathbb{R}^k$ und rechts das euklidische Skalarprodukt auf $\mathbb{R}^n$.

und erhalten

$$C := T^\top BT = \begin{pmatrix} \lambda_1 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & \lambda_n & 0 \\ 0 & \dots & 0 & \lambda_{n+1} \end{pmatrix} \qquad \text{mit} \qquad \lambda_{n+1} := b_n - \frac{b_1^2}{\lambda_1} - \ldots - \frac{b_n^2}{\lambda_n}.$$

Wegen $\det(T) = 1$ gilt $0 < \det(B) = \det(C) = \lambda_1 \cdot \dots \cdot \lambda_{n+1}$, also $\lambda_{n+1} > 0$ wegen $\lambda_1, \dots, \lambda_n > 0$. Somit ist $C = (ST)^\top A \cdot (ST)$ positiv definit, also auch A wegen Bemerkung 102.b). □

Der Vorteil des Hurwitzkriteriums ist, dass man die Eigenwerte nicht kennen muss um auszurechnen, ob eine Matrix positiv oder negativ definit ist.

Korollar 62. *Sei $A = (A_{ij})_{1 \le i,j \le n}$ eine symmetrische (2×2)-Matrix. Zeigen Sie:*

1) *A indefinit* $\Leftrightarrow$ $\det(A) < 0$ *(nicht Hurwitz)*
2) *A positiv definit* $\Leftrightarrow$ $\det(A) > 0 \ \wedge \ A_{11} > 0$
3) *A negativ definit* $\Leftrightarrow$ $\det(A) > 0 \ \wedge \ A_{11} < 0$

Beweis. 1) Ist $\det(A) < 0$, so haben wegen (479) die beiden Eigenwerte von A verschiedene Vorzeichen. Also ist A indefinit.

2) Folgt wegen $\det(A_1) = A_{11}$ und $\det(A_2) = \det(A)$ direkt aus dem Hurwitzkriterium.

3) Folgt wegen $\det(-A) = \det(A)$ (beachte $n = 2$) direkt aus Teil 2).[125] □

Satz 95 (Hinreichende Bedingung für Extrema)**.** *Sei $U \subseteq \mathbb{R}^n$ eine offene Menge, $f \in C^2(U, \mathbb{R})$ und $p \in U$ ein kritischer Punkt von f. Dann gilt:*

1) *Ist $H_p(f)$ positiv definit, so ist p ein isoliertes lokales Minimum von f.*
2) *Ist $H_p(f)$ negativ definit, so ist p ein isoliertes lokales Maximum von f.*
3) *Ist $H_p(f)$ indefinit, so ist p ein Sattelpunkt von f.*

Beweis. Gemäß Terminologie 45 gilt (vgl. (475) mit $\mathrm{d}_p f = 0$ links, sowie (476) rechts)

$$f = \underbrace{\mathrm{T}_p^2(f)}_{f(p) \, + \, \frac{1}{2} \cdot \mathcal{H}_p(f)(\cdot - p)} + \mathrm{R}_p^2(f) \qquad \text{mit} \qquad \lim_{x \to p} \frac{|\mathrm{R}_p^2(f)(x)|}{\|x - p\|_{\text{euk}}^2} = 0. \tag{480}$$

Zu $\varepsilon > 0$ vorgegeben, existiert daher ein $\delta > 0$ mit $(x \equiv p + h)$

$$\mathsf{B}[\mathrm{d}_{\|\cdot\|_{\text{euk}}}]_\delta(p) \subseteq U \quad \text{und} \quad |\mathrm{R}_p^2(f)(p+h)| \le \varepsilon \cdot \|h\|_{\text{euk}}^2 \quad \text{für alle} \quad h \in \mathsf{B}[\mathrm{d}_{\|\cdot\|_{\text{euk}}}]_\delta(0). \tag{481}$$

1) Sei $H_p(f)$ positiv definit, und $v_1, \dots, v_n$ eine Orthonormalbasis von Eigenvektoren von $H_p(f)$ mit Eigenwerten $\lambda_1, \dots, \lambda_n > 0$ (vgl. Bemerkung 102.a)).

   • Wir setzen $\lambda := \min(\lambda_1, \dots, \lambda_n) > 0$, und erhalten für $\mathbb{R}^n \ni h = \sum_{j=1}^n \nu_j \cdot v_j$:

$$\mathcal{H}_p(f)(h) = \langle h, H_p(f) \cdot h \rangle_{\text{euk}} \overset{(478)}{=} \textstyle\sum_{j=1}^n \lambda_j \cdot |\nu_j|^2 \ge \lambda \cdot \sum_{j=1}^n |\nu_j|^2 = \lambda \cdot \|h\|_{\text{euk}}^2.$$

$$\Big(\|h\|_{\text{euk}}^2 = \langle h, h \rangle_{\text{euk}} = \langle \textstyle\sum_{i=1}^n \nu_i \cdot v_i, \sum_{j=1}^n \nu_j \cdot v_j \rangle_{\text{euk}} = \sum_{i,j=1}^n \nu_i \cdot \nu_j \cdot \langle v_i, v_j \rangle_{\text{euk}} = \sum_{i=1}^n |\nu_i|^2 \Big)$$

[125] A negativ definit $\overset{\text{def.}}{\Longleftrightarrow}$ $-A$ positiv definit

- Sei $\varepsilon := \frac{\lambda}{4}$, und $\delta > 0$ wie in (481). Für $0 \neq h \in \mathsf{B}[\mathrm{d}_{\|\cdot\|_{\mathrm{euk}}}]_\delta(0)$, gilt dann

$$f(p+h) - f(p) \overset{(480)}{=} \frac{1}{2} \cdot \mathcal{H}_p(f)(h) + \underbrace{\mathrm{R}_p^2(f)(p+h)}_{\geq -\frac{\lambda}{4} \cdot \|h\|^2_{\mathrm{euk}}} \geq \frac{\lambda}{2} \cdot \|h\|^2_{\mathrm{euk}} - \frac{\lambda}{4} \cdot \|h\|^2_{\mathrm{euk}} = \frac{\lambda}{4} \cdot \|h\|^2_{\mathrm{euk}} > 0.$$

Also ist p ein isoliertes lokales Minimum.

2) Folgt analog zu 1), bzw. durch Anwendung von 1) auf $-f$.

3) Per Voraussetzung existieren $0 \neq h_\pm \in \mathbb{R}^n$ mit

$$\mu_+ := \langle h_+, H_p(f) \cdot h_+ \rangle_{\mathrm{euk}} > 0 \qquad \text{und} \qquad \mu_- := \langle h_-, H_p(f) \cdot h_- \rangle_{\mathrm{euk}} < 0. \tag{482}$$

Wir wählen $\delta > 0$ wie in (481) für $\varepsilon := \min(|\mu_+| \cdot \|h_+\|^{-2}_{\mathrm{euk}},\ |\mu_-| \cdot \|h_-\|^{-2}_{\mathrm{euk}})/4$; also

$$\begin{aligned} |t| < \tilde{\delta} := \tfrac{\delta}{\max(\|h_+\|_{\mathrm{euk}}, \|h_-\|_{\mathrm{euk}})} \quad &\Longrightarrow \quad \|t \cdot h_\pm\|_{\mathrm{euk}} < \delta \\ &\Longrightarrow \quad |\mathrm{R}_p^2(f)(p + t \cdot h_\pm)| \overset{(481)}{\leq} \varepsilon \cdot \underbrace{\|t \cdot h_\pm\|^2_{\mathrm{euk}}}_{t^2 \cdot \|h_\pm\|^2_{\mathrm{euk}}} \leq t^2 \cdot \tfrac{|\mu_\pm|}{4}. \end{aligned}$$

Für $0 < t < \tilde{\delta}$ erhalten wir mit der linken Seite von (480):

$$\begin{aligned} f(p + t \cdot h_+) - f(p) &= \frac{t^2}{2} \cdot \overbrace{\langle h_+, H_p(f) \cdot h_+ \rangle_{\mathrm{euk}}}^{\mu_+} + \mathrm{R}_p^2(f)(p + t \cdot h_+) \geq \frac{t^2}{2} \cdot \mu_+ - t^2 \cdot \frac{|\mu_+|}{4} > 0 \\ f(p + t \cdot h_-) - f(p) &= \frac{t^2}{2} \cdot \underbrace{\langle h_-, H_p(f) \cdot h_- \rangle_{\mathrm{euk}}}_{\mu_-} + \mathrm{R}_p^2(f)(p + t \cdot h_-) \leq \frac{t^2}{2} \cdot \mu_- + t^2 \cdot \frac{|\mu_-|}{4} < 0, \end{aligned}$$

womit p wegen (477) ein Sattelpunkt von f ist. □

Das selbe Argument wie im Beweis von Satz 95.3) liefert eine notwendige Bedingung für das Vorliegen von Extremstellen.

Korollar 63 (Notwendige Bedingung für Extrema)**.** *Sei $U \subseteq \mathbb{R}^n$ offen, $f \in C^2(U, \mathbb{R})$ und $p \in U$ ein kritischer Punkt von f. Dann gilt:*[126]

a) *Ist p ein lokales Maximum, so ist $H_p(f)$ negativ semidefinit.*

b) *Ist p ein lokales Minimum, so ist $H_p(f)$ positiv semidefinit.*

Beweis durch Kontraposition. Sei $H_p(f)$ nicht negativ semidefinit (bzw. nicht positiv semidefinit). Dann existiert ein $0 \neq h_+ \in \mathbb{R}^n$ (bzw. ein $0 \neq h_- \in \mathbb{R}^n$) wie in (482). Wie im Beweis von Satz 95.3) finden wir ein $\delta > 0$ mit $f(p + t \cdot h_+) - f(p) > 0$ (bzw. $f(p + t \cdot h_-) - f(p) < 0$) für alle $0 < t < \delta$, womit $f(p)$ kein lokales Maximum (bzw. kein lokales Minimum) sein kann. □

Bemerkung 103. *Die Bedingung in Korollar 63 ist üblicherweise nicht hinreichend, d.h., ist p ein kritischer Punkt und ist die Hessematrix semidefinit, so lassen sich keine allgemeinen Aussagen tätigen. Beispielsweise ist (der Nullpunkt) $(0,0)$ ein kritischer Punkt der Funktionen $f, g, h\colon \mathbb{R}^2 \to \mathbb{R}$:*

$$f(x,y) = x^2 + y^4, \qquad g(x,y) = x^2 \qquad \textit{und} \qquad h(x,y) = x^2 + y^3,$$

welche im Nullpunkt alle die positiv semidefinite Hessematrix $\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ besitzen:

- *f hat im Nullpunkt ein isoliertes globales Minimum,*
- *g hat im Nullpunkt ein (nichtisoliertes) globales Minimum,*
- *h hat im Nullpunkt einen Sattelpunkt (kein Extremum).*

[126] Ist also p ein kritischer Punkt von f und $H_p(f)$ weder positiv semidefinit noch negativ semidefinit, so ist p automatisch ein Sattelpunkt. Dies ist auch deswegen klar, weil dann $H_p(f)$ sowohl einen positiven als auch einen negativen Eigenwert haben muss; denn es sind ja weder alle Eigenwerte ≤ 0 noch sind alle Eigenwert ≥ 0 (vgl. Bemerkung 102.a)).

16 Der Satz über die Umkehrfunktion

In diesem Kapitel untersuchen wir, inwiefern sich Differenzierbarkeitseigenschaften bijektiver Abbildungen auf deren Umkehrabbildung übertragen. In Abschnitt 16.3.1 verallgemeinern wir den Satz über die Differenzierbarkeit der Umkehrfunktion (Satz 42) auf Funktionen in mehreren Veränderlichen (siehe Satz 98). Der entsprechende Satz über die lokale Umkehrbarkeit (Satz 99) bzw. der Satz über die Umkehrfunktion (globale Variante) ist ein zentrales Ergebnis der Differentialrechnung mehrerer Veränderlicher. Er hat wichtige Anwendungen auf die Beschreibung von Lösungsmengen von Gleichungen, denen wir uns in Abschnitt 16.4 zuwenden. Für den Beweis besagter Aussagen benötigen allerdings noch etwas Vorbereitung; insbesondere werden wir sogleich den Banachschen Fixpunktsatz beweisen, welcher bereits für sich genommen ein zentrales Ergebnis der Analysis ist:

16.1 Der Banachsche Fixpunktsatz

In diesem Abschnitt beweisen wir den Banachschen Fixpunktsatz, der das zentrale Werkzeug zur Lösung nichtlinearer Gleichungen ist, d.h. Gleichungen der Form $g(x) = y$ mit einer vorgegebenen Funktion $g\colon X \to Y$, einem fixierten Element $y \in Y$, und dem zu ermittelnden Element $x \in X$.

Genauer betrachten wir die Situation, in der Y ein $\mathbb{K}$-Vektorraum ist (genauer $Y = \mathbb{R}^n$ mit $n \in \mathbb{N}_{>0}$) und $X \subseteq Y$ eine Teilmenge. Die grundlegende Idee ist dann, dass man die Gleichung $g(x) = y$ oft als Fixpunktgleichung

$$f_y(x) = x \qquad \text{mit} \qquad f_y\colon X \to X, \quad x \mapsto x + (y - g(x)) \tag{483}$$

umschreiben kann, wobei natürlich $f_y(X) \subseteq X$ gewährleistet werden muss.

Notation 33. *Ist X eine Menge und $f\colon X \to X$ eine Abbildung, so bezeichnet im Folgenden f^n für $n \in \mathbb{N}_{>0}$ die n-fache Verkettung von f mit sich selbst. Weiterhin notieren wir $f^1 := f$ und $f^0 := \mathrm{id}_X$.*

Definition 76. *Sei (X, d) ein metrischer Raum. Eine Abbildung $f\colon X \to X$ heißt* Kontraktion *(oder* kontraktive Selbstabbildung*) mit* Kontraktionskonstante $0 \leq \lambda < 1$, *wenn gilt:*

$$\mathrm{d}(f(x), f(y)) \leq \lambda \cdot \mathrm{d}(x, y) \qquad \forall\, x, y \in X. \tag{484}$$

Kontraktionen sind also Lipschitz-stetige Abbildungen mit Lipschitz-Konstante $0 \leq L \equiv \lambda < 1$.

Beachte: *Ist $(V, \|\cdot\|)$ ein normierter Raum, so liest sich (484) natürlich in der Form*

$$\underbrace{\|f(x) - f(y)\|}_{\mathrm{d}_{\|\cdot\|}(f(x), f(y)))} \leq \lambda \cdot \underbrace{\|x - y\|}_{\mathrm{d}_{\|\cdot\|}(x,y)} \qquad \forall\, x, y \in V.$$

Der folgende grundlegende Satz garantiert die Existenz und Eindeutigkeit von Lösungen von Fixpunktgleichungen der Form (483) im Falle vollständiger metrischer Räume.

Satz 96 (Banachscher Fixpunktsatz)**.** *Sei (X, d) ein vollständiger metrischer Raum. Dann hat jede Kontraktion $f\colon X \to X$ genau einen Fixpunkt, d.h., es existiert genau ein $x \in X$ mit $f(x) = x$. Genauer gilt:*

$$x = \lim\nolimits_n f^n(y) \qquad \textit{für jedes fixierte} \qquad y \in X. \tag{485}$$

Beweis. Sei $0 \leq \lambda < 1$ die Kontraktionskonstante von f.

- Der Fixpunkt ist eindeutig sofern er existiert; denn sind $X \ni x \neq y \in X$ beides Fixpunkte von f, so erhalten wir den Widerspruch: (d$(x, y) > 0$ wegen $x \neq y$)

$$\mathrm{d}(x, y) = \mathrm{d}(f(x), f(y)) \leq \lambda \cdot \mathrm{d}(x, y) \overset{\mathrm{d}(x,y)>0}{<} \mathrm{d}(x, y).$$

- Für die Existenzaussage sowie (485) genügt es nachzuweisen, dass für jedes fixiertes $y \in X$, die Folge $(f^n(y))_{n\in\mathbb{N}}$ gegen ein $x_y \in X$ konvergiert. Wegen der Stetigkeit (Lipschitzstetigkeit) von f, gilt dann nämlich automatisch:
$$x_y = \lim_n f^{n+1}(y) = \lim_n f(f^n(y)) = f(\lim_n f^n(y)) = f(x_y).$$
Da (X, d) vollständig ist, müssen nun lediglich zeigen, dass $(f^n(y))_{n\in\mathbb{N}}$ eine Cauchy-Folge ist ($y \in X$ fixiert). Seien hierfür $n, k \in \mathbb{N}_{>0}$. Wir erhalten durch iterierte Anwendung von (484) (erster und dritter Schritt), Mehrfachanwendung der Dreiecksungleichung (zweiter Schritt), sowie mit der geometrischen Summenformel (Korollar 22.1) im vorletzten Schritt): $(f^0 = \mathrm{id}_X,\ f^1 = f)$
$$\begin{aligned}
\mathrm{d}\big(f^n(y), f^{n+k}(y)\big) &\leq \lambda^n \cdot \mathrm{d}\big(y, f^k(y)\big) \\
&\leq \lambda^n \cdot \overbrace{\big(\mathrm{d}(y, f(y)) + \mathrm{d}(f(y), f^2(y)) + \cdots + \mathrm{d}(f^{k-1}(y), f^k(y))\big)}^{\sum_{\ell=0}^{k-1} \mathrm{d}\left(f^\ell(y), f^{\ell+1}(y)\right)} \\
&\leq \lambda^n \cdot \textstyle\sum_{\ell=0}^{k-1} \lambda^\ell \cdot \mathrm{d}\big(y, f(y)\big) \\
&= \lambda^n \cdot \big(\textstyle\sum_{\ell=0}^{k-1} \lambda^\ell\big) \cdot \mathrm{d}(y, f(y)) \\
&= \lambda^n \cdot \tfrac{1-\lambda^k}{1-\lambda} \cdot \mathrm{d}(y, f(y)) \\
&\leq \tfrac{\lambda^n}{1-\lambda} \cdot \mathrm{d}(y, f(y)).
\end{aligned}$$
Wegen $0 \leq \lambda < 1$ gilt nun $\lim_n \lambda^n = 0$ (Übung 60 und Proposition 2), womit die Cauchy-Eigenschaft unmittelbar folgt. □

Bemerkung 104. *Die Bedingung der Vollständigkeit von (X, d) in Satz 96 ist wesentlich. Wir erinnern hierfür an Lemma 72:*
$$(X, \mathrm{d}) \text{ vollständig und } Y \subseteq X\colon \qquad (Y, \mathrm{d}_Y) \text{ vollständig} \quad \Leftrightarrow \quad Y \text{ abgeschlossen in } (X, \mathrm{d}).$$
Beispielsweise ist $(X, \mathrm{d}) := (\mathbb{R}, \mathrm{d}_{|\cdot|})$ vollständig wegen Satz 17; aber $Y := (0, 1]$ ist nicht abgeschlossen in $(\mathbb{R}, \mathrm{d}_{|\cdot|})$, also ist (Y, d_Y) nicht vollständig. Entsprechend existiert die Kontraktion
$$f_Y\colon Y = (0, 1] \to (0, 1] = Y, \qquad y \mapsto \frac{y}{2},$$
die keinen Fixpunkt in $Y = (0, 1]$ hat, denn $x = \frac{x}{2}$ für $x \in \mathbb{R}$ impliziert ja $x = 0 \notin (0, 1] = Y$.[127]

Beispiel 108. *Sei $f \in C^1([a, b], \mathbb{R})$ mit $f([a, b]) \subseteq [a, b]$ und $|f'|_\infty =: \lambda < 1$. Der Hauptsatz liefert*
$$\mathrm{d}_{|\cdot|}(f(x), f(y)) = |f(x) - f(y)| = \big|\textstyle\int_x^y f'\big| \leq \big|\int_x^y |f'|\big| \leq \big|\int_x^y \lambda\big| = \lambda \cdot |x - y| = \lambda \cdot \mathrm{d}_{|\cdot|}(x, y).$$
Somit ist f eine kontraktive Selbstabbildung mit Kontraktionskonstanten λ, hat also einen eindeutigen Fixpunkt $x \in [a, b]$.[128]

Beispiel 109. *Der Banachsche Fixpunktsatz hat eine anschauliche Konsequenz:*

- *Wir interpretieren $X := \{x \in \mathbb{R}^2 \mid \|x\|_{\mathrm{euk}} \leq 1\}$ als Teilmenge der „lokal flachen" Erdoberfläche. (Wegen Lemma 72 ist $\big(X, (\mathrm{d}_{\|\cdot\|_{\mathrm{euk}}})_X\big)$ ein vollständiger metrischer Raum.)*
- *Sei $z \in X$ und $0 < \lambda < 1$ mit $K := \mathsf{B}_\lambda(z) \subseteq X$: Wir interpretieren K als eine um den Maßstab λ (um den Mittelpunkt $(0, 0) \in X$) verkleinerte Karte von X, die wir am Punkt z im Bereich X auf den Boden gelegt haben.*

[127] Alternativ: Die Kontraktion $f\colon X = \mathbb{R} \ni x \mapsto \frac{x}{2} \in \mathbb{R} = X$ hat den Fixpunkt $x = 0$; und wegen $f_Y = f|_Y$ ist jeder Fixpunkt von f_Y automatisch auch ein Fixpunt von f. Wegen der Eindeutigkeit des Fixpunktes von f, kann also f_Y wegen $x \notin (0, 1]$ keinen Fixpunkt haben.

[128] Beachte: Für $f = \mathrm{id}_{[a,b]}$ gilt $|f'| = f' = 1$, und jedes $x \in [a, b]$ ist Fixpunkt von f. Die Bedingung $\lambda < 1$ ist also tatsächlich essentiell; in diesem Fall für die Eindeutigkeit des Fixpunktes.

Die folgende Abbildung ist offensichtlich kontraktiv:

$$f\colon X \to X, \qquad x \mapsto z + \lambda \cdot x.$$

Anschaulich ordnet sie einem Punkt $x \in X$ denjenigen Punkt $y \in X$ zu, welcher direkt unter dem x in der Karte markierenden Punkt liegt. Satz 96 zeigt nun, dass f genau einen Fixpunkt hat. Dieser ist gerade der (eindeutige) Punkt $x \in X$ (der sich wegen $f(X) \subseteq K$ zwangsweise unter der auf dem Boden liegenden Karte befindet), der unter demjenigen Punkt auf der Karte liegt, welcher ihn in der Karte markiert. (Markiert man x auf der Karte, und sticht am eingetragenen Punkt mit einem Degen durch die Karte in den Boden, so befindet sich die Einstichstelle im Boden gerade bei x.)

16.2 Invertierbare Lineare Abbildung

In diesem Abschnitt beschäftigen wir uns mit einigen grundlegenden Eigenschaften invertierbarer linearer Abbildungen.

16.2.1 Die Neumannsche Reihe*

Notation 34. *Gegeben ein normierter Raum $(V, \|\cdot\|)$, so bezeichnet $\mathsf{Aut}_{\mathcal{C}}(V) \subseteq \mathsf{Aut}(V)$ die Menge aller stetigen Automorphismen $\phi\colon V \to V$, deren Umkehrabbildung ϕ^{-1} ebenfalls stetig ist.*

Satz 97. *Sei $(V, \|\cdot\|)$ ein Banachraum, und sei $\phi \in \mathsf{Aut}_{\mathcal{C}}(V)$ mit $\|\phi\|_{\mathrm{op}} < 1$. Dann gilt $(\mathrm{id}_V - \phi) \in \mathsf{Aut}_{\mathcal{C}}(V)$. Das stetige Inverse von $(\mathrm{id}_V - \phi)$ ist gegeben durch die sogenannte* Neumannsche Reihe*:*

$$(\mathrm{id}_V - \phi)^{-1} = \textstyle\sum_{n=0}^{\infty} \phi^n \qquad \textit{mit} \qquad \|(\mathrm{id}_V - \phi)^{-1}\|_{\mathrm{op}} \leq \frac{1}{1-\|\phi\|_{\mathrm{op}}}. \tag{486}$$

D.h., die Folge $(S_n)_{n\in\mathbb{N}}$ der stetigen linearen Abbildungen[129] *$S_n := \sum_{k=0}^{n} \phi^n$ für $n \in \mathbb{N}$ konvergiert bezüglich der Operatornorm $\|\cdot\|_{\mathrm{op}}$ gegen $(\mathrm{id}_V - \phi)^{-1}$ mit*

$$(\mathrm{id}_V - \phi)^{-1}\colon V \to V, \qquad v \mapsto \lim_n S_n(v) = \textstyle\sum_{n=0}^{\infty} \phi^n(v). \tag{487}$$

Beweis. Wir erhalten induktiv aus (381):

$$\|\phi^n(v)\| \overset{(381)}{\leq} \|\phi\|_{\mathrm{op}}^n \cdot \|v\| \qquad \forall\, n \in \mathbb{N},\ v \in V \tag{488}$$

$$\overset{\|\phi\|_{\mathrm{op}}<1}{\Longrightarrow} \qquad \lim_n \phi^n(v) = 0 \qquad \forall\, v \in V. \tag{489}$$

- Wir zeigen zunächst, dass die Abbildung (vgl. (487))

$$\Theta\colon V \to V, \qquad v \mapsto \lim_n S_n(v) = \textstyle\sum_{n=0}^{\infty} \phi^n(v)$$

  definiert ist. Für $v \in V$, $\ell \in \mathbb{N}_{>0}$ und $n \in \mathbb{N}_{>0}$ gilt (Korollar 22.1) im vierten Schritt):

$$\begin{aligned}\|S_{n+\ell}(v) - S_{n-1}(v)\| &\leq \textstyle\sum_{k=n}^{n+\ell} \|\phi^k(v)\| \overset{(488)}{\leq} \sum_{k=n}^{n+\ell} \|\phi\|_{\mathrm{op}}^k \cdot \|v\| = \|v\| \cdot \|\phi\|_{\mathrm{op}}^n \cdot \sum_{\ell=0}^{n} \|\phi\|_{\mathrm{op}}^\ell \\ &= \|v\| \cdot \|\phi\|_{\mathrm{op}}^n \cdot \tfrac{1-\|\phi\|_{\mathrm{op}}^{n+1}}{1-\|\phi\|_{\mathrm{op}}} \leq \|v\| \cdot \tfrac{\|\phi\|_{\mathrm{op}}^n}{1-\|\phi\|_{\mathrm{op}}}\end{aligned}$$

  mit $\lim_n(\|\phi\|_{\mathrm{op}})^n = 0$ wegen $0 \leq \|\phi\|_{\mathrm{op}} < 1$. Somit ist $(S_n(v))_{n\in\mathbb{N}}$ für jedes $v \in V$ eine Cauchy-Folge, konvergiert also wegen der Vollständigkeit von V gegen ein Element in V.

- Die Abbildung Θ ist linear wegen Proposition 6.1) (Grenzwertsätze für Folgen bzw. Reihen), denn $S_n = \sum_{k=0}^{n} \phi^n$ ist linear für jedes $n \in \mathbb{N}$.

[129] Als Verkettung stetiger linearer Abbildungen ist ϕ^n linear und stetig für jedes $n \in \mathbb{N}$. Somit ist auch $\sum_{k=0}^{n} \phi^n$ als Summe stetiger linearer Abbildung für jedes $n \in \mathbb{N}$ stetig.

- Die Abbildung Θ ist stetig mit $\|\Theta\|_{\mathrm{op}} \leq \frac{1}{1-\|\phi\|_{\mathrm{op}}}$: Mit der Stetigkeit von $\|\cdot\|$ (erster Schritt), Lemma 30 (Erhalt von Ungleichungen unter Limiten im dritten Schritt), und Korollar 22.1) (letzter Schritt) folgt

$$\|\Theta(v)\| = \lim_n \|S_n(v)\| \overset{(488)}{\leq} \lim_n \sum_{k=0}^n \|\phi\|_{\mathrm{op}}^k \cdot \|v\| = \|v\| \cdot \lim_n \sum_{k=0}^n \|\phi\|_{\mathrm{op}}^k = \|v\| \cdot \frac{1}{1-\|\phi\|_{\mathrm{op}}}$$

für alle $v \in V$, also $\|\Theta\|_{\mathrm{op}} \leq \frac{1}{1-\|\phi\|_{\mathrm{op}}}$.

- Es gilt $\Theta \circ \Lambda = \mathrm{id}_V = \Lambda \circ \Theta$ für die stetige lineare Abbildung $\Lambda := \mathrm{id}_V - \phi$. In der Tat erhalten wir für $v \in V$ mit (489) im jeweils letzten Schritt: (ϕ^k ist linear für alle $k \in \mathbb{N}$)

$$\begin{aligned}
(\Theta \circ \Lambda)(v) - \mathrm{id}_V(v) &= \Theta(\Lambda(v)) - \mathrm{id}_V(v) = \lim_n S_n(\Lambda(v)) - \mathrm{id}_V(v) \\
&= \lim_n \sum_{k=0}^n \phi^k\big((\mathrm{id}_V - \phi)(v)\big) - \mathrm{id}_V(v) \\
&= \lim_n \sum_{k=0}^n \big(\phi^k(v) - \phi^{k+1}(v)\big) - \mathrm{id}_V(v) = -\lim_n \phi^{n+1}(v) = 0, \\
(\Lambda \circ \Theta)(v) - \mathrm{id}_V(v) &= \Lambda(\Theta(v)) - \mathrm{id}_V(v) = \lim_n \Lambda(S_n(v)) - \mathrm{id}_V(v) \\
&= \lim_n \sum_{k=0}^n (\mathrm{id}_V - \phi)\big(\phi^k(v)\big) - \mathrm{id}_V(v) \\
&= \lim_n \sum_{k=0}^n \big(\phi^k(v) - \phi^{k+1}(v)\big) - \mathrm{id}_V(v) = -\lim_n \phi^{n+1}(v) = 0.
\end{aligned}$$

Dies zeigt, dass $\mathrm{id}_V - \phi = \Lambda$ invertierbar ist mit stetiger linearer Umkehrabbildung $(\mathrm{id}_V - \phi)^{-1} = \Theta$ (also (487)). Schließlich konvergiert $(S_n)_{n\in\mathbb{N}}$ bezüglich der Operatornorm gegen $(\mathrm{id}_V - \phi)^{-1}$:

Für $v \in V$ und $n \in \mathbb{N}$ erhalten wir mit der Stetigkeit von $\|\cdot\|$ (zweiter Schritt) sowie Korollar 22.1) (letzter Schritt):

$$\begin{aligned}
\|((\mathrm{id}_V - \phi)^{-1} - S_n)(v)\| &= \|\Theta(v) - S_n(v)\| = \lim_k \|S_{n+k}(v) - S_n(v)\| \\
&= \lim_k \|\textstyle\sum_{\ell=n+1}^k \phi^\ell(v)\| \overset{(488)}{\leq} \|v\| \cdot \|\phi\|_{\mathrm{op}}^{n+1} \cdot \lim_k \sum_{\ell=0}^k \|\phi\|_{\mathrm{op}}^\ell \\
&= \|v\| \cdot \|\phi\|_{\mathrm{op}}^{n+1} \cdot \frac{1}{1-\|\phi\|_{\mathrm{op}}};
\end{aligned}$$

also $0 \leq \|(\mathrm{id}_V - \phi)^{-1} - S_n\|_{\mathrm{op}} \leq \|\phi\|_{\mathrm{op}}^{n+1} \cdot \frac{1}{1-\|\phi\|_{\mathrm{op}}}$, mit $\lim_n \|\phi\|_{\mathrm{op}}^{n+1} = 0$ wegen $0 \leq \|\phi\|_{\mathrm{op}} < 1$. Die Behauptung folgt nun aus dem Quetschlemma (Lemma 31). □

16.2.2 Invertierbare lineare Abbildungen

Sei $m, n \in \mathbb{N}_{>0}$. Im Folgenden werden wir den Raum $\mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ häufig mit dem Raum $\mathrm{M}_{n,m}(\mathbb{R})$ der $n \times m$-Matrizen identifizieren; und zwar, wie in Terminologie 34, vermöge Darstellung bezüglich der jeweiligen kanonischen Basis $(e_1, \dots, e_m)$ von $\mathbb{R}^m$ und $(\tilde{e}_1, \dots, \tilde{e}_n)$ von $\mathbb{R}^n$, d.h. (vgl. (383))

$$\begin{aligned}
\Omega[m,n]\colon \quad & \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \ni \phi \mapsto \big(\tilde{\mathrm{pr}}_i\big(\phi(e_j)\big)\big)_{(i,j)\in\mathcal{I}(n,m)} \in \mathrm{M}_{n,m}(\mathbb{R}) \\
\mho[m,n]\colon \quad & \mathrm{M}_{n,m}(\mathbb{R}) \ni (A_{ij})_{(i,j)\in\mathcal{I}(n,m)} \mapsto \big[\phi\colon v \mapsto \textstyle\sum_{i=1}^n \big(\sum_{j=1}^m A_{ij} \cdot v_j\big) \cdot \tilde{e}_i\big] \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)
\end{aligned} \tag{490}$$

mit $\mho[m,n] = \Omega[m,n]^{-1}$. Weiterhin werden wir den Raum $\mathrm{M}_{n,m}(\mathbb{R})$ mit dem Raum $\mathbb{R}^{m\cdot n}$ identifizieren, und zwar vermöge dem Isomorphismus (vgl. (382))

$$\Theta[m,n]\colon \mathrm{M}_{n,m}(\mathbb{R}) \to \mathbb{R}^{m\cdot n}, \qquad \begin{pmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nm} \end{pmatrix} \mapsto (A_{11}, \dots A_{1m}, \dots, A_{n1}, \dots, A_{nm}).$$

Bemerkung 105 (Differenzierbarkeit)**.** *Sei (X, d) ein metrischer Raum. Wir hatten in Bemerkung 89 eingesehen, dass wir durch die Identifizierung* $\mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \cong \mathbb{R}^{m \cdot n}$ *vermöge des Isomorphismus* $\Upsilon[m, n] := \Theta[m, n] \circ \Omega[m, n] \colon \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \to \mathbb{R}^{m \cdot n}$

- *eine stetige Abbildung* $\mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \to X$ *auch als stetige Abbildung* $\mathbb{R}^{m \cdot n} \to X$ *auffassen können.*
- *eine stetige Abbildung* $X \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ *auch als stetige Abbildung* $X \to \mathbb{R}^{m \cdot n}$ *auffassen können.*

Selbige Identifizierung erlaubt es nun auch, in diesem Kontext von (stetige) differenzierbaren bzw. k-mal (stetige) partielle differenzierbaren ($k \in \mathbb{N}$) Abbildungen zu sprechen. Sei genauer $\alpha \colon U \to V$ eine Abbildung mit Mengen U, V. Die folgenden Fälle sind von Relevanz:[130]

- $U \subseteq \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ *ist offen mit $m, n \in \mathbb{N}_{>0}$, und* $V \subseteq \mathrm{Hom}(\mathbb{R}^{m'}, \mathbb{R}^{n'})$ *ist offen mit $m', n' \in \mathbb{N}_{>0}$. Wir definieren:*

$$\widetilde{\alpha} := \Upsilon[m', n'] \circ \alpha \circ \Upsilon[m, n]^{-1} \colon \mathbb{R}^{m \cdot n} \supseteq U' := \Upsilon[m, n](U) \to \Upsilon[m', n'](V) =: V' \subseteq \mathbb{R}^{m' \cdot n'}.$$

- $U \subseteq \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ *ist offen mit $m, n \in \mathbb{N}_{>0}$, und $V \subseteq \mathbb{R}^\ell$ ist offen mit $\ell \in \mathbb{N}_{>0}$. Wir definieren:*

$$\widetilde{\alpha} := \alpha \circ \Upsilon[m, n]^{-1} \colon \mathbb{R}^{m \cdot n} \supseteq U' := \Upsilon[m, n](U) \to V \subseteq \mathbb{R}^\ell.$$

- $U \subseteq \mathbb{R}^\ell$ *ist offen mit $\ell \in \mathbb{N}_{>0}$, und* $V \subseteq \mathrm{Hom}(\mathbb{R}^{m'}, \mathbb{R}^{n'})$ *ist offen mit $m', n' \in \mathbb{N}_{>0}$. Wir definieren:*

$$\widetilde{\alpha} := \Upsilon[m', n'] \circ \alpha \colon \mathbb{R}^\ell \supseteq U \to \Upsilon[m', n'](V) =: V' \subseteq \mathbb{R}^{m' \cdot n'}.$$

In den entsprechenden Fällen heißt α (stetige) differenzierbar *bzw.* k-mal (stetig) partiell differenzierbar *($k \in \mathbb{N}$) genau dann, wenn Entsprechendes für $\widetilde{\alpha}$ gilt.*

Analoge Konventionen gelten, wenn man die auftretenden Homomorphismenräume durch entsprechende Matrizenräume ersetzt: $\mathrm{Hom}(\mathbb{R}^p, \mathbb{R}^q) \to \mathrm{M}_{p,q}(\mathbb{R})$ *sowie* $\Upsilon[p, q] \leftrightarrow \Theta[p, q]$ *mit* $p, q \in \mathbb{N}_{>0}$.

Terminologie 46. *Sei $n \in \mathbb{N}_{>0}$ vorgegeben:*

- *Unter der Identifizierung* (490) *(also vermöge $\Omega[n, n]$) gilt* $\mathrm{End}(\mathbb{R}^n) \cong \mathrm{M}_{n \times n}(\mathbb{R})$*. Ebenso ist die* allgemeine lineare Gruppe *(general linear group)*

$$\mathrm{GL}_n(\mathbb{R}) := \mathsf{Aut}(\mathbb{R}^n) = \{\phi \in \mathrm{End}(\mathbb{R}^n) \mid \phi \text{ ist invertierbar } \}$$

identisch zu den invertierbaren $n \times n$-Matrizen

$$\mathrm{M}^{\mathrm{inv}}_{n \times n}(\mathbb{R}) := \{A \in \mathrm{M}_{n \times n}(\mathbb{R}) \mid \exists!\, A^{-1} \equiv B \in \mathrm{M}_{n \times n}(\mathbb{R})\colon\ B \cdot A = \mathbf{1}_n = A \cdot B\}.$$

Beachte: *Sind $\phi, \psi \in \mathrm{End}(\mathbb{R}^n)$ mit Darstellungsmatrizen $M_\phi = \Omega[n, n](\phi)$ und $M_\psi = \Omega[n, n](\psi)$, so zeigt Bemerkung 95.*a)*, dass die Darstellungmatrix von $\phi \circ \psi$ gegeben ist durch:*

$$M_{\phi \circ \psi} = \Omega[n, n](\phi \circ \psi) = \Omega[n, n](\phi) \cdot \Omega[n, n](\psi) = M_\phi \cdot M_\psi. \qquad (\star)$$

Insbesondere gilt: $\Omega[n, n](\chi^{-1}) = \Omega[n, n](\chi)^{-1}$ *für alle* $\chi \in \mathrm{GL}_n(\mathbb{R})$. (491)

$(\,\Omega[n, n](\chi^{-1}) \cdot \Omega[n, n](\chi) = \Omega[n, n](\chi^{-1} \circ \chi) = \mathbf{1}_n = \Omega[n, n](\chi \circ \chi{-}1) = \Omega[n, n](\chi) \cdot \Omega[n, n](\chi^{-1}))$

[130] Die auftretenden Mengen U', V' sind offen, denn die $\Upsilon[m, n], \Upsilon[m, n]^{-1}, \Upsilon[m', n'], \Upsilon[m', n']^{-1}$ sind stetig (da linear); und für eine Bijektion $\eta\colon X \to Y$ zwischen Mengen X, Y, gilt ja $\eta(W) = (\eta^{-1})^{-1}(W)$ für jedes $W \subseteq X$.

- *Die* Determinantenfunktion *ist gegeben durch*

$$\det\colon \mathbb{R}^{n^2} \overset{\Theta[n,n]}{\cong} \mathrm{M}_{n,n}(\mathbb{R}) \to \mathbb{R}, \qquad (A_{ij})_{1\le i,j\le n} \mapsto \textstyle\sum_{\sigma\in \mathsf{S}_n} \operatorname{sgn}(\sigma)\cdot A_{1,\sigma(1)}\cdots A_{n,\sigma(n)},$$

wobei S_n *die Menge aller Permutationen von* $\{1,\dots,n\}$ *(alle Bijektionen* $\{1,\dots,n\}\to\{1,\dots,n\}$*) bezeichnet. Somit ist* $\widetilde{\det} = \det\circ\Theta[n,n]^{-1}$ *eine Polynomfunktion, und somit glatt (stetig).*

$*$Bemerkung: *Die Determinantenfunktion ist vermöge* $\det\circ\Omega[n,n]\colon \mathrm{End}(\mathbb{R}^n)\to\mathbb{R}$ *natürlich auch auf den Endomorphismen definiert. Wir wollen an dieser Stelle kurz anmerken, dass der Wert* $(\det\circ\Omega[n,n])(\phi)$ *wegen der Multiplikativität der Determinantenfunktion nicht von der expliziten Wahl des Isomorphismus* $\Omega[n,n]$ *abhängt. Sind genauer* $v_1,\dots,v_m$ *und* $w_1,\dots,w_n$ *zwei vorgegebene Basen von* $\mathbb{R}^m$ *und* $\mathbb{R}^n$*, so kann man ein* $\phi\in\mathrm{Hom}(\mathbb{R}^m,\mathbb{R}^n)$ *ja immer auch bezüglich dieser fixierten Basen als* $n\times m$*-Matrix darstellen. Ist nun* $m=n$ *(die obigen zwei Basen dürfen weiterhin auch unterschiedlich sein), und* $\hat\Omega[n,n]\colon \mathrm{End}(\mathbb{R}^n)\to\mathrm{M}_{n,n}(\mathbb{R})$ *die zugehörige Darstellungsabbildung, so gilt* $\det\circ\Omega[n,n]=\det\circ\hat\Omega[n,n]$ *wegen der Multiplikativität der Determinantenfunktion, und weil* $\hat\Omega[n,n]=C\cdot\Omega[n,n]\cdot C^{-1}$ *für eine entsprechende Basiswechselmatrix* $C\in\mathrm{GL}_n(\mathbb{R})$ *gilt:*

$$\begin{aligned} & \det(A\cdot B)=\det(A)\cdot\det(B) && \forall\, A,B\in\mathrm{M}_{n,n}(\mathbb{R})\\ \Longrightarrow\quad & \det(C^{-1})=\det(C)^{-1} && \forall\, C\in\mathrm{GL}_n(\mathbb{R})\\ \Longrightarrow\quad & \det(C^{-1}\cdot A\cdot C)=\det(A) && \forall\, A\in\mathrm{M}_{n,n}(\mathbb{R}),\ C\in\mathrm{GL}_n(\mathbb{R}). \end{aligned}$$

Bemerkung 106. *Aus der Linearen Algebra kennen Sie die folgenden Charakterisierungen der Automorphismengruppe* $\mathrm{GL}_n(\mathbb{R})\cong\mathrm{M}^{\mathrm{inv}}_{n\times n}(\mathbb{R})$*. Sei* $\phi\in\mathrm{End}(\mathbb{R}^n)$ *und* $A:=\Omega[n,n](\phi)\in\mathrm{M}_{n,n}(\mathbb{R})$*. Dann sind äquivalent:*

a) $\phi\in\mathrm{GL}_n(\mathbb{R})$ *(ϕ ist bijektiv), hat also ein inverses.* *($A\in\mathrm{M}^{\mathrm{inv}}_{n\times n}(\mathbb{R})$)*

b) $\det(A)\neq 0$.

c) ϕ *ist surjektiv, hat also ein Rechtsinverses.* *($\exists\, B_R\in\mathrm{M}_{n,n}(\mathbb{R})\colon\ A\cdot B_R=\mathbf{1}_n$)*

d) ϕ *ist injektiv, hat also ein Linksinverses.* *($\exists\, B_L\in\mathrm{M}_{n,n}(\mathbb{R})\colon\ B_L\cdot A=\mathbf{1}_n$)*

Proposition 25. *Sei* $n\in\mathbb{N}_{>0}$*. Dann gelten die folgenden Aussagen:*[131]

1) $\mathrm{GL}_n(\mathbb{R})$ *ist eine offene Teilmenge von* $\mathrm{End}(\mathbb{R}^n)$*.* *($U:=\Upsilon[n,n](\mathrm{GL}_n(\mathbb{R}))\subseteq\mathbb{R}^{n^2}$ offen)*

2) *Die Inversion* $\bullet^{-1}\colon\mathrm{GL}_n(\mathbb{R})\to\mathrm{GL}_n(\mathbb{R})$ *ist glatt.* *($\widetilde{\bullet^{-1}}\colon\mathbb{R}^{n^2}\supseteq U\to U\subseteq\mathbb{R}^{n^2}$ ist glatt)*

Beweis. 1) Wegen Bemerkung 106.b) gilt

$$\mathrm{GL}_n(\mathbb{R})=(\underbrace{\det\circ\Omega[n,n]}_{=:\,\Lambda})^{-1}(\mathbb{R}_\times)\quad\text{mit}\quad \Lambda=\underbrace{\det\circ\Theta[n,n]^{-1}}_{=\,\widetilde{\det}}\circ\Upsilon[n,n]\colon\mathrm{End}(\mathbb{R}^n)\to\mathbb{R}\quad\text{stetig.}$$

Da $\mathbb{R}_\times\subseteq\mathbb{R}$ offen ist, folgt hiermit die Behauptung aus Proposition 9.2).

2) Sei $A\in\mathrm{M}^{\mathrm{inv}}_{n\times n}(\mathbb{R})$ invertierbar, d.h. $\det(A)\neq 0$. Die Cramersche Regel (Lineare Algebra) besagt

$$\left(A^{-1}\right)_{ij}=\frac{(-1)^{i+j}}{\det(A)}\cdot\det\left(A[ji]\right)\qquad\forall\,1\le i,j\le n,\tag{492}$$

[131] Gemäß Korollar 54 sind $\Upsilon[n,n],\Upsilon[n,n]^{-1}$ als lineare Abbildung zwischen endlichdimensionalen $\mathbb{R}$-Vektorräumen stetig; bilden also wegen Proposition 9.2) offene Teilmengen auf offene Teilmengen ab. Beispielsweise ist ja das Bild einer Teilmenge unter $\Upsilon[n,n]$ gerade das Urbild selbiger Teilmenge unter der stetigen Umkehrabbildung $\Upsilon[n,n]^{-1}$.

wobei die Matrix $A[ji]$ durch Streichen der j-ten Zeile und der i-ten Spalte von A entsteht. Hiermit sieht man, dass jede Komponentenfunktionen von $\widetilde{\bullet^{-1}}$ gegeben ist als ein Quotient zweier Polynomfunktionen, also glatt wegen Übung 161. Wegen Lemma 84.1) ist somit $\widetilde{\bullet^{-1}}$ glatt. Genauer errechnet man für $x \in U \subseteq \mathbb{R}^{n^2}$:[132]

$$\begin{aligned}\big(\Theta[n,n]^{-1} \circ \widetilde{\bullet^{-1}}\big)(x)_{ij} &= \big(\Omega[n,n] \circ \bullet^{-1} \circ \Upsilon[n,n]^{-1}\big)(x)_{ij} \overset{(491)}{=} \Big(\Omega[n,n]\big(\Upsilon[n,n]^{-1}(x)\big)^{-1}\Big)_{ij} \\ &= \Big(\big(\Theta[n,n]^{-1}(x)\big)^{-1}\Big)_{ij} \overset{(492)}{=} \frac{(-1)^{i+j}}{\det(\Theta[n,n]^{-1}(x))} \cdot \det\big(\Theta^{-1}[n,n](x)[ji]\big) \\ &= \frac{(-1)^{i+j}}{\widetilde{\det}(x)} \cdot \det\big(\Theta^{-1}[n,n](x)[ji]\big). \qquad \square\end{aligned}$$

16.3 Der Satz über die Umkehrfunktion

In diesem Abschnitt untersuchen wir die Differbarkeitseigenschaften von Umkehrabbildung differenzierbarer Abbildungen. Wir beginnen mit den folgenden grundlegenden Begrifflichkeiten:

Definition 77 (Homöomorphismus)**.** *Seien (X_1, d_1) und (X_2, d_2) metrische Räume, sowie $Y_1 \subseteq X_1$ und $Y_2 \subseteq X_2$ Teilmengen.*

- *Eine Abbildung $f\colon X_1 \to X_2$ heißt* Homöomorphismus*, wenn f stetig und bijektiv ist, und die Umkehrabbildung $f^{-1}\colon X_2 \to X_1$ ebenfalls stetig ist.*
- *Eine Abbildung $g\colon Y_1 \to Y_2$ heißt* Homöomorphismus*, wenn g aufgefasst als Abbildung zwischen den metrischen Unterräumen $(Y_1, (\mathrm{d}_1)_{Y_1})$ und $(Y_2, (\mathrm{d}_2)_{Y_2})$ ein Homöomorphismus ist.*

Definition 78 (Diffeomorphismus)**.** *Seien $n \in \mathbb{N}_{>0}$, $k \in \mathbb{N} \cup \{\infty\}$, und $U, V \subseteq \mathbb{R}^n$ offen. Eine Bijektion $f\colon U \to V$ heißt* C^k-Diffeomorphismus*, wenn f, f^{-1} beides C^k-Abbildungen sind.*[133]

Beachte: *Ein C^k-Diffeomorphismus ist insbesondere ein Homöomorphismus.*

Das folgende Korollar zu Bemerkung 96.2) klärt, warum man bei „in beide Richtungen differenzierbaren" Abbildungen $f\colon \mathbb{R}^m \supseteq U \to V \subseteq \mathbb{R}^n$ (bspw. bei Diffeomorphismen) automatisch immer $m = n$ annehmen muss.

Korollar 64. *Seien $m, n \in \mathbb{N}_{>0}$, sowie $U \subseteq \mathbb{R}^m$ und $V \subseteq \mathbb{R}^n$ offen. Sei $f\colon U \to V$ bijektiv und differenzierbar in $p \in U$. Ist f^{-1} differenzierbar in $f(p)$, so gilt:*

$$m = n \qquad \text{mit} \qquad \mathrm{d}_p f \in \mathrm{GL}_n(\mathbb{R}) \qquad \text{mit} \qquad \mathrm{d}_{f(p)} f^{-1} = (\mathrm{d}_p f)^{-1}. \tag{493}$$

Beweis. Sei $g := f$ und $h := f^{-1}$. Dann gilt $g \circ h = \mathrm{id}_V$ und $h \circ g = \mathrm{id}_U$. Bemerkung 96.2) (Kettenregel) liefert für $q = f(p)$ die folgenden beide Aussagen:

$$\begin{aligned}\mathrm{id}_{\mathbb{R}^n} &= \mathrm{d}_q \mathrm{id}_V = \mathrm{d}_{h(q)} g \circ \mathrm{d}_q h = \mathrm{d}_p f \circ \mathrm{d}_{f(p)} f^{-1} \qquad \text{mit} \qquad n \leq m \\ \mathrm{id}_{\mathbb{R}^m} &= \mathrm{d}_p \mathrm{id}_U = \mathrm{d}_{g(p)} h \circ \mathrm{d}_p g = \mathrm{d}_{f(p)} f^{-1} \circ \mathrm{d}_p f \qquad \text{mit} \qquad m \leq n,\end{aligned}$$

womit (493) unmittelbar folgt. $\square$

Bemerkung* 19 (Jordan-Brouwerscher-Zerlegungssatz)**.** *Ergänzenden zu Korollar 64 wollen wir an dieser Stelle noch die folgenden beiden Resultate anführen, die man aus dem sogenannten Jordan-Brouwerschen-Zerlegungssatz erhält:*[134]

[132] Beachte: Die Komponentenfunktionen von $\widetilde{\bullet^{-1}}$ sind gegeben durch entsprechende die Matrixeinträge von $\Theta[n,n]^{-1} \circ \widetilde{\bullet^{-1}}$.

[133] Formal korrekter wäre es zu fordern, dass f und $(f|^V)^{-1}$ beides C^k-Abbildungen sind, wobei beide Abbildungen als Abbildungen nach $\mathbb{R}^n$ aufzufassen wären.

[134] Für den Jordan-Brouwerschen-Zerlegungssatz siehe 14.19 in [5]. Siehe weiterhin 14.13.d) in [5] für a); sowie 14.21 in [5] für b).

a) Satz von der Invarianz der Dimension: *Seien $U \subseteq \mathbb{R}^m$ und $V \subseteq \mathbb{R}^n$ mit $m, n \in \mathbb{N}_{>0}$ nichtleer und offen. Existiert ein Homöomorphismus $f\colon U \to V$, so gilt notwendig $m = n$.*

   Beachte: *Die Offenheit beider Teilmengen ist hier eine essentielle Voraussetzung; denn beispielsweise ist $f\colon U := \mathbb{R} \ni x \mapsto (x, 0) \in \mathbb{R} \times \{0\} =: V$ ein Homöomorphismus zwischen der in $\mathbb{R}$ offenen Teilmenge $U = \mathbb{R}$, sowie der in $\mathbb{R}^2$ abgeschlossenen Teilmenge $V = \mathbb{R} \times \{0\}$.*[135]

   Beachte: *Ist $f\colon \mathbb{R}^m \supseteq U \to V \subseteq \mathbb{R}^n$ mit $m, n \in \mathbb{N}_{>0}$ ein Diffeomorphismus (f bijektiv und C^k mit f^{-1} ebenfalls C^k), so sind U, V gemäß unseren Definitionen notwendig offen. Da f ein Homöomorphismus ist, erzwingt obiger Satz (in Übereinstimmung mit Korollar 64) $m = n$.*

b) Satz von der Invarianz offener Mengen: *Sei $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^n$ injektiv und stetig. Dann ist $V := f(U) \subseteq \mathbb{R}^n$ ebenfalls offen, und $f|^V\colon U \to V$ ist ein Homöomorphismus.*

   Beachte: *Ist $f\colon \mathbb{R}^n \supseteq A \to B \subseteq \mathbb{R}^n$ mit $n \in \mathbb{N}_{>0}$ ein Homöomorphismus, so zeigt obiger Satz insbesondere: A offen $\Leftrightarrow$ B offen.*

   Beweis obigen Satzes für den Fall $n = 1$ mit einem offenen Intervall $I \equiv U \subseteq \mathbb{R}$: Wegen Satz 37 ist f streng monoton, $J := V = f(U)$ ein Intervall, und $(f|^V)^{-1}$ stetig (also $f|^V$ ein Homöomorphismus). Für die Offenheit von J sei f streng monoton wachsend (den anderen Fall behandelt man analog). Für $y \in J$, sei $x \in I$ mit $f(x) = y$. Wegen der Offenheit von I existiert ein $\varepsilon > 0$ mit $(x - 2\varepsilon, x + 2\varepsilon) \subseteq I$; also $f(x - \varepsilon) < f(x) < f(x + \varepsilon)$, da f streng monoton wachsend. Die Intervalleigenschaft von J impliziert $y = f(x) \in (f(x - \varepsilon), f(x + \varepsilon)) \subseteq J$. $\square$

16.3.1 Differenzierbarkeit der Umkehrfunktion

Satz 98 (Differenzierbarkeit der Umkehrfunktion)**.** *Sei $n \in \mathbb{N}_{>0}$, $U, V \subseteq \mathbb{R}^n$ offen, und $f\colon U \to V$ ein Homöomorphismus. Ist f differenzierbar in $p \in U$ mit $\mathrm{d}_p f \in \mathrm{GL}_n(\mathbb{R})$, so ist f^{-1} differenzierbar in $f(p)$ mit $\mathrm{d}_{f(p)} f^{-1} = (\mathrm{d}_p f)^{-1}$.*[136]

Beachte: Ist $n \equiv 1$ und $U \equiv I \subseteq \mathbb{R}$ ein offenes Intervall, so ist wegen Satz 37 (Stetigkeit der Umkehrfunktion) die Homöomorphismeneigenschaft von f gleichbedeutend mit der Stetigkeit und Bijektivität von f (d.h. f ist stetig und injektiv mit $f(I) = V$); und $\mathrm{d}_p f \in \mathrm{GL}_n(\mathbb{R})$ bedeutet $f'(p) \neq 0$. Satz 98 stellt somit eine direkte Verallgemeinerung von Satz 42 (Differenzierbarkeit der Umkehrfunktion) dar.

Beweis. Sei $g := f^{-1}$. Da f in p differenzierbar ist, existiert gemäß Proposition 22 eine Abbildung $\Phi\colon U - p \to \mathrm{End}(\mathbb{R}^n)$ mit

$$f(p+h) - f(p) = \Phi(h)(h) \quad \forall\, h \in U - p, \qquad \Phi \text{ stetig in } 0, \qquad \Phi(0) = \mathrm{d}_p f.$$

- Per Annahme gilt $\Phi(0) = \mathrm{d}_p f \in \mathrm{GL}_n(\mathbb{R})$, wobei $\mathrm{GL}_n(\mathbb{R}) \subseteq \mathrm{End}(\mathbb{R}^n)$ wegen Proposition 25.1) offen ist. Da Φ stetig in 0 ist, existiert somit ein $\delta > 0$ mit $\mathsf{B}_\delta(0) \subseteq U - p$, sodass $\Phi(\mathsf{B}_\delta(0)) \subseteq \mathrm{GL}_n(\mathbb{R})$ gilt. Wir erhalten für $h \in \mathsf{B}_\delta(0)$:

$$g(f(p+h)) - g(f(p)) = p + h - p = h = \Phi(h)^{-1}(\Phi(h)(h)) = \Phi(h)^{-1}(f(p+h) - f(p)). \qquad (494)$$

- Wegen der Stetigkeit von g ist

$$W := f(\mathsf{B}_\delta(p)) = g^{-1}(\mathsf{B}_\delta(p)) \subseteq V$$

[135] Die Umkehrabbildung von f ist die Einschränkung $\mathrm{pr}_1|_{\mathbb{R}\times\{0\}}$ der (stetigen) Projektion auf den ersten Faktor, also ebenfalls stetig wegen Lemma 40.2) .

[136] Die Kernaussage dieses Satzes ist, dass f^{-1} in $f(p)$ differenzierbar ist. Die explizite Formel $\mathrm{d}_{f(p)} f^{-1} = (\mathrm{d}_p f)^{-1}$ folgt dann ja auch aus (493) in Korollar 64 (vgl. Bemerkung 57 für den eindimensionalen Fall).

offen in V, also wegen der Offenheit von V auch offen in $\mathbb{R}^n$ (vgl. (140) in Bemerkung 36.a)). Wegen der Stetigkeit von g, sowie $\Phi(\mathsf{B}_\delta(0)) \subseteq \mathrm{GL}_n(\mathbb{R})$ und der Stetigkeit der Inversion $\bullet^{-1}\colon \mathrm{GL}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ (Proposition 25.2)) ist

$$\Psi\colon W - f(p) \to \mathrm{GL}_n(\mathbb{R}) \subseteq \mathrm{End}(\mathbb{R}^n), \qquad \tilde{h} \mapsto \Phi\big(\overbrace{g(\underbrace{f(p)+\tilde{h}}_{\in W})}^{\in \mathsf{B}_\delta(p)} - p\big)^{-1} = (\bullet^{-1} \circ \Phi)\big(g(f(p)+\tilde{h}) - p\big)$$

stetig in 0. Für $\tilde{h} \in W - f(p)$ und $h := g(f(p)+\tilde{h}) - p \in \mathsf{B}_\delta(0)$, gilt dann $f(p)+\tilde{h} = f(p+h)$; und wir erhalten

$$\begin{aligned} g|_W(f(p)+\tilde{h}) - g|_W(f(p)) &= g(f(p+h)) - g(f(p)) \overset{(494)}{=} \Phi(h)^{-1}\big(\overbrace{f(p+h) - f(p)}^{=\tilde{h}}\big) \\ &= \Phi\big(\underbrace{g(f(p)+\tilde{h}) - p}_{=h}\big)^{-1}(\tilde{h}) = \Psi(\tilde{h})(\tilde{h}). \end{aligned}$$

Proposition 22 zeigt somit, dass $g|_W$ differenzierbar in $f(p)$ ist, mit

$$\mathrm{d}_{f(p)}(g|_W) = \Psi(0) = \Phi(\overbrace{g(f(p)) - p}^{=p-p=0})^{-1} = \Phi(0)^{-1} = (\mathrm{d}_p f)^{-1}.$$

Da $W \subseteq \mathbb{R}^n$ offen ist mit $W \subseteq V$, folgt die Behauptung nun aus Terminologie 41.(1).c) (Lokalität von Differenzierbarkeit). □

Wir erinnern an Lemma 85, welches besagt, dass eine Abbildung $f\colon \mathbb{R}^m \supseteq U \to \mathbb{R}^n$ für $k \in \mathbb{N}_{>0}$ genau dann eine C^k-Abbildung ist, wenn f differenzierbar und $\widetilde{\mathrm{d}f}\colon U \to \mathbb{R}^{m\cdot n}$ eine C^{k-1}-Abbildung ist. Zusammen mit Satz 98 erhalten wir die folgende Aussage:

Korollar 65. *Seien $n, k \in \mathbb{N}_{>0}$, $U, V \subseteq \mathbb{R}^n$ offen, und $f\colon U \to V$ ein Homöomorphismus. Ist f eine C^k-Abbildung mit* $\mathrm{im}(\mathrm{d}f) \subseteq \mathrm{GL}_n(\mathbb{R})$, *so ist f ein C^k-Diffeomorphismus. (Insbesondere ist f^{-1} glatt falls f glatt ist.)*

Beweis. Sei $g := f^{-1}$. Wegen Satz 98 ist g differenzierbar mit $\mathrm{d}_q g = (\mathrm{d}_{g(q)} f)^{-1}$ für alle $q \in V$, also stetig wegen Lemma 78. Somit ist

$$\mathrm{d}g = (\bullet^{-1} \circ \mathrm{d}f) \circ g\colon V \to \mathrm{GL}_n(\mathbb{R}) \tag{495}$$

als Verkettung stetiger Abbildungen ebenfalls stetig (Proposition 25.2)). Hiermit gilt die Aussage bereits für $k = 1$. Sei nun also $k \geq 2$. Wegen Lemma 85 müssen wir dann nur nachweisen, dass $\widetilde{\mathrm{d}g} \in C^{k-1}(V, \mathbb{R}^{n^2})$ gilt. Nun gilt gemäß unseren Konventionen in Bemerkung 105:

$$\widetilde{\mathrm{d}g} = \Upsilon[n,n] \circ \mathrm{d}g \overset{(495)}{=} (\Upsilon[n,n] \;\circ\; \bullet^{-1} \circ \Upsilon[n,n]^{-1}) \circ (\Upsilon[n,n] \circ \mathrm{d}f) \circ g = \underbrace{\widetilde{\bullet^{-1}} \circ \widetilde{\mathrm{d}f}}_{=:\,\alpha} \circ\, g.$$

Wegen Lemma 85 ist $\widetilde{\mathrm{d}f}$ eine C^{k-1} Abbildung, und wegen Proposition 25.2) ist $\widetilde{\bullet^{-1}}$ glatt. Wegen Lemma 88 ist somit α eine C^{k-1}-Abbildung. Die Behauptung folgt nun per Induktion:

- (IA) Wir hatten bereits gezeigt, dass g eine C^1-Abbildung ist.
- (IV) Sei g eine C^ℓ-Abbildung für ein $1 \leq \ell \leq k-1$.
- (IS) $\widetilde{\mathrm{d}g}$ ist als Verkettung der C^ℓ-Abbildungen α und g gemäß Lemma 88 ebenfalls eine C^ℓ-Abbildung. Also ist g wegen Lemma 85 eine $C^{\ell+1}$-Abbildung. □

16.3.2 Lokale Umkehrbarkeit

In Satz 98 (bzw. Korollar 65) hatten wir gesehen, dass die Homöomorphismeneigenschaft zusammen mit der Invertierbarkeit des Differentials bereits die Differenzierbarkeit der Umkehrabbildung impliziert (bzw., dass im C^k-Falle die Umkehrabbildung automatisch ebenfalls eine C^k-Abbildung ist). Man kann sich nun fragen, inwiefern die eher schwer nachprüfbare Homöomorphismenvoraussetzung auch abgemildert werden kann. Die folgende Definition ist hierfür zentral:

Definition 79 (Lokale Umkehrbarkeit)**.** *Sei $n \in \mathbb{N}_{>0}$, $k \in \mathbb{N}_{>0} \cup \{\infty\}$, $U \subseteq \mathbb{R}^n$ offen. Eine C^k-Abbildung $f\colon U \to \mathbb{R}^n$ heißt:*

1) lokal um $p \in U$ umkehrbar, *wenn eine offene Umgebungen W von p mit $W \subseteq U$ existiert, sodass $V := f(W) \subseteq \mathbb{R}^n$ offen ist und*
$$f_W := (f|_W)|^V \colon W \to V$$
ein C^k-Diffeomorphismus. Es heißt $(f_W)^{-1}\colon V \to W$ lokale Umkehrfunktion von f um p.

   Beachte: *Ist f lokal um p umkehrbar, so ist $f(U)$ automatisch eine Umgebung von $f(p)$; denn es gilt ja $f(p) \in V \subseteq f(U)$ (und V ist offen in $\mathbb{R}^n$).*

2) lokal umkehrbar, *wenn f um jeden Punkt $p \in U$ lokal umkehrbar ist.*

   Beachte: *Ist f lokal umkehrbar, so ist $f(U)$ wegen Teil* 1) *automatisch offen; denn für jedes vorgegebene $y \in f(U)$, existiert ja ein $p \in U$ mit $y = f(p)$.*

Das Kernresultat dieses Abschnittes ist der folgende Satz:

Satz 99 (Lokale Umkehrbarkeit um einem Punkt)**.** *Sei $n \in \mathbb{N}_{>0}$, $U \subseteq \mathbb{R}^n$ offen, $p \in U$ und $f\colon U \to \mathbb{R}^n$ eine C^k-Abbildung mit $k \in \mathbb{N}_{>0} \cup \{\infty\}$. Es ist f genau dann lokal um p umkehrbar, wenn $\mathrm{d}_p f \in \mathrm{GL}_n(\mathbb{R})$ gilt.*

Beweis der Implikation „⇒“ in Satz 99. Ist f lokal um p umkehrbar, so gilt $\mathrm{d}_p f \in \mathrm{GL}_n(\mathbb{R})$ wegen Korollar 64 und Terminologie 41.(1).c) (Lokalität von Differenzierbarkeit). □

Bevor wir nun die (weitaus schwierigere) andere Implikation „⇐“ in Satz 99 zeigen, wollen wir zunächst noch ein paar Schlussfolgerungen ziehen und Beispiele geben. Zunächst folgt unmittelbar:

Satz 100 (lokale Umkehrbarkeit)**.** *Sei $n \in \mathbb{N}_{>0}$, $U \subseteq \mathbb{R}^n$ offen, und $f\colon U \to \mathbb{R}^n$ eine C^k-Abbildung mit $k \in \mathbb{N}_{>0} \cup \{\infty\}$. Es ist f genau dann lokal umkehrbar, wenn $\mathrm{im}(\mathrm{d}f) \subseteq \mathrm{GL}_n(\mathbb{R})$ gilt.*

Beweis. Folgt unmittelbar aus Satz 99. □

Aus Satz 99 erhalten wir schließlich noch die folgende Aussage, die es sozusagen erlaubt die Diffeomorphismeneigenschaft einer gegebenen C^k-Abbildung einfach anhand ihrer Injektivitätseigenschaft und an der Determinante ihrer Jacobimatrix festzustellen.

Satz 101 (Satz über die Umkehrfunktion)**.** *Sei $n \in \mathbb{N}_{>0}$, $k \in \mathbb{N}_{>0} \cup \{\infty\}$, $U \subseteq \mathbb{R}^n$ offen, und $f\colon U \to \mathbb{R}^n$ eine injektive C^k-Abbildung mit $\mathrm{im}(\mathrm{d}f) \subseteq \mathrm{GL}_n(\mathbb{R})$. Dann ist $V := f(U) \subseteq \mathbb{R}^n$ offen, und $f|^V\colon U \to V$ ist ein C^k-Diffeomorphismus.*

Beweis. • Wegen der Injektivität von f existiert die Umkehrabbildung $(f|^V)^{-1}\colon V \to U \subseteq \mathbb{R}^n$.

- Wegen $V = f(U)$ existiert zu jedem $y \in V$ ein $p \in U$ mit $y = f(p)$.
- Zu jedem $p \in U$ liefert Satz 99 offene Teilmengen $W_p, V_{f(p)} \subseteq \mathbb{R}^n$ mit $p \in W_p \subseteq U$ und $f(p) \in V_{f(p)}$, sodass $f_{W_p} = (f|_{W_p})|^{V_{f(p)}}\colon W_p \to V_{f(p)}$ ein C^k-Diffeomorphismus ist, d.h. $V_{f(p)} \subseteq f(U) = V$.

Folglich ist $V = \bigcup_{p\in U} V_{f(p)}$ offen. Nun ist $(V_{f(p)})_{p\in U}$ eine offene Überdeckung von V mit
$$(f|^V)^{-1}|_{V_{f(p)}} = \underbrace{(f_{W_p})^{-1}}_{C^k\text{-Abbildung}} \qquad \forall\, p \in U,$$
also zeigt Lemma 83.2), dass $(f|^V)^{-1}$ eine C^k-Abbildung ist. □

Beispiel 110. *Seien $n, k \in \mathbb{N}_{>0}$, $U, V \subseteq \mathbb{R}^n$ offen, und $f\colon U \to V$ eine surjektive ($f(U) = V$) C^k-Abbildung mit $\mathrm{im}(\mathrm{d}f) \subseteq \mathrm{GL}_n(\mathbb{R})$.*

a) *Sei $n = 1$ und $I \equiv U$ ein offenes Intervall. Dann ist f bereits injektiv (Zwischenwertsatz); also ist f wegen Satz 101 ein C^k-Diffeomorphismus. Insbesondere ist $J \equiv V$ ein offenes Intervall (f ist stetig; also ist $J = f(I)$ wegen Korollar 28 ein Intervall*[137]*).*

   Beweis. Es ist $\mathrm{im}(\mathrm{d}f) \subseteq \mathrm{GL}_1(\mathbb{R})$ gleichbedeutend mit $\mathrm{im}(f') \subseteq \mathbb{R}_\times$. Wegen der Stetigkeit von f' zeigt der Zwischenwertsatz (Satz 36), dass $f' > 0$ oder $f' < 0$ gilt. Wegen Korollar 32.3) ist f streng monoton; also injektiv wegen Lemma 43. □

   Beachte: *Die Voraussetzung, dass U ein Intervall ist, ist hier essentiell; denn beispielsweise ist die auf $\mathbb{R}_\times$ eingeschränkte Betragsfunktion $f\colon \mathbb{R}_\times \ni x \mapsto |x| \in (0, \infty)$ glatt mit $f(\mathbb{R}_\times) = (0, \infty)$ offen und $f' \neq 0$, aber sicher nicht injektiv.*

b) *Sei $n \geq 2$. Wir betrachten die glatte, und wegen Satz 52 (Polardarstellung) surjektive Abbildung:*

$$f\colon U \equiv (0, \infty) \times \mathbb{R} \to \mathbb{R}^2 \setminus \{(0, 0)\} \equiv V, \qquad (r, \phi) \mapsto (r \cdot \cos(\phi), r \cdot \sin(\phi)).$$

   Es gilt $\mathrm{im}(\mathrm{d}f) \subseteq \mathrm{GL}_2(\mathbb{R})$, denn für $(r, \phi) \in (0, \infty) \times \mathbb{R}$ ergibt sich die Jacobimatrix zu:

$$J_{(r,\phi)}(f) = \begin{pmatrix} \cos(\phi) & -r\sin(\phi) \\ \sin(\phi) & r\cos(\phi) \end{pmatrix} \qquad \Longrightarrow \qquad \det\big(J_{(r,\phi)}(f)\big) = r \neq 0.$$

   - *f ist nicht injektiv (Satz 101 ist nicht anwendbar) wegen*

$$f(r, 2\pi n + \phi) = f(r, \phi) \qquad \forall\, n \in \mathbb{Z},\ (r, \phi) \in (0, \infty) \times \mathbb{R},$$

     kann also auch kein Hömöomorphismus (also auch kein C^∞-Diffeomorphismus) sein.
   - *Wegen Satz 100 ist f aber lokal umkehrbar (beachte, dass V tatsächlich offen ist).*

     *Beispielsweise: *Für $p = (t_0, \phi_0) \in U$ fixiert ist*

$$W_p := (0, \infty) \times (\phi_0 - \pi, \phi_0 + \pi) \subseteq \mathbb{R}^2$$

     offen mit $p \in W_p$. Zudem ist $f_{W_p}\colon W_p \to V_{f(p)} := f(W_p)$ injektiv (Eindeutigkeit der Polardarstellung Satz 52 und Periodizitätseigenschaften von Sinus und Kosinus), wobei

$$V_{f(p)} = \mathbb{R}^2 \setminus \big(\underbrace{[0, \infty) \cdot (\cos(\phi_0 + \pi), \sin(\phi_0 + \pi))}_{=:\,A}\big)$$

     wegen Beispiel 31 offen in $\mathbb{R}^2$ (man macht sich leicht klar, dass das Liniensegment A abgeschlossen in $\mathbb{R}^2$ ist). Somit ist f_{W_p} bijektive und glatt; und es lässt sich zudem zeigen, dass $(f_{W_p})^{-1}$ ebenfalls glatt ist: Für $\phi_0 = \frac{\pi}{2}$ ist beispielsweise die Einschränkung von $(f_{W_p})^{-1}$ auf $O := \{(x, y) \in V \mid y > 0\}$ (Schnitt von $V_{f(p)}$ mit der oberen Halbebene) explizit gegeben durch:

$$O \to (0, \infty) \times (0, \pi), \qquad (x, y) \mapsto \begin{pmatrix} \sqrt{x^2 + y^2} \\ \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) \end{pmatrix}.$$

Bemerkung 107 (Vereinfachung der Voraussetzungen von Satz 99)**.** *Seien $n, k \in \mathbb{N}_{>0}$, $U \subseteq \mathbb{R}^n$ offen, $p \in U$, sowie $f \in C^k(U, \mathbb{R}^n)$ mit $\mathrm{d}_p f \in \mathrm{GL}_n(\mathbb{R})$. Wir betrachten die Abbildung*

$$\tilde{f}\colon U - p =: \tilde{U} \to \mathbb{R}^n, \qquad x \mapsto (\mathrm{d}_p f)^{-1}(f(p + x) - f(p))$$

$$\text{also} \tag{496}$$

$$f = \mathrm{d}_p f(\tilde{f}(\cdot - p)) + f(p).$$

Es gilt $\tilde{f} \in C^k(\tilde{U}, \mathbb{R}^n)$ mit $\tilde{f}(0) = 0$ und $\mathrm{d}_0 \tilde{f} = \mathrm{id}_{\mathbb{R}^n}$; wobei f lokal um p umkehrbar ist, falls $\tilde{f}$ lokal um 0 umkehrbar ist.

[137] Die Offenheit von $f(I)$ folgt alternativ auch aus der strengen Monotonie von f – vgl. Bemerkung* 19.b).

- *Es gilt $\tilde f \in C^k(\tilde U, \mathbb{R}^n)$ mit $\tilde f(0) = 0$ und $\mathrm{d}_0 \tilde f = \mathrm{id}_{\mathbb{R}^n}$.*

  $(*)$*Beweis der Behauptung.* Offensichtlich gilt $\tilde f(0) = 0$. Sei nun $g := (\mathrm{d}_p f)^{-1} \colon \mathbb{R}^n \to \mathbb{R}^n$. Dann gilt $\mathrm{d}g = \mathrm{d}_p f$ wegen Übung 153 (g ist linear), d.h. g ist eine konstante Abbildung. Somit ist für $1 \le j \le n$ auch $\partial_j g = (\mathrm{d}_p f)^{-1}(e_j)$ konstant, woraus umgehend $g \in C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ folgt (alle partiellen Ableitungen der Ordnung ≥ 2 sind konstant 0). Lemma 88 (Verkettungen) zeigt nun $\tilde f = g \circ (f(p + \cdot) - f(p)) \in C^k(\tilde U, \mathbb{R}^n)$ und Satz 88.2) (Kettenregel) liefert:

  $$\mathrm{d}_0 \tilde f = (\mathrm{d}_p f)^{-1} \circ \mathrm{d}_p f = \mathrm{id}_{\mathbb{R}^n}. \qquad \square$$

- *Ist $\tilde f$ lokal um 0 umkehrbar, so ist f lokal um p umkehrbar.*

  $(*)$*Beweis der Behauptung.* Sei $\tilde W \subseteq \mathbb{R}^n$ offen mit $0 \in \tilde W \subseteq \tilde U$ sowie $\tilde V \subseteq \mathbb{R}^n$ offen mit $0 = \tilde f(0) \in \tilde V$, sodass $\tilde f_{\tilde W} \colon \tilde W \to \tilde V$ ein C^k-Diffeomorphismus ist:

  - $V := \mathrm{d}_p f(\tilde V) + f(p) \subseteq \mathbb{R}^n$ und $W := p + \tilde W \subseteq \mathbb{R}^n$ sind offen; denn $\mathbb{R}^n \ni x \mapsto x + p \in \mathbb{R}^n$ und $\mathbb{R}^n \ni x \mapsto \mathrm{d}_p f(x) + f(p) \in \mathbb{R}^n$ sind offensichtlich Homöomorphismen. Wir erhalten

    $$f(W) \overset{(496)}{=} \mathrm{d}_p f(\tilde f(\tilde W)) + f(p) = \mathrm{d}_p f(\tilde V) + f(p) = V,$$

    d.h. $f_W := (f|_W)|^V \colon W \to V$ ist definiert (und eine C^k-Abbildung).
  - Ähnlich wie im vorangegangenen Punkt folgt, dass

    $$g \colon V \to W, \qquad y \mapsto p + (\tilde f_{\tilde W})^{-1}\big((\mathrm{d}_p f)^{-1}(y - f(p))\big)$$

    C^k-Abbildung ist. Zudem folgt $f_W \circ g = \mathrm{id}_V$ sowie $g \circ f_W = \mathrm{id}_W$, also $g = (f_W)^{-1}$:

    $$g(f_W(x)) = \underbrace{g\big(\overbrace{\mathrm{d}_p f(\tilde f(x-p)) + f(p)}^{\overset{(496)}{=} f(x)}\big)}_{p + (\tilde f_{\tilde W})^{-1}(\tilde f(x-p))} = x \qquad \forall\, x \in W = p + \tilde W,$$

    $$f_W(g(y)) = \mathrm{d}_p f(\tilde f(g(y) - p)) + f(p) = y \qquad \forall\, y \in V = f(p) + \mathrm{d}_p f(\tilde V). \qquad \square$$

Wir beenden nun schließlich den Beweis von Satz 99:

Beweis der Implikation „$\Leftarrow$“ in Satz 99. Sei $n \in \mathbb{N}_{>0}$, $U \subseteq \mathbb{R}^n$ offen, $p \in U$ und $f \colon U \to \mathbb{R}^n$ eine C^k-Abbildung für $k \in \mathbb{N}_{>0} \cup \{\infty\}$ mit $\mathrm{d}_p f \in \mathrm{GL}_n(\mathbb{R})$. Wir müssen zeigen, dass f lokal um p umkehrbar ist. Wegen Bemerkung 107 können wir hierfür

$$p = 0, \qquad f(0) = 0, \qquad \mathrm{d}_0 f = \mathrm{id}_{\mathbb{R}^n}$$

annehmen. Sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$.

- Es existiert ein $\delta > 0$ mit

  $$\mathrm{d}f(O) \subseteq \mathrm{GL}_n(\mathbb{R}) \qquad \text{sowie} \qquad \|\mathrm{id}_{\mathbb{R}^n} - \mathrm{d}_z f\|_{\mathrm{op}} \le \tfrac{1}{2} \qquad \forall\, z \in O := \mathsf{B}_{2\delta}(0), \tag{497}$$

  denn $\mathrm{d}f$ ist stetig (in 0) mit $\mathrm{d}_0 f = \mathrm{id}_{\mathbb{R}^n} \in \mathrm{GL}_n(\mathbb{R})$, und $\mathrm{GL}_n(\mathbb{R}) \subseteq \mathrm{End}(\mathbb{R}^n)$ ist offen wegen Proposition 25.1)).

Es reicht nun zu zeigen, dass offene Teilmengen

$$W, V \subseteq \mathbb{R}^n \quad \text{mit} \quad O = \mathsf{B}_{2\delta}(0) \supseteq W \ni 0 \in V \quad \text{und} \quad f(W) = V \quad \text{existieren,}$$

sodass $f_W = (f|_W)^V : W \to V$ bijektiv mit stetiger Umkehrabbildung $(f_W)^{-1} : V \to W$ ist:

In der Tat, wegen der Stetigkeit von f (also auch f_W wegen Lemma 40), ist dann nämlich f_W ein Homöomorphismus (und eine C^k-Abbildung); und wegen $\operatorname{im}(\mathrm{d}f_W) \subseteq \mathrm{d}f(\mathsf{B}_{2\delta}(0)) = \mathrm{d}f(O) \subseteq \mathrm{GL}_n(\mathbb{R})$, zeigt dann Korollar 65, dass f_W ein C^k-diffeomorphismus ist.

Für $y \in \mathbb{R}^n$ definieren wir

$$g_y : O \to \mathbb{R}^n, \qquad x \mapsto x + (y - f(x)).$$

a) Für $x \in U$ gilt:[138] $\qquad \underbrace{g_y(x) = x}_{x \text{ ist Fixpunkt von } g_y} \quad \Leftrightarrow \quad f(x) = y$

(* Die Idee ist nun also geeignete offene Umgebungen $W, v \subseteq O$ von 0 zu finden, sodass $\Gamma : V \ni y \mapsto \Gamma(y) \in W$ definiert und stetig ist; wobei $\Gamma(y)$ einen fixierten (den eindeutigen) Fixpunkt von g_y bezeichnen soll.)

b) g_y ist Lipschitzstetig mit Lipschitzkonstante $\frac{1}{2}$, d.h. (Schrankensatz und (497))

$$\|g_y(x) - g_y(x')\| \leq \tfrac{1}{2} \cdot \|x - x'\| \qquad \forall\, x, x' \in O. \tag{498}$$

Beweis von (498). Es gilt $g \in C^1(O, \mathbb{R}^n)$ wegen

$$\mathrm{d}g_y : O \ni z \mapsto \mathrm{d}_z g_y = \mathrm{id}_{\mathbb{R}^n} - \mathrm{d}_z f \in \mathrm{End}(\mathbb{R}^n) \qquad \text{also} \qquad \|\mathrm{d}g_y\|_{\mathrm{op}} \overset{(497)}{\leq} \tfrac{1}{2}.$$

Für $x, x' \in O$ vorgegeben, gilt nun $x + [a, b] \cdot v \subseteq O$ für $a := 0$, $b := 1$ und $v := x' - x$ mit

$$\|\mathrm{d}_z g_y(v)\| \leq \|\mathrm{d}_z g_y\|_{\mathrm{op}} \cdot \|v\| \leq \tfrac{1}{2} \cdot \|x - x'\| =: L. \tag{499}$$

Somit zeigt Satz 90 (Schrankensatz):

$$\|g_y(x) - g_y(x')\| = \|g_y(x + a \cdot v) - g_y(x + b \cdot v)\| \leq (b - a) \cdot L = \tfrac{1}{2} \cdot \|x - x'\|. \qquad \square$$

Sei nun $X := \overline{\mathsf{B}}_\delta(0) \subseteq O$ und $B := \overline{\mathsf{B}}_{\frac{\delta}{2}}(0)$.

c) Für $y \in B$ hat g_y einen eindeutigen Fixpunkt $\Gamma(y) \in X$ (Beweis unten); d.h.,

$$x \in X,\, y \in B: \qquad f(x) = y \quad \overset{\text{a)}}{\Longleftrightarrow} \quad g_y(x) = x \quad \overset{\text{Eindeutigkeit}}{\Longleftrightarrow} \quad x = \Gamma(y). \tag{500}$$

Hiermit folgt:

i) $\Gamma : B \to X, \quad y \mapsto \Gamma(y)$ ist definiert, (Bezeichnung für Γ: Fixpunktabbildung)

ii) $f \circ \Gamma = \mathrm{id}_B$, $\qquad (y \in B: \ \Gamma(y)$ Fixpkt. von $g_y \overset{\text{a)}}{\Rightarrow} f(\Gamma(y)) = y)$

[138] Man kann sich plausibel machen, dass für x, y nahe bei 0, die iterierte Anwendung von g_y ein Urbild von y unter f liefert – Wir argumentieren nun informell ohne Berücksichtigung von Definitionsbereichen, und benutzen das Symbol $\approx$ im Sinne „ungefähr“: Wegen $\mathrm{d}_0 f = \mathrm{id}_{\mathbb{R}^n}$ und der Stetigkeit von $\mathrm{d}f$ in 0 gilt $\mathrm{d}_x f \approx \mathrm{id}_{\mathbb{R}^n}$ für x nahe bei 0. Weiterhin ist $\mathrm{R}^1_0 f(y - f(x)) \approx 0$ klein für x, y nahe bei 0; denn wegen der Stetigkeit von f in 0 mit $f(0) = 0$, ist dann ja auch $f(x)$ klein. Wir erhalten $f(g_y(x)) = f(x + (y - f(x)) \approx f(x) + \mathrm{d}_x f(y - f(x)) \approx f(x) + \mathrm{id}_{\mathbb{R}^n}(y - f(x)) = y$; wobei wir im ersten Schritt das Taylorsche Restglied erster Ordnung haben wegfallen lassen, und im zweiten Schritt $\mathrm{d}_x f \approx \mathrm{id}_{\mathbb{R}^n}$ eingesetzt haben. Folglich ist nicht unplausibel, dass $f(g_y(x))$ (für x, y klein) bereits näher am Wert y liegt, als dies für $f(x)$ der Fall ist.

iii) $f|_{f^{-1}(B)\cap X}$ ist injektiv wegen (500) (Eindeutigkeit des Fixpunktes).[139]

Beweis der Existenz und Eindeutigkeit des Fixpunktes.

- Für $y \in B = \overline{\mathsf{B}}_{\frac{\delta}{2}}(0)$ gilt $g_y(X) \subseteq X$, d.h.,

  $$\tilde{g}_y := g_y|_X \colon X \to X \qquad \text{ist eine Selbstabbildung;}$$

  denn für $x \in X = \overline{\mathsf{B}}_\delta(0)$ gilt ja: $\quad (g_y(0) = y + (0 - f(0)) = y$, da $f(0) = 0$ mit $0 \in X)$

  $$\|g_y(x)\| = \|g_y(0) + (g_y(x) - g_y(0))\| \leq \|y\| + \|g_y(x) - g_y(0)\| \overset{(498)}{\leq} \frac{\delta}{2} + \frac{1}{2} \cdot \|x\| \leq \delta.$$

  Wegen (498) (beachte $X \subseteq O$), ist $\tilde{g}_y$ zudem kontraktiv mit Kontraktionskonstante $\lambda = \frac{1}{2}$.
- Wegen (498) in Teil b) ist $\tilde{g}_y$ (für $y \in B$) kontraktiv (mit Kontraktionskonstante $\frac{1}{2}$).
- $X = \overline{\mathsf{B}}_\delta(0)$ ist abgeschlossen in $\mathbb{R}^n$, also wegen wegen Korollar 50.2) (Vollständigkeit von $\mathbb{R}^n$) und Lemma 72 ein vollständiger metrischer Raum (mit der eingeschränkten Normmetrik). Der Banachsche Fixpunktsatz (Satz 96) zeigt daher, dass $\tilde{g}_y$ (für $y \in B$) genau einen Fixpunkt (in X) hat. □

Wir betrachten nun die offenen Mengen (f ist stetig)

$$V := \mathsf{B}_{\frac{\delta}{2}}(0) \subseteq B \qquad \text{und} \qquad W := f^{-1}(V) \cap \mathsf{B}_\delta(0) \subseteq f^{-1}(B) \cap X \subseteq X.$$

Dann ist $f|_W \colon W \to V$ injektiv, wegen Punkt c).iii). Wir müssen nun noch zeigen, dass

$$f_W = (f|_W)|^V \colon W \to V$$

surjektiv ist; und dass die somit existente Umkehrabbildung $(f_W)^{-1} \colon V \to W$ stetig ist.[140] Hierfür genügt es nachzuweisen, dass

$$\|x - x'\| \leq 2 \cdot \|f(x) - f(x')\| \qquad \forall\, x, x' \in X \tag{501}$$

gilt; denn hiermit schlussfolgert man:

- f_W ist surjektiv: $\qquad$ (Punkt c).ii) im ersten Schritt; $0 \in X$ und $f(0) = 0$ im zweiten Schritt)

  $y \in V = \mathsf{B}_{\frac{\delta}{2}}(0) \subseteq B$

  $$\begin{aligned}
  &\overset{f\circ\Gamma = \mathrm{id}_B}{\Longrightarrow} \quad f(\Gamma(y)) = y \quad \text{mit} \quad \Gamma(y) \in X \quad \text{also} \quad \Gamma(y) \in f^{-1}(y) \subseteq f^{-1}(V)\\
  &\Longrightarrow \quad \|\Gamma(y)\| = \|\Gamma(y) - 0\| \overset{(501)}{\leq} 2 \cdot \|f(\Gamma(y)) - f(0)\| = 2 \cdot \|f(\Gamma(y))\| = 2 \cdot \|y\| < 2 \cdot \frac{\delta}{2} = \delta\\
  &\Longrightarrow \quad \Gamma(y) \in f^{-1}(V) \cap \mathsf{B}_\delta(0) = W \quad \text{mit} \quad f(\Gamma(y)) = y.
  \end{aligned}$$

  Somit ist f_W bijektiv, d.h., $(f_W)^{-1}$ existiert.
- $(f_W)^{-1}$ ist Lipschitzstetig, also stetig: $\qquad$ (beachte $W \subseteq X$ im ersten Schritt)

  $$\|(f_W)^{-1}(y) - (f_W)^{-1}(y')\| \overset{(501)}{\leq} 2 \cdot \big\|f\big((f_W)^{-1}(y)\big) - f\big((f_W)^{-1}(y')\big)\big\| = 2 \cdot \|y - y'\| \qquad \forall\, y, y' \in V.$$

[139] Bemerkung: Es gilt $M := f^{-1}(B) \cap X = \Gamma(B)$. In der Tat gilt $\Gamma(B) \subseteq M$ wegen c).ii); und für $x \in M$ gilt $x \in X$ und $f(x) =: y \in B$, also $x = \Gamma(y) \in \Gamma(B)$ wegen (500).

[140] Man überlegt sich auch leicht, dass dann notwendig $(f_W)^{-1} = \Gamma|_V$ gilt.

Um schließlich (501) nachzuweisen beachten wir zunächst, dass $\tilde{g}_0 + f|_X = \mathrm{id}_X$ gilt:

$$\tilde{g}_0(z) + f(z) = (0 + (z - f(z))) + f(z) = z \qquad \forall\, z \in X.$$

Für $x, x' \in X \subseteq O$ erhalten wir

$$\begin{aligned} \|x - x'\| &= \|\big(\tilde{g}_0(x) + f(x)\big) - \big(\tilde{g}_0(x') + f(x')\big)\| \\ &\leq \|\tilde{g}_0(x) - \tilde{g}_0(x')\| + \|f(x) - f(x')\| \\ &\overset{(498)}{\leq} \tfrac{1}{2} \cdot \|x - x'\| + \|f(x) - f(x')\|, \end{aligned}$$

und Umstellen liefert (501). □

Bemerkung* 20. *Mit leichter Modifikation zeigt der Beweis von Satz 99 eigentlich sogar die folgende tollere Aussage:*

Satz* 1. *Sei $n \in \mathbb{N}_{>0}$, $U \subseteq \mathbb{R}^n$ offen, $p \in U$ und $f\colon U \to \mathbb{R}^n$ eine differenzierbare Abbildung mit $\mathrm{d}_p f \in \mathrm{GL}_n(\mathbb{R})$, sodass $\mathrm{d}_p f$ stetig in p ist. Dann existieren offene Teilmengen $W, V \subseteq \mathbb{R}^n$ mit $p \in W \subseteq U$ und $f(p) \in V \subseteq \mathbb{R}^n$, sodass $f_W = (f|_W)|^V \colon W \to V$ ein Homöomorhismus ist.*

- *Satz 99 folgt dann im C^k-Falle auch sofort aus Satz* 1 und Korollar 65.*
- *Mit Satz* 1 lässt sich beispielsweise auch die Homöomorphismenvoraussetzung in Satz 98 austauschen durch die Forderung, dass f differenzierbar ist mit $\mathrm{d}_p f \in \mathrm{GL}_n(\mathbb{R})$ sowie $\mathrm{d}_p f$ stetig in p. Genauer gilt dann ja wegen Satz* 1 die Voraussetzung von Satz 98 lokal um p; wenn man also U und V durch entsprechend kleinere offene Umgebungen $\tilde{U}$ von p und $\tilde{V}$ von $f(p)$ ersetzt, sowie f durch $\tilde{f} := (f|_{\tilde{U}})|^{\tilde{V}}$.*
- *Im Falle $n = 1$ lässt sich Satz* 1 wie folgt formulieren und beweisen:*

  Satz* 2. *Sei $U \subseteq \mathbb{R}$ offen, $p \in U$, und $f\colon U \to \mathbb{R}$ differenzierbar mit $f'(p) \neq 0$, sodass f' stetig in p ist. Dann existiert ein offenes Intervall $I \subseteq U$ mit $p \in I$, sodass $J := f(I)$ ein offenes Intervall ist, und $(f|_I)|^J \colon I \to J$ ein Homöomorphismus.*

  Beweis. Sei o.B.d.A. $f'(p) > 0$ (der Fall $f'(p) < 0$ folgt analog). Wegen der Stetigkeit von f' in p (und der Offenheit von U), finden wir ein offenes Intervall $I \subseteq U$ mit $p \in I$, sodass $f'|_I > 0$ gilt. Für $\tilde{f} := f|_I$ gilt dann $\tilde{f}' = f'|_I$ wegen Korollar 31; und weiterhin ist $\tilde{f}$ stetig wegen Lemma 40.2). Wegen Korollar 32.3) ist $\tilde{f}$ streng monoton wachsend; also injektiv wegen Lemma 43. Wie in Bemerkung* 19.b) folgt nun, dass $J = f(I) = \tilde{f}(I)$ ein offenes Intervall ist, sowie $(f|_I)|^J = \tilde{f}|^J \colon I \to J$ ein Homöomorphismus. □

Beweis(skizze) von Satz 1.* Wie in Bemerkung 107 sieht man, dass man sich wieder nur um den Fall $p = 0$, $f(0) = 0$, $\mathrm{d}_0 f = \mathrm{id}_{\mathbb{R}^n}$ zu kümmern braucht. Die gleiche Argumentation wie im Beweis von Satz 99 (ohne die Differenzierbarkeitsaussagen) zeigt dann die Behauptung, denn:

- Für (497) wurde nur die Stetigkeit von $\mathrm{d}f$ in 0 benutzt.
- Für (498) wurde (497) sowie der Schrankensatz (Satz 90) benutzt, welcher die Stetigkeit von $\mathrm{d}f$ voraussetzt. Allerdings folgt mit Hilfe von Lemma* 1 in Bemerkung* 18, aus (497) ebenfalls (498): Für x, x', a, b, v und $L \geq 0$ wie in b) (Beweis von Satz 99), sei $\gamma\colon (-\epsilon, 1+\epsilon) \ni t \mapsto g_y(x + t\cdot v) \in O$ (mit $\epsilon > 0$ passend gewählt). Wegen Bemerkung 97 (und (419)) ist γ differenzierbar mit $\gamma'\colon (-\epsilon, 1+\epsilon) \ni t \mapsto D_v g_y(x + t \cdot v) = \mathrm{d}_{x+t\cdot v} g_y(v) \in \mathbb{R}^n$. Wegen (499) gilt dann für jedes $\varepsilon > 0$:

  $$\mathrm{im}(\gamma') \subseteq \mathsf{B}_{L+\varepsilon}(0) \quad \overset{(441)}{\Longrightarrow} \quad \gamma(b) - \gamma(a) = \frac{\gamma(b) - \gamma(a)}{b - a} \in \mathsf{B}_{L+\varepsilon}(0) \quad \Longrightarrow \quad \|g_y(x) - g_y(x')\| < L + \varepsilon,$$

  also $\|g_y(x) - g_y(x')\| \leq L = \frac{1}{2} \cdot \|x - x'\|$. □

16.4 Der Satz über Implizite Funktionen

In diesem Abschnitt wollen wir noch eine wichtige Folgerung aus Satz 99 (lokale Umkehrbarkeit um einem Punkt) behandeln, den sogenannten Satz über die impliziten Funktionen. Prinzipiell geht es darum, Niveaumengen

$$\Phi^{-1}(u) \subseteq \underbrace{\mathbb{R}^m \times \mathbb{R}^n}_{\cong\, \mathbb{R}^{m+n}} \qquad \text{mit} \qquad m, n \in \mathbb{N}_{>0} \qquad \text{und} \qquad u \in \operatorname{im}(\Phi)$$

einer C^k-Abbildung mit $k \in \mathbb{N}_{>0}$[141]

$$\Phi \colon \mathbb{R}^m \times \mathbb{R}^n \supseteq U \times V \to \mathbb{R}^n \quad \text{mit} \quad U \subseteq \mathbb{R}^m \quad \text{sowie} \quad V \subseteq \mathbb{R}^n \quad \text{offen} \tag{502}$$

lokal um gewisse Punkte in $(x, y) \in \Phi^{-1}(u)$ durch m Veränderliche (Koordinaten) zu parametrisieren. Eine alternative Anschauung ist, dass man für ein „$\tilde{x}$ nahe bei x" ein eindeutiges „$\tilde{y}$ nahe bei y" finden will, sodass ebenfalls $(\tilde{x}, \tilde{y}) \in \Phi^{-1}(u)$ gilt. (Die Zuordnung $\eta \colon \tilde{x} \mapsto \tilde{y}$ nennt sich implizite Funktion, da sie implizit (indirekt) durch Φ festgelegt wird; und die zugehörige „Parametrisierung" (injektiv) ist dann gegeben durch $\Psi \colon \tilde{x} \mapsto (\tilde{x}, \eta(\tilde{x}))$ – Details siehe Satz 102.)

Notation 35. *Wir identifizieren wieder $\mathbb{R}^{m+n} \cong \mathbb{R}^m \times \mathbb{R}^n$. Wir schreiben im Folgenden aber eher $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ mit $x \in \mathbb{R}^m$ und $y \in \mathbb{R}^n$ anstelle $z \in \mathbb{R}^{m+n}$, da dann die Anordnung der Koordinaten auch unmittelbar klar ist.*

Bemerkung 108. *In Beispiel 111 werden wir die Folgenden Situationen diskutieren:*

a) *Sei $m, n = 1$, $u = 1$, sowie*

$$\Phi \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \qquad (x, y) \mapsto x^2 + y^2.$$

$\Phi^{-1}(u) \subseteq \mathbb{R}^{m+n} = \mathbb{R}^2$ ist der Einheitskreis, welcher lokal „am Nordpol" $(0, 1)$ durch eine offene Teilmenge des $\mathbb{R}^m = \mathbb{R}$ parametrisiert werden soll („1-dimensionales Objekt" im $\mathbb{R}^2$).

b) **(∗)** *Sei $m = 2$, $n = 1$, $u = 1$, sowie*

$$\Phi \colon \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}, \qquad ((x, y), z) \mapsto x^2 + y^2 + z^2.$$

$\Phi^{-1}(u) \subseteq \mathbb{R}^{m+n} = \mathbb{R}^3$ ist die Einheitssphäre, welche lokal „am Nordpol" $(0, 0, 1)$ durch eine offene Teilmenge des $\mathbb{R}^m = \mathbb{R}^2$ parametrisiert werden soll („2-dimensionales Objekt" im $\mathbb{R}^3$).

c) **(∗)** *Sei $m = 1$, $n = 2$, $u = (1, 0)$, sowie*

$$\Phi \colon \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2, \qquad (x, (y, z)) \mapsto \big(x^2 + y^2 + z^2, z\big).$$

$\Phi^{-1}(u) \subseteq \mathbb{R}^{m+n} = \mathbb{R}^3$ ist der Einheitskreis, der in der Ebene $H = \mathbb{R}^2 \times \{0\} \subseteq \mathbb{R}^3$ liegt ($\Phi^{-1}(u)$ ist der Schnitt der Einheitssphäre mit H), welche lokal um $(0, 1, 0)$ durch eine offene Teilmenge des $\mathbb{R}^m = \mathbb{R}$ parametrisiert werden soll („1-dimensionales Objekt" im $\mathbb{R}^3$).

Für Φ wie in (502) und $(x, y) \in U \times V \subseteq \mathbb{R}^m \times \mathbb{R}^n$ definieren wir die linearen (stetigen) Abbildungen:

$$\begin{aligned} \mathrm{d}\Phi[x, y]_{\boldsymbol{L}} &: \mathbb{R}^m \to \mathbb{R}^n, \qquad (v_1, \dots, v_m) \mapsto \mathrm{d}_{(x,y)}\, \Phi(v_1, \dots, v_m, \overbrace{0, \dots, 0}^{n\text{-mal}}) \\ \mathrm{d}\Phi[x, y]_{\boldsymbol{R}} &: \mathbb{R}^n \to \mathbb{R}^n, \qquad (w_1, \dots, w_n) \mapsto \mathrm{d}_{(x,y)}\, \Phi(\underbrace{0, \dots, 0}_{m\text{-mal}}, w_1, \dots, w_n), \end{aligned} \tag{503}$$

[141]Beachte, dass $U \times V \subseteq \mathbb{R}^m \times \mathbb{R}^n \cong \mathbb{R}^{m+n}$ gemäß Bemerkung 87.b) offen ist.

sodass wegen der Linearität von $\mathrm{d}_{(x,y)}\Phi$ gilt:[142]

$$\mathrm{d}_{(x,y)}\Phi = \mathrm{d}\Phi[x,y]_{\boldsymbol{L}} \circ \mathrm{pr}_{\boldsymbol{L}} + \mathrm{d}\Phi[x,y]_{\boldsymbol{R}} \circ \mathrm{pr}_{\boldsymbol{R}} \qquad \forall\,(x,y) \in U \times V \tag{504}$$

mit den Projektionen $\mathrm{pr}_{\boldsymbol{L}}\colon \mathbb{R}^m \times \mathbb{R}^n \ni (v,w) \mapsto v \in \mathbb{R}^m$ und $\mathrm{pr}_{\boldsymbol{R}}\colon \mathbb{R}^m \times \mathbb{R}^n \ni (v,w) \mapsto w \in \mathbb{R}^n$.
Die zugehörigen Darstellungsmatrizen (bezüglich der kanonischen Basen) sind gegeben durch:[143]

$$\widehat{\mathrm{d}\Phi}[x,y]_{\boldsymbol{L}} = \begin{pmatrix} \partial_1\Phi_1(x,y) & \cdots & \partial_m\Phi_1(x,y) \\ \vdots & & \vdots \\ \partial_1\Phi_n(x,y) & \cdots & \partial_m\Phi_n(x,y) \end{pmatrix}, \qquad \widehat{\mathrm{d}\Phi}[x,y]_{\boldsymbol{R}} = \begin{pmatrix} \partial_{m+1}\Phi_1(x,y) & \cdots & \partial_{m+n}\Phi_1(x,y) \\ \vdots & & \vdots \\ \partial_{m+1}\Phi_n(x,y) & \cdots & \partial_{m+n}\Phi_n(x,y) \end{pmatrix},$$

also $J_{(x,y)}(\Phi) = \big(\, \widehat{\mathrm{d}\Phi}[x,y]_{\boldsymbol{L}} \;\; \widehat{\mathrm{d}\Phi}[x,y]_{\boldsymbol{R}} \,\big)$ (wenn man beide Matrizen zusammenklebt). Der Satz über die impliziten Funktionen besagt nun Folgendes:

Satz 102 (Satz über Implizite Funktionen)**.** *Seien $m,n,k \in \mathbb{N}_{>0}$, $U \subseteq \mathbb{R}^m$ und $V \subseteq \mathbb{R}^n$ offen, $\Phi \in C^k(U \times V, \mathbb{R}^n)$, sowie $u \in \mathbb{R}^n$ und $(x,y) \in \Phi^{-1}(u)$. Es gelte zudem $\mathrm{d}\Phi[x,y]_{\boldsymbol{R}} \in \mathrm{GL}_n(\mathbb{R})$, d.h.*

$$\det\big(\widehat{\mathrm{d}\Phi}[x,y]_{\boldsymbol{R}}\big) \neq 0 \qquad \textit{für} \qquad \widehat{\mathrm{d}\Phi}[x,y]_{\boldsymbol{R}} = \begin{pmatrix} \partial_{m+1}\Phi_1(x,y) & \cdots & \partial_{m+n}\Phi_1(x,y) \\ \vdots & & \vdots \\ \partial_{m+1}\Phi_n(x,y) & \cdots & \partial_{m+n}\Phi_n(x,y) \end{pmatrix}.$$

1) *Es existieren $\tilde{U} \subseteq \mathbb{R}^m$ und $\tilde{V} \subseteq \mathbb{R}^n$ offen mit $x \in \tilde{U} \subseteq U$ und $y \in \tilde{V} \subseteq V$, sowie eine C^k-Abbildung $\eta\colon \tilde{U} \to \tilde{V}$ (* implizite Funktion zu (x,y) *) mit*[144]

$$\Phi^{-1}(u) \cap (\tilde{U} \times \tilde{V}) = \Gamma_\eta \qquad \textit{d.h.} \qquad \Phi(\tilde{x},\tilde{y}) = u \;\; \textit{für} \;\; (\tilde{x},\tilde{y}) \in \tilde{U} \times \tilde{V} \quad \Leftrightarrow \quad \tilde{y} = \eta(\tilde{x}). \tag{505}$$

(Zu jedem $\tilde{x} \in \tilde{U}$ existiert genau ein $\tilde{y} \in \tilde{V}$ mit $\Phi(\tilde{x},\tilde{y}) = u$, und es gilt $\tilde{y} = \eta(\tilde{x})$.)

Somit gilt $(x,y) \in \mathrm{im}(\Psi) = \Gamma_\eta = \Phi^{-1}(u) \cap (\tilde{U} \times \tilde{V})$ für die injektive C^k-Abbildung[145]

$$\Psi := \mathrm{id}_{\tilde{U}} \times \eta\colon \tilde{U} \to \tilde{U} \times \tilde{V} \subseteq \mathbb{R}^m \times \mathbb{R}^n, \qquad \tilde{x} \mapsto (\tilde{x}, \eta(\tilde{x})).$$

(„Lokale C^k-Parametrisierung der Niveaumenge $\Phi^{-1}(u) \subseteq \mathbb{R}^m \times \mathbb{R}^n$ um (x,y).“)

2) *Für jedes $\tilde{x} \in \tilde{U}$ mit $\mathrm{d}\Phi_{\boldsymbol{R}}[\tilde{x}, \eta(\tilde{x})] \in \mathrm{GL}_n(\mathbb{R})$ gilt die Formel:*

$$\mathrm{d}_{\tilde{x}}\eta = -(\mathrm{d}\Phi[\tilde{x},\eta(\tilde{x})]_{\boldsymbol{R}})^{-1} \circ \mathrm{d}\Phi[\tilde{x},\eta(\tilde{x})]_{\boldsymbol{L}}\colon \mathbb{R}^m \to \mathbb{R}^n. \tag{506}$$

Beachte:

- Die Implizite Funktion $\eta\colon \tilde{U} \to \tilde{V}$ muss weder injektiv noch surjektiv sein (vgl. Beispiel 111). Die explizite Angabe der offenen Teilmenge $\tilde{V} \subseteq \mathbb{R}^n$ dient sozusagen der Sicherstellung der Bedingung (505) (d.h. $\Phi^{-1}(u) \cap (\tilde{U} \times \tilde{V}) = \Gamma_\eta$); also der Eindeutigkeit des Elementes $\tilde{y} \in \tilde{V}$ mit $\Phi(\tilde{x},\tilde{y}) = u$, zu einem vorgegebenen $\tilde{x} \in \tilde{U}$.
- Ist $\hat{U} \subseteq \mathbb{R}^m$ offen mit $x \in \hat{U} \subseteq \tilde{U}$, so gilt (505) offensichtlich entsprechend für $\hat{\eta} := \eta|_{\hat{U}}$ und $\hat{V} := \tilde{V}$, d.h. $\hat{\eta}$ ist ebenfalls eine implizite Funktion zu (x,y).

[142] Es gilt $\mathrm{d}_{(x,y)}\Phi(v,w) = \mathrm{d}_{(x,y)}\Phi(v,0) + \mathrm{d}_{(x,y)}\Phi(0,w) = (\mathrm{d}\Phi[x,y]_{\boldsymbol{L}} \circ \mathrm{pr}_{\boldsymbol{L}})(v,w) + (\mathrm{d}\Phi[x,y]_{\boldsymbol{R}} \circ \mathrm{pr}_{\boldsymbol{R}})(v,w)$ für $v \in \mathbb{R}^m$ und $w \in \mathbb{R}^n$, wegen $\mathbb{R}^m \times \mathbb{R}^n \ni (v,w) = (v,0) + (0,w)$.

[143] Beispielsweise gilt gemäß Terminologie 34 (erster und vierten Schritt) für $1 \leq i \leq m$ und $1 \leq j \leq n$: $(\widehat{\mathrm{d}\Phi}[x,y]_{\boldsymbol{L}})_{ij} = \Omega[m,n](\mathrm{d}\Phi[x,y]_{\boldsymbol{L}})_{ij} = \tilde{\mathrm{pr}}_i(\mathrm{d}\Phi[x,y]_{\boldsymbol{L}}(e_j)) = \tilde{\mathrm{pr}}_i(\mathrm{d}_{(x,y)}\Phi(e_j,0)) = \Omega[m+n,n](\mathrm{d}_{(x,y)}\Phi)_{ij} = J_{(x,y)}(\Phi)_{ij} = \partial_j\Phi_i(x,y)$ mit Terminologie 41.(1).d) im letzten Schritt.

[144] Insbesondere gilt $\eta(x) = y$ wegen $(x,y) \in \Phi^{-1}(u) \cap (\tilde{U} \times \tilde{V})$.

[145] Die Injektivität von Ψ folgt umgehend aus der Injektivität von $\mathrm{id}_{\tilde{U}}$; und die C^k-Eigenschaft ist klar wegen Lemma 84 und Übung 153 ($\mathrm{id}_{\tilde{U}}$ ist glatt).

- Insbesondere kann durch Verkleinern von $\tilde{U}$ um x (Einschränken von η bei Beibehaltung von $\tilde{V}$ als Bildbereich) immer erreicht werden, dass (506) für alle $\tilde{x} \in \tilde{U}$ gilt. In der Tat ist ja $\mathrm{GL}_n(\mathbb{R}) \subseteq \mathrm{End}(\mathbb{R}^n)$ offen und $\bullet^{-1}$ stetig (glatt) wegen Proposition 25; und wegen der Stetigkeit von Φ und η ist
$$\Lambda\colon \tilde{U} \to \mathrm{End}(\mathbb{R}^n), \qquad \tilde{x} \mapsto \big[\mathrm{d}\Phi[\tilde{x},\eta(\tilde{x})]_{\boldsymbol{R}}\colon v \mapsto \mathrm{d}_{(\tilde{x},\tilde{y})}\Phi(0,v)\big]$$
ebenfalls stetig mit $\Lambda(x) = \mathrm{d}\Phi[x,\eta(x)]_{\boldsymbol{R}} = \mathrm{d}\Phi[x,y]_{\boldsymbol{R}} \in \mathrm{GL}_n(\mathbb{R})$. In der Tat gilt ja allgemein für $z, z' \in \tilde{U}$:
$$\begin{aligned}\|\Lambda(z) - \Lambda(z')\|_{\mathrm{op}} &= \sup\nolimits_{\mathbb{R}} \big(\|\mathrm{d}_{(z,\eta(z))}\Phi(0,w) - \mathrm{d}_{(z',\eta(z'))}\Phi(0,w)\| \,\big|\, w \in \mathbb{R}^n\big)\\ &\leq \sup\nolimits_{\mathbb{R}} \big(\|\mathrm{d}_{(z,\eta(z))}\Phi(v,w) - \mathrm{d}_{(z',\eta(z'))}\Phi(v,w)\| \,\big|\, (v,w) \in \mathbb{R}^m \times \mathbb{R}^n\big)\\ &= \|\mathrm{d}_{(z,\eta(z))}\Phi - \mathrm{d}_{(z',\eta(z'))}\Phi\|_{\mathrm{op}}.\end{aligned}$$

Beweis. 2) Ψ ist differenzierbar mit $\mathrm{d}_{\tilde{x}}\Psi = \mathrm{id}_{\mathbb{R}^m} \times \mathrm{d}_{\tilde{x}}\eta$, wegen Lemma 79, Übung 153, und der Kettenregel (Satz 88.2)). Wegen $\Phi \circ \Psi = u$ erhalten wir (Übung 153 mit $\phi \equiv 0$ im ersten Schritt, und die Kettenregel Satz 88.2) im zweiten Schritt):
$$\begin{aligned} 0 &= \mathrm{d}_{\tilde{x}}(\Phi \circ \Psi) = \mathrm{d}_{(\tilde{x},\eta(\tilde{x}))}\Phi \circ \mathrm{d}_{\tilde{x}}\Psi \\ &\overset{(504)}{=} (\mathrm{d}\Phi[\tilde{x},\eta(\tilde{x})]_{\boldsymbol{L}} \circ \mathrm{pr}_{\boldsymbol{L}} + \mathrm{d}\Phi[\tilde{x},\eta(\tilde{x})]_{\boldsymbol{R}} \circ \mathrm{pr}_{\boldsymbol{R}}) \circ (\mathrm{id}_{\mathbb{R}^m} \times \mathrm{d}_{\tilde{x}}\eta) \\ &= \mathrm{d}\Phi[\tilde{x},\eta(\tilde{x})]_{\boldsymbol{L}} + \mathrm{d}\Phi[\tilde{x},\eta(\tilde{x})]_{\boldsymbol{R}} \circ \mathrm{d}_{\tilde{x}}\eta.\end{aligned}$$
Umstellen liefert (506).

1) Wegen Lemma 84 und der Glattheit von $\mathrm{id}_U \circ \mathrm{pr}_{\boldsymbol{L}}$ (Übung 153 sowie die Kettenregel Lemma 88 bzw. Satz 88.2)) ist
$$\Theta := (\mathrm{id}_U \circ \mathrm{pr}_{\boldsymbol{L}}) \times \Phi\colon \mathbb{R}^m \times \mathbb{R}^n \supseteq U \times V \to \mathbb{R}^m \times \mathbb{R}^n, \qquad (\tilde{x},\tilde{y}) \mapsto (\tilde{x}, \Phi(\tilde{x},\tilde{y})) \tag{507}$$
eine C^k-Abbildung mit
$$\mathrm{d}_{(\tilde{x},\tilde{y})}\Theta = (\mathrm{id}_{\mathbb{R}^m} \circ \mathrm{pr}_{\boldsymbol{L}}) \times \mathrm{d}_{(\tilde{x},\tilde{y})}\Phi\colon (v,w) \mapsto \big(v, \mathrm{d}_{(\tilde{x},\tilde{y})}\Phi(v,w)\big) \qquad \forall\, (\tilde{x},\tilde{y}) \in \tilde{U} \times \tilde{V}.$$
Dann gilt $\mathrm{d}_{(x,y)}\Theta \in \mathrm{GL}_{m+n}(\mathbb{R})$, denn mit (504) folgt ($\mathbf{0}_{mn}$ ist die $m \times n$-Matrix, die nur aus Nullen besteht):
$$J_{(x,y)}(\Theta) = \begin{pmatrix} \mathbf{1}_m & \mathbf{0}_{mn} \\ \widehat{\mathrm{d}\Phi}[x,y]_{\boldsymbol{L}} & \widehat{\mathrm{d}\Phi}[x,y]_{\boldsymbol{R}} \end{pmatrix} \overset{\text{Kästchensatz (LA)}}{\Longrightarrow} \det\big(J_{(x,y)}(\Theta)\big) = \det\big(\widehat{\mathrm{d}\Phi}[x,y]_{\boldsymbol{R}}\big) \neq 0.$$
Wegen Satz 99 (lokale Umkehrbarkeit) existieren offene Umgebungen $O \subseteq \mathbb{R}^m \times \mathbb{R}^n$ von (x,y) und $Q \subseteq \mathbb{R}^m \times \mathbb{R}^n$ von $\Theta(x,y) = (x, \Phi(x,y)) = (x,u)$, sodass
$$\alpha := (\Theta|_O)|^Q\colon O \to Q \tag{508}$$
ein C^k-Diffeomorphismus ist, mit $(x,u) = \alpha(x,y) \in Q$. Wir erhalten aus der Definition von Θ (erste Implikation), sowie der Injektivität von α (zweite Implikation):
$$\begin{aligned}\Phi(\tilde{x},\tilde{y}) = \Phi(\tilde{x},\tilde{y}') \quad \text{für} \quad & (\tilde{x},\tilde{y}) \in O \ni (\tilde{x},\tilde{y}') \\ \Longrightarrow \qquad & \alpha(\tilde{x},\tilde{y}) = \Theta(\tilde{x},\tilde{y}) = \Theta(\tilde{x},\tilde{y}') = \alpha(\tilde{x},\tilde{y}') \\ \Longrightarrow \qquad & (\tilde{x},\tilde{y}) = (\tilde{x},\tilde{y}') \\ \Longrightarrow \qquad & \tilde{y} = \tilde{y}'.\end{aligned} \tag{509}$$

Bemerkung:

- Wir können O um (x,y) beliebig verkleinern (Q entsprechend verkleinern, und α einschränken sowie koeinschränken), ohne die bisher erhaltenen Eigenschaften zu verlieren.

  (Ist nämlich $\tilde{O} \subseteq \mathbb{R}^m \times \mathbb{R}^n$ offen mit $(x,y) \in \tilde{O} \subseteq O$, so ist $\tilde{Q} := \alpha(\tilde{O}) \subseteq \mathbb{R}^m \times \mathbb{R}^n$ ebenfalls offen mit $(x,u) = \alpha(x,y) \in \tilde{Q}$; denn $\alpha\colon O \to Q$ ist ein Homöomorphismus, und O, Q sind offen in $\mathbb{R}^m \times \mathbb{R}^n$ (siehe (140) in Bemerkung 36.a)).[146]
- Ist $\hat{W} \subseteq \mathbb{R}^m \times \mathbb{R}^n$ offen und $(\hat{x}, \hat{y}) \in \hat{W}$, so existieren offene Umgebungen $\hat{W}_1 \subseteq \mathbb{R}^m$ von $\hat{x}$ und $\hat{V} \subseteq \mathbb{R}^n$ von $\hat{y}$, sodass $\hat{W}_1 \times \hat{W}_2 \subseteq \hat{W}$ gilt. (Bemerkung 87.b) und Terminologie 31.2) [147])

Wir Schlussfolgern:

- Wir können $O = \tilde{U} \times \tilde{V}$ für offene Teilmengen $\tilde{U} \subseteq \mathbb{R}^m$, $\tilde{V} \subseteq \mathbb{R}^n$ mit $x \in \tilde{U} \subseteq U$ und $y \in \tilde{V} \subseteq V$ annehmen.[148]

  Offensichtlich gilt dann (die rechte Seite gilt wegen (507))

$$\mathrm{pr}_{\boldsymbol{R}} \circ \alpha^{-1}\colon Q \to \tilde{V} \qquad \text{sowie} \qquad (\mathrm{pr}_{\boldsymbol{L}} \circ \alpha^{-1})(\tilde{x}, w) = \tilde{x} \qquad \forall\, (\tilde{x}, w) \in Q. \tag{510}$$

- Wir können $\tilde{U} \times \{u\} \subseteq Q$ annehmen, indem wir $\tilde{U}$ um x verkleinern.

  Beweis der Behauptung. Wegen $(x,u) = \alpha(x,y) \in Q$, existieren $O_1 \subseteq \mathbb{R}^m$ offen mit $x \in O_1$ sowie $O_2 \subseteq \mathbb{R}^n$ offen mit $u \in O_2$, sodass $O_1 \times O_2 \subseteq Q$ gilt. Wir müssen jetzt nur $\tilde{U}$ durch $\tilde{U} \cap O_1$ (offen) ersetzen; denn per Konstruktion gilt ja $x \in \tilde{U} \cap O_1$ mit $(\tilde{U} \cap O_1) \times \{u\} \subseteq O_1 \times O_2 \subseteq Q$. $\square$

  Nun ist die konstante Abbildung $\chi_u\colon \tilde{U} \ni x \mapsto u \in Q$ glatt. Also ist (beachte die linke Seite von (510), sowie $\mathrm{im}(\mathrm{id}_{\tilde{U}} \times \chi_u) = \tilde{U} \times \{u\} \subseteq Q$)

$$\eta := \mathrm{pr}_{\boldsymbol{R}} \circ \alpha^{-1} \circ \big(\mathrm{id}_{\tilde{U}} \times \chi_u\big)\colon \tilde{U} \to \tilde{V}, \qquad \tilde{x} \mapsto \mathrm{pr}_{\boldsymbol{R}}(\alpha^{-1}(\tilde{x}, u)) \tag{511}$$

  wegen (Lemma 84, Übung 153, Lemma 88) eine C^k-Abbildung.

Für $\tilde{x} \in \tilde{U}$ erhalten wir (mit der Definition von $\eta/\Theta/\alpha$ im ersten/dritten/vierten Schritt)

$$\begin{aligned}
\Phi(\tilde{x}, \eta(\tilde{x})) &\overset{(510),(511)}{=} \Phi\Big(\underbrace{\overbrace{\big(\mathrm{pr}_{\boldsymbol{L}} \circ \alpha^{-1}\big)(\tilde{x}, u)}^{\overset{(510)}{=}\ \tilde{x}}, \overbrace{\big(\mathrm{pr}_{\boldsymbol{R}} \circ \alpha^{-1}\big)(\tilde{x}, u)}^{\overset{(511)}{=}\ \eta(\tilde{x})}}_{= \alpha^{-1}(\tilde{x}, u) \ \text{ wegen } \ \alpha^{-1} = (\mathrm{pr}_{\boldsymbol{L}} \circ \alpha^{-1},\, \mathrm{pr}_{\boldsymbol{R}} \circ \alpha^{-1})} \Big) \\
&= \quad \Phi(\alpha^{-1}(\tilde{x}, u)) \overset{(507)}{=} (\mathrm{pr}_{\boldsymbol{R}} \circ \Theta)(\alpha^{-1}(\tilde{x}, u)) \overset{(508)}{=} \mathrm{pr}_{\boldsymbol{R}}(\tilde{x}, u) = u,
\end{aligned}$$

also $\Gamma_\eta \subseteq \Phi^{-1}(u) \cap (\tilde{U} \times \tilde{V})$; und wegen der Eindeutigkeitsaussage (509) folgt hiermit (505). $\square$

Bemerkung 109. *Es seien die Voraussetzungen von Satz 102 erfüllt. Für* $n = 1$ *schreibt sich* (506) *in Termen der Darstellungsmatrizen der involvierten linearen Abbildungen wegen der Rechenregeln* $(\star)$ *und* (491) *für Darstellungsmatrizen in der Form*

$$\nabla\eta(\tilde{x}) = -\frac{\big(\partial_1\Phi(\tilde{x}, \eta(\tilde{x})), \dots, \partial_m\Phi(\tilde{x}, \eta(\tilde{x}))\big)}{\partial_{m+1}\Phi(\tilde{x}, \eta(\tilde{x}))} \qquad \forall\, \tilde{x} \in \tilde{U} \tag{512}$$

denn es gilt ja $J_\bullet(\eta) = \nabla\eta$ *sowie* $\widehat{\mathrm{d}\Phi}[\tilde{x}, \eta(\tilde{x})]_{\boldsymbol{R}} = \partial_{m+1}\Phi(\tilde{x}, \eta(\tilde{x}))$.

[146](1) $\tilde{O}$ ist offen in O (als metrischer Unterraum) da $O, \tilde{O}$ offen in $\mathbb{R}^m \times \mathbb{R}^n$ sind. (2) $\tilde{Q}$ ist dann offen in Q (als metrischer Unterraum) da α ein Homöomorphismus ist. (3) Da Q offen in $\mathbb{R}^m \times \mathbb{R}^n$ ist folgt, dass $\tilde{Q}$ auch offen in $\mathbb{R}^m \times \mathbb{R}^n$ ist.

[147]Man sieht das auch anhand der Gestalt der offenen Bälle in $\mathbb{R}^m \times \mathbb{R}^n$ bezüglich der Maximumsnorm.

[148]Beachte: Sind $A \subseteq \mathbb{R}^m$, $B \subseteq \mathbb{R}^n$ gegeben mit $A \times B \subseteq O$, so gilt wegen $O \subseteq U \times V$ automatisch $A \subseteq U$ und $B \subseteq V$.

Man beachte hierbei, dass zwar $\eta(\tilde{x})$ für beliebiges $\tilde{x} \in \tilde{U}$ nicht explizit gegeben ist, für $\tilde{x} = x$ wegen $\eta(x) = y$ aber schon. Somit lässt sich $\nabla\eta(x)$ mit (512) tatsächlich auch explizit ausrechnen.

(*Anmerkung: *Die rechte Seite lässt sich schöner schreiben als* $-\frac{\nabla\Phi_{\eta(\tilde{x})}(\tilde{x})}{\Phi_{\tilde{x}}(\eta(\tilde{x}))'}$*, mit den* C^k*-Abbildungen (Details siehe Bemerkung* 22):*

$$\Phi_{\eta(\tilde{x})}\colon \mathbb{R}^m \ni x \mapsto \Phi(x, \eta(\tilde{x})) \in \mathbb{R}^n \qquad \text{und} \qquad \Phi_{\tilde{x}}\colon \mathbb{R} = \mathbb{R}^n \ni y \mapsto \Phi(\tilde{x}, y) \in \mathbb{R}^n. \Big)$$

Beispiel 111. a) *Sei* $m, n = 1$, $u = 1$, *sowie* ($J_{(x,y)}\Phi = (2x\ 2y)$)

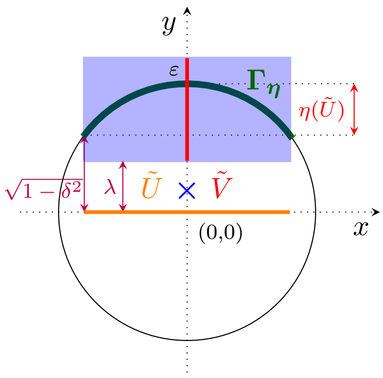
$$\Phi\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \qquad (x, y) \mapsto x^2 + y^2.$$

Dann ist $\Phi^{-1}(u) \subseteq \mathbb{R}^2$ *der Einheitskreis; und wir betrachten den Nordpol* $(x, y) = (0, 1)$*:*

- $\widehat{\mathrm{d}\Phi}[\tilde{x}, \tilde{y}]_{\boldsymbol{R}} = \partial_2\Phi(\tilde{x}, \tilde{y}) = 2\tilde{y}$ *ist eindimensional, also gilt* $\det(\widehat{\mathrm{d}\Phi}[\tilde{x}, \tilde{y}]_{\boldsymbol{R}}) = 2\tilde{y} > 0$ *für* $\tilde{y} > 0$.
- *Die Aussage in Satz 102 gilt bspw. für:*

$$\tilde{U} = (-\delta, \delta), \qquad \tilde{V} = (\lambda, 1 + \varepsilon), \qquad \eta\colon \tilde{U} \ni \tilde{x} \mapsto \sqrt{1 - \tilde{x}^2} \in \tilde{V}$$

mit $0 < \delta \leq 1$, $-\sqrt{1 - \delta^2} \leq \lambda \leq \sqrt{1 - \delta^2}$, $\varepsilon > 0$ *beliebig aber fixiert.*



Skizze für $\delta = 0.8$, $\lambda = 0.4$, $\varepsilon = 0.2$.

Sei beispielsweise $\delta = 1$, *also* $\lambda = 0$ *und* $\varepsilon > 0$*:*

– $\Psi\colon \tilde{U} \ni \tilde{x} \mapsto (\tilde{x}, \eta(\tilde{x})) \in \tilde{U} \times \tilde{V}$ *parametrisiert den oberen Einheitshalbkreis (ohne* $(\pm 1, 0)$*).*
– η *ist nicht injektiv, denn es gilt* $\eta(\tau) = \eta(-\tau)$ *für alle* $\tau \in (0, 1)$.
– η *ist nicht surjektiv (für kein* $\varepsilon > 0$*), denn es gilt* $\operatorname{im}(\eta) = (0, 1] \subset (0, 1 + \varepsilon) = \tilde{V}$.
– *Es gilt* $\eta'(\tilde{x}) = -\frac{\tilde{x}}{\eta(\tilde{x})}$ *für jedes* $\tilde{x} \in \tilde{V} = (-1, 1)$*; und mit (512) erhalten wir ebenfalls:*

$$\eta'(\tilde{x}) = -\frac{\partial_1\Phi\big(\tilde{x}, \sqrt{1-\tilde{x}^2}\big)}{\partial_2\Phi\big(\tilde{x}, \sqrt{1-\tilde{x}^2}\big)} = -\frac{\tilde{x}}{\sqrt{1-\tilde{x}^2}} = -\frac{\tilde{x}}{\eta(\tilde{x})}.$$

Beachte: *Für* $\tilde{x} \in \tilde{V} = (-1, 1)$ *gilt* $\det\big(\widehat{\mathrm{d}\Phi}[\tilde{x}, \eta(\tilde{x})]_{\boldsymbol{R}}\big) = 2\eta(\tilde{x}) = \sqrt{1 - \tilde{x}^2} > 0$.

(∗) *Da* $\tilde{U}$ *wegzusammenhängen ist, gilt zudem die Eindeutigkeitsaussage in Bemerkung* 21.*

b) **(∗)** *Sei* $m = 2$, $n = 1$, $u = 1$, *sowie*

$$\Phi\colon \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}, \qquad ((x_1, x_2), z) \mapsto x_1^2 + x_2^2 + z^2.$$

Dann ist $\Phi^{-1}(u) \subseteq \mathbb{R}^3$ *die Einheitssphäre; und wir betrachten den Nordpol* $((x_1, x_2), z) = ((0, 0), 1)$*:*

- $\widehat{\mathrm{d}\Phi}[(\tilde{x}_1, \tilde{x}_2), \tilde{z}]_{\boldsymbol{R}} = \partial_3\Phi((\tilde{x}, \tilde{y}), \tilde{z}) = 2\tilde{z} \quad \Rightarrow \quad \det(\widehat{\Phi}[(\tilde{x}_1, \tilde{x}_2), \tilde{z}]_{\boldsymbol{R}}) = 2\tilde{z} > 0$ *für* $\tilde{z} > 0$.

- *Die Aussage in Satz 102 gilt bspw. für:* ($\varepsilon > 0$ *beliebig aber fixiert,* $\tilde{U}$ *ist wegzshd.*)

$$\tilde{U} = \{w \in \mathbb{R}^2 \mid \|w\|_{\text{euk}} < 1\} \subseteq \mathbb{R}^2, \qquad \tilde{V} = (0, 1+\varepsilon), \qquad \eta \colon \tilde{U} \ni \tilde{w} \mapsto \sqrt{1 - \|\tilde{w}\|_{\text{euk}}^2} \in \tilde{V}.$$

  - $\Psi \colon \tilde{U} \ni \tilde{w} \mapsto (\tilde{w}, \eta(\tilde{w})) \in \mathbb{R}^3$ *parametrisiert die obere Einheitshalbsphäre.*
  - η *ist nicht injektiv, denn es gilt* $\eta(\tilde{w}) = \eta(\tilde{w}')$ *für alle* $\tilde{w}, \tilde{w}' \in \tilde{U}$ *mit* $\|\tilde{w}\|_{\text{euk}} = \|\tilde{w}'\|_{\text{euk}}$.
  - η *ist nicht surjektiv (für kein* $\varepsilon > 0$*), denn es gilt* $\operatorname{im}(\eta) = (0,1] \subset (0, 1+\varepsilon)$.
  - *Es gilt* $\nabla\eta(\tilde{x}_1, \tilde{x}_2) = -\frac{(\tilde{x}_1, \tilde{x}_2)}{\eta(\tilde{x}_1, \tilde{x}_2)}$ *für jedes* $(\tilde{x}_1, \tilde{x_2}) \in \tilde{U}$*; und mit* (512) *erhalten wir ebenfalls:*

$$\nabla\eta(\tilde{x}_1, \tilde{x}_2) = -\frac{\left(\partial_1\Phi\left((\tilde{x}_1,\tilde{x}_2), \sqrt{1-\tilde{x}_1^2-\tilde{x}_2^2}\right), \partial_2\Phi\left((\tilde{x}_1,\tilde{x}_2), \sqrt{1-\tilde{x}_1^2-\tilde{x}_2^2}\right)\right)}{\partial_3\Phi\left((\tilde{x}_1,\tilde{x}_2), \sqrt{1-\tilde{x}_1^2-\tilde{x}_2^2}\right)} = -\frac{(\tilde{x}_1,\tilde{x}_2)}{\sqrt{1-\tilde{x}_1^2-\tilde{x}_2^2}} = -\frac{(\tilde{x}_1,\tilde{x}_2)}{\eta(\tilde{x}_1,\tilde{x}_2)}.$$

c) **(∗)** *Sei* $m = 1$, $n = 2$, $u = (1, 0)$, *sowie*

$$\Phi \colon \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2, \qquad (x, (y_1, y_2)) \mapsto \left(x^2 + y_1^2 + y_2^2, y_2\right).$$

Dann ist $\Phi^{-1}(u) \cong \{(a, b, 0) \mid a^2 + b^2 = 1\}$ *der Einheitskreis, welcher in der Ebene* $\mathbb{R}^2 \times \{0\}$ *liegt. Wir betrachten den Punkt* $(x, (y_1, y_2)) = (0, (1, 0))$*:*

- *Es gilt* $\det\left(\widehat{\mathrm{d}\Phi}[0, (1, 0)]\right) = 2$, *wegen* $\begin{pmatrix} \partial_2\Phi_1(x,(y_1,y_2)) & \partial_3\Phi_1(x,(y_1,y_2)) \\ \partial_2\Phi_2(x,(y_1,y_2)) & \partial_3\Phi_2(x,(y_1,y_2)) \end{pmatrix} = \begin{pmatrix} 2y_1 & 2y_2 \\ 0 & 1 \end{pmatrix}.$
- *Die Aussage in Satz 102 gilt bspw. für:* ($\varepsilon > 0$ *beliebig aber fixiert,* $\tilde{U}$ *ist wegzshd.*)

$$\tilde{U} = (-1, 1), \qquad \tilde{V} = (0, 1+\varepsilon) \times \underbrace{(-\varepsilon, \varepsilon)}_{(\#)}, \qquad \eta \colon \tilde{U} \ni \tilde{x} \mapsto \left(\sqrt{1 - \tilde{x}^2}, 0\right) \in \tilde{V}.$$

((#): Das offene Intervall $(-\varepsilon, \varepsilon)$ *dient hier nur der Sicherstellung, dass* $\tilde{V} \subseteq \mathbb{R}^2$ *offen ist.)*

Übung 164. *Wir betrachten die glatte Abbildung*

$$\Phi \colon \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}, \qquad ((x_1, x_2), y) \mapsto y^3 + 4y - x_1^2 + x_1x_2^2 + 8x_2 - 7$$

sowie $u \in \mathbb{R}$ *mit* $\Phi^{-1}(u) \neq \emptyset$,[149] *und fixieren* $(x, y) \in \Phi^{-1}(u)$ *mit* $x = (x_1, x_2) \in \mathbb{R}^2$.

a) *Zeigen Sie, dass die Voraussetzungen von Satz 102 im Punkt* (x, y) *erfüllt sind, dass also* $\mathrm{d}\Phi[x, y]_{\boldsymbol{R}} \in \mathrm{GL}_1(\mathbb{R})$ *gilt.*

   (Salopp ausgedrückt: Die Urbildmenge $\Phi^{-1}(u)$ *ist lokal um jedem Punkt gegeben durch den Graphen einer Impliziten Funktion.)*

b) *Berechnen Sie den Gradienten* $\nabla\eta(x)$ *für eine durch Satz 102 bereitgestellte (glatte) implizite Funktion* $\eta \colon \tilde{U} \to \tilde{V}$ *für* $\tilde{U} \subseteq \mathbb{R}^2$ *offen mit* $x \in \tilde{U}$ *sowie* $\tilde{V} \subseteq \mathbb{R}$ *offen mit* $y \in \tilde{V}$*, d.h.,*

$$\Phi^{-1}(u) \cap (\tilde{U} \times \tilde{V}) = \Gamma_\eta.$$

Übung 165. *Wir betrachten die glatte Abbildung*

$$\Phi \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \qquad (x, y) \mapsto y^3 + x^5 + 2y,$$

und setzen $u = 0$. *Zeigen Sie, dass* Φ *für jedes* $(x, y) \in \Phi^{-1}(u)$ *die Voraussetzungen von Satz 102 erfüllt, und berechnen Sie* $\eta'(x)$ *für eine zu* (x, y) *gehörige (glatte) implizite Funktion* η.

[149] Beispielsweise $u = -7$ wegen $\Phi((0, 0), 0) = -7$.

Übung 166 (Satz 99 aus Satz 102). *Sei $n \in \mathbb{N}_{>0}$, $W \subseteq \mathbb{R}^n$ offen, $p \in U$ und $f\colon W \to \mathbb{R}^n$ eine C^k-Abbildung mit $k \in \mathbb{N}_{>0} \cup \{\infty\}$, sodass $\mathrm{d}_p f \in \mathrm{GL}_n(\mathbb{R})$ gilt. Folgern Sie aus Satz 102 (mit $m = n$), dass f lokal um p umkehrbar ist.*

Anleitung: *Betrachten Sie hierfür die C^k- Abbildung*

$$\Phi := \big(\mathrm{id}_{\mathbb{R}^n} \circ \mathrm{pr}_{\boldsymbol{R}} - f \circ \mathrm{pr}_{\boldsymbol{L}}\big)\colon \overbrace{\mathbb{R}^n}^{=:\,U} \times \overbrace{W}^{=:\,V} \to \mathbb{R}^n, \qquad (x, y) \mapsto x - f(y)$$

mit $u := 0 \in \mathbb{R}^n$ und $(x, y) := (f(p), p)$, verifizieren Sie $\phi(x, y) = u$ sowie $\widehat{\mathrm{d}\Phi}[x, y]_{\boldsymbol{R}} = J_p f$, und werten Sie schließlich die Bedingung (505) *aus. (Möglicherweise müssen Sie den Bildbereich einer durch Satz 102 bereitgestellten impliziten Funktion trickreich verkleinern.)*

16.4.1 Zusatzmaterial*

Bemerkung* 21 (Eindeutigkeit der Impliziten Funktion). *Es seien die Voraussetzungen von Satz 102 erfüllt. Die Implizite Funktion η ist eindeutig in dem folgenden Sinne:*

- **Definition:** *Eine Teilmenge $O \subseteq \mathbb{R}^m$ heißt* wegzusammenhängend*, wenn zu je zwei Punkten $z, z' \in O$ eine stetige Kurve (Weg) $\gamma\colon [0, 1] \to O$ existiert mit $\gamma(0) = z$ und $\gamma(1) = z'$.*

  Durch verkleinern von $\tilde{U}$ können wir erreichen, dass $\tilde{U}$ wegzusammenhängend ist. In der Tat finden wir ja ein $\varepsilon > 0$ mit $\mathsf{B}_\varepsilon(x) \subseteq \tilde{U}$ (bzgl. einer fixierten Norm $\|\cdot\|$ auf $\mathbb{R}^m$), wobei $\mathsf{B}_\varepsilon(x)$ wegen der Dreiecksungleichung wegzusammenhängend ist:

  Beweis der Behauptung. Für $\gamma\colon [0, 1] \ni t \mapsto (1 - t) \cdot z + t \cdot z' \in \mathbb{R}^m$ mit $z, z' \in \mathsf{B}_\varepsilon(x)$ gilt wegen der Dreiecksungleichung:

  $$\begin{aligned} \|\gamma(t) - x\| &= \|(1 - t) \cdot (z - x) + t \cdot (z' - x)\| \\ &\leq \ (1 - t) \cdot \|z - x\| + t \cdot \|z' - x\| \\ &< \ (1 - t) \cdot \varepsilon + t \cdot \varepsilon = \varepsilon. \end{aligned}$$ □

- **Eindeutigkeit:** *Seien $\eta\colon \tilde{U} \to \tilde{V}$ und $\eta'\colon \tilde{U} \to \tilde{V}'$ beides stetige Abbildungen,*[150] *für die* (505) *gilt (also implizite Funktionen zu (x, y)). Ist $\tilde{U}$ wegzusammenhängend, so gilt $\eta = \eta'$ (formell präziser $\mathrm{im}(\eta) = \mathrm{im}(\eta')$ und $\eta|^{\mathrm{im}(\eta)} = \eta'|^{\mathrm{im}(\eta')}$).*

  Beweis der Eindeutigkeit. Sei $\tilde{x} \in \tilde{U}$ vorgegeben, und $\gamma\colon [0, 1] \to \mathbb{R}^m$ stetig mit $\mathrm{im}(\gamma) \subseteq \tilde{U}$, $\gamma(0) = x$, $\gamma(1) = \tilde{x}$. Wegen $(x, y) \in (\tilde{U} \times \tilde{V}) \cap (\tilde{U} \times \tilde{V}')$ gilt $\eta(x) = \eta'(x)$; also ist

  $$\tau := \sup\nolimits_{\mathbb{R}} \Big(\underbrace{\big\{\sigma \in [0, 1] \ \big|\ \eta(\gamma(t)) = \eta'(\gamma(t)) \ \text{für alle} \ t \in [0, \sigma]\,\big\}}_{=:\,M} \Big) \in [0, 1]$$

  definiert ($0 \in M \neq \emptyset$). Dann gilt $\tau \in M$; denn wegen Übung 64 existiert eine Folge $\{\tau_n\}_{n \in \mathsf{N}}$ in M mit $\lim_n \tau_n = \tau$, und mit der Stetigkeit (Folgenstetigkeit) von $\eta \circ \gamma, \eta' \circ \gamma$ folgt:

  $$\eta(\gamma(\tau)) = \lim\nolimits_n \gamma(\eta(\tau_n)) = \lim\nolimits_n \gamma(\eta'(\tau_n)) = \eta'(\gamma(\tau)).$$

  Dann ist $O := \tilde{V} \cap \tilde{V}'$ offen mit $\eta(\gamma([0, \tau])) \subseteq O \supseteq \eta'(\gamma([0, \tau]))$.

[150] Für die Eindeutigkeitsaussage genügt es in der Tat, nur die Stetigkeit vorauszusetzen.

Gilt nun $\tau < 1$, so existiert daher ein $0 < \varepsilon \leq 1 - \tau$ mit $\eta(\gamma([0, \tau + \varepsilon))) \subseteq O \supseteq \eta'(\gamma([0, \tau + \varepsilon)))$. Für jedes $t \in [0, \tau + \varepsilon/2]$, gilt dann:

$$(\gamma(t), \eta(\gamma(t))) \in \underbrace{\overbrace{\tilde{U} \times O}^{\subseteq \tilde{U} \times \tilde{V}'}}_{\subseteq \tilde{U} \times \tilde{V}} \ni (\gamma(t), \eta'(\gamma(t))) \qquad \stackrel{(505)}{\Longrightarrow} \qquad \eta(\gamma(t)) = \eta'(\gamma(t)).$$

Wir erhalten den Widerspruch $\tau < \tau + \varepsilon/2 \in M \leq \tau$; also gilt $\tau = 1$, und somit

$$\eta(\tilde{x}) = \eta(\gamma(1)) = \eta'(\gamma(1)) = \eta'(\tilde{x}). \qquad \square$$

Bemerkung* 22. *Sei Φ wie in (502), sowie $x \in U$, $y \in V$. Per Induktion folgt aus Übung 144, dass*

$$\begin{aligned} \Phi_y &\colon \mathbb{R}^m \to \mathbb{R}^n, \qquad & x &\mapsto \Phi(x, y) \\ \Phi_x &\colon \mathbb{R}^n \to \mathbb{R}^n, \qquad & y &\mapsto \Phi(x, y) \end{aligned}$$

ebenfalls C^k-Abbildungen sind, mit

$$\begin{aligned} \partial_{j_1} \ldots \partial_{j_p} \Phi_y &= (\partial_{j_1} \ldots \partial_{j_p} \Phi)|_{U \times \{y\}} \qquad && \forall\, 1 \leq p \leq k,\ \{j_1, \ldots, j_p\} \subseteq \{1, \ldots, n\} \\ \partial_{j_1} \ldots \partial_{j_p} \Phi_x &= (\partial_{j_1} \ldots \partial_{j_p} \Phi)|_{\{x\} \times V} \qquad && \forall\, 1 \leq p \leq k,\ \{j_1, \ldots, j_p\} \subseteq \{m+1, \ldots, m+n\}. \end{aligned}$$

Insbesondere besteht dann der Zusammenhang

$$\begin{array}{ccc} \mathrm{d}\Phi[x, y]_{\boldsymbol{L}} = \mathrm{d}_x \Phi_y & \textit{und} & \mathrm{d}\Phi[x, y]_{\boldsymbol{R}} = \mathrm{d}_y \Phi_x, \\ & \textit{also} & \\ \widehat{\mathrm{d}\Phi}[x, y]_{\boldsymbol{L}} = J_x(\Phi_y) & \textit{und} & \widehat{\mathrm{d}\Phi}[x, y]_{\boldsymbol{R}} = J_y(\Phi_x). \end{array}$$

Beweis. Wir zeigen nur die Aussage für Φ_x (die entsprechende Aussage für Φ_y folgt analog): Die Aussage gilt für $k = 0$, denn wegen Übung 144 ist Φ_x stetig. Wir zeigen per Induktion:

$$\partial_{j_1} \ldots \partial_{j_p} \Phi_x = (\partial_{j_1} \ldots \partial_{j_p} \Phi)|_{\{x\} \times V} \qquad \forall\, 1 \leq p \leq k,\ \{j_1, \ldots, j_p\} \subseteq \{m+1, \ldots, m+n\}. \tag{513}$$

Die Stetigkeit der partiellen Ableitungen folgt dann automatisch aus Übung 144.

- Für $k = 1$ ist (513) klar; denn für $y \in V$, $m+1 \leq j \leq m+n$, und $t \in \mathbb{R}$ mit $(x, y) + t \cdot e_j \subseteq U \times V$, gilt $\Phi((x, y) + t \cdot e_j) = \Phi_x(y + t \cdot e_j)$, also

$$\partial_j \Phi(x, y) = \lim_{t \to 0} \frac{\Phi((x, y) + t \cdot e_j) - \Phi(x, y)}{t} = \lim_{t \to 0} \frac{\Phi_x(y + t \cdot e_j) - \Phi_x(y)}{t} = \partial_j \Phi_x(y).$$

- Es gelte die Aussage für ein $k \in \mathbb{N}_{>0}$, und es sei Φ eine C^{k+1}-Abbildung. Für $m+1 \leq j_0, \ldots, j_k \leq m+n$ und $y \in V$ erhalten wir aus der Induktionsvoraussetzung im zweiten Schritt:

$$\partial_{j_0} \ldots \partial_{j_k} \Phi(x, y) = \partial_{j_0}(\partial_{j_1} \ldots \partial_{j_k} \Phi)(x, y) = \partial_{j_0}(\partial_{j_1} \ldots \partial_{j_k} \Phi_x)(y) = \partial_{j_0} \ldots \partial_{j_k} \Phi_x(y). \qquad \square$$

17 Gewöhnliche Differentialgleichungen

In diesem Kapitel beschäftigen wir uns mit der Existenz und Eindeutigkeit von Lösungen von gewöhnlichen Differentialgleichungen erster Ordnung. Wir diskutieren zunächst den eindimensionalen Fall (hier gelten besondere Aussagen) und widmen uns danach der höher-dimensionalen Situation.

17.1 Differentialgleichungen erster Ordnung

In diesem Abschnitt behandeln wir gewöhnliche Differentialgleichungen erster Ordnung:

Definition 80. *Sei $M \subseteq \mathbb{R} \times \mathbb{R} \cong \mathbb{R}^2$ und $f\colon M \to \mathbb{R}$ eine stetige Abbildung.*[151] *Dann heißt*

$$\varphi' = f(\cdot, \varphi) \qquad (\textit{salopp } x' = f(t,x)) \tag{514}$$

(gewöhnliche) Differentialgleichung erster Ordnung (auf M). *Eine Bedingung der Gestalt*

$$\varphi(t_0) = x_0 \qquad \textit{für} \qquad (t_0, x_0) \in M \tag{515}$$

heißt Anfangsbedingung zu der Differentialgleichung (514).

- *Eine* (lokale) Lösung von (514) *ist eine differenzierbare Abbildung $\varphi\colon D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}$ ein nichtentartetes Intervall, sodass gilt:*

  $$\Gamma_\varphi = \{(t, \varphi(t)) \mid t \in D\} \subseteq M \qquad \textit{sowie} \qquad \varphi'(t) = f(t, \varphi(t)) \qquad \forall\, t \in D. \tag{516}$$

  (Beachte, dass die rechte Seite nur formuliert werden kann, wenn die linke Seite gilt.)
- *Gilt zusätzlich zu* (516) *noch*

  $$(t_0, x_0) \in \Gamma_\varphi \quad \textit{für ein} \quad (t_0, x_0) \in M \qquad \big(t_0 \in D \quad \textit{mit} \quad \varphi(t_0) = x_0 \quad \textit{und} \quad (t_0, x_0) \in M\big), \tag{517}$$

  so wird φ als (lokale) Lösung von (514) zur Anfangsbedingung (515) *bezeichnet.*[152]

Bemerkung 110. (a) *Die linke Seite von* (516) *erzwingt natürlich*

$$D \subseteq \mathrm{pr}_1(M) = \{t \in \mathbb{R} \mid \exists\, x \in \mathbb{R}\colon (t,x) \in M\}.$$

Im Allgemeinen ist aber D eine echte Teilmenge der rechten Seite (vgl. Beispiel 114).

(b) *Angenommen f hängt nicht von x ab; genauer*

$$f\colon M = A \times \mathbb{R} \to \mathbb{R}, \qquad (t,x) \mapsto h(t)$$

mit einer stetigen Abbildung $h\colon A \to \mathbb{R}$. Dann schreibt sich (514) *in der Form $\varphi' = h$. Ist A ein nichtentartetes Intervall sowie $(t_0, x_0) \in A \times \mathbb{R}$, so ist gemäß dem Hauptsatz*

$$\varphi(t) = x_0 + \textstyle\int_{t_0}^t h \qquad \forall\, t \in A.$$

die eindeutige Lösung von $\varphi' = h = f(\cdot, \varphi)$ mit $\varphi(t_0) = x_0$ und $\mathrm{dom}(\varphi) = A$.

(c) *Mit Hilfe des Hauptsatzes sieht man Folgendes (vgl. Lemma 93 in Abschnitt 17.3.3):*

Sei $\varphi\colon D \to \mathbb{R}$ stetig mit $\Gamma_\varphi \subseteq M$ und $(t_0, x_0) \in M$, sowie $D \subseteq \mathbb{R}$ nichtentartet. Dann ist φ eine Lösung von $\varphi' = f(\cdot, \varphi)$ mit $\varphi(t_0) = x_0$ (φ differenzierbar) genau dann, wenn gilt:

$$\varphi(t) = x_0 + \int_{t_0}^t f(s, \varphi(s))\,\mathrm{d}s \qquad \forall\, t \in D.$$

Mit dieser Integralgleichung lässt sich aber ohne Weiteres keine Lösung direkt gewinnen, da ja φ auf beiden Seite auftritt. In Abschnitt 17.3.4 werden wir dennoch obige Gleichung benutzen, um Lösungen zu konstruieren (und zwar iterativ, und unter Zusatzbedingung an f).

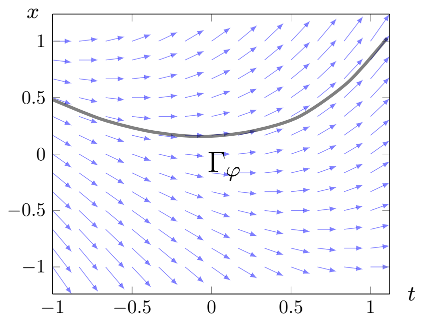
---

[151] Wir versehen M wieder mit einer eingeschränkten Normmetrik. Wir hatten ja in Kapitel 11.3 geklärt, dass es bezüglich Stetigkeit von Abbildungen etc. hierbei nicht auf die explizite Wahl der Norm auf $\mathbb{R}^2$ ankommt.

[152] Gilt (516), so ist also φ für jedes $(t_0, x_0) \in \Gamma_\varphi$ eine (lokale) Lösung von (516) zur Anfangsbedingung (515).

Bemerkung 111 (Richtungsfeld)**.** *Sei f wie in Definition 80. Anschaulich beschreibt f ein Anstiegsfeld (Richtungsfeld). Gesucht ist dann eine differenzierbare Abbildung $\varphi\colon \mathbb{R} \supseteq D \to \mathbb{R}$, sodass die Tangente an den Graphen $\Gamma_\varphi \subseteq M$ in $(t, \varphi(t))$ $(t \in D)$ gerade den Anstieg $f(t, \varphi(t))$ hat; d.h., für den Geschwindigkeitsvektor $\gamma'_\varphi = (1, \varphi')$ der Graphenkurve*

$$\gamma_\varphi\colon D \ni t \mapsto (t, \varphi(t)) \in M \subseteq \mathbb{R}^2 \qquad \textit{soll} \qquad \gamma'_\varphi = (1, f(\cdot, \varphi)) \qquad \textit{gelten.}$$



Am Punkt (t, x) ist der entsprechende blaue Richtungspfeil gegeben durch $(1, f(t, x))$.
Am Punkt $(t, \varphi(t))$ ist der entsprechende blaue Richtungspfeil gegeben durch $(1, f(t, \varphi(t)))$.

Manchmal kann man anhand des Richtungsfeldes auch Lösungen erraten:

Beispiel A: *Wir betrachten die Abbildung*

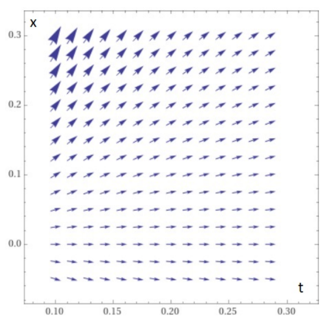
$$f\colon (0, \infty) \times \mathbb{R} \to \mathbb{R}, \qquad (t, x) \mapsto \frac{x}{t}. \tag{518}$$

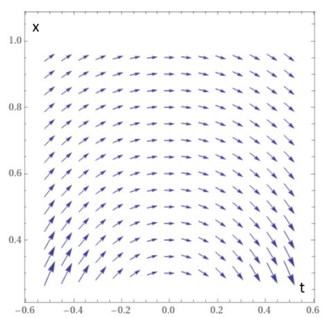
Das Richtungsfeld von f zeigt radial nach außen, und ist entlang der Geraden durch den Ursprung sogar konstant. Wie zu vermuten wäre, ist $\varphi_c\colon (0, \infty) \ni t \mapsto c \cdot t \in \mathbb{R}$ für jedes fixierte $c \in \mathbb{R}$ eine Lösung von $\varphi' = f(\cdot, \varphi)$ (nachrechnen). Gibt man nun eine Anfangsbedingung $\varphi_c(t_0) = x_0$ mit $(t_0, x_0) \in (0, \infty) \times \mathbb{R}$ vor, so erzwingt diese $c = \frac{x_0}{t_0}$.

Beispiel B: *Wir betrachten die Abbildung*

$$\tilde{f}\colon \mathbb{R} \times (0, \infty) \to \mathbb{R}, \qquad (t, x) \mapsto -\frac{t}{x}. \tag{519}$$

Das Richtungsfeld von g zeigt entlang von Halbkreisen. Wie zu vermuten wäre ist $\varphi_c\colon (-c, c) \ni t \mapsto \sqrt{c^2 - t^2} \in \mathbb{R}$ für jedes fixierte $c > 0$, eine Lösung von $\varphi' = \tilde{f}(\cdot, \varphi)$ (nachrechnen).[153] *Gibt man nun eine Anfangsbedingung $\varphi_c(t_0) = x_0$ mit $(t_0, x_0) \in \mathbb{R} \times (0, \infty)$ vor, so erzwingt diese $c = \sqrt{x_0^2 + t_0^2}$ (der Radius der Kreises in dem Γ_ϕ liegt).*





Links das Richtungsfelds $(1, f(t, x))$ von f, und rechts das Richtungsfeld $(1, \tilde{f}(t, x))$ von $\tilde{f}$.
(Auf dem gemeinsamen Definitionsbereich $(0, \infty) \times (0, \infty)$ gilt $\langle (1, f), (1, \tilde{f}) \rangle_{\mathrm{euk}} = 0$.)

[153] Es gilt $\Gamma_{\varphi_c} \subseteq M = \mathbb{R} \times (0, \infty)$ wegen $\mathrm{im}(\varphi) \subseteq (0, \infty)$. Es wäre nun zwar φ_c prinzipiell auch auf $[-c, c]$ definiert, aber wegen $\varphi_c(\pm c) = 0 \notin (0, \infty)$ wäre dies dann keine richtige Lösung von $\varphi' = \tilde{f}(\cdot, \varphi)$.

Terminologie 47 (Zeitunabhängigkeit). *Unter dem „zeitunabhängigen Fall“ verstehen wir im Folgenden die Situation, in der f nicht von t abhängt, d.h.*

$$f(s,x) = f(t,x) \quad \textit{für alle} \quad s,t,x \in \mathbb{R} \quad \textit{mit} \quad (s,x),(t,x) \in M.$$

Ist Beispielsweise $g\colon \mathbb{R} \supseteq B \to \mathbb{R}$ eine stetige Abbildung und $A \subseteq \mathbb{R}$ eine Teilmenge, so ist $f\colon M := A \times B \ni (t,x) \mapsto g(x) \in \mathbb{R}$ stetig und zeitunabhängig.

Wir geben nun zunächst einige zeitunabhängige Beispiele; und diskutieren dann einige Sonderfälle (spezielle Gestalt von f), die dann auch Zeitabhängigkeiten beinhalten.

Beispiel 112 (Natürliches Wachstum und Natürlicher Zerfall). *Sei $\lambda \in \mathbb{R}$ und* (zeitunabhängig)

$$f\colon M = \mathbb{R}^2 \to \mathbb{R}, \qquad (t,x) \mapsto \lambda \cdot x.$$

Sei $D \subseteq \mathbb{R}$ nichtentartet, sowie $(t_0, x_0) \in M$ mit $t_0 \in D$. Gemäß Korollar 33 ist

$$\varphi\colon D \to \mathbb{R}, \qquad t \mapsto x_0 \cdot \mathrm{e}^{\lambda \cdot (t-t_0)}$$

die eindeutige von $\varphi' = f(\cdot, \varphi)$ mit $\varphi(t_0) = x_0$ und $\mathrm{dom}(\varphi) = D$, d.h.

$$\varphi' = \lambda \cdot \varphi \qquad \textit{mit} \qquad \varphi(t_0) = x_0. \tag{520}$$

Die Gleichung (520) *wird auch als Gleichung des natürlichen Wachstums ($\lambda > 0$) bzw. des natürlichen Zerfalls ($\lambda < 0$) bezeichnet (oder auch als Populationsgleichung):*

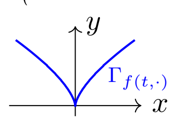
- *Gegeben sei eine Bakterienpopulation in einer Petrischale mit Grundfläche F zum Zeitpunkt $t \in \mathbb{R}$. Bei ausreichendem Nährstoffangebot, wird sich jede Bakterie in zeitlich periodischen Zyklen durch Zellteilung vermehren. Hat sich eine Bakterie gerade geteilt, so wird sich diese nach dem Zeitintervall Δt wieder teilen. Bei großer Bakterienanzahl N wird die Flächendichte $\eta(t) := N(t)/F$ ein sinnvolles Maß für die Bakterienanzahl $N(t)$ (zum Zeitpunkt t) in der Petrischale sein. Da sich nicht alle Bakterien zur gleichen Zeit teilen, sondern im Zeitintervall Δt statistisch gleichverteilte Teilungszeitpunkte haben, werden wir η in Abhängigkeit der Zeit wohl als differenzierbare Funktion $\mathbb{R} \to \mathbb{R}$ modellieren dürfen. Ist $0 < s \leq 1$, so werden sich nach dem Zeitinterval $s \cdot \Delta t$, gemessen zum Zeitpunkt t, gerade $s \cdot N(t)$ Bakterien vermehrt haben – Die Dichte $\eta(t)$ wird also ebenfalls um $s \cdot \eta(t)$ angewachsen sein, also $\eta(t + s \cdot \Delta t) = \eta(t) + s \cdot \eta(t)$. Wir erhalten:*

  $$\frac{\eta(t + s\cdot\Delta t) - \eta(t)}{s \cdot \Delta t} = \frac{s \cdot \eta(t)}{s \cdot \Delta t} = \frac{\eta(t)}{\Delta t} \qquad \forall\ 0 < s \leq 1.$$

  Grenzwertbildung $s \to 0$ liefert $\eta'(t) = \frac{\eta(t)}{\Delta t}$, also Gleichung (520) *sofern wir $\lambda := (\Delta t)^{-1} > 0$, $x_0 := \eta(t_0)$ und $\varphi := \eta$ setzen.*

- *Eine ähnliche Argumentation führt auf die Gleichung* (520) *mit negativem $\lambda < 0$, wenn man den Zerfall der Atome eines radioaktiven Materials beschreiben will (vorausgesetzt, die Anzahl der Atome ist sehr groß). Jedes Atom zerfällt zufällig, hat aber eine gewisse Zerfallswahrscheinlichkeit pro Zeiteinheit. Die Anzahl der pro Zeiteinheit zerfallenden Atome ist daher proportional zur Anzahl der noch nicht zerfallenen Atome.*

Beispiel 113 (Uneindeutigkeit von Lösungen). *Lösungen von Differentialgleichungen (erster Ordnung) sind nicht immer eindeutig.*[154] *Sei beispielsweise* (zeitunabhängig)



$$f\colon M = \mathbb{R}^2 \to \mathbb{R}, \qquad (t,x) \mapsto 3 \cdot \sqrt[3]{x^2}.$$

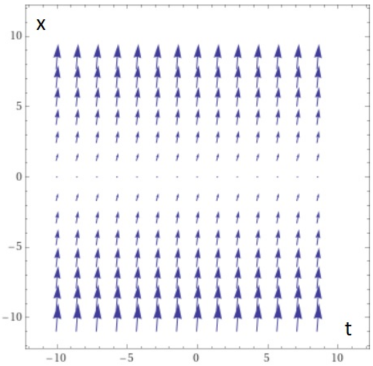
[154] Wir werden in Abschnitt 17.3 noch eine Zusatzvoraussetzung an f kennenlernen (lokale Lipschitz-Bedingung), die zumindest lokale Eindeutigkeit garantiert.

Die differenzierbaren Abbildungen:

$$\varphi_1\colon \mathbb{R}\ni t\mapsto 0\in\mathbb{R} \qquad \textit{sowie} \qquad \varphi_2\colon \mathbb{R}\ni t\mapsto (t-t_0)^3\in\mathbb{R}$$

$$\varphi_3\colon \mathbb{R}\ni t\mapsto \begin{cases} 0 & \textit{für } t\le t_0 \\ (t-t_0)^3 & \textit{für } t>t_0 \end{cases} \qquad \textit{sowie} \qquad \varphi_4\colon \mathbb{R}\ni t\mapsto \begin{cases} (t-t_0)^3 & \textit{für } t\le t_0 \\ 0 & \textit{für } t>t_0 \end{cases} \tag{521}$$

sind alle Lösungen von $\varphi' = f(\cdot,\varphi)$ *zur Anfangsbedingung* $\varphi(t_0)=x_0$ *für* $t_0\in\mathbb{R}$ *und* $x_0=0$.[155]



Das Richtungsfeld $(1, f(t,x))$ *von* f.

Beispiel 114 (Lokalität von Lösungen). *Zu einer vorgegeben Differentialgleichung samt Anfangsbedingungen gibt es oft nur lokale Lösungen, d.h., der Definitionsbereich einer Lösung kann nicht notwendig (durch Fortsetzen der Lösung) auf ein größeres Intervall erweitert werden:*

a) *Wir betrachten die stetige Abbildung* (zeitunabhängig)*:* $(\mathrm{pr}_1(M)=\mathbb{R})$

$$f\colon M=\mathbb{R}^2\to\mathbb{R}, \qquad x\mapsto 1+x^2.$$

Die differenzierbare Abbildung

$$\varphi\colon (-\tfrac{\pi}{2},\tfrac{\pi}{2})\to\mathbb{R}, \qquad t\mapsto \tan(t)$$

ist eine Lösung von $\varphi'=f(\cdot,\varphi)$ *mit* $\varphi(0)=0$ $(\tan' = 1+\tan^2)$*. Ihr Definitionsbereich lässt sich nicht erweitern; denn ist beispielsweise* $\tilde\varphi\colon (-\frac{\pi}{2},\frac{\pi}{2}]\to\mathbb{R}$ *differenzierbar (bzw. einfach nur stetig) mit* $\tilde\varphi|_{(-\frac{\pi}{2},\frac{\pi}{2})}=\varphi$*, so erhalten wir den Widerspruch*

$$\mathbb{R}\ni\tilde\varphi(\tfrac{\pi}{2}) = \lim_{t\uparrow\frac{\pi}{2}}\tilde\varphi(t) = \lim_{t\uparrow\frac{\pi}{2}}\varphi(t) = \lim_{t\uparrow\frac{\pi}{2}}\tan(t)=\infty.$$

b) *Wir betrachten den Kegel* $M:=\{(t,x)\in\mathbb{R}^2 \mid |x|\le t\}$*, und hierauf die konstante Abbildung* $f\colon M\ni(t,x)\mapsto\lambda\in\mathbb{R}$ *mit* $\lambda\in\mathbb{R}$ (zeitunabhängig)*. Wir fixieren* $(t_0,x_0)\in M$ *und setzen*

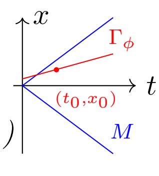
$$\phi\colon\mathbb{R}\to\mathbb{R}, \qquad t\mapsto x_0+\lambda\cdot(t-t_0).$$

Offensichtlich gilt dann $(\mathrm{pr}_1(M)=\mathbb{R})$

$$\phi'=\lambda=f(\cdot,\phi) \qquad \textit{mit} \qquad \phi(t_0)=x_0. \tag{522}$$

- *Sei* $|\lambda|\le 1$*, und* $\varphi:=\phi|_{[t_0,\infty)}$*. Dann ist* φ *eine Lösung von* $\varphi'=f(\cdot,\varphi)$ *mit* $\varphi(t_0)=x_0$*; und zwar wegen* (522)*, denn es gilt* $\Gamma_\varphi\subseteq M$*:*

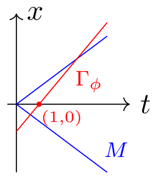
$$|\varphi(t)|\le|x_0|+|\lambda|\cdot(t-t_0) \overset{(t_0,x_0)\in M}{\le} t_0+(t-t_0)=t \qquad \forall\, t\in[t_0,\infty).$$



(Wie im Bild angedeutet, lässt sich φ *unter Umständen auf Zeiten* $<t_0$ *erweitern.)*

[155] Grob gesagt gibt es hier Probleme mit der Eindeutigkeit, weil $\lim_{x\to 0}\frac{|f(t,x)-f(t,0)|}{|x|}=\infty$ gilt (vgl. Beispiel 120.b)).

- *Sei $\lambda > 1$, sowie $t_0 = 1$ und $x_0 = 0$. Dann gilt*



$$(t, \phi(t)) \in M \quad \textit{für} \quad t \geq 1 \qquad \Leftrightarrow \qquad \lambda \cdot (t-1) \leq t \qquad \Leftrightarrow \qquad t \leq \tfrac{\lambda}{\lambda - 1},$$

  (Wie im Bild angedeutet, lässt sich φ auch auf Zeiten < 1 erweitern.)

  und somit wegen (522)*:*

  - $\varphi := \phi|_{[1,\lambda/(\lambda-1)]}$ *ist eine Lösung von $\varphi' = f(\cdot, \varphi)$ mit $\varphi(t_0) = x_0$.*
  - $\varphi_\varepsilon := \phi|_{[1,\lambda/(\lambda-1)+\varepsilon]}$ *mit $\varepsilon > 0$ ist keine Lösung von $\varphi' = f(\cdot, \varphi)$ mit $\varphi(t_0) = x_0$.*

Unter obiger Menge M mag man sich eine mit konstanter Rate in beide Richtungen ausdehnende/zufrierende (eindimensional) Eisfläche vorstellen, auf welcher zu einem Zeitpunkt t_0 ein Schlittschuhläufer in eine der beiden Richtungen bei x_0 startet und sich mit konstanter Geschwindigkeit fortbewegt. Lösungen obiger Differentialgleichung sind dann diejenigen φ's, bei denen sich der Schlittschuhläufer zu allen Zeiten (Definitionsbereich von φ) auf der (eindimensionalen) Eisfläche befindet (und nicht in den See fällt, und dann von der zufrierenden Eisdecke eingeschlossen wird ...). Dieses Beispiel verdeutlicht auch, dass es durchaus sinvoll sein kann, beliebige Teilmengen $M \subseteq \mathbb{R}^2$ zuzulassen (wie wir es ja getan haben) anstatt beispielsweise nur Teilmengen der Form $D_1 \times D_2$, mit Intervallen $D_1, D_2 \subseteq \mathbb{R}$ zu betrachten. Die spezielle Wahl der Menge M erlaubt es sozusagen gewisse Randbedingungen vorzugeben, die beispielsweise gewisse Eigenarten eines durch eine Differentialgleichung zu modellierenden Realsystems (physikalischen Systems) kodieren.

17.2 Spezielle Typen von Differentialgleichung erster Ordnung

Wir diskutieren nun einige Situationen, in denen die Abbildung f eine gewisse Form aufweist.

17.2.1 Differentialgleichungen mit getrennten Variablen

In diesem Abschnitt betrachten wir den Spezialfall, dass f faktorisiert:

Terminologie 48 (Differentialgleichung mit getrennten Variablen)**.** *Seien $D, B \subseteq \mathbb{R}$ nichtentartete Intervalle, sowie $h\colon D \to \mathbb{R}$ und $g\colon B \to \mathbb{R}$ stetige Abbildungen mit $g(x) \neq 0$ für alle $x \in B$. Wir betrachten die stetige Abbildung:*[156] *(beachte $\mathrm{pr}_1(M) = D$ und $\mathrm{pr}_2(M) = B$)*

$$f\colon M = D \times B \to \mathbb{R}, \qquad (t, x) \mapsto h(t) \cdot g(x).$$

Die Differentialgleichung (514) *ließt sich nun in der Form*

$$\varphi' = \underbrace{h \cdot (g \circ \varphi)}_{f(\cdot,\varphi)} \qquad (\textit{salopp } x' = h(t) \cdot g(x)), \tag{523}$$

und (515) *weiterhin als $\varphi(t_0) = x_0$ für $(t_0, x_0) \in M$.*

- *Eine Differentialgleichung der Form* (523) *heißt* Differentialgleichung mit getrennten Variablen*.*
- *Eine differenzierbare Abbildung $\varphi\colon \tilde{D} \to \mathbb{R}$ ist genau eine Lösung von* (523) *zur Anfangsbedingung $\varphi(t_0) = x_0$, wenn gilt:*

$$\begin{aligned}
\Gamma_\varphi \subseteq M \qquad &\overset{M = D \times B}{\Longleftrightarrow} \qquad \tilde{D} \subseteq D \quad \wedge \quad \mathrm{im}(\varphi) \subseteq B \\
(t_0, x_0) \in \Gamma_\varphi \qquad &\overset{M = D \times B}{\Longleftrightarrow} \qquad \varphi(t_0) = x_0 \quad \textit{mit} \quad t_0 \in \tilde{D},\ x_0 \in B \\
\varphi'(t) = h(t) \cdot g(\varphi(t)) \qquad &\qquad \forall\, t \in \tilde{D}.
\end{aligned}$$

[156] $f = (h \circ \mathrm{pr}_1|_M) \cdot (g \circ \mathrm{pr}_2|_M)$ ist als Produkt der stetigen Abbildungen $h \circ \mathrm{pr}_1|_M\colon M \to \mathbb{R}$ und $g \circ \mathrm{pr}_2|_M\colon M \to \mathbb{R}$ ebenfalls stetig, vgl. Beispiel 97.

Bemerkung 112. *Einen Lösungsansatz für* (523) *erhält man durch „Saloppumformungen":*

$\frac{\mathrm{d}x}{\mathrm{d}t} = h(t) \cdot g(x)$ „$\Rightarrow$" $\frac{\mathrm{d}x}{g(x)} = h(t)\,\mathrm{d}t$ „$\Rightarrow$" $\int_{\varphi(t_0)}^{\varphi(t)} \frac{\mathrm{d}x}{g(x)} = \int_{t_0}^{t} h(t)\,\mathrm{d}t$ „$\Rightarrow$" $\varphi = \big(\int_{x_0}^{\bullet} \frac{1}{g}\big)^{-1} \circ \big(\int_{t_0}^{\bullet} h\big)$.

Wir betrachten nun also die Stammfunktionen

$$H\colon D \ni t \mapsto \int_{t_0}^{t} h \in \mathbb{R} \qquad \text{und} \qquad G\colon B \ni x \mapsto \int_{x_0}^{x} \frac{1}{g} \in \mathbb{R}, \tag{i}$$

und bemerken Folgendes:

a) *Nach dem Hauptsatz sind H und G stetig differenzierbar (also stetig) mit $H' = h$ und $G' = \frac{1}{g}$. Zudem ist G streng monoton; also injektiv wegen Lemma 43:*
   - *Es gilt $\frac{1}{g} > 0$ oder $\frac{1}{g} < 0$, denn ansonsten liefert der Zwischenwertsatz einen Widerspruch zur Nullstellenfreiheit von $\frac{1}{g}$.*
   - *Gilt $\frac{1}{g} > 0/\frac{1}{g} < 0$, so ist G streng monoton wachsend/fallend wegen Korollar 32.3)* $\big(G' = \frac{1}{g}\big)$.

b) *$\tilde{B} := G(B)$ ist ein nichtentartetes Intervall mit $0 = G(x_0) \in \tilde{B}$, und*

$$L := \big(G|^{\tilde{B}}\big)^{-1}\colon \tilde{B} \to B \text{ ist stetig differenzierbar mit } L'(q) = g(L(q)) \text{ für alle } q \in \tilde{B}. \tag{ii}$$

In der Tat:
   - *Wegen Korollar 28 ist $\tilde{B} = G(B)$ ein Intervall; und da G injektiv und B nichtentartet ist, ist auch $\tilde{B}$ nichtentartet. Zudem ist L stetig wegen Satz 37.*
   - *Wegen Satz 42 ist L differenzierbar ($G' = \frac{1}{g}$ hat keine Nullstellen) mit*

$$L'(G(p)) = \Big(\big(G|^{\tilde{B}}\big)^{-1}\Big)'(G(p)) = \frac{1}{G'(p)} = g(p) \qquad \forall\, p \in B.$$

   Die Substitution $p \equiv L(q)$ mit $q \in \tilde{B}$ liefert (ii). *Also ist $L' = g \circ L$ als Verkettung stetiger Funktionen ebenfalls stetig.*

c) *Ist B offen, so existiert $\tilde{D} \subseteq D$ nichtentartet mit $t_0 \in \tilde{D}$ und $H(\tilde{D}) \subseteq \tilde{B} = G(B)$:*[157]

Beweis. • Es gilt $H(t_0) = 0 = G(x_0) \in \tilde{B}$.
   - $\tilde{B}$ ist offen, da G streng monoton ist.[158] Wegen $0 = G(x_0) \in \tilde{B}$, existiert $\varepsilon > 0$ mit $(-\varepsilon, \varepsilon) \subseteq \tilde{B}$.
   - Da $H\colon D \to \mathbb{R}$ stetig ist mit $H(t_0) = 0 \in \tilde{B}$, existiert $\delta > 0$ mit $H(\tilde{D}) \subseteq (-\varepsilon, \varepsilon) \subseteq \tilde{B}$, für das (wegen Übung 91) nichtentartete Intervall $\tilde{D} := (t_0 - \delta, t_0 + \delta) \cap D$. □

Satz 103. *Seien $D, B \subseteq \mathbb{R}$ nichtentartete Intervalle, sowie $h\colon D \to \mathbb{R}$ und $g\colon B \to \mathbb{R}$ stetige Abbildungen mit $g(x) \neq 0$ für alle $x \in D$. Sei $(t_0, x_0) \in D \times B$ vorgegeben, und seien H, G, L wie in* (i) *und* (ii) *definiert. Sei weiterhin $\tilde{D} \subseteq D$ nichtentartet mit $t_0 \in \tilde{D}$ und $H(\tilde{D}) \subseteq G(B) = \tilde{B}$ (vgl. Bemerkung 112.c)). Die C^1-Abbildung*

$$\varphi := L \circ H|_{\tilde{D}}\colon \tilde{D} \to B$$

ist die eindeutige Lösung der Differentialgleichung $\varphi' = h \cdot (g \circ \varphi)$ zur Anfangsbedingung $\varphi(t_0) = x_0$ mit $\mathrm{dom}(\varphi) = \tilde{D}$.

Beweis. Es gilt $\varphi \in C^1(\tilde{D}, \mathbb{R})$ da L und H stetig differenzierbar sind.

[157] Beispielsweise existiert für $D = [0,1]$, $B = [0,1]$, $t_0 = 1$, $x_0 = 0$, $h = 1$, $g = 1$ kein derartiges Intervall $\tilde{D} \subseteq D$; denn es gilt ja $H\colon [0,1] \ni t \mapsto t - 1 \in [-1, 0]$ und $G\colon [0,1] \ni t \mapsto t \in [0,1]$; also $H([0,1]) \cap G([0,1]) = \{0\}$.

[158] Sei o.B.d.A. G streng monoton wachsend. Für $y \in \tilde{B}$ sei $x \in B$ mit $G(x) = y$. Wegen der Offenheit von B existiert ein $\varepsilon > 0$ mit $(x - 2\varepsilon, x + 2\varepsilon) \subseteq B$; also $G(x - \varepsilon) < G(x) < G(x + \varepsilon)$. Die Intervalleigenschaft von $\tilde{B}$ impliziert $y = G(x) \in (G(x - \varepsilon), G(x + \varepsilon)) \subseteq \tilde{B}$.

- Lösungseigenschaft: Es gilt $\varphi(t_0) = L(H(t_0)) = L(0) = G^{-1}(0) = x_0$, und die Kettenregel liefert:

$$\varphi'(t) = H'(t) \cdot L'(H(t)) \overset{\text{(i),(ii)}}{=} h(t) \cdot g((L \circ H)(t)) = h(t) \cdot g(\varphi(t)) \qquad \forall\, t \in \tilde{D}.$$

  Weiterhin gilt $\Gamma_\varphi \subseteq \tilde{D} \times B \subseteq D \times B = M$.
- Eindeutigkeit: Sei $\tilde{\varphi}\colon \tilde{D} \to \mathbb{R}$ mit $\Gamma_{\tilde{\varphi}} \subseteq D \times B$ sowie $\tilde{\varphi}' = h \cdot (g \circ \tilde{\varphi})$ und $\tilde{\varphi}(t_0) = x_0$ $((t_0, x_0) \in \Gamma_{\tilde{\varphi}})$. Die Kettenregel liefert:

$$(G \circ \tilde{\varphi})'(t) = G'(\tilde{\varphi}(t)) \cdot \tilde{\varphi}'(t) \overset{\text{(i)}}{=} \frac{1}{g(\tilde{\varphi}(t))} \cdot h(t) \cdot g(\tilde{\varphi}(t)) = h(t) \qquad \forall\, t \in \tilde{D}.$$

  Wegen $(G \circ \tilde{\varphi})(t_0) = G(x_0) = 0$ zeigt der Hauptsatz, dass $G \circ \tilde{\varphi} = H|_{\tilde{D}}$ gilt; also $\tilde{\varphi} = L \circ H = \varphi$ wegen $H(\tilde{D}) \subseteq \tilde{B}$. □

Beispiel 115. *Wir betrachten die Funktionen aus Bemerkung 111:*

a) *Sei $D = (0, \infty) = B$, $(t_0, x_0) \in D \times B$, sowie (Einschränkung von* (518) *auf $D \times B$):*

$$f\colon D \times B \to \mathbb{R}, \qquad (t, x) \mapsto \frac{x}{t} = h(t) \cdot g(x)$$

mit $h\colon D \ni t \mapsto \frac{1}{t} \in D$ und $g = \mathrm{id}_B\colon B \to B$. Dann gilt mit $\tilde{B} := \mathbb{R}$:

$$\begin{aligned} H\colon D &\to \mathbb{R}, & t &\mapsto \log(t) - \log(t_0) \\ G\colon B &\to \tilde{B}, & x &\mapsto \log(x) - \log(x_0) \\ L\colon \tilde{B} &\to B, & y &\mapsto \mathrm{e}^{y + \log(x_0)}. \end{aligned}$$

Für $\tilde{D} := D$ gilt dann $t_0 \in \tilde{D}$ und $H(\tilde{D}) = H(D) = \mathbb{R} = \tilde{B}$. Satz 103 zeigt, dass

$$\varphi = L \circ H|_{\tilde{D}}\colon (0, \infty) \to (0, \infty), \qquad t \mapsto \mathrm{e}^{H(t) + \log(x_0)} = \frac{x_0}{t_0} \cdot t$$

die eindeutige Lösung von $\varphi' = h \cdot (g \circ \varphi)$ mit $\varphi(t_0) = x_0$ und $\mathrm{dom}(\varphi) = \tilde{D} = D = (0, \infty)$ ist – Im Einklang mit Beispiel A in Bemerkung 111.[159]

b) $*$ *Sei $D = \mathbb{R}$, $B = (0, \infty)$, $(t_0, x_0) \in D \times B$, sowie (siehe* (519)*):*

$$\tilde{f}\colon D \times B \to \mathbb{R}, \qquad (t, x) \mapsto -\frac{t}{x} = \tilde{h}(t) \cdot \tilde{g}(x)$$

mit $\tilde{h} = -\mathrm{id}_D\colon D \to D$ und $\tilde{g}\colon B \ni x \mapsto \frac{1}{x} \in B$. Dann gilt mit $\tilde{B} := \big(-\frac{x_0^2}{2}, \infty\big)$:

$$\begin{aligned} \tilde{H}\colon D &\to \mathbb{R}, & t &\mapsto -\tfrac{t^2}{2} + \tfrac{t_0^2}{2} \\ \tilde{G}\colon B &\to \tilde{B}, & x &\mapsto \tfrac{x^2}{2} - \tfrac{x_0^2}{2} \\ \tilde{L}\colon \tilde{B} &\to B, & y &\mapsto (2y + x_0^2)^{\frac{1}{2}} \end{aligned}$$

Für $c := \sqrt{x_0^2 + t_0^2} \subseteq \mathbb{R}$ und $\tilde{D} := (-c, c)$, gilt dann $t_0 \in \tilde{D}$ und $\tilde{H}(\tilde{D}) = \big(-\frac{x_0^2}{2}, \frac{t_0^2}{2}\big) \subseteq \tilde{B}$. Satz 103 zeigt, dass

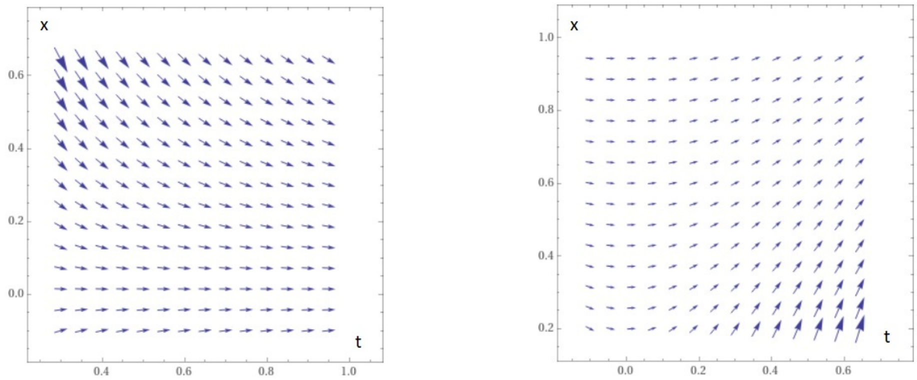
$$\varphi = \tilde{L} \circ \tilde{H}|_{\tilde{D}}\colon (-c, c) \to (0, c), \qquad t \mapsto \sqrt{c^2 - t^2}$$

die eindeutige Lösung von $\varphi' = \tilde{f}(\cdot, \varphi)$ mit $\varphi(t_0) = x_0$ und $\mathrm{dom}(\varphi) = (-c, c)$ ist – Im Einklang mit Beispiel B in Bemerkung 111.

[159] Analog kann man natürlich auch den Fall $D = (0, \infty)$ und $B = (-\infty, 0)$ behandeln.

Übung 167. *Sei Sei $D = (0, \infty) = B$, $(t_0, x_0) \in D \times B$, sowie*

$$f \colon D \times B \to \mathbb{R}, \qquad (t, x) \mapsto -\frac{x}{t}$$
$$\tilde{f} \colon D \times B \to \mathbb{R}, \qquad (t, x) \mapsto \frac{t}{x}.$$



Links das Richtungsfeld $(1, f(t,x))$ von f, und rechts das Richtungsfeld $(1, \tilde{f}(t,x))$ von $\tilde{f}$. (Auf dem gemeinsamen Definitionsbereich $(0,\infty) \times (0,\infty)$ gilt $\langle (1,f),(1,\tilde{f})\rangle_{\mathrm{euk}} = 0$.)

a) Für $(t_0, x_0) \in D \times B$ vorgegeben, bestimmen Sie die eindeutige Lösung von $\varphi' = f(\cdot, \varphi)$ mit $\varphi(t_0) = x_0$ und $\mathrm{dom}(\varphi) = D = (0, \infty)$.

b) Für $(t_0, x_0) \in D \times B$ vorgegeben, bestimmen Sie die eindeutige Lösung von $\varphi' = \tilde{f}(\cdot, \varphi)$ mit $\varphi(t_0) = x_0$ und

- $x_0 \geq t_0$: $\mathrm{dom}(\varphi) = \tilde{D} := D = (0, \infty)$,
- $x_0 < t_0$: $\mathrm{dom}(\varphi) = \tilde{D} := \left(\sqrt{t_0^2 - x_0^2}, \infty\right)$.

Machen Sie sich die beiden Fälle anhand einer Skizze des Graphen von φ klar.

17.2.2 Lineare Differentialgleichungen

In diesem Abschnitt betrachten wir den Spezialfall linearer Differentialgleichungen:

Terminologie 49 (Lineare Differentialgleichung)**.** *Sei $D \subseteq \mathbb{R}$ nichtentartet, und $\alpha, \beta \colon D \to \mathbb{R}$ stetig. Dann ist die Abbildung*

$$f \colon M = D \times \mathbb{R} \to \mathbb{R}, \qquad (t, x) \mapsto \alpha(t) \cdot x + \beta(t)$$

ebenfalls stetig; und die Differentialgleichung (514) *ließt sich in der Form:*

$$\varphi' = \underbrace{\alpha \cdot \varphi + \beta}_{f(\cdot,\varphi)} \qquad (\textit{salopp } x' = \alpha(t) \cdot x + \beta(t)). \tag{524}$$

Eine differenzierbare Abbildung $\varphi \colon \tilde{D} \to \mathbb{R}$ ist genau eine Lösung von (524) *zur Anfangsbedingung $\varphi(t_0) = x_0$, wenn gilt:*

$$\begin{array}{ccl} \Gamma_\varphi \subseteq M & \overset{M = D\times\mathbb{R}}{\Longleftrightarrow} & \tilde{D} \subseteq D \\ (t_0, x_0) \in \Gamma_\varphi & \overset{M = D\times\mathbb{R}}{\Longleftrightarrow} & \varphi(t_0) = x_0 \quad \textit{mit} \quad t_0 \in \tilde{D},\ x_0 \in \mathbb{R} \\ \varphi'(t) = \alpha(t) \cdot \varphi(t) + \beta(t) & & \forall\, t \in \tilde{D}. \end{array}$$

Eine Differentialgleichung der Form (524) *heißt* lineare Differentialgleichung erster Ordnung*; und zwar*

- homogen *falls* $\beta = 0$ *gilt,*
- inhomogen *falls* $\beta \neq 0$ *gilt.*

Lemma 91 (Homogene lineare Differentialgleichung)**.** *Sei $D \subseteq \mathbb{R}$ nichtentartet, $\alpha\colon D \to \mathbb{R}$ stetig, sowie $\tilde{D} \subseteq D$ nichtentartet und $(t_0, x_0) \in \tilde{D} \times \mathbb{R}$. Die C^1-Abbildung*

$$\varphi\colon \tilde{D} \to \mathbb{R}, \qquad t \mapsto x_0 \cdot \exp_{\mathbb{R}}\big(\textstyle\int_{t_0}^{t} \alpha\big) \tag{525}$$

ist die eindeutige Lösung der homogenen linearen Differentialgleichung $\varphi' = \alpha \cdot \varphi$ zur Anfangsbedingung $\varphi(t_0) = x_0$, mit $\mathrm{dom}(\varphi) = \tilde{D}$.

Beweis. Offensichtlich gilt $\varphi(t_0) = x_0$; und der Hauptsatz sowie die Kettenregel liefern $\varphi \in C^1(\tilde{D}, \mathbb{R})$ mit $\varphi' = \alpha \cdot \varphi$. Für die Eindeutigkeit sei $\tilde{\varphi}\colon \tilde{D} \to \mathbb{R}$ differenzierbar mit $\tilde{\varphi}' = \alpha \circ \tilde{\varphi}$ und $\tilde{\varphi}(t_0) = x_0$. Die Abbildung $\eta := \tilde{\varphi} \cdot \exp_{\mathbb{R}}\big(-\int_{t_0}^{\bullet} \alpha\big)$ ist differenzierbar (übliche Ableitungsregeln); und wir erhalten

$$\begin{aligned} \eta' = \alpha \cdot \tilde{\varphi} \cdot \exp_{\mathbb{R}}\big(-\textstyle\int_{t_0}^{\bullet} \alpha\big) + \tilde{\varphi} \cdot (-\alpha) \cdot \exp_{\mathbb{R}}\big(-\textstyle\int_{t_0}^{\bullet} \alpha\big) = 0 \qquad &\Longrightarrow \qquad \eta = \eta(t_0) = x_0 \\ &\Longrightarrow \qquad \tilde{\varphi} = x_0 \cdot \exp_{\mathbb{R}}\big(\textstyle\int_{t_0}^{\bullet} \alpha\big) = \varphi \end{aligned}$$

mit Korollar 32.2) (erste Implikation) sowie Satz 30.b) (zweite Implikation). □

Beispiel 116. *a) Den Fall $\alpha = \lambda \in \mathbb{R}$ konstant hatten wir bereits in Beispiel 112 diskutiert.*

*b) Im Einklang mit mit Beispiel A in Bemerkung 111 (bzw. Beispiel 115.*a)*), also*

$$f\colon (0, \infty) \times \mathbb{R} \to \mathbb{R}, \qquad (t, x) \mapsto \frac{x}{t} = \alpha(t) \cdot x$$

mit $\alpha\colon \mathbb{R} \ni t \mapsto \frac{1}{t} \in (0, \infty)$, liefert Lemma 91 für $(t_0, x_0) \in \tilde{D} \times \mathbb{R}$ mit $\tilde{D} \subseteq D = (0, \infty)$ die eindeutige Lösung

$$\varphi\colon \tilde{D} \to \mathbb{R}, \qquad t \mapsto x_0 \cdot \exp_{\mathbb{R}}\big(\textstyle\int_{t_0}^{t} \frac{1}{s}\, \mathrm{d}s\big) = x_0 \cdot \exp_{\mathbb{R}}\big(\log(t) - \log(t_0)\big) = \frac{x_0}{t_0} \cdot t$$

zur Differentialgleichung $\varphi' = \alpha \cdot \varphi$ mit $\varphi(t_0) = x_0$ sowie $\mathrm{dom}(\varphi) = \tilde{D}$.

Den allgemeinen inhomogenen Fall behandelt man, indem man sich vom homogenen Fall inspirieren lässt: Man ersetzt die Konstante x_0 in (525) durch eine differenzierbare Funktion $u\colon D \to \mathbb{R}$ mit $u(t_0) = x_0$ (Variation der Konstanten), und leitet mutig ab und stellt um:

$$\begin{aligned} \varphi := u \cdot \underbrace{\exp_{\mathbb{R}}\big(\textstyle\int_{t_0}^{\bullet} \alpha\big)}_{=:\, \varphi_0} \qquad &\Longrightarrow \qquad \alpha \cdot \varphi + \beta \overset{!}{=} \varphi' = u' \cdot \varphi_0 + \alpha \cdot \underbrace{u \cdot \varphi_0}_{\varphi} \\ &\Longrightarrow \qquad u' = \varphi_0^{-1} \cdot \beta \\ &\overset{u(t_0) = x_0}{\Longrightarrow} \qquad u = x_0 + \textstyle\int_{t_0}^{\bullet} (\varphi_0^{-1} \cdot \beta) \quad \text{mit} \quad \varphi_0^{-1} = \exp_{\mathbb{R}}\big(-\textstyle\int_{t_0}^{\bullet} \alpha\big). \end{aligned} \tag{526}$$

Satz 104 (Inhomogene lineare Differentialgleichung)**.** *Sei $D \subseteq \mathbb{R}$ nichtentartet, $\alpha, \beta\colon D \to \mathbb{R}$ stetig und $\tilde{D} \subseteq D$ nichtentartet und $(t_0, x_0) \in \tilde{D} \times \mathbb{R}$. Dann ist die C^1-Abbildung*

$$\varphi\colon \tilde{D} \to \mathbb{R}, \qquad t \mapsto \underbrace{\big(x_0 + \textstyle\int_{t_0}^{t} (\varphi_0^{-1} \cdot \beta)\big)}_{=:\, u} \cdot \underbrace{\exp_{\mathbb{R}}\big(\textstyle\int_{t_0}^{t} \alpha\big)}_{=:\, \varphi_0} \tag{527}$$

die eindeutige Lösung der Differentialgleichung $\varphi' = \alpha \cdot \varphi + \beta$ zu der Anfangsbedingung $\varphi(t_0) = x_0$ mit $\mathrm{dom}(\varphi) = \tilde{D}$.

Beweis. Es gilt $\varphi \in C^1(\tilde{D}, \mathbb{R})$ wegen der Produktregel; denn es gilt $u, \varphi_0, \varphi_0^{-1} \in C^1(\tilde{D}, \mathbb{R})$ wegen dem Hauptsatz, der Produktregel, der Kettenregel, und der Glattheit von $\exp_{\mathbb{R}}$. Offensichtlich gilt $\varphi(t_0) = x_0$; und die Lösungseigenschaft verifiziert man mit den üblichen Ableitungsregeln:

$$\varphi' = u \cdot \varphi_0' + u' \cdot \varphi_0 = \alpha \cdot \underbrace{u \cdot \varphi_0}_{= \varphi} + (\varphi_0^{-1} \cdot \beta) \cdot \varphi_0 = \alpha \cdot \varphi + \beta.$$

Für die Eindeutigkeit, sei $\tilde{\varphi} \colon \tilde{D} \to \mathbb{R}$ differenzierbar mit $\tilde{\varphi}' = \alpha \cdot \tilde{\varphi} + \beta$ und $\tilde{\varphi}(t_0) = x_0$. Wir setzen $\tilde{u} := \tilde{\varphi} \cdot \varphi_0^{-1}$, also $\tilde{\varphi} = \tilde{u} \cdot \varphi_0$, und erhalten wegen $\tilde{u}(t_0) = \tilde{\varphi}(t_0) = x_0$ mit dem Hauptsatz (Satz 57):

$$\tilde{u}' = \overbrace{(\alpha \cdot \tilde{\varphi} + \beta)}^{= \tilde{\varphi}'} \cdot \varphi_0^{-1} + \tilde{\varphi} \cdot \overbrace{(-\alpha \cdot \varphi_0^{-1})}^{= (\varphi_0^{-1})'} = \beta \cdot \varphi_0^{-1} \quad \overset{\text{Satz 57}}{\Rightarrow} \quad \tilde{u} = \overbrace{x_0 + \textstyle\int_{t_0}^{\bullet} \varphi_0^{-1} \cdot \beta}^{= u} \quad \Rightarrow \quad \tilde{\varphi} = \overbrace{u \cdot \varphi_0}^{= \varphi}. \quad \square$$

Beispiel 117. *Sei $\tilde{D} = \mathbb{R} = D$, $t_0, x_0 \in \mathbb{R}$, sowie $\alpha \colon \mathbb{R} \ni t \mapsto 2t \in \mathbb{R}$ und $\beta \colon \mathbb{R} \ni t \mapsto t^3 \in \mathbb{R}$. Wir betrachten die inhomogene lineare Differentialgleichung*

$$\varphi' = \alpha \cdot \varphi + \beta \quad \textit{mit} \quad \varphi(t_0) = x_0 \qquad\qquad (\textit{salopp } x' = 2t \cdot x + t^3).$$

Wir erhalten $\varphi_0(t) = \exp_{\mathbb{R}}\big(\int_{t_0}^t 2s \,\mathrm{d}s\big) = \exp_{\mathbb{R}}\big(t^2 - t_0^2\big) = \frac{\mathrm{e}^{t^2}}{\mathrm{e}^{t_0^2}}$ für alle $t \in \mathbb{R}$, also

$$\begin{aligned}\varphi(t) = \frac{\mathrm{e}^{t^2}}{\mathrm{e}^{t_0^2}} \cdot \Bigg(x_0 + \mathrm{e}^{t_0^2} \cdot \underbrace{\int_{t_0}^t \frac{s^3}{\mathrm{e}^{s^2}} \,\mathrm{d}s}_{\left[-\frac{s^2+1}{2} \cdot \mathrm{e}^{-s^2}\right]_{t_0}^t} \Bigg) \quad &= \quad \frac{\mathrm{e}^{t^2}}{\mathrm{e}^{t_0^2}} \cdot x_0 + \mathrm{e}^{t^2} \cdot \Bigg(-\frac{t^2+1}{2} \cdot \mathrm{e}^{-t^2} + \frac{t_0^2+1}{2} \cdot \mathrm{e}^{-t_0^2} \Bigg) \\ &= \quad \frac{\mathrm{e}^{t^2}}{\mathrm{e}^{t_0^2}} \cdot \Big(x_0 + \frac{t_0^2+1}{2} \Big) - \frac{t^2+1}{2}.\end{aligned}$$

17.2.3 Homogene Differentialgleichungen*

Terminologie 50 (Homogene Differentialgleichung)**.** *Sei $B \subseteq \mathbb{R}$ ein nichtentartetes Intervall und $g \colon B \to \mathbb{R}$ stetig. Dann ist die Abbildung*

$$f \colon M_B := \Big\{ (t,x) \in \mathbb{R}_{\times} \times \mathbb{R} \;\Big|\; \frac{x}{t} \in B \Big\}, \qquad (t,x) \mapsto g\Big(\frac{x}{t}\Big)$$

ebenfalls stetig; und die Differentialgleichung (514) *ließt sich in der Form:*

$$\varphi' = f(\cdot, \varphi) = g\Big(\frac{\varphi}{\cdot}\Big) \qquad\qquad \Big(\textit{salopp } x' = \Big(\frac{x}{t}\Big)\Big). \tag{528}$$

Eine differenzierbare Abbildung $\varphi \colon \tilde{D} \to \mathbb{R}$ ($\tilde{D}$ nichtentartet) ist somit eine Lösung von (528) *zur Anfangsbedingung $\varphi(t_0) = x_0$ genau dann, wenn gilt:*

$$\begin{aligned} \Gamma_\varphi \subseteq M_B \qquad \textit{also} \qquad \tilde{D} \subseteq \mathbb{R}_{\times} \;\; \wedge \;\; \big\{ \tfrac{\varphi(t)}{t} \;\big|\; t \in \tilde{D} \big\} \subseteq B \\ \varphi(t_0) = x_0 \quad \textit{für} \quad (t_0, x_0) \in \tilde{D} \times \mathbb{R} \quad \textit{mit} \quad \tfrac{x_0}{t_0} \in B \\ \varphi'(t) = g\big(\tfrac{\varphi(t)}{t}\big) \qquad \forall\, t \in \tilde{D}. \end{aligned} \tag{529}$$

Eine Differentialgleichung der Form (528) *heißt* homogene Differentialgleichung erster Ordnung.[160]

Die Lösungstheorie homogener Differentialgleichungen lässt sich auf die Lösungstheorie von Differentialgleichungen mit getrennten Variablen zurückführen:

[160]Homogene Differentialgleichungen haben prinzipiell nichts mit homogenen linearen Differentialgleichungen zu tun (Verwechslungsgefahr). Die Bezeichnung „homogen“ kommt daher, dass $f(\lambda t, \lambda x) = f(t,x)$ für $\lambda \in \mathbb{R}_{\times}$ und $(t,x) \in M$ mit $(\lambda t, \lambda x) \in M$ gilt, d.h. f ist eine sogenannte homogene Funktion.

Kurzfassung: Gegeben sei eine Differentialgleichung in Saloppform $x' = g(\frac{x}{t})$ mit $g\colon B \to \mathbb{R}$ stetig ($B \subseteq \mathbb{R}$ nichtentartet). Man erhält eine Lösung $\varphi\colon \mathbb{R}_\times \supseteq \tilde{D} \to B$ zur Anfangsbedingung $\varphi(t_0) = x_0$ für $(t_0, x_0) \in \tilde{D} \times \mathbb{R}$ mit $\frac{x_0}{t_0} \in B$, indem man zunächst durch Substitution informell umformt:

$$z := \frac{x}{t} \qquad \Rightarrow \qquad z + t \cdot z' = x' = g\Big(\frac{x}{t}\Big) = g(z) \qquad \Rightarrow \qquad z' = \frac{g(z) - z}{t}; \tag{530}$$

dann eine Lösung $\hat{\varphi}\colon \mathbb{R}_\times \supseteq \tilde{D} \to B$ mit $\hat{\varphi}(t_0) = \frac{x_0}{t_0}$ der rechten Seite (Differentialgleichung mit getrennten Variablen) berechnet, und schließlich $\varphi\colon \tilde{D} \ni t \mapsto t \cdot \hat{\varphi} \in \mathbb{R}$ setzt. In der Tat gilt dann ja $\varphi(t_0) = t_0 \cdot \frac{x_0}{t_0} = x_0$, sowie für $t \in \tilde{D}$:

$$\begin{aligned} (t, \varphi(t)) = (t, t \cdot \hat{\varphi}(t)) \in M_B \qquad &\text{wegen} \qquad \tfrac{\varphi(t)}{t} = \hat{\varphi(t)} \in B \\ \varphi'(t) = \hat{\varphi}(t) + t \cdot \hat{\varphi}'(t) = \hat{\varphi}(t) &+ g(\hat{\varphi}(t)) - \varphi(t) = g\big(\tfrac{\varphi(t)}{t}\big). \end{aligned}$$

Beispiel: Sei $(t_0, x_0) \in \mathbb{R}_\times \times \mathbb{R}$, sowie $g\colon B = \mathbb{R} \ni z \mapsto 1 + z + z^2 \in \mathbb{R}$. Dann ließt sich (528) salopp

$$x' = 1 + \tfrac{x}{t} + \big(\tfrac{x}{t}\big)^2 \qquad \text{sowie die rechte Seite von (530) in der Form} \qquad \tfrac{\mathrm{d}z}{\mathrm{d}t} = z' = \tfrac{1+z^2}{t}.$$

Eine Saloppumformung wie in Bemerkung 112 liefert:

$$\begin{aligned} \tfrac{\mathrm{d}z}{1+z^2} = \tfrac{\mathrm{d}t}{t} \quad &\text{„}\Rightarrow\text{“} \quad \arctan(\hat{\varphi}(t)) - \arctan(\hat{\varphi}(t_0)) = \int_{t_0}^{\hat{\varphi}(t)} \tfrac{\mathrm{d}z}{1+z^2} = \int_{t_0}^{t} \tfrac{1}{t}\,\mathrm{d}t = \overbrace{\log(t) - \log(t_0)}^{\log(\frac{t}{t_0})} \\ &\text{„}\Rightarrow\text{“} \quad \hat{\varphi}(t) = \tan\big(\log\big(\tfrac{t}{t_0}\big) + \arctan\big(\tfrac{x_0}{t_0}\big)\big) \\ &\text{„}\Rightarrow\text{“} \quad \varphi(t) = t \cdot \tan\big(\underbrace{\log\big(\tfrac{t}{t_0}\big) + \arctan\big(\tfrac{x_0}{t_0}\big)}_{=:\,\Lambda(t)}\big) \end{aligned}$$

für alle t in einem nichtentarteten Intervall $\tilde{D} \subseteq \mathbb{R}_\times$ mit $t_0 \in \tilde{D}$ derart, dass $\Lambda(\tilde{D}) \subseteq \big(-\frac{\pi}{2}, \frac{\pi}{2}\big)$ gilt. $\big($Beachte: $\Lambda\colon \mathbb{R}_\times \ni t \mapsto \log\big(\frac{t}{t_0}\big) + \arctan\big(\frac{x_0}{t_0}\big) \in \mathbb{R}$ ist stetig mit $\Lambda(t_0) = \arctan\big(\frac{x_0}{t_0}\big) \in \big(-\frac{\pi}{2}, \frac{\pi}{2}\big)$.$\big)$ □

Ausführliche Darstellung: In der Situation von Terminologie 50 sei:

$$\begin{aligned} \hat{f}\colon \mathbb{R}_\times \times B \to \mathbb{R}, \qquad &(t, x) \mapsto \hat{h}(t) \cdot \hat{g}(x) = \frac{g(x) - x}{t} \\ &\text{mit} \\ \hat{h}\colon \mathbb{R}_\times \ni t \mapsto \tfrac{1}{t} \in \mathbb{R} \qquad &\text{und} \qquad \hat{g} := g - \mathrm{id}_B\colon B \ni x \mapsto g(x) - x \in \mathbb{R}. \end{aligned} \tag{531}$$

Sei nun $\tilde{D} \subseteq \mathbb{R}_\times$ nichtentartet, sowie $\varphi, \hat{\varphi}\colon \tilde{D} \to \mathbb{R}$ mit

$$\hat{\varphi}(t) = \tfrac{\varphi(t)}{t} \quad \forall\, t \in \tilde{D} \qquad \text{bzw.} \qquad \varphi(t) = t \cdot \hat{\varphi}(t) \quad \forall\, t \in \tilde{D}. \tag{532}$$

Dann gilt:

- φ (stetig) differenzierbar $\Leftrightarrow$ $\hat{\varphi}$ (stetig) differenzierbar, (Quotientenregel und Produktregel)
- $\Gamma_\varphi \subseteq M_B \quad \Leftrightarrow \quad \Gamma_{\hat{\varphi}} \subseteq \tilde{D} \times B$, (erste Zeile in (529))
- $\varphi(t_0) = x_0$ für $(t_0, x_0) \in \tilde{D} \times \mathbb{R}$ mit $\frac{x_0}{t_0} \in B \quad \Leftrightarrow \quad \hat{\varphi}(t_0) = \frac{x_0}{t_0}$ mit $\big(t_0, \frac{x_0}{t_0}\big) \in \tilde{D} \times B$.
- Für jedes $t \in \tilde{D}$ gilt:

$$\varphi'(t) = g\big(\tfrac{\varphi(t)}{t}\big) = f(t, \varphi(t)) \quad \Leftrightarrow \quad \hat{\varphi}(t) + t \cdot \hat{\varphi}'(t) = g(\hat{\varphi}(t)) \quad \Leftrightarrow \quad \hat{\varphi}'(t) = \tfrac{g(\hat{\varphi}(t)) - \hat{\varphi}(t)}{t} = \hat{f}(t, \hat{\varphi}(t))$$

Dies zeigt die erste Aussage des folgenden Korollares zu Satz 103:

Korollar 66. *Sei $B \subseteq \mathbb{R}$ ein nichtentartetes Intervall und $g\colon B \to \mathbb{R}$ stetig. Sei $\tilde{D} \subseteq \mathbb{R}_\times =: D$ ein nichtentartetes Intervall, $(t_0, x_0) \in M_B$ mit $t_0 \in \tilde{D}$, und $\varphi, \hat{\varphi}\colon \tilde{D} \to \mathbb{R}$ miteinander wie in* (532) *verknüpft. Dann gilt:*

(1) *Die folgenden Aussagen sind äquivalent zueinander:*

- φ *ist eine Lösung von* $\varphi' = f(\cdot, \varphi) = g\big(\frac{\varphi}{\cdot}\big)$ *mit* $\varphi(t_0) = x_0$. *(siehe* (528)*)*
- $\hat{\varphi}$ *ist eine Lösung von* $\hat{\varphi}' = \hat{f}(\cdot, \varphi) = \hat{h} \cdot (\hat{g} \circ \hat{\varphi}) = \frac{(g - \mathrm{id}_B) \circ \hat{\varphi}}{\cdot}$ *mit* $\hat{\varphi}(t_0) = \frac{x_0}{t_0}$.

(2) *Seien* $\hat{H}, \hat{G}, \hat{L}$ *bezüglich der Abbildungen* (531) *wie in* (i) *und* (ii) *definiert mit* $(\hat{t}_0, \hat{x}_0) = (t_0, \frac{x_0}{t_0})$*; und es gelte* $\hat{H}(\tilde{D}) \subseteq \hat{G}(B) = \tilde{B}$ *(durch verkleinern von* $\tilde{D}$ *um* t_0 *lässt sich dies immer erreichen falls* B *offen ist):*

- *Gemäß Satz 103 ist die* C^1*-Abbildung*

$$\hat{\varphi} := \hat{L} \circ \hat{H}|_{\tilde{D}} \colon \tilde{D} \to B$$

die eindeutige Lösung der Differentialgleichung $\hat{\varphi}' = \hat{h} \cdot (\hat{g} \circ \hat{\varphi})$ *zu der Anfangsbedingung* $\hat{\varphi}(t_0) = \frac{x_0}{t_0}$*, sodass* $\mathrm{dom}(\hat{\varphi}) = \tilde{D}$ *gilt.*

- *Gemäß Teil* (1) *ist die* C^1*-Abbildung* $\varphi \colon \tilde{D} \ni t \mapsto t \cdot \hat{\varphi}(t) \in \mathbb{R}$ *die eindeutige Lösung der Differentialgleichung* $\varphi' = g\big(\frac{\varphi}{\cdot}\big)$ *zur Anfangsbedingung* $\varphi(t_0) = x_0$ *mit* $\mathrm{dom}(\varphi) = \tilde{D}$.

Beweis. Klar. □

17.3 Differentialgleichungen in mehreren Veränderlichen

In diesem Abschnitt verallgemeinern wir die in Abschnitt 17.1 eingeführten Begrifflichkeiten auf die mehrerer Variablen.

17.3.1 Zeitabhängige Vektorfelder

Eine zentrale Begrifflichkeit im Rahmen gewöhnlicher Differentialgleichungen erster Ordnung in mehreren Veränderlichen ist das sogenannte zeitabhängige Vektorfeld:

Terminologie 51 (zeitabhängiges Vektorfeld)**.** *Sei* $n \in \mathbb{N}_{>0}$ *und* $O \subseteq \mathbb{R} \times \mathbb{R}^n$ *eine offene Teilmenge. Eine Abbildung* $\mathcal{V} \colon O \to \mathbb{R}^n$ *heißt* zeitabhängiges Vektorfeld auf O*. Man sagt:*

a) $\mathcal{V}$ erfüllt eine lokale Lipschitz-Bedingung in $(t, x) \in O$*, wenn ein* $L \geq 0$ *und eine offenen Teilmenge* $V \subseteq \mathbb{R} \times \mathbb{R}^n$ *mit* $(t, x) \in V \subseteq O$ *existiert, sodass*

$$\|\mathcal{V}(t', x_1) - \mathcal{V}(t', x_2)\| \leq L \cdot \|x_1 - x_2\| \qquad \forall\, (t', x_1), (t', x_2) \in V \tag{533}$$

für eine Norm $\|\cdot\|$ *auf* $\mathbb{R}^n$ *gilt.*

Beachte: *Gilt* (533) *für eine Norm auf* $\mathbb{R}^n$*, so mit entsprechend modifizierter Konstante* L *auch für jede andere Norm auf* $\mathbb{R}^n$ *(Äquivalenz aller Normen auf* $\mathbb{R}^n$*).*

b) $\mathcal{V}$ erfüllt eine lokale Lipschitz-Bedingung*, wenn* $\mathcal{V}$ *eine lokale Lipschitz-Bedingung in jedem* $(t, x) \in O$ *erfüllt.*

$\mathcal{V}$ *heißt* zeitabhängiges C^k-Vektorfeld *für* $k \in \mathbb{N} \cup \{\infty\}$*, wenn* $\mathcal{V} \in C^k(O, \mathbb{R}^n)$ *gilt.*

Die lokale Lipschitz-Bedingung (533) ist von zentraler Wichtigkeit für die Lösbarkeit und Eindeutigkeit der, einem Vektorfeld zugeordneten gewöhnlichen Differentialgleichung, welche wir im nächsten Abschnitt behandeln werden.

Bemerkung 113. *Ist $\mathcal{V}$ ein zeitabhängiges C^1-Vektorfeld, so erfüllt $\mathcal{V}$ eine lokale Lipschitz-Bedingung:*

In der Tat, wegen Übung 157 ist $\mathcal{V}$ ja sogar lokal Lipschitzstetig; d.h. für $(t,x) \in O$ vorgegeben, existiert $V \subseteq \mathbb{R} \times \mathbb{R}^n$ offen mit $(t,x) \in V \subseteq O$, sodass

$$\|\mathcal{V}(t_1,x_1) - \mathcal{V}(t_2,x_2)\|_1 \leq L \cdot \|(t_1,x_1) - (t_2,x_2)\|_2 \qquad \forall\,(t_1,x_1),\ (t_2,x_2) \in V. \tag{534}$$

für fixierte Normen $\|\cdot\|_1$ auf $\mathbb{R}^n$ und $\|\cdot\|_2$ auf $\mathbb{R}^{n+1}$ gilt. Setzt man $\|\cdot\| := \|\cdot\|_1$ und wählt

$$\|\cdot\|_2 \colon \mathbb{R} \times \mathbb{R}^n \to [0,\infty), \qquad (t,x) \mapsto \max(|t|, \|x\|),$$

so ließt sich die rechte Seite von (534) in der Form $L \cdot \max(|t_1 - t_2|, \|x_1 - x_2\|)$; und man erhält (533) im Falle $t_1 = t' = t_2$.

Beispiel 118. *Sei $n = 3$ und $O = \mathbb{R} \times \mathbb{R}^3$. Stellt man sich den $\mathbb{R}^3$ mit einer strömenden Flüssigkeit gefüllt vor, so kann man diesem System ein zeitabhängiges Vektorfeld $\mathcal{V}$ dadurch zuordnen, dass man $\mathcal{V}(t,x)$ als die Geschwindigkeit des Flüssigkeitspartikels definiert, welches sich zum Zeitpunkt t am Ort x befindet.*[161]

Beispiel 119 (Rotierende Einheitskreisscheibe)**.** *Sei $n = 2$, $v = (v_1, v_2) \in \mathbb{R}^2$, sowie*

$$O = \{(t,x) \in \mathbb{R} \times \mathbb{R}^2 \mid \|x - t \cdot v\|_{\text{euk}} < 1\} \subseteq \mathbb{R} \times \mathbb{R}^2.$$

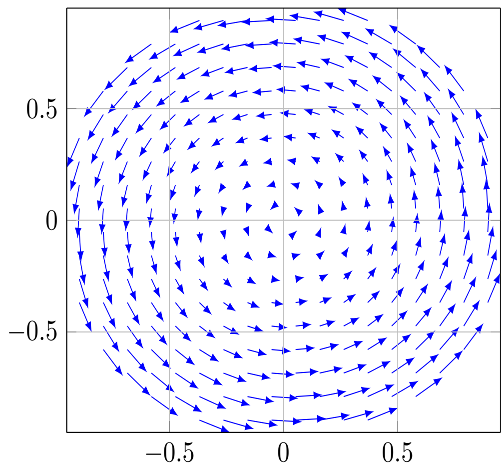
Wir betrachten das glatte zeitabhängige Vektorfeld:[162]

$$\mathcal{V} \colon O \to \mathbb{R}^2, \qquad (t,(x,y)) \mapsto v + (-(y - t \cdot v_2), (x - t \cdot v_1)).$$

*Wie wir in Beispiel 120.*a) *noch genauer sehen werden, beschreibt es die Geschwindigkeitsverteilung der Punkte auf einer offenen und mit der Winkelgeschwindigkeit 1 entgegen dem Uhrzeigersinn (um ihren Mittelpunkt) rotierenden Einheitsscheibe, deren Mittelpunkt sich geradlinig mit konstanter Geschwindigkeit v in der Ebene bewegt und sich zum Zeitpunkt $t = 0$ im Ursprung befindet. Zum Zeitpunkt $t \in \mathbb{R}$ ist dann $\mathcal{V}(t,\cdot)$ auf der offenen Einheitskreischeibe um den Punkt $t \cdot v$ definiert, denn es gilt:*

$$\big(\{t\} \times \mathbb{R}^2\big) \cap O = \big\{x \in \mathbb{R}^2 \;\big|\; \|x - t \cdot v\|_{\text{euk}} < 1\big\}.$$

Für $t = 0$ ergibt sich die folgenden Skizze für $\mathcal{V}(0,\cdot)$:



[161] In der Realität enthält jeder beschränkte Raumbereich natürlich nur endlich viele Flüssigkeitspartikel, sodass man $\mathcal{V}(t,x)$ eher als die mittlere Geschwindigkeit aller Partikel definieren würde, die sich zum Zeitpunkt t bspw. in einem ε-Ball ($\varepsilon > 0$ einmal fixiert) um x tummeln.

[162] Offensichtlich erfüllt $\mathcal{V}$ eine lokale Lipschitz-Bedingung, denn es gilt ja $\mathcal{V}(t,(x,y)) - \mathcal{V}(t,(x',y')) = (y' - y, x - x')$ für $t \in \mathbb{R}$, $(x,y),(x',y') \in \mathbb{R}^2$.

Man verifiziert leicht (Übung), dass die Menge O tatsächlich offen in $\mathbb{R}\times\mathbb{R}^2$ ist; allerdings nicht in der Form $I \times U$, mit $I \subseteq \mathbb{R}$ sowie $O \subseteq \mathbb{R}^2$ offen, geschrieben werden kann. Die allgemeine Formulierung in unseren Definitionen lässt es also durchaus wieder zu, dass zeitabhängige Vektorfelder nicht zu allen Zeiten auf denselben Raumbereichen definiert sein müssen.

Eine ähnliche Argumentation wie in Übung 157 liefert das folgende Kriterium für die Erfülltheit lokaler Lipschitz-Bedingungen:

Proposition 26 (Kriterium für Lokale Lipschitzbedingung)**.** *Sei $n \in \mathbb{N}_{>0}$, $O \subseteq \mathbb{R} \times \mathbb{R}^n$ offen, und $\mathcal{V}\colon O \to \mathbb{R}^n$ ein zeitabhängiges Vektorfeld derart, dass für jedes $1 \leq j \leq n$ die partielle Ableitung (Richtungsableitung)*

$$\tilde{\partial}_j\mathcal{V}\colon O \to \mathbb{R}^n, \qquad (t,x) \mapsto D_{(0,e_j)}\mathcal{V}(t,x) =: \tilde{\partial}_j\mathcal{V}(t,x)$$

existiert und stetig ist. Dann erfüllt $\mathcal{V}$ eine lokale Lipschitz-Bedingung; genauer existieren zu jedem fixierten $(t,x) \in O$ gewisse $L \geq 0$, $\delta > 0$, sowie $U \subseteq \mathbb{R}^n$ offen mit $x \in U$, sodass gilt:

$$\|\mathcal{V}(t',x_1) - \mathcal{V}(t',x_2)\|_\infty \leq L \cdot \|x_1 - x_2\|_\infty \qquad \forall\,(t',x_1),\ (t',x_2) \in (t-\delta,t+\delta) \times U.$$

$*$*Beweis.* Wir betrachten die jeweiligen Maximumsnormen auf $\mathbb{R} \times \mathbb{R}^n$ und $\mathbb{R}^n$. Sei $(t,x) \in O$ fixiert. Wir wählen $\delta > 0$ mit

$$V := (t-\delta,t+\delta) \times \overbrace{\mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_\delta(x)}^{=:\,U} = \mathsf{B}[\mathrm{d}_{\|\cdot\|_\infty}]_\delta(t,x) \subseteq O.$$

- Für jedes fixierte $\tau \in (t-\delta,t+\delta)$ ist

  $$f_\tau\colon U \to \mathbb{R}^n, \qquad x \mapsto \mathcal{V}(\tau,x)$$

  eine C^1-Abbildung wegen Satz 87, denn f_τ ist einmal stetig partiell differenzierbar wegen[163]

  $$\partial_j f_\tau(y) = \lim_{h\to 0}\frac{f_\tau(y+h\cdot e_j) - f_\tau(y)}{h} = \lim_{h\to 0}\frac{\mathcal{V}((\tau,y)+h\cdot(0,e_j)) - \mathcal{V}(\tau,y)}{h} = \underbrace{D_{(0,e_j)}\mathcal{V}(\tau,y)}_{\tilde{\partial}_j\mathcal{V}(\tau,y)}$$

  für alle $1 \leq j \leq n$ und $y \in U$. Für $y \in U$ gilt dann:

  $$\mathrm{d}_y f_\tau(v) \overset{(421)}{=} \textstyle\sum_{j=1}^n v_j \cdot \underbrace{\tilde{\partial}_j\mathcal{V}(\tau,y)}_{\mathrm{d}_y f_\tau(e_j)} \qquad \forall\, v = (v_1,\dots,v_n) \in \mathbb{R}^n \quad \text{also} \tag{535}$$
  $$\|\mathrm{d}_y f_\tau(v)\|_\infty \leq \|v\|_\infty \cdot n \cdot \max\big(\|\tilde{\partial}_j\mathcal{V}(\tau,y)\|_\infty \,\big|\, 1 \leq j \leq n\big).$$

- Für $1 \leq j \leq n$ ist $\tilde{\partial}_j\mathcal{V}$ stetig in (t,x). Durch verkleinern von $\delta > 0$ können wir somit erreichen, dass für ein $L \geq 0$ und alle $(\tau,y) \in \mathsf{B}[\mathrm{d}_{\|\cdot\|}]_\delta(t,x) = V$ gilt:[164]

  $$\|\tilde{\partial}_j\mathcal{V}(\tau,y)\|_\infty \leq \frac{L}{n} \quad \forall\, 1 \leq j \leq n \qquad \overset{(535)}{\Longrightarrow} \qquad \|\mathrm{d}_y f_\tau(v)\|_\infty \leq \overbrace{L \cdot \|v\|_\infty}^{=:\,L_v} \quad \forall\, v \in \mathbb{R}^n. \tag{536}$$

- Für $(t',x_1),(t',x_2) \in V$ vorgegeben, sei $a := 0$, $b := 1$ und $v := x_2 - x_1$. Dann gilt $x_1 + [a,b]\cdot v \subseteq U$ (da U konvex); und Satz 90 (Schrankensatz) zusammen mit (536) liefert:

  $$\|\mathcal{V}(t',x_1) - \mathcal{V}(t',x_2)\|_\infty = \|f_{t'}(x_1 + a\cdot v) - f_{t'}(x_1 + b\cdot v)\|_\infty \leq (b-a)\cdot L_v = L \cdot \|x_1 - x_2\|_\infty. \quad \square$$

[163] Für die Stetigkeit von $\partial_j f_\tau$ benötigen wir an dieser Stelle tatsächlich nicht die Stetigkeit der vollen Abbildung $\tilde{\partial}_j\mathcal{V}$ (die ja per Voraussetzung gegeben ist), sondern nur die Stetigkeit von $U \ni y \mapsto \tilde{\partial}_j\mathcal{V}(\tau,y) \in \mathbb{R}^n$ (welche ja durch Übung 144 bereitgestellt wird). Allerdings werden wir die Stetigkeit von $\tilde{\partial}_j\mathcal{V}$ tatsächlich weiter unten noch explizit verwenden.

[164] Man setze beispielsweise $L := n\cdot\big(1+\max\big(\|\tilde{\partial}_j\mathcal{V}(t,x)\|_\infty \,|\, 1 \leq j \leq n\big)\big)$ und wähle $\delta > 0$ entsprechend klein (Stetigkeit von $\tilde{\partial}_j\mathcal{V}$ in (t,x)).

Bemerkung 114. *Sei $n \in \mathbb{N}_{>0}$ und $U \subseteq \mathbb{R}^n$ offen, und $\mathcal{W}\colon U \to \mathbb{R}^n$ eine Abbildung. Wir können $\mathcal{W}$ vermöge*

$$\mathcal{V}_{\mathcal{W}}\colon O := \mathbb{R} \times U \to \mathbb{R}^n, \qquad (t,x) \mapsto \mathcal{W}(x)$$

immer auch als zeitabhängiges Vektorfeld auffassen.[165] *Offensichtlich gilt dann:*

- *$\mathcal{V}_{\mathcal{W}}$ erfüllt eine lokale Lipschitz-Bedingung in $(t,x) \in O$ gdw. $\mathcal{W}$ lokal Lipschitzstetig in x ist.*
- *$\mathcal{V}_{\mathcal{W}}$ erfüllt eine lokale Lipschitz-Bedingung gdw. $\mathcal{W}$ lokal Lipschitzstetig ist (also lokal Lipschitzstetig in jedem $x \in U$).*
- *Für $k \in \mathbb{N}$ gilt $\mathcal{V}_{\mathcal{W}} \in C^k(O,\mathbb{R}^n)$ gdw. $\mathcal{W} \in C^k(U,\mathbb{R}^n)$ gilt (einfache Induktion).*

17.3.2 Gewöhnliche Differentialgleichung

Definition 81. *Sei $O \subseteq \mathbb{R} \times \mathbb{R}^n$ offen, und $\mathcal{V}\colon O \to \mathbb{R}^n$ ein stetiges zeitabhängiges Vektorfeld auf O. Dann heißt*

$$\gamma' = \mathcal{V}(\cdot,\gamma) \tag{537}$$

(gewöhnliche) Differentialgleichung erster Ordnung (auf O). *Eine Bedingung der Gestalt*

$$\gamma(t_0) = x_0 \qquad \textit{für} \qquad (t_0,x_0) \in O, \tag{538}$$

heißt Anfangsbedingung zu der Differentialgleichung (537)*.*

(1) • *Eine* lokale Lösung von (537) *ist eine differenzierbare Kurve $\gamma\colon I \to \mathbb{R}^n$ auf einem nichtleeren offenen Intervall $I \subseteq \mathbb{R}$, mit*

$$\Gamma_\gamma = \{\,(t,\gamma(t)) \mid t \in I\,\} \subseteq O \qquad \textit{und} \qquad \gamma'(t) = \mathcal{V}(t,\gamma(t)) \quad \forall\, t \in I. \tag{539}$$

*(Beachte, dass die rechte Seite nur formuliert werden kann, wenn die linke Seite gilt. Beachte weiterhin, dass γ wegen Terminologie 36.*b) *stetig ist, da differenzierbar.)*

• *Gilt zusätzlich zu* (539) *noch*

$$(t_0,x_0) \in \Gamma_\gamma \quad \textit{für ein} \quad (t_0,x_0) \in O \qquad \big(t_0 \in I \quad \textit{mit} \quad \gamma(t_0) = x_0 \quad \textit{und} \quad (t_0,x_0) \in O\big),$$

so wird γ als lokale Lösung von (514) zur Anfangsbedingung (538) *bezeichnet.*

(2) • *Die Differentialgleichung* (537) *heißt* lokal eindeutig lösbar in $(\mathsf{t}_0,\mathsf{x}_0) \in \mathsf{O}$, *wenn für je zwei lokale Lösungen $\gamma_1\colon I_1 \to \mathbb{R}^n$ und $\gamma_2\colon I_2 \to \mathbb{R}^n$ von* (537) *zur Anfangsbedingung* (538) *(d.h. $t_0 \in I_1 \cap I_2$ mit $\gamma_1(t_0) = x_0 = \gamma_2(t_0)$) gilt:*

$$\exists\, J \subseteq I_1 \cap I_2 \;\; \textit{offenes Intervall mit} \;\; t_0 \in J \;\; \textit{und} \;\; \gamma_1|_J = \gamma_2|_J. \tag{540}$$

• *Die Differentialgleichung* (537) *heißt* lokal eindeutig lösbar, *wenn für je zwei lokale Lösungen $\gamma_1\colon I_1 \to \mathbb{R}^n$ und $\gamma_2\colon I_2 \to \mathbb{R}^n$ von* (537) *gilt:*[166]

$$\underbrace{\gamma_1(t_0) = \gamma_2(t_0) \quad \textit{für ein} \quad t_0 \in I_1 \cap I_2}_{\textit{gleichbedeutend zu } \Gamma_{\gamma_1} \cap \Gamma_{\gamma_2} \neq \emptyset} \qquad \Longrightarrow \qquad \gamma_1|_{I_1 \cap I_2} = \gamma_2|_{I_1 \cap I_2}. \tag{541}$$

Beachte: *Lemma 92 weiter unten zeigt, dass* (537) *genau dann lokal eindeutig lösbar ist, wenn* (537) *in jedem Punkt lokal eindeutig lösbar ist.*

[165] Gemäß unseren Definitionen ist $\mathcal{V}_{\mathcal{W}}$ ein zeitabhängiges Vektorfeld, da es ja auf der offenen Teilmenge $O \subseteq \mathbb{R} \times \mathbb{R}^n$ definiert ist. Natürlich ist aber $\mathcal{V}_{\mathcal{W}}$ zeitunabhängig in dem Sinne, dass $\mathcal{V}_{\mathcal{W}}(s,x) = \mathcal{V}_{\mathcal{W}}(t,x)$ für alle $s,t \in \mathbb{R}$ und $x \in \mathbb{R}^n$ gilt (vgl. Terminologie 47).

[166] Nochmals anders formuliert bedeutet die linke Seite von (541), dass ein $(t_0,x_0) \in O$ existiert, sodass γ_1,γ_2 beides lokale Lösungen zum selben Anfangswertproblem (538) sind (setze $x_0 := \gamma_1(t_0) = \gamma_2(t_0)$).

Bemerkung 115. a) *Beachte, dass in Teil* (2) *von Definition 81 die Existenz von lokalen Lösungen nicht gefordert wird (nur die entsprechende Eindeutigkeitseigenschaften im Falle der Existenz).*

b) *Sei* (537) *lokal eindeutig lösbar, sowie* $I \subseteq \mathrm{pr}_1(O)$ *ein nichtleeres offenes Intervall und* $(t_0, x_0) \in O$ *mit* $t_0 \in I$. *Aus* (541) *folgt unmittelbar:*

Ist $\gamma\colon I \to \mathbb{R}^n$ *eine lokale Lösung von* $\gamma' = \mathcal{V}(\cdot, \gamma)$ *mit* $\gamma(t_0) = x_0$, *so ist* γ *bereits die eindeutige lokale Lösung von* $\gamma' = \mathcal{V}(\cdot, \gamma)$ *mit* $\gamma(t_0) = x_0$, *sodass* $\mathrm{dom}(\gamma) = I$ *gilt.*

c) *Setzen wir* $n \equiv 1$, $\gamma \equiv \varphi$ *und* $\mathcal{V} \equiv f$ *in Definition 81.*(1), *so erhalten wir Definition 80 aus Abschnitt 17.1 zurück; und zwar für den Spezialfall, dass dort* M *offen ist, wobei auch nur (lokale) Lösungen auf offenen Intervallen betrachtet werden. (Wir beschränken uns im höherdimensionalen Fall auf offene Definitionsbereiche, weil hierdurch die Gesamtdiskussion einheitlicher und untechnischer/unkomplizierter wird.)*

Lemma 92. *Sei* $O \subseteq \mathbb{R} \times \mathbb{R}^n$ *offen, und* $\mathcal{V}\colon O \to \mathbb{R}^n$ *ein stetiges zeitabhängiges Vektorfeld auf* O. *Die Differentialgleichung* $\gamma' = \mathcal{V}(\cdot, \gamma)$ *ist genau dann lokal eindeutig lösbar, wenn sie in jedem Punkt lokal eindeutig lösbar ist.*

Beweis. Die eine Implikation ist klar, denn (541) impliziert offensichtlich (540). Für die andere Richtung argumentieren wir ähnlich wie im Beweis von Satz 74:

Seien $\gamma_1\colon I_1 \to \mathbb{R}^n$ und $\gamma_2\colon I_2 \to \mathbb{R}^n$ lokale Lösungen von $\gamma' = \mathcal{V}(\cdot, \gamma)$ mit $(t_0, x_0) \in \Gamma_{\gamma_1} \cap \Gamma_{\gamma_2}$, d.h. $\gamma_1(t_0) = x_0 = \gamma_2(t_0)$ mit $t_0 \in I_1 \cap I_2$. Es bezeichne $\mathcal{J}$ die Menge aller (nichtleeren) offenen Intervalle $J \subseteq I_1 \cap I_2$ mit $t_0 \in J$ und $\gamma_1|_J = \gamma_2|_J$.

- Es gilt $\mathcal{J} \neq \emptyset$, da (537) lokal eindeutig lösbar in (t_0, x_0) ist. (beachte $\gamma_1(t_0) = x_0 = \gamma_2(t_0)$)
- Es ist $I := \bigcup_{J \in \mathcal{J}} J \subseteq I_1 \cap I_2$ als Vereinigung offener Intervalle offen wegen Proposition 3.OM3), und ein Intervall wegen Übung 22 (beachte $t_0 \in J$ für alle $J \in \mathcal{J}$). Zudem gilt offensichtlich $t_0 \in I$.
- Es gilt $\gamma_1|_I = \gamma_2|_I$, denn jedes $t \in I$ ist in einem $J \in \mathcal{J}$ enthalten.

Wir wollen nun zeigen, dass $I \overset{!}{=} I_1 \cap I_2 =: \mathcal{T}$ gilt ($\mathcal{T}$ ist ein Intervall wegen Übung 22, sowie offen endlicher als Schnitt offener Mengen), und schreiben

$$\mathcal{T} = (\tau_-, \tau_+) \qquad \text{sowie} \qquad I = (\iota_-, \iota_+) \qquad \text{mit} \qquad \overline{\mathbb{R}} \ni \tau_- \leq \iota_- < \iota_+ \leq \tau_+ \in \overline{\mathbb{R}}.$$

Angenommen es gilt $\iota_+ < \tau_+$ (den Fall $\tau_- < \iota_-$ behandelt man analog), also $\iota_+ \neq \infty$. Wir wählen eine Folge $(t_n)_{n \in \mathbb{N}}$ in I mit $\lim_n t_n = \iota_+ =: t_0'$. Dann gilt $\gamma_1(t_n) = \gamma_2(t_n)$ für alle $n \in \mathbb{N}$, also wegen der Stetigkeit (Folgenstetigkeit) von γ_1, γ_2 (γ_1, γ_2 sind differenzierbar, also stetig):

$$x_0' := \gamma_1(t_0') = \lim_n \gamma_1(t_n) = \lim_n \gamma_2(t_n) = \gamma_2(t_0') \qquad \Longrightarrow \qquad (t_0', x_0') \in \Gamma_{\gamma_1} \cap \Gamma_{\gamma_2}.$$

Wegen der lokal eindeutigen Lösbarkeit von $\gamma' = \mathcal{V}(\cdot, \gamma)$ in (t_0', x_0'), existiert ein offenes Intervall $J \subseteq I_1 \cap I_2$ mit $I \not\ni t_0' \in J$ und $\gamma_1|_J = \gamma_2|_J$ (vgl. (540)). Wegen der Offenheit von J und $\iota_+ = t_0' \in J$ gilt $I \cap J \neq \emptyset$; also ist $\tilde{I} := I \cup J \subseteq I_1 \cap I_2$ ein offenes Intervall (Übung 22) mit $\gamma_1|_{\tilde{I}} = \gamma_2|_{\tilde{I}}$. Wegen $t_0 \in I \subseteq \tilde{I}$ gilt somit $\tilde{I} \in \mathcal{J}$, also $\tilde{I} \subseteq I$, was $I \not\ni t_0' \in J \subseteq \tilde{I}$ widerspricht. □

17.3.3 Lokale Eindeutigkeit

Der nächste Satz (Satz 105) garantiert die lokale Eindeutigkeit im Lipschitzfall. Die lokale Existenz werden wir gleich im Anschluss beweisen. Wir bemerken zunächst Folgendes (vgl. Bemerkung 111.(c)):

Lemma 93. *Sei* $O \subseteq \mathbb{R} \times \mathbb{R}^n$ *offen,* $(t_0, x_0) \in O$, $I \subseteq \mathbb{R}$ *ein offenes Intervall,* $\gamma\colon I \to \mathbb{R}^n$ *eine Kurve mit* $\Gamma_\gamma \subseteq O$, *sowie* $\mathcal{V}\colon O \to \mathbb{R}^n$ *ein stetiges zeitabhängiges Vektorfeld auf* O. *Dann sind äquivalent:*

a) γ *ist differenzierbar mit* $\gamma' = \mathcal{V}(\cdot,\gamma)$ *und* $(t_0,x_0) \in \Gamma_\gamma$.

(γ ist eine lokale Lösung der Differentialgleichung $\gamma' = \mathcal{V}(\cdot,\gamma)$ zur Anfangsbedingung $\gamma(t_0) = x_0$.)

b) γ *ist stetig mit*

$$\gamma(t) = x_0 + \int_{t_0}^{t} \mathcal{V}(s,\gamma(s))\,\mathrm{d}s \qquad \forall\, t \in I.$$

Beachte: *Der Integrand* $\mathcal{V} \circ (\mathrm{id}_I \times \gamma)\colon I \ni s \mapsto \mathcal{V}(s,\gamma(s)) \in \mathbb{R}^n$ *ist stetig wegen Lemma 74; denn* id_I, γ *sind ja beide stetig.*

Beweis. Dies folgt sofort aus dem Hauptsatz für stetige Kurven Korollar 55. □

Satz 105 (Lokale Eindeutigkeit von Lösungen)**.** *Sei* $n \in \mathbb{N}_{>0}$, $O \subseteq \mathbb{R} \times \mathbb{R}^n$ *offen, und* $\mathcal{V}\colon O \to \mathbb{R}^n$ *ein stetiges zeitabhängiges Vektorfeld auf* O. *Dann gelten die folgenden Aussagen:*

1) *Erfüllt* $\mathcal{V}$ *eine lokale Lipschitz-Bedingung in* $(t_0,x_0) \in O$, *so ist die Differentialgleichung* $\gamma' = \mathcal{V}(\cdot,\gamma)$ *lokal eindeutig lösbar in* (t_0,x_0).
2) *Erfüllt* $\mathcal{V}$ *eine lokale Lipschitz-Bedingung, so ist die Differentialgleichung* $\gamma' = \mathcal{V}(\cdot,\gamma)$ *lokal eindeutig lösbar.*

Beweis. Teil 2) folgt sofort aus Teil 1) und Lemma 92. Für Teil 1), sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$; sowie $L > 0$ und $V \subseteq \mathbb{R} \times \mathbb{R}^n$ offen mit $(t_0,x_0) \in V \subseteq O$, sodass gilt: (vgl. (533))

$$\|\mathcal{V}(t',x_1) - \mathcal{V}(t',x_2)\| \leq L \cdot \|x_1 - x_2\| \qquad \forall\, (t',x_1),(t',x_2) \in V. \tag{542}$$

Seien $\gamma_1\colon I_1 \to \mathbb{R}^n$, $\gamma_2\colon I_2 \to \mathbb{R}^n$ zwei lokale Lösungen von $\gamma' = \mathcal{V}(\cdot,\gamma)$ mit $\gamma_1(t_0) = x_0 = \gamma_2(t_0)$:

- Indem wir I_1 und I_2 um t_0 verkleinern, können wir wegen $\gamma_1(t_0) = x_0 = \gamma_2(t_0)$ und der Stetigkeit von γ_1, γ_2 (da differenzierbar) annehmen, dass $\Gamma_{\gamma_1}, \Gamma_{\gamma_2} \subseteq V$ gilt.[167]
- Wegen Lemma 93 gilt:

$$\alpha\colon I := I_1 \cap I_2 \to \mathbb{R}^n, \qquad t \mapsto (\gamma_1 - \gamma_2)(t) = \int_{t_0}^{t} \big(\mathcal{V}(s,\gamma_1(s)) - \mathcal{V}(s,\gamma_2(s))\big)\,\mathrm{d}s,$$

wobei $t_0 \in I \neq \emptyset$ wegen Übung 21 ein offenes Intervall ist, und α stetig (sogar C^1).
- Wegen $\alpha(t_0) = 0$, existiert $0 < \delta \leq \frac{1}{2L}$ mit $J := (t_0 - \delta, t_0 + \delta) \subseteq I$ und $\|\alpha|_J\| \leq 1$. Wir setzen

$$C := \sup\nolimits_{\mathbb{R}} \big(\underbrace{\{\, s \in J \mid \|\alpha|_J(s)\| \,\}}_{:= M}\big) \leq 1,$$

und erhalten (beachte $\Gamma_{\gamma_1}, \Gamma_{\gamma_2} \subseteq V$ im zweiten Schritt)

$$\begin{aligned} \|\alpha|_J(t)\| &\overset{(399)}{\leq} \Big| \int_{t_0}^{t} \|\mathcal{V}(s,\gamma_1(s)) - \mathcal{V}(s,\gamma_2(s))\|\,\mathrm{d}s \Big| \overset{(542)}{\leq} \Big| \int_{t_0}^{t} L \cdot \overbrace{\|\gamma_1(s) - \gamma_2(s)\|}^{\|\alpha|_J(s)\| \leq C}\,\mathrm{d}s \Big| \\ &\leq \Big| \int_{t_0}^{t} L \cdot C\,\mathrm{d}s \Big| = |t - t_0| \cdot L \cdot C < \delta \cdot L \cdot C \leq C/2 \end{aligned}$$

für alle $t \in J$. Dies zeigt $M < C/2$, also $C = \sup_{\mathbb{R}}(M) \leq C/2$, also $C = 0$. Somit gilt $\alpha|_J = 0$, also $\gamma_1|_J = \gamma_2|_J$. □

Beispiel 120.

[167]Für $p \in \{1,2\}$ ist $\mu_p := \mathrm{id}_{I_p} \times \gamma_p\colon I_p \ni t \mapsto (t,\gamma_p(t)) \in O$ stetig mit $\mu_p(t_0) = (t_0,x_0) \in V$.

a) *Das Vektorfeld aus Beispiel 119 erfüllt eine lokale Lipschitz-Bedingung. Für* $(t_0, x_0) \in O$ *vorgegeben, ist die zugehörige (in diesem Fall globale Lösung) von* $\gamma' = \mathcal{V}(\cdot, \gamma)$ *mit* $\gamma(t_0) = x_0$ *gegeben durch (wir identifizieren* $\mathbb{R}^2 \cong \mathbb{C}$*):*

$$\gamma\colon \mathbb{R} \to \mathbb{R}^2, \qquad t \mapsto t \cdot v + r \cdot \underbrace{\mathrm{e}^{\mathrm{i} \cdot \overbrace{((t - t_0) \ + \ \phi_0)}^{=:\, \phi(t)}}}_{\cong\, (\cos(\phi(t)), \sin(\phi(t)))},$$

sofern wir $r := \|z\|_{\mathrm{euk}}$ *für* $z := x_0 - t_0 \cdot v$ *setzen und* $\phi_0 \in \mathbb{R}$ *derart wählen, dass* $z = r \cdot \mathrm{e}^{\mathrm{i} \cdot \phi_0}$ *gilt. Dies ist die Bahnkurve eines fixierten Punktes auf einer offenen Einheitsscheibe, welche sich mit der Winkelgeschwindigkeit* 1 *entgegen dem Uhrzeigersinn (um ihren Mittelpunkt) rotiert, und deren Mittelpunkt sich geradlinig mit konstanter Geschwindigkeit* v *in der Ebene bewegt und sich zum Zeitpunkt* $t = 0$ *im Ursprung befindet.*

b) *Wir betrachte die Situation aus Beispiel 113, d.h.,* $n = 1$, $O = M = \mathbb{R} \times \mathbb{R}$ *und*

$$\mathcal{V} = f\colon O = \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \qquad (t, x) \mapsto 3 \cdot \sqrt[3]{x^2}.$$

Wir setzen $A := \mathbb{R} \times \{0\}$ *(abgeschlossen) sowie* $\tilde{O} := (\mathbb{R} \times \mathbb{R}) \setminus A$ *(offen):*

- $\tilde{\mathcal{V}} := \mathcal{V}|_{\tilde{O}}$ *erfüllt eine lokale Lipschitz-Bedingung da glatt (vgl. Proposition 26).*
- $\mathcal{V}$ *erfüllt in keinem* $(t_0, x_0) \in A$ *eine lokale Lipschitz-Bedingung; denn für jedes* $t' \in \mathbb{R}$ *und* $0 \neq x \in \mathbb{R}$ *gilt ja:*[168]

$$\frac{|\mathcal{V}(t', x) - \mathcal{V}(t', 0)|}{|x - 0|} = \frac{|\mathcal{V}(t', x)|}{|x|} = 3 \cdot \frac{\sqrt[3]{|x|^2}}{|x|} = \frac{3}{\sqrt[3]{|x|}} \overset{x \to 0}{\longrightarrow} \infty.$$

In Beispiel 113 hatten wir ja auch bereits die vier verschiedenen Lösungen (521) *(setze dort* $\varphi = \gamma$*) mit* $\gamma(t_0) = 0$ *kennengelernt.*

Seien nun $t_0, x_0 > 0$ *vorgegeben, also* $(t_0, x_0) \in \tilde{O}$*. Dann ist* *(beachte* $\Gamma_{\tilde{\gamma}} \subseteq \tilde{O}$*)*

$$\tilde{\gamma}\colon \tilde{I} := \big(\underbrace{-\sqrt[3]{x_0} + t_0}_{=:\, \tau}, \infty\big) \to (0, \infty) \subseteq \mathbb{R}, \qquad t \mapsto x_0 + (t - t_0)^3$$

*die gemäß Bemerkung 115.*b) *eindeutige Lösung von* $\tilde{\gamma}' = \tilde{\mathcal{V}}(\cdot, \tilde{\gamma})$ *auf* $\tilde{I}$ *mit* $\tilde{\gamma}(t_0) = x_0$*. Will man allerdings* $\tilde{\gamma}$ *(„nach links“) auf* $\mathbb{R}$ *ausdehnen, so hat man wieder die zwei Möglichkeiten:*

$$\gamma_1\colon \mathbb{R} \ni t \mapsto x_0 + (t - t_0)^3 \in I_1 := \mathbb{R} \qquad \textit{sowie} \qquad \gamma_2\colon I_2 := \mathbb{R} \ni t \mapsto \begin{cases} x + (t - t_0)^3 & \textit{für } t > \tau \\ 0 & \textit{für } t \leq \tau. \end{cases}$$

Beachte: *Es gilt* $(\tau, 0) \in A \cap \Gamma_{\gamma_1} \cap \Gamma_{\gamma_2}$ *und* $\mathcal{V}$ *erfüllt in* $(\tau, 0)$ *keine lokale Lipschitz-Bedingung.*

Insbesondere sieht man an diesem Beispiel nun auch explizit, dass die lokale Lipschitz-Bedingung in einem Punkt tatsächlich auch nur die lokale Lösbarkeit in diesem Punkt erzwingt, d.h., es gilt hier zwar $\gamma_1|_J = \gamma_2|_J$ *für* $J := (\tau, \infty) \subset \mathbb{R} = I_1 \cap I_2$*, aber eben nicht* $\gamma_1|_{I_1 \cap I_2} = \gamma_2|_{I_1 \cap I_2}$*.*

17.3.4 Lokale Existenz: Der Satz von Picard Lindelöf

Bemerkung 116 (Erinnerung: Funktionenräume)**.** *Sei* (Z, d) *ein metrischer Raum und* $(V, \|\cdot\|)$ *ein Banachraum (über* $\mathbb{R}$ *oder* $\mathbb{C}$*). Wir hatten in Satz 38 in Abschnitt 7.3.1 gesehen, dass die Menge*

$$C_b(Z, V) = \{f \in C^0(Z, V) \mid \exists\, C \geq 0\colon \|f(z)\| \leq C \textit{ für alle } z \in Z\}$$

[168] Womit für **kein** $L, \varepsilon > 0$ gelten kann: $|\mathcal{V}(t', x_1) - \mathcal{V}(t', x_2)| \leq L \cdot |x_1 - x_2|$ für alle $(t', x_1), (t', x_2) \in (t_0 - \delta, t_0 + \delta) \times (x_0 - \delta, x_0 + \delta)$.

der stetigen beschränkten Funktionen $Z \to V$ versehen mit der Supremumsnorm (vgl. (223))

$$\|\cdot\|_\infty\colon C_b(Z,V) \to [0,\infty), \qquad f \mapsto \sup_{\mathbb{R}}(\{\|f(z)\| \mid z \in Z\})$$

ebenfalls ein Banachraum ist, also per definitionem $(C_b(Z,V), \mathrm{d}_{\|\cdot\|_\infty})$ ein vollständiger metrischer Raum. Für $\delta > 0$ und $f \in C_b(Z,V)$ vorgegeben, ist nun[169]

$$\begin{aligned} B := \overline{\mathsf{B}}[\mathrm{d}_{\|\cdot\|_\infty}]_\delta(f) &= \big\{g \in C_b(Z,V) \mid \|f-g\|_\infty = \mathrm{d}_{\|\cdot\|_\infty}(f,g) \leq \delta\big\} \\ &= \big\{g \in C^0(Z,V) \mid \|f(z) - g(z)\| \leq \delta \ \ \textit{für alle} \ \ z \in Z\big\} \end{aligned} \tag{543}$$

abgeschlossen wegen Lemma 28; also zeigt Lemma 72, dass der metrische (Unter)raum $(B, (\mathrm{d}_{\|\cdot\|_\infty})_B)$ ebenfalls vollständig ist. Wir werden diesen Sachverhalt nun im Beweis von Satz 106 für den Fall anwenden, dass $Z \equiv I \subseteq \mathbb{R}$ ein nichtleeres offenes Intervall (mit der eingeschränkten Betragsmetrik), $V \equiv \mathbb{R}^n$ mit $n \in \mathbb{N}_{>0}$ (versehen mit einer fixierten Norm), und $f\colon I \ni t \mapsto x_0 \in \mathbb{R}^n$ eine konstante Abbildung ist.

Satz 106 (Picard Lindelöf)**.** *Sei $n \in \mathbb{N}_{>0}$, $O \subseteq \mathbb{R} \times \mathbb{R}^n$ offen, $(t_0, x_0) \in O$, und $\mathcal{V}\colon O \to \mathbb{R}^n$ ein stetiges zeitabhängiges Vektorfeld auf O. Erfüllt $\mathcal{V}$ in (t_0,x_0) eine lokale Lipschitz-Bedingung, so existiert eine lokale Lösung der Differentialgleichung $\gamma' = \mathcal{V}(\cdot,\gamma)$ zur Anfangsbedingung $\gamma(t_0) = x_0$.*[170]

Beweis. Wir fixieren auf $\mathbb{R}^n, \mathbb{R} \times \mathbb{R}^n$ die jeweiligen Maximumsnormen, und bezeichnen beide ausnahmsweise wieder mit $\|\cdot\|_{\max}$, um diese im Folgenden von der Supremumsnorm $\|\cdot\|_\infty$ (bezüglich der Maximumsnorm auf $\mathbb{R}^n$) zu unterscheiden.[171] Wegen Lemma 93 müssen wir nur zeigen, dass eine stetige Kurve $\gamma\colon I \to \mathbb{R}^n$ ($\emptyset \neq I \subseteq \mathbb{R}$ ein offenes Intervall) existiert, mit $\Gamma_\gamma \subseteq O$ und

$$\gamma(t) = x_0 + \textstyle\int_{t_0}^t \mathcal{V}(s, \gamma(s))\,\mathrm{d}s \qquad \forall\, t \in I. \tag{544}$$

Sei hierfür $L > 0$ sowie $V \subseteq O$ offen mit $(t_0,x_0) \in V$, sodass gilt: (vgl. Lipschitz-Bedingung (533)):

$$\|\mathcal{V}(t', x_1) - \mathcal{V}(t', x_2)\|_{\max} \leq L \cdot \|x - y\|_{\max} \qquad \forall\, (t', x_1), (t', x_2) \in V. \tag{545}$$

Sei $\delta > 0$ mit $\mathsf{B}_{2\delta}(t_0,x_0) \subseteq V$, also $K := [t_0 - \delta, t_0 + \delta] \times \overline{\mathsf{B}}_\delta(x_0) = \overline{\mathsf{B}}_\delta(t_0, x_0) \subseteq V$. Sei zudem $0 < \lambda < 1$ fixiert. Wegen der Stetigkeit von $\mathcal{V}$ (in (t_0,x_0)) können wir $\delta > 0$ derart verkleinern, dass für ein $M > 0$ gilt:[172]

$$\|\mathcal{V}(z)\|_{\max} \leq M \qquad \forall\, z \in K. \tag{546}$$

Sei $\varepsilon := \min\big(\delta, \frac{\delta}{M}, \frac{\lambda}{L}\big)$. Wir setzen: (vgl. (543))

$$\begin{array}{lll} I := (t_0 - \varepsilon, t_0 + \varepsilon) & \text{sowie} & B := \big\{\gamma \in C^0(I, \mathbb{R}^n) \mid \|\gamma(t) - x_0\|_\infty \leq \delta \ \ \text{für alle} \ \ t \in I\big\} \\ & \text{also} & \\ \Gamma_\gamma \subseteq K \quad \forall\, \gamma \in B & \text{wegen} & I \overset{\varepsilon \leq \delta}{\subseteq} [t_0 - \delta, t_0 + \delta] \quad \text{und} \quad \|\gamma(t) - x_0\|_{\max} \leq \delta \quad \forall\, t \in I. \end{array} \tag{547}$$

Gemäß Bemerkung 116 ist $(B, (\mathrm{d}_{\|\cdot\|_\infty})_B)$ vollständig; und zudem ist

$$\Theta\colon B \to B, \qquad \gamma \mapsto x_0 + \textstyle\int_{t_0}^{\bullet} \mathcal{V}(s, \gamma(s))\,\mathrm{d}s \tag{548}$$

eine kontraktive Selbstabbildung mit Kontraktionskonstante λ:

[169] Für die zweite Identität beachte $\|g(z)\| \leq \|f(z)\| + \|f(z) - g(z)\|$ für alle $z \in Z$.

[170] Es existiert also eine differenzierbare Kurve $\gamma\colon I \to \mathbb{R}^n$ mit einem offenen Intervall $I \subseteq \mathbb{R}$, sodass $(t_0, x_0) \in \Gamma_\gamma \subseteq O$ und $\gamma' = \mathcal{V}(\cdot, \gamma')$ gilt.

[171] Beachte: Für $\delta > 0$ und $(t,x) \in \mathbb{R} \times \mathbb{R}^n$ gilt $\overline{\mathsf{B}}_\delta(t,x) = [t - \delta, t + \delta] \times \overline{\mathsf{B}}_\delta(x)$ für den abgeschlossenen Maximumsnorm δ-Ball um (t,x) in $\mathbb{R} \times \mathbb{R}^n$ auf der linken Seite und dem abgeschlossenen Maximumsnorm δ-Ball um x in $\mathbb{R}^n$ auf der rechten Seite.

[172] Man setzte $M := 1 + \|\mathcal{V}(t_0, x_0)\|_{\max}$, und verkleinere $\delta > 0$ entsprechend. Alternativ folgt die Existenz eines derartigen $M \geq 0$ wegen Satz 83 (allgemeiner Satz vom Maximum) auch direkt aus der Kompaktheit von K; denn $K \subseteq \mathbb{R}^{n+1}$ ist abgeschlossen und beschränkt, also kompakt wegen dem Satz von Heine-Borel (Satz 84).

- $\Theta(\mathsf{B}) \subseteq \mathsf{B}$: Sei $\gamma \in B$. Dann ist $\mathcal{V}(\cdot,\gamma)|_I$ definiert wegen $\Gamma_\gamma \subseteq K \subseteq V \subseteq O$, und offensichtlich stetig (vgl. Lemma 93.b)). Der Hauptsatz zeigt $\Theta(\gamma) \in C^1(I,\mathbb{R}^n) \subseteq C^0(I,\mathbb{R}^n)$; und zudem folgt für jedes $t \in I$: (Definition von ε im letzten Schritt)

$$\|\Theta(\gamma)(t) - x_0\|_{\max} = \Big\| \textstyle\int_{t_0}^t \mathcal{V}(s,\gamma(s))\,\mathrm{d}s\Big\| \overset{(399)}{\leq} \Big| \textstyle\int_{t_0}^t \|\mathcal{V}(s,\gamma(s))\|_{\max}\,\mathrm{d}s\Big| \overset{(546)}{\leq} |t-t_0|\cdot M < \varepsilon \cdot M \leq \delta.$$

- Kontraktionseigenschaft: Für $\gamma_1,\gamma_2 \in B$ und $t \in I$ gilt (beachte $\Gamma_{\gamma_1},\Gamma_{\gamma_2} \subseteq K \subseteq V$ im dritten Schritt):

$$\begin{aligned}
\|\Theta(\gamma_1)(t) &- \Theta(\gamma_2)(t)\|_{\max} \\
&= \Big\| \textstyle\int_{t_0}^t \big(\mathcal{V}(s,\gamma_1(s)) - \mathcal{V}(s,\gamma_2(s))\big)\,\mathrm{d}s\,\Big\|_{\max} \overset{(399)}{\leq} \Big| \textstyle\int_{t_0}^t \|\mathcal{V}(s,\gamma_1(s)) - \mathcal{V}(s,\gamma_2(s))\|_{\max}\,\mathrm{d}s\,\Big| \\
&\overset{(545)}{\leq} \Big| \textstyle\int_{t_0}^t L\cdot\|\gamma_1(s)-\gamma_2(s)\|_{\max}\,\mathrm{d}s\,\Big| \leq \Big| \textstyle\int_{t_0}^t L\cdot\|\gamma_1-\gamma_2\|_\infty\,\Big| = |t-t_0|\cdot L\cdot\|\gamma_1-\gamma_2\|_\infty \\
&< \varepsilon\cdot L\cdot\|\gamma_1-\gamma_2\|_\infty \leq \lambda\cdot\|\gamma_1-\gamma_2\|_\infty,
\end{aligned}$$

  also $(\mathrm{d}_{\|\cdot\|_\infty})_B(\Theta(\gamma_1),\Theta(\gamma_2)) = \|\Theta(\gamma_1)-\Theta(\gamma_2)\|_\infty \leq \lambda\cdot\|\gamma_1-\gamma_2\|_\infty = \lambda\cdot(\mathrm{d}_{\|\cdot\|_\infty})_B(\gamma_1,\gamma_2)$.

Wegen Satz 96 (Banachscher Fixpunktsatz) existiert (genau) ein $\gamma \in B$ mit

$$\gamma = \Theta(\gamma) \overset{(548)}{=} x_0 + \textstyle\int_{t_0}^{\bullet} \mathcal{V}(s,\gamma(s))\,\mathrm{d}s,$$

was wegen $\Gamma_\gamma \subseteq K \subseteq V \subseteq O$ und $\gamma(t_0) = x_0$ die gesuchte lokale Lösung ist (vgl. (544)). □

A Appendix

Zusatzmaterial VL1, Analysis II WS21/22: Dr. Maximilian Hanusch: Polynome vs. Polynomfunktionen:[173]

Korollar 30 besagt allgemein für $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$: Ist $f := \mathsf{Q}|_{\mathrm{K}[\mathsf{Q}]}$ mit $\mathsf{Q}(x) = \sum_{n=0}^{\infty} c_n \cdot (x-p)^n$ eine Potenzreihe mit $\mathsf{R}[\mathsf{Q}] > 0$, so sind äquivalent:

1) Es gilt $c_n = 0$ für alle $n \in \mathbb{N}$.
2) Es gilt $f = 0$.
3) Es existiert eine Folge $(x_n)_{n\in\mathsf{N}}$ in $\mathrm{K}[\mathsf{Q}] \setminus \{p\}$ mit $\lim_n x_n = p$, sowie $f(x_n) = 0$ für alle $n \in \mathsf{N}$.

Dies gilt natürlich auch für Polynomfunktionen; d.h, es existiert $n \in \mathbb{N}$ mit $c_k = 0$ für alle $k \geq n+1$.[174] Der formale Raum der Polynome ist (generell für jeden Körper $\mathbb{K}$) definiert durch

$$\mathcal{P}_{\mathbb{K}} := \{(c_k)_{k\in\mathbb{N}} \text{ ist Folge in } \mathbb{K} \,|\, \exists\, n \in \mathbb{N}\colon c_k = 0 \text{ für alle } k \geq n+1\}.$$

Für $p \in \mathbb{K}$ fixiert, betrachten wir die Abbildung

$$\Omega_{\mathbb{K},p}\colon \mathcal{P}_{\mathbb{K}} \to \mathsf{Abb}(\mathbb{K},\mathbb{K}), \qquad (c_k)_{k\in\mathbb{N}} \mapsto \big[\overbrace{x \mapsto \textstyle\sum_{k=0}^{n} c_k \cdot (x-p)^k}^{\text{Polynomfunktion}}\big].$$

Im Falle $\mathbb{K} = \{\mathbb{R}, \mathbb{C}\}$ ist dann $\Omega_{\mathbb{K},p}$ wegen Korollar 30 (zweite Implikation) injektiv:

$$\begin{aligned} \Omega_{\mathbb{K},p}((a_k)_{k\in\mathbb{N}}) = \Omega_{\mathbb{K},p}((b_k)_{k\in\mathbb{N}}) \qquad &\Longrightarrow \qquad \Omega_{\mathbb{K},p}((a_k - b_k)_{k\in\mathbb{N}}) = 0 \\ &\Longrightarrow \qquad a_k - b_k = 0 \text{ für alle } k \in \mathbb{N}. \end{aligned} \tag{549}$$

Besagte Injektivität ist gleichbedeutend mit der Eindeutigkeit der Darstellung einer Polynomfunktion, die Sie für den Fall $p = 0$ und $\mathbb{K} = \mathbb{R}$ bereits in Übung 95 nachgewiesen hatten.

Der Hintergrund zu den obigen Betrachtungen ist der, dass in der Algebra mit Polynom üblicherweise ein Element von $\mathcal{P}_{\mathbb{K}}$ gemeint ist, welches man dann (meistens wird $p = 0$ betrachtet) in der Form $\sum_{k=0}^{n} a_k \cdot x^k$ **notiert/darstellt/hinschreibt**, aber prinzipiell erst einmal nicht als Abbildungen $\mathbb{K} \to \mathbb{K}$ interpretiert.

Anmerkung: Es wird $\mathcal{P}_{\mathbb{K}}$ zu einem $\mathbb{K}$-Vektorraum, wenn man die Addition und $\mathbb{K}$-Multiplikation auf $\mathcal{P}_{\mathbb{K}}$ gliedweise definiert. Hiermit ist dann $\Omega_{\mathbb{K},p}$ eine $\mathbb{K}$-lineare Abbildung, was nun gerade in der ersten Implikation in (549) benutzt wurde. Zudem definiert man das Produkt zweier Polynome derart, dass es dem Produkt von Polynomfunktionen entspricht – also formell durch die Cauchy-Produktformel. Hiermit wird dann $\Omega_{\mathbb{K},p}$ sogar zu einem Homomorphismus von $\mathbb{K}$-Algebren (eine $\mathbb{K}$-Algebra ist ein $\mathbb{K}$-Vektorraum V zusammen mit einer $\mathbb{K}$-bilinearen Verknüpfung $*\colon V \times V \to V$). ‡

Im Falle $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ geht dann sozusagen wegen der Injektivität von $\Omega_{\mathbb{K},p}$ keine Information verloren, wenn man ein Polynom einfach als Abbildung (also als Polynomfunktion) auffasst. In der Situation endlicher Körper sieht die Sache allerdings ganz anders aus – Hier ist es wichtig, sich zunächst immer klar zu machen, ob von Polynomen oder Polynomfunktionen die Rede ist:

Sei $\mathbb{K} = \{0_{\mathbb{K}}, 1_{\mathbb{K}}\}$ der zweielementige Körper, d.h. $0_{\mathbb{K}} + 0_{\mathbb{K}} = 0_{\mathbb{K}}$, $0_{\mathbb{K}} + 1_{\mathbb{K}} = 1_{\mathbb{K}}$, $0_{\mathbb{K}} \cdot 1_{\mathbb{K}} = 0_{\mathbb{K}} = 0_{\mathbb{K}} \cdot 0_{\mathbb{K}}$, $1_{\mathbb{K}} \cdot 1_{\mathbb{K}} = 1_{\mathbb{K}}$, und zudem $1_{\mathbb{K}} + 1_{\mathbb{K}} = 0_{\mathbb{K}}$. Dann ist

$$P(x) = 1_{\mathbb{K}} \cdot x + 1_{\mathbb{K}} \cdot x^2$$

[173] Dies ist ein Zusatzmaterial: Der hier erläuterte Sachverhalt wurde in der ersten Zentralübung am 15.10.2021 erklärt.
[174] Wir bezeichnen im Folgenden auch Funktionen der Form $P(x) = \sum_{k=0}^{n} c_k \cdot (x-p)^k$ mit $p \in \mathbb{K}$ als Polynomfunktion.

nicht das Nullpolynom (nicht alle Koeffizienten a_k sind $0_{\mathbb{K}}$), aber $\Omega_{\mathbb{K},0}(P)\colon \mathbb{K} \to \mathbb{K}$ ist dennoch die Nullabbildung wegen (beachte, dass $\mathbb{K}$ nur die zwei Elemente $0_{\mathbb{K}}, 1_{\mathbb{K}}$ hat, die man einsetzen kann)

$$\begin{aligned}
\Omega_{\mathbb{K},0}(P)(0_{\mathbb{K}}) &= 1_{\mathbb{K}} \cdot 0_{\mathbb{K}} + 1_{\mathbb{K}} \cdot (0_{\mathbb{K}})^2 = 0_{\mathbb{K}} + 0_{\mathbb{K}} = 0_{\mathbb{K}} \\
\Omega_{\mathbb{K},0}(P)(1_{\mathbb{K}}) &= \underbrace{1_{\mathbb{K}} \cdot 1_{\mathbb{K}}}_{1_{\mathbb{K}}} + \underbrace{1_{\mathbb{K}} \cdot \underbrace{1_{\mathbb{K}}^2}_{=1_{\mathbb{K}}}}_{1_{\mathbb{K}}} = 1_{\mathbb{K}} + 1_{\mathbb{K}} = 0_{\mathbb{K}}.
\end{aligned}$$

Man beachte, dass das Bild des Nullpolynoms $0_{\mathbb{K}}$ unter $\Omega_{\mathbb{K},0}$ ebenfalls die Nullabbildung ist.

Anmerkung: Versucht man Korollar 30 (in Polynomvariante) auf den endlichen Körper $\mathbb{K} = \{0_{\mathbb{K}}, 1_{\mathbb{K}}\}$ zu übertragen, so scheitert man daran, dass bspw. für $p = 0_{\mathbb{K}}$ keine Folge wie in Korollar 30.3) existieren kann. In der Tat wäre besagte Folge bereits konstant $1_{\mathbb{K}}$, und dann wegen Axiom M1) sogar bezüglich keiner Metrik auf $\mathbb{K}$ gegen $0_{\mathbb{K}}$ konvergent. Für $\mathbb{K} = \mathbb{Q}$ ist Korollar 30 allerdings übertragbar, und $\Omega_{\mathbb{K},p}$ ist ebenfalls injektiv.

***Zusatz:** Alternativ erhält man die Injektivität von $\Omega_{\mathbb{K},p}$ auch aus Satz 33, welcher in analoger Form auch für Polynomfunktionen in beliebigen Entwicklungspunkten $p \in \mathbb{K}$ gilt:[175]

- Ist P ein Polynom vom Grad $n \geq 1$ in p, so ist $\Omega_{\mathbb{K},p}(P)$ eine Polynomfunktion vom Grad n in p, kann also wegen Satz 33 nicht die Nullabbildung sein, da $\mathbb{K}$ unendlich viele Elemente hat.
- Ist P ein Polynom vom Grad $n \geq 0$ in p mit $\Omega_{\mathbb{K},p}(P) = 0$, so muss wegen dem vorherigen Punkt bereits $n = 0$ gelten, d.h. $P = a_0 \in \mathbb{K}$ ist ein konstantes Polynom. Indem wir ein $x \in \mathbb{K}$ einsetzen, erhalten wir $a_0 = \Omega_{\mathbb{K},p}(P)(x) = 0$, d.h. P ist das Nullpolynom.
- Somit hat die lineare Abbildung $\Omega_{\mathbb{K},p}$ trivialen Kern, ist also injektiv.

Literatur

[1] Th. Bröcker, “Analysis I”, Spektrum Akad. Verlag, Heidelberg, 2. Auflage, 1995.

[2] Th. Bröcker, “Analysis II”, Spektrum Akad. Verlag, Heidelberg, 2. Auflage, 1995.

[3] K. Endl, W. Luh, “Analysis I”, Akad. Verlagsgesellschaft, Wiesbaden, 6. Auflage, 1980.

[4] O. Forster, “Analysis 1”, Springer Spektrum, 12. Auflage, 2015.

[5] L. Führer, “Allgemeine Topologie mit Anwendung”, Vieweg, Braunschweig 1977.

[6] H. Heuser, “Lehrbuch der Analysis, Teil I”, B. G. Teubner, Stuttgart, 1984.

[7] H. Holdgrün, “Analysis 1”, Leins Verlag, Göttingen, 1998.

[8] L. Hörmander, “The Analysis of Linear Partial Differential Operators I”, Springer-Verlag Berlin Heidelberg, 2. Aufl., 2003.

[9] K. Jänich, “Funktionentheorie: Eine Einführung”, Springer; 6. Aufl., 2004.

[10] K. Königsberger, “Analysis I und II”, Springer-Verlag, 6. Auflage, 2004.

[11] W. Rudin, “Principles of Mathematical Analysis”, McGraw-Hill Education Ltd, 3. Edition, 1976.

[175] Ist P eine Polynomfunktion in $p \in \mathbb{K}$, so ist $Q\colon \mathbb{K} \ni x \mapsto P(p+x)$ eine Polynomfunktion in 0; und es gilt für $z \in \mathbb{K}$: $Q(z) = 0 \quad \Leftrightarrow \quad P(p+z) = 0$.