Skript zu den Vorlesungen

Analysis 1 und 2

Universität Augsburg

akademisches Jahr 2012/13

Vorwort

Dieses Skript entstand zu den Anfängervorlesungen „Analysis 1“ und „Analysis 2“ im akademischen Jahr 2012/13 an der Universität Augsburg.

Der Inhalt entspricht einer im deutschsprachigen Raum durchaus üblichen Vorlesung über Funktionen einer und mehrerer reeller Veränderlicher für das erste Studienjahr. Aufbau, Darstellung und Argumentation lehnen sich stark an die klassischen und beliebten Lehrbücher [Fo1, Fo2] und [K1, K2] an. Das einführende Kapitel 0 orientiert sich hingegen stark an [A1], während dem Kapitel über das Riemann–Stieltjes–Integral die entsprechenden Abschnitte in [W2] und [GBGW2] zugrunde liegen. Die Lebensdaten einiger erwähnter großer Mathematiker sind, wenn nicht allgemein bekannt, größtenteils dem Buch [GIS] entnommen.

Üblicherweise wird am Anfang eines Kapitels oder Abschnitts kurz angegeben an welchen Lehrbüchern sich die Darstellung orientiert. So können auch anhand der Literatur den Stoff nacharbeiten und bei Bedarf weiter vertiefen. Am Ende des Skriptes befindet sich ein Schlagwortverzeichnis, welches den Gebrauchswert dieses Textes erhöhen soll. Ebenfalls am Ende finden Sie ein kurzes Literaturverzeichnis. An der Universität Augsburg online verfügbare Bücher sind entsprechend verlinkt.

Zentrale Aussagen werden als „Theorem“ bezeichnet und sind oft mit Namen (z. B. „Satz von Bolzano-Weierstraß“) oder mit thematischen Bezeichnungen (z. B. „Mittelwertsatz“) verbunden.

Der Text enthält blaue Einschübe. In diesen Einschüben versuche ich anschauliche, aber dafür oft nicht sehr präzise Erklärungen zu geben, also quasi „aus dem Nähkästchen zu plaudern“. Die Argumentationen und Erläuterungen im blauen Text genügen also sehr oft nicht strengen mathematischen Anforderungen. *Insbesondere sind Argumentationen nach der Art des blauen Textes für einen Beweis und somit zur Lösung von Übungs- oder Klausuraufgaben häufig nicht hinreichend präzise und somit oftmals ungeeignet.* Dieser blaue Text sollte noch weniger als der Rest dieses Skriptes auf die Goldwaage gelegt werden. Ich hoffe allerdings, einigen Studierenden mit dem blauen Text eine Hilfestellung zum Verständnis des Stoffes geben zu können.

Zahlreiche Druckfehler wurden bisher von Studierenden entdeckt und mir mitgeteilt.

Herzlichen Dank dafür! Leider sind sicher noch sehr sehr viele Fehler in diesem Skript enthalten, welche ihrer Entdeckung harren. Dafür bitte ich um Entschuldigung. Für Hinweise auf Fehler in diesem Skript bin ich immer sehr dankbar.

Inhaltsverzeichnis

Erster Teil

0 Grundlagen

In diesem Kapitel orientieren wir uns vor allem an [A1, Kap. 1].

0.1. Mathematische Logik

Aussagen und ihre Verknüpfungen

In diesem Abschnitt führen wir kurz einige grundlegende Begriffe der mathematischen Logik ein. Hierbei orientieren wir uns an [A1, Abschnitt 1.1]. Den Einen mag dieser Abschnitt ein wenig wie Trockenschwimmübungen vorkommen, und damit haben sie sicherlich nicht ganz unrecht. Den Anderen mag es in diesem Abschnitt an der ein oder anderen Stelle an Präzision oder Tiefgang fehlen. Auch dieser Eindruck ist nicht von der Hand zu weisen. Wer mehr über mathematische Logik lernen möchte, sei auf die übliche Fachliteratur, z. B. auf [EFT] verwiesen.

Unter einer *Aussage* (im aristotelischen[1] Sinn) verstehen wir ein „sprachliches Konstrukt“, welches entweder wahr (w) oder falsch (f), aber nicht beides ist. Eine Aussage muss also einen der beiden *Wahrheitswerte* (w/f) annehmen.

Bemerkung 0.1. (i) Der Wahrheitswert einer Aussage A hängt weder von der persönlichen Meinung, Überzeugung, dem eigenen Geschmack, noch von einer Interpretation oder Auslegung der Begriffe ab. Er ist in einem gewissen Sinne universell. Sätze wie „Die Musik von Mozart ist schön“ oder „Bachs ≪Kunst der Fuge≫ ist die Krone dieser Kompositionstechnik“ sind also für uns keine Aussagen.

(ii) Es ist bei einer Aussage nicht wichtig, ob wir entscheiden können, ob eine Aussage wahr oder falsch ist. Der Satz „Jede gerade natürliche Zahl, welche größer als 2 ist, ist die Summe zweier Primzahlen.“ (starke Goldbachsche[2]-Vermutung) kann nur entweder wahr oder falsch sein und ist somit eine Aussage. Zur Zeit ist jedoch der Wahrheitswert dieser Aussage nicht bekannt.

[1] Aristoteles (384 v. Chr. - 322 v. Chr.), antiker griechischer Philosoph
[2] Christian Goldbach (1690-1764), deutscher Gelehrter

Zwei Aussagen A und B können miteinander verknüpft werden. Die wichtigsten *Verknüpfungen* sind hierbei:

Fachbegriff	**Notation**	**Sprechweise**
Negation	$\neg A$	„nicht A"
Konjunktion	$A \wedge B$	„A und B"
Disjunktion	$A \vee B$	„A oder B"
Implikation	$A \implies B$ $B \impliedby A$	„aus A folgt B"/ „A impliziert B"/ „wenn A, dann B" „A ist hinreichend für B"/ „B ist notwendig für A"
Bijunktion	$A \iff B$	„A ist äquivalent zu B"/ „A ist gleichwertig mit B" „A genau dann, wenn B"/ „A dann und nur dann, wenn B" „A ist notwendig und hinreichend für B"

Diese Wahrheitswerte dieser Verknüpfungen in Abhängigkeit von den Wahrheitswerten von A und B sind durch die Zeilen folgender *Wahrheitstafel* festgelegt:

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \implies B$	$A \iff B$
w	w	f	w	w	w	w
w	f	f	f	w	f	f
f	w	w	f	w	w	f
f	f	w	f	f	w	w

Bemerkung 0.2. (i) Wenn A und B beide wahr sind, so ist auch $A \vee B$ wahr. Die Disjunktion ist also kein ausschließendes „entweder, oder".
Das Symbol $\vee$ für die Disjunktion erinnert an das lateinische Wort „vel" für „oder".

(ii) Wenn A falsch ist, so ist die Implikation $A \implies B$ immer (unabhängig vom Wahrheitswert von B) wahr.

(iii) Die Bijunktion $A \iff B$ ist genau dann wahr, wenn die Wahrheitswerte von A und B übereinstimmen.

(iv) Ist A eine Aussage, so ist die Aussage $(A \vee (\neg A))$ unabhängig von den Wahrheitswerten von A immer wahr und $(A \wedge (\neg A))$ immer falsch.

Ein Aussage, welche sich aus Teilaussagen zusammensetzt, und welche für alle möglichen Kombinationen von Wahrheitswerten der Teilaussagen immer wahr ist, heißt *Tautologie.* Für die Beweistechnik in der Mathematik wichtige Tautologien sind:

Lemma 0.3. *Für drei Aussagen A, B und C sind folgende Aussagen Tautologien:*

(i) $(A \iff B) \iff ((A \implies B) \wedge (B \implies A))$.[3]

[3] Diese Tautologie bedeutet, dass man den Beweis der Aussage $(A \iff B)$ in zwei Schritte zerlegen kann. Man kann nämlich getrennt die Aussagen $(A \implies B)$ und $(B \implies A)$ zeigen, was in der Praxis oft gemacht wird.

(ii) $(A \implies B) \iff ((\neg B) \implies (\neg A))$.[4]

(iii) Transitivität der Implikation: $((A \implies B) \wedge (B \implies C)) \implies (A \implies C)$.

(iv) $(\neg(A \implies B)) \iff (A \wedge (\neg B))$.[5]

Beweis. Wir zeigen für (i) mit Hilfe einer Wahrheitstafel exemplarisch, dass die Aussage

$$((A \iff B) \iff ((A \implies B) \wedge (B \implies A))),$$

welche wir abkürzend mit C bezeichnen, für alle Kombinationen der Wahrheitswerte von A und B wahr ist:

A	B	$A \iff B$	$A \implies B$	$B \implies A$	$(A \implies B) \wedge (B \implies A)$	C
w	w	w	w	w	w	w
w	f	f	f	w	f	w
f	w	f	w	f	f	w
f	f	w	w	w	w	w

Die Beweise der anderen Tautologien ist dem Leser als Übungsaufgabe überlassen. □

Lemma 0.4. *Seien A, B und C Aussagen. Dann sind folgende Aussagen Tautologien:*

(i) $(A \wedge A) \iff A$.

(ii) $(A \vee A) \iff A$.

(iii) $\neg(\neg A) \iff A$.

(iv) Kommutativgesetze für Konjunktion und Disjunktion:

$$\begin{aligned} (A \wedge B) &\iff (B \wedge A), \\ (A \vee B) &\iff (B \vee A). \end{aligned}$$

(v) Assoziativgesetze für Konjunktion und Disjunktion:

$$\begin{aligned} (A \wedge (B \wedge C)) &\iff ((A \wedge B) \wedge C), \\ (A \vee (B \vee C)) &\iff ((A \vee B) \vee C). \end{aligned}$$

(vi) Distributivgesetze für Konjunktion und Disjunktion:

$$\begin{aligned} (A \vee (B \wedge C)) &\iff ((A \vee B) \wedge (A \vee C)), \\ (A \wedge (B \vee C)) &\iff ((A \wedge B) \vee (A \wedge C)). \end{aligned}$$

---

[4] Diese Tautologie führt zum *Beweis durch Kontraposition:* Anstelle $(A \implies B)$ zu zeigen, beweist man $((\neg B) \implies (\neg A))$.

[5] Diese Tautologie führt zum *Beweis durch Widerspruch:* Möchte man $(A \implies B)$ zeigen, so kann man zeigen, dass $(A \wedge (\neg B))$ falsch ist, denn dann ist auch $(\neg(A \implies B))$ falsch und damit $(A \implies B)$ wahr.

(vii) De Morgansche[6] Regeln :

$$\begin{aligned}(\neg(A \wedge B)) &\iff (\neg A) \vee (\neg B),\\ (\neg(A \vee B)) &\iff (\neg A) \wedge (\neg B).\end{aligned}$$

(viii) Abtrennungsregeln:

$$\begin{aligned}(A \wedge (A \implies B)) &\implies B,\\ ((\neg B) \wedge (A \implies B)) &\implies (\neg A).\end{aligned}$$

Der Beweis ist wiederum eine Übung.

Quantoren

Wir folgen weiterhin dem empfehlenswerten Buch [A1, Abschnitt 1.1].

Eine *Aussageform* ist ein „sprachliches Gebilde“, in welchem eine oder mehrere variable Größen (Variable) auftreten, welche jedoch einem festgelegten *Geltungsbereich* entstammen. Zudem fordern wir: Setzt man für alle variable Größen, die in einer gegebenen Aussageform vorkommen, konkrete Werte aus dem Geltungsbereich ein, so erhält man eine Aussage.

Alle Werte aus dem Geltungsbereich der Variablen einer Aussageform, welche zu einer wahren Aussagen führen, bilden zusammen den *Erfüllungsbereich* der Aussageform.

Beispiel 0.5. Die Aussageform „Das Quadrat einer natürlichen Zahl x ist gleich einer geraden natürlichen Zahl y“, ist z. B. wahr für $x = 2$ und $y = 4$, falsch für $x = 2$ und $y = 8$ und falsch für $x = 3$ und jede beliebige gerade Zahl y.

Eine andere Art aus Aussageformen Aussagen zu gewinnen stellen die beiden *Quantoren*

- der *Allquantor* $\forall$ und
- der *Existenzquantor* $\exists$

dar. Ist $A(x)$ eine Aussageform in der nur die variable Größe x vorkommt, so erhalten wir folgende Aussagen:

$(\forall\, x :\ A(x))$	steht für	„Für *jeden* Wert x aus dem Geltungsbereich gilt $A(x)$.“,
$(\exists\, x :\ A(x))$	steht für	„Es existiert (mindestens ein) Wert x aus dem Geltungsbereich, sodass $A(x)$ gilt.“.

[6] Augustus De Morgan (1806 - 1871), britischer Mathematiker. Diese Regeln waren aber wohl bereits im 14. Jahrhundert bekannt.

Enthält eine Aussage mehrere Quantoren hintereinander, so gilt folgenden Konvention: *Jede variable Größe in einer Aussage, welche zu einem Existenzquantor gehört, hängt von allen links von ihr in der Aussage bereits aufgetretenen Variablen ab.*

Quantoren sind nicht kommutativ, denn in der Aussage

$$\forall\, \varepsilon\, \forall\, x\, \exists\, \delta : A(\varepsilon, \delta, x)$$

hängt nun δ von ε wie auch von x ab. In der Aussage

$$\forall\, \varepsilon\, \exists\, \delta\, \forall\, x : A(\varepsilon, \delta, x)$$

hängt hingegen δ nur von ϵ, nicht aber von x ab. Dieser formal feine, aber in der Tat große Unterschied, wird in dieser Vorlesung noch eine wichtige Rolle spielen. Folgende umgangssprachliche Formulierungen sollen die Nichtkommutativität von Quantoren noch einmal verdeutlichen:

„Für jeden Einwohner Bayerns gibt es eine (Verwaltungs-)Gemeinde, in welcher sein erster Wohnsitz liegt.“

„Es gibt eine (Verwaltungs-)Gemeinde, sodass alle Einwohner Bayerns dort ihren ersten Wohnsitz haben.“

Für die Negation einer Aussage mit Quantoren gilt:

$$\begin{aligned} (\neg(\forall\, x : A(x))) &\iff (\exists\, x : (\neg A(x)), \\ (\neg(\exists\, x : A(x))) &\iff (\forall\, x : (\neg A(x)). \end{aligned}$$

Um also nachzuweisen, dass die Aussage $(\forall\, x : A(x))$ falsch ist, genügt es ein *Gegenbeispiel* zu finden, d. h. ein spezielles x_0 aus dem Geltungsbereich anzugeben, für welches $(\neg A(x_0))$ wahr ist.

Treten mehrere Quantoren in einer Aussage auf, so erhält man die Negation iterativ indem man jeweils den Quantor $\forall$ durch $\exists$ und den Quantor $\exists$ durch $\forall$ ersetzt und dann die Aussageform am Ende negiert. So gilt zum Beispiel:

$$(\neg(\forall\, \varepsilon\, \exists\, \delta\, \forall\, x : A(\varepsilon, \delta, x))) \iff (\exists\, \varepsilon\, \forall\, \delta\, \exists\, x : (\neg A(\varepsilon, \delta, x)))$$

Man kann den Existenzquantor noch verschärfen, indem man fordert, dass nicht nur mindestens ein Wert im Geltungsbereich existiert, sondern dass es genau einen Wert gibt im Geltungsbereich gibt. Man schreibt dann kurz:

$$\exists! \quad \text{lies} \quad \text{„es existiert genau ein“}$$

Zwar werden wir in den folgenden Kapiteln aus stilistischen Gründen unsere mathematischen Sätze meistens in gewöhnlicher Sprache niederschreiben, jedoch ist es manchmal hilfreich, diese Aussagen dann noch einmal in eine formale Sprache zu übersetzen, da man die formale Sprache besser manipulieren kann (z. B. bei Negationen). Eine gute und klar verständliche mathematische Darstellung ist für die Kommunikation mit anderen notwendig (z. B. in Prüfungen (Staatsexamina) und bei mathematischen Veröffentlichungen). Das Erlernen einer guten mathematischen Darstellung, sozusagen der mathematischen Sprache, ist ein wichtiges Ziel der ersten Studiensemester.

0.2. Naive Mengenlehre

In diesem Abschnitt orientieren wir uns an [A1, Abschnitt 1.2].

Der Begriff der Menge

Mengen sind grundlegende Objekte der Mathematik. Eine exakte Definition des Begriffs *Menge* ist etwas aufwendig. Für unsere Bedürfnisse genügt die (etwas problematische) Vorstellung Cantors[7] von einer Menge als eine Sammlung gewisser Objekte, welche *Elemente* der Menge genannt werden[8]. Eine übliche Art den anschaulichen Begriff einer Menge mathematisch präzise zu erfassen ist die axiomatische Mengenlehre[9] nach Zermelo[10] und Fraenkel[11].

Ist X eine Menge und x ein Element von X, so drücken wir diese Tatsache durch die Notation

$$x \in X$$

aus. Ist für ein Objekt x und eine Menge X die Aussage $x \in X$ falsch, so schreiben wir kurz

$$x \notin X.$$

Unserem naiven Mengenbegriff liegt die Vorstellung zugrunde, dass eine Menge durch ihre Elemente eindeutig bestimmt ist. Daher ist $y \in X$ für ein gegebenes Objekt y und eine gegebene Menge X immer eine Aussage. Zwei Mengen X und Y nennen wir (unserem naiven Mengenbegriff entsprechend) *gleich*[12], notiert

$$X = Y,$$

[7] Georg Ferdinand Ludwig Philipp Cantor (1845 - 1918), deutscher Mathematiker

[8] Cantor formulierte, was er unter einer Menge versteht, wie folgt:

„Unter einer „Menge" verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die „Elemente" von M genannt werden) zu einem Ganzen."

(siehe G. Cantor, Beiträge zur Begründung der transfiniten Mengenlehre, *Math. Ann.* **46** (1895)). Dieser anschauliche Mengenbegriff ist jedoch keinesfalls unproblematisch. In seinem Buch „Principles of Mathematics“ (Cambridge 1903), machte Bertrand Russel auf ein Paradoxon in der naiven Mengenlehre aufmerksam (Russellsche Antinomie), indem er die Menge M derjenigen Mengen betrachtete, welche sich selbst nicht als Element enthalten. Wenn man nun annimmt, dass M selbst ein Element von M ist, so folgt, dass M nicht in M enthalten sein kann, ein Widerspruch. Nimmt man hingegen an, dass M sich nicht selbst als Element enthält, dann ist M aber ein Element von M, wiederum ein Widerspruch. Man muss also den Begriff der Menge enger und mathematisch präziser fassen.

[9] Hier stellt sich nun ein anderes Problem: Nach dem zweiten Unvollständigkeitssatz von Kurt Gödel (1906 - 1978) kann die Widerspruchsfreiheit (Konsistenz) des Zermelo-Fraenkelschen Axiomssystems nur dann aus diesem selbst heraus bewiesen werden, wenn es inkonsistent wäre. Für eine detailliertere Beschreibung der Geschichte der Mengenlehre und deren axiomatische Fassungen verweisen wir auf das auch ansonsten interessante Buch [Ebb, Kap. 14].

[10] Ernst Friedrich Ferdinand Zermelo (1871 - 1953), deutscher Mathematiker

[11] Abraham Halevi Fraenkel (1891 - 1965), aus Bayern stammender israelischer Mathematiker

[12] Diese Definition der Gleichheit von Mengen entspricht im Axiomssystem von Zermelo-Fraenkel dem *Extensionalitätsaxiom*.

wenn jedes Element von X auch ein Element von Y ist und umgekehrt, d.h.

$$(X = Y) :\iff ((\forall\, x \in X :\ x \in Y) \wedge (\forall\, y \in Y :\ y \in X))^{13}.$$

Sind zwei Mengen X und Y nicht gleich, so schreibt man $X \neq Y$, d. h.

$$(X \neq Y) :\iff (\neg(X = Y)).$$

Eine Menge X heißt *Teilmenge* einer Menge Y, notiert $X \subset Y$ oder gleichbedeutend $X \subseteq Y$, wenn jedes Element von X auch ein Element von Y ist, formal

$$(X \subset Y) :\iff (\forall x \in X : x \in Y).$$

Damit ist die Aussage

$$(X = Y) \iff ((X \subset Y) \wedge (Y \subset X))$$

eine Tautologie[14]. Wenn man in einer Notation ausdrücken möchte, dass X eine Teilmenge von Y ist, aber X nicht gleich Y ist, so schreibt man $X \subsetneq Y$ und es gilt

$$(X \subsetneq Y) \iff ((X \subset Y) \wedge (X \neq Y)).$$

Man sagt dann auch: X ist eine *echte Teilmenge* von Y.

Mengen können in verschiedener Weise beschrieben werden und werden gewöhnlich mit geschweiften Klammern $\{\ldots\}$ dargestellt. Manchmal kann man die Elemente einfach der Reihe nach angeben, z. B.

$$\{\diamondsuit, \heartsuit, \spadesuit, \clubsuit\} \quad \text{oder} \quad \{0, 1, 2, 3, 4, \ldots\}.$$

Nach unserer Definition der Gleichheit von Mengen spielt bei einer solchen Aufzählung die Reihenfolge keine Rolle. Auch ändert sich die Menge nicht, wenn man Elemente doppelt zählt. So gilt z. B.

$$\{\diamondsuit, \heartsuit, \spadesuit, \clubsuit\} = \{\spadesuit, \clubsuit, \diamondsuit, \heartsuit\} = \{\diamondsuit, \heartsuit, \spadesuit, \clubsuit, \diamondsuit, \heartsuit\}.$$

Eine besondere Menge ist die *leere Menge* $\emptyset = \{\}$, welche kein Element enthält[15]. Die leere Menge ist Teilmenge einer jeden Menge Y, da die definierende Eigenschaft für eine Teilmenge, nämlich hier $\forall x \in \emptyset : x \in Y$, in diesem Fall als „leere“ Aussage wahr ist.

Sei $A(x)$ eine Aussageform deren Geltungsbereich G eine Menge ist, dann bilden alle Elemente aus G, für die $A(x)$ wahr ist (also der Erfüllungsbereich der Aussageform $A(x)$) eine Menge, welche kurz mit

$$\{x \in G \mid A(x)\}$$

[13] Hier haben wir eine neue Notation eingeführt, nämlich $:\iff$. Dieses Symbol besagt, dass die links stehende Notation (welche noch nicht erklärt ist) durch die rechte Aussage (welche bereits bekannt ist) definiert wird, sodass eine Tautologie entsteht.

[14] Diese spielt in der Beweistechnik eine wichtige Rolle. Die Gleichheit zweier Mengen X und Y kann man in zwei Schritten zeigen, indem man sowohl $X \subset Y$ wie auch $Y \subset X$ beweist.

[15] In der axiomatischen Mengenlehre nach Zermelo-Fraenkel wird die Existenz einer solchen Menge im *Leeremengenaxiom* gefordert.

bezeichnet wird[16].

Sie kennen bereits einige Mengen aus Ihrer Schulzeit, z. B.

$$\begin{aligned}
\mathbb{N} &= \{0,1,2,\ldots\} \quad \text{(nach DIN 5473)}\\
\mathbb{N}^* &= \{1,2,\ldots\},\\
\mathbb{Z} &= \{0,\pm1,\pm2,\ldots\},\\
\mathbb{Z}^* &= \{\pm1,\pm2,\ldots\}.
\end{aligned}$$

Operationen mit Mengen

Die Menge aller Teilmengen einer Menge Y heißt *Potenzmenge*[17] von Y, notiert

$$\mathcal{P}(Y) := \{X \ : \ X \subset Y\}.$$

Definition 0.6. Für zwei Mengen X und Y definieren wir:

- die *Vereinigung*[18] $X \cup Y$ von X und Y als

$$X \cup Y := \{x \mid (x \in X) \vee (x \in Y)\},$$

- den *Durchschnitt* $X \cap Y$ von X und Y als

$$X \cap Y := \{x \mid (x \in X) \wedge (x \in Y)\},$$

- die *Differenzmenge* $X \setminus Y$ durch

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

Zwei Mengen X und Y heißen *disjunkt*, wenn $X \cap Y = \emptyset$.

Sei I eine beliebige Indexmenge und X_i, $i \in I$, Mengen, so definieren wir deren Vereinigung durch

$$\bigcup_{i\in I} X_i := \{x \mid (\exists\, i \in I : x \in X_i)\}$$

und deren Durchschnitt durch

$$\bigcap_{i\in I} X_i := \{x \mid (\forall\, i \in I : x \in X_i)\}.$$

Zum Beispiel erhalten wir für $X_i = \{x \in \mathbb{Z} \mid x \geq i\}$, $i \in \mathbb{N}$, somit $\bigcup_{i\in\mathbb{N}} X_i = \mathbb{N}$ und $\bigcap_{i\in\mathbb{N}} X_i = \emptyset$.

[16] Dies entspricht dem *Aussonderungsaxiom* der Zermelo-Fraenkelschen axiomatischen Mengenlehre.

[17] In der Zermelo-Fraenkelschen Mengenlehre wird die Existenz dieser Potenzmenge im *Potenzmengenaxiom* gefordert.

[18] In der Zermelo-Fraenkelschen Mengenlehre fordert das *Vereinigungsaxiom* die Existenz einer solchen Menge.

Lemma 0.7. *Seien X, Y, Z Mengen, dann gilt:*

(i) $X \cap X = X = X \cup X$,

(ii) Kommutativgesetze: $X \cup Y = Y \cup X$ *sowie* $X \cap Y = Y \cap X$,

(iii) Assoziativgesetze: $(X \cup Y) \cup Z = X \cup (Y \cup Z)$ *sowie* $(X \cap Y) \cap Z = X \cap (Y \cap Z)$,

(iv) Distributivgesetze:
$X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$ *sowie* $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$.

Der Beweis dieses Lemmas sind dem Leser zur Übung überlassen.

Aufgabe 0.8. Seien X und Y Mengen. Zeigen Sie:

$$((X \cup Y) = (X \cap Y)) \iff (X = Y).$$

Definition 0.9. Sei X eine Menge und $U \subset X$ eine Teilmenge, dann heißt die Menge

$$C_X(U) = X \setminus U$$

das *Komplement* von U in X.

Lemma 0.10. *Sei X eine Menge und U, V zwei Teilmengen von X. Dann gilt*

(i) $U \cup C_X(U) = X$ *und* $U \cap C_X(U) = \emptyset$,

(ii) $C_X(C_X(U)) = U$,

(iii) De Morgansche *Regeln*[19]*:* $C_X(U \cap V) = C_X(U) \cup C_X(V)$ *sowie* $C_X(U \cup V) = C_X(U) \cap C_X(V)$.

Der Beweis dieses Lemmas ist wiederum ein Übungsaufgabe.

Das *kartesische Produkt* von einer Menge X mit einer Menge Y, notiert $X \times Y$, ist definiert als

$$X \times Y := \{(x, y) \mid (x \in X) \wedge (y \in Y)\}$$ [20].

Wenn $X \neq Y$, dann gilt $X \times Y \neq Y \times X$. Für endlich viele Mengen $X_1, \ldots, X_n$, $n \in \mathbb{N}$, kann man das kartesische Produkt auch einfach definieren:

$$X_1 \times \cdots \times X_n = \{(x_1, \ldots, x_n) \mid (\forall i \in \{1, \ldots, n\} : x_i \in X_i)\}$$ [21].

Wie am Beispiel $\mathbb{R}^n$ bereits gesehen, kann man auch das n-fache kartesische Produkt

$$X^n := \underbrace{X \times \cdots \times X}_{n\text{-mal}} = \{(x_1, \ldots, x_n) \mid x_1, \ldots, x_n \in X\}$$

einer Menge X mit sich selbst bilden.

[19] Man beachte die formale Analogie zu den De Morganschen Regeln der Logik (siehe Lemma 0.4).

[20] Hierbei gilt $(x, y) = (x', y')$ genau dann, wenn $x = x'$ und $y = y'$.

[21] Hierbei gilt dann $(x_1, \ldots, x_n) = (x'_1, \ldots, x'_n)$ genau dann, wenn $x_i = x'_i$ für alle $i \in \{1, \ldots, n\}$.

Aufgabe 0.11. Seien U, X, Y und Z Mengen. Zeigen Sie:

(i) $(X \times Y = \emptyset) \iff ((X = \emptyset) \vee (Y = \emptyset))$,

(ii) $(X \times Z) \cup (Y \times Z) = (X \cup Y) \times Z$,

(iii) $(U \cap X) \times (Y \cap Z) = (U \times Y) \cap (X \times Z)$.

0.3. Ordnungsrelationen

Wir orientieren uns in diesem Abschnitt an [A1, Abschn. 1.3.3].

Seien X und Y Mengen, so heißt eine Teilmenge $R \subset X \times Y$ *Relation* zwischen X und Y. Im Fall $X = Y$ sprechen wir einfach von einer Relation auf X. Ist $R \subset X \times Y$, so heißt die Teilmenge

$$R^{-1} = \{(y, x) \in Y \times X \mid (x, y) \in R\} \subset Y \times X$$

die *Umkehrrelation* von R.

Eine Relation $R \subset X \times X$ auf einer Menge X heißt:

- *reflexiv,* wenn gilt: $\forall x \in X : \ (x, x) \in R$,
- *transitiv,* wenn gilt: $\forall x, y, z \in X : ((x, y) \in R \wedge (y, z) \in R) \implies (x, z) \in R$,
- *symmetrisch,* wenn gilt: $\forall x, y \in X : (x, y) \in R \implies (y, x) \in R$,
- *antisymmetrisch,* wenn gilt: $\forall x, y \in X : ((x, y) \in R \wedge (y, x) \in R) \implies (x = y)$.

Definition 0.12. Sei X eine Menge. Eine reflexive, transitive und antisymmetrische Relation $O \subset X \times X$ heißt *Ordnungsrelation* auf X.

Eine Menge X mit einer Ordnungsrelation $O \subset X \times X$ heißt *geordnete Menge*, wenn gilt:

$$\forall x, y \in X : ((x, y) \in O \vee (y, x) \in O).$$

Ist O eine Ordnungsrelation auf X, so schreiben wir auch

$$x \preceq y :\iff (x, y) \in O.$$

Beispiel 0.13. (a) Sei M eine Menge und $X = \mathcal{P}(M)$ die Potenzmenge von M, dann ist

$$O = \{(U, V) \in \mathcal{P}(M) \times \mathcal{P}(M) : \ U \subseteq V\}$$

eine Ordnungsrelation auf X. Wenn M mindestens zwei (verschiedene) Elemente hat, so ist X mit der Ordnungsrelation O aber *keine* geordnete Menge.

(b) Die Menge $\mathbb{N}$ der natürlichen Zahlen ist bzgl. der Ordnungsrelation

$$O = \{(x, y) \in \mathbb{N} \times \mathbb{N} : \ x \leq y\}$$

eine geordnete Menge[22].

[22]Hierbei steht $\leq$ für das bereits aus der Schule bekannte „kleiner oder gleich“.

Sei X eine Menge mit einer Ordnungsrelation $O \subset X \times X$ und $x, y \in X$, so schreiben wir

$$x \prec y :\iff ((x \preceq y) \wedge (x \neq y)).$$

Theorem 0.14. *Sei X eine geordnete Menge, dann gilt*

(i) Für je zwei Elemente $x, y \in X$ gilt genau eine der drei Möglichkeiten:
- $x \prec y$,
- $x = y$,
- $y \prec x$.

(ii) $\forall x, y, z \in X : \ ((x \prec y) \wedge (y \prec z)) \implies (x \prec z)$.

Beweis. ad (i): Wir unterscheiden zwei Fälle:

a) Sei $x = y$, dann ist $(x \prec y) \vee (y \prec x)$ nach Definition von $\prec$ falsch und die Aussage bewiesen.

b) Sei $x \neq y$. Aufgrund der Antisymmetrie einer Ordnungsrelation ist $((x \prec y) \wedge (y \prec x))$ immer falsch. Daher genügt es in unserem Fall die Aussage

$$((x \prec y) \vee (y \prec x))$$

zu beweisen.
Da X geordnet ist gilt $((x \preceq y) \vee (y \preceq x))$. Zusammen mit unserer Voraussetzung trifft also $(x \neq y) \wedge ((x \preceq y) \vee (y \preceq x))$ zu. Aufgrund des Distributivgesetzes für Aussagen (siehe 0.4) erhalten wir

$$((x \preceq y) \wedge (x \neq y)) \vee ((y \preceq x) \wedge (x \neq y))$$

also

$$(x \prec y) \vee (y \prec x).$$

ad (ii): Zunächst folgt aus $(x \prec y) \wedge (y \prec x)$ sofort $((x \preceq y) \wedge (y \preceq x))$ und damit wegen der Transitivität von Ordnungsrelationen $x \preceq z$. Wir müssen also noch $x \neq z$ zeigen. Sei also $x = z$, so erhalten wir $(x = z) \wedge (x \prec y)$ und damit $(z \prec y)$. Jedoch ist $(z \prec y) \wedge (y \prec z)$ nach Teil (i) falsch.

□

0.4. Äquivalenzrelationen

Manchmal möchte man eine Menge in Teilmengen zerlegen, zum Beispiel um Elemente nach gewissen Eigenschaften zu sortieren. Zum Beispiel möchte man vielleicht die Menge aller Früchte in einem Supermarkt zerlegen in die Menge der Äpfel, der Birnen, der Zitrusfrüchte, etc..

Eine *Zerlegung* einer Menge $X \neq \emptyset$ ist eine Menge $\{X_i : \ i \in I\} \subset \mathcal{P}(X)$ von Teilmengen sodass $X = \bigcup_{i \in I} X_i$. Eine Zerlegung $\{X_i : i \in I\}$ von X heißt *disjunkt*, wenn zudem gilt:

(i) $X_i \neq \emptyset$ für alle $i \in I$,

(ii) X_i und X_j, $i \neq j$, sind immer *disjunkte* Teilmengen von X, d. h. $\forall i, j \in I, i \neq j :$ $X_i \cap X_j = \emptyset$.

Um deutlich zu machen, dass X die Vereinigung disjunkter Mengen ist schreibt man dann auch $X = \bigsqcup_{i \in I} X_i$. Zum Beispiel gilt:

- $\{1, 2, 3, 4\} = \{1, 2\} \sqcup \{3, 4\}$,
- $\mathbb{N}^2 = \bigsqcup_{y \in \mathbb{N}} \{(x, y) : \ x \in \mathbb{N}\}$.

Wenn eine disjunkte Zerlegung $X = \bigsqcup_{i \in I} X_i$ gegeben ist, so sagen wir $x \in X$ ist äquivalent zu $x' \in X$, kurz $x \sim x'$, wenn es ein $i \in I$ gibt, sodass $x, x' \in X_i$. Die Relation $x \sim x'$ hat folgende drei Eigenschaften:

(Ä1) *Reflexivität:* $\forall x \in X : \ x \sim x$,

(Ä2) *Symmetrie:* $\forall x, x' \in X : \ (x \sim x' \implies x' \sim x)$,

(Ä3) *Transitivität:* $\forall x, x', x'' \in X : \ (x \sim x' \text{ und } x' \sim x'' \implies x \sim x'')$.

Etwas formaler definieren wir eine Äquivalenzrelation als eine Relation $R \subset X \times X$ auf X, welche aus allen Paaren von Elementen aus X besteht, welche wir als äquivalent ansehen möchten.

Definition 0.15. Eine *Äquivalenzrelation* auf einer Menge $X \neq \emptyset$ ist eine reflexive, symmetrische und transitive Relation $R \subset X \times X$ auf X.

Sei R eine Äquivalenzrelation auf X, so nennen wir zwei Elemente x und x' aus X *äquivalent,* kurz $x \sim x'$, wenn $(x, x') \in R$. Weiter bezeichnen wir mit

$$[x] := \{x' \in X \ : \ x \sim x'\} = \{x' \in X : \ (x, x') \in R\} \in \mathcal{P}(X)$$

die *Äquivalenzklasse* von x. Die Menge

$$X/\!\sim \, := \{[x] \ : \ x \in X\} \subset \mathcal{P}(X)$$

heißt der *Quotient* von X nach der Äquivalenzrelation R.

An dieser Stelle möchten wir ein oft verwendetes Beispiel angeben, welches die Situation anschaulich verdeutlicht: Sei X die Menge aller Schüler einer Grundschule. Wir nennen zwei Schüler äquivalent, wenn beide in der selben Schulklasse sind. Eine Äquivalenzklasse entspricht also einer Schulklasse und der Quotient $X/\!\sim$ ist die Menge aller Schulklassen der Grundschule.

Wir haben bereits gesehen, dass eine disjunkte Zerlegung $X = \bigsqcup_{i \in I} X_i$ eine Äquivalenzrelation auf X definiert, nämlich

$$R = \{(x, x') : \ (\exists i \in I : \ x, x' \in X_i)\}.$$

Betrachten wir noch einmal das Beispiel $X = \{1,2,3,4\} = \{1,2\} \sqcup \{3,4\}$. Hier wäre $R = \{(1,1),(2,2),(3,3),(4,4),(1,2),(2,1),(3,4),(4,3)\}$ und $[1] = \{1,2\} = [2]$, $[3] = \{3,4\} = [4]$ und $X/{\sim} = \{[1],[3]\}$.

Lemma 0.16. *Sei R eine Äquivalenzrelation auf $X \neq \emptyset$, dann gilt für alle $x, x' \in X$ entweder $[x] = [x']$ oder $[x] \cap [x'] = \emptyset$.*

Beweis. Wir müssen zeigen, dass aus $[x] \cap [x'] \neq \emptyset$ folgt $[x] = [x']$.

Sei also $[x] \cap [x'] \neq \emptyset$, dann gibt es ein $x'' \in [x] \cap [x']$ also $x \sim x''$ sowie $x'' \sim x'$. Aufgrund der Transitivität folgt dann $x \sim x'$. Sei nun $y \in [x]$, dann gilt $y \sim x$ und, wie zuvor gesehen, $x \sim x'$, also $y \sim x'$ und damit $y \in [x']$. Damit haben wir $[x] \subset [x']$ gezeigt.

Analog kann man $[x'] \subset [x]$ beweisen. Insgesamt folgt also $[x] = [x']$. $\square$

Lemma 0.16 zeigt, dass für eine Äquivalenzrelation R auf $X \neq \emptyset$ der Quotient $X/{\sim}$ eine disjunkte Zerlegung von X ist, d. h.

$$X = \bigsqcup_{[x] \in X/\sim} [x].$$[23]

Für Quotienten möchten wir noch zwei instruktive Beispiele geben (siehe [He1]):

1. **Rationale Zahlen:** Wir betrachten auf $\mathbb{Z} \times \mathbb{Z}^*$ die Äquivalenzrelation
$$(z,n) \sim (z',n') :\iff zn' = z'n.$$
(Es ist eine Übungsaufgabe nachzuweisen, dass es sich tatsächlich um eine Äquivalenzrelation handelt.) Der Quotient $(\mathbb{Z} \times \mathbb{Z}^*)/{\sim}$ heißt Menge der *rationalen Zahlen* und ein Element $[(z,n)] \in (\mathbb{Z} \times \mathbb{Z}^*)/{\sim}$ heißt *rationale Zahl*. Die rationale Zahl $[(z,n)]$ schreibt man auch als $\frac{z}{n}$[24].

2. **Restklassen:** Bereits aus der Schule kennen Sie das Rechnen mit Resten bei Divisionen ganzer Zahlen, so gilt z. B. $4 \equiv 12 \mod 2$. Sei nun $n \in \mathbb{N}^*$, so definieren wir auf $\mathbb{Z}$ die Äquivalenzrelation
$$x \sim_n x' :\iff x = x' \mod n \iff n \text{ teilt } x - x' \iff \exists a \in \mathbb{Z} : x - x' = an.$$

[23] Wenn Ihnen diese Schreibweise zu abstrakt ist, so können Sie auch aus jeder von R definierten Äquivalenzklasse genau ein Element auswählen. Dieses Element wird oft *Repräsentant* seiner Äquivalenzklasse genannt. Dass Sie gleichzeitig aus allen (auch aus unendlich vielen) Äquivalenzklassen jeweils einen Repräsentaten wählen dürfen, garantiert das *Auswahlaxiom* der Zermelo-Fraenkelschen Mengenlehre. Die Menge dieser Repäsentaten bildet dann die Indexmenge I und es gilt

$$X = \bigsqcup_{x \in I} [x].$$

[24] Überlegen Sie sich, was unsere Definition mit der Ihnen aus der Schule bekannten Definition einer rationalen Zahl zu tun hat.

Es gibt offenbar genau n Äquivalenzklassen, nämlich $[0]$, $[1]$, $\ldots$, $[n-1]$. Den entsprechenden Quotienten bezeichnen wir mit

$$\mathbb{Z}_n := \mathbb{Z}/\sim_n = \{[0], [1], \ldots, [n-1]\}.$$

So ist z. B. $\mathbb{Z}_2 = \{[0], [1]\} = \{[-428], [975]\}$, wobei $[0] = [-428] = \{2z : \; z \in \mathbb{Z}\}$ die Menge aller geraden ganzen Zahlen und $[1] = [975] = \{2z+1 : \; z \in \mathbb{Z}\}$ die Menge aller ungeraden ganzen Zahlen ist.

Aufgabe 0.17. Sei $X \neq \emptyset$ eine Menge und $R_1, R_2 \subset X \times X$ zwei Äquivalenzrelationen auf X.

(a) Zeigen Sie: $R_1 \cap R_2$ ist eine Äquivalenzrelation auf X.

(b) Ist auch $R_1 \cup R_2$ eine Äquivalenzrelation auf X? Begründen Sie Ihre Antwort.

0.5. Abbildungen

Hier orientieren wir uns an [He1] und an [A1].

Definition 0.18. Eine *Abbildung* oder *Funktion* f von einer nichtleeren Menge X in eine nichtleere Menge Y ist eine Zuordnung, welche *jedem* Element x aus X ein *eindeutiges* Element $f(x)$ in Y zuordnet, notiert:

$$f : X \to Y, \; x \mapsto f(x).$$

Man nennt die nichtleere Menge X die *Definitionsmenge* (Definitionsbereich) von f und Y die *Zielmenge* von f.

Die Zielmenge einer Abbildung ist automatisch nicht leer.

Im Folgenden sind alle Definitionsmengen von Abbildungen nicht leer, auch wenn wir das nicht immer wieder angeben. Manchmal schreiben wir $X \xrightarrow{f} Y$ anstelle von $f : X \to Y$.

Auch wenn die oben gegebene Definition einer Abbildung Ihnen aus der Schule vertraut sein mag, erscheint sie vielleicht doch ein wenig unpräzise. So ist insbesondere der Begriff „Zuordnung“ nicht zuvor definiert worden. Um den Begriff Abbildung auf Begriffe aus der Mengenlehre zurückzuführen verwenden wir die Idee des Funktionsgraphen. Für eine Abbildung $f : X \to Y$ ist das die Menge $\{(x, f(x)) : \; x \in X\}$ und damit eine Teilmenge von $X \times Y$.

Formal gesprochen ist eine Funktion (oder ein *Funktionsgraph*) von einer Menge $X \neq \emptyset$ in eine Menge Y eine Relation $\Gamma_f \subset X \times Y$ mit der Eigenschaft

$$\forall x \in X \; \exists! y \in Y : (x, y) \in \Gamma_f,$$

oder äquivalent mit den beiden Eigenschaften

(i) $\forall x \in X\ \exists y \in Y:\ (x, y) \in \Gamma_f$,

(ii) $\forall x \in X\ \forall (y, z) \in Y \times Y:\ (((x, y) \in \Gamma_f) \wedge (x, z) \in \Gamma_f) \implies (y = z))$.

Aus Γ_f gewinnt man eine Abbildung $f : X \to Y$ indem man für $x \in X$ setzt $f(x) = y$, wobei $(x, y) \in \Gamma_f$. Die Relation Γ_f kann dann als

$$\Gamma_f = \{(x, f(x)) \mid x \in X\}$$

umgeschrieben werden und ist somit der bereits aus der Schule bekannte Graph der Funktion f.

Zwei Abbildungen $f_1 : X_1 \to Y_1$ und $f_2 : X_2 \to Y_2$ heißen gleich, wenn $X_1 = X_2$, $Y_1 = Y_2$ und $f_1(x) = f_2(x)$ für alle $x \in X_1 = X_2$ gilt, oder äquivalent, wenn $\Gamma_{f_1} = \Gamma_{f_2}$ und $Y_1 = Y_2$. Somit muss man in der Notation einer Abbildung nicht nur eine Zuordnungsvorschrift, sondern immer auch die Definitionsmenge und die Zielmenge angeben. Die Abbildungen

$$\begin{aligned} \mathbb{N} \to \mathbb{N},&\quad x \mapsto x^2 + 1,\\ \mathbb{Z} \to \mathbb{Z},&\quad x \mapsto x^2 + 1,\\ \mathbb{Z} \to \mathbb{N},&\quad x \mapsto x^2 + 1 \end{aligned}$$

sind nach unserer Definition nicht gleich.

Die Teilmenge

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

von Y heißt *Bild* der Abbildung $f : X \to Y$.

Wir führen nun drei sehr wichtige Begriffe ein:

Definition 0.19. Eine Abbildung $f : X \to Y$ heißt:

(i) *injektiv,* wenn f verschiedene Elemente aus X auf verschiedene Elemente in Y abbildet, d. h.

$$\forall x, x' \in X\ :\ (f(x) = f(x') \implies x = x')^{25},$$

Diese Definition besagt also, dass eine injektive Abbildung verschiedene Elemente des Definitionsbereichs auf verschiedene Elemente der Zielmenge abbilden muss.

(ii) *surjektiv,* wenn $f(X) = Y$, d. h.

$$\forall y \in Y\ \exists x \in X:\ f(x) = y^{26},$$

Diese Definition besagt also, dass eine surjektive Abbildung jedes Element der Zielmenge „trifft“.

(iii) *bijektiv,* wenn f injektiv und surjektiv ist[27]. Eine bijektive Abbildung wird bisweilen auch *Identifikation* (zwischen Mengen) genannt.
Eine bijektive Abbildung identifiziert also die Elemente des Definitionsbereichs eins zu eins mit den Elementen der Zielmenge.

[25] Anders ausgedückt hat Γ_f die Eigenschaft $(((x_1, y) \in \Gamma_f) \wedge ((x_2, y) \in \Gamma_f)) \implies (x_1 = x_2)$.

[26] Anders ausgedückt hat Γ_f die Eigenschaft $\forall y \in Y\ \exists x \in X : (x, y) \in \Gamma_f$.

[27] Um zu zeigen, dass eine Abbildung bijektiv ist, muss man also sowohl Injektivität wie auch Surjektivität nachweisen.

Injektivität und Surjektivität sind verschiedene Eigenschaften einer Abbildung und sollten nicht verwechselt werden. Die Surjektivität hängt stark von der Wahl der gewählten Zielmenge ab. Ist $f : X \to Y$ eine beliebige Abbildung, so ist die Abbildung $\hat{f} : X \to f(X),\ x \mapsto \hat{f}(x) = f(x)$ immer surjektiv. Bei gegebener Abbildungsvorschrift kann die Injektivität durchaus von der Definitionsmenge von f abhängen. Das folgende Beispiel von bereits aus der Schule bekannten Funktionen soll das verdeutlichen: Sei $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \ : \ x \geq 0\}$, dann gilt:

$$\begin{array}{lll}
\mathbb{R} \to \mathbb{R}, & x \mapsto x^2 & \text{ist weder injektiv noch surjektiv,} \\
\mathbb{R}_{\geq 0} \to \mathbb{R}, & x \mapsto x^2 & \text{ist injektiv aber nicht surjektiv,} \\
\mathbb{R} \to \mathbb{R}_{\geq 0}, & x \mapsto x^2 & \text{ist surjektiv aber nicht injektiv,} \\
\mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, & x \mapsto x^2 & \text{ist bijektiv.}
\end{array}$$

Aufgabe 0.20. Geben Sie *alle* injektiven Abbildungen $\{1, 2, 3\} \to \{1, 2, 3\}$ an und begründen Sie Ihre Antwort.

Aufgabe 0.21. Geben Sie Beispiele für

(a) Abbildungen $\mathbb{Z} \to \mathbb{Z}$, die injektiv aber nicht surjektiv ist,

(b) Abbildungen $\mathbb{N} \to \mathbb{N}$, die surjektiv aber nicht injektiv ist,

(c) Abbildungen $\mathbb{Z} \to \mathbb{N}$, die bijektiv ist,

und begründen Sie Ihre Antwort.

Ist $f : X \to Y$ eine Abbildung und $U \subset X$ eine nichtleere Teilmenge, dann heißt die Abbildung

$$f|_U : U \to Y,\ u \mapsto f|_U(u) = f(u)$$

Einschränkung von f auf U.

Wenn $f : X \to Y$ eine beliebige Abbildung ist, so definieren wir eine neue Abbildung

$$f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X),\ U \mapsto f^{-1}[U] := \{x \in X \ : \ f(x) \in U\}. \tag{0.1}$$

Die Menge $f^{-1}[U] \subset X$ heißt *Urbild* von $U \subset Y$ unter f. Nach Definition einer Abbildung $f : X \to Y$ gilt immer $f^{-1}[Y] = X$ und $f^{-1}[\emptyset] = \emptyset$.

Mit Hilfe des Urbilds können wir auch die Begriffe injektiv und surjektiv erklären: So ist $f : X \to Y$ genau dann injektiv, wenn $f^{-1}[\{y\}]$ höchstens ein Element enthält und f ist dann und nur dann surjektiv, wenn $f^{-1}[\{y\}] \neq \emptyset$ für alle $y \in Y$. Schließlich ist also f genau dann bijektiv, wenn $f^{-1}[\{y\}]$ für alle $y \in Y$ einelementig ist.

Aufgabe 0.22. Sei $f : X \to Y$ eine Abbildung und seien $U, V \subset Y$. Beweisen Sie:

(a) $f^{-1}[U \cup V] = f^{-1}[U] \cup f^{-1}[V]$,

(b) $f^{-1}[U \cap V] = f^{-1}[U] \cap f^{-1}[V]$.

Ist $f : X \to Y$ bijektiv, so kann man auch die *Umkehrabbildung*

$$f^{-1} : Y \to X,\ y \mapsto f^{-1}(y), \tag{0.2}$$

wobei $f^{-1}(y) = x$ genau dann gilt, wenn $y = f(x)$ ist, definieren[28]. Es gilt also für die Umkehrabbildung einer bijektiven Abbildung $y = f(f^{-1}(y))$ und $f^{-1}(f(x)) = x$.

Selbst wenn $\frac{1}{f}$ definiert sein sollte, so ist $\frac{1}{f}$ im Allgemeinen nicht die Umkehrabbildung von f.

Lemma 0.23. *Sei $f : X \to Y$ bijektiv, dann ist auch $f^{-1} : Y \to X$ bijektiv.*

Beweis. Sei $f^{-1}(y) = f^{-1}(y')$, dann folgt $y = f(f^{-1}(y)) = f(f^{-1}(y')) = y'$. Also ist f^{-1} injektiv. Wenn $x \in X$, dann gilt $x = f^{-1}(f(x))$. Somit ist f^{-1} auch surjektiv. □

Seien $g : X \to Y$ und $f : Y \to Z$ zwei Abbildungen, wobei die Definitionsmenge Y von f mit der Zielmenge von g übereinstimme, dann definieren wir die *Hintereinanderschaltung* $f \circ g$ von f und g als

$$f \circ g : X \to Z,\ x \mapsto f(g(x)).$$

Lemma 0.24. *Eine Abbildung $f : X \to Y$ ist genau dann bijektiv, wenn es eine Abbildung $g : Y \to X$ gibt, sodass $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$ ist. Es gilt dann $g = f^{-1}$.*

Beweis. Ist f bijektiv, so hat $g = f^{-1}$ die gewünschten Eigenschaften.

Nehmen wir nun an, es gibt eine Abbildung $g : Y \to X$ mit den angegeben Eigenschaften, dann ist f injektiv, denn aus $f(x) = f(x')$ folgt $x = g(f(x)) = g(f(x')) = x'$. Anderseits ist f surjektiv, denn für alle $y \in Y$ gilt $f(g(y)) = y$.

Wegen $f(g(y)) = y$ für alle $y \in Y$ folgt $f^{-1}(y) = g(y)$ für alle $y \in Y$ und somit $g = f^{-1}$. □

Aufgabe 0.25. Sei $f : X \to Y$ eine Abbildung und $U, V \subset X$. Zeigen Sie, dass gilt:

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

(b) $f(U \cap V) \subset f(U) \cap f(V)$, jedoch ist $f(U \cap V) = f(U) \cap f(V)$ im Allgemeinen falsch.

(c) $f(U \cup V) = f(U) \cup f(V)$.

(d) $f(U) \setminus f(V) \subset f(U \setminus V)$, jedoch ist $f(U) \setminus f(V) = f(U \setminus V)$ im Allgemeinen falsch.

Aufgabe 0.26. Seien $f : X \to Y$ und $g : Y \to Z$ zwei Abbildungen.

[28] An dieser Stelle ist ein Hinweis angebracht um Missverständnisse zu vermeiden. Die Notation f^{-1} haben wir nun für zwei völlig verschiedene Abbildungen verwendet, einmal für eine Urbild-Abbildung in Formel (0.1) welche für jede beliebige Abbildung definiert ist und einmal für die Umkehrabbildung in Formel (0.2), welche nur existiert, wenn f bijektiv ist. Da die Definitions- und Zielmengen beider Abbildungen aber unterschiedlich sind, sollte eine Verwechslung nicht auftreten. Für eine bijektive Abbildung f gilt $f^{-1}[\{y\}] = \{x\}$ genau dann, wenn $f^{-1}(y) = x$.

(a) Zeigen Sie:

$$\begin{array}{rcl} f \text{ injektiv und } g \text{ injektiv} & \Longrightarrow & g \circ f \text{ injektiv,} \\ f \text{ surjektiv und } g \text{ surjektiv} & \Longrightarrow & g \circ f \text{ surjektiv.} \end{array}$$

(b) Gelten auch die umgekehrten Implikationen? Beweisen Sie Ihre Antwort.

Aufgabe 0.27. Seien X und Y zwei nichtleere Mengen und $f : X \to Y$ eine Abbildung. Zeigen Sie:

(a) Durch $x \sim x'$, wenn $f(x) = f(x')$, $x, x' \in X$, ist auf X eine Äquivalenzrelation definiert.

(b) Die Abbildung

$$\hat{f} : X/\sim \to f(X) \subset Y, \quad [x] \mapsto f(x)$$

ist *wohldefiniert*[29] und bijektiv.

0.6. Mächtigkeit von Mengen

In diesem Abschnitt folgen wir [A1, Abschn. 1.4.4].

Definition 0.28. Zwei nichtleere Mengen X und Y heißen *gleichmächtig*, wenn es eine bijektive Abbildung $X \to Y$ gibt.
Der „gleichmächtig“ bedeutet anschaulich „gleich viele Elemente besitzen“. Wir möchten dieser Idee auch einen Sinn geben, wenn Mengen unendlich viele Elemente haben. Dazu benutzen wir eine bijektive Abbildung, die die Elemente von X eins zu eins mit den Elementen von Y identifiziert und umgekehrt. Wir werden später sehen, dass es Mengen mit unendlich vielen Elementen gibt, die aber nicht gleichmächtig sind. So gibt es z. B. „mehr“reelle Zahlen als rationale Zahlen (siehe Satz 0.31 und Satz 1.30).

Wir sagen X hat $n \in \mathbb{N}^*$ Elemente, notiert $|X| = n$, wenn X und $\{x \in \mathbb{N}^*;\ 1 \leq x \leq n\}$ gleichmächtig sind.

Die leere Menge hat 0 Elemente, notiert $|\emptyset| = 0$.

Definition 0.29. Eine Menge X heißt

- *endlich*, wenn es ein $n \in \mathbb{N}$ gibt, sodass $|X| = n$.
- *unendlich*, wenn X nicht endlich ist.
- *abzählbar unendlich* , wenn X und $\mathbb{N}$ gleichmächtig sind.
- *höchstens abzählbar* oder kurz *abzählbar*, wenn X endlich oder abzählbar ist.
- *überabzählbar*, wenn X nicht höchstens abzählbar ist.

[29] d. h. die Abbildungsvorschrift hängt nicht von der Wahl des Repräsentanten einer Äquivalenzklasse ab.

Sei X eine abzählbar unendliche Menge, dann heißt eine bijektive Abbildung $f : \mathbb{N} \to X$ *Abzählung* von X und X kann als

$$X = \{f(j) \mid j \in \mathbb{N}\}$$

geschrieben werden.
Der Begriff „abzählbar unendlich“ ist immer noch recht anschaulich. Obwohl die Menge unendlich viele Elemente hat, können wir an der Menge gewissermaßen entlanggehen und die Elemente eines nach dem anderen nummerieren. Das macht gerade die bijektive Abbildung $f : \mathbb{N} \to X$.

Aufgabe 0.30. Zeigen Sie, dass $\mathbb{Z}$ abzählbar unendlich ist (vgl. Aufgabe 0.21(c)).

Satz 0.31. *Folgende Mengen sind abzählbar unendlich:*

(i) $\mathbb{N} \times \mathbb{N}$,

(ii) $\mathbb{Q}$.

Beweis. Cantorsches erstes Diagonalargument, siehe [A1, Abschn. 1.4.4]. Ein schönes Bild des Cantorschen Abzählungsschemas für $\mathbb{Q}$ finden Sie in [K1, Abschn. 2.4]. □

Satz 0.32. *Sei* $\{X_j \mid j \in \mathbb{N}\}$ *ein abzählbares System von Mengen, wobei jede Menge* X_j *selbst abzählbar ist, d. h. für jedes* $j \in \mathbb{N}$ *gibt es eine bijektive Abbildung* $f_j : \mathbb{N} \to X_j$. *Dann ist die Vereinigung* $X = \bigcup_{j\in\mathbb{N}} X_j$ *abzählbar.*

Beweis. Cantorsches erstes Diagonalargument, siehe [A1, Abschn. 1.4.4]. □

Satz 0.33. *Sei* X *eine nichtleere Menge, so gibt es keine surjektive Abbildung* $f : X \to \mathcal{P}(X)$.

Beweis. Nehmen wir an, es gäbe eine surjektive Abbildung

$$f : X \to \mathcal{P}(X),\ x \mapsto f(x) \subset X.$$

Betrachten wir nun die Menge

$$Y = \{x \in X \mid x \notin f(x)\} \subset X.$$

Da f surjektiv ist, gibt es ein $y \in X$ mit $f(y) = Y$. Nun gibt es nur folgende zwei Möglichkeiten:

(1) Wenn $y \in f(y)$, dann gilt $y \notin Y$ und damit $y \in f(y) \setminus Y$, d. h. $f(y) \setminus Y \neq \emptyset$ und damit $f(y) \neq Y$, ein Widerspruch.

(2) Wenn $y \notin f(y)$, dann gilt $y \in Y$ und damit $y \in Y \setminus f(y)$ also $f(y) \neq Y$, wiederum ein Widerspruch.

□

Betrachten wir nun die Menge $X = \mathbb{N}$, so erhalten wir, zumal $\mathcal{P}(\mathbb{N})$ sicherlich nicht endlich ist, als Konsequenz:

Korollar 0.34. $\mathcal{P}(\mathbb{N})$ *ist überabzählbar.*

1 Zahlen

1.1. Körper

In diesem Abschnitt orientieren wir uns an [A1, Abschn. 1.5], [B1, Kap. I, § 2] und [K1, Kap. 2].

In der Schule haben Sie zunächst die natürlichen Zahlen $0, 1, 2, 3, \ldots$ kennen gelernt. Zur Subtraktion wurde es nötig auch die negativen ganzen Zahlen $-1, -2, -3, \ldots$ einzuführen. Die Division motivierte dann die rationalen Zahlen („Bruchzahlen“). Bereits in der Schule haben Sie gesehen, dass in der Natur Zahlen vorkommen, welche werder ganzzahlig sind, noch sich durch ein Verhältnis von ganzen Zahlen (Brüche) beschreiben lassen. Ein Beispiel ist die Länge d der Diagonale des Quadrats mit Seitenlänge 1. Nach dem Satz von Pythagoras gilt nämlich $d^2 = 2$ und die positive Lösung dieser Gleichung, nämlich $d = \sqrt{2}$ ist nicht rational. Ein weiteres Beispiel ist der Umfang π des Kreises mit Durchmesser 1. Auch die Eulersche Zahl $e = \exp(1)$ ist nicht rational. In der Schule wurde daher von reellen Zahlen gesprochen, ohne jedoch diesen Begriff genauer zu erläutern. Überhaupt stellt sich heraus, dass „fast alle“ reellen Zahlen nicht rational, also keine Brüche sind. Wir gehen zwar davon aus, dass die reellen Zahlen gegeben sind, geben jedoch ihre axiomatische Definition auch an, um uns auf sicheren Grund zu stellen.

Zunächst benötigen wir den algebraischen Begriff eines Körpers. Dabei handelt es sich um Mengen auf denen eine Addition und eine Multiplikation definiert ist, die den bereits in der Schule gelernten Rechengesetzen genügen.

Definition 1.1. Ein *Körper* ist ein Tripel $(\mathbb{K}, +, \cdot)$ bestehend aus einer Menge $\mathbb{K}$ und zwei Verknüpfungen (Abbildungen)

$$\begin{aligned} \mathbb{K} \times \mathbb{K} \to \mathbb{K}, \quad (k, k') \mapsto k + k' \\ \mathbb{K} \times \mathbb{K} \to \mathbb{K}, \quad (k, k') \mapsto k \cdot k' \end{aligned}$$

mit folgenden Eigenschaften:

(A) Addition[30]:

[30] Diese Eigenschaften besagen, dass $(\mathbb{K}, +)$ eine abelsche Gruppe ist.

- (A1) Assoziativität der Addition: $\forall k, k', k'' \in \mathbb{K}: \ (k + k') + k'' = k + (k' + k'')$.
- (A2) Existenz eines neutralen Elements der Addition: $\exists 0 \in \mathbb{K}\ \forall k \in \mathbb{K}: \ 0 + k = k$.
- (A3) Existenz eines Inversen Elements bzgl. der Addition: $\forall k \in \mathbb{K}\ \exists {-k} \in \mathbb{K}:$ $-k + k = 0$.
- (A4) Kommutativität der Addition: $\forall k, k' \in \mathbb{K}: \ k + k' = k' + k$.

(M) Multiplikation: Sei $\mathbb{K}^* := \mathbb{K} \setminus \{0\}$.

- (M1) Assoziativität der Multiplikation: $\forall k, k', k'' \in \mathbb{K}: \ (k \cdot k') \cdot k'' = k \cdot (k' \cdot k'')$.
- (M2) Kommutativität der Multiplikation: $\forall k, k' \in \mathbb{K}: \ k \cdot k' = k' \cdot k$.
- (M3) Existenz eine neutralen Elements der Multiplikation: $\exists 1 \in \mathbb{K}^*\ \forall k \in \mathbb{K}:$ $1 \cdot k = k$ (insbesondere ist $1 \neq 0$).
- (M4) Existenz eines Inversen der Multiplikation: $\forall k \in \mathbb{K}^*\ \exists k^{-1} \in \mathbb{K}: \ k^{-1} \cdot k = 1$ (Für $0 \in \mathbb{K}$ existiert also *kein* Inverses bzgl. der Multiplikation).

(D) Distributivgesetz: $\forall k, k', k'' \in \mathbb{K}: \ (k + k') \cdot k'' = (k \cdot k'') + (k' \cdot k'')$.

Beispiel 1.2. Sei Menge $\mathbb{Q}$ der Brüche mit der üblichen bereits in der Schule erlernten Addition und Multiplikation ist ein Körper

Folgende Rechenregeln in einem Körper folgen aus den Axiomen und werden üblicherweise in der Vorlesung Lineare Algebra 1 bewiesen. In der Vorlesung Algebra werden Körper dann noch genauer untersucht.

Lemma 1.3 (Rechenregeln in einem Körper)**.** *Sei* $(\mathbb{K}, +, \cdot)$ *ein Körper und* $k, k', k'' \in \mathbb{K}$. *Dann gilt:*

(a) Die Elemente 0 *und* 1 *sind eindeutig bestimmt.*

(b) $(k + k' = k + k'') \implies (k' = k'')$. *Insbesondere gilt* $(k + k' = k) \implies (k' = 0)$ *und* $(k + k' = 0) \implies (k' = -k)$.

(c) $-(-k) = k$.

(d) Wenn $k \neq 0$, *dann* $(k \cdot k' = k \cdot k'') \implies k' = k''$. *Insbesondere gilt für* $k \neq 0$ $(k \cdot k' = k) \implies (k' = 1)$, *sowie* $(k \cdot k' = 1) \implies k' = k^{-1}$.

(e) Für $k \neq 0$ *gilt* $(k^{-1})^{-1} = k$.

(f) Nullteilerfreiheit: $(k \cdot k' = 0) \iff ((k = 0) \vee (k' = 0))$.

(g) $(-k) \cdot k' = -(k \cdot k') = k \cdot (-k')$.

(h) $(-k) \cdot (-k') = k \cdot k'$.

Wir schreiben oft kk' anstelle von $k \cdot k'$. und $\frac{k}{k'}$ anstelle von $k \cdot (k'^{-1})$. Zudem sprechen wir einfach von einem Körper $\mathbb{K}$ und erwähnen die implizit gegebene Addition und Multiplikation nicht extra.

In einem Körper $\mathbb{K}$ definieren wir nun auch noch eine Subtraktion durch $k - k' := k + (-k')$ für alle $k, k' \in \mathbb{K}$, sowie eine Division durch $\frac{k'}{k} = k' \cdot (k^{-1})$. für alle $k \in \mathbb{K}^*$ und alle $k' \in \mathbb{K}$.

Zum Einsparen von Klammern verwenden wir die übliche Konvention „Inversion vor Punktrechnung vor Strichrechnung". Desweiteren sprechen wir künftig von einem Körper $\mathbb{K}$ und unterschlagen in der Notation einfach die zum Tripel gehörenden Abbildungen der Addition und der Multiplikation.

Fazit: In einem Körper hat man eine Addition und eine Multiplikation, die sich so verhalten, wie man es aus der Schule von Addition und Multiplikation kennt.

1.2. Geordnete Körper

In diesem Abschnitt orientieren wir uns vor allem an [A1, Abschn. 1.5], aber auch an [B1, Kap. I, § 2].

Wir möchten weiterhin die reellen Zahlen axiomatisch charakterisieren. Wie bereits aus der Schule bekannt, ist es für zwei reelle Zahlen sinnvoll zu sagen, dass die eine größer als die andere ist. Wir möchten dieses jetzt axiomatisch fassen:

Definition 1.4. Ein Körper $\mathbb{K}$ der gleichzeitig (bzgl. einer gegebenen Ordnungsrelation) eine geordnete Menge ist, heißt *geordneter Körper*, wenn die Ordnungsrelation und die Körperstruktur in folgendem Sinne verträglich sind:

(O1) *Monotonie der Addition:* Für alle $x, y, z \in \mathbb{K}$ gilt $(x \prec y) \implies (x + z \prec y + z)$.

(O2) *Monotonie der Multiplikation:* Für alle $x, y, z \in \mathbb{K}$ gilt $((x \prec y) \wedge (0 \prec z)) \implies ((x \cdot z) \prec (y \cdot z))$.

Ein Element x eines geordneten Körpers $\mathbb{K}$ heißt

- *positiv*, wenn $0 \prec x$,
- *negativ*, wenn $x \prec 0$,
- *nicht positiv*, wenn $x \preceq 0$,
- *nicht negativ*, wenn $0 \preceq x$.

Ein typisches Beispiel eines geordneten Körpers ist $\mathbb{Q}$ (mit der bereits bekannten Addition und Multiplikation) und der üblichen Ordnungsrelation $\leq$, d. h. bei gleichnamigen Brüchen mit natürlichem Nennen gilt $\frac{a}{c} \leq \frac{b}{c}$, $c \in \mathbb{N}^*$, wenn $a \leq b$ (hierbei haben wir die übliche Ordnungsrelation von $\mathbb{Z}$ genutzt).

Satz 1.5 (Rechenregeln für Ungleichungen)**.** *Für alle Elemente x, y, z eines geordneten Körpers $\mathbb{K}$ gilt:*

(i) $(x \prec y) \implies (0 \prec (y - x))$,

(ii) $((x \prec y) \wedge (z \prec 0)) \implies ((y \cdot z) \prec (x \cdot z))$,

(iii) $((x \prec 0) \vee (0 \prec x)) \implies (0 \prec x^2 := x \cdot x)$,[31]

(iv) $(0 \prec x \prec y) \implies 0 \prec y^{-1} \prec x^{-1}$.

Um Klammern zu sparen verwenden wir die Konvention „Rechenoperation vor Ordnungsrelation“.

Beweis. ad (i): Setze $z := -x$, dann folgt mit $x \prec y$ aus (O1) sofort $0 = x - x \prec y - x$.

ad (ii): Mit (O1) erhalten wir $0 \prec y - x$ und $0 = z - z \prec -z$. Wegen (O2) folgt $0 = 0 \cdot (-z) \prec (y-x) \cdot (-z)$ und damit wegen (O1) $yz = yz + 0 \prec yz + (y-x) \cdot (-z) = yz - yz + xz = xz$, also die Aussage.

ad (iii): Wenn $0 \prec x$, so erhalten wir aus (O2) sofort $0 \prec x^2$. Wenn $x \prec 0$, dann gilt mit (i) $0 \prec -x$ und damit nach (O2) $0 \prec (-x)^2 = x^2$.

ad (iv): Nach (iii) gilt $0 \prec 1^2 = 1 = y \cdot \frac{1}{y}$. Wenn $\frac{1}{y} \prec 0$, dann folgt nach (O1) $0 \prec -\frac{1}{y}$ und damit nach (O2) $0 \prec -\frac{1}{y} \cdot y = -1$, ein Widerspruch. Wenn $\frac{1}{y} = 0$, so auch $1 = y \cdot \frac{1}{y} = 0$, wieder ein Widerspruch. Damit muss also $0 \prec \frac{1}{y}$ gelten. Mit der gleichen Argumentation erhalten wir auch $0 \prec \frac{1}{x}$ und damit $0 \prec \frac{1}{x} \cdot \frac{1}{y}$. Aus $x \prec y$ folgt mit (O2) sofort $\frac{1}{y} \prec \frac{1}{x}$.

$\square$

Wenn Sie den Körper $\mathbb{Q}$ mit der ihnen vertrauten Ordnungrelation „kleiner gleich“ betrachten, so sind Ihnen die Rechenregeln aus Satz 1.5 plausibel.

Aufgabe 1.6. Auf der Menge $\mathbb{Z}_2 := \{0, 1\}$ definieren wir eine Addition durch $0 + 1 = 1 + 0 = 1$ und $0 + 0 = 1 + 1 = 0$ sowie eine Multiplikation durch $0 \cdot 1 = 1 \cdot 0 = 0 \cdot 0 = 0$ und $1 \cdot 1 = 1$.

(a) Zeigen Sie, dass $\mathbb{Z}_2$ bzgl. dieser Verknüpfungen ein Körper ist.

(b) Zeigen Sie, dass es keine Ordnungsrelation auf $\mathbb{Z}_2$ gibt, die $\mathbb{Z}_2$ zu einem geordneten Körper macht.

(c) Zeigen Sie allgemein, dass ein Körper mit endlich vielen Elementen kein geordneter Körper sein kann.

Sei $\mathbb{K}$ ein geordneter Körper. Für zwei Teilmengen $U, V \subset \mathbb{K}$ definieren wir

$$U + V := \{u + v \mid u \in U,\ v \in V\} \quad \text{und} \quad U \cdot V := \{u \cdot v \mid u \in U,\ v \in V\}.$$

[31] Damit erhalten wir auch $0 \prec 1$. Denn aus $1 \prec 0$ folgt $0 \prec 1^2 = 1$ und damit ein Widerspruch.

Sei weiterhin $x \in \mathbb{K}$, so definieren wir

$$\begin{aligned} x \preceq U &:\iff \forall u \in U : \ x \preceq u, \\ U \preceq x &:\iff \forall u \in U : \ u \preceq x. \end{aligned}$$

Desweiteren definieren wir für eine Teilmenge $U \subset \mathbb{K}$ das *Maximum*, notiert $\max(U)$, und das *Minimum*, notiert $\min(U)$, von U durch

$$x = \max(U) :\iff (x \in U \text{ und } U \preceq x)$$

bzw.

$$x = \min(U) :\iff (x \in U \text{ und } x \preceq U).$$

Beachten Sie, dass $\max(U)$ bzw. $\min(U)$ nicht immer existieren müssen.

Aufgabe 1.7. Sei U eine Teilmenge eines geordneten Körpers $\mathbb{K}$.

(a) Beweisen Sie: $\max(U)$ und $\min(U)$ sind, wenn sie existieren, eindeutig bestimmt.

(b) Beweisen Sie: Ist U endlich, dann existieren $\max(U)$ und $\min(U)$.

(c) Seien $U, V \subset \mathbb{K}$, sodass $\max(U)$, $\max(V)$, $\min(U)$, $\min(V)$ existieren. Beweisen Sie:

   (i) $U \subset V \implies \max(U) \preceq \max(V)$,

   (ii) $\max(U + V) = \max(U) + \max(V)$,

   (iii) Wenn $0 \preceq U$ und $0 \preceq V$, dann $\max(U \cdot V) = \max(U) \cdot \max(V)$,

   (iv) $\max(-U)$ existiert und es gilt $\min(U) = -\max(-U)$,

   (v) $\max(U \cup V)$ existiert und es gilt $\max(U \cup V) = \max(\{\max(U), \max(V)\})$,

   (vi) $\min(U \cup V)$ existiert und es gilt $\min(U \cup V) = \min(\{\min(U), \min(V)\})$.

Definition 1.8. Sei $\mathbb{K}$ ein geordneter Körper. Für ein Element $x \in \mathbb{K}$ definieren wir den *Betrag* von x, notiert $|x|$, durch

$$|x| := \max\{x, -x\}.$$

Aufgabe 1.9. Beweisen Sie, dass für alle Elemente x, y eines geordneten Körpers $\mathbb{K}$ gilt:

(a) $0 \preceq |x|$, sowie $|x| = 0 \iff x = 0$.

(b) $|x \cdot y| = |x| \cdot |y|$ und damit auch $|-x| = |x|$.

(c) Die *Dreiecksungleichung* $|x+y| \preceq |x| + |y|$ und damit auch $||x| - |y|| \preceq |x - y|$ sowie $||x| - |y|| \preceq |x + y|$.

   Zusammenfassend gilt also

$$||x| - |y|| \preceq |x + y| \preceq |x| + |y| \tag{1.1}$$

Die Dreiecksungleichung (1.1) wird uns in dieser Vorlesung immer wieder als wichtiges Hilfsmittel bei Abschätzungen begegnen.

1.3. Vollständigkeit und reelle Zahlen

In diesem Abschnitt orientieren wir uns vor allem an [A1, Abschn. 1.5], aber auch an [B1, Kap. I, § 2]. Wir führen nun eine Eigenschaft eines geordneten Körpers ein, der die reellen Zahlen charakterisiert (siehe Satz 1.17) und sie vom geordneten Körper $\mathbb{Q}$ unterscheidet.

Definition 1.10. Sei $\mathbb{K}$ ein geordneter Körper. Eine Teilmenge $V \subset \mathbb{K}$ heißt

- *nach oben beschränkt*, wenn gilt: $\exists o \in \mathbb{K} : V \preceq o$. Ein solches o heißt dann *obere Schranke* von V.
- *nach unten beschränkt*, wenn gilt: $\exists u \in \mathbb{K} : u \preceq V$. Ein solches u heißt dann *untere Schranke* von V.
- *beschränkt*, wenn sie sowohl nach oben, wie auch nach unten beschränkt ist.

Obere bzw. untere Schranken einer nach oben bzw. nach unten beschränkten Menge sind nicht eindeutig bestimmt.

Definition 1.11. (i) Sei V eine nach oben beschränkte Teilmenge eines geordneten Körpers $\mathbb{K}$. Eine obere Schranke ω von V heißt *Supremum* (oder kleinste obere Schranke) von V, notiert $\omega = \sup(V)$, wenn gilt:

$$\forall \varepsilon \in \mathbb{K}_{\succ 0}\ \exists v \in V : \omega - \varepsilon \prec v.$$

Hierbei bezeichnet $\mathbb{K}_{\succ 0} := \{x \in \mathbb{K} : 0 \prec x\}$ die Menge aller positiven Elemente in $\mathbb{K}$.

(ii) Sei V eine nach unten beschränkte Teilmenge eines geordneten Körpers $\mathbb{K}$. Eine untere Schranke α von V heißt *Infimum* (oder größte untere Schranke) von V, notiert $\alpha = \inf(V)$, wenn gilt:

$$\forall \varepsilon \in \mathbb{K}_{\succ 0}\ \exists v \in V : v \prec \alpha + \varepsilon.$$

Aufgabe 1.12. Sei V eine Teilmenge eines geordneten Körpers $\mathbb{K}$. Zeigen Sie, dass gilt:

(a) Wenn $\max(V)$ existiert, dann gilt $\max(V) = \sup(V)$.

(b) Wenn $\min(V)$ existiert, dann gilt $\min(V) = \inf(V)$.

(c) Wenn $\sup(V)$ existiert, so gilt $\sup(V) = \min(\{x \in \mathbb{K} :\ V \preceq x\})$.

(d) Wenn $\inf(V)$ existiert, so gilt $\inf(V) = \max(\{x \in \mathbb{K} :\ x \preceq V\})$.

Satz 1.13 (Eindeutigkeit des Supremums und des Infimums)**.** *Sei V eine Teilmenge eines geordneten Körpers $\mathbb{K}$. Dann hat V höchstens ein Supremum und höchstens ein Infimum*[32].

[32] Jedoch muss V weder ein Infimum noch ein Supremum besitzen

Beweis. Nehmen wir an V hätte zwei verschiedene Suprema, nämlich ω_1 und ω_2, wobei ohne Einschränkung $\omega_1 \prec \omega_2$ gilt. Dann ist $0 \prec \varepsilon := \omega_2 - \omega_1$ und $\omega_1 = \omega_2 - \varepsilon$. Da ω_2 ein Supremum von V ist, gibt es ein $v \in V$ mit $\omega_1 = \omega_2 - \varepsilon \prec v$. Damit ist aber ω_1 keine obere Schranke von V und insbesondere auch kein Supremum. Widerspruch.

Auf ähnliche Weise zeigt man die Eindeutigkeit des Infimums (Übung). □

Definition 1.14. Ein geordneter Körper $\mathbb{K}$ heißt *(ordnungs)vollständig*, wenn er folgende *Supremumseigenschaft* besitzt:

Jede nichtleere, nach oben beschränkte Teilmenge von $\mathbb{K}$ besitzt ein Supremum.

Beispiel 1.15. Ein typisches Beispiel für einen geordneten Körper, der *nicht* ordnungsvollständig ist, ist $\mathbb{Q}$. Wie wir wissen ist $\sqrt{2}$ keine rationale Zahl und die Teilmenge $\{q \in \mathbb{Q} : \ q^2 \leq 2\} \subset \mathbb{Q}$ ist sehr wohl beschränkt, z. B. durch die obere Schranke 2, besitzt jedoch als Teilmenge von $\mathbb{Q}$ kein Supremum.

Satz 1.16. *Ein geordneter Körper $\mathbb{K}$, der die Supremumseigenschaft hat, besitzt auch die folgende* Infimumseigenschaft*:*

Jede nichtleere, nach unten beschränkte Teilmenge von $\mathbb{K}$ besitzt ein Infimum.

Beweis. Sei $V \subset \mathbb{K}$ nichtleer und nach unten beschränkt, sowie $\rho \in \mathbb{K}$ eine untere Schranke von V, also $\rho \preceq V$ Dann ist die Menge $-V := \{x \in \mathbb{K} : \ -x \in V\}$ durch $-\rho$ nach oben beschränkt, da $-V \preceq -\rho$ und besitzt nach Voraussetzung ein Supremum $\omega \in \mathbb{K}$. Damit ist $-\omega$ eine untere Schranke von V. Um zu sehen, dass $-\omega$ das Infimum von V ist, wählen wir beliebig ein $\varepsilon \in \mathbb{K}$ mit $0 \prec \varepsilon$. Da ω ein Supremum von $-V$ ist, gibt es ein Element $v \in V$ mit $\omega - \varepsilon \prec -v$ und damit nach Satz 1.5 $v \prec (-\omega) + \varepsilon$. Damit ist $-\omega$ ein Infimum von V. □

Satz 1.17. *Bis auf Isomorphie gibt es nur einen ordnungsvollständigen geordneten Körper*[33].

Wir werden diesen Satz nicht beweisen und verweisen zur Konstruktion z. B. durch Äquivalenzklassen rationaler Cauchyfolgen[34] oder durch Dedekindsche[35] Schnitte eines ordnungsvollständigen geordneten Körpers, wie auch zum Beweis von dessen Eindeutigkeit auf [Ebb, Kap. 2] von K. Mainzer.

Definition 1.18. Als $\mathbb{R}$ bezeichnen wir den (bis auf Isomorphie) eindeutig bestimmten ordnungsvollständigen geordneten Körper. Wie nennen ihn *Körper der reellen Zahlen.* Elemente von $\mathbb{R}$ heißen *reelle Zahlen*.

Wir stellen uns die Menge der reellen Zahlen als eine (nach beiden Seiten unendliche) Gerade vor (Zahlenstrahl, reelle Achse). In dieser Vorstellung gilt $x \prec y$, wenn x auf dem Zahlenstrahl links von y liegt. Wir schreiben benutzen fortan die bereits aus der Schule bekannte Notation für die Ordnungsrelation auf $\mathbb{R}$:

$$x < y \quad \text{oder gleichbedeutend} \quad y > x.$$

[33] d. h. seien $\mathbb{K}_1$ und $\mathbb{K}_2$ zwei ordnungsvollständige geordnete Körper, so gibt es eine bijektive Abbildung $f : \mathbb{K}_1 \to \mathbb{K}_2$, sodass sowohl f wie auch die Umkehrabbildung f^{-1} mit der Addition, der Multiplikation und der Ordnungsrelation verträglich sind.

[34] Augustin Louis Cauchy (1789 - 1857), französischer Mathematiker

[35] Julius Wilhelm Richard Dedekind (1831 - 1916), deutscher Mathematiker

Natürliche, ganze und rationale Zahlen

Wir können nun die Menge $\mathbb{N}$ der natürlich Zahlen als Teilmenge von $\mathbb{R}$ auch axiomatisch definieren:

Definition 1.19. Die Menge $\mathbb{N}$ der natürlichen Zahlen ist die kleinste[36] Teilmenge von $\mathbb{R}$, mit den folgenden Eigenschaften

(N1) $0 \in \mathbb{N}$,

(N2) Aus $n \in \mathbb{N}$ folgt $n+1 \in \mathbb{N}$.

Diese Definition der natürlichen Zahlen als Teilmenge der reellen Zahlen führt uns zum *Induktionsprinzip:*

Gilt für eine Teilmenge $V \subset \mathbb{R}$, welche 0 enthält, dass $v \in V$ immer $v+1 \in V$ nach sich zieht, dann ist $\mathbb{N} \subseteq V$.

Wir erhalten $\mathbb{N}$ als Durchschnitt aller Teilmengen $V \subset \mathbb{R}$, welche 0 enthalten und für die aus $v \in V$ immer $v+1 \in V$ folgt. Setzen wir $2 := 1+1,\ 3 := 1+1+1,\ 4 := 1+1+1+1,$ etc., so überlegt man sich, dass

$$\mathbb{N} = \{0, 1, 2, ...\}.$$

Aus dieser Definition von $\mathbb{N}$ gewinnen wir eine Beweistechnik, das *Prinzip der vollständigen Induktion*:

Möchte man nun für eine Aussageform A über der Grundmenge $\mathbb{N}$ beweisen, dass $A(n)$ für ein beliebiges $n \in \mathbb{N}$ wahr ist, so kann man etwa wie folgt vorgeben:

(1) *Induktionsanfang:* Man beweist zunächst, dass $A(0)$ wahr ist.

(2) *Induktionsschluß:*Für beliebig aber fest gewähltes $n \in N$ zeigt man die Aussage. $(A(n) \implies A(n+1))$.[37]

Zudem können wir Abbildung $f : \mathbb{N} \to X$ *rekursiv* definieren indem man z. B. wie folgt vorgeht:

(1) Definiere $f(0)$ (oder $f(0), f(1), \ldots, f(k)$).

(2) Gebe eine Vorschrift V an, mit welcher man $f(n+1)$ aus $f(0),\ f(1), \ldots, f(n)$ bestimmen kann:
$$f(n+1) = V(f(0), f(1), \ldots, f(n+1)).$$

---

[36] d. h. es gilt das Induktionsprinzip: Für jede Teilmenge $V \subset \mathbb{R}$, die die geforderten Eigenschaften hat, gilt $V \subseteq \mathbb{N}$.

[37] Man kann hier auch Varianten verwenden, z. B. kann man auch zeigen, wenn für alle Vorgänger k von $n+1$ die Aussage $A(k)$ wahr ist, dann ist auch $A(n+1)$ wahr.

Beispiel 1.20. Die *Fibonacci*[38] *Zahl* $f(n)$, $n \in \mathbb{N}$, ist rekursiv definiert über

$$f(0) = 0,\ f(1) = 1 \text{ und } f(n+2) = f(n) + f(n+1).$$

Nun aber fahren wir mit der Untersuchung von $\mathbb{R}$ fort. Mit $\mathbb{N}$ ist auch $\mathbb{Z}$ eine Teilmenge von $\mathbb{R}$, nämlich

$$\mathbb{Z} = \{x \in \mathbb{R} :\ x \in \mathbb{N} \text{ oder } -x \in \mathbb{N}\} \subset \mathbb{R}.$$

Damit können wir auch $\mathbb{Q}$ als Teilmenge von $\mathbb{R}$ betrachten durch

$$\mathbb{Q} := \{x \cdot y^{-1} :\ x, y \in \mathbb{Z},\ y \neq 0\} \subset \mathbb{R}.$$

Satz 1.21 (Prinzip von Archimedes[39])**.** *Seien $x, y \in \mathbb{R}$ mit $y > 0$, dann gibt es $n \in \mathbb{N}^*$, sodass $n \cdot y > x$.*

Beweis. Wäre diese Aussage falsch, so wäre x eine obere Schranke der Menge

$$M := \{n \cdot y :\ n \in \mathbb{N}^*\} \subset \mathbb{R}.$$

Damit hätte M auch ein Supremum. Setzen wir $\varepsilon = y > 0$, so gibt es ein Element $n_0 \in \mathbb{N}^*$ sodass $\sup(M) - y < n_0 \cdot y$ und damit $\sup(M) < n_0 \cdot y + y = (n_0 + 1)y$. Da aber $(n_0 + 1)y \in M$ ist $\sup(M)$ keine obere Schranke von M, Widerspruch. □

Korollar 1.22. *Sei $y \in \mathbb{R}$ mit $y > 0$, dann gibt es ein $n \in \mathbb{N}^*$, sodass $\frac{1}{n} < y$.*

Beweis. Nach dem Prinzip von Archimedes gibt es ein $n \in \mathbb{N}^*$, sodass $1 < ny$, also $\frac{1}{n} < y$. □

Satz 1.23. *Jede nichtleere Teilmenge von $\mathbb{N}$ hat ein Minimum.*

Beweis. Sei nun $A \subset \mathbb{N}$ eine nichtleere Teilmenge, welche *kein* kleinstes Element enthält, dann ist für alle $n \in \mathbb{N}$ die Menge $A \cap \{0, 1, \dots, n\}$ leer.

In der Tat stimmt das für $n = 0$, da sonst $\min(A) = 0$. Wenn nun aber $A \cap \{0, 1, \dots, n\}$ leer ist, so ist auch $A \cap \{0, 1, \dots, n+1\}$ leer, da sonst $\min(A) = n + 1$.

Nach der Definition von $\mathbb{N}$ ist aber damit A bereits leer, ein Widerspruch. □

Korollar 1.24. *Jede nach oben (unten) beschränkte nichtleere Teilmenge von $\mathbb{Z}$ enthält ein größtes (kleinstes) Element.*

Beweis. Übungsaufgabe. □

Satz 1.25 (Dichtheit von $\mathbb{Q}$ in $\mathbb{R}$)**.** *Seien $x, y \in \mathbb{R}$ mit $x < y$. Dann gibt es $q \in \mathbb{Q}$ mit $x < q < y$.*

[38] Leonardo Fibonacci oder Leonardo da Pisa, italienischer Gelehrter, 12. und 13. Jahrhundert
[39] Archimedes von Syrakus (3. Jahrh. v. Chr.) antiker griechischer Gelehrter

Beweis. Wähle $n \in \mathbb{N}^*$, sodass $\frac{1}{n} < y - x$ und betrachte die Menge $A = \{k \in \mathbb{Z} : \ k > nx\} \subset \mathbb{Z}$. Nach dem Prinzip von Archimedes ist A nicht leer und nach Konstruktion nach unten beschränkt. Damit enthält A aber ein kleinstes Element $m \in \mathbb{Z}$ (Korollar 1.22) Aus $m > nx$ und $nx \leq m - 1$ ergibt sich nun

$$x < \frac{m}{n} = \frac{m-1}{n} + \frac{1}{n} < x + y - x = y.$$

Damit ist $q := \frac{m}{n} \in \mathbb{Q}$ die gesucht rationale Zahl. $\square$

Anschaulich bedeutet Satz 1.25, das wir mit unserem menschlichen Auge die Menge $\mathbb{Q}$ als Teilmenge der reellen Zahlengeraden nicht von $\mathbb{Q}$ unterscheiden können.

Satz 1.26 (Bernoullische[40] Ungleichung)**.** *Sei $x \in \mathbb{R}$ mit $x \geq -1$ und $n \in \mathbb{N}$, dann gilt*

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

Beweis. Durch vollständige Induktion:

- Induktionsanfang: $n = 0 : \ (1+x)^0 = 1 = 1 + 0x$.
- Induktionsschluss: Aus $(1+x)^n \geq 1 + nx$ folgt mit $1 + x \geq 0 :$

  $$(1+x)^{n+1} = (1+x)^n(1+x) \geq (1+nx)(1+x) = 1+(n+1)x+nx^2 \geq 1+(n+1)x.$$

$\square$

Korollar 1.27. *(i) Sei $b \in \mathbb{R}$ mit $b > 1$. Dann gilt: Für alle $K \in \mathbb{R}$ gibt es ein $n \in \mathbb{N}$, sodass $b^n > K$.*

(ii) Sei $0 < q < 1$. Dann gilt: Für alle $\varepsilon > 0$ gibt es ein $n \in \mathbb{N}$, sodass $q^n < \varepsilon$.

Beweis. Setze $b = 1 + x$ mit $x > 0$. Dann besagt die Bernoullische Ungleichung $b^m \geq 1 + mx$ für alle $m \in \mathbb{N}$. Nach dem Prinzip von Archimedes gibt es ein $n \in \mathbb{N}$, sodass $nx > K$ uns somit $b^n \geq 1 + nx > 1 + K > K$. Damit gilt (i).

Setzt man $b = q^{-1}$ und $K = \varepsilon^{-1}$ in (i) ein, so folgt (ii). $\square$

1.4. Intervallschachtelung

In diesem Abschnitt orientieren wir uns an [K1, Abschn. 2.3 f.] Wir lernen eine Technik kennen, die sich aus der Vollständigkeit von $\mathbb{R}$ ergibt, um reelle Zahlen mit gewissen Eigenschaften zu konstruieren, indem wir sie quasi „eingrenzen“.

Seien $x, y \in \mathbb{R}$ mit $x \leq y$, dann heißt

- $[x, y] := \{a \in \mathbb{R} : \ x \leq a \leq y\}$ *kompaktes Intervall*,

[40] Jakob I. Bernoulli (1655 - 1705), schweizerischer Mathematiker

- $(x, y) =]x, y[:= \{a \in \mathbb{R} : \ x < a < y\}$ *offenes beschränktes Intervall,*
- $[x, y) = [x, y[:= \{a \in \mathbb{R} : \ x \leq a < y\}$ (nach rechts) *halboffenes Intervall,*
- $(x, y] =]x, y] := \{a \in \mathbb{R} : \ x < a \leq y\}$ (nach links) *halboffenes Intervall.*

Für ein beschränktes Intervall $I \in \{[x, y], (x, y), [x, y), (x, y]\}$ heißt $|I| := y - x$ die *Länge* von I. Neben den oben stehenden beschränkten Intervallen gibt es noch die unbeschränkten Intervalle:

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

Auch $\mathbb{R}$ selbst fassen wir bei Bedarf als Intervall $]-\infty, \infty[$ auf.

Definition 1.28. Eine reelle *Intervallschachtelung* ist eine Familie $(I_n) := \{I_n : \ n \in \mathbb{N}\}$ kompakter Intervalle in $\mathbb{R}$, sodass gilt

(I1) $I_{n+1} \subset I_n$ für alle $n \in \mathbb{N}$,

(I2) $\forall \varepsilon > 0 \ \exists n \in \mathbb{N} : \ |I_n| < \varepsilon$.

Bei einer Intervallschachtelung liegen also nach (I1) die Intervalle ineinander, wie das Wort „Schachtelung“ bereits suggeriert. Die Intervalllängen werden also mit wachsendem n immer kleiner. In (I2) verwenden wir die Formulierung „$\forall \varepsilon > 0 \ \ldots \ \ldots \ < \varepsilon$“. Solche Formulierungen werden uns noch häufig begegnen, z. B. bei Begriff der Konvergenz oder der Stetigkeit. Sie sind wie folgt zu verstehen: „Für *beliebig kleine* positive $\varepsilon \ldots$“. Hier in $(I2)$ heißt dass, dass die Intervalllängen mit wachsendem n beliebig klein, also kleiner als jede vorgegebene Schranke ε werden.

Satz 1.29 (Prinzip der Intervallschachtelung)**.** *Sei (I_n) eine reelle Intervallschachtelung, dann gibt es genau ein Element $a \in \mathbb{R}$, sodass für alle $n \in \mathbb{N}$ gilt $a \in I_n$, oder kurz* $\bigcap\limits_{n \in \mathbb{N}} I_n = \{a\}$.[41]

Beweis. Sei $I_n = [x_n, y_n]$, dann ist die Menge $A := \{x_n : \ n \in \mathbb{N}\}$ nach (I1) durch alle y_n nach oben beschränkt und hat somit ein Supremum. Für dieses Supremum gilt aber damit $x_n \leq \sup(A) \leq y_n$ für alle $n \in \mathbb{N}$, also $a := \sup(A) \in I_n$ für alle $n \in \mathbb{N}$.

Wenn nun zwei verschiedene reelle Zahlen $a_1 < a_2$ in allen I_n liegen, so gilt $[a_1, a_2] \subset I_n$ für alle $n \in \mathbb{N}$ und damit $|I_n| \geq (a_2 - a_1)$ im Widerspruch zu (I_2). $\square$

[41] Die Gültigkeit dieser Aussage ist äquivalent zur Ordnungsvollständigkeit von $\mathbb{R}$ (siehe [K1, Abschn. 2.3]).

Das Prinzip der Intervallschachtelung wird uns erlauben reelle Zahlen mit gewissen Eigenschaften zu „konstruieren“ indem wir sie (ihre Lage auf dem reellen Zahlenstrahl) durch eine Intervallschachtelung immer weiter eingrenzen, bis am Ende nur noch eine Zahl übrigbleibt.

Satz 1.30. *Die Menge $\mathbb{R}$ ist überabzählbar.*

Beweis. Zunächst ist $\mathbb{R}$ nicht endlich, da $\mathbb{N} \subset \mathbb{R}$. Angenommen $\mathbb{R}$ wäre abzählbar unendlich, dann gibt es eine bijektive Abbildung $f : \mathbb{N} \to \mathbb{R}$.

Wir konstruieren nun eine reelle Intervallschachtelung (I_n), sodass für alle $n \in \mathbb{N}$ gilt $f(n) \notin I_n$ rekursiv:

Sei $I_0 := [f(0) + 1, f(0) + 2] = [x_0, y_0]$. Sei nun $I_n = [x_n, y_n]$ bereits konstruiert. Wir betrachten $I_{n,1} := \left[x_n, x_n + \frac{y_n - x_n}{3}\right]$, $I_{n,2} := \left[x_n + \frac{y_n - x_n}{3},\ x_n + \frac{2(y_n - x_n)}{3}\right]$ und $I_{n,3} : \left[x_n + \frac{2(y_n - x_n)}{3},\ y_n\right]$. Dann gilt $I_n = I_{n,1} \cup I_{n,2} \cup I_{n,3}$ und wir können ein $j_0 \in 1, 2, 3$ wählen, sodass $f(n+1) \notin I_{n,j_0}$. Wir setzen $I_{n+1} := I_{n,j_0}$. Man überlege sich, weshalb wir so eine reelle Intervallschachtelung konstruiert haben (Übung).

Nach Satz 1.29 gibt es ein Element $y \in \mathbb{R}$ mit $y \in I_n$ für alle $n \in \mathbb{N}$. Damit gilt aber $y \notin \mathrm{Bild}(f)$. Somit ist f nicht surjektiv. Widerspruch. $\square$

Dieser Satz besagt, dass, obwohl $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt, es „sehr sehr viel“ mehr reelle Zahlen als Brüche gibt.

Satz 1.31 (Existenz von Wurzeln)**.** *Sei $x \in \mathbb{R}_{>0} = \{a \in \mathbb{R} : \ a > 0\}$ und $k \in \mathbb{N}^*$, dann gibt es genau ein $y \in \mathbb{R}_{>0}$ sodass $y^k = x$. Wir notieren $y = x^{\frac{1}{k}}$ bzw. $y = \sqrt[k]{x}$.*

Beweis. Nehmen wir an, es gäbe zwei verschiedene solche Zahlen, z. B. $y_1, y_2 \in \mathbb{R}_{>0}$. Wir nehmen ohne Einschränkung $y_1 < y_2$ an. Dann folgt aber $x = y_1^k < y_2^k = x$, ein Widerspruch.

Nun konstruieren wir y durch eine Intervallschachtelung. Hierbei genügt es $x > 1$ anzunehmen, denn wenn $0 < x < 1$, so betrachten wir $\tilde{x} := \frac{1}{x}$ anstelle von x. Für $x = 1$ erhalten wir sofort $y = 1$.

Sei $I_0 := [1, x]$. Ist $I_n = [a_n, b_n]$ bereits konstruiert, so betrachten wir den Mittelpunkt von $m := \frac{1}{2}(a_n + b_n)$ von I_n und setzten

$$I_{n+1} = [a_{n+1}, b_{n+1}] := \begin{cases} [a_n, m], & \text{wenn} \quad m^k \geq x \\ [m, b_n], & \text{wenn} \quad m^k < x \end{cases}$$

Es gilt somit

- $a_n^k \leq x \leq b_n^k$,
- $|I_{n+1}| = \frac{1}{2}|I_n|$, also $|I_n| = \frac{1}{2}^n |I_0| = \frac{1}{2}^n (x - 1)$.

Nach Konstruktion gilt $I_{n+1} \subset I_n$. Sei nun $\varepsilon > 0$, dann gibt es nach Korollar 1.27 eine Zahl $n \in \mathbb{N}$, sodass $\frac{1}{2}^n < \varepsilon \cdot |I_0|^{-1}$. gilt. Damit ist $|I_n| < \varepsilon$.

Nach Satz 1.29 gibt es genau eine Zahl $y \in \mathbb{R}$, welche in allen Intervallen I_n liegt.

Wir müssen nun nachweisen, dass $y^k = x$ gilt. Dazu betrachten wir die Intervalle $I_n^k := [a_n^k, b_n^k]$. Sie bilden eine Intervallschachtelung. In der Tat folgt aus $I_{n+1} \subset I_n$ auch $I_{n+1}^k \subset I_n^k$. Desweiteren gilt

$$|I_n^k| = (b_n^k - a_n^k) = (b_n - a_n)(b_n^{k-1}a_n^0 + b_n^{k-2}a_n + ... + b_n a_n^{k-2} + b_n^0 a_n^{k-1}) < |I_n| k b_1^{k-1}$$

Sei nun $\varepsilon > 0$. Da die I_n eine Intervallschachtelung bilden, gibt es ein $m \in \mathbb{N}$, sodass $|I_m| < \frac{\varepsilon}{kb_1^{k-1}}$. Damit gilt $|I_m^k| < \varepsilon$.

Nach Konstruktion gilt für alle $n \in \mathbb{N}$ einerseits $a_n^k \leq x \leq b_n^k$, also $x \in I_n^k$, und andererseits $y \in I_n$, also $y^k \in I_n^k$. Damit liegen sowohl x und y^k beide für alle $n \in \mathbb{N}$ in I_n. Aus Satz 1.29 folgt $x = y^k$. □

1.5. Komplexe Zahlen

In Satz 1.31 haben wir gesehen, dass man in $\mathbb{R}$ Quadratwurzeln aus positiven Zahlen ziehen kann. Aus negativen reellen Zahlen ist dieses in $\mathbb{R}$ jedoch nicht möglich, denn Satz 1.5 folgt $x^2 \geq 0$ für alle $x \in \mathbb{R}$.

Daher erweitern wir noch einmal den Zahlbegriff und definieren eine neue Zahl i mit der Eigenschaft

$$i^2 = -1.$$

Diese Zahl i darf mit einer Zahl $y \in \mathbb{R}$ zu $yi = iy$ multipliziert werden und wir verwenden die üblichen Rechenregeln. Daher erfüllt yi die Gleichung $(yi)^2 = y^2 i^2 = -y^2$. Insgesamt betrachten wir nun die Menge der *komplexen Zahlen*

$$\mathbb{C} := \{x + yi : \ x, y \in \mathbb{R}\}.$$

Der *Realteil* einer komplexen Zahl $z = x + yi$ ist die reelle Zahl $\Re(z) = x$ und der *Imaginärteil* von z ist die reelle Zahl $\Im(z) = y$.

Wir rechnen mit komplexen Zahlen in gewohnter Weise. Die unbekümmerte Rechnung

$$\begin{aligned}
(x + yi) + (x' + y'i) \quad &„=“ \quad (x + x') + (y + y')i \\
(x + yi)(x' + y'i) \quad &„=“ \quad x(x' + y'i) + yi(x' + y'i) \\
&„=“ \quad xx' + xy'i + yx'i + yy'i^2 \\
&„=“ \quad xx' + xy'i + yx'i - yy' \\
&„=“ \quad (xx' - yy') + (xy' + yx')i
\end{aligned}$$

motiviert die folgende Definition der Addition und der Multiplikation komplexer Zahlen:

$$\begin{aligned}
(x + yi) + (x' + y'i) &:= (x + x') + (y + y')i \\
(x + yi)(x' + y'i) &:= (xx' - yy') + (xy' + yx')i.
\end{aligned}$$

Man beachte, dass man $\mathbb{R}$ auf natürliche Weise in $\mathbb{C}$ einbetten kann,

$$\mathbb{R} \hookrightarrow \mathbb{C}, \quad x \mapsto (x + 0i),$$

die mit der Addition und der Multiplikation verträglich ist (Übungsaufgabe), d. h.

$$x + x' \mapsto (x + 0i) + (x' + 0i) \quad \text{und} \quad xx' \mapsto (x + 0i)(x' + 0i).$$

Daher fassen wir jede reelle Zahl auch als komplexe Zahl auf, so z. B. $0 = 0 + 0i$ oder $1 = 1 + 0i$. Dann gilt $0z = 0$ und $1z = z$ für alle $z \in \mathbb{C}$.

Wir können $\mathbb{C}$ auch mit $\mathbb{R}^2$ identifizieren:

$$\mathbb{C} \to \mathbb{R}^2, \quad (x + yi) \mapsto \begin{pmatrix} x \\ y \end{pmatrix}.$$

Die Menge der reellen Zahlen in $\mathbb{C}$ entspricht damit der x-Achse. Diese Identifikation ist auch mit der Addition und der Multiplikation mit reellen Zahlen verträglich (Übungsaufgabe), d. h.

$$(x + yi) + (x' + y'i) \mapsto \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x' \\ y' \end{pmatrix} \quad \text{und} \quad \lambda(x + yi) \mapsto \lambda \begin{pmatrix} x \\ y \end{pmatrix}, \ \lambda \in \mathbb{R}.$$

Der folgende Satz zeigt, dass beim Rechnen mit komplexen Zahlen die üblichen Rechenregeln gelten:

Satz 1.32. *Mit der angegeben Addition und Multiplikation ist $\mathbb{C}$ ein Körper.*

Beweis. Der Beweis der Körpereigenschaften ist Routine.

Am schwierigsten zu sehen, ist die Existenz eines multiplikativen Inversen: Sei $z = (x + yi) \neq 0$. Dann ist auch $x^2 + y^2 \neq 0$ und wir setzen $z' = \frac{x}{x^2+y^2} + \frac{-y}{x^2+y^2} i$. Man rechnet dann leicht $zz' = 1$ nach. Damit haben wir die Existenz eines multiplikativen Inversen gezeigt. $\square$

Für eine komplexe Zahl $z = x + yi$ definieren wir $\overline{z} := x - yi$, die zu z *konjugierte* komplexe Zahl. Die Abbildung $\mathbb{C} \to \mathbb{C}$, $z \mapsto \overline{z}$ heißt *komplexe Konjugation*. Identifizieren wir $\mathbb{C}$ mit $\mathbb{R}^2$ wie zuvor, so entspricht die komplexe Konjugation gerade der Spiegelung an der x-Achse.

Lemma 1.33. *Die komplexe Konjugation ist mit der Addition und der Multiplikation verträglich, d. h. es gilt für alle $z_1, z_2 \in \mathbb{C}$*

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} \quad \textit{und} \quad \overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}.$$

Beweis. Übungsaufgabe $\square$

Man rechnet leicht nach, dass für $z = x + yi$ gilt:

$$z\,\overline{z} = x^2 + y^2 \in \mathbb{R}.$$

Wir definieren den *Betrag* von z als die reelle Zahl

$$|z| := \sqrt{z\,\overline{z}} = \sqrt{x^2 + y^2}.$$

Dieser Betrag von $z = x + iy$ entspricht der Länge des Vektors $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, den Sie bereits aus der Schule kennen.

Aufgabe 1.34. Zeigen Sie, dass für $z \in \mathbb{C}$ gilt:

(a) $z + \overline{z} = 2\Re(z)$,

(b) $z - \overline{z} = 2i\Im(z)$,

(c) $(z = \overline{z}) \iff z \in \mathbb{R}$,

(d) $|z| = |\overline{z}|$,

(e) $|\Re(z)| \leq |z|$ sowie $|\Im(z)| \leq |z|$,

(f) Ist $z \neq 0$, so gilt $z^{-1} = \frac{1}{|z|^2}\overline{z}$.

Lemma 1.35. *Der Betrag hat folgende Eigenschaften*[42]*:*

(i) $|z| \geq 0$ *für alle* $z \in \mathbb{C}$ *und* $|z| = 0 \iff z = 0$.

(ii) $|z_1 z_2| = |z_1| \cdot |z_2|$ *für alle* $z_1, z_2 \in \mathbb{C}$.

(iii) Es gilt die Dreiecksungleichung

$$|z_1 + z_2| \leq |z_1| + |z_2|$$

für alle $z_1, z_2 \in \mathbb{C}$.

Beweis. Die Behauptung (i) ist einfach nachzuweisen.

Für (ii) betrachtet man

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

und zieht die Wurzel aus den nicht negativen Zahlen.

Aus $|\Re(z)| \leq |z|$ erhalten wir zunächst

$$\Re(z_1\overline{z_2}) \leq |z_1\overline{z_2}| = |z_1| \cdot |\overline{z_2}| = |z_1| \cdot |z_2|$$

und damit

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)(\overline{z_1} + \overline{z_2}) = z_1\overline{z_1} + z_1\overline{z_2} + z_2\overline{z_1} + z_2\overline{z_2} \\ &= |z_1|^2 + 2\Re(z_1\overline{z_2}) + |z_2|^2 \\ &\leq |z_1|^2 + 2|z_1| \cdot |z_2| + |z_2|^2 = (|z_1| + |z_2|)^2. \end{aligned}$$

Daraus folgt die Dreiecksungleichung. □

[42] Ein Körper $\mathbb{K}$ auf dem eine Funktion $\mathbb{K} \to \mathbb{R}$, $x \to |x|$ definiert ist, welche diese Eigenschaften hat heißt *bewerteter Körper*.

Wir haben komplexe Zahlen eingeführt, um Wurzeln aus beliebigen reellen Zahlen ziehen zu können. Wurzeln kann man aber ebenso aus beliebigen komplexen Zahlen ziehen und jede quadratische Gleichung

$$z^2 + a = 0,\ a \in \mathbb{C}$$

hat eine komplexe Lösung (Übungsaufgabe). Dasselbe gilt auch für Gleichungen höherer Ordnung, z. B. kubische Gleichungen. Allgemein gilt:

Theorem 1.36 (Fundamentalsatz der Algebra)**.** *Sei $n \in \mathbb{N}^*$ dann hat jede Gleichung der Form*

$$a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0 = 0 \quad \textit{mit} \quad a_0, \ldots, a_n \in \mathbb{C},\ a_n \neq 0$$

eine komplexe Lösung.

Wir haben nun einen wichtigen algebraischen Vorteil von $\mathbb{C}$ gegenüber $\mathbb{R}$ angesprochen. Dieser ist jedoch mit einem Nachteil erkauft. Im Gegensatz zu reellen Zahlen gilt:

Satz 1.37. *Es gibt keine Ordnungsrelation auf $\mathbb{C}$, die $\mathbb{C}$ zu einem geordneten Körper macht.*

Beweis. Angenommen $\mathbb{C}$ sei mit $\preceq$ ein geordneter Körper. Dann gilt nach Satz 1.5 $0 \prec 1^2 = 1$ und damit $-1 = 0 + (-1) \prec 1 + (-1) = 0$. Nun erhalten wir wieder mit Satz 1.5, da entweder $0 \prec i$ oder $i \prec 0$ gilt, den Widerspruch $0 \prec i^2 = -1$. □

Ungleichungen zwischen komplexen Zahlen, die nicht reell sind, haben also keinen Sinn. So ist die Aussage $|z_1| \leq |z_2|$ für beliebige $z_1, z_2 \in \mathbb{C}$ erklärt, *nicht* jedoch $z_1 \leq z_2$.

2 Folgen, Reihen, Konvergenz

2.1. Summenzeichen und binomischer Lehrsatz

In diesem Abschnitt stellen wir einige Notationen und Aussagen bereit, die wir später in diesem Kapitel verwenden werden. Wir orientieren uns an [Fo1, § 1] und an [K1, Kap. 1].

Seien k und n ganze Zahlen mit $k \leq n$ und $a_k, a_{k+1}, \ldots, a_n$ komplexe Zahlen. Dann notieren wir die Summe von $a_k, a_{k+1}, \ldots, a_n$ durch

$$\sum_{j=k}^{n} a_j := a_k + a_{k+1} + \cdots + a_n$$

und das Produkt von $a_k, a_{k+1}, \ldots, a_n$ durch

$$\prod_{j=k}^{n} a_j := a_k \cdot a_{k+1} \cdot \cdots \cdot a_n.$$

Manchmal ist es nützlich für die leere Summe und das leere Produkt festzulegen:

$$\sum_{j=k}^{k-1} a_j = 0 \quad \text{und} \quad \prod_{j=k}^{k-1} a_j = 1.$$

Für eine ganze Zahl m mit $k-1 \leq m \leq n$ gilt dann

$$\sum_{j=k}^{m} a_j + \sum_{j=m+1}^{n} a_j = \sum_{j=k}^{n} a_j \quad \text{und} \quad \prod_{j=k}^{m} a_j \cdot \prod_{j=m+1}^{n} a_j = \prod_{j=k}^{n} a_j.$$

Beispiel 2.1 (Geometrische Summenformel)**.** Für alle $z \in \mathbb{C} \setminus \{1\}$ und $n \in \mathbb{N}^*$ gilt

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

Hierbei gilt $z^0 := 1$.

Beweis. Vollständige Induktion:

- Für $n=1$ gilt $\frac{1-z^2}{1-z} = \frac{(1-z)(1+z)}{1-z} = 1+z$.
- Induktionschluß: $\sum\limits_{j=0}^{n+1} z^j = \sum\limits_{j=0}^{n} z^j + z^{n+1} = \frac{1-z^{n+1}}{1-z} + z^{n+1} = \frac{1-z^{n+2}}{1-z}$.

□

Für eine natürliche Zahl $n \in \mathbb{N}$ definieren wir $n!$, genannt *n-Fakultät*, durch

$$n! := \begin{cases} 1, & \text{für} \quad n = 0 \\ \prod\limits_{j=1}^{n} j, & \text{für} \quad n \neq 0 \end{cases}.$$

Bemerkung 2.2. Man kann sich leicht überlegen, dass $n!$ für $n \in \mathbb{N}^*$ die Anzahl aller Anordnungen von n Elementen ist.

Für zwei natürliche Zahlen n, k definieren wir den *Binomialkoeffizienten* $\binom{n}{k}$ als

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

Wir erkennen die Symmetrie $\binom{n}{k} = \binom{n}{n-k}$ für $k \leq n$.

Lemma 2.3 (Pascalsche[43] Gleichung)**.** *Für $n, k \in \mathbb{N}$ mit $1 \leq k \leq n$ gilt*

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

Beweis. Für $k = n$ ist die Formel richtig. Für $1 \leq k \leq n-1$ ist der Nachweis eine einfache Rechnung. □

Sie kennen bereits den Binomialkoeffizienten aus der Stochastik in der Schule:

Satz 2.4. *Die Anzahl der k-elementigen Teilmengen einer n-elementigen Menge ist $\binom{n}{k}$.*

Einen Beweis finden Sie z. B. in [Fo1, § 1]. Diese Aussage zeigt auch, dass Binomialkoeffizienten natürliche Zahlen sind.

Der Begriff „Binomialkoeffizient“ wird am Binomischen Lehrsatz deutlich. Die Binomialkoeffizienten sind gerade die Koeffizienten in der Summendarstellung des binomischen Lehrsatzes:

Satz 2.5 (Binomischer Lehrsatz)**.** *Seien $z, w \in \mathbb{C}$ und $n \in \mathbb{N}$, dann gilt*

$$(z+w)^n = \sum_{j=0}^{n} \binom{n}{j} z^{n-j} w^j.$$

Man kann den Binomischen Lehrsatz durch vollständige Induktion mit Hilfe der Pascalschen Gleichung beweisen (siehe z. B. [Fo1, § 1]). Diesen haben wir auch im Vorkurs vorgestellt. Sollten Sie nicht am Vorkurs teilgenommen haben, so ist der Beweise eine gute Übungsaufgabe. Ein anderer Beweis verwendet kombinatorische Argumente (siehe z. B. [K1, Kap. 1].

[43] Blaise Pascal (1623-1662) französischer Gelehrter

2.2. Komplexe Folgen und Grenzwerte

In diesem Abschnitt orientieren wir uns an [Fo1, § 4] und [K1, § 5]. In diesem Abschnitt lernen wir das zentrale und grundlegende Konzept der Analysis kennen: die *Konvergenz.*

Die hinter dem Konvergenzbegriff stehende Idee ist eine *beliebig genaue* Annäherung (in einem zu definierenden Sinn). Sie kennen das bereits von irrationalen Zahlen, wie z. B. π. Da π als irrationale Zahl nicht als abbrechende oder schließlich periodische Dezimalzahl darstellbar ist, wird in praktischen Anwendungen π oftmals angenähert, z. B. indem man deren Dezimaldarstellung an irgendeiner Stelle abbricht und die letzte Zahl rundet. Je später man abbricht um so genauer ist diese Approximation. Für komplexe Zahlen wird diese Idee der beliebig genauen Annäherung in diesem Abschnitt präzisiert und untersucht.

Wenn nichts anderes angegeben ist,
so bezeichnet von nun an n immer eine natürliche Zahl.

Definition 2.6. Sei X eine Teilmenge von $\mathbb{C}$. Eine *Folge in X* ist eine Abbildung

$$\mathbb{N} \to X, \quad n \mapsto a_n.$$

Hat diese Abbildung Werte in $\mathbb{C}$, ohne dass X näher spezifiziert ist, so sprechen wir von einer *komplexen Folge* oder kurz von einer *Folge.*

Hat diese Abbildung nur Werte in $\mathbb{R}$, so sprechen wir von einer *reellen Folge*.

Die Terminologie *Folge* wird deutlich, wenn wir uns eine solche Abbildung als „Auflistung“ Ihrer Werte vorstellen. Daher schreiben für eine Folge üblicherweise

$$(a_0, a_1, a_2, \dots) \quad \text{oder} \quad (a_n)_{n\in\mathbb{N}} \quad \text{oder kurz} \quad (a_n)$$

Wir nennen a_n auch das n-te Glied der Folge a.

Wir können die Abbildungsvorschrift für eine Folge explizit angeben, z. B. $a_n = c$ für die konstante Folge c oder $a_n = \frac{1}{n}$, oder rekursiv, wie beim Beispiel 1.20 der Fibonacci-Zahlen.

Definition 2.7. Eine komplexe Folge $(a_n)_{n\in\mathbb{N}}$ heißt *konvergent gegen* $a \in \mathbb{C}$, notiert $\lim\limits_{n\to\infty} a_n = a$ oder $a_n \xrightarrow{n\to\infty} a$, wenn gilt:

$$\forall \varepsilon > 0 \quad \exists N = N_\varepsilon \in \mathbb{N} \quad \forall n \geq N_\varepsilon : \quad |a_n - a| < \varepsilon.^{44}$$

Eine Folge (a_n) heißt *konvergent*, wenn es eine Zahl $a \in \mathbb{C}$ gibt, mit $a = \lim\limits_{n\to\infty} a_n$. Die Zahl a heißt dann *Grenzwert* oder *Limes* der Folge (a_n). Konvergiert eine Folge gegen 0 so spricht man auch von einer *Nullfolge*.

Eine Folge heißt *divergent*, wenn sie nicht konvergent ist.

[44] Beachten Sie, dass $\varepsilon > 0$ bedeutet, dass $\varepsilon \in \mathbb{R}_{>0}$ liegt.

Um die obigen Definitionen besser zu verstehen betrachten wir die Menge

$$D_\varepsilon(a) := \{z \in \mathbb{C} : \; |z - a| < \varepsilon\} \tag{2.1}$$

Da der Betrag in $\mathbb{C}$ gerade die übliche Länge von Vektoren im $\mathbb{R}^2 \cong \mathbb{C}$ beschreibt, ist $D_\varepsilon(a)$ die Menge aller Elemente in $\mathbb{C}$, welche zu $a \in \mathbb{C}$ einen Abstand von weniger als ε haben. Damit ist $D_\varepsilon(a)$ die Kreisscheibe (ohne Rand) mit Mittelpunkt a und Radius ε. Wir nennen diese Kreisscheibe $D_\varepsilon(a)$ um a auch *ε-Umgebung* von a.

Wenn eine Folge $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert, so heißt dass, egal wie klein wir den Radius ε auch wählen, sich immer *fast alle*, d. h. alle bis auf endlich viele[45], Folgeglieder a_n in der Kreisscheibe um a mit Radius ε liegen. Somit nähern sich die Folgeglieder a_n der Zahl a für große n „immer mehr und beliebig nah an“.

Für das Konvergenzverhalten spielen also *endlich viele* Folgeglieder keine Rolle.

Satz 2.8. *Sei (a_n) eine konvergente Folge, so ist ihr Grenzwert eindeutig bestimmt.*

Anschaulich ist dieser Satz einleuchtend. Angenommen es gäbe zwei verschiedene Grenzwerte a und $\tilde{a}$ von (a_n). Dann gibt es ein $\varepsilon > 0$ (z. B. $\varepsilon = \frac{|a-\tilde{a}|}{3}$), sodass die Umgebungen $D_\varepsilon(a)$ und $D_\varepsilon(\tilde{a})$ disjunkt sind. Wenn nun (a_n) einerseits gegen a und andererseits gegen $\tilde{a}$ konvergieren würde, so müssten fast alle Folgeglieder von a_n einerseits in $D_\varepsilon(a)$ und andererseits in $D_\varepsilon(\tilde{a})$ liegen, was absurd ist. Die Tatsache, dass man zwei verschiedene Punkte durch (offene) Umgebungen trennen kann, wird auch *Hausdorff*[46]*-Eigenschaft* genannt.

Beweis. Angenommen (a_n) hätte zwei verschiedene Grenzwerte a und $\tilde{a} \neq a$. Setze $\varepsilon := \frac{|a-\tilde{a}|}{3}$. Dann gibt es ein $N_\varepsilon \in \mathbb{N}$, sodass $|a_n - a| < \varepsilon$ für alle $n \geq N_\varepsilon$. Ebenso gibt es ein $\tilde{N}_\varepsilon \in \mathbb{N}$, sodass $|a_n - \tilde{a}| < \varepsilon$ für alle $n \geq \tilde{N}_\varepsilon$. Für $n \geq \max(N_\varepsilon, \tilde{N}_\varepsilon)$ erhalten wir mit der Dreiecksungleichung (Lemma 1.35) den Widerspruch:

$$|a - \tilde{a}| = |(a - a_n) + (a_n - \tilde{a})| \leq \underbrace{|a - a_n|}_{<\varepsilon} + \underbrace{|a_n - \tilde{a}|}_{<\varepsilon} < 2\varepsilon = \frac{2}{3}|a - \tilde{a}|.$$

□

Beispiel 2.9.

- Die konstante Folge $a_n = c$ ist offenbar konvergent gegen c.
- Die Folge $a_n = (-1)^n$ ist divergent.

  Angenommen a_n konvergiere gegen $a \in \mathbb{C}$, dann gibt es für $\varepsilon = \frac{1}{2}$ ein N_ε, sodass $|a - a_n| < \varepsilon$ für alle $n \geq N_\varepsilon$. Damit ergibt für $n \geq N_\varepsilon$ der Widerspruch:

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

[45] nämlich alle bis auf eventuell $a_0, a_1, \ldots, a_{N_\varepsilon - 1}$

[46] Felix Hausdorff (1868 - 1942), deutscher Mathematiker

Definition 2.10. Eine Folge komplexer Zahlen (a_n) heißt *beschränkt*, wenn die Menge $\{|a_n| : n \in \mathbb{N}\}$ beschränkt ist, d. h.

$$\exists R \in \mathbb{R}_{>0}\, \forall n \in \mathbb{N} : \ |a_n| \le R$$

also $a_n \in D_R(0)$ für alle $n \in \mathbb{N}$.

Satz 2.11. *Eine konvergente Folge (a_n) ist beschränkt.*[47]

Auch dieser Satz ist recht anschaulich: Bei einer konvergenten Folge liegen schliesslich fast alle Folgeglieder in einer kleinen Umgebung des Grenzwertes.

Beweis. Da (a_n) konvergent ist, gibt es ein $a \in \mathbb{C}$ und ein $N \in \mathbb{N}$, sodass $|a - a_n| < 1$ für alle $n \ge N$ gilt. Für $R := \max\{|a_0|, |a_1|, \ldots, |a_{N-1}|, |a+1|\}$ gilt dann $|a_n| \le R$ für alle $n \in \mathbb{N}$. □

Lemma 2.12 (Rechenregeln für Grenzwerte von Folgen)**.** *Seien (a_n) und (b_n) Folgen mit $a = \lim_{n\to\infty} a_n$ und $b = \lim_{n\to\infty} b_n$. Dann gilt:*

(i) $(a_n + b_n)$ konvergiert gegen $a + b$, also $\lim_{n\to\infty}(a_n + b_n) = a + b$.

(ii) $(a_n b_n)$ konvergiert gegen ab, also $\lim_{n\to\infty}(a_n b_n) = ab$. Für $b_n = c$ erhalten wir insbesondere $\lim_{n\to\infty}(c a_n) = ca$.

(iii) $(a_n - b_n)$ konvergiert gegen $a - b$, also $\lim_{n\to\infty}(a_n - b_n) = a - b$.

(iv) Wenn $b \ne 0$, so gilt $b_n \ne 0$ für alle bis auf endlich viele $n \in \mathbb{N}$, und $\frac{a_n}{b_n}$ konvergiert gegen $\frac{a}{b}$, also $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}$.

Beweis. ad (i): Sei $\varepsilon > 0$, so gibt es $N^a, N^b \in \mathbb{N}$, sodass $|a - a_n| < \frac{1}{2}\varepsilon$ für alle $n \ge N^a$ und $|b - b_n| < \frac{1}{2}\varepsilon$ für alle $n \ge N^b$. Wenn $n \ge \max(N^a, N^b) =: N$, dann erhalten wir

$$|(a+b) - (a_n + b_n)| \le |a - a_n| + |b - b_n| \le \varepsilon.$$

ad (ii): Als konvergente Folge ist (a_n) beschränkt, d. h. es gibt $R_a \in \mathbb{R}$ sodass $|a_n| \le R_a$. Setze $R := \max(R_a, |b|)$.

Sei nun $\varepsilon > 0$. Da $a = \lim_{n\to\infty} a_n$ und $b = \lim_{n\to\infty} b_n$ gibt es $N_a, N_b \in \mathbb{N}$, sodass gilt:

$$\forall n \ge N_a : \ |a - a_n| < \frac{\varepsilon}{2R} \quad \text{und} \quad \forall n \ge N_b : \ |b - b_n| < \frac{\varepsilon}{2R}.$$

Setze $N := \max(N_a, N_b)$. Dann gilt für alle $n \ge N$:

$$\begin{aligned} |a_n b_n - ab| &= |a_n(b_n - b) + (a_n - a)b| \le |a_n(b_n - b)| + |(a_n - a)b| \\ &= |a_n| \cdot |b_n - b| + |a_n - a| \cdot |b| \\ &< 2R\frac{\varepsilon}{2R} = \varepsilon. \end{aligned}$$

[47] Die Umkehrung von Satz 2.11 gilt nicht, wie das Beispiel $a_n = (-1)^n$ zeigt.

ad (iii): Folgt fast unmittelbar aus (i) und (ii).

ad (iv): Wegen (ii) genügt es zu zeigen, dass gilt: Wenn $b \neq 0$, so gilt $b_n \neq 0$ für alle bis auf endlich viele $n \in \mathbb{N}$, und $\frac{1}{b_n}$ konvergiert gegen $\frac{1}{b}$.

Für $\tilde{\varepsilon} := \frac{|b|}{2} > 0$, gibt es ein $\tilde{N} \in \mathbb{N}$, sodass $|b_n - b| < \tilde{\varepsilon} = \frac{|b|}{2}$ für alle $n \geq \tilde{N}$. Damit gilt insbesondere $|b_n| \geq \frac{|b|}{2} > 0$ für alle $n \geq \tilde{N}$.

Sei nun $\varepsilon > 0$ vorgegeben. Dann gibt es ein $N \in \mathbb{N}$, sodass für alle $n \geq N$ gilt:

$$|b_n - b| < \frac{\varepsilon |b|^2}{2}.$$

Für $n \geq \max(\tilde{N}, N)$ gilt dann

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \left|\frac{b - b_n}{b b_n}\right| = \frac{|b - b_n|}{|b| \cdot |b_n|} < \frac{2\varepsilon |b|^2}{2|b|^2} = \varepsilon.$$

□

Lemma 2.13. *Sei (a_n) eine komplexe Folge, die gegen a konvergiert, dann gilt:*

(i) $\lim\limits_{n\to\infty} |a_n| = |a|$,

(ii) $\lim\limits_{n\to\infty} \overline{a_n} = \overline{a}$,

(iii) $\lim\limits_{n\to\infty} \Re(a_n) = \Re(a)$ *und* $\lim\limits_{n\to\infty} \Im(a_n) = \Im(a)$,

(iv) $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} \Re(a_n) + i \lim\limits_{n\to\infty} \Im(a_n)$,

(v) Der Grenzwert einer konvergenten reellen Folge (a_n) ist reell.

Der Beweis ist eine Übungsaufgabe. Die folgenden Funktionen der Auf- bzw. der Abrundung heißen auch *Gaußklammern*[48]:

$$\begin{aligned} \lceil\ \rceil :\ \ & \mathbb{R} \to \mathbb{Z} \quad x \mapsto \lceil x \rceil := \min\{k \in \mathbb{Z} :\ k \geq x\}, \\ \lfloor\ \rfloor :\ \ & \mathbb{R} \to \mathbb{Z} \quad x \mapsto \lfloor x \rfloor := \max\{k \in \mathbb{Z} :\ k \leq x\}. \end{aligned}$$

Beispiel 2.14 (Grenzwerte nützlicher Folgen)**.**

(i) $\lim\limits_{n\to\infty} \frac{1}{n^q} = 0$ für alle $q \in \mathbb{Q}_{>0} = \{x \in \mathbb{Q} :\ x > 0\}$[49].

(ii) $\lim\limits_{n\to\infty} \sqrt[n]{x} = 1$ für alle $x \in \mathbb{R}_{>0}$.

(iii) $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$.

[48] Johann Carl Friedrich Gauß (1777 - 1855), deutscher Mathematiker und Naturwissenschaftler

[49] Für $k \in \mathbb{N}$ und $j \in \mathbb{N}^*$ gilt $n^{\frac{k}{j}} := (n^k)^{\frac{1}{j}} = \sqrt[j]{n^k}$ und $n^{-\frac{k}{j}} = \frac{1}{n^{\frac{k}{j}}}$.

(iv) $\lim\limits_{n\to\infty} z^n = 0$ für alle $z \in D_1(0) = \{w \in \mathbb{C} : \ |w| < 1\}$.

(v) $\lim\limits_{n\to\infty} \frac{n^k}{z^n} = 0$ für alle $k \in \mathbb{N}$ und alle $z \in \mathbb{C}$ mit $|z| > 1$.

Beweis. ad (i): Für $\varepsilon > 0$, wähle $N := \left\lceil \varepsilon^{-\frac{1}{q}} + 1 \right\rceil$. Dann gilt für alle $n \geq N$:

$$\left|0 - \frac{1}{n^q}\right| = \left(\frac{1}{n}\right)^q \leq \left(\frac{1}{N}\right)^q = \left(\frac{1}{\left\lceil \varepsilon^{-\frac{1}{q}} + 1 \right\rceil}\right)^q \leq \left(\frac{1}{\varepsilon^{-\frac{1}{q}} + 1}\right)^q < \left(\frac{1}{\varepsilon^{-\frac{1}{q}}}\right)^q = \varepsilon.$$

ad (ii): Wegen Aussage (iv) von Lemma 2.12 genügt es $x \geq 1$ zu betrachten, da man im Fall $0 < x < 1$ einfach $\frac{1}{x}$ anstelle von x betrachtet.

Aus $x > 1$ folgt zunächst $\sqrt[n]{x} > 1$ und damit $x_n := \sqrt[n]{x} - 1 > 0$. Wir erhalten mit der Bernoullischen Ungleichung (Satz 1.26) $x = (1 + x_n)^n \geq 1 + nx_n > nx_n$ und damit $x_n < \frac{x}{n}$. Somit gilt für alle $n \geq N := \left\lceil \frac{x}{\varepsilon} \right\rceil$:

$$|\sqrt[n]{x} - 1| = x_n < \varepsilon.$$

ad (iii): Wir nehmen ohne Einschränkung $n \geq 2$ an. Da $x_n := \sqrt[n]{n} - 1 > 0$ folgt aus dem Binomischen Lehrsatz (Satz 2.5):

$$n = (1 + x_n)^n = \sum_{j=0}^{n} \binom{n}{j} x_n^j > 1 + \binom{n}{2} x_n^2 = 1 + \frac{n(n-1)}{2} x_n^2$$

und damit $x_n < \sqrt{2/n}$.

Für gegebenes $\varepsilon > 0$ wählen wir $N = \max\left\{2, \left\lceil \frac{2}{\varepsilon^2} \right\rceil\right\}$ und erhalten für $n \geq N$:

$$|\sqrt[n]{n} - 1| = x_n < \varepsilon.$$

ad (iv): Eine direkte Konsequenz von Korollar 1.27(ii).

ad (v): Wir setzen $x = |z| - 1 > 0$ und wählen ein $p \in \mathbb{N}$ mit $p > k$. Wenn $n \geq 2p$ erhalten wir aus dem Binomischen Lehrsatz (Satz 2.5):

$$|z|^n = (1 + x)^n > \binom{n}{p} x^p = \frac{n!}{(n-p)!p!} x^p > \frac{n^p}{2^p p!} x^p.$$

Da $p > k$ gilt auch $n^{p-k} \geq n$ und damit

$$\left|\frac{n^k}{z^n}\right| < \frac{2^p p!}{x^p n^{p-k}} \leq \frac{2^p p!}{x^p} \cdot \frac{1}{n}.$$

Für $n \geq N := \max\left(\left\lceil \frac{2^p p!}{x^p} \frac{1}{\varepsilon} \right\rceil, 2p\right)$ erhalten wir

$$\left|\frac{n^k}{z^n}\right| < \varepsilon.$$

□

2.3. Konvergenzsätze für reelle Folgen

Im Gegensatz zu $\mathbb{C}$ ist $\mathbb{R}$ ein vollständiger geordneter Körper. Diese Anordnung können wir verwenden, um nützliche Aussagen zu machen, welche besonders auf die Situation reeller Folgen zugeschnitten sind. Wir orientieren uns weiter an [Fo1, § 4] und [K1, § 5].

Lemma 2.15 (Schwache Monotonie von Grenzwerten). *Seien (a_n) und (b_n) zwei konvergente reelle Folgen mit $a = \lim_{n\to\infty} a_n$ und $b = \lim_{n\to\infty} b_n$. Wenn $a_n \leq b_n$ für alle $n \in \mathbb{N}$ mit $n \geq N_0$ gilt, dann folgt $a \leq b$.*[50]

Beweis. Für $\varepsilon > 0$ gibt es einerseits $N_a \in \mathbb{N}$, sodass $a - \varepsilon < a_n$ für alle $n \geq N_a$ und andererseits ein $N_b \in \mathbb{N}$, sodass $b_n < b + \varepsilon$ für alle $n \geq N_b$. Wenn $n \geq N := \max(N_a, N_b, N_0)$, dann gilt $a - \varepsilon < a_n \leq b_n < b + \varepsilon$ und damit $a - b < 2\varepsilon$. Also muss $a - b \leq 0$ gelten. □

Korollar 2.16. *Sei (a_n) eine konvergente reelle Folge mit $a_n \in [x, y]$ für alle $n \in \mathbb{N}$ mit $n \geq N_0$, dann gilt auch $\lim_{n\to\infty} a_n \in [x, y]$.*[51]

Satz 2.17 (Einschließungsregel, Sandwichlemma). *Seien (a_n) und (c_n) zwei konvergente reelle Folgen mit dem selben Grenzwert $L = \lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n$. Sei (b_n) eine reelle Folge mit*

$$a_n \leq b_n \leq c_n \quad \textit{für alle}\ n \in \mathbb{N}, n \geq N_0.$$

Dann konvergiert auch b_n gegen L.

Beweis. Sei $\varepsilon > 0$ vorgegeben, so gibt es ein $N \in \mathbb{N}$, $N \geq N_0$ sodass für alle $n \geq N$ gilt $a_n, c_n \in (L - \varepsilon, L + \varepsilon)$. Mit $a_n \leq b_n \leq c_n$ folgt dann auch $b_n \in (L - \varepsilon, L + \varepsilon)$ für alle $n \geq N$ und somit die Aussage. □

Wir haben bereits gesehen, dass jede konvergente Folge beschränkt ist. Jedoch ist eine beschränkte Folge keinesfalls auch zwingend konvergent, wie das Beispiel $a_n = (-1)^n$ zeigt.

Definition 2.18. Eine reelle Folge (a_n) heißt

- *schließlich monoton wachsend* bzw. *schließlich streng monoton wachsend*, wenn es ein $N_0 \in \mathbb{N}$ gibt, sodass für alle $n \geq N_0$ gilt

  $$a_n \leq a_{n+1}, \quad \text{bzw.} \quad a_n < a_{n+1}.$$

  Ist $N_0 = 0$ in dieser Definition, so sprechen wir von einer monoton wachsenden Folge.

[50] Aus $a_n < b_n$ kann man hingegen *nicht* $a < b$, sondern nur $a \leq b$ folgern, wie das Beispiel $a_n = 0$ und $b_n = \frac{1}{n}$ zeigt.

[51] Hierbei ist es wichtig, dass das Intervall $[x, y]$ abgeschlossen ist. Denn für $a_n = \frac{1}{n}$ gilt zwar $a_n \in (0, 1]$ für alle $n \in \mathbb{N}^*$, aber $0 = \lim_{n\to\infty} a_n \notin (0, 1]$.

- *schließlich monoton fallend*, bzw. *schließlich streng monoton fallend*,wenn es ein $N_0 \in \mathbb{N}$ gibt, sodass für alle $n \geq N_0$ gilt

$$a_n \geq a_{n+1}, \quad \text{bzw.} \quad a_n > a_{n+1}.$$

  Ist $N_0 = 0$ in dieser Definition, so sprechen wir von einer monoton fallenden Folge.

- *(schließlich) monoton*, wenn (a_n) (schließlich) monoton wachsend oder (schließlich) monoton fallend ist.

- *nach oben beschränkt*, wenn die Menge $\{a_n : n \in \mathbb{N}\}$ nach oben beschränkt ist.

- *nach unten beschränkt*, wenn die Menge $\{a_n : n \in \mathbb{N}\}$ nach unten beschränkt ist.

Satz 2.19. *(i) Jede schließlich wachsende und nach oben beschränkte reelle Folge ist konvergent.*

(ii) Jede schließlich fallende und nach unten beschränkte reelle Folge ist konvergent.

Beweis. Wir beweisen nur die Aussage (i). Der Beweis von (ii) geht analog (Übung). Sei (a_n) eine schließlich wachsende und nach oben beschränkte reelle Folge und $N_0 \in \mathbb{N}$ so, dass $a_{n+1} \geq a_n$ für alle $n \geq N_0$ gilt.

Wir betrachten nun die Menge $A := \{a_n : n \in \mathbb{N}, n \geq N_0\}$. Da die Folge a_n nach oben beschränkt ist, ist auch A eine nach oben beschränkte Menge. Damit existiert wegen der Ordnungsvollständigkeit von $\mathbb{R}$ das Supremum $s = \sup(A)$. Sei nun $\varepsilon > 0$ vorgegeben. Da s die kleinste obere Schranke von A ist, gibt es ein $N \in \mathbb{N}, N \geq N_0$, sodass $s - \varepsilon < a_N$. Für alle $n \geq N$ erhalten wir nun

$$s - \varepsilon < a_N \leq a_n \leq s < s + \varepsilon.$$

□

2.4. Der Satz von Bolzano-Weierstraß

In diesem Abschnitt folgen wir der Darstellung in [K1, Abschn. 5.5].

Definition 2.20. Sei (a_n) eine Folge komplexer Zahlen. Eine Zahl $h \in \mathbb{C}$ heißt *Häufungswert*[52] von (a_n), wenn gilt: Für alle $\varepsilon > 0$ liegen unendlich viele Folgeglieder in $D_\varepsilon(h) = \{z \in \mathbb{C} : |z - h| < \varepsilon\}$.

Beispiel 2.21. (i) Der einzige Häufungswert einer konvergenten Folge ist ihr Grenzwert (Übung).

[52] In der Literatur ist hierfür der Begriff *Häufungspunkt* einer Folge gebräuchlicher. Da dieser Begriff jedoch etwas doppeldeutig ist, da man zwischen Häufungspunkten von Folgen und Mengen unterscheiden muss, schließe ich mich gerne dem Vorschlag von Cédric Martin an und spreche vom Häufungswert einer Folge.

(ii) Die Folge $a_n = (-1)^n$ hat zwei Häufungswerte, nämlich 1 und -1.

(iii) Sei $f : \mathbb{N} \to \mathbb{Q}$ eine bijektive Abbildung (Cantors Abzählungsschema), dann ist jede reelle Zahl ein Häufungswert der Folge $a_n = f(n)$ (vgl. Satz 1.25.

(iv) Die Folge $a_n = n$ hat keinen Häufungswert.

Aufgabe 2.22. Sei (a_n) eine komplexe Folge und $h \in \mathbb{C}$. Zeigen Sie, dass folgende Aussagen äquivalent sind:

- h ist ein Häufungswert der Folge (a_n).
- $\forall \varepsilon > 0 \, \forall N \in \mathbb{N} \, \exists n \geq N : \; |a_n - h| < \varepsilon$.

Definition 2.23. Sei (a_n) eine Folge und $n_0 < n_1 < ...$ eine streng monoton wachsende Folge natürlicher Zahlen, dann heißt die Folge

$$k \mapsto a_{n_k}$$

Teilfolge von (a_n).

Lemma 2.24. *Sei (a_{n_k}) eine Teilfolge einer gegen a konvergenten Folge (a_n), dann konvergiert auch a_{n_k} gegen a.*

Der Beweis ist eine Übungsaufgabe.

Lemma 2.25. *Für eine komplexe Folge (a_n) und eine Zahl $h \in \mathbb{C}$ sind folgende Aussagen äquivalent:*

1. *h ist ein Häufungswert von (a_n).*
2. *Es gibt eine Teilfolge von (a_n) die gegen h konvergiert.*

Beweis. Ist h ein Häufungswert von (a_n), so liegen in jeder ε-Umgebung von h unendlich viele Folgeglieder. Für $\varepsilon = 1$ wählen wir ein $n_0 \in \mathbb{N}$, sodass $|a_{n_0} - h| < 1$. Seien nun $n_0, \ldots, n_k$ bereits gewählt. Dann wählen wir n_{k+1} so, dass einerseits $n_{k+1} > n_k$ und andererseits $|a_{n_{k+1}} - a| < \frac{1}{k+1}$ gilt. Die so konstruierte Teilfolge (a_{n_k}) konvergiert nach Konstruktion gegen h (vgl. dazu Korollar 1.22).

Sei nun (a_{n_k}) eine Teilfolge von (a_n), die gegen h konvergiert, so liegen in jeder ε-Umgebung $D_\varepsilon(h)$ unendlich viele Glieder der Teilfolge (a_{n_k}) und damit auch unendlich viele Glieder der Folge (a_n). □

Theorem 2.26 (Satz von Bolzano[53]-Weierstraß[54] für reelle Folgen)**.** *Jede beschränkte Folge reeller Zahlen hat eine konvergent Teilfolge.*

[53] Bernardus Placidus Johann Nepomuk Bolzano (1781 - 1848), böhmischer Mathematiker und Philosoph, römisch-katholischer Priester

[54] Karl Theodor Wilhelm Weierstraß (1815 - 1897), deutscher Mathematiker

Beweis. Sei (a_n) eine reelle beschränkte Folge, d. h. es gibt ein $R > 0$, sodass $a_n \in [-R, R]$ für alle $n \in \mathbb{N}$ gilt.

Wir konstruieren nun eine Intervallschachtelung und eine konvergente Teilfolge. Wir setzen $I_0 := [-R, R] = [b_0, c_0]$ und $n_0 := 0$.

Seien nun $I_0, I_1, \dots, I_k$ und $n_0, n_1, \dots, n_k$ bereits konstruiert, wobei I_k unendlich viele Glieder der Folge (a_n) enthält, so konstruieren wir I_{k+1} wie folgt: Wir teilen $I_k = [b_k, c_k]$ in zwei gleich große Teilintervalle $[b_k, (b_k + c_k)/2]$ und $[(b_k + c_k)/2, c_k]$ und wählen aus diesen beiden eines aus, in welchem unendlich viele Folgeglieder von (a_n) liegen und nennen es $I_{k+1} := [b_{k+1}, c_{k+1}]$. Dann wählen $n_{k+1} > n_k$ mit $a_{n_{k+1}} \in I_{k+1}$.

Die so konstruierte Familie I_k ist eine Intervallschachtelung, da einerseits $I_{k+1} \subset I_k$ für alle $k \in \mathbb{N}$ gilt. Andererseits folgt aus $|I_{k+1}| = \frac{|I_k|}{2}$ induktiv:

$$|I_k| = \frac{1}{2^k}|I_0| = \frac{1}{2^k}2R.$$

Wegen Korollar 1.27 gibt es damit für jedes $\varepsilon > 0$ ein $K \in \mathbb{N}$, sodass $|I_K| < \varepsilon$.

Sei $h \in \mathbb{R}$ die eindeutige Zahl mit $h \in I_k$ für alle $k \in \mathbb{N}$. Da $a_{n_k} \in I_k$ gibt es für jedes $\varepsilon > 0$ ein $K \in \mathbb{N}$, sodass $|I_k| \leq \varepsilon$ für alle $k \geq K$ gilt. Wir erhalten somit für alle $k \geq N$

$$|a_{n_k} - h| \leq |I_k| < \varepsilon.$$

Damit konvergiert die Teilfolge (a_{n_k}) gegen h. □

Theorem 2.27 (Satz von Bolzano-Weierstraß (komplexe Folgen))**.** *Jede beschränkte Folge komplexer Zahlen hat einen Häufungswert.*

Beweis. Sei (a_n) eine beschränkte Folge komplexer Zahlen, d. h. es gibt ein $R > 0$, sodass $|a_n| \leq R$ und somit $\Re(a_n), \Im(a_n) \in [-R, R]$ und für alle $n \in \mathbb{N}$ gilt.

Nach Theorem 2.26 gibt es ein natürliche Zahlen $n_0 < n_1 < \cdots < n_k < \dots$, sodass die Folge $(\Re(a_{n_k}))$ gegen eine reelle Zahl h_1 konvergiert.

Nun ist die Teilfolge (a_{n_k}) immer noch beschränkt und es gibt natürliche Zahlen $n_{k_0} < n_{k_1} < \cdots < n_{k_j} < \dots$ mit $n_{k_j} \in \{n_k : k \in \mathbb{N}\}$ für alle $j \in \mathbb{N}$, sodass $\left(\Im\left(a_{n_{k_j}}\right)\right)$ gegen eine reelle Zahl h_2 konvergiert.

Die Teilfolge $\left(a_{n_{k_j}}\right)$ von (a_n) konvergiert nach Lemma 2.24 und 2.13 gegen $h_1 + ih_2$. Die Aussage folgt nun mit Lemma 2.25. □

2.5. Cauchyfolgen und Banachscher Fixpunktsatz

Wir folgen in diesem Abschnitt [K1, Abschn. 5.6] sowie [A1, Abschn. 2.4]

Definition 2.28. Eine komplexe Folge (a_n) heißt *Cauchyfolge*, wenn gilt:

$$\forall \varepsilon > 0\ \exists N_\varepsilon > 0\ \forall n, m \geq N : \ |a_n - a_m| < \varepsilon$$

(hierbei sind n und m selbstverständlich natürliche Zahlen).

Satz 2.29. *Für eine komplexe Folge (a_n) sind folgende Aussagen äquivalent:*

(i) (a_n) *ist konvergent*

(ii) (a_n) *ist eine Cauchyfolge.*

Beweis. Konvergiere zunächst (a_n) gegen $a \in \mathbb{C}$ und sei ε gegeben. So gibt es ein $N \in \mathbb{N}$, sodass $|a_k - a| < \frac{\varepsilon}{2}$ für alle $k \geq N$ gilt. Für $n, m \geq N$ erhalten wir mit der Dreiecksungleichung

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

Somit ist (a_n) eine Cauchyfolge.

Sei nun (a_n) eine Cauchyfolge. Dann gibt es ein $N \in \mathbb{N}$, sodass für alle $n \geq N$ gilt

$$|a_n| - |a_N| \leq |a_n - a_N| < 1.$$

Damit ist aber (a_n) durch $\max\{|a_0|, ..., |a_{N-1}|, |a_N| + 1\}$ beschränkt.
Damit besitzt (a_n) eine konvergente Teilfolge (a_{n_k}) (Satz von Bolzano-Weierstraß). Der Grenzwert der Teilfolge (a_{n_k}) sei mit a bezeichnet. Sei nun $\varepsilon > 0$ gegeben, da gibt es ein $N_1 \in \mathbb{N}$, sodass für alle $n, m \geq N_1$ gilt: $|a_n - a_m| < \frac{\varepsilon}{2}$. Ferner gibt es ein $N_2 \in \mathbb{N}$, sodass für alle $n_k \geq N_2$ gilt: $|a_{n_k} - a| < \frac{\varepsilon}{2}$. Für $\tilde{N} = \max(N_1, N_2)$ erhalten wir mit einem gewählten $n_k \geq \tilde{N}$ für alle $n \geq \tilde{N}$:

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

Damit konvergiert (a_n) gegen a. □

Bemerkung 2.30. Man kann ein Modell von $\mathbb{R}$ aus $\mathbb{Q}$ wie folgt konstruieren: Man betrachte die Menge aller Äquivalenzklassen von Cauchyfolgen in $\mathbb{Q}$. Hierbei heißen zwei Cauchyfolgen (a_n) und (b_n) in $\mathbb{Q}$ äquivalent, wenn die Differenzfolge $(a_n - b_n)$ gegen Null konvergiert. Auf diesem Quotienten kann man dann in natürlicher Weise eine Körperstruktur und eine Ordnungsrelation definieren, sodass man einen ordnungsvollständigen geordneten Körper erhält (siehe z. B. den Abschnitt [Ebb, Kap. 2, § 3] von K. Mainzer).

Definition 2.31. Sei $X = \mathbb{R}$ oder $X = \mathbb{C}$. Eine Abbildung $f : X \to X$ heißt *Kontraktion* von X, wenn gilt:

$$\exists K \in [0, 1)\ \forall x, y \in X : \ |f(x) - f(y)| \leq K|x - y|.$$

Definition 2.32. Ein Punkt $x \in X$ heißt *Fixpunkt* einer Abbildung $f : X \to X$, wenn $f(x) = x$.

Theorem 2.33 (Banachscher[55] Fixpunktsatz). *Sei $X = \mathbb{R}$ oder $X = \mathbb{C}$, dann hat jede Kontraktion $f : X \to X$ genau einen Fixpunkt.*

[55] Stefan Banach (1892 - 1945), polnischer Mathematiker

Beweis. Unser Beweis liefert sogar eine Konstruktion des Punktes a. Sei $a_0 \in X$ beliebig. Wir definieren rekursiv $a_{n+1} := f(a_n)$.

Behauptung 1: (a_n) *ist eine Cauchyfolge*

Da $|f(x) - f(y)| \leq K|x - y|$ für ein $K \in [0, 1)$ und für alle $x, y \in X$ gilt, erhalten wir für $n \geq 1$:

$$|a_{n+1} - a_n| = |f(a_n) - f(a_{n-1})| \leq K|a_n - a_{n-1}|.$$

Durch vollständige Induktion beweist man damit für alle $n \geq 1$:

$$|a_{n+1} - a_n| \leq K^n |a_1 - a_0|.$$

Seien nun n, m mit $n > m$ beliebig gewählt, so folgt mit der Dreiecksungleichung:

$$|a_n - a_m| \leq |a_n - a_{n-1}| + \cdots + |a_{m+1} - a_m| \leq (K^{n-1} + \cdots + K^m)|a_1 - a_0|.$$

Wegen

$$(1 - K)(K^{n-1} + \cdots + K^m) = K^m - K^n \leq K^m$$

erhalten wir

$$|a_n - a_m| \leq \frac{K^m}{1 - K}|a_1 - a_0|.$$

Da $K \in [0, 1)$ folgt aus Beispiel 2.14(iv):

$$\lim_{m\to\infty} \frac{K^m}{1 - K}|a_1 - a_0| = 0.$$

Sei also $\varepsilon > 0$, dann gibt es ein $N \in \mathbb{N}$, sodass für alle $n > m \geq N$ gilt:

$$|a_n - a_m| \leq \varepsilon.$$

Damit erkennt man, dass (a_n) eine Cauchyfolge ist.

Behauptung 2: *Der Grenzwert* a *von* (a_n) *ist ein Fixpunkt von* f.

Sei $\varepsilon > 0$, so gibt es ein $N \in \mathbb{N}$ mit $|a_n - a| < \varepsilon$ für alle $n \geq N$. Somit gilt aber für $n \geq N$:

$$|a_{n+1} - f(a)| = |f(a_n) - f(a)| \leq K|a_n - a| < K\varepsilon < \varepsilon.$$

Die Teilfolge (a_{n+1}) der konvergenten Folge (a_n) konvergiert also gegen $f(a)$. Damit konvergiert aber auch (a_n) gegen $f(a)$. Da aber (a_n) ebenso gegen a konvergiert, ist $f(a) = a$ wegen Satz 2.8.

Behauptung 3: *a ist der einzige Fixpunkt on f.* Sei $b \neq a$ ein weiterer Fixpunkt von f, so gilt:

$$|a - b| = |f(a) - f(b)| \leq K|a - b| < |a - b|,$$

ein Widerspruch. □

Ausblick

Für den Begriff der Konvergenz einer Folge in einer Menge X benötigt man eine Struktur, die feststellt, welche Punkte „eng benachbart“ sind. Eine solche Struktur heißt *Topologie* auf X. In der Analysis treten häufig Topologien auf, die durch eine *Abstandsfunktion*, auch *Metrik* genannt, induziert sind. Eine *Abstandsfunktion* auf X ist eine Abbildung $d : X \times X \to X$, welche folgende Eigenschaften hat:

- *positive Definitheit:* $\forall x, y \in X : \ d(x, y) \geq 0$ und $d(x, y) = 0 \iff x = y$,
- *Symmetrie:* $\forall x, y \in X : \ d(x, y) = d(y, x)$,
- *Dreiecksungleichung:* $\forall x, y, z \in X : \ d(x, z) + d(z, y) \geq d(x, y)$.

Eine Menge X mit einer Abstandfunktion d heißt auch *metrischer Raum*.

Typische Beispiele wären $X = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ mit $d(x, y) = |x - y|$. Eine Folge $a_0, \ a_1, \ldots$ mit $a_n \in X$, wobei X ein metrischer Raum ist, konvergiert dann gegen $a \in X$, wenn gilt

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall n \geq N : \ d(a_n, a) < \varepsilon.$$

Der Grenzwert konvergenter Folgen in einem metrischen Raum ist wieder eindeutig bestimmt. Auch der Begriff einer Cauchyfolge überträgt sich auf metrische Räume indem man sagt: Eine Folge (a_n) in X ist eine Cauchyfolge, wenn gilt:

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall n, m \geq N : \ d(a_n, a_m) < \varepsilon.$$

Wie bereits das Beispiel $\mathbb{Q}$ mit der üblichen Abstandsfunktion zeigt, müssen Cauchyfolgen in metrischen Räumen nicht immer konvergieren. Metrische Räume, in welchen jede Cauchyfolge konvergent ist, heißen *vollständig*. Beispiele hierfür sind $\mathbb{R}$ und $\mathbb{C}$ mit den üblichen Abstandsfunktionen.

Der oben gegebene Beweis des Banachschen Fixpunktsatzes kann sogar einfach auf Kontraktionen eines sogenannten vollständigen metrischen Raumes übertragen werden (siehe z. B. [A1, Satz 2.4.7]), sodass der Banachsche Fixpunktsatz für Kontraktionen eines beliebigen vollständigen metrischen Raumes gilt.

2.6. Bestimmte Divergenz

Wir verweisen auf [K1, Abschn. 5.7] und [Fo1, § 4].

Bestimmte Divergenz komplexer Folgen

Definition 2.34. Eine *komplexe* Folge (a_n) *divergiert bestimmt* gegen Unendlich[56], notiert $\lim\limits_{n \to \infty} a_n = \infty$, wenn gilt:

$$\forall R > 0 \ \exists N \in \mathbb{N} \ \forall n \geq N : \ |a_n| > R.$$

[56] in manchen Lehrbüchern spricht man auch von *uneigentlicher Konvergenz* gegen Unendlich

Bestimmte Divergenz reeller Folgen

Definition 2.35. Eine *reelle* Folge (a_n)

- *divergiert bestimmt* gegen Unendlich[57], notiert $\lim_{n\to\infty} a_n = \infty$, wenn gilt:

$$\forall R > 0\ \exists N \in \mathbb{N}\ \forall n \geq N :\ a_n > R.$$

- *divergiert bestimmt* gegen $-\infty$[58], notiert $\lim_{n\to\infty} a_n = -\infty$, wenn gilt:

$$\forall R > 0\ \exists N \in \mathbb{N}\ \forall n \geq N :\ a_n < -R.$$

Vorsicht. Eine reelle Folge, die in $\mathbb{R}$ bestimmt divergiert, divergiert auch als komplexe Folge bestimmt. Umgekehrt gilt das aber *nicht*. Die reelle Folge $a_n = (-2)^n$ divergiert als komplexe Folge bestimmt gegen ∞, als reelle Folge divergiert sie zwar, aber nicht bestimmt.

Beobachtung 2.36. *Jede nicht beschränkte monotone reelle Folge divergiert bestimmt.*

Aufgabe 2.37. Sei (a_n) eine reelle Folge. Zeigen Sie:

(a) Wenn $\lim_{n\to\infty} a_n = \pm\infty$, dann gilt $a_n \neq 0$ für fast alle $n \in \mathbb{N}$, und es gilt $\lim_{n\to\infty} \frac{1}{a_n} = 0$.

(b) Wenn $a_n > 0$ für fast alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} a_n = 0$, dann gilt $\lim_{n\to\infty} \frac{1}{a_n} = \infty$.

(c) Wenn $a_n < 0$ für fast alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} a_n = 0$, dann gilt $\lim_{n\to\infty} \frac{1}{a_n} = -\infty$.

Definition 2.38. Für eine beschränkte reelle Folge (a_n) betrachten wir die Folgen (s_k) und (i_k) mit $s_k := \sup\{a_n :\ n \geq k\}$ und $i_k := \inf\{a_n :\ n \geq k\}$ und definieren[59] den

- *Limes superior* von (a_n), notiert $\limsup_{n\to\infty} a_n$, durch

$$\limsup_{n\to\infty} a_n := \lim_{k\to\infty} s_k.$$

- *Limes inferior* von (a_n), notiert $\liminf_{n\to\infty} a_n$, durch

$$\liminf_{n\to\infty} a_n := \lim_{k\to\infty} i_k.$$

Ist eine Folge (a_n) nicht nach oben beschränkt, so definieren wir $\limsup_{n\to\infty} a_n := \infty$. Ist eine Folge (a_n) nicht nach unten beschränkt, so definieren wir $\liminf_{n\to\infty} a_n := -\infty$.

---

[57] in manchen Lehrbüchern spricht man auch von *uneigentlicher Konvergenz* gegen Unendlich

[58] In manchen Lehrbüchern spricht man auch von *uneigentlicher Konvergenz* gegen $-\infty$.

[59] Da die Folge (s_k) monoton fallend und die Folge (i_k) monoton wachsend ist und beide beschränkt sind, existiert der Limes superior und der Limes inferior.

Aufgabe 2.39 (Eigenschaften des Limes superior und des Limes inferior)**.** Sei (a_n) eine beschränkte reelle Folge und H die Menge aller Häufungswerte von (a_n). Zeigen Sie:

(a) $\limsup\limits_{n\to\infty} a_n = \inf\{\sup\{a_k : k \geq n\} : n \in \mathbb{N}\}$.

(b) $\liminf\limits_{n\to\infty} a_n = \sup\{\inf\{a_k : k \geq n\} : n \in \mathbb{N}\}$.

(c) $\sup(H)$ und $\inf(H)$ existieren, und es gilt $\sup(H) \in H$ sowie $\inf(H) \in H$.

(d) $\limsup\limits_{n\to\infty} a_n = \sup(H)$ und $\liminf\limits_{n\to\infty} a_n = \inf(H)$.

(e) (a_n) ist genau dann konvergent, wenn $\liminf\limits_{n\to\infty} a_n = \limsup\limits_{n\to\infty} a_n$.

(f) $\limsup\limits_{n\to\infty} a_n \geq \liminf\limits_{n\to\infty} a_n$.

(g) $\liminf\limits_{n\to\infty}(-a_n) = -\limsup\limits_{n\to\infty} a_n$.

2.7. Reihen

Für diesen Abschnitt vergleiche auch [K1, Kap. 6] und [Fo1, § 4, § 7, § 8].

Ist eine komplexe Folge (a_n) gegeben, so heißt (s_n) definiert durch

$$s_n := \sum_{j=0}^{n} a_j = a_0 + a_1 + \cdots + a_{n-1} + a_n$$

Folge der *Partialsummen*. Diese Folge wird im Folgenden auch durch $\sum a_n$ abgekürzt und als (unendliche) *Reihe* bezeichnet. Im Fall der Konvergenz oder der bestimmten Divergenz der Folge $(s_n) = \sum a_n$ definieren wir

$$\sum_{j=0}^{\infty} a_j := \lim_{n\to\infty} s_n.$$

Aufgrund der besonderen Struktur der Folge (s_n) kann man bereits aus dem Verhalten der Folge (a_n) Rückschlüsse auf die Konvergenz der Reihe $\sum a_n$ ziehen. Dieses ist das zentrale Anliegen dieses Abschnitts.

Lemma 2.40 (Nullfolgenkriterium)**.** *Wenn $\sum a_n$ konvergiert, so gilt* $\lim\limits_{n\to\infty} a_n = 0$.

Beweis. Aus $\lim\limits_{n\to\infty} s_n = s$ folgt wegen $a_n = s_n - s_{n-1}$ sofort $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} s_n - \lim\limits_{n\to\infty} s_{n-1} = s - s = 0$. ☐

Bemerkung 2.41. Die Umkehrung dieses Satzes gilt nicht, wie die harmonische Reihe zeigt. Mit Hilfe des Nullfolgenkriteriums kann man also nicht die Konvergenz einer Reihe beweisen, sondern nur ihre Divergenz feststellen, wenn (a_n) nicht gegen 0 konvergiert.

Als direkte Folgerung aus Satz 2.29 erhalten wir:

Satz 2.42 (Cauchykriterium für Reihen). *Eine (reelle oder komplexe) Reihe* $\sum a_n$ *konvergiert dann und nur dann, wenn sie eine Cauchyfolge ist, d. h.*

$$\forall \varepsilon > 0 \, \exists N \in \mathbb{N} \, \forall n > m \geq N : \; |s_n - s_m| = \left| \sum_{j=m+1}^{n} a_j \right| < \varepsilon.$$

Satz 2.43 (Konvergenzkriterium für Reihen mit nur nicht positiven oder nur nicht negativen Summanden). *Ist* (a_n) *eine nicht negative (bzw. nicht positive) reelle Folge, d. h.* $a_n \geq 0$ $(a_n \leq 0)$ *für alle* $n \in \mathbb{N}$, *dann ist* $s_n = \sum_{j=0}^{n} a_n$ *monoton wachsend (bzw. monoton fallend) und es gilt:*

- *Ist* s_n *beschränkt, so konvergiert* $\sum a_n$.
- *Ist* s_n *unbeschränkt, so divergiert* $\sum a_n$ *bestimmt gegen* $+\infty$ *(bzw. gegen* $-\infty$*).*

Beispiel 2.44 (Nützliche Reihen). (i) Sei $z \in \mathbb{C}$ mit $|z| < 1$, dann gilt für die *geometrische Reihe*

$$\sum_{j=0}^{\infty} z^j = 1 + z + z^2 + z^3 + \cdots = \frac{1}{1-z}.$$

Für $|z| \geq 1$ ist die Folge $a_n = z^n$ keine Nullfolge und somit $\sum z^n$ divergent.

(ii) Es gilt

$$\sum \left(\frac{1}{n}\right)^r \text{ ist } \begin{cases} \text{konvergent,} & \text{für } r > 1 \\ \text{bestimmt divergent gegen } +\infty, & \text{für } r \leq 1. \end{cases}.$$

Insbesondere divergiert die *harmonische Reihe* $(r = 1)$ bestimmt:

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

Beweis. ad (i): Nach Beispiel 2.1 gilt

$$s_n = \sum_{j=0}^{n} z^j = \frac{1 - z^{n+1}}{1-z}.$$

Da nach Beispiel 2.44 gilt $\lim_{n\to\infty} z^{n+1} = 0$ für $|z| < 1$ und damit

$$\sum_{j=0}^{\infty} z^j = \lim_{n\to\infty} s_n = \frac{1}{1-z}.$$

ad (ii): Sei zunächst $r > 1$. Nach Satz 2.43 genügt es zu zeigen, dass die Folge (s_n) beschränkt ist. Sei nun $k \in \mathbb{N}$, sodass $2^k - 1 \geq n$, dann gilt

$$\begin{aligned} s_n &\leq s_{2^k-1} = 1 + \underbrace{\left(\frac{1}{2^r} + \frac{1}{3^r}\right)}_{\leq 2\cdot\frac{1}{2^r}} + \cdots + \underbrace{\left(\frac{1}{2^{(k-1)r}} + \cdots + \frac{1}{(2^k-1)^r}\right)}_{\leq 2^{k-1}\frac{1}{2^{(k-1)r}}} \\ &\leq 1 + \frac{1}{2^{r-1}} + \cdots + \frac{1}{2^{(k-1)(r-1)}} \\ &\leq \sum_{j=0}^{\infty}\left(\frac{1}{2^{r-1}}\right)^j = \frac{1}{1-2^{1-r}}. \end{aligned}$$

In der letzten Gleichung verwendeten wir Teil (i), da für $r > 1$ gilt $\frac{1}{2^{r-1}} < 1$.

Sei $r \leq 1$, dann gilt für $k \in \mathbb{N}$ mit $n \geq 2^k$

$$\begin{aligned} s_n &= 1 + \frac{1}{2^r} + \cdots + \frac{1}{n^r} \geq 1 + \frac{1}{2} + \cdots + \frac{1}{n} \geq 1 + \frac{1}{2} + \cdots + \frac{1}{2^k} \\ &= 1 + \frac{1}{2} + \underbrace{\left(\frac{1}{3} + \frac{1}{4}\right)}_{\geq 2\frac{1}{4} = \frac{1}{2}} + \cdots + \underbrace{\left(\frac{1}{2^{k-1}+1} + \cdots + \frac{1}{2^k}\right)}_{\geq 2^{k-1}\frac{1}{2^k}} \\ &\geq 1 + \frac{k}{2}. \end{aligned}$$

Da die Folge $(1 + \frac{k}{2})$ nicht beschränkt ist, gilt gleiches auch für s_n.

□

Satz 2.45 (Leibniz[60]-Kriterium)**.** *Sei (a_n) eine reelle positive und monoton fallende Folge mit $\lim\limits_{n\to\infty} a_n = 0$. Dann gilt:*

(i) Die alternierende[61] Reihe $\sum(-1)^n a_n$ konvergiert.

(ii) $\left|\sum\limits_{j=0}^{\infty}(-1)^j a_j - \sum\limits_{j=0}^{n}(-1)^j a_j\right| \leq a_{n+1}$.

Beweis. Wir betrachten die beiden Teilfolgen s_{2k} und s_{2k+1} der Folge der Partialsummen $s_n := \sum\limits_{j=0}^{n}(-1)^j a_j$.

Da (a_n) monoton fallend ist, ist die Teilfolge s_{2k} wegen

$$s_{2k+2} - s_{2k} = a_{2k+2} - a_{2k+1} \leq 0$$

[60] Gottfried Wilhelm Leibniz (1646 - 1716), deutscher Gelehrter. Namensgeber des gleichnamigen Butterkekses

[61] d. h. die Summanden der Reihe haben abwechselnde Vorzeichen

ebenso monoton fallend. Die Teilfolge s_{2k+1} ist hingegen wegen

$$s_{2k+3} - s_{2k+1} = a_{2k+2} - a_{2k+3} \geq 0$$

monoton wachsend. Zudem gilt $s_{2k} - s_{2k-1} = a_{2k} > 0$ für $k \geq 1$. Für die Intervalle

$$I_k := [s_{2k-1}, s_{2k}]$$

folgt also:

- $I_{k+1} \subset I_k$,
- $|I_k| = s_{2k} - s_{2k-1} = a_{2k} \xrightarrow{k\to\infty} 0$.

Somit ist I_k eine Intervallschachtelung und es gilt:

$$\left\{\sum_{j=0}^{\infty}(-1)^j a_j\right\} := \bigcap_{k\in\mathbb{N}^*} I_k,$$

also (i).

Der Grenzwert $s := \sum_{j=0}^{\infty}(-1)^j a_j$ liegt immer zwischen s_n und s_{n+1}, da s_{2k} monoton gegen s fällt und s_{2k+1} monoton gegen s wächst. Damit gilt $|s - s_n| \leq |s_{n+1} - s_n| = a_{n+1}$. Damit gilt (ii). $\square$

Definition 2.46. Eine (reelle oder komplexe) Reihe $\sum_{j=0}^{\infty} a_n$ heißt *absolut konvergent*, wenn die Reihe $\sum_{j=0}^{\infty} |a_n|$ konvergiert.

Bemerkung 2.47. Nullfolgen (a_n), sodass $\sum a_n$ absolut konvergiert sind also Nullfolgen, welche in einem gewissen Sinne besonders „rasch“ gegen 0 konvergieren. Man spricht auch von *summierbaren (Null)folgen.* Kriterien für absolut konvergente Reihen, besagen also etwas über die „Geschwindigkeit“ der Konvergenz von (a_n) gegen 0.

Beispiel 2.48. Nach dem Leibniz-Kriterium ist die Reihe $\sum(-1)^n \frac{1}{n}$ konvergent, jedoch *nicht* absolut konvergent, da die harmonisch Reihe divergiert.

Proposition 2.49. *Eine absolut konvergente Reihe ist konvergent.*

Beweis. Konvergiert $\sum_{j=0}^{\infty} |a_n|$, so gilt nach der Dreiecksungleichung und dem Cauchykriterium für konvergente Reihen gibt es zu einem vorgegebenen $\varepsilon > 0$ ein $N \in \mathbb{N}$, sodass für alle $n > m \geq N$ gilt

$$\sum_{j=m+1}^{n} |a_n| < \varepsilon.$$

Mithilfe der Dreiecksungleichung erhalten wir sodann

$$\left|\sum_{j=m+1}^{n} a_n\right| \leq \sum_{j=m+1}^{n} |a_n| < \varepsilon.$$

Nach dem Cauchykriterium für konvergente Reihen konvergiert also auch $\sum_{j=0}^{\infty} a_n$. □

Satz 2.50 (Majorantenkriterium)**.** *Sei (c_n) eine positive reelle Folge und (a_n) eine komplexe Folge.*

(i) Wenn

- *$|a_n| \leq c_n$ für fast alle $n \in \mathbb{N}$ gilt und*
- *$\sum c_n$ konvergiert,*

dann konvergiert auch $\sum a_n$ absolut.

(ii) Falls zudem $|a_n| \leq c_n$ sogar für alle $n \in \mathbb{N}$ gilt, so folgt

$$\sum_{j=0}^{\infty} |a_j| \leq \sum_{j=0}^{\infty} c_j.$$

In diesem Fall heißt die Reihe $\sum_{j=0}^{\infty} c_j$ Majorante *von $\sum_{j=0}^{\infty} |a_j|$.*

Beweis. Sei $K \in \mathbb{N}$, sodass $|a_n| \leq c_n$ für alle $n \geq K$ gilt. Sei zudem $\varepsilon > 0$ vorgegeben. Dann gibt es nach dem Cauchykriterium für Reihen ein $N \geq K$, sodass für alle $n > m \geq N$ gilt

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

Damit erhalten wir für diese n, m sofort

$$\sum_{j=m+1}^{n} |a_j| \leq \sum_{j=m+1}^{n} c_j < \varepsilon.$$

Damit erfüllt auch $\sum |a_n|$ das Cauchykriterium für Reihen und ist somit konvergent.

Die zweite Aussage folgt mit Lemma 2.15 aus

$$\sum_{j=0}^{n} |a_j| \leq \sum_{j=0}^{n} c_j$$

für alle $n \in \mathbb{N}$. □

Aus dem Majorantenkriterium erhält man durch einen Widerspruchsbeweis unmittelbar:

Korollar 2.51 (Minorantenkriterium). *Seien (a_n) und (c_n) reelle Folgen. Wenn*

- *$a_n \geq c_n > 0$ für fast alle $n \in \mathbb{N}$ gilt und*
- *$\sum c_n$ divergiert,*

dann divergiert auch $\sum a_n$.

Satz 2.52 (Quotientenkriterium, d'Alembert[62]–Kriterium). *Sei (a_n) eine komplexe Folge, wobei $a_n \neq 0$ für fast alle $n \in \mathbb{N}$ angenommen wird. Wenn $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| =: q$ existiert, dann gilt:*

(i) Falls $q < 1$ konvergiert die Reihe $\sum a_n$ absolut.

(ii) Falls $q > 1$ divergiert die Reihe $\sum a_n$.

Beweis. Im Fall (i) sei $\kappa \in \mathbb{R}$ so gewählt, dass $0 \leq q < \kappa < 1$ gilt. Da $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = q < \kappa$, gibt es ein $N \in \mathbb{N}$, sodass für alle $n \geq N$ gilt $\left|\frac{a_{n+1}}{a_n}\right| < \kappa$, also $|a_{n+1}| < \kappa |a_n|$. Mittels Induktion erhält man so

$$|a_n| < \kappa^{n-N} |a_N| = \underbrace{\kappa^{-N} |a_N|}_{\text{Konstante}} \cdot \kappa^n.$$

Da die geometrische Reihe $\sum \kappa^n$ konvergiert, konvergiert auch $\kappa^{-N}|a_N| \sum \kappa^n$. Nach dem Majorantenkriterium ist damit auch $\sum |a_n|$ konvergent.

Im Fall (ii) wählen wir $\kappa \in \mathbb{R}$ mit $1 < \kappa < q$. Dann gibt es ein $N \in \mathbb{N}$, sodass für alle $n \geq N$ gilt $\left|\frac{a_{n+1}}{a_n}\right| > \kappa$. Da fast alle $a_n \neq 0$, ist die Folge $(|a_n|)$ schließlich streng monoton wachsend, konvergiert insbesondere nicht gegen Null. Damit ist $\sum a_n$ aber divergent. □

Bemerkung 2.53. Wenn $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = q = 1$, so erhält man aus dem Quotientenkriterium *keine* Aussage über die Konvergenz der Reihe $\sum a_n$. So liefert das Quotientenkriterium für die Reihe $\sum \left(\frac{1}{n}\right)^r$ immer $q = 1$, aber nur für $r > 1$ konvergiert auch $\sum a_n$ (siehe Bsp. 2.44).

Analog wie Satz 2.52 kann man folgende Variante des Quotientenkriteriums beweisen:

Satz 2.54 (Variante des Quotientenkriteriums). *Sei (a_n) eine komplexe Folge, wobei $a_n \neq 0$ für fast alle $n \in \mathbb{N}$ angenommen wird. Wenn es ein $0 < q < 1$ gibt, sodass $\left|\frac{a_{n+1}}{a_n}\right| \leq q$ für fast alle $n \in \mathbb{N}$ gilt, dann konvergiert die Reihe $\sum a_n$ absolut.*

Satz 2.55 (Wurzelkriterium (von Cauchy)). *Sei (a_n) eine komplexe Folge und $L := \limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}$.*

[62] d'Alembert, eigentlich Jean-Baptiste le Rond (1717–1783), französische Gelehrter

(i) *Wenn $L < 1$, dann konvergiert $\sum a_n$ absolut.*

(ii) *Wenn $L > 1$, dann divergiert $\sum a_n$.*

Beweis. Im Fall (i) sei $0 \leq L < \kappa < 1$. Dann gibt es ein $N \in \mathbb{N}$, sodass für alle $n \geq N$ gilt $\sqrt[n]{|a_n|} \leq \kappa$ und somit $|a_n| \leq \kappa^n$. Da die geometrische Reihe $\sum \kappa^n$ mit $0 < \kappa < 1$ konvergiert, ist auch $\sum a_n$ nach dem Majorantenkriterium konvergent.

Im Fall (ii) gilt $|a_n| \geq 1$ für unendlich viele $n \in \mathbb{N}$. Damit konvergiert a_n aber nicht gegen 0 und die Reihe $\sum a_n$ ist divergent. $\square$

Bemerkung 2.56. Wenn $\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|} = L = 1$, so erhält man aus dem Wurzelkriterium *keine* Aussage über die Konvergenz der Reihe $\sum a_n$. Das Wurzelkriterium ergibt für die Reihe $\sum \left(\frac{1}{n}\right)^r$ immer $L = 1$, aber nur für $r > 1$ konvergiert auch $\sum a_n$ (siehe Bsp. 2.44).

Aufgabe 2.57 (Vergleich zwischen Wurzel- und Quotientenkriterium)**.** Sei (a_n) eine komplexe Folge mit $a_n \neq 0$ für fast alle $n \in \mathbb{N}$, sodass $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| =: q$ existiert.
Zeigen Sie: $L := \limsup\limits_{n\to\infty} \sqrt[n]{|a_n|} \leq q$.

Bemerkung 2.58 (Vergleich zwischen Wurzel- und Quotientenkriterium)**.** Aufgabe 2.57 zeigt, dass wenn für eine Reihe $\sum a_n$ die Konvergenz mit Hilfe des Quotientenkriteriums nachgewiesen werden kann, so kann man das auch mit Hilfe des Wurzelkriteriums. Umgekehrt gilt das nicht. Ein extremes Beispiel ist z. B.

$$a_n = \begin{cases} \frac{1}{2^n}, & \text{falls } n \text{ gerade} \\ \frac{1}{3^n}, & \text{falls } n \text{ ungerade} \end{cases}.$$

Aus dem Wurzelkriterium $\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}$ folgt für dieses Beispiel die Konvergenz, während das Quotientenkriterium nicht angewandt werden kann, da die Folge

$$\left|\frac{a_{n+1}}{a_n}\right| = \begin{cases} \frac{1}{3}\left(\frac{2}{3}\right)^n, & \text{falls } n \text{ gerade} \\ \frac{1}{2}\left(\frac{3}{2}\right)^n, & \text{falls } n \text{ ungerade} \end{cases}$$

nicht einmal beschränkt ist.

Aus Satz 2.55 folgt:

Korollar 2.59 (Variante des Wurzelkriteriums)**.** *Sei (a_n) eine komplexe Folge. Wenn es ein $0 < L < 1$ gibt, sodass $\sqrt[n]{|a_n|} \leq L$ für fast alle $n \in \mathbb{N}$ gilt, dann konvergiert die Reihe $\sum a_n$ absolut.*

Bemerkung 2.60 (Integralvergleichskriterium)**.** Wir werden später noch ein weiteres Konvergenzkriterium, das Integralvergleichskriterium kennenlernen, siehe Satz 5.52.

Definition 2.61. Sei $\sum a_n$ eine Reihe und $\tau : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung, dann heißt die Reihe $\sum a_{\tau(n)}$ *Umordnung* von $\sum a_n$.

Satz 2.62 (Umordnungssatz für absolut konvergente Reihen)**.** *Ist $\sum a_n$ eine absolut konvergente reelle oder komplexe Reihe, so konvergiert auch jede Umordnung $\sum a_{\tau(n)}$ und es gilt*

$$\sum_{j=0}^{\infty} a_j = \sum_{j=0}^{\infty} a_{\tau(j)}.$$

Beweis. Sei also $\tau : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Da $\sum |a_n|$ konvergiert, gibt es für ein vorgegebenes $\varepsilon > 0$ ein $N \in \mathbb{N}$, sodass

$$\left| \sum_{j=0}^{\infty} |a_j| - \sum_{j=0}^{N-1} |a_j| \right| = \sum_{j=N}^{\infty} |a_j| < \frac{\varepsilon}{2}.$$

Somit erhalten wir auch

$$\left| \sum_{j=0}^{\infty} a_j - \sum_{j=0}^{N-1} a_j \right| = \left| \sum_{j=N}^{\infty} a_j \right| \leq \sum_{j=N}^{\infty} |a_j| < \frac{\varepsilon}{2}.$$

Sei nun $M \in \mathbb{N}$ so groß gewählt, dass

$$\{0, \ldots, N-1\} \subset \{\tau(0), \tau(1), \ldots, \tau(M)\}$$

gilt. Damit erhalten wir für $m \geq M$:

$$\begin{aligned}
\left| \sum_{j=0}^{m} a_{\tau(j)} - \sum_{j=0}^{\infty} a_j \right| &\leq \left| \sum_{j=0}^{m} a_{\tau(j)} - \sum_{j=0}^{N-1} a_j \right| + \underbrace{\left| \sum_{j=0}^{N-1} a_j - \sum_{j=0}^{\infty} a_j \right|}_{< \frac{\varepsilon}{2}} \\
&< \left| \sum_{j \in J} a_j \right| + \frac{\varepsilon}{2} \quad \text{mit} \quad J := \{\tau(0), \ldots, \tau(m)\} \setminus \{0, \ldots N-1\} \\
&\leq \sum_{j \in J} |a_j| + \frac{\varepsilon}{2} \leq \sum_{j=N}^{\infty} |a_j| + \frac{\varepsilon}{2} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

□

Im Umordnungssatz (Satz 2.62) ist die Voraussetzung absolut konvergent essentiell. In der Tat gilt:

Aufgabe 2.63. Sei $\sum a_n$ eine konvergente aber nicht absolut konvergente *reelle* Reihe und $x \in \mathbb{R}$. Zeigen Sie: Es gibt eine bijektive Abbildung $\tau : \mathbb{N} \to \mathbb{N}$ für die $\sum_{j=0}^{\infty} a_{\tau(j)} = x$ gilt.

Während eine Vertauschung endlich vieler Summanden einer konvergenten Reihe den Grenzwert natürlich nicht ändern kann, ändert sich der Grenzwert einer konvergenten reellen Reihe bei Vertauschung von unendlich vielen Summanden nur dann nicht, wenn die Reihe absolut konvergiert. Eine Art „unendliches Kommutativgesetz“ gilt also nur für absolut konvergente Reihen.

Den folgenden Satz schieben wir ein, da wir ihn später beim Beweis der Funktionalgleichung der Exponentialfunktion benötigen. Er besagt, wie man Reihen miteinander multiplizieren kann:

Satz 2.64 (Satz von Mertens[63] / Cauchy-Produkt). *Sei $\sum a_n$ eine absolut konvergente komplexe Reihe und $\sum b_n$ eine konvergente komplexe Reihe, sowie $c_n := \sum\limits_{j=0}^{n} a_j b_{n-j}$. Dann konvergiert auch $\sum c_n$ und es gilt zudem*

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n\right) \cdot \left(\sum_{j=0}^{\infty} b_j\right).$$

Beweis. [64] Sei $a := \sum\limits_{n=0}^{\infty} a_n$, $b := \sum\limits_{n=0}^{\infty} b_n$ sowie $s_n^a := \sum\limits_{j=0}^{n} a_j$, $s_n^b := \sum\limits_{j=0}^{n} b_j$ und $s_n^c := \sum\limits_{j=0}^{n} c_j$. Dann gilt

$$\begin{aligned}
ab - s_n^c &= ab - \sum_{j=0}^{n} c_j &&= ab - \sum_{j=0}^{n}\sum_{k=0}^{j} a_k b_{j-k} \\
&= ab - \sum_{j=0}^{n} a_j \sum_{k=0}^{n-j} b_k &&= ab - \sum_{j=0}^{n} a_j s_{n-j}^b \\
&= (a - s_n^a)b + s_n^a b - \sum_{j=0}^{n} a_j s_{n-j}^b &&= (a - s_n^a)b + \sum_{j=0}^{n} a_j b - \sum_{j=0}^{n} a_j s_{n-j}^b \\
&= (a - s_n^a)b + \sum_{j=0}^{n} a_j (b - s_{n-j}^b)
\end{aligned} \tag{2.2}$$

Wir betrachten

$$\sum_{j=0}^{n} a_j (b - b_{n-j}) = \sum_{j=0}^{\left\lfloor \frac{n}{2} \right\rfloor} a_j (b - s_{n-j}^b) + \sum_{j=\left\lfloor \frac{n}{2} \right\rfloor + 1}^{n} a_j (b - s_{n-j}^b),$$

[63] Franz Carl Josef Mertens (1840 - 1927), deutsch und österreichischer Mathematiker

[64] siehe `http://de.wikipedia.org/wiki/Satz_von_Mertens_(Cauchy-Produkt)` Fassung vom 21.07.2012, 14:09 Uhr

dann gilt

$$\begin{aligned}
\left|\sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} a_j(b - s^b_{n-j})\right| &\leq \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} |a_j| \cdot |b - s^b_{n-j}| \\
&\leq \underbrace{\max\left\{|b - s^b_k| : n - \left\lfloor \frac{n}{2} \right\rfloor \leq k \leq n\right\}}_{\xrightarrow{n\to\infty} 0} \cdot \underbrace{\sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} |a_j|}_{\leq \sum_{j=0}^{\infty} |a_j| = \text{konst.}} \\
&\xrightarrow{n\to\infty} 0.
\end{aligned}$$

und

$$\begin{aligned}
\left|\sum_{j=\lfloor \frac{n}{2} \rfloor + 1}^{n} a_j(b - s^b_{n-j})\right| &\leq \sum_{j=\lfloor \frac{n}{2} \rfloor + 1}^{n} |a_j| \cdot \underbrace{|b - s^b_{n-j}|}_{\leq R} \\
&\leq R \cdot \underbrace{\sum_{j=\lfloor \frac{n}{2} \rfloor + 1}^{n} |a_j|}_{\xrightarrow{n\to\infty} 0 \text{ (Cauchykrit.)}} \\
&\xrightarrow{n\to\infty} 0.
\end{aligned}$$

und somit $\lim\limits_{n\to\infty}\left(\sum\limits_{j=0}^{n} a_j(b - b_{n-j})\right) = 0$. Da zudem auch $\lim\limits_{n\to\infty}((a - s^a_n)b) = 0$, erhalten wir $\lim\limits_{n\to\infty}(ab - s^c_n) = 0$ also $\lim\limits_{n\to\infty} s^c_n = ab$. Dieses ist aber die Aussage. □

Bemerkung 2.65. Wenn im Satz 2.64 die Reihe $\sum b_n$ auch absolut konvergiert, so konvergiert auch $\sum c_n$ absolut (siehe z. B. [Fo1, § 8]).

Wenn in der beim Satz von Mertens verwendeten Notation die Reihen $\sum a_n$ und $\sum b_n$ lediglich konvergieren, so muss $\sum c_n$ nicht konvergent sein, wie das Beispiel $a_n = b_n = \frac{(-1)^n}{\sqrt{n+1}}$ zeigt (Übung). Falls jedoch zudem auch $\sum c_n$ konvergiert, so gilt $\sum\limits_{n=0}^{\infty} c_n = \left(\sum\limits_{n=0}^{\infty} a_n\right)\left(\sum\limits_{n=0}^{\infty} b_n\right)$ nach einem Satz von Abel[65].

2.8. Potenzreihen

Wir orientieren uns an [K1, Abschn. 6.4].

[65]Niels Henrik Abel (1802 - 1829), norwegischer Mathematiker.

Sei (a_n) eine komplexe Folge so heißt die Folge der Partialsummen

$$s_n(z) = \sum_{j=0}^{n} a_j z^j,$$

welche wir auch mit $\sum a_n z^n$ bezeichnen, *Potenzreihe*.

Wir nehmen an, dass die Variable z eine komplexe Zahl ist. Die Frage ob $\sum a_n z^n$ konvergiert hängt natürlich stark vom Wert von z ab. Wir möchten nun die Menge

$$\mathrm{Konv}(a_n) := \left\{ z \in \mathbb{C} : \sum a_n z^n \text{ konvergent} \right\},$$

genannt *Konvergenzbereich* der Potenzreihe $\sum a_n z^n$ näher untersuchen. Es wird sich herausstellen, dass $\mathrm{Konv}(a_n)$ eine „Kreisscheibe“um 0 ist, wobei einige Randpunkte von $\mathrm{Konv}(a_n)$ zu $\mathrm{Konv}(a_n)$ gehören können und andere Randpunkte nicht.

Zunächst ist klar, dass $0 \in \mathrm{Konv}(a_n)$ und damit $\mathrm{Konv}(a_n) \neq \emptyset$.

Lemma 2.66. *Sei $z_o \in \mathrm{Konv}(a_n)$, dann konvergiert $\sum a_n z^n$ für alle $z \in \mathbb{C}$ mit $|z| < |z_o|$ absolut.*

Beweis. Da $\sum a_n z_o^n$ konvergiert gilt $\lim\limits_{n\to\infty} a_n z_o^n = 0$. Damit ist die Folge $(a_n z_o^n)$ beschränkt, d. h. es gibt ein $R > 0$ mit $|a_n z_o^n| < R$. Damit folgt

$$|a_n z^n| = \left| a_n z_o^n \left(\frac{z}{z_o} \right)^n \right| = |a_n z_o^n| \cdot \left(\frac{|z|}{|z_o|} \right)^n \leq R q^n$$

mit $q = \frac{|z|}{|z_o|} < 1$. Da die geometrische Reihe $\sum q^n$ mit $0 \leq q < 1$ konvergiert, ist $\sum (R q^n)$ eine konvergent Majorante von $\sum a_n z^n$. Damit konvergiert $\sum a_n z^n$ absolut nach Satz 2.50. □

Definition 2.67. Sei (a_n) eine komplexe Folge, dann heißt

$$R := R(a_n) := \sup\{ r \in \mathbb{R} : \sum a_n r^n \text{ konvergiert} \}$$

Konvergenzradius der Potenzreihe $\sum a_n z^n$.

Da $\sum a_n 0^n = a_0$ gilt $R \geq 0$. Der Begriff Konvergenz*radius* wird durch folgenden Satz veranschaulicht:

Satz 2.68. *Die Potenzreihe $\sum a_n z^n$ ist*

- *für alle $z \in \mathbb{C}$ mit $|z| < R$ absolut konvergent und*
- *für alle $z \in \mathbb{C}$ mit $|z| > R$ divergent.*

Beweis. Sei $|z| < R$, dann gibt es nach der Definition des Supremums ein $r \in \mathbb{R}$ mit $|z| < r \leq R$, sodass $\sum a_n r^n$ konvergiert. Nach Lemma 2.66 konvergiert dann $\sum a_n z^n$ absolut.

Sei $|z| > R$ und $\sum a_n z^n$ konvergent dann konvergiert nach Lemma 2.66 $\sum a_n r^n$ für alle $r \in \mathbb{R}$ mit $0 \leq r < |z|$, also auch für alle $R < r < |z|$, im Widerspruch zur Definition von R. □

Bemerkung 2.69. Für $z \in \mathbb{C}$ mit $|z| = R$ kann $\sum a_n z^n$ konvergieren, aber auch divergieren (siehe unten).

Satz 2.70 (Formeln für den Konvergenzradius)**.** *Es gilt*

- *Euler*[66]*:* $R(a_n) = \frac{1}{q}$ *mit* $q := \lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$, *wenn dieser Grenzwert existiert,*
- *Cauchy-Hadamard*[67]*:* $R(a_n) = \frac{1}{L}$ *mit* $L := \limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}$,

wobei wir hier $\frac{1}{\infty} = 0$ *und* $\frac{1}{0} = \infty$ *verstehen.*

Beweis. Für $z \neq 0$ gilt mit $b_n := a_n z^n$

$$\hat{q} := \lim_{n\to\infty} \left|\frac{b_{n+1}}{b_n}\right| = |z| \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = |z|q \begin{cases} < 1, & \text{wenn} \quad |z| < \frac{1}{q} \\ > 1, & \text{wenn} \quad |z| > \frac{1}{q} \end{cases}$$

Nach dem Quotientenkriterium (Satz 2.52) konvergiert damit (absolut) $\sum a_n z^n$, wenn $|z| < \frac{1}{q}$, und $\sum a_n z^n$ divergiert, falls $|z| > \frac{1}{q}$. Damit gilt $R(a_n) = \frac{1}{q}$. Wenn $q = 0$, dann konvergiert $\sum a_n z^n$ für alle $z \in \mathbb{C}$. Wenn $q = \infty$, dann konvergiert $\sum a_n z^n$ nur für $z = 0$.

Der Beweis der Formel von Cauchy-Hadamard geht analog mit Hilfe des Wurzelkriteriums, Satz 2.55 (siehe z. B. [K1, Abschn. 6.4]). □

Beispiel 2.71. (i) Der Konvergenzradius von $\sum z^n$ ist $R = 1$. Für alle $z \in \mathbb{C}$ mit $|z| = 1$ divergiert jedoch $\sum z^n$.

(ii) Der Konvergenzradius von $\sum \frac{1}{n} z^n$ ist $R = 1$. Für $z = 1$ divergiert die Potenzreihe als harmonische Reihe. Für $z = -1$ konvergiert die Reihe (als alternierende harmonische Reihe) nach dem Leibnizkriterium, Satz 2.45.

(iii) Der Konvergenzradius von $\sum \frac{1}{n^2} z^n$ ist $R = 1$. Für alle $z \in \mathbb{C}$ mit $|z| = 1$ konvergiert die Potenzreihe absolut, weil $\sum \frac{1}{n^2}$ konvergiert.

Lemma 2.72 (Restgliedabschätzung)**.** *Sei* $\sum a_n z^n$ *eine Potenzreihe mit Konvergenzradius* $R > 0$. *Sei weiterhin* $0 < r < R$ *und* $n \in \mathbb{N}^*$, *dann gibt es eine Konstante* $c_{r,n} > 0$, *sodass gilt:*

$$\left|\sum_{j=n}^{\infty} a_j z^j\right| \leq c_{r,n}|z|^n \quad \textit{für alle } |z| \leq r.$$

Beweis. Da $\sum\limits_{j=0}^{\infty} |a_j r^j|$ existiert, existiert auch $\sum\limits_{j=n}^{\infty} |a_j| r^j = r^n \sum\limits_{j=n}^{\infty} |a_j| r^{j-n}$ und damit auch $c_{r,n} := \sum\limits_{j=n}^{\infty} |a_j| r^{j-n}$. Wir erhalten so für $|z| \leq r$:

$$\left|\sum_{j=n}^{\infty} a_j z^j\right| \leq \sum_{j=n}^{\infty} |a_j| \cdot |z|^j \leq |z|^n \cdot \sum_{j=n}^{\infty} |a_j| \cdot r^{j-n} = c_{r,n}|z|^n.$$

□

[66] Leonhard Euler (1707 - 1783), schweizerischer Mathematiker

[67] Jacques Salomon Hadamard (1865 - 1963), französischer Mathematiker

Proposition 2.73. *Sei* (a_n) *eine komplexe Folge, welche nicht konstant Null ist, sodass die Potenzreihe* $\sum a_n z^n$ *Konvergenzradius* $R > 0$ *hat. Wir setzen* $f : D_R(0) \to \mathbb{C},\ z \mapsto f(z) := \sum\limits_{n=0}^{\infty} a_n z^n$*. Dann gibt es ein* $0 < \delta < R$*, sodass* $f|_{D_\delta(0)}$ *nur endlich viele Nullstellen*[68] *hat.*

Beweis. Sei $N \in \mathbb{N}$, sodass $a_N \neq 0$ und $a_n = 0$ für $n < N$. Sei $0 < r < R$ gewählt, dann gibt es ein $c_{r,N+1} > 0$ sodass

$$\left|f(z) - a_N z^N\right| = \left|\sum_{j=N+1}^{\infty} a_j z^j\right| \leq c_{r,N+1}|z|^{N+1} \quad \text{für alle } |z| \leq r. \tag{2.3}$$

Angenommen die Aussage ist falsch, dann gibt es in jeder Kreisscheibe um 0 mit Radius $\frac{r}{m}$, $m \in \mathbb{N}^*$, unendlich viele Nullstellen von f und daher eine Folge (z_m) mit $f(z_m) = 0$, $z_m \neq 0$ und $\lim\limits_{m\to\infty} z_m = 0$. Mit (2.3) erhalten wir dann $|a_N| \cdot |z_m|^N \leq c_{r,N+1}|z_m|^{N+1}$ und damit

$$|a_N| \leq c_{r,N+1}|z_m| \xrightarrow{m\to\infty} 0,$$

also $a_N = 0$, im Widerspruch zu $a_N \neq 0$. □

Satz 2.74 (Identitätssatz für Potenzreihen)**.** *Sei* $\sum a_n z^n$ *eine Potenzreihe mit Konvergenzradius* $R_a > 0$ *und* $\sum b_n z^n$ *eine Potenzreihe mit Konvergenzradius* $R_b > 0$*. Sei* $R := \min(R_a, R_b)$ *und* $f_a : D_R(0) \to \mathbb{C},\ z \mapsto f_a(z) := \sum\limits_{n=0}^{\infty} a_n z^n$ *sowie* $f_b : D_R(0) \to \mathbb{C},\ z \mapsto f_b(z) := \sum\limits_{n=0}^{\infty} b_n z^n$*. Wenn es eine Folge* (z_m) *mit* $\lim\limits_{m\to\infty} z_m = 0$ *gibt, sodass* $f_a(z_m) = f_b(z_m)$ *und* $z_m \neq 0$ *für alle* $m \in \mathbb{N}$ *gilt, dann gilt*

$$a_j = b_j \quad \textit{für alle } j \in \mathbb{N}$$

und somit $f_a = f_b$*.*

Beweis. Man betrachte $f - g$ und wende Proposition 2.73 an (Übung). □

Ausblick

Konvergenz ist ein fundamentales Konzept der *mengentheoretischen Topologie*. Unsere Definition der Konvergenz von Folgen erweitert sich direkt auf Folgen in metrischen Räumen. Die Eigenschaft eines metrischen Raumes, dass jede Cauchy-Folge in ihm konvergiert, wird als (topologische) Vollständigkeit bezeichnet.

Für allgemeinere topologische Räume wird der Begriff der Folge zum Begriff eines Netzes verallgemeinert.

Eine gute Einführung in die mengentheoretische Topologie finden Sie in dem Buch [Q], welches eine Gruppe Mathematiker an der Universität Bochum (deren Campus im Stadtteil Querenburg liegt) unter dem Pseudonym „Boto von Querenburg“ herausgegeben hat. Der Vorname Boto steht hierbei für „Bochumer Topologen“.

[68]Sei $f : X \to \mathbb{C}$ eine Abbildung. Ein Element $x \in X$ heißt *Nullstelle* von f, wenn $f(x) = 0$ gilt.

3

Stetige Abbildungen und ihre Eigenschaften

In diesem Kapitel betrachten wir Abbildungen, deren Definitionbereich X eine Teilmenge der reellen oder komplexen Zahlen sind und welche „benachbarte Punkte auf benachbarte Punkte“ abbilden. Solche Abbildungen heißen stetig[69]. Aus der Schule mag Ihnen der Begriff der Stetigkeit für reelle Funktionen bereits bekannt sein. Dort wurde die Anschauung vermittelt, dass eine Funktion $f : [a, b] \to \mathbb{R}$ stetig ist, wenn „man den Graphen zeichnen kann ohne mit dem Stift abzusetzen“. Diese Anschauung werden wir formalisieren und wichtige Eigenschaften stetiger Abbildungen kennenlernen. Wir verwenden in diesem Abschnitt die Begriffe „Abbildung“ und „Funktion“synonym.

3.1. Stetige Funktionen

Wir richten uns in diesem Abschnitt nach [K1, Kap. 7.1].

In diesem Abschnitt bezeichne X immer eine Teilmenge von $\mathbb{R}$ oder $\mathbb{C}$.

Definition 3.1. [70] Wir nennen eine Abbildung (Funktion) $f : X \to \mathbb{C}$ *stetig in einem Punkt* $x_o \in X$, wenn gilt:

$$\forall \varepsilon > 0 \, \exists \delta = \delta_{\varepsilon, x_o} > 0 \, \forall x \in X : \quad (|x - x_o| < \delta) \implies (|f(x) - f(x_o)| < \varepsilon).$$

oder anders ausgedrückt (siehe Gleichung (2.1)):

$$\forall \varepsilon \in \mathbb{R}_{>0} \, \exists \delta = \delta_{\varepsilon, x_o} \in \mathbb{R}_{>0} : \ f(D_\delta(x_o) \cap X) \subseteq D_\varepsilon(f(x_o)).$$

Die Abbildung $f : X \to \mathbb{C}$ heißt *stetig auf* X, wenn f in *jedem* Punkt von X stetig ist, d. h.

$$\forall x_o \in X \quad \forall \varepsilon > 0 \quad \exists \delta = \delta_{\varepsilon, x_o} > 0 \quad \forall x \in X : \quad (|x - x_o| < \delta) \implies (|f(x) - f(x_o)| < \varepsilon)$$

bzw.

$$\forall x_o \in X \quad \forall \varepsilon \in \mathbb{R}_{>0} \quad \exists \delta = \delta_{\varepsilon, x_o} \in \mathbb{R}_{>0} : \quad f(D_\delta(x_o) \cap X) \subseteq D_\varepsilon(f(x_o)).$$

[69] Stetigkeit kann man sehr allgemein für Abbildungen zwischen topologischen Räumen definieren.
[70] nach K. Weierstraß

Bemerkung 3.2.

- Beachten Sie, dass in der Definition der Stetigkeit von f auf X δ sowohl vom vorgegebenen ε wie aber auch vom betrachteten Punkt x_o abhängt. Um das zu verdeutlichen haben wir $\delta = \delta_{\varepsilon,x_o}$ geschrieben.
- Ob eine Funktion stetig ist, hängt auch von Ihrem Definitionsbereich ab, so ist z. B. *jede* Funktion $f : \mathbb{N} \to \mathbb{C}$ stetig (Übung). Daher muss der Definitionsbereich bei einer Aussage über die Stetigkeit einer Funktion immer unbedingt spezifiziert werden.

Beispiel 3.3.

(i) Konstante Funktionen sind offenbar stetig.

(ii) Die Identitäts-Abbildung $f : X \to \mathbb{C},\ x \mapsto x$ ist stetig. Für gegebenes $\varepsilon > 0$ hat $\delta = \varepsilon$ die gewünschten Eigenschaft.

(iii) Die *Dirichlet*[71]*-Funktion*

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} 1, & \text{wenn} \quad x \in \mathbb{Q} \\ 0, & \text{wenn} \quad x \notin \mathbb{Q} \end{cases}$$

ist in keinem Punkt von $\mathbb{R}$ stetig.

(iv) Die *Thomae*[72]*-Funktion*

$$f : [0,1] \to \mathbb{R}, \quad x \mapsto \begin{cases} 0, & \text{wenn} \quad x \notin \mathbb{Q} \cap [0,1] \\ 1, & \text{wenn} \quad x = 0 \\ \frac{1}{q}, & \text{wenn} \quad x = \frac{p}{q} \text{ für teilerfremde } p, q \in \mathbb{N}^*,\ p < q \end{cases}$$

ist nur in den irrationalen Punkten von $\mathbb{R}$ stetig.

Definition 3.4. Eine Funktion $f : X \to \mathbb{C}$ heißt *folgenstetig in* $x_o \in X$, wenn gilt: Für *jede* Folge (x_n) mit $x_n \in X$ und $\lim_{n\to\infty} x_n = x_o$ gilt:

$$\lim_{n\to\infty} f(x_n) = f(x_o).$$

Die Funktion f heißt *folgenstetig auf* X, wenn f in jedem Punkt von X folgenstetig ist.

Satz 3.5.

(i) Ist $f : X \to \mathbb{C}$ stetig in $x_o \in X$, dann ist f in x_o auch folgenstetig.[73]

(ii) Ist $f : X \to \mathbb{C}$ folgenstetig in $x_o \in X$, dann ist f in x_o auch stetig.[74]

[71] Johann Peter Gustav Lejeune Dirichlet (1805 - 1859), deutscher Mathematiker

[72] Carl Johannes Thomae (1840 - 1921), deutscher Mathematiker

[73] Diese Aussage gilt sehr allgemein für stetige Abbildungen zwischen topologischen Räumen.

[74] Diese Aussage gilt auch für Abbildungen zwischen metrischen Räumen und anderen „gutartigen“ topologischen Räumen, jedoch nicht allgemein.

Beweis. Sei (x_n) eine Folge mit $x_n \in X$, welche gegen x_o konvergiert und $\varepsilon > 0$ vorgegeben. Da f in x_o stetig ist, gibt es ein $\delta > 0$, sodass $|f(x) - f(x_o)| < \varepsilon$ für alle $x \in D_\delta(x_o)$. Da $\lim\limits_{n\to\infty} x_n = x_o$ gibt es ein $N > 0$, sodass $x_n \in D_\delta(x_o)$ für alle $n \geq N$ gilt. Damit erhalten wir $|f(x_n) - f(x_o)| < \varepsilon$ für alle $n \geq N$. Da $\varepsilon > 0$ beliebig war folgt $\lim\limits_{n\to\infty} f(x_n) = f(x_o)$. Somit ist f in x_o folgenstetig und wir haben (i) gezeigt.

Um Aussage (ii) zu zeigen nehmen wir an, f sei aber nicht stetig in x_o, d. h.

$$\exists \varepsilon_o > 0 \; \forall \delta > 0 \; \exists x_\delta \in X \text{ mit } |x_\delta - x_o| < \delta \text{ und } |f(x_\delta) - f(x_o)| \geq \varepsilon.$$

Wählen wir speziell $\delta = \frac{1}{n}$, $n \in \mathbb{N}^*$, so gilt:

$$\exists \varepsilon_o > 0 \; \forall n \in \mathbb{N}^* \; \exists x_n \in X \text{ mit } |x_n - x_o| < \frac{1}{n} \text{ und } |f(x_n) - f(x_o)| \geq \varepsilon.$$

Die so erhaltene Folge (x_n) konvergiert gegen x_o, aber $f(x_n)$ konvergiert nicht gegen $f(x_o)$. Damit ist f auch nicht folgenstetig in x_o. □

Korollar 3.6. *Eine Abbildung $f : X \to \mathbb{C}$ ist genau dann stetig auf X, wenn f folgenstetig auf X ist.*

Definition 3.7. Eine Funktion $f : X \to \mathbb{C}$ heißt *Lipschitz*[75]*-stetig* auf X, wenn f die Abstände nur beschränkt verzerrt, d. h. genauer

$$\exists L \geq 0 \; \forall x, y \in X : \quad |f(x) - f(y)| \leq L|x - y|.$$

Die Konstante L heißt *Lipschitz-Konstante*[76] zu f.

Satz 3.8. *Eine auf X Lipschitz-stetige Funktion $f : X \to \mathbb{C}$ ist auf X auch stetig.*

Beweis. Sei $L > 0$ eine Lipschitz-Konstante zu f und ε vorgegeben, dann setze $\delta := \frac{\varepsilon}{L}$. □

Beispiel 3.9. Die Funktionen

(a) $\mathrm{abs} : \mathbb{C} \to \mathbb{R}, \quad z \mapsto |z|$,

(b) $\Re : \mathbb{C} \to \mathbb{R}, \quad z \mapsto \Re(z)$,

(c) $\Im : \mathbb{C} \to \mathbb{R}, \quad z \mapsto \Im(z)$,

(d) $\mathrm{Konj} : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \overline{z}$

sind Lipschitz-stetig auf $\mathbb{C}$ mit Lipschitz-Konstante $L = 1$ und somit auf ganz $\mathbb{C}$ stetig.

Aufgabe 3.10. Zeigen Sie: Die Funktion $w_k : \mathbb{R}_{\geq 0} \to \mathbb{R}$, $x \mapsto \sqrt[k]{x}$, $k \in \mathbb{N}_{\geq 2}$, ist zwar auf ganz $\mathbb{R}_{\geq 0}$ stetig, jedoch nicht auf $\mathbb{R}_{\geq 0}$ Lipschitz-stetig.

[75] Rudolf Otto Sigismund Lipschitz (1832 - 1903), deutscher Mathematiker

[76] Für $L < 1$ ist f ein Kontraktion im Sinne von Definition 2.31.

Lemma 3.11 (Rechenregeln für stetige Funktionen)**.** *Sei $X \subset \mathbb{C}$, $x_o \in X$, und $f, g : X \to \mathbb{C}$ Funktionen, welche in x_o stetig sind. Sei nun $Y \subset \mathbb{C}$ mit $f(X) \subseteq Y$ und $h : Y \to \mathbb{C}$ eine Funktion, die in $f(x_o)$ stetig ist. Dann gilt:*

(i) $f + g$ ist stetig in x_o[77]*.*

(ii) $f \cdot g$ ist stetig in x_o[78]*.*

(iii) Wenn $g(x_o) \neq 0$, dann gibt es ein $r > 0$, sodass $g(x) \neq 0$ für alle $x \in D_r(x_o) \cap X$. Die Abbildung $\frac{f}{g} : D_r(x_o) \cap X \to \mathbb{C}$ ist dann auch stetig in x_o[79]*.*

(iv) $h \circ f$ ist stetig in x_o.

Beweis. Sei x_n eine beliebige Folge in X mit $\lim\limits_{n\to\infty} x_n = x_o$. Da f und g stetig in x_o sind, gilt mit Satz 3.5 $\lim\limits_{n\to\infty} f(x_n) = f(x_o)$ wie auch $\lim\limits_{n\to\infty} g(x_n) = g(x_o)$.

Nach den üblichen Rechenregeln für Grenzwerte (Lemma 2.12) folgt damit auch

$$\lim_{n\to\infty} \big((f+g)(x_n)\big) = (f+g)(x_o) \quad \text{und} \quad \lim_{n\to\infty} \big((f\cdot g)(x_n)\big) = (f\cdot g)(x_o).$$

Da die gegen x_o konvergierende Folge (x_n) in X beliebig war, folgen die Aussagen (i) und (ii) unmittelbar aus Satz 3.5.

Sei nun $g(x_o) \neq 0$, dann gibt es für $\varepsilon := \frac{|g(x_o)|}{2}$ ein $\delta =: r > 0$, sodass für alle $x \in D_r(x_o) \cap X$ gilt: $|g(x) - g(x_o)| \leq \frac{|g(x_o)|}{2}$. Damit gilt

$$|g(x)| = |g(x_o) + g(x) - g(x_o)| \geq |g(x_o)| - |g(x) - g(x_o)| > \frac{|g(x_o)|}{2} > 0.$$

für alle $x \in D_r(x_o) \cap X$. Betrachten wir nun eine beliebige Folge in $D_r(x_o) \cap X$ mit $\lim\limits_{n\to\infty} x_n = x_o$. Dann gilt mit Lemma 2.12

$$\lim_{n\to\infty} \left(\frac{f}{g}(x_n)\right) = \frac{f}{g}(x_o).$$

Nach Satz 3.5 ist damit auch $\frac{f}{g}$ in x_o stetig. Das zeigt (iii).

Sei nun x_n wieder eine beliebige gegen x_o konvergierende Folge in X. Da f stetig ist, konvergiert die Folge $y_n := f(x_n)$ gegen $y_o := f(x_o)$. Da h stetig ist, konvergiert somit auch $h(y_n) = (h \circ f)(x_n)$ gegen $h(y_o) = (h \circ f)(x_o)$. Mit Satz 3.5 folgt also auch (iv). □

Beispiel 3.12. Da konstante Funktionen und die Identität auf ganz $\mathbb{C}$ stetig sind, sind auch alle *polynomialen Funktionen*

$$p : \mathbb{C} \to \mathbb{C}, z \mapsto \sum_{j=0}^{n} a_j z^j \quad \text{mit } a_0, \ldots, a_n \in \mathbb{C}$$

---

[77]punktweise Addition von komplexwertigen Funktionen: $f + g : X \to \mathbb{C}$, $x \mapsto f(x) + g(x)$.

[78]punktweise Multiplikation von komplexwertigen Funktionen: $f \cdot g : X \to \mathbb{C}$, $x \mapsto f(x)g(x)$.

[79]punktweise Multiplikation von komplexwertigen Funktionen: $\frac{f}{g} : X \to \mathbb{C}$, $x \mapsto \frac{f(x)}{g(x)}$, wobei g keine Nullstelle in X hat.

auf ganz $\mathbb{C}$ stetig.

Ebenso sind alle *rationalen Funktionen*, das sind Quotienten $\frac{p_1}{p_2}$ polynomialer Funktionen $p_1, p_2 : \mathbb{C} \to \mathbb{C}$ auf der Menge $\mathbb{C} \setminus \{z \in \mathbb{C} : p_2(z) = 0\}$ stetig.

3.2. Zwischenwertsatz für stetige reelle Funktionen

Der folgende Satz ist ein wichtiges Hilfsmittel der reellen Analysis. Für Werte:

Theorem 3.13 (Zwischenwertsatz)**.** *Jede stetige Funktion $f : [a, b] \to \mathbb{R}$ nimmt jeden Wert zwischen $f(a)$ und $f(b)$ an*[80]*. Genauer:*

Wenn $f(a) \le f(b)$, dann gilt:

$$\forall y \in [f(a), f(b)]\ \exists x \in [a, b] : \quad f(x) = y.$$

Wenn $f(a) \ge f(b)$, dann gilt:

$$\forall y \in [f(b), f(a)]\ \exists x \in [a, b] : \quad f(x) = y.$$

Beweis. Wir betrachten hier nur den Fall $f(a) \le f(b)$. Der andere Fall ist dem geneigten Leser gerne als Übungsaufgabe überlassen. Sei $y \in [f(a), f(b)]$. Wir definieren durch Intervallhalbierung rekursiv eine Intervallschachtelung wie folgt: $I_0 = [a_0, b_0] := [a, b]$. Sei $I_n = [a_n, b_n]$ bereits konstruiert, dann ist

$$I_{n+1} = [a_{n+1}, b_{n+1}] := \begin{cases} \left[\frac{a_n+b_n}{2},\ b_n\right], & \text{falls } f\left(\frac{a_n+b_n}{2}\right) \le y \le f(b_n) \\ \left[a_n,\ \frac{a_n+b_n}{2}\right], & \text{falls } f(a_n) \le y < f\left(\frac{a_n+b_n}{2}\right) \end{cases}$$

Sei $x \in \mathbb{R}$ der eindeutige Punkt mit

$$\{x\} = \bigcap_{n \in \mathbb{N}} I_n.$$

Dann gilt $\lim\limits_{n\to\infty} a_n = x = \lim\limits_{n\to\infty} b_n$. Wegen $f(a_n) \le y$ und $y \le f(b_n)$ folgt mit der Stetigkeit von f einerseits $f(x) = \lim\limits_{n\to\infty} f(a_n) \le y$ und andererseits $f(x) = \lim\limits_{n\to\infty} f(b_n) \ge y$, also $f(x) = y$. ☐

Aufgabe 3.14 (Staatsexamen 63910, Herbst 2010, Thema 1, Aufg. 1(d))**.** Seien $a, b \in \mathbb{R}$ mit $a < b$ und sei $f : [a, b] \to \mathbb{R}$ eine Funktion mit folgender Eigenschaft: Für alle $c, d \in \mathbb{R}$ mit $a \le c < d \le b$ gilt: $f|_{[c,d]}$ nimmt alle Werte zwischen $f(c)$ und $f(d)$ an.

Ist f dann stetig auf $[a, b]$? Beweisen Sie Ihre Antwort.

Korollar 3.15. *Jede stetige Funktion $f : [a, b] \to [a, b]$ hat einen Fixpunkt.*

Beweis. Die Funktion $g : [a, b] \to \mathbb{R},\ x \mapsto x - f(x)$ ist stetig und es gilt $g(a) \le 0$ und $g(b) \ge 0$, da nach Voraussetzung $f(a) \ge a$ und $f(b) \le b$. Damit gibt es einen Punkt $x \in [a, b]$ mit $g(x) = 0 = f(x) - x$, d. h. $f(x) = x$. ☐

[80] Dieser Satz ist ein Spezialfall eines Satzes aus der Topologie, welcher besagt, dass Bilder zusammenhängender Mengen unter stetigen Abbildungen wieder zusammenhängend sind.

Aufgabe 3.16. Seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a, b] \to [a, b]$ eine (nicht unbedingt stetige) Funktion. Entscheiden Sie, ob folgende Aussagen wahr oder falsch sind, und beweisen Sie die Richtigkeit Ihrer Antwort[81].

(a) Ist f monoton wachsend[82], so hat f einen Fixpunkt.

(b) Ist f monoton fallend[83], so hat f einen Fixpunkt.

3.3. Stetige Funktionen auf Kompakta

Wir folgen hier [K1, Abschn. 7.5] und [BF1, Abschn. 7.3].

Definition 3.17. Eine Teilmenge A von $\mathbb{C}$ bzw. $\mathbb{R}$ heißt *abgeschlossen*, wenn der Grenzwert *jeder* konvergenten Folge in A wieder in A liegt, d. h. ist (a_n) eine Folge mit $a_n \in A$ und $a = \lim\limits_{n\to\infty} a_n$, dann gilt $a \in A$.

Eine Teilmenge $K \subset \mathbb{C}$ bzw. $K \subset \mathbb{R}$, heißt *kompakt*, wenn K abgeschlossen und beschränkt[84] ist.

Beispiel 3.18. (i) Die Mengen $\mathbb{R}$ und $\mathbb{C}$ sind abgeschlossen.

(ii) Wenn $f_1, \dots, f_m : \mathbb{C} \to \mathbb{R}$ reellwertige stetige Funktionen sind und $k_1, \dots k_m \in \mathbb{R}$, dann ist die Menge $\{z \in \mathbb{C} : \ f_1(z) \le k_1, \dots f_m(z) \le k_m\}$ abgeschlossen (Übung).

(iii) kompakte Intervalle $[a, b] \subset \mathbb{R}$ sind kompakt.

Aufgabe 3.19. Ist $A \subset \mathbb{R}$ abgeschlossen, dann gilt $\sup(A) \in A$ und $\inf(A) \in A$.

Lemma 3.20. *Seien $A_1, ..., A_m$ abgeschlossene Teilmengen von $\mathbb{C}$ bzw. $\mathbb{R}$, dann ist auch $A := \bigcup\limits_{j=1}^{m} A_j$ abgeschlossen.*

Beweis. Sei (x_n) eine gegen x konvergierende Folge mit $x_n \in A$. Dann gibt es ein $k \in \{1, \dots m\}$ und eine Teilfolge (x_{n_j}) welche in A_k liegt. Damit gilt nach Lemma 2.24 $x = \lim\limits_{j\to\infty} x_{n_j}$. Da A_k abgeschlossen ist, gilt $x \in A_k$ und damit $x \in A$. □

Korollar 3.21. *Seien $K_1, ..., K_m$ kompakte Teilmengen von $\mathbb{C}$ bzw. $\mathbb{R}$, dann ist auch $K := \bigcup\limits_{j=1}^{m} K_j$ kompakt.*

Lemma 3.22. *Sei I eine Indexmenge und $\{A_j : \ j \in I\}$ eine Menge von abgeschlossenen Teilmengen $A_j \subset \mathbb{C}$. Dann ist auch $A := \bigcap\limits_{j\in I} A_j$ abgeschlossen.*

---

[81] Dankenswerterweise hat mein Kollege Ingo Blechschmidt mich darauf hingewiesen, dass Teil (a) ein Spezialfall des Fixpunktsatzes von Knaster-Tarski ist.

[82] $f : X \to \mathbb{R}$, $X \subset \mathbb{R}$, heißt *monoton wachsend,* wenn gilt: $\forall x, y \in D : \ (x < y) \implies (f(x) \le f(y))$.

[83] $f : X \to \mathbb{R}$, $X \subset \mathbb{R}$, heißt *monoton fallend,* wenn gilt: $\forall x, y \in D : \ (x < y) \implies (f(x) \ge f(y))$.

[84] Erinnerung: K ist beschränkt, wenn es ein $R > 0$ gibt, sodass für alle $x \in K$ gilt $|x| < R$.

Beweis. Sei (x_n) eine gegen x konvergierende Folge mit $x_n \in A$. Dann gilt $x_n \in A_j$ für alle $n \in \mathbb{N}$ und alle $j \in I$. Da A_j abgeschlossen ist, folgt $x \in A_j$ für alle $j \in I$ und damit $x \in A$. □

Korollar 3.23. *Sei I eine Indexmenge und $\{K_j : j \in I\}$ eine Menge von kompakten Teilmengen $K_j \subset \mathbb{C}$. Dann ist auch $K := \bigcap_{j \in I} K_j$ kompakt.*

Satz 3.24 (Bolzano-Weierstraß-Charakterisierung kompakter Mengen)**.** *Für eine Teilmenge $K \subset \mathbb{C}$ sind folgende Aussagen äquivalent:*

(i) K ist kompakt.

(ii) Jede Folge in K besitzt eine Teilfolge, welche gegen einen Punkt von K konvergiert.

Beweis.
- $(i) \implies (ii)$: Wenn K kompakt ist, dann ist K beschränkt und somit jede Folge in K beschränkt. Nach Theorem 2.27 hat jede Folge in K damit eine konvergente Teilfolge. Da K abgeschlossen ist, liegt der Grenzwert einer solchen Teilfolge ebenfalls in K.

- $(ii) \implies (i)$: Nun habe K die Eigenschaft (ii).

  Wäre K nicht beschränkt, so gäbe es eine Folge (x_n) in K mit $|x_n| \geq n$. Im Widerspruch zu (ii) hat eine solche Folge keine konvergente Teilfolge.

  Sei nun (x_n) eine gegen x konvergente Folge in K und (x_{n_j}) die in der Eigenschaft (ii) erwähnte konvergente Teilfolge von (x_n). Nach Lemma 2.24 gilt $\lim_{j \to \infty} x_{n_j} = x$. Nach Eigenschaft (ii) liegt damit $x \in K$. Somit ist K abgeschlossen.

□

Korollar 3.25. *Das Bild kompakter Mengen unter stetigen Abbildungen ist kompakt*[85]*, genauer: Sei $K \subset \mathbb{C}$ kompakt und $f : K \to \mathbb{C}$ stetig, dann ist $f(K)$ kompakt.*

Beweis. Sei (y_n) eine Folge in $f(K)$, dann gibt es eine Folge (x_n) in K mit $f(x_n) = y_n$ für alle $n \in \mathbb{N}$. Da K kompakt ist gibt es nach Satz 3.24 eine Teilfolge (x_{n_j}) von (x_n), welche gegen $x \in K$ konvergiert. Da f stetig ist, konvergiert aber auch $(y_{n_j}) := (f(x_{n_j}))$ gegen $y := f(x) \in f(K)$. Nach Satz 3.24 ist damit auch $f(K)$ kompakt. □

Theorem 3.26 (Weierstraßscher Satz vom Extremum)**.** *Jede reellwertige stetige Funktion auf einem Kompaktum nimmt ein Maximum und ein Minimum an*[86]*. Genauer: Sei $K \subset \mathbb{C}$ kompakt und $f : K \to \mathbb{R}$ eine stetige Funktion, dann gibt es $x_1, x_2 \in K$, sodass für alle $x \in K$ gilt*

$$f(x_1) \leq f(x) \leq f(x_2).$$

[85] Dieser Satz gilt allgemein.

[86] Diese Aussage gilt für stetige reellwertige Funktionen auf beliebigen kompakten topologischen Räumen

Beweis. Nach Korollar 3.25 ist $f(K)$ kompakt. Damit gilt $\sup(f(K)) \in f(K)$ und $\inf(f(K)) \in f(K)$ und es gibt $x_1, x_2 \in K$ mit $f(x_1) = \sup(f(K))$ und $f(x_2) = \inf(f(K))$. □

Satz 3.27. *Sei $K \subset \mathbb{C}$ kompakt und $f : K \to \mathbb{C}$ stetig und injektiv. Dann ist die Umkehrabbildung $f^{-1} : f(K) \to K$ wiederum stetig*[87].

Beweis. Angenommen f^{-1} wäre nicht stetig. Dann gibt es eine Folge (y_n) in $f(K)$, welche gegen y konvergiert[88] und für die Folge (x_n) mit $x_n := f^{-1}(y_n)$ nicht gegen $x := f^{-1}(y)$ konvergiert. Nach Satz 3.24, gibt es eine konvergente Teilfolge (x_{n_j}) von x_n. Wir können sogar annehmen, dass diese Teilfolge nicht gegen x konvergiert[89], d. h. $x' := \lim\limits_{j\to\infty} x_{n_j} \neq x$.

Da f stetig ist konvergiert die Teilfolge (y_{n_j}) mit $y_{n_j} := f(x_{n_j})$ gegen $f(x')$. Wegen der Injektivität von f ist $f(x') \neq f(x) = y$.

Andererseits konvergiert (y_{n_j}) gegen y, da (y_n) gegen y konvergiert (Lemma 2.24). Damit gilt also $f(x') = y$, ein Widerspruch. □

Bemerkung 3.28. (i) Die Voraussetzung, dass der Definitionsbereich K von f kompakt ist, darf in Satz 3.27 nicht einfach weggelassen werden, wie folgendes Beispiel zeigt: Die Funktion

$$f : [0, 2\pi[\to \mathbb{C}, \quad t \mapsto \cos(t) + i\sin(t).$$

ist injektiv und stetig, jedoch ist f^{-1} in 1 nicht stetig.

(ii) Ist I ein Intervall und $f : I \to \mathbb{R}$ stetig und injektiv. Dann ist f streng monoton[90] und $f^{-1} : f(I) \to \mathbb{R}$ ist wiederum stetig (Übung).

(iii) Ist I ein nichtleeres Intervall und $f : I \to \mathbb{R}$ streng monoton sowie stetig in x_o, dann ist die Umkehrfunktion stetig in $y_o := f(x_o)$. (Übung, siehe auch [K1, Abschn. 7.2]).

---

[87] Diese Aussage stimmt für alle injektiven stetigen Abbildungen von einem Kompaktum in einen topologischen Hausdorffraum.

[88] nach Korollar 3.25 ist $f(K)$ kompakt. Damit gilt $y \in f(K)$.

[89] Denn wenn jede konvergente Teilfolge der beschränkten Folge (x_n) gegen x konvergiert, dann gilt $\liminf\limits_{n\to\infty}(x_n) = \limsup\limits_{n\to\infty}(x_n) = x$. Damit würde (x_n) gegen x konvergieren (siehe Aufgabe 2.39.)

[90] Sei X eine Teilmenge von $\mathbb{R}$ und $f : X \to \mathbb{R}$ eine Funktion. Dann heißt f

- *monoton wachsend*, wenn $\forall x, y \in X : (x < y) \implies (f(x) \leq f(y))$,
- *streng monoton wachsend*, wenn $\forall x, y \in X : (x < y) \implies (f(x) < f(y))$,
- *monoton fallend*, wenn $\forall x, y \in X : (x < y) \implies (f(x) \geq f(y))$,
- *streng monoton fallend*, wenn $\forall x, y \in X : (x < y) \implies (f(x) > f(y))$,
- *monoton,* wenn f monoton wachsend oder fallend ist,
- *streng monoton,* wenn f streng monoton wachsend oder fallend ist.

(iv) Ist $X \subset \mathbb{R}$ *offen*[91] und $f : X \to \mathbb{R}$ eine stetige injektive Funktion, dann ist $f^{-1} : f(X) \to \mathbb{R}$ stetig (siehe z. B. [BF1, Abschn. 7.3]).

(v) Die Funktion $f :]-\infty, 0[\cup[2,\infty[\to \mathbb{R},\ x \mapsto \begin{cases} x, & \text{wenn} \quad x \in]-\infty, 0[\\ x-2, & \text{wenn} \quad x \in [2,\infty[\end{cases}$ ist stetig und bijektiv. Ihre Umkehrabbildung f^{-1} ist jedoch nicht stetig in 0 (Übung).

Definition 3.29. Sei $X \subset \mathbb{C}$. Eine Funktion $f : X \to \mathbb{C}$ heißt *gleichmäßig stetig auf* X, wenn gilt:

$$\forall \varepsilon > 0 \quad \exists \delta = \delta_\varepsilon \quad \forall x_o, x \in X : \quad (|x - x_o| < \delta) \implies (|f(x) - f(x_o)| < \varepsilon).$$

Bemerkung 3.30. Im Gegensatz zur einfachen Stetigkeit auf X gibt es bei vorgegebenem $\varepsilon > 0$ ein globales $\delta > 0$, d. h. ein δ, welches nur vom vorgegebenen ε *nicht* aber von x_o abhängt. Gleichmäßig stetige Funktion sind natürlich insbesondere stetig.

Lemma 3.31. *Ist $f : X \to \mathbb{C}$ Lipschitz-stetig, so ist f gleichmäßig stetig.*

Beweis. Sei $L > 0$ eine Lipschitz-Konstante für f und $\varepsilon > 0$ gegeben, so wähle $\delta = \frac{\varepsilon}{L}$ unabhängig von x_o. □

Beispiel 3.32. (1) Die Funktion $f : \mathbb{R} \to \mathbb{R},\ x \mapsto x^2$ ist auf $\mathbb{R}$ *nicht* gleichmäßig stetig (Übung).

(2) Die Funktion $w_k : [0,\infty[\to \mathbb{R},\ x \mapsto \sqrt[k]{x},\ k \in \mathbb{N}_{\geq 2}$, ist zwar auf $\mathbb{R}$ gleichmäßig stetig, jedoch nicht Lipschitz-stetig (Übung).

Theorem 3.33 (Satz von Heine[92]). *Sei $K \subset \mathbb{C}$ kompakt und $f : K \to \mathbb{C}$ stetig, dann ist f gleichmäßig stetig auf K.*[93]

Beweis. Angenommen f wäre nicht gleichmäßig stetig, dann gilt:

$$\exists \varepsilon' > 0 \quad \forall \delta > 0 \quad \exists x_\delta, x'_\delta \in K \text{ mit } |x_\delta - x'_\delta| < \delta \text{ und } |f(x_\delta) - f(x'_\delta)| \geq \varepsilon'.$$

Insbesondere gilt:

$$\exists \varepsilon' > 0 \quad \forall n \in \mathbb{N}^* \quad \exists x_n, x'_n \in K \text{ mit } |x_n - x'_n| < \frac{1}{n} \text{ und } |f(x_n) - f(x'_n)| \geq \varepsilon'.$$

Da K kompakt ist, hat die so erhaltene Folge (x_n) eine gegen ein $x \in K$ konvergierende Teilfolge (x_{n_j}) (Satz 3.24). Für die entsprechende Teilfolge (x'_{n_j}) der Folge (x'_n) gilt $\lim\limits_{j\to\infty} x'_{n_j} = x$, weil $|x_{n_j} - x'_{n_j}| < \frac{1}{n_j}$. Da f stetig ist gilt $\lim\limits_{j\to\infty} f(x_{n_j}) = f(x) = \lim\limits_{j\to\infty} f(x'_{n_j})$. Dieses widerspricht aber $|f(x_{n_j}) - f(x'_{n_j})| \geq \varepsilon'$. □

Bemerkung 3.34. Die Kompaktheitsvoraussetzung an den Definitionsbereich von f im Satz von Heine ist notwendig, denn die Funktion $f :]0,1] \to \mathbb{R},\ x \mapsto \frac{1}{x}$ ist stetig aber nicht gleichmäßig stetig auf $]0,1]$ (Übung).

---

[91] eine Menge $X \subset \mathbb{R}$ heißt *offen*, wenn gilt: $\forall x \in X\ \exists \varepsilon > 0 : (x - \varepsilon, x + \varepsilon) \subset X$.

[92] Heinrich Eduard Heine (1821 - 1881), deutscher Mathematiker

[93] Dieser Satz gilt allgemein für stetige Abbildungen auf einem kompakten metrischen Räumen mit Werten in metrischen Räumen (siehe Theorem 6.70).

3.4. Funktionenfolgen

Wir orientieren uns hier an [Fo1, § 21].

Seien X und Y nichtleere Mengen, dann bezeichnen wir mit

$$\mathrm{Abb}(X,Y) = Y^X := \{f : X \to Y\}$$

die Menge aller Abbildungen von X nach Y.

Definition 3.35. Sei X eine Menge. Eine *Folge komplexwertiger Funktionen auf* X ist eine Abbildung

$$\mathbb{N} \to \mathrm{Abb}(X,\mathbb{C}),\ n \mapsto f_n.$$

Oft schreiben wir auch kurz (f_n) für diese Folge von Funktionen.

Definition 3.36. Eine Folge komplexwertiger Funktionen (f_n) auf X *konvergiert auf* X *punktweise* gegen $f \in \mathrm{Abb}(X,\mathbb{C})$, notiert $f_n \xrightarrow[\text{pkt}]{n\to\infty} f$, wenn gilt:

$$\forall x \in X \quad \forall \varepsilon > 0 \quad \exists N = N_{x,\varepsilon} \in \mathbb{N} \quad \forall n \geq N : \quad |f_n(x) - f(x)| < \varepsilon.$$

Bemerkung 3.37. In der Definition der punktweisen Konvergenz hängt also $N = N_{x,\varepsilon}$ sowohl vom betrachteten Punkt $x \in X$ wie auch vom vorgegebenen $\varepsilon > 0$ ab.

Angenommen alle Funktionen f_n haben eine bestimmte gemeinsame Eigenschaft und die (f_n) konvergieren (punktweise) gegen f. Dann stellt sich die Frage ob f ebenfalls diese Eigenschaft hat, oder nicht. Hierbei stellt sich heraus, dass bei punktweiser Konvergenz Eigenschaften oftmals *nicht* an ihre Grenzfunktion vererbt werden.

Sind z. B. die Funktion f_n alle stetig, so muss die Grenzfunktion f bei punktweiser Konvergenz nicht stetig sein:

Betrachten wir die Folge stetiger Funktionen

$$f_n : [0,1] \to \ C, \quad x \mapsto x^n.$$

Dann gilt $f_n(1) = 1$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} f_n(x) = 0$ für alle $x \in [0,1)$. Damit konvergiert f_n punktweise gegen die Funktion

$$f : [0,1] \to \mathbb{C}, \quad x \mapsto \begin{cases} 0, & \text{falls} \quad 0 \leq x < 1 \\ 1, & \text{falls} \quad x = 1 \end{cases},$$

welche in $x = 1$ offenbar nicht stetig ist. Wir möchten nun einen Konvergenzbegriff für Folgen von Funktionen einführen, welcher Stetigkeit an die Grenzfunktion vererbt.

Definition 3.38. Eine Folge komplexwertiger Funktionen (f_n) auf X *konvergiert auf* X *gleichmäßig* gegen $f \in \mathrm{Abb}(X,\mathbb{C})$, notiert $f_n \xrightarrow[\text{glm}]{n\to\infty} f$, wenn gilt:

$$\forall \varepsilon > 0 \quad \exists N = N_\varepsilon \in \mathbb{N} \quad \forall n \geq N \quad \forall x \in X : \quad |f_n(x) - f(x)| < \varepsilon.$$

Bemerkung 3.39. Im Vergleich zur Definition der punktweisen Konvergenz haben wir bei der gleichmäßigen Konvergenz den Quantor $\forall x \in X$ von vorne nach hinten verschoben. Dieses ändert aber die Definition in einem wesentlichen Punkt. Unser N hängt bei gleichmäßiger Konvergenz zwar noch vom vorgegebenen $\varepsilon > 0$ ab, jedoch nicht mehr vom Punkt $x \in X$.

Theorem 3.40. *Sei $X \subset \mathbb{C}$ und (f_n) eine Folge komplexwertiger stetiger Funktionen auf X, welche auf X gleichmäßig gegen $f \in \mathrm{Abb}(X, \mathbb{C})$ konvergiert. Dann ist auch f stetig auf X.*

Beweis. Seien $x_o \in X$ und $\varepsilon > 0$ beliebig gewählt. Da (f_n) gleichmäßig gegen f konvergiert, gibt es ein $N = N_\varepsilon$, sodass für alle $x' \in X$ gilt:

$$|f_N(x') - f(x')| < \frac{\varepsilon}{3}.$$

Weil aber f_N stetig ist, gibt es ein $\delta = \delta_\varepsilon$, sodass für alle $x \in X$ mit $|x - x_o| < \delta$ gilt:

$$|f_N(x) - f_N(x_o)| < \frac{\varepsilon}{3}.$$

Mit der Dreiecksungleichung erhalten wir also für alle $x \in X$ mit $|x - x_o| < \delta$:

$$|f(x) - f(x_o)| \leq \underbrace{|f(x) - f_N(x)|}_{<\frac{\varepsilon}{3}} + \underbrace{|f_N(x) - f_N(x_o)|}_{<\frac{\varepsilon}{3}} + \underbrace{|f_N(x_o) - f(x_o)|}_{<\frac{\varepsilon}{3}} < \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, ist f stetig in x_o, und da $x_o \in X$ beliebig war, ist f stetig auf X. □

Definition 3.41. Sei $f \in \mathrm{Abb}(X, \mathbb{C})$, dann heißt

$$\|f\|_{\mathrm{sup}} := \sup\{|f(x)| : \ x \in X\} \in \mathbb{R} \cup \{\infty\}$$

die *Supremumsnorm*[94] von f.[95]

---

[94]Eine *Norm* auf einem Vektorraum V über $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ ist eine Abbildung $\| \dots \| : V \to \mathbb{R}_{\geq 0}$, $v \mapsto \|v\|$ mit folgenden drei Eigenschaften:

- $\|v\| = 0 \implies v = 0$,
- $\|\lambda v\| = |\lambda| \cdot \|v\|$ für alle $v \in V$ und alle $\lambda \in \mathbb{K}$,
- $\|v + w\| \leq \|v\| + \|w\|$ für alle $v, w \in V$.

Ein Vektorraum mit einer Norm heißt auch *normierter Vektorraum*. Durch eine Norm ist auf V eine Abstandsfunktion $d : V \times V \to \mathbb{R}_{\geq 0}$ definiert, indem man $d(v, w) := \|v - w\|$ setzt. Ein elementares Beispiel ist die Betragsfunktion auf dem Vektorraum $V = \mathbb{C}$ über $\mathbb{K}$. Unendlich dimensionale normierte Vektorräume werden vor allem in der *Funktionalanalysis* untersucht. Es sei Ihnen als Übung überlassen nachzuweisen, dass die Supremumsnorm eine Norm ist (auf welchem Vektorraum?).

[95]In der Literatur sind je nach genauem Kontext auch die Bezeichnungen $\|f\|_\infty$, $\|f\|_X$ oder $\|f\|_{\mathcal{C}^0(X,\mathbb{C})}$ für die Supremumsnorm gebräuchlich.

Beobachtung 3.42. *Sei* (f_n) *eine Folge in* $\mathrm{Abb}(X, \mathbb{C})$ *und* $f \in \mathrm{Abb}(X, \mathbb{C})$*. Dann sind folgende Aussagen äquivalent:*

- (f_n) *konvergiert auf* X *gleichmäßig gegen* f*.*
- $\lim\limits_{n \to \infty} \|f_n - f\|_{\sup} = 0.$

Ähnlich wie man aus Folgen komplexer Zahlen Reihen bilden kann, kann man das auch aus Folgen in $\mathrm{Abb}(X, \mathbb{C})$, da man ja komplexwertige Funktionen auf X punktweise addieren kann: Für eine Folge (f_n) in $\mathrm{Abb}(X, \mathbb{C})$ bezeichne $\sum f_n$ die Folge der Partialsummen

$$n \mapsto \sum_{j=0}^{n} f_j$$

in $\mathrm{Abb}(X, \mathbb{C})$ und $\sum\limits_{j=0}^{\infty} f_j$ die Grenzfunktion $x \mapsto \sum\limits_{j=0}^{\infty} f_j(x),\ x \in X$, falls $\sum f_n$ punktweise konvergiert.

Theorem 3.43 (Konvergenzkriterium von Weierstraß)**.** *Sei* (f_n) *eine Folge von Funktionen in* $\mathrm{Abb}(X, \mathbb{C})$ *sodass* $\sum \|f_n\|_{\sup}$ *konvergiert. Dann gilt:*

(i) Für alle $x \in X$ *konvergiert* $\sum f_n(x)$ *absolut.*

(ii) Die Folge $F_n := \sum\limits_{j=0}^{n} f_j$ *konvergiert auf* X *gleichmäßig gegen* $F := \sum\limits_{j=0}^{\infty} f_j$*.*

Beweis. Für $x \in X$ gilt $|f_n(x)| \leq \|f_n\|_{\sup}$. Da

$$\sum \|f_n\|_{\sup}$$

konvergiert, konvergiert nach dem Majorantenkriterium (Satz 2.50) $\sum f_n(x)$ absolut. Das zeigt (i).

Insbesondere konvergiert damit $\sum f_n$ punktweise gegen die Abbildung $F := \sum\limits_{j=0}^{\infty} f_j :$ $X \to \mathbb{C},\ x \mapsto \sum\limits_{j=0}^{\infty} f_n(x)$. Sei $\varepsilon > 0$ beliebig gewählt. Da $\sum \|f_n\|_{\sup}$ konvergiert, gibt es ein $N \in \mathbb{N}$, sodass für alle $n \geq N$ gilt:

$$\left| \sum_{j=0}^{\infty} \|f_j\|_{\sup} - \sum_{j=0}^{n} \|f_j\|_{\sup} \right| = \sum_{j=n+1}^{\infty} \|f_j\|_{\sup} < \varepsilon.$$

Somit erhalten wir für alle $x \in X$ und alle $n \geq N$:

$$|F_n(x) - F(x)| = \left| \sum_{j=n+1}^{\infty} f_n(x) \right| \leq \sum_{j=n+1}^{\infty} |f_n(x)| \leq \sum_{j=n+1}^{\infty} \|f_n(x)\|_{\sup} < \varepsilon$$

□

Potenzreihen entstehen durch Summation von Funktionenfolgen vom Typ

$$f_n(z) = a_n z^n.$$

Korollar 3.44. *Sei (a_n) eine Folge komplexer Zahlen und der Konvergenzradius R der Potenzreihe $\sum a_n z^n$ strikt positiv. Sei $0 < \rho < R$ und $\overline{D_\rho(0)} := \{z \in \mathbb{C} : |z| \leq \rho\}$. Dann konvergiert die Folge polynomialer Funktionen*

$$F_n : \overline{D_\rho(0)} \to \mathbb{C}, \quad z \mapsto \sum_{j=0}^{n} a_j z^j$$

auf $\overline{D_\rho(0)}$ gleichmäßig gegen

$$F : \overline{D_\rho(0)} \to \mathbb{C}, \quad z \mapsto \sum_{j=0}^{\infty} a_j z^j.$$

Insbesondere ist F auf ganz $\overline{D_\rho(0)}$ stetig.

Beweis. Wir schreiben kurz $f_n : \overline{D_\rho(0)} \to \mathbb{C}$, $z \mapsto a_n z^n$. Sei $z_o \in D_R(0)$ mit $\rho < |z_o| < R$. Dann konvergiert $\sum a_n z_o^n$ absolut (Satz 2.68). Damit ist $|a_n z_o^n|$ eine Nullfolge und inbesondere beschränkt. Es gibt also ein $K > 0$ mit $|a_n z_o^n| < K$ für alle $n \in \mathbb{N}$ Damit gilt für $z \in \overline{D_\rho(0)}$:

$$|f_n(z)| = |a_n z^n| = |a_n z_o^n| \cdot \left|\frac{z}{z_o}\right|^n.$$

Da $\left|\frac{z}{z_o}\right| \leq \frac{\rho}{|z_o|} =: \kappa$ und $\kappa \in (0,1)$ gilt $|f_n(z)| \leq K\kappa^n$ für alle $z \in \overline{D_\rho(0)}$ und damit $\|f_n\|_{\sup} \leq K\kappa^n$ auf $\overline{D_\rho(0)}$. Da die geometrische Reihe $\sum \kappa^n$ konvergiert, konvergiert auch $\sum \|f_n\|_{\sup}$ auf $\overline{D_\rho(0)}$. Mit Theorem 3.43 folgt nun die gewünschte Aussage. □

Bemerkung 3.45. Die Folge $F_n : D_R(0) \to \mathbb{C}$, $z \mapsto \sum_{j=0}^{n} a_j z^j$ konvergiert zwar punktweise aber im Allgemeinen nicht gleichmäßig auf der offenen Kreisscheibe $D_R(0)$ gegen $F : D_R(0) \to \mathbb{C}$, $z \mapsto \sum_{j=0}^{\infty} a_j z^j$, wohl aber ist die Konvergenz auf jedem Kompaktum in $D_R(0)$ gleichmäßig.

Korollar 3.46. *Sei $\sum a_n z^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann ist die Funktion*

$$f : D_R(0) \to \mathbb{C}, \quad z \mapsto \sum_{j=0}^{\infty} a_j z^j$$

auf $D_R(0)$ stetig.

Beweis. Sei $z_o \in D_R(0)$ beliebig gewählt, dann gibt es ein $\rho > 0$ mit $|z_o| < \rho < R$. Da die Funktion $\overline{D_\rho(0)} \to \mathbb{C}$, $z \mapsto \sum_{j=0}^{\infty} a_j z^j$ auf ganz $\overline{D_\rho(0)}$ stetig ist, ist f auch in z_o stetig. □

Korollar 3.47 (verschobene Potenzreihen). *Sei* $\sum a_n z^n$ *eine Potenzreihe mit Konvergenzradius* $R > 0$ *und* $a \in \mathbb{C}$. *Dann konvergiert die Reihe* $\sum\limits_{j=0}^{\infty} a_j(z_o - a)^j$ *für alle* $z_o \in D_R(a) = \{z \in \mathbb{C} : |z - a| < R\}$ *absolut, und die Funktion*

$$f : D_R(a) \to \mathbb{C}, \quad z \mapsto \sum_{j=0}^{\infty} a_j(z-a)^j$$

ist auf $D_R(a)$ *stetig.*

Lemma 3.48 (Potenzreihen mit reellen Koeffizienten). *Sei* $\sum a_n z^n$ *eine Potenzreihe mit Konvergenzradius* $R > 0$. *Zudem gelte* $a_n \in \mathbb{R}$ *für alle* $n \in \mathbb{N}$. *Dann gilt für die stetige Funktion* $f : D_R(0) \to \mathbb{C}$, $z \mapsto \sum\limits_{j=0}^{\infty} a_j z^j$:

$$f(\overline{z}) = \overline{f(z)} \quad \textit{für alle } z \in D_R(0).$$

Insbesondere ist $f(x) \in \mathbb{R}$ *für alle* $x \in D_R(0) \cap \mathbb{R}$.

Beweis. Sei $z \in D_R(0)$, dann gilt auch $\overline{z} \in D_R(0)$ und die Folge $F_n(\overline{z}) := \sum\limits_{j=0}^{n} a_j \overline{z}^j$ konvergiert gegen $f(\overline{z})$. Andererseits konvergiert $F_n(\overline{z})$ auch gegen $\overline{f(z)}$, da die komplexe Konjugation stetig ist, und wegen $\overline{a_n} = a_n$ gilt:

$$F_n(\overline{z}) := \sum_{j=0}^{n} a_j \overline{z}^j = \sum_{j=0}^{n} \overline{a_j}\,\overline{z^j} = \overline{\sum_{j=0}^{n} a_j z^j} = \overline{F_n(z)}.$$

□

3.5. Exponentialfunktion

Dieser Abschnitt ist sehr stark inspiriert von [Fo1, § 8] und von [K1, Kap. 8].

Lemma 3.49. *Die* Exponentialreihe

$$\sum \frac{1}{j!} z^j$$

hat unendlichen Konvergenzradius.

Beweis. Nach der Eulerschen Formel (siehe Satz 2.70) gilt

$$R = \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} = \lim_{n\to\infty} \frac{(n+1)!}{n!} = \lim_{n\to\infty} (n+1) = \infty.$$

□

Bemerkung 3.50. Insbesondere ist für ein beliebiges aber festes $z \in \mathbb{C}$ die Folge

$$a_n := \frac{z^n}{n!}$$

eine Nullfolge.

Die Exponentialreihe ist also auf jeder kompakten Teilmenge von $\mathbb{C}$ gleichmäßig konvergent, nicht jedoch auf ganz $\mathbb{C}$.

Definition 3.51. Die Abbildung

$$\exp : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \exp(z) := \sum_{j=0}^{\infty} \frac{1}{j!} z^j$$

heißt (komplexe) *Exponentialfunktion*.

Korollar 3.52. *Die Exponentialabbildung ist auf ganz* $\mathbb{C}$ *stetig und es gilt* $\exp(0) = \frac{1}{0!} = 1$.

Beweis. Sei $z_o \in \mathbb{C}$, dann gibt es ein $r > 0$ mit $|z_o| < r$. Auf $\overline{D_r(0)}$ ist die Exponentialreihe gleichmäßig konvergent und stellt damit dort eine stetige Funktion dar. Also ist exp in z_o stetig. $\square$

Definition 3.53. Die reelle Zahl $e := \exp(1) = \sum\limits_{j=0}^{\infty} \frac{1}{j!}$ heißt *Eulersche Zahl*

Die Eulersche Zahl ist irrational, wie folgende Aufgabe zeigt:

Aufgabe 3.54. Zeigen Sie:

(a) Für alle $n \in \mathbb{N}$ gilt $0 < e - s_n < \frac{2}{(n+1)!}$, wobei $s_n := \sum\limits_{j=0}^{n} \frac{1}{j!}$.

(b) Für alle $n \in \mathbb{N}$ gilt $n! \cdot s_n \in \mathbb{N}$.

(c) Zeigen Sie $e \notin \mathbb{Q}$. (Hinweis: Wäre $e \in \mathbb{Q}$, dann gäbe es ein $N \in \mathbb{N}$ mit $N! \cdot e \in \mathbb{N}$.)

Satz 3.55 (Funktionalgleichung der Exponentialfunktion)**.** *Für* $z, w \in \mathbb{C}$ *gilt*

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

Beweis. Nach dem Satz von Mertens über das Cauchy-Produkt (Satz 2.64) absolut konvergenter Reihen gilt

$$\exp(z)\exp(w) = \left(\sum_{n=0}^{\infty} \frac{z^n}{n!}\right) \cdot \left(\sum_{j=0}^{\infty} \frac{w^j}{j!}\right) = \sum_{n=0}^{\infty} c_n$$

mit

$$c_n = \sum_{j=0}^{n} \left(\frac{z^j}{j!} \cdot \frac{w^{n-j}}{(n-j)!} \right) = \sum_{j=0}^{n} \frac{n!}{n!j!(n-j)!} z^j w^{n-j} = \frac{1}{n!} \sum_{j=0}^{n} \binom{n}{j} z^j w^{n-j} = \frac{1}{n!}(z+w)^n,$$

wobei wir in der letzten Gleichung den Binomischen Lehrsatz (Satz 2.5) verwendet haben. Zusammenfassend gilt also

$$\exp(z)\exp(w) = \sum_{n=0}^{\infty} \frac{1}{n!}(z+w)^n = \exp(z+w).$$

□

Bemerkung 3.56. Die komplexe Exponentialfunktion hat offenbar die Eigenschaft $\exp(\mathbb{R}) \subset \mathbb{R}$ (siehe Lemma 3.48). Die auf ganz $\mathbb{R}$ stetige Funktion $\exp|_{\mathbb{R}}$ heißt *reelle* Exponentialfunktion und wird üblicherweise auch einfach mit exp bezeichnet.

Satz 3.57 (Eigenschaften der Exponentialfunktion)**.** *(i)* $\exp(z) \neq 0$ *für alle* $z \in \mathbb{C}$.

(ii) $\exp(-z) = \frac{1}{\exp(z)}$ *für alle* $z \in \mathbb{C}$.

(iii) $\exp(x) > 0$ *für alle* $x \in \mathbb{R}$.

(iv) $\exp(n) = e^n$ *für alle* $n \in \mathbb{Z}$.

(v) $\exp\left(\frac{m}{n}\right) = (\sqrt[n]{e})^m$ *für alle* $m \in \mathbb{Z}$ *und* $n \in \mathbb{N}^*$.[96]

(vi) $|\exp(ix)| = 1$ *für alle* $x \in \mathbb{R}$.

(vii) Die reelle Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ *ist streng monoton wachsend und somit injektiv.*

(viii) $\exp(\mathbb{R}) = \mathbb{R}_{>0}$.

Beweis.

- Zu (i) und (ii): $\exp(z)\exp(-z) = \exp(z-z) = \exp(0) = 1$. Damit ist $\exp(z) \neq 0$ für alle $z \in \mathbb{C}$ (Nullteilerfreiheit von Körpern), und es gilt $\exp(-z) = \frac{1}{\exp(z)}$.
- Für $x \geq 0$ gilt $\exp(x) = \sum\limits_{j=0}^{\infty} \frac{x^j}{j!} > 0$, und für $x < 0$ gilt $\exp(x) = \frac{1}{\exp(-x)} > 0$. Das zeigt (iii).
- Um (iv) zu zeigen, genügt es wegen (ii) $\exp(n) = e^n$ für alle $n \in \mathbb{N}$ mit vollständiger Induktion nachzuweisen:

  Für $n = 0$ ist die Aussage wegen $\exp(0) = 1 = e^0$ sicherlich war. Für $n = 1$ ist die Aussage gerade die Definition von e.

  Zum Induktionsschlus betrachten wir

  $$\exp(n+1) = \exp(n)\exp(1) = e^n \cdot e = e^{n+1}.$$

[96] Hierbei verwenden wir für eine Zahl $x^{-k} = \frac{1}{x^k}$ für $k \in \mathbb{N}$.

- Für $q = \frac{1}{n}$, $n \in \mathbb{N}^*$, gilt
$$e = \exp(1) = \exp\left(\frac{n}{n}\right) = \left(\exp\left(\frac{1}{n}\right)\right)^n,$$
also $\exp\left(\frac{1}{n}\right) = \sqrt[n]{e}$.
Ist $q = \frac{m}{n} \in \mathbb{Q}_{>0}$ mit $m, n \in \mathbb{N}^*$, so gilt
$$\exp\left(\frac{m}{n}\right) = \exp\left(\frac{1}{n}\right)^m = (\sqrt[n]{e})^m.$$
Ist $q = -\frac{m}{n} \in \mathbb{Q}_{<0}$ mit $m, n \in \mathbb{N}^*$, so erhalten wir
$$\exp\left(-\frac{m}{n}\right) = \left(\exp\left(\frac{m}{n}\right)\right)^{-1} = (\sqrt[n]{e})^{-m}.$$
Damit ist (v) gezeigt.
- Mit Lemma 3.48 gilt $\exp(-ix) = \exp(\overline{ix}) = \overline{\exp(ix)}$ und daher (vi):
$$|\exp(ix)|^2 = \exp(ix)\overline{\exp(ix)} = \exp(ix)\exp(-ix) = \exp(0) = 1.$$
- Für alle $x \in \mathbb{R}$ und alle $h > 0$ gilt
$$\exp(x+h) = \exp(x)\exp(h) = \exp(x) \cdot \underbrace{\sum_{j=0}^{\infty} \frac{h^j}{j!}}_{> \frac{h^0}{0!} = 1} > \exp(x).$$
Das beweist (vii).
- Sei $y > 1$. Da $\exp(0) = 1$ und $\exp(y) > 1 + y$ gibt es nach dem Zwischenwertsatz ein $x \in [0, y]$ mit $\exp(x) = y$. Ist $y \in]0, 1[$, so ist $\frac{1}{y} > 1$ und es gibt es $x \in \mathbb{R}$ mit $\exp(x) = \frac{1}{y}$, also $\exp(-x) = y$. Schliesslich ist $\exp(0) = 1$. Damit gilt (viii).

□

Proposition 3.58. *Sei (z_n) eine Folge in $\mathbb{C}$ mit $\lim\limits_{n\to\infty} z_n = z$, dann gilt*
$$\exp(z) = \lim_{n\to\infty} \left(\left(1 + \frac{z_n}{n}\right)^n\right).$$

Beweis. Sei $\varepsilon > 0$. Da $\lim\limits_{n\to\infty} z_n = z$ und die Exponentialreihe auf ganz $\mathbb{C}$ konvergiert, kann man ein $N_1 \in \mathbb{N}$ wählen, sodass für alle $n \geq N_1$ gilt
$$\sum_{j=N_1+1}^{\infty} \frac{(|z|+1)^j}{j!} < \frac{\varepsilon}{3} \quad \text{und} \quad |z_n| < |z| + 1.$$

Dann erhalten wir für $n \geq N_1$ mit dem Binomischen Lehrsatz

$$\begin{aligned} \left|\left(1+\frac{z_n}{n}\right)^n - \exp(z)\right| &= \left|\left(1+\frac{z_n}{n}\right)^n - \sum_{j=0}^{\infty}\frac{z^j}{j!}\right| \\ &\leq \underbrace{\sum_{j=0}^{N_1}\left|\binom{n}{j}\frac{z_n^j}{n^j} - \frac{z^j}{j!}\right|}_{=:A} + \underbrace{\sum_{j=N_1+1}^{n}\binom{n}{j}\frac{|z_n|^j}{n^j}}_{:=B} + \underbrace{\sum_{j=N_1+1}^{\infty}\frac{|z|^j}{j!}}_{<\frac{\varepsilon}{3}}. \end{aligned}$$

Wegen

$$\begin{aligned} \binom{n}{j}\frac{1}{n^j} &= \frac{n!}{j!(n-j)!n^j} \\ &= \frac{n(n-1)(n-2)\dots(n-j+1)}{j!n^j} \\ &= \frac{1}{j!}\underbrace{\underbrace{\left(1-\frac{1}{n}\right)}_{<1}\underbrace{\left(1-\frac{2}{n}\right)}_{<1}\dots\underbrace{\left(1-\frac{j-1}{n}\right)}_{<1}}_{(j-1)\text{ Faktoren}} \\ &< \frac{1}{j!} \end{aligned}$$

gilt $\lim\limits_{n\to\infty}\binom{n}{j}\frac{1}{n^j} = \frac{1}{j!}$ und zudem

$$B = \sum_{j=N_1+1}^{n}\binom{n}{j}\frac{|z_n|^j}{n^j} \leq \sum_{j=N_1+1}^{n}\frac{(|z|+1)^j}{j!} < \frac{\varepsilon}{3}.$$

Da $\lim\limits_{n\to\infty}\binom{n}{j}\frac{z_n^j}{n^j} = \frac{z^j}{j!}$ konvergiert die endliche aus N_1+1 Summanden bestehende Summe A gegen 0. Somit gibt es gibt ein $N_2 \in \mathbb{N},\ N_2 > N_1$, sodass für alle $n \geq N_2$ gilt $A < \frac{\varepsilon}{3}$. Insgesamt erhalten wir $\left|\left(1+\frac{z_n}{n}\right)^n - \exp(z)\right| < \varepsilon$ für $n \geq N_2$. $\square$

Korollar 3.59. $e = \lim\limits_{n\to\infty}\left(\left(1+\frac{1}{n}\right)^n\right)$.

Aufgabe 3.60. Zeigen Sie: Für vorgegebenes $c \in \mathbb{C}$ ist $f : \mathbb{C} \to \mathbb{C},\ z \mapsto \exp(cz)$ die einzige Funktion, für die gilt

- $f(z_1 + z_2) = f(z_1)f(z_2)$ für alle $z_1, z_2 \in \mathbb{C}$ und
- $\lim\limits_{n\to\infty}\frac{f(z_n)-1}{z_n} = c$ für alle komplexen Nullfolgen (z_n).

3.6. Trigonometrische Funktionen

Dieser Abschnitt ist sehr stark inspiriert von [Fo1, § 14] und von [K1, Kap. 8].

Definition 3.61. Die Funktionen

$$\begin{aligned}\cos : \mathbb{R} \to \mathbb{R}, \quad & x \mapsto \Re(\exp(ix)) = \frac{1}{2}(\exp(ix) + \exp(-ix)) \quad \text{und} \\ \sin : \mathbb{R} \to \mathbb{R}, \quad & x \mapsto \Im(\exp(ix)) = \frac{1}{2i}(\exp(ix) - \exp(-ix))\end{aligned}$$

heißen (reeller) *Cosinus* (Kosinus) und (reeller) *Sinus*.

Bemerkung 3.62.

- Für alle $x \in \mathbb{R}$ gilt

$$\exp(ix) = \cos(x) + i\sin(x).$$

  Diese Gleichung wir auch *Eulersche Formel* genannt.

- $\cos(-x) = \cos(x)$ und $\sin(-x) = -\sin(x)$.
- $\sin^2(x) + \cos^2(x) = 1$.

Da die Multiplikation mit der imaginären Einheit i, die komplexe Exponentialfunktion $\exp : \mathbb{C} \to \mathbb{C}$ und die Funktionen Realteil $\Re : \mathbb{C} \to \mathbb{R}$ und Imaginärteil $\Im : \mathbb{C} \to \mathbb{R}$ auf ganz $\mathbb{C}$ stetig sind, folgt:

Satz 3.63. *Die reellen Funktionen Sinus,* $\sin : \mathbb{R} \to \mathbb{R}$, *und Cosinus,* $\cos : \mathbb{R} \to \mathbb{R}$, *sind auf ganz* $\mathbb{R}$ *stetig.*

Satz 3.64 (Additionstheoreme)**.** *Sei* $x, y \in \mathbb{R}$, *dann gilt*

$$\begin{aligned}\cos(x+y) &= \cos(x)\cos(y) - \sin(x)\sin(y), \\ \sin(x+y) &= \sin(x)\cos(y) + \cos(x)\sin(y).\end{aligned}$$

Der Beweis ist eine Übungsaufgabe.

Korollar 3.65. *Sei* $x, y \in \mathbb{R}$, *dann gilt*

$$\begin{aligned}\sin(x) - \sin(y) &= 2\cos\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right), \\ \cos(x) - \cos(y) &= -2\sin\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right).\end{aligned}$$

Der Beweis ist wiederum eine Übungsaufgabe.

Satz 3.66 (Potenzreihendarstellung von Sinus und Cosinus)**.** *Sei* $x \in \mathbb{R}$ *beliebig, dann gilt*

$$\sin(x) = \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!} x^{2j+1} \quad \textit{und} \quad \cos(x) = \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j)!} x^{2j}.$$

Beide Potenzreihen konvergieren absolut für alle $x \in \mathbb{R}$ *und gleichmäßig auf jedem Kompaktum in* $\mathbb{R}$.

Beweis. Wir teilen die Exponentialreihe in eine Potenzreihe mit geraden und eine mit ungeraden Exponenten auf. Mit $i^{2j} = (-1)^j$ und $i^{2j+1} = (-1)^j i$ erhalten wir

$$\exp(ix) = \sum_{j=0}^{\infty} \frac{1}{j!}(ix)^j = \sum_{j=0}^{\infty} \frac{i^j}{j!} x^j = \underbrace{\sum_{j=0}^{\infty} \frac{(-1)^j}{(2j)!} x^{2j}}_{=\Re(\exp(ix))\in\mathbb{R}} + i \cdot \underbrace{\sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!} x^{2j+1}}_{=\Im(\exp(ix))\in\mathbb{R}}$$

und damit die Aussage.

Die absolute und auf jedem Kompaktum in $\mathbb{R}$ gleichmäßige Konvergenz der angegebenen Potenzreihen für Sinus und Cosinus folgt sofort aus den entsprechenden Eigenschaften der komplexen Exponentialreihe. □

Bemerkung 3.67. Da die Potenzreihen des reellen Sinus und Cosinus konvergieren, haben die komplexen Potenzreihen $\sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!} z^{2j+1}$ und $\cos(x) = \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j)!} z^{2j}$ mit $z \in \mathbb{C}$ Konvergenzradius $R = \infty$ (vgl. Lemma 2.66). Wir können damit den Sinus und den Cosinus auf ganz $\mathbb{C}$ erweitern indem wir setzten

$$\begin{aligned} \sin : \mathbb{C} \to \mathbb{C}, \qquad & z \mapsto \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!} z^{2j+1} \quad \text{und} \\ \cos : \mathbb{C} \to \mathbb{C}, \qquad & z \mapsto \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j)!} z^{2j}. \end{aligned}$$

Beide Funktionen sind auf ganz $\mathbb{C}$ stetig und heißen *komplexe* Sinus bzw. Cosinusfunktion. Die Additionstheoreme aus Satz 3.64 gelten auch für die komplexen Cosinus- und Sinusfunktionen.

Satz 3.68 (Einschließungslemma)**.** *Ist $x \in (0, 2]$, so gilt*

$$\begin{aligned} 1 - \tfrac{x^2}{2} &< \cos(x) < 1 - \tfrac{x^2}{2} + \tfrac{x^4}{24}, \\ x(1 - \tfrac{x^2}{6}) = x - \tfrac{x^3}{6} &< \sin(x) < x. \end{aligned}$$

Beweis. Wir beweisen die Aussage lediglich für den Sinus. Die Argumentation für den Cosinus wird analog geführt (siehe z. B. [K1, Abschn. 8.7]).

Die Sinus-Reihe

$$\sum (-1)^j \frac{x^{2j+1}}{(2j+1)!}$$

ist eine alternierende Reihe. Und für $x \in (0, 2]$ ist die Folge (a_n) mit $a_n := \frac{x^{2n+1}}{(2n+1)!}$ eine strikt positive Nullfolge (da ja die Sinusreihe konvergiert). Zudem ist a_n streng monoton fallend, da

$$\frac{a_{n+1}}{a_n} = \frac{x^2}{(2n+2)(2n+3)} < 1, \quad \text{für} \quad x \in (0, 2].$$

Wie im Beweis des Leibniz-Kriteriums gesehen, gilt für $s_n(x) := \sum_{j=0}^{n} (-1)^j \frac{x^{2j+1}}{(2j+1)!}$

$$s_{2k+1}(x) \leq \sin(x) \quad \text{und} \quad \sin(x) \leq s_{2k}$$

für alle $k \in \mathbb{N}$. Da nun aber die Nullfolge (a_n) *streng* monoton fällt, gilt sogar

$$s_1 = x - \frac{x^3}{6} < \sin(x) < s_0 = x.$$

$\square$

Korollar 3.69. *(i) Ist* $x \in (0, 2]$, *dann gilt* $\sin(x) > 0$.

(ii) $\cos|_{[0,2]}$ *ist eine streng monoton fallende Funktion.*

Beweis. Die Aussage (i) ergibt sich aus $x(1 - \frac{x^2}{6}) < \sin(x)$, da für $x \in (0, 2]$ sowohl x wie auch $1 - \frac{x^2}{6}$ strikt positiv sind.

Für die Aussage (ii) verwenden wir Korollar 3.65. Seien $x, y \in [0, 2]$ mit $x > y$, dann gilt

$$\cos(x) - \cos(y) = -2 \cdot \underbrace{\sin\left(\frac{x-y}{2}\right)}_{>0} \underbrace{\sin\left(\frac{x+y}{2}\right)}_{>0} < 0.$$

$\square$

Korollar 3.70. *Die Funktion* $\cos|_{[0,2]}$ *hat genau eine Nullstelle.*

Beweis. Da $\cos|_{[0,2]}$ stetig ist und $\cos(0) = 1$ und $\cos(2) < 1 - \frac{2^2}{2} + \frac{2^4}{24} = -\frac{1}{3}$ hat $\cos|_{[0,2]}$ nach dem Zwischenwertsatz mindestens eine Nullstelle. Da $\cos|_{[0,2]}$ streng monoton fallend ist, gibt es höchstens Nullstelle. $\square$

Definition 3.71. Wir bezeichnen mit π die eindeutige Zahl in $[0, 4]$, sodass $\cos\left(\frac{\pi}{2}\right) = 0$ gilt. Diese Zahl π wird auch *Kreiszahl* genannt.

Korollar 3.72. *Es gilt* $\sin\left(\frac{\pi}{2}\right) = 1$.

Beweis. Wegen $1 = \cos^2\left(\frac{\pi}{2}\right) + \sin^2\left(\frac{\pi}{2}\right) = \sin^2\left(\frac{\pi}{2}\right)$ gilt $\sin\left(\frac{\pi}{2}\right) = \pm 1$. Mit $\sin(x) > 0$ für alle $x \in (0, 2]$, folgt die Behauptung. $\square$

Aus $\exp\left(i\frac{\pi}{2}\right) = \cos\left(\frac{\pi}{2}\right) + i \sin\left(\frac{\pi}{2}\right) = i$ folgt mit der Funktionalgleichung der Exponentialfunktion und der Eulerschen Formel folgende Wertetabelle:

x	0	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π
$\exp(ix)$	1	i	-1	$-i$	1
$\cos(x)$	1	0	-1	0	1
$\sin(x)$	0	1	0	-1	0

Damit erhalten wir aus der Funktionalgleichung der Exponentialfunktion und der Eulerschen Gleichung:

$$\begin{array}{rclrclrcl}
\exp\left(z+i\frac{\pi}{2}\right) &=& i\exp(z), & \exp(z+i\pi) &=& -\exp(z), & \exp(z+i2\pi) &=& \exp(z),\\
\cos\left(x+\frac{\pi}{2}\right) &=& -\sin(x), & \cos(x+\pi) &=& -\cos(x), & \cos(z+2\pi) &=& \cos(x),\\
\sin\left(x+\frac{\pi}{2}\right) &=& \cos(x), & \sin(x+\pi) &=& -\sin(x), & \sin(z+2\pi) &=& \sin(x),
\end{array}$$

für alle $z \in \mathbb{C}$ und alle $x \in \mathbb{R}$.

Definition 3.73. Eine Funktion $f : \mathbb{C} \to \mathbb{C}$ bzw. $f : \mathbb{R} \to \mathbb{R}$ heißt *periodisch* mit *Periode* $\rho \in \mathbb{C}^*$ bzw. $\rho \in \mathbb{R}^*$, wenn

$$f(z+\rho) = f(z) \quad \text{bzw.} \quad f(x+\rho) = f(x)$$

für alle $z \in \mathbb{C}$ bzw. alle $x \in \mathbb{R}$ gilt.

Wir haben gesehen, dass cos und sin periodisch mit Periode 2π sind.

Satz 3.74. *Die Menge aller Nullstellen des reellen Cosinus ist* $\left\{\frac{\pi}{2} + k\pi : \ k \in \mathbb{Z}\right\}$.
Die Menge aller Nullstellen des reellen Sinus ist $\{k\pi : \ k \in \mathbb{Z}\}$.

Beweis. Die einzige Nullstelle von $\cos|_{[0,\pi/2]}$ ist $\frac{\pi}{2}$. Wegen $\cos(x) = \cos(-x)$ ist $\frac{\pi}{2}$ auch die einzige Nullstelle von $\cos|_{(-\pi/2,\,\pi/2]}$. Wegen $\cos(x+\pi) = -\cos(x)$ sind $\frac{\pi}{2}$ und $\frac{\pi}{2}+\pi$ die einzigen Nullstellen von cos im Intervall $(-\pi/2, \pi/2+\pi]$ der Länge 2π. Da cos periodisch mit Periode 2π ist, folgt die Aussage für die Nullstellen des reellen Cosinus.
Aus $\cos\left(x+\frac{\pi}{2}\right) = -\sin(x)$ folgt die entsprechende Aussage für den Sinus. □

Korollar 3.75. *Die kleinste positive Periode der reellen Sinus- und Cosinusfunktion ist* 2π.

Beweis. Da die Nullstellen von sin und cos einen Abstand von $k\pi,\ k \in \mathbb{Z}$, haben, ist jede Periode von sin und cos von der Form $k\pi,\ k \in \mathbb{Z}^*$. Da aber π wegen $\cos(0) = 1$ und $\cos(\pi) = -1$ keine Periode von cos ist ist 2π die kleinste positive Periode vom Cosinus. Analoge Argumente führen zur entsprechenden Aussage für den Sinus. □

Aufgabe 3.76. Sei $z \in \mathbb{C}$. Zeigen Sie: Wenn $\exp(z) = 1$, dann gilt $z \in \{2k\pi i : \ k \in \mathbb{Z}\}$.

Aufgabe 3.77. Zeigen Sie:

(a) $\cos|_{[-\pi/2,0]}$ ist streng monoton wachsend.

(b) $\cos|_{[0,\pi]}$ ist streng monoton fallend.

(c) $\sin|_{[-\pi/2,\,\pi/2]}$ ist streng monoton wachsend.

Mit $\cos(0) = 1$, $\cos(\pi) = -1$, $\sin(-\pi/2) = -1$ und $\sin(\pi/2) = 1$ sind nach dem Zwischenwertsatz für stetige Funktionen die streng monotonen Funktionen $\cos : [0,\pi] \to [-1,1]$ und $\sin : [-\pi/2, \pi/2] \to [-1,1]$ bijektiv. Deren Umkehrfunkionen heißen Hauptzweig des *Arkuscosinus*, notiert

$$\arccos : [-1,1] \to [0,\pi]$$

bzw. Hauptzweig des *Arkussinus*, notiert

$$\arcsin : [-1,1] \to [-\pi/2, \pi/2].$$

Definition 3.78. Die Abbildungen

$$\tan : \mathbb{R} \setminus \left\{\frac{\pi}{2} + k\pi : \; k \in \mathbb{Z}\right\} \to \mathbb{R}, \quad x \mapsto \tan(x) := \frac{\sin(x)}{\cos(x)}$$

bzw.

$$\cot : \mathbb{R} \setminus \{k\pi : \; k \in \mathbb{Z}\} \to \mathbb{R}, \quad x \mapsto \cot(x) := \frac{\cos(x)}{\sin(x)}$$

heißen reeller *Tangens* bzw. reeller *Cotangens* (Kotangens).

Aufgabe 3.79. Zeigen Sie:

(a) $\tan|_{(-\pi/2,\, \pi/2)}$ ist stetig.

(b) $\tan|_{(-\pi/2, 0)}$ ist negativ.

(c) $\tan|_{(0,\, \pi/2)}$ ist positiv.

(d) $\tan|_{(-\pi/2,\, \pi/2)}$ ist streng monoton wachsend.

(e) $\tan((-\pi/2,\, \pi/2)) = \mathbb{R}$.

Wiederum mithilfe des Zwischenwertsatzes sehen wir, dass die streng monotone stetige Funktion

$$\tan : (-\pi/2,\, \pi/2) \to \mathbb{R}$$

bijektiv ist. Deren Umkehrfunktion heißt Hauptzweig des *Arkustangens*, notiert

$$\arctan : \mathbb{R} \to (-\pi/2,\, \pi/2).$$

Satz 3.80 (Polarkoordinaten)**.** *Die Abbildung*

$$\mathbb{R}_{>0} \times \left[-\frac{\pi}{2}, \frac{3\pi}{2}\right) \to \mathbb{C}^* := \mathbb{C} \setminus \{0\}, \quad (r,\varphi) \mapsto r\exp(i\varphi) = r\cos(\varphi) + ir\sin(\varphi)$$

ist bijektiv.

Beweis. Offenbar gilt $|r\exp(i\varphi)| = r|\exp(i\varphi)| = r$.

Zur Surjektivität: Sei $z \in \mathbb{C}^*$, dann zerlegen wir $\frac{z}{|z|}$ in Real- und Imaginärteil, d. h. wir schreiben $\frac{z}{|z|} = x + iy$ mit $x, y \in \mathbb{R}$ und $x^2 + y^2 = 1$. Wir definieren

$$\varphi := \begin{cases} \arcsin(y), & \text{wenn } x \geq 0, \\ \pi - \arcsin(y), & \text{wenn } x < 0. \end{cases}$$

Wenn $x \geq 0$, dann ist $\varphi \in [-\pi/2,\ \pi/2]$ und $\sin(\varphi) = y$. In diesem Fall ist $\cos(\varphi) \geq 0$ und es gilt $\cos(\varphi) = x$, da $\sin(\varphi)^2 + \cos(\varphi)^2 = 1$.

Wenn $x < 0$, dann ist $\varphi \in [\pi/2,\ 3\pi/2]$ und $\sin(\varphi) = \sin(\pi - \arcsin(y)) = -\sin(-\arcsin(y)) = \sin(\arcsin(y)) = y$. Desweiteren gilt offenbar $\cos(\varphi) \leq 0$ und daher wegen $\sin(\varphi)^2 + \cos(\varphi)^2 = 1$ auch $\cos(\varphi) = x$.

Insgesamt gilt also mit $r := |z|$:

$$z = |z| \cdot \frac{z}{|z|} = r(x + iy) = r(\cos(\varphi) + i\sin(\varphi)) = r\exp(i\varphi).$$

Zur Injektivität: Sei $z = r_1\exp(i\varphi_1) = r_2\exp(i\varphi_2)$ für $r_{1,2} \in \mathbb{R}_{>0}$ und $\varphi_{1,2} \in \left[-\frac{\pi}{2}, \frac{3\pi}{2}\right)$. dann gilt $r_1 = |z| = r_2$ und damit $\exp(i\varphi_1) = \exp(i\varphi_2)$, also $1 = \exp(i\varphi_1)\exp(-i\varphi_2) = \exp(i(\varphi_1 - \varphi_2))$. Somit gilt $\varphi_1 - \varphi_2 \in 2\mathbb{Z}\pi$ und damit $\varphi_1 - \varphi_2 = 0$. $\square$

Die Darstellung einer komplexen Zahl $z \neq 0$ als $z = r\exp(i\varphi)$ heißt *Polarkoordinaten*. Hierbei ist $r = |z|$ der Betrag von z, und $\varphi \in \mathbb{R}$ heißt das *Argument* von z, wobei φ nur bis auf Addition eines Elementes von $2\mathbb{Z}\pi$ durch z bestimmt ist. Oftmals wird daher einfach $\varphi \in [0, 2\pi)$ angenommen.

Ausblick

Der Begriff einer stetigen Abbildung gehört genuin in den Bereich der *mengentheoretischen Topologie*. Lineare Abbildungen zwischen topologischen Räumen spielen in der Topologie eine ähnliche Rolle wie lineare Abbildungen zwischen Vektorräumen in der Linearen Algebra.

Unsere Definition einer stetigen Abbildung erweitert sich in natürlicher Weise auch auf Abbildungen zwischen metrischen Räumen. Die Äquivalenz von Stetigkeit und Folgenstetigkeit gilt auch für Abbildungen zwischen metrischen Räumen.

Wer mehr über den Begriff der stetigen Abbildung zwischen metrischen oder sogar allgemeinen topologischen Räumen erfahren möchte, sei auf die ersten beiden Kapitel des Buches [Q].

4 Differentialrechnung

4.1. Grenzwerte bei Funktionen

Wir orientieren uns hier an [Fo1, S. 60 f.].

Definition 4.1. Sei $X \subset \mathbb{C}$ eine nichtleere Teilmenge, dann bezeichnet $\overline{X}$ die Menge aller Elemente $a \in \overline{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ mit der Eigenschaft: Es gibt eine Folge (z_n) von Elementen aus X mit $\lim\limits_{n\to\infty} z_n = a$. Die Menge $\overline{X}$ heißt *Abschluß* von X in $\overline{\mathbb{C}}$.

Sei $f : X \to \mathbb{C}$ eine Funktion. Wir schreiben

- $\lim\limits_{\substack{z\to a\\ z\in X}} f(z) = c$ oder kurz $\lim\limits_{z\to a} f(z) = c$
  mit $a \in \overline{X}$, wenn für *alle* Folgen (z_n) in X mit $\lim\limits_{n\to\infty} z_n = a$ gilt $\lim\limits_{n\to\infty} f(z_n) = c$.

- $\lim\limits_{\substack{z\to \infty\\ z\in X}} f(z) = c$ oder kurz $\lim\limits_{z\to \infty} f(z) = c$,
  wenn $\infty \in \overline{X}$ und für *alle* Folgen (z_n) in X mit $\lim\limits_{n\to\infty} z_n = \infty$ gilt $\lim\limits_{n\to\infty} f(z_n) = c$.

Definition 4.2. Sei $X \subset \mathbb{R}$ eine nichtleere Teilmenge, dann bezeichnet $\overline{X}$ die Menge aller Elemente $a \in \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty, -\infty\}$ mit der Eigenschaft: Es gibt eine Folge (x_n) von Elementen aus X mit $\lim\limits_{n\to\infty} x_n = a$. Die Menge $\overline{X}$ heißt *Abschluß* von X in $\overline{\mathbb{R}}$.

Sei $f : X \to \mathbb{C}$ eine Funktion. Wir schreiben

- $\lim\limits_{\substack{x\to a\\ x\in X}} f(x) = c$ oder kurz $\lim\limits_{x\to a} f(x) = c$
  mit $a \in \overline{X}$, wenn für *alle* Folgen (x_n) in X mit $\lim\limits_{n\to\infty} x_n = a$ gilt $\lim\limits_{n\to\infty} f(x_n) = c$.

- $\lim\limits_{\substack{x\uparrow a\\ x\in X}} f(x) = c$ oder kurz $\lim\limits_{x\uparrow a} f(x) = c$
  mit $a \in \overline{X}$, wenn für *alle* Folgen (x_n) in X mit $x_n < a$ für alle $n \in \mathbb{N}$ und $\lim\limits_{n\to\infty} x_n = a$ gilt $\lim\limits_{n\to\infty} f(x_n) = c$.

- $\lim\limits_{\substack{x \downarrow a \\ x \in X}} f(x) = c$ oder kurz $\lim\limits_{x \downarrow a} f(x) = c$
  mit $a \in \overline{X}$, wenn für *alle* Folgen (x_n) in X mit $x_n > a$ für alle $n \in \mathbb{N}$ und $\lim\limits_{n\to\infty} x_n = a$ gilt $\lim\limits_{n\to\infty} f(x_n) = c$.

- $\lim\limits_{\substack{x \to \infty \\ x \in X}} f(x) = c$ oder kurz $\lim\limits_{x \to \infty} f(x) = c$,
  wenn $\infty \in \overline{X}$ und für *alle* Folgen (x_n) in X mit $\lim\limits_{n\to\infty} x_n = \infty$ gilt $\lim\limits_{n\to\infty} f(x_n) = c$.

- $\lim\limits_{\substack{x \to -\infty \\ x \in X}} f(x) = c$ oder kurz $\lim\limits_{x \to -\infty} f(x) = c$,
  wenn $-\infty \in \overline{X}$ und für *alle* Folgen (x_n) in X mit $\lim\limits_{n\to\infty} x_n = -\infty$ gilt $\lim\limits_{n\to\infty} f(x_n) = c$.

Bemerkung 4.3. Eine Funktion $f : X \to \mathbb{C}$ ist offenbar genau dann stetig in $a \in X$, wenn $\lim\limits_{x\to a} f(x) = f(a)$ gilt.

4.2. Differenzierbare Funktionen

In diesem Abschnitt folgen wir der Darstellung in [K1, Abschn. 9.1] und [Fo1, § 15].

Definition 4.4. Sei $I \subset \mathbb{R}$ ein offenes Intervall und $f : I \to \mathbb{C}$ eine Funktion. Für $x_o \in I$ heißt die Abbildung

$$I \setminus \{x_o\} \to \mathbb{C}, \quad x \mapsto \frac{f(x) - f(x_o)}{x - x_o}$$

Differenzenquotient von f in x_o.

Wir nennen f

- *differenzierbar (ableitbar) in* $x_o \in I$, wenn der Grenzwert

$$\lim_{\substack{x \to x_o \\ x \in I \setminus \{x_o\}}} \frac{f(x) - f(x_o)}{x - x_o} = \lim_{x \to x_o} \frac{f(x) - f(x_o)}{x - x_o}$$

  existiert und eine komplexe Zahl ist. In diesem Fall heißt

$$f'(x_o) := \lim_{x \to x_o} \frac{f(x) - f(x_o)}{x - x_o}$$

  Ableitung oder *Differential* von f im Punkt x_o.

- auf ganz I *differenzierbar* oder kurz *differenzierbar* oder *ableitbar*, wenn f in *jedem* Punkt $x \in I$ differenzierbar ist. In diesem Fall heißt die Funktion

$$f' : I \to \mathbb{C}, \quad x \mapsto f'(x)$$

  Ableitung von f.

Bemerkung 4.5.

- Die angegebene Definition von Differenzierbarkeit kann man tel quel auch für Abbildungen von $f : \mathbb{C} \to \mathbb{C}$ übernehmen. Jedoch führt das Studium dieses Ableitungsbegriffes für Funktionen einer komplexen Variablen zu einer sehr schönen aber eigenen Theorie, der sogenannten Funktionentheorie.
- Ist $f : I \to \mathbb{R}$ reellwertig, so hat der Differenzenquotient $\frac{f(x)-f(x_o)}{x-x_o}$ eine geometrische Interpretation: Er gibt die Steigung der Sekante des Graphen von f zwischen $(x, f(x))$ und $(x_o, f(x_o))$ an.
- Alternativ kann man auch sagen: $f : I \to \mathbb{C}$ ist in $x_o \in I$ differenzierbar, wenn der Limes
$$\lim_{h\to 0} \frac{f(x_o + h) - f(x_o)}{h}$$
existiert. In diesem Fall ist natürlich
$$\lim_{h\to 0} \frac{f(x_o + h) - f(x_o)}{h} = f'(x_o) = \lim_{x\to x_o} \frac{f(x) - f(x_o)}{x - x_o}.$$
- Anstelle von $f'(x_o)$ sind auch die Notationen $\dot{f}(x_o)$, $\frac{df}{dx}(x_o)$ $D_{x_o}f$ oder $Df(x_o)$ üblich.

Satz 4.6 (Charakterisierung von Differenzierbarkeit)**.** *Für eine Funktion $f : I \to \mathbb{C}$ sind folgende Aussagen äquivalent:*

(i) f ist in $x_o \in I$ differenzierbar.

(ii) Es gibt eine Funktion $\phi : I \to \mathbb{C}$, welche in x_o stetig ist, sodass für alle $x \in I$ gilt:
$$f(x) - f(x_o) = (x - x_o) \cdot \phi(x).$$

(iii) Es gibt eine $\mathbb{R}$-lineare Abbildung $L : \mathbb{R} \to \mathbb{C}$, sodass
$$\lim_{h\to 0} \frac{f(x_o + h) - f(x_o) - L(h)}{h} = 0.$$

Gegebenenfalls gilt $f'(x_o) = \phi(x_o)$ bzw. $L(h) = f'(x_o) \cdot h$.

Beweis. Ist f in x_o differenzierbar, dann ist die Funktion
$$\phi : I \to \mathbb{C}, \quad x \mapsto \begin{cases} \frac{f(x)-f(x_o)}{x-x_o}, & \text{wenn} \quad x \neq x_o \\ f'(x_o), & \text{wenn} \quad x = x_o \end{cases}$$
stetig, da $\lim\limits_{x\to x_o} \phi(x) = \lim\limits_{x\to x_o} \frac{f(x)-f(x_o)}{x-x_o} = f'(x_o) = \phi(x_o)$. Also gilt $(i) \implies (ii)$.

Wenn $f(x) - f(x_o) = (x - x_o) \cdot \phi(x)$ für eine in x_o stetige Funktion ϕ gilt, dann existiert auch
$$\lim_{x\to x_o} \frac{f(x) - f(x_o)}{x - x_o} = \lim_{x\to x_o} \phi(x) = \phi(x_o).$$

Somit gilt $(ii) \implies (i)$.

Ist f in x_o differenzierbar, so setze $L(h) = f'(x_o) \cdot h$. Damit folgt (iii) aus (i).

Sei $L(h) = c \cdot h$ mit $c \in \mathbb{C}$ eine Lineare Abbildung mit $\lim\limits_{h \to 0} \frac{f(x_o+h)-f(x_o)-L(h)}{h} = 0$. Dann folgt

$$0 = \lim_{h \to 0} \frac{f(x_o + h) - f(x_o) - c \cdot h}{h} = \lim_{h \to 0} \frac{f(x_o + h) - f(x_o)}{h} - c,$$

also

$$f'(x_o) = \lim_{h \to 0} \frac{f(x_o + h) - f(x_o)}{h} = c.$$

So erhalten wir aus (iii) die Aussage (i). □

Ist f in x_o differenzierbar, dann heißt die affin-lineare Abbildung

$$T_{x_o} : I \to \mathbb{C}, \quad x \mapsto f(x_o) + L(x - x_o) = f(x_o) + f'(x_o)(x - x_o)$$

(affin) lineare Approximation von f in x_o. Ist f reellwertig, dann ist T_{x_o} die Tangente am Graphen von f im Punkt $(x_o, f(x_o))$.

Aus der Charakterisierung von Differenzierbarkeit in x_o, welche in Satz 4.6 (ii) gegeben ist, folgt sofort:

Korollar 4.7. *Ist $f : I \to \mathbb{C}$ in $x_o \in I$ differenzierbar, dann ist f in x_o stetig.*

Beweis. Da f in x_o differenzierbar ist, gilt

$$f(x) = (x - x_o) \cdot \phi(x) + f(x_o)$$

für eine in x_o stetige Funktion $\phi : I \to \mathbb{C}$ (Satz 4.6 (ii)). Damit folgt

$$\lim_{x \to x_o} f(x) = \lim_{x \to x_o} \underbrace{(x - x_o) \cdot \phi(x)}_{\to 0 \cdot \phi(x_o) = 0} + f(x_o) = f(x_o).$$

Also ist f stetig in x_o. □

Bemerkung 4.8. Die Umkehrung von Korollar 4.7 ist falsch: Eine in x_o stetige Funktion $f : I \to \mathbb{C}$ muss nicht in x_o differenzierbar sein, wie folgende Beispiele zeigen:

(a) $f : R \to \mathbb{R},\ x \mapsto |x|$ ist in $x_o = 0$ stetig, aber nicht differenzierbar.

(b) $f : R \to \mathbb{R},\ x \mapsto \begin{cases} x \cdot \sin\left(\frac{1}{x}\right), & \text{für} \quad x \neq 0, \\ 0, & \text{für} \quad x = 0 \end{cases}$ ist in $x_o = 0$ stetig, aber nicht differenzierbar.

Ja, es gibt sogar Funktionen $f : \mathbb{R} \to \mathbb{R}$, welche auf ganz $\mathbb{R}$ stetig sind, aber in *keinem* Punkt von $\mathbb{R}$ differenzierbar sind (siehe z. B. [K1, Abschn. 3.11]).

Satz 4.9. *Seien $f, g : I \to \mathbb{C}$ in x_o differenzierbare Funktionen und $\lambda \in \mathbb{C}$, dann gilt:*

(i) $f + g$ ist in x_o differenzierbar mit

$$(f+g)'(x_o) = f'(x_o) + g'(x_o).$$

(ii) $\lambda \cdot f$ ist in x_o differenzierbar mit

$$(\lambda \cdot f)'(x_o) = \lambda \cdot f'(x_o).$$

Der Beweis ergibt sich einfach aus den Eigenschaften von Grenzwerten.

Die Mengen $\{f : I \to \mathbb{C} :\ f \text{ differenzierbare Funktion}\}$ und $\{f : I \to \mathbb{C} :\ f \text{ Funktion}\}$ sind bzgl. der punktweisen Addition und skalaren Multiplikation komplexe Vektorräume und die Abbildung

$$\{f : I \to \mathbb{C} :\ f \text{ differenzierbare Funktion}\} \to \{f : I \to \mathbb{C} :\ f \text{ Funktion}\}, \quad f \mapsto f'$$

ist $\mathbb{C}$-linear. Dieses besagt gerade der Satz 4.9.

Satz 4.10 (Produkt– und Quotientenregel)**.** *Seien $f, g : I \to \mathbb{C}$ in x_o differenzierbare Funktion und $\lambda \in \mathbb{C}$, dann gilt:*

(i) $f \cdot g$ ist in x_o differenzierbar mit

$$(f \cdot g)'(x_o) = f'(x_o) \cdot g(x_o) + f(x_o) \cdot g'(x_o) \qquad \textit{(Produktregel).}$$

(ii) Ist $g(x) \neq 0$ für alle $x \in I$, dann ist $\frac{f}{g}$ in x_o differenzierbar mit

$$\left(\frac{f}{g}\right)'(x_o) = \frac{f'(x_o)g(x_o) - f(x_o)g'(x_o)}{(g(x_o))^2} \qquad \textit{(Quotientenregel).}$$

Beweis. Durch einfache Umformungen sieht man, dass der Differenzenquotient von $f \cdot g$ in x_o als

$$\frac{f(x_o+h) - f(x_o)}{h} g(x_o+h) + \frac{g(x_o+h) - g(x_o)}{h} f(x_o)$$

und der Differenzenquotient von $\frac{f}{g}$ in $x_o$als

$$\frac{1}{g(x_o+h)g(x_o)} \left(\frac{f(x_o+h) - f(x_o)}{h} g(x_o) - \frac{g(x_o+h) - g(x_o)}{h} f(x_o) \right)$$

geschrieben werden kann. Mit $h \to 0$ erhalten wir daraus die Differenzierbarkeitsaussagen und die Gleichung für die Ableitung. □

Definition 4.11. Sei $k \in \mathbb{N}^*$ und $f : I \to \mathbb{C}$. Wir setzen $f^{(0)} := f$ und $f^{(k)} := \left(f^{(k-1)}\right)'$, wenn $f^{(k-1)}$ differenzierbar ist. f heißt *n-mal differenzierbar,* wenn $f^{(n)}$ existiert.

Aufgabe 4.12. Seien $n \in \mathbb{N}^*$, I ein offenes Intervall in $\mathbb{R}$ und $f, g : I \to \mathbb{C}$ zwei n-mal differenzierbare Funktionen. Zeigen Sie $f \cdot g$ ist n-mal differenzierbar mit

$$(f \cdot g)^{(n)} = \sum_{j=0}^{n} \left(\binom{n}{j} f^{(j)} \cdot g^{(n-j)} \right).$$

Satz 4.13 (Kettenregel)**.** *Sei $f : I \to \mathbb{R}$ mit $f(I) \subset J$ in $x_o \in I$ differenzierbar und $g : J \to \mathbb{C}$ in $y_o := f(x_o)$ differenzierbar, dann gilt $g \circ f : I \to \mathbb{C}$ ist in x_o differenzierbar mit*

$$(g \circ f)'(x_o) = g'(f(x_o)) \cdot f'(x_o) \qquad \textit{(Kettenregel).}$$

Beweis. Wir verwenden hier die Charakterisierung (ii) der Differenzierbarkeit in Satz 4.6 und schreiben

$$f(x) - f(x_o) = (x - x_o)\phi(x) \qquad \text{bzw.} \qquad g(y) - g(y_o) = (y - y_o)\psi(y)$$

wobei $\phi : I \to \mathbb{R}$ stetig in x_o und $\psi : J \to \mathbb{C}$ stetig in $\psi(y_o)$ mit $\phi(x_o) = f'(x_o)$ und $\psi(y_o) = g'(y_o) = g'(f(x_o))$ sind.

Wir erhalten durch eine einfache Rechnung

$$(g \circ f)(x) - (g \circ f)(x_o) = (x - x_o) \cdot (\psi \circ f)(x) \cdot \phi(x).$$

Da die Funktion $(\psi \circ f) \cdot \phi$ in x_o stetig ist mit

$$((\psi \circ f) \cdot \phi)(x_o) = g'(f(x_o)) \cdot f'(x_o)$$

folgt die Behauptung mit Satz 4.6. □

Beispiel 4.14. (a) Für $f_n : \mathbb{R} \to \mathbb{R},\ x \mapsto x^n,\ n \in \mathbb{N}$, gilt $f_n'(x) = nx^{n-1}$.

Die Fälle $n = 0, 1$ sind einfach von Hand nachzurechnen. Zum Induktionsschluss erhalten wir aus der Produktformel mit $f_{n+1} = f_n \cdot f_1$ sofort nach Induktionsvoraussetzung

$$f_{n+1}'(x) = f_n'(x) \cdot f_1(x) + f_n(x) \cdot f_1'(x) = nx^{n-1}x + x^n = (n+1)x^n.$$

(b) Sei $n \in \mathbb{N}^*$ und $f_{-n} : \mathbb{R} \setminus \{0\} \to \mathbb{R},\ x \mapsto x^{-n}$, dann gilt

$$f_{-n}'(x) = -nx^{-n-1}.$$

(c) Für die reelle Exponentialfunktion exp gilt $\exp' = \exp$.

Mit Übungsaufgabe 53 (Übungsblatt 11) gilt

$$\frac{\exp(x+h) - \exp(x)}{h} = \exp(x) \cdot \underbrace{\frac{\exp(h) - 1}{h}}_{\xrightarrow{h \to 0} 1} \xrightarrow{h \to 0} \exp(x).$$

(d) Nach der Kettenregel gilt für jedes $c \in \mathbb{C}$ und $f_c : \mathbb{R} \to \mathbb{C},\ x \to \exp(cx)$:

$$f_c'(x) = c\exp(cx).$$

(e) Es gilt $\sin'(x) = \cos(x)$ und $\cos'(x) = -\sin(x)$. Mit $\cos(x) = \frac{1}{2}(\exp(ix) + \exp(-ix))$ erhalten wir aus den Ableitungsregeln

$$\begin{aligned} \cos'(x) &= \frac{1}{2}(i\exp(ix) - i\exp(-ix)) \\ &= -\frac{1}{2i}(\exp(ix) - \exp(-ix)) \\ &= -\sin(x). \end{aligned}$$

Der Beweis für die Ableitung von sin kann analog geführt werden.

Satz 4.15 (Ableitung der Umkehrfunktion)**.** *Sei $f : I \to \mathbb{R}$ eine streng monotone, reellwertige Funktion, welche in $x_o \in I$ differenzierbar ist mit $f'(x_o) \neq 0$. Sei $g : f(I) \to I$ die Umkehrfunktion von f. Dann ist g in $y_o := f(x_o)$ differenzierbar und es gilt*

$$g'(y_o) = \frac{1}{f'(x_o)} = \frac{1}{f'(g(y_o))} \qquad \textit{(Ableitung der Umkehrfunktion).}$$

Beweis. Da f in x_o differenzierbar ist, gibt es nach Satz 4.6 eine Funktion $\phi : I \to \mathbb{R}$, welche in x_o stetig ist mit $\phi(x_o) = f'(x_o) \neq 0$, sodass

$$f(x) - f(x_o) = \phi(x)(x - x_o). \tag{4.1}$$

Die Funktion ϕ hat keine Nullstelle: Für $y \neq y_o$ gilt nämlich $f(y) \neq f(y_o)$ wegen der strengen Monotonie von f, und für $y = y_o$ ist $\phi(x_o) = f'(x_o) \neq 0$.

Setzen wir nun $y = f(x)$ und $x = g(y)$ in Gleichung (4.1) ein, so folgt $y - y_o = \phi(g(y))(g(y) - g(y_o))$ und damit

$$g(y) - g(y_o) = \frac{1}{\phi(g(y))}(y - y_o)$$

für alle $y \in f(I)$. Da f streng monoton und in x_o stetig ist, ist die Umkehrfunktion g in y_o stetig. Weil zudem ϕ in $x_o = \phi(y_o)$ stetig ist und nicht verschwindet, ist auch die Funktion $\frac{1}{\phi \circ g}$ in y_o stetig. Nach Satz 4.6 ist damit g in y_o differenzierbar mit $g'(y_o) = \frac{1}{\phi(g(y_o))} = \frac{1}{\phi(x_o)} = \frac{1}{f'(x_o)}$. □

4.3. Der Mittelwertsatz und seine Anwendungen

Wir orientieren uns in diesem Abschnitt an [B1, Kap. V], [Fo1, § 16] und an [K1, Abschn. 9.3 u. 9.4].

Definition 4.16. Sei $f : I \to \mathbb{R}$ eine Funktion auf einem offenen Intervall $I \subset \mathbb{R}$. Wir sagen f hat in $x_o \in I$ ein

- *lokales Maximum*, wenn gilt:

$$\exists \varepsilon > 0 \ \ \forall x \in I \cap (x_o - \varepsilon, x_o + \varepsilon) : \quad f(x) \le f(x_o).$$

- *echtes (oder isoliertes) lokales Maximum*, wenn gilt:

$$\exists \varepsilon > 0 \ \ \forall x \in (I \setminus \{x_o\}) \cap (x_o - \varepsilon, x_o + \varepsilon) : \quad f(x) < f(x_o).$$

- *lokales Minimum*, wenn gilt:

$$\exists \varepsilon > 0 \ \ \forall x \in I \cap (x_o - \varepsilon, x_o + \varepsilon) : \quad f(x) \ge f(x_o).$$

- *echtes (oder isoliertes) lokales Minimum*, wenn gilt:

$$\exists \varepsilon > 0 \ \ \forall x \in (I \setminus \{x_o\}) \cap (x_o - \varepsilon, x_o + \varepsilon) : \quad f(x) > f(x_o).$$

- *lokales Extremum*, wenn f in x_o ein lokales Maximum oder Minimum hat.

Satz 4.17 (Satz von Fermat[97]). *Sei $f : I \to \mathbb{R}$ eine Funktion, die in $x_o \in I$ ein lokales Extremum hat und in diesem Punkt x_o differenzierbar ist, dann gilt $f'(x_o) = 0$.*

Beweis. Wir beweisen diesen Satz für lokale Minima. Für lokale Maxima ist der Beweis analog.

Für genügend kleine $h \in \mathbb{R}$ gilt dann $f(x_o+h) > f(x_o)$ und somit $f(x_o+h)-f(x_o) \ge 0$. Da f in x_o differenzierbar ist, existiert $f'(x_o) = \lim\limits_{h\to 0} \frac{f(x_o+h)-f(x_o)}{h}$ und es gilt

$$\lim_{h\uparrow 0} \frac{f(x_o+h)-f(x_o)}{h} = \lim_{h\to 0} \frac{f(x_o+h)-f(x_o)}{h} = \lim_{h\downarrow 0} \frac{f(x_o+h)-f(x_o)}{h}.$$

Mit $\lim\limits_{h\uparrow 0} \underbrace{\frac{f(x_o+h)-f(x_o)}{h}}_{\le 0} \le 0$ und $\lim\limits_{h\downarrow 0} \underbrace{\frac{f(x_o+h)-f(x_o)}{h}}_{\ge 0} \ge 0$ erhalten wir $f'(x_o) = \lim\limits_{h\to 0} \frac{f(x_o+h)-f(x_o)}{h} = 0$. □

Im Folgenden seien a und b zwei reelle Zahlen mit $a < b$.

Theorem 4.18 (Satz von Rolle[98]). *Sei $f : [a,b] \to \mathbb{R}$ eine stetige und auf (a,b) differenzierbare Funktion mit $f(a) = f(b)$, dann gibt es ein $\xi \in (a,b)$ mit $f'(\xi) = 0$.*

Beweis. Ist f eine konstante Funktion, dann ist die Aussage des Satzes klar.

Sei f nun nicht konstant. Als stetige Funktion nimmt f auf dem Kompaktum $[a,b]$ (globales) Maximum und ein (globales) Minimum an (Theorem 3.26), welche nicht gleich sind, da f nicht konstant ist. Damit wird mindestens eines dieser Extrema in einem Punkt $\xi \in (a,b)$ angenommen. Nach Satz 4.17 gilt $f'(\xi) = 0$. □

[97] Pierre de Fermat (17. Jahrh.), französischer Jurist und Mathematiker
[98] Michel Rolle (1652-1719), französischer Mathematiker

Theorem 4.19 (Mittelwertsatz der Differentialrechnung)**.** *Sei $f : [a, b] \to \mathbb{R}$, $a < b$, eine stetige und auf (a, b) differenzierbare Funktion. Dann gibt es ein $\xi \in (a, b)$, sodass gilt:*

$$\frac{f(b) - f(a)}{b - a} = f'(\xi)$$

Beweis. Die Hilfsfunktion

$$h : [a, b] \to \mathbb{R}, \quad x \mapsto f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

ist stetig und auf (a, b) differenzierbar und erfüllt $h(b) = f(a) = h(a)$. Nach dem Satz von Rolle (Theorem 4.18) gibt es ein $\xi \in (a, b)$ mit $0 = h'(\xi) = f'(\xi) - \frac{f(b)-f(a)}{b-a}$, also $f'(\xi) = \frac{f(b)-f(a)}{b-a}$. □

Theorem 4.20 (Verallgemeinerter Mittelwertsatz)**.** *Zwei stetige Funktionen $f, g : [a, b] \to \mathbb{R}$ seien auf (a, b) differenzierbar, wobei zudem g' auf (a, b) keine Nullstelle hat. Dann gibt es ein $\xi \in (a, b)$, sodass*[99]

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

Beweis. Wir betrachten die offenbar stetige und auf (a, b) differenzierbare Hilfsfunktion

$$h : [a, b] \to \mathbb{R}, \quad x \mapsto f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)).$$

Diese erfüllt $h(b) = h(a)$. Gemäß des Satzes von Rolle (Theorem 4.18) gibt es also ein $\xi \in (a, b)$ mit $0 = h'(\xi) = f'(\xi) - \frac{f(b)-f(a)}{g(b)-g(a)} g'(\xi)$ und damit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

□

Korollar 4.21 (Schrankensatz)**.** *Sei $f : [a, b] \to \mathbb{R}$, $a < b$, eine stetige und auf (a, b) differenzierbare Funktion. Angenommen es gibt Zahlen $m, M \in \mathbb{R}$, sodass für alle $x \in (a, b)$ gilt.*

$$m \leq f'(x) \leq M,$$

dann gilt

$$m(y - x) \leq f(y) - f(x) \leq M(y - x)$$

für alle $a \leq x \leq y \leq b$. Insbesondere ist f dann auf $[a, b]$ Lipschitz-stetig.

[99] Da g' auf (a, b) keine Nullstelle hat, folgt nach dem Mittelwertstatz $g(b) - g(a) \neq 0$.

Beweis. Für $x = y$ ist die Aussage trivial. Sei also $a \le x < y \le b$. Dann betrachten wir $f|_{[x,y]}$ und erhalten nach dem Mittelwertsatz

$$\frac{f(y) - f(x)}{y - x} = f'(\xi)$$

für ein passendes $\xi \in (x, y)$. Mit $m \le f'(\xi) \le M$ folgt sofort $m \le \frac{f(y)-f(x)}{y-x} \le M$ und damit auch die Aussage. $\square$

Als direkte Konsequenz erhalten wir:

Korollar 4.22. *Ist $f : [a, b] \to \mathbb{R}$ stetig und auf (a, b) differenzierbar, sodass $f' : (a, b) \to \mathbb{R}$ stetig ist, dann ist f auf jedem kompakten Teilintervall von (a, b) Lipschitz–stetig.*

Korollar 4.23. *Jede stetige Funktion $f : [a, b] \to \mathbb{R}$, $a < b$, die auf (a, b) differenzierbar ist und $f'(x) = 0$ für alle $x \in (a, b)$ erfüllt, ist konstant.*

Korollar 4.24 (Ableitung und Monotonie)**.** *Sei $f : [a, b] \to \mathbb{R}$, $a < b$, eine stetige und auf (a, b) differenzierbare Funktion mit Ableitung $f' : (a, b) \to \mathbb{R}$. Dann gilt:*

(i) Wenn $f' > 0$, d. h. wenn $f'(x) > 0$ für alle $x \in (a, b)$, dann ist f streng monoton wachsend.

(ii) f ist genau dann monoton wachsend, wenn $f' \ge 0$, d. h. $f'(x) \ge 0$ für alle $x \in (a, b)$.

(iii) Wenn $f' < 0$, d. h. wenn $f'(x) < 0$ für alle $x \in (a, b)$, dann ist f streng monoton fallend.

(iv) f ist genau dann monoton fallend, wenn $f' \le 0$, d. h. $f'(x) \le 0$ für alle $x \in (a, b)$.

Beweis. Wir zeigen nur die Aussage (i) und (ii). Die Aussagen (iii) und (iv) zeigt man analog oder durch Übergang zu $-f$.

Sei $f' > 0$ und angenommen f sei nicht streng monoton wachsend, dann gibt es $x, y \in [a, b]$ mit $x < y$ und $f(x) \ge f(y)$. Nach dem Mittelwertsatz für $f|_{[x,y]}$ gibt es $\xi \in (x, y)$ mit $0 \ge \frac{f(y)-f(x)}{y-x} = f'(\xi)$, im Widerspruch zu $f'(\xi) > 0$. Das zeigt (i).

Ist f monoton wachsend und auf (a, b) differenzierbar, dann gilt für alle $x \in (a, b)$:

$$f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x} = \lim_{y \downarrow x} \underbrace{\frac{f(y) - f(x)}{y - x}}_{\ge 0} \ge 0.$$

Wenn nun $f' \ge 0$ und f nicht monoton wachsend ist, dann gibt es $x, y \in [a, b]$ mit $x < y$ und $f(x) > f(y)$. Nach dem Mittelwertsatz für $f|_{[x,y]}$ gibt es $\xi \in (x, y)$ mit $0 < \frac{f(y)-f(x)}{y-x} = f'(\xi)$, im Widerspruch zu $f'(\xi) \ge 0$. Das zeigt (ii). $\square$

Bemerkung 4.25. Die Funktion $f : [-1, 1] \to \mathbb{R}$, $x \to x^3$ ist streng monoton wachsend mit $f'(0) = 0$. Damit impliziert streng monoton wachsend nicht eine strikt positive Ableitung. Analog impliziert streng monoton fallend nicht eine strikt negative Ableitung.

Definition 4.26. Sei $f : I \to \mathbb{C}$ eine Funktion, wir definieren rekursiv die n-te Ableitung $f^{(n)}$ von f durch $f^{(0)} = f$ und $f^{(n)} = \left(f^{(n-1)}\right)'$, $n \in \mathbb{N}^*$, wenn $f^{(n-1)}$ existiert und ableitbar ist. Die zweite Ableitung $f^{(2)}$ von f wird oft auch mit f'' bezeichnet.

Eine Funktion $f : I \to \mathbb{C}$ heißt n-mal *stetig differenzierbar,* wenn die n-te Ableitung $f^{(n)}$ von f existiert und auch stetig ist.

Satz 4.27. *Sei $f : I \to \mathbb{R}$ eine differenzierbare Funktion und in $x_o \in I$ ein Punkt mit $f'(x_o) = 0$. Angenommen in x_o existiere sogar die zweite Ableitung von f. Dann gilt:*

(i) Wenn $f''(x_o) > 0$, dann hat f in x_o ein echtes Minimum.

(ii) Wenn $f''(x_o) < 0$, dann hat f in x_o ein echtes Maximum.

Beweis. Wir beweisen nur die Aussage (ii). Die Aussage (i) zeigt man auf ähnliche Weise. Aus $0 > f''(x_o) = \lim\limits_{x \to x_o} \frac{f'(x)-f'(x_o)}{x-x_o} = \lim\limits_{x \to x_o} \frac{f'(x)}{x-x_o}$ schliessen wir, dass es ein $\varepsilon > 0$ gibt, sodass $\frac{f'(x)}{x-x_o} < 0$ für alle $x \in (x_o - \varepsilon, x_o + \varepsilon) \cap (I \setminus \{x_o\})$.

Gilt nun $x \in (x_o, x_o + \varepsilon)$, so ist $x - x_o > 0$ und damit $f'(x) < 0$ Damit ist $f|_{(x_o, x_o+\varepsilon)}$ streng monoton fallend.

Gilt nun $x \in (x_o - \varepsilon, x_o)$, so ist $x - x_o < 0$ und damit $f'(x) > 0$ Damit ist $f|_{(x_o-\varepsilon, x_o)}$ streng monoton wachsend.

Zusammenfassend folgt, dass f in x_o ein echtes Maximum haben muss. □

Wir möchten nun einige Sätze diskutieren, die als *Regeln von de l'Hôpital*[100] bezeichnet werden.

Satz 4.28 (1. Regel von de l'Hôpital). *Seien $f, g : [a,b] \to \mathbb{R}$, $a, b \in \mathbb{R}, a < b$, zwei stetige und auf (a,b) differenzierbare Funktionen mit $f(a) = g(a) = 0$, wobei g' auf (a,b) keine Nullstelle habe. Wenn der Grenzwert $\lim\limits_{x \downarrow a} \frac{f'(x)}{g'(x)}$ (im eigentlichen Sinn) existiert*[101]*, dann existiert auch der Grenzwert $\lim\limits_{x \downarrow a} \frac{f(x)}{g(x)}$ und es gilt*

$$\lim_{x \downarrow a} \frac{f(x)}{g(x)} = \lim_{x \downarrow a} \frac{f'(x)}{g'(x)}.$$

Beweis. Nach dem verallgemeinerten Mittelwertsatz gibt es für jedes $x \in [a,b]$ eine Zahl $\lambda_x \in (0,1)$, sodass

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(a + \lambda_x(x-a))}{g'(a + \lambda_x(x-a))}$$

gilt. Da $\lim\limits_{x \downarrow a} \lambda_x(x-a) = 0$ folgt die Aussage. □

[100] Guillaume François Antoine de L'Hôpital (1661 - 1704), französischer Mathematiker

[101] Wenn dieser Grenzwert nicht existiert, so kann nichts gefolgert werden.

Korollar 4.29 (2. Regel von de l'Hôpital)**.** *Seien $f, g : [a, \infty) \to \mathbb{R}$ zwei stetige und auf (a, ∞) differenzierbare Funktionen mit $\lim\limits_{x\to\infty} f(x) = 0 = \lim\limits_{x\to\infty} g(x)$, wobei g' auf (a, ∞) keine Nullstelle habe. Wenn der Grenzwert $\lim\limits_{x\to\infty} \frac{f'(x)}{g'(x)}$ (im eigentlichen Sinn) existiert*[102]*, dann existiert auch der Grenzwert $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)}$ und es gilt*

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)}.$$

Beweis. Zum Beweis können wir ohne Einschränkung $a > 0$ voraussetzen. Die Aussage folgt dann im Wesentlichen durch die Substitution $x \to \frac{1}{t}$ aus der 1. Regel von de l'Hôpital. Die Details sind eine gute Übung. □

Satz 4.30 (3. Regel von de l'Hôpital)**.** *Seien $f, g : (a, b) \to \mathbb{R}, a \in \mathbb{R}, a < b$, zwei differenzierbare Funktionen mit $\lim\limits_{x\downarrow a} g(x) = \infty = \lim\limits_{x\downarrow a} f(x)$, wobei g' keine Nullstelle habe. Wenn der Grenzwert $\lim\limits_{x\downarrow a} \frac{f'(x)}{g'(x)}$ (im eigentlichen Sinn) existiert*[103]*, dann existiert auch der Grenzwert $\lim\limits_{x\downarrow a} \frac{f(x)}{g(x)}$ und es gilt*

$$\lim_{x\downarrow a} \frac{f(x)}{g(x)} = \lim_{x\downarrow a} \frac{f'(x)}{g'(x)}.$$

Beweis. Sei $L := \lim_{x\downarrow a} \frac{f'(x)}{g'(x)}$ und $\varepsilon > 0$ beliebig. Dann gibt es ein $\delta > 0$ mit $\delta < b - a$, sodass für alle $t \in (a, a + \delta)$ gilt:

$$\left| \frac{f'(t)}{g'(t)} - L \right| < \frac{\varepsilon}{2}.$$

Da $\lim\limits_{x\downarrow a} g(x) = \infty$ können wir zudem $g(t) > 0$ für alle $t \in (a, a + \delta)$ annehmen. Sei nun $y \in (a, a + \delta)$, so folgt für alle $x \in (a, y)$ nach dem verallgemeinerten Mittelwertsatz

$$\left| \frac{f(x) - f(y)}{g(x) - g(y)} - L \right| < \frac{\varepsilon}{2}.$$

Aus

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(y)}{g(x) - g(y)} \cdot \frac{1 - \frac{g(y)}{g(x)}}{1 - \frac{f(y)}{f(x)}}$$

folgt

$$\frac{f(x)}{g(x)} - \frac{f(x) - f(y)}{g(x) - g(y)} = \underbrace{\frac{f(x) - f(y)}{g(x) - g(y)}}_{\text{beschränkt}} \cdot \left(\underbrace{\frac{1 - \frac{g(y)}{g(x)}}{1 - \frac{f(y)}{f(x)}}}_{\xrightarrow{x\downarrow a} 1} - 1 \right) \xrightarrow{x\downarrow a} 0.$$

[102]Wenn dieser Grenzwert nicht existiert, so kann nichts gefolgert werden.
[103]Wenn dieser Grenzwert nicht existiert, so kann nichts gefolgert werden.

Nach eventueller Verkleinerung von $\delta > 0$ können wir also annehmen, dass zudem für alle $x \in (a, a+\delta)$ gilt:

$$\left|\frac{f(x)}{g(x)} - \frac{f(x)-f(y)}{g(x)-g(y)}\right| < \frac{\varepsilon}{2}.$$

Insgesamt gilt also:

$$\left|\frac{f(x)}{g(x)} - L\right| \leq \left|\frac{f(x)}{g(x)} - \frac{f(x)-f(y)}{g(x)-g(y)}\right| + \left|\frac{f(x)-f(y)}{g(x)-g(y)} - L\right| < \varepsilon.$$

□

Satz 4.31. *Sei $c \in \mathbb{R}$ und $f : \mathbb{R} \to \mathbb{R}$ eine differenzierbare Funktion mit $f'(x) = cf(x)$ und für alle $x \in \mathbb{R}$ und $f(0) = 1$. Dann gilt $f(x) = \exp(cx)$ für alle $x \in \mathbb{R}$.*

Beweis. Sei $g : \mathbb{R} \to \mathbb{R}, x \mapsto f(x)\exp(-cx)$. Dann ist g überall differenzierbar und damit auch überall stetig. Wegen

$$g'(x) = f'(x)\exp(-cx) - cf(x)\exp(-cx) = cf(x)\exp(-cx) - cf(x)\exp(-cx) = 0$$

ist g nach Korollar 4.23 auf jedem kompakten Intervall in $\mathbb{R}$ und damit auf ganz $\mathbb{R}$ konstant. Mit $g(0) = f(0)\exp(0) = 1$ folgt $1 = g(x) = f(x)\exp(-cx)$ für alle $x \in \mathbb{R}$ und somit $f(x) = \exp(cx)$. □

Wir erhalten somit eine weitere Charakterisierung der reellen Exponentialfunktion:

Korollar 4.32 (Charakterisierung der Exponentialfunktion)**.** *Die Exponentialfunktion ist die einzige differenzierbare Funktion $f : \mathbb{R} \to \mathbb{R}$, welche $f' = f$ und $f(0) = 1$ erfüllt.*

Wir führen nun noch zur Vollständigkeit folgende üblichen Notationen ein:
Sei $\mathbb{K} = \mathbb{R}, \mathbb{C}$ und $I \subset \mathbb{R}$ ein offenes Intervall sowie $n \in \mathbb{N}$, dann definieren wir

$$\mathcal{C}^n(I, \mathbb{K}) = \mathcal{C}^n(I) := \left\{f : I \to \mathbb{K} :\ f^{(n)} \text{ existiert und ist stetig}\right\}.$$

Die Menge $\mathcal{C}^0(I, \mathbb{K})$ ist also die Menge der stetigen $\mathbb{K}$-wertigen Funktionen auf I. Offenbar gilt $\mathcal{C}^{n+1}(I, \mathbb{K}) \subset \mathcal{C}^n(I, \mathbb{K})$ für alle $n \in \mathbb{N}$. Wir setzen

$$\mathcal{C}^\infty(I, \mathbb{K}) = \mathcal{C}^\infty(I) := \bigcap_{n=0}^{\infty} \mathcal{C}^n(I, \mathbb{K}) = \{f : I \to \mathbb{K} :\ f \text{ beliebig oft stetig differenzierbar}\}.$$

Die Mengen $\mathcal{C}^j(I, \mathbb{K}),\ j \in \mathbb{N} \cup \{\infty\}$ sind bzgl. der natürlichen punktweisen Addition und skalaren Multiplikation (unendlich dimensionale) $\mathbb{K}$-Vektorräume. Bezüglich der punktweisen Multiplikation sind sie sogar Algebren.

Aufgabe 4.33.

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} \exp(-1/x), & \text{wenn} \quad x > 0 \\ 0, & \text{wenn} \quad x \leq 0 \end{cases}$$

(a) Zeigen Sie: Es gibt polynomiale Funktionen $p_n : \mathbb{R} \to \mathbb{R}$, sodass für alle $n \in \mathbb{N}$ gilt[104]:

$$f^{(n)}(x) = \begin{cases} p_n(1/x)\exp(-1/x), & \text{wenn} \quad x > 0 \\ 0, & \text{wenn} \quad x \leq 0 \end{cases}$$

(b) Skizzieren Sie den Graphen von f.

(c) Zeigen Sie: Für $F : \mathbb{R} \to \mathbb{R},\ x \mapsto \exp(e) \cdot f\big(f(1) - f(1-x)\big)$ gilt:

  (i) F ist beliebig oft differenzierbar.

  (ii) $F|_{(-\infty,0]} = 0$ und $F|_{[1,\infty)} = 1$.

  (iii) $F|_{[0,1]}$ ist streng monoton wachsend.

(d) Skizzieren Sie den Graphen von F.

(e) Zeigen Sie: Sei $a < b$ und $\varepsilon > 0$. Es gibt eine beliebig oft differenzierbare Funktion $h : \mathbb{R} \to \mathbb{R}$, sodass

$$h(x) \begin{cases} = 1, & \text{wenn} \quad x \in [a,b], \\ = 0, & \text{wenn} \quad x \notin [a-\varepsilon, b+\varepsilon], \\ \in [0,1], & \text{sonst.} \end{cases}$$

Hinweis: Skizzieren Sie einen möglichen Graphen von h und verwenden Sie F aus Teil (b) um eine passende Funktion h anzugeben.

4.4. Ableitungen von Reihen

In diesem Abschnitt orientieren wir uns an [K1, Abschn. 9.5] und [Fo1, § 21].

Satz 4.34. *Gegeben seien Funktionen $f_n : I \to \mathbb{C}$, $n \in \mathbb{N}$, und ein Punkt $x_o \in I$. Wir nehmen an, dass gilt:*

(i) $\sum f_n$ konvergiert punktweise auf I ,

(ii) für alle $n \in \mathbb{N}$ ist $f_n : I \to \mathbb{C}$ Lipschitz-stetig mit Lipschitz-Konstanten L_n, sodass $\sum L_n$ konvergiert.

(iii) f_n ist für alle $n \in \mathbb{N}$ in x_o differenzierbar und

(iv) $\sum f_n'(x_o)$ konvergiert

Dann ist $f := \sum f_n : I \to \mathbb{C},\ x \mapsto \sum_{j=0}^{\infty} f_j(x)$ in x_o differenzierbar mit

$$f'(x_o) = \sum_{j=0}^{\infty} f_j'(x_o).$$

[104]Betrachten Sie insbesondere die Ableitungen in $x = 0$

Beweis. Sei $\varepsilon > 0$ beliebig vorgegeben. Wir wählen $N \in \mathbb{N}$, sodass sowohl

$$\sum_{j=N+1}^{\infty} L_n < \frac{\varepsilon}{3}$$

als auch

$$\left|\sum_{j=N+1}^{\infty} f_n'(x_o)\right| < \frac{\varepsilon}{3}.$$

Wir erhalten so für $x \in I \setminus \{x_o\}$:

$$\begin{aligned}
&\left|\frac{f(x)-f(x_o)}{x-x_o} - \sum_{j=0}^{\infty} f_j'(x_o)\right| \\
= \quad &\left|\sum_{j=0}^{\infty} \frac{f_j(x)-f_j(x_o)}{x-x_o} - \sum_{j=0}^{\infty} f_j'(x_o)\right| \\
= \quad &\left|\sum_{j=0}^{\infty} \left(\frac{f_j(x)-f_j(x_o)}{x-x_o} - f_j'(x_o)\right)\right| \\
\leq \quad &\left|\sum_{j=0}^{N} \left(\frac{f_j(x)-f_j(x_o)}{x-x_o} - f_j'(x_o)\right)\right| + \sum_{j=N+1}^{\infty} \underbrace{\left|\frac{f_j(x)-f_j(x_o)}{x-x_o}\right|}_{\leq L_j} + \left|\sum_{j=N+1}^{\infty} f_j'(x_o)\right| \\
\leq \quad &\left|\sum_{j=0}^{N} \frac{f_j(x)-f_j(x_o)}{x-x_o} - \sum_{j=0}^{N} f_j'(x_o)\right| + \frac{2}{3}\varepsilon.
\end{aligned}$$

Da die endliche Summe $\sum\limits_{j=0}^{N} f_j$ in x_o ableitbar ist, mit $\left(\sum\limits_{j=0}^{N} f_j\right)'(x_o) = \sum\limits_{j=0}^{N} f_j'(x_o)$, gibt es ein $\delta > 0$, sodass für alle $x \in (I \setminus \{x_o\}) \cap (x_o - \delta,\ x_o + \delta)$

$$\left|\sum_{j=0}^{N} \frac{f_j(x)-f_j(x_o)}{x-x_o} - \sum_{j=0}^{N} f_j'(x_o)\right| \leq \frac{\varepsilon}{3}.$$

und somit

$$\left|\frac{f(x)-f(x_o)}{x-x_o} - \sum_{j=0}^{\infty} f_j'(x_o)\right| < \varepsilon$$

gilt. Damit folgt die Aussage. $\square$

Korollar 4.35. *Gegeben seien Funktion $f_n : I \to \mathbb{R}$, $n \in \mathbb{N}$, sodass gilt:*

- *für alle $n \in \mathbb{N}$ ist f_n ableitbar,*
- *$\sum f_n$ konvergiert auf I punktweise,*

- $\sum \|f_n'\|_{\sup}$ *konvergiert.*

Dann ist $f := \sum f_n : I \to \mathbb{R},\ x \mapsto \sum\limits_{j=0}^{\infty} f_j(x)$ *differenzierbar mit* $f' = \sum_{j=0}^{\infty} f_j'$.

Beweis. Die Voraussetzungen von Satz 4.34 sind offenbar in jedem Punkt $x_o \in I$ erfüllt, denn es gilt:

- Wenn $\sum \|f_n\|_{\sup}$ konvergiert, dann konvergiert auch $\sum\limits_{j=0}^{\infty} f_j'(x_o)$ (Majoranten-Kriterium).
- Wenn $\sum \|f_n\|_{\sup}$ konvergiert, dann ist (vgl. Schrankensazu (Korollar 4.21)) $L_n := \|f_n\|_{\sup}$ für jedes $n \in \mathbb{N}$ eine Lipschitz-konstante für f_n.

$\square$

Satz 4.36 (Ableitung von Potenzreihen 1)**.** *Sei* (a_n) *eine reelle Folge, sodass die Potenzreihe* $\sum a_n x^n$ *Konvergenzradius* $R > 0$ *hat. Dann ist die Funktion*

$$f : (-R, R) \to \mathbb{R},\ x \mapsto \sum_{j=0}^{\infty} a_j x^j$$

auf dem ganzen Intervall $(-R, R)$ *differenzierbar mit*

$$f'(x) = \sum_{j=0}^{\infty} j a_j x^{j-1} = \sum_{j=1}^{\infty} j a_j x^{j-1} = \sum_{j=0}^{\infty} (j+1) a_{j+1} x^j.$$

Beweis. Wir zeigen zunächst, dass die Potenzreihe $\sum \underbrace{(n+1)a_{n+1}}_{:=b_n} x^n$ ebenso Konvergenzradius R hat. Dieses ergibt sich sofort aus dem Quotientenkriterium:

$$\left|\frac{b_n}{b_{n+1}}\right| = \left|\frac{(n+1)a_{n+1}}{(n+2)a_{n+2}}\right| = \frac{1+\frac{1}{n}}{1+\frac{2}{n}} \cdot \left|\frac{a_{n+1}}{a_{n+2}}\right| \longrightarrow R.$$

Auf jedem Intervall $[-r, r]$ mit $0 < r < R$ konvergiert somit die Reihe $\sum \|(n+1)a_{n+1}x^n\|_{\sup}$ (vgl. Korollar 3.44). Aus Korollar 4.35 folgt sofort die Aussage. $\square$

Korollar 4.37 (Ableitung von Potenzreihen 2)**.** *Sei* (a_n) *eine reelle Folge, sodass die Potenzreihe* $\sum a_n (x - x_o)^n$ *Konvergenzradius* $R > 0$ *hat. Dann ist die Funktion*

$$f : (x_o - R, x_o + R) \to \mathbb{R},\ x \mapsto \sum_{j=0}^{\infty} a_j (x - x_o)^j$$

auf dem ganzen Intervall $(x_o - R, x_o + R)$ *differenzierbar mit*

$$f'(x) = \sum_{j=0}^{\infty} j a_j (x - x_o)^{j-1} = \sum_{j=1}^{\infty} j a_j (x - x_o)^{j-1} = \sum_{j=0}^{\infty} (j+1) a_{j+1} (x - x_o)^j.$$

Durch einfache Rechnung erhalten wir aus dem Identitätssatz für Potenzreihen (Satz 2.74):

Korollar 4.38 (Ableitung von Potenzreihen 3)**.** *Sei* (a_n) *eine reelle Folge, sodass die Potenzreihe* $\sum a_n(x-x_o)^n$ *Konvergenzradius* $R>0$ *hat. Dann ist die Funktion*

$$f:(x_o-R,x_o+R)\to\mathbb{R},\ x\mapsto\sum_{j=0}^{\infty}a_j(x-x_o)^j$$

auf dem ganzen Intervall (x_o-R,x_o+R) *beliebig oft differenzierbar. Zudem gilt*

$$a_n=\frac{f^{(n)}(x_o)}{n!}.$$

4.5. Logarithmus und allgemeine Exponentialfunktionen

Wir orientieren uns in diesem Abschnitt an [Fo1, §12]

Nach den letzten beiden Aussagen von Satz 3.57 ist die Abbildung $\exp:\mathbb{R}\to\mathbb{R}_{>0}$ bijektiv.

Definition 4.39. Die Umkehrabbildung

$$\ln:\mathbb{R}_{>0}\to\mathbb{R}$$

von $\exp:\mathbb{R}\to\mathbb{R}_{>0}$ wir *natürlicher Logarithmus* genannt.

Aufgabe 4.40. (a) Zeigen Sie: Für alle $x>0$ gilt $\ln'(x)=\frac{1}{x}$.

(b) Zeigen Sie: Der Logarithmus ln ist streng monoton wachsend.

(c) Zeigen Sie: Der Logarithmus ln hat genau eine Nullstelle in $x=1$.

(d) Zeigen Sie: Für alle $x,y\in\mathbb{R}_{>0}$ gilt $\ln(xy)=\ln(x)+\ln(y)$.

Definition 4.41 (Exponentialfunktion zu einer anderen Basis)**.** Sei $b>0$, dann heißt die Funktion

$$\exp_b:\mathbb{R}\to\mathbb{R},\quad x\mapsto\exp_b(x):=\exp(x\ln(b))$$

Exponentialfunktion zur Basis b.

Beobachtung 4.42.
- *Als Verknüpfung differenzierbarer Funktionen ist* $\exp_b$ *offenbar wieder differenzierbar.*
- *Aus der Funktionalgleichung der Exponentialfunktion* exp *folgt*

$$\exp_b(x+y)=\exp_b(x)\exp_b(y)$$

für alle $x,y\in\mathbb{R}$.

Aufgabe 4.43. Sei $b > 0$.

(a) Sei $x \in \mathbb{R}$. Zeigen Sie (z. B. durch vollständige Induktion): Für alle $n \in \mathbb{N}^*$ gilt $\exp_b(nx) = (\exp_b(x))^n$.

(b) Zeigen Sie: Für alle $k \in \mathbb{Z}$ gilt: $\exp_b(k) = b^k$.

(c) Zeigen Sie: Für alle $m \in \mathbb{Z}$ und alle $n \in \mathbb{N}^*$ gilt $\exp_b\left(\frac{m}{n}\right) = (\sqrt[n]{b})^m$.

Motiviert duch die Gleichung $\exp_b\left(\frac{m}{n}\right) = (\sqrt[n]{b})^m$ für alle $m \in \mathbb{Z}$ und alle $n \in \mathbb{N}^*$ setzen wir für $b > 0$ und $x \in \mathbb{R}$:

$$b^x := \exp_b(x) = \exp(x \ln(b)). \tag{4.2}$$

Aufgabe 4.44. Beweisen Sie: Seien $a, b > 0$ und $x, y \in \mathbb{R}$, dann gilt:

(a) $b^{x+y} = b^x \cdot b^y$.

(b) $\ln(b^x) = x \ln(b)$.

(c) $(b^x)^y = b^{(x \cdot y)}$.

(d) $a^x \cdot b^x = (a \cdot b)^x$.

(e) $\left(\frac{1}{b}\right)^x = b^{-x}$.

4.6. Konvexe Funktionen und wichtige Ungleichungen

In diesem Abschnitt folgen wir [Fo1, § 16] und [K1, Abschn. 9.7 u. 9.8].

Definition 4.45. Eine Funktion $f : I \to \mathbb{R}$ auf einem (offenen) Intervall I heißt *konvex*, wenn gilt:

$$\forall x_1, x_2 \in I \; \forall \lambda \in (0,1) : \quad f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2).$$

Eine Funktion f heißt *konkav*, wenn $-f$ konvex ist.

Bemerkung 4.46.

- Linearkombinationen der Form $ax_1 + bx_2$ mit $a + b = 1$ heißen Konvexkombinationen.
- Ist $x_1 < x_2$ und f konvex, dann liegt der Graph von $f|_{[x_1,x_2]}$ unter der Geraden durch die Punkte $(x_1, f(x_1))$ und $(x_2, f(x_2))$.

Aufgabe 4.47. Zeigen Sie: Eine Funktion $f : I \to \mathbb{R}$ ist genau dann konkav, wenn gilt:

$$\forall x_1, x_2 \in I \; \forall \lambda \in (0,1) : \quad f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda) f(x_2).$$

Aufgabe 4.48. Konvexe und konkave Funktionen $f : I \to \mathbb{R}$ sind immer auch stetig.

Satz 4.49. *Für eine zweimal differenzierbare Funktion $f : I \to \mathbb{R}$ sind folgende Aussage äquivalent:*

(i) *f ist konvex.*

(ii) $f'' \geq 0$.

Beweis. • *(i)* $\implies$ *(ii)* : Sei f und konvex und $x_o \in I$ mit $f''(x_o) < 0$. Wir betrachten die zweimal differenzierbare Funktion

$$h : I \to \mathbb{R}, \quad x \mapsto f(x) - f'(x_o)(x - x_o).$$

Dann gilt offenbar $h'(x_o) = 0$ und $h''(x_o) = f''(x_o) < 0$ und h hat gemäß Satz 4.27 in x_o ein echtes lokales Maximum. Somit gibt es ein $\delta > 0$, sodass gilt:

– $[x_o - \delta, x_o + \delta] \subset I$ und

– $h(y) < h(x_o)$ für alle $y \in [x_o - \delta, x_o + \delta] \setminus \{x_o\}$.

Wir erhalten insbesondere:

$$f(x_o) = h(x_o) > \frac{1}{2}\Big(h(\underbrace{x_o - \delta}_{=:x_1}) + h(\underbrace{x_o + \delta}_{=:x_2})\Big) = \frac{1}{2}\left(f(x_1) + f(x_2)\right).$$

Dieses widerspricht jedoch der Konvexität von f, denn mit $\lambda = \frac{1}{2}$ erhalten wir $x_o = \lambda x_1 + (1 - \lambda)x_2$ und somit

$$f\Big(\lambda x_1 + (1 - \lambda)x_2\Big) = f(x_o) > \frac{1}{2}\left(f(x_1) + f(x_2)\right) = \lambda f(x_1) + (1 - \lambda)f(x_2).$$

• *(ii)* $\implies$ *(i)* : Wenn $f'' \geq 0$, dann ist f' monoton wachsend. Seien nun $x_1, x_2 \in I$ mit $x_1 < x_2$ und $\lambda \in (0, 1)$. Wir setzen $x_1 < x = \lambda x_1 + (1 - \lambda)x_2 < x_2$. Nach dem Mittelwertsatz gibt es $\xi_1 \in (x_1, x)$ und $\xi_2 \in (x, x_2)$, sodass

$$\frac{f(x_1) - f(x)}{x_1 - x} = f'(\xi_1) \leq f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x}.$$

Mit $x - x_1 = (1 - \lambda)(x_2 - x_1)$ und $x_2 - x = \lambda(x_2 - x_1)$ erhalten wir

$$\frac{f(x) - f(x_1)}{1 - \lambda} \leq \frac{f(x_2) - f(x)}{\lambda},$$

also

$$f(\lambda x_1 + (1 - \lambda)x_2) = f(x) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

□

Korollar 4.50. *Für eine zweimal differenzierbare Funktion $f : I \to \mathbb{R}$ sind folgende Aussage äquivalent:*

(i) *f ist konkav.*

(ii) $f'' \leq 0$.

Definition 4.51. Sei $p \in \mathbb{R}_{\geq 1}$. Die Abbildung

$$\| \dots \|_p : \mathbb{C}^n \to \mathbb{R}, \quad z = (z_1, \dots, z_n) \mapsto \|z\|_p := \left(\sum_{j=1}^{n} |z_j|^p \right)^{\frac{1}{p}}.$$

heißt *p-Norm.*

Bemerkung 4.52. Die p-Norm ist eine Norm im Sinne der Definition (siehe Fußnote 94). Die Dreicksungleichung für die p-Norm werden wir weiter unten nachweisen (siehe Theorem 4.55). Sie heißt Minkowski-Ungleichung.

Für $p = 2$ ist die 2-Norm gerade die übliche Euklidische Norm.

Als Anwendung der Konvexität zeigen wir die Hölder[105]- und die Minkowski[106]-Ungleichung im $\mathbb{C}^n$. Diese Ungleichungen verallgemeinern sich auf L^p-Räume.

Satz 4.53 (Young[107]-Ungleichung)**.** *Seien $x, y \geq 0$ beliebig und $p, q > 1$ zwei Zahlen mit $\frac{1}{p} + \frac{1}{q} = 1$, dann gilt*

$$x^{\frac{1}{p}} y^{\frac{1}{q}} \leq \frac{x}{p} + \frac{y}{q}.$$

Beweis. Ist $x = 0$ oder $y = 0$ so ist die Aussage offenbar wahr.

Sei also $x, y > 0$. Wegen $\ln''(x) = -\frac{1}{x^2} < 0$ ist der natürliche Logarithmus offenbar konkav. Somit gilt

$$\ln\left(\frac{1}{p}x + \frac{1}{q}y\right) \geq \frac{1}{p}\ln(x) + \frac{1}{q}\ln(y) = \ln\left(x^{\frac{1}{p}} y^{\frac{1}{q}}\right).$$

Da die Exponentialfunktion streng monoton wächst folgt

$$\frac{x}{p} + \frac{y}{q} = \exp\left(\ln\left(\frac{1}{p}x + \frac{1}{q}y\right)\right) \geq \exp\left(\ln\left(x^{\frac{1}{p}} y^{\frac{1}{q}}\right)\right) = x^{\frac{1}{p}} y^{\frac{1}{q}}.$$

□

Theorem 4.54 (Hölder-Ungleichung[108])**.** *Seien $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$ und $z = (z_1, \dots, z_n)$, $w = (w_1, \dots, w_n) \in \mathbb{C}^n$, dann gilt*

$$\sum_{j=1}^{n} |z_j w_j| \leq \|z\|_p \|w\|_q.$$

[105] Otto Ludwig Hölder (1859–1937), deutscher Mathematiker

[106] Hermann Minkowski (1864–1909), deutscher Mathematiker

[107] William Henry Young (1863 – 1942), englischer Mathematiker

[108] Diese Hölder-Ungleichung verallgemeinert die Ungleichung von Cauchy–Schwarz für das hermitesche Standardskalarprodukt auf $\mathbb{C}^n$. Für $p = q = 2$ ist diese Hölder-Ungleichung gerade die Cauchy–Schwarz-Ungleichung.

Beweis. Wenn $\|z\|_p = 0$ oder $\|y\|_q = 0$, so ist die Aussage klar. Sei also $\|x\|_p \neq 0$ und $\|y\|_q \neq 0$, setzen wir für $j = 1, \dots n$:

$$\zeta_j = \frac{|z_j|^p}{(\|z\|_p)^p} \quad \text{und} \quad \omega_j = \frac{|w_j|^q}{(\|w\|_q)^q}.$$

Wir erhalten nun mit der Young-Ungleichung für alle $j \in \{1, \dots, n\}$:

$$\frac{|z_j w_j|}{\|z\|_p \|w\|_q} = (\zeta_j)^{\frac{1}{p}} (\omega_j)^{\frac{1}{q}} \leq \frac{\zeta_j}{p} + \frac{\omega_j}{q}.$$

Da $\sum\limits_{j=1}^{n} \zeta_j = 1 = \sum\limits_{j=1}^{n} \omega_j$, erhalten wir

$$\frac{1}{\|z\|_p \|w\|_q} \sum_{j=1}^{n} |z_j w_j| \leq \sum_{j=1}^{n} \frac{\zeta_j}{p} + \sum_{j=1}^{n} \frac{\omega_j}{q} = \frac{1}{p} + \frac{1}{q} = 1,$$

also

$$\sum_{j=1}^{n} |z_j w_j| \leq \|z\|_p \|w\|_q.$$

□

Theorem 4.55 (Minkowski-Ungleichung)**.** *Die p-Norm,* $p \geq 1$, *ist subadditiv, d. h. sei* $p \geq 1$ *und* $z, w \in \mathbb{C}^n$ *beliebig, dann gilt*

$$\|z + w\|_p \leq \|z\|_p + \|w\|_p \qquad \textit{(Dreiecksungleichung für die p-Norm).}$$

Beweis.

- $p = 1$: In diesem Fall folgt die Aussage einfach aus der Dreiecksungleichung.
- Sei $p > 1$ gegeben und $q > 1$, sodass $\frac{1}{p} + \frac{1}{q} = 1$, also $q(p-1) = p$. Sei $u = (u_1, \dots, u_n) \in \mathbb{R}^n \subset \mathbb{C}^n$ mit $u_j := |z_j + w_j|^{p-1} \geq 0$ für $j \in \{1, \dots, n\}$. Dann gilt $u_j^q = |z_j + w_j|^{q(p-1)} = |z_j + w_j|^p$. Mit der Dreiecksungleichung sowie der Hölderschen Ungleichung folgt:

$$\begin{aligned} (\|z + w\|_p)^p &= \sum_{j=1}^{n} (|z_j + w_j|)^p = \sum_{j=1}^{n} (|z_j + w_j| \cdot |z_j + w_j|^{p-1}) = \sum_{j=1}^{n} (|z_j + w_j| \cdot u_j) \\ &\leq \sum_{j=1}^{n} (|z_j| \cdot u_j + |w_j| \cdot u_j) = \sum_{j=1}^{n} (|z_j u_j|) + \sum_{j=1}^{n} (|w_j u_j|) \\ &\leq \|z\|_p \|u\|_q + \|w\|_p \|u\|_q = (\|z\|_p + \|w\|_p) \|u\|_q \\ &= (\|z\|_p + \|w\|_p)(\|z + w\|_p)^{\frac{p}{q}}. \end{aligned}$$

Wir erhalten die Aussage da $1 = p - \frac{p}{q}$, wobei der Fall $z = -w$ sofort einsichtig ist.

□

4.7. Taylor-Polynome und Taylor-Reihen

Wir folgen in diesem Abschnitt der Darstellung in [BF1, Abschn. 8.4].

Durch die Ableitung haben wir eine Funktion lokal durch eine affin lineare Funktion, das sind polynomiale Funktionen von Grad höchstens 1, approximiert. In diesem Abschnitt untersuchen wir, wie man Funktionen lokal durch passende polynomiale Funktionen von höherem Grad approximieren kann.

Definition 4.56. Sei $f : I \to \mathbb{R}$ n-mal differenzierbar, $k \in \{0, 1, \ldots, n\}$ und $x_o \in I$. Die polynomiale Funktion

$$T_k = T_k^{f,x_o} : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \sum_{j=0}^{k} \frac{f^{(j)}(x_o)}{j!}(x - x_o)^j$$

heißt *Taylor*[109]*–Polynom* von f von Ordnung[110] k in x_o.[111]

Bemerkung 4.57. Durch eine einfache Rechnung verifiziert man, dass für eine n-mal stetig differenzierbare Funktion f und ihr Taylor–Polynom T_k^{f,x_o}, $0 \le k \le n$ gilt für die Ableitungen in x_o:

$$\left(T_k^{f,x_o}\right)^{(j)}(x_o) = f^{(j)}(x_o)$$

für alle $j \in \{0, \ldots, n\}$.

Eine wichtige Frage ist, wie gut approximiert ein Taylor–Polynom die dazugehörige Funktion. Dazu betrachten wir die Differenz, welche Restglied genannt wird:

Definition 4.58. Sei $f : I \to \mathbb{R}$ n-mal differenzierbar, $k \in \{0, 1, \ldots, n\}$ und $x_o \in I$. Die Funktion

$$R_k = R_k^{f,x_o} : I \to \mathbb{R}, \qquad x \mapsto f(x) - T_k^{f,x_o}(x)$$

heißt *Restglied* zum Taylor–Polynom T_k^{f,x_o}

Satz 4.59 (Taylor–Formel mit Lagrange[112]–Restglied[113].)**.** *Sei $f : I \to \mathbb{R}$ $(n+1)$-mal differenzierbar und $x_o, x_1 \in I$ mit $x_1 \neq x_o$, dann gibt es ein ξ zwischen*[114] *x_o und x_1, sodass*

$$R_n^{f,x_o}(x_1) = f(x_1) - T_n^{f,x_o}(x_1) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x_1 - x_o)^{n+1}.$$

Diese Formel für das Restglieds heißt Lagrange–Form *des Restglieds oder kurz* Lagrange–Restglied.

[109] Brook Taylor (1685 – 1731), englischer Mathematiker

[110] Ist $f^{(k)}(x_o) = 0$, so stimmt die Ordnung nicht mit dem Grad der polynomialen Funktion überein.

[111] Manchmal wird T_k^{f,x_o} auch als *k-Jet* von f in x_o bezeichnet.

[112] Joseph–Louis de Lagrange (1736 – 1813), italienisch-französischer Mathematiker

[113] Eine andere Form des Restglieds, die sogenannte Integralform, finden Sie in Satz 5.48

[114] d. h. wenn $x_o < x_1$, dann gilt $\xi \in (x_o, x_1)$, wenn aber $x_o > x_1$, dann gilt $\xi \in (x_1, x_o)$.

Beweis. Wir betrachten die Hilfsfunktion

$$h : I \to \mathbb{R}, \quad x \mapsto f(x_1) - T_n^{f,x}(x_1) - c\frac{(x_1 - x)^{n+1}}{(n+1)!},$$

also

$$h(x) = f(x_1) - f(x) - f'(x)(x_1 - x) - \frac{f''(x)}{2}(x_1 - x)^2 - \cdots - \frac{f^{(n)}(x)}{n!}(x_1 - x)^n - c\frac{(x_1 - x)^{n+1}}{(n+1)!},$$

wobei $c \in \mathbb{R}$ so gewählt ist, dass $h(x_0) = 0$. Da $h(x_1) = 0$ gilt, und h nach den Voraussetzungen differenzierbar ist, gibt es nach dem Satz von Rolle (siehe Theorem 4.18) ein ξ zwischen x_o und x_1, sodass gilt:

$$0 = h'(\xi) = -\frac{f^{(n+1)}(\xi)}{n!}(x_1 - \xi)^n + c\frac{(x_1 - \xi)^n}{n!},$$

also $c = f^{(n+1)}(\xi)$. Hierbei haben wir die Ableitung $h'(x) = -\frac{f^{(n+1)}(x)}{n!}(x_1 - x)^n + c\frac{(x_1 - x)^n}{n!}$ explizit ausgerechnet. Wir erhalten also

$$0 = h(x_o) = f(x_1) - T_n^{f,x_o}(x_1) - f^{(n+1)}(\xi)\frac{(x_1 - x_o)^{n+1}}{(n+1)!},$$

d. h.

$$R_n^{f,x_o}(x_1) := f(x_1) - T_n^{f,x_o}(x_1) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x_1 - x_o)^{n+1}.$$

□

Bemerkung 4.60. Für $n = 0$ ist die Lagrange–Form des Restglieds der Mittelwertsatz.

Satz 4.61. *Sei $f \in \mathcal{C}^{n+1}(I, \mathbb{R})$ und $x_o \in I$ mit*

$$f'(x_o) = f''(x_o) = \cdots = f^{(n)}(x_o) = 0 \quad \textit{und} \quad f^{(n+1)}(x_o) \neq 0.$$

Dann gilt:

(i) Ist n gerade, dann hat f in x_o kein Extremum.

(ii) Ist n ungerade, so gilt:
- *Wenn $f^{(n+1)}(x_o) > 0$, dann hat f in x_o ein isoliertes Minimum.*
- *Wenn $f^{(n+1)}(x_o) < 0$, dann hat f in x_o ein isoliertes Maximum.*

Beweis. Nach der Taylor–Formel mit Lagrange–Restglied gilt:

$$f(x) = f(x_o) + \frac{f^{(n+1)}(\xi_x)}{(n+1)!}(x - x_o)^{n+1},$$

wobei ξ_x zwischen x und x_o liegt. Da $f^{(n+1)}$ stetig ist mit $f^{(n+1)}(x_o) \neq 0$ und weil I offen ist, gibt es ein $\delta > 0$, sodass $(x_o - \delta, x_o + \delta) \subset I$ und $f^{(n+1)}(x) \neq 0$ für alle

$x \in (x_o - \delta, x_o + \delta)$. Somit gilt entweder $f^{(n+1)}|_{(x_o-\delta,x_o+\delta)} > 0$ oder $f^{(n+1)}|_{(x_o-\delta,x_o+\delta)} < 0$, je nach Vorzeichen von $f^{(n+1)}(x_o)$. Betrachten wir nun die Gleichung

$$f(x) - f(x_o) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!}(x - x_o)^{n+1}$$

mit ξ_x zwischen x und x_o für $x \in (x_o-\delta, x_o+\delta) \subset I$. Ist n gerade, so wechselt $f(x)-f(x_o)$ im Punkt x_o das Vorzeichen. Also hat f in x_o kein Extremum. Ist n ungerade, so gilt

- $f(x) > f(x_o)$ für alle $x \in (x_o - \delta, x_o + \delta) \setminus \{x_o\}$, wenn $f^{(n+1)}(x_o) > 0$, und $f(x) < f(x_o)$ für alle $x \in (x_o - \delta, x_o + \delta) \setminus \{x_o\}$, wenn $f^{(n+1)}(x_o) < 0$.

Die Aussage folgt nun unmittelbar. $\square$

Bemerkung 4.62.
- Satz 4.61 verallgemeinert Satz 4.27.
- In der nachfolgenden Aufgabe 4.63 wird eine Funktion f vorgestellt, welche ein isoliertes Extremum in einem Punkt hat, in welchem alle Ableitungen von f verschwinden.

Aufgabe 4.63.

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} \exp\left(-\frac{1}{x^2}\right), & \text{wenn} \quad x \neq 0 \\ 0, & \text{wenn} \quad x = 0 \end{cases}$$

Zeigen Sie:

(a) Sei $n \in \mathbb{N}$, dann gibt es polynomiale Funktionen p_n, sodass für alle $n \in \mathbb{N}$ und alle $x \in \mathbb{R} \setminus \{0\}$ gilt:
$$f^{(n)}(x) = p_n(1/x)\exp(-1/x^2).$$
(Vgl. auch Aufgabe 4.33)

(b) f ist beliebig oft differenzierbar.

(c) $f^{(n)}(0) = 0$ für alle $n \in \mathbb{N}$.

(d) f hat in Null ein isoliertes Minimum.

(e) Skizzieren Sie den Graph von f.

Definition 4.64. Sei $f \in \mathcal{C}^\infty(I, \mathbb{R})$ und $x_o \in I$, dann heißt die formale Potenzreihe

$$T(x) = T^{f,x_o}(x) := \sum \frac{f^{(n)}(x_o)}{n!}(x - x_o)^n$$

Taylor–Reihe von f in x_o.[115]

[115] Wenn $x_o = 0$ ist, dann nennt man $T^{f,0}$ auch manchmal *Maclaurin–Reihe*, nach dem schottischen Mathematiker Colin Maclaurin (1698–1746).

Bemerkung 4.65.

- Es gilt $f(x_o) = T^{f,x_o}(x_o)$.
- Die Taylorreihe einer beliebige oft differenzierbaren Funktion $f : I \to \mathbb{R}$ in einen Punkt $x_o \in I$ kann Konvergenzradius $R = 0$ haben (siehe Aufgabe 4.66).
- Die Taylorreihe einer beliebige of differenzierbaren Funktion $f : I \to \mathbb{R}$ in einen Punkt $x_o \in I$ kann Konvergenzradius $R > 0$ haben, jedoch gilt $f(x) - T^{f,x_o}(x) \neq 0$ für alle $x \in I \cap (x_o - R, x_o + R)$ mit $x \neq x_o$.

  Ein Beispiel ergibt sich aus Aufgabe 4.63. Hier verschwinden alle Ableitungen der Funktion
  $$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} \exp\left(-\frac{1}{|x|}\right), & \text{wenn} \quad x \neq 0 \\ 0, & \text{wenn} \quad x = 0 \end{cases}$$
  im Punkt $x_o = 0$. Damit gilt $R = \infty$ und $T^{f,x_o}(x) = 0$ für alle $x \in \mathbb{R}$. Jedoch ist $f(x) \neq 0$ für alle $x \in \mathbb{R} \setminus \{0\}$.

Aufgabe 4.66. Zeigen Sie:

(a) Die Reihe $\sum\limits_{n=1}^{\infty} \frac{1}{2^n} \cos(n^2 x)$ konvergiert für alle $x \in \mathbb{R}$.

(b) Die Funktion
$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \sum_{n=1}^{\infty} \frac{1}{2^n} \cos(n^2 x)$$
ist beliebig oft differenzierbar.

(c) Die Taylor–Reihe von f im Entwicklungspunkte $x_o = 0$ konvergiert nur in $x_o = 0$ und in keinem anderen Punkt $x \in \mathbb{R} \setminus \{x_o\}$.

Definition 4.67. Eine beliebig oft differenzierbare Funktion $f : I \to \mathbb{R}$ auf einem offenen Intervall I heißt *analytisch in* $x_o \in I$, wenn es ein $\delta > 0$ gibt, sodass gilt:

- T^{f,x_o} konvergiert auf $(x_o - \delta, x_o + \delta)$ und
- für alle $x \in (x_o - \delta, x_o + \delta) \cap I$ gilt: $f(x) = T^{f,x_o}(x)$.

Ist f in jedem Punkt x_o analytisch, so nennen wir f eine *analytische* Funktion auf I.

Wir bezeichnen mit $\mathcal{C}^w(I, \mathbb{R})$ die Menge aller analytischen reelwertigen Funktionen auf I.

Bemerkung 4.68.

- $\mathcal{C}^w(I, \mathbb{R})$ ist ein $\mathbb{R}$-Vektorraum und sogar eine reelle Algebra.
- Sei $f : I \to \mathbb{R}$ in $x_o \in I$ analytisch. Dann ist die Taylorreihe von f um x_o die einzige Potenzreihe um x_o, welche f um x_o lokal darstellt (siehe Satz 2.74 und auch Korollar 4.38).

Aus der Definition des Restgliedes des Taylor-Polynoms ergibt sich sofort:

Proposition 4.69. *Sei* $f \in \mathcal{C}^\infty(I,\mathbb{R})$ *und* $x_o \in I$, *sodass der Konvergenzradius* R *der Taylor-Reihe* T^{f,x_o} *strikt positiv ist. Wenn für* $x \in I \cap (x_o - R, x_o + R)$ *gilt* $\lim_{n\to\infty} R_n^{f,x_o}(x) = 0$, *wobei* $R_n^{f,x_o}(x) := f(x) - T_n^{f,x_o}(x)$, *dann gilt für dieses* x *auch* $f(x) = T^{f,x_o}(x)$.

Mithilfe des Majorantenkriteriums, der Euler–Formel für den Konvergenzradius einer Potenzreihe (Satz 2.70), der Form des Restgliedes von Lagrange (Satz 4.59) und der Bemerkung 3.50 folgt:

Korollar 4.70. *Sei* $f \in \mathcal{C}^\infty(I,\mathbb{R})$. *Wenn es eine stetige Funktion* $h : I \to \mathbb{R}$ *gibt, sodass gilt:*

$$\forall x \in I \; \forall n \in \mathbb{N} : \; |f^{(n)}(x)| \leq \big(h(x)\big)^n,$$

dann ist $f \in \mathcal{C}^w(I,\mathbb{R})$.

4.8. Ausblick

Der Begriff der Ableitung als lokale affin lineare Approximation überträgt sich auch auf Abbildungen von zwischen Banach–Räumen. Ein Banach–Raum ist hierbei ein Vektorraum mit einer Norm, welcher (topologisch) vollständig ist, d. h. jede Cauchy–Folge in ist konvergent. Eine gute Darstellung finden Sie in Kapitel VIII des Buches [D].

Zweiter Teil

5 Integration

Sei $f : [t_0, t_1] \to \mathbb{R},\ x \mapsto f(x)$ die Funktion, die den (kontinuierlichen) Temperaturverlauf (z. B. in Grad Celsius[116]) an einem fest gewählten Ort zwischen der Anfangszeit t_0 und der Endzeit t_1 angibt, d. h. $f(x) \in \mathbb{R}$ ist die Temperatur an dem Ort in Grad Celsius zum Zeitpunkt $x \in [t_0, t_1]$. Nun möchte man gerne die Durchschnittstemperatur im Zeitraum $[t_0, t_1]$ an diesem Ort wissen. Bei einer Messung von endlich vielen Temperaturdaten würde man hierzu zunächst die gemessenen Temperaturen addieren und das Ergebnis anschließend durch die Anzahl der Messungen dividieren. Bei kontinuierlichen Messungen wird die Rolle der Summation in einem gewissen Sinn von der Integration übernommen. In gewisser Weise werden „infinitesimal dünne Streifen" der (mit einem Vorzeichen behafteten) Länge $f(x)$ auf dem Intervall $[t_0, t_1]$ der x-Achse aufaddiert. Wie bei der Summe endlich vieler Messwerte, wird der zu $f(x)$ gehörenden „infinitesimal dünne Streifen" abgezogen, wenn der Messwert $f(x)$ negativ ist. Auch das auf Leibniz zurückgehende Integralsymbol $\int$, welches einem lang gezogenen s in einigen alten Schrifttypen, z. B. ſ, entlehnt ist, verweist auf diese Grundidee der Summation.

Geometrisch beschreibt das Integral von f über ein Intervall $[t_0, t_1]$ den Flächeninhalt der gesamten durch den Graphen von f, der x-Achse und den Geraden $x = t_0$ und $x = t_1$ eingeschlossenen beschränkten Fläche. Hierbei wird der Inhalt der Fläche oberhalb der x-Achse positiv und der Inhalt der Fläche unterhalb der x-Achse negativ gewertet.

In der Mathematik gibt es verschiedene Integralbegriffe, welche für teilweise unterschiedliche Klassen von Funktionen definiert sind.

Historisch entstand die Integralrechnung aus dem Wunsch Flächeninhalte bzw. Volumina berechnen zu können. Nachdem die Differentialrechnung von Newton[117] und gleichzeitig und unabhängig auch von Leibniz eingeführt worden war, erkannte man die Integration als Umkehroperation der Ableitung (Hauptsatz der Differential- und Integralrechnung). Ein gut fundierter Integralbegriff für Regelfunktionen geht auf Cauchy zurück. Verallgemeinert wird dieser Integralbegriff duch das hier vorgestellte Riemann[118]–

[116] Anders Celsius (1701–1744), schwedischer Naturwissenschaftler
[117] Sir Isaac Newton (1642–1727), englischer Gelehrter
[118] Georg Friedrich Bernhard Riemann (1826–1866), deutscher Mathematiker

Integral. Durch Hinzufügen von Gewichtsfunktionen erweiterte Stieltjes[119] das Riemann–Integral zum Riemann-Stieltjes–Integral[120], welches in der Stochastik eine große Rolle spielt. Die Idee von Gewichten führte schließlich zum Begriff des Maßes von Lebesgue[121] und zum Lebesgue–Integral, welches der üblicherweise in der modernen Analysis und Stochastik verwendete Integralbegriff ist. Eine Übersicht über verschiedene Integralbegriffe finden Sie z. B. im Buch [HS].

In diesem Kapitel führen wir das Riemann–Integral ein. Hierbei verwenden wir den Zugang über Ober– und Unterintegrale, wie er z. B. in [Fo1, § 18] dargestellt ist.

5.1. Riemann–Integral für Treppenfunktionen

Zum Riemann–Integral gibt es verschiedene Zugänge. Hier halten wir uns eng an [Fo1, § 18].

Definition 5.1. Eine *Zerlegung* eines kompakten Intervalls $[a,b] \subset \mathbb{R}$, $a < b$, ist eine Menge $Z = \{x_0, x_1, \dots, x_n\} \subset \mathbb{R}$ mit

$$a = x_0 < x_1 < \dots < x_{n-1} < x_n = b.$$

Sei $Z = \{x_0, x_1, \dots, x_n\}$ eine Zerlegung von $[a,b]$. Die Zahl

$$\mu_Z := \max\{x_j - x_{j-1} : \ j \in \{1, \dots, n\}\}$$

heißt *Feinheit* der Zerlegung Z.

Seien Z und Z' zwei Zerlegungen von $[a,b]$. Die Zerlegung Z' heißt *Verfeinerung* von Z, falls $Z \subseteq Z'$ gilt[122].

Eine Funktion $\phi : [a,b] \to \mathbb{R}$ heißt *Treppenfunktion* bezüglich einer Zerlegung $Z = \{x_0, x_1, \dots, x_n\} \subset \mathbb{R}$ von $[a,b]$, wenn für alle $j \in \{1, \dots, n\}$ die Einschränkung $\phi|_{(x_{j-1}, x_j)}$ von ϕ auf das offene Intervall $(x_{j-1},\, x_j)$ konstant ist.

Bemerkung 5.2.
- Ist eine Funktion $\phi : [a,b] \to \mathbb{R}$ Treppenfunktion bzgl. einer Zerlegung Z von $[a,b]$, dann ist ϕ auch Treppenfunktion bzgl. jeder Verfeinerung von Z.
- Sind $\phi_1 : [a,b] \to \mathbb{R}$ eine Treppenfunktion bzgl. einer Zerlegung Z_1 und $\phi_2 : [a,b] \to \mathbb{R}$ eine Treppenfunktion bzgl. einer Zerlegung Z_2 von $[a,b]$, dann sind beide ϕ_1 und ϕ_2 Treppenfunktionen bzgl. der Zerlegung $Z_1 \cup Z_2$. Bei zwei gegebenen Treppenfunktionen $\phi_{1,2} : [a,b] \to \mathbb{R}$ können wir also immer annehmen, dass sie beide Treppenfunktionen zu einer gemeinsamen Zerlegung von $[a,b]$ sind.

[119] Thomas Joannes Stieltjes jr. (1856–1894), niederländischer Mathematiker
[120] Es sei kurz darauf hingewiesen, dass es verschiedene nicht-äquivalente Fassungen von Stieltjes–Integralen mit teilweise auch verschiedenen Bezeichnungen gibt (siehe z. B. [W2, § 6], [HS])
[121] Henri-Léon Lebesgue (1875–1941), französischer Mathematiker
[122] Insbesondere ist Z eine Verfeinerung von sich selbst.

Wir bezeichnen mit $\mathcal{T}_{[a,b]}$ die Menge aller Treppenfunktionen $\phi : [a,b] \to \mathbb{R}$ auf dem kompakten Intervall $[a,b]$.

Lemma 5.3. *$\mathcal{T}_{[a,b]}$ ist ein Untervektorraum des reellen Vektorraums aller Funktionen $f : [a,b] \to \mathbb{R}$.*

Dieser einfache Beweis ist eine Übungsaufgabe.

Definition 5.4. Die Abbildung

$$\int_a^b : \mathcal{T}_{[a,b]} \to \mathbb{R}, \quad \phi \mapsto \int\limits_a^b \phi(x)dx,$$

heißt *Integral*. Für eine Treppenfunktion $\phi \in \mathcal{T}_{[a,b]}$ bzgl. einer Zerlegung $\{x_0, x_1, \dots, x_n\}$ von $[a,b]$ ist dieses *Integral* durch

$$\int\limits_a^b \phi(x)dx := \sum_{j=1}^n c_j(x_j - x_{j-1})$$

definiert, wobei c_j der Wert der konstanten Funktion $\phi|_{(x_{j-1},x_j)}$ ist.

Bemerkung 5.5. Damit die Definition des Integrals einer Treppenfunktion ϕ wirklich vernünftig ist, müssen wir zeigen, dass diese Definition nicht von der Zerlegung abhängt, bzgl. welcher ϕ eine Treppenfunktion ist.

Seien also $Z = \{x_0, x_1, \dots, x_n\}$ und $Z' = \{x'_0, x'_1, \dots, x'_m\}$ zwei Zerlegungen bzgl. derer ϕ eine Treppenfunktion ist, d. h. $\phi|_{(x_{j-1},x_j)}$ und $\phi|_{(x'_{k-1},x'_k)}$ sind für alle $j \in \{1, \dots, n\}$ und alle $k \in \{1, \dots, m\}$ konstant. Wir setzen

$$c_j := \phi(x),\ x \in (x_{j-1},\, x_j) \qquad \text{und} \qquad c'_k := \phi(x),\ x \in (x'_{k-1},\, x'_k).$$

Wir müssen nun

$$\sum_{j=1}^n c_j(x_j - x_{j-1}) = \sum_{k=1}^m c'_k(x'_k - x'_{k-1}) \tag{5.1}$$

nachweisen. Zunächst zeigen wir:

Wenn Z' eine Verfeinerung von Z ist, dann ist die Gleichung (5.1) wahr.

Sei nun $j \in \{1, \dots, n\}$, dann gibt es ein $x'_{k_{j-1}} \in Z'$ mit $x_{j-1} = x'_{k_{j-1}}$ und ein $x'_{k_j} \in Z'$ mit $x_j = x'_{k_j}$ und wir erhalten

$$x_{j-1} = x'_{k_{j-1}} < x'_{k_{j-1}+1} < \dots < x'_{k_j} = x_j$$

Desweiteren gilt

$$c'_l = c_j \quad \text{für alle} \quad k_{j-1} + 1 \le l \le k_j.$$

Damit erhalten wir

$$\begin{aligned}\sum_{k=1}^{m} c'_k(x'_k - x'_{k-1}) &= \sum_{j=1}^{n} \sum_{l=k_{j-1}+1}^{k_j} c'_l(x'_l - x'_{l-1}) = \sum_{j=1}^{n} c_j \left(\sum_{l=k_{j-1}+1}^{k_j} (x'_l - x'_{l-1}) \right) \\ &= \sum_{j=1}^{n} c_j(x_j - x_{j-1}).\end{aligned}$$

Seien nun Z und Z' beliebig, dann gilt Gleichung (5.1).

Die Zerlegung $\tilde{Z} = \{\tilde{x}_0, \tilde{x}_1, \ldots, \tilde{x}_s\} := Z \cup Z'$ ist sowohl eine Verfeinerung von Z als auch von Z'. Für alle $l \in \{1, \ldots, s\}$ bezeichnen wir mit $\tilde{c}_l$ den Wert der konstanten Funktion $\phi|_{(\tilde{x}_{l-1},\, \tilde{x}_l)}$. Nach dem zuvor bewiesenen gilt:

$$\sum_{j=1}^{n} c_j(x_j - x_{j-1}) = \sum_{l=1}^{s} c'_k(\tilde{x}_l - \tilde{x}_{l-1}) = \sum_{k=1}^{m} c'_k(x'_k - x'_{k-1}).$$

Wir erkennen folgende Eigenschaften des Integrals von Treppenfunktion:

Satz 5.6 (Eigenschaften des Integrals von Treppenfunktionen)**.**

(a) Das Integral $\int\limits_a^b : \mathcal{T}_{[a,b]} \to \mathbb{R}$ *ist $\mathbb{R}$-linear, d. h. für zwei Treppenfunktionen $\phi_1, \phi_2 \in \mathcal{T}_{[a,b]}$ und alle $\lambda \in \mathbb{R}$ gilt:*

$$\int_a^b (\lambda \cdot \phi_1(x) + \phi_2(x))dx = \lambda \cdot \int_a^b \phi_1(x)dx + \int_a^b \phi_2(x)dx.$$

(b) Das Integral ist monoton, *d. h. für zwei Treppenfunktionen $\phi_1, \phi_2 \in \mathcal{T}_{[a,b]}$ gilt: Aus $\phi_1 \leq \phi_2$ (d. h. wenn $\phi_1(x) \leq \phi_2(x)$ für alle $x \in [a, b]$ gilt), folgt*

$$\int\limits_a^b \phi_1(x)dx \leq \int\limits_a^b \phi_2(x)dx.$$

5.2. Riemann-Integral für beschränkte Funktionen

In diesem Abschnitt orientieren wir uns weiter an [Fo1, § 18].

Definition 5.7. Für eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ definieren wir das *Oberintegral*

$$\overline{\int_a^b} f(x)dx := \inf \left\{ \int_a^b \phi(x)dx : \ \phi \in \mathcal{T}_{[a,b]},\ \phi \geq f \right\}$$

und das *Unterintegral*

$$\underline{\int_a^b} f(x)dx := \sup \left\{ \int_a^b \phi(x)dx : \ \phi \in \mathcal{T}_{[a,b]},\ \phi \leq f \right\}.$$

Bemerkung 5.8. (i) Da f beschränkt ist, sind $\overline{\int_a^b} f(x)dx$ und $\underline{\int_a^b} f(x)dx$ endlich.

(ii) Für eine Treppenfunktionen $\phi \in \mathcal{T}_{[a,b]}$ gilt

$$\overline{\int_a^b} \phi(x)dx = \int_a^b \phi(x)dx = \underline{\int_a^b} \phi(x)dx.$$

(iii) Für jede beschränkte Funktion $f : [a,b] \to \mathbb{R}$ gilt

$$\overline{\int_a^b} f(x)dx \geq \underline{\int_a^b} f(x)dx.$$

Lemma 5.9 (Beziehung zwischen Ober– und Unterintegral)**.** *Für eine beschränkte Funktion $f : [a,b] \to \mathbb{R}$ gilt*

$$\underline{\int_a^b} f(x)dx = -\overline{\int_a^b} (-f(x))dx.$$

Beweis. Dieser Beweis ist eine Übungsaufgabe. □

Proposition 5.10 (Eigenschaften des Oberintegrals)**.** *Für beschränkte Funktionen $f_1, f_2, f : [a,b] \to \mathbb{R}$ und $\lambda \geq 0$ gilt:*

(i) $\overline{\int_a^b} (f_1 + f_2)(x)dx \leq \overline{\int_a^b} f_1(x)dx + \overline{\int_a^b} f_2(x)dx.$

(ii) $\overline{\int_a^b} (\lambda f)(x)dx = \lambda \overline{\int_a^b} f(x)dx.$

Beweis. Sei $\varepsilon > 0$ beliebig. Es gibt nun Treppenfunktionen $\phi_1, \phi_2 \in \mathcal{T}_{[a,b]}$ mit $\phi_1 \geq f_1$ und $\phi_2 \geq f_2$, sodass gilt

$$\int_a^b \phi_1(x)dx \leq \frac{\varepsilon}{2} + \overline{\int_a^b} f_1(x)dx \quad \text{und} \quad \int_a^b \phi_2(x)dx \leq \frac{\varepsilon}{2} + \overline{\int_a^b} f_2(x)dx.$$

Aus $\phi_1 + \phi_2 \geq f_1 + f_2$ folgt dann sofort

$$\overline{\int_a^b} (f_1 + f_2)(x)dx \leq \int_a^b (\phi_1 + \phi_2)(x)dx \leq \varepsilon + \overline{\int_a^b} f_1(x)dx + \overline{\int_a^b} f_2(x)dx.$$

Damit gilt für alle $\varepsilon > 0$

$$\overline{\int_a^b} (f_1 + f_2)(x)dx \leq \varepsilon + \overline{\int_a^b} f_1(x)dx + \overline{\int_a^b} f_2(x)dx.$$

Daraus folgt die Aussage (i).

Für $\lambda = 0$ ist die Aussage (ii) sicher richtig. Um (ii) für $\lambda > 0$ nachzuweisen genügt es zu zeigen, dass für alle $\varepsilon > 0$ gilt

$$-\varepsilon + \lambda \overline{\int_a^b} f(x)dx \leq \overline{\int_a^b} (\lambda f)(x)dx \leq \varepsilon + \lambda \overline{\int_a^b} f(x)dx,$$

wobei $\lambda > 0$. Wir zeigen zunächst

$$\lambda \overline{\int_a^b} f(x)dx \leq \varepsilon + \overline{\int_a^b} (\lambda f)(x)dx.$$

Es gibt eine Treppenfunktion $\psi \in \mathcal{T}_{[a,b]}$ mit $\psi \geq (\lambda f)$ sowie auch $\int\limits_a^b \psi(x)dx \leq \varepsilon + \overline{\int\limits_a^b}(\lambda f)(x)dx$. Wegen $\frac{1}{\lambda}\psi \geq f$ gilt

$$\overline{\int_a^b} f(x)dx \leq \int_a^b \left(\frac{1}{\lambda}\psi\right)(x)dx = \frac{1}{\lambda}\int_a^b \psi(x)dx \leq \frac{1}{\lambda}\left(\varepsilon + \overline{\int_a^b}(\lambda f)(x)dx\right).$$

Somit gilt

$$\lambda \overline{\int_a^b} f(x)dx \leq \varepsilon + \overline{\int_a^b} (\lambda f)(x)dx.$$

Als letztes zeigen wir die Ungleichung

$$\overline{\int_a^b} (\lambda f)(x)dx \leq \varepsilon + \lambda \overline{\int_a^b} f(x)dx.$$

Es gibt nun eine Treppenfunktion $\phi \geq f$ mit

$$\int_a^b \phi(x)dx \leq \frac{\varepsilon}{\lambda} + \overline{\int_a^b} f(x)dx.$$

Wegen $\lambda\phi \geq \lambda f$ gilt schließlich:

$$\begin{aligned}\overline{\int_a^b} (\lambda f)(x)dx &\leq \int_a^b (\lambda\phi)(x)dx = \lambda \int_a^b \phi(x)dx \leq \lambda\left(\frac{\varepsilon}{\lambda} + \overline{\int_a^b} f(x)dx\right) \\ &= \varepsilon + \lambda \overline{\int_a^b} f(x)dx.\end{aligned}$$

□

Aus Lemma 5.9 und Proposition 5.10 folgt durch elementare Überlegungen:

Proposition 5.11 (Eigenschaften des Unterintegrals). *Für beschränkte Funktionen* $f_1, f_2, f : [a, b] \to \mathbb{R}$ *gilt:*

(i) $\underline{\int_a^b}(f_1 + f_2)(x)dx \geq \underline{\int_a^b} f_1(x)dx + \underline{\int_a^b} f_2(x)dx.$

(ii) Für $\lambda \geq 0$ *gilt* $\underline{\int_a^b}(\lambda f)(x)dx = \lambda \underline{\int_a^b} f(x)dx$.

(iii) Für $\lambda < 0$ *gilt* $\overline{\int_a^b}(\lambda f)(x)dx = \lambda \underline{\int_a^b} f(x)dx$ *und* $\underline{\int_a^b}(\lambda f)(x)dx = \lambda \overline{\int_a^b} f(x)dx$.

Definition 5.12. Eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ auf einem kompakten Intervall $[a, b] \subset \mathbb{R}$ heißt *Riemann–integrierbar* oder kurz *integrierbar*, wenn gilt:

$$\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx.$$

In diesem Fall nennen wir die reelle Zahl

$$\int_a^b f(x)dx := \overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx$$

das *Riemann–Integral* oder oft auch kurz das *Integral* von f.[123]

Bemerkung 5.13. Für eine Treppenfunktion $\phi \in \mathcal{T}_{[a,b]}$ stimmt das Riemann-Integral mit dem zuvor definierten Integral überein.

Aus der Definition des Riemann-Integrals und aus den Propositionen 5.10 und 5.11 ergibt sich:

Satz 5.14 (Eigenschaften des Riemann-Integrals). *(i) Die Menge* $\mathcal{R}_{[a,b]}$ *aller Riemann–integrierbaren Funktionen* $f : [a, b] \to \mathbb{R}$ *ist bzgl. der üblichen Addition und skalaren Multiplikation ein reeller Vektorraum und das Riemann–Integral ist eine lineare Abbildung von* $\mathcal{R}_{[a,b]}$ *nach* $\mathbb{R}$*, d. h. für alle* $f, g \in \mathcal{R}_{[a,b]}$ *und für alle* $\lambda \in \mathbb{R}$ *gilt:*

- $\int_a^b (f + g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx,$
- $\int_a^b (\lambda f)(x)dx = \lambda \int_a^b f(x)dx.$

(ii) Das Riemann–Integral ist monoton, d. h.: Sind $f, g \in \mathcal{R}_{[a,b]}$ *mit* $f \leq g$*, dann folgt*

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx.$$

Beispiel 5.15 (Das klassische Beispiel einer nicht Riemann–integrierbaren Funktion)**.** Die Funktion

$$f : [0, 1] \to \mathbb{R}, \quad x \mapsto \begin{cases} 1, & \text{wenn} \quad x \in \mathbb{Q} \cap [0, 1] \\ 0, & \text{sonst} \end{cases}$$

ist nicht Riemann–integrierbar (Übung).

[123] Die Notation $f(x)dx$ wird in der Theorie der Differentialformen klar. Hier nehmen wir diese Notation einfach zur Kenntnis und beachten, dass $\int_a^b f(x)dx = \int_a^b f(t)dt = \int_a^b f(\tau)d\tau$ gilt.

Satz 5.16 (Charakterisierung Riemann–integrierbarer Funktionen)**.** *Eine Funktion $f : [a,b] \to \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn gilt:*

$$\forall \varepsilon > 0 \; \exists \phi_1, \phi_2 \in \mathcal{T}_{[a,b]} : \quad \phi_1 \leq f \leq \phi_2 \text{ und } \int_a^b \phi_2(x)dx - \int_a^b \phi_1(x)dx \leq \varepsilon.$$

Der Beweis ist eine Übungsaufgabe, welche mit Standardmethoden gelöst werden kann.

Satz 5.17 (Riemann-Integrierbarkeit monotoner Funktionen)**.** *Ist $f : [a,b] \to \mathbb{R}$ monoton (und damit auch beschränkt), dann ist f Riemann-integrierbar.*

Beweis. Wir nehmen an, dass $f : [a,b] \to \mathbb{R}$ monoton fallend ist. Sei $\varepsilon > 0$ vorgegeben. Wir wählen $n \in \mathbb{N}$ so groß, dass

$$\frac{(f(a) - f(b)) \cdot (b-a)}{n} \leq \varepsilon$$

gilt und betrachten die äquidistante Zerlegung $Z = \{x_0, x_1 \ldots, x_n\}$ von $[a,b]$ definiert durch

$$x_j = a + \frac{j}{n}(b-a) \quad \text{für } j = 0, 1, 2, \ldots, n.$$

Wir betrachten nun die beiden Treppenfunktionen $\phi_{1,2}$ bzgl. Z, welche durch $\phi_1(a) = \phi_2(a) = f(a)$ und

$$\begin{aligned} \phi_1(x) &:= f(x_j), && \text{für } x \in (x_{j-1}, x_j] \\ \phi_2(x) &:= f(x_{j-1}), && \text{für } x \in (x_{j-1}, x_j] \end{aligned}$$

definiert sind. Da f monoton fallend ist, gilt offenbar

$$\phi_1 \leq f \leq \phi_2.$$

Desweiteren erhalten wir

$$\begin{aligned} \int_a^b \phi_2(x)dx - \int_a^b \phi_1(x)dx &= \sum_{j=1}^n f(x_{j-1}) \underbrace{(x_j - x_{j-1})}_{=\frac{b-a}{n}} - \sum_{j=1}^n f(x_j) \underbrace{(x_j - x_{j-1})}_{=\frac{b-a}{n}} \\ &= \frac{b-a}{n} \left(\sum_{j=1}^n f(x_{j-1}) - \sum_{j=1}^n f(x_j) \right) \\ &= \frac{b-a}{n}(f(x_0) - f(x_n)) = \frac{b-a}{n}(f(a) - f(b)) \\ &\leq \varepsilon. \end{aligned}$$

Nach Satz 5.16 ist f somit Riemann–integrierbar.

Der Beweis für den Fall, dass f monoton wachsend ist, ist dem Leser überlassen. □

Satz 5.18 (Riemann-Integrierbarkeit stetiger Funktionen)**.** *Ist $f : [a,b] \to \mathbb{R}$ stetig, dann ist f auch Riemann–integrierbar.*

Um Satz 5.18 zu beweisen, benötigen wir zunächst folgende Aussage (s. [Fo1, § 11, Satz 5])

Proposition 5.19 (Approximation stetiger Funktionen durch Treppenfunktionen). *Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion und $\varepsilon > 0$. Dann gibt es Treppenfunktionen $\phi_{1,2} : [a, b] \to \mathbb{R}$, welche die folgenden beiden Eigenschaften haben:*

(i) $\phi_1 \leq f \leq \phi_2$ und

(ii) für alle $x \in [a, b]$ gilt $(\phi_2(x) - \phi_1(x)) \leq \varepsilon$.

Beweis. Da $[a, b]$ ein kompaktes Intervall ist, ist $f : [a, b] \to \mathbb{R}$ als stetige Funktion nach dem Satz von Heine (Theorem 3.33) auch gleichmäßig stetig. Damit gibt es ein $\delta > 0$, sodass für alle $x, y \in [a, b]$ mit $|x - y| < \delta$ folgt

$$|f(x) - f(y)| < \frac{\varepsilon}{2}.$$

Wir wählen nun $n \in \mathbb{N}$ hinreichend groß, sodass

$$\frac{b-a}{n} < \delta$$

gilt und betrachten die äquidistante Zerlegung $Z = \{x_0, x_1 \ldots, x_n\}$ von $[a, b]$ definiert durch

$$x_j = a + \frac{j}{n}(b - a) \quad \text{für } j = 0, 1, 2, \ldots, n.$$

Wir definieren nun $\phi_{1,2}$ durch $\phi_1(a) = \phi_2(a) := f(a)$ und

$$\begin{array}{rcll} \phi_1(x) & := & f(x_j) - \frac{\varepsilon}{2}, & \text{für} \quad x \in (x_{j-1}, x_j] \\ \phi_2(x) & := & f(x_j) + \frac{\varepsilon}{2}, & \text{für} \quad x \in (x_{j-1}, x_j]. \end{array}$$

Nach Definition gilt für alle $x \in [a, b]$:

$$\phi_2(x) - \phi_1(x) \leq \varepsilon.$$

Zudem gilt

$$\phi_1(a) = f(a) = \phi_2(a).$$

Für $x \in (x_{j-1}, x_j],\ j = \{1, \ldots, n\}$, folgt aus $|x - x_j| < \delta$ sofort

$$-\frac{\varepsilon}{2} < f(x) - f(x_j) < \frac{\varepsilon}{2}$$

und somit

$$\phi_1(x) = f(x_j) - \frac{\varepsilon}{2} < f(x) < f(x_j) + \frac{\varepsilon}{2} = \phi_2(x).$$

Damit gilt aber auch

$$\phi_1(x) \leq f(x) \leq \phi_2(x)$$

für alle $x \in [a, b]$, also $\phi_1 \leq f \leq \phi_2$. □

Nun können wir zum Beweis von Satz 5.18 schreiten.

Beweis. (von Satz 5.18)
Sei $f : [a,b] \to \mathbb{R}$ stetig und $\varepsilon > 0$ beliebig gewählt. Nach Proposition 5.19 gibt es $\phi_{1,2} \in \mathcal{T}_{[a,b]}$ mit folgenden Eigenschaften

- $\phi_1 \le f \le \phi_2$ und
- $(\phi_2(x) - \phi_1(x)) < \frac{\varepsilon}{b-a}$ für alle $x \in [a,b]$.

Da $\phi_2 - \phi_1$ wieder eine Treppenfunktion ist, folgt aus der Linearität und der Monotonie des Integrals

$$\int_a^b \phi_2(x)dx - \int_a^b \phi_1(x)dx = \int_a^b (\phi_2(x) - \phi_1(x))dx \le \int_a^b \frac{\varepsilon}{b-a}dx = \varepsilon.$$

Hierbei haben wir $\frac{\varepsilon}{b-a}$ als die konstante Funktion $[a,b] \to \mathbb{R},\ x \mapsto \frac{\varepsilon}{b-a}$ aufgefasst. Nach Satz 5.16 ist f somit Riemann–integrierbar. □

Aufgabe 5.20. Zeigen Sie: Sei $f : [a,b] \to [0,\infty)$ eine nicht-negative stetige Funktion mit $\int\limits_a^b f(x)dx = 0$, dann gilt $f = 0$.

Satz 5.21. *Sei $f, g \in \mathcal{R}_{[a,b]}$, dann sind auch folgende Funktionen Riemann–integrierbar:*

(a) $f_+ : [a,b] \to \mathbb{R},\ x \mapsto \max\{f(x), 0\}$ *und* $f_- : [a,b] \to \mathbb{R},\ x \mapsto \max\{-f(x), 0\}$,

(b) $|f|^p : [a,b] \to \mathbb{R},\ x \mapsto |f(x)|^p$, *für* $p \ge 1$,

(c) $fg : [a,b] \to \mathbb{R}$.

Beweis. Da mit f auch $-f$ integrierbar ist und $f_- = (-f)_+$ gilt, genügt es nachzuweisen, dass f_+ integrierbar ist. Sei $\varepsilon > 0$. Nach Satz 5.16 gibt es $\phi_{1,2} \in \mathcal{T}_{[a,b]}$ mit $\phi_1 \le f \le \phi_2$ und $\int_a^b (\phi_2 - \phi_1)(x)dx \le \varepsilon$. Dann gilt $\phi_{1+} \le f_+ \le \phi_{2+}$ und ϕ_{1+}, ϕ_{2+} sind Treppenfunktionen auf $[a,b]$. Aus $\phi_{2+} - \phi_{1+} \le \phi_2 - \phi_1$ folgt nun

$$\int_a^b (\phi_{2+} - \phi_{1+})(x)dx \le \int_a^b (\phi_2 - \phi_1)(x)dx \le \varepsilon.$$

Nach Satz 5.16 ist also auch f_+ Riemann–integrierbar. Damit gilt die Aussage (a).

Wir beachten, dass $|f| = f_+ + f_-$ gilt. Somit ist mit f auch $|f|$ Riemann–integrierbar. Um zu zeigen, dass $|f|^p,\ p \ge 1$ Riemann–integrierbar ist, genügt es, nach geeigneter Multiplikation mit einer Konstanten, anzunehmen, dass $0 \le f \le 1$ gilt. Wieder wählen wir ein $\varepsilon > 0$. Da f Riemann–integrierbar ist, gibt es $\phi_{1,2} \in \mathcal{T}_{[a,b]}$, welche

$$0 \le \phi_1 \le f \le \phi_2 \le 1$$

und

$$\int_a^b (\phi_2 - \phi_1)(x)dx \le \frac{\varepsilon}{p}.$$

Für die Treppenfunktionen ϕ_1^p und ϕ_2^p gilt also $\phi_1^p \leq f^p \leq \phi_2^p$. Für $0 \leq y_1 \leq y_2 \leq 1$ folgt aus dem Mittelwertsatz der Differentialrechnung für die Funktion $x \mapsto x^p$ mit $p \geq 1$:

$$\frac{y_2^p - y_1^p}{y_2 - y_1} \leq p \cdot \xi^{p-1} \leq p, \qquad \xi \in [y_1, y_2] \subset [0,1].$$

Wählen wir nun $y_{1,2} = \phi_{1,2}(x)$, so folgt:

$$\phi_2^p - \phi_1^p \leq p(\phi_2 - \phi_1).$$

Aus der Monotonie des Integrals folgt

$$\int_a^b (\phi_2^p - \phi_1^p)(x)dx \leq p \int_a^b (\phi_2 - \phi_1)(x)dx \leq \varepsilon.$$

Somit ist auch f^p Riemann–integrierbar. Das beweist (b).

Die Aussage (c) ergibt sich sofort aus

$$fg = \frac{1}{4}((f+g)^2 - (f-g)^2).$$

□

Vorsicht. Das Riemann–Integral ist *nicht* multiplikativ, d. h. im Allgemeinen gilt für $f, g \in \mathcal{R}_{[a,b]}$:

$$\int_a^b (fg)(x)dx \neq \left(\int_a^b f(x)dx\right)\left(\int_a^b g(x)dx\right).$$

Theorem 5.22 (Erster Mittelwertsatz für das Riemann–Integral)**.** *Seien $f, g : [a,b] \to \mathbb{R}$ stetige Funktionen wobei g nicht negativ ist, d. h. $g(x) \geq 0$ für alle $x \in [a,b]$. Dann gilt:*

$$\exists x_o \in [a,b] : \quad \int_a^b f(x)g(x)dx = f(x_o)\int_a^b g(x)dx.$$

Wählt man speziell $g(x) = 1$ für alle $x \in [a,b]$, so erhält man:

$$\exists x_o \in [a,b] : \quad \int_a^b f(x)dx = f(x_o)(b-a).$$

Beweis. Die stetige Funktion f hat auf dem kompakten Intervall $[a,b]$ ein Minimum m und ein Maximum M (Weierstraßscher Satz vom Extremum (Theorem Satz 3.26). Da $g \geq 0$ erhalten wir damit $mg \leq fg \leq Mg$. Aufgrund der Monotonie und der Linearität des Riemann–Integrals erhalten wir

$$m\int_a^b g(x)dx \leq \int_a^b f(x)g(x)dx \leq M\int_a^b g(x)dx.$$

Man beachte, dass $\int_a^b g(x)dx \geq 0$ gilt. Es existiert nun ein $\lambda \in [m,M]$, sodass

$$\int_a^b f(x)g(x)dx = \lambda\int_a^b g(x)dx.$$

Nach der Zwischenwerteigenschaft stetiger Funktionen gibt es nun ein $x_o \in [a,b]$ mit $f(x_o) = \lambda \in [m,M]$. □

5.3. Riemannsche und Darbouxsche Summen

Andere Zugänge zum Riemann–Integral sind die Riemannschen Summen und auch Darbouxschen Summen. In diesem Abschnitt stellen wir diese kurz dar. Wir lehnen uns an [Fo1, § 18], [GBGW2, 6.2], [H1, Kap. 8] und [W1, § 9] an.

Definition 5.23. Sei $f : [a, b] \to \mathbb{R}$ eine beschränkte Funktion und $Z = \{x_0, x_1, \dots, x_n\}$ eine Zerlegung von $[a, b]$. Dann heißen die Summen

- $$R(f, Z, \xi_1, \xi_2, \dots, \xi_n) := \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1})$$

  Riemannsche Summe von f bzgl. der Zerlegung Z von $[a, b]$ und den *Stützstellen* $\xi_j \in [x_{j-1}, x_j]$ für $j \in \{1, \dots, n\}$.

- $$O(f, Z) := \sum_{j=1}^{n} (\sup\{f(x) : \ x \in [x_{j-1}, x_j]\})(x_j - x_{j-1})$$

  (Darbouxsche[124]) Obersumme von f zgl. Z.

- $$U(f, Z) := \sum_{j=1}^{n} (\inf\{f(x) : \ x \in [x_{j-1}, x_j]\})(x_j - x_{j-1})$$

  (Darbouxsche) Untersumme von f zgl. Z.

Beobachtung 5.24. *Es gilt offenbar*

$$U(f, Z) \leq R(f, Z, \xi_1, \xi_2, \dots, \xi_n) \leq O(f, Z)$$

für eine beliebige Wahl von Stützstellen $\xi_j \in [x_{j-1}, x_j]$ für $j \in \{1, \dots, n\}$.

Folgende Beobachtung wird durch eine Skizze sofort klar. Die formale Ausformulierung des Beweises ist dann eine Übungsaufgabe:

Beobachtung 5.25. *Ist Z eine Verfeinerung von $\tilde{Z}$, dann gilt*

$$O(f, Z) \leq O(f, \tilde{Z}) \quad \text{und} \quad U(f, Z) \geq U(f, \tilde{Z}).$$

Lemma 5.26. *Seien Z_1 und Z_2 zwei beliebige Zerlegungen von $[a, b]$, dann gilt*

$$U(f, Z_1) \leq O(f, Z_2).$$

Beweis. $Z := Z_1 \cup Z_2$ ist sowohl eine Verfeinerung von Z_1 wie auch von Z_2. Damit erhalten wir $U(f, Z_1) \leq U(f, Z) \leq O(f, Z) \leq O(f, Z_2)$. ☐

[124] Jean Gaston Darboux (1842–1917), französischer Mathematiker

Theorem 5.27 (Charakterisierung des Riemann–Integrals durch Riemannsche und Darbouxsche Summen)**.**
Für eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ sind folgende drei Aussagen äquivalent:

(i) f ist Riemann-integrierbar.

(ii) Für alle $\varepsilon > 0$ gibt es eine Zerlegung Z von $[a, b]$, sodass

$$O(f, Z) - U(f, Z) < \varepsilon.$$

(iii) Es gibt eine Zahl $\iota \in \mathbb{R}$ mit folgender Eigenschaft: Für alle $\varepsilon > 0$ gibt es ein $\delta > 0$, sodass für jede Zerlegung $Z = \{x_0, x_1, \ldots, x_n\}$ von $[a, b]$ der Feinheit $\mu_Z \leq \delta$ und jede Wahl von Stützstellen $\xi_1, \ldots, \xi_n$ gilt

$$|\iota - R(f, Z, \xi_1, \ldots, \xi_n)| \leq \varepsilon.$$

(iv) Es gibt eine Zahl $\iota \in \mathbb{R}$ mit folgender Eigenschaft: Für alle $\varepsilon > 0$ gibt es eine Zerlegung $\tilde{Z}_\varepsilon = \{\tilde{x}_0, \ldots \tilde{x_m}\}$ von $[a, b]$, sodass für jede Verfeinerung $Z = \{x_0, \ldots, x_n\}$ von $\tilde{Z}_\varepsilon$ und jede Wahl von Stützstellen $\xi_1, \ldots, \xi_n,\ \xi_k \in [x_{k-1}, x_k],\ k = 1, \ldots, n$, gilt:

$$|\iota - R(f, Z, \xi_1, \ldots, \xi_n)| \leq \varepsilon.$$

Bemerkung 5.28.

- Die Zahl ι in den Aussagen (iii) und (iv) von Theorem 5.27 ist eindeutig bestimmt. Es gilt $\iota = \int_a^b f(x)dx$.
- Der Beweis der Äquivalenz der Aussagen (i), (ii) und (iii) von Theorem 5.27 ist eine Übungsaufgabe auf dem Übungsblatt 2.
- Die Aussagen (iii) und (iv) von Theorem 5.27 klingen sehr ähnlich. Die Äquivalenz der Aussagen (iii) und (iv) für die Riemannsche Summe und das Riemann–Integral überträgt sich aber bereits nicht mehr auf analoge Konstruktionen für das Stieltjes-Integral. Nähere Erläuterungen hierzu finden Sie in [W2, §6].

Beweis. Wir zeigen hier nur die Implikationen $(i) \implies (iii) \implies (iv) \implies (ii)$ des Satzes 5.27. Um unsere Schlusskette zu schließen fehlt noch die Implikation $(ii) \implies (i)$, welche eine Übungsaufgabe ist.

- *Wir zeigen zunächst $(i) \implies (iii)$, also den vielleicht schwierigsten Teil Ihrer Übungsaufgabe:*

  Sei ε gegeben und $\iota = \int_a^b f(x)dx$. Gemäß Satz 5.16 gibt es nun $\phi_{1,2} \in \mathcal{T}_{[a,b]}$ zu einer (ohne Einschränkung) gemeinsamen Zerlegung $\tilde{Z} = \{\tilde{x}_0, \tilde{x}_1, \ldots, \tilde{x}_m\}$ von $[a, b]$, sodass $\phi_1 \leq f \leq \phi_2$ und

$$\int_a^b \phi_2(x)dx - \int_a^b \phi_1(x)dx \leq \frac{\varepsilon}{2}$$

gilt. Somit erhalten wir

$$\text{(5.2)} \quad \int_a^b f(x)dx \le \int_a^b \phi_1(x)dx + \frac{\varepsilon}{2} \qquad \text{und} \qquad \int_a^b \phi_2(x)dx \le \int_a^b f(x)dx + \frac{\varepsilon}{2}.$$

Wir setzen nun

$$\delta := \frac{\varepsilon}{8Mm}$$

mit $M := \sup\{|f(x)| : \ x \in [a,b]\}$. M ist eine positive reelle Zahl, da f beschränkt ist.

Sei nun $Z = \{x_0, x_1, \ldots, x_n\}$ eine Zerlegung von $[a,b]$ der Feinheit $\mu_Z \le \delta$ und $\{\xi_1, \ldots \xi_n\}$ eine Menge von Stützstellen mit $\xi_j = [x_{j-1}, x_j], \ j = 1, \ldots, n$.

Wir betrachten nun die durch

$$\psi(x_j) = 0, \ j = 0, 1, \ldots, n, \qquad \text{und} \qquad \psi(x) = f(\xi_j) \text{ für } x \in (x_{j-1}, x_j), \ j = 1, \ldots, n,$$

definierte Treppenfunktion $\psi \in \mathcal{T}_{[a,b]}$ zur Zerlegung Z. Wir betrachten die Mengen

$$\begin{aligned} J &:= \big\{j \in \{1, \ldots n\} \mid \exists k \in \{1, \ldots, m\} : [x_{j-1}, x_j] \subset (\tilde{x}_{k-1}, \tilde{x}_k)\big\} \\ &= \big\{j \in \{1, \ldots n\} \mid [x_{j-1}, x_j] \cap \tilde{Z} = \emptyset\big\}, \\ U &:= \bigcup_{j \in J} (x_{j-1}, x_j) \ \subset \ [a,b]. \end{aligned}$$

Damit gilt

- $R(f, Z, \xi_1, \ldots, \xi_n) = \sum\limits_{j=1}^{n} f(\xi_j)(x_j - x_{j-1}) = \int\limits_a^b \psi(x)dx,$
- $\phi_1 - 2M \le \psi \le \phi_2 + 2M,$
- $\phi_1|_U \le \psi|_U \le \phi_2|_U.$

Wir definieren nun eine Treppenfunktion $\eta \in \mathcal{T}_{[a,b]}$ bzgl. Z durch

$$\eta(x) := \begin{cases} 0, & \text{wenn} \ \ x \in U \\ 2M, & \text{wenn} \ \ x \in [a,b] \setminus U \end{cases}.$$

Aus den Eigenschaften von ψ folgt

$$\phi_1 - \eta \le \psi \le \phi_2 + \eta.$$

Betrachten wir nun die Menge

$$I := \{1, \ldots, n\} \setminus J \ = \ \big\{j \in \{1, \ldots n\} : \ [x_{j-1}, x_j] \cap \tilde{Z} \neq \emptyset\big\}$$

Jedes der $m-1$ Elemente $\tilde{x}_k$ mit $k = 1, \ldots m-1$ liegt in genau einem Intervall der Form $[x_{j-1}, x_j]$, wenn $\tilde{x}_k \notin Z$, und in genau zwei Intervallen der Form $[x_{j-1}, x_j]$, wenn $\tilde{x}_k \in Z$. Die beiden Endpunkte $\tilde{x}_0 = a$ bzw. $\tilde{x}_m = b$ liegen natürlich genau in den Intervallen $[x_0, x_1]$ bzw. $[x_{n-1}, x_n]$. Also gilt $1, n \in I$. Somit hat I höchstens $2m$ Elemente. Wir bemerken zudem, dass

$$[a,b] \setminus U = \bigcup_{i \in I} [x_{i-1}, x_i].$$

Wir erhalten, da η in höchstens $2m$ Intervallen der Form $[x_{k-1}, x_k]$ nicht verschwindet, folgt

$$\int_a^b \eta(x)dx \leq 2M(2m)\delta = 4Mm\delta \leq \frac{\varepsilon}{2}.$$

Somit erhalten wir

$$-\frac{\varepsilon}{2} + \int_a^b \phi_1(x)dx \leq \int_a^b \psi(x)dx \leq \frac{\varepsilon}{2} + \int_a^b \phi_2(x)dx$$

Mit (5.2) folgt

$$\left| \int_a^b f(x)dx - R(f, Z, \xi_1, \ldots, \xi_n) \right| = \left| \int_a^b f(x)dx - \int_a^b \psi(x)dx \right| \leq \varepsilon.$$

- *Wir zeigen nun (iii) $\Longrightarrow$ (iv):*

  Sei $\varepsilon > 0$ gegeben, sowie ι und $\delta > 0$ aus Eigenschaft (iii). Sei $\tilde{Z}_\varepsilon$ die äquidistante Zerlegung von $[a,b]$ mit $\mu_{\tilde{Z}_\varepsilon} = \delta$. Dann gilt für jede Verfeinerung $Z = \{x_0, x_1, \ldots, x_n\}$ von $\tilde{Z}_\varepsilon$ auch $\mu_Z \leq \mu_{\tilde{Z}_\varepsilon} = \delta$. Damit folgt aus (iii) sofort

  $$|\iota - R(f, Z, \xi_1, \ldots, \xi_n)| \leq \varepsilon.$$

  für jede Wahl von Stützstellen $\xi_1, \ldots, \xi_n,\ \xi_k \in [x_{k-1}, x_k],\ k = 1, \ldots, n$.

- *Zum Schluß beweisen wir noch (iv) $\Longrightarrow$ (ii).*

  Sei $\varepsilon > 0$. Nach (iv) gibt es ein $\iota \in \mathbb{R}$ und eine Zerlegung $\tilde{Z}_{\varepsilon/4}$, sodass für jede Verfeinerung $Z = \{x_0, \ldots, x_n\}$ von $\tilde{Z}_{\varepsilon/4}$ und *jede* Wahl der Stützstellen $\xi_1, \ldots, \xi_n$ gilt:

  $$\iota - \frac{\varepsilon}{4} \leq R(f, Z, \xi_1, \ldots, \xi_n.) \leq \iota + \frac{\varepsilon}{4}.$$

  Wir wählen nun ein solche Verfeinerung $Z = \{x_0, \ldots, x_n\}$ von $\tilde{Z}_{\varepsilon/4}$. Hierzu gibt es Stützstellen $\xi_1^u, \ldots, \xi_n^u$ und $\xi_1^o, \ldots, \xi_n^o$, sodass

  $$R(f, Z, \xi_1^u, \ldots, \xi_n^u) - \frac{\varepsilon}{4} \leq U(f, Z) \qquad \text{und} \qquad O(f, Z) \leq R(f, Z, \xi_1^o, \ldots, \xi_n^o) + \frac{\varepsilon}{4}.$$

  Dann gilt

  $$\begin{aligned} \iota - \frac{\varepsilon}{2} &= \iota - \frac{\varepsilon}{4} - \frac{\varepsilon}{4} \leq R(f, Z, \xi_1^u, \ldots, \xi_n^u.) - \frac{\varepsilon}{4} \leq U(f, Z) \\ &\leq O(f, Z) \leq R(f, Z, \xi_1^o, \ldots, \xi_n^o) + \frac{\varepsilon}{4} \leq \iota + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \iota + \frac{\varepsilon}{2}. \end{aligned}$$

  Also $\iota - \frac{\varepsilon}{2} \leq U(f, Z) \leq O(f, Z) \leq \iota + \frac{\varepsilon}{2}$ und somit $O(f, Z) - U(f, Z) \leq \varepsilon$.

Der noch fehlende Beweis von $(ii) \implies (i)$ ist eine Übungsaufgabe. $\square$

Korollar 5.29. *Eine beschränkte Funktion $f : [a,b] \to \mathbb{R}$ ist dann und nur dann Riemann–integrierbar, wenn*

$$\sup\{U(f,Z) : \ Z \text{ Zerlegung von } [a,b]\} = \inf\{O(f,Z) : \ Z \text{ Zerlegung von } [a,b]\}.$$

In diesem Fall gilt

$$\begin{aligned}\int_a^b f(x)dx &= \sup\{U(f,Z) : \ Z \text{ Zerlegung von } [a,b]\} \\ &= \inf\{O(f,Z) : \ Z \text{ Zerlegung von } [a,b]\}.\end{aligned}$$

Satz 5.30. *Seien $a < b < c$, dann gilt*

(a) $f \in \mathcal{R}_{[a,c]}$ dann und nur dann, wenn $f|_{[a,b]} \in \mathcal{R}_{[a,b]}$ und $f|_{[b,c]} \in \mathcal{R}_{[b,c]}$.

(b) Ist $f \in \mathcal{R}_{[a,c]}$, dann gilt

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx.$$

Hierbei haben wir die Notationen

$$\int_a^b f(x)dx := \int_a^b f|_{[a,b]}(x)dx \qquad \text{sowie} \qquad \int_b^c f(x)dx := \int_b^c f|_{[b,c]}(x)dx$$

verwendet.

Der Beweis dieses Satzes ist eine Übungsaufgabe. Für Teil (a) können Sie z. B. die Charakterisierung aus Satz 5.16 verwenden, und für Teil (b) z. B. die Aussage (iii) oder (iv) von Theorem 5.27 benutzen.

Definition 5.31. Sei $f \in \mathcal{R}_{[a,c]}$, dann definieren wir

(i) $\int\limits_b^b f(x)dx := 0$ für $b \in [a,c]$,

(ii) $\int\limits_{b_2}^{b_1} f(x)dx := -\int\limits_{b_1}^{b_2} f(x)dx$ für $a \leq b_1 \leq b_2 \leq c$.

Aufgabe 5.32. Für $f \in \mathcal{R}_{[a,b]}$ und $p \geq 1$ definieren wir

$$\|f\|_p := \left(\int_a^b |f(x)|^p\right)^{\frac{1}{p}}.$$

Zeigen Sie: Seien $f, g \in \mathcal{R}_{[a,b]}$, dann gilt:

(a) $\|f+g\|_p \leq \|f\|_p + \|g\|_p$, falls $p \geq 1$ (Dreiecksungeichung).

(b) Falls $p, q > 1$ und $\frac{1}{p} + \frac{1}{q} = 1$, dann $\int_a^b |f(x)g(x)|dx \leq \|f\|_p \cdot \|g\|_q$.

Bemerkung 5.33. Für $p = 1$ erhalten wir mit Aufgabe 5.32, Teil (a), sofort die „Dreiecksungleichung“ für $f, g \in \mathcal{R}_{[a,b]}$:

$$\int_a^b |f(x) + g(x)|dx \;\leq\; \int_a^b |f(x)|dx + \int_a^b |g(x)|dx \tag{5.3}$$

Analog zur üblichen „umgekehrten Dreiecksungleichung“ erhalten wir

$$\int_a^b |f(x) - g(x)|dx \;\geq\; \int_a^b |f(x)|dx - \int_a^b |g(x)|dx \tag{5.4}$$

Zum Beweis wendet man die Ungleichung (5.3) auf die Riemann-integrierbaren Funktion $h_1 = f - g$ und $h_2 = g$ an.

Theorem 5.34 (Vertauschung von Integration und Limes bei gleichmäßiger Konvergenz)**.** *Seien $f_n : [a, b] \to \mathbb{R}$, $n \in \mathbb{N}$ Riemann–integrierbare Funktionen, welche gleichmäßig gegen eine Funktionen $f : [a, b] \to \mathbb{R}$ konvergieren. Dann gilt*

(a) f ist Riemann–integrierbar.

(b) $\int\limits_a^b f(x)dx = \lim\limits_{n\to\infty} \left(\int\limits_a^b f_n(x)dx \right)$.

Beweis. Sei $\varepsilon > 0$ und $\tilde{\varepsilon} := \frac{\varepsilon}{2(1+b-a)}$. Da f_n gleichmäßig gegen f konvergiert, gibt es ein $N \in \mathbb{N}$, sodass für alle $x \in [a, b]$ gilt $f_N(x) - \tilde{\varepsilon} \leq f(x) \leq f_N(x) + \tilde{\varepsilon}$. Da f_N Riemann–integrierbar ist, gibt es Treppenfunktionen $\phi, \psi \in \mathcal{T}_{[a,b]}$ mit

$$\phi \leq f_N \leq \psi \qquad \text{und} \qquad \int_a^b (\psi(x) - \phi(x))dx \leq \varepsilon.$$

Damit gilt für die Treppenfunktionen $\phi - \tilde{\varepsilon}$ und $\psi + \tilde{\varepsilon}$ auch

$$\phi - \tilde{\varepsilon} \;\leq\; f_N - \tilde{\varepsilon} \;\leq\; f \;\leq\; f_N + \tilde{\varepsilon} \;\leq\; \psi + \tilde{\varepsilon}$$

und

$$\begin{aligned} \int_a^b \big((\psi(x) + \tilde{\varepsilon}) - (\phi(x) - \tilde{\varepsilon})\big)dx \;&\leq\; 2\tilde{\varepsilon}(b - a) + \int_a^b \big(\psi(x) - \phi(x)\big)dx \\ &\leq\; 2\tilde{\varepsilon}(b - a) + \tilde{\varepsilon} \leq 2\tilde{\varepsilon}(1 + b - a) = \epsilon. \end{aligned}$$

Somit ist auch f nach Satz 5.16 Riemann–integrierbar.

Um nun das Integral von f zu berechnen, verwenden wir die Monotonie des Integrals und erhalten so

$$\left| \int_a^b f(x)dx - \int_a^b f_n(x)dx \right| \leq \int_a^b |f(x) - f_n(x)|dx \leq (b - a)\|f - f_n\|_{\sup}.$$

Da f_n gleichmäßig gegen f konvergiert, gilt $\lim\limits_{n\to\infty}(\|f-f_n\|_{\sup})=0$, also

$$\int\limits_a^b f(x)dx = \lim_{n\to\infty}\left(\int\limits_a^b f_n(x)dx\right).$$

□

Bemerkung 5.35. Die gleichmäßige Konvergenz in Theorem 5.34 ist essentiell notwendig, wie das folgende Beispiel zeigt (vgl. Staatsexamen 43910, Frühjahr 2010,Thema 3, Aufgabe 4): Die Funktionenfolge

$$f_n : [0,1] \to \mathbb{R}, \quad x \mapsto \frac{2nx}{(1+nx^2)^2}$$

konvergiert punktweise, aber *nicht* gleichmäßig gegen die konstante Funktion 0 auf $[0,1]$. Jedoch konvergiert die Folge $a_n := \int\limits_0^1 f_n(x)dx$ gegen 1 und nicht gegen 0 (Übungsaufgabe).

Bemerkung 5.36. Einer der ersten mathematisch gut fundierten Integralbegriffe geht auf Cauchy zurück. Dieser Cauchysche Integralbegriff gilt für Regelfunktionen. *Regelfunktionen* auf $[a,b]$ sind hierbei Funktionen $f:[a,b]\to\mathbb{R}$, welche gleichmäßige Grenzwerte von Treppenfunktionen sind, d. h. $f:[a,b]\to\mathbb{R}$ ist genau dann eine Regelfunktion, wenn es eine Folge $\phi_n \in \mathcal{T}_{[a,b]}$ gibt, sodass $\lim\limits_{n\to\infty}\|f-\phi_n\|_{\sup}=0$ gilt[125]. Das Integral einer solchen Regelfunktion f mit $\lim\limits_{n\to\infty}\|f-\phi_n\|_{\sup}=0$ ist dann definiert als[126] $\int_a^b f(x)dx := \lim\limits_{n\to\infty}\phi_n(x)dx$. Theorem 5.34 zeigt, dass das Riemann–Integral für Regelfunktionen mit dem Cauchyschen Integralbegriff übereinstimmt. Somit ist das Riemann–Integral eine Verallgemeinerung dessen Cauchy–Integrals für Regelfunktionen.

5.4. Hauptsatz der Differential- und Integralrechnung

Zwar ist es bisweilen grundsätzlich möglich das Riemann–Integral einer integierbaren Funktion $f:[a,b]\to\mathbb{R}$ mittels Riemannscher oder Darbouxscher Summen zu berechnen, jedoch ist dieser Weg meistens mit viel Mühen verbunden. Für stetige Funktionen werden wir feststellen, dass die Integration die Umkehroperation der Ableitung ist. Mit Hilfe dieses Wissens und einiger Integrationsregeln kann man das Integral einer gegebenen stetigen Funktion in zahlreichen Fällen explizit ausrechnen. Jedoch ist integrieren viel schwieriger als ableiten. Um mit meinen hervorragenden Mathematiklehrer Wolfgang Roth zu sprechen: „Ableiten ist Technik, integrieren ist Kunst!“

Wir folgen in diesem Abschnitt [Fo1, § 19].

In diesem Abschnitt bezeichnen I und J jeweils offene Intervalle in $\mathbb{R}$.

[125] Stetige oder monotone Funktionen sind Regelfunktionen

[126] Eine Einführung in die Integralrechnung mittels Regelfunktionen finden Sie z. B. in [BF1, Kap. 10] oder [K1, Kap. 12].

Definition 5.37. Sei $I \subset \mathbb{R}$ ein offenes Intervall. Eine differenzierbare Funktion $F : I \to \mathbb{R}$ heißt *Stammfunktion* von f, wenn die Ableitung von F gerade f ist, also $f = F'$ gilt.

Aus Korollar 4.23 folgt einfach:

Lemma 5.38. *Ist $F : I \to \mathbb{R}$ eine Stammfunktion von $f : I \to \mathbb{R}$ und $c \in \mathbb{R}$, dann ist auch $F+c$ eine Stammfunktion von f. Umgekehrt unterscheiden sich zwei Stammfunktion $F_1, F_2 : I \to \mathbb{R}$ von $f : I \to \mathbb{R}$ gerade um eine additive Konstante*[127]*, d. h. $F_2 = F_1 + c$ für ein passendes $c \in \mathbb{R}$.*

Theorem 5.39 (Hauptsatz der Differential- und Integralrechnung, 1. Teil)**.**
Sei $I \subset \mathbb{R}$ ein offenes Intervall, $a \in I$ und $f : I \to \mathbb{R}$ eine stetige *Funktion, dann ist die Funktion*

$$F : I \to \mathbb{R}, \quad x \mapsto \int_a^x f(t)dt$$

eine Stammfunktion von f, d. h. F ist differenzierbar und es gilt $F' = f$.[128]

Beweis. Für $h \neq 0$ hinreichend klein und $x \in I$ betrachten wir den Differenzenquotienten

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h}\left(\int_a^{x+h} f(t)dt - \int_a^x f(t)dt\right) = \frac{1}{h}\int_x^{x+h} f(t)dt$$

Nach dem Mittelwertsatz der Integralrechnung, Theorem 5.22, gibt es ein ξ_h zwischen x und $x+h$, d. h. $\xi_h \in [x, x+h]$ für $h > 0$ bzw. $\xi_h \in [x+h, x]$ für $h < 0$, sodass gilt

$$\int_x^{x+h} f(t)dt = h \cdot f(\xi_h).$$

Da $\lim\limits_{h\to 0}(\xi_h) = x$, erhalten wir wegen der Stetigkeit von f :

$$\lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to 0} f(\xi_h) = f(x).$$

Damit ist F in x differenzierbar mit $F'(x) = f(x)$. $\square$

Theorem 5.40 (Hauptsatz der Differential- und Integralrechnung, 2. Teil)**.**
Sei $I \subset \mathbb{R}$ ein offenes Intervall, $f : I \to \mathbb{R}$ eine stetige *Funktion und F eine Stammfunktion von f. Dann gilt für alle $a, b \in I$:*

$$\int_a^b f(x)dx = F(b) - F(a).$$

Bemerkung 5.41. Für $F(b) - F(a)$ sind auch die Bezeichnungen $[F(x)]_a^b$ und $F(x)\big|_a^b$ gebräuchlich.

[127] Hierbei ist wichtig, dass ein Intervall immer eine zusammenhängende Teilmenge von $\mathbb{R}$ ist.

[128] Die $F : I \to \mathbb{R}$, $x \mapsto \int_a^x f(t)dt$ wird oft *unbestimmtes Integral* von f genannt.

Beweis. Die Funktion
$$F_0 : I \to \mathbb{R}, \quad x \mapsto \int_a^x f(t)dt$$
ist nach Theorem 5.39 eine Stammfunktion von f und es gilt $F_0(a) = 0$.

Ist F eine beliebige Stammfunktion, so gilt $F = F_0 + c$ für ein passendes $c \in \mathbb{R}$ und damit
$$\int_a^b f(t)dt = F_0(b) = F_0(b) - F_0(a) = (F_0(b) + c) - (F_0(a) + c) = F(b) - F(a).$$
□

Bemerkung 5.42. Für eine Stammfunktion F einer stetigen Funktion $f : [a, b] \to \mathbb{R}$ schreibt man bisweilen
$$\int f(x)dx.$$
Diese Notation ist natürlich etwas unpräzise, da Stammfunktionen nur bis auf eine additive Konstante eindeutig bestimmt sind.

Aus Theorem 5.34 kann man mit Hilfe des Hauptsatzes der Differential- und Integralrechnung zeigen:

Satz 5.43 (Vertauschung von Grenzwerten und Ableitungen)**.** *Sei $f_n : [a, b] \to \mathbb{R}$ eine Folge stetig differenzierbarer Funktionen, welche punktweise gegen eine Funktion $f : [a, b] \to \mathbb{R}$ konvergiert. Wenn die Folge der Ableitungen f_n' gleichmäßig gegen eine Funktion $f^* : [a, b] \to \mathbb{R}$ konvergiert, dann gilt:*

(a) f ist differenzierbar.

(b) $f' = f^*$, d. h. für alle $x \in [a, b]$ gilt: $f'(x) = \lim_{n\to\infty} f_n'(x)$.

Der Beweis ist eine Übungsaufgabe.

Während es schwierig ist zu einer Funktion f eine Stammfunktion zu finden, ist es leicht nachzuweisen, dass eine Funktion F eine Stammfunktion von f ist, denn es genügt dazu F abzuleiten und $F' = f$ zu sehen. Viele Formelsammlungen bieten ein lange Tabelle von Stammfunktionen. Eine Liste von Stammfunktionen einiger weniger elementarer Funktionen ist:

Funktion $f(x)$	Stammfunktion $F(x)$
$x^n,\ n \in \mathbb{Z} \setminus \{-1\}$	$\frac{1}{n+1}x^{n+1}$
$\frac{1}{x}$	$\ln(\lvert x\rvert)$
$\sin(x)$	$-\cos(x)$
$\cos(x)$	$\sin(x)$
$\exp(x)$	$\exp(x)$
$\frac{1}{1+x^2}$	$\arctan(x)$

Vorsicht! Bei der Anwendung von Theorem 5.40 zur Bestimmung von $\int_a^b f(x)dx$ ist zu beachten, dass f auf ganz $[a,b]$ definiert und stetig sein muss.

Da nach Korollar 3.44 Potenzreihen im Inneren ihres Konvergenzbereichs absolut konvergieren folgt mit Theorem 5.34:

Satz 5.44. *Sei (a_n) eine reelle Zahlenfolge, sodass die Potenzreihe $\sum a_n x^n$ einen positiven Konvergenzradius $R > 0$ hat. Dann ist die Stammfunktion $F : (-R,R) \to \mathbb{R}$ der Funktion $f : (-R,R) \to \mathbb{R},\ x \mapsto \sum_{n=0}^{\infty} a_n x^n$ durch $F : (-R,R) \to \mathbb{R},\ x \mapsto \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}$ gegeben.*

In Theorem 5.39 haben wir gesehen, dass für stetige Funktionen die Integration in einem gewissen Sinne die Umkehroperation der Ableitung ist. Wir erhalten somit Analoga zu den Ableitungsregeln.

Als Folgerung aus der Kettenregel für Ableitungen erhalten wir:

Satz 5.45 (Substitutionsregel)**.** *Sei $f : I \to \mathbb{R}$ stetig und sei $g : J \to I$ stetig differenzierbar, wobei I und J offenen Intervalle in $\mathbb{R}$ sind. Für $a, b \in J$ gilt dann:*

$$\int_a^b f\big(g(t)\big)g'(t)dt = \int_{g(a)}^{g(b)} f(x)dx.$$

Beweis. Für eine Stammfunktion $F : I \to \mathbb{R}$ von f erhalten wir mit der Kettenregel

$$(F \circ g)'(t) = F'(g(t))g'(t) = f(g(t))g'(t)$$

und somit wegen Theorem 5.40:

$$\int_a^b f\big(g(t)\big)g'(t)dt = (F \circ g)(b) - (F \circ g)(a) = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(x)dx.$$

□

Als Konsequenz aus der Produktregel für Ableitungen erhalten wir

Satz 5.46 (Partielle Integration)**.** *Seien $f, g : I \to \mathbb{R}$ stetig differenzierbare Funktionen und $a, b \in I$, dann gilt:*

$$\int_a^b f(x)g'(x)dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x)dx.$$

Beweis. Nach der Produktregel gilt

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x)$$

Da alle vorkommenden Funktionen stetig sind, folgt aus Theorem 5.40 mit der Additivität des Integrals sofort:

$$\int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx = [f(x)g(x)]_a^b.$$

□

Wir erweitern die Liste der möglichen Integrationsverfahren noch um ein weiteres, die *Partialbruchzerlegung*. Dieses findet bei rationalen Funktionen, also Quotienten $f = \frac{p}{q}$ zweier polynomialer Funktionen p und q, Anwendung. Die Methode besteht darin, den Quotienten $f = \frac{p}{q}$ als Summe einfacher rationaler Funktionen zu schreiben und dann jeden der Summanden von f (aufgrund der Linearität des Integrals) einzeln zu integrieren. Jede reelle polynomiale Funktion q kann als Produkt irreduzibler Polynome über $\mathbb{R}$ geschrieben werden. Irreduzible Polynome haben keine reellen Nullstellen und entweder Grad 0, 1 oder 2. Den Beweis des Satzes über die Partialbruchzerlegung reeller rationaler Funktionen finden Sie z. B. in [H1, Abschnitt 69]:

Satz 5.47 (Partialbruchzerlegung)**.** *Sei $f = \frac{p}{q}$ eine reelle rationale Funktion, wobei wir annehmen, dass für die reellen polynomialen Funktionen p und q gilt* $1 \leq \mathrm{Grad}(p) < \mathrm{Grad}(q)$. *Sei*

$$q(x) = c \cdot \left(\prod_{i=1}^{r} (x - x_i)^{n_i} \right) \cdot \left(\prod_{k=1}^{s} (x^2 + a_k x + b_k)^{m_k} \right)$$

die Zerlegung von q in über $\mathbb{R}$ irreduzible Faktoren. Dann gibt es reelle Zahlen A_{ij}, B_{kl} und C_{kl}, sodass gilt

$$f(x) = \sum_{i=1}^{r} \sum_{j=1}^{n_i} \frac{A_{ij}}{(x - x_i)^j} + \sum_{k=1}^{s} \sum_{l=1}^{m_k} \frac{B_{kl}x + C_{kl}}{(x^2 + a_k x + b_k)^l}.$$

Die Koeffizienten A_{ij}, B_{kl} und C_{kl}, ergeben sich durch Koeffizientenvergleich.

Eine weitere Konsequenz aus dem Hauptsatz der Differential- und Integralrechnung ist die Integralform des Restgliedes in der Taylor–Formel (für die Lagrange–Form siehe Satz 4.59).

Satz 5.48 (Taylor–Formel mit Integralform des Restglieds.)**.** *Sei $n \in \mathbb{N}$ und $f : I \to \mathbb{R}$ $(n+1)$-mal stetig differenzierbar und $x_o \in I$. Dann gilt für alle $x \in \mathbb{R}$:*

$$R_n^{f,x_o}(x) = f(x) - T_n^{f,x_o}(x) = \frac{1}{n!} \int_{x_o}^{x} (x - t)^n f^{(n+1)}(t)dt.$$

Diese Form des Restglieds heißt Integralform.

Beweis. Für $n = 0$ ist unsere Aussage gerade die Behauptung des Satzes der Hauptsatz der Differential- und Integralrechnung:

$$f(x) - f(x_o) = \int_{x_o}^{x} f'(t)dt.$$

Angenommen, die Aussage ist für ein beliebiges aber festes $n - 1 \in \mathbb{N}$ wahr, dann gilt mit partieller Integration:

$$\begin{aligned}
f(x) - T^{f,x_o}_{n-1}(x) = R^{f,x_o}_{n-1}(x) &= \frac{1}{(n-1)!}\int_{x_o}^{x}(x-t)^{n-1}f^{(n)}(t)dt \\
&= -\int_{x_o}^{x} f^{(n)}(t)\frac{d}{dt}\left(\frac{(x-t)^n}{n!}\right)dt \\
&= \left[-f^{(n)}(t)\frac{(x-t)^n}{n!}\right]_{x_o}^{x} + \frac{1}{n!}\int_{x_o}^{x}(x-t)^n f^{(n+1)}(t)dt \\
&= \frac{f^{(n)}(x_o)}{n!}(x-x_o)^n + \frac{1}{n!}\int_{x_o}^{x}(x-t)^n f^{(n+1)}(t)dt.
\end{aligned}$$

Damit erhalten wir

$$f(x)-T^{f,x_o}_{n-1}(x)-\frac{f^{(n)}(x_o)}{n!}(x-x_o)^n = f(x)-T^{f,x_o}_{n}(x) = R^{f,x_o}_{n}(x) = \frac{1}{n!}\int_{x_o}^{x}(x-t)^n f^{(n+1)}(t)dt.$$

□

Bemerkung 5.49 (Riemann–Integral für $\mathbb{C}$-wertige Funktionen einer reellen Variable)**.** Eine komplexwertige Funktion $f : [a,b] \to \mathbb{C}$ heißt Riemann–integrierbar, wenn ihr Realteil $\Re(f)$ und ihr Imaginärteil $\Im(f)$ Riemann–integrierbar sind. Wir setzen in diesem Fall:

$$\int_a^b f(x)dx = \int_a^b \Re(f(x))dx + i\int_a^b \Im(f(x))dx.$$

Da $\Re(f') = (\Re(f))'$ und $\Im(f') = (\Im(f))'$ gilt auch der Hauptsatz der Differential- und Integralrechnung: Ist $f : [a,b] \to \mathbb{C}$ stetig, dann ist

$$F : [a,b] \to \mathbb{C}, \quad x \mapsto \int_a^x f(t)dt$$

eine Stammfunktion von F, d. h. $F' = f$. Stammfunktionen von f sind bis auf komplexe additive Konstanten eindeutig bestimmt. Der Beweis folgt komponentenweise für Real- und Imaginärteil aus dem Hauptsatz der Differential- und Integralrechnung für reellwertige stetige Funktionen. Ebenso gilt

$$\int_a^b f(x)dx = F(b) - F(a)$$

für eine Stammfunktion $F : [a,b] \to \mathbb{C}$ der stetigen Funktion $f : [a,b] \to \mathbb{C}$. Die Rechenregeln gelten analog. Die Beweise erfolgen komponentenweise.

5.5. Uneigentliche Riemann–Integrale

Wir folgen in diesem Abschnitt der Darstellung in [Fo1, § 20].

Definition 5.50.

- Sei $a < b$ mit $b \in \mathbb{R} \cup \{\infty\}$ und $f : [a, b) \to \mathbb{R}$ eine Funktion, sodaß $f|_{[a,R]}$ für alle $R \in (a, b)$ Riemann–integrierbar ist. Falls der Grenzwert $\lim\limits_{R\uparrow b} \int\limits_a^R f|_{[a,R]}(x)dx = \lim\limits_{R\uparrow b} \int\limits_a^R f(x)dx$ existiert, so setzen wir
$$\int\limits_a^b f(x)dx := \lim_{R\uparrow b} \int_a^R f(x)dx.$$
- Sei $a < b$ mit $a \in \mathbb{R} \cup \{-\infty\}$ und $f : (a, b] \to \mathbb{R}$ eine Funktion, sodaß $f|_{[R,b]}$ für alle $R \in (a, b)$ Riemann–integrierbar ist. Falls der Grenzwert $\lim\limits_{R\downarrow a} \int\limits_R^b f|_{[R,b]}(x)dx = \lim\limits_{R\downarrow a} \int\limits_R^b f(x)dx$ existiert, so setzen wir
$$\int\limits_b^a f(x)dx := \lim_{R\downarrow a} \int\limits_R^b f(x)dx.$$
- Sei $a < b$ mit $a \in \mathbb{R} \cup \{-\infty\}$ und $b \in \mathbb{R} \cup \{\infty\}$ und $c \in (a, b)$. Sei $f : (a, b) \to \mathbb{R}$ eine Funktion, sodass für alle $a < R_1 < c < R_2 < b$ die Grenzwerte $\lim\limits_{R_2\uparrow b} \int\limits_c^{R_2} f(x)dx$ und $\lim\limits_{R_1\downarrow a} \int\limits_{R_1}^c f(x)dx$ beide existieren, dann setzen wir
$$\int_a^b f(x)dx = \lim_{R_1\downarrow a} \int\limits_{R_1}^c f(x)dx + \lim_{R_2\uparrow b} \int\limits_c^{R_2} f(x)dx$$
Diese Definition ist unabhängig von der Wahl von $c \in (a, b)$.

Bemerkung 5.51. Für eine stetige Funktion $f : \mathbb{R} \to \mathbb{R}$ ist der *Cauchysche Hauptwert* von f, notiert $\mathrm{CH}(f)$, definiert als
$$\mathrm{CH}(f) := \lim_{R\to\infty} \int_{-R}^R f(x)dx,$$
falls dieser Grenzwert existiert.

Wenn $\int\limits_{-\infty}^{\infty} f(x)dx$ erklärt ist, so existiert auch der Cauchysche Hauptwert von f und es gilt $\mathrm{CH}(f) = \int\limits_{-\infty}^{\infty} f(x)dx$. Umgekehrt hingegen kann der Cauchysche Hauptwert existieren, ohne dass $\int\limits_{-\infty}^{\infty} f(x)dx$ erklärt sein muss:

Für $f : \mathbb{R} \to \mathbb{R},\ x \mapsto \sin(x)$ existiert $\lim\limits_{R_2 \uparrow \infty} \int\limits_0^{R_2} f(x)dx$ offenbar *nicht*. Somit ist auch $\int\limits_{-\infty}^{\infty} \sin(x)dx$ nicht erklärt. Hingegen gilt $\mathrm{CH}(\sin) = 0$.

Wir lernen nun noch ein Konvergenzkriterium für Reihen kennen, welches die Integralrechnung verwendet.

Satz 5.52 (Integralvergleichskriterium). *Sei $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ eine monoton fallende nicht-negative Funktion, dann konvergiert die Reihe $\sum f(n)$ genau dann, wenn der $\lim\limits_{R\to\infty} \int\limits_0^R f(x)dx$ konvergiert, also $\int\limits_0^\infty f(x)dx$ existiert. Es gilt in diesem Fall für alle $n \in \mathbb{N}$:*

$$\sum_{j=n+1}^{\infty} f(j) \leq \int_n^\infty f(x)dx \leq \sum_{j=n}^{\infty} f(j).$$

Beweis. Da f monoton fallend ist, gilt für alle $n \in \mathbb{N}$:

$$f(n+1) = \int_n^{n+1} f(n+1)dx \leq \int_n^{n+1} f(x)dx \leq f(n).$$

Somit ergibt sich für alle $n \in \mathbb{N}$ und alle $N \in \mathbb{N}^*$:

$$\sum_{j=n+1}^{n+N} f(j) \leq \int_n^{n+N} f(x)dx \leq \sum_{j=n}^{n+N} f(j).$$

Durch Übergang zum Grenzwert ergibt sich die Abschätzung. Für $n = 0$ gilt, falls $\int_0^\infty f(x)dx = \lim_{N\to\infty} \int_0^N f(x)dx$ existiert, so konvergiert auch $\sum f(n)$, da die Folge der Partialsummen monoton wachsend und nach oben beschränkt ist.

Konvergiert nun $\sum f(n)$, so ist $R \mapsto \int_0^R f(x)dx$ monoton wachsend und beschränkt, also konvergent. $\square$

Aufgabe 5.53. Betrachte $\ln_1 := \ln : (0,\infty) \to \mathbb{R}$, $\ln_2 := \ln \circ \ln_1 : D_2 = (1,\infty) \to \mathbb{R}$ und $\ln_{p+1} := \ln \circ \ln_p : D_{p+1} = \{x \in \mathbb{R}_{\geq 0} : \ln_p(x) > 0\}$ für $p \in \mathbb{N}_{\geq 2}$.

(a) Zeigen Sie: Für $p \in \mathbb{N}_{\geq 2}$ gilt $(\ln_p)'(x) = \left(x \cdot \prod\limits_{j=1}^{p-1} \ln_j(x)\right)^{-1}$, $x \in D_p$.

(b) Sei $p \in \mathbb{N}^*$. Zeigen Sie: Die Reihe $\sum\limits_{n=N}^{\infty} \left(n \cdot (\ln_p(n))^\alpha \cdot \prod\limits_{j=1}^{p-1} \ln_j(n)\right)^{-1}$, $N \in D_p$, konvergiert genau für $\alpha > 1$. Was bedeutet das speziell für $p = 1$?

5.6. Riemann–Stieltjes–Integral

Das Riemann–Stieltjes–Integral steht in gewisser Weise zwischen dem Riemann–Integral und dem modernen Lebesgue–Integral. Das Riemann–Stieltjes–Integral wird nicht in allen einführenden Analysisbüchern behandelt. Wir orientieren uns an [GBGW2, Abschn. 6.4], [H1, Kap. XI] und insbesondere an [W2, § 6].

Definition 5.54. Für zwei Funktionen $f, g : [a, b] \to \mathbb{R}$, eine Zerlegung $Z = \{a = x_o, x_1, \dots, x_n = b\}$ von $[a, b]$ und Stützstellen $\xi_1, \dots, \xi_n$ bzgl. Z, d. h. $\xi_j \in [x_{j-1}, x_j]$, heißt die Summe

$$S(f, dg, Z, \xi_1, \dots, \xi_n) := \sum_{j=1}^{n} f(\xi_j)\big(g(x_j) - g(x_{j-1})\big)$$

Riemann–Stieltjes–Summe von f bezüglich g und der Zerlegung Z mit Stützstellen $\xi_1, \dots, \xi_n$.

Definition 5.55. Seien $f, g : [a, b] \to \mathbb{R}$ zwei Funktionen, wobei $a < b$. Wir nennen f *Riemann–Stieltjes–integrierbar* bezüglich der *Gewichtsfunktion*[129] g oder kurz *RS–integrierbar* bzgl. g, wenn gilt:

Es gibt eine Zahl $\iota \in \mathbb{R}$ mit folgender Eigenschaft: Für alle $\varepsilon > 0$ gibt es eine Zerlegung $\tilde{Z}_\varepsilon$, sodass für jede Verfeinerung $Z = \{a = x_0, \dots, x_n = b\}$ von $\tilde{Z}_\varepsilon$ und jede Wahl von Stützstellen $\xi_j \in [x_{j-1}, x_j],\ j = 1, \dots, n$ bzgl. Z gilt:

$$|\iota - S(f, dg, Z, \xi_1, \dots, \xi_n)| \le \varepsilon.$$

Die (eindeutig bestimmte[130]) Zahl ι heißt in diesem Fall *Riemann–Stieltjes–Integral von f bezüglich g*, oder kurz *RS-Integral* von f bzgl. g.

Bemerkung 5.56.

1. Gemäß Theorem 5.27 stimmt für die Identitätsfunktion $g(x) = x$ das Riemann–Stieltjes–Integral mit dem Riemann–Integral überein.
2. Ist die Gewichtsfunktion $g : [a, b] \to \mathbb{R}$ konstant, so ist *jede* Funktion $f : [a, b] \to \mathbb{R}$ RS-integrierbar bzgl. g und es gilt

$$\int_a^b f(x)dg(x) = 0,$$

da ja jede Riemann–Stieltjes-Summe von f bzgl. g verschwindet.

Aufgabe 5.57. Sei $H : [-1, 1] \to \mathbb{R}$ die Heaviside[131]–Funktion definiert durch $H|_{[-1,0]} := 0$ und $H|_{(0,1]} := 1$. Zeigen Sie: Ist $f : [-1, 1] \to \mathbb{R}$ in $x_o = 0$ rechtsseitig stetig[132], dann ist f bzgl. H Riemann–Stieltjes–integrierbar mit

$$\int_{-1}^{1} f(x)dH(x) = f(0).$$

[129] Die Funktion g wird auch *Integrator* genannt.
[130] Der Beweis der Eindeutigkeit ist eine Übung.
[131] Oliver Heaviside (1850–1925), britischer Physiker)
[132] d. h. $\lim\limits_{x \downarrow x_o} f(x) = f(x_o)$

Bemerkung 5.58. Stieltjes ursprüngliche Definition des Riemann–Stieltjes–Integral war anders formuliert. Er forderte hierzu die Existenz einer Zahl $\tilde{\iota}$ mit der Eigenschaft, dass es für jedes $\varepsilon > 0$ es ein $\delta > 0$ gäbe, sodass für alle Zerlegungen $Z = \{x_0, \ldots, x_n\}$ von $[a, b]$ der Feinheit $\mu_Z \leq \delta$ und jede Wahl $\xi_1, \ldots, \xi_n$ von Stützstellen bzgl. Z gilt

$$|\tilde{\iota} - S(f, dg, Z, \xi_1, \ldots, \xi_n)| \leq \varepsilon.$$

Man sieht leicht, dass jede Funktion, die nach Stieltjes ursprünglicher Definition integrierbar ist auch nach unserer Definition 5.55 Riemann–Stieltjes–integrierbar ist. Für das Riemann–Integral, also für $g(x) = x$ stimmen beide Begriffe auch überein (Theorem 5.27). Pollard[133] beobachtete, dass, anders als für das Riemann–Integral, die von uns oben gewählte Definition 5.55 des Riemann–Stieltjes–Integrals allgemeiner als Stieltjes ursprünglicher Begriff ist, weshalb der von uns in Definition 5.55 als Riemann–Stieltjes–Integral eingeführte Integralbegriff in der Literatur auch manchmal „verallgemeinertes Riemann–Stieltjes–Integral“ genannt wird[134]. Eine eingehendere Diskussion dieses Phänomens finden Sie in [W2, § 6].

Mit Hilfe der Dreiecksungleichung erhalten wir folgende Charakterisierung Riemann–Stieltjes–integrierbarer Funktion:

Lemma 5.59 (Cauchy–Kriterium für Riemann–Stieltjes–Integrierbarkeit)**.** *Eine Funktion $f : [a, b] \to \mathbb{R}$ ist dann und nur dann bzgl. $g : [a, b] \to \mathbb{R}$ Riemann–Stieltjes–integrierbar, wenn gilt: Für alle $\varepsilon > 0$ gibt es eine Zerlegung $\tilde{Z}_\varepsilon$ von $[a, b]$, sodass für alle Verfeierungen Z und Z' von $\tilde{Z}_\varepsilon$ und jede Wahl von Stützstellen $\xi_1, \ldots, \xi_n$ bzgl. Z und $\xi'_1, \ldots, \xi'_{n'}$ bzgl. Z' gilt:*

$$|S(f, dg, Z, \xi_1, \ldots, \xi_n) - S(f, dg, Z', \xi'_1, \ldots, \xi'_{n'})| \leq \varepsilon.$$

Der Beweis ist eine Übungsaufgabe.

Sind $\lambda_1, \lambda_2 \in \mathbb{R}$ und f_1, f_2, f, g_1, g_2 und g reellwertige Funktionen auf $[a, b]$, so gilt für jede Zerlegung $Z = \{a = x_0, x_1, \ldots, x_n = b\}$ von $[a, b]$ und jede Wahl $\xi_1, \ldots, \xi_n$ von Stützstellen bzgl. Z :

$$S(\lambda_1 f_1 + \lambda_2 f_2, dg, Z, \xi_1, \ldots, \xi_n) = \lambda_1 S(f_1, dg, Z, \xi_1, \ldots, \xi_n) + \lambda_2 S(f_2, dg, Z, \xi_1, \ldots, \xi_n)$$

und

$$S(f, d(\lambda_1 g_1 + \lambda_2 g_2), Z, \xi_1, \ldots, \xi_n) = \lambda_1 S(f, dg_1, Z, \xi_1, \ldots, \xi_n) + \lambda_2 S(f, dg_2, Z, \xi_1, \ldots, \xi_n).$$

Daraus ergibt sich:

Satz 5.60 (Linearität des Riemann–Stieltjes–Integrals)**.**

[133] S. Pollard, The Stieltjes integral and its generalizations, *Quarterly J. Math.*, **49** (1920-23), 73–138

[134] Bezeichne $H : [-1, 1] \to \mathbb{R}$ die Heaviside–Funktion. Während nach unserer Definition 5.55 die Funktion $f : [-1, 1] \to \mathbb{R}$, $x \mapsto H(-x)$ bzgl. $g = H$ Riemann–Stieltjes–integrierbar ist (Übungsaufgabe), stimmt das jedoch nicht für Stieltjes ursprüngliche Definition.

(i) *Seien* $f_1, f_2 : [a,b] \to \mathbb{R}$ *bzgl.* $g : [a,b] \to \mathbb{R}$ *Riemann–Stieltjes–integrierbar und* $\lambda_1, \lambda_2 \in \mathbb{R}$. *Dann ist auch* $\lambda_1 f_1 + \lambda_2 f_2$ *bzgl.* g *Riemann–Stieltjes–integrierbar und es gilt*

$$\int_a^b (\lambda_1 f_1(x) + \lambda_2 f_2(x))dg(x) = \lambda_1 \int_a^b f_1(x)dg(x) + \lambda_2 \int_a^b f_2(x)dg(x).$$

(ii) *Seien* $f : [a,b] \to \mathbb{R}$ *bzgl.* $g_1 : [a,b] \to \mathbb{R}$ *und* $g_2 : [a,b] \to \mathbb{R}$ *Riemann–Stieltjes–integrierbar und* $\lambda_1, \lambda_2 \in \mathbb{R}$. *Dann ist* f *auch bzgl.* $\lambda_1 g_1 + \lambda_2 g_2$ *Riemann–Stieltjes–integrierbar und es gilt*

$$\int_a^b f(x)d(\lambda_1 g_1(x) + \lambda_2 g_2(x)) = \lambda_1 \int_a^b f(x)dg_1(x) + \lambda_2 \int_a^b f(x)dg_2(x).$$

Definition 5.61. Für zwei Funktionen $f, g : [a,b] \to \mathbb{R}$ mit $a < b$ setzen wir

$$\int_a^a f(x)dg(x) := 0$$

und, falls f bzgl. g Riemann–Stieltjes–Integrierbar ist, so definieren wir

$$\int_b^a f(x)dg(x) := -\int_a^b f(x)dg(x).$$

Satz 5.62. *Seien* $f, g : [a,b] \to \mathbb{R}$ *Funktionen und* $c \in \mathbb{R}$ *mit* $a < c < b$, *dann ist* f *genau dann bzgl.* g *Riemann–Stieltjes–integrierbar, wenn die Funktionen* $f|_{[a,c]}$ *bgzl.* $g|_{[a,c]}$ *und* $f|_{[c,b]}$ *bgzl.* $g|_{[c,b]}$ *Riemann–Stieltjes–integrierbar sind. In diesem Falle gilt*

$$\int_a^b f(x)dg(x) = \int_a^c f(x)dg(x) + \int_c^b f(x)dg(x),$$

wobei $\int_a^c f(x)dg(x) := \int_a^c f|_{[a,c]}(x)dg|_{[a,c]}(x)$ *und* $\int_c^b f(x)dg(x) := \int_a^c f|_{[c,b]}(x)dg|_{[c,b]}(x)$.

Auch dieser Beweis ist eine Übungsaufgabe.

5.7. Rechenregeln für Riemann–Stieltjes–Integrale

Wir bereits im Fall des Riemann–Integrals ist es oft recht aufwendig, konkrete Integrale mittels ihrer Definition zu berechenen. In diesem Abschnitt geben wir einige wichtige Sätze, die bei der expliziten Bestimmung von Riemann–Stieltjes–Integralen nützlich sein können. Wir verweisen wieder auf [W2, § 6.3 u. 6.4], aber auch auf [GBGW2, Abschn. 6.4] und [H1, Kap. XI].

Satz 5.63 (Partielle Integration beim Riemann–Stieltjes–Integralen)**.**
Sei $f : [a,b] \to \mathbb{R}$ *bzgl.* $g : [a,b] \to \mathbb{R}$ *Riemann–Stieltjes–integrierbar, dann ist auch* g *bzgl.* f *Riemann–Stieltjes–integrierbar und es gilt*

$$\int_a^b f(x)dg(x) = [f(x)g(x)]_a^b - \int_a^b g(x)df(x).$$

Beweis. Sei $Z = \{a = x_0, x_1, \ldots, x_n = b\}$ eine Zerlegung von $[a, b]$ und $\xi_1, \ldots, \xi_n$ Stützstellen bzgl. Z, d. h. $\xi_j \in [x_{j-1}, x_j]$, $j = 1, \ldots, n$, dann gilt offenbar

$$\begin{aligned}[f(x)g(x)]_a^b &= \underbrace{\sum_{j=1}^n g(\xi_j)(f(x_j) - f(x_{j-1}))}_{=S(g,df,Z,\xi_1,\ldots,\xi_n)} \\ &\quad + \sum_{j=1}^n \big(f(x_{j-1})(g(\xi_j) - g(x_{j-1}) + f(x_j)(g(x_j) - g(\xi_j))\big).\end{aligned} \tag{5.5}$$

Sei $\varepsilon > 0$. Da f bzgl. g Riemann–Stieltjes–integrierbar ist, gibt es eine Zerlegung $\tilde{Z}_\varepsilon$, sodass für jede Verfeinerung Z von $\tilde{Z}_\varepsilon$ und jede Wahl von Stützstellen gilt $\xi_1, \ldots, \xi_n$ bzgl. Z

$$\left|\int_a^b f(x)dg(x) - S(f, dg, Z, \xi_1, \ldots, \xi_n)\right| \leq \varepsilon.$$

Wählen wir nun eine solche Verfeinerung $Z = \{x_0, \ldots, x_n\}$ von $\tilde{Z}_\varepsilon$ und Stützstellen $\xi_1, \ldots, \xi_n$ bzgl. Z. Damit ist $Z' = Z \cup \{\xi_1, \ldots, \xi_n\}$ ebenso eine Verfeinerung von $\tilde{Z}_\varepsilon$. Wählen wir als Menge der Stützstellen ξ' jeweils die linke Intervallgrenze in $[x_{j-1}, \xi_j]$ und jeweils die rechte Intervallgrenze in $[\xi_j, x_j]$, so erhalten wir[135]:

$$S(f, dg, Z', \xi') = \sum_{j=1}^n \big(f(x_{j-1})(g(\xi_j) - g(x_{j-1}) + f(x_j)(g(x_j) - g(\xi_j))\big).$$

Mit Gleichung (5.5) erhalten wir so

$$\begin{aligned} & \left|-\int_a^b f(x)dg(x) + [f(x)g(x)]_a^b - S(g, df, Z, \xi_1, \ldots, \xi_n)\right| \\ = & \left|\int_a^b f(x)dg(x) - [f(x)g(x)]_a^b + S(g, df, Z, \xi_1, \ldots, \xi_n)\right| \\ = & \left|\int_a^b f(x)dg(x) - \left(\sum_{j=1}^n \big(f(x_{j-1})(g(\xi_j) - g(x_{j-1}) + f(x_j)(g(x_j) - g(\xi_j))\big)\right)\right| \\ = & \left|\int_a^b f(x)dg(x) - S(f, dg, Z', \xi')\right| \\ \leq & \ \varepsilon. \end{aligned}$$

Somit ist auch g bzgl. f Riemann–Stieltjes-integrierbar mit

$$\int_a^b g(x)df(x) = [f(x)g(x)]_a^b - \int_a^b f(x)dg(x).$$

□

[135] Wenn $\xi_j \in \{x_{j-1}, x_j\}$ so tritt das Intervall $[x_{j-1}, \xi_j]$ bzw. $[\xi_j, x_j]$ nicht auf, aber der entsprechende Summand verschwindet.

Satz 5.64 (Riemann–Stieltjes–Integral und Riemann–Integral)**.** *Sei $f \in \mathcal{R}_{[a,b]}$ und $g : [a,b] \to \mathbb{R}$ stetig differenzierbar*[136]*, dann ist f bzgl. g Riemann–Stieltjes–integrierbar mit*

$$\int_a^b f(x)dg(x) = \int_a^b f(t)g'(t)dt.$$

Beweis. Sei $\varepsilon > 0$ und $\tilde{\varepsilon} = \frac{\varepsilon}{2K(b-a)}$.

Da f Riemann–integrierbar und g' als stetige Funktion ebenfalls Riemann–integriebar ist, gilt $fg' \in \mathcal{R}_{[a,b]}$. Nach Theorem 5.27 gibt es also ein $\delta_1 > 0$, sodass für jede Zerlegung Z mit $\mu_Z \leq \delta_1$ und jede Menge $\xi = \{\xi_1, \ldots, \xi_n\}$ von Stützstellen bzgl. Z gilt

$$\left| \int_a^b f(t)g'(t)dt - R(fg', Z, \xi) \right| \leq \frac{\varepsilon}{2}.$$

Als Riemann–integrierbare Funktion ist f beschränkt. Es gibt also ein $K \in \mathbb{R}$ mit $\|f\|_{\sup} \leq K$. Als stetige Funktion auf einem kompakten Intervall ist g' gleichmäßig stetig (Satz von Heine, Theorem 3.33). Damit gibt es ein $\delta_2 > 0$, sodass $|g'(x) - g'(y)| \leq \tilde{\varepsilon}$ für alle $x, y \in [a,b]$ mit $|y - x| \leq \delta_2$ gilt. Sei $Z = \{a = x_0, x_1, \ldots, x_n = b\}$ eine Zerlegung von $[a,b]$ der Feinheit $\mu_Z \leq \delta := \min\{\delta_1, \delta_2\}$ und $\xi = \{\xi_1, \ldots, \xi_n\}$ eine Menge von Stützstellen bzgl. Z. Nach dem Mittelwertsatz der Differentialrechnung, Theorem 4.19, gibt es Punkte $\tilde{\xi}_j \in [x_{j-1}, x_j]$, $j = 1, \ldots, n$, sodass

$$S(f, dg, Z, \xi) = \sum_{j=1}^n f(\xi_j)(g(x_j) - g(x_{j-1})) = \sum_{j=1}^n f(\xi_j)g'(\tilde{\xi}_j)(x_j - x_{j-1}).$$

Für die Riemann–Summe von $f \cdot g'$ erhalten wir

$$R(fg', Z, \xi) = \sum_{j=1}^n f(\xi_j)g'(\xi_j)(x_j - x_{j-1}).$$

Damit gilt

$$|R(fg', Z, \xi) - S(f, dg, Z, \xi)| \leq \sum_{j=1}^n \underbrace{|f(\xi_j)|}_{\leq K} \cdot \underbrace{|g'(\tilde{\xi}_j) - g'(\xi_j)|}_{\leq \tilde{\varepsilon}} \cdot (x_j - x_{j-1}) \leq K\tilde{\varepsilon}(b-a) = \frac{\varepsilon}{2}.$$

Die Ungleichung $|R(fg', Z, \xi) - S(f, dg, Z, \xi)| \leq \frac{\varepsilon}{2}$ gilt also für alle Zerlegungen von $[a,b]$ der Feinheit $\mu_Z \leq \delta = \min\{\delta_1, \delta_2\}$ und eine beliebige Menge ξ von Stützstellen bzgl. Z. Wir wählen nun als $\tilde{Z}_\varepsilon$ eine Zerlegung mit $\mu_{\tilde{Z}_\varepsilon} \leq \delta$. Sei Z eine Verfeinerungen von $\tilde{Z}_\varepsilon$ und ξ eine beliebige Menge von Stützstellen bzgl. Z, dann ist $\mu_Z \leq \delta$, und wir erhalten

$$\begin{aligned} &\left| \int_a^b f(t)g'(t)dt - S(f, dg, Z, \xi) \right| \\ \leq\ &\left| \int_a^b f(t)g'(t)dt - R(fg', Z, \xi) \right| + |R(fg', Z, \xi) - S(f, dg, Z, \xi)| \\ \leq\ &\varepsilon. \end{aligned}$$

[136] d. h. g ist Einschränkung einer stetig differenzierbaren Funktion $(a - \varepsilon, b + \varepsilon) \to \mathbb{R}$.

Somit ist f bzgl. g Riemann–Stieltjes–integrierbar, und das Riemann–Integral $\int_a^b f(t)g'(t)dt$ stimmt mit dem Riemann–Stieltjes–Integral $\int_a^b f(x)dg(x)$ überein. □

Korollar 5.65. *Ist $f : [a,b] \to \mathbb{R}$ stetig differenzierbar und $g \in \mathcal{R}_{[a,b]}$, dann ist f bzgl. g Riemann–Stieltjes–integrierbar.*

Bemerkung 5.66. Nach dem Hauptsatz der Differential- und Integralrechnung gilt also: Ist $f \in \mathcal{R}_{[a,b]}$ und $h : [a,b] \to \mathbb{R}$ stetig, dann gilt

$$\int_a^b f(t)h(t)dt = \int_a^b f(x)dg(x)$$

mit

$$g(t) = \int_a^t h(\tau)d\tau.$$

5.8. Funktionen von beschränkter Variation

Wir orientieren uns hier vor allem an [W2, § 5 u. 6], aber auch an [GBGW2, Kap. 6] und [H1, Kap. XI].

Zunächst haben wir das Riemann–Stieltjes–Integral ohne weitere Voraussetzungen an die Gewichtsfunktion g definiert. In diesem Abschnitt betrachten wir nun eine besondere Klasse von Gewichtsfunktionen g mit der wichtigen Eigenschaft, dass jede *stetige* Funktion f bzgl. g Riemann–Stieltjes–integrierbar ist.

Definition 5.67. Die *Variation* einer Funktion $g : [a,b] \to \mathbb{R}$ bzgl. einer Zerlegung $Z = \{x_0, \ldots, x_n\}$ von $[a,b]$ ist die nicht-negative Zahl

$$V(g,Z) := \sum_{j=1}^{n} |g(x_j) - g(x_{j-1})|.$$

Die *Totalvariation* einer Funktion $g : [a,b] \to \mathbb{R}$ ist

$$V_a^b(g) := \sup \big\{ V(g,Z) : \ Z \text{ Zerlegung von } [a,b] \big\} \in \mathbb{R}_{\geq 0} \cup \{\infty\}.$$

Falls $V_a^b(g) < \infty$ heißt g von *beschränkter Variation*.

Bemerkung 5.68.

- Ist $g : [a,b] \to \mathbb{R}$ monoton, so ist g von beschränkter Variation.

  Wenn g monoton wachsend ist, dann gilt für eine beliebige Zerlegung $Z = \{a = x_0, \ldots, x_n = b\}$ von $[a,b]$

  $$V(g,Z) = \sum_{j=1}^{n} |g(x_j) - g(x_{j-1})| = \sum_{j=1}^{n} (g(x_j) - g(x_{j-1})) = g(b) - g(a).$$

  Also $V_a^b(g) = g(b) - g(a)$. Wenn g monoton fallend ist, so erhalten wir analog $V_a^b(g) = g(a) - g(b)$.

- Sei $g : [a, b] \to \mathbb{R}$ Lipschitz-stetig mit Lipschitzkonstante L, dann ist g von beschränkter Variation.

  Sei $Z = \{a = x_0, \ldots, x_n = b\}$ eine beliebige Zerlegung von $[a, b]$, dann gilt:

$$V(g, Z) = \sum_{j=1}^{n} |g(x_j) - g(x_{j-1})| \leq L \sum_{j=1}^{n} (x_j - x_{j-1}) = L(b - a).$$

Lemma 5.69. *Sei $g : [a, b] \to \mathbb{R}$ von beschränkter Variation und $c \in \mathbb{R}$ mit $a < c < b$, dann sind auch $g|_{[a,c]}$ und $g|_{[c,b]}$ von beschränkter Variation, und es gilt*

$$V_a^b(g) = V_a^c(g) + V_c^b(g)$$

mit $V_a^c(g) := V_a^c(g|_{[a,c]})$ und $V_c^b(g) := V_c^b(g|_{[c,b]})$.

Beweis. Sei Z_1 eine beliebige Zerlegung von $[a, c]$ und Z_2 eine beliebige Zerlegung von $[c, b]$, dann ist $Z = Z_1 \cup Z_2$ eine Zerlegung von $[a, b]$ und wir erhalten

$$V(g|_{[a,c]}, Z_1) + V(g|_{[c,b]}, Z_2) = V(g, Z) \leq V_a^b(g).$$

Da $V(g|_{[a,c]}, Z_1)$ und $V(g|_{[c,b]}, Z_2)$ nicht-negativ sind, folgt sofort, dass $g|_{[a,c]}$ und $g|_{[c,b]}$ von beschränkter Variation sind. Da Z_1 und Z_2 beliebige gewählt waren erhalten wir durch Übergang zum Supremum sofort

$$V_a^c(g|_{[a,c]}) + V_c^b(g|_{[c,b]}) \leq V_a^b(g). \tag{5.6}$$

Sei andererseits Z eine beliebige Zerlegung von $[a, b]$ und $Z' = Z \cup \{c\}$. Wir setzen $Z'_1 := Z' \cap [a, c]$ und $Z'_2 := Z' \cap [c, b]$. Dann gilt $Z' = Z_1 \cup Z_2$. Aus der Dreiecksungleichung folgt $V(g, Z) \leq V(g, Z')$ und damit

$$V(g, Z) \leq V(g, Z') = V(g|_{[a,c]}, Z'_1) + V(g|_{[c,b]}, Z'_2) \leq V_a^c(g|_{[a,c]}) + V_c^b(g|_{[c,b]}).$$

Da Z beliebig gewählt war, erhalten wir durch Übergang zum Supremum:

$$V_a^b(g) \leq V_a^c(g|_{[a,c]}) + V_c^b(g|_{[c,b]}). \tag{5.7}$$

Die Ungleichungen (5.6) und (5.7) zusammen ergeben die Behauptung. □

Aufgabe 5.70. (a) Zeigen Sie: Seien $g_{1,2} : [a, b] \to \mathbb{R}$ zwei Funktionen von beschränkter Variation, dann gilt

$$V_a^b(g_1 + g_2) \leq V_a^b(g_1) + V_a^b(g_2).$$

(b) Zeigen Sie: Die Menge $\mathrm{BV}([a, b]) := \{g : [a, b] \to \mathbb{R} :\ g \text{ von beschränkter Variation}\}$ ist ein reeller Untervektorraum des Vektorraums aller Funktionen von $[a, b]$ nach $\mathbb{R}$.

Theorem 5.71. *Sei $g : [a, b] \to \mathbb{R}$ von beschränkter Variation, dann ist jede stetige Funktion $f : [a, b] \to \mathbb{R}$ bzgl. g Riemann–Stieltjes–integrierbar und es gilt*

$$\left| \int_a^b f(x) dg(x) \right| \leq \|f\|_{\sup} \cdot V_a^b(g).$$

Beweis. Sei $\varepsilon > 0$ und $\varepsilon' := \frac{\varepsilon}{2V_a^b(g)}$.

Als stetige Funktion auf einem kompakten Intervall ist f gleichmäßig stetig (Theorem 3.33). Damit gibt es ein $\delta > 0$, sodass für alle $x, y \in [a, b]$ mit $|x - y| \leq \delta$ gilt

$$|f(x) - f(y)| \leq \varepsilon'.$$

Als $\tilde{Z}_\varepsilon = \{\tilde{x}_0, \ldots, \tilde{x}_m\}$ wählen wir eine Zerlegung von $[a, b]$ der Feinheit $\mu_{\tilde{Z}_\varepsilon} \leq \delta$. Zudem wählen wir als Stützstellen bzgl. $\tilde{Z}_\varepsilon$ die Menge $\tilde{\xi} = \left\{\tilde{\xi}_1 = \tilde{x}_1, \ldots, \tilde{\xi}_m = \tilde{x}_m\right\}$. Damit erhalten wir als Riemann–Stieltjes–Summe

$$S(f, dg, \tilde{Z}_\varepsilon, \tilde{\xi}) = \sum_{j=1}^{m} f(\tilde{x}_j)(g(\tilde{x}_j) - g(\tilde{x}_{j-1})).$$

Sei $Z = \{x_0, \ldots, x_n\}$ eine beliebige Verfeinerung von $\tilde{Z}_\varepsilon$, dann gilt insbesondere $\mu_Z \leq \delta$. Zudem sei $\xi = \{\xi_1, \ldots, \xi_n\}$ eine Menge von Stützstellen bzgl. Z. Dann ist die Riemann–Stieltjes–Summe

$$S(f, dg, Z, \xi) = \sum_{j=1}^{n} f(\xi_j)(g(x_j) - g(x_{j-1})).$$

Da Z eine Verfeinerung ist, gibt es für alle $j = 0, \ldots m$ ein $k_j \in \{0, \ldots, n\}$, sodass $x_{k_j} = \tilde{x}_j$ gilt. Für die Differenz zwischen dem ersten Summanden

$$f(\tilde{x}_1)(g(\tilde{x}_1) - g(\tilde{x}_0)) = \sum_{j=1}^{k_1} f(\tilde{x}_1)(g(x_j) - g(x_{j-1}))$$

von $S(f, dg, \tilde{Z}_\varepsilon, \tilde{\xi})$ und den ersten k_1 Summanden von $S(f, dg, Z, \xi)$ erhalten wir mit $|f(\xi_j) - f(\tilde{x}_1)| \leq \varepsilon'$ für $j = 1, \ldots, k_1$ (da $|\xi_j - \tilde{x}_1| \leq |\tilde{x}_2 - \tilde{x}_1| \leq \delta$) und Lemma 5.69:

$$\begin{aligned}
&\left| f(\tilde{x}_1)(g(\tilde{x}_1) - g(\tilde{x}_0)) - \sum_{j=1}^{k_1} f(\xi_j)(g(x_j) - g(x_{j-1})) \right| \\
= \;& \left| \sum_{j=1}^{k_1} (f(\tilde{x}_1) - f(\xi_j)) \cdot (g(x_j) - g(x_{j-1})) \right| \\
\leq \;& \sum_{j=1}^{k_1} \underbrace{|f(\tilde{x}_1) - f(\xi_j)|}_{\leq \varepsilon'} \cdot |g(x_j) - g(x_{j-1})| \\
\leq \;& \varepsilon' \sum_{j=1}^{k_1} |g(x_j) - g(x_{j-1})| \\
\leq \;& \varepsilon' V_{\tilde{x}_0}^{\tilde{x}_1}\big(g|_{[\tilde{x}_0, \tilde{x}_1]}\big).
\end{aligned}$$

Analoge Abschätzungen erhält man für die übrigen Teilintervalle $[\tilde{x}_1, \tilde{x}_2] = [x_{k_1}, x_{k_2}]$, …, $[\tilde{x}_{m-1}, \tilde{x}_m] = [x_{k_{m-1}}, x_{k_m}]$ von $[a,b]$ bzgl. der Zerlegung $\tilde{Z}_\varepsilon$. Mit Lemma 5.69 und der Dreiecksungleichung folgt

$$\begin{aligned}
&\left|S(f, dg, Z, \xi) - S(f, dg, \tilde{Z}_\varepsilon, \tilde{\xi})\right| \\
\leq\; & \sum_{j=1}^{m} \left| f(\tilde{x}_j)(g(\tilde{x}_j) - g(\tilde{x}_{j-1})) - \sum_{l=k_{j-1}+1}^{k_j} f(\xi_l)(g(x_l) - g(x_{l-1}))\right| \\
\leq\; & \varepsilon' \cdot \sum_{j=1}^{m} V_{\tilde{x}_{j-1}}^{\tilde{x}_j}\left(g|_{[\tilde{x}_{j-1}, \tilde{x}_j]}\right) \;\leq\; \varepsilon' V_a^b(g) \;=\; \frac{\varepsilon}{2}.
\end{aligned}$$

Seien nun Z und Z' zwei Verfeinerungen von $\tilde{Z}_\varepsilon$ und ξ bzw. ξ' jeweils Mengen von Stützstellen bzgl. Z bzw. Z'. Dann erhalten wir mit der Dreiecksungleichung und dem Vorherigen:

$$\begin{aligned}
& |S(f, dg, Z, \xi) - S(f, dg, Z', \xi')| \\
=\; & \left|\left(S(f, dg, Z, \xi) - S(f, dg, \tilde{Z}_\varepsilon, \tilde{\xi})\right) + \left(S(f, dg, \tilde{Z}_\varepsilon, \tilde{\xi}) - S(f, dg, Z', \xi')\right)\right| \\
\leq\; & \underbrace{\left|S(f, dg, Z, \xi) - S(f, dg, \tilde{Z}_\varepsilon, \tilde{\xi})\right|}_{\leq \frac{\varepsilon}{2}} + \underbrace{\left|S(f, dg, \tilde{Z}_\varepsilon, \tilde{\xi}) - S(f, dg, Z', \xi')\right|}_{\leq \frac{\varepsilon}{2}} \\
\leq\; & \varepsilon.
\end{aligned}$$

Nach dem Cauchy–Kriterium 5.59 ist damit f bzgl. g Riemann–Stieltjes–integrierbar. Aus

$$|S(f, dg, Z, \xi)| \leq \|f\|_{\sup} \sum_{j=1}^{n} |g(x_j) - g(x_{j-1})| \leq \|f\|_{\sup} \cdot V_a^b(g)$$

folgt die Ungleichung

$$\left|\int_a^b f(x) dg(x)\right| \leq \|f\|_{\sup} \cdot V_a^b(g).$$

□

5.9. Mittelwertsätze für Riemann–Stieltjes–Integrale

Wir folgen hier [W2, Abschn. 6.7] auch [GBGW2, Kap. 6].

Ist $g : [a, b] \to \mathbb{R}$ monoton wachend und $Z = \{x_0, x_1, \ldots, x_n\}$ eine beliebige Zerlegung von $[a, b]$. Aus $x_{j-1} < x_j$ für $j = 1, \ldots n$ folgt dann $g(x_j) - g(x_{j-1}) \geq 0$. Sei nun ξ eine beliebige Menge von Stützstellen bzgl. Z. Für zwei Funktionen $f_{1,2} : [a, b] \to \mathbb{R}$ folgt aus $f_1 \leq f_2$ dann für eine monoton wachsende Gewichtsfunktion g sofort

$$S(f_1, dg, Z, \xi) \leq S(f_2, dg, Z, \xi).$$

Theorem 5.72 (Erster Mittelwertsatz für Riemann–Stieltjes–Integrale)**.** *Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion, welche bzgl. einer monoton wachsenden Gewichtsfunktion $g : [a,b] \to \mathbb{R}$ Riemann–Stieltjes-integrierbar ist. Dann gibt es μ mit $m := \inf\{f(x) : x \in [a,b]\} \le \mu \le \sup\{f(x) : x \in [a,b]\} =: M$, sodass*

$$\int_a^b f(x)dg(x) = \mu(g(b) - g(a)).$$

Falls f stetig ist, gibt es ein $c \in [a,b]$ mit $\mu = f(c)$ (Zwischenwerteigenschaft stetiger Funktionen).

Beweis. Sei $Z = \{x_0, \ldots, x_n\}$ eine Zerlegung von $[a,b]$ und $\xi = \{\xi_1, \ldots, \xi_n\}$ eine Menge von Stützstellen bzgl. Z. Dann gilt, da g monton wachsend ist:

$$m(g(b) - g(a)) \le \sum_{j=1}^{n} f(\xi_j)(g(x_j) - g(x_{j-1})) = S(f, dg, Z, \xi) \le M(g(b) - g(a))$$

und damit

$$m(g(b) - g(a)) \le \int_a^b f(x)dg(x) \le M(g(b) - g(a)).$$

Die Aussage ergibt sich nun unmittelbar. □

Theorem 5.73 (Zweiter Mittelwertsatz für Riemann–Stieltjes–Integrale)**.** *Ist $f : [a,b] \to \mathbb{R}$ eine monotone Funktion und $g : [a,b] \to \mathbb{R}$ stetig, dann ist f bzgl. g Riemann–Stieltjes–integrierbar und es gibt $c \in \mathbb{R}$ mit $a \le c \le b$, sodass*

$$\int_a^b f(x)dg(x) = f(a)(g(c) - g(a)) + f(b)(g(b) - g(c)).$$

Beweis. Wir nehmen ohne Einschränkung an, dass f monton wachsend ist (ansonsten betrachte $-f$ anstelle von f). Da f von beschränkter Variation ist, ist die stetige Funktion g nach Theorem 5.71 bzgl. f Riemann–Stieltjes–integrierbar. Nach Theorem 5.72 gibt es ein $c \in [a,b]$, sodass

$$\int_a^b g(x)df(x) = g(c)(f(b) - f(a)).$$

Nach Satz 5.63 ist dann auch f bzgl. g Riemann–Stieltjes–integrierbar, und es gilt

$$\begin{aligned}
\int_a^b f(x)dg(x) &= [f(x)g(x)]_a^b - \int_a^b g df(x) \\
&= f(b)g(b) - f(b)g(b) - g(c)(f(b) - f(a)) \\
&= f(a)(g(c) - g(a)) + f(b)(g(b) - g(c)).
\end{aligned}$$

□

Korollar 5.74 (Zweiter Mittelwertsatz für Riemann–Integrale)**.** *Ist* $f : [a, b] \to \mathbb{R}$ *eine monotone und* $h : [a, b] \to \mathbb{R}$ *eine stetig Funktion, dann gibt es ein* $c \in [a, b]$, *sodass*

$$\int_a^b f(t)h(t)dt = f(a)\int_a^c h(t)dt + f(b)\int_c^b h(t)dt.$$

Beweis. Diese Aussage folgt aus dem Hauptsatz der Differential- und Integralrechnung zusammen mit Theorem 5.73 und Satz 5.64, indem man als Gewichtsfunktion $g : [a, b] \to \mathbb{R}$, $x \mapsto \int_a^x h(s)ds$ betrachtet. □

5.10. Konvergenzsätze für Riemann–Stieltjes–Integrale

Der Autor lernte die Beweise aus Vorlesungsnotizen von Herrn Professor L. Heinrich.

Satz 5.75. *Sei* $f_n : [a, b] \to \mathbb{R}$ *eine Folge* stetiger *Funktionen, welche gleichmäßig gegen eine (stetige) Funkion* $f : [a, b] \to \mathbb{R}$ *konvergiert. Sei weiterhin* $g : [a, b] \to \mathbb{R}$ *eine Funktion von beschränkter Variation, dann gilt*

$$\int_a^b f(x)dg(x) = \lim_{n\to\infty}\left(\int_a^b f_n(x)dg(x)\right).$$

Beweis. Mit Theorem 5.71 erhalten wir

$$\left|\int_a^b f_n(x)dg(x) - \int_a^b f(x)dg(x)\right| \le \left|\int_a^b (f_n(x) - f(x))dg(x)\right| \le \|f_n - f\|_{\sup} V_a^b(g)$$

Da (f_n) gleichmäßig gegen f konvergiert, gilt $\lim\limits_{n\to\infty} \|f_n - f\|_{\sup} = 0$. Somit folgt die Aussage. □

Satz 5.76 (Satz von Helly[137]-Bray[138])**.** *Sei* $f : [a, b] \to \mathbb{R}$ *eine* stetige *Funktion, und sei* $g_n : [a, b] \to \mathbb{R}$ *eine Folge von Funktionen von beschränkter Variation, sodass es eine Konstante* $c > 0$ *mit* $V_a^b(g_n) \le c$ *für alle* $n \in \mathbb{N}$ *gibt. Wir nehmen an, dass* g_n *punktweise gegen eine Funktion* $g : [a, b] \to \mathbb{R}$ *konvergiert*[139]*. Dann gilt*

$$\int_a^b f(x)dg(x) = \lim_{n\to\infty}\left(\int_a^b f(x)dg_n(x)\right).$$

Beweis. Sei $C > 0$ so gewählt, dass $\max\limits_{x\in[a,b]}(|f(x)|) < C$ und $C \ge c$ gilt.

Sei $\varepsilon > 0$ und $\tilde{\varepsilon} = \frac{\varepsilon}{4C}$. Da f als stetige Funktion auf einem Kompaktum gleichmäßig stetig ist (Satz von Heine), gibt es ein $\delta > 0$, sodass für alle $x, y \in [a, b]$ mit $|x - y| \le \delta$

[137] Eduard Helly (1884 – 1943), österreichischer Mathematiker
[138] Hubert Evelyn Bray (1889(?) – 1978(?)), amerikanischer Mathematiker
[139] Dann ist auch g von beschränkter Variation mit $V_a^b(g) \le c$ (Übungsaufgabe).

gilt $|f(x) - f(y)| \le \tilde{\varepsilon}$. Wir wählen nun eine Zerlegung $Z = \{x_0, \dots, x_N\}$ von $[a, b]$ der Feinheit $\mu_Z \le \delta$. Mit Satz 5.62 erhalten wir

$$\int\limits_a^b f(x)dg(x) = \sum_{j=1}^N \int\limits_{x_{j-1}}^{x_j} f(x)dg(x) = \sum_{j=1}^N \int\limits_{x_{j-1}}^{x_j} (f(x) - f(x_j))dg(x) + \sum_{j=1}^N \underbrace{\int\limits_{x_{j-1}}^{x_j} f(x_j)dg(x)}_{=f(x_j)(g(x_j)-g(x_{j-1}))}$$

und analog für jedes $n \in \mathbb{N}$:

$$\int\limits_a^b f(x)dg_n(x) = \sum_{j=1}^N \int\limits_{x_{j-1}}^{x_j} (f(x) - f(x_j))dg_n(x) + \sum_{j=1}^N \underbrace{\int\limits_{x_{j-1}}^{x_j} f(x_j)dg_n(x)}_{=f(x_j)(g_n(x_j)-g_n(x_{j-1}))} \quad .$$

Da g_n punktweise gegen g konvergiert, gibt es ein $n_\varepsilon \in \mathbb{N}$, sodass für alle $n \ge n_\varepsilon$ gilt

$$\max\{|g_n(x_k) - g(x_k)| : \ k = 0, 1, \dots, N\} \le \frac{\tilde{\varepsilon}}{N}.$$

Mit Theorem 5.71 und Lemma 5.69 erhalten wir für $n \ge n_\varepsilon$ wegen $V_a^b(g_n - g) \le V_a^b(g_n) + V_a^b(g) \le 2C$:

$$\begin{aligned}
& \left| \int_a^b f(x)dg_n(x) - \int_a^b f(x)dg(x) \right| \\
\le \ & \sum_{j=1}^N \left| \int_{x_{j-1}}^{x_j} (f(x) - f(x_j))d(g_n(x) - g(x)) \right| \\
& + \sum_{j=1}^N |f(x_j)| \cdot |g(x_j) - g(x_{j-1}) - g_n(x_j) + g_n(x_{j-1})| \\
\le \ & \sum_{j=1}^N \underbrace{\|(f - f(x_j))|_{[x_{j-1},x_j]}\|_{\sup}}_{\le \tilde{\varepsilon}} V_{x_{j-1}}^{x_j}(g_n - g) \\
& + \sum_{j=1}^N \underbrace{|f(x_j)|}_{\le C} \cdot (\underbrace{|g(x_j) - g_n(x_j)|}_{\le \frac{\tilde{\varepsilon}}{N}} + \underbrace{|g_n(x_{j-1}) - g(x_{j-1})|}_{\frac{\tilde{\varepsilon}}{N}}) \\
\le \ & \underbrace{\tilde{\varepsilon} \sum_{j=1}^N V_{x_{j-1}}^{x_j}(g_n - g)}_{=V_a^b(g_n - g) \le 2C} + 2\tilde{\varepsilon}C \ = \ 4C\tilde{\varepsilon} \ = \ \varepsilon.
\end{aligned}$$

□

6 Differentialrechnung in mehreren Variablen

6.1. Metrische Räume

Topologie metrischer Räume

Ein zentraler Begriff der Analysis ist „Konvergenz“. Wir möchten diesen Begriff nun in höheren Dimensionen untersuchen. Zur Definition der Konvergenz benötigen wir eine Idee von der „Nachbarschaft“ oder „Umgebung“ eines Punktes. Der Begriff der Umgebung eines Punktes gehört in die mengentheoretische Topologie[140]. Oftmals wird dieser Begriff der Umgebung eines Punktes mit Hilfe einer Abstandsfunktion erklärt. Mengen mit einer Abstandsfunktion werden wir in diesem Abschnitt näher betrachten.

Wir orientieren uns hier an [Fo2, § 1,2], [K2, Kap. 1] und [A1, Abschn. 1.8].

Definition 6.1. Ein *metrischer Raum* (X, d) ist ein Tupel bestehend aus einer Menge X und einer Abbildung $d : X \times X \to \mathbb{R}$, welche *Metrik* oder *Abstandsfunktion* auf X genannt wird und folgende Eigenschaften erfüllt:

(M1) $(d(x, y) = 0) \iff (x = y)$.

(M2) Symmetrie: $\forall x, y \in X : \ d(x, y) = d(y, x)$.

(M3) Dreiecksungleichung: $\forall x, y, z \in X : d(x, y) + d(y, z) \geq d(x, z)$.

Beobachtung 6.2. *Ist (X, d) ein metrischer Raum, so gilt für alle $x, y \in X : \ d(x, y) \geq 0$. In der Tat gilt*

$$0 = d(x, x) \leq d(x, y) + d(y, x) = d(x, y) + d(x, y) = 2d(x, y).$$

Beispiel 6.3. (i) Sei X eine beliebige Menge, dann ist

$$d : X \times X \to \mathbb{R}, \quad (x, y) \mapsto \begin{cases} 0, & \text{wenn} \quad x = y \\ 1, & \text{wenn} \quad x \neq y \end{cases}$$

eine Metrik auf X, genannt die *diskrete Metrik* auf X.

[140] im 19. Jahrhundert oft als „Analysis situs“ bezeichnet

(ii) Sei (X, d) ein metrischer Raum und $Y \subset X$, dann ist $(Y, d|_{Y \times Y})$ ein metrischer Raum. Die Metrik $d|_{Y \times Y}$ auf Y wird die von d *induzierte Metrik* genannt.

Sprechweise 6.4. Sei (X, d) ein metrischer Raum, $m \in X$, $r \in \mathbb{R}$ sowie $Y \subset X$. Dann heißt

- $$B_r(m) = B_r^d(m) = \{x \in X : \ d(x, m) < r\}$$

  offener Ball oder *offene Kugel* um m (oder mit Mittelpunkt m) von Radius r bzgl. d.

- Notation geändert!

  $$B_r^a(m) = B_r^{a,d}(m) = \{x \in X : \ d(x, m) \leq r\}$$

  abgeschlossener Ball oder *abgeschlossene Kugel* um m (oder mit Mittelpunkt m) von Radius r bzgl. d. Für $r < 0$ erhalten wir $B_r(m) = B_r^a(m) = \emptyset$.

- Y eine *Umgebung* von m bzgl. d, wenn gilt:

  $$\exists \varepsilon > 0 : \ B_\varepsilon(m) \subset Y.$$

Definition 6.5. Sei (X, d) ein metrischer Raum.

- Eine Menge $U \subset X$ heißt *offen* in (X, d), falls gilt:

  $$\forall u \in U \ \exists \varepsilon_u > 0 : \ B_{\varepsilon_u}(u) \subset U.$$

  Anders ausgedrückt: U ist eine Umgebung aller $u \in U$. Notation: $U \Subset X$.

- Eine Menge $A \subset X$ heißt *abgeschlossen* in (X, d), falls $X \setminus A$ offen ist. Notation: $A \Subset X$.

- Sei $Y \subset X$. Ein Punkt $x \in X$ heißt *Randpunkt* von Y (bzgl. d), falls gilt

  $$\forall \varepsilon > 0 : \quad \big(B_\varepsilon(x) \cap Y \neq \emptyset \quad \text{und} \quad B_\varepsilon(x) \cap (X \setminus Y) \neq \emptyset\big).$$

  Anders ausgedrückt: Jede Umgebung U von x hat nichtleeren Schnitt sowohl mit Y wie auch mit $X \setminus Y$. Der *Rand* von Y in X (bzgl. d) ist

  $$\partial Y := \{x \in X : \ x \text{ ist Randpunkt von } Y\}.$$

Bemerkung 6.6. Ist (X, d) ein metrischer Raum.

- Dann sind X und $\emptyset$ sowohl offen als auch abgeschlossen. Insbesondere ist die Aussage „$Y \subset X$ ist offen in X“ keineswegs die Negation von „$Y \subset X$ ist abgeschlossen in X“ oder umgekehrt.

- Ist $Y \subset X$, dann gilt $\partial Y = \partial(X \setminus Y)$.

Aufgabe 6.7. Sei (X, d) ein metrischer Raum, $x \in X$ und $\varepsilon > 0$. Zeigen Sie:

(a) $B_\varepsilon(x) = \{y \in X: \ d(x,y) < \varepsilon\}$ ist offen in (X, d).

(b) $B^a_\varepsilon(x) = \{y \in X: \ d(x,y) \leq \varepsilon\}$ ist abgeschlossen in (X, d).

Satz 6.8. *Sei (X, d) ein metrischer Raum und $\mathcal{O}_d \subset \mathcal{P}(X)$ die Menge aller bzgl. d in X offenen Teilmengen von X, dann gilt*[141]*:*

(i) $X, \emptyset \in \mathcal{O}_d$.

(ii) $(U_1, U_2 \in \mathcal{O}_d) \implies (U_1 \cap U_2 \in \mathcal{O}_d)$

(iii) Sei I eine beliebige Indexmenge und U_i, $i \in I$, eine Familie von Elementen aus $\mathcal{O}_d$, dann gilt

$$\bigcup_{i \in I} U_i \in \mathcal{O}_d.$$

Die Teilmenge $\mathcal{O}_d \subset \mathcal{P}(X)$ heißt *Topologie* des metrischen Raumes (X, d).
Der Beweis ist eine Übungsaufgabe.

Beispiel 6.9 (Diskrete Metrik)**.** Sei X eine nichtleere Menge mit mindestens zwei Elementen und $d(x,y) = 1$ für $x \neq y$ die diskrete Metrik auf X. Dann ist der offene Ball $B_1(x) = \{y \in X : \ d(x,y) < 1\} = \{x\}$ und der abgeschlossene Ball ist $B^a_1(x) = \{y \in X : \ d(x,y) \leq 1\} = X$. Die Menge aller in (X, d) offenen Mengen ist $\mathcal{P}(X)$. Damit ist der offene Ball $B_1(x)$ auch abgeschlossen in (X, d). Weiterhin gilt $\partial B_1(x) = \emptyset \neq X \setminus \{x\} = \{y \in X : \ d(x,y) = 1\}$. Somit ist der Abschluss von $B_1(x)$ gerade $B_1(x)$ selbst und nicht etwa $B^a_1(x) = \{y \in X : \ d(x,y) \leq 1\} = X$. Somit gilt im Allgemeinen

$$\partial B_r(x) \neq \{y \in X : \ d(x,y) = r\}.$$

Satz 6.10. *Sei (X, d) ein metrischer Raum und $Y \subset X$, dann gilt:*

(i) $Y \setminus \partial Y \overset{\circ}{\subset} X$.

(ii) $Y \cup \partial Y \overset{\bullet}{\subset} X$.

(iii) $\partial Y \overset{\bullet}{\subset} X$.

Beweis. ad (i): Wenn $Y \setminus \partial Y = \emptyset$, so sind wir fertig. Ansonsten sei $x \in Y \setminus \partial Y$. Da $x \notin \partial Y$, gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \cap (X \setminus Y) = \emptyset$. Desweiteren gilt für dieses ε zudem $B_\varepsilon(x) \cap \partial Y = \emptyset$, sonst wäre ja $B_\varepsilon(x)$ eine Umgebung eines jeden Punktes von $B_\varepsilon(x) \cap \partial Y$ und somit $B_\varepsilon(x) \cap (X \setminus Y) \neq \emptyset$. Damit gilt aber $B_\varepsilon(x) \subset Y \setminus \partial Y$. Da x ein beliebiger Punkt in $Y \setminus \partial Y$ war, folgt die Aussage.

[141] Diese Eigenschaften von $\mathcal{O}_d$ sind die definitorischen Eigenschaften einer sogenannten *Topologie* auf X. Diese Definition wurde 1922 etwa zeitgleich von K. Kuratowski und auch H. Tietze vorgeschlagen.

ad (ii): Nach (i) ist $(X \setminus Y) \setminus \partial(X \setminus Y) = (X \setminus Y) \setminus \partial Y$ offen in X und somit

$$X \setminus ((X \setminus Y) \setminus \partial Y) = (X \setminus (X \setminus Y)) \cup \partial Y = Y \cup \partial Y$$

abgeschlossen in (X, d).

ad (iii): Da $\partial Y = (Y \cup \partial Y) \setminus (Y \setminus \partial Y)$ gilt, und $(Y \cup \partial Y)$ abgeschlossen sowie $Y \setminus \partial Y$ offen in (X, d) ist, ist auch

$$X \setminus \partial Y = (X \setminus (Y \cup \partial Y)) \cup (Y \setminus \partial Y)$$

offen und ∂Y somit abgeschlossen in (X, d).

□

Sprechweise 6.11. Sei (X, d) ein metrischer Raum und $Y \subset X$. Dann heißt

- $Y^\circ := Y \setminus \partial Y$ das *Innere* oder der *offenen Kern* von Y.
- $\overline{Y} := Y \cup \partial Y$ der *Abschluß* oder die *abgeschlossene Hülle* von Y.

Bemerkung 6.12. Wie das Beispiel der diskreten Metrik (Beispiel 6.9) zeigt, ist der Abschluß des offenen Balls von Radius r im Allgemeinen nicht der abgeschlossene Ball von Radius r, d. h.

$$B_r^a(x) \neq \overline{B_r(x)}.$$

Es gilt aber immer

$$\overline{B_r(x)} \subset B_r^a(x).$$

Aufgabe 6.13. Sei (X, d) ein metrischer Raum und $Y \subset X$. Zeigen Sie:

(a) Y ist genau dann offen in (X, d), wenn $Y \cap \partial Y = \emptyset$ gilt.

(b) Y ist genau dann abgeschlossen in (X, d), wenn $\partial Y \subset Y$ gilt.

Satz 6.14 (Metrische Räume sind hausdorffsch)**.** *Ein metrischer Raum (X, d) hat die Hausdorff-Eigenschaft, d. h. sind $x, y \in X$ mit $x \neq y$, dann gibt es offene Teilmengen $U_x, U_y \Subset X$ mit $x \in U_x$, $y \in U_y$ und $U_x \cap U_y = \emptyset$.*

Beweis. Sei $\varepsilon = \frac{d(x,y)}{3}$, dann haben $U_x = B_\varepsilon(x)$ und $U_y = B_\varepsilon(y)$ die gewünschten Eigenschaften. □

Konvergenz in metrischen Räumen

Wir verallgemeinern nun den Begriff der Konvergenz in den komplexen Zahlen mit dem üblichen euklidischen Abstand zur Konvergenz in einem beliebigen metrischen Raum, bzw. später in einem normierten Vektorraum. Viele Aussagen und Argumente übertragen sich direkt auf diesen verallgemeinerten Fall. Daher entscheiden sich manche Dozenten, diesen Stoff ziemlich an den Anfang des Analysis–Zyklus zu stellen, um Verdoppelungen in der Argumentation zu vermeiden. Obwohl dieser Weg effizient, recht elegant und im Prinzip auch nicht wirklich schwieriger ist, haben wir uns gegen diesen entschieden, um allzu viel Abstraktion am Anfang des Studiums zu vermeiden. Hingegen werden wir Argumente nicht immer im Detail ausführen, wenn sie analog zu bereits bekannten Argumenten sind und im Wesentlichen durch eine Änderung der Notation erhalten werden können.

Definition 6.15. Sei (X, d) ein metrischer Raum und $\mathbb{N} \to X, n \mapsto x_n$ eine Folge in X. Die Folge (x_n) heißt *konvergent* in (X, d), wenn gilt

$$\exists x \in X \, \forall \varepsilon > 0 \, \exists N \in \mathbb{N} \, \forall n \geq N : \; d(x_n, x) \leq \varepsilon.$$

Die Zahl x heißt dann *Grenzwert* oder *Limes* der Folge (x_n), notiert $x = \lim\limits_{n \to \infty} x_n$.

Bemerkung 6.16.

- Äquivalent können wir die Konvergenz einer Folge (x_n) in (X, d) auch durch den Begriff der Umgebung definieren: Die Folge (x_n) konvergiert gegen x, wenn es für jede Umgebung U_x von x ein $N \in \mathbb{N}$ gibt, sodass für alle $n \geq N$ gilt $x_n \in U_x$.
- Wenn eine Folge (x_n) in einem metrischen Raum einen Grenzwert besitzt, so ist dieser aufgrund der Hausdorff–Eigenschaft, Satz 6.14, eindeutig.

Satz 6.17 (Folgenkriterium für Abgeschlossenheit)**.** *Sei (X, d) ein metrischer Raum und sei $A \subset X$. Dann sind folgende Aussagen äquivalent:*

(i) A ist abgeschlossen in (X, d).

(ii) Für jede in X konvergente Folge (x_n) mit $x_n \in A$ für alle $n \in \mathbb{N}$ gilt $\lim\limits_{n \to \infty} x_n \in A$.

Beweis.

- Wir zeigen zunächst $(i) \implies (ii)$. Angenommen es gibt eine Folge (x_n) in X mit $x_n \in A$ für alle $n \in \mathbb{N}$, welche in (X, d) gegen $x \in X \setminus A$ konvergiert. Da A abgeschlossen ist, ist $X \setminus A$ offen und somit eine Umgebung von x. Also gibt es ein $N \in \mathbb{N}$, sodass für alle $n \geq N$ gilt $x_n \in X \setminus A$, Widerspruch.
- Wir zeigen nun $(ii) \implies (i)$. Wir behaupten:

$$\forall x \in X \setminus A \quad \exists \varepsilon_x > 0 : \; B_{\varepsilon_x}(x) \subset X \setminus A. \tag{6.1}$$

Damit wäre $X \setminus A$ offen und A abgeschlossen in (X, d). Nehmen wir an, dass die Aussage (6.1) sei falsch. Dann gibt es ein $x \in X \setminus A$, sodass für alle $\varepsilon > 0$ gilt $B_\varepsilon(x) \cap A \neq \emptyset$. Wir wählen nun für jedes $n \in \mathbb{N}$ ein Element $x_n \in B_{\frac{1}{n+1}}(x) \cap A$. Damit ist (x_n) eine Folge von Elementen aus A, welche gegen x konvergiert. Damit gilt nach (ii) auch $x \in A$, Widerspruch.

$\square$

Oft kennt man den Grenzwert einer konvergenten Folge nicht explizit. In vollständigen Räumen hilft da das Cauchy–Kriterium.

Definition 6.18. Eine Folge (x_n) in einem metrischen Raum (X, d) heißt *Cauchyfolge* in (X, d), wenn gilt:

$$\forall \varepsilon > 0 \, \exists N \in \mathbb{N} \, \forall m, n \geq N : \; d(x_m, x_n) \leq \varepsilon.$$

Satz 6.19. *Jede konvergente Folge (x_n) in einem metrischen Raum (X, d) ist eine Cauchyfolge.*

Beweis. Sei $x = \lim\limits_{n\to\infty} x_n$ und $\varepsilon > 0$. Dann gibt es ein $N \in \mathbb{N}$, sodass

$$d(x_n, x) \leq \frac{\varepsilon}{2}$$

für alle $n \geq N$ gilt. Für $m, n \geq N$ folgt dann durch die Dreiecksungleichung

$$d(x_m, x_n) \leq d(x_m, x) + d(x, x_n) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Die Umkehrung, nämlich daß jede Cauchyfolge in einem metrischen Raum konvergiert, ist im Allgemeinen falsch, wie folgendes Beispiel zeigt:

Beispiel 6.20. Wir betrachten den metrischen Raum $X = (0, 2)$ mit der Metrik $d(x, y) = |x - y|$. Die Folge $x_n = \frac{1}{n+1}$ ist eine Cauchyfolge in (X, d), welche jedoch in (X, d) nicht konvergiert.

Definition 6.21. Ein metrischer Raum (X, d) heißt *vollständig*, wenn jede Cauchyfolge in (X, d) auch in (X, d) konvergiert.

Satz 6.22 (verallgemeinertes Intervallschachtelungsprinzip)**.** *Sei (X, d) ein vollständiger metrischer Raum und*

$$A_0 \supset A_1 \supset A_2 \supset \ldots$$

eine Folge ineinander geschachtelter nichtleerer abgeschlossener Teilmengen von X, sodass

$$\lim_{n\to\infty} \operatorname{diam}(A_n) = 0$$

gilt, wobei $\operatorname{diam}(A_n) := \sup\{d(x, y) : \; x, y \in A_n\}$ *den* Durchmesser *von A_n bezeichnet. Dann gibt es genau ein $x \in X$ mit*

$$\{x\} = \bigcap_{n\in\mathbb{N}} A_n.$$

Beweis. Da A_n nicht leer ist, wählen wir für jedes $n \in \mathbb{N}$ ein x_n mit $x_n \in A_n$. Für $m, n \geq N$ folgt aufgrund der Ineinanderschachtelung somit $x_n, x_m \in A_N$ und damit

$$d(x_m, x_n) \leq \operatorname{diam}(A_N).$$

Da $\lim\limits_{n\to\infty} \operatorname{diam}(A_n) = 0$ ist (x_n) eine Cauchyfolge, welche gegen ein Element $x \in X$ konvergiert, weil X vollständig ist. Sei $n \in \mathbb{N}$ beliebig, dann gilt $x_n \in A_k$ für alle $k \leq n$. Aus Satz 6.17 folgt somit $x \in A_k$ für alle $k \in \mathbb{N}$.

Zu guter letzt folgt aus $\lim\limits_{n\to\infty} \operatorname{diam}(A_n) = 0$, dass $\bigcap\limits_{n\in\mathbb{N}} A_n$ höchstens ein Element enthalten kann. □

6.2. Normierte Räume

Wir lehnen uns in diesem Abschnitt an [Fo2, § 1,2], [K2, Kap. 1] und [A1, Abschn. 1.8] an.

In vielen Situationen, die in der Analysis betrachtet werden, ist $X = V$ ein reeller Vektorraum und die Abstandsfunktion d wird durch eine Norm auf V erzeugt:

Definition 6.23. Eine *Norm* auf einem reellen Vektorraum V ist eine Abbildung

$$\|\ldots\| : V \to \mathbb{R}, \quad x \mapsto \|x\|,$$

welche die folgenden Eigenschaften erfüllt:

(N1) $(\|x\| = 0) \iff (x = 0)$.

(N2) $\forall x \in V\, \forall \lambda \in \mathbb{R} : \ \|\lambda x\| = |\lambda| \cdot \|x\|$.

(N3) Dreiecksungleichung $\forall x, y \in V : \|x + y\| \leq \|x\| + \|y\|$.

Das Tupel $(V,\ \|\ldots\|)$ heißt *normierter Vektorraum.*

Beobachtung 6.24. *In einem normierten Vektorraum gilt* $\|x\| \geq 0$, *da*

$$0 = \|0\| = \|x - x\| \leq \|x\| + \|-x\| = \|x\| + |-1| \cdot \|x\| = 2\|x\|.$$

Beispiel 6.25. (i) Die übliche *Euklidische Norm* auf $\mathbb{R}^n$ ist gegeben durch

$$\left\| \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\|_{\text{eukl}} := \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Für $n = 1$ ist diese euklidische Norm einfach der Betrag.

(ii) Die *Maximumsnorm* auf $\mathbb{R}^n$ ist definiert durch

$$\left\| \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\|_{\max} := \max\{|x_1|, \ldots, |x_n|\}.$$

Es gilt hierbei für alle $x \in \mathbb{R}^n$ offenbar:

$$\|x\|_{\max} \leq \|x\|_{\text{eukl}} \leq \sqrt{n}\|x\|_{\max}. \tag{6.2}$$

(iii) Sei X eine beliebige nichtleere Menge und $V = \left\{ f : X \to \mathbb{R} : \sup\limits_{x \in X} |f(x)| < \infty \right\}$, dann ist

$$\|f\|_{\sup} := \sup\{|f(x)| :\ x \in X\}$$

eine Norm auf V (Übungsaufgabe).

(iv) Sei $V = \mathcal{C}([a,b],\mathbb{R})$ der Vektorraum der reellwertigen stetigen Funktionen auf $[a,b]$ und sei $p \geq 1$. Dann ist die p-Norm

$$\|f\|_p := \left(\int_a^b |f(x)|^p dx\right)^{\frac{1}{p}}$$

eine Norm auf $V = \mathcal{C}([a,b],\mathbb{R})$ (siehe Aufgaben 5.20 und 5.32).

(v) Seien $(V, \|\ldots\|_V)$ und $(W, \|\ldots\|_W)$ zwei normierte (reelle) Vektorräume, dann ist auch $\mathrm{Hom}(V,W) := \{f : V \to W : \ f \text{ linear}\}$ ein reeller Vektorraum bzgl. der punktweisen Addition und skalaren Multiplikation. Die Norm

$$\|f\|_{\mathrm{op}} := \sup\left\{\frac{\|f(x)\|_W}{\|x\|_V} : \ x \in V \setminus \{0\}\right\} = \sup\{\|f(x)\|_W : \ x \in V, \ \|x\|_V = 1\}$$

auf $\mathrm{Hom}(V,W)$ heißt *Operatornorm*. Der Nachweis, dass es sich tatsächlich um eine Norm auf $\mathrm{Hom}(V,W)$ handelt, ist eine Übungsaufgabe.

Für $(V, \|\ldots\|_V) = (\mathbb{R}^n, \|\ldots\|_{\mathrm{eukl}})$ und $(W, \|\ldots\|_W) = (\mathbb{R}^m, \|\ldots\|_{\mathrm{eukl}})$ ist der Vektorraum $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ als reeller Vektorraum isomorph zum Vektorraum $\mathbb{R}^{m\times n}$ der reellen $(m \times n)$-Matrizen und wir haben für eine Matrix $A \in \mathbb{R}^{m\times n}$:

$$\|A\|_{\mathrm{op}} := \sup\left\{\frac{\|Ax\|_{\mathrm{eukl}}}{\|x\|_{\mathrm{eukl}}} : \ x \in \mathbb{R}^n \setminus \{0\}\right\} = \sup\{\|Ax\|_{\mathrm{eukl}} : \ x \in \mathbb{R}^n, \ \|x\|_{\mathrm{eukl}} = 1\}.$$

Lemma 6.26. *Sei* $(V, \ \|\ldots\|)$ *ein normierter Vektorraum, dann ist*

$$d_{\|\ldots\|} : V \times V \to \mathbb{R}, \quad (x,y) \mapsto \|x - y\|$$

eine Metrik auf V*. Diese Metrik heißt von der Norm* $\|\ldots\|$ induzierte Metrik *auf* V*.*

Der Beweis ist eine einfache Übungsaufgabe.

Wir fassen nun einen normierten Vektorraum $(V, \ \|\ldots\|)$ insbesondere als metrischen Raum $(V, d_{\|\ldots\|})$ auf. Damit übertragen sich Begriffe wie offene/abgeschlossene Teilmenge, Konvergenz von Folgen, Cauchyfolge oder Vollständigkeit automatisch auf normierte Vekorräume.

Definition 6.27. Ein vollständiger[142] normierter Vektorraum $(V, \|\ldots\|)$ heißt *Banachraum*[143].

Aufgabe 6.28. Sei $(V, \|\ldots\|)$ ein normierter reeller Vektorraum und $d = d_{\|\ldots\|}$ die induzierte Metrik. Sei $v \in V$ und $r > 0$. Zeigen Sie, dass gilt[144]:

(a) $\partial B_r^d(v) := \partial\{x \in V : \ d(x,v) < r\} = \{x \in V : \ d(x,v) = r\}$.

[142] d. h. der metrische Raum $(V, d_{\|\ldots\|})$ ist vollständig

[143] Stefan Banach (1892–1945), polnischer Mathematiker.

[144] Beide Aussagen sind in allgemeinen metrischen Räumen falsch, wie die diskrete Metrik zeigt.

(b) Der Abschluss von $B_r^d(v) := \{x \in V : d(x,v) < r\}$ ist gleich dem abgeschlossenen Ball $B_r^{a,d}(v) := \{x \in V : d(x,v) \le r\}$.

Beobachtung 6.29. *Eine Folge (x_n) in $(\mathbb{R}^m, \|\dots\|_{\text{eukl}})$ mit $x_n = \begin{pmatrix} x_{n1} \\ \vdots \\ x_{nm} \end{pmatrix}$. Dann konvergiert die Folge x_n genau dann gegen $x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, wenn für jedes $k = \{1, \dots, m\}$ die reelle Zahlenfolge x_{nk} gegen x_k konvergiert.*

Konvergenz im $(\mathbb{R}^m, \|\dots\|_{\text{eukl}})$ bedeutet also Konvergenz in jeder Komponente.

Diese Behauptung ergibt sich sofort aus der Abschätzung

$$|x_{nk} - x_k| \le \|x_n - x\|_{\text{eukl}} = \sqrt{\sum_{j=1}^{m} |x_{nj} - x_j|^2} \le \sum_{j=1}^{m} |x_{nj} - x_j|$$

für $j = 1, \dots, m$.

Theorem 6.30 (Satz von Bolzano–Weierstraß)**.** *Für eine Folge (x_n) in $(\mathbb{R}^m, \|\dots\|_{\text{eukl}})$ gilt:*

(i) Ist (x_n) beschränkt, d. h. es gibt ein $C > 0$, sodass $\|x_n\|_{\text{eukl}} \le C$ für alle $n \in \mathbb{N}$, dann hat (x_n) eine konvergente Teilfolge.

(ii) Ist (x_n) eine Cauchyfolge, so ist (x_n) konvergent.
Mit anderen Worten der normierte Vektorraum $(\mathbb{R}^m, \|\dots\|_{\text{eukl}})$ ist vollständig.

Beweis. ad (i): Der Beweis erfolgt durch Induktion über n. Der Induktionsanfang ist Theorem 2.26. Das Argument des Induktionsschlußes ist analog zum Beweis von Theorem 2.27. Die detaillierte Ausführung ist eine Übungsaufgabe.

ad (ii): Ist (x_k) eine Cauchyfolge, so sind wegen $|x_{n_1k} - x_{n_2k}| \le \|x_{n_1} - x_{n_2}\|_{\text{eukl}}$ auch die m Folgen der Komponenten (x_{nk}) für $k = 1, \dots m$ Cauchyfolgen in $\mathbb{R}$. Nach Satz 2.29 sind diese Komponentenfolgen also konvergent. Sei $x_k := \lim\limits_{n\to\infty} x_{nk},\ k = 1, \dots, m$, so konvergiert (x_n) gegen $x = (x_1, \dots, x_m)$.

□

Definition 6.31. Sei V ein reeller Vektorraum. Zwei Normen $\|\dots\|_1$ und $\|\dots\|_2$ auf V heißen *äquivalent*[145] , wenn es $c, C > 0$ gibt mit

$$c\|x\|_2 \le \|x\|_1 \le C\|x\|_2$$

für alle $x \in V$.

[145] Dieser Begriff definiert eine Äquivalenzrelation auf der Menge aller Normen auf V (Übung).

Bemerkung 6.32. Für einen normierten Vektorraum $(V, \ \|\ldots\|)$ bezeichnen wir mit $\mathcal{O}_{\|\ldots\|}$ die Topologie des metrischen Raumes $(V, d_{\|\ldots\|})$. Sind zwei Normen $\|\ldots\|_1$ und $\|\ldots\|_2$ auf V äquivalent, so gilt $\mathcal{O}_{\|\ldots\|_1} = \mathcal{O}_{\|\ldots\|_2}$, auch wenn die Normen als Funktionen selbst verschieden sind. Eine Folge in V konvergiert in diesem Fall genau dann bzgl. der durch die Norm $\|\ldots\|_1$ induzierten Metrik, wenn sie bzgl. der durch die Norm $\|\ldots\|_2$ induzierten Metrik konvergiert. Genauere Ausführungen finden Sie in [K2, Abschn. 1.2].

Theorem 6.33. *Alle Normen auf $\mathbb{R}^n$ sind äquivalent.*

Beweis. Seien $\|\ldots\|_1$ und $\|\ldots\|_2$ zwei Normen auf $\mathbb{R}^n$. Da die Äquivalenz von Normen eine Äquivalenzrelation auf der Menge der Normen auf $\mathbb{R}^n$ ist, genügt es anzunehmen, dass $\|x\|_2$ die euklidische Norm ist. Sei $\{e_1, \ldots, e_n\}$ die Standardbasis[146] des $\mathbb{R}^n$ und $x = \sum\limits_{j=1}^{n} x_j e_j$. Dann gilt nach der Definition einer Norm und der Cauchy–Schwarz–Ungleichung[147] für die euklidische Norm:

$$\|x\|_1 \leq \sum_{j=1}^{n} |x_j| \cdot \|e_j\|_1 \leq \underbrace{\left(\sqrt{\sum_{j=1}^{n} \|e_j\|_1^2}\right)}_{=:C} \|x\|_{\mathrm{eukl}} = C\|x\|_{\mathrm{eukl}}. \tag{6.3}$$

Sei S^{n-1} die $(n-1)$-dimensionale euklidische Einheitssphäre in $\mathbb{R}^n$, d. h.

$$S^{n-1} := \{x \in \mathbb{R}^n : \ \|x\|_{\mathrm{eukl}} = 1\} = \partial B_1^{\mathrm{eukl}}(0),$$

wobei $B_1^{\mathrm{eukl}}(0)$ die offene Kugel um 0 von Radius 1 im $\mathbb{R}^n$ bzgl. der durch die euklidische Norm induzierten Metrik. Insbesondere ist S^{n-1} nach Satz 6.10 abgeschlossen in $(\mathbb{R}^n, \|\ldots\|_{\mathrm{eukl}})$. Wir setzen

$$c := \inf\{\|x\|_1 : \ x \in S^{n-1}\}.$$

Wir zeigen zunächst $c > 0$. Falls $c = 0$, so gibt es eine Folge (x_m) in $\mathbb{R}^n$ mit $x_m \in S$ für alle $m \in \mathbb{N}$ und $\lim\limits_{m\to\infty} \|x_m\|_1 = 0$. Nach dem Satz von Bolzano–Weierstraß (Theorem 6.30) besitzt (x_n) eine in $(\mathbb{R}^n, \|\ldots\|_{\mathrm{eukl}})$ konvergente Teilfolge (x_{m_k}). Nach Satz 6.17 liegt $a := \lim\limits_{k\to\infty} x_{m_k}$ (Grenzwert in $(\mathbb{R}^n, \|\ldots\|_{\mathrm{eukl}})$) ebenfalls in S. Mit Gleichung (6.3) erhalten wir

$$\|a\|_1 \leq \|a - x_{m_k}\|_1 + \|x_{m_k}\|_1 \leq C\|a - x_{m_k}\|_{\mathrm{eukl}} + \|x_{m_k}\|_1.$$

[146] d. h. alle Komponenten von e_j verschwinden außer der j-ten Komponente, welche gleich 1 ist.

[147] Für das euklische Standardskalarprodukt besagt die Cauchy–Schwarz–Ungleichung (siehe z. B. [Fi1, Kap. 5])

$$\langle x, y\rangle^2 \leq \|x\|_{\mathrm{eukl}}^2 \|y\|_{\mathrm{eukl}}^2,$$

oder in Komponenten

$$\left(\sum_{j=1}^{n} x_j y_j\right)^2 \leq \left(\sum_{j=1}^{n} x_j^2\right)\left(\sum_{j=1}^{n} y_j^2\right).$$

Für $k \to \infty$ folgt nun aber $\|a\|_1 = 0$, also $a = 0$ und somit $a \notin S$, Widerspruch. Somit gilt $c > 0$. Für $x \neq 0$ erhalten wir also $c \leq \left\| \frac{x}{\|x\|_{\text{eukl}}} \right\|_1 = \frac{\|x\|_1}{\|x\|_{\text{eukl}}}$ und somit

$$c\|x\|_{\text{eukl}} \leq \|x\|_1.$$

Letztere Ungleichung ist für $x = 0$ trivialerweise erfüllt. □

Korollar 6.34. *Auf einem endlich dimensionalen reellen Vektorraum V sind alle Normen äquivalent.*[148]

Beweis. Sei $n = \dim_{\mathbb{R}}(V)$ und seien $\|\ldots\|_1$ und $\|\ldots\|_2$ zwei Normen auf V. Aus der Linearen Algebra wissen wir, dass es eine Vektorraum–Isomorphismus $\phi : \mathbb{R}^n \to V$ gibt. Wir wählen einen solchen und setzten

$$\|x\|_1^\phi = \|\phi(x)\|_1 \quad \text{und} \quad \|x\|_2^\phi = \|\phi(x)\|_2$$

für alle $x \in \mathbb{R}^n$. Dann sind $\|\ldots\|_1^\phi$ und $\|\ldots\|_2^\phi$ zwei Normen auf $\mathbb{R}^n$ und nach Theorem 6.33 gibt es $C, c > 0$ mit

$$c\|x\|_2^\phi \leq \|x\|_1^\phi \leq C\|x\|_2^\phi.$$

Aufgrund der Surjektivität von ϕ folgt daraus die Behauptung. □

6.3. Stetige Abbildungen

Wir folgen in diesem Abschnitt vor allem [Fo2, §2].

Definition 6.35. Seien (X, d_X) und (Y, d_Y) zwei metrische Räume. Eine Abbildung $f : X \to Y$ heißt *stetig* in $a \in X$, wenn gilt:

$$\forall \varepsilon > 0\ \exists \delta_a > 0\ \forall x \in X:\ d_X(x, a) < \delta_a \implies d_Y(f(x), f(a)) < \varepsilon.$$

Die Abbildung f heißt (auf X) *stetig*, wenn sie in jedem Punkt $a \in X$ stetig ist.

Definition 6.36. Seien (X, d_X) und (Y, d_Y) zwei metrische Räume. Eine Abbildung $f : X \to Y$ heißt *folgenstetig* in $a \in X$, wenn gilt:

$$\lim_{x \to a} f(x) = f(a),$$

d. h. für jede Folge (x_n) in X mit $\lim\limits_{n \to \infty} x_n = a$ gilt $\lim\limits_{n \to \infty} f(x_n) = f(a)$.

Analog zu Satz 3.5 erhalten wir:

Satz 6.37. *Eine Funktion $f : X \to Y$ zwischen zwei metrischen Räumen (X, d_X) und (Y, d_Y) ist genau dann stetig in einem Punkt $a \in X$, wenn sie in a auch folgenstetig ist.*

[148] Diese Aussage gilt *nicht* für unendlich dimensionale reelle Vektorräume (siehe Übungsaufgaben).

Beweis. Nehmen wir an, f ist in a stetig. Sei (x_n) eine Folge in X mit $\lim_{n\to\infty} x_n = a$ und $\varepsilon > 0$ vorgegeben. Aufgrund der Stetigkeit von f in a, gibt es ein $\delta > 0$, sodass für alle $x \in X$ mit $d_X(x,a) < \delta$ gilt $d_Y(f(x), f(a)) < \varepsilon$. Wegen $\lim_{n\to\infty} x_n = a$ gibt es ein $N \in \mathbb{N}$, sodass $d_X(x_n, a) < \delta$ für alle $n \geq N$ gilt. Damit gilt aber für $n \geq N$ auch $d(f(x_n), f(a)) < \varepsilon$. Also $\lim_{n\to\infty} f(x_n) = f(a)$.

Nehmen wir an, f ist in a folgenstetig, aber nicht stetig, dann gibt es ein $\varepsilon > 0$, sodass für alle $\delta > 0$ ein $x \in X$ mit $d_X(x,a) < \delta$ und $d_Y(f(x), f(a)) \geq \varepsilon$ existiert. Sei $n \in \mathbb{N}$, so wählen wir für $\delta = \frac{1}{n}$ ein solches x_n. Dann gilt $\lim_{n\to\infty} x_n = a$ und $d_Y(f(x_n), f(a)) \geq \varepsilon$, im Widerspruch zur Folgenstetigkeit von f in a.

□

Beispiel 6.38. Betrachten wir $(\mathbb{R}^n, \|\ldots\|_{\text{eukl}})$, $n \in \mathbb{N}^*$, und $(\mathbb{R}, |\ldots|)$, so ist für jedes $j \in \{1, \ldots, n\}$ die Projektion $\pi_j : \mathbb{R}^n \to \mathbb{R}$, $(x_1, \ldots, x_n) \to x_j$ auf die j-te Komponente stetig.

Satz 6.39. *Seien (X, d_X), (Y, d_Y) und (Z, d_Z) drei metrische Räume und $f : X \to Y$ sowie $g : Y \to Z$ Abbildungen. Wenn f stetig in $a \in X$ und g stetig in $f(a) \in Y$ ist, dann ist $g \circ f : X \to Z$ stetig in a.*

Beweis. Sei (x_n) eine Folge in X mit $\lim_{n\to\infty} x_n = a \in X$. Dann ist $f(x_n)$ eine Folge in Y, und wegen der Stetigkeit von f in a gilt $\lim_{n\to\infty} f(x_n) = f(a)$. Da g in $f(a)$ stetig ist, folgt

$$\lim_{n\to\infty}((g \circ f)(x_n)) = \lim_{n\to\infty}(g(f(x_n))) = g(f(a)) = (g \circ f)(a).$$

Also ist $g \circ f$ ebenfalls in a stetig. □

Sei $f : X \to \mathbb{R}^n$ eine Abbildung, so schreiben wir oft

$$f = (f_1, f_2, \ldots, f_n).$$

Die Funktionen $f_j : X \to \mathbb{R}$, $j = 1, \ldots, n$ werden auch *Komponentenfunktionen* von f genannt. Mit der Projektion π_j auf die j-te Komponente erhalten wir $f_j = \pi_j \circ f$.

Satz 6.40. *Sei (X, d) ein metrischer Raum. Eine Abbildung $f = (f_1, \ldots, f_n) : X \to \mathbb{R}^n$ ist genau dann (bzgl. der durch $\|\ldots\|_{\text{eukl}}$ induzierten Metrik auf $\mathbb{R}^n$) stetig, wenn für jedes $j \in \{1, \ldots, n\}$ die Komponentenfunktion $f_j : X \to \mathbb{R}$ stetig ist (bzgl. der durch den Betrag definierten Metrik auf $\mathbb{R}$).*

Beweis. Ist f stetig, so auch $f_j = \pi_j \circ f$ für alle $j = 1, \ldots n$. Seien umgekehrt alle $f_j : X \to \mathbb{R}$ stetig und (x_k) eine Folge in X mit $\lim_{k\to\infty} x_k = a \in X$, dann gilt mit Beobachtung 6.29

$$\begin{aligned}
\lim_{k\to\infty} f(x_k) &= \lim_{k\to\infty}(f_1(x_k), \ldots, f_n(x_k)) \\
&= \left(\lim_{k\to\infty} f_1(x_k), \ldots, \lim_{k\to\infty} f_n(x_k)\right) \\
&= (f_1(a), \ldots, f_n(a)) = f(a).
\end{aligned}$$

□

Lemma 6.41. *Folgende drei Abbildungen sind bzgl. der durch die euklidische Norm auf* $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ *induzierten Metrik auf* $\mathbb{R} \times \mathbb{R}$ *bzw.* $\mathbb{R} \times \mathbb{R}^*$ *und dem Betrag auf* $\mathbb{R}$ *stetig:*

$$\begin{aligned} \textit{Addition}: &\quad \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (x,y) \mapsto x + y; \\ \textit{Multiplikation}: &\quad \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (x,y) \mapsto x \cdot y = xy; \\ \textit{Division}: &\quad \mathbb{R} \times \mathbb{R}^* \to \mathbb{R}, \quad (x,y) \mapsto \frac{x}{y}. \end{aligned}$$

Beweis. Sei $\big((x_n, y_n)\big)$ eine Folge in $\mathbb{R} \times \mathbb{R}$, welche gegen (x,y) konvergiert. Mit $\lim_{n\to\infty} x_n = x$ und $\lim_{n\to\infty} y_n = y$ erhalten wir nach den Rechenregeln für konvergente Folgen sofort $\lim_{n\to\infty} (x_n + y_n) = x + y$ sowie $\lim_{n\to\infty} (x_n \cdot y_n) = x \cdot y$.

Falls $y_n \neq 0$ für alle $n \in \mathbb{N}$ und $y \neq 0$, so erhalten wir auch $\lim_{n\to\infty} \frac{x_n}{y_n} = \frac{x}{y}$. $\square$

Korollar 6.42. *Sei* (X,d) *ein metrischer Raum und* $\mathbb{R}$ *versehen mit der üblichen durch den Betrag induzierten Metrik, und seien* $f, g : X \to \mathbb{R}$ *zwei stetige Funktionen, dann sind auch die Abbildungen*

$$f + g : X \to \mathbb{R},\ x \mapsto f(x) + g(x) \quad \textit{und} \quad f \cdot g : X \to \mathbb{R},\ x \mapsto f(x)g(x)$$

stetig. Falls $0 \notin \operatorname{Bild}(g)$, *gilt dieses ebenso für*

$$\frac{f}{g} : X \to \mathbb{R},\ x \mapsto \frac{f(x)}{g(x)}.$$

Satz 6.43 (Stetigkeit linearer Abbildungen zwischen normierten Vektorräumen)**.** *Seien* $(V, \|\ldots\|_V)$ *und* $(W, \|\ldots\|_W)$ *zwei normierte Vektorräume und sei* $f : V \to W$ *eine* lineare *Abbildung. Dann sind folgende Aussagen äquivalent:*

(i) f *ist stetig*

(ii) Es gibt eine Konstante $C > 0$, *sodass für alle* $x \in V$ *gilt* $\|f(x)\|_W \leq C\|x\|_V$. *Anders ausgedrückt:* $\|f\|_{\mathrm{op}} \leq C$.

Beweis. Sei f stetig, insbesondere in 0. Da $f(0) = 0$ gilt, gibt es zu $\varepsilon = 1$ ein $\delta > 0$, sodass für alle $v \in V$ mit $\|v\|_V < \delta$ gilt

$$\|f(v)\|_W < 1.$$

Sei $C := \frac{2}{\delta}$ und $x \in V \setminus \{0\}$, dann gilt für $v := \frac{1}{C\|x\|_V} x$ wegen $\|v\| = \frac{1}{C\|x\|_V}\|x\|_V = \frac{\delta}{2} < \delta$ also

$$1 > \|f(v)\|_W = \left\| f\left(\frac{1}{C\|x\|_V} x\right)\right\|_W = \frac{1}{C\|x\|_V}\|f(x)\|_W.$$

Somit gilt (auch für $x = 0$):

$$\|f(x)\|_W < C\|x\|_V.$$

Angenommen es gibt eine Konstante $C > 0$, sodass $\|f(x)\|_W \leq C\|x\|_V$ für alle $x \in V$ gilt. Sei nun (x_n) eine Folge in V mit $\lim_{n\to\infty} x_n = x$. Dann gilt

$$\|f(x_n) - f(x)\|_W = \|f(x_n - x)\|_W \leq C\|x_n - x\|_v.$$

Daraus folgt $\lim_{n\to\infty} \|f(x_n) - f(x)\|_W = 0$, also $\lim_{n\to\infty} f(x_n) = f(x)$. Damit ist f stetig. $\square$

Korollar 6.44. *Jede lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ ist bzgl. den euklidischen Normen stetig.*

Beweis. Jede lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ ist von der Form $f(x) = A \cdot x$ für eine Matrix $A = (a_{jk}) \in \mathbb{R}^{m\times n}$. Betrachten wir die Operatornorm, so gilt per Definition

$$\|Ax\|_{\mathrm{eukl}} \leq \|A\|_{\mathrm{op}} \cdot \|x\|_{\mathrm{eukl}}$$

für alle $x \neq 0$, und für $x = 0$ ist diese Ungleichung evident. Die Operatornorm $C := \|A\|_{\mathrm{op}}$ ist endlich, da gilt:

$$\|A\|_{\mathrm{op}} \leq \sqrt{nm}\max\{|a_{jk}| : \ j = 1, \dots, m; \ k = 1, \dots, n\} = \sqrt{nm}\max_{j,k} |a_{jk}|.$$

Um diese Ungleichung zu beweisen erinnern wir uns, dass $\|A\|_{\mathrm{op}} = \sup\{\|Ax\|_{\mathrm{eukl}} : \|x\|_{\mathrm{eukl}} = 1\}$ gilt und wählen ein $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ mit $\|x\|_{\mathrm{eukl}} = 1$. Für $y = (y_1, \dots, y_m) = Ax$, erhalten wir dann $y_j = \sum\limits_{k=1}^{n} a_{jk}x_k$. Mit der Ungleichung von Cauchy–Schwarz folgt dann

$$|y_j| = \left|\sum_{k=1}^{n} a_{jk}x_k\right| \leq \sqrt{\sum_{k=1}^{n} a_{jk}^2} \cdot \underbrace{\sqrt{\sum_{l=1}^{n} x_l^2}}_{=1} \leq \sqrt{\sum_{k=1}^{n} (\max_{j,k} |a_{jk}|)^2} = \sqrt{n} \cdot \max_{j,k} |a_{jk}|.$$

Also gilt $\|y\|_{\mathrm{eukl}} \leq \sqrt{nm} \cdot \max\limits_{j,k} |a_{jk}|$. □

Korollar 6.45. *Jede lineare Abbildung zwischen endlich dimensionalen normierten reellen Vektorräumen ist stetig*[149].

Definition 6.46. Sei X eine Menge und (Y, d_Y) ein metrischer Raum. Sei $f_n : X \to Y$, $n \in \mathbb{N}$, eine Folge von Abbildungen. Wir sagen f_n konvergiert *gleichmäßig* gegen eine Funktion $f : X \to Y$, wenn gilt

$$\forall \varepsilon > 0 \, \exists N \in \mathbb{N} \, \forall n \geq N : \ \sup\{d_Y(f_n(x), f(x)) : x \in X\} \leq \varepsilon.$$

Wir verallgemeinern nun Theorem 3.40:

Theorem 6.47. *Seien (X, d_X) und (Y, d_Y) zwei metrische Räume und sei $f_n : X \to Y$, $n \in \mathbb{N}$, eine Folge stetiger Abbildungen, welche gleichmäßig gegen eine Abbildung $f : X \to Y$ konvergiert. Dann ist f stetig.*

Beweis. Der Beweis geht analog zum Beweis von Theorem 3.40. Sei $a \in X$ beliebig, und sei $\varepsilon > 0$. Da die Folge von Abbildungen (f_n) gleichmäßig gegen f konvergiert, gilt:

$$\exists N \in \mathbb{N} \, \forall n \geq N \, \forall x \in X : \ d_Y(f_n(x), f(x)) < \frac{\varepsilon}{3}.$$

[149] Für lineare Abbildungen zwischen unendlich dimensionalen normierten Vektorräumen ist diese Aussage im Allgemeinen falsch (siehe Übungsaufgaben).

Da nun f_N in a stetig ist, gibt es ein $\delta > 0$, sodass für alle $x \in X$ mit $d_X(x, a) < \delta$ gilt

$$d_Y(f_N(x), f_N(a)) \leq \frac{\varepsilon}{3}.$$

Somit erhalten wir für alle $x \in X$ mit $d_X(x, a) < \delta$:

$$d_Y(f(x), f(a)) \leq \underbrace{d_Y(f(x), f_N(x))}_{<\frac{\varepsilon}{3}} + \underbrace{d_Y(f_N(x), f_N(a))}_{\frac{\varepsilon}{3}} + \underbrace{d_Y(f_N(a), f(a))}_{<\frac{\varepsilon}{3}} < \varepsilon.$$

Damit ist f aber stetig in a und, da a beliebig war, überall stetig. $\square$

Bemerkung 6.48. Aus der Definition der gleichmäßigen Konvergenz sieht man sofort, dass gilt: Eine Folge $f_n : X \to (\mathbb{R}, |\ldots|)$ konvergiert genau dann gleichmäßig gegen eine Funktion $f : X \to \mathbb{R}$, wenn gilt

$$\lim_{n\to\infty} \|f_n - f\|_{\sup} = \lim_{n\to\infty} (\sup\{|f_n(x) - f(x)| : \ x \in X\}) = 0.$$

Korollar 6.49. *Der normierte Vektorraum* $\big(\mathcal{C}([a, b], \mathbb{R}),\ \|\ldots\|_{\sup}\big)$ *der stetigen reellwertigen Funktionen auf* $[a, b], a < b$, *versehen mit der Supremumsnorm ist vollständig.*

Beweis. Sei (f_n) eine Cauchyfolge von Funktionen in $\mathcal{C}([a, b], \mathbb{R})$. Für alle $x \in [a, b]$ ist damit $(f_n(x))$ eine reelle Cauchyfolge und somit konvergent. Wir betrachten die Funktion

$$f : [a, b] \to \mathbb{R}, \quad x \mapsto f(x) := \lim_{n\to\infty} f_n(x).$$

Wir müssen nun zeigen dass $\lim\limits_{n\to\infty} \|f_n - f\|_{\sup} = 0$ gilt. Damit konvergiert dann (f_n) gleichmäßig gegen f und f ist somit auch stetig.

Sei also $\varepsilon > 0$. Da (f_n) bzgl. $\|\ldots\|_{\sup}$ eine Cauchyfolge ist, gibt es ein $N \in \mathbb{N}$, sodass für *alle* $n, m \geq N$ gilt

$$\|f_n - f_m\|_{\sup} < \frac{\varepsilon}{2}.$$

Andererseits gibt es für jedes $x \in [a, b]$ auch ein $m_x \geq N$, sodass

$$|f_{m_x}(x) - f(x)| < \frac{\varepsilon}{2}.$$

Damit gilt für alle $x \in [a, b]$ und für alle $n \geq N$:

$$|f_n(x) - f(x)| \leq \underbrace{|f_n(x) - f_{m_x}(x)|}_{<\frac{\varepsilon}{2}} + \underbrace{|f_{m_x}(x) - f(x)|}_{<\frac{\varepsilon}{2}} < \varepsilon$$

und damit

$$\|f_n - f\|_{\sup} := \sup\{|f_n(x) - f(x)| : \ x \in [a, b]]\} \leq \varepsilon.$$

Folglich gilt $\lim\limits_{n\to\infty} \|f_n - f\|_{\sup} = 0$. $\square$

Bemerkung 6.50. Man sieht sofort, dass auch der Vektorraum $\mathcal{C}(X, \mathbb{R})$ aller stetigen reelwertigen Funktionen auf einem metrischen Raum (X, d) bzgl. der Supremumsnorm vollständig ist.

Aufgabe 6.51. Auf $\mathcal{C}^1([0,1], \mathbb{R})$ betrachten wir die Normen

$$\begin{aligned} \|f\|_{\mathcal{C}^1,i} &:= \|f\|_{\sup} + \|f'\|_{\sup}, \qquad \|f\|_{\mathcal{C}^1,ii} := \max\{\|f\|_{\sup},\ \|f'\|_{\sup}\} \\ \text{und} \quad \|f\|_{\mathcal{C}^1,iii} &:= \sup\{|f(x)| + |f'(x)| : x \in [0,1]\} \end{aligned}$$

Zeigen Sie:

(a) Die Normen $\|\ldots\|_{\mathcal{C}^1,i}$, $\|\ldots\|_{\mathcal{C}^1,ii}$ und $\|\ldots\|_{\mathcal{C}^1,iii}$ auf $\mathcal{C}^1([0,1], \mathbb{R})$ sind äquivalent.

(b) Die Normen $\|\ldots\|_{\sup}$ und $\|\ldots\|_{\mathcal{C}^1,i}$ auf $\mathcal{C}^1([0,1], \mathbb{R})$ sind *nicht* äquivalent.

(c) $\big(\mathcal{C}^1([0,1], \mathbb{R}),\ \|\ldots\|_{\mathcal{C}^1,i}\big)$ ist ein vollständiger normierter Vektorraum (Banachraum).

(d) Die lineare Abbildung

$$I : \mathcal{C}([0,1], \mathbb{R}) \to \mathbb{R},\ f \mapsto \int_0^1 f(x)dx$$

ist bzgl. der Norm $\|\ldots\|_{\sup}$ auf $\mathcal{C}([0,1], \mathbb{R})$ und dem Betrag auf $\mathbb{R}$ stetig.

(e) Die lineare Abbildung

$$D : \mathcal{C}^1([0,1], \mathbb{R}) \to \mathcal{C}([0,1], \mathbb{R}),\ f \mapsto f'$$

ist

(i) bzgl. der Norm $\|\ldots\|_{\sup}$ auf $\mathcal{C}^1([0,1], \mathbb{R})$ und auf $\mathcal{C}([0,1], \mathbb{R})$ *nicht* stetig.

(ii) bzgl. der Norm $\|\ldots\|_{\mathcal{C}^1,i}$ auf $\mathcal{C}^1([0,1], \mathbb{R})$ und der Norm $\|\ldots\|_{\sup}$ auf $\mathcal{C}([0,1], \mathbb{R})$ stetig.

6.4. Stetige Abbildungen und Kompaktheit

In diesem Abschnitt lehnen wir uns eng an [Fo2, §3] an.

Wir verallgemeinern nun den Begriff des kompakten Intervalls und der kompakten Menge in $\mathbb{C}$ auf allgemeine metrische Räume.

Definition 6.52. Sei $W \subset X$ eine Teilmenge eines metrischen Raumes[150] (X, d). Eine Familie offener Teilmengen $\{U_i \Subset X :\ i \in I\}$ heißt *offene Überdeckung* von W, wenn gilt

$$W \subseteq \bigcup_{i \in I} U_i.$$

[150] Hier könnten wir sogar einen beliebigen topologischen Raum annehmen.

Bemerkung 6.53. Jede $A \subset X$ Teilmenge eines metrischen Raumes[150] (X, d) besitzt eine offene Überdeckung, da ja X selbst offen ist.

Definition 6.54. Sei (X, d) ein metrischer Raum[150]. Eine Teilmengen $K \subset X$ heißt *kompakt* in (X, d), wenn gilt: *Jede* offene Überdeckung $\{U_i \overset{\circ}{\subset} X : \ i \in I\}$ von K besitzt eine endliche offene Teilüberdeckung, d. h. es gibt eine endliche Teilmenge $J \subset I$, sodass $K \subseteq \bigcup\limits_{j \in J} U_j$ gilt.

Beispiel 6.55. (i) Offene nichtleere Intervalle $(a, b) \subset \mathbb{R}$ mit sind in $(\mathbb{R}, |\ldots|)$ *nicht* kompakt, denn die offene Überdeckung mit $U_n = \left(a + \frac{b-a}{3n}, b - \frac{b-a}{3n}\right)$, $n \in \mathbb{N}^*$, von (a, b) besitzt keine endliche Teilüberdeckung.

(ii) Die Menge $\left\{\frac{1}{n} : \ n \in \mathbb{N}_{\geq 2}\right\}$ ist in $(\mathbb{R}, |\ldots|)$ *nicht* kompakt, denn die offene Überdeckung mit $U_n = \left(\frac{1}{n+1}, \frac{1}{n-1}\right)$, $n \in \mathbb{N}_{\geq 2}$ enthält keine endliche Teilüberdeckung, da der Punkt $\frac{1}{n}$ nur in der Teilmenge U_n der offenen Überdeckung enthalten ist.

(iii) Sei (x_n) eine konvergente Folge in einem metrischen Raum (X, d) und $x := \lim\limits_{n \to \infty} x_n$. Dann ist die Menge
$$K = \{x_n : \ n \in \mathbb{N}\} \cup \{x\}$$
kompakt in (X, d). Sei nämlich $(U_i)_{i \in I}$ eine offene Überdeckung von K. Dann gibt es ein i_0 mit $x \in U_{i_0}$. Da $x := \lim\limits_{n \to \infty} x_n$ liegen alle bis auf endlich viele Folgeglieder, sagen wir $x_{n_1}, \ldots, x_{n_k}$, der Folge (x_n) in U_{i_0}. Für jedes $j \in \{1, \ldots, k\}$ gibt es natürlich ein U_{i_j} mit $x_{n_j} \in U_{i_j}$. Somit gilt $K \subseteq \bigcup\limits_{j=0}^{k} U_{i_j}$.

Satz 6.56. *Eine kompakte Teilmenge K eines metrischen Raumes (X, d) ist beschränkt*[151] *und abgeschlossen.*

Beweis. Sei $x \in X$, dann gilt
$$K \subseteq X = \bigcup_{n \in \mathbb{N}^*} B_n(x).$$
Damit ist $\{B_n(x) : \ n \in \mathbb{N}^*\}$ eine offene Überdeckung von K. Da K kompakt ist, gibt es $n_1, \ldots, n_k \in \mathbb{N}^*$ mit $K \subseteq \bigcup\limits_{j=1}^{k} B_{n_j}(x)$. Für $m := \max\{n_1, \ldots, n_k\}$ folgt also $K \subseteq B_m(x)$. Also ist K beschränkt.

Wir behaupten $X \setminus K$ ist offen. Falls $K = X$ ist die Aussage wahr. Ansonsten gibt es ein $y \in X \setminus K$. Wir betrachten die offenen Mengen
$$U_n := \left\{z \in X : \ d(y, z) > \frac{1}{n}\right\} = X \setminus B^a_{1/n}(y), \quad n \in \mathbb{N}^*.$$

[151] d. h. K liegt in einer offenen Kugel von endlichem Radius

Wegen

$$K \subseteq X \setminus \{y\} = \bigcup_{n \in \mathbb{N}^*} U_n$$

ist $\{U_n : n \in \mathbb{N}^*\}$ eine offene Überdeckung von K. Wegen der Kompaktheit gibt es nun $n_1, \ldots, n_l \in \mathbb{N}^*$ mit

$$K \subseteq \bigcup_{j=1}^{l} U_{n_j}.$$

Für $M := \max\{n_1, \ldots, n_l\}$ gilt somit $B_{1/M}(y) \subseteq X \setminus K$. Somit ist $X \setminus K$ offen und K abgeschlossen. □

Bemerkung 6.57. Die Umkehrung von Satz 6.56 ist falsch. Betrachten wir $\mathbb{R}$ mit der Metrik $d(x,y) = 1$ für alle $x \neq y$. Dann ist jede Einpunktmenge offen, da $\{x\} = B_{\frac{1}{2}}(x)$ und somit jede Teilmenge von $\mathbb{R}$ offen. Nun ist $\mathbb{R}$ abgeschlossen und beschränkt, da $\mathbb{R} = B_2(0)$. Aber $\mathbb{R}$ ist nicht kompakt in $(\mathbb{R}, d)$, da $\mathbb{R} = \bigcup\limits_{x \in \mathbb{R}} \{x\}$ eine offene Überdeckung von $\mathbb{R}$ ist, welche keine endliche Teilüberdeckung von $\mathbb{R}$ enthält.

Korollar 6.58. *Eine konvergente Folge in einem metrischen Raum ist beschränkt.*

Satz 6.59. *Sei K eine kompakte Teilmenge eines metrischen Raumes (X, d) und $A \subset K$ eine abgeschlossene Teilmenge von X, welche ganz in K liegt. Dann ist A ebenfalls kompakt.*

Beweis. Sei $(U_i)_{i \in I}$ eine offene Überdeckung von A, dann ist $\{U_i : i \in I\} \cup (X \setminus A)$ eine offene Überdeckung von X und somit auch von K. Da K kompakt ist, gibt es $i_1, \ldots, i_k \in I$ mit

$$K \subset (X \setminus A) \cup \bigcup_{j=1}^{k} U_{i_j}.$$

Damit gilt $A \subset \bigcup\limits_{j=1}^{k} U_{i_j}$. □

Satz 6.60. *Seien $[a_j, b_j],\ a_j < b_j,\ j = 1, \ldots, n$, kompakte Intervalle in $\mathbb{R}$, dann ist der abgeschlossene Quader im $\mathbb{R}^n$*

$$Q := [a_1, b_1] \times \cdots \times [a_n, b_n] = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : \ a_j \leq x_j \leq b_j,\ j = 1, \ldots, n\}$$

kompakt im $(\mathbb{R}^n, \|\ldots\|_{\text{eukl}})$.

Beweis. Sei $(U_i)_{i \in I}$ eine offene Überdeckung von Q. Angenommen Q kann nicht von endlich vielen Elementen aus $\{U_i : i \in I\}$ überdeckt werden. Dann können wir eine unendliche Folge (Q_m) von abgeschlossenen Quadern in $\mathbb{R}^n$ mit folgenden Eigenschaften konstruieren:

(i) $Q_0 \supset Q_1 \supset \cdots \supset Q_m \supset \ldots$

(ii) Kein Q_m kann durch endlich viele Elemente aus $\{U_i : i \in I\}$ überdeckt werden.

(iii) Für alle $m \in \mathbb{N}$ gilt $\operatorname{diam}(Q_m) = 2^{-m}\operatorname{diam}(Q)$.

Wir konstruieren die Folge (Q_m) rekursiv. Dazu setzen wir $Q_0 := Q$ und nehmen nun an, ein solches $Q_m = I_1 \times \cdots \times I_n$ sei bereits konstruiert. Hierbei sind $I_j \subset \mathbb{R}$, $j = 1, \ldots, n$, kompakte Intervalle in $\mathbb{R}$, die mehr als zwei Punkte enthalten. Wir schreiben jedes dieser Intervalle als Vereinigung zweier Intervalle der halben Länge.

$$I_j = I_j^1 \cup I_j^2$$

und definieren

$$Q_m^{(s_1,\ldots,s_n)} := I_1^{s_1} \times I_2^{s_2} \times \cdots \times I_n^{s_n}, \quad s_j \in \{1,2\},\ j = 1, \ldots, n.$$

Wir erhalten so 2^n abgeschlossene Teilquader von Q_m mit

$$Q_m = \bigcup_{s_1,\ldots,s_n} Q_m^{(s_1,\ldots,s_n)} = Q_m.$$

Da nach Voraussetzung Q_m nicht von endlich vielen Elementen aus $\{U_i : \ i \in I\}$ überdeckt werden kann, gibt es (mindestens) einen Quader $Q_m^{(s_1,\ldots,s_n)}$, welcher ebenso nicht von endlich vielen Elementen aus $\{U_i : i \in I\}$ überdeckt werden kann. Ein solcher sei Q_{m+1}. Nach Konstruktion gilt $\operatorname{diam}(Q_{m+1}) = \frac{1}{2}\operatorname{diam}(Q_m)$ und somit nach Voraussetzung

$$\operatorname{diam}(Q_{m+1}) = \frac{1}{2}\operatorname{diam}(Q_m) = \frac{1}{2}2^{-m}\operatorname{diam}(Q) = 2^{-m-1}\operatorname{diam}(Q).$$

Nach dem verallgemeinerten Intervallschachtelungsprinzip, Satz 6.22, gilt

$$\bigcap_{m\in\mathbb{N}} Q_m = \{x\}.$$

Da $\{U_i : i \in I\}$ eine offene Überdeckung von Q ist, gibt es ein $i_0 \in I$ mit $x \in U_{i_0}$. Als offene Menge gibt es zudem ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset U_{i_0}$. Nun gibt es aber ein $M \in \mathbb{N}$, sodass für alle $m \geq M$ gilt $\operatorname{diam}(Q_m) < \varepsilon$. Da $x \in Q_m$ folgt $Q_m \subset U_{i_0}$ für $m \geq M$. Damit ist U_{i_0} eine offene Überdeckung von Q_m für alle $m \geq M$, im Widerspruch zur Eigenschaft (ii). $\square$

Theorem 6.61 (Satz von Heine–Borel[152]). *Eine Teilmenge $K \subset (\mathbb{R}^n, \|\ldots\|_{\mathrm{eukl}})$ ist genau dann kompakt, wenn sie beschränkt[153] und abgeschlossen ist.*

Beweis. Eine Implikation ist gerade Satz 6.56. Sei nun $K \subset \mathbb{R}^n$ beschränkt und abgeschlossen. Wegen der Beschränktheit von K gibt es einen Quader $Q \subset \mathbb{R}^n$ mit $K \subset Q$. Dieser ist aber nach Satz 6.60 kompakt. Da K abgeschlossen ist, ist auch K nach Satz 6.59 kompakt. $\square$

[152] Émile Borel (1871–1956), französischer Mathematiker
[153] d. h. es gibt $r > 0$ mit $K \subset B_r(0)$.

Korollar 6.62. *Sei $(V, \|\dots\|)$ ein endlich dimensionaler reeller normierter Vektorraum*[154]*. Eine Teilmenge $K \subset V$ ist genau dann kompakt, wenn K abgeschlossen und beschränkt ist.*

Definition 6.63. Eine Teilmenge $A \subset X$ eines metrischen Raumes (X, d) heißt *folgenkompakt*, wenn jede Folge (x_n) in A eine konvergente Teilfolge besitzt, deren Grenzwert in A liegt.

Wir können nun den Satz von Bolzano–Weierstraß (Theorem 6.30) weiter verallgemeinern:

Theorem 6.64 (Satz von Bolzano–Weierstraß)**.** *Jede kompakte Teilmenge eines metrischen Raumes ist folgenkompakt.*

Beweis. Sei K eine kompakte Teilmenge eines metrischen Raumes (X, d) und (x_n) eine Folge in K. Wir müssen zeigen, dass es eine Teilfolge (x_{n_j}) von (x_n) gibt, welche gegen ein Element aus K konvergiert. Nehmen wir an keine Teilfolge von (x_n) konvergiert gegen einen Punkt in K. Dann gibt es für alle $y \in K$ eine offene Umgebung $U_y \Subset X$ mit $y \in U_x$, sodass $x_n \in U_y$ nur für endlich viele $n \in \mathbb{N}$ gilt. Da $(U_y)_{y \in \mathbb{K}}$ eine offene Überdeckung von K ist, gibt es endlich viele $y_1, \dots y_k \in K$ mit

$$K \subseteq \bigcup_{j=1}^{k} U_{y_j}.$$

Damit gilt aber $x_n \in K$ für nur endlich viele $n \in \mathbb{N}$, im Widerspruch zur Annahme, dass (x_n) eine Folge in K ist. □

Bemerkung 6.65. Es gilt auch die Umkehrung von Theorem 6.64 (siehe z. B. [Q, Satz 8.28]). Eine Teilmenge eines metrischen Raumes ist somit genau dann kompakt, wenn sie folgenkompakt ist.

Beispiel 6.66. Wir betrachten den unendlich dimensionalen reellen Vektorraum

$$V := \{a : \mathbb{N} \to \mathbb{R} \mid \exists N_a \in \mathbb{N}\, \forall n \geq N_a : a(n) = 0\}$$

aller reellen Zahlenfolgen, deren Folgenglieder schließlich alle verschwinden. Dann definiert

$$\|a\| := \max\{|a(n)| : \ n \in \mathbb{N}\}, \quad a \in V$$

eine Norm auf V. Wir betrachten nun den abgeschloßenen Ball

$$B_2^a(0).$$

[154]Diese Aussage ist für unendlich dimensionale normierte Vektorräume falsch wie Beispiel 6.66 zeigt. Man kann zeigen, dass ein normierter reeller Vektorraum $(V, \|\dots\|)$ genau dann endlich dimensional ist, wenn der (offenbar beschränkte) abgeschlossene Einheitsball um 0 kompakt ist. Denn ist V nicht endlich dimensional, so kann man rekursiv eine Folge (x_n) linear unabhängiger Einheitsvektoren in V konstruieren, sodass $\|x_j - x_k\| \geq \frac{1}{2}$ für $j \neq k$ gilt. Diese Folge hat dann keine konvergente Teilfolge. Somit ist der abgeschlossene Einheitsball nicht folgenkompakt und damit auch nicht kompakt.

Diese Menge $B_2^a(0)$ ist offenbar beschränkt und abgeschlossen in V. Betrachten wir nun die Folge (a_n) in V mit

$$a_n(j) = \begin{cases} 0; & j \neq n, \\ 1; & j = n. \end{cases}$$

in $B_2^a(0)$. Diese Folge (a_n) hat keine konvergente Teilfolge (Übung). Somit kann die abgeschlossene und beschränkte Menge $B_2^a(0)$ nach dem Satz von Bolzano–Weierstraß (Theorem 6.64) nicht kompakt sein.

Somit erweitert sich das Korollar 6.62 nicht auf unendlich dimensionale reelle normierte Vektorräume.

Wir verallgemeinern nun die Aussage des Korollars 3.25.

Satz 6.67. *Seien (X, d_X) und (Y, d_Y) zwei metrische Räume und $f : X \to Y$ eine stetige Abbildung. Sei $K \subset X$ eine kompakte Teilmenge, dann ist $f(K) \subset Y$ kompakt.*

Beweis. Sei $(U_i)_{i\in I}$ eine offene Überdeckung von $f(K)$. Dann ist $V_i = f^{-1}(U_i)$ für alle $i \in I$ offen[155] in X. Da $(V_i)_{i\in I}$ eine offene Überdeckung von K ist, und weil K kompakt ist, gibt es eine endliche Teilmenge $J \subset I$ mit $K \subset \bigcup_{j\in J} V_j$. Damit gilt $f(K) \subset \bigcup_{j\in J} U_j$. □

Korollar 6.68 (Weierstraßscher Satz vom Extremum)**.** *Sei (X, d) ein metrischer Raum und $K \subset X$ eine nichtleere kompakte Teilmenge. Sei $f : X \to \mathbb{R}$ stetig, dann gibt es ein $M \in K$ und ein $m \in K$, sodass gilt*

$$f(M) = \sup\{f(x) : x \in X\} \quad \textit{und} \quad f(m) = \inf\{f(x) : x \in X\}.$$

Kurz: Eine reellwertige stetige Funktion nimmt auf einem Kompaktum ein Maximum und ein Minimum an.

Der Beweis ist eine Übungsaufgabe (siehe auch [Fo2, §3, Satz 7]).

Definition 6.69. Seien (X, d_X) und (Y, d_Y) zwei metrische Räume. Eine Abbildung $f : X \to Y$ heißt *gleichmäßig stetig*, wenn gilt:

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x, y \in X : \ \big(d_X(x, y) < \delta\big) \implies \big(d_Y(f(x), f(y)) < \varepsilon\big).$$

Als Verallgemeinerung des Theorems 3.33 erhalten wir:

Theorem 6.70 (Satz von Heine)**.** *Seien (X, d_X) und (Y, d_Y) zwei metrische Räume, wobei X kompakt ist, und sei $f : X \to Y$ eine stetige Abbildung. Dann ist f gleichmäßig stetig.*

[155] In der Tat: Sei $i \in I$ und sei $x \in V_i$. Da U_i offen ist, gibt es ein $\varepsilon > 0$, sodass $B_\varepsilon^{d_Y}(f(x)) \subset U_i$ gilt. Da f stetig in x ist, gibt es ein $\delta > 0$ mit $f\big(B_\delta^{d_X}(x)\big) \subset B_\varepsilon^{d_Y}(f(x))$. Somit gilt $B_\delta^{d_X}(x) \subset V_i$.

Allgemein gilt: $f : X \to Y$ ist genau dann (auf ganz X) stetig, wenn das Urbild $f^{-1}(U)$ jeder offenen Menge $U \subset Y$ offen in X ist. So wird dann Stetigkeit für Abbildungen zwischen beliebigen topologischen Räumen definiert.

Beweis. Sei $\varepsilon > 0$. Da f stetig ist, gibt es für jedes $\xi \in X$ ein $r_\xi > 0$ sodass für alle $\zeta \in X$ gilt

$$\big(d_X(\zeta, \xi) < r_\xi\big) \implies \left(d_Y(f(\zeta), f(\xi)) < \frac{\varepsilon}{2}\right).$$

Nun ist $\left(B_{\frac{1}{2}r_\xi}(\xi)\right)_{\xi \in X}$ eine offene Überdeckung von X. Da X kompakt ist, gibt es endlich viele Punkte $\xi_1, \ldots, \xi_k \in X$, sodass

$$X = \bigcup_{j=1}^{k} B_{\frac{1}{2}r_{\xi_j}}(\xi_j).$$

Sei $\delta := \frac{1}{2}\min\{r_{\xi_j} : \quad j = 1, \ldots, k\}$ $x, y \in X$ mit $d_X(x, y) < \delta$. Dann gibt es ein $j \in \{1, \ldots, k\}$ mit $x \in B_{\frac{1}{2}r_{\xi_j}}(\xi_j)$. Dann gilt natürlich $y \in B_{r_{\xi_j}}(\xi_j)$ und wir erhalten

$$d_Y(f(x), f(y)) \leq \underbrace{d_Y(f(x), f(\xi_j))}_{<\frac{\varepsilon}{2}} + \underbrace{d_Y(f(\xi_j), f(y))}_{<\frac{\varepsilon}{2}} < \varepsilon.$$

□

6.5. Kurven

Wir folgen in diesem Abschnitt [Fo2, §4]. Wer mehr über die Geometrie von Kurven im $\mathbb{R}^n$ erfahren möchte, sei z. B. auf [EJ, Kap. 2] verwiesen.

In diesem Abschnitt sei $\mathbb{R}^n$, $n \in \mathbb{N}^*$, immer mit der durch die euklidische Norm $\|\ldots\| = \|\ldots\|_{\text{eukl}}$ induzierten Metrik versehen. Wir bezeichnen mit $I \subset \mathbb{R}$ ein Intervall, welches mindestens zwei Punkte enthält.

Definition 6.71. Eine stetige Abbildung

$$f : I \to \mathbb{R}^n,\ n \in \mathbb{N}^*,$$

heißt *Kurve* im $\mathbb{R}^n$.

Der Begriff einer Kurve entspricht der Anschauung, denn das Bild von f beschriebt im Fall $n = 3$ eine „Kurve“ (im anschaulichen Sinne) im Raum $\mathbb{R}^3$. Wir möchten nun einer Kurve eine Länge zuordnen. Die natürliche Idee ist dabei, die Kurve durch Polygonzüge[156] zu approximieren. Sei $I = [a, b]$, dann wählen wir hierzu eine Zerlegung $Z = \{a = t_0, t_1, \ldots, t_m = b\}$ von I und ersetzen den Graphen von f zwischen allen Punkten $f(t_j)$ und $f(t_{j-1})$ durch das Geradenstück zwischen diesen Punkten. Der so erhaltenen Polygonzug $P_f(Z)$ hat dann die Länge

$$L(P_f(Z)) = \sum_{j=1}^{m} \|f(t_j) - f(t_{j-1})\|.$$

[156] Polygonzüge sind stetige Kurven, die sich stückweise aus Teilstücken von Geraden (eindimensionalen affinen Unterräumen) zusammensetzen.

Definition 6.72. Eine Kurve $f : [a,b] \to \mathbb{R}^n$, $n \in \mathbb{N}^*$, heißt *rektifizierbar*, wenn gilt: Es gibt ein $L \in \mathbb{R}$, sodass es für alle $\varepsilon > 0$ ein $\delta > 0$ gibt, sodass für jede Zerlegung Z von $[a,b]$ der Feinheit $\mu_Z \le \delta$ gilt:

$$|L - L(P_f(Z))| \le \varepsilon.$$

Die Zahl L heißt dann *Länge* oder *Bogenlänge* der Kurve f.

Die Länge einer rektifizierbaren Kurve ist eindeutig bestimmt (Übung).

Beispiel 6.73. Die stetige Kurve

$$f : [0,1] \to \mathbb{R}^2,\ t \mapsto \begin{cases} \big(t, t \cdot \sin(\pi/t)\big), & t \neq 0 \\ (0,0), & t = 0 \end{cases}$$

ist *nicht* rektifizierbar (Übungsaufgabe).

Definition 6.74. Sei $I = (a,b) \subset \mathbb{R}$ ein offenes Intervall, Eine Kurve $f : I \to \mathbb{R}^n$ heißt

- in $t_o \in I$ *differenzierbar* oder *ableitbar*, wenn der Limes

  $$f'(t_o) := \lim_{t \to t_o} \frac{f(t) - f(t_o)}{t - t_o}$$

  existiert. In diesem Fall heißt $f'(t_o) \in \mathbb{R}^n$ *Ableitung* oder *Differential* von f im Punkt t_o[157].

- auf ganz I *differenzierbar* oder kurz *differenzierbar* oder *ableitbar*, wenn f in *jedem* Punkt $t \in I$ differenzierbar ist. In diesem Fall heißt die Funktion

  $$f' : I \to \mathbb{R}^n, \quad t \mapsto f'(t)$$

  Ableitung von f.

Ist I kein offenes Intervall, so heißt eine Kurve $f : I \to \mathbb{R}^n$ *differenzierbar*, wenn es ein offenes Intervall $J \subset \mathbb{R}$ mit $I \subset J$ und eine differenzierbare Kurve $F : J \to \mathbb{R}^n$ mit $F|_I = f$ gibt.

Bemerkung 6.75. Sei $f = (f_1, \ldots, f_n) : I \to \mathbb{R}^n$ eine Kurve, wobei $f_j,\ j = 1, \ldots, n$, die j-te Komponentenfunktion von f ist. Mit der Beobachtung 6.29 erhalten wir: Die Kurve f ist genau dann in einem Punkt t differenzierbar, wenn jede Komponentenfunktion $f_j,\ j = 1, \ldots, n$, in t ableitbar ist. In diesem Fall gilt

$$f'(t) = \left(f_1'(t), \ldots f_n'(t)\right).$$

[157] Diese Definition verallgemeinert die Definition der Differenzierbarkeit von Funktionen einer reellen Variablen mit Werten in $\mathbb{C} \cong \mathbb{R}^2$.

Satz 6.76. *Jede stetig differenzierbare Kurve $f : [a,b] \to \mathbb{R}$ ist rektifizierbar mit Länge*

$$L = \int_a^b \|f'(t)\| dt.$$

Beweis. Sei $f : [a,b] \to \mathbb{R}^n$ stetig differenzierbar.

1. Als Erstes zeigen wir:

$$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall t, \tau \in [a,b] : \; (0 < |t-\tau| \leq \delta) \implies \left(\left\| \frac{f(t) - f(\tau)}{t - \tau} - f'(t) \right\| \leq \varepsilon \right).$$

Sei zunächst $n = 1$, und sei $\varepsilon > 0$. Nach dem Satz von Heine ist $f' : [a,b] \to \mathbb{R}$ gleichmäßig stetig. Damit gibt es ein $\delta > 0$, sodass für alle $t, s \in [a,b]$ gilt:

$$(|t - s| \leq \delta) \implies (|f'(t) - f'(s)| \leq \varepsilon).$$

Für $t, \tau \in [a,b]$ mit $0 < |t - \tau| \leq \delta$ gibt es nach dem Mittelwertsatz ein s zwischen t und τ, sodass

$$\frac{f(t) - f(\tau)}{t - \tau} = f'(s).$$

Damit erhalten wir

$$\left| \frac{f(t) - f(\tau)}{t - \tau} - f'(t) \right| = |f'(s) - f'(t)| \leq \varepsilon.$$

Sei nun $f' = (f_1', \ldots, f_n')$ und $\varepsilon > 0$. Für jedes $j \in \{1, \ldots, n\}$ gibt es nach dem zuvor gezeigten ein $\delta_j > 0$, sodass für alle $t, \tau \in [a,b]$ mit $0 < |t - \tau| \leq \delta_j$ gilt $\left| \frac{f_j(t) - f_j(\tau)}{t - \tau} - f_j'(t) \right| \leq \frac{\varepsilon}{\sqrt{n}}$. Setze $\delta = \min\{\delta_j : \; j = 1, \ldots, n\}$. Dann erhalten wir aufgrund der Abschätzung der euklidischen Norm durch die Maximumsnorm, Gleichung (6.2), für alle $t, \tau \in [a,b]$ mit $0 < |t - \tau| \leq \delta$:

$$\left\| \frac{f(t) - f(\tau)}{t - \tau} - f'(t) \right\| \leq \sqrt{n} \max \left\{ \left| \frac{f_j(t) - f_j(\tau)}{t - \tau} - f_j'(t) \right| : \; j = 1, \ldots, n \right\} \leq \varepsilon.$$

2. Wir beweisen nun die eigentliche Aussage des Satzes. Dazu sei $\varepsilon > 0$. Gemäß Theorem 5.27 gibt es $\delta' > 0$, sodass für jede Zerlegung $Z = \{a = t_0, t_1 \ldots, t_m = b\}$ von $[a,b]$ der Feinheit $\mu_Z \leq \delta'$ gilt

$$\left| \int_a^b \|f'(t)\| dt - \sum_{j=1}^m \|f'(t_j)\|(t_j - t_{j-1}) \right| \leq \frac{\varepsilon}{2}.$$

Nach Teil 1 dieses Beweises gibt es ein δ mit $0 < \delta \leq \delta'$, sodass für jede Zerlegung $Z = \{a = t_0, t_1 \ldots, t_m = b\}$ der Feinheit $\mu_Z \leq \delta$ und alle $j = 1, \ldots, m$ gilt:

$$\left\| \frac{f(t_j) - f(t_{j-1})}{t_j - t_{j-1}} - f'(t_j) \right\| \leq \frac{\varepsilon}{2(b-a)}$$

und mit der Dreiecksungleichung

$$\Big| \|f(t_j) - f(t_{j-1})\| - \|f'(t_j)\|(t_j - t_{j-1}) \Big| \leq \frac{\varepsilon}{2(b-a)}(t_j - t_{j-1}),$$

also insgesamt

$$\left| \sum_{j=1}^{m} \|f(t_j) - f(t_{j-1})\| - \int_a^b \|f'(t)\| dt \right| \leq \varepsilon.$$

□

Aufgabe 6.77. Sei $\phi : [a, b] \to [c, d]$ eine bijektive stetig differenzierbare Abbildung und $f : [c, d] \to \mathbb{R}^n$ eine stetig differenzierbare Kurve. Dann gilt für die Längen L_f und $L_{f\circ\phi}$ der Kurven[158] f und $f \circ \phi$:

$$L_f = L_{f\circ\phi}.$$

Definition 6.78. Eine beschränkte[159] Funktion $f : [a, b] \to \mathbb{R}^n$ auf einem kompakten Intervall $[a, b] \subset \mathbb{R}$ heißt *Riemann–integrierbar* oder kurz *integrierbar*, wenn gilt: Es gibt ein $\iota \in \mathbb{R}^n$, sodass es für alle $\varepsilon > 0$ ein $\delta > 0$ gibt, sodass für jede Zerlegung $Z = \{a = t_0, t_1, \ldots, t_n = b\}$ der Feinheit $\mu_Z \leq \delta$ gilt

$$\|\iota - R(f, Z, \xi_1, \xi_2, \ldots, \xi_n)\| \leq \varepsilon,$$

wobei $R(f, Z, \xi_1, \xi_2, \ldots, \xi_n) := \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1})$ die *Riemannsche Summe* von f bzgl. der Zerlegung Z und den *Stützstellen* $\xi_j \in [x_{j-1}, x_j]$ genannt wird.

Der Vektor $\iota \in \mathbb{R}^n$ heißt dann *Riemann–Integral* von f, notiert

$$\iota = \int_a^b f(t) dt.$$

Die Überlegung in 6.29 zeigt uns, dass gilt:

Beobachtung 6.79. *Eine Funktion $f = (f_1, \ldots, f_n) : [a, b] \to \mathbb{R}^n$ ist genau dann Riemann–integrierbar, wenn jede Komponentenfunktion $f_j, j = 1, \ldots, n$, Riemann–integrierbar ist. In diesem Fall gilt*

$$\int_a^b f(t) dt = \left(\int_a^b f_1(t) dt, \ldots, \int_a^b f_n(t) dt \right).$$

Insbesondere sind stetige Funktionen $f : [a, b] \to \mathbb{R}^n$ stets Riemann–integrierbar.

Bis zum Ende dieses Abschnitts orientieren wir uns an [EJ, Abschn. 2.1].

[158] Man nennt $f \circ \phi$ eine *Umparametrisierung* von f.
[159] d. h. es gibt $r > 0$, sodass $f([a, b]) \subset B_r(0)$

Satz 6.80. *Sei $f : [a,b] \to \mathbb{R}^n$ stetig, dann gilt (für die euklidische Norm $\| \dots \|$)*

$$\left\| \int_a^b f(t)dt \right\| \leq \int_a^b \|f(t)\|dt.$$

Hierbei gilt Gleichheit, falls alle $f(t)$ gleichgerichtet[160] *sind.*

Beweis. Sei $w = \int_a^b f(t)dt \in \mathbb{R}^n$. Wenn $w = 0$ ist, so gilt obige Aussage und die Gleichheit tritt wegen der Stetigkeit von f genau dann ein, wenn f die Nullabbildung ist.

Sei nun $w \neq 0$, dann gilt mit dem euklidischen Standardskalarprodukt $\langle \cdot, \cdot \rangle$:

$$\|w\|^2 = \langle w, w \rangle = \left\langle w, \int_a^b f(t)dt \right\rangle = \int_a^b \langle w, f(t) \rangle dt.$$

Nach der Ungleichung von Cauchy-Schwarz gilt nun $\langle w, f(t) \rangle \leq \|w\| \cdot \|f(t)\|$ für alle $t \in [a,b]$, wobei Gleichheit genau dann eintritt, wenn $f(t)$ und w immer gleichgerichtet sind.

Aufgrund der Monotonie des Integrals erhalten wir

$$\|w\|^2 \leq \int_a^b \|w\| \cdot \|f(t)\|dt = \|w\| \int_a^b \|f(t)\|dt$$

und somit, wegen $w \neq 0$:

$$\|w\| \leq \int_a^b \|f(t)\|dt$$

Hierbei gilt Gleichheit genau dann, wenn w und $f(t)$ immer gleichgerichtet sind. □

Definition 6.81. Eine Kurve $f : [a,b] \to \mathbb{R}^n$ heißt *regulär*, wenn sie stetig differenzierbar ist und die Ableitung f' keine Nullstelle hat.

Korollar 6.82. *Seien $x, y \in \mathbb{R}^n$ zwei verschiedene Punkte und $f : [a,b] \to \mathbb{R}^n$ eine reguläre Kurve mit $f(a) = x$ und $f(b) = y$. Dann gilt für die Länge L_f von f :*

$$L_f \geq \|x - y\|.$$

Wenn $L_f = \|x - y\|$, dann gibt es eine stetig differenzierbare bijektive Abbildung $\phi : [a,b] \to [0,1]$, sodass

$$f = c_{xy} \circ \phi,$$

wobei $c_{xy} : [0,1] \to \mathbb{R}^n,\ t \mapsto x + t(y - x)$ die Strecke von x nach y ist.

Kurz: Bis auf Umparametrisierung ist nur die Strecke die kürzeste Verbindung zweier Punkte im $\mathbb{R}^n$.

[160] d. h. es gibt ein $z \in \mathbb{R}^n \setminus \{0\}$, sodass $f(t) \in \mathbb{R}_{\geq 0} z$ für alle $t \in [a,b]$ gilt

Beweis. Es gilt $L_f = \int_a^b \|f'(t)\|dt \geq \left\|\int_a^b f'(t)dt\right\| = \|f(b) - f(a)\| = \|x - y\|$. Die Gleichheit gilt genau dann, wenn alle $f'(t)$ zu $\int_a^b f'(t)dt = y - x$ gleichgerichtet sind, d. h. $f'(t) = \mu(t)(y - x)$ für eine stetige strikt positive Funktion $\mu : [a, b] \to \mathbb{R}$. Nach dem Hauptsatz der Differential- und Integralrechnung erhalten wir nun

$$f(t) = x + \int_a^t f'(s)ds = x + \left(\int_a^t \mu(s)ds\right)(y - x).$$

Die Behauptung folgt indem wir $\phi(t) := \int_a^t \mu(s)ds$ setzen. □

6.6. Partielle Ableitungen

Dieser Abschnitt ist an [Fo2, §5,6] angelehnt.

Wir bezeichnen mit $\{e_1, \ldots, e_n\}$ die Standardbasis des $\mathbb{R}^n$, d. h. e_j ist der Vektor in $\mathbb{R}^n$ dessen Komponenten alle verschwinden, bis auf die j-te Komponente, welche eins ist. Weiterhin nehmen wir an, dass $\mathbb{R}^n$ mit der von der eukidischen Norm $\|\ldots\| = \|\ldots\|_{\text{eukl}}$ induzierten Metrik versehen sei.

Definition 6.83. Sei $U \Subset \mathbb{R}^n$ und $v \in \mathbb{R}^n$ mit $\|v\| = 1$. Eine Funktion $f : U \to \mathbb{R}^m$, $U \Subset \mathbb{R}^n$, heißt in einem Punkt u *in Richtung* v *differenzierbar (ableitbar)*, wenn der Grenzwert

$$D_v f(u) := \lim_{h \to 0} \frac{f(u + hv) - f(u)}{h}, \quad h \in \mathbb{R} \setminus \{0\}$$

existiert[161]. In diesem Fall heißt $D_v f(u)$ die *Richtungsableitung* von f im Punkt $u \in U$ in Richtung $v \in \mathbb{R}^n$.

Bemerkung 6.84. Sei $f : U \to \mathbb{R}^m$, $U \Subset \mathbb{R}^n$, in $u \in U$ in Richtung $v \in \mathbb{R}^n$, $\|v\| = 1$, ableitbar. Da $U \subset \mathbb{R}^n$ offen ist, gibt es ein $\varepsilon > 0$, sodass für alle $t \in (-\varepsilon, \varepsilon)$ der Vektor $u + tv$ in U liegt. Betrachten wir das Geradenstück

$$\gamma_v : (-\varepsilon, \varepsilon) \to U, \quad t \mapsto u + tv,$$

dann ist die Abbildung $f \circ \gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^m$ in $t = 0$ differenzierbar mit

$$(f \circ \gamma)'(0) = D_v f(u).$$

Definition 6.85. Sei $U \Subset \mathbb{R}^n$ eine offene Teilmenge des $\mathbb{R}^n$ und $j \in \{1, \ldots, n\}$. Eine Abbildung $f : U \to \mathbb{R}^m$ heißt

- im Punkt $u \in U$ *bzgl. der j-ten Koordinatenrichtung partiell ableitbar (partiell differenzierbar)*, falls die Richtungsableitung

$$D_j f(u) = \frac{\partial f}{\partial x_j}(u) := D_{e_j} f \in \mathbb{R}^m$$

[161] Hierbei betrachten wir nur die $h \in \mathbb{R} \setminus \{0\}$ mit $u + hv \in U$.

existiert[162]. In diesem Fall heißt $D_j f(u) = \frac{\partial f}{\partial x_j}(u) \in \mathbb{R}^m$ die *j-te partielle Ableitung* von f in u.

- (auf U) *bzgl. der j-ten Koordinatenrichtung partiell ableitbar (partiell differenzierbar)*, wenn f in jedem Punkt $u \in U$ bzgl. der j-ten Koordinatenrichtung partiell ableitbar ist.
- im Punkt $u \in U$ *partiell ableitbar (partiell differenzierbar)*, wenn f für alle $j \in \{1, \ldots, n\}$ im Punkt $u \in U$ bzgl. der j-ten Koordinatenrichtung partiell ableitbar ist.
- (auf U) *partiell ableitbar (partiell differenzierbar)*, wenn f in jedem Punkt $u \in U$ partiell ableitbar ist.

Bemerkung 6.86. Sei $u = (u_1, \ldots, u_n) \in U \Subset \mathbb{R}^n$. Dann gibt es ein $\varepsilon > 0$, sodass für alle $t \in (u_j - \varepsilon, u_j + \varepsilon)$ gilt $(u_1, \ldots, u_{j-1}, t, u_{j+1}, \ldots, u_n) \in U$. Die Funktion f ist genau dann im Punkt u bzgl. der j-ten Koordinatenrichtung partiell ableitbar, wenn die Funktion

$$F_{j,u} : (u_j - \varepsilon, u_j + \varepsilon) \to \mathbb{R}^m, \quad x \mapsto f(u_1, \ldots, u_{j-1}, t, u_{j+1}, \ldots, u_n)$$

in $t = u_j$ differenzierbar ist. Es gilt dann

$$D_j f(u) = \frac{\partial f}{\partial x_j}(u) = F'_{j,u}(u_j).$$

Somit ist die j-partielle Ableitung in u auf eine gewöhnlich Ableitung nach einer Variablen zurückgeführt. Aus diesem Grund gelten für partielle Ableitungen analoge Rechenregeln wir für gewöhnliche Ableitungen von Funktionen einer reellen Veränderlichen.

Beispiel 6.87. Die Funktion

$$r : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}, \quad x = (x_1, \ldots, x_n) \mapsto \|x\|_{\text{eukl}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

ist auf ganz $\mathbb{R}^n \setminus \{0\}$ partiell differenzierbar mit

$$\frac{\partial r}{\partial x_j}(x_1, \ldots, x_n) = \frac{x_j}{r(x_1, \ldots, x_n)}.$$

Vorsicht. Ist $U \Subset \mathbb{R}^n$ und $n \geq 2$, so muß eine Funktion $f : U \to \mathbb{R}$, welche in $u \in U$ partiell differenzierbar ist, in u *keineswegs* stetig sein, wie folgendes Beispiel zeigt:

Die Funktion

$$f : \mathbb{R}^n \to \mathbb{R}, \quad (x_1, \ldots, x_n) \mapsto \begin{cases} \frac{x_1 \cdot x_2 \cdots x_n}{(x_1^2 + x_2^2 + \cdots + x_n^2)^n}, & \text{falls } (x_1, \ldots, x_n) \neq 0 \\ 0, & \text{falls } (x_1, \ldots, x_n) = 0 \end{cases} \quad n \geq 2, \tag{6.4}$$

[162] d. h. der Grenzwert

$$D_j f(u) = \frac{\partial f}{\partial x_j}(u) := \lim_{h \to 0} \frac{f(u + he_j) - f(u)}{h}$$

existiert. Wobei bei diesem Limes natürlich nur $h \in \mathbb{R}$ mit $h \neq 0$ und $u + he_j \in U$ betrachtet werden.

ist in $x = 0$ partiell differenzierbar, da für jedes $j = 1, \dots, n$ und alle $h \in \mathbb{R} \setminus \{0\}$ mit $f(0 + he_j) = 0$ auch

$$\frac{f(0 + he_j) - f(0)}{h} = 0$$

gilt. Somit ist $\frac{\partial f}{\partial x_j}(0) = 0$.

Hingegen ist f in 0 nicht stetig, da für die Folge $a_j = \left(\frac{1}{j}, \dots, \frac{1}{j}\right)$ gilt

$$f(a_j) = \left(\frac{j}{n}\right)^n$$

und somit $\lim\limits_{j \to \infty} f(a_j) = \infty$ folgt.

Definition 6.88. Ist $f : U \to \mathbb{R}^m$, $U \Subset \mathbb{R}^n$, partiell differenzierbar, so erhalten wir eine Abbildung

$$D_j f = \frac{\partial f}{\partial x_j} : U \to \mathbb{R}^m, \quad x \mapsto D_j f(x) = \frac{\partial f}{\partial x_j}(x).$$

Wir sagen f ist *zweimal partiell differenzierbar*, wenn alle Abbildungen $D_j f$, $j = 1, \dots, n$, wieder partiell differenzierbar sind, also für alle $j, k \in \{1, \dots, n\}$ die Abbildungen

$$D_k D_j f := D_k(D_j f) : U \to \mathbb{R}^m$$

existieren. Anstelle von $D_k D_j f(u)$ schreibt man auch

$$\frac{\partial^2 f}{\partial x_k \partial x_j}(u).$$

Rekursiv definieren wir nun: Eine l-mal partiell differenzierbare $f : U \to \mathbb{R}^m$ heißt $(l+1)$*-mal partiell differenzierbar*, wenn *jede* l-te partielle Ableitung

$$D_{j_l} D_{j_{l-1}} \dots D_{j_2} D_{j_1} f, \quad j_1, \dots, j_l \in \{1, \dots, n\},$$

wiederum partiell differenzierbar ist.

Anstelle von $D_{j_l} D_{j_{l-1}} \dots D_{j_2} D_{j_1} f$ schreibt man auch

$$\frac{\partial^l f}{\partial x_{j_l} \partial x_{j_{l-1}} \dots \partial x_{j_1}}.$$

Theorem 6.89 (Satz von Schwarz[163] / Satz von Clairaut[164])**.** *Sei $U \Subset \mathbb{R}^n$, $u \in U$ und $f : U \to \mathbb{R}^m$ sowie $j, k \in \{1, \dots, n\}$. Wenn die ersten partiellen Ableitungen $D_j f$, $D_k f$ und die zweite partiellen Ableitungen $D_j D_k f$ sowie $D_k D_j f$ auf U existieren und wenn zudem die zweiten partiellen Ableitungen $D_j D_k f$ und $D_k D_j f$ im Punkt $u \in U$ stetig sind, dann gilt*[165]

$$D_j D_k f(u) = D_k D_j f(u).$$

[163] Hermann Amandus Schwarz (1843 – 1921), deutscher Mathematiker
[164] Alexis-Claude Clairaut (1713 – 1765), französischer Gelehrter
[165] Eine etwas verschärfte Formulierung dieses Satzes finden Sie in [K2, Abschn. 2.4].

Beweis. Ohne Beschränkung der Allgemeinheit dürfen wir $n = 2,\ m = 1,\ j = 1$ und $k = 2$ sowie $u = (0,0) \in U$ annehmen. Um Indizes zu vermeiden schreiben wir für die Komponenten der Variablen $x = x_1$ und $y = x_2$. Da $U \subset \mathbb{R}^2$ offen ist und $0 \in U$ gilt, gibt es ein $\delta > 0$, sodass

$$\{(x,y) \in \mathbb{R}^2 :\ |x| < 2\delta,\ |y| < 2\delta\} \subset U$$

gilt.

(i) Für $|y| < \delta$, betrachten wir die differenzierbare Funktion

$$F_y : (-\delta,\delta) \to \mathbb{R}, \quad x \mapsto f(x,y) - f(x,0).$$

Nach dem Mittelwertsatz der Differentialrechnung (Theorem 4.19) gibt es ein ξ_{xy} zwischen 0 und x, sodass

$$F_y(x) - F_y(0) = F_y'(\xi_{xy})x = \big(D_1 f(\xi_{xy},y) - D_1 f(\xi_{xy},0)\big)x$$

gilt. Wenden wir nun den Mittelwertsatz auf die differenzierbare Funktion

$$(-\delta,\delta) \to \mathbb{R}, \quad y \mapsto D_1 f(\xi_{xy},y)$$

an, so gibt es ein η_{xy} zwischen 0 und y, sodass

$$F_y'(\xi_{xy}) = D_1 f(\xi_{xy},y) - D_1 f(\xi_{xy},0) = D_2 D_1(\xi_{xy},\eta_{xy})y.$$

Zusammenfassend erhalten wir:

$$\begin{aligned} f(x,y) - f(x,0) - f(0,y) + f(0,0) &= F_y(x) - F_y(0) \\ &= F_y'(\xi_{xy})x \\ &= D_2 D_1(\xi_{xy},\eta_{xy})xy. \end{aligned} \tag{6.5}$$

(ii) Wir führen nun die Überlegungen aus Teil (i) noch einmal mit vertauschten Rollen von x und y durch: Für $|x| < \delta$, betrachten wir also die differenzierbare Funktion

$$\tilde{F}_x : (-\delta,\delta) \to \mathbb{R}, \quad y \mapsto f(x,y) - f(0,y).$$

Nach dem Mittelwertsatz der Differentialrechnung gibt es nun ein $\tilde{\eta}_{xy}$ zwischen 0 und y, sodass

$$\tilde{F}_x(y) - \tilde{F}_x(0) = \tilde{F}_x'(\tilde{\eta}_{xy})y = \big(D_2 f(x,\tilde{\eta}_{xy}) - D_2 f(0,\tilde{\eta}_{xy})\big)y.$$

Durch Anwendung des Mittelwertsatzes auf die differenzierbare Funktion

$$(-\delta,\delta) \to \mathbb{R}, \quad x \mapsto D_2 f(x,\tilde{\eta}_{xy})$$

finden wir ein $\tilde{\xi}_{xy}$ zwischen 0 und x, sodass

$$\tilde{F}_x'(\tilde{\eta}_{xy}) = D_2 f(x,\tilde{\eta}_{xy}) - D_2 f(0,\tilde{\eta}_{xy}) = D_1 D_2(\tilde{\xi}_{xy},\tilde{\eta}_{xy})x.$$

Insgesamt folgt

$$\begin{aligned} f(x,y) - f(x,0) - f(0,y) + f(0,0) &= \tilde{F}_x(y) - \tilde{F}_x(0) \\ &= \tilde{F}_x'(\tilde{\eta}_{xy})x \\ &= D_1 D_2(\tilde{\xi}_{xy},\tilde{\eta}_{xy})xy. \end{aligned} \tag{6.6}$$

Wenn $xy \neq 0$, so erhalten wir aus den Gleichungen (6.5) und (6.6) sofort

$$D_2D_1(\xi_{xy}, \eta_{xy}) = D_1D_2(\tilde{\xi}_{xy}, \tilde{\eta}_{xy}).$$

Da ξ_{xy} und $\tilde{\xi}_{xy}$ zwischen 0 und x liegen und η_{xy} sowie $\tilde{\eta}_{xy}$ zwischen 0 und y liegen, folgt für $xy \neq 0$ und $x, y \to 0$ aufgrund der vorausgesetzten Stetigkeit der zweiten partiellen Ableitungen von f in $(0,0)$ die zu zeigende Behauptung

$$D_1D_2(0,0) = D_2D_1(0,0).$$

□

Die Stetigkeitsvoraussetzung an mindestens eine der betrachteten zweiten partiellen Ableitungen von f in Theorem 6.89 ist notwendig, wie folgendes Beispiel zeigt:

Aufgabe 6.90.

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2}, & \text{falls} \quad (x,y) \neq (0,0) \\ 0, & \text{falls} \quad (x,y) = (0,0) \end{cases}$$

(a) Zeigen Sie: f ist zweimal partiell differenzierbar.

(b) Zeigen Sie: $\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0)$.

(c) Welche Voraussetzung von Theorem 6.89 ist nicht erfüllt?

Durch vollständige Induktion folgt aus Theorem 6.89 sofort:

Korollar 6.91 (Satz von Schwarz / Satz von Clairaut)**.** *Sei $U \Subset \mathbb{R}^n$, und sei $f : U \to \mathbb{R}^m$ eine k-mal differenzierbare Funktion, sodass alle l-ten partiellen Ableitungen mit $l \leq k$ auf U stetig sind. Sei $\pi : \{1, \ldots, k\} \to \{1, \ldots, k\}$ eine Permutation[166] von $\{1, \ldots, k\}$, dann gilt auf U für alle $j_1, \ldots, j_k \in \{1, \ldots, n\}$:*

$$D_{j_k} D_{j_{k-1}} \ldots D_{j_2} D_{j_1} f = D_{j_{\pi(k)}} D_{j_{\pi(k-1)}} \ldots D_{j_{\pi(2)}} D_{j_{\pi(1)}} f.$$

6.7. Die totale Ableitung

In diesem Abschnitt möchten wir die Ableitung einer Funktion $f : \mathbb{R}^n \to \mathbb{R}^m$ definieren. Da wir nicht durch Vektoren dividieren können, ist eine naive Definition über den Differenzenquotienten nicht unmittelbar möglich. Die Grundidee der lokalen Approximation der Funktion f durch eine (affin) lineare Abbildung (siehe Satz 4.6, Teil (iii)) verallgemeinert sich aber in natürlicher Weise. Wir folgen in diesem Kapitel wieder dem Buch Analysis 2 von O. Forster, siehe [Fo2, §6], und auch [Esch, § 7,8].

In diesem Abschnitt betrachten wir $\mathbb{R}^n$ mit der durch die euklidische Norm induzierten Metrik.

[166] d. h. π ist eine bijektive Abbildung von {1,. . . ,k} in sich selbst.

Definition 6.92. Sei $U \Subset \mathbb{R}^n$ eine offene Teilmenge. Eine Abbildung $f : U \to \mathbb{R}^m$ heißt in einem Punkt $u \in U$ *(total) differenzierbar* oder *(total) ableitbar*, wenn gilt: Es gibt eine $\mathbb{R}$-lineare Abbildung $A_u : \mathbb{R}^n \to \mathbb{R}^m$ und eine Abbildung $\phi_u : B_{r_u}(0) \to \mathbb{R}^m$ für ein hinreichend kleines $r_u > 0$, sodass einerseits

$$\lim_{\eta \to 0} \frac{\phi_u(\eta)}{\|\eta\|} = 0$$

und für alle $\xi \in B_{r_u}(0)$ gilt $u + \xi \in U$ sowie

$$f(u+\xi) = f(u) + A_u(\xi) + \phi_u(\xi). \tag{6.7}$$

Die $\mathbb{R}$-lineare Abbildung $A_u : \mathbb{R}^n \to \mathbb{R}^m$ heißt in diesem Fall das *totale Differential* von f in u. Wir schreiben auch

$$A_u = Df(u).$$

Ist eine Abbildung $f : U \to \mathbb{R}^m$ in jedem Punkt $u \in U$ total differenzierbar, dann nennen wir f kurz *(total) differenzierbar* oder *(total) ableitbar*.

Bemerkung 6.93. Nach Satz 4.6 stimmt diese Definition für $n = m = 1$ mit der üblichen Definition überein. Gleiches gilt auch für $n = 1$ und m beliebig.

Beobachtung 6.94. *Man sieht leicht (Übungsaufgabe), dass die Menge $X(U,u,m)$ aller in $u \in U \Subset \mathbb{R}^n$ total differenzierbaren Funktionen $f : U \to \mathbb{R}^m$ ein reeller Vektorraum ist und für $f_1, f_2 \in X(U,u,m)$ und $\lambda \in \mathbb{R}$ gilt*

$$D(\lambda f_1 + f_2)(u) = \lambda Df_1(u) + Df_2(u).$$

Lemma 6.95. *Ist $f : U \to \mathbb{R}^m$ in einem Punkt $u \in U$ total differenzierbar, dann ist f in diesem Punkt u stetig.*

Beweis. Sei (u_j) eine Folge in U mit $\lim\limits_{j\to\infty} u_j = u$. Wir können ohne Einschränkung $u_j - u \in B_{r_u}(0)$ für alle $j \in \mathbb{N}$ annehmen. Dann gilt nach der Definition der totalen Differenzierbarkeit

$$\lim_{j\to\infty} f(u_j) = \lim_{j\to\infty} f(u + (u_j - u)) = \lim_{j\to\infty} (f(u) + A(u_j - u) + \phi_u(u_j - u)) = f(u),$$

da eine lineare Abbildung zwischen endlich dimensionalen Vektorräumen stetig ist. □

Bemerkung 6.96. Da es in einem Punkt $u \in U$ unstetige aber in u partiell differenzierbare Abbildungen $f : U \to \mathbb{R}^m$ gibt (z. B. die in (6.4) definierte Funktion), sind in u partiell differenzierbare Abbildungen nicht unbedingt in u auch total differenzierbar.

$$(f \text{ partiell differenzierbar}) \not\Rightarrow (f \text{ total differenzierbar})$$

Selbst wenn in einem Punkt $u \in U$ *alle* Richtungsableitungen von f existieren, so muss f in diesem Punkt u unbedingt total differenzierbar sein, wie folgendes Beispiel zeigt:

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto \begin{cases} \frac{x^2 y}{x^4+y^2}, & \text{wenn} \quad (x,y) \neq (0,0) \\ 0, & \text{wenn} \quad (x,y) = (0,0) \end{cases}$$

(a) Zeigen Sie: In $(0,0)$ existieren alle Richtungsableitungen von f.

(b) Zeigen Sie: f ist in $(0,0)$ unstetig[167].

Lemma 6.97. *Sei $f : U \to \mathbb{R}^m$ eine in einem Punkt $u \in U$ total differenzierbare Funktion und sei $v \in \mathbb{R}^n$ mit $\|v\| = 1$ beliebig. Dann ist f in u in Richtung v ableitbar und es gilt*

$$D_v f(u) = Df(u)(v).$$

Insbesondere ist f in u partiell differenzierbar, und für alle $j \in \{1, \dots, n\}$ gilt

$$\frac{\partial f}{\partial x_j}(u) = D_j f(u) = Df(u)(e_j).$$

Beweis. Wenn f in u total differenzierbar ist, so gilt offenbar

$$\begin{aligned} D_v f(u) &= \lim_{h\to 0} \frac{f(u+hv) - f(u)}{h} = \lim_{h\to 0} \frac{Df(u)(hv) + \phi(hv)}{h} \\ &= \lim_{h\to 0} \frac{hDf(u)(v) + \phi(hv)}{h} = Df(u)(v) + \underbrace{\lim_{h\to 0} \frac{\phi(hv)}{h}}_{=0} \\ &= Df(u)(v). \end{aligned}$$

□

Bemerkung 6.98. Aus der Linearen Algebra wissen wir, dass eine lineare Abbildung $L : \mathbb{R}^n \to \mathbb{R}^m$ durch die Bilder einer Basis von $\mathbb{R}^n$ eindeutig bestimmt ist. Wählen wir in $\mathbb{R}^n$ die Standardbasis $\{e_1, \dots, e_n\}$ und betrachten wir die reelle $(m \times n)$–Matrix $A_L \in \mathbb{R}^{m\times n}$ deren k-te Spalte gerade $L(e_k)$ ist, so erhalten wir für alle $x \in \mathbb{R}^n$:

$$L(x) = A_L \cdot x.$$

Die Matrix A_L wird darstellende Matrix der linearen Abbildung $L : \mathbb{R}^n \to \mathbb{R}^m$ (bzgl. der Standardbasen) genannt. Wir werden in Zukunft in der Notation nicht mehr zwischen der linearen Abbildung L und ihrer darstellenden Matrix (bzgl. den Standardbasen) A_L unterscheiden. Einige weitere Details dazu sind im Anhang B zusammengestellt.

Wenden wir nun diese Vorüberlegung auf das totale Differential $Df(u)$ einer in u total differenzierbaren Abbildung $f : U \to \mathbb{R}^m$ ($U \Subset \mathbb{R}^n$) an, so sehen wir zunächst, dass $Df(u)$ nach Lemma 6.97 eindeutig bestimmt ist, da das Bild von e_j unter $Df(u)$ gerade $Df(u)(e_j) = \frac{\partial f}{\partial x_j}(u) = D_j f(u) \in \mathbb{R}^m$ sein muß. Die j-te Spalte der darstellenden Matrix $J_u f$ von $Df(u)$ ist also $\frac{\partial f}{\partial x_j}(u) = D_j f(u) \in \mathbb{R}^m$. Somit erhalten wir

$$J_u f = Df(u) = \left(\tfrac{\partial f}{\partial x_1}(u), \quad \dots, \quad \tfrac{\partial f}{\partial x_n}(u)\right) = \big(D_1 f(u), \dots, D_n f(u)\big) \in \mathbb{R}^{m\times n}. \tag{6.8}$$

Die Matrix $J_u f$ heißt *Jacobimatrix*[168] oder *Funktionalmatrix* von f im Punkt u.

[167] und damit in $(0,0)$ auch nicht total differenzierbar

[168] Carl Gustav Jacob Jacobi (1804–1851), deutscher Mathematiker

Aufgabe 6.99. Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge, $u \in U$ und $f = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$ eine Abbildung. Zeigen Sie, dass folgende Aussagen äquivalent sind:

(i) Die Funktion f ist in u total differenzierbar.

(ii) Alle Komponentenfunktionen $f_j : U \to \mathbb{R}$ $(j = 1 \ldots, m)$ von f sind in u total differenzierbar.

Satz 6.100. *Sei $f : U \to \mathbb{R}^m$, $U \Subset \mathbb{R}^n$, eine partiell differenzierbare Funktion und seien alle partiellen Ableitungen von f im Punkt $u = (u_1, \ldots, u_n) \in U$ stetig. Dann ist f in u total differenzierbar.*

Beweis. Nach Aufgabe 6.99 genügt es $m = 1$ anzunehmen. Aufgrund der Offenheit von U gibt es ein $\delta > 0$, sodass $B_\delta(u) \subset U$ gilt. Sei $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$ mit $0 < \|\xi\| < \delta$. Wir betrachten nun folgende $n + 1$ Punkte in U :

$$\eta_j := u + \sum_{k=1}^{j} \xi_k e_k = \begin{pmatrix} u_1 + \xi_1 \\ \vdots \\ u_j + \xi_j \\ u_{j+1} \\ \vdots \\ u_n \end{pmatrix}, \quad j = 0, \ldots, n.$$

Es gilt $\eta_{j+1} - \eta_j = \xi_{j+1} e_{j+1}$ sowie $\eta_0 = u$ und $\eta_n = u + \xi$. Durch diesen „Trick“ unterschieden sich η_{j+1} und η_j nur in einer Komponente und wir können den Mittelwertsatz der Differentialrechnung für Funktionen einer reellen Veränderlichen (Theorem 4.19) auf die differenzierbaren Funktionen

$$h_j : [0, 1] \to \mathbb{R}, \quad t \mapsto f(\eta_{j-1} + t\xi_j e_j), \quad j = 1, \ldots, n$$

anwenden und erhalten: Für alle $j = 1, \ldots, n$ gibt es ein $\phi_{j\xi} \in [0, 1]$, sodass

$$f(\eta_j) - f(\eta_{j-1}) = h_j(0) - h_j(1) = h'(\phi_{j\xi}) = D_j f\big(\underbrace{\eta_{j-1} + \phi_{j\xi}\xi_j e_j}_{=: y_{j\xi}}\big)\xi_j$$

(Die letzte Gleichung ergibt sich hierbei aus Bemerkung 6.86). Durch Summation folgt

$$f(u + \xi) - f(u) = f(\eta_n) - f(\eta_0) = \sum_{j=1}^{n} (f(\eta_j) - f(\eta_{j-1})) = \sum_{j=1}^{n} D_j f(y_{j\xi})\xi_j.$$

Wir setzen $A = (D_1 f(u), \ldots, D_n f(u)) \in \mathbb{R}^{m \times n}$ und erhalten mit

$$\phi(\xi) := \sum_{k=1}^{n} \big(D_k f(y_{k\xi}) - D_k f(u)\big)\xi_k$$

nun

$$f(u + \xi) = f(u) + A\xi + \phi(\xi).$$

Es bleibt zu zeigen, dass
$$\lim_{\xi\to 0}\frac{\phi(\xi)}{\|\xi\|}=0$$
gilt. Da $\|\xi\|\geq|\xi_k|$ für alle ξ und alle $k=1,\ldots,n$ gilt, erhalten wir $\frac{|\xi_j|}{\|\xi\|}\leq 1$, falls $\xi\neq 0$. Aus $\xi\to 0$ folgt $y_{k\xi}\to u$ für alle $k\in\{1,\ldots,n\}$ und wegen der Stetigkeit aller partiellen Ableitungen in u auch $D_kf(y_{k\xi})\to D_kf(u)$ für alle $k=1,\ldots,n$ und somit
$$0\leq\left|\frac{\phi(\xi)}{\|\xi\|}\right|\leq\sum_{k=1}^{n}\frac{|\xi_k|}{\|\xi\|}\,|D_kf(y_{k\xi})-D_kf(u)|\leq\sum_{k=1}^{n}|D_kf(y_{k\xi})-D_kf(u)|\to 0,$$
also
$$\lim_{\xi\to 0}\frac{\phi(\xi)}{\|\xi\|}=0.$$
□

Aufgabe 6.101 (4+4)**.**
$$f:\mathbb{R}^2\to\mathbb{R},\quad (x,y)\mapsto\begin{cases}(x^2+y^2)\sin\left(\frac{1}{\sqrt{x^2+y^2}}\right), & \text{wenn}\quad (x,y)\neq(0,0)\\ 0, & \text{wenn}\quad (x,y)=(0,0)\end{cases}$$

(a) Zeigen Sie: f ist $(0,0)$ total differenzierbar.

(b) Zeigen Sie: Alle ersten partiellen Ableitungen von f sind in $(0,0)$ nicht stetig.

Definition 6.102. Sei $U \Subset \mathbb{R}^n$. Eine Funktion $f:U\to\mathbb{R}^m$ heißt k-mal *stetig differenzierbar*, wenn *alle* k-ten partiellen Ableitung von f auf U existieren[169] und stetig sind.
$$\mathcal{C}^k(U,\mathbb{R}^m):=\{f:U\to\mathbb{R}^m:\ f \text{ ist } k\text{-mal stetig differenzierbar}\}$$

Bemerkung 6.103. Wir haben gesehen, dass gilt:
$$\begin{array}{ccc} f \text{ einmal stetig partiell differenzierbar} & & \\ \Downarrow & & \\ f \text{ ist total differenzierbar} & \Longrightarrow & f \text{ stetig}\\ \Downarrow & & \\ f \text{ ist partiell differenzierbar} & & \end{array}$$

Die Umkehrungen dieser Schlußrichtungen gelten jeweils im Allgemeinen *nicht.*

Satz 6.104. *Sei $f:U\to\mathbb{R}^m$ k-mal stetig partiell differenzierbar mit $k\in\mathbb{N}^*$. Sei $1\leq l\leq k$, dann sind alle l-ten partiellen Ableitungen von f stetig.*

[169] Damit müssen natürlich auch alle r-ten partiellen Ableitungen von f für $1\leq r\leq k$ existieren.

Beweis. Es genügt zu zeigen, dass alle $(k-1)$-ten partiellen Ableitungen von f stetig sind. Für $1 \leq l < k-1$ folgt die Aussage durch wiederholte Anwendung des folgenden Arguments:

Seien $j_1, \ldots, j_{k-1} \in \{1, \ldots, n\}$ beliebig vorgegeben und $g := D_{j_{k-1}} D_{j_{k-2}} \ldots D_{j_1} f$. Da g nach Voraussetzung einmal stetig partiell differenzierbar ist, ist g total differenzierbar und somit stetig. Also ist $D_{j_{k-1}} D_{j_{k-2}} \ldots D_{j_1} f$ stetig. Da $j_1, \ldots, j_{k-1}$ beliebig waren, ist f auch $(k-1)$-mal stetig differenzierbar. □

Satz 6.105 (Kettenregel)**.** *Sei $U \Subset \mathbb{R}^n$ und $V \Subset \mathbb{R}^m$ offene Teilmengen sowie $g : U \to V \subset \mathbb{R}^m$ und $f : V \to \mathbb{R}^l$ Abbildungen. Wenn g in $u \in U$ total differenzierbar ist, und f in $g(u) \in V$ total differenzierbar ist, dann ist auch die Abbildung*

$$f \circ g : U \to \mathbb{R}^l$$

im Punkt u total differenzierbar und es gilt

$$D(f \circ g)(u) = Df(g(u)) \circ Dg(u),$$

oder durch Jacobimatrizen ausgedrückt

$$J_u(f \circ g) = J_{g(u)} f \cdot J_u g.$$

Beweis. Sei $\xi \in \mathbb{R}^n \setminus \{0\}$ mit $\|\xi\|$ klein, dann gilt, da f in $g(u)$ differenzierbar ist:

$$\begin{aligned}
(f \circ g)(u+\xi) - (f \circ g)(u) &= f(g(u+\xi)) - f(g(u)) \\
&= f\big(g(u) + \underbrace{(g(u+\xi) - g(u))}_{=:h(\xi)}\big) - f(g(u)) \\
&= f(g(u) + h(\xi)) - f(g(u)) \\
&= Df(g(u))(h(\xi)) + \phi_f(h(\xi)).
\end{aligned}$$

Da g in u differenzierbar ist, folgt

$$h(\xi) = g(u+\xi) - g(u) = Dg(u)\xi + \phi_g(\xi)$$

und somit wegen der Linearität von $Df(g(u))$:

$$\begin{aligned}
(f \circ g)(u+\xi) - (f \circ g)(u) &= Df(g(u))(Dg(u)\xi) + \underbrace{Df(g(u))(\phi_g(\xi)) + \phi_f(h(\xi))}_{=:\psi(\xi)} \\
&= (Df(g(u)) \circ Dg(u))(\xi) + \psi(\xi).
\end{aligned}$$

Um den Beweis abzuschliessen, müssen wir zeigen, dass

$$\lim_{\xi \to 0} \frac{\psi(\xi)}{\|\xi\|} = \lim_{\xi \to 0} \frac{Df(g(u))(\phi_g(\xi)) + \phi_f(h(\xi))}{\|\xi\|} = 0$$

gilt. Aufgrund der Linearität von $Df(g(u))$ und seiner Stetigkeit in 0 erhalten wir für $\xi \to 0$:

$$\frac{Df(g(u))(\phi_g(\xi))}{\|\xi\|} = Df(g(u))\Bigg(\underbrace{\frac{\phi_g(\xi)}{\|\xi\|}}_{\to 0}\Bigg) \to 0.$$

Desweiteren gilt für $\xi \to 0$ wegen der Stetigkeit von g in u auch $h(\xi) = (g(u+\xi)-g(u)) = (Dg(u)(\xi) + \phi_g(\xi)) \to 0$. Wir folgern so für $\xi \to 0$:

$$\frac{\phi_f(h(\xi))}{\|\xi\|} = \frac{\phi_f(h(\xi))}{\|h(\xi)\|}\frac{\|h(\xi)\|}{\|\xi\|} \leq \underbrace{\frac{\phi_f(h(\xi))}{\|h(\xi)\|}}_{\to 0}\Bigg(\underbrace{\frac{\|Dg(u)(\xi)\|}{\|\xi\|}}_{\leq \|Dg(u)\|_{\text{op}}} + \underbrace{\frac{\|\phi_g(\xi)\|}{\|\xi\|}}_{\to 0}\Bigg) \to 0,$$

wobei wir die Funktion $\frac{\phi_f(h(\xi))}{\|h(\xi)\|}$ in den Nullstellen von h durch 0 stetig fortgesetzt haben. □

Bemerkung 6.106 (Produktregel)**.** Es gibt auch ein höher dimensionales Analogon zur Produktregel für Ableitungen reellwertiger Funktionen einer reelle Variable: Sei $U \Subset \mathbb{R}^n$ und seien $f, g : U \to \mathbb{R}^m$ zwei Funktionen sowie $b : \mathbb{R}^m \to \mathbb{R}^m \to \mathbb{R}^k$ eine bilineare Abbildung. Wir betrachten die Funktion

$$f \times_b g : U \to \mathbb{R}^k, \quad u \mapsto b(f(u), g(u)).$$

Sind f und g in $u \in U$ (total) differenzierbar, dann ist auch $f \times_b g$ in u (total) differenzierbar mit

$$D(f \times_b g)(u)(\xi) = b(Df(u)\xi, g(u)) + b(f(u), Dg(u)\xi).$$

Zum Beweis und für weitere Details und Erläuterungen verweisen wir auf [K2, Abschn. 3.1].

Theorem 6.107 (Mittelwertsatz)**.** *Sei $U \Subset \mathbb{R}^n$ und $f = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$ eine stetig differenzierbare Funktion, und seien $u \in U$ und $\xi \in \mathbb{R}^n$ gegeben, sodass das Bild der Strecke $[0,1] \to \mathbb{R}^n$, $t \mapsto u+t\xi$ von u nach $u+\xi$ ganz in U liegt, d. h. $u+[0,1]\xi \subset U$. Dann gilt*[170]

$$f(u+\xi) - f(u) = \left(\int_0^1 (J_{u+t\xi} f)dt\right) \cdot \xi = \int_0^1 ((J_{u+t\xi} f) \cdot \xi)dt.$$

[170]Identifizieren wir den Vektorraum $\mathbb{R}^{m\times n}$ aller reellen $(m \times n)$-Matrizen mit $\mathbb{R}^{mn}$, so können wir wie Beobachtung 6.79 das Integral einer stetige Abbildung $A : [t_0, t_1] \to \mathbb{R}^{m\times n}$, $t \mapsto A(t) = (a_{jk}(t))$ (d. h. alle Koeffizientenfunktionen a_{jk} sind stetig) definieren durch

$$\int_{t_0}^{t_1} A(t)dt = \begin{pmatrix} \int_{t_0}^{t_1} a_{11}(t)dt & \ldots & \int_{t_0}^{t_1} a_{1n}(t)dt \\ \vdots & \ddots & \vdots \\ \int_{t_0}^{t_1} a_{m1}(t)dt & \ldots & \int_{t_0}^{t_1} a_{mn}(t)dt \end{pmatrix} \in \mathbb{R}^{m\times n}.$$

Beweis. Wir betrachten die Funktionen

$$g_j : [0,1] \to \mathbb{R},\ t \mapsto f_j(u+t\xi), \qquad j = 1,\dots,m.$$

Nach der Kettenregel ist $g_j,\ j = 1,\dots,m$, stetig differenzierbar. Mit dem Hauptsatz der Differential- und Integralrechnung und der Kettenregel erhalten wir:

$$\begin{aligned} f_j(u+\xi) - f_j(u) &= g_j(1) - g_j(0) = \int_0^1 g_j'(t)dt \\ &= \int_0^1 \left(\sum_{k=1}^n \frac{\partial f_j}{\partial x_k}(u+t\xi)\xi_k \right) dt = \sum_{k=1}^n \left(\int_0^1 \frac{\partial f_j}{\partial x_k}(u+t\xi)dt \right) \xi_k. \end{aligned}$$

Die Aussage folgt nun aus der Definition der Matrixmultiplikation und der Gleichung (6.8) für die Jacobimatrix. □

Korollar 6.108 (Schrankensatz)**.** *Sei $U \Subset \mathbb{R}^n$ und $f = (f_1,\dots,f_m) : U \to \mathbb{R}^m$ eine stetig differenzierbare Funktion, und seien $u \in U$ und $\xi \in \mathbb{R}^n$ gegeben, sodass das Bild der Strecke $[0,1] \to \mathbb{R}^n,\ t \mapsto u+t\xi$ von u nach $u+\xi$ ganz in U liegt, d. h. $u+[0,1]\xi \subset U$. Sei*

$$M := \sup\{\|J_{u+t\xi} f\|_{\mathrm{op}} :\ t \in [0,1]\},$$

dann gilt

$$\|f(u+\xi) - f(u)\| \le M\|\xi\|.$$

Beweis. Mit Theorem 6.107 und Satz 6.80 erhalten wir:

$$\begin{aligned} \|f(u+\xi) - f(u)\| &= \left\| \int_0^1 (J_{u+t\xi} f \cdot \xi)dt \right\| \\ &\le \int_0^1 \|(J_{u+t\xi} f) \cdot \xi\| dt \\ &\le \int_0^1 (\underbrace{\|J_{u+t\xi} f\|_{\mathrm{op}}}_{\le M} \cdot \|\xi\|)dt \\ &\le M\|\xi\|. \end{aligned}$$

□

6.8. Taylorformel und „Kurvendiskussion“ in mehreren Veränderlichen

Dieses Kapitel orientiert sich stark an [Fo2, § 7] und an [K2, Absdchn. 2.5]. Haben wir zuvor (total) differenzierbare Funktionen lokal durch affin-lineare Funktionen approximiert, so möchten wir nun eine Taylorformel für Funktionen mehrerer Veränderlicher herleiten und so hinreichend oft stetig differenzierbare Funktionen lokal durch Polynome in mehreren Variablen approximieren.

Sei $U \Subset \mathbb{R}^n$ und $f : U \to \mathbb{R}$ eine $(p+1)$-mal stetig differenzierbare Abbildung. Ferner sei $u \in U$ sowie $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$, sodass für alle $t \in [0,1]$ gilt

$$h(t) := u + t\xi \in U.$$

Wir betrachten nun die Funktion

$$F : [0,1] \to \mathbb{R}, \quad t \mapsto (f \circ h)(t) = f(u + t\xi).$$

Dann erhalten rekursiv mit der Kettenregel für $k = 1, \dots p+1$:

$$\begin{aligned}
F'(t) &= Df(u+t\xi)h'(t) \;=\; \sum_{j=1}^{n} D_j f(u+t\xi)\xi_j,\\
F''(t) &= \sum_{j_2=1}^{n}\sum_{j_1=1}^{n} D_{j_2}D_{j_1} f(u+t\xi)\xi_{j_1}\xi_{j_2},\\
&\;\vdots\\
F^{(k)}(t) &= \sum_{j_k=1}^{n}\cdots\sum_{j_1=1}^{n}(D_{j_k}\cdots D_{j_1} f(u+t\xi))\xi_{j_1}\dots\xi_{j_k}.
\end{aligned}$$

Insbesondere ist mit f auch F eine $(p+1)$-mal stetig differenzierbare Abbildung.

Notation 6.109. Sei $f : U \to \mathbb{R}$ k-mal stetig differenzierbar, sei $u \in U$ und $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$, so setzen wir

$$d^k f(u)\xi^k := \sum_{j_1=1}^{n}\cdots\sum_{j_k=1}^{n}(D_{j_k}\cdots D_{j_1} f(u))\xi_{j_1}\dots\xi_{j_k}.$$

Zudem definieren wir

$$d^0 f(u)\xi^0 := f(u).$$

Mit Satz 4.59 gibt es ein $\tau \in [0,1]$, sodass gilt:

$$f(u+\xi) = F(1) = \sum_{k=0}^{p}\frac{1}{p!}F^{(k)}(0)1^k + \frac{1}{(p+1)!}F^{(p+1)}(\tau)1^{p+1}.$$

Somit erhalten wir:

Satz 6.110 (Taylorformel in mehreren Veränderlichen)**.** *Sei $U \Subset \mathbb{R}^n$ und $f : U \to \mathbb{R}$ eine $(p+1)$-mal stetig differenzierbare Funktion. Ferner sei $u \in U$ und $\xi \in \mathbb{R}^n$, sodass für alle $t \in [0,1]$ gilt $h(t) := u + t\xi \in U$. Dann gibt es ein $\tau \in [0,1]$, sodass*

$$f(u+\xi) = F(1) = \Big(\sum_{k=0}^{p}\frac{1}{k!}d^k f(u)\xi^k\Big) + \frac{1}{(p+1)!}d^{p+1} f(u+\tau\xi)\xi^{p+1}. \tag{6.9}$$

Bemerkung 6.111. Setzen wir $x = u+\xi$ und bezeichnen wir $[u, x]$ die Strecke zwischen u und x, d. h. $[u, x] = \{u+tx,\ t \in [0, 1]\}$, so können wir die Gleichung (6.9) in Anlehnung an Abschnitt 4.7 auch schreiben als:

$$f(x) = \underbrace{\Big(\sum_{k=0}^{p} \frac{1}{k!} d^k f(u)(x-u)^k\Big)}_{=:T_p^{f,u}(x)} + \underbrace{\frac{1}{(p+1)!} d^{p+1} f(\eta)(x-u)^{p+1}}_{=:R_p^{f,u}(x)}, \tag{6.10}$$

für einen passenden Punkt $\eta \in [u, x]$. Wir nennen

$$T_p^{f,u}(x) := \sum_{k=0}^{p} \frac{1}{k!} d^k f(u)(x-u)^k$$

Taylor–Polynom von f von Ordnung p in u (mit Entwicklungspunkt u)[171] und

$$R_p^{f,u}(x) = f(x) - T_p^{f,u}(x) = \frac{1}{(p+1)!} d^{p+1} f(\eta)(x-u)^{p+1}$$

Lagrange-Form des *Restgliedes* von $T_p^{f,u}(x)$. Das Taylor-Polynom $T_p^{f,u}(x)$ ist ein Polynom in den Komponenten $x_1, \dots, x_n$ von x.

Korollar 6.112. *Sei $U \Subset \mathbb{R}^n$ und $f : U \to \mathbb{R}$ eine p-mal stetig differenzierbare Funktion $(p \geq 1)$ und sei $u \in U$. Sei $x \in U$, sodass $\{u + t(x-u) :\ t \in [0, 1]\}$ ganz in U liegt, dann gilt für $R_p^{f,u}(x) = f(x) - T_p^{f,u}(x)$ einerseits $R_p^{f,u}(u) = 0$ und*

$$\lim_{x \to u} \frac{R_p^{f,u}(x)}{\|x-u\|^p} = 0.$$

Beweis. Sei $\varepsilon > 0$. Da f p-mal stetig differenzierbar ist, gibt es ein $\delta > 0$ mit $B_\delta(u) \subset U$, sodass für alle $y \in B_\delta(u)$ gilt:

$$H(y) := \frac{1}{p!} \sum_{j_1=1}^{n} \cdots \sum_{j_p=1}^{n} \left| D_{j_1} \cdots D_{j_p} f(y) - D_{j_1} \cdots D_{j_p} f(u) \right| < \varepsilon.$$

Zu jedem $x \in B_\delta(u)$ gibt es nach der Taylorformel ein $\eta_x \in [u, z] = \{u + tx :\ t \in [0, 1]\}$, sodass

$$\begin{aligned} f(x) &= T_{p-1}^{f,u}(x) + \frac{1}{p!} d^p f(\eta_x)(x-u)^p \\ &= T_p^{f,u}(x) + \underbrace{\frac{1}{p!}\left(d^p f(\eta_x)(x-u)^p - d^p f(u)(x-u)^p\right)}_{=R_p^{f,u}(x)}. \end{aligned}$$

[171] Manchmal wird $T_p^{f,u}$ auch als *k-Jet* von f in u bezeichnet.

Somit erhalten wir

$$\begin{aligned}
\left|R_p^{f,u}(x)\right| &= \frac{1}{p!}\left|\big(d^p f(\eta_x) - d^p f(u)\big)(x-u)^p\right| \\
&= \frac{1}{p!}\left|\sum_{j_1=1}^{n}\cdots\sum_{j_p=1}^{n}\big(D_{j_1}\cdots D_{j_p}f(\eta_x) - D_{j_1}\cdots D_{j_p}f(u)\big)\cdot(x_{j_1}-u_{j_1})\ldots(x_{j_p}-u_{j_p})\right| \\
&\leq \frac{1}{p!}\sum_{j_1=1}^{n}\cdots\sum_{j_p=1}^{n}\Big(\big|D_{j_1}\cdots D_{j_p}f(\eta_x) - D_{j_1}\cdots D_{j_p}f(u)\big|\cdot\underbrace{|x_{j_1}-u_{j_1}|}_{\leq\|x-u\|}\ldots\underbrace{|x_{j_p}-u_{j_p}|}_{\leq\|x-u\|}\Big) \\
&\leq \|x-u\|^p\underbrace{\frac{1}{p!}\sum_{j_1=1}^{n}\cdots\sum_{j_p=1}^{n}\big|D_{j_1}\cdots D_{j_p}f(\eta_x) - D_{j_1}\cdots D_{j_p}f(u)\big|}_{=H(\eta_x)} \\
&\leq \varepsilon\|x-u\|^p.
\end{aligned}$$

Damit folgt die Aussage. □

Bemerkung 6.113 (Taylorformel für $p = 1$)**.** Für $p = 1$ erhalten wir aus Gleichung 6.9:

$$f(u+\xi) = f(u) + \underbrace{\sum_{j=1}^{n} D_j f(u)\xi_j}_{=Df(u)(\xi)} + \underbrace{R_1^{f,u}(u+\xi)}_{=:\phi(\xi)},$$

wobei nach Korollar 6.112 gilt

$$\lim_{\xi\to 0}\frac{\phi(\xi)}{\|\xi\|} = 0.$$

Für $p = 1$ beschreibt also die Taylorformel gerade die lokale affin-lineare Approximation durch das totale Differential.

Bemerkung 6.114 (Taylorformel für $p = 2$)**.** Sei $f : U \to \mathbb{R}$ zweimal stetig differenzierbar. Wegen

$$d^2 f(u)\xi^2 := \sum_{j=1}^{n}\sum_{k=1}^{n}(D_j D_k f(u))\xi_j \ldots \xi_k.$$

erhalten wir mit

$$(6.11)\qquad (\mathrm{Hess} f)(u) := \big(D_j D_k f(u)\big)_{j,k} = \begin{pmatrix} D_1D_1f(u) & \ldots & D_1D_nf(u) \\ D_2D_1f(u) & \ldots & D_2D_nf(u) \\ \vdots & \ddots & \vdots \\ D_nD_1f(u) & \ldots & D_nD_nf(u) \end{pmatrix} \in \mathbb{R}^{n\times n}$$

die Taylorformel für $p = 2$:

$$(6.12)\quad f(u+\xi) = f(u) + \sum_{j=1}^{n} Df(u)\xi + \underbrace{\frac{1}{2}(\xi_1,\ldots\xi_n)\cdot(\mathrm{Hess} f)(u)\cdot\begin{pmatrix}\xi_1\\ \vdots \\ \xi_n\end{pmatrix}}_{\frac{1}{2}\langle\xi,(\mathrm{Hess} f)(u)\xi\rangle} + R_2^{f,u}(u+\xi).$$

Die Matrix $(\mathrm{Hess} f)(u) \in \mathbb{R}^{n\times n}$ in Gleichung (6.11) der zweiten partiellen Ableitungen von f in u wird *Hesse-Matrix*[172] von f in u genannt. Ist $f : U \to \mathbb{R}$ zweimal stetig partiell differenzierbar, so ist die Hesse-Matrix von f nach dem Satz von Schwarz (Theorem 6.89) symmetrisch.

Definition 6.115. Sei $U \Subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ partiell differenzierbar. Ein Punkt $u \in U$ heißt *kritischer Punkt* von f, wenn

$$D_j f(u) = 0 \in \mathbb{R}^{1\times n}$$

für alle $j = 1, \dots, n$ gilt.

Satz 6.116 (Notwendige Bedingung für lokale Extrema)**.** *Sei $U \Subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine partiell differenzierbare Funktion. Hat f in $u \in U$ ein lokales Extremum*[173]*, dann ist u ein kritischer Punkt von f.*

Beweis. Wir betrachten die n differenzierbaren Funktionen

$$h_j : (-\varepsilon, \varepsilon) \to \mathbb{R}, \quad t \mapsto f(u + te_j).$$

Hat f in u ein lokales Extremum, so haben auch alle h_j in $t = 0$ ein lokales Extremum. Nach dem Satz von Fermat, Satz 4.17, folgt dann

$$0 = h_j'(0) = D_j f(u)$$

für alle $j = 1, \dots, n$. □

Bereits aus der Schule kennen Sie ein hinreichendes Kriterium für das Vorliegen lokaler Extrema einer zweimal stetig differenzierbaren Funktionen $f : (a, b) \to \mathbb{R}$. Ist $x_o \in (a, b)$ ein kritischer Punkt von f mit $f''(x_o) > 0$, so hat f in x_o ein lokales Minimum. Ist x_o ein kritischer Punkt von f mit $f''(x_o) < 0$, so hat f in x_o ein lokales Maximum. In mehreren Veränderlichen ($n \geq 2$) kommt eine neue Situation hinzu: Längs gewisser Kurven durch einen kritischen Punkt u_o kann $f : U \to \mathbb{R}$ bei u_o ein lokales Minimum haben, und längs anderer Kurven durch u_o ein lokales Maximum bei u_o haben. Für $U \Subset \mathbb{R}^2$ sieht der Graph von f dann um u_o einem Sattel nicht unähnlich. Längs des Pferde- oder Trampeltierrückens liegt in u_o ein lokales Minimum vor. Damit rutscht der Reiter weder nach vorne noch nach hinten herunter. Senkrecht dazu (zwischen den Beinen des Reiters) hat f in u_o hingegen ein Maximum. Man nennt daher einen kritischen Punkt u_o von f in welchem kein lokales Extremum von f vorliegt auch *Sattelpunkt* von f. Alpinisten mag ein Sattelpunkt auch an eine Passhöhe erinnern.

Definition 6.117. Eine reelle symmetrische Matrix[174] $A \in \mathbb{R}^{n\times n}$ heißt

[172] Ludwig Otto Hesse (1811–1874), deutscher Mathematiker

[173] Erinnerung: f hat in $u \in U$ ein lokales Extremum, falls f in u ein lokales Maximum oder ein lokales Minimum annimmt. Hierbei hat f in u ein lokales Maximum (bzw. Minimum), falls es ein $\varepsilon > 0$ gibt, sodass $B_\varepsilon(u) \subset U$ und $f(y) \leq f(u)$ (bzw. $f(y) \geq f(u)$) für alle $y \in B_\varepsilon(u)$ gilt.

[174] Aus der Linearen Algebra wissen wir, dass A dann über $\mathbb{R}$ diagonalisierbar ist.

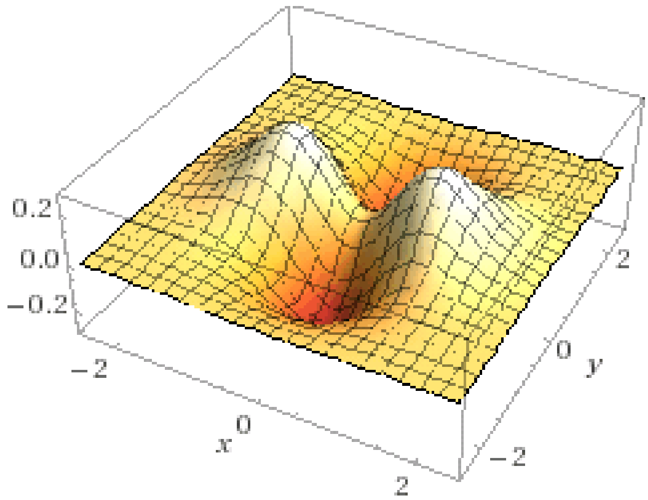

Abbildung 6.1.: $f(x,y) = \frac{3}{4}\exp(-(x^2+y^2))(x^2-y^2)$, Graphik erstellt mit `http://www.wolframalpha.com`

- *degeneriert*, wenn $\det(A) = 0$ gilt. Äquivalent ist A degeneriert, wenn 0 ein Eigenwert von A ist.
- *positiv definit*, wenn für alle $\xi \in \mathbb{R}^n \setminus \{0\}$ gilt
$$\langle \xi, a\xi \rangle = \xi^T A \xi > 0.$$
Äquivalent ist A positiv definit, wenn alle Eigenwerte von A positiv sind[175].
- *negativ definit*, wenn $-A$ positiv definit ist (oder wenn alle Eigenwerte von A negativ sind).
- *indefinit*, wenn A nicht degeneriert und weder positiv noch negativ definit ist, d. h. kein Eigenwert von A verschwindet und A besitzt sowohl positive wie auch negative Eigenwerte.

Satz 6.118 (Hinreichende Bedingung für lokale Extrema). *Sei $U \Subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ zweimal stetig differenzierbar und $u \in U$ ein kritischer Punkt von f. Dann gilt:*

(i) Ist $(\mathrm{Hess} f)(u)$ *positiv definit, so hat f in u ein isoliertes lokales Minimum*[176].

(ii) Ist $(\mathrm{Hess} f)(u)$ *negativ definit, so hat f in u ein isoliertes lokales Maximum*[177].

(iii) Besitzt $(\mathrm{Hess} f)(u)$ *sowohl einen positiven wie auch einen negativen Eigenwert, so hat f in u kein lokales Extremum (also einen Sattelpunkt).*

[175] In der linearen Algebra werden Sie auch noch das *Hauptminorenkriterium* kennen lernen: Eine reelle symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann positiv definit, wenn $\det(A_k) > 0$ für alle $k = 1, \dots, n$ gilt. Hierbei ist A_k der linke obere $(k \times k)$-Block von A.

[176] d. h. es gibt ein $\delta > 0$, sodass $B_\delta(u) \subset U$ und $f(x) > f(u)$ für alle $x \in B_\delta(u) \setminus \{u\}$ gilt.

[177] d. h. es gibt ein $\delta > 0$, sodass $B_\delta(u) \subset U$ und $f(x) < f(u)$ für alle $x \in B_\delta(u) \setminus \{u\}$ gilt.

Beweis. Aus Korollar 6.112 und Bemerkung 6.114 folgt wegen $Df(u) = 0$ für $\xi \in \mathbb{R}^n$ mit $\|\xi\|$ hinreichend klein:

$$f(u+\xi) = f(u) + \frac{1}{2}\langle \xi,\ A\xi \rangle + \phi(\xi) \tag{6.13}$$

mit $A := (\mathrm{Hess} f)(u)$ und $\lim\limits_{\xi \to 0} \frac{\phi(\xi)}{\|\xi\|^2} = 0$, d. h.

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall \xi \in \mathbb{R}^n:\ (\|\xi\| < \delta) \implies \big(|\phi(\xi)| \leq \varepsilon \|\xi\|^2\big). \tag{6.14}$$

ad (i). Sei $A := (\mathrm{Hess} f)(u)$ positiv definit und $S = S^{n-1} = \{\xi \in \mathbb{R}^n :\ \|\xi\| = 1$ die Einheitssphäre in $\mathbb{R}^n$ (mit der euklidischen Norm). Die offenbar stetige Funktion

$$h : S \to \mathbb{R}, \quad \xi \mapsto \langle \xi, A\xi \rangle$$

nimmt ein Minimum an, da die Einheitssphäre $S \subset \mathbb{R}^n$ kompakt ist (vgl. Korollar 6.62). Wegen $h(\xi) = \langle \xi, A\xi \rangle > 0$ für alle $\xi \in S$ gilt auch

$$\mu := \min\{h(\xi) : \xi \in S\} > 0.$$

Es gilt dann[178]

$$\forall \xi \in \mathbb{R}^n: \quad \langle \xi, A\xi \rangle \geq \mu \|\xi\|^2. \tag{6.15}$$

Sei nun $\delta > 0$ hinreichend klein, sodass zudem für alle $\xi \in \mathbb{R}^n$ mit $\|\xi\| < \delta$ gilt

$$|\phi(\xi)| \leq \frac{\mu}{4}\|\xi\|^2$$

(vgl. (6.14)). Somit erhalten wir aus (6.13) und (6.15) für alle $\xi \in \mathbb{R}^n$ mit $0 < \|\xi\| < \delta$:

$$f(u+\xi) \geq f(u) + \frac{\mu}{4}\|\xi\|^2 > f(u).$$

Also hat f in u ein lokales Minimum.

ad (ii). Folgt aus (i) durch Übergang zu $-f$.

ad (iii). Sei $\alpha > 0$ ein positiver und $\beta < 0$ ein negativer Eigenwert von A und $\xi \in S$ ein Eigenvektor von A zum Eigenwert α und $\eta \in S$ ein Eigenvektor von A zum Eigenwert β, d. h.

$$A\xi = \alpha\xi \qquad \text{und} \qquad A\eta = \beta\eta.$$

[178] In der Tat: Für $\xi = 0$ ist diese Ungleichung offensichtlich wahr. Wenn $\xi \neq 0$, so erhalten wir mit $\frac{\xi}{\|\xi\|} \in S$:

$$\langle \xi, A\xi \rangle = \left\langle \|\xi\|\frac{\xi}{\|\xi\|},\ A\|\xi\|\frac{\xi}{\|\xi\|} \right\rangle = \|\xi\|^2 \underbrace{\left\langle \frac{\xi}{\|\xi\|},\ A\frac{\xi}{\|\xi\|} \right\rangle}_{\geq \mu} \geq \mu\|\xi\|^2.$$

Für alle $t \in \mathbb{R}$ gilt wegen $\|\xi\| = 1 = \|\eta\|$:

$$\langle t\xi,\ At\xi\rangle = \alpha t^2 \qquad \text{und} \qquad \langle t\eta,\ At\eta\rangle = \beta t^2.$$

Nach (6.13) erhalten wir somit für $|t|$ hinreichend klein

$$f(u+t\xi) = f(u) + \frac{\alpha}{2}t^2 + \phi(t\xi) \quad \text{und} \quad f(u+t\eta) = f(u) + \frac{\beta}{2}t^2 + \phi(t\eta).$$

Wegen (6.14) gibt es ein $\delta > 0$, sodass für alle $|t| < \delta$ zudem gilt

$$|\phi(t\xi)| \le \frac{\alpha}{4}t^2 \quad \text{und} \quad |\phi(t\eta)| \le \frac{-\beta}{4}t^2.$$

Insgesamt folgt für $0 < |t| < \delta$ also

$$f(u+t\xi) > f(u) \qquad \text{und} \qquad f(u+t\eta) < f(u)$$

und somit die Behauptung.

□

Bemerkung 6.119 (Vorsicht!)**.** Sei $f : U \to \mathbb{R}$ zweimal stetig partiell differenzierbar und sei $u \in U$ ein kritischer Punkt von f. Hat $(\mathrm{Hess} f)(u)$ keine negativen, wohl aber verschwindende Eigenwerte (oder keine positiven, wohl aber verschwindende Eigenwerte), so kann keine Aussage über das Vorliegen eines lokalen Extremums in u gemacht werden, wie folgende Beispiele zeigen (Übung): Alle Funktionen

$$\begin{aligned} f_1 :&\ \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto x^2 + y^4, \\ f_2 :&\ \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto x^2, \\ f_3 :&\ \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto x^2 + y^3, \end{aligned}$$

haben in $(0,0)$ einen kritischen Punkt mit

$$(\mathrm{Hess} f_j)(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \qquad j = 1,2,3.$$

Jedoch hat f_1 in $(0,0)$ ein isoliertes lokales Minimum, während f_2 in $(0,0)$ ein nicht-isoliertes lokales Minimum hat und f_3 in $(0,0)$ überhaupt kein lokales Extremum besitzt.

Aufgabe 6.120. Zeigen Sie für

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto 3x^2 - 4xy^2 + y^4.$$

(a) Sei $(a,b) \neq (0,0)$ und $g_{a,b} : \mathbb{R} \to \mathbb{R}^2$, $t \mapsto (at, bt)$, dann hat $f \circ g$ in $t = 0$ ein lokales Minimum.

(b) Die Funktion f hat in $(0,0)$ *kein* lokales Minimum.

Bemerkung 6.121 (Strategie zur Bestimmung globaler Extrema auf Kompakta)**.** Sei $K \subset \mathbb{R}^n$ ein Kompaktum mit nichtleerem Inneren, d. h. $K^\circ \neq \emptyset$, und sei $f : K \to \mathbb{R}$ eine stetige und auf K° partiell differenzierbare Funktion. Als stetige Funktion auf einem Kompaktum nimmt f ein Maximum und ein Minimum an. Wenn man dieses explizit bestimmen möchte, kann man versuchen, wie folgt vorzugehen:

- 1. Schritt: Man bestimme alle kritischen Stellen von $f|_{K^\circ}$.

- 2. Schritt: Man bestimme alle Extrema von f auf dem Rand ∂K von K.

- 3. Schritt: Man vergleiche die Funktionswerte von f in den kritischen Stellen von $f|_{K^\circ}$ mit dem Extremwerten von $f|_{\partial K}$.

Bemerkung 6.122 (Strategie zur Bestimmung globaler Extrema)**.** Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine partiell differenzierbare Funktion.

- 1. Schritt: Man bestimme die Funktionswerte von f in alle kritischen Stellen von f. M sei der größte Funktionswert und m der kleinste Funktionswert von f an einer kritischen Stelle.

- 2. Schritt: Wenn es ein $R > 0$ gibt, sodass $f|_{\mathbb{R}^n \setminus B_R(0)}$ nur Werte kleiner als M (bzw. nur Werte größer als m) annimmt, dann ist M das globale Maximum (bzw. m das gobale Minimum) von f.

6.9. Satz von der Umkehrfunktion

Wir kommen nun zu einem Höhepunkt dieser Vorlesung, dem Satz von der Umkehrfunktion, auch Umkehrsatz genannt. In diesem Abschnitt folgen wir [Esch, Abschn. 13] und [K2, Abschn. 3.5].

Definition 6.123. Sei $U \Subset \mathbb{R}^n$ und $V \Subset \mathbb{R}^m$. Eine Abbildung $f : U \to V$ heißt *Diffeomorphismus* oder genauer $(\mathcal{C}^1)$-*Diffeomorphismus*, wenn f invertierbar ist[179] und sowohl f wie auch f^{-1} beide stetig partiell differenzierbar sind.

Bemerkung 6.124. Sei $f : U \to V$ ein Diffeomorphimus, wobei $U \Subset \mathbb{R}^n$ und $V \Subset \mathbb{R}^m$, dann existiert die Umkehrabbildung $f^{-1} : V \to U$, mit $f \circ f^{-1} = \mathrm{id}_V$ und $f^{-1} \circ f = \mathrm{id}_U$ Mit der Kettenregel erhalten wir also für $u \in U$

$$\mathrm{I}_n = J_u(f^{-1} \circ f) = J_{f(u)}(f^{-1}) \circ J_u f \quad \text{und} \quad \mathrm{I}_m = J_{f(u)}(f \circ f^{-1}) = J_u f \circ J_{f(u)}(f^{-1}).$$

Somit ist $J_{f(u)}(f^{-1})$ die zu $J_u f$ inverse Matrix, also $(J_u f)^{-1} = J_{f(u)}(f^{-1})$. Insbesondere gilt auch $n = m$.

[179] d. h. es gibt eine Abbildung $g : V \to U$, mit $f \circ g = \mathrm{id}_V$ und $g \circ f = \mathrm{id}_U$, notiert $g = f^{-1}$.

Theorem 6.125 (Satz von der lokalen Umkehrfunktion). *Sei $U \Subset \mathbb{R}^n$ und $u \in U$ sowie $f : U \to \mathbb{R}^n$ ein stetig partiell differenzierbare Abbildung. Wenn $Df(u)$ invertierbar ist, so gibt es offene Umgebungen $X, Y \Subset \mathbb{R}^n$ mit $u \in X \subset U$ und $f(u) \in Y$, sodass $f(X) = Y$ gilt und*

$$f|_X : X \to Y$$

ein Diffeomorphismus ist

Der hier präsentierte Beweis des Umkehrsatzes benutzt den Banachschen Fixpunktsatz in höheren Dimensionen (vgl. auch Theorem 2.33):

Theorem 6.126 (Banachscher Fixpunktsatz). *Sei $K \subset \mathbb{R}^n$ kompakt und $\psi : K \to K$ ein Kontraktion, d. h. es gibt eine Konstante κ mit $0 < \kappa < 1$, sodass für alle $x, y \in K$ gilt*

$$\|\psi(x) - \psi(y)\| \leq \kappa \|x - y\|.$$

Dann hat ψ genau einen Fixpunkt in K.

Beweis. Der Beweis von Theorem 2.33 läßt sich direkt auf Theorem 6.126 übertragen. Wir müssen nur X durch K und $|\dots|$ durch $\|\dots\|$ ersetzen. Genauer wählen wir ein $x_0 \in K$ und definieren rekursiv $x_{j+1} = \psi(x_j)$. So erhalten wir eine Cauchyfolge (x_j), welche gegen ein Element $x \in K$ konvergiert, da K kompakt ist. Dieser Limes $x \in K$ ist dann der einzige Fixpunkt von ψ in K. □

Beweis. (von Theorem 6.125): Indem wir ggf. zur Abbildung $h(x) = f(x + u) - f(u)$ übergehen können wir $u = 0$ und $f(u) = 0$ annehmen. Zudem können wir anstatt f auch

$$h(x) = (Df(u))^{-1}(f(x))$$

betrachten und somit ohne Einschränkung zudem $J_0 f = I_n$ annehmen.

1. *Konstruktion der lokalen Umkehrabbildung:*

   Wir haben nun die Gleichung $y = f(x)$ für kleine x und kleine y nach x aufzulösen. Hierzu formulieren wir dieses Fragestellung als Fixpunktproblem und betrachten den Ausdruck

   $$\psi_y(x) := y + x - f(x).$$

   Ist x_y ein Fixpunkt von $\psi_y(x)$, also $x_y = \psi_y(x_y) = y + x_y - f(x_y)$, dann gilt $f(x_y) = y$.

   Für alle $x \in U$ erhalten wir $J_x \psi_y = I_n - J_x f$. Wegen der Stetigkeit der partiellen Ableitungen, also der Spaltenvektoren der Jacobimatrix von f, gibt es ein $r > 0$, sodass folgende Aussagen gelten:

   - $B^a_{2r}(0) \subset U$
   - für alle $x \in B^a_{2r}(0)$ gilt $\|J_x \psi_y\|_{\text{op}} = \|I_n - J_x f\|_{\text{op}} \leq \frac{1}{2}$.

- für alle $x \in B^a_{2r}(0)$ gilt $\det(J_x f) \neq 0$[180].

Nach dem Schrankensatz (Korollar 6.108) gilt für alle $x_1, x_2 \in B^a_{2r}(0)$:

$$\|\psi_y(x_1) - \psi_y(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\|. \tag{6.16}$$

Mit der umgekehrten Dreiecksungleichung erhalten wir für $x \in B^a_{2r}(0)$ und $\|y\| < r$:

$$\|\psi_y(x)\| \leq \|\psi_y(x) - \underbrace{\psi_y(0)}_{=y}\| + \|y\| < 2r. \tag{6.17}$$

Fazit: Ist $\|y\| < r$, dann gilt $\psi_y(B^a_{2r}(0)) \subset B^a_{2r}(0)$ und die Abbildungen

$$\psi_y : B^a_{2r}(0) \to B^a_{2r}(0)$$

sind Kontraktionen mit Lipschitz-Konstante $\kappa = \frac{1}{2} \in (0,1)$. Da $B^a_{2r}(0)$ kompakt ist[181], hat ψ_y für jedes $y \in B_r(0)$ nach dem Banachschen Fixpunktsatz, Theorem 6.126, genau einen Fixpunkt $x_y \in B_{2r}(0)$ (vgl. (6.17)). Es gilt dann $y = f(x_y)$. Setzen wir $Y_0 := B_r(0) \Subset \mathbb{R}^n$ und $X_0 := f^{-1}\big(B_r(0)\big) \cap B_{2r}(0) \Subset \mathbb{R}^n$ (vgl. auch Fußnote 155), dann erfüllt

$$g : Y_0 \to X_0,\ y \mapsto x_y$$

die Gleichung

$$f(g(y)) = y. \tag{6.18}$$

2. *Nachweis der (Lipschitz)-Stetigkeit von g:*
   Seien $y_1, y_2 \in Y_0$ und $x_1 = g(y_1)$ sowie $x_2 = g(y_2)$. Dann gilt nach Konstruktion mit (6.16):

$$\begin{aligned}
\|g(y_1) - g(y_2)\| &= \|x_1 - x_2\| \\
&= \|\psi_0(x_1) - \psi_0(x_2) + f(x_1) - f(x_2)\| \\
&\leq \|\psi_0(x_1) - \psi_0(x_2)\| + \|\underbrace{f(x_1)}_{=y_1} - \underbrace{f(x_2)}_{=y_2}\| \\
&\leq \frac{1}{2}\|\underbrace{x_1}_{=g(y_1)} - \underbrace{x_2}_{=g(y_2)}\| + \|y_1 - y_2\|,
\end{aligned}$$

also insgesamt

$$\|g(y_1) - g(y_2)\| \leq 2\|y_1 - y_2\|. \tag{6.19}$$

Somit ist g offenbar stetig , ja sogar Lipschitz-stetig mit Lipschitz-Konstante 2.

[180] Hier verwenden wir die Stetigkeit der Determinantenfunktion einer Matrix, welche ein Polynom in den Koeffizienten der Matrix ist (Leibniz-Formel), sowie die Tatsache, dass für die invertierbare Matrix $J_0 f$ gilt $\det(J_0 f) \neq 0$. Diese letzte Forderung an r wird erst in einem späteren Schritt wichtig.

[181] da $\mathbb{R}^n$ endlich dimensional ist

3. *Zu zeigen: g ist stetig partiell differenzierbar:*
Sei $y \in Y_0$ und $x := g(y) \in X_0$. Wir setzen $k := g(y+h) - g(y)$ mit $h \in \mathbb{R}^n$, sodass $\|h\|$ hinreichend klein ist. Da g stetig ist, gilt $k \to 0$, wenn $h \to 0$. Also kann auch $\|k\|$ als hinreichend klein angenommen werden. Da f in x differenzierbar ist, gilt
$$h = f(x+k) - f(x) = Df(x)k + \phi_f(k)$$
mit $\lim\limits_{k\to 0} \frac{\phi_f(k)}{\|k\|} = 0$. Weil $x \in B^a_{2r}(0)$ ist $Df(x)$ invertierbar, und es folgt
$$g(y+h) - g(y) = k = (Df(x))^{-1}h - (Df(x))^{-1}\phi_f(k).$$
Wir setzen $\phi_g(h) := -(Df(x))^{-1}\phi_f(k)$ und erhalten
$$g(y+h) = g(y) + (Df(x))^{-1}h + \phi_g(h).$$
Mit (6.19) gilt
$$\frac{\|k\|}{\|h\|} = \frac{\|g(y+h) - g(y)\|}{\|h\|} \leq 2$$
also
$$\begin{aligned} \frac{\|\phi_g(h)\|}{\|h\|} &= \frac{\|(Df(x))^{-1}\phi_f(k)\|}{\|h\|} \\ &\leq \|(Df(x))^{-1}\|_{\text{op}} \frac{\|\phi_f(k)\|}{\|k\|}\frac{\|k\|}{\|h\|} \\ &\leq 2\|(Df(x))^{-1}\|_{\text{op}} \frac{\|\phi_f(k)\|}{\|k\|}. \end{aligned}$$
und somit $\frac{\|\phi_g(h)\|}{\|h\|} \to 0$ für $h \to 0$, zumal $h \to 0$ auch $k \to 0$ impliziert. Somit ist aber auch g in y differenzierbar mit
$$Dg(y) = (Df(x))^{-1} = (Df(g(y)))^{-1}.$$
Da die Abbildungen g und Df und die Inversion von Matrizen[182] stetig sind, ist auch die Abbildung $Y_0 \to \mathbb{R}^{n\times n}$, $y \mapsto Dg(y) = (Df(g(y)))^{-1}$ stetig. Damit ist g stetig partiell differenzierbar.

4. *Definition von X und Y:*
Wir setzen $X := f^{-1}(Y_0) \cap X_0 \subset X_0$ und $Y := g^{-1}(X) \subset Y_0$. Beide Mengen sind offenbar offen. Sei $x \in X$, dann gilt $y := f(x) \in Y_0$ und $x' := g(f(x)) \in X_0 \subset B^a_{2r}(0)$ Wir erhalten mit (6.18) $f(x') = f(x) = y$. Somit sind x wie auch x' Fixpunkte von ψ_y, welche in $B^a_{2r}(0)$ liegen. Also gilt $x = x'$ und damit auch
$$g(f(x)) = x.$$
Somit ist $Y = f(X)$ und $f|_X : X \to Y$ ein Diffeomorphismus mit Umkehrabbildung $g|_Y$.

[182]Die Stetigkeit der Inversion von invertierbaren Matrizen sieht man z. B. an der Cramerschen Regel für das Inverse einer invertierbaren Matrix A. Die Koeffizienten von A^{-1} sind demnach rationale Funktionen in den Koeffizienten von A.

□

Bemerkung 6.127. In Theorem 6.125 ist die Voraussetzung, dass f stetig partiell differenzierbar unbedingt notwendig, wie an folgendem Beispiel deutlich wird:

Die Abbildung

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} x/2 + x^2 \sin(1/x), & \text{wenn} \quad x \neq 0 \\ 0, & \text{wenn} \quad x = 0 \end{cases}$$

ist auf ganz $\mathbb{R}$ differenzierbar mit

$$f' : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} 1/2 + 2x \sin(1/x) - \cos(1/x), & \text{wenn} \quad x \neq 0 \\ 1/2, & \text{wenn} \quad x = 0 \end{cases}$$

Damit ist $f'(0) = 1/2 \neq 0$ invertierbar, die Ableitung f' jedoch in 0 nicht stetig. Da f' in jedem offenen Intervall um 0 unendlich oft das Vorzeichen wechselt, ist f' auf keinem offenen Intervall um 0 monoton und somit auch auf keinem offenen Intervall um 0 lokal invertierbar.

Korollar 6.128 (Offenheitssatz)**.** *Sei $U \Subset \mathbb{R}^n$ und $f : U \to \mathbb{R}^n$ eine stetig partiell differenzierbare Abbildung. Wenn für alle $u \in U$ die Differentiale $Df(u)$ invertierbar sind, dann ist $f(U)$ eine offene Teilmenge von $\mathbb{R}^n$.*

Beweis. Nach Theorem 6.125 gibt es für jedes $u \in U$ eine offene Menge $X_u \Subset U$ mit $u \in X_u$, sodass $f(X_u) \Subset \mathbb{R}^n$ offen ist. Als Vereinigung offener Mengen ist auch $f(U) = \bigcup\limits_{u \in U} f(X_u)$ offen in $\mathbb{R}^n$. □

Korollar 6.129 (Diffeomorphiesatz)**.** *Sei $U \Subset \mathbb{R}^n$ und $f : U \to \mathbb{R}^n$ eine injektive stetig partiell differenzierbare Abbildung, sodass für alle $u \in U$ die Differentiale $Df(u)$ invertierbar sind. Dann ist f ein Diffeomorphismus auf sein Bild, d. h. $f : U \to f(U)$ ist ein Diffeomorphismus.*

Beweis. Sei $g : f(U) \to U$ die Umkehrabbildung von $f : U \to f(U)$ und $\tilde{U} \Subset U$ eine offene Teilmenge von U, dann ist $g^{-1}(\tilde{U}) = f(\tilde{U})$ nach Korollar 6.128 offen in $\mathbb{R}^n$. Damit ist g stetig (siehe auch Fußnote 155) und $f : U \to f(U)$ ein Homöomorphismus[183]. Da die Differenzierbarkeit in einem Punkt jeweils eine lokale Eigenschaft ist, ist g nach Theorem 6.125 auch in jedem Punkt stetig partiell differenzierbar. Also ist f ein Diffeomorphismus auf sein Bild. □

6.10. Satz über implizite Funktionen

In diesem Abschnitt folgen wir [Esch, Abschn. 14] und [K2, Abschn. 3.6].

[183] d. h. eine stetige umkehrbare Abbildungen, deren Umkehrabbildung wiederum stetig ist

Notation 6.130. Seien $U \Subset \mathbb{R}^n$ und $V \Subset \mathbb{R}^p$ offene Teilmengen, dann ist $U \times V$ eine offene Teilmenge von $\mathbb{R}^n \times \mathbb{R}^p = \mathbb{R}^{n+p}$. Sei $f : U \times V \to \mathbb{R}^q$ eine stetig differenzierbare Abbildung Für festes $u \in U$ betrachten wir die Abbildung

$$f_u : V \to \mathbb{R}^q, \quad v \mapsto f(u,v),$$

und für festes $v \in V$ betrachten wir die Abbildung

$$f^v : U \to \mathbb{R}^q, \quad u \mapsto f(u,v).$$

Wir definieren

$$D_U f(u,v) := D(f^v)(u) : \mathbb{R}^n \xrightarrow{\text{linear}} \mathbb{R}^q,$$

bzw. in Matrixschreibweise

$$J_U f(u,v) := J_u(f^v) \in \mathbb{R}^{q\times n},$$

sowie

$$D_V f(u,v) := D(f_u)(v) : \mathbb{R}^p \xrightarrow{\text{linear}} \mathbb{R}^q,$$

bzw. in Matrixschreibweise

$$J_V f(u,v) := J_v(f_u) \in \mathbb{R}^{q\times p}.$$

Damit gilt offenbar für $u \in U$ und $v \in V$

$$J_{(u,v)} f = \big(J_U f(u,v), \quad J_V f(u,v)\big) \in \mathbb{R}^{q\times(n+p)},$$

d. h. ist $x \in \mathbb{R}^n$ und $y \in \mathbb{R}^p$, also $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n\times p}$, so gilt

$$J_{(u,v)} f \begin{pmatrix} x \\ y \end{pmatrix} = J_U f(u,v)x + J_V f(u,v)y.$$

Wir führen noch folgende Projektionsabbildungen ein

$$\pi_1 : U \times V \to U,\ (u,v) \mapsto u \qquad \text{und} \qquad \pi_2 : U \times V \to V,\ (u,v) \mapsto v.$$

Theorem 6.131. *Seien $U \Subset \mathbb{R}^n$ und $V \Subset \mathbb{R}^p$ offene Teilmengen, und sei $f : U \times V \to \mathbb{R}^p$ eine stetig partiell differenzierbare Abbildung ($p = q$), welche in einem Punkt $(u_o, v_o) \in U\times V$ eine Nullstelle haben, d. h. $f(u_o, v_o) = 0 \in \mathbb{R}^p$. Wenn in diesem Punkt $J_V f(u_o, v_o) \in \mathbb{R}^{p\times p}$ invertierbar ist, dann gibt es*

- *eine offene Menge $X \Subset U \times V$ mit $(u_o, v_o) \in X$,*
- *eine offene Menge $Y \Subset \mathbb{R}^n \times \mathbb{R}^p$ mit $(u_o, 0) \in Y$,*
- *einen Diffeomorphismus $G : Y \to X$ mit $G(u_o, 0) = (u_o, v_o)$,*

sodass

$$f \circ G = \pi_2.$$

Theorem 6.132 (Satz über implizite Funktionen)**.**

(i) Seien $U \Subset \mathbb{R}^n$ und $V \Subset \mathbb{R}^p$ offene Teilmengen, und sei $f : U \times V \to \mathbb{R}^p$ eine stetig partiell differenzierbare Abbildung ($p = q$), welche in einem Punkt $(u_o, v_o) \in U \times V$ eine Nullstelle hat, d. h. $f(u_o, v_o) = 0 \in \mathbb{R}^p$. Wenn in diesem Punkt $J_V f(u_o, v_o) \in \mathbb{R}^{p\times p}$ invertierbar ist, dann gibt es

- *eine offene Menge $X \Subset U \times V$ mit $(u_o, v_o) \in X$,*
- *eine offene Menge $\tilde{U} \Subset U$ mit $u_o \in \tilde{U}$*
- *eine stetig partiell differenzierbare Abbildung $g : \tilde{U} \to \mathbb{R}^p$,*

sodass

$$f^{-1}(0) \cap X = \mathrm{Graph}(g) = \{(u, g(u)) : \ u \in \tilde{U}\},$$

d. h. für alle $(u, v) \in X$ gilt

$$f(u, v) = 0 \iff v = g(u).$$

(ii) Die Funktion g aus Teil (i) erfüllt

$$g(u_o) = v_o \quad \textit{und} \quad J_{u_o} g = -\big(J_V f(u_0, v_0)\big)^{-1} \cdot J_U f(u_o, v_o).$$

Beweis. (zu Theorem 6.131 und Theorem 6.132): Wir betrachten die stetig partiell differenzierbare Abbildung

$$F := \pi_1 \times f : U \times V \to \mathbb{R}^n \times \mathbb{R}^p, \qquad \begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} \pi_1(u, v) \\ f(u, v) \end{pmatrix} = \begin{pmatrix} u \\ f(u, v) \end{pmatrix}.$$

Damit gilt

$$J_{(u,v)} F = \begin{pmatrix} I_n & 0 \\ J_U f(u, v) & J_V f(u, v) \end{pmatrix}.$$

Da $J_V f(u_o, v_o)$ nach Voraussetzung invertierbar ist, gilt nach der Regel für die Berechnung von Determinanten von block–diagonalen Matrizen

$$\det(J_{(u_o,v_o)} F) = \underbrace{\det(I_n)}_{=1} \cdot \underbrace{\det(J_V f(u_o, v_o))}_{\neq 0} \neq 0.$$

Somit ist $J_{(u_o,v_o)} F$ invertierbar. Nach dem Umkehrsatz, Theorem 6.125, gibt es offene Mengen $X \Subset U \times V$ und $Y \in \mathbb{R}^{n+p}$ mit $(u_o, v_o) \in X$ und $F(u_o, v_o) = (u_o, 0) \in Y$, sodass $F|_X : X \to Y$ ein Diffeomorphismus ist. Sei $G = (F|_X)^{-1} : Y \to X$ die Umkehrabbildung von F. Aus $F \circ G = \mathrm{id}_Y$ erhalten wir für alle $(s, t) \in Y \subset \mathbb{R}^n \times \mathbb{R}^p$:

$$(s, t) = (F \circ G)(s, t) = \big(\pi_1(G(s, t)), \ f(G(s, t))\big).$$

Insbesondere gilt wegen der zweiten Komponente

$$f(G(s,t)) = t = \pi_2(s,t),$$

womit Theorem 6.131 bewiesen ist.

Aus $G \circ F = \mathrm{id}_X$ erhalten wir für alle $(u,v) \in X \subset \mathbb{R}^n \times \mathbb{R}^p$:

$$(u,v) = (G \circ F)(u,v) = G\big(u, f(u,v)\big) = \Big(\pi_1\big(G\big(u, f(u,v)\big)\big),\ \pi_2\big(G\big(u, f(u,v)\big)\big)\Big),$$

also insbesondere

$$v = \pi_2\big(G\big(u, f(u,v)\big)\big).$$

Betrachten wir $\tilde{U} := \pi_1(X)$ und $g : \tilde{U} \to \mathbb{R}^p,\ u \mapsto \pi_2(G(u,0))$, so sehen wir, dass für alle $(u,v) \in X$ gilt

$$f(u,v) = 0 \iff G(u,0) = (u,v) \iff g(u) = v.$$

Damit ist auch Teil (i) von Theorem 6.132 bewiesen.

Da $(u_o, v_o) \in X$ folgt aus $f(u_o, v_o) = 0$ wie zuvor $g(u_o) = v_o$. Weiterhin ist g als Verknüpfung differenzierbarer Abbildungen differenzierbar. Um die Jacobimatrix von g in u_o zu berechnen, betrachten wir die Hilfsfunktion

$$h : \tilde{U} \to X \subset \mathbb{R}^n \times \mathbb{R}^p, \quad u \mapsto \begin{pmatrix} u \\ g(u) \end{pmatrix}.$$

Für $u \in \tilde{U}$ gilt $(f \circ h)(u) = f(u, g(u)) = 0$ und somit

$$\begin{aligned} 0 &= J_u(f \circ h) \\ &= J_{h(u)} f \cdot J_u h \\ &= \big(J_U f(h(u)),\quad J_V f(h(u))\big) \cdot \begin{pmatrix} I_n \\ J_u g \end{pmatrix} \\ &= J_U f(h(u)) + J_V f(h(u)) \cdot J_u g \end{aligned}$$

also

$$J_V f(h(u)) \cdot J_u g = -J_U f(h(u)).$$

Insbesondere gilt für u_o mit $h(u_o) = (u_o, g(u_o)) = (u_o, v_o)$, da $J_U f(h(u_o)) = J_U f(u_o, v_o)$ invertierbar ist:

$$J_{u_o} g = -\big(J_V f(u_o, v_o)\big)^{-1} \cdot J_U f(u_o, v_o).$$

Somit ist auch Teil (ii) von Theorem 6.132) gezeigt. □

Beispiel 6.133. Wir betrachten das Gleichungssystem

$$\begin{array}{ccccccc} x^3 & + & y^3 & + & z^3 & = & 7, \\ xy & + & yz & + & zx & = & -2. \end{array}$$

Eine Lösung ist offenbar $\begin{pmatrix}2, & -1, & 0\end{pmatrix}$. Die Abbildung

$$f : U \times V \to \mathbb{R}^2, \qquad (x,(y,z)) \mapsto \begin{pmatrix} x^3+y^3+z^3-7 \\ xy+yz+zx+2 \end{pmatrix}$$

mit $U = \mathbb{R}$ und $V = \mathbb{R}^2$ hat also in $(2,(-1,0)) \in \mathbb{R}\times\mathbb{R}^2$ eine Nullstelle und jede Lösung des obigen Gleichungssystems ist eine Nullstelle von f und umgekehrt. Es gilt

$$D_V f(x,(y,z)) = \left(\tfrac{\partial f}{\partial y}(x,(y,z)), \ \ \tfrac{\partial f}{\partial z}(x,(y,z))\right) = \begin{pmatrix} 3y^2 & 3z^2 \\ x+z & y+x \end{pmatrix}$$

und somit

$$D_V f(2,(-1,0)) = \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}.$$

Da $\det\left(D_V f(2,(-1,0))\right) = 3 \neq 0$, gibt es nach Theorem 6.132 $a < 2 < b$ und eine stetig partiell differenzierbare Funktion

$$g : (a,b) \to \mathbb{R}^2$$

mit $g(2) = (-1,0)$, sodass $f(x,g(x)) = 0$ für alle $x \in (a,b)$. Nahe $(2,-1,0)$ sind also alle Lösungen des Gleichungssystems von der Form

$$(x, g(x)).$$

Ferner gilt

$$g'(2) = (-4, 9).$$

6.11. Untermannigfaltigkeiten des $\mathbb{R}^n$ und Extrema unter Nebenbedingungen

Wir orientieren uns in diesem Abschnitt an [Esch, Abschn. 15 f.] und [K2, Abschn. 3.7 f.].

Definition 6.134. Eine Teilmenge $M \subset \mathbb{R}^n$ heißt m-dimensionale *Untermannigfaltigkeit* von $\mathbb{R}^n$, wenn gilt: Für alle $u \in M$ gibt es eine offene Teilmenge $U \Subset \mathbb{R}^n$ mit $u \in U$ und eine offene Teilmenge $V \Subset \mathbb{R}^n$ mit $0 \in V$, sowie einen Diffeomorphismus $\Phi : U \to V$ mit $\Phi(u) = 0$, sodass gilt:

$$\begin{aligned} \Phi(M \cap U) &= V \cap \{(x_1,\ldots,x_m,0,\ldots,0) \in \mathbb{R}^n : \ x_1,\ldots x_m \in \mathbb{R}\} \\ &\cong V \cap \mathbb{R}^m. \end{aligned}$$

Die Abbildung Φ heißt *Karte* von M um den Punkt u.

Sei $c : (-\varepsilon,\varepsilon) \to M \subset \mathbb{R}^n$ eine differenzierbare Kurve mit $c(0) = u \in M$, deren Bild ganz in M liegt, dann heißt der Vektor $c'(0) \in \mathbb{R}^n$ *Tangentialvektor* an M in $u \in M$. Für $u \in M$ setzen wir

$$T_u M := \{v \in \mathbb{R}^n : \ v \text{ Tangentialvektor an } M \text{ in } u\}.$$

Die Menge $T_u M$ heißt *Tangentialraum* von M im Punkt u.

Lemma 6.135. *Ist $M \subset \mathbb{R}^n$ eine m-dimensionale Untermannigfaltigkeit von $\mathbb{R}^n$ und $u \in M$, dann ist T_uM ein m-dimensionaler Untervektorraum von $\mathbb{R}^n$.*

Beweis. Wir zeigen: Sei $\Phi : U \to V$ eine Karte von M um u, dann gilt

$$T_uM = \underbrace{D(\Phi^{-1})(0)}_{\mathbb{R}^n \xrightarrow[\text{linear}]{\cong} \mathbb{R}^n} \Big(\underbrace{\{(x \in \mathbb{R}^n : \ x_{m+1} = x_{m+2} = \cdots = x_n = 0\}}_{m-\text{dim. Untervekorraum von } \mathbb{R}^n} \Big).$$

Da die Bilder m-dimensionaler Untervektorräume unter Vektorraumisomorphismen wieder m-dimensionale Unterräume sind, haben wir damit dieses Lemma bewiesen.

Sei einerseits $c : (-\varepsilon, \varepsilon) \to M \subset \mathbb{R}^n$ eine differenzierbare Kurve mit $c(0) = u$. Dann können wir (durch Verkleinerung von $\varepsilon > 0$) ohne Einschränkung annehmen, dass Bild$(c) \subset U$ gilt. Damit ist $\Phi \circ c : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ eine differenzierbare Kurve deren Bild ganz im Untervektorraum $\{x \in \mathbb{R}^n : \ x_{m+1} = \cdots = x_n = 0\}$ liegt. Also liegt auch $(\Phi \circ c)'(0) = D\Phi(u)(c'(0))$ in $\{x \in \mathbb{R}^n : \ x_{m+1} = \cdots = x_n = 0\}$. Damit liegt aber der Tangentialvektor $c'(0)$ von M in u in der Menge

$$D(\Phi^{-1})(0)\big(\{x \in \mathbb{R}^n : \ x_{m+1} = \cdots = x_n = 0\}\big).$$

Sei nun andererseits $w \in \{x \in \mathbb{R}^n : \ x_{m+1} = \cdots = x_n = 0\}$, dann ist $w = \gamma'(0)$ mit $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n, \ t \mapsto tw$, wobei hier $\varepsilon > 0$ hinreichend klein sei, sodass das Bild von γ ganz in $\Phi(U) \cap \{x \in \mathbb{R}^n : \ x_{m+1} = \cdots = x_n = 0\}$ liegt. Dann ist $c := \Phi^{-1} \circ \gamma$ eine differenzierbare Kurve in $\mathbb{R}^n$, deren Bild ganz in $U \cap M$ liegt und $\gamma(0) = u$ erfüllt. Somit ist $c'(0) = D(\Phi^{-1})(0)(w)$ ein Tangentialvektor von M in u. □

Definition 6.136. Sei $M \subset \mathbb{R}^n$ eine m-dimensionale Untermannigfaltigkeit und $u \in M$. Das orthogonale Komplement (bzgl. des Standardskalarprodukts $\langle \ldots, \ldots \rangle$)

$$N_uM = (T_uM)^\perp$$

von T_uM in $\mathbb{R}^n$ heißt *Normalraum* von M im Punkt u. Es gilt also

$$\mathbb{R}^n = N_uM \oplus T_uM.$$

Wichtige Beispiele von Untermannigfaltigkeiten sind reguläre Urbilder:

Definition 6.137. Sei $U \Subset \mathbb{R}^n$ eine offene Teilmenge, und sei $f : U \to \mathbb{R}^p$ mit $p \leq n$ eine stetig partiell differenzierbare Abbildung. Ein Punkt $u \in U$ wird *regulärer Punkt* von f genannt, wenn die lineare Abbildung $Df(u) : \mathbb{R}^n \to \mathbb{R}^p$ Rang p hat (also surjektiv ist).

Sei $y \in \mathbb{R}^p$, dann heißt sein Urbild

$$f^{-1}(\{y\}) = \{u \in U : \ f(u) = y\}$$

reguläres Urbild oder *reguläre Niveaumenge*, wenn *alle* Punkte in $f^{-1}(\{y\})$ reguläre Punkte von f sind.

Satz 6.138 (Reguläre Urbilder sind Untermannigfaltigkeiten)**.** *Sei $U \Subset \mathbb{R}^n$, und sei $f : U \to \mathbb{R}^p$, $p \leq n$, eine stetig partiell differenzierbare Abbildung. Ist $y \in \mathrm{Bild}(f)$ und $M := f^{-1}(\{y\})$ ein reguläres Urbild, dann ist M eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$, wobei $m = n - p$.*

Beweis. Sei $u \in M$ beliebig. Dann hat $Df(u)$ Rang p, d.h. p Spaltenvektoren von $J_u f$ sind linear unabhängig. Nach Umnummerierung der Variablen (Koordinatenachsen) können wir annehmen, dass die letzen p Spalten von $J_u f$ linear unabhängig sind. Wir schreiben nun $\mathbb{R}^m = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n : \; x_{m+1} = \cdots = x_n = 0\}$, $m = n - p$, und $\mathbb{R}^p = \{x \in \mathbb{R}^n : \; x_1 = \cdots = x_m = 0\}$ und erhalten

$$\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^p.$$

Nun gibt es einen offenen Quader $Q \Subset U$ mit $u \in Q$. Wir wählen offene Quader $V \Subset \mathbb{R}^m$ und $W \Subset \mathbb{R}^p$, sodass $V \times W = Q$. Wir wählen $v \in V$ und $w \in W$, sodass $u = (v, w)$. Sei $g := f|_Q - y$, dann gilt $g(u) = g(v, w) = 0$. Die Matrix $J_W g(v, w)$ besteht dann aus den p letzten Spalten von $J_u f$ und ist damit invertierbar. Theorem 6.131 sagt, dass es

- eine offene Menge $X \Subset Q$ mit $u = (v, w) \in X$,
- eine offene Menge $Y \Subset \mathbb{R}^n$ mit $(v, 0) \in Y$
- einen Diffeomorphismus $G : Y \to X$, mit $G(v, 0) = (v, w) = u$.

gibt, sodass

$$g \circ G = \pi_2.$$

Betrachten wir die offene Menge

$$\tilde{Y} := Y - (v, 0) = \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^p = \mathbb{R}^n : \; (x + v, y) \in Y\} \Subset \mathbb{R}^n$$

und den Diffeomorphismus $h : \tilde{Y} \to Y, (x, y) \mapsto (x + v, y)$, so erhalten wir $0 \in \tilde{Y}$ und

$$g \circ G \circ h = \pi_2 : \tilde{Y} \to \mathbb{R}^p.$$

Wir setzen $\Phi := (G \circ h)^{-1}$. Dann gilt

$$\begin{aligned} & x \in M \cap X \\ \iff \quad & g((G \circ h) \circ \Phi)(x)) = g(x) = 0 \\ \iff \quad & \pi_2(\Phi(x)) = 0 \\ \iff \quad & \Phi(x) \in \mathbb{R}^m \times \{0\}. \end{aligned}$$

Somit gilt

$$\Phi(M \cap X) = (\mathbb{R}^m \times \{0\}) \cap Y.$$

Somit ist Φ eine Karte von M um 0. $\square$

Beispiel 6.139. Ist $U \Subset \mathbb{R}^n$, dann ist $U = f^{-1}(\{0\})$ reguläres Urbild der Abbildung

$$f : U \to \mathbb{R}^0 = \{0\}, \quad u \mapsto 0$$

(Übung).

Beispiel 6.140. Betrachten wir die Abbildung

$$g : \mathbb{R}^n \to \mathbb{R}, \quad x = (x_1, \dots, x_n) \mapsto \|x\|^2 = x_1^2 + \cdots + x_n^2,$$

dann ist die Sphäre $S_r = \{x \in \mathbb{R}^n : \ \|x\| = r\}$ um den Ursprung mit Radius $r > 0$ reguläres Urbild, nämlich $S_r = g^{-1}(r^2)$ (Übung).

Beispiel 6.141. Wir betrachten den Vektorraum $\mathbb{R}^{n\times n}$ der reellen $(n \times n)$-Matrizen, welche wir mit $\mathbb{R}^{n^2}$ identifizieren können und den Vektorraum Sym_n der reellen symmetrischen $(n\times n)$-Matrizen, welcher Dimension $d = 1+2+\cdots+n = \frac{n(n+1)}{2}$ hat und somit mit $\mathbb{R}^d$ identifiziert werden kann.

Wir betrachten die Abbildung

$$f : \mathbb{R}^{n\times n} \to \mathrm{Sym}_n, \quad A \mapsto A \cdot A^T$$

Das Urbild der Einheitsmatrix I_n ist dann offenbar die orthogonale Gruppe $O_n \subset \mathbb{R}^{n\times n}$, da

$$f^{-1}(I_n) = \{A \in \mathbb{R}^{n\times n} : \ AA^T = I_n\} = \{A \in \mathbb{R}^{n\times n} : \ A^T = A^{-1}\} = O_n.$$

Wir möchten zeigen, dass $O_n = f^{-1}(I_n)$ ein reguläres Urbild ist. Dazu berechenen wir zunächst die Ableitung von f in einem beliebigen Punkt $A \in O_n$. Für $X \in \mathbb{R}^{n\times n} \setminus \{0\}$ gilt

$$\begin{aligned} f(A+X) &= (A+X)(A+X)^T \ = \ (A+X)(A^T + X^T) \\ &= \underbrace{AA^T}_{=f(A)} + (AX^T + XA^T) + XX^T \end{aligned}$$

Da für die Operatornorm $\|A \cdot B\|_{\mathrm{op}} \le \|A\|_{\mathrm{op}} \|B\|_{\mathrm{op}}$ gilt, erhalten wir

$$\frac{\|XX^T\|_{\mathrm{op}}}{\|X\|_{\mathrm{op}}} \le \frac{\|X\|_{\mathrm{op}} \cdot \|X^T\|_{\mathrm{op}}}{\|X\|_{\mathrm{op}}} = \frac{\|X\|_{\mathrm{op}} \cdot \|X\|_{\mathrm{op}}}{\|X\|_{\mathrm{op}}} = \|X\|_{\mathrm{op}} \xrightarrow[X\to 0]{} 0.$$

Weil zudem $X \mapsto AX^T + XA^T$ linear ist, gilt

$$Df(A) : \mathbb{R}^{n\times n} \to \mathrm{Sym}_n, \quad X \mapsto AX^T + XA^T.$$

Nun müssen wir zeigen, dass die Abbildung $Df(A)$ für $A \in O_n$ surjektiv ist:
Sei $Y \in \mathrm{Sym}_n$, dann folgt aus $AA^T = I_n$ sofort

$$Df(A)\left(\frac{1}{2} Y A\right) = Y.$$

Zusammenfassend erkennen wir, dass O_n ein reguläres Urbild und somit eine Untermannigfaltigkeit von $\mathbb{R}^{n\times n}$ ist.

Definition 6.142. Sei $U \Subset \mathbb{R}^n$ und $g : U \to \mathbb{R}$ eine partiell differenzierbare reellwertige Funktion, dann heißt die Abbildung

$$\nabla g : U \to \mathbb{R}^n, \quad u \mapsto \begin{pmatrix} D_1 g(u) \\ \vdots \\ D_n g(u) \end{pmatrix} = \begin{pmatrix} \frac{\partial g}{\partial x_1}(u) \\ \vdots \\ \frac{\partial g}{\partial x_n}(u) \end{pmatrix}$$

Gradient von g. Ist g in u differenzierbar, so gilt

$$\nabla g(u) = (J_u g)^T.$$

Lemma 6.143. *Sei $U \Subset \mathbb{R}^n$, und sei $f = \begin{pmatrix} f_1 \\ \vdots \\ f_p \end{pmatrix} : U \to \mathbb{R}^p$, $p \leq n$, eine stetig partiell differenzierbare Abbildung. Ist $y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} \in \operatorname{Bild}(f)$ und $M := f^{-1}(\{y\})$ ein reguläres Urbild sowie $u \in M$, dann ist $\{\nabla f_1(u), \ldots, \nabla f_p(u)\}$ eine Basis von $N_u M$.*

Beweis. Da M eine $(n-p)$-dimensionale Untermannigfaltigkeit von $\mathbb{R}^n$ ist, hat der Untervektorraum $T_u M$ Dimension $n-p$ und $N_u M$ somit Dimension p. Es genügt also zu zeigen:

(i) $\nabla f_j(u) \perp T_u M$ für alle $u \in M$ und für alle $j = 1, \ldots, p$,

(ii) $\{\nabla f_1(u), \ldots, \nabla f_p(u)\}$ ist linear unabhängig.

Wir zeigen zunächst Teil (i). Sei $v \in T_u M$ und $c : (-\varepsilon, \varepsilon) \to M \subset \mathbb{R}^n$ eine differenzierbare Kurve mit $c(0) = u$ und $c'(0) = v$. Dann gilt $f_j(c(t)) = y_j$ für alle $t \in (-\varepsilon, \varepsilon)$ und alle $j = 1, \ldots, p$ und damit

$$0 = (f_j \circ c)'(0) = Df_j(u)(v) = \langle \nabla f_j(u), v \rangle.$$

Zu (ii) bemerken wir, dass, da M ein reguläres Urbild ist, $Df(u) = (\nabla f_1(u), \ldots \nabla f_p(u))^T$ für alle $u \in M$ Rang p hat. Somit ist $\{\nabla f_1(u), \ldots \nabla f_p(u)\}$ für alle $u \in M$ linear unabhängig (der Spaltenrang ist gleich dem Zeilenrang). $\square$

Theorem 6.144 (Lokale Extrema unter Nebenbedingungen)**.** *Sei $U \Subset \mathbb{R}^n$ offen, und sei $f : U \to \mathbb{R}$ eine differenzierbare Funktion. Ferner sei $M \subset U \Subset \mathbb{R}^n$ eine Untermannigfaltigkeit des $\mathbb{R}^n$ und $u_o \in M$ ein Punkt, an welchem $f|_M$ ein lokales Extremum annimmt*[184]*. Dann gilt $\nabla f(u_o) \in N_{u_o} M$.*[185]

[184] d. h. Es gibt ein $\varepsilon > 0$, sodass für den offenen Ball $B_\varepsilon(u_o) \Subset \mathbb{R}^n$ gilt entweder

$$\forall u \in M \cap B_\varepsilon(u_o) : \; f(u) \leq f(u_o)$$

(lokales Maximum) oder

$$\forall u \in M \cap B_\varepsilon(u_o) : \; f(u) \geq f(u_o)$$

(lokales Minimum). Wenn $f|_M$ in u_o ein lokales Extremum hat, so muss f in u_o kein lokales Extremum haben. Die Einschränkung von f auf M wird auch *Nebenbedingung* genannt.

[185] d. h. die Komponente von $\nabla f(u_o)$ tangential zu M in u_o verschwindet. Falls $M = U$, so gilt

Ist M sogar ein reguläres Urbild einer stetig partiell differenzierbaren Abbildung

$$g := \begin{pmatrix} g_1 \\ \vdots \\ g_p \end{pmatrix} : U \to \mathbb{R}^p,$$

also $M = g^{-1}(\{y\})$ mit $y \in \mathrm{Bild}(g)$, dann gibt es eindeutig bestimmte Zahlen $\lambda_1, \ldots, \lambda_p \in \mathbb{R}$, sodass

$$\nabla f(u_o) = \sum_{j=1}^{p} \lambda_j \nabla g_j(u_o).$$

Die Zahlen $\lambda_1, \ldots, \lambda_p \in \mathbb{R}$ heißen Lagrange–Multiplikatoren.

Beweis. Indem wir ggf. zu $-f$ übergehen, können wir ohne Beschränkung der Allgemeinheit annehmen, dass $f|_M$ in u_o ein lokales Maximum hat, d. h. es gibt ein $\varepsilon > 0$, sodass für alle $u \in M \cap B_\varepsilon(u_o)$ gilt $f(u) \leq f(u_o)$.

Sei $v \in T_{u_o}M$ und $c : (-\delta, \delta) \to M \cap B_\varepsilon(u_o) \subset \mathbb{R}^n$ eine differenzierbare Kurve mit $c(0) = u_o$ und $c'(0) = v$. Dann gilt $f(c(t)) \leq f(c(0)) = f(u_o)$ für alle $t \in (-\delta, \delta)$. Somit hat die differenzierbare Funktion $f \circ c : (-\delta, \delta) \to \mathbb{R}$ in $t = 0$ ein lokales Maximum. Also gilt $0 = (f \circ c)'(0) = Df(u_o)(c'(0)) = \langle \nabla f(u_o), v\rangle$. Also gilt $\nabla f(u_o) \perp v$. Da $v \in T_{u_o}M$ beliebig war, folgt

$$\nabla f(u_o) \in N_{u_o}M.$$

Ist M das reguläre Urbild $M = g^{-1}(\{y\})$, so ist $\{\nabla g_1(u_o), \ldots, \nabla g_p(u_o)\}$ nach Lemma 6.143 eine Basis von $N_{u_o}M$. Damit gibt es eindeutig bestimmte $\lambda_1, \ldots, \lambda_p \in \mathbb{R}$ mit

$$\nabla f(u_o) = \sum_{j=1}^{p} \lambda_j \nabla g_j(u_o).$$

□

Satz 6.145. *Eine reelle symmetrische Matrix A hat einen reellen Eigenwert.*[186]

Beweis. Jede reelle symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ definiert eine quadratische Form auf $\mathbb{R}^n$, nämlich

$$Q_A : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto x^T A x.$$

Wir möchten nun die Richtung bestimmen, in welche Q_A ein Extremum aufweist. Dazu schränkten wir Q_A auf die kompakte Einheitssphäre

$$S_1 = \{x \in \mathbb{R}^n : \ \|x\|_{\mathrm{eukl}} = 1\}$$

insbesondere $\nabla f(u_o) = 0$.

[186] In der Vorlesung Lineare Algebra 2 haben Sie diesen Satz mit Hilfe des Fundamentalsatzes der Algebra bewiesen. Die Beweise des Fundamentalsatzes der Algebra verwenden aber immer auch Argumente aus der Topologie oder der Analysis.

ein. Da $Q_A|_{S_1}$ eine stetige polynomiale Funktion ist, gibt es ein $x_o \in S_1$, sodass $Q_A|_{S_1}$ in x_o ein Maximum und ein Minimum annimmt. Nun ist $S_1 = g^{-1}(1)$ das reguläre Urbild der Abbildung $g : \mathbb{R}^n \to \mathbb{R},\ x \mapsto \|x\|^2$ (siehe Beispiel 6.140) und wir erhalten $\nabla g(x_o) = 2x_o$.

Wie berechnen die Richtungsableitung von Q_A in x_o in Richtung $v \in \mathbb{R}^n$ mit $\|v\| = 1$:

$$\begin{aligned} D_v Q_A(x_o) &= \left.\frac{d}{dt}\right|_{t=0} DQ_A(x_o + tv) \\ &= \left.\frac{d}{dt}\right|_{t=0} (x_o + tv)^T A(x_o + tv) \\ &= \left.\frac{d}{dt}\right|_{t=0} (x_o^T A x_o + t x_o^T A v + t v^T A x_o + t^2 v^T A v) \\ &= x_o^T A v + v^T A x_o. \end{aligned}$$

Wegen $A^T = A$ erhalten wir $D_v Q_A(x_o) = x_o^T A v + v^T A x_o = 2x_o^T A v$. Damit ist Q_A insbesondere stetig partiell differenzierbar mit $J_{x_o} Q_a = 2x_o^T A \in \mathbb{R}^{1\times n}$, also

$$\nabla Q_A(x_o) = (J_{x_o} Q_A)^T = 2Ax_o \in \mathbb{R}^n.$$

Da $Q_A|_{S_1}$ in $x_o \in S_1$ ein Extremum annimmt, gilt mit Theorem 6.144:

$$\nabla Q_A(x_o) = 2Ax_o = \lambda \nabla g(x_o) = \lambda 2 x_o.$$

Somit ist $x_o \in \mathbb{R}^n$ ein Eigenwert von A zum Eigenwert λ. $\square$

$\mathcal{A}$ Exkurs: Fourierreihen

Ähnlich wie wir für nur „wenige" reelle Zahlen Namen haben, so können wir auch nur „wenige" Funktionen explizit benennen. Daher approximieren wir Funktionen oft. Ein Beispiel hierfür sind die Ihnen bereits bekannten Taylorpolynome. Für periodische Funktionen, hat eine polynomiale Approximation den Nachteil, dass die wichtige Information der Periodizität bei Approximation durch Polynome verloren geht, da nicht konstante Polynome nie periodisch sind.

Es gibt jedoch zwei Ihnen sehr gut bekannte periodische Funktionen, nämlich Sinus und Kosinus. In diesem Abschnitt beschreiben wir, wie man periodische Funktionen durch Summen von Sinus- und Kosinusschwingungen verschiedener ganzzahliger Frequenzen und unterschiedlicher Amplituden approximieren kann. Im Gegensatz zum Taylorpolynom benötigen wir keine Voraussetzungen an die Differenzierbarkeit, sondern an die Integrierbarkeit, was natürlich viel schwächer ist.

Diese Idee, periodische Funktionen durch (endliche oder unendliche) Linearkombinationen von $\cos(kx)$ und $\sin(kx)$, $k \in \mathbb{N}$ (trigonometrische Polynome oder trigonometrische Reihen) zu approximieren ist relativ alt und hat noch heute viele Anwendungen in der Physik (Analyse von Schwingungen) und den Ingenieurwissenschaften (Signalverarbeitung).

Bereits im 18. Jahrhundert verwendeten Daniel Bernoulli, Leonard Euler u. A. trigonometrische Reihen zum Studium der partiellen Differentialgleichung der „schwingende Saite"

$$\frac{\partial^2 f}{\partial t^2} = c^2 \frac{\partial^2 f}{\partial x^2}$$

beim Ansatz $f(x,t) = u(t)v(x)$. Im Buch ≪*Théorie analytique de la chaleur*≫ (Fermin Didot, Paris 1822) von Joseph Fourier finden sich bereits systematische mathematische Untersuchungen über trigonometrische Reihen. Die Konvergenz trigonometrischer Reihen wurde dann im 19. und 20. Jahrhundert von Johann Peter Gustav Lejeune Dirichlet, Friedrich Wilhelm Bessel, Nikolai Nikolajewitsch Lusin, Lennart Carleson u. A. weiter erforscht.

Wir folgen in diesem Exkurs eng der Darstellung in [Fo1, § 23], verweisen aber auch auf [K1, Kap. 17] und [H2, Kap. XVII].

Definition A.1. Eine Funktion $f : \mathbb{R} \to \mathbb{C}$ heißt *periodisch*, wenn es ein $P > 0$ gibt, sodass für alle $x \in \mathbb{R}$ gilt:

$$f(x+P) = f(x)$$

Die Zahl P heißt *Periode* von f.

Bemerkung A.2. Hat $f : \mathbb{R} \to \mathbb{C}$ die Periode P, dann hat

$$\tilde{f} : \mathbb{R} \to \mathbb{C}, \quad x \mapsto \tilde{f}(x) := f\left(\frac{P}{2\pi}x\right)$$

die Periode 2π und umgekehrt (Übung). Wir werden unsere Betrachtungen daher auf (2π)-periodische Funktionen einschränken.

Definition A.3 (Trigonometrische Polynome)**.** Eine Funktion

$$p : \mathbb{R} \to \mathbb{C},\ x \mapsto \frac{a_0}{2} + \sum_{k=1}^{n} \big(a_k \cos(kx) + b_k \sin(kx)\big),\ a_k, b_k \in \mathbb{C}$$

heißt *trigonometrisches Polynom*.

Bemerkung A.4. Aus den Additionstheoremen erhält man die *Orthogonalitätsrelationen*:

$$\begin{aligned}
\int_0^{2\pi} \sin(kx)\sin(lx)dx &= \int_0^{2\pi} \cos(kx)\cos(lx)dx = \begin{cases} 0, & k \neq l \\ \pi, & k = l \geq 1 \end{cases} \\
\int_0^{2\pi} \sin(kx)\cos(lx)dx &= 0 \quad \text{für alle } k, l \in \mathbb{N}
\end{aligned}$$

(siehe auch Übungsaufgabe 10, Übungsblatt 3). Für ein trigonometrisches Polynom $p(x) = \frac{a_0}{2} + \sum\limits_{k=1}^{n} \big(a_k \cos(kx) + b_k \sin(kx)\big)$ folgt dann (Übungsaufgabe)

$$\begin{aligned}
a_k &= \frac{1}{\pi}\int_0^{2\pi} p(x)\cos(kx)dx, \quad k = 0, 1, ..., n \\
b_k &= \frac{1}{\pi}\int_0^{2\pi} p(x)\sin(kx)dx, \quad k = 1, ..., n.
\end{aligned}$$

Bemerkung A.5 (Komplexe Schreibweise trigonometrischer Polynome)**.** Mit

$$\begin{aligned}
\cos(kx) &= \Re(e^{ikx}) = \tfrac{1}{2}\big(e^{ikx} + e^{-ikx}\big) \\
\sin(kx) &= \Im(e^{ikx}) = \tfrac{1}{2i}\big(e^{ikx} - e^{-ikx}\big)
\end{aligned}$$

kann man ein trigonometrisches Polynom $p(x) = \frac{a_0}{2} + \sum\limits_{k=1}^{n} \big(a_k \cos(kx) + b_k \sin(kx)\big)$ auch schreiben als

$$p(x) = \sum_{k=-n}^{n} c_k e^{ikx}$$

mit $c_0 = \frac{a_0}{2}, \quad c_k = \frac{1}{2}(a_k - ib_k), \quad c_{-k} = \frac{1}{2}(a_k + ib_k), \quad k \in \mathbb{N}$.
Diese komplexe Schreibweise ist oft rechnerisch elegant und wird daher auch in der Elektrotechnik und in anderen technischen Disziplinen gerne verwendet. Es gilt offenbar die Orthogonalitätsrelation

$$\int\limits_0^{2\pi} e^{ikx} dx = \begin{cases} 0, & k \neq 0 \\ 2\pi, & k = 0 \end{cases}$$

und damit

$$c_k = \frac{1}{2\pi} \int\limits_0^{2\pi} p(x) e^{-ikx} dx, \; k = 0, \pm 1, ..., \pm n,$$

also

$$p(x) = \sum_{k=-n}^{n} c_k e^{ikx} \text{ mit } c_k = \frac{1}{2\pi} \int_0^{2\pi} p(x) e^{-ikx} dx, \; k = 0, \pm 1, ..., \pm n.$$

Aus der komplexen Darstellung erhalten wir mit

$$\begin{aligned} a_0 &= 2c_0, \\ a_k &= c_k + c_{-k}, \\ b_k &= i(c_k - c_{-k}), \quad k = 1, ..., n \end{aligned}$$

die trigonometrische Schreibweise zurück. Die Menge aller trigonometrischen Polynome in reeller Schreibweise und auch die Menge aller trigonometrischen Polynome in komplexer Schreibweise bildet offenbar jeweils einen reellen bzw. komplexen Vektorraum.

Definition A.6 (Fourierpolynom und Fourierreihe)**.** Sei $f : \mathbb{R} \to \mathbb{C}$ eine 2π-periodische und auf $[0, 2\pi]$ Riemann–integrierbar Funktion, dann heißt

$$\begin{aligned} \mathcal{F}_n^f(x) &:= \frac{a_0}{2} + \sum_{k=1}^{n} \big(a_k \cos(kx) + b_k \sin(kx)\big) \\ &= \sum_{k=-n}^{n} c_k e^{ikx} \end{aligned}$$

mit

$$\begin{aligned} a_k &:= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx \quad \text{und} \quad b_k := \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx, \qquad \text{bzw.} \\ c_k &:= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx \end{aligned}$$

Fourierpolynom[187] von f vom Grad n. Die formale Reihe

$$\begin{aligned}\mathcal{F}^f(x) &:= \frac{a_0}{2} + \sum_{k=1}^{\infty} \big(a_k \cos(kx) + b_k \sin(kx)\big) \\ &= \sum_{k=-\infty}^{\infty} c_k e^{ikx}\end{aligned}$$

mit den oben abgegebenen Koeffizienten a_k und b_k bzw. c_k wird *Fourierreihe* von f genannt.

Notation A.7.

$$\hat{f}(k) := \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx = c_k$$

und damit

$$\mathcal{F}_n^f(x) := \sum_{k=-n}^{n} \hat{f}(k) e^{ikx} \quad \text{und} \quad \mathcal{F}^f(x) := \sum_{k=-\infty}^{\infty} \hat{f}(k) e^{ikx}$$

Beobachtung A.8 (Rechenregeln)**.** *Die Linearität des Integrals in der Definition von $\hat{f}(k)$ impliziert*

$$(\widehat{\alpha f + g})(k) = \alpha \hat{f}(k) + \hat{g}(k)$$

und damit

$$\mathcal{F}_n^{\alpha f + g} = \alpha \mathcal{F}_n^f + \mathcal{F}_n^g \quad \textit{und} \quad \mathcal{F}^{\alpha f + g} = \alpha \mathcal{F}^f + \mathcal{F}^g$$

für zwei 2π-periodische Funktionen $f, g : \mathbb{R} \to \mathbb{C}$, die auf $[0, 2\pi]$ Riemann–integrierbar sind und $\alpha \in \mathbb{C}$.

Aufgabe A.9. Zeigen Sie:

(a) Ist f gerade, d.h. $f(-x) = f(x)$, so ist $b_k = 0$.

(b) Ist f ungerade, d.h. $f(-x) = -f(x)$, so ist $a_k = 0$.

Beispiel A.10 (Fourierpolynome der Rechteckschwingung)**.** Sei f die periodische Fortsetzung von

$$f(x) = \begin{cases} 1, & 0 \leq x < \pi \\ -1, & \pi \leq x < 2\pi \end{cases},$$

dann gilt

$$\begin{aligned} a_k &= 0 \quad (f \text{ ungerade}), \\ b_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx = \frac{2}{\pi} \int_0^{\pi} \sin(kx) dx = \begin{cases} 0, & k \text{ gerade} \\ \frac{4}{k\pi}, & k \text{ ungerade} \end{cases} \end{aligned}$$

und damit

$$\mathcal{F}_{2n-1}^f(x) = \frac{4}{\pi} \sum_{k=1}^{n} \frac{1}{2k-1} \sin((2k-1)x).$$

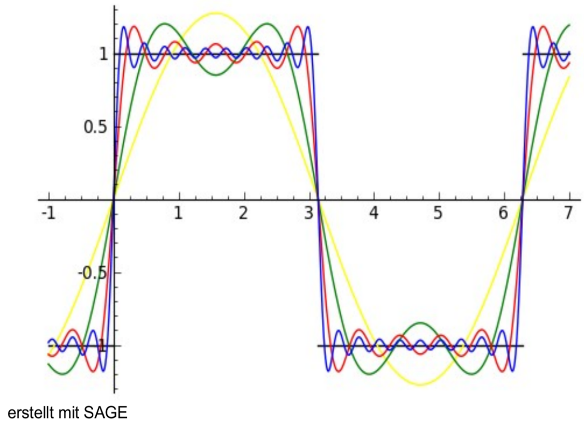

Abbildung A.1.: $\mathcal{F}_1^f$, $\mathcal{F}_3^f$, $\mathcal{F}_9^f$, $\mathcal{F}_{19}^f$ (Graphik erstellt mit SAGE)

Wir beobachten, dass die Oszillationen (Amplituden) in der Nähe der Sprungstellen höher werden und somit die Approximation von f durch ihre Fourierpolynome in der Nähe der Sprungstellen schlechter wird. Dieses Phänomen bleibt grundsätzlich auch bei Erhöhung des Grades des Fourierpolynoms bestehen, es lokalisiert sich jedoch immer mehr bei den Sprungstellen. Das beschriebene Verhalten der Fourierpolynome wird *Gibbssches*[188] *Phänomen* genannt. Anschaulich gesprochen müssen die stetigen Fourierpolynome die Unstetigkeitsstellen von f „überspringen", was zu den beobachteten höheren Oszillation nahe der Sprungsstellen von f führt. Dieses Gibbssche Phänomen tritt grundsätzlich bei Fourierpolynomen von stückweise stetig differenzierbaren Funktionen in deren Unstetigkeitsstellen auf.

In diesem Beispiel nehmen die Fourierpolynome an den Sprungstellen das arithmetische Mittel der rechts- und linksseitigen Grenzwerte von f in der Sprungstelle an (vgl. dazu [K1, Abschn. 17.4].

Beispiel A.11 (Fourierpolynome der Kippschwingung)**.** Sei f die periodische Fortsetzung von

$$f(x) = \begin{cases} x, & 0 \leq x < \pi \\ x - 2\pi, & \pi \leq x < 2\pi \end{cases} .$$

Dann gilt

$$\begin{aligned} a_k &= 0 \quad (\text{da } f \text{ ungerade}), \\ b_k &= \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(kx)dx = \frac{2}{\pi}\int_0^{\pi} x\sin(kx)dx = \overset{\text{part. Int.}}{\ldots} = (-1)^{n+1}\frac{2}{n} \end{aligned}$$

und damit

$$\mathcal{F}_n^f(x) = 2\sum_{k=1}^{n}\frac{(-1)^{k+1}}{k}\sin(kx) = 2\left(\sin(x) - \frac{\sin(2x)}{2} \pm \ldots \pm \frac{\sin(nx)}{n}\right).$$

[187] Jean-Baptiste-Joseph Fourier (1768–1830), französischer Physiker und Mathematiker
[188] Josiah Willard Gibbs (1839–1903), amerikanischer Physiker

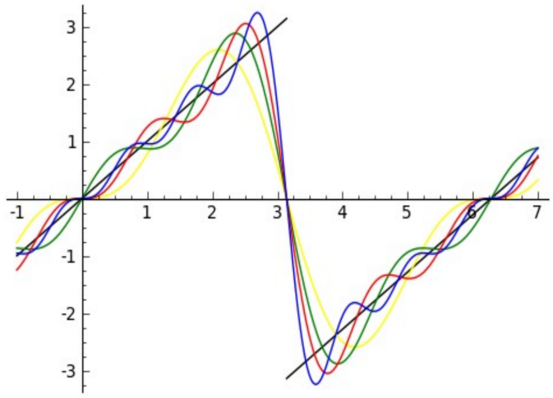

Abbildung A.2.: $\mathcal{F}_2^f$, $\mathcal{F}_3^f$, $\mathcal{F}_4^f$, $\mathcal{F}_6^f$ (Graphik erstellt mit SAGE)

Bemerkung A.12 (Hilfsmittel aus der linearen Algebra)**.** Wir betrachten den $\mathbb{C}$-Vektorraum

$$V := \{f : \mathbb{R} \to \mathbb{C} : \ f \ 2\pi\text{-periodisch}, \ f|_{[0,2\pi]} \text{ Riemann–integrierbar}\}$$

versehen mit der positiv–semidefiniten Hermiteschen Sesquilinearform

$$\langle f, g\rangle_{L^2} := \frac{1}{2\pi}\int_0^{2\pi} f(x)\overline{g(x)}dx, \qquad f, g \in V,$$

welche auch L^2–*Skalarprodukt* genannt wird. Es gilt also

(i) $\langle f, g\rangle_{L^2} = \overline{\langle g, f\rangle_{L^2}}$

(ii) $\langle \lambda f + f', g\rangle_{L^2} = \lambda\langle f, g\rangle_{L^2} + \langle f', g\rangle_{L^2}$ und $\langle f, \lambda g + g'\rangle_{L^2} = \overline{\lambda}\langle f, g\rangle_{L^2} + \langle f, g'\rangle_{L^2}$

(iii) $\langle f, f\rangle_{L^2} \geq 0$

Für positiv–semidefinite Hermitesche Sesquilinearform gilt die Ungleichung von Cauchy-Schwarz:

$$|\langle f, g\rangle_{L^2}|^2 \leq \langle f, f\rangle_{L^2}\langle g, g\rangle_{L^2}.$$

Das L^2-Skalarprodukt induziert auf V die L_2–*Halbnorm*:

$$\|f\|_{L^2} := \sqrt{\langle f, f\rangle_{L^2}} := \sqrt{\frac{1}{2\pi}\int_0^{2\pi} f(x)\overline{f(x)}dx} = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2 dx} \geq 0.$$

Es gilt also

(i) $\|\lambda f\|_{L^2} = |\lambda| \cdot \|f\|_{L^2}$

(ii) Dreiecks-Ungleichung: $\|f + g\|_{L^2} \leq \|f\|_{L^2} + \|g\|_{L^2}$

Proposition A.13. *Sei $f : \mathbb{R} \to \mathbb{C}$ eine 2π-periodische Funktio und $f|_{[0,2\pi]}$ Riemann-integrierbar. Sei $\mathcal{T}_n(x) = \sum\limits_{k=-n}^{n} t_k e^{ikx}$ ein trigonometrisches Polynom von Grad n mit $\mathcal{T}_n \neq \mathcal{F}_n^f$. Dann gilt:*

(i) $\|f - \mathcal{F}_n^f\|_{L^2} < \|f - \mathcal{T}_n\|_{L^2}$,
d. h. das Fourierpolynom von f von Grad n ist das eindeutige trigonometrische Polynom von Grad n, welches f in der L^2-Halbnorm[189] am besten approximiert.

(ii) $\|f - \mathcal{F}_n^f\|_{L^2}^2 = \|f\|_{L^2}^2 - \sum\limits_{k=-n}^{n} |\hat{f}(k)|^2$

Notation A.14. Die Funktionen

$$e_k : \mathbb{R} \to \mathbb{C},\ x \mapsto e^{ikx}, \qquad k \in \mathbb{Z}$$

bilden ein Orthonormalsystem von Funktionen in V, denn es gilt:

$$\langle e_k, e_l \rangle_{L^2} = \frac{1}{2\pi} \int\limits_0^{2\pi} e^{ikx} e^{-ilx} dx = \frac{1}{2\pi} \int\limits_0^{2\pi} e^{i(k-l)x} dx = \delta_{kl} = \begin{cases} 1, & k = l \\ 0, & k \neq l \end{cases}$$

Wir erhalten in dieser Notation:

$$\hat{f}(k) = \frac{1}{2\pi} \int\limits_0^{2\pi} f(x) e^{-ikx} dx = \frac{1}{2\pi} \int\limits_0^{2\pi} f(x) \overline{e^{ikx}} dx = \langle f, e_k \rangle_{L^2}.$$

Beweis. Sei $\mathcal{T}_n = \sum\limits_{k=-n}^{n} t_k e_k$ ein beliebiges trigonometrisches Polynom von Grad n, dann gilt:

$$\begin{aligned}
\|f - \mathcal{T}_n\|_{L^2}^2 &= \langle f - \mathcal{T}_n, f - \mathcal{T}_n \rangle_{L^2} \\
&= \langle f, f \rangle_{L^2} - \langle \mathcal{T}_n, f \rangle_{L^2} - \langle f, \mathcal{T}_n \rangle_{L^2} + \langle \mathcal{T}_n, \mathcal{T}_n \rangle_{L^2} \\
&= \|f\|_{L^2}^2 - \sum_{k=-n}^{n} t_k \underbrace{\langle e_k, f \rangle_{L^2}}_{=\overline{\hat{f}(k)}} - \sum_{k=-n}^{n} \overline{t_k} \underbrace{\langle f, e_k \rangle_{L^2}}_{=\hat{f}(k)} + \sum_{k,l=-n}^{n} t_k \overline{t_l} \underbrace{\langle e_k, e_l \rangle_{L^2}}_{=\delta_{kl}} \\
&= \|f\|_{L^2}^2 + \underbrace{\sum_{k=-n}^{n} |t_k|^2 - \sum_{k=-n}^{n} t_k \overline{\hat{f}(k)} - \sum_{k=-n}^{n} \overline{t_k} \hat{f}(k)}_{= \sum_{k=-n}^{n} |t_k - \hat{f}(k)|^2 - \sum_{k=-n}^{n} |\hat{f}(k)|^2} \\
&= \|f\|_{L^2}^2 - \sum_{k=-n}^{n} |\hat{f}(k)|^2 + \sum_{k=-n}^{n} |t_k - \hat{f}(k)|^2
\end{aligned}$$

[189] Anschaulich bedeutet das „im Sinne der kleinsten Summe der Quadrate der Fehler (Abweichungen)“

Damit ist $\|f - \mathcal{T}_n\|_{L^2}^2$ genau dann minimal, wenn $\sum\limits_{k=-n}^{n} |t_k - \hat{f}(k)|^2 = 0$, also $t_k = \hat{f}(k)$ für alle $k = -n, ..., 0, ..., n$ gilt. Das bedeutet aber $\mathcal{T} = \mathcal{F}_n^f$. □

Nach Proposition A.13 gilt für alle n

$$0 \leq \|f - \mathcal{F}_n^f\|_{L^2}^2 = \|f\|_{L^2}^2 - \sum_{k=-n}^{n} |\hat{f}(k)|^2$$

und damit

$$\sum_{k=-n}^{n} |\hat{f}(k)|^2 \leq \|f\|_{L^2}^2.$$

Daraus folgt:

Korollar A.15 (Besselsche[190] Ungleichung)**.** *Für alle 2π-periodischen Funktion, sodass $f|_{[0,2\pi]}$ Riemann-integrierbar ist, gilt*

$$\sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2 \leq \|f\|_{L^2}^2 \quad \text{(Besselsche Ungleichung).}$$

Die Fourierpolynome von f konvergieren genau dann in der L^2–Halbnorm (oder im quadratischen Mittel*) gegen f, d.h. $\lim_{n\to\infty} \|f - \mathcal{F}_n^f\|_{L^2} = 0$, wenn die* Parsevalsche[191] Gleichung

$$\sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2 = \|f\|_{L^2}^2$$

gilt.

Lemma A.16. *Ist $f : \mathbb{R} \to \mathbb{R}$ eine 2π-periodische Funktion, sodass $f|_{[0,2\pi]}$ eine Treppenfunktion ist, dann gilt*

$$\lim_{n\to\infty} \|f - \mathcal{F}_n^f\|_{L^2} = 0.$$

Beweis. (i) *Behauptung: Die Aussage ist, wenn*

$$f|_{[0,2\pi]}(x) = \begin{cases} 0, & \textit{wenn} \quad 0 \leq x < a \\ 1, & \textit{wenn} \quad a \leq x \leq b \\ 0, & \textit{wenn} \quad b < x < 2\pi \end{cases},$$

mit $0 \leq a \leq b \leq 2\pi$ gilt.

Wir erhalten als Koeffizienten der Fourierreihe von f :

$$\begin{aligned} \hat{f}(0) &= \frac{b-a}{2\pi}, \\ \hat{f}(k) &= \frac{1}{2\pi}\int_a^b \exp(-ikx)dx = \frac{i}{2\pi k}\big(\exp(-ikb) - \exp(-ika)\big), \quad k \in \mathbb{Z} \setminus \{0\} \end{aligned}$$

[190] Friedrich Wilhelm Bessel (1784–1846), deutscher Naturwissenschaftler
[191] Marc-Antoine Parseval (1755–1836), französischer Mathematiker

und damit für $k \neq 0$: $|\hat{f}(k)|^2 = \frac{1-\cos(k(b-a))}{2\pi^2 k^2}$. Wir erhalten also

$$\begin{aligned}
\sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2 &= \frac{(b-a)^2}{4\pi^2} + \sum_{k=1}^{\infty} \frac{1-\cos(k(b-a))}{\pi^2 k^2} \\
&= \frac{(b-a)^2}{4\pi^2} + \frac{1}{\pi^2} \underbrace{\sum_{k=1}^{\infty} \frac{1}{k^2}}_{=\frac{\pi^2}{6}} - \frac{1}{\pi^2} \underbrace{\sum_{k=1}^{\infty} \frac{\cos(k(b-a))}{k^2}}_{\substack{= \frac{((b-a)-\pi)^2}{4} - \frac{\pi^2}{12} \\ \text{siehe [Fo1, Bsp. 21.7]}}} \\
&= \frac{b-a}{2\pi} \\
&= \|f\|_{L^2}^2.
\end{aligned}$$

Nach Korollar A.15 gilt dann $\lim\limits_{n\to\infty} \|f - \mathcal{F}_n^f\|_{L^2} = 0$.

(ii) Ist $f|_{[0,2\pi]}$ eine Treppenfunktion, so gibt es konstanten $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ und (2π)-periodische Funktionen $f_1, \ldots, f_m$ von der in Teil (i) behandelten Art, sodass

$$f = \sum_{j=1}^{m} \alpha_j f_j.$$

Somit gilt

$$\mathcal{F}_n^f = \sum_{j=1}^{m} \alpha_j \mathcal{F}_n^{f_j}$$

und wir erhalten

$$\left\|f - \mathcal{F}_n^f\right\|_{L^2} = \left\| \sum_{j=1}^{m} \alpha_j (f_j - \mathcal{F}_n^{f_j}) \right\|_{L^2} \leq \sum_{j=1}^{m} |\alpha_j| \cdot \|f_j - \mathcal{F}_n^{f_j}\|_{L^2}.$$

Nach Teil (i) erhalten wir $\lim\limits_{n\to\infty} \|f - \mathcal{F}_n^f\|_{L^2} = 0$.

$\square$

Theorem A.17. *Ist $f : \mathbb{R} \to \mathbb{R}$ eine 2π-periodische Funktion, sodass $f|_{[0,2\pi]}$ Riemann–integrierbar ist, dann gilt*

$$\lim_{n\to\infty} \|f - \mathcal{F}_n^f\|_{L^2} = 0.$$

Beweis. Nach ggf. entsprechender Stauchung können wir $|f(x)| \leq 1$ für alle $x \in \mathbb{R}$ annehmen. Sei nun $\varepsilon > 0$. Dann gibt es periodische Funktionen $\phi_1, \phi_2 : \mathbb{R} \to \mathbb{R}$, sodass gilt:

- $\phi_1|_{[0,2\pi]}, \phi_2|_{[0,2\pi]}$ sind Treppenfunktionen,
- $-1 \leq \phi_1 \leq f \leq \phi_2 \leq 1$,

- $\int_0^{2\pi} \phi_2(x)dx - \int_0^{2\pi} \phi_1(x)dx \leq \frac{\pi}{4}\varepsilon^2$

(siehe Satz 5.16). Setzen wir $g := f - \phi_1$, dann erhalten wir

$$\begin{aligned} |g(x)|^2 &= |f(x) - \phi_1(x)|^2 \leq |\phi_2(x) - \phi_1(x)|^2 \\ &= \underbrace{(\phi_2(x) - \phi_1(x))}_{\leq 2} \cdot (\phi_2(x) - \phi_1(x)) \\ &\leq 2(\phi_2(x) - \phi_1(x)) \end{aligned}$$

und somit

$$\|g\|_{L^2}^2 = \frac{1}{2\pi} \int_0^{2\pi} |g(x)|^2 dx \leq \frac{1}{\pi} \int_0^{2\pi} (\phi_2(x) - \phi_1(x))dx \leq \frac{\varepsilon^2}{4}.$$

Nach Lemma A.16 gibt es ein $N \in \mathbb{N}$, sodass für alle $n \geq N$ gilt

$$\|\phi_1 - \mathcal{F}_n^{\phi_1}\|_{L^2} \leq \frac{\varepsilon}{2}.$$

Nach Proposition A.13 gilt

$$\|g - \mathcal{F}_n^g\|_{L^2}^2 \leq \|g\|_{L^2}^2 \leq \frac{\varepsilon^2}{4}.$$

Mit $\mathcal{F}_n^f = \mathcal{F}_n^g + \mathcal{F}_n^{\phi_1}$ folgt für $n \geq N$:

$$\|f - \mathcal{F}_n^f\|_{L^2} \leq \|\phi_1 - \mathcal{F}_n^{\phi_1}\|_{L^2} + \|g - \mathcal{F}_n^g\|_{L^2} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

B Lineare Abbildungen

Für diesen Anhang, welcher meinem Skript zur Linearen Algebra 1 vom SS 2011 und dem Skript zum Brückenkurs im SS 2013 entnommen ist, verweisen wir insbesondere auf [He1].

Nach konstanten Abbildungen sind lineare Abbildungen in einem gewissen Sinne in einem gewissen Sinne die „einfachsten“ Abbildung. Daher werden sie in der Mathematik und ihren Anwendungen vielfach verwendet. Die Definitions- und Zielmengen linearer Abbildungen sind Vektorräume. Sie kennen bereits einige wichtige Beispiele von Vektorräumen aus der Schule:

- $\mathbb{R}^0 := \{0\}$,
- $\mathbb{R}^1 = \mathbb{R}$, die Menge der reellen Zahlen,
- $\mathbb{R}^2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} ;\ x_1, x_2 \in \mathbb{R} \right\}$,
- $\mathbb{R}^3 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} ;\ x_1, x_2, x_3 \in \mathbb{R} \right\}$.

Eine wichtige gemeinsame Eigenschaft dieser drei Mengen ist, dass man deren Elemente miteinander addieren kann (Vektoraddition) und auch jedes Element mit einer Zahl λ (Streckungsfaktor) multiplizieren kann, wobei die üblichen Rechenregeln gelten. Für Elemente aus $\mathbb{R}^3$ geht das komponentenweise:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix} \quad \text{und} \quad \lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \end{pmatrix}.$$

Wir können diese komponentenweisen Rechenoperationen einfach auf die Menge

$$\mathbb{R}^n := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} ;\ x_1, ..., x_n \in \mathbb{R} \right\}$$

aller n-Tupel reeller Zahlen verallgemeinern, wobei nun n eine beliebige positive natürliche Zahl ist:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \text{und} \quad \lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

Auch hierbei gelten die üblichen Rechenregeln. Mengen mit derartigen Rechenoperationen, welche den üblichen Regeln gehorchen, heißen in der Mathematik (reelle) *Vektorräume*.

Ein ausgezeichnetes Element von $\mathbb{R}^n$ ist $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, welches wir in Zukunft kurz mit $0 \in \mathbb{R}^n$ bezeichnen werden.

Nun kennen Sie schon einige wichtige Vektorräume, nämlich für jede natürliche Zahl n den Vektorraum $\mathbb{R}^n$.

Definition B.1. Eine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ heißt *linear*, wenn f mit der Addition sowie mit der Multiplikation mit reellen Zahlen verträglich ist. Genauer muss gelten:

(i) $f(x + y) = f(x) + f(y)$ für alle $x, y \in \mathbb{R}^n$ und

(ii) $f(\lambda \cdot x) = \lambda \cdot f(x)$ für alle $\lambda \in \mathbb{R}$ und alle $x \in \mathbb{R}^n$.

Aufgabe B.2. Seien n und m positive natürliche Zahlen und $f : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung. Zeigen Sie folgende Aussage: Für jede natürliche Zahl $k \geq 2$ und k beliebige Elemente $v_1, ..., v_k \in \mathbb{R}^n$ gilt

$$f(v_1 + ... + v_k) = f(v_1) + ... + f(v_k),$$

oder in Summennotation

$$f\left(\sum_{j=1}^{k} v_j\right) = \sum_{j=1}^{k} f(v_j).$$

Satz B.3. *Eine Abbildung $f : \mathbb{R} \to \mathbb{R}$ ist genau dann linear, wenn es ein $a \in \mathbb{R}$ gibt, sodass $f(x) = a \cdot x$ für alle $x \in \mathbb{R}$.*

Beweis. Zunächst rechnet man leicht nach, dass für beliebiges $a \in \mathbb{R}$ die Abbildung $f : \mathbb{R} \to \mathbb{R},\ x \mapsto a \cdot x$ linear ist.
Sei andererseits $f : \mathbb{R} \to \mathbb{R}$ linear, so gilt $f(x) = f(x \cdot 1) = x \cdot f(1) = f(1) \cdot x = a \cdot x$ mit $a = f(1)$. $\square$

Unser Ziel ist es nun lineare Abbildungen zwischen $\mathbb{R}^n$ und $\mathbb{R}^m$ zu studieren. Betrachtet man obigen Beweis, so möchte man in $\mathbb{R}^n$ Elemente finden, welche die Rolle der $1 \in \mathbb{R}^1$ übernehmen können.

Definition B.4. Die n Vektoren $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, ..., e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ in $\mathbb{R}^n$ bilden die *Standardbasis* des $\mathbb{R}^n$. Hierbei ist e_j der Vektor im $\mathbb{R}^n$ dessen Komponenten alle verschwinden, bis auf die Komponente in der j-ten Zeile, welche gleich eins ist.

Man beachte, dass die Anzahl der Komponenten der Vektoren durch die Notation e_j alleine nicht klar wird, denn $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \mathbb{R}^2$ und $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{R}^3$.

Jedes beliebige Element $x \in \mathbb{R}^n$ kann man eindeutig als *Linearkombination* der Elemente $e_1, ..., e_n$ schreiben, denn es gilt:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + ... + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \sum_{j=1}^{n} x_j \cdot e_j.$$

Satz B.5. *Eine Abbildung $f : \mathbb{R}^n \to \mathbb{R}$ ist genau dann linear, wenn es reelle Konstanten $a_1, ..., a_n$ gibt, sodass*

$$f(x) = a_1 x_1 + .. + a_n x_n = \sum_{j=1}^{n} a_j x_j$$

für alle $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ gilt.

Beweis. Als Übungsaufgabe rechne man nach, dass die Abbildung

$$f : \mathbb{R}^n \to \mathbb{R},\ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto a_1 x_1 + .. + a_n x_n = \sum_{j=1}^{n} a_j x_j$$

linear ist.
Sei andererseits $f : \mathbb{R}^n \to \mathbb{R}$ linear, so gilt nach Aufgabe B.2 und der Definition der Linearität $f(x) = f\left(\sum_{j=1}^{n} x_j \cdot e_j\right) = \sum_{j=1}^{n} f(x_j \cdot e_j) = \sum_{j=1}^{n} x_j \cdot \underbrace{f(e_j)}_{=:a_j} = \sum_{j=1}^{n} a_j \cdot x_j$. □

Jede Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ besteht aus m Komponentenfunktion $f_1, ..., f_m : \mathbb{R}^n \to \mathbb{R}$ und wir schreiben:

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : \mathbb{R}^n \to \mathbb{R}^m, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} = \begin{pmatrix} f_1(x_1, ..., x_n) \\ \vdots \\ f_m(x_1, ..., x_n) \end{pmatrix},$$

wobei wir $f_j(x_1, ..., x_n)$ anstelle von $f_j\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right)$ schreiben.

Lemma B.6. *Eine Abbildung* $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : \mathbb{R}^n \to \mathbb{R}^m$ *ist genau dann linear, wenn alle Komponentenfunktionen* $f_1, ..., f_m : \mathbb{R}^n \to \mathbb{R}$ *linear sind.*

Beweis. (i) Folgende Aussagen sind äquivalent:

- $f(x+y) = f(x) + f(y)$ für alle $x, y \in \mathbb{R}^n$,
- $\begin{pmatrix} f_1(x+y) \\ \vdots \\ f_m(x+y) \end{pmatrix} = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} + \begin{pmatrix} f_1(y) \\ \vdots \\ f_m(y) \end{pmatrix} = \begin{pmatrix} f_1(x) + f_1(y) \\ \vdots \\ f_m(x) + f_m(y) \end{pmatrix} \quad \forall x, y \in \mathbb{R}^n$,
- $f_j(x+y) = f_j(x) + f_j(y)$ für alle $x, y \in \mathbb{R}^n$ und alle $j \in \{1, ..., m\}$.

(ii) Folgende Aussagen sind äquivalent:

- $f(\lambda \cdot x) = \lambda \cdot f(x)$ für alle $x \in \mathbb{R}^n$ und alle $\lambda \in \mathbb{R}$,
- $\begin{pmatrix} f_1(\lambda \cdot x) \\ \vdots \\ f_m(\lambda \cdot x) \end{pmatrix} = \lambda \cdot \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$ für alle $x \in \mathbb{R}^n$ und alle $\lambda \in \mathbb{R}$,
- $f_j(\lambda \cdot x) = \lambda \cdot f_j(x)$ für alle $x \in \mathbb{R}^n$, alle $\lambda \in \mathbb{R}$ und alle $j \in \{1, ..., m\}$.

$\square$

Nach Satz B.5 können wir die linearen Komponentenfunktionen f_j von f als

$$f_j(x_1, ..., x_n) = a_{j1}x_1 + a_{j2}x_2 + ... + a_{jn}x_n = \sum_{k=1}^{n} a_{jk}x_k$$

mit reellen Konstanten $a_{j1}, ..., a_{jn}$ schreiben. Somit erhalten wir

Theorem B.7 (Beschreibung aller linearen Abbildungen)**.** *Eine Abbildung* $f : \mathbb{R}^n \to \mathbb{R}^m$ *ist genau dann linear, wenn es reelle Konstanten* $a_{11}, ..., a_{1n}, a_{21} ..., a_{2n}, ..., a_{m1}, ..., a_{mn}$ *gibt, sodass*

$$f(x) = \begin{pmatrix} a_{11}x_1 + .. + a_{1n}x_n \\ a_{21}x_1 + .. + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + .. + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} a_{1k}x_k \\ \sum_{k=1}^{n} a_{2k}x_k \\ \vdots \\ \sum_{k=1}^{n} a_{mk}x_k \end{pmatrix} \tag{B.1}$$

für alle $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ *gilt.*

Zur Übersichtlichkeit schreiben wir diese Konstanten in eine *Matrix*

$$A_f := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = (a_{ij}),$$

das ist ein rechteckiges Zahlenschema. In diesem Fall hat A_f m Zeilen und n Spalten, hat also $m \cdot n$ Einträge. Wir nennen A_f eine reelle $(m \times n)$-*Matrix*.

Nach Satz B.7 gehört zu jeder linearen Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ eine reelle $(m \times n)$-Matrix A_f. Andererseits gehört auch zu jeder reellen $(m \times n)$-*Matrix*

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

die lineare Abbildung

$$f_A : \mathbb{R}^n \to \mathbb{R}^m, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} \sum\limits_{k=1}^{n} a_{1k}x_k \\ \sum\limits_{k=1}^{n} a_{2k}x_k \\ \vdots \\ \sum\limits_{k=1}^{n} a_{mk}x_k \end{pmatrix}.$$

Sollten Sie bereits die Matrixmultiplikation kennen, so werden Sie erkannt haben, dass gilt:

$$f_A : \mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto A \cdot x.$$

Diese Beziehung zwischen linearen Abbildungen und Matrizen wird im folgenden Satz noch einmal verdeutlicht:

Theorem B.8. *Die j-te Spalte der zur linearen Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ gehörenden Matrix A_f ist gleich $f(e_j)$ $(j = 1, ..., n)$.*

Beweis. Aus Gleichung B.1 ergibt sich sofort $f(e_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$. □

Damit ist eine lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ bereits durch die Bilder $f(e_1), ..., f(e_n) \in \mathbb{R}^m$ der Standardbasis $\{e_1, ..., e_n\}$ des $\mathbb{R}^n$ eindeutig festgelegt. Hierbei können die Bilder $f(e_1), ..., f(e_n) \in \mathbb{R}^m$ beliebig vorgegeben werden.

Literaturverzeichnis

[A1] B. Aulbach, *Analysis I,* Universität Augsburg 2001

[A2] B. Aulbach, *Analysis II,* Universität Augsburg 2002

[BF1] M. Barner und F. Flohr, *Analysis I,* de Gruyter, Berlin 1974

[BE] R. Busam und T. Epp, *Prüfungstrainer Analysis,* 2. Auflage, Spektrum Verlag, Berlin 2007

[B1] T. Bröcker, *Analysis I,* 2. Auflage, Spektrum akademischer Verlag, Heidelberg 1995 (siehe `http://www.mathematik.uni-regensburg.de/broecker/index.htm`)

[B2] T. Bröcker, *Analysis II,* 2. Auflage, Spektrum akademischer Verlag, Heidelberg 1995 (siehe `http://www.mathematik.uni-regensburg.de/broecker/index.htm`)

[D] J. Dieudonné, *Foundations of modern Analysis* Academic Press, New York 1960

[Ebb] H.-D. Ebbingshaus (Hrsg.), *Zahlen,* 3. Auflage, Springer-Verlag, Berlin 1992

[EFT] H.-D. Ebbingshaus, J. Flum und W. Thomas, *Einführung in die mathematische Logik,* 4. Auflage, Spektrum akademischer Verlag, Heidelberg 1996

[Esch] J.-H. Eschenburg, *Reelle Analysis mehrerer Veränderlicher*, Vorlesungsskript, Universität Augsburg, 1995 u. 2008

[EJ] J.-H. Eschenburg und J. Jost, *Differentialgeometrie und Minimalflächen*, Springer-Verlag, Berlin 2007

[Fi1] G. Fischer, *Lineare Algebra,* 17. Auflage, Vieweg+Teubner, Wiesbaden 2010 (16. Auflage http://dx.doi.org/10.1007/978-3-8348-9574-5)

[Fo1] O. Forster, *Analysis 1,* 4. Auflage, Vieweg, Braunschweig 1983 (aktuellere Auflage unter http://dx.doi.org/10.1007/978-3-8348-8139-7)

[Fo2] O. Forster, *Analysis 2,* 5. Auflage, Vieweg, Braunschweig 1984 (aktuellere Auflage unter http://dx.doi.org/10.1007/978-3-8348-8103-8)

[FW1] O. Forster und R. Wessoly, *Übungsbuch zur Analysis 1,* 5. Auflage, Vieweg+Teubner, Wiesbaden 2011 (siehe http://dx.doi.org/10.1007/978-3-8348-8228-8)

[FT2] O. Forster und T. Szymczak , *Übungsbuch zur Analysis 2,* 7. Auflage, Vieweg+Teubner, Wiesbaden 2011 (siehe http://dx.doi.org/10.1007/978-3-8348-8140-3)

[GIS] S. Gottwald, H.-J. Ilgauds und K.-H. Schlote (Herausg.), *Lexikon bedeutender Mathematiker,* Bibliographisches Institut, Leipzig 1990

[GBGW1] P. Günther, K. Beyer, S. Gottwald und V. Wünsch, *Grundkurs Analysis, Teil 1,* 2. Auflage, Teubner, Leipzig 1972

[GBGW2] P. Günther, K. Beyer, S. Gottwald und V. Wünsch, *Grundkurs Analysis, Teil 2,* Teubner, Leipzig 1973

[He1] E. Heintze, *Lineare Algebra I,* Manuskript, Universität Augsburg, Sommersemester 2008

[H1] H. Heuser, *Lehrbuch der Analysis, Teil 1,* 11. Auflage, B. G. Teubner, Stuttgart 1994

[H2] H. Heuser, *Lehrbuch der Analysis, Teil 2,* 9. Auflage, B. G. Teubner, Stuttgart 1995

[HS] D. Hoffmann und F.-W. Schäfke, *Integrale,* Bibliographisches Institut, Mannheim 1992

[K1] K. Königsberger, *Analysis 1,* 3. Auflage, Springer-Verlag, Berlin 1995

[K2] K. Königsberger, *Analysis 2,* Springer-Verlag, Berlin 1993

[Q] B. v. Querenburg, *Mengentheoretische Topologie,* 3. Auflage, Springer-Verlag, Berlin 2008

[S1] R. Scherz, *Protokoll der Vorlesung Analysis I,* Universität Augsburg, WS 2000/01

[ST] I.M. Singer und J.A. Thorpe, *Lecture Notes on elementary topology and geometry,* Springer-Verlag, Berlin 1967

[W1] W. Walter, *Analysis 1,* 3. Auflage, Springer-Verlag, Berlin 1992

[W2] W. Walter, *Analysis 2,* 5. Auflage, Springer-Verlag, Berlin 2002

Da an deutschsprachigen Universitäten ein Grundvorlesungszyklus in Analysis zum Pflichtkanon gehört, gibt es eine große Fülle deutschsprachiger Lehrbücher. Hier kann daher nur eine sehr kleine Auswahl angeben werden, mit der natürlich kein Werturteil verbunden ist.

Bei Büchern, welche die Universitätsbibliothek Augsburg online abonniert hat, ist die URL angegeben und verlinkt. Von Computern der Universität Augsburg oder über VPN-Client sollten diese Bücher online zugänglich sein.

Diese Vorlesung orientiert sich vornehmlich an [Fo1, Fo2, K1], in einigen Abschnitten auch an [A1, BF1, GBGW2, K2].

Zu den Büchern [Fo1, Fo2] von O. Forster gibt es auch Übungsbücher [FW1, FT2], in welchen Sie viele Aufgaben zum Selbststudium und zur Vertiefung des Stoffes finden können. Zur Vorbereitung auf mündliche Prüfungen sei noch auf [BE] verwiesen.

Index

Schlagwortverzeichnis