

Analysis für Elektrotechnik

Prof. Dr. Volker Bach

Technische Universität Braunschweig

Sommersemester 2026

Diese Vorlesung ist eine Einführung in die Analysis einer und mehrerer reellen Veränderlichen. Jedes Kapitel enthält am Schluss neben dem in der Vorlesung behandelten und somit prüfungsrelevanten Stoff dargestellt noch einen Abschnitt mit (nicht prüfungsrelevanten) Ergänzungen, d.h. Beweise und Beispiele, die wegen der Fülle des Materials in der Vorlesung nicht behandelt werden können.

Inhaltsverzeichnis

I.	Grundlagen, Konventionen und Notationen	5
I.1.	Quantoren und Logik	5
I.2.	Mengen	8
I.3.	Ordnungsrelationen	12
I.4.	Funktionen	13
I.5.	Beweistechniken	15
	I.5.1. Vollständige Induktion	15
	I.5.2. Beweis durch Kontraposition	16
I.6.	Notationen	16
I.7.	Ergänzungen	18
	I.7.1. Äquivalenzrelationen	18
	I.7.2. Das griechische Alphabet	20
II.	Reelle und komplexe Zahlen	21
II.1.	Körper	21
II.2.	Das Supremumsaxiom	22
II.3.	Ergänzung: Konstruktion der Reellen Zahlen	25
	II.3.1. Äquivalenz rationaler Cauchy-Folgen	25
	II.3.2. Die Grundrechenarten	27
	II.3.3. Das Supremumsaxiom	30
III.	Endliche, abzählbare und überabzählbare Mengen	32
III.1.	Abzählbare Mengen	32
III.2.	Dezimaldarstellung von Zahlen	33
III.3.	Ergänzungen	35
	III.3.1. Teilmengen abzählbarer Mengen sind abzählbar	35
	III.3.2. Abzählbare Vereinigungen abzählbarer Mengen sind abzählbar	36
	III.3.3. Beweis der Gleichmächtigkeit der reellen Zahlen und der Dezimal- darstellungen (Lemma III.8)	36
IV.	Folgen und Konvergenz	39
IV.1.	Konvergenz von Zahlenfolgen	39
IV.2.	Reelle Folgen und Monotonie	41
IV.3.	Cauchy-Folgen	44
IV.4.	Ergänzungen	46
	IV.4.1. Vertauschung von Limiten mit Produkt und Quotient	46
	IV.4.2. Limes Superior/Inferior als größter/kleinsten Häufungswert	47

V.	Reihen	49
V.1.	Definitionen und Beispiele	49
V.2.	Absolute Konvergenz	50
V.3.	Konvergenzkriterien	53
V.4.	Exponentialfunktion und trigonometrische Funktionen	55
V.5.	Ergänzungen	58
	V.5.1. Der Satz von Cauchy zur Konvergenz von Reihen nichtnegativer, monotoner Summanden	58
	V.5.2. Beweis des Großen Umordnungssatzes V.7	59
	V.5.3. Beweis von Satz V.9 zum Cauchy-Produkt	60
	V.5.4. Rückführung des Quotientenkriteriums auf das Wurzelkriterium	61
VI.	Stetigkeit	63
VI.1.	Offene, abgeschlossene und kompakte Mengen	63
VI.2.	Stetige Abbildungen	65
VI.3.	Erhaltung topologischer Eigenschaften durch stetige Funktionen	67
VII.	Differentiation	69
VII.1.	Differenzierbare Funktionen	69
VII.2.	Minima, Maxima und Mittelwertsatz	73
VII.3.	Höhere Ableitungen und der Satz von Taylor	76
VII.4.	Ergänzungen	81
	VII.4.1 Beweis der Kettenregel	81
	VII.4.2 Zwischenwertsatz für die Ableitung	82
	VII.4.3 $0/0$ und ∞/∞ – Der Satz von L’Hospital	82
VIII.	Integration	84
VIII.1.	Partitionen, Ober- und Untersummen	84
VIII.2.	Riemann-integrierbare Funktionen	86
VIII.3.	Integration und Differentiation	89
VIII.4.	Absolute Integrabilität und Uneigentliche Integrale	93
VIII.5.	Ergänzungen	97
	VIII.5.1 Beweis von Lemma VIII.6	97
	VIII.5.2 Monotonie und Linearität des Integrals	97
	VIII.5.3 Integrabilität stetiger Funktionen	99
	VIII.5.4 Beweis von Satz VIII.9	100
	VIII.5.5 Variablensubstitution – Beweis von Satz VIII.14	102
	VIII.5.6 Der Satz von Taylor mit Integralrestglied	103
IX.	Normen und Banachräume	106
IX.1.	Normierte Räume und Banachräume	107
IX.2.	Topologie und Stetigkeit in normierten Räumen	109
IX.3.	Ergänzungen	111
	IX.3.1 Vektorräume und lineare Abbildungen	111

IX.3.2.	Inneres und Abschluss von Mengen	115
X.	Differentiation vektorwertiger Funktionen	117
X.1.	Partielle und stetige Differenzierbarkeit	117
X.2.	Totale Differenzierbarkeit	122
X.3.	Ergänzungen	125
X.3.1.	Beschränkte lineare Operatoren	125
X.3.2.	Eigenschaften der Operatornorm	126
X.3.3.	Äquivalenz von Stetigkeit und Beschränktheit linearer Operatoren . .	126
X.3.4.	Beweis der Kettenregel	127
X.3.5.	Äquivalenz stetiger partieller und stetiger totaler Differenzierbarkeit	128
XI.	Richtungsableitung und Extrema	130
XI.1.	Der Satz von Taylor in mehreren Veränderlichen	130
XI.2.	Gradient, Richtungsableitung und Hessesche Matrix	133
XI.3.	Kritische Punkte und Extremalwerte bei mehreren Veränderlichen	134
XI.4.	Ergänzungen	139
XI.4.1.	Beweis von Korollar XI.2	139
XI.4.2.	Das Minimax-Prinzip für Matrizen (Lemma XI.7)	140
XII.	Maßtheoretische Grundbegriffe	142
XII.1.	Ringe und Figuren	142
XII.2.	Sigma-Algebren und Maße	144
XII.3.	Ergänzungen	148
XII.3.1	Ringe und Figuren	148
XII.3.2	Prämaße	151
XII.3.3	Sigma-Algebren	155
XII.3.4	Das äußere Maß	157
XII.3.5	Maße	161
XIII.	Integration über Maße	164
XIII.1.	Messbare Funktionen und Definition des Integrals	164
XIII.2.	Eigenschaften des Integrals	167
XIII.3.	Die Konvergenzsätze	167
XIII.4.	Ergänzungen	172
XIII.4.	Riemann-Integrabilität monotoner Funktionen	172
XIII.4.	Messbare Funktionen	175
XIII.4.	Die Konvergenzsätze	178
XIII.4.	Eigenschaften des Integrals	180

I. Grundlagen, Konventionen und Notationen

Dieses Kapitel stellt eine Übersicht über in der Mathematik häufig gebrauchte Begriffe, Konventionen und Notationen dar. Der Inhalt wurde schon im vergangenen Wintersemester in der Vorlesung *Lineare Algebra für Elektrotechnik* behandelt und wird hier nur für die Studienanfänger des laufenden Sommersemesters vorgestellt. In der Vorlesung wird dieses Kapitel nicht behandelt, und Leser:innen sind gehalten, sich den Stoff selbst anzueignen.

I.1. Quantoren und Logik

- Eine Aussage im mathematischen Sinne ist ein Wortgebilde, dem man entweder den Wert wahr (w) oder falsch (f) zuordnen kann.
 - „*Es regnet jetzt*“, „*Braunschweig ist die schönste Stadt Deutschlands*“ oder „ $1 \cdot 1 = 4$ “ sind jeweils Aussagen,
 - „*Grün*“, „*Fünfzehn Mann auf des toten Mannes Kiste -Johoo, johoo, johoo- und 'ne Buddel mit Rum!*“ oder „ $15x + 7y$ “ sind jeweils keine Aussagen.

Dabei geht es nur um die prinzipielle Zuordnung von w und f und weder um die praktische Nachprüfbarkeit der Aussage als Fakt noch um deren Objektivität noch um die Frage, ob sie eine Wertung beinhaltet.

- Das Zeichen $=$ bedeutet Gleichheit und ist in seiner Bedeutung evident. Das Zeichen $:=$ bedeutet, dass die linke Seite durch die rechte definiert wird, das Zeichen $\hat{=}$ bedeutet, dass die linke Seite durch die rechte (im Sinn einer Namensgebung) abgekürzt wird.

Beispiel:

$$a := f(0), \quad f(1) \hat{=} b. \tag{I.1}$$

bedeutet

$$a \text{ wird als Wert der Funktion } f \text{ bei } 0 \text{ definiert, der Wert der Funktion } f \text{ bei } 1 \text{ wird hingegen mit } b \text{ bezeichnet.} \tag{I.2}$$

- Das Zeichen $\forall$ bedeutet **für alle** (umgedrehtes A wie Alle).

Beispiel:

$$\forall x, y > 0 : \quad x \cdot y > 0 \quad (I.3)$$

bedeutet

$$\text{Für alle } x > 0 \text{ und } y > 0 \text{ gilt } x \cdot y > 0. \quad (I.4)$$

- Das Zeichen $\exists$ bedeutet **es existiert** (umgedrehtes E wie E existiert).

Beispiel:

$$\forall x > 0 \exists y < 0 : \quad x + y = 0 \quad (I.5)$$

bedeutet

$$\text{Für jedes } x \text{ größer Null existiert ein } y \text{ kleiner Null, so dass } x + y = 0 \text{ gilt. (Das gesuchte } y \text{ ist natürlich } -x.) \quad (I.6)$$

- Man beachte, dass Quantoren im Allgemeinen nicht vertauscht werden dürfen. Dazu betrachten wir folgende Beispiele:

$$\forall n \in \mathbb{N} \exists m \in \mathbb{N} : \quad m > n$$

bedeutet

$$\text{Zu jeder natürlichen Zahl } n \text{ gibt es eine natürliche Zahl } m \text{ so, dass } m \text{ größer als } n \text{ ist. (Diese Aussage ist wahr).} \quad (I.7)$$

$$\exists m \in \mathbb{N} \forall n \in \mathbb{N} : \quad m > n$$

bedeutet

$$\text{Es gibt eine natürliche Zahl } m, \text{ so dass alle natürlichen Zahlen } n \text{ kleiner sind als } m. \text{ (Diese Aussage ist falsch).} \quad (I.8)$$

- Das Zeichen $\Rightarrow$ bedeutet **impliziert**.

Beispiel:

$$A \Rightarrow B \quad (I.9)$$

bedeutet

$$\text{Aussage } A \text{ impliziert Aussage } B, \text{ d.h.: Ist } A \text{ wahr, so ist auch } B \text{ wahr.} \quad (I.10)$$

- Wie oben gesagt, kann eine mathematische Aussage A ein Satz, eine Bedingung oder auch eine Behauptung sein. In jedem Fall ist sie aber wahr oder falsch, $A \in \{w, f\}$.

Beispiel:

$$A := \text{Es regnet.}, \quad B := \text{Die Erde wird nass.} \quad (I.11)$$

Dann gilt die Implikation $A \Rightarrow B$, was gelesen werden muss als

$$(A = w) \Rightarrow (B = w).$$

Dieses Beispiel wirkt etwas künstlich. Darum geben wir ein weiteres Beispiel, in dem die Aussagen von Platzhaltern abhängen:

$$A(x) := \begin{cases} w, & \text{falls } x \geq 5, \\ f, & \text{falls } x < 5, \end{cases} \quad B(y) := \begin{cases} w, & \text{falls } y \geq 7, \\ f, & \text{falls } y < 7, \end{cases} \quad C(z) := \begin{cases} w, & \text{falls } z \geq 33, \\ f, & \text{falls } z < 33, \end{cases}$$

dann gilt die folgende Implikation (s. (I.18)–(I.19)):

$$A(x) \wedge B(y) \Rightarrow C(x \cdot y). \quad (\text{I.12})$$

- Das Zeichen $\iff$ bedeutet **ist gleichwertig mit** oder **ist äquivalent zu**, d.h.

$$A \iff B \quad (\text{I.13})$$

bedeutet

$$A \text{ ist genau dann wahr, wenn } B \text{ wahr ist.} \quad (\text{I.14})$$

- Das Zeichen $\vee$ ist ein logisches **oder**, d.h.

$$A \vee B \quad (\text{I.15})$$

$$\begin{aligned} &\text{ist wahr, falls } A \text{ oder } B \text{ oder beide wahr sind und falsch,} \\ &\text{falls } A \text{ und (gleichzeitig auch) } B \text{ falsch sind.} \end{aligned} \quad (\text{I.16})$$

Beispiel:

$$\{x \cdot y = 0\} \iff \{(x = 0) \vee (y = 0)\}. \quad (\text{I.17})$$

- Das Zeichen $\wedge$ ist ein logisches **und**, d.h.

$$A \wedge B \quad (\text{I.18})$$

$$\begin{aligned} &\text{ist wahr, falls } A \text{ und (gleichzeitig auch) } B \text{ wahr sind und} \\ &\text{sonst falsch.} \end{aligned} \quad (\text{I.19})$$

Beispiel 1:

$$\{(x = 0) \wedge (y = 0)\} \Rightarrow \{x + y = 0\}. \quad (\text{I.20})$$

Beispiel 2:

$$A \iff B \text{ ist gleichwertig mit } (A \Rightarrow B) \wedge (B \Rightarrow A) \quad (\text{I.21})$$

d.h. (um die Verwirrung komplett zu machen)

$$(A \iff B) \iff [(A \Rightarrow B) \wedge (B \Rightarrow A)]. \quad (\text{I.22})$$

- Die logische **Negation** wird mit $\neg$ bezeichnet, also

$$\neg A \quad (\text{I.23})$$

$$\text{ist wahr, falls } A \text{ falsch ist und falsch, falls } A \text{ wahr ist.} \quad (\text{I.24})$$

- Eine wichtige Beobachtung ist die *Kontraposition*, d.h. dass

$$(A \Rightarrow B) \iff (\neg B \Rightarrow \neg A). \quad (\text{I.25})$$

Dies versteht man intuitiv sofort an folgendem Beispiel:

Es regnet. $\Rightarrow$ Die Erde wird nass.

ist gleichwertig mit

$$\text{Die Erde ist nicht nass.} \Rightarrow \text{Es regnet nicht.} \quad (\text{I.26})$$

- Es ist nützlich, sich die Werte der Aussagen A , B , $A \Rightarrow B$, $A \Leftrightarrow B$, $A \vee B$, $A \wedge B$ und $\neg A$, in einer Wertetabelle zu verdeutlichen:

A	B	$A \Rightarrow B$	$A \Leftrightarrow B$	$A \vee B$	$A \wedge B$	$\neg A$
w	w	w	w	w	w	f
w	f	f	f	w	f	f
f	w	w	f	w	f	w
f	f	w	w	f	f	w

(I.27)

Wir sehen, dass $[A \Rightarrow B] \Leftrightarrow [(\neg A) \vee B] \Leftrightarrow [\neg(A \wedge (\neg B))]$ gilt, da die Werte von $A \Rightarrow B$ und $(\neg A) \vee B$ für alle vier möglichen Werte des Paares $(A, B) \in \{(w, w), (w, f), (f, w), (f, f)\}$ übereinstimmen.

- Für logische Verknüpfungen gelten

$$\text{das Kommutativgesetz:} \quad (\text{I.28})$$

$$A \vee B = B \vee A,$$

$$A \wedge B = B \wedge A,$$

$$\text{das Assoziativgesetz:} \quad (\text{I.29})$$

$$A \vee (B \vee C) = (A \vee B) \vee C,$$

$$A \wedge (B \wedge C) = (A \wedge B) \wedge C,$$

$$\text{das Distributivgesetz:} \quad (\text{I.30})$$

$$A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C),$$

$$A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C).$$

$$\text{sowie:} \quad (\text{I.31})$$

$$\neg(A \vee B) = (\neg A) \wedge (\neg B),$$

$$\neg(A \wedge B) = (\neg A) \vee (\neg B).$$

I.2. Mengen

Mengen sind (endliche, abzählbare oder sogar überabzählbare) Sammlungen mathematischer Objekte.

- $\{1, 5, 9\}$ ist die Menge, die die Zahlen 1, 5 und 9 enthält,
- $\{x \in \mathbb{N} \mid x < 5\}$ enthält alle natürlichen Zahlen, die kleiner als 5 sind (also $\{x \in \mathbb{N} \mid x < 5\} = \{1, 2, 3, 4\}$).
- Jedes Element einer Menge wird nur einmal aufgeführt, beispielsweise ist $\{1, 5, 9, 9, 5\} = \{1, 5, 9\}$.
- $x \in M$ heißt x **ist Element der Menge** M .
- $x \notin M$ heißt x **ist nicht in der Menge** M **enthalten**.
- Die **Anzahl der Elemente** einer Menge M wird mit $|M|$ oder auch $\#M$ bezeichnet. Beispielsweise ist $|\{1, 5, 9\}| = 3$.
- $A \subseteq B$ bedeutet, dass die Menge A in der Menge B enthalten ist, also dass A **Teilmenge** von B ist. Umgekehrt heißt $A \supseteq B$, dass die Menge A die Menge B enthält:

$$(A \subseteq B) \iff (B \supseteq A) \iff (x \in A \Rightarrow x \in B). \quad (\text{I.32})$$

- $\emptyset = \{ \}$ ist die leere Menge, die kein Element enthält.
- Gleichheit von Mengen bedeutet elementweise Übereinstimmung,

$$(A = B) \iff (x \in A \iff x \in B). \quad (\text{I.33})$$

- $\{x \mid E(x)\}$ und $\{x\}_{E(x)}$ bezeichnen die *Menge aller x , die die Eigenschaft $E(x)$ besitzen*.

Beispiel:

$$U := \{n \mid \exists k \in \mathbb{N} : n = 2k - 1\} \quad (\text{I.34})$$

ist

$$\text{die Menge aller } n, \text{ für die es eine natürliche Zahl } k \text{ gibt, so dass } n = 2k - 1 \text{ gilt}, \quad (\text{I.35})$$

d.h.

$$U = \{2k - 1 \mid k \in \mathbb{N}\} = 2\mathbb{N} - 1 \quad (\text{I.36})$$

ist die Menge aller ungeraden Zahlen.

Die Eigenschaft $E(x)$ kann auch durch eine **Indexmenge** $\mathcal{I}$ charakterisiert sein. Beispiel: Mit $\mathcal{I} := \{1, 3, 5, 7\}$ ist

$$\{x_i \mid i \in \mathcal{I}\} = \{x_i\}_{i \in \mathcal{I}} = \{x_1, x_3, x_5, x_7\}. \quad (\text{I.37})$$

- Haben wir mehrere Mengen, etwa A_1 , A_2 und A_3 , so bildet $M = \{A_1, A_2, A_3\}$ wieder eine Menge – eine Menge von Mengen. Dies kann man so fortsetzen und kommt zu Mengen von Mengen von Mengen u.s.w. Der Übersichtlichkeit halber hat sich deshalb im Sprachgebrauch bewährt, die übergeordnete Menge M als **Familie**, **System**, **Kollektion** oder auch **Klasse** zu bezeichnen. Somit ist M *die Familie der Mengen A_1 , A_2 und A_3* .

- Die Familie aller Teilmengen einer Menge M bezeichnet man als ihre **Potenzmenge** $\mathfrak{P}(M)$. Dabei zählen auch die leere Menge $\emptyset$ und M selbst als Teilmenge von M . Für $|M| < \infty$ ist $|\mathfrak{P}(M)| = 2^{|M|}$. (Warum?)

Beispiel:

$$M := \{1, 2, 3\} \Rightarrow \mathfrak{P}(M) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\}. \quad (\text{I.38})$$

- Die **Vereinigung**, der **Durchschnitt** und die **Differenz** zweier Mengen A, B werden wie folgt bezeichnet:

$$\text{Vereinigung:} \quad A \cup B = \{x \mid (x \in A) \vee (x \in B)\}, \quad (\text{I.39})$$

$$\text{Durchschnitt:} \quad A \cap B = \{x \mid (x \in A) \wedge (x \in B)\}, \quad (\text{I.40})$$

$$\text{Differenz:} \quad A \setminus B = \{x \mid (x \in A) \wedge (x \notin B)\}. \quad (\text{I.41})$$

- Ist A eine Teilmenge einer **Obermenge** M , d.h. $A \subseteq M$, so bezeichnet

$$A^c := M \setminus A \quad (\text{I.42})$$

das **Komplement von A bezüglich M** . (Vorsicht, der Notation A^c für das Komplement sieht man die Grundmenge M , auf die sie sich bezieht nicht mehr an!)

Beispiel: Seien $A = \{1, 2, 3\}$, $B = \{3, 4, 5\}$, $M = \{1, 2, \dots, 10\}$. Dann sind

$$A \cup B = \{1, 2, 3, 4, 5\}, \quad A \cap B = \{3\}, \quad (\text{I.43})$$

$$A \setminus B = \{1, 2\}, \quad A^c = \{4, 5, 6, 7, 8, 9, 10\}. \quad (\text{I.44})$$

- Vereinigungen und Durchschnitte können auch über Familien $\{A_i\}_{i \in \mathcal{I}}$ von Mengen A_i gebildet werden, wobei i eine Indexmenge $\mathcal{I}$ durchläuft.

Beispiel:

$$\bigcup_{i \in \mathcal{I}} A_i := \{x \mid \exists i \in \mathcal{I} : x \in A_i\}, \quad (\text{I.45})$$

ist

$$\text{die Vereinigung der Mengen } A_i, \text{ d.h. die } x, \text{ die in (mindestens) einer Menge } A_i \text{ mit } i \in \mathcal{I} \text{ enthalten sind;} \quad (\text{I.46})$$

$$\bigcap_{i \in \mathcal{I}} A_i := \{x \mid \forall i \in \mathcal{I} : x \in A_i\}, \quad (\text{I.47})$$

ist

$$\text{der Durchschnitt der Mengen } A_i, \text{ d.h. die } x, \text{ die in allen Mengen } A_i \text{ mit } i \in \mathcal{I} \text{ enthalten sind.} \quad (\text{I.48})$$

- Für Vereinigung, Durchschnitt und Komplementbildung von Mengen gelten Kommutativ-, Assoziativ-, und Distributionsgesetze, analog zu den entsprechenden Gesetzen für die logischen Verknüpfungen $\vee$, $\wedge$ und $\neg$.

- Sind A und B zwei nichtleere Mengen, so bezeichnet

$$A \times B = \{(a, b) \mid a \in A, b \in B\} \quad (\text{I.49})$$

das kartesische Produkt von A und B , d.h. die Menge aller Paare (a, b) , die sich mit Elementen a aus A und b aus B bilden lässt. (I.50)

Allgemeiner ist

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \dots, a_n) \mid a_1 \in A_1, \dots, a_n \in A_n\} \quad (\text{I.51})$$

das (n -fache) kartesische Produkt der Mengen $A_1, A_2, \dots, A_n$, d.h. die Menge aller n -Tupel $(a_1, a_2, \dots, a_n)$, die sich mit Elementen a_i aus A_i , für $i \in \{1, 2, \dots, n\}$ bilden lässt. (I.52)

Vorsicht! Oft wird $A_1 \times A_2 \times \cdots \times A_n$ mit $A_1 \cup A_2 \cup \cdots \cup A_n$ verwechselt. Der Unterschied wird aber schon deutlich, wenn man die Zahl der Elemente für $A := A_1 = A_2$ betrachtet:

$$|A \times A| = |A|^2, \quad (\text{I.53})$$

$$|A \cup A| = |A|. \quad (\text{I.54})$$

- Sind $a_1, a_2, a_3, \dots$ Zahlen, so bezeichnen wir $(a_n)_{n \in \mathbb{N}}$ als **Zahlenfolge**. Man beachte auch hier den Unterschied zwischen dem Tupel $(a_n)_{n \in \mathbb{N}}$ und der Menge $\{a_n\}_{n \in \mathbb{N}}$. So ist beispielsweise für die Zahlenfolge $a_1 = a_2 = a_3 = \dots = 1$, die konstant gleich eins ist,

$$(a_n)_{n \in \mathbb{N}} = (1, 1, 1, \dots), \quad (\text{I.55})$$

aber

$$\{a_n\}_{n \in \mathbb{N}} = \{1, 1, 1, \dots\} = \{1\}. \quad (\text{I.56})$$

- Häufig wiederkehrende Mengen haben in der Mathematik eine eigene Bezeichnung bekommen. Wir listen die Symbole für die wichtigsten Zahlenmengen auf:

$$\text{die natürlichen Zahlen: } \mathbb{N} := \{1, 2, 3, \dots\}, \quad (\text{I.57})$$

$$\text{die natürlichen Zahlen mit Null: } \mathbb{N}_0 := \{0, 1, 2, 3, \dots\}, \quad (\text{I.58})$$

$$\text{die ganzen Zahlen: } \mathbb{Z} := \{0, 1, -1, 2, -2, 3, -3, \dots\}, \quad (\text{I.59})$$

$$\text{die rationalen Zahlen: } \mathbb{Q} := \left\{ \frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N} \right\}, \quad (\text{I.60})$$

$$\text{die reellen Zahlen: } \mathbb{R} \quad (\text{I.61})$$

$$\text{die komplexen Zahlen: } \mathbb{C} \quad (\text{I.62})$$

Die präzise Definition der reellen oder gar der komplexen Zahlen geht über den üblichen Schulstoff hinaus. Wir werden dies in den kommenden Wochen in der Vorlesung behandeln.

- Weiterhin führen wir für $m, n \in \mathbb{N}_0$ mit $m \leq n$ noch die Bezeichnung

$$\mathbb{Z}_m^n := \{m, m+1, m+2, \dots, n\} \quad (\text{I.63})$$

für die natürlichen Zahlen zwischen m und n ein.

I.3. Ordnungsrelationen

Die Zeichen $<$, $>$, $\leq$, $\geq$ haben wir in verschiedenen Beispielen in den vorigen Abschnitten wie selbstverständlich benutzt.

- $a \leq b$ heißt a **ist kleiner als oder gleich** b .
- $a \geq b$ heißt a **ist größer als oder gleich** b .
- $a < b$ heißt a **ist kleiner als** b . Zur Unterscheidung dieser Relation von $a \leq b$ sagt man auch a **ist echt kleiner als** b oder a **ist strikt kleiner als** b .
- $a > b$ heißt a **ist (echt, strikt) größer als** b .
- Offenbar gilt

$$a < b \iff b > a, \quad (\text{I.64})$$

$$a \leq b \iff (a < b) \vee (a = b), \quad (\text{I.65})$$

$$a \geq b \iff (a > b) \vee (a = b). \quad (\text{I.66})$$

Die Ordnungsrelation $<$ lässt sich aber auch auf andere Mengen, als den uns vertrauten Zahlen übertragen. Deshalb ist es zweckmäßig, den Begriff einer geordneten Menge präzise zu definieren.

Definition I.1. Eine Menge $S \neq \emptyset$ heißt **(total) geordnet** bezüglich “ $<$ ” $:\Leftrightarrow$

$$(i) \quad \text{Sind } a, b \in S, \text{ so gilt genau eine der drei Relationen } a < b, a = b \text{ oder } a > b. \quad (\text{I.67})$$

$$(ii) \quad \text{Sind } a, b, c \in S, \text{ und gilt } a < b \text{ und } b < c, \text{ dann gilt auch } a < c. \quad (\text{I.68})$$

Das Symbol “ $<$ ” heißt **Ordnungsrelation** auf S .

Beispiele für total geordnete Mengen sind $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$. Auf den komplexen Zahlen gibt es keine (mit den Verknüpfungen verträgliche) Ordnungsrelation. Ebenso gibt es keine (mit den Verknüpfungen verträgliche) Ordnungsrelation auf den Vektoren in $\mathbb{R}^3$.

Mit Hilfe der Ordnungsrelation kann man **Intervalle** in $\mathbb{R}$ definieren. Seien $a, b \in \mathbb{R}$ und

$a \leq b$. Dann heißen

$$(a, b) := \{x \in \mathbb{R} \mid a < x < b\} \quad (\text{I.69})$$

das **offene Intervall** von a nach b ,

$$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\} \quad (\text{I.70})$$

das **abgeschlossene Intervall** von a nach b ,

$$[a, b) := \{x \in \mathbb{R} \mid a \leq x < b\} \quad (\text{I.71})$$

das **rechts halboffene Intervall** von a nach b ,

$$(a, b] := \{x \in \mathbb{R} \mid a < x \leq b\} \quad (\text{I.72})$$

das **links halboffene Intervall** von a nach b

und insbesondere

$$\mathbb{R}^+ := \{x \in \mathbb{R} \mid x > 0\}, \quad \mathbb{R}^- := \{x \in \mathbb{R} \mid x < 0\}, \quad (\text{I.73})$$

$$\mathbb{R}_0^+ := \{x \in \mathbb{R} \mid x \geq 0\}, \quad \mathbb{R}_0^- := \{x \in \mathbb{R} \mid x \leq 0\}. \quad (\text{I.74})$$

I.4. Funktionen

Funktionen, auch **Abbildungen** genannt, sind die wichtigsten Objekte der Mathematik. Eine Funktion f ordnet jedem Element x seiner **Definitions-*m*enge** $\mathcal{D}$ genau ein Element $f(x)$ seiner **Wertemenge** $\mathcal{W}$ zu. Die symbolische Schreibweise dafür ist

$$f : \mathcal{D} \rightarrow \mathcal{W}, \quad x \mapsto f(x). \quad (\text{I.75})$$

Dabei ist die Definitionsmenge zwar voll ausgeschöpft, denn $f(x)$ ist für jedes $x \in \mathcal{D}$ definiert. Für die Wertemenge muss das aber nicht der Fall sein. Sind $\mathcal{D}' \subseteq \mathcal{D}$ und $\mathcal{W}' \subseteq \mathcal{W}$ Teilmengen von $\mathcal{D}$ bzw. $\mathcal{W}$, so bezeichnen wir mit

$$f(\mathcal{D}') := \{f(x) \in \mathcal{W} \mid x \in \mathcal{D}'\} \quad (\text{I.76})$$

die **Bildmenge** (oder das Bild) von $\mathcal{D}'$ und

$$f^{-1}(\mathcal{W}') := \{x \in \mathcal{D} \mid f(x) \in \mathcal{W}'\} \quad (\text{I.77})$$

die **Urbildmenge** (oder das Urbild) von $\mathcal{W}'$.

Es kann also $f(\mathcal{D}) \subset \mathcal{W}$ durchaus eine *echte* Teilmenge des Wertebereichs sein. (Wir sollten hier aber erwähnen, dass diese Konvention nicht einheitlich akzeptiert ist. Manche Autoren verlangen, dass für $f : \mathcal{D} \rightarrow \mathcal{W}$ auch stets $f(\mathcal{D}) = \mathcal{W}$ gilt, andere fordern noch nicht einmal, dass $f^{-1}(\mathcal{W}) = \mathcal{D}$.)

Definition I.2. Seien $\mathcal{D}, \mathcal{W} \neq \emptyset$ und $f : \mathcal{D} \rightarrow \mathcal{W}$ eine Abbildung.

$$f \text{ heißt } \mathbf{surjektiv} \quad :\Leftrightarrow \quad f(\mathcal{D}) = \mathcal{W}, \quad (\text{I.78})$$

$$f \text{ heißt } \mathbf{injektiv} \quad :\Leftrightarrow \quad \forall x, x' \in \mathcal{D} : (f(x) = f(x')) \Rightarrow (x = x'), \quad (\text{I.79})$$

$$f \text{ heißt } \mathbf{bijektiv} \quad :\Leftrightarrow \quad f \text{ ist surjektiv und injektiv.} \quad (\text{I.80})$$

Bemerkungen und Beispiele.

- $\exp : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto e^x$ ist nicht surjektiv, (wegen $e^x > 0$) aber injektiv (wegen der strengen Monotonie der Exponentialfunktion).
- $\sin : \mathbb{R} \rightarrow [-1, 1]$, $x \mapsto \sin x$ ist surjektiv aber nicht injektiv.
- $\tan : (-\pi/2, \pi/2) \rightarrow \mathbb{R}$, $x \mapsto \tan x$ ist bijektiv.

Definition I.3. Für $g : A \rightarrow B$ und $f : g(A) \rightarrow C$ ist die **Verkettung** oder **Komposition** oder auch **Hintereinanderschaltung** $f \circ g$ von g und f wie folgt definiert:

$$f \circ g : A \rightarrow C, \quad x \mapsto f(g(x)). \quad (\text{I.81})$$

Satz I.4. Seien $g : A \rightarrow B$ und $f : B \rightarrow C$.

- (i) Sind f und g surjektiv, so ist auch $f \circ g : A \rightarrow C$ surjektiv.
- (ii) Sind f und g injektiv, so ist auch $f \circ g : A \rightarrow C$ injektiv.
- (iii) Sind f und g bijektiv, so ist auch $f \circ g : A \rightarrow C$ bijektiv.

Beweis.

Zu (i): Sei $c \in C$. Weil f surjektiv ist, gibt es ein $b \in B$ mit $f(b) = c$. Weil g surjektiv ist, gibt es ein $a \in A$ mit $g(a) = b$. Also gilt $(f \circ g)(a) = c$.

Zu (ii): Seien $a, a' \in A$ mit $(f \circ g)(a) = (f \circ g)(a')$. Weil f injektiv ist, folgt $g(a) = g(a')$. Weil g injektiv ist, folgt dann auch $a = a'$.

Zu (iii): Folgt aus (i) und (ii). □

Wichtig ist also zu beachten, dass der Definitionsbereich von f mit dem Bildbereich von g übereinstimmt. Man beachte auch die Reihenfolge: obwohl die Komposition $f \circ g$ heißt, wird erst g auf $x \in A$ angewandt und danach f auf das Ergebnis $g(x) \in B$.

Die Bedeutung der Bijektivität liegt darin, dass sie die Existenz und Eindeutigkeit der Umkehrfunktion sichert, wie der folgende Satz zeigt.

Satz I.5. Seien A, B zwei nichtleere Mengen und $f : A \rightarrow B$ eine bijektive Abbildung. Dann gibt es eine eindeutige Abbildung $g : B \rightarrow A$ so, dass $g \circ f = \mathbb{1}_A$ und $f \circ g = \mathbb{1}_B$ gelten, d.h. dass

$$\forall x \in A : g[f(x)] = x \quad \text{und} \quad \forall y \in B : f[g(y)] = y \quad (\text{I.82})$$

gelten. In diesem Fall heißt $g : B \rightarrow A$ **Umkehrabbildung zu f** , und wir schreiben $g =: f^{-1}$.

Beweis. Sei $y \in B$. Weil f surjektiv ist, gibt es ein $x \in A$ so, dass $y = f(x)$, und weil f injektiv ist, ist x das einzige Element in A , für das $y = f(x)$ ist. Also definiert $g(y) := x$ eine Abbildung $g : B \rightarrow A$. Diese Abbildung hat die Eigenschaft, dass $f[g(y)] = f(x) = y$ für alle $y \in B$ gilt. Ist umgekehrt $x \in A$ beliebig, so setzen wir $y := f(x)$ und beobachten, dass $f(x) = y = f[g(y)] = f(g[f(x)])$ gilt. Aus der Injektivität von f folgt nun jedoch $x = g[f(x)]$. □

Eine wichtige Klasse von Funktionen ist die der charakteristischen Funktionen, auch Indikatorfunktionen genannt. Ist $\mathcal{D}$ eine nichtleere Menge und $A \subseteq \mathcal{D}$ eine Teilmenge, so ist die **charakteristische Funktion** von A gegeben als

$$\mathbb{1}_A : \mathcal{D} \rightarrow \{0, 1\}, \quad x \mapsto \begin{cases} 1 & \text{falls } x \in A, \\ 0 & \text{falls } x \notin A. \end{cases} \quad (\text{I.83})$$

Mit anderen Worten: $\mathbb{1}_A[x]$ ist genau dann gleich 1, wenn x in A liegt und anderenfalls gleich 0.

I.5. Beweistechniken

I.5.1. Vollständige Induktion

Eine häufig verwendete Beweistechnik ist die **vollständige Induktion**. Zunächst stellen wir das Verfahren abstrakt vor. Nehmen wir an, wir wollten Aussagen $A(1), A(2), A(3), \dots$ beweisen. Dann können wir folgende Tatsache verwenden.

Satz I.6. *Gibt es ein $n_0 \in \mathbb{N}$, so dass $A(n_0)$ wahr ist, und gilt die Implikation*

$$A(n) \Rightarrow A(n+1), \quad (\text{I.84})$$

für jedes $n \geq n_0$, $n \in \mathbb{N}$, so ist $A(m)$ wahr, für jedes $m \geq n_0$, $m \in \mathbb{N}$.

Beweis. Wendet man (I.84) $(m - n_0)$ -mal an, so erhält man

$$A(n_0) \Rightarrow A(n_0 + 1) \Rightarrow A(n_0 + 2) \Rightarrow \dots \Rightarrow A(m - 1) \Rightarrow A(m). \quad (\text{I.85})$$

□

Der Beweis durch vollständige Induktion wird an einem Beispiel am deutlichsten. Wir wollen für $n \in \mathbb{N}$ die Summe

$$F(n) := 1 + 2 + \dots + n \quad (\text{I.86})$$

berechnen, und wir haben die Vermutung, dass $F(n) = G(n)$, wobei

$$G(n) := \frac{n(n+1)}{2}. \quad (\text{I.87})$$

Nun gilt es, die Aussage

$$A(n) = w \quad :\Leftrightarrow \quad F(n) = G(n) \quad (\text{I.88})$$

für alle $n \in \mathbb{N}$ zu beweisen.

- **Induktionsanfang:** Wähle $n_0 := 1$. Dann ist

$$F(n_0) = F(1) = 1 = \frac{1(1+1)}{2} = G(1) = G(n_0), \quad (\text{I.89})$$

und $A(n_0) = A(1) = w$.

- **Induktionsannahme:** Seien $n \geq 1$ und gelte $A(n) = w$, also $F(n) = G(n)$.
- **Induktionsschritt:** Wir zeigen, dass aus $A(n) = w$ auch $A(n+1) = w$ folgt. Dazu beobachten wir, dass unter Verwendung von $F(n) = G(n)$ auch

$$\begin{aligned} F(n+1) &= n+1 + F(n) = n+1 + G(n) = n+1 + \frac{n(n+1)}{2} \\ &= \frac{(n+1)(n+2)}{2} = G(n+1) \end{aligned} \quad (\text{I.90})$$

gilt, dass somit also $A(n+1) = w$ richtig ist.

Nach Satz I.6 ist damit $A(n) = w$ für alle $n \in \mathbb{N}$ bewiesen.

I.5.2. Beweis durch Kontraposition

Neben der vollständigen Induktion ist auch der Beweis durch Kontraposition eine häufig verwendete Methode, die auf (I.25) beruht,

$$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A). \quad (\text{I.91})$$

Wir illustrieren dies wieder mit einem Beispiel: Seien die Aussagen A und B definiert durch

$$A = w :\Leftrightarrow \sqrt{2} \in \mathbb{Q}, \quad (\text{I.92})$$

$$B = w :\Leftrightarrow \exists p, q \in \mathbb{N}, \text{ ggT}(p, q) = 1 : 2q^2 = p^2 \quad (\text{I.93})$$

wobei $\text{ggT}(a, b) \in \mathbb{N}$ den größten gemeinsamen Teiler zweier natürlichen Zahlen $a, b \in \mathbb{N}$ bezeichne.

Ist nun $A = w$, so gibt es Zahlen $p', q' \in \mathbb{N}$ so, dass $\sqrt{2} = \frac{p'}{q'}$ gilt. Enthalten p' und q' einen gemeinsamen ganzzahligen Faktor $r \in \mathbb{N}$, sodass also $p' = pr$ und $q' = qr$ gelten, so können wir r herauskürzen und erhalten $\sqrt{2} = \frac{p}{q}$ mit teilerfremden p und q , d.h. $\text{ggT}(p, q) = 1$. Damit ist auch $B = w$, und wir erhalten

$$A \Rightarrow B. \quad (\text{I.94})$$

Die Aussage B ist jedoch stets falsch, weil dann der Primfaktor 2 in $2q^2$ in ungerader Anzahl und in p^2 in gerader Anzahl auftreten müsste.

$$\{B = f\} \Rightarrow \{\neg B = w\} \Rightarrow \{\neg A = w\} \Rightarrow \{\sqrt{2} \notin \mathbb{Q}\}. \quad (\text{I.95})$$

Wir bemerken, dass der Beweis durch Kontraposition sehr ähnlich zur Methode des *Widerspruchsbeweises* ist, letzterer beruht auf $[A \Rightarrow B] = (\neg A) \vee B = \neg[A \wedge (\neg B)]$.

I.6. Notationen

Seien $m, n \in \mathbb{Z}$ mit $m \leq n$ und

$$A = \{a_m, a_{m+1}, a_{m+2}, \dots, a_n\} \quad (\text{I.96})$$

eine Menge von Zahlen. Dann ist das Summenzeichen wie folgt definiert,

$$\sum_{i=m}^n a_i := a_m + a_{m+1} + a_{m+2} + \dots + a_n. \quad (\text{I.97})$$

Wir bemerken, dass der *Summationsindex* i durch irgend einen anderen Buchstaben außer m oder n ersetzt werden kann,

$$\sum_{i=m}^n a_i = \sum_{k=m}^n a_k = \sum_{j=m}^n a_j. \quad (\text{I.98})$$

Für $m > n$ wird $\sum_{i=m}^n a_i := 0$ definiert.

Mit $\mathcal{I} := \{m, m+1, \dots, n\}$ und A wird die Summe auch oft noch anders geschrieben:

$$\sum_{i=m}^n a_i = \sum_{i \in \mathcal{I}} a_i = \sum_{a \in A} a. \quad (\text{I.99})$$

I.7. Ergänzungen

I.7.1. Äquivalenzrelationen

Häufig lässt sich eine Menge in eine Familie disjunkter Teilmengen zerlegen, deren Elemente jeweils ähnliche Eigenschaften haben. Beispiel:

Wir zerlegen $\mathbb{Z}$ in $\mathbb{Z} = A_0 \cup A_1 \cup A_2$, wobei

$$A_0 := 3\mathbb{Z} = \{3k \mid k \in \mathbb{Z}\}, \quad (\text{I.100})$$

$$A_1 := 3\mathbb{Z} + 1 = \{3k + 1 \mid k \in \mathbb{Z}\}, \quad (\text{I.101})$$

$$A_2 := 3\mathbb{Z} + 2 = \{3k + 2 \mid k \in \mathbb{Z}\}. \quad (\text{I.102})$$

Offensichtlich sind $A_0 \cap A_1 = A_1 \cap A_2 = A_0 \cap A_2 = \emptyset$. Die Elemente in A_j ($j = 0, 1, 2$) lassen sich dadurch charakterisieren, dass sie einen Rest j beim Teilen durch 3 ergeben.

Wir formalisieren nun diese Überlegungen.

Definition I.7. Sei A eine Menge. Eine Abbildung $R : A \times A \rightarrow \{w, f\}$ heißt **Relation auf A** . Für $R(a, b) = w$ schreiben wir auch $a \sim b$.

Definition I.8. Eine Relation $R : A \times A \rightarrow \{w, f\}$ auf einer Menge A , mit $R(a, b) = w \Leftrightarrow a \sim b$ heißt **Äquivalenzrelation**, falls folgende drei Eigenschaften gelten:

$$\text{Reflexivität} \quad \forall a \in A : \quad a \sim a, \quad (\text{I.103})$$

$$\text{Symmetrie} \quad \forall a, b \in A : \quad a \sim b \Leftrightarrow b \sim a, \quad (\text{I.104})$$

$$\text{Transitivität} \quad \forall a, b, c \in A : \quad (a \sim b \wedge b \sim c) \Rightarrow (a \sim c). \quad (\text{I.105})$$

Satz I.9. Eine Äquivalenzrelation auf einer Menge A bewirkt eine Zerlegung von A in disjunkte Teilmengen. Dabei sind zwei Elemente aus A genau dann äquivalent, wenn sie derselben Teilmenge angehören.

Beweis. Zu $a \in A$ definieren wir

$$[a]_{\sim} := \{x \in A \mid a \sim x\}. \quad (\text{I.106})$$

Wegen $a \in [a]_{\sim}$ ist $[a]_{\sim}$ nicht leer. Wir zeigen nun für $a, b \in A$, dass

$$\text{entweder} \quad [a]_{\sim} \cap [b]_{\sim} = \emptyset \quad (\text{I.107})$$

$$\text{oder} \quad [a]_{\sim} = [b]_{\sim} \quad (\text{I.108})$$

gilt. (Wegen $[a]_{\sim} \neq \emptyset$ können (I.107) und (I.108) nicht gleichzeitig gelten.)

Sei $[a]_{\sim} \cap [b]_{\sim} \neq \emptyset$. Dann gibt es also ein gemeinsames Element $c \in [a]_{\sim}, c \in [b]_{\sim}$. Damit gelten $a \sim c$ und $c \sim b$, also auch $a \sim b$. Ist nun $x \in [a]_{\sim}$, dann gilt $x \sim a$ und mit $a \sim b$ auch $x \sim b$, also $x \in [b]_{\sim}$. Es folgt, dass $[a]_{\sim} \subseteq [b]_{\sim}$. Genauso erhält man $[b]_{\sim} \subseteq [a]_{\sim}$, also $[a]_{\sim} = [b]_{\sim}$. Damit ist (I.107)–(I.108) gezeigt.

Schreiben wir jetzt

$$A = \bigcup_{a \in A} [a]_{\sim}, \quad (\text{I.109})$$

folgt die Aussage unmittelbar durch Zusammenfassen gleicher $[a]_{\sim}$. □

Definition I.10. Die Teilmengen $[a]_{\sim}$ heißen **Äquivalenzklassen**. Die Familie der Äquivalenzklassen bezeichnet man mit

$$A/\sim \quad (\text{sprich: " } A \text{ modulo } \sim "). \quad (\text{I.110})$$

Liegt a in einer Äquivalenzklasse, so heißt a **Repräsentant der Klasse**.

Bemerkungen und Beispiele.

Sind $A := \mathbb{Z}$ die ganzen Zahlen und $p \in \mathbb{N}$ eine natürliche Zahl, so sind

$$m \sim n \Leftrightarrow \exists k \in \mathbb{Z} : m - n = kp. \quad (\text{I.111})$$

$$\Rightarrow \mathbb{Z} = [0]_{\sim} \cup [1]_{\sim} \cup \dots \cup [p-1]_{\sim}, \quad (\text{I.112})$$

$$[j]_{\sim} = \{kp + j \mid k \in \mathbb{Z}\}. \quad (\text{I.113})$$

$\mathbb{Z}/\sim$ bezeichnet man auch mit $\mathbb{Z}/p\mathbb{Z}$ oder $\mathbb{Z}_p$ und $[j]_{\sim} =: [j] \bmod p$.

Definition I.11. Sei $\sim$ eine Äquivalenzrelation auf einer Menge A . Eine Teilmenge $S \subseteq A$ heißt ein **vollständiges Repräsentantensystem zu $\sim$** , falls folgende zwei Eigenschaften gelten:

- (i) Jedes Element aus A ist zu einem Element aus S äquivalent.
- (ii) Die Elemente aus S sind paarweise nicht äquivalent.

Bemerkungen und Beispiele.

$$A := \{g \subseteq \mathbb{R}^2 \mid g \text{ ist eine Gerade}\}, \quad (\text{I.114})$$

$$g_1 \sim g_2 \Leftrightarrow g_1 \text{ und } g_2 \text{ sind parallel.}$$

$$\Rightarrow S = \{g \in A \mid g \cap \{\vec{0}\} = \{\vec{0}\}\} \quad (\text{I.115})$$

$$\text{ist ein vollständiges Repräsentantensystem.} \quad (\text{I.116})$$

I.7.2. Das griechische Alphabet

Das griechische Alphabet wird in der Mathematik häufig verwendet. Zum Abschluss geben wir noch eine Liste der gebräuchlichsten griechischen Buchstaben:

KLEIN					
α alpha	β beta	γ gamma	δ delta	ϵ epsilon	
ε epsilon	ζ zeta	η eta	θ theta	ϑ theta	
ι jota	κ kappa	λ lambda	μ mü	ν nü	
ξ xi	$\omicron$ o	π pi	φ phi	ρ rho	
ϱ rho	σ sigma	ς sigma	τ tau	υ upsilon	
ϕ phi	φ phi	χ chi	ψ psi	ω omega	
GROSS					
Γ Gamma	Δ Delta	Θ Theta	Λ Lambda	Ξ Xi	
Π Pi	Σ Sigma	Υ Upsilon	Φ Phi	Ψ Psi	
Ω Omega					

II. Reelle und komplexe Zahlen

II.1. Körper

Im vorigen Kapitel I haben wir die wichtigsten Zahlenmengen bereits genannt: die *natürlichen* Zahlen,

$$\mathbb{N} := \{1, 2, 3, \dots\}, \quad (\text{II.1})$$

die *ganzen* Zahlen,

$$\mathbb{Z} := \{0, +1, -1, +2, -2, \dots\}, \quad (\text{II.2})$$

die *rationalen* Zahlen,

$$\mathbb{Q} := \left\{ \frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N} \right\}, \quad (\text{II.3})$$

sowie die *reellen* und die *komplexen* Zahlen,

$$\mathbb{R} \quad \text{und} \quad \mathbb{C}. \quad (\text{II.4})$$

Wir wenden uns zunächst $\mathbb{N}$, $\mathbb{Z}$ und $\mathbb{Q}$ zu.

- Für $a, b \in \mathbb{N}$ ist auch $a + b \in \mathbb{N}$. Diese Tatsache bezeichnet man als *Abgeschlossenheit* von $\mathbb{N}$ bezüglich Addition.
- I.A. gilt $a - b \in \mathbb{N}$ jedoch nicht. Dafür geht man von $\mathbb{N}$ zu $\mathbb{Z}$ über; für $a, b \in \mathbb{Z}$ sind $a + b$ und $a - b \in \mathbb{Z}$. Insbesondere ist 0 das *neutrale Element* bezüglich Addition in $\mathbb{Z}$: $a + 0 = 0 + a = a$. Man sagt, dass $\mathbb{Z}$ bezüglich der Addition $+$ eine *Gruppe* bildet; diesen Begriff haben wir in der Vorlesung *Lineare Algebra für Elektrotechnik* ausführlich behandelt.
- Weiterhin ist $\mathbb{Z}$ auch bezüglich Multiplikation abgeschlossen, d.h. für $a, b \in \mathbb{Z}$ ist auch $a \cdot b \in \mathbb{Z}$, und es gilt das Distributivgesetz, $a(b + c) = ab + bc$. Somit ist $\mathbb{Z}$ bezüglich der Addition $+$ und der Multiplikation $(\cdot)$ ein *Ring*; diesen Begriff haben wir in der Vorlesung *Lineare Algebra für Elektrotechnik* ebenfalls ausführlich behandelt.
- Schließlich gelangt man von $\mathbb{Z}$ zu $\mathbb{Q}$ durch die Forderung, dass auch Abgeschlossenheit bezüglich Division gelten soll: Für $a, b \in \mathbb{Q}$ sind $a + b, a - b, a \cdot b \in \mathbb{Q}$ und $\frac{a}{b} \in \mathbb{Q}$, falls $b \neq 0$. Diese Eigenschaften von $\mathbb{Q}$ stehen auch exemplarisch für die allgemeine Definition eines *Körpers*; auch diesen Begriff haben wir in der Vorlesung *Lineare Algebra für Elektrotechnik* behandelt, knüpfen aber nochmal neu daran an:

Definition II.1. Eine nichtleere Menge $\mathbb{F} \neq \emptyset$ heißt **Körper**¹ : $\Leftrightarrow$

Auf $\mathbb{F}$ sind **Addition** $+: \mathbb{F} \times \mathbb{F} \rightarrow \mathbb{F}$ und **Multiplikation** $(\cdot): \mathbb{F} \times \mathbb{F} \rightarrow \mathbb{F}$ definiert, die die folgenden Eigenschaften besitzen:

- (i) $\mathbb{F}$ ist bezüglich Addition eine kommutative Gruppe mit neutralem Element Null $0 \in \mathbb{F}$, d.h. für es gelten $0 + a = a$, für alle $a \in \mathbb{F}$, und weiterhin

$$\forall a, b, c \in \mathbb{F} : \quad (a + b) + c = a + (b + c), \quad a + b = b + a, \quad (\text{II.5})$$

$$\forall a \in \mathbb{F} \exists (-a) \in \mathbb{F} : \quad a + (-a) = 0; \quad (\text{II.6})$$

- (ii) $\mathbb{F} \setminus \{0\}$ ist bezüglich Multiplikation eine kommutative Gruppe mit neutralem Element Eins $1 \in \mathbb{F} \setminus \{0\}$, d.h. es gelten $1 \cdot a = a$, für alle $a \in \mathbb{F} \setminus \{0\}$, und weiterhin

$$\forall a, b, c \in \mathbb{F} \setminus \{0\} : \quad (a \cdot b) \cdot c = a \cdot (b \cdot c), \quad a \cdot b = b \cdot a, \quad (\text{II.7})$$

$$\forall a \in \mathbb{F} \setminus \{0\} \exists \frac{1}{a} \in \mathbb{F} \setminus \{0\} : \quad a \cdot \frac{1}{a} = 1; \quad (\text{II.8})$$

- (iii) Es gilt das Distributivgesetz

$$\forall a, b, c \in \mathbb{F} : \quad a \cdot (b + c) = a \cdot b + a \cdot c. \quad (\text{II.9})$$

Bemerkungen und Beispiele.

- Aus $0 + 0 = 0$ folgt $0 \cdot a + 0 \cdot a = (0 + 0) \cdot a = 0 \cdot a$, woraus wir $0 \cdot a = 0$ durch Addition von $-(0 \cdot a)$ für jedes $a \in \mathbb{F} \setminus \{0\}$ gewinnen. Insbesondere gilt dann $1 \cdot 0 = 0$, und (II.7) lässt sich auf alle $a, b, c \in \mathbb{F}$ erweitern.
- Gemäß Definition II.1 enthält jeder Körper $\mathbb{F}$ mindestens Null und Eins, $0, 1 \in \mathbb{F}$, und diese sind voneinander verschieden.
- Der kleinst(möglich)e Körper ist $\mathbb{Z}_2 = \{0, 1\}$, und in der Tat ist dies ein Körper mit $0 + 0 = 0$, $0 + 1 = 1 + 0 = 1$, $1 + 1 = 0$ und $0 \cdot 0 = 0$, $0 \cdot 1 = 1 \cdot 0 = 0$, $1 \cdot 1 = 1$.
- Die Mengen $\mathbb{N}$ und $\mathbb{Z} \supseteq \mathbb{N}$ der natürlichen und der ganzen Zahlen sind keine Körper.
- Die Menge $\mathbb{Q} \supseteq \mathbb{Z}$ der rationalen Zahlen ist ein Körper, und zwar der kleinste, der $\mathbb{Z}$ enthält.

II.2. Das Supremumsaxiom

Wir erinnern als Nächstes an den Begriff der *totalen Ordnung*, den wir schon in Definition I.1 kennengelernt haben.

Definition II.2. Eine Menge $S \neq \emptyset$ heißt **total geordnet** bezüglich „ $<$ “ : $\Leftrightarrow$

- (i) Sind $a, b \in S$, so gilt genau eine der drei Relationen $a < b$, $a = b$ oder $a > b$. (II.10)

- (ii) Sind $a, b, c \in S$, und gilt $a < b$ und $b < c$, dann gilt auch $a < c$. (II.11)

¹engl.: field

Das Symbol „ $<$ “ heißt **Ordnungsrelation** auf S .

Beispiele für total geordnete Mengen sind $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ und -wie wir später sehen werden-, auch $\mathbb{R}$.

Definition II.3. Seien $S \neq \emptyset$ eine bezüglich „ $<$ “ total geordnete Menge und $T \subseteq S$ eine Teilmenge.

- (i) Gibt es ein $b \in S$ so, dass $t \leq b$ für alle $t \in T$ gilt, so heißt T **nach oben beschränkt**, und b heißt **obere Schranke** an T .
- (ii) Gibt es ein $b' \in S$ so, dass $t \geq b'$ für alle $t \in T$ gilt, so heißt T **nach unten beschränkt**, und b' heißt **untere Schranke** an T .
- (iii) Ist T nach oben und nach unten beschränkt, so heißt T **beschränkt**.

Definition II.4. Seien $S \neq \emptyset$ eine bezüglich „ $<$ “ total geordnete Menge, $T \subseteq S$ eine nach oben beschränkte Teilmenge. Ein Element $b \in S$ heißt **Supremum von T** : $\Leftrightarrow$

$$(i) \quad \forall t \in T : \quad t \leq b, \quad (II.12)$$

$$(ii) \quad \forall a \in S, a < b \exists t \in T : \quad a < t. \quad (II.13)$$

Bemerkungen und Beispiele.

- Besitzt eine Teilmenge $T \subseteq S$ einer total geordneten Menge S ein Supremum $b \in S$, so ist dieses eindeutig und wir schreiben

$$b =: \sup \{T\}, \quad (II.14)$$

für das Supremum von T .

- Sind $S := \mathbb{Q}$ und $T := \{a \in \mathbb{Q} \mid a^2 < 2\}$, so ist S total geordnet bezüglich der üblichen Relation „ $<$ “ zwischen Zahlen, und T ist nach oben beschränkt - z.B. ist $2 \in \mathbb{Q}$ eine obere Schranke.

Es existiert jedoch kein Supremum von T in S : Man sieht nämlich leicht ein, dass $b^2 = 2$ für das Supremum $b = \sup\{T\} \in S = \mathbb{Q}$ gelten müsste, diese Gleichung besitzt aber keine rationale Lösung, wie wir schon in der neunten Klasse lernen ($\sqrt{2}$ ist irrational).

- Wir folgern, dass $\mathbb{Q}$ das unten formulierte Supremumsaxiom nicht erfüllt.

Definition II.5. Eine total geordnete Menge $S \neq \emptyset$ erfüllt das **Supremumsaxiom**, falls jede nach oben beschränkte Teilmenge $T \subseteq S$ ein Supremum $\sup\{T\} \in S$ besitzt.

Satz II.6. Es gibt einen eindeutigen total geordneten Körper $\mathbb{F} \supseteq \mathbb{Q}$, der das Supremumsaxiom erfüllt und in dem für alle $a, b, c \in \mathbb{F}$: $a < b \Rightarrow a + c < b + c$ und $a, b > 0 \Rightarrow a \cdot b > 0$ gelten. Wir nennen diesen Körper $\mathbb{F}$ die **reellen Zahlen** und bezeichnen ihn mit $\mathbb{R}$.

Bemerkungen und Beispiele.

- Der Beweis von Satz II.6 beinhaltet die Konstruktion der reellen aus den rationalen Zahlen, und für sie fehlt uns in der Vorlesung leider die Zeit. Diese Konstruktion ist aber ein Juwel der Mathematik und soll niemandem vorenthalten werden; wir fügen sie als Ergänzung zu diesem Kapitel unten an.
- Die komplexen Zahlen $\mathbb{C}$ beinhalten die reellen Zahlen $\mathbb{R}$, wenn wir $\mathbb{R}$ mit $\operatorname{Re}\{\mathbb{C}\} = \{z \in \mathbb{C} \mid \operatorname{Im}(z) = 0\}$ den komplexen Zahlen mit verschwindendem Imaginärteil identifizieren. Und auf $\mathbb{C}$ lässt sich auch eine Ordnungsrelation „ $\prec$ “ definieren, bezüglich der $\mathbb{C}$ total geordnet ist. Mit dieser Ordnungsrelation ist jedoch das Supremumsaxiom verletzt oder die Eigenschaften $a \prec b \Rightarrow a + c \prec b + c$ und $a, b \succ 0 \Rightarrow a \cdot b \succ 0$ gelten nicht für alle $a, b, c \in \mathbb{C}$. Nur die reellen Zahlen $\mathbb{R}$ besitzen alle diese Eigenschaften.
- Die komplexen Zahlen $\mathbb{C}$ haben wir im vergangenen Semester in der Vorlesung *Lineare Algebra für Elektrotechnik* definiert und ihre Eigenschaften dort ausführlich diskutiert. Wir wiederholen dies nicht verweisen dazu auf das Vorlesungsmanuskript.

Die im folgenden Lemma festgehaltene Konsequenz des Supremumsaxioms spielt in den Beweisen der Analysis eine zentrale Rolle

Lemma II.7. *Seien $T_1 \subseteq \mathbb{R}$ eine nach oben und $T_2 \subseteq \mathbb{R}$ eine nach unten beschränkte Menge reeller Zahlen. Dann gelten*

$$(i) \quad \forall \varepsilon > 0 \exists t_1 \in T_1 : \quad \sup T_1 - \varepsilon < t_1, \quad (\text{II.15})$$

$$(ii) \quad \forall \varepsilon > 0 \exists t_2 \in T_2 : \quad \inf T_2 + \varepsilon > t_2. \quad (\text{II.16})$$

Beweis. Folgt unmittelbar aus den Definitionen von Supremum und Infimum. □

Bemerkungen und Beispiele.

- Lemma II.7 besagt, dass man das Supremum $\sup(T_1)$ einer nach oben beschränkten Teilmenge $T_1 \subseteq \mathbb{R}$ und das Infimum $\inf(T_2)$ einer nach unten beschränkten Teilmenge $T_2 \subseteq \mathbb{R}$ durch Elemente aus T_1 und T_2 beliebig gut nähern kann: Zu vorgegebenem $\varepsilon > 0$ -und sei es auch noch so klein- findet man $t_1 \in T_1$ und $t_2 \in T_2$, so dass

$$0 \leq \sup(T_1) - t_1 < \varepsilon \quad \text{und} \quad 0 \leq t_2 - \inf(T_2) < \varepsilon. \quad (\text{II.17})$$

- Es sind $\sup[0, 5] = \sup(0, 5) = 5$ und $\inf[0, 5] = \inf(0, 5) = 0$.
- Mit $T := \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$ sind $\sup\{T\} = 1$ und $\inf\{T\} = 0$.

II.3. Ergänzung: Konstruktion der Reellen Zahlen

In dieser Ergänzung wollen wir die Konstruktion der reellen Zahlen vorstellen. Dabei kann man verschiedene Wege beschreiten - etwa den über die Dedekindschen Schnitte. Wir wählen jedoch einen andere Weg, der in die Funktionalanalysis weist: Wir führen die reellen Zahlen als Äquivalenzklassen von Cauchy-Folgen ein - ganz analog zur Vervollständigung normierter Vektorräume.

II.3.1. Äquivalenz rationaler Cauchy-Folgen

Definition II.8. Sei

$$\mathcal{R} := \left\{ (a_n)_{n=1}^{\infty} \in \mathbb{Q}^{\mathbb{N}} \mid \forall k \in \mathbb{N} \exists n_0 \in \mathbb{N} \forall m, n \geq n_0 : |a_m - a_n| \leq 10^{-k} \right\} \quad (\text{II.18})$$

die Menge aller **rationalen Cauchy-Folgen**. Wir definieren eine Relation $\sim: \mathcal{R} \times \mathcal{R} \rightarrow \{w, f\}$ durch

$$(a_n)_{n=1}^{\infty} \sim (b_n)_{n=1}^{\infty} \quad :\Leftrightarrow \quad \forall k \in \mathbb{N} \exists n_0 \in \mathbb{N} \forall m, n \geq n_0 : |a_m - b_n| \leq 10^{-k}. \quad (\text{II.19})$$

Gilt (II.18), so heißen $\underline{a} := (a_n)_{n=1}^{\infty}$ und $\underline{b} := (b_n)_{n=1}^{\infty}$ **äquivalent**.

Lemma II.9. $\sim: \mathcal{R} \times \mathcal{R} \rightarrow \{w, f\}$ ist eine Äquivalenzrelation.

Beweis. Seien $\underline{a} := (a_n)_{n=1}^{\infty}$, $\underline{b} := (b_n)_{n=1}^{\infty}$, $\underline{c} := (c_n)_{n=1}^{\infty} \in \mathcal{R}$.

Reflexivität: $\underline{a} \sim \underline{a}$ ist gleichwertig mit $\underline{a} \in \mathcal{R}$.

Symmetrie: Offensichtlich ist $\underline{a} \sim \underline{b}$ gleichwertig mit $\underline{b} \sim \underline{a}$.

Transitivität: Sei $k \in \mathbb{N}$, und gelten $\underline{a} \sim \underline{b}$ und $\underline{b} \sim \underline{c}$. Dann gibt es $n'_0, n''_0 \in \mathbb{N}$, so dass

$$\forall m, n \geq n'_0 : |a_m - b_n| \leq 10^{-k-1}, \quad (\text{II.20})$$

$$\forall n, \ell \geq n''_0 : |b_n - c_\ell| \leq 10^{-k-1}. \quad (\text{II.21})$$

Mit $n_0 := \max(n'_0, n''_0)$ ist damit

$$\forall m, \ell \geq n_0 : |a_m - c_\ell| \leq |a_m - b_m| + |b_m - c_\ell| \leq 10^{-k-1} + 10^{-k-1} \leq 10^{-k}. \quad (\text{II.22})$$

und daher gilt auch $\underline{a} \sim \underline{c}$. □

Bemerkungen und Beispiele.

- Somit zerfällt $\mathcal{R}$ in disjunkte Äquivalenzklassen, die wir als **reelle Zahlen** bezeichnen,

$$\mathbb{R} := \mathcal{R} / \sim. \quad (\text{II.23})$$

- Ist $\underline{a} \in \mathcal{R}$, so bezeichnen wir die zugehörige Äquivalenzklasse mit $[a]$.

- Die rationalen Zahlen sind in folgender Weise in $\mathbb{R}$ eingebettet: Zu $q \in \mathbb{Q}$ betrachten wir die konstante Folge $(q, q, q, \dots) = (q)_{n=1}^\infty$. Offenbar ist $(q)_{n=1}^\infty \in \mathcal{R}$, und wir identifizieren $q \in \mathbb{Q}$ mit der Äquivalenzklasse $[(q)_{n=1}^\infty] \in \mathbb{R}$.
- Wir schreiben insbesondere $[0] := [(0)_{n=1}^\infty]$.

Lemma II.10. Sei $\underline{a} = (a_n)_{n=1}^\infty \in \mathcal{R}$ und $[\underline{a}] \neq [0]$. Dann gilt

$$\text{entweder } [\underline{a}] > [0] \quad :\Leftrightarrow \quad \exists L, n_0 \in \mathbb{N} \quad \forall n \geq n_0 : a_n \geq 10^{-L}, \quad (\text{II.24})$$

$$\text{oder } [\underline{a}] < [0] \quad :\Leftrightarrow \quad \exists L, n_0 \in \mathbb{N} \quad \forall n \geq n_0 : a_n \leq -10^{-L}. \quad (\text{II.25})$$

Beweis. Da $[\underline{a}] \neq [0]$, ist $(a_n)_{n=1}^\infty \not\sim (0)_{n=1}^\infty$, und es gilt

$$\exists L \in \mathbb{N} \quad \forall N_0 \in \mathbb{N} \quad \exists \tilde{n} \geq N_0 : |a_{\tilde{n}} - 0| = |a_{\tilde{n}}| \geq 10^{-L+1}. \quad (\text{II.26})$$

Wählen wir nun $k := L + 1$, so ergibt sich aus der Tatsache, dass $\underline{a} \in \mathcal{R}$,

$$\exists n_0 \in \mathbb{N} \quad \forall m, n \geq n_0 : |a_m - a_n| \leq 10^{-L}. \quad (\text{II.27})$$

Jetzt wählen wir $N_0 := n_0$ in (II.26) und erhalten

$$\exists \tilde{n} \geq n_0 : |a_{\tilde{n}}| \geq 10^{-L+1}. \quad (\text{II.28})$$

Wir setzen dann $n := \tilde{n}$ in (II.27) und sehen, dass

$$\forall m \geq n_0 : |a_m| \geq |a_{\tilde{n}}| - |a_m - a_{\tilde{n}}| \geq 10^{-L+1} - 10^{-L} = 9 \cdot 10^{-L}. \quad (\text{II.29})$$

Ist $a_{n_0} > 0$, so ist $a_{n_0} = |a_{n_0}|$ und nach (II.26) gilt für alle $m \geq n_0$:

$$\begin{aligned} a_m &= a_{n_0} + a_m - a_{n_0} = |a_{n_0}| + a_m - a_{n_0} = |a_{n_0}| - |a_m - a_{n_0}| \\ &\geq 9 \cdot 10^{-L} - 10^{-L} = 8 \cdot 10^{-L} \geq 10^{-L}. \end{aligned} \quad (\text{II.30})$$

Ist umgekehrt $a_{n_0} < 0$, so ist $a_{n_0} = -|a_{n_0}|$, und analog folgt

$$\forall m \geq n_0 : a_m \leq -10^{-L}. \quad (\text{II.31})$$

□

Lemma II.11. Seien $\underline{a} := (a_n)_{n=1}^\infty, \underline{b} := (b_n)_{n=1}^\infty \in \mathcal{R}$. Dann gilt

$$(\underline{a} \sim \underline{b}) \quad \Leftrightarrow \quad (\underline{a} - \underline{b} \sim \underline{0}), \quad (\text{II.32})$$

wobei $\underline{a} - \underline{b} := (a_n - b_n)_{n=1}^\infty$.

Beweis.

" $\Rightarrow$ ": Ist $\underline{a} \sim \underline{b}$, so gilt insbesondere

$$\forall k \in \mathbb{N} \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : |a_n - b_n| \leq 10^{-k}, \quad (\text{II.33})$$

also auch

$$\forall k \in \mathbb{N} \exists n_0 \in \mathbb{N} \forall m, n \geq n_0 : |a_m - b_n| \leq 10^{-k}, \quad (\text{II.34})$$

wobei 0 das m . Element der Folge $\underline{0}$ ist. Mit anderen Worten: $(\underline{a} - \underline{b}) \sim \underline{0}$.

” $\Leftarrow$ ”: Ist umgekehrt $(\underline{a} - \underline{b}) \sim \underline{0}$, so gilt (II.33). Weiterhin ist $(a_n)_{n=1}^\infty \in \mathcal{R}$, und daher gibt es $n'_0 \in \mathbb{N}$, so dass

$$\forall m, n \geq n'_0 : |a_m - a_n| \leq 10^{-k}. \quad (\text{II.35})$$

Für $m, n \geq \max(n_0, n'_0)$ erhalten wir somit

$$|a_m - b_n| \leq |a_m - a_n| + |a_n - b_n| \leq 2 \cdot 10^{-k}, \quad (\text{II.36})$$

d.h. $\underline{a} \sim \underline{b}$. □

II.3.2. Die Grundrechenarten

Definition II.12. Seien $\underline{a} = (a_n)_{n=1}^\infty$, $\underline{b} = (b_n)_{n=1}^\infty \in \mathcal{R}$. Wir definieren

$$\underline{a} \pm \underline{b} := (a_n \pm b_n)_{n=1}^\infty, \quad (\text{II.37})$$

$$\underline{a} \cdot \underline{b} := (a_n \cdot b_n)_{n=1}^\infty \quad (\text{II.38})$$

und, für $\underline{b} \not\sim \underline{0}$,

$$\frac{1}{\underline{b}} := \left(\frac{1}{b_n + \text{sgn}(b_n) \cdot 10^{-n}} \right)_{n=1}^\infty, \quad (\text{II.39})$$

wobei $\text{sgn} : \mathbb{Q} \rightarrow \{-1, 1\}$,

$$\text{sgn}(q) := \begin{cases} 1, & \text{falls } q \geq 0, \\ -1, & \text{falls } q < 0. \end{cases} \quad (\text{II.40})$$

Lemma II.13. Seien $\underline{a} = (a_n)_{n=1}^\infty$, $\underline{b} = (b_n)_{n=1}^\infty \in \mathcal{R}$. Dann sind $\underline{a} \pm \underline{b}$, $\underline{a} \cdot \underline{b} \in \mathcal{R}$, und für $\underline{b} \not\sim \underline{0}$ ist auch $\frac{1}{\underline{b}} \in \mathcal{R}$.

Beweis. Da $\underline{a}, \underline{b} \in \mathcal{R}$, gibt es zu $k \in \mathbb{N}$ zwei Zahlen $n'_0, n''_0 \in \mathbb{N}$, sodass

$$\forall m, n \geq n'_0 : |a_m - a_n| \leq 10^{-k-1}, \quad (\text{II.41})$$

$$\forall m, n \geq n''_0 : |b_m - b_n| \leq 10^{-k-1}. \quad (\text{II.42})$$

somit gilt $\forall m, n \geq \max(n'_0, n''_0)$:

$$\begin{aligned} |(a_m \pm b_m) - (a_n \pm b_n)| &= |(a_m - a_n) \pm (b_m - b_n)| \\ &\leq |a_m - a_n| + |b_m - b_n| \leq 10^{-k-1} + 10^{-k-1} \leq 10^{-k}, \end{aligned} \quad (\text{II.43})$$

also $\underline{a} \pm \underline{b} \in \mathcal{R}$. Mit $k := 1$ folgt aus $\underline{a}, \underline{b} \in \mathcal{R}$, dass

$$\exists n_0 \in \mathbb{N} \forall m, n \geq n_0 : |a_m - a_n|, |b_m - b_n| \leq 10^{-1}. \quad (\text{II.44})$$

Setzen wir $m := n_0$, so folgt daraus, dass

$$\forall m, n \geq n_0 : |a_m| + |b_m| \leq |a_{n_0}| + |b_{n_0}| + 2 \cdot 10^{-1} \leq |a_{n_0}| + |b_{n_0}| + 1. \quad (\text{II.45})$$

Mit

$$C := \left(\sum_{r=1}^{n_0} |a_r| + |b_r| \right) + 1, \quad (\text{II.46})$$

erhalten wir

$$\sup_{n \in \mathbb{N}} |a_n| \leq C, \quad \sup_{n \in \mathbb{N}} |b_n| \leq C. \quad (\text{II.47})$$

Wir beobachten nun, dass

$$\begin{aligned} |a_m b_m - a_n b_n| &= |(a_m - a_n) b_m + a_n (b_m - b_n)| \\ &\leq |a_m - a_n| \cdot |b_m| + |a_n| \cdot |b_m - b_n| \\ &\leq C \cdot (|a_m - a_n| + |b_m - b_n|). \end{aligned} \quad (\text{II.48})$$

Seien nun $k \in \mathbb{N}$ und $\ell \in \mathbb{N}$ so, dass $C \leq 10^\ell$. Dann gibt es $n'_0, n''_0 \in \mathbb{N}$, so dass

$$\forall m, n \geq n'_0 : |a_m - a_n| \leq 10^{-k-\ell}, \quad (\text{II.49})$$

$$\forall m, n \geq n''_0 : |b_m - b_n| \leq 10^{-k-\ell}. \quad (\text{II.50})$$

Also ist mit (II.48)

$$\forall m, n \geq \max(n'_0, n''_0) : |a_m b_m - a_n b_n| \leq 10^{-k}, \quad (\text{II.51})$$

d.h. $\underline{a} \cdot \underline{b} \in \mathcal{R}$.

Ist schließlich $\underline{b} \sim \underline{0}$, so existieren nach Lemma II.10 zwei Zahlen $L, n'_0 \in \mathbb{N}$, sodass

$$\forall n \geq n'_0 : |b_n| \geq 10^{-L}. \quad (\text{II.52})$$

Für $k \in \mathbb{N}$ gibt es $n''_0 \in \mathbb{N}$, so dass

$$\forall m, n \geq n''_0 : |b_m - b_n| \leq 10^{-k-1-2L}, \quad (\text{II.53})$$

und so erhalten wir für alle $m, n \geq \max(n_0, n'_0) + 2L + k + 1 =: n_0$.

$$\begin{aligned} &|(b_m + \operatorname{sgn}(b_m) \cdot 10^{-m})^{-1} - (b_n + \operatorname{sgn}(b_n) \cdot 10^{-n})^{-1}| \\ &= \left(\frac{1}{|b_m| + 10^{-m}} \right) \cdot \left(\frac{1}{|b_n| + 10^{-n}} \right) \cdot |b_m - b_n + \operatorname{sgn}(b_m) 10^{-m} - \operatorname{sgn}(b_n) 10^{-n}| \\ &\leq 10^{-2L} \cdot (10^{-k-1-2L} + 10^{-k-1-2L} + 10^{-k-1-2L}) \leq 3 \cdot 10^{-k-1} \leq 10^{-k}. \end{aligned} \quad (\text{II.54})$$

Also ist $\frac{1}{\underline{b}} \in \mathcal{R}$. □

Lemma II.14. Seien $\underline{a} = (a_n)_{n=1}^\infty$, $\hat{\underline{a}} = (\hat{a}_n)_{n=1}^\infty$, $\underline{b} = (b_n)_{n=1}^\infty$, $\hat{\underline{b}} = (\hat{b}_n)_{n=1}^\infty \in \mathcal{R}$ mit $\underline{a} \sim \hat{\underline{a}}$ und $\underline{b} \sim \hat{\underline{b}}$. Dann sind

$$\underline{a} \pm \underline{b} \sim \hat{\underline{a}} \pm \hat{\underline{b}} \quad (\text{II.55})$$

$$\underline{a} \cdot \underline{b} \sim \hat{\underline{a}} \cdot \hat{\underline{b}} \quad (\text{II.56})$$

und, falls $\underline{b} \not\sim \underline{0}$,

$$\frac{1}{\underline{b}} \sim \frac{1}{\hat{\underline{b}}}. \quad (\text{II.57})$$

Beweis. Wie in Lemma II.13 gibt es $L \in \mathbb{N}$, so dass

$$\forall n \in \mathbb{N} : |a_n|, |\hat{a}_n|, |b_n|, |\hat{b}_n| \leq 10^L. \quad (\text{II.58})$$

Zu $k \in \mathbb{N}$ gibt es dann $n_0 \in \mathbb{N}$, so dass

$$\forall m, n \geq n_0 : |a_m - \hat{a}_n|, |b_m - \hat{b}_n| \leq 10^{-k-1-L}. \quad (\text{II.59})$$

Also sind für alle $m, n \geq n_0$

$$\begin{aligned} |(a_m \pm b_m) - (\hat{a}_n \pm \hat{b}_n)| &= |(a_m - \hat{a}_n) \pm (b_m - \hat{b}_n)| \\ &\leq 2 \cdot 10^{-k-1-L} \leq 10^{-k} \end{aligned} \quad (\text{II.60})$$

und

$$\begin{aligned} |a_m \cdot b_m - \hat{a}_n \cdot \hat{b}_n| &= |(a_m - \hat{a}_n)b_m + \hat{a}_n(b_m - \hat{b}_n)| \\ &\leq |b_m| \cdot |a_m - \hat{a}_n| + |\hat{a}_n| \cdot |b_m - \hat{b}_n| \\ &\leq 2 \cdot 10^L \cdot 10^{-k-1-L} \leq 10^{-k}. \end{aligned} \quad (\text{II.61})$$

Dies beweist (II.55) und (II.56). Glg (II.57) ist analog. $\square$

Korollar II.15. Die Verknüpfungen $+$, $-$, $\cdot$: $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$,

$$[\underline{a}] \pm [\underline{b}] := [\underline{a} \pm \underline{b}], \quad (\text{II.62})$$

$$[\underline{a}] \cdot [\underline{b}] := [\underline{a} \cdot \underline{b}], \quad (\text{II.63})$$

und die Abbildung $1/(\cdot) : \mathbb{R} \setminus \{[0]\} \rightarrow \mathbb{R}$,

$$\frac{1}{[\underline{b}]} := \left[\frac{1}{\underline{b}} \right] \quad (\text{II.64})$$

sind wohldefiniert, d.h. unabhängig von den gewählten Repräsentanten $\underline{a}, \underline{b} \in \mathcal{R}$.

Lemma II.16. $\mathbb{R}$ ist bezüglich der in (II.62)–(II.63) und durch

$$\forall [\underline{a}] \in \mathbb{R}, [\underline{b}] \in \mathbb{R} \setminus \{[0]\} : \frac{\underline{a}}{\underline{b}} := \left[\underline{a} \cdot \frac{1}{\underline{b}} \right] \quad (\text{II.65})$$

definierten Verknüpfungen ein Körper.

Beweis. Nachprüfen der Körperaxiome; dabei sind $[0]$ und $[1] = [(1)_{n=1}^\infty]$ die neutralen Elemente der Addition bzw. Multiplikation. $\square$

Definition II.17. Für je zwei Elemente $[a], [b] \in \mathbb{R}$ mit $[a] \neq [b]$ definieren wir

$$[a] < [b] \quad :\Leftrightarrow [a - b] < 0, \quad (\text{II.66})$$

$$[a] > [b] \quad :\Leftrightarrow [a - b] > 0, \quad (\text{II.67})$$

wobei die rechten Seiten in Lemma II.10 definiert sind.

Lemma II.18. $\mathbb{R}$ ist bezüglich " $<$ " total geordnet, und für alle $a, b, c \in \mathbb{R}$ gelten $a < b \Rightarrow a + c < b + c$ und $a, b > 0 \Rightarrow a \cdot b > 0$.

Beweis. Nachprüfen der Definition einer totalen Ordnung und der Eigenschaften $a < b \Rightarrow a + c < b + c$ und $a, b > 0 \Rightarrow a \cdot b > 0$. $\square$

II.3.3. Das Supremumsaxiom

Satz II.19. $\mathbb{R}$ erfüllt das Supremumsaxiom.

Beweis. Sei $A \subseteq \mathbb{R}$ eine durch $[\underline{c}] \in \mathbb{R}$, $\underline{c} = (c_n)_{n=1}^\infty \in \mathcal{R}$ nach oben beschränkte Teilmenge, wobei wir der Einfachheit halber annehmen, dass $A \cap \mathbb{R}^+ \neq \emptyset$.

Da $\underline{c} \in \mathcal{R}$, ist auch $\underline{c}$ beschränkt, d.h. es gibt ein $N \in \mathbb{N}$, sodass

$$\forall [a] \in A : [a] < [(10^N)_{n=1}^\infty]. \quad (\text{II.68})$$

Wir setzen nun $b_1 := \beta_1 \cdot 10^{N-1}$, wobei $\beta_1 \in \{0, 1, \dots, 9\}$ so gewählt ist, dass

$$\exists x_1 \in A : x_1 \geq [(\beta_1 \cdot 10^{N-1})_{n=1}^\infty] \text{ und} \quad (\text{II.69})$$

$$\forall x \in A : x < [((\beta_1 + 1) \cdot 10^{N-1})_{n=1}^\infty]. \quad (\text{II.70})$$

Anschließend setzen wir $b_2 := b_1 + \beta_2 \cdot 10^{N-2}$, wobei $\beta_2 \in \{0, 1, \dots, 9\}$ so gewählt ist, dass

$$\exists x_2 \in A : x_2 \geq [(b_2)_{n=1}^\infty] \text{ und} \quad (\text{II.71})$$

$$\forall x \in A : x < [(b_2 + 10^{N-2})_{n=1}^\infty], \quad (\text{II.72})$$

u.s.w. Für allgemeines $k \in \mathbb{N}$ setzen wir $b_k := b_{k-1} + \beta_k \cdot 10^{N-k}$, wobei $\beta_k \in \{0, 1, \dots, 9\}$ so gewählt ist, dass

$$\exists x_k \in A : x_k \geq [(b_k)_{n=1}^\infty] \text{ und} \quad (\text{II.73})$$

$$\forall x \in A : x < [(b_k + 10^{N-k})_{n=1}^\infty]. \quad (\text{II.74})$$

Die so gebildete Folge $(b_n)_{n=1}^\infty =: \underline{b}$ ist offenbar eine Cauchy-Folge, denn für $m > n \geq n_0$ ist

$$|b_m - b_n| = \left| \sum_{\ell=n+1}^m \beta_\ell \cdot 10^{N-\ell} \right| \leq 10^{N-n} \leq 10^{N-n_0}. \quad (\text{II.75})$$

Sei nun $\underline{d} := (d_n)_{n=1}^\infty \in \mathcal{R}$ mit $[\underline{d}] < [\underline{b}]$, d.h. $[\underline{b} - \underline{d}] > [0]$. Nach Lemma II.10 gibt es dann $m_0, M \in \mathbb{N}$, sodass

$$\forall m > n \geq m_0 : \quad d_m < b_m - 10^{-M} \leq b_n - (10^{-M} - 10^{N-n}), \quad (\text{II.76})$$

wie in (II.75). Für $n_0 := m_0 + N + M + 1$ ergibt dies

$$\forall m > n_0 : \quad d_m \geq b_{n_0} - 10^{-M-1}, \quad (\text{II.77})$$

was

$$[\underline{d}] \leq \left[(b_{n_0} - 10^{-M-1})_{n=1}^\infty \right] \leq x_{n_0} - \left[(10^{-M-1})_{n=1}^\infty \right] < x_{n_0} \in A \quad (\text{II.78})$$

impliziert. Also ist $[\underline{d}]$ keine obere Schranke an A .

Sei schließlich $[\underline{a}] \in A$, und nehmen wir an, dass $[\underline{a}] > [\underline{b}]$, also $[\underline{a} - \underline{b}] > [0]$. Dann gibt es $m_0, M \in \mathbb{N}$ so, dass

$$\forall m > k \geq m_0 : \quad a_m > b_m + 10^{-M} \geq b_k + 10^{-M} - 10^{N-k}. \quad (\text{II.79})$$

Für $k := M + N + 1 + m_0$ bedeutet dies, dass für alle $m > k$

$$\begin{aligned} a_m &\geq b_k + 10^{-M} - 10^{-M-1} \geq b_k + 9 \cdot 10^{-M-1} \\ &= b_k + 9 \cdot 10^{m_0+N-1} \geq b_k + 10^{N-k}, \end{aligned} \quad (\text{II.80})$$

was

$$[\underline{a}] \geq \left[(b_k + 10^{N-k})_{k=1}^\infty \right] \quad (\text{II.81})$$

impliziert. Gleichung (II.81) steht jedoch in Widerspruch zu (II.74). Also gilt $[\underline{a}] \leq [\underline{b}]$.

Zusammenfassend erhalten wir

$$\forall [\underline{a}] \in A : \quad [\underline{a}] \leq [\underline{b}], \quad (\text{II.82})$$

$$\forall [\underline{d}] < [\underline{b}] \exists [\underline{a}] \in A : \quad [\underline{d}] < [\underline{a}]. \quad (\text{II.83})$$

Also ist $[\underline{b}] = \sup A$. □

III. Endliche, abzählbare und überabzählbare Mengen

Wir haben schon einige Mengen in den Kapiteln I und II kennengelernt, etwa die Zahlenmengen $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ und $\mathbb{R}$. Jede dieser Zahlenmengen enthält unendlich viele Elemente,

$$|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = |\mathbb{R}| = \infty. \quad (\text{III.1})$$

Um mengentheoretische Unterschiede zwischen ihnen auszumachen, müssen wir unsere bisherigen Begriffe über Mengen verfeinern.

III.1. Abzählbare Mengen

Definition III.1. Zwei Mengen A, B heißen **gleichmächtig**

$$:\Leftrightarrow \exists f : A \rightarrow B : f \text{ ist bijektiv.} \quad (\text{III.2})$$

Bemerkungen und Beispiele.

- Jede N -elementige Menge, $\{a_1, a_2, \dots, a_N\}$ ist gleichmächtig zu $\{1, 2, \dots, N\}$, denn

$$f : \{1, 2, \dots, N\} \rightarrow \{a_1, a_2, \dots, a_N\}, \quad k \mapsto a_k, \quad (\text{III.3})$$

ist eine Bijektion.

- $\mathbb{N}$ ist gleichmächtig zu $\mathbb{Z}$, denn

$$f : \mathbb{N} \rightarrow \mathbb{Z}, \quad k \mapsto \begin{cases} 0, & \text{falls } k = 1, \\ \frac{1}{2}k, & \text{falls } k \text{ gerade} \wedge k \geq 2, \\ -\frac{1}{2}(k-1), & \text{falls } k \text{ ungerade} \wedge k \geq 2, \end{cases} \quad (\text{III.4})$$

ist eine Bijektion,

$$f(1) = 0, \quad f(2) = 1, \quad f(3) = -1, \quad f(4) = 2, \quad f(5) = -2, \quad \dots \quad (\text{III.5})$$

Definition III.2. Eine Menge A heißt **abzählbar** $:\Leftrightarrow$

$$(a) \quad A \text{ ist } \mathbf{endlich}, \text{ d.h. } |A| \in \mathbb{N} (\Leftrightarrow |A| < \infty) \text{ oder} \quad (\text{III.6})$$

$$(b) \quad A \text{ ist } \mathbf{unendlich}, |A| = \infty, \text{ und } A \text{ ist gleichmächtig zu } \mathbb{N} \quad (\text{III.7})$$

Ist A nicht abzählbar, so heißt A **überabzählbar**.

Lemma III.3. Seien A eine abzählbare Menge und $B \subseteq A$. Dann ist B abzählbar.

Satz III.4. $\mathbb{N} \times \mathbb{N}$ ist abzählbar.

Beweis. Wir zeichnen die Elemente $(m, n) \in \mathbb{N} \times \mathbb{N}$ in eine Tabelle und definieren eine Abbildung $J : \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$, indem wir den Pfeilen folgen,

$$\begin{array}{ccccc}
 J(1) := (1, 1) & \rightarrow & J(2) := (1, 2) & & J(6) := (1, 3) \quad \dots \\
 & & \swarrow & & \nearrow \\
 J(3) := (2, 1) & & J(5) := (2, 2) & & \\
 \downarrow & & \nearrow & & \\
 J(4) := (3, 1) & & & &
 \end{array} \tag{III.8}$$

Die Bijektivität von J ist offensichtlich. □

Satz III.5. $\mathbb{Q}$ ist abzählbar.

Beweis. Wir führen den Beweis für $\mathbb{Q}_+$. Jedes Element $q \in \mathbb{Q}_+$ lässt sich eindeutig durch $(m, n) \in \mathbb{N} \times \mathbb{N}$ als

$$q = \frac{m}{n} \tag{III.9}$$

darstellen, wenn man voraussetzt, dass m und n teilerfremd sind (d.h. man kann m/n nicht kürzen). Also ist

$$J : \mathbb{Q}_+ \rightarrow A := \{(m, n) \in \mathbb{N} \times \mathbb{N} \mid m, n \text{ teilerfremd}\}, \quad \frac{m}{n} \mapsto (m, n), \tag{III.10}$$

eine Bijektion. Nach Satz III.4 ist $\mathbb{N} \times \mathbb{N}$ und nach Lemma III.3 somit auch $A \subseteq \mathbb{N} \times \mathbb{N}$ abzählbar. Also ist $\mathbb{Q}_+$ abzählbar. □

III.2. Dezimaldarstellung von Zahlen

Definition III.6. Eine **Folge (in A)** ist eine Abbildung

$$a : \mathbb{N} \rightarrow A, \quad n \mapsto a_n, \tag{III.11}$$

wobei $A \neq \emptyset$ eine nichtleere Menge ist. Statt (III.11) schreibt man auch $a_1, a_2, a_3, \dots$ oder $(a_n)_{n=1}^\infty$. Die Menge aller Folgen in A bezeichnet man mit $A^\mathbb{N}$.

Wir wollen nun die Dezimaldarstellung von Zahlen zwischen 0 und 1 genauer untersuchen.

Definition III.7. Eine Folge $a : \mathbb{N} \rightarrow \{0, 1, 2, \dots, 9\}$ heißt **Dezimaldarstellung** der Zahl $0, a_1 a_2 a_3 \dots$

$$:\Leftrightarrow \quad \forall n \in \mathbb{N} \exists k \in \mathbb{N}, k > n : \quad a_k < 9. \tag{III.12}$$

Die Menge der Dezimaldarstellungen bezeichnen wir mit $\mathcal{D}$.

Die Bedingung (III.12) schließt Zahlen mit Periode 9, wie z.B.

$$x = 0,1729999 \dots \quad (\text{III.13})$$

aus, denn diese ist ja bereits durch

$$0,173000 \dots$$

dezimal dargestellt, und wir möchten, dass die Dezimaldarstellung eindeutig ist, siehe auch Abschnitt III.3.3.

Lemma III.8. *Die Menge der Dezimaldarstellungen $\mathcal{D}$ und die Menge $[0, 1) := \{x \in \mathbb{R} \mid 0 \leq x < 1\}$ der reellen Zahlen zwischen Null und Eins sind gleichmächtig.*

Beweis. Den ausführlichen Beweis findet man in den Ergänzungen im Abschnitt III.3.3. Wir bemerken hier nur, dass eine Bijektion $\tilde{J} : \mathcal{D} \rightarrow [0, 1)$ durch $\tilde{J}[(a_n)_{n=1}^\infty] := \sum_{n=1}^\infty a_n \cdot 10^{-n}$ gegeben ist. Allerdings wissen wir noch nicht, wie man unendlich viele Summanden addiert und müssen den Beweis in Abschnitt III.3.3 den Mitteln führen, die uns zur Verfügung stehen, nämlich dem Supremumsaxiom. $\square$

Satz III.9. *$\mathcal{D}$ ist nicht abzählbar.*

Beweis. Nehmen wir an,

$$\mathcal{D} = \{(a_n^{(1)})_{n \in \mathbb{N}}, (a_n^{(2)})_{n \in \mathbb{N}}, \dots\} \quad (\text{III.14})$$

wäre eine Abzählung. Definiere nun eine Folge $(b_n)_{n \in \mathbb{N}}$ durch

$$b_n := \begin{cases} 1, & \text{falls } a_n^{(n)} \in \{5, 6, 7, 8, 9\}, \\ 7, & \text{falls } a_n^{(n)} \in \{0, 1, 2, 3, 4\}, \end{cases} \quad (\text{III.15})$$

so dass

$$\forall n \in \mathbb{N} : b_n \neq a_n^{(n)}. \quad (\text{III.16})$$

Offenbar wäre

$$(b_n)_{n \in \mathbb{N}} \in \mathcal{D}, \quad (\text{III.17})$$

da alle b_n verschieden von 9 sind. Andererseits impliziert aber (III.16), dass

$$\forall k \in \mathbb{N} : (b_n)_{n \in \mathbb{N}} \neq (a_n^{(k)})_{n \in \mathbb{N}}, \quad (\text{III.18})$$

also

$$(b_n)_{n \in \mathbb{N}} \notin \{(a_n^{(1)})_{n \in \mathbb{N}}, (a_n^{(2)})_{n \in \mathbb{N}}, \dots\} = \mathcal{D}, \quad (\text{III.19})$$

was in Widerspruch zu (III.17) steht. $\square$

Satz III.10. *$\mathbb{R}$ ist überabzählbar.*

Beweis. Wäre $\mathbb{R}$ abzählbar, so wäre auch $[0, 1) \subseteq \mathbb{R}$ als Teilmenge abzählbar. $[0, 1)$ ist aber gleichmächtig zur überabzählbaren Menge $\mathcal{D}$ der Dezimaldarstellungen und somit selbst überabzählbar. Also kann $\mathbb{R}$ nicht abzählbar sein. $\square$

III.3. Ergänzungen

III.3.1. Teilmengen abzählbarer Mengen sind abzählbar

Beweis. Ist $|B|$ endlich, so gilt die Behauptung automatisch. Wir können also o.B.d.A. voraussetzen, dass B und somit auch $A \supseteq B$ unendlich sind,

$$|B| = |A| = \infty. \quad (\text{III.20})$$

Da A abzählbar ist, gibt es eine Bijektion

$$x : \mathbb{N} \rightarrow A, \quad n \mapsto x_n, \quad (\text{III.21})$$

und

$$A = \{x_1, x_2, x_3, \dots\}. \quad (\text{III.22})$$

Da $B \subseteq A$, gibt es eine Zahl $n_1 \in \mathbb{N}$, so dass

$$x_1 \notin B, x_2 \notin B, \dots, x_{n_1-1} \notin B, x_{n_1} \in B. \quad (\text{III.23})$$

Weiterhin gibt es eine Zahl $n_2 \in \mathbb{N}, n_2 > n_1$, so dass

$$x_{n_1+1} \notin B, \dots, x_{n_2-1} \notin B, x_{n_2} \in B. \quad (\text{III.24})$$

Führen wir dieses Verfahren so fort, erhalten wir eine Abbildung

$$y : \mathbb{N} \rightarrow B, \quad j \mapsto y_j := x_{n_j}, \quad (\text{III.25})$$

wobei $n_j \in \mathbb{N}, j \leq n_j < n_{j+1}$. Weil $x : \mathbb{N} \rightarrow A$ eine Bijektion ist, gilt $x_m \neq x_n$ falls $m < n$.

Insbesondere ist für $i < j$ auch $n_i < n_j$ und somit $x_{n_i} \neq x_{n_j}$,

$$i < j \Rightarrow n_i < n_j \Rightarrow y_i = x_{n_i} \neq x_{n_j} = y_j. \quad (\text{III.26})$$

Also ist y injektiv. Ist nun $b \in B \subseteq A$, so gibt es ein $n \in \mathbb{N}$ mit

$$b = x_n. \quad (\text{III.27})$$

Mit der oben beschriebenen Prozedur erhält man dann $n = n_j$, für ein gewisses $j \in \mathbb{N}, j \leq n$ (nach endlich vielen Schritten, das ist hier der Punkt!), also

$$b = x_{n_j} = y_j. \quad (\text{III.28})$$

Somit ist y auch surjektiv und damit bijektiv □

III.3.2. Abzählbare Vereinigungen abzählbarer Mengen sind abzählbar

Satz III.11. *Jede Vereinigung abzählbar vieler abzählbarer Mengen ist abzählbar.*

Beweis. Seien die Mengen $A_1, A_2, A_3, \dots$ gegeben als

$$A_j = \{x_{j,1}, x_{j,2}, x_{j,3}, \dots\}. \quad (\text{III.29})$$

Dann ist

$$A := \bigcup_{j \in \mathbb{N}} A_j = \{x_{j,k} \mid \exists (j,k) \in \mathbb{N} \times \mathbb{N} : x_{j,k} \in A_j\}. \quad (\text{III.30})$$

Wie im Beweis von Satz III.5 folgern wir nun, dass A gleichmächtig zu einer Teilmenge von $\mathbb{N} \times \mathbb{N}$ und somit abzählbar ist. $\square$

III.3.3. Beweis der Gleichmächtigkeit der reellen Zahlen und der Dezimaldarstellungen (Lemma III.8)

Beweis. Seien $(a_n)_{n=1}^\infty \in \mathcal{D}$ eine Dezimaldarstellung und, für $N \in \mathbb{N}$,

$$q_N := \sum_{n=1}^N a_n \cdot 10^{-n} \in \mathbb{Q} \subseteq \mathbb{R}. \quad (\text{III.31})$$

Offensichtlich gilt immer $q_N < 1$, deshalb ist die Menge

$$A := \{q_1, q_2, q_3, \dots\} \subseteq \mathbb{R} \quad (\text{III.32})$$

nach oben beschränkt und hat ein Supremum $\sup A \in [0, 1]$. Wegen (III.12) gilt sogar $q_N \leq 1 - 10^{-\tilde{n}}$, für ein gewisses $\tilde{n} \in \mathbb{N}$, und daher

$$\sup A \in [0, 1). \quad (\text{III.33})$$

Wir erhalten somit eine Abbildung

$$J : \mathcal{D} \rightarrow [0, 1), \quad (a_n)_{n \in \mathbb{N}} \mapsto \sup A. \quad (\text{III.34})$$

Seien nun $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}} \in \mathcal{D}$ und $(a_n)_{n \in \mathbb{N}} \neq (b_n)_{n \in \mathbb{N}}$. Wir zeigen, dass dann auch $\sup A \neq \sup B$, wobei $B = \{p_1, p_2, p_3, \dots\}$ und $p_N = \sum_{n=1}^N b_n \cdot 10^{-n}$. Dann gibt es ein $m \in \mathbb{N}$, so dass

$$a_1 = b_1, \quad a_2 = b_2, \quad \dots, \quad a_{m-1} = b_{m-1}, \quad a_m \leq b_m - 1 \quad (\text{III.35})$$

(oder umgekehrt, dann vertausche man die Rollen von $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$).

Weiterhin gibt es nach (III.12) ein $\tilde{n} \geq m + 1$ mit

$$a_{\tilde{n}} \leq 8. \quad (\text{III.36})$$

Bilden wir nun q_N , für $N \geq \tilde{n}$, so gilt

$$\begin{aligned} q_N &= \sum_{n=1}^N a_n \cdot 10^{-n} = \sum_{n=1}^{m-1} b_n \cdot 10^{-n} + \sum_{n=m}^{\tilde{n}-1} a_n \cdot 10^{-n} + \underbrace{\sum_{n=\tilde{n}}^N a_n \cdot 10^{-n}}_{\leq 9 \cdot 10^{-\tilde{n}}} \\ &\leq \sum_{n=1}^{m-1} b_n \cdot 10^{-n} a_m \cdot 10^{-m} + \underbrace{\sum_{n=m+1}^{\tilde{n}} 9 \cdot 10^{-\tilde{n}}}_{=10^{-m}-10^{-\tilde{n}}}, \end{aligned} \quad (\text{III.37})$$

und andererseits, für $N \geq m$,

$$p_N := \sum_{n=1}^N b_n \cdot 10^{-n} \geq \sum_{n=1}^{m-1} b_n \cdot 10^{-n} + b_m \cdot 10^{-m}. \quad (\text{III.38})$$

Mit

$$A := \{q_1, q_2, q_3, \dots\}, \quad B := \{p_1, p_2, p_3, \dots\} \quad (\text{III.39})$$

und

$$q_1 \leq q_2 \leq q_3 \leq \dots, \quad p_1 \leq p_2 \leq p_3 \leq \dots \quad (\text{III.40})$$

erhalten wir dann

$$\forall k \in \mathbb{N} : \quad \sup A = \sup\{q_k, q_{k+1}, q_{k+2}, \dots\} \quad (\text{III.41})$$

$$\forall \ell \in \mathbb{N} : \quad \sup B = \sup\{q_\ell, q_{\ell+1}, q_{\ell+2}, \dots\}. \quad (\text{III.42})$$

Also ist

$$\begin{aligned} \sup A &= \sup\{q_{\tilde{n}}, q_{\tilde{n}+1}, q_{\tilde{n}+2}, \dots\} \\ &\leq \sum_{n=1}^{m-1} b_n \cdot 10^{-n} + (a_m + 1) \cdot 10^{-m} - 10^{-\tilde{n}} \\ &\leq \sum_{n=1}^{m-1} b_n \cdot 10^{-n} + b_m \cdot 10^{-m} - 10^{-\tilde{n}} \\ &\leq \sup\{p_m, p_{m+1}, p_{m+2}, \dots\} - 10^{-\tilde{n}} \\ &= \sup B - 10^{-\tilde{n}} < \sup B. \end{aligned} \quad (\text{III.43})$$

Insbesondere ist

$$J[(a_n)_{n \in \mathbb{N}}] = \sup A < \sup B = J[(b_n)_{n \in \mathbb{N}}], \quad (\text{III.44})$$

und somit J *injektiv*.

Um die Surjektivität zu zeigen, wählen wir eine Zahl $x_0 \in [0, 1)$ und bestimmen die natürliche Zahl $a_1 \in \{0, 1, \dots, 9\}$ so, dass

$$0 \leq x_1 := x_0 - a_1 \cdot 10^{-1} < 10^{-1}. \quad (\text{III.45})$$

Anschließend wählen wir $a_2 \in \{0, 1, \dots, 9\}$ so, dass

$$0 \leq x_2 := x_1 - a_2 \cdot 10^{-2} < 10^{-2} \quad (\text{III.46})$$

und allgemein $a_k \in \{0, \dots, 9\}$ so, dass

$$0 \leq x_k := x_{k-1} - a_k \cdot 10^{-k} < 10^{-k}, \quad (\text{III.47})$$

für vorher gewählte $a_1, a_2, \dots, a_{k-1} \in \{0, \dots, 9\}$. Auf diese Weise erhalten wir eine Folge $(a_n)_{n \in \mathbb{N}} \in \mathcal{D}$. Bilden wir wieder

$$q_N := \sum_{n=1}^N a_n 10^{-n}, \quad A := \{q_1, q_2, \dots\}, \quad (\text{III.48})$$

so ist mit (III.45)-(III.47), für alle $N, k \in \mathbb{N}$

$$q_N \leq x_0 < q_N + 10^{-N} \leq q_{N+k} + 10^{-N}. \quad (\text{III.49})$$

Also gilt, für alle $N \in \mathbb{N}$,

$$\sup A \leq x_0 < \sup\{q_N, q_{N+1}, \dots\} + 10^{-N} = \sup A + 10^{-N}. \quad (\text{III.50})$$

Aus

$$\forall N \in \mathbb{N} : \quad \sup A \leq x_0 < \sup A + 10^{-N} \quad (\text{III.51})$$

folgt aber

$$x_0 = \sup A = J[(a_n)_{n \in \mathbb{N}}], \quad (\text{III.52})$$

und J ist auch *surjektiv* und somit bijektiv. $\square$

IV. Folgen und Konvergenz

IV.1. Konvergenz von Zahlenfolgen

Wir erinnern an den Begriff der Zahlenfolge, den wir schon im Kapitel I eingeführt haben. Im Allgemeinen ist eine **Folge** $(a_n)_{n=1}^\infty \in A^\mathbb{N}$ in A eine Abbildung $a_{(\cdot)} : \mathbb{N} \rightarrow A$, $n \mapsto a_n$. Folgen mit Werten in $A \subseteq \mathbb{K}$ nennen wir **Zahlenfolgen**.

Definition IV.1.

- (i) Eine Zahlenfolge $(a_n)_{n=1}^\infty \in \mathbb{K}^\mathbb{N}$ heißt **konvergent** $:\Leftrightarrow$

$$\exists a \in \mathbb{K} \forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : |a_n - a| \leq \varepsilon. \quad (\text{IV.1})$$

In diesem Fall heißt a **Grenzwert** oder **Limes** von $(a_n)_{n=1}^\infty$, und wir schreiben statt (IV.1) auch abkürzend

$$a_n \rightarrow a, \quad n \rightarrow \infty, \quad (\text{IV.2})$$

oder

$$\lim_{n \rightarrow \infty} \{a_n\} = a. \quad (\text{IV.3})$$

- (ii) Eine Zahlenfolge $(a_n)_{n=1}^\infty \in \mathbb{K}^\mathbb{N}$ heißt **divergent** $:\Leftrightarrow$

$$(a_n)_{n=1}^\infty \text{ ist nicht konvergent.} \quad (\text{IV.4})$$

- (iii) Enthält eine Zahlenfolge $(a_n)_{n=1}^\infty \in \mathbb{K}^\mathbb{N}$ eine konvergente Teilfolge $(a_{n_k})_{k=1}^\infty$, so heißt deren Grenzwert $\lim_{k \rightarrow \infty} \{a_{n_k}\}$ **Häufungswert** von $(a_n)_{n=1}^\infty$.

Bemerkungen und Beispiele.

- Wir bemerken, dass der Limes einer konvergenten Folge eindeutig ist. Ist nämlich

$$a_n \rightarrow a \quad \text{und} \quad a_n \rightarrow a', \quad n \rightarrow \infty, \quad (\text{IV.5})$$

so gibt es für jedes $\varepsilon > 0$ zwei Indizes $n_0, n'_0 \in \mathbb{N}$, so dass

$$\forall n \geq n_0 : |a_n - a| \leq \varepsilon, \quad (\text{IV.6})$$

$$\forall n \geq n'_0 : |a_n - a'| \leq \varepsilon. \quad (\text{IV.7})$$

Für $n \geq \max\{n_0, n'_0\}$ ist demnach

$$|a - a'| \leq |a_n - a| + |a_n - a'| \leq 2\varepsilon. \quad (\text{IV.8})$$

Weil $\varepsilon > 0$ beliebig klein gewählt werden kann, folgt daraus

$$a = a'. \quad (\text{IV.9})$$

- Seien $\mathbb{K} = \mathbb{R}$ und $a_n := 1/n$. Dann konvergiert $a_n \rightarrow 0$, für $n \rightarrow \infty$.
- Seien $\mathbb{K} = \mathbb{R}$ und $a_n := (-1)^n + \frac{1}{n}$. Dann ist $(a_n)_{n=1}^\infty$ divergent und hat die Häufungswerte $\{-1, 1\}$.
- Seien $\mathbb{K} = \mathbb{C}$ und $a_n := \alpha + \frac{1}{n} + i\beta - \frac{i}{n}$. Dann konvergiert $a_n \rightarrow \alpha + i\beta$, für $n \rightarrow \infty$.
- Seien $\mathbb{K} = \mathbb{C}$ und $a_n := \cos(n) + i \sin(n)$. Dann ist $(a_n)_{n=1}^\infty$ divergent.
- Jede konvergente Zahlenfolge $(a_n)_{n=1}^\infty$ in $\mathbb{K}$ ist auch **beschränkt**. Genauer gesagt, ist dann die Menge $\{a_n\}_{n=1}^\infty \subseteq \mathbb{K}$ der Folgeglieder beschränkt, d.h.

$$\exists R < \infty \forall n \in \mathbb{N} : |a_n| \leq R. \quad (\text{IV.10})$$

Ist nämlich $a := \lim_{n \rightarrow \infty} a_n$, so können wir $\varepsilon := 1$ in (IV.1) wählen und erhalten ein $n_0 \in \mathbb{N}$, sodass

$$\forall n \geq n_0 : |a_n| \leq |a| + |a_n - a| \leq |a| + 1. \quad (\text{IV.11})$$

Also gilt (IV.10) mit

$$R := \max\{|a_1|, |a_2|, \dots, |a_{n_0-1}|, |a| + 1\} < \infty. \quad (\text{IV.12})$$

- Die Konvergenz von komplexen Folgen ist gleichbedeutend mit der Konvergenz ihres Real- und Imaginärteils,

$$(z_n)_{n=1}^\infty \in \mathbb{C}^\mathbb{N} \text{ ist konvergent} \Leftrightarrow \quad (\text{IV.13})$$

$$(\operatorname{Re}\{z_n\})_{n=1}^\infty \in \mathbb{R}^\mathbb{N} \text{ und } (\operatorname{Im}\{z_n\})_{n=1}^\infty \in \mathbb{R}^\mathbb{N} \text{ sind beide konvergent.}$$

Ist nämlich $|z_n - z| = \sqrt{\operatorname{Re}\{z_n - z\}^2 + \operatorname{Im}\{z_n - z\}^2} \leq \varepsilon$, so sind auch $|\operatorname{Re}\{z_n - z\}| \leq \varepsilon$ und $|\operatorname{Im}\{z_n - z\}| \leq \varepsilon$.

Sind umgekehrt $|\operatorname{Re}\{z_n - z\}| \leq \varepsilon$ und $|\operatorname{Im}\{z_n - z\}| \leq \varepsilon$, so ist $|z_n - z| \leq \sqrt{2}\varepsilon$.

Satz IV.2. Seien $\alpha \in \mathbb{K}$ und $(a_n)_{n=1}^\infty, (b_n)_{n=1}^\infty \in \mathbb{K}^\mathbb{N}$ zwei konvergente Zahlenfolgen und

$$a := \lim_{n \rightarrow \infty} \{a_n\}, \quad b := \lim_{n \rightarrow \infty} \{b_n\}. \quad (\text{IV.14})$$

Dann sind auch $(a_n + b_n)_{n=1}^\infty$ und $(\alpha a_n)_{n=1}^\infty$ konvergent, und es gilt

$$\lim_{n \rightarrow \infty} \{a_n + b_n\} = a + b = \lim_{n \rightarrow \infty} \{a_n\} + \lim_{n \rightarrow \infty} \{b_n\}, \quad (\text{IV.15})$$

$$\lim_{n \rightarrow \infty} \{\alpha a_n\} = \alpha a = \alpha \cdot \lim_{n \rightarrow \infty} \{a_n\}. \quad (\text{IV.16})$$

Beweis. Sei $\varepsilon > 0$. Dann ist auch $\varepsilon/(2 + |\alpha|) > 0$. Weil $(a_n)_{n=1}^\infty$ konvergent ist, gibt es $n'_0 \in \mathbb{N}$, sodass

$$\forall n \geq n'_0 : |a_n - a| \leq \frac{\varepsilon}{2 + |\alpha|}, \quad (\text{IV.17})$$

und weil $(b_n)_{n=1}^\infty$ konvergent ist, gibt es $n''_0 \in \mathbb{N}$, sodass

$$\forall n \geq n''_0 : |b_n - b| \leq \frac{\varepsilon}{2}. \quad (\text{IV.18})$$

Setzen wir nun

$$n_0 := \max\{n'_0, n''_0\}, \quad (\text{IV.19})$$

dann gilt für alle $n \geq n_0$, dass

$$\begin{aligned} |(a_n + b_n) - (a + b)| &= |(a_n - a) + (b_n - b)| \\ &\leq |a_n - a| + |b_n - b| \leq \frac{\varepsilon}{2 + |\alpha|} + \frac{\varepsilon}{2} \leq \varepsilon \end{aligned} \quad (\text{IV.20})$$

und

$$|\alpha a_n - \alpha a| = |\alpha| \cdot |a_n - a| \leq \frac{|\alpha| \cdot \varepsilon}{2 + |\alpha|} \leq \varepsilon. \quad (\text{IV.21})$$

□

Satz IV.3. Seien $(a_n)_{n=1}^\infty$ und $(b_n)_{n=1}^\infty \in \mathbb{K}^\mathbb{N}$ zwei konvergente Zahlenfolgen und

$$a := \lim_{n \rightarrow \infty} \{a_n\}, \quad b := \lim_{n \rightarrow \infty} \{b_n\}. \quad (\text{IV.22})$$

Dann ist $(a_n \cdot b_n)_{n=1}^\infty \in \mathbb{K}^\mathbb{N}$ konvergent, und es gilt

$$\lim_{n \rightarrow \infty} \{a_n \cdot b_n\} = a \cdot b = \lim_{n \rightarrow \infty} \{a_n\} \cdot \lim_{n \rightarrow \infty} \{b_n\}. \quad (\text{IV.23})$$

Sind außerdem $b_n \neq 0, \forall n \in \mathbb{N}$ und $b \neq 0$, so ist auch $(a_n/b_n)_{n=1}^\infty$ konvergent, und es gilt

$$\lim_{n \rightarrow \infty} \left\{ \frac{a_n}{b_n} \right\} = \frac{a}{b} = \frac{\lim_{n \rightarrow \infty} \{a_n\}}{\lim_{n \rightarrow \infty} \{b_n\}}. \quad (\text{IV.24})$$

IV.2. Reelle Folgen und Monotonie

Im vorigen Abschnitt haben wir gesehen, dass der Konvergenzbegriff in $\mathbb{C}$ auf den in $\mathbb{R}$ zurückgeführt werden kann. Aus der Ordnungsstruktur in $\mathbb{R}$ erhalten wir allerdings noch weitere Aussagen, die keine Entsprechung in $\mathbb{C}$ haben.

Definition IV.4. Eine reelle Zahlenfolge $(a_n)_{n=1}^\infty \in \mathbb{R}^\mathbb{N}$

$$\text{heißt } \left\{ \begin{array}{l} \text{monoton steigend} \\ \text{streng monoton steigend} \\ \text{monoton fallend} \\ \text{streng monoton fallend} \end{array} \right\} :\Leftrightarrow \forall n \in \mathbb{N} : \left\{ \begin{array}{l} a_n \leq a_{n+1}, \\ a_n < a_{n+1}, \\ a_n \geq a_{n+1}, \\ a_n > a_{n+1}. \end{array} \right. \quad (\text{IV.25})$$

Satz IV.5. Sei $(a_n)_{n=1}^\infty$ eine Folge in $\mathbb{R}$.

$$(i) \quad \begin{array}{l} \text{Ist } (a_n)_{n=1}^\infty \text{ nach oben beschränkt und monoton steigend,} \\ \text{so ist } (a_n)_{n=1}^\infty \text{ konvergent.} \end{array} \quad (\text{IV.26})$$

$$(ii) \quad \begin{array}{l} \text{Ist } (a_n)_{n=1}^\infty \text{ nach unten beschränkt und monoton fallend,} \\ \text{so ist } (a_n)_{n=1}^\infty \text{ konvergent.} \end{array} \quad (\text{IV.27})$$

Beweis. Offensichtlich sind (i) und (ii) äquivalent, wie man aus Ersetzung von a_n durch $-a_n$ ersieht. Wir zeigen nur (i). Die Menge $A := \{a_1, a_2, a_3, \dots\}$ ist nach oben beschränkt, und wir setzen

$$a := \sup A. \quad (\text{IV.28})$$

Sei nun $\varepsilon > 0$. Weil a das Supremum von A ist, gibt es nach Lemma II.7 ein $a_{n_0} \in A$, so dass

$$a_{n_0} \leq a \leq a_{n_0} + \varepsilon. \quad (\text{IV.29})$$

Da $(a_n)_{n \in \mathbb{N}}$ monoton wachsend ist, gilt $a_{n_0} \leq a_n$ für $n \geq n_0$, und daher

$$\forall n \geq n_0 : \quad 0 \leq a - a_n \leq a - a_{n_0} \leq \varepsilon. \quad (\text{IV.30})$$

Also ist $(a_n)_{n=1}^\infty$ konvergent und $\lim_{n \rightarrow \infty} \{a_n\} = a$. $\square$

Sei $(a_n)_{n=1}^\infty \in [-R, R]^\mathbb{N}$ eine nach oben und unten durch $\pm R$, $0 < R < \infty$ beschränkte Folge. Setzen wir

$$A_m := \{a_m, a_{m+1}, a_{m+2}, \dots\}, \quad (\text{IV.31})$$

so gilt

$$[-R, R] \supseteq A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots \quad (\text{IV.32})$$

Setzen wir weiter, für $m \in \mathbb{N}$,

$$b_m := \inf A_m \quad \text{und} \quad c_m := \sup A_m, \quad (\text{IV.33})$$

so gilt

$$\forall m \in \mathbb{N} : \quad -R \leq b_m \leq c_m \leq R. \quad (\text{IV.34})$$

Außerdem impliziert (IV.31)-(IV.32), dass

$$b_m = \inf A_m = \min \{a_m, \inf A_{m+1}\} \leq \inf A_{m+1} = b_{m+1}, \quad (\text{IV.35})$$

$$c_m = \sup A_m = \max \{a_m, \sup A_{m+1}\} \geq \sup A_{m+1} = c_{m+1}. \quad (\text{IV.36})$$

Also sind beide Folgen, $(b_m)_{m=1}^\infty$ und $(c_m)_{m=1}^\infty$, beschränkt und monoton und daher auch konvergent in $\mathbb{R}$.

Definition IV.6. Sei $(a_n)_{n=1}^\infty$ eine Folge in $\mathbb{R}$.

(i.a) Ist $(a_n)_{n=1}^\infty$ nicht nach oben beschränkt, dann setzen wir

$$\limsup_{n \rightarrow \infty} \{a_n\} := \overline{\lim}_{n \rightarrow \infty} \{a_n\} := \infty. \quad (\text{IV.37})$$

(Dabei ist “ ∞ ” nur als Symbol zu verstehen. $\infty, -\infty \notin \mathbb{R}$.)

- (i.b) Ist $(a_n)_{n=1}^\infty$ nach oben beschränkt, und ist $(c_m)_{m=1}^\infty$, mit $c_m := \sup\{a_m, a_{m+1}, \dots\}$, nicht nach unten beschränkt, so setzen wir

$$\limsup_{n \rightarrow \infty} \{a_n\} := -\infty. \quad (\text{IV.38})$$

- (i.c) Ist $(a_n)_{n=1}^\infty$ nach oben beschränkt, und ist $(c_m)_{m=1}^\infty$, mit $c_m := \sup\{a_m, a_{m+1}, \dots\}$, nach unten beschränkt, so setzen wir

$$\limsup_{n \rightarrow \infty} \{a_n\} := \lim_{m \rightarrow \infty} \{c_m\} = \lim_{m \rightarrow \infty} \left\{ \sup\{a_n \mid n \geq m\} \right\}. \quad (\text{IV.39})$$

(ii)

$$\liminf_{n \rightarrow \infty} \{a_n\} := \underline{\lim}_{n \rightarrow \infty} \{a_n\} := -\limsup_{n \rightarrow \infty} \{-a_n\}. \quad (\text{IV.40})$$

$\limsup_{n \rightarrow \infty} \{a_n\}$ heißt **Limes superior** von $(a_n)_{n=1}^\infty$,
 $\liminf_{n \rightarrow \infty} \{a_n\}$ heißt **Limes inferior** von $(a_n)_{n=1}^\infty$.

Bemerkungen und Beispiele.

- Sei $a_n := n$ für $n \in \mathbb{N}$. Dann ist $(a_n)_{n=1}^\infty$ nicht nach oben beschränkt, und es gilt (i.a), also $\limsup_{n \rightarrow \infty} \{a_n\} = \infty$.
- Sei $a_n := -n$ für $n \in \mathbb{N}$. Dann ist $(a_n)_{n=1}^\infty$ nach oben beschränkt. Weiterhin ist $c_m = \sup\{-m, -m-1, -m-2, \dots\} = -m$, deshalb ist $(c_m)_{m=1}^\infty$ nicht nach unten beschränkt, und es gilt (i.b), also $\limsup_{n \rightarrow \infty} \{a_n\} = -\infty$.
- Sei $a_n := (-1)^n \left(1 + \frac{1}{n}\right)$. Wegen $-2 \leq a_n \leq 2$ ist dann $(a_n)_{n=1}^\infty$ beschränkt, und es gilt (i.c). Weiterhin ist

$$c_{2k-1} = \sup \left\{ -\left(\frac{2k}{2k-1}\right), +\left(\frac{2k+1}{2k}\right), -\left(\frac{2k+2}{2k+1}\right), +\left(\frac{2k+3}{2k+2}\right), \dots \right\} = \frac{2k+1}{2k}, \quad (\text{IV.41})$$

$$c_{2k} = \sup \left\{ \frac{2k+1}{2k}, \frac{-2k-2}{2k+1}, \frac{2k+2}{2k+1}, \dots \right\} = \frac{2k+1}{2k},$$

also $c_m \rightarrow 1$, für $m \rightarrow \infty$, und $\limsup_{n \rightarrow \infty} \{a_n\} = 1$.

- Genauso sieht man, dass $\liminf_{n \rightarrow \infty} \{a_n\} = -1$ für $a_n := (-1)^n \left(1 + \frac{1}{n}\right)$. Wir beobachten, dass -1 und 1 Häufungswerte der Folge $(a_n)_{n=1}^\infty$ sind.
- Sei $(a_n)_{n=1}^\infty$ eine Abzählung von $\mathbb{Q} \cap (0, 1)$.
 Dann sind $\liminf_{n \rightarrow \infty} \{a_n\} = 0$ und $\limsup_{n \rightarrow \infty} \{a_n\} = 1$.

Satz IV.7. Sei $(a_n)_{n=1}^\infty \in \mathbb{R}^\mathbb{N}$ eine Folge in $\mathbb{R}$.

(i) Folgende Charakterisierungen sind gleichwertig:

$$\left\{ \limsup_{n \rightarrow \infty} \{a_n\} =: \bar{a} \in \mathbb{R} \text{ existiert} \right\} \quad (\text{IV.42})$$

$$\Leftrightarrow \left\{ \begin{array}{l} \forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : a_n \leq \bar{a} + \varepsilon \quad \text{und} \\ \forall \varepsilon > 0 \forall n_0 \in \mathbb{N} \exists n \geq n_0 : a_n \geq \bar{a} - \varepsilon \end{array} \right\} \quad (\text{IV.43})$$

$$\Leftrightarrow \left\{ (a_n)_{n=1}^{\infty} \text{ ist nach oben beschränkt und } \bar{a} \text{ ist ihr größter Häufungswert} \right\}. \quad (\text{IV.44})$$

(ii) Folgende Charakterisierungen sind gleichwertig:

$$\left\{ \liminf_{n \rightarrow \infty} \{a_n\} =: \underline{a} \in \mathbb{R} \text{ existiert} \right\} \quad (\text{IV.45})$$

$$\Leftrightarrow \left\{ \begin{array}{l} \forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : a_n \geq \underline{a} - \varepsilon \quad \text{und} \\ \forall \varepsilon > 0 \forall n_0 \in \mathbb{N} \exists n \geq n_0 : a_n \leq \underline{a} + \varepsilon \end{array} \right\} \quad (\text{IV.46})$$

$$\Leftrightarrow \left\{ (a_n)_{n=1}^{\infty} \text{ ist nach unten beschränkt und } \underline{a} \text{ ist ihr kleinster Häufungswert} \right\}. \quad (\text{IV.47})$$

Korollar IV.8. Sei $(a_n)_{n=1}^{\infty}$ eine Folge in $\mathbb{R}$.

$$\left[(a_n)_{n=1}^{\infty} \text{ ist konvergent in } \mathbb{R} \right] \Leftrightarrow \left[\limsup_{n \rightarrow \infty} \{a_n\} = \liminf_{n \rightarrow \infty} \{a_n\} \in \mathbb{R} \right], \quad (\text{IV.48})$$

und in diesem Fall gilt

$$\limsup_{n \rightarrow \infty} \{a_n\} = \liminf_{n \rightarrow \infty} \{a_n\} = \lim_{n \rightarrow \infty} \{a_n\}. \quad (\text{IV.49})$$

IV.3. Cauchy-Folgen

Definition IV.9. Eine Zahlenfolge $(a_n)_{n=1}^{\infty} \in \mathbb{K}^{\mathbb{N}}$ heißt **Cauchy-Folge**

$$:\Leftrightarrow \forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall m, n \geq n_0 : |a_m - a_n| \leq \varepsilon. \quad (\text{IV.50})$$

Cauchy-Folgen spielen in der Analysis eine große Rolle, weil man mit ihrer Hilfe den Konvergenzbegriff einführen kann, ohne expliziten Bezug auf den Grenzwert zu nehmen.

Lemma IV.10. Sei $(a_n)_{n=1}^{\infty} \in \mathbb{K}^{\mathbb{N}}$ eine konvergente Folge. Dann ist $(a_n)_{n=1}^{\infty}$ eine Cauchy-Folge.

Beweis. Seien $\varepsilon > 0$ und $a := \lim_{n \rightarrow \infty} \{a_n\}$. Dann gibt es ein $n_0 \in \mathbb{N}$, so dass

$$\forall n \geq n_0 : |a_n - a| \leq \varepsilon/2. \quad (\text{IV.51})$$

Also ist

$$\forall m, n \geq n_0 : |a_m - a_n| \leq |a_m - a| + |a_n - a| \leq \varepsilon. \quad (\text{IV.52})$$

□

Satz IV.11 (Cauchy-Kriterium). *Ist $(a_n)_{n=1}^\infty \in \mathbb{R}^\mathbb{N}$ eine reelle Cauchy-Folge, so ist $(a_n)_{n=1}^\infty$ konvergent.*

Beweis. Sei $(a_n)_{n=1}^\infty \in \mathbb{R}^\mathbb{N}$ eine reelle Cauchy-Folge. Wählen wir $\varepsilon := 1$, dann gibt es nach (IV.50) ein $n_0 \in \mathbb{N}$, so dass (mit $m := n_0$)

$$\forall n \geq n_0 : |a_n - a_{n_0}| \leq 1. \quad (\text{IV.53})$$

Also ist

$$\forall n \geq n_0 : |a_n| \leq |a_{n_0}| + |a_n - a_{n_0}| \leq |a_{n_0}| + 1, \quad (\text{IV.54})$$

und daher

$$\forall n \in \mathbb{N} : |a_n| \leq 1 + \max\{|a_1|, |a_2|, \dots, |a_{n_0}|\}, \quad (\text{IV.55})$$

d.h. $(a_n)_{n=1}^\infty$ ist nach oben und nach unten beschränkt. Nach Definition IV.6 sind deshalb

$$\bar{a} := \limsup_{n \rightarrow \infty} \{a_n\}, \quad \underline{a} := \liminf_{n \rightarrow \infty} \{a_n\} \in \mathbb{R}, \quad (\text{IV.56})$$

und es genügt zu zeigen, dass $\bar{a} = \underline{a}$. Sei dazu $\varepsilon > 0$ gewählt. Weil $(a_n)_{n=1}^\infty$ eine Cauchy-Folge ist, gibt es ein $n_0 \in \mathbb{N}$, so dass

$$\forall n \geq n_0 : |a_n - a_{n_0}| \leq \varepsilon, \quad (\text{IV.57})$$

was

$$\forall n \geq n_0 : a_{n_0} - \varepsilon \leq a_n \leq a_{n_0} + \varepsilon \quad (\text{IV.58})$$

impliziert. Daher gilt auch,

$$a_{n_0} - \varepsilon \leq \inf_{n \geq n_0} \{a_n\} =: b_{n_0} \leq c_{n_0} := \sup_{n \geq n_0} \{a_n\} \leq a_{n_0} + \varepsilon. \quad (\text{IV.59})$$

Aus (IV.59) und der Tatsache, dass b_m monoton steigt und c_m monoton sinkt, erhalten wir

$$a_{n_0} - \varepsilon \leq b_{n_0} \leq \lim_{n \rightarrow \infty} \{b_m\} = \underline{a} \quad (\text{IV.60})$$

und

$$\bar{a} = \lim_{n \rightarrow \infty} \{c_m\} \leq c_{n_0} \leq a_{n_0} + \varepsilon, \quad (\text{IV.61})$$

also

$$0 \leq \bar{a} - \underline{a} \leq (a_{n_0} + \varepsilon) - (a_{n_0} - \varepsilon) = 2\varepsilon. \quad (\text{IV.62})$$

Da $\varepsilon > 0$ beliebig klein gewählt werden kann, folgt daraus, dass

$$\underline{a} = \bar{a}, \quad (\text{IV.63})$$

was nach Korollar IV.8 die Konvergenz von $(a_n)_{n=1}^\infty$ zur Konsequenz hat. □

IV.4. Ergänzungen

IV.4.1. Vertauschung von Limiten mit Produkt und Quotient

Beweis. (Beweis von Satz IV.3) Als konvergente Folgen sind $(a_n)_{n=1}^\infty$ und $(b_n)_{n=1}^\infty$ beschränkt. Es gibt also ein $R < \infty$, so dass

$$\max \left\{ \sup_n |a_n|, \sup_n |b_n|, |a|, |b| \right\} \leq R. \quad (\text{IV.64})$$

Sei $\varepsilon > 0$. Dann gibt es $n'_0, n''_0 \in \mathbb{N}$, so dass

$$\forall n \geq n'_0 : |a_n - a| \leq \frac{\varepsilon}{2R}, \quad (\text{IV.65})$$

$$\forall n \geq n''_0 : |b_n - b| \leq \frac{\varepsilon}{2R}. \quad (\text{IV.66})$$

Setzen wir $n_0 := \max\{n'_0, n''_0\}$, dann gilt, für alle $n \geq n_0$

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - ab_n + ab_n - ab| \\ &\leq |a_n - a| \cdot |b_n| + |a| \cdot |b_n - b| \leq \frac{\varepsilon}{2R} \cdot R + R \cdot \frac{\varepsilon}{2R} = \varepsilon. \end{aligned} \quad (\text{IV.67})$$

Also ist $\lim_{n \rightarrow \infty} \{a_n b_n\} = ab$.

Für den Beweis von (IV.24) zeigen wir zunächst, dass die Folge $(\frac{1}{b_n})_{n=1}^\infty$ beschränkt ist. Dazu setzen wir $\varepsilon' := \frac{|b|}{2}$. Aus der Konvergenz von $(b_n)_{n=1}^\infty$ folgt dann die Existenz von $\tilde{n}_0 \in \mathbb{N}$, so dass

$$\forall n \geq \tilde{n}_0 : |b_n - b| \leq \varepsilon' = \frac{|b|}{2}. \quad (\text{IV.68})$$

Damit ist aber

$$\forall n \geq \tilde{n}_0 : |b_n| = |b + b_n - b| \geq |b| - |b_n - b| \geq \frac{|b|}{2}. \quad (\text{IV.69})$$

Also ist

$$\max \left\{ \frac{|1|}{|b|}, \sup_n \left| \frac{1}{b_n} \right| \right\} \leq R := \frac{2}{|b|} + \max \left\{ \frac{|1|}{|b_k|} \mid 1 \leq k \leq \tilde{n}_0 \right\} < \infty. \quad (\text{IV.70})$$

Für $\varepsilon > 0$ impliziert wiederum die Konvergenz von $(b_n)_{n=1}^\infty$ die Existenz von $n_0 \in \mathbb{N}$, so dass

$$\forall n \geq n_0 : |b - b_n| \leq \frac{\varepsilon}{R^2}. \quad (\text{IV.71})$$

Somit ist dann

$$\forall n \geq n_0 : \left| \frac{1}{b} - \frac{1}{b_n} \right| = \frac{|b - b_n|}{|b| \cdot |b_n|} \leq R^2 \cdot \frac{\varepsilon}{R^2} = \varepsilon. \quad (\text{IV.72})$$

Also konvergiert $(\frac{1}{b_n})_{n=1}^\infty$ in $\mathbb{K}$ gegen $\frac{1}{b}$. Die Behauptung (IV.24) folgt nun aus (IV.23). $\square$

IV.4.2. Limes Superior/Inferior als größter/kleinsten Häufungswert

Beweis. (Beweis von Satz IV.7) Aussagen (i) und (ii) sind offenbar wieder äquivalent, und wir zeigen nur (i). Dazu zeigen wir

$$(IV.42) \Rightarrow (IV.43) \Rightarrow (IV.44) \Rightarrow (IV.42). \quad (IV.73)$$

(IV.42) $\Rightarrow$ (IV.43): Sei $\limsup_{n \rightarrow \infty} \{a_n\} = \bar{a} \in \mathbb{R}$, und sei $\varepsilon > 0$. Nehmen wir an, es gäbe unendlich viele $a_n \geq \bar{a} + \varepsilon$, also eine Teilfolge $(a_{n_j})_{j=1}^{\infty}$ mit $a_{n_j} \geq \bar{a} + \varepsilon$. Wegen $n_j \rightarrow \infty$, für $j \rightarrow \infty$, gibt es zu jedem $m \in \mathbb{N}$ ein $n_j \geq m$ und deshalb ist

$$\forall m \in \mathbb{N} : \sup_{n \geq m} \{a_n\} \geq a_{n_j} \geq \bar{a} + \varepsilon, \quad (IV.74)$$

also

$$\limsup_{n \rightarrow \infty} \{a_n\} \geq \bar{a} + \varepsilon = \limsup_{n \rightarrow \infty} \{a_n\} + \varepsilon. \quad (IV.75)$$

Widerspruch. Daraus folgt die Existenz eines $n_0 \in \mathbb{N}$, so dass

$$\forall n \geq n_0 : a_n \leq \bar{a} + \varepsilon. \quad (IV.76)$$

Gäbe es nur endlich viele $n \in \mathbb{N}$ mit $a_n \geq \bar{a} - \varepsilon$, so müsste es $n'_0 \in \mathbb{N}$ geben, so dass

$$\forall n_0 \geq n'_0 : a_n \leq \bar{a} - \varepsilon. \quad (IV.77)$$

Dann wäre aber

$$\limsup_{n \rightarrow \infty} \{a_n\} \leq \sup_{n \geq n'_0} \{a_n\} \leq \bar{a} - \varepsilon = \limsup_{n \rightarrow \infty} \{a_n\} - \varepsilon. \quad (IV.78)$$

Widerspruch. Daraus folgt die Existenz einer Teilfolge $(a_{n_j})_{j=1}^{\infty}$ mit $a_{n_j} \geq \bar{a} - \varepsilon$, für alle $j \in \mathbb{N}$.

(IV.43) $\Rightarrow$ (IV.44): Seien $\varepsilon > 0$ und $n_0 \in \mathbb{N}$ und $(a_{n_j})_{j=1}^{\infty}$ so, dass

$$\left(\forall m \geq n_0 : a_m \leq \bar{a} + \varepsilon \right) \wedge \left(\forall j \in \mathbb{N} : a_{n_j} \geq \bar{a} - \varepsilon \right). \quad (IV.79)$$

Dann gilt

$$\forall j \in \mathbb{N}, n_j \geq n_0 : |a_{n_j} - \bar{a}| \leq \varepsilon, \quad (IV.80)$$

und $\bar{a}$ ist ein Häufungswert von $(a_n)_{n=1}^{\infty}$. Außerdem ist $(a_n)_{n=1}^{\infty}$ nach (IV.79) offensichtlich nach oben beschränkt.

Ist nun $b > \bar{a}$, so wählen wir $\varepsilon := \frac{b-\bar{a}}{3} > 0$. Nach (IV.43) gibt es ein $n_0 \in \mathbb{N}$, so dass

$$\forall m \geq n_0 : a_m \leq \bar{a} + \varepsilon = b - 2\varepsilon. \quad (IV.81)$$

Also kann b kein Häufungswert von $(a_n)_{n=1}^{\infty}$ sein. Somit ist $\bar{a}$ der größte Häufungswert von $(a_n)_{n=1}^{\infty}$.

(IV.44) $\Rightarrow$ (IV.42): Seien $(a_n)_{n=1}^{\infty}$ eine nach oben beschränkte Folge und $\bar{a}$ ihr größter Häufungswert. Weil $(a_n)_{n=1}^{\infty}$ nach oben beschränkt ist, ist $\limsup_{n \rightarrow \infty} \{a_n\} < \infty$. Ist

andererseits $\varepsilon > 0$, so gibt es eine Teilfolge $(a_{n_j})_{j=1}^\infty$ mit $a_{n_j} \geq \bar{a} - \varepsilon$, für alle $j \in \mathbb{N}$. Dann ist

$$\limsup_{n \rightarrow \infty} \{a_n\} = \lim_{n \rightarrow \infty} \left(\sup_{m \geq n} \{a_m\} \right) \geq \lim_{n \rightarrow \infty} \left(\sup_{j \in \mathbb{N}: n_j \geq n} \{a_{n_j}\} \right) \geq \bar{a} - \varepsilon, \quad (\text{IV.82})$$

und mit $\varepsilon \rightarrow 0$ folgt

$$\bar{a} \leq \limsup_{n \rightarrow \infty} \{a_n\} < \infty. \quad (\text{IV.83})$$

Sei nun $c_n := \sup_{m \geq n} \{a_m\}$. Dann ist c_n monoton fallend und $\limsup_{n \rightarrow \infty} \{a_n\} = \lim_{n \rightarrow \infty} \{c_n\}$. Weiterhin gibt es zu jedem $n \in \mathbb{N}$ einen Index $m(n) \geq n$, so dass

$$c_n - \frac{1}{n} \leq a_{m(n)} \leq c_n, \quad (\text{IV.84})$$

nach Definition des Supremums. O.B.d.A. können wir $m(n) < m(n+1)$ annehmen und erhalten eine konvergente Teilfolge $(a_{m(n)})_{n=1}^\infty$ mit

$$\lim_{n \rightarrow \infty} \{a_{m(n)}\} = \lim_{n \rightarrow \infty} \{c_n\} = \limsup_{n \rightarrow \infty} \{a_n\}. \quad (\text{IV.85})$$

Also ist $\limsup_{n \rightarrow \infty} \{a_n\}$ ein Häufungswert von $(a_n)_{n=1}^\infty$. Weil $\bar{a}$ der größte Häufungswert ist, folgt

$$\limsup_{n \rightarrow \infty} \{a_n\} = \bar{a}. \quad (\text{IV.86})$$

□

V. Reihen

V.1. Definitionen und Beispiele

Definition V.1. Sei $(a_n)_{n=1}^\infty \in \mathbb{K}^\mathbb{N}$ eine Zahlenfolge. Dann heißt die Folge $(s_m)_{m=1}^\infty \in \mathbb{K}^\mathbb{N}$, mit

$$s_m := \sum_{n=1}^m a_n, \quad (\text{V.1})$$

Reihe in $\mathbb{K}$ und s_m nennt man die **m. Partialsumme** (dieser Reihe). Ist $(s_m)_{m=1}^\infty$ konvergent, so schreiben wir

$$\sum_{n=1}^\infty a_n := \lim_{m \rightarrow \infty} \{s_m\} = \lim_{m \rightarrow \infty} \left\{ \sum_{n=1}^m a_n \right\}. \quad (\text{V.2})$$

Eine Umschreibung des Cauchy-Kriteriums, Satz IV.11, und des Monotoniekriteriums, Satz IV.5, von Folgen auf Reihen liefert

Lemma V.2 (Cauchy-Kriterium für Reihen). *Sei $(\sum_{n=1}^m a_n)_{m=1}^\infty \in \mathbb{K}^\mathbb{N}$ eine Reihe in $\mathbb{K}$. Dann gilt*

$$\left(\sum_{n=1}^m a_n \right)_{m=1}^\infty \text{ ist konvergent} \Leftrightarrow \forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall m \geq n \geq n_0 : \left| \sum_{k=n}^m a_k \right| \leq \varepsilon. \quad (\text{V.3})$$

Lemma V.3 (Monotoniekriterium für Reihen). *Sei $(a_n)_{n=1}^\infty \in (\mathbb{R}_0^+)^\mathbb{N}$ eine Folge nicht-negativer Zahlen. Dann gilt*

$$\left\{ \left(\sum_{n=1}^m a_n \right)_{m=1}^\infty \text{ ist konvergent in } \mathbb{R}_0^+ \right\} \Leftrightarrow \left\{ \left(\sum_{n=1}^m a_n \right)_{m=1}^\infty \text{ ist nach oben beschränkt} \right\}. \quad (\text{V.4})$$

Bemerkungen und Beispiele.

- Wir beachten, dass man nicht nur Reihen als spezielle Folgen gewinnen kann, sondern dass man auch umgekehrt Folgen als spezielle Reihen auffassen kann, nämlich mit $a_n := s_n - s_{n-1}$.

- Aus der Anwendung des Cauchy-Kriteriums für Reihen, Lemma V.2, gewinnt man mit der Wahl $m = n \geq n_0$ sofort auch das **Nullfolgenkriterium**: Ist $(\sum_{n=1}^m a_n)_{m=1}^\infty \in \mathbb{K}^\mathbb{N}$ eine konvergente Reihe, so ist $(a_n)_{n=1}^\infty$ auch eine Nullfolge, $\lim_{n \rightarrow \infty} a_n = 0$. Meistens wendet man die Umkehrung dieser Aussage an: Wenn $a_n \rightarrow 0$, $n \rightarrow \infty$, nicht gilt, muss die Reihe $(\sum_{n=1}^m a_n)_{m=1}^\infty$ divergieren.
- Sei $\lambda \in \mathbb{K}$ mit $0 \leq |\lambda| < 1$. Dann ist die **geometrische Reihe** $(\sum_{n=0}^m \lambda^n)_{m=0}^\infty$ konvergent in $\mathbb{K}$, und es gilt

$$\sum_{n=0}^{\infty} \lambda^n = \frac{1}{1 - \lambda}. \quad (\text{V.5})$$

Aus

$$(1 - \lambda) \left(\sum_{n=0}^m \lambda^n \right) = \sum_{n=0}^m \lambda^n - \sum_{n=0}^m \lambda^{n+1} = \sum_{n=0}^m \lambda^n - \sum_{\tilde{n}=1}^{m+1} \lambda^{\tilde{n}} = 1 - \lambda^{m+1} \quad (\text{V.6})$$

folgt nämlich

$$s_m := \sum_{n=0}^m \lambda^n = \frac{1 - \lambda^{m+1}}{1 - \lambda} \rightarrow \frac{1}{1 - \lambda}, \quad (\text{V.7})$$

für $m \rightarrow \infty$.

- Die **harmonische Reihe** $(\sum_{n=1}^m \frac{1}{n})_{m=1}^\infty \in (\mathbb{R}_0^+)^\mathbb{N}$ ist divergent, dies folgt aus Satz V.13 von Cauchy. In der Tat können wir später leicht mit Hilfe der Integralrechnung zeigen, dass $\sum_{n=1}^m \frac{1}{n} \approx \ln(m)$ unbeschränkt wächst, falls $m \rightarrow \infty$.
- Satz V.13 liefert aber auch die Konvergenz der Reihe $(\sum_{n=1}^m \frac{1}{n^p})_{m=1}^\infty \in (\mathbb{R}_0^+)^\mathbb{N}$, für jedes $p > 1$.

V.2. Absolute Konvergenz

Definition V.4. Eine Reihe $(\sum_{n=1}^m a_n)_{m=1}^\infty$ in $\mathbb{K}$ heißt **absolut konvergent**

$$:\Leftrightarrow \quad \text{Die Reihe } \left(\sum_{n=1}^m |a_n| \right)_{m=1}^\infty \text{ ist konvergent in } \mathbb{R}_0^+. \quad (\text{V.8})$$

Lemma V.5. Jede absolut konvergente Reihe ist konvergent.

Beweis. Ist die Reihe $(\sum_{n=1}^m a_n)_{m=1}^\infty \in \mathbb{K}^\mathbb{N}$ absolut konvergent, so ist $(\sum_{n=1}^m |a_n|)_{m=1}^\infty \in (\mathbb{R}_0^+)^\mathbb{N}$ konvergent. Zu jedem $\varepsilon > 0$ gibt es also ein $n_0 \in \mathbb{N}$, so dass

$$\forall m \geq n \geq n_0 : \sum_{k=n}^m |a_k| \leq \varepsilon. \quad (\text{V.9})$$

Mit der Dreiecksungleichung gilt dann auch

$$\forall m \geq n \geq n_0 : \left| \sum_{k=n}^m a_k \right| \leq \sum_{k=n}^m |a_k| \leq \varepsilon, \quad (\text{V.10})$$

also ist nach Lemma V.2 $(\sum_{n=1}^m a_n)_{m=1}^\infty$ konvergent in $\mathbb{K}$. $\square$

Bemerkungen und Beispiele.

- Nicht jede konvergente Reihe ist auch absolut konvergent. Um dies zu sehen betrachten wir die Reihe $(\sum_{n=1}^m a_n)_{m=1}^\infty$ mit $a_n := (-1)^{n+1} \frac{1}{n} \in \mathbb{R}$.

Wir beobachten, dass für $m = 2k$ gilt

$$\begin{aligned} s_{2k} &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2k-1} - \frac{1}{2k} \\ &\leq 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2k-1} - \frac{1}{2k} + \underbrace{\frac{1}{2k+1} - \frac{1}{2k+2}}_{\geq 0} = s_{2k+2}. \end{aligned} \quad (\text{V.11})$$

Also ist die Folge $(s_{2k})_{k=1}^\infty$ der Partialsummen mit gerader Zahl von Summanden monoton wachsend. Außerdem ist wegen $2\ell(2\ell-1) \geq \ell^2$

$$s_{2k} = \sum_{\ell=1}^k \left(\frac{1}{2\ell-1} - \frac{1}{2\ell} \right) = \sum_{\ell=1}^k \frac{1}{2\ell(2\ell-1)} \leq \sum_{\ell=1}^k \frac{1}{\ell^2} \leq \text{const} < \infty \quad (\text{V.12})$$

beschränkt, da die Reihe $(\sum_{\ell=1}^k \frac{1}{\ell^2})_{k=1}^\infty$ konvergent ist. Somit ist die Folge $(s_{2k})_{k=1}^\infty$ konvergent.

Sei nun $\varepsilon > 0$. Da $(s_{2k})_{k=1}^\infty$ konvergent ist, gibt es ein $k_0 \in \mathbb{N}_0$ so, dass

$$\forall \ell \geq k \geq k_0 : \quad |s_{2\ell} - s_{2k}| \leq \frac{\varepsilon}{3}. \quad (\text{V.13})$$

Wir wählen $n_0 \in \mathbb{N}$ so groß, dass $n_0 \geq 2k_0 + 2$ und $\frac{1}{n_0} \leq \frac{\varepsilon}{3}$ ist. Sind nun $m, n \in \mathbb{N}$ mit $m > n \geq n_0$, so gibt es eindeutige Zahlen $\ell, k \in \mathbb{N}$, $\ell \geq k \geq k_0$ und $\sigma, \tau \in \{0, 1\}$, sodass $m = 2\ell + \sigma$ und $n = 2k + \tau$ gelten. Damit erhalten wir aus (V.13)

$$\begin{aligned} |s_m - s_n| &= |s_{2\ell+\sigma} - s_{2k+\tau}| \leq |s_{2\ell+\sigma} - s_{2\ell}| + |s_{2\ell} - s_{2k}| + |s_{2k} - s_{2k+\tau}| \\ &\leq \frac{1}{2\ell} + \frac{\varepsilon}{3} + \frac{1}{2k} \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \leq \varepsilon. \end{aligned} \quad (\text{V.14})$$

Also ist $(s_n)_{n=1}^\infty$ eine Cauchy-Folge und daher auch konvergent.

- Andererseits ist $(\sum_{n=1}^m (-1)^{n+1} \frac{1}{n})_{m=1}^\infty$ nicht absolut konvergent, denn $(\sum_{n=1}^m |(-1)^{n+1} \frac{1}{n}|)_{m=1}^\infty = (\sum_{n=1}^m \frac{1}{n})_{m=1}^\infty$ ist divergent.

Eine fundamentale Eigenschaft absolut konvergenter Reihen ist die Tatsache, dass auch “Umordnungen” absolut konvergieren, und zwar alle gegen denselben Grenzwert.

Definition V.6. Seien $(a_n)_{n=1}^\infty, (b_n)_{n=1}^\infty \in \mathbb{K}^\mathbb{N}$ Zahlenfolgen. Die Reihe $(\sum_{n=1}^m b_n)_{m=1}^\infty$ heißt **Umordnung der Reihe** $(\sum_{n=1}^m a_n)_{m=1}^\infty : \Leftrightarrow$

$$\exists \text{ Bijektion } \sigma : \mathbb{N} \rightarrow \mathbb{N} \quad \forall n \in \mathbb{N} : \quad b_n = a_{\sigma(n)}. \quad (\text{V.15})$$

Satz V.7 (Großer Umordnungssatz). Seien $(\sum_{n=1}^m a_n)_{m=1}^\infty$ eine Reihe in $\mathbb{K}$ und $(\sum_{n=1}^m b_n)_{m=1}^\infty$ eine Umordnung von $(\sum_{n=1}^m a_n)_{m=1}^\infty$. Dann gilt

$$(i) \quad \left\{ \left(\sum_{n=1}^m a_n \right)_{m=1}^\infty \text{ ist abs. konv.} \right\} \Leftrightarrow \left\{ \left(\sum_{n=1}^m b_n \right)_{m=1}^\infty \text{ ist abs. konv.} \right\}, \quad (V.16)$$

$$(ii) \quad \left\{ \left(\sum_{n=1}^m a_n \right)_{m=1}^\infty \text{ ist absolut konvergent} \right\} \Rightarrow \left\{ \sum_{n=1}^\infty a_n = \sum_{n=1}^\infty b_n \right\}. \quad (V.17)$$

Umgekehrt zeigt der nun folgende Satz, dass jede konvergente, aber nicht absolut konvergente Reihe in $\mathbb{R}$ so umgeordnet werden kann, dass sie gegen jeden beliebigen vorgegebenen Grenzwert konvergiert.

Satz V.8. Sei $(\sum_{n=1}^m a_n)_{m=1}^\infty$ eine Reihe in $\mathbb{R}$, die konvergent, aber nicht absolut konvergent ist, und seien $\alpha, \beta \in \mathbb{R}, \alpha \leq \beta$, beliebig. Dann gibt es eine Umordnung $(\sum_{n=1}^m b_n)_{m=1}^\infty$ von $(\sum_{n=1}^m a_n)_{m=1}^\infty$, so dass

$$\liminf_{m \rightarrow \infty} \{t_m\} = \alpha, \quad \limsup_{m \rightarrow \infty} \{t_m\} = \beta, \quad (V.18)$$

wobei $(t_m)_{m=1}^\infty$ die Folge in $\mathbb{R}$ mit $t_m := \sum_{n=1}^m b_n$ ist.

Eine Anwendung von Satz V.7 ist das *Cauchy-Produkt*.

Dazu betrachten wir zwei Zahlenfolgen $(a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty \in \mathbb{K}^\mathbb{N}$ und die zugehörigen Reihen $(\sum_{n=0}^m a_n)_{m=1}^\infty, (\sum_{n=0}^m b_n)_{m=1}^\infty$ in $\mathbb{K}$, die wir als absolut konvergent annehmen. Wir bilden nun

$$c_0 := a_0 b_0, \quad (V.19)$$

$$c_1 := a_0 b_1 + a_1 b_0, \quad (V.20)$$

$$c_2 := a_0 b_2 + a_1 b_1 + a_2 b_0, \quad (V.21)$$

$$\vdots$$

$$c_n := a_0 b_n + a_1 b_{n-1} + \dots + a_n b_0 = \sum_{k=0}^n a_k b_{n-k}, \quad (V.22)$$

für $n \in \mathbb{N}_0$, und betrachten die Reihe $(\sum_{n=0}^m c_n)_{m=1}^\infty$, das Cauchy-Produkt der Reihen $(\sum_{n=0}^m a_n)_{m=1}^\infty$ und $(\sum_{n=0}^m b_n)_{m=1}^\infty$.

Satz V.9 (Cauchy-Produkt). Seien $(\sum_{n=0}^m a_n)_{m=1}^\infty, (\sum_{n=0}^m b_n)_{m=1}^\infty$ zwei absolut konvergente Reihen in $\mathbb{K}$ und $c_n := \sum_{k=0}^n a_k b_{n-k}$, für $n \in \mathbb{N}_0$. Dann ist das **Cauchy-Produkt** $(\sum_{n=0}^m c_n)_{m=1}^\infty$ absolut konvergent in $\mathbb{K}$, und es gilt

$$\sum_{n=0}^\infty c_n = \sum_{n=0}^\infty \sum_{k=0}^n a_k b_{n-k} = \left(\sum_{n=0}^\infty a_n \right) \cdot \left(\sum_{n=0}^\infty b_n \right). \quad (V.23)$$

V.3. Konvergenzkriterien

Für die Überprüfung, ob eine Reihe konvergent oder divergent ist, diskutieren wir drei Konvergenzkriterien: das Majorantenkriterium, das Wurzelkriterium und das Quotientenkriterium.

Satz V.10 (Majorantenkriterium).

- (i) Ist $(a_n)_{n=1}^\infty \in \mathbb{K}^\mathbb{N}$ eine Zahlenfolge und ist $(b_n)_{n=1}^\infty \in (\mathbb{R}_0^+)^\mathbb{N}$ eine Folge nichtnegativer Zahlen, sodass $(\sum_{n=1}^m b_n)_{m=1}^\infty$ (absolut) konvergent ist und $|a_n| \leq b_n$ für alle $n \in \mathbb{N}$ gilt, so ist auch $(\sum_{n=1}^m a_n)_{m=1}^\infty$ absolut konvergent in $\mathbb{K}$.
- (ii) Sind $(a_n)_{n=1}^\infty \in (\mathbb{R}_0^+)^\mathbb{N}$ und $(b_n)_{n=1}^\infty \in (\mathbb{R}_0^+)^\mathbb{N}$ zwei Folgen nichtnegativer Zahlen, sodass $(\sum_{n=1}^m b_n)_{m=1}^\infty$ divergent ist und $a_n \geq b_n$ für alle $n \in \mathbb{N}$ gilt, so ist auch $(\sum_{n=1}^m a_n)_{m=1}^\infty$ divergent.

Beweis.

(ii) Die Reihen über $(a_n)_{n=1}^\infty \in (\mathbb{R}_0^+)^\mathbb{N}$ und $(b_n)_{n=1}^\infty \in (\mathbb{R}_0^+)^\mathbb{N}$ sind jeweils genau dann konvergent, wenn $(s_m)_{m=1}^\infty$ bzw. $(t_m)_{m=1}^\infty$ nach oben beschränkt ist, wobei $s_m := \sum_{n=0}^m a_n$ und $t_m := \sum_{n=0}^m b_n$. Wegen der Divergenz von $(\sum_{n=1}^m b_n)_{m=1}^\infty$ ist $(t_m)_{m=1}^\infty$ unbeschränkt somit auch $(s_m)_{m=1}^\infty$, da $s_m \geq t_m \geq 0$. Also ist auch $(\sum_{n=1}^m a_n)_{m=1}^\infty$ divergent.

Zu (i): Für $\varepsilon > 0$ gibt es, wegen der Konvergenz von $(\sum_{n=1}^m b_n)_{m=1}^\infty$, ein $n_0 \in \mathbb{N}$, so dass

$$\forall m \geq n \geq n_0 : \sum_{k=n}^m |a_k| \leq \sum_{k=n}^m b_k \leq \varepsilon. \quad (\text{V.24})$$

Also ist $(\sum_{n=1}^m a_n)_{m=1}^\infty$ absolut konvergent und somit auch konvergent, nach Lemma V.5. □

Satz V.11 (Wurzelkriterium). Sei $(a_n)_{n=1}^\infty \in \mathbb{K}^\mathbb{N}$ eine Zahlenfolge.

(i) Gilt

$$\limsup_{n \rightarrow \infty} \{ \sqrt[n]{|a_n|} \} < 1, \quad (\text{V.25})$$

so ist die Reihe $(\sum_{n=1}^m a_n)_{m=1}^\infty$ absolut konvergent in $\mathbb{K}$.

(ii) Gilt

$$\limsup_{n \rightarrow \infty} \{ \sqrt[n]{|a_n|} \} > 1, \quad (\text{V.26})$$

so ist die Reihe $(\sum_{n=1}^m a_n)_{m=1}^\infty$ divergent in $\mathbb{K}$.

Beweis.

Zu (i) Sei $\alpha := \limsup_{n \rightarrow \infty} \{ \sqrt[n]{|a_n|} \} < 1$. Für $\varepsilon := \frac{1-\alpha}{2} > 0$ gibt es dann nach Satz IV.7 (ii) ein $n_0 \in \mathbb{N}$, so dass

$$\forall n \geq n_0 : \sqrt[n]{|a_n|} \leq \alpha + \varepsilon = \frac{1+\alpha}{2} < 1. \quad (\text{V.27})$$

Also ist, für alle $m \geq n \geq n_0$,

$$\begin{aligned} \sum_{k=n}^m |a_k| &= \sum_{k=n}^m (\sqrt[k]{|a_k|})^k \leq \sum_{k=n}^m \left(\frac{1+\alpha}{2}\right)^k \leq \left(\frac{1+\alpha}{2}\right)^n \cdot \sum_{j=0}^{\infty} \left(\frac{1+\alpha}{2}\right)^j \\ &= \left(\frac{1+\alpha}{2}\right)^n \cdot \frac{1}{1 - \frac{1+\alpha}{2}} = \left(\frac{1+\alpha}{2}\right)^n \cdot \left(\frac{2}{1-\alpha}\right). \end{aligned} \quad (\text{V.28})$$

Wegen $\frac{1+\alpha}{2} < 1$ folgt $\lim_{n \rightarrow \infty} \left\{ \left(\frac{1+\alpha}{2}\right)^n \right\} = 0$, und somit ist nach dem Cauchy-Kriterium für Reihen $\left(\sum_{n=1}^m a_n\right)_{m=1}^{\infty}$ absolut konvergent.

(Gemäß (V.27) ist $|a_n| \leq \left(\frac{1+\alpha}{2}\right)^n$ eine summierbare Majorante und die Behauptung hätte sich nach (V.27) auch direkt aus Satz V.10 (i) ergeben.)

Zu (ii) Seien $\alpha := \lim_{n \rightarrow \infty} \sup \left\{ \sqrt[n]{|a_n|} \right\} > 1$ und $\varepsilon := \frac{\alpha-1}{2} > 0$. Dann gibt es eine Teilfolge $(a_{n_j})_{j=1}^{\infty}$ mit

$$\forall j \in \mathbb{N} : \quad \sqrt[n_j]{|a_{n_j}|} > \alpha - \varepsilon = \frac{\alpha+1}{2} > 1. \quad (\text{V.29})$$

Dann sind

$$\forall j \in \mathbb{N} : \quad |a_{n_j}| \geq \left(\frac{\alpha+1}{2}\right)^{n_j} \geq 1. \quad (\text{V.30})$$

Um zu zeigen, dass $\left(\sum_{n=1}^m a_n\right)_{m=1}^{\infty}$ divergiert, beweisen wir, dass das Cauchy-Kriterium nicht erfüllt ist. Seien nämlich $\delta := \frac{1}{2}$ und $n_0 \in \mathbb{N}$ irgend eine natürliche Zahl. Dann gibt es ein $j \in \mathbb{N}$, so dass $n_j \geq n_0$, und

$$\left| \sum_{k=n_j}^{n_j} a_{n_j} \right| = |a_{n_j}| \geq 1 > \delta. \quad (\text{V.31})$$

Für $m := n := n_j \geq n_0$ ergibt sich daraus ein Widerspruch zum Cauchy-Kriterium. $\square$

Bemerkungen und Beispiele.

- Für $\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = 1$ gibt es keine allgemeine Aussage. Beispielsweise ist $\sqrt[n]{\frac{1}{n^p}} \rightarrow 1$, mit $n \rightarrow \infty$, für alle $p > 0$, es ist aber $\left(\sum_{n=1}^m \frac{1}{n^2}\right)_{m=1}^{\infty}$ (absolut) konvergent in $\mathbb{R}$ und $\left(\sum_{n=1}^m \frac{1}{n}\right)_{m=1}^{\infty}$ divergent in $\mathbb{R}$.

Satz V.12 (Quotientenkriterium). *Sei $(a_n)_{n=1}^{\infty} \in \mathbb{K}^{\mathbb{N}} \setminus \{0\}$ eine Zahlenfolge.*

(i) Für

$$\limsup_{n \rightarrow \infty} \left\{ \frac{|a_{n+1}|}{|a_n|} \right\} < 1 \quad (\text{V.32})$$

ist die Reihe $\left(\sum_{n=1}^m a_n\right)_{m=1}^{\infty}$ absolut konvergent.

(ii) Gibt es ein $n_0 \in \mathbb{N}$, so dass

$$\forall n \geq n_0 : \frac{|a_{n+1}|}{|a_n|} \geq 1, \quad (\text{V.33})$$

so ist die Reihe $(\sum_{n=1}^m a_n)_{m=1}^\infty$ divergent.

Bemerkungen und Beispiele.

- Ein Vergleich mit Bedingung (V.26) aus dem Wurzelkriterium für die Divergenz einer Reihe würde, übertragen auf das Quotientenkriterium, die Bedingung

$$\limsup_{n \rightarrow \infty} \left\{ \frac{|a_{n+1}|}{|a_n|} \right\} > 1 \quad (\text{V.34})$$

für die Divergenz einer Reihe suggerieren. Dies ist jedoch falsch,¹ wie das folgende Gegenbeispiel illustriert.

- Für

$$a_n := \begin{cases} 3^{-n}, & \text{falls } n \text{ ungerade ist,} \\ 2^{-n+1}, & \text{falls } n \text{ gerade ist,} \end{cases} \quad (\text{V.35})$$

ist $(a_n)_{n=1}^\infty$ sicher konvergent, da $|a_n|$ die summierbare Majorante 2^{-n+1} besitzt, aber

$$\limsup_{n \rightarrow \infty} \left\{ \frac{|a_{n+1}|}{|a_n|} \right\} \geq \limsup_{k \rightarrow \infty} \left\{ \frac{|a_{2k+1}|}{|a_{2k}|} \right\} = \limsup_{k \rightarrow \infty} \left\{ \frac{2^{-2k}}{3^{-2k}} \right\} = \limsup_{k \rightarrow \infty} \left\{ \left(\frac{3}{2}\right)^{2k} \right\} = \infty. \quad (\text{V.36})$$

V.4. Exponentialfunktion und trigonometrische Funktionen

Eine der wichtigsten Anwendungen der obigen Konvergenzkriterien ist die Darstellung bzw. Definition elementarer Funktionen durch Potenzreihen. Wir beginnen mit der Exponentialfunktion.

- Seien $z \in \mathbb{C} \setminus \{0\}$ und $a_n := z^n/n!$. Wir beobachten, dass

$$\frac{|a_{n+1}|}{|a_n|} = \frac{|z|^{n+1} \cdot n!}{(n+1)! \cdot |z|^n} = \frac{|z|}{(n+1)} \rightarrow 0, \quad n \rightarrow \infty. \quad (\text{V.37})$$

Also ist

$$\limsup_{n \rightarrow \infty} \left\{ \frac{|a_{n+1}|}{|a_n|} \right\} = \limsup_{n \rightarrow \infty} \left\{ \frac{|z|}{n+1} \right\} = \lim_{n \rightarrow \infty} \left\{ \frac{|z|}{n+1} \right\} = 0 < 1. \quad (\text{V.38})$$

¹Ich danke Ch. Brauer und B. Komander für diesen Hinweis.

Mit Hilfe des Quotientenkriteriums, Satz V.12, sehen wir also, dass die Reihe $(\sum_{n=0}^m \frac{z^n}{n!})_{m=1}^\infty$ absolut konvergiert. Den Limes dieser Reihe nennen wir **Exponentialfunktion**,

$$\exp : \mathbb{C} \rightarrow \mathbb{C}, \quad \exp[z] := \sum_{n=0}^{\infty} \frac{z^n}{n!}. \quad (\text{V.39})$$

Dabei vereinbaren wir, dass $\exp[0] := 1$, d.h. hier ist Konvention $0^0 = 1$ sinnvoll und die Reihe für $\exp[0]$ besitzt nur einen nicht-verschwindenden Summanden.

- Analog seien $z \in \mathbb{C} \setminus \{0\}$ und $c_k := (-1)^k z^{2k} / (2k)!$ sowie $s_k := (-1)^k z^{2k+1} / (2k+1)!$. Dann sind

$$\limsup_{k \rightarrow \infty} \left\{ \frac{|c_{k+1}|}{|c_k|} \right\} = \limsup_{k \rightarrow \infty} \left\{ \frac{|z|^2}{(2k+2)(2k+1)} \right\} = 0 < 1, \quad (\text{V.40})$$

$$\limsup_{k \rightarrow \infty} \left\{ \frac{|s_{k+1}|}{|s_k|} \right\} = \limsup_{k \rightarrow \infty} \left\{ \frac{|z|^2}{(2k+3)(2k+2)} \right\} = 0 < 1, \quad (\text{V.41})$$

und die zugehörigen Reihen, die **Kosinusreihe** $\cos[z]$ und die **Sinusreihe** $\sin[z]$,

$$\cos[z] := \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k}}{(2k)!}, \quad \sin[z] := \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k+1}}{(2k+1)!}, \quad (\text{V.42})$$

sind absolut konvergent für $z \in \mathbb{C} \setminus \{0\}$. Für $z = 0$ ergibt sich $\cos[0] = 1$ und $\sin[0] = 0$.

- Weiterhin beobachten wir, dass wegen $i^{2k} = (-1)^k$

$$\begin{aligned} e^{iz} &:= \exp[iz] = \sum_{n=0}^{\infty} \frac{i^n z^n}{n!} = \sum_{k=0}^{\infty} \frac{i^{2k} z^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{i^{2k+1} z^{2k+1}}{(2k+1)!} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k+1}}{(2k+1)!} = \cos[z] + i \sin[z] \end{aligned} \quad (\text{V.43})$$

für alle $z \in \mathbb{C}$ gilt.

- Insbesondere sind dann

$$\cos[-z] = \sum_{k=0}^{\infty} \frac{(-1)^k (-z)^{2k}}{(2k)!} = \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k}}{(2k)!} = \cos[z], \quad (\text{V.44})$$

$$\sin[-z] = \sum_{k=0}^{\infty} \frac{(-1)^k (-z)^{2k+1}}{(2k+1)!} = - \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k+1}}{(2k+1)!} = -\sin[z], \quad (\text{V.45})$$

woraus

$$\frac{1}{2}(e^{iz} + e^{-iz}) = \frac{1}{2}(\cos[z] + i \sin[z] + \cos[-z] + i \sin[-z]) = \cos[z], \quad (\text{V.46})$$

$$\frac{1}{2i}(e^{iz} - e^{-iz}) = \frac{1}{2i}(\cos[z] + i \sin[z] - \cos[-z] - i \sin[-z]) = \sin[z] \quad (\text{V.47})$$

folgen.

- So erhalten wir für alle $z \in \mathbb{C}$ die hyperbolischen Funktionen aus den trigonometrischen durch

$$\cosh[z] := \cos[iz] = \frac{1}{2}(e^z + e^{-z}) \quad \textbf{Kosinus Hyperbolicus}, \quad (\text{V.48})$$

$$\sinh[z] := \frac{1}{i} \sin[iz] = \frac{1}{2}(e^z - e^{-z}) \quad \textbf{Sinus Hyperbolicus}. \quad (\text{V.49})$$

- Definieren wir außerdem die **Binominalkoeffizienten**

$$\forall m, n \in \mathbb{N}_0, m \geq n : \quad \binom{m}{n} := \frac{m!}{n!(m-n)!}, \quad (\text{V.50})$$

so gilt

$$\forall m \in \mathbb{N}_0, z, w \in \mathbb{C} : \quad (z+w)^m = \sum_{n=0}^m \binom{m}{n} z^n w^{m-n}. \quad (\text{V.51})$$

Setzen wir $a_n := \frac{z^n}{n!}$ und $b_n := \frac{w^n}{n!}$, dann erhalten wir also für das Cauchy-Produkt

$$\begin{aligned} c_m &:= \sum_{n=0}^m a_n b_{m-n} = \sum_{n=0}^m \frac{z^n}{n!} \frac{w^{m-n}}{(m-n)!} \\ &= \frac{1}{m!} \sum_{n=0}^m \underbrace{\frac{m!}{n!(m-n)!}}_{=\binom{m}{n}} z^n w^{m-n} = \frac{(z+w)^m}{m!}. \end{aligned} \quad (\text{V.52})$$

Also gilt das Additionstheorem für die Exponentialfunktion

$$\forall z, w \in \mathbb{C} : \quad \exp[z] \cdot \exp[w] = \left(\sum_{n=0}^{\infty} a_n \right) \left(\sum_{m=0}^{\infty} b_m \right) = \left(\sum_{m=0}^{\infty} c_m \right) = \exp[z+w]. \quad (\text{V.53})$$

- Schließlich erhalten wir die Additionstheoreme für die trigonometrischen Funktionen und komplexe Argumente $z, w \in \mathbb{C}$, etwa

$$\begin{aligned} &\cos[z] \cos[w] - \sin[z] \sin[w] \\ &= \frac{1}{4}(e^{iz} + e^{-iz})(e^{iw} + e^{-iw}) - \frac{1}{4i^2}(e^{iz} - e^{-iz})(e^{iw} - e^{-iw}) \\ &= \frac{1}{4}(e^{iz+iw} + e^{-iz+iw} + e^{iz-iw} + e^{-iz-iw} + e^{iz+iw} - e^{-iz+iw} - e^{iz-iw} + e^{-iz-iw}) \\ &= \frac{1}{2}(e^{i(z+w)} + e^{-i(z+w)}) = \cos[z+w], \end{aligned} \quad (\text{V.54})$$

und genauso $\sin[z+w] = \sin[z] \cos[w] + \cos[z] \sin[w]$.

V.5. Ergänzungen

V.5.1. Der Satz von Cauchy zur Konvergenz von Reihen nichtnegativer, monotoner Summanden

Satz V.13 (Cauchy). Sei $(a_n)_{n=1}^\infty \in (\mathbb{R}_0^+)^{\mathbb{N}}$ eine monoton fallende Folge nichtnegativer Zahlen. Dann gilt

$$\left\{ \left(\sum_{n=1}^m a_n \right)_{m=1}^\infty \text{ ist konvergent in } \mathbb{R}_0^+ \right\} \Leftrightarrow \left\{ \left(\sum_{k=1}^\ell 2^k \cdot a_{2^k} \right)_{\ell=1}^\infty \text{ ist konvergent in } \mathbb{R}_0^+ \right\}. \quad (\text{V.55})$$

Beweis. Wir setzen

$$s_m := \sum_{n=1}^m a_n \quad \text{und} \quad t_\ell := \sum_{k=1}^\ell 2^k a_{2^k}. \quad (\text{V.56})$$

Konvergenz von $(\sum_{n=1}^m a_n)_{m=1}^\infty$ bzw. $(\sum_{k=1}^\ell 2^k a_{2^k})_{\ell=1}^\infty$ ist dann nach Lemma V.3 äquivalent zur Beschränktheit von $(s_m)_{m=1}^\infty$ und $(t_\ell)_{\ell=1}^\infty$ nach oben. Es genügt also zu zeigen, dass

$$\left[(s_m)_{m=1}^\infty \text{ ist nach oben beschränkt.} \right] \Leftrightarrow \left[(t_\ell)_{\ell=1}^\infty \text{ ist nach oben beschränkt.} \right] \quad (\text{V.57})$$

Für $m < 2^\ell$ ist

$$s_m \leq a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + \dots + \underbrace{(a_{2^\ell} + a_{2^\ell+1} + \dots + a_{2^{\ell+1}-1})}_{2^\ell \text{ Summanden}}. \quad (\text{V.58})$$

Wegen $a_n \geq a_{n+1}$ sind

$$\begin{aligned} a_1 &\leq a_1, & a_2 + a_3 &\leq 2a_2, & a_4 + a_5 + a_6 + a_7 &\leq 4a_4, \\ & & \dots, a_{2^\ell} + a_{2^\ell+1} + \dots + a_{2^{\ell+1}-1} &\leq 2^\ell \cdot a_{2^\ell}. \end{aligned} \quad (\text{V.59})$$

Also gilt

$$s_m \leq a_1 + 2a_2 + 4a_4 + \dots + 2^\ell a_{2^\ell} = t_\ell \quad (\text{V.60})$$

und somit

$$\sup \{s_m \mid m \in \mathbb{N}\} \leq \sup \{t_\ell \mid \ell \in \mathbb{N}\}. \quad (\text{V.61})$$

Für $m > 2^\ell$ ist jedoch

$$\begin{aligned} s_m &\geq a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) \\ &\quad + \dots + \underbrace{(a_{2^{\ell-1}+1} + a_{2^{\ell-1}+2} + \dots + a_{2^\ell})}_{2^{\ell-1} \text{ Summanden}} \\ &\geq \frac{1}{2}a_1 + a_2 + 2a_4 + 4a_8 + \dots + 2^{\ell-1}a_{2^\ell} \\ &= \frac{1}{2} (a_1 + 2a_2 + 4a_4 + \dots + 2^\ell a_{2^\ell}) = \frac{1}{2} t_\ell, \end{aligned} \quad (\text{V.62})$$

wiederum wegen $a_n \geq a_{n+1}$. Somit ist

$$\sup \{s_m \mid m \in \mathbb{N}\} \geq \frac{1}{2} \sup \{t_\ell \mid \ell \in \mathbb{N}\}. \quad (\text{V.63})$$

□

V.5.2. Beweis des Großen Umordnungssatzes V.7

Zu (i): Seien $\sigma : \mathbb{N} \rightarrow \mathbb{N}, \sigma^{-1} : \mathbb{N} \rightarrow \mathbb{N}$ Bijektionen, so dass

$$\forall n \in \mathbb{N} : \quad a_{\sigma(n)} = b_n, \quad a_n = b_{\sigma^{-1}(n)}. \quad (\text{V.64})$$

Für $m \in \mathbb{N}$ setzen wir

$$A(m) := \max \{\sigma(1), \sigma(2), \dots, \sigma(m)\}, \quad (\text{V.65})$$

$$B(m) := \max \{\sigma^{-1}(1), \sigma^{-1}(2), \dots, \sigma^{-1}(m)\}. \quad (\text{V.66})$$

Wir betrachten nun die monoton steigenden Folgen $(s_m)_{m=1}^\infty, (t_m)_{m=1}^\infty$ in $\mathbb{R}$, wobei

$$s_m := \sum_{n=1}^m |a_n|, \quad t_m := \sum_{n=1}^m |b_n|. \quad (\text{V.67})$$

Wir beobachten, dass, für jedes $m \in \mathbb{N}$,

$$t_m := \sum_{n=1}^m |a_{\sigma(n)}| \leq \sum_{n=1}^{A(m)} |a_n| = s_{A(m)}, \quad (\text{V.68})$$

$$s_m := \sum_{n=1}^m |b_{\sigma^{-1}(n)}| \leq \sum_{n=1}^{B(m)} |b_n| = t_{B(m)}. \quad (\text{V.69})$$

Somit ist

$$\sup_{m \in \mathbb{N}} \{t_m\} \leq \sup_{m \in \mathbb{N}} \{s_{A(m)}\} \leq \sup_{m \in \mathbb{N}} \{s_m\} \leq \sup_{m \in \mathbb{N}} \{t_{B(m)}\} \leq \sup_{m \in \mathbb{N}} \{t_m\}, \quad (\text{V.70})$$

also

$$\sup_{m \in \mathbb{N}} \{s_m\} = \sup_{m \in \mathbb{N}} \{t_m\}, \quad (\text{V.71})$$

und $(s_m)_{m=1}^\infty$ ist genau dann konvergent, wenn $(t_m)_{m=1}^\infty$ konvergent ist, nach Satz IV.5 (i).

Zu (ii): Seien $m \in \mathbb{N}$ und $B(m)$ wie in (V.66). Dann ist $B(m) \geq m$ und

$$\{\sigma^{-1}(1), \sigma^{-1}(2), \dots, \sigma^{-1}(m)\} \subseteq \{1, 2, \dots, B(m)\}, \quad (\text{V.72})$$

was

$$\begin{aligned} \{1, 2, \dots, m\} &= \{\sigma(\sigma^{-1}(1)), \sigma(\sigma^{-1}(2)), \dots, \sigma(\sigma^{-1}(m))\} \\ &\subseteq \{\sigma(n) \mid n = 1, 2, \dots, B(m)\} \subseteq \{1, 2, \dots, A(B(m))\} \end{aligned} \quad (\text{V.73})$$

impliziert. Also ist $Q_m := \{\sigma(n)\}_{n=1}^{B(m)} \supseteq \{1, \dots, m\}$

$$\sum_{n=1}^{B(m)} b_n - \sum_{n=1}^m a_n = \sum_{k \in Q_m} a_k - \sum_{k \in \{1, \dots, m\}} a_k = \sum_{k \in Q_m \setminus \{1, \dots, m\}} a_k, \quad (\text{V.74})$$

und wegen $Q_m \subseteq \{1, 2, \dots, A \circ B(m)\}$ ist dann

$$\left| \sum_{n=1}^{B(m)} b_n - \sum_{n=1}^m a_n \right| \leq \sum_{k \in Q_m \setminus \{1, \dots, m\}} |a_k| \leq \sum_{n=m+1}^{A \circ B(m)} |a_n| \leq \sum_{n=m+1}^{\infty} |a_n|. \quad (\text{V.75})$$

Ist nun $\varepsilon > 0$, und ist $m_0 \in \mathbb{N}$ so groß, dass

$$\forall m \geq m_0 : \sum_{n=m_0+1}^m |a_n|, \sum_{n=m_0+1}^m |b_n| \leq \frac{\varepsilon}{3}, \quad (\text{V.76})$$

dann sind, wegen $B(m) \geq m$,

$$\left| \sum_{n=1}^{\infty} a_n - \sum_{n=1}^{m_0} a_n \right|, \left| \sum_{n=1}^{\infty} b_n - \sum_{n=1}^{B(m_0)} b_n \right| \leq \frac{\varepsilon}{3}. \quad (\text{V.77})$$

Also ist mit (V.77) und (V.75))

$$\begin{aligned} \left| \sum_{n=1}^{\infty} b_n - \sum_{n=1}^{\infty} a_n \right| &\leq \left| \sum_{n=1}^{\infty} b_n - \sum_{n=1}^{B(m_0)} b_n \right| + \left| \sum_{n=1}^{B(m_0)} b_n - \sum_{n=1}^{m_0} a_n \right| + \left| \sum_{n=1}^{m_0} a_n - \sum_{n=1}^{\infty} a_n \right| \\ &\leq \frac{2\varepsilon}{3} + \sum_{n=m_0+1}^{\infty} |a_n| \leq \varepsilon. \end{aligned} \quad (\text{V.78})$$

Da $\varepsilon > 0$ beliebig klein gewählt werden kann, folgt $\left| \sum_{n=1}^{\infty} b_n - \sum_{n=1}^{\infty} a_n \right| = 0$, also

$$\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} a_n. \quad (\text{V.79})$$

V.5.3. Beweis von Satz V.9 zum Cauchy-Produkt

Seien $a := \sum_{n=0}^{\infty} a_n$, $b := \sum_{n=0}^{\infty} b_n$, $A := \sum_{n=0}^{\infty} |a_n|$, $B := \sum_{n=0}^{\infty} |b_n|$. Sei weiterhin $\varepsilon > 0$. Wegen der absoluten Konvergenz von $(\sum_{n=0}^{\infty} a_n)_{m=0}^{\infty}$ und $(\sum_{n=0}^{\infty} b_n)_{m=0}^{\infty}$ gibt es ein $n_0 \in \mathbb{N}$, so dass, für alle $n \geq n_0$,

$$\left| a - \sum_{k=0}^n a_k \right| \leq \varepsilon, \quad \left| b - \sum_{k=0}^n b_k \right| \leq \varepsilon, \quad (\text{V.80})$$

$$\left| A - \sum_{k=0}^n |a_k| \right| \leq \frac{\varepsilon}{A+B}, \quad \left| B - \sum_{k=0}^n |b_k| \right| \leq \frac{\varepsilon}{A+B}. \quad (\text{V.81})$$

Wir beobachten nun, dass für $M \geq 2m$, $M, m \in \mathbb{N}$

$$\sum_{n=2m}^M |c_n| = \sum_{n=2m}^M \sum_{k=0}^n |a_k| |b_{n-k}| \leq \sum_{\alpha=0}^m \sum_{\beta=m}^M |a_\alpha| |b_\beta| + \sum_{\alpha=m}^M \sum_{\beta=0}^m |a_\alpha| |b_\beta|. \quad (\text{V.82})$$

Also ist, für $M \geq 2m \geq m \geq n_0$,

$$\begin{aligned} \sum_{n=2m}^M |c_n| &\leq \underbrace{\left(\sum_{\alpha=0}^m |a_\alpha| \right)}_{\leq A} \underbrace{\left(\sum_{\beta=m}^M |b_\beta| \right)}_{\leq \varepsilon/A+B} + \underbrace{\left(\sum_{\alpha=m}^M |a_\alpha| \right)}_{\leq \varepsilon/A+B} \underbrace{\left(\sum_{\beta=0}^m |b_\beta| \right)}_{\leq B} \\ &\leq (A+B) \cdot \left(\frac{\varepsilon}{A+B} \right) = \varepsilon. \end{aligned} \quad (\text{V.83})$$

Somit ist $(\sum_{n=0}^m c_n)_{m=0}^\infty$ absolut konvergent. Außerdem ist

$$\begin{aligned} \left| ab - \sum_{n=0}^{2m} c_n \right| &\leq \left| ab - \left(\sum_{\alpha=0}^m a_\alpha \right) \left(\sum_{\beta=0}^m b_\beta \right) \right| + \left| \left(\sum_{\alpha=0}^m a_\alpha \right) \left(\sum_{\beta=0}^m b_\beta \right) - \sum_{n=0}^{2m} c_n \right| \\ &\leq \left| a - \sum_{\alpha=0}^m a_\alpha \right| \cdot |b| + \left| \sum_{\alpha=0}^m a_\alpha \right| \cdot \left| b - \sum_{\beta=0}^m b_\beta \right| \\ &\quad + \left(\sum_{\alpha=0}^m |a_\alpha| \right) \left(\sum_{\beta=m}^{2m} |b_\beta| \right) + \left(\sum_{\alpha=m}^{2m} |a_\alpha| \right) \left(\sum_{\beta=0}^m |b_\beta| \right) \\ &\leq \varepsilon(|b| + A + A + B) \leq 2(A+B)\varepsilon, \end{aligned} \quad (\text{V.84})$$

also gilt auch (V.23).

V.5.4. Rückführung des Quotientenkriteriums auf das Wurzelkriterium

Beweis von Satz V.12 [Quotientenkriterium].

Zu (i) Sei $\beta := \limsup_{n \rightarrow \infty} \left\{ \frac{|a_{n+1}|}{|a_n|} \right\} \in \mathbb{R}_0^+$. Dann gibt es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass

$$\forall n \geq n_0 : \quad \frac{|a_{n+1}|}{|a_n|} \leq \beta + \varepsilon, \quad (\text{V.85})$$

nach Satz IV.7 (ii). Also ist

$$\forall n \geq n_0 : \quad |a_n| \leq \frac{|a_n|}{|a_{n-1}|} \cdot \frac{|a_{n-1}|}{|a_{n-2}|} \cdots \frac{|a_{n_0+1}|}{|a_{n_0}|} \cdot |a_{n_0}| \leq (\beta + \varepsilon)^{n-n_0} |a_{n_0}|, \quad (\text{V.86})$$

und daher

$$\sqrt[n]{|a_n|} \leq (\beta + \varepsilon) \cdot \sqrt[n]{|a_{n_0}| \cdot (\beta + \varepsilon)^{-n_0}} \rightarrow \beta + \varepsilon, \quad \text{für } n \rightarrow \infty. \quad (\text{V.87})$$

Somit erhalten wir, dass

$$\limsup_{n \rightarrow \infty} \left\{ \sqrt[n]{|a_n|} \right\} \leq \beta + \varepsilon. \quad (\text{V.88})$$

Im Limes $\varepsilon \rightarrow 0$ folgt damit, dass

$$\limsup_{n \rightarrow \infty} \left\{ \sqrt[n]{|a_n|} \right\} \leq \beta = \limsup_{n \rightarrow \infty} \left\{ \frac{|a_{n+1}|}{|a_n|} \right\}, \quad (\text{V.89})$$

und (i) folgt nun aus dem Wurzelkriterium.

Zu (ii) Sind umgekehrt $n_0 \in \mathbb{N}$ und

$$\forall n \geq n_0 : \quad \frac{|a_{n+1}|}{|a_n|} \geq 1, \quad (\text{V.90})$$

so ist

$$\forall n \geq n_0 : \quad |a_n| \geq \frac{|a_n|}{|a_{n-1}|} \cdot \frac{|a_{n-1}|}{|a_{n-2}|} \cdots \frac{|a_{n_0+1}|}{|a_{n_0}|} \cdot |a_{n_0}| \geq |a_{n_0}| > 0. \quad (\text{V.91})$$

Also ist $(a_n)_{n=1}^{\infty}$ keine Nullfolge, und das Nullfolgenkriterium ergibt die Divergenz der zugehörigen Reihe. $\square$

VI. Stetigkeit

VI.1. Offene, abgeschlossene und kompakte Mengen

Definition VI.1.

(i) Für alle $x \in \mathbb{K}$ und $r > 0$ heißt

$$B_{\mathbb{K}}(x, r) := \left\{ y \in \mathbb{K} \mid |y - x| < r \right\} \quad (\text{VI.1})$$

offene Kugel vom Radius r um x .

(ii) Sei $A \subseteq \mathbb{K}$ eine Teilmenge. Ein Punkt $x \in A$ heißt **innerer Punkt (I.P.) von A**

$$:\Leftrightarrow \quad \exists r > 0 : \quad B_{\mathbb{K}}(x, r) \subseteq A. \quad (\text{VI.2})$$

(iii) Sei $D \subseteq \mathbb{K}$. Ein Punkt $x \in \mathbb{K}$ heißt **Häufungspunkt (H.P.) von D**

$$:\Leftrightarrow \quad \forall r > 0 : \quad B_{\mathbb{K}}(x, r) \cap (D \setminus \{x\}) \neq \emptyset \quad (\text{VI.3})$$

$$\Leftrightarrow \quad \forall r > 0 \exists y \in D, y \neq x : \quad |y - x| < r.$$

Bemerkungen und Beispiele.

- Für $\mathbb{K} = \mathbb{R}$, $x \in \mathbb{R}$ und $r > 0$ ist $B_{\mathbb{R}}(x, r) = (x - r, x + r)$ das *offene Intervall der Länge $2r$ um x* .
- Für $\mathbb{K} = \mathbb{C}$, $z \in \mathbb{C}$ und $r > 0$ bezeichnet man $B_{\mathbb{C}}(z, r) \equiv D(z, r) := \{w \in \mathbb{C} : |w - z| < r\}$ als *offene Kreisscheibe vom Radius r um z* .
- Für $\mathbb{K} = \mathbb{R}$ und $a, b \in \mathbb{R}$, mit $a < b$, sind

$$[a, b] = \{x \in \mathbb{R} \mid x \text{ ist H.P. von } [a, b]\} = \{x \in \mathbb{R} \mid x \text{ ist H.P. von } (a, b)\}, \quad (\text{VI.4})$$

$$(a, b) = \{x \in \mathbb{R} \mid x \text{ ist I.P. von } [a, b]\} = \{x \in \mathbb{R} \mid x \text{ ist I.P. von } (a, b)\}. \quad (\text{VI.5})$$

- Für $\mathbb{K} = \mathbb{R}$ und $A = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$ ist die Menge der inneren Punkte leer, und 0 ist der einzige Häufungspunkt von A .

Definition VI.2.

(i) Eine Teilmenge $A \subseteq \mathbb{K}$ heißt **offen**

$$:\Leftrightarrow \quad \text{Jeder Punkt in } A \text{ ist ein I.P. oder } A = \emptyset$$

$$\Leftrightarrow \quad \forall x \in A \exists r > 0 : \quad B_{\mathbb{K}}(x, r) \subseteq A. \quad (\text{VI.6})$$

(ii) Eine Teilmenge $D \subseteq \mathbb{K}$ heißt **abgeschlossen**

$$\begin{aligned} &:\Leftrightarrow D \text{ enthält alle seine H.P. oder } D = \emptyset \\ &\Leftrightarrow \forall x \in \mathbb{K} : (x \text{ ist H.P. von } D \Rightarrow x \in D). \end{aligned} \quad (\text{VI.7})$$

(iii) Eine Teilmenge $C \subseteq \mathbb{K}$ heißt **beschränkt**

$$:\Leftrightarrow \exists R < \infty : C \subseteq B_{\mathbb{K}}(0, R). \quad (\text{VI.8})$$

(iv) Eine Teilmenge $K \subseteq \mathbb{K}$ heißt **kompakt**

$$:\Leftrightarrow K \text{ ist beschränkt und abgeschlossen.} \quad (\text{VI.9})$$

Bemerkungen und Beispiele.

- $\mathbb{K} \subseteq \mathbb{K}$ und $\emptyset \subseteq \mathbb{K}$ sind stets sowohl offen als auch abgeschlossen. Es gibt also Mengen, die beide Eigenschaften besitzen.
- Für $\mathbb{K} = \mathbb{R}$ und $a, b \in \mathbb{R}$, mit $a < b$, ist (a, b) offen und nicht abgeschlossen, hingegen ist $[a, b]$ abgeschlossen und nicht offen.
- Für $\mathbb{K} = \mathbb{R}$ und $a, b \in \mathbb{R}$, mit $a < b$, sind $(a, b] = \{x \in \mathbb{R} | a < x \leq b\}$ und $[a, b) = \{x \in \mathbb{R} | a \leq x < b\}$ weder offen noch abgeschlossen. Es gibt also Mengen, die keine der beiden Eigenschaften besitzen.
- Für $\mathbb{K} = \mathbb{R}$ und $a, b \in \mathbb{R}$, mit $a < b$, ist $[a, b]$ abgeschlossen und beschränkt, da $[a, b] \subseteq (-R, R) = B_{\mathbb{R}}(0, R)$ mit $R := |a| + |b| + 1$. Also ist $[a, b]$ kompakt.

Satz VI.3. Für $A \subseteq \mathbb{K}$ sind folgende Aussagen gleichwertig:

$$A \text{ ist offen} \quad \Leftrightarrow \quad A^c := \mathbb{K} \setminus A \text{ ist abgeschlossen.} \quad (\text{VI.10})$$

Lemma VI.4.

(i) Sind $J \neq \emptyset$ eine Indexmenge und $\{A_j\}_{j \in J} \subseteq \mathfrak{P}(\mathbb{K})$ eine Familie offener Teilmengen von $\mathbb{K}$, so ist auch ihre Vereinigung

$$\bigcup_{j \in J} A_j \quad \text{offen.} \quad (\text{VI.11})$$

(ii) Sind $A_1, A_2, \dots, A_N \subseteq \mathbb{K}$ endliche viele offene Teilmengen von $\mathbb{K}$, so ist auch ihr Durchschnitt

$$A_1 \cap A_2 \cap \dots \cap A_N \quad \text{offen.} \quad (\text{VI.12})$$

Bemerkungen und Beispiele.

- Glg. (VI.12) ist im Allgemeinen für den Durchschnitt unendlich vieler offener Mengen nicht richtig. Sind beispielsweise $A_n := (-\frac{1}{n}, \frac{1}{n})$, für $n \in \mathbb{N}$, so ist $\bigcap_{n=1}^{\infty} A_n = \{0\}$ nicht offen.

VI.2. Stetige Abbildungen

Definition VI.5. Seien $U \subseteq \mathbb{K}$, $U \neq \emptyset$ eine nichtleere Teilmenge reeller oder komplexer Zahlen, $x_0 \in U$ und $f : U \rightarrow \mathbb{K}$ eine Abbildung.

(i) f heißt **stetig in x_0**

$$:\Leftrightarrow \forall \varepsilon > 0 \exists \delta > 0 \forall x \in U \cap B_{\mathbb{K}}(x_0, \delta) : f(x) \in B_{\mathbb{K}}(f(x_0), \varepsilon). \quad (\text{VI.13})$$

$$\Leftrightarrow \forall \varepsilon > 0 \exists \delta > 0 \forall x \in U, |x - x_0| < \delta : |f(x) - f(x_0)| < \varepsilon.$$

(ii) f heißt **stetig auf U** $:\Leftrightarrow$

$$\forall x_0 \in U : f \text{ ist stetig in } x_0. \quad (\text{VI.14})$$

(iii) Wir schreiben

$$\lim_{x \rightarrow x_0} f(x) = y \quad :\Leftrightarrow \quad (\text{VI.15})$$

$$\forall (x_n)_{n=1}^{\infty} \in (U \setminus \{x_0\})^{\mathbb{N}}, \lim_{n \rightarrow \infty} \{x_n\} = x_0 : \lim_{n \rightarrow \infty} \{f(x_n)\} = y.$$

Bemerkungen und Beispiele.

Für $U = \mathbb{K} = \mathbb{R}$ sind folgende Funktionen stetig auf $\mathbb{R}$:

- Polynome $p(x) = \sum_{n=0}^m c_n x^n$,
- trigonometrische Funktionen $\sin(x), \cos(x)$,
- die Exponentialfunktion e^x .

(Begründungen kommen erst später.)

Satz VI.6. Seien $U \subseteq \mathbb{K}$, $U \neq \emptyset$ eine nichtleere Teilmenge reeller oder komplexer Zahlen, $x_0 \in U$ und $f : U \rightarrow \mathbb{K}$ eine Abbildung. Dann sind folgende Aussagen gleichwertig:

$$\left[f \text{ ist stetig in } x_0 \in U \right] \Leftrightarrow \left[\lim_{x \rightarrow x_0} \{f(x)\} = f(x_0) \right]. \quad (\text{VI.16})$$

In diesem Fall nennt man die Eigenschaft $\lim_{x \rightarrow x_0} \{f(x)\} = f(x_0)$ **Folgenstetigkeit von f in x_0** .

Beweis.

“ $\Rightarrow$ ”: Seien $(x_n)_{n=1}^{\infty} \in U^{\mathbb{N}}$ eine Folge in U mit $x_n \rightarrow x_0$ und $\varepsilon > 0$. Nach (VI.13) gibt es ein $\delta > 0$, so dass mit $|x - x_0| < \delta$ auch $|f(x) - f(x_0)| < \varepsilon$ gilt. Da $x_n \rightarrow x_0$ konvergiert, gibt es zu obigem $\delta > 0$ auch ein $n_0 \in \mathbb{N}$, sodass

$$\forall n \geq n_0 : |x_n - x_0| < \delta, \quad (\text{VI.17})$$

und daher gilt auch

$$\forall n \geq n_0 : |f(x_n) - f(x_0)| < \varepsilon. \quad (\text{VI.18})$$

Zusammenfassend erhalten wir somit

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : |f(x_n) - f(x_0)| < \varepsilon, \quad (\text{VI.19})$$

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

“ $\Leftarrow$ ”: Sei f nicht stetig in x_0 , d.h.

$$\exists \varepsilon > 0 \forall \delta > 0 \exists x \in U, |x - x_0| < \delta : |f(x) - f(x_0)| \geq \varepsilon. \quad (\text{VI.20})$$

Sei $\varepsilon > 0$ nun eine solche Zahl. Setzen wir $\delta := 1/n$, so gibt es gemäß (VI.20) ein $x_n \in U$ mit $|x_n - x_0| < 1/n$ und $|f(x_n) - f(x_0)| \geq \varepsilon$, für alle $n \in \mathbb{N}$. Damit ist $(x_n)_{n=1}^\infty \in U^\mathbb{N}$ eine konvergente Zahlenfolge in $\mathbb{K}$ mit $x_n \rightarrow x_0$, weil $|x_n - x_0| < 1/n$. Andererseits ist aber $|f(x_n) - f(x_0)| \geq \varepsilon > 0$, und die Zahlenfolge $(f(x_n))_{n=1}^\infty \in \mathbb{K}^\mathbb{N}$ konvergiert nicht gegen $f(x_0)$. $\square$

Satz VI.7. Seien $U \subseteq \mathbb{K}$, $f, g : U \rightarrow \mathbb{K}$ stetig in $x_0 \in U$ und $\alpha \in \mathbb{K}$.

(i) Dann sind auch $f + \alpha g, f \cdot g : U \rightarrow \mathbb{K}$ stetig in x_0 , und es gelten

$$\lim_{x \rightarrow x_0} \{f(x) + \alpha g(x)\} = \lim_{x \rightarrow x_0} \{f(x)\} + \alpha \lim_{x \rightarrow x_0} \{g(x)\} = f(x_0) + \alpha g(x_0). \quad (\text{VI.21})$$

$$\lim_{x \rightarrow x_0} \{f(x) \cdot g(x)\} = \lim_{x \rightarrow x_0} \{f(x)\} \cdot \lim_{x \rightarrow x_0} \{g(x)\} = f(x_0) \cdot g(x_0). \quad (\text{VI.22})$$

(ii) Ist $g(x_0) \neq 0$, so gibt es ein $\delta > 0$, sodass

$$\forall x \in U \cap B_{\mathbb{K}}(x_0, \delta) : |g(x)| \geq \frac{1}{2}|g(x_0)| \geq 0. \quad (\text{VI.23})$$

Weiterhin ist die Abbildung $f/g : U \cap B_{\mathbb{K}}(x_0, \delta) \rightarrow \mathbb{K}$ stetig in x_0 , und es gilt

$$\lim_{x \rightarrow x_0} \left\{ \frac{f(x)}{g(x)} \right\} = \frac{\lim_{x \rightarrow x_0} \{f(x)\}}{\lim_{x \rightarrow x_0} \{g(x)\}} = \frac{f(x_0)}{g(x_0)}. \quad (\text{VI.24})$$

Beweis. Folgt direkt aus den Sätzen VI.6, IV.2 und IV.3. $\square$

Satz VI.8. Seien $U \subseteq \mathbb{K}$, $U \neq \emptyset$, und $g : U \rightarrow \mathbb{K}$ stetig in $x_0 \in U$. Seien weiterhin $V := g(U) \subseteq \mathbb{K}$ und $f : V \rightarrow \mathbb{K}$ stetig in $g(x_0) \in V$. Dann ist $f \circ g : U \rightarrow \mathbb{K}$ stetig in $x_0 \in U$, und es gilt

$$\lim_{x \rightarrow x_0} \{f(g(x))\} = f(g(x_0)). \quad (\text{VI.25})$$

Beweis. Auch dieser Satz folgt unmittelbar aus Satz VI.6. $\square$

Bemerkungen und Beispiele.

- Offensichtlich sind $f_0(x) = 1$ und $f_1(x) = x$ stetige Funktionen auf $\mathbb{K}$. Die Stetigkeit eines Polynoms

$$\begin{aligned} p(x) &= c_0 + c_1 x + c_2 x^2 + \dots + c_N x^N \\ &= (c_0 \cdot f_0 + c_1 \cdot f_1 + c_2 \cdot [f_1]^2 + \dots + c_N \cdot [f_1]^N)[x] \end{aligned} \quad (\text{VI.26})$$

folgt nun aus Satz VI.7.

- Seien $z_0 \in \mathbb{C}$ und $z \in D(z_0, 1)$. Dann ist

$$\begin{aligned}
 |e^z - e^{z_0}| &= |e^{z-z_0} - 1| \cdot |e^{z_0}| = \left| \sum_{n=1}^{\infty} \frac{(z-z_0)^n}{n!} \right| \cdot \left| \sum_{n=0}^{\infty} \frac{z_0^n}{n!} \right| \\
 &\leq \left(\sum_{n=1}^{\infty} \frac{|z-z_0|^n}{n!} \right) \cdot \left(\sum_{n=0}^{\infty} \frac{|z_0|^n}{n!} \right) = |z-z_0| \left(\sum_{k=0}^{\infty} \frac{|z-z_0|^k}{(k+1)!} \right) \cdot \left(\sum_{n=0}^{\infty} \frac{|z_0|^n}{n!} \right) \\
 &\leq |z-z_0| \left(\sum_{k=0}^{\infty} \frac{|z-z_0|^k}{k!} \right) \cdot \left(\sum_{n=0}^{\infty} \frac{|z_0|^n}{n!} \right) \leq |z-z_0| e^{|z-z_0|} e^{|z_0|} \leq |z-z_0| e^{|z_0|+1}.
 \end{aligned} \tag{VI.27}$$

Somit ist $\lim_{z \rightarrow z_0} \{e^z\} = e^{z_0}$, für alle $z_0 \in \mathbb{C}$, und $\exp : \mathbb{C} \rightarrow \mathbb{C}$ ist stetig auf $\mathbb{C}$.

- Mit $\mathbb{K} = \mathbb{C}$, $g(z) = -z^2$ und $f(w) = e^w$ folgt aus der Stetigkeit von f und g auf $\mathbb{C}$ folgt mit Satz VI.8 auch die Stetigkeit von $(f \circ g)[z] = \exp[-z^2]$ auf $\mathbb{C}$.

VI.3. Erhaltung topologischer Eigenschaften durch stetige Funktionen

Satz VI.9. Sei $f : \mathbb{K} \rightarrow \mathbb{K}$ eine Abbildung. Dann sind folgende drei Aussagen gleichwertig:

$$f \text{ ist stetig auf } \mathbb{K}; \tag{VI.28}$$

$\Leftrightarrow$ Ist $V \subseteq \mathbb{K}$ offen, so ist auch das Urbild $f^{-1}[V]$ offen, d.h.

$$\{V \subseteq \mathbb{K} \text{ ist offen} \Rightarrow f^{-1}[V] = \{x \in \mathbb{K} | f(x) \in V\} \text{ ist offen}\}; \tag{VI.29}$$

$\Leftrightarrow$ Ist $V \subseteq \mathbb{K}$ abgeschlossen, so ist auch das Urbild $f^{-1}[V]$ abgeschlossen, d.h.

$$\{V \subseteq \mathbb{K} \text{ ist abgeschlossen} \Rightarrow f^{-1}[V] = \{x \in \mathbb{K} | f(x) \in V\} \text{ ist abgeschlossen}\}. \tag{VI.30}$$

Satz VI.10. Sind $f : \mathbb{K} \rightarrow \mathbb{K}$ eine auf $\mathbb{K}$ stetige Abbildung und $K \subseteq \mathbb{K}$ kompakt, so ist auch die Bildmenge $f(K) := \{f(x) | x \in K\}$ von K kompakt.

Korollar VI.11. Sind $f : \mathbb{K} \rightarrow \mathbb{R}$ eine auf $\mathbb{K}$ stetige Abbildung und $K \subseteq \mathbb{K}$ kompakt, so gibt es $x_{\min}, x_{\max} \in K$, so dass

$$f(x_{\min}) = \inf\{f(K)\} \quad \text{und} \quad f(x_{\max}) = \sup\{f(K)\}. \tag{VI.31}$$

Eine weiterer wichtiger Fakt ist, dass stetige Abbildungen auf Kompakta auch stets *gleichmäßig stetig* sind, wie der folgende Satz belegt.

Satz VI.12. Sind $f : \mathbb{K} \rightarrow \mathbb{K}$ eine auf $\mathbb{K}$ stetige Abbildung und $K \subseteq \mathbb{K}$ kompakt, so ist f auf K auch **gleichmäßig stetig** : $\Leftrightarrow$

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, x' \in K : \left\{ |x - x'| < \delta \Rightarrow |f(x) - f(x')| < \varepsilon \right\}. \tag{VI.32}$$

Bemerkungen und Beispiele.

- Der Unterschied der gleichmäßigen Stetigkeit zur (einfachen) Stetigkeit liegt in der Unabhängigkeit von $\delta > 0$ von der auf Stetigkeit zu untersuchenden Stelle $x \in K$. Kurz gesagt hängt δ bei einfacher Stetigkeit von ε und x ab, während δ bei gleichmäßiger Stetigkeit nur von ε abhängt.
- Für Korollar VI.11 hat sich auch folgende Sprechweise eingebürgert:
Eine stetige Funktion nimmt auf einem Kompaktum ihr Maximum und Minimum an.
- Außerdem schreibt man in dem Fall, dass für $A \subseteq \mathbb{R}$ auch $\inf\{A\} \in A$ bzw. $\sup\{A\} \in A$ gilt,

$$\inf\{A\} =: \min\{A\} \quad \text{und} \quad \sup\{A\} =: \max\{A\}. \quad (\text{VI.33})$$

- Sind $K \subseteq \mathbb{K}$ kompakt und $f : K \rightarrow \mathbb{R}^+$ stetig auf K , so ist

$$\inf\{f(K)\} = \min\{f(K)\} > 0. \quad (\text{VI.34})$$

- Für $f : \mathbb{R}_0^+ \rightarrow \mathbb{R}^+$, $f(x) = e^{-x}$, ist $U := \mathbb{R}_0^+$ nicht kompakt, und daher überrascht es auch nicht, dass

$$\inf\{f(U)\} = 0 \notin f(U). \quad (\text{VI.35})$$

- Für $a > 0$ ist $K := [0, a]$ kompakt, und mit $f : K \rightarrow \mathbb{R}_0^+$ und $f(x) = e^{-x}$ gilt

$$\inf\{f(K)\} = \min\{f(K)\} = e^{-a} \in f(K). \quad (\text{VI.36})$$

VII. Differentiation

In diesem Kapitel wollen wir den Begriff der Ableitung für reelle Funktionen einer reellen Variablen diskutieren. Der allgemeine Fall einer vektorwertigen Funktion wird später noch ausführlich behandelt.

VII.1. Differenzierbare Funktionen

Definition VII.1. Seien $U \subseteq \mathbb{R}$ offen und $f : U \rightarrow \mathbb{R}$ eine Abbildung.

(i) f ist bei $x_0 \in U$ **differenzierbar** $:\Leftrightarrow$

$$f'(x_0) := \lim_{x \rightarrow x_0} \left\{ \frac{f(x) - f(x_0)}{x - x_0} \right\} \in \mathbb{R} \quad (\text{VII.1})$$

existiert. In diesem Fall heißt $f'(x_0)$ **Ableitung von f bei x_0** , und wir schreiben auch

$$f'(x_0) =: \frac{df(x_0)}{dx} =: \left(\frac{d}{dx} f \right)(x_0). \quad (\text{VII.2})$$

Der Quotient $(f(x) - f(x_0))/(x - x_0)$, für $x \neq x_0$ heißt **Differenzenquotient**.

(ii) f ist auf U differenzierbar $:\Leftrightarrow$

$$\forall x_0 \in U : \quad f \text{ ist bei } x_0 \text{ differenzierbar.} \quad (\text{VII.3})$$

Lemma VII.2. Seien $U \subseteq \mathbb{R}$ offen und $f : U \rightarrow \mathbb{R}$ eine Abbildung. Dann gilt folgende Äquivalenz:

$$\begin{aligned} \left[f \text{ ist bei } x_0 \in U \text{ differenzierbar} \right] &\Leftrightarrow \\ &\left\{ \exists c \in \mathbb{R} \forall \varepsilon > 0 \exists \delta > 0 \forall |x - x_0| \leq \delta : \right. \\ &\quad \left. \left| f(x) - (f(x_0) + c \cdot (x - x_0)) \right| \leq \varepsilon \cdot |x - x_0| \right\}. \end{aligned} \quad (\text{VII.4})$$

In diesem Fall ist $c = f'(x_0)$.

Beweis. Analog zum Beweis von Satz VI.6 (Äquivalenz von Stetigkeit und Folgenstetigkeit). $\square$

Bemerkungen und Beispiele.

- Glg. (VII.4) sagt, dass $f(x_0) + f'(x) \cdot (x - x_0)$ die beste lineare Approximation für $f(x)$ in einer Umgebung von x_0 ist.

Die Gleichung lässt sich auch schreiben als

$$\lim_{x \rightarrow x_0} \left| \frac{f(x) - (f(x_0) + c \cdot (x - x_0))}{x - x_0} \right| = 0.$$

Lemma VII.3. Seien $U \subseteq \mathbb{R}$ offen und $f : U \rightarrow \mathbb{R}$ bei $x_0 \in U$ differenzierbar. Dann ist f in $x_0 \in U$ stetig.

Beweis. Sei $(x_n)_{n=1}^\infty$ eine Folge in $U \setminus \{x_0\}$ mit $x_n \rightarrow x_0$, $n \rightarrow \infty$. Dann ist

$$\begin{aligned} \lim_{n \rightarrow \infty} \{f(x_n)\} &= \lim_{n \rightarrow \infty} \left\{ \left(\frac{f(x_n) - f(x_0)}{x_n - x_0} \right) \cdot (x_n - x_0) + f(x_0) \right\} \\ &= \lim_{n \rightarrow \infty} \left\{ \frac{f(x_n) - f(x_0)}{x_n - x_0} \right\} \cdot \lim_{n \rightarrow \infty} \{x_n - x_0\} + f(x_0) \\ &= f'(x_0) \cdot 0 + f(x_0) = f(x_0). \end{aligned} \tag{VII.5}$$

□

Satz VII.4. Seien $U \subseteq \mathbb{R}$ offen, $f, g : U \rightarrow \mathbb{R}$ differenzierbar bei $x_0 \in U$ und $\alpha \in \mathbb{R}$. Dann sind $f + \alpha g$, $f \cdot g : U \rightarrow \mathbb{R}$ differenzierbar bei x_0 , und es gelten

(i) die **Linearität**, d.h. es gilt

$$(f + \alpha g)'(x_0) = f'(x_0) + \alpha g'(x_0), \tag{VII.6}$$

(ii) die **Produktregel**,

$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0)g'(x_0), \tag{VII.7}$$

(iii) und, für $g(x_0) \neq 0$, die **Quotientenregel**

$$\left(\frac{f}{g} \right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}. \tag{VII.8}$$

Beweis. Gleichung (VII.6) folgt aus

$$\begin{aligned} \lim_{x \rightarrow x_0} \left\{ \frac{(f + \alpha g)(x) - (f + \alpha g)(x_0)}{x - x_0} \right\} &= \lim_{x \rightarrow x_0} \left\{ \frac{f(x) - f(x_0)}{x - x_0} \right\} + \alpha \lim_{x \rightarrow x_0} \left\{ \frac{g(x) - g(x_0)}{x - x_0} \right\} \\ &= f'(x_0) + \alpha g'(x_0). \end{aligned} \tag{VII.9}$$

Für den Beweis von (VII.7) benutzen wir die Stetigkeit von f und g bei x_0 , nach Lemma VII.3. Damit erhalten wir

$$\begin{aligned}
 (f \cdot g)'(x_0) &= \lim_{x \rightarrow x_0} \left\{ \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \right\} \\
 &= \lim_{x \rightarrow x_0} \left\{ \frac{f(x)[g(x) - g(x_0)] + [f(x) - f(x_0)]g(x_0)}{x - x_0} \right\} \\
 &= \lim_{x \rightarrow x_0} \{f(x)\} \cdot \lim_{x \rightarrow x_0} \left\{ \frac{g(x) - g(x_0)}{x - x_0} \right\} + \lim_{x \rightarrow x_0} \left\{ \frac{f(x) - f(x_0)}{x - x_0} \right\} \cdot g(x_0) \\
 &= f(x_0)g'(x_0) + f'(x_0) \cdot g(x_0).
 \end{aligned} \tag{VII.10}$$

Für $g(x_0) \neq 0$ ist

$$\begin{aligned}
 \left(\frac{1}{g}\right)'(x_0) &= \lim_{x \rightarrow x_0} \left\{ \frac{1}{(x - x_0)} \left(\frac{1}{g(x)} - \frac{1}{g(x_0)} \right) \right\} \\
 &= \lim_{x \rightarrow x_0} \left\{ \frac{1}{g(x)g(x_0)} \cdot \frac{g(x_0) - g(x)}{x - x_0} \right\} = \frac{-g'(x_0)}{g(x_0)^2},
 \end{aligned} \tag{VII.11}$$

was mit Hilfe der Produktregel über $\frac{f}{g} = f \cdot \frac{1}{g}$ sofort die Quotientenregel (VII.8) impliziert. $\square$

Satz VII.5 (Kettenregel). *Seien $U, V \subseteq \mathbb{R}$ offen, $g : U \rightarrow V$ differenzierbar bei $x_0 \in U$ und $f : V \rightarrow \mathbb{R}$ differenzierbar bei $g(x_0) \in V$. Dann ist $[f \circ g] : U \rightarrow \mathbb{R}$ differenzierbar bei $x_0 \in U$, und es gilt*

$$(f \circ g)'(x_0) = f'(g(x_0)) \cdot g'(x_0). \tag{VII.12}$$

Beweis. (Beweisidee) Da g bei x_0 differenzierbar ist, ist g in x_0 auch stetig, und mit $x_n \rightarrow x_0$ konvergiert auch $g(x_n) \rightarrow g(x_0)$. Damit gilt

$$\begin{aligned}
 (f \circ g)'(x_0) &= \lim_{x \rightarrow x_0} \left\{ \frac{f(g(x)) - f(g(x_0))}{x - x_0} \right\} \\
 &= \lim_{x \rightarrow x_0} \left\{ \left(\frac{f(g(x)) - f(g(x_0))}{g(x) - g(x_0)} \right) \cdot \left(\frac{g(x) - g(x_0)}{x - x_0} \right) \right\} \\
 &= \lim_{x \rightarrow x_0} \left\{ \frac{f(g(x)) - f(g(x_0))}{g(x) - g(x_0)} \right\} \cdot \lim_{x \rightarrow x_0} \left\{ \frac{g(x) - g(x_0)}{x - x_0} \right\} = f'(g(x_0)) \cdot g'(x_0).
 \end{aligned} \tag{VII.13}$$

Leider bereitet das Dividieren durch $g(x) - g(x_0)$ Schwierigkeiten, weil wir a priori nicht wissen, dass $g(x) - g(x_0) \neq 0$ gilt. Der volle Beweis ist daher formal etwas aufwändiger (aber ohne echte zusätzliche Schwierigkeiten). $\square$

Bemerkungen und Beispiele.

- Seien $U := \mathbb{R}$ und $f_n(x) := x^n$, $n \in \mathbb{N}_0$. Dann gilt

$$f'_n(x) = n x^{n-1}. \tag{VII.14}$$

Um dies einzusehen, betrachten wir zunächst $n = 0, 1$. $f_0(x) - f_0(x_0) = 1 - 1 = 0$, für alle $x, x_0 \in \mathbb{R}$, also gilt

$$f'_0 = 0 \quad (= 0 \cdot x^{-1}). \quad (\text{VII.15})$$

Weiterhin ist

$$\lim_{x \rightarrow x_0} \left\{ \frac{f_1(x) - f_1(x_0)}{x - x_0} \right\} = \lim_{x \rightarrow x_0} \left\{ \frac{x - x_0}{x - x_0} \right\} = 1, \quad (\text{VII.16})$$

also

$$f'_1(x) = 1 \quad (= 1 \cdot x^0). \quad (\text{VII.17})$$

Gelte nun (VII.14) für alle $n \in \{0, 1, 2, \dots, N\}$. Dann ist nach der Produktregel

$$\begin{aligned} f'_{N+1}(x) &= (f_1 \cdot f_N)'(x) = f'_1(x) \cdot f_N(x) + f_1(x) \cdot f'_N(x) \\ &= 1 \cdot x^N + x \cdot N \cdot x^{N-1} = (N+1) \cdot x^N. \end{aligned} \quad (\text{VII.18})$$

Somit folgt (VII.14) für alle $n \in \mathbb{N}_0$ durch Induktion.

- Seien $U := \mathbb{R} \setminus \{0\}$ und $f_n(x) = x^n, n \in \mathbb{Z}$. Dann gilt wiederum

$$f'_n(x) = n x^{n-1}. \quad (\text{VII.19})$$

Für $n < 0$ folgt dies aus der Quotientenregel,

$$\begin{aligned} f'_n(x) &= \left(\frac{f_0}{f_{-n}} \right)'(x) = \frac{1}{f_{-n}(x)^2} \left(f'_0(x) \cdot f_{-n}(x) - f_0(x) \cdot f'_{-n}(x) \right) \\ &= \frac{1}{x^{-2n}} \cdot \left(0 - 1 \cdot (-n) \cdot x^{-n-1} \right) = n \cdot x^{n-1}. \end{aligned} \quad (\text{VII.20})$$

- Sei $U = \mathbb{R}$, dann gilt

$$(e^x)' = e^x. \quad (\text{VII.21})$$

Für $x \neq x_0$ ist nämlich

$$\frac{e^x - e^{x_0}}{x - x_0} - e^{x_0} = e^{x_0} \left(\frac{e^{x-x_0} - 1}{x - x_0} - 1 \right) = e^{x_0} \left(\frac{e^{x-x_0} - 1 - (x - x_0)}{x - x_0} \right) \quad (\text{VII.22})$$

und

$$\begin{aligned} |e^{x-x_0} - 1 - (x - x_0)| &\leq \sum_{k=2}^{\infty} \frac{|x - x_0|^k}{k!} = \sum_{n=0}^{\infty} \frac{|x - x_0|^{n+2}}{(n+2)!} \\ &\leq |x - x_0|^2 \cdot \sum_{n=0}^{\infty} \frac{|x - x_0|^n}{n!} = |x - x_0|^2 \cdot e^{|x-x_0|}. \end{aligned} \quad (\text{VII.23})$$

Also ist

$$\begin{aligned} \lim_{x \rightarrow x_0} \left| \frac{e^x - e^{x_0}}{x - x_0} - e^{x_0} \right| &= e^{x_0} \cdot \lim_{x \rightarrow x_0} \left| \frac{e^{x-x_0} - 1 - (x - x_0)}{x - x_0} \right| \\ &\leq e^{x_0} \cdot \lim_{x \rightarrow x_0} \{ |x - x_0| \cdot e^{|x-x_0|} \} = 0. \end{aligned} \quad (\text{VII.24})$$

VII.2. Minima, Maxima und Mittelwertsatz

Definition VII.6. Seien $X \subseteq \mathbb{K}$ eine Menge und $f : X \rightarrow \mathbb{R}$ eine reelle Funktion auf X .

(i) Ein Punkt $x_0 \in X$ heißt **lokales Minimum** von $f : \Leftrightarrow$

$$\exists \delta > 0 \forall x \in B_{\mathbb{K}}(x_0, \delta) \cap X : \quad f(x) \geq f(x_0). \quad (\text{VII.25})$$

(ii) Ein Punkt $x_0 \in X$ heißt **globales Minimum** von $f : \Leftrightarrow$

$$\forall x \in X : \quad f(x) \geq f(x_0). \quad (\text{VII.26})$$

(iii) Ein Punkt $x_0 \in X$ heißt **lokales Maximum** von $f : \Leftrightarrow$

$$\exists \delta > 0 \forall x \in B_{\mathbb{K}}(x_0, \delta) \cap X : \quad f(x) \leq f(x_0). \quad (\text{VII.27})$$

(iv) Ein Punkt $x_0 \in X$ heißt **globales Maximum** von $f : \Leftrightarrow$

$$\forall x \in X : \quad f(x) \leq f(x_0). \quad (\text{VII.28})$$

Satz VII.7. Seien $X \subseteq \mathbb{R}$, $f : X \rightarrow \mathbb{R}$, x_0 ein I.P. von X und f bei $x_0 \in X$ differenzierbar. Dann gilt:

$$\left[x_0 \text{ ist ein lokales Maximum oder Minimum} \right] \Rightarrow \left[f'(x_0) = 0 \right]. \quad (\text{VII.29})$$

Beweis. Wir beweisen (VII.29) nur im Fall, dass x_0 ein lokales Minimum ist (die Behandlung des Falls eines lokalen Maximums ist analog). Es gibt also ein $\delta' > 0$ so, dass

$$\forall x \in (x_0 - \delta', x_0 + \delta') \cap X : \quad f(x) \geq f(x_0). \quad (\text{VII.30})$$

Weil $x_0 \in X$ ein I.P. ist, können wir durch Wahl eines genügend kleinen $0 < \delta \leq \delta'$ erreichen, dass $(x_0 - \delta, x_0 + \delta) \subseteq X$ und daher sogar

$$\forall x \in (x_0 - \delta, x_0 + \delta) : \quad f(x) \geq f(x_0) \quad (\text{VII.31})$$

gilt. Für $n \in \mathbb{N}$ seien nun $x_n^{(\pm)} := x_0 \pm \frac{\delta}{2n}$ definiert, sodass $(x_n^{(\pm)})_{n=1}^\infty \in [(x_0 - \delta, x_0 + \delta) \setminus \{x_0\}]^\mathbb{N}$ zwei Folgen in $(x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$ sind, die beide gegen x_0 konvergieren. Somit sind einerseits

$$f'(x_0) = \lim_{n \rightarrow \infty} \left\{ \frac{f(x_n^{(+)}) - f(x_0)}{x_n^{(+)} - x_0} \right\} = \lim_{n \rightarrow \infty} \left\{ \underbrace{\frac{2n}{\delta}}_{>0} \cdot \underbrace{[f(x_n^{(+)}) - f(x_0)]}_{\geq 0} \right\} \geq 0 \quad (\text{VII.32})$$

und andererseits

$$f'(x_0) = \lim_{n \rightarrow \infty} \left\{ \frac{f(x_n^{(-)}) - f(x_0)}{x_n^{(-)} - x_0} \right\} = \lim_{n \rightarrow \infty} \left\{ \underbrace{\frac{-2n}{\delta}}_{<0} \cdot \underbrace{[f(x_n^{(-)}) - f(x_0)]}_{\geq 0} \right\} \leq 0, \quad (\text{VII.33})$$

was zusammen $f'(x_0) = 0$ ergibt. $\square$

Eine Anwendung dieses Satzes ist der nun folgende Mittelwertsatz.

Satz VII.8 (Mittelwertsatz der Differentialrechnung). *Seien $a, b \in \mathbb{R}$, $a < b$, $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b) . Dann gibt es einen Punkt $x_0 \in (a, b)$, so dass*

$$\frac{f(b) - f(a)}{b - a} = f'(x_0). \quad (\text{VII.34})$$

Beweis. Wir setzen $g : [a, b] \rightarrow \mathbb{R}$,

$$\forall x \in [a, b] : \quad g(x) := \frac{f(b) - f(a)}{b - a}x - f(x). \quad (\text{VII.35})$$

Dann ist auch g stetig auf $[a, b]$ und differenzierbar auf (a, b) .

Wir unterscheiden nun zwei Fälle:

1. Ist $g(x) \equiv c$ konstant auf $[a, b]$, so gilt

$$\frac{f(b) - f(a)}{b - a} - f'(x) = g'(x) = 0, \quad (\text{VII.36})$$

sogar für alle $x \in (a, b)$. (In diesem Fall ist nämlich der Graph von f eine Gerade.)

2. Sei g nicht konstant auf $[a, b]$. Weil $[a, b]$ kompakt und g stetig auf $[a, b]$ sind, nimmt g in $[a, b]$ sein Minimum und Maximum an, d.h. es gibt $x_{\min}, x_{\max} \in [a, b]$, so dass

$$g(x_{\min}) = \inf \{g([a, b])\}, \quad g(x_{\max}) = \sup \{g([a, b])\}, \quad (\text{VII.37})$$

und weil g nicht konstant ist, gilt

$$g(x_{\min}) < g(x_{\max}). \quad (\text{VII.38})$$

Außerdem ist

$$\begin{aligned} g(a) &= \frac{f(b) - f(a)}{b - a} \cdot a - f(a) = \frac{f(b)a - bf(a)}{b - a} \\ &= \frac{f(b) - f(a)}{b - a} \cdot b - f(b) = g(b). \end{aligned} \quad (\text{VII.39})$$

Daher muss

$$\text{entweder:} \quad x_{\max} \in (a, b) \quad \text{und} \quad g(x_{\min}) = g(a) = g(b) < g(x_{\max}) \quad (\text{VII.40})$$

$$\text{oder:} \quad x_{\min} \in (a, b) \quad \text{und} \quad g(x_{\min}) < g(a) = g(b) = g(x_{\max}) \quad (\text{VII.41})$$

$$\text{oder (sogar):} \quad x_{\min}, x_{\max} \in (a, b) \quad \text{und} \quad g(x_{\min}) < g(a) = g(b) < g(x_{\max}) \quad (\text{VII.42})$$

gelten.

Gilt etwa (VII.41), so ist $x_{\min}$ ein innerer Punkt von (a, b) und ein lokales Minimum von g und deshalb nach Satz VII.7

$$0 = g'(x_{\min}) = \frac{f(b) - f(a)}{b - a} - f'(x_{\min}), \quad (\text{VII.43})$$

und $x_{\min}$ ist der gesuchte Punkt in (VII.34). Analog beweist man (VII.34) für (VII.40) und (VII.42). $\square$

Definition VII.9. Seien $U \subseteq \mathbb{R}$ und $f : U \rightarrow \mathbb{R}$.

$$f \text{ hei\ss}t \left\{ \begin{array}{l} \text{monoton steigend} \\ \text{streng monoton steigend} \\ \text{monoton fallend} \\ \text{streng monoton fallend} \end{array} \right\} \text{ in } U$$

$$:\Leftrightarrow \forall x, x' \in U, x < x' : \left\{ \begin{array}{l} f(x) \leq f(x') \\ f(x) < f(x') \\ f(x) \geq f(x') \\ f(x) > f(x') \end{array} \right. . \quad (\text{VII.44})$$

Satz VII.10. Seien $a, b \in \mathbb{R}, a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar in (a, b) . Dann gelten

$$\left(\forall x \in (a, b) : f'(x) \left\{ \begin{array}{l} \geq \\ > \\ \leq \\ < \\ = \end{array} \right\} 0 \right) \Rightarrow \left(f \text{ ist } \left\{ \begin{array}{l} \text{monoton steigend} \\ \text{streng monoton steigend} \\ \text{monoton fallend} \\ \text{streng monoton fallend} \\ \text{konstant} \end{array} \right\} \text{ in } [a, b] \right). \quad (\text{VII.45})$$

Beweis. Seien $x, x' \in [a, b]$ und $x < x'$. Dann ist die Einschränkung von f auf $[x, x']$ auch stetig auf $[x, x']$ und differenzierbar auf (x, x') . Nach dem Mittelwertsatz VII.5 gibt es also einen Punkt $x_0 \in (x, x')$, so dass

$$f(x') - f(x) = (x' - x) \cdot f'(x_0). \quad (\text{VII.46})$$

Somit ist wegen $x' - x > 0$

$$\left(f'(x_0) \left\{ \begin{array}{l} > \\ < \\ = \end{array} \right\} 0 \right) \Rightarrow \left(f(x) \left\{ \begin{array}{l} < \\ > \\ = \end{array} \right\} f(x') \right). \quad (\text{VII.47})$$

$\square$

Bemerkungen und Beispiele.

- Die Umkehrung “ $\Leftarrow$ ” gilt in (VII.45) im Allgemeinen nicht. So ist beispielsweise $f(x) := x^3$ auf ganz $\mathbb{R}$ streng monoton steigend, aber $f'(0) = 0$.
- Es gilt aber eine etwas schwächere Aussage als Umkehrung von (VII.45), nämlich

$$\left(f \text{ ist } \begin{cases} \text{monoton steigend} \\ \text{monoton fallend} \\ \text{konstant} \end{cases} \text{ in } [a, b] \right) \Rightarrow \left(\forall x \in (a, b) : f'(x) \begin{cases} \geq \\ \leq \\ = \end{cases} 0 \right). \quad (\text{VII.48})$$

VII.3. Höhere Ableitungen und der Satz von Taylor

Definition VII.11. Seien $U \subseteq \mathbb{R}$ offen, $f : U \rightarrow \mathbb{R}$ eine Abbildung und $n \in \mathbb{N}$.

(i) f ist **bei $x_0 \in U$ n -mal differenzierbar**.

$:\Leftrightarrow$ Es gibt ein $\delta > 0$, so dass für alle $x \in (x_0 - \delta, x_0 + \delta)$:

$$\begin{aligned} f &\text{ ist differenzierbar bei } x \text{ und } f'(x) := \frac{df}{dx}(x), \\ f'(x) &\text{ ist differenzierbar bei } x \text{ und } f''(x) := \frac{d}{dx}[f'](x), \\ f''(x) &\text{ ist differenzierbar bei } x \text{ und } f'''(x) := \frac{d}{dx}[f''](x), \\ &\vdots \\ f^{(n-2)}(x) &\text{ ist differenzierbar bei } x \text{ und } f^{(n-1)}(x) := \frac{d}{dx}[f^{(n-2)}](x), \\ f^{(n-1)}(x) &\text{ ist differenzierbar bei } x_0 \text{ und } f^{(n)}(x_0) := \frac{d}{dx}[f^{(n-1)}](x_0). \end{aligned} \quad (\text{VII.49})$$

Wir schreiben die k . Ableitung auch als

$$\frac{d^k f}{dx^k}(x) := f^{(k)}(x). \quad (\text{VII.50})$$

(ii) f ist **auf U n -mal differenzierbar**.

$:\Leftrightarrow$

$$\forall x_0 \in U : f \text{ ist bei } x_0 \text{ } n\text{-mal differenzierbar.} \quad (\text{VII.51})$$

Satz VII.12 (Taylor). Seien $a, b, x_0 \in \mathbb{R}$ mit $a < x_0 < b$, $n \in \mathbb{N}$, und $f : (a, b) \rightarrow \mathbb{R}$ auf (a, b) $(n+1)$ -mal differenzierbar. Dann gibt es zu jedem $x \in (a, b) \setminus \{x_0\}$ ein $x' \in (x, x_0)$ [falls $x < x_0$] bzw. $x' \in (x_0, x)$ [falls $x > x_0$], sodass

$$f(x) = T_n[f, x_0; x] + R_{n+1}[f, x_0; x], \quad (\text{VII.52})$$

wobei

$$\begin{aligned} T_n[f, x_0; x] &:= \sum_{k=0}^n \left\{ \frac{(x-x_0)^k}{k!} f^{(k)}(x_0) \right\} \\ &= f(x_0) + (x-x_0) \cdot f'(x_0) + \frac{(x-x_0)^2}{2!} f''(x_0) + \dots + \frac{(x-x_0)^n}{n!} f^{(n)}(x_0) \end{aligned} \quad (\text{VII.53})$$

das **Taylorpolynom n. Ordnung** und

$$R_{n+1}[f, x_0; x] := \frac{(x - x_0)^{n+1}}{(n+1)!} f^{(n+1)}(x') \quad (\text{VII.54})$$

das **Taylor'sches Restglied (n + 1). Ordnung** sind.

Beweis. Wir können $x > x_0$ annehmen, der Beweis für $x < x_0$ ist analog. Für alle $t \in [x_0, x]$ setzen wir

$$p(t) := \sum_{k=0}^n \left\{ \frac{(t - x_0)^k}{k!} f^{(k)}(x_0) \right\}, \quad (\text{VII.55})$$

$$\lambda := \frac{(n+1)!}{(x - x_0)^{n+1}} (f(x) - p(x)) \quad (\text{VII.56})$$

und

$$g(t) := f(t) - p(t) - \frac{\lambda}{(n+1)!} (t - x_0)^{n+1}. \quad (\text{VII.57})$$

Beachte, dass für $\ell = 0, 1, 2, \dots, n$,

$$\begin{aligned} \frac{d^\ell p(t)}{dt^\ell} &= \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} \frac{d^\ell [(t - x_0)^k]}{dt^\ell} = \sum_{k=1}^n \frac{f^{(k)}(x_0)}{k!} \cdot k \cdot \frac{d^{\ell-1} [(t - x_0)^{k-1}]}{dt^{\ell-1}} \\ &= \dots = \sum_{k=\ell}^n \frac{f^{(k)}(x_0)}{(k - \ell)!} (t - x_0)^{k-\ell}. \end{aligned} \quad (\text{VII.58})$$

Somit ist

$$g^{(\ell)}(t) = f^{(\ell)}(t) - \sum_{m=0}^{n-\ell} \frac{f^{(m+\ell)}(x_0)}{m!} (t - x_0)^m - \frac{\lambda (t - x_0)^{n+1-\ell}}{(n+1-\ell)!}. \quad (\text{VII.59})$$

Insbesondere sind

$$g(x_0) = f(x_0) - \frac{f(x_0)}{0!} + 0 - 0 = 0, \quad (\text{VII.60})$$

$$g^{(1)}(x_0) = f'(x_0) - \frac{f'(x_0)}{1!} + 0 - 0 = 0, \quad (\text{VII.61})$$

$\vdots$

$$g^{(n)}(x_0) = f^{(n)}(x_0) - f^{(n)}(x_0) = 0 \quad (\text{VII.62})$$

und

$$g^{(n+1)}(t) = f^{(n+1)}(t) - \lambda. \quad (\text{VII.63})$$

Nach Wahl von λ in (VII.56) folgt außerdem $g(x) = 0$, also

$$g(x_0) = g(x) = 0. \quad (\text{VII.64})$$

Nach dem Mittelwertsatz gibt es ein $x_1 \in (x_0, x)$, so dass $g'(x_1) = 0$. Also gilt

$$g'(x_0) = g'(x_1) = 0, \quad (\text{VII.65})$$

und der Mittelwertsatz impliziert abermals die Existenz eines $x_2 \in (x_0, x_1)$ mit $g''(x_2) = 0$, also

$$g''(x_0) = g''(x_2) = 0, \quad (\text{VII.66})$$

und damit wieder $g'''(x_3) = 0$ für ein gewisses $x_3 \in (x_0, x_2)$ u.s.w. Schließlich erhalten wir ein $x_{n+1} \in (x_0, x_n) \subseteq (x_0, x_{n-1}) \subseteq \dots \subseteq (x_0, x)$ mit

$$0 = g^{(n+1)}(x_{n+1}) = f^{(n+1)}(x_{n+1}) - \lambda, \quad (\text{VII.67})$$

also

$$f(x) = p(x) + \frac{(x - x_0)^{n+1}}{(n+1)!} \cdot f^{(n+1)}(x_{n+1}), \quad (\text{VII.68})$$

und $x_{n+1} \in (x_0, x)$ ist der gesuchte Punkt x' . $\square$

Bemerkungen und Beispiele.

- Für $n \in \mathbb{N}_0$ und $x \in \mathbb{R}$ gilt mit $f(x) := e^x$

$$f^{(n+1)}(x) = f^{(n)}(x) = \dots = f'(x) = f(x) = e^x, \quad (\text{VII.69})$$

und insbesondere

$$f(0) = f'(0) = f''(0) = f^{(n)}(0) = 1. \quad (\text{VII.70})$$

Also ist

$$\forall x \in \mathbb{R}: \quad \left| e^x - \left\{ 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} \right\} \right| \leq \frac{x^{n+1}}{(n+1)!} \cdot e^{|x|}, \quad (\text{VII.71})$$

denn für $x' \in (-x, x)$ ist $e^{x'} \leq e^{|x|}$.

- Für $x = 1$ und $n = 10$ erhält man

$$\left| e - \left\{ 2 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{10!} \right\} \right| = |e - (2,718281801)| \leq \frac{e}{11!} \leq e \cdot 0,000002505, \quad (\text{VII.72})$$

also

$$2,71827 \leq \frac{2,718281801}{1,000002505} \leq e \leq \frac{2,718281801}{0,999997495} \leq 2,71829. \quad (\text{VII.73})$$

- Sei folgende Aufgabe gestellt: *Berechne $\sin(\frac{1}{10})$ mit einer Genauigkeit von mindestens 10^{-6} .*

Dazu wenden wir Satz VII.12 von Taylor auf $f(x) = \sin(x)$ mit $a = -\infty$, $b = \infty$ und $x_0 := 0$ an und berechnen die Ableitungen von f ,

$$f^{(0)}(x) = \sin(x), \quad f^{(1)}(x) = \cos(x), \quad f^{(2)}(x) = -\sin(x), \quad f^{(3)}(x) = -\cos(x), \quad (\text{VII.74})$$

u.s.w. Insbesondere sind dann

$$f^{(0)}(0) = 0, \quad f^{(1)}(0) = 1, \quad f^{(2)}(0) = 0, \quad f^{(3)}(0) = -1, \quad \dots \quad (\text{VII.75})$$

und wir erhalten

$$T_6[\sin, 0; x] = 0 + \frac{1}{1!}x + 0 + \frac{-1}{3!}x^3 + 0 + \frac{1}{5!}x^5 + 0 = x - \frac{x^3}{6} + \frac{x^5}{120} \quad (\text{VII.76})$$

sowie

$$R_7[\sin, 0; x] = \frac{-\cos(x')}{7!}x^7. \quad (\text{VII.77})$$

Wegen $|\cos(x')| \leq 1$ ist nach dem Satz VII.12 von Taylor dann für alle $x \in \mathbb{R}$, $x \neq 0$,

$$\left| \sin(x) - \left(x - \frac{x^3}{6} + \frac{x^5}{120} \right) \right| = |R_7[\sin, 0; x]| \leq \frac{|x|^7}{7!} = \frac{|x|^7}{5040}. \quad (\text{VII.78})$$

Wählen wir $x = \frac{1}{10}$, so erhalten wir

$$\begin{aligned} & \left| \sin\left(\frac{1}{10}\right) - (0,1-0,0001\bar{6} + 0,00000008\bar{3}) \right| \\ &= \left| \sin\left(\frac{1}{10}\right) - (0,09983341\bar{6}) \right| \leq \frac{10^{-7}}{5000} = 2 \cdot 10^{-11}, \end{aligned} \quad (\text{VII.79})$$

d.h. mit nur drei Termen approximieren wir die Sinusfunktion von $\frac{1}{10}$ sogar mit einer Genauigkeit von $2 \cdot 10^{-11}$!

- Setzen wir in (VII.53) $n = \infty$, so erhalten wir $f(x)$ *formal* aus einer Potenzreihe um x_0 , der Taylorreihe,

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k. \quad (\text{VII.80})$$

Für viele Funktionen kann man die Konvergenz dieser Reihe durch Abschätzung der Taylorschen Restglieder in beliebig hoher Ordnung gewinnen. Aus (VII.71) sieht man, dass dies z.B. für $f(x) = e^x$ gelingt, und in der Tat ist die e^x definierende Potenzreihe V.6 gerade die Taylorreihe um $x_0 = 0$.

- Es gibt aber auch viele Funktionen, für die (VII.80) falsch ist. Betrachten wir z.B.

$$f(x) := \begin{cases} e^{-1/x^2} & \text{für } x \neq 0, \\ 0 & \text{für } x = 0, \end{cases} \quad (\text{VII.81})$$

so ist $f(x) > 0$, für $x \neq 0$, aber

$$\forall n \in \mathbb{N}_0 : f^{(n)}(0) = 0, \quad (\text{VII.82})$$

d.h. die Taylorreihe von f um $x_0 = 0$,

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \equiv 0, \quad (\text{VII.83})$$

verschwindet identisch. Damit konvergiert die Taylorreihe von f zwar, hat jedoch mit der Funktion f nichts zu tun!

- Eine Anwendung der Taylorschen Formel (VII.52) ist die Kurvendiskussion.

Ist $f : U \rightarrow \mathbb{R}$ eine differenzierbare Funktion, so verschwindet zwar bei lokalem Maximum oder Minimum die Ableitung f' , es ist aber noch nicht klar, wie man entscheiden kann, ob es sich auch wirklich um ein Maximum oder Minimum handelt. Dies lässt sich erst bei Betrachtung der höheren Ableitungen sagen; Darüber gibt Satz VII.13 Auskunft.

Satz VII.13. *Seien $U \subseteq \mathbb{R}$ offen, $n \in \mathbb{N}$, $n \geq 2$, und $f : U \rightarrow \mathbb{R}$ eine auf U n -mal differenzierbare Abbildung. Sei weiterhin $x_0 \in U$, und gelte*

$$f'(x_0) = f''(x_0) = \dots = f^{(n-1)}(x_0) = 0, \quad f^{(n)}(x_0) \neq 0, \quad (\text{VII.84})$$

und sei $f^{(n)}$ stetig in x_0 . Dann liegt folgende Situation lokal bei x_0 vor:

- Ist n gerade und $f^{(n)}(x_0) > 0$, so ist x_0 ein lokales Minimum;
- Ist n gerade und $f^{(n)}(x_0) < 0$, so ist x_0 ein lokales Maximum;
- Ist n ungerade, so ist x_0 ein Wendepunkt von f .

Beweis. Wir zeigen nur, dass für gerades $n = 2k \geq 2$, also $k \in \mathbb{N}$ und $f^{(2k)}(x_0) > 0$ ein lokales Minimum vorliegt. Seien $x \in (x_0 - \delta, x_0 + \delta)$ und $0 < \delta \ll 1$. Nach dem Satz von Taylor gibt es ein x' mit $|x' - x_0| \leq |x - x_0|$, so dass

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + \dots + \frac{(x - x_0)^{2k-1}}{(2k-1)!} f^{(2k-1)}(x_0) + \frac{(x - x_0)^{2k}}{(2k)!} f^{(2k)}(x') \\ &= f(x_0) + \frac{(x - x_0)^{2k}}{(2k)!} f^{(2k)}(x'). \end{aligned} \quad (\text{VII.85})$$

Da $f^{(2k)}$ in x_0 stetig ist, gibt es ein $\delta' > 0$, so dass

$$\forall |x' - x_0| < \delta' : |f^{(2k)}(x') - f^{(2k)}(x_0)| \leq \frac{1}{2} f^{(2k)}(x_0), \quad (\text{VII.86})$$

für $|x' - x_0| < \delta'$ also

$$f^{(2k)}(x') \geq f^{(2k)}(x_0) - |f^{(2k)}(x') - f^{(2k)}(x_0)| \geq \frac{1}{2} f^{(2k)}(x_0) > 0. \quad (\text{VII.87})$$

Somit gilt für alle $x \in \mathbb{R}$ mit $|x' - x_0| < \min\{\delta, \delta'\}$ auch

$$f(x) \geq f(x_0) + \frac{(x - x_0)^{2k}}{(2k)!} \frac{f^{(2k)}(x_0)}{2} \geq f(x_0). \quad (\text{VII.88})$$

□

VII.4. Ergänzungen

VII.4.1. Beweis der Kettenregel

Beweis. (Beweis von Satz VII.5) Sei $\alpha > 0$. Dann gibt es gemäß Lemma VII.2 ein $\beta > 0$, so dass, für alle $y \in V$ mit $|y - g(x_0)| \leq \beta$ auch

$$\left| f(y) - \{f(g(x_0)) + f'(g(x_0))(y - g(x_0))\} \right| \leq \alpha |y - g(x_0)| \quad (\text{VII.89})$$

gilt. Wir setzen $\gamma := \min\{\beta, \alpha, 1\} > 0$. Weiterhin gibt es zu $\gamma > 0$ ein $\eta > 0$, so dass, für alle $x \in U$ mit $|x - x_0| \leq \eta$ auch

$$\left| g(x) - \{g(x_0) + g'(x_0) \cdot (x - x_0)\} \right| \leq \gamma |x - x_0| \quad (\text{VII.90})$$

gilt. Wir setzen

$$\kappa := \min \left\{ \eta, \frac{\gamma}{|g'(x_0)| + \gamma} \right\} > 0. \quad (\text{VII.91})$$

Dann gilt für alle $|x - x_0| \leq \kappa$:

$$\left| g(x) - g(x_0) \right| \leq \left(|g'(x_0)| + \beta \right) |x - x_0| \leq \gamma, \quad (\text{VII.92})$$

also auch

$$\begin{aligned} & \left| f[g(x)] - \{f[g(x_0)] + f'[g(x_0)] \cdot g'(x_0) \cdot (x - x_0)\} \right| \\ & \leq \left| f[g(x)] - \{f[g(x_0)] + f'[g(x_0)] \cdot (g(x_0) - g(x_0))\} \right| \\ & \quad + \left| f'[g(x_0)] \right| \cdot \left| g(x) - \{g(x_0) + g'(x_0)(x - x_0)\} \right|, \end{aligned} \quad (\text{VII.93})$$

und somit

$$\begin{aligned} & \left| f[g(x)] - \{f[g(x_0)] + f'(g(x_0))g'(x_0)(x - x_0)\} \right| \\ & \leq \alpha |g(x) - g(x_0)| + \gamma |x - x_0| \leq \left[\alpha |g'(x_0)| + \alpha\gamma + \gamma \right] \cdot |x - x_0| \\ & \leq \alpha \left(|g'(x_0)| + 2 \right) \cdot |x - x_0|, \end{aligned} \quad (\text{VII.94})$$

da $\gamma \leq \min\{1, \alpha\}$.

Für vorgegebenes $\varepsilon > 0$ wählen wir $\alpha := \varepsilon \cdot \left(|g'(x_0)| + 2 \right)^{-1}$ und anschließend β, γ, η , und $\kappa =: \delta > 0$, wie in (VII.89)–(VII.91). Somit gilt für alle $|x - x_0| \leq \delta$:

$$\left| f[g(x)] - \{f[g(x_0)] + f'[(g(x_0))] \cdot g'(x_0) \cdot (x - x_0)\} \right| \leq \varepsilon \cdot |x - x_0|. \quad (\text{VII.95})$$

□

VII.4.2. Zwischenwertsatz für die Ableitung

Für eine differenzierbare Funktion f ist die Ableitung f' im Allgemeinen nicht stetig, es gilt aber die folgende Aussage:

Lemma VII.14. *Seien $a, b \in \mathbb{R}$, $a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b) . Seien $x, x' \in (a, b)$ mit $x < x'$, und $f'(x) < f'(x')$. Dann gibt es für jedes $f'(x) < \lambda < f'(x')$ ein $x < t < x'$, so dass $f'(t) = \lambda$.*

Beweis. Wir setzen

$$\forall x < t < x' : \quad g(t) := f(t) - \lambda \cdot t, \quad (\text{VII.96})$$

so dass,

$$g'(x) = f'(x) - \lambda < 0, \quad g'(x') = f'(x') - \lambda > 0. \quad (\text{VII.97})$$

Daher gibt es ein $\delta > 0$, so dass

$$\forall t \in (x, x + \delta) : \quad g(x) > g(t), \quad (\text{VII.98})$$

$$\forall t \in (x' - \delta, x') : \quad g(t) < g(x'). \quad (\text{VII.99})$$

Somit muss g in (x, x') ein lokales Minimum haben, d.h.

$$\exists t \in (x, x') : \quad 0 = g'(t) = f'(t) - \lambda. \quad (\text{VII.100})$$

□

VII.4.3. $0/0$ und ∞/∞ – Der Satz von L'Hospital

Satz VII.15 (L'Hospital). *Seien $U \subseteq \mathbb{R}$, $x_0 \in U$ ein innerer Punkt und $f, g : U \rightarrow \mathbb{R}$ differenzierbar bei $x_0 \in U$. Seien weiterhin*

$$f(x_0) = g(x_0) = 0, \quad \forall 0 < |x - x_0| \leq \delta : g(x) \neq 0 \quad (\text{VII.101})$$

für ein gewisses $\delta > 0$, und $g'(x_0) \neq 0$. Dann lässt sich $f(x)/g(x)$ stetig in x_0 fortsetzen mit

$$\lim_{x \rightarrow x_0} \left\{ \frac{f(x)}{g(x)} \right\} = \frac{f'(x_0)}{g'(x_0)}. \quad (\text{VII.102})$$

Beweis. Für jedes $\alpha > 0$ gibt es ein $\delta > 0$, so dass für alle $0 < |x - x_0| \leq \delta$

$$|f(x) - f(x_0) - f'(x_0)(x - x_0)| \leq \alpha |x - x_0|, \quad (\text{VII.103})$$

$$|g(x) - g(x_0) - g'(x_0)(x - x_0)| \leq \alpha |x - x_0|. \quad (\text{VII.104})$$

Ist $0 < \alpha \leq \frac{1}{2}|g'(x_0)|$, so gilt auch insbesondere

$$|g(x)| \geq (|g'(x_0)| - \alpha)|x - x_0| \geq \frac{g'(x_0)}{2} |x - x_0| > 0, \quad (\text{VII.105})$$

für alle $0 < |x - x_0| \leq \delta$. Damit erhalten wir

$$\left| \frac{f(x)}{g(x)} - \frac{f'(x_0)}{g'(x_0)} \right| = \left| \frac{(f'(x_0) + \alpha_f(x))(x - x_0)}{(g'(x_0) + \alpha_g(x))(x - x_0)} - \frac{f'(x_0)}{g'(x_0)} \right|, \quad (\text{VII.106})$$

wobei $|\alpha_f(x)|, |\alpha_g(x)| \leq \alpha$ sind. Also ist, für alle $0 < |x - x_0| \leq \delta$,

$$\begin{aligned} \left| \frac{f(x)}{g(x)} - \frac{f'(x_0)}{g'(x_0)} \right| &= \left| \frac{f'(x_0) + \alpha_f(x)}{g'(x_0) + \alpha_g(x)} - \frac{f'(x_0)}{g'(x_0)} \right| \\ &= \frac{1}{|g'(x_0)| \cdot |g'(x_0) + \alpha_g(x)|} \cdot |\alpha_f(x) \cdot g'(x_0) - \alpha_g(x) \cdot f'(x_0)| \\ &\leq \frac{2\alpha}{|g'(x_0)|^2} (|g'(x_0)| + |f'(x_0)|). \end{aligned} \quad (\text{VII.107})$$

Ist nun $\varepsilon < 0$ vorgegeben, so wählen wir

$$\alpha := \frac{|g'(x_0)|^2}{2(|g'(x_0)| + |f'(x_0)|)} \cdot \min\{\varepsilon, 1\} > 0, \quad (\text{VII.108})$$

und beobachten, dass $\alpha \leq |g'(x_0)|/2$. Also ist,

$$\forall 0 < |x - x_0| \leq \delta : \quad \left| \frac{f(x)}{g(x)} - \frac{f'(x_0)}{g'(x_0)} \right| \leq \varepsilon \quad (\text{VII.109})$$

und somit

$$\lim_{x \rightarrow x_0} \left\{ \frac{f(x)}{g(x)} \right\} = \frac{f'(x_0)}{g'(x_0)}. \quad (\text{VII.110})$$

□

VIII. Integration

In diesem Kapitel führen wir das *Riemann-Integral* für reelle Funktionen einer reellen Veränderlichen ein. Das Lebesgue-Integral wird in einer späteren Vorlesung behandelt.

VIII.1. Partitionen, Ober- und Untersummen

Definition VIII.1. Seien $a, b \in \mathbb{R}, a < b$.

- (i) Eine endliche Teilmenge $P = \{x_1, x_2, \dots, x_L\} \subseteq (a, b)$ von (a, b) heißt **Partition von $[a, b]$** . Wir wollen stets annehmen, dass die Punkte in P aufsteigend geordnet sind, also

$$a =: x_0 < x_1 < x_2 < \dots < x_L < x_{L+1} := b. \quad (\text{VIII.1})$$

Die Menge der Partitionen von $[a, b]$ bezeichnen wir mit $\mathcal{P}[a, b]$.

- (ii) Seien $P, P' \in \mathcal{P}[a, b]$ zwei Partitionen von $[a, b]$.

$$P \text{ heißt } \mathbf{feiner} \text{ als } P' \quad :\Leftrightarrow \quad P' \subseteq P. \quad (\text{VIII.2})$$

Wir sagen auch, dass P eine **Verfeinerung von P'** ist.

Bemerkungen und Beispiele.

- Wir bemerken, dass zwei Partitionen nicht immer vergleichbar bezüglich ihrer Feinheit sind. Sind nämlich $P \setminus P'$ und $P' \setminus P$ beide nicht leer, so ist weder P feiner als P' noch P' feiner als P . Es lässt sich aber immer eine Partition P'' finden, die sowohl feiner als P als auch feiner als P' ist, nämlich

$$P'' = P \cup P'. \quad (\text{VIII.3})$$

Eine Ordnungsrelation mit dieser Eigenschaft bezeichnet man als *partielle Ordnung*.

Definition VIII.2. Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion, also $f([a, b]) \subseteq [-R, R]$, für genügend großes $R < \infty$. Sei weiterhin $P = \{x_1, x_2, \dots, x_L\} \subseteq (a, b)$ eine Partition.

- (i) Wir definieren die **Untersumme** $\underline{\mathcal{I}}(f; P)$ und die **Obersumme** $\overline{\mathcal{I}}(f; P)$ durch

$$\underline{\mathcal{I}}(f; P) := \sum_{j=0}^L (x_{j+1} - x_j) \cdot \inf \{f(x) \mid x_j \leq x \leq x_{j+1}\}, \quad (\text{VIII.4})$$

$$\overline{\mathcal{I}}(f; P) := \sum_{j=0}^L (x_{j+1} - x_j) \cdot \sup \{f(x) \mid x_j \leq x \leq x_{j+1}\}. \quad (\text{VIII.5})$$

- (ii) Wir definieren das **Unterintegral** $\int_a^b f(x) dx$ und das **Oberintegral** $\int_a^b \overline{f(x)} dx$ von f durch

$$\int_a^b f(x) dx := \sup \{ \underline{\mathcal{I}}(f; P) \mid P \in \mathcal{P}[a, b] \}, \quad (\text{VIII.6})$$

$$\int_a^b \overline{f(x)} dx := \inf \{ \overline{\mathcal{I}}(f; P) \mid P \in \mathcal{P}[a, b] \}. \quad (\text{VIII.7})$$

Lemma VIII.3. Seien $a, b \in \mathbb{R}$, $a < b$, $f : [a, b] \rightarrow \mathbb{R}$ beschränkt und $P, P' \in \mathcal{P}[a, b]$ zwei Partitionen von $[a, b]$. Ist P' feiner als P , so gilt

$$\underline{\mathcal{I}}(f; P) \leq \underline{\mathcal{I}}(f; P'), \quad \text{und} \quad \overline{\mathcal{I}}(f; P) \geq \overline{\mathcal{I}}(f; P'). \quad (\text{VIII.8})$$

Beweis. Wir beweisen nur $\underline{\mathcal{I}}(f; P) \leq \underline{\mathcal{I}}(f; P')$. Sind $P = \{x_1, x_2, \dots, x_L\}$, $a =: x_0 < x_1 < \dots < x_L < x_{L+1} := b$, und $y \in (a, b) \setminus P$, dann gibt es ein $k \in \{0, 1, 2, \dots, L\}$, so dass $x_k < y < x_{k+1}$, und es gilt

$$\begin{aligned} & (x_{k+1} - x_k) \cdot \inf \{ f(x) \mid x_k \leq x \leq x_{k+1} \} \\ &= (x_{k+1} - y) \cdot \inf \{ f(x) \mid x_k \leq x \leq x_{k+1} \} + (y - x_k) \cdot \inf \{ f(x) \mid x_k \leq x \leq x_{k+1} \} \\ &\leq (x_{k+1} - y) \cdot \inf \{ f(x) \mid y \leq x \leq x_{k+1} \} + (y - x_k) \cdot \inf \{ f(x) \mid x_k \leq x \leq y \}. \end{aligned} \quad (\text{VIII.9})$$

Daher ist

$$\begin{aligned} & \underline{\mathcal{I}}(f, P \cup \{y\}) - \underline{\mathcal{I}}(f, P) \\ &= (y - x_k) \cdot \inf \{ f(x) \mid x_k \leq x \leq y \} + (x_k - y) \cdot \inf \{ f(x) \mid y \leq x \leq x_{k+1} \} \\ &\quad - (x_{k+1} - x_k) \cdot \inf \{ f(x) \mid x_k \leq x \leq x_{k+1} \} \\ &\geq 0. \end{aligned} \quad (\text{VIII.10})$$

Für $P' \setminus P = \{y_1, y_2, \dots, y_M\}$ erhalten wir also in M Schritten

$$\begin{aligned} \underline{\mathcal{I}}(f; P) &\leq \underline{\mathcal{I}}(f; P \cup \{y_1\}) \leq \underline{\mathcal{I}}(f; P \cup \{y_1, y_2\}) \\ &\leq \dots \leq \underline{\mathcal{I}}(f; P \cup \{y_1, y_2, \dots, y_M\}) = \underline{\mathcal{I}}(f, P'). \end{aligned} \quad (\text{VIII.11})$$

□

Lemma VIII.4. Seien $a, b \in \mathbb{R}$, $a < b$, und $f : [a, b] \rightarrow \mathbb{R}$ beschränkt. Dann ist

$$\int_a^b f(x) dx \leq \int_a^b \overline{f(x)} dx. \quad (\text{VIII.12})$$

Beweis. Nach Lemma VIII.3 gilt für je zwei Partitionen $P, P' \in \mathcal{P}[a, b]$ von $[a, b]$, dass

$$\underline{\mathcal{I}}(f; P) \leq \underline{\mathcal{I}}(f; P \cup P') \leq \overline{\mathcal{I}}(f; P \cup P') \leq \overline{\mathcal{I}}(f; P'). \quad (\text{VIII.13})$$

Da P' eine beliebige Partition ist, folgt daraus auch

$$\forall P : \quad \underline{\mathcal{I}}(f; P) \leq \inf \{ \overline{\mathcal{I}}(f; P') \mid P' \in \mathcal{P}[a, b] \}, \quad (\text{VIII.14})$$

also

$$\sup_{P \in \mathcal{P}[a, b]} \{ \underline{\mathcal{I}}(f, P) \} \leq \inf_{P' \in \mathcal{P}[a, b]} \{ \overline{\mathcal{I}}(f, P') \}. \quad (\text{VIII.15})$$

□

Definition VIII.5. Seien $a, b \in \mathbb{R}$ mit $a < b$. Eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt **Riemann-integrierbar auf $[a, b]$** : $\Leftrightarrow$

$$\int_a^b f(x) dx := \int_a^b f(x) dx = \overline{\int_a^b f(x) dx}. \quad (\text{VIII.16})$$

In diesem Fall schreiben wir $f \in \mathcal{R}[a, b]$ und $\int_a^b f(x) dx$ heißt **(Riemann-)Integral von f von a nach b** .

Lemma VIII.6. Seien $a, b \in \mathbb{R}, a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ beschränkt. Dann gilt folgende Äquivalenz:

$$\left[f \in \mathcal{R}[a, b] \right] \Leftrightarrow \left[\forall \varepsilon > 0 \exists P \in \mathcal{P}[a, b] : \overline{\mathcal{I}}(f; P) - \underline{\mathcal{I}}(f; P) \leq \varepsilon \right]. \quad (\text{VIII.17})$$

Satz VIII.7. Seien $a, b, \alpha \in \mathbb{R}, a < b$ und $f, g \in \mathcal{R}[a, b]$.

(i) Dann ist $(f + \alpha g) \in \mathcal{R}[a, b]$, und es gilt

$$\int_a^b (f + \alpha g)(x) dx = \int_a^b f(x) dx + \alpha \int_a^b g(x) dx. \quad (\text{VIII.18})$$

(ii) Ist $f(x) \leq g(x)$, für alle $x \in [a, b]$, so gilt

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx. \quad (\text{VIII.19})$$

(iii) Mit $f \in \mathcal{R}[a, b]$ ist auch $|f| \in \mathcal{R}[a, b]$, und es gilt

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx. \quad (\text{VIII.20})$$

VIII.2. Riemann-integrierbare Funktionen

In diesem Abschnitt wollen wir eine (möglichst große) Familie Riemann-integrierbarer Funktionen identifizieren. Dafür stellen wir zunächst fest, dass stetige Funktionen auf $[a, b]$ Riemann-integrierbar sind.

Satz VIII.8. Seien $a, b \in \mathbb{R}, a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$. Dann ist $f \in \mathcal{R}[a, b]$.

Beweis. Den Beweis findet man bei den Ergänzungen im Anhang VIII.5.3. $\square$

Weiterhin beobachten wir, dass ein “Ausreißer” als Funktionswert das Integral nicht verändert.

Satz VIII.9. Seien $a, b, c \in \mathbb{R}, a < c < b$ und $f \in \mathcal{R}[a, b]$. Dann ist $f \in \mathcal{R}[a, c] \cap \mathcal{R}[c, b]$, und es gilt

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx. \quad (\text{VIII.21})$$

Sind umgekehrt $f_{ac} \in \mathcal{R}[a, c]$ und $f_{cb} \in \mathcal{R}[c, b]$, dann ist mit

$$f(x) := \begin{cases} f_{ac}(x) & \text{für } x \in [a, c), \\ \text{beliebig, } \in \mathbb{R} & \text{für } x = c, \\ f_{cb}(x) & \text{für } x \in (c, b] \end{cases} \quad (\text{VIII.22})$$

auch $f \in \mathcal{R}[a, b]$ und es gilt

$$\int_a^b f(x) dx = \int_a^c f_{ac}(x) dx + \int_c^b f_{cb}(x) dx. \quad (\text{VIII.23})$$

Aus den Sätzen VIII.8 und VIII.9 ergibt sich durch eine einfache Induktion sofort folgendes Korollar.

Korollar VIII.10. Seien $a_0, a_1, \dots, a_M \in \mathbb{R}, a_j < a_{j+1}$, und $f : [a_0, a_M] \rightarrow \mathbb{R}$ stückweise stetig, d.h. f ist stetig auf (a_j, a_{j+1}) und die Grenzwerte $\lim_{x \searrow a_j} f(x)$ und $\lim_{x \nearrow a_j} f(x)$ existieren für $j = 0, 1, 2, \dots, M-1$. Dann ist $f \in \mathcal{R}[a_0, a_M]$.

Bemerkungen und Beispiele.

- Sei $(a_n)_{n=1}^\infty$ eine Abzählung von $\mathbb{Q} \cap (0, 1)$.
- Seien $M \in \mathbb{N}$ und $f_M : [0, 1] \rightarrow \{0, 1\}$,

$$f_M(x) := \begin{cases} 1, & \text{falls } x \in \{0, a_1, a_2, \dots, a_M, 1\}, \\ 0 & \text{sonst.} \end{cases} \quad (\text{VIII.24})$$

Mit einer geeigneten Umordnung ist dann $0 < a_{i_1} < a_{i_2} < \dots < a_{i_M} < 1$, und f_M ist stückweise stetig. Nach Korollar VIII.10 ist also $f_M \in \mathcal{R}[0, 1]$ und

$$\int_0^1 f_M(x) dx = 0. \quad (\text{VIII.25})$$

- Die **Dirichlet-Funktion** $f_\infty : [0, 1] \rightarrow \mathbb{R}$ ist definiert durch

$$f_\infty(x) := \begin{cases} 1, & \text{falls } x \in \mathbb{Q}, \\ 0, & \text{falls } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases} \quad (\text{VIII.26})$$

Seien $P = \{x_1, x_2, \dots, x_L\} \subseteq (0, 1)$, $x_0 := 0 < x_1 < x_2 < \dots < x_L < x_{L+1} := 1$, eine Partition von $[0, 1]$. Zu jedem $k \in \{0, 1, \dots, L\}$ gibt es zwei Zahlen,

$$\alpha \in \mathbb{Q} \cap (x_k, x_{k+1}), \quad \beta \in (\mathbb{R} \setminus \mathbb{Q}) \cap (x_k, x_{k+1}). \quad (\text{VIII.27})$$

Dies hat zur Folge, dass

$$0 \leq \inf_{x \in I_k} \{f_\infty(x)\} \leq f_\infty(\beta) = 0, \quad (\text{VIII.28})$$

$$1 \geq \sup_{x \in I_k} \{f_\infty(x)\} \geq f_\infty(\alpha) = 1. \quad (\text{VIII.29})$$

Somit ist, für alle $P \in \mathcal{P}[0, 1]$,

$$\overline{\mathcal{I}}(f_\infty; P) = \underbrace{\sum_{k=0}^L (x_{k+1} - x_k)}_{=1} \cdot \underbrace{\sup_{x \in I_k} \{f_\infty(x)\}}_{=1} = 1, \quad (\text{VIII.30})$$

$$\underline{\mathcal{I}}(f_\infty; P) = \underbrace{\sum_{k=0}^L (x_{k+1} - x_k)}_{=1} \cdot \underbrace{\inf_{x \in I_k} \{f_\infty(x)\}}_{=0} = 0, \quad (\text{VIII.31})$$

also

$$\int_0^1 f_\infty(x) dx = 0 \neq 1 = \overline{\int_0^1 f_\infty(x) dx}. \quad (\text{VIII.32})$$

Die Dirichlet-Funktion ist also *nicht* Riemann-integrierbar.

- In der Dirichlet-Funktion manifestiert sich die Schwäche des Riemann-Integralbegriffs. Nicht, dass der Wert des Integrals dieser pathologischen Funktion irgend eine Bedeutung hätte. Vielmehr ist f_∞ als Limes von f_M anzusehen:

Für festes $x \in [0, 1]$ betrachten wir die Folge $(f_M(x))_{M=1}^\infty$ in $\{0, 1\}$.

- Ist $x \in \mathbb{R} \setminus \mathbb{Q}$, dann ist $f_M(x) = 0 = f_\infty(x)$, $\forall M \in \mathbb{N}$.
- Ist umgekehrt $x \in \mathbb{Q}$, dann gibt es ein $K(x) \in \mathbb{N}$, so dass $x = a_{K(x)}$. Für alle $M \geq K(x)$ ist dann $f_M(x) = 1 = f_\infty(x)$.

Zusammenfassend erhalten wir, dass $(f_M(x))_{M=1}^\infty$ eine monoton steigende Funktionenfolge ist, $f_M(x) \leq f_{M+1}(x)$, und dass f_M *punktweise* gegen f_∞ konvergiert, d.h.

$$\forall x \in [0, 1] : \lim_{M \rightarrow \infty} \{f_M(x)\} = f_\infty(x). \quad (\text{VIII.33})$$

Die Folge $(f_M)_{M=1}^\infty$ und ihr Limes f_∞ bilden also ein *Gegenbeispiel* für die Vermutung, dass

$$\left[f_M \xrightarrow{pktw.} f_\infty \right] \Rightarrow \left[\lim_{M \rightarrow \infty} \int_0^1 f_M(x) dx = \int_0^1 f_\infty(x) dx \right], \quad (\text{VIII.34})$$

ja sogar, dass

$$\left[f_M \xrightarrow{pktw.} f_\infty, f_M \in \mathcal{R}[0, 1] \right] \Rightarrow \left[f_\infty \in \mathcal{R}[0, 1] \right], \quad (\text{VIII.35})$$

d.h. Aussagen (VIII.34) und (VIII.35) sind i.A. beide falsch! Aus diesem Grund wurden vor über 100 Jahren die Maßtheorie und das Lebesgue-Integral entwickelt – eine sehr fruchtbare Entwicklung, wie sich später gezeigt hat.

- Schließlich beobachten wir, dass sich aus der Dirichlet-Funktion leicht ein Gegenbeispiel für die Umkehrung von Satz VIII.7 (iii) konstruieren lässt: Setzen wir $f := f_\infty - \frac{1}{2}$, so ist $|f| \equiv \frac{1}{2} \in \mathcal{R}[0, 1]$ aber $f \notin \mathcal{R}[0, 1]$.

VIII.3. Integration und Differentiation

Satz VIII.11. Seien $a, b \in \mathbb{R}, a < b$ und $f \in \mathcal{R}[a, b]$. Sei $F : [a, b] \rightarrow \mathbb{R}$ definiert durch

$$\forall x \in (a, b] : \quad F(x) := \int_a^x f(t) dt, \quad F(a) := 0. \quad (\text{VIII.36})$$

Dann gelten folgende Aussagen:

(i) F ist stetig auf $[a, b]$.

(ii) Ist f stetig in $x_0 \in (a, b)$, dann ist F differenzierbar bei x_0 , und es gilt

$$F'(x_0) = f(x_0). \quad (\text{VIII.37})$$

Beweis.

Zu (i): Aus der Riemann-Integrierbarkeit von f folgt auch dessen Beschränktheit,

$$\exists M \in \mathbb{R}^+ \quad \forall x \in [a, b] : \quad |f(x)| \leq M. \quad (\text{VIII.38})$$

Damit ist aber nach Satz VIII.7

$$\forall a \leq y < x \leq b : \quad |F(x) - F(y)| = \left| \int_y^x f(t) dt \right| \leq M(x - y), \quad (\text{VIII.39})$$

und F ist (sogar gleichmäßig) stetig auf $[a, b]$.

Zu (ii): Sei $\varepsilon > 0$. Da f stetig in $x_0 \in (a, b)$ ist, gibt es ein $\delta > 0$, so dass

$$\forall t \in [x_0 - \delta, x_0 + \delta] : \quad |f(t) - f(x_0)| \leq \varepsilon. \quad (\text{VIII.40})$$

Für $x_0 < x \leq x_0 + \delta$ gilt also nach Satz VIII.7 (iii)

$$\begin{aligned}
 \left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| &= \frac{1}{|x - x_0|} \cdot |F(x) - F(x_0) - (x - x_0) \cdot f(x_0)| \\
 &= \frac{1}{|x - x_0|} \cdot \left| \int_{x_0}^x f(t) dt - \int_{x_0}^x f(x_0) dt \right| = \frac{1}{|x - x_0|} \cdot \left| \int_{x_0}^x (f(t) - f(x_0)) dt \right| \\
 &\leq \frac{1}{|x - x_0|} \int_{x_0}^x |f(t) - f(x_0)| dt \leq \frac{1}{|x - x_0|} \cdot \varepsilon \cdot \int_{x_0}^x dt = \varepsilon.
 \end{aligned} \tag{VIII.41}$$

Analog zu (VIII.41) erhält man dieselbe Ungleichung, falls $x_0 - \delta \leq x < x_0$ und somit insgesamt

$$\lim_{x \rightarrow x_0} \left\{ \frac{F(x) - F(x_0)}{x - x_0} \right\} = f(x_0). \tag{VIII.42}$$

□

Satz VIII.12 (Hauptsatz der Integral- und Differentialrechnung¹). *Seien $a, b \in \mathbb{R}$, $a < b$, $f \in \mathcal{R}[a, b]$ und $F : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$, differenzierbar auf (a, b) und derart, dass*

$$\forall x \in (a, b) : F'(x) = f(x). \tag{VIII.43}$$

Dann heißt F **Stammfunktion von f** ², und es gilt

$$\int_a^b f(x) dx = F(b) - F(a). \tag{VIII.44}$$

Beweis. Sei $P = \{x_1, \dots, x_L\} \in \mathcal{P}[a, b]$ eine Partition mit $x_0 := a < x_1 < x_2 < \dots < x_L < x_{L+1} := b$. Nach dem Mittelwertsatz der Differentialrechnung, Satz VII.8, gibt es zu jedem $k \in \{0, 1, 2, \dots, L\}$ ein $y_k \in (x_k, x_{k+1})$, so dass

$$F(x_{k+1}) - F(x_k) = (x_{k+1} - x_k) \cdot f(y_k). \tag{VIII.45}$$

Also ist

$$\begin{aligned}
 F(b) - F(a) &= F(x_{L+1}) - F(x_0) \\
 &= [F(x_{L+1}) - F(x_L)] + [F(x_L) - F(x_{L-1})] + \dots + [F(x_1) - F(x_0)] \\
 &= \sum_{k=0}^L \{F(x_{k+1}) - F(x_k)\} = \sum_{k=0}^L (x_{k+1} - x_k) \cdot f(y_k).
 \end{aligned} \tag{VIII.46}$$

Da

$$\inf_{x_k \leq x \leq x_{k+1}} \{f(x)\} \leq f(y_k) \leq \sup_{x_k \leq x \leq x_{k+1}} \{f(x)\}, \tag{VIII.47}$$

¹engl.: “fundamental theorem of calculus”

²engl.: “integral of f ”.

folgt aus (VIII.46), dass

$$\underline{\mathcal{I}}(f; P) \leq F(b) - F(a) \leq \overline{\mathcal{I}}(f; P). \quad (\text{VIII.48})$$

Also ist

$$\int_a^b f(x) dx = \sup_{P \in \mathcal{P}[a,b]} \{\underline{\mathcal{I}}(f; P)\} \leq F(b) - F(a) \leq \inf_{P \in \mathcal{P}[a,b]} \{\overline{\mathcal{I}}(f; P)\} = \int_a^b f(x) dx. \quad (\text{VIII.49})$$

□

Bemerkungen und Beispiele.

- Wir berechnen $I = \int_1^2 x^2 dx$. Für $F, f : \mathbb{R} \rightarrow \mathbb{R}$, mit $F(x) := \frac{1}{3}x^3$ und $f(x) := x^2$ sind F und f stetig und differenzierbar auf $\mathbb{R}$, deswegen ist auch $f \in \mathcal{R}[1, 2]$, und weiterhin gilt $F' = f$ auf $\mathbb{R}$. Daher ist

$$\int_1^2 x^2 dx = \int_1^2 f(x) dx = F(2) - F(1) = \frac{2^3 - 1^3}{3} = \frac{7}{3}. \quad (\text{VIII.50})$$

- Wir berechnen $J = \int_0^5 \cos(x) dx$. Für $F, f : \mathbb{R} \rightarrow \mathbb{R}$, mit $F(x) := \sin(x)$ und $f(x) := \cos(x)$ sind F und f stetig und differenzierbar auf $\mathbb{R}$, deswegen ist auch $f \in \mathcal{R}[0, 5]$, und es gilt $F' = f$ auf $\mathbb{R}$. Daher ist

$$\int_0^5 \cos(x) dx = \int_0^5 f(x) dx = F(5) - F(0) = \sin(5) - \sin(0) = \sin(5). \quad (\text{VIII.51})$$

Satz VIII.13 (Partielle Integration).³ Seien $a, b \in \mathbb{R}$, $a < b$ und $F, G : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$, differenzierbar auf (a, b) und $F', G' \in \mathcal{R}[a, b]$. Dann sind $F'G, G'F \in \mathcal{R}[a, b]$, und es gilt

$$\int_a^b F'(x)G(x) dx = F(b)G(b) - F(a)G(a) - \int_a^b F(x)G'(x) dx. \quad (\text{VIII.52})$$

Beweis. Man sieht leicht, dass mit $f, g \in \mathcal{R}[a, b]$ auch $f \cdot g \in \mathcal{R}[a, b]$. Da F und G stetig auf $[a, b]$ sind, gilt insbesondere $F, G \in \mathcal{R}[a, b]$. Demnach sind also auch $F'G, G'F \in \mathcal{R}[a, b]$. Setzen wir

$$\forall x \in [a, b] : \quad H(x) := F(x) \cdot G(x), \quad (\text{VIII.53})$$

so folgt (VIII.52) direkt aus Satz VIII.12 und der Produktregel, Satz VII.4. □

Bemerkungen und Beispiele.

³engl.: “integration by parts” und nicht “partial integration”

- Wir wollen $I = \int_1^2 x^2 \sin(x) dx$ mit Hilfe der partiellen Integration berechnen. Dabei stellt sich zunächst die Frage, welche Wahl wir für F und G treffen. Definieren wir $F'(x) := x^2$ und $G(x) := \sin(x)$, so sind $F(x) = \frac{1}{3}x^3 + C$ und $G'(x) = \cos(x)$, wobei $C \in \mathbb{R}$ eine frei wählbare Konstante ist, die wir gleich Null wählen, also $F(x) = \frac{1}{3}x^3$. Damit sind F, F', G, G' alle integrierbar auf $[1, 2]$, und

$$\begin{aligned} I &= \int_1^2 F'(x) G(x) dx = \left[F(x) G(x) \right]_1^2 - \int_1^2 F(x) G'(x) dx \\ &= \left[\frac{1}{3} x^3 \sin(x) \right]_1^2 - \frac{1}{3} \int_1^2 x^3 \cos(x) dx, \end{aligned} \quad (\text{VIII.54})$$

wobei wir wie üblich $[h(x)]_a^b := h(b) - h(a)$ notieren. Nun ist jedoch das ursprünglich zu berechnende Integral I in ein komplizierteres umgewandelt worden. Die obige Wahl von F und G ist also offensichtlich ungeeignet.

- Wir machen einen zweiten Versuch zur Berechnung von $I = \int_1^2 x^2 \sin(x) dx$ mit Hilfe der partiellen Integration. Diesmal setzen wir $F_1'(x) := \sin(x)$ und $G_1(x) := x^2$, sodass $F_1(x) = -\cos(x) + C$ und $G_1'(x) = 2x$, wobei $C \in \mathbb{R}$ eine frei wählbare Konstante ist, die wir gleich Null wählen, also $F_1(x) = -\cos(x)$. Damit sind F_1, F_1', G_1, G_1' alle integrierbar auf $[1, 2]$, und

$$\begin{aligned} I &= \int_1^2 F_1'(x) G_1(x) dx = \left[F_1(x) G_1(x) \right]_1^2 - \int_1^2 F_1(x) G_1'(x) dx \\ &= \left[-\cos(x) x^2 \right]_1^2 + 2 \int_1^2 x \cos(x) dx. \end{aligned} \quad (\text{VIII.55})$$

Der Grad des Monoms ist also um eins reduziert worden. Wir wenden nochmals partielle Integration an, und zwar mit $F_2'(x) := \cos(x)$ und $G_2(x) := x$, sodass $F_2(x) = \sin(x)$ und $G_2'(x) = 1$. Damit wird

$$\begin{aligned} I &= \left[-x^2 \cos(x) \right]_1^2 + 2 \int_1^2 x \cos(x) dx \\ &= \left[-x^2 \cos(x) \right]_1^2 + 2 \left[x \sin(x) \right]_1^2 - 2 \int_1^2 \sin(x) dx \\ &= \left[-x^2 \cos(x) + 2x \sin(x) + 2 \cos(x) \right]_1^2 \\ &= -2 \cos(2) + 4 \sin(2) - \cos(1) - 2 \sin(1). \end{aligned} \quad (\text{VIII.56})$$

Satz VIII.14 (Variablensubstitution). *Seien $a, b \in \mathbb{R}, a < b$ und $\varphi : [a, b] \rightarrow \mathbb{R}$, stetig auf $[a, b]$, differenzierbar auf (a, b) und strikt monoton steigend, $\varphi'(t) > 0, \forall t \in (a, b)$. Sei weiterhin $f \in \mathcal{R}[\varphi_a, \varphi_b]$, wobei $\varphi_a := \varphi(a), \varphi_b := \varphi(b)$. Dann ist $(t \mapsto (f \circ \varphi)(t) \cdot \varphi'(t)) \in \mathcal{R}[a, b]$, und es gilt*

$$\int_a^b f[\varphi(t)] \varphi'(t) dt = \int_{\varphi_a}^{\varphi_b} f(s) ds. \quad (\text{VIII.57})$$

Beweis. Hat f eine Stammfunktion F (d.h. es gilt $F' = f$), dann folgt die Aussage direkt aus dem Hauptsatz der Differential- und Integralrechnung und der Kettenregel.

Einerseits gilt nach dem Hauptsatz

$$\int_{\varphi_a}^{\varphi_b} f(s) \, ds = F(\varphi_b) - F(\varphi_a) = F(\varphi(b)) - F(\varphi(a)).$$

Andererseits ist

$$(F \circ \varphi)'(t) = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t)$$

und daher

$$\int_a^b f[\varphi(t)] \varphi'(t) \, dt = \int_a^b (F \circ \varphi)'(t) \, dt = (F \circ \varphi)(b) - (F \circ \varphi)(a) = F(\varphi(b)) - F(\varphi(a)),$$

was die Behauptung zeigt.

Den ausführlichen Beweis im allgemeinen Fall findet man im Anhang. $\square$

Bemerkungen und Beispiele.

- Wir bemerken, dass es einen Satz VIII.14 entsprechenden Satz für strikt monoton fallendes φ gibt.
- Merkregel: Substituiere $s := \varphi(t)$ und “erweitere” mit dt :

$$\int_{\varphi(a)}^{\varphi(b)} f(s) \, ds = \int_a^b f[\varphi(t)] \frac{d\varphi(t)}{dt} \, dt. \quad (\text{VIII.58})$$

Wir bemerken, dass (VIII.58) auch im Prinzip für nicht-monotone $\varphi : [a, b] \rightarrow \mathbb{R}$ gilt. Allerdings sind die Voraussetzungen stärker, und man muss auf formale Manipulation achten.

- Wir berechnen $J = \int_0^5 \frac{ds}{1+s^2}$. Dazu definieren wir $\varphi : (-\pi/2, \pi/2) \rightarrow \mathbb{R}$, $\varphi(t) := \tan(t) = \frac{\sin(t)}{\cos(t)}$ und beobachten, dass

$$\varphi'(t) = \frac{\sin^2(t) + \cos^2(t)}{\cos^2(t)} = 1 + \frac{\sin^2(t)}{\cos^2(t)} = 1 + \tan^2(t) = 1 + \varphi^2(t). \quad (\text{VIII.59})$$

Nach Satz VIII.14 ist damit

$$J = \int_{\arctan(0)}^{\arctan(5)} \frac{\varphi'(t) \, dt}{1 + \varphi^2(t)} = \int_0^{\arctan(5)} dt = \arctan(5). \quad (\text{VIII.60})$$

VIII.4. Absolute Integrabilität und Uneigentliche Integrale

In Kapitel V.2 haben wir gesehen, dass der Begriff der *absoluten Konvergenz* von Reihen eine zentrale Bedeutung besitzt; unter anderem wegen des großen Umordnungssatzes, Satz V.7. Wir bauen nun analog den entsprechenden Begriff der *absolut integrierbaren*

Funktionen auf. Dazu verwenden wir folgende Notation: Sind $f : \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion und $\lambda > 0$, so definieren wir $f_\lambda : [-\lambda, \lambda] \rightarrow [-\lambda, \lambda]$ durch

$$f_\lambda(x) := \mathbb{1}_{\{x : |f(x)| \leq \lambda\}}(x) f(x) - \lambda \mathbb{1}_{\{x : f(x) < -\lambda\}}(x) + \lambda \mathbb{1}_{\{x : f(x) > \lambda\}}(x), \quad (\text{VIII.61})$$

d.h. anschaulich gesprochen setzen wir einen quadratischen, um den Ursprung zentrierten Rahmen der Kantenlänge 2λ auf den Graphen von f und schneiden alles außerhalb des Rahmens ab. Analog zu (VIII.100) sieht man, dass für jedes Intervall $J \subseteq [-\lambda, \lambda]$

$$\sup_J \{f_\lambda\} - \inf_J \{f_\lambda\} \leq \sup_J \{f\} - \inf_J \{f\} \quad (\text{VIII.62})$$

gilt, woraus wiederum

$$0 \leq \overline{\mathcal{I}}(f_\lambda; P) - \underline{\mathcal{I}}(f_\lambda; P) \leq \overline{\mathcal{I}}(f; P) - \underline{\mathcal{I}}(f; P) \quad (\text{VIII.63})$$

für jede Partition $P \in \mathcal{P}[-\lambda, \lambda]$ folgt. Nach Lemma VIII.6 ist also $f_\lambda \in \mathcal{R}[-\lambda, \lambda]$, falls $f \in \mathcal{R}[-\lambda, \lambda]$.

Definition VIII.15. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ heißt **absolut (Riemann-)integrabel**

$$:\Leftrightarrow \forall \lambda > 0 : f_\lambda \in \mathcal{R}[-\lambda, \lambda], \quad (\text{VIII.64})$$

$$\sup_{\lambda > 0} \left\{ \int_{-\lambda}^{\lambda} |f_\lambda(x)| dx \right\} < \infty. \quad (\text{VIII.65})$$

In diesem Fall heißt

$$\int_{-\infty}^{\infty} f(x) dx := \lim_{\lambda \rightarrow \infty} \left\{ \int_{-\lambda}^{\lambda} f_\lambda(x) dx \right\} \quad (\text{VIII.66})$$

uneigentliches Integral von f über $\mathbb{R}$.

Satz VIII.16. Seien $f, g : \mathbb{R} \rightarrow \mathbb{R}$ absolut integrabel und $\alpha, a \in \mathbb{R}$. Dann ist auch $f + \alpha g$ absolut integrabel und es gelten

$$\int_{-\infty}^{\infty} (f + \alpha g)(x) dx = \int_{-\infty}^{\infty} f(x) dx + \alpha \int_{-\infty}^{\infty} g(x) dx, \quad (\text{VIII.67})$$

$$\int_{-\infty}^{\infty} f(x) dx = \lim_{m \rightarrow \infty} \left\{ \lim_{n \rightarrow \infty} \int_{-n}^m f(x) dx \right\} = \lim_{n \rightarrow \infty} \left\{ \lim_{m \rightarrow \infty} \int_{-n}^m f(x) dx \right\}, \quad (\text{VIII.68})$$

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^a f(x) dx + \int_a^{\infty} f(x) dx. \quad (\text{VIII.69})$$

Bemerkungen und Beispiele.

- Die Abbildungen $f, g, h : \mathbb{R} \rightarrow \mathbb{R}$, die durch

$$f(x) := \frac{\mathbb{1}_{(0,1)}(x)}{\sqrt{x}} = \begin{cases} \frac{1}{\sqrt{x}}, & 0 < x < 1 \\ 0, & \text{sonst,} \end{cases} \quad (\text{VIII.70})$$

$$g(x) := \frac{1}{1+x^2}, \quad (\text{VIII.71})$$

$$h(x) := \frac{\mathbb{1}_{\mathbb{R} \setminus \{0\}}(x)}{\sqrt{|x|} + x^2} = \begin{cases} \frac{1}{\sqrt{|x|} + x^2}, & x \neq 0, \\ 0, & x = 0, \end{cases} \quad (\text{VIII.72})$$

gegeben sind, sind alle absolut integrierbar.

Für f gilt

$$f_\lambda(x) = \begin{cases} \frac{1}{\sqrt{x}}, & \frac{1}{\lambda^2} < x < 1, \\ \lambda, & 0 < x \leq \frac{1}{\lambda^2}, \\ 0, & \text{sonst.} \end{cases} \quad (\text{VIII.73})$$

Also ist für $\lambda > 1$

$$\int_{-\lambda}^{\lambda} f_\lambda(x) dx = \int_0^{1/\lambda^2} \lambda dx + \int_{1/\lambda^2}^1 \frac{1}{\sqrt{x}} dx. \quad (\text{VIII.74})$$

Es ist

$$\int_0^{1/\lambda^2} \lambda dx + \int_{1/\lambda^2}^1 \frac{dx}{\sqrt{x}} = \lambda \cdot \frac{1}{\lambda^2} + \left[2\sqrt{x} \right]_{1/\lambda^2}^1 = \frac{1}{\lambda} + 2\left(1 - \frac{1}{\lambda}\right) \rightarrow 2, \quad (\text{VIII.75})$$

im Limes $\lambda \rightarrow \infty$.

Zur Untersuchung von g bemerken wir, dass $g_\lambda(x) = \frac{1}{1+x^2}$ für $\lambda > 1$ und daher

$$\int_{-\lambda}^{\lambda} g_\lambda(x) dx = \arctan(\lambda) - \arctan(-\lambda) \rightarrow \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi. \quad (\text{VIII.76})$$

im Limes $\lambda \rightarrow \infty$ gilt.

Für h ist das Argument etwas aufwändiger: Für $|x| < 1$ mit $x \neq 0$ gilt $h(x) \leq \frac{1}{\sqrt{|x|}}$, und für $|x| \geq 1$ gilt $h(x) \leq \frac{1}{1+x^2}$. Insgesamt ist also

$$0 \leq h(x) \leq \begin{cases} 0, & x = 0, \\ \frac{1}{\sqrt{|x|}}, & 0 < |x| < 1, \\ \frac{1}{1+x^2}, & |x| \geq 1. \end{cases} \quad (\text{VIII.77})$$

Für $|x| < 1$ argumentieren wir wie für f , und für $|x| \geq 1$ argumentieren wir wie bei der Untersuchung von g . Insgesamt sehen wir, dass für $\lambda > 1$

$$\int_{-\lambda}^{\lambda} |h_\lambda(x)| dx \leq \int_{-\lambda}^{-1} \frac{dx}{1+x^2} + \int_{-1}^{-1/\lambda^2} \frac{dx}{\sqrt{x}} + \int_{1/\lambda^2}^1 \frac{dx}{\sqrt{x}} + \int_1^{\lambda} \frac{dx}{1+x^2}, \quad (\text{VIII.78})$$

wobei wir $\int_a^b \frac{dx}{q(x)} := \int_a^b \frac{1}{q(x)} dx$ schreiben. Aus den vorigen Beispielen wissen wir, dass alle vier Integrale für $\lambda > 0$ beschränkt sind und daher folgt die absolute Integrierbarkeit von h . (Beachte: Den Wert des uneigentlichen Integrals $\int_{-\infty}^{\infty} h(x) dx$ haben wir nicht berechnet, aber wir wissen, dass er existiert und eindeutig ist.)

- Es ist zwar für alle $\lambda > 0$

$$\int_{-\lambda}^{\lambda} \sin(x) dx = 0, \quad (\text{VIII.79})$$

da $\sin(-x) = -\sin(x)$, und daher ist auch

$$\lim_{\lambda \rightarrow \infty} \int_{-\lambda}^{\lambda} \sin(x) \, dx = 0, \quad (\text{VIII.80})$$

aber dennoch ist $x \mapsto \sin(x)$ nicht absolut integrabel, denn für jedes $k \in \mathbb{Z}$ gilt

$$\int_{k\pi}^{(k+1)\pi} |\sin(x)| \, dx = 2 \quad (\text{VIII.81})$$

und daher

$$\int_{-\lambda}^{\lambda} |\sin(x)| \, dx \geq 2k, \quad (\text{VIII.82})$$

für $\lambda \geq k\pi$. Also ist

$$\sup_{\lambda > 0} \int_{-\lambda}^{\lambda} |\sin(x)| \, dx \geq \sup_{\lambda > 0} \left\{ \frac{2(\lambda - 1)}{\pi} \right\} = \infty. \quad (\text{VIII.83})$$

Wir bemerken, dass in diesem Fall auch die beiden Grenzwerte $\lim_{n \rightarrow \infty} \int_{-n}^n \sin(x) \, dx$ und $\lim_{m \rightarrow \infty} \int_{-m}^m \sin(x) \, dx$ nicht existieren.

VIII.5. Ergänzungen

VIII.5.1. Beweis von Lemma VIII.6

Beweis.

“ $\Leftarrow$ ”: Gemäß der rechten Seite in (VIII.17) gilt

$$0 \leq \overline{\int_a^b} f(x) dx - \underline{\int_a^b} f(x) dx \leq \overline{\mathcal{I}}(f; P) - \underline{\mathcal{I}}(f; P) \leq \varepsilon, \quad (\text{VIII.84})$$

für jedes $\varepsilon > 0$. Also muss

$$\overline{\int_a^b} f(x) dx = \underline{\int_a^b} f(x) dx \quad (\text{VIII.85})$$

und somit $f \in \mathcal{R}[a, b]$ sein.

“ $\Rightarrow$ ”: Nach Definition von Infimum und Supremum gibt es zu jedem $\varepsilon > 0$ zwei Partitionen $P', P'' \in \mathcal{P}[a, b]$ von $[a, b]$, so dass

$$\underline{\mathcal{I}}(f; P') \geq \underline{\int_a^b} f(x) dx - \frac{\varepsilon}{2}, \quad (\text{VIII.86})$$

$$\overline{\mathcal{I}}(f; P'') \leq \overline{\int_a^b} f(x) dx + \frac{\varepsilon}{2}. \quad (\text{VIII.87})$$

Da $f \in \mathcal{R}[a, b]$, gilt also für die Partition $P := P' \cup P''$ von $[a, b]$, dass

$$\begin{aligned} 0 &\leq \overline{\mathcal{I}}(f; P) - \underline{\mathcal{I}}(f; P) \leq \overline{\mathcal{I}}(f; P'') - \underline{\mathcal{I}}(f; P') \\ &\leq \left(\overline{\int_a^b} f(x) dx + \frac{\varepsilon}{2} \right) - \left(\underline{\int_a^b} f(x) dx - \frac{\varepsilon}{2} \right) = \varepsilon. \end{aligned} \quad (\text{VIII.88})$$

□

VIII.5.2. Monotonie und Linearität des Integrals

Beweis von Satz VIII.7. Seien $c, d \in [a, b]$ mit $c < d$, sodass $J := [c, d] \subseteq [a, b]$ ein Teilintervall von $[a, b]$ ist. Wir beobachten zunächst, dass für je zwei beschränkte Funktionen $f_1, f_2 : [a, b] \rightarrow \mathbb{R}$

$$\sup_J \{f_1 + f_2\} \leq \sup_J \{f_1\} + \sup_J \{f_2\}, \quad (\text{VIII.89})$$

$$\inf_J \{f_1 + f_2\} \geq \inf_J \{f_1\} + \inf_J \{f_2\} \quad (\text{VIII.90})$$

gilt, wobei wir hier und im Weiteren $\sup_J \{f\} := \sup_{x \in J} \{f(x)\}$ schreiben. Für jede Partition $P \in \mathcal{P}[a, b]$ erhalten wir daraus

$$\overline{\mathcal{I}}(f_1 + f_2; P) \leq \overline{\mathcal{I}}(f_1; P) + \overline{\mathcal{I}}(f_2; P), \quad (\text{VIII.91})$$

$$\underline{\mathcal{I}}(f_1 + f_2; P) \geq \underline{\mathcal{I}}(f_1; P) + \underline{\mathcal{I}}(f_2; P). \quad (\text{VIII.92})$$

Zu (i): Wegen $\underline{\mathcal{I}}(-g, P) = -\overline{\mathcal{I}}(g, P)$ ist mit $g \in \mathcal{R}[a, b]$ auch $-g \in \mathcal{R}[a, b]$, und es gilt

$$\int_a^b -g(x) dx = - \int_a^b g(x) dx. \quad (\text{VIII.93})$$

Daher reicht es, (VIII.18) für $\alpha > 0$ zu zeigen.

Sind nun $P', P'' \in \mathcal{P}[a, b]$ so, dass

$$\overline{\mathcal{I}}(f, P') \leq \int_a^b f(x) dx + \frac{\varepsilon}{2} \quad \text{und} \quad \overline{\mathcal{I}}(g, P'') \leq \int_a^b g(x) dx + \frac{\varepsilon}{2(\alpha + 1)}, \quad (\text{VIII.94})$$

so setzen wir $P := P' \cup P''$. Lemma VIII.3 und (VIII.89) implizieren nun, dass

$$\begin{aligned} \overline{\mathcal{I}}(f + \alpha g; P) &\leq \overline{\mathcal{I}}(f; P) + \overline{\mathcal{I}}(\alpha g; P) = \overline{\mathcal{I}}(f; P) + \alpha \overline{\mathcal{I}}(g; P) \\ &\leq \overline{\mathcal{I}}(f; P') + \alpha \overline{\mathcal{I}}(g; P'') \leq \int_a^b f(x) dx + \int_a^b g(x) dx + \varepsilon. \end{aligned} \quad (\text{VIII.95})$$

Mit $\varepsilon \rightarrow 0$ folgt

$$\overline{\int_a^b f(x) + \alpha g(x) dx} \leq \int_a^b f(x) dx + \alpha \int_a^b g(x) dx. \quad (\text{VIII.96})$$

Analog erhält man

$$\underline{\int_a^b f(x) + \alpha g(x) dx} \geq \int_a^b f(x) dx + \alpha \int_a^b g(x) dx, \quad (\text{VIII.97})$$

und somit (i).

Zu (ii) Glg. (VIII.19) ergibt sich aus (VIII.18) und der Tatsache, dass offensichtlich, für $h \in \mathcal{R}[a, b]$ mit $h(x) \geq 0$, für $x \in [a, b]$, auch

$$\int_a^b h(x) dx \geq 0, \quad (\text{VIII.98})$$

da $\underline{\mathcal{I}}(h; P) \geq 0$, für alle $P \in \mathcal{P}[a, b]$. Mit (VIII.18) und (VIII.98) erhalten wir nämlich

$$\int_a^b g(x) dx - \int_a^b f(x) dx = \int_a^b (g - f)(x) dx \geq 0. \quad (\text{VIII.99})$$

Zu (iii) Sei $J := [c, d] \subseteq [a, b]$ ein Teilintervall von $[a, b]$ wie oben. Wir zeigen zunächst, dass

$$\sup_J \{f\} - \inf_J \{f\} \geq \sup_J \{|f|\} - \inf_J \{|f|\} \quad (\text{VIII.100})$$

gilt. Für $f \geq 0$ und somit $|f| \equiv f$ auf J ist dies offensichtlich, ebenso für $f \leq 0$, d.h. $|f| \equiv -f$ auf J . Sind hingegen $\inf_J \{f\} < 0 < \sup_J \{f\}$, so folgt mit $f_{\pm} := \pm f \cdot \mathbf{1}[\pm f > 0] \geq 0$ und

$$f = f_+ - f_- \quad \text{und} \quad |f| = f_+ + f_-, \quad (\text{VIII.101})$$

dass $\sup_J \{f\} = \sup_J \{f_+\}$, $\inf_J \{f\} = -\sup_J \{f_-\}$ und deshalb

$$\begin{aligned} \sup_J \{f\} - \inf_J \{f\} &= \sup_J \{f_+\} + \sup_J \{f_-\} \geq \sup_J \{f_+ + f_-\} = \sup_J \{|f|\} \\ &\geq \sup_J \{|f|\} - \inf_J \{|f|\}. \end{aligned} \quad (\text{VIII.102})$$

Somit gilt (VIII.100) allgemein für alle $f \in \mathcal{R}[a, b]$. Aus (VIII.100) folgt jedoch unmittelbar für jede Partition $P \in \mathcal{P}[a, b]$, dass

$$0 \leq \overline{\mathcal{I}}(|f|; P) - \underline{\mathcal{I}}(|f|; P) \leq \overline{\mathcal{I}}(f; P) - \underline{\mathcal{I}}(f; P), \quad (\text{VIII.103})$$

und Lemma VIII.6 impliziert sofort, dass auch $|f| \in \mathcal{R}[a, b]$. Die Ungleichung (VIII.20) ist nun eine unmittelbare Konsequenz aus (ii) und $|f| \geq \pm f$. $\square$

VIII.5.3. Integrierbarkeit stetiger Funktionen

Beweis von Satz VIII.8. Da $[a, b]$ kompakt ist, nimmt nach Korollar VI.11 f auf $[a, b]$ ein Minimum und Maximum an. Mit $M := \max \{|\sup f([a, b])|, |\inf f([a, b])|\} < \infty$ gilt also

$$\forall x \in [a, b] : \quad f(x) \in [-M, M], \quad (\text{VIII.104})$$

und f ist beschränkt. Außerdem ist f gemäß Satz VI.12 sogar *gleichmäßig stetig* auf $[a, b]$, das heißt, zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass

$$\forall x, x' \in [a, b], \quad |x - x'| \leq \delta : \quad |f(x) - f(x')| \leq \varepsilon. \quad (\text{VIII.105})$$

Sei nun $\varepsilon > 0$. Dann existiert $\delta > 0$ wie oben und ohne Einschränkung nehmen wir $\delta \leq b - a$. Dann setzen wir

$$x_0 := a, \quad x_1 := a + \delta, \quad x_2 := a + 2\delta, \dots, x_L := a + L\delta, \quad x_{L+1} := b, \quad (\text{VIII.106})$$

wobei $L \in \mathbb{N}$ so gewählt wird, dass

$$a + L\delta < b \leq a + (L + 1)\delta. \quad (\text{VIII.107})$$

Weiterhin ist $f : I_k := [x_k, x_{k+1}] \rightarrow \mathbb{R}$ stetig, für jedes $k = 0, 1, 2, \dots, L$, und aus der Kompaktheit von I_k folgt die Existenz von $x_k^{\min}, x_k^{\max} \in I_k$, sodass

$$f(x_k^{\min}) = \inf_{x \in I_k} \{f(x)\}, \quad f(x_k^{\max}) = \sup_{x \in I_k} \{f(x)\}. \quad (\text{VIII.108})$$

Dann ist $|x_k^{\min} - x_k^{\max}| \leq \delta$ und nach (VIII.105)

$$0 \leq \sup_{x \in I_k} \{f(x)\} - \inf_{x \in I_k} \{f(x)\} = f(x_k^{\max}) - f(x_k^{\min}) \leq \varepsilon. \quad (\text{VIII.109})$$

Definieren wir durch $P := \{x_1, x_2, \dots, x_L\} \in \mathcal{P}[a, b]$ eine Partition, so erhalten wir damit

$$\begin{aligned} 0 \leq \overline{\mathcal{I}}(f; P) - \underline{\mathcal{I}}(f; P) &= \sum_{k=0}^L (x_{k+1} - x_k) \cdot \underbrace{\left(\sup_{x \in I_k} \{f(x)\} - \inf_{x \in I_k} \{f(x)\} \right)}_{\leq \varepsilon} \\ &\leq \varepsilon \cdot \sum_{k=0}^L (x_{k+1} - x_k) = \varepsilon \cdot (b - a). \end{aligned} \quad (\text{VIII.110})$$

Also ist

$$0 \leq \overline{\int_a^b f(x) dx} - \underline{\int_a^b f(x) dx} \leq \overline{\mathcal{I}}(f; P) - \underline{\mathcal{I}}(f; P) \leq \varepsilon \cdot (b - a), \quad (\text{VIII.111})$$

und dies gilt für jedes $\varepsilon > 0$. Es folgt also $\overline{\int_a^b f(x) dx} = \underline{\int_a^b f(x) dx}$. $\square$

VIII.5.4. Beweis von Satz VIII.9

Beweis von Satz VIII.9. Für den Beweis führen wir etwas neue Notation ein. Ist $P \in \mathcal{P}[a, b]$ eine Partition, $P = \{x_1, \dots, x_L\} \subseteq (a, b)$, $x_0 := a < x_1, x_2, \dots < x_L < x_{L+1} := b$, und ist $f : [a, b] \rightarrow \mathbb{R}$ beschränkt, so schreiben wir

$$\overline{\mathcal{I}}_a^b(f; P) := \sum_{j=0}^L (x_{j+1} - x_j) \cdot \sup_{x_j \leq x \leq x_{j+1}} \{f(x)\}, \quad (\text{VIII.112})$$

$$\underline{\mathcal{I}}_a^b(f; P) := \sum_{j=0}^L (x_{j+1} - x_j) \cdot \inf_{x_j \leq x \leq x_{j+1}} \{f(x)\}. \quad (\text{VIII.113})$$

Seien $\varepsilon > 0$ und $P \in \mathcal{P}[a, b]$ eine Partition derart, dass

$$\overline{\mathcal{I}}_a^b(f; P) - \underline{\mathcal{I}}_a^b(f; P) \leq \varepsilon. \quad (\text{VIII.114})$$

Dann ist $P \cap (a, c) \in \mathcal{P}[a, c]$ eine Partition von $[a, c]$, und es gilt

$$\begin{aligned} & \overline{\mathcal{I}}_a^c[f; P \cap (a, c)] - \underline{\mathcal{I}}_a^c[f; P \cap (a, c)] \\ & \leq \overline{\mathcal{I}}_a^c[f; P \cap (a, c)] - \underline{\mathcal{I}}_a^c[f; P \cap (a, c)] + \overline{\mathcal{I}}_c^b[f; P \cap (c, b)] - \overline{\mathcal{I}}_c^b[f; P \cap (c, b)] \\ & = \overline{\mathcal{I}}_a^b[f; P \cup \{c\}] - \underline{\mathcal{I}}_a^b[f; P \cup \{c\}] \\ & \leq \overline{\mathcal{I}}_a^b(f; P) - \underline{\mathcal{I}}_a^b(f; P) \leq \varepsilon, \end{aligned} \quad (\text{VIII.115})$$

wobei wir

$$\underline{\mathcal{I}}_a^b[f; P \cup \{c\}] = \underline{\mathcal{I}}_a^c[f; P \cap (a, c)] + \underline{\mathcal{I}}_c^b[f; P \cap (c, b)] \quad (\text{VIII.116})$$

benutzen. Aus (VIII.115) folgt aber nach Definition VIII.5, dass $f \in \mathcal{R}[a, c]$. Analog zeigt man, dass $f \in \mathcal{R}[c, b]$. Aus (VIII.116) erhalten wir außerdem, dass

$$\begin{aligned} \overline{\int_a^b f(x) dx} &= \inf_{P \in \mathcal{P}[a, b]} \{\overline{\mathcal{I}}_a^b(f; P)\} = \inf_{P \in \mathcal{P}[a, b]} \{\overline{\mathcal{I}}_a^b(f; P \cup \{c\})\} \\ &= \inf_{P \in \mathcal{P}[a, b]} \{\overline{\mathcal{I}}_a^c(f; P \cap (a, c)) + \overline{\mathcal{I}}_c^b(f; P \cap (c, b))\} \\ &= \inf_{P \in \mathcal{P}[a, c]} \{\overline{\mathcal{I}}_a^c(f; P)\} + \inf_{P \in \mathcal{P}[c, b]} \{\overline{\mathcal{I}}_c^b(f; P)\} \\ &= \overline{\int_a^c f(x) dx} + \overline{\int_c^b f(x) dx}, \end{aligned} \quad (\text{VIII.117})$$

und analog zeigt man die Gleichheit der Unterintegrale.

Für den Beweis von (VIII.23) bemerken wir zunächst, dass mit f_{ac}, f_{cb} und $f(c) \in \mathbb{R}$ auch die Funktion $f : [a, b] \rightarrow \mathbb{R}$ in (VIII.22) beschränkt ist, also $|f(x)| \leq M$, für ein gewisses $M < \infty$ und alle $x \in [a, b]$ gilt.

Seien $\varepsilon > 0$ und $P' \in \mathcal{P}[a, c]$, $P'' \in \mathcal{P}[c, b]$ so, dass

$$\bar{\mathcal{I}}_a^c(f_{ac}, P') - \int_a^c f_{ac}(x) dx \leq \frac{\varepsilon}{4}, \quad (\text{VIII.118})$$

$$\bar{\mathcal{I}}_c^b(f_{cb}, P'') - \int_c^b f_{cb}(x) dx \leq \frac{\varepsilon}{4}. \quad (\text{VIII.119})$$

Wählen wir nun $\delta > 0$ so klein, dass

$$P' \cap (a, c - \delta) = P'' \cap (c + \delta, b) = \emptyset, \quad (\text{VIII.120})$$

so folgt mit (VIII.117) und

$$P := P' \cup \{c - \delta, c + \delta\} \cup P'', \quad (\text{VIII.121})$$

dass

$$\begin{aligned} \bar{\mathcal{I}}_a^b(f; P) &= \bar{\mathcal{I}}_a^{c-\delta}(f; P') + \bar{\mathcal{I}}_{c-\delta}^{c+\delta}(f; \emptyset) + \bar{\mathcal{I}}_{c+\delta}^b(f; P'') \\ &= \bar{\mathcal{I}}_a^{c-\delta}(f_{ac}; P') + \bar{\mathcal{I}}_{c-\delta}^{c+\delta}(f; \emptyset) + \bar{\mathcal{I}}_{c+\delta}^b(f_{cb}; P'') \\ &\leq \bar{\mathcal{I}}_a^c(f_{ac}; P') + 6\delta M + \bar{\mathcal{I}}_a^b(f_{cb}; P'') \\ &\leq \int_a^c f_{ac}(x) dx + \int_c^b f_{cb}(x) dx + \frac{\varepsilon}{2} + 6\delta M. \end{aligned} \quad (\text{VIII.122})$$

Wählen wir nun $\delta \leq \varepsilon/12M$, so folgt aus (VIII.122), dass

$$\overline{\int_a^b f(x) dx} \leq \bar{\mathcal{I}}_a^b(f; P) \leq \int_a^c f_{ac}(x) dx + \int_c^b f_{cb}(x) dx + \varepsilon, \quad (\text{VIII.123})$$

für jedes $\varepsilon > 0$. Also ist

$$\overline{\int_a^b f(x) dx} \leq \int_a^c f_{ac}(x) dx + \int_c^b f_{cb}(x) dx. \quad (\text{VIII.124})$$

Genauso zeigt man, dass

$$\underline{\int_a^b f(x) dx} \geq \int_a^c f_{ac}(x) dx + \int_c^b f_{cb}(x) dx, \quad (\text{VIII.125})$$

und daraus folgt die Behauptung, denn das Unterintegral ist stets kleiner als das Oberintegral. $\square$

VIII.5.5. Variablensubstitution – Beweis von Satz VIII.14

Beweis.

(i): Wir nehmen zunächst an, f wäre konstant auf $[\varphi_a, \varphi_b]$, also

$$\forall \varphi_a \leq s \leq \varphi_b : f(s) = \lambda, \quad (\text{VIII.126})$$

für eine geeignete Konstante $\lambda \in \mathbb{R}$. Dann ist die Behauptung sicher richtig, denn nach dem Hauptsatz der Differential- und Integralrechnung gilt

$$\int_a^b f[\varphi(t)] \varphi'(t) dt = \lambda \int_a^b \varphi'(t) dt = \lambda[\varphi(b) - \varphi(a)] = \lambda \int_{\varphi_a}^{\varphi_b} ds = \lambda \int_{\varphi_a}^{\varphi_b} f(s) ds. \quad (\text{VIII.127})$$

(ii): Als Nächstes beobachten wir, dass sich diese Aussage nach Korollar VIII.10 auf stückweise konstante Funktionen überträgt. Sind nämlich $P = \{y_1, y_2, \dots, y_L\} \in \mathcal{P}[\varphi_a, \varphi_b]$, mit $y_0 := \varphi(a) < y_1 < y_2 < \dots < y_{L+1} := \varphi(b)$, und

$$f(s) := \begin{cases} \lambda_k, & \text{falls } s \in (y_k, y_{k+1}), \\ \mu_k, & \text{falls } s = y_k, \end{cases} \quad (\text{VIII.128})$$

mit $\lambda_0, \lambda_2, \dots, \lambda_L, \mu_0, \mu_1, \dots, \mu_{L+1} \in \mathbb{R}$, so gilt nach (VIII.127))

$$\begin{aligned} \int_a^b f[\varphi(t)] \varphi'(t) dt &= \sum_{k=0}^L \lambda_k \int_{\varphi^{-1}(y_k)}^{\varphi^{-1}(y_{k+1})} \varphi'(t) dt \\ &= \sum_{k=0}^L \lambda_k [\varphi(\varphi^{-1}(y_{k+1})) - \varphi(\varphi^{-1}(y_k))] = \sum_{k=0}^L \lambda_k (y_{k+1} - y_k) \\ &= \sum_{k=0}^L \int_{y_k}^{y_{k+1}} f(s) ds = \int_{\varphi_a}^{\varphi_b} f(s) ds. \end{aligned} \quad (\text{VIII.129})$$

(iii): Seien nun $f \in \mathcal{R}[\varphi_a, \varphi_b]$ ohne weitere Zusatzannahmen, $\varepsilon > 0$ und $P \in \mathcal{P}[\varphi_a, \varphi_b]$, mit $P = \{y_1, \dots, y_L\}$ so, dass $y_0 := \varphi_a < y_1 < y_2 < \dots < y_L < y_{L+1} = \varphi_b$ und

$$\overline{\mathcal{I}}_{\varphi_a}^{\varphi_b}(f; P) - \underline{\mathcal{I}}_{\varphi_a}^{\varphi_b}(f; P) \leq \varepsilon. \quad (\text{VIII.130})$$

Wir definieren nun $f_{gr}, f_{kl} \in \mathcal{R}[a, b]$ durch $\bar{\lambda}_0 := \bar{\lambda}_{L+1} := -\infty, \underline{\lambda}_0 := \underline{\lambda}_{L+1} := \infty$,

$$f_{gr}(s) := \begin{cases} \bar{\lambda}_k := \sup_{y_k \leq s \leq y_{k+1}} \{f(s)\}, & \text{falls } s \in (y_k, y_{k+1}) \\ \max\{\bar{\lambda}_{k-1}, \bar{\lambda}_k\}, & \text{falls } s = y_k, \end{cases} \quad (\text{VIII.131})$$

und

$$f_{kl}(s) := \begin{cases} \underline{\lambda}_k := \inf_{y_k \leq s \leq y_{k+1}} \{f(s)\}, & \text{falls } s \in (y_k, y_{k+1}) \\ \min\{\bar{\lambda}_{k-1}, \bar{\lambda}_k\}, & \text{falls } s = y_k, \end{cases} \quad (\text{VIII.132})$$

Damit sind

$$\forall s \in [\varphi_a, \varphi_b] : f_{kl}(s) \leq f(s) \leq f_{gr}(s). \quad (\text{VIII.133})$$

Außerdem sind

$$\begin{aligned}\bar{\mathcal{I}}_{\varphi_a}^{\varphi_b}(f; P) &= \bar{\mathcal{I}}_{\varphi_a}^{\varphi_b}(f_{gr}; P) \\ &= \int_{\varphi_a}^{\varphi_b} f_{gr}(s) ds = \int_a^b f_{gr}(\varphi(t)) \varphi'(t) dt \geq \overline{\int_a^b f(\varphi(t)) \varphi'(t) dt}\end{aligned}\quad (\text{VIII.134})$$

und analog

$$\begin{aligned}\underline{\mathcal{I}}_{\varphi_a}^{\varphi_b}(f; P) &= \underline{\mathcal{I}}_{\varphi_a}^{\varphi_b}(f_{kl}; P) \\ &= \int_{\varphi_a}^{\varphi_b} f_{kl}(s) ds = \int_a^b (f_{kl} \circ \varphi)(t) \varphi'(t) dt \leq \underline{\int_a^b f(\varphi(t)) \varphi'(t) dt}.\end{aligned}\quad (\text{VIII.135})$$

Also ist

$$\overline{\int_a^b f(\varphi(t)) \varphi'(t) dt} \leq \bar{\mathcal{I}}_{\varphi_a}^{\varphi_b}(f; P) \leq \underline{\mathcal{I}}_{\varphi_a}^{\varphi_b}(f; P) + \varepsilon \leq \int_{\varphi_a}^{\varphi_b} f(s) ds + \varepsilon \quad (\text{VIII.136})$$

und genauso

$$\underline{\int_a^b f(\varphi(t)) \varphi'(t) dt} \geq \underline{\mathcal{I}}_{\varphi_a}^{\varphi_b}(f; P) \geq \bar{\mathcal{I}}_{\varphi_a}^{\varphi_b}(f; P) - \varepsilon \geq \int_{\varphi_a}^{\varphi_b} f(s) ds - \varepsilon, \quad (\text{VIII.137})$$

für jedes $\varepsilon > 0$. Mit $\varepsilon \rightarrow 0$ erhalten wir also

$$\int_{\varphi_a}^{\varphi_b} f(s) ds \leq \underline{\int_a^b f(\varphi(t)) \varphi'(t) dt} \leq \overline{\int_a^b f(\varphi(t)) \varphi'(t) dt} \leq \int_{\varphi_a}^{\varphi_b} f(s) ds, \quad (\text{VIII.138})$$

was die Behauptung ergibt. $\square$

VIII.5.6. Der Satz von Taylor mit Integralrestglied

Wir geben noch eine Formulierung des Satzes von Taylor, in der das Taylorsche Restglied $R_{n+1}[f, x_0; x]$ aus Satz VII.12 durch das *Taylorsche Integralrestglied* $\tilde{R}_{n+1}[f, x_0; x]$ ersetzt wird. Dies hat den Vorteil, dass $\tilde{R}_{n+1}[f, x_0; x]$ genau quantifiziert ist, während bei $R_{n+1}[f, x_0; x] = \frac{(x-x_0)^{n+1}}{(n+1)!} f^{(n+1)}(x')$ die durch Anwendung des Mittelwertsatzes resultierende Zwischenstelle x' unbestimmt bleibt.

Allerdings sind in Satz VIII.17 die Voraussetzungen etwas stärker als in Satz VII.12, denn zusätzlich zur $(n+1)$ -maligen Differenzierbarkeit von f wird die Stetigkeit der $(n+1)$. Ableitung $f^{(n+1)}$ gefordert.

Satz VIII.17 (Taylor). *Seien $a, b, x_0 \in \mathbb{R}, a < x_0 < b, n \in \mathbb{N}$ und $f : (a, b) \rightarrow \mathbb{R}$ auf (a, b) $(n+1)$ -mal differenzierbar und $f^{(n+1)}$ stetig auf (a, b) . Dann ist, für alle $x \in (a, b)$*

$$f(x) = T_n[f, x_0; x] + \tilde{R}_{n+1}[f, x_0; x] \quad (\text{VIII.139})$$

$$= \sum_{k=0}^n \left\{ \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \right\} + \tilde{R}_{n+1}[f, x_0; x], \quad (\text{VIII.140})$$

wobei das **Taylorische Integralrestglied** $(n+1)$. **Ordnung** durch

$$\tilde{R}_{n+1}[f, x_0; x] := \int_{x_0}^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt, \quad \text{falls } x \geq x_0, \text{ bzw.} \quad (\text{VIII.141})$$

$$\tilde{R}_{n+1}[f, x_0; x] := \int_x^{x_0} \frac{(t-x)^n}{n!} f^{(n+1)}(t) dt, \quad \text{falls } x < x_0, \quad (\text{VIII.142})$$

gegeben ist.

Beweis. Wir beweisen durch Induktion, dass für alle $m = 0, 1, 2, \dots, n$

$$f(x) = \sum_{k=0}^m \left\{ \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k \right\} + \tilde{R}_{m+1}(x), \quad (\text{VIII.143})$$

gilt, wobei wir der Einfachheit halber $x > x_0$ annehmen wollen. (Der Fall $x < x_0$ ist analog, und $x = x_0$ ist trivial.) Für $m = 0$ ergibt sich (VIII.143) aus dem Hauptsatz der Differential- und Integralrechnung,

$$f(x) = f(x_0) + f(x) - f(x_0) = f(x_0) + \int_{x_0}^x f'(t) dt = f(x_0) + \tilde{R}_1(x). \quad (\text{VIII.144})$$

Gilt nun (VIII.143) für $\tilde{m} = m-1 \in \{0, 1, 2, \dots, n-1\}$, so erhalten wir

$$\begin{aligned} f(x) - \sum_{k=0}^m \left\{ \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k \right\} &= \left(f(x) - \sum_{k=0}^{m-1} \left\{ \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k \right\} \right) - \frac{f^{(m)}(x_0)}{m!} (x-x_0)^m \\ &= \tilde{R}_m(x) - \frac{f^{(m)}(x_0)}{m!} (x-x_0)^m \\ &= \int_{x_0}^x \frac{(x-t)^{m-1}}{(m-1)!} (f^{(m)}(t) - f^{(m)}(x_0)) dt, \end{aligned} \quad (\text{VIII.145})$$

da $\int_{x_0}^x (x-t)^{m-1} dt = \frac{(x-x_0)^m}{m}$ ist. Durch abermalige Anwendung des Hauptsatzes der Differential- und Integralrechnung erhalten wir also, dass

$$f(x) - \sum_{k=0}^m \left\{ \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k \right\} = \int_{x_0}^x \frac{(x-t)^{m-1}}{(m-1)!} \left(\int_{x_0}^t f^{(m+1)}(s) ds \right) dt. \quad (\text{VIII.146})$$

Setzen wir nun

$$F(t) := \int_{x_0}^t f^{(m+1)}(s) ds \quad \text{und} \quad G(t) := \frac{-(x-t)^m}{m!}, \quad (\text{VIII.147})$$

so folgt dass

$$F'(t) = f^{(m+1)}(t) \quad \text{und} \quad G'(t) = \frac{(x-t)^{m-1}}{(m-1)!}, \quad (\text{VIII.148})$$

und erhalten wir durch partielle Integration, dass

$$\begin{aligned}
 \int_{x_0}^x \frac{(x-t)^{m-1}}{(m-1)!} \left(\int_{x_0}^t f^{(m+1)}(s) ds \right) dt &= \int_{x_0}^x F(t) G'(t) dt \\
 &= F(x) \underbrace{G(x)}_{=0} - \underbrace{F(x_0) G(x_0)}_{=0} - \int_{x_0}^x F'(t) G(t) dt \\
 &= \int_{x_0}^x f^{(m+1)}(t) \frac{(x-t)^m}{m!} dt.
 \end{aligned} \tag{VIII.149}$$

□

IX. Normen und Banachräume

Zur Vorbereitung der Differenzialrechnung für vektorwertige Funktionen rufen wir einige Begriffe aus der Vorlesung zur *Linearen Algebra* in Erinnerung. Weitere Grundbegriffe der linearen Algebra, die für diese Vorlesung wichtig sind, finden Studienanfänger im ergänzenden Abschnitt IX.3.1. Ebenso wiederholen wir die schon bekannte Definition einer *Norm* und ergänzen diese mit zugehörigen topologischen Begriffen, die uns erlauben, Konvergenz auch für vektorwertige Folgen zu definieren.

Zunächst erinnern wir an den Begriff des Vektorraums selbst: Für $d \in \mathbb{N}$ und $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ ist

$$\mathbb{K}^d := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} \mid x_1, x_2, \dots, x_d \in \mathbb{K} \right\} \quad (\text{IX.1})$$

mit komponentenweise definierten Vektorraumverknüpfungen

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} + \alpha \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{pmatrix} := \begin{pmatrix} x_1 + \alpha y_1 \\ x_2 + \alpha y_2 \\ \vdots \\ x_d + \alpha y_d \end{pmatrix} \quad (\text{IX.2})$$

ein Vektorraum über $\mathbb{K}$ der Dimension d ist. Mit Hilfe der kanonischen Basisvektoren

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad e_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \dots, \quad e_d := \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in \mathbb{K}^d \quad (\text{IX.3})$$

kann man

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} = x_1 e_1 + x_2 e_2 + \dots + x_d e_d \quad (\text{IX.4})$$

schreiben. Die Darstellung (IX.4) eines Vektors $x \in \mathbb{K}^d$ als Linearkombination $x = \sum_{i=1}^d x_i e_i$ ist eindeutig.

Auf $\mathbb{R}^d$ ist durch

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \middle| \begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix} \right\rangle_{\text{eukl}} := x_1 y_1 + x_2 y_2 + \dots + x_d y_d \quad (\text{IX.5})$$

das euklidische Skalarprodukt definiert, und es gilt die Cauchy-Schwarzsche Ungleichung,

$$|\langle x|y \rangle| \leq \sqrt{\langle x|x \rangle} \sqrt{\langle y|y \rangle}. \quad (\text{IX.6})$$

IX.1. Normierte Räume und Banachräume

Definition IX.1. Sei X ein $\mathbb{K}$ -Vektorraum ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$).

Eine Abbildung $\|\cdot\| : X \rightarrow \mathbb{R}_0^+$ heißt **Norm (auf X)** $:\Leftrightarrow$

$$(i) \quad \forall x \in X : \quad \{\|x\| = 0 \Leftrightarrow x = 0\} \quad (\text{IX.7})$$

$$(ii) \quad \forall x \in X, \lambda \in \mathbb{K} : \quad \|\lambda x\| = |\lambda| \cdot \|x\|, \quad (\text{IX.8})$$

$$(iii) \quad \forall x, y \in X : \quad \|x + y\| \leq \|x\| + \|y\|. \quad (\text{IX.9})$$

In diesem Fall bezeichnet man $(X, \|\cdot\|)$ als **normierten Vektorraum (über $\mathbb{K}$)**.

Bemerkungen und Beispiele.

- Ist $(X, \|\cdot\|)$ ein normierter Vektorraum über $\mathbb{K}$, so definiert die Norm $\|\cdot\| : X \rightarrow \mathbb{R}_0^+$ einen Abstandsbegriff auf X . Für $x, y \in X$ kann nämlich $\|x - y\|$ als Abstand von x und y interpretiert werden. Mit diesem Abstandsbegriff können viele in $\mathbb{R}^2$ oder $\mathbb{R}^3$ angestellten geometrische Überlegungen auf X übertragen werden.
- Ist X ein $\mathbb{K}$ -Vektorraum mit Skalarprodukt $\langle \cdot | \cdot \rangle : X \times X \rightarrow \mathbb{K}$, so induziert das Skalarprodukt durch $\|x\| := \sqrt{\langle x|x \rangle}$ eine Norm auf X .
- Für $d \in \mathbb{N}$ ist $(\mathbb{R}^d, \|\cdot\|_{\text{eukl}})$ bezüglich der **euklidischen Norm** $\|x\|_{\text{eukl}} = \sqrt{\langle x|x \rangle_{\text{eukl}}}$ ein normierter Vektorraum über $\mathbb{R}$.
- Für $d \in \mathbb{N}$ ist $(\mathbb{C}^d, \|\cdot\|_{\text{unit}})$ bezüglich der **unitären Norm** $\|x\|_{\text{unit}} = \sqrt{\langle x|x \rangle_{\text{unit}}}$ ein normierter Vektorraum über $\mathbb{C}$.
- Für $d \in \mathbb{N}$ und $1 \leq p < \infty$ definiert $\|x\|_p := (|x_1|^p + |x_2|^p + \dots + |x_d|^p)^{\frac{1}{p}}$ eine Norm auf $\mathbb{K}^d$; ebenso $\|x\|_\infty := \max\{|x_1|, |x_2|, \dots, |x_d|\}$. Diese Normen sind alle äquivalent zueinander, d.h. zu $1 \leq p, q \leq \infty$ gibt es eine Konstante $C_{d,p,q} < \infty$ so, dass

$$\forall x \in X : \quad \frac{1}{C_{d,p,q}} \cdot \|x\|_p \leq \|x\|_q \leq C_{d,p,q} \cdot \|x\|_p. \quad (\text{IX.10})$$

Die durch das euklidische bzw. unitäre Skalarprodukt induzierte Norm ist $\|x\|_2$.

Definition IX.2. Sei $(X, \|\cdot\|)$ ein normierter Vektorraum über $\mathbb{K}$.

(i) Eine Folge $(x_n)_{n=1}^\infty \in X^\mathbb{N}$ heißt **konvergent** $:\Leftrightarrow$

$$\exists x \in X \quad \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \|x - x_n\| \leq \varepsilon. \quad (\text{IX.11})$$

In diesem Fall schreiben wir

$$\lim_{n \rightarrow \infty} \{x_n\} = x \quad \text{oder} \quad x_n \rightarrow x, \quad n \rightarrow \infty. \quad (\text{IX.12})$$

(ii) Eine Folge $(x_n)_{n=1}^\infty \in X^\mathbb{N}$ heißt **Cauchy-Folge** $:\Leftrightarrow$

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall m, n \geq n_0 : \quad \|x_m - x_n\| \leq \varepsilon. \quad (\text{IX.13})$$

(iii) Ein normierter Vektorraum $(X, \|\cdot\|)$ heißt **vollständig** $:\Leftrightarrow$

$$\text{Jede Cauchy-Folge in } (X, \|\cdot\|) \text{ ist konvergent.} \quad (\text{IX.14})$$

Einen vollständigen normierten Vektorraum bezeichnet man als **Banachraum**.

Lemma IX.3. Seien $d \in \mathbb{N}$ und $(x^{(n)})_{n=1}^\infty \in \mathbb{K}^d$ eine Folge von Vektoren in $(\mathbb{K}^d, \|\cdot\|_{\text{eukl/unit}})$ mit

$$x^{(n)} = \begin{pmatrix} x_1^{(n)} \\ x_2^{(n)} \\ \vdots \\ x_d^{(n)} \end{pmatrix}. \quad (\text{IX.15})$$

Dann sind folgende Aussagen gleichwertig:

$$\left\{ (x^{(n)})_{n=1}^\infty \text{ ist konvergent in } \mathbb{K}^d \right\} \quad (\text{IX.16})$$

$$\Leftrightarrow \left\{ \forall i \in \mathbb{Z}_1^d : (x_i^{(n)})_{n=1}^\infty \text{ ist konvergent in } \mathbb{K} \right\}$$

$$\Leftrightarrow \left\{ \forall i \in \mathbb{Z}_1^d : (x_i^{(n)})_{n=1}^\infty \text{ ist eine Cauchy-Folge in } \mathbb{K} \right\}$$

$$\Leftrightarrow \left\{ (x^{(n)})_{n=1}^\infty \text{ ist eine Cauchy-Folge in } \mathbb{K}^d \right\}, \quad (\text{IX.17})$$

und in diesem Fall gilt

$$\lim_{n \rightarrow \infty} \{x^{(n)}\} = \begin{pmatrix} \lim_{n \rightarrow \infty} \{x_1^{(n)}\} \\ \vdots \\ \lim_{n \rightarrow \infty} \{x_d^{(n)}\} \end{pmatrix}. \quad (\text{IX.18})$$

Bemerkungen und Beispiele.

- Gemäß Lemma IX.3 sind $(\mathbb{R}^d, \|\cdot\|_{\text{eukl}})$ und $(\mathbb{C}^d, \|\cdot\|_{\text{unit}})$ für alle $d \in \mathbb{N}$ reelle bzw. komplexe Banachräume.

IX.2. Topologie und Stetigkeit in normierten Räumen

Definition IX.4. Seien $(X, \|\cdot\|)$ ein normierter Vektorraum über $\mathbb{K}$ und $A \subseteq X$ eine Teilmenge.

(i) Sind $r > 0$ und $x \in X$, so heißt

$$B(x, r) := \{y \in X \mid \|x - y\| < r\} \quad (\text{IX.19})$$

offene Kugel um x vom Radius r .

(ii) Ein Punkt $x \in A$ heißt **innerer Punkt (I.P.)**

$$:\Leftrightarrow \quad \exists r > 0 : \quad B(x, r) \subseteq A. \quad (\text{IX.20})$$

(iii) Ein Punkt $x \in X$ heißt **Häufungspunkt (H.P.) von A**

$$:\Leftrightarrow \quad \forall r > 0 : \quad (B(x, r) \cap A) \setminus \{x\} \neq \emptyset. \quad (\text{IX.21})$$

(iv) Eine Teilmenge $U \subseteq X$ heißt **offen**

$$:\Leftrightarrow \quad \text{Jeder Punkt in } U \text{ ist ein innerer Punkt} \quad (\text{IX.22})$$

$$\Leftrightarrow \quad \forall x \in U \exists r > 0 : \quad B(x, r) \subseteq U. \quad (\text{IX.23})$$

(v) Eine Teilmenge $V \subseteq X$ heißt **abgeschlossen**

$$:\Leftrightarrow \quad V \text{ enthält alle seine Häufungspunkte} \quad (\text{IX.24})$$

$$\Leftrightarrow \quad \left\{ x \text{ ist H.P. von } V \Rightarrow x \in V \right\}. \quad (\text{IX.25})$$

(vi) Eine Teilmenge $K \subseteq X$ heißt **kompakt**

$$:\Leftrightarrow \quad \text{Jede offene Überdeckung von } K \text{ enthält eine endliche offene Überdeckung}$$

$$\begin{aligned} :\Leftrightarrow \quad & \text{Ist } \mathfrak{U} \text{ eine Familie offener Mengen mit } K \subseteq \bigcup_{U \in \mathfrak{U}} U \\ & \Rightarrow \exists U_1, U_2, \dots, U_n \in \mathfrak{U} : \quad K \subseteq U_1 \cup U_2 \cup \dots \cup U_n. \end{aligned} \quad (\text{IX.26})$$

Bemerkungen und Beispiele.

- Ist $(X, \|\cdot\|)$ ein normierter Vektorraum so sind die leere Menge $\emptyset \subseteq X$ und auch $X \subseteq X$ selbst sowohl offen als auch abgeschlossen.
- Sind $(X, \|\cdot\|)$ ein normierter Vektorraum und $A \subseteq X$ eine Teilmenge, so ist $x \in X$ genau dann ein Häufungspunkt von A , wenn es eine konvergente Folge $(x_n)_{n=1}^\infty \in (A \setminus \{x\})^\mathbb{N}$ gibt mit $\lim_{n \rightarrow \infty} x_n = x$.
- Sind $(X, \|\cdot\|)$ ein normierter Vektorraum und $A \subseteq X$ eine Teilmenge, so gilt wie in $\mathbb{K}$ folgende Äquivalenz:

$$\{A \text{ ist offen}\} \Leftrightarrow \{A^c = X \setminus A \text{ ist abgeschlossen}\}. \quad (\text{IX.27})$$

- Für $d \in \mathbb{N}$ und $(X, \|\cdot\|) = (\mathbb{R}^d, \|\cdot\|_{\text{eukl}})$ oder $(X, \|\cdot\|) = (\mathbb{K}^d, \|\cdot\|_{\text{unit}})$ gilt wie für $d = 1$ folgende, als **Satz von Heine-Borel** bekannte Äquivalenz:

$$\{K \subseteq X \text{ ist kompakt}\} \Leftrightarrow \{K \subseteq X \text{ ist beschränkt und abgeschlossen}\}. \quad (\text{IX.28})$$

Definition IX.5. Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ zwei normierte Vektorräume über $\mathbb{K}$ und $M \subseteq X$ und $N \subseteq Y$ zwei Teilmengen.

- (i) Eine Abbildung $f : M \rightarrow N$ heißt **stetig in $x \in M$**

$$:\Leftrightarrow \forall \varepsilon > 0 \exists \delta > 0 : f[B_X(x, \delta)] \subseteq B_Y(f[x], \varepsilon) \quad (\text{IX.29})$$

$$\Leftrightarrow \forall \varepsilon > 0 \exists \delta > 0 \forall x' \in M : \|x' - x\|_X < \delta \Rightarrow \|f(x') - f(x)\|_Y < \varepsilon.$$

Ist f stetig in x , für alle $x \in M$, so nennen wir f **stetig**.

- (ii) Eine Abbildung $f : M \rightarrow N$ heißt **folgenstetig in $x \in M$**

$$:\Leftrightarrow \forall (x_n)_{n=1}^\infty \in M^\mathbb{N} : x = \lim_{n \rightarrow \infty} \{x_n\} \Rightarrow f(x) = \lim_{n \rightarrow \infty} \{f(x_n)\}. \quad (\text{IX.30})$$

Ist f folgenstetig in x , für alle $x \in M$, so nennen wir f **folgenstetig**.

Satz IX.6. Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ zwei normierte Vektorräume über $\mathbb{K}$, $M \subseteq X$ und $N \subseteq Y$ zwei Teilmengen, $f : M \rightarrow N$ und $x \in M$. Dann gelten folgende Aussagen.

$$(i) \quad f \text{ ist stetig in } x \Leftrightarrow f \text{ ist folgenstetig in } x. \quad (\text{IX.31})$$

$$(ii) \quad f \text{ ist stetig} \Leftrightarrow f \text{ ist folgenstetig} \quad (\text{IX.32})$$

$$\Leftrightarrow \forall U \subseteq N : U \text{ ist offen} \Rightarrow f^{-1}(U) \text{ ist offen} \quad (\text{IX.33})$$

$$\Leftrightarrow \forall V \subseteq N : V \text{ ist abgeschlossen} \Rightarrow f^{-1}(V) \text{ ist abgeschlossen}. \quad (\text{IX.34})$$

(iii) Sind $f : M \rightarrow N$ stetig und $M \subseteq X$ kompakt, so ist auch $f(M) \subseteq Y$ kompakt.

(ii) Sind $f : M \rightarrow \mathbb{R}$ stetig und $M \subseteq X$ kompakt, so nimmt f auf M sein Minimum und Maximum an, d.h. es existieren $x_{\min}, x_{\max} \in M$ so, dass

$$\forall x \in M : f(x_{\min}) \leq f(x) \leq f(x_{\max}). \quad (\text{IX.35})$$

IX.3. Ergänzungen

IX.3.1. Vektorräume und lineare Abbildungen

Die *Lineare Algebra* ist die Theorie der Vektorräume und der linearen Abbildungen, die strukturerhaltenden Abbildungen zwischen Vektorräumen. In diesem Abschnitt führen wir einige Begriffe der linearen Algebra ein, ohne die wir die Inhalte aus *Analysis 2* nicht formulieren können und die deshalb unabdingbar sind. Wir beschränken uns hier auf Vektorräume über die Körper $\mathbb{R}$ der reellen Zahlen und $\mathbb{C}$ der komplexen Zahlen, die wir wie schon in *Analysis 1* einheitlich mit $\mathbb{K}$ bezeichnen, sofern die sich darauf beziehende Betrachtung für beide Fälle durchgeführt werden kann.

Definition IX.7. Eine Menge X heißt **Vektorraum über $\mathbb{K}$** oder **$\mathbb{K}$ -Vektorraum** $:\Leftrightarrow$

Auf X sind Addition $(+): X \times X \rightarrow X$ und Multiplikation mit einem Skalar $(\cdot): \mathbb{K} \times X \rightarrow X$ definiert, die die folgenden Eigenschaften besitzen:

$$(i) \quad \forall \vec{a}, \vec{b}, \vec{c} \in X: \quad \vec{a} + \vec{b} = \vec{b} + \vec{a}, \quad (\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c}), \quad (\text{IX.36})$$

$$(ii) \quad \exists \vec{0} \in X \quad \forall \vec{a} \in X: \quad \vec{a} = \vec{a} + \vec{0} = \vec{0} + \vec{a}, \quad (\text{IX.37})$$

$$(iii) \quad \forall \vec{a} \in X \quad \exists (-\vec{a}) \in X: \quad \vec{a} + (-\vec{a}) = (-\vec{a}) + \vec{a} = \vec{0}, \quad (\text{IX.38})$$

$$(iv) \quad \forall \vec{a}, \vec{b} \in X, \alpha, \beta \in \mathbb{K}: \quad \begin{cases} \alpha \cdot (\vec{a} + \vec{b}) = \alpha \cdot \vec{a} + \beta \cdot \vec{b}, \\ (\alpha + \beta) \cdot \vec{a} = \alpha \cdot \vec{a} + \beta \cdot \vec{a}, \end{cases} \quad (\text{IX.39})$$

$$(v) \quad \forall \vec{a} \in X, \alpha, \beta \in \mathbb{K}: \quad 1 \cdot \vec{a} = \vec{a}, \quad \alpha(\beta \vec{a}) = (\alpha\beta)\vec{a}. \quad (\text{IX.40})$$

Die Elemente von X heißen **Vektoren**, $\vec{0} \in X$ ist der **Nullvektor**. Eine Teilmenge $Z \subseteq X$, die selbst ein Vektorraum ist, bezeichnet man als **Teilraum** oder **Unterraum** von X . Vektorräume über $\mathbb{R}$ bezeichnet man auch als **reelle Vektorräume**, Vektorräume über $\mathbb{C}$ als **komplexe Vektorräume**.

Bemerkungen und Beispiele.

- Gleichungen (IX.36), (IX.37) und (IX.38) besagen zusammen, dass X bezüglich der Addition $(+)$ eine abelsche Gruppe bildet.
- $\{\vec{0}\} \subseteq X$ ist der kleinste und $X \subseteq X$ der größte Unterraum von X .
- Sind $n \in \mathbb{N}$, $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K}$ und $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n \in X$ Vektoren, so bezeichnet man Summen der Form $\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \dots + \alpha_n \vec{x}_n \in X$ als **Linearkombination** der $\vec{x}_j$ und die Zahlen α_j in diesem Zusammenhang als ihre **Koeffizienten**.
- *Multiplikation mit einem Skalar* und *Skalarprodukt* ist nicht dasselbe.
- $\mathbb{K}$ ist ein Vektorraum über $\mathbb{K}$. (Dieses Beispiel ist allerdings etwas künstlich.)

- Für $d \in \mathbb{N}$ ist

$$\mathbb{K}^d = \left\{ \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} \mid x_1, x_2, \dots, x_d \in \mathbb{K} \right\} \quad (\text{IX.41})$$

bezüglich komponentenweiser Addition und Multiplikation mit einem Skalar,

$$\gamma \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix} := \begin{pmatrix} \gamma x_1 + y_1 \\ \vdots \\ \gamma x_d + y_d \end{pmatrix}, \quad (\text{IX.42})$$

ein Vektorraum über $\mathbb{K}$. Dabei sind

$$0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{und} \quad - \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} = \begin{pmatrix} -x_1 \\ \vdots \\ -x_d \end{pmatrix}. \quad (\text{IX.43})$$

- Für $d \in \mathbb{N}$ definieren wir die **kanonischen Basisvektoren** durch

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \dots, \quad \vec{e}_d = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \quad (\text{IX.44})$$

Die Menge $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_d\} \subseteq \mathbb{K}^d$ bezeichnet man auch als **Standardbasis**.

- Mit der Standardbasis wird

$$\vec{x} := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + \dots + x_d \vec{e}_d. \quad (\text{IX.45})$$

d.h. jeder Vektor kann als Linearkombination der d kanonischen Basisvektoren $\vec{e}_1, \dots, \vec{e}_d \in \mathbb{K}^d$ mit seinen Komponenten $x_1, \dots, x_d \in \mathbb{K}$ als Koeffizienten geschrieben werden.

- Die Darstellung (IX.45) ist eindeutig in dem Sinne, dass für $\alpha_1, \beta_1, \alpha_2, \beta_2, \dots, \alpha_d, \beta_d \in \mathbb{K}$ aus $\alpha_1 \vec{e}_1 + \alpha_2 \vec{e}_2 + \dots + \alpha_d \vec{e}_d = \beta_1 \vec{e}_1 + \beta_2 \vec{e}_2 + \dots + \beta_d \vec{e}_d$ auch $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \dots, \alpha_d = \beta_d$ folgen.
- Für $d = 1, 2, 3$ gewinnt man durch Zeichnungen eine gewisse Vorstellung von den Vektoren in $\mathbb{R}^d$.
- Schränkt man bei den komplexen Zahlen die Multiplikation auf die reellen Elemente ein, so ist $\mathbb{C} = \mathbb{R}^2$ ein Vektorraum über $\mathbb{R}$ (der Dimension 2).

- Ein weiteres Beispiel eines Vektorraums ist die Menge der reellen Funktionen auf $[\alpha, \beta]$, $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$,

$$\mathcal{F} := \{f : [\alpha, \beta] \rightarrow \mathbb{R}\}. \quad (\text{IX.46})$$

Mit $f, g \in \mathcal{F}$ und $\gamma \in \mathbb{R}$ wird $\mathcal{F}$ durch

$$\forall x \in [\alpha, \beta] : (f + \gamma \cdot g)(x) := f(x) + \gamma \cdot g(x) \quad (\text{IX.47})$$

zum Vektorraum (“punktweise Addition und Multiplikation mit einem Skalar”). Dabei sind

$$\forall x \in [\alpha, \beta] : 0(x) := 0, \quad (-f)(x) := -f(x). \quad (\text{IX.48})$$

Für diesen Vektorraum $\mathcal{F}$ liefert eine Zeichnung eine gewisse Anschauung.

- Das obige Beispiel lässt sich noch wie folgt verallgemeinern. Sind $A \neq \emptyset$ eine Menge und X ein Vektorraum über $\mathbb{K}$, so wird durch punktweise Addition und punktweise Multiplikation mit einem Skalar wie in (IX.47) auch der Funktionenraum

$$\mathcal{F}' := \{f : A \rightarrow X\} \quad (\text{IX.49})$$

zum Vektorraum über $\mathbb{K}$.

- Insbesondere ist $\mathcal{F}'$ für $A := \mathbb{Z}_1^d$ und $X := \mathbb{K}^1 = \mathbb{K}$ gleich (genauer: *isomorph zu*) $\mathbb{K}^d$,

$$\mathbb{K}^d = \{\vec{x} : \mathbb{Z}_1^d \rightarrow \mathbb{K}, \nu \mapsto x_\nu\}. \quad (\text{IX.50})$$

- Die Notation für Vektoren mit Vektorpfeilen ist etwas schwerfällig, und wir schreiben im Folgenden einfach $x, y, e_m^{(X)}$ und $e_n^{(Y)}$ statt $\vec{x}, \vec{y}, \vec{e}_m^{(X)}$ und $\vec{e}_n^{(Y)}$.

Definition IX.8. Seien X und Y zwei $\mathbb{K}$ -Vektorräume. Eine Abbildung $A : X \rightarrow Y$ heißt **linear** oder **(linearer) Operator**

$$:\Leftrightarrow \forall x, y \in X, \alpha \in \mathbb{K} : A(\alpha x + y) = \alpha A(x) + A(y). \quad (\text{IX.51})$$

Die Menge der linearen Abbildungen $X \rightarrow Y$ bezeichnen wir mit $\mathcal{L}(X; Y)$.

Bemerkungen und Beispiele.

- Die Menge $\mathcal{L}(X; Y)$ der linearen Abbildungen ist selbst ein $\mathbb{K}$ -Vektorraum bezüglich punktweiser Operationen, d.h. für $\alpha \in \mathbb{K}$ und $A, B \in \mathcal{L}(X; Y)$ ist $(\alpha A + B) \in \mathcal{L}(X; Y)$ definiert durch

$$\forall x \in X : (\alpha A + B)(x) := \alpha A(x) + B(x). \quad (\text{IX.52})$$

- Durch vollständige Induktion erhält man aus (IX.51) leicht, dass

$$A\left(\sum_{i=1}^k \alpha_i x_i\right) = \sum_{i=1}^k \alpha_i A(x_i), \quad (\text{IX.53})$$

für alle $x_1, \dots, x_k \in X$ und $\alpha_1, \dots, \alpha_k \in \mathbb{K}$.

- $A(0_X) = A(0_{\mathbb{K}} \cdot 0_X) = 0_{\mathbb{K}} \cdot A(0_X) = 0_Y$.
- Seien $M, N \in \mathbb{N}$ und $X = \mathbb{K}^M$ und $Y = \mathbb{K}^N$. Wir bezeichnen mit

$$e_1^{(X)}, e_2^{(X)}, \dots, e_M^{(X)} \subseteq X \quad \text{und} \quad e_1^{(Y)}, e_2^{(Y)}, \dots, e_N^{(Y)} \subseteq Y \quad (\text{IX.54})$$

die jeweiligen Standardbasen in $\mathbb{K}^M$ und $\mathbb{K}^N$.

- Sei $A \in \mathcal{L}(\mathbb{K}^M; \mathbb{K}^N)$ eine lineare Abbildung. Für jeden Vektor $x \in \mathbb{K}^M$ lässt sich der zugehörige Bildvektor $A(x) \in \mathbb{K}^N$ in eindeutiger Weise als Linearkombination $A(x) = \gamma_1 e_1^{(Y)} + \gamma_2 e_2^{(Y)} + \dots + \gamma_N e_N^{(Y)}$ der kanonischen Basisvektoren in $\mathbb{K}^N$ darstellen. Insbesondere gibt es $N \cdot M$ eindeutige Zahlen $\{a_{n,m} | n \in \mathbb{Z}_1^N, m \in \mathbb{Z}_1^M\}$ so, dass

$$\begin{aligned} A(e_1^{(X)}) &= a_{1,1} e_1^{(Y)} + a_{2,1} e_2^{(Y)} + \dots + a_{N,1} e_N^{(Y)}, \\ A(e_2^{(X)}) &= a_{1,2} e_1^{(Y)} + a_{2,2} e_2^{(Y)} + \dots + a_{N,2} e_N^{(Y)}, \\ &\vdots \\ A(e_M^{(X)}) &= a_{1,M} e_1^{(Y)} + a_{2,M} e_2^{(Y)} + \dots + a_{N,M} e_N^{(Y)} \end{aligned} \quad (\text{IX.55})$$

gelten.

- Für

$$x = x_1 e_1^{(X)} + x_2 e_2^{(X)} + \dots + x_M e_M^{(X)} \in X \quad (\text{IX.56})$$

ist dann

$$A(x) = y_1 e_1^{(Y)} + y_2 e_2^{(Y)} + \dots + y_N e_N^{(Y)} \in Y, \quad (\text{IX.57})$$

wobei

$$\forall n \in \mathbb{Z}_1^N : \quad y_n = \sum_{m=1}^M a_{n,m} x_m. \quad (\text{IX.58})$$

- Die zu A gehörige Zahlentabelle

$$\mathcal{M}[A] := \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,M} \\ a_{2,1} & a_{2,2} & \dots & a_{2,M} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} & \dots & a_{N,M} \end{pmatrix} \in \mathbb{K}^{M \cdot N} \quad (\text{IX.59})$$

bezeichnet man als Matrix und genauer als Matrixdarstellung von A bezüglich der Standardbasen in $\mathbb{K}^M$ und $\mathbb{K}^N$.

- Für $M, N \in \mathbb{N}$ ist die Matrixdarstellung $\mathcal{M} : \mathcal{L}(\mathbb{K}^M; \mathbb{K}^N) \rightarrow \mathbb{K}^{M \cdot N}$ eine lineare Bijektion zwischen den Vektorräumen $\mathcal{L}(\mathbb{K}^M; \mathbb{K}^N)$ und $\mathbb{K}^{M \cdot N}$. D.h. $\mathcal{M}$ ist bijektiv und es gilt $\mathcal{M}[\alpha A + B] = \alpha \mathcal{M}[A] + \mathcal{M}[B]$ für alle $\alpha \in \mathbb{K}$ und $A, B \in \mathcal{L}(\mathbb{K}^M; \mathbb{K}^N)$.

- Konkret bedeutet dies, dass man die Matrixdarstellung als riesigen Spaltenvektor interpretiert, den man aus Platzmangel in eine Tabelle mit N Zeilen und M Spalten schreibt. Dann ergibt sich

$$\begin{aligned} \mathcal{M}[A] &= \begin{pmatrix} a_{1,1} & \dots & a_{1,M} \\ \vdots & & \vdots \\ a_{N,1} & \dots & a_{N,M} \end{pmatrix}, \quad \mathcal{M}[B] = \begin{pmatrix} b_{1,1} & \dots & b_{1,M} \\ \vdots & & \vdots \\ b_{N,1} & \dots & b_{N,M} \end{pmatrix} \\ \Rightarrow \mathcal{M}[\alpha A + B] &= \begin{pmatrix} \alpha a_{1,1} + b_{1,1} & \dots & \alpha a_{1,M} + b_{1,M} \\ \vdots & & \vdots \\ \alpha a_{N,1} + b_{N,1} & \dots & \alpha a_{N,M} + b_{N,M} \end{pmatrix}. \end{aligned} \quad (\text{IX.60})$$

- Obwohl dies aus Sicht der linearen Algebra etwas ungenau ist, werden wir im Weiteren nicht mehr zwischen $\mathcal{L}(\mathbb{K}^M; \mathbb{K}^N)$ und $\mathbb{K}^{M \times N}$ unterscheiden und die lineare Abbildung A mit ihrer Matrixdarstellung gleichsetzen,

$$A := \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,M} \\ a_{2,1} & a_{2,2} & \dots & a_{2,M} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} & \dots & a_{N,M} \end{pmatrix}. \quad (\text{IX.61})$$

- Sind X, Y, Z drei $\mathbb{K}$ -Vektorräume und $A \in \mathcal{L}(Y; Z)$ sowie $B \in \mathcal{L}(X; Y)$, so ist auch deren Komposition linear, $(A \circ B) \in \mathcal{L}(X; Z)$.
- Sind $M, N, P \in \mathbb{N}$ und $X = \mathbb{K}^M$, $Y = \mathbb{K}^N$, $Z = \mathbb{K}^P$, so ist für $A \in \mathcal{L}(Y; Z)$ und $B \in \mathcal{L}(X; Y)$ die Matrixdarstellung von $(A \circ B) \in \mathcal{L}(X; Z)$ durch das *Matrixprodukt* der Matrixdarstellungen von A und B gegeben.

IX.3.2. Inneres und Abschluss von Mengen

Durch Nachprüfen der Definition IX.4 erhält man leicht folgende Aussage.

Lemma IX.9. *Sei $(X, \|\cdot\|)$ ein normierter Vektorraum.*

- (i) *Sind I eine Indexmenge und $\{U_i\}_{i \in I} \subseteq \mathfrak{P}(X)$ eine Familie offener Mengen, so ist auch deren Vereinigung $\bigcup_{i \in I} U_i$ offen.*
- (ii) *Sind $m \in \mathbb{N}$ und $U_1, U_2, \dots, U_m \subseteq X$ endlich viele offene Teilmengen von X , so ist auch ihr Durchschnitt $U_1 \cap U_2 \cap \dots \cap U_m$ offen.*
- (iii) *Sind J eine Indexmenge und $\{V_j\}_{j \in J} \subseteq \mathfrak{P}(X)$ eine Familie abgeschlossener Mengen, so ist auch deren Durchschnitt $\bigcap_{j \in J} V_j$ abgeschlossen.*
- (iv) *Sind $n \in \mathbb{N}$ und $V_1, V_2, \dots, V_n \subseteq X$ endlich viele abgeschlossene Teilmengen von X , so ist auch ihre Vereinigung $V_1 \cup V_2 \cup \dots \cup V_n$ abgeschlossen.*

Definition IX.10. Seien $(X, \|\cdot\|)$ ein normierter Vektorraum und $A \subseteq X$ eine Teilmenge.

(i) Die Vereinigung

$$A^\circ := \bigcup \left\{ U \mid U \subseteq A, U \text{ offen} \right\} \quad (\text{IX.62})$$

aller offenen Teilmengen von A heißt **Inneres von A** .

(ii) Der Durchschnitt

$$\overline{A} := \bigcap \left\{ V \mid A \subseteq V \subseteq X, V \text{ abgeschlossen} \right\} \quad (\text{IX.63})$$

aller abgeschlossenen Obermengen A heißt **Abschluss von A** .

Bemerkungen und Beispiele.

- Nach Lemma IX.9 (i) ist das Innere A° einer Teilmenge $A \subseteq X$ eines normierten Vektorraums $(X, \|\cdot\|)$ stets offen.
- Analog ist nach Lemma IX.9 (iii) der Abschluss $\overline{A}$ einer Teilmenge $A \subseteq X$ stets abgeschlossen.
- Das Innere A° einer Teilmenge $A \subseteq X$ ist die größte offene Teilmenge von A , d.h. für jede offene Teilmenge $B \subseteq A$ folgt $B \subseteq A^\circ$.
- Außerdem kann A° als die Menge aller inneren Punkte von A charakterisiert werden,

$$A^\circ = \{x \in A \mid x \text{ ist I.P. von } A\}. \quad (\text{IX.64})$$

- Der Abschluss $\overline{A}$ einer Teilmenge $A \subseteq X$ ist die kleinste abgeschlossene Obermenge von A , d.h. für jede abgeschlossene Obermenge $C \supseteq A$ folgt $C \supseteq \overline{A}$.
- Außerdem kann $\overline{A}$ als die Menge aller Punkte aus A und aller Häufungspunkte von A charakterisiert werden,

$$\overline{A} = \{x \in X \mid x \in A \text{ oder } x \text{ ist H.P. von } A\}. \quad (\text{IX.65})$$

X. Differentiation vektorwertiger Funktionen

X.1. Partielle und stetige Differenzierbarkeit

Definition X.1. Seien $M \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$ offen, $x \in U$, $F : U \rightarrow \mathbb{R}$ und $j \in \mathbb{Z}_1^M$.

(i) Die Funktion F heißt **bei $x \in U$ partiell nach x_j differenzierbar**

$$:\Leftrightarrow \quad \frac{\partial F(x)}{\partial x_j} := \lim_{t \rightarrow 0} \left\{ \frac{F(x + te_j) - F(x)}{t} \right\} \text{ existiert,} \quad (\text{X.1})$$

wobei $e_j = (0, \dots, 0, 1, 0, \dots, 0)^T \in \mathbb{R}^M$ der j . kanonische Basisvektor ist. In diesem Fall heißt $\frac{\partial F(x)}{\partial x_j} =: \partial_{x_j} F(x)$ **partielle Ableitung von F nach x_j bei x .**

(ii) Die Funktion F heißt **bei $x \in U$ partiell differenzierbar**

$$:\Leftrightarrow \quad \forall j \in \mathbb{Z}_1^M : \quad F \text{ ist bei } x \text{ partiell nach } x_j \text{ differenzierbar.} \quad (\text{X.2})$$

In diesem Fall heißt der reelle Vektor

$$\nabla F(x) := \begin{pmatrix} \frac{\partial F(x)}{\partial x_1} \\ \frac{\partial F(x)}{\partial x_2} \\ \vdots \\ \frac{\partial F(x)}{\partial x_M} \end{pmatrix} \in \mathbb{R}^M \quad (\text{X.3})$$

der partiellen Ableitungen **Gradient von F bei x .**

(iii) Die Funktion F heißt **auf U partiell differenzierbar**

$$:\Leftrightarrow \quad \forall x \in U : \quad F \text{ ist bei } x \text{ partiell differenzierbar.} \quad (\text{X.4})$$

Bemerkungen und Beispiele.

- Da U eine offene Menge ist, ist $x \in U$ ein innerer Punkt, und es gibt $r > 0$, sodass $B_{\text{eukl}}(x, r) \subseteq U$. Diese Eigenschaft sichert, dass $F(x + te_j)$ und damit der Differenzenquotient in (X.1) für $t \in (-r, r)$ überhaupt definiert ist. Dies unterstreicht die Bedeutung des topologischen Begriffs der *offenen Menge*.
- Die partielle Ableitung einer Funktion F nach einer Variablen x_j bedeutet konkret, dass man $F(x_1, x_2, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_M)$ nach x_j differenziert und alle anderen Variablen, $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_M$, als Parameter betrachtet.
- Wir betrachten $U = \mathbb{R}^3$ und $F(x) := x_1 x_2^2 + 10 x_3^5 x_2$, wobei $x = (x_1, x_2, x_3)^T$. Dann sind

$$\frac{\partial F(x)}{\partial x_1} = x_2^2, \quad \frac{\partial F(x)}{\partial x_2} = 2x_1 x_2 + 10 x_3^5, \quad \frac{\partial F(x)}{\partial x_3} = 50 x_3^4 x_2, \quad (\text{X.5})$$

die partiellen Ableitungen, und der Gradient von F bei x ist

$$\nabla F(x) := \begin{pmatrix} x_2^2 \\ 2x_1 x_2 + 10 x_3^5 \\ 50 x_3^4 x_2 \end{pmatrix} \in \mathbb{R}^3. \quad (\text{X.6})$$

- Wir betrachten $U = \mathbb{R}^2$ und $F(x) := \sin(x_1^2 - x_2^2)$, wobei $x = (x_1, x_2)^T$. Dann sind

$$\frac{\partial F(x)}{\partial x_1} = 2x_1 \cos(x_1^2 - x_2^2), \quad \frac{\partial F(x)}{\partial x_2} = -2x_2 \cos(x_1^2 - x_2^2), \quad (\text{X.7})$$

die partiellen Ableitungen, und der Gradient von F bei x ist

$$\nabla F(x) := \begin{pmatrix} 2x_1 \cos(x_1^2 - x_2^2) \\ -2x_2 \cos(x_1^2 - x_2^2) \end{pmatrix} \in \mathbb{R}^2. \quad (\text{X.8})$$

Definition X.2. Seien $M, N \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$ offen, $V \subseteq \mathbb{R}^N$ offen und $F = (F_1, F_2, \dots, F_N) : U \rightarrow V$.

(i) F ist **bei $x \in U$ partiell differenzierbar**

$$:\Leftrightarrow \quad \forall i \in \mathbb{Z}_1^N : \quad F_i \text{ ist bei } x \text{ partiell differenzierbar.} \quad (\text{X.9})$$

(ii) F ist **auf U partiell differenzierbar**

$$:\Leftrightarrow \quad \forall x \in U : \quad F \text{ ist bei } x \text{ partiell differenzierbar.} \quad (\text{X.10})$$

In diesem Fall heißt die reelle $N \times M$ -Matrix

$$J_F(x) := \begin{pmatrix} \frac{\partial F_1(x)}{\partial x_1} & \cdots & \frac{\partial F_1(x)}{\partial x_M} \\ \vdots & & \vdots \\ \frac{\partial F_N(x)}{\partial x_1} & \cdots & \frac{\partial F_N(x)}{\partial x_M} \end{pmatrix} \in \mathfrak{M}_{N \times M}(\mathbb{R}) \quad (\text{X.11})$$

der partiellen Ableitungen **Jacobimatrix von F bei x** .

Bemerkungen und Beispiele.

- Wir betrachten $U = \mathbb{R}^+ \times (-\pi, \pi) \ni (r, \varphi)$ und $F : U \rightarrow \mathbb{R}^3$,

$$F(r, \varphi) := \begin{pmatrix} F_1(r, \varphi) \\ F_2(r, \varphi) \end{pmatrix} := \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{pmatrix}. \quad (\text{X.12})$$

Dann ist

$$J_F(x) = \begin{pmatrix} \frac{\partial F_1(r, \varphi)}{\partial r} & \frac{\partial F_1(r, \varphi)}{\partial \varphi} \\ \frac{\partial F_2(r, \varphi)}{\partial r} & \frac{\partial F_2(r, \varphi)}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos(\varphi) & -r \sin(\varphi) \\ \sin(\varphi) & r \cos(\varphi) \end{pmatrix}. \quad (\text{X.13})$$

Definition X.3. Seien $M, N \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$, $V \subseteq \mathbb{R}^N$ offene Teilmengen und $F : U \rightarrow V$.

- (i) F heißt **stetig differenzierbar auf U** $\Leftrightarrow$

$$F \text{ ist partiell differenzierbar auf } U \text{ und } \forall i \in \mathbb{Z}_1^N, j \in \mathbb{Z}_1^M : \frac{\partial F_i}{\partial x_j} \text{ ist stetig auf } U. \quad (\text{X.14})$$

Die Klasse der auf U stetig differenzierbaren Funktionen mit Werten in V bezeichnet man mit

$$C^1(U; V) := \{F : U \rightarrow V \mid F \text{ ist stetig differenzierbar auf } U\}. \quad (\text{X.15})$$

- (ii) Sei $k \in \mathbb{N}$. F heißt **k -mal stetig differenzierbar auf U** $\Leftrightarrow$

F ist stetig differenzierbar auf U und

$$\forall j_1 \in \mathbb{Z}_1^M : \frac{\partial F}{\partial x_{j_1}} \text{ ist stetig differenzierbar auf } U \text{ und} \quad (\text{X.16})$$

$$\forall j_1, j_2 \in \mathbb{Z}_1^M : \frac{\partial}{\partial x_{j_2}} \left(\frac{\partial F}{\partial x_{j_1}} \right) \text{ ist stetig differenzierbar auf } U \text{ und}$$

$\vdots$

$$\forall j_1, \dots, j_{k-1} \in \mathbb{Z}_1^M : \frac{\partial}{\partial x_{j_{k-1}}} \left(\dots \frac{\partial}{\partial x_{j_2}} \left(\frac{\partial F}{\partial x_{j_1}} \right) \right) \text{ ist stetig differenzierbar auf } U.$$

In diesem Fall bezeichnet

$$\frac{\partial^k F(x)}{\partial x_{j_k} \partial x_{j_{k-1}} \dots \partial x_{j_1}} := \frac{\partial}{\partial x_{j_k}} \left[\frac{\partial}{\partial x_{j_{k-1}}} \left(\dots \left(\frac{\partial}{\partial x_{j_1}} F \right) \right) \right](x) \quad (\text{X.17})$$

die **k . partielle Ableitung von F nach $x_{j_1}, \dots, x_{j_k}$** . Die Klasse der auf U k -mal stetig differenzierbaren Funktionen mit Werten in V bezeichnet man mit

$$C^k(U; V) := \{F : U \rightarrow V \mid F \text{ ist } k\text{-mal stetig differenzierbar auf } U\}. \quad (\text{X.18})$$

- (iii) Die Klasse der **unendlich oft (stetig) differenzierbaren** oder auch **glatten Funktionen auf U mit Werten in V** wird mit

$$C^\infty(U; V) := \bigcap_{k=1}^{\infty} C^k(U; V) \quad (\text{X.19})$$

bezeichnet, und

$$\begin{aligned} C_0^\infty(\mathbb{R}^M; V) &:= \{F \in C^\infty(\mathbb{R}^M; V) \mid \text{supp}(F) \text{ ist kompakt}\} \\ &= \{F \in C^\infty(\mathbb{R}^M; V) \mid \exists R < \infty \ \forall \vec{x} \in \mathbb{R}^M, |\vec{x}| \geq R: F(\vec{x}) = 0\} \end{aligned} \quad (\text{X.20})$$

sind die **unendlich oft differenzierbaren Funktionen mit kompaktem Träger**, wobei **Träger** $\text{supp}(F)$ **von F** die Menge $\text{supp}(F) := \{x \in \mathbb{R}^M \mid F(x) \neq 0\}$ bezeichnet, auf der F nicht verschwindet.

Bemerkungen und Beispiele.

- Wir betrachten abermals $U = \mathbb{R}^3$, $N = 1$, $V = \mathbb{R}$ und $F(x) := x_1 x_2^2 + 10 x_3^5 x_2$, wobei $x = (x_1, x_2, x_3)^T$. Dann sind

$$\frac{\partial F(x)}{\partial x_2} = 2x_1 x_2 + 10 x_3^5, \quad \frac{\partial^2 F(x)}{\partial x_1 \partial x_2} = 2x_2, \quad \frac{\partial^2 F(x)}{\partial x_1 \partial x_1 \partial x_2} = 0, \quad (\text{X.21})$$

$$\frac{\partial F(x)}{\partial x_1} = x_2^2, \quad \frac{\partial^2 F(x)}{\partial x_2 \partial x_1} = 2x_2, \quad \frac{\partial^2 F(x)}{\partial x_1 \partial x_2 \partial x_1} = 0. \quad (\text{X.22})$$

Der folgende Satz zeigt, dass stetige Differenzierbarkeit die Vertauschbarkeit der partiellen Ableitungen garantiert.

Satz X.4. Seien $M, k \in \mathbb{N}$, $k \geq 2$, $U \subseteq \mathbb{R}^M$ und $V \subseteq \mathbb{R}$ offene Teilmengen und $F \in C^k(U; V)$. Seien weiterhin $j_1, j_2, \dots, j_k \in \mathbb{Z}_1^M$ und sei $\pi \in \mathcal{S}_k$ eine Permutation von k Elementen. Dann gilt für jedes $x \in U$ dass

$$\frac{\partial^k F(x)}{\partial x_{j_1} \partial x_{j_2} \cdots \partial x_{j_k}} = \frac{\partial^k F(x)}{\partial x_{j_{\pi(1)}} \partial x_{j_{\pi(2)}} \cdots \partial x_{j_{\pi(k)}}}. \quad (\text{X.23})$$

Beweis. Wir zeigen (X.23) zunächst für $k = 2$. Seien dazu $F \in C^2(U; V)$ sowie $\alpha, \beta \in \mathbb{Z}_1^M$ mit $\alpha < \beta$. Dann ist

$$\begin{aligned} \frac{\partial^2 F(x)}{\partial x_\alpha \partial x_\beta} &= \lim_{t \rightarrow 0} \left\{ \frac{1}{t} \left(\frac{\partial F}{\partial x_\beta}(x + te_\alpha) - \frac{\partial F}{\partial x_\beta}(x) \right) \right\} \\ &= \lim_{t \rightarrow 0} \lim_{s \rightarrow 0} \left\{ \frac{1}{st} \left[F(x + te_\alpha + se_\beta) - F(x + te_\alpha) - F(x + se_\beta) + F(x) \right] \right\} \\ &= \lim_{t \rightarrow 0} \lim_{s \rightarrow 0} \{G(t, s)\}, \end{aligned} \quad (\text{X.24})$$

und mit Vertauschen von α und β auch $\frac{\partial^2 F(x)}{\partial x_\beta \partial x_\alpha} = \lim_{s \rightarrow 0} \lim_{t \rightarrow 0} \{G(t, s)\}$, wobei

$$G(t, s) := \frac{1}{st} \left[F(x + te_\alpha + se_\beta) - F(x + se_\beta) - F(x + te_\alpha) + F(x) \right], \quad (\text{X.25})$$

für festes $x \in U$. Wir definieren weiterhin, für festes $s \neq 0$,

$$\tilde{G}(t) := F(x + te_\alpha + se_\beta) - F(x + te_\alpha), \quad (\text{X.26})$$

so dass,

$$\begin{aligned} G(t, s) &= \frac{1}{st} [\tilde{G}(t) - \tilde{G}(0)] = \frac{1}{s} \frac{d\tilde{G}(\theta_t t)}{dt} \\ &= \frac{1}{s} \left[\frac{\partial F}{\partial x_\alpha}(x + \theta_t te_\alpha + se_\beta) - \frac{\partial F}{\partial x_\alpha}(x + \theta_t te_\alpha) \right] \\ &= \frac{\partial^2 F}{\partial x_\beta \partial x_\alpha}(x + \theta_t te_\alpha + \theta_s se_\beta), \end{aligned} \quad (\text{X.27})$$

nach zweimaliger Anwendung des Mittelwertsatzes der Differenzialrechnung, wobei $\theta_t, \theta_s \in (0, 1)$. Gleichung (X.27) zeigt, dass $G(t, s)$ stetig fortsetzbar in (t, s) bei $(0, 0)$ ist, also gleichzeitig in $t \rightarrow 0$ und $s \rightarrow 0$ und unabhängig von der Reihenfolge der Limiten, da $F \in C^2(U; V)$. Also ist

$$\frac{\partial^2 F(x)}{\partial x_\alpha \partial x_\beta} = \lim_{t \rightarrow 0} \lim_{s \rightarrow 0} \{G(t, s)\} = \lim_{(t, s) \rightarrow (0, 0)} G(t, s) = \lim_{s \rightarrow 0} \lim_{t \rightarrow 0} \{G(t, s)\} = \frac{\partial^2 F(x)}{\partial x_\beta \partial x_\alpha}. \quad (\text{X.28})$$

Seien nun $k \geq 3$, $F \in C^k(U; V)$ und $1 \leq \ell \leq k - 1$. Dann ist

$$\hat{F} := \frac{\partial^{k-\ell-1} F}{\partial x_{j_{\ell+2}} \partial x_{j_{\ell+3}} \dots \partial x_{j_k}} \in C^{\ell+1}(U; V) \subseteq C^2(U; V), \quad (\text{X.29})$$

und damit gilt nach (X.28)

$$\frac{\partial^2 \hat{F}}{\partial x_{j_\ell} \partial x_{j_{\ell+1}}} = \frac{\partial^2 \hat{F}}{\partial x_{j_{\ell+1}} \partial x_{j_\ell}}. \quad (\text{X.30})$$

Setzen wir $\hat{F}$ ein, so erhalten wir

$$\frac{\partial^k F}{\partial x_{j_1} \dots \partial x_{j_\ell} \partial x_{j_{\ell+1}} \dots \partial x_{j_k}} = \frac{\partial^k F}{\partial x_{j_1} \dots \partial x_{j_{\ell+1}} \partial x_{j_\ell} \dots \partial x_{j_k}}, \quad (\text{X.31})$$

also die Gültigkeit von (X.23) im Fall, dass π eine Transposition ist. Da sich jede Permutation als Komposition von Transpositionen schreiben lässt, folgt somit (X.23) auch für den allgemeinen Fall $\pi \in \mathcal{S}_K$. $\square$

Bemerkungen und Beispiele.

- Für $k = 2$ und $\alpha, \beta \in \{1, \dots, M\}$ behauptet (X.23), dass

$$\frac{\partial^2 F(x)}{\partial x_\alpha \partial x_\beta} = \frac{\partial^2 F(x)}{\partial x_\beta \partial x_\alpha}. \quad (\text{X.32})$$

- Seien $M = k = 2$ und $f(x_1, x_2) = x_1^3 x_2 + 5x_1 x_2^6$. Dann sind

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 3x_1^2 x_2 + 5x_2^6, \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = x_1^3 + 30x_1 x_2^5, \quad (\text{X.33})$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2) = 3x_1^2 + 30x_2^5 = \frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2}. \quad (\text{X.34})$$

X.2. Totale Differenzierbarkeit

Im Folgenden legen wir stets die euklidische Norm zugrunde und notieren für $x = (x_1, x_2, \dots, x_M)^T \in \mathbb{R}^M$ und $y = (y_1, y_2, \dots, y_N)^T \in \mathbb{R}^N$ mit

$$\|x\|_M := \sqrt{x_1^2 + x_2^2 + \dots + x_M^2} \quad \text{und} \quad \|y\|_N := \sqrt{y_1^2 + y_2^2 + \dots + y_N^2} \quad (\text{X.35})$$

stets die euklidischen Norm von x bzw. y . Die Indizes geben die Dimension des Raums an.

Definition X.5. Seien $M, N \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$ und $V \subseteq \mathbb{R}^N$ zwei nichtleere offene Teilmengen und $x \in U$.

(i) Eine Abbildung $F : U \rightarrow V$ heißt **total differenzierbar bei x** .

$$:\Leftrightarrow \quad \exists A \in \mathbb{R}^{N \times M} : \quad \lim_{z \rightarrow 0} \left\{ \frac{\|F(x+z) - F(x) - Az\|_N}{\|z\|_M} \right\} = 0. \quad (\text{X.36})$$

In diesem Fall heißt die Matrix

$$F'(x) := dF(x) := A \quad (\text{X.37})$$

totale Ableitung von F bei x .

(ii) Die Abbildung $F : U \rightarrow V$ heißt **total differenzierbar auf U**

$$:\Leftrightarrow \quad \forall x \in U : \quad F \text{ ist differenzierbar bei } x. \quad (\text{X.38})$$

In diesem Fall ist $F' : U \rightarrow \mathbb{R}^{N \times M}$.

Bemerkungen und Beispiele.

- Für $M = N = 1$ sind die linearen Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$ reelle 1×1 -Matrizen, also reelle Zahlen. Nach Definition X.5 ist $f : \mathbb{R} \rightarrow \mathbb{R}$ somit genau dann total differenzierbar bei $x \in \mathbb{R}$, wenn es eine Zahl $f'(x) \in \mathbb{R}$ gibt, sodass

$$\lim_{z \rightarrow 0} \left\{ \frac{|f(x+z) - f(x) - f'(x) \cdot z|}{|z|} \right\} = \left| \lim_{z \rightarrow 0} \left\{ \frac{f(x+z) - f(x)}{z} \right\} - f'(x) \right| = 0, \quad (\text{X.39})$$

da der Betrag $|\cdot| : \mathbb{R} \rightarrow \mathbb{R}_0^+$ gerade die euklidische Norm auf den reellen Zahlen $\mathbb{R}$ ist. Dies ist aber offensichtlich gleichwertig mit der Existenz von

$$f'(x) := \lim_{z \rightarrow 0} \left\{ \frac{f(x+z) - f(x)}{z} \right\} \in \mathbb{R}. \quad (\text{X.40})$$

Für $M = N = 1$ fällt also totale Differenzierbarkeit mit dem üblichen Begriff der Differenzierbarkeit einer reellen Funktion einer reellen Variablen zusammen.

- Weiterhin bemerken wir, dass die Ableitung einer total differenzierbaren Funktion eindeutig ist. Gilt nämlich (X.36) sowohl für $A \in \mathbb{R}^{N \times M}$ als auch für $\tilde{A} \in \mathbb{R}^{N \times M}$, so setzen wir $B := A - \tilde{A}$. Sind nun $z \in \mathbb{R}^M \setminus \{0\}$ und $t > 0$, so beobachten wir, dass

$$\frac{\|Bz\|_N}{\|z\|_M} = \frac{\|B(tz)\|_N}{\|tz\|_M} = \lim_{t \rightarrow 0} \frac{\|B(tz)\|_N}{\|tz\|_M}, \quad (\text{X.41})$$

da der zweite Ausdruck gar nicht von t abhängt. Somit ist aber

$$\begin{aligned} \frac{\|B(z)\|_N}{\|z\|_M} &= \lim_{t \rightarrow 0} \left\{ \frac{\|(A - \tilde{A})(tz)\|_N}{\|tz\|_M} \right\} \\ &\leq \lim_{t \rightarrow 0} \left\{ \frac{\|F(x + tz) - F(x) - A(tz)\|_N}{\|tz\|_M} \right\} \\ &\quad + \lim_{t \rightarrow 0} \left\{ \frac{\|F(x + tz) - F(x) - \tilde{A}(tz)\|_N}{\|tz\|_M} \right\} \\ &= 0 \end{aligned} \quad (\text{X.42})$$

also ist $\|Bz\|_N = 0$, für alle $z \in X$, was $Bz = 0$ und daher $B = A - \tilde{A} = 0$ impliziert.

- Wir stellen fest, dass die totale Differenzierbarkeit (X.36) von $F : U \rightarrow V$ bei $x \in U$ auch äquivalent durch

$$\begin{aligned} \exists F'(x) \in \mathcal{B}(X; Y), \quad \delta > 0, \quad r_x : B(0, \delta) \rightarrow Y \quad \forall z \in B(0, \delta) : \\ F(x + z) - F(x) = F'(x)z + r_x(z) \quad \|z\|_M, \quad \lim_{z \rightarrow 0} \|r_x(z)\|_N = 0, \end{aligned} \quad (\text{X.43})$$

dargestellt werden kann.

Lemma X.6. Sind $M, N \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$, $V \subseteq \mathbb{R}^N$ nichtleere offene Teilmengen und $F : U \rightarrow V$ total differenzierbar bei $x \in U$, so ist F auch stetig in $x \in U$.

Satz X.7 (Kettenregel). Seien $M, N, L \in \mathbb{N}$ und $U \subseteq \mathbb{R}^M$, $V \subseteq \mathbb{R}^N$, $W \subseteq \mathbb{R}^L$ nichtleere offene Teilmengen. Seien weiterhin $G : U \rightarrow V$ total differenzierbar bei $x \in U$ und $F : V \rightarrow W$ total differenzierbar bei $y := G(x) \in V$. Dann ist auch $F \circ G : U \rightarrow W$ total differenzierbar bei x , und es gilt

$$(F \circ G)'(x) = F'[G(x)] \circ G'(x). \quad (\text{X.44})$$

Wir stellen nun den Zusammenhang zwischen der partiellen und der totalen Differenzierbarkeit einer Funktion her. Der folgende Satz zeigt, dass diese Begriffe zusammenfallen, wenn man zusätzlich die Stetigkeit der Ableitung fordert.

Satz X.8. Seien $M, N \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$ offen, $V \subseteq \mathbb{R}^N$ offen und $F : U \rightarrow V$. Dann sind folgende Aussagen äquivalent:

$$F \text{ ist total differenzierbar auf } U, \text{ und } F' : U \rightarrow \mathcal{B}(\mathbb{R}^M; \mathbb{R}^N) \text{ ist stetig in } U. \quad (\text{X.45})$$

$$\Leftrightarrow F \in C^1(U; V), \text{ d.h. } F \text{ ist stetig differenzierbar auf } U. \quad (\text{X.46})$$

Gilt (X.45) oder (X.46), so ist

$$\forall x \in U : F'(x) = J_F(x), \quad (\text{X.47})$$

d.h. die Ableitung $F'(x)$ von F ist gegeben durch die Matrix $(\partial_{x_j} F_i)_{i \in \mathbb{Z}_1^N, j \in \mathbb{Z}_1^M}$ der partiellen Ableitungen, der Jacobimatrix $J_F(x)$ von F bei x .

Bemerkungen und Beispiele.

- Satz X.8 zeigt, dass totale und partielle Differenzierbarkeit zusammenfallen, sofern die Ableitungen stetig sind.
- Im Allgemeinen ist jedoch partielle Differenzierbarkeit der schwächere Begriff, und es gibt Funktionen, die partiell, aber nicht total differenzierbar in einem zu betrachteten Punkt $x \in U$ sind. Es gilt also nur

$$F \text{ ist total differenzierbar in } x \in U \Rightarrow F \text{ ist partiell differenzierbar in } x \in U, \quad (\text{X.48})$$

aber nicht notwendig die Umkehrung.

Satz X.8 zeigt auch, dass stetige Differenzierbarkeit auch im allgemeinen Kontext das richtige Konzept ist, und wir definieren:

Definition X.9. Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ zwei Banachräume über $\mathbb{K}$ und $U \subseteq X$ und $V \subseteq Y$ zwei nichtleere, offene Teilmengen. Die Menge der **stetig auf U differenzierbaren Funktionen mit Werten in V** ist definiert als

$$C^1(U; V) := \left\{ F : U \rightarrow V \mid F \text{ ist total differenzierbar auf } U \text{ und } F' \in C[U; \mathcal{B}(X; Y)] \right\}. \quad (\text{X.49})$$

X.3. Ergänzungen

X.3.1. Beschränkte lineare Operatoren

Definition X.10. Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ zwei normierte $\mathbb{K}$ -Vektorräume. Wir bezeichnen mit

$$\mathcal{B}(X; Y) := \{ A : X \rightarrow Y \mid A \text{ linear, } \|A\|_{\mathcal{B}(X; Y)} < \infty \} \quad (\text{X.50})$$

den Raum der **beschränkten linearen Operatoren (von X nach Y)**, wobei

$$\|A\|_{\mathcal{B}(X; Y)} := \sup_{x \in X \setminus \{0\}} \left\{ \frac{\|Ax\|_Y}{\|x\|_X} \right\} = \sup_{x \in X, \|x\|_X=1} \{ \|Ax\|_Y \} \quad (\text{X.51})$$

die **Operatornorm von A** ist.

Bemerkungen und Beispiele.

- Seien $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ und $(Z, \|\cdot\|_Z)$ drei Banachräume und $x \in X$, $A \in \mathcal{B}(X; Y)$ sowie $B \in \mathcal{B}(Y; Z)$. Aus der Definition (X.51) der Operatornorm folgt sofort, dass

$$\|Ax\|_Y \leq \|A\|_{\mathcal{B}(X; Y)} \cdot \|x\|_X, \quad (\text{X.52})$$

$$\|BA\|_{\mathcal{B}(X; Z)} \leq \|B\|_{\mathcal{B}(Y; Z)} \cdot \|A\|_{\mathcal{B}(X; Y)}. \quad (\text{X.53})$$

- Sind $(X, \|\cdot\|_X)$ ein normierter Raum und $(Y, \|\cdot\|_Y)$ ein Banachraum, so ist auch $(\mathcal{B}(X; Y), \|\cdot\|_{\mathcal{B}(X; Y)})$ ein Banachraum.
- Sind $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ zwei Banachräume über $\mathbb{K}$ und $A : X \rightarrow Y$ linear, so gilt folgende Äquivalenz:

$$\{ \|A\|_{\mathcal{B}(X; Y)} < \infty \} \quad \Leftrightarrow \quad \{ A : X \rightarrow Y \text{ ist stetig} \}. \quad (\text{X.54})$$

- Seien $M, N \in \mathbb{N}$ und $X = \mathbb{K}^M$, $Y = \mathbb{K}^N$. Wir wählen auf X und Y jeweils die ℓ^1 -Norm

$$\|x\|_X := |x_1| + \dots + |x_M|, \quad \|y\|_Y := |y_1| + \dots + |y_N|, \quad (\text{X.55})$$

wobei $x = (x_1, \dots, x_M)^T \in \mathbb{K}^M$ und $y = (y_1, \dots, y_N)^T \in \mathbb{K}^N$. Sei nun $A : \mathbb{K}^M \rightarrow \mathbb{K}^N$ eine lineare Abbildung mit der Matrixdarstellung $(A_{k,\ell})_{k \in \mathbb{Z}_1^N, \ell \in \mathbb{Z}_1^M} \in \mathfrak{M}_{N \times M}(\mathbb{K})$ bezüglich der Standardbasen in $\mathbb{K}^M$ und $\mathbb{K}^N$. Für $x = (x_1, \dots, x_M)^T \in \mathbb{K}^M$ ist dann

$$\begin{aligned} \|Ax\|_Y &= \sum_{k=1}^N |(Ax)_k| = \sum_{k=1}^N \left| \sum_{\ell=1}^M A_{k,\ell} x_\ell \right| \leq \sum_{\ell=1}^M \sum_{k=1}^N |A_{k,\ell}| |x_\ell| \\ &\leq \left(\max_{1 \leq m \leq M} \sum_{k=1}^N |A_{k,m}| \right) \sum_{\ell=1}^M |x_\ell| = \left(\max_{1 \leq m \leq M} \left\{ \sum_{k=1}^N |A_{k,m}| \right\} \right) \|x\|_X. \end{aligned} \quad (\text{X.56})$$

Also ist

$$\|A\|_{\mathcal{B}(X;Y)} = \sup_{x \in X \setminus \{0\}} \left\{ \frac{\|Ax\|_Y}{\|x\|_X} \right\} \leq \max_{1 \leq m \leq M} \left\{ \sum_{k=1}^N |A_{k,m}| \right\}. \quad (\text{X.57})$$

- Insbesondere ist jede lineare Abbildung von $\mathbb{K}^M$ nach $\mathbb{K}^N$ ein beschränkter linearer Operator,

$$\mathcal{B}(\mathbb{K}^M; \mathbb{K}^N) = \mathcal{L}(\mathbb{K}^M; \mathbb{K}^N), \quad (\text{X.58})$$

und zwar unabhängig von der auf $\mathbb{K}^M$ und $\mathbb{K}^N$ definierten Norm.

X.3.2. Eigenschaften der Operatornorm

Seien $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ und $(Z, \|\cdot\|_Z)$ drei Banachräume und $x \in X$, $A \in \mathcal{B}(X; Y)$ sowie $B \in \mathcal{B}(Y; Z)$. Wir zeigen folgenden Eigenschaften:

$$\|Ax\|_Y \leq \|A\|_{\mathcal{B}(X;Y)} \cdot \|x\|_X, \quad (\text{X.59})$$

$$\|BA\|_{\mathcal{B}(X;Z)} \leq \|B\|_{\mathcal{B}(Y;Z)} \cdot \|A\|_{\mathcal{B}(X;Y)}. \quad (\text{X.60})$$

Um (X.59) zu zeigen, schätzen wir für $x \neq 0$ wie folgt ab,

$$\begin{aligned} \|Ax\|_Y &= \frac{\|Ax\|_Y}{\|x\|_X} \cdot \|x\|_X \leq \sup_{x' \neq 0} \left\{ \frac{\|Ax'\|_Y}{\|x'\|_X} \right\} \cdot \|x\|_X \\ &= \|A\|_{\mathcal{B}(X;Y)} \cdot \|x\|_X. \end{aligned} \quad (\text{X.61})$$

Weiterhin gilt für $x \in X \setminus \{0\}$ mit $Bx \neq 0$, dass

$$\begin{aligned} \frac{\|ABx\|_Z}{\|x\|_X} &= \frac{\|A(Bx)\|_Z}{\|Bx\|_Y} \cdot \frac{\|Bx\|_Y}{\|x\|_X} \leq \sup_{y \in Y \setminus \{0\}} \left\{ \frac{\|Ay\|_Z}{\|y\|_Y} \right\} \cdot \sup_{x \in X \setminus \{0\}} \left\{ \frac{\|Bx\|_Y}{\|x\|_X} \right\} \\ &= \|A\|_{\mathcal{B}(Y;Z)} \cdot \|B\|_{\mathcal{B}(X;Y)}, \end{aligned} \quad (\text{X.62})$$

was natürlich auch für $Bx = 0$ richtig ist und somit (X.60) impliziert.

X.3.3. Äquivalenz von Stetigkeit und Beschränktheit linearer Operatoren

Wir zeigen, dass für zwei Banachräume $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ und einen linearen Operator $A : X \rightarrow Y$ folgende Eigenschaften gleichwertig sind:

$$\left\{ \|A\|_{\mathcal{B}(X;Y)} < \infty \right\} \quad \Leftrightarrow \quad \left\{ A : X \rightarrow Y \text{ ist stetig} \right\}. \quad (\text{X.63})$$

“ $\Rightarrow$ ”: Wir kürzen $\|A\|_{\mathcal{B}(X;Y)}$ mit $a := \|A\|_{\mathcal{B}(X;Y)}$ ab. Wegen

$$\|Ax - Ay\|_Y = \|A(x - y)\|_Y \leq a \cdot \|x - y\|_X \quad (\text{X.64})$$

ist

$$A(B_X(x, r)) \subseteq B_Y(Ax, ar), \quad (\text{X.65})$$

für alle $x \in X$ und $r > 0$. Für $\varepsilon > 0$ ist mit $\delta := \varepsilon a^{-1}$ also

$$\forall x \in X : \quad A(B_X(x, \delta)) \subseteq B_Y(Ax, \varepsilon). \quad (\text{X.66})$$

“ $\Leftarrow$ ”: Das Urbild der offenen Kugel $B_Y(0, 1) := \{y \in Y \mid \|y\| < 1\}$ um 0 in Y ist wegen der Stetigkeit von A und $A(0_X) = 0_Y$ eine offene Umgebung $U \subseteq X$ von $0 \in U$. Also gibt es ein $r > 0$, so dass $B_X(0, r) \subseteq U = A^{-1}[B_Y(1, 0)]$. D.h.

$$A[B_X(0, r)] \subseteq B_Y(0, 1) \quad \Leftrightarrow \quad \forall \tilde{x} \in X, \quad \|\tilde{x}\|_X < r : \quad \|A\tilde{x}\|_Y < 1. \quad (\text{X.67})$$

Ist nun $x \in X \setminus \{0\}$, so setzen wir

$$\tilde{x} := \frac{r}{2\|x\|_X} x \quad (\text{X.68})$$

und beobachten, dass

$$\|\tilde{x}\|_X = \frac{r}{2\|x\|_X} \cdot \|x\|_X = \frac{r}{2} < r, \quad (\text{X.69})$$

also $\tilde{x} \in B_X(0, r)$. Somit ist $\|A\tilde{x}\|_Y < 1$ gemäß (X.67). Aus der Linearität von A erhalten wir schließlich

$$\frac{\|Ax\|_Y}{\|x\|_X} = \frac{\left(\frac{r}{2\|x\|_X}\right) \cdot \|Ax\|_Y}{\left(\frac{r}{2\|x\|_X}\right) \cdot \|x\|_X} = \frac{\|A\tilde{x}\|_Y}{\|\tilde{x}\|_X} \leq \frac{2}{r}, \quad (\text{X.70})$$

und somit

$$\|A\|_{\mathcal{B}(X;Y)} \leq \frac{2}{r} < \infty. \quad (\text{X.71})$$

X.3.4. Beweis der Kettenregel

Wir setzen (X.43) sukzessiv ein und erhalten

$$\begin{aligned} & (F \circ G)(x + z) - (F \circ G)(x) \\ &= F'[G(x)](G(x + z) - G(x)) + r_{F,G(x)}[G(x + z) - G(x)] \cdot \|G(x + z) - G(x)\|_Y \\ &= (F'[G(x)] \circ G'(x))(z) + F'[G(x)](r_{G,x}(z)) \cdot \|z\|_X + \\ & \quad r_{F,G(x)}[G(x + z) - G(x)] \cdot \|G(x + z) - G(x)\|_Y \\ &= (F'[G(x)] \circ G'(x))(z) + r_{F \circ G, x}(z) \|z\|_X, \end{aligned} \quad (\text{X.72})$$

mit

$$r_{F \circ G, x}(z) := F'[G(x)](r_{G,x}(z)) + r_{F,G(x)}[G(x + z) - G(x)] \frac{\|G(x + z) - G(x)\|_Y}{\|z\|_X}, \quad (\text{X.73})$$

und es gilt zu zeigen, dass

$$\lim_{z \rightarrow 0} \|r_{F \circ G}(x, z)\|_Z = 0. \quad (\text{X.74})$$

Da $F'(G(x)) \in \mathcal{B}(Y, Z)$, können wir (X.43) anwenden und erhalten

$$\lim_{z \rightarrow 0} \|F'[G(x)](r_{G,x}(z))\|_Z \leq \|F'(G(x))\|_{\mathcal{B}(Y;Z)} \cdot \lim_{z \rightarrow 0} \|r_G(x, z)\|_Y = 0, \quad (\text{X.75})$$

wobei wir die Differenzierbarkeit von G bei x ausnutzen. Abermals wegen der Differenzierbarkeit von G bei x gilt

$$\limsup_{z \rightarrow 0} \left\{ \frac{\|G(x+z) - G(x)\|_Y}{\|z\|_X} \right\} = \|G'(x)\|_{\mathcal{B}(X;Y)}. \quad (\text{X.76})$$

Nach Lemma X.6 ist G stetig bei x , und es gilt $\lim_{z \rightarrow 0} \{G(x+z) - G(x)\} = 0$. Daher ist

$$\lim_{z \rightarrow 0} r_{F,G(x)}[G(x+z) - G(x)] = 0, \quad (\text{X.77})$$

und (X.75)–(X.77) ergeben (X.44).

X.3.5. Äquivalenz stetiger partieller und stetiger totaler Differenzierbarkeit

Wir beweisen Satz X.8 in diesem Abschnitt.

Beweis. Wir beweisen zunächst (X.46) $\Rightarrow$ (X.45), und zwar für den Spezialfall $N = 1$. Dazu wählen wir für $\mathbb{R}^M$ wieder die ℓ^1 -Norm, $\|(x_1, \dots, x_M)^T\| := |x_1| + \dots + |x_M|$. Wir schreiben $x = \sum_{j=1}^M x_j e_j$, $z = \sum_{j=1}^M z_j e_j$ und $w_m = \sum_{j=1}^m z_j e_j$, für alle $0 \leq m \leq M$, wobei $e_i \in \mathbb{R}^M$ die kanonischen Basivektoren sind, sowie $\partial_m F(x) := \frac{\partial F(x)}{\partial x_m}$. und beachten, dass $w_M = z$ und $w_0 = 0$. Damit ist

$$\begin{aligned} & \left| F(x+z) - F(x) - \sum_{m=1}^M \partial_m F(x) z_m \right| \\ &= \left| \{F(x+w_M) - F(x+w_{M-1})\} + \{F(x+w_{M-1}) - F(x+w_{M-2})\} \right. \\ & \quad \left. + \dots + \{F(x+w_1) - F(x+w_0)\} - \sum_{m=1}^M \partial_m F(x) z_m \right| \\ &\leq \left| \sum_{m=1}^M \{F(x+w_m) - F(x+w_{m-1}) - \partial_m F(x) z_m\} \right| \\ &\leq \sum_{m=1}^M |F(x+w_{m-1} + z_m e_m) - F(x+w_{m-1}) - \partial_m F(x) z_m| \end{aligned} \quad (\text{X.78})$$

Weiterhin sind alle $\partial_m F$ auf einer offenen Umgebung $U \ni x$ stetig, und daher sichert eine genügend kleine Wahl von $\delta > 0$, dass

$$\forall 1 \leq m \leq M : \quad \partial_m F \text{ ist stetig auf } x + K_\delta, \quad (\text{X.79})$$

wobei $K_\delta := [-\delta, \delta]^M$. Für $1 \leq m \leq M$ und $z \in K_\delta$ impliziert der Mittelwertsatz der Differentialrechnung die Existenz einer Zahl $0 < \vartheta_m < 1$, so dass

$$F(x + w_{m-1} + z_m e_m) - F(x + w_{m-1}) = \partial_m F(x + w_{m-1} + \vartheta_m z_m e_m) \cdot z_m, \quad (\text{X.80})$$

und wegen $w_{m-1} + \vartheta_m z_m e_m \in K_\delta$ erhalten wir daraus

$$|F(x + w_{m-1} + z_m e_m) - F(x + w_{m-1}) - \partial_m F(x) z_m| \leq Q(\delta) |z_m|, \quad (\text{X.81})$$

wobei

$$Q(\delta) := \max_{1 \leq m \leq M} \max_{y \in K_\delta} |\partial_m F(x + y) - \partial_m F(x)| \rightarrow 0, \quad (\text{X.82})$$

für $\delta \rightarrow 0$, wegen der Stetigkeit von $\partial_m F$ auf $U \supset x + K_\delta$. Setzen wir (X.82) in (X.78) ein, so erhalten wir

$$\frac{1}{\|z\|} \left| F(x + z) - F(x) - \sum_{m=1}^M \partial_m F(x) z_m \right| \leq \frac{Q(\delta)}{\|z\|} \left(\sum_{m=1}^M |z_m| \right) = Q(\delta) \rightarrow 0, \quad (\text{X.83})$$

für $\delta \rightarrow 0$. Dies beweist die totale Differenzierbarkeit von F bei x und $F'(x) = J_F(x)$ für $N = 1$. Letztere Gleichung ergibt auch die Stetigkeit von $x \mapsto F'(x)$, da $x \mapsto J_F(x)$ nach (X.46) stetig ist.

Um (X.46) $\Rightarrow$ (X.45) für $N \geq 1$ zu beweisen, wenden wir (X.83) auf alle Komponenten $F_1, F_2, \dots, F_N$ an und erhalten

$$\forall 1 \leq n \leq N : \quad \lim_{z \rightarrow 0} \left\{ \frac{1}{\|z\|} \left| F_n(x + z) - F_n(x) - \sum_{m=1}^M \partial_m F_n(x) z_m \right| \right\} = 0. \quad (\text{X.84})$$

Damit erhalten wir aber

$$\lim_{z \rightarrow 0} \left\{ \frac{1}{\|z\|_{\mathbb{R}^M}} \left\| F_n(x + z) - F_n(x) - \sum_{m=1}^M \partial_m F_n(x) z_m \right\|_{\mathbb{R}^N} \right\} \quad (\text{X.85})$$

$$= \lim_{z \rightarrow 0} \left\{ \frac{1}{\|z\|_{\mathbb{R}^M}} \sum_{n=1}^N \left| F_n(x + z) - F_n(x) - \sum_{m=1}^M \partial_m F_n(x) z_m \right| \right\} = 0.$$

Dies beweist für $N \geq 1$ die totale Differenzierbarkeit von F bei x , $F'(x) = J_F(x)$ und die Stetigkeit von F' wie für $N = 1$.

Um (X.45) $\Rightarrow$ (X.46) zu beweisen, wählen wir $j \in \mathbb{Z}_1^M$ und setzen $z := t e_j$, sodass

$$\begin{aligned} 0 &= \lim_{z \rightarrow 0} \left\{ \frac{\|F(x + z) - F(x) - F'(x) \cdot z\|_{\mathbb{R}^N}}{\|z\|_{\mathbb{R}^M}} \right\} \\ &= \lim_{t \rightarrow 0} \left\{ \frac{1}{t} \|F(x + t e_j) - F(x) - t F'(x) e_j\| \right\}. \end{aligned} \quad (\text{X.86})$$

Daher muss auch, für jedes $i \in \mathbb{Z}_1^N$

$$\begin{aligned} 0 &= \lim_{t \rightarrow 0} \left\{ \frac{1}{t} [F_i(x + t e_j) - F_i(x) - (F'(x))_{i,j} \cdot t] \right\} \\ &= \lim_{t \rightarrow 0} \left\{ \frac{F_i(x + t e_j) - F_i(x)}{t} \right\} - (F'(x))_{i,j} \end{aligned} \quad (\text{X.87})$$

gelten. Also ist F partiell differenzierbar auf U , die Jacobimatrix $J_F = F'$ ist gegeben durch die Ableitung von F , die gemäß (X.45) stetig von x abhängt. $\square$

XI. Richtungsableitung und Extrema

In der Vorlesung über Differentiation einer reellen Funktion f einer reellen Veränderlichen haben wir notwendige und hinreichende Bedingungen für die Existenz von Extrema, Minima und Maxima gefunden: An einer solchen Stelle gilt notwendig $f' = 0$, und $f'' > 0$ bzw. $f'' < 0$ sind hinreichend dafür, dass diese Stelle ein Minimum bzw. Maximum ist. Ziel dieses Abschnitts ist die Formulierung eines entsprechenden Kriteriums für reelle Funktionen mehrerer reeller Veränderlicher.

XI.1. Der Satz von Taylor in mehreren Veränderlichen

Dazu verallgemeinern wir als erstes den Satz von Taylor auf mehrere Variablen.

Satz XI.1 (Taylor). *Seien $M \in \mathbb{N}$, $k \in \mathbb{N}_0$ und $U \subset \mathbb{R}^M$ eine offene Teilmenge. Seien weiterhin $x_0 \in U$ und $z \in \mathbb{R}^M$, so dass*

$$S := \{x_0 + tz \mid 0 \leq t \leq 1\} \subseteq U, \quad (\text{XI.1})$$

und sei $F \in C^{k+1}(U; \mathbb{R})$. Dann gibt es ein $\vartheta \in (0, 1)$ so, dass

$$\begin{aligned} F(x_0 + z) &= \sum_{p=0}^k \frac{1}{p!} \sum_{j_1, j_2, \dots, j_p=1}^M \frac{\partial^p F(x_0)}{\partial x_{j_1} \dots \partial x_{j_p}} \cdot z_{j_1} \cdot z_{j_2} \cdots z_{j_p} \\ &\quad + \sum_{j_1, \dots, j_{k+1}=1}^M \frac{1}{(k+1)!} \left(\frac{\partial^{k+1} F(x_0 + \vartheta z)}{\partial x_{j_1} \dots \partial x_{j_{k+1}}} \right) \cdot z_{j_1} \cdot z_{j_2} \cdots z_{j_{k+1}}, \end{aligned} \quad (\text{XI.2})$$

wobei $z = (z_1, \dots, z_M)^T = \sum_{j=1}^M z_j e_j$.

Beweis. Wir definieren

$$f(t) := F(x_0 + tz). \quad (\text{XI.3})$$

Da U offen ist und $\{x_0 + tz \mid 0 \leq t \leq 1\}$ enthält, ist für genügend kleine $\eta > 0$ auch $\tilde{S} := \{x_0 + tz \mid t \in I_\eta\} \subseteq U$, wobei $I_\eta := (-\eta, 1 + \eta)$. Durch Anwendung der Kettenregel, Satz X.7, auf $f := F \circ G$, wobei

$$G : I_\eta \rightarrow U, \quad t \mapsto x_0 + tz, \quad (\text{XI.4})$$

erhält man nun leicht, dass

$$f \in C^{k+1}(I_\eta; \mathbb{R}). \quad (\text{XI.5})$$

Die Kettenregel sichert nämlich, dass mit $F \in C^1(U; \mathbb{R})$ und $G \in C^1(I_\eta; U)$ auch $f \in C^1(I_\eta; \mathbb{R})$ mit

$$\begin{aligned} f'(s) &= F'(x_0 + sz) \cdot G'(s) \\ &= \left(\frac{\partial F(x_0 + sz)}{\partial x_1}, \dots, \frac{\partial F(x_0 + sz)}{\partial x_M} \right) \cdot \begin{pmatrix} \frac{\partial G_1(s)}{\partial s} \\ \vdots \\ \frac{\partial G_M(s)}{\partial s} \end{pmatrix} \\ &= \left(\frac{\partial F(x_0 + sz)}{\partial x_1}, \dots, \frac{\partial F(x_0 + sz)}{\partial x_M} \right) \cdot \begin{pmatrix} z_1 \\ \vdots \\ z_M \end{pmatrix} \\ &= \sum_{j_1=1}^M \frac{\partial F(x_0 + sz)}{\partial x_{j_1}} \cdot z_{j_1}. \end{aligned} \quad (\text{XI.6})$$

Setzen wir nun

$$F_1(x) := \sum_{j_1=1}^M \frac{\partial F(x)}{\partial x_{j_1}} z_{j_1}, \quad (\text{XI.7})$$

so ist $F_1 \in C^k(U; \mathbb{R})$ und

$$f'(s) = (F_1 \circ G)(s). \quad (\text{XI.8})$$

Abermals erhalten wir dann aus der Kettenregel, dass $f' \in C^1(I_\eta; \mathbb{R})$ mit

$$\begin{aligned} f''(s) &= (f'(s))' = F_1'(x_0 + sz) \cdot G'(s) \\ &= \sum_{j_2=1}^M \frac{\partial F_1(x_0 + sz)}{\partial x_{j_2}} \cdot z_{j_2} = \sum_{j_1, j_2=1}^M \frac{\partial^2 F(x_0 + sz)}{\partial x_{j_1} \partial x_{j_2}} \cdot z_{j_1} z_{j_2}, \end{aligned} \quad (\text{XI.9})$$

wobei wir die Vertauschbarkeit der partiellen Ableitungen gemäß Satz X.4 verwenden.

Nach $(k+1)$ -maliger Anwendung dieses Arguments erhalten wir dann, dass $f \in C^{k+1}(I_\eta; \mathbb{R})$ -wie in (XI.5) behauptet- und dass

$$f^{(p)}(s) = \sum_{j_1, j_2, \dots, j_p=1}^M \frac{\partial^p F(x_0 + sz)}{\partial x_{j_1} \partial x_{j_2} \dots \partial x_{j_p}} z_{j_1} \cdot z_{j_2} \dots z_{j_p}, \quad (\text{XI.10})$$

für alle $p = 0, 1, \dots, k+1$. Nun wenden wir den Satz VII.12 von Taylor auf f an und erhalten daraus eine Zwischenstelle $\vartheta \in (0, 1)$, sodass

$$f(1) = \sum_{p=0}^K \frac{1}{p!} f^{(p)}(0) + \frac{1}{k!} f^{(k+1)}(\vartheta) \quad (\text{XI.11})$$

gilt, was die Behauptung ergibt. □

Satz XI.1 besitzt folgende alternative Formulierung:

Korollar XI.2. Seien $M \in \mathbb{N}$, $k \in \mathbb{N}_0$ und $U \subseteq \mathbb{R}^M$ eine offene Teilmenge. Seien $x_0 \in U$ und $z \in \mathbb{R}^M$, so dass

$$S := \{x_0 + tz \mid 0 \leq t \leq 1\} \subseteq U, \quad (\text{XI.12})$$

und sei $F \in C^{k+1}(U; \mathbb{R})$. Dann gibt es ein $\vartheta \in (0, 1)$ so, dass

$$\begin{aligned} F(x_0 + z) &= \sum_{p=0}^k \sum_{\substack{p_1, \dots, p_M=0 \\ p_1 + \dots + p_M = p}}^p \left(\frac{\partial^p F(x_0)}{\partial x_1^{p_1} \dots \partial x_M^{p_M}} \right) \cdot \left(\frac{z_1^{p_1} \dots z_M^{p_M}}{p_1! \dots p_M!} \right) \\ &\quad + \sum_{\substack{p_1, \dots, p_M=0 \\ p_1 + \dots + p_M = k+1}}^{k+1} \left(\frac{\partial^{k+1} F(x_0 + sz)}{\partial x_1^{p_1} \dots \partial x_M^{p_M}} \right) \cdot \frac{z_1^{p_1}}{p_1!} \frac{z_2^{p_2}}{p_2!} \dots \frac{z_M^{p_M}}{p_M!}. \end{aligned} \quad (\text{XI.13})$$

Bemerkungen und Beispiele.

Für $M = 2$, $U := \mathbb{R}^2$ und $F(x_1, x_2) := \cos(x_1) + \cos(x_2)$ ist offensichtlich $F \in C^\infty(\mathbb{R}^2; \mathbb{R})$. Der Graph von F ähnelt einem Eierkarton. Wir wenden nun Satz XI.1 für $k = 2$ an und erhalten mit

$$\frac{\partial F}{\partial x_1} = -\sin x_1, \quad \frac{\partial F}{\partial x_2} = -\sin x_2, \quad (\text{XI.14})$$

$$\frac{\partial^2 F}{\partial x_1^2} = -\cos x_1, \quad \frac{\partial^2 F}{\partial x_1 \partial x_2} = \frac{\partial^2 F}{\partial x_2 \partial x_1} = 0, \quad \frac{\partial^2 F}{\partial x_2^2} = -\cos x_2, \quad (\text{XI.15})$$

$$\frac{\partial^3 F}{\partial x_1^3} = \sin x_1, \quad \frac{\partial^3 F}{\partial x_1^2 \partial x_2} = 0, \quad \frac{\partial^3 F}{\partial x_1 \partial x_2^2} = 0, \quad \frac{\partial^3 F}{\partial x_2^3} = \sin x_2, \quad (\text{XI.16})$$

dass

$$\begin{aligned} F(x + z) &= \cos(x_1) + \cos(x_2) - z_1 \cdot \sin(x_1) - z_2 \cdot \sin(x_2) \\ &\quad - \frac{1}{2} z_1^2 \cdot \cos(x_1) - \frac{1}{2} z_2^2 \cdot \cos(x_2) + R_3 \cdot \|z\|_2^3, \end{aligned} \quad (\text{XI.17})$$

wobei

$$\begin{aligned} R_3 &:= \frac{1}{3!} \sum_{j_1, \dots, j_3=1}^2 \left| \frac{\partial^3 F(x + \vartheta z)}{\partial x_{j_1} \partial x_{j_2} \partial x_{j_3}} \right| \cdot \frac{|z_{j_1}| \cdot |z_{j_2}| \cdot |z_{j_3}|}{\|z\|_2^3} \\ &:= \frac{1}{6} \left| \sin(x_1 + \vartheta z_1) \right| \cdot \frac{|z_1|^3}{\|z\|_2^3} + \frac{1}{6} \left| \sin(x_2 + \vartheta z_2) \right| \cdot \frac{|z_2|^3}{\|z\|_2^3} \\ &\leq \frac{1}{6} \frac{|z_1|^3 + |z_2|^3}{\|z\|_2^3} \leq \frac{1}{3}, \end{aligned} \quad (\text{XI.18})$$

wobei wir noch verwenden, dass $|z_j|^3 \leq (z_1^2 + z_2^2)^{3/2}$ gilt.

XI.2. Gradient, Richtungsableitung und Hessesche Matrix

Definition XI.3. Seien $M \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$ eine nichtleere offene Teilmenge und $x \in U$.

(i) Für $F \in C^1(U; \mathbb{R})$ bezeichnen wir den Spaltenvektor

$$\nabla F(x) := (F'(x))^T = \begin{pmatrix} \frac{\partial F(x)}{\partial x_1} \\ \vdots \\ \frac{\partial F(x)}{\partial x_M} \end{pmatrix} \quad (\text{XI.19})$$

als **Gradient von F bei x** und

$$\langle \nabla F(x) | z \rangle_{\text{eukl}} = \sum_{j=1}^M \frac{\partial F(x)}{\partial x_j} \cdot z_j \quad (\text{XI.20})$$

heißt **Richtungsableitung von F bei x in Richtung z** , wobei $\langle x | y \rangle_{\text{eukl}} := \sum_{j=1}^M x_j y_j$ das euklidische Skalarprodukt auf $\mathbb{R}^M$ bezeichnet.

(ii) Für $F \in C^2(U; \mathbb{R})$ heißt $\text{Hess } F(x) \in \mathfrak{M}_{M \times M}(\mathbb{R})$,

$$(\text{Hess } F(x))_{i,j} \equiv (F''(x))_{ij} := \frac{\partial^2 F(x)}{\partial x_i \partial x_j} \quad (\text{XI.21})$$

Hessesche Matrix von F bei x .

Bemerkungen und Beispiele.

- Wir betrachten nun einen Vektor $z = (z_1, \dots, z_M) \in \mathbb{R}^M$ der Länge 1, d.h. $z_1^2 + z_2^2 + \dots + z_M^2 = 1$, und bilden für $F \in C^1(U; \mathbb{R})$, $U \subseteq \mathbb{R}^M$ offen und $x \in U$,

$$\lim_{t \rightarrow 0} \left\{ \frac{F(x + tz) - F(x)}{t} \right\} = \sum_{j=1}^M \frac{\partial F}{\partial x_j}(x) \cdot z_j = \langle \nabla F(x) | z \rangle. \quad (\text{XI.22})$$

Es gilt also

$$F(x + tz) = F(x) + t[\langle \nabla F(x) | z \rangle + r(t)], \quad (\text{XI.23})$$

mit $\lim_{t \rightarrow 0} r(t) = 0$. Die Richtungsableitung gibt also an, wie stark F an der Stelle x in Richtung z wächst.

- Mit der Cauchy-Schwarz-Ungleichung erhalten wir weiter, dass

$$|\langle \nabla F(x) | z \rangle| \leq \|\nabla F(x)\| \cdot \|z\| = \|\nabla F(x)\|. \quad (\text{XI.24})$$

Andererseits hat aber $\|\nabla F(x)\|^{-1} \cdot \nabla F(x)$, für $\nabla F(x) \neq 0$, die Länge 1, und es gilt

$$\left\langle \nabla F(x) \left| \frac{\nabla F(x)}{\|\nabla F(x)\|} \right. \right\rangle = \frac{\|\nabla F(x)\|^2}{\|\nabla F(x)\|} = \|\nabla F(x)\|. \quad (\text{XI.25})$$

Also ist

$$\|\nabla F(x)\| = \max \left\{ \langle \nabla F(x) | z \rangle \mid z \in \mathbb{R}^M, \|z\| = 1 \right\}, \quad (\text{XI.26})$$

und somit ist die Richtungsableitung maximal, wenn z parallel zum Gradienten gewählt wird. Anders ausgedrückt, *liegt der Gradient einer Funktion bei x in Richtung des größtmöglichen Anstiegs von F .*

- Außerdem steht der Gradient von F bei x senkrecht zur zu x gehörigen Niveau-
menge. Ist nämlich $\alpha \in C^\infty((a, b); \mathcal{N}_E)$ eine Kurve in

$$\mathcal{N}_E = \{x \in U \mid F(x) = E\} = F^{-1}(E), \quad (\text{XI.27})$$

so folgt sofort, dass

$$0 = \frac{d}{ds}(E) = \frac{d}{ds}[(F \circ \alpha)(s)] = \sum_{j=1}^M \frac{\partial F(x)}{\partial x_j} \cdot \frac{d\alpha_j(s)}{ds} = \left\langle \nabla F(x) \mid \frac{d\alpha(s)}{ds} \right\rangle. \quad (\text{XI.28})$$

XI.3. Kritische Punkte und Extremalwerte bei mehreren Veränderlichen

Definition XI.4. Seien $(X, \|\cdot\|)$ ein normierter Raum, $U \subseteq X$ eine nichtleere Teilmenge und $F : U \rightarrow \mathbb{R}$. Dann heißt

$$\begin{aligned} x_0 \in U \text{ lokales } \left\{ \begin{array}{l} \text{Minimum} \\ \text{Maximum} \end{array} \right\} \text{ von } F \\ \Leftrightarrow \quad \exists \delta > 0 \forall x \in B(x_0, \delta) \cap U : \quad \left\{ \begin{array}{l} F(x) \geq F(x_0) \\ F(x) \leq F(x_0) \end{array} \right\}. \end{aligned} \quad (\text{XI.29})$$

Definition XI.5. Seien $M \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$ eine nichtleere offene Teilmenge und $F \in C^1(U; \mathbb{R})$. Ein Punkt $x_0 \in U$ heißt **kritisch** (oder **extremal**)

$$\Leftrightarrow \quad \nabla F(x_0) = 0. \quad (\text{XI.30})$$

Satz XI.6. Seien $M \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$ eine nichtleere offene Teilmenge und $F \in C^1(U; \mathbb{R})$. Dann gelten

$$(i) \quad \left(x_0 \in U \text{ ist lokales Minimum von } F \right) \implies \left(\nabla F(x_0) = 0 \right). \quad (\text{XI.31})$$

$$(ii) \quad \left(x_0 \in U \text{ ist lokales Maximum von } F \right) \implies \left(\nabla F(x_0) = 0 \right). \quad (\text{XI.32})$$

Beweis. Wir zeigen nur (i), (ii) ist analog. Seien also $x_0 \in U$ ein lokales Minimum von F und $\delta > 0$ wie in (XI.29). Da $U \subseteq \mathbb{R}^M$ offen ist, gibt es ein $\delta' > 0$, sodass $B(x_0, \delta') \subseteq U$. Nach (XI.29) gilt also mit $\delta'' := \min\{\delta, \delta'\} > 0$, dass

$$\forall z \in B(x_0, \delta'') : \quad F(x_0 + z) - F(x_0) \geq 0. \quad (\text{XI.33})$$

Für $z \in \mathbb{R}^M$ mit $\|z\|_M = 1$ und $t \in (0, \delta'')$ ist $\pm tz \in B(x_0, \delta'')$ und deshalb

$$\pm F'(x_0)z = \frac{1}{t} F'(x_0)(\pm tz) = \lim_{t \searrow 0} \left\{ \frac{F(x_0 \pm tz) - F(x_0)}{t} \right\} \geq 0. \quad (\text{XI.34})$$

Für alle $z \in \mathbb{R}^M$ mit $\|z\|_M = 1$ muss daher $\nabla F(x_0) = 0$ gelten. Ist nun $w \in \mathbb{R}^M \setminus \{0\}$, so setzen wir $z_w := \|w\|^{-1}w$ und beobachten, dass $\|z_w\|_M = 1$. Deshalb gilt

$$\forall w \in \mathbb{R}^M \setminus \{0\} : \quad \langle \nabla F(x_0) | w \rangle = \|w\| \langle \nabla F(x_0) | z_w \rangle = 0, \quad (\text{XI.35})$$

was für $w = 0$ trivial richtig ist. $\square$

Bemerkungen und Beispiele.

Wir kommen nun auf das Beispiel (XI.14)–(XI.18)

$$F : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x_1, x_2) \mapsto F(x_1, x_2) = \cos(x_1) + \cos(x_2) \quad (\text{XI.36})$$

zurück. Die Menge $\mathcal{E}$ der Extrempunkte von F ist gegeben durch

$$\mathcal{E} := \left\{ x = (x_1, x_2) \in \mathbb{R}^2 \mid \nabla F(x) = \begin{pmatrix} -\sin(x_1) \\ -\sin(x_2) \end{pmatrix} = 0 \right\} = \pi\mathbb{Z} \times \pi\mathbb{Z}. \quad (\text{XI.37})$$

Dabei sind etwa $(0, 0) \in \mathcal{E}$ ein lokales Maximum und $(\pi, \pi) \in \mathcal{E}$ ein lokales Minimum, jedoch sind z.B. $(0, \pi)$, $(\pi, 0) \in \mathcal{E}$ weder lokale Minima noch lokale Maxima, sondern Sattelpunkte.

Wie in der Kurvendiskussion für reelle Funktionen einer reellen Veränderlichen, lassen sich hinreichende Bedingungen für die Existenz eines lokalen Maximums oder Minimums durch die zweiten (oder höheren) Ableitungen formulieren.

Sei $F \in C^2(U; \mathbb{R})$, wobei $U \subseteq \mathbb{R}^M$ eine offene Teilmenge ist. Dann ist nach Satz X.4

$$\forall x \in U, \quad i, j \in \mathbb{N}_1^M : \quad (F''(x))_{i,j} = \frac{\partial^2 F(x)}{\partial x_i \partial x_j} = \frac{\partial^2 F(x)}{\partial x_j \partial x_i} = (F''(x))_{j,i} \in \mathbb{R}, \quad (\text{XI.38})$$

d.h. die Hessesche Matrix von F ist reell und symmetrisch und somit *selbstadjungiert*.

Lemma XI.7 (Minimax-Prinzip). *Seien $M \in \mathbb{N}$ und $A = A^T \in \mathfrak{M}_{M \times M}(\mathbb{R})$ eine reelle, selbstadjungierte Matrix mit Eigenwerten $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_M$. Dann gilt*

$$\lambda_1 = \min \left\{ \frac{\langle x | Ax \rangle}{\langle x | x \rangle} \mid x \in \mathbb{R}^M \setminus \{0\} \right\} \leq \max \left\{ \frac{\langle x | Ax \rangle}{\langle x | x \rangle} \mid x \in \mathbb{R}^M \setminus \{0\} \right\} = \lambda_M. \quad (\text{XI.39})$$

Definition XI.8. Seien $M \in \mathbb{N}$ und $A = A^T \in \mathfrak{M}_{M \times M}(\mathbb{R})$ eine selbstadjungierte Matrix mit Eigenwerten $\lambda_1, \dots, \lambda_M \in \mathbb{R}$.

$$A \text{ heißt positiv definit} \quad :\Leftrightarrow \quad \forall \alpha \in \mathbb{Z}_1^M : \lambda_\alpha > 0; \quad (\text{XI.40})$$

$$A \text{ heißt positiv semidefinit} \quad :\Leftrightarrow \quad \forall \alpha \in \mathbb{Z}_1^M : \lambda_\alpha \geq 0; \quad (\text{XI.41})$$

$$A \text{ heißt negativ definit} \quad :\Leftrightarrow \quad \forall \alpha \in \mathbb{Z}_1^M : \lambda_\alpha < 0; \quad (\text{XI.42})$$

$$A \text{ heißt negativ semidefinit} \quad :\Leftrightarrow \quad \forall \alpha \in \mathbb{Z}_1^M : \lambda_\alpha \leq 0; \quad (\text{XI.43})$$

$$A \text{ heißt indefinit} \quad :\Leftrightarrow \quad \exists \alpha, \beta \in \mathbb{Z}_1^M : \lambda_\alpha > 0, \lambda_\beta < 0. \quad (\text{XI.44})$$

Satz XI.9. Seien $M \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$ eine offene und nichtleere Teilmenge, $F \in C^2(U; \mathbb{R})$ und $x_0 \in U$ ein Extrempunkt, $\nabla F(x_0) = 0$. Dann gelten:

- (i) Ist die Hessesche Matrix $F''(x_0)$ von F bei x_0 positiv definit, so ist x_0 ein lokales Minimum.
- (ii) Ist die Hessesche Matrix $F''(x_0)$ von F bei x_0 negativ definit, so ist x_0 ein lokales Maximum.

Beweis. Wir beweisen nur (i), der Beweis für (ii) ist analog. Zunächst bemerken wir, dass die positive Definitheit von $F''(x_0)$ und die Stetigkeit von F'' in U impliziert, dass auch $F''(x)$ positiv definit ist, falls $\|x - x_0\|$ genügend klein ist. Sind $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_M$ die Eigenwerte von $F''(x_0)$ und $z \in \mathbb{R}^M$, so gilt nämlich nach Lemma XI.7

$$\begin{aligned} \langle z | F''(x) z \rangle &= \langle z | F''(x_0) z \rangle + \langle z | [F''(x) - F''(x_0)] z \rangle \\ &\geq \lambda_1 \|z\|^2 - \|z\| \cdot \|[F''(x) - F''(x_0)] z\| \\ &\geq \left\{ \lambda_1 - \|F''(x) - F''(x_0)\|_{B(\mathbb{R}^M; \mathbb{R}^M)} \right\} \cdot \|z\|^2. \end{aligned} \quad (\text{XI.45})$$

Wählen wir nun $\delta > 0$ so klein, dass $\|F''(x) - F''(x_0)\| \leq \lambda_1/2$, für $x \in B_\delta(x_0)$, so folgt dann

$$\forall x \in B(x_0, \delta) \quad \forall z \in \mathbb{R}^M : \quad \langle z | F''(x) z \rangle \geq \frac{\lambda_1}{2} \|z\|^2. \quad (\text{XI.46})$$

Nun wenden wir den Satz XI.1 von Taylor mit $k = 1$ auf F für x_0 und $z \in B(0, \delta)$ an und erhalten ein $\vartheta \in (0, 1)$ so, dass

$$\begin{aligned} F(x_0 + z) - F(x_0) &= \langle \nabla F(x_0) | z \rangle + \frac{1}{2} \sum_{i,j=1}^M \left(\frac{\partial^2 F(x_0 + \vartheta z)}{\partial x_i \partial x_j} \right) z_i z_j \\ &= \left\langle z \left| F''(x_0 + \vartheta z) z \right. \right\rangle \geq \frac{\lambda_1}{4} \|z\|^2 > 0, \end{aligned} \quad (\text{XI.47})$$

gemäß (XI.46), da $(x_0 + \vartheta z) \in B_\delta(x_0)$. □

Bemerkungen und Beispiele.

- Wir bemerken, dass aus (XI.47) ersichtlich ist, dass das Minimum bei x_0 auch nicht entartet ist, d.h. dass $F(x_0 + z) > F(x_0)$ für alle $z \in B(x_0, \delta) \setminus \{0\}$ gilt.
- Betrachten wir abermals $F(x_1, x_2) = \cos(x_1) + \cos(x_2)$ auf $\mathbb{R}^2$ mit kritischen Punkten

$$\mathcal{E} = \{(n_1\pi, n_2\pi) \mid n_1, n_2 \in \mathbb{Z}\}. \quad (\text{XI.48})$$

Mit

$$\frac{\partial^2 F}{\partial x_1^2} = -\cos(x_1), \quad \frac{\partial^2 F}{\partial x_1 \partial x_2} = \frac{\partial^2 F}{\partial x_2 \partial x_1} = 0, \quad \frac{\partial^2 F}{\partial x_2^2} = -\cos(x_2) \quad (\text{XI.49})$$

ist dann

$$F''(x_1, x_2) = \begin{pmatrix} -\cos(x_1) & 0 \\ 0 & -\cos(x_2) \end{pmatrix}. \quad (\text{XI.50})$$

Nun ist

$$-\cos(n\pi) = \begin{cases} -1, & \text{falls } n \in 2\mathbb{Z}, \\ +1, & \text{falls } n \in 2\mathbb{Z} + 1. \end{cases} \quad (\text{XI.51})$$

Also zerfällt $\mathcal{E} = \pi\mathbb{Z} \times \pi\mathbb{Z}$ in disjunkte Teilmengen,

$$\begin{aligned} \mathcal{E} = & (2\pi\mathbb{Z} \times 2\pi\mathbb{Z}) \dot{\cup} (2\pi\mathbb{Z} \times 2\pi\mathbb{Z} + \pi) \dot{\cup} \\ & (2\pi\mathbb{Z} + \pi \times 2\pi\mathbb{Z}) \dot{\cup} (2\pi\mathbb{Z} + \pi \times 2\pi\mathbb{Z} + \pi), \end{aligned} \quad (\text{XI.52})$$

und

$$F'' \text{ ist } \begin{cases} \text{positiv definit auf} & (2\pi\mathbb{Z} + \pi \times 2\pi\mathbb{Z} + \pi), \\ \text{negativ definit auf} & (2\pi\mathbb{Z} \times 2\pi\mathbb{Z}), \\ \text{indefinit auf} & (2\pi\mathbb{Z} \times 2\pi\mathbb{Z} + \pi) \cup (2\pi\mathbb{Z} + \pi \times 2\pi\mathbb{Z}). \end{cases} \quad (\text{XI.53})$$

Entsprechend sind

- $(0, 0) \in (2\pi\mathbb{Z} \times 2\pi\mathbb{Z})$ ein lokales Maximum, da $F''(0, 0)$ negativ definit ist,
 - $(\pi, \pi) \in (2\pi\mathbb{Z} + \pi \times 2\pi\mathbb{Z} + \pi)$ ein lokales Minimum, da $F''(\pi, \pi)$ positiv definit ist und
 - $(0, \pi), (\pi, 0) \in (2\pi\mathbb{Z} + \pi \times 2\pi\mathbb{Z}) \cup (2\pi\mathbb{Z} \times 2\pi\mathbb{Z} + \pi)$ sind Sattelpunkte, da $F''(0, \pi)$ und $F''(\pi, 0)$ indefinit sind.
- Tatsächlich kann man im Fall, dass x_0 ein kritischer Punkt mit *indefiniten* Hesseschen Matrix $F''(x_0)$ zeigen, dass die Eigenvektoren $\{\varphi_\alpha\}$ zu negativen Eigenwerten die Richtungen angeben, in denen F lokal maximal ist. Genauso geben die Eigenvektoren $\{\varphi_\alpha\}$ zu positiven Eigenwerten die Richtungen an, in denen F lokal minimal ist.

Wegen dieser geometrischen Interpretation bezeichnet man die durch die Eigenvektoren Hesseschen Matrix $F''(x_0)$ definierten Raumrichtungen als *Hauptachsen* und die orthogonale Matrix U , die die Hessesche Matrix diagonalisiert, als *Hauptachsentransformation*.

- Sind $U = \mathbb{R}^2$ und $F(x_1, x_2) := x_1^2 - x_2^4$, so ist

$$\frac{\partial F}{\partial x_1} = 2x_1, \quad \frac{\partial F}{\partial x_2} = -4x_2^3 \quad \Rightarrow \quad \mathcal{E} = \{(0, 0)\}, \quad (\text{XI.54})$$

$$\frac{\partial^2 F(x)}{\partial x_1^2} = 2, \quad \frac{\partial^2 F(x)}{\partial x_1 \partial x_2} = \frac{\partial^2 F(x)}{\partial x_2 \partial x_1} = 0, \quad \frac{\partial^2 F(x)}{\partial x_2^2} = -12x_2^2. \quad (\text{XI.55})$$

Somit ist

$$F''(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \quad (\text{XI.56})$$

positiv semidefinit. Der Punkt $x = (0, 0)$ ist aber kein lokales Minimum. Wir sehen an diesem Beispiel, dass sich Satz XI.9 nicht auf kritische Punkte mit positiv semidefiniten Hesseschen Matrix verallgemeinern lässt.

Definition XI.10. Seien $M \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$ eine offene und nichtleere Teilmenge und $F \in C^2(U; \mathbb{R})$. Ein Punkt $x_0 \in U$ heißt **Sattelpunkt**

$$:\Leftrightarrow x_0 \text{ ist kritisch, } \nabla F(x_0) = 0, \text{ und} \tag{XI.57}$$

$$\exists \delta > 0, z_{\pm} \in \mathbb{R}^M \quad \forall t \in (-\delta, \delta), t \neq 0: \quad F(x_0 + tz_-) < F(x_0) < F(x_0 + tz_+).$$

Bemerkungen und Beispiele.

Ist x_0 ein Extrempunkt von F mit indefiniter Hesseschen Matrix $F''(x_0)$, so ist x_0 auch ein Sattelpunkt von F .

XI.4. Ergänzungen

XI.4.1. Beweis von Korollar XI.2

Korollar XI.2 folgt unmittelbar aus Satz XI.1 und der Identität

$$\frac{1}{p!} \sum_{j_1, \dots, j_p=1}^M \frac{\partial^p F(x)}{\partial x_{j_1} \dots \partial x_{j_p}} = \sum_{\substack{p_1, \dots, p_M=0 \\ p_1 + \dots + p_M = p}}^p \frac{1}{p_1! \dots p_M!} \cdot \frac{\partial^p F(x)}{\partial x_1^{p_1} \dots \partial x_M^{p_M}}, \quad (\text{XI.58})$$

die wir nun für alle $x \in U$ und $p \in \{0, 1, \dots, k+1\}$ durch Induktion in p zeigen wollen. Offenbar sind beide Seiten in (XI.58) gleich $F(x)$, für $p = 0$. Gilt nun (XI.58) für $p \in \{0, 1, \dots, k\}$, so erhalten wir

$$\begin{aligned} \frac{1}{p!} \sum_{j_1, \dots, j_{p+1}=1}^M \frac{\partial^{p+1} F(x)}{\partial x_{j_1} \dots \partial x_{j_p} \partial x_{j_{p+1}}} &= \sum_{j_{p+1}=1}^M \frac{\partial}{\partial x_{j_{p+1}}} \left[\frac{1}{p!} \sum_{j_1, \dots, j_p=1}^M \frac{\partial^p F(x)}{\partial x_{j_1} \dots \partial x_{j_p}} \right] \\ &= \sum_{j=1}^M \frac{\partial}{\partial x_j} \left[\sum_{\substack{p_1, \dots, p_M=0 \\ p_1 + \dots + p_M = p}}^p \frac{1}{p_1! \dots p_M!} \cdot \frac{\partial^p F(x)}{\partial x_1^{p_1} \dots \partial x_j^{p_j} \dots \partial x_M^{p_M}} \right] \\ &= \sum_{j=1}^M \left[\sum_{\substack{p_1, \dots, p_M=0 \\ p_1 + \dots + p_M = p}}^p \frac{1}{p_1! \dots p_M!} \cdot \frac{\partial^{p+1} F(x)}{\partial x_1^{p_1} \dots \partial x_j^{p_j+1} \dots \partial x_M^{p_M}} \right] \\ &\stackrel{\tilde{p}_j := p_{j+1}}{=} \sum_{j=1}^M \left[\sum_{\substack{p_1, \dots, \tilde{p}_j, \dots, p_M=0 \\ p_1 + \dots + \tilde{p}_j + \dots + p_M = p+1}}^p \frac{\mathbb{1}[\tilde{p}_j \geq 1]}{p_1! \dots (\tilde{p}_j - 1)! \dots p_M!} \cdot \frac{\partial^{p+1} F(x)}{\partial x_1^{p_1} \dots \partial x_j^{\tilde{p}_j} \dots \partial x_M^{p_M}} \right] \\ &\stackrel{p_j := \tilde{p}_j}{=} \sum_{j=1}^M \left[\sum_{\substack{p_1, \dots, p_M=0 \\ p_1 + \dots + p_M = p+1}}^p \frac{p_j}{p_1! \dots p_j! \dots p_M!} \cdot \frac{\partial^{p+1} F(x)}{\partial x_1^{p_1} \dots \partial x_j^{p_j} \dots \partial x_M^{p_M}} \right], \quad (\text{XI.59}) \end{aligned}$$

wobei wir

$$\frac{\mathbb{1}[\tilde{p}_j \geq 1]}{(\tilde{p}_j - 1)!} = \frac{\mathbb{1}[p_j \geq 1]}{(p_j - 1)!} = \frac{p_j}{p_j!} \quad (\text{XI.60})$$

verwendet haben. Nun vertauschen wir die Summationen und erhalten (XI.58) für $p+1$:

$$\begin{aligned} \frac{1}{p+1} \sum_{j=1}^M \left[\sum_{\substack{p_1, \dots, p_M=0 \\ p_1 + \dots + p_M = p+1}}^{K+1} \frac{p_j}{p_1! \dots p_M!} \cdot \frac{\partial^{p+1} F(x)}{\partial x_1^{p_1} \dots \partial x_M^{p_M}} \right] & \quad (\text{XI.61}) \\ &= \sum_{\substack{p_1, \dots, p_M=0 \\ p_1 + \dots + p_M = p+1}}^{K+1} \underbrace{\left\{ \frac{1}{p+1} \left[\sum_{j=1}^M p_j \right] \right\}}_{=1} \frac{1}{p_1! \dots p_M!} \cdot \frac{\partial^{p+1} F(x)}{\partial x_1^{p_1} \dots \partial x_M^{p_M}}. \end{aligned}$$

Nach Induktion gilt damit (XI.58) für alle $p \in \{0, 1, \dots, k+1\}$.

XI.4.2. Das Minimax-Prinzip für Matrizen (Lemma XI.7)

Zum Beweis von Lemma XI.7 verwenden wir im Folgenden eine Variante des Spektralsatzes für selbstadjungierte Matrizen, der in der linearen Algebra bewiesen wird. Wir beschränken uns hier auf den Fall reeller, symmetrischer Matrizen.

Satz XI.11 (Spektralsatz). *Seien $M \in \mathbb{N}$ und $A \in \mathfrak{M}_{M \times M}(\mathbb{R})$ eine reelle, selbstadjungierte Matrix, $A_{i,j} = A_{j,i} \in \mathbb{R}$, $\forall i, j \in \{1, \dots, M\}$. Dann sind alle Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_M \in \mathbb{R}$ reell, und es gibt eine ONB $\{\varphi_\alpha\}_{\alpha=1, \dots, M} \subseteq \mathbb{R}^M$ aus zugehörigen Eigenvektoren von A , so dass*

$$\forall \alpha = 1, \dots, M : \quad A \varphi_\alpha = \lambda_\alpha \varphi_\alpha. \quad (\text{XI.62})$$

Explizit gilt mit $\varphi_\alpha = \sum_{i=1}^M \varphi_\alpha(i) e_i$ sogar

$$\forall i, j = 1, \dots, M : \quad A_{i,j} = \sum_{\alpha=1}^M \lambda_\alpha \varphi_\alpha(i) \varphi_\alpha(j). \quad (\text{XI.63})$$

Bemerkungen und Beispiele.

- Wir bemerken, dass die Eigenwerte $\lambda_1, \dots, \lambda_M$ von A durchaus nicht alle paarweise verschieden sein müssen.
- Weiterhin bemerken wir, dass die Transformationsmatrix U , die die kanonische Basis $\{e_1, \dots, e_M\}$ mit $Ue_i = \varphi_i$ auf die ONB $\{\varphi_1, \dots, \varphi_M\}$ aus Eigenvektoren von A abbildet, **orthogonal** ist, $U \in \mathcal{O}(M)$, d.h. $UU^T = U^T U = \mathbb{1}$. Wegen

$$U_{ij} = \varphi_j(i), \quad (U^T)_{kl} = \varphi_k(l), \quad (\text{XI.64})$$

folgt daraus, dass

$$\sum_{\alpha=1}^M \varphi_\alpha(i) \varphi_\alpha(\ell) = (UU^T)_{i\ell} = (\mathbb{1})_{i,\ell} = \delta_{i,\ell}, \quad (\text{XI.65})$$

$$\sum_{j=1}^M \varphi_\alpha(j) \varphi_\beta(j) = (U^T U)_{\alpha\beta} = (\mathbb{1})_{\alpha\beta} = \delta_{\alpha,\beta}. \quad (\text{XI.66})$$

Wir kommen nun zum Beweis des Minimax-Prinzips (XI.39). Aus (XI.63) erhalten wir, mit $x = \sum_{i=1}^M x_i e_i$,

$$\begin{aligned} \langle x | Ax \rangle &= \sum_{i,j=1}^M x_i A_{i,j} x_j = \sum_{\alpha=1}^M \sum_{i,j=1}^M \lambda_\alpha x_i \varphi_\alpha(i) \varphi_\alpha(j) x_j \\ &= \sum_{\alpha=1}^M \lambda_\alpha \underbrace{\left(\sum_{i=1}^M x_i \varphi_\alpha(i) \right)^2}_{\geq 0} \leq \lambda_M \sum_{\alpha=1}^M \left(\sum_{i=1}^M x_i \varphi_\alpha(i) \right)^2 \\ &= \lambda_M \sum_{i,j=1}^M x_i \underbrace{\left(\sum_{\alpha=1}^M \varphi_\alpha(i) \varphi_\alpha(j) \right)}_{=\delta_{i,j}} x_j = \lambda_M \sum_{i=1}^M x_i^2 = \lambda_M \|x\|^2. \end{aligned} \quad (\text{XI.67})$$

Genauso beweist man $\langle x | Ax \rangle \geq \lambda_1 \|x\|^2$. Somit gilt

$$\lambda_1 \leq \min \left\{ \frac{\langle x | Ax \rangle}{\langle x | x \rangle} \mid x \in \mathbb{R}^M \setminus \{0\} \right\} \leq \max \left\{ \frac{\langle x | Ax \rangle}{\langle x | x \rangle} \mid x \in \mathbb{R}^M \setminus \{0\} \right\} \leq \lambda_M. \quad (\text{XI.68})$$

Die in (XI.39) behaupteten Gleichungen folgen nun durch Einsetzen speziell von $x := \varphi_1$ bzw. $x := \varphi_M$.

XII. Maßtheoretische Grundbegriffe

XII.1. Ringe und Figuren

Definition XII.1. Sei Ω eine Menge und $\mathfrak{P}(\Omega) := \{A : A \subseteq \Omega\}$ ihre Potenzmenge. Ein Mengensystem $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ heißt **Ring** $:\Leftrightarrow$

$$\emptyset \in \mathfrak{R}, \quad (\text{XII.1})$$

$$A, B \in \mathfrak{R} \Rightarrow A \setminus B \in \mathfrak{R}, \quad (\text{XII.2})$$

$$A, B \in \mathfrak{R} \Rightarrow A \cup B \in \mathfrak{R}. \quad (\text{XII.3})$$

Wir bemerken, dass aus (XII.2) mit $A, B \in \mathfrak{R}$ auch

$$\begin{aligned} A \cap B &= \{\omega \in A \mid \omega \in B\} = \{\omega \in A \mid \omega \notin (A \setminus B)\} \\ &= A \setminus (A \setminus B) \in \mathfrak{R}. \end{aligned} \quad (\text{XII.4})$$

Damit folgt, dass ein Ring mit endlich vielen Mengen auch deren Vereinigung und Durchschnitt enthält,

$$A_1, A_2, \dots, A_n \in \mathfrak{R} \Rightarrow \bigcup_{j=1}^n A_j, \bigcap_{j=1}^n A_j \in \mathfrak{R}. \quad (\text{XII.5})$$

Bemerkungen und Beispiele.

- Die leere Menge bildet einen Ring, $\mathfrak{R} = \{\emptyset\}$.
- Ebenso erhält man einen Ring, wenn man noch die volle Menge hinzufügt, $\mathfrak{R}' = \{\emptyset, \Omega\}$.
- Natürlich ist die Potenzmenge $\mathfrak{P}(\Omega)$ selbst ein Ring.
- Für das nächste Beispiel erinnern wir an die Halbordnung $\leq$ auf $\mathbb{R}^d$, die für $a, b \in \mathbb{R}^d$ gegeben ist durch

$$a \leq b :\Leftrightarrow a_1 \leq b_1, a_2 \leq b_2, \dots, a_d \leq b_d. \quad (\text{XII.6})$$

Definition XII.2. Sei $\Omega = \mathbb{R}^d$ und $d \geq 1$. Für $a, b \in \mathbb{R}^d$ mit $a \leq b$ heißt

$$[a, b) := [a_1, b_1) \times [a_2, b_2) \times \cdots \times [a_d, b_d) \quad (\text{XII.7})$$

(nach rechts halboffener) **Quader**. Die Menge aller Quader bezeichnen wir mit $\mathfrak{Q}_d$. Die Vereinigung endlich vieler Quader sammeln wir im System der **Figuren**,

$$\mathfrak{F}_d := \{q_1 \cup \cdots \cup q_N \mid q_1, q_2, \dots, q_N \in \mathfrak{Q}_d, N < \infty\}. \quad (\text{XII.8})$$

Lemma XII.3. $\mathfrak{F}_d$ ist ein Ring.

Mit einer Partition analog zur Konstruktion der Quader q_ξ in (XII.38) kann man zeigen, dass sich jede Figur als endliche Vereinigung *disjunkter* Quader schreiben lässt.

Lemma XII.4.

$$\mathfrak{F}_d = \{q_1 \cup \cdots \cup q_N \mid q_1, q_2, \dots, q_N \in \mathfrak{Q}_d, N < \infty, \forall i < j: q_i \cap q_j = \emptyset\}, \quad (\text{XII.9})$$

Beweis. Übungsaufgabe. □

Wir kommen nun zum Begriff des (Raum-)Inhalts.

Definition XII.5. Sei $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ ein Ring. Eine Abbildung $\mu : \mathfrak{R} \rightarrow \mathbb{R}_0^+$ heißt (**endlicher**) **Inhalt** (auf $\mathfrak{R}$) $:\Leftrightarrow$

- (i) $\mu(\emptyset) = 0$.
- (ii) Sind $A_1, A_2, \dots, A_N \in \mathfrak{R}$ paarweise disjunkt, so ist μ **additiv**,

$$\mu\left(\bigcup_{i=1}^N A_i\right) = \sum_{i=1}^N \mu(A_i). \quad (\text{XII.10})$$

Im Allgemeinen bezeichnet man Abbildungen $\mu : \mathfrak{R} \rightarrow [0, \infty] = \mathbb{R}_0^+ \cup \{\infty\}$, die (i) und (ii) erfüllen, als Inhalte. Ein endlicher Inhalt ist dann ein solcher, bei dem $\mu(A) < \infty$ für alle $A \in \mathfrak{R}$ gilt. Wir betrachten hier nur endliche Inhalte, deshalb ignorieren wir in dieser Vorlesung den Unterschied, aber in den Ergänzungen wird zwischen Inhalten und endlichen Inhalten unterschieden.

Das wichtigste Beispiel für Definition XII.5 in unserem Zusammenhang ist das übliche, d -dimensionale Volumen: Für $a, b \in \mathbb{R}^d$, mit $a \leq b$, ist $[a, b) \in \mathfrak{Q}_d$ ein Quader mit Volumen

$$\mu_d([a, b)) := \prod_{\nu=1}^d (b_\nu - a_\nu). \quad (\text{XII.11})$$

Sind weiterhin $q_1, q_2, \dots, q_N \in \mathfrak{Q}_d$ paarweise disjunkte Quader, so setzen wir

$$\mu_d\left(\bigcup_{i=1}^N q_i\right) = \sum_{i=1}^N \mu_d(q_i). \quad (\text{XII.12})$$

Nach Lemma XII.4 ist jede Figur darstellbar als eine endliche Vereinigung disjunkter Quader, und insofern lässt sich jeder Figur durch (XII.12) ein Volumen zuordnen.

Lemma XII.6. μ_d ist ein (endlicher) Inhalt auf $\mathfrak{F}_d$.

XII.2. Sigma-Algebren und Maße

Wir beginnen diesen Abschnitt mit der Einführung einer Vorstufe von Maßen - den Prämaßen.

Definition XII.7. Seien Ω eine Menge und $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ ein Ring. Eine Abbildung $\mu : \mathfrak{R} \rightarrow [0, \infty]$ heißt **Prämaß** (auf $\mathfrak{R}$) : $\Leftrightarrow$

- (i) $\mu(\emptyset) = 0$.
- (ii) Die Abbildung μ ist **Sigma-additiv**, d. h., ist $(A_n)_{n=1}^\infty \in \mathfrak{R}^\mathbb{N}$ eine Folge paarweise disjunkter Mengen in $\mathfrak{R}$, so dass auch ihre Vereinigung in $\mathfrak{R}$ liegt, $\bigcup_{n=1}^\infty A_n \in \mathfrak{R}$, so gilt

$$\mu\left(\bigcup_{n=1}^\infty A_n\right) = \sum_{n=1}^\infty \mu(A_n). \quad (\text{XII.13})$$

Gibt es außerdem eine Folge $(A_n)_{n=1}^\infty$ in $\mathfrak{R}$ mit $\mu(A_n) < \infty$, für alle $n \in \mathbb{N}$, und $A_n \nearrow \Omega$, d.h. $A_n \subseteq A_{n+1}$ und $\bigcup_{n=1}^\infty A_n = \Omega$, so heißt das Prämaß μ **Sigma-endlich**.

Wir sehen, dass jedes Prämaß auf $\mathfrak{R}$ auch einen Inhalt auf $\mathfrak{R}$ definiert. Der folgende Satz gibt eine notwendige Bedingung für die Umkehrung.

Lemma XII.8. Seien Ω eine Menge, $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ ein Ring und $\mu : \mathfrak{R} \rightarrow \mathbb{R}_0^+$ ein Inhalt. Dann gilt folgende Äquivalenz:

$$\mu \text{ ist ein Prämaß auf } \mathfrak{R} \quad (\text{XII.14})$$

$$\Leftrightarrow \mu \text{ ist } \emptyset\text{-stetig} \quad (\text{XII.15})$$

$$: \Leftrightarrow \forall (A_n)_{n=1}^\infty \in \mathfrak{R}^\mathbb{N}, A_n \searrow \emptyset : \lim_{n \rightarrow \infty} \mu(A_n) = 0,$$

wobei $A_n \searrow \emptyset$ bedeutet, dass $A_n \supseteq A_{n+1}$ und $\bigcap_{n=1}^\infty A_n = \emptyset$.

Satz XII.9. μ_d ist ein Prämaß auf $\mathfrak{F}_d$.

Die entscheidende Einsicht, die den Grundbaustein der Maßtheorie darstellt, ist, dass der Begriff des Mengenrings (hier: die Figuren $\mathfrak{F}_1$) nicht flexibel genug ist und durch den der Sigma-Algebra ersetzt werden muss.

Definition XII.10. Sei Ω eine Menge und $\mathfrak{P}(\Omega) := \{A : A \subseteq \Omega\}$ ihre Potenzmenge. Ein Mengensystem $\mathfrak{A} \subseteq \mathfrak{P}(\Omega)$ heißt **Sigma-Algebra** (über Ω) : $\Leftrightarrow$

$$\Omega \in \mathfrak{A}, \quad (\text{XII.16})$$

$$A \in \mathfrak{A} \Rightarrow A^c := \Omega \setminus A \in \mathfrak{A}, \quad (\text{XII.17})$$

$$(A_n)_{n=1}^\infty \subseteq \mathfrak{A} \Rightarrow \bigcup_{n=1}^\infty A_n \in \mathfrak{A}. \quad (\text{XII.18})$$

Bemerkungen und Beispiele.

- Wir bemerken, dass $\mathfrak{A} = \{\emptyset, \Omega\}$ die kleinstmögliche und $\mathfrak{A} = \mathfrak{P}(\Omega)$ die größtmögliche Sigma-Algebra über Ω sind. Der Ausgangspunkt der Maßtheorie ist jedoch die Erkenntnis, dass für $\Omega = \mathbb{R}^d$ der Ring der Figuren $\mathfrak{F}_d$ zu klein und die Potenzmenge $\mathfrak{P}(\Omega)$ zu groß sind.
- Weiterhin bemerken wir, dass jede Sigma-Algebra $\mathfrak{A}$ auch ein Ring ist.

Lemma XII.11. *Jeder (endliche, abzählbare oder auch überabzählbare) Durchschnitt von Sigma-Algebren ist wieder eine Sigma-Algebra.*

Beweis. Der Beweis erfolgt durch Nachprüfen der Eigenschaften (XII.16)–(XII.18). $\square$

Ist nun $\mathcal{E} \subseteq \mathfrak{P}(\Omega)$ ein System von Teilmengen von Ω , so setzen wir

$$\mathfrak{A}(\mathcal{E}) := \bigcap \left\{ \mathfrak{A} \subseteq \mathfrak{P}(\Omega) \mid \mathcal{E} \subseteq \mathfrak{A}, \mathfrak{A} \text{ ist eine Sigma-Algebra} \right\}. \quad (\text{XII.19})$$

Nach Lemma XII.11 ist $\mathfrak{A}(\mathcal{E})$ wieder eine Sigma-Algebra, und nach Konstruktion ist sie die kleinste Sigma-Algebra, die $\mathcal{E}$ enthält.

Definition XII.12.

- Zu $\mathcal{E} \subseteq \mathfrak{P}(\Omega)$ heißt $\mathfrak{A}(\mathcal{E})$ die von $\mathcal{E}$ erzeugte Sigma-Algebra.
- Für $\Omega = \mathbb{R}^d$ heißt die von den Quadern $\mathfrak{Q}_d$ bzw. den Figuren $\mathfrak{F}_d$ erzeugte Sigma-Algebra die Sigma-Algebra der **Borelmengen**,

$$\mathfrak{B}_d := \mathfrak{A}(\mathfrak{Q}_d) = \mathfrak{A}(\mathfrak{F}_d). \quad (\text{XII.20})$$

Bemerkungen und Beispiele.

- Wir bemerken, dass, wenn $\hat{\mathfrak{A}}$ bereits eine Sigma-Algebra ist, die von $\hat{\mathfrak{A}}$ erzeugte Sigma-Algebra mit $\hat{\mathfrak{A}}$ übereinstimmt. Insbesondere ist

$$\mathfrak{A}(\mathfrak{A}(\mathcal{E})) = \mathfrak{A}(\mathcal{E}), \quad (\text{XII.21})$$

für jedes $\mathcal{E} \subseteq \mathfrak{P}(\Omega)$.

- Sind $a, b \in \mathbb{R}$ mit $a < b$, so sind $\left([a, b + \frac{1}{n}]\right)_{n=1}^{\infty}, \left([a + \frac{1}{n}, b)\right)_{n=1}^{\infty} \in \mathfrak{Q}_1^{\mathbb{N}}$ zwei Folgen halboffener Intervalle, und daher sind auch

$$(a, b) = \bigcup_{n=1}^{\infty} [a + \frac{1}{n}, b) \in \mathfrak{B}_1 \quad \text{und} \quad [a, b] = \bigcap_{n=1}^{\infty} [a, b + \frac{1}{n}] \in \mathfrak{B}_1, \quad (\text{XII.22})$$

und somit alle offenen und abgeschlossenen Intervalle Elemente der Sigmaalgebra der Borelmengen $\mathfrak{B}_1$.

- Für $a \in \mathbb{R}$ ist auch die einelementige Menge

$$\{a\} = \bigcap_{n=1}^{\infty} [a, a + \frac{1}{n}) \in \mathfrak{B}_1, \quad (\text{XII.23})$$

eine Borelmenge. Ist nun $(q_n)_{n=1}^{\infty} \in \mathbb{Q}^{\mathbb{N}}$ eine Abzählung der rationalen Zahlen zwischen 0 und 1, also $\{q_n \mid n \in \mathbb{N}\} = \mathbb{Q}$, so sind somit auch

$$\mathbb{Q} = \bigcup_{n=1}^{\infty} \{q_n\} \in \mathfrak{B}_1 \quad \text{und} \quad \mathbb{Q} \cap (0, 1) \in \mathfrak{B}_1. \quad (\text{XII.24})$$

Borelmengen.

Satz XII.13. Seien $\mathcal{O}_d \subseteq \mathfrak{P}(\mathbb{R}^d)$ die offenen, $\mathcal{C}_d \subseteq \mathfrak{P}(\mathbb{R}^d)$ die abgeschlossenen und $\mathcal{K}_d \subseteq \mathfrak{P}(\mathbb{R}^d)$ die kompakten Mengen in $\mathbb{R}^d$. Die von ihnen erzeugten Sigma-Algebren stimmen alle mit den Borelmengen überein,

$$\mathfrak{B}_d = \mathfrak{A}(\mathcal{O}_d) = \mathfrak{A}(\mathcal{C}_d) = \mathfrak{A}(\mathcal{K}_d). \quad (\text{XII.25})$$

Definition XII.14. Sei Ω eine Menge und $\mathfrak{A} \subseteq \mathfrak{P}(\Omega)$ eine Sigma-Algebra. Eine Abbildung $\mu : \mathfrak{A} \rightarrow [0, \infty]$ heißt **Maß** (auf $\mathfrak{A}$): $\Leftrightarrow$

- (i) $\mu(\emptyset) = 0$.
- (ii) Die Abbildung μ ist σ -additiv, d. h., ist $(A_n)_{n=1}^{\infty}$ eine Folge paarweise disjunkter Mengen in $\mathfrak{A}$, so gilt

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n). \quad (\text{XII.26})$$

Gibt es außerdem eine Folge $(A_n)_{n=1}^{\infty}$ in $\mathfrak{A}$ mit $\mu(A_n) < \infty$, für alle $n \in \mathbb{N}$, und $A_n \nearrow \Omega$, d.h. $A_n \subseteq A_{n+1}$ und $\bigcup_{n=1}^{\infty} A_n = \Omega$, so heißt das Maß μ **Sigma-endlich**.

Ein Maß ist also ein Prämaß auf einer Sigma-Algebra. Wir kommen nun zum Hauptresultat dieses Kapitels.

Satz XII.15. Seien Ω eine Menge, $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ ein Ring und $\mu : \mathfrak{R} \rightarrow \mathbb{R}_0^+$ ein Sigma-endliches Prämaß auf $\mathfrak{R}$. Dann besitzt μ eine eindeutige Fortsetzung zu einem Sigma-endlichen Maß $\tilde{\mu} : \mathfrak{A}(\mathfrak{R}) \rightarrow [0, \infty]$, d.h. $\tilde{\mu}(A) = \mu(A)$, für alle $A \in \mathfrak{R}$.

Korollar XII.16. Das auf den Figuren $\mathfrak{F}_d$ definierte Prämaß μ_d hat eine eindeutige Fortsetzung zu einem Sigma-endlichen Maß auf $\mathfrak{B}_d$, das **Lebesgue-Borel-Maß**, das wir abermals mit μ_d bezeichnen.

Bemerkungen und Beispiele.

- Für $a \in \mathbb{R}$ ist

$$\mu_1(\{a\}) = \mu_1([a, a]) = \lim_{n \rightarrow \infty} \mu_1([a, a + \frac{1}{n}]) = \lim_{n \rightarrow \infty} \frac{1}{n} = 0. \quad (\text{XII.27})$$

Einelementige Borelmengen haben also Maß null.

- Aus der Sigmaadditivität des Lebesgue-Borelmaßes folgt aber auch, dass

$$\mu_1(\mathbb{Q}) = \mu_1\left(\bigcup_{n=1}^{\infty} \{q_n\}\right) = \sum_{n=1}^{\infty} \mu_1(\{q_n\}) = 0, \quad (\text{XII.28})$$

d.h. die Borelmenge der rationalen Zahlen hat ebenfalls Maß null.

Definition XII.17. Sind Ω eine Menge und $\mathfrak{A} \subseteq \mathfrak{P}(\Omega)$ eine Sigma-Algebra, so bezeichnet man das Paar $(\Omega, \mathfrak{A})$ als **Messraum** oder **messbaren Raum**, und die Mengen in $\mathfrak{A}$ heißen **messbar**. Ist weiterhin $\mu : \mathfrak{A} \rightarrow [0, \infty]$ ein Maß, so bezeichnet man das Tripel $(\Omega, \mathfrak{A}, \mu)$ als **Maßraum**.

Insbesondere ist nach Korollar XII.16 das Tripel $(\mathbb{R}^d, \mathfrak{B}_d, \mu_d)$ ein Maßraum.

Wir bemerken noch, dass sich das sogenannte *Lebesgue-Maß* vom Lebesgue-Borel-Maß in einem wichtigen Punkt unterscheidet: Ist $A \in \mathfrak{B}_d$ eine messbare Menge mit Maß Null, $\mu_d(A) = 0$, so gilt wegen der Monotonie des (in den Ergänzungen eingeführten) äußeren Maßes auch $\mu_d^*(A') = 0$, für jede Teilmenge $A' \subseteq A$. Trotzdem muss A' nicht notwendig messbar sein, und obwohl 0 ein natürlicher Kandidat für $\mu_d(A')$ ist, wäre dann $\mu_d(A')$ nicht definiert. Die Beseitigung dieses Nachteils führt auf das *Lebesgue-Maß*. Wir verzichten aber hier darauf, da es die Einführung neuer Begriffe erforderlich macht. Vielmehr ignorieren wir im Folgenden diesen Unterschied und sprechen vom *Lebesgue-Maß*, obwohl wir stets das *Lebesgue-Borel-Maß* meinen.

XII.3. Ergänzungen

XII.3.1. Ringe und Figuren

Definition XII.18. Sei Ω eine Menge und $\mathfrak{P}(\Omega) := \{A : A \subseteq \Omega\}$ ihre Potenzmenge. Ein Mengensystem $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ heißt **Ring** : $\Leftrightarrow$

$$\emptyset \in \mathfrak{R}, \quad (\text{XII.29})$$

$$A, B \in \mathfrak{R} \Rightarrow A \setminus B \in \mathfrak{R}, \quad (\text{XII.30})$$

$$A, B \in \mathfrak{R} \Rightarrow A \cup B \in \mathfrak{R}. \quad (\text{XII.31})$$

Wir bemerken, dass aus (XII.30) mit $A, B \in \mathfrak{R}$ auch

$$\begin{aligned} A \cap B &= \{\omega \in A \mid \omega \in B\} = \{\omega \in A \mid \omega \notin (A \setminus B)\} \\ &= A \setminus (A \setminus B) \in \mathfrak{R}. \end{aligned} \quad (\text{XII.32})$$

Damit folgt, dass ein Ring mit endlich vielen Mengen auch deren Vereinigung und Durchschnitt enthält,

$$A_1, A_2, \dots, A_n \in \mathfrak{R} \Rightarrow \bigcup_{j=1}^n A_j, \bigcap_{j=1}^n A_j \in \mathfrak{R}. \quad (\text{XII.33})$$

Bemerkungen und Beispiele.

- Die leere Menge bildet einen Ring, $\mathfrak{R} = \{\emptyset\}$.
- Ebenso erhält man einen Ring, wenn man noch die volle Menge hinzufügt, $\mathfrak{R}' = \{\emptyset, \Omega\}$.
- Natürlich ist die Potenzmenge $\mathfrak{P}(\Omega)$ selbst ein Ring.
- Für das nächste Beispiel erinnern wir an die Halbordnung $\leq$ auf $\mathbb{R}^d$, die für $a, b \in \mathbb{R}^d$ gegeben ist durch

$$a \leq b :\Leftrightarrow a_1 \leq b_1, a_2 \leq b_2, \dots, a_d \leq b_d. \quad (\text{XII.34})$$

Definition XII.19. Sei $\Omega = \mathbb{R}^d$ und $d \geq 1$. Für $a, b \in \mathbb{R}^d$ mit $a \leq b$ heißt

$$[a, b) := [a_1, b_1) \times [a_2, b_2) \times \dots \times [a_d, b_d) \quad (\text{XII.35})$$

(nach rechts halboffener) **Quader**. Die Menge aller Quader bezeichnen wir mit $\mathfrak{Q}_d$. Die Vereinigung endlich vieler Quader sammeln wir im System der **Figuren**,

$$\mathfrak{F}_d := \{q_1 \cup \dots \cup q_N \mid q_1, q_2, \dots, q_N \in \mathfrak{Q}_d, N < \infty\}. \quad (\text{XII.36})$$

Lemma XII.20. $\mathfrak{F}_d$ ist ein Ring.

Beweis. Offensichtlich gelten (XII.29) und (XII.31) trivialerweise, und die einzig nachzuweisende Eigenschaft ist (XII.30).

Dazu zeigen wir zunächst, dass die Differenz zweier Quader eine Figur ist. Seien $a, b, a', b' \in \mathbb{R}^d$ mit $a \leq b$ und $a' \leq b'$. Wir zeigen, dass $[a, b) \setminus [a', b') \in \mathfrak{F}_d$. Dazu ordnen wir die Koordinaten $a_\nu, b_\nu, a'_\nu, b'_\nu$ nach der Größe, d. h., für alle $\nu = 1, 2, \dots, d$, definieren wir $-\infty < x_\nu^{(1)} < \dots < x_\nu^{(k_\nu)} < \infty$, durch $\{x_\nu^{(1)}, \dots, x_\nu^{(k_\nu)}\} = \{a_\nu, b_\nu, a'_\nu, b'_\nu\}$, wobei $1 \leq k_\nu \leq 4$, je nachdem, wie viele Koordinaten doppelt auftreten. Anschließend setzen wir

$$\mathcal{K} := \{1, \dots, k_1 - 1\} \times \{1, \dots, k_2 - 1\} \times \dots \times \{1, \dots, k_d - 1\} \quad (\text{XII.37})$$

und, für $\xi := (\xi_1, \dots, \xi_d) \in \mathcal{K}$,

$$q_\xi := [x_1^{(\xi_1)}, x_1^{(\xi_1+1)}) \times [x_2^{(\xi_2)}, x_2^{(\xi_2+1)}) \times \dots \times [x_d^{(\xi_d)}, x_d^{(\xi_d+1)}). \quad (\text{XII.38})$$

Für $\xi \neq \xi'$ gilt offensichtlich $q_\xi \cap q_{\xi'} = \emptyset$. Außerdem sind

$$[a, b) = \bigcup_{\xi \in \mathcal{K}: q_\xi \cap [a, b) \neq \emptyset} q_\xi, \quad [a', b') = \bigcup_{\xi \in \mathcal{K}: q_\xi \cap [a', b') \neq \emptyset} q_\xi, \quad (\text{XII.39})$$

und insbesondere ist die Differenz der Quader $[a, b)$ und $[a', b')$ eine Figur,

$$[a, b) \setminus [a', b') = \bigcup_{\xi \in \mathcal{K}: q_\xi \cap ([a, b) \setminus [a', b')) \neq \emptyset} q_\xi \in \mathfrak{F}_d. \quad (\text{XII.40})$$

Seien nun $\check{H} \in \mathfrak{F}_d$ eine Figur und $\hat{q}_1, \dots, \hat{q}_M \in \mathfrak{Q}_d$ so, dass $\check{H} = \bigcup_{i=1}^M \hat{q}_i$.

Sei weiterhin $q \in \mathfrak{Q}_d$. Nach (XII.40) und (XII.31) ist dann

$$\check{H} \setminus q = \left(\bigcup_{i=1}^M \hat{q}_i \right) \setminus q = \bigcup_{i=1}^M (\hat{q}_i \setminus q) \in \mathfrak{F}_d, \quad (\text{XII.41})$$

d.h., die Differenz einer Figur und eines Quaders ist eine Figur. Sind nun $f \in \mathfrak{F}_d$ eine Figur und $q_1, \dots, q_N \in \mathfrak{Q}_d$ so, dass $f = \bigcup_{j=1}^N q_j$, dann ist nach (XII.41) auch

$$\check{H} \setminus f = \check{H} \setminus \left(\bigcup_{j=1}^N q_j \right) = (\dots (\check{H} \setminus q_1) \setminus q_2) \setminus \dots \setminus q_N \in \mathfrak{F}_d. \quad (\text{XII.42})$$

□

Mit einer Partition analog zur Konstruktion der Quader q_ξ in (XII.38) kann man zeigen, dass sich jede Figur als endliche Vereinigung *disjunkter* Quader schreiben lässt.

Lemma XII.21.

$$\mathfrak{F}_d = \left\{ q_1 \cup \dots \cup q_N \mid q_1, q_2, \dots, q_N \in \mathfrak{Q}_d, N < \infty, \quad \forall i < j: q_i \cap q_j = \emptyset \right\}, \quad (\text{XII.43})$$

Beweis. Übungsaufgabe.

□

Wir kommen nun zum Begriff des (Raum-)Inhalts.

Definition XII.22. Sei $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ ein Ring. Eine Abbildung $\mu : \mathfrak{R} \rightarrow \mathbb{R}_0^+$ heißt **Inhalt** (auf $\mathfrak{R}$) $:\Leftrightarrow$

- (i) $\mu(\emptyset) = 0$.
- (ii) Sind $A_1, A_2, \dots, A_N \in \mathfrak{R}$ paarweise disjunkt, so ist μ **additiv**,

$$\mu\left(\bigcup_{i=1}^N A_i\right) = \sum_{i=1}^N \mu(A_i). \quad (\text{XII.44})$$

Das wichtigste Beispiel für Definition XII.22 in unserem Zusammenhang ist das übliche, d -dimensionale Volumen: Für $a, b \in \mathbb{R}^d$, mit $a \leq b$, ist $[a, b) \in \mathfrak{Q}_d$ ein Quader mit Volumen

$$\mu_d([a, b)) := \prod_{\nu=1}^d (b_\nu - a_\nu). \quad (\text{XII.45})$$

Sind weiterhin $q_1, q_2, \dots, q_N \in \mathfrak{Q}_d$ paarweise disjunkte Quader, so setzen wir

$$\mu_d\left(\bigcup_{i=1}^N q_i\right) = \sum_{i=1}^N \mu_d(q_i). \quad (\text{XII.46})$$

Nach Lemma XII.21 ist jede Figur darstellbar in eine endliche Vereinigung disjunkter Quader, und insofern lässt sich jeder Figur durch (XII.46) ein Volumen zuordnen.

Lemma XII.23. μ_d ist ein Inhalt auf $\mathfrak{F}_d$.

Beweis. Definition XII.22(i) und (ii) gelten trivialerweise, falls die Wohldefiniertheit von $\mu_d : \mathfrak{F}_d \rightarrow [0, \infty]$ gesichert ist. Dazu ist zu bemerken, dass die Zerlegung einer Figur in disjunkte Quader keineswegs eindeutig ist. Sind $q_1, \dots, q_M \in \mathfrak{Q}_d$ und $\hat{q}_1, \dots, \hat{q}_N \in \mathfrak{Q}_d$ jeweils paarweise disjunkt, ergeben aber in der Vereinigung dieselbe Menge,

$$\bigcup_{i=1}^M q_i = \bigcup_{j=1}^N \hat{q}_j, \quad (\text{XII.47})$$

so bleibt zu zeigen, dass auch die zugeordneten Volumina gleich sind,

$$\sum_{i=1}^M \mu_d(q_i) = \mu_d\left(\bigcup_{i=1}^M q_i\right) = \mu_d\left(\bigcup_{j=1}^N \hat{q}_j\right) = \sum_{j=1}^N \mu_d(\hat{q}_j). \quad (\text{XII.48})$$

Dafür bemerken wir, dass $\tilde{q}_{ij} := q_i \cap \hat{q}_j \in \mathfrak{Q}_d$ eine Familie disjunkter Quader definiert, und dass

$$q_i = \bigcup_{j=1}^N \tilde{q}_{ij} \quad \text{und} \quad \hat{q}_j = \bigcup_{i=1}^M \tilde{q}_{ij} \quad (\text{XII.49})$$

für alle $i = 1, \dots, M$ und alle $j = 1, \dots, N$ gilt. Also ist

$$\begin{aligned} \sum_{i=1}^M \mu_d(q_i) &= \sum_{i=1}^M \mu_d\left(\bigcup_{j=1}^N \tilde{q}_{ij}\right) = \sum_{i=1}^M \sum_{j=1}^N \mu_d(\tilde{q}_{ij}) \\ &= \sum_{j=1}^N \mu_d\left(\bigcup_{i=1}^M \tilde{q}_{ij}\right) = \sum_{j=1}^N \mu_d(\hat{q}_j), \end{aligned} \quad (\text{XII.50})$$

und μ_d ist auf den Figuren wohldefiniert. $\square$

XII.3.2. Prämaße

Definition XII.24. Seien Ω eine Menge und $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ ein Ring. Eine Abbildung $\mu : \mathfrak{R} \rightarrow [0, \infty]$ heißt **Prämaß** (auf $\mathfrak{R}$) $:\Leftrightarrow$

- (i) $\mu(\emptyset) = 0$.
- (ii) Die Abbildung μ ist **Sigma-additiv**, d. h., ist $(A_n)_{n=1}^\infty \in \mathfrak{R}^\mathbb{N}$ eine Folge paarweise disjunkter Mengen in $\mathfrak{R}$, so dass auch ihre Vereinigung in $\mathfrak{R}$ liegt, $\bigcup_{n=1}^\infty A_n \in \mathfrak{R}$, so gilt

$$\mu\left(\bigcup_{n=1}^\infty A_n\right) = \sum_{n=1}^\infty \mu(A_n). \quad (\text{XII.51})$$

Gibt es außerdem eine Folge $(A_n)_{n=1}^\infty$ in $\mathfrak{R}$ mit $\mu(A_n) < \infty$, für alle $n \in \mathbb{N}$, und $A_n \nearrow \Omega$, d.h. $A_n \subseteq A_{n+1}$ und $\bigcup_{n=1}^\infty A_n = \Omega$, so heißt das Prämaß μ **Sigma-endlich**.

Wir sehen, dass jedes Prämaß auf $\mathfrak{R}$ auch einen Inhalt auf $\mathfrak{R}$ definiert. Der folgende Satz gibt eine notwendige Bedingung für die Umkehrung.

Lemma XII.25. Seien Ω eine Menge, $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ ein Ring und $\mu : \mathfrak{R} \rightarrow \mathbb{R}_0^+$ ein Inhalt. Dann gilt folgende Äquivalenz:

$$\mu \text{ ist ein Prämaß auf } \mathfrak{R} \quad (\text{XII.52})$$

$$\Leftrightarrow \mu \text{ ist } \emptyset\text{-stetig} \quad (\text{XII.53})$$

$$: \Leftrightarrow \forall (A_n)_{n=1}^\infty \in \mathfrak{R}^\mathbb{N}, A_n \searrow \emptyset : \lim_{n \rightarrow \infty} \mu(A_n) = 0,$$

wobei $A_n \searrow \emptyset$ bedeutet, dass $A_n \supseteq A_{n+1}$ und $\bigcap_{n=1}^\infty A_n = \emptyset$.

Beweis.

(i): Sei $(B_n)_{n=1}^\infty$ eine Folge in $\mathfrak{R}$ mit $B_n \searrow \emptyset$, d.h. $B_n \supseteq B_{n+1}$ und $\bigcap_{n=1}^\infty B_n = \emptyset$. Bilden wir $A_n := B_n \setminus B_{n+1}$, so ist $(A_n)_{n=1}^\infty$ eine Folge paarweise disjunkter Mengen in $\mathfrak{R}$, und B_1 lässt sich als disjunkte Vereinigung für alle $N \in \mathbb{N}$ wie folgt darstellen,

$$B_1 = \bigcup_{n=1}^N A_n \cup B_{N+1} = \bigcup_{n=1}^\infty A_n. \quad (\text{XII.54})$$

Weil μ ein Prämaß ist, folgt daraus, dass

$$\mu(B_1) = \mu\left(\bigcup_{n=1}^N A_n\right) + \mu(B_{N+1}) = \sum_{n=1}^N \mu(A_n) + \mu(B_{N+1}) \quad \text{und} \quad (\text{XII.55})$$

$$\mu(B_1) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) + \mu(B_{N+1}) = \sum_{n=1}^{\infty} \mu(A_n). \quad (\text{XII.56})$$

Insbesondere ist die Reihe $\sum_{n=1}^{\infty} \mu(A_n)$ konvergent, und deshalb muss

$$\mu(B_{N+1}) = \sum_{n=N+1}^{\infty} \mu(A_n) \rightarrow 0 \quad (\text{XII.57})$$

konvergieren, für $N \rightarrow \infty$.

(ii): Sei $(A_n)_{n=1}^{\infty}$ eine Folge paarweise disjunkter Mengen in $\mathfrak{R}$, so dass auch $A := \bigcup_{n=1}^{\infty} A_n \in \mathfrak{R}$. Setzen wir $\tilde{B}_N := \bigcup_{n=1}^N A_n$ und $B_N := \bigcup_{n=N+1}^{\infty} A_n$, so sind $\tilde{B}_N$ und B_N disjunkt und $A = \tilde{B}_N \cup B_N$. Offensichtlich sind $\tilde{B}_N \in \mathfrak{R}$ und somit auch $B_N = A \setminus \tilde{B}_N \in \mathfrak{R}$. Da μ ein endlicher Inhalt ist, sind

$$\mu(A), \mu(B_N), \mu(\tilde{B}_N) < \infty \quad (\text{XII.58})$$

endlich, und die Disjunktheit von $\tilde{B}_N$ und B_N impliziert, dass $\mu(A) = \mu(\tilde{B}_N) + \mu(B_N)$ bzw.

$$\forall N \in \mathbb{N} : \quad \mu(B_N) = \mu(A) - \mu(\tilde{B}_N). \quad (\text{XII.59})$$

Außerdem impliziert die Disjunktheit der Mengen A_n , dass $B_N \searrow \emptyset$, für $N \rightarrow \infty$, denn zu jedem $x \in A$ gibt es genau ein $n(x) \in \mathbb{N}$, so dass $x \in A_{n(x)}$, und somit ist $x \notin B_N$, für alle $N > n(x)$. Aus der $\emptyset$ -Stetigkeit und (XII.59) folgt deshalb

$$0 = \lim_{N \rightarrow \infty} \mu(B_N) = \lim_{N \rightarrow \infty} \mu(A \setminus \tilde{B}_N) \quad (\text{XII.60})$$

$$\begin{aligned} &= \mu(A) - \lim_{N \rightarrow \infty} \mu(\tilde{B}_N) = \mu(A) - \lim_{N \rightarrow \infty} \mu\left(\bigcup_{n=1}^N A_n\right) \\ &= \mu(A) - \lim_{N \rightarrow \infty} \sum_{n=1}^N \mu(A_n) = \mu(A) - \sum_{n=1}^{\infty} \mu(A_n), \end{aligned}$$

was gerade die Sigma-Additivität ergibt. □

In dem nun folgenden Satz verwenden wir noch folgende elementare Identität,

Lemma XII.26. *Seien Ω eine Menge, $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ ein Ring, $\mu : \mathfrak{R} \rightarrow [0, \infty]$ ein Inhalt und $A, B \in \mathfrak{R}$. Dann gilt*

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B). \quad (\text{XII.61})$$

Beweis. Wir beobachten, dass in

$$A \cup B = A \cup (B \setminus A) \quad \text{und} \quad B = (A \cap B) \cup (B \setminus A) \quad (\text{XII.62})$$

auf den rechten Seiten jeweils zwei disjunkte Mengen vereinigt sind. Aus der Additivität des Inhalts folgt damit

$$\mu(A \cup B) - \mu(A) = \mu(B \setminus A) = \mu(B) - \mu(A \cap B), \quad (\text{XII.63})$$

woraus sich sofort Glg. (XII.61) ergibt. $\square$

Satz XII.27. μ_d ist ein Prämaß auf $\mathfrak{F}_d$.

Beweis. Da wir in Lemma XII.23 bereits gezeigt haben, dass μ_d einen (offensichtlich endlichen) Inhalt auf $\mathfrak{F}_d$ definiert, bleibt nach Lemma XII.25 nur die $\emptyset$ -Stetigkeit von μ_d zu zeigen. Sei dazu $(A_n)_{n=1}^\infty$ eine monoton fallende Folge in $\mathfrak{F}_d$, d.h. $A_n \supseteq A_{n+1}$. Es reicht zu zeigen, dass

$$\delta := \lim_{n \rightarrow \infty} \mu_d(A_n) = \inf_{n \in \mathbb{N}} \mu_d(A_n) > 0, \quad (\text{XII.64})$$

die Aussage

$$\bigcap_{n=1}^\infty A_n \neq \emptyset \quad (\text{XII.65})$$

impliziert, wobei wir o. B. d. A. $\delta \leq 1$ annehmen können. Wir fixieren nun $n \in \mathbb{N}$ und wählen nichtleere, disjunkte Quader $q^{(1)}, \dots, q^{(L)}$ in $\mathfrak{Q}_d$ so, dass

$$A_n = \bigcup_{\ell=1}^L q^{(\ell)}, \quad (\text{XII.66})$$

wobei $q^{(\ell)} = x^{(\ell)} + [-\varepsilon^{(\ell)}, \varepsilon^{(\ell)})$, für geeignete $x^{(\ell)} \in \mathbb{R}^d$ und $\varepsilon^{(\ell)} \in (\mathbb{R}^+)^d$. Anschließend definieren wir etwas kleinere Quader

$$\hat{q}^{(\ell)} := x^{(\ell)} + \left(1 - \frac{\delta 2^{-n-1-\ell}}{1 + \mu_d(q^{(\ell)})}\right)^{1/d} [-\varepsilon^{(\ell)}, \varepsilon^{(\ell)}) \in \mathfrak{Q}_d \quad (\text{XII.67})$$

und setzen

$$B_n := \bigcup_{\ell=1}^L \hat{q}^{(\ell)}. \quad (\text{XII.68})$$

Wir beobachten, dass der Abschluss von B_n kompakt ist und dass

$$\overline{B_n} = \bigcup_{\ell=1}^L \overline{\hat{q}^{(\ell)}} \subseteq A_n. \quad (\text{XII.69})$$

Weiterhin beobachten wir, dass

$$\begin{aligned}
 \mu_d(A_n) - \mu_d(B_n) &= \sum_{\ell=1}^L \left\{ \mu_d(q^{(\ell)}) - \left(1 - \frac{\delta 2^{-n-1-\ell}}{1 + \mu_d(q^{(\ell)})} \right) \mu_d(q^{(\ell)}) \right\} \\
 &= \sum_{\ell=1}^L \delta 2^{-n-1-\ell} \left(\frac{\mu_d(q^{(\ell)})}{1 + \mu_d(q^{(\ell)})} \right) \\
 &\leq \sum_{\ell=1}^L \delta 2^{-n-1-\ell} \leq \delta 2^{-n-1}.
 \end{aligned} \tag{XII.70}$$

Im nächsten Schritt definieren wir mit Hilfe der Mengen $B_n \in \mathfrak{F}_d$ eine weitere Folge $(C_N)_{N=1}^\infty$ von Figuren in $\mathfrak{F}_d$ durch

$$C_N := \bigcap_{n=1}^N B_n = C_{N-1} \cap B_N. \tag{XII.71}$$

Daraus ergibt sich mit (XII.61), dass

$$\mu_d(C_N) = \mu_d(C_{N-1} \cap B_N) = \mu_d(C_{N-1}) + \mu_d(B_N) - \mu_d(C_{N-1} \cup B_N). \tag{XII.72}$$

Nun ist $C_{N-1} \subseteq B_{N-1} \subseteq A_{N-1}$ und auch $B_N \subseteq A_N \subseteq A_{N-1}$, also $C_{N-1} \cup B_N \subseteq A_{N-1}$. Setzen wir dies und (XII.70) in (XII.72) ein, so erhalten wir

$$\begin{aligned}
 \mu_d(C_N) &\geq \mu_d(C_{N-1}) + \mu_d(B_N) - \mu_d(A_{N-1}) \\
 &\geq \mu_d(C_{N-1}) + \mu_d(A_N) - \mu_d(A_{N-1}) - \delta 2^{-N-1},
 \end{aligned} \tag{XII.73}$$

woraus wir schließen, dass

$$\begin{aligned}
 \mu_d(C_N) - \mu_d(A_N) &\tag{XII.74} \\
 &\geq \mu_d(C_{N-1}) - \mu_d(A_{N-1}) - \delta 2^{-(N-1)-2} \\
 &\geq \mu_d(C_{N-2}) - \mu_d(A_{N-2}) - \delta (2^{-(N-1)-2} + 2^{-(N-2)-2}) \\
 &\cdots \geq \mu_d(C_1) - \mu_d(A_1) - \delta \sum_{\ell=1}^{N-1} 2^{-2-\ell} \\
 &= \mu_d(B_1) - \mu_d(A_1) - \delta \sum_{\ell=1}^{N-1} 2^{-2-\ell} \\
 &\geq -\delta \sum_{\ell=0}^{\infty} 2^{-2-\ell} = -\delta 2^{-1},
 \end{aligned}$$

da $C_1 = B_1$. Wegen $\mu_d(A_N) \geq \delta$ impliziert dies, dass

$$\forall N \in \mathbb{N} : \quad \mu_d(C_N) \geq \frac{\delta}{2} > 0. \tag{XII.75}$$

was wiederum zur Folge hat, dass $C_N \neq \emptyset$, für alle $N \in \mathbb{N}$.

Schließlich stellen wir fest, dass $(\overline{C_N})_{N=1}^\infty$ eine monotone Folge nichtleerer, kompakter Mengen ist, $\overline{C_N} \supseteq \overline{C_{N+1}} \neq \emptyset$, was nach dem Satz von Heine-Borel impliziert, dass auch der Durchschnitt aller $\overline{C_N}$ nichtleer ist. Mit $\overline{C_n} \subseteq \overline{B_n} \subseteq A_n$ folgt also

$$\bigcap_{n=1}^\infty A_n \supseteq \bigcap_{n=1}^\infty \overline{B_n} \supseteq \bigcap_{n=1}^\infty \overline{C_n} \neq \emptyset. \quad (\text{XII.76})$$

□

XII.3.3. Sigma-Algebren

Die entscheidende Einsicht, die den Grundbaustein der Maßtheorie darstellt, ist, dass der Begriff des Mengengerings (hier: die Figuren $\mathfrak{F}_1$) nicht flexibel genug ist und durch den der Sigma-Algebra ersetzt werden muss.

Definition XII.28. Sei Ω eine Menge und $\mathfrak{P}(\Omega) := \{A : A \subseteq \Omega\}$ ihre Potenzmenge. Ein Mengensystem $\mathfrak{A} \subseteq \mathfrak{P}(\Omega)$ heißt **Sigma-Algebra** (über Ω) : $\Leftrightarrow$

$$\Omega \in \mathfrak{A}, \quad (\text{XII.77})$$

$$A \in \mathfrak{A} \Rightarrow A^c := \Omega \setminus A \in \mathfrak{A}, \quad (\text{XII.78})$$

$$(A_n)_{n=1}^\infty \subseteq \mathfrak{A} \Rightarrow \bigcup_{n=1}^\infty A_n \in \mathfrak{A}. \quad (\text{XII.79})$$

Bemerkungen und Beispiele.

- Wir bemerken, dass $\mathfrak{A} = \{\emptyset, \Omega\}$ die kleinstmögliche und $\mathfrak{A} = \mathfrak{P}(\Omega)$ die größtmögliche Sigma-Algebra über Ω sind. Der Ausgangspunkt der Maßtheorie ist jedoch die Erkenntnis, dass für $\Omega = \mathbb{R}^d$ der Ring der Figuren $\mathfrak{F}_d$ zu klein und die Potenzmenge $\mathfrak{P}(\Omega)$ zu groß sind.
- Weiterhin bemerken wir, dass jede Sigma-Algebra $\mathfrak{A}$ auch ein Ring ist. Denn mit $\Omega \in \mathfrak{A}$ ist auch $\emptyset = \Omega^c \in \mathfrak{A}$; für $A, B \in \mathfrak{A}$ setzen wir $A_1 := A$, $A_2 := B$ und $A_n := \emptyset$, für alle $n \geq 2$. Dann ist $(A_n)_{n=1}^\infty$ eine Folge in $\mathfrak{A}$, und deshalb ist auch $A \cup B = \bigcup_{n=1}^\infty A_n \in \mathfrak{A}$. Schließlich sind mit $A, B \in \mathfrak{A}$ auch $A^c, B^c \in \mathfrak{A}$. Dann ist auch $A^c \cup B^c \in \mathfrak{A}$ und somit auch

$$A \setminus B = A \cap B^c = [A^c \cup B]^c \in \mathfrak{A}. \quad (\text{XII.80})$$

Lemma XII.29. Jeder (endliche, abzählbare oder auch überabzählbare) Durchschnitt von Sigma-Algebren ist wieder eine Sigma-Algebra.

Beweis. Der Beweis erfolgt durch Nachprüfen der Eigenschaften (XII.77)–(XII.79). □

Ist nun $\mathcal{E} \subseteq \mathfrak{P}(\Omega)$ ein System von Teilmengen von Ω , so setzen wir

$$\mathfrak{A}(\mathcal{E}) := \bigcap \left\{ \mathfrak{A} \subseteq \mathfrak{P}(\Omega) \mid \mathcal{E} \subseteq \mathfrak{A}, \mathfrak{A} \text{ ist eine Sigma-Algebra} \right\}. \quad (\text{XII.81})$$

Nach Lemma XII.29 ist $\mathfrak{A}(\mathcal{E})$ wieder eine Sigma-Algebra, und nach Konstruktion ist sie die kleinste Sigma-Algebra, die $\mathcal{E}$ enthält.

Definition XII.30.

- (i) Zu $\mathcal{E} \subseteq \mathfrak{P}(\Omega)$ heißt $\mathfrak{A}(\mathcal{E})$ die **von $\mathcal{E}$ erzeugte Sigma-Algebra**.
- (ii) Für $\Omega = \mathbb{R}^d$ heißt die von den Quadern $\mathfrak{Q}_d$ bzw. den Figuren $\mathfrak{F}_d$ erzeugte Sigma-Algebra die Sigma-Algebra der **Borelmengen**,

$$\mathfrak{B}_d := \mathfrak{A}(\mathfrak{Q}_d) = \mathfrak{A}(\mathfrak{F}_d). \quad (\text{XII.82})$$

Wir bemerken, dass, wenn $\widehat{\mathfrak{A}}$ bereits eine Sigma-Algebra ist, die von $\widehat{\mathfrak{A}}$ erzeugte Sigma-Algebra mit $\widehat{\mathfrak{A}}$ übereinstimmt. Insbesondere ist

$$\mathfrak{A}(\mathfrak{A}(\mathcal{E})) = \mathfrak{A}(\mathcal{E}), \quad (\text{XII.83})$$

für jedes $\mathcal{E} \subseteq \mathfrak{P}(\Omega)$.

Satz XII.31. Seien $\mathcal{O}_d \subseteq \mathfrak{P}(\mathbb{R}^d)$ die offenen, $\mathcal{C}_d \subseteq \mathfrak{P}(\mathbb{R}^d)$ die abgeschlossenen und $\mathcal{K}_d \subseteq \mathfrak{P}(\mathbb{R}^d)$ die kompakten Mengen in $\mathbb{R}^d$. Die von ihnen erzeugten Sigma-Algebren stimmen alle mit den Borelmengen überein,

$$\mathfrak{B}_d = \mathfrak{A}(\mathcal{O}_d) = \mathfrak{A}(\mathcal{C}_d) = \mathfrak{A}(\mathcal{K}_d). \quad (\text{XII.84})$$

Beweis. In $\mathbb{R}^d$ ist eine Menge genau dann kompakt, wenn sie beschränkt und abgeschlossen ist; das ist die Aussage des Satzes von Heine-Borel. Also ist $\mathcal{K}_d \subseteq \mathcal{C}_d$, und deshalb auch $\mathfrak{A}(\mathcal{K}_d) \subseteq \mathfrak{A}(\mathcal{C}_d)$. Sind umgekehrt $A \in \mathcal{C}_d$ eine abgeschlossene Menge und $\bar{B}_r \in \mathcal{K}_d$ die abgeschlossene Kugel in $\mathbb{R}^d$ um den Ursprung und mit Radius $r \in \mathbb{R}^+$, so sind auch $A_n := A \cap \bar{B}_n \in \mathcal{K}_d$, für alle $n \in \mathbb{N}$, und

$$A = \bigcup_{n=1}^{\infty} A_n \in \mathfrak{A}(\mathcal{K}_d). \quad (\text{XII.85})$$

Also ist $\mathcal{C}_d \subseteq \mathfrak{A}(\mathcal{K}_d)$, und somit gilt nach (XII.83))

$$\mathfrak{A}(\mathcal{K}_d) \subseteq \mathfrak{A}(\mathcal{C}_d) \subseteq \mathfrak{A}(\mathfrak{A}(\mathcal{K}_d)) \subseteq \mathfrak{A}(\mathcal{K}_d). \quad (\text{XII.86})$$

Also ist $\mathfrak{A}(\mathcal{K}_d) = \mathfrak{A}(\mathcal{C}_d)$.

Ist $A \in \mathcal{C}_d$ abgeschlossen, so ist $A^c \in \mathcal{O}_d$ offen. Weil $\mathfrak{A}(\mathcal{O}_d)$ mit A^c auch dessen Komplement A enthält, folgt also, dass $\mathcal{C}_d \subseteq \mathfrak{A}(\mathcal{O}_d)$, was nach (XII.83) wiederum $\mathfrak{A}(\mathcal{C}_d) \subseteq \mathfrak{A}(\mathcal{O}_d)$ nach sich zieht. Dasselbe Argument liefert aber auch $\mathfrak{A}(\mathcal{O}_d) \subseteq \mathfrak{A}(\mathcal{C}_d)$, und somit erhalten wir

$$\mathfrak{A}(\mathcal{K}_d) = \mathfrak{A}(\mathcal{C}_d) = \mathfrak{A}(\mathcal{O}_d). \quad (\text{XII.87})$$

Wir zeigen nun noch, dass $\mathfrak{B}_d = \mathfrak{A}(\mathcal{O}_d)$. Seien dazu $a, b \in \mathbb{R}^d$ mit $a < b$, d.h. $a_1 < b_1, \dots, a_d < b_d$. Zu $n \in \mathbb{N}$ definieren wir

$$A_n := \left(a_1 + \frac{b_1 - a_1}{2n}, b_1 \right) \times \cdots \times \left(a_d + \frac{b_d - a_d}{2n}, b_d \right). \quad (\text{XII.88})$$

Dann ist $A_n \in \mathcal{O}_d$, für alle $n \in \mathbb{N}$, und

$$[a, b) = \bigcup_{n=1}^{\infty} A_n \in \mathfrak{A}(\mathcal{O}_d). \quad (\text{XII.89})$$

Also ist $\mathfrak{Q}_d \subseteq \mathfrak{A}(\mathcal{O}_d)$, woraus abermals mit (XII.83) auch $\mathfrak{A}(\mathfrak{Q}_d) \subseteq \mathfrak{A}(\mathcal{O}_d)$ folgt.

Um die umgekehrte Inklusion $\mathfrak{A}(\mathcal{O}_d) \subseteq \mathfrak{A}(\mathfrak{Q}_d)$ zu beweisen, betrachten wir eine offene Menge $A \in \mathcal{O}_d$. Da jeder Punkt $x \in A$ ein innerer Punkt ist, gibt es ein $r(x) > 0$, so dass $B_{r(x)}(x) \subseteq A$. Wir wählen nun $\rho(x) \in \mathbb{Q}^+$ so, dass $(2d)^{-1}r(x) < \rho(x) < d^{-1}r(x)$ und dann einen Punkt $y(x) \in \mathbb{Q}^d$, mit $\|y(x) - x\| < (4d)^{-1}r(x)$. Definieren wir nun noch $\varepsilon(x) \in \mathbb{Q}^d$ durch $\varepsilon_\nu(x) := \rho(x)$, für alle $\nu = 1, \dots, d$, so beobachten wir dass $|x_\nu - y_\nu(x)| < \varepsilon_\nu(x)$ und somit dass

$$x \in [y(x) - \varepsilon(x), y(x) + \varepsilon(x)) \subseteq B_{r(x)}(x) \subseteq A \quad (\text{XII.90})$$

Inbesondere ist

$$A = \bigcup_{x \in A} \{x\} \subseteq \bigcup_{x \in A} [y(x) - \varepsilon(x), y(x) + \varepsilon(x)) \subseteq A, \quad (\text{XII.91})$$

also

$$A = \bigcup_{x \in A} [y(x) - \varepsilon(x), y(x) + \varepsilon(x)). \quad (\text{XII.92})$$

Die Vereinigung auf der rechten Seite in (XII.92) ist abzählbar, denn die Menge der Quader $[y(x) - \varepsilon(x), y(x) + \varepsilon(x)) \in \mathfrak{Q}_d$ ist mit $y(x), \varepsilon(x) \in \mathbb{Q}^d$ abzählbar. Es folgt also, dass $\mathcal{O}_d \subseteq \mathfrak{A}(\mathfrak{Q}_d)$ und somit auch $\mathfrak{A}(\mathcal{O}_d) \subseteq \mathfrak{A}(\mathfrak{Q}_d) = \mathfrak{B}_d$. $\square$

XII.3.4. Das äußere Maß

Wir führen nun den auf Caratheodory zurückgehenden Begriff des äußeren Maßes ein.

Definition XII.32. Seien Ω eine Menge, $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ ein Ring und $\mu : \mathfrak{R} \rightarrow \mathbb{R}_0^+$ ein Prämaß auf $\mathfrak{R}$. Für $Q \in \mathfrak{P}(\Omega)$ sei

$$\mathfrak{U}(Q) := \left\{ (A_n)_{n=1}^{\infty} \subseteq \mathfrak{R} \mid Q \subseteq \bigcup_{n=1}^{\infty} A_n \right\}, \quad (\text{XII.93})$$

wobei $(A_n)_{n=1}^{\infty} \subseteq \mathfrak{R}$ eine Folge in $\mathfrak{R}$ bezeichne. Das **äußere Maß** (bezüglich μ) ist die Abbildung $\mu^* : \mathfrak{P}(\Omega) \rightarrow [0, \infty]$, die durch $\mu^*(Q) := \infty$, falls $\mathfrak{U}(Q) = \emptyset$, und

$$\mu^*(Q) := \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) \mid (A_n)_{n=1}^{\infty} \in \mathfrak{U}(Q) \right\}, \quad (\text{XII.94})$$

falls $\mathfrak{U}(Q) \neq \emptyset$, definiert ist.

Lemma XII.33. Seien Ω eine Menge, $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ ein Ring und $\mu : \mathfrak{R} \rightarrow [0, \infty]$ ein Prämaß auf $\mathfrak{R}$. Das zugehörige äußere Maß besitzt folgende Eigenschaften:

$$\mu^* \geq 0, \quad \mu^*(\emptyset) = 0, \quad (\text{XII.95})$$

$$Q_1 \subseteq Q_2 \Rightarrow \mu^*(Q_1) \leq \mu^*(Q_2), \quad (\text{XII.96})$$

$$\mu^*\left(\bigcup_{n=1}^{\infty} Q_n\right) \leq \sum_{n=1}^{\infty} \mu^*(Q_n). \quad (\text{XII.97})$$

Ist $A \in \mathfrak{R}$, so gelten weiterhin noch

$$\forall Q \in \mathfrak{P}(\Omega) : \mu^*(Q) = \mu^*(Q \cap A) + \mu^*(Q \setminus A), \quad (\text{XII.98})$$

$$\mu^*(A) = \mu(A). \quad (\text{XII.99})$$

Beweis.

Glg. (XII.95): Die Nichtnegativität von μ^* ist offensichtlich. Mit $A_n := \emptyset$ ist $(A_n)_{n=1}^{\infty} \in \mathfrak{U}(\emptyset)$ und deshalb $0 \leq \mu^*(\emptyset) \leq 0$.

Glg. (XII.96): Mit $Q_1 \subseteq Q_2$ ist $\mathfrak{U}(Q_1) \supseteq \mathfrak{U}(Q_2)$ und daher $\mu^*(Q_1) \leq \mu^*(Q_2)$.

Glg. (XII.97): Seien $(Q_n)_{n=1}^{\infty}$ eine Folge in $\mathfrak{P}(\Omega)$ und $Q := \bigcup_{n=1}^{\infty} Q_n$. Sei weiterhin $\varepsilon > 0$. Wir können für alle $n \in \mathbb{N}$ annehmen, dass $\mathfrak{U}(Q_n) \neq \emptyset$, denn anderenfalls ist die rechte Seite in (XII.97) unendlich, und (XII.97) ist somit gültig. Nach Definition von μ^* gibt dann es zu jedem $Q_n \in \mathfrak{P}(\Omega)$ eine Folge $(A_{n,m})_{m=1}^{\infty} \in \mathfrak{U}(Q_n)$, so dass

$$\forall n \in \mathbb{N} : \quad \mu^*(Q_n) \geq \sum_{m=1}^{\infty} \mu(A_{n,m}) - \varepsilon 2^{-n}. \quad (\text{XII.100})$$

Da $Q_n \subseteq \bigcup_{m=1}^{\infty} A_{n,m}$, ist auch $Q \subseteq \bigcup_{m,n=1}^{\infty} A_{n,m}$, d.h., die (abzählbare) Doppelfolge überdeckt Q also $(A_{n,m})_{m,n=1}^{\infty} \in \mathfrak{U}(Q)$. Somit ist

$$\mu^*(Q) \leq \sum_{m,n=1}^{\infty} \mu(A_{n,m}) \leq \sum_{n=1}^{\infty} \{\mu^*(Q_n) + \varepsilon 2^{-n}\} = \varepsilon + \sum_{n=1}^{\infty} \mu^*(Q_n). \quad (\text{XII.101})$$

Da $\varepsilon > 0$ beliebig klein gewählt werden kann, impliziert dies (XII.97).

Glg. (XII.98): Seien $Q \in \mathfrak{P}(\Omega)$ und $A \in \mathfrak{R}$. Setzen wir $Q_1 := Q \cap A$, $Q_2 := Q \setminus A$ und $Q_3 := Q_4 := \dots := \emptyset$, so folgt aus (XII.95) und (XII.97), dass

$$\mu^*(Q) \leq \mu^*(Q \cap A) + \mu^*(Q \setminus A). \quad (\text{XII.102})$$

Es verbleibt also die umgekehrte Ungleichung zu zeigen. Dazu können wir wieder $\mathfrak{U}(Q) \neq \emptyset$ annehmen, denn anderenfalls gilt $\mu^*(Q \cap A) + \mu^*(Q \setminus A) \leq \infty = \mu^*(Q)$ trivialerweise. Ist also $(A_n)_{n=1}^{\infty} \in \mathfrak{U}(Q)$, dann beobachten wir zunächst, dass $\mu(A_n) = \mu(A_n \cap A) + \mu(A_n \setminus A)$, weil μ ein Inhalt ist. Addieren wir diese Gleichungen, so erhalten wir

$$\sum_{n=1}^{\infty} \mu(A_n \cap A) + \sum_{n=1}^{\infty} \mu(A_n \setminus A) = \sum_{n=1}^{\infty} \mu(A_n). \quad (\text{XII.103})$$

Außerdem sind $(A_n \cap A)_{n=1}^\infty \in \mathfrak{U}(Q \cap A)$ und $(A_n \setminus A)_{n=1}^\infty \in \mathfrak{U}(Q \setminus A)$, daher sind

$$\mu^*(Q \cap A) \leq \sum_{n=1}^\infty \mu(A_n \cap A) \quad \text{und} \quad \mu^*(Q \setminus A) \leq \sum_{n=1}^\infty \mu(A_n \setminus A), \quad (\text{XII.104})$$

woraus mit (XII.103)

$$\mu^*(Q \cap A) + \mu^*(Q \setminus A) \leq \sum_{n=1}^\infty \mu(A_n) \quad (\text{XII.105})$$

für alle $(A_n)_{n=1}^\infty \in \mathfrak{U}(Q)$ folgt. Also gilt auch

$$\mu^*(Q \cap A) + \mu^*(Q \setminus A) \leq \inf_{(A_n)_{n=1}^\infty \in \mathfrak{U}(Q)} \left\{ \sum_{n=1}^\infty \mu(A_n) \right\} = \mu^*(Q). \quad (\text{XII.106})$$

Glg. (XII.99): Mit $A_1 := A$ und $A_2 := A_3 := \dots := \emptyset$ ist $(A_n)_{n=1}^\infty \in \mathfrak{U}(A)$, und deshalb ist

$$\mu^*(A) \leq \mu(A). \quad (\text{XII.107})$$

Um die umgekehrte Ungleichung zu beweisen, verwenden wir, dass μ auf $\mathfrak{R}$ ein Prämaß und somit $\emptyset$ -stetig ist. Wir wissen bereits, dass $\mathfrak{U}(A) \ni (A, \emptyset, \emptyset, \dots)$ nicht leer ist. Seien also $(A_n)_{n=1}^\infty \in \mathfrak{U}(A)$ beliebig und $\varepsilon > 0$. Wir setzen $B_N := A \setminus (\bigcup_{n=1}^N A_n) \in \mathfrak{R}$ und beobachten, dass

$$\mu(A) = \mu\left(A \cap \bigcup_{n=1}^N A_n\right) + \mu(B_N) \leq \sum_{n=1}^N \mu(A_n) + \mu(B_N). \quad (\text{XII.108})$$

Mit $N \rightarrow \infty$ ergibt dies insbesondere

$$\mu(A) \leq \sum_{n=1}^\infty \mu(A_n) + \limsup_{N \rightarrow \infty} \mu(B_N). \quad (\text{XII.109})$$

Nun bemerken wir, dass $B_N \searrow \emptyset$, da $A \subseteq \bigcup_{n=1}^\infty A_n$. Weil μ $\emptyset$ -stetig ist, folgt damit, dass $\limsup_{N \rightarrow \infty} \mu(B_N) = \lim_{N \rightarrow \infty} \mu(B_N) = 0$, und wir haben

$$\mu(A) \leq \sum_{n=1}^\infty \mu(A_n), \quad (\text{XII.110})$$

für alle $(A_n)_{n=1}^\infty \in \mathfrak{U}(A)$, was $\mu(A) \leq \mu^*(A)$ nach sich zieht. $\square$

Satz XII.34 (Caratheodory). *Seien Ω eine Menge, $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ ein Ring, $\mu : \mathfrak{R} \rightarrow [0, \infty]$ ein Prämaß auf $\mathfrak{R}$ und $\mu^* : \mathfrak{P}(\Omega) \rightarrow [0, \infty]$ das zugehörige äußere Maß. Dann ist*

$$\mathfrak{A}^* := \left\{ A \in \mathfrak{P}(\Omega) \mid \forall Q \in \mathfrak{P}(\Omega) : \mu^*(Q) = \mu^*(Q \cap A) + \mu^*(Q \setminus A) \right\} \quad (\text{XII.111})$$

eine Sigma-Algebra, die $\mathfrak{R}$ umfasst, also

$$\mathfrak{R} \subseteq \mathfrak{A}(\mathfrak{R}) \subseteq \mathfrak{A}^* \subseteq \mathfrak{P}(\Omega). \quad (\text{XII.112})$$

Beweis. Es ist nur zu zeigen, dass $\mathfrak{A}^*$ eine Sigma-Algebra ist, denn nach (XII.98) ist $\mathfrak{A} \subseteq \mathfrak{A}^*$. Weiterhin ist trivialerweise $\Omega \in \mathfrak{A}^*$, und auch die Implikation $A \in \mathfrak{A}^* \Rightarrow A^c \in \mathfrak{A}^*$ gilt trivial, wenn wir $Q \setminus A = Q \cap A^c$ beachten. Es ist also nur nachzuweisen, dass $\mathfrak{A}^*$ unter abzählbaren Vereinigungen abgeschlossen ist. Dazu betrachten wir zunächst $A, B \in \mathfrak{A}^*$. Dann ist

$$\begin{aligned} \mu^*(Q) &= \mu^*(Q \cap A) + \mu^*(Q \cap A^c) \\ &= \mu^*(Q \cap A \cap B) + \mu^*(Q \cap A^c \cap B) \\ &\quad + \mu^*(Q \cap A \cap B^c) + \mu^*(Q \cap A^c \cap B^c) \\ &= \mu^*(Q \cap A \cap B) + \mu^*(Q \cap A^c \cap B) \\ &\quad + \mu^*(Q \cap A \cap B^c) + \mu^*(Q \cap [A \cup B]^c). \end{aligned} \quad (\text{XII.113})$$

Ersetzen wir in diese Identität Q durch $Q \cap [A \cup B]$, so fällt der letzte Term heraus, und die anderen drei bleiben unverändert,

$$\begin{aligned} \mu^*(Q \cap [A \cup B]) &= \mu^*(Q \cap [A \cup B] \cap A \cap B) + \mu^*(Q \cap [A \cup B] \cap A^c \cap B) \\ &\quad + \mu^*(Q \cap [A \cup B] \cap A \cap B^c) \\ &= \mu^*(Q \cap A \cap B) + \mu^*(Q \cap A^c \cap B) + \mu^*(Q \cap A \cap B^c). \end{aligned} \quad (\text{XII.114})$$

Ersetzen wir die ersten 3 Terme auf der rechten Seite von (XII.113) durch die linke Seite von (XII.114), so erhalten wir

$$\mu^*(Q) = \mu^*(Q \cap [A \cup B]) + \mu^*(Q \cap [A \cup B]^c), \quad (\text{XII.115})$$

also $A \cup B \in \mathfrak{A}^*$. Sind nun $A_1, A_2 \in \mathfrak{A}^*$ disjunkt, so sind $A_1 \cap A_2 = \emptyset$, $A_1 \cap A_2^c = A_1$ und $A_1^c \cap A_2 = A_2$, und aus (XII.114) folgt dann außerdem, dass

$$\mu^*(Q \cap [A_1 \cup A_2]) = \mu^*(Q \cap A_1) + \mu^*(Q \cap A_2). \quad (\text{XII.116})$$

Durch Induktion übertragen sich (XII.115) und (XII.116) sofort auf endlich viele Mengen:

Ist $(A_n)_{n=1}^\infty \subseteq \mathfrak{A}^*$ eine Folge paarweise disjunkter Mengen und

$$B_N := \bigcup_{n=1}^N A_n \quad \text{sowie} \quad A := \bigcup_{n=1}^\infty A_n, \quad (\text{XII.117})$$

so folgt, dass $B_N \in \mathfrak{A}^*$ und

$$\mu^*(Q \cap B_N) = \sum_{n=1}^N \mu^*(Q \cap A_n). \quad (\text{XII.118})$$

Wegen $B_N \subseteq A$ ist $Q \setminus B_N \supseteq Q \setminus A$ und daher $\mu^*(Q \setminus B_N) \geq \mu^*(Q \setminus A)$, was mit (XII.118) und $B_N \in \mathfrak{A}^*$ zusammen

$$\mu^*(Q) = \mu^*(Q \cap B_N) + \mu^*(Q \setminus B_N) \geq \sum_{n=1}^N \mu^*(Q \cap A_n) + \mu^*(Q \setminus A) \quad (\text{XII.119})$$

ergibt. Im Limes $N \rightarrow \infty$ erhalten wir daraus

$$\begin{aligned} \mu^*(Q) &\geq \sum_{n=1}^{\infty} \mu^*(Q \cap A_n) + \mu^*(Q \setminus A) \\ &\geq \mu^*(Q \cap A) + \mu^*(Q \setminus A) \geq \mu^*(Q), \end{aligned} \quad (\text{XII.120})$$

wobei wir zusätzlich (XII.97) verwenden. Also muss tatsächlich Gleichheit überall in (XII.120) gelten, und es folgt, dass $A \in \mathfrak{A}^*$. $\square$

XII.3.5. Maße

Definition XII.35. Sei Ω eine Menge und $\mathfrak{A} \subseteq \mathfrak{P}(\Omega)$ eine Sigma-Algebra. Eine Abbildung $\mu : \mathfrak{A} \rightarrow [0, \infty]$ heißt **Maß** (auf $\mathfrak{A}$): $\Leftrightarrow$

- (i) $\mu(\emptyset) = 0$.
- (ii) Die Abbildung μ ist σ -additiv, d. h., ist $(A_n)_{n=1}^{\infty}$ eine Folge paarweise disjunkter Mengen in $\mathfrak{A}$, so gilt

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n). \quad (\text{XII.121})$$

Gibt es außerdem eine Folge $(A_n)_{n=1}^{\infty}$ in $\mathfrak{A}$ mit $\mu(A_n) < \infty$, für alle $n \in \mathbb{N}$, und $A_n \nearrow \Omega$, d.h. $A_n \subseteq A_{n+1}$ und $\bigcup_{n=1}^{\infty} A_n = \Omega$, so heißt das Maß μ **Sigma-endlich**.

Ein Maß ist also ein Prämaß auf einer Sigma-Algebra. Wir kommen nun zum Hauptresultat dieses Kapitels.

Satz XII.36. Seien Ω eine Menge, $\mathfrak{R} \subseteq \mathfrak{P}(\Omega)$ ein Ring und $\mu : \mathfrak{R} \rightarrow \mathbb{R}_0^+$ ein Prämaß auf $\mathfrak{R}$. Dann definiert die Restriktion des äußeren Maßes $\mu^* : \mathfrak{P}(\Omega) \rightarrow [0, \infty]$ eine Fortsetzung von μ zu einem Maß $\tilde{\mu} := \mu^*|_{\mathfrak{A}(\mathfrak{R})} : \mathfrak{A}(\mathfrak{R}) \rightarrow [0, \infty]$, so dass also $\tilde{\mu}|_{\mathfrak{R}} = \mu$. Ist das Prämaß μ Sigma-endlich, so ist $\tilde{\mu}$ die einzige Fortsetzung von μ zu einem Maß auf $\mathfrak{A}(\mathfrak{R})$.

Beweis. Wie bereits in (XII.95) bemerkt, gilt $\mu^*(\emptyset) = 0$. Um zu zeigen, dass die Restriktion von μ^* auf $\mathfrak{A}(\mathfrak{R})$ ein Maß definiert, brauchen wir also nur die Sigma-Additivität von μ^* auf $\mathfrak{A}(\mathfrak{R})$ nachzuweisen. Dazu beobachten wir, dass $\mathfrak{A}(\mathfrak{R}) \subseteq \mathfrak{A}^*$, nach Satz XII.34. Die Sigma-Additivität von μ^* erhalten wir aus (XII.120) sogar auf $\mathfrak{A}^*$, indem wir $Q := A$ setzen,

$$\mu^*(A) = \sum_{n=1}^{\infty} \mu^*(A \cap A_n) + \mu^*(A \setminus A) = \sum_{n=1}^{\infty} \mu^*(A_n). \quad (\text{XII.122})$$

Zum Beweis der Eindeutigkeit der Fortsetzung von μ würden wir noch die Begriffe der *Dynkin-Systeme* bzw. der *monotonen Klassen* benötigen. Aus Zeitmangel verzichten wir auf diese Konstruktion und verweisen den interessierten Leser auf die Bücher von Bauer¹ und von Lieb und Loss². $\square$

¹ *Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie*, 3. Auflage, de Gruyter 1978.

² *Analysis*, AMS Publications, 1997

Korollar XII.37. *Das auf den Figuren $\mathfrak{F}_d$ definierte Prämaß μ_d hat eine eindeutige Fortsetzung zu einem Sigma-endlichen Maß auf $\mathfrak{B}_d$, das **Lebesgue-Borel-Maß**, das wir abermals mit μ_d bezeichnen.*

Wir bemerken, dass jedes Maß $\hat{\mu}$ auf $\mathbb{R}^d$ auch den Wert $\hat{\mu}(\mathbb{R}^d)$ festlegen muss, da $\mathbb{R}^d$ nach (XII.77) Element einer jeden Sigma-Algebra über $\mathbb{R}^d$ ist. Für das Lebesgue-Borel-Maß muss $\mu_d(\mathbb{R}^d) = \infty$ sein, da $\mu_d(\mathbb{R}^d) \geq \mu_d([-L, L]^d) = 2^d L^d$, für jedes $L > 0$.

Damit ist klar, dass die Zulässigkeit des Wertes “ ∞ ” für Maße $\mu : \mathfrak{A} \rightarrow [0, \infty] = \overline{\mathbb{R}_0^+} = \mathbb{R}_0^+ \cup \{\infty\}$ eine unabdingbare Forderung ist.

Definition XII.38. Sind Ω eine Menge und $\mathfrak{A} \subseteq \mathfrak{P}(\Omega)$ eine Sigma-Algebra, so bezeichnet man das Paar $(\Omega, \mathfrak{A})$ als **Messraum** oder **messbaren Raum**, und die Mengen in $\mathfrak{A}$ heißen **messbar**. Ist weiterhin $\mu : \mathfrak{A} \rightarrow [0, \infty]$ ein Maß, so bezeichnet man das Tripel $(\Omega, \mathfrak{A}, \mu)$ als **Maßraum**.

Insbesondere ist nach Korollar XII.37 das Tripel $(\mathbb{R}^d, \mathfrak{B}_d, \mu_d)$ ein Maßraum.

Wir bemerken noch, dass sich das sogenannte *Lebesgue-Maß* vom Lebesgue-Borel-Maß in einem wichtigen Punkt unterscheidet: Ist $A \in \mathfrak{B}_d$ eine messbare Menge mit Maß Null, $\mu_d(A) = 0$, so gilt wegen der Monotonie des äußeren Maßes auch $\mu_d^*(A') = 0$, für jede Teilmenge $A' \subseteq A$. Trotzdem muss A' nicht notwendig messbar sein, und obwohl 0 ein natürlicher Kandidat für $\mu_d(A')$ ist, wäre dann $\mu_d(A')$ nicht definiert. Diesen Nachteil kann man noch beseitigen, indem man zu den Borelmengen noch die Mengen in $\mathfrak{A}^*$ mit äußerem Maß Null hinzufügt. Genauer gesagt definieren wir die **Nullmengen** durch

$$\mathfrak{N}_d := \left\{ A \in \mathfrak{A}^*(\mathfrak{F}_d) \mid \mu_d^*(A) = 0 \right\} \quad (\text{XII.123})$$

und anschließend die Sigma-Algebra der **Lebesgue-messbaren Mengen** durch

$$\mathfrak{L}_d := \mathfrak{A}(\mathfrak{B}_d \cup \mathfrak{N}_d). \quad (\text{XII.124})$$

Da $\mathfrak{L}_d \subseteq \mathfrak{A}^*(\mathfrak{F}_d)$, kann man sich leicht überzeugen, dass das Lebesgue-Borel-Maß auf $\mathfrak{B}_d$ in eindeutiger Weise zum **Lebesgue-Maß** auf $\mathfrak{L}_d$ fortgesetzt werden kann. Wir ignorieren im Folgenden diesen Unterschied und sprechen vom *Lebesgue-Maß*, obwohl wir stets das *Lebesgue-Borel-Maß* meinen.

Schließlich stellt sich die Frage, ob $\mathfrak{A}^* \subseteq \mathfrak{P}(\Omega)$ eine echte Teilmenge darstellt. Die Antwort auf diese Frage hängt von Ω , dem Ring $\mathfrak{R}$ und dem darauf definierten Inhalt μ ab. Ist etwa Ω abzählbar, und enthält $\mathfrak{R}$ die Elemente von Ω als einelementige Mengen, so ist sofort $\mathfrak{A}(\mathfrak{R}) = \mathfrak{P}(\Omega)$ und somit auch $\mathfrak{A}^* = \mathfrak{P}(\Omega)$. Für den Ring $\mathfrak{F}_d$ der Figuren mit Prämaß μ_d gilt allerdings

$$\mathfrak{A}^*(\mathfrak{F}_d) \neq \mathfrak{P}(\mathbb{R}^d). \quad (\text{XII.125})$$

Die Konstruktion von Mengen in $\mathfrak{P}(\mathbb{R}^d) \setminus \mathfrak{A}^*(\mathfrak{F}_d)$ ist allerdings anspruchsvoll, unintuitiv und führt auf alle möglichen (scheinbar) paradoxe Sachverhalte. Eine interessante Diskussion dieser Mengen findet man bei Oxtoby³.

³Maß und Kategorie Springer-Verlag, 1982.

Satz XII.39. *Das Lebesgue-Maß $\mu_d : \mathcal{B}_d \rightarrow [0, \infty]$ besitzt die Eigenschaften der **äußeren Regularität** und der **inneren Regularität**, d.h. für jedes $A \in \mathcal{B}_d$ gilt*

$$\mu_d(A) = \inf \{ \mu_d(U) \mid U \supseteq A, U \text{ ist offen} \} \quad (\text{XII.126})$$

$$= \sup \{ \mu_d(C) \mid C \subseteq A, C \text{ ist kompakt} \}. \quad (\text{XII.127})$$

Beweis. Wir zeigen nur (XII.126)), bemerken aber zuvor, dass $\mathbb{R}^d \supseteq A$ offen und $\emptyset \subseteq A$ abgeschlossen sind. Für $\mu_d(A) = \infty$ gilt (XII.126) trivialerweise, sei also $\mu_d(A) < \infty$. Da $\mu_d = \mu^*|_{\mathcal{L}_d}$, wobei μ^* das zum Ring $\mathfrak{F}_d$ der Figuren gehörige äußere Maß notiert, gibt es gemäß (XII.111) zu jedem $\varepsilon > 0$ eine Folge $(A_n)_{n=1}^\infty$ in $\mathfrak{F}_d$, so dass

$$A \subseteq \bigcup_{n=1}^\infty A_n \quad \text{und} \quad \mu_d(A) \geq \left(\sum_{n=1}^\infty \mu_d(A_n) \right) - \varepsilon. \quad (\text{XII.128})$$

Jedes A_n ist aber die Vereinigung endlich vieler halboffener und paarweise disjunkter Quader. Daher gibt es Folgen $(a_k)_{k=1}^\infty, (b_k)_{k=1}^\infty$ in $\mathbb{R}^d$, mit $a_k < b_k$, so dass

$$A \subseteq \bigcup_{k=1}^\infty [a_k, b_k) \quad \text{und} \quad \mu_d(A) \geq \left(\sum_{k=1}^\infty \mu_d([a_k, b_k)) \right) - \varepsilon. \quad (\text{XII.129})$$

Wählen wir nun $c_k < a_k$ so, dass

$$\mu_d((c_k, b_k)) \leq \mu_d([a_k, b_k]) + \varepsilon \cdot 2^{-k}, \quad (\text{XII.130})$$

so ist

$$U := \bigcup_{k=1}^\infty (c_k, b_k) \supseteq \bigcup_{k=1}^\infty [a_k, b_k) \supseteq A \quad (\text{XII.131})$$

offen und

$$\begin{aligned} \mu_d(A) &\geq \left(\sum_{k=1}^\infty \left\{ \mu_d((c_k, b_k)) - \varepsilon \cdot 2^{-k} \right\} \right) - \varepsilon \\ &= \left(\sum_{k=1}^\infty \mu_d((c_k, b_k)) \right) - 2\varepsilon \geq \mu_d(U) - 2\varepsilon. \end{aligned} \quad (\text{XII.132})$$

Da $\varepsilon > 0$ beliebig klein gewählt werden kann, folgt daraus (XII.126). $\square$

XIII. Integration über Maße

XIII.1. Messbare Funktionen und Definition des Integrals

Wir beginnen dieses Kapitel mit dem Begriff der Messbarkeit, genauer der reellen Messbarkeit. Dazu fügen wir zu den reellen Zahlen noch die Werte “ $-\infty$ ” und “ ∞ ” hinzu, denn wir haben schon im vorigen Kapitel gesehen, dass Maße den Wert “ ∞ ” annehmen können, etwa $\mu_d(\mathbb{R}^d) = \infty$. Wir setzen also $\overline{\mathbb{R}} := [-\infty, \infty] := \mathbb{R} \cup \{-\infty, \infty\}$ und $\overline{\mathfrak{B}}_1 := \mathfrak{A}(\overline{\mathbb{R}})$.

Definition XIII.1. Sei $(\Omega, \mathfrak{A})$ ein Messraum. Eine Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ heißt **reell-messbar** : $\Leftrightarrow$

$$\forall A \in \overline{\mathfrak{B}}_1 : \quad f^{-1}(A) \in \mathfrak{A}. \quad (\text{XIII.1})$$

Die reell-Messbarkeit einer Funktion lässt sich auf ein einfacheres Kriterium zurückführen: es reicht, in (XIII.1) für A halboffene Quader einzusetzen. Es gilt sogar folgende Aussage.

Lemma XIII.2. Sei $(\Omega, \mathfrak{A})$ ein Messraum. Eine Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ ist genau dann reell-messbar, wenn ihre **Niveaumengen** $A(f, \lambda) := f^{-1}((\lambda, \infty]) = \{\omega \in \Omega \mid f(\omega) > \lambda\} \in \mathfrak{A}$ für alle $\lambda \in \mathbb{R}$ in $\mathfrak{A}$ liegen.

Das folgende Lemma besagt, dass die Menge der reell-messbaren Funktionen unter abzählbaren Supremums- und Infimumsbildungen, sowie Bildungen des Limes Superior und des Limes Inferior abgeschlossen ist.

Lemma XIII.3. Seien $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $(f_k)_{k=1}^\infty$ eine Folge reell-messbarer Funktionen. Dann sind $\sup_k f_k$, $\inf_k f_k$, $\limsup_k f_k$, $\liminf_k f_k$, und, falls für alle $\omega \in \Omega$ existent, auch $\lim_k f_k$ reell-messbar.

Die Klasse der reell-messbaren Funktionen ist genügend groß um alle erdenklichen Beispiele miteinzuschließen - auch die meisten “pathologischen” Beispiele, wie etwa die Dirichlet-Funktion. In der Tat ist die Konstruktion konkreter nicht-messbarer Funktionen genauso schwierig wie etwa die Konstruktion von Mengen in $\mathfrak{P}(\mathbb{R}^d) \setminus \mathfrak{A}^*(\mathfrak{F}_d)$.

Wir definieren nun das Integral für nichtnegative Funktionen eines Maßraums.

Definition XIII.4. Seien $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $f : \Omega \rightarrow [0, \infty]$ eine nichtnegative, reell-messbare Funktion. Für $\lambda > 0$ definieren wir die **Niveaumengen von f** durch

$$A(f, \lambda) := f^{-1}((\lambda, \infty]) = \{\omega \in \Omega \mid f(\omega) > \lambda\}. \quad (\text{XIII.2})$$

Nach Definition XIII.1 ist $A(f, \lambda) \in \mathfrak{A}$, und

$$F(f, \lambda) := \mu(A(f, \lambda)) = \mu[f^{-1}((\lambda, \infty])] \quad (\text{XIII.3})$$

definiert eine Abbildung $F(f, \cdot) : \mathbb{R}^+ \rightarrow [0, \infty]$.

Offensichtlich ist $F(f, \lambda)$ stets eine in λ monoton fallende Funktion, denn für $\lambda < \lambda'$ ist $(\lambda, \infty] \supseteq (\lambda', \infty]$, daher auch

$$A(f, \lambda) = f^{-1}((\lambda, \infty]) \supseteq f^{-1}((\lambda', \infty]) = A(f, \lambda'), \quad (\text{XIII.4})$$

und somit

$$F(f, \lambda) = \mu(A(f, \lambda)) \geq \mu(A(f, \lambda')) = F(f, \lambda'). \quad (\text{XIII.5})$$

Das nun folgende Lemma besagt, dass monoton fallende Funktionen $\mathbb{R}^+ \rightarrow \mathbb{R}_0^+$ stets Riemann-integrierbar sind.

Lemma XIII.5. Ist $F : \mathbb{R}^+ \rightarrow \mathbb{R}_0^+$ monoton fallend, so ist F auf $[a, b]$ Riemann-integrierbar, für alle $0 < a < b < \infty$. Ist darüber hinaus

$$I_F := \sup_{0 < a < b < \infty} \int_a^b F(x) dx < \infty, \quad (\text{XIII.6})$$

so ist F auch auf $\mathbb{R}^+$ Riemann-integrierbar, und $\int_0^\infty F(x) dx = I_F$.

Nach Lemma XIII.5 ist dann $F(f, \cdot)$ auf $[a, b]$ Riemann-integrierbar, falls $0 < a < b < \infty$ so gewählt sind, dass $F(f, a) < \infty$.

Definition XIII.6. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum.

- (i) Eine reell-messbare Funktion $f : \Omega \rightarrow [0, \infty]$ heißt **(μ) -integrierbar**, falls $F(f, \cdot)$ auf $\mathbb{R}^+$ Riemann-integrierbar ist. In diesem Fall heißt

$$\int_\Omega f(\omega) d\mu(\omega) := \int_0^\infty F(f, \lambda) d\lambda \quad (\text{XIII.7})$$

Integral von f .

- (ii) Eine reell-messbare Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ heißt **(μ) -integrierbar**, falls $f_+ := \max\{f, 0\}$ und $f_- := \max\{-f, 0\}$ jeweils μ -integrierbar sind. Das **Integral** von f ist dann definiert durch

$$\int_\Omega f(\omega) d\mu(\omega) := \int_\Omega f_+(\omega) d\mu(\omega) - \int_\Omega f_-(\omega) d\mu(\omega). \quad (\text{XIII.8})$$

- (iii) Eine komplexe Funktion $f : \Omega \rightarrow \mathbb{C}$ heißt **messbar**, falls $\operatorname{Re}\{f\}$ und $\operatorname{Im}\{f\}$ jeweils reell-messbar sind. Sind darüber hinaus $\operatorname{Re}\{f\}$ und $\operatorname{Im}\{f\}$ jeweils μ -integrabel, so ist f (μ -)**integrabel**, und ihr **Integral** ist definiert durch

$$\int_{\Omega} f(\omega) d\mu(\omega) := \int_{\Omega} \operatorname{Re}\{f(\omega)\} d\mu(\omega) + i \int_{\Omega} \operatorname{Im}\{f(\omega)\} d\mu(\omega). \quad (\text{XIII.9})$$

Bemerkungen und Beispiele.

- Sind $M \in \mathfrak{A}$ eine messbare Menge eines Maßraums $(\Omega, \mathfrak{A}, \mu)$ und

$$\mathbb{1}_M(\omega) \equiv \mathbb{1}[\omega \in M] := \begin{cases} 1 & \text{falls } \omega \in M, \\ 0 & \text{falls } \omega \notin M, \end{cases} \quad (\text{XIII.10})$$

so ist, für $\eta > 0$,

$$A(\mathbb{1}_M, \eta) = \{\omega \in \Omega \mid \mathbb{1}_M(\omega) > \eta\} = \begin{cases} \emptyset, & \text{falls } \eta \geq 1, \\ M, & \text{falls } 0 < \eta < 1. \end{cases} \quad (\text{XIII.11})$$

Somit erhalten wir

$$F(\mathbb{1}_M, \eta) = \mu(M) \cdot \mathbb{1}[0 < \eta \leq 1] \quad (\text{XIII.12})$$

und

$$\int_{\Omega} \mathbb{1}_M d\mu = \mu(M). \quad (\text{XIII.13})$$

- Insbesondere folgt aus (XIII.13), mit $M := \{\omega \in \Omega \mid f(\omega) > \lambda\}$, wobei $\lambda > 0$ und f messbar sind, dass

$$\int_{\Omega} \mathbb{1}[f(\omega) > \lambda] d\mu(\omega) = \mu(\{f(\omega) > \lambda\}) = F(f, \lambda). \quad (\text{XIII.14})$$

- Zur Rechtfertigung von (XIII.7) dient folgende formale “Herleitung”, die eine nützliche *Merkhilfe* (aber auch nicht mehr als das) darstellt:

$$\begin{aligned} \int_0^{\infty} F(f, \lambda) d\lambda &= \int_0^{\infty} \left(\int_{\Omega} \mathbb{1}[f(\omega) > \lambda] d\mu(\omega) \right) d\lambda \\ &= \int_{\Omega} \left(\int_0^{\infty} \mathbb{1}[f(\omega) > \lambda] d\lambda \right) d\mu(\omega) \\ &= \int_{\Omega} \left(\int_0^{f(\omega)} d\lambda \right) d\mu(\omega) = \int_{\Omega} f(\omega) d\mu(\omega). \end{aligned} \quad (\text{XIII.15})$$

(Problematisch ist in (XIII.15) die Vertauschung der Integrationen, d.h. die zweite Gleichung.)

XIII.2. Eigenschaften des Integrals

Um die Eigenschaften des Integrals zu formulieren, definieren wir den Begriff *fast überall*.

Definition XIII.7. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum. Eine Eigenschaft $\mathcal{E} : \Omega \rightarrow \{w, f\}$ (w = wahr, f = falsch) gilt **μ -fast überall** $:\Leftrightarrow$

$$\mathcal{E}^{-1}(f) \in \mathfrak{A} \quad \text{und} \quad \mu(\mathcal{E}^{-1}(f)) = 0, \quad (\text{XIII.16})$$

d.h., die Menge, auf der $\mathcal{E}$ nicht gilt, hat Maß Null.

Satz XIII.8. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum.

- (i) Ist $f : \Omega \rightarrow \mathbb{C}$ eine messbare Funktion, dann ist auch $|f| : \Omega \rightarrow \mathbb{R}_0^+$ reell-messbar. f ist genau dann μ -integrabel, wenn $|f|$ μ -integrabel ist.
- (ii) Sind $f, g : \Omega \rightarrow \mathbb{C}$ μ -integrabel und $\alpha \in \mathbb{C}$, so ist auch $f + \alpha g : \Omega \rightarrow \mathbb{C}$ μ -integrabel, und es gilt

$$\int_{\Omega} (f + \alpha g)(\omega) d\mu(\omega) = \int_{\Omega} f(\omega) d\mu(\omega) + \alpha \int_{\Omega} g(\omega) d\mu(\omega). \quad (\text{XIII.17})$$

Der nun folgende Satz zeigt, dass das Lebesgue-Integral auf $\mathbb{R}$ wirklich eine Verallgemeinerung des Riemann-Integrals ist.

Satz XIII.9. Ist $f \in \mathcal{R}(\mathbb{R}; \mathbb{R}_0^+)$ eine nichtnegative Riemann-integrable Funktion, so gibt es eine reell-messbare und Lebesgue-integrable Funktion $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}_0^+$, die μ_1 -fast überall mit f auf $\mathbb{R}$ übereinstimmt und für die gilt, dass

$$\underbrace{\int_{-\infty}^{\infty} f(x) dx}_{\text{Riemann-Integral}} = \underbrace{\int_{\mathbb{R}} \tilde{f}(x) d\mu_1(x)}_{\text{Lebesgue-Integral}}. \quad (\text{XIII.18})$$

Um zu sehen, dass das Lebesgue-Integral eine echte Fortsetzung des Riemann-Integrals ist, stellen wir als Nächstes die Konvergenzsätze vor.

XIII.3. Die Konvergenzsätze

Der große Vorteil der Definition XIII.6 des Integrals liegt in der Gültigkeit der *Konvergenzsätze* für Integrale, die in der Analysis von überaus großer Bedeutung sind¹. Der erste Konvergenzsatz ist der *Satz über monotone Konvergenz*.

Satz XIII.10 (Monotone Konvergenz). Seien $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $(f_k)_{k=1}^{\infty}$ eine monoton steigende Folge nichtnegativer, μ -integrabler Funktionen, d.h., $f_k : \Omega \rightarrow [0, \infty]$,

¹Lieb und Loss schreiben in *Analysis* (AMS Publications, 1997): "...convergence theorems for integrals, whose importance can hardly be overrated."

$f_k \leq f_{k+1}$, und $\int_{\Omega} f_k(\omega) d\mu(\omega) < \infty$, für alle $k \in \mathbb{N}$. Dann ist $f := \lim_k f_k = \sup_k f_k$ reell-messbar und genau dann μ -integrabel, wenn

$$\lim_{k \rightarrow \infty} \int_{\Omega} f_k(\omega) d\mu(\omega) < \infty, \quad (\text{XIII.19})$$

und in diesem Fall gilt

$$\int_{\Omega} f(\omega) d\mu(\omega) = \lim_{k \rightarrow \infty} \int_{\Omega} f_k(\omega) d\mu(\omega). \quad (\text{XIII.20})$$

Der Beweis des Satzes XIII.10 über monotone Konvergenz ist mit unserer Definition des Integrals erstaunlich einfach und beruht im Wesentlichen auf Lemma XIII.14, das eine elementare Aussage über Riemann-Integrale ist. Lemma XIII.14 und sein Beweis findet man in den Ergänzungen.

Der zweite Konvergenzsatz, *Fatous Lemma*, baut auf dem Satz über monotone Konvergenz auf.

Lemma XIII.11 (Fatous Lemma). Seien $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $(f_k)_{k=1}^{\infty}$ eine Folge nichtnegativer, μ -integrabler Funktionen, d.h., $f_k : \Omega \rightarrow [0, \infty]$ und $\int_{\Omega} f_k(\omega) d\mu(\omega) < \infty$, für alle $k \in \mathbb{N}$. Ist darüber hinaus $\liminf_{k \rightarrow \infty} \int_{\Omega} f_k(\omega) d\mu(\omega) < \infty$, so ist $f := \liminf_k f_k$ reell-messbar, μ -integrabel, und es gilt

$$\liminf_{k \rightarrow \infty} \int_{\Omega} f_k(\omega) d\mu(\omega) \geq \int_{\Omega} f(\omega) d\mu(\omega). \quad (\text{XIII.21})$$

Wir kommen nun zu einem der wichtigsten Sätze der Analysis überhaupt, dem *Satz über majorisierte Konvergenz*².

Satz XIII.12 (Majorisierte Konvergenz). Seien $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $(f_k)_{k=1}^{\infty}$ eine punktweise, μ -fast überall konvergente Folge komplexer, μ -integrabler Funktionen, die durch eine μ -integrable Funktion $g : \Omega \rightarrow [0, \infty]$ majorisiert wird, d.h. $f_k : \Omega \rightarrow \mathbb{C}$, $f(\omega) := \lim_k f_k(\omega) \in \mathbb{C}$, μ -fast überall, $|f_k| \leq g$, für alle $k \in \mathbb{N}$, und $\int_{\Omega} g(\omega) d\mu(\omega) < \infty$. Dann ist f μ -integrabel, und es gilt

$$\int_{\Omega} f(\omega) d\mu(\omega) = \lim_{k \rightarrow \infty} \int_{\Omega} f_k(\omega) d\mu(\omega). \quad (\text{XIII.22})$$

Bemerkungen und Beispiele.

- Der Beweis des Satzes XIII.12 über majorisierte Konvergenz beruht auf dem Satz XIII.10 über monotone Konvergenz und Fatous Lemma, Lemma XIII.11.
- Die Monotonievoraussetzungen im Satz XIII.10 über monotone Konvergenz sind stark und in vielen Anwendungen nicht erfüllt. Dafür ist aber auch die Aussage stark, nämlich eine Äquivalenz und nicht nur eine Implikation: Mit $f_k \nearrow f$ ist die Grenzfunktion f genau dann integrabel, wenn $\sup_k \int f_k d\mu < \infty$.

²engl.: dominated convergence

- Der Satz XIII.12 über majorisierte Konvergenz setzt im Gegensatz dazu nur die punktweise Konvergenz $f = \lim_k f_k$ der Funktionenfolge $(f_k)_{k=1}^\infty$ und die Existenz einer integrierbaren Majorante g voraus - eine Monotonie der Funktionenfolge wie in Satz XIII.10 ist nicht notwendig. Allerdings kann man etwa aus einer erfolglosen Suche nach einer integrierbaren Majoranten nicht schließen, dass die Grenzfunktion nicht integrierbar ist.
- Wir betrachten die Dirichlet-Funktion $f_\infty : \mathbb{R} \rightarrow \mathbb{R}$,

$$f_\infty(x) := \mathbb{1}_{\mathbb{Q}}(x) := \begin{cases} 1 & \text{falls } x \in \mathbb{Q}, \\ 0 & \text{falls } x \in \mathbb{R} \setminus \mathbb{Q}, \end{cases} \quad (\text{XIII.23})$$

von der wir schon wissen, dass sie nicht Riemann-integrierbar ist. Wir nehmen nun an, dass $(q_n)_{n=1}^\infty \in \mathbb{Q}^\mathbb{N}$ eine Abzählung der rationalen Zahlen ist, d.h. $\{q_1, q_2, q_3, \dots\} = \mathbb{Q}$. Wir definieren $f_n : \mathbb{R} \rightarrow \mathbb{R}$ für $n \in \mathbb{N}$ durch

$$f_n(x) := \begin{cases} 1 & \text{falls } x \in \{q_1, q_2, q_3, \dots, q_n\}, \\ 0 & \text{falls } x \in \mathbb{R} \setminus \{q_1, q_2, q_3, \dots, q_n\}, \end{cases} \quad (\text{XIII.24})$$

und beobachten, dass $f_n \nearrow f_\infty$ punktweise monoton gegen f_∞ konvergiert, genauer dass

$$\forall x \in \mathbb{R} : \quad f_n(x) \leq f_{n+1}(x) \rightarrow f_\infty(x), \quad n \rightarrow \infty. \quad (\text{XIII.25})$$

- Für $n \in \mathbb{N}$ ist f_n eine stückweise stetige Funktion mit endlich vielen Sprungstellen bei $q_1, q_2, \dots, q_n$ und daher Riemann-integrierbar, $f_n \in \mathcal{R}(\mathbb{R})$. Nehmen wir o.B.d.A. an, dass alle q_i verschieden sind, so gibt es eine Permutation $\pi \in \mathcal{S}_n$, sodass $q_{\pi(1)} < q_{\pi(2)} < \dots < q_{\pi(n)}$, und wir erhalten

$$\begin{aligned} \int_{\mathbb{R}} f_n(x) d\mu_1(x) &= \int_{-\infty}^{\infty} f_n(x) dx \\ &= \int_{-\infty}^{q_{\pi(1)}} f_n(x) dx + \int_{q_{\pi(1)}}^{q_{\pi(2)}} f_n(x) dx + \dots + \int_{q_{\pi(n-1)}}^{q_{\pi(n)}} f_n(x) dx + \int_{q_{\pi(n)}}^{\infty} f_n(x) dx \\ &= 0 + 0 + \dots + 0 + 0 = 0, \end{aligned} \quad (\text{XIII.26})$$

da f_n nur bei q_i von Null verschieden ist. Nach Satz XIII.10 über monotone Konvergenz ist demnach auch $f_\infty = \lim_{n \rightarrow \infty} f_n$ Lebesgue-integrierbar, und es gilt

$$\int_{\mathbb{R}} f_\infty(x) d\mu_1(x) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} f_n(x) d\mu_1(x) = \lim_{n \rightarrow \infty} \{0\} = 0. \quad (\text{XIII.27})$$

- Wir illustrieren die Anwendung von Satz XIII.12 über majorisierte Konvergenz. Dazu nehmen wir an, dass $f : \mathbb{R} \rightarrow \mathbb{C}$ eine μ_1 -integrierbare Funktion ist, d.h. dass f messbar und $|f| : \mathbb{R} \rightarrow \mathbb{R}_0^+$ μ_1 -integrierbar ist (s. Satz XIII.8). Außerdem soll auch $x \mapsto x f(x)$ integrierbar sein. Wir definieren die **Fouriertransformierte** $\check{H} := \mathcal{F}[f] : \mathbb{R} \rightarrow \mathbb{C}$ von f durch

$$\forall \xi \in \mathbb{R} : \quad \check{H}(\xi) := \int_{\mathbb{R}} e^{-i\xi x} f(x) \frac{d\mu_1(x)}{(2\pi)^{1/2}}. \quad (\text{XIII.28})$$

Dabei verwenden wir, dass $x \mapsto e^{-i\xi x} f(x)$ als Produkt der messbaren Funktionen $x \mapsto e^{-i\xi x}$ und $x \mapsto f(x)$ selbst messbar und wegen $|e^{-i\xi x} f(x)| = |f(x)|$ auch integrabel ist.

Ziel ist nun die Berechnung der Ableitung $\check{H}'$ von $\check{H}$. Wir erwarten, dass

$$\begin{aligned}\check{H}'(\xi) &= \frac{d}{d\xi} \int e^{-i\xi x} f(x) \frac{d\mu_1(x)}{(2\pi)^{1/2}} = \int \frac{d}{d\xi} (e^{-i\xi x}) f(x) \frac{d\mu_1(x)}{(2\pi)^{1/2}} \\ &= -i \int e^{-i\xi x} x f(x) \frac{d\mu_1(x)}{(2\pi)^{1/2}},\end{aligned}\quad (\text{XIII.29})$$

da auch $x \mapsto x \cdot f(x)$ integrabel ist. Die Rechnung (XIII.29) ist jedoch rein formal und kein Beweis, da sowohl das Integral wie auch die Ableitung Operationen sind, die Grenzprozesse involvieren, deren Vertauschbarkeit ("Ableitung unter Integral ziehen") unklar ist und für jeden Einzelfall überprüft werden muss.

In Erwartung des Ergebnisses (XIII.29) fixieren wir $\xi \in \mathbb{R}$ und berechnen für $h \neq 0$

$$\begin{aligned}R_\xi(h) &:= \frac{\check{H}(\xi+h) - \check{H}(\xi)}{h} + i \int e^{-i\xi x} x f(x) \frac{d\mu_1(x)}{(2\pi)^{1/2}} \\ &= \int \left(\frac{e^{-i(\xi+h)x} - e^{-i\xi x}}{h} + i e^{-i\xi x} x \right) f(x) \frac{d\mu_1(x)}{(2\pi)^{1/2}} \\ &= \int e^{-i\xi x} \left(\frac{e^{-ihx} - 1 + ihx}{h} \right) f(x) \frac{d\mu_1(x)}{(2\pi)^{1/2}} = \int f_h(x) \frac{d\mu_1(x)}{(2\pi)^{1/2}},\end{aligned}\quad (\text{XIII.30})$$

wobei (für festes $\xi \in \mathbb{R}$)

$$f_h(x) := e^{-i\xi x} \left(\frac{e^{-ihx} - 1 + ihx}{h} \right) f(x). \quad (\text{XIII.31})$$

Nach dem Satz von Taylor (1. Ordnung und Entwicklungspunkt Null sowie 2. Ordnung und Entwicklungspunkt Null) gibt es $\vartheta_1, \vartheta_2 \in (0, 1)$, sodass

$$e^{-ihx} = 1 + h \frac{\partial[e^{-isx}]}{\partial s} \Big|_{s=\vartheta_1 h} = 1 - ihx e^{-i\vartheta_1 hx}, \quad (\text{XIII.32})$$

$$e^{-ihx} = 1 - ihx + \frac{h^2}{2} \frac{\partial^2[e^{-isx}]}{\partial s^2} \Big|_{s=\vartheta_2 h} = 1 - ihx - \frac{h^2 x^2}{2} e^{-i\vartheta_2 hx}, \quad (\text{XIII.33})$$

und wegen (XIII.33) ist

$$\left| \frac{e^{-ihx} - 1 + ihx}{h} \right| = \frac{h x^2}{2}. \quad (\text{XIII.34})$$

Damit konvergiert $f_h \rightarrow 0$ punktweise, für $h \rightarrow 0$, da für festes $x \in \mathbb{R}$

$$\lim_{h \rightarrow 0} |f_h(x)| = \lim_{h \rightarrow 0} \left\{ \frac{1}{2} h x^2 f(x) \right\} = 0. \quad (\text{XIII.35})$$

Außerdem ist $|x| |f(x)|$ eine integrable Majorante an $f_h(x)$, denn

$$|f_h(x)| = \left| \frac{e^{-ihx} - 1 + ihx}{h} \right| |f(x)| = |e^{-i\vartheta_1 hx} - 1| |x| |f(x)| \leq 2 |x| |f(x)|. \quad (\text{XIII.36})$$

Nach dem Satz XIII.12 über majorisierte Konvergenz erhalten wir also

$$\lim_{h \rightarrow 0} R_\xi(h) = \lim_{h \rightarrow 0} \int f_h(x) \frac{d\mu_1(x)}{(2\pi)^{1/2}} = \int \left\{ \lim_{h \rightarrow 0} f_h(x) \right\} \frac{d\mu_1(x)}{(2\pi)^{1/2}} = 0 \quad (\text{XIII.37})$$

und damit die behauptete Gleichung (XIII.29), die abgekürzt besagt, dass

$$\frac{d}{d\xi} \left(\mathcal{F}[f] \right) = \mathcal{F}[-ix f]. \quad (\text{XIII.38})$$

XIII.4. Ergänzungen

XIII.4.1. Riemann-Integrierbarkeit monotoner Funktionen

Zu Beginn dieses Kapitels beweisen wir zwei wichtige Lemmata über das Riemann-Integral monotoner Funktionen auf $\mathbb{R}^+$. Dazu rufen wir in Erinnerung, dass, für $a, b \in \mathbb{R}$ mit $a < b$, die Menge der Partitionen des Intervalls (a, b) gegeben ist durch

$$\mathcal{P}[a, b] := \bigcup_{N=1}^{\infty} \left\{ (x_0, x_1, x_2, \dots, x_N) \mid a =: x_0 < x_1 < \dots < x_{N-1} < x_N := b \right\}. \quad (\text{XIII.39})$$

Für eine beschränkte Funktion $F : [a, b] \rightarrow \mathbb{R}$ und $P = (x_n)_{n=0}^N \in \mathcal{P}[a, b]$ sind dann

$$\overline{\mathcal{I}}[F, P] := \sum_{n=1}^N (x_n - x_{n-1}) \sup_{x_{n-1} \leq x \leq x_n} \{F(x)\}, \quad (\text{XIII.40})$$

$$\underline{\mathcal{I}}[F, P] := \sum_{n=1}^N (x_n - x_{n-1}) \inf_{x_{n-1} \leq x \leq x_n} \{F(x)\}, \quad (\text{XIII.41})$$

und *Obersumme* und *Untersumme* berechnen sich aus

$$\overline{\int_a^b} F := \inf_{P \in \mathcal{P}[a, b]} \overline{\mathcal{I}}[F, P] \quad \text{und} \quad \underline{\int_a^b} F := \sup_{P \in \mathcal{P}[a, b]} \underline{\mathcal{I}}[F, P]. \quad (\text{XIII.42})$$

Stimmen Ober- und Untersumme überein, so heißt die beschränkte Funktion $F : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar, $F \in \mathcal{R}[a, b]$, und die Obersumme heißt das Integral von F ,

$$\int_a^b F(x) dx := \overline{\int_a^b} F = \underline{\int_a^b} F. \quad (\text{XIII.43})$$

Das uneigentliche Integral, für $a = -\infty$, $b = \infty$ oder unbeschränktes F , ist durch einen geeigneten Grenzprozess definiert.

Wir wollen hier die spezielle Situation studieren, dass $F : \mathbb{R}^+ \rightarrow \mathbb{R}_0^+$ monoton fallend ist. Dann ist die Einschränkung von F auf $[a, b]$ beschränkt, für alle $0 < a < b < \infty$.

Lemma XIII.13. *Ist $F : \mathbb{R}^+ \rightarrow \mathbb{R}_0^+$ monoton fallend, so ist F auf $[a, b]$ Riemann-integrierbar, für alle $0 < a < b < \infty$. Ist darüber hinaus*

$$I_F := \sup_{0 < a < b < \infty} \int_a^b F(x) dx < \infty, \quad (\text{XIII.44})$$

so ist F auch auf $[0, \infty)$ Riemann-integrierbar, und $\int_0^\infty F(x) dx = I_F$.

Beweis. Seien $0 < a < b < \infty$, dann ist $F : [a, b] \rightarrow [0, F(a)]$ beschränkt. Für $N \in \mathbb{N}$, $x_0 := a$, $x_N := b$ und $P = (x_n)_{n=1}^{N-1} \in \mathcal{P}[a, b]$ sind

$$\sup_{x_{n-1} \leq x \leq x_n} \{F(x)\} = F(x_{n-1}) \quad \text{und} \quad \inf_{x_{n-1} \leq x \leq x_n} \{F(x)\} = F(x_n) \quad (\text{XIII.45})$$

Wählen wir speziell $x_n := a + \frac{n}{N}(b-a)$, für $n = 0, 1, \dots, N$, so folgt damit, dass

$$\begin{aligned} 0 &\leq \overline{\int_a^b F} - \underline{\int_a^b F} \leq \overline{\mathcal{I}}[F, P] - \underline{\mathcal{I}}[F, P] \\ &= \frac{b-a}{N} \left(\sum_{n=1}^N F(x_{n-1}) - \sum_{n=1}^N F(x_n) \right) = \frac{(b-a)[F(a) - F(b)]}{N} \rightarrow 0, \end{aligned} \quad (\text{XIII.46})$$

für $N \rightarrow \infty$. Somit ist F auf $[a, b]$ Riemann-integrierbar. Da $F \geq 0$, gilt außerdem

$$\forall 0 < a' \leq a < b \leq b' : \quad \int_a^b F(x) dx \leq \int_{a'}^{b'} F(x) dx. \quad (\text{XIII.47})$$

Daher ist die Folge $(I_{k,\ell})_{k,\ell=1}^\infty$, wobei

$$I_{k,\ell} := \int_{1/k}^\ell F(x) dx, \quad (\text{XIII.48})$$

monoton wachsend in k und ℓ , und es gilt

$$I_F = \lim_{k \rightarrow \infty} \lim_{\ell \rightarrow \infty} I_{k,\ell} = \lim_{\ell \rightarrow \infty} \lim_{k \rightarrow \infty} I_{k,\ell}. \quad (\text{XIII.49})$$

□

Das nächste Lemma zeigt, dass für monotone Funktionen nicht nur die Riemann-Integrierbarkeit leicht zu zeigen ist, sondern dass auch die Vertauschung des Integrals mit dem Grenzwert für monotone Folgen monotoner Funktionen unproblematisch ist.

Lemma XIII.14. *Sei $(F_k)_{k=1}^\infty$ eine monoton wachsende, punktweise konvergente Folge nichtnegativer, monoton fallender, Riemann-integrierbarer Funktionen auf $\mathbb{R}^+$, deren Integrale konvergieren, d.h. $F_k : \mathbb{R}^+ \rightarrow \mathbb{R}_0^+$ besitze folgende Eigenschaften:*

$$\forall 0 < x < x' < \infty : \quad F_k(x) \geq F_k(x') \geq 0, \quad (\text{XIII.50})$$

$$\forall x \in \mathbb{R}^+ : \quad F_k(x) \leq F_{k+1}(x) \leq F(x) := \lim_{k \rightarrow \infty} F_k(x) < \infty. \quad (\text{XIII.51})$$

Dann gelten

(i) Die Funktion $F := \lim_{k \rightarrow \infty} F_k : \mathbb{R}^+ \rightarrow \mathbb{R}_0^+$ ist genau dann Riemann-integrierbar auf $\mathbb{R}^+$, wenn

$$\lim_{k \rightarrow \infty} \int_0^\infty F_k(x) dx < \infty. \quad (\text{XIII.52})$$

(ii) Gilt (XIII.52), so stimmen auch die Integrale überein,

$$\int_0^\infty F(x) dx = \int_0^\infty \left(\lim_{k \rightarrow \infty} F_k(x) \right) dx = \lim_{k \rightarrow \infty} \int_0^\infty F_k(x) dx. \quad (\text{XIII.53})$$

Beweis. Zunächst stellen wir fest, dass mit F_k auch F monoton fallend ist. Sind nämlich $0 < x < x' < \infty$, so ist $F_k(x) \geq F_k(x')$, und deshalb gilt

$$F(x) = \lim_{k \rightarrow \infty} F_k(x) \geq \lim_{k \rightarrow \infty} F_k(x') = F(x'). \quad (\text{XIII.54})$$

Daher ist $F : [a, b] \rightarrow [0, F(a)]$, für $0 < a < b < \infty$, beschränkt, monoton fallend und nach Lemma XIII.13 auch Riemann-integrierbar. Wählen wir wieder $N \in \mathbb{N}$, $x_0 := a$, $x_N := b$ und die Partition $P = (x_n)_{n=1}^{N-1} \in \mathcal{P}[a, b]$ mit $x_n := a + \frac{n}{N}(b - a)$, für $n = 0, 1, \dots, N$, so folgt, für alle $k \in \mathbb{N}$, dass

$$\begin{aligned} 0 &\leq \overline{\int_a^b F} - \underline{\int_a^b F_k} \leq \overline{\mathcal{I}}[F, P] - \underline{\mathcal{I}}[F_k, P] \\ &= \frac{b-a}{N} \sum_{n=1}^N \{F(x_{n-1}) - F_k(x_n)\} \\ &= \frac{b-a}{N} \left(\sum_{n=1}^N \{F(x_{n-1}) - F(x_n)\} + \sum_{n=1}^N \{F(x_n) - F_k(x_n)\} \right) \\ &\leq \frac{(b-a)[F(a) - F(b)]}{N} + (b-a) \max_{1 \leq n \leq N} \{F(x_n) - F_k(x_n)\}. \end{aligned} \quad (\text{XIII.55})$$

Ist nun $\varepsilon > 0$, dann wählen wir erst $N \in \mathbb{N}$ so groß, dass

$$\frac{(b-a)[F(a) - F(b)]}{N} \leq \frac{\varepsilon}{2}, \quad (\text{XIII.56})$$

und anschließend wählen wir $k_0 \in \mathbb{N}$ so groß, dass

$$\forall k \geq k_0 : \quad (b-a) \max_{1 \leq n \leq N} \{F(x_n) - F_k(x_n)\} \leq \frac{\varepsilon}{2}. \quad (\text{XIII.57})$$

Zu jedem $\varepsilon > 0$ gibt es also ein $k_0 \in \mathbb{N}$, so dass

$$\forall k \geq k_0 : \quad 0 \leq \overline{\int_a^b F} - \underline{\int_a^b F_k} \leq \varepsilon. \quad (\text{XIII.58})$$

Da $F_k \leq F$ auf $[a, b]$, folgt aus (XIII.58) insbesondere, dass F auf $[a, b]$ Riemann-integrierbar ist und dass

$$\int_a^b F(x) dx = \lim_{k \rightarrow \infty} \int_a^b F_k(x) dx, \quad (\text{XIII.59})$$

für alle $0 < a < b < \infty$.

Existiert nun $\lim_{k \rightarrow \infty} \int_0^\infty F_k(x) dx < \infty$, so folgt aus (XIII.59), dass

$$\int_a^b F(x) dx \leq \lim_{k \rightarrow \infty} \int_0^\infty F_k(x) dx, \quad (\text{XIII.60})$$

für alle $0 < a < b < \infty$, also

$$\int_0^\infty F(x) dx = \sup_{0 < a < b < \infty} \int_a^b F(x) dx \leq \lim_{k \rightarrow \infty} \int_0^\infty F_k(x) dx < \infty, \quad (\text{XIII.61})$$

und F ist auf $\mathbb{R}^+$ Riemann-integrierbar. Ist umgekehrt F auf $\mathbb{R}^+$ Riemann-integrierbar, so folgt sofort aus $F_k \nearrow F$, dass

$$\lim_{k \rightarrow \infty} \int_0^\infty F_k(x) dx \leq \int_0^\infty F(x) dx < \infty \quad (\text{XIII.62})$$

und $\lim_{k \rightarrow \infty} \int_0^\infty F_k(x) dx < \infty$. Aus (XIII.61)-(XIII.62) ergibt sich unmittelbar die Äquivalenz (i) und die Identität (ii). $\square$

XIII.4.2. Messbare Funktionen

Nach diesen vorbereitenden Lemmata über Riemann-Integrale monotoner Funktionen kommen wir nun zur Definition des Messbarkeitsbegriffs.

Definition XIII.15. Seien $(\Omega, \mathfrak{A})$ und $(\Omega', \mathfrak{A}')$ zwei Messräume. Eine Abbildung $f : \Omega \rightarrow \Omega'$ heißt **messbar** $:\Leftrightarrow$

$$\forall A' \in \mathfrak{A}' : f^{-1}(A') \in \mathfrak{A}. \quad (\text{XIII.63})$$

Ist speziell $(\Omega', \mathfrak{A}') = (\overline{\mathbb{R}}, \overline{\mathfrak{B}}_1)$, wobei wir vereinbaren, dass $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ und $\overline{\mathfrak{B}}_1 := \mathfrak{A}(\mathfrak{B}_1 \cup \{\{-\infty\}, \{\infty\}\})$, so heißt eine messbare Abbildung $\Omega \rightarrow \overline{\mathbb{R}}$ **reell-messbar**.

Wir werden in dieser Vorlesung nur *reell-messbare* Funktionen behandeln, es wird also stets $(\Omega', \mathfrak{A}') = (\overline{\mathbb{R}}, \overline{\mathfrak{B}}_1)$ sein. Das folgende, wichtige Lemma besagt, dass die Menge der reell-messbaren Funktionen unter abzählbaren Supremums- und Infimumsbildung, sowie den Limes Superior und Limes Inferior abgeschlossen ist.

Lemma XIII.16. Seien $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $(f_k)_{k=1}^\infty$ eine Folge reell-messbarer Funktionen. Dann sind $\sup_k f_k$, $\inf_k f_k$, $\limsup_k f_k$, $\liminf_k f_k$, und, falls für alle $\omega \in \Omega$ existent, auch $\lim_k f_k$ reell-messbar.

Beweis. Wir beweisen nur exemplarisch, dass $f := \sup_k f_k$ reell-messbar ist, und dazu schreiben wir $f^{-1}(A) = \{\omega \in \Omega | f(\omega) \in A\} =: \{f(\omega) \in A\}$ u. s. w. Für $b \in \mathbb{R}$ ist

$$f^{-1}([-\infty, b)) = \left\{ \sup_k f_k(\omega) < b \right\} = \bigcap_{k=1}^\infty \{f_k(\omega) < b\}. \quad (\text{XIII.64})$$

Da f_k reell-messbar ist, ist $\{f_k(\omega) < b\} \in \mathfrak{A}$. Weil $\mathfrak{A}$ unter Differenzbildung und abzählbarer Durchschnittsbildung abgeschlossen ist, ist also auch $f^{-1}([-\infty, b)) \in \mathfrak{A}$. Somit ist auch

$$f^{-1}([a, b)) = f^{-1}([-\infty, b)) \setminus f^{-1}([-\infty, a)) \in \mathfrak{A}, \quad (\text{XIII.65})$$

für alle $a < b$. Ebenso ist

$$f^{-1}(\{\infty\}) = f^{-1}([0, \infty]) \setminus \bigcup_{n=1}^{\infty} f^{-1}([0, n]) \in \mathfrak{A}, \quad (\text{XIII.66})$$

und analog $f^{-1}(\{-\infty\}) \in \mathfrak{A}$. Wir definieren nun

$$\mathfrak{C}_1 := \left\{ A \in \mathfrak{P}(\overline{\mathbb{R}}) \mid f^{-1}(A) \in \mathfrak{A} \right\}, \quad (\text{XIII.67})$$

und erhalten also aus (XIII.65) und (XIII.66), dass

$$\Omega_1 \cup \{\{\infty\}, \{-\infty\}\} \subseteq \mathfrak{C}_1. \quad (\text{XIII.68})$$

Weiterhin beobachten wir, dass $\mathfrak{C}_1$ eine Sigma-Algebra ist, weil dies auf $\mathfrak{A}$ zutrifft und das Urbild f^{-1} mit mengentheoretischen Operationen vertauscht. Da die halboffenen Intervalle Ω_1 zusammen mit $\{\infty\}$ und $\{-\infty\}$ die Sigma-Algebra $\overline{\mathfrak{B}_1}$ erzeugen, folgt also $\overline{\mathfrak{B}_1} \subseteq \mathfrak{C}_1$. $\square$

Die Klasse der reell-messbaren Funktionen ist genügend groß um alle erdenklichen Beispiele miteinzuschließen – auch die meisten “pathologischen” Beispiele, wie etwa die Dirichlet-Funktion. In der Tat ist die Konstruktion konkreter nicht-messbarer Funktionen genauso schwierig wie etwa die Konstruktion von Mengen in $\mathfrak{P}(\mathbb{R}^d) \setminus \mathfrak{A}^*(\mathfrak{F}_d)$.

Wir wollen nun das Integral für nichtnegative Funktionen eines Maßraums definieren.

Definition XIII.17. Seien $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $f : \Omega \rightarrow [0, \infty]$ eine reell-messbare Funktion. Für $\lambda > 0$ definieren wir die **Niveaumengen** von f durch

$$A(f, \lambda) := f^{-1}((\lambda, \infty]) = \{\omega \in \Omega \mid f(\omega) > \lambda\}. \quad (\text{XIII.69})$$

Nach Definition XIII.15 ist $A(f, \lambda) \in \mathfrak{A}$, und

$$F(f, \lambda) := \mu(A(f, \lambda)) = \mu[f^{-1}((\lambda, \infty])] \quad (\text{XIII.70})$$

definiert eine Abbildung $F(f, \cdot) : \mathbb{R}^+ \rightarrow [0, \infty]$.

Offensichtlich ist $F(f, \lambda)$ stets eine in λ monoton fallende Funktion, denn für $\lambda < \lambda'$ ist $(\lambda, \infty] \supseteq (\lambda', \infty]$, daher auch

$$A(f, \lambda) = f^{-1}((\lambda, \infty]) \supseteq f^{-1}((\lambda', \infty]) = A(f, \lambda'), \quad (\text{XIII.71})$$

und somit

$$F(f, \lambda) = \mu(A(f, \lambda)) \geq \mu(A(f, \lambda')) = F(f, \lambda'). \quad (\text{XIII.72})$$

Nach Lemma XIII.13 ist dann $F(f, \cdot)$ auf $[a, b]$ Riemann-integrierbar, falls $0 < a < b < \infty$ so gewählt sind, dass $F(f, a) < \infty$.

Definition XIII.18. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum.

- (i) Eine reell-messbare Funktion $f : \Omega \rightarrow [0, \infty]$ heißt **(μ -)integrabel**, falls $F(f, \cdot)$ auf $\mathbb{R}^+$ Riemann-integrabel ist. In diesem Fall heißt

$$\int_{\Omega} f(\omega) d\mu(\omega) := \int_0^{\infty} F(f, \lambda) d\lambda \quad (\text{XIII.73})$$

Integral von f .

- (ii) Eine reell-messbare Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ heißt **(μ -)integrabel**, falls $f_+ := \max\{f, 0\}$ und $f_- := \max\{-f, 0\}$ jeweils μ -integrabel sind. Das **Integral** von f ist dann definiert durch

$$\int_{\Omega} f(\omega) d\mu(\omega) := \int_{\Omega} f_+(\omega) d\mu(\omega) - \int_{\Omega} f_-(\omega) d\mu(\omega). \quad (\text{XIII.74})$$

- (iii) Eine komplexe Funktion $f : \Omega \rightarrow \mathbb{C}$ heißt **messbar**, falls $\operatorname{Re}\{f\}$ und $\operatorname{Im}\{f\}$ jeweils reell-messbar sind. Sind darüber hinaus $\operatorname{Re}\{f\}$ und $\operatorname{Im}\{f\}$ jeweils μ -integrabel, so ist f **(μ -)integrabel**, und ihr **Integral** ist definiert durch

$$\int_{\Omega} f(\omega) d\mu(\omega) := \int_{\Omega} \operatorname{Re}\{f(\omega)\} d\mu(\omega) + i \int_{\Omega} \operatorname{Im}\{f(\omega)\} d\mu(\omega). \quad (\text{XIII.75})$$

Bemerkungen und Beispiele.

- Sind $M \in \mathfrak{A}$ eine messbare Menge eines Maßraums $(\Omega, \mathfrak{A}, \mu)$ und

$$\mathbb{1}_M(\omega) \equiv \mathbb{1}[\omega \in M] := \begin{cases} 1 & \text{falls } \omega \in M, \\ 0 & \text{falls } \omega \notin M, \end{cases} \quad (\text{XIII.76})$$

so ist, für $\eta > 0$,

$$A(\mathbb{1}_M, \lambda) = \{\omega \in \Omega \mid \mathbb{1}_M(\omega) > \eta\} = \begin{cases} \emptyset, & \text{falls } \eta > 1, \\ M, & \text{falls } 0 < \eta \leq 1. \end{cases} \quad (\text{XIII.77})$$

Somit erhalten wir

$$F(\mathbb{1}_M, \lambda) = \mu(M) \cdot \mathbb{1}[0 < \eta \leq 1] \quad (\text{XIII.78})$$

und

$$\int_{\Omega} \mathbb{1}_M d\mu(\omega) = \mu(M). \quad (\text{XIII.79})$$

- Insbesondere folgt aus (XIII.79), mit $M := \{\omega \in \Omega \mid f(\omega) > \lambda\}$, wobei $\lambda > 0$ und f messbar sind, dass

$$\int_{\Omega} \mathbb{1}[f(\omega) > \lambda] d\mu(\omega) = \mu(\{f(\omega) > \lambda\}) = F(f, \lambda). \quad (\text{XIII.80})$$

- Zur Rechtfertigung von (XIII.73) dient folgende formale “Herleitung”, die eine nützliche *Merkhilfe* (aber auch nicht mehr als das) darstellt:

$$\begin{aligned}
 \int_0^\infty F(f, \lambda) d\lambda &= \int_0^\infty \left(\int_\Omega \mathbf{1}[f(\omega) > \lambda] d\mu(\omega) \right) d\lambda \\
 &= \int_\Omega \left(\int_0^\infty \mathbf{1}[f(\omega) > \lambda] d\lambda \right) d\mu(\omega) \\
 &= \int_\Omega \left(\int_0^{f(\omega)} d\lambda \right) d\mu(\omega) = \int_\Omega f(\omega) d\mu(\omega). \quad (\text{XIII.81})
 \end{aligned}$$

(Problematisch ist in (XIII.81) die Vertauschung der Integrationen, d.h. die zweite Gleichung.)

XIII.4.3. Die Konvergenzsätze

Der große Vorteil der Definition XIII.18 liegt aber in der Mühelosigkeit, mit der wir mit ihrer Hilfe die *Konvergenzsätze* beweisen können. Das ist der Gegenstand der folgenden Untersuchungen. Dafür definieren wir noch zuvor den Begriff “fast überall”.

Definition XIII.19. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum. Eine Eigenschaft $\mathcal{E} : \Omega \rightarrow \{w, f\}$ (w = wahr, f = falsch) gilt μ -**fast überall** $:\Leftrightarrow$

$$\mathcal{E}^{-1}(f) \in \mathfrak{A} \quad \text{und} \quad \mu(\mathcal{E}^{-1}(f)) = 0, \quad (\text{XIII.82})$$

d.h., die Menge, auf der $\mathcal{E}$ nicht gilt, hat Maß Null.

Wir kommen zum ersten Konvergenzsatz.

Satz XIII.20 (Monotone Konvergenz). *Seien $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $(f_k)_{k=1}^\infty$ eine monoton steigende Folge nichtnegativer, μ -integrabler Funktionen, d.h., $f_k : \Omega \rightarrow [0, \infty]$, $f_k \leq f_{k+1}$, und $\int_\Omega f_k(\omega) d\mu(\omega) < \infty$, für alle $k \in \mathbb{N}$. Dann ist $f := \lim_k f_k = \sup_k f_k$ reell-messbar und genau dann μ -integrabel, wenn*

$$\lim_{k \rightarrow \infty} \int_\Omega f_k(\omega) d\mu(\omega) < \infty, \quad (\text{XIII.83})$$

und in diesem Fall gilt

$$\int_\Omega f(\omega) d\mu(\omega) = \lim_{k \rightarrow \infty} \int_\Omega f_k(\omega) d\mu(\omega). \quad (\text{XIII.84})$$

Beweis. Die Reell-Messbarkeit von f haben wir bereits in Lemma XIII.16 bewiesen. Sei nun $\lambda > 0$. Aus der Monotonie von $(f_k)_{k=1}^\infty$ folgt sofort, dass $\{f_k > \lambda\} \subseteq \{f_{k+1} > \lambda\} \subseteq \{f > \lambda\}$. Sei andererseits $\omega \in \{f > \lambda\}$, was gleichwertig ist mit $\sup_k f_k(\omega) > \lambda$. Dann gibt es aber ein $k_0 \equiv k_0(\omega) \in \mathbb{N}$, so dass $f_{k_0}(\omega) > \lambda$, und damit $\omega \in \{f_{k_0} > \lambda\}$. Also ist

$$\bigcup_{k=1}^\infty \{f_k > \lambda\} \subseteq \{f > \lambda\} \subseteq \bigcup_{k=1}^\infty \{f_k > \lambda\}, \quad (\text{XIII.85})$$

und es muss tatsächlich Gleichheit überall in (XIII.85) gelten. Schreiben wir nun

$$B_k := \{f_k > \lambda\} \setminus \bigcup_{j=1}^{k-1} \{f_j > \lambda\}, \quad (\text{XIII.86})$$

so ist $(B_k)_{k=1}^\infty$ eine Folge paarweise disjunkter Mengen in $\mathfrak{A}$, und es gilt

$$F(f_k, \lambda) = \mu(\{f_k > \lambda\}) = \mu\left(\bigcup_{j=1}^k B_j\right), \quad (\text{XIII.87})$$

$$F(f, \lambda) = \mu(\{f > \lambda\}) = \mu\left(\bigcup_{j=1}^\infty B_j\right). \quad (\text{XIII.88})$$

Somit folgt aus der Sigma-Additivität von μ , dass

$$F_k(\lambda) := F(f_k, \lambda) = \sum_{j=1}^k \mu(B_j) \nearrow \sum_{j=1}^\infty \mu(B_j) = F(f, \lambda) =: F(\lambda), \quad (\text{XIII.89})$$

für $k \rightarrow \infty$. Damit ist $(F_k)_{k=1}^\infty$ eine monoton steigende Folge nichtnegativer, monoton fallender Funktionen auf $\mathbb{R}^+$ mit punktwisen Limes $F = \lim_k F_k$. Im Fall, dass $F < \infty$ auf $\mathbb{R}^+$, folgt nun die Behauptung direkt aus Lemma XIII.14. Gibt es umgekehrt ein $\lambda_0 > 0$, so dass $F(\lambda_0) = \infty$, so sind beide Seiten der behaupteten Gleichung (XIII.84)) unendlich, und die Behauptung gilt trivialerweise. $\square$

Lemma XIII.21 (Lemma von Fatou). *Seien $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $(f_k)_{k=1}^\infty$ eine Folge nichtnegativer, μ -integrabler Funktionen, d.h., $f_k : \Omega \rightarrow [0, \infty]$ und $\int_\Omega f_k(\omega) d\mu(\omega) < \infty$, für alle $k \in \mathbb{N}$. Ist darüber hinaus $\liminf_{k \rightarrow \infty} \int_\Omega f_k(\omega) d\mu(\omega) < \infty$, so ist $f := \liminf_k f_k$ reell-messbar, μ -integrabel, und es gilt*

$$\liminf_{k \rightarrow \infty} \int_\Omega f_k(\omega) d\mu(\omega) \geq \int_\Omega f(\omega) d\mu(\omega). \quad (\text{XIII.90})$$

Beweis. Definieren wir $\tilde{f}_k(\omega) := \inf_{j \geq k} f_j(\omega)$, so ist $(\tilde{f}_k)_{k=1}^\infty$ eine monoton steigende Folge reell-messbarer Funktionen mit $\lim_{k \rightarrow \infty} \tilde{f}_k(\omega) = \liminf_{k \rightarrow \infty} f_k(\omega) = f(\omega)$. Außerdem ist $\tilde{f}_k$ μ -integrabel, da $\tilde{f}_k \leq f_k$. Nach Satz XIII.20 der monotonen Konvergenz gilt deshalb

$$\begin{aligned} \int_\Omega f(\omega) d\mu(\omega) &= \lim_{k \rightarrow \infty} \int_\Omega \tilde{f}_k(\omega) d\mu(\omega) = \liminf_{k \rightarrow \infty} \int_\Omega \tilde{f}_k(\omega) d\mu(\omega) \\ &\leq \liminf_{k \rightarrow \infty} \int_\Omega f_k(\omega) d\mu(\omega), \end{aligned} \quad (\text{XIII.91})$$

wobei letztere Ungleichung abermals aus $\tilde{f}_k \leq f_k$ folgt. $\square$

Wir kommen nun zu einem der wichtigsten Sätze der Analysis, dem Satz über majorisierte Konvergenz³.

³engl.: dominated convergence

Satz XIII.22 (Majorisierte Konvergenz). Seien $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $(f_k)_{k=1}^\infty$ eine punktweise, μ -fast überall konvergente Folge komplexer, μ -integrabler Funktionen, die durch eine μ -integrable Funktion $g : \Omega \rightarrow [0, \infty]$ majorisiert wird, d.h. $f_k : \Omega \rightarrow \mathbb{C}$, $f(\omega) := \lim_k f_k(\omega) \in \mathbb{C}$, μ -fast überall, $|f_k| \leq g$, für alle $k \in \mathbb{N}$, und $\int_\Omega g(\omega) d\mu(\omega) < \infty$. Dann ist f μ -integrabel, und es gilt

$$\int_\Omega f(\omega) d\mu(\omega) = \lim_{k \rightarrow \infty} \int_\Omega f_k(\omega) d\mu(\omega). \quad (\text{XIII.92})$$

Beweis. Es reicht, die Behauptung für nichtnegative f_k zu zeigen, denn wir können stets $f_k = (\operatorname{Re} f_k)_+ - (\operatorname{Re} f_k)_- + i(\operatorname{Im} f_k)_+ - i(\operatorname{Im} f_k)_-$ zerlegen, wobei jeder der vier Terme auf der rechten Seite aus einer nichtnegativen Funktion resultiert.

Sei also f_k nichtnegativ, konvergent, $f_k \rightarrow f$, μ -fast überall, und majorisiert durch g , d. h. $f_k \leq g$. Die Nullmenge, auf der f_k nicht konvergiert, beeinflusst das Integral nicht, und wir können das Lemma von Fatou auf $(f_k)_{k=1}^\infty$ und $f = \lim_k f_k = \liminf_k f_k$ anwenden. Damit ist f integrabel, und wir erhalten

$$\int_\Omega f(\omega) d\mu(\omega) \leq \liminf_{k \rightarrow \infty} \int_\Omega f_k(\omega) d\mu(\omega). \quad (\text{XIII.93})$$

Anschließend können wir das Lemma von Fatou auf $(g - f_k)_{k=1}^\infty$ und $g - f = \lim_k (g - f_k) = \liminf_k (g - f_k)$ anwenden, da $g - f_k$ dieselben Voraussetzungen wie f_k erfüllt. Beachten wir noch die Linearität (XIII.100) des Integrals (für deren Beweis wir zwar den Satz XIII.20 über die monotone Konvergenz brauchen, aber nicht die hier zu beweisende majorisierte Konvergenz), so erhalten wir

$$\begin{aligned} \int_\Omega f(\omega) d\mu(\omega) &= \int_\Omega g(\omega) d\mu(\omega) - \int_\Omega \{g(\omega) - f(\omega)\} d\mu(\omega) \\ &\geq \int_\Omega g(\omega) d\mu(\omega) - \liminf_{k \rightarrow \infty} \int_\Omega \{g(\omega) - f_k(\omega)\} d\mu(\omega) \\ &= \limsup_{k \rightarrow \infty} \int_\Omega f_k(\omega) d\mu(\omega). \end{aligned} \quad (\text{XIII.94})$$

Somit müssen $\liminf_{k \rightarrow \infty} \int_\Omega f_k(\omega) d\mu(\omega)$ und $\limsup_{k \rightarrow \infty} \int_\Omega f_k(\omega) d\mu(\omega)$ mit $\int_\Omega f(\omega) d\mu(\omega)$ übereinstimmen. Dies impliziert aber die Konvergenz von $\int_\Omega f_k(\omega) d\mu(\omega)$ gegen $\int_\Omega f(\omega) d\mu(\omega)$. $\square$

XIII.4.4. Eigenschaften des Integrals

Es gibt verschiedene Wege zur Einführung des Integrals $\int_\Omega f(\omega) d\mu(\omega)$ einer reell-messbaren Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$, der hier vorgestellte folgt Lieb und Loss⁴. Ein anderer Weg zum Integral führt über *Treppenfunktionen*, d.h. Funktionen, die auf Ω nur endlich viele Werte annehmen. Das folgende Lemma zeigt, dass sich in der Tat jede (nichtnegative) messbare Funktion als (monotoner) Limes von Treppenfunktionen darstellen lässt.

Lemma XIII.23. Seien $(\Omega, \mathfrak{A})$ ein messbarer Raum und $f : \Omega \rightarrow [0, \infty]$ reell-messbar. Dann gibt es eine Folge $(f_L)_{L=1}^\infty$ reell-messbarer Funktionen $f_L : \Omega \rightarrow \mathbb{R}_0^+$ mit den folgenden Eigenschaften:

⁴Analysis, AMS Publications, 1997.

- (i) Zu jedem $L \in \mathbb{N}$ gibt es nichtnegative Zahlen $\alpha_L^{(1)}, \alpha_L^{(2)}, \dots, \alpha_L^{(2^L)} \in \mathbb{R}_0^+$ und paarweise disjunkte Mengen $A_L^{(1)}, A_L^{(2)}, \dots, A_L^{(2^L)} \in \mathfrak{A}$, so dass

$$f_L = \alpha_L^{(1)} \mathbb{1}_{A_L^{(1)}} + \alpha_L^{(2)} \mathbb{1}_{A_L^{(2)}} + \dots + \alpha_L^{(2^L)} \mathbb{1}_{A_L^{(2^L)}}. \quad (\text{XIII.95})$$

Funktionen dieser Form heißen **Treppenfunktionen**.

- (ii) Für alle $\omega \in \Omega$ gilt $f_L(\omega) \leq f_{L+1}(\omega) \nearrow f(\omega)$, im Limes $L \rightarrow \infty$.

Beweis. Für $L \in \mathbb{N}$ setzen wir

$$f_L(\omega) := \left(\sum_{\ell=1}^{2^L} \left(\frac{\ell-1}{2^L} \right) \cdot \mathbb{1} \left[\frac{\ell-1}{2^L} \leq f(\omega) < \frac{\ell}{2^L} \right] \right) + 2^L \cdot \mathbb{1}[f(\omega) \geq 2^L]. \quad (\text{XIII.96})$$

Dann ist f_L offensichtlich eine Treppenfunktion und es gilt $f_L \leq f$.

Seien $\omega \in \Omega$ so, dass $f(\omega) < \infty$ und $k \in \mathbb{N}$ und $\tau \in \{0, 1\}$ so, dass $2^{-L-1}(2k - \tau - 1) \leq f(\omega) < 2^{-L-1}(2k - \tau)$. Dann ist $f(\omega) < 2^{-L}k$ und somit

$$f_L(\omega) \leq \frac{k-1}{2^L} \leq \frac{k - \frac{1}{2}(1+\tau)}{2^L} = \frac{2k - \tau - 1}{2^{L+1}} = f_{L+1}(\omega). \quad (\text{XIII.97})$$

Da offensichtlich $f_L(\omega) = 2^L \leq 2^{L+1} = f_{L+1}(\omega)$ gilt, falls $f(\omega) = \infty$, folgt, dass

$$f_L \leq f_{L+1}. \quad (\text{XIII.98})$$

Ist $f(\omega) = \infty$, so ist $f_L(\omega) = 2^L \nearrow \infty$, für $L \rightarrow \infty$. Ist hingegen $f(\omega) < \infty$, so gilt für $L > \log_2[f(\omega)]$, dass

$$f_L(\omega) \leq f(\omega) < f_L(\omega) + 2^{-L}, \quad (\text{XIII.99})$$

und somit gilt auch in diesem Fall $f_L(\omega) \nearrow f(\omega)$. $\square$

Satz XIII.24. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum.

- (i) Ist $f : \Omega \rightarrow \mathbb{C}$ eine messbare Funktion, dann ist auch $|f| : \Omega \rightarrow \mathbb{R}_0^+$ reell-messbar. f ist genau dann μ -integrabel, wenn $|f|$ μ -integrabel ist.
- (ii) Sind $f, g : \Omega \rightarrow \mathbb{C}$ μ -integrabel und $\alpha \in \mathbb{C}$, so ist auch $f + \alpha g : \Omega \rightarrow \mathbb{C}$ μ -integrabel, und es gilt

$$\int_{\Omega} (f + \alpha g)(\omega) d\mu(\omega) = \int_{\Omega} f(\omega) d\mu(\omega) + \alpha \int_{\Omega} g(\omega) d\mu(\omega). \quad (\text{XIII.100})$$

Beweis. Wir beweisen nur (ii) und auch dies nur für $\alpha = 1$ und für $f, g : \Omega \rightarrow \mathbb{R}_0^+$. Dazu beweisen wir (ii) zunächst für zwei Treppenfunktionen

$$\tilde{f} := \sum_{i=1}^M \alpha_i \mathbb{1}_{A_i}, \quad \tilde{g} := \sum_{j=1}^N \beta_j \mathbb{1}_{B_j}, \quad (\text{XIII.101})$$

mit $\alpha_1, \dots, \alpha_M, \beta_1, \dots, \beta_N \in \mathbb{R}^+$ und $A_1, \dots, A_M, B_1, \dots, B_N \in \mathfrak{A}$ mit $\mu(A_i), \mu(B_j) < \infty$ und $A_k \cap A_\ell = B_k \cap B_\ell = \emptyset$, für $k \neq \ell$. Setzen wir

$$\forall i \in \mathbb{N}_1^M, j \in \mathbb{N}_1^N : \quad \alpha_{i,j} := \alpha_i, \beta_{i,j} := \beta_j \quad (\text{XIII.102})$$

und $C_{ij} := A_i \cap B_j$, so sind auch die Mengen $C_{ij} \in \mathfrak{A}$ paarweise disjunkt, und es gelten

$$\tilde{f} = \sum_{i=1}^M \sum_{j=1}^N \alpha_{i,j} \mathbb{1}_{C_{ij}}, \quad \tilde{g} = \sum_{i=1}^M \sum_{j=1}^N \beta_{i,j} \mathbb{1}_{C_{ij}}, \quad (\text{XIII.103})$$

und

$$\tilde{f} + \tilde{g} = \sum_{i=1}^M \sum_{j=1}^N (\alpha_{i,j} + \beta_{i,j}) \mathbb{1}_{C_{ij}}. \quad (\text{XIII.104})$$

Beachte nun, dass

$$\begin{aligned} \mu\left(\{\omega \in \Omega \mid \tilde{f}(\omega) > \lambda\}\right) &= \mu\left(\left\{\omega \in \Omega \mid \sum_{i,j} \alpha_{i,j} \mathbb{1}_{C_{ij}}(\omega) > \lambda\right\}\right) \\ &= \sum_{i=1}^M \sum_{j=1}^N \mathbb{1}[\alpha_{i,j} > \lambda] \cdot \mu(C_{ij}). \end{aligned} \quad (\text{XIII.105})$$

Also ist

$$\begin{aligned} \int_{\Omega} \tilde{f}(\omega) d\mu &= \int_0^{\infty} \sum_{i=1}^M \sum_{j=1}^N \mu(C_{ij}) \mathbb{1}[\alpha_{i,j} > \lambda] d\lambda \\ &= \sum_{i=1}^M \sum_{j=1}^N \mu(C_{ij}) \left(\int_0^{\alpha_{i,j}} d\lambda \right) = \sum_{i=1}^M \sum_{j=1}^N \alpha_{i,j} \cdot \mu(C_{ij}). \end{aligned} \quad (\text{XIII.106})$$

Genauso folgen

$$\int_{\Omega} \tilde{g}(\omega) d\mu(\omega) = \sum_{i=1}^M \sum_{j=1}^N \beta_{i,j} \mu(C_{ij}), \quad (\text{XIII.107})$$

$$\int_{\Omega} (\tilde{f} + \tilde{g})(\omega) d\mu(\omega) = \sum_{i=1}^M \sum_{j=1}^N (\alpha_{i,j} + \beta_{i,j}) \mu(C_{ij}), \quad (\text{XIII.108})$$

und daher auch

$$\int_{\Omega} (\tilde{f} + \tilde{g})(\omega) d\mu(\omega) = \int_{\Omega} \tilde{f}(\omega) d\mu(\omega) + \int_{\Omega} \tilde{g}(\omega) d\mu(\omega). \quad (\text{XIII.109})$$

Sind nun $f, g : \Omega \rightarrow \mathbb{R}_0^+$ beliebige, μ -integrable Funktionen, so gibt es gemäß Lemma XIII.23 Folgen $(f_L)_{L=1}^{\infty}$ und $(g_L)_{L=1}^{\infty}$ von Treppenfunktionen mit $f_L \nearrow f$ und $g_L \nearrow g$,

μ -fast überall. Nach dem Satz über monotone Konvergenz und mit (XIII.109) folgt somit, dass

$$\begin{aligned} \int_{\Omega} (f + g)(\omega) d\mu(\omega) &= \lim_{L \rightarrow \infty} \left\{ \int_{\Omega} (f_L + g_L)(\omega) d\mu(\omega) \right\} \\ &= \lim_{L \rightarrow \infty} \left\{ \int_{\Omega} f_L(\omega) d\mu(\omega) + \int_{\Omega} g_L(\omega) d\mu(\omega) \right\} \\ &= \int_{\Omega} f(\omega) d\mu(\omega) + \int_{\Omega} g(\omega) d\mu(\omega). \end{aligned} \quad (\text{XIII.110})$$

□

Der nun folgende Satz zeigt, dass das Lebesgue-Integral auf $\mathbb{R}$ wirklich eine Verallgemeinerung des Riemann-Integrals ist.

Satz XIII.25. *Ist $f \in \mathcal{R}(\mathbb{R}; \mathbb{R}_0^+)$ eine nichtnegative Riemann-integrierbare Funktion, so gibt es eine reell-messbare und Lebesgue-integrierbare Funktion $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}_0^+$, die μ_1 -fast überall mit f auf $\mathbb{R}$ übereinstimmt und für die gilt, dass*

$$\underbrace{\int_{-\infty}^{\infty} f(x) dx}_{\text{Riemann-Integral}} = \underbrace{\int_{\mathbb{R}} \tilde{f}(x) d\mu_1(x)}_{\text{Lebesgue-Integral}}. \quad (\text{XIII.111})$$

Beweis. Wir beweisen Satz XIII.25 nur für den Fall, dass $\text{supp}(f) \subseteq (a, b)$, für geeignete $-\infty < a < b < \infty$ gilt, d.h. dass $f \in \mathcal{R}[a, b]$. Dann gibt es zu jedem $k \in \mathbb{N}$ eine Partition $\hat{P}_k \in \mathcal{P}[a, b]$ von (a, b) , so dass

$$\bar{I}(f, \hat{P}_k) - \underline{I}(f, \hat{P}_k) \leq \frac{1}{k}. \quad (\text{XIII.112})$$

Setzen wir $P_k := \hat{P}_1 \cup \hat{P}_2 \cup \dots \cup \hat{P}_k$, so gilt

$$\bar{I}(f, P_k) - \underline{I}(f, P_k) \leq \frac{1}{k} \quad \text{und} \quad P_k \subseteq P_{k+1}, \quad (\text{XIII.113})$$

für jedes $k \in \mathbb{N}$. Wir schreiben nun $P_k =: \{x_0 = a < x_1 < \dots < x_L < x_{L+1} = b\}$ und bilden nun zwei Funktionenfolgen auf $[a, b]$ durch

$$\tilde{f}_k := \sum_{\ell=1}^L \alpha_{\ell} \mathbf{1}_{Q_{\ell}}, \quad \bar{f}_k := \sum_{\ell=1}^L \beta_{\ell} \mathbf{1}_{Q_{\ell}}, \quad (\text{XIII.114})$$

$$\alpha_{\ell} := \min_{x_{\ell-1} \leq x \leq x_{\ell}} \{f(x)\}, \quad \beta_{\ell} := \max_{x_{\ell-1} \leq x \leq x_{\ell}} \{f(x)\}, \quad Q_{\ell} := [x_{\ell-1}, x_{\ell}).$$

Offensichtlich sind $\tilde{f}_k$ und $\bar{f}_k$ Treppenfunktionen, für die nach (XIII.106) Lebesgue-Integral und Riemann-Integral übereinstimmen; genauer gelten sogar

$$\int_{\mathbb{R}} \tilde{f}_k(x) d\mu_1(x) = \sum_{\ell=1}^L \alpha_{\ell} \mu_1(Q_{\ell}) = \sum_{\ell=1}^L \alpha_{\ell} (x_{\ell} - x_{\ell-1}) = \underline{I}(f, P_k), \quad (\text{XIII.115})$$

$$\int_{\mathbb{R}} \bar{f}_k(x) d\mu_1(x) = \sum_{\ell=1}^L \beta_{\ell} \mu_1(Q_{\ell}) = \sum_{\ell=1}^L \beta_{\ell} (x_{\ell} - x_{\ell-1}) = \bar{I}(f, P_k). \quad (\text{XIII.116})$$

Weiterhin sind

$$\tilde{f}(x) := \lim_{k \rightarrow \infty} \underline{f}_k(x) = \sup_k \underline{f}_k(x) \leq f(x), \quad (\text{XIII.117})$$

$$\bar{f}(x) := \lim_{k \rightarrow \infty} \bar{f}_k(x) = \inf_k \bar{f}_k(x) \geq f(x), \quad (\text{XIII.118})$$

für alle $x \in [a, b)$. Nach Lemma XIII.16 sind $\tilde{f}$ und $\bar{f}$ auf $\mathbb{R}$ reell-messbar, und der Satz über monotone Konvergenz sowie (XIII.112) und (XIII.115) implizieren, dass

$$\int_{\mathbb{R}} \tilde{f}(x) d\mu_1(x) = \lim_{k \rightarrow \infty} \left\{ \int_{\mathbb{R}} \tilde{f}_k(x) d\mu_1(x) \right\} = \lim_{k \rightarrow \infty} \underline{I}(f, P_k) = \int_a^b f(x) dx, \quad (\text{XIII.119})$$

$$\int_{\mathbb{R}} \bar{f}(x) d\mu_1(x) = \lim_{k \rightarrow \infty} \left\{ \int_{\mathbb{R}} \bar{f}_k(x) d\mu_1(x) \right\} = \lim_{k \rightarrow \infty} \bar{I}(f, P_k) = \int_a^b f(x) dx. \quad (\text{XIII.120})$$

Insbesondere ist

$$\int_{\mathbb{R}} \{\bar{f}(x) - \tilde{f}(x)\} d\mu_1(x) = 0, \quad (\text{XIII.121})$$

was wegen $\bar{f} - \tilde{f} \geq 0$ auch $\bar{f} = \tilde{f}$ fast überall nach sich zieht. Da $\tilde{f} \leq f \leq \bar{f}$, gilt somit auch $f = \tilde{f}$ fast überall. $\square$